

Document downloaded from:

<http://hdl.handle.net/10251/50302>

This paper must be cited as:

Prats Montalbán, JM.; Ferrer Riquelme, AJ. (2014). Statistical Process Control based on Multivariate Image Analysis: A new proposal for monitoring and defect detection. *Computers and Chemical Engineering*. 71:501-511.
doi:10.1016/j.compchemeng.2014.09.014.



The final publication is available at

Copyright Elsevier

Statistical Process Control based on Multivariate Image Analysis: a new proposal for monitoring and defect detection

J. M. Prats-Montalbán* and A. Ferrer

Multivariate Statistical Engineering Group.

Department of Applied Statistics, Operations Research and Quality.

Universidad Politécnica de Valencia.

Cno. De Vera s/n, Edificio 7A, 46022, Valencia, Spain.

* Departamento de Estadística e IO Aplicadas y Calidad. Universidad Politécnica de Valencia. Cno. De Vera s/n, Edificio 7A, 46022, Valencia, Spain. Tlf: +34.96.387.70.07 ext. 74949, Fax: +34.96.387.74.99. E-mail: jopraron@eio.upv.es

Abstract

The monitoring, fault detection and visualization of defects are a strategic issue for product quality. This paper presents a novel methodology based on the integration of textural Multivariate image analysis (MIA) and multivariate statistical process control (MSPC) for process monitoring. The proposed approach combines MIA and p -control charts, as well as T^2 and RSS images for defect location and visualization. Simulated images of steel plates are used to illustrate the monitoring performance of it. Both approaches are also applied on real clover images.

Key Words: Multivariate Image Analysis (MIA), ARL, Control charts, RSS image, T^2 image, Wavelets.

1 INTRODUCTION

During the last 40 years visual inspection has been introduced in different industrial and technological fields for defect detection and classification. Two issues are critical for these purposes. One is the time required, because in processes usually many parts or samples are to be inspected, and inspection speed is a critical concern due to the high production rates (*e.g.* in ceramic tiles manufacturing). Another issue is the objectivity and consistency in the classification of the manufactured parts or in the monitoring of a process when establishing if it is in control or not. Visual inspection has been traditionally performed by well-trained operators in order to determine if the manufactured part meets the quality standards, presents any type of defect, fits a specific class, or if it comes from an in-control process. Examples can be found in the steel industry (Cardin et al., 2011); in the mining industry (Liu et al., 2005), or in the ceramic tile industry (López et al., 2007; Prats-Montalbán et al., 2008).

However, this type of visual inspection is not exempt of subjective criteria, no matter the experience of the operator. The reasons for this subjectivity are related to (i) the variability in the classifications made by the same operator, due to *e.g.* fatigue or loss of attention (repeatability), and (ii) the variability in the classifications made by different operators, related to the different visual perception characteristics of each person (reproducibility). Thus, it is necessary to develop reliable image-based monitoring systems with low repeatability and reproducibility variation.

Nowadays, cameras-based visual systems capture information from the process without interfering with the product due to their non-destructive and non-invasive nature, avoiding the aforementioned

problems of human visual inspection. This way, three different objectives are reached: (1) reducing expensive on-line sensors and laboratory sampling analyses; (2) avoiding tests that cannot be applied on the whole production as is the case of destructive tests to determine the stress endurance of a metallic, ceramic or any other kind of part; (3) allowing on-line characterization of products and processes that are not always amenable via traditional on-line instrumentation (Yu and MacGregor, 2003a, 2003b).

Once an image has been acquired from the process, it is possible to apply image analysis techniques to analyze it. This information is used for classification, defect detection or image object characterization. A well established image analysis technique in process monitoring is multivariate image analysis (MIA) (Geladi and Grahn, 1996; Prats-Montalbán et al., 2011). Different works have treated the monitoring and classification problems based on MIA. An excellent review about the application of MIA in industrial environments for monitoring, prediction and control can be found in Duchesne et al. (2012).

In a recent work, Megahed et al. (2011) have performed a review on the use of control charts with image data. Many different methods are presented and commented. Their main drawback is that they do not provide specific spatial control charts able to precisely localize the defect detected by the (whole) image analysis control systems. The paper claims for further study in this topic, in particular *“the study of the statistical properties and performance of many control charting methods”*. It also states that there is a lack of study of the *“performance of existing methods as well as understand the differences between competing approaches.”*

The present work tries to address these last two goals. In particular the paper is focused on the development of a MIA-based statistical process control (SPC) monitoring scheme able to check the

adequacy of the product with regards to the established standards and, furthermore, locate and isolate any kind of defect. We do not address the problems related to illuminating conditions changes. Different ways for dealing with this problem are treated by Reis (2009), and very recently by Ottavian, Barolo and García-Muñoz (2013a, 2013b). This paper presents and develops a novel Multivariate Image Analysis technique, and deeply analyzes its potential in comparison with a competing approach, in terms of control charts performance and statistical properties; trying to understand the differences between them.

Steel plate simulated images are used to illustrate the monitoring performance of the proposed approach. Section 2 presents the data structure linked to these images, the different image preprocessing techniques required for on-line implementation and the proposed MIA-based SPC monitoring scheme. Section 3 illustrates the statistical comparison study with a sound competing approach (Liu and MacGregor, 2005, 2006) in terms of average run length (ARL), based on simulated images of steel plates. Section 4 presents the results of both methodologies applied on a real data set (clover images). Finally, Section 5 outlines the conclusions of this paper.

