Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 51 (2007) 6395–6403
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Detecting defects with image data

Mónica Benito[a,*], Daniel Peña[b]

[a] *Universidad de Alcalá, Department of Statistics, Spain*
[b] *Universidad Carlos III de Madrid, Department of Statistics, Spain*

Received 5 September 2005; received in revised form 5 February 2007; accepted 5 February 2007
Available online 28 February 2007

Abstract

Quality control using continuous monitoring from images is emerging as an active research area. These applications require adaptive statistical techniques in order to detect and isolate process abnormalities. A novel approach is introduced for monitoring schemes in the setting of image data when the quality is associated with uniform pixel gray-scales. The proposed approach requires the definition of a statistic which takes into account both the spatial dependency and the changes in local variability. An application on paper surface demonstrates how the monitoring scheme performs in practical applications.

Keywords: Bootstrap; Spatial variability; Variance ratio; Uniform gray-scales

1. Introduction

In many industrial applications process monitoring by image analysis is emerging as a useful tool in quality control. For inspection of defects in material surfaces one of the most important properties is texture. Detecting a class of defects in a monochrome image is similar to separating certain objects of interest from the background. Work in this area includes a wide range of Bayesian restoration methods (Geman and Geman, 1984; Bendoudjema and Pieczynski, 2005). These methods produce satisfactory results for high-level problems such as object recognition but they generally fail to capture small objects of interest (defects). Other approaches to texture defect detection are oriented toward applying different filters to the textured images (Heiler and Schnörr, 2005; Mandriota et al., 2004a, b). These methods have in common the choice of particular attributes (features) for the filter response (i.e. the filter applied to the image) which are considered the most important for characterizing texture. For our purpose, the extracted features are critical in identifying the presence of a particular defect, and they may be very sensitive to the size of the defect. Another useful approximation is kriging (Conradsen and Nilsson, 1987), where the prediction error at every pixel is used to detect irregularities in the image. However, this method is not easy to apply in defining a general quality control scheme for the production process, which is our main interest. Other texture analysis methods include the use of texture features derived from the image (Conradsen and Nielsen, 1987; Liu and Jernigan, 1990; Ersboll and Conradsen, 1992; Siew and Hodgson, 1988; Epifanio and Ayala, 2002; Mandriota et al., 2004a, b). However, while these methods are useful in many applications they are less suitable for the detection of local variations associated with small defects.

* Corresponding author. Tel.: +34 91 885 42 75; fax: +34 91 885 42 06.
E-mail address: monica.benito@uah.es (M. Benito).

0167-9473/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2007.02.016

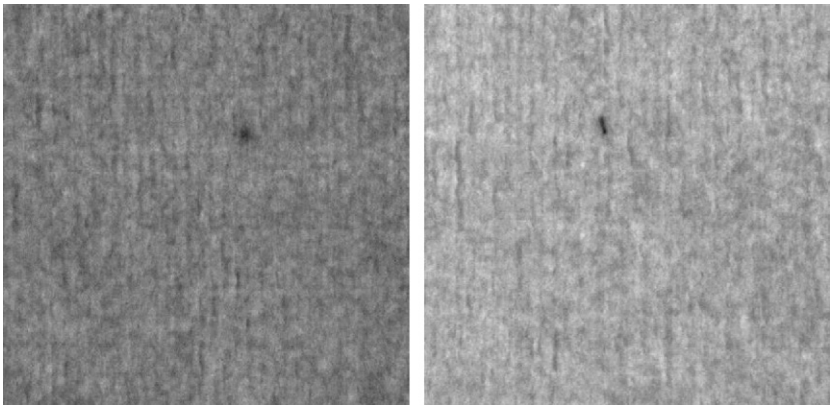

Fig. 1. Textured paper surface with a class of defect.

