



# Wavelet texture analysis of on-line acquired images for paper formation assessment and monitoring

Marco S. Reis<sup>a,\*</sup>, Armin Bauer<sup>b</sup>

<sup>a</sup> Department of Chemical Engineering, University of Coimbra, Rua Sílvio Lima, 3030-790, Coimbra, Portugal

<sup>b</sup> Voith Paper Automation GmbH & Co KG, Linzer Strasse 55, 3100, St. Pölten, Austria

## ARTICLE INFO

### Article history:

Received 16 June 2008

Received in revised form 29 September 2008

Accepted 30 September 2008

Available online 11 October 2008

### Keywords:

Multivariate image analysis

Wavelet texture analysis

Paper formation

Wavelets

Multivariate statistical process control

Principal components analysis

## ABSTRACT

Paper formation (the distribution and intermixing of fibres in a paper sheet), plays a central role in paper products, and is usually evaluated off-line, with a significant delay relative to the high production rates achieved in modern paper machines. In this paper, we address an approach for evaluating and monitor paper formation using images acquired with an especially designed sensor, in-line, in-situ and in real time. The methodology essentially consists of applying wavelet texture analysis to raw images, in order to compute a wavelet signature for each image, based on which their discrimination, according to the formation quality level, can be made. A PCA analysis of such features confirms the different formation quality levels defined a priori after visual inspection, and, furthermore, suggests a new subclass for abnormal samples, related to the bulkiness of fibre flocks. A multivariate statistical process control framework, based on such PCA description (PCA-MSPC), is proposed to monitor formation quality, which provides quite good results when applied to the available images, as analyzed with the ROC curve for the method and confirmed with a Monte Carlo simulation study using subimages with 1/4 of the size of the original ones.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

Paper surface properties play a key role in the definition of the overall quality level of fine paper products, affecting features such as general appearance (optical properties, flatness, etc.), printability (e.g. the absorption of ink) and friction (related to the ability of feeding paper to printers without causing jams, as well as to the wearing of printer components). A central aspect of paper surface quality is the level of uniformity in the way fibres are distributed across its surface, known as paper formation. A paper with good formation has a surface free of fibre flocks or of any high fibre density regions, whereas one having bad formation does present a strong and noticeable heterogeneity in the surface distribution of fibres. Papers with a poor formation can cause many problems to the final user. For instance, the average strength of a sheet with poor formation is less than that for a comparable sheet with a better formation, the quality of printing operations is also inferior, and the coating consumption increases as paper formation gets worse. For these reasons, it constitutes a matter of great concern to paper producers, being routinely measured in the quality control labs in modern paper production facilities, either through visual inspection and comparison with a series of standards representing different quality levels of formation or, more frequently, using instrumentation that scan the paper sheet and measure the

variation in light transmission through it. However, such an inspection activity is carried too late in the process and any abnormality is only detected after large amounts of paper have already been produced (a modern paper machine can produce paper at a linear speed of approximately 100 km/h, meaning that, in a single day of nominal production, a 9 m wide paper web is produced that covers the distance from Moscow to Rome). Thus, it is both opportune and of foremost importance to move such an inspection activity to the production facilities, where paper is actually being made, turning it into a fast and effective method for performing on-line process monitoring of paper formation. This is not however an easy task to accomplish and relevant technological challenges must be properly dealt with first, before such a system can be implemented. One aspect, for instance, is the way quality images are extracted over a fast moving web in the wire section of a paper machine, in a high-humidity environment, where drops erratically fly around, potentially interfering with the camera objective. In this paper, we analyze a set of real images acquired with an experimental apparatus developed at Voith Paper Automation that successfully circumvent these difficulties, and evaluate the potential for conducting image-based process monitoring as well as product evaluation and classification activities, in-line, in situ and in real time, with such technology.

Image processing is a mature field of knowledge, with many powerful techniques developed for dealing with issues such as image enhancement (so that image features become better resolved and clear, by adjusting contrast and brightness, enhancing object edges,

\* Corresponding author. Tel.: +351 239 798700.

E-mail address: [marco@eq.uc.pt](mailto:marco@eq.uc.pt) (M.S. Reis).

smoothing noise, etc.) and segmentation (partitioning the image in meaningful compartments, usually called regions or segments, through pixel, region, edge or model-based segmentation algorithms), which quite often are preliminary stages for further qualitative or quantitative analysis [1–3]. Another wave of applications that has emerged during the last 15 years or so, but gaining significant momentum in the last 5 years, is linked to the application of multivariate statistical analysis techniques, such as principal component analysis (PCA), along with other multiway, regression (partial least squares, PLS) and classification (PLS-DA) extensions, as well as other data processing techniques (e.g., transforms such as wavelets and Fourier), to image analysis problems, exploring the correlation structures existing between different wavelength channels (multivariate or hyperspectral images) while adequately handling pixel-to-pixel, and also sometimes, depending on the approach followed, image-to-image variability. A multivariate image is a set of congruent images stacked together, each one being relative to a particular wavelength band (in the most common case of hyperspectral images) and providing the numerical intensity values (normally integers from 0 to 255) for such channel (congruent images are images that, after proper pre-processing, each pixel maintains the same meaning and provides the same type of information, regarding the elements to be analyzed). Such multivariate images have typically three modes, two for the position coordinates ( $x, y$ ) and one for the channel wavelength ( $\lambda$ ). These are adequately handled by unfolding the 3-way raw image into a 2-way array preserving the wavelength dimension [4] or directly applying multiway approaches (such as PARAFAC). Multivariate Image Analysis (MIA) approaches have already been applied to different types of problems, involving congruent/non-congruent images and following pixel-based/image-based approaches. Some examples of successful applications found in literature include the assessment of the surface quality of steel [5] and artificial stones [6], monitoring flames in industrial boilers [7] and softwood lumber grading [8]. As further examples of image-based tasks, now involving predictive modelling (regression analysis), one can refer the estimation of: coating content in snack food production [9], fruit yield from airborne images [10], concentration of ingredients in cheese [11] and of pulp properties [12]. Applications involving congruent images are less common, but can also be found in well controlled experimental settings [13].

