



UNIVERSITÀ DEGLI STUDI DELLA TUSCIA DI VITERBO

Dipartimento di Geologia e Ingegneria Meccanica, Naturalistica e Idraulica per il
Territorio

Corso di Dottorato di Ricerca in
MECCANICA AGRARIA - XXIII Ciclo

INNOVAZIONI BIOFOTONICHE IN AGRICOLTURA: ANALISI IPERSPETTRALE E DELLA FORMA PER IMMAGINI

Settore scientifico disciplinare
AGR/09 MECCANICA AGRARIA

Coordinatore

Prof. Ing. Danilo MONARCA

Firma

Tutor

Ing. Giovanni SANTORO

Firma

Dottorando

Dr. Paolo MENESATTI

Firma

Viterbo, 20 Dicembre 2010

RIASSUNTO

La biofotonica comprende l'insieme delle discipline tecnologiche e ingegneristiche che utilizzano quale principale vettore informativo la radiazione elettromagnetica (fotonica - 50 mld €, 246mila addetti nella UE nel 2005), nelle applicazioni nelle scienze della terra e nei biosistemi. Le attività scientifiche di biofotonica nell'ingegneria agraria sono in forte crescita in relazione all'elevato tasso di innovatività (basato su comparti high-tech: ottica, elettronica, informatica e algoritmica) e alle esigenze della ricerca e della produzione di sistemi di monitoraggio o controllo altamente informativi, multiparametrici, speditivi, non distruttivi, per applicazioni in campo (on-the-go) o in linea di processo (on-line). Numerose ricerche sono state analizzate e riportate per applicazioni nell'ingegneria agraria: robotica (visione artificiale) e macchine motrici e semoventi (guida automatica e assistita), macchine operatrici in lavorazioni di campo (fertilizzazione e distribuzione fitosanitaria di precisione) e raccolta, testing e certificazione di caratteristiche operative e della qualità di lavoro, macchine per il postraccolta (selezionatrici per qualità esterna e interna dei frutti), qualità dei prodotti agricoli (carne, ortofrutticoli, pesce) e alimentari (formaggi, pane).

Tra i molteplici, sono stati considerati due tra i più innovativi ambiti tecnico-metodologico-applicativi: l'analisi della forma per immagini e l'analisi di immagini iperspettrali, di cui sono state riportate applicazioni sperimentali pubblicate o in corso di pubblicazione (principalmente su riviste internazionali ISI).

L'analisi della forma di prodotti agricoli è stata analizzata nella sua complessità scientifica e applicativa attraverso una specifica review (*Shape analysis of agricultural products by computer vision – a review of recent research advances*), caratterizzandosi come tecnica oggettiva utile per la selezione qualitativa varietale (per esempio, distinguendo frutti di produzioni DOP o IGP), la determinazione di soglie e limiti di difetti e danni di forma e come supporto ai genetisti nel confronto clonale e varietale. Sperimentalmente, la tecnica è stata applicata combinando l'acquisizione dei profili delle immagini di prodotti agricoli, con l'analisi modellistica classificatoria multivariata supervisionata (Partial Least Square Discriminant Analysis – PLSDA) applicata sui coefficienti delle equazioni delle armoniche che descrivono matematicamente il profilo, calcolate attraverso l'algoritmo EFA (Analisi

Ellittica di Fourier). Tale sistema tecnico-metodologico, per la prima volta applicato dagli autori in ambito agroalimentare, si è rivelato molto potente e discriminante, consentendo di classificare frutti di mandorlo (in guscio e sgusciati) distinguendo mediamente al 90% tra gruppi di forma simile, su un totale di 18 cultivar (*Quantitative method for shape description of almond cultivars [Prunus amygdalus Batsch]*). La stessa analisi (*Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars*) è stata applicata su quattro varietà tradizionali di nocciole (tonde: Tonda di Giffoni, Tonda Romana; oblunghe: San Giovanni, Mortarella) con risultati di classificazione superiori al 95% tra le cv all'interno di ciascun gruppo. Anche 17 differenti genotipi di arance Tarocco (*Discrimination of Tarocco sweet orange [Citrus sinensis (L.) Osbeck] varieties using opto-electronic elliptic Fourier based analysis of fruit shape*), sono stati analizzati con la tecnica EFA e PLSDA, riuscendo a classificare 5 gruppi morfologici, in modo più oggettivo rispetto agli standard descrittivi dei manuali varietali (IPGRI o Citrus Industry). In ultimo, questa tecnica innovativa è stata applicata, con buon grado di successo (media 69%), per valutare il grado attacco dei due patogeni micotossinogeni responsabili della fusariosi della spiga del frumento, misurando variazioni di forma delle cariossidi (striminzitura) (*Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels*).

La spettrometria di immagine è una delle tecniche fotoniche più potenti risolvendo nella dimensione spettrale tutti i punti di un'immagine. Il soggetto di analisi è caratterizzato da un'impronta spettrale e da proprietà spaziali (morfologiche, strutturali). Tralasciando il remote sensing, le applicazioni spaziano dal controllo ambientale (proximal sensing), alle macchine per selezione e alla qualificazione dei prodotti alimentari. La prima attività sperimentale riportata, è inerente lo studio del grado di maturazione delle mele golden (*Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples*) attraverso determinazione del tradizionale indice di amido, calcolato innovativamente attraverso analisi di immagini iperspettrali nel vicino infrarosso (NIR). I buoni risultati ottenuti indicano la possibilità di sostituire il test che prevede l'uso di prodotti chimici (iodio-ioduro) con la tecnica rapida e di non contatto. La seconda attività (*Quality evaluation of fish by hyperspectral imaging*), oltre a indicare la

maggior parte delle ricerche applicative sul prodotto ittico, è stata finalizzata alla determinazione dei giorni di conservazione in refrigerazione di pesci allevati.

Attraverso l'uso, per la prima volta documentato, della combinazione dell'analisi iperspettrale e della forma per immagini, è stato possibile estrarre le bande spettrali e l'area della superficie del pesce più informative ai fini della stima della freschezza. Il sistema, veloce e di non contatto, si è dimostrato in grado di rilevare lo stato di freschezza, distinguendo con un valore pari al 79.4% pesci conservati solo 1-2 giorni (prodotto considerato fresco), da pesci conservati 3-5 giorni (pesce non fresco, ma ancora perfettamente edibile).

ABSTRACT

Biophotonics includes all the technological and engineering disciplines that use the electromagnetic radiation as the main carrier of information (photonics - 50 billion €, 246.000 workers in the EU in 2005) in earth sciences and biosystems applications.

The use of biophotonics in agricultural engineering discipline is growing in relation to the high innovation rate (based on high-tech sectors: optics, electronics, computers and algorithms) and the needs of research, production monitoring and control systems, highly informative, multiparametric, rapid, non destructive, for applications in the field (on-the-go) or in the process line (on-line). Numerous studies have been analyzed and reported for agricultural applications in: robotics (vision), tractors and self-propelled machines (automatic or assisted guide), agricultural field machines (fertilizers and pesticide precise distribution), test and certification of operational characteristics and work quality, post-harvest machines (fruit selection based on external and internal qualities), quality of agricultural products (meat, fruit and vegetables, fish) and food (cheese, bread).

Among such a variety of applications this work consider two of the most innovative technical-methodological sector of application, shape imaging and hyperspectral imaging, which have been reported in experimental applications published or submitted for publication (mainly on international ISI journals).

A specific review (*Shape Analysis of Agricultural products by computer vision - a review of recent research advances*) explored the technical-methodological complexity and the scientific applications of the shape analysis of agricultural products. This technique revealed to be useful for the objective selection of samples variety quality (for example, distinguishing the geographical origin of fruits), the determination of thresholds for damage and defects of shape and form and as a support to genetists for clonal and varietal comparison. Experimentally, the technique has been applied by combining the sample profiles, extracted from the images of agricultural products, with the supervised multivariate classificatory model analysis (Partial Least Square Discriminant Analysis - PLSDA) applied to the harmonic coefficients of the equations that mathematically describe the profile, calculated through the algorithm EFA (Elliptical Fourier analysis). This system was the first

time applied by the authors in the food industry. It has proved very powerful and discriminating, allowing the correct classification of almond fruit (kernel and in-shell) with an average value of 90% between groups of similar shape on a total of 18 cultivars (*Quantitative method for shape description of almond cultivars [Prunus amygdalus Batsch]*). The same analysis (*Shape-based methodology for multivariate Discrimination Among Italian hazelnut cultivars*) was applied on four traditional hazelnut cultivars (round: Tonda di Giffoni, Tonda Romana, oblong: San Giovanni, Mortarella). The classification results with over 95% between the cultivar within each group. Moreover, 17 different genotypes of Tarocco oranges (*Tarocco Discrimination of sweet orange [Citrus sinensis (L.) Osbeck] varieties using Elliptic Fourier based opto-electronic analysis of fruit shape*) were analyzed with this technique (EFA + PLSDA), to classify 5 morphological groups in a more objective way than the reference selection technique which sorts them manually (IPGRI or Citrus Industry). Finally, this novel technique has been applied, with a good success (average 69%), to assess the degree of attack of two phungi micotoxins responsible for Fusarium head blight of wheat by measuring variations in the kernels shape (*Application of morphometric image analysis system to evaluate the incidence of fusarium head blight infected wheat kernels*).

Hyperspectral imaging is one of the most powerful techniques to resolve in spectral dimension all the pixels of an image. The subject of analysis is characterized by spectral fingerprint and spatial properties (morphological, structural). Apart from remote sensing, the applications range from environmental monitoring (proximal sensing), to field machinery or to the agricultural products and food quality analysis and selection. An innovative determination of the maturity degree of Golden apples (*Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples*) is the first experimental work reported. It was based on the determination of the traditional starch index, calculated innovatively through a hyperspectral near infrared (NIR) imaging. The obtained good results indicate the possibility to replace the reference test, involving the use of chemicals (iodine-iodide), with a rapid and non-contact technique. The second activity (*Quality evaluation of fish by hyperspectral imaging*) reports a large overview of hyperspectral imaging application for fish quality analysis. The experimental part aimed to estimate

fish freshness of farmed fish in terms of days of chilled conservation. The article reports the first literature example of the combination of shape and hyperspectral imaging in agricultural and food systems. In this way, it was possible to extract the more informative spectral bands and surface area of the fish, for freshness estimation. The fast non-contacting system was able to detect the fish freshness, correctly classifying (79.4%) fish stored only 1-2 days (fresh product) from fish stored 3-5 days (fish not fresh, but still perfectly edible).

Ringraziamenti

Un ringraziamento sentito va al prof. Danilo Monarca che mi ha supportato e sopportato in questa operazione accademica raggiunta in piena maturità anagrafica.

Ringrazio sentitamente il Direttore del CRA-ING e tutor del presente dottorato, Ing. Giovanni Santoro che mi ha saggiamente consigliato in tutti questi anni e continuamente spronato all'attività di ricerca da svolgere con serietà e onestà.

Un ringraziamento assolutamente doveroso va a tutti il collaboratori del gruppo afferente al Laboratorio per le Applicazioni Ingegneristiche Innovative in Agricoltura e Ambiente (AgriTechLab) che, immeritadamente, mi pregio di coordinare dall'anno 2000. Mi hanno coadiuvato in tutti questi anni, lasciandomi la soddisfazione di veder crescere professionalità e qualità umane ed etiche.

Riporto di seguito un elenco, spero il più completo possibile, in ordine di prossimità temporale:

Corrado Costa

Jacopo Aguzzi

Federico Pallottino

Francesca Antonucci

Silvia Solaini

Stefano Giorgi

Emiliano Canali

Graziella Paglia

Stefano D'Andrea

Renata Lucarini

Iliana Niciarelli

Marcos Pincu

Matteo Cegna

Gigliola Urbani

Vorrei ringraziare per il prezioso aiuto i colleghi:

Alessandro Infantino del Centro di ricerca per la patologia vegetale (Roma) del

Consiglio per la Ricerca e la Sperimentazione in Agricoltura (CRA-PAV)

Francesco Intrigliolo, Giuseppe Russo, Santo Recupero e Reforgiato Recupero Centro

di ricerca per l'agrumicoltura e le colture mediterranee (Acireale CT) del del

Consiglio per la Ricerca e la Sperimentazione in Agricoltura (CRA-ACM)

Roberto De Salvador e Pasquale Piccirillo, rispettivamente del Centro di ricerca per la

Frutticoltura di Roma e di Caserta del Consiglio per la Ricerca e la

Sperimentazione in Agricoltura (CRA-FRU e CRA-FRC)

Angelo Zanella del Centro di Sperimentazione Agraria di Laimburg (Bolzano)

Cristina Lo fazio e Gilberto Ferrari di Federcoopesca (Roma)

Un ringraziamento particolare e ricordo inteso, va al compianto collega e amico

Giacomo Lanza del CRA-ACM di Acireale che alleviava con il suo buonumore e la

sua tenacia professionale anche i momenti più complessi di un'attività spesso intensa.

Sommario

RIASSUNTO	2
ABSTRACT	5
Ringraziamenti	8
Parte I	14
Introduzione e Stato dell'arte	14
1. Introduzione e obiettivi	15
1.2 Obiettivi e piano di attività della tesi	16
1.3. Collaborazioni esterne e finanziamenti alle attività sperimentali	17
2. STATO DELL'ARTE	21
2.1 Fotonica e Biofotonica (photonics and biophotonics)	21
2.2 Elementi costitutivi dei sistemi di analisi fotonica	25
2.2.1 <i>Visione artificiale e acquisizione del segnale immagine</i>	28
<i>L'immagine</i>	29
<i>Acquisizione dell'immagine</i>	32
<i>I sistemi di illuminazione</i>	39
<i>Elettronica di digitalizzazione, controllo e memorizzazione</i>	42
<i>Estensione dei sistemi di analisi di immagine al dominio del tempo</i>	44
2.2.2 <i>Processamento delle immagini (elaborazione ed analisi)</i>	46
<i>Elaborazione delle immagini</i>	50
<i>Analisi delle immagini</i>	52
2.3 Analisi multivariata delle immagini e delle informazioni	57
<i>Principal Component Analysis - PCA</i>	59
<i>Partial Least Squares (PLS) e Partial Least Squares Discriminant Analysis (PLS-DA)</i>	60
2.4 Applicazioni biofotoniche	63
<i>Imaging ambiente, suolo, piante, animali</i>	64
<i>Animali zootecnici</i>	64
<i>Suolo agrario</i>	66
<i>Piante</i>	66
<i>Robotica e macchine agricole</i>	68
<i>Robot e macchine autonome</i>	69
<i>Applicazioni nelle fasi di preraccolta</i>	71
<i>Applicazioni sulle macchine agricole</i>	72
<i>Macchine motrici</i>	72
<i>Macchine operatrici</i>	72
<i>Testing e analisi delle caratteristiche operative e costruttive delle macchine</i>	74
<i>Macchine Motrici</i>	74
<i>Macchine operatrici</i>	75
<i>Food processing e qualità dei prodotti agroalimentari</i>	76

2.5 Bibliografia consultata	81
Parte II	110
Attività e risultati sperimentali pubblicati	110
3. ANALISI DELLA FORMA PER IMMAGINI (Shape imaging)	111
3.1 Introduzione	111
3.1.1 La forma e la taglia	111
3.1.2 Rilevanza nell'ingegneria agraria, agroindustria e alimentare	112
3.1.3 Relazione forma-qualità	114
3.1.4 Dalla valutazione soggettiva alla misura quantitativa della forma	116
3.1.5 Analisi della forma per immagini: innovazione in ambito biofotonico	117
3.1.6 Bibliografia	122
3.2 La morfometria	123
3.2.1. Morfometria tradizionale	124
3.2.2 La Morfometria Geometrica	125
3.2.3 Analisi dei profili – Analisi Ellittica di Fourier	129
3.2.4 Bibliografia	132
3.3 Pubblicazioni presentate	134
3.3.1 ABSTRACT	134
3.4 Shape analysis of agricultural products by computer vision – a review of recent research advances	139
Costa C, Antonucci F, Pallottino F, Aguzzi J, Sun DW, Menesatti P, ACCEPTED. Shape analysis of agricultural products by computer vision – a review of recent research advances. ACCEPTED TO FOOD AND BIOPROCESS TECHNOLOGY	139
3.5 Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars	189
Menesatti P, Costa C, Paglia G, Pallottino F, D'Andrea S, Rimatori V, Aguzzi J, 2008. Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars. BIOSYSTEMS ENGINEERING, 101(4): 417-424.	189
3.6 Quantitative method for shape description of almond cultivars (<i>Prunus amygdalus Batsch</i>)	208
Antonucci F, Costa C, Pallottino F, Paglia G, Rimatori V, De Giorgio D, Menesatti P, ACCEPTED. Quantitative method for shape description of almond cultivars (<i>Prunus amygdalus Batsch</i>). IN PRESS TO FOOD AND BIOPROCESS TECHNOLOGY	208
3.7 Discrimination of Tarocco sweet orange [<i>Citrus sinensis</i> (L.) Osbeck] varieties using opto-electronic elliptic Fourier based analysis of fruit shape	241
Costa C, Menesatti P, Paglia G, Pallottino F, Aguzzi J, Rimatori V, Russo G, Recupero S, Reforgiato Recupero G, 2009. Discrimination of Tarocco sweet orange [<i>Citrus sinensis</i> (L.) Osbeck] varieties using opto-electronic elliptic Fourier based analysis of fruit shape. POSTHARVEST BIOLOGY AND TECHNOLOGY, 54: 38-47.	241

3.8 Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels	266
<i>Menesatti P, Antonucci F, Costa C, Santori A, Niciarelli I, Infantino A, 2009, Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels. 1st International Workshop on Computer Image Analysis in Agriculture, Potsdam, Germany 27 – 28 August 2009, Bornimer Agrartechnische Berichte - Heft 69, ISSN 0947-7314, Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V. (ATB)</i>	
4. ANALISI DI IMMAGINE IPERSPETTRALE (hyperspectral imaging)	273
4.1 Introduzione	273
4.1.1 Proprietà ottiche dei materiali	276
4.1.2 Le Proprietà della luce	276
4.1.3 Le leggi sperimentali della propagazione luminosa: la riflessione e rifrazione	280
4.1.4 Superficie diffondente ideale	282
4.1.5 Misure spettrali	282
4.1.6 Spettrofotometri	283
4.1.7 Spettrometri e spettrofotometri puntuali	286
4.1.8 Spettrometri per immagini	287
4.1.9 Scanner spettrali	289
4.1.10 Processamento dell'immagine spettrale	296
4.1.11 Analisi d'immagine iperspettrale	297
4.1.12 Analisi di immagine iperspettrale: innovazione in ambito biofotonico e agroalimentare	303
4.1.13 Bibliografia	306
4.2. Pubblicazioni presentate	313
4.2.1 ABSTRACT	313
4.3. Quality evaluation of fish by hyperspectral imaging	316
<i>Menesatti P, Costa C, Aguzzi J, 2010. Quality evaluation of fish by hyperspectral imaging. In: Hyperspectral imaging for food quality: analysis and control. In D.-W. Sun (Ed.), Hyperspectral imaging for food quality analysis and control. London, Burlington, San Diego: ACADEMIC PRESS, ELSEVIER: 273-294. ISBN: 9780123747532</i>	
4.4 Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples	342
<i>Menesatti P, Zanella A, D'Andrea S, Costa C, Paglia G, Pallottino F, 2009, Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples. FOOD BIOPROCESS TECHNOL 2:308–314 - DOI 10.1007/s11947-008-0120-8</i>	
5. CONCLUSIONI GENERALI	359
6. LISTA DELLE PUBBLICAZIONI NEL TRIENNIO (2008-2010) DI DOTTORATO	361
6.1 Pubblicazioni su riviste scientifiche	361
6.1.1 Articoli su riviste peer-reviewed ISI (con IF)	361
6.1.2 Articoli su riviste internazionali peer-reviewed	364

<i>6.1.3 Capitoli di libri a diffusione internazionale</i>	365
<i>6.1.4 Articoli su riviste nazionali</i>	365
6.2 Brevetti	367
6.3 Atti di convegni	367

Parte I

Introduzione e stato dell'arte

1. Introduzione e obiettivi

La biofotonica comprende l'insieme delle discipline tecnologiche e ingegneristiche che utilizzano quale principale vettore informativo la luce e più in generale la radiazione elettromagnetica (fotonica), nelle applicazioni in ambito biologico e delle scienze della terra, comprendendo in esse l'agricoltura, l'agroambiente e l'agroalimentare. La fotonica è uno dei settori scientifici tecnologicamente più all'avanguardia, in modo paragonabile alle nanotecnologie e genera un indotto economico (50 mld €, UE 2005) e sociale, rilevante (246mila addetti, UE 2005).

Le attività di ricerca nel comparto disciplinare sono in forte sviluppo, in quanto sfruttano i continui miglioramenti dell'information and communication technology (ICT) e dell'elettronica (*aspetto tecnologico*), utilizzano la più avanzata analisi digitale e modellistica inferenziale, predittiva, multivariata (*aspetto metodologico*) e si estendono fino ai comparti ambientale, ecologico, biologico, agricolo, food (*aspetto applicativo*).

I tassi di crescita economica della fotonica sono molto elevati, tra il 7-8% l'anno, così come la produzione scientifica: mezzo milione di riferimenti dal 2000 al 2007 per tutti i comparti disciplinari, oltre 60mila per i soli settori agroalimentare e biologico-ambientale (2007 +100% rispetto al 2000).

Ciononostante, esiste un ampio margine di sviluppo applicativo della biofotonica avanzata, rispetto ai settori agro-alimentare e biologico-ambientale, dove è minore o ancora poco sviluppata l'integrazione tra le diverse discipline.

1.2 Obiettivi e piano di attività della tesi

I settori di applicazione delle tecniche e metodiche che sono proposte nella presente tesi, pur essendo scientificamente molto avanzati, scontano, in genere, una minore “contaminazione” multidisciplinare, soprattutto a svantaggio delle discipline ingegneristiche e fisiche più innovative, rispetto al altri ambiti scientifici (chimico, medico).

Il dottorando ha indirizzato la propria attività di ricerca verso l'applicazione in campo agroalimentare e bioambientale di tecniche e metodiche di biofotoniche. Tale aspetto, connesso con le esigenze di cui sopra, ha portato alla formulazione del piano di attività del presente dottorato, il cui obiettivo generale è quello di spingere verso una forte intergrazione multidisciplinare tra tecnologie e sistemi innovativi in ambito biofotonico e potenziali applicazioni nei settori dell'ingegneria agraria.

Il fine è quello di fornire la base tecnologica e metodologica più innovativa, ad applicazioni parimenti innovative in ambito agroalimentare e ambientale, potenzialmente molto importanti per la loro valenza scientifica, ambientale, economica e sociale.

Per conferire supporto applicativo di elevata valenza scientifica ai metodi indagati dal dottorando, le attività sperimentale sono state inserite nell'ambito di diversi finanziamenti e supportate anche da una serie di collaborazioni esterne, come di specificato nel paragrafo 1.3.

Gli obiettivi specifici delle presente tesi sono:

- 1) presentare una overview su rilevanza economica, tecniche e metodi della biofotonica
- 2) illustrare uno stato dell'arte sulle ricerche e applicazioni sperimentali della biofotonica nell'ambito dell'ingegneria agraria e agroalimentare;
- 3) riportare attività sperimentali condotte e pubblicate dal dottorando nell'ambito della biofotonica applicata all'ingegneria agroalimentare, scegliendo due ambiti tecnico-metodologici di elevata innovatività, comunemente denominati:
 - a. analisi della forma per immagini
 - b. analisi di immagini iperspettrali

La ripartizione dei due settori è puramente di indirizzo, in quanto entrambi i sistemi si basano su tecnologie comuni e si compongono di tecniche e metodi simili o intersecantesi.

L'analisi della forma, misurata e parametrata attraverso complessi algoritmi, è una tecnica molto interessante, per il processo di qualificazione del prodotto agroalimentare. La sua applicazione pratica è, tuttavia, relativamente, semplice e poco costosa, richiedendo sistemi tecnologici di largo consumo (videocamere o fotocamere CCD o CMOS).

L'analisi di immagine iperspettrale, invece, è una tecnica molto più complessa sia dal punto di vista metodologico che tecnologico. Necessita di specifici strumenti, il cui costo, tuttavia, è in costante diminuzione, in virtù del notevole progresso tecnologico del settore. Associando informazioni spettrali e spaziali, si conferma come una delle tecniche fotoniche più informative. È ampiamente studiata e applicata per la qualificazione multiparametrica di prodotti, processi e fattori di produzione.

Per entrambe le tecniche sono state riportate delle attività sperimentali innovative, pubblicate su riviste di impatto, sia dal punto di vista applicativo che metodologico.

Si sottolinea come per la prima volta in ambito agroalimentare, sia stata applicata l'analisi della forma, attraverso l'analisi ellittica di Fourier (EFA) e una tecnica combinata di analisi di immagini iperspettrali e geometria morfometrica.

1.3. Collaborazioni esterne e finanziamenti alle attività sperimentali

L'attività sperimentale esposta negli articoli scientifici riportati nella presente tesi si è stata sviluppata dallo scrivente dottorando e dal gruppo di ricerca del CRA-ING afferente al Laboratorio per le Applicazioni Ingegneristiche Innovative in Agricoltura e Ambiente (AgriTechLab) coordinato dallo stesso dottorando, grazie ai finanziamenti sotto elencati:

- RAVAGRU; (2008-2010) DM MIPAAF 187/7303/07 DEL 02/08/2007; Titolo Progetto: Ricerche avanzate in agrumicoltura e loro applicazione; Ente finanziatore: MIPAF; Coordinamento scientifico: CRA-ACM, Consiglio per la Ricerca in Agricoltura, Centro di ricerca per l'agrumicoltura e le colture mediterranee (Acireale CT), dr. Francesco Intrigliolo; Responsabile U.O. ricerca: dott. Paolo Menesatti; titolo ricerca: Sviluppo di metodiche strumentali speditive per la valutazione non distruttiva della consistenza dei frutti di agrume e sperimentazione di tecnologie innovative di conservazione (attestato CRAING prot. 2874 del 18/11/2009);
- AGRUQUAL; 2006-2008 DM MIPAAF 213/7303/05 DEL 22/07/2005; Titolo Progetto: Innovazioni per il raggiungimento della qualità globale in agrumicoltura; Programma finalizzato Ministero delle Politiche Agricole e Forestali;- Coordinatore scientifico: CRA-ACM (Agrumicoltura, Acireale - CT), dr. Francesco Intrigliolo; Responsabile convenzione scientifica: dott. Paolo Menesatti; titolo ricerca: Applicazioni di sistemi termografici e olfattometrici per il mantenimento qualitativo dei prodotti in post-raccolta (attestato CRAING prot. 2875 del 18/11/2009);
- FRUMED, VAFRUSEME; 2006-2008 DM MIPAAF 212/7303/05 DEL 22/07/2005; “Valorizzazione della frutta secca del meridione: germoplasma, tecniche innovative nella produzione e conservazione per il miglioramento della qualità”; Ente finanziatore: MiPAF, Programma di sviluppo per il Mezzogiorno d'Italia: ricerca ed innovazione tecnologica. (accantonamento delibera CIPE n. 17/2003); Coordinatore scientifico: CRA-FRU (Frutticoltura, RM) prof. Carlo Fideghelli; Responsabile U.O. ricerca: dott. Paolo Menesatti; titolo ricerca: Valutazione delle caratteristiche qualitative della frutta secca con metodi non distruttivi
- HIGHVISION; Studio e prime applicazioni di nuove metodologie biofotoniche nell'agroalimentare, biologia e ambiente; Ente finanziatore: MIPAF, Bando CRA Interno DM 19177/7303/08 del 19/12/2008; COORDINATORE scientifico: CRA-ING (dott. Paolo Menesatti); Responsabile U.O. ricerca N° 1: dott. Paolo Menesatti, titolo ricerca: Studio e applicazione di nuove

metodiche strumentali ed elaborative di image analysis, artificial vision e morfometria

- POR Puglia - pesce; 2005-2006; Determinazione di indici di freschezza e provenienza di prodotti ittici attraverso applicazioni optoelettroniche non invasive; Programma Operativo Regionale - Puglia; Proponente e responsabile del finanziamento: GESTICOM srl; Coordinatore e responsabile scientifico: dott. Paolo Menesatti;
- MELANIR - Miglioramento della qualita' e determinazione strumentale degli indici di maturazione delle mele attraverso analisi d'immagine iperspettrale e multispettrale nel visibile e nel vicino infrarosso; Finanziamento da parte del Centro di Sperimentazione Agraria e Forestale di Laimburg (BZ) – su fatturazione; COORDINATORE e Responsabile tecnico-scientifico dell'incarico: dott. Paolo Menesatti

La stessa attività si è avvalsa di importanti e fondamentali collaborazioni esterne, tra cui si citano:

- Instituto de ciencias del Mar (ICM-CSIC) - Departement Recursos Marinos Renovables (Barcelona, Spain), dr. Jacopo Aguzzi (Postdoctoral Fellow of the “JAE” Program - Education and Science Ministry-MEC, Spain);
- Centro di ricerca per la patologia vegetale (Roma) del Consiglio per la Ricerca e la Sperimentazione in Agricoltura (CRA-PAV), dr. Alessandro Infantino
- Centro di ricerca per l'agrumicoltura e le colture mediterranee (Acireale CT) del del Consiglio per la Ricerca e la Sperimentazione in Agricoltura (CRA-ACM), Dr. Francesco Introgliolo, Dr. Giuseppe Russo, Dr. Santo Recupero, Dr. Reforgiato Recupero e il compianto Dr. Giacomo Lanza
- Centro di ricerca per la frutticoltura (Roma) del Consiglio per la Ricerca e la Sperimentazione in Agricoltura (CRA-FRU), Dr. Flavio Roberto De Salvador
- Centro di ricerca per la frutticoltura (Caserta) del Consiglio per la Ricerca e la Sperimentazione in Agricoltura (CRA-FRC), Dr. Pasquale Piccirillo
- Centro di Sperimentazione Agraria di Laimburg (Bolzano), Dr. Angelo Zanella

- Federazione delle Cooperative della Pesca, FEDERCOOPESCA (Roma), Dr. Gilberto Ferrari e Dr.ssa Cristina Lo fazio

2. STATO DELL'ARTE

In questa sezione è riportata una breve analisi bibliografica relativa alle principali metodologie e applicazioni di biofotonica nei settori disciplinari ingegneristico agrario. Al fine di indirizzare l'argomento trattato verso le applicazioni sperimentali successivamente indicate, lo stato dell'arte è stato circoscritto alle metodologie e applicazioni dei settori: biologico, ambientale, ristretta agli ambiti di proximal sensing, applicazioni di laboratorio e sistemi operativi e produttivi in macchine, industrie, prototipi.

2.1 Fotonica e Biofotonica (*photonics and biophotonics*)

La visione artificiale (Jain, 2005; Steger, 2007), l'analisi di immagine e la spettrofotometria di immagini sono, nell'insieme, tecniche optoelettroniche (Lamela, 2007; Hall, 1984; Urbani, 2002), i cui elementi operativi caratterizzanti sono fondamentalmente tre: l'acquisizione di un segnale immagine, l'ampiezza di banda dello spettro elettromagnetico utilizzato nella generazione dell'immagine, la tipologia di elaborazione e analisi dell'immagine. L'insieme delle discipline tecnologiche e ingegneristiche che utilizzano quale principale vettore informativo la radiazione elettromagnetica (luce=photos, in tutta la sua estensione e bande) è anche indicato con il termine di fotonica (Vukusic, 2003; Anderson, 2006; Prasad, 2004).

Volendo comprendere anche le applicazioni in ambito biologico e delle scienze della terra, comprendendo in esse l'agricoltura, l'agroambiente, l'agroalimentare, è stato utilizzato anche il termine di biofotonica (Leahy, 2007; Wang, 2007; Science, 2005; Kincade, 2008; Jovin, 2004; Birge, 2004; Philipp, 2006; Prasad, 2003).

La fotonica è uno dei settori scientifici tecnologicamente più all'avanguardia, in modo paragonabile alle nanotecnologie, rispetto alle quali, tra l'altro, ha forti margini di sovrapposizione. Secondo uno studio commissionato dalla UE (Photonics in Europe - Economic impact, 2005) alla Piattaforma Tecnologica Europea (Photonics21), l'industria e l'indotto del settore ha fatturato (2005) 228 miliardi di

EUR nel mondo (stimato per 439 miliardi di EUR nel 2015), circa 50 nella UE, occupando in Europa 246mila lavoratori. I tassi di crescita del settore sono molto elevati, tra il 7-8% l'anno.

Il settore tecnologico dell'optoelettronica e quello applicativo della biofotonica è in forte crescita scientifica come indica la ricerca di citazioni bibliografiche, relative alla parola chiave "image analysis", effettuata su "google scholar" per soggetti di area agricola, biologica e ambientale (anni 2000 - 2007). Essa fornisce l'indicazione di circa 64000 referenze, che diventano 555000 estendendo la stessa analisi ai settori Chemistry, Engineering, Computer science e Medicine. Inoltre, la produzione bibliografica è cresciuta negli ultimi anni, in relazione alle maggiori disponibilità tecnologiche e prestazioni H&S, quasi raddoppiando, per i soli riferimenti all'area agricola, biologica e ambientale, passando dall'anno 2000 al 2007.

La piattaforma tecnologica UE, Photonics 21, fondata a Dicembre 2005 è una struttura di incontro e interscambio tra le industrie del settore e gli operatori della ricerca e innovazione, con il supporto dalla Commissione Europea.

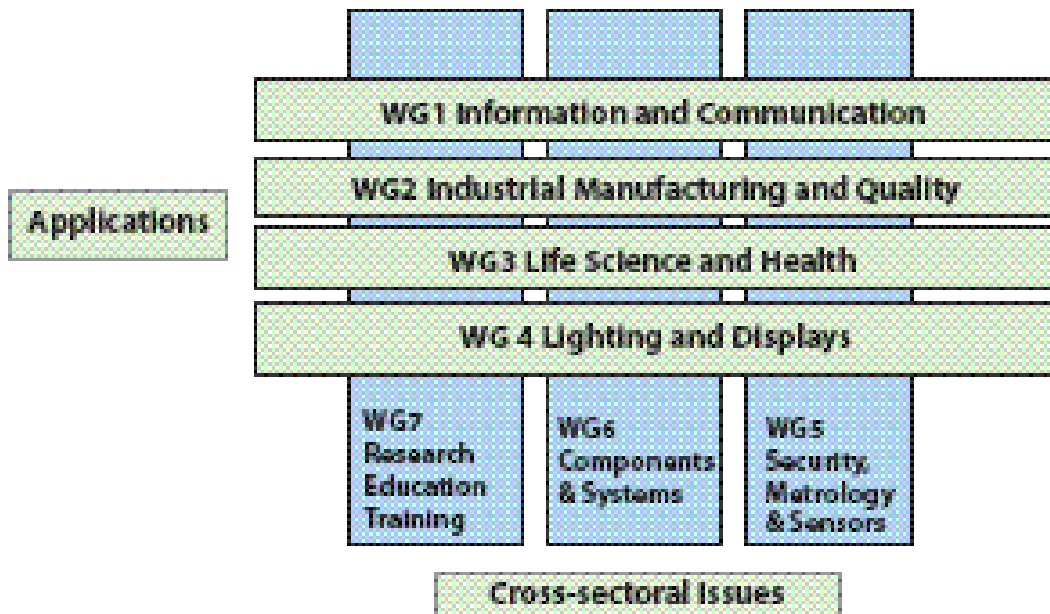
La fotonica è ritenuta un fattore chiave per l'innovazione tecnologica e uno delle più importanti tecnologie per i mercati emergenti nel 21° secolo. Le aree di maggiore interesse sono considerate:

- Informazione e comunicazione - la banda larga di prossima generazione sarà possibile solo attraverso soluzioni ottiche
- Produzione industriale - tecnologia fondamentale per la creazione di prossima generazione di prodotti ad alta efficienza energetica nelle industrie europee;
- Scienze della Vita e Salute - dispositivi fotonici e le procedure nuovi metodi di diagnosi, monitoraggio e trattamento malattie per le società che invecchia
- generazione di energia attraverso fotonica – Fotovoltaico per contribuire in modo significativo sia per il futuro approvvigionamento di energia elettrica a livello mondiale sia per la mitigazione dei cambiamenti climatici
- Lighting & Display - illuminazione a risparmio energetico e display soluzioni di ridurre notevolmente il consumo di potenza e emissioni di CO2 e sono un settore ad alta crescita economica

Photonic 21, ritiene che il 21 ° secolo sarà il secolo del fotone. Senza una leadership europea nelle tecnologie fotoniche, l'industria europea sarà vulnerabile alla forte concorrenza degli Stati Uniti e dall'Asia.

La struttura operativa del working groups (WG) della piattaforma tecnologica è schematizzata in fig. 1.

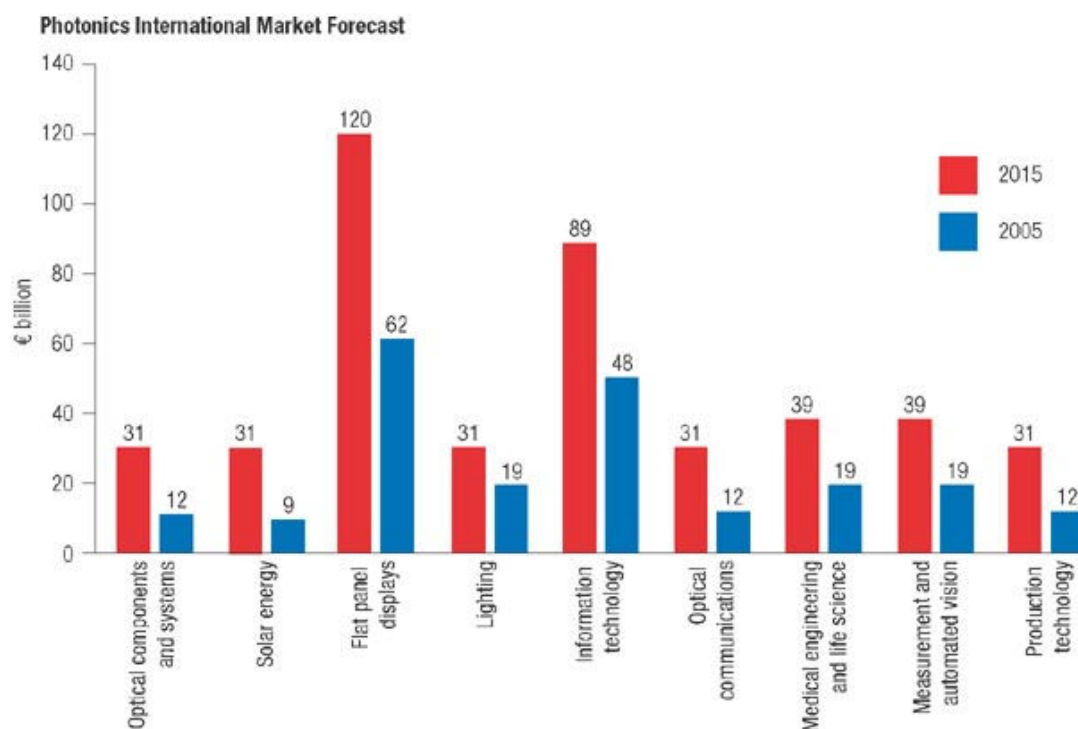
Fig. 1. Schema della struttura di attività dei working group che compongono la piattaforma tecnologica sUE sulla fotonica (Photonics 21) (fonte UE: http://cordis.europa.eu/fp7/ict/photonics/p21_en.html)



Più in dettaglio, sono stimati (Fig. 2) in forte crescita per il 2015, i ricavi inerenti le applicazioni industriali fotoniche sui monitor e display (+100%)e nell'information technology (+85%).

Anche l'incremento del mercato di applicazioni fotoniche connesse con il medicale e le scienze della vita è atteso essere molto forte (+105%), passando in valori assoluti da 19 a 39 miliardi di EUR.

Fig 2. Previsione del mercato internazionale di attività legate alla fotonica (Gevaux D, 2008, http://www.nature.com/nphoton/journal/v2/n3/box/nphoton.2008.07_BX1.html)



Come è stato accennato, l'interesse verso le applicazioni della fotonica nelle scienze della vita (biofotonica) è molto elevato, anche a livello comunitario, come indica l'esistenza di un portale della UE (Photonics4Life) per la creazione e potenziamento del network di eccellenze europee in campo della biofotonica applicata alle scienze della vita (<http://www.photonics4life.eu/P4L/Activities>).

Il network è inserito nell'ICT research field in FP7 e la prevalenza di applicazioni il *life science* è verso il campo biomedico. Tuttavia, l'importanza nelle scienze biologiche e in agricoltura sta crescendo in relazione alla diffusione delle tecnologie e delle potenziali applicazioni.

Il portale *Photonics4Life* raccoglie le risorse umane, le infrastrutture, le capacità, l'esperienza e la conoscenza di più di 50 istituti di alta ricerca nel campo della Biofotonica, provenienti da Stati membri CE e della Russia. Tra essi ci sono 13 membri di base, che hanno costituito la rete nel 2008 e gli Associati P4L, che hanno successivamente aderito. *Photonics4Life* ha l'obiettivo di colmare il divario tra le diverse discipline che vanno dalla fisica e dalla chimica, alla biologia attraverso l'ingegneria e la medicina. Si vuole collegare l'esperienza degli istituti di ricerca con

le PMI e grandi imprese, al fine di promuovere la ricerca Biofotonica e rafforzare la competitività economica dell'Europa nel mercato globale biofotonica.

2.2 Elementi costitutivi dei sistemi di analisi fotonica

La visione rappresenta per l'uomo il fondamentale mezzo esplorativo e cognitivo. Anche nella prassi scientifica, nell'industria e nel lavoro professionale, molte analisi e valutazioni sono affidate alla percezione visiva umana, che, tuttavia, può indurre ad una notevole soggettività di misura e limitarne l'affidabilità, la ripetibilità e la "numerosità". Per questo, da quando l'elettronica e l'informatica hanno messo a disposizione le prime soluzioni tecnologiche, si è enormemente ampliato l'impegno di investimento e di ricerca nel settore della visione artificiale e dell'analisi dell'immagine. Queste sono, attualmente, alla base di numerose applicazioni civili, industriali e scientifiche. I più diffusi e noti settori di sviluppo e utilizzo sono la medicina e la biologia, le scienze dei materiali, le applicazioni industriali, il controllo di produzione e la sorveglianza.

Anche in agricoltura, nell'agro-alimentare e nell'acquacoltura e pesca, queste discipline hanno trovato estesa applicazione, collocandosi, sotto il profilo metodologico e strumentale, nell'ambito delle discipline afferenti all'ingegneria agraria.

Si può inoltre evidenziare il particolare aspetto di "orizzontalità" che caratterizza, dal punto di vista applicativo agricolo, la visione artificiale e la successiva elaborazione e analisi di immagine. Queste discipline hanno difatti trovato utilizzo in tutti i settori del comparto agricolo e alimentare, sebbene con profonde differenze applicative e di risultati. Ad esempio, si sono dimostrate molto più promettenti ed affidabili le applicazioni a punto fisso di analisi o di controllo e selezione, alcune delle quali già operano produttivamente nel post-raccolta e nell'industria di trasformazione, rispetto alla raccolta robotizzata della frutta, dove, non per solo limite della visione artificiale, l'applicazione è ancora in fase sperimentale.

Nei settori di indagine del agroalimentare, biologia e ambiente, le finalità applicative delle tecnologie fotoniche possono essere ricondotte a tre elementi:

1. visione artificiale e machine vision per lo sviluppo di sensoristica di controllo e monitoraggio sulle macchine (selezionatrici dei frutti, robotica);
2. sviluppo di sistemi di misura ed analisi (spettrofotometri, analizzatori di immagine, ecc) di parametri produttivi ed aspetti quali-quantitativi;
3. applicazioni specifiche nell'ambito della ricerca finalizzata o di base, contribuendo con nuove forme di analisi e fornendo informazioni multiparametriche, multisensore, altrimenti difficilmente acquisibili (analisi delle forme vegetali o animali, colore o caratteristiche spettrali, ecc.).

Non è spesso possibile distinguere i diversi ambiti tecnologici e/o operativi di sistemi evoluti di analisi delle immagini. Soprattutto per applicazioni in ambito industriale, si utilizza spesso il termine di *machine vision* (Jain, 2005)

Altro termine molto utilizzato è quello di *computer vision*, tradotto spesso con il termine di visione artificiale, che rimanda a sistemi complessi di acquisizione/analisi delle immagini, ma per applicazioni o in contesti non esclusivamente industriali.

L'EMVA (Associazione Europea di Machine Vision -

<http://www.emva.org/cms/index.php>) definisce la *machine vision* come il processo di trattamento di immagini o sequenze di immagini con l'obiettivo di manipolarle e analizzarle, al fine di

- migliorare la qualità delle immagini (contrasto, colore, ecc);
- ripristinare immagini (ad esempio la riduzione del rumore)
- codificare le immagini (compressione dei dati, per esempio)
- comprendere e interpretare le immagini (analisi di immagine, riconoscimento di forme).

In tal modo, la tecnologia di visione può essere applicata ovunque le immagini vengano generate e debbano essere analizzate: in biologia (cellule di conteggio), in medicina (interpretazione risultati della scansione CT), nel settore delle costruzioni (analisi termografiche di edifici) o nella sicurezza (verifica delle dimensioni biometriche). La tecnologia della visione artificiale è una tecnologia interdisciplinare che combina l'illuminazione, l'ottica, l'elettronica, la tecnologia dell'informazione e

della comunicazione, il software, l'algoritmica e la modellistica. In ambito industriale la comprensione e l'interpretazione di immagini digitali, sono utilizzate per il controllo dei processi di produzione e tale ambito sensoristico è evoluto in una delle tecnologie chiave nel settore dell'automazione industriale, che viene utilizzato in quasi tutte le industrie manifatturiere. Sebbene il termine di *machine vision*, fosse originariamente riferito alle applicazioni industriali della tecnologia di visione, essa è spesso utilizzata per riferirsi anche ad applicazioni non industriali (Jahr, 2003).

Un sistema complesso di visione artificiale e analisi delle immagini è fondamentalmente costituito dai seguenti sottosistemi:

- apparato di visione e acquisizione del segnale immagine;
- sottosistema di processamento ed elaborazione del segnale, ricostruzione e sintesi dell'immagine;
- sottosistema di analisi dei dati a fini interpretativi, inferenziali, classificatori o predittivi.

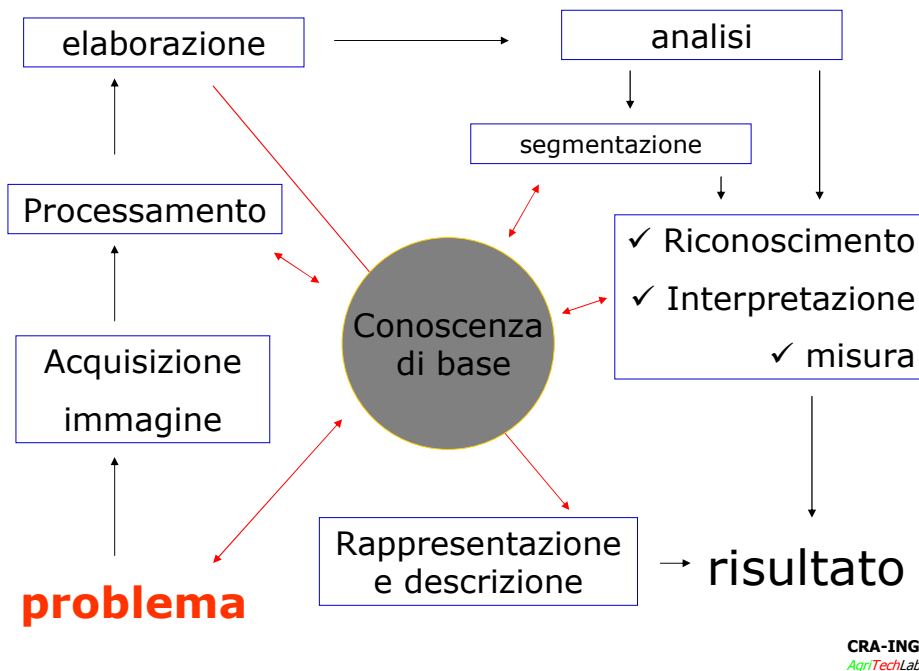
L'immagine è la sorgente informativa (input) attraverso cui analizzare e/o interagire (output) con l'ambiente reale.

L'intero processo che parte dal problema da risolvere, e arriva al risultato operativo è schematizzato in Fig. 3.

Centro di tutto il processo è la conoscenza di base di tutte le fasi coinvolte, a partire dalla competenza nell'utilizzo dei dispositivi tecnici, a tutti i passaggi elaborativi, di processo e di analisi.

Di seguito, saranno brevemente illustrati i principali sottosistemi operanti in un sistema di *machine vision*.

Fig. 3. Schematizzazione di un processo completo di elaborazione ed analisi di immagine (Zeiss, 1990, rielaborato)



2.2.1 Visione artificiale e acquisizione del segnale immagine

La visione artificiale prende avvio come potenziamento funzionale e strumentale dell'interpretazione visiva da parte dell'uomo. In effetti il primo e tuttora più sofisticato sistema di acquisizione, elaborazione ed analisi di immagine è il complesso formato dall'occhio, dal nervo ottico e dal cervello. L'analisi delle immagini è in natura un processo altamente sofisticato e specializzato in tutti gli organismi dotati di apparati per la percezione visiva. Impegna una considerevole parte delle strutture cerebrali, in quanto la complessità informativa, anche della più semplice delle immagini, è notevolissima (tabella 1).

E' sufficiente considerare che un'immagine tridimensionale, come la vediamo noi, fornisce informazioni circa il colore, le ombre, le forme, gli oggetti, le distanze. Anche in agricoltura, l'analisi di immagine soggettiva, ovvero quella affidata all'uomo, ha avuto e tuttora possiede un'importanza fondamentale, nel senso che l'osservazione serve ad aumentare la comprensione di fenomeni agro-biologici,

spesso molto complessi. Infatti, da sempre, l'agricoltore osserva con attenzione l'ambiente e il frutto del proprio lavoro. Il fine è di carpire dall'immagine visiva quelle informazioni che, integrate con l'esperienza e la sensibilità personale, possono consentire di superare le difficoltà, di perfezionare il lavoro e, in definitiva, di migliorare la produttività. L'analisi di immagine numerica eseguita con l'ausilio di strumenti elettronici e informatici, è la trasposizione scientifica, tecnologica di questo processo, eseguita con l'obiettivo di esprimere misure quantitative e "oggettive" e non solo valutazioni qualitative e "soggettive".

Tabella 1. Capacità di informazione degli organi di senso dell'uomo (P. Perin, www.unipv.it/magistretti/didattica/neurofiscell/perin/lezione1.ppt)

Modalita'	Numero recettori	Capacita' di informazione (bit/s)
Vista	$2 \cdot 10^8$	10^7
Udito	$3 \cdot 10^4$	10^5
Tatto	10^7	10^6
Gusto	$3 \cdot 10^7$	10^3
Olfatto	$7 \cdot 10^7$	10^5

L'immagine

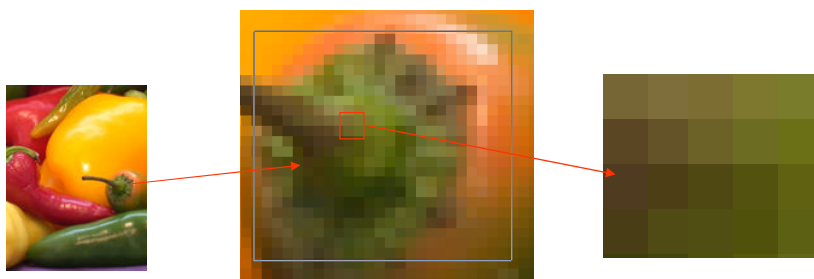
Nell'immagine, la parametrizzazione misuristica o informativa dell'ambiente o dell'oggetto di indagine – tipico di uno strumento di misura – si affida a rilievi non puntuali o monocanale (come ad esempio il colore di un singolo punto), ma avviene per matrici di singoli punti di misura (pixels) contemporaneamente rilevati, come nel caso della colorimetria di immagine, dove è possibile apprezzare, contemporaneamente, diversi punti di diversa colorazione e rilevare caratteristiche topologiche dell'area analizzata.

Una definizione (Wikipedia), indica l'immagine come la metodica di rappresentazione secondo coordinate spaziali indipendenti, di un oggetto o di una scena contenente informazioni descrittive e quantitative riferite all'oggetto e alla scena. Nella comunicazione, il linguaggio delle immagini è intrinsecamente indeterminato, evocativo, dotato di segni che assumono valore simbolico, nell'analisi computerizzata invece risulta ben definita e formalizzata.

L'elemento quantico dell'immagine è il pixel, contrazione della locuzione inglese *picture element* che indica ciascuno degli elementi puntiformi che compongono la rappresentazione di una immagine digitale.

In genere i punti sono così piccoli e numerosi da non essere distinguibili ad occhio nudo e sono caratterizzati dalla propria posizione sulla matrice immagine e da valori che indicano colore e intensità in relazione al sistema di rappresentazione adottato (fig. 4)

Fig 4. Esempio di visualizzazione di singoli pixel a partire da un'immagine digitale a colori



Le immagini digitali possono essere rappresentate numericamente come matrici. La dimensionalità di tali matrici è normalmente indicata come segue:

- Array immagine lineare
- Matrice 2D immagine XY, in Bianco/nero, a livelli di grigio
- Matrice nD immagine XYZ, colore, video, multi o hyperspectral

Per convenzione, si indicano le dimensioni delle matrici in termini di coordinate cartesiane (X, Y, Z): l'array avrà dunque dimensione $X*Y$, con X o $Y = 1$; la matrice 2D, dimensioni $X*Y$; la matrice nD, dimensioni $X*Y*Z$.

Nelle matrici nD, tipicamente nelle immagini iperspettrali, per ogni valore di coordinata Z, avremo un piano XY, chiamato “strato” o “layer”. Queste immagini saranno indicate con il termine di multilayer.

Il contenuto informativo (risoluzione cromatica o profondità di colore) di ogni pixel dell’immagine è, in termini digitali, suddiviso in livelli (semi)continui, il cui numero totale possibile segue la scala delle potenze binarie (tabella 2).

Tabella 2. Principali livelli di digitalizzazione dell’informazione in bit e corrispondenti livelli totali di digitalizzazione espressi come potenze binarie o numeri decimali; un bit è l’unità di misura dell’informazione (dall’inglese "binary unit"), definita come la quantità minima di informazione che serve a discernere tra due possibili alternative equiprobabili.

n bit	2ⁿ livelli	n livelli totali
1	2 ¹	2
8	2 ⁸	256
10	2 ¹⁰	1 024
12	2 ¹²	4 096
16	2 ¹⁶	65 536
24	2 ²⁴	16 777 216
48	2 ⁴⁸	281 474 976 710 656

I livelli di digitalizzazione normalmente più utilizzati nelle immagini sono 1, 8, 24 bit, ma bisogna considerare che le combinazioni informative possibili per l’intera immagine, sono il prodotto della risoluzione di ciascuno strato. Consideriamo, ad esempio, un’immagine XYZ a colori RGB (red, green, blue – standard molto utilizzato nell’imaging di largo consumo), che si estende su tre strati nella dimensione Z, uno strato per ciascuno dei tre canali RGB. Se ciascuno strato avesse una risoluzione di 8 bit (256 livelli), le possibili combinazioni colorimetriche dell’immagine sarebbero $2^8 \times 2^8 \times 2^8 = 2^{24}$. Quindi un valore molto elevato.

Considerando, invece, immagini a 2D, ossia XY, le risoluzioni più comuni sono: 1 bit, 2 livelli cromatici, in genere bianco/nero; 8 bit, 256 livelli (di grigio), dal nero (livello 0) al bianco (livello 255).

Le caratteristiche fondamentali che definiscono le proprietà dell'immagine sono:

- risoluzione spaziale;
- risoluzione di segnale (cromatica, spettrale, ecc.);
- tipologia di formato di salvataggio dell'immagine;
- l'eventuale compressione dell'informazione (compressione che può essere anche alterativa).

Quando l'immagine è utilizzata ai fini misuristici quantitativi, ossia quando l'immagine inquadra una scena o in oggetto di cui sia importante conoscere le dimensioni, sarà necessario definire:

- il campo immagine (field of view – FOV), ossia l'ampiezza (larghezza e altezza) della scena inquadrata dall'immagine;
- il campo istantaneo di immagine (instantaneous field of view – IFOV), ossia l'ampiezza (larghezza e altezza) di parte della scena o dell'oggetto risolta da un singolo pixel.

Acquisizione dell'immagine

Un sistema di acquisizione dell'immagine digitale comprende, in genere, quattro sottosistemi:

- l'apparato ottico di visione e focalizzazione;
- il sensore di acquisizione dell'immagine;
- l'apparato di illuminazione;
- il sistema di conversione analogico/digitale del segnale, di memorizzazione e di gestione digitale complessiva dell'operazione.

In fig. 5 è rappresentato uno schema di massima, relativo ad un sistema di acquisizione di immagini a colori operanti in linea di processo per la selezione di prodotti ortofrutticoli. Sono ben identificabili i quattro sottosistemi indicati. Gli stessi sottosistemi sono visibili in figura 6, relativamente ad un apparecchio sperimentale utilizzato per l'acquisizione delle immagini iperspettrali nel visibile e vicino infrarosso. Il sistema sarà dettagliatamente illustrato nel capitolo inerente le applicazioni di analisi di immagini iperspettrali.

I sistemi di imaging, soprattutto quelli più avanzati e raffinati, possono operare acquisizione e riconoscimento di immagine nel visibile integrato (sistemi real color, o monocromatici: scanner, fotocamere o videocamere), oppure per singole bande spettrali (spettrometri di immagine) del visibile (400-700 nm) o del vicino infrarosso (700-1000 nm) o nell'infrarosso termico, (termocamere).

Esistono altri sistemi che generano immagini da segnali diversi: ultrasuoni, NMR (risonanza magnetica nucleare), raggi X.

Un ulteriore esempio di sistema di acquisizioni delle immagini, è rappresentato da strumenti di rilevamento senza contatto di immagini termiche (termocamere – Fig. 7). Attraverso le termocamere è possibile rappresentare graficamente (funzione immagine) e in molti casi misurare (funzione radiometrica), la radiazione infrarossa totale proveniente da un corpo.

Qualsiasi corpo, per temperature superiori allo zero assoluto ($-273\text{ }^{\circ}\text{C}$), emette energia, sotto forma di radiazione elettromagnetica, che è funzione principalmente della sua temperatura (T^4) e della sua capacità di emissione o emissività (legge di Stefan-Boltzman per i corpi reali) (Maldague, 1994).

Per temperature inferiori ai $500\text{ }^{\circ}\text{C}$, il massimo di irradiazione si osserva per lunghezze d'onda dell'infrarosso medio. Disponendo di sensori in grado di rilevare, con una certa accuratezza, l'irradiazione a queste lunghezze d'onda - termocamere *long wave* $7.5\text{-}14\text{ }\mu\text{m}$ -, è possibile risalire alla temperatura superficiale del corpo, operando delle correzioni dovute a molteplici fattori: fisici, ambientali e strumentali.

Fig. 5. Rappresentazione schematica di un sistema di elaborazione ed analisi delle immagini utilizzato per la selezione di prodotti ortofrutticoli a punto fisso (Menesatti, 1995)

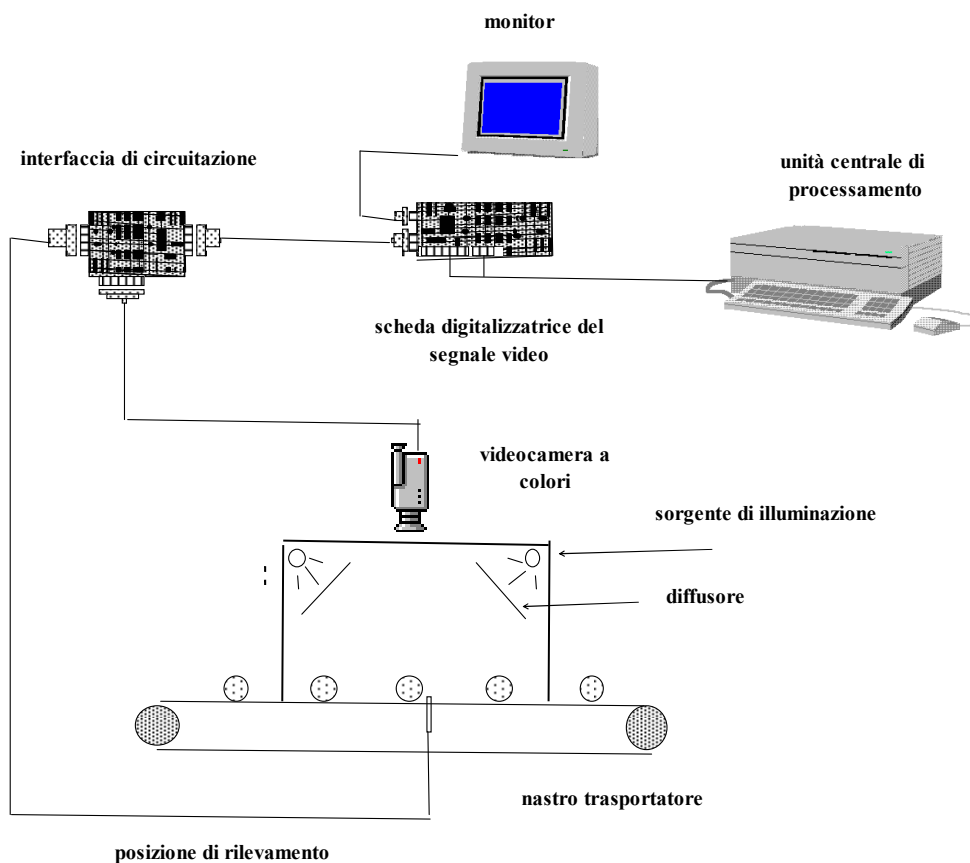
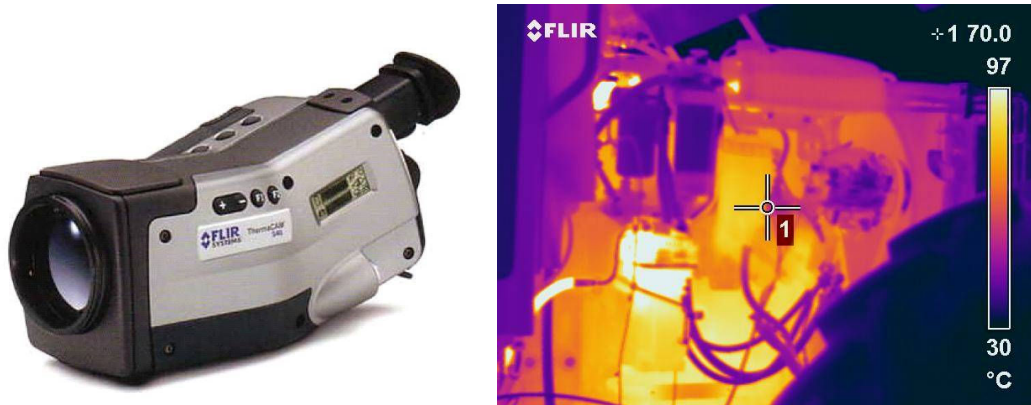


Fig. 6. esempio di sistema di acquisizione di immagini iperspettrali nel visibile e vicino infrarosso (CRA-ING, AgriTechLab, Monterotondo Roma)



Fig. 7, Strumento (termocamera) per l'acquisizione senza contatto di immagini nell'infrarosso termico (IR) e immagine termica di sottocofano di trattore (foto CRA-ING)



L'ottica di visualizzazione e focalizzazione ha un ruolo molto importante nell'acquisizione delle immagini. I sistemi ottici sono i più vari e sono strettamente dipendenti da:

- tipologia di indagine
- l'ampiezza del FOV
- caratteristiche dell'oggetto o degli oggetti da osservare/misurare
- caratteristiche dell'illuminazione (naturale/artificiale), campo chiaro, scuro
- bande spettrali di osservazione;

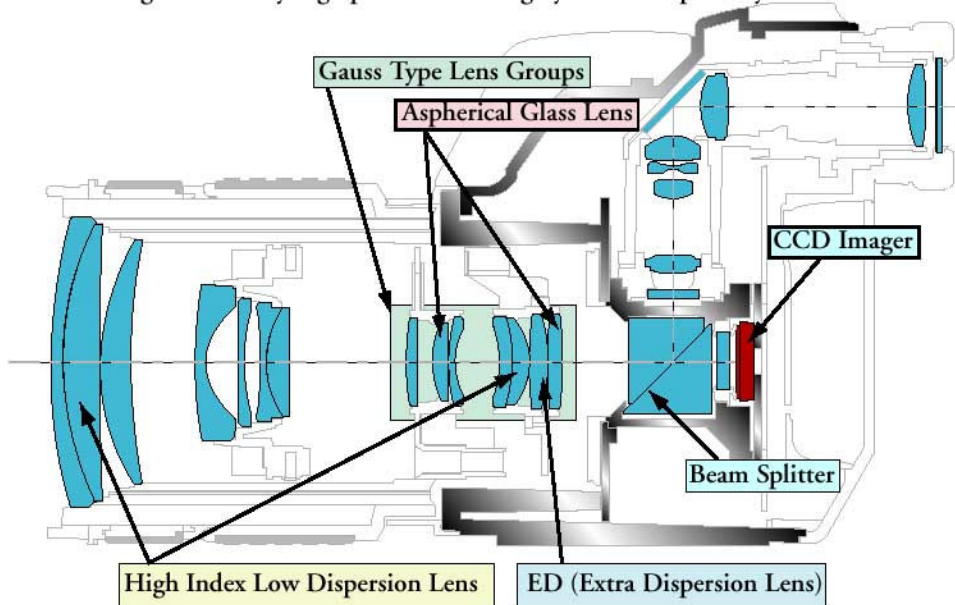
Il costo dei sistemi ottici può essere ampiamente superiore al sensore di acquisizione, soprattutto nel caso di lenti ad elevata trasmissività luminosa o per particolari bande spettrali (infrarosso termico), dove è necessario utilizzare materiali speciali (ottiche al germanio).

L'approfondimento di questi aspetti tecnici, molto complessi, esula dalle finalità della presente tesi. Si accenna solo alla generale necessità di considerare una coerente strutturazione del sistema ottico di visualizzazione con gli obiettivi tecnici ed economici dell'intero sistema di analisi delle immagini.

Nei dispositivi di *imaging* di largo consumo (es. fotocamere), il sistema ottico può essere integrato e solidale con il sensore di acquisizione e con l'elettronica di controllo (fig. 8). In questo caso, le caratteristiche principali da considerare degli obiettivi sono: lunghezza focale, apertura o luminosità, angolo di campo, messa a fuoco.

Fig. 8. Schema ottico di una fotocamera reflex digitale (<http://www.imaging-resource.com>)

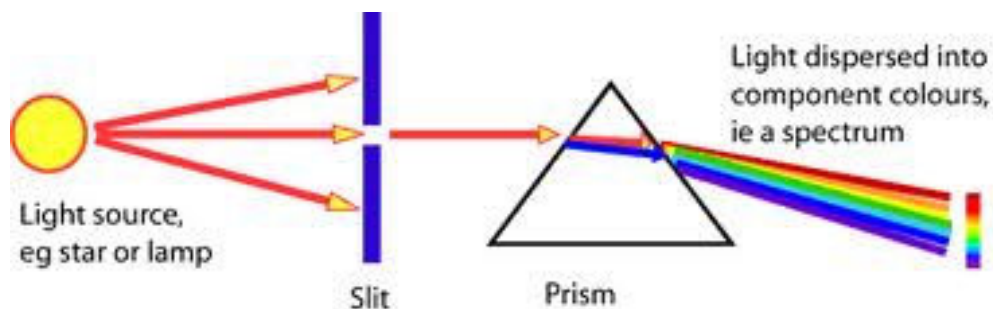
Exclusively developed 14 element, 11 group 4 X zoom for the E-10, designed as a very high precision and highly accurate optical system.



Original Image ©2000 Olympus. Modifications ©2000, Imaging Resource, all rights reserved

Un tipo particolare di ottica, che è stata utilizzata nelle attività sperimentali riportate nella presente tesi, è quella definibile spettrometrica (spettrometri). Questa ottica consente di separare le componenti spettrali incidenti e di disperderle sul sensore di acquisizione. Il principio di base di funzionamento è rappresentato dalla dispersione spettrale di un fascio di luce passante per un prisma trasparente (Fig. 9).

Fig. 9. Dispersione spettrale della luce attraverso un prisma



Anche per quanto riguarda i sensori, la forte evoluzione tecnologica in campo fotonico, mette a disposizione delle possibili applicazioni numerosi tipologie e modelli dalla svariate caratteristiche tecniche e costi.

Ci limitiamo a riportate alcune informazioni riguardo i sistemi più diffusi e strettamente relazionabili alle attività sperimentali riportate nella presente tesi.

La principale e più diffusa “famiglia” di sensori di acquisizione delle immagini digitali può essere definita con il termine di sensori cromatici. Sono sensori formati da elementi fotosensibili a semiconduttori in grado di trasformare un segnale elettromagnetico (luminoso) in un segnale elettrico che è convertito in digitale attraverso un convertitore analogico/digitale.

Le due tecnologie più diffuse sono denominate:

- CCD (Charge Coupled Device);
- CMOS (Complementary Metal Oxide Semiconductor).

I sensori basati su tecnologia CMOS sono più economici dei CCD, ma offrono una qualità dell'immagine minore. Essi rappresentano la maggioranza dei apparati di visione/acquisizione installati su telefonini, videocamere e fotocamere di basso costo. I sistemi a CCD (fig. 10), invece, sono in genere utilizzati per le video-fotocamere di qualità o industriali per applicazioni molto specializzate e per apparati di elevata qualità/risoluzione ottica.

Un sensore particolare per l'acquisizione di immagini termiche, ossia immagini basate sulla radianza termica emessa in range spettrale dell'infrarosso (tipicamente 7.5-14 μm), è il microbolometro (Fig. 11).

Fig. 10, Schema di sensore matriciale CCD (charge coupling device) per l'acquisizione digitale delle immagini (<http://www.sensorcleaning.com/whatisasensor.php>)

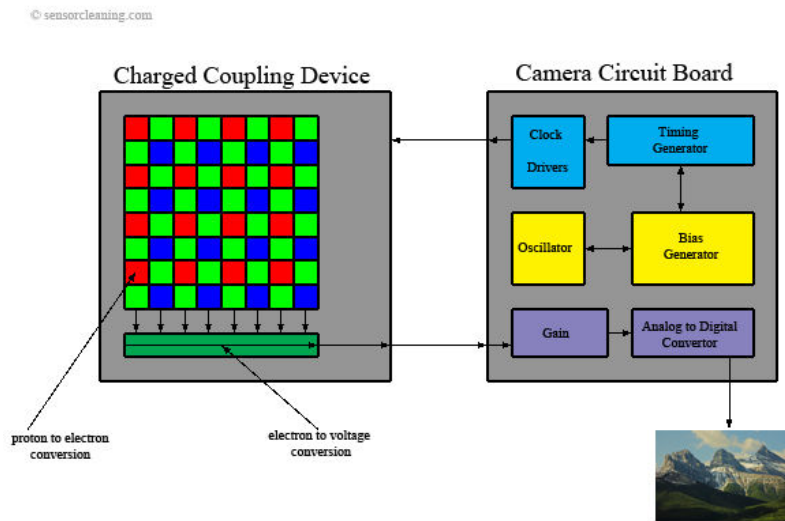
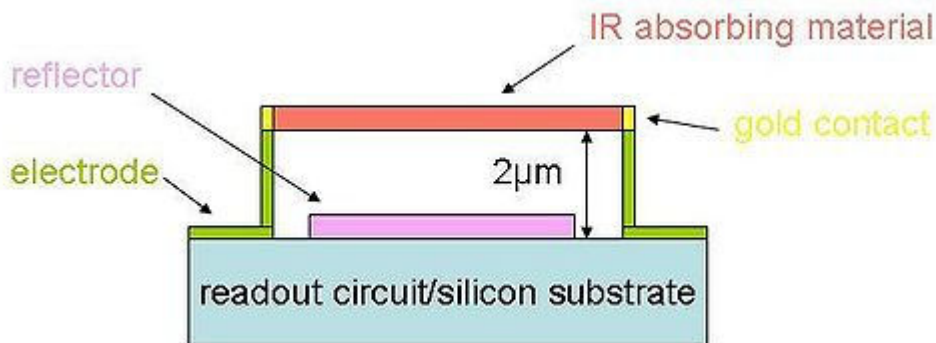


Fig. 11, Schema di funzionamento di un elemento fotosensibili (pixel) di un sensore microbolometrico non raffreddato per l'acquisizioni di immagini termiche



Il microbolometro è un sensore a stato solido non raffreddato che consiste di un array di pixel. Ogni pixel è costituito da diversi strati: un materiale assorbente la radiazione nell'infrarosso termico che riscaldandosi genera un segnale elettrico, acquisito, tramite gli elettrodi, da un substrato di silicio e un circuito integrato di lettura (ROIC). Un riflettore, ad esempio uno specchio in titanio, viene creato sotto il materiale assorbente l'infrarosso termico. Poiché una parte della radiazione IR è in grado di passare attraverso lo strato assorbente, il riflettore ridirige questa radiazione verso lo strato assorbente per assicurare la maggiore assorbimento possibile, permettendo quindi la generazione di un segnale più forte. La qualità delle immagini

create da microbolometri ha continuato ad aumentare. La matrice microbolometrica si trova comunemente in due dimensioni, 320×240 pixel o meno costosa 160×120 pixel. La tecnologia attuale ha portato alla produzione di dispositivi con 640×480 o 1024×768 pixel. C'è stata anche una riduzione nelle dimensioni dei pixel individuali. La dimensione dei pixel era tipicamente di 45 micron nei vecchi dispositivi ed è stata diminuito a 17 micron nei dispositivi attuali.

I sistemi di illuminazione

Gli apparati di illuminazione sono componenti fondamentali in molti sistemi di acquisizione delle immagini. In effetti, alcuni sistemi operano su luce o radiazione elettromagnetica naturale, come nel caso delle termocamere, ma in moltissimi casi, al fine di ottenere la visualizzazione del fenomeno e/o un'immagine più nitida, è indispensabile utilizzare una sorgente di illuminazione artificiale. Con il termine illuminazione si intende generalmente un emissione nello spettro visibile e infrarosso, ma è possibile considerare illuminanti diverse (raggi X) o segnali non elettromagnetici (es. ultrasuoni).

Il sistema di illuminazione è a volte così importante nel processo di acquisizione delle immagini, da risultare spesso molto più costoso del sensore o delle ottiche. In effetti, soprattutto nella pratica industriale (machine vision) o in quella videofotografica da interno o laboratorio, la tipologia, la disposizione e l'intensità di illuminazione, condizionano fortemente la qualità dell'immagine acquisita.

Esula dagli obiettivi della presente tesi, l'approfondimento di teorie e pratiche illuminotecniche. Ci limiteremo ad indicare alcuni parametri e caratteristiche di base che saranno utili per una migliore comprensione delle attività sperimentali di seguito riportate.

I principali elementi (caratteristiche illuminotecniche) da considerare in un sistema di illuminazione ai fini di acquisizione e analisi delle immagini digitali, sono:

- Tipo e disposizione dell'apparato acquisizione dell'immagine
 - o sensibilità
 - o geometria
- Caratteristiche della scena e degli oggetti da illuminare

- Caratteristiche ottiche della superficie (riflettente, diffondente, traslucida)
- Caratteristiche geometriche (planarità, elevata tridimensionalità) dello sfondo e degli oggetti
- Velocità di acquisizione o di movimento della scena
- Caratteristiche dell'illuminante
 - geometria di illuminazione
 - tempo di esposizione all'illuminazione (luce costante, luce pulsata)
 - curva spettrale;
 - temperatura di colore;
 - intensità

Il flusso luminoso (Lumen, lm) identifica l'intera potenza irradiata dalla sorgente di luce. Una sorgente luminosa irradia il suo flusso luminoso generalmente in diverse direzioni con diversa intensità. L'intensità della luce irradiata in una determinata direzione viene definita intensità luminosa.

L'intensità luminosa emessa da un apparecchio illuminante non è la stessa in tutte le direzioni dipendendo dal tipo di lampada e forma. Le case costruttrici ricavano empiricamente dei diagrammi detti curve fotometriche, che forniscono l'intensità luminosa nelle varie direzioni.

L'illuminamento (lux, lx) è dato dal rapporto tra il flusso luminoso irradiato e la superficie illuminata, rappresenta quindi il flusso luminoso per unità di superficie illuminata. L'illuminamento prodotto da un flusso luminoso di 1 lm che cade in modo uniforme su una superficie di 1m² è pari a 1 lx. Si dimostra che l'illuminamento è inversamente proporzionale al quadrato della distanza fra sorgente e punto illuminato

In merito alle geometrie di illuminazione, possiamo distinguere:

- diretta: è un sistema nel quale il flusso luminoso incide direttamente sugli oggetti illuminati;
- indiretta: sistema nel quale il flusso luminoso incide sugli oggetti illuminati solo dopo riflessione a strutture prossime alla scena, in misura superiore al

90%; si ottiene così un'illuminazione dolce, priva di ombre, detta anche diffusa, di elevato costo in relazione alla luce assorbita

- semidiretta: mista o diretta-indiretta, semindiretta, sistemi di illuminazione intermedi fra i due precedenti, con percentuale di flusso luminoso che incide direttamente sugli oggetti illuminati;
- senza ombre: quella prodotta da una sorgente di grande area, con forte riduzione delle ombre;
- retroilluminazione (backlighting): dove la scena e gli oggetti sono illuminati dalla parte opposta rispetto al sistema di acquisizione delle immagini (campo chiaro);

La retroilluminazione è utilizzata quando sia necessario acquisire il contorno di oggetti con elevato dettaglio. Ponendo gli oggetti su una superficie trasparente illuminata, essi appariranno completamente scuri su fondo chiaro.

Questa tecnica è stata utilizzata, ad esempio, per l'acquisizione del contorno di frutti di nocciole, mandorle o cariossidi al fine del successivo calcolo matematico del profilo secondo la tecnica di analisi ellittica di Fourier (EFA).

Nell'illuminazione diretta deve essere considerata anche la legge di Lambert, che indica come la luce si rifletta con intensità proporzionale al rapporto tra la direzione del raggio incidente e la normale alla superficie in quel punto.

Lo spettro luminoso emesso da un corpo presenta un picco di emissione determinato, in base alla legge di Wien, esclusivamente dalla sua temperatura. La temperatura di colore, di una certa radiazione luminosa, è pari alla temperatura che dovrebbe avere il corpo affinché la radiazione luminosa emessa da quest'ultimo appaia cromaticamente più vicina possibile alla radiazione considerata.

Una temperatura bassa (sempre però nell'incandescenza, intorno ai 2000 K) corrisponde ad un colore giallo-arancio. Scendendo si passa al rosso ed all'infrarosso, non più visibile, mentre salendo di temperatura la luce si fa dapprima più bianca, quindi azzurra, violetta ed ultravioletta. Quando comunemente si dice che una luce è calda, in realtà questa corrisponde ad una temperatura di colore bassa, viceversa una temperatura maggiore produce una luce definita comunemente fredda. Tale definizione ha una motivazione puramente psicologica, poiché la nostra mente tende

ad associare a colori come il rosso o il giallo-arancio l'idea di caldo ed a colori come il bianco o l'azzurro l'idea di freddo.

Elettronica di digitalizzazione, controllo e memorizzazione

Il processo di visualizzazione e acquisizione delle immagini in formato digitale è necessariamente controllato e gestito da sistemi elettronici e informatici specificamente realizzati.

I componenti principali sono:

- i sistemi di controllo dell'illuminazione e delle ottiche (quando integrate)
- i sistemi di controllo del sensore di acquisizione delle immagini
- il convertitore analogico digitale
- la memoria, temporanea e permanente di salvataggio delle immagini
- il software di interfaccia con l'utente

In funzione del grado di complessità dell'apparato di visione artificiale, l'elettronica di controllo potrà essere *embedded* su specifici moduli *hardware*, oppure richiedere veri e propri computer. In campo industriale, sono spesso utilizzati i cosiddetti PLC (controllore logico programmabile o *programmable logic controller*) (Fig. 12), che sono dei computer industriali con proprie strutture e interfacce *hardware* e *software*. In effetti, l'*imaging*, soprattutto se *real time*, genera un flusso di dati molto elevato che può essere gestito e memorizzato solo da strutture molto performanti. Inoltre, molti sistemi di machine vision, richiedono il processamento e l'elaborazione in tempo reale di tali immagini per pilotare gli organi effettuatori delle operazioni di output previste. Anche in questo caso è necessario che il sistema disponga di notevoli risorse e velocità computazionali per consentire

Fig. 12, Esempio di computer industriale (PLC) modulare, specificamente utilizzato per la gestione di un sistema di machine vision e il controllo degli organi elettromeccanici (es. nastri trasportatori, convogliatori, effettuatori)



Un altro elemento elettronico di fondamentale importanza nei sistemi di visione artificiale è il convertitore analogico/digitale (Analog to Digital Converter -ADC). L'ADC è un circuito elettronico in grado di convertire un segnale analogico con andamento continuo (ad es. una tensione) in una serie di valori discreti. La risoluzione di un ADC indica il numero di valori discreti che può produrre ed è, normalmente, usualmente espressa in bit. Negli apparati foto-video più comuni, l'ADC è embedded con l'elettronica di bordo dell'apparato stesso. Infatti, molte video o fotocamere, già forniscono in uscita immagini in formato digitale. I sistemi più sofisticati con elevati flussi dati, dovuti ad alte velocità di scansione e/o grandi dimensionalità delle immagini (es. immagini iperspettrali), prevedono, invece, l'utilizzo di convertitori analogico-digitali a bordo di schede indipendenti dal sistema dall'apparato di visione (Fig. 13). Queste schede (frame grabbers), dispongono di banchi propri di memoria temporanea per migliorare le prestazioni di transfer rate e sincronizzazione del flusso dati/immagini e necessitano di un alloggiamento su computer.

Fig. 13. Scheda elettronica per la conversione digitale di segnale analogico da videocamera e trasferimento su computer (frame grabber)



Estensione dei sistemi di analisi di immagine al dominio del tempo

I sistemi di image analysis, quando impiegati ai fini misuristici quanti-qualitativi, sono normalmente considerati come sistemi di misura multipunto e multiparametrici. Questo fatto è in relazione alla estensione spaziale dell'immagine e alle componenti morfologiche e/o cromatiche dell'immagine. Esiste, tuttavia, una terza dimensione misuristica, spesso sottovalutata o non sempre ben considerata, che si estende al dominio del tempo (*time domain image analysis*).

Dal punto di vista tecnologico, le più moderne videocamere possono acquisire immagini temporalmente spaziate anche di pochi ns, avendo frame-rate che vanno da 24 fps (frames per second) fino a 100.000 fps (high-speed cameras). La più performante telecamera arriva ad una stratosferica velocità di acquisizione pari a 200 milioni di immagini al secondo

(<http://advance.uri.edu/pacer/september2000/story9.htm> last access 23-09-2010).

Tuttavia, la capacità informativa delle videoimmagini legate al dominio del tempo sono da mettere in relazione alla specificità del fenomeni agrobiologici o industriali esplorati. Questo è particolarmente vero negli studi di cronobiologia (Aguzzi et al., 2009a, Aguzzi et al., 2009b, Menesatti et al., 2009), dove per l'impiego della (video)image analysis è necessario considerare due aspetti fondamentali: i) il ciclo temporale di potenziale presenza/assenza degli organismi o dei fenomeni indagati (*object detection*) e ii) il loro tempo di permanenza nel field of view (FOV)

determinato dalle traiettorie seguite e dalla velocità di movimento (*object tracking*). Nel primo caso, la distribuzione dell'evento può seguire andamenti di tipo completamente stocastico e per in molti casi (organismi rari) distribuzioni di tipo Poisson (*law of rare events*). In questi casi, per ottenere un campionamento efficace si dovrà estendere il tempo di acquisizione delle immagini (durata dell'acquisizione). Per l'*object tracking* è invece importante considerare una frequenza di acquisizione delle immagini che sia n volte (con $n > 4$ o 5) superiore alla probabilità media di presenza temporale dell'organismo intero nel FOV dell'immagine, nota a priori la velocità di movimento dello stesso (frequenza di acquisizione). Per la stima della frequenza minima di acquisizione, da cui derivare il valore minimo di fps della videocamera, si utilizza un approccio di campionamento basato sulla teoria di Nyquist-Shannon.

In termini più generali, possiamo quindi classificare i sistemi di image analysis rispetto al dominio del tempo, in:

- *still images*, quando la componente temporale non è considerata;
- *video imaging*, quando questa è considerata

In senso stretto, pertanto, considereremo videoimaging anche quelle situazioni a piccolo frame rate (i.e 1 frame/day o 1 frame/week) sempre che il risultato dell'analisi dell'immagine venga rapportato alla durata e alla frequenza temporale di acquisizione.

L'analisi delle immagini in situazioni di video imaging, potrà poi considerare su principali aspetti:

- *single frame acquisition*: una collezione temporalmente distanziata di immagini statiche che consentano una il rilevamento della presenza/assenza di soggetti nel tempo;
- *motion vision*: dopo l'intervallo temporale tra le immagini determinato dal frame-rate, consente non solo di rilevare presenze o assenze dei soggetti di indagine, ma anche i loro movimento nel campo immagine nel tempo.

Applicazioni frequenti di *video imaging* si trovano in ambito biologico e etologico (monitoraggio dei comportamenti di animali di interesse zootecnico) (Cangar *et al.*, 2008).

Nel primo modo, la frequenza di acquisizione è più bassa del tasso di movimento

degli animali, nel secondo, invece, è il contrario e la frequenza di acquisizione delle immagini è notevolmente più alta della velocità di spostamento entro il FOV dell'animale. Dal punto di vista computazionale, nella *single frame acquisition* avremo solo un problema di riconoscimento, mentre nella *motion vision* è necessario disporre di processamento algoritmico tale da riconoscere sia il soggetto analizzato sia le sue traiettorie di movimento (*tracking*). In particolare, il tracking richiede un calcolo attento delle traiettorie in modo da evitare il riconteggio (re-sampling) dello stesso individuo durante il flusso temporale delle immagini al fine di evitare, ad esempio, sovrastime della sua presenza e in definitiva errori di campionamento popolazionele o di biomassa. Inoltre, un attento settaggio della frequenza di acquisizione della videocamera deve essere effettuata come in precedenza indicato.

2.2.2 Processamento delle immagini (elaborazione ed analisi)

Il processamento delle immagini (*image processing*) comprende tutte le fasi elaborative ed analitiche che dal segnale/immagine portano fino all'espressione numerica o alla restituzione qualitativa cercata.

In genere, sono distinte due fasi di processamento, spesso non separate nei moderni software di image processing:

- l'elaborazione delle immagini: trattamento del segnale per ottenere una nuova immagine migliorata;
- l'analisi delle immagini: sintesi informativa dell'immagine con indicazioni quantitative o qualitative relativamente al contenuto delle stesse.

In molti casi, l'elaborazione delle immagini è propedeutica alla fase di analisi. E' il caso, ad esempio, dell'aumento del contrasto, della riduzione del rumore o delle tecniche per far risaltare i contorni di oggetti interni all'immagine.

Le finalità generali del processamento di immagine sono:

il miglioramento della qualità delle immagini ai fini di una ispezione ed interpretazione visuale da parte di esperti (analisi qualitativa) (Sheikh, 2006)

l'analisi per il riconoscimento automatico dei contenuti di un'immagine (analisi automatica quantitativa).

Nell'analisi quantitativa delle immagini (cromatiche o di derivazione iperspettrale), si succedono differenti steps elaborativi finalizzati, nel complesso, al miglioramento delle caratteristiche cromatiche (contrasti, equalizzazione, amplificazione, denoising, smoothing, ecc), all'estrazione delle parti sensibili dell'immagine stessa per differenziare oggetti e sfondo (segmentation e thresholding), per arrivare alla misura degli oggetti interessanti (morfometriche o densitometriche, texture mapping) (Bharati et al., 2004; Li et Yang, 2008; Lillholm et al., 2003).

In particolare le misure morfometriche (bidimensionali), basate sull'analisi digitale delle immagini, sono finalizzate a studiare le differenze di forma (shape) di organismi o prodotti, considerando (allometria) (Gould, 1986) o a prescindere dalla taglia (size) relativa. I principali sistemi di analisi della forma sono la morfometria geometrica (Bookstein, 1991; Rohlf, 1996) e l'analisi dei profili. La prima è basata su punti di omologia biologica o strutturale (landmarks) e, a differenza della morfometria classica che è basata sulle relazioni multiple di grandezze lineari, si basa su relazioni multivariate di distanze e angoli tra i vari punti delle shapes. L'analisi dei profili è meno potente della morfometria geometrica, viene utilizzata quando il numero di landmarks è nullo o molto ridotto (0-4 all'incirca). L'estrazione dei punti del profilo è una procedura automatizzabile. Tra le diverse tecniche si indica l'Elliptic Fourier Analysis (EFA) per studiare profili chiusi (Crampton, 1995; Daegling, 2000; Kuhl and Giardina, 1982; Lestrel, 1997).

La morfometria tridimensionale, rispetto alla morfometria tradizionale, è in grado di estrarre informazioni sulla dimensione relativa e la disposizione di punti nello spazio (Gunz et al., 2005; Slice, 2007). L'estrazione e l'analisi dei profili chiusi tridimensionali è ancora in una fase molto sperimentale. Attualmente, alcuni autori hanno cercato di adattare tecniche 2D (come la EFA) ai sistemi 3D, con tutta una serie di problemi dovuti alla complessità del calcolo e alla disponibilità di sistemi efficienti di mappatura tridimensionale di organismi e prodotti (scanner 3D).

Un'altra importante branca della visione artificiale è rappresentata dall'analisi di immagine dinamica, ossia dall'analisi di immagini in sequenza spaziale o temporale. Applicazioni importanti sono, nel primo caso, la stereovisione (Costa et al., 2006), nel secondo, l'analisi del movimento (2D o 3D) di organismi o oggetti (motion vision and detection). In particolare, l'analisi del movimento di organismi, può spesso, basarsi su

acquisizione di sequenze temporali di immagini in ambiente naturale (anche subacqueo), imponendo così limiti e difficoltà di tipo ottico (risoluzione, sensibilità), elettronico (frame rate, lunghezza della ripresa, rumore del segnale), illuminotecnico (alternanza dei cicli luce/buio, variabilità dovuta alle condizioni ambientali) e operativo (protezione dalle accidentalità naturali, esposizione agli agenti atmosferici). La procedura di elaborazione ed analisi è in questo caso, più complessa, rispetto a sequenze acquisite in ambiente controllato, e prevede una serie di operazioni (filtraggi, analisi localizzate, differenze, classificazioni) (Bruni et al., 2008; Hecker et al., 2008; Knossow et al., 2008; Peters, 2008; Scantlebury, 2007).

Un esempio di processamento delle immagini è schematizzato in figura 14.

Fig. 14. Tipica sequenza di misura in fase di analisi di immagine



Vediamo di seguito alcuni dei principali processi di elaborazione e analisi delle immagini.

Elaborazione delle immagini

La definizione di elaborazione numerica delle immagini, indica un insieme di azioni compiute da appositi algoritmi su una immagine numerica per modificarla in modo da raggiungere un prefissato obiettivo, producendo un risultato che è ancora una immagine (Ardizzone, 2010, web resource). Questa definizione è limitativa e non riscuote l'approvazione di tutti gli studiosi. Sono oggi considerati temi di elaborazione di immagine anche:

- Compressione di immagini(image coding)
- Analisi di immagini(image analysis)
- Riconoscimento di forme (pattern recognition)
- Trattamento di video (video processing)

Tra le funzioni principali dell'elaborazione delle immagini, troviamo:

- ottimizzazione del contrasto
- filtraggi
- correzione dello shading
- trasformazioni spaziali
- trasformazioni geometriche
- operazioni algebriche
- filtraggio in frequenza

L'ottimizzazione del contrasto è una tecnica molto utilizzata per correggere problemi di limitata dinamica dei livelli di grigio. In genere, un'immagine poco contrastata è stata acquisita in condizioni non idonee di luce. Tra procedimenti più utilizzati per la correzione del contrasto, possiamo considerare l'accentuazione lineare del contrasto, che viene usata anche come normalizzazione del contrasto. I punti più chiari e quelli più scuri dell'immagine vengono calcolati automaticamente. Il campo dinamico che ne risulta viene poi esteso alla maggior quantità possibile di livello di grigio del sistema.

Nel caso dell'accentuazione non lineare del contrasto, determinate regioni d'intensità, entro le quali si trovano le strutture di interesse, possono venir opportunamente estese.

Nella linearizzazione dell'istogramma, le intensità presenti nell'immagine vengono valutate in base alla loro frequenza.

I livelli di grigio, le cui aree sono maggiormente diffuse, vengono contrastati più fortemente di quelli, che presentano solo una partecipazione ridotta all'immagine nel suo complesso. In pratica, accade molto spesso che le strutture di interesse si presentino molto scure rispetto allo sfondo. Modificando la caratteristica del livello di grigio (la cosiddetta funzione gamma), il contrasto può venir rinforzato in forma selettiva nei livelli di grigio più scuro e ridotto nello sfondo, che riveste poca importanza.