2 MATERIALS AND METHODS

2.1 Image data structure

The first type of information when dealing with digital images is formed by the intensities of the $I = n_1 \times n_2$ pixels arranged in the two dimensions that shape the images. In the case of grey level images, each pixel summarizes the global level of spectral information in it. The second type is the spatial information associated to the variations of intensity levels existing in a local zone (a neighborhood) of the image (i.e. texture). Bharati (2002) and Bharati and MacGregor (2000) provide an approach for easily gathering this type of information.

Bharati and MacGregor's approach consists of unfolding each image into just one column vector having the intensity levels of all the pixels, and registering, for each one of them, the intensities of the neighboring ones. This is illustrated in Figure 1, where a 3x3 neighborhood window is used. In this study, it is sensible to this window size, since the unspecific defects analyzed can be of different sizes and shapes, and also random textured. However, in those cases where one is more or less aware about the type of defects (nature, size, shape, orientation, etc.) that may appear in any process, a rational way to approach would be to design an experiment with different window sizes and directions (image analysis kernels) being analyzed, afterwards deciding which treatment to use for optimizing some or various types-of-defects detections.

[INSERT FIG. 1 ABOUT HERE]

This leads to a highly correlated and complex data structure suitable to be analyzed by multivariate statistical projection methods such as principal component analysis (PCA) (Jackson, 2003). This approach can be applied in a straightforward way to color images by simply applying this procedure to each color band, afterwards stacking each data structure related to each color one beside the other (Prats-Montalbán and Ferrer, 2007).

2.2 MIA-based SPC scheme

The MIA-based SPC monitoring scheme proposed in this work, as any SPC scheme, is carried out in two phases. In Phase I (model building) monitoring charts are built according to a set of historical in-control data, once the performance of the process has been modeled, and the assumptions of its behavior are checked.

In Phase II (model exploitation) these charts are used to monitor the process using on-line images, assuming the form of the distribution of the monitoring statistics to be known along with its values of the in-control parameters (Woodall, 2000). This monitoring scheme constitutes what the authors have named in the MIA field as the *fit to a pattern model approach* (FPM) (Prats-Montalbán and Ferrer, 2007, 2011). In this paper we have improved the FPM approach and named as *Percentage-based FPM*, PFPM. For a general overview of the application of PCA to multivariate SPC see *e.g.* Ferrer (2007). The basics of this scheme are explained in the following.

2.2.1 Phase I (model building): off-line process monitoring

Phase I of PFPM starts by building a PCA model (Jackson, 2003). PCA projects the original variables onto new ones, called latent variables, orthogonal and arranged according to their eigenvalues. Applying a PCA model to the textural matrix \mathbf{X} ($I \times J$) can be expressed as:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (1)$$

, where \mathbf{T} ($I \times R$) and \mathbf{P} ($J \times R$) are the scores and loading matrices for R principal components ($R \leq \text{rank}(\mathbf{X})$), respectively; and \mathbf{E} ($I \times J$) is the residual matrix of the PCA model.

The number of R components to extract is a non-trivial issue to deal with, and it depends on the problem at hand (Camacho and Ferrer, 2012, 2014). One possible approach is trying to extract as many components as necessary for appropriate modeling of the main spatial features of the images, *e.g.* by scree plot and further loadings inspection if looking for any deviation from these behaviors; or by looking for those components gathering some specific defects, when a priori known; or to choose that

number of components that maximizes detection capacity (or classification accuracy) when defect images are available in addition to NOC images (Prats-Montalbán et al., 2008). Note that the widely used cross-validation approach is not appropriate to address this part of the modeling stage (Camacho and Ferrer, 2012, 2014), since minimizing the squared prediction error in crossvalidation (objective function in CV) is not necessarily related to maximizing the fault detection power (objective function in process monitoring and fault detection).

The model is built using one or some in-control or process pattern images. These images are collected when the process is operated in-control and assuming independence. Once the loading matrix \mathbf{P} of the in-control or pattern model has been obtained, the scores of a new image \mathbf{X}_{new} are calculated as:

$$\mathbf{T}_{new} = \mathbf{X}_{new} \mathbf{P} \quad (2)$$

From the scores, the residuals can be computed:

$$\mathbf{E}_{new} = \mathbf{X}_{new} - \mathbf{T}_{new} \mathbf{P}^T \quad (3)$$

From the scores and the residuals associated to each pixel, two statistics are derived: T_R^2 and RSS (residual sum of squares) (Kourti and MacGregor, 1996). These two statistics differ in their conceptual meaning.

T_R^2 is the Hotelling- T^2 statistic when a reduced subspace with R components is used, instead the original variables space. Under the assumption that the scores are normally distributed (they are linear

combinations of random variables), control limits at significance level (Type I risk) α_1 can be computed using, e.g. Tracy et al. (1992).

The residuals sum of squares $RSS = \mathbf{e}^T \mathbf{e}$, where \mathbf{e} is the residual vector of a given pixel. Assuming that residuals are normally distributed, several approximations have been proposed for the RSS distribution (Ferrer, 2007). In this paper, approximate critical values at significance level α_l have been calculated following Jackson and Mudholkar (1979).

Values of the T_R^2 -statistic over the control limits indicate that the corresponding pixels present extreme values, even though maintaining the internal correlation structure of the model fitted. On the other hand, RSS -statistic values over the control limits are related to pixels that do not behave in the same way as the ones used to create the model, in the sense that correlation structure of the model is broken.

These critical values represent approximate $(1-\alpha_1)\times 100$ control limits for both statistics. The normality assumptions done are usually quite reasonable because reference models are built from in-control or process pattern images (Nomikos and MacGregor, 1995). This can indeed be verified using normal probability plots and other available tools. Also, when large sets of NOC images are available, as is the case in the first simulation study, one can use reference distributions (if not normal). Anyway, control limits can also be computed from distribution free methods by repeated sampling from the T_R^2 and RSS values of the pixels of the in-control images, as it has been done for further control limits computations presented later.