The analysis of paper formation belongs to the general class of “texture analysis” problems, for which Wavelet Texture Analysis (WTA) has been considered a state of the art approach in the extraction of relevant information from images [5,14]. WTA enables the computation of a variety of features from the 2-D wavelet transform of each image that, after proper selection, are usually strongly connected to the type of texture phenomenon under analysis, providing a wavelet features signature of it. Other approaches based upon wavelets were also applied in the context of MIA, exploring the ability of such transform to accumulate both spatial and frequency information in each of its transform coefficients, and to bring some necessary congruency to the analysis in the image domain [13]. In what concerns to the analysis of paper formation, Bouydaï et al. [15] applied a 2-D wavelet transform to light transmission images of paper sheets, focusing on the details of the second decomposition, whose overall energy (sum of the energies obtained for each detail image) was taken as a good indicator of the quality level of paper formation. Even though the analysis only used a limited number of decompositions, which might hinder its discrimination ability for different abnormal formation patterns, the authors found it more sensitive than the standard formation assessment technique. Scharcanski [16] developed an approach based on stochastic modelling of the wavelet coefficients at different resolutions, and, based on such probabilistic descriptions, proposed a multiresolution distance measure for stochastic textures, showing ability to discriminate between different levels of quality of paper formation. However, this procedure do rely

somewhat on the adequacy of the hypothesised distributions, and the use of equal weights for all scales in the definition of the distance measure may not be an adequate strategy, as shown in our analysis further ahead in this paper. To sum up, the procedural simplicity and underlying clarity of the WTA-based methodology, its computation efficiency and effectiveness, are strong reasons for its selection, especially in situations such as the one we are considering in this paper, where the relevant image features are likely to be present in any region of the image space (random texture).

It is perhaps worth noticing that the approach pursued in this paper do present strong analogies to that followed before by one of the authors when addressing the issue of simultaneously monitor roughness (a fine length-scale phenomena, regarding very small irregularities in the surface of paper) and waviness (larger scale deviations from the flat shape), in an integrated way, using accurate surface profiles obtained through a mechanical stylus profilometer [17]. Such 1-D profiles were analyzed using a wavelet-based technique, in order to identify and separate the scales corresponding to paper roughness and waviness, for which features were computed that allow for their proper monitoring. The fact that the present approach for addressing a rather different paper surface quality problem also calls for a multiscale framework, is a clear manifestation of the multiscale nature of paper structure, mentioned in [17], and, in a certain sense, the current approach does represent an extension of the 1-D methodology followed in such reference, to the analysis of signals with higher number of modes (images).

This paper is organized as follows. In Section 2, the essential nomenclature and concepts underlying wavelet multiresolution decomposition and WTA are briefly addressed, as well as the basic steps of the methodology to be followed. Then, in Section 3, we describe the measuring sensor and the image set to be analyzed. The results of their analysis are present in the next section, along with those for the image-based PCA-MSPC approach for paper formation, which is evaluated both using the original images, as well as through a Monte Carlo study of their subimages. The essential conclusions of this work are referred in the final section of this paper.

## 2. Wavelet and wavelet texture analysis (WTA) basics

### 2.1. Wavelet transforms

The wavelet transform of a signal,  $x$ , consists of an alternative representation of it, given by a coarser resolution version plus all the details lost across the several decomposition stages performed. This so called multiresolution decomposition [18], is conducted by successively applying quadrature mirror filters to the coarser resolution version obtained after each decomposition stage, from which another (higher level) coarser resolution results plus a detail signal, i.e., a signal that represents what is lost when passing from one resolution level to next coarser one. After a few decomposition stages, say  $J_{dec}$ , a single coarser resolution approximation vector is obtained, say  $\mathbf{a}_{J_{dec}}$ , along with all the details lost through the several decomposition stages,  $\mathbf{d}_1, \dots, \mathbf{d}_{J_{dec}}$ . They convey the same overall information as the original signal  $x$ , but organized in a different way. In particular, each coefficient accumulates information from a certain region (box) in the time-frequency space, a property that is very useful for a number of applications, in particular for WTA.

One can also compute the wavelet transform of a 2-D signal (a image), using a 2-D wavelet transform. This can be accomplished through the alternating application of the 1-D filtering operations to the rows and columns of the matrix of pixel intensities, implicitly using a separable 2-D wavelet basis set (tensor products of the 1-D basis functions), or through the application of non-separable 2-D wavelet functions [19–21]. In this paper, we follow the first, more conventional and widely used approach. Applying such a 2-D wavelet transform to an image, say  $\mathbf{X}$ , results, after  $J_{dec}$  decomposition stages,

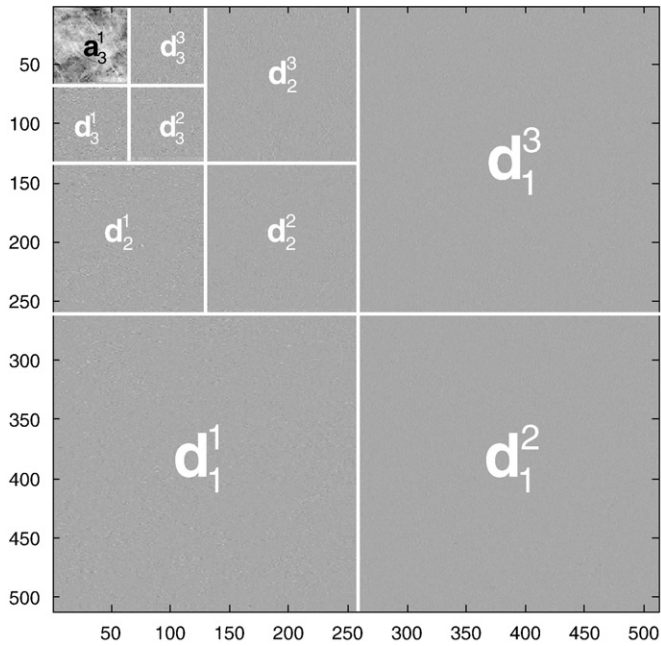


Fig. 1. 2-D wavelet transform of an image, with the different detail images and the resulting coarser resolution approximation image, after three decomposition stages ( $J_{dec}=3$ ). The transform has the same number of coefficients as the size of the original image.