Il filtraggio digitale appartiene ai metodi di maggiore efficacia per l'ottimizzazione delle immagini. I filtri sopprimono principalmente il rumore, evidenziano strutture di interesse o rivelano automaticamente i contorni. Si tratta di operazioni di matrice, vale a dire che ogni punto immagine del risultato finale è calcolato in base ai punti immagine contigui al punto stesso sull'originale. I filtri usati sono di tipo passa-alto, passa-basso e passa-banda, lineari e non, o sono combinazioni di filtri per la soppressione del rumore, l'accentuazione del dettaglio o l'eliminazione di disturbi. La correzione dello "shading" serve a ridurre dalle immagini disomogeneità e strutture estranee, che falserebbero il risultato di una misurazione automatica. Può trattarsi di una illuminazione poco omogenea, di ombreggiature o riproduzioni di impurità nel percorso ottico. Per la correzione dello shading viene generata una cosiddetta immagine di riferimento (background), in cui è rappresentata solo la funzione degli errori. Può essere ad esempio un'immagine "vuota" o un originale sottoposto a filtraggio passa-basso (Zeiss, 1990).

Il filtro molto efficiente per la background correction, quando non sia disponibile un'immagine di riferimento è il cosiddetto filtraggio *Top Hat* che utilizza un filtro morfologico per estrarre la "grana" di sfondo dell'immagine associata ad una subtraction tra l'immagine originale e quella filtrata. Molto importante la dimensione del filtro morfologico: un valore piccolo esalta i dettagli e quindi la subtraction creerà un'immagine praticamente nera. Più è alto il valore (in proporzione alla dimensione

dell'immagine) e più lo sfondo non considera i piccoli oggetti; il filtraggio è più generale, ma meno effettivo (Russ, 1995).

Le funzioni di variazione delle dimensioni delle immagini includono funzioni come ingrandimento o impiccolimento (zoom), trasposizione di immagini e copiatura di porzioni di immagini in diverse memorie.

La trasformazione geometrica viene usata per la correzione di deformazioni (ad esempio distorsioni prospettiche) e di aberrazioni ottiche nella formazione delle immagini.

Attraverso le operazioni aritmetiche è possibile associare tra loro numerose immagini. Mediante addizione le immagini vengono sovrapposte, con la sottrazione vengono filtrate le differenze. Moltiplicazione e divisione vengono usate per la manipolazione delle caratteristiche o per la mascheratura di determinate regioni.

I filtri digitali sono tutti realizzati in funzione locale, vale a dire che si tratta di operazioni di matrice, che vengono eseguite sull'originale. Una variante è invece rappresentata dal filtraggio in "funzione della frequenza". La trasformazione più nota - che viene usata spesso per l'elaborazione monodimensionale del segnale - è la trasformazione di Fourier (FFT -Fast Fourier Transformation). Essa converte un'immagine in una serie di oscillazioni sinusoidali e cosinusoidali, aventi una frequenza inversamente proporzionale alla risoluzione dell'immagine stessa, ottenendo lo spettro della frequenza. In tal modo è possibile riconoscere meglio determinate caratteristiche dell'oggetto. Questo spettro può inoltre essere elaborato successivamente (mascheratura, filtraggio), permettendo così, mediante un trattamento selettivo dei coefficienti di Fourier, la realizzazione di funzioni atte al miglioramento ed al restauro dell'immagine, che non sarebbero state possibili nella normale immagine grigia.

Un sistema di analisi efficiente deve essere in grado di eseguire numerose altre funzioni di elaborazione di immagini.

Analisi delle immagini

Le operazioni di analisi delle immagini, comprendono, invece, funzioni la cui principale finalità è quella di ridurre e sintetizzare la complessità informativa dell'immagine, estraendo i dati più significativi finalizzati all'obiettivo prefissato.

A differenza dell'elaborazione, con l'analisi si parte da un'immagine per ottenere un singolo valore o al più un array di valori, statistici o descrittivi.

Tra le principali operazioni possiamo citare:

- statistiche
 - o estrazione di valori di pixel
 - o istogrammi di frequenza
 - o correlazioni tra immagini
 - o valori medi, mediani e deviazione standard
- binarizzazione dell'immagine
 - o segmentazione oggetti e sfondo
 - o riconoscimento di strutture (tracciamento di contorni, identificazione di bordi, estrazioni di linee, misure di diametri, grandezze)
- analisi della texture
 - o entropia (locale o dell'intera immagine)
 - o graylevel Gray-level co-occurrence matrix
 - o intensità cromatiche e colorimetriche di oggetti

Uno dei procedimenti di sintesi informativa delle immagini, più importante è la segmentazione. La segmentazione delle immagini è l'identificazione e l'isolamento della regione di interesse all'interno dell'immagine; in termini pratici, la suddivisione dell'immagine in oggetti (da misurare) e sfondo.

Il risultato della segmentazione è di regola un'immagine binaria, bivalente a due livelli cromatici (1 bit), in genere, bianco e nero.

Il passaggio dall'immagine grigia a quella bianca può avvenire in molti modi diversi; a seconda della complessità dell'immagine, possono essere utilizzati vari tipi di segmentazione.

Il procedimento più semplice è la cosiddetta discriminazione del livello di grigio. In questo caso si determina, a mezzo di una o più soglie del livello di grigio, in quale regione di intensità si trovino gli oggetti da misurare; tutto il resto viene definito sfondo.

E' condizione essenziale che gli oggetti risaltino sullo sfondo, grazie alla loro luminosità che potrà essere migliorata in fase di elaborazione preliminare. Questo tipo di sogliatura fissa (*fixed thresholding*) può essere anche applicata ad immagini a colori (RGB), ma presenta un problema di corretta scelta della soglia quando si debba automatizzare il processo di segmentazione.

Un algoritmo di segmentazione completamente autonomo fornisce segmentazione obiettive e coerenti, ma in generale la segmentazione automatica è uno dei compiti più difficili nella elaborazione delle immagini (Gunasekaran, 1996). Tan (2004) ha rilevato che la segmentazione affidabile e coerente di immagini di prodotti alimentari (carne) è una sfida continua. In molti casi segmentazione manuale è a volte necessaria per ottenere risultati di elevata qualità. Altri autori, Jackman et al (2008) hanno sviluppato metodi automatici combinati ad operazioni manuali per completare una segmentazione efficiente di immagini complesse di prodotti alimentari. Lo svantaggio di segmentazione manuale è che i dati successivi è fortemente dipendente dalla precisione e la riproducibilità di questa operazione manuale (Jackman et al. 2009).

Diverse tecniche e algoritmi sono stati proposti per effettuare una segmentazione automatica delle immagini.

I principali metodi possono essere distinti in:

- metodi basati sul pixel (*Pixel-based*), che utilizzano i valori di luminosità dei singoli pixel;
 - o *histogram based*: la soglia è scelta sulla base dei valori di minimo o massimo, assoluto o locale della distribuzione di frequenza dei livelli di grigio dei singoli pixel;
 - o metodo OTSU (Otsu, 1979), che identifica la soglia di segmentazione analizzando la qualità della distribuzione dei livelli di grigio in modo che sia massima la varianza interclasse e minima quella intraclasse tra la due classi (oggetto e sfondo) considerate
 - o metodi di clustering non supervisionati: sono metodi di analisi multivariata che effettuano raggruppamenti di pixel in base a similarità specifiche (es. valori sopra o sotto soglia di distanze euclidee) oppure di distanza o prossimità da centroidi di classe (*kmeans clustering* – Menesatti et al., 2008);

- metodi basati sulla regione (Region-based), dove si analizzano i valori di luminosità di aree dell'immagine;
 - o region growing” si basa sul raggruppamento automatico di regioni di pixel che hanno una proprietà in comune; ad esempio gruppi di pixel che hanno la medesima luminosità;
- metodi basati sul contorno (Edge-based), si estraggono i contorni per delimitare gli oggetti;
 - o algoritmo di gradiente-based, che individua bruschi cambiamenti locali di una particolare proprietà di pixel e quindi collega i pixel in un bordo, con i pixel all'interno o all'esterno del bordo che sono considerati sfondo o oggetti;
- metodi basati su un modello (model-based), si parte dalla conoscenza a priori del modello dell'oggetto
 - o algoritmo KNN (K Nearest Neighbours), operante su due classi, sfondo e oggetti (Shakhnarovich *et al.* 2005).

Il KNN è un classificatore multivariato supervisionato che assegna il valore di classe ad un oggetto non classificato, sulla base della classe più frequente tra i primi k più vicini in termini di distanza euclidea, estratti da un insieme le cui classi sono note con certezza (training set). L'algoritmo KNN, è stato utilizzato per la segmentazione efficiente di immagini a colori RGB, relative alle impronte di gocce di trattamenti fitosanitari in prove di qualità di distribuzione di irroratrici ad aeroconvezione (Pergher et al., 2008). Attraverso tale algoritmo, ogni pixel di ciascuna immagine è stato classificato, per k=3, appartenente ad una delle due classi, sulla base del valore delle tre coordinate RGB. Il training set è stato costruito calcolando il valore medio di ogni coordinata RGB per alcune aree omogenee per classe (sfondo e impronte delle gocce) estratte manualmente da un operatore su una immagine a caso. Lo stesso algoritmo è stato utilizzato per la determinazione attraverso image analysis del grado di pelatura di diverse cultivar di nocciole (Pallottino et al., IN PRESS).

L'immagine segmentata rappresenta il punto di partenza per una valutazione misuristica degli oggetti identificati o serve da maschera per una successiva combinazione di immagini. Di regola, l'"oggetto" (che deve venir misurato) viene rappresentato in bianco e lo "sfondo" in nero.

Dopo la segmentazione può essere necessario modificare ulteriormente l'immagine binaria, se ad esempio non è possibile rilevare con precisione gli oggetti, a causa del contenuto troppo complesso dell'immagine.

In questo caso, le operazioni più importanti sono:

- erosione (erode): eliminazione di un determinato numero di pixel lungo i bordi degli oggetti
- dilatazione (dilation): opposta all'erosione
- apertura (opening) viene effettuata dapprima un'erosione, seguita da una dilatazione; questa operazione viene usata per separare strutture o eliminare piccoli oggetti di disturbo dall'immagine (pulizia dell'immagine binaria);
- chiusura (closing): funzione opposta all'opening che si compone di una dilatazione seguita da un'erosione; il closing viene impiegato per riunire strutture separate erroneamente o per chiudere piccoli buchi ed aperture;
- riempimento (fill): con essa vengono riempiti automaticamente fori presenti nell'oggetto; per effettuare una classificazione geometrica dello stesso (es. area);
- scheletrizzazione (skeletoning): le strutture vengono ridotte a linee.

Quando, nel corso del procedimento di analisi, si ottiene una immagine binaria ottimale, l'analisi vera e propria può essere considerata conclusa. Ad essa seguono le funzioni tecniche di misurazione. Per i parametri, che devono venir misurati, esistono diversi criteri di differenziazione.

Può trattarsi di parametri specifici, che vengono determinati per ogni regione dell'immagine, indipendentemente da quante strutture singole vi siano contenute. Parametri specifici possono essere ad esempio il "numero delle strutture singole", la "incidenza percentuale delle aree delle fasi rilevate", la "somma di tutti i perimetri" o la "distribuzione di lunghezza delle corde".

Una seconda categoria di parametri è quella specifica dell'oggetto; in essa i parametri vengono determinati per ogni struttura singola dell'immagine. Ne sono un esempio l'"area", il "perimetro" ed il "diametro massimo". Oltre a questi parametri geometrici, esistono anche grandezze densitometriche e fotometriche, che possono venir dedotte dall'immagine, a patto di poter disporre dell'informazione del livello di grigio.

Parametri tipici in questo caso sono la "trasmissione", l'"estinzione", la "densità ottica integrata" e la "riflessione".

Accanto a questi parametri esistono anche dei procedimenti speciali, che si avvalgono di metodi statistici, per rendere possibili indicazioni anche complesse di caratteristiche geometriche e topologiche.

Tra queste indichiamo le misure morfometriche di base e quelle derivata dall'analisi di geometria morfometrica o analisi matematica dei profili.

L'insieme di queste tecniche è descritto con maggiore dettaglio nel capitolo inerente l'analisi della forma attraverso le immagini.

In ultimo, un accenno all'analisi della texture delle immagini. Il significato di 'struttura' (texture) nella computer vision è completamente diverso dal significato usuale di texture negli alimenti (Iqbal et al, 2010).

La texture visiva può essere definita come l'organizzazione spaziale delle variazioni di intensità in un'immagine a lunghezze d'onda diverse, come ad esempio le porzioni visibili e infrarosse dello spettro (Haralick et al., 1973).

Per un prodotto alimentare (ad esempio la carne), le caratteristiche tessiturali definite dalle mappe di intensità dell'immagine, possono essere collegate alle caratteristiche visive sensoriale dell'alimento (ruvidità consistenza, regolarità, ecc), ma anche a proprietà interne o visivamente non percepibili (Basset et al., 2000; Quevedo & Aguilera, 2008). La texture di un'immagine è una proprietà prevalente di tutte le superfici alimentari, dunque è di fondamentale importanza disporre di metodi robusti ed efficienti per l'elaborazione delle immagini con texture (Tuceryan, 1998).

2.3 Analisi multivariata delle immagini e delle informazioni

Le immagini, come le informazioni da esse estratte, sono complessi array di dati, spesso a più variabili (multivariati).

Tra le moderne tecniche di analisi dei dati, anche sotto forma di immagini, abbiamo quelle complessivamente indicate come tecniche di analisi multivariata o chemometriche (Forina, 2006).

Queste tecniche sono state utilizzate in tutti i lavori sperimentali riportati nella presente tesi, per tale motivo si riporta di seguito una breve presentazione delle principali analisi multivariate.

Le tecniche di analisi multivariate sono generalmente suddivise in due principali categorie: quelle non supervisionate e quelle supervisionate. Nelle tecniche non supervisionate il raggruppamento o la suddivisione in cluster degli elementi multivariati (Xblock) è effettuata sulla base di relazioni funzionali tra gli elementi stessi (distanze, varianze). Esse non necessitano nella fase di training della conoscenza del campione analizzato. Nelle tecniche supervisionate, invece, il riferimento di attribuzione di classe è dato da una variabile risposta singola o multipla (Yblock). In tal senso, i modelli multivariati forzano l'attribuzione verso una prestabilita classe e sono, in genere, più performanti dei sistemi non supervisionati. Le tecniche non supervisionate, sono maggiormente utilizzate in senso esplorativo, quando l'obiettivo sia quello di analizzare o visualizzare (es. PCA) le relazioni di aggregazione non forzata (supervisionata) tra gli elementi.

Relativamente alle tecniche di classificazione e clustering supervisionate, è possibile distinguere due approcci di base: il modellamento e la classificazione vera e propria (Forina, 2006i). Nella classificazione, il modello attribuisce l'elemento analizzato ad una delle classi predeterminate. E' necessario che siano previste almeno due classi di attribuzione. Nel modellamento, invece, è possibile attribuire l'elemento considerato ad una delle classi predeterminate o a nessuna di esse; in questo caso se il campione non è attribuito alla classe si indica come outlier. Nel modellamento, si calcola la "probabilità di predizione" e una soglia di classificazione per ciascuna classe modellata.

L'efficienza di modellamento è indicata dai parametri statistici di sensibilità e specificità. La sensibilità (sensitivity) è la percentuale di specie della categoria accettata dal modello di classe. La specificità (specificity) è la percentuale di specie di categorie differenti dal modello, quindi rigettate dal modello di classe.

Nella classificazione, invece, si utilizza la confusion matrix o matrice di corretta classificazione.

I modelli supervisionati includono generalmente una fase di calibrazione (training) ed una di test di validazione, e per entrambe le fasi, possono essere calcolati i parametri di classificazione ed efficienza.

Per questioni di sintesi, si riportano di seguito alcuni cenni relativi ai due principali classificatori utilizzati nelle tecniche di analisi delle immagini: la PCA – *Principal component analysis* – tipo non supervisionato e la PLS *partial least square* – tipo supervisionato.

La PLS è una tecnica utilizzata per operazioni di predizione quantitativa, ma nella funzione PLS-DA (*partial least square – Discriminant Analysis*) è impiegata per la classificazione qualitativa.

Principal Component Analysis - PCA

Lo scopo della PCA (Wold, 1987) è di descrivere un pattern di dati in uno spazio le cui dimensioni siano più piccole del numero di features (variabili). La riduzione deve limitarsi a uno spazio a due o tre dimensioni per visualizzare facilmente similarità o differenze. Le differenti coordinate dello spazio ridotto sono non correlate e ortogonali e rappresentano informazioni completamente indipendenti.

La tecnica della PCA consiste nella trasformazione lineare della matrice X di dati in un sub-set che rappresenta l'intera informazione di X. Ciò è reso possibile dal fatto che i dati sono "centrati" rispetto alla media (normalizzati) e che quindi tutta l'informazione è trasferita nella matrice di covarianza. Si assume anche che il data-set sia distribuito normalmente. Attraverso la matrice di covarianza è possibile ridistribuire i vettori dati (pattern) in un sottospazio i cui assi sono combinazione dello spazio di origine (Principal component). Praticamente la PCA lavora sulle varianze interne dei vettori misura (pattern).

Rispetto la matrice di dati la PCA fornisce due informazioni principali: *loading* e *score*. Lo score riguarda la misura ed è la coordinata del vettore di misura nello spazio delle componenti principali. Il loading concerne la singola variabile x (la features) e indica quanto questa contribuisce alla varianza delle PC (componenti principali). Il Residual rappresenta la variabilità non spiegata, ossia indica il rumore. Attraverso la PCA è, pertanto, possibile verificare il grado di distribuzione in

sottogruppi del data set, operata in maniera autonoma dal modello, senza apportare conoscenza *a priori* in merito alla ipotesi sperimentale di separazione in gruppi.

Partial Least Squares (PLS) e Partial Least Squares Discriminant Analysis (PLS-DA)

La Partial Least Squares (PLS) (Geladi and Kowalski, 1986) rappresenta una delle tecniche statistiche più affidabili, utilizzate per cercare le correlazioni tra i segnali di uscita di uno strumento multicanale e le informazioni contenute nelle N misure stesse. In particolare il modello è utilizzato in senso predittivo per stimare quantitativamente il valore dei parametri chimici di riferimento.

Il modello opera attraverso apposito algoritmo (SIMPLS, De Jong 1993) sull'insieme delle informazioni, che nel nostro caso erano di tipo spettrale (variabili di input, matrice X, Xblock) e sia sui valori quantitativi di riferimento (parametri chimici) sia sulle classi funzionali osservate (classe di stagionatura, Y). Poiché i valori appartenenti ad uno stesso data-set sono legati da complessi rapporti funzionali, il modello, attraverso una procedura iniziale di calcolo di massime correlazioni, determina, tramite un processo ricorsivo, l'insieme minimo di n variabili di stima (*latent and score variable*) le quali sono rappresentabili in uno spazio n-dimensionale e che sono utilizzate dalla PLS per calcolare la miglior matrice di regressione (B) tra le X e le Y. La matrice B ottima è quella che minimizza la matrice degli errori di stima del modello (E), secondo l'equazione matriciale:

$$Y = BX + E \quad (1)$$

Il modello comprende una fase di *trainig* e una fase di *test di validazione*, e per entrambe è possibile calcolare gli errori per ciascuna fase. L'andamento degli errori è, in genere, decrescente per la fase di training (RMSEC) e crescente per la fase di crossvalidazione (RMSECV). Il numero di variabili latenti utilizzato si basa sulla minimizzazione contemporanea di questi due tipi di errori, oppure sulla migliore performance del test.

Attraverso la matrice B, sono stimati i nuovi valori di Y (valori stimati, matrice Y') espressi o come etichetta di classe - class label oppure come valore quantitativo. Il confronto tra i valori osservati Y e i calcolati Y' consente di valutare il grado di corretta classificazione a priori o l'efficienza di predizione dei valori osservati.

L'algoritmo di base PLS può essere utilizzato per due tipologie analitiche:

- 1) predizioni quantitative
- 2) classificazioni o modellamenti.

Il primo caso, è rappresentato tipicamente dalla stima di un parametro quantitativo, chimico o fisico (misurato in modo tradizionale), tramite le misure multicanale (in questo caso, spettrali) che sono state rilevate sul medesimo campione. Il grado di correttezza della stima deve essere dedotto tramite il confronto diretto delle Y misurate con quello delle Y' stimate, calcolando differenti parametri indicativi della efficienza di predizione, come:

- coefficiente di correlazione (r) tra valori predetti e misurati
- RMSE (Root Mean Square Error)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{I_p} (\hat{y}_i - y_i)^2} \quad (2)$$

- SEP (standard error of prevision)

$$SEP = \sqrt{\frac{1}{n} \sum_{i=1}^{I_p} (\hat{y}_i - y_i - bias)^2} \quad (3)$$

dove:

- ✓ \hat{y}_i - è il valore predetto di ogni osservazione;
- ✓ y_i - è il valore misurato di ogni osservazione;
- ✓ n - numero di osservazioni nel set dei valori predetti;
- ✓ Bias - differenza sistematica tra I valori misurati e predetti.

essendo:

$$bias = \frac{1}{n} \sum_{i=1}^{I_p} (\hat{y}_i - y_i) \quad (4)$$

In altre situazioni, si ha la necessità di sviluppare un modello PLS che sulla base delle informazioni multicanale (spettro) possa attribuire il campione di misura a:

- 1) ad una ed una sola delle classi predeterminate (classificazione);
- 2) ad una delle classi predeterminate o a nessuna di esse (modellamento); in questo caso se il campione non è attribuito alla classe si indica come *outlier*.

Per questa tipologia di analisi, il modello multivariato è più correttamente chiamato PLSDA, Partial least Square Discriminate Analysis.

La PLSDA non è altro che un'analisi di regressione PLS nella quale la variabile "risposta" è categorica (y-block; sostituito da un set di dummy-variables (1,0) per ogni categoria) esprimendo la classe di appartenenza delle unità statistiche (Costa et.al., 2008).

In classificazione, la PLSDA non ammette, come variabili risposta, altre che non siano quelle che definiscono i gruppi di individui (i modelli costruiti dalla PLS sono resi, dalla PLSDA, categorici, quindi discriminanti). Come conseguenza, tutte le variabili misurate giocano lo stesso ruolo rispetto alla classe assegnata. La PLSDA stabilisce la percentuale di corretta classificazione e il peso di ogni pixel in ciascun latent vector (LV) (vale a dire nei parametri principali che descrivono il modello). L'abilità di predizione della PLSDA dipende anche dal numero (k) di LV usati nel modello. Il valore ottimale di k viene determinato dalla predizione dei risultati per gruppi indipendenti di campioni (test set) per differenti valori di k , e determinando quel valore per il quale la più alta percentuale di corretta classificazione è stata trovata nel test set.

Nel modellamento, la PLSDA calcola la "probabilità di predizione" e una soglia di classificazione per ciascuna classe modellata. Queste sono calcolate usando il metodo Bayesiano che essenzialmente prende i valori y predetti dal modello PLSDA, adatta ad essi una distribuzione normale e poi usa ciò per calcolare la probabilità di osservare un dato valore y . Il metodo stabilisce la percentuale di corretta classificazione ed il peso di ciascuna specie in ogni LV. Questa analisi esprime anche i parametri statistici indicanti l'efficienza di modellazione segnalata dalla sensibilità e dalla specificità dei parametri. La sensibilità (*sensitivity*) è la percentuale di specie della categoria accettata dal modello di classe. La specificità (*specificity*) è la

percentuale di specie di categorie differenti dal modello, quindi rigettate dal modello di classe.

Anche nella PLSDA, il modello include una fase di calibrazione (*training*) ed una di test di validazione, e per entrambe le fasi, può essere determinata la percentuale di corretta classificazione calcolando gli errori residui.

2.4 Applicazioni biofotoniche

Il settore tecnologico dell'optoelettronica e quello applicativo della biofotonica è in forte crescita (Urbani et al., 2003) come indica la ricerca di citazioni bibliografiche, relative alla parola chiave "image analysis", effettuata su "google scholar" per soggetti di area agricola, biologica e ambientale (anni 2000 – 2007). Essa fornisce l'indicazione di circa 64000 referenze, che diventano 555000 estendendo la stessa analisi ai settori Chemistry, Engineering, Computer science e Medicine. Inoltre, la produzione bibliografica è cresciuta negli ultimi anni, in relazione alle maggiori disponibilità tecnologiche e prestazioni H&S, quasi raddoppiando, per i soli riferimenti all'area agricola, biologica e ambientale, passando dall'anno 2000 al 2007.

Come già indicato la tecniche di *machine vision*, si distinguono, da altre tecniche e tecnologie puntuali, per offrire specifiche caratteristiche operative (Urbani and Menesatti, 2002):

- elevata produttività (efficienze ed efficacia);
- non-distruttività delle misure, spesso di non contatto;
- multiparametricità;
- ampia banda informativa;
- trasportabilità e oprtabilità;
- speditività;
- integrabilità con altri sistemi elettronici e/o con i sistemi informativi aziendali.

L'utilizzo di questi sistemi nell'ambito dell'agricoltura, ingegneria e meccanica agraria, sono in forte crescita. Anche la letteratura scientifica e le attività divulgative

e congressuali, considerano questo settore disciplinare con sempre maggiore interesse (Tillet, 1991; Menesatti, 1995a; Menesatti, 1995b; Menesatti, 1996; Menesatti, 1997; Fedrizzi and Menesatti, 1997). Nel 2009 è stato istituito, ad esempio, il gruppo di lavoro *Image Analysis*, nell'ambito dell'Associazione Internazionale di Ingegneria Agraria (CIGR) (<http://www.atb-potsdam.de/CIGR-ImageAnalysis/board.htm>).

Data l'elevata numerosità di lavori scientifici pubblicati e la varietà di applicazioni, si riporta di seguito una sintesi delle principali attività citate in letteratura, ripartite per i seguenti settori afferenti all'ingegneria agroalimentare e dei biosistemi (Menesatti et al., 2000;):

- applicazioni per l'ambiente, il suolo, gli animali zootecnici e le piante agrarie
- macchine agricole
 - o robotica
 - o motrici e operatrici (in funzioni pre e raccolta)
 - applicazioni dirette
 - testing e certificazione
- macchine per il post-raccolta
- food processing e qualità dei prodotti agroalimentari

Imaging ambiente, suolo, piante, animali

Animali zootecnici

In ambito zootecnico, oltre ai robot per la mungitura, la computer vision è stata ampiamente utilizzata in ricerca e attività inerenti lo studio del benessere (Barbera, 2003), del comportamento e della stima del movimento, del peso e della conformazione degli animali. La termografia all'infrarosso è stata utilizzata per monitorare variazioni spaziali e temporali della temperatura superficiale di animali di interesse zootecnico, al fine di studiare stati fisiopatologici legati alla definizione di benessere animale (Menesatti et al, 2004; Shao, 2008; Xin, 1999) o come tecnologie speditive e di non contatto per l'identificazione di stati di alterazione patologica (Edgar et al., 2009). Sistemi di computer vision o image analysis nel visibile sono

invece stati studiati nelle applicazioni per lo studio comportamentale di galline ovaiole (Barbosa et al., 2006; Barbosa et al., 2007; Barbosa et al., 2008) e nella loro distribuzione spaziale negli allevamenti a terra in funzione di situazioni di comfort o stress termico (Rodrigues et al., 2008). Cronin (2008) ha studiato un sistema per immagini per contare le galline in gabbia e per controllare l'integrità delle uova sui nastri di raccolta. Lo stesso autore indica come attualmente ci sia un notevole dispendio di personale per la valutazione quotidiana soggettiva dello stato di benessere delle ovaiole. Il sistema di computer vision è riuscito a riconoscere correttamente gli animali al 79% e identificare il 95% dei corpi estranei tra le uova.

Leuschner (2005) ha utilizzato la computer image analysis per lo studio della crescita e del comportamento dei suinetti, Ipema et al. (2004) hanno monitorato le condizioni di comfort delle vacche da latte durante la mungitura, mentre Souza et al. (2008) hanno studiato il comportamento di vacche da latte in relazione a variazioni microclimatiche con una serie di microcamere dimostrando come la computer vision sia uno strumento molto efficace per tali studi. Duarte et al (2009), invece, attraverso image processing, hanno invece studiato il comportamento di pesci allevati (sogliole), monitorando l'attività di movimento degli organismi e trovano una elevata corrispondenza tra la procedura automatizzata e l'osservazione diretta di operatori esperti. Brandl (1997) e Pastorelli et. al (2006) hanno studiato i movimenti e il posizionamento dei maiali in allevamento come sistemi per studiare il comportamento ed associarlo alle funzioni vitali degli animali.

Molto importanti sono gli studi volti a definire e applicare metodologie per la stima indiretta del peso degli animali zootenici al fine di monitorare lo stato di ingrasso senza interventi stressanti sugli animali che possono essere faticosi e potenzialmente pericolosi per gli operatori. La maggior parte degli studi riguarda l'analisi del peso dei maiali (Yang et al., 2006; Fu et al, 2008; Rydberg and Gilbertsson, 2005; Kollis et al, 2007), sebbene Vranken et al (2005) abbiano sviluppato un metodo per la pesatura del polli che riduce dal 14% al 5% l'errore di stima del peso riferito all'ultimo giorno di ingrasso. Infine, l' image processing è stata anche utilizzata per l'analisi dimensionali delle carcasse e la relazione della loro composizione con le dimensioni dei maiali vivi (Doeschl-Wilson et al., 2008) e dell'influenza delle caratteristiche

genetiche sulla conformazione e sul punteggio di distribuzione del grasso (fat score) di agnelli (Rius-Vilarrasa et al., 2009, Rius-Vilarrasa et al., 2010).

Suolo agrario

Le applicazioni di computer vision e soprattutto di analisi di immagine multi o iperspettrale sono relativamente numerose in campo agroambientale, sia per il monitoraggio e il controllo del suolo che delle piante e delle acque.

La maggior parte di questi studi sono inerenti il remote sensing, ossia un controllo ambientale da remoto (satellite o aereo) che data la particolare specificità non saranno trattati in questa sede.

Tuttavia, con riferimento all'analisi iperspettrale di immagine, il rapido progresso tecnologico ha messo recentemente a disposizione sistemi potenti e sufficientemente portatili da poter essere utilizzati anche in studi di proximal sensing "ground based". Tra questi possiamo citare i primi studi di caratterizzazione della parte superficiale di suoli contaminati da rame (Antonucci et al., 2010), in cui la stima del contenuto in rame è effettuata attraverso un sistema di analisi di immagine iperspettrale VIS-Nir. La determinazione speditiva dell'inquinamento da rame può rappresentare una importante informazione per la corretta gestione dei vigneti, soprattutto di indirizzo biologico. In precedenza, Bonifazi et al. (2003) aveva utilizzato il medesimo sistema spettrale per immagini per la determinazione di stati di alterazione microbiologica di suoli (Bonifazi et al., 2004), di loro contaminazione chimica (Bonifazi et al., 2005) o per la caratterizzazione di pietre ornamentali (Bonifazi et al., 2005). Più recentemente, O'Rourke et al (2010) ha dimostrato la possibilità di ottenere buone predizioni del contenuto in carbonio organico di suoli forestali ottenendo in validazione in relativamente basso, errore quadratico medio pari a 14.6 g kg⁻¹.

Piante

In ambito vegetale, le tecniche biofotoniche (Menesatti et al., 2007) sono state ampiamente utilizzate per la caratterizzazione e l'identificazione specifica e varietale, per la stima speditiva dello stato di carenza nutrizionale o dello stress idrico, o, ancora, per la misura delle dimensioni fogliari o della chioma (canopy).

Molte di queste attività, soprattutto l'identificazione specifica o varietale Camargo Neto et al, 2006) sono state basate su tecniche di rilievo morfologico (tradizionale, geometrico o analisi dei profili) e saranno trattate con maggiore dettaglio nel prossimo capitolo.

La termografia, ad esempio, è stata utilizzata per monitorare la temperatura della chioma (Wang et al., 2010) con il fine di predisporre sistemi automatici per il monitoraggio dello stress idrico, mentre Chaerle et al, (1999) hanno utilizzato questa tecnica per la visualizzazione presintomatica delle interazioni pianta-virus su tabacco (tobacco mosaic virus - TMV). Gli stessi autori, successivamente, hanno valutato la termografia in combinazione con fluorescenza della clorofilla per immagini per giungere a un maggiore dettaglio di analisi per la stessa problematica (Chaerle et al, 2004).

Tecniche spettrali estese agli intervalli del visibile e vicino infrarosso (400-1000 nm), sono state invece utilizzate per la stima del contenuto fogliare in azoto di piante di arancio (Menesatti et al, 2008) e pomodoro (Ulissi et al, 2010) con risultati estimativi superiori alle tecniche ottiche consolidate (SPAD), predisponendo la possibilità di sviluppare sistemi precisi e speditivi per indirizzare le scelte di fertilizzazione delle colture.

Karimi et al, 2006, hanno applicato una modellistica multivariata basata sulle support vector machines alle immagini iperspettrali per la stima (efficiente) dello status nutrizione in azoto del mais. Menesatti et al, 2010, hanno ottenuto buoni risultati anche nella stima spettrale multicomponente (N, P, K, Ca, Mg, Fe, Zn, Mn) su foglie di arancio, con correlazioni in validazione tra valori osservati e predetti che andavano da $r=0.883$ per Mg a $r=0.481$ per P con un errore standard di predizione compresi tra 0.01 per P a 12.418 per Fe. Gli stessi autori hanno utilizzato sistemi di spettrofotometria di immagine per la selezione qualitativa di foglie di tabacco curato (Menesatti et al., 2001) e sistemi spettrali VIS-NIR puntuali portatili, utilizzati in campo per la stima dello stato di maturità di uve da vino di differenti vitigni (Menesatti, 2007).

Tecnologia del legno

Anche nella tecnologia e qualità del legno, si trovano differenti e interessanti applicazioni. Lebow et al, (1996) hanno classificato alcune caratteristiche della superficie del legno attraverso riflettanza spettrale, mentre Schimleck et al, (2004) hanno sviluppato le calibrazioni spettrali nel vicino infrarosso per alcune proprietà del legno. Altri autori hanno applicato metodi fotonici non distruttivi di indagine sul legno basati su fluorescenza (per eucalipto, Billa et al, 2000), sul vicino infrarosso per legnami duri (Kelley et al, 2004a) o più morbidi (Kelley et al, 2004b), o delle proprietà della polpa (Bharati et al., 2004). Hagman, già nel 1997, ha invece utilizzato con successo un sistema di spettrometria di immagine per l'analisi e lo sviluppo di modelli predittivi delle proprietà del legno.

Anche la computer vision, è stata utilizzata con successo come ausilio nelle fasi di lavorazione del legno (Brunner et al, 1990), per la classificazione delle caratteristiche specifiche superficiali (Butler et al., 2001; Butler et al., 2002) e per l'analisi della decolorazione del legno in relazione all'esposizione luminosa prolungata (Chang and Chang, 2001).

Robotica e macchine agricole

L'optoelettronica, rappresenta uno degli indirizzi evolutivi più complessi e raffinati, delle applicazioni elettroniche attualmente diffuse anche sulle macchine agricole (Urbani et al., 2003).

Fino agli anni '60, lo sviluppo della meccanizzazione agricola è stato rappresentato principalmente da innovazioni delle componenti meccaniche degli attrezzi agricoli e di conseguenza dall'affermarsi di tecniche colturali più articolate.

Dagli anni '70, si è verificato un progressivo coinvolgimento delle tecnologie elettroniche nella componentistica e nelle macchine agricole.

Nel nuovo millennio, l'innovazione meccanico-agricola sarà, probabilmente, legata alla totale integrazione elettronica e informatica nella meccanica e del trasferimento di questa innovazione sul campo della tecnica di meccanizzazione.

Il termine stesso di meccanica agraria è ormai riduttivo per descrivere lo scenario tecnologico prossimo venturo: più propriamente si potrebbe parlare di mecatronica

agricola (Marchant, 1991), fino ad azzardare il termine della (futuribile) cibernetica agraria (Urbani and Menesatti, 2002).

Quest'ultimo termine non deve stupire più di tanto, in quanto le macchine future dovranno essere non solo a "misura" dell'uomo (sicurezza, ergonomia e funzionalità), ma anche in qualche modo ricalcare l'operato umano in termini di adattabilità, flessibilità e complessità funzionale ed operativa. Ciò significa, dotare le macchine dell'elevata capacità di interazione ambientale propria dell'uomo. Questa, a sua volta, è basata sull'acquisizione e scambio informativo e sull'elaborazione di queste informazioni secondo costrutti logico-razionali, al fine di mettere in esecuzione strategie operative complesse (Menesatti, 1996; Menesatti, 1997).

L'attività di ricerca in *computer o machine vision* (Chen et al, 2002) nel settore della meccanica agraria, può essere a sua volta distinta nei seguenti principali indirizzi:

- sviluppo di tecnologie, metodologie e sensoristica per
 - o potenziali applicazioni sulle macchine (pre e postraccolta)
 - o testing e analisi delle caratteristiche operative e costruttive
- applicazioni sulle macchine
 - o operazioni in campo
 - o lavorazione e selezione postraccolta
- robotica
- interazione con sistemi ICT (information and communication technology)

Robot e macchine autonome

Le prime applicazioni di computer vision e image analysis che trovano ampio e concreta applicazione in ambito industriale sono per costituire gli apparati di visione artificiale dei robot o di macchine fortemente automatizzate. I primi sistemi ad essere utilizzati in agricoltura, a partire dalla fine degli anni 80', sono stati i robot per la mungitura (Application of robots in agriculture - the example of milking robots, 1990; Artmann and Schillingmann, 1989). Come nel settore industriale, ruolo essenziale nello sviluppo robotico è rappresentato dal sistema di fotonico di visione artificiale (Gouws, 1993; Komiya and Kawakami, 1994). Attualmente, i robot per la mungitura sono tra i pochi esempi in campo agricolo di macchine robotiche ad ampia

diffusione ed utilizzo (Robot milking rapidly increasing. 2010), sebbene l'effetto delle diverse variabili ambientali sulla loro performance sia ancora molto studiato (Caudrillier et al, 2009; Bonsels and Milchpraxis, 2010).

Un altro importante settore di applicazioni robotiche in agricoltura riguarda gli ambienti protetti e le serre (García et al., 2007). Le condizioni ambientali dei sistemi protetti, soprattutto le temperature e la distribuzione di fitochimici, possono essere potenzialmente dannosi per l'uomo. L'applicazione robotica è finalizzata all'automazione delle operazioni colturali per una minimizzazione dei rischi e una riduzione dei costi. Questi due fattori sono principalmente legate alle operazioni di distribuzione di antiparassitari e alla raccolta del prodotto (Wan Ishak et al., 2010; Fujiura et al., 2010). In entrambi i casi, i sistemi fotonici, soprattutto di visione e orientamento artificiale (Buemi et al, 1994; Åstrand and Baerveldt, 2002; Van Henten et al, 2002) sono di grande importanza per lo sviluppo di sistemi realmente automatici in grado di operare in contesti ambientali non sempre standardizzabili a livelli industriali. Il sistema di visione (Noordam et al., 2005), infatti, associato a metodi di intelligenza artificiale (Morimoto and Hashimoto, 1998) può garantire alla macchina un discreto grado di autonomia comportamentale (Mehta et al, 2008) in contesti variabili.

Un altro settore nel quale si è cimentata la ricerca per i robot in agricoltura, riguarda lo sviluppo di macchine a guida automatica (An Oiu and Chang, 2008) e l'agricoltura sostenibile (Hagras et al., 2002) e di precisione (Hemming et al., 2008). I sistemi di visione artificiale sono essenziali (Foglia et al, 2003) per una parametrizzazione ad ampia banda dell'ambiente agricolo e devono interagire con altri sistemi di posizionamento (Tillet et al., 2002) e navigazione (An Oiu and Chang, 2008). Alcune ricerche si spingono fino a considerare possibili sviluppi di sistemi robotici con locomozione a zampe (Hoffman, 1991) per consentire un movimento più agile in ambienti accidentati.

Altre interessanti esperienze di ricerca in ambito robotico, sono inerenti a macchine autonome per l'effettuazione di fondamentali operazioni agricole ad alta intensità di lavoro umano (raccolta), protezione delle colture, trapianto (Ryu et al, 2001; Ting et al., 1996) o impatto ambientale (Maldonado and Nakaji, 2008). Belforte et al (2004) hanno studiato una raccogliitrice robotizzata per l'uva da tavola e il kiwi, Tillet et al

(1995) un robot per il confezionamento di ortofrutticoli, mentre Singh e Montemerlo (1998) hanno sperimentato interventi di potatura automatizzata attraverso sistemi di computer vision. Per superare il problema della visione tridimensionale dell'ambiente apprezzando non solo le dimensioni relative tra oggetti giacenti sullo stesso piano focale, ma anche la distanza tra oggetti su piani focali diversi (profondità), sono stati proposti differenti metodi tra cui la stereovisione. Questo metodo è completamente basato sulla visione artificiale e si compone di un sistema di due video o fotocamere poste a distanza nota (Menesatti, 1996). La stereovisione è stata applicata a sistemi robotici per aumentare le prestazioni della raccolta automatica, migliorando il riconoscimento dei frutti (Si et al. 2010) e la loro localizzazione spaziale (Zhou et al., 2010).

La protezione delle colture è un altro settore nel quale si è indirizzata la ricerca finalizzata allo sviluppo di sistemi autonomi o robotici. La visione artificiale è stata sperimentata con successo per l'applicazione accurata di microdosi di erbicidi durante la semina (Gaard and Lund, 2007), la guida automatica delle macchine irroratrici (Cho, 2002; Sanchiz and Marchant, 1998), il riconoscimento di piante e semi attraverso immagini all'infrarosso termico (Segmentation of plants and weeds for a precision crop protection robot using infrared images, 1996).

Applicazioni nelle fasi di preraccolta

L'impiego di tecniche e metodiche di computer vision e image analysis nel comparto agromeccanico si è esteso anche al settore delle macchine agricole di pieno campo, operatrici e motrici. Abbiamo distinto queste applicazioni come sistemi di imaging in pre-raccolta e raccolta, da quelli, ugualmente molto diffusi, in post-raccolta, che comprendono per lo più macchine a punto fisso per la lavorazione dei prodotti. Nell'ambito dei sistemi in preraccolta possiamo ulteriormente distinguere studi e applicazioni per implementazioni migliorative dirette sulle macchine motrici o operatrici o finalizzate al testing e all'analisi delle caratteristiche operative e costruttive di queste.

Applicazioni sulle macchine agricole

Macchine motrici

Come è stato già indicato nella sezione delle applicazioni per la robotica, i sistemi di machine o computer vision sono stati sperimentati per le macchine motrici con il principale obiettivo di sviluppare sistemi di guida assistita o autonoma (Vision-guidance of agricultural vehicles, 1995; Gerrish et al, 1997). Come indicato nel paragrafo relativo alle applicazioni nella robotica, una delle principali difficoltà che incontra l'analisi di immagini bidimensionali nel parametrare lo spazio di ambienti naturali o poco standardizzati è l'apprezzamento della terza dimensione o profondità (Menesatti, 1998). Un sistema di visione basato su un solo sensore, che inquadra scene con oggetti di dimensioni non note non ha possibilità di calcolare le distanze tra il sensore (e dunque la macchina) e questi o le distanze assolute tra oggetti stessi. I sistemi di guida automatica devono quindi risolvere prima questo problema. Molti ricercatori, hanno utilizzato sviluppato sistemi 3D, associando all'immagine 2D, l'informazione della profondità rilevata con distanziometri laser (Subramanian et al , 2006). Le immagini tuttavia offrono un'ampia e complessa parametrizzazione dell'ambiente circostante la macchina, in relazione alla disponibilità di informazioni spazio-topologiche e temporali. L'adozione di tecniche di stereovisione, anche sui trattori ha consentito di a Kise et al., (2005) e Kise and Zhang (2008) si sviluppare un sistema di visione 3D, basato su una stereocamera installata su un trattore, impiegato per la misura dell'altezza e del volume di colture a file in pieno campo e l'identificazione delle file stesse. Secondo gli autori, il sistema potrebbe essere validamente per lo sviluppo di un sistema di guida automatico in grado di seguire le file della coltura. Morimoto et al (2000) hanno invece utilizzato sistemi fotonici per l'identificazione di ostacoli durante il movimento di macchine motrici agricole. Il sistema potenzialmente applicabile su trattori , mietitrebbie o altre macchine semoventi genera, attraverso l'analisi di immagini digitali dell'ostacolo, un comando di stop di emergenza con un errore di 30 cm e una velocità di detection di 5.9 Hz.

Macchine operatrici

L'applicazione di sistemi e metodi fotonici sulle macchine agricole operatrici ha il principale scopo di migliorare le prestazioni operative delle macchine stesse, attraverso il controllo più raffinato di fasi complesse come l'identificazione del target (ad es. la pianta, la foglia o il frutto) e il posizionamento della macchina rispetto al target. Le principali applicazioni riguardano macchine per la raccolta, con il fine di

riduzione dei costi di manodopera, e macchine per il trattamento irriguo o fitosanitario, con il fine di ottimizzare la distribuzione. I sistemi di computer vision sono considerati in queste applicazioni come sensori, “in linea su mezzi in movimento”, (on-the-go machine vision sensing, McCarthy et al, 2008), al pari dell’ambito agroindustriale dove si parla di sensori in linea di processo (on-line machine vision sensing, Cheng et al, 2003) e rivestono un importante ruolo nella configurazione di macchine agricole per l’agricoltura di precisione (Lee et al, 2010; Menesatti et al, 2002). Le implementazioni su macchine per la raccolta riguardano in molti casi sistemi di visione ambientale e identificazione dei frutti per lo sviluppo per attività automatiche o completamente robotizzate (Jiménez et al, 2000; Hayashi et al., 2010). Altri ricercatori, invece, hanno concentrato la propria attenzione verso l’applicazione di sistemi e metodi di visione artificiale per l’automazione della sola fase di raccolta. Zeng et al (2009) hanno affrontato il problema del riconoscimento di singoli frutti in immagini in cui compaiono sovrapposti e quindi non ben separati. Zang et al (2010) hanno utilizzato un sistema di machine vision assistito da laser per una migliore localizzazione dei frutti in una macchina prototipale per la raccolta semiautomatica di fragole. Al-Mallahi et al (2010), hanno sviluppato un sistema di visione artificiale basato su immagini ultraviolette per rilevare tuberi di patata sulla raccogliitrice ed eliminare zolle e tuberi indesiderati, nel nastro trasportatore prima di entrare in un contenitore di stoccaggio. La rilevazione si è basata sulla riflessione ultravioletta dei tuberi rispetto al loro background che comprende pezzi di zolle. I risultati hanno mostrato che il 98,79% dei tuberi e 98,28% delle zolle sono stati rilevati correttamente. Inoltre, il tempo di elaborazione necessario per il segmento di oggetti in ogni fotogramma è stato di circa 94ms che indicava la possibilità di applicare il sistema di visione sviluppato in tempo reale all’interno della normale velocità del trasportatore della raccogliitrice. Un sistema di machine vision per il conteggio dei frutti di agrume è stato sviluppato da Chinchuluun et al (2009) per una macchina raccogliitrice a scuotimento. Il sistema testato a banco e in campo, consisteva in una camera a 3CCD, quattro lampade alogene, un encoder e un computer portatile. La stima del peso effettivo dei frutti attraverso analisi delle immagini è risultata molto buona con coefficiente di determinazione tra stimati e osservati di 0,962. La correlazione tra conteggio manuale e conteggio automatico dei frutti in caduta ha fornito un valore di $R^2 = 0,891$. Sempre su frutti di agrume, Bulanon et al (2010) hanno incrementato il riconoscimento dei frutti in pianta, in diverse condizioni di luce, attraverso un sistema multispettrale per immagini. Leemans and Destain (2006) hanno messo a confronto due metodi basati sulla visione artificiale per fornire assistenza alla guida di orientamento per la semina, al

fine di migliorare la spaziatura precisione nei passaggi contigui. Utilizzando un trattamento delle immagini basato sulla trasformata di Hough si è ottenuta una notevole precisione nella rilevazione delle file (errore min di 30 mm). Per le macchine per la distribuzione dei prodotti fitosanitari, i sistemi di computer vision sono stati applicati per la detection delle piante da trattare (Tellaeché et al, 2001), delle malerbe (Bossu et al., 2008), della parte terminale della chioma (Zhang et al, 2007) , di frutti e foglie in vigneto per lo sviluppo di una irroratrice selettiva automatica (Berenstein et al, 2010). Sena et al, (2003) hanno elaborato una tecnica per il riconoscimento della nottua autunnale (*Spodoptera frugiperda*) che è uno dei parassiti del mais più importanti del Brasile. Immagini di piante di mais danneggiate e non danneggiate sono state prese in otto diverse tappe e in tre diverse intensità di luce. L'algoritmo ha classificato correttamente il 94,72% delle 720 immagini. Kim and Slaughter (2008), hanno invece applicato le tecniche di machine vision per migliorare l'efficienza di displocazione dei sistemi di irrorazione nei confronti dei target.

Anche le operazioni colturali di irrigazione e fertilizzazione, possono essere migliorate attraverso l'applicazione di sistemi di computer vision (McCarthy et al, 2008). La termografia consente di stimare in circa 3 s il contenuto idrico del terreno (Millozza et al., 2005) con buona approssimazione.

Testing e analisi delle caratteristiche operative e costruttive delle macchine

Macchine Motrici

Il testing e la certificazione delle macchine agricole motrici e operatrici rappresenta una importante branca dell'attività di ricerca e sperimentazione in ambito agromeccanico mondiale. Il testing, in genere, ha come finalità l'analisi delle performance e delle caratteristiche operative e costruttive delle macchine che gli operatori del settore agricolo si troveranno ad utilizzare.

I sistemi di image analysis sono anche in questo caso, sempre più utilizzati per fornire informazioni multiparametriche in modo speditivo ed affidabile.

Per le macchine motrici, la computer vision è stata applicata per la certificazione dei sedili (Braikov and Rusev, 1996) osservandone le oscillazioni al banco prova, per valutare la resistenza meccanica delle cabine di sicurezza (Mizgalski et al., 1992), per misurare le traiettorie del movimento dei trattori (Martin, 2008) o valutare le performance dei sistemi di guida satellitare (Easterly et al., 2010) o come parte di un sistema per la certificazione di trattori e macchine semoventi sottoposte a riparazioni (Kolchin, 2004). Altro settore, principalmente legato alle macchine motrici, dove

L'analisi di immagine ha trovato interessanti applicazioni è quello dello studio degli pneumatici: al banco prova (Castillo et al., 2006) e nei test di certificazione (Kozyrev et al, 1991), nelle deformazioni a seguito del superamento di ostacoli (System for measuring the coordinates of tire surfaces in transient conditions when rolling over obstacles: Description of the system and performance analysis, 2008) o per la valutazione delle conseguenti deformazioni del terreno (Shen and Yu, 1991).

Macchine operatrici

Molto ampio è anche il settore di indagine relativo al testing e la certificazione delle macchine agricole operatrici, data la numerosità e la varietà di tipologie e operazioni condotte con queste macchine. In prove di certificazione delle macchine per la lavorazione del terreno, condotte attraverso l'ENTAM (European Network for Testing of Agricultural Machines, <http://www.entam.net/>), Menesatti et al. (2005) hanno sviluppato un metodo basato sull'analisi di immagini digitali per la misura obiettiva e speditiva dei residui vegetali sul e nel terreno, per valutare quantitativamente la capacità di interrimento delle macchine (Bonifazi et al., 2007). Gli stessi autori (Menesatti et al, 2007) hanno valutato l'effetto di fertilizzazioni attraverso residui urbani su mais, misurando dimensioni della canopy, delle foglie e delle cariossidi.

Akdemr et al, (2008) hanno sviluppato applicato sistema di analisi delle immagini per contare i numeri di semi, come, piantati su seminatrice di precisione, diminuendo l'errore sulle distanze nella fase della semina di grano, orzo, barbabietola, girasole e cipolla. Nell'ambito dei trattamenti fitosanitari, sistemi di computer vision sono stati utilizzati con successo soprattutto per l'analisi della qualità di distribuzione effettuata attraverso cartine idrosensibili (Banai et al, 2010). Infatti, le gocce che colpiscono la superficie di queste cartine, lasciano un'impronta (stain) di colore diverso dallo sfondo in modo da poter essere rilevate attraverso l'analisi delle immagini (Menesatti and Seth, 1997). La digitalizzazione delle immagini ha consentito di confrontare diversi metodi di processamento utilizzando una piattaforma web (Menesatti, 1998). Smith et al (2000) sono riusciti a mettere in relazione la grandezza delle gocce di prodotto irrorato e la morfologia delle foglie sulla deposizione dell'irrorato, mentre Menesatti et al (2008) hanno applicato un metodo di analisi di immagini termografiche per valutare direttamente in campo su una siepe campione la qualità di distribuzione di un atomizzatore in funzione di diverse regolazioni. Il sistema di analisi, basato sull'irrorazione di acqua più fredda della superficie irrorata, è stato confrontato con i metodi delle cartine idrosensibili e del banco verticale (metodi

quantitativi non spazializzati), dimostrando una buona rispondenza dei risultati e soprattutto la possibilità di osservare spazialmente variazioni anche limitate dell'acqua distribuita.

Food processing e qualità dei prodotti agroalimentari

Nel food processing troviamo la più numerosa produzione scientifica in merito ad applicazioni biofotoniche per l'analisi di processo e di diversi aspetti della qualità dei prodotti (organolettica, sanitaria, nutrizionale) (Menesatti et. al 2000b). Dal punto di vista tecnologico, le applicazioni studiate spaziano dai più avanzati sistemi di computer vision (Menesatti, 1999), all'analisi di immagine multi e iperspettrale (Park et al, 2002), all'imaging in risonanza magnetica nucleare, raggi X e altro ancora. Dal punto di vista metodologico, sono applicate le più avanzate tecniche di elaborazione e processing di analisi di immagine (Brosnan and Sun, 2004) e analisi multivariate dei dati.

Per l'analisi di parametri di qualità del frumento sono state utilizzate sia tecniche di image analysis (Zayas et al, 1996) sia tecniche spettroscopiche (VIS e Nir). Sistemi spettrali operanti nel vicino infrarosso (Nir), sono stati utilizzati per la misura speditiva e non distruttiva del contenuto proteico delle cariossidi (Delwiche and Hruschka, 2000), per differenziare la presenza di vitrosità nel grano duro (Dowell, 2000) e per l'analisi della durezza (Maghirang and Dowell, 2003). Le tecniche spettroscopiche Nir si sono rivelate utili nella detection automatica non distruttiva di entomo-infestazioni interne alle cariossidi. Le tecniche e i sistemi di analisi di immagine (Ruan et al, 1998) e di spettroscopia sia nel visibile sia nel vicino infrarosso, sono state utilmente sperimentate anche nella determinazione di infezioni da patogeni (*Fusarium*, Ruan et al, 1998; *Septoria*, Cavelier et al, 1995) e per l'analisi di contaminazioni da micotossine (*Deoxynivalenol*, Delwiche et al, 2005; *vomitoxin*, Dowell et al, 1999; *aflatoxina e fumonisina*, Pearson et al, 2004).

Le tecniche spettroscopiche e in particolare quelle basate sul vicino infrarosso, presentano comunque alcuni limiti come il relativo costo e, spesso, una ridotta risoluzione spaziale di misura. Per tale motivi, parte dell'attività di ricerca si è indirizzata verso lo studio di metodiche ottiche di costo tecnologico e applicativo minore, ma con maggiore risoluzione ottica, in grado di effettuare misure in continuo

su singole cariossidi, anche su ampie partite di prodotto (Delwiche et al, 2005; Yang et al 2009).

Le tecniche ottiche per immagini non spettrali (*chromatic imaging*) operanti nel visibile attraverso tecnologie a minor costo (CMOS e CCD), si basano sulla stima del parametro FDK (Fusarium-damaged kernels) (Delwiche et al, 2005; Yang et al, 2009) che, in pratica, rileva il grado di striminzitura della cariosside. Attraverso tale parametro, ampiamente utilizzato per l'analisi campionaria visuale di grandi partite (Schaafsma et al, 2004), è possibile predire il contenuto in micotossine, specificamente il deossinivalenolo (DON) (Wrigley, 1999).

La stima dell'FDK attraverso l'impiego di tecniche di *chromatic imaging* si basa sull'analisi della forma della cariosside. Tale analisi può essere sviluppata utilizzando dei parametri descrittivi della forma di tipo morfologico (lunghezze, aree, estensioni del perimetro, rapporti tra assi) e/o descrittori più complessi derivati da tecniche algoritmiche di analisi dei profili chiusi. La maggioranza dei lavori in letteratura, utilizza parametri morfologici per la misura dell'FDK (Delwiche et al, 2005; Yang et al, 2009). L'efficacia di tale approccio metodologico è limitata dalla bassa capacità discriminativa dei parametri. Ad esempio, i parametri diretti (area, diametri, ecc.) sono dipendenti, non solo dalla eventuale striminzitura, ma anche dalla taglia della cariosside e quindi sono fortemente influenzati dalla variabilità biologica, varietale e agropedoclimatica della coltura. Inoltre, il limitato spettro di variabili discriminanti, orienta verso l'impiego di modellistica predittiva monovariata o inferenziale.

Le tecniche più innovative della forma dei prodotti agroalimentari (Costa et al, ACCEPTED) sono invece basate su analisi di approssimazioni matematiche dei profili chiusi, tra queste, una delle più promettenti è quella basata sull'Analisi Ellittica di Fourier (EFA).

Con questa tecnica è stata sperimentata con risultati promettenti, proprio per l'analisi della striminzitura del frumento tenero, da cui derivare una stima dell'incidenza di fusariosi (Menesatti, 2009).

Per quanto riguarda il settore carne, troviamo numerose applicazioni sia sul prodotto pronto al consumo che sulle carcasse e mezzene (Menesatti et al, 2007). Park et al (2008) hanno studiato l'effetto della lavorazione della carcassa e successivamente del metodo di cottura sulla palabilità al consumatore della carne di tre muscoli. Oliver et

al (2010) hanno utilizzato un sistema di analisi delle immagini per migliorare e completare il sistema di classificazione delle carcasse SEUROP, attualmente in uso in Europa. Dopo la macellazione, 91 carcasse di razze da carne locali spagnoli sono stati pesati e classificati secondo il sistema SEUROP. Due fotografie digitali (una laterale e una vista dorsale) sono stati presi dei lati carcassa sinistra, e sono state fatte un totale di 33 misurazioni morfometriche (lunghezze, perimetri, aree). Una maggiore
avc La classificazione effettuata con le regole SEUROP è stata confrontata con quella sviluppata attraverso image analysis che ha dimostrato maggiore accuratezza. I risultati suggeriscono che il metodo di analisi delle immagini potrebbe essere utile come mezzo di integrazione e miglioramento del sistema SEUROP di classificazione delle carcasse bovine. Anche Moore et al (2010) hanno lavorato sulla comparazione tra classificazioni commerciali delle carcasse (punteggio di mazzatura o marbling score) e parametri oggettivi di video image analysis (VIA) e computer vision system (CVS). E' stata valutata la precisione e la ripetibilità del sistema VIA-CVS allo scopo di cercare l'approvazione ufficiale USDA per l'utilizzo di uno strumento commerciale. Sulla base dei positivi risultati, dopo il termine di un'ampia prova ripetuta coinvolgendo anche valutatori esperti, il sistema innovativo ha ottenuto l'approvazione dell'USDA e 2006 strumenti sono stati assegnati ad altrettanti valutatori per l'attività routinaria di classificazione.

Rius-Vilarrasa et al (2009) hanno invece confrontato la tecnica VIA con il metodo di classificazione standard EUROP per le carcasse ovine, trovando che la video image analysis è stata capace di incrementare il valore di corretta predizione delle principali caratteristiche della carcassa rispetto lo standard applicato soggettivamente dagli esperti.

D'Andrea et al (2007) hanno valutato la qualità e l'effetto di conservazione della carne di coniglio attraverso sistemi di spettrometria di immagine. Gli stessi autori (Menesatti et al. 2007) con la stessa tecnica sono stati in grado di classificare correttamente la tipologia di allevamento, biologico o industriale, dei conigli. Pincu et al (2005) hanno analizzato gli effetti sulle caratteristiche colorimetriche e spettrali di carne chianina in relazione a diversi sistemi di allevamento e tempi di conservazione, mentre Liu et al (2010) sono riusciti a categorizzare la qualità di carne di maiale utilizzando uno speciale filtraggio bidimensionale (filtro di Gabor) applicato alle

immagini iperspettrali di fettine di muscolo. Valuos et al (2010) hanno utilizzato complesse funzioni di texture imaging per descrivere quantitativamente le caratteristiche superficiali di prosciutti di maiale. Tecniche di analisi di immagine iperspettrale sono state utilizzate con successo per l'ispezione delle carcasse di pollo (Park, 2008) (Park & Chen, 2001), la detection di contaminanti superficiali sulle stesse (Lawrence et al., 2004) o per la separazione di petti di pollo derivanti da animali sani o setticemici (Hsieh et al., 2002).

El Jabri et al (2010) hanno studiato attraverso image analysis con luce visibile e ultravioletta, la distribuzione di connettivo intramuscolare in muscoli semimembranosi ottenendo un elevato grado di predizione della tenerezza della carne. Analisi molto simili sono state condotte anche da Iacurto et al (1999). Infine, Kröger et al (2008) hanno studiato il parametro qualitative della tenerezza della carne attraverso analisi di immagini ottenute ai raggi X, mentre Prieto et al (2010) hanno utilizzato la tomografia computerizzata a raggi X combinato con tecniche di predizione multivariata PLS, per la stima della qualità e della composizione di carne bovina.

Per quanto riguarda i pesci di allevamento, Costa et al (ACCEPTED) ha utilizzato tecniche di imaging combinate con l'analisi morfologica Ellittica di Fourier per differenziare la qualità di spigola allevata in differenti condizioni. Lo stesso autore (Costa et al, 2010a), con la medesima tecnica ha analizzato l'influenza genetica e dell'ambiente sulle malformazioni di spigola. Lo stesso autore, ha applicato diverse tecniche di calibrazione colorimetrica di immagini RGB al fine di valutare la freschezza di spigola (Costa et al 2010b) o per quantificare differenze tra pesci allevati e selvaggi (Costa et al 2010c). Menesatti et al (2010) hanno invece applicato una tecnica combinata di analisi di immagine iperspettrale nel visibile e vicino infrarosso e morfometria geometrica per uniformare geometricamente i soggetti analizzati determinando non solo le lunghezze d'onda più informative ai fini classificatori della freschezza (minor tempo di conservazione rispetto la pesca), ma anche l'area del pesce maggiormente suscettibile di fornire tali informazioni.

L'analisi di immagine iperspettrale era stata usata come sistema di predizione della freschezza di orate e spigole (Menesatti et al, 2001; Menesatti et al, 2007; Menesatti et al, 2006), cefali e trote (Costa et al, IN PRESS). Aguzzi et al (2007) hanno

utilizzato un sistema di analisi del movimento per immagini al fine di monitorare i ritmi biologici dei pesci. Costa et al, (2008 e 2009) hanno anche studiato la forma di popolazioni geneticamente diverse del mollusco bivalve vongola verace, riuscendo ad attribuire la corretta popolazione e provenienza geografica con notevole precisione.

Le tecniche di image analysis sono state utilizzate come tecniche innovative per la valutazione della shelf life del formaggio (Milanello, 2005) e della cagliata (Ni and Guansekar, 2004), per l'ispezione della distribuzione degli ingredienti in formaggi pastorizzati (Jeliński et al, 2007) e per la misura delle tipiche ocellature da gas (Caccamo et al 2004).

Per i prodotti da forno come il pane, le tecniche di computer vision sono state applicate per lo studio della crescita durante la cottura (Bai, 2006), la descrizione della texture apparente (Bertrand, 1992), per la caratterizzazione di differenti pani tipici (Brescia, 2007), la caratterizzazione della porosità e struttura (Colucci, 2004; Paoletti, 1998) e per la modellizzazione matematica della lievitatura (Romano et al, 2007).

2.5 Bibliografia consultata

- Aguzzi J, Costa C, Fujiwara Y, Iwase R, Ramirez-Llorda E, Menesatti P, 2009b. A novel morphometry-based protocol of automated video-image analysis for species recognition and activity rhythms monitoring in deep-sea fauna. SENSORS, (ISSN 1424-8220), 9: 8438-8455. (IF2008=1.870)
- Aguzzi J, Costa C, Menesatti P, Manuel A, Del Río J, García JA, Fujiwara Y, Iwase R, 2009a. Potential applications of automated video-image analysis in the pelagic and demersal environment including the deep-sea. INSTRUMENTATION VIEWPOINT, 8: 72-73. Paper version ISSN 1697-2562 DL B-51.702-03; Electronic version ISSN 1886-4864 DL B-32814-2006
- Aguzzi, J.; Caprioli, R.; Costa, C.; Menesatti, P.; Scardi, M., 2007. Real Time Motion Analysis as a useful tool to monitor behavioural rhythms and activity statuses in fishes. INSTRUMENTATION VIEWPOINT, 6.
http://www.cdsarti.org/nova_web/ang/recerca/publi_butlletins.asp
- Akdemr, B.; Kaysoglu, B.; Benet, B, 2008, Developing image analysis system for laboratory tests of precision and bulk sowing machines.; Central theme, technology for all: sharing the knowledge for development. Proceedings of the International Conference of Agricultural Engineering, XXXVII Brazilian Congress of Agricultural Engineering, International Livestock Environment Symposium - ILES VIII, Iguassu Falls City, Brazil, 31st August to 4th September, 2008; Bonn: International Commission of Agricultural Engineering (CIGR), Institut fur Landtechnik, 2008
- Al-Mallahi, A., Kataoka, T., Okamoto, H., & Shibata, Y. (2010). Detection of potato tubers using an ultraviolet imaging-based machine vision system. *Biosystems Engineering*, 105(2), 257-265. doi:10.1016/j.biosystemseng.2009.11.004.
- An Qiu; Ji ChangYing; Zhou Jun; Zhao XiaoBo, 2008, Design of an autonomous guided robot for agriculture.; Central theme, technology for all: sharing the knowledge for development. Proceedings of the International Conference of Agricultural Engineering, XXXVII Brazilian Congress of Agricultural Engineering, International Livestock Environment Symposium - ILES VIII, Iguassu Falls City, Brazil, 31st August to 4th September, 2008; Bonn:

- International Commission of Agricultural Engineering (CIGR), Institut für Landtechnik, 2008
- Anderson S. Photonics advances on all fronts. *Laser Focus World*. December 2006:9, 9.
- Antonucci F, Menesatti P, Canali E, Giorgi S, Maienza A, Stazi S R, 2010, Hyperspectral Imaging Characterization Of Agricultural Topsoil Copper Concentration. 2nd International Workshop on Image Analysis in Agriculture. 26-27 August 2010, Budapest, Hungary.
- Application of robots in agriculture - the example of milking robots. Proceedings of a conference held in Braunschweig, Germany, 5-6 December 1990. / Robotereinsatz in der Landwirtschaft am Beispiel des Melkens VDI/MEG Kolloquium Landtechnik; (9), Düsseldorf: VDI, 1990, 252pp
- Ardizzone, Uni Palermo,
http://www.pa.icar.cnr.it/infantino/slides_eid/EID_Parte_1.pdf, ultimo accesso 1 Nov 2010
- Artmann, R.; Schillingmann, D.; 1989, Automation of milking by using robots and electronics , *Agrotique 89*. Proceedings of the second international conference, Bordeaux, France, 26-28 September 1989.; Marseilles: TEKNEA, 1989, 331-348
- Åstrand, Björn; Baerveldt, Albert-Jan. An 2002, Agricultural Mobile Robot with Vision-Based Perception for Mechanical Weed Control, *Autonomous Robots*, July 2002, Vol. 13 Issue: Number 1 p21-35
- Automated visual assessment of plant condition, specifically foliage wilting, reflectance and growth parameters, using
- Bai X, Zhou W. Study of the bread oven rise by online image analysis. *Asia-Pacific Journal of Chemical Engineering* . October 2006;1(1-2):104-109. .
- Banaj, .., Tadic´, V., Lukinac, J., & Horvat, D. (2010). The use of water sensitive paper for the evaluation of spray coverage in an apple orchard. *Poljoprivreda / Agriculture*, 16(1), 43-49
- Barbera, S., 2003 Video image analysis and animal welfare on farm.; *Animal Welfare*; 12 (4), Wheathampstead: Universities Federation for Animal Welfare (UFAW), 2003, 513-515

- Barbosa Filho, J. A. D.; Silva, I. J. O.; Silva, M. A. N.; Silva, C. J. M., 2007, Behavior evaluation of laying hens using image sequences. / Avaliação dos comportamentos de aves poedeiras utilizando seqüência de imagens. By: Engenharia Agrícola; 27 (1), Jaboticabal: Sociedade Brasileira de Engenharia Agrícola (SBEA), 2007, 93-99
- Barbosa Filho, J. A. D.; silva, I. J. O.; Silva, M. A. N., 2006, Welfare evaluation of laying hens kept in different housing systems - a study using image analysis. / Avaliação do bem-estar de aves poedeiras criadas em diferentes sistemas - um estudo utilizando análise de imagens. Agronegócio, tecnologia e inovação. Anais do V Congresso Brasileiro de Agroinformática e II Simpósio Brasileiro de Tecnologia da Informação no Agronegócio Cooperativo, Londrina, PR, 28 a 30 de setembro de 2005. Tome II; Londrina: Sociedade Brasileiro de Informática Aplicada a Agropecuária e Agroindústria (SBI-Agro), 2006, 1121-1128
- Barbosa Filho, J. A. D.; Silva, I. J. O.; Silva, M. A. N., 2008, Welfare evaluation by image analysis of laying hens in different breeding systems and environmental. Proceedings of the International Conference of Agricultural Engineering, XXXVII Brazilian Congress of Agricultural Engineering, Iguassu Falls City, Brazil, 31st August to 4th September, 2008
- Basset, O., Buquet, B., Abouelkaram, S., Delachartre, P., & Culioli, J. (2000). Application of texture image analysis for the classification of bovine meat. Food Chemistry, 69(4), 437-445.
- Belforte G, G.Eula, T.Raparelli, G.Turco 2004 Automatic table grape and kiwifruit harvester, IMG04 Conference, Genova, July 2004;
- Berenstein, Ron; Shahar, Ohad; Shapiro, Amir; Edan, Yael. 2010, Grape clusters and foliage detection algorithms for autonomous selective vineyard sprayer, Intelligent Service Robotics, 1-11
- Bertrand D, Le Guerneve C, Marion D, Devaux M, Robert P. Description of the textural appearance of bread crumb by video image analysis. Cereal chemistry . May 1992;69(3):257-261. .

- Bharati, Manish H., MacGregor, John F., and Marc Champagne. 2004. "Using near-infrared multivariate image regression to predict pulp properties." *Tappi Journal* (2002) 3, no. 5: 8-14.
- Bharati, Manish H.; Liu, J. Jay; MacGregor, John F. Image texture analysis: methods and comparisons. *Chemometrics & Intelligent Laboratory Systems*, Jun2004, Vol. 72 Issue 1, p57-71, 15p;
- Billa, E., Pastou, A., Monties, B., Romero, J., and E.G. Koukios. 2000. "Multivariate chemometric analysis of the fluorescence spectra of eucalyptus wood." *Industrial crops and products* 11, no. 2/3: 187-196.
- Birge R. Introduction To Biophotonics. *Quarterly Review of Biology*. March 2004;79(1):66-66.
- Bonifazi G., Menesatti P., 2004. Digital imaging-based spectrophotometric techniques for soil pathogen monitoring and detection. *Proc. SPIE Vol. 5271*, p. 206-211, *Monitoring Food Safety, Agriculture, and Plant Health*; Bent S. Bennedsen, Yud-Ren Chen, George E. Meyer, Andre G. Senecal, Shu-I Tu; Eds.
- Bonifazi G., Menesatti P., Millozza M. 2004. Multi and hyperspectral digital imaging based techniques for agricultural soil characterization. *Proc. SPIE (The International Society for Optical Engineering) Vol. 5587*, p. 1-9,. *Nondestructive Sensing for Food Safety, Quality, and Natural Resources*; Yud-Ren Chen, Shu-I Tu; Eds
- Bonifazi G., Menesatti P., Millozza M., 2007. Imaging based procedures to evaluate harrowing actions on soil vegetal waste, *IS&T/SPIE 19th Annual Symposium "Electronic Imaging" 28/1 -1/2/2007 San Jose, California USA*. Abstract N° 6501-38, p 146 abstract proceeding.
- Bonifazi G., Piga L., Serranti S., Menesatti P. 2005. Hyperspectral imaging based techniques in contaminated soils characterization. *Proc. SPIE (The International Society for Optical Engineering) Vol. 5996*, p. 215-223, *Optical Sensors and Sensing Systems for Natural Resources and Food Safety and Quality*; Yud-Ren Chen, George E. Meyer, Shu-I Tu; Eds.
- Bonifazi G., Serranti S., Menesatti P. 2005. Hyperspectral imaging based techniques in ornamental stone characterization. *Proc. SPIE (The International Society for*

- Optical Engineering) Vol. 5996, p. 203-214, Optical Sensors and Sensing Systems for Natural Resources and Food Safety and Quality; Yud-Ren Chen, George E. Meyer, Shu-I Tu; Eds.
- Bonsels, T.; Milchpraxis; 2010, Milking robots: optimizing animal observation and data management / *Melkroboter: Tierbeobachtung und Datenmanagement optimieren* 48 (1), Gelsenkirchen: Verlag Th. Mann, 2010, 16-20
- Bookstein, F. L., 1991. Morphometric tool for landmark data. Cambridge Univ. Press.
- Bossu, J.; Gée, C.; Truchetet, F.; 2008 Development of a machine vision system for a real time precision sprayer. Agricultural and biosystems engineering for a sustainable world. International Conference on Agricultural Engineering, Hersonissos, Crete, Greece, 23-25 June, 2008; Silsoe: European Society of Agricultural Engineers (AgEng), 2008, OP-1820
- Braikov, I., & Rusev, R. (1996). Basis of a method for certification testing of tractor seats. *Selskostopanska Tekhnika*, 33(2), 40-43. Retrieved from CAB Abstracts database.
- Brandl, N., 1997, Measuring pig travel by image analysis; *Livestock environment* 5, Volume 1. Proceedings of the Fifth International Symposium, Bloomington, Minnesota, USA, 29-31 May, 1997; St Joseph: American Society of Agricultural Engineers (ASAE), 1997, 461-467
- Brescia M, Sacco D, Sgaramella A, et al. Characterisation of different typical Italian breads by means of traditional, spectroscopic and image analyses. *Food Chemistry* . September 2007;104(1):429-438. .
- Brosnan T, Sun D. Improving quality inspection of food products by computer vision a review. *Journal of Food Engineering*, January 2004;61(1):3. .
- Bruni C, Ferrone C, Koch G, Lucchetti M. Motion identification from image sequences: Information theory and pixel selection. *Computers & Mathematics with Applications*. July 2008;56(2):375-386.
- Brunner, C.C., Shaw, G.B., Butler, D.A., and J.W. Funck. 1990. "Using color in machine vision systems for wood processing." *Wood and fiber science : journal of the Society of Wood Science and Technology* 22, no. 4: 413-428.
- Buemi, F.; Magrassi, M.; Mannucci, A.; Massa, M.; Sandini, G.; 1994, The vision system for the AGROBOT project. *Computers in agriculture. Proceedings of the*

- 5th international conference held in Orlando, Florida, USA, 6-9 February 1994.;
St. Joseph, Michigan: American Society of Agricultural Engineers, 1994, 93-98
- Bulanon, D., Burks, T., & Alchanatis, V. (2010). A multispectral imaging analysis for enhancing citrus fruit detection. *Environment Control in Biology*, 48(2), 81-91
- Butler, D. A., Brunner, C. C., and J. W. Funck. 2001. "Wood-surface feature classification using extended-color information." *Holz als Roh-und Werkstoff* 59, no. 6: 475.
- Butler, D.A., Brunner, C.C., and J.W. Funck. 2002. "Wood-surface feature classification via extended-color imagery." *Forest products journal* 52, no. 6: 80-84.
- Caccamo M, Melilli C, Barbano D, Portelli G, Marino G, Licitra G. 2004, Measurement of gas holes and mechanical openness in cheese by image analysis. *Journal of dairy science* . March 2004;87,:739-748.
- Camargo Neto J, George E. Meyer, David D. Jones, Ashok K. Samal. 2006, Plant species identification using Elliptic Fourier leaf shape analysis. *Computers and Electronics in Agriculture* 50 (2006) 121–124.
- Cangar O, T. Leroy, M. Guarino, E. Vranken, R. Fallon, J. Lenehan, J. Mee, D. Berckmans, 2008, Automatic real-time monitoring of locomotion and posture behaviour of pregnant cows prior to calving using online image analysis, *Computers and Electronics in Agriculture*, 64, 53-60
- Castillo, J., De La Blanca, A., Cabrera, J., & Simón, A. (2006). An optical tire contact pressure test bench. *Vehicle System Dynamics*, 44(3), 207-221.
Retrieved from E-Journals database.
- Caudrillier, J.; Morin, C.; Bouton, L.; Lensink, J.; Bizeray-Filoche, D.; 2009, Impact of grazing on the robotic milking system. / *Impact du pâturage sur le système de traite robotisée* Ethnozootecnie; (87), Clermont-Ferrand: Société d'Ethnozootecnie, 2009, 173-174
- Cavelier, M.; Bollen, L.; Dardenne, P. Use of near infrared spectroscopy for determining the rates of wheat seed infection by *Septoria nodorum* Berk 47th International symposium on crop protection, Gent, Belgium, 9 May, 1995.

- Chaerle, L., Hagenbeek, D., De Bruyne, E., Valcke, R., & Van Der Straeten, D. (2004). Thermal and Chlorophyll-Fluorescence Imaging Distinguish Plant-Pathogen Interactions at an Early Stage. *Plant and Cell Physiology*, 45(7), 887-896
- Chaerle, L., Van Caeneghem, W., Messens, E., Lambers, H., Van Montagu, M., & Van Der Straeten, D. (1999). Presymptomatic visualization of plant-virus interactions by thermography. *Nature Biotechnology*, 17(8), 813-816
- Chang HuiTing, and Chang ShangTzen. 2001. "Correlation between softwood discoloration induced by accelerated lightfastness testing and by indoor exposure." *Polymer Degradation and Stability* 72, no. 2: 361.
- Chen, Y. R.; Chao, K. L.; Kim, M. S., 2002, Machine vision technology for agricultural applications. *Computers and Electronics in Agriculture*; 36 (2/3), 173-191
- Cheng, X., Tao, Y., Chen, Y.R. 2003. Nir/Mir Dual-Sensor Machine Vision System For On-Line Apple Stem-End/Calyx Recognition. *Transactions Of The Asae*. 46(2):551-558.
- Chinchuluun, R., Lee, W., & Ehsani, R. (2009). Machine vision system for determining citrus count and size on a canopy shake and catch harvester. *Applied Engineering in Agriculture*, 25(4), 451-458. Retrieved from CAB Abstracts database.
- Cho S I, N.H.Ki, 2002, Automonous speed sprayer guidance using machine vision and fuzzy logic", *Trans.of the ASAE*
- Colucci F, D'Egidio M. 2004, Use of image analysis to evaluate the crumb structure of bread / Applicazione dell'analisi d'immagine per la valutazione della struttura della mollica del pane. *Tecnica Molitoria* . 2004;55(10):967. .
- Costa C, Aguzzi J, Menesatti P, Antonucci F, Rimatori V, Mattoccia M, 2008. Shape analysis of different populations of clams in relation to their geographical structure. *JOURNAL OF ZOOLOGY*, 276(1): 71-80. (IF2007=1.516)
- Costa C, Antonucci F, Pallottino F, Aguzzi J, Sun DW, Menesatti P, ACCEPTED. Shape analysis of agricultural products by computer vision – a review of recent research advances. *ACCEPTED TO FOOD AND BIOPROCESS TECHNOLOGY* (IF2009= 2.238)

- Costa C, Antonucci F, Pallottino F, Boglione C, Cataudella S, Menesatti P, 2010b. Color-warping imaging: a non destructive technique to evaluate gilthead seabream (*Sparus aurata*, Linnaeus 1758) freshness. 2nd International Workshop on Image Analysis in Agriculture. 26-27 August 2010, Budapest, Hungary.
- Costa C, Antonucci F, Pallottino F, Canali E, Boglione C, Cataudella S, Menesatti P, 2010c. An innovative colorimetric calibration method used to quantify differences among wild and reared gilthead seabream (*Sparus aurata*, Linnaeus 1758). XVIIth World Congress of the International Commission of Agricultural Engineering (CIGR). June 13-17, 2010, Québec City, Canada: 251
- Costa C, D'Andrea S, Russo R, Antonucci F, Pallottino F, Menesatti P, ACCEPTED. Application of non invasive techniques to differentiate sea bass (*Dicentrarchus labrax*, L. 1758) quality cultured under different conditions. ACCEPTED TO AQUACULTURE INTERNATIONAL
- Costa C, Loy A, Cataudella S, Davis D, Scardi M (2006) Extracting fish size using dual underwater cameras. *Aquacultural Engineering* 35(3): 218-227.
- Costa C, Menesatti P, Aguzzi J, D'Andrea S, Antonucci F, Rimatori V, Pallottino P, Mattoccia M, 2009. External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*. *FOOD AND BIOPROCESS TECHNOLOGY*, 3(1): 43-48. DOI: 10.1007/s11947-008-0068-8, (IF2009= 2.238)
- Costa C, Vandeputte M, Antonucci F, Boglione C, Menesatti P, Cenadelli S, Parati K, Chavanne H, Chatain B, 2010a. Genetic and environmental influences on shape variation in the European sea bass (*Dicentrarchus labrax*). *BIOLOGICAL JOURNAL OF THE LINNEAN SOCIETY*, 2010, 101, 427–436.
- Costa C: Menesatti P, Antonucci F, Costa C, Pallottino F, Paglia G, Niciarelli I, Papetti P, IN PRESS. Analisi dello stato di freschezza di trota e cefalo attraverso sistemi non-distruttivi e sperimentazione di sistemi innovativi di conservazione refrigerata. IN PRESS BY EUROFISHMARKET.
- Crampton, J. S., 1995. Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *LETHAIA*, 28: 179-186.

- Cronin, G.M., 2008, Using video image analysis to count hens in cages and reduce egg breakage on collection belts, , Australian journal of experimental agriculture, 2008, v. 48, no. 6-7, p. 768-772
- D'Andrea S, M. Cegna, F. Pallottino, I. Niciarelli, P. Menesatti. 2007. La qualità e la conservazione della carne di coniglio: rilevamenti per mezzo delle tecniche opto-spetttrali. LA RIVISTA DI SCIENZA DELL'ALIMENTAZIONE. Fo.S.A.N., anno 36, pp.1-7.
- Daegling, D. J.; Jungers, W. L., 2000. Elliptical Fourier analysis of symphyseal shape in great ape mandibles. J. HUM. EVOL., 39(1): 107-122.
- De Jong, S. (1993), "SIMPLS: An alternative approach to partial least squares regression," Chemometrics and Intelligent Laboratory Systems, 18, 251-263
- Delwiche S. R. , Thomas C. Pearson and Daniel L. Brabec. 2005. High-Speed Optical Sorting of Soft Wheat for Reduction of Deoxynivalenol. Plant Disease / Vol. 89 No. 11
- Delwiche, S. R., and Hruschka, W. R. 2000. Protein content of bulk wheat from near-infrared reflectance of individual kernels. Cereal Chem. 77:86-88.
- Doeschl-Wilson, A., Green, D., Fisher, A., Carroll, S., Schofield, C., & Whittemore, C. (2005). The relationship between body dimensions of living pigs and their carcass composition. Meat Science, 70(2), 229-240
- Dowell, F. E. 1998. Automated color classification of single wheat kernels using visible and near-infrared reflectance. Cereal Chem. 75:142-144.
- Dowell, F. E. 2000. Differentiating vitreous and nonvitreous durum wheat kernels by using near-infrared spectroscopy. Cereal Chem. 77:155-158.
- Dowell, F. E., Ram, M. S., and Seitz, L. M. 1999. Predicting scab, vomitoxin, and ergosterol in single wheat kernels using nearinfrared spectroscopy. Cereal Chem. 76:573-576.
- Duarte, S.; Reig, L.; Oca, J., 2009, Measurement of sole activity by digital image analysis, Aquacultural Engineering, Jul2009, Vol. 41 Issue 1, p22-27
- Easterly, D., Adamchuk, V., Kocher, M., & Hoy, R. (2010). Using a vision sensor system for performance testing of satellite-based tractor auto-guidance. Computers & Electronics in Agriculture, 72(2), 107-118. doi:10.1016/j.compag.2010.03.004.

- Edgar, J. L.; Paul, E. S.; Nicol, C. J.; 2009, Thermal imaging as a non-invasive tool to assess mild distress in chickens. World Poultry Science Association (WPSA), Proceedings of the 8th European Symposium on Poultry Welfare, Cervia, Italy, 18-22 May, 2009; Beekbergen: World's Poultry Science Association (WPSA), 2009, 96 (Conference paper) AN: 20093297698
- El Jabri, M.; Abouelkaram, S.; Damez, J.L.; Berge, P.. 2010, Image analysis study of the perimysial connective network, and its relationship with tenderness and composition of bovine meat. Journal of Food Engineering, Jan2010, Vol. 96 Issue 2, p316-322
- EMVA (Associazione Europea di Machine Vision) - <http://www.emva.org/cms/index.php>
- Fedrizzi M., Menesatti P. 1997. Visione artificiale e analisi di immagine. LAGUNA. n. 4. pp.12-21.
- Foglia M, A.Gentile, A.Millella, G.Reina 2003, A computer vision application for the kinematic measurement of a cylindrical mobile robot", Raad'03 12th Int.Workshop on Robotics in Alpe-Adria-Danube Region, 2003;
- Forina M (2006) I metodi chemometrici di classificazione: a cosa servono e come sono utilizzati. Proceedings of the Nir Italia 2006, 25-26 May 2006 Ferrara.
- Fu, W., Teng, G., & Wang, Y. (2008). Web-based method for inspecting the pig body size and height using binocular stereovision technology. Central theme, technology for all: sharing the knowledge for development. Proceedings of the International Conference of Agricultural Engineering, XXXVII Brazilian Congress of Agricultural Engineering, International Livestock Environment Symposium - ILES VIII, Iguassu Falls City, Brazil, 31st August to 4th September, 2008
- Fujiura, T.; Wada, T.; Nishiura, Y.; Ikeda, H.; Lin TingXi; Banba, Y.; Dohi, M., 2010, Robot for harvesting cherry tomatoes in mobile cultivation facilities (Part 1) - Outline of the robot and 3-D image recognition. Journal of the Japanese Society of Agricultural Machinery; 72 (2), Saitama: Japanese Society of Agricultural Machinery, 2010, 143-151

- Gaard H. T.; Lund, I., 2007, Application accuracy of a machine vision-controlled robotic micro-dosing system, *Biosystems Engineering*; 96 (3), Oxford: Elsevier, 2007, 315-322
- García, M. A.; Gutiérrez, S.; López, H. C.; Rivera, S.; Ruiz, A. C., 2007, State of the art of robot technology applied to greenhouses. *Avances en Investigacion Agropecuaria*, sep-dic2007, Vol. 11 Issue 3, p53-61
- Geladi, P, and Kowalski, B. (1986), “Partial least-squares regression: A tutorial,” *Analytica Chimica Acta*, 185, 1-17.
- Gerrish, J., Fehr, B., Ee, G., & Welch, D. (1997). Self-steering tractor guided by computer-vision. *Applied Engineering in Agriculture*, 13(5), 559-563
- Gevaux D, (2008) Charting the future for Europe, *Nature Photonics* **2**, 125 - 127 doi:10.1038/nphoton.2008.07
- Gonzalez R. C., Woods R. E., *Digital Image Processing*, pp. 443–465, Addison-Wesley 1992.
- Gouws, Johan., 1993, The systematic development of a machine vision based milking robot, ed **Gouws**, xvi, 180 p. Wageningen University Dissertation
- Gunasekaran, S. (1996). Computer vision technology for food quality assurance. *Trends in Food Science and Technology*, 7(2): 245-255.
- Gunz P, Mitteroecker P, Bookstein FL. 2005. Semilandmarks in three dimensions. In: *Modern Morphometrics In Physical Anthropology*. New York: Kluwer. Acad./Plenum pp. 73–98.
- Hagman, O. 1997. "Multivariate prediction of wood surface features using an imaging spectrograph." *Holz als Roh- und Werkstoff* 55, no. 6: 377.
- Hagras, Hani; Colley, Martin; Callaghan, Victor; Carr-West, Malcolm. 2002, Online Learning and Adaptation of Autonomous Mobile Robots for Sustainable Agriculture by Autonomous Robots, July 2002, Vol. 13 Issue: Number 1 p37-52
- Hall D. *Optoelectronics: An Introduction* (Book). *Physics Today*. April 1984;37(4):81
- Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6), 610-621
- Hayashi, S., Shigematsu, K., Yamamoto, S., Kobayashi, K., Kohno, Y., Kamata, J., et al. (2010). Evaluation of a strawberry-harvesting robot in a field

- test. *Biosystems Engineering*, 105(2), 160-171.
doi:10.1016/j.biosystemseng.2009.09.011.
- Hecker D, Kappler J, Glassmann A, Schilling K, Alt W. Image analysis of time-lapse movies "A precision control guided approach to correct motion artefacts. *Journal of Neuroscience Methods*. July 2008;172(1):67-73.
- Hemming, J.; Tuijl, B. A. J. van; Balendonck, J.; Henten, E. J. van; Janssen, H. J. J. 2008, Cropscout II, a modular mini field robot for research on precision agriculture.; *Agricultural and biosystems engineering for a sustainable world. International Conference on Agricultural Engineering, Hersonissos, Crete, Greece, 23-25 June, 2008*; Silsoe: European Society of Agricultural Engineers (AgEng), 2008, OP-1225
- Hoffman, R.; 1991, Autonomous legged robots in agriculture. *Automated agriculture for the 21st century. Proceedings of a conference held in Chicago, Illinois, USA, 16-17 December 1991.*; St. Joseph, Michigan: American Society of Agricultural Engineers, 1991, 218-225
- Hsieh C; Chen Y R; Dey B P; Chan D E (2002) Separating septicemic and normal chicken livers by visible/near-infrared spectroscopy and back-propagation neural networks. *Transaction of the ASAE*, vol 45 (2), 459-469
- Iacurto M., Menesatti P., Gigli S., Marcelli S., Paglia G. 1999 Artificial vision techniques for indirect determination of fat and fibre on semitendinosus muscle: a first approach. oral communication at XIII Congresso Nazionale Associazione Scientifica di produzione Animale (ASPA), Piacenza (Italy) 21-24/06/1999 – proceedings pp. 629 -631
- Inspirations from Biological Optics for Advanced Photonic Systems. *Science*. November 2005;310(5751):1148-1150.
- Ipema, A. H.; Haan, L. de; Hemming, J.; Schouten, W. G. P., 2004, Image analysis for monitoring unrest or discomfort of cows during milking.; *Automatic milking: a better understanding. Conference Proceedings, Lelystad, Netherlands, March 2004*; Wageningen: Wageningen Academic Publishers, 2004, 420-421 (Book chapter; Conference paper) AN: 20043140289

- Iqbal A, Valous N A, Mendoza F, Sun D W, Allen P, 2010 Classification of pre-sliced pork and turkey ham qualities based on image colour and textural features and their relationships with consumer responses *Meat Science*, 84 (3), 455-465
- Jackman P, Sun Da-Wen, Allen P, Automatic segmentation of beef longissimus dorsi muscle and marbling by an adaptable clustering and contrast enhancement algorithm. *Meat Science* Volume 83, Issue 2, 2009, Pages 187-194
- Jackman, P., Sun, D.-W., Du, C.-J., Allen, P. & Downey, G. (2008). Prediction of beef eating quality from colour, marbling and wavelet texture features. *Meat Science*, 80(4): 1273-1281.
- Jahr, Ingmar, 2003, *Lexikon der industriellen Bildverarbeitung*, 1150 Stichwörter, Vision Academy
- Jain R, Kasturi R, Schunck BG (2005) *Machine vision*. McGraw-Hill Science (Eds.).
- Jain R, Kasturi R, Schunck BG (2005) *Machine vision*. McGraw-Hill Science (Eds.).
- Jeliński, Tomasz; Du, Cheng-Jin; Sun, Da-Wen; Fornal, Józef , 2007, Inspection of the distribution and amount of ingredients in pasteurized cheese by computer vision. *Journal of Food Engineering*, Nov2007, Vol. 83 Issue 1, p3-9
- Jiménez, A.R., R. Ceres, and J.L. Pons. 2000. "A vision system based on a laser range-finder applied to robotic fruit harvesting." *Machine Vision and Applications* 11, no. 6: 321-329.
- Jovin T. Introduction to Biophotonics (Book). *Journal of the American Chemical Society*. September 29, 2004;126(38):12191-12191.
- Karimi, Y.; Prasher, S.O.; Patel, R.M.; Kim, S.H., 2006, Application of support vector machine technology for weed and nitrogen stress detection in corn, *Computers and Electronics in Agriculture*, April 2006, Vol. 51 Issue: Number 1-2 p99-109, 11p
- Kelley, S. S., Rials, T. G., Snell, R., Groom, L. H., and A. Sluiter. 2004a. "Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood." *Wood Science and Technology* 38,
- Kelley, T.G. Rials, L.R. Groom, and C.-L. So. 2004b. "Use of near infrared spectroscopy to predict the mechanical properties of six softwoods." *Holzforschung: International Journal of the Biology, Chemistry, Physics, & Technology of Wood* 58, no. 3: 252-260.