This way, for each image, it is possible to obtain the percentage of pixels (p) that exceed the established $(1-\alpha_1)\times 100$ control limits for both T_R^2 and RSS statistics. These quantities will be referred

as $p(T_R^2)$ and $p(RSS)$ values, respectively. Therefore, from a set of in-control images, it is possible to build p -based control charts (Montgomery, 2001) for each statistic, yielding the $p(T_R^2)$ and $p(RSS)$ charts. This means that a new Type I risk α_2 should be defined. Only $(1-\alpha_2)\times 100$ upper control limits (threshold percentages) will be defined in both charts to detect special events leading to some abnormal behavior of the process due to changes in normal operating conditions; or the appearance of problems, defects and other special causes of variability in the images taken from the process being monitored. This way, it is possible to minimize, or at least control, the percentage of false positives (*i.e.* NOC images being signaled as defective) of the monitoring process; hence improving the performance of the FPM approach presented in 2007.

2.2.2 *Phase II (model exploitation): on-line process monitoring*

Once the control charts have been built from in-control images, new images from the process can be monitored on-line. This way, when a new image is projected onto the pattern model, the percentage of pixels beyond the control limits for T_R^2 ($p(T^2)$) and RSS ($p(RSS)$), are determined. If any of these percentages exceeds its corresponding threshold percentage (upper control limits for the $p(T_R^2)$ and $p(RSS)$ control charts) the new image is considered different from the in-control or pattern images; whereas if these percentages are lower than their respective thresholds, the new image is classified as *fitting the model* (in-control). This is why this scheme will be named as *percentage-based fit to a pattern model* approach (PFPM).

This approach is useful in the case of classification issues. However, there are scenarios where we are not only interested in classifying the whole image as belonging to one specific class, but also in detecting and isolating those pixels associated with any non-established defect type. This isolation

(defect localization) does not only depend on the percentage of pixels, but also their location and spatial distribution in the image.

This second goal (defect localization) can be fulfilled by using the so-called T^2 (for simplicity instead of T_R^2) and RSS images (Prats-Montalbán, 2005); which have been used in a wide variety of examples (Prats-Montalbán and Ferrer, 2007; Prats-Montalbán et al., 2011; López et al. 2006; Lopez-García et al., 2010; Liang et al., 2009; or Kruse et al., 2014) . These pictures are built by simply locating the original positions of those pixels in the image whose T_R^2 and RSS values are beyond their established control limits. Because only the abnormal pixels are displayed, one can easily detect and isolate those areas in the image linked to possible defects.

The drawback is that, in some cases, when the defect is very small or of very slight color grade, it might happen that, when the $p(T_R^2)$ and $p(RSS)$ of an image would not exceed the threshold percentages established for their respective control charts, the defect would not be displayed (although detected if located in the T^2 and RSS images). However, these cases are by concept related to unimportant defects, so the approach seems sensible.

Summarizing, the sequential steps for the monitoring procedure proposed is:

Phase I:

1. Select one or several in-control (pattern) images and build the textural matrix \mathbf{X} .
2. Fit a PCA model on \mathbf{X} .
3. Calculate the corresponding scores and residuals for all the pixels.
4. Calculate the T_R^2 and RSS statistics for all the pixels.

5. Determine the critical values (upper control limits) for T_R^2 and RSS statistics, at some confidence level $(1-\alpha_1)\times 100$.
6. Project a set of in-control images onto the built PCA model; compute for each of them the percentage of pixels beyond the control limits for T_R^2 and RSS statistics: $p(T_R^2)$ and $p(RSS)$; and build the $p(T_R^2)$ and $p(RSS)$ control charts with $(1-\alpha_2) \times 100$ upper control limits (threshold percentages).

Phase II:

7. Project a new image onto the reference in-control PCA model of step 1.
8. Calculate the scores and residuals, and their corresponding T_R^2 and RSS values for all the pixels of the new image.
9. Compute for the new image the percentage of pixels beyond the control limits for T_R^2 and RSS statistics: $p(T_R^2)$ and $p(RSS)$; and plot these values in their corresponding $p(T_R^2)$ and $p(RSS)$ control charts.
10. When the percentage of pixels for either the T_R^2 or RSS exceeds the corresponding control limits, display the T^2 and RSS images by locating into their original coordinates in the new image the T_R^2 and RSS values of those pixels surpassing the control limits for T_R^2 and RSS statistics, respectively.

This approach holds important practical advantages that make it suitable for on-line implementation. First, it is capable of dealing with any kind of unpredictable defect. These types of defects are inherent to the production processes for most products. Second, only few in-control or pattern sample images are enough to build the in-control model. This is a desirable feature in flexible production environments characterized by producing low amounts of highly diverse products. Finally, the approach does not require high computational cost, which makes it suitable for real-time requirements.

2.3 COMPETING APPROACH

In order to validate the proposed methodology, its statistical performance in terms of average run length (ARL) has been compared to a competing approach based on Liu & MacGregor's (LMcG) (Liu and MacGregor, 2006). This approach has been selected because it is also a MIA-based technique, and the only one to our knowledge that uses control charts for detecting the defective images and image location graphs for identifying any type of defect. In the original paper, this approach consists of two steps. First it applies morphological operations for defect detection that might be present in the images; afterwards performing a discrete wavelet transform (DWT) decomposition (Walczak, 2000) for characterizing the images once the defective ones detected in the previous step have been removed from the process.