in a coarser version of it,  $a_{j_{dec}}^1$ , as happened in the 1-D case, plus all the detail images that are lost across the different decomposition stages. However, in the 2-D situation, there are three types of details that are obtained at each scale, regarding the different combinations of decomposition operations successively applied to the rows and columns of the previous coarser approximation image, corresponding to the three types of basis functions used. Such detail images contain information regarding different spatial orientations: horizontal, vertical and diagonal [22]. For instance, for scale  $j$ , we can find the detail images:  $d_j^1$  (details with sensitivity North-South, or vertical),  $d_j^2$  (details with diagonal sensitivities) and  $d_j^3$  (details with sensitivity East-West, or horizontal). Fig. 1 shows an example of the wavelet transform of an image, after three decomposition stages ( $J_{dec}=3$ ). The final number of 2-D wavelet coefficients in all detail and approximation images is equal to the number of pixels in the original image.

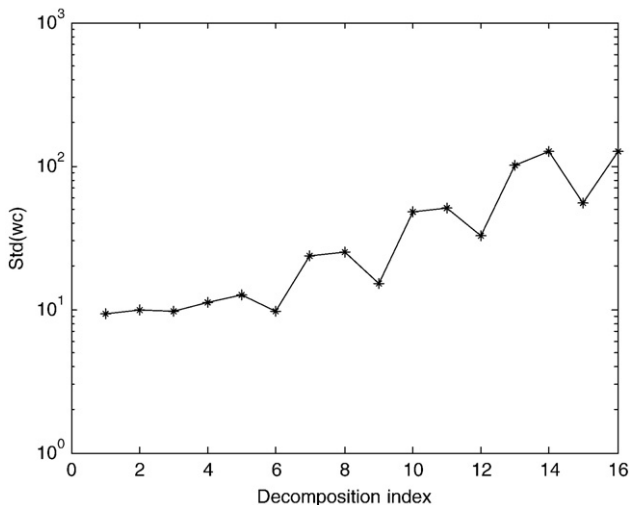


Fig. 2. A typical features profile (wavelet signature) for a paper sample image, computed using the standard deviation metric.

2.2. Wavelet texture analysis (WTA)

The wavelet coefficients of the transformed image carry local information about the energy found in certain frequencies ranges and with some orientation (indexed to the scale and type of details). Therefore, by computing a metric for each scale and orientation, which is a function of all the coefficients at that scale/orientation combination, one is indeed summarizing all the relevant information for a particular range of frequencies (along a given orientation). WTA consists on computing features in this way, and analyse their discriminating power regarding the textural phenomena under analysis. Some examples of WTA features used in practice are the energy ( $E_{jk} = \|d_j^k\|_F^2$ , where  $\|\cdot\|_F$  stands for the Frobenius norm of a matrix), entropy, averaged  $l_1$ -norm or the standard deviation of the set of coefficients. In this work we adopted the last way of computing features, the standard deviation of coefficients:  $s_{jk} = \text{std}(d_j^k)$ , where std represents the sample standard deviation of the wavelet coefficients contained in the argument (as before,  $j=1, \dots, J_{dec}$  and  $k=\{1,2,3\}$ ). Fig. 2 presents one example of a features vector computed for a given image collected from the paper production process, using  $J_{dec}=5$ . In this case one has 16 features per image:  $J_{dec} \times 3 + 1 = 16$ , as we have 3 types of detail images per decomposition level, plus the final coarser approximation (which is sometimes omitted from this type of analysis, in order to minimize the effects of non-uniform illumination of the sample area). WTA uses such parsimonious features profiles (also called its wavelet signatures) summarizing the fundamental information contained in the raw images, to analyze and classify the distribution of samples with different texture patterns or characteristics.

2.3. Methodology for paper formation analysis and evaluation

The essential steps underlying the proposed methodology for the analysis of paper formation based upon WTA features are outlined in Fig. 3. After acquisition, the image may be pre-processed, for instance

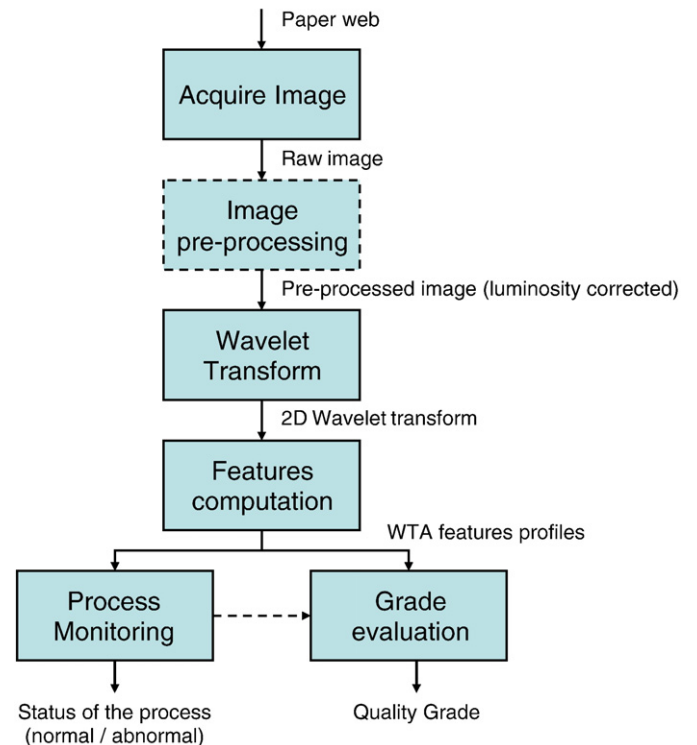


Fig. 3. WTA-based methodology for the evaluation of the formation quality of paper and/or for its monitoring.