- Kim, D., & Slaughter, D. (2008). Image-based real-time displacement measurement system. *Biosystems engineering*, 101(4), 388-395
- Kincade K. International congress lays the groundwork for biophotonics roadmap. *Laser Focus World*. April 2008;44(4):21-25.
- Kise, M., & Zhang, Q. (2008). Development of a stereovision sensing system for 3D crop row structure mapping and tractor guidance. *Biosystems Engineering*, 101(2), 191-198. doi:10.1016/j.biosystemseng.2008.08.001.
- Kise, M., Zhang, Q., & Más, F. (2005). A stereovision-based crop row detection method for tractor-automated guidance. *Biosystems Engineering*, 90(4), 357-367. doi:10.1016/j.biosystemseng.2004.12.008.
- Knossow D, Ronfard R, Horaud R. Human Motion Tracking with a Kinematic Parameterization of Extremal Contours. *International Journal of Computer Vision*. September 2008;79(3):247-269.
- Kolchin, A. (2004). Instrumental control system for certification of tractors and self-propelled farm machines subjected to repair. *Traktory i Sel'skokhozyaistvennyye Mashiny*, (12), 40-42
- Kollis, K., Phang, C., Banhazi, T., & Searle, S. (2007). Weight estimation using image analysis and statistical modelling: a preliminary study. *Applied Engineering in Agriculture*, 23(1), 91-96
- Komiya, M.; Kawakami, K.; 1994, Development of automated milking equipment. Detection of teat location by image processing *Journal of the Japanese Society of Agricultural Machinery*; 56 (3), 1994, 85-92
- Kozyrev, S., Stankevich, E., Beloslyudov, A., & Goncharenko, S. (1991). Tyres for traction vehicles: certification and input control of exploitation properties. *Traktory i Sel'skokhozyaistvennyye Mashiny*, (5), 25-27. Retrieved from CAB Abstracts database.
- Kröger, Chris; Devine, Carrick; Bartle, C. Murray; West, John G. 2008, Meat Tenderness Evaluation by X-Ray Image Analysis. *AIP Conference Proceedings*, 11/6/2008, Vol. 1060 Issue 1, p51-54
- Kuhl, F. P.; Giardina, C. R., 1982. Elliptic Fourier features of a closed contour. *COMP. GRAPH. IMAGE PROCESS.*, 18: 236-258.

- Lamela H, Ruiz-Llata M. Optoelectronic neural processor for smart vision applications. *Imaging Science Journal*. 2007;55(4):197-205.
- Lawrence K C, William R. Windham, Bosoon Park, Douglas P. Smith, Gavin H. Poole 2004. Comparison between visible/NIR spectroscopy and hyperspectral imaging for detecting surface contaminants on poultry carcasses. *Proc. SPIE* Vol. 5271, p. 35-42
- Leahy M, Enfield J, Clancy N, O' Doherty J, McNamara P, Nilsson G. Biophotonic methods in microcirculation imaging. *Medical Laser Application*. October 2007;22(2):105-126.
- Lebow, P. K., Brunner, C. C., Maristany, A. G., and D. A. Butler. 1996."Classification of wood surface features by spectral reflectance." *Wood and Fiber Science* 28, no. 1: 74.
- Lee, W. S., V. Alchanatis, C. Yang, M. Hirafuji, D. Moshou, and C. Li. 2010. "Sensing technologies for precision specialty crop production." *Computers and Electronics in Agriculture* 74, no. 1: 2-33
- Leemans, V., & Destain, M. (2006). Application of the Hough transform for seed row localisation using machine vision. *Biosystems Engineering*, 94(3), 325-336. doi:10.1016/j.biosystemseng.2006.03.014.
- Lestrel, P. E., 1997. *Fourier descriptors and their applications in biology*. Cambridge Univ. Press: 466 p.
- Leuschner, P. (2005). Computer image analysis in pig husbandry. / *Computer - Bildanalyse in der Schweinehaltung: optiSORT Grossgruppenhaltung von Hölscher+Leuschner*. Landtechnik, 60(3), 177.
- Li S, Yang B. 2008. Multifocus image fusion using region segmentation and spatial frequency. *Image & Vision Computing*. July 2008;26(7):971-979.
- Lillholm, Martin; Nielsen, Mads; Griffin, Lewis D. Feature-Based Image Analysis. *International Journal of Computer Vision*, May/Jun2003, Vol. 52 Issue 2/3, p73-95, 23p;
- Liu, L.; Ngadi, M.O.; Prasher, S.O.; Gariépy, C.. 2010 Categorization of pork quality using Gabor filter-based hyperspectral imaging technology. *Journal of Food Engineering*, Aug2010, Vol. 99 Issue 3, p284-293

- Maghirang, E. B., and Dowell, F. E. 2003. Hardness measurement of bulk wheat by single-kernel visible and near-infrared reflectance spectroscopy. *Cereal Chem.* 80:316-322.
- Maldague X P V (1994). *Infrared Methodology and Technology (Nondestructive Testing Monographs and Tracts, Vol 7)*, CRC Press
- Maldonado, A. I. L.; Nakaji, K.; 2008, Development of an intelligent robot for an agricultural production ecosystem - new concept of robot and dynamics of a golden apple snail in paddy, *Journal of the Faculty of Agriculture, Kyushu University*; 53 (1), Fukuoka: Faculty of Agriculture Publications, Kyushu University, 2008, 115-119
- Marchant J A, 1991 *Mechatronics in agricultural engineering, Mechatronics, volume 1, Issue 1, 1991, Pages 11-18,*
- Martin, J., & Gil, J. (2008). Calibration and evaluation of a machine vision algorithm to detect agricultural tractor movement. *Bulgarian Journal of Agricultural Science*, 14(4), 436-444. Retrieved from CAB Abstracts database.
- McCarthy, C.; Hancock, N.; Raine, S., 2008, Applied machine vision of plants: a review with implications for field deployment in automated farming operations. *Intelligent Service Robotics-9*
- Mehta, S.S.; Burks, T.F.; Dixon, W.E.. 2008, Vision-based localization of a wheeled mobile robot for greenhouse applications: A daisy-chaining approach, *Computers and Electronics in Agriculture*, August 2008, Vol. 63 Issue: Number 1 p28-37
- Menesatti P, Aguzzi J, Costa C, García JA, Sardà F, 2009. A new morphometric implemented video-image analysis protocol for the study of social modulation in activity rhythms of marine organisms. *JOURNAL OF NEUROSCIENCE METHODS*, 184: 161-168.
- Menesatti P, Antonucci F, Costa C, Santori A, Niciarelli I, Infantino A, 2009, Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels. 1st International Workshop on Computer Image Analysis in Agriculture, Potsdam, Germany 27 – 28 August 2009, *Bornimer Agrartechnische Berichte - Heft 69, ISSN 0947-7314, Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V. (ATB)*

- Menesatti P, Costa C, Aguzzi J, 2010. Quality evaluation of fish by hyperspectral imaging. In: Hyperspectral imaging for food quality: analysis and control. In D.-W. Sun (Ed.), Hyperspectral imaging for food quality analysis and control. London, Burlington, San Diego: ACADEMIC PRESS, ELSEVIER: 273-294. ISBN: 9780123747532
- Menesatti P, G. Urbani, M. Millozza, S. D'Andrea, S. Solaini, G. Paglia, I. Niciarelli 2006. Prediction of the Chilling Storage Time of Fresh Salt-water Fishes by Means of Non-destructive Techniques. 2006 CIGR Section VI International Symposium on Future Of Food Engineering, Warsaw, Poland, 26-28 April 2006, oral communication, abstract p. 145.
- Menesatti P, M. Biocca, S. D'Andrea, M. Pincu. 2008. Thermography to analyze distribution of agricultural sprayers. QIRT JOURNAL (QUANTITATIVE INFRARED THERMOGRAPHY JOURNAL) vol. 5, n° 1/2008, pp. 81-96
- Menesatti P, M. Biocca, S. D'Andrea, M. Pincu. 2008. Thermography to analyze distribution of agricultural sprayers. QIRT JOURNAL (QUANTITATIVE INFRARED THERMOGRAPHY JOURNAL) vol. 5, n° 1/2008, pp. 81-96
- Menesatti P, Roccuzzo G, Cegna M, Torrisi B, Niciarelli I., Allegra M, Intrigliolo F. 2008. Estimation of citrus leaves N content by VIS-Nir portable spectrophotometer. 11th Int Citrus Congress Wuhan 26-30/10/08 (China), pp. 174-175
- Menesatti P. 1995a. L'analisi di immagine nell'ingegneria agraria. MACCHINE E MOTORI AGRICOLI. n. 4; pp. 27-30.
- Menesatti P. 1995b. Analisi di immagine, quali applicazioni in agricoltura. MACCHINE E MOTORI AGRICOLI. n. 5; pp. 18-22.
- Menesatti P. 1996. La visione artificiale e l'analisi dell'immagine in agricoltura. MCMICROCOMPUTER. n. 160. pp. 158-166.
- Menesatti P. 1996. Measure accuracy of stereovision systems based on CCD video-photographic equipment in application to agricultural and environmental survey. SPIE's (International Society for Optical Engineering) International Conference PHOTONICS EAST '96 of Boston (USA) 18-22/11/96; proceedings of conference on "Optics in Agriculture, Forestry and Biological Processing II"; vol. SPIE n. 2907; pp. 268-272

- Menesatti P. 1997. La ricerca sull'analisi di immagine in agricoltura. MONDO MACCHINA n. 11-12 pp. 16-22
- Menesatti P. 1998. Integrazione di sistemi VIA (video image analysis) attraverso Internet: valutazione comparativa dello spettro di distribuzione di fitofarmaci. INFORMATORE FITOPATOLOGICO n. 9/98. pp27-30
- Menesatti P. 1998. Visione artificiale stereoscopica per il telerilevamento agro-meccanico e ambientale da terra. RIVISTA ITALIANA DI INGEGNERIA AGRARIA. n. 2 pp. 74-78.
- Menesatti P. 1999. Applicazione dell'analisi di immagine per la valutazione qualitativa di prodotti agro-alimentari. oral communication at Convegno AIIA 1999 - L'innovazione tecnologica per l'agricoltura di precisione e la qualità produttiva. Grugliasco (TO) 22-24/6/99 – estratti pp. 389 - 400
- Menesatti P., 2004. Photonics technologies to develop meat quality control systems. 39° Simposio Internazionale di Zootecnia. Meat science & research. CNR Roma 10 Giugno 2004. SIPZOO (Soc. Italiana per il Progresso della Zootecnia). Atti pp. 193-214. main lectures (comunicazione orale su invito)
- Menesatti P., Antonucci F., Pallottino F., Rocuzzo G., Allegra M., Stagno F., Intrigliolo F. (2010). Estimation of plant nutritional status by VIS-NIR spectrophotometric analysis on orange leaves [Citrus sinensis (L) Osbeck cv Tarocco]. BIOSYSTEMS ENGINEERING doi:10.1016/j.biosystemseng.2010.01.003; 105(2010), 448-454
- Menesatti P., Bonifazi G., Piazzolla P. 2002. Hyperspectral imaging application to estimate agricultural soil quality parameters for precision farming. Int. Conf. Agr. Engineering - AgEng Budapest (Ungheria) 30/6-4/7/2002 - book of abstract part 1 pp. 140-141
- Menesatti P., Fedrizzi M., Paglia G., Solaini S., Sperduti M., Uniformi M., 2001. Prime valutazioni su applicazioni di imaging spettrale per la selezione qualitativa di foglie di tabacco curato. VII Convegno Nazionale Di Ingegneria Agraria – AIIA 11-13/9/01 Vieste (FG) – book of abstract 2000
- Menesatti P., Greppi G., 2004. Spettrofotometria VIS-NIR e qualità della carne: introduzione ai sistemi, metodologie, analisi dei dati ed esempi di applicazioni sperimentali – in Sistemi e metodologie per la definizione della qualità della

- carne: IL COLORE – ed. Istituto Sperimentale Italiano “Lazzaro Spallanzani”, Fondazione Cariplo, 2005, pp. 55-92.
- Menesatti P., Iacurto M., D’Andrea S., Pincu M., Saltalamacchia F. 2004. La termografia all’infrarosso per rilevare il benessere animale. L’INFORMATORE AGRARIO. 45/2004. pp. 49-53.
- Menesatti P., Millozza M., Pincu M., Cervellini C. 2005. Misura della qualità di erpicatura mediante image analysis sui residui vegetali nel terreno. AIIA 2005: Catania, 27-30 giugno 2005 L’ingegneria agraria per lo sviluppo sostenibile dell’area Mediterranea. oral communication. Atti convegno. P. 85. paper available on CD-Rom - Paper n 3018, 11 pagine
- Menesatti P., Paglia G., Sperduti M., Uniformi M.. 2000b. Opto-informatica per la qualità dei prodotti e dell'ambiente. Convegno Internazionale Produzioni alimentari e qualità della vita Sassari, 4 - 8 Settembre 2000 – proceedings pp. 659-666
- Menesatti P., Seth S. 1997. VIA (video image analysis) systems integration by the Internet: comparative evaluation of spray droplet spectrum analysis systems. WMC '97 (World Manufacturing Congress) - ISMT '97 (International Symposium on manufacturing Technology); Massey University; Auckland, New Zealand in November 18-21, pp. 182-188.
- Menesatti P., Urbani G., Solaini S., Paglia G. 2001 Visible spectrophotometric imaging for noninvasive prediction of fish (*Dicentrarchus labrax*) freshness. Oral communication at the International Conference Physical methods in Agriculture (PMA01) Prague Czech Rep. 27-30/8/2001 – proceedings pp. 204-209.
- Menesatti, P.. 2007. In-field spectrophotometric measurement to estimate maturity stage of wine grapes. Optics for Natural Resources, SPIE Agriculture, and Foods II, Yud-Ren Chen; George E. Meyer; Shu-I Tu, Editors, 676118, Proceeding Vol. 6761
- Menesatti, P.; D’Andrea, S.; Costa, C., 2007. Spectral and thermal imaging for meat quality evaluation. In: New developments in evaluation of carcass and meat quality in cattle and sheep. C. Lazzaroni, S. Gigli, D. Gabina (Eds.).

WAGENINGEN ACADEMIC PUBLISHERS ISSN 0071-2477. EAAP 123: 115-134.

- Menesatti, P.; Millozza, M.; Pincu, M.; Pallottino, F.; Paglia, G.; Cegna, M.; Niciarelli, I.; Stagno, F.; 2007. Metodi Fotonici per la determinazione speditiva non-distruttiva in campo dello stato nutrizionale e dello stress di piante di agrumi. *ITALUS HORTUS*, 14 (6), 2007: 33-34
- Menesatti, P.; S. D'Andrea; P. Negretti. 2007. Classification of rabbit meat obtained with industrial and organic breeding by means of spectrophotometric technique. *SPIE Optics for Natural Resources, Agriculture, and Foods II*, Yud-Ren Chen; George E. Meyer; Shu-I Tu, Editors, 67611A, Proceeding Vol. 6761
- Menesatti, P.; S. D'Andrea; S. Socciarelli. 2007. Application of image analysis techniques to evaluate the effect of urban residuals fertilization on corn (*Zea mays*) production. *SPIE Optics for Natural Resources, Agriculture, and Foods II*, Yud-Ren Chen; George E. Meyer; Shu-I Tu, Editors, 67611B, Proceeding Vol. 6761
- Menesatti, P.; Urbani, G.; Pallottino, F.; D'Andrea, S.; Costa, C., 2007. Non-destructive multi-parametric instruments for fish freshness estimation. *INSTRUMENTATION VIEWPOINT*, 6.
http://www.cdsarti.org/nova_web/ang/research/publi_bulletins.asp
- Milanello S. 2005 Innovative organoleptic tests for the evaluation of shelf life / Tecniche sensoristiche innovative per la valutazione della: shelf-life. *Latte* . 2005;30(5):52.
- Millozza M., Menesatti P., Solaini S. 2005 Stima termografica del contenuto idrico del terreno agrario.– Convegno Nazionale - AIIA 2005: Catania, 27-30 giugno 2005 L'ingegneria agraria per lo sviluppo sostenibile dell'area Mediterranea. oral communication. Atti convegno. P. 122. paper available on CD-Rom - Paper n 3058, 12 pagine
- Mizgalski, A., Kruszewski, A., Majczak, W., & Karaś, J. (1992). Use of video techniques in strength testing of cab frameworks and safety cages. / Wykorzystanie techniki video w badaniach wytrzymałościowych szkieletów kabin i ram ochronnych. *Zeszyty Problemowe Postępów Nauk Rolniczych*, 404217-222. Retrieved from CAB Abstracts database.

- Moore, C. B.; Bass, P. D.; Green, M. D.; Chapman, P. L.; O'Connor, M. E.; Yates, L. D.; Scanga, J. A.; Tatum, J. D.; Smith, G. C.; Belk, K. E. 2010. Establishing an appropriate mode of comparison for measuring the performance of marbling score output from video image analysis beef carcass grading systems. *Journal of Animal Science*, Jul2010, Vol. 88 Issue 7, p2464-2475
- Morimoto, E., Suguri, M., & Umeda, M. (2000). Obstacle finding system for autonomous vehicle for Japanese paddy farm road. 2000 ASAE Annual International Meeting, Milwaukee, Wisconsin, USA, 9-12 July 2000
- Morimoto, T., and Y. Hashimoto. 1998. "An intelligent control technique based on fuzzy controls, neural networks and genetic algorithms for greenhouse automation." *Artificial intelligence in agriculture 1998. A Proceedings volume from the 3rd IFAC/CIGR Workshop, Makuhari, Chiba, Japan, 24-26 April 1998* 61-66.
- Ni H, Guansekaran S. Image processing algorithm for cheese shred evaluation. *Journal of food engineering* . January 2004;61,:37-45.
- Noordam, J. C.; Hemming, J.; Heerde, C. van; Golbach, F.; Soest, R. van; Wekking, E.; 2005, Automated rose cutting in greenhouses with 3D vision and robotics: analysis of 3D vision techniques for stem detection. *Acta Horticulturae*; 691(Vol 2), Leuven: International Society for Horticultural Science (ISHS), 2005, 885-892
- O'Rourke S, Holden N1, Menesatti P, Giorgi S, Canali E, Predictions of soil organic carbon in topsoil from forest soils using VIS-NIR hyperspectral imaging, Conference IASIM-10 (International Association for Spectral Imaging), Dublin 17th Nov 2010
- Oliver, A.; Mendizabal, J.A.; Ripoll, G.; Albertí, P.; Purroy, A., 2010 Predicting meat yields and commercial meat cuts from carcasses of young bulls of Spanish breeds by the SEUROP method and an image analysis system.. *Meat Science*, Apr2010, Vol. 84 Issue 4, p628-633
- Otsu N (1979). A threshold selection method from gray-level histograms". *IEEE Trans. Sys., Man., Cyber.* 9 : 62–66. doi : 10.1109/TSMC.1979.4310076

- Pallottino F, Menesatti P, Costa C, Paglia G, De Salvador F R, Lolletti D. IN PRESS
Image analysis techniques for automated hazelnut peeling determination. FOOD
AND BIOPROCESS TECHNOLOGY. DOI 10.1007/s11947-009-0211-1.
- Paoletti F., Menesatti P., Quaglia G.B., Di Giacomo G., Aloisi L., Terreri M. 1998
Crumb grain and texture evaluation of bread obtained with *Kluiveromyces lactis*
addition. 13th International Congress on Agricultural Engineering (CIGR)
Rabat (Morocco). 2-6/2/98. vol. 6. pp. 197-202.
- Park B. and Y. R. Chen. 2001. Co-occurrence matrix texture features of multi.spectral
images on poultry carcasses. Journal agricultural engineering researcher (2001)
78 (2), 127-139.
- Park B. Quality inspection of poultry carcasses. Computer vision technology for food
quality evaluation / edited by Da-Wen Sun [e-book].Amsterdam ; Boston :
Elsevier/Academic Press, 2008; 2008:157-187
- Park B., K.C. Lawrence, W.R. Windham, D.P. Smith and P.W. Feldner,
Hyperspectral imaging for food processing automation, Proceedings of the
SPIE, Volume 4816 (2002), pp. 308-16.
- Park BY, Hwang IH, Cho SH, Yoo YM, Kim JH, Lee JM, Polkinghorne R,
Thompson JM (2008). Effect of carcass suspension and cooking method on the
palatability of three beef muscles as assessed by Korean and Australian
consumers. Australian Journal of Experimental Agriculture 48, 1396-1404.
- Pastorelli, G.; Musella, M.; Zaninelli, M.; Tangorra, F.; Corino, C.;, 2006, Static
spatial requirements of growing-finishing and heavy pigs. Livestock Science;
105 (1/3), Amsterdam: Elsevier, 2006, 260-264
- Pearson T, Brabec D, Haley S, Color image based sorter for separating red and white
wheat, Sensing and Instrumentation for Food Quality and Safety Dec 2008,
Volume 2, Number 4: 280-288
- Pearson, T. C., Wicklow, D. T., and Pasikatan, M. C. 2004. Reduction of aflatoxin
and fumonisin contamination in yellow corn by highspeed dual wavelength
sorting. Cereal Chem. 81:490-498.
- Pergher G., Petris R., Biocca M., Menesatti P. 2007. Prime verifiche di campo di
un'irroratrice schermata ad aeroconvezione di nuova concezione. Giornate
Fitopatologiche 2008.

- Peters R. Image motion environments: background noise for movement-based animal signals. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*. May 2008;194(5):441-456.
- Philipp C, Berlien H. The future of biophotonics in medicine “A proposal. *Medical Laser Application*. June 2006; 21(2):115-122.
- Photonics 21 (European Technology Platform on Photonics), 2005, Photonics in Europe - Economic impact,
<http://www.photonics21.org/download/Photonics21FlyerGeneralInformation.pdf>
f
- Pincu M., D’Andrea S., Menesatti P., Iacurto M. 2005. Effetti sulle caratteristiche colorimetriche e spettrali di carne chianina in relazione a diversi sistemi di allevamento e tempi di conservazione. 4th World Italian Beef Cattle Congress, Gubbio (PG) Italy, April 29th - May 1th, 2005. oral communication. Proceedings pp. 537-541. abstract cited in *Taurus speciale* anno XVI n° 6, p. 80
- Prasad P. Polymer science and technology for new generation photonics and biophotonics. *Current Opinion in Solid State & Materials Science*. January 2004;8(1):11-19.
- Prieto, N.; Navajas, E.A.; Richardson, R.I.; Ross, D.W.; Hyslop, J.J.; Simm, G.; Roehle, R.. 2010, Predicting beef cuts composition, fatty acids and meat quality characteristics by spiral computed tomography. *Meat Science*, Nov2010, Vol. 86 Issue 3, p770-779
- Quevedo, R.A., & Aguilera, J.M. (2008). Computer Vision and Stereoscopy for Estimating Firmness in the Salmon (Salmon salar) Fillets. *Food and Bioprocess Technology*, doi 547 10.1007/s11947-008-0097-3.
- Rius-Vilarrasa, E., Bünger, L., Brotherstone, S., Macfarlane, J., Lambe, N., Matthews, K., et al. (2010). Genetic parameters for carcass dimensional measurements from Video Image Analysis and their association with conformation and fat class scores. *Livestock Science*, 128(1-3), 92-100.
- Rius-Vilarrasa, E., Bünger, L., Brotherstone, S., Matthews, K., Haresign, W., Macfarlane, J., et al. (2009). Genetic parameters for carcass composition and performance data in crossbred lambs measured by Video Image Analysis. *Meat Science*, 81(4), 619-625

- Rius-Vilarrasa, E.; Bünger, L.; Maltin, C.; Matthews, K.R.; Roehe, R.. 2009, Evaluation of Video Image Analysis (VIA) technology to predict meat yield of sheep carcasses on-line under UK abattoir conditions. *Meat Science*, May2009, Vol. 82 Issue 1, p94-100
- Robot milking rapidly increasing. 2010, *Veepro Magazine*; 74, Arnhem: Veepro Holland, 2010, 4-5
- Rodrigues, V. C.; Barbosa Filho, J. A. D.; Vieira, A. M. C.; Vieira, F. M. C.; Silva, I. J. O.; 2008, Spatial distribution of laying hens in different environmental conditions by image processing and correspondence analysis. Central theme, technology for all: sharing the knowledge for development. Proceedings of the International Conference of Agricultural Engineering, XXXVII Brazilian Congress of Agricultural Engineering, International Livestock Environment Symposium - ILES VIII, Iguassu Falls City, Brazil, 31st August to 4th September, 2008; Bonn: International Commission of Agricultural Engineering (CIGR), Institut für Landtechnik, 2008
- Rohlf, F. J., 1996. Morphometric spaces, shape components, and the effects of linear transformations. In: *Advances in Morphometrics*. Marcus, L. F., M.Corti, A. Loy, G. J. P. Naylor, and D.E.Slice eds. New York: Plenum.
- Romano A, Toraldo G, Cavella S, Masi P. 2007 Description of leavening of bread dough with mathematical modelling. *Journal of Food Engineering* . November 2007;83(2):142-148. .
- Ruan, R., Ning, S., Song, A., Ning, A., Jones, R., and Chen, P. 1998. Estimation of Fusarium scab in wheat using machine vision and a neural network. *Cereal Chem.* 75:455-459.
- Russ, J C, *The Image Processing Handbook ed 2*. 1995, Boca Raton: CRC Press; 674 p.
- Rydberg, A., & Gilbertsson, M. (2005). Monitoring live pig weight with a mobile imaging system
- Ryu KH , G.Kim, J.S.Han, 2001, Development of a robotic transplanter for bedding plants", *J.of Agric.Eng.Res*
- Sanchiz J M, F.Pla, J.A.Marchant, 1998, An approach to the vision tasks involved in an autonomous crop protection vehicle" *Eng.Appl. of Artificial Intelligence*

- Scantlebury N, Sajic R, Campos A. Kinematic Analysis of Drosophila Larval Locomotion in Response to Intermittent Light Pulses. *Behavior Genetics*. May 2007;37(3):513-524.
- Schaafsma, A.W.; Dexter, J.; Clear, R.; Savard, M.E., 2004 Methods and issues regarding detection of deoxynivalenol, Fusarium-damaged kernels, and Fusarium spp. in commercial grain in Canada. *Canadian journal of plant pathology = Revue Canadienne de phytopathologie*, 2004 Oct-Dec, v. 26, no. 4, p. 443-452. 26 4
- Schimleck, L. R., Sturzenbecher, R., Jonesa, P. D., and R. Evans. 2004. "Development of wood property calibrations using near infrared spectra having different spectral resolutions." *Journal of Near Infrared Spectroscopy* 12, no. 1: 55.
- Segmentation of plants and weeds for a precision crop protection robot using infrared images. 1996 *IEE Proceedings - Vision, Image and Signal Processing* 143, no. 2: 118-124.
- Sena Jr, D.G; Pinto, F.A.C.; Queiroz, D.M.; Viana, P.A., 2003, Fall Armyworm Damaged Maize Plant Identification using Digital Images *Biosystems Engineering*, 2003, Vol. 85 Issue: Number 4 p449-454, 6p;
- Shakhnarovich, Darrell, Indyk (editor) 2005, *Nearest-Neighbor Methods in Learning and Vision*, The MIT Press, 2005, ISBN 0-262-19547-X
- Shao, Bin , 2008, A real-time computer vision assessment and control of thermal comfort for group-housed pigs., , *Computers and electronics in agriculture*, 2008 June, v. 62, no. 1, p. 15-21.
- Sheikh, Hamid Rahim; Bovik, Alan C. *Image Information and Visual Quality*. *IEEE Transactions on Image Processing*, Feb2006, Vol. 15 Issue 2, p430-444, 15p;
- Shen, J., & Yu, Q. (1991). Application of image processing technique in the determination of soil deformation. *International Agricultural Mechanization Conference*. Proceedings of a conference held in Beijing, China, 16-20 October 1991, 2.163-168. Retrieved from CAB Abstracts database.
- Si, YongSheng, Jun Qiao, Gang Liu, Rui Gao, and Bei He. 2010. "Recognition and location of fruits for apple harvesting robot." *Nongye Jixie Xuebao* =

- Transactions of the Chinese Society for Agricultural Machinery* 41, no. 9: 148-153.
- Singh S, M.Montemerlo 1998, Grading of vegetative cuttings using computer vision, *Advanced Robotics*, Vol 12, No. 5, pp 551-564
- Slice D (2007). Geometric morphometrics. *Annu. Rev. Anthropol.* 36:261–81.
- Smith D B, S.D.Askew, W.H.Morris, D.R.Shaw, M.Boyette 2000, Droplet size and leaf morphology effects on pesticide spray deposition, *Transactions of the ASAE*; 43(2), 255-259;
- Souza, S. R. L.; Naas, I. A.; Moura, D. J. 2008, Computational vision use for evaluation of confined dairy cows behavior.; Central theme, technology for all: sharing the knowledge for development. *Proceedings of the International Conference of Agricultural Engineering, XXXVII Brazilian Congress of Agricultural Engineering, International Livestock Environment Symposium - ILES VIII, Iguassu Falls City, Brazil, 31st August to 4th September, 2008; Bonn: International Commission of Agricultural Engineering (CIGR), Institut für Landtechnik, 2008*
- Steger C, Ulrich M, Wiedemann C (2007) *Machine vision algorithms and applications*. Wiley-VCH Ed., Berlin.
- Subramanian, V., Burks, T., & Arroyo, A. (2006). Development of machine vision and laser radar based autonomous vehicle guidance systems for citrus grove navigation. *Computers and Electronics in Agriculture*, 53(2), 130-143. doi:10.1016/j.compag.2006.06.001.
- System for measuring the coordinates of tire surfaces in transient conditions when rolling over obstacles: Description of the system and performance analysis. 2008, *Review of Scientific Instruments*, 79(6), 065105
- Tan, J. (2004). Meat quality evaluation by computer vision. *Journal of Food Engineering*, 323, 61(1): 27-35.
- Tellaeche, Alberto; Pajares, Gonzalo; Burgos-Artizzu, Xavier P.; Ribeiro, Angela 2001 A computer vision approach for weeds identification through Support Vector Machines,. *Applied Soft Computing*, January 2011, Vol. 11 Issue: Number 1 p908-915, 8p;

- Tillet N D , T.Hague, S.J.Miles, 2002, Inter-row vision guidance for mechanical weed control in sugar beet, *Comp.and Electr. in Agric.*
- Tillet N D, W.He, R.D.Tillet, 1995, Development of a vision guided robot manipulator for packing horticultural produce, *J. of Agric.Eng.Res.* (61):3, 145-154
- Tillett R.D. 1991. Image analysis for agricultural processes: a review of potential opportunities. *Journal of agricultural Engineering research.* 50. pp.247-258. 1991.
- Ting, K C, P.P.Ling, G.A.Giacomelli 1996, Research on flexible automation and robotics for plant production at rutgers university, , *Adv.Space Res*
- Tuceryan, M. (1998). Texture analysis. In C.H. Chen, L. F. Pau, and P. S. P. Wang (Eds.), *The Handbook of Pattern Recognition and Computer Vision* 2nd Ed. (pp. 207-248), New Jersey: World Scientific Publishing Co.
- Ulissi V, Menesatti P, Benincasa P., Guiducci M, 2010, Estimation of tomato nutritional status by VIS-Nir portable spectrophotometric system. XVIIth World Congress of the International Commission of Agricultural Engineering (CIGR), Québec City, Canada June 13-17, 2010: 108
- Urbani G., Menesatti P., Niciarelli I., D'Andrea S. 2003. Tecnologie optoelettroniche nel mondo agricolo: sistemi di spettrofotometria di immagine. *MONDO MACCHINA.* N. 7-8/2003. pp. 34-38.
- Urbani G., Menesatti P., Solaini S. 2002. Tecnologie opto-elettroniche in agricoltura: nuove prospettive e indirizzi di ricerca. *Mondo Macchina.* N. 6/2002. pp. 35-38.
- Valous, Nektarios A.; Drakakis, Konstantinos; Sun, Da-Wen. 2010, Detecting fractal power-law long-range dependence in pre-sliced cooked pork ham surface intensity patterns using Detrended Fluctuation Analysis. *Meat Science,* Oct2010, Vol. 86 Issue 2, p289-297
- Van Henten, E.J.; Hemming, J.; van Tuijl, B.A.J.; Kornet, J.G.; Meuleman, J.; Bontsema, J. 2002, An Autonomous Robot for Harvesting Cucumbers in Greenhouses *Autonomous Robots,* November 2002, (13), 241-258
- Vision-guidance of agricultural vehicles. 1995 *Autonomous Robots,* 2(1), 65-76.

- Vranken, E.; Chedad, A.; Aerts, J. M.; Berckmans, D.; 2005, Improving the accuracy of automatic broiler weighing by image analysis. Precision livestock farming '05; Wageningen: Wageningen Academic Publishers, 2005, 265-271
- Vukusic P, Sambles J. Photonic structures in biology. Nature. August 14, 2003;424(6950):852.
- Wan Ishak, W. I.; Kit, W. H.; Awal, M. A.; Pertanika, 2010, Design and development of eggplant harvester for gantry system. Journal of Science & Technology; 18 (2), UPM Selangor: Universiti Putra Malaysia Press, 2010, 231-242
- Wang B. Biophotonics: Visions for Better Health Care. Medical Physics. April 2007;34(4):1522-1523.
- Wang, X., Yang, W., Wheaton, A., Cooley, N., & Moran, B. (2010). Automated canopy temperature estimation via infrared thermography: A first step towards automated plant water stress monitoring. Computers and Electronics in Agriculture, 73(1), 74-83
- Wold, S., K. Esbensen, and P. Geladi. 1987. Principal component analysis. Chemometric and intelligent Lab. Sys 2: 37-52.
- Wrigley, C. W., Potential methodologies and strategies for the rapid assessment of feed-grain quality. Special issue: Premium grains for livestock. Australian Journal of Agricultural Research 50 (5), 1999, p 789-805
- Xin, H., 1999, Assessing swine thermal comfort by image analysis of postural behaviors.; Journal of Animal Science; 77 1999, 1-9
- Yang I-Chang, Delwiche S R, Chen S, Lo Y. Martin, 2009 Enhancement of Fusarium head blight detection in free-falling wheat kernels using a bichromatic pulsed LED design, Optical Engineering Jan 2009, Volume 48, Number 2: 023602
- Yang, Y., Teng, G., & Li, B. (2006). Determination of pig weight from 2D images. Journal of China Agricultural University, 11(3), 61-64.
- Zayas, I. Y., Martin, C. R., Steele, J. L., and Katsevich, A. 1996. Wheat classification using image analysis and crush-force parameters. Trans ASAE 39:2199-2204.
- ZEISS S.p.A. 1990. Elaborazione ed analisi d'immagine. Documento informativo ZEISS
- Zeng, Qingbing; Miao, Yubin; Liu, Chengliang; Wang, Shiping , 2009 Algorithm based on marker-controlled watershed transform for overlapping plant fruit

segmentation, *Optical Engineering*, February 2009, Vol. 48 Issue: Number 2
p027201-027201-10

Zhang, K., Yang, L., & Zhang, T. (2010). Object locating method of laser-assisted machine vision for strawberry-harvesting. *Nongye Jixie Xuebao = Transactions of the Chinese Society for Agricultural Machinery*, 41(4), 151-156.

Zhang, L., Wang Shu, M., Chen Bing, Q., & Zhang Hong, X. (2007). Crop-edge detection based on machine vision. *New Zealand Journal of Agricultural Research*, 50(5), 1367-1374

Zhou, Jun, Rui Liu, and GaoYang Zhang. 2010. "Design of fruit harvesting robot based on stereo vision." *Nongye Jixie Xuebao = Transactions of the Chinese Society for Agricultural Machinery* 41, no. 6: 158-162.

Parte II

Attività e risultati sperimentali pubblicati

3. ANALISI DELLA FORMA PER IMMAGINI (*Shape imaging*)

3.1 Introduzione

3.1.1 La forma e la taglia

The study of form may be descriptive merely, or it may become analytical. We begin by describing the shape of an object in simple words of common speech: we end by defining it in the precise language of mathematics; and the one method we end by defining it in the precise language of mathematics; and the one method tends to follow the other in strict scientific order and historical continuity.

(D'Arcy Thompson - 1915)

L'idea di forma è uno dei basilari concetti che pervadono le scienze sociali, fisiche, ingegneristiche e biologiche.

La funzionalità dei manufatti industriali, il comportamento etologico o biologico delle specie, o il grado di preferenza di un prodotto alimentare, sono tutti aspetti influenzati dalla forma.

In termini molto generali, la forma è rappresentata dalla sagoma dell'aspetto esteriore di un oggetto, ossia il suo contorno esterno. Essa è indipendente dalla dimensione dell'oggetto che invece è descritta dalla taglia.

Oltre alla taglia, esistono altri numerosi descrittori che includono la configurazione, il colore, la disposizione, la simmetria e la relazionalità tra le parti. Alcuni di questi elementi possono essere considerati propri della forma, nel senso che ne completano la specifica descrizione; altri, come ad esempio il colore, sono spesso confusi o integrati in una valutazione complessiva, soprattutto in ambito descrittivo e qualitativo.

In realtà, l'abilità umana nel discriminare facilmente le forme notandone le differenze è così ben integrata che le risposte comportamentali sono spesso inconse. L'occhio

carpisce differenze sottili osservando i contrasti, prima attraverso le differenze maggiori cosicché facilmente può avvenire il processo di discriminazione (Lestrel, 1997).

3.1.2 Rilevanza nell'ingegneria agraria, agroindustria e alimentare

La forma può essere considerato uno degli aspetti caratteristici e caratterizzanti di:

- oggetti e manufatti artificiali, comprendendo con essi le macchine e attrezzature agricole;
- prodotti agricoli e alimentari;
- organismi viventi di interesse agrario.

Si possono, complessivamente, attribuire alla forma (e alla taglia) di questi soggetti così differenti, funzioni importati e disparate, legate a:

- funzionalità tecnica e operativa;
- ergonomia ed estetica;
- specificità e adattabilità ecologica, etologica e ambientale;
- capacità produttiva e riproduttiva;
- riconoscibilità e identificabilità dei prodotti e degli alimenti, sia in termini più prettamente normativi che legati alla sfera culturale e agrosociologica;
- gradimento, preferenza e accettabilità da parte del consumatore con importanti risvolti in termini di scelte economiche, marketing e redditività.

Per quanto riguarda le macchine e le attrezzature agricole, la forma rappresenta un importante fattore di funzionalità, con ampie connessioni con gli aspetti operativi, produttivi, ergonomici, della sicurezza e di gradimento (Rajvir, 1998).

Possiamo pensare, quanto influenzi la forma di un organo di lavorazione (ad es. il vomere dell'aratro), la qualità e la produttività dell'operazione stessa; oppure come la forma di un automezzo, sia essa automobile o trattore, eserciti un maggiore o minore *appeal* sul potenziale acquirente, in relazione al giusto mix di forma, dimensioni e colore (Fig 1). Dunque la forma è spesso interpretata o considerata come

intrinsecamente legata alla funzionalità, ossia alla sostanza operativa del bene o prodotto.

Fig. 1. Stylists queue up to shape the tractors of the future (dal sito Farmers Guardian, 21 December 2009, <http://www.farmersguardian.com/home/machinery/machinery-features/stylists-queue-up-to-shape-the-tractors-of-the-future/29438.article>)



Nelle scienze biologiche, lo studio della forma rappresenta una disciplina consolidata e in forte espansione, grazie alle forti relazioni di questa peculiare caratteristica con lo sviluppo, la differenziazione, l'adattamento e il comportamento delle specie e degli ecotipi. E' proprio in ambito biologico che lo studio della morfometria, come di seguito sarà meglio specificato, prende avvio e si sviluppa come importante disciplina scientifica (Costa, 2004).

Molto complessa è anche la funzione della forma per i prodotti agricoli e alimentari (Costa *et al*, accepted)

L'analisi della forma di prodotti agricoli è spesso un attributo necessario in campo applicativo per la descrizione delle cultivar e/o dei genotipi, la valutazione della loro

qualità, la stima della biodiversità a livello varietale, la valutazione delle preferenze dei consumatori, l'eventuale ereditabilità delle caratteristiche morfologiche ed infine la valutazione di anomalie di sviluppo (sia in campo morfologico che colorimetrico). Per esempio, ad oggi, la descrizione delle cultivar e dei genotipi avviene solo da un punto di vista visivo: ciò permette l'analisi di un limitato numero di campioni e, in alcuni casi, anche la non oggettività di tali analisi.

3.1.3 Relazione forma-qualità

Come è stato accennato, la forma di un prodotto o di un oggetto, contribuisce, spesso significativamente, a stabilire la sua qualità funzionale e merceologica.

Nel caso ad esempio dei prodotti ortofrutticoli, la forma insieme alla dimensione (taglia), definisce un attributo basilare della proprietà organolettica relativa all'*aspetto* del prodotto.

Nella percezione sensoriale, la forma può essere integrata al colore (*base, sovracolore, luminosità, brillantezza*) e alla superficie (*struttura, tessitura*).

Le alterazioni qualitative a carico di questo attributo, in termini di *difetti e danni della forma, asimmetrie, deformità*, sono quasi sempre inserite nei disciplinari per la produzione e commercializzazione dei prodotti ortofrutticoli (figura 2).

Un altro esempio, che sarà illustrato nel seguente paragrafo come risultato di una specifica attività sperimentale, riguarda la relazione tra forma e danno, conseguente all'attacco di un patogeno fungino (*Fusarium culmorum*) su cariosside di frumento tenero. In questo caso, la fusariosi determina, come effetto secondario, un raggrinzimento e restringimento della cariosside (*striminzitura*) che, se opportunamente misurata, potrebbe rappresentare un indice indiretto dell'attacco patogenico.

Tabella 1. Proprietà e caratteristiche organolettiche dei prodotti ortofrutticoli (Zerbini, 1989 – rielaborato)

proprietà organolettica	caratteristica descrittiva	condizione di appagamento del consumatore
Gusto	odore, aroma, sapore	<i>indice stato freschezza, conservabilità</i>
Consistenza	primaria	<i>composizione chimica, tessitura e porosità</i>
	secondaria	<i>caratteristiche reologiche (elasticità, viscosità e adesività)</i>
	terziaria	<i>sensazione alla masticazione</i>
Colore	tinta, luminosità saturazione intensità	
Aspetto	forma	<i>uniformità presentazione, difetti o danni</i>
	dimensione (o taglia)	<i>calibro, pezzatura</i>

Fig. 2. Esempi di difetti di forma su carota (Neri et.al., 2002)



Fig. 3. Esempio di variazioni di forma di cariossidi di frumento tenero in relazione ad attacchi di fusariosi (Mernesatti et al., 2009)



3.1.4 Dalla valutazione soggettiva alla misura quantitativa della forma

La valutazione della forma, soprattutto nella sua funzionalità estetica e di gradimento da parte del consumatore, è molto spesso effettuata attraverso un'analisi soggettiva, utilizzando descrittori semantici.

Tale descrizione semantica è ancora molto utilizzata sia nella descrizione dei caratteri morfologici di specie e cultivar o razze vegetali e animali, sia nella definizione dei difetti o danni nei protocolli o disciplinari di produzione.

L'analisi soggettiva della forma, a differenza di altri parametri qualitativi come ad esempio il colore o la consistenza, ha quindi ancora una ampia applicazione e notevole rilevanza economica e scientifica.

Questo dipende principalmente da due fattori concomitanti:

- la difficoltà di descrivere quantitativamente forme molto articolate
- la complessità di identificare e misurare appropriati descrittori delle forme complesse.

Ciononostante, l'attività scientifica e quella produttivo-commerciale richiede con sempre maggiore insistenza, lo sviluppo e l'applicazione di sistemi strumentali di misura, possibilmente automatizzabili, rapidi e di limitato costo.

Questo al fine di aumentare i punti e i campioni di misura, garantendo standardizzabilità e ripetibilità dell'analisi, per arrivare, in alcuni casi, allo sviluppo di sistemi di selezione online (Pallottino *et al.*, in press).

A tale richiesta, è possibile attualmente fornire risposte, sempre più precise ed adeguate, in relazione a due fattori risolutivi:

- - lo sviluppo e l'applicazione di modellistica e algoritmica specifica per la descrizione delle forme complesse, attraverso rappresentazione numerica e geometrica di queste;
- - la disponibilità tecnologica crescente di sistemi strumentali rapidi, efficienti e di costo limitato in grado di rilevare i parametri della forma necessari per la modellizzazione quantitativa.

Tale disponibilità è in riferimento alle tecnologie strumentali a base fisica, che integrano controlli elettronici e informazione multicanale. Esse stanno assumendo un ruolo sempre più importante nell'industria agroalimentare in relazione alla loro capacità di descrivere in modo più specifico le caratteristiche qualitative dei prodotti, indicando anche importanti informazioni sulla produzione e sull'origine o anche come metodi per una definizione rapida e oggettiva di stati (alterati o meno) di qualità (Menesatti, 2000a). Queste tecniche sono associate ad una elevata capacità analitica (multiparametrica e multicanale), e, spesso, presentano importanti caratteristiche di non distruttività del campione di misura, relativa semplicità di impiego, rapidità, portabilità in ambienti operativi, e interfacciabilità con altra sensoristica elettronica.

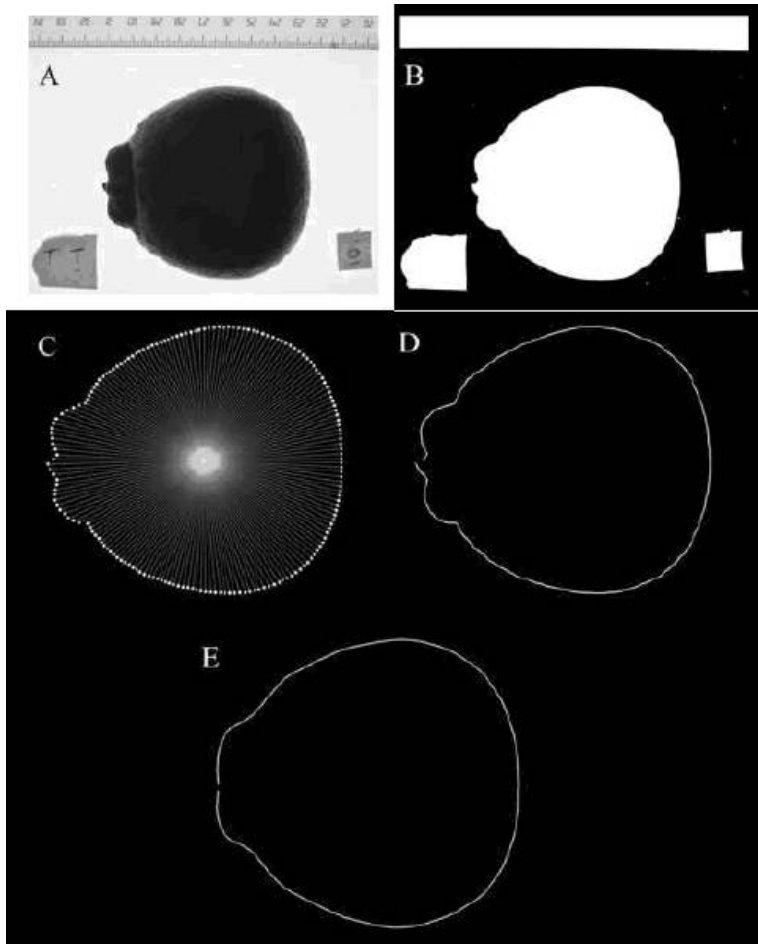
3.1.5 Analisi della forma per immagini: innovazione in ambito biofotonico

Tra le tecnologie a base fisica, più interessanti e ed efficaci per l'analisi strumentale della forma di prodotti agricoli, sono sicuramente da considerare le tecnologie fotoniche, soprattutto quelle che integrano l'informazione dell'immagine (Menesatti, 2000b).

L'immagine digitale che grazie ai sistemi di ultima generazione raggiunge risoluzione ottiche e cromatiche elevatissime, consente di determinare con altissima precisione l'esatto contorno o profilo di oggetti o prodotti. L'estrazione del profilo è condizione basilare per la successiva analisi quantitativa della forma.

Un esempio di processo di estrazione e analisi della forma di un prodotto ortofrutticolo, attraverso immagini digitali, è schematizzato in figura 4.

Fig. 4. Esempio di processo di analisi della forma di un prodotto ortofruitticolo attraverso analisi di immagine e modellizzazione matematica del profilo (Studio quantitativo per differenziare la forma di 17 genotipi di tarocchi: da Costa *et al.*, 2009)



- A. Immagine originale
- B. Estrazione del profilo
- C. Estrazione dei 180 punti sul profilo
- D. Conversione dei punti in outline
- E. Profilo dopo analisi Ellittica di Fourier

Altri fattori importanti nella configurazione di un sistema di analisi della forma per immagini, sono:

- caratteristiche del sensore ottico di acquisizione delle immagini (lineare o matriciale, CCD o CMOS, risoluzione ottica, risoluzione di colore, frame rate, rapporto segnale rumore, calibrazione del bianco, shuttering, ecc.);
- caratteristiche dell'impianto di illuminazione (potenza, intensità luminosa, spettro e temperatura di colore dell'illuminante);
- geometria di illuminazione-acquisizione rispetto all'oggetto di misura (illuminazione e ripresa diretta, retroilluminazione, luce pulsata, luce tangenziale, campo chiaro o campo scuro);
- disposizione del prodotto (posizione casuale o visualizzazione di specifiche parti);
- caratteristiche del profilo estratto (equiangolato o equispaziato);
- tipo di analisi quantitativa ,morfologica: tradizionale, geometrica, analisi ellittica di fourier, da cui derivare specifici descrittori del forma, semplici o complessi;
- tipologia di analisi classificatoria o predittiva: monovariata o multivariata, non supervisionata o supervisionata.

Il sistema, complessivamente, deve poi essere valutato in base a:

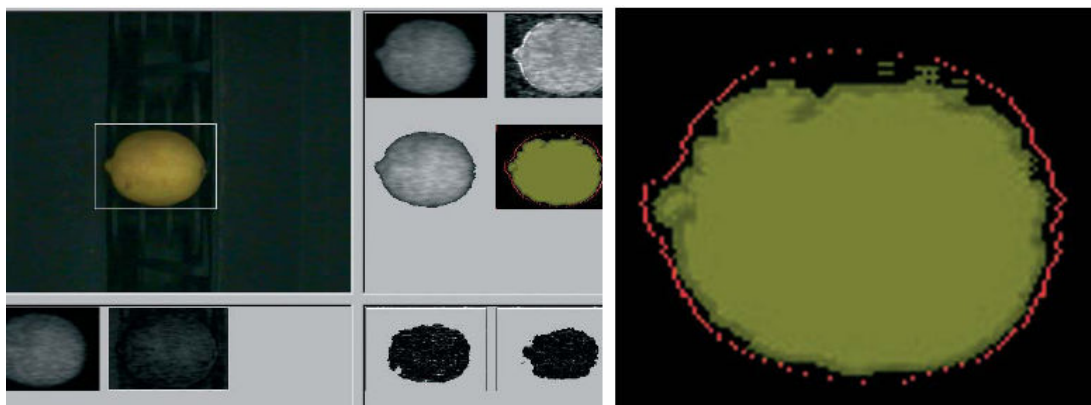
- efficacia ed efficienza di identificazione e misura della forma
- costo operativo;
- velocità di misura;
- accuratezza di misura;
- ripetibilità di misura;
- manutenibilità del sistema;
- livello di complessità operativa
- integrabilità con altri sistemi, soprattutto digitali e con sistemi informativi esterni.

Allo stato attuale delle conoscenze metodologiche e delle disponibilità tecnologiche è ben possibile ipotizzare lo sviluppo di siffatti sistemi, che possano operare con successo tecnico e vantaggio economico in differenti contesti soprattutto dell'agroindustria e dell'alimentare.

In effetti, i sistemi di selezione optoelettronica, soprattutto per gli ortofrutticoli, sono già molto diffusi nella pratica delle industrie di lavorazione e in quelle alimentari,

dove, tuttavia, la principale operazione richiesta alle macchine è la selezione dimensionale del prodotto, insieme a quella della identificazione di danni. Esistono, tuttavia, nuove generazioni di macchine, basate su sistemi ottici molto evoluti che acquisiscono in continuo immagini di frutti in rotazione, in grado di determinare anche variazioni significative del profilo, da cui derivare l'eventuale presenza di danni o difetti di forma (Fig. 5).

Fig. 5. Esempio di identificazione di difetto di forma su limone effettuato attraverso una macchina per la selezione di caratteristiche di qualità esterna di prodotti ortofrutticoli (fonte: Sistema Ultravision, Unitec Group – Lugop (RA) – Italia, <http://www.novatec-sa.gr/pics/upload/BrochureQualita.pdf>)



Tuttavia, anche nei sistemi industriali più avanzati, non è ancora stata sviluppata una integrazione tra l'apparato di visione artificiale, in grado di rilevare molto precisamente il profilo del frutto da diversi punti di osservazione e nuove metodiche di modellizzazione complessa della forma, che saranno di seguito meglio indicate sia a livello teorico che applicativo-sperimentale.

Ad esempio, uno dei metodi più avanzati di misura complessa della forma è rappresentato dall'Analisi Ellittica di Fourier (EFA), che sarà spiegata con maggior dettaglio nei prossimi paragrafi.

L'EFA consente di modellizzare matematicamente dei profili chiusi anche molto complessi a partire dalle coordinate estratte dal perimetro di un'immagine di un oggetto (frutto o prodotto), acquisita in modo automatico.

L'analisi EFA approssima il perimetro chiuso a una serie di equazioni armoniche ciascuna descritta da diversi coefficienti, ed è indipendente dalla dimensione (taglia) dell'oggetto e dal suo orientamento nell'immagine. L'insieme dei valori dei coefficienti di tutte le armoniche considerate rappresenta il dataset (multivariato) dei descrittori morfologici dei prodotti da classificare in base alla loro forma.

La combinazione di questa tecnica con metodi di classificazione multivariata, permette (Fig. 4):

- di parametrare numericamente forme complesse;
- di rappresentare geometricamente tali forme;
- calcolare e rappresentare una "forma media" di riferimento, indipendentemente dalla variabilità dimensionale dei prodotti;
- misurare eventuali variazioni di ciascun individuo da tale forma media;
- stabilire quindi soglie e percentuali per difetti e danni;
- classificare e distinguere cultivar, genotipi o ecotipi diversi, anche quando, caratterizzati da lievi variazioni di forma.

Una macchina per la selezione basata su tale sistema, dovrebbe quindi implementare tre principali sistemi:

- l'apparato di visione (traslatore, illuminazione, videocamera, convertitore A/D);
- l'algoritmo di estrazione del profilo e di computo delle EFA;
- il modello di predizione o classificazione multivariato, precedentemente sviluppato sulla base di attività sperimentali.

Una macchina di queste caratteristiche non è ancora disponibile sul mercato e il sottoscritto dottorando, insieme ad altri coautori, ha provveduto nel 2009 a presentare domanda di brevetto industriale relativa a tale sistema (Menesatti *et al.*, 2009).

Si ritiene, pertanto, che l'analisi quantitativa della forma possa rappresentare un interessante segmento innovativo nell'ambito del comparto che è stato complessivamente indicato con il termine di biofotonica.

La crescente importanza dell'analisi quantitativa della forma per differenti settori scientifici e industriali (biologia, ingegneria, agricoltura), associata alla disponibilità tecnologica di sistemi evoluti di visione artificiale e alla conoscenza di

modellizzazione numerica di forme complesse e metodi di classificazione multivariata, identifica, a mio avviso, una nicchia di forte innovazione tecnologica, con elevate potenzialità applicative soprattutto nel contesto agroindustriale italiano, storicamente caratterizzato dalla presenza di operatori molto specializzati e fortemente orientati alle richieste del comparto agroalimentare nazionale e internazionale.

3.1.6 Bibliografia

Costa C, Antonucci F, Pallottino F, Aguzzi J, Sun DW, Menesatti P, ACCEPTED.

Shape analysis of agricultural products by computer vision – a review of recent research advances. ACCEPTED TO FOOD AND BIOPROCESS TECHNOLOGY

Costa C, Menesatti P, Paglia G, Pallottino F, Aguzzi J, Rimatori V, Russo G, Recupero S, Reforgiato Recupero G, 2009. Discrimination of Tarocco sweet orange [*Citrus sinensis* (L.) Osbeck] varieties using opto-electronic elliptic Fourier based analysis of fruit shape. POSTHARVEST BIOLOGY AND TECHNOLOGY, 54: 38-47.

Costa C. 2004. Lo studio della forma in ittiologia: Morfometria geometrica e analisi dei profili. Ph.D. thesis in evolutionary biology and Ecology. University Rome Tor Vergata. 177 p.

Menesatti P, Antonucci F, Costa C, Santori A, Niciarelli I, Infantino A, 2009, Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels. 1st International Workshop on Computer Image Analysis in Agriculture, Potsdam, Germany 27 – 28 August 2009, Bornimer Agrartechnische Berichte - Heft 69, ISSN 0947-7314, Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V. (ATB)

Menesatti P, Costa C, Aguzzi J, Antonucci F. "Apparato di riconoscimento di prodotti agroalimentari", Verbale di Deposito Domanda di Brevetto per Invenzione Industriale, presso la Camera di Commercio Industria, Artigianato e Agricoltura di ROMA, Numero domanda: RM2009A000325, Data di deposito: 25/06/2009,

CCIAA di deposito: ROMA, Richiedente: Consiglio per la Ricerca e la Sperimentazione in Agricoltura (C.R.A.)

- Menesatti P. 2000a, Misura strumentale della qualità in ortofruitticoltura ITALUS HORTUS vol. 7. n. 2, pp. 12-19.
- Menesatti P. 2000b. Misura strumentale di aspetti della qualità degli alimenti attraverso tecnologie opto-elettroniche INDUSTRIE ALIMENTARI, n. 2, pp. 128-136
- Neri U., Beni C., Aromolo R., Miraglia R., Menesatti P., 2002. Metodologie per la valutazione della fertilità dei suoli e della qualità dei prodotti tipici: la carota di Maccarese – Convegno Nazionale VI Giornate Scientifiche SOI (Società Orticola Italiana) – Spoleto 23-25/4/2002 – Atti. pp. 493-494.
- Pallottino F, Menesatti P, Costa C, Paglia G, De Salvador F R, Lolletti D. IN PRESS Image analysis techniques for automated hazelnut peeling determination. FOOD AND BIOPROCESS TECHNOLOGY. DOI 10.1007/s11947-009-0211-1. (IF2009= 2.238)
- Yadav Rajvir, Tewa V. K., 1998 Tractor operator workplace design—A review; Journal of Terramechanics, Volume 35, Issue 1, January 1998, Pages 41-53.

3.2 La morfometria

L'analisi e la descrizione quantitativa (misura) della forma è complessivamente denominata morfometria (Blackith e Reyment, 1971).

L'esigenza di espressione quantitativa, soprattutto nella identificazione e classificazione di oggetti, prodotti o specie, è duplice:

1. rappresentare la forma e le sue variazioni graficamente e geometricamente (espressione geometrica);
2. misurare la forma, analizzandola quindi in modo quantitativo (espressione numerica).

L'analisi può essere sviluppata a differenti livelli di complessità e accuratezza descrittiva, in funzione del numero e della tipologia di descrittori utilizzati.

In generale, possiamo distinguere:

1. morfometria tradizionale (analisi attraverso descrittori semplici diretti e relativi);
2. morfometria geometrica;
3. analisi dei profili chiusi.

3.2.1. Morfometria tradizionale

La morfometria tradizionale affida l'analisi della forma alla misura di una serie di parametri descrittivi di tipo assoluto o relativo.

I descrittori assoluti più utilizzati sono: l'area, il perimetro, estensioni di parti o lati, aperture angolari, gli assi dimensionali diretti (minore, maggiore o medio) o di inscrizione o circoscrizione in figure geometriche regolari.

Questi parametri offrono spesso una descrizione della forma osservata, accurata, riproducibile e facilmente intuibile. Tuttavia, tale analisi è dipendente dalla taglia degli oggetti, rappresentandone, spesso, la principale finalità di questa tipologia di misura.

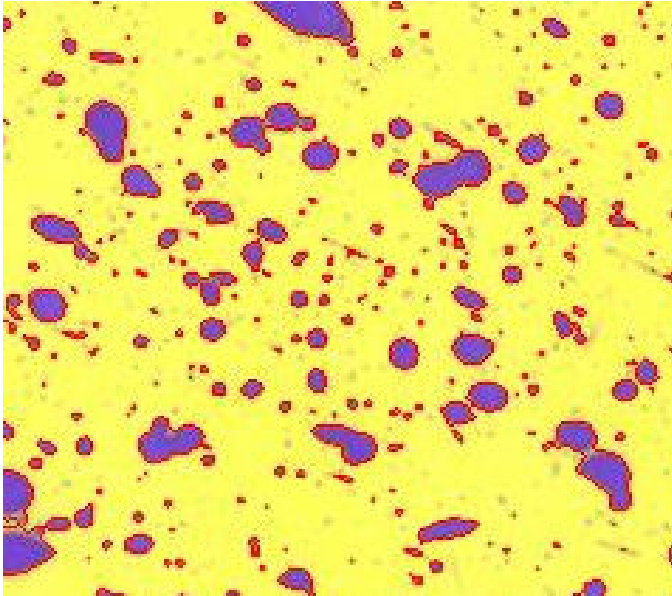
Ad esempio, la principale operazione di selezione qualitativa di prodotti ortofrutticoli in postraccolta, è rappresentata dalla calibratura che è effettuata sulla base di descrittori semplici assoluti (diametri) ed è finalizzata all'ottenimento di classi dimensionali omogenee di prodotto (Menesatti, 2000).

I descrittori relativi sono invece rappresentati da rapporti dimensionali dello stesso oggetto o di oggetti di riferimento.

Un esempio, in campo agromeccanico, è fornito dai descrittori di misura della forma e della taglia delle singole impronte delle gocce di irrorato (figura 1), al fine dell'analisi della qualità di distribuzione delle macchine irroratrici, effettuata attraverso il metodo delle cartine idrosensibili (Menesatti, 1998). In questo caso, i principali descrittori sono relativi all'area e ai diametri minimi e massimi (diretti) e, per quelli relativi, al rapporto tra essi e alla circolarità della goccia (rapporto tra il diametro medio e il diametro del cerchio di area equivalente all'oggetto).

Anche per l'analisi della porosità di pane artigianale ottenuto con nuovo lieviti, sono stati utilizzati parametri morfologici diretti e relativi (Paoletti et al., 1998).

Fig. 1. Esempio di rilievo automatico per immagini digitali, di forma e dimensioni di impronte di gocce di prodotto fitoiatrico in fase acquosa liquida, asperso su cartina idrosensibile



L'applicazione di tecniche di analisi statistica, soprattutto multivariata, ha potenziato la capacità analitica e discriminatoria dell'analisi morfologica tradizionale.

Rimane, tuttavia, esclusa l'organizzazione geometrica degli elementi misurati, con l'impossibilità, quindi, di affrontare un livello di analisi superiore, collegato alla forma come avente delle peculiari proprietà geometriche: nella morfometria tradizionale non c'è spazio per alcuna informazione di carattere geometrico (Bookstein, 1993).

3.2.2 La Morfometria Geometrica

La morfometria geometrica ha rappresentato un significativo passo in avanti, rispetto alla morfometria tradizionale, in quanto consente l'intergrazione della descrizione numerica e di quella geometrica della forma.

La morfometria geometrica si sviluppa in ambito biologico a partire dal 1970 con la principale finalità di evidenziare analogie e differenze tra le specie, considerando la variabilità morfologica individuale.

Questa disciplina si occupa dello studio della forma attraverso l'utilizzo, come dati su cui effettuare le elaborazioni, di punti definiti da coordinate in due o tre dimensioni, i *landmarks*. I landmarks sono coordinate di punti omologhi nello spazio cartesiano (Bookstein, 1991), rilevabili su ciascun organismo o oggetto e scelti come rappresentativi della geometria della struttura studiata (Rohlf e Marcus, 1993). L'omologia può essere intesa in senso geometrico o in senso strettamente biologico a seconda delle finalità dello studio morfometrico.

I diversi organismi vengono così rappresentati da diverse configurazioni di landmarks, che, se scelti appropriatamente, rappresentano la base per una futura interpretazione biologica dei risultati delle analisi. In generale i landmarks vengono classificati sotto tre tipologie:

- TIPO 1: indicativi di omologie supportate da consistenti prove da un punto di vista biologico;
- TIPO 2: rappresentano omologie definite solo con criterio geometrico;
- TIPO 3: sono definiti con una funzione matematica.

I landmarks consentono una descrizione immediata ed un'espressione sintetica della forma di un individuo. Attraverso i landmarks si possono effettuare elaborazioni statistiche, ma la vera potenzialità della morfometria geometrica è la capacità di questa tecnica di descrivere quantitativamente la deformazione, definibile come la trasformazione tra una configurazione di landmarks e un'altra, spesso di riferimento generale o medio (Bookstein, 1993).

In termini di morfometria geometrica, la forma (*form*) si compone di dimensioni (*size*) e configurazione (*shape*) che è data dalla posizione reciproca che i landmarks rappresentano. Le due componenti sono analizzate separatamente (Corti, 1996), con la duplice finalità di:

- Confrontare configurazioni (*shapes*) diverse e quantificare la differenza tra individui come distanza tra essi;
- Definire quelle variabili attraverso cui descrivere la deformazione fra diverse forme.

Il primo obiettivo si può raggiungere realizzando una sovrapposizione ottimale di un individuo su un altro in modo che i suoi landmarks si allineino il più possibile con le

posizioni dei corrispondenti landmarks del secondo. Le differenze in *shape* possono così essere visualizzate e quantificate come differenze, che chiamiamo residui, nella relativa posizione dei landmarks corrispondenti (Rohlf e Slice, 1990). Si sviluppano su questo criterio i cosiddetti “metodi di sovrapposizione”, attraverso i quali ottenere una sovrapposizione di forme eliminando informazioni circa la variabilità in orientamento, posizione e dimensioni, attraverso la realizzazione di una sovrapposizione ottimale dei landmarks dei diversi individui.

Il punto di partenza per tale standardizzazione consiste nel trovare il centroide di ciascun individuo che abbiamo rappresentato con la configurazione dei landmarks. Il centroide non è altro che la media aritmetica delle coordinate per ciascuna delle dimensioni dei punti presi in esame: esso è quindi individuato come una coppia di coordinate. In un sistema di assi cartesiani ortogonali possiamo così traslare gli individui fino a sovrapporne i centroidi nell’origine.

Le dimensioni degli individui vengono rappresentate attraverso il *centroid size* (CS) definito (Bookstein, 1991) come la radice quadrata della sommatoria delle distanze al quadrato tra il centroide ed ogni punto omologo; grandezza questa che utilmente riesce a rappresentare una caratteristica unitaria di una intera configurazione di landmarks, cioè la sua taglia.

$$CS = \sqrt{\sum_{i=1}^k (x_c - x_i)^2}$$

Le prime due operazioni prefigurate (sovrapposizione ed uniformazione per taglia) possono essere quindi effettuate senza equivoci nel sistema cartesiano attraverso la definizione di matrici di trasformazione della nostra matrice di coordinate. Per l’ultima delle operazioni previste, quella della rotazione, si possono applicare diverse matrici, a seconda del criterio utilizzato.

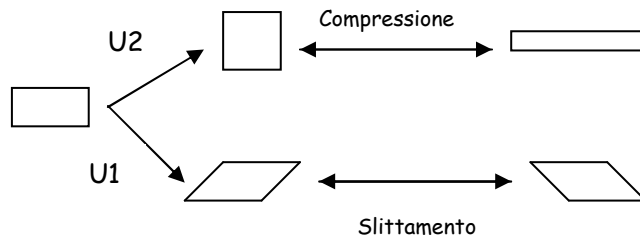
Attraverso il metodo dei minimi quadrati (GLS, *Generalized Least Square* – Rohlf e Slice, 1990) si sceglie un angolo θ per definire la matrice di rotazione costituita da seni e coseni dello stesso, in modo che la somma dei quadrati delle distanze di ciascun landmark di un individuo dal corrispondente landmark di un altro individuo preso come riferimento siano minime.

Un altro metodo è quello definito “*resistant fit*” o GRF (*General Resistant Fit* – Rohlf e Slice, 1990) che consiste nel considerare le mediane ripetute fra i rapporti fra le

differenze nelle posizioni dei landmarks per scegliere l'angolo ottimale secondo cui ruotare il valore per uniformare la taglia. Questo ultimo metodo permette di ottenere risultati più precisi allorquando si abbia di fronte un campione in cui le differenze fra forme siano più legate a pochi landmarks, e cioè a una porzione limitata della figura. I risultati delle suddette analisi sono rappresentate da grafici in cui viene visualizzata la sovrapposizione, in ciascun landmark, di tutte le configurazioni su quella di riferimento, detta anche *consenso*.

Entrambe i metodi possono essere ulteriormente perfezionati quando si voglia scomporre la differenza, descritta come trasformazione, fra due o più configurazioni, in una componente uniforme ed una localizzata o non uniforme (Rohlf e Slice, 1990). La variazione uniforme si riscontra nel momento in cui le forme confrontate siano differenziabili unicamente da variazioni su ampia scala (vedi fig. 2), riguardanti la forma nella sua interezza, mentre una trasformazione è tanto meno uniforme tanto più è localizzata in una porzione limitata della figura.

Fig. 2. Trasformazioni uniformi (Costa, 2004)



Le trasformazioni uniformi tra una configurazione e l'altra sono le compressioni (*stretching*) lungo una delle direzioni degli assi ortogonali, ma anche gli slittamenti (*shearing*) di una parte sull'altra.

Il modello che consente di descrivere e approssimare la quantificazione della deformazione nel suo complesso, è il *thin-plate spline* (Bookstein, 1991). Questo metodo permette di visualizzare la deformazione attraverso griglie di deformazione dette *splines*. Il modello definisce inoltre una nuova base funzionale, a partire dalla configurazione di consenso. Su questa nuova base è possibile proiettare tutti i residui relativi agli individui nello spazio definito dalla base funzionale stessa. Queste proiezioni vanno a comporre una nuova matrice che viene detta *Weight Matrix* (W).

Se sono presenti entrambe le componenti, uniforme e non uniforme, la matrice viene detta W' . Tale matrice contiene tutte le informazioni contenute nelle configurazioni originali dei landmarks, tranne quelle relative alla taglia (analizzabile indipendentemente come CS), la componente uniforme, le informazioni relative all'orientamento degli individui. Si possono effettuare a questo punto sulla *Weight Matrix* tutte le analisi statistiche necessarie. Ogni riga della *Weight Matrix* rappresenterà un individuo, ogni colonna una variabile. Una delle analisi statistiche che si possono compiere a partire dalla *Weight Matrix* consiste in una Analisi delle Componenti Principali. Tale analisi applicata a dati morfometrici di questo tipo viene detta *Relative Warp Analysis* (Rohlf, 1992). Gli assi della RWA possono essere messi in relazioni con differenti tipi di variabili come la taglia o variabili legate alla disposizione degli individui nello spazio o a differenti tipi di variabili come l'ecologia o il trofismo. La *Weight Matrix* può anche essere utilizzata per condurre un'Analisi delle Variate Canoniche che ha invece una maggior potenza nella discriminazione tra i gruppi. Il modello permette una grande versatilità per la rappresentazione grafica aumentando notevolmente la facilità di presentazione dei risultati, poiché rende possibile la visualizzazione della deformazione attraverso gli *splines* in relazione alla posizione degli individui lungo gli assi multivariati, siano essi componenti principali o assi canonici.

3.2.3 Analisi dei profili – Analisi Ellittica di Fourier

Uno dei principali problemi dell'applicazione della morfometria geometrica in contesti misuristici industriali o agroalimentari, riguarda la difficoltà nell'automatizzare l'estrazione dei landmarks. Inoltre, in tali contesti, molto spesso, la forma dei prodotti o degli oggetti da misurare non presenta particolari o numerosi punti di omologia.

Immaginiamo, ad esempio, di dover studiare modificazioni della forma per discriminare frutti di cultivar di nocciolo o l'effetto della striminzitura sulle cariossidi di frumento. In entrambi i casi, che saranno oggetto di trattazione specifica nei seguenti paragrafi, gli oggetti di misura presentano pochi punti di omologia (uno o

due al massimo) non facilmente identificabili attraverso tecniche di analisi di immagine automatica.

Per l'analisi morfologica complessa, in questi casi, si può ricorrere alle metodiche di analisi dei profili aperti o chiusi.

Nell'ambito di tali metodiche, quella che è considerata la più efficace, quando si analizzano profili chiusi, è l'analisi di Fourier, in particolare l'analisi ellittica di Fourier (EFA) (Rohlf e Archie, 1984; Ferson et al., 1985).

Questo metodo (sviluppato da Kuhl e Giardina, 1982) è basato su decomposizioni separate di Fourier delle prime differenze delle coordinate x e y (Δx_i and Δy_i) come una funzione parametrica della distanza cordale cumulata, t , dei punti lungo il profilo. La distanza t è scalata per andare da 0 a 2π radianti. I coefficienti di Fourier per la k^{th} armonica della proiezione di x sul profilo sono:

$$A_k = \frac{T}{2p^2\pi^2} \sum_{i=1}^p \frac{\Delta x_i}{\Delta t_i} \left(\cos \frac{2\pi k t_i}{T} - \cos \frac{2\pi k t_{i-1}}{T} \right)$$

$$B_k = \frac{T}{2p^2\pi^2} \sum_{i=1}^p \frac{\Delta x_i}{\Delta t_i} \left(\sin \frac{2\pi k t_i}{T} - \sin \frac{2\pi k t_{i-1}}{T} \right)$$

dove p è il numero di distanze lungo il profilo, $\Delta x_i = x_i - x_{i-1}$, Δt_i è la distanza cordale delle distanze tra i punti $i-1$ e i , t_i è la lunghezza cumulata di tale distanza fino alla distanza i , e $T (=t_p)$ è la lunghezza totale del contorno del profilo. E' da notare che $\Delta x_1 = x_1 - x_p$. Il termine costante per x è:

$$A_0 = x_0 + \frac{1}{T} \sum_{i=1}^p \left(\frac{\Delta x_i}{2\Delta t_i} (t_i^2 - t_{i-1}^2) + \left(\sum_{j=1}^{i-1} \Delta x_j - \frac{\Delta x_i}{\Delta t_i} t_{i-1} \right) \Delta t_i \right)$$

dove la somma con i limiti di $i-1$ sono uguali a zero quando $i = 1$. Questa equazione rappresenta una integrazione numerica di $\frac{1}{T} \int_0^T x(t) dt$.

Dato un set di coefficienti dell'ellittica di Fourier, una curva può essere disegnata utilizzando le equazioni (per $t = 0$ a 2π) e n armoniche:

$$x(t) = A_0 + \sum_{k=1}^n A_k \cos kt_k + B_k \sin kt_{k-1}$$

$$y(t) = C_0 + \sum_{k=1}^n C_k \cos kt_k + D_k \sin kt_{k-1}$$

I coefficienti per le proiezioni y , C_i , D_i e C_0 sono calcolabili nello stesso modo (semplicemente sostituendo le x con y nell'equazione sopra riportata. Questi quattro parametri sono stimati per ogni armonica.

Il numero totale di armoniche può essere calcolato per ogni profilo ed è uguale alla metà del numero totale di punti estratti lungo il profilo ('Nyquist frequency'). Come per tutti i metodi di Fourier, comunque, la maggior parte dell'informazione è racchiusa nelle prime armoniche. Inoltre, essendo le armoniche ortogonali tra loro (quindi indipendenti) quelle meno informative possono essere escluse dall'analisi senza influire sulle alter e sul risultato finale (Crampton, 1995).

Il numero di armoniche corretto viene stimato dal 'Fourier power spectrum' medio. La 'Fourier power' di un armonica è proporzionale alla sua ampiezza e fornisce una misura della quantità di informazione sulla forma descritta da tale armonica. Per la k^{th} armonica la 'Fourier power' è data dall'espressione:

$$\text{Fourier power} = (A_k^2 + B_k^2 + C_k^2 + D_k^2)/2$$

Dove A B C D sono i coefficienti di Fourier delle k^{th} armonica descritti precedentemente. La serie di Fourier può essere troncata al valore di k per il quale la media cumulate della 'Fourier power' è pari, ad esempio al 99.99% della media totale della 'Fourier power'. per ogni profilo, la 'Fourier power' totale è calcolata come la somma, da 1 a k , delle 'Fourier power' individuali e dove k è uguale alla 'Nyquist frequency' (Crampton, 1995).

L'analisi Ellittica di Fourier consente di analizzare profili chiusi anche molto complessi. Ha una maggiore performance rispetto alla morfometria geometrica laddove siano pochi i landmarks o non ben identificabili.

Inoltre, l'EFA è completamente automatizzabile in sistemi di visione artificiale e il trattamento delle numerose informazioni conferite dai coefficienti delle armoniche, con modellistica multivariata incrementa fortemente la capacità discriminatoria di questa potente tecnica di analisi della forma.

3.2.4 Bibliografia

- Blackith, R.; Reyment, R., 1971. Multivariate morphometrics. Academic Press
- Bookstein, F. L., 1991. Morphometric tool for landmark data. Cambridge Univ. Press.
- Bookstein, F. L., 1993. A brief history of the morphometrics synthesis. In: Marcus, F. L.; Bello, E.; Garcia-Valdecasas (Eds.), 1993. Contribution to morphometrics. Monografias del Museo Nacional de Ciencias Naturales, Madrid: 15-40.
- Corti, M., 1996. Forma e dimensioni: la nuova sintesi. LVII Congresso UZI, San Benedetto del Tronto.
- Costa C. 2004. Lo studio della forma in ittiologia: Morfometria geometrica e analisi dei profili. Ph.D. thesis in evolutionary biology and Ecology. University Rome Tor Vergata. 177 p.
- Crampton JS (1995) Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia*, 28, 179-186.
- Ferson, S.; Rohlf, F. J.; Koehn, R. K., 1985. Measuring shape variation of two-dimensional outlines. *SYST. ZOOL.*, 34(1): 59-68.
- Kuhl, F. P.; Giardina, C. R., 1982. Elliptic Fourier features of a closed contour. *COMP. GRAPH. IMAGE PROCESS.*, 18: 236-258.
- Menesatti P. 2000 Misura strumentale della qualità in ortofrutticoltura ITALUS HORTUS vol. 7. n. 2, pp. 12-19.
- Menesatti P. 1998. Integrazione di sistemi VIA (video image analysis) attraverso Internet: valutazione comparativa dello spettro di distribuzione di fitofarmaci. *INFORMATORE FITOPATOLOGICO* n. 9/98. pp27-30
- Paoletti F., Menesatti P., Quaglia G.B., Di Giacomo G., Aloisi L., Terreri M. 1998 Crumb grain and texture evaluation of bread obtained with *Kluiveromyces lactis*

addition. 13th International Congress on Agricultural Engineering (CIGR)
Rabat (Morocco). 2-6/2/98. vol. 6. pp. 197-202.

Rohlf, F. J., 1992. The analysis of shape variation using ordinations of fitted functions. In: Sorensen, J. T.; Foottit, R. Ordination in the study of morphology, evolution and systematics of insects: applications and quantitative genetic rationals., Ed. Elsevier.: 95-110.

Rohlf, F. J.; Archie, J. W., 1984. A comparison of Fourier methods for description of wing shape in mosquitoes (Diptera: Culicidae). SYST. ZOOL., 33(3): 302-317.

Rohlf, F. J.; Marcus, L. F., 1993. A revolution in morphometrics. TREE, 8(4): 129-132.

3.3 Pubblicazioni presentate

Nella presente sezione sono presentati cinque lavori scientifici, di cui il dottorando è coautore, inerenti review o applicazioni sperimentali dell'analisi della forma per immagini in differenti contesti agroalimentari.

Si riportano di seguito gli abstract delle singole pubblicazioni:

3.3.1 ABSTRACT

Shape analysis of agricultural products by computer vision – a review of recent research advances. Costa C, Antonucci F, Pallottino F, Aguzzi J, Sun DW, Menesatti P, ACCEPTED TO FOOD AND BIOPROCESS TECHNOLOGY

Appearance of agricultural products deeply condition their marketing. Appearance is normally evaluated by considering size, shape, form, colour, freshness condition, and finally the absence of visual defects. Among these features, the shape plays a crucial role. Description of agricultural product shape is often necessary in research fields for a range of different purposes, including the investigation of shape traits heritability for cultivar descriptions, plant variety or cultivar patents and evaluation of consumer decision performance. This review reports the main applications of shape analysis on agricultural products such as relationships between shape and i. genetic, ii. conformity/condition, iii. product characterization, iv. product sorting and v. clone selection. Shape can be a protagonist of evaluation criteria only if an appreciable level of image shape processing and automation and data are treated with solid multivariate statistic. In this context image-processing algorithms have been increasingly developed in the last decade in order to objectively measure the external features of agricultural products. Grading and sorting of agricultural products using machine vision in conjunction with pattern recognition techniques offers many advantages over the conventional optical or mechanical sorting devices. With this aims we propose a new automated shape processing system (ASPS) which could be useful for both scientific and industrial purposes, forming the bases of a common language for the scientific community. We applied such a processing scheme to

morphologically discriminate nuts fruit of different species. Operative Matlab codes for shape analysis are reported.

Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars, 2008, Menesatti P, Costa C, Paglia G, Pallottino F, D'Andrea S, Rimatori V, Aguzzi J, BIOSYSTEM ENGINEERING, 101(4): 417-424.

Cultivar discrimination during on-line quality selection is required by high quality food industries. The aim of this work was to evaluate the potential use and efficacy of shape-based techniques in order to discriminate among four traditional Italian cultivars (Tonda di Giffoni, San Giovanni, Mortarella and Tonda Romana). Tonda di Giffoni and Tonda Romana are very similar having a spherical shape, while the other two cultivars are elongated. RGB images of about 400 hazel-nuts were analysed with a morphological method based on the elliptic Fourier approximation to closed contours in a two-dimensional plane. This method was applied on the three outlines obtained by the polar, lateral and random plane positioning view of in-shell and unblanched kernel. The coefficients of the harmonic equations were analysed via PLSDA multivariate classification and mean outline for each group was graphically extracted. Results show higher percentage of correct classification for the lateral view (from 77.5% to 98.8% in the independent test). Also the random positioning view, in particular for in-shell kernels between the two rounded cultivars and between the two oblong cultivars, shown good classification results (respectively 95.1 and 97.6). This preliminary study demonstrates the potential of modern multivariate techniques using shape-based methods on digital images to achieve high efficiency performance in fruit grading and classification.

Quantitative method for shape description of almond cultivars (*Prunus amygdalus* Batsch). Antonucci F, Costa C, Pallottino F, Paglia G, Rimatori V, De Giorgio D, Menesatti P, IN PRESS IN FOOD AND BIOPROCESS TECHNOLOGY

The aim of the present work was to propose a rapid, non-invasive and quantitative image analysis method based on Elliptic Fourier Analysis (EFA) and on carpological measurements, to discriminate between 18 cultivars and shape groups of almonds kernels and in-shell fruit. The shape groups were identified using two clustering techniques: a non-hierarchical method (k-means) and a hierarchical one (Ward's method). Both methods found the same numbers of groups for in-shell fruit and kernels. The obtained results indicate that such differences can be used to discriminate among shape groups. This method wasn't efficient in discriminating single cultivars. In order to classify fruit into shape groups a Partial Least Squares Discriminant Analysis was applied. This analysis applied on the 18 cultivars groups showed low percentages of correct classification for both, in-shell (38.58%) and kernels (31.36%). The same analysis computed on shape groups shows percentages of correct classification higher than 89%. Merging Elliptic Fourier Analysis, clustering methods and modelling techniques, set the base for the implementation of an automated online fruit sorting. A Matlab script was developed to determine the right number of clusters in k-means clustering.

Discrimination of Tarocco sweet orange [*Citrus sinensis* (L.) Osbeck] varieties using opto-electronic elliptic Fourier based analysis of fruit shape. 2009. Costa C, Menesatti P, Paglia G, Pallottino F, Aguzzi J, Rimatori V, Russo G, Recupero S, Reforgiato Recupero G, POSTHARVEST BIOLOGY AND TECHNOLOGY, 54: 38-47.

Blood orange cultivars of the sweet orange [*Citrus sinensis* (L.) Osbeck] differ from the common sweet orange group (*Valencia Late*, *Washington navel*, *Navelina*) by the presence in the flesh and sometimes in the rind, of red anthocyanin pigments. Among

blood orange varieties, Tarocco is the most variable due to its particular characteristics. The presence of several Tarocco varieties, often characterized by similar maturation periods, necessitates accurate postharvest fruit evaluation, particularly appearance, since this is a primary criterion of consumer preference. In this work a total of 929 fruit belonging to 17 different Tarocco genotypes were analyzed. Optoelectronic techniques were used to discriminate among fruit shapes using Elliptic Fourier Analysis (EFA) to analyse fruit lateral shapes. Fruit shape for different genotypes was classified according to the IPGRI e Citrus Industry classification. The efficiency of these methods was tested by reclassifying fruit shape typologies by k-means analysis. We also computed the best number of k (4) by implementing a suited script in MatLab. Results were screened by multivariate classification techniques (i.e., PSLDA) in order to evaluate the efficiency of the group classifications. The combined EFA and k-means analysis increased the efficiency of genotype classification based on fruit shape in comparison with reported descriptive methods. For example, comparing the two models with 5 groups (Citrus Industry and k-means-5), the percentage of correct classification in the independent test dataset was higher in the k-means-5 model (respectively, 46.6% vs 26.0% compared to a random probability of classification of 20%). EFA could measure single fruit shape allowing the comparison of their conformity within a standard of reference. The results set the basis for a shape description of different Tarocco varieties based on quantitative morphological statistics, a practice that, until now, has been carried out exclusively in a descriptive fashion. Consequently, our work represents the first discrimination of genetically different cultivars of the same species based on fruit shape.

Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels. 2009. Menesatti P, Antonucci F, Costa C, Santori A, Niciarelli I, Infantino A, 1st International Workshop on Computer Image Analysis in Agriculture, Potsdam, Germany 27 – 28 August 2009, Bornimer Agrartechnische Berichte - Heft 69, ISSN 0947-7314, Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V. (ATB)

Fusarium Head Blight (FHB) is a disease of complex aetiology affecting wheat and barley worldwide. The disease has a great impact on yield, but mostly on health, due to the ability of several *Fusarium* species involved in the diseases to produce mycotoxins dangerous to human and cattle. Grading of cereals for industry by means of visual estimation of disease incidence is not always so accurate in predicting mycotoxin levels on wheat samples. Rapid and accurate analyses of large wheat samples by means of non-destructive methods are needed. Among several techniques available, image-analyses, taking into account the entire external kernels shapes, have been developed. Preliminary experiments using durum wheat kernels artificially infected with *Fusarium graminearum* and *F. culmorum* have been performed in Italy. Samples were acquired with a high resolution scanner. After an automated thresholding procedure performed by an edge detection Sobel filtering, 90 points (x, y) equally angularly spaced (one point every 4°) from the centroid were digitized along the outline. Elliptic Fourier Analysis (EFA) was performed to extract shape data to be analyzed via Partial least squares discriminant analysis (PLSDA). Three classes of infection were considered: healthy, shrivelled and chalky. *F. culmorum* infection is better distinguishable than *F. graminearum*. Percentages of correct classification resulted to be 76.19% for *F. culmorum* infection and 59.38% for *F. graminearum*. In general the intermediate class (shrivelled) is always badly classified. Image analysis of *Fusarium* infected kernels showed promising results for future practical applications.

3.4 Shape analysis of agricultural products by computer vision – a review of recent research advances

FULL REFERENCE OF THE PUBLICATION

Costa C, Antonucci F, Pallottino F, Aguzzi J, Sun DW, Menesatti P, ACCEPTED. Shape analysis of agricultural products by computer vision – a review of recent research advances. ACCEPTED TO FOOD AND BIOPROCESS TECHNOLOGY

1. Introduction

The shape of agricultural products such as fruit, vegetables and grain is one of the most important factors for their classification and grading in relation to commercial quality and organoleptic properties (Morimoto et al., 2000). Moreover, the appearance of fresh agricultural products is a primary criterion in making purchasing decisions (Kays, 1991). In this context, the appearance of unities of products is evaluated by considering their size, shape, form, colour, freshness condition, and finally the absence of visual defects. All these characteristics contribute to the overall appearance, which is globally evaluated either in a metric or a subjective manner as an important quality indicator throughout the commercial-utilization chain, from the production, to the storage, the marketing, and finally down to the consumer (Kays, 1999). Among appearance features of agricultural products, the shape plays a central role. Shape is an important factor in distinguishing between different cultivars on a genetic basis, or among the same cultivar depending on farming conditions. For example, irregularities in shape are a critical factor in consumer decision. Less pronounced shape defects are not perceived, while on the contrary, more extreme variations may deeply influence purchasing decision, leading to the ultimate rejection of a product (Kays, 1999).

Description of agricultural product shape is often necessary in research fields for a range of different purposes, including investigating heritability shape traits (Tanaka et al., 1955; Peterson, 1959; Wall, 1967; Curtis, 1968; Currie et al., 2000; White et al.,

2000; Nunome et al., 2001; van der Knaap & Tanksley, 2003; Zygier et al., 2005; Brewer et al., 2007; Menesatti et al., 2008; Costa et al., 2009a; Antonucci et al., 2010), cultivar descriptions for plant variety or cultivar patents (Beyer et al., 2002), evaluation of consumer decision performance (Jahns et al., 2001), physical key parameters in packaging and shipping (Sadrunia et al., 2007; Pallottino et al., 2010b), or determining misshaped fruit in a cultivar, etc. Despite all these applications, to date shape is still the object of subjective classification in the majority of cases.

In general terms, shape is represented by the outline of an agricultural product, being independent from the effect of size at the final stage of growth when commercialization or cultivar selection is of relevance. Shape can be a protagonist of evaluation criteria only if an appreciable level of image processing and automation is achieved. Shape description, quantification and discrimination in relation to different conditions of cultivation, storage, and etc. can be very specific and precise, when image processing and automation are developed on the basis of precise mathematical relationships and corresponding solid statistic treatments. These can be found by coupling video-image analysis with multivariate statistic taken from the domain of geometric morphometry. This mixed approach represents a novel focus in automated object recognition and classification (Costa et al., 2008; Aguzzi et al., 2009a; 2009b) for animal forms that can be now implemented and applied for agricultural processing purposes.

Aim of this paper is to review the state of the art of shape analysis of agricultural products by means of computer vision. Moreover, we propose a new automated shape processing system (ASPS) which could be useful for both scientific and industrial purposes, forming the bases of a common language for the scientific community.

1.1 Shape and genetic

Agricultural product shapes are determined both on genetic and environmental basis (Sadrunia et al., 2007). Quantitative-trait loci have been identified as DNA regions controlling particular morphological differences among different cultivars of important agricultural species (e.g. Doganlar et al., 2002; Brewer et al., 2007). The similarity existing between fruit morphology and inheritance allows using shape differences as discriminating factors for taxonomical purposes and cultivar/stain

origin assessment (Cannon & Manos, 2001). Accordingly, the efficiency of any shape-based morphometric analysis will depend on the established level of correlation between shape descriptors and underlying genetic variation. In this sense, the combined use of digital image and statistic analyses could be a successful tool for identifying certain shape traits as a result of determined genetic conformations (Currie et al., 2000). The quantitative study of the variation in these traits may disclose the contribution of the genetic *versus* the environmental control on the way of appearing of different strains and cultivars.

Social and economic implications of studies linking inheritance to agricultural products morphology are abundant. The process of domestication often occurs on fruit appearance, where shape is a central criterion. This implies that parents are chosen on the base of phenotype features that maximise desirable commercial characters, without reference to underlying genetic information (White et al., 2000). The domestication of wild varieties into actual cultured forms can result from the mutation and subsequent selection of single loci as in the case of pepper (*Capsicum* spp.; Peterson, 1958; Zygier et al., 2005), seed plant (*Cucumis melo*; Wall, 1967), watermelon (*Citrullus lanatus*; Tanaka et al., 1955), rice (*Oryza sativa*; Zheng et al., 2007), and different tomato strains (Alpert et al. 1995; Ku et al., 1999; Van der Knaap & Tanksley, 2003; Xiao et al., 2008).

Shape differences result from differences in genes which control the relative rates of growth between the polar and equatorial dimensions (reviewed by Peterson, 1958). Therefore, different shapes could be typically attributed to homozygous recessive/dominant as well as heterozygous allelic combinations (reviewed by Tanaka et al., 1955). In some cases, the establishment of linkage between genetic and morphology is more complicated. For example, pears (*Pyrus* spp.) fruit shape is under polygenic control (White et al., 2000). The identification of molecular markers responsible for morphological variation may assist the selection of traits for breeding programs in relation to market needs (Nunome et al., 2001). The introduction of new morphological characteristics in commercially important cultivars by genetic manipulation in order to increase their market value is presently a reality.

