The final publication is available at

http://dx.doi.org/10.1016/j.compag.2010.02.001

Automatic detection of skin defects in citrus fruits using a multivariate image analysis approach

F. López-García[1], G. Andreu-García[1], J. Blasco[2], N. Aleixos[3], J.M. Valiente[1]

[1] Instituto de Automática e Informática Industrial. Universidad Politécnica de Valencia. Camino de Vera s/n, 46022, Valencia, Spain. {flopez,gandreu,jvalient}@disca.upv.es

[2] Centro de Agroingeniería. Instituto Valenciano de Investigaciones Agrarias. Cra. Moncada-Náquera, Km. 5, 46113 Moncada, Spain. blasco_josiva@gva.es

[3] Instituto en Bioingeniería y Tecnología Orientada al Ser Humano. Universidad Politécnica de Valencia. Camino de Vera s/n, 46022 Valencia, Spain. naleixos@dig.upv.es

Abstract
One of the main problems in the citrus postharvest is the detection of visual defects in order to classify the fruit depending on their appearance. Species and cultivars of citrus present a great unpredictability in textures and colours that makes difficult to develop a general and unsupervised method able to perform this task. In this paper we study the use of a general approach originally developed for the detection of defects in random colour textures. It is based on a Multivariate Image Analysis strategy and uses Principal Component Analysis to extract a reference eigenspace from a matrix built by unfolding colour and spatial data from samples of defect-free peel. Test images are also unfolded and projected onto the reference eigenspace and the result is a score matrix which is used to compute defect maps based on the T2 statistic. In addition, a multiresolution scheme is also introduced in the original method to speed up the process. Unlike the techniques commonly used for the detection of defects in fruits, this is an unsupervised method that needs few samples to be trained. It is also a simple approach suitable for real-time compliance. Experimental work was performed on 120 samples of oranges and mandarins from four different cultivars. The success ratio for the detection of individual defects was 91.5%, while the classification ratio of damaged/sound samples was 94.2%. These results show that the studied method can be suitable for the task of citrus inspection.

1. Introduction

The automatic inspection of quality in the agro-industry is becoming of paramount importance in order to decrease production costs and increase quality standards as stated by Chen et al. (2002). On the packing lines, where external quality attributes are currently inspected visually, machine vision is providing a way to perform this task automatically. Vízhányó and Felföldi (2000) and Aleixos et al. (2002) reported automatic systems for grading the quality of mushrooms and citrus respectively using machine vision. Blasco et al. (2009) developed a prototype to inspect Satsuma (Citrus unshiu) segments and classify performing a morphological analysis of each segment. The reduction in the price of machine vision equipment, together with the increased processing capacity of computers now allow more complex analyses of images. The detection of defects and the quality control of the product is one of the major applications of machine vision in this field (Brosnan and Sun, 2002). The scientific literature contains a number of examples of the application of various techniques of image analysis to find external defects or the presence of foreign objects in fruit. This automatic detection of defects or blemishes in fruits is particularly important on packing lines, where the solution usually adopted to inspect the whole surface of the fruit is to force a rotational movement by means of bi-conical rollers (Bennedsen et al., 2005). One of the main problems related with these systems is how to distinguish between true defects and other natural elements in the fruits as the stem or the calyx (Penman, 2001). Moreover, to differentiate among different types of defects during postharvest inspection is increasingly important in order to enhance the possibilities of marketing the fruit according to their actual quality. In this sense, Blasco et al. (2007) presented a work to identify 11 different types of defects of citrus by analysing different spectral areas like ultraviolet or near infrared (NIR). The process is simpler in those fruits that have a uniform colour and becomes more

complicated when multiple colours appear (Leemans et al., 1999) or in fruits whose skin presents a rough texture, as in the case of citrus. One of the main problems arising from the analysis of natural objects is the wide diversity of colours and textures that may be present in them. Fruits belonging to the same commercial batch can have different colours depending on the stage of maturity. The colour of particular defects in some fruits can even be similar to the colour of the sound peel of other fruits. This variability means that automatic inspection systems have to be frequently retrained in order to adapt the inspection to these changes, and this is a process that needs to be carried out by expert operators who must spend long periods of time performing this task. Unsupervised techniques are important in these cases. Several attempts to develop unsupervised systems have been reported. Blasco et al. (2007) develop a computer vision system for the inspection of citrus fruits based on unsupervised growing region algorithms, assuming that in a fruit with defects in the peel, the larger region corresponds to the sound peel and smaller areas were therefore defects. Moshou et al. (2004) report a system based on self-organising maps (SOM) and multilayer perceptron to analyse hyperspectral images of fields of wheat and detect the infestation of yellow rust. Using a quadratic discriminating model, they were able to distinguish between diseased and healthy spectra with a success rate of 95%, and by means of neural networks, the detection rate reached 99% correct classification. A back propagation artificial neural networks (BPANN) algorithm with one hidden layer was applied by Li et al. (2002) to differentiate between stem-calyx and true defect areas of apples; the test results showed that the accuracy of the network classifier was over 93%. However, in general, but in particular in citrus fruits, the development of methods that do not require supervised training, which are also able to operate independently with different batches and cultivars of fruits, remains a challenge.

