ivia

Institut Valencià
d'Investigacions Agràries

Segmentation techniques in image analysis: a comparative study

Raffaele Vitale[a,*], José Manuel Prats-Montalbán[a,*], Fernando López-García[b], José Blasco[c], Alberto Ferrer[a]

[a]*Grupo de Ingeniería Estadística Multivariante, Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain*
[b]*Instituto de Automática e Informática Industrial, Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain*
[c]*Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Cra. Moncada-Náquera km 5, 46113, Moncada, Spain*

Abstract

Nowadays, the detection, localisation and quantification of different kinds of features in an RGB image (*segmentation*) is extremely helpful for e.g. process monitoring or customer product acceptance. In this article some of the most commonly used RGB image segmentation approaches are compared in an orange quality control case-study. ANalysis Of VAriance (ANOVA) and Correspondence Analysis (CA) are combined for determining their most relevant differences and highlighting their pros and cons.

Keywords: segmentation, colour information, textural information, graphs, Multivariate Image Analysis (MIA)

1. Introduction

The quality evaluation of products and processes mostly depends on the identification of features which are distinct from standard patterns. Nowadays, such an identification is more and more often carried out by Computer Vision Systems (CVS) comparing images collected along the production chain with reference ones. Nevertheless, if these references are not available, the aforementioned quality assessment cannot be addressed by direct comparison or pattern recognition. This is generally the case in e.g. fruit industry, whose products may exhibit completely

---

*Corresponding authors:
Telephone number: +34963877007 ext. 74949
Email address:* rvitale86@gmail.com (Raffaele Vitale);
jopramon@eio.upv.es (José Manuel Prats-Montalbán)

different shapes, colours and/or defects even if collected in the same area or from the same tree. Considering in addition that a single inspection line commonly controls about 9 tons of fruits per hour flowing at a very high speed, the problem seems far from being simple to solve. For this reason, CSV have had limited success and diffusion in this specific field, where Manual Visual Inspection (MVI) still plays a predominant role. On the other hand, MVI clearly lacks objectivity and is deeply influenced by the mood and/or the fatigue of the operators. Furthermore, it is biased by both between- and within-inspector variability. Therefore, it is important to develop automatic image processing techniques, capable of coping with these kinds of issues.

One of the first steps in image processing is the so-called *segmentation*. In the computer vision field, segmentation is usually defined as the process of partitioning a digital image into multiple segments (sets of pixels also known as super-pixels) and is aimed at distinguishing the different objects or regions of interest present in a picture. More precisely, it permits to assign each one of its pixels to a specific class or category so that those belonging to the same subgroup share certain characteristics [1]. Such a task can be accomplished by numerous disparate methods, which are commonly classified according to various criteria [2, 3], for instance whether they are supervised (they take advantage on the *a priori* knowledge on the class-belonging of a specific set of pixels for the assignation of new ones) or unsupervised (they look for clusters or groups of pixels, which are not known beforehand) [4]. Here, a classification based on the nature of the information exploited for assigning each pixel to a specific class or region and on the data modelling approach is proposed. Specifically, i) colour analysis-based, ii) texture analysis-based, iii) graph-based and iv) Multivariate Image Analysis (MIA)-based techniques are distinguished.

The main objective of this article is to compare some of the segmentation strategies representative of these four categories (see Table 1) and, concretely, determine which ones enable a correct identification of sound and green areas as well as of scale blemishes on the surface of several orange samples. Such strategies, namely Nearest Centroid (NC), K-Means (K-M), Standard Deviation, Range, Entropy, Felzenszwalb-Huttenlocher approach (F-H), Partial Least Squares Discriminant Analysis (PLSDA), Kernel Partial Least Squares Discriminant Analysis (K-PLSDA), Q-statistic and D-statistic, are probably the most popular, widespread and commonly used (in the forms and configurations described in the following section) to tackle problems of this type [2, 16, 17] and

Table 1 - Overview of the image segmentation techniques under comparison

Colour analysis-based segmentation techniques	Texture analysis-based segmentation techniques	Graph-based segmentation techniques	Multivariate Image Analysis (MIA)-based segmentation techniques
Nearest Centroid (NC) [5]	Standard deviation [7]	**Felzenszwalb-Huttenlocher approach (F-H)** [8]	Partial Least Squares Discriminant Analysis (PLSDA) [9, 10]
K-**Means (*K*-M)** [6]	Range [7]		Kernel Partial Least Squares Discriminant Analysis (K-PLSDA) [11, 12]
	Entropy [7]		Q-statistic [13, 14]
			D-statistic [13–15]

Bold text indicates unsupervised segmentation techniques

3

can be found coded in many ready-to-use routines for programming suites like Matlab [18] and R [19]. This study does, however, not aim at identifying which of them constitute the best segmentation methods to be applied in fruit quality control case-studies, but is rather an attempt of determining their most relevant differences and highlighting their pros and cons by means of classical statistical approaches, i.e. ANalysis Of VAriance (ANOVA) and Correspondence Analysis (CA). As far as the authors are concerned no similar works have been reported in the scientific literature before.

2. Methods

As detailed before, segmentation consists of assigning each pixel of an image to a specific category or class, which defines its nature. In this case, pixels belonging to sound, green or blemished orange peel areas are to be discriminated. This can be achieved according to different classification rules, depending on the algorithmic procedure underlying the adopted methodology, and exploiting different kinds of information that can be extracted directly from the picture.

2.1. Colour analysis-based segmentation techniques

These techniques directly deal with the three colour intensity values (red, green and blue) of the pixels of the various images under study. These images are unfolded into two-dimensional arrays so that each one of their rows contains the intensity values associated to a single specific pixel and subsequently analysed as such [3].

2.1.1. Nearest Centroid (NC)

The Nearest Centroid (NC) classification, also known as Nearest Prototype or Rocchio classification, is a non-parametric approach usually exploited for pattern recognition purposes [5]. Unlabelled pixels are classified as belonging to the category to whose centroid (estimated from a set of training data) their distance (here, euclidean) is minimum.