1.2 Shape and conformity/condition

Shape variation is an inherent factor in the production of agricultural products. Due to its characteristics, some portion of the total of each commodity to be harvested will deviate from what is considered optimum for one or more quality components.

Products in this defective category display quality defects, and are commercially undesirable, which prevents them to be considered as an optimum in terms of quality. While we consider defects as distinctly atypical and externally imposed alterations, such as insect or hail damage, substandard product in terms of appearance categories (size, shape, form, colour and condition) is also defective.

In this context, the conformity of agricultural products (i.e. in terms of morphological-organoleptic homogeneity) is represented by the sum of several biological parameters that must satisfy a quality standard criterion in order to be considered acceptable by consumer. The conformity sets a level of restriction, since it implies a certain number and type of valuable parameters, which depend upon the commercial context and its legislation. Once the international law is satisfied, each state, region or large organized distribution chain can apply a more restrictive version of it. Indeed, the chains of trading enterprises that count on several average or big centres deal with the large –scale distribution and need to provide products with constant characteristics during time, namely, with a defined conformity in relation to a reference standard. The absence of visual defects represents an important factor normally used as an attribute of conformity.

Condition is a central aspect of product conformity but its definition is still subjective. Kays (1999) proposes to use a reduced but variable set of different parameters that encompass a wide range of the products properties, a fact that complicate the definition of its condition.

1.3 Shape and product characterization

The process of domestication of plant populations undoubtedly constitutes a considerable selective factor in their evolution (Diamond, 2002). Varietal inheritance of numerous cultivated plants results from a long history of peoples and anthropogenic activities, and it nowadays consists of thousands of varieties (so-called cultivars) even when their chronological and geographical origins are not yet fully

established. The understanding of the history and evolution of plants under domestication, the identification of varietal inheritance, and the reconstruction of the beginnings and the exploitation of crop plants must be carried out by conjoint biological, palaeobotanical and archaeological studies.

Nowadays, a great variety of foodstuffs can be labelled for origin determination, industrial processing specification or genetic characterization, according to the different methods of production. With the increasing number of the partners within the European Community (EC) as well as outside Europe itself, products of improved and impoverished quality are in competition together, often bearing the same name. This unfair competition has negative effects on marketing since a misleading consumers decision can occur, therefore discouraging producers. This is the reason why many EC countries created in 1992 a certification system to promote and protect agricultural food products in relation to certain standard of production known as Protected Designation of Origin (PDO), Protected Geographical Indication (PGI) and finally, Traditional Speciality Guaranteed (TSG).

Presently, the certification system is missing of shape based criteria. Cultivar certification can be based on a general and qualitative shape description (Paulus & Schrevens, 1999). Shape analysis may represent a valid tool for foodstuffs origin certification, as required to protect the interest of producers and to identify fraudulent products (Costa et al., 2010). In this context image-processing algorithms have been increasingly developed in the last decade in order to objectively measure the external features of horticultural products (Menesatti et al., 2008; Costa et al., 2009a; Antonucci et al., 2010).

1.4 Shape, product sorting and clone selection

Shape is one of the most important attribute that the consumer evaluates when buying (Scott, 1998). According to consumer's interests in relation to shape-based choice, several laboratories nowadays carry out panel tests. Panel testing procedure is performed by a group of persons trained for the evaluation of products and guided by a leader who chooses the most appropriate methodology of evaluation and typology of data elaboration. Sensory analysis is often implemented and used within panel test trials (Lawless & Heymann, 1998). Such a technique is of great help in the market

analysis, since it considers many appreciable parameters including shape. Shape analysis can be used in order to create provisional models (Antonucci et al., 2007; Windhager et al., 2008). These models can be used to predict consumer's choices and lead the market toward specific direction (e.g. cultivars clone selection).

Shape evaluation is a crucial element for industrial processing when sorting online fruit or vegetable stock, as computerised shape analysis represents a fast and cheap solution to select products among the whole stock according to certain requirements (e.g. dividing this by following the merceological classes) once a suitable reference model is developed.

Machine vision systems are replacing the process of manual inspection of products in different industries. Inspection operations may include defect detection, dimensional measurement and product spatial orientation, as well as grading, sorting, and finally counting. Machine vision has several advantages over the conventional methods of inspection. It can be tuned and then adjusted in order to work with other on-line processing tasks acting over 24 hours. This procedure can be programmed in order to take dimensional measurements more accurately and consistently than a human being and finally it can give an objective measure of colour and morphology of the object which an inspector could only assess subjectively (Batchelor et al., 1985). As there is no physical contact involved, this method is hygienic and the possibility of damage to the fragile biological products is consistently reduced at inspection.

It is difficult for a computer algorithm to identify and classify size and colour of the biological entities due to the natural variation in shape. Grading and sorting of agricultural products using machine vision in conjunction with pattern recognition techniques, including neural networks, offers many advantages over the conventional optical or mechanical sorting devices (Menesatti et al., 2008). Multiple sensors can be used to gather the necessary information from the kernels and send suitable signals to a computer where they can be decoded for multi-category classification (Pallottino et al., 2010a). Image-processing algorithms can be used to extract higher level information from the input signals for improved classification performance. The classification parameters can be easily modified to take into account annual variations in the product. When neural networks are used as classifiers, the sorting device can be

equipped with a training option through which the machine can be trained for recognizing new grades or for different products.

Concerning the activity of the genetic research the importance of shape evaluation in the selection of cultivars is well-known, as discussed previously. The analysis of the shape through the acquisition of RGB images allows the detection of minimum strain differences not always perceivable by human eyes, thus being of help in speeding up genotypes improvements. Ohsawa et al. (1998) defines kernel shape as one of the important descriptors for the evaluation of buckwheat genetic resources (IPGRI, 1994). Moreover, kernel shape has been considered as closely correlated with agroecotypes, which is one of the important agronomic traits in buckwheat breeding (Matano & Ujihara, 1973; Namai, 1980; Uehara & Namai, 1994; Hirose et al., 1995). Kernel shape has been mainly evaluated to date by visual inspection and categorical classifications, the latter being based on rough estimates of the form such as triangle type, ovate type and conoidal type in most cases (IPGRI, 1994; Hirose et al., 1995; Namai, 1980; Uehara & Namai, 1994). Aspect ratio and the area of two-dimensional projection of kernels are a few examples of quantitative measures of kernel shape (Matano & Ujihara, 1973; NIAR, 1992). However, it has been difficult to carry out statistical analysis on the relationships among kernel shape and other agronomic traits, such as plant height, days to flowering, seed weight, and volume weight, because of the lack of objective and quantitative point indices. Therefore, the development of a new quantitative evaluation method is highly advisable. The same consideration can be applied to other cases. Regarding fruit selection in new apple or Tarocco sweet orange varieties, Paulus and Schrevens (1999) and Costa et al. (2009a) respectively, underlined the importance to evaluate the phenotypic appearance through the characterized of cultivar shape differences.

2. Methods for shape analysis

To date, various types of evaluation techniques for the agricultural product shape determination have been studied. However, the potentials for quantitative evaluation of the agricultural product shape have not been fully exploited because it is characterized by complexity and uncertainty (Morimoto et al., 2000). Measurements

of fruit shape are generally subjective if based on comparison with silhouettes of standard cultivars (i.e. in descriptive methods). Some descriptors incorporate ratios calculated from measurements of length and width. Various authors have measured several parameters to give a more accurate comparison among cultivar shapes (White et al., 2000). Other more Accurate morphometric descriptors are based on the analysis of profiles by Elliptic Fourier and Bazier analyses, as well as on the definition of landmarks by geometric morphometry (Crampton, 1995; Loy et al., 2000; Zelditch et al., 2004).

Conceptually shape descriptors and the classification techniques are different. The present review is centered on shape descriptors and classification methods. The shape descriptors can be processed with different statistical approaches for inferential, ordering, modelling and classification purposes. Many different techniques, univariate or multivariate are presented in literature. The following sub-chapters show the main shape-descriptors associated with their principal analytical methods. An additional sub-chapter introduces and explains in details the multivariate techniques owing higher degree of complexity and innovation. Table 1 lists papers on shape analysis of agricultural products.

Table 1: Shape analysis and representation of agricultural products.

	Species/Crop	Shape Analysis	Representation	Ref.
Fruits	Apple	EFD in relation with genetics, PCA	Mean \pm Max/min for each PC axis	Currie et al., 2000
	Apple	Ratio	None	Noè & Eccher, 1996
	Apple	Fourier expansion and PCA	Mean for each cluster individual	Paulus & Schrevens, 1999

samples on the

PCA

<i>Citrus</i>	Spherical descriptors	harmonic	Mean \pm SD for each PC axis	Ding et al., 2000
Tarocco sweet orange	EFA, PLSDA	clustering and	Mean \pm SD for each cv and cluster	Costa et al., 2009a
Pear	Ratios in relation with genetics		None	White et al., 2000
Strawberries	Descriptive		Reference drawings	Khanizadeh, 1994
Sweet cherry	PCA and clustering on shape		Mean for each cluster	Beyer et al., 2002
<i>Cucumis</i>	Ratios and shape indices in relation with genetics		None	Wall, 1967
Watermelon	Descriptive and regression	and ratio,	Reference drawings	Sadrnia et al., 2007
Watermelon	Descriptive indices in relation with genetics	and shape	None	Tanaka et al., 1955
Vegetables <i>Beta</i>	Descriptive in relation with genetics		None	Curtis, 1968
Cucumber	Ratios in relation with plant position		None	Marcelis, 1994

Pickling cucumber	Shape index	None	Wehner & Cramer, 1996
Eggplant	Descriptive in relation with genetics	Reference drawings	Nunome et al., 2001
<i>Capsicum</i>	Descriptive and shape indices in relation with genetics	None	Peterson, 1959
Pepper	Descriptive	None	Aloni et al., 1999
Pepper	Shape index in relation with genetics	None	Zygier et al., 2005
Squash	Descriptive	None	Nerson, 2005
Tomato	Descriptive and Shape index, PCA	Ratios, None	Brewer et al., 2006
Tomato	Descriptive and Shape index in relation with genetics	Ratios, Individual samples	Brewer et al., 2007
Tomato	Ratio	None	Jahns al., 2001
Tomato	Ratios in relation with genetics	None	Ku et al., 1999
Tomato	Profile data and ANN	None	Morimoto et al., 2000
Tomato	Ratios in relation with	None	van der Knaap

		genetics		& Tanksley, 2003
	Tomato	Descriptive in relation with genetics	Individual samples	Xiao et al., 2008
	White-flowered gourd	EFA and PCA	Mean \pm 2SD for each PC axis	Morimoto et al., 2005
Nuts	Almond	EFA, clustering and PLSDA	Mean \pm SD for each cv cluster	Antonucci et al., 2010
	Hazelnut	EFA and PLSDA	Mean for each LV	Menesatti et al., 2008
	Pistachio	FD, decision tree and ANN	None	Ghazanfari et al., 1997
Cereals	Rice	Ratios in relation with genetics	None	Rabiei et al., 2004
	Rice	Shape factors	None	Sakai et al., 1996
	Rice	Ratio	None	Webb, 1991
	Rice	Ratio	None	Yadav & Jindal, 2001
	Rice	Ratios in relation with genetics	None	Zheng et al., 2007

Other	Common buckwheat	EFA and PCA	Mean for each cv and mean \pm 1998 2 σ for each PC axis	Ohsawa et al.,
	Olive	Geometric morphometry (landmarks) and CVA	None	Terral et al., 2004
	<i>Thlaspi</i>	Descriptive in relation with genetics	None	Mummenhoff et al., 1997

2.1 Descriptive methods and visual comparisons

The most common measurements that are made on objects are those that describe shape. Shape features are physical dimensional measures that characterize the appearance of an object. Area, perimeter, major and minor axes lengths, and aspect ratio are some of the most commonly measured morphological features.

Morphological features are widely used in automated grading, sorting and detection of objects in the industry.

Nowadays different procedures have been used for characterising fruit shape. Simple techniques include ratings based on visual comparison of shape with respect to reference drawings (Beyer et al., 2002). These drawing serve as reference in classifying cultivars and normalised fruit shapes. For example, for sour and sweet cherry, drawings of fruit contours (in front view) of five fruit shape categories (i.e. kidney-shaped, flat-round, round, oblong, and cordate) have been published by the “Union Internationale pour la Protection des Obtentions Végétales” (UPOV, 1976) (Schmidt et al., 1985). Another example refers to watermelons (Sadriani et al., 2007), where the United States Department of Agriculture (USDA, 1997) grade standard provides as reference two shapes in three classifications (U.S. Fancy, U.S. No. 1, U.S. No. 2) based on visual comparison of fruit shape relative to reference drawing . Ratings based on visual comparison do not require any electronic equipment. However, the method is biased by observers’ subjective judgement. Also, rating

scores may be biased by confounding variables such as fruit size or colour. Therefore, this procedure runs very slowly and it is not sufficient for classification fruit in distribution terminals prior to marketing.

Referring to grain, the visual classification of cultivar types is a suitable procedure of classification (Webb, 1991). The various grain types are objectively classified according to length, width, length/width ratio, thickness, and grain weight. Length is a measure of the largest dimension of the rough-, brown-, milled-rice grain. Shape is determined by a ratio among three dimensions (i.e. length, width, and thickness).

Quantifying and describing fruit shape can be necessary in agricultural research for a variety of purposes. These include cultivar description in applications such as plant variety rights (PVR) or cultivar registers (Khanizadeh 1994; Beyer et al., 2002), evaluation of consumer preference (Kays, 1999), investigating heritability of fruit shape traits (Cannon & Manos, 2001), or analyzing shape abnormalities (Brewer et al., 2007). Furthermore, shape and size can be a great discriminator in order to select fruit for different commercial and industrial purposes. Similar shapes and sizes of fruit allow the industry to work contemporarily with several different cultivars in the transformation phase, being the instruments calibrated to process a homogenous stock of product (Antonucci et al., 2010).

Generally, a shape descriptor is important for different reasons, such as a method of selection in post-harvest activities. In this scenario, methods to identify shape groups are very functional. In Antonucci et al. (2010) shape clustering was conducted on 18 cultivars of almond (*Prunus amygdalus*, Batsch) on the basis of shape profile. Such groups were determined by means of hierarchical- (Ward's method) and non-hierarchical methods (k-means). Both methods found the same numbers of groups for in-shell fruit and kernels starting from the mean coefficients of all harmonic equations, as extracted by the Elliptic Fourier Analysis. Results indicated that such differences can be used to discriminate among shape groups but not single cultivars, as reported by De Giorgio et al. (1996). These results demonstrated that apart from a semantic description of cultivar shape, as shown by IPGRI (1999) for *citrus*, a new method to quantify differences in shape was found. In fact, in the study of Antonucci et al. (2010) an attempt to compare the classification developed by the International Board for Plant Genetic Resources (IBPGR) (Gülcan, 1985) and the in-shell fruit

mean outline per group of shape was conducted. Such shape discrimination was found to be complex due to the extremely general description given by IBPGR. In relation to this, the proposed method possesses the following advantages: 1) it can classify the shape of each single fruit within established groups; 2) it can describe the general fruit shape as well as the shape of particular regions such as the apex and the base; and finally, 3) it can quantitatively measure the shape variability among different cultivars.

Also in Costa et al. (2009a), a total of 929 fruit belonging to 17 different Tarocco (*Citrus siensis*) genotypes were analyzed by means of elliptic Fourier analysis with opto-electronic techniques, in order to discriminate among fruit shapes. Fruit shape for different genotypes was classified according to IPGRI (1999) and *Citrus industry* (Hodgson, 1967) schemes. The efficiency of these methods was tested by reclassifying fruit shape typologies by k-means analysis. Results were screened by multivariate classification techniques such as the Partial Least Squared Analysis (PLSDA) in order to evaluate the efficiency of the group classifications. Also in this case, the obtained results indicated that not only singular genotype descriptions are important for industrial activity but also the description of shape groups.

2.2 Ratios and shape indices

Measurement of fruit shape is generally subjective based on comparison with silhouettes of standard cultivars. Various authors have measured several parameters to give a more accurate comparison of shape (Heijden & Vossepoel, 1994), particularly at the extremities of the fruit (Kajiura & Suzuki, 1980; White & Bailey, 1995). Some descriptors incorporate ratios calculated from measurements of length and width (Tufts & Hansen, 1931; Thibault et al., 1983; White & Alspach, 1996). The use of ratios enables a comparison of fruit shape between fruit of differing sizes. Generally, the length/ circumference ratio consistently contributes to the determination of the fruit shape, which is an important quality factor. Best quality fruits have a high length/ circumference ratio (Marcelis, 1994).

For apple fruit shape categories results of ratio analysis such as oblate, globose, conical and oblong have been widely used to describe the cultivar registers or new cultivar announcements (Hedrick, 1938; Smith, 1971; Lane & McDonald, 1987;

Cripps, 1993; Morgan et al., 1993). In fact, fruit aspect-related ratios are the first quantitative shape trait applied by breeders of apple industry. The aspect ratio is quantitatively inherited in different apple strains (Spinks, 1936; Brown, 1960). However, intermediate shapes can not be easily identified and categories were not ranked in the past, which made shape selection of apples very difficult. To overcome some of the disadvantages of a discrete shape category, Smith (1971) sorted into groups the similar shape of fruit from different strains according to common traits of their aspect: flat, intermediate or tall. These authors also defined a conic shape trait with rectangular, truncate-conic, and a round aspect with convex or straight form. A similar research was done on peaches to investigate the alignment of clingstones and to identify surface defects (Currie et al., 2000).

Within the classification framework of similar shape of cultivars, Sarkar and Wolfe (1985) studied which algorithms are required to efficiently sort tomatoes by computer vision. Tomato was classified according to 10 shape categories such as rounded, high-rounded, ellipsoid, or pyriform. Additionally, the distal end of the fruit was categorized as indented, flat, or pointed, whereas the proximal end of the fruit was categorized as flat or indented (International Plant Genetic Resources Institute, 1996). While these classifications are useful to group tomato varieties and to describe the respective cultivars, the classification scheme cannot be utilized to conduct precise quantitative measurements in a reliable and systematic manner. In addition, the terminology referring to features of fruit shape is not enough detailed, hence tending to be too taxon-specific. While the taxon-specific terminology may not pose a problem for intraspecific comparison, the cross-species comparisons within element of the same genus may be hampered by the lack of agreement upon terms that should be used to designate common discriminating attributes (Brewer et al., 2006). In this scenario, the development of structured and controlled vocabularies arranged in ontologies, would provide great benefit to botanists and agronomists (Bruskiewich et al., 2002). The specific terminological vocabulary presented by Brewer et al. (2006, 2007) was hence proposed to consistently facilitate the use of discriminating shape features within and across taxa.

Research focusing more on shape characterization was done on bell peppers (*Capsicum annuum*) (Paulus & Schrevens, 1999). Nevertheless, most of the shape

algorithms are used to quantify the roundness, the rectangularity, the triangularity or the elongation of the product, by calculating ratios of the projected area to width of the product.

Referring to rice, the milling quality has become increasingly important because it is the final part of grain yield, making it fit for eating and so directly related to the net income of farmers. For rough rice, milling quality is defined as brown rice rate (i.e. the percentage of brown rice), milled rice rate (i.e. the percentage of milled rice), and head rice rate (i.e. the percentage of head rice), as estimated by several morphological traits such as grain shape and size, physical and chemical properties of rice kernel, etc. (Zheng et al. 2006). Anyway, the genetic components underlying the ratio relationships are to date unclear. Grain shape includes morphological traits under direct selection that are highly correlated with milling quality. Typically, increased grain length, the length-width or length-thickness ratios are negatively associated with grain milling quality, while increased grain width and thickness tend to result in increased milling quality (Xu et al., 2004; Wang et al., 2005). Rice milling quality is known to be as a complex trait determined by both genetic and environmental components. Results from classical genetic analyses indicate that rice milling quality correlates with grain shape traits (Shi & Zhu, 1997). In the study of Zheng et al. (2006) an effort was made to identify loci associated with quantitative traits at the base of milling quality and grain shape. The authors used a large set of introgression lines and molecular markers in order to give a better definition of the genetic basis of milling quality and its relationship with grain shape traits and to obtain useful information for the strains improvement.

2.3 Outline based methods

There are several approaches that can be used to deal with shape outline data. These methods involve the fitting of some type of curve to the object outline and then, the resulting coefficients are used as variables for statistical analysis (Rohlf & Bookstein, 1990). The most common approach is the fitting with polynomial functions or trigonometric series (e.g. the group of Fourier analyses), splines, etc.

Within Fourier analysis the most common method is the Elliptic Fourier Analysis (EFA) on the contour coordinates (Rohlf & Archie, 1984). This method decomposes a

curve into a set of harmonically related ellipses (Crampton, 1995; Lestrel, 1997; Loy et al., 2000; Jensen et al., 2002; Costa et al., 2010). Another method is represented by the Fourier descriptors (FD) analysis. Descriptors represent the boundary of a region of the object and these can be used to quantify the shape as a periodic function that can be expanded in a Fourier series (Goto et al., 2005). The obtained information is a spectrum given by the frequencies and amplitudes of the waves approximating the contour. In the case of fast Fourier transform (FFT) the object outline is represented by the fitting of an arbitrary set of trigonometric functions. The mathematical expression is hence dependent on the function to be approximated (Jayas et al., 2000). In general, FFT is a more efficient descriptive tool if shapes to be classified are very different, and this method has found some applications for online sorting and classification (Aguzzi et al., 2009b). The Bezier polynomial analysis (Bezier, 1970) is another method used to discriminate object outlines. Bezier parameters, also called “vertices”, are estimated from x,y coordinates according to the least-square method of Engels (Engels, 1986; Loy et al., 2000).

Other alternative methods to Fourier analyses are based on landmarks configurations such as the Procrustes analysis (i.e. the rotational fit of an object in relation to a reference one according to a set of landmarks on a spline grid) (Bookstein, 1996; Jensen et al., 2002; Terral et al., 2004). In a recent development, wavelets analysis has been used for outline morphological classification in substitution to Fourier analytical sets. Wavelets are mathematic functions for identifying spot of elevated variation in linearized object outlines. These functions can be obtained by the fitting of outlines with different frequency components, so each component can be studied with a resolution matched to its scale (Parisi-Baradad et al., 2005). The output signal indicates morphological high variability as sharp spikes within output frequency graphs (Jayas et al., 2000). The drawback of this technique is the need to have a common starting point (Capoccioni et al., 2009).

Examples of Fourier processing for morphological discrimination among cultivars are abundant (see Table 1). Ohsawa et al. (1998), Cannon and Manos (2001), and Morimoto et al. (2005) used this method to study the shape variation of common buckwheat (*Fagopyrum esculentum*) fruit, Bornean *Lithocarpus* (Fagaceae), and white-flowered gourd (*Lagenaria siceraria*). The coefficients obtained by harmonic

fitting onto fruit profiles were statistically treated with the principal component analysis (PCA).

Currie et al. (2000) used Fourier descriptors in characterizing digitized cross-sections of apple fruit contours. A PCA was applied on Fourier descriptors. Also Goto et al. (2005) used elliptic FD to study fruit shape variation in *Fraxinus mandshurica* var. *japonica* using PCA. Ghazanfari et al. (1997) employed the same FD-based method for machine vision-grading of pistachio nuts, by using a decisional tree and ANN. Ding et al. (2000) employed a CCD laser displacement sensor to obtain three-dimensional fruit contour data of a range of *Citrus* species and fruit shape was characterized using spherical harmonic descriptors. In order to characterize the shape of new apple cultivars, Paulus and Schrevens (1999) used a Fourier expansion coupled with PCA.

Beyer et al. (2002) analyzed sweet cherry (*Prunus avium*) fruit shape belonging to 40 cultivars by digitizing the front and side outlines (three-dimensions); Cartesian coordinates were subsequently normalised for differing fruit size. After a PCA they characterized cultivars into 5 standard fruit categories.

Recently, EFA was also used in association with other different and more complex multivariate techniques. EFA coefficients were extracted for classification and modelling by Partial Least Squared Analysis (PLSDA) and clustering on hazelnuts (Menesatti et al., 2008), Tarocco sweet orange fruits (Costa et al., 2009a), and almonds (Antonucci et al., 2010).

To our knowledge, the only article on agricultural products shape analysis in which a landmark-based approach was used is the one of Terral et al. (2004) who studied archaeological and modern stones of olives. They used only 2 landmarks and 18 semi-landmarks.

2.4 Multivariate analyses

Multivariate analyses are generally divided into two main categories: unsupervised and supervised (Menesatti et al., 2007). For unsupervised techniques grouping or clustering methods for multivariate elements (x-block) are based on functional relationships among the same elements (distances, variances). They do not need for an *a priori* knowledge of the class categories. Differently, in supervised techniques

the class attribution is given by a single or multiple variables (y-block). In this way, multivariate methods are forced to cluster into *a priori* established classes.

Unsupervised methods are mainly applied in an exploratory sense, when the aim is to analyze or visualize non-forced aggregating relationships (unsupervised) among elements (Forina, 2006).

Concerning supervised techniques it is possible to distinguish two main analytical approaches: modelling and classification. Supervised methods are derived from the observation and then the use of known classes, called the training set. The derived classification criteria can then be used to classify each new object within a test set. This can be applied for both classification and the computing of efficiency parameters. Classification analysis needs a decision rule, called the “classification criterion”, to distinguish objects into classes on the basis of selected quantitative features (Jayas et al., 2000). For modelling, it is instead possible to attribute objects into one or more classes, but also to none (i.e. in this case the object is an outlier). Modelling techniques calculate the “prediction probability” with a classification threshold for each modelled class (Aguzzi et al., 2009a). The modelling efficiency is indicated by statistical parameters such as “sensitivity” and “specificity”. Sensitivity represents the percentage of the objects of a category accepted by the modelled class. Specificity is the percentage of objects different from the modelled classes, as rejected by this classification criterion (Costa et al., 2009a). On the other hand, for the classification, a matrix of correct classification can be used (Forina, 2006).

The statistics used to investigate ratios and shape indices are normally descriptive and represented by simple regression (Li et al., 2004), ANOVA (Brewer et al., 2007), principal components analysis (PCA), and canonical discriminant analysis (CDA) (Brewer et al., 2007). Many other studies uses instead shape-based methods in association with PCA (Ohsawa et al., 1998; Paulus & Schrevens, 1999; Currie et al., 2000; Cannon & Manos, 2001; Beyer et al., 2002; Goto et al., 2005; Morimoto et al., 2005; Brewer et al., 2006). This is because different shapes exhibit a certain level of quantitative variation related to genotypic and environmental effects. While shape can be categorized in some way for species, sub-species, cultivar, merceologic classes, crops, and etc., the quantitative variation in fruit shape can be analysed by methods based on classification and modelling such as PLS-based (PLS, PLSDA; Sjöström et

al., 1986; Sabatier et al., 2003; Menesatti et al., 2008; Costa et al., 2009a; Antonucci et al., 2010), soft independent modelling of class analogy (SIMCA; Casale et al., 2007; Aguzzi et al., 2009) or clustering of the Fourier coefficients (Costa et al., 2009; Antonucci et al., 2010).

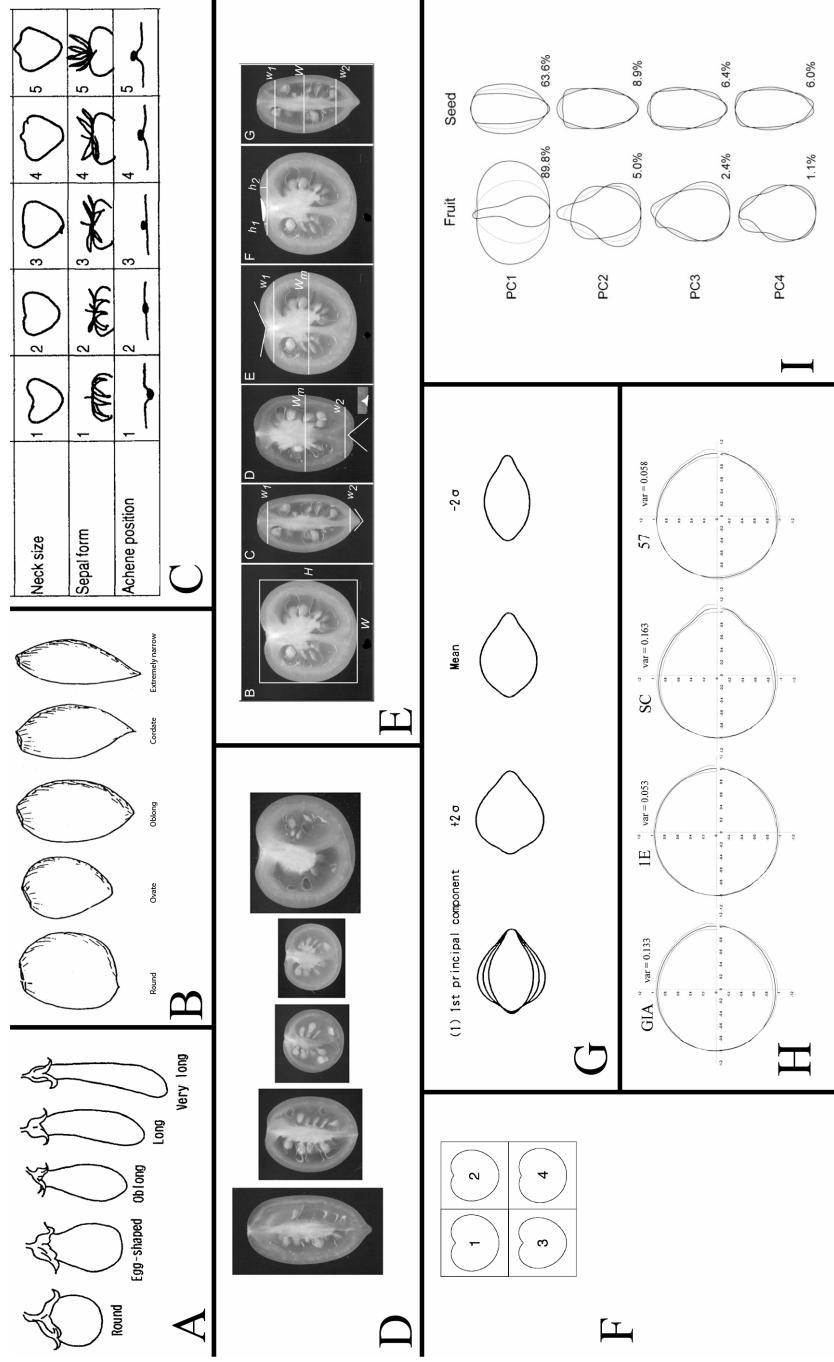
Also ANNs have been widely used for quantifying the variation in the shape of fruits (Ghazanfari et al., 1997; Morimoto et al., 2000). ANNs are very effective in many applications and are particularly useful as generalized non-linear regression tools (Masters, 1994; Costa et al., 2006, 2009b). They can perform arbitrary non-linear mappings in patterns of information.

2.5 Shape representation

In order to visually represent shapes it is crucial to find a way of graphically reporting these differences. Authors reported many different ways of representing shape variability.

The simplest way of representing shape differences is to use an individual reference shape model. This method was employed by Brewer et al. (2007) and Xiao et al. (2008) for genetic studies on tomato and by Cannon and Manos (2001) on *Lithocarpus*, Khanizadeh (1994) for strawberries, Nunome et al. (2001) for eggplants and by Sadrnia et al. (2007) for watermelons.

A more complex way of representing shape variation is represented by the use of configurations figures made by an average configuration (Ohsawa et al., 1998; Paulus & Schrevens, 1999; Beyer et al., 2002; Menesatti et al., 2008) under the form of a line with the maximum/minimum values (Currie et al., 2000), the \pm standard deviation (SD) (Ding et al., 2000; Costa et al., 2009a; Antonucci et al., 2010) or \pm 2SD (Ohsawa et al., 1998; Goto et al., 2005; Morimoto et al., 2005). Some examples of shape representation are given in Figure 1.



(Beyer et al., 2002); (G). mean $\pm 2\sigma$ (standard error) for the first PCA axis for apples (Ohsawa et al., 1998); (H). mean \pm standard deviation (SD) and variance for each cultivar of Tarocco sweet orange (Costa et al., 2009); (I). mean $\pm 2SD$ for each PCA axes of white-flowered gourd (Morimoto et al., 2005).

3. Future perspectives: an Automated Shape Processing System (ASPS)

The application of automated processing systems, within other research fields, is consistently reported in literature (Simigiana and Starkeya, 1986; Schneider et al., 1995; Rodriguez et al., 2006; Costa et al., 2009b). The aim of this chapter is to propose a flexible technological tool based on the shape analyses techniques which return better results, as reviewed above.

Given the morphometric and biometric analytic background for industrial cultivar or genetic strain processing and implementation, Menesatti and his co-workers (Menesatti et al., 2008; Costa et al., 2009a; Antonucci et al., 2010) recently proposed a new automated shape processing system (ASPS). The purpose of this proposed procedure is to provide a new common measuring and language procedure, giving at the same time an efficient informatic support for automated video-image analysis processing (i.e. see the Appendix no. 1 on Matlab programming codes) either for scientific and industrial purposes. The proposed procedure is fully automated and it could be implementable within the framework of Matlab codes (are provided as appendixes in the paper).

3.1 The procedure

Figure 2 summaries the steps in the ASPS procedure. The first step (Figure 2A) is represented by the digital-image acquisition of the samples at a high resolution through the use of a standard commercial digital camera. Samples, as single units or in groups, can be placed on an illuminated dashboard with a back background lighting which enhances the object contrast, a condition that facilitates the outline detection.

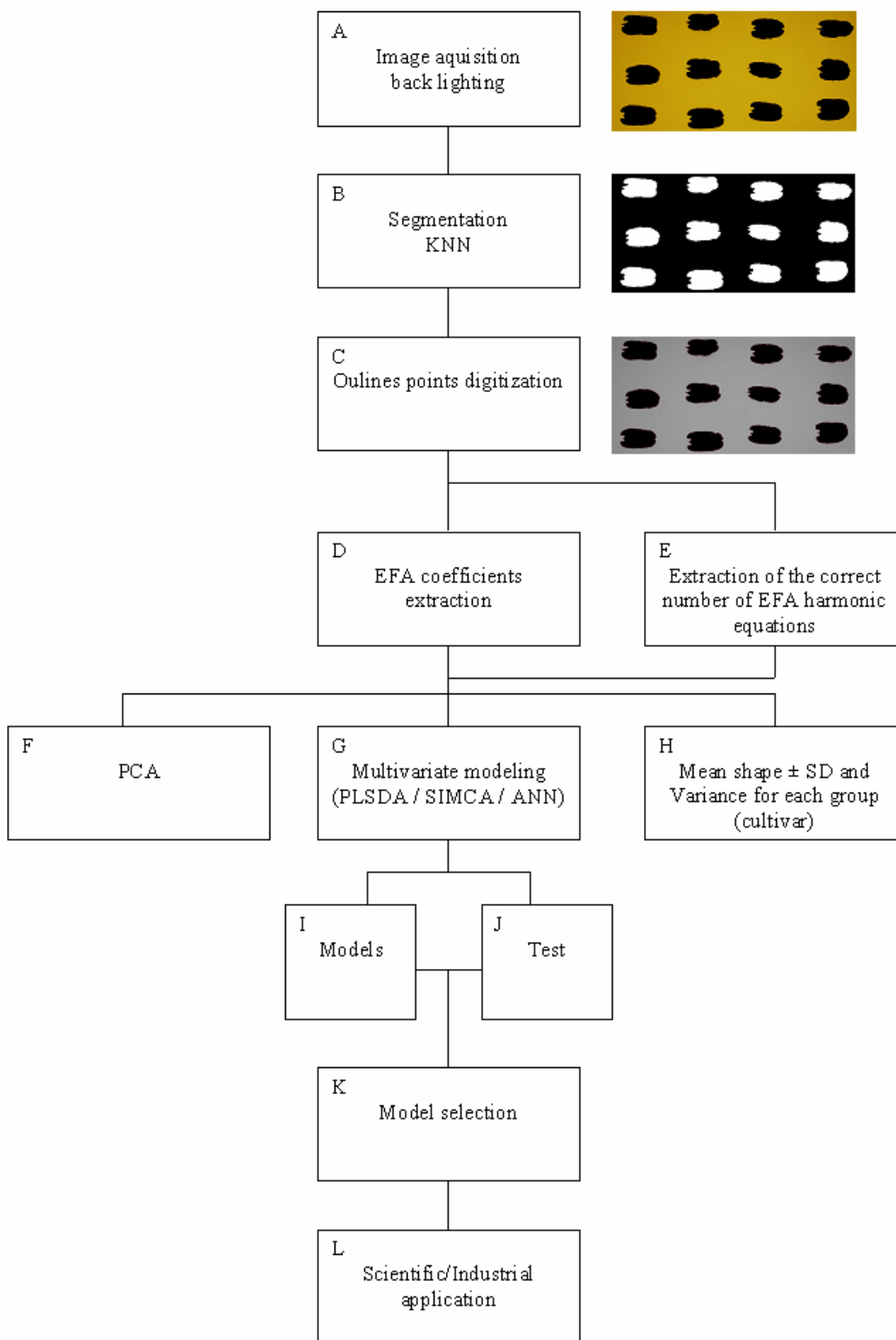


Figure 2: Diagram of the automated procedure for fruit shape extraction. On the left side of A, B, C examples of consecutive steps of image processing of the pecans Kiowa cultivar kernels are shown.

Then, each image can be segmented (Figure 2B) into black (i.e. background) and white (i.e. object) by applying the KNN procedure as formalized in Pallottino et al. (2010b). With this supervised multivariate clustering method, image binarization can be carried out since each pixel of the RGB input image is classified by a majority score given by its position/similarity in relation to its neighbours, being that pixel assigned to the most common class among its k-nearest ($k=3$). The neighbours are taken from a set of pixels for which the correct classification is known (i.e. the training set as directly created by an observer). The training set is built by calculating the mean value for each RGB channel on some extracted and representative patch of the original image.

After the KNN image binarization, a common number of points (x,y) equally angularly spaced from the centroid can be digitized along the object outline (Figure 2C) by using the Matlab procedure (Menesatti et al., 2008; Costa et al., 2009a; Antonucci et al., 2010), as reported in Appendix 1. Coordinates are aligned by Generalised Procrustes Analysis (GPA), a procedure that consists of three steps: the translation of points coordinates to a common centroid located at the origin $(0,0)$ of a reference system of coordinates; the scaling of each outline at the unitary centroid size; and finally, the rotation of coordinates to minimize the sum of square distances between correspondent landmarks (Bookstein, 1991). The 180 aligned coordinates can be then treated as outlines data (Menesatti et al., 2008; Costa et al., 2009a; Antonucci et al., 2010).

The overall shape can be at this point studied by EFA on the outline coordinates (Figure 2E). The harmonic coefficients describe the size, shape, and orientation of each harmonic ellipse and represent the input required by the multivariate statistics. The total number of harmonics that can be computed for any outline is equal to half of the total number of outline coordinates (i.e. the 'Nyquist frequency' rule). The Fourier series can be truncated at a number of coefficients (i.e. the k value) at which the average cumulative power of shape outline fitting is 99.999% of the total average power (Menesatti et al., 2008; Costa et al., 2009a; Antonucci et al., 2010). In order to do so, for any outline, the total power is calculated as the sum, from 1 to k , of individual harmonic powers, where k is equal to the Nyquist frequency (Crampton, 1995). The procedure to extract the correct number of harmonic equations is

implemented in Matlab and reported in Appendix 2. Once the correct number of harmonic equations to be used for shape extraction is obtained, their coefficients can be automatically obtained using a Matlab procedure (Figure 2D), as reported by Thomas (2006) and modified by Eom (2008).

EFA coefficients can then be treated for object morphological discrimination according to different types of multivariate analysis. These analyses are the PCA (Figure 2F) (Currie et al., 2000; Loy et al., 2000; Cannon & Manos, 2001; Beyer et al., 2002; Goto et al., 2005) or alternatively, different modelling technique (Figure 2G) such as PLSDA (Costa et al., 2008; Menesatti et al., 2008), SIMCA (Aguzzi et al., 2009a) or ANN (Ghazanfari et al., 1997; Morimoto et al., 2000; Costa et al., 2006; 2009b). These two latter can be used to graphically extract the mean outline and its standard deviation ranges (Figure 2H) as a graphical measure of the extent of morphological variation (Costa et al., 2009a; Antonucci et al., 2010).

3.2 A Case study with nuts fruits

As an example of the elaborated ASPS procedure, Menesatti and co-workers (Menesatti et al., 2008; Antonucci et al., 2010) focused on the morphological discrimination among different cultivars of walnuts, pecans, and pistachios which are commercially important commodity for the tree nut industry (INC, 2002).

In the proposed procedural example, a total number of in-shell fruits and kernels of different commercially important cultivars were used (Table 2). The selected group includes 180 walnuts from 3 commercially important cultivars; 72 pecans belonging to 2 different pecans cultivars; and finally, 120 pistachios from two different 2 cultivars.

Table 2: Species, cultivars number of in-shell fruit and kernels.

Species	Cultivars	N° in-shell fruit	N° kernels
Pecan	Kiowa	48	48
Pecan	Wichita	24	24
Pistachio	Bianca	60	60
Pistachio	Gloria	60	60
Walnut	B1.F2.P14	60	-
Walnut	B1.F7.P5	60	-
Walnut	Sorrento	60	-

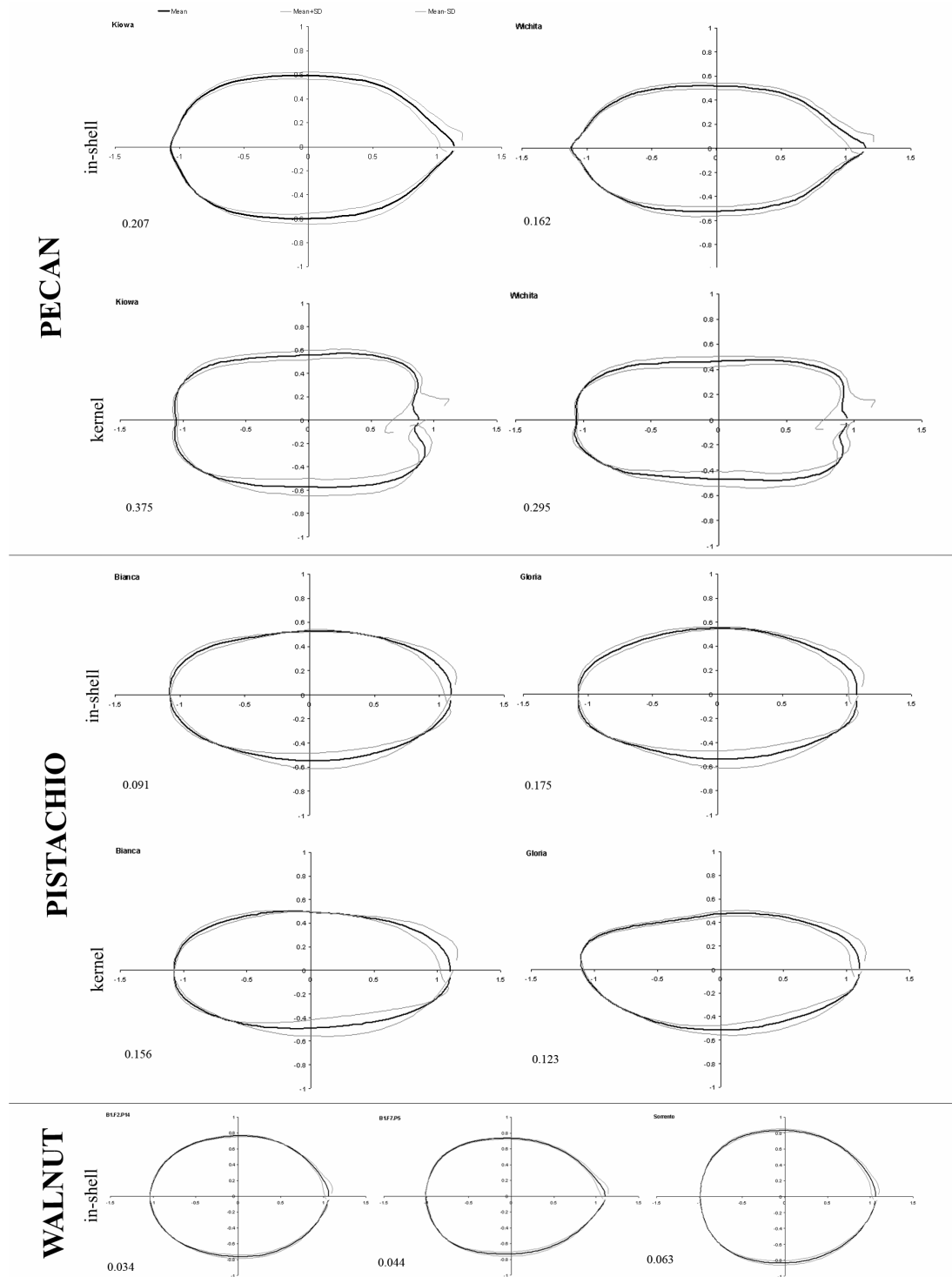
In these samples, walnuts were provided by the “Germoplasm Orchard” of the Fruitculture Research Unit of the Agriculture Research Council (CRA-FRC) of Caserta; pistachios and pecan were provided by the “Germoplasm Orchard” of the Fruitculture Centre of the Agriculture Research Council (CRA-FRU) of Rome, and All nuts were harvested between December 2006 and April 2007 during the optimal harvesting time for each species and cultivar.

According to the ASPS procedure previously described (see Figure 2), a group of fruits was placed on an illuminated dashboard to increase the contrast. Digital images were then acquired by a high resolution Nikon digital camera (model D1X) equipped with the Nikkor lens AF type (24mm f/2.8D). After the image acquisition, the KNN image binarization was conducted and a total number of 180 points (x,y) equally angularly spaced (one point every 2°) from the centroid were digitized. The overall shape was then studied by EFA on the outline coordinates, with resulting harmonic coefficients being treated with the PLSDA modelling technique. For each analysis, every sub-group corresponding to a different species strain was divided into two subsets: i) 75% of specimens for the class modelling and validation; ii) 25% of specimens for the independent test. Specimens were optimally attributed to each one of these two subsets based on the Euclidean distances. These distances were calculated according to the Kennard and Stone (1969) algorithm that selects objects without a priori knowledge of a regression model (i.e. the hypothesis is that the true model requires a uniform distribution of objects in the information space). PLSDA was performed using Matlab (rel. 7.1; PLSToolbox Eigenvectorb 4.0) on the shape

variables (X-block; EFA coefficients). The X-block values were pre-processed with the 'autoscale' Matlab procedure. The models were chosen so that the number of latent variables resulted in a higher percentage of correct classification in the relative independent test.

Results indicated that the correct number of harmonic equations was equal to 30 for pecan in-shell and 37 for pecan kernels; 19 for pistachio in-shell and 26 for pistachio kernels; and finally, 17 for walnut in-shell and, respectively. For in-shell and kernels of each species, the mean outline the standard deviation range are shown in Figure 3 along with the correct number of harmonic equations as well as the values of variance.

Figure 3: Mean outline (black) and the standard deviation range (grey) for in-shell and kernels of each species with the correct number of harmonics as obtained by EFA. The values of variance for each variety are shown on the bottom-left corner.



PLSDA was conducted on EFA coefficients to discriminate different cultivars of the same in-shell/kernel/species. Table 3 reports the characteristics of the models obtained. All these models have high percentages of both specificity and sensitivity (i.e. ranging from 89.7% for pistachio in-shell to 100% for pecan kernels). Also the percentages of correct classification are always high (100% in the independent test for pistachio in-shell).

Table 3: Characteristics and principal results of the PLSDA models performed on EFA coefficients.

Note: N is the number of samples. n° units (Y-Block) is the number of units to be discriminated by the PLSDA. n° LV is the number of latent vectors for each model. Random Probability (%) is the probability of random assignment of an individual into a unit.

	Pecan in-shell	Pecan kernel	Pistachio in-shell	Pistachio kernels	Walnut in-shell
N	72	72	120	120	180
n° units (Y-block)	2	2	2	2	3
n° LV	4	6	4	4	12
% Cumulated Variance X-block	30.33	30.30	35.01	21.58	60.23
Mean Specificity (%)	97.2	100	89.7	97.8	97.43
Mean Sensitivity (%)	97.2	100	89.7	97.8	97.80
Random Probability (%)	50	50	50	50	33.3
Mean Class. Err. (%)	0.028	0	0.103	0.023	0.024
Mean RMSEC	0.5430	0.5112	0.5813	0.5443	0.4069
Mean % Corr. Class. Model	98.1	100	90.1	97.8	99.3
Mean % Corr. Class. Independent Test	93.8	83.3	100	93.1	95.6

4. Conclusions

The present review analyzed the state of the art of the agricultural product shape analysis. This subject is growing in importance due to many factors including consumers choices, industrial online processing, cultivar description and selection. The technological advances in computer vision allow the use of even more efficient sensors, mathematical and statistical tools.

The proposed automated shape processing system (ASPS) and the attached Matlab codes give an operative protocol of analysis which could be useful for both scientific and industrial purposes, forming the bases of a common language for the scientific community. The applications on nuts fruit showed the high efficiency and applicability of the proposed ASPS protocol.

This review showed how shape appeared to be a measurable parameter of great importance as aid to geneticists as well as for industrial purposes (acceptability following different specific utilizations, fraud prevention).

References

- Aguzzi J, Costa C, Antonucci F, Company JB, Menesatti P & Sardá F (2009a) Influence of diel behaviour in the morphology of decapod natantia. *Biological Journal of the Linnaean Society*, 96, 517-532.
- Aguzzi J, Costa C, Fujiwara Y, Iwase R, Ramirez-Llorda E & Menesatti P (2009b) A novel morphometry-based protocol of automated video-image analysis for species recognition and activity rhythms monitoring in deep-sea fauna. *Sensors*, 9, 8438-8455.
- Aloni B, Pressman E & Karni L (1999) The effect of fruit load, defoliation and night temperature on the morphology of pepper flowers and on fruit shape. *Annals of Botany*, 83, 529-534.
- Alpert KB, Grandillo S & Tanksley SD (1995) fw2.2: a major QTL controlling fruit weight is common to both red- and green fruited tomato species. *Theoretical and Applied Genetics*, 91, 994-1000.

- Antonucci F, Costa C, Scardi M & Cataudella S (2007) Linking shape, taxonomy and function in teleost fish: a machine learning approach. *Instrumentation Viewpoint*, 6.
http://www.cdsarti.org/nova_web/ang/recerca/publi_butlletins.asp.
- Antonucci F, Costa C, Pallottino F, Paglia G, Rimatori V, De Giorgio D & Menesatti P (2010) Quantitative method for shape description of almond cultivars (*Prunus amygdalus* Batsch). *Food and Bioprocess Technology*, in press DOI: 10.1007/s11947-010-0389-2.
- Batchelor BC, Hill DA & Hodgson DC (1985) *Automated Visual Inspection*. IFS (Publications) Ltd, London, UK.
- Beyer M, Hahn R, Peschel S, Harz M & Knoche B (2002) Analyzing fruit shape in sweet cherry (*Prunus avium* L.). *Scientia Horticulturae*, 96, 139-150.
- Bezier PE (1970) *Emploi des machines a` comande nume´rique*. Ed. Masson, Paris, France.
- Bookstein FL (1991) *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge Univ. Press, New York. 435 pp.
- Bookstein FL (1996) Landmark method for forms without landmarks: morphometrics of group difference in outline shape. *Medical Image Analysis*, 1(3), 225-243.
- Brewer MT, Lang L, Fujimura K, Dujmovic N, Gray S & van der Knaap E (2006) Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other plant species. *Plant Physiology*, 141, 15-25.
- Brewer MT, Moyseenko JB, Monforte AJ & van der Knaap E (2007) Morphological variation in tomato: a comprehensive study of quantitative trait loci controlling fruit shape and development. *Journal of Experimental Botany*, 58(6), 1339-1349.
- Brown AG (1960) The inheritance of shape, size and season in progenies of the cultivated apple. *Euphytica*, 9, 327-337.
- Bruskiewich R, Coe EH, Jaiswal P, McCouch S, Polacco M, Stein L, Vincent L & Ware D (2002) The Plant Ontology™ Consortium and Plant Ontologies. *Comparative and Functional Genomics*, 3(2), 137-142.

- Cannon CH & Manos PS (2001) Combining and comparing morphometric shape descriptors with a molecular phylogeny: the case of fruit type evolution in *Bornean Lithocarpus* (Fagaceae). *Systematic Biology*, 50(6), 860-880.
- Capoccioni F, Costa C, Menesatti P, Lombarte A, Aguzzi J & Ciccotti E (2009) Otolith growth allometry measurements in the European eel. *Instrumentation Viewpoint*, 8, 71-72.
- Casale M, Armanino C, Casolino C & Forina M (2007) Combining information from headspace mass spectrometry and visible spectroscopy in the classification of the Ligurian olive oils. *Analytical Chimica Acta*, 589, 89-95.
- Costa C, Loy A, Cataudella S, Davis D & Scardi M (2006) Extracting fish size using dual underwater cameras. *Aquacultural Engineering*, 35(3), 218-227.
- Costa C, Aguzzi J, Menesatti P, Antonucci F, Rimatori V & Mattoccia M (2008) Shape analysis of different populations of clams in relation to their geographical structure. *Journal of Zoology*, 276, 71-80.
- Costa C, Menesatti P, Paglia G, Pallottino F, Aguzzi J, Rimatori V, Russo G, Recupero S & Reforgiato Recupero G (2009a) Quantitative evaluation of Tarocco sweet orange fruit shape using opto-electronic elliptic Fourier based analysis. *Postharvest Biology and Technonology*, 54, 38-47.
- Costa C, Scardi M, Vitalini V & Cataudella S (2009b) A dual camera system for counting and sizing Northern Bluefin Tuna (*Thunnus thynnus*; Linnaeus, 1758) stock, during transfer to aquaculture cages, with a semi automatic Artificial Neural Network tool. *Aquaculture*, 291(3-4), 161-167.
- Costa C, Menesatti P, Aguzzi J, D'Andrea S, Antonucci F, Rimatori V, Pallottino P & Mattoccia M, 2010. External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*. *Food and Bioprocess Technology*, 3(1), 43-48.
- Cripps JEL, Richards LA & Mairata AM (1993) 'Pink Lady' Apple. *HortScience*, 28(10), 1057.
- Currie AJ, Ganeshanandam S, Noiton DA, Garrick D, Shelbourne CJA & Oraguzie N (2000) Quantitative evaluation of apple (*Malus x domestica* Borkh.) fruit shape by principal component analysis and Fourier descriptors. *Euphytica*, 111, 219-227.

- Curtis GJ (1968) Observations of fruit shape and other characters in the species of the section patellares, genus Beta. *Euphytica*, 17, 485-491.
- De Giorgio D, Macchia M, Stelluti M & Lanza F (1996) Collection of 205 almond tree cultivars in “seedling rootstocks” field of Istituto Sperimentale Agronomico (BA), descriptive monograph. *Agricoltura Ricerca XVII*, ISMEA, 162, 55-161.
- Diamond J (2002) Evolution, consequences and future of plant and animal domestication. *Nature*, 418, 700-707.
- Ding W, Nesumi H, Takano Y & Ukai Y (2000) Quantitative evaluation of three-dimensional fruit shape and size of Citrus species based on spherical harmonic descriptors. *Euphytica*, 114, 103-115.
- Doganlar S, Frary A, Daunay MC, Lester RN & Tanksley SD (2002) Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. *Genetics* 161, 1713-1726.
- Engels H (1986) A least squares method for estimation of Bezier curves and surface and its applicability to multivariate analysis. *Mathematical Bioscience*, 79(2), 1-16.
- Eom S (2008)
<http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=12746&objectType=FILE>
- Forina M (2006) I metodi chemometrici di classificazione: a cosa servono e come sono utilizzati. Proceedings of the Nir Italia 2006, 25-26 May 2006 Ferrara.
- Ghazanfari A, Irudayaraj J, Kusalik A & Romaniuk M (1997) Machine Vision Grading of Pistachio Nuts Using Fourier Descriptors. *Journal of Agricultural Engineering Research*, 68, 247-252.
- Goto S, Iwata H, Shibano S, Ohya K, Suzuki A & Ogawa H (2005) Fruit shape variation in *Fraxinus mandshurica* var. *japonica* characterized using elliptic Fourier descriptors and the effect on flight duration. *Ecological Research*, 20, 733-738.
- Gülcan R (1985) Descriptors list for Almond (*Prunus amygdalus*) (Revised). International Board for Plant Genetic Resources (IBPGR), 30 pp.
- Hedrick UP (1938). *Cyclopedia of hardy fruits*, 2nd ed. The MacMillan Company, New York.

- Heijden GW, Van Der AM & Vossepoel AM (1994) On defining length and width for botanical object discrimination. *Aspects of visual form processing*, pp. 562-573.
- Hirose T, Yoshida M, Nemoto H, Kitabayashi H, Minami M, Matano T, Sherchand KK & Ujihara A (1995) Diversity of grain character of tartary buckwheat in Nepal. In: Matano T, Ujihara A (Eds.), *Current Advances in Buckwheat Research*, pp. 385-388. Shinshu Univ Press.
- Hodgson RW (1967) Horticultural varieties of citrus. In: Reuther W, Webber HJ, Bachelor LD (Eds), *The Citrus Industry*, vol. 1. University of California, Berkley, pp. 431-59.
- INC (2002) Official Response to Draft Report of Joint World Health Organization (WHO) Food and Agriculture Organization (FAO) Expert Consultation on Diet, Nutrition and the Prevention of Chronic Diseases. Available from: <http://www.nuthealth.org/nutrition/INC-ResponsetoWHO-FAOExpertPanelReport-June15.doc>
- IPGRI (1994) Descriptors for buckwheat (*Fagopyrum* spp.). International Plant Genetic Resources Institute, Rome, Italy, pp 48.
- IPGRI (1996) Descriptors for tomato (*Lycopersicon* spp.). IPGRI, Rome, Italy.
- IPGRI (1999) Descriptors for *Citrus*. International Plant Genetic Resources Institute, Rome, Italy. ISBN 92-9043-425-2 (URL: <<http://www.cgiar.org/ipgri/>>).
- Jahns G, Nielsen HM & Wolfgang P (2001) Measuring image analysis attributes and modelling fuzzy consumer aspects for tomato quality. *Computers and Electronics in Agriculture*, 31, 17-29.
- Jayas DS, Paliwal J & Visen NS (2000) *Journal of Agricultural Engineering Research*, 77(2), 119-128.
- Jensen RJ, Ciofani KM & Miramontes LC (2002) Lines, outlines, and landmarks: morphometric analyses of leaves of *Acer rubrum*, *Acer saccharinum* (Aceraceae) and their hybrid. *Taxon*, 51, 475-492.
- Kajiura I & Suzuki S (1980) Variations in fruit shapes of Japanese pear cultivars; geographic differentiations and changes by breeding. *Japanese Journal of Breeding*, 30(4), 309-328.

- Kays SJ (1991) Postharvest physiology of perishable plant products. Van Nostrand Reinhold, New York.
- Kays SJ (1999) Preharvest factors affecting appearance. *Postharvest Biology and Technology*, 15, 233-247.
- Kennard RW & Stone LA (1969) Computer aided design of experiments. *Technometrics*, 11, 137-148.
- Khanizadeh S (1994) Breeding strawberries for Eastern Central Canada. *Euphytica*, 77, 45-49.
- Ku HM, Doganlar S, Chen KY & Tanksley SD (1999) The genetic basis of pear-shaped tomato fruit. *Theoretical and Applied Genetics*, 9, 844-850.
- Lane WD & MacDonald RA (1987) 'Shamrock' apple. *HortScience*, 22(3), 515-516.
- Lawless H & Heymann H (1998) Sensory evaluation of food – principles and practices. Chapman & Hall, New York, USA.
- Lazzaroni, S. Gigli, D. Gabina (Eds.). Wageningen Academic Publishers ISSN 0071–2477. EAAP 123, 115-134.
- Lestrel PE (1997) Fourier descriptors and their applications in biology. Cambridge University Press, Cambridge (UK).
- Li J, Xiao J, Grandillo S, Jiang L, Wan Y, Deng Q, Yuan L & McCouch SR (2004) QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (*O. sativa* L.) and African (*O. glaberrima* S.) rice. *Genome*, 47, 697-704.
- Loy A, Busilacchi S, Costa C, Ferlin L & Cataudella S (2000) Comparing geometric morphometrics and outlines fitting methods to monitor fish shape variability of *Diplodus puntazzo* (Teleostea, Sparidae). *Aquacultural Engineering*, 21(4), 271-283.
- Marcelis LFM (1994) Fruit shape in cucumber as influenced by position within the plant, fruit load and temperature. *Scientia Horticulturae*, 56, 299-308.
- Masters T (1994) Signal and Image Processing with Neural Networks: a C++ Sourcebook. Wiley, New York.
- Matano T & Ujihara A (1973) Agroecotypes of *Fagopyrum* II. Classification of common buckwheat (*Fagopyrum esculentum* Moench.) by kernel shape and

- these geographical distribution in Japan (in Japanese). *Japanese Journal of Crop Science*, 42(Suppl 2), 29-30.
- Menesatti P, D'Andrea S & Costa C (2007) Spectral and thermal imaging for meat quality evaluation. In: Lazzaroni S, Gigli D, Gabina C (eds) *New developments in evaluation of carcass and meat quality in cattle and sheep*. Wageningen Academic Publishers ISSN 0071-2477. EAAP 123, 115-134.
- Menesatti P, Costa C, Paglia G, Pallottino F, D'Andrea S, Rimatori V & Aguzzi J (2008) Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars. *Biosystem Engineering*, 101(4), 417-424.
- Morgan J, Richards A & Dowle E (1993) *The Book of Apples*. Ebury Press, London.
- Morimoto T, Takeuchi T, Miyata H & Hashimoto Y (2000) Pattern recognition of fruit shape based on the concept of chaos and neural networks. *Computers and Electronics in Agriculture*, 26, 171-186.
- Morimoto Y, Maundu P, Fujimaki H & Morishima H (2005) Diversity of landraces of the white-flowered gourd (*Lagenaria siceraria*) and its wild relatives in Kenya: fruit and seed morphology. *Genetic Resources and Crop Evolution*, 52, 737-747.
- Mummenhoff K, Franzke A & Koch M (1997) Molecular data reveal convergence in fruit characters used in the classification of *Thlaspi s. l.* (Brassicaceae). *Botanical Journal of the Linnean Society*, 125, 183-199.
- Namai H (1980) Studies on the pollination ecology in crops. 5. Variation of gene frequency of buckwheat population under different conditions of seed multiplication culture (in Japanese). *Japanese Journal of Breeding*, 30(Suppl 2), 274-275.
- Nerson H (2005) Effects of fruit shape and plant density on seed yield and quality of squash. *Scientia Horticulturae*, 105, 293-304.
- NIAR (1992) *The manual for characterization and evaluation in plant genetic resources. Part 1. Rice, wheat, barley, legume, root and tuber crop, millet and other cereals, industrial crop*. National Institute of Agrobiological Resources, Ministry of Agriculture, Forestry and Fisheries, Japan, 72-77.
- Noè N & Eccher T (1996) 'Golden Delicious' apple fruit shape and russeting are affected by light conditions. *Scientia Horticulturae*, 65, 209-213.

- Nunome T, Ishiguro K, Yoshida T & Hirai M (2001) Mapping of fruit shape and color development traits in eggplant (*Solanum melongena* L.) based on RAPD and AFLP markers. *Breeding Science*, 51, 19-26.
- Ohsawa R, Tsutsumi T, Uehara H, Namai H & Ninomiya S (1998) Quantitative evaluation of common buckwheat (*Fagopyrum esculentum* Moench) kernel shape by elliptic Fourier descriptor. *Euphytica*, 101, 175-183.
- Pallottino F, Costa C, Menesatti P & Moresi M (2010a) Use of image analysis to assess the effective surface area of orange fruit submitted to uniaxial compression tests. Submitted to *Journal of Food Engineering*.
- Pallottino F, Menesatti C, Costa C, Paglia G, De Salvador FR & Lolletti D (2010b) Image analysis techniques for automated hazelnut peeling determination. *Food and Bioprocess Technology*, 3(1), 155-159.
- Parisi-Baradad V, Lombarte A, Garcia-Ladona E, Cabestany J, Piera J & Chic O (2005) Otolith shape contour analysis using affine transformation invariant wavelet transforms and curvature scale space representation. *Marine and Freshwater Research*, 56, 795-804.
- Paulus I & Schrevens E (1999) Shape Characterization of New Apple Cultivars by Fourier Expansion of Digitized Images. *Journal of Agricultural Engineering Research*, 72, 113-118.
- Peterson PA (1959) Linkage of fruit shape and color genes in *Capsicum*. *Genetics*, 44: 407-419.
- Wall JR (1967) Correlated inheritance of sex expression and fruit shape in *Cucumis*. *Euphytica*, 16, 199-208.
- Rabiei B, Valizadeh M, Ghareyazie B, Moghaddam M & Ali AJ (2004) Identification of QTLs for rice grain size and shape of Iranian cultivars using SSR markers. *Euphytica*, 137, 325-332.
- Rodriguez A, Ehlenberger DB, Hof PR & Wearne SL (2006) Rayburst sampling, an algorithm for automated three-dimensional shape analysis from laser scanning microscopy images. *Nature Protocols*, 1, 2152-2161.
- Rohlf FJ & Archie JW (1984). A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Systematic Zoology*, 3, 302-317.

- Rohlf FJ & Bookstein FL (1990) Proceedings of the Michigan Morphometrics Workshop. Special Publication No. 2 The University of Michigan Museum of Zoology, Ann Arbor, Michigan, USA.
- Sabatier R, Vivein M & Amenta P (2003) Two approaches for Discriminant Partial Least Square. In: Schader, M., Gaul, W., Vichi, M. (Eds.), *Between data science and applied data analysis*. Springer-Verlag, Berlin, Germany.
- Sadnia H, Rajabipour A, Jafari A, Javadi A & Mostofi Y (2007) Classification and Analysis of Fruit Shapes in Long Type Watermelon Using Image Processing. *International Journal Of Agriculture & Biology*, 9(1).
- Sakai N, Yonekawa S, Matsuzaki A & Morishima H (1996) Two-dimensional image analysis of the shape of rice and its application to separating varieties. *Journal of Food Engineering*, 21, 397-407.
- Sarkar N & Wolfe RR (1985) Feature extraction techniques for sorting tomatoes by computer vision. *Transactions of the ASAE*, 28(3), 970-974.
- Schmidt H, Vittup-Christensen J, Watkins R & Smith RA (1985) Cherry descriptor list. IBPGR. Rome, Italy.
- Schneider G, Artmann G & Henning G (1995) Automated image processing system for shape recognition of single red blood cells based on out-of-focus images. *Biotechnology*, 32(2), 237-238.
- Scott D (1998) Horticulture fashion food conference. *Orchardist NZ*, Feb, 37-50.
- Shi CH & Zhu J (1997) Analysis of genetic covariances between plant agronomic traits and milling quality traits of indica rice. *Journal of the Zhejiang Agriculture University*, 23, 331-337.
- Simigiana S & Starkey J (1986) Automated grain shape analysis. *Journal of Structural Geology*, 8(5), 589-592.
- Sjöström M, Wold S & Söderström B (1986) PLS Discrimination plots. In: Gelsema ES, Kanals LN (eds) *Pattern recognition in practice II*. Elsevier, Amsterdam, The Netherlands.
- Smith MWG (1971) *National Apple Register of the United Kingdom*. Ministry of Agriculture, Fisheries and Food, London.

- Spinks GT (1936) Apple breeding investigations. I. Results obtained from certain families of seedlings. Annual Report of the Long Ashton Research Station, 1935, 19-49.
- Tanaka T, Wimol S & Mizutani T (1955) Inheritance of fruit shape and seed size of watermelon. Journal of the Japanese Society of Horticultural Science, 64(3), 543-548.
- Terral JF, Alonso N, Buxò i Capdevila R, Chatti N, Fabre L, Fiorentino G, Marinval P, Pérez Jordà G, Pradat B, Rovira R & Alibert P (2004) Historical biogeography of olive domestication (*Olea europaea* L.) as revealed by geometrical morphometry applied to biological and archaeological material. Journal of Biogeography, 31, 63-77.
- Thibault B, Watkins R & Smith RA (1983) Descriptor list for pear. International Board for Plant Genetic Resources. IBPGR/82/57.
- Thomas D (2006) fEfourier: Matlab function for Elliptical Fourier shape descriptors. <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=12746&objectType=FILE>
- Tufts WP & Hansen CJ (1931) Variations in shape of Bartlett pears. Journal of the American Society of Horticulture 28, 627-633.
- Uehera H & Namai H (1994) Relations among morphological traits of seed and cotyledon and wering habit in common buckwheat (*Fagopyrum esculentum*) (in Japanese). Breeding Science, 44(Suppl 1), 150.
- United States Department of Agriculture (USDA) (1997) United States Standards for Grades of Watermelons. Washington, D.C.
- UPOV (1976) Guidelines for the conduct of test for distinctness, homogeneity and stability of the cherry. UPOV, TG/35/3.
- van der Knaap E & Tanksley SD (2003) The making of a bell pepper-shaped tomato fruit: identification of loci controlling fruit morphology in Yellow Stuffer tomato. Theoretical and Applied Genetics, 107, 139-147.
- Wall JR (1967) Correlated inheritance of sex expression and fruit shape in Cucumis. Euphytica, 16, 199-208.

- Wang DY, Zhang XF, Zhu ZW, Chen N, Min J, Yao Q, Yan JL & Liao XY (2005) Correlation analysis of rice grain quality characteristics. *Acta Agronomica Sinica*, 31, 1086-1091.
- Webb BD (1991) Rice Quality and Grades. In: Rice utilization Volume II (second edition) (Luh BS Ed.) AVI book, New York, USA. pp. 89-119
- Wehner TC & Cramer CS (1996) Gain for picklig cucumber yield and fruit shape using recurrent selection. *Crop Science*, 36(6), 1538-1544.
- White AG & Alspach PA (1996) Variation in fruit shape in three pear hybrid progenies. *New Zealand Journal of Crop and Horticultural Science* 24, 409-413.
- White AG & Bailey DG (1995) Digital Imaging: a useful technique for analysing fruit shape in pears. *Fruit Varieties Journal*, 49(4), 224-226.
- White AG, Alspach PA, Weskett RH & Brewer LR (2000) Heritability of fruit shape in pears. *Euphytica*, 112, 1-7.
- Windhager S, Slice D, Schaefer K, Oberzaucher E, Thorstensen T & Grammer K (2008) Face to Face The Perception of Automotive Designs. *Hum Nat* DOI 10.1007/s12110-008-9047-z.
- Xu ZJ, Chen WF, Ma DR, Lu YN, Zhou SQ & Liu LX (2004) Correlations between rice grain shapes and main qualitative characteristics. *Acta Agronomica Sinica*, 30, 894-900.
- Xiao H, Jiang N, Schaffner E, Stockinger EJ & Van der Knaap E (2008) A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science*, 319, 1527-1530.
- Yadav BK & Jindal VK (2001) Monitoring milling quality of rice by image analysis. *Computers and Electronics in Agriculture*, 33, 19-33.
- Zelditch ML, Swiderski DL, Sheets HD & Fink WL. (2004) *Geometric Morphometrics for Biologists: A primer*. San Diego: Elsevier Academic Press. 443 p.
- Zheng TQ, Xu JL, Li ZK, Zhai HQ & Wan JM (2007) Genomic regions associated with milling quality and grain shape identified in a set of random introgression lines of rice (*Oryza sativa* L.). *Plant Breeding*, 126, 158-163.
- Zygier S, Chaim AB, Efrati A, Kaluzky G, Borovsky Y & Paran I (2005) QTLs mapping for fruit size and shape in chromosomes 2 and 4 in pepper and a

comparison of the pepper QTL map with that of tomato. Theoretical and Applied Genetics, 111, 437-445.

Appendix 1

Matlab script for the extraction of n equally angularly spaced points (see Figure 2C).

% the input images should be at two levels, type: logical, uint8, double

% Image Processing Toolbox is requested

%%INPUT VARIABLES

DirIn='C:\BW\';

DirOut='C:\OUTLINE\';

ImageExtention='bmp';

nPoints=180; *%insert the requested number of equally angularly spaced n points along the profile*

%%END OF INPUT VARIABLES

outl2=[str2cell('file name'), str2cell('angle °'), str2cell('X coord'), str2cell('Y coord'), str2cell('X coord original'), str2cell('Y coord original')];

passo=360/nPoints;

L=2;

files=dir(strcat(DirIn,'*.', ImageExtention));

for ii=1:size(files, 1); *%open all the files inside the input directory*

 fname=files(ii).name

 bw=imread (strcat(DirIn,fname));

 bw2 = bwperim(bw);

 h=figure('Visible', 'Off');, imshow(bw2, 'InitialMagnification', 30);, title ('perimeter');

```

S = regionprops(bwlabel(bw),'Perimeter', 'Centroid', 'FilledImage', 'PixelList');

outline=[];
for cnt =1:size(S,1)

S2=S(cnt).FilledImage;           %single object extraction
S3= regionprops(bwlabel(S2),'Perimeter', 'Centroid', 'PixelList');

CentroideOriginale=S(cnt).Centroid;
xcOrig=CentroideOriginale(1,1);
ycOrig=CentroideOriginale(1,2);

Perim=S3.PixelList; % x, y
Centre=S3.Centroid; % x, y

%% TRIGONOMETRIC approach

xc=Centre(1,1); yc=Centre(1,2);
text(xcOrig,ycOrig,'!', 'color','red');
bw2=S2;
BW=zeros(size(bw2,1), size(bw2,2));

for a=0:passo:360-passo

if a<90
    C=xc*tand(a);
if C<=yc
    c=[1 xc xc 1];
    r=[yc-C yc yc-L yc-C-L];
    bww = roipoly(BW,c,r);
    res=bw2+bww;
    [yi xi]=find(res==2);

```

```

ind=find(xi==min(xi));
yi=yi(ind(1,1),1);
xi=xi(ind(1,1),1);

else
C=yc;
AA=xc-C*cotd(a);
c=[AA AA+L xc+L xc];
r=[1 1 yc yc];
bww = roipoly(BW,c,r);
res=bw2+bww;
[yi xi]=find(res==2);
ind=find(yi==min(yi));
yi=yi(ind(1,1),1);
xi=xi(ind(1,1),1);
end
end

if a==90
c=[xc xc+L xc+L xc];
r=[1 1 yc yc];
bww = roipoly(BW,c,r);
res=bw2+bww;
[yi xi]=find(res==2);
ind=find(yi==min(yi));
yi=yi(ind(1,1),1);
xi=xi(ind(1,1),1);
end

if a>90 & a<=180
C=yc*tand(a-90);
xxc=size(BW,2)-xc;

```

```

if C<=xxc
    c=[C+xc C+xc+L xc+L xc];
    r=[1 1 yc yc];
    bww = roipoly(BW,c,r);
    res=bw2+bww;
    [yi xi]=find(res==2);
    ind=find(yi==min(yi));
    yi=yi(ind(1,1),1);
    xi=xi(ind(1,1),1);

else
    C=xxc;
    AA=yc-C*cotd(a-90);
    c=[size(BW,2) size(BW,2) xc xc];
    r=[AA AA+L yc+L yc];
    bww = roipoly(BW,c,r);
    res=bw2+bww;
    [yi xi]=find(res==2);
    ind=find(xi==max(xi));
    yi=yi(ind(1,1),1);
    xi=xi(ind(1,1),1);

end
end

if a>180 & a<270
    C=xxc*tand(a-180);
    yyc=size(BW,1)-yc;

if C<=yyc
    c=[size(BW,2) size(BW,2) xc xc];
    r=[yc+C yc+C+L yc+L yc];

```

```

    bww = roipoly(BW,c,r);
    res=bw2+bww;
    [yi xi]=find(res==2);
    ind=find(xi==max(xi));
    yi=yi(ind(1,1),1);
    xi=xi(ind(1,1),1);
end

if C>yyc
    C=yyc;
    AA=xc+C*cotd(a-180);
    c=[xc xc+L AA+L AA];
    r=[yc yc size(BW,1) size(BW,1)];
    bww = roipoly(BW,c,r);
    res=bw2+bww;
    [yi xi]=find(res==2);
    ind=find(yi==max(yi));
    yi=yi(ind(1,1),1);
    xi=xi(ind(1,1),1);
end
end

if a==270
    c=[xc xc+L xc+L xc];
    r=[yc yc size(BW,1) size(BW,1)];
    bww = roipoly(BW,c,r);
    res=bw2+bww;
    [yi xi]=find(res==2);
    ind=find(yi==max(yi));
    yi=yi(ind(1,1),1);
    xi=xi(ind(1,1),1);
end

```

end

if a>270 & a<=360

 C=xc*tand(360-a);

if C<=yyc

 c=[1 1 xc xc];

 r=[yc+C yc+C+L yc+L yc];

 bww = roipoly(BW,c,r);

 res=bw2+bww;

 [yi xi]=find(res==2);

 ind=find(xi==min(xi));

 yi=yi(ind(1,1),1);

 xi=xi(ind(1,1),1);

end

if C>yyc

 C=yyc;

 AA=xc-C*cotd(360-a);

 c=[AA AA+L xc+L xc];

 r=[size(BW,1) size(BW,1) yc yc];

 bww = roipoly(BW,c,r);

 res=bw2+bww;

 [yi xi]=find(res==2);

 ind=find(yi==max(yi));

 yi=yi(ind(1,1),1);

 xi=xi(ind(1,1),1);

end

end %if

outline=[outline; a, xi, yi, xi+(xcOrig-xc), yi+(ycOrig-yc)];

```

end    %for 0:360°

end

%%graphical output
xf=outline(:,4); yf=outline(:,5);
line(xf,yf,'LineWidth',3, 'color','red');
text(xf,yf,'o', 'color','green');

%save the resulting image with the marked outline
saveas(h, strcat(DirOut,'Outline_', fname,'.jpg'),'jpg');

close all

outl1=[repmat(str2cell(fname),size(outline,1),1), num2cell(outline)];
outl2=[outl2; outl1];

end          %for ii

Outline=outl2;
ObjNum=(size(outl2,1)-1)/nPoints;
ImgList=unique(outl2(2:end,1));

%%SAVE the results

save (strcat(DirOut, 'Outline.mat'), 'Outline', 'nPoints', 'ObjNum', 'ImgList');

```

Appendix 2

Matlab script for the extraction of the correct number of harmonics EFA harmonics equations following the procedure proposed by Crampton (1995), Menesatti et al. (2008), Costa et al. (2009) and Antonucci et al. (in press) (see Figure 2E).

```
%The script loads the OUTLINE.mat file produced with the appendix 1 script
%INPUT VARIABLES
DirIn='C:\Outline\';
DirOut='C:\EFA\';
FileIn='Outline.mat';
ThresholdPcum=99.999;    %insert the threshold value for cumulated variance
%%END OF INPUT VARIABLES

load (strcat(DirIn, FileIn));

iNoOfHarmonicsAnalyse=nPoints/2;
bNormaliseSizeState=1;
bNormaliseOrientationState=1;

EFAcoef=[];

for o=2:nPoints:size(Outline,1)
    outline=cell2mat([Outline(o:o+nPoints-1,3), Outline(o:o+nPoints-1,4)]);

    %the function fEfourier2.m was built following Thomas (2006) fEfourier.m and
    %modified by Eom (2008)
    rFSDs = fEfourier2(outline, iNoOfHarmonicsAnalyse, bNormaliseSizeState,
        bNormaliseOrientationState);
    EFAcoef=[EFAcoef; rFSDs];
end          %for o=2

EFAcoef=EFAcoef(:, 2:end);    %delete the zeroth harmonic
```

```

for d=1:4:size(EFAcoef,1)
    EFAcoef(d,1)=1;
    EFAcoef(d+1,1)=0;
    EFAcoef(d+2,1)=0;
end

Q=EFAcoef.*EFAcoef;
SQ=[];

for d=1:4:size(Q,1)
    SQ2=sum(Q(d:d+3,:),1);
    SQ=[SQ; SQ2];
end

mSQ=mean(SQ,1);
SmSQ = sum(mSQ, 2);
P=mSQ./SmSQ*100;
Pcum=cumsum(P,2)';      %cumulative sum of P

iNoOfHarmonicsAnalyse=find(Pcum>=ThresholdPcum, 1, 'first')+1;
RecordFileNames=[];
EFAdb=[];

for o=2:nPoints:size(Outline,1)
    outline=cell2mat([Outline(o:o+nPoints-1,3), Outline(o:o+nPoints-1,4)]);
    RecordFileNames=[RecordFileNames; Outline(o,1)];
    rFSDs =fEfourier2(outline, iNoOfHarmonicsAnalyse, bNormaliseSizeState,
bNormaliseOrientationState);
    out=[];

for s=1:size(rFSDs,2)

```

```
        out=[out, rFSDs(:,s)'];  
end  
  
EFAdb=[EFAdb; out(:,8:end)];  
  
end          %for o=2  
  
NoHarmonic=iNoOfHarmonicsAnalyse-1;  
  
%%SAVE the result as matlab file  
save (strcat(DirOut, 'EFAdb.mat'), 'EFAdb', 'RecordFileNames', 'NoHarmonic');
```

3.5 Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars

FULL REFERENCE OF THE PUBLICATION

Menesatti P, Costa C, Paglia G, Pallottino F, D'Andrea S, Rimatori V, Aguzzi J, 2008. Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars. BIOSYSTEMS ENGINEERING, 101(4): 417-424.

Introduction

According to FAO statistics the main countries producing hazelnut (*Corylus avellana*, L. 1753) are Turkey, Italy, Spain, USA and Greece. Turkey covers respectively 70% and 82% of the world's production and export, with Italy following with nearly 20% of the production and the 15% in terms of export (FAO, 2000). The world hazelnut production shows fluctuations depending on climatic conditions from year to year. In Italy, the most important cultivars are the native "Tonda Gentile Romana" in Lazio region, "Tonda di Giffoni" "Mortarella", "San Giovanni", "Camponica", "Riccia di Talanico", "Tonda Bianca" and "Tonda Rossa" in Campania, "Tonda Gentile delle Langhe" in Piemonte and "Santa Maria di Gesù" in Sicily.

Shelled hazelnut accounted for 79% on the total amount of world's hazelnut production while in-shell hazelnuts accounted for 21%. The food industry applies a qualitative selection only on shelled nuts. Kernels for confectionery should be plump and free from shrivelled kernels, of mould and extraneous matter.

In Italy, the confectionery industry and high quality markets have increased their demand for entire shelled nuts of the same cultivar to emphasize the traditional product or in order to guarantee local production. Quality improvement plans have primarily pointed to genetic and cultural improvements of products grown in plantations which seek to produce more homogeneous fruits and are not scattered among other plantations. The improvement in quality was identified in the conception of the improvement plan scheme as an important target which would pave the way to improved competitiveness (SEC, 2002). For this reason cultivar discrimination during on-line quality selection is required by high quality food industries.

Actually, non-destructive online selection of nuts using image analysis and optoelectronic techniques mainly leads to damage identification or sorting by morphological characteristics, for example opening extent of pistachio shells (Ghazanfari et al., 1997; Pearson & Toyofuku, 2000). These systems are able to detect external fruit damages (especially insect damage and moulds) (Kim & Schatzki, 1998). Shape-based analysis using machine vision could help in grading and selecting morpho-types (Nagata & Cao, 2000). Various methods for quantitatively evaluating shapes have been suggested in biological and agronomical context. The most common is based on elliptic Fourier descriptors, which have been successfully applied to the evaluation of several plant organs such as leaves (White et al., 1998; McLellan, 1993; Iwata & Ukai, 2002; Jensen et al., 2002; Neto et al., 2006), leaflets (Furuta et al., 1995), kernels (Ohsawa et al., 1998), roots (Iwata et al., 1998), flowers (Uga et al., 2003; Yoshioka et al., 2004) and fruits (Goto et al., 2005). This method describes the entire shape mathematically by transforming coordinate information concerning the contours into Fourier coefficients (Rohlf & Archie, 1984). The aim of this work was to evaluate the potential use and efficacy of shape-based techniques in order to discriminate among 4 traditional Italian cultivars (Tonda di Giffoni, San Giovanni, Mortarella and Tonda Romana). Tonda di Giffoni and Tonda Romana are very similar and are classified by FAO (2000) as round shaped, while the other two cultivars are classified as oblong. Multivariate classification methods (Partial Least Square Discriminant Analysis, PLSDA) was applied on elliptic Fourier analysis (EFA) coefficients extracted from the digital images of hazelnuts either in-shell or as kernels. The main goal was to demonstrate an innovative combined approach based on optoelectronic techniques (image analysis, morphometry and multivariate statistical analysis) that could lead to future application, at industrial level, of rapid, effective and non-destructive cultivar selection.

Materials and methods

About 100 hazelnuts per cultivar (Tonda di Giffoni, San Giovanni, Mortarella and Tonda Romana) (Table 1) were randomly chosen from three producing plants.

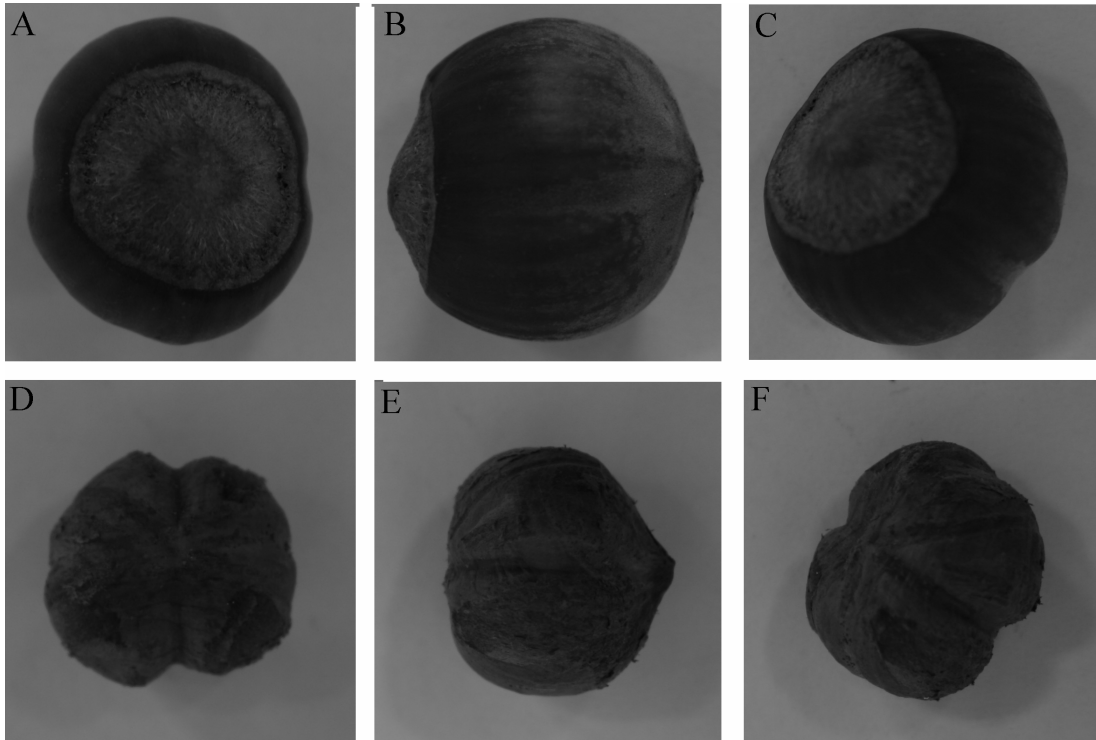
Table 1. Numbers and weights of the hazelnuts analyzed for each cultivar.

Cultivar	Label	Number of fruits analyzed	Weight in-shell (g) \pm SD	Weight kernel (g) \pm SD
Mortarella	MORTARELLA	104	1.93 \pm 0.35	0.84 \pm 0.21
San Giovanni	GIOVANNI	104	2.43 \pm 0.55	1.08 \pm 0.38
Tonda Giffoni	GIFFONI	102	2.58 \pm 0.45	1.22 \pm 0.31
Tonda Romana	ROMANA	104	2.85 \pm 0.54	1.28 \pm 0.39

All the fruit were provided by the cultivar collection of the Fruit Grown Research Institute of Caserta and harvested in the production season 2006. Being a methodological study, in order to avoid the influence of several factors, which can affect the size and shape of the kernels, samples came from a single growing location, field, crop year and from a selected research station that certified the genotype and the growing conditions – the present study would need to be extended to a more extensive range of material before it could be applied in practice. RGB images were analysed using a morphological method based on the elliptic Fourier approximation to closed contours in a two-dimensional plane.

Digital images of hazelnuts were acquired using a background illumination in order to maximize the object outline. Digital images, with high optical resolution, were acquired for each fruit in-shell (I) and unblanched kernel (K) from different viewpoints: polar view (PV), lateral view (LV) and the random plane positioning view (RV) (Fig.1).

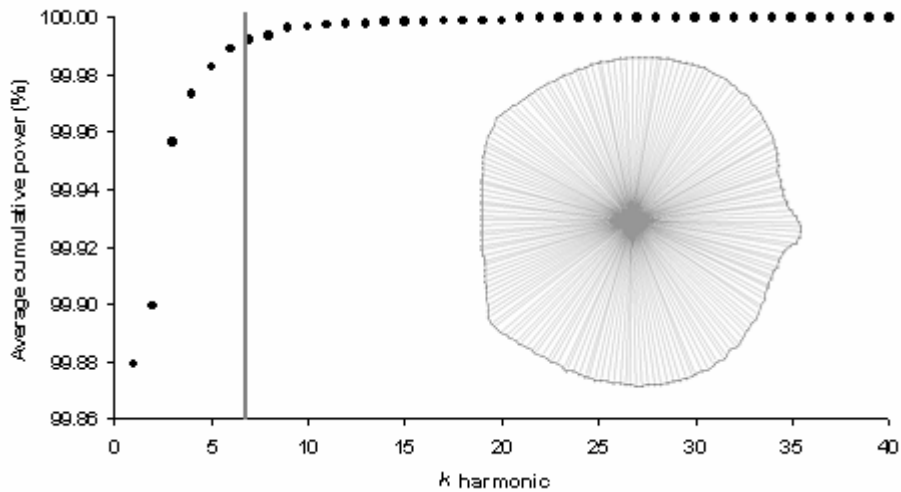
Figure 1. Examples of the three different view. A) polar view of the in-shell; B) lateral view of the in-shell; C) random plane positioning view of the in-shell; D) polar view of the kernel; E) lateral view of the kernel; F) random plane positioning view of the kernel.



Elliptic Fourier analysis

The overall shell shape was studied through elliptic Fourier analysis (EFA) of the contour coordinates (Rohlf & Archie, 1984). This method consists of decomposing a curve into a sum of harmonically related ellipses (e.g. Lestrel, 1997; Loy et al., 2000). The EFA was applied on the three series of outlines. The profiles were extracted with the software TPSdig2 (Rohlf, 2006) obtaining 180 equally angularly spaced points for each individual (Fig. 2).

Figure 2: Average cumulative power of the average total power at different k harmonics for the random plane view of the unblanched kernel (LV-I). The gray line represent the value of 99.99%. The outline, constructed on 180 equi-angular spaced coordinates (one every 2°), is also shown.



There are several approaches that can be used to deal with outline data ; they involve fitting some type of curve to the outline and then using the parameters of the curve for subsequent analysis (Rohlf, 1996). Of these, elliptic Fourier decomposition (Kuhl & Giardina, 1982) has been shown to be a powerful taxonomic descriptor (Rohlf & Archie, 1984; Ferson et al., 1985; Loy et al., 2000; Sheets et al., 2006) and have several advantages over other methods such as that it does not require explicit definition of a biologically homologous or mathematically-determined outline centroid. Given an outline approximated as a polygon described by a series of xy-coordinates, Fourier analysis is used to reduce the dimensionality of the data set and to eliminate the redundant information. This redundancy is determined by the high correlation between adjacent coordinates in the raw data set. Elliptic Fourier Analysis (EFA) is based on the separate Fourier decomposition of the incremental changes of the x- and y-coordinates as functions of the cumulative chordal length of the outline polygon. It yields the spectrum of the closed contour in terms of harmonically related trigonometric curves. For each harmonic equation, two Fourier coefficients are

computed for both the x- and y-projections, and thus the total number of coefficients is $4n$, where n is the number of harmonics fitted to the outline (Crampton, 1995). The correct number of harmonics was calculated using the method proposed by Crampton (1995). The Fourier series was truncated for both the view and the merceological status (I or K) at the value of k at which the average cumulative power was 99.99% of the average total power. As recommended by Rohlf & Archie (1984), Elliptic Fourier coefficients were mathematically normalized in order to avoid bias in results due to the different size, location, rotation and starting position of specimens. The coefficients of the harmonic equations were extracted with the software Morpheus (Slice, 1998) and analysed via multivariate classification. Mean outline for each group was graphically extracted using Morpheus.

Statistical classification and modelling

In order to build models discriminating between the cultivars based on different view and status, a multivariate classification analysis PLS-DA (Partial Least Square Discriminant Analysis) was applied.