Textural analysis can improve the detection of defects, since the roughness appearance of fruit surface in sound and defective areas is commonly different. Texture can be used to differentiate between these two regions of interest. For instance, Kavdir and Guyer (2004) use textural features based on spatial grey-level dependence matrices (SGDM) to process monochromatic images of apples. In order to detect bruises in the skin of the apples and discriminate them from the stem/calyx, they compare different techniques such as k-NN (nearest neighbour), BPANN, DT (decision trees) and a Bayesian classifier. A similar technique based on SGDM was used by Pydipati et al. (2006) to inspect defective and sound leaves from citrus trees using computer vision. A total of 39 textural features were used to identify different types of defects affecting the leaves.

Currently, in the area of Computer Vision, there are several methods oriented to the detection of defects in random colour textures (of natural or artificial materials) that can be useful for the task of citrus inspection. Among these methods Mäenpää et al. (2003) present a fast method that combines colour and texture properties. They chose colour percentiles as colour features and simple statistics computed in the spatial domain as texture descriptors. Texture features included multidimensional distributions of co-occurrences and signed differences, local binary patterns (LBP), edge-based features and Hough transform features. In the sets of colour percentiles, LBP and edge-based features, a feature selection procedure was performed. For classification they used a k-NN approach with parameter k equal to 3. An important drawback of the method is that it uses supervised training and large numbers of samples are required to build the reference model. Also, the model is based on the knowledge of defective and sound samples compiled in the training stage and thus it could prove inadequate for the detection of new types of defects (unpredictable new defects).

More recently, the Texems approach was proposed by Xie and Mirmehdi (2007). This method is based on the use of texture primitives representing the texture layout of normal samples. Textural primitive information is extracted, in an unsupervised manner, from a small number of training samples using Gaussian Mixture Modelling (GMM). Instead of generating them directly from the RGB (red, green, blue) channels, they use a reference eigenspace built from the colour data of pixels belonging to a small set of defect-free textures. RGB channels of a training image are then projected onto the reference eigenspace to achieve three decorrelated eigenchannels from which texems are extracted. This is performed to overcome the problem of high correlation among the RGB channels. In the defect detection stage, testing images are projected onto the reference eigenspace. Then, textures in the resulting eigenchannels are compared with the texems collected in the training stage and, finally,

defects are detected using a probability function built from Gaussian distributions. This method is relatively complex and therefore the computational costs can be also relatively expensive.

In general, a desirable feature for the automatic inspection methods is the capability to deal with new unpredictable defects that may occur during production. In this case, the proper approach should be to build a model of sound areas and compare new test images against this model, that is, perform novelty detection. Simple unsupervised training using few samples is another desirable feature, as conditions at the production lines can change frequently. In addition, real-time compliance is an important issue so that the overall production can be inspected at on-line rates. Thus, approaches with low computational costs are valuable.

In particular, for the inspection of citrus fruits in postharvest a desirable feature is the capability to easily deal with new types of samples. This idea fits with the problems found during the inspection. Due to the large variety of colours and textures present in citrus, inspection systems need frequent tuning to adjust the classification to the features of each batch of fruit, an also, if the cultivar changes the system needs to be fully retrained. Therefore, unsupervised and easy to train systems are needed.

The method we study in this paper (López-García et al., 2006) belongs to the group of methods oriented to the detection of defects in random colour textures. It is based on a Multivariate Image Analysis strategy, developed in recent years in the field of applied statistics, and uses the T2 statistic to detect defective pixels and build the corresponding map of defects. Principal Component Analysis is used to extract a reference eigenspace from a matrix built by unfolding colour and spatial data from samples of sound peel. Test images are unfolded and projected onto this reference eigenspace, and then the result, a matrix of scores, is used to compute the T2 values of pixels and build the map of defects. This method is consistent with the desirable features mentioned above for inspection methods. It builds a model of defect-free colour textures using unsupervised training and few samples. It considers any pixel that do not match the sound model as defective, thus it is adequate for the detection of new unpredictable defects. And also, it uses a simple approach that makes it suitable for real-time compliance. In the present work we extend the original method and include a multiresolution framework to speed-up the overall process and a post-processing stage to improve detection results. Extensive experimental work involving 462 different experiments is performed on the ground truth of oranges and mandarins, 120 samples coming from four different cultivars. Achieved results show that the method can be adequate for the task of visual quality inspection of citrus. It correctly detected 91.5% of individual defects present in fruits and achieved a classification ratio of damaged/sound samples of 94.2%.