2.1.2. K-Means (K-M)

K-Means (K-M) is a vector quantisation algorithm, popular for unsupervised cluster analysis [6]. It aims at partitioning all the pixels of an image into K classes in an iterative fashion:

1. The centroids of the K classes are initialised;

2. Every pixel of the image is assigned to the class with the nearest centroid;

3. The centroids of the K classes are recalculated after the classification step;

4. The procedure is repeated until a convergence criterion is met.

Clearly, there is no need of training data to carry out a segmentation by K-M[i].

2.2. Texture analysis-based segmentation techniques

Texture analysis provides information about pixel intensity changes within pre-defined spatial domains [7]. In this work, the original RGB images are converted to grey-scale ones. Afterwards, each new intensity value is substituted by a first-order statistic, derived from a neighbouring window surrounding its corresponding pixel and capturing the aforementioned local variability[ii]. If an unlabelled pixel is found to be characterised by a value of this first-order statistic within a particular interval, determined based on training data and typical of one of the considered classes, it is assigned to such a category.

2.2.1. Standard deviation

Probably the best known index to perform a textural transform of RGB images is standard deviation. Generally, it applies that the smoother (more homogeneous) the examined image area, the lower its respective standard deviation values, and *vice versa*.

2.2.2. Range

As for standard deviation, also range permits to catch the textural information contained in RGB images. Smooth textures (homogeneous image regions) usually result in lower range values and *vice versa*.

2.2.3. Entropy

Entropy relates to the non-homogeneity of a scene [20]. It is estimated as:

$$E = -\sum_{i \in H} p_i \log p_i \qquad \sum_i p_i = 1 \tag{1}$$

---

[i]Each image is handled separately and independently from the others.

[ii]Specifically, 3×3 pixel windows were circumscribed to compute standard deviation, range and entropy.

being p_i the relative frequency of the i-th intensity value in the neighbouring window of interest and H its global intensity range.

Homogeneous image regions are known to feature low entropy, and *vice versa*.

2.3. Graph-based segmentation techniques

Graph-based segmentation techniques look at RGB images as kinds of networks, also known as *graphs*, i.e. sets of edges connecting certain pairs of adjacent pixels (vertices or nodes). Each one of these edges is associated to a weight, which is function of the dissimilarity between the pixels it connects (e.g. the difference in their intensity, location or some other spatial attribute). Segmentation is then addressed by finding a subgroup of meaningful edges, separating image regions encompassing pixels with different properties. Several decision criteria can be taken into account to achieve the identification of such a subgroup. Here, the algorithmic procedure proposed by Felzenszwalb and Huttenlocher in [8] (F-H) is applied, which attains the selection of the significant edge weights by adaptively assessing both pixel intensity differences across boundaries (to be maximised) and intensity differences between neighbouring pixels within single delimited areas (to be minimised)[iii]. Once performed a single segmentation per colour channel separately, the three of them are intersected in order to achieve the final pixel discrimination.

2.4. Multivariate Image Analysis (MIA)-based segmentation techniques

Multivariate Image Analysis (MIA) stands for the study of the image-intrinsic information through multivariate models. The basic principle of MIA is to unfold the investigated images into a matrix, say $\mathbf{X}$, whose row and column dimensions relate to the pixel mode and the colour and/or texture mode, respectively, and analyse the resulting data structure by means of e.g. Principal Component Analysis (PCA) [21] or Partial Least Squares regression (PLS) [22]. Details about the various ways of accomplishing this unfolding step[iv] are provided in [3].

[iii]Not even this segmentation strategy requires training data.

[iv]Here, both colour and textural information were combined to conduct the comparison.

2.4.1. Partial Least Squares Discriminant Analysis (PLSDA)

Partial Least Squares Discriminant Analysis (PLSDA) [9, 10] is the direct extension of PLS, developed for classification purposes: $\mathbf{X}$ is regressed via PLS on a dummy binary-coded response matrix, $\mathbf{Y}$, made up by a set of piled Z-dimensional row vectors[v], constructed so that, if their corresponding pixels are members of the z-th class, they have a 1-value in their z-th entry and 0-values in all the other ones. Whenever new images are collected and unfolded, the *a posteriori* probabilities that each one of their pixels belongs to each category are calculated. The assignation is finally carried out according to the highest-probability rule.

2.4.2. Kernel Partial Least Squares Discriminant Analysis (K-PLSDA)

Especially when more than two classes of pixels are dealt with, their possible overlap may dramatically jeopardise the quality of the final PLSDA-based segmentation. One option to overcome such an issue is resorting to a non-linear variant of PLSDA, Kernel Partial Least Squares Discriminant Analysis (K-PLSDA). K-PLSDA first converts $\mathbf{X}$ through a specific transformation function into a squared array, i.e. the so-called kernel matrix, which permits to map the original data onto a new feature space, allowing possible non-linear relationships to be described in a linear way [23]. The transformation function may be of different types (linear, polynomial, gaussian *etc.*) and its final form is commonly chosen in order to maximise the discriminant or predictive ability of the fitted model[vi]. The kernel matrix is then double-centred to preserve its symmetry [11, 12, 24] and subjected to classical PLSDA.

2.4.3. Q- and D-statistic

Alternatively to PLSDA, PCA can be applied to every single subset of $\mathbf{X}$ (namely $\mathbf{X}_z$), including only the information associated to the pixels belonging to the z-th category, so that Z

---

[v]Z equals the number of categories to be discriminated.

[vi]In this case, several functions with a growing degree of non-linearity were tested: linear, 2nd-9th order polynomial and Gaussian. Notice that the last one embraces a supplementary parameter, σ, which needs to be adjusted. Values of σ varying from 0.5 (yielding an extremely non-linear transformation) to 5000 (yielding an approximately linear transformation) were spanned.

independent class models are subsequently built. Thus, unlabelled pixels can be discriminated according to two distance indices:

- Q-statistic, reflecting their perpendicular distance to each model hyperplane [13, 14];

- D-statistic, reflecting the distance from the origin of each model hyperplane to their projection onto it [13–15].

For a fair comparison, such pixels are assigned to the class, for whose model they show the lowest $\frac{Q_z}{Q_{z,95}}$ or $\frac{D_z}{D_{z,95}}$ ratio, where $Q_{z,95}$ and $D_{z,95}$ denote an empirical 95% confidence threshold for Q and D, respectively, estimated from training data.