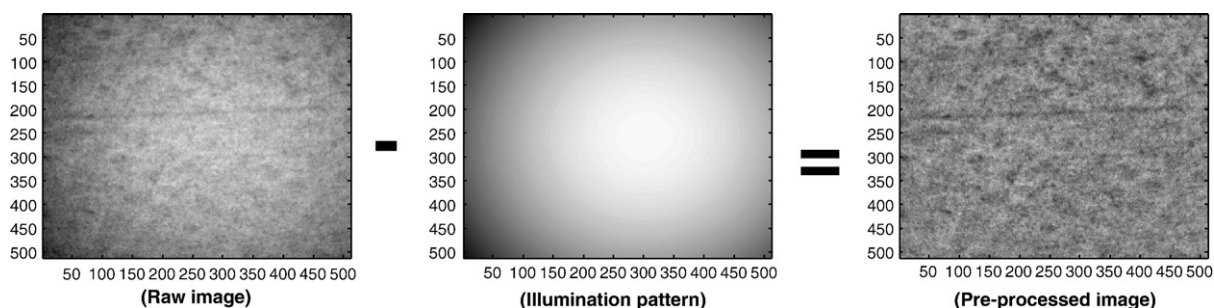


Fig. 4. Image pre-processing by removal of the non-uniform illumination effect.

by correcting non-uniform illumination effects that may disturb the subsequent analysis (Fig. 4).

This can be easily achieved by subtracting the estimated illumination pattern from the image. However, WTA is often found to be rather robust to this type of effects, and this is the reason why we keep this stage optional in our procedure. One reason to explain such a robustness arises from the fact that non-uniform illumination is a low frequency feature of the image, which therefore does not interfere with coefficients relative to higher frequency bands, but only with those for coarser scales (low frequency bands), in particular with the ones belonging to the coarsest approximation signal. Therefore, by removing the last approximation image from the analysis, one is implicitly removing or minimizing the effect of non-uniform illumination. Afterwards, the 2-D wavelet transform is computed as well as the WTA features that will be used for grade evaluation and/or process monitoring. Process monitoring can also make use of a grade classifier, namely for establishing the type of abnormality present when an alarm is triggered in the monitoring control charts. These issues will be described more thoroughly in the results section.

### 3. Materials and methods

#### 3.1. Formation sensor

The formation sensor consists of a digital camera within a housing with a rotating head at high speed, in order to prevent dirt accumulation on the housing surface. This design protects the camera from the harsh environmental conditions (steamy and wet) prevailing in the forming section of a paper machine, ensuring that the sensor will function properly with little maintenance (Fig. 5).

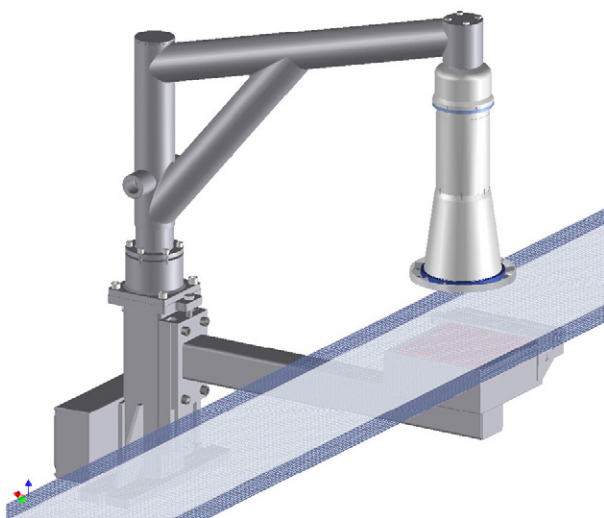


Fig. 5. Scheme illustrating the experimental image acquisition apparatus for on-line formation monitoring.

With this setup in use, a transmission image of the wet web on the forming wire is obtained. The formation sensor is usually employed to monitor and optimize the forming section of a paper machine. It is especially useful for multi-ply paper machines where the formation of an individual ply cannot be determined with standard formation sensors that are placed close to the reel. The images acquired for this article were taken from pilot plant test trials, where blade settings as well as the flow around the blades were changed. The camera used was a Jai A10 CL, with a Navitar DO-2595 lens. The strobe light consists of a led array emitting red light (CCS LDL-TP).

#### 3.2. Description of the image set

A set of 24 images, representing different formation patterns, were collected for analysis. Before any data processing was made, these raw images were visually assessed in order to verify whether the subsequent analysis does provide meaningful results, i.e. to enable the confirmation that samples clustering together in a given features subspace, do represent similar textural patterns as they are visually perceived by the final user of paper. The preliminary evaluation scores obtained by visual inspection are presented in Table 1, with image samples representative from the main quality grades (three) presented in Fig. 6. From this table, we can also notice that a dubious sample was assigned the class label “1.5” (borderline between classes with scores 1 and 2; in fact it does seem to present more similarity with samples from class 1, but shows more texture variation than a typical grade “1” sample) and an outlier was identified, labelled with score “9” (Fig. 7).

### 4. Results

In this section we present the main results obtained from the analysis of the sample images from paper with different quality

Table 1

Classification scores for the images analyzed. Score “1” represents paper with good formation, score “2” paper with poor formation, and score “3” stands for paper with bad formation

Sample	Score	Sample	Score
1	1	13	1
2	1	14	1
3	1	15	1
4	2	16	2
5	2	17	2
6	1.5	18	2
7	1	19	3
8	1	20	3
9	1	21	3
10	2	22	2
11	9	23	2
12	2	24	2

A score “1.5” was given to a sample which is borderline between classes “1” and “2”, from a visual inspection, while “9” represents a clear outlier (perhaps a problem during image acquisition).