2. Materials and methods
2.1. Acquisition of the images
The set of fruit used to carry out the experiments consisted of a total of 120 citrus (oranges and mandarins) coming from four different cultivars: Clemenules, Marisol, Fortune, and Valencia (30 samples per variety). This fruit was randomly collected from a citrus packing house and included early cultivars of mandarins (Marisol and Clemenules) and late cultivars (Fortune and Valencia) trying to generalise the experiments. A total of five fruits of each cultivar belonged to the extra category (MAPA, 1993) which means that they were fully free of defects. Other 25 fruits of each cultivar fitted secondary commercial categories and had several skin defects, trying to represent the cause of most important looses during post-harvesting. Each piece of this latest fruits contained one or more skin defects and diseases caused by anthracnose, stem-end injury, green mould, oleocellosis (rind-oil spots), presence of scales, wind scarring, thrips scarring, sooty mould and phytotoxicity. The fruits were inspected visually and any defects that were found were identified, labelled and numbered. Figure 1 shows samples of several fruits used in the experiments.
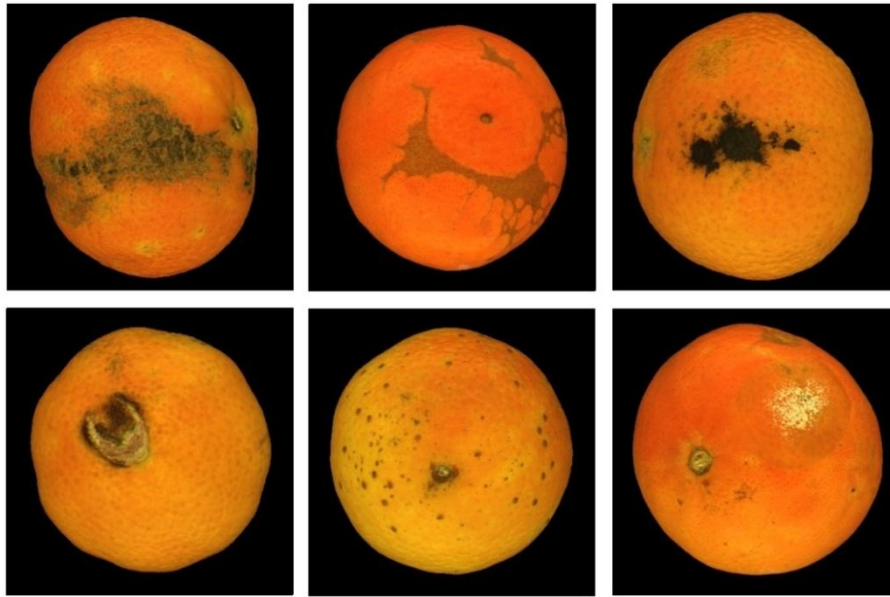
Figure 1. Samples of defect types found in the fruit. From top to bottom and left to right: wind scarring, thrips scarring, sooty mould, stem-end injury, scale infestation and green mould.

The colour image acquisition system used in the tests consisted of a 3-CCD (charge-coupled device) camera (Sony XC-003P) capable of acquiring images with a size of 768 x 576 pixels. The lighting system was composed of eight fluorescent tubes (daylight type, 25 w) and polarised filters were placed in front of the lamps and coupled to the camera lens in order to avoid bright spots in the scene caused by cross-polarisation. Images containing one fruit, with a resolution of 0.17 mm/pixel, were acquired off-line, the side of the fruit that contained the defects being presented manually to the camera.

2.2 Approach based on Multivariate Image Analysis

Multivariate Image Analysis (MIA), based on Principal Component Analysis (PCA), has been used to analyse colour (Geladi and Granh, 1996) and texture (Bharati and MacGregor, 2000) separately. However, López-García et al. (2006) propose a strategy that combines colour and texture information in the MIA approach by extending the input data structure introduced in the PCA to include colour and local spatial information (pixel's vicinity) in one array. In the matrix of unfolded data built from the image, each row is composed of the RGB values of one pixel and its vicinity. Pixel's vicinity is defined using a neighbourhood window with can be of different shapes and sizes. Figure 2 shows the unfolded RGB raw data of the image that corresponds to a square neighbourhood of size 3x3. In this case, for a given pixel *ith*, its *R* value is translated first, next the *R* value of the top-left neighbour and then the rest following a clockwise direction. By following the same approach with the *G* and *B* channels, each row of the matrix of unfolded data is created.
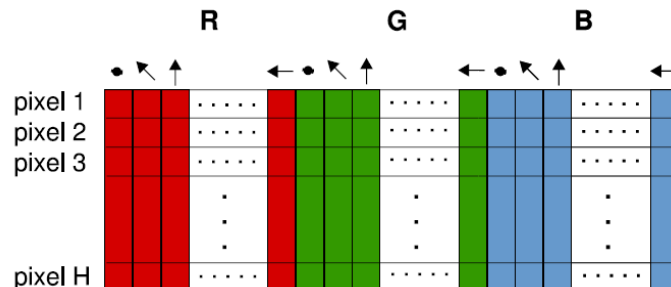

Figure 2. Matrix of unfolded RGB data using a 3x3 neighbourhood window.

To define the pixel's vicinity other window shapes apart of squares are suitable for use, such as hexagons or crosses. However, the most common is the square shape with size *WxW* with *W* varying in odd numbers, e.g. 3x3, 5x5, etc. Other unfolding orders are possible although it is not relevant because PCA compression is applied to data.

The method using the T^2 statistic works in the following way. A training image free of defects is used to compile a reference eigenspace that will be used to build the reference model of pixels belonging to defect-free areas and to perform defect detection in test images. Let I$_{NxM}$ be the training image and let

$$x_i = \begin{bmatrix} r_{(n,m)}, r_{(n-1,m-1)}, r_{(n-1,m)}, \cdots, r_{(n,m-1)} \\ g_{(n,m)}, g_{(n-1,m-1)}, g_{(n-1,m)}, \cdots, g_{(n,m-1)} \\ b_{(n,m)}, b_{(n-1,m-1)}, b_{(n-1,m)}, \cdots, b_{(n,m-1)} \end{bmatrix} \qquad (1)$$

be the colour-spatial feature of each pixel in I$_{NxM}$. Let X = $\{x_i \in R^K, i{=}1,2,\ldots,q\}$ be the set of q vectors from the pixels of the training image, where K is the number of pixels in the neighbourhood window multiplied by the number of colour channels. Let $\bar{x} = \frac{1}{q}\sum_{x \in X} x$ be the mean vector of X. An eigenspace is obtained by applying PCA to the mean-centred colour-spatial feature matrix X. The eigenvectors E = $[e_1, e_2, \ldots, e_L]$, $e_j \in R^K$ are extracted using Singular Value Decomposition. L is the number of selected principal components (L $\leq$ $rank$(X)). Now, by projecting the training image onto the reference eigenspace, a score matrix A is computed.

$$X = AE^T + R \rightarrow A = XE \qquad (2)$$

Where X is the unfolded image, E is the eigenspace and R is the residual matrix.