3. Dataset

30 RGB images of sweet oranges, collected from the Citrus Germplasm Bank at the Instituto Valenciano de Investigaciones Agrarias, were recorded by a Canon EOS 550D digital camera with a resolution of 0.0625 mm/pixel, installed in an inspection chamber (see Figure 1) internally illuminated by eight fluorescent tubes (OSRAM L 18W/965 BIOLUX, colour temperature = 6500 K, colour rendering index > 90%).

The samples were collected at different ripening stages and, therefore, their colour varied from green to orange depending on their maturity status. Most of them were also characterised by particular dark spots, mainly generated by California red scale (*Aonidiella aurantii*), long mussel scale (*Lepidosaphes gloverii*) and purple mussel scale (*Lepidosaphe backii*) infestations, which
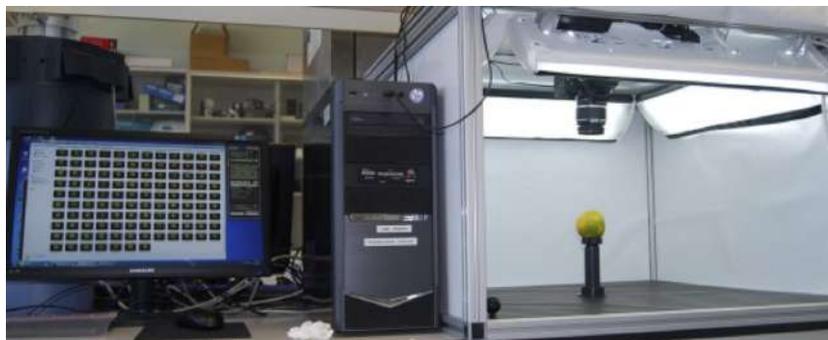

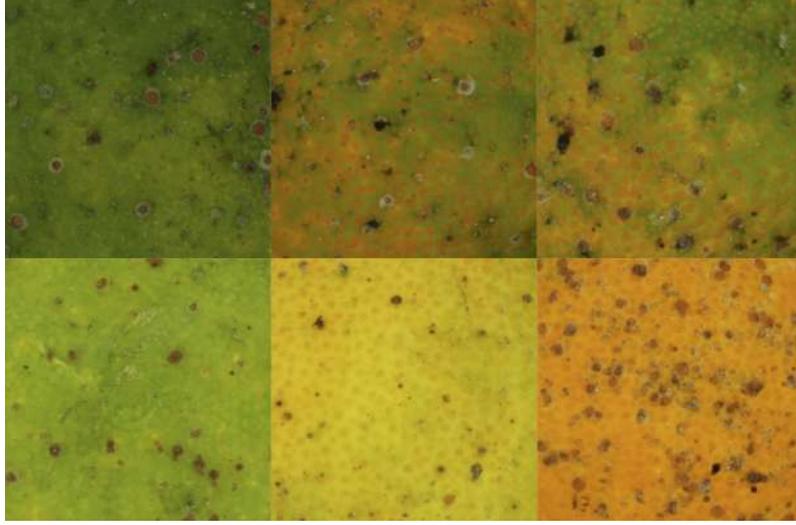Figure 1 - Image capturing computer vision system

Figure 2 - Details of 6 different orange images highlighting sound, green-coloured and blemished peel areas

depreciated their commercial value and whose detection is then of utmost importance from a quality control perspective. Hence, three distinctive features were observed: sound orange-coloured, green-coloured and blemished peel areas (see Figure 2).

Except for K-M and Graphs, 3330 training pixels per each one of the aforementioned regions were randomly selected from 5 images to establish the segmentation criteria or calibrate the segmentation/classification models. The potential of the single methods illustrated in the previous section was evaluated by exploiting the remaining 25 test pictures.

4. Comparative study

Once determined the membership class of all the test pixels, the image segmentations led to by the various adopted approaches were compared to reference ones manually elaborated by an expert operator. Segmentation accuracy degree was assessed in terms of F − score, computed for the n-th image and the z-th category as:

$$\text{F} - \text{score}_{n,z} = 2 \times \frac{\text{precision}_{n,z} \times \text{recall}_{n,z}}{\text{precision}_{n,z} + \text{recall}_{n,z}} \qquad \forall n = 1, 2, \ldots, 25 \qquad \forall z = 1, 2, 3 \qquad (2)$$

being

$$\text{precision}_{n,z} = \frac{\text{TP}_{n,z}}{\text{TP}_{n,z} + \text{FP}_{n,z}} \qquad (3)$$

9

$$\text{recall}_{n,z} = \frac{\text{TP}_{n,z}}{\text{TP}_{n,z} + \text{FN}_{n,z}} \tag{4}$$

where $\text{TP}_{n,z}$, $\text{FP}_{n,z}$ and $\text{FN}_{n,z}$ stand for *True Positives* (the number of pixels of the *n*-th image correctly identified as belonging to the *z*-th category), *False Positives* (the number of pixels of the *n*-th image mistakenly identified as belonging to the *z*-th category) and *False Negatives* (the number of pixels of the *n*-th image mistakenly identified as not belonging to the *z*-th category), respectively.

For every *z*-th class, statistically significant differences among the F-score indices associated to the techniques under study and constituting measures of their general performance were detected by a two-way ANalysis of VAriance (ANOVA) taking into account one fixed factor (i.e. segmentation method) and one blocking factor (i.e. orange sample). For a more comprehensive overview of their power, pros and cons, the contingency tables made up by the total number per class of True Positives, True Negatives (TN, the number of pixels correctly identified as not belonging to the *z*-th category), False Positives and False Negatives, resulting from the distinct methodologies, were investigated by means of Correspondence Analysis (CA) [25], which is conceptually similar to PCA, but was proposed for categorical rather than continuous data processing.

5. Results

Tables 2, 3 and 4 list the global values of TP, FP, FN, TN, precision, recall and F-score, related to the 25 test images, yielded by the different segmentation strategies for the sound peel area, green peel area and surface blemishes class, respectively. As the effect of the two factors included in the ANOVA models was found to be statistically significant ($p-$values $\ll 0.05$), the 95% Least Significant Difference (LSD) intervals were calculated for each single considered category. They are displayed in Figures 3a), 3b) and 3c)[vii]. Clearly, PLSDA and K-PLSDA outmatched the other approaches under study (F-H was as effective as them only when detecting surface blemishes). Colour analysis- (NC and *K*-M) and texture analysis-based techniques (Standard deviation, Range and Entropy) were not able to satisfactorily identify green and blemished peel areas. NC worked

---

[vii]Notice that the LSD intervals are centred around the average F-scores calculated over the 25 test images, which do not necessarily correspond to the global values shown in Tables 2, 3 and 4.