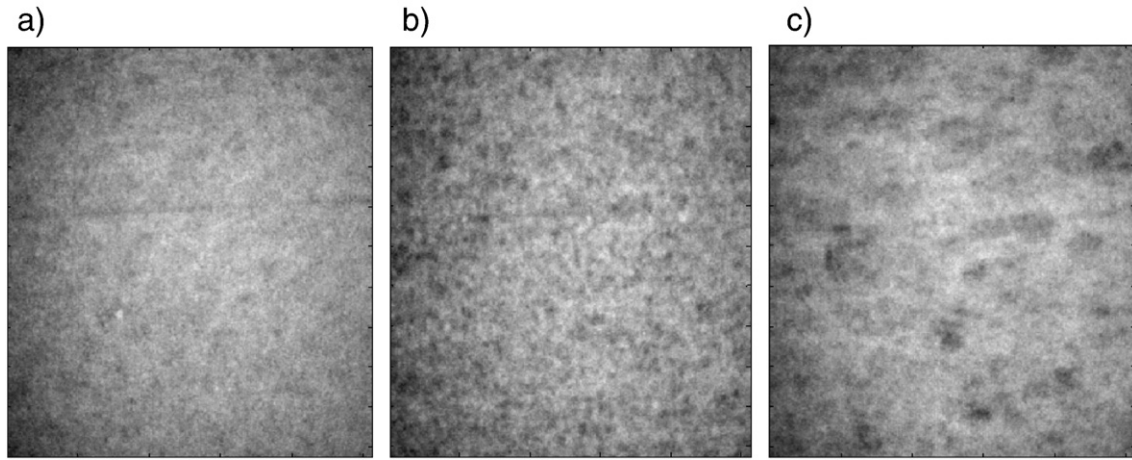


Fig. 6. Images representative from the three main quality classes: a) class with scores 1, b) class with scores 2, c) class with scores 3.

grades. We begin by addressing the structure of the WTA features computed from such images, and their separation power with respect to the preliminary classification task. Then we focus on how such a procedure can easily give rise to a process monitoring scheme, with the potential for simultaneously classify abnormal samples. Computations were conducted in the Matlab environment (The MathWorks, Inc.), using in-house developed code and the package WaveLab 8.02 (available at <http://www-stat.stanford.edu/~wavelab/>) for performing the wavelet transformation computations.

#### 4.1. Analysis of the image set: WTA-based analysis and classification of paper formation

We first computed the WTA features from all the 24 images, from which the profile presented in Fig. 2 is a particular case. The resulting wavelet signatures are presented in Fig. 8. Such features were computed after image pre-processing for illumination correction, but the effect of such correction is only marginal, and do not alter in any meaningful way, the results and conclusions presented here (this was checked by direct comparison of the results with and without illumination correction).

From this figure we can see that apparently not all scales present the same discriminating power regarding the different quality grades. In fact features 1 to 6, regarding the two first decomposition stages,

are quite insensitive to the different formation grades present in the image set. On the other hand, features for decomposition stages higher than 2 (index 7 and above) do seem to cluster the samples in different groups. An explanation of this behaviour can be found in the multiscale structure of paper surface: finer scales are connected with surface phenomena such as wire marks, whereas coarser scales are linked with formation and other coarser scale structural features. Decomposing the  $(24 \times 16)$  wavelet features matrix (24 image samples with 16 wavelet features computed for each one of them), using principal component analysis (PCA) with mean correction [23,24],

$$X_{WTA} = T \cdot P^T + E \quad (1)$$

where  $X_{WTA}$  is the  $(n \times m)$  wavelet features matrix,  $T$  the  $(n \times a)$  matrix of PCA scores,  $P$  the  $(m \times a)$  matrix of PCA loadings and  $E$  the  $(n \times m)$  residual matrix, we can indeed see that the loadings for the first two principal components do not weight significantly the first 6 features (Fig. 9). On the other hand, coefficients for the coarser scales tend to present high magnitudes and preferentially opposite signs. These results are consistent with what is expected from prior knowledge about paper formation, which is essentially a coarser scale phenomenon, and the fact that such behaviour is being properly captured by the first two PCA components does support their “visual significance”. Furthermore, only the first two PCA components were analyzed,

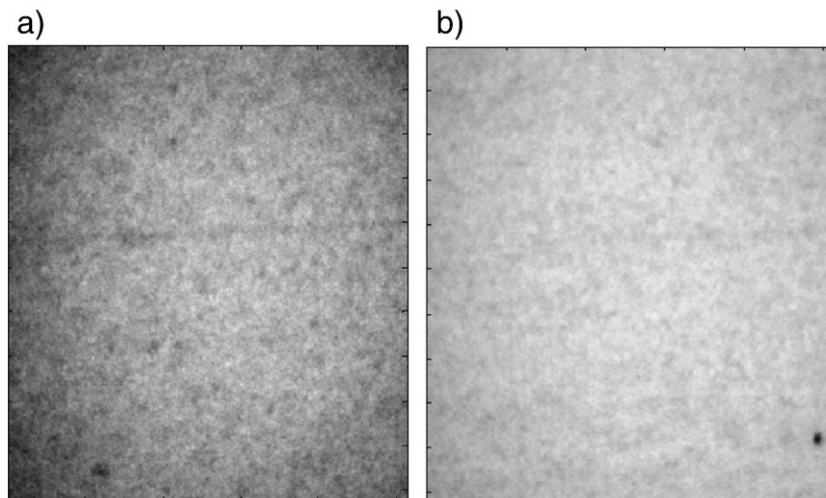


Fig. 7. Sample images with score 1.5 (a) and with score 9 (b).

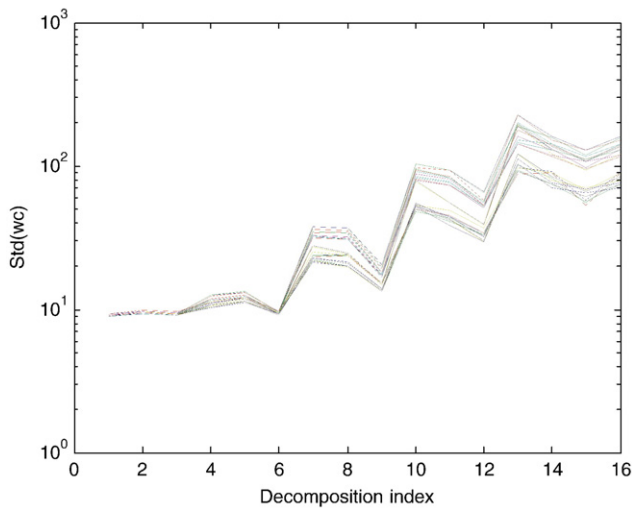


Fig. 8. Wavelet features profiles for the 24 image samples.