T^2 values of pixels are computed from the score matrix.

$$T_i^2 = \sum_{l=1}^{L} \frac{t_{il}^2}{s_l^2} \qquad (3)$$

where t_{il} is the score value of a given ith pixel in the lth principal component (lth eigenvalue) with variance s_l^2.

A behaviour model of normal pixels belonging to defect-free areas is created by computing the T^2 statistic for every pixel in a training image. T_i^2 is, in fact, the Mahalanobis distance of the projection of the pixel neighbourhood onto the eigenspace with respect to the model's centre of gravity (mean), and represents a measure of the variation of each pixel inside the model. In order to achieve a threshold level of T^2 defining the normal behaviour of pixels, a cumulative histogram is computed from the T^2 values of the training image. The threshold is then determined by choosing an extreme percentile in the histogram, commonly 90% or 95%. Any pixel with a T^2 value greater than the threshold will be considered a pixel belonging to a defective area. More than one training image can be used to achieve the reference eigenspace and compute the cumulative histogram.

Once the T^2 threshold has been set, it is used to inspect new images. For a new test image, the RGB data is unfolded as in figure 2 to achieve X$_{new}$. Then, X$_{new}$ is projected onto the reference eigenspace and a new score matrix is obtained, A$_{new}$ = X$_{new}$E. The T^2 value of each pixel is computed and whether the pixel is sound or defective is determined using the T^2 threshold. The T^2 image is created by locating only those pixels that exceed the threshold level in their original coordinates (Prats-Montalbán and Ferrer, 2007). Thus, since only the pixels belonging to defective areas are displayed, the defects in the image can be detected and isolated. Alternatively, the T^2 image can be translated into a map of defects by changing the T^2 values as follows:

$$\forall p_i \in I_{new}, p_i \in D_{map} \leftrightarrow T_i^2(p_i) > T_H^2 \qquad (4)$$

Where p$_i$ are the pixels, I$_{new}$ is the testing image, D$_{map}$ is the image map of defects and T_H^2 is the threshold for a given percentile. A flow diagram of the MIA method is shown in Figure 3.
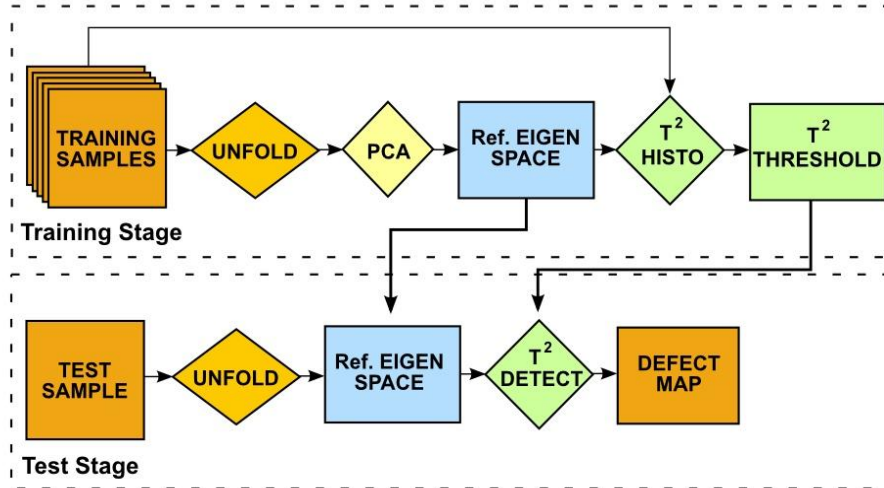
Figure 3. Flow diagram of the MIA approach.

2.3 Multiresolution and Post-Processing

In order to speed-up the process and improve the results of defect detection, we have extended the original method by including a multiresolution scheme and a post-processing stage. Multiresolution is introduced to capture defects, and parts of defects, of different sizes with minimum computational costs. In this case, we fix the size of the vicinity window to the minimum size of 3x3, and apply the method to the test sample at several scales (lower of equal to 1.0). This way, bigger defects are collected at lower scales, where the 3x3 window covers a bigger area than in the original scale. Lower scales lead to smaller matrices of unfolded data and then the process is accelerated because the major computational cost of the method is concentrated in the projection of data matrix onto the reference eigenspace (a matrix multiplication). By reducing the size of matrices the computational cost is significantly reduced. The final map of defects is built by combining the defect maps computed at each scale and then resized to the original size of samples (scale 1.0). It is noteworthy that when we use a 0.5 scale the size of matrices is reduced to 1/4 and the multiplication cost is reduced to 1/8.

