Food Bioprocess Technol (2011) 4:673–692
DOI 10.1007/s11947-011-0556-0

REVIEW PAPER

Shape Analysis of Agricultural Products: A Review of Recent Research Advances and Potential Application to Computer Vision

Corrado Costa · Francesca Antonucci · Federico Pallottino · Jacopo Aguzzi · Da-Wen Sun · Paolo Menesatti

Received: 1 June 2010 / Accepted: 9 March 2011 / Published online: 30 March 2011

Abstract The appearance of agricultural products deeply conditions 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: (1) genetic; (2) conformity and condition ratios; (3) products characterization; (4) product sorting and finally, (5) 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 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.

Keywords Image analysis · Shape analysis · Agricultural products · Multivariate statistics

Source of support Finalized research projects "FRUMED2" and "HIGHVISION"; this study is supported by the Italian Ministry of Agriculture and Forestry.

C. Costa (✉) · F. Antonucci · F. Pallottino · P. Menesatti
Agricultural Engineering Research Unit of the Agriculture Research Council, CRA-ING, Via della Pascolare, 16, 00015, Monterotondo Scalo,
Rome, Italy
e-mail: corrado.costa@entecra.it

J. Aguzzi
Instituto de Ciencías del Mar (ICM-CSIC),
Paseo Marítimo de la Barceloneta, 37-49,
08003 Barcelona, Spain

D.-W. Sun
Food Refrigeration and Computerized Food Technology,
Agriculture and Food Science Centre,
University College Dublin, Belfield,
Dublin 4, Ireland

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 organolectic 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. 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; Currie et al. 2000; White et al. 2000; Nunome et al. 2001; van der Knaap and Tanksley 2003; Zygier et al. 2005; Brewer et al. 2007; Menesatti et al. 2008; Costa et al. 2009a; Antonucci et al. 2011), 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 (Sadrnia et al. 2007; Pallottino et al. 2010), 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 this context, the aim of this review is to characterize the present state-of-the-art of shape analysis research with morphometric tools of agricultural products and potential future applications in relation to the computer vision. Accordingly, 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.

Shape and Genetic

Agricultural product shapes are determined both on genetic and environmental basis (Sadrnia 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 and 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 1959; Zygier et al. 2005), 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 and 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 1959). 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.

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 (Riyadi et al. 2008). 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.

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 (Milanesi et al. 2011). 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, Protected Geographical Indication and finally, Traditional Speciality Guaranteed.

Presently, the certification system is missing of shape-based criteria. Cultivar certification can be based on a general and qualitative shape description (Paulus and 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 (Venora et al. 2007, 2009; Menesatti et al. 2008; Costa et al. 2009a; Antonucci et al. 2011).

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 and Heymann 1998).

Machine vision systems are replacing the process of manual inspection of products in different industries (Ruiz-Altisent et al. 2010). Fruit shape is one of the most important factors for classifying and grading fresh horticultural products. Inspection operations may include defect detection, dimensional measurement and product spatial orientation, as well as grading, sorting and finally counting, but in many packinghouses, fruit shape is still manually determined (Xiaobo et al. 2008). Machine vision has several advantages over the conventional methods of inspection (Blasco et al. 2009). It can be tuned and then adjusted in order to work with other on-line processing tasks acting over 24 h. 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; Jarimopas and Jaisin 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. 2011). 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 agro-ecotypes, which is one of the important agronomic traits in buckwheat breeding (Matano and Ujihara 1973; Uehera and 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; Uehera and Namai 1994). Aspect ratio and the area of two-dimensional projection of kernels are a few examples of quantitative measures of kernel shape (Matano and 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.

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 centred 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 subchapters show the main shape descriptors associated with their principal analytical methods. An additional subchapter 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.

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, as well as the 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 (Sadrnia et al. 2007), where the United States Department of Agriculture (USDA 1997) grade standard provides as reference two shapes in three classifications (US Fancy, US No. 1, 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