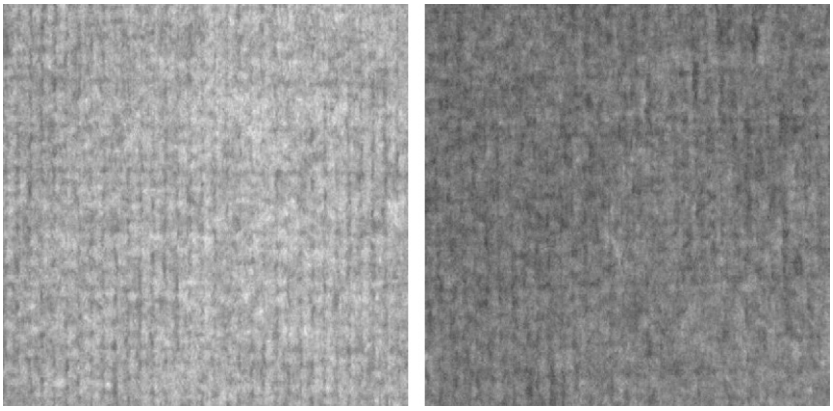

Fig. 2. Textured paper surface.

Consider, for instance, the monitoring of paper surface production. Figs. 1 and 2 show images extracted from a long strip of paper surface during the production process. In Fig. 1 the two images show a defect, whereas in Fig. 2 both images correspond to non-defective surfaces. These figures also show that the intensity levels among images can be different due to external factors in the digitalization process. Under such conditions, histogram gray-level processing techniques (Guo and Pandit, 1998; Sahoo et al., 1988) may fail to provide vital information about clusters in the histogram corresponding to certain objects of interest. Fig. 3 shows the histograms of gray values in Figs. 1 and 2. The four images have a similar unimodal histogram and consequently, traditional control charts based on the mean and standard deviation of the histogram data (Luceño and Box, 1997) are unable to detect defects such as those shown in the two images in Fig. 1.

Traditional texture analysis methods are more concerned with the problem of extracting textural features than with the problem of detecting local variations associated with a class of small defects. In these cases, the contribution of the defect to any statistic computed using all the pixels in the image is expected to be small, and, therefore, statistics based on the whole image are not expected to be powerful. An alternative way that overcomes this problem is dividing the image into blocks and defining some numerical descriptor of the spatial dependencies between the pixels inside each block. This block statistic should be able to discriminate between blocks which contain defects, that will be called defect blocks, and blocks which do not contain defects, or regular blocks. This work proposes a statistic defined as a ratio of the maximum and average variability within each block, which has been found to be useful in practical applications.

The rest of this paper is organized as follows. Section 2 describes briefly the notation for the blocks in the image and presents the proposed statistic for defect detection in textured surfaces. Section 3 illustrates its performance in a sample

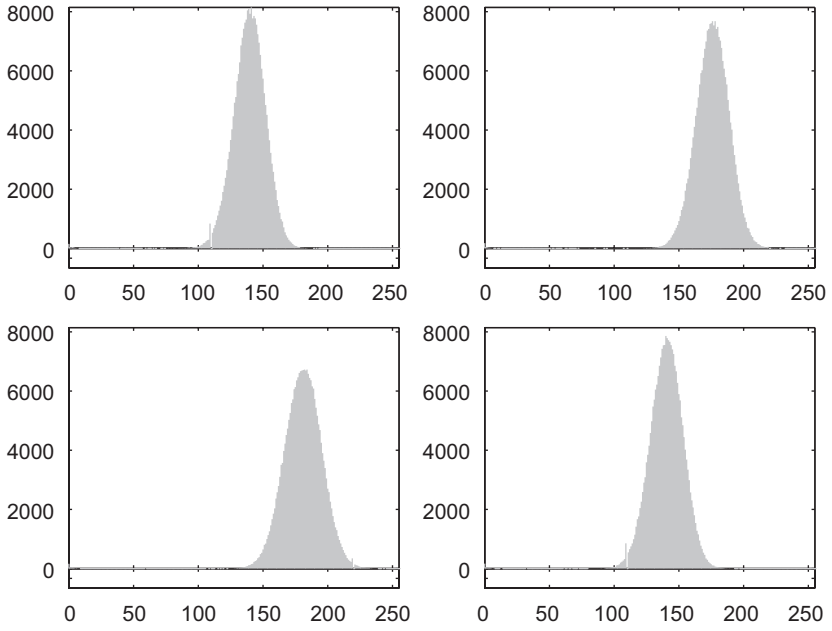

Fig. 3. Histogram for the paper surfaces of Fig. 1 (top panels) and Fig. 2 (bottom panels).

of paper surface images during the production of a long strip of paper surface. Section 4 contains some concluding remarks.