PLS-DA (Sjöström et al., 1986; Sabatier et al., 2003) consists of a classical partial least squares analysis (PLS) regression where the response variable is a categorical one (Y-block; replaced by the set of dummy variables describing the categories) expressing the class membership of the statistical units. Therefore, PLS-DA does not allow for response variables other than those that define the groups of individuals. As a consequence, all measured variables play the same role with respect to the class (cultivar) assignment. PLS components result from the compromise between two purposes: describing the set of explanatory variables and predicting the response ones. The PLS-based classification benefits from such a property as it builds typologies with intrinsic prediction power. The model includes a calibration phase and a validation phase for which residual errors (root mean square error) were calculated (RMSEC, RMSECV).

The prediction ability of PLS-DA depends also on the number (k) of the LV used in the model. The optimal k value was determined by predicting the results for independent sets of samples (test set) for different values of k , and determining the

value of k for which the highest percentage of correct classification was found in the test set.

PLSDA calculates a "prediction probability" and a classification threshold for each class modelled. These are calculated using a Bayesian method which takes the predicted y values from the PLSDA model, fits a normal distribution to them, then uses that to calculate the probability of observing a given y-value.

The PLS-DA analysis provides the percentage of correct classification and the loadings of each species on each LV. This analysis expressed also the statistical parameters indicating the modelling efficiency indicated by sensitivity and specificity parameters. The sensitivity is the percentage of the species of a category accepted by the class model. The specificity is the percentage of the species of the categories different from the modelled one, rejected by the class model.

This analysis was performed using Matlab (rel. 7.1, PLSToolbox Eigenvector rel. 4.0) on the shape variables (X-block; harmonic coefficients). The X block (EFA coefficients) values were pre-processed with an abs procedure (takes the absolute values of the data). Each dataset, consisting of the harmonic coefficients of each view and the merceological status (I or K), was divided into two subsets: the first, containing 80% of fruits, was used for the class modelling and validation; the second (83 fruits) was used for the independent test. To optimally select the 20% test set, the Kennard-Stone algorithm (1969) was applied. This algorithm belongs to the family of space-filling algorithms and is based on Euclidean distances between data. These algorithms select objects without the *a priori* knowledge of a regression model. The hypothesis is that as the true model is rather complex, it requires an uniform distribution of objects in the information space (see also Costa and collaborators (2008) for further detail on methodology).

PLSDA was performed in order to discriminate each viewpoints and merceological status of: a) the 4 studied cultivars (Tonda di Giffoni, San Giovanni, Mortarella and Tonda Romana); b) the differences between the two rounded shape cultivars (Tonda di Giffoni and Tonda Romana); c) the differences between the two oblong shape cultivars (San Giovanni and Mortarella); d) The differences between rounded and oblong cultivars.

Results

The total number of harmonics (the ‘Nyquist frequency’) is equal to 90. The Fourier series was truncated at different k values, depending on the viewpoints and merceological status (I or K), at which the average cumulative power reached 99.99% (Table 2).

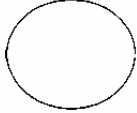
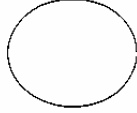
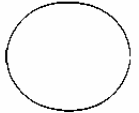
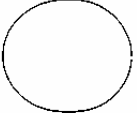
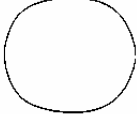
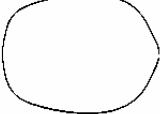
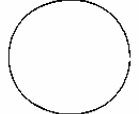
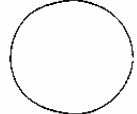
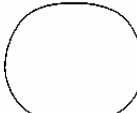

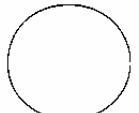
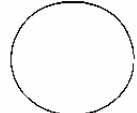
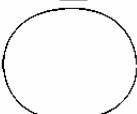
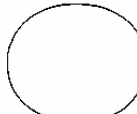
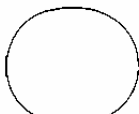
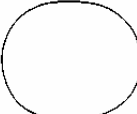


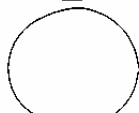
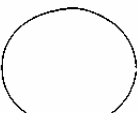


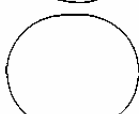

Table 2. k values at different view and status.

View	Status	Label	k
Polar	In-shell	PV-I	5
Lateral	In-shell	LV-I	7
Random plane	In-shell	RV-I	6
Polar	Unblanched kernel	PV-K	12
Lateral	Unblanched kernel	LV-K	10
Random plane	Unblanched kernel	RV-K	10

Figure 2 does represents an example of cumulative power of the Fourier series for the random plane view of the unblanched kernel (LV-I) where the Fourier series were truncated at 7 k harmonic equations.

In figure 3 the mean configurations of the different cultivars with different view and status are represented. The number of harmonics for each view and status depends on the one calculated and reported in table 2. It is possible to observe general differences between round shaped hazelnuts (Tonda di Giffoni and Tonda Romana) and oblong shaped hazelnuts (San Giovanni and Mortarella); meanwhile round shaped cultivars appear similar.

Figure 3: Mean configurations of the different cultivars with different view and status.

	MORTARELLA	SAN GIOVANNI	TONDA GIFFONI	TONDA ROMANA
PV- I				
LV- I				
RV- I				
PV- K				
LV- K				
RV- K				

a) PLSDA to discriminate the four cultivars (Tonda di Giffoni, San Giovanni, Mortarella and Tonda Romana)

In Table 3 (Section A) it is possible to observe the higher percentage of correct classification in the independent test for the in-shell (I) views, particularly for lateral view (97.5%) and random plane view (87.8%). In figure 4 the scatter plot of the first 3 LVs scores obtained by the PLSDA on the four cultivars on the LV-I is reported. Differences between cultivars are clearly observable. The PLSDA model for LV-I has high sensitivity (0.90) and specificity (0.88).

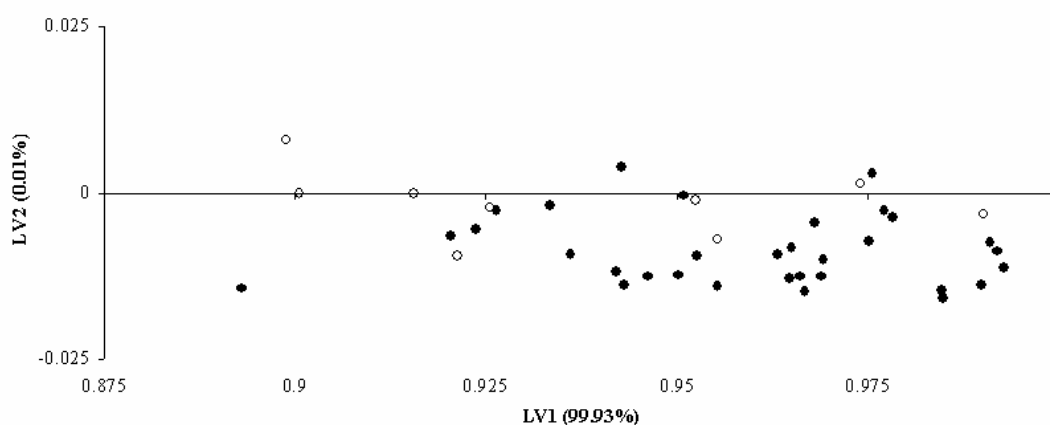
Table 3. PLS-DA summarized results; A) Discrimination between the four cultivars (cv); B) Discrimination between the two rounded shaped cv (Tonda Romana and Tond Giffoni); C) Discrimination between the two oblong shaped cv (San Giovanni and Mortarella; D) Discrimination between the rounded and oblong cv.

PLS DA	cv-view	N° LVs	% Classif Model	% Classif Test	Cumulated %			
					explained Variance (X- block)	Mean RMSECV	Mean Sensitivity	Mean Specificity
A	4C_LV-I	10	83.2	97.5	99.98	0.29	0.90	0.88
	4C_LV-K	11	74.5	77.5	99.95	0.33	0.79	0.85
	4C_PV-I	8	59.9	59.8	99.99	0.38	0.72	0.75
	4C_PV-K	6	53.0	34.7	99.87	0.40	0.63	0.74
	4C_RV-I	11	79.2	87.8	99.99	0.31	0.86	0.85
	4C_RV-K	10	65.7	69.0	99.94	0.34	0.82	0.78
B	RO_LV-I	4	85.6	90.0	99.96	0.35	0.86	0.86
	RO_LV-K	11	84.7	82.4	99.96	0.35	0.84	0.85
	RO_PV-I	6	75.2	73.2	99.98	0.43	0.74	0.73
	RO_PV-K	5	78.9	67.6	99.85	0.41	0.76	0.77
	RO_RV-I	7	78.8	95.1	99.97	0.40	0.74	0.76
	RO_RV-K	4	74.5	50.0	99.86	0.42	0.76	0.75
C	OB_LV-I	8	97.0	97.6	99.98	0.24	0.96	0.96
	OB_LV-K	10	87.2	89.2	99.94	0.33	0.85	0.87
	OB_PV-I	5	76.6	87.8	99.97	0.41	0.76	0.78
	OB_PV-K	9	80.5	73.0	99.92	0.40	0.80	0.82
	OB_RV-I	11	97.0	97.6	99.99	0.24	0.99	0.99
	OB_RV-K	9	91.3	91.9	99.93	0.30	0.91	0.90
D	OR_LV-I	11	91.1	98.8	99.98	0.27	0.95	0.94
	OR_LV-K	10	91.3	95.8	99.94	0.28	0.91	0.92
	OR_PV-I	7	78.6	75.6	99.99	0.40	0.78	0.78
	OR_PV-K	10	77.0	60.0	99.93	0.41	0.76	0.76
	OR_RV-I	10	92.5	97.6	99.98	0.29	0.92	0.93
	OR_RV-K	6	87.4	94.4	99.89	0.33	0.86	0.88

b) PLSDA to discriminate the two rounded shape cultivars (Tonda di Giffoni and Tonda Romana)

In Table 3 (Section B) it is possible to observe a higher percentage of correct classification in the independent test for the in-shell (I) views, particularly random plane view (95.1%) and lateral view (90.0%). In figure 5 differences between cultivars are not very visible on the first three axes even though the efficacy of discrimination of the PLSDA model with 7 LV is very high. However this model doesn't have high values of sensitivity (0.74) and specificity (0.76) indicating a limited efficacy and efficiency in modelling even with high performance of classification.

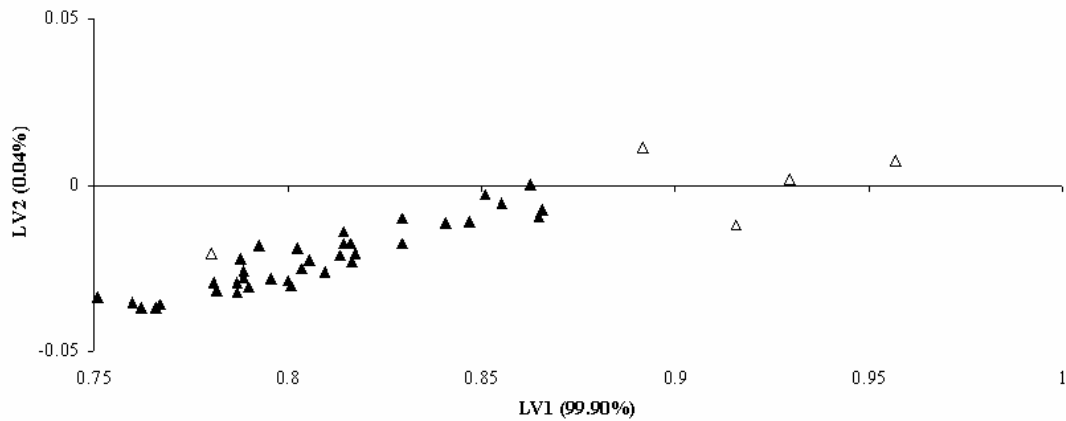
Figure 5: Scatter plot of the first 2 LVs obtained by the PLSDA on the rounded cultivars on the RV-I. Tonda di Giffoni (black circle); Tonda Romana (white circle).



c) PLSDA to discriminate the two oblong shape cultivars (San Giovanni and Mortarella)

In Table 3 (Section C) it is possible to observe the higher percentage of correct classification to discriminate between the two oblong cultivars in random plane view (RV-I=97.6% ; RV-K=91.9%) with respect to the other views, except LV-I (97.6%). In figure 6 differences between the two cultivars are clearly observable. The PLSDA model for RV-I has very high values of sensitivity (0.99) and specificity (0.99).

Figure 6: Scatter plot of the first 2 LVs obtained by the PLSDA on the oblong cultivars on the RV-I. San Giovanni (black triangle); Mortarella (white triangle).

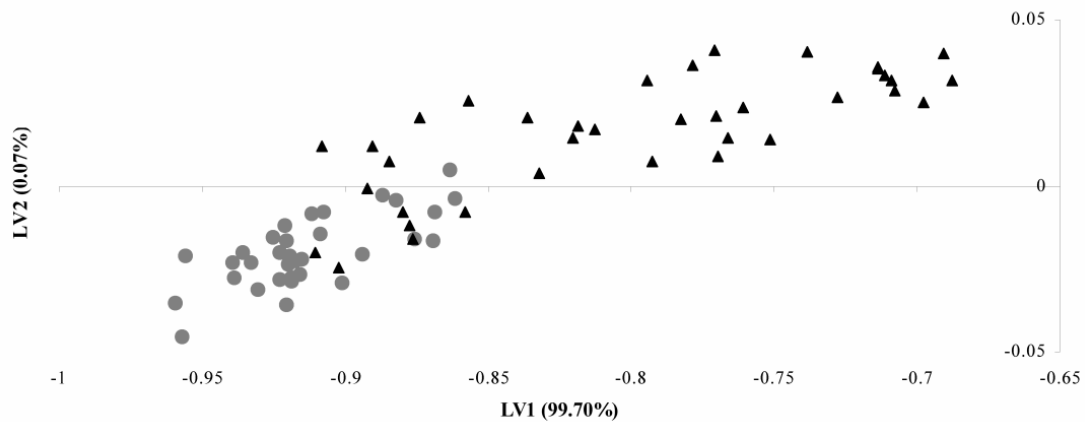


d) PLSDA to discriminate rounded and oblong cultivars

In Table 3 (Section D) it is possible to observe the higher percentage of correct classification to discriminate between oblong and rounded cultivars in both lateral view (LV-I=98.8% ; LV-K=95.8%) and random plane view (RV-I=97.6% ; RV-K=94.4%) with respect to the polar view. In figure 7 differences between the two cultivars are clearly observable. The PLSDA model for RV-K has high values of sensitivity (0.86) and specificity (0.88).

In general the lateral view and random plane view were better for classification, particularly for the in-shell hazelnuts.

Figure 7: Scatter plot of the first 2 LVs obtained by the PLSDA on the oblong (black triangle) vs rounded (gray circle) cultivars on the RV-K.



Discussion

Describing fruit shape is often necessary in agricultural research for a range of different purposes including cultivar description in applications for plant variety rights or cultivar registers (Paulus & Schrevens, 1999; Beyer et al., 2002; Paglia et al., 2008), evaluation of consumer preference (Kays, 1999), investigating heritability of fruit shape traits (Cannon & Manos, 2001), or analyzing shape abnormalities (Brewer et al., 2007).

Considering the reduced sample size and the controlled conditions of growth and harvest, the proposed methodology suggested high classification performance. In this study two different approaches were applied. The first is based on the appropriate and regular positioning of the fruit in the lateral and polar view. The second is based on the random positioning of the fruit on the plane.

Lateral and polar view

In this study the best results on classification were obtained with the lateral view positioning in particular of the in-shell fruit (LV-I; from 90.0 to 98.8%), followed by the lateral view of the kernel (LV-K; from 77.5 to 95.8%). Meanwhile the polar view positioning in hazelnuts did not produce good results (PV; from 34.7 to 87.8%). Shape is considered a fundamental property of a fruit together with size, colour, condition and absence of defects (Kays, 1999). Cultivar certification is based also on

a general and qualitative shape description (see for example Paulus & Schrevens, 1999). This study has demonstrated the technical feasibility of a quantitative shape description and determination of four hazelnuts cultivars. For example, the application of similar methods on different fruit leads to cultivar description (Paulus & Schrevens, 1999; Dubey et al., 2006) and fruit grading (Tao et al., 1995; Ghazanfari et al., 1997).

Random plane view

In this study the results obtained with the random plane view positioning are good especially for the in-shell fruits (RV-I; from 87.8 to 97.6%). This approach has the potential to be implemented in the future to develop new online sorting systems with high speed and minimum damage. However, it is necessary to undertake a validation on a wider sample size to cover the possible variations expected from growing hazelnuts in different regions, crop years and agro-pedo-climatic conditions.

Among the optoelectronic non-destructive methods, to our knowledge, none of spectroscopic techniques have been used for cultivar selection, while there are few shape-based studies for the same purpose (Tao et al., 1995; Paglia et al., 2008). In other research fields, especially botany and agronomy, there have been some shape-based studies that describe different varieties and species through analysis of leaves (Chi et al., 2002; Jensen et al., 2002; McLellan, 2005; Neto et al., 2006), fruits (Cannon and Manos, 2001; Beyer et al., 2002) or flowers (Dominguez et al., 1998; Yoshioka et al., 2004).

Conclusion

The implementation of modern multivariate techniques, such as PLS-DA, on shape-based methods, such as elliptic Fourier analysis, on digital images, demonstrates, in this preliminary study, a high efficiency and performances in fruit grading and classification.

The use of an image analysis system based on common chromatic digital standard (RGB or grayscale), instead of spectroscopic analysis, matches the general agro-industrial aim to develop selection machines based on consolidated and cheaper technology.

Acknowledgements

The research project "FRUMED" was supported by the Italian Ministry of Agriculture and Forestry, Paper no. 30. The authors would like to thank Mrs. Iliana Niciarelli and Mr Matteo Cegna for the support during the laboratory operation.

References

- Beyer M; Hahn R; Peschel S; Harz M; Knoche M (2002). Analysing fruit shape in sweet cherry (*Prunus avium* L.). *Scientia Horticulturae*, 96, 139-150
- Brewer M T; Moyseenko J B; Monforte A J; van der Knaap E (2007). Morphological variation in tomato: a comprehensive study of quantitative trait loci controlling fruit shape and development. *Journal of Experimental Botany*, 58(6), 1339–1349
- Cannon C H; Manos P S (2001). Combining and comparing morphometric shape descriptors with a molecular phylogeny: the case of fruit type evolution in Bornean *Lithocarpus* (Fagaceae). *Systematic Biology*, 50(6), 860–880
- Chi Y T; Chien C F; Lin T T (2002). Leaf shape modelling and analysis using geometric descriptors derived from Bezier curves. *Transactions of the ASAE*, 46(1), 175–185
- Costa C; Menesatti P; Aguzzi J; D'Andrea S; Antonucci F; Rimatori V; Pallottino F; Mattoccia M (2008). External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*. In Press *Food and Bioprocess Technology*.
- Crampton J S (1995). Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia*, 28, 179-186.
- Dominguez C A; Eguiarte L E; Nunez-Farfa J; Dirzo R (1998). Flower morphometry of *Rhizophora mangle* (Rhizophoraceae): geographical variation in Mexican populations. *American Journal of Botany*, 85(5), 637–643
- Dubey B P; Bhagwat S G; Shouche S P; Sainis J K (2006). Potential of artificial neural networks in varietal identification using morphometry of wheat grains. *Biosystems Engineering* 95(1), 61-67
- FAO (2000). Inventory of hazelnut research, germoplasm and references. REU Technical Series No.56

- Ferson S; Rohlf F J; Koehn R K (1985). Measuring shape variation of two-dimensional outlines. *Systematic Zoology*, 34, 59-68
- Furuta N; Ninomiya S; Takahashi S; Ohmori H; Ukai Y (1995). Quantitative evaluation of soybean (*Glycine max* L., Merr.) leaflet shape by principal component scores based on elliptic Fourier descriptor. *Breeding Science*, 45, 315–320
- Ghazanfari A; Irudayaraj J; Kusalik A; Romaniuk M, (1997). Machine vision grading of pistachio nuts using Fourier descriptors. *Journal of Agricultural Engineering Research*, 68, 247–252
- Goto S; Iwata H; Shibano S; Ohya K; Suzuki A; Ogawa H (2005). Fruit shape variation in *Fraxinus mandshurica* var. *japonica* characterized using elliptic Fourier descriptors and the effect on flight duration. *Ecological Research*, 20, 733–738
- Iwata H; Ukai Y (2002). SHAPE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *Journal of Heredity*, 93, 384–385
- Iwata H; Niikura S; Matsuura S; Takano Y; Ukai Y (1998). Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors. *Euphytica*, 102, 143–149
- Jensen R J; Ciofani K M; Miramontes L C (2002). Lines, outlines, and landmarks: morphometric analyses of leaves of *Acer rubrum*, *Acer saccharinum* (Aceraceae) and their hybrid. *Taxon*, 51, 475-492
- Kays S J (1999). Preharvest factors affecting appearance. *Postharvest Biology and Technology*, 15, 233–247
- Kennard R W; Stone L A (1969). Computer aided design of experiments. *Technometrics*, 11, 137-148
- Kim S W; Schatzki T F (1998). Detection of insect damages in almonds. *Proceedings of SPIE*, 3543, 101-110
- Kuhl F P; Giardina C R (1982). Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, 18, 236–258
- Lestrel P E (1997). *Fourier descriptors and their applications in biology*. Cambridge Univ. Press.

- Loy A; Busilacchi S; Costa C; Ferlin L; Cataudella S (2000). Comparing geometric morphometrics and outlines fitting methods to monitor fish shape variability of *Diplodus puntazzo* (Teleostea , Sparidae). *Aquacultural Engineering*, 21(4), 271-283
- McLellan T (1993). The roles of heterochrony and heteroblasty in the diversification of leaf shape in *Begonia dregei* (Begoniaceae). *American Journal of Botany*, 80, 796–804
- McLellan T (2005). Correlated evolution of leaf shape and trichomes in *Begonia dregei* (Begoniaceae). *American Journal of Botany*, 92(10), 1616–1623
- Nagata M; Cao Q (2000). Study on grade judgment of fruit vegetables using machine vision. *Japan Agricultural Research Quarterly*, 32(4), <http://www.jircas.affrc.go.jp/kankoubutsu/jarq/32-4/nagata/nagata.htm>
- Neto J C; Meyer G E; Jones D D; Samal A K (2006). Plant species identification using Elliptic Fourier leaf shape analysis. *Computers and Electronics in Agriculture*, 50(2), 121-134
- Ohsawa R; Tsutsumi T; Uehara H; Namai H; Ninomiya S (1998). Quantitative evaluation of common buckwheat (*Fagopyrum esculentum* Moench) kernel shape by elliptic Fourier descriptor. *Euphytica*, 101, 175–183
- Paglia G; Antonucci F; Costa C; De Giorgio D; Menesatti P (2008). Studio sulle differenze di forma in frutti di cultivar di mandorlo (*Prunus amygdalus* Batsch) attraverso l’analisi ellittica di Fourier. VIII Convegno Nazionale “La biodiversità – una risorsa per sistemi multifunzionali”, 21-23 Aprile 2008 Lecce.
- Paulus I; Schrevens E (1999). Shape characterization of new apple cultivars by Fourier expansion of digitized images. *Journal of Agricultural Engineering Research*, 72, 113–118
- Pearson T; Toyofuku N (2000). Automated sorting of pistachio nuts with closed shells. *Applied Engineering in Agriculture*, 16(1), 91-94
- Rohlf F J (1996). Fitting curves to outlines. *Proceedings of Michigan Morphometric Workshop, Special Publication*, 2, 167-177
- Rohlf F J (2006). TpsDig Ver. 2.10, digitalized landmarks and outlines. Dept. of Ecology and Evolution, State Univ. of New York at Stony Brook.

- Rohlf F J; Archie J W (1984). A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Systematic Zoology*, 3, 302-317
- SEC (2002). Analysis of the nut sector. Brussels, 10.7.2002. SEC (2002) 797
- Sabatier R; Vivein M; Amenta P (2003). Two approaches for Discriminant Partial Least Square. In: Schader M., Gaul W., Vichi, M. (Eds.): *Between Data Science and Applied Data Analysis*, Springer-Verlag, Berlin
- Sheets H D; Covino K M; Panasiewicz J M; Morris S R (2006). Comparison of geometric morphometric outline methods in the discrimination of age-related differences in feather shape. *Frontiers in Zoology*, 3, 15
- Sjöström M; Wold S; Söderström B (1986). PLS Discrimination plots. In: Gelsema E.S, Kanals L.N. (Eds.): *Pattern Recognition in Practice II*. Elsevier, Amsterdam
- Slice D E (1998). Morpheus et al.: software fore morphometric research. Revision 01-30-98-Beta. Department of ecology and evolution, State University of New York, Stony Brook, New York
- Tao Y; Morrow C; Heinemann P H; Sommer H J (1995). Fourier-based separation technique for shape grading of potatoes using machine vision. *Transactions of the ASAE*, 38(3), 949-957
- Uga Y; Fukuta Y; Cai H W; Iwata H; Ohsawa R; Morishima H; Fujimura T (2003). Mapping QTLs influencing rice floral morphology using recombinant inbred lines derived from a cross between *Oryza sativa* L. and *Oryza rufipogon* Griff. *Theoretical and Applied Genetics*, 107, 218–226
- White R J; Prentice H C; Verwijst T (1988). Automated image acquisition and morphometric description. *Canadian Journal of Botany*, 66, 450–459
- Yoshioka Y; Iwata H; Ohsawa R; Ninomiya S (2004). Analysis of petal shape variation of *Primula sieboldii* by elliptic Fourier descriptors and principal component analysis. *Annals of Botany*, 94, 1–8

3.6 Quantitative method for shape description of almond cultivars (*Prunus amygdalus* Batsch)

FULL REFERENCE OF THE PUBLICATION

Antonucci F, Costa C, Pallottino F, Paglia G, Rimatori V, De Giorgio D, Menesatti P, ACCEPTED. Quantitative method for shape description of almond cultivars (*Prunus amygdalus* Batsch). IN PRESS TO FOOD AND BIOPROCESS TECHNOLOGY

Introduction

The cultivated almond, *Prunus amygdalus* Batsch, apparently originated from one or more of the many wild species that evolved in the deserts and Lower Mountain slopes of central and southwest Asia (Kester & Ross, 1996). Nowadays it is cultivated either intensively or semi-intensively in many areas of the world. Almond trees have been extremely distributed over the warm temperate region (Agunbiade & Olanlokun, 2006). Almonds are grown as orchard crops, highly nutritious and expensive, representing a luxury food.

Fruit from new cultivars can command a premium (Scott, 1998) creating incentives for breeders to continuously develop new types to maintain premium returns. The growing economic and commercial interests in this species lead the variety selection towards the development of new cultivars. The required attributes are an increased productivity, adaptability to different agro-climatical and agro-pedological contexts, better resistance to diseases, and higher qualitative, organoleptic and nutritional characteristics. New varieties must be tested in different agro-pedoclimatical contexts and described through standard techniques.

A descriptor list regarding just in-shell almond fruit with shape charts was developed by International Board for Plant Genetic Resources (IBPGR) (Gülcan, 1985). This document provides an international format producing a universally understood “language” for plant genetic resources data collection. This approach assists with the standardization of descriptor definitions. Descriptors should be used when they are useful to both, the researcher for the management and maintenance of the collection

and to the users of the plant genetic resources. Moreover a collection of carpological measurements and a productive evaluation of 205 almond tree cultivars were reported by the CRA-SCA Research Unit for Cropping Systems Agriculture in Dry Environments (De Giorgio et al., 1996). The appearance of agricultural products is a primary criterion in the purchasing decisions of consumers, being at the base of selection judgement of the product's individual unit quality (Kays, 1999). Fruit and vegetable appearances are defined by size, shape, form, colour and absence of visual defects. Appearance represents a fundamental attribute to select throughout all the market production chain, storage, marketing, and finally, utilization (Kays, 1999). Nowadays carpological measurements are utilised as descriptors of cultivars. Specific qualitative indices describing differences among cultivars, built on shape description, are present nowadays for many other fruit.

Sensory evaluation of agricultural products is a growing sector of particular interest, not only for typical products such as those having a Protected Designation of Origin (PDO) or Protected Geographical Indication (PGI), but also for the ones that wish to obtain a quality label or recognition. Moreover these non-destructive, fast, repeatable and cheap techniques represent an opportunity for identifying market sector preferences. Turner et al. (2008) found that both size and shape influenced the consumer preference in sweet cherry. For pears, Kappel et al. (1995) found that the highest preference was for a specific shape (pyriform) and carpological descriptor (ratio between length and diameter). Also, for apple fruit Hampson & Quamme (2000) conclude that method able to identify tolerance limits for fruit shape and size attributes that could be used by breeders to assess the preferences of their target consumer population.

Furthermore such techniques can be used as verification techniques and protect the typicity against imitations or frauds.

Quantifying and describing fruit shape can be necessary in agricultural research for a variety of purposes. These include cultivar description in applications such as Plant Variety Rights (PVR) or cultivar registers (Khanizadeh 1994; Beyer et al., 2002), evaluation of consumer preference (Kays, 1999), investigating heritability of fruit shape traits (Cannon & Manos, 2001), or analyzing shape abnormalities (Brewer et al., 2007).

In this scenario various methods for quantitatively evaluating botanical shapes have been suggested. The most common methodology is based on Elliptic Fourier descriptors, which was successfully used on leaves (White et al., 1988; Jensen et al., 2002; Neto et al., 2006), leaflets (Furuta et al., 1995), kernels (Ohsawa et al., 1998), roots (Iwata et al., 1998), flowers (Yoshioka et al., 2004), and fruit (Currie et al., 2000; Goto et al., 2005). This method mathematically describes the entire shape of an object by transforming the contour into Fourier coefficients. Few works ever considered the analysis of fruit shape by means of optoelectronic techniques approach (i.e. digital images are analysed by means of combined morphometry and multivariate statistics) (Menesatti et al., 2008). In botanical studies, PCA (Principal Components Analysis) approach was used to observe shape differences between groups at different levels (cultivar or species). A recent study by Menesatti et al. (2008) aimed to create a classification and modelling system able to attribute existing shape classes to new genotypes for industrial sorting equipments implementation. For this kind of approach, supervised multivariate modelling such as PLS-DA (Partial Least Squares Discriminant Analysis) (Menesatti et al., 2008) or SIMCA (Soft Independent Modelling of Class Analogy) (Aguzzi et al., 2009) are suggested.

The aim of the present work was to propose a rapid, non-invasive and quantitative image analysis method based on Elliptic Fourier Analysis (EFA) and on carpological measurements, to discriminate between 18 cultivars and shape groups of almonds kernels and in-shell fruit. The shape groups were identified using two clustering techniques: a non-hierarchical method (k-means) and a hierarchical one (Ward's method).

Results were screened by PLS-DA. The proposed approach offers the possibility to operate selecting the fruit stock identifying two levels of shape differences.

Materials and methods

Data collection

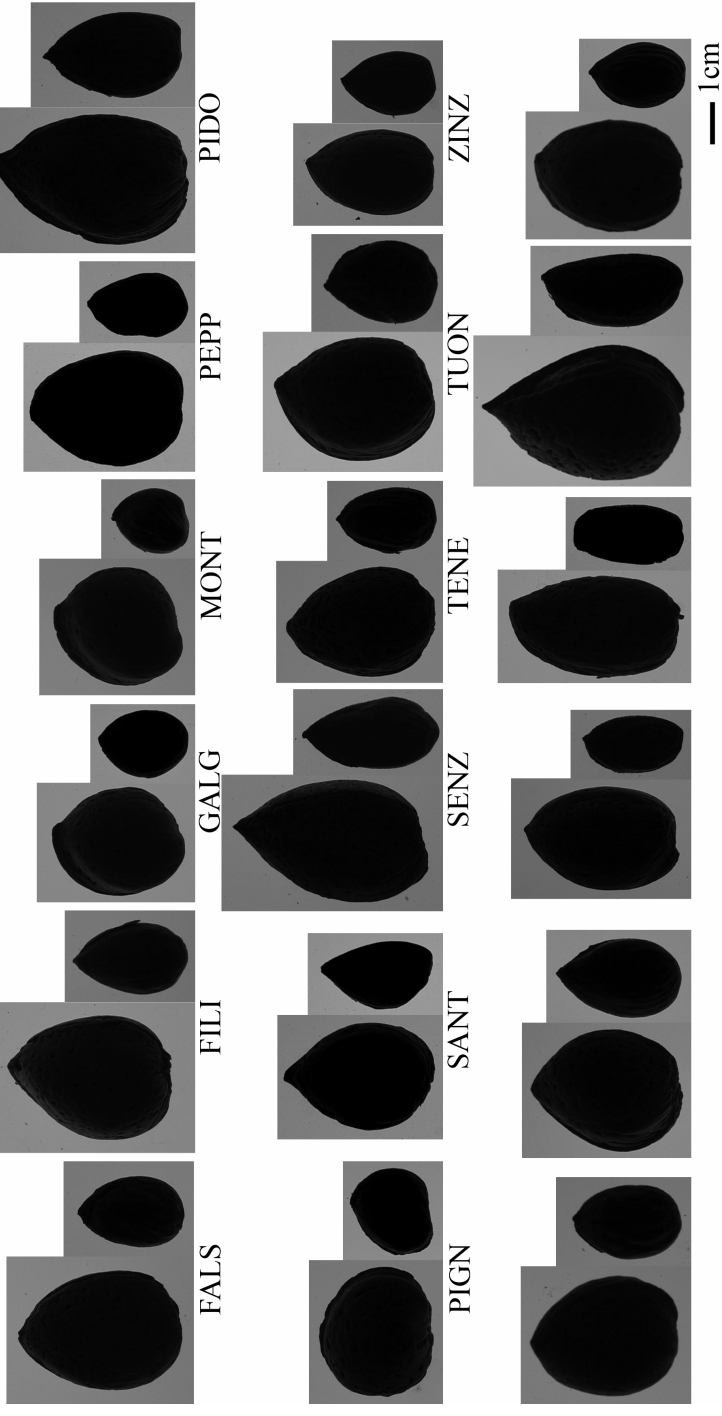
A total number of 864 in-shell fruit (48 *per* cultivar), among which 850 kernels, belonging to 18 commercially important Italian almonds cultivars (Banchiere, Barlettana, Catuccia, Catucedda, Cosimo di Bari, Cristomorto, Falsa Catuccia, Filippo Ceo, Galgano, Montrone, Pepparuddo, Pidocchioso, Pignatidde Tardiva,

Santoro, Senz'Arte, Tenente, Tuono, Zin Zin) were analyzed in this study (Tab. 1, Fig. 1). All fruit were provided by the “almonds germoplasm orchard” (CRA-SCA of Bari) and were sampled 5 times between December 2006 and April 2007. To avoid the possible influence of maturity stage on shape variation, fruit samplings were conducted in two moments: before and during the optimal harvesting time for each cultivar.

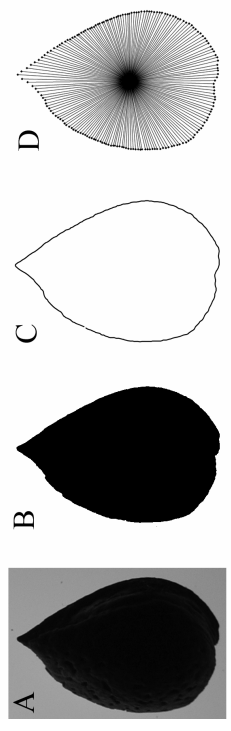
Table 1: Cultivars, Labels, Number of in-shell fruit and Number of kernels.

Cultivars	Label	N° in-shell fruit	N° kernels
Banchiere	BANC	48	48
Barlettana	BARL	48	42
Catuccia	CATC	48	47
Catucedda	CATD	48	48
Cosimo di Bari	COSI	48	48
Cristo Morto	CRIS	48	45
Falsa Catuccia	FALS	48	48
Filippo Ceo	FILI	48	48
Galgano	GALG	48	47
Montrone	MONT	48	48
Pepparuddo	PEPP	48	48
Pidocchioso	PIDO	48	46
Pignatidde Tardiva	PIGN	48	48
Santoro	SANT	48	48
Senz'arte	SENZ	48	48
Tenente	TENE	48	47
Tuono	TUON	48	48
Zin zin	ZINZ	48	48

Fig. 1 Top part of the figure: image examples for the 18 cultivars utilised with metric reference and labels (on the left the in-shell fruit, on the right the kernels).



Bottom part from the left to the right: A) original image acquired converted in greyscale; B) image binarized; C) outline; D) 180° points (x,y) equally and angularly spaced (one point every 2°) from the centroid along the outline.



Digital image analysis

For each fruit a digital image was acquired at high resolution with a digital camera Nikon (D1X) equipped with the lens Nikon AF (24mm f/2.8D). Each fruit was placed on an illuminated dashboard, increasing the contrast, in a lateral position directed to the right side with a metric reference. The metric mark of reference was placed on a side of each portrayed fruit in order to record its size. Each image was primarily binarized and a total number of 180 points (x, y) equally angularly spaced (one point every 2°) from the centroid were digitized, along the outline with the software TPSdig2 (Rohlf, 2006). Coordinates were aligned by Generalised Procrustes Analysis (GPA), a procedure that consists in three steps: the translation of points coordinates to a common centroid located at the origin (0,0) of a reference system of coordinates; the scaling of each outline with the unitary centroid size; and finally, the rotation of coordinates to minimize the sum of square distances between correspondent landmarks (Bookstein, 1991). The 180 aligned coordinates were treated as outlines data (Menesatti et al., 2008).

Elliptic Fourier Analysis (EFA)

The overall shape (of in-shell fruit and of kernels) was studied by Elliptic Fourier Analysis on the outline coordinates (Rohlf & Archie, 1984). Fruit shape outline can be approximated by a polygon of xy -coordinates. EFA is based on the separate Fourier decompositions of the incremental changes of the x - and y -coordinates as functions of the cumulative chordal length of the outline polygon. It yields to the spectrum of the fruit shape closed contour in terms of harmonically related trigonometric curves. For each harmonic equation, two Fourier coefficients are computed for both the x - and y -projections, thus the total number of coefficients is $4n$, where n is the number of harmonics fitted to the outline (Crampton, 1995). The total number of harmonics that can be computed for any outline is equal to half of the total number of outline coordinates (the 'Nyquist frequency'). The Fourier series was truncated at the value of k at which the average cumulative power is 99.999% of the average total power (Menesatti et al., 2008). For any outline, the total power was calculated as the sum, from 1 to k , of individual harmonic powers where k is equal to the Nyquist frequency (Crampton, 1995). The harmonic coefficients describe the size, shape and orientation of each harmonic ellipse and form the input to multivariate

statistics. The coefficients of the harmonic equations were extracted with the software Morpheus (Slice, 1998) and analysed via multivariate classification. Mean outline, standard deviation (SD) ranges were graphically reported for each cv and variance was reported.

Clustering

K-means clustering classifies a given multivariate dataset through a certain number of clusters (k) fixed a priori (Hill & Lewicki, 2007; Chitta & Narasimha Murty, 2010). K-means uses an iterative algorithm that minimizes the sum of distances from each object to its cluster centroid, over all clusters. This algorithm moves objects between clusters until the sum cannot be decreased further. The result is a set of clusters that are as compact and well-separated as possible.

Since the k-means method aims to minimize the sum of squared distances from all points to their cluster centres, this should result in compact clusters. It is possible to use the distances of the points from their cluster centre to determine whether the clusters are compact. For this purpose, the intra-cluster distance measure, which is simply the distance between a point and its cluster centre are calculated; the average of all of these distances, defined as:

$$\text{int } ra = \frac{1}{N} \sum_{i=1}^K \sum_{x \in C_i} \|x - z_i\|^2 \quad (1)$$

where N is the number of pixels in the image, K is the number of clusters, and z_i is the cluster centre of cluster C_i . We obviously want to minimize this measure. The inter-cluster distance, or the distance between clusters, should be as big as possible. It is calculated as the distance between cluster centres and takes the minimum of this value, defined as

$$\text{int } er = \min \left(\|z_i - z_j\|^2 \right)_{i=1,2,\dots,K-1, j=i+1,\dots,K} \quad (2)$$

Only the minimum of this value was taken. Since both of these measures help us to determine a good clustering the ratio between these two measures was calculated for each k:

$$\text{validity} = \text{mean int } ra / \text{mean int } er \quad (3)$$

The clustering which gives a minimum value for the validity measure is the ideal value of k in the k-means procedure. This procedure is proposed by Ray & Turi (1999) and implemented in Matlab (Attached 1) introducing an iterative procedure (1000 cycles) to smooth the stochastic attitude of k-means procedure, was applied on the entire samples of in-shell fruit and kernels from $k=2$ to $k=25$. The coefficients of the harmonic equations and of the carpological measurements for each cv were clustered (k-means) with the unsupervised k-means clustering technique, attributing the group following the mode value for each cv. To obtain the mean configuration of each group the mean of the coefficients of each fruit belonging to each group was calculated and graphically reported.

The hierarchical clustering routine produces a dendrogram showing how data points can be clustered. As suggested by Doering & Ludwig (1990), the Ward's method algorithm was used (Ward, 1963). Clusters are joined such that increase in within-group variance is minimized. For Ward's method, a Euclidean distance measure is inherent to the algorithm.

Classification analysis

The groups obtained from the k-means clustering on EFA coefficients and carpological measurements were analyzed via PLS-DA. For each analysis each subgroup was subdivided in two groups: i) 75% of specimens for the class modelling and validation; ii) 25% of specimens for the independent test, optimally chosen with the Euclidean distances based on the algorithm of Kennard & Stone (1969) that selects objects without the *a priori* knowledge of a regression model (i.e. the hypothesis is that a flat distribution of the data is preferable for a regression model) (Maesschalck et al., 1999). PLS-DA (Sabatier et al., 2003) is a partial least squares analysis (PLS) regression where the response variable is categorical (Y-block; replaced by a set of dummy variables describing the categories) expressing the class membership of the statistical units. For the methodology of this analysis see Costa et al. (2008; 2010) and Menesatti et al. (2008). This analysis also expresses the statistical parameters indicating the modelling efficiency in terms of sensitivity and specificity of the parameters. The sensitivity is the percentage of the samples of a category accepted by the class model. The specificity is the percentage of the samples of the categories

different from the modelled one, rejected by the class model. Generally, the trend of the residual errors is decreasing in the calibration phase (Root Mean Square Error of Calibration; RMSEC) and increasing for the validation phase (Root Mean Square Error of Cross-Validation; RMSECV) (Pallottino et al., 2010). This analysis was performed using Matlab (rel. 7.1; PLSToolbox Eigenvectorb 4.0) on the shape variables (X-block; EFA coefficients and carpological measurements). The X-block values were pre-processed with an “autoscale” Matlab procedure. This procedure centres the columns to a zero mean value and scales it to the unit variance. The best performing models were selected among the ones with a number of LV ranging from 1 to 20.

Results and discussion

The correct number of harmonics to be used for the computation of the lateral profile of in-shell fruit and kernels was calculated on the whole samplings (‘Nyquist frequency’ = 90). The value selected for the analysis (i.e. the first value exceeding the 99.999%), was equal to 27 harmonics for the in-shell fruit and to 31 for the kernels.

Cultivar discrimination

Figure 2 and Figure 3 report respectively the mean configuration (black line), the standard deviation range (grey lines) and the relative variance, for each cultivar of in-shell fruit and kernels.

Fig. 2 Mean outline (black line) and the standard deviation range (gray lines) for each cv of in-shell fruit extracted with 27 harmonics. On the top-right corner for each configuration are reported the values of variance for each cv. See Table 1 for cv codes.

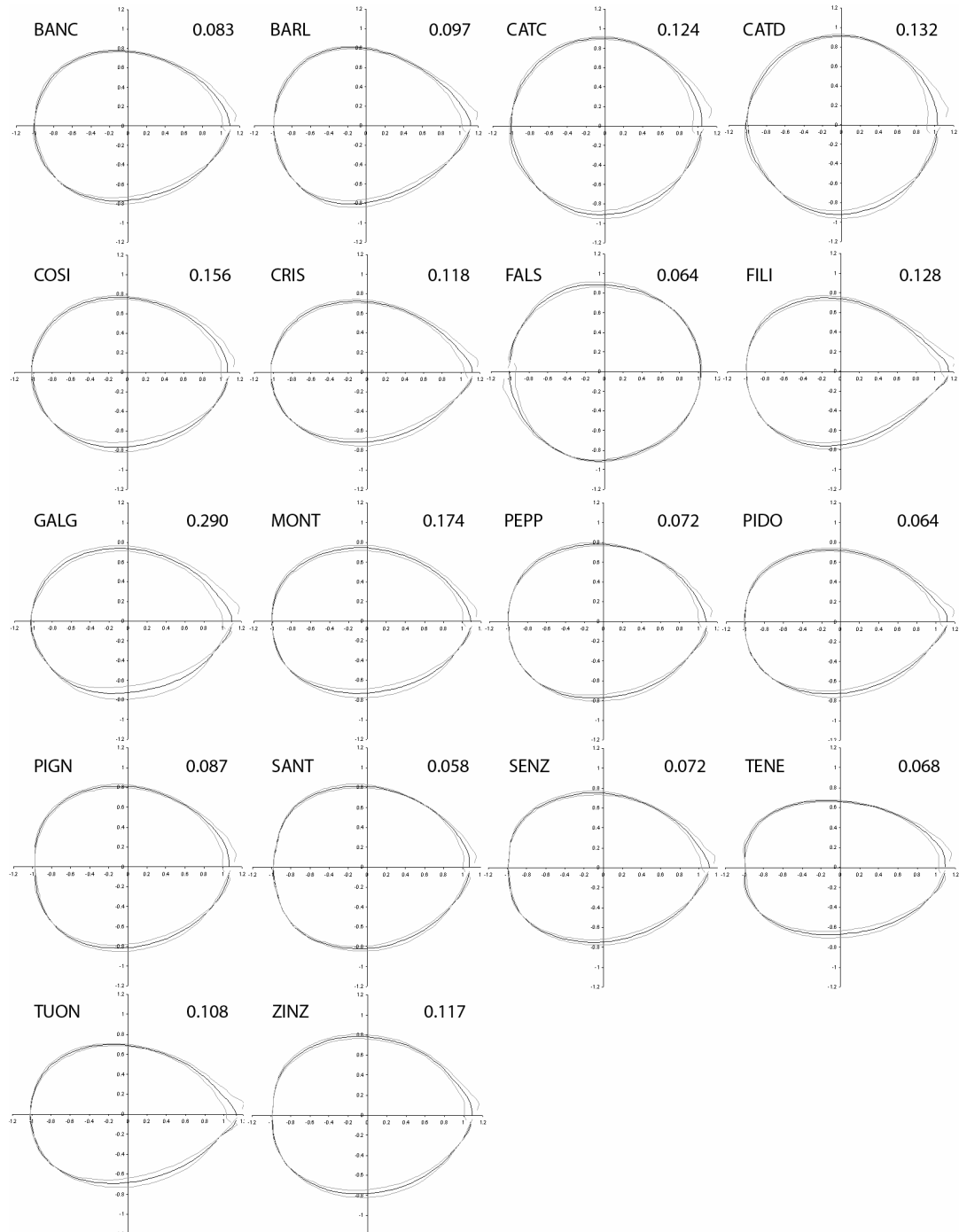
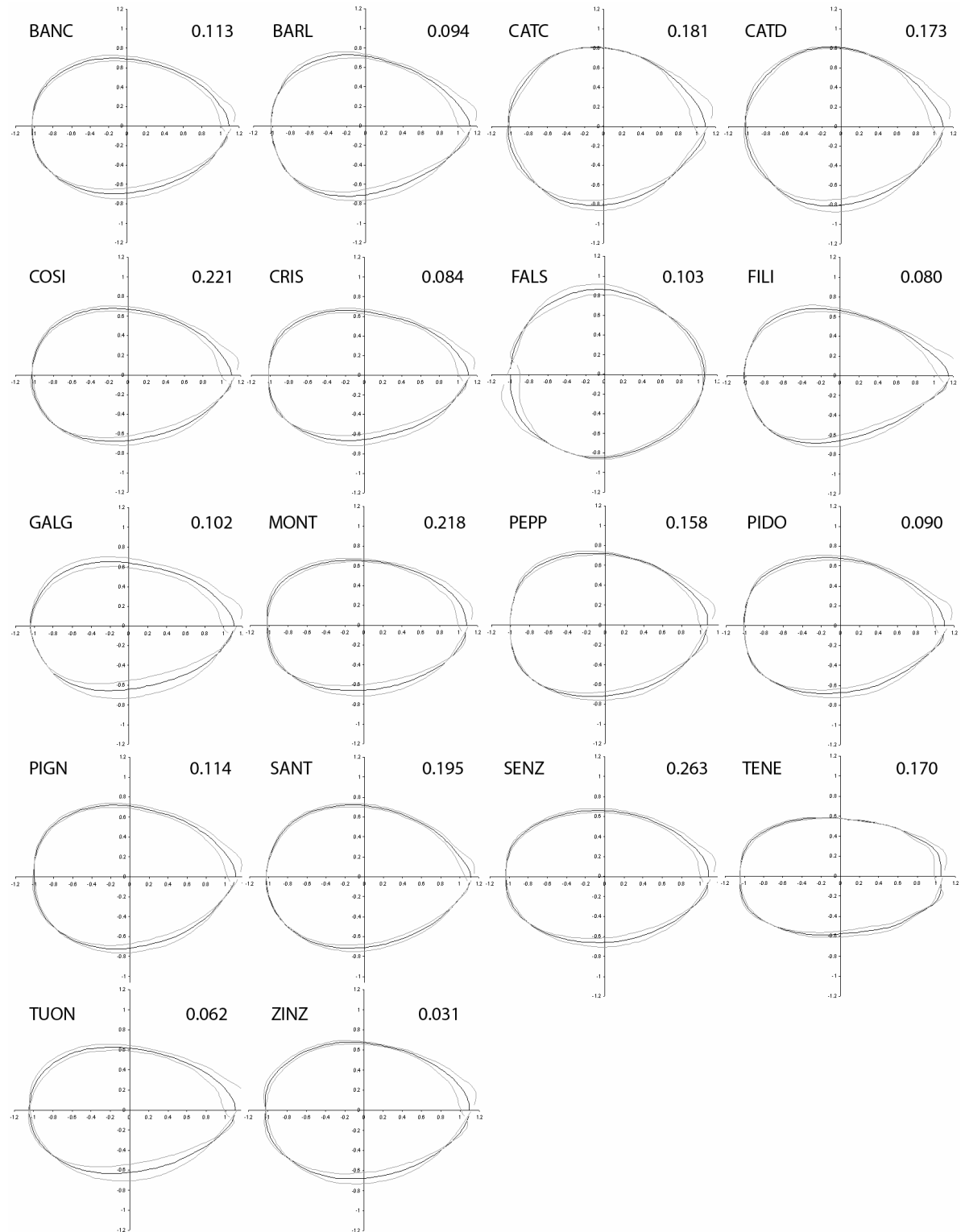


Fig. 3 Mean outline (black line) and the standard deviation range (gray lines) for each cultivar of kernels extracted with 31 harmonics. On the top-right corner for each configuration are reported the values of variance for each cv. See Table 1 for cv codes.



In Figure 2 is possible to observe as different well distinguishable shapes belong to different cv. CATC, CATD and FALS appear more spherical. BARL, CRIS, FILI, PIDO, SENZ and TUON have an ovoid shape with an acute apex and a convex base. A more squared shape belongs to TENE. It is also possible to notice a higher variance in GALG (0.290) while this is lower in SANT (0.058).

The kernels show similar characterization. CATC, and FALS appear more spherical. BARL, FILI, SANT and PIGN have an ovoid shape with an acute apex and a convex base. TENE and MONT are characterized by a squared shape. It is also possible to observe a higher variance in GALG (0.593) and the lower of the cv TENE (0.038).

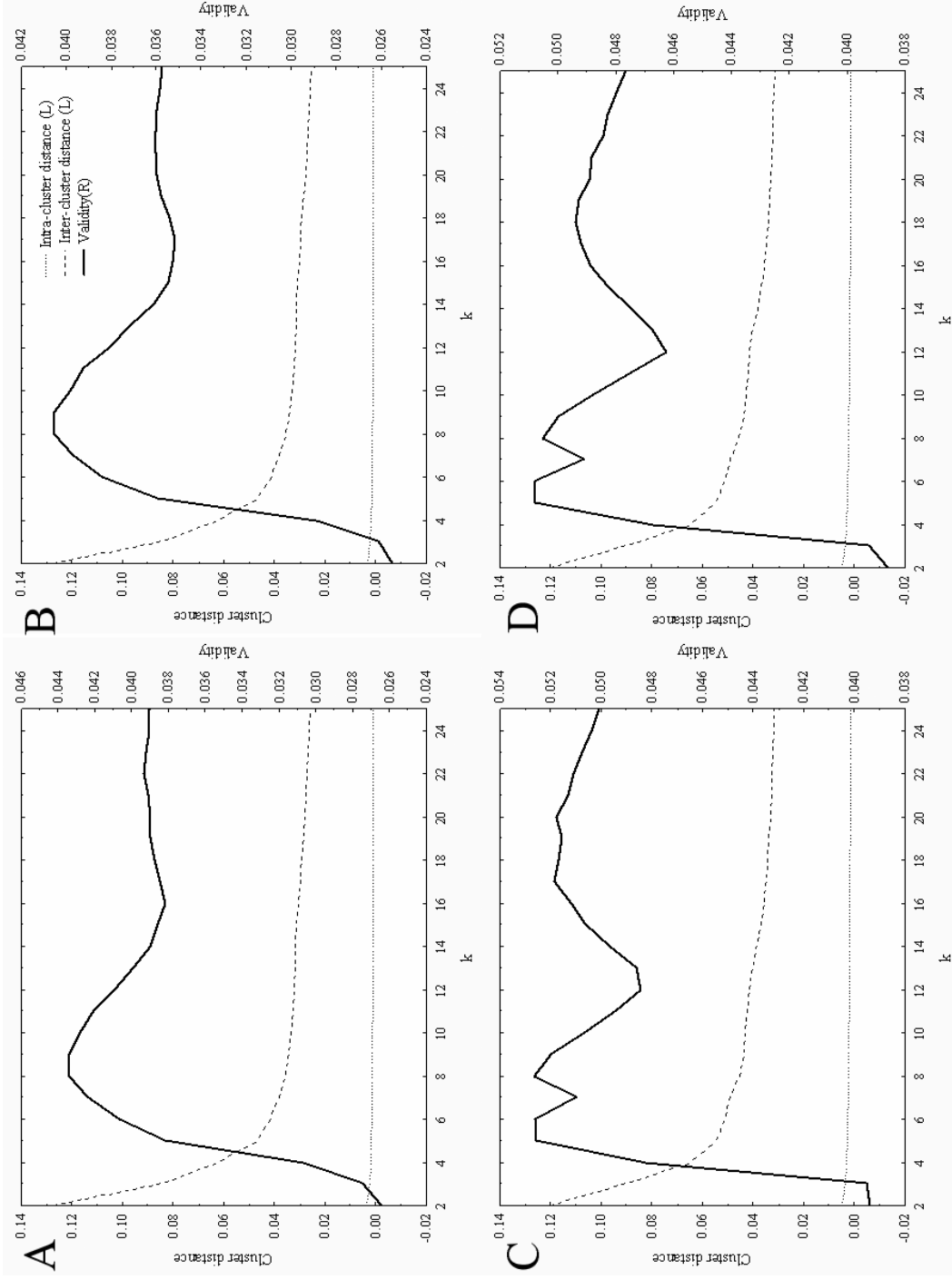
Shape groups

Non-hierarchical method (k-means)

The results of the procedure of validation to find the best number of k-clusters are reported in Figure 4. The best number of clusters to be used on this dataset was chosen above the plateau (0.09 regarding both, in-shell fruit and kernels). For the in-shell fruit the analysis extracted a value of k=2 and k=5 for the EFA coefficients (Fig. 4A) and carpological measurements (Fig. 4B). Similar results were found regarding the kernels with k values of 2 and 4 (Fig. 4C, 4D) respectively.

Fig. 4 Results of the

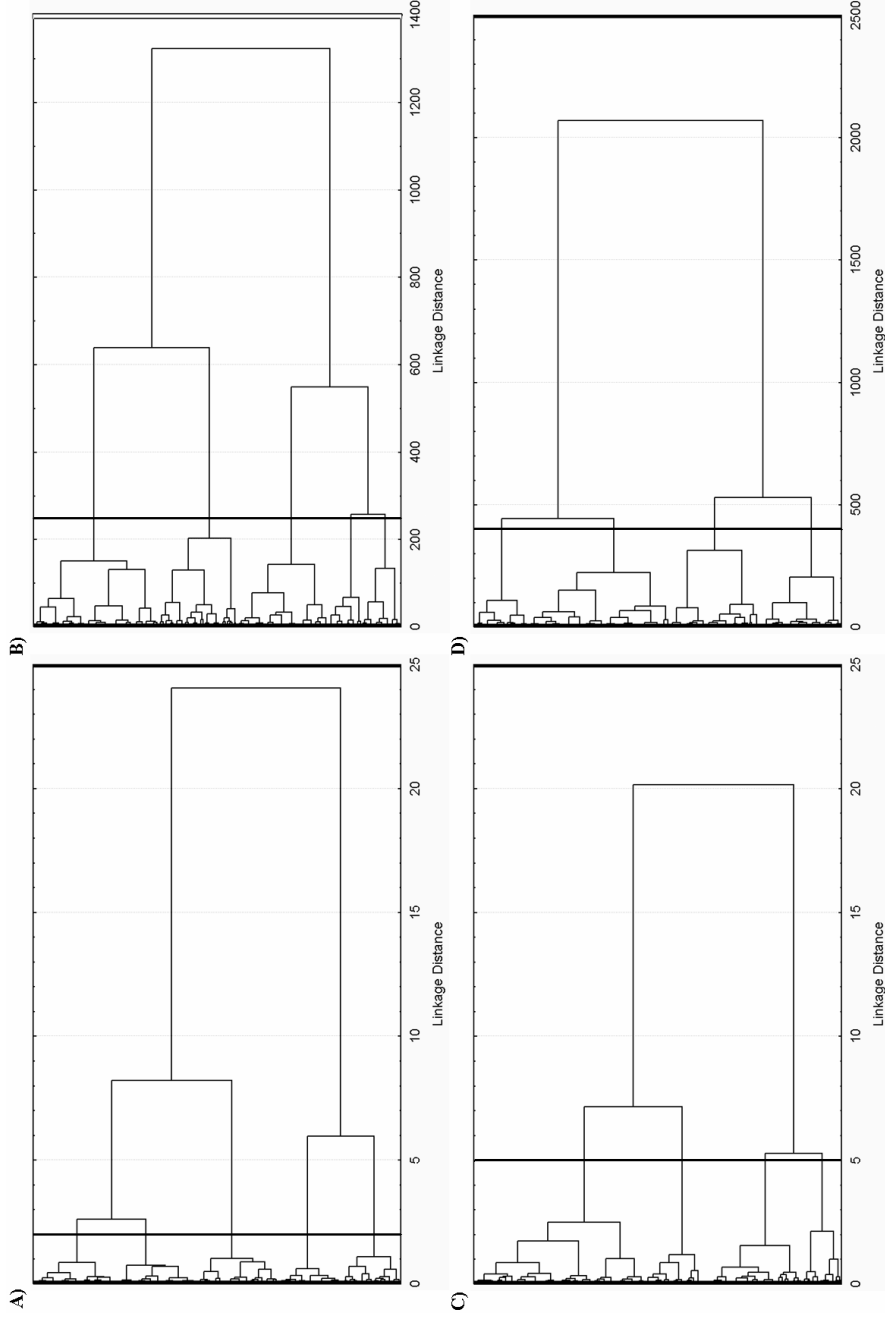
validation procedure to find the best number of k clusters for the in-shell fruit (A, EFA coefficients; B, carpological measurements) and kernels (C, EFA coefficients; D, carpological measurements).



Hierarchical one (Ward's method)

In Figure 5 are reported the dendrograms built with the hierarchical clustering routine (Ward's method). Figure 5A and 5B show the dendrograms based on the EFA coefficients and carpological measurements for the in-shell while Figure 5C and 5D for the kernels. Generally each dendrogram shows as larger ramifications divided the specimens in two principal groups. For the in-shell fruit five groups were obtained cutting the dendrograms at the linkage distance of 2 for EFA coefficients (Fig. 5A) and of 250 for carpological measurements (Fig. 5B). Regarding the kernels, four groups were defined cutting the dendrograms at the linkage distance of 5 for EFA coefficients (Fig. 5C) and of 400 for carpological measurements (Fig. 5D).

Fig. 5 Dendrograms built with a hierarchical clustering routine. A and B respectively represent dendrograms of EFA coefficients and carpological measurements of in-shell fruit; C and D dendrograms of EFA coefficients and carpological measurements of kernels. On the y axis are represented the single individual belonging to each cv. Vertical lines for each dendrogram represent the thresholds values at which the branches were cut to extract the correct number of shape groups.



Shapes description

In Figure 6 are reported the mean configurations (black line) and the standard deviation ranges (grey lines). In this case both, in-shell fruit and kernels, were extracted using a value of $k=2$. Each cluster in the same group follows the mode. For the in-shell fruit: Group 1 (CATC, CATD, FALS, PIGN, and SANT) has a spheroid/ellipsoid shape with a rounded apex and a convex base; Group 2 (BANC, BARL, COSI, CRIS, FILI, GALG, MONT, PEPP, PIDO, SENZ, TENE, TUON and ZINZ) has an ovoid/elongated shape with an acute apex and a convex base. For the kernels: Group 1 (CATC, CATD, FALS, PEPP, PIGN and SANT) is described by a quite ellipsoid shape with an acute apex and a convex base; Group 2 (BANC, BARL, COSI, CRIS, FILI, GALG, MONT, PIDO, SENZ, TENE, TUON and ZINZ) has an ovoid/elongated shape with an acute apex and convex base. The only difference between the cv grouping, based on in-shell fruit and kernels, is the cv PEPP. In the kernels clustering this cv is associated with Group 2 while, in in-shell fruit clustering it is associated with Group 1.

Fig. 6 Mean outline (black line) and the standard deviation range (grey lines) for in-shell fruit (upper side of Figure) and kernels (bottom side of Figure) extracted with $k=2$.

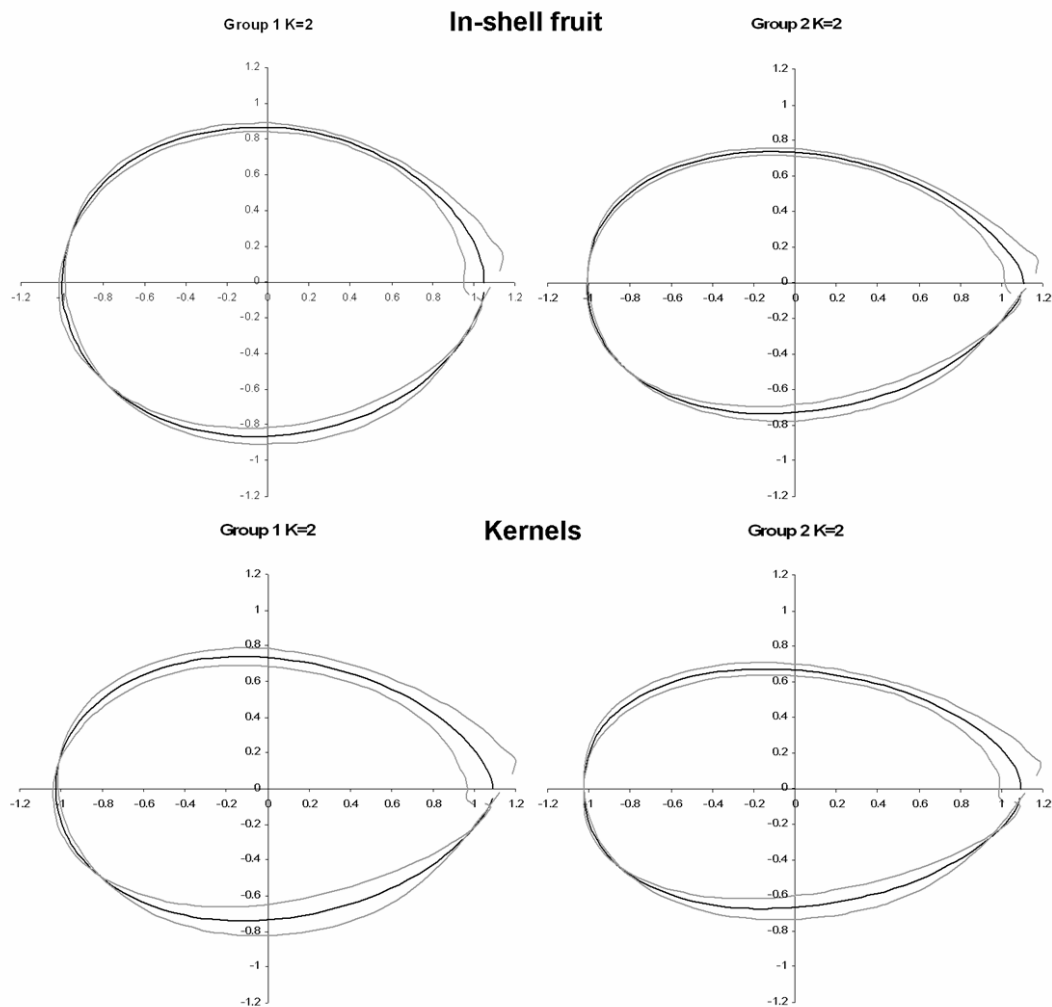
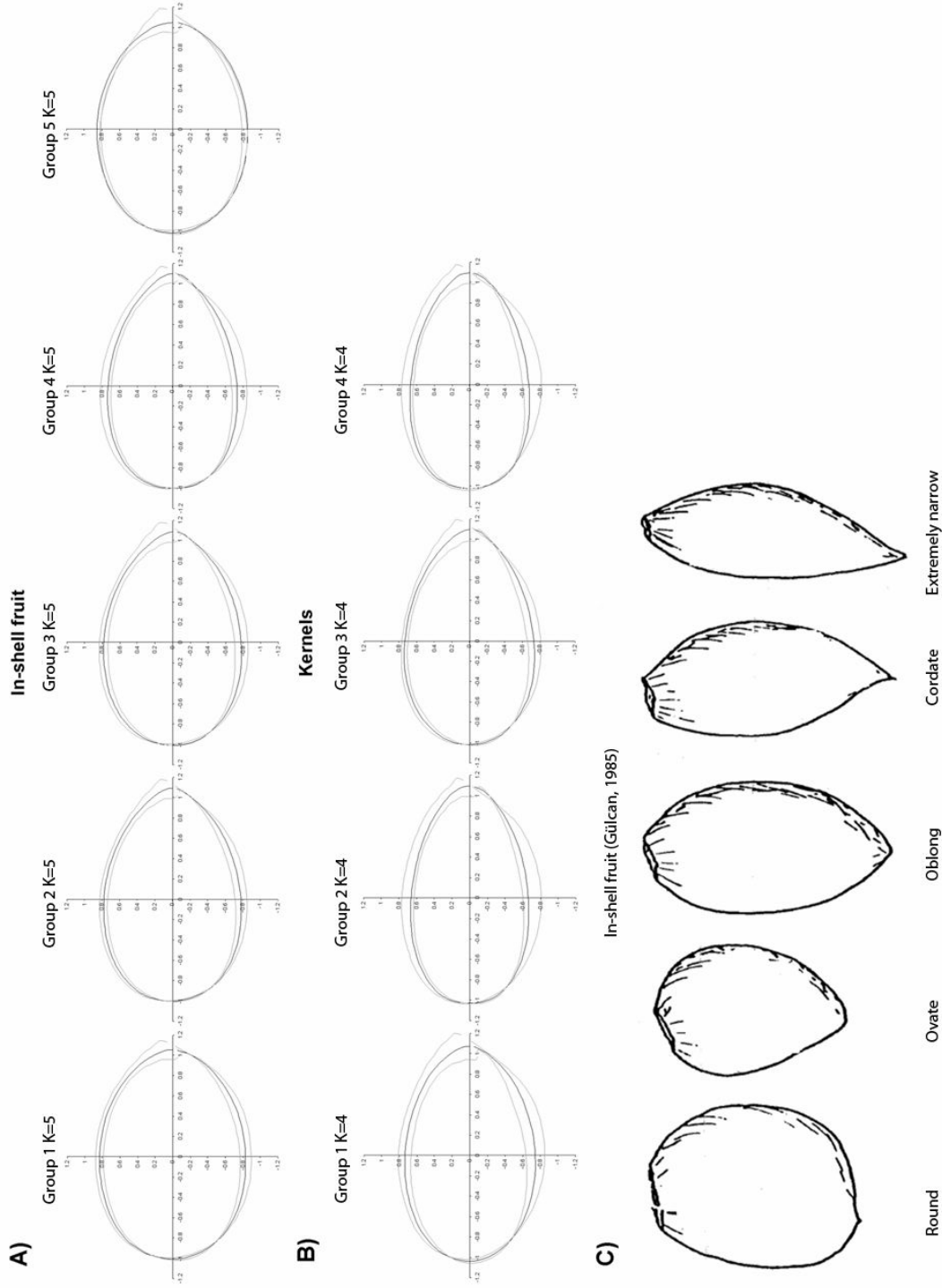


Figure 7 reports the mean configurations and the standard deviation ranges extracted with $k=5$ for in-shell fruit and with $k=4$ for kernels. For in-shell fruit: Group 1 (CATD and FALS) has a slightly spheroid shape with a rounded apex and a convex base; Group 2 (BANC, COSI, FILI, GALG, MONT, PEPP, PIDO, SENZ and ZINZ) is characterized by an ellipsoid shape with an acute apex and a convex base; Group 3 (BARL, PIGN and SANT) has a moderately ellipsoid shape with an acute apex and a convex base; Group 4 (CRIS, TENE and TUON) has a marked ellipsoid shape with an acute apex and a convex base; Group 5 (CATC) presents a spheroid shape with rounded apex and a convex base. For the kernels: Group 1 (CATC, CATD and FALS)

is defined by an ellipsoid shape with an acute apex and a convex base; Group 2 (COSI, CRIS, FILI, NONT, PIDO, SENZ and ZINZ) has a marked ellipsoid shape with an acute apex and a truncate base; Group 3 (BANC, BARL, PEPP, PIGN and SANT) presents an ellipsoid shape with an acute apex and a convex base; Group 4 (GALG, TENE and TUON) is characterized by a marked ellipsoid shape with an acute apex and a truncate base; Cultivars of Group 5 for the in-shell fruit clustering belong to Group 1 for the kernels clustering. BANC and PEPP cultivars belong to Group 3 for the in-shell fruit clustering, meanwhile for the kernels clustering belong to Group 3. CRIS belongs to Group 2 for the in-shell fruit clustering, meanwhile for the kernels clustering belongs to Group 4. GALG belongs to Group 4 for the in-shell fruit clustering, meanwhile for the kernels clustering belongs to Group 2.

Fig. 7 Mean outline (black line) and the standard deviation range (grey lines) for in-shell fruit (A) and kernels (B) extracted with $k=5$ for in-shell fruit and with $k=4$ for kernels. Almond from the International Board for Plant Genetic Resources (IBPGR) (Gülcan, 1985) (C).



Classification analysis

PLS-DA was conducted to discriminate models based on the groups of k-means (2 and 4 or 5). Table 2 reports the characteristics of the models performed on EFA coefficients for k=2 (both in-shell fruit and kernels), for k= 5 (in-shell fruit), k=4 (kernels) and on the EFA coefficients of the 18 cv (both in-shell fruit and kernels).

Table 2: Characteristics and principal results of the PLSDA models performed on EFA coefficients with k=2 (both in-shell fruit and kernels) and with k=5 (for in-shell fruit) and with k=4 (for kernels) and on EFA coefficients of the 18 cv (both in-shell fruit and kernels). N is the number of samples. n° units (Y-Block) is the number of units to be discriminated by the PLSDA. n° LV is the number of latent vectors for each model. Random Probability (%) is the probability of random assignment of an individual into a unit.

	EFA coefficients k-means			EFA coefficients 18 cv		
	K=2	K=5	K=4	in-shell fruit	in-shell fruit	kernels
	<i>in-shell fruit</i>	<i>in-shell fruit</i>	<i>kernels</i>	<i>in-shell fruit</i>	<i>in-shell fruit</i>	<i>kernels</i>
N	864	864	850	864	864	850
n° units (Y-block)	2	5	2	5	18	18
n° LV	6	7	3	7	8	9
% Cumulated Variance X-block	24.66	27.52	10.68	27.52	29.04	99.89
Mean Specificity (%)	96.1	84.84	96.45	84.84	77.14	72.53
Mean Sensitivity (%)	96.1	89.88	96.45	89.88	79.49	80.87
Random Probability (%)	50	20	50	20	5.56	5.56
Mean Class. Err. (%)	3.38	12.65	3.54	12.65	21.7	23.3
Mean RMSEC	0.5606	0.354	0.5713	0.354	0.2202	0.2152
Mean % Corr. Class. Model	93.22	67.8	94.84	67.8	38.58	31.36
Mean % Corr. Class. Independent Test	90.23	60.47	89.1	60.47	43.98	30.62

The model on $k=2$ were based on 6 Latent Variables for in-shell fruit and on 3 for kernels. Both, in-shell fruit and kernels models, have a high sensitivity, sensibility, and a low mean classification error. Similarly, the percentage of the model correct classification (93.22% for in-shell fruit and 94.84% for kernels) and of the independent test (90.23% for in-shell fruit and 89.1% for kernels) shown high values. In line with these results the two models with $k=5$ (in-shell fruit) and $k=4$ (kernels) have comparable values of mean sensitivity, specificity, classification error and RMSEC; the percentage of correct classification in the model is slightly higher for kernels (71.05%).

The models performed on EFA coefficients of the 18 cv were based on 8 Latent Variables for in-shell fruit and on 9 LV for kernels. Both, in-shell fruit and kernels models have a quite high sensitivity, sensibility and RMSEC. The percentage of correct classification of the model is 38.58% for in-shell fruit and 31.36% for kernels, while, in the independent test is 43.98% for in-shell fruit and 30.62% for kernels. Considering the random probability (5.56%) it represents a high value of discrimination.

Tables 3 ($k=2$) and Table 4 ($k=5$) report the specimens correctly classified from the PLS-DA (independent test) performed on EFA coefficient belonging to in-shell fruit. Tables 5 ($k=2$) and 6 ($k=4$) report the same results relatively to the kernels.

Table 3: Results of the predicted specimens belonging to in-shell fruit of the independent test of the PLSDA models performed on the EFA coefficients with $k=2$.

	Group 1	Group 2	TOT
Group 1	74	0	74
Group 2	21	120	141

Table 4: Results of the predicted specimens belonging to in-shell fruit of the independent test of the PLSDA models performed on the EFA coefficients with k=5.

	Group 1	Group 2	Group 3	Group 4	Group 5	TOT
Group 1	15	0	0	0	1	16
Group 2	2	26	11	37	2	78
Group 3	3	5	33	0	11	52
Group 4	0	7	0	45	0	52
Group 5	3	0	3	0	11	17

Table 5: Results of the predicted specimens belonging to kernels of the independent test of the PLSDA models performed on the EFA coefficients with k=2.

	Group 1	Group 2	TOT
Group 1	105	23	128
Group 2	0	83	83

Table 6: Results of the predicted specimens belonging to kernels of the independent test of the PLSDA models performed on the EFA coefficients with k=4.

	Group 1	Group 2	Group 3	Group 4	TOT
Group 1	37	10	1	13	61
Group 2	3	25	0	0	28
Group 3	0	0	40	1	41
Group 4	12	6	28	35	81

Finally Tables 6 and 7 report PLS-DA (independent test) performed on the EFA coefficients of the 18 cv for in-shell fruits and kernels respectively. For the in-shell fruit it is possible to observe as some cv are optimally predicted (FALS and TENE), meanwhile others score lower prediction values (BANC, CRIS, MONT, PEPP, PIDO and ZINZ). For kernels, again, only FALS and TENE are optimally predicted.

Table 7: Results of the predicted specimens belonging to in-shell fruit of the independent test of the PLSDA models performed on the EFA coefficients of the 18 cv.

	BANC	BARL	CATC	CATD	COSI	CRIS	FALS	FILI	GALG	MONT	PEPP	PIDO	PIGN	SANT	SENZ	TENE	TUON	ZINZ	TOT
BANC	0	1	1	1	1	1	1	3	0	0	0	0	1	1	1	0	1	0	12
BARL	0	10	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	12
CATC	0	0	4	5	0	0	3	0	0	0	0	0	0	0	0	0	0	0	12
CATD	0	0	3	6	0	0	1	2	0	0	0	0	0	0	0	0	0	0	12
COSI	0	2	0	2	2	0	1	1	1	0	0	0	0	0	0	3	0	0	12
CRIS	0	0	0	0	0	1	0	4	0	0	0	1	0	0	0	5	1	0	12
FALS	0	0	1	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	12
FILI	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	3	0	12
GALG	0	0	0	0	1	0	1	0	6	0	0	1	0	0	0	1	2	0	12
MONT	0	0	0	0	0	1	1	2	1	0	0	1	0	0	1	5	0	0	12
PEPP	0	0	1	0	0	0	3	2	1	0	1	0	0	0	0	1	2	1	12
PIDO	0	0	0	0	0	0	0	4	0	0	0	1	0	0	1	0	6	0	12
PIGN	0	1	1	3	0	0	0	0	0	0	0	0	5	2	0	0	0	0	12
SANT	0	0	0	0	0	0	0	0	0	0	0	0	1	10	1	0	0	0	12
SENZ	0	0	0	0	0	0	0	0	0	0	0	1	0	0	8	1	2	0	12
TENE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	12
TUON	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	2	8	0	12
ZINZ	0	0	0	0	0	0	0	3	0	0	0	0	2	1	2	2	1	1	12

Table 8: Results of the predicted specimens belonging to kernels of the independent test of the PLSDA models performed on the EFA coefficients of the 18 cv.

	BANC	BARL	CATC	CATD	COSI	CRIS	FALS	FILI	GALG	MONT	PEPP	PIDO	PIGN	SANT	SENZ	TENE	TUON	ZINZ	TOT
BANC	0	0	2	0	0	0	3	0	3	1	1	0	0	0	0	2	0	0	12
BARL	0	0	0	1	0	0	0	7	1	0	0	0	1	0	0	0	0	0	10
CATC	0	0	2	8	0	0	1	0	0	0	0	0	0	0	0	0	0	0	11
CATD	0	0	4	3	0	0	2	0	2	0	0	0	1	0	0	0	0	0	12
COSI	0	0	0	0	0	0	1	4	3	1	0	0	1	0	1	1	0	0	12
CRIS	0	0	0	0	0	0	1	3	0	0	1	2	1	0	0	3	0	0	11
FALS	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	12
FILI	0	0	0	0	0	0	0	10	1	0	0	1	0	0	0	0	0	0	12
GALG	0	0	0	0	0	0	1	2	3	0	2	0	0	0	0	3	0	0	11
MONT	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	9	0	0	12
PEPP	0	0	0	0	0	0	0	0	0	0	8	3	0	0	0	1	0	0	12
PIDO	0	0	1	2	0	0	0	0	0	0	1	6	0	0	0	1	0	0	11
PIGN	0	1	0	1	0	0	2	2	0	1	4	0	0	1	0	0	0	0	12
SANT	0	0	0	2	0	0	3	3	0	1	0	1	0	1	0	0	0	1	12
SENZ	0	0	0	0	0	0	1	1	0	1	0	2	0	0	4	2	0	1	12
TENE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	11
TUON	0	0	0	0	0	0	0	4	3	0	0	0	0	0	0	4	1	0	12
ZINZ	0	0	0	0	0	0	1	0	0	0	0	2	1	0	1	5	0	2	12

Automated shape identification is a concrete possibility, being an objective, repeatable, fast and cheap tool to help genetists in the selection process (Costa et al., 2008). Furthermore almond shape and size can be a great discriminator in order to select fruit for different commercial and industrial purposes. Similar shapes and size of fruit allow the industry to work contemporarily with several different cv in the transformation phase, being the instruments calibrated to process a homogenous stock of product. Generally a shape descriptor is important for different reasons, such as a method of selection in postharvest activities. The carpological measurements are often applied as important cultivars descriptors.

This study, for the first time, instrumentally identifies qualitative shape descriptors for almond fruit. This attribute description was possible following the clusters attribution for the different cv (in-shell and kernels). The Elliptic Fourier analysis is applied as an innovative tool for cultivar discrimination and description (Menesatti et al., 2008). Our results indicate that fruit shape differences can be used to discriminate different cv of almond. These results indicate that apart from a semantic description of cv shape, such as the IPGRI (1999) for *Citrus*, a new method to quantify differences in shape was found. An attempt to compare the classification developed by the International Board for Plant Genetic Resources (IBPGR) (Gülcan, 1985) and the in-shell fruit mean outline *per cv* (Fig. 2) was made. Such attribution was found to be complex due to the extremely general description given by the IBPGR. With respect to this last, the proposed method posses the following advantages: i. it can attribute the shape of each singular fruit; ii. It describes the general fruit shape as well as the apex and the base ones; iii. It can quantitatively measure the shape variability among different cv.

This method has not the efficiency to allow the identification of each singular cv, but it classifies cv into groups of shapes. In fact, on the base of the results of PLS-DA performed on the 18 cv, we obtained absolute low percentage of correct classification in the model (38.58% for in-shell fruit and 31.36% for kernels), however these values can be considered high if compared with the random probability equal to 5.6%.

Moreover, such method could also return information based on the conformity parameter (variance, Fig. 2 and 3) being this a primary request of the breeders (Kays, 1999). For example, considering the most homogeneous cv, SANT for in-shell fruit

and TENE for kernels, the variance has respectively values of 0.058 and 0.068 (Figs. 2, 3 and Tabs. 7, 8).

The coupling of EFA analysis with PLS-DA modelling technique, based on classification and prediction models, allows performing a more efficient class attribution of new cultivars. A procedure to estimate the optimal number of clusters to be used in the k-means method was also presented by applying the validation procedures of Ray & Turi (1999) on the EFA coefficients dataset. This algorithm can be used for other image analysis in-field applications.

Conclusion

In this study the shape appeared to be a measurable parameter of great importance as aid to genetists (for cv comparison) as well as for industrial purposes (acceptability following different specific utilizations, fraud prevention). The proposed method allows the implementation of a fully Automated Shape Processing System (ASPS). The Matlab source code for the identification of the number of k clusters was successfully used for almond shape classification representing a useful tool in many different research fields. At k=2 the system have the best performance but with scarce applicative sense, meanwhile with k=4 (kernels) and k=5 (in-shell) there are good classification performances having an operative sense. The image analysis techniques used could be implemented for labels attribution (PDO and PGI) allowing a high speed of analysis, giving the possibility to evaluate the whole stock of product, overpassing the subjectivity of the expert evaluation and increasing its repeatability.

References

- Agunbiade SO & Olanlokun JO (2006) Evaluation of some nutritional characteristics of indian almond (*Prunus amygdalus*) nut. Pakistan Journal of Nutrition 5(4), 316-318.
- Aguzzi J, Costa C, Antonucci F, Company JB, Menesatti P & Sardá F (2009) Influence of diel behaviour in the morphology of decapod natantia. Biological Journal of the Linnean Society, 96, 517-532.
- Beyer M, Hahn R, Peschel S, Harz M & Knoche B (2002) Analyzing fruit shape in sweet cherry (*Prunus avium* L.). Scientia Horticulturae, 96, 139-150.

- Bookstein FL (1991) Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge Univ. Press, New York. 435 pp.
- Brewer MT, Moyseenko JB, Monforte AJ & Van Der Knaap E (2007) Morphological variation in tomato: a comprehensive study of quantitative trait loci controlling fruit shape and development. *Journal of Experimental Botany*, 58(6), 1339-1349.
- Cannon CH & Manos PS (2001) Combining and comparing morphometric shape descriptors with a molecular phylogeny: the case of fruit type evolution in Bornean *Lithocarpus* (Fagaceae). *Systematic Biology*, 50(6), 860-880.
- Chitta R & Narasimha Murty M (2010) Two-level k-means clustering algorithm for k- \square relationship establishment and linear-time classification. *Pattern Recognition* 43(3), 796-804.
- Costa C, Menesatti P, Aguzzi J, D'Andrea S, Antonucci F, Rimatori V, Pallottino P & Mattoccia M (2010) External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*. *Food and Bioprocess Technology* 3(1), 43-48.
- Costa C, Aguzzi J, Menesatti P, Antonucci F, Rimatori V & Mattoccia M (2008) Shape analysis of different populations of clams in relation to their geographical structure. *Journal of Zoology*, 276, 71-80.
- Crampton JS (1995) Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia*, 28, 179-186.
- Currie AJ, Ganeshanandam S, Noiton DA, Garrick D, Shelbourne CJA & Oraguzie N (2000) Quantitative evaluation of apple (*Malus x domestica* Borkh.) fruit shape by principal component analysis and Fourier descriptors. *Euphytica*, 111, 219-227.
- De Giorgio D, Macchia M, Stelluti M & Lanza F (1996) Collection of 205 almond tree cultivars in "seedling rootstocks" field of Istituto Sperimentale Agronomico (BA), descriptive monograph. *Agricoltura Ricerca XVII, ISMEA*, 162, 55-161.
- Doering P & Ludwig J (1990) Shape analysis of otoliths - a tool for indirect ageing of Eel, *Anguilla anguilla* (L.)? *Internationale Revue der gesamten Hydrobiologie und Hydrographie*, 75(6), 737-743.

- Furuta N, Ninomiya S, Takahashi S, Ohmori H & Ukai Y (1995) Quantitative evaluation of soybean (*Glycine max* L., Merr.) leaflet shape by principal component scores based on elliptic Fourier descriptor. *Breeding Science*, 45, 315-320.
- Goto S, Iwata H, Shibano S, Ohya K, Suzuki A & Ogawa H (2005) Fruit shape variation in *Fraxinus mandshurica* var. japonica characterized using elliptic Fourier descriptors and the effect on flight duration. *Ecological Research*, 20, 733-738.
- Gülcan R (1985) Descriptors list for Almond (*Prunus amygdalus*) (Revised). International Board for Plant Genetic Resources (IBPGR), 30 pp.
- Hampson CR & Quamme HA (2000) Use of preference testing to identify tolerance limits for fruit visual attributes in apple breeding. *Horticultural Science*, 35(5), 921-924.
- IPGRI (1999) Descriptors for Citrus. International Plant Genetic Resources Institute, Rome, Italy. ISBN 92-9043-425-2 (URL: <<http://www.cgiar.org/ipgri/>>).
- Hill T & Lewicki P (2007) STATISTICS Methods and Applications. StatSoft, Tulsa, USA.
- Iwata H, Niikura S, Matsuura S, Takano Y & Ukai Y (1998) Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors. *Euphytica*, 102, 143-149.
- Jensen RJ, Ciofani KM & Miramontes LC (2002) Lines, outlines, and landmarks: morphometric analyses of leaves of *Acer rubrum*, *Acer saccharinum* (Aceraceae) and their hybrid. *Taxon*, 51, 475-492.
- Kappel F, Fisher-Fleming R & Hogue EJ (1995) Ideal pear sensory attributes and fruit characteristics. *Horticultural Science* 30(5), 955-972.
- Kays SJ (1999) Preharvest factors affecting appearance. *Postharvest Biology and Technology*, 15, 233-247.
- Kennard RW & Stone LA (1969) Computer aided design of experiments. *Technometrics*, 11, 137-148.
- Kester DE & Ross NW (1996) History. In: Micke, W.C. (Ed.), *Almond Production Manual*, Division of Agriculture and Natural Resources, University of California, 1-2.

- Khanizadeh S (1994) Breeding strawberries for Eastern Central Canada. *Euphytica*, 77, 45-49.
- Maesschalck RD, Estienne F, Verdù-Andres J, Candolfi A, Centner V, Despagne F, Jouan-Rimbaud D, Walczak B, Massart DL, De Jong S, De Noord OE, Puel C & Vandeginste BMG (1999) The Development of Calibration Models for Spectroscopic Data Using Principal Component Regression. *Internet J. Chem.* 2, 19. <http://www.ijc.com/articles/1999v2/19/>.
- Menesatti P, Costa C, Paglia G, Pallottino F, D'Andrea S, Rimatori V & Aguzzi J (2008) Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars. *Biosystem Engineering*, 101(4), 417-424.
- Neto JC, Meyer GE, Jones DD & Samal AK (2006) Plant species identification using Elliptic Fourier leaf shape analysis. *Computer and Electronic in Agriculture*, 50(2), 121-134.
- Ohsawa R, Tsutsumi T, Uehara H, Namai H & Ninomiya S (1998) Quantitative evaluation of common buckwheat (*Fagopyrum esculentum* Moench) kernel shape by elliptic Fourier descriptor. *Euphytica*, 101, 175-183.
- Pallottino F, Menesatti C, Costa C, Paglia G, De Salvador FR & Lolletti D (2010) Image analysis techniques for automated hazelnut peeling determination. *Food and Bioprocess Technology*, 3(1), 155-159.
- Ray S & Turi RH (1999) Determination of number of clusters in k-means clustering and application in colour image segmentation. *Proc of the 4th International Conference on Advances in Pattern Recognition and Digital Techniques (ICAPRDT'99)*, Calcutta, India, 137-143.
- Rohlf FJ (2006) TpsDig Ver. 2.10, digitalized landmarks and outlines. Dept. of Ecology and Evolution, State Univ. of New York at Stony Brook.
- Rohlf FJ & Archie JW (1984) A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Systematic Zoology*, 3, 302-317.
- Scott D (1998) Horticulture fashion food conference. *Orchardist New Zealand*, Feb, 37-50.
- Sabatier R, Vivein M & Amenta P (2003) Two approaches for Discriminant Partial Least Square. In: Schader, M., Gaul, W., Vichi, M., (Eds.), *Between data science and applied data analysis*. Springer Verlag, Berlin, 100-108.

- Slice DE (1998) Morpheus et al.: software fore morphometric research. Revision 01-30-98-Beta. Department of ecology and evolution, State University of New York, Stony Brook, New York.
- Ward JH (1963) Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58, 236-244.
- White RJ, Prentice HC & Verwijst T (1988) Automated image acquisition and morphometric description. Canadian Journal of Botany, 66, 450-459.
- Yoshioka Y, Iwata H, Ohsawa R & Ninomiya S (2004) Analysis of petal shape variation of *Primula sieboldii* by elliptic Fourier descriptors and principal component analysis. Annals of Botany, 94, 1-8.

Appendix 1

Script to determine the Right Number of Clusters in k-Means clustering (request Matlab V7.0 and Statistical Toolbox)

```
% validity = average(intra-class distance) / min(inter-class distance)
% average(intra-class distance), mean distance of each object from its centroid
% min(inter-class distance), minimum distance between the cluster centroids
```

```
clear all
```

```
close all
```

```
DirIn="";
```

```
DirOut="";
```

```
files=dir(strcat(DirIn,'*.mat'));
```

```
for ii=1:size(files, 1);
```

```
fname=files(ii).name
```

```

load (strcat(DirIn, fname));

X=x;          %Rows of X correspond to points, columns correspond to variables.

DistFinale2=[];

for k=2:25          %size(X,1)-1
k
DistFinale=[];
for ww=1:1000;

[IDX,C, sumd] = kmeans(X, k, 'emptyaction', 'singleton');

Classi=unique(IDX);
NClasse=[];
for cx=1:numel(Classi)
NClasse1= numel(find(IDX==Classi(cx,1)));
NClasse=[NClasse;NClasse1];
end
ObsMatrix=[Classi, NClasse];

intra=sum(sumd)/size(X,1); %originale

interDist=[];
for c1=1:size(C,1)-1          %centroids
for c2=c1+1:size(C,1)
cd1=C(c1,:);
cd2=C(c2,:);

% Euclidean distance of xtest from each k
interDist=[interDist;sqrt(sum((cd1-cd2).^2))];

```

```

end
end

inter=min(interDist);
validity=intra/inter;

DistFinale=[DistFinale; intra, inter];

end

DistFinale2=[DistFinale2; k, mean(DistFinale(:,1)), mean(DistFinale(:,2)),
mean(DistFinale(:,1))/mean(DistFinale(:,2))];

end          % for k

figure, plot (DistFinale2(:,1), DistFinale2(:,4)), title (fname)

save (strcat(DirOut, 'BestKmeans-', fname), 'DistFinale2');

end

```

3.7 Discrimination of Tarocco sweet orange [*Citrus sinensis* (L.) Osbeck] varieties using opto-electronic elliptic Fourier based analysis of fruit shape

FULL REFERENCE OF THE PUBLICATION

Costa C, Menesatti P, Paglia G, Pallottino F, Aguzzi J, Rimatori V, Russo G, Recupero S, Reforgiato Recupero G, 2009. Discrimination of Tarocco sweet orange [*Citrus sinensis* (L.) Osbeck] varieties using opto-electronic elliptic Fourier based analysis of fruit shape. POSTHARVEST BIOLOGY AND TECHNOLOGY, 54: 38-47.

Introduction

Tarocco blood orange fruit have some unique characteristics such as better taste and relatively high concentrations of vitamin C and anthocyanins (Rapisarda and Russo, 2000). Recently isolated selections of Tarocco allow for a long harvest season (from December to May) (Reforgiato Recupero and Russo, 2002). These newly selected genotypes also have variable fruit quality traits in particular rind and flesh anthocyanin contents, fruit shape and size, peelability and shelf-life. Moreover, the postharvest chain, including long transport transit times, storage and selection processes can cause fruit quality loss in terms of appearance, the latter being a primary criterion for consumer preference (Kays, 1999).