Table 1 Shape analysis and representation of agricultural products

	Species/crop	Shape analysis	Representation	Ref.
Fruits	Apple	EFD in relation with genetics, PCA	Mean±Max/min for each PC axis	Currie et al. 2000
	Apple	Fourier expansion and PCA	Mean for each cluster and individual samples on the PCA	Paulus and Schrevens 1999
	Apple	Fourier expansion and decision tree	None	Xiaobo et al. 2008
	Averrhoa carambola	Thresholding FD	None	Abdullah et al. 2006
	Citrus	Spherical harmonic descriptors	Mean±SD for each PC axis	Ding et al. 2000
	Citrus unshiu	FFT	Radius signature of the contour	Blasco et al. 2009
	Tarocco sweet orange	EFA, clustering and PLSDA	Mean±SD for each cv and cluster	Costa et al. 2009a
	Papaya	Wavelets and LDA	None	Riyadi et al. 2008
	Pear	Ratios in relation with genetics	None	White et al. 2000
	Strawberries	Descriptive	Reference drawings	Khanizadeh 1994
	Strawberries	Lines relative dimensions	None	Liming and Yanchao 2010
	Sweet cherry	PCA and clustering on 3D shape	Mean for each cluster	Beyer et al. 2002
	Watermelon	Descriptive and ratio, regression	Reference drawings	Sadrnia et al. 2007
	Watermelon	Descriptive and shape indices in relation with genetics	None	Tanaka et al. 1955
Vegetables	Cucumber	Inflection points of the skeletonized fruit area	None	Van Eck et al. 1998
	Eggplant	Descriptive in relation with genetics	Reference drawings	Nunome et al. 2001
	Capsicum	Descriptive and shape indices in relation with genetics	None	Peterson 1959
	Bell pepper	Ratios in relation with volume	None	Ngouajio et al. 2003
	Pepper	Shape index in relation with genetics	None	Zygier et al. 2005
	Squash	Descriptive	None	Nerson 2005
	Tomato	Descriptive and ratios, shape index, PCA	None	Brewer et al. 2006
	Tomato	Descriptive and ratios, shape index in relation with genetics	Individual samples	Brewer et al. 2007
	Tomato	Ratio	None	Jahns et 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 genetics	None	van der Knaap and Tanksley 2003
	Tomato	Descriptive in relation with genetics	Individual samples	Xiao et al. 2008
	Watermelon	Ellipsoid approximation	None	Koc 2007
	White-flowered gourd	EFA and PCA	Mean±2SD for each PC axis	Morimoto et al. 2005
Nuts	Almond	EFA, clustering and PLSDA	Mean±SD for each cv and cluster	Antonucci et al. 2011
	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 and Jindal 2001
	Rice	Ratios in relation with genetics	None	Zheng et al. 2007
	Various	Wavelet analysis with linear discriminant analysis (LDA) and quadratic discriminant Analysis (QDA)	None	Choudhary et al. 2008
Other	Common buckwheat	EFA and PCA	Mean for each cv and mean±2σ for each PC axis	Ohsawa et al. 1998
	Lentils	Shape ratio and LDA	None	Venora et al. 2007
	Olea europaea	Geometric morphometry (landmarks) and CVA	None	Terral et al. 2004
	Olea europaea	EFA and clustering	Mean±SD for each cv	Milanesi et al. 2011
	Phaseolus vulgaris	Shape ratio and LDA	None	Venora et al. 2009
	Sweet tamarind	Automated detection of center and extraction of the pods curvature	Center and radii	Jarimopas and Jaisin 2008

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 or cultivar registers (Khanizadeh 1994; Beyer et al. 2002), evaluation of consumer preference (Kays 1999), investigating hereditability of fruit shape traits (Cannon and 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. 2011).

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. (2011), 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. (2011), 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.

Ratios and Shape Indices

The 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 (Koc 2007; Liming and Yanchao 2010), particularly at the extremities of the fruit (White and Bailey 1995). Some descriptors incorporate ratios calculated from measurements of length and width (Tufts and Hansen 1931; Thibault et al. 1983; White and 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 (Venora et al. 2007, 2009). Length/diameter ratios were studied in order to extract fruit volume (Sabliov et al. 2002; Wang and Nguang 2007) and attribute shape classes as described, for example, in bell peppers by Ngouajio et al. (2003).

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; 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 cannot be easily identified, and categories were not ranked in the past, which made shape selection of apples very difficult. 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 and 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. 2007). 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 and Zhu 1997). In the study of Zheng et al. (2007), 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.

Van Eck et al. (1998) used an image analysis procedure to skeletonize the cucumbers shape, by using the inflection points to calculate the shape of the fruit neck. Inflection points can be found as local extremes of the first-order derivative of the width function. Another interesting application is the one presented by Jarimopas and Jaisin (2008). These authors used the automated extraction of the centre and the circle of 55 pixel radius to determine the shape of the sweet tamarind pod, in order to identify three shape categories (i.e. curved, slightly curved and straight) from its curvature. This method was implemented into an experimental sorting machine vision system.

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 and 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 and 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 on-line 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 (Choudhary et al. 2008; Riyadi et al. 2008). Wavelets are mathematical 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). Abdullah et al. (2006) developed an automated inspection system using machine vision technology for detecting the quality features of golden delicious starfruits of *Averrhoa carambola* through a direct thresholding method on FD.

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 PCA, they characterized cultivars into five 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 PLSDA and clustering on hazelnuts (Menesatti et al. 2008), Tarocco sweet orange fruits (Costa et al. 2009a), almonds (Antonucci et al. 2011) and olives (Milanesi et al. 2011).

Blasco et al. (2009) proposed and tested an automated machine for inspecting and sorting processed mandarin segments. They used a combined matrix composed by morphological parameters (shape factor, compactness, elongation, length, area and symmetry) and the first 10 harmonics of the FFT of the radius signature. A standard non-linear Bayesian discriminant analysis was then applied in order to separate the segments into four commercial categories, having a global efficiency of 93.2%.