Table 2 - Global TP, FP, FN, TN, precision, recall and F-score yielded by the various approaches under study for the sound peel area class. Bold characters indicate the highest precision, recall and F-score values

Segmentation method	Sound peel area						
	True Positives (TP)	False Positives (FP)	False Negatives (FN)	True Negatives (TN)	Precision	Recall	F-score
NC	1310877	127934	480670	648080	0.9111	0.7317	0.8116
K-M	615968	269077	1175579	506937	0.6960	0.3438	0.4603
Standard deviation	825091	299396	966456	476618	0.7337	0.4605	0.5659
Range	540394	197470	1251153	578544	0.7324	0.3016	0.4273
Entropy	854303	324538	937244	451476	0.7247	0.4769	0.5752
F-H	1568223	583715	223324	192299	0.7287	**0.8753**	0.7953
PLSDA	1562519	20410	229028	755604	**0.9871**	0.8722	**0.9261**
K-PLSDA	156610	20506	224937	755508	**0.9871**	**0.8744**	**0.9274**
Q-statistic	733050	88726	1058497	687288	0.8920	0.4092	0.5610
D-statistic	1148340	27108	643207	748906	0.9769	0.6410	0.7741

11

Table 3 - Global TP, FP, FN, TN, precision, recall and F-score yielded by the various approaches under study for the green peel area class. Bold characters indicate the highest precision, recall and F-score values

Segmentation method	True Positives (TP)	False Positives (FP)	False Negatives (FN)	True Negatives (TN)	Precision	Recall	F-score
			Green peel area				
NC	280877	158046	413699	1476671	0.6399	0.4044	0.4956
K-M	230189	567488	464387	1067229	0.2886	0.3314	0.3085
Standard deviation	277198	598843	417378	1035874	0.3164	0.3991	0.3530
Range	347590	778806	346986	855911	0.3086	0.5004	0.3818
Entropy	282777	529609	411799	1105108	0.3481	0.4071	0.3753
F-H	509836	1276225	184740	358492	0.2855	0.7340	0.4111
PLSDA	634050	124372	60526	1510345	**0.8360**	**0.9129**	**0.8727**
K-PLSDA	639258	121954	55318	1512763	**0.8398**	**0.9204**	**0.8782**
Q-statistic	328220	210111	366356	1424606	0.6097	0.4725	0.5324
D-statistic	449446	155983	245130	1478734	0.7424	0.6471	0.6915

N.B. 2 out of the 25 test images did not contain pixels belonging to the green peel area class and were then excluded from the comparative study for this particular category.

Table 4 - Global TP, FP, FN, TN, precision, recall and F-score yielded by the various approaches under study for the surface blemishes class. Bold characters indicate the highest precision, recall and F-score values

| | Surface blemishes | | | | | | |
Segmentation method	True Positives (TP)	False Positives (FP)	False Negatives (FN)	True Negatives (TN)	Precision	Recall	F-score
NC	63125	625429	18300	1860707	0.0917	0.7753	0.1640
K-M	29949	703116	51476	1783020	0.0409	0.3678	0.0735
Standard deviation	66787	393741	14638	2092395	0.1450	0.8202	0.2465
Range	70302	492197	11123	1993939	0.1250	0.8634	0.2184
Entropy	58780	400583	22645	2085553	0.1280	0.7219	0.2174
F-H	50767	224394	30658	2261742	0.1845	0.6235	0.2847
PLSDA	69446	156762	11979	2329374	**0.3070**	0.8529	**0.4515**
K-PLSDA	69602	152471	11823	2333665	**0.3134**	0.8548	**0.4587**
Q-statistic	78198	1118405	3227	1367731	0.0653	**0.9604**	0.1224
D-statistic	76330	709570	5095	1776566	0.0971	**0.9374**	0.1760
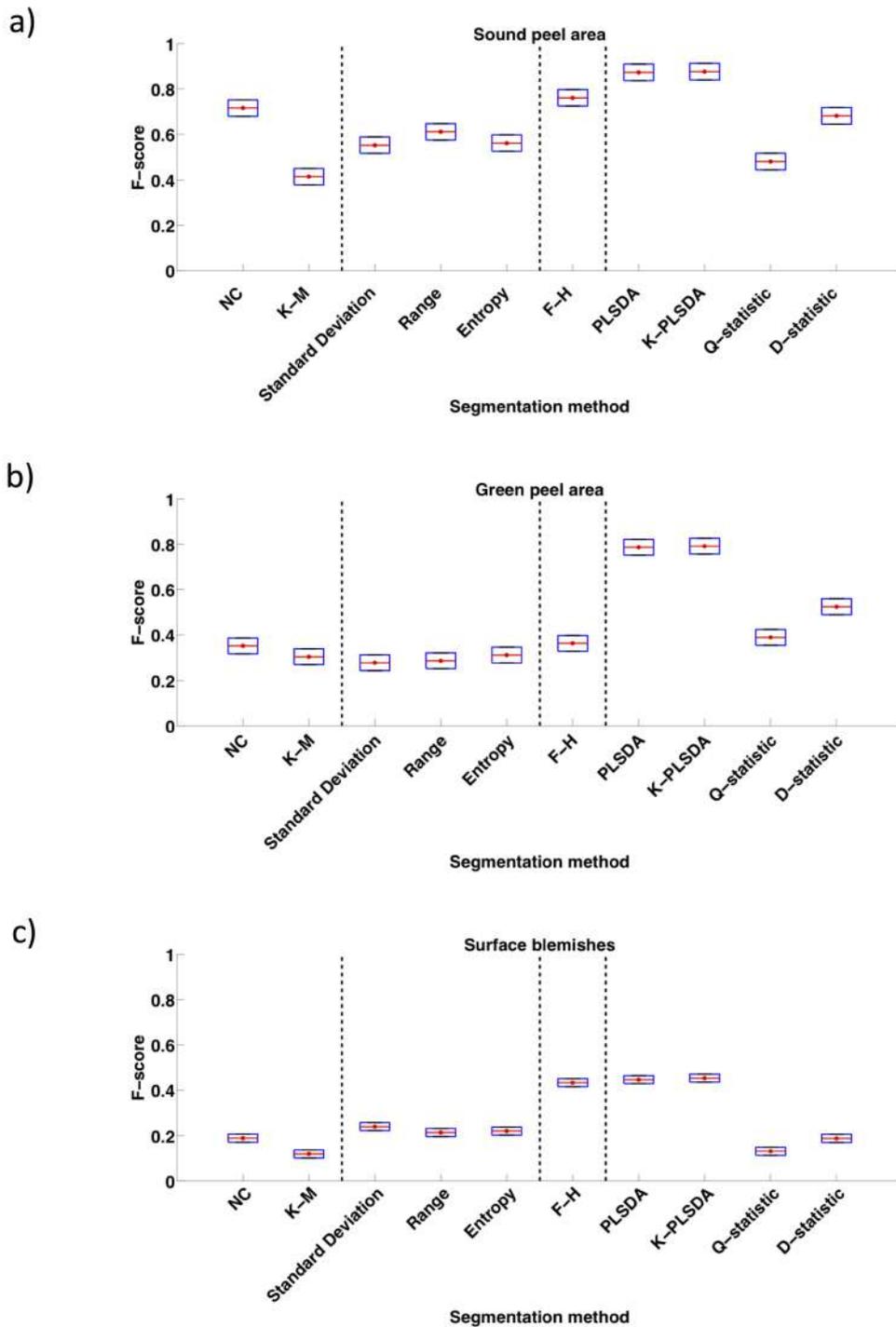
Figure 3 - 95% Least Significant Difference (LSD) intervals resulting from the ANOVA models built on the F-score indices yielded by the compared segmentation approaches for a) the sound peel area class, b) the green peel area class and c) the surface blemishes class. The vertical dotted lines separate the different categories of methods (see Table 1)