Appearance, including fruit shape, is particularly used as a selection criterion throughout the supply and consumer chain (Kays, 1999). Fruit shape discrimination is still dependent upon subjective criteria because of the absence of suitable numerical methods for quantifying discrimination and morphological differences. Indeed, descriptors such as IPGRI (1999) and *Citrus industry* (Hodgson, 1967) classify citrus shape into standard classes allowing a certain amount of shape variability which is not measurable. The development of an objective and repeatable method would increase the efficacy of a system that could be used both for consumer acceptability analysis and as an automated quality assessment tool.

To date, different methods for quantitative shape evaluation have been suggested to distinguish genotypes. The most common methodology is based on Elliptic Fourier descriptors, which has been successfully used to assess differences among leaves (White et al., 1988; Jensen et al., 2002; Neto et al., 2006), leaflets (Furuta et al., 1995), kernels (Ohsawa et al., 1998), roots (Iwata et al., 1998), flowers (Yoshioka et al., 2004), and fruit (Currie et al., 2000; Goto et al., 2005). This method mathematically describes the entire shape of an object by transforming the coordinate information concerning its contour into Fourier coefficients (Rohlf and Archie, 1984). Few studies, however, have considered the analysis of fruit shape by means of optoelectronic techniques. Menesatti et al. (2008) suggested a method based on the combined use of morphometry and multivariate statistics to discriminate among different Italian hazelnut genotypes. Most of the studies available use a PCA approach to observe shape differences between groups at different levels (cultivar or species). The present study aimed to create a classification and modelling system able to attribute existing shape classes to new genotypes to be coupled with industrial sorting equipment. For this kind of approach, supervised multivariate modelling such as PLSDA (Menesatti et al., 2008) or SIMCA (Aguzzi et al., 2009) was suggested. The aim of this work, therefore, was to develop a new method for shape discrimination through image analysis performed on sweet orange fruit lateral shapes by means of Elliptic Fourier Analysis (EFA). The method was tested on 17 different Tarocco genotypes in order to assess the ability of the system to perceive minimum expected differences between similar genotypes. Results were screened by multivariate classification techniques (i.e. PLSDA) in order to assemble groups corresponding to different fruit shape. The output of optoelectronic measurements combined with EFA on processed fruit shapes provided the basis for future experimental protocols and validated the analytic capacity of this method. Due to the methodological focus of this paper, the main intention was to evaluate an innovative, combined approach to the study of the fruit shape based on optoelectronic techniques including image analysis, morphometry and multivariate statistical analysis.

Materials and methods

Biological samples

A total of 929 fruit belonging to 17 genotypes (Table 1) were analyzed. All fruit were provided by the CRA-ACM experimental farm Palazzelli (Lentini-Siracusa, Sicily) in the 2007 harvest season.

Table 1. Variety, code and quantity of fruit (N) used for the morphometric analysis (*nucellar selection; **old line, virus-free by micrografting). Shape classification by IPGRI (1999) and Citrus Industry (Hodgson, 1967).

Label	Cultivar or Selection	N	IPGRI	Citrus industry
S-N1	Scirè D 2062*	52	Fruit form spheroid with base truncate	Fruit form globose with base truncate
S-N2	Scirè D 2071*	51	Fruit form spheroid with base truncate	Fruit form globose with base truncate
S-V	Scirè VCR**	60	Fruit form spheroid with base truncate	Fruit form globose with base truncate
G-N	Gallo C 898*	38	Fruit form spheroid-ellipsoid with base small concave collared	Fruit form obovoid with base low collared and depressed
G-V	Gallo VCR**	44	Fruit form spheroid-ellipsoid with base small concave collared	Fruit form obovoid with base low collared and depressed
TDV	TDV-P2 *	68	Fruit form spheroid with base small concave collared	Fruit form obovoid with base low collared and depressed
TAP	Tapi*	32	Fruit form spheroid with base small concave collared	Fruit form obovoid with base low collared and depressed
57	57-1E-1*	27	Fruit form spheroid with base small	Fruit form obovoid with base low collared

			concave collared	and depressed
1E	1E VCR**	28	Fruit form spheroid with base small concave collared	Fruit form obovoid with base low collared and depressed
L-N1	Lempso C 5787*	78	Fruit form spheroid with base small concave collared	Fruit form obovoid with base low collared and depressed
L-N2	Lempso C*	75	Fruit form spheroid with base small concave collared	Fruit form obovoid with base low collared and depressed
L-V	Lempso VCR**	79	Fruit form spheroid with base small concave collared	Fruit form obovoid with base low collared and depressed
SC	Sciara C 1882*	71	Fruit form spheroid- ellipsoid with base concave collared	Fruit form obovoid with base high collared and depressed
M	Meli C 8158*	57	Fruit form ellipsoid with base small concave collared	Fruit form ovoid with base low collared and depressed
TRI	Trincali C 3169*	26	Fruit form obloid with base truncate	Fruit form oblate with base flat
GIA	Giarretta C 5681*	74	Fruit form spheroid with base small concave collared	Fruit form obovoid with base low collared and depressed
F	Fondaconovo VCR**	69	Fruit form spheroid- ellipsoid with base small concave collared	Fruit form obovoid with base low collared and depressed

Table 1 reports the shape classification for each genotype based on the IPGRI descriptors (1999) and Citrus Industry (Hodgson, 1967) classification. According to IPGRI (1999), 6 different fruit shapes can be attributed to different genotypes as

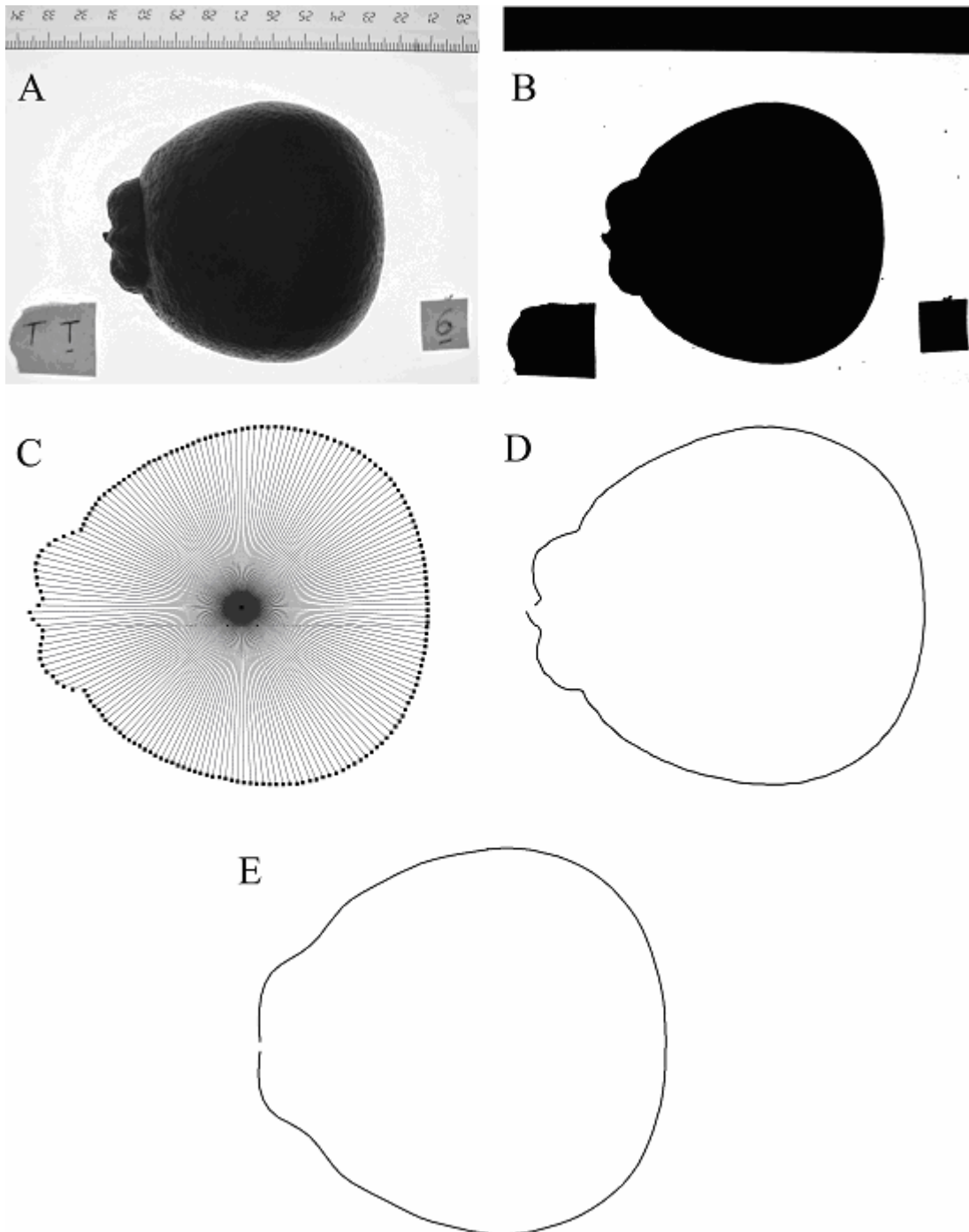
follows: (i) spheroid with truncated base; (ii) spheroid-ellipsoid with base small concave collared; (iii) spheroid with base small concave collared; (iv) spheroid-ellipsoid with base concave collared; (v) ellipsoid with base small concave collared; (vi) obloid with base truncate. In contrast, according to the Citrus Industry descriptors (Hodgson, 1967), 5 different fruit shapes can be identified for the same genotypes: (i) globose with base truncate; (ii) obloid with base low collared and depressed; (iii) obovoid with base high collared and depressed; (iv) ovoid with base low collared and depressed; (v) oblate with base flat. We did not consider the IPGRI and Citrus Industry apex shape classifications due to their uniformity (rounded) for all the genotypes studied.

Seventeen genotypes, belonging to 11 different nucellar or VCR selections (Sciré, Gallo, TDV-P2, Tapi, 57-1E, Lempso, Sciara, Meli, Trincoli, Giarretta and Fondaconuovo) were analysed.

Digital image analysis

A digital high resolution photograph per fruit was acquired with a digital camera Nikon (D1X) equipped with a *Nikon AF lens (24mm f/2.8D)*. Each fruit was laterally placed on an illuminated dashboard, increasing the contrast, with the stem directed to the right side. A metric reference mark was placed by the side of each fruit in order to record its size (Fig. 1A). Each image was binarized (Fig. 1B); a total number of 180 points (x,y) equally angularly spaced from the centroid (one point every 2°) were digitized along the outline (Fig. 1C) utilizing the software TPSdig2 (Rohlf, 2006). The coordinates were aligned by Generalised Procrustes Analysis, a procedure that consists of three steps: translation of the coordinates point to a common centroid located at the origin (0,0) of a reference system of coordinates; the scaling of each outline to the unitary centroid size; and finally, the rotation of coordinates to minimize the sum of square distances between correspondent landmarks. The 180° aligned coordinates were treated as outline data (Fig. 1D; Menesatti et al., 2008). The experimental setup is shown in Fig. 1.

Figure 1: Experimental setup and digital video-image processing of a single fruit (SC genotype in the example): A. original image with metric reference and labels; B. binarized image; C. 180° points (x,y) equally and angularly spaced (one point every 2°) from the centroid along the outline; D. resultant outline; E. outline reconstructed with the elliptic Fourier analysis (i.e. 14 coefficients).



Elliptic Fourier analysis

Given that the outline for any fruit shape can be approximated by a polygon of x, y -coordinates, Elliptic Fourier Analysis (EFA) analysis was used to reduce the dimensionality of the data set and eliminate redundant information (as determined by high correlation between adjacent coordinates in the raw data set). EFA is based on the separate Fourier decompositions of the incremental changes of the x - and y -coordinates as functions of the cumulative chordal length of the outline polygon. EFA yields the spectrum of the fruit shape's closed contour in terms of harmonically related trigonometric curves. For each harmonic equation, two Fourier coefficients are computed for both the x - and y -projections; thus the total number of coefficients is $4n$, where n is the number of harmonics fitted to the outline (Crampton, 1995). The total number of harmonics that can be computed for any outline is equal to half of the total number of outline coordinates (i.e. the 'Nyquist frequency'). However, as with all Fourier methods much of the shape information can be summarized by the first few harmonics. Furthermore, the following harmonics are orthogonal (i.e. independent) and hence some of coefficients can be excluded from the analysis without affecting the others (Crampton, 1995). The number of harmonics required was estimated from the average Fourier power (or variance) spectrum. The Fourier power of a harmonic is proportional to its amplitude and provides a measure of the amount of 'shape information' described by that harmonic. For the n^{th} harmonic, the Fourier power is given by the expression:

$$\text{Fourier power} = (A_n^2 + B_n^2 + C_n^2 + D_n^2)/2$$

where A , B , C , and D are the Fourier coefficients of the n^{th} harmonic. The Fourier series was truncated at the value k , at which the average cumulative power was 99.99% of the total averaged power. For any outline, the total power was calculated as the sum, from 1 to k , of individual harmonic powers, where k is equal to the Nyquist frequency (Crampton, 1995). The harmonic coefficients describe the size, shape and orientation of each harmonic ellipse and form the input for multivariate statistics.

Mean outline and standard deviation (SD) ranges were graphically extracted for each genotype. Pairwise-comparisons (post-hoc) (i.e., Hotelling's tests PAST software; Hammer et al., 2001) between the harmonic coefficients of the fruit of each group were performed *via* MANOVA.

Clustering

Mean coefficients of all harmonic equations for each genotype were clusterized by an unsupervised k-means clustering technique into 6 and 5 groups (k-means-6, k-means-5) in order to be compared with IPGRI (1999) and Citrus Industry (Hodgson, 1967) classifications.

K-means (MacQueen, 1967) clustering classifies a multivariate dataset through a certain number of clusters (k) designated a priori. K-means uses an iterative algorithm that minimizes the sum of distances from each object to its cluster centroid, over all clusters. This algorithm moves objects between clusters until the sum cannot be decreased further. The result is a set of clusters that are as compact and well-separated as possible.

Moreover, another clustering was performed on the basis of the optimal number of clusters (k-means-best) obtained by the procedure of Ray and Turi (1999). This validity measure was based on the intra- and inter-cluster distance which simplifies discriminate element membership based on their reciprocal positioning. The basic procedure involves producing all clusters from 2 up to Kmax clusters, where Kmax represents an upper limit on the number of clusters. Then the validity measure is calculated to determine which is the best clustering by finding the minimum value for this measure. This procedure consists in minimizing the ratio between intra-cluster and inter-cluster distances. At the minimum value of this ratio there is the right number of k clusters. This procedure was replicated 1000 times on the whole EFA coefficients dataset, in order to obtain a more stable mean value.

Classification analysis

The results of fruit shape clustering according to the 5 different selected methods (k-means-best, IPGRI, k-means-6, Citrus Industry, and k-means-5) were compared together *via* the multivariate supervised method, Partial Least Squares Discriminant

Analysis (PLSDA), in order to: (i) test their different efficiency classification, (ii) create shape classification models of reference applicable in future works for new or already existing genotypes, and (iii) classify single fruit on the base of its appearance quality.

To test the different classification efficiency, the entire dataset were divided in 2 subgroups. Dataset 1 as the modelling group was composed of 7 genotypes (SN-1, G-V, TRI, TAP, S-N2, SC, and M) representing all the possible shapes classified by both IPGRI (1999) and Citrus Industry (Hodgson, 1967). This subgroup was used to create a model that considers each kind of shape class. Dataset 2, the testing group, was made from the remaining 10 genotypes (TDV, 57, GIA, L-N1, S-V, F, G-N, L-V, 1E, and L-N2). This was used to test the efficiency of the created models.

PLSDA expresses the class membership of the statistical units (Sjöström *et al.*, 1986; Sabatier *et al.*, 2003) of the response variables (Y-block) in a categorical fashion.

PLSDA was performed on the dataset 1 to create the models tested on dataset 2. Y-block classes were replaced by a set of dummy variables describing the categories (for methodological details see Costa *et al.*, 2008a,b; Menesatti *et al.*, 2008). The model includes a calibration phase and a validation phase, for both of which the percentages of correct classification were calculated. The prediction ability in the test phase depends also on the number of the Latent Variables used in the model.

Therefore, we chose the models for which the number of Latent Variables resulted in a higher percentage of correct classification in the relative independent test. PLSDA calculates a 'prediction probability' and a classification threshold for each class modelled. These were calculated using a Bayesian method (i.e. `plsdthres` function of Matlab 7.1. PLStoolbox) that basically takes the predicted *y*-values from the PLSDA model, fits a normal distribution onto them and then uses the results of that fitting to calculate the probability of observation for a given *y*-value. The PLSDA analysis provides the percentage of correct classifications and the loadings per species on each Latent Variable. This analysis also expresses the statistical parameters indicating the modelling efficiency in terms of sensitivity and specificity parameters where the first represents the percentage of the species of a category accepted by the class model and the second the percentage of the species of the categories different from the modelled one, rejected by the class model. Such analysis was performed using Matlab (rel. 7.1;

PLSToolbox Eigenvector 4.0) on the shape variables (X-block; harmonic coefficients). The X-block values were pre-processed with an 'abs' procedure that took into account the absolute value of the data.

Results

The EFA of the lateral fruit profile from different Tarocco genotypes showed a mean cumulative power of 99.99% at the 14th harmonic fitting, indicating that fruit shape is efficiently described by the coefficients of the first 14 harmonics. These coefficients were successively considered for the analysis.

For each genotype, the mean configuration (black line), the standard deviation range with 14 harmonics (grey line) and the values of variance are graphically reported in Fig. 2.

The following different shape categories were observed: spheroid for L-V and 1E; obloid for TRI; ellipsoid with base concave collared for G-N and M; or, with an evident neck for SC. A greater genotype variance is also visible in pictures of G-N, GIA, F, and S-N1. Greater variance, indicating a lower degree of shape conformity, was found for genotypes G-N (0.215), SC (0.163), and F (0.125); less variance, indicating a higher degree of shape conformity, was found for genotypes 1E (0.053), 57 (0.058), and TRI (0.063). In order to observe significant differences between strains ($p < 0.05$), a MANOVA analysis was computed on the EFA coefficients (see Table 2). Bonferroni corrected values (i.e., multiplied by the number of pairwise comparisons) are given below the diagonal. This analysis indicated that the genotypes TRI, SC, and M were not significantly different in comparison to G-V, TAP, G-N, 57, and G-N. In contrast, G-N was significantly different only from L-V.

Figure 2: Mean outline (black line) and the standard deviation range (grey line) for each variety extracted with 14 harmonics by Elliptic Fourier analysis. On the top-right corner for each configuration are reported values of variance for each variety (see Table 1 for variety codes). The stem is directed to the right side in all schemes where a discontinuity in fruit contour appears.

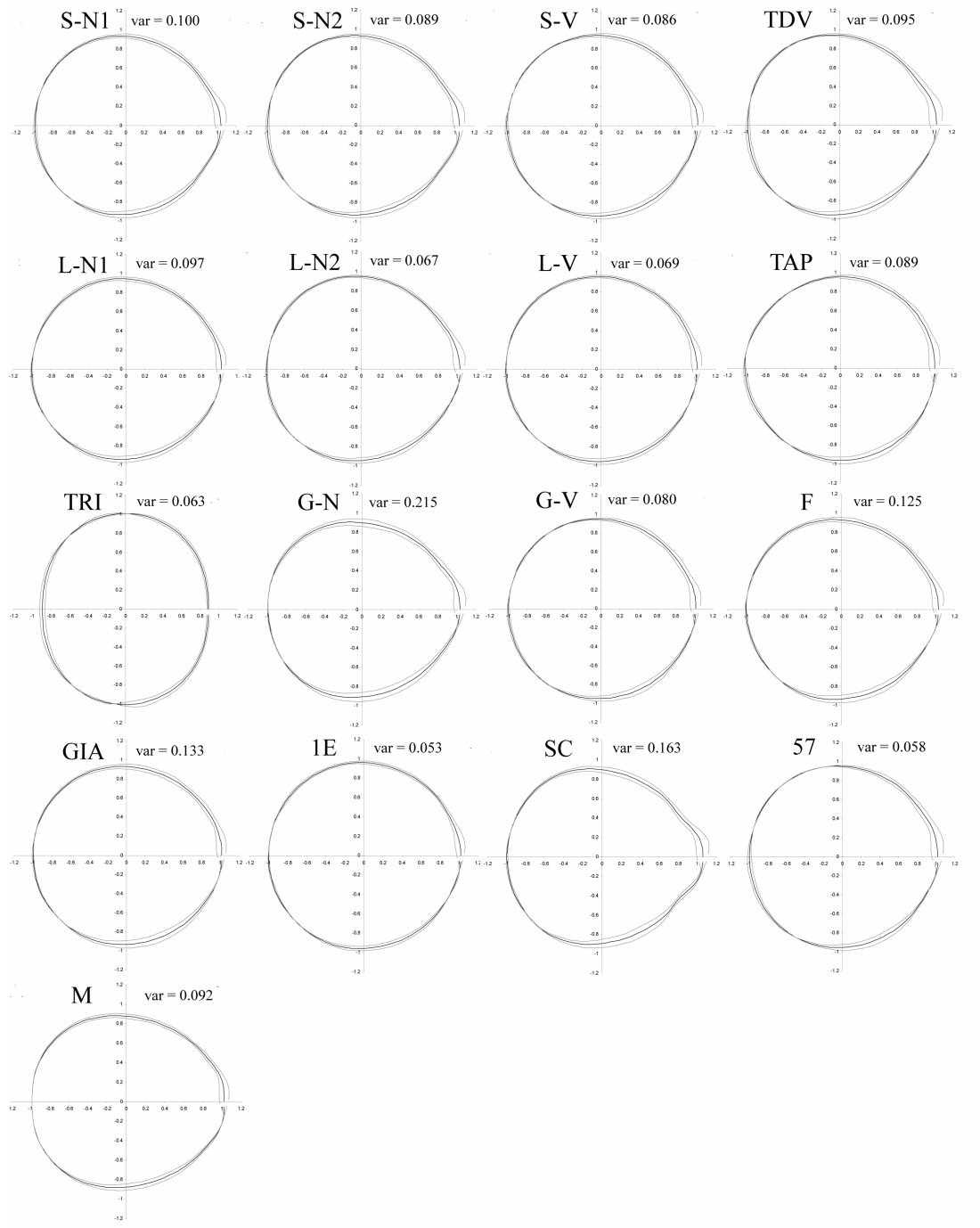


Table 2: Results of the MANOVA analysis on EFA coefficients. Hotelling's pairwise comparisons; Bonferroni corrected. In grey, significant ($p < 0.05$) differences. See Table 1 for variety codes.

	S-N1	G-V	TRI	TDV	57	GIA	L-N1	TAP	S-V	S-N2	F	G-N	L-V	1E	SC	L-N2	M
S-N1																	
G-V	50.742																
TRI	0.000	0.745															
TDV	1.297	2.676	0.000														
57	7.845	0.320	0.000	59.062													
GIA	0.007	15.789	0.000	0.002	0.938												
L-N1	0.222	11.103	0.000	0.002	0.119	62.229											
TAP	0.219	14.098	35.529	1.250	14.460	0.705	3.969										
S-V	56.593	51.072	0.000	0.012	0.002	0.169	0.168	1.344									
S-N2	100.557	3.187	0.000	7.842	0.163	0.001	0.000	0.031	9.481								
F	20.852	27.076	0.000	0.010	4.143	0.514	4.092	6.221	1.398	0.000							
G-N	1.181	0.387	0.572	0.452	67.705	8.093	0.176	16.550	2.621	1.366	0.962						
L-V	1.124	45.184	0.000	0.052	8.826	0.000	0.849	22.541	0.679	0.001	0.204	0.012					
1E	90.416	106.799	0.000	7.187	109.686	15.288	15.969	93.671	39.911	2.140	19.173	0.478	134.931				
SC	0.013	0.001	0.000	0.005	0.073	0.000	0.000	0.000	0.000	0.000	0.000	0.330	0.000	0.023			
L-N2	29.740	79.539	0.000	2.855	14.448	0.000	0.715	0.013	0.147	10.901	1.731	0.055	0.112	73.918	0.000		
M	0.000	0.000	0.000	0.000	0.115	0.000	0.000	0.000	0.000	0.000	0.000	11.524	0.000	0.000	0.000	0.000	0.000

The mean configuration (black line), standard deviation range with 14 harmonics (grey line) for the IPGRI shape classification and Citrus Industry classifications are respectively reported in Figs 3 and 4. Both shape classifications include in the same cluster the 3 genotypes SC, M and TRI. Similarly group A (for both) is represented by the same 3 Scirè genotypes (S-N1, S-N2, and S-V). The group B for the Citrus Industry classification (Fig. 4) corresponded to two similar shaped groups in the IPGRI classification (Fig. 3).

Figure 3: Mean outline (black line) and the standard deviation range (grey line) for fruit shapes classified by IPGRI (1999) on the present dataset (see Table 1 for genotypes attribution); A. spheroid with base truncate; B. spheroid with base small concave collared; C. spheroid-ellipsoid with base small concave collared; D. spheroid-ellipsoid with base concave collared; E. ellipsoid with base small concave collared; F. obloid with base truncate.

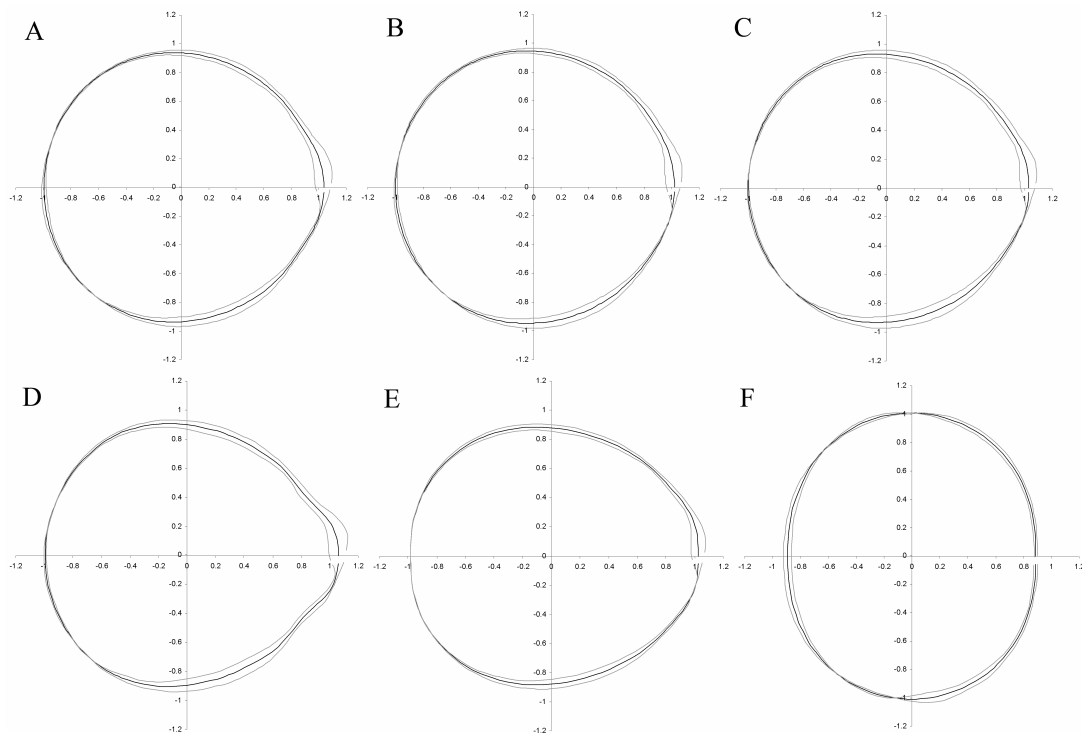
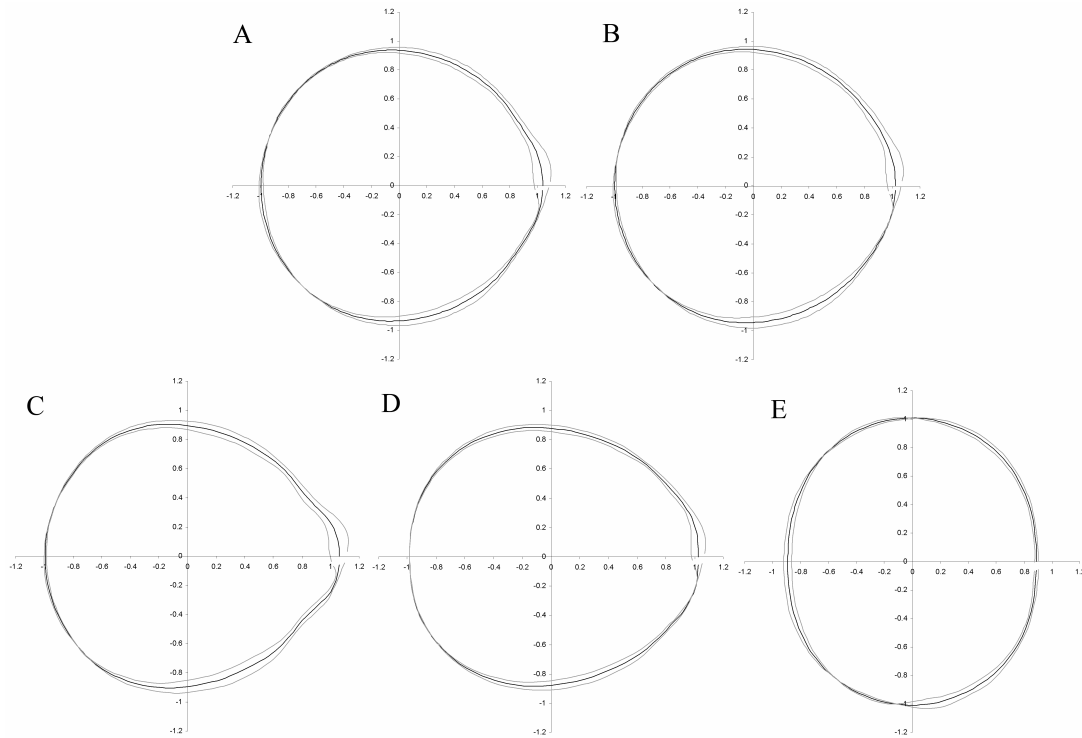
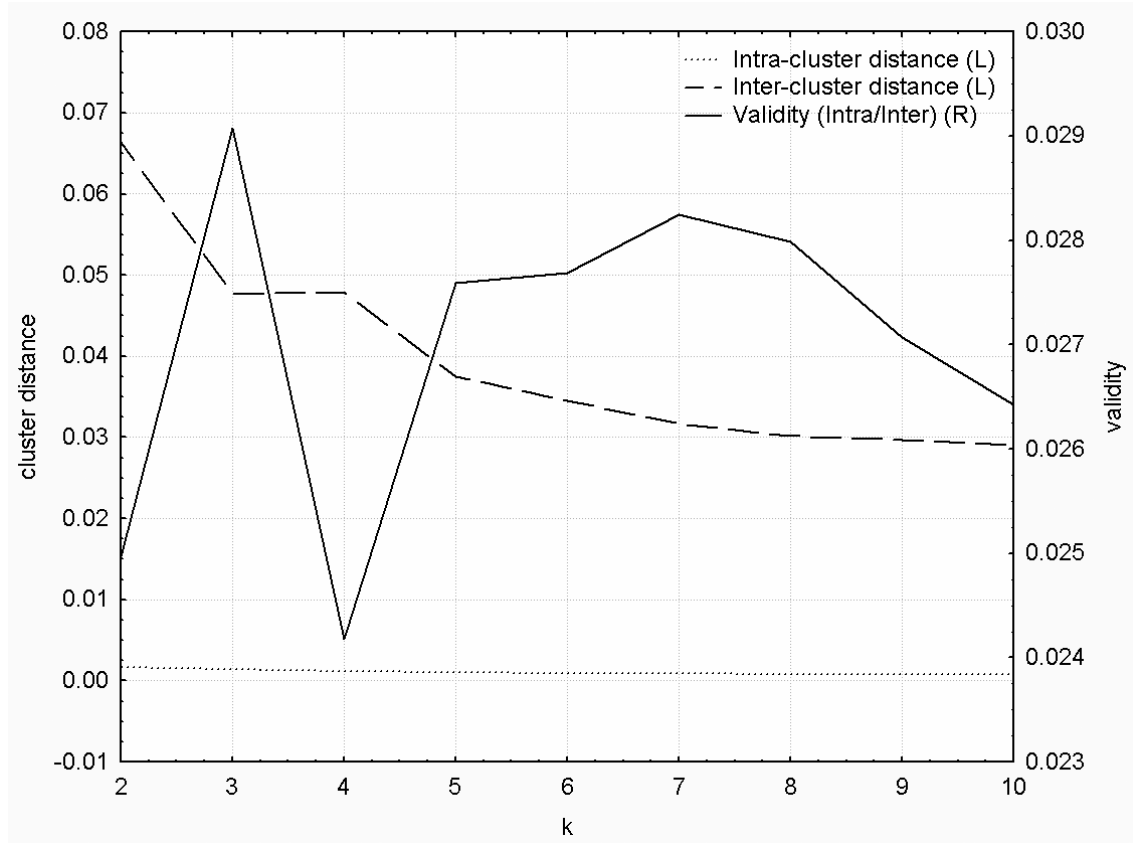


Figure 4: Mean outline (black line) and the standard deviation range (grey line) for fruit shapes classified by Citrus Industry (Hodgson, 1967) on the present dataset (see Table 1 for genotypes attribution); A. globose with base truncate; B. obovoid with base low collared and depressed; C. obovoid with base high collared and depressed; D. ovoid with base low collared and depressed; E. oblate with base flat.



The results of the validation procedure to find the best number of k-clusters are reported in Fig. 5. In this graphic output, the best number of clusters to be used in this dataset (i.e. the minimum of the validity curve) was 4. Values of k=4 indicating the mean configurations (black line) and the standard deviation ranges (grey line) are reported in Fig. 5. Group A genotypes (F, GIA, L-N2, S-N1, S-N2, S-V, and TDV) have a spheroid shape. Group B genotypes (1E, 57, G-V, L-N1, L-V, and TAP) have a spheroid/obovoid shape with base low collared. Group C genotypes (G-N, M, and SC) have an ellipsoid shape with base concave collared. The genotype TRI is classified alone in a distinct group (i.e. D) having an obloid shape with a flat base.

Figure 5: Results of the procedure of validation to find the best number of k clusters.



PLSDA was conducted to discriminate models based on the groups of k-means-4 (i.e. the best number of k clusters), IPGRI, Citrus Industry, k-means-5, k-means-6, and to compare the classification efficiency between 5-groups models (Citrus Industry and k-means-5) and 6-groups models (IPGRI and k-means-6). Table 3 reports the characteristics of the resulting models, all based on 9 Latent Variables.

Table 3: Characteristics and principal results of the PLSDA models. N is the number of samples. n° units (Y-Block) is the number of units to be discriminated by the PLSDA. n° LV is the number of latent vectors for each model. Random Probability (%) is the probability of random assignment of an individual into a unit.

	k-means 4	Citrus industry	k-means 5	IPGRI	k-means 6
N model (dataste 1)	333	333	333	333	333
N test (dataste 2)	596	596	596	596	596
cv model	7	7	7	7	7
cv test	10	10	10	10	10
n° units (Y-block)	4	5	5	6	6
n° LV	9	9	9	9	9
Cumulated Variance X-block (%)	99.99	99.99	99.99	99.99	99.99
Cumulated Variance Y-block (%)	57.64	50.72	53.77	46.59	45.59
Mean Sensitivity (%)	0.864	85.0	83.8	85.1	85.9
Mean Specificity (%)	0.839	82.9	81.5	81.6	77.8
Mean classification error	0.149	0.160	0.174	0.167	0.182
Mean RMSEC	0.314	0.305	0.296	0.292	0.295
Random Probability (%)	25	20	20	16.7	16.7
Corr. Class. Model (%) on dataset 1	74.5	69.7	70.3	64.3	57.1
n° misclassified model	85	101	99	119	143
Corr. Class. Indep. Test (%)on dataset 2	58.2	26.0	46.6	9.2	27.0
n° misclassified test	249	441	318	541	435

The model k-means-4 had the best scoring with the procedure of validation the higher sensitivity, RMSEC, and finally the lowest mean classification error (Fig. 5). This is confirmed by the high percentage of correct classifications for both the model

(74.5%) and the independent test (58.2%). Looking at the two models with 5 groups (Citrus Industry and k-means-5), mean sensitivity, specificity classification error, and RMSEC produced comparable values; the percentage of correct classification in the model dataset results were slightly higher for k-means-based model while the percentage of correct classification in the independent test dataset was higher in the k-means-5 model (respectively, 46.6% vs 26.0% compared to the random probability of classification equals to 20%). Comparing the two models with 6 groups (IPGRI and k-means-6), mean specificity appeared higher for the IPGRI model (81.6% vs 77.8%); the other model efficiency indicators (sensitivity, classification error, and RMSEC) were all similar. The percentage of correct classification in the model dataset was higher for IPGRI with respect to the k-means-based model (64.3% vs 57.1%), but the percentage of correct classification in the independent test dataset was comparatively higher in the k-means-6 model (27.0% vs 9.2% with respect to the random probability of classification equals to 16.4%).

Discussion

In this study, the screening of Tarocco orange genotypes was successfully carried out applying a morphological analysis of their fruit shape with Elliptic Fourier Analysis (i.e. EFA) in combination with unsupervised clustering methods (k-means) and multivariate modelling techniques Partial Least Squared Analysis (i.e. PSLDA). Our approach proved its efficacy in comparison with traditional and more subjective methods currently in use, mostly based on human subjective evaluation (i.e. IPGRI and Citrus Industry). Citrus descriptors from IPGRI (1999) and Citrus Industry (Hodgson, 1967) define fruit shape on the basis of its whole contour, collar and apex. In several cases, according to these classification systems, it is not possible to effectively discriminate some genotypes which cluster together by using k-means analysis.

Fruit shape is a primary selection criterion for consumers and this qualitative attribute of cultivars is chiefly used as discriminator for scientific and consumer purposes (Kays, 1999; Menesatti et al., 2008). Selection criteria used for fruit choice in marketing are partly based on evaluation of the form. Citrus fruit with shape defects are often rejected or placed in a lower value category (see Organizzazione di

Produttori Agridea, 2009). However, fruit shape variability, within the same or among different cultivars, proposes a challenge to operate quality discrimination at key stages in the supply chain and postharvest operations. In particular, cultivar selection is currently still carried out by visual inspection. This is also the case of Tarocco oranges where not only fruit shape but also fruit skin, rind colours, and maturation timing are all considered as crucial characteristics defining fruit quality. Several statistical approaches can be used to discriminate biological objects on the base of the morphological study of their contours. These studies are based on the fitting of some type of circular function to the outline itself and on the consequent computation of curve parameters for subsequent analysis (Rohlf, 1996). Among these methods, the elliptical Fourier decomposition in this study provided better results (Table 3). This classification system has proved to be a powerful descriptor in different taxonomic groups (e.g., Rohlf and Archie, 1984; Lestrel, 1997; Loy et al., 2000; Sheets et al., 2006; Menesatti et al., 2008). The efficacy of this method has also been tested in cases where the genetic attribution of individuals was certain (Costa et al., 2008a,b). In this context, we applied a similar methodology to quantify morphological differences in several Tarocco genotypes. To do so, we considered that the explicit definition of a biologically homologous or mathematically determined outline made by the definition of several landmarks (i.e., homologous points) is not required. The study of fruit shape where the object outline of a whole fruit is considered may represent a consistent advantage for Elliptic Fourier Analysis (EFA) in comparison with more traditional morphometric landmark-based methods, (Crampton, 1995; Costa et al., 2008a; Menesatti et al., 2008).

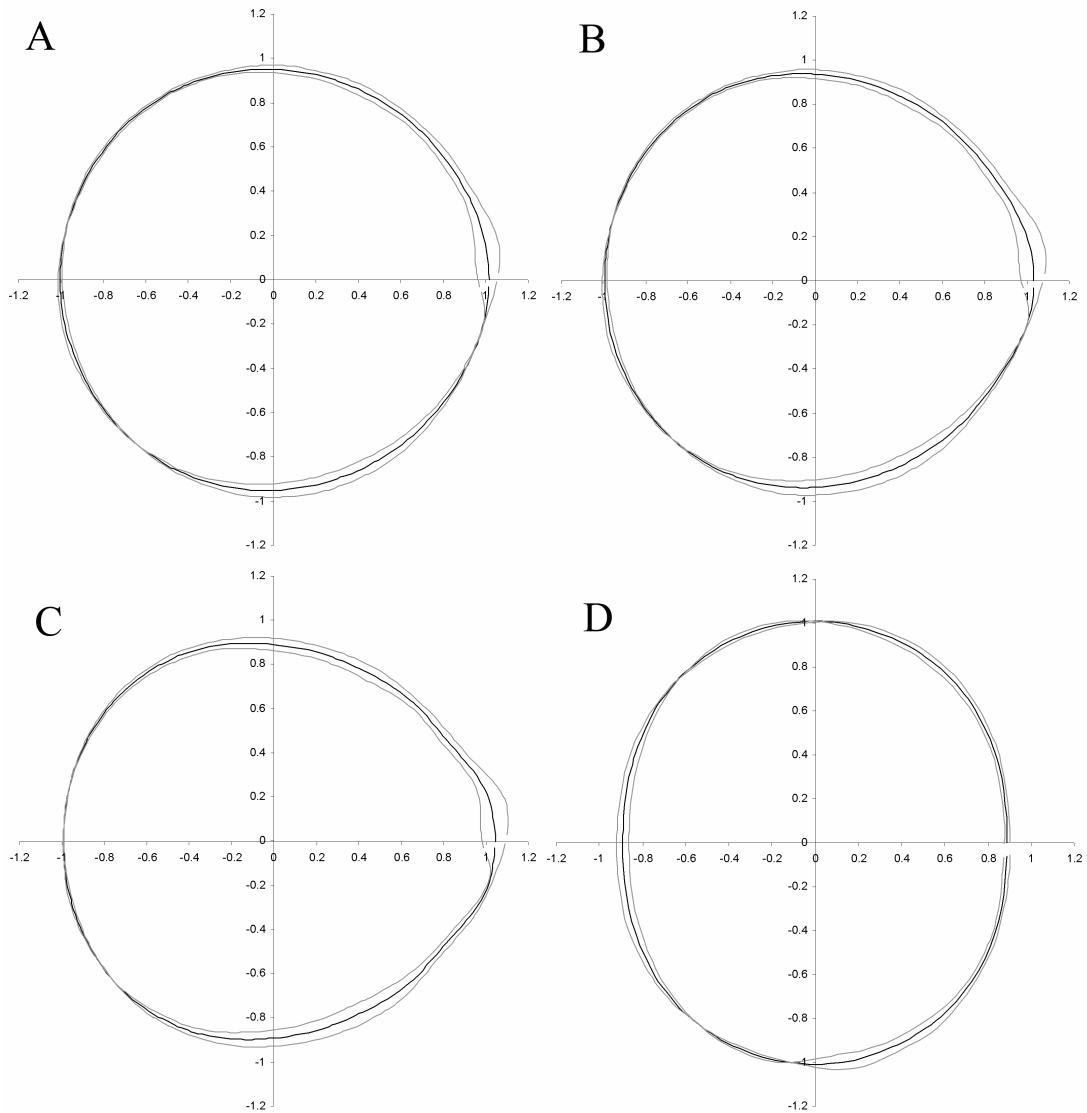
Our data indicate that an instrumental optoelectronic method combining EFA and multivariate morphometric is an efficient tool to statistically discriminate and classify different cultivars and genotypes based on the average fruit shape. The coupling of EFA analysis and the PLSDA modelling technique based on classification and prediction models allows for a more efficient class attribution according to new genotypes and commercial batches of fruit with different appearance qualities. Conducting the same procedure on all orange fruit, as on IPGRI (1999), or on the whole Citrus Industry dataset (Hodgson, 1967), shape groups were not only qualitatively described, but their difference was also instrumentally and numerically

quantified. Moreover, a procedure to estimate the optimal number of clusters with the k-means method was also presented by applying the validation procedures (Ray and Turi, 1999) on the EFA coefficients dataset. This procedure could be applied to other product species such as nuts and apples, identifying in this case 4 clusters.

Our optoelectronic method presents several points of interest in relation to postharvest fruit processing. Firstly, fruit shape was measured in an automated manner and processed by video-image analysis. The required hardware and software to perform these operations are broadly available at low cost and already used today in quality control. Secondly, digitized fruit shape was modelled by means of EFA analysis, which works in a size-independent manner (Menesatti et al., 2008). Shape description was hence made by detailed mathematic configuration allowing not only distinguishing of minimum shape differences, but also quantification of that difference by providing a numerical coefficient used to set a threshold for the acceptance or rejection of the product. The EFA analysis and the statistical clustering techniques (i.e. k-means) provided a quantitative description of fruit shapes which allows the computing of an average shape value through averaged EFA coefficients. These were used to efficiently discriminate different genotypes (see Figs. 5, 6), setting the base for a more generally usable protocol of classification.

In sweet orange, fruit shape is influenced by flowering. Fruit resulting from late blooming may have a more oblong shape, with a drier rind. In this context, our results confirm the capacity of multivariate techniques to discriminate among genotypes of Tarocco, despite the high morphological variability of their fruit. M and SC genotypes, as well as TRI, showed the greatest differences in fruit shape, as confirmed by their separate clustering in all shape classifications (Figs. 3, 4, and 6). In contrast, genotypes of groups 1 and 2 (Fig. 6) were morphologically similar and almost undistinguishable one from the other. All the Scirè genotypes (i.e. S-N1, S-N2, and S-V) and 1E (i.e. 57 and 1E) belong to the same groups: group 1 and group 2, respectively (Fig. 6). Moreover 1E varieties and Tringali are genotypes with the highest degree of shape conformity.

Figure 6: Mean outline (black line) and the standard deviation range (grey line) for fruit shapes classified with k-means on the mean EFA coefficients of each genotype ($k=4$); A. F, GIA, L-N2, S-N1, S-N2, S-V, TDV; B. 1E, 57, G-V, L-N1, L-V, TAP; C. G-N, M, SC; D. TRI.



In our analysis, genotypes derived by nucellar selections were also included. While for the Scirè and 1E varieties, the nucellar selections ranked in the same reference groups (i.e., group 1 and group 2, respectively, see Fig. 6), nucellar selections of Gallo and Lempsò fell into different groups: group 2 and group 3 for Gallo; group 1

and group 2 for Lempso. This underlines the observation that these nucellar selections present significant differences in fruit shape in comparison to their old lines.

In conclusion, our results provide the basis for a new method for sweet orange shape discrimination through image analysis performed on fruit lateral shapes by means of Elliptic Fourier Analysis (EFA). The method may become particularly important when new varieties need to be patented. The IPGRI (1999) and Citrus Industry (Hodgson, 1967) procedures already stress that citrus fruit shape is an important parameter for classification, recognizing several contours such as spheroid, ellipsoid, pyriform, oblique, obloid, ovoid, etc. However, this classification uses subjective evaluation, its power of discrimination being consistently lower than one using multivariate morphological techniques, especially for groups with intermediate morphological characters. Such a subjective classification cannot successfully distinguish all the morphological differentiation presented by clonal selections which represent the source of variability of Tarocco selections. When applied to other citrus fruit, this work may help in defining a new class of descriptors with the strength of a mathematical basis. The system is able to discriminate small differences such as within similar genotypes. Thus, it could successfully be used to recognise shape variations due to events such as deformation caused by storage and handling during shipment or from surface defects negatively influencing fruit quality. Finally it is interesting to note that the system used is size independent.

Acknowledgements

This study was supported by a grant from the Ministero Italiano per le Politiche Agricole (Law 41/82). entitled “Ricerche avanzate in agrumicoltura e loro applicazioni” Paper No. 29. The authors would like to thank Mrs. Iliana Niciarelli and Mr Matteo Cegna for the support during the laboratory operation. Jacopo Aguzzi is a fellow of the Juan de la Cierva Postdoctoral Program (MEC, Spain).

References

- Aguzzi, J., Costa, C., Antonucci, F., Company, J.B., Menesatti, P., Sardá, F., 2009. Influence of diel behaviour in the morphology of decapod natantia.. *Biol. J. Linn. Soc.* 96, 517–532.

- Bonina, F., Saija, A., Tomaino, A., Lo Cascio, R., Rapisarda, P., Dederen, J.C., 1998. In vitro antioxidant activity and in vivo photoprotective effect of a red orange extract. *Int. J. Cosmet. Sci.* 20, 331-342.
- Costa, C., Menesatti, P., Aguzzi, J., D'Andrea, S., Antonucci, F., Rimatori, V., Pallottino, P., Mattoccia, M., 2008a. External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*. In Press *Food and Bioprocess. Technol.* DOI: 10.1007/s11947-008-0068-8
- Costa, C., Aguzzi, J., Menesatti, P., Antonucci, F., Rimatori, V., Mattoccia, M., 2008b. Shape analysis of different populations of clams in relation to their geographical structure. *J. Zool.* 276, 71-80.
- Currie, A.J., Ganeshanandam, S., Noiton, D.A., Garrick, D., Shelbourne, C.J.A., Oraguzie, N., 2000. Quantitative evaluation of apple (*Malus x domestica* Borkh.) fruit shape by principal component analysis and Fourier descriptors. *Euphytica* 111, 219-227.
- Crampton, J.S., 1995. Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia* 28, 179-186.
- Fauconneau, B., Waffo-Teguo, P., Huguet, F., Barrier, L., Decendit, A., Merillon, J.M., 1997. Comparative study of radical scavenger and antioxidant properties of phenolic compounds from *Vitis vinifera* cell cultures using in vitro tests. *Life Sci.* 61, 2103–2110.
- Furuta, N., Ninomiya, S., Takahashi, S., Ohmori, H., Ukai, Y. 1995. Quantitative evaluation of soybean (*Glycine max* L., Merr.) leaflet shape by principal component scores based on elliptic Fourier descriptor. *Breed. Sci.* 45, 315–320.
- Goto, S., Iwata, H., Shibano, S., Ohya, K., Suzuki, A., Ogawa, H., 2005. Fruit shape variation in *Fraxinus mandshurica* var. *japonica* characterized using elliptic Fourier descriptors and the effect on flight duration. *Ecol. Res.* 20, 733–738.
- Hammer, Ø., Harper, D.A.T., Ryan, P.D., 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, 4, 9pp.

- Hodgson, R.W., 1967. Horticultural varieties of citrus. In: Reuther, W., Webber, H.J., Bachelor, L.D. (Eds.), *The Citrus Industry*, vol. 1. University of California, Berkley, pp. 431–591.
- IPGRI, 1999. *Descriptors for Citrus*. International Plant Genetic Resources Institute, Rome, Italy. ISBN 92-9043-425-2 (URL: <<http://www.cgiar.org/ipgri/>>).
- Iwata, H., Niikura, S., Matsuura, S., Takano, Y., Ukai, Y., 1998. Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors. *Euphytica* 102, 143–149.
- Jensen, R.J., Ciofani, K.M., Miramontes, L., C. 2002. Lines, outlines, and landmarks: morphometric analyses of leaves of *Acer rubrum*, *Acer saccharinum* (Aceraceae) and their hybrid. *Taxon* 51, 475-492.
- Kays, S.J., 1999. Preharvest factors affecting appearance. *Postharvest Biol. Technol.* 15, 233–247.
- Lestrel, P.E. (Ed.), 1997. *Fourier descriptors and their applications in biology*. Cambridge Univ. Press, Cambridge.
- Loy, A., Busilacchi, S., Costa, C., Ferlin, L., Cataudella, S., 2000. Comparing geometric morphometrics and outlines fitting methods to monitor fish shape variability of *Diplodus puntazzo* (Teleostea, Sparidae). *Aquacult. Eng.* 21, 271-283.
- Maccarone, E., Rapisarda, P., Fanella, F., Arena, E., Mondello, L., 1998. Cyanidin-3-(6''-malonyl)-glucoside. An important anthocyanin of blood orange juice. *Ital. J. Food Sci.*, 10, 367-372.
- MacQueen, J.B., 1967. Some Methods for classification and analysis of multivariate observations. *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press. 1, 281-297.
- Menesatti, P., Costa, C., Paglia, G., Pallottino, F., D'Andrea, S., Rimatori, V., Aguzzi, J., 2008. Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars. *Biosystem Eng.* 101, 417-424.
- Neto, J.C., Meyer, G.E., Jones, D.D., Samal, A.K., 2006. Plant species identification using Elliptic Fourier leaf shape analysis. *Comp. Electron. Agricult.* 50, 121-134.

- Nijveldt, R.J., van Nood, E., van Hoorn, D.E., Boelens, P.G., van Norren, K., van Leeuwen, P.A., 2001. Flavonoids: a review of probable mechanisms of action and potential applications. *Am. J. Clin. Nutr.* 74, 418–425.
- Ohsawa, R., Tsutsumi, T., Uehara, H., Namai, H., Ninomiya, S., 1998. Quantitative evaluation of common buckwheat (*Fagopyrum esculentum* Moench) kernel shape by elliptic Fourier descriptor. *Euphytica* 101, 175–183.
- Organizzazione di Produttori Agridea, 2009. Disciplinare di produzione. (URL: http://www.agri-idea.com/My_Homepage_Files/disciplinare.pdf).
- Rapisarda, P., Russo, G., 2000. Fruit quality of five Tarocco selections grown in Italy. *Proc. Int. Soc. Citriculture* 2, 1149-1153.
- Rapisarda, P., Bellomo, S.E., Intrigliolo, F., 2001. Anthocyanins in blood oranges: Composition and biological activity. In: Pandalai, S.G. (Ed), *Recent Res. Devel. Agric. Food Chem.; Research Signpost* 5, 217-230.
- Ray, S., Turi, R.H., 1999. Determination of number of clusters in k-means clustering and application in colour image segmentation. Proceedings of the 4th International Conference on Advances in Pattern Recognition and Digital Techniques (ICAPRDT'99), Calcutta, India, 137-143.
- Reforgiato Recupero, G., Russo, G., 2002. La propagazione dell'arancio Tarocco e del clementine Comune. *Italus hortus* 9, 81-84.
- Reforgiato Recupero G., Russo G., Recupero S., Zurru R. Deidda B., Mulas M., 2009. Horticultural evaluation of new *Citrus latipes* hybrids as rootstocks for citrus. *HortScience* 44, 1–4.
- Rohlf, F.J., 1996. Fitting curves to outlines. *Proc. Michigan Morphom. Work., Spec. Publ.* 2, 167-177.
- Rohlf, F.J., 2006. TpsDig Ver. 2.10, digitalized landmarks and outlines. Dept. of Ecology and Evolution, State Univ. of New York at Stony Brook.
- Rohlf, F.J., Archie, J.W., 1984. A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Syst. Zool.* 3, 302-317.
- Sabatier, R., Vivein, M., Amenta, P., 2003. Two approaches for Discriminant Partial Least Square. In: Schader, M., Gaul, W., Vichi, M. (Eds), *Between data science and applied data analysis*. Springer-Verlag, Berlin.

- Sheets, H.D., Covino, K.M., Panasiewicz, J.M., Morris, S.R., 2006. Comparison of geometric morphometric outline methods in the discrimination of age-related differences in feather shape. *Frontiers in Zoology* 3, 15.
- Sjöström, M., Wold, S., Söderström, B., 1986. PLS Discrimination plots. In: Gelsema, E.S., Kanals, L.N. (Eds), *Pattern recognition in practice II*. Elsevier, Amsterdam.
- White, R.J., Prentice, H.C., Verwijst, T., 1988. Automated image acquisition and morphometric description. *Can. J. Bot.* 66, 450–459.
- Yoshioka, Y., Iwata, H., Ohsawa, R., Ninomiya, S., 2004. Analysis of petal shape variation of *Primula sieboldii* by elliptic Fourier descriptors and principal component analysis. *Ann. Bot.* 94, 1–8.

3.8 Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels

FULL REFERENCE OF THE PUBLICATION

Menesatti P, Antonucci F, Costa C, Santori A, Niciarelli I, Infantino A, 2009, Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels. 1st International Workshop on Computer Image Analysis in Agriculture, Potsdam, Germany 27 – 28 August 2009, Bornimer Agrartechnische Berichte - Heft 69, ISSN 0947-7314, Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V. (ATB)

Introduction

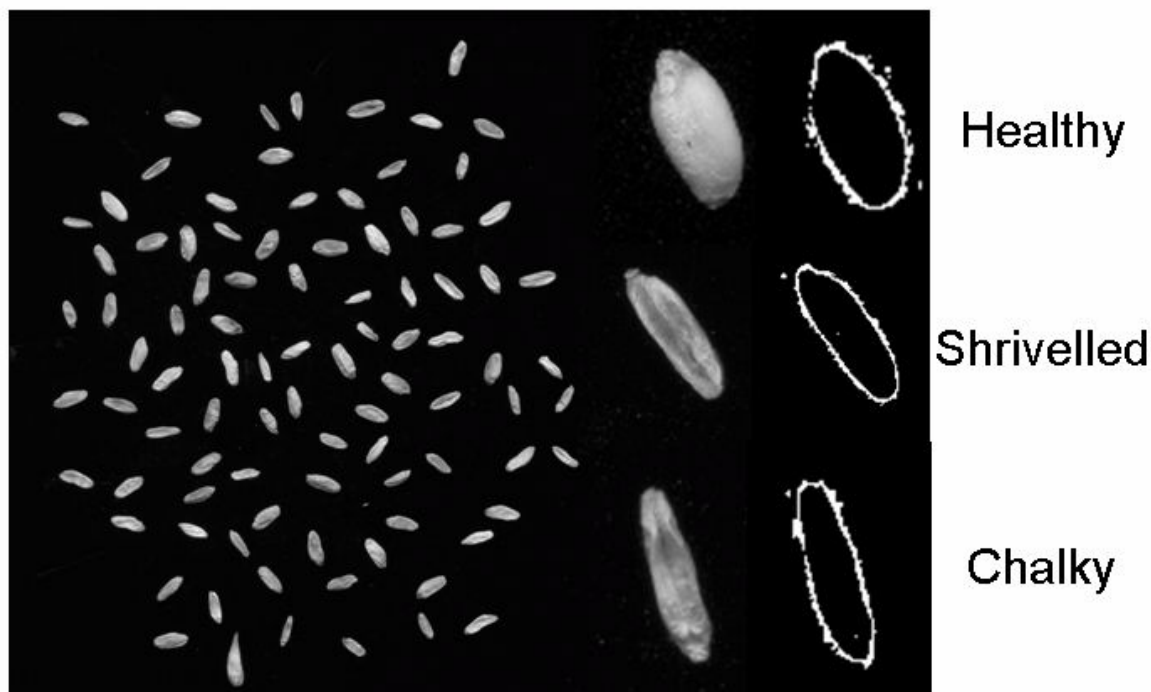
Fusarium head blight (FHB), also known as scab, is a fungal disease that occurs in several winter cereals worldwide. The disease has a great impact on yield, but mostly on health, due to the ability of several *Fusarium* species involved in the diseases to produce mycotoxins, particularly deoxynivalenol (DON), dangerous to human and cattle. These pathogens cause yield and quality losses due to sterility of the florets and formation of fusarium damaged kernels (FDK) typically shrivelled, light in weight and white or pink in colour (Bushnell et al., 2003). The relationships between FDK, fungal biomass and DON content have been extensively studied, but the results are sometimes conflicting (Paul et al., 2005). Most of these methods are destructive, time consuming, expensive and laborious. The proposed Optoelectronic non-contact systems are based on near infrared spectroscopy (Nir) (Dowell et al., 1999). The aim of this work is to suggest a quantitative, non-destructive, cost-effective and rapid image analysis method, based on i) simple low-cost greyscale CMOS video camera for profile acquisition; ii) Elliptic Fourier analysis (EFA) of the kernel outlines; iii) multivariate classification techniques (i.e. PLSDA), to discriminate three different classes of FHB wheat infected kernels in order to correlate estimated and actual FDK.

Materials and Methods

Data collection

Fusarium infected wheat kernels were obtained by artificially inoculating durum wheat plants, of highly FHB susceptible cultivars Simeto, grown at the CRA-PAV farm (Rome, Italy). At flowering, plants were sprayed with a suspensions of *Fusarium culmorum* (ER-1480) or *F. graminearum* (ER-1481) at the concentration of 1×10^5 conidia ml⁻¹ and then covered with plastic sheets for two days after inoculation. At harvest, kernels inoculated were visually grouped in three classes of FHB infection on the basis of their shape (healthy, shrivelled and chalky kernels covered by white mycelium). A total number of 613 kernels, randomly disposed on the plane of a scanner, were acquired with a high resolution scanner (600 d.p.i. 24 bit color). An automated thresholding procedure was performed by an edge detection Sobel filtering in order to detect the edges. Consequently 90 points (x, y) equally angularly spaced from the centroid (one point every 4°), were automatically digitized along the outlines using the software Matlab (rel. 7.1; PLSToolbox Eigenvectorb 4.0) (Fig. 1).

Figure 1. On the left side: samples randomly disposed on the plane of a scanner. On the right side from the top to the bottom: kernels at the three classes of infection (healthy, shrivelled and chalky) before and then the automated outlines acquisition.



Data processing

The overall shape of the artificially infected kernels was studied by EFA on the outline coordinates previously extracted. EFA is based on the separate Fourier decompositions of the incremental changes of the x - and y -coordinates as functions of the cumulative chordal length of an outline polygon. The kernels shape closed contour is expressed in terms of harmonically related trigonometric curves. For each harmonic equation, two Fourier coefficients are computed for both the x - and y -projections thus, the total number of coefficients is $4n$, where n is the number of harmonics fitted to the outline (Crampton, 1995). The total number of harmonics for any outline is equal to half of the total number of outline coordinates (the ‘Nyquist frequency’). The Fourier series was truncated at the value of k at which the average cumulative power is 99.999% of the average total power (Menesatti et al., 2008). The harmonic coefficients, automatically extracted with Matlab, describe the size, shape and orientation of each harmonic ellipse form the input to multivariate statistics. The

classification procedure of the three different classes of infection (healthy, shrivelled and chalky), and affected by the two pathogens, was performed in Matlab on the shape variables (X-block; EFA coefficients) and the shape data were analyzed *via* partial least squares discriminant analysis (PLSDA) (Costa et al. 2009; Menesatti et al. 2008). The X-block values were pre-processed with different Matlab procedures. For the analysis each group was subdivided in: i) 75% of specimens for the class modelling and validation; ii) 25% of specimens for the independent test. PLSDA (Sabatier et al., 2003) is a partial least squares analysis (PLS) regression where the response variable is categorical (Y-block; infection status) expressing the class membership of the statistical units.

Results and Discussion

For the kernels lateral profile the correct number of harmonics (13) was calculated on the whole samplings ('Nyquist frequency' = 45). The PLSDA model performed on EFA coefficients for kernels affected by *Fusarium culmorum* is based on 16 Latent Variables (LV), presents a preprocessing for the X-block as absolute value and a Root Mean Square Error of Calibration (RMSEC) of 0.294. The total percentage of correct classification is 68.4%. Table 1 shows the results relative to the independent test (25%) of the classification procedure performed by PLSDA. In this case the intermediate class (shrivelled) is poorly represented and badly classified (21.1%).

Table 1. Confusion matrix relative to the independent test (25%) of the classification procedure of the three different classes of infection (healthy, shrivelled and chalky) affected by *Fusarium culmorum*. The Mean (%) represents the percentage of cases correctly allocated. The percentage of cases correctly allocated derived by relating the number of cases correctly allocated to the class (in bold) to the total number of cases of that class (Total).

Kernels affected by <i>Fusarium culmorum</i>					
	Chalky	Shrivelled	Healthy	Total	Mean (%)
Chalky	21	0	2	23	91.3%
Shrivelled	5	4	10	19	21.1%
Healthy	2	1	39	42	92.9%

The PLSDA model performed on EFA coefficients for kernels affected by *F. graminearum* is based on 7 LV, presents a “gls weighting” preprocessing for the *X*-block and a RMSEC of 0.401. The total percentage of correct classification is 56.9%. Table 2 shows the results relative to the independent test (25%) of the classification procedure performed by PLSDA. The “healthy” class is the best classified (65.6%). The results suggest that *F. culmorum* infection is better distinguishable than *F. graminearum*, probably because this specie is more aggressive than *F. graminearum* and is able to produce stronger damages on kernels.

Table 2. Confusion matrix relative to the independent test (25%) of the classification procedure of the three different classes of infection (healthy, shrivelled and chalky) affected by *Fusarium graminearum*. The Mean (%) represents the percentage of cases correctly allocated. The percentage of cases correctly allocated may be derived by relating the number of cases correctly allocated to the class (in bold) to the total number of cases of that class (Total).

Kernels affected by <i>Fusarium graminearum</i>					
	Chalky	Shrivelled	Healthy	Total	Mean (%)
Chalky	6	3	3	12	50%
Shrivelled	6	11	3	20	55%
Healthy	1	10	21	32	65.6%

Conclusions

This work was developed to verify the technical and methodological feasibility of the proposed method in order to evaluate the infective incidence of FHB. The shape analysis appeared an important technique to discriminate different classes of FHB wheat infected kernels. The results obtained will be validated on a higher numbers of both, artificially and naturally FHB infected wheat seed lots. Image analysis allows a higher speed of analysis, an increase of the evaluated seeds number and the repeatability of the measure and overpassing the subjectivity of the expert view.

Acknowledgements

This work was funded by the project HighVision (DM 19177/7303/08) from the Italian Ministry of Agricultural, Food and Forestry Politics.

References

- Bushnell, W. R., Hazen, B. E. & Pritsch, C. 2003. Histology and physiology of Fusarium head blight. In *Fusarium head blight of wheat and barley*, pp. 44-83, K. J. Leonard & W. R. Bushnell eds. APS Press, St. Paul.
- Costa, C., Menesatti, P., Aguzzi, J., D'Andrea, S., Antonucci, F., Rimatori, V., Pallottino, F. & Mattoccia, M. 2009. External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*. *Food and Bioprocess Technol* doi: 10.1007/s11947-008-0068-8.
- Crampton, J. S. 1995. Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia* **28**: 179-186.
- Dowell, F. E., Ram, M. S. & Seitz, L. M. 1999. Predicting scab, vomitoxin, and ergosterol in single wheat kernels using near infrared spectroscopy. *Cereal Chem* **76**: 573-576.
- Menesatti, P., Costa, C., Paglia, G., Pallottino, F., D'Andrea, S., Rimatori, V. & Aguzzi, J. 2008. Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars. *Biosystem Eng* **101**(4): 417-424.
- Paul, P. A., Lipps, P. E. & Madden, L. V. 2005. Relationship between visual estimates of Fusarium head blight intensity and deoxynivalenol accumulation in harvested wheat grain: a meta-analysis. *Phytopathology* **95**(10): 1225-1236.
- Sabatier, R., Vivein, M. & Amenta, P. 2003. Two approaches for Discriminant Partial Least Square. In *Between Data Science and Applied Data Analysis*, Schader M., Gaul W., Vichi, M. eds. Berlin, Springer-Verlag.

4. ANALISI DI IMMAGINE IPERSPETTRALE

(hyperspectral imaging)

4.1 Introduzione

L'analisi di immagine spettrale o spettrofotometria d'immagine è una moderna e raffinata tecnica d'analisi optoelettronica che consente di ottenere informazioni sulle caratteristiche chimico-fisiche di un campione attraverso la misura e l'interpretazione del diverso comportamento della superficie dei corpi, relativamente all'assorbimento o alla riflessione delle onde elettromagnetiche (luce) (Elmasry and Sun, 2010).

I sistemi di rilievo spettrofotometrico di corpi solidi possono essere di tipo puntuale, in genere relative all'ampiezza della sonda di rilievo, che forniscono come misura i valori spettrali mediati su tutta l'area di acquisizione, e di immagine che forniscono invece informazioni di tipo più dettagliato (Urbani and Menesatti, 2002).

La spettrofotometria d'immagine, infatti, consente di acquisire, attraverso speciali detector, immagini più o meno ampie dei soggetti di analisi, rilevando simultaneamente lo spettro associato a ciascun punto dell'immagine stessa (pixel). Questo termine indica più propriamente la tecnica di acquisizione e elaborazione di immagini spettrali, ma la tecnica è spesso indicata come analisi di immagine iperspettrale (hyperspectral imaging, HI). Questa tecnica consente di acquisire, attraverso speciali detector, immagini più o meno ampie dei soggetti di analisi, rilevando simultaneamente lo spettro visibile (VIS) o infrarosso vicino (NIR) associato a ciascun punto dell'immagine stessa (pixel), per ciascuna banda spettrale rilevata (Qin, 2010).

Tali sistemi, dunque, consentono non solo una misura spettrale differenziata per punti interessanti dell'immagine, ma anche l'applicazione di tecniche di trattamento elaborativo e di analisi di elementi morfologici significativi dell'immagine (Elmasry and Sun, 2010).

Alla capacità analitica – multiparametrica e multicanale – le tecniche spettrofotometriche associano interessanti caratteristiche di non distruttività della misura, relativa semplicità e velocità, portabilità in ambienti operativi,

interfacciabilità informatica e elettronica e con altri sistemi sensoriali strumentali (Qin, 2010).

Per questo motivo, negli ultimi anni sono cresciuti notevolmente sia l'interesse scientifico sia quello applicativo in merito all'impiego delle tecnologie spettrofotometriche per la misura di aspetti qualitativi dei prodotti agricoli e agroalimentari (Gowen *et al.*, 2007)

I principi fisico-chimici di base alla spettrofotometria, si fondano sulla capacità che ha la materia di assorbire o riflettere, una certa quantità di energia luminosa incidente. Atomi e molecole di un qualsiasi corpo sono, nel loro "stato fondamentale" o non eccitato, in una condizione di equilibrio energetico, che è la risultante della differente distribuzione di energia tra le diverse particelle elementari.

Questo equilibrio, è mantenuto da interscambi di energia che risultano dalle vibrazioni delle molecole e dalle collisioni tra le molecole.

Quando poi una certa quantità di luce s'irradia su di un corpo, questa in parte viene riflessa in parte assorbita. La quantità di luce assorbita dal corpo perturba l'equilibrio intramolecolare, le particelle infatti assorbendo energia passano in un momentaneo "stato eccitato", ossia gli elettroni passano su di un livello orbitale di energia superiore.

Poiché tali stati energetici sono quantizzati, ossia, per passare da uno stato all'altro, gli elettroni richiedono dei precisi valori dell'energia fotonica, verranno assorbiti solo i fotoni che avranno lunghezze d'onda appropriate, tutte le altre verranno riflesse.

Lo studio delle radiazioni luminose assorbite o riflesse, la determinazione cioè del cosiddetto spettro di assorbimento o di riflettanza mediante spettroscopi e spettrofotometri, è quindi un valido ed raffinatissimo mezzo di analisi quantitativa per la caratterizzazione e la identificazione di numerosissimi composti.

Il dato di riflettanza o di assorbanza spettrale può essere tipico, soprattutto quando riferito a specie chimiche altamente concentrate, può pertanto essere utilizzato come indice di uno specifico stato chimico-fisico di un oggetto di analisi.

In tal senso, questo tipo di metodica riscuote elevato interesse nel settore dell'industria agroalimentare soprattutto come metodo per una definizione rapida e oggettiva di stati di qualità dei prodotti commerciali (Menesatti, 2000; Guidetti *et al.*,

1998), ed in campo ecologico come alternativo, rapido e non distruttivo metodo per un monitoraggio di ambienti sia terrestri che marini.

I sistemi di analisi per spettrofotometria d'immagine possono suddividersi in due categorie relative alla distanza di acquisizione:

- *Remote sensing*, che raccoglie informazioni spettrali da satellite o aereo o da terra;
- *Proximal sensing*, che raccoglie informazioni spettrali su oggetti posti a distanza ravvicinata;

Pur non essendo netta la distinzione tra le dimensioni della scena inquadrata, ai fini di una separazione tra le due categorie, è possibile indicare, come il primo sistema permetta di effettuare analisi su aree a più vasta estensione (macroaree territoriali), come ad esempio l'intera chioma di un albero (rilevamento da terra), intere aree coltivate (telerilevamento da aereo) od altri spazi vegetazionali ad elevata copertura come ad esempio foreste, boschi o praterie subacquee di posidonia (telerilevamento aereo o satellitare).

Con le tecniche di *proximal sensing*, invece, si acquisiscono informazioni a distanza più ravvicinata. Tale distanza, in linea teorica, può andare dai pochi centimetri ai manometri, ma nella accezione del presente articolo, si riferisce a rilievi ravvicinati che non coinvolgono il livello microscopico.

In generale, il *remote sensing* fornisce elementi riguardo lo stato generale di un'area e pertanto viene per lo più utilizzata per studi ecofisiologici di monitoraggio ambientale (Hege et al., 2003; Goetz et al., 1985; Ustin et al., 1988; Muirhead, 1998; Pasqualini et al., 1996; Yacobi et al., 1995; Bricaud et al., 1998; Frouin et al., 1996; Anderson and Perry, 1996; Luoheng et al., 1997; Zavattini et al., 2003)), o come strumento di monitoraggio dei sistemi agrari per un'agricoltura di precisione (Bocchi and Boschetti, 2002; Lamb, 2000).

Il *proximal sensing*, invece fornisce analisi più particolareggiate ed è anche utilizzato a livello industriale per un'identificazione e selezione del singolo prodotto (es. nei sistemi di movimentazione in linea) (Kim, 2001; Park et al., 2002))

4.1.1 Proprietà ottiche dei materiali

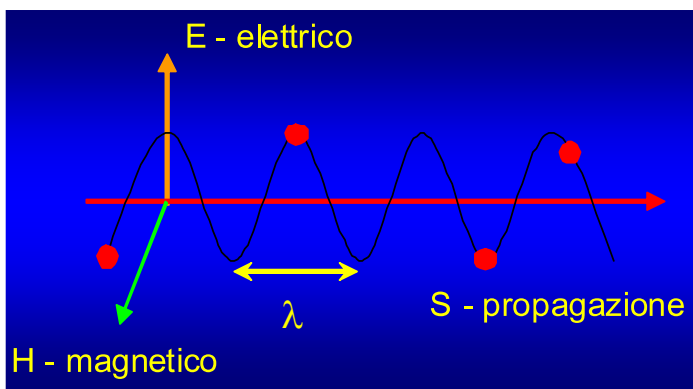
La spettrofotometria e la colorimetria sono sistemi che attraverso opportune tecnologie e metodologie consentono di valutare alcune proprietà dei corpi irraggiati da onde elettromagnetiche di determinate caratteristiche (luce) (Oleari, 1998).

4.1.2 Le Proprietà della luce

La luce caratterizza il mondo nel quale viviamo, è infatti attraverso la luce che percepiamo la maggior parte delle informazioni su quanto ci circonda. E' di primaria importanza, quindi conoscerne le proprietà per comprenderne meglio gli utilizzi nel campo scientifico, infatti in questi ultimi anni si stanno sviluppando sempre di più le tecniche di analisi d'immagine sia qualitativa che quantitativa nei più svariati campi applicativi.

La luce è costituita da onde elettromagnetiche (Figura 1) come dimostrò Maxwell verso la fine dell'Ottocento studiando l'equazioni fondamentali dell'elettromagnetismo. Egli trovò delle soluzioni aventi le proprietà di onde trasversali che si propagano con una velocità che coincide, entro gli errori, alla velocità della luce calcolata sperimentalmente.

Fig. 1. Luce, radiazione elettromagnetica – campo elettromagnetico che si propaga nel tempo e nello spazio



Molti fenomeni luminosi, i più comuni, si possono spiegare con un semplice modello di raggi rettilinei (ottica geometrica); questo approccio più approssimativo è

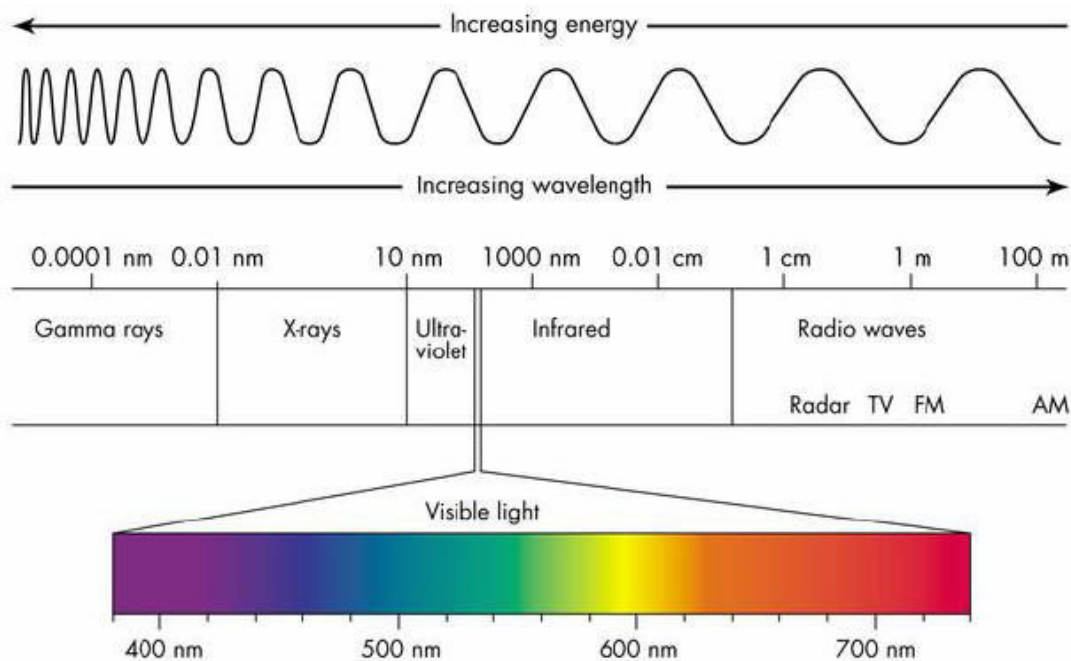
applicabile quando le onde non incontrano ostacoli di dimensioni simili alla lunghezza d'onda della radiazione.

Se osserviamo l'intero spettro di onde elettromagnetiche fantasiosamente detto "l'arcobaleno di Maxwell", risulta chiaro come il genere umano si sia evoluto in un ambiente permeato da radiazioni elettromagnetiche, di cui il sole è la sorgente principale ma non l'unica, pensiamo ai segnali radiotelevisivi, microonde dei sistemi radar, comunicazioni radiofoniche, apparecchiature a raggi X, sostanze radiattive, le stelle e le galassie intorno a noi, ma anche semplici oggetti di uso comune come lampadine flash fotografici e motori (Halliday et al., 2001).

La scala della lunghezze d'onda e delle frequenze è concepita in modo che ogni trattino rappresenta una variazione di un fattore 10, le lunghezze d'onda delle onde elettromagnetiche non hanno limiti intrinseci superiori o inferiori e la scala risulta illimitata. Le componenti principali sono (Fig. 2):

- Le *onde lunghe* vanno da pochi Hz a 10^4
- Le *onde a frequenza radio* vanno da circa 10^4 a 10^9 Hz e comprendono le radiazioni da linee elettriche, le onde radio AM e Fm e le onde TV.
- Le *microonde* vanno da circa 10^9 a 10^{11} Hz e costituiscono le onde utilizzate per le telecomunicazioni e le onde radar.
- L'*infrarosso* va da circa 10^{11} a 10^{14} Hz.
- La *regione del visibile* è di particolare interesse anche se risulta assai ristretta nell'intero spettro elettromagnetico va infatti da 10^{14} a 10^{15} i limiti dello spettro visibile non sono ben definiti perchè dipendono dalla sensibilità dell'occhio umano che è asintotica sia per le lunghezze d'onda corte che lunghe nella è riportata la sensibilità relativa ad un ipotetico occhio standard la massima sensibilità si ha a circa 555 nm corrispondente alla luce del colore giallo-verde.
- L'*ultravioletto* va da circa 10^{14} a 10^{17} la maggior parte delle radiazioni emesse dal sole vengono assorbite dagli strati più alti dell'atmosfera terrestre.
- I *raggi X* vanno da 10^{17} a 10^{19} , oltre si trovano i *raggi gamma*.

Fig. 2. Spettro elettromagnetico



Ci sono corpi come il sole che emettono direttamente la luce e sono detti corpi luminosi “sorgenti di luce”, gli altri oggetti non emettono direttamente la luce, ma diffondono in tutte le direzioni la luce che ricevono dai corpi luminosi.

Alcune proprietà di interazione luce-materia sono(Oleari, 1998):

- per i corpi *non autoluminosi*
 - Riflessione, rifrazione e diffusione: cambiamento direzione della luce;
 - Assorbimento: l’energia assorbita è riemessa in forma radiativa non visibile (es. IR termico).
- Per i corpi *autoluminosi*
 - Fotoluminescenza: fenomeno di eccitazione a data lunghezza d’ondal (es.UV) e remissione ad altra (es. Visibile);
 - Fluorescenza: fotoluminescenza per periodo breve;
 - Fosforescenza: fotoluminescenza per periodo più lungo.

I meccanismi attraverso i quali un corpo può emettere luce sono molteplici, ma l'emissione avviene in pacchetti finiti di energia determinando una modifica dell'energia di un atomo o di una molecola. Gli elettroni degli atomi e delle molecole passano da un livello minimo a stati di energia superiore attraverso il riscaldamento della sostanza o per l'attraversamento di una corrente elettrica.

Un corpo può emettere luce quindi divenire una sorgente di luce se è portato a una temperatura sufficientemente elevata. La luce diverrà sempre più intensa e chiara mano a mano che la temperatura del corpo aumenta.

La tipologia delle radiazioni emesse sono legate alla posizione degli elettroni negli orbitali. L'emissione di luce visibile si ha quando gli elettroni che compiono le transazioni sono elettroni delle orbite esterne. Se le transazioni coinvolgono elettroni di orbite interne si ha l'emissione nell'ultravioletto o di raggi X, mentre le transizioni tra stati eccitati di tipo rotazionale o vibrazionale delle molecole causate dall'agitazione termica a temperature superiori un migliaio di kelvin danno le emissioni nell'infrarosso.

La determinazione della temperatura superficiale di una stella si fa appunto in base ad una valutazione del colore della luce che essa irradia, attraverso un apparecchio speciale detto spettroscopio.

I corpi che sono colpiti dalla luce possono lasciarla passare al loro interno oppure fermarla. Nel primo caso si dice che sono trasparenti, nel secondo caso si parla di corpi opachi. Tra queste due categorie ce n'è una intermedia, quella dei corpi traslucidi.

La luce si propaga in linea retta nei materiali omogenei ma anche nel vuoto, tale comportamento si discosta da quello del suono che si propaga solo nella materia. Per quanto detto fino ad ora la propagazione rettilinea della luce è verificabile solamente fintanto che essa non incontra degli ostacoli con dimensioni lineari confrontabili con la sua lunghezza d'onda, in questo caso la natura ondulatoria diventa dominante.

4.1.3 Le leggi sperimentali della propagazione luminosa: la riflessione e rifrazione

Quando un raggio di luce incontra la superficie di separazione di due mezzi trasparenti si produce un raggio riflesso e uno rifratto, entrambi i raggi rimangono sul piano di incidenza (Fig. 3).

L'angolo di riflessione è uguale all'angolo di incidenza, mentre l'angolo di rifrazione è in rapporto con l'angolo di incidenza secondo la formula:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \quad \text{Legge di Snell} \quad (1)$$

dove:

n_1 = indice di rifrazione del mezzo 1

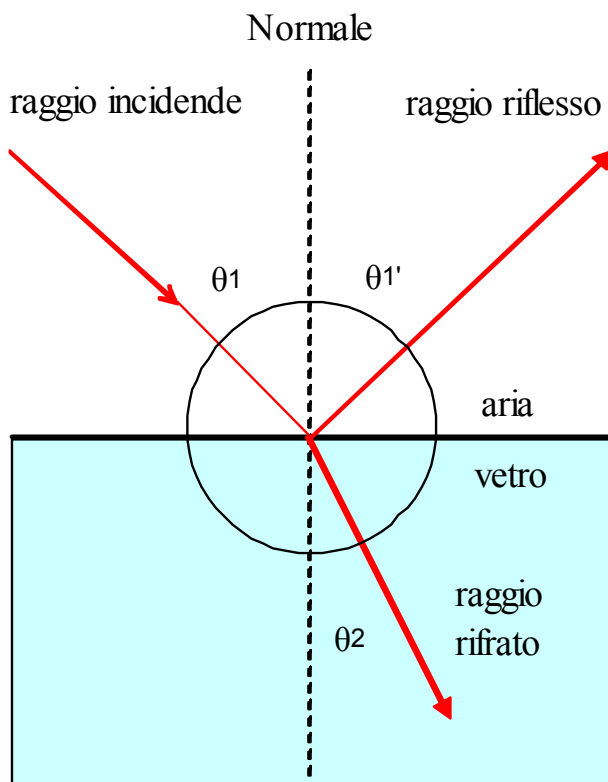
n_2 = indice di rifrazione del mezzo 2

$$n = c/v \quad (2)$$

c = velocità della luce nel vuoto

v = velocità della luce nel mezzo

Fig. 3. Comportamento di un raggio incidente che intercetta una superficie piana di vetro, (riflessione e rifrazione).



Per il vuoto $n = 1$, l'aria si avvicina a tale valore mentre nessun mezzo ha un valore inferiore ad 1.

L'indice di rifrazione dipende dalla lunghezza d'onda della luce ad eccezione del vuoto. Quindi, nello stesso mezzo, radiazioni di lunghezza d'onda diverse hanno anche velocità e angoli di rifrazione differenti. Una luce formata da più lunghezze d'onda si scompone nelle sue componenti fondamentali, quando subisce una rifrazione del suo raggio sulla superficie che separa i due mezzi. Tale effetto è detto *dispersione cromatica*.

Differenti condizioni influenzano il colore e lo spettro misurato, tra queste:

- le caratteristiche fisiche dei materiali;
- le caratteristiche fisico-geometriche della superficie irradiata;
- le caratteristiche della luce incidente;
- le geometrie di illuminazione dei campioni;
- la geometria della raccolta delle radiazioni emessa dai campioni.

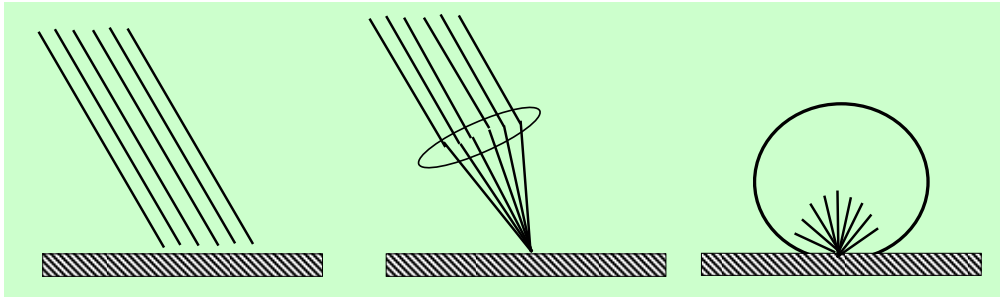
Tra le grandezze spettrali di corpi non autoluminosi, importanti ai fini della presente trattazione, si possono indicare (Oleari, 1998):

- Riflettanza spettrale: rapporto tra flusso radiante riflesso e flusso radiante incidente date condizioni geometriche e lunghezza d'onda
- Trasmittanza spettrale: di un corpo di spessore s è il rapporto tra flusso radiante trasmesso e flusso radiante incidente date condizioni geometriche e lunghezza d'onda
- Riflettanza e trasmittanza sono dette speculari o regolari in assenza di deviazione dei raggi dovuta a diffusione o diffrazione

Le proprietà ottiche dei materiali, dipendono fondamentalmente anche dalle geometrie di irradiazione e visione, tra le principali (fig. 4): parallele, coniche o emisferiche.

Per il principio di reciprocità, i flussi incidenti e riflessi possono essere completamente invertiti senza che la riflettanza cambi.

Fig 4. Esempio di possibili geometrie di irradiazione o visione



4.1.4 Superficie diffondente ideale

Il diffusore “ideale” riflette con un'intensità radiante che cambia proporzionalmente al coseno dell'angolo tra la direzione della riflessione e la normale alla superficie (legge coseno Lambert).

Il massimo della intensità riflessa è per la normale alla superficie

E' necessario tuttavia che la superficie “lambertiana” sia irradiata in modo uniforme, come nel caso della sfera integratrice.

4.1.5 Misure spettrali

Esistono sostanzialmente due categorie misuristiche spettrali (Oleari, 1998):

- *Spettroradiometria (metodo assoluto)*: misura quantitativa di spettri di energia radiante
- *Spettrofotometria (metodo relativo)*: misura comparativa di due spettri di cui uno può essere di riferimento

Le grandezze radiometriche sono di grande utilità, possono però essere sostituite da altre grandezze quando, come nella colometria, si debba quantificare l'emissione luminosa in termini di risposta del sistema visivo umano. Tali grandezze sono dette *fotometriche* e sono definite a partire dalle grandezze radiometriche mediante pesatura con la curva di risposta spettrale dell'occhio umano (*curva fotopica relativa*).

Gli apparecchi per l'effettuazione delle misure spettrali, in base a quanto sopra indicato, possono essere classificati, in:

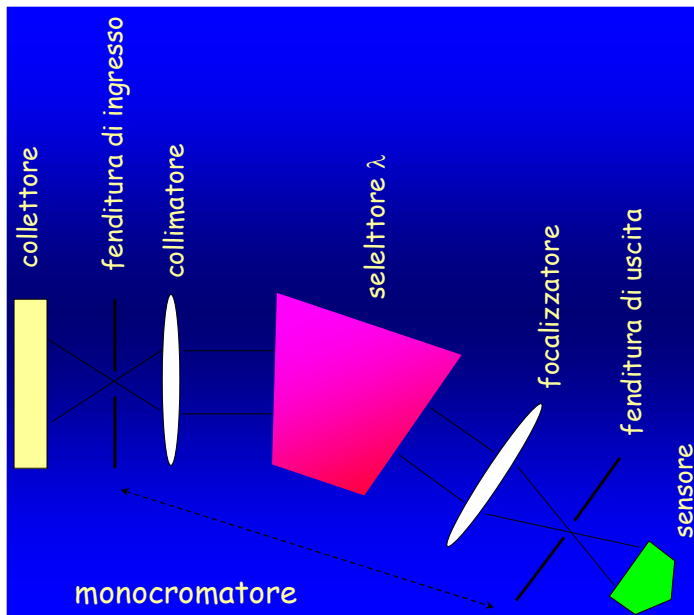
- Fotometri: misurano grandezze luminose con risposta spettrale secondo curva fotopica relativa;
- Colorimetri: misurano caratteristiche tricromatiche a partire dalla curva fotopica relativa;
- Spettrofotometri: misurano grandezze radiometriche spettrali in misura comparativa;
- spettroradiometri: misurano grandezze radiometriche spettrali in misura assoluta;

4.1.6 Spettrofotometri

In fig. 5 è indicato schematicamente un sistema spettrofotometrico.

Il collettore raccoglie la luce e la proietta sulla fenditura di ingresso. Il collimatore indirizza la luce collimata al monocromatore (a prisma o a reticolo) che serve per separare le differenti componenti spettrali della luce. Il focalizzatore ha la funzione di indirizzare in modo ordinato le componenti spettrali sul sensore. La fenditura di uscita ha la funzione di selezionare l'intervallo di lunghezze d'onda, nonché determina la larghezza di banda, ossia la minima ampiezza spettrale di risoluzione delle lunghezze d'onda passanti. Le dimensioni della fenditura determinano una misura discreta dello spettro radiante continuo della luce e per ogni lunghezza d'onda "discretizzata" dal sistema il valore è relativo alla media pesata sull'intervallo dato dall'ampiezza di banda.

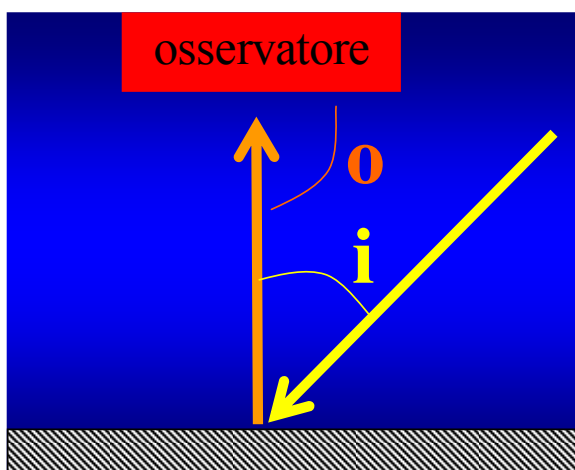
Fig. 5 Schema di base di uno spettrofotometro



Di fondamentale importanza è che la misura della grandezza radiometrica o fotometrica dipende dalla geometria dello strumento (fig. 6), in termini di:

- Angolo di incidenza dell'illuminante (i);
- Angolo di osservazione (o).

Fig. 6. Geometria dello strumento in termini di angolo di incidenza dell'illuminante (i) e angolo di osservazione (o)



Negli strumenti e rilevazioni a norma con le standardizzazioni internazionali (CIE) questi angoli devono essere indicati (Oleari, 1998).

In genere, si indica con β la valutazione in riflettanza, con 1° pedice il valore dell'angolo i , e secondo pedice il valore dell'angolo o .