2. Quantifying spatial dependencies

Data collected from a two-dimensional space, such as image data, usually have spatial dependency. For image data, the sample points correspond to pixel values taken from a discrete two-dimensional space regularly distributed (lattice data). When analyzing homogeneous texture, such as paper surface, the local spatial variation of the pixels by rows and columns must be uniform. To illustrate this, Fig. 4 shows a subset of pixels in a row of a regular block and Fig. 5 in a defect block. In Fig. 5 the gray-level of some elements of the row (15, 16 and 17) is not similar to its neighbors. This departure from uniformity implies that the local variation of these subsets of pixels is greater than the ones from Fig. 4. For instance, the standard deviation of the pixels in these rows is 0.07 in Fig. 4 and 0.26 in Fig. 5. The same difference is found when considering the columns of the block. This result suggests that the relative variances by rows and columns could be used as a measure of the local spatial variability in the block.

Suppose that we divide each image X_n, from the sample $X_1, \ldots, X_N$ of digitized images from the production process, into a set of K non-overlapping $m_1 \times m_2$ pixel blocks, where m_1 and m_2 are integers. For the sake of simplicity we consider $m_1 = m_2 = m$. If m is not an exact divisor of the number of rows (columns) in the image, the last row (column) of blocks in X_n is obtained using some of the previous rows (columns) of pixels in order to construct $m \times m$ blocks. The overlap between these last blocks is assumed to be insignificant. The value m is chosen so that $m = cD$, where D is the maximum width (or height) of the defect and $c > 1$. Our experience suggests choosing c in $2 < c < 4$.

2.1. Measuring the local spatial variability

Suppose we compute within a block the variances of the pixel intensities by rows

$$s_{ri}^2(k) = (m-1)^{-1} \sum_{j=1}^{m} (x_{ij}^{(k)} - \overline{x}_{i\cdot}^{(k)})^2, \tag{1}$$

Fig. 4. Row of pixels in a homogeneous texture.

Fig. 5. Row of pixels that shows a defect in columns 15–17.

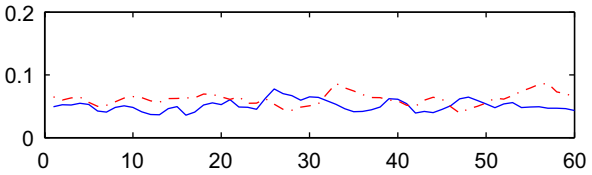

Fig. 6. Function of variances by rows (dash-dot) and variances by columns (solid).

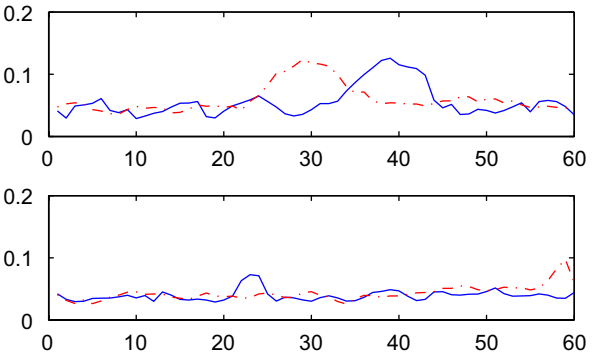

Fig. 7. Function of variances by rows (dash-dot) and variances by columns (solid) in two blocks which contain one defect. Each panel corresponds to a block in two different images.

where $\overline{x}_{i.}^{(k)} = m^{-1}\sum_{j=1}^{m} x_{ij}^{(k)}$ is the mean of the pixels in the ith row in the kth block, $i = 1, \ldots, m$. Let $S_r^2(k) = (s_{r1}^2(k), \ldots, s_{rm}^2(k))$ be the vector of variances of the rows in the kth block. We can consider $S_r^2(k)$ as a function of k and compare these functions for different blocks. In the same way we can compute the variances by columns, $s_{cj}^2(k)$, $j = 1, \ldots, m$, and the function $S_c^2(k)$ of the variance of the columns. A homogeneous texture provides variance functions, $S_r^2(k)$ and $S_c^2(k)$, with values similar to each other, and Fig. 6 shows these functions for a regular block with no defects. However, when the block contains a defect, the variance functions will be expected to be different, as shown in Fig. 7.