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 two landmarks and 18 semi-landmarks.

Multivariate Analyses

Multivariate analyses are generally divided into two main categories: unsupervised and supervised. 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 not only 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. 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. 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), PCA and canonical discriminant analysis (Brewer et al. 2007). Many other studies uses instead shape-based methods in association with PCA (Ohsawa et al. 1998; Paulus and Schrevens 1999; Currie et al. 2000; Cannon and 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, 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; Bylesjo et al. 2006; Tominaga 2006; Menesatti et al. 2008; Costa et al. 2009a; Antonucci et al. 2011), soft independent modelling of class analogy (SIMCA; Bylesjo et al. 2006; Tominaga 2006; Casale et al. 2007; Aguzzi et al. 2009a, b), clustering of the Fourier coefficients (Xiaobo et al. 2008; Costa et al. 2009a, b) and standard non-linear Bayesian discriminant analysis (Blasco et al. 2009).

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.

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 and 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 ± standard deviation (SD; Ding et al. 2000; Costa et al. 2009a; Antonucci et al. 2011; Milanesi et al. 2011) or ±2SD (Ohsawa et al. 1998; Goto et al. 2005; Morimoto et al. 2005). Some examples of shape representation are given in Fig. 1.

Future Perspectives: An Automated Shape Processing System

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. 2011) recently proposed a new 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 Appendix 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).

The Procedure

Figure 2 summaries the steps in the ASPS procedure. The first step (Fig. 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.

Then, each image can be segmented (Fig. 2b) into black (i.e. background) and white (i.e. object) by applying the KNN procedure as formalized in Pallottino et al. (2010). With this supervised multivariate clustering method, image binarization can be carried out since each pixel of the RBG 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 (Fig. 2c) by using the