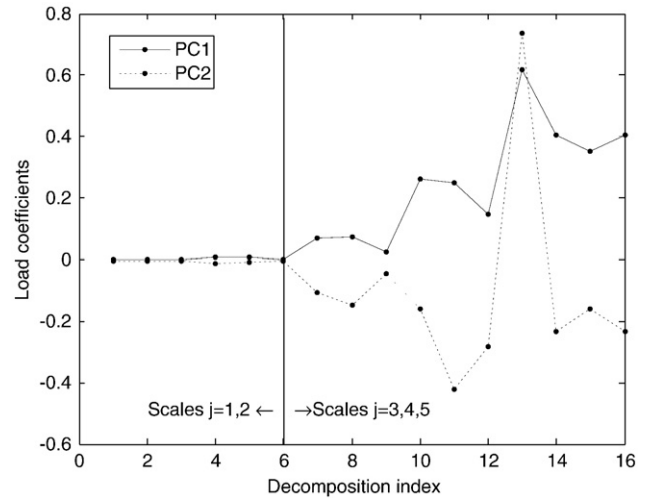


Fig. 9. PCA loadings of the wavelet features profile matrix for the first 2 principal components.

because they concentrate most of the overall variation in  $X_{WTA}$  (97.96%), and correspond to the minimum number of components that must be retained in order to obtain a proper separation of the different formation grades (as we confirmed through cross-validation with a linear classifier).

Analysing the scores of the PCA decomposition ( $T_1$  and  $T_2$ ), it is possible to see that the samples with different formation quality grades do appear in different, easily distinguishable, clusters (Fig. 10). Samples with good formation appear on the left side of the plot, well separated from those with poor formation, appearing on the right side. Therefore, the first principal component seems to be doing a good separation job between good and poor formation samples. The second component, on the other hand, is discriminating the several types of abnormalities, separating those labelled a priori with score “2” from those labelled with score “3”.

Furthermore, from the analysis of Fig. 10, it is apparent that what was previously considered similar samples with score “2”, could in fact represent somewhat different texture patterns. Analysing samples from the cluster “2A” and “2B” (Fig. 11) one can check that they do present in fact some distinguishing features. Namely, samples from cluster “2B” tend to show more “cloudy” patterns, while the size of irregularities in those from cluster “2A” are smaller and appearing more frequently in the image. Therefore, the unsupervised analysis of WTA features has not only provided for a good separation of samples according to the quality grades defined a priori (but not used in the analysis), validating their pertinence as discriminating descriptors, but also enable for the discovery of new classes, allowing for a better resolved definition of different quality grades.

Furthermore, we can now perhaps understand better our classification uncertainty regarding sample labelled “1.5”: this sample, in spite of clearly belonging to cluster “1” (to which it is indeed in general more similar), is almost as far from the farthest sample in this cluster, than from the closest sample from cluster “2”, which may indicate significant differences from some other representatives from the score “1” cluster.

#### 4.2. An image-based process monitoring scheme of paper formation based on WTA features profiles

Following the good separation results obtained between images samples with different formation quality grades, using the WTA features, and provided that the image acquisition system is to be operated on-line, in-situ, we found it both opportune and natural to address the issue of implementing an on-line process monitoring scheme for paper forma-

tion. Given the highly correlated nature of the data matrix of wavelet features structure,  $X_{WTA}$ , a suitable monitoring scheme, with well established results, is the so called multivariate statistical process monitoring (MSPC) based on PCA (PCA-MSPC) [25–28]. This methodology consists of modelling the systems operation under normal conditions using a PCA model, in order to capture the normal correlation structure and variability relative to the so called “common” causes. In this framework, for each new multivariate observation ( $x_i$ ), two statistics are computed, one relative to the (Mahalanobis) distance between its projection in the PCA subspace and the centroid of the normal operation conditions (NOC) cluster, also in the PCA subspace, usually referred as  $T_i^2$ , and another relative to its (squared) Euclidean distance to such a subspace,  $Q_i$ :

$$T_i^2 = t_i \Lambda^{-1} t_i^T \quad (2)$$

$$Q_i = e_i e_i^T \quad (3)$$

where  $t_i$  stands for the scores of the  $i$ th observation ( $i$ th row of the scores matrix,  $T$ ), whose NOC covariance matrix is given by  $\Lambda$ ;  $e_i$  is the  $i$ th residual vector ( $i$ th row of the residual matrix,  $E$ ). In simple terms, the  $T^2$  statistic monitors the variability in the PCA subspace, whereas

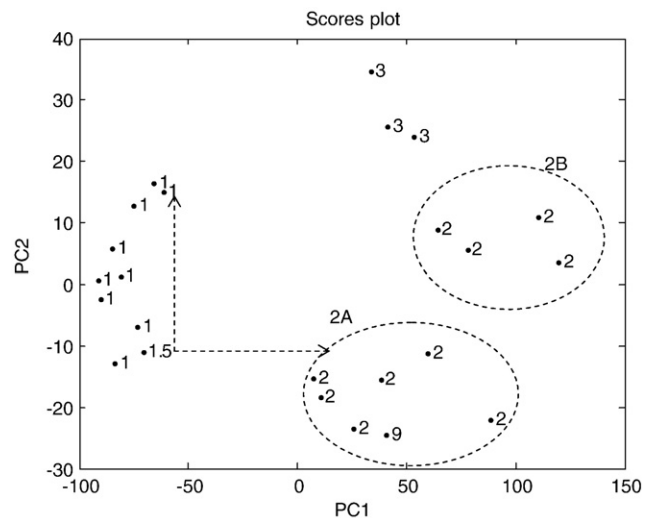


Fig. 10. PCA scores of the wavelet features profile matrix for the first 2 principal components.