Note that when the block has a defect these functions will have a clear maximum, which corresponds roughly to the center of the defect. The relative size of the maximum of the function depends on its average level, which is defined by the variances for rows or columns of the regular block. For instance, in the top panel of Fig. 7 the maximum variance is larger than 0.1 and the average variance is about 0.05. On the other hand, in the bottom panel in this figure, the maximum is smaller, but the average variance is also smaller. Thus we propose to measure the size of the maximum with respect to the average level of variability. In order to have a robust measure of this average level we use the median of the variances. This leads to the following summary statistic of the heterogeneity in the variability for rows:

$$\frac{\max(S_r^2(k))}{\mathrm{med}(S_r^2(k))} = \frac{v_r^2(k)}{\overline{s}_r^2(k)},$$

where $v_r^2(k)$ is the maximum and $\bar{s}_r^2(k)$ the median of $(s_{r1}^2(k), \ldots, s_{rm}^2(k))$. A similar statistic, $v_c^2(k)/\bar{s}_c^2(k)$, can be obtained for columns.

When a defect appears in several rows it will increase the variability of all of them. Thus we may have a more powerful test procedure if we take this property into account. Suppose that in a given block the maximum variability for rows occurs at row i^*. If the block has a defect we expect that this maximum corresponds to a row which includes the defect. In addition, we expect that the neighboring rows, $i^* \pm h$, will also have a relatively large variability. On the other hand, if the block is regular this maximum will happen by chance, and will not give information about the relative variability of neighbor rows. We can take this possible dependency into account by considering not only the variability at the maximum, but also the relative variability in the neighboring rows or columns.

Define the distance weight of the $(i^* + h)$th row with respect to the maximum value in the i^*th row in the kth block as $\widetilde{w}_h^{(k)} = 1/h$ if $h \neq 0$ and $\widetilde{w}_0^{(k)} = 2$ if $h = 0$, where the subscript h indicates the distance between the rows. If we standardize these values so that they add up to one across the rows we obtain the weights,

$$w_h^{(k)} = \frac{\widetilde{w}_h^{(k)}}{2(1 + \sum_{h=1}^{H} \widetilde{w}_h^{(k)})}, \tag{2}$$

where $H = [m/2]$. Then we can compute the weighted relative variability of the rows in the kth block by

$$t_r^{(k)} = \frac{\sum_{h=1}^{H} w_h^{(k)} s_{ri^*\pm h}^2(k)}{\bar{s}_r^2(k)}. \tag{3}$$

We can carry out the same analysis for the columns and obtain $t_c^{(k)}$. The value

$$t^{(k)} = t_r^{(k)} + t_c^{(k)} \tag{4}$$

describes the heterogeneity in the block. The heterogeneity in the image can be measured by

$$t = \max_{k=1,\ldots,K} (t_r^{(k)} + t_c^{(k)}). \tag{5}$$

This statistic computes the maximum of a weighted average of relative variances and, as we show in the next section, it can be useful in identifying heterogeneity in many applications. Note that when a defect image is found, because the statistic (5) is larger than a given threshold, the block with maximum variability will be expected to include the defect, or the largest part of it, and the coordinates (i^*, j^*) of the rows and columns of maximum variability in this block will provide an estimate of the position of the center of the defect.

In some applications, such as the one we will analyze next, we know in advance that we are searching for dark groups of pixels and so a white spot, caused by illumination effects, may be detected wrongly as a defect. This can be avoided by checking that the pixel which defines the rows and columns of maximum variability corresponds to a dark spot, instead of a white spot.

2.2. Sampling distribution of the statistic to monitor

We will estimate the sampling distribution of t using the bootstrap. The procedure is non-parametric, and its validity does not rely on special assumptions such as normality. The only assumption we need is that the data are independently and randomly chosen from some population with unknown distribution.