generally better than K-M while D-statistic always guaranteed a statistically significantly larger F-score than Q-statistic. It is also important to notice that the outcomes associated to the surface blemishes class are on average negatively biased: according to Table 4 most of the methods generated an acceptable quantity of TP and FN (high recall) but too many FP (low precision). This is due to the extreme variability affecting the intensity of the pixels belonging to it.

By looking at the CA bi-plots in Figures 4a), 4b) and 4c), some additional insights can be gleaned. As the two CA dimensions relate to TP/FN and TN/FP, respectively, the best segmentation strategies can be found lying in the TP/TN quadrant (highlighted). This is always the case for both PLSDA and K-PLSDA. In comparison with the other approaches, F-H produced a very large number of FP for the sound and the green peel area categories, but not for the surface blemishes one. That means F-H cannot adequately discriminate the pixels of the first two classes. Regarding the surface blemishes, Q-statistic, D-statistic, NC and K-M delivered the highest number of FP. Conversely, K-M returned a more considerable amount of FN than the other techniques for all the three categories as well as Q-statistic and Range for the sound peel area class and NC for the green peel area class.

6. Illustration case

The practical consequences of the previous results can be easily visualised by an illustrative example. Figures 5, 6 and 7 show the segmentations accomplished manually and by the different concerned methodologies for the sound peel areas, the green peel areas and the surface blemishes of one of the 25 test images. In addition to what remarked before, it is worth pointing out that:

- although K-PLSDA is the only non-linear statistical technique applied here, a strict similarity between it and classical (linear) PLSDA was observed. This is justified by the fact that, for the former, an approximately linear kernel function (Gaussian with a σ-parameter value around 500) was selected as the most appropriate to transform the analysed data structure, giving rise to a discrimination among the pixels of the three classes analogous to that enabled by PLSDA. It is then possible to conclude that no strong non-linear relationships needed to be modelled for a good classification, probably due to the limited number of han-
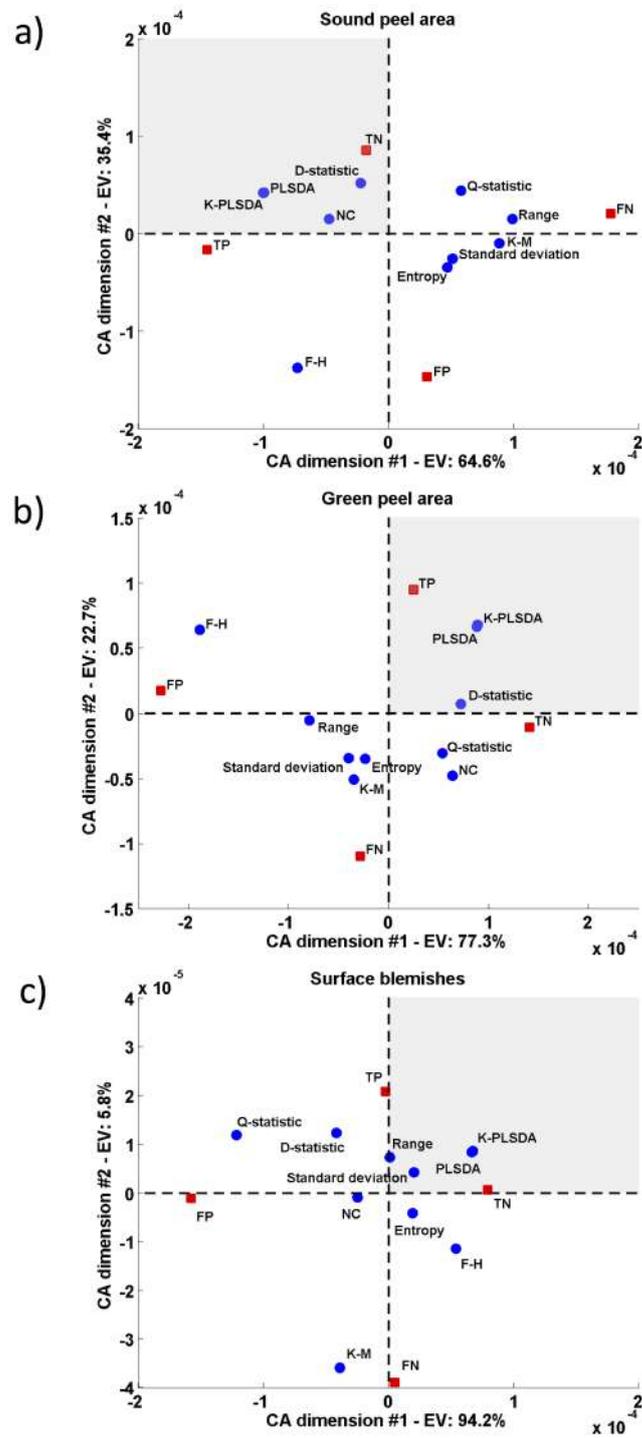
15