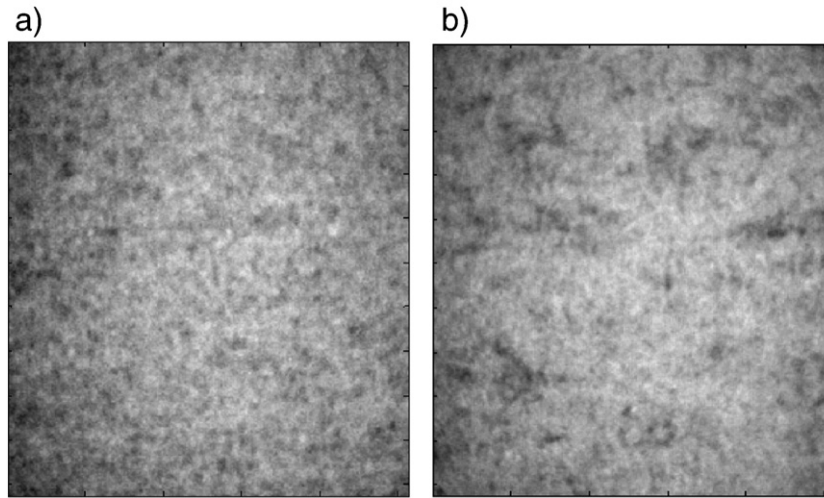


Fig. 11. Images representative from clusters 2A (a) and 2B (b).

$Q$  addresses variability around this subspace. Control limits for these statistics have been derived for the multivariate normal case. The upper critical value for the  $T^2$  statistic is:

$$T_{lim}^2(a, n, \alpha) = \frac{a(n-1)}{(n-a)} F(a, n-a, \alpha) \quad (4)$$

where  $F(v_1, v_2, \alpha)$  is the upper  $\alpha \times 100\%$  percentage point for the  $F$  distribution with  $v_1$  and  $v_2$  degrees of freedom ( $\alpha$  is the significance level selected for conducting the monitoring task),  $n$  is the number of observations in the NOC data set and  $a$  is the number of principal components. As for the  $Q$  statistic, the upper control limit is given by:

$$Q_{lim}(a, \alpha) = \theta_1 \left( \frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right)^{\frac{1}{h_0}} \quad (5)$$

where

$$\theta_i = \sum_{j=a+1}^m \lambda_j^i, \quad i = 1 : 3 \quad (6)$$

$$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2} \quad (7)$$

$\lambda_j^i$  stands for the  $j$ th eigenvalue (following a decreasing magnitude ordering, obtained from the PCA analysis of NOC data), raised to the  $i$ th power, and  $c_\alpha$  is the upper  $\alpha \times 100\%$  percentage point for the standard normal distribution ( $m$  is the number of WTA features). Individual scores can also be monitored using univariate SPC charts, but in this work we base the monitoring scheme around the  $T^2$  and  $Q$  statistics. Applying the PCA-MSPC methodology to the present case of monitoring paper formation, encompasses the definition of a NOC set for estimating the PCA subspace and monitoring statistics. We assumed that, under normal operation, only samples with grade “1” formation quality are being produced. Therefore a small subset of four grade “1” samples was selected for the purposes of estimating the procedure parameters and statistics,  $I_{NOC} = \{1-3, 13\}$ . The remaining samples were used to test the procedure:  $I_{TEST, NORMAL} = \{7-9, 14, 15\}$ ,  $I_{TEST, ABNORMAL} = \{4, 5, 10, 12, 16-24\}$ . As the training and test set are rather small, we focused firstly our analysis on the ability of PCA-MSPC based on WTA features to separate the normal samples in the test set from the abnormal ones. This is essentially a two-class classification problem, where the classifier is built using the statistical process control methodology. For such a type of classification problems, the receiver-operating characteristic curves (ROC curves) provide a suitable tool for analyzing the classifier performance, namely its

sensitivity (probability of detecting the abnormality, when it happens) and specificity (probability of not detecting the abnormality, when it is absent). These concepts are connected to the so called False Positive Rate (FPR) and True Positive Rate (TPR), where the TPR is the fraction of observations violating the monitoring detection limits, relatively to all abnormal samples available in the test set, and FPR is the fraction of normal samples violating the detection limits (the sensitivity is equal to the TPR, while specificity is given by  $1 - \text{FPR}$ ). The ROC curve is a plot of FPR versus TPR, and the higher the values obtained for TPR at a given level of FPR, the better is the classifier or monitoring scheme (this is also some times quantified through the area below the ROC curves: the higher the area the better the classifier performance). These curves are computed by varying the classifier thresholds ( $Th_i$ ) that define the regions corresponding to the two classes, and for each threshold value, a FPR and TPR pair is calculated, which will then be plotted along all the others:  $\text{FPR}(Th_i)$  vs.  $\text{TPR}(Th_i)$ ,  $Th_i \in [Th_{min}, Th_{max}]$ . In the present case of statistical process monitoring, this corresponds to vary the significance level of the control charts. The ROC curve obtained is presented in Fig. 12, where we can see that all abnormal samples were always correctly detected at all levels of false alarms contemplated (FPR), which is the “ideal curve” for a classifier. This result really confirms the separation power of the PCA-MSPC procedure based on the WTA features of the acquired images, now regarding abnormal vs. normal samples.

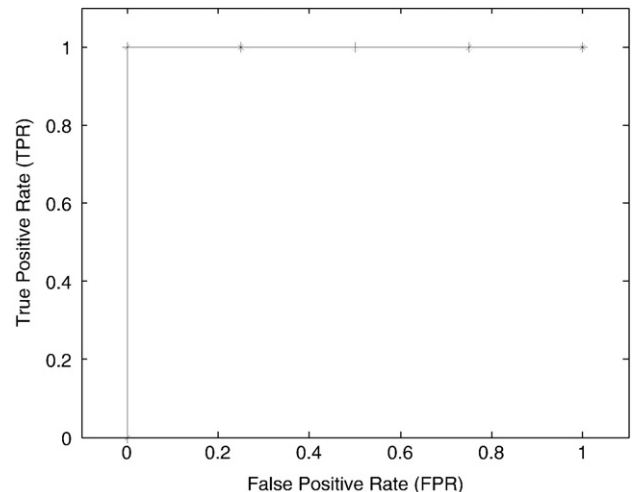
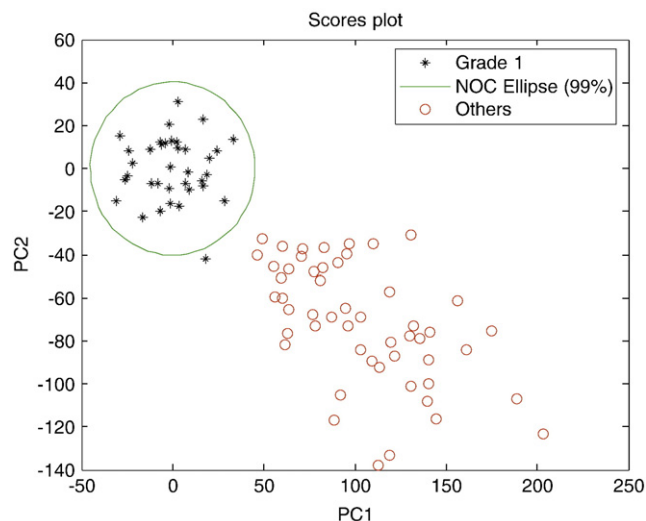