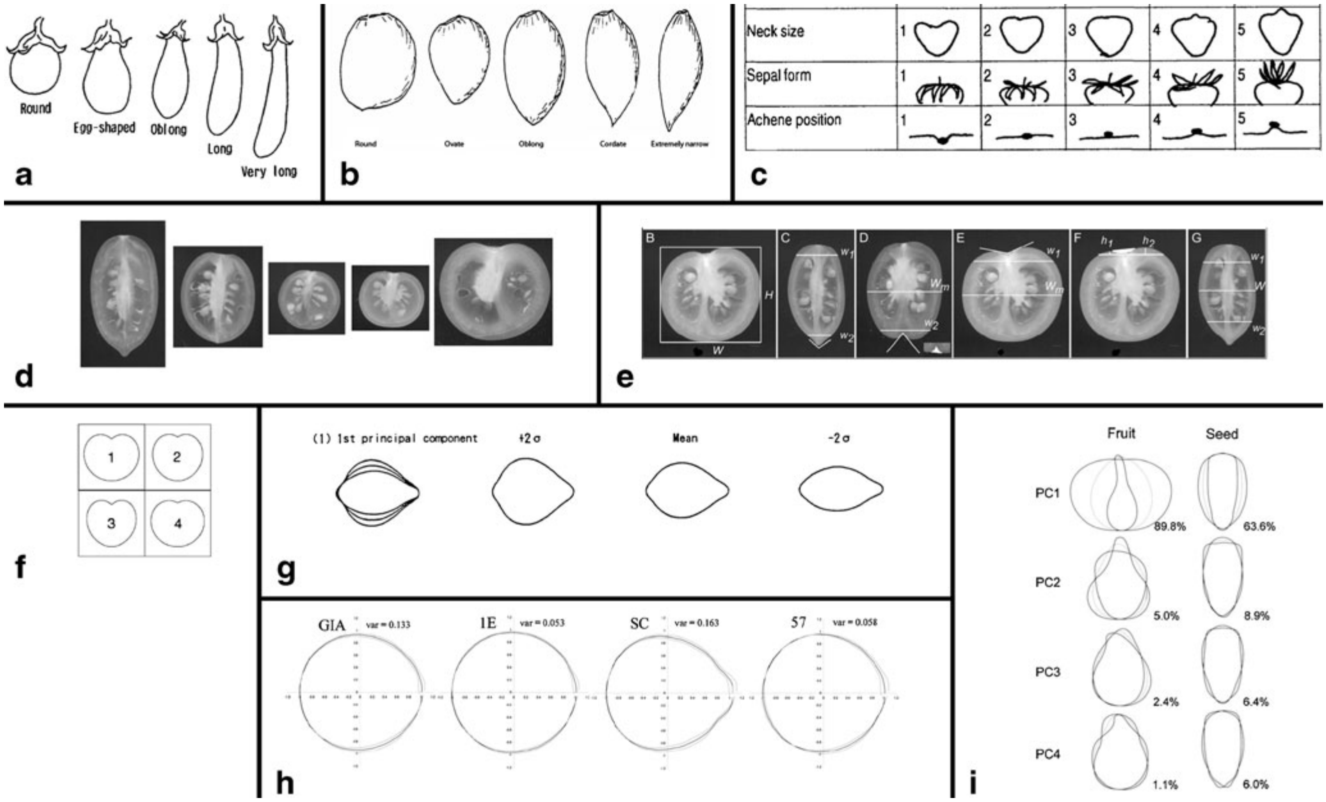

Fig. 1 Examples of shape representation. **a** Eggplant (Nunome et al. 2001); **b** almonds (Gülcan 1985); **c** strawberries (Khanizadeh 1994); **d** individual samplings of tomatoes (Brewer et al. 2007); **e** aspect ratios of tomatoes (Brewer et al. 2007); **f** EFA representation of sweet cherry for each cluster (Beyer et al. 2002); **g** mean±2σ (standard error) for the first PCA axis for apples (Ohsawa et al. 1998); **h** mean±standard deviation (SD) and variance for each cultivar of Tarocco sweet orange (Costa et al. 2009a, b); **i** mean±2SD for each PCA axes of white-flowered gourd (Morimoto et al. 2005)

Matlab procedure (Menesatti et al. 2008; Costa et al. 2009a; Antonucci et al. 2011; Milanesi et al. 2011), as reported in Appendix 1. Coordinates are aligned by Generalised Procrustes Analysis, 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. 2011).

The overall shape can be at this point studied by EFA on the outline coordinates (Fig. 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. 2011). 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 (Fig. 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 multivariant analysis. These analyses are the PCA (Fig. 2f; Currie et al. 2000; Loy et al. 2000; Cannon and Manos 2001; Beyer et al. 2002; Goto et al. 2005) or alternatively, different modelling technique (Fig. 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 (Fig. 2h) as a graphical measure of the extent of morphological variation (Costa et al. 2009a; Antonucci et al. 2011).

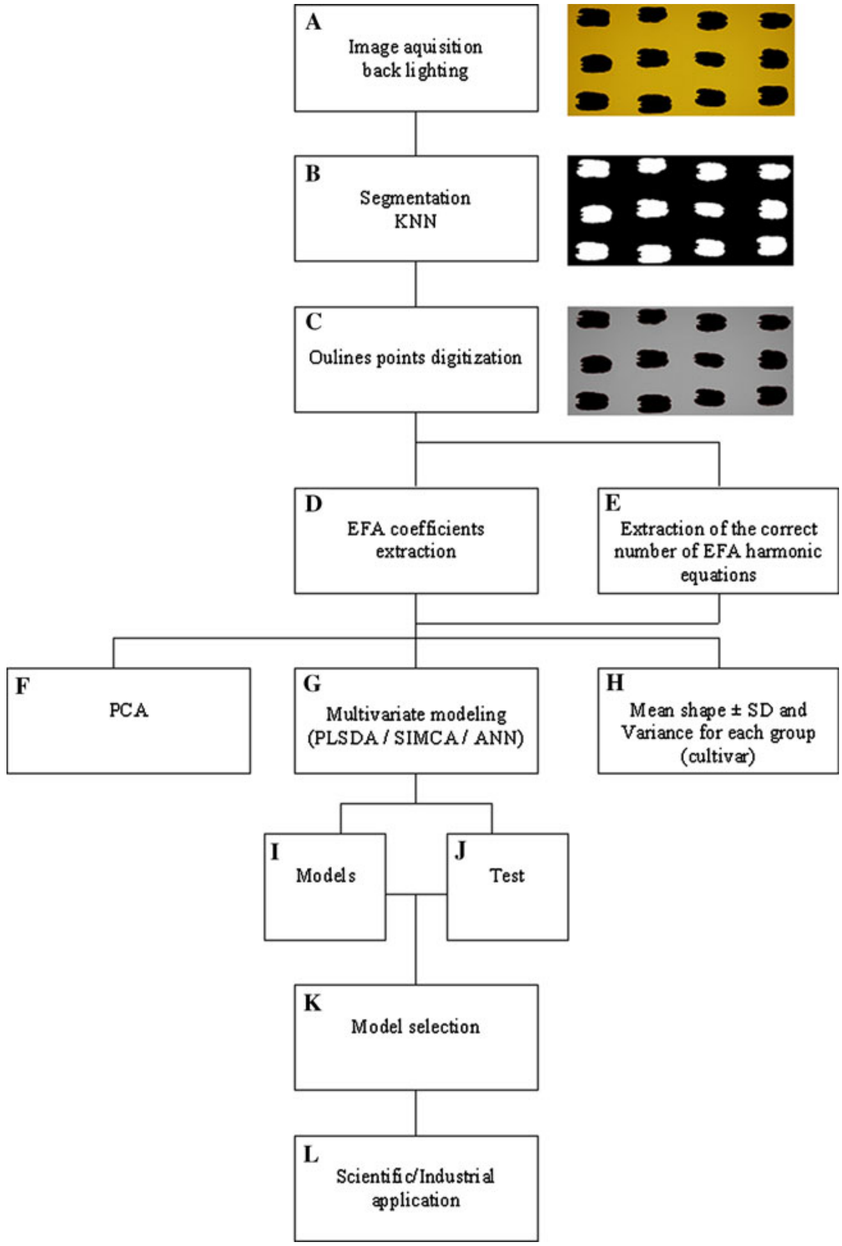

Fig. 2 Diagram of the automated procedure for fruit shape extraction. On the *left side* of **a**–**c** examples of consecutive steps of image processing of the pecans Kiowa cultivar kernels are shown