In the original method, to capture defects of different sizes it is necessary to use different window sizes (e.g. 3 x 3, 5 x 5, 7 x 7, etc.) and join the resulting maps of defects. This approach implies high computing costs since the size of the rows in the unfolded matrices grows exponentially with the window size. The size of a row in the matrix of unfolded data is $rowsize = N^2*C$, when we use a neighbourhood window of NxN pixels and C different colour channels. Thus, for RGB and a 3x3 window the $rowsize$ is 27, for a 5x5 window it is 75, 147 for a 7x7 window, etc. Consequently, computational costs of handling the matrix of unfolded data increase exponentially with N.
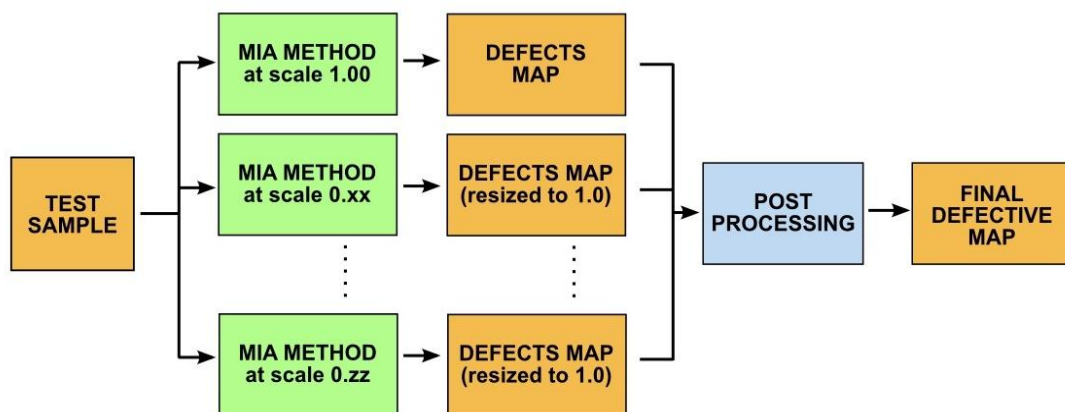

Figure 4. Global scheme of MIA approach including multiresolution scheme and post-processing.

Furthermore, we added a new post-processing stage to erase the fruits borders captured by the method (see Section 2.4) and improve the defective maps by cleaning noise and compact results. Post-processing is performed through simple morphological operations. The global scheme of the method, including multiresolution and post-processing, is shown in figure 4.

2.4 Experimental Work

First step in the experimental work was to select a set of defect-free samples for each cultivar in order to build the corresponding model of sound colour textures, which is built by extracting a reference eigenspace later used to compute the defects in test samples. For each cultivar we collected a set of 64 different sound patches from its corresponding samples in the ground truth (see figure 5). We used patches instead of complete sound samples to introduce in the model more different types of sound peels and collect as much as possible the wide variability of colours and textures present in the samples of the cultivar.
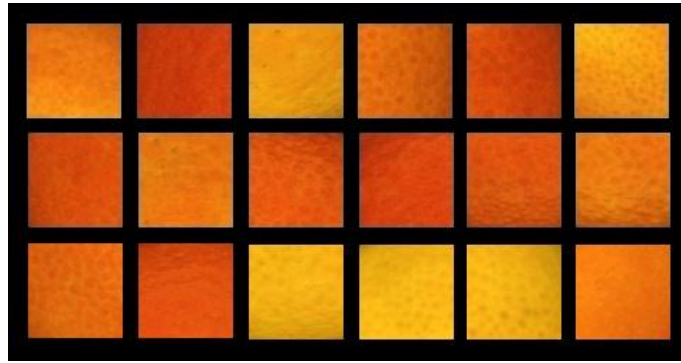

Figure 5. Several patches of defect-free peel areas extracted to build the model of sound colour textures corresponding to Clemenules cultivar.

From the patches we built a mosaic training image that was used to compute the reference eigenspace (the eigenspace compiles the data model of sound colour textures). Nevertheless, this approach introduces a problem involving the borders of fruits; they can be detected as defects because borders are not included within the collection of sound patches. We minimise this problem by erasing the borders in the defective maps by means of a morphological procedure introduced in the post-processing stage. This solution has been already proposed in previous works (Unay and Gosselin, 2007) and does not represent a major loose of information since, in an automatic sorting system, several images of each fruit would be acquired and analysed, and so the border areas could be inspected in other images.

Once the collection of patches was fixed, in order to tune the parameters of the method for each cultivar, we designed a set of experiments that involved applying the method to the ground truth of the cultivar and extract defect maps but varying in each experiment; the number of principal eigenvectors chosen to build the reference model, the percentile used to achieve the T^2 threshold, and the set of scales used in the multiresolution scheme. We varied the number of principal eigenvectors in the range [1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27], the T^2 threshold (the percentile chosen in the cumulative histogram) in the range [90%, 95%, 99%], and the set of scales used for the multiresolution scheme within the following possible combinations; (0.25, 0.12), (0.50, 0.12), (0.50, 0.25), (1.00, 0.12), (1.00, 0.25), (1.00, 0.50), (0.50, 0.25, 0.12), (1.00, 0.25, 0.12), (1.00, 0.50, 0.12), (1.00, 0.50, 0.25), (1.0, 0.50, 0.25, 0.12). This led to a total number of 462 different experiments carried out for each cultivar. To tune the parameters, that is, to choose those values that maximise the quality of the maps of defects, we marked manually the defective areas in the samples of the corresponding cultivar and then compared with the defective maps obtained by the method using the measures of Precision, Recall and F-score, which are defined as:

$$Precision = \frac{tp}{tp + fp} \qquad (5)$$

$$Recall = \frac{tp}{tp + fn} \qquad (6)$$

$$F = 2 * \frac{Precision * Recall}{Precision + Recall} \qquad (7)$$

Where tp (true positives) is the number of pixels marked and correctly detected, fp (false positives) is the number of pixels not marked but detected, and fn (false negatives) corresponds to the number of pixels marked but not detected. Precision is defined as the ratio between pixels correctly detected and the total number of pixels detected, whereas Recall is the ratio between pixels correctly detected over the total number of pixels marked. Precision can be seen as a measure of exactness (fidelity), Recall is a measure of completeness, and the F-score is a measure that combines Precision and Recall through their harmonic mean.