Wavelets can be considered as one of the state-of-the-art tools in image texture analysis. They are pairs of filters that transform the images into a new representation by splitting their features at different resolution scales, yielding the so-called approximation and detail images; while maintaining the textural structure of the analyzed images. This way, they constitute a multiscale representation of the image that can be used to analyze the texture present in an image at different scales. When dealing with digital images, the discrete wavelet transform (DWT) is usually applied. In this paper, the Daubechies 4 wavelet (Daubechies, 1988) has been used due to its computational simplicity (Porter and Canagarajah, 1997) and good localization properties (Van der Wouwer, 1998). Nevertheless, it is possible to optimize the selection of the wavelet family and their settings (filter length, decomposition level, etc.), as proposed by Prats-Montalbán, Cocchi and Ferrer (2014).

In this study we have inverted the order of the steps of LMcG approach, because our goal is not to monitor the aesthetic quality property of the final product, but defect detection (monitoring and location of the defect). Hence, first the DWT decomposition is carried out for image feature characterization and the variances (*i.e.* energies) of the different horizontal, vertical and diagonal detail images and of the last scale approximation image are computed, from one up to four scales. Then, a PCA model from the features extracted is built on a set of NOC images, and the T_R^2 and RSS $(1-\alpha)\times 100$ control limits, computed by resampling methods, are worked out. Second, when projecting any new image onto the PCA model, whenever either of its corresponding T_R^2 or RSS statistics exceeds their respective control limits, a morphological operation is applied in order to detect the potential abnormal behavior (*i.e.* defect) in the image.

3 RESULTS

3.1 Comparison study

The types of images to be analyzed have been simulated so we can control to which extent they are damaged, which are the characteristics of the defects (number of defects and injury grade), and where they are located. The images were of size 512×512 , compressed with DWT one level to 256×256 due to computational restrictions. One example of a NOC simulated image and a real steel plate image is presented in Figure 2, so it is possible to appreciate the similarities between them.

[INSERT FIG. 2 ABOUT HERE]

We have simulated two different types of defects, spots and scratches, with different levels in their injury grade. Since one goal is to analyze the ability of the two methodologies to detect any defect

present in an image, by using either the Hotelling T^2 or RSS statistics, a design of experiments (DoE) was developed for each of the two types of defects analyzed.

Regarding the spots defects, a 30 run ($3 \times 5 \times 2$) full factorial design was performed using as factors:

- Level of the defect, obtained as a range from 1 to 3 times the standard deviation in steps of 1 (3 levels)
- Area of the defect, ranging from 10×10 to 50×50 pixels² in steps of 10 (5 levels).
- Approach: Percentage-based Fit to a Pattern Model (PFPM) or Liu&MacGregor's (LMcG) (2 levels).

For the scratches defect, a 720 run ($6 \times 5 \times 2 \times 6 \times 2$) full factorial design was performed:

- Number of scratches (NrSc): from 1 to 26 in steps of 5 (6 levels)
- Length of the scratch: from 10 to 50 pixels in steps of 10 (5 levels)
- Width of the scratch: two or four pixels (2 levels)
- Level of the defect, obtained as a range from 0.5 to 3 times the standard deviation in intervals of .5 (6 levels)
- Approach: Percentage-based Fit to a Pattern Model (PFPM) or Liu&MacGregor's (LMcG) (2 levels)

3.1.1 Phase I: model building

One remarkable difference between the proposed approach and Liu and MacGregor's (LMcG) approach is that whereas in the former one or some images are required for building the PCA model, in the latter some set of NOC images are required (25 images in this study). This is due to the fact that the proposed approach works within the image domain, so the model is built with as many pixels as

the image is formed by. In this work, because these pixels come from a simulated NOC image, they can be assumed to join a considerable set of representative samples for our PCA model. In other cases, depending of the process and images characteristics, using more than one NOC image would be a preferable choice, in order to deal with the NOC variation. From the PCA model, the computation of new $(1-\alpha_2)\times 100$ control limits for the percentage of pixels exceeding the $(1-\alpha_1)\times 100$ control limits (steps 2 to 5) is carried out by projecting a set of in-control images as explained in step 6.

In our case, from one NOC image, the T_R^2 and RSS control limits were computed using Tracy et al.'s and Jackson-Mudholkar formulas, respectively, for a Type I risk $\alpha_1=0.01$. Afterwards, from a set of 10000 new NOC images, the $p(T_R^2)$ and $p(\text{RSS})$ control limits for a Type I risk $\alpha_2=0.0027$, related to an in-control average run length (ARL_0) of 370, were computed by calculating the 99.73% percentile of the sample statistical distributions ($UCL_{p(T_R^2)}^{\alpha_2} = 1.07\%$ and $UCL_{p(\text{RSS})}^{\alpha_2} = 1.31\%$, respectively). Similar results (not shown) were obtained assuming normality.

In order to check whether these values provided reasonable in-control ARL's of about 370, a bootstrap resampling procedure (Efron, 1979; Efron and Tibshirani, 1986) was performed in a similar way as in González and Sánchez (2008). From the 10000 NOC images, we computed 2000 in-control run lengths (RL's), obtaining the in-control ARL for these 2000 values. This procedure was repeated 2000 times in a bootstrap simulation (Seppala *et al.*, 1995), hence obtaining a 2000 sample distribution of the in-control ARL. This procedure was performed for both PFPM and LMcG approaches, for each of the two statistics analyzed.

For the proposed procedure, the 95% uncertainty intervals of the in-control ARL sampling distribution were [368.02-402.14] for $p(T_R^2)$ and [367.73-401.68] for $p(\text{RSS})$. This way, the values used as $p(T_R^2)$

and $p(RSS)$ limits were assumed to be valid. Corresponding results were found for LMCG methodology: 95% uncertainty intervals [354.10-386.84] for T_R^2 , and [355.84-387.65] for RSS.