Nell'esempio della figura, la codifica è $\beta_{45/0}$.

E' possibile inoltre distinguere nei sistemi di irradiazione e osservazione, due tipologie:

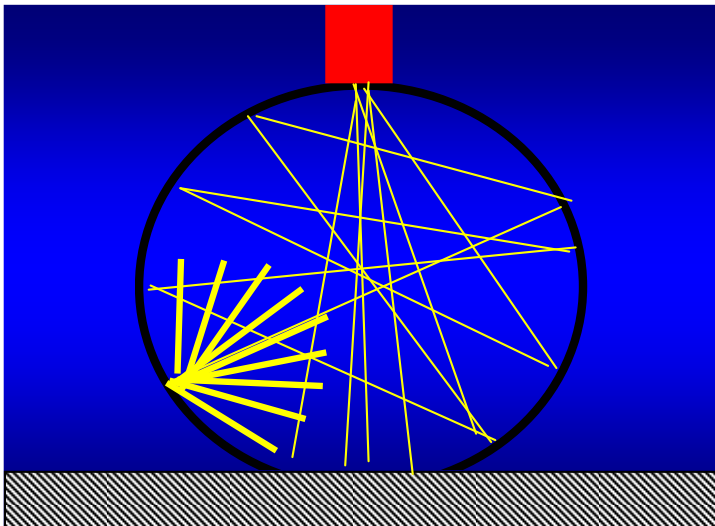
- luce direzionale;
- luce diffusa.

Nel secondo caso, per la codifica di indicazione si utilizza il pedice “ d ”.

Il dispositivo utilizzato per l'irradiazione e/o l'osservazione in luce diffusa è la sfera integratrice, un cui schema è indicato nella figura 7.

Per entrambe le tipologie, le misure spettrali possono includere (SPINC) o escludere (SPEX) la componente di luce riflessa specularmente.

Fig. 7. Schema di dispositivo utilizzato per l'irradiazione e/o l'osservazione in luce diffusa denominato sfera integratrice



4.1.7 Spettrometri e spettrofotometri puntuali

Lo spettrofotometro puntuale è uno strumento atto all'acquisizione della componente spettrale della luce riflessa o trasmessa da oggetti. Il valore di misura è mediato sull'area dell'oggetto inquadrata dalla sonda di acquisizione. Nell'ambito di tale area, il sistema non è in grado di distinguere il contributo spettrale di strutture diverse.

Il sistema è basilarmente costituito da uno spettrometro – apparato ottico passivo – e dal corredo elettronico di generazione di luce e di acquisizione, elaborazione e trasmissione cavo del segnale luminoso convertito in segnale elettrico.

Un esempio di sistema puntuale è quello illustrato nelle figure 8 e 9.

Fig. 8. Esempio di sistema spettrofotometrico puntuale portatile

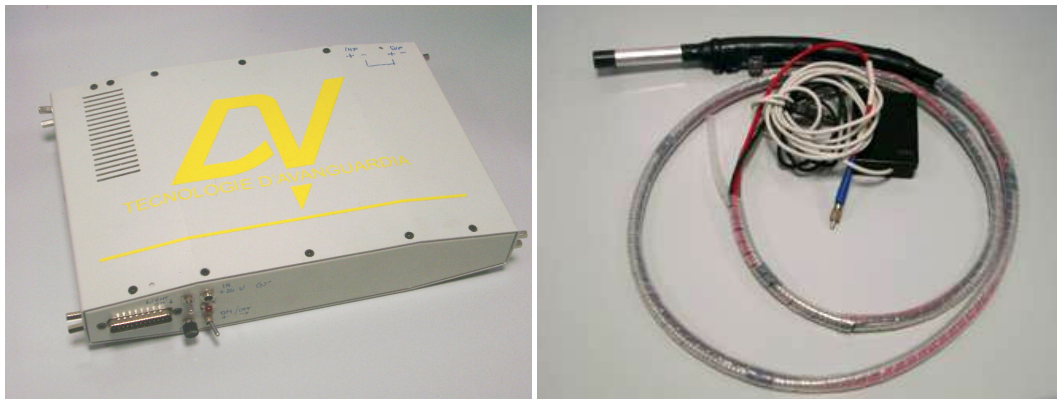
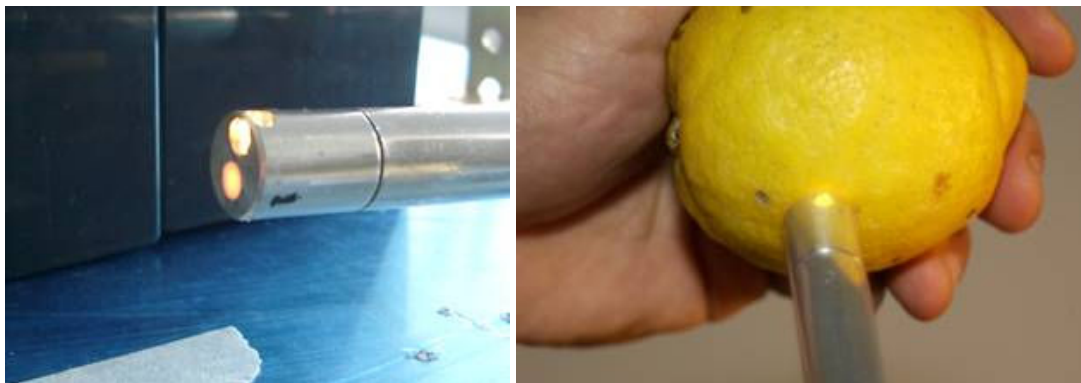


Fig. 9. Dettaglio dell'apparato di illuminazione dell'oggetto di misura e acquisizione del segnale spettrale del sistema spettrofotometrico puntuale illustrato in fig 8.



Il sistema indicato come esempio, disponibile presso l'AgriTechLab (Laboratorio per le Applicazioni Ingegneristiche Innovative in Agricoltura) del CRA-ING (Unità di Ingegneria Agraria del Consiglio per la Sperimentazione e la Ricerca in Agricoltura), ha un range spettrale tra 280 nm e i 1180 nm. In tal senso, questo spettrometro è sensibile all'ultravioletto (UV: 280-400 nm), al visibile (VIS: 400-700 nm) e all'infrarosso vicino (NIR:700-1180). Per l'intero range di sensibilità, il sistema fornisce dati spettrali ogni 3 nm in media (circa 300 lunghezze d'onda), con una dinamica massima di segnale pari a 14 bit (65000 livelli).

Tuttavia, il reale range spettrale di operatività è determinato dalla tipologia di illuminante impiegata, ossia dalle componenti spettrali emesse dalla luce di illuminazione, come indicato dalle curve di risposta spettrale di seguito indicate e riferite a lampade alogene con parabole diecriche (non riflettenti infrarosso) o in alluminio (maggiormente riflettenti la componente Ir).

4.1.8 Spettrometri per immagini

Gli spettrometri per immagini sono sistemi altamente sofisticati che, a differenza dei sistemi puntuali, consentono di rilevare contemporaneamente le componenti spettrali su più punti del soggetto di misura (Qin, 2010). In genere, i sistemi più avanzati operano su una matrice relativamente densa di punti di rilievo (immagine) e per ciascun punto (pixel) acquisiscono l'informazione spettrale sia nel visibile che nell'infrarosso. In tal senso si parla di immagine risolta nello spettro o *spectrally-resolved image* (Qin, 2010).

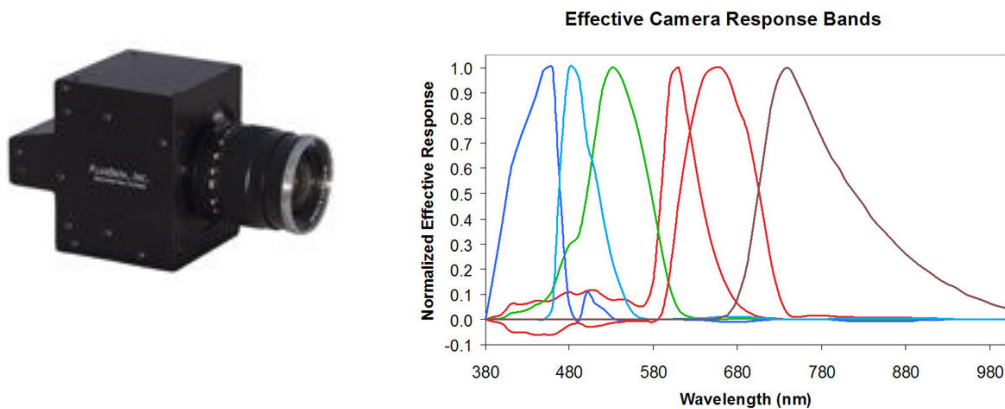
Sempre rispetto ai sistemi puntuali, gli spettrometri per immagine si caratterizzano per la possibilità di rilevare variazioni spettrali legate alla conformazione spaziale (topologica) della superficie del soggetto di misura. In pratica, è possibile osservare per oggetti non omogenei, variazioni spazio-spettrali, altrimenti determinabili in via approssimativa solo misurando ripetutamente per zone contigue con un sistema puntuale.

Tra i diversi sistemi disponibili (Qin, 2010), i meno complessi, sono quelli cosiddetti a (video) camera filtrata (filtered camera), dove l'immagine è acquisita attraverso una

comune camera a stato solido (CCD) e l'informazione spettrale è determinata filtrando lo spettro in ingresso con filtri passabanda diversi. Per ottenere immagini a diverse lunghezze d'onda, o meglio a diversi intervalli di lunghezze d'onda, è necessario acquisire l'immagine più volte variando il range di filtraggio. Questo sistema, ancorché implementi filtri tunabili elettronici, non consente una elevata risoluzione spettrale (minore di 10 intervalli) e fornisce immagini multispettrali (Aleixos et al., 2002).

Un sistema simile, ma più evoluto è rappresentato dalla camera multispettrale (fig. 10). In questo caso il sistema di spettrometria di immagine è costituito tra tre diversi sensori di immagine (CCD), ciascuno dei quali è filtrato per alcune bande dello spettro elettromagnetico del visibile e del vicino infrarosso.

Fig. 10. Esempio di camera multispettrale a 3 CCD (FluxData, **FD-1665 3CCD**, <http://www.fluxdata.com/products/high-resolution-3-ccd-multispectral-camera/?gclid=CIXttK6IgKUCFVcj3wodtmAw8Q>)



I due sistemi indicati rappresentano in molti casi un buon compromesso tra le necessità di velocità di acquisizione (elevata), una buona risoluzione spaziale e la risoluzione spettrale (molto limitata).

Quando invece sia necessario disporre di spettrometri con elevate risoluzioni spettrali, si ricorre normalmente a sistemi più complessi indicati con il termine di sistemi per immagine ad integrazione di linee (per i dispositivi aerei o satellitari) o scanner spettrali per i dispositivi da laboratorio.

Di seguito saranno illustrati due scanner spettrali da laboratorio a disposizione presso l'AgriTechLab del CRA-ING e utilizzati per le attività sperimentali riportate nei successivi paragrafi.

4.1.9 Scanner spettrali

Un sistema spettrometria di immagine, strutturato come scanner spettrale consta di quattro parti fondamentali:

- l'apparato di acquisizione delle immagini spettrali costituito da uno spettrometro accoppiato ad una videocamera CCD e dalle ottiche di ingrandimento e focalizzazione;
- un tavolino traslatore su cui è posizionato l'oggetto di cui acquisire l'immagine spettrale;
- l'apparato di illuminazione in luce stabilizzata, collimata e con una specifica geometria;
- il sistema di acquisizione del segnale, elaborazione e analisi, comprensivo di convertitore A/D, PC e software.

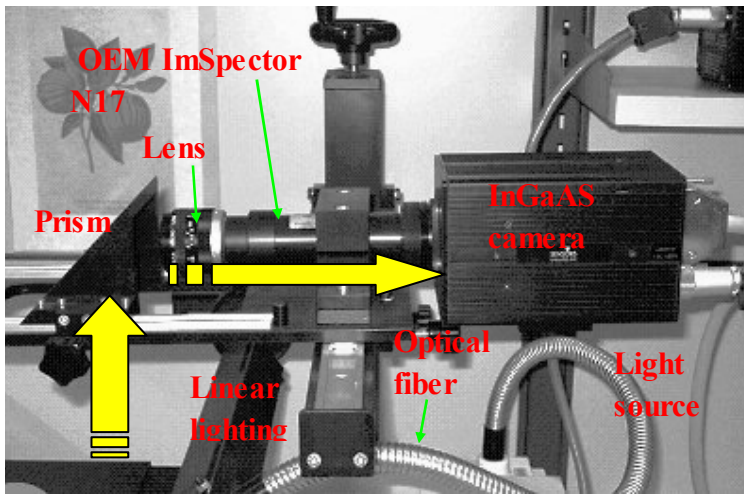
Per descrivere nel dettaglio l'apparato di acquisizione delle immagini spettrali, si fa riferimento al sistema *ImSpector*, brevettato dalla società *Specim* (Oulu, Finland). Questo sistema, nell'utilizzo sperimentale indicato nei paragrafi di seguito, è stato utilizzato in due configurazioni:

- spettrometro per immagini VIS-NIR – 400- 980 nm, (Mod. V10)
- spettrometro per immagini NIR -1000-1700 nm, (Mod. N17)

L'architettura di base dei due modelli è simile e prevede i seguenti elementi comuni (Fig. 11):

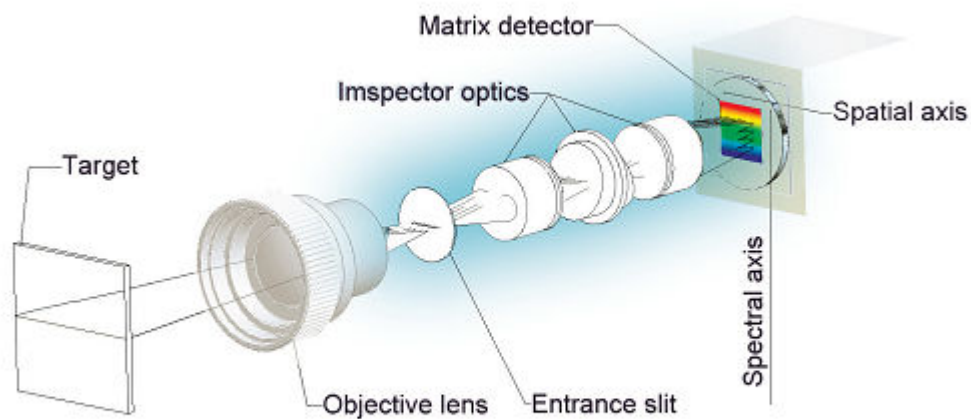
- uno spettrometro ottico;
- una telecamera video;
- le ottiche di visualizzazione e focalizzazione, compreso il prisma di riflessione a 45°

Fig. 11. Architettura del sistema di acquisizione delle immagini spettrali (spectral scanner – N17, Inspector- Specim – Oulu Findand)



Il principio di funzionamento (Fig. 12) si basa sulla dispersione della luce proveniente dal target e passante per lo spettrometro (freccie gialle, in Fig. 11), attraverso la fessura di entrata (entrance slit), operata dal sistema brevettato PGP (Prism-grating-prism; prisma-reticolo-prisma) (Vaarala *et al.*, 1997).

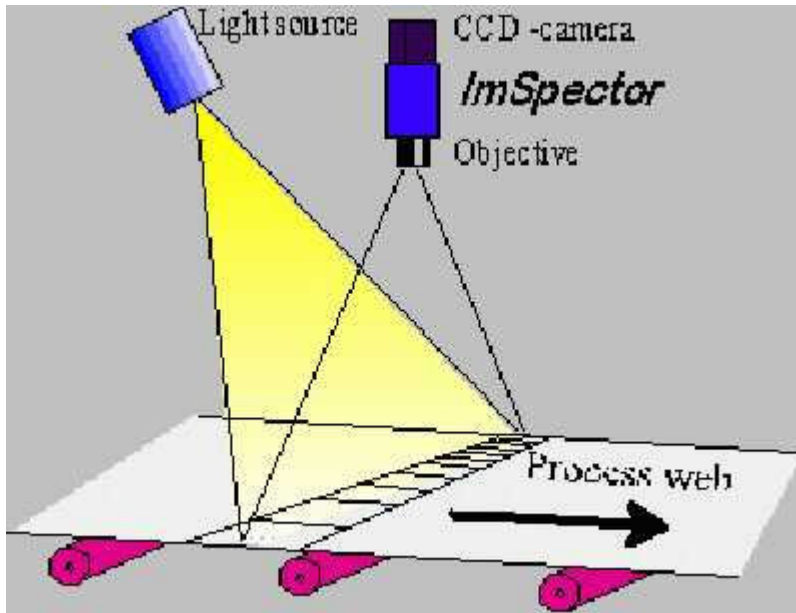
Fig 12. Principio di funzionamento del sistema spettrometrico proprietario ImSpector (http://www.autovision.net/docs/ImSpector_brochure1.6.pdf)



Operating principle of ImSpector

La luce proveniente dall'oggetto è relativa ad una porzione lineare di questo (fig 13) parallela all'*entrace slit* dello spettrometro (immagine lineare).

Fig. 13. Esempio di acquisizione di una linea di pixel nel sistema ImSpector



Essa è dispersa sul sensore matriciale della telecamera utilizzando un asse del sensore per l'informazione spaziale e l'altro per l'informazione spettrale (Fig. 12).

Questo fa dell'ImSpector il primo spettrometro d'immagine a visione diretta nel quale la luce viaggia direttamente attraverso l'asse ottico dello strumento.

Vi sono notevoli vantaggi dati da questa configurazione

(<http://www.vinci.it/pop3.htm>):

- lo spettrometro si installa interponendolo tra il corpo della telecamera e l'ottica di un sistema già esistente senza bisogno di dover riconfigurare il sistema optomeccanico
- produce un' immagine di alta qualità senza distorsioni;
- non vi sono ottiche riflettive ed il sistema è indipendente dalla polarizzazione;
- Il sistema è molto robusto e compatto (L135 x W70 x 1160) mm ed è ideale per l'applicazione in ambiente industriale;

- il sistema è molto facile da customizzare per specifici impieghi OEM nella regione spettrale 380-2400 nm.

Le caratteristiche specifiche dei due spettrometri utilizzati per le prove sperimentali sono riportate nella tabella 1.

Tabella 1. Specifiche tecniche e operative dei due spettrometri per immagini utilizzati nelle prove sperimentali successivamente descritte.

Producer	SPECIM	SPECIM
model	V10	N17
Spectral range	400 - 980 nm	1000-1700 nm
Dispersion	90.9 nm/mm	120 nm/mm
Spectral resolution	3 nm	12 nm
Image size	6.6 (spectral) x 8.8 (spatial) mm, corresponding to standard 1, 2/3" image sensor.	6.6 (spectral) x 8.8 (spatial) mm, corresponding to standard 2/3" image sensor.
Spatial resolution	15 line-pairs/mm, rms spot radius < 60 μm within 2/3" image area	15 line-pairs/mm, rms spot radius < 60 μm within 2/3 " image area
Aberrations	Insignificant astigmatism	Insignificant astigmatism
Bending of spectral lines across spatial axis (smile)	< 50 μm (1/2 " area), < 80 μm (2/3 " area)	< 50 μm (1/2 " area), < 80 μm (2/3 " area)
Bending of spatial lines across spectral axis (keystone)	< 16 μm (1/2 " area), < 25 μm (2/3 " area)	< 16 μm (1/2 " area), < 25 μm (2/3 " area)
Numerical aperture	F/2.8	F/2.8
Slit width	25	80 micron
Effective slit length	9.8 mm	9.8 mm
Total efficiency (typical)	> 50%, independent of	> 50%, independent of

	polarization	polarization
Stray light	< 0.5 % (halogen lamp, 633 nm notch filter)	< 0.5 % (halogen lamp, 633 nm notch filter)
Matricial Sensors for spatial/spetral dispersion	Sony CCD 2/3"	Sensors Unlimited, INC. SU128 Focal Plane Array – matrix resolution 128 x 128 pixel
A/D	10bit	12bit

Il sistema spettrometrico fin qui descritto, opera come una telecamera spettrale lineare, acquisendo, perciò, di un oggetto statico una immagine caratterizzata da una sola riga di pixel.

Per ottenere una immagine spaziale bidimensionale completa è necessario movimentare l'oggetto inquadrato dal sistema spettrale sincronizzando l'acquisizione spettrale con la traslazione dell'oggetto.

A tal fine è stato realizzato un dispositivo di movimentazione e sincronizzazione, denominato scanner spettrale (Fig. 14) o spectral scanner (DV, Padova), utilizzabile per entrambi i sistemi spettrali indicati in tabella 1.

L'intera apparecchiatura comandata attraverso personal computer consente l'acquisizione immagini complete di informazioni spaziali e spettrali di prodotti posizionati sul tavolino di traslazione, rimanendo nell'ambito delle dimensioni dettate dall'ottica utilizzata e dalla distanza tra l'oggetto e il sensore spettrale. La velocità di traslazione, come pure il numero di linee dell'immagine, sono variabili a seconda della necessità di acquisire più o meno velocemente immagini più o meno risolte. La risoluzione spettrale non è invece modificabile.

Fig 14. Sistema scanner spettrale o spectral scanner (DV srl, Padova) che associa un tavolino di traslazione orizzontale (2) al sistema ImSpector (1) consente l'acquisizione e il salvataggio su PC trasportabile (3) di immagini spettrali complete dei prodotti



Trattandosi di immagini la cui risoluzione ottica è di circa 300 000 pixel che contengono informazioni spettrali a 8 bit per 120 lunghezze d'onda per ciascun pixel, si comprende l'enorme quantità di informazioni contenuta in ciascun in ciascun singolo file e quindi le dimensioni informatiche del file stesso. Per limitare lo spazio di storicizzazione di queste immagini, pur mantenendo un'elevata risoluzione spaziale, è stata impiegata una velocità media di scannerizzazione, acquisendo circa 600 linee per la coordinata x e dalle 300 alle 500 linee per la coordinata y, a seconda della grandezza dell'oggetto o per gli oggetti da visualizzare. Ciò nonostante l'intera analisi effettuata ha comportato da storicizzazione di più di 35 Gb di immagini spettrali.

Lo scanner spettrale dispone di una sorgente di illuminazione stabilizzata a fibra ottica a spettro solare che illumina trasversalmente il tavolino di traslazione.

L'illuminante è posizionata a quarantacinque gradi rispetto al piano del tavolino, ottemperando in tal modo alle condizioni geometriche poste dagli standard colorimetrici CIELab o anche CIE 1976 ($L^*a^*b^*$) (Oleari, 1999).

Operazioni preliminari alla misura molto importanti sono quelle relative alla taratura di fondo scala del nero e del bianco. Infatti la misura in termini spettrali è fornita dal parametro di riflettanza (R), che è pari alla quota riflessa della luce inviata (radianza) calibrata rispetto a un valore 0 (nero) è un valore di 100% (bianco standard).

Il calcolo della riflettanza è effettuato attraverso la seguente formula:

$$R = \frac{r_s - r_b}{r_w - r_b} \quad (1)$$

dove

R la riflettanza relativa del campione per ogni lunghezza d'onda;

r_s è il valore assoluto del segnale (radianza) misurata per il campione ad ogni lunghezza d'onda;

r_b è il valore assoluto del segnale (radianza) misurata ad ogni lunghezza d'onda per il nero (background noise)

r_w è il valore assoluto del segnale (radianza) misurata ad ogni lunghezza d'onda per il bianco standard (100% di riflettanza).

In tal modo, si ottengono informazioni standardizzate e conoscendo lo spettro della luce emessa e lo spettro riflesso dal bianco standard, a sua volta precedentemente calibrato, è possibile confrontare dati di riflettanza e di colorimetria si con altri sistemi di misura.

L'insieme delle informazioni spettrali e colorimetriche contenute nelle immagini spettrali è gestito da un apposito software, che completa lo spectral scanner. Le immagini spettrali storicizzate dopo l'acquisizione, possono essere analizzate successivamente in modalità offline utilizzando il tools di misura messi a disposizione dal software.

4.1.10 Processamento dell'immagine spettrale

Le immagini spettrali, ottenute con i sistemi sopra descritti, possono essere processate attraverso due fondamentali approcci (Menesatti *et al.*, 2010):

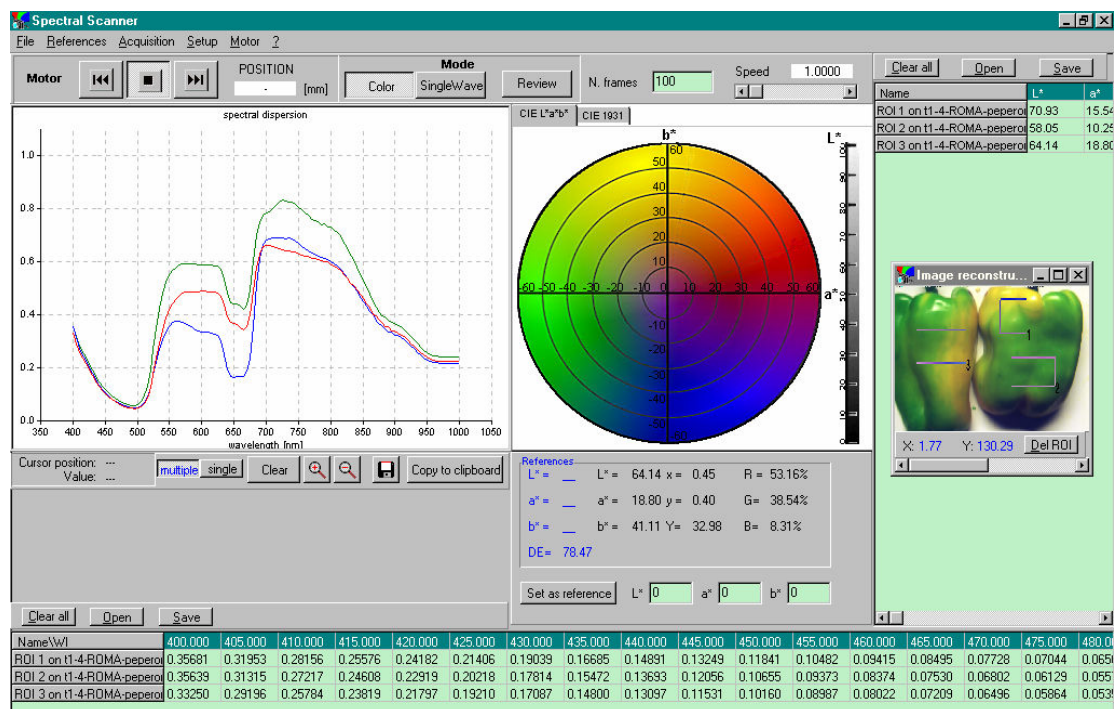
- ispezione e analisi visuale da operatore esperto;
- analisi di immagine multi o iperspettrale in automatico.

La procedura di estrazione da operatore delle informazioni spettrali dalle immagini acquisite è passaggio cruciale per l'analisi effettuata. Infatti disponendo dell'intera immagine spettrale dei prodotti è possibile scegliere zone o aree differenti di prodotti stessi cui corrispondono spesso risposte spettrali e colorimetriche diverse. Inoltre devono essere evitate dall'analisi le zone non propriamente appartenenti al prodotto (es. frutto, peduncoli o foglie) o le zone più in ombra o meno visibili dovute alla conformazione geometriche del prodotto stesso.

Data la difficoltà di scelta, questa è in genere effettuata da operatori esperti che, in un secondo momento, possono riaprire l'immagine spettrale storicizzata e effettuare con estrema cura la misura in zone confacenti e prestabilite. Utilizzando un tool dal software, l'operatore traccia una ROI (region of interest) delimitando l'area quadrata o rettangolare interessante per la misura (Fig. 15). In genere, tranne casi di prodotti molto piccoli come gli acini d'uva, l'area tracciata ha una superficie approssimativa di 10-15 centimetri quadrati, corrispondente a circa 3000-5000 pixel. Per ottenere una misura mediata più rappresentativa, è consigliabile tracciare più ROI in zone diverse del prodotto analizzato.

Dell'intera area prescelta il software calcola il valore medio su tutti pixel relativo ai valori di riflettanza per tutte le lunghezze d'onda e alle coordinate colorimetriche. L'insieme delle informazioni colorimetriche e spettrali, sotto forma numerica, è poi esportato in formato testo e successivamente importando in foglio elettronico dove costruire un data base per associare alle informazioni spettrali, i valori di riferimento qualitativi o quantitativi.

Fig. 15. Schermata del software di analisi dello spectral scanner con in evidenza le curve di riflettanza (a sinistra) relative alle aree di misura (ROI) contrassegnate da operatore, su prodotti ortofrutticoli (a destra), il diagramma colorimetrico (al centro), i dati colorimetrici (in alto a destra) e i dati della riflettanza spettrale per le diverse lunghezze d'onda (in basso).



4.1.11 Analisi d'immagine iperspettrale

Le procedure di analisi in automatico delle immagini multispettrali (quando meno di 10 lunghezze d'onda) o iperspettrali (> 10 lunghezze d'onda) (Menesatti *et al.*, 2010), comprendono due principali elementi, intrinsecamente legati alla duplice natura, spettrale e di immagine, della tecnica:

- elaborazione ed analisi dell'informazione spettrale
- elaborazione ed analisi dell'informazione spaziale o topologica.

Elaborazione ed analisi dell'informazione spettrale

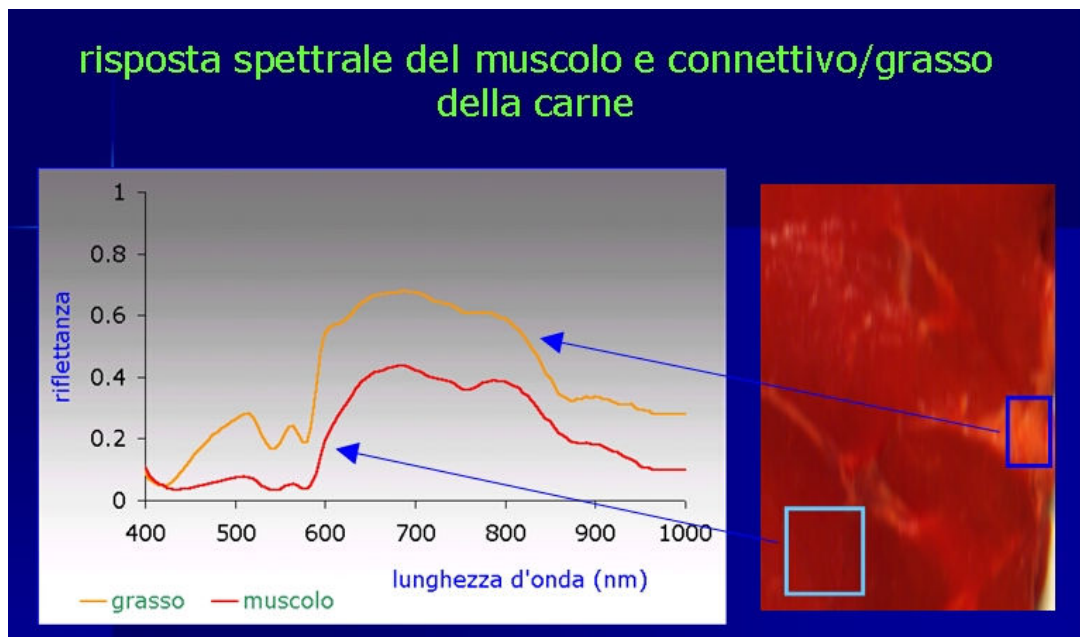
Lo spettrofotometro è uno strumento che dal punto di vista misuristico, è detto multicanale. Ossia, per ciascuna lettura o misura, fornisce in uscita non un solo valore, ma un insieme (array) di valori associati a quella misura, ma riferiti a

differenti lunghezze d'onda. Altra particolarità è che, per una serie di misure, i valori spettrali ad una data lunghezza d'onda sono, spesso, correlati con i valori rilevati ad altre lunghezze, in genere quelle più vicine.

In effetti, soprattutto per la spettrofotometria su campioni solidi non trattati - come nei casi esposti nella presente trattazione - quasi mai sono evidenziabili risposte spettrali molto “piccate”, ossia risposte con picchi molto stretti per specifiche lunghezze d'onda, quanto piuttosto andamenti costanti di crescita o calo per ampi intervalli.

Se, per esempio, la riflettanza media di una serie di campioni di carne, evidenzia un andamento crescente massimo per 700 nm a partire dai 600, è altamente probabile che le risposte spettrali per lunghezze d'onda maggiori di 600 nm e minori di 700 nm siano altamente correlate tra loro (Fig. 16).

Fig. 16. Esempio di risposta spettrale di diverse aree di una fettina di carne (immagine spettrale acquisita presso il CRA-ING con sistema scanner spettrale VIS-NIR)



Questa proprietà, riscontrabile anche in altri strumenti multicanale, agevola e indirizza l'applicazione di particolari metodiche di analisi statistica dei dati,

utilizzando la capacità di sintesi informativa ottenibile sfruttando le relazioni funzionali tra le risposte (misure) dei differenti canali (Di Natale et al., 1997). Tali metodiche sono complessivamente iscrivibili alla categoria dei sistemi statistici multivariati, principalmente con finalità discriminatorie di classificazione o predizione di risposte quantitative (Chevallier et al., 2006; Haaland and Thomas, 1988).

Altro intervento molto importante, ai fini analitici, è il trattamento o pre-processing dei dati spettrali (Burger and Geladi, 2006; Burger and Geladi, 2007)

Numerose e, in molti casi, complesse e specifiche, sono le tecniche di elaborazione del segnale, preliminare all'analisi discriminatoria, tra queste: l'applicazione di derivate prime o seconde (Davies, 2007), filtraggi passa banda basso o alto, applicazione di algoritmi molto specifici (Burger and Geladi, 2007).

Tali tecniche sono fortemente consolidate in ambito chimico e chemometrico e sono state brevemente descritte nell'introduzione generale della presente tesi..

Elaborazione ed analisi dell'informazione spaziale

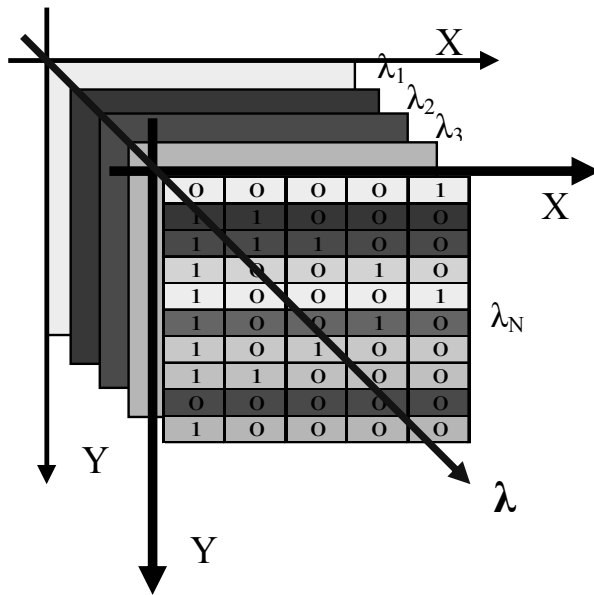
Nelle immagini multi o iperspettrali, l'informazione spettrale è relativa ad ogni pixel dell'immagine stessa. Poiché l'informazione spaziale dell'immagine è rappresentata da una matrice di pixel estesa nelle due dimensioni X e Y, l'immagine iperspettrale comprenderà tre dimensioni X, Y, Z (Elmasry and Sun, 2010), dove la terza dimensione è rappresentata dalle 114 (VIS-Nir) o 70 (Nir) differenti lunghezze d'onda.

Si parla in questo caso di ipercubo o matrice cubica iperspettrale le cui caratteristiche sono esemplificate in figura 17.

Ogni livello (layer) dell'asse Z (l'asse delle lunghezze d'onda) corrisponderà ad una immagine spaziale ad quella specifica lunghezza d'onda.

In pratica, l'immagine iperspettrale è una immagine spaziale a multistrato spettrale.

Fig. 17. Esempio schematico di ipercubo o matrice cubica iperspettrale



Ai fini di una interpretazione operativa spazio-spettrale è necessario condensare e sintetizzare le due dimensionalità informative fino ad arrivare ad immagini facilmente interpretabili o a numeri o indici complessivi.

La strategia di “sintesi” informativa è in relazione alla tipologia interpretativa prescelta: se il fenomeno indagato non presenta variazioni nella sua distribuzione spaziale (ad es. il colore di una soluzione liquida omogenea) tutta l’informazione dell’immagine potrà essere condensata in pochi numeri senza considerare differenze di area, altrimenti, come nel caso di immagini di prodotti agricoli, sarà importante valutare anche differenze di distribuzione del fenomeno (ad es. la presenza non omogenea di un contaminante superficiale) anche per le diverse zone dell’immagine (Menesatti *et al.*, 2009).

Questo secondo caso è quello che maggiormente interessa l’analisi di immagine iperspettrale propriamente detta.

La strategia di analisi di immagini iperspettrali si articola, normalmente, nelle seguenti fasi (Elmasry and Sun, 2010):

- 1) calibrazione del dato spettrale in riflettanza o assorbanza
- 2) preprocessamento del dato spettrale
- 3) analisi del dato spettrale

4) processamento dell'immagine

Le prime due fasi sono state già brevemente indicate nei paragrafi precedenti. Tali fasi sono, in realtà, preliminari al trattamento del dato spettrale che deve essere sintetizzato per estrarre l'informazione significativa. Questo può essere fatto ricorrendo a diverse tecniche:

- estrazione di singole lunghezze d'onda ritenute significative o altamente informative (ad es. l'assorbimento nella banda dei 1400-1440 nm per il contenuto di acqua) (Osborne *et al.*, 1997);
- confronto tra diverse lunghezze d'onda (sottrazioni, rapporti spettrali);
- clustering multivariato;
- classificazione multivariata (Cheng *et al.*, 2004);
- predizione multivariata (Naganathana *et al.*, 2008).

Le tecniche di analisi multivariata applicabili possono essere non supervisionate (es PCA) (Cheng *et al.*, 2004), dove la sintesi informativa è realizzata a partire dalla variabilità interna dei valori spettrali, o supervisionata (PLS, PLSDA) (Haaland and Thomas, 1988; Chevallier *et al.*, 2006), dove tale sintesi è indirizzata dalla risposta nota della variabile dipendente, sia essa categorica o quantitativa.

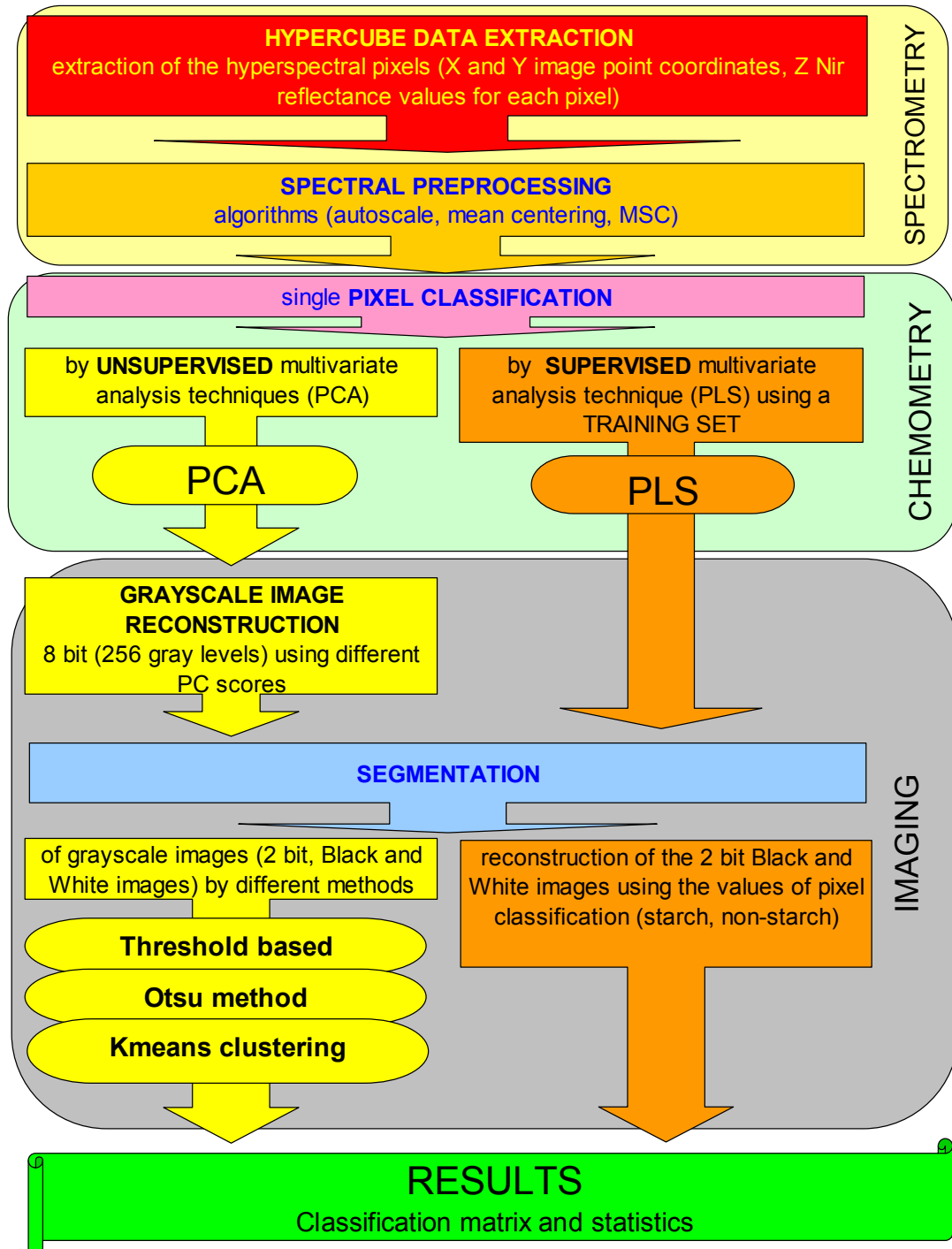
L'obiettivo è di ottenere, in base alla sintesi informativa della dimensione spettrale, una immagine bidimensionale più facilmente interpretabile.

In funzione della tipologia di analisi spettrale è possibile ottenere immagini spaziali, con informazioni a due livelli (bianco/nero) quindi direttamente interpretabili o a più livelli (es. 8 bit pari a 256 livelli di grigio). In tal caso, è necessario procedere ad un ulteriore passaggio di analisi e sintesi informativa, effettuato attraverso tecniche di image processing (Zwiggelaar *et al.*, 1996).

Nel complesso, la procedura di analisi di immagini iperspettrali può essere sviluppata come nell'esempio indicato nello schema di figura 18 (Menesatti *et al.*, 2009).

In questo esempio, si riporta l'approccio utilizzato per l'analisi di immagini iperspettrali NIR per la valutazione dell'indice di amido delle mele. La parte sperimentale sarà più dettagliatamente descritta nei seguenti paragrafi.

Fig. 18. Esempio di procedura di analisi di immagini iperspettrali attraverso tecniche chemometriche differenti (non supervisionate e supervisionate) e applicazione di tecniche di analisi di immagine



Ulteriori possibili passaggi elaborativi intermedi o finali sono:

- Identificazione della o delle variabili più informative, selezione delle variabili e modellizzazione differenziata (Osborne *et al.*, 1997);
- analisi dei loadings e degli scores derivanti da tale analisi per identificare eventuali range spettrali o singole lunghezze d'onda da utilizzare nella ricostruzione dell'immagine Cheng *et al.*, 2004);
- applicazione di tecniche di augmentation e segmentation dei dati (Zwiggelaar *et al.*, 1996);
- Ricostruzione dell'immagine dai loadings, scores o dai ratio tra lunghezze d'onda;

Molto importanti sono anche le tecniche di image processing.

Tra quelle più avanzate possiamo citare :

- misura di caratteristiche di densità ottica
- misura di parametri di imaging texture analysis
- rotation invariant features
- self-organizing map
- histogram 1st order statistics
- object morphological parameters
- tecniche frattali

Una menzione particolare spetta alla possibilità, concretamente verificata per la prima volta in letteratura, con la pubblicazione del lavoro (Menesatti *et al.*, 2010) riportato nel seguente paragrafo, della combinazione delle due tecniche di analisi della forma e analisi di immagine iperspettrale.

4.1.12 Analisi di immagine iperspettrale: innovazione in ambito biofotonico e agroalimentare

Uno dei problemi al momento maggiormente affrontato dall'industria agroalimentare è quello di studiare un sistema strumentale che effettui, in modo oggettivo, un controllo sulla qualità del prodotto commerciale, al fine di determinare maggiori garanzie sull'effettive caratteristiche organolettiche ed igienico-sanitarie.

In tal senso, l'analisi di spettrofotometria di immagine VIS-NIR, risulta essere una tra le tecniche innovative e non distruttive più interessanti, in quanto permette, riferendosi ad una vasta gamma di prodotti, di dare valutazioni sulla qualità rapide, oggettive e multianalitiche.

L'analisi spettrofotometrica, infatti, è in grado attraverso lo studio delle differenze dello stato chimico-fisico dei prodotti agricoli di definire dei validi criteri di oggettività utili per la determinazione di una stima della qualità del prodotto agricolo e alimentare.

Gli aspetti di qualità del prodotto, che possono essere misurati con tale metodologia, si riferiscono a quelli che maggiormente influenzano la scelta del consumatore e pertanto il mercato delle vendite: aspetto, colore e gusto, ossia le proprietà organolettiche e merceologiche, ma anche difetti, danni o contaminazioni.

L'analisi di immagine multi o iperspettrale, può essere eseguita in riflettanza, in trasmittanza o in fluorescenza. La maggior parte degli studi pubblicati sono state effettuate in modalità di riflettanza (Gowen et al., 2010), che è di solito eseguita nelle bande spettrali Vis-NIR (400-1000 nm) e NIR (1000-1700 nm).

Per quanto riguarda i rilievi dei difetti o danni sui frutti, Cheng et al. (2004) e Liu et al. (2005) hanno utilizzato un sistema di imaging iperspettrale che operava nella regione 430-930 nm in riflettanza per individuare i danni da freddo su cetrioli. Nicolai et al. (2006) hanno utilizzato un sistema NIR (900-1.700 nm) per individuare i difetti sulle mele. Il sistema è stato in grado di identificare le lesioni invisibili a occhio nudo, anche se la ridotta luminosità ai bordi dell'immagine ha causato alcuni errori di classificazione. Elmasry et al. (2009) hanno studiato l'uso di un sistema di Vis HSI NIR (400-1000 nm) e di reti neurali per la rilevazione di danni da freddo nelle mele. Gowen et al., (2008a), hanno utilizzato un sistema di analisi di immagine iperspettrale VIS-NIR, in combinazione con classificazione multivariata non supervisionata (PCA) per l'identificazione del danno meccanico sui funghi coltivati. Lo stesso gruppo, con simili tecniche spettrali, ha analizzato il deterioramento dei funghi durante la conservazione (Gowen et al., 2008b) e il congelamento (Gowen et al., 2009).

Xing et al. (2007) hanno sviluppato un sistema per immagini Vis-NIR (400-1000 nm), testato su diverse varietà (Xing et al., 2005; Xing and De Baerdemaeker, 2005; Xing et al., 2006), per identificare i danni meccanici sulle mele. Un sistema a quattro

lunghezze d'onda multispettrali ha identificato con precisione i danni meccanici sulle mele. Elmasry et al. (2008) hanno sviluppato un sistema operativo nella stessa regione lunghezza d'onda, ma su varietà di mele diverse, e hanno scoperto che, selezionando un sistema a tre lunghezza d'onda multispettrali, possono essere ben distinti i frutti danneggiati da quelli sani.

Per quanto riguarda, invece, la rilevazione di patologie sui prodotti: Qin et al. (2009) hanno sviluppato un sistema di analisi di immagine iperspettrale (450-930 nm) per individuare il cancro degli agrumi in pompelmi, che è una malattia grave che può colpire la buccia della maggior parte delle varietà commerciali di agrumi. L'uso di un metodo di classificazione basato sull'informazione di divergenza spettrale, basato sulla quantificazione delle analogie spettrale utilizzando una predeterminata spettro di riferimento del cancro, ha permesso una corretta classificazione della patologia al 96,2%. Gómez-Sanchis et al. (2008) hanno studiato la fattibilità di individuare il marciume causato da *Penicillium digitatum* in mandarini con un sistema operativo spettrale per immagini nell'intervallo 320-1.100 nm. Lo studio ha concluso che venti era il numero minimo di bande per ottimizzare la classificazione con un valore del 91% di corretta classificazione. Il gruppo di ricerca di Kim et al. (2002) ha sviluppato un sistema da laboratorio con una gamma spettrale di 430-930 nm per analizzare aspetti della qualità e di sicurezza degli alimenti. Il sistema è stato utilizzato in esperimenti di riflettanza per immagini, per il rilevamento di difetti e di contaminanti fecali su superfici di mele (Kim et al, 2002; Liu et al, 2007). Algoritmi semplificati sono stati sviluppati per rilevare i contaminanti, seppure con scarse performance quando i livelli dei contaminanti erano sottili. Park et al. (2006) hanno studiato le prestazioni di un sistema di riflettanza per immagini iperspettrali VIS-NIR utilizzato per il rilevamento superficiale di contaminanti fecali del pollame. Il sistema si è rilevato relativamente efficace e ha consentito l'identificazione delle lunghezze d'onda più informative, predisponendo la costruzione di un sistema di imaging multispettrale. In ulteriori indagini (Park et al., 2007), lo stesso sistema è stato impiegato per identificare il tipo e la fonte dei contaminanti.

In merito alla valutazione della qualità interna dei prodotti agricoli: Polder et al. (2002) hanno dimostrato che un sistema iperspettrale in riflettanza nella regione spettrale 396-736 nm è stato più efficace di un sistema per immagini RGB, di

discriminare grado di maturazione dei pomodori, indipendentemente dalle condizioni di illuminazione. Elmasry et al. (2007) hanno utilizzato un sistema Vis-NIR (400-1000 nm) per la determinazione non distruttiva della qualità di fragola. Le lunghezze d'onda ottimali sono stati ottenute da analisi multivariate in Partial Least Square e poi utilizzate per la stima del contenuto di umidità, contenuto di solidi totali e il pH. Sistemi simili sono stati utilizzati per valutare la qualità e il livello di marezza della carne di maiale (Qiao et al, 2007; Liu et al, 2010). L'uso di immagini iperspettrali nella valutazione della qualità del pesce è stato recentemente riportato in letteratura. Nel successivo capitolo, questo aspetto sarà specificamente approfondito.

La letteratura riportata è solo una parte delle decine di migliaia di referenze complessive, che comprendono sia gli aspetti tecnologici che quelli applicativi nelle scienze della vita e nelle scienze fisiche. Di queste referenze, circa un migliaio sono ascrivibili ad applicazioni nei biosistemi e quindi riferibili alla biofotonica.

L'incremento negli ultimi anni è impressionante e sta ad indicare la crescente importanza della tecnologia in questo settore *high-tech*.

E' atteso inoltre, una sensibile diminuzione dei costi strumentali anche dei sistemi più avanzati, che potrebbero consentire di raggiungere il punto di break-even in merito alla sostenibilità economica all'adozione di tecniche di analisi di immagine multi o iperspettrale anche nell'agroindustria.

4.1.13 Bibliografia

Aleixos, N., Blasco, J., Navarrón, F. and Moltó, E. (2002). Multispectral inspection of citrus in real-time using machine vision and digital signal processors.

Computers and Electronics in Agriculture, 33(2), 121-137.

Anderson J.E., Perry J. E., 1996.Characterization of wetland plant stress using leaf spectral reflectance: Implications for wetland remote sensing. Wetlands 1996 vol. 16, no. 4, pp. 477-487.

Bocchi S., Boschetti M., 2002. Aspetti scientifici e tecnologici dell'agricoltura di precisione. Supplemento a l'Informatore Agrario no.13, pp. 37-43.

- Bricaud A., Morel A., Babin M., Allali K., Claussure H., 1998. Variations of light absorption by suspended particles with chlorophyll a concentration in oceanic (case 1) waters: Analysis and implications for bio-optical models. *J. Geophysical Res. C Oceans* 1998 vol. 47, no. 103, no. C13, pp.31, 033-31, 044.
- Burger, J. and Geladi, P. (2006). Hyperspectral NIR image regression part II: Dataset preprocessing diagnostics. *Journal of Chemometrics*, 20(3), 106-119.
- Burger, J. and Geladi, P. (2007). Spectral pre-treatments of hyperspectral near infrared images: analysis of diffuse reflectance scattering. *Journal of Near Infrared Spectroscopy*, 15(1), 29-38.
- Cheng, X., Chen, Y. R., Tao, Y., Wang, C. Y., Kim, M. S. and Lefcourt, A. M. (2004). A novel integrated PCA and FLD method on hyperspectral image feature extraction for cucumber chilling damage inspection. *Transactions of the ASAE*, 47(4), 1313-1320.
- Chevallier, S., Bertrand, D., Kohler, A. and Courcoux, P. (2006). Application of PLS-DA in multivariate image analysis. *Journal of Chemometrics*, 20(5), 221-229.
- Davies, A. M. C. (2007). Back to basics: spectral pre-treatments - derivatives. *Spectroscopy Europe*, 19(2), 32-33.
- Di Natale C, Davide F, Magagnano A, Cristalli G, D'Amico A, Sberveglieri G, 1997 "Matrici di sensori: Analisi Multicomponente", *Il nuovo saggiaatore* 13, N 4 1997 32-46
- Elmasry, G. and Sun, D.-W. (2010). Principles of hyperspectral imaging technology. In Sun, D.-W. (Eds), *Hyperspectral imaging for food quality analysis and control*, pp. 3-43. Academic Press.
- Elmasry, G., Wang, N. and Vigneault, C. (2009). Detecting chilling injury in Red Delicious apple using hyperspectral imaging and neural networks. *Postharvest Biology and Technology*, 52(1), 1-8.
- Elmasry, G., Wang, N., ElSayed, A. and Ngadi, M. (2007). Hyperspectral imaging for the nondestructive determination of some quality attributes for strawberry. *Journal of Food Engineering*, 81(2), 98-107.
- Elmasry, G., Wang, N., Vigneault, C., Qiao, J. and ElSayed, A. (2008). Early detection of apple bruises on different background colors using hyperspectral imaging. *LWT - Food Science and Technology*, 41(2), 337-345.

- Frouin R., Schwindling M., Deschamps P. Y., 1996. Spectral reflectance of sea foam in the visible and near-infrared: In situ measurements and remote sensing implications. *J. Geophys. Res. C Oceans* 1996 vol. 101, no. C6, pp. 14361-14371.
- Goetz, A. F. H., Vane, G., Solomon, T. E. and Rock, B. N. (1985). Imaging spectrometry for earth remote sensing. *Science*, 228, 1147-1153.
- Gómez-Sanchis, J., Gómez-Chova, L., Aleixos, N., Camps-Valls, G., Montesinos-Herrero, C., Moltó, E. and Blasco, J. (2008). Hyperspectral system for early detection of rottenness caused by *Penicillium digitatum* in mandarins. *Journal of Food Engineering*, 89(1), 80-86.
- Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G. and Frías, J. M. (2007). Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science and Technology*, 18, 590-598.
- Gowen, A. A., O'Donnell, C. P., Taghizadeh, M., Cullen, P. J. and Downey, G. (2008a). Hyperspectral imaging combined with principal component analysis for bruise damage detection on white mushrooms (*Agaricus bisporus*). *Journal of Chemometrics*, 22(3-4), 259-267.
- Gowen, A. A., O'Donnell, C. P., Taghizadeh, M., Gaston, E., O'Gorman, A., Cullen, P. J., Frías, J. M., Esquerre, C. and Downey, G. (2008b). Hyperspectral imaging for the investigation of quality deterioration in sliced mushrooms (*Agaricus bisporus*) during storage. *Sensing and Instrumentation for Food Quality and Safety*, 2(3), 133-143.
- Gowen, A. A., Taghizadeh, M. and O'Donnell, C. P. (2009). Identification of mushrooms subjected to freeze damage using hyperspectral imaging. *Journal of Food Engineering*, 93(1), 7-12.
- Gowen, A. A., Taghizadeh, M. and O'Donnell, C. P. (2010). Using hyperspectral imaging for quality evaluation of mushrooms. In Sun, D.-W. (Eds), *Hyperspectral imaging for food quality analysis and control*, pp. 403-430. Academic Press.
- Guidetti R., Mignani I., Oberti R., 1998. Analisi dell'immagine per la valutazione della qualità dei frutti: misura dell'intensità di fluorescenza come indicatore del grado di maturazione. *Italus Hortus* – vol. 5, n. 5-6, settembre-dicembre 1998.

- Haaland, D. M. and Thomas, E. V. (1988). Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data. *Analytical Chemistry*, 60, 1202-1208.
- Halliday, D. - Resnick, R. - Walker, J. 2001, *Fondamenti di fisica: elettrologia - magnetismo - ottica*. - 5. ed. - Milano: Casa editrice Ambrosiana, 2001
- Hege, E., O'Connell, D., Johnson, W., Basty, S. and Dereniak, E. (2003). Hyperspectral imaging for astronomy and space surveillance. *Proceedings of SPIE*, 5159, 380-391.
- Keefe P. D., Draper S. R., 1986. The measurement of new characters for cultivar identification in wheat using machine vision. *Seed Science and Technology*, 14: 715-724.
- Kim, M. S., Chen, Y. R. and Mehl, P. M. (2001). Hyperspectral reflectance and fluorescence imaging system for food quality and safety. *Transactions of the ASAE*, 44(3), 721-729.
- Kim, M. S., Lefcourt, A. M., Chao, K., Chen, Y. R., Kim, I. and Chan, D. E. (2002). Multispectral detection of fecal contamination on apples based on hyperspectral imagery: Part I. Application of visible and near-infrared reflectance imaging. *Transactions of the ASAE*, 45(6), 2027-2037.
- Lamb, D.W. (2000). The use of qualitative airborne multispectral imaging for managing agricultural crops – A case study in south eastern Australia. *Australian Journal of Experimental Agriculture*, 40: 725-38.
- Liu, L., Ngadi, M. O., Prasher, S. O. and Gariépy, C. (2010). Categorization of pork quality using Gabor filter-based hyperspectral imaging technology. *Journal of Food Engineering*, 99(3), 284-193.
- Liu, Y., Chen, Y.-R., Kim, M. S., Chan, D. E. and Lefcourt, A. M. (2007). Development of simple algorithms for the detection of fecal contaminants on apples from visible/near infrared hyperspectral reflectance imaging. *Journal of Food Engineering*, 81(2), 412-418.
- Luoheng H. and Donald C. Rundquist, 1997. Comparison of NIR/RED ratio derivative of reflectance in estimating algal-chlorophyll concentration: a case study in turbid reservoir. *Rem. Sens. Environ.* 62: 253-261.

- Menesatti P, Costa C, Aguzzi J, 2010. Quality evaluation of fish by hyperspectral imaging. In: Hyperspectral imaging for food quality: analysis and control. In D.-W. Sun (Ed.), Hyperspectral imaging for food quality analysis and control. London, Burlington, San Diego: ACADEMIC PRESS, ELSEVIER: 273-294. ISBN: 9780123747532
- Menesatti P, Zanella A, D'Andrea S, Costa C, Paglia G, Pallottino F, 2009, Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples. FOOD BIOPROCESS TECHNOL 2:308–314 - DOI 10.1007/s11947-008-0120-8
- Menesatti P. 2000. Misura strumentale di aspetti della qualità degli alimenti degli alimenti attraverso tecnologie opto-elettroniche. Industrie Alimentari – XXXIX (2000) febbraio.
- Muirhead D.L., 1998. Color Measurement and Reflectance Spectroscopy of Total Suspend Solids for Remote sensing and Environmental Monitoring of Aquatic Ecosystems. Dissertation Abstracts International Part B: Science and Engineering 1998 vol. 59, no. 6, p. 2971.
- Naganathana, G. K., Grimesb, L. M., Subbiaha, J., Calkinsb, C. R., Samalc, A. and Meyera, G. E. (2008). Visible/near-infrared hyperspectral imaging for beef tenderness prediction. Computers and Electronics in Agriculture, 64(2), 225-233.
- Nicolaï, B. M., Lötze, E., Peirs, A., Scheerlinck, N. and Theron, K. I. (2006). Non-destructive measurement of bitter pit in apple fruit using NIR hyperspectral imaging. Postharvest Biology and Technology, 40(1), 1-6.
- Oleari Claudio (a cura di). 1998. Misurare il colore. Milano. Hoepli.
- Osborne, S. D., Jordan, R. B. and Künnemeyera, R. (1997). Method of wavelength selection for partial least squares. Analyst, 122, 1531-1537.
- Park B., K.C. Lawrence, W.R. Windham, D.P. Smith and P.W. 2002, Feldner, Hyperspectral imaging for food processing automation, Proceedings of the SPIE, Volume 4816 (2002), pp. 308-16
- Park, B., Lawrence, K. C., Windham, W. R. and Smith, D. P. (2006). Performance of hyperspectral imaging system for poultry surface fecal contaminant detection. Journal of Food Engineering, 75(3), 340-348.

- Park, B., Windham, W. R., Lawrence, K. C. and Smith, D. (2007). Contaminant classification of poultry hyperspectral imagery using a spectral angle mapper algorithm. *Biosystems Engineering*, 96(3), 323-333
- Pasqualini, -V.; Pergent-Martini, -C.; Pergent, -G., 1996. The use of aerial teledection for the characterization of the Mediterranean coastal environment: *Posidonia oceanica*. *Proceedings-of-the-2nd-international-conference-on-the-mediterranean-coastal-environment-medcoast-95,-october-24-27-1995,-tarragona,-spain*. Ozhan-E.-ed. 1996.
- Polder, G., Heijden, G. and Young, I. (2002). Spectral image analysis for measuring ripeness of tomatoes. *Transactions of the ASAE*, 45, 1155-1161.
- Qiao, J., Ngadi, M., Wang, N., Gariépy, C. and Prasher, S. (2007). Pork quality and marbling level assessment using a hyperspectral imaging system. *Journal of Food Engineering*, 83(1), 10-16.
- Qin, J. (2010). Hyperspectral imaging instruments. In Sun, D.-W. (Eds), *Hyperspectral imaging for food quality analysis and control*, pp. 129-172. Academic Press.
- Qin, J., Burks, T. F., Ritenour, M. A. and Bonn, W. G. (2009). Detection of citrus canker using hyperspectral reflectance umaging with spectral information divergence. *Journal of Food Engineering*, 93(2), 183-191.
- Urbani G., Menesatti P., Solaini S. 2002. Tecnologie opto-elettroniche in agricoltura: nuove prospettive e indirizzi di ricerca. *Mondo Macchina*. N. 6/2002. pp. 35-38.
- Ustin S. L, Curtiss B. & Martens S., 1998. Remote sensing of forest Condition in the western U.S., EPA Forest Response Program Annual Report. Corpus Christrie Texas, 1988.
- Vaarala T, M. Aikio and H. Keraenen, An advanced prism-grating-prism imaging spectrograph in online industrial applications, *Proceedings of the SPIE*, Volume 3101 (1997), pp. 322-30.
- Xing, J. and De Baerdemaeker, J. (2005). Bruise detection on 'Jonagold' apples using hyperspectral imaging. *Postharvest Biology and Technology*, 37(2), 152-162.
- Xing, J., Bravo, C., Jancsó, P., Ramon, H. and De Baerdemaeker, J. (2005). Detecting bruises on 'Golden Delicious' apples using hyperspectral imaging with multiple wavebands. *Biosystems Engineering*, 90(1), 27-36.

- Xing, J., Jancsó, P. and De Baerdemaeker, J. (2006). Stem-end/calyx identification on apples using contour analysis in multispectral images. *Biosystems Engineering*, 96(2), 231-237.
- Xing, J., Saeys, W. and De Baerdemaeker, D. (2007). Combination of chemometric tools and image processing for bruise detection on apples. *Computers and Electronics in Agriculture*, 56(1), 1-13.
- Yacobi Y.Z., Gitelson A., Mayo M., 1995. Remote sensing of chlorophyll in Lake Kinneret using-spectral-resolution radiometer and Landsat TM: Spectral features of reflectance and algorithm development. *J. Plankton Res.* 1995 vol. 17, no. 11, pp. 2155-2173.
- Zavattini G., S. Vecchi, R.M. Leahy, D.J. Smith, D.J., and S.R Cherry, A hyperspectral fluorescence imaging system for biological applications, *IEEE Nuclear Science Symposium*, Piscataway, USA, Oct. 19-25, 2003
- Zwiggelaar, R., Yang, Q., Garcia-Pardo, E. and Bull, C. R. R. (1996). Use of spectral information and machine vision for bruise detection on peaches and apricots. *Journal of Agricultural Engineering Research*, 63(4), 323-331.

4.2. Pubblicazioni presentate

Nella presente sezione sono presentati due lavori scientifici, di cui il dottorando è coautore, inerenti review e applicazioni sperimentali dell'analisi di immagine iperspettrale in differenti contesti agroalimentari.

Si riportano di seguito gli abstract delle singole pubblicazioni:

4.2.1 ABSTRACT

Quality evaluation of fish by hyperspectral imaging

Menesatti P, Costa C, Aguzzi J, 2010. Quality evaluation of fish by hyperspectral imaging. In: Hyperspectral imaging for food quality: analysis and control. In D.-W. Sun (Ed.), Hyperspectral imaging for food quality analysis and control. London, Burlington, San Diego: ACADEMIC PRESS, ELSEVIER: 273-294. ISBN: 9780123747532

We will discuss the use of hyperspectral imaging as a method to provide an objective and qualitative evaluation of the state of the fish freshness. We focus on establishing a correlation between the spectral reflectance of selected areas of the epidermis and the time of storage in standard refrigeration procedures. We will also discuss the possibility of finding objective parameters for the good prediction of fish freshness that considers products stored for more than three days as “non-fresh but still edible”. Case studies corresponding to two different analytic procedures are described, including subjective ROI (region of interest) identification in hyperspectral images and morphometric superimposition for automated topographical hyperspectral image analysis. The first method is based on the subjective choice of the sampling areas within the hyperspectral images that bear the most interesting information according to a subjective criterion of observed evaluation. The operator can delimit the region to be analysed. For that region, an average value of spectral reflectance can be computed within the VIS-NIR or NIR range. The second method is based on the first method, and it represents an evolution of the technique, with the use of geometric

morphometric tools for the superimposition of hyperspectral cubes from image pixels of different samples.

The first technique used obtaining an elevated degree in the assessment of fish freshness. The second technique used represented an important methodological evolution of the first one technique. Based on combined hyperspectral and geometric morphometric techniques, spectral information from pixels was associated with their topographic location for the first time. This novel approach is based on the *a priori* discrimination of which wavelength areas are more discriminating in relation to fish freshness, considering the unity of the fish product as spectrally inhomogeneous. In the second case study the proposed technique represents an important methodological development by combining hyperspectral imaging and geometric morphometric tools. This technique was applied in the hyperspectral field, resulting in a new innovation allowing the association of topologic spectral information. Notwithstanding, an automated method for the extraction of the fish outline can be implemented in the next future.

Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples

Menesatti P, Zanella A, D'Andrea S, Costa C, Paglia G, Pallottino F, 2009, Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples. FOOD BIOPROCESS TECHNOL 2:308–314 - DOI 10.1007/s11947-008-0120-8

Fruit maturity indexes are crucial in harvest time determination and commercial context. The harvest time of apples, matching the desired commercial characteristics, is assessed through starch-iodine test in practice. Fruit halves are dipped into iodine solution and patterns are visually evaluated by experts comparing them to reference charts. Aim of the work was to study the relationships of near infrared (NIR) spectral images (1000-1700 nm), starch/starch-free patterns visually assessed, and RGB color images. Spectral images of 88 Golden Delicious Klon B apples were sampled at 7

different maturity stages. Partial least-squares discriminant analysis (PLSDA) technique was used on hyperspectral NIR images to classify single pixels using its NIR reflectance spectrum. The response variable (i.e. the classification for each pixel) was identified through the matching of single pixel obtained with the color images, segmented in two classes (starch and starch-free), and the NIR hyperspectral matrix. Mean hyperspectral classification obtained through PLSDA modeling on individual apple correctly classified 80.81% of the total pixels, while the unique model, i.e. a single model including all the fruits, resulted in 66.33%. In the latter case, the relationship with the RGB classification showed high values (Pearson correlation coefficient $r = 0.95$). The present work shows the feasibility of NIR imaging spectroscopy as a tool for apple fruit maturity determination, avoiding expert's subjective interpretation by traditional starch index assignments.

4.3. Quality evaluation of fish by hyperspectral imaging

FULL REFERENCE OF THE PUBLICATION

Menesatti P, Costa C, Aguzzi J, 2010. Quality evaluation of fish by hyperspectral imaging. In: Hyperspectral imaging for food quality: analysis and control. In D.-W. Sun (Ed.), Hyperspectral imaging for food quality analysis and control. London, Burlington, San Diego: ACADEMIC PRESS, ELSEVIER: 273-294. ISBN: 9780123747532

Introduction

Quality is an important factor in enhancing competitiveness in agricultural or fish production. The concept of quality is related to safety, nutritional or nutraceutical value, and to organoleptic properties such as freshness. In order to ensure the appropriate food quality and safety for the health of consumers, legal requirements and new quality standards are constantly developed according to EU Directives (Knaflawska & Pospiech, 2007). Especially in Europe, there is an increasing interest in labeling the quality of agro-fish products for human consumption.

Quality evaluation, therefore, progressively began to be a central aspect in agro-food and fish production and industrial processing. In this context, it is important to consider that the term "quality" in commercial, scientific, and the related legislation fields may refer to different aspects for different operators. Moreover, the actual trend is to relate "quality" to each specific product type (species, origin, rearing technique) and each individual organism (Costa *et al.*, 2009c). For example, chemical composition differences in fish flesh between wild and farmed sea bass from Greece and Italy have been reported by Alasalvar *et al.* (2002) and Orban *et al.* (2002). According to this example, it is important to find efficient analytic methods to attribute a differential quality to captured or farmed stock fish with poorer meat quality.

In relation to fish products coming from aquaculture facilities or commercial fisheries, each category of product is characterized by size, shape, colour, freshness,

and finally by the absence of visual morphological defects (Costa *et al.*, 2009c). In particular, the appearance is an easily treatable criterion utilized to select the piece of product throughout the market chain from its production to its storage, marketing, and finally, to users (Kays, 1999). In that context, an important aspect of quality is related to the concept of freshness.

Freshness, in relation to fish quality, represents a pivotal aspect in its socioeconomic usage and economic value. Scientific methods for the evaluation of freshness may be conveniently divided into two categories: sensory and instrumental. Since the consumer is the ultimate judge of quality, most methods must be correlated with measures related to sight, touch or odour perception (Menesatti *et al.*, 2002, 2006; Menesatti, Urbani *et al.*, 2007). While sensory-based methods of measurements must be performed under carefully controlled scientific conditions in order to allow a trustable reproduction of results so that the effects of the testing environment, personal bias, etc., can be reduced (Huss, 1995), instrumental techniques are less subjective. The bias introduced by observer-based evaluation, as well as the surrounding environment, is comparatively higher.

Instrumental methods can be divided into biochemical–chemical methods, microbiological methods, and physical methods (Menesatti, Urbani *et al.*, 2007). The appeal of biochemical–chemical methods for fish quality evaluation is related to their ability to establish quantitative standards for freshness based on tolerance levels in chemical spoilage. The aim of microbiological examinations of fish products is to evaluate the possible presence of bacteria that are of public health concern and to maintain hygienic quality in terms of temperature and cleanness during handling and processing. Methods of a microbiological nature are correlated with sensory quality evaluations of chemical compounds in relation to spoilage or to modifications associated with the industrial processing itself (e.g., the breakdown of amines or nucleotides in the canning process as a result of high temperatures). The microbiological aspects affecting fish quality are mostly related to public health and the obtained data on its assessment do not provide information about freshness. Finally, there are physical methods that are based on the testing of softness/hardness of food texture. These methods are particularly appreciated for their rapid and non-destructive approach. One method in particular is based on the changes in the

electrical properties of skin and tissue after fish death (Jason & Richards, 1975). Changes in conductance properties are associated with variations in meat quality post mortem in relation to bacterial spoilage. Also, the evaluation of firmness can be considered as an indicator of good quality, with a good correlation with sensory and chemical properties (Alasalvar *et al.*, 2001; Menesatti, Pallottino *et al.*, 2009).

Texture in fish meat can be instrumentally measured by techniques based on puncturing, compression, cutting, or stretching (i.e., tension) (Menesatti & Urbani, 2004). Among all of these, the most widely used techniques are cutting force and compression (Sigurgisladottir *et al.*, 1999).

The analysis of the optical properties of food has recently been assuming a greater relevance in product and organoleptic assessments of quality within the physical methods. Spectrophotometric applications are particularly relevant in the outputs from meat under a light beam of a single spectral band and the components contribute to a more detailed and refined evaluation of quality characteristics, giving important indicators on the method of production of the meat and its origin (Menesatti, D'Andrea *et al.*, 2007). Spectrophotometric techniques are associated with a high analytic ability based on their non-destructiveness, relative simplicity, speed, and portability in operative environments during measurements. Also, the high level of informatization and hardware development (in terms of interfacing with other instrumental sensors) have progressively shifted scientific and applied interest in quality assessment procedures in meat production toward spectrophotometric applications.

In recent years near-infrared reflectance spectroscopy (NIRS) has been increasingly used as a non-destructive and rapid technique in the assessment of food quality (Chen & He, 2007; Xiccato *et al.*, 2004) and associated health issues (Kim *et al.*, 2002).

NIRS has been also used in fish meat quality tests. The spectral variation among fish meat samples depends on the feeding regime of the fishes, as well as water quality, growth pattern, and muscular activity (Karoui *et al.*, 2007). NIRS has been successfully used in salmon, trout, cod, halibut, and sea bass in relation to chemical composition prediction, protein content, and levels of humidity (i.e., moisture) (Cozzolino *et al.*, 2002; Mathias *et al.*, 1987; Nortvedt *et al.*, 1998; Solberg & Fredriksen, 2001; Xiccato *et al.*, 2004).

Particular spectroscopic applications are conducted with advanced technologies of VIS/NIR and NIR spectral imaging. These instruments are able to acquire spectral images at a high-density resolution (150–250 k-pixels) where each pixel possesses the entirety of the spectral information (VIS and NIR). Thus, this technique integrates conventional imaging and spectroscopy to obtain both spatial and spectral information from an object (Gowen *et al.*, 2007; Menesatti, Zanella *et al.*, 2009).

Multi- or hyperspectral analysis (≤ 10 and > 10 spectral bands, respectively) is the new frontier of optical imaging. Hyperspectral imaging, within the VIS/NIRS techniques, is useful to analyse the spectra of inhomogeneous materials that contain a wide range of spectral (Mehl *et al.*, 2002) and spatial information (Park *et al.*, 2006).

Hyperspectral images can be considered as hypercube matrices; three-dimensional blocks of data made by two spatial plane coordinates and one wavelength dimension (Gowen *et al.*, 2007).

Multi- and hyperspectral optical imaging has been successfully used in vegetable and meat quality discrimination in recent years because of its high capability for the detailed analysis of food product structure (Menesatti, D'Andrea *et al.*, 2007). In fruit post-harvesting treatment, this technique has been successfully used for the detection of quality defects in cucumbers, tomatoes, pears, and apples (e.g., Li *et al.*, 2002; Liu *et al.*, 2006; Polder *et al.*, 2002). It has also been used in food applications in relation to the biochemical properties of sugar contents (Bellon *et al.*, 1993), moisture content (Katayama *et al.*, 1996), and acidity (Lammertyn *et al.*, 1998).

Referring to the use of hyperspectral imaging in fish production, few applications have been reported in the literature to date. Published data mostly refer to the detection of fat and water content in fish fillet products (ElMasry & Wold, 2008), for production line sampling (Wold *et al.*, 2006) or for fish freshness detection (Menesatti, Urbani *et al.*, 2007). In the near future, however, the technological evolutions of photonics will reach a break-even point where the spectroscopic technology could be broadly adopted given its low price and safety of use (Menesatti, D'Andrea *et al.*, 2007; Park *et al.*, 2004; Yang *et al.*, 2005). In this context, it is possible that hyperspectral imaging will provide a valid contribution in relation to the monitoring of the organoleptic and merceologic properties of fish production during all steps along the production chain.

In this chapter we will discuss the use of hyperspectral imaging as a method to provide an objective and qualitative evaluation of fish freshness. We focus on establishing a correlation between the spectral reflectance of selected areas of the epidermis and the time of storage in standard refrigeration procedures. We will also discuss the possibility of finding objective parameters for the good prediction of fish freshness that consider products stored for more than three days as “non-fresh but still edible”.

Case studies corresponding to two different analytic procedures will be described, including subjective ROI (region of interest) identification in hyperspectral images and morphometric superimposition for automated topographical hyperspectral image analysis. The first method is based on the subjective choice of the sampling areas within the hyperspectral images that bear the most interesting information according to a subjective criterion of observed evaluation. The operator can delimit the region to be analysed. For that region, an average value of spectral reflectance can be computed within the VIS/NIR or NIR range. The second method is based on the first method, and it represents an evolution of the technique, with the use of geometric morphometric tools for the superimposition of hyperspectral cubes from image pixels of different samples.

Subjective ROI on Hyperspectral Images for Fish Freshness Identification

Fifty wild chub mackerel (*Scomber japonicus*) and 80 hatchery-reared sea bass (*Dicentrarchus labrax*) were used in this case study. Chub mackerel were fished in the mid-low Adriatic Sea (Manfredonia). Sea bass were cultured intensively in concrete tanks (CT) or in sea cages (SC). All the fish were collected in May 2005 from three fish farms in the same southern Italian region (Puglia). Their rearing conditions were as follow:

- Peschiere Tarantine farm: CT with a stocking condition at $19\text{ }^{\circ}\text{C}$ with a density of 50 kg/m^3 .
- Panittica Pugliese farm: Both CT and SC; in CT with a stocking condition at $19.5\text{ }^{\circ}\text{C}$ with a density of 35 kg/m^3 .
- Tortuga farm: SC (Ionian Sea).

Before harvesting, all fish were fasted for 24 hours. They were killed by immersion in chilled water and then covered with ice. All fish were transported to CRA-ING (Monterotondo, Rome) in a refrigerated unit maintaining a constant temperature of between 0 and 4{ts}°C. Specimens were analysed at 1, 2, 4, and 6 days post mortem (d.p.m.).

Before the spectral scanner analysis the fish were taken from the refrigerator and left at room temperature for 30 min to eliminate the dry film on their skin created by the freezing process. A hyperspectral imaging system was used to integrate the spectroscopic and spatial imaging information of the fish. This system, in addition to the spatial information, can provide information at hyper/multiple wavelengths for each pixel of the sample.

The hyperspectral system used was composed of four parts (Figure 1): a sample transportation plate (spectral scanner DV, Padua, Italy); a collimated illumination device (Fiber-lite, Dolan-Jenner, USA) composed of one 150W halogen lamp as the source light and one illumination opening in the optical fibre measuring 200{ts}mm long and 2{ts}mm width, positioned at 45° in relation to the transportation plate (i.e., bearing the sample) and presenting a minimum light divergence; and an imaging spectrograph (ImSpec V10, Specim Ltd, Oulu, Finland) coupled with a standard C-mount zoom lens and a Teli CCD monochrome camera (Toshiba-Teli CS8310BC, Japan).

Figure 1: The VIS-NIR hyperspectral system used.



The ImSpec is based on a patented prism-grating-prism (PGP) construction (a holographic transmission grating). The incoming line image (frame) is projected and dispersed onto the charge-coupled device (CCD). Each frame contained the line pixels in one dimension (spatial axis) and the spectral pixels in the other dimension (spectral axis), providing full spectral information for each line pixel. The reconstruction of the entire hyperspectral image of the sample was performed by scanning the sample line-by-line as the transportation plate moved it through the field of view. The resolution of the line image was 700 pixels by 10 bits. The number of frames (image resolution in Y-axes) was variable, from 10 to 500, depending on the speed and the accuracy of the transportation plate line scanning. The system was operated in a dark laboratory to minimize interference from ambient light. Other basic characteristics of the system were: spectral range, 400–970 nm; spectral

resolution, 5 {ts}nm; dispersion, 90.9 {ts}nm/mm; sensor image size, 6.6 (spectral) × 8.8 (spatial) mm, corresponding to standard 2/3 {ts}in. image sensor; spatial resolution, 15 line-pairs/mm; rms spot radius <60 {ts}μm within 2/3 image area; aberrations, insignificant astigmatism; slit width, 25; effective slit length, 9.8 {ts}mm; total efficiency (typical) >50%; and it was independent of polarization.

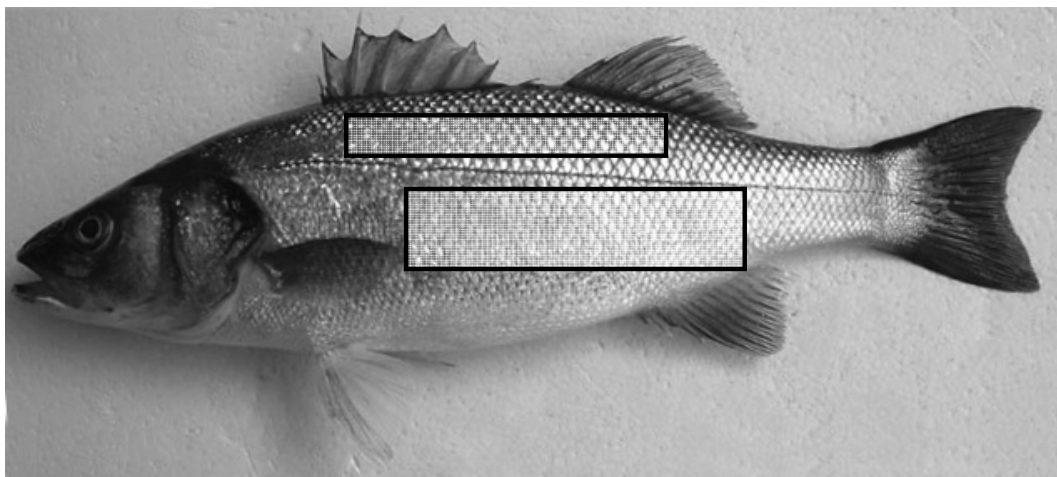
Spectral values were expressed in terms of relative reflectance (R), by applying the following equation:

$$R = \frac{r_s - r_b}{r_w - r_b} \quad (1)$$

where R is the relative reflectance of the sample at each wavelength; r_s is the absolute signal value (radiance) measured for the sample at each wavelength; r_b is the absolute signal value (radiance) measured at each wavelength for black (background noise); and r_w is absolute signal value (radiance) measured at each wavelength for the standard white (100% of reflectance).

Hyperspectral images of the lateral side of the fish (from the area under the attachment of the first dorsal fin to the area above the end of the anal fin) were analysed with the software Spectral Scanner (ver. 1.4.1) (DV Optics, Padua, Italy). On each hyperspectral image a trained operator selected two ROIs (Figure 2) to measure the mean VIS/NIR spectral reflectance.

Figure 2: Examples of ROIs applied by operators for the spectral image analysis.



A supervised multivariate technique such as partial least squares discriminant analysis (PLSDA) was applied to observe freshness differences (<3d.p.m. vs. >3d.p.m.) in relation to mean spectral reflectance values. The PLSDA (Sabatier *et al.*, 2003; Sjöström *et al.*, 1986) consists of a classic partial least squares (PLS) analysis regression where the response variable is a categorical one (Y-block; replaced by the set of dummy variables describing the categories) expressing the class membership of the statistical units (Aguzzi *et al.*, 2009; Costa *et al.*, 2008; 2009b). The PLSDA does not allow for response variables other than those that define the groups of individuals, fresh (<3 d.p.m.) or non-fresh (>3d.p.m.). The model includes a calibration phase and a cross-validation phase; during both phases the percentages of correct classification were calculated. The prediction ability in the test phase also depends on the number of latent variables (LV) used in the model. The optimal number of LV was chosen on the basis of the highest percentage of correct classification. The PLSDA analysis provides the percentage of correct classification of the entire model as well as for the two classes considered. This analysis was performed using Matlab 7.1 (The Math Works, Natick, USA) and PLS Toolbox 4.0 (Eigenvector Research Inc, Wenatchee, USA) for all the combinations of different preprocessing treatments (none, autoscale, mean centre, Savitzky–Golay, ECC) and LV (2-20).

Figures 8.3 and 8.4 show the mean values in spectral reflectance at different d.p.m. (1, 2, 4, and 6) for chub mackerel and sea bass, respectively. In Figure 8.3 it is possible to observe that the mean reflectance values for fresh individuals (<3 d.p.m) that are between 450 and 600{ts}nm are well separated from values for the others. In Figure 8.4 (at 6 d.p.m.) the spectral reflectance values are always different from and lower than all the others. This first result shows how spectral information can be used to discriminate freshness status, although such information cannot be used for the topographic evaluation of which areas are more informative than others in the determination of freshness status. The selection of ROI is still subject to operator evaluation, its automation being at present difficult if the spectral-topographic contribution of the different areas is not established.

Figure 3. Average spectral reflectance over consecutive days of freezing conservation as measured on Chub mackerel side. d.p.m., days post mortem.

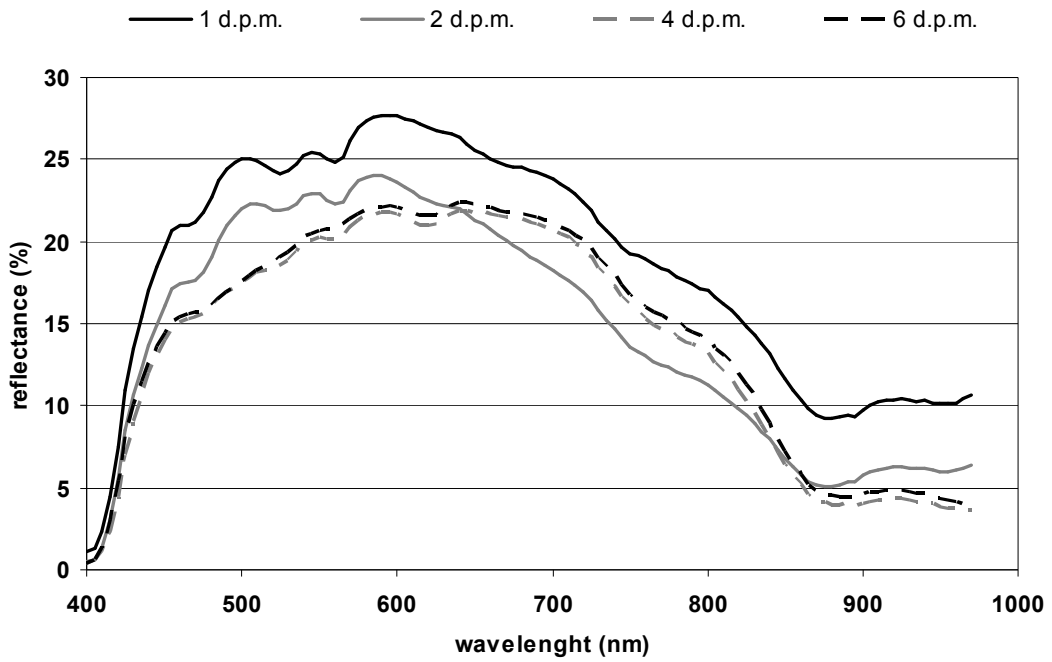
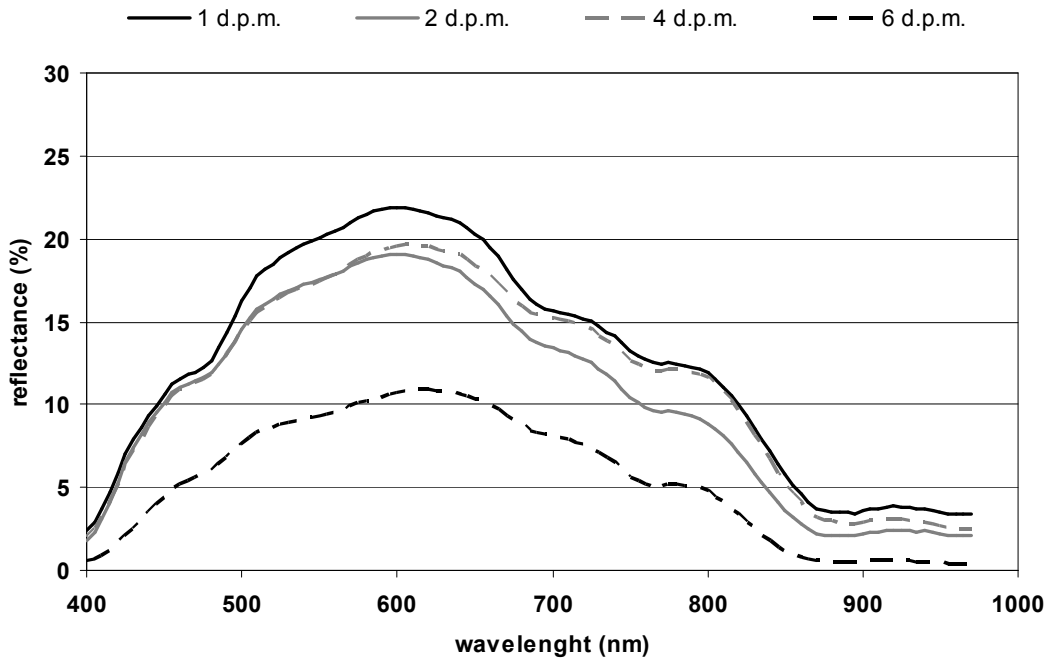


Figure 4: Average spectral reflectance over consecutive days of freezing conservation ass measured on Sea bass side. d.p.m., days post mortem.



The present results agree well with observations on the progressive freshness deterioration of fish after death within 6 days. From the results, the skin brilliancy is a major element influencing the spectral analysis. This confirms the role of skin brilliancy in quality judgment on the integrity of the product. In fact, it is possible to notice a consistent reduction of spectral reflectance values within 450–650{ts}nm in samples with more days of frozen conservation. Also, a modification in the spectral quality response occurs since the reflectance curve is more smoothed in fish that is less fresh. A similar trend, although with more variation, was also observed in the sea bass over a longer period of conservation time.

Results of the two PLSDA models built for the two studied species are reported in Table 1.

Table 1. Characteristics and principal results of the two PLSDA models built from Chub mackerels and Sea bass reflectance data.

	Chub mackerel	Sea bass
N	50	80
n° units (x-block)	101	101
n° units (y-block)	2	2
Preprocessing	none	None
cross validation	leave one out	venetian blinds
n° LV	4	6
Cumulated Variance X-block (%)	99.7	99.9
Cumulated Variance Y-block (%)	25.3	22.7
Mean RMSEC	0.605	0.591
Mean RMSECV	0.628	0.626
% Corr. Class.	88.0	82.5

Note: N is the number of samples; n° units (Y-Block) is the number of units (fresh \leq 3d.p.m.; non fresh $>$ 3d.p.m.) to be discriminated by the PLSDA, and n° LV is the number of latent vectors for each model.

A high percentage of correct classification between fresh (≤ 3 d.p.m.) and non fresh (> 3 d.p.m.) fish is reported at values of 88% for chub mackerels and 82.5% for sea bass. With the use of a multivariate approach on the hyperspectral results we efficiently show the ordination of samples according to the number of days after the death of the fish.

The present data indicate that better performing models possess lower numbers of LV (i.e., 4 and 6 for the chub mackerel and the sea bass, respectively). These samples were also those that did not undergo preprocessing. Our results can be explained by assuming that differences in the reported values of spectral reflectance intensity are sufficiently neat, especially in the range of visible wavelengths. Also, a relatively more simplified PSLDA model can still discriminate well among classes. The highest value of correct classification of the chub mackerel in comparison to the sea bass should be attributed to a clearer distinction in average spectral reflectance among samples of both considered classes. The variability in the spectral response is evidenced by the fact that the lowest value of cumulated variance in the Y-block depends not only upon samples' variability but also on the subjectivity in the ROI selection. Subjectivity problems in ROI selection were treated by Peirs *et al.* (2002), who found different ROI values depending on observer attribution.

Morphometric Superimposition for Topographical Fish Freshness Comparison

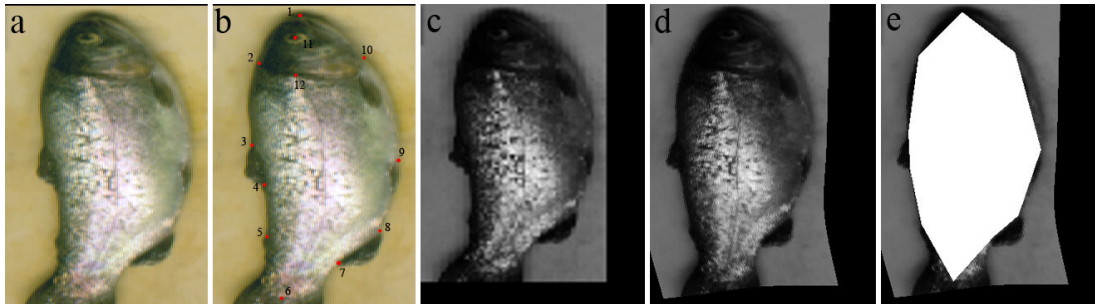
In order to limit analytic errors in hyperspectral evaluation given the subjective choice of areas by the operator, an automatic topographic approach was developed. This represents a forward step in the analysis of quality, the importance of which has not been studied. In this case study, five specimens of rainbow trout (*Oncorhynchus mykiss* Walbaum, 1792) were used that came from Azienda Agricola Sterpo (Rivignano, North-Eastern Italy). After collection the fish were killed by immersion in water and ice and then stored in refrigerated tanks for the duration of their transport to the laboratory facilities (CRA-ING, Rome). In the laboratory the fish were stored according to traditional market techniques, i.e., in an industrial refrigerator at $2\text{ }^{\circ}\text{C}$, and in polystyrene boxes with holes on all sides. The fish were also covered by ice both beneath and on top. Direct contact with ice, which causes potential damage to fish tissues, was prevented by using plastic parafilm. Each single trout was

used three times at T0 = 1 d.p.m. (days post mortem), T1 = 3 d.p.m., T2 = 7 d.p.m., and finally T3 = 10 d.p.m. Before use in the spectral scanner, fish were taken from the refrigerator and left at room temperature for 30 min to eliminate the dry ice film on the fish surface.