Let $Y_1, Y_2, \ldots, Y_M$ be a sample of M homogeneous texture images. By splitting each image into K blocks we obtain a sample of $L = MK$ independent regular blocks. Given a parameter that is defined as a function $\theta = T(X)$ of the values in the population of homogeneous textured blocks (regular blocks) and a statistic that is the same function of the observations, $t = T(Y)$, the bootstrap estimates the sampling distribution $F_t(X)$ of the statistic of interest by resampling (Efron and Tibshirani, 1993). The resampling is done many times, B. Each time b, $b = 1, \ldots, B$, a random sample of K blocks is chosen with replacement from the set of L blocks. The statistic (5) is evaluated at each bootstrap sample, giving a set of bootstrap values t_b^*, $b = 1, \ldots, B$. A bootstrap sample is indicated by $*$. The empirical distribution of these bootstrap values $\widehat{F}_t(X)$ estimates the theoretical sampling distribution $F_t(X)$. Bootstrap confidence intervals can be computed by using the $\alpha/2$ and $1 - \alpha/2$ quantiles of $\widehat{F}_t(X)$ as a $(1 - \alpha)$ level confidence interval for the parameter.

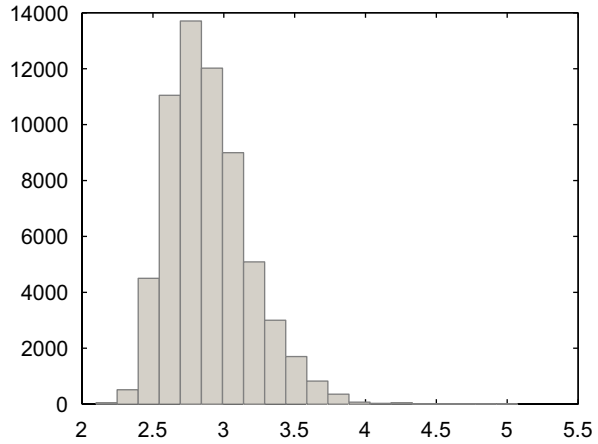

Fig. 8. Bootstrap distribution of θ.

As we are interested in one-sided confidence intervals, we use the $(1-\alpha)$ quantile as the critical value for the estimator. One of the reviewers had suggested that a Chi-squared distribution might provide a good approximation to the null distribution of the proposed statistic. This was investigated and found not to be the case. For this study, $M = 50$ images are arbitrarily selected from the homogeneous texture paper surface. Using $B = 1000$, the statistic (5) is estimated for each bootstrap sample, giving the set of values t_b^*, for $b = 1, \ldots, B$. The histogram of these values is plotted in Fig. 8.

3. Experimental results

In this section, we present the experimental results from a number of textured paper surfaces for evaluating the validity of the proposed algorithm. The sample to monitor is composed of a set of $N = 92$ digitized images acquired during the production process of a long paper surface. The images are 492×496 pixels and are all displayed as gray-level images. It is known that the maximum surface of the defect is 20×20 pixels. We recommend that m be at least three times the maximum width (or height) of the defect, and thus blocks of size 60×60 ($m = 60$) have been chosen. A 99% confidence interval for θ is estimated with $B = 1000$ bootstrap samples, and the endpoint is used as the threshold θ^c of product acceptability. In this experiment, we obtain the value $\theta^c = 3.7$. We have chosen a significance level of $\alpha = 0.01$ to avoid detecting wrongly defects in homogeneous textured images.

The quality control scheme is implemented as follows. Let $t^{(1)}, t^{(2)}, \ldots, t^{(K)}$, be the values of the statistic (4) by blocks in an image X. Assume that these statistics are ordered, so that $t^{(1)} \geqslant t^{(2)} \geqslant \cdots \geqslant t^{(K)}$. Then if $t^{(1)}$ is smaller than the critical value $\theta^c = 3.7$, we conclude that there are no defects in the image. Otherwise we assume that the image has at least a defect. As in a given image we may have more than one defect, we test for an additional defect by comparing the next largest value $t^{(2)}$ with θ^c. If a new defect is detected in another block, then continue this process until some $t^{(k)}$ is lower than the estimated critical value, and we conclude that there are no more defects in X. It may happen that the image has only one defect but it is located between two consecutive blocks. Then we expect that the two highest values of the statistic, $t^{(1)}$ and $t^{(2)}$, should be greater than θ^c. Moreover, the corresponding blocks must be located together.