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. 2011) 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 three commercially important cultivars; 72 pecans belonging to two different pecans cultivars and finally, 120 pistachios from two different cultivars.

In these samples, walnuts were provided by the "Germoplasm Orchard" of the Fruitculture Research Unit

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

Species	Cultivars	No. in-shell fruit	No. of 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	-

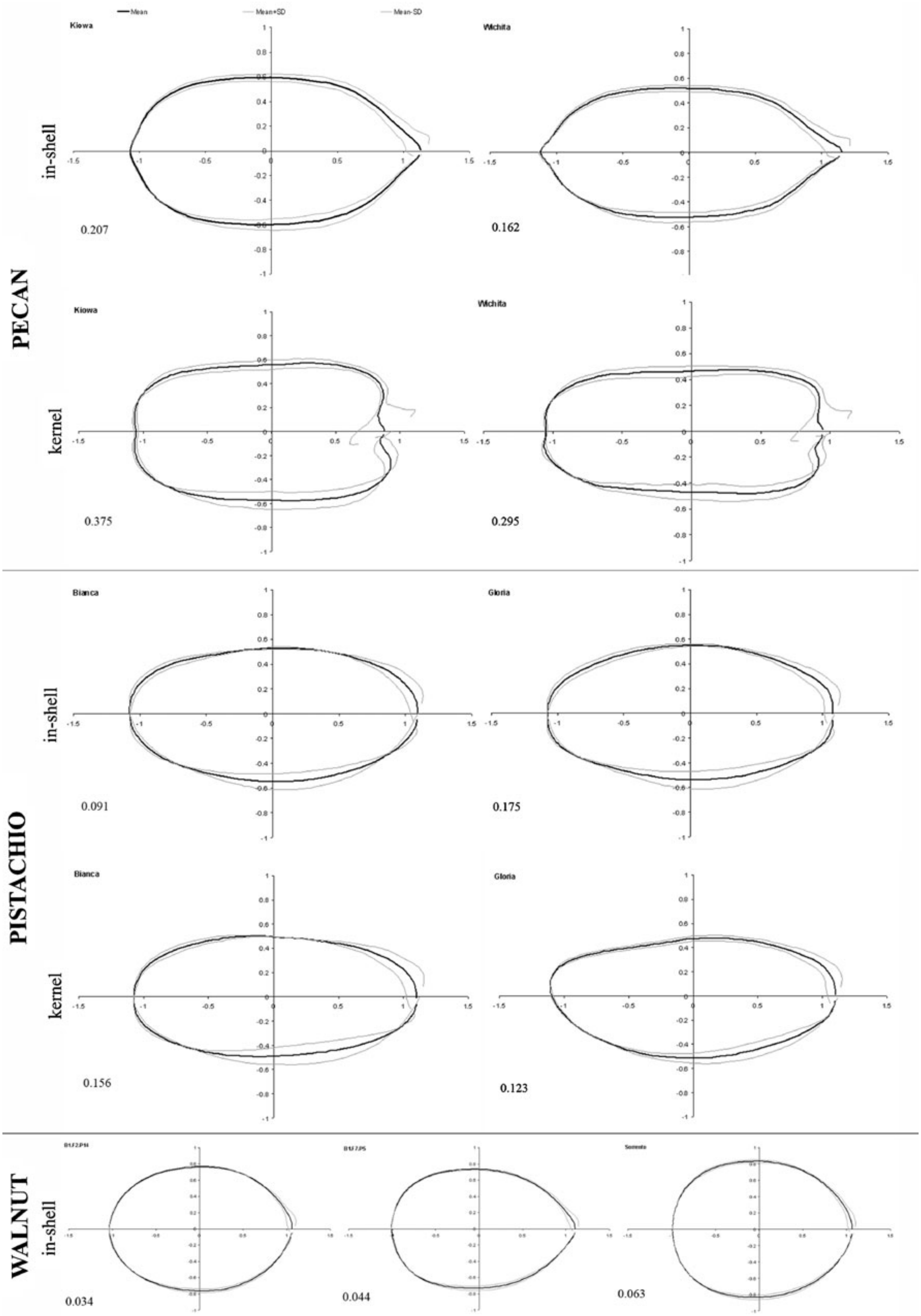

Fig. 3 Mean outline (*black*) and the standard deviation range (*grey*) for in-shell and kernels of each species with the correct number of harmonics as obtained by EFA. The values of variance for each variety are shown on the *bottom-left corner*

of the Agriculture Research Council of Caserta; pistachios and pecan were provided by the "Germoplasm Orchard" of the Fruitculture Centre of the Agriculture Research Council 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 Fig. 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 (24 mm 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: (1) 75% of specimens for the class modelling and validation and (2) 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 and the standard deviation range are shown in Fig. 3 along with the correct number of harmonic equations as well as the values of variance.