The trout were scanned with the same spectral system used in the previous case study (see Figure 1). The hyperspectral VIS/NIR image acquisition time lasted about 8{s}. For each acquired pixel in each image wavelength layer, the spectral reflectance value was measured and computed accordingly to Equation (1).

The image-warping protocol adapted for spectral matrixes was used to superimpose the RGB images of all sampled individuals taken on four different occasions (d.p.m.) (Costa *et al.*, 2009a). Images were warped to a standard view by fixing a set of reference points on the surfaces of the animal body. Using this method, the shape and color pattern of each individual was morphologically adjusted to the shape of the consensus configuration of the entire sample, as calculated via geometric morphometric tools. Geometric morphometric methods were developed to quantify and visualize deformations of morphometric points (landmarks) in a coordinate space of reference, as conceptualized by D'Arcy Thompson (1917). Landmarks are defined as homologous points that bear information on the geometry of biological forms (Bookstein, 1991). Using the consensus configuration of all specimens as the starting form, landmark configurations for each individual were aligned, translated, rotated, and scaled to a unit centroid size by the generalized Procrustes analysis (GPA) (Rohlf & Slice, 1990). Residuals from the fitting were modeled with the thin-plate spline interpolating function (Antonucci *et al.*, 2009; Bookstein, 1991; Costa *et al.*, 2006; Rohlf & Bookstein, 1990; Rohlf & Slice, 1990; Zelditch, *et al.*, 2004). This warping procedure involves standardizing the shape and size of each wavelength layer image with a generalized orthogonal least-squares Procrustes (GPA) superimposition (translation, scaling, and rotation) conducted on the set of 12 reference points (Figure 5b) (Rohlf, 1999).

Figure 5: Comparison of original fish image with various processed images. a) original image; b) image with 12 landmarks; c) image of one hyperspectral layer (650 nm); d) image after the warping procedure; e) ROI based on the first 10 landmarks (ROI_L) positioned on the outline of the fish..



A supervised multivariate classification technique, such as PLSDA, was used to observe freshness differences (<4 d.p.m. vs. >4 d.p.m.). Such an approach has never before been used in similar studies with hyperspectral methodology. Three different multivariate classification approaches (i.e., AP1, AP2, and AP3) were used for this occasion:

AP1: In order to evaluate the ROI topographic positioning based on the first 10 landmarks (ROI_L) and on the contribution of selected wavelengths, a data set was built for each tested individual at each time of sampling (d.p.m.) by considering as X-block variable of each pixel for each wavelength layer at its topographic position. In order to reduce the matrix dimension the images were resized 3641 pixels (i.e., 1:0.3) and the number of wavelengths considered was 61 (500–800{ts}nm; step-frequency: 5{ts}nm).

AP2: In order to verify the classification capacity of this system each ROI_L pixel from each image of the fish at different d.p.m. was individually classified based on the dichotomic categorization “fresh”/“not-fresh” valid for the entire fish.

AP3: The proportion of ROI_L (small ROI_L named ROI_S) and the wavelengths resulted as more informative (620-780nm) based on AP1, was therefore applied using AP2. The results indicate that the consensus ROI_L encompassed 40{ts}420 pixels when images were not rescaled. For a rescaling equal to 1:0.3, the ROI_L encompassed 3641 pixels for each tested fish. All of these pixels present reflectance values for at least 61

wavelengths. Taken together, these represent a great quantity of data for each individual tested fish to be classified upon. This result is important since, for example, Farzam *et al.* (2008) showed how hyperspectral methodology requires huge calculation resources for data treatment.

Based on AP1, which considers both the spectral values of each pixel as well as its topographic position, more than 20{ts}000 variables were obtained for each of the individuals, to be considered within the X-block ($3,641 * 61 = 222,101$). Based on AP2, by considering each pixel within the ROI_L, 72{ts}820 samples were obtained. With AP3, by reducing the ROI_L at ROI_S and by considering only 17 variables, the samples number was reduced to 5580 (Table 2). Data reduction is important for data treatment in quality analytic procedures (Farzam *et al.*, 2008). Accordingly, the present result set is based on considering this discrimination important for the industrial processing of cultured fish. These findings represent the first step in filling a gap in the existing technology in food industrial processing, as already identified by Menesatti, Zanella *et al.* (2009).

The PLSDA results based on the three proposed approaches are reported in Table 2.

Table 2. Characteristics and principal results of the PLSDA models. N is the number of samples.

	AP1	AP2	AP3
N samples	20	72820	5580
N units (X-block)	222101	61	33
N units (Y-block)	2	2	2
N LV	10	16	9
Pre-processing X-block	Baseline	Median center	Median center
Cumulated Variance X-block (%)	95.28	99.95	99.91
Cumulated Variance Y-block (%)	99.99	51.42	19.29
Mean Sensitivity (%)	1	55.4	81.0
Mean Specificity (%)	1	63.2	77.9
Mean classification error	0	0.39794	0.20645
Mean RMSEC	0.0061	0.49251	0.63524
Random Probability (%)	50	50	50
Corr. Class. Model (%)	100	66.77	79.4

Note: N is the number of samples; n° is the number of units (Y-Block) to be discriminated by the PLSDA; n° LV is the number of latent vectors for each model; and Random Probability (%) is the probability of random assignment of an individual into a unit.

As expected, AP1 reached a percentage of individuals correct classified of 100, since it is based on many variables and comparatively few samples. The AP2 and AP3 methods present a percentage of single pixel correct classification that is high (66.77% for AP2 and 79.40% for AP3). The selection of ROIs and the concomitant reduction of the most significant variables, as in AP3, created an interesting increment in the percentage of correct pixel classification (from 66.77% in AP2, to 79.4% in AP3).

Considering AP1, the first LV explains the major variance for X-block and Y, achieving 83.62% and 48.54%, respectively. For this reason, we use the loadings for each pixel of LV1 to observe the main topographic contribution of the wavelengths.

In Figure 6 the graphic output of AP1 is shown for the selected ROI_L, with the loading contribution for LV1 exceeding the 90th percentile for each wavelength layer.

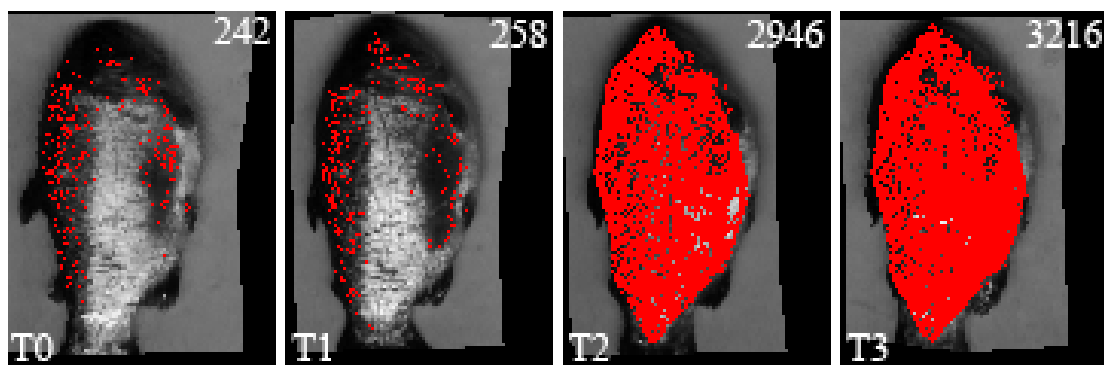
Figure 6: Results of AP1 for the selected ROIL for each wavelength layer (from 500 to 800nm, on the top right of each image) with white pixels reported as the loadings LV1 contribution exceeding the 90th percentile.



It can be noticed that the most important contribution is evident with wavelengths within the range 600–800 nm. From the topographic point of view, pixels that present higher levels of load contribution to the output of the PLSDA are all located within the central area of the fish body, close to the anal fin, at the level of the lateral line (Figure 6). Pixels of this area, and associated with wavelengths within the 600–800 nm range, were chosen to implement a model of classification according to AP3.

According to AP2, each pixel per selected ROI_L in all fish was classified as fresh/non-fresh (Figure 7).

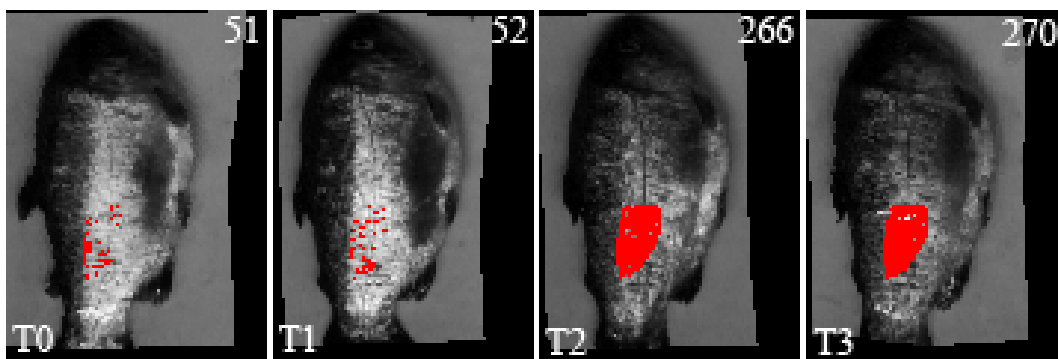
Figure 7. Example of classification based on AP2 where one individual is pictured for one hyperspectral layer (i.e. at 650 nm) at different d.p.m. Pixels in red are those classified as ‘non-fresh’. The number of red pixel is reported on the top right of each figure and the total number of pixels of the ROI is 3641.



With the same ROI_L used for all fish, because of previous superimposing with the geometric morphometry, all hyperspectral images are topographically comparable. Hence, each fish can be classified as fresh/non-fresh based on the number of pixels. A threshold of 50% in the ROI_L pixels was used to discriminate inclusion of the sample into one of the two chosen class-statuses. All of the fresh-category individuals were well classified and 9 out of 10 (90%) non-fresh individuals were correctly classified. Globally, with AP2 a percentage of correct classification of 95% was reached.

According to AP3, each pixel within the selected ROIs was classified as fresh and non-fresh separately (Figure 8). Again, with the ROIs being equal for each fish, all hyperspectral images can be topographically compared. Accordingly, each sample could be classified as fresh or non-fresh based on the area extension of pixels within that class. As stated before, in order to classify each individual into the two class-statuses a threshold of 50% in the ROIs pixels was used. Globally, the AP3 percentage of correct classification was 95%.

Figure 8. Example of classification based on AP3 where one individual is pictured for one hyperspectral layer (i.e. at 650 nm) at different d.p.m. Pixels in red are those classified as ‘non-fresh’. The number of red pixels is reported on the top right of each figure and the total number of pixels of the ROI is 279.



Conclusions

Hyperspectral imaging is a technique of high technological and methodological complexity, but with elevated applicative power. In the market, fish freshness is defined and regulated by EU Directive No. 103/76, which classifies the product on the basis of quality parameters, such as the consistency of the meat, the visual aspect (colour of the eye and the gill, the brightness of the skin), and, finally, odour. It has been demonstrated that the quality of fish from both fishery and aquaculture can be evaluated using the hyperspectral video-image morphometric-based analysis.

In particular, two different methods were used on the acquired images that allow for both subjective and objective analysis. The first technique gave an elevated degree in the assessment of fish freshness. The second technique represented an important methodological evolution of the first technique. Based on combined hyperspectral and geometric morphometric techniques, spectral information from pixels was associated with their topographic location for the first time. This novel approach is based on the a priori discrimination of which wavelength areas are more discriminating in relation to fish freshness, considering the unity of the fish product as spectrally inhomogeneous.

In the second case study the proposed technique represents an important methodological development by combining hyperspectral imaging and geometric morphometric tools. This technique was applied in the hyperspectral field, resulting in an innovation allowing the association of topological spectral information. An automated method for the extraction of the fish outline can be implemented in the next future.

Nomenclature

Symbols

R	relative reflectance of the sample at each wavelength
r_s	absolute signal value (radiance) measured for the sample at each wavelength
r_b	absolute signal value (radiance) measured at each wavelength for the black (background noise)
r_w	absolute signal value (radiance) measured at each wavelength for the standard white (100% of reflectance)

Abbreviations

AP	approach
CT	concrete tanks
d.p.m.	days post mortem
GPA	generalized Procrustes analysis
LV	latent variables
NIR	near-infrared

NIRS near-infrared reflectance spectroscopy
PGP patented prism-grating-prism
PLS partial least squares analysis
PLSDA partial least squares discriminant analysis
ROI region of interest
ROI_L ROI large (i.e. topographic positioning based on the first 10 landmarks)
ROI_S ROI small (i.e. the proportion of ROI_L and the wavelengths resulted in more informative
SC sea cages
VIS/NIR visible/near-infrared

Acknowledgment

This work was funded by the project HighVision (DM 19177/7303/08) from the Italian Ministry of Agricultural, Food and Forestry Politics and by Programma Operativo Regionale – Puglia (Gesticom srl) and Friuli Venezia Giulia (Federcoopescas). Jacopo Aguzzi is a Postdoctoral Fellow of the “JAE” Program (Education and Science Ministry-MEC, Spain).

References

- Aguzzi J, Costa C, Antonucci F, Company JB, Menesatti P, Sardá F (2009) Influence of diel behaviour in the morphology of decapod natantia. *Biological Journal of the Linnean Society*, 96: 517-532.
- Alasalvar C, Garthwaite T, Alexis MN, Grigorakis K, (2001) Freshness assessment of cultured sea bream (*Sparus aurata*) by chemical, physical and sensory methods. *Food Chemistry*, 72: 33-40.
- Alasalvar C, Taylor KDA, Oksuz A, Shahidi F, Alexis M, (2002) Comparison of freshness quality of cultured and wild sea bass (*Dicentrarchus labrax*). *Journal of Food Science*, 67: 3220-3226.
- Antonucci F, Costa C, Aguzzi J, Cataudella S, (2009) Ecomorphology of morpho-functional relationships in the family of sparidae: a quantitative statistic approach. *Press Journal of Morphology* doi: 10.1002/jmor.10725.

- Bellon V, Vigneau JL., Leclercq M, (1993) Feasibility and performance of a new, multiplexed, fast and low-cost fiber-optic NIR spectrometer for the on-line measurement of sugars in fruits. *Applied Spectroscopy*, 47 (7): 1079–1083.
- Bookstein FL (1991) *Morphometric tools for landmark data*. Cambridge: Cambridge University Press.
- Chen H, He Y, (2007) Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science & Technology*, 18: 72-83.
- Costa C, Aguzzi J, Menesatti P, Antonucci F, Rimatori V, Mattoccia M, (2008) Shape analysis of different populations of clams in relation to their geographical structure. *Journal of Zoology*, 276: 71-80.
- Costa C, Angelini C, Scardi M, Menesatti P, Utzeri C, (2009a) Using image analysis on the ventral colour pattern in *Salamandrina perspicillata* (Savi, 1821) (Amphibia, Salamandridae) to discriminate among populations. *Biological Journal of the Linnean Society*, 96 (1): 35-43.
- Costa C, Menesatti P, Aguzzi J, D'Andrea S, Antonucci F, Rimatori V, Pallottino P, Mattoccia M, (2009b) External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*. *Food and Bioprocess Technology*, doi: 10.1007/s11947-008-0068-8.
- Costa C, Menesatti P, Paglia G, Pallottino F, Aguzzi J, Rimatori V, Russo G, Recupero S, Reforgiato Recupero G, (2009c) Quantitative evaluation of Tarocco sweet orange fruit shape using opto-electronic elliptic Fourier based analysis. *Postharvest Biology and Technology*, 54: 38-47.
- Costa C, Tibaldi E, Pasqualetto L, Loy A (2006) Morphometric comparison of the cephalic region of cultured *Acipenser baerii* (Brandt, 1869), *Acipenser naccarii* (Bonaparte, 1836) and their hybrid. *Journal of Ichthyology*, 22 (1): 8–14.
- Cozzolino D, Murray I, Scaife JR, (2002) Near infrared reflectance spectroscopy in the prediction of chemical characteristics of minced raw fish. *Aquaculture Nutrition*, 8: 1–6.
- ElMasry G, Wold JP (2008) High-speed assessment of fat and water content distribution in fish fillets using online imaging spectroscopy. *Journal of Agricultural and Food Chemistry*, 56 (17): 7672-7.

- FAO Fisheries Department, Fishery Information, Data and Statistics Unit (2007) FISHSTAT Plus: Universal software for fishery statistical time series. Version 2.32.
- Farzam M, Beheshti S, Raahemifar K, (2008) Calculation of abundance factors in hyperspectral imaging using genetic algorithm. *Canadian Conference on Electrical and Computer Engineering*, pp. 837-842.
- Gowen AA, O'Donnell CP, Cullen PJ, Downey G, Frias JM, (2007) Hyperspectral imaging: an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, 18: 590-598.
- Huss HH, (1995) Quality and quality changes in fresh fish. FAO Fisheries Technical Paper 348. <http://www.fao.org/docrep/V7180E/V7180E00.HTM#Contents>
- Karoui R, Schoonheydt R, Cecuypere E, Nicolai B, De Baedemaeker J, (2007) Front face fluorescence spectroscopy as a tool for the assessment of egg freshness during storage at a temperature of 12.2 degrees C and 87% relative humidity. *Analytica Chimica Acta*, 582: 83-91.
- Katayama K, Komaki K, Tamiya S, (1996) Prediction of starch, moisture, and sugar in sweetpotato by near infrared transmittance. *HortScience*, 31 (6): 1003–1006.
- Kays SJ, (1999) Preharvest factors affecting appearance. *Postharvest Biology and Technology*, 15: 233–247.
- Knaflewska J, Pospiech E, (2007) Quality assurance systems in food industry and health security of food. *Acta Scientiarum Polonorum, Technologia Alimentaria*, 6 (2): 75-85.
- Lammertyn J, Nicolai B, Ooms K, De Smedt V, De Baerdemaeker J, (1998) Non-destructive measurement of acidity, soluble solids, and firmness of Jonagold apples using NIR-spectroscopy. *Transactions of the ASAE*, 41 (4): 1089–1094.
- Li Q, Wang M, Gu W, (2002) Computer vision based system for apple surface defect detection. *Computers and Electronics in Agriculture*, 36 (2): 215–223.
- Liu Y, Chen YR, Wang CY, Chan DE, Kim MS, (2006) Development of hyperspectral imaging technique for the detection of chilling injury in cucumbers; spectral and image analysis. *Applied Engineering in Agriculture*, 22 (1): 101–111.

- Loy A, Busilacchi S, Costa C, Ferlin L, Cataudella S, (2000) Comparing geometric morphometrics and outlines fitting methods to monitor fish shape variability of *Diplodus puntazzo* (Teleostea, Sparidae). *Aquacultural Engineering*, 21 (4): 271–283.
- Lu R, Park B, (2008) Hyperspectral and multispectral imaging for food quality and safety. *Sensing and Instrumentation for Food Quality and Safety*, 2 (3): 131–132.
- Mathias JA, Williams PC, Sobering DC (1987) The determination of lipid and protein in freshwater fish by using nearinfrared reflectance spectroscopy. *Aquaculture*, 61: 303–311.
- Mehl PM, Chao K, Kim MS, Chen YR (2002) Detection of contamination on selected apple cultivars using hyperspectral and multispectral image analysis. *Applied Engineering in Agriculture* 18 (2): 219–226.
- Menesatti P, D'Andrea S, Costa C, (2007) Spectral and thermal imaging for meat quality evaluation. In: C. Lazzaroni, S. Gigli, D. Gabina (Eds), *New developments in evaluation of carcass and meat quality in cattle and sheep*. Wageningen: Wageningen Academic Publishers, EAAP 123: 115–134.
- Menesatti P, Paglia G, Solaini S, Zanella A, Stainer R, Costa C, Cecchetti M, (2002) Non-linear multiple regression models to estimate the drop damage index of fruit. *Biosystems Engineering*, 83 (3): 319–326.
- Menesatti P, Pallottino F, Lanza G, Paglia G, (2009) Prediction of blood orange MT firmness by multivariate modelling of low alterative penetrometric data set: a preliminary study. *Postharvest Biology and Thechnology*, 51: 434–436.
- Menesatti P, Urbani G (2004) Prediction of the chilling storage time of fresh salt-water fishes by soft modelling (PLS) of low-alterative penetrometric test. Oral presentation International and European Agricultural Engineering Conference – AgEng2004 – Workshop “Physical Methods in Agriculture”, Leuven, Belgium, 12-16 September 2004. Book of abstracts, 940–941.
- Menesatti P, Urbani G, Millozza M, D'Andrea S, Solaini S, Paglia G, Niciarelli I, (2006) Prediction of the chilling storage time of fresh salt-water fishes by means of non-destructive techniques. Oral communication CIGR Section VI

- International Symposium on Future of Food Engineering, Warsaw, Poland, 26-28 April 2006. Abstract 145.
- Menesatti P, Urbani G, Pallottino F, D'Andrea S, Costa C (2007) Non-destructive multi-parametric instruments for fish freshness estimation. *Instrumentation Viewpoint*, 6.
- Menesatti P, Zanella A, D'Andrea S, Costa C, Paglia G, Pallottino F (2009) Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples. *Food and Bioprocess Technology*, 2 (3): 308–314.
- Nicosevici T, Garcia R, Gracias N (2007) Identification of geometrically consistent interest points for the 3D scene reconstruction. *Instrumentation Viewpoint*, 6.
- Nortvedt R, Torrissen OJ, Tuene S (1998) Application of near infrared transmittance spectroscopy in the determination of fat, protein, dry matter in Atlantic halibut fillets. *Chemometrics and Intelligent Laboratory Systems*, 42: 199–207.
- Orban E, Di Lena G, Navigato T, Casini I, Santaroni G, Marxetti A, Caproni R (2002) Quality characteristics of sea bass intensively reared and from a lagoon as affected by growth conditions and the aquatic environment. *Journal of Food Science*, 67: 542-546.
- Park B, Lawrence KC, Windham WR, Smith D, (2006). Performance of hyperspectral imaging system for poultry surface fecal contaminant detection. *Journal of Food Engineering*, 75 (3), 340-348.
- Park B, Windham WR, Lawrence KC, Smith DP (2004) Hyperspectral image classification for fecal and ingesta identification by spectral angle mapper. ASAE/CSAE meeting, Ottawa, Ontario, Canada, ASAE Paper No. 043032.
- Peirs A, Scheerlinck N, Perez AB, Jancsok P, Nicolai BM (2002) Uncertainty analysis and modelling of the starch index during apple fruit maturation. *Postharvest Biology and Technology*, 26 (2), 199-207.
- Polder G, Heijden G, Young I (2002) Spectral image analysis for measuring ripeness of tomatoes. *Transactions of the ASAE*, 45: 1155-1161.
- Rohlf FJ, (1999) Shape statistics: Procrustes superimpositions and tangent spaces. *Journal of Classification*, 16: 197-223.

- Rohlf FJ, Bookstein FL, (1990) Proceedings of the Michigan morphometric workshop. Special Publication No. 2. Ann Arbor, MI: University of Michigan Museum of Zoology.
- Rohlf FJ, Slice D, (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, 39: 40–59.
- Sabatier R, Vivein M, Amenta P, (2003) Two approaches for discriminant partial least square. In: M. Schader, *et al.* (Ed.), *Between Data Science and Applied Data Analysis*. Berlin, Germany: Springer.
- Sigurgisladottir S, Hafsteinsson H, Jonsson A, Lie O, Nortvedt R, Thomassen M, Torrissen O (1999) Textural properties of raw salmon fillets as related to sampling method. *Journal of Food Science* 64: 99-104.
- Sjöström M, Wold S, Söderström B (1986) PLS discrimination plots. In: E.S. Gelsema & L.N. Kanals (Eds.), *Pattern recognition in practice II*. Amsterdam, The Netherlands: Elsevier.
- Solberg C, Fredriksen G, 2001. Analysis of fat and dry matter in capelin by near infrared transmission spectroscopy. *Journal of Near Infrared Spectroscopy*, 9: 221–228.
- Thompson D'AW, (1917) *On Growth and Form*. London, UK: Cambridge University Press.
- Wold JP; Johansen, I. R.; Haugholt, K. H.; Tschudi, J.; Thielemann, J.; Segtnan, V. H.; Narum, B.; Wold, E. (2006) Non-contact transreflectance near infrared imaging for representative on-line sampling of dried salted coalfish (bacalao). *Journal of Near Infrared Spectroscopy*; 14 (1), 59-66.
- Xiccato GA, Trocino, F. Tulli and E. Tibaldi (2004) Prediction of chemical composition and origin identification of European sea bass (*Dicentrarchus labrax* L.) by near infrared reflectance spectroscopy (NIRS). *Food Chemistry*, 86: 275–281.
- Yang CC, Chen YR and Chao K (2005) Development of multispectral image processing algorithms for identification of wholesome, septicemic, and inflammatory process chickens. *Journal of Food Engineering*, 69 (2): 225-234.
- Zelditch ML, Swiderski DL, Sheets HD, Fink WL (2004) *Geometric Morphometrics for Biologists: A Primer*. San Diego, CA: Elsevier Academic Press.

4.4 Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples

FULL REFERENCE OF THE PUBLICATION

Menesatti P, Zanella A, D'Andrea S, Costa C, Paglia G, Pallottino F, 2009, Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples. FOOD BIOPROCESS TECHNOL 2:308–314 - DOI 10.1007/s11947-008-0120-8

Introduction

In post-harvest biology a fruit is considered mature at "that stage at which a commodity has reached a sufficient stage of development that after harvesting and post-harvest handling, its quality will be at least the minimum acceptable to the ultimate consumer" (Reid, 1992). Mitchell (1991) suggested that a minimum maturity index should allow individual shippers to have the opportunity to produce and market a product at a higher maturity compared to the base level that was set up for the industry. The definition of apple maturity corresponding to the stage of fruit development giving minimum acceptable quality to the ultimate consumer implies measurable points in the commodity's development, and the need for techniques to measure maturity. The maturity indices are important to trade regulation, marketing strategy and to the efficient use of labor and resources (Crisosto, 1994). Maturity measurements carried out by producers, handlers, and quality control personnel need to be simple, readily performable in the field or at a certain inspection point. The data should preferably be objective (an instrumental measurement) rather than subjective (an evaluation by human beings) and it is beneficial if the maturity index can be determined by non-destructive measurement (Herold et al., 2005). Several different parameters are used to assess fruit maturity: starch, firmness, juice sugar and acid content, seed color, flesh color, background color, and internal ethylene concentration (Knee, 1993; Crisosto, 1994; Olmo et al, 2000). Some papers compare and combine different parameters for maturity analysis. For example Zude-Sasse et al. (2001)

tested spectral transmittance recordings and elastic behaviour for their potential to describe the maturity of apples. Noh et al. (2007) developed an integrated hyperspectral reflectance and fluorescence imaging technique for measuring apple maturity.

Presently, for apple the most reliable index is the starch degradation pattern of the pulp. Starch conversion into sugars is provide an important indicators for predicting the optimal harvest time in apples production (Smith et al., 1979; Brookfield et al, 1997; Peirs et al., 2002). The starch conversion is usually assessed using an iodine test, in which cut fruits are dipped in iodine solution and stain patterns are rated. Peirs et al. (2002) analyzed the subjective starch scores observing that the variability encountered in the traditional visual starch index scores could be attributed to: a) differences within replicates done by the same inspector, b) inaccurate measurement between experts, c) the subjective interpretation of the reference charts.

Researchers have been developed optical techniques to non destructively determine maturity level and quality of apples within the visible wavelength range of the fruit reflectance and transmittance spectra and backscattering imaging (Zude, 2003; Herold, 2005; Peng & Lu, 2006). Regarding the starch-iodine index, Peirs et al. (2002) developed an imaging technique to reduce the uncertainty due to subjective interpretation of different inspectors. Hyperspectral techniques were also studied with respect to existing machine vision techniques to avoid the use of toxic iodine solution (Peirs et al., 2003).

The hyperspectral imaging is a complex technique that involves automated processing of images providing the spectral data in the visible or near infrared range. Additional methods allow the detection of X-ray, electron energy loss, Raman signal, cathode-luminescence spectroscopy. This technique has the advantage to combine the spectral information on each pixel of the image (Menesatti et al., 2007).

Hyperspectral imaging was used for bruise detection on 'Jonagold' (Juan and Baerdemaeker, 2005) and 'Golden Delicious' apples (Xing et al., 2005), to determine surface defects and contaminations (Mehl et al., 2004) and for non-destructive measurement of bitter pit in apple fruit (Nicolai et al., 2006). Peirs et al. (2003) used an hyperspectral near infrared reflectance imaging system for the determination of the starch index of apple fruit. The hyperspectral images were processed by a

unsupervised multivariate techniques such as principle component analysis (PCA) and the difference between the starch area and non-starch area was classified using a threshold value of the first principle component score plot.

Based on these experiences, the authors applied a different methodology using the supervised multivariate technique of classification by means of partial least-squares discriminant analysis (PLSDA) to elaborate hyperspectral NIR images of cut in halves 'Golden Delicious' apples. Aim of the work was to study the relationships between NIR spectral images and starch/starch-free patterns, indirectly measuring apples maturity stages, in order to develop an accurate, objective and fast method to assess starch index avoiding subjective assessment and usage of toxic iodine solution.

Materials and Methods

The experiment was performed at the Research Centre for Agriculture and Forestry of Laimburg on 88 apples (Golden Delicious Klon B) sampled 7 times within the ripening period at the tree to obtain a wide range of maturity stages. The apples were transported into the laboratory and stored under controlled temperature conditions (2 °C). Twenty-four hours before the analyses apples were thermalized under ambient conditions (21 °C) in order to avoid influences on the NIR measurement. An equatorial central slice of each fruit, the position of which was marked with a reference needle, was used for different analyses:

- a) NIR image acquisition (without any treatment); b) RGB image acquisition after iodine treatment;
- c) visual expert's evaluation after iodine treatment (classical method).

Starch-iodine test

Iodine test represent the actual reference method to assess apple's maturity index. It is based on the principle that the starch, not converted into sugar during the ripening process, getting in contact with an iodine solution turns to blue-black. Fruit coloring is done dipping half apple in the solution prepared dissolving 10 grams of iodine crystals and 25 grams of potassium iodide in 1 liter of water. The pattern of starch

disappearance is specific for each variety (Blanpied and Silsby, 1992). The maturity index was evaluated using a reference color chart.

Experts assigned a starch index score (from 0% to 100%) to each image by comparison with the reference color chart. The chart shows how the maturity depends not only on the percentage of starch-free (white), but also on its pattern on the apple slice surface.

Color image segmentation

After starch-iodine test an image was acquired with a flatbed photometric scanner at 300 DPI 24 bit color. The color images, subtracted by background and internal carpellar cavity, were segmented into 2 levels grayscale: white corresponding to starch-free; mid-grey corresponding to starch, using the k-nearest neighbors (KNN) supervised multivariate clustering method (Dasarathy, 1995).

Each pixel of the RGB input image was classified by a majority vote of its neighbors, with the pixel being assigned the class most common amongst its k-nearest neighbors (k=3). The neighbors were taken from a set of pixels for which the correct classification was known (training set). The training set was built by calculating the mean value for each RGB channel on some extracted original image patches. The classification procedure and the reconstruction of the final images were performed all in Matlab environment. For the KNN training set, 50 areas for starch and 50 for starch-free from 20 apples images randomly chosen were used. From these areas all the images of the dataset (88 apples) were classified.

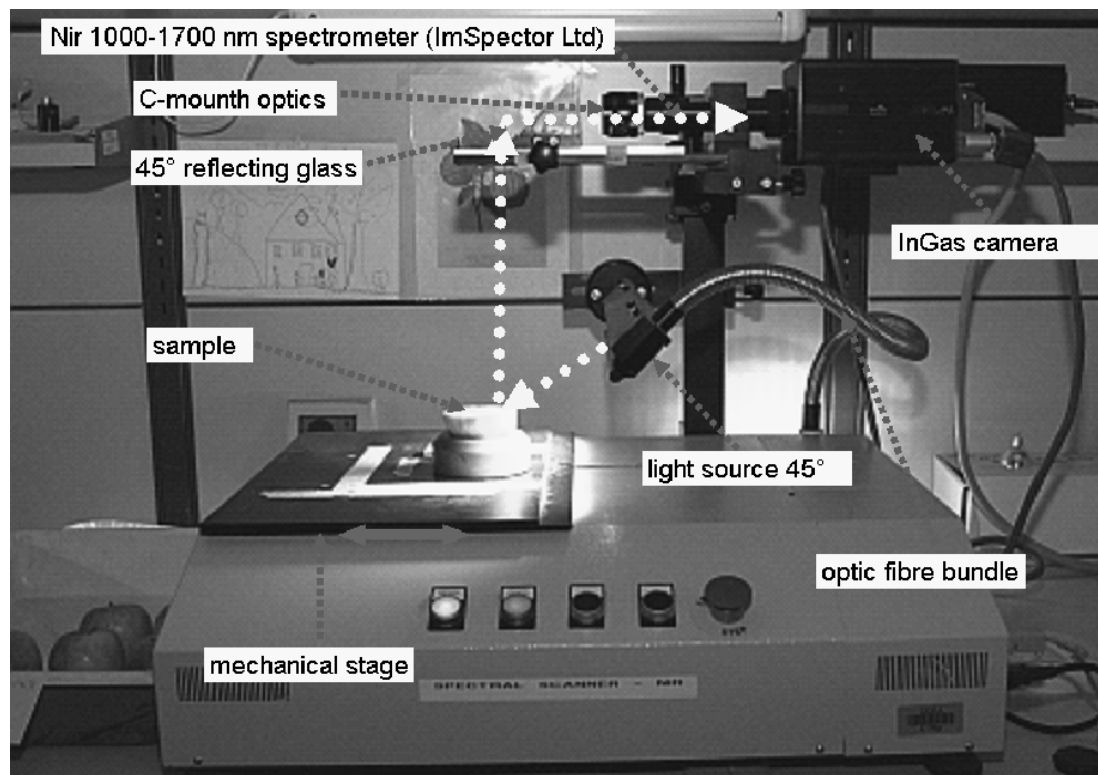
NIR Imaging spectrophotometer and hyperspectral imaging

The NIR imaging system equipped with a spectral scanner (DV, Padova, Italy) and spectrophotometer (ImSpector N17, Specim, Oulu, Finland) is based on the utilization of an integrated hardware and software architecture able to digitally capture and handle spectra as an image sequence. The incoming image line (frame) is projected and dispersed onto InGaAs array, for NIR analysis. Each frame contains the pixels line in one dimension (spatial axis) and the spectral pixels in the other dimension (spectral axis), providing full spectral information for each pixels line. The reconstruction of a whole hyperspectral image of the sample was performed scanning

the sample line-by-line as the entire field of view defined by the optical system is acquired (Bonifazi et al., 2005).

The detection architecture system (Fig. 1) consists of four elements: i) a sample transportation plate ii) a collimated illumination device composed by one 150W halogen lamp, as source light, and one illumination opening in optical fiber of 200 mm long and 2 mm width, positioned at 45° in respect to the transportation plate (i.e. to the sample) and presenting a minimum light divergence, iii) an imaging spectrophotometer coupled with a standard C-mount zoom lens and iv) an array detector, a Pixelvision SU128 InGaAs IR camera.

Figure 1: NIR hyperspectral imaging system.



The imaging spectrometer was used to acquire NIR images ranging from 1000 to 1700 nm. The resolution of the image line is 128 pixels by 12 bits. Three-hundred-fifty lines (frames) compose each image. The spatial resolution for each apple image

(without the boundaries) resulted to be equal to about 100 x 100 pixels depending from the apple size.

All spectral values were expressed in terms of relative reflectance (R), applying the following equation:

$$R = \frac{r_s - r_b}{r_w - r_b} \quad (1)$$

and being R the relative reflectance of the sample at each wavelength, r_s the absolute signal value (radiance) measured for the sample at each wavelength, r_b the absolute signal value (radiance) measured at each wavelength for the black (background noise) and r_w the absolute signal value (radiance) measured at each wavelength for the standard white (100% of reflectance). The NIR image acquisition time lasts about 8 seconds. The hyperspectral classification once the model has been developed and tested lasts 2 seconds.

In the present work, to increase the potential ability of starch index classification, supervised multivariate techniques such as partial least-squares discriminant analysis (PLSDA) was applied, instead of unsupervised one (Peirs et al., 2003).

PLSDA (Sjöström et al., 1986; Sabatier et al., 2003) consists in a classical partial least-squares analysis (PLS) regression where the response variable is a categorical one (Y-block; replaced by the set of dummy variables describing the categories) expressing the class membership of the statistical units (Costa et al., 2008a; 2008b). PLSDA does not allow for response variables other than those that define the groups of individuals (starch or starch-free). The model includes a calibration phase and a cross-validation phase; during both the percentages of correct classification were calculated. The prediction ability in the test phase depends also from the number of the latent variables (LV) used in the model. Therefore, on a sub-sample of 20 apples, we chose the optimal number of LV on the base of the highest percentage of correct classification of the model.

The PLSDA analysis provides the percentage of correct classification of the whole model as well as for the two classes considered (starch-free and starch). This analysis was performed using Matlab (version 7.1, PLS Toolbox Eigenvector version 4.0).

NIR spectra values were pre-processed with the Matlab preprocessing tool 'autoscale'. PLSDA technique was used on hyperspectral NIR images to classify each

single pixel on the base of its NIR reflectance spectrum. The response variable (i.e. the classification for each pixel) was identified through the matching within each single pixel between the color images, segmented in two classes (starch-free and starch), and the NIR hyperspectral matrix.

To obtain a proper superimposition between color (starch-iodine test image: SIT_i) and NIR images (NIR_i), each SIT_i was warped (rotated, scaled and translated) with the same angle of each NIR_i, having the two images the same reference point (a needle). After the warping procedure the two images had the same angle, size and resolution (Costa et al., 2008c). A bounding box of both SIT_i and NIR_i was extracted; the background and internal carpellar spaces were consequently eliminated from the analysis.

Two types of image comparison were performed:

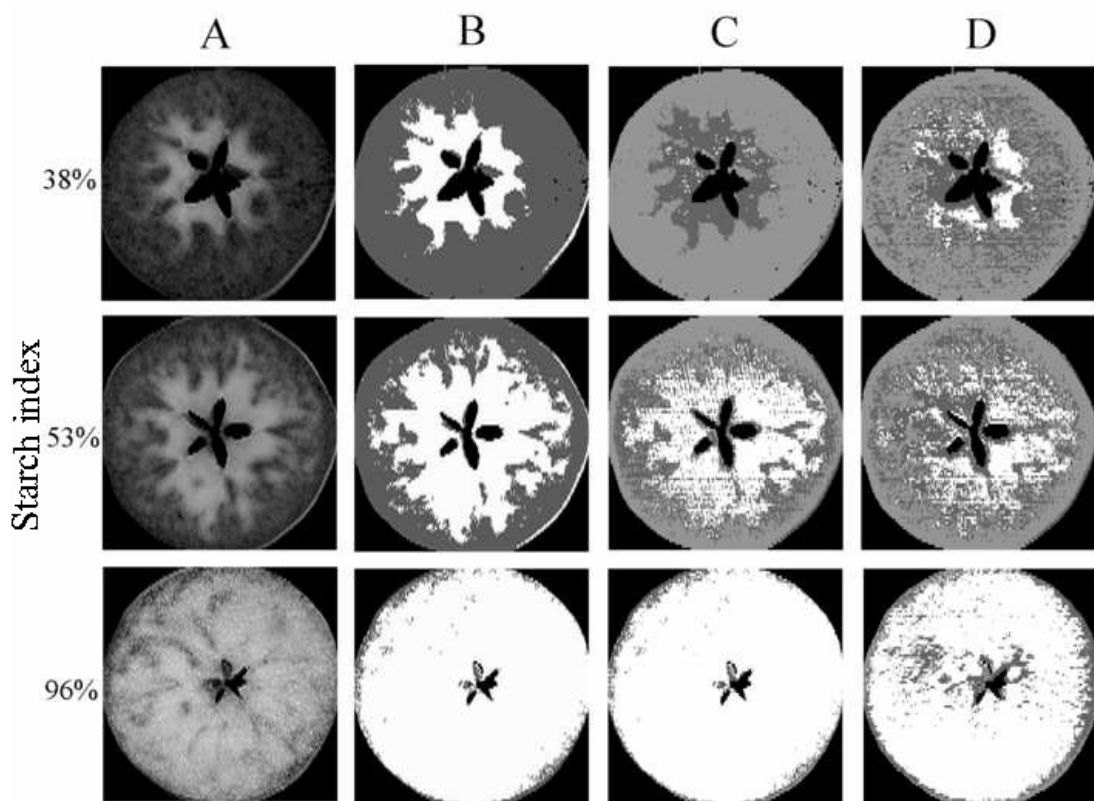
- 1) classification based on PLSDA models for individual apples (i.e. one model for each single fruit);
- 2) PLSDA unique model classification (i.e. all the fruits classified by a single model).

Results and discussion

Classification based on PLSDA models for individual apples

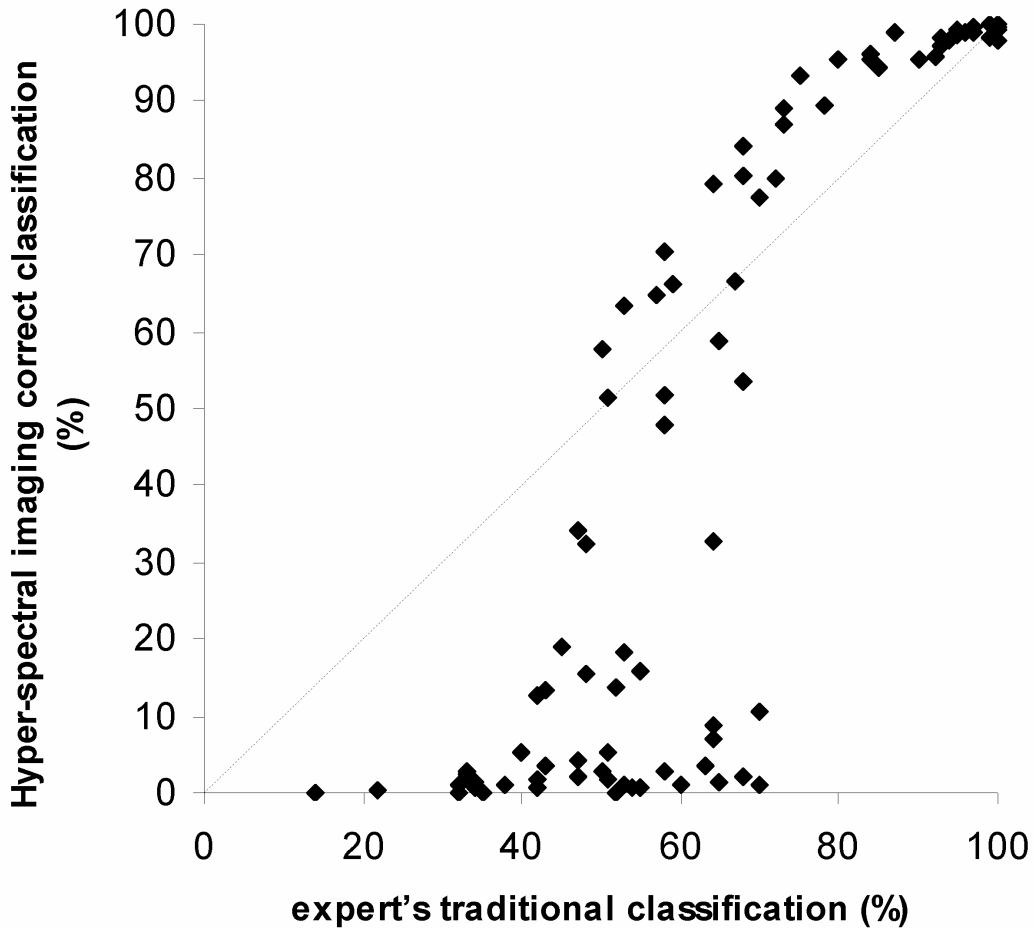
The first analysis was conducted on each apple NIR matrix using as y-variable the classification on the correspondent SIT_i KNN-based. It was possible to observe the correctness of the PLSDA classification (Fig. 2C) with respect to the KNN-segmented images (Fig. 2B) exemplarily for three apples at different starch indices (Fig. 2A). Considering all the fruits, the mean percentage of the pixels correctly classified by NIR_i, with respect to SIT_i and after KNN, was 80.81±10.72%. The starch classification of each apple was based on the percentage of classified starch-free areas on the total (starch-free and starch) classified pixels. This percentage was compared with the expert's evaluation and with the SIT_i starch classification KNN-based.

Figure 2: Three examples of apples at different starch index (assessed by expert's traditional classification): A) original SITi; B) SITi after KNN classification (black= non considered areas, white=starch-free areas, dark gray=starch areas) ; C) NIRi after classification based on PLSDA models for each apple (black= non considered areas, white=starch-free areas, pale gray=starch areas, dark grey=misclassified areas); D) NIRi after classification based on PLSDA unique model (black= non considered areas, white=starch-free areas, pale gray=starch areas, dark grey=misclassified areas).



The regression between the classification obtained by expert's traditional starch index and the SITi classification KNN-based showed a Pearson correlation coefficient of $r = 0.90$ and non linear regression $r = 0.92$ (Fig. 3).

Figure 3: Classification based on PLSDA models for each apple: relationship between the expert's traditional classification and the SITi classification KNN-based.



Good performances was observed of classification for higher maturity percentages. Expert's traditional classification, especially on low values, is based not only on the percentage of starch-free but also on the positioning of starch-free areas on the sliced surface. This fact explain the inferior performances obtained for lower maturity values. Figure 4 displays the relationship between the expert's traditional starch index classification and the one obtained by hyperspectral imaging by means of PLSDA (Pearson correlation coefficient $r = 0.86$; non linear regression $r = 0.90$). A trend similar to the one shown in figure 3 is observable.

Figure 4: Classification based on PLSDA models for each apple: relationship between the expert's traditional classification and the hyper-spectral imaging PLSDA correct classification.

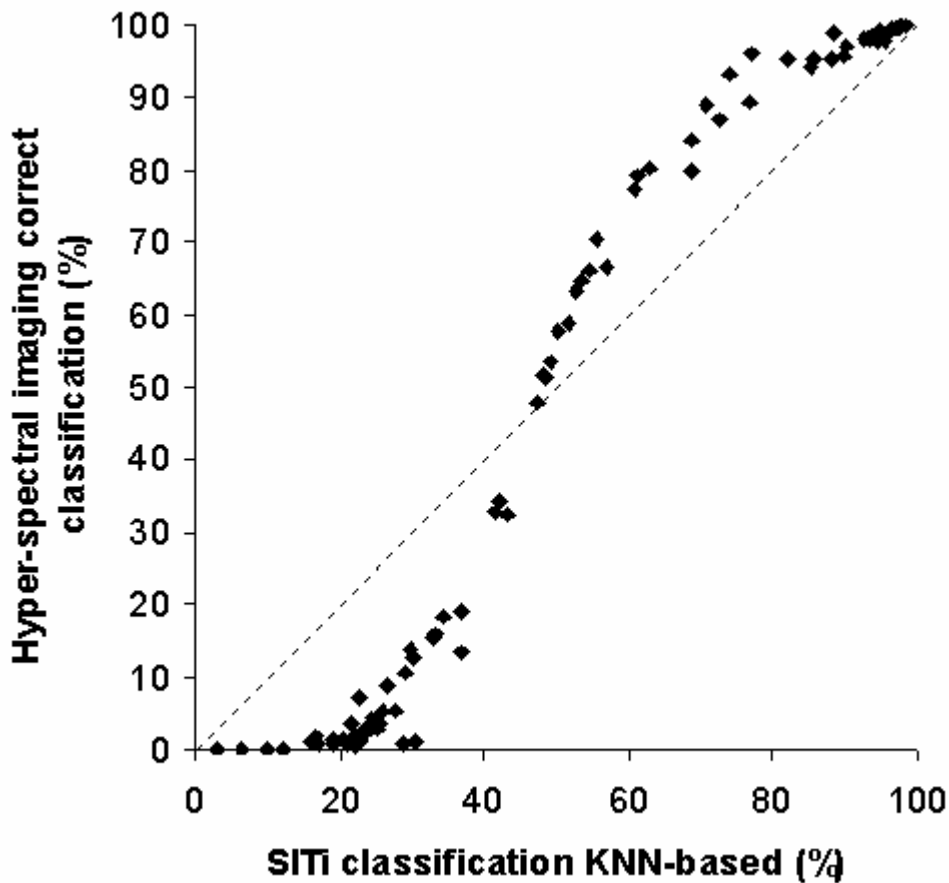
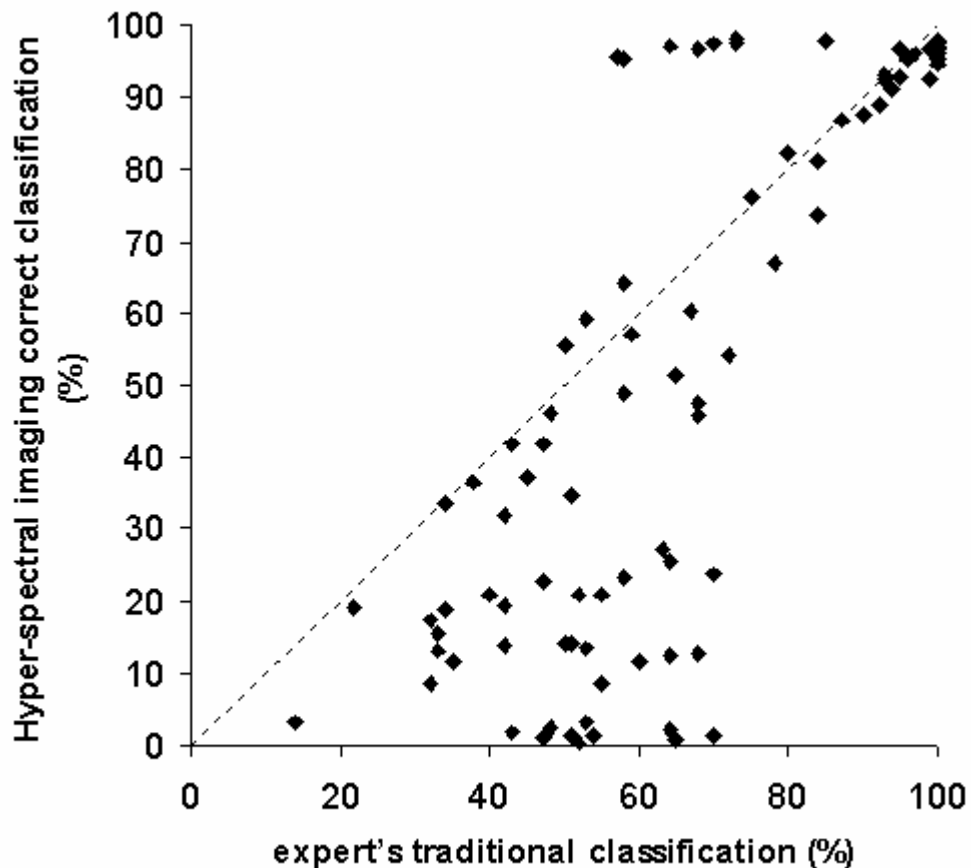


Figure 5 reports the regression between the starch index SITi classification KNN-based and the one found by hyperspectral imaging PLSDA classification (Pearson correlation coefficient $r = 0.97$; non linear regression $r = 0.98$). Here for lower values of maturity, hyperspectral imaging appear to be less efficient (zero) for starch-free detection. Generally the present classification indices became interesting from an applicative point of view after a certain percentage of starch conversion (Peirs et al., 2002).

Figure 5: Classification based on PLSDA models for each apple: relationship between the SITi classification KNN-based and the hyper-spectral imaging PLSDA correct classification.



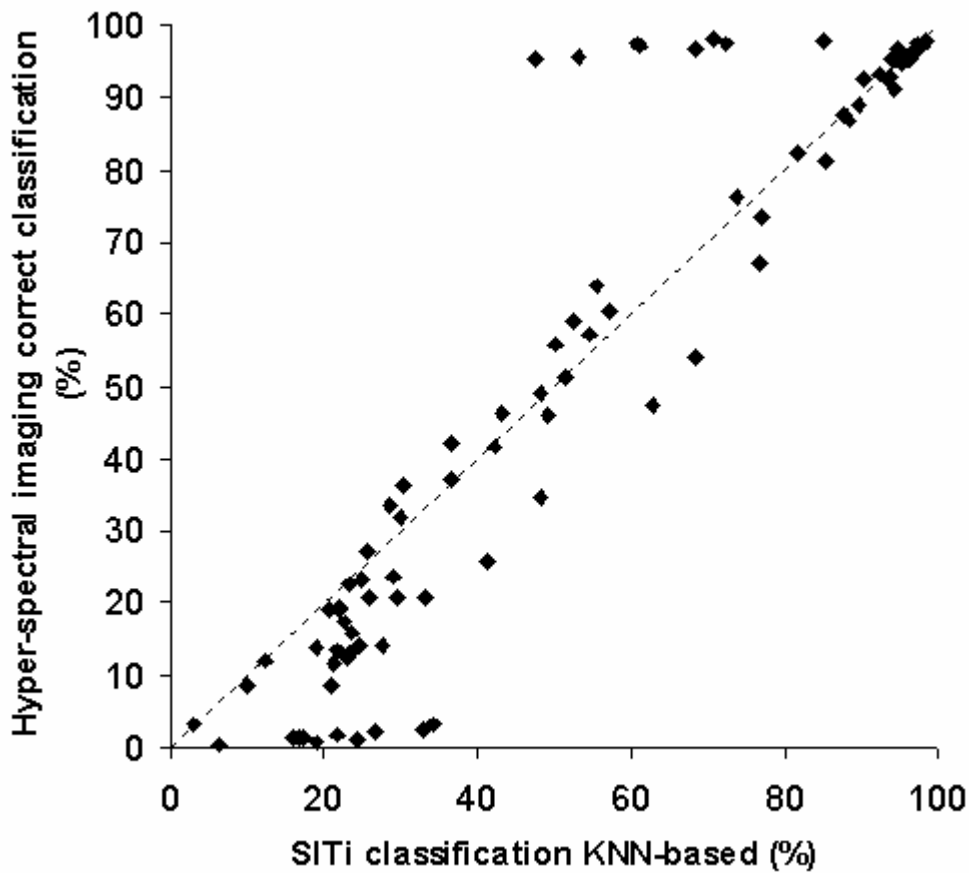
PLSDA unique model classification

A unique model was built on a training dataset (1000 starch and starch-free NIR spectra) manually extracted from the NIR hyper-spectral image warped on the SITi (Table 1). The percentage of correct classification of the model based on 15 latent variables (LV) resulted equal to 72.7%.

Applying the unique model, the starch level of each fruit was classified with respect to the SITi classification in terms of percentage of classified starch-free on the total

(starch-free and starch) classified pixels. The mean percentage of correct classification was $66.33 \pm 9.53\%$.

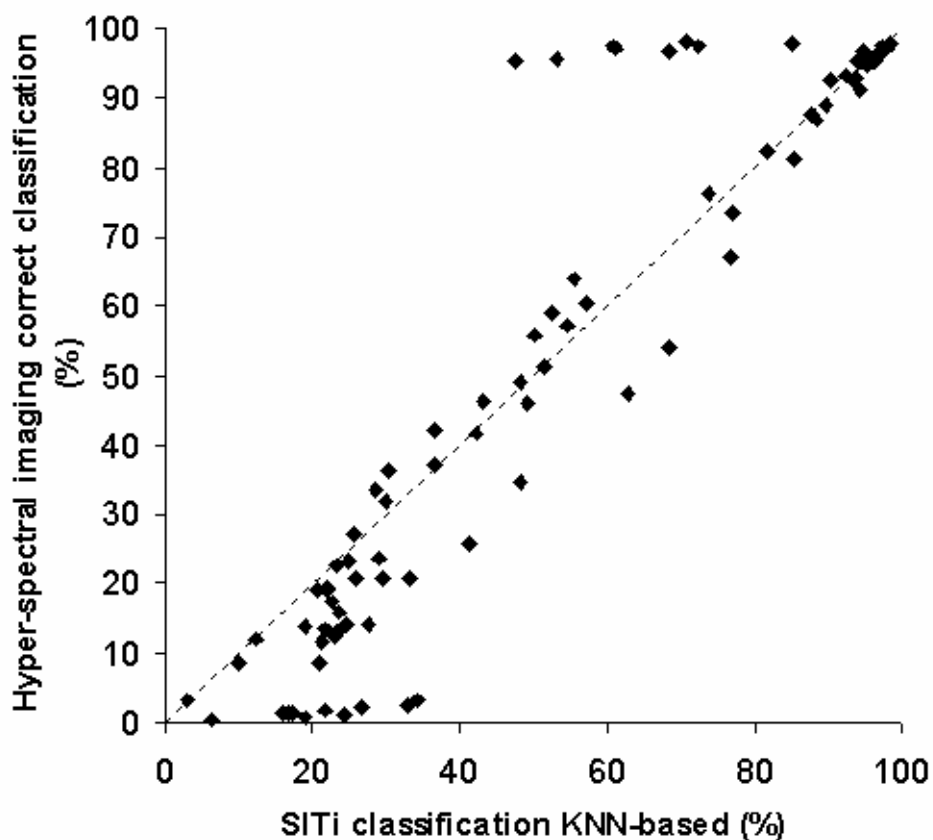
Figure 6: PLSDA unique classification model classification: relationship between the expert's traditional classification and the hyper-spectral imaging PLSDA correct classification.



The scatter plot of the regression between the expert's traditional starch classification index and the one obtained by the unique hyper-spectral imaging PLSDA-based resulted in Pearson correlation coefficient $r = 0.79$ and non linear regression $r = 0.82$ (Fig. 6). The results obtained with the unique PLSDA model underlines a lower correlation performance respect to the one performed for each apple. Also here similar trend appears compared to figure 4 for lower maturity values, while for higher values outliers were present. Similar observations were obtained comparing the

relationship between the SITi classification and unique PLSDA classification resulting in Pearson correlation coefficient $r = 0.95$ and non linear regression $r = 0.94$ (Fig. 7) with figure 5.

Figure 7: PLSDA unique classification model classification: relationship between the SITi classification KNN-based and the hyper-spectral imaging PLSDA correct classification.



In this case hyperspectral imaging (based on PLSDA unique model) miss-classified few fruits with respect SITi classification KNN-based.

The classification based on PLSDA models for individual apple resulted in lower classification errors. This can mainly be attributed to the high heterogeneity of individual apples. The second PLSDA unique model, on the contrary, is based on the small training set and a wide test set. This approach showed good results even if lower than the previous one, based on single fruit classification. Even taking into

account the presence of few outliers, this method has a high potential of applicability. Moreover it must be considered that over 2300000 spectra (hyperspectral pixels) were classified using a training dataset composed by only 1000 starch and starch-free NIR spectra (0.043%), while usually the ratio between training and test set ranges from 50% to 80%. The classification performance of this unique model is both good and highly related to the SITi classification KNN based.

Conclusions

This work shows the feasibility of NIR imaging spectroscopy for apple fruit starch determination.

NIR imaging spectroscopy comprises three main advantages with respect to traditional wet-chemistry analyses and SITi imaging:

- It could avoid the subjective interpretation of the expert's traditional starch index assignments, that are subject to a certain amount of inconsistency and inaccuracy (Peirs et al., 2002)
- the measurement could be faster and this implies that more measurements might be performed during the same time span
- this procedure avoids the use of toxic substances even in comparison with the color image analysis

Future target of this work is to develop a system that could operate in uncontrolled ambient conditions, such as auction hangars, and that might operated by unskilled staff.

Acknowledgements

This study was funded by the Research Centre for Agriculture and Forestry Laimburg.

References

Blanpied GD & Silsby KJ (1992) Predicting harvest date windows for apples. Cornell Cooperative Extension Publication. Information Bulletin, 221, 1-12.

- Bonifazi G, Serranti S & Menesatti P (2005) Hyperspectral imaging based techniques in ornamental stone characterization. In: Yud-Ren Chen et al (eds) Proceedings of the SPIE (The International Society for Optical Engineering) on Optical Sensors and Sensing Systems for Natural Resources and Food Safety and Quality, Vol. 5996, pp 203-214.
- Brookfield P, Murphy P, Harker R & MacRae E (1997) Starch degradation and starch pattern indices; interpretation and relationship to maturity. *Postharvest Biology and Technology*, 1(I), 23-30.
- Costa C, Menesatti P, Aguzzi J, D'Andrea S, Antonucci F, Rimatori V, Pallottino P & Mattoccia, M (2008a) External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*. In Press, *Food and Bioprocess Technology*, DOI: 10.1007/s11947-008-0068-8.
- Costa C, Aguzzi J, Menesatti P, Antonucci F, Rimatori V & Mattoccia M (2008b) Shape analysis of different populations of *Ruditapes decussatus* and *R. philippinarum* in relation to their geographical structure. Accepted by, *Journal of Zoology*.
- Costa C, Angelini C, Scardi M, Menesatti P & Utzeri C (2008c) Using image analysis on the ventral colour pattern in *Salamandrina perspicillata* (Savi, 1821) (Amphibia, Salamandridae) to discriminate among populations. Accepted by, *Biological Journal of the Linnean Society*.
- Crisosto CH (1994) Stone fruit maturity indices: a descriptive review. *Postharvest News and Information*, 5(6), 65-68.
- Dasarathy BV (1995) Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques. IEEE Computer Society Press, Los Alamitos, CA, USA.
- Herold B, Truppel I, Zude M & Geyer M (2005) Spectral measurements on 'Elstar' apples during fruit development on the tree. *Biosystems Engineering*; 91(2), 173-182.
- Juan X & Baerdemaeker J (2005) Bruise detection on 'Jonagold' apples using hyperspectral imaging. *Postharvest Biology and Technology*, 37(2), 152-162.
- Knee M (1993) Pome fruits. In: GB Seymour et al (eds) *Biochemistry of Fruit Ripening*, pp 325-346. Chapman and Hall, New York, USA.

- Mehl P, Chen Y, Kim M, & Chan D (2004) Development of hyperspectral imaging technique for the detection of apple surface defects and contaminations. *Journal of Food Engineering*, 61, 67-81.
- Menesatti P, D'Andrea S & Costa C (2007) Spectral and thermal imaging for meat quality evaluation. In: C Lazzaroni et al (eds) *New developments in evaluation of carcass and meat quality in cattle and sheep*, 123, pp 115-134.. Wageningen Academic Publishers, Wageningen, The Netherlands.
- Mitchell FG, Mayer G, Saenz M, Slaughter D, Johnson RS, Biasi B & Delwiche M (1991) Selecting and handling high quality stone fruit for fresh market. In: *1991 Research Reports for California Peaches and Nectarines*, California Tree Fruit Agreement, pp1-11. Sacramento, California, USA.
- Nicolai B, Lotze E, Peirs A, Scheerlinck N & Theron K (2006) Non-destructive measurement of bitter pit in apple fruit using NIR hyperspectral imaging. *Postharvest Biology and Technology*, 40, 1-6.
- Noh HK, Peng Y & Lu R (2007) Integration of hyperspectral reflectance and fluorescence imaging for assessing apple maturity. *Transactions of the ASABE*, 50(3), 963-971.
- Olmo M, Nadas A & García JM (2000) Nondestructive methods to evaluate maturity level of oranges. *Journal of Food Science*, 65(2), 365–369.
- Peirs A, Scheerlinck N, Baerdemaeker J & Nicolai B (2003) Starch index determination of apple fruit by means of a hyperspectral near infrared reflectance imaging system. *Journal of Near Infrared Spectroscopy*, 11(5), 379-389.
- Peirs A, Scheerlinck N, Perez AB, Jancsok P & Nicolai BM (2002) Uncertainty analysis and modelling of the starch index during apple fruit maturation. *Postharvest Biology and Technology*, 26(2), 199-207.
- Peirs A (2002) Measurement of apple quality and maturity by means of NIR spectroscopy and imaging. PhD Thesis n°535. Department of Agriculture, Katholieke Universiteit, Leuven, Belgium.
- Peng YK & Lu RF (2006) Improving apple fruit firmness predictions by effective correction of multispectral scattering images. *Postharvest Biology and Technology*, 41, 266-274.

- Reid MS (1992) Maturation and maturity indices. In: La Rue JH & Johnson RS (eds) Peaches, Plums and Nectarines: Growing and Handling for Fresh Market, 3331, pp 21-28. University of California Department of Agriculture and Natural Resources Publication, California, USA.
- Sabatier R, Vivein M & Amenta P (2003) Two approaches for Discriminant Partial Least Square. In: Schader M et al (eds) Between Data Science and Applied Data Analysis. Springer-Verlag, Berlin, Germany.
- Sjöström M, Wold S & Söderström B (1986) PLS Discrimination plots. In: Gelsema ES & Kanals LN (eds) Pattern Recognition in Practice II. Elsevier, Amsterdam, The Netherlands.
- Smith RB, Loughheed EC, Franklin EW & McMillan I (1979) The starch iodine test for determining stage of maturation in apples. Canadian Journal of Plant Science, 59, 725-735.
- Xing J, Bravo C, Jancsok P, Ramon H & Baerdemaeker J (2005) Detecting bruises on 'Golden Delicious' apples using hyperspectral imaging with multiple wavebands. Biosystems Engineering, 90, 27-36.
- Zude M (2003) Comparison of indices and multivariate models to non-destructively predict the fruit chlorophyll by means of visible spectrometry in apples. Analytica Chimica Acta, 481, 119-126.
- Zude-Sasse M, Herold B, Geyer M & Huyskens-Keil S (2001) Influence of maturity stage on physical properties in apple. Acta Horticulturae, 553(1), 109-110.

5. CONCLUSIONI GENERALI

La fotonica rappresenta una delle frontiere tecnologie ed economiche del nostro tempo. Tecnologie sempre più performanti e metodologie sempre più potenti spingono verso applicazioni che solo pochi anni or sono sembravano fantascienza. La ricaduta delle innovazioni fotoniche nelle applicazioni in agricoltura (biofotonica) è sicuramente importante, come dimostra la crescita esponenziale di pubblicazioni scientifiche nel settore. Anche le applicazioni operative non mancano, in campo (harvester), come in stalla (robot per la mungitura) o in linea di lavorazione (selezionatrici).

Tuttavia, nuove e ampie possibilità applicative si offrono in relazione alla migliore disponibilità del combinato tecnologia/metodi e delle mutate esigenze del settore primario (qualità, rispetto ambientale, tutela della salute, certificazione).

In ogni caso, il motore dell'innovazione è la ricerca. Essa deve fornire continui sviluppi al settore agroindustriale, ricercando tra le numerosissime possibilità tecnico-metodologiche, le soluzioni potenzialmente più efficienti e più efficaci.

L'attività presentata nella tesi, cerca di fornire un minimo contributo a questo processo di ricerca-innovazione-sviluppo-applicazione, così importante per la crescita economica del settore agroalimentare. Gli aspetti tecnico-metodologici presentati, classificati in estrema sintesi con il titolo di *analisi della forma e iperspettrale per immagini*, sono tra i più innovativi e promettenti nel settore della biofotonica.

L'analisi della forma, misurata e parametrata attraverso complessi algoritmi, è una tecnica molto interessante, per il processo di qualificazione del prodotto agroalimentare. La forma, infatti, contribuisce fortemente alla definizione di quella complessa proprietà della qualità organolettica che è l'aspetto. Oltre la qualità organolettica, la forma di un oggetto o un prodotto definisce la sua funzionalità, ma anche la specificità varietale, la provenienza geografica o la tipicità. Nonostante la notevole complessità metodologica, l'analisi della forma è relativamente semplice e poco costosa da applicare nella pratica, richiedendo sistemi tecnologici oramai di larga diffusione (videocamere o fotocamere).

L'analisi di immagine iperspettrale, invece, è una tecnica molto più complessa sia dal punto di vista metodologico che tecnologico. Necessita di specifici strumenti, il cui

costo, tuttavia, è in costante diminuzione, in virtù del notevole progresso tecnologico del settore. Associando informazioni spettrali e spaziali, si conferma come una delle tecniche fotoniche più informative. E' ampiamente studiata e applicata per la qualificazione multiparametrica di prodotti, processi e fattori di produzione.

Per entrambe le tecniche sono state riportate delle attività sperimentali innovative, pubblicate su riviste di impatto, sia dal punto di vista applicativo che metodologico. Per la prima volta in ambito agroalimentare, è stata applicata l'analisi della forma, attraverso l'analisi ellittica di Fourier (EFA) e una tecnica combinata di analisi di immagini iperspettrali e geometria morfometrica.

Le attività sperimentali in questi due specifici settori sono attualmente in prosecuzione, come pure in altri similari: termografia, motion vision.

Le nuove tecnologie e i nuovi metodi che l'intensa attività di ricerca e sviluppo a livello mondiale, sviluppano in ambito (bio)fotonico, aprono indirizzi verso nuove e entusiasmanti attività sperimentali che dovranno attentamente considerare anche le reali ricadute operative in ambito agroalimentare e ambientale.

6. LISTA DELLE PUBBLICAZIONI NEL TRIENNIO (2008-2010) DI DOTTORATO

6.1 Pubblicazioni su riviste scientifiche

6.1.1 Articoli su riviste peer-reviewed ISI (con IF)

1. Pallottino F, Costa C, Menesatti P, Moresi M, ACCEPTED. Use of image analysis to assess the effective surface area of orange fruit submitted to uniaxial compression tests. ACCEPTED TO JOURNAL OF FOOD ENGINEERING
2. Costa C, D'Andrea S, Russo R, Antonucci F, Pallottino F, Menesatti P, In press. Application of non invasive techniques to differentiate sea bass (*Dicentrarchus labrax*, L. 1758) quality cultured under different conditions. In press by AQUACULTURE INTERNATIONAL. DOI: 10.1007/s10499-010-9393-9
3. Antonucci F, Pallottino F, Paglia G, Palma A, D'Aquino S, Menesatti P. 2010. Non-destructive Estimation of Mandarin Maturity Status Through Portable VIS-NIR Spectrophotometer. ACCEPTED FOOD AND BIOPROCESS TECHNOLOGY (IF2009=2.238) DOI: 10.1007/s11947-010-0414-5
4. Costa C, Antonucci F, Pallottino F, Aguzzi J, Sun DW, Menesatti P, ACCEPTED. Shape analysis of agricultural products by computer vision – a review of recent research advances. ACCEPTED TO FOOD AND BIOPROCESS TECHNOLOGY (IF2009= 2.238)
5. Burchi G, Prisa D, Ballarin A, Menesatti P, 2010, Improvement of flower color by means of leaf treatments in lily, SCIENTIA HORTICULTURAE, 125 (2010) p. 456-460, (IF2010=1.197)
6. Costa C, Vandeputte M, Antonucci F, Boglione C, Menesatti P, Cenadelli S, Parati K, Chavanne H, Chatain B, 2010. Genetic and environmental influences on shape variation in the European sea bass (*Dicentrarchus labrax*). BIOLOGICAL JOURNAL OF THE LINNEAN SOCIETY, 2010, 101, 427–436. (IF2008=2.019)

7. Antonucci F, Costa C, Pallottino F, Paglia G, Rimatori V, De Giorgio D, Menesatti P, ACCEPTED. Quantitative method for shape description of almond cultivars (*Prunus amygdalus* Batsch). ACCEPTED TO FOOD AND BIOPROCESS TECHNOLOGY (IF2009= 2.238)
8. Zude M, Menesatti P, 2009, Editorial Note, FOOD AND BIOPROCESS TECHNOLOGY (2009) 2:239, DOI 10.1007/s11947-008-0159-6
9. Angelini C, Costa C, Raimondi S, Menesatti P, Utzeri C, 2010, Image analysis of the ventral colour pattern discriminates between Spectacled Salamanders, *Salamandrina perspicillata* and *S. terdigitata* (Amphibia, Salamandridae). AMPHIBIA-REPTILIA 31 (2010): 273-282. (IF2008=0.784)
10. Aguzzi J, Company JB, Costa C, Menesatti P, Garcia JA, Bahamon N, Puig P, Sardà F, ACCEPTED. Activity rhythms in deep-sea demersal crustacean decapods: chronobiological challenges and potential technological scenarios. ACCEPTED BY FRONTIERS IN BIOSCIENCE. (IF2008=3.308)
11. Pallottino F, Menesatti P, Costa C, Paglia G, De Salvador F R, Lolletti D. IN PRESS Image analysis techniques for automated hazelnut peeling determination. FOOD AND BIOPROCESS TECHNOLOGY. DOI 10.1007/s11947-009-0211-1. (IF2009= 2.238)
12. Menesatti P., Antonucci F., Pallottino F., Rocuzzo G., Allegra M., Stagno F., Intrigliolo F. (2010). Estimation of plant nutritional status by VIS-NIR spectrophotometric analysis on orange leaves [*Citrus sinensis* (L) Osbeck cv Tarocco]. BIOSYSTEMS ENGINEERING doi:10.1016/j.biosystemseng.2010.01.003; 105(2010), 448-454
13. Aguzzi J, Costa C, Fujiwara Y, Iwase R, Ramirez-Llorda E, Menesatti P, 2009. A novel morphometry-based protocol of automated video-image analysis for species recognition and activity rhythms monitoring in deep-sea fauna. SENSORS, (ISSN 1424-8220), 9: 8438-8455. (IF2008=1.870)
14. Menesatti P, Aguzzi J, Costa C, García JA, Sardà F, 2009. A new morphometric implemented video-image analysis protocol for the study of social modulation in activity rhythms of marine organisms. JOURNAL OF NEUROSCIENCE METHODS, 184: 161-168. (IF2008=2.092)

15. Costa C, Menesatti P, Paglia G, Pallottino F, Aguzzi J, Rimatori V, Russo G, Recupero S, Reforgiato Recupero G, 2009. Discrimination of Tarocco sweet orange [*Citrus sinensis* (L.) Osbeck] varieties using opto-electronic elliptic Fourier based analysis of fruit shape. *POSTHARVEST BIOLOGY AND TECHNOLOGY*, 54: 38-47.
16. Menesatti P., Pallottino F., Lanza G., Paglia G. 2009. Prediction of blood orange MT firmness by multivariate modelling of low alterative penetrometric data set: A preliminary study. *POSTHARVEST BIOLOGY AND TECHNOLOGY*, 51 (2009) 434–436
17. Aguzzi J, Costa C, Antonucci F, Company JB, Menesatti P, Sardá F, 2009. Influence of diel behaviour in the morphology of decapod natantia. *BIOLOGICAL JOURNAL OF THE LINNEAN SOCIETY*, 96: 517-532.
18. Costa, C.; Angelini, C.; Scardi, M.; Menesatti, P.; Utzeri, C., 2009. Using image analysis on the ventral colour pattern in *Salamandrina perspicillata* (Savi, 1821) (Amphibia, Salamandridae) to discriminate among populations. *BIOLOGICAL JOURNAL OF THE LINNEAN SOCIETY*, 96, 35-43 (IF2006=2.066)
19. Menesatti P, Costa C, Paglia G, Pallottino F, D'Andrea S, Rimatori V, Aguzzi J, 2008. Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars. *BIOSYSTEMS ENGINEERING*, 101(4): 417-424.
20. Costa C, Aguzzi J, Menesatti P, Antonucci F, Rimatori V, Mattoccia M, 2008. *Shape analysis of different populations of clams in relation to their geographical structure*. *JOURNAL OF ZOOLOGY*, 276(1): 71-80. (IF2007=1.516)
21. Menesatti P, Zanella A, D'Andrea S, Costa C, Paglia G, Pallottino F, 2009, Supervised multivariate analysis of hyperspectral NIR images to evaluate the starch index of apples. *FOOD AND BIOPROCESS TECHNOLOGY* 2:308–314 - DOI 10.1007/s11947-008-0120-8, (IF2009= 2.238)
22. Costa C, Menesatti P, Aguzzi J, D'Andrea S, Antonucci F, Rimatori V, Pallottino P, Mattoccia M, 2009. External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*.

6.1.2 Articoli su riviste internazionali peer-reviewed

1. Paglia, G., Costa, C., D'Andrea, S., Niciarelli, I., Pallottino, F. and Menesatti, P. 2009. INSTRUMENTAL TECHNIQUES FOR POMOLOGICAL CHARACTERIZATION AND HAZELNUT CULTIVAR CLASSIFICATION. *Acta Hort. (ISHS)* 845:701-706; http://www.actahort.org/books/845/845_110.htm
2. Costa, C., Paglia, G., De Salvador, F.R., Lolletti, D., Rimatori, V. And Menesatti, P. 2009. HAZELNUT CULTIVAR IDENTIFICATION WITH LEAF MORPHOMETRIC ANALYSIS: PRELIMINARY RESULTS. *Acta Hort. (ISHS)* 845:245-248; http://www.actahort.org/books/845/845_34.htm
3. Aguzzi J, Costa C, Menesatti P, Manuel A, Del Río J, García JA, Fujiwara Y, Iwase R, 2009. Potential applications of automated video-image analysis in the pelagic and demersal environment including the deep-sea. *INSTRUMENTATION VIEWPOINT*, 8: 72-73. Paper version ISSN 1697-2562 DL B-51.702-03; Electronic version ISSN 1886-4864 DL B-32814-2006
4. Capoccioni F, Costa C, Menesatti P, Lombarte A, Aguzzi J, Ciccotti E, 2009. Otolith growth allometry measurements in the European eel. *INSTRUMENTATION VIEWPOINT*, 8: 71-72. Paper version ISSN 1697-2562 DL B-51.702-03; Electronic version ISSN 1886-4864 DL B-32814-2006
5. Costa C, Pallottino F, Angelini C, Proietti M, Capoccioni F, Aguzzi J, Antonucci F, Menesatti P, 2009. Colour calibration for quantitative biological analysis: a novel automated multivariate approach. *INSTRUMENTATION VIEWPOINT*, 8: 70-71. Paper version ISSN 1697-2562 DL B-51.702-03; Electronic version ISSN 1886-4864 DL B-32814-2006
6. Aguzzi J, Costa C, Company JB, Antonucci F, Pallottino F, Menesatti P, Canali E, Giorgi S, Angelini C, Ketmaier V, 2009. Application of geometric-morphometric, hyperspectral imaging and molecular markers to the study of

- depth-driven differences in populations of Decapods (Crustacea).
INSTRUMENTATION VIEWPOINT, 8: 73-74. Paper version ISSN 1697-
2562 DL B-51.702-03; Electronic version ISSN 1886-4864 DL B-32814-2006
7. Aguzzi J, Costa C, Menesatti P, Garcia JA, Chiesa JJ, Sardá F, 2009.
Monochromatic blue light entrains diel activity cycles in the Norway lobster,
Nephrops norvegicus (L.) as measured by automated video-image analysis.
SCIENTIA MARINA, 73(4): 773-783.
 8. Pallottino F, Costa C, Menesatti P and Moresi M, (2009), Compression testing of
orange fruits, CHEMICAL ENGINEERING TRANSACTIONS, 17, 885-890
 9. Costa C, Menesatti P, Raimondi S, Angelini C, Utzeri C, 2008. Using image
analysis on the ventral colour pattern to discriminate between *Salamandrina
perspicillata* and *Salamandrina terdigitata*. In: Corti C. (ed.).
HERPETOLOGIA SARDINIAE. Societas Herpetologica Italica/Edizioni
Belvedere, Latina, "le scienze" 8: 163-166..
 10. Menesatti P, M. Biocca, S. D'Andrea, M. Pincu. 2008. Thermography to analyze
distribution of agricultural sprayers. QIRT JOURNAL (QUANTITATIVE
INFRARED THERMOGRAPHY JOURNAL) vol. 5, n° 1/2008, pp. 81-96

6.1.3 Capitoli di libri a diffusione internazionale

1. Menesatti P, Costa C, Aguzzi J, 2010. Quality evaluation of fish by hyperspectral
imaging. In: Hyperspectral imaging for food quality: analysis and control. In
D.-W. Sun (Ed.), Hyperspectral imaging for food quality analysis and control.
London, Burlington, San Diego: ACADEMIC PRESS, ELSEVIER: 273-294.
ISBN: 9780123747532

6.1.4 Articoli su riviste nazionali

1. Menesatti P, Antonucci F, Costa C, Pallottino F, Paglia G, Niciarelli I, Papetti P,
IN PRESS. Analisi dello stato di freschezza di trota e cefalo attraverso sistemi

- non-distruttivi e sperimentazione di sistemi innovativi di conservazione refrigerata. IN PRESS BY EUROFISHMARKET
2. Canali E, Costa C, Giorgi S, Antonucci F, Menesatti P. Florovivaismo: sensibile aumento della SAU e tenuta delle esportazioni. COLTURE PROTETTE, 7-8, 2010, pp. 54-61
 3. Corrado Costa: Menesatti P, Antonucci F, Costa C, Pallottino F, Paglia G, Niciarelli I, Papetti P, IN PRESS. Analisi dello stato di freschezza di trota e cefalo attraverso sistemi non-distruttivi e sperimentazione di sistemi innovativi di conservazione refrigerata. IN PRESS BY EUROFISHMARKET.
 4. Aguzzi J, Sardà F, García JA, Costa C, Menesatti P, Sarrià D, Del Rio J, Manuel A, 2009. When to buy Lobster. EUROFISHMARKET, 2: 62-66.
 5. Pallottino F, Menesatti P, Strano MC, Calandra M, Lanza G. 2009, Metodiche strumentali per la valutazione della consistenza dei frutti di agrume. RIVISTA DI FRUTTICOLTURA, 10, pp.76-81
 6. Pallottino, F., Antonucci, F., Canali, M., Giorgi, S., Menesatti, P. 2009. PRS per prodotti deperibili. COLTURE PROTETTE, ORTICOLTURA E FLOROVIVAISMO. Edagricole. anno XXXVIII/ottobre 2009, pp. 50-57
 7. Antonucci, F., Pallottino, F., Canali, M., Giorgi, S., Menesatti, P. 2009. RFID, fresh cut più efficiente. COLTURE PROTETTE, ORTICOLTURA E FLOROVIVAISMO. Edagricole. anno XXXVIII/ottobre 2009, pp.58-61.
 8. Pallottino F, Menesatti P, Paglia G, Strano MC, Calandra M, Lanza G, 2009 Effetti sulla qualità di arance Tarocco del sistema di trasporto a refrigerazione passiva (PRSTM). RIVISTA DI FRUTTICOLTURA, vol 71, N. 1-2, pp. 28-33
 9. D'Andrea S, M. Cegna, F. Pallottino, I. Niciarelli, P. Menesatti. 2007. La qualità e la conservazione della carne di coniglio: rilevamenti per mezzo delle tecniche opto-spetttrali. LA RIVISTA DI SCIENZA DELL'ALIMENTAZIONE. Fo.S.A.N., anno 36, pp.1-7.
 10. Pansini L., Cegna M., Fedrizzi M., Menesatti P., Pari L. 2008. Legatura meccanica della vite, scegliere il legaccio giusto. L'INFORMATORE AGRARIO, 4/2008, p. 93-96.

6.2 Brevetti

1. Menesatti P, Costa C, Aguzzi J, Antonucci F. "Apparato di riconoscimento di prodotti agroalimentari", Verbale di Deposito Domanda di Brevetto per Invenzione Industriale, presso la Camera di Commercio Industria, Artigianato e Agricoltura di ROMA, Numero domanda: RM2009A000325, Data di deposito: 25/06/2009, CCIAA di deposito: ROMA, Richiedente: Consiglio per la Ricerca e la Sperimentazione in Agricoltura (C.R.A.)

6.3 Atti di convegni

1. O'Rourke S, Holden N, Menesatti P, Giorgi S, Canali E, Predictions of soil organic carbon in topsoil from forest soils using VIS-NIR hyperspectral imaging, Conference IASIM-10 (International Association for Spectral Imaging), Dublin 17th Nov 2010
2. Pallottino F, Moresi M, Lanza M C, Antonucci F, Menesatti P. SUBMITTED. Classificazione reometrica di arance rosse: analisi strumentale e sensoriale a confronto, III convegno nazionale SISS, Portici (NA), 1-2 Dicembre 2010
3. Antonucci F, Menesatti P, Canali E, Giorgi S, Maienza A, Stazi SR. 2010. Tecniche di spettrofotometria d'immagine nella valutazione del quantitativo di rame in suoli contaminati ad "hoc". Poster al XXVIII Convegno Nazionale della Società Italiana di Chimica Agraria (SICA) presso l'Università Cattolica del Sacro Cuore, Istituto di Chimica Agraria ed Ambientale, Piacenza 20-21 Settembre 2010
4. Aguzzi J, Costa C, Menesatti P, Company JB, 2010. New challenges to deep-water stock and biodiversity assessments: the chronobiological scenario and automated video-imaging solutions. Presentazione al Automated Image Processing Workshop organized by National Oceanic and Atmospheric Administration (NOAA), September 7-9, 2010, Seattle, USA.
5. Matabos M, Aguzzi J, Robert K, Costa C, Menesatti P, Company JB, Juniper K, 2010. A preliminary multi-parametric study on behavioural modulation in

- demersal decapods of Saanich Inlet by VENUS platform. Presentation at the 12th Deep-sea Biology Symposium, June 7-11, 2010, Reykjavík, Iceland
6. Aguzzi J, Furushima Y, Chiesa JJ, Costa C, Menesatti P, Fujiwara Y, Company JB, 2010. Behavioural rhythms in deep-water continental margins and the deep-sea. Presentazione al the 12th Deep-sea Biology Symposium, June 7-11, 2010, Reykjavík, Iceland.
 7. Antonucci F, Menesatti P, Canali E, Giorgi S, Maienza A, Stazi S R, 2010, Hyperspectral Imaging Characterization Of Agricultural Topsoil Copper Concentration. 2nd International Workshop on Image Analysis in Agriculture. 26-27 August 2010, Budapest, Hungary.
 8. Costa C, Antonucci F, Pallottino F, Boglione C, Cataudella S, Menesatti P, 2010. Color-warping imaging: a non destructive technique to evaluate gilthead seabream (*Sparus aurata*, Linnaeus 1758) freshness. 2nd International Workshop on Image Analysis in Agriculture. 26-27 August 2010, Budapest, Hungary.
 9. Matabos M, Aguzzi J, Robert K, Costa C, Menesatti P, Company JB, Tunnicliffe V, Juniper K, SUBMITTED. A preliminary multi-parametric study on behavioural modulation in demersal decapods of Saanich Inlet by VENUS platform. 31st Pacific Evolution Ecology Conference (PEEC), March 5-7, 2010, Bamfield Marine Sciences Center, University of Victoria (Canada)
 10. Pallottino F, Moresi M, Stefano G, Menesatti P, Rheometrical characterization of orange fruit under stress-relaxation tests. XVIIth World Congress of the International Commission of Agricultural Engineering (CIGR), Québec City, Canada June 13-17, 2010: 256
 11. Ulissi V, Menesatti P, Benincasa P., Guiducci M, 2010, Estimation of tomato nutritional status by VIS-Nir portable spectrophotometric system. XVIIth World Congress of the International Commission of Agricultural Engineering (CIGR), Québec City, Canada June 13-17, 2010: 108
 12. Costa C, Antonucci F, Pallottino F, Canali E, Boglione C, Cataudella S, Menesatti P, 2010. An innovative colorimetric calibration method used to quantify differences among wild and reared gilthead seabream (*Sparus aurata*, Linnaeus 1758). XVIIth World Congress of the International Commission of

Agricultural Engineering (CIGR). June 13-17, 2010, Québec City, Canada:
251

13. Biocca M, Gallo P, Menesatti P. 2009 Aerodynamic and Morphological Properties of Organo-mineral Fertiliser Particles. Proceedings of "XXXIII CIOSTA CIGR V Conference 2009: Technology and management to ensure sustainable agriculture, agrosystems, forestry and safety" 17-19 June, Reggio Calabria. Giametta G., Zimbalatti G. Eds., Vol. 2, 1399-1403.
14. Lai A, Pallottino F, Sighicelli M, Menesatti P, 2009, OPTOELECTRONICAL SYSTEMS FOR FRESH-CUT SALADS EVALUATIONS: A PRELIMINARY STUDY, 5CIGR Potsdam
15. Menesatti P, D'andrea S, Cavalieri A, Baldi M, Lanini M, Vitali A, Lacetera N, Bernabucci U, Nardone A, 2009, Microclimatic indexes to compare different dairy cows housing systems in different climatic conditions, EAAP (European association of animal production) – 60th Annual Meeting, Barcelona, Spain, 24-27 August 2009, book of abstract n. 15, p. 117
16. Aguzzi J, Vitali A, D'andrea S, Cavalieri A, Menesatti P, 2009, Variation of milk yield in relation to temperature humidity index in different dairy cow housing systems: a preliminary time-lagged approach, EAAP (European association of animal production) – 60th Annual Meeting, Barcelona, Spain, 24-27 August 2009, book of abstract n. 15, p. 118
17. Menesatti P, Antonucci F, Costa C, Santori A, Niciarelli I, Infantino A, 2009, Application of morphometric image analysis system to evaluate the incidence of fusarium head blight wheat infected kernels. 1st International Workshop on Computer Image Analysis in Agriculture, Potsdam, Germany 27 – 28 August 2009, Bornimer Agrartechnische Berichte - Heft 69, ISSN 0947-7314, Leibniz-Institut für Agrartechnik Potsdam-Bornim e.V. (ATB)
18. Menesatti P, Rocuzzo G, Cegna M, Torrisi B, Niciarelli I., Allegra M, Intrigliolo F. 2008. Estimation of citrus leaves N content by VIS-Nir portable spectrophotometer. 11th Int Citrus Congress Wuhan 26-30/10/08 (China), pp. 174-175
19. Costa C, Menesatti P, Raimondi S, Angelini C, Utzeri C, 2008. Using image analysis on the ventral colour pattern to discriminate between Salamandrina

- perspicillata and Salamandrina terdigitata. Presentazione al VII Congresso Nazionale della Societas Herpetologica Italica, 1–5 Ottobre 2008, Oristano (Italy)
20. Paglia G, Antonucci F, Costa C, De Giorgio D, Menesatti P, 2008. Studio sulle differenze di forma in frutti di cultivar di mandorlo (*Prunus amygdalus* Batsch) attraverso l'analisi ellittica di Fourier. Poster al VIII Convegno Nazionale "La biodiversità – una risorsa per sistemi multifunzionali", 21–23 Aprile 2008, Lecce (Italy).
21. Menesatti P, F. Pallottino, S. D'Andrea. 2008. Sistemi strumentali innovativi per l'analisi e la gestione di aspetti della qualità organolettica di ortaggi di IV gamma. Oral communication, Convegno "Cibi di ieri e di domani: qualità e sicurezza tra tradizione e innovazione", QUALICIBI, Positano 28-30 Maggio 2008.
22. Paglia G., Costa C., D'Andrea S., Niciarelli I., Pallottino F., Menesatti P. 2008. Instrumental techniques for pomological characterization and classification of hazelnut cultivars. 7th Int. Congress on Hazelnut, Viterbo 23 Giugno 2008
23. Costa C., Paglia G., De Salvador F.R., Lolletti D., Rimatori V., Menesatti P.. 2008. Hazelnuts cultivar identification with leaves morphometric analysis: preliminary results. 7th Int. Congress on Hazelnut, Viterbo 23 Giugno 2008
24. Pallottino F., Cegna M, Paglia G, Menesatti P. 2008. Evaluation of fresh-cut salads organoleptic aspects through optoelectronic and physical systems. International Conference on Agricultural Engineering (AGENG08), 23-25 June 2008 Hersonissos, Crete – Greece, p. 75
25. Costa, C; Menesatti, P.; Lolletti, D.; Paglia, G.; Pallottino, F.; De Salvador, F. R., 2008. Image analysis techniques for the automated evaluation of toasted hazelnuts peeling. International Conference on Agricultural Engineering (AGENG08), 23-25 June 2008 Hersonissos, Crete – Greece. ORAL presentation, pp.54-55
26. Pergher G., Petris R., Biocca M., Menesatti P. 2007. Prime verifiche di campo di un'irroratrice schermata ad aeroconvezione di nuova concezione. Giornate Fitopatologiche 2008.

27. D. Prisa, A. Ballarin, F. Pierandrei, S. D'Andrea, P. Menesatti, G. Burchi. 2008.
Improvement of flower color by means of leaf treatments in Asiatic Lily. 9th
International ISHS Symposium on Postharvest Quality of Ornamental Plants.
Odense (DK), 11-14 August 2008