Once the set of experiments was performed for each cultivar, the mean values of previous measures were computed. Then we selected the most balanced result taking into account the achieved measures but also the corresponding computational cost. In table 1 we show for the Clemenules cultivar the combinations of factors that achieved the best results of Precision, Recall and F-score.

Table 1. Best results achieved in Clemenules experiments.

Scales	Eigenvectors	Percentile	Precision	Recall	F-score
0.50, 0.12	13	95	0.57	0.65	0.55
0.50, 0.25	**11**	**95**	**0.60**	**0.61**	**0.54**
0.50, 0.25	13	95	0.59	0.64	0.55
1.00, 0.50	9	95	0.66	0.59	0.55
1.00, 0.50	13	95	0.57	0.69	0.55
1.00, 0.50	23	99	0.66	0.61	0.57
1.00, 0.50, 0.25	9	95	0.62	0.62	0.55

In order to selected the most balanced result we needed to choose the combination of factors that achieved the best F-score value with a good balance between Precision and Recall. But also we have to take into account the computational cost, which is significantly reduced when we use lower scales and less principal eigenvectors. Therefore, for the Clemenules cultivar we chose the following combination of factors; scales (0.50, 0.25), 11 principal eigenvectors and percentile 95. This is not the best result if we only take into account Precision, Recall and F-score, but it uses only two lower scales, a small number of principal eigenvectors (11 eigenvectors), and also Precision and Recall values are quite balanced.

Table 2. Best combinations of factors for each cultivar and the corresponding mean values of Precision, Recall and F-Score.

Cultivar	Eigenvectors	Percentile	Scales	Precision	Recall	F-Score
Clemenules	11	95	0.50, 0.25	0.60	0.61	0.54
Fortune	17	90	0.50, 0.25	0.54	0.69	0.56
Marisol	23	90	0.50, 0.25	0.62	0.58	0.53
Valencia	27	95	0.50, 0.12	0.64	0.67	0.62
Global Mean				0.60	0.64	0.56

Table 2 shows the best combination of factors selected for each cultivar. Averaged values of Precision and Recall may seem low but it does not mean that defects are not properly estimated. In figure 7

several examples show the difference between the original defects, the defects marked manually and the defects detected using the MIA approach. We can see that although the method does not fit precisely nor extensively the marked defects, it provides an estimation of them that is good enough to carry out the inspection task.

3 Results

Once we tuned the parameters, from the corresponding marked defects and defective maps we counted the actual defects, the correctly detected defects and the false detections for each cultivar of citrus. These results are shown in table 3 (percentage of false detections is provided with regards to the total number of defects plus the false detections). In this case the method achieves a correct classification ratio of 91.5%. Current European standards for citrus quality (MAPA, 1993) are subjective setting the size of the defects that fits each commercial category. Only imprecise definitions like "only very slight defects in the skin of the fruits are allowed in the extra category", are given. However, they allow several tolerances such as 5% of misclassification for extra category and 10% of misclassification for the rest of categories. For the fruits used in the experiments (those which contain defects), the misclassification allowed is 10% which fits in the achieved results.

Table 3. Detection results of individual defects.

Cultivar	Defects	Detected	False detections
Clemenules	238	211 (88.7%)	4 (1.7%)
Fortune	172	159 (92.4%)	10 (5.5%)
Marisol	195	185 (94.9%)	7 (3.5%)
Valencia	138	125 (90.6%)	6 (4.2%)
Total	**743**	**680 (91.5%)**	**27 (3.5%)**

An example of the analysis made by the system is shown in figure 6. In the left image, the defects visually found in a mandarin cv. Fortune has been marked using a graphical tool. The right image shows the defects detected by the algorithm in the same image.
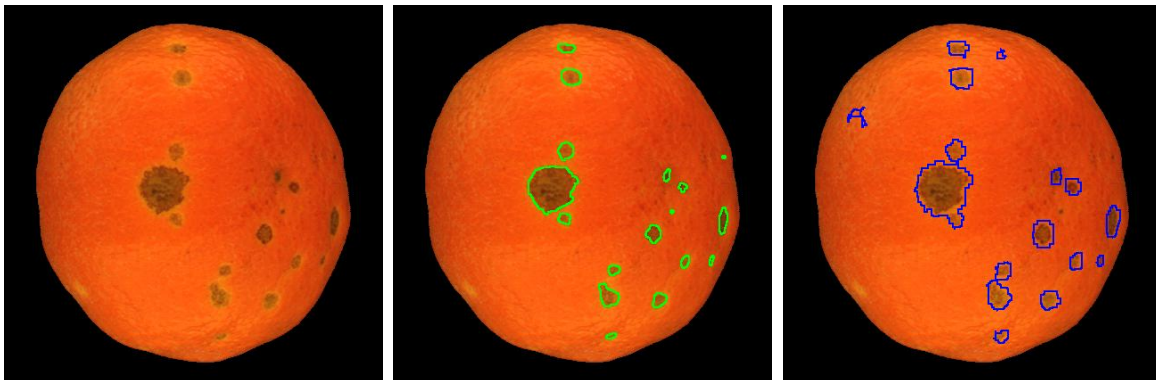

Figure 6. Left) original image. Middle) fruit with 17 defects encountered trough visual inspection and marked manually. Right) 15 defects and 2 false detections were detected by the vision system.