Figure 4 - Correspondence Analysis (CA) bi-plots for a) the sound peel area class, b) the green peel area class and c) the surface blemishes class. EV stands for Explained Variance. The TP/TN quadrant is highlighted
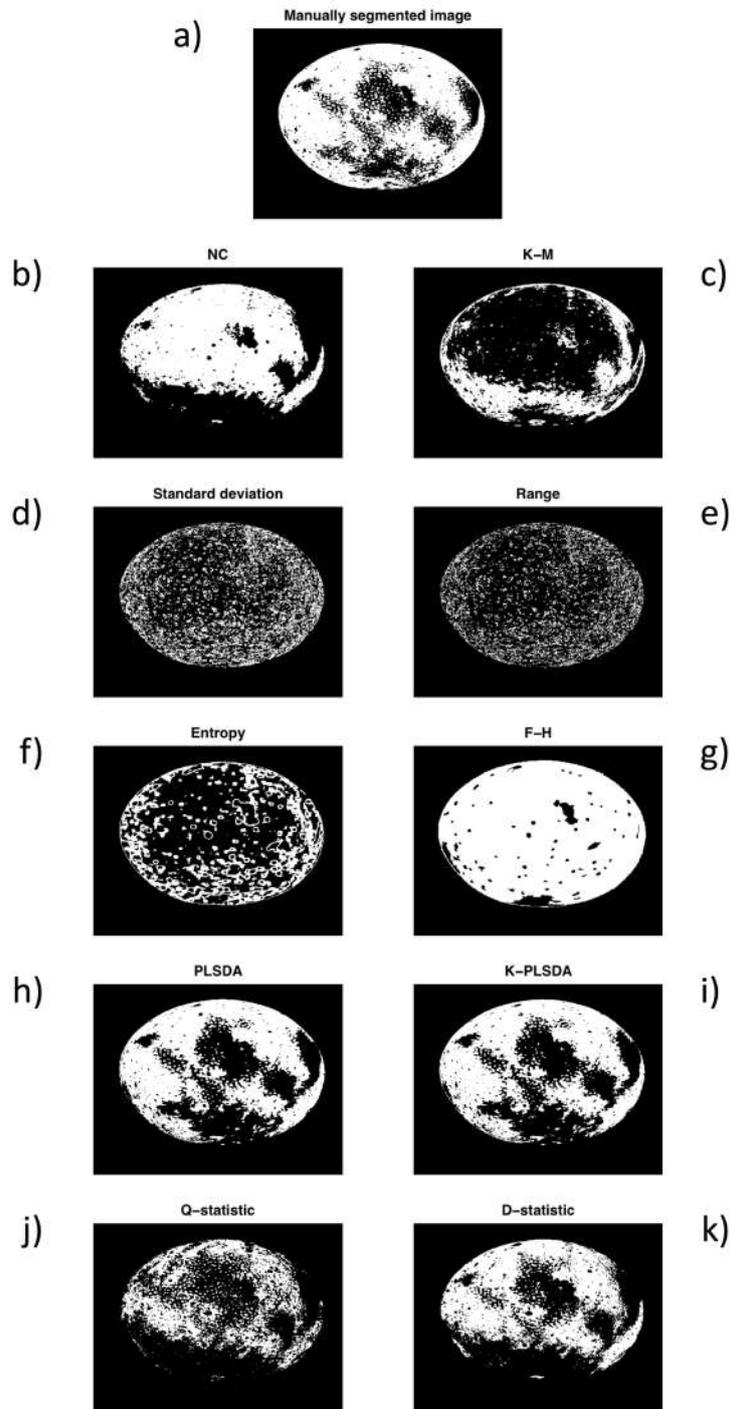
Figure 5 - Sound peel area: segmentations accomplished a) manually (reference) and by b) NC, c) *K*-M, d) Standard deviation, e) Range, f) Entropy, g) F-H, h) PLSDA, i) K-PLSDA, j) *Q*-statistic and k) *D*-statistic. White colour identifies the pixels assigned to the concerned class
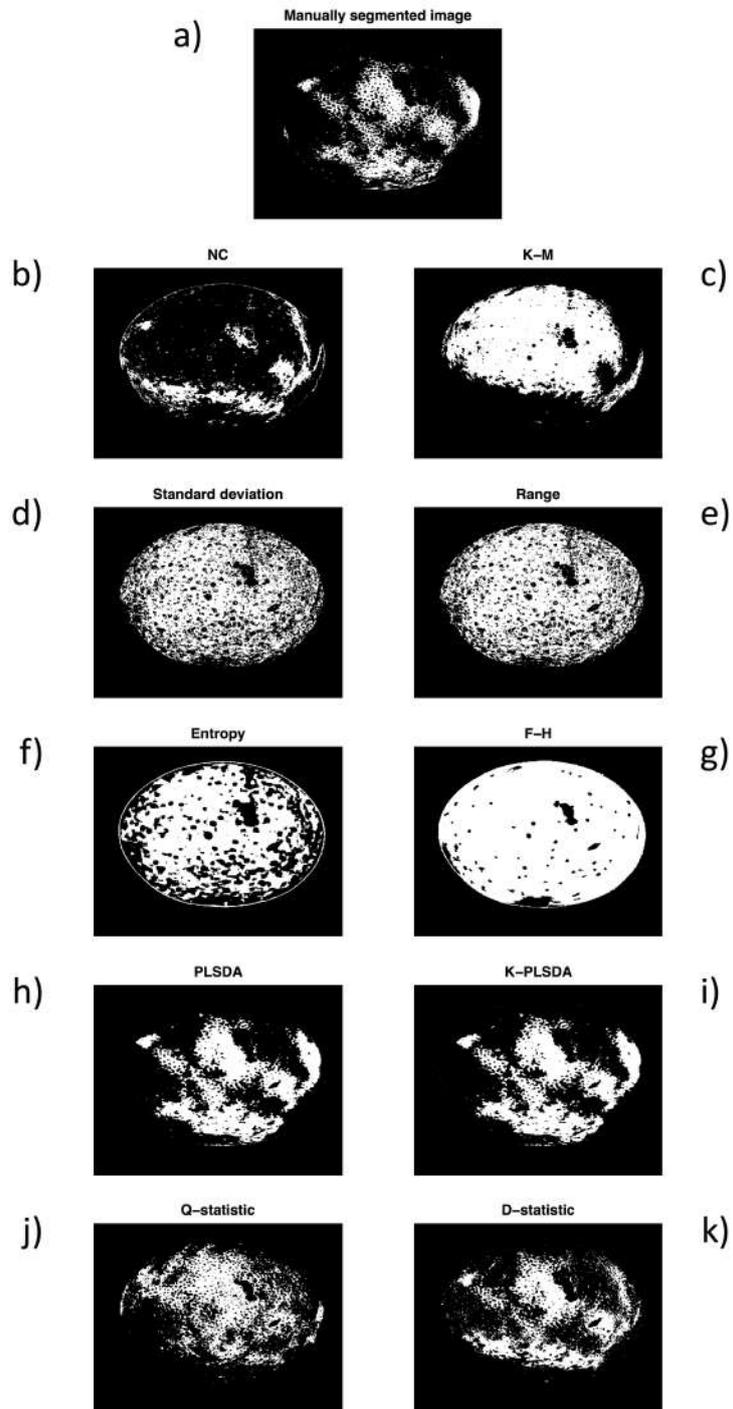
17