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).

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 on-line 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 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).

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

	Pecan in-shell	Pecan kernels	Pistachio in-shell	Pistachio kernels	Walnut in-shell
N	72	72	120	120	180
No. units (Y-block)	2	2	2	2	3
No. 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

N number of samples, *No. units* (*Y-Block*) number of units to be discriminated by the PLSDA, *No. LV* number of latent vectors for each model, *Random probability* (%) probability of random assignment of an individual into a unit

Appendix 1

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

```
% the input images should be at two levels, type: logical, uint8, double
% Image Processing Toolbox is requested

%%INPUT VARIABILES
DirIn='C:\BW\';
DirOut='C:\OUTLINE\';
ImageExtention='bmp';
nPoints=180;            %insert the requested number of equally angulary 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


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,'°', '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. (2009a, b) and Antonucci et al. (2011) (see Fig. 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 buitl 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');
```

References

Abdullah, M. Z., Mohamad-Saleh, J., Fathinul-Syahir, A. S., & Mohd-Azemi, B. M. N. (2006). Discrimination and classification of fresh-cut starfruits (*Averrhoa carambola* L.) using automated machine vision system. *Journal of Food Engineering, 76*, 506–523.

Aguzzi, J., Costa, C., Antonucci, F., Company, J. B., Menesatti, P., & Sardá, F. (2009). 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. (2009). 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.

Alpert, K. B., Grandillo, S., & Tanksley, S. D. (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., Pallottino, F., Paglia, G., Rimatori, V., De Giorgio, D., Menesatti, P. (2011). Quantitative method for shape description of almond cultivars (*Prunus amygdalus* Batsch). *Food and Bioprocess Technology*. doi:10.1007/s11947-010-0389-2.

Batchelor, B. C., Hill, D. A., & Hodgson, D. C. (1985). *Automated visual inspection*. London: IFS.

Beyer, M., Hahn, R., Peschel, S., Harz, M., & Knoche, B. (2002). Analyzing fruit shape in sweet cherry (*Prunus avium* L.). *Scientia Horticolturae, 96*, 139–150.

Bezier, P. E. (1970). *Emploi des machines a` comande nume´rique*. Paris: Ed. Masson.

Blasco, J., Aleixos, N., Cubero, S., Gómez-Sanchís, J., & Moltó, E. (2009). Automatic sorting of satsuma (*Citrus unshiu*) segments using computer vision and morphological features. *Computers and Electronics in Agriculture, 66*, 1–8.

Bookstein, F. L. (1991). *Morphometric tools for landmark data: Geometry and biology*. New York: Cambridge Univ. Press. 435 pp.

Bookstein, F. L. (1996). Landmark method for forms without landmarks: morphometrics of group difference in outline shape. *Medical Image Analysis, 1*(3), 225–243.

Brewer, M. T., 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, 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.

Brown, A. G. (1960). The inheritance of shape, size and season in progenies of the cultivated apple. *Euphytica, 9*, 327–337.

Bruskiewich, R., Coe, E. H., Jaiswal, P., McCouch, S., Polacco, M., Stein, L., et al. (2002). The Plant Ontology™ consortium and plant ontologies. *Comparative and Functional Genomics, 3*(2), 137–142.

Bylesjo, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., & Trygg, J. (2006). OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics, 20*, 341–351.

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.

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.

Choudhary, R., Paliwal, J., & Jayas, D. S. (2008). Classification of cereal grains using wavelet, morphological, colour, and textural features of non-touching kernel images. *Biosystems Engineering, 99*, 330–337.

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., et al. (2009). 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. (2009). 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., et al. (2010). External shape differences between sympatric populations of commercial clams *Tapes decussatus* and *T. philippinarum*. *Food and Bioprocess Technology, 3*(1), 43–48.

Crampton, J. S. (1995). Elliptic Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia, 28*, 179–186.

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.

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, M. C., Lester, R. N., & Tanksley, S. D. (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, U. P. (1938). *Cyclopediea of hardy fruits* (2nd ed.). New York: MacMillan.

Hirose, T., Yoshida, M., Nemoto, H., Kitabayashi, H., Minami, M., Matano, T., et al. (1995). Diversity of grain character of tartary buckwheat in Nepal. In T. Matano & A. Ujihara (Eds.), *Current advances in buckwheat research* (pp. 385–388). Nagano: Shinshu University Press.