Results in table 3, show that although they are slightly lower than those obtained by Blasco et al. (2007) (94.2% of correct detections, 4.0% of false detections), the method can be adequate for the task of citrus inspection, having the advantage that it does not need to assume the hypothesis that most of the orange is a sound area, as it is necessary in the aforementioned work to carry out the unsupervised training. Most of false detections correspond to small areas, sometimes nearer to an actual defect (figure 7). This is explained by the wide variability of colours and textures among the samples within each cultivar. However, the capability of the method to deal with this wide variability is one of its strengths.

Table 4. Correct detection of each type of defect in all cultivars.

Type of defects	Correct detection
Scales	92.8%
Wind scarring	91.5%
Oleocellosis	87.2%
Phytotoxicity	93.8%
Thrips scarring	90.5%
Anthracnose	92.3%
Decay with spores	91.7%
Stem-end injury	100.0%
Sooty mould	100.0%
Total	**91.5%**
Stem end	100.0%

All types of defects present in the samples were detected in at least 87% of cases. Table 4 shows the ratio of correct detection of defects in all cultivars by defect types. Oleocellosis is the defect that had most problems of detection due their variability of appearance from light green to brown. Other defects like sooty mould, characterised by black stains on the skin, or stem-end injury were detected in 100% of cases. One of the current inconveniences is that the system does not discriminate the stem-end, which is detected in 100% of cases, from the true defects. However, further processing can allow making this discrimination and, even, among each type of defect. Having a large variability in the textures, colours and shapes of the defects, it is important to highlight that all of them have been detected with a relative high success rate, which makes the MIA approach a promising method to be implemented in an automatic sorting system. Since the MIA approach works with texture information instead of single pixel information, this method becomes appropriated in this type of scenes with great variability.

Table 5. Classification results of complete samples in damaged/sound classes.

Cultivar	Correct	Incorrect	Accuracy
Clemenules	28	2	93.3%
Fortune	28	2	93.3%
Marisol	29	1	96.7%
Valencia	28	2	93.3%
Total	**113**	**7**	**94.2%**

Table 5 shows the classification results of complete samples of oranges and mandarins between two classes; sound and damaged. As it was commented previously in Section 2.1, within the ground truth of each cultivar there were 25 samples with defects (damaged sample) and 5 samples free of defects (sound sample). From the point of view of the inspection algorithm, damaged samples are those with presence of defects after the application of the MIA method and sound samples are those without the presence of defects. In this sense a good accuracy ratio is achieved by the method in all cultivars (all of them are over 93% of success ratio) and in total (94.2% of global performance).
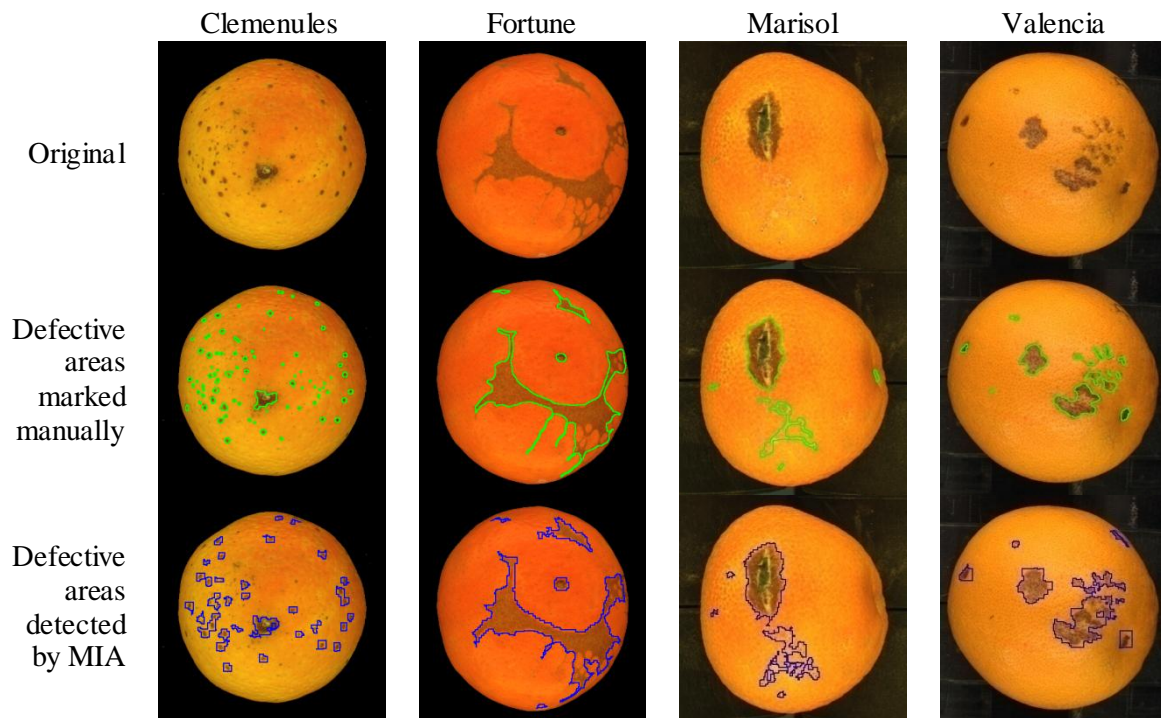
Figure 7. Original samples (top), manually marked defects in green (middle) and defective areas detected using the method in dark blue (bottom).