Figure 6 - Green peel area: segmentations accomplished a) manually (reference) and by b) NC, c) *K*-M, d) Standard deviation, e) Range, f) Entropy, g) F-H, h) PLSDA, i) K-PLSDA, j) *Q*-statistic and k) *D*-statistic. White colour identifies the pixels assigned to the concerned class
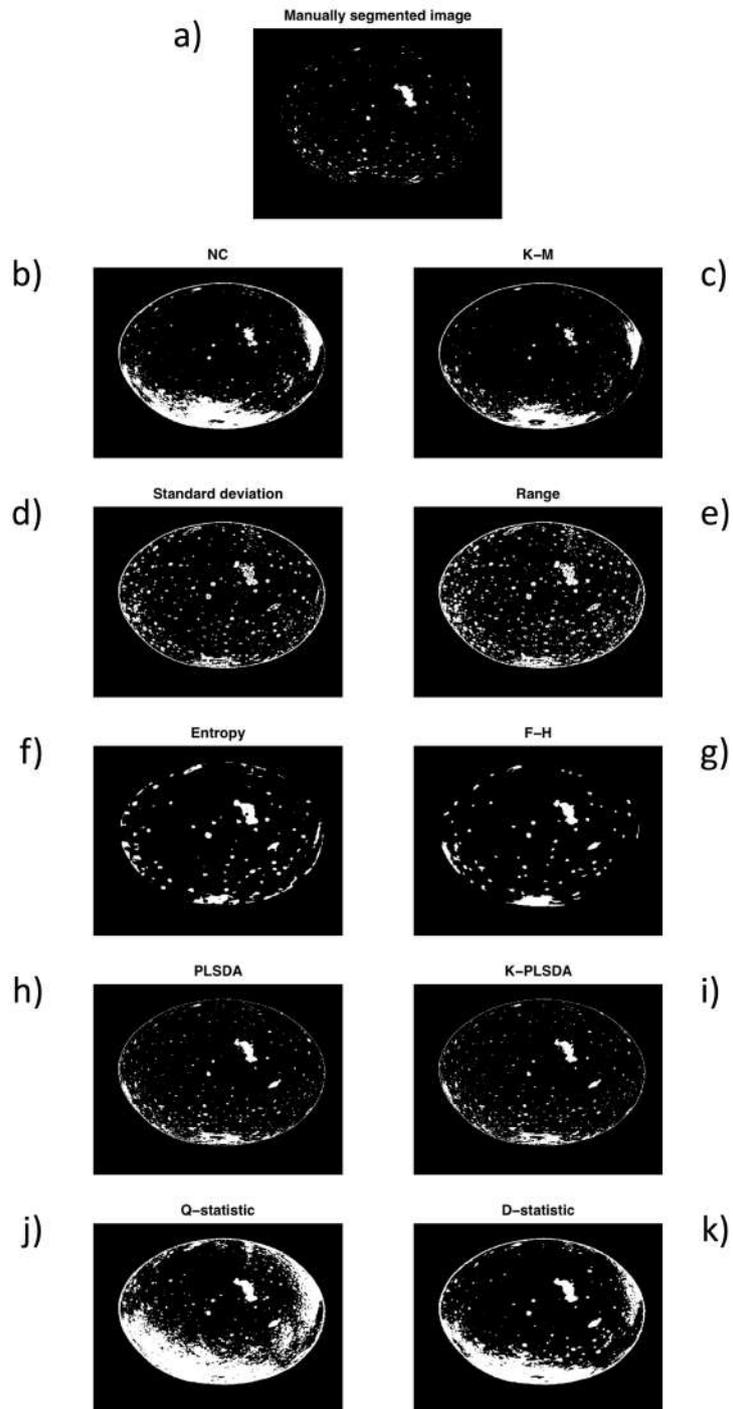
18

Figure 7 - Surface blemishes: segmentations accomplished a) manually (reference) and by b) NC, c) K-M, d) Standard deviation, e) Range, f) Entropy, g) F-H, h) PLSDA, i) K-PLSDA, j) Q-statistic and k) D-statistic. White colour identifies the pixels assigned to the concerned class

dled categories;

- generally, *K*-M exhibited a similar segmentation performance to NC. However, across all the test images, the three recognised clusters were not always corresponding to the same pixel classes. I.e. in the various images the first cluster sometimes captured the sound peel area category, sometimes the green peel area category and sometimes the blemished area category. This is a consequence of the unsupervised nature of the iterative algorithmic procedure *K*-M is based on, which makes the segmentation of series of different images not univocal. That might constitute a critical issue when image segmentation has to be performed automatically and sequentially. The reported outcomes also highlight this effect: Figures 5c) and 6c) clearly show that in such an illustrative example the first and the second cluster of pixels are inverted;

- solely for F-H sound and green peel areas could not be discriminated at all (Figures 5g) and 6g) are identical). This was found to be the main limitation of F-H and is the root cause of the high amount of FP it led to;

- all the methods suffered from the same issue when dealing with the surface blemishes class: too many pixels were incorrectly classified as belonging to it, confirming its rather large internal variability, which prevented a more satisfactory segmentation.