3.1.2 Phase II: model exploitation

Once all the control limits were tuned to the same in-control $ARL=370$, simulations for both types of defects (spots and scratches) following the DoE already commented were developed. Monte Carlo (Metropolis and Ulam, 1949) distributions of the RL's were obtained in the following way.

For each treatment (defect characteristics), a set of 10000 images was generated. These images were projected with replacement onto the PCA models built in Phase I. For PFPM, this projection was repeated until one image presented a value above the upper control limits for both $p(T_R^2)$ and $p(RSS)$ control charts. This procedure was repeated 2000 times, generating a RL distribution for each DoE treatment from where the average of the log run length (AlogRL) was computed. For LMCG approach, an analogous procedure was followed, using the T_R^2 and RSS values of each projected image as the statistic to be compared to the control limits established. For these types of simulated images, and for the two types of simulated defects (spots and scratches), RSS statistic outperforms T_R^2 , so only the results from the former are presented.

Spots

Table 1 presents the ANOVA results for the RSS statistic. Area and Level factors and their interactions with Method are statistically significant ($p\text{-value}<0.05$). On average, both methodologies provide similar AlogRL for detecting spots in the images. The differences between both methodologies rely on the way the AlogRL's evolve with the size of the area, and the degree of the defect (measured in terms

of standard deviations). Fig. 3 a) and b) displays 95% least significance difference (LSD) intervals for Method *vs* Area, and Method *vs* Level interactions, respectively. These figures show that PFPM is more sensitive to the size of the defect while LMcG is more sensitive to the level of the defect.

These two behaviors are related to the abnormalities the methods are trying to detect. PFPM is focused on detecting a percentage of pixels above the established limits. Thus, it is more sensitive to an increase in the percentage of defective pixels rather than to an increase in the level of the defect in the pixel, which may be or not surpassing the established limits.

[INSERT TABLE 1 ABOUT HERE]

On the other hand, LMcG approach is related to the energy (variance) gathered by each wavelet detail image. These images only detect changes in intensity of the pixel, i.e. the border (perimeter) of the defects. An increment in the area of the defect yields a small increment in the perimeter of the defect. This may explain why AlogRL of LMcG is not seriously affected by an increment in the area of the defect (Fig. 3 a). This way, for this approach it is more important (in average for all the areas) to have a high change in the intensity that would lead to a high detail pixel value, hence a high variance (detail images have zero mean); rather than a larger perimeter with low change, which does not end in such high variance. This means that, in general, LMcG is more influenced by the value of the pixel (related to **standard deviations**) in the border than by the percentage of pixels (related to **area**), since the higher the value is, the more it influences the variance.

[INSERT FIG. 3 ABOUT HERE]

Scratches

ANOVA results for the RSS statistic are shown in Table 2. In this case all factors and all interactions with Method are statistically significant ($p\text{-value} < 0.05$). Figure 4 shows 95% LSD intervals for the interaction of the Method with the other factors.

[INSERT TABLE 2 ABOUT HERE]

For the scratches defect, there is no inner area, since a scratch is just like a border. So basically, both PFPM and LMcG are extracting information from the same pixels. As shown in Figure 4, PFPM approach uniformly outperforms LMcG approach.

Summarizing, for spots defects, the approach selection depends on the size of the areas (PFPM for bigger areas, LMcG for smaller) and their intensity level (PFPM for lower intensities, LMcG for higher); while for scratches defects, it is better to use PFPM approach.

[INSERT FIG. 4 ABOUT HERE]

Finally, the reader must be aware that these results are obtained for the type of images and defects simulated, as commented before. For other type of images and defects, results might be different, requiring new studies.

3.2 Application on some defective images

Once the comparison between the detection stages has been performed for the two methodologies, the next step involves analyzing the visual aspect of the defects location step. Hence, the location

procedure of each methodology is first introduced. Afterwards, three representative examples from the defects analyzed are presented and commented.

Regarding the proposed PFFM approach, it must be remembered that it provides a combined detection-location framework based on a multilevel control charting that, using the same primary features (i.e. percentage of pixels beyond the limits), is able to:

- First, detect whenever a new image has a percentage of pixels beyond the maximum admitted for some $(1-\alpha_2)\times 100$ control limit, as in any typical unilateral univariate statistical control chart.
- Second, display the T^2 and RSS images by locating into their original coordinates in the new image the T_R^2 and RSS values of those pixels surpassing the control limits for T_R^2 and RSS statistics, respectively; hence unveiling the shape of the defect detected, no matter its position and form.

However, LMcG needs for a further additional image analysis technique for locating any defect previously detected by the wavelet feature extraction-based MSPC system. LMcG proposes to use mathematical morphology (Haralick et al., 1987; Serra, 1988) basic operators: dilation and erosion. These are used for finding the location in an image of some feature of interest or defect, by trying to identify in the image the pattern of the defect by a priori known shapes called structuring elements. If a feature or defect disappears after several applications of one of these operations, it is considered as noise and should be ignored. Otherwise, it will remain in the image. Moreover, the morphological operation applied needs for some manually selected thresholding that is not needed in the PFFM approach, because the latter applies well established $(1-\alpha_2)\times 100$ statistical control limits.

Figure 5 (left) shows an example of an image with a defect related to a spot of dimensions 30×30 pixels², consisting of an increment in the intensity of the affected pixels of one standard deviation. In this case the defect has been only detected by the RSS statistic for the percentage of pixels beyond the limits (1.63% vs the 1.31% established control limit) of the PFPM approach; i.e. not by LMcG approach nor by the T_R^2 statistic of the PFPM approach. Fig. 5 (middle) shows the corresponding RSS image, which correctly locates the defect.