4   Conclusions

In this paper we have studied the detection of defects during the post-harvest quality inspection of citrus using a Multivariate Image Analysis approach. The method uses PCA analysis to extract a reference eigenspace that models sound colour textures from a matrix built by unfolding colour and spatial data of defect-free samples. Test images are also unfolded and projected onto the reference eigenspace to obtain a score matrix that is used to compute T^2 defect maps. To speed up the process a multiresolution scheme was introduced, and also, a post-processing stage to improve detection results.

The method is able to perform novelty detection and detect new unpredictable defects because it uses a model of sound colour textures and considers as defective any pixel that do not fit the model sufficiently. It also needs only a few samples to carry out the unsupervised training. For this reason, it is suitable for citrus fruit inspection as these systems need frequent tuning to adjust to the inspection of new cultivars and even the features of each batch of fruit within the same cultivar.

Experimental work was performed using 120 samples (images) of randomly selected oranges and mandarins belonging to four different cultivars. First, a set of experiments were carried out to tune the parameters of the method for each cultivar. These included the number of principal eigenvectors used to define the reference eigenspace, the T^2 threshold (percentile in the T^2 cumulative histogram) used to determine if pixels in test samples are sound or defective and, finally, the set of scales used in the multiresolution framework. Once the parameters were tuned, we compiled the results of defect detection, achieving 91.5% of correct defect detections and 3.5% of false detections. By using spectral and textural features, the main contribution of this method is the capability of detecting external defects in different cultivars of citrus that present different textures carrying out only a single previous unsupervised training. The method achieves a performance ratio of 94.2% to classify complete samples of fruit as damaged or sound. These results show that the studied method can be adequate for the task of citrus inspection.

References

Aleixos N; Blasco J; Navarrón F; 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.

Bennedsen, B.S., Peterson, D.L., Tabb, A. (2005) Identifying defects in images of rotating apples. Computers and Electronics in Agriculture 48 (2005) 92–102.

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.

Blasco, J., Aleixos, N., Moltó, E. (2007). Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm. Journal of Food Engineering, 81:535–543.

Bharati, M.H., MacGregor, J.F (2000). Texture analysis of images using Principal Component Analysis. In SPIE/Photonics Conference on Process Imaging for Automatic Control, (2000):27–37.

Brosnan, T., Sun, D.W. (2002). Inspection and grading of agricultural and food products by computer vision systems, a review. Computers and Electronics in Agriculture. 36, 193–213.

Chen, Y.R., Chao, K., Kim, M.S., (2002). Machine vision technology for agricultural applications. Computers and Electronics in Agriculture. 33, 173–191.

Geladi, P., Granh, H. (1996). Multivariate Image Analysis. Wiley, Chichester, England.

Kavdır, I., Guyer D.E. (2004). Comparison of Artificial Neural Networks and Statistical Classifiers in Apple Sorting using Textural Features. Biosystems Engineering 89 (3), 331–344.

Leemans, V., Magein, H., Destain, M.F. (1998). Defect segmentation on 'Jonagold' apples using colour vision and a Bayesian classification method. Computers and Electronics in Agriculture, 23(1) 43-53.

Li, Q., Wang, M., & Gu, W. (2002). Computer vision based system for apple surface defect detection. Computers and Electronics in Agriculture, 36(2–3), 215–223.

López-García, F., Prats, J.M., Ferrer, A., Valiente, J.M. (2006) Defect Detection in Random Colour Textures using the MIA T2 Defect Maps. Lecture Notes in Computer Science, 4142:752–763.

Mäenpää, T., Viertola, J., Pietikäinen, M. (2003). Optimising colour and texture features for real-time visual inspection. Pattern Analysis and Applications, 6(2003):169–175.

MAPA (1993). Norma de calidad para los cítricos (Citrus quality standards). Ministerio de Agricultura Pesca y alimentación, Spain.

Moshou, D., Bravo, C., West, J., McCartney, A., Ramon, H. (2004). Automatic detection of 'yellow rust' in wheat using reflectance measurements and neural networks. Computers and Electronics in Agriculture, 44 (3), 173-188.

Penman, D.W., (2001). Determination of stem and calyx location on apples using automatic visual inspection. Computers and Electronics in Agriculture 33, 7–18.

Prats-Montalbán, J.M., Ferrer, A. (2007). Integration of colour and textural information in multivariate image analysis: defect detection and classification issues. Journal of Chemometrics, 21(1-2):10–23.

Pydipati, R., Burks, T.F., Lee, W.S. (2006). Identification of citrus disease using color texture features and discriminant analysis Computers and Electronics in Agriculture 52, 49–59

Throop, J.A., Aneshansley, D.J., Anger, W.C., Peterson, D.L. (2005) Quality evaluation of apples based on surface defects: development of an automated inspection system. Postharvest Biology and Technology 36 281–290.

Unay D., Gosselin B. (2007). Stem and calyx recognition on 'Jonagold' apples by pattern recognition. Journal of Food Engineering 78, 597–605.

Vízhányó T, Felföldi, J. (2000) Enhancing colour differences in images of diseased mushrooms Computers and Electronics in Agriculture 26 187–198

Xie, X., Mirmehdi, M. (2007) TEXEMS: Texture Exemplars for Defect Detection on Random Textured Surfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(8):1454–1464.