In order to check the efficiency of the algorithm we compute the size and the power of the test. For this task, two different sets of images were used. The first one, composed of 50 non-defective images, was used to compute the empirical size. The second one, composed of 92 defective images, was used to compute the power. We report in Table 1 the results obtained in the experiment when varying the size of m. Note that best results were obtained when $m = 60$.

As an illustration we show in Figs. 9–11 the defect detection procedure in four defective images choosing $m = 60$. In Fig. 9 the defect can be seen clearly. In the middle panel the values of the statistic $t^{(k)}$ are plotted for the 64 blocks, and the circle point corresponds to the maximum value, $t^{(1)}$, which occurs in the 29th block. The horizontal line is the threshold $\theta^c = 3.7$. The bottom panel shows the function of variances for the defective block, $S_r^2(29)$ and $S_c^2(29)$. From this graph we can observe that the spatial location of the spot is around the 30th row and 40th column. Fig. 10

Table 1
Results of the quality control monitoring when varying the size of the window m

N	50			
	$m = 40$	$m = 60$	$m = 80$	$m = 100$
Homogeneous surface				
Images detected as defectives	0	0	0	0
Percentage of defectives (%)	0	0	0	0
N	92			
	$m = 40$	$m = 60$	$m = 80$	$m = 100$
Surface with a class of defect				
Images detected as defectives	57	89	78	72
Percentage of defectives (%)	61.9	96.7	84.8	78.3

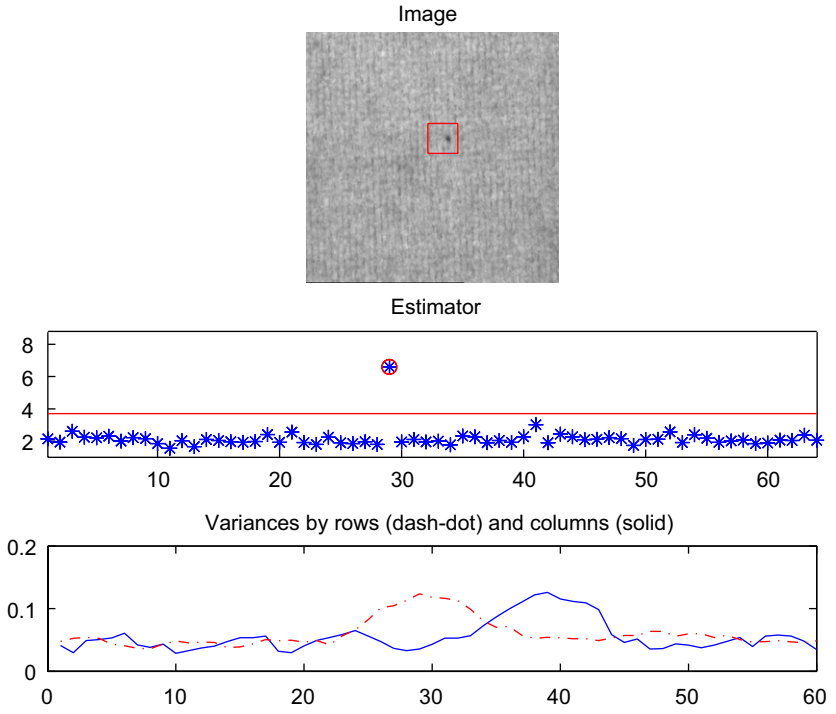

Fig. 9. Quality control monitoring for an image with a defect.

shows an image where the defect is difficult to detect at first sight although the algorithm is able to do it successfully. In Fig. 11 we show an example where the defect appears between two blocks. The two panels in the top of Fig. 11 show the blocks 36 and 44 as defect blocks. In the middle panel the values of the statistic $t^{(k)}$ are plotted and the circle points correspond to these defect blocks. The 36th block has the defect at the end of the last row, hence the last components of $S_r^2(36)$ should be high, as well as those of $S_c^2(36)$. This is observed in the lower left panel of Fig. 11. Block 44 has the defect in the first row, so that the first elements of $S_r^2(44)$ have high variances, as well as those at the end of $S_c^2(44)$ (see the right side of the figure). The images wrongly detected as non-defective have in common the small size of the defect.