Hodgson, R. W. (1967). Horticultural varieties of citrus. In W. Reuther, H. J. Webber, & L. D. Bachelor (Eds.), *The Citrus Industry* (Vol. 1, pp. 431–459). Berkeley: University of California.

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.)* (p. 48). Rome: International Plant Genetic Resources Institute.

IPGRI. (1996). *Descriptors for tomato (Lycopersicon spp.)*. Rome: IPGRI.

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, H. M., & Wolfgang, P. (2001). Measuring image analysis attributes and modelling fuzzy consumer aspects for tomato quality. *Computers and Electronics in Agriculture, 31*, 17–29.

Jarimopas, B., & Jaisin, N. (2008). An experimental machine vision system for sorting sweet tamarind. *Journal of Food Engineering, 89*, 291–297.

Jayas, D. S., Paliwal, J., & Visen, N. S. (2000). *Journal of Agricultural Engineering Research, 77*(2), 119–128.

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. (1991). *Postharvest physiology of perishable plant products*. New York: Van Nostrand Reinholt.

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.

Khanizadeh, S. (1994). Breeding strawberries for Eastern Central Canada. *Euphytica, 77*, 45–49.

Koc, A. B. (2007). Determination of watermelon volume using ellipsoid approximation and image processing. *Postharvest Biology and Technology, 45*, 366–371.

Ku, H. M., Doganlar, S., Chen, K. Y., & Tanksley, S. D. (1999). The genetic basis of pear-shaped tomato fruit. *Theoretical and Applied Genetics, 9*, 844–850.

Lawless, H., & Heymann, H. (1998). *Sensory evaluation of food—principles and practices*. New York: Chapman & Hall.

Lestrel, P. E. (1997). *Fourier descriptors and their applications in biology.* Cambridge: Cambridge University Press.

Li, J., Xiao, J., Grandillo, S., Jiang, L., Wan, Y., Deng, Q., et al. (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.

Liming, X., & Yanchao, Z. (2010). Automated strawberry grading system based on image processing. *Computers and Electronics in Agriculture, 71*, 32–39.

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.

Masters, T. (1994). *Signal and image processing with neural networks: a C++ sourcebook*. New York: Wiley.

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., Costa, C., Paglia, G., Pallottino, F., D'Andrea, S., Rimatori, V., et al. (2008). Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars. *Biosystem Engineering, 101*(4), 417–424.

Milanesi, C., Sorbi, A., Paolucci, E., Antonucci, F., Menesatti, P., Costa, C., et al. (2011). Pomology observations, morphometric analysis, ultrastructural study and allelic profiles of "olivastra Seggianese" endocarps from ancient olive trees (*Olea europaea* L.). *Comptes Rendus Biologies, 334*(1), 39–49.

Morgan, J., Richards, A., & Dowle, E. (1993). *The Book of Apples*. London: Ebury.

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.

Nerson, H. (2005). Effects of fruit shape and plant density on seed yield and quality of squash. *Scientia Horticulturae, 105*, 293–304.

Ngouajio, M., Kirk, W., & Goldy, R. (2003). A simple model for rapid and nondestructive estimation of bell pepper fruit volume. *HortScience, 38*(4), 509–511.

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* (pp. 72–77). Tokyo: National Institute of Agrobiological Resources, Ministry of Agriculture, Forestry and Fisheries.

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., Menesatti, C., Costa, C., Paglia, G., De Salvador, F. R., & Lolletti, D. (2010). Image analysis techniques for automated hazelnut peeling determination. *Food and Bioprocess Technology, 3*(1), 155–159.

Pallottino, F., Costa, C., Menesatti, P., & Moresi, M. (2011). Use of image analysis to assess the effective surface area of orange fruit submitted to uniaxial compression tests. *Journal of Food Engineering, 103*, 308–316.

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, P. A. (1959). Linkage of fruit shape and color genes in Capsicum. *Genetics, 44*, 407–419.

Rabiei, B., Valizadeh, M., Ghareyazie, B., Moghaddam, M., & Ali, A. J. (2004). Identification of QTLs for rice grain size and shape of Iranian cultivars using SSR markers. *Euphytica, 137*, 325–332.

Riyadi, S., Ishak, A. J., Mustafa, M. M., & Hussain, A. (2008). Wavelet-based feature extraction technique for fruit shape classification. In: 5th International Symposium Mechatronics and Its Applications, 2008. 27–29 May 2008: 1–5.

Rodriguez, A., Ehlenberger, D. B., Hof, P. R., & Wearne, S. L. (2006). Rayburst sampling, an algorithm for automated three-dimensional shape analysis from laser scanning microscopy images. *Nature Protocols, 1*, 2152–2161.

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.

Rohlf, F. J., & Bookstein, F. L. (1990). *Proceedings of the Michigan Morphometrics Workshop. Special Publication No. 2*. Ann Arbor: The University of Michigan Museum of Zoology.