- it can also be said that both the techniques which performed the best in this particular application (i.e. PLSDA and K-PLSDA) are supervised. Notice this has to be looked at as a circumscribed rather than general conclusion. In fact, in many other fields of interest (e.g. remote sensing) that could easily not be the case.

7. Concluding remarks

A comprehensive comparative study among various segmentation methodologies (namely Nearest Centroid, *K*-Means, Standard Deviation, Range, Entropy, Felzenszwalb-Huttenlocher approach, Partial Least Squares Discriminant Analysis, Kernel Partial Least Squares Discriminant Analysis, *Q*-statistic and *D*-statistic) was carried out to determine which strategies could enable a correct

discrimination of sound orange-coloured, green-coloured and blemished areas on the peel of several orange samples. ANOVA-based LSD intervals highlighted that PLSDA and K-PLSDA outmatched the other techniques in terms of F-score, a general measure of segmentation accuracy and precision. Furthermore, CA permitted to more specifically appraise their pros and cons and recognise in a very simple and direct graphical way those strategies yielding higher/lower than average quantities of TP, FP, FN and TN, respectively. The final outcomes revealed that resorting to both colour and textural information in combination with Multivariate Image Analysis (MIA)-based segmentation approaches (i.e. PLSDA and K-PLSDA) might represent a suitable option when dealing with complex problems like the one at hand in this specific case.

8. Acknowledgements

9. References

[1] L. Shapiro, G. Stockman, Computer vision, 1st Edition, Prentice Hall Inc., Upper Saddle River, USA, 2001.

[2] R. Szeliski, Computer vision - Algorithms and applications, 1st Edition, Springer-Verlag Ltd., London, UK, 2011.

[3] J. Prats-Montalbán, A. de Juan, A. Ferrer, Multivariate image analysis: a review with applications, Chemometr. Intell. Lab 107 (2011) 1–23.

[4] M. Bevilacqua, R. Bucci, A. Magrì, A. Magrì, R. Nescatelli, F. Marini, Chemometrics in food chemistry, 1st Edition, Elsevier B.V., Oxford, UK, 2013, Ch. Classification and class-modelling, pp. 171–233.

[5] C. Manning, P. Raghavan, H. Schütze, Introduction to information retrieval, 1st Edition, Cambridge University Press, Cambridge, UK, 2008.

[6] J. MacQueen, Some methods for classification and analysis of multivariate observations, in: Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, University of California Press, Berkeley, USA, 1967, pp. 281–297.

[7] R. Haralick, Statistical and structural approaches to texture, P. IEEE 67 (1979) 786–804.

[8] P. Felzenszwalb, D. Huttenlocher, Efficient graph-based image segmentation, Int. J. Comput. Vision 59 (2004) 167–181.

[9] S. Wold, C. Albano, W. Dunn III, K. Esbensen, E. Hellberg, E. Johansson, M. Sjöström, Food research and data analysis, 1st Edition, Elsevier Applied Science, Barking, UK, 2009, Ch. Pattern recognition: finding and using regularities in multivariate data, pp. 147–188.

[10] M. Barker, W. Rayens, Partial Least Squares for discrimination, J. Chemometr. 17 (2003) 166–173.

[11] G. Postma, P. Krooshof, L. Buydens, Opening the kernel of kernel partial least squares and support vector machines, Anal. Chim. Acta 705 (2011) 123–134.

[12] R. Vitale, O. de Noord, A. Ferrer, A kernel-based approach for fault diagnosis in batch processes, J. Chemometr. 28 (2014) 697–707.

[13] J. Prats-Montalbán, A. Ferrer, Integration of colour and textural information in multivariate image analysis: defect detection and classification issues, J. Chemometr. 21 (2007) 10–23.

[14] J. Prats-Montalbán, Control estadístico de procesos mediante análisis multivariante de imágenes, Ph.D. thesis, Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universitat Politècnica de València (2005).

[15] F. López, J. Prats, A. Ferrer, J. Valiente, Defect detection in random colour textures using the MIA T^2 defect maps, Lect. Notes Comput. Sc. 4142 (2006) 752–763.

[16] P. Ho, Image segmentation, 1st Edition, In Tech d.o.o., Rijeka, Croatia, 2011.

[17] N. Pal, S. Pal, A review on image segmentation techniques, Pattern Recogn. 26 (1993) 1277–1294.

[18] MATLAB R2012b (8.0.0.783), The MathWorks Inc., Natick, USA.

[19] R Development Core Team, R: a language and environment for statistical computing, R Foundation for Statistical Computing, Wien, Austria (2008).

[20] R. Gonzalez, R. Woods, Digital image processing, 3rd Edition, Prentice Hall Inc., Upper Saddle River, USA, 2007.

[21] S. Wold, K. Esbensen, P. Geladi, Principal Component Analysis, Chemometr. Intell. Lab 2 (1987) 37–52.

[22] P. Geladi, B. Kowalski, Partial Least Squares Regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.

[23] D. Cao, Y. Liang, Q. Xu, Q. Hu, L. Zhang, G. Fu, Exploring nonlinear relationships in chemical data using kernel-based methods, Chemometr. Intell. Lab 107 (2011) 106–115.

[24] R. Vitale, O. de Noord, A. Ferrer, Pseudo-sample based contribution plots: innovative tools for fault diagnosis in kernel-based batch process monitoring, Chemometr. Intell. Lab 149B (2015) 40–52.

[25] H. Hirschfeld, A connection between correlation and contingency, Math. Proc. Cambridge 31 (1935) 520–524.