[INSERT FIG. 5 ABOUT HERE]

On the other hand, for LMcG approach, using a square-based structuring element similar to the simulated defect that we want to locate, the result is shown in Fig. 5 (right). Note that, not only it has not been detected by LMcG approach, but also the image is noisier than the one obtained by the proposed PFPM approach.

The second example shown in Figure 6 (left) is related to scratches defect. Fig. 6 (middle) shows that the location performed by the RSS image is really good, and slightly better than that achieved by morphological operations (Fig. 6 (right)). Furthermore, in accordance with the simulation study results, it has only been detected by PFPM approach.

[INSERT FIG. 6 ABOUT HERE]

Nevertheless, when the degree of the defect is low, PFPM performs better than LMcG not only in terms of detection, but also in locating and visualizing the defects. This can be observed in the third example shown in Figure 7, where the original image (Fig. 7 (left)) has five very slight scratches to be detected. PFPM approach is able to detect the faulty image and also to correctly locate the scratches, as

shown in RSS image (Fig. 7 (middle)). In this case, nor the detection or the location can be achieved by LMcG approach.

[INSERT FIG. 7 ABOUT HERE]

4 APPLICATION ON A REAL IMAGE DATA SET

This last section deals with the application of both methodologies on a real data set. It consists of 66 clover images (300×300 pixels), 30 related to healthy leaves (i.e. NOC), plus 36 related to ozone affected leaves, kindly provided by the Norwegian University of Life Sciences (Dept. of Mathematic Sciences and Technology). Figure 8 shows an example of a healthy clover leaf (a), and an example of an ozone affected clover leaf (b). In order to deal with the average color differences between both healthy and ozone affected leaves, different prior pre-processings have been applied, such as wavelet based background elimination (by eliminating some deep scale A smoothed image), subtraction of the minimum value from each color band, or no pre-processing; afterwards selecting the one providing the best results for each approach at hand.

[INSERT FIG. 8 ABOUT HERE]

4.1 PFPM approach

From the 30 healthy leaves, 5 were randomly selected in order to build the PCA model for the PFPM approach, following Phase I steps 1 to 6. This way, the corresponding $(1-\alpha_1)$ and $(1-\alpha_2)$ control limits for T_R^2 and RSS were computed. Afterwards, the 36 ozone affected images were projected onto the model, obtaining the images with a percentage of pixels over the $(1-\alpha_1)$ control limit over the

maximum permitted by the $(1-\alpha_2)$ control limit. Figure 9 shows the control charts for T_R^2 and RSS statistic.

[INSERT FIG. 9 ABOUT HERE]

As can be seen, in this case, it was the T_R^2 statistic the one providing the best detection results, in comparison with the ones obtained from the simulated images, which strengthens the idea that image analysis is problem dependent. Figure 10 provides the results for the 20th image projected onto the PCA model. It seems that the detection performed by the T_R^2 statistic is even better than that performed by the expert manually (Mask). In general, the results provided tend to overestimate (30 out of the 36 images projected) the percentages determined manually, although in many cases the difference is not big, and can be due to subjective criteria. In those cases where differences are big (6 out of the 36), possible reasons are: on one hand, the fact that we are dealing with natural images with big variations between them that cannot be completely handled by the models (unless being tricky during the calibration process, which could lead to overfitting and bad validation results on different validation datasets). On the other hand, the images come from not completely flat surfaces, which may produce some illumination problems, hence false detections for the problem at hand. It must be noted that in this case the best pre-processing was subtracting the minimum value at each color band prior to further modeling. However, in all cases, even for that not signaled by the $(1-\alpha_2)$ control limit, the localization of the defects has been successful.

[INSERT FIG. 10 ABOUT HERE]

4.2 LMcG approach

For the Liu & MacGregor approach, since we had only 30 healthy images, in order to have a better distribution of the T^2 and RSS values for the calibration data set, we generated a free distribution of the T_R^2 and RSS values by resampling of 25 images, creating a PCA model and projecting the 30 images whole data set. This procedure was repeated 20 times, therefore generating a 600 values vector for each statistic. Afterwards obtaining the $(1-\alpha_2)$ control limit from these vectors (there is no $(1-\alpha_1)$ control limit in this case), and looking for the number of PC's that maximize the detection capability of the approach (either for the T_R^2 or RSS), the 36 defective images were projected onto the PCA model built on the 30 NOC images.

In this case, the control charts for both T_R^2 and RSS statistics (Fig. 11) show that 30 images are signaled by the RSS statistic, which is the best for the LMcG approach, using the wavelet-based background elimination. Once the images have been signaled, the following step is to apply the morphological operations. Fig. 12 shows the corresponding results for image 34, which has been signaled out by the LMcG approach.

[INSERT FIG. 11 ABOUT HERE]

[INSERT FIG. 12 ABOUT HERE]

This way, also in this case the PFPM provides slightly better results than LMcG approach, which in this case tends to underestimate (10 out of the 30 detected) the number of pixels affected by the defect (ozone).

5 CONCLUSIONS

A monitoring scheme based on Multivariate Image Analysis (MIA), Percentage-based Fit to a Pattern Model (PFPM), has been developed for controlling the adequacy of new images to some in-control images, and locating and isolating any kind of defect. The p -charts and T^2 and RSS images related to the PFPM approach have performed well when applied on steel plate simulated images, providing an efficient integral tool for monitoring processes where visual inspection is a critical task, such as those related to ceramics, metal surfaces, artificial stone plates, etc.

The proposed methodology has been compared in an intensive study with the one proposed by Liu and MacGregor (2006). Results show that the proposed methodology obtains similar results for spots defects, whereas for scratches defects the proposed methodology performs better; for the type of images analyzed.