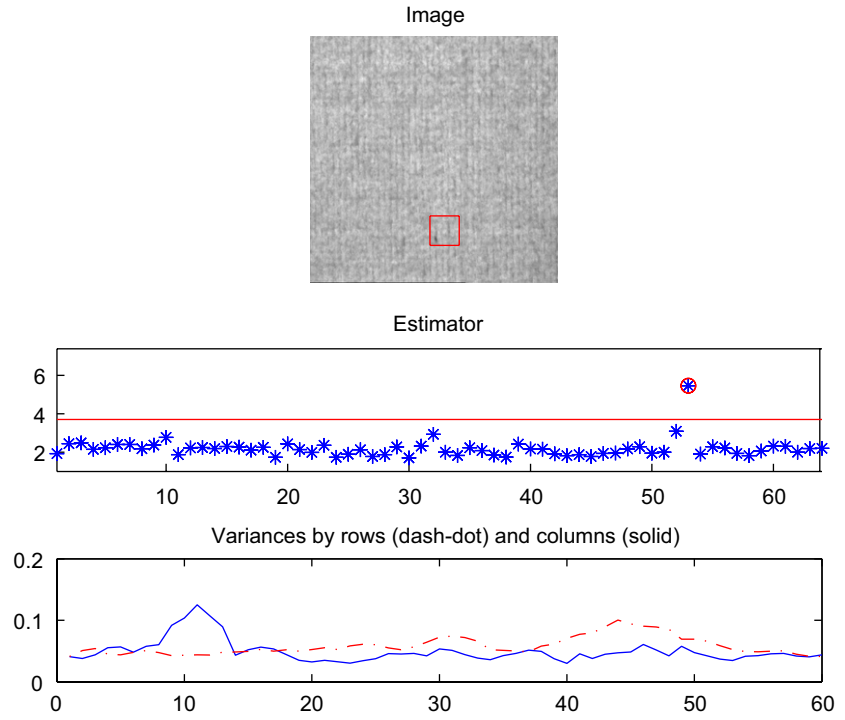

Fig. 10. Quality control monitoring for an image with a defect.