Ruiz-Altisent, N., Ruiz-Garcia, L., Moreda, G. P., Lu, R., Hernandez-Sanchez, N., Correa, E. C., et al. (2010). Sensors for product characterization and quality of specialty crops - A review. *Computers and Electronics in Agriculture, 74*, 176–194.

Sabatier, R., Vivein, M., & Amenta, P. (2003). Two approaches for discriminant partial least square. In M. Schader, W. Gaul, & M. Vichi (Eds.), *Between data science and applied data analysis*. Berlin: Springer.

Sabliov, C. M., Boldor, D., Keener, K. M., & Farkas, B. E. (2002). Image processing method to determine surface area and volume of axi-symmetric agricultural products. *International Journal of Food Properties, 5*(3), 641–653.

Sadrnia, H., Rajabipour, A., Jafary, 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), 68–70.

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, R. R. (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, R. A. (1985). *Cherry descriptor list*. Rome: IBPGR.

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). Hortcouture Fashion Food Conference. Orchardist NZ, Feb, 37–50.

Shi, C. H., & 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., & Starkeya, 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 E. S. Gelsema & L. N. Kanals (Eds.), *Pattern recognition in practice II*. Amsterdam: Elsevier.

Smith, M. W. G. (1971). *National apple register of the United Kingdom*. London: Ministry of Agriculture, Fisheries and Food.

Spinks, G. T. (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, J. F., Alonso, N., Buxò i Capdevila, R., Chatti, N., Fabre, L., Fiorentino, G., et al. (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, R. A. (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.

Tominaga, Y. (2006). Comparative study of class data analysis with PCA-LDA, SIMCA, PLS, ANNs, and k-NN. *Chemometrics and Intelligent Laboratory Systems, 49*(1), 105–115.

Tufts, W. P., & Hansen, C. J. (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: USDA.

UPOV (1976). Guidelines for the conduct of test for distintness, homogeneity and stability of the cherry. UPOV, TG/35/3.

van der Knaap, E., & Tanksley, S. D. (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.

van Eck, J. W., van der Heijden, G. W. A. M., & Polder, G. (1998). Accurate measurement of size and shape of cucumber fruits with image analysis. *Journal of Agricultural Engineering Research, 70*, 335–343.

Venora, G., Grillo, O., Shahin, M. A., & Symons, S. J. (2007). Identification of Sicilian landraces and Canadian cultivars of lentil using an image analysis system. *Food Research International, 40*, 161–166.

Venora, G., Grillo, O., Ravalli, C., & Cremonini, R. (2009). Identification of Italian landraces of bean (*Phaseolus vulgaris* L.) using an image analysis system. *Scientia Horticulturae, 121*, 410–418.

Wang, T. Y., & Nguang, S. K. (2007). Low cost sensor for volume and surface area computation of axi-symmetric agricultural products. *Journal of Food Engineering, 79*, 870–877.

Wang, D. Y., Zhang, X. F., Zhu, Z. W., Chen, N., Min, J., Yao, Q., et al. (2005). Correlation analysis of rice grain quality characteristics. *Acta Agronomica Sinica, 31*, 1086–1091.

Webb, B. D. (1991). Rice quality and grades. In B. S. Luh (Ed.), *Rice utilization, Volume II* (2nd ed., pp. 89–119). New York: AVI book.

White, A. G., & Alspach, P. A. (1996). Variation in fruit shape in three pear hybrid progenies. *New Zealand Journal of Crop and Horticultural Science, 24*, 409–413.

White, A. G., & Bailey, D. G. (1995). Digital imaging: a useful technique for analysing fruit shape in pears. *Fruit Varieties Journal, 49*(4), 224–226.

White, A. G., Alspach, P. A., Weskett, R. H., & Brewer, L. R. (2000). Heritability of fruit shape in pears. *Euphytica, 112*, 1–7.

Xu, Z. J., Chen, W. F., Ma, D. R., Lu, Y. N., Zhou, S. Q., & Liu, L. X. (2004). Correlations between rice grain shapes and main qualitative characteristics. *Acta Agronomica Sinica, 30*, 894–900.

Xiao, H., Jiang, N., Schaffner, E., Stockinger, E. J., & Van der Knaap, E. (2008). A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science, 319*, 1527–1530.

Xiaobo, Z., Jiewen, Z., Yanxiao, L., Jiyong, S., & Xiaoping, Y. (2008). Apples shape grading by Fourier expansion and genetic program algorithm. In: 4th International Conference on Natural Computation IEEE, 18–20 Oct. 2008: 85–90.

Yadav, B. K., & Jindal, V. K. (2001). Monitoring milling quality of rice by image analysis. *Computers and Electronics in Agriculture, 33*, 19–33.

Zelditch, M. L., Swiderski, D. L., Sheets, H. D., & Fink, W. L. (2004). *Geometric morphometrics for biologists: A primer* (p. 443). San Diego: Elsevier.

Zheng, T. Q., Xu, J. L., Li, Z. K., Zhai, H. Q., & Wan, J. M. (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, A. B., 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.