The benefits of PFPM is not only in terms of the ARL, but also in being able to directly relate the spatially located defects to the detection made by the SPC system. Thus, these located defects can be directly pointed out as being responsible for the out of control signal in the detection first stage. This way, the proposed methodology reaches the two searched goals (detection and location) by using just one technique, instead of having to apply further additional techniques.

The proposed methodology has been illustrated on grey-level images. However, as shown in the application to real leaves images, when other types of features are related to the process and quality attributes (e.g. color, spectral features, and combinations of spectral and textural features), the monitoring scheme can be applied equally well, just using color and texture, multispectral or even hyperspectral wavelengths for building the NOC PCA model.

ACKNOWLEDGEMENTS

The authors want to thank Ole Mathis Kruse and Prof. Cecilia Futsaether, from the Norwegian University of Life Sciences (Dept. of Mathematic Sciences and Technology), for providing the real image data set. This research work was partially supported by the Spanish Ministry of Economy and Competitiveness under the project DPI 2011- 28112-C04-02.

REFERENCES

- Bharati, M. H., and MacGregor J.F. (1998). Multivariate Image Analysis for process monitoring and control. *Ind. Eng. Chem. Res.*, 37, 4715-4724.
- Bharati, M.H, and MacGregor J.F. (2000). Texture analysis of images using Principal Component Analysis. *SPIE/Photonics Conference on Process Imaging for Automatic Control*, Boston, MA. 27-37.
- Bharati, M. H. (2002). *Multivariate Image Analysis and Regression for Industrial Process Monitoring and Product Quality Control, Ph.D. Thesis*, McMaster Univesity, Hamilton (Ontario), Canada.
- Bharati, M.H., MacGregor, J.F., and Tropper W. (2003). Softwood Lumber Grading through On-line Multivariate Image Analysis Techniques. *Ind. Eng. Chem. Res.*, 42, 5345-5353.
- Bharati, M.H., Liu, J.J., and MacGregor, J.F. (2004). Image texture analysis: methods and comparisons. *Chemom. Intell. Lab. Syst.*, 72, 57-71.

Boukouvalas, C., DeNatale, F., DeToni, G., Kittler, J., Marik, R., Mirmehdi, M., and Petrou, M. (1998). Automatic system for surface inspection and sorting of tiles. *J. Mater. Process. Technol.*, 82, 179-188.

Camacho, J. and Ferrer, A. (2012) Cross-validation in PCA models with the element wise k-fold (ekf) algorithm: Theoretical aspects. *J. Chemometr.*, 26 (7), 361-373.

Camacho, J. and Ferrer, A. (2014) Cross-validation in PCA models with the element wise k-fold (ekf) algorithm: Practical aspects. *Chemom. Intell. Lab. Syst.*, 131, 37-50.

Cardin, M., MacGregor, J.F., Miletic, I.P., Bruwer, M.J. (2011) Image-Based Endpoint Carbon Prediction for a Basic Oxygen Furnace. *Iron Steel Technol.*, 8 (11); 79-85.

Daubechies, I. (1988). Orthonormal Bases of Compactly Supported Wavelets. *Commun. Pure Appl. Math.*, 41, 909-996.

Duchesne, C., Liu, J.J. and MacGregor, J.F. (2012). Multivariate image analysis in the process industries: A review. *Chemom. Intell. Lab. Syst.*, 117(1), 116-128.

Efron, B. (1979). Bootstrap Methods: Another look at the jackknife. *Ann. Stat.*, 7, 1-26.

Efron, B., and Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Stat. Sci.*, 1(1), 54-77.

Ferrer, A. (2007). Multivariate Statistical Process Control based on Principal Component Analysis (MSPC-PCA): Some Reflections and a Case Study in an Autobody Assembly Process. *Qual. Eng.*, 19, 311-325.

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

González, I., Sánchez, I. (2008). Principal Alarms in Multivariate Statistical Process Control. *J. Qual. Technol.*, 40(1), 19-30.

Huang, J., and Esbensen, K.H. (2000). Applications of the Angle Measure Technique (AMT) in image analysis Part 1. A new methodology for *in situ* powder characterization. *Chemom. Intell. Lab. Syst.*, 54, 1-19.

Jackson, J.E., and Mudholkar, G.S. (1979). Control Procedures for Residuals Associated With Principal Component Analysis. *Technometrics*, 21(3), 341-349.

Jackson, J.E. (2003). *A User's guide to Principal Components*. Ed. Wiley, New York.

Jones, A.L., and Woodall, W.H. (1998). The performance of Bootstrap Control Charts, *J. Qual. Technol.*, 30(4), 362-375.

Kourti, T., and MacGregor, J.F. (1996). Multivariate SPC Methods for Process and Product Monitoring. *J. Qual. Technol.*, 28(4) 409-428.

Kruse, O.M., Prats-Montalbán, J.M.; Indahl, U.G.; Kvaal, K.; Ferrer, A.; Futsaether, C.M. (2014) Pixel classification methods for identifying and quantifying leaf surface injury from digital images. *Comput. Electron. Agr.* (2014), in revision.

Haralick, R.M., Sternberg, S.R., and Zhuang, X. (1987). Image Analysis Using Mathematical Morphology. *IEEE T. Pattern Anal.*, 9(4), 532-550.

Liang, D., Deng, W., Wang, X. Zhang, Y. (2009) Multivariate Image Analysis in Gaussian Multi-Scale Space for Defect Detection. *J. Bionic Eng.*, 6(3), 298–305

Liu, J.J., Bharati, M.H., Dunn, K.G., and MacGregor, J.F. (2005). Automatic masking in multivariate image analysis using support vector machines. *Chemom. Intell. Lab. Syst.*, 79, 42-54.