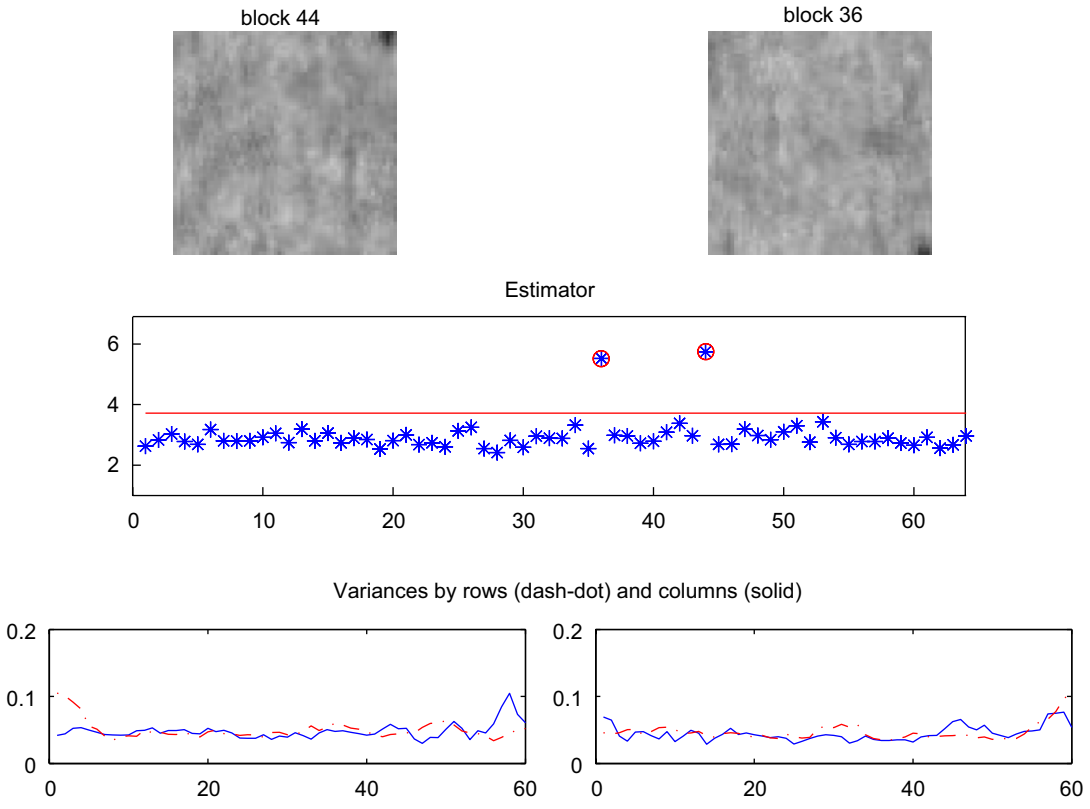

Fig. 11. Quality control monitoring for an image when the defect is between two blocks.

4. Conclusions

In this paper we have introduced a procedure based on a new statistic for defect detection in textured surfaces. This procedure can be used to develop a quality control system for automatically locating and identifying defects on textured surfaces. The main idea of the proposed procedure is to divide the image into a set of non-overlapping blocks, compute the relative variability in each block and use as a measure of the heterogeneity in the image the maximum relative variability in the blocks. The relative variability is measured by comparing the maximum variability for rows or columns to the average variability, estimated in a robust way. Also we make our statistic more powerful by making a local smoothing to take into account spatial dependence. The sampling distribution of the statistic is obtained by bootstrap and we have explored the use of parametric approximation, such as the Chi-squared, that, while not appropriate for our data, can be useful in other applications. Based on experimental results, the proposed method performs reliably for homogeneous textures such as paper surfaces. Further investigation may extend this technique to the inspection of defects in any homogeneous texture, such as leather, wood, and other industrial products.

Acknowledgments

This research has been supported by Spanish MEC Grant SEJ2004-03303.

References

Bendoudjema, D., Pieczynski, W., 2005. Unsupervised image segmentation using triplet Markov fields. Comput. Vision Image Understanding 99 (3), 476–498.

Conradsen, K., Nielsen, B.K., 1987. Textural features in classification of digital images. In: Proceedings of the Second International Tampere Conference in Statistics, pp. 143–159.

Conradsen, K., Nilsson, G., 1987. Data dependent filters for edge enhancement of Landsat images. Comput. Vision, Graphics Image Process. 38, 101–121.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York.

Epifanio, I., Ayala, G., 2002. A random set view of texture classification. IEEE Trans. Image Processing 11 (8), 859–867.

Ersboll, B.K., Conradsen, K., 1992. A strategy for grading natural materials using a two step classification procedure. SPIE 1821, 318–329.

Geman, G., Geman, G., 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell. 6, 721–741.

Guo, R., Pandit, S.M., 1998. Automatic threshold selection based on histogram modes and a discriminant criterion. Mach. Vision Appl. 10, 331–338.

Heiler, M., Schnörr, C., 2005. Natural image statistics for natural image segmentation. Internat. J. Comput. Vision 63 (1), 5–19.

Liu, S.S., Jernigan, M., 1990. Texture analysis and discrimination in additive noise. Comput. Vision Graphics Image Understanding 49, 52–67.

Luceño, A., Box, G., 1997. Statistical Control by Monitoring and Feedback Adjustment. Wiley, New York.

Mandriota, C., Niti, C., Ancona, N., Stella, E., Distante, R.M., 2004a. Texture measures for carpet wear assessment. IEEE Trans. Pattern Anal. Mach. Intell. 10, 92–105.

Mandriota, C., Niti, C., Ancona, N., Stella, E., Distante, R.M., 2004b. Filter-based feature selection for rail defect detection. Mach. Vision Appl. 15, 179–185.

Sahoo, P., Soltani, S., Wong, A., Cheng, Y., 1988. A survey of thresholding techniques. Comput. Vision Graphics Image Process. 41 (2), 223–260.

Siew, L.H., Hodgson, R.M., 1988. Texture measures for carpet wear assessment. IEEE Trans. Pattern Anal. Mach. Intell. 10, 92–105.