Liu, J.J., MacGregor, J.F., Duchesne, C., and Bartolacci, G. (2005). Flotation froth monitoring using multiresolutional multivariate image analysis. *Miner. Eng.*, 18, 65-76.

Liu, J.J., and MacGregor, J.F. (2006). Estimation and monitoring of product aesthetics: application to manufacturing of “engineered stone” countertops. *Mach. Vision Appl.*, 16, 374-383.

Liu, J.J., and MacGregor J.F. (2005). Modeling and Optimization of Product Appearance: Application to Injection-Molded Plastic Panels. *Ind. Eng. Chem. Res.*, 44, 4687-4696.

Lopez-Garcia, F., Andreu-Garcia, G., Blasco, J., Aleixos, N. & Valiente, J. M. (2010). Automatic detection of skin defects in citrus fruits using a multivariate image analysis approach. *Comput. Electron. Agr.*, 71 (2): 189-197.

López, F., Prats, J. M., Ferrer, A. & Valiente, J. M. (2006). Defect detection in random colour textures using the MIA T-2 defect maps. In Campilho, A. & Kamel, M. (eds) *Lect. Notes Comput. Sc.*, vol. 4142 *Image Analysis and Recognition, Pt 2*, pp. 752-763. Berlin: Springer-Verlag Berlin.

López, F., Valiente, J.M., Prats-Montalbán, J.M., and Ferrer, A. (2008). Performance evaluation of soft color texture descriptors for surface grading using experimental design and logistic regression. *Pattern Recogn.*, 41, 1744-1755.

Megahed, F.M.; Woodall, W.H.; and Camelio, J. A. (2011). A Review and Perspective on Control Charting with Image Data. *J. Qual. Technol.*, 43(2), 83-98.

Metropolis, N., and Ulam. S. (1949). The Monte Carlo Method. *J. Am. Stat. Assoc.*, 44(247), 335-341.

Montgomery, D.C. (2001). *Introduction to Statistical Quality Control, 4th Edition*, Ed. Wiley, New York.

Nomikos, P., and MacGregor, J.F. (1995). Multivariate SPC Charts For Monitoring Batch Processes. *Technometrics*, 37(1), 41-59.

Ottavian, M; Barolo, M; García-Muñoz, S. (2013). Maintenance of Machine Vision Systems for Product Quality. Assessment. Part I. Addressing Changes in Lighting Conditions. *Ind. Eng. Chem. Res.*, 52, 12309-12318

Ottavian, M; Barolo, M; García-Muñoz, S. (2014) Maintenance of Machine Vision Systems for Product Quality. Assessment. Part II: addressing camera replacement. *Ind. Eng. Chem. Res.*, 53(4) 1529-1536.

Porter, R., and Canagarajah, N. (1997). Robust rotation-invariant texture classification: wavelet, Gabor filter and GMRF based schemes. *IEE P. Vis. Image Sign.*, 144(3), 180-188.

Prats-Montalbán (2005) Control Estadístico de Procesos mediante Análisis Multivariante de Imágenes. Tesis Doctoral. Universidad Politécnica de Valencia

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

Prats-Montalbán, J.M., López, F., Valiente, J.M., and Ferrer, A (2008). Multivariate statistical projection methods to perform robust feature extraction and classification in surface grading, *J. Electron. Imaging*, 17(3), 031106, 1-10.

Prats-Montalbán, J.M., Ferrer, A., and de Juan, A. (2011). Multivariate Image Analysis: a review with applications. *Chemom. Intell. Lab. Syst.*, 107, 1-23.

Prats-Montalbán, Cocchi and Ferrer (2014). N-way modeling for wavelet filter determination in Multivariate Image Analysis. Submitted to *Chemom. Intell. Lab. Syst.*

Reis, M.; Bauer, A. (2009). Wavelet texture analysis of on-line acquired images for paper formation assessment and monitoring. *Chemom. Intell. Lab. Syst.*, 95, 129–137.

Seppala, T., Moskowitz, H., Plante, R., and Tang, J. (1995). Statistical Process Control via the Subgroup Bootstrap. *J. Qual. Technol.*, 27(2), 139-153.

Serra, J. (1988). *Image analysis and mathematical morphology*. Academic press, London.

Tracy, N.D.; Young, J.C.; and Mason, R.L. (1992). *Multivariate Control Charts for Individual Observations*. *J. Qual. Technol.*, 24, 88-95.

Van der Wouwer G. (1998) *Wavelets for multiscale texture analysis, PhD Thesis*, University of Antwerpen, Antwerpen, Belgium.

Walczak, B. (2000) *Wavelets in Chemistry*. Elsevier Science. Amsterdam.

Woodall, W.H. (2000). Controversies and Contradictions in Statistical Process Control. *J. Qual. Technol.*, 32(4), 341-350

Yu, H., and MacGregor, J.F. (2003a). Multivariate image analysis and regression for prediction of coating and distribution in the production of snack foods. *Chemom. Intell. Lab. Syst.*, 72, 57-71.

Yu, H., and MacGregor, J.F. (2003b). Monitoring Turbulent Nonpremixed Flames in an Industrial Boiler Using Multivariate Image Analysis (MIA). *Proceedings of the Seventeenth International Forum. Process Analytical Chemistry. January 21-24, 2003*; 597-602.

Yu, H., MacGregor, J.F., Haarsma, G., and Bourg, W. (2003). Digital Imaging for On-line Monitoring and Control of Industrial Snack Food Processes. *Ind. Eng. Chem. Res.*, 42, 3036-3044.