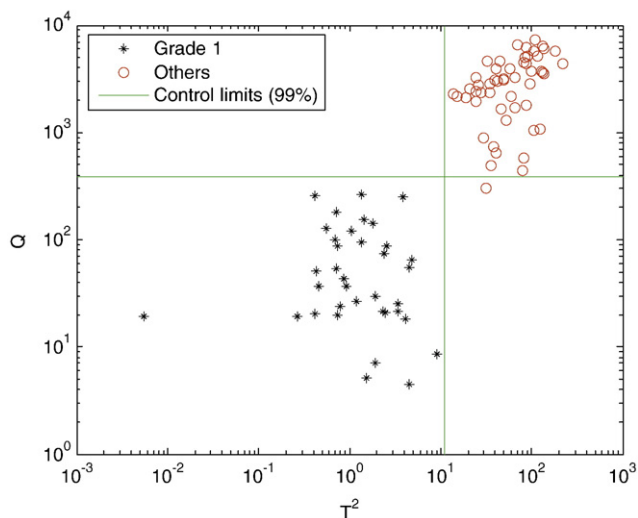
Fig. 12. ROC curve for the PCA-MSPC monitoring of paper formation using WTA features.



**Fig. 13.** PCA model estimated from all subimages with grade “1” (\*). Samples presenting formation problems were also projected in this subspace (O). The 99% ellipse for the grade “1” samples is also shown, representing normal operation conditions (NOC) region.

As the number of images used in this analysis was quite low, we refined this study by using “more” images, obtained by dividing the original ones in 4 parts (i.e., each original image give rise to 4 sub-images of equal size, 1/4 the size of the original image, corresponding to its four quadrants). Using all such subimages from grade “1” samples to define the NOC conditions, resulted in the PCA scores plot presented in Fig. 13 and in the plot summarizing the values obtained for two PCA-MSPC statistics ( $T^2$  and  $Q$ ) presented in Fig. 14 (note the use of log–log scales in this plot, as the separation ability obtained with this methodology was quite high).

However, the use of all grade “1” samples for constructing the NOC region is a rather optimistic situation, necessarily leading to an improved separation performance between normal and abnormal samples. Therefore, to circumvent this limitation in the analysis, we have randomly sampled half of the grade “1” subimages to be used as a reference set for estimating the PCA model and the PCA-MSPC statistics, and applied it to assess the status of the remaining subimages not contained in the test set (the remaining 50% of normal subimages, and all the abnormal ones), from which the FPR and TPR



**Fig. 14.** Combined plot for the  $T^2$  and  $Q$  statistics, obtained after considering all grade “1” subimages as representing normal operation conditions.

**Table 2**

Results for the Monte Carlo study regarding the evaluation of normal and abnormal subimages

	Mean	Standard deviation
FPR	0.056	0.075
TPR	1.00	0.00

were computed. By repeating this procedure 1000 times, the results presented in Table 2 were obtained. These results show that abnormal samples were always correctly detected, whereas a small fraction of the normal ones were misclassified (note the relatively high standard deviation obtained when compared to the mean value for the FPR, which can be attributed to the low number of samples used in each Monte Carlo trial to establish the normal operation conditions necessary to estimate the model and statistics, leading to an increased rate of false alarms).

## 5. Conclusions

In this paper, an approach for evaluating and monitor paper formation was presented, which uses real images acquired with an especially designed sensor, located in the end of the wire section of a paper machine. The methodology essentially consists of applying wavelet texture analysis to the raw images, in order to compute a features profile or wavelet signature for each image, based on which their discrimination can be made. We found it to be quite robust to illumination non-uniformity effects, as the presence/absence of illumination correction pre-processing did not led to any assignable difference in the analysis. The length scales more relevant for discriminating the different levels of quality regarding paper formation, were also analyzed using plots of the WTA profiles and of the loadings vectors obtained after PCA modelling. The clusters obtained in the scores plot of such PCA description do confirm the preliminary visual assessment made of the samples, and, furthermore, suggest the existence of two subclasses of grade “2” abnormal samples, differing in the bulkiness of fibre flocks. The good discrimination results obtained prompted us to develop an approach for implementing statistical process control of paper formation, given the importance of such characteristic for the final quality of paper products. In this regard, a PCA-MSPC approach was applied to WTA features computed for each image. Such an approach led to quite good results when applied to the available images, as analyzed with the method’s ROC curve, and confirmed with a Monte Carlo study using subimages with 1/4 of the size of the original ones. These results attest the potential of the proposed methodology for integrating an image-based statistical process control of paper formation, in-line, in-situ and in real time.

## References

- [1] F. van der Heijden, Image Based Measurement Systems - Object Recognition and Parameter Estimation, Wiley, Chichester, 1994.
- [2] B. Jähne, Digital Image Processing, Springer-Verlag, Berlin, 1993.
- [3] J.C. De Anda, X.Z. Wang, K.J. Roberts, Chem. Eng. Sci. 60 (2005) 1053–1065.
- [4] P. Geladi, H. Grahn, Multivariate Image Analysis, Wiley, Chichester, 1996.
- [5] M.H. Bharati, J.J. Liu, J.F. MacGregor, Chemom. Intell. Lab. Syst., Lab. Inf. Manag. 72 (2004) 57–71.
- [6] J.M. Prats-Montalbán, A. Ferrer, J. Chemom. 21 (2007) 10–23.
- [7] H. Yu, J.F. MacGregor, AIChE J. 50 (7) (2004) 1474–1483.
- [8] M.H. Bharati, J.F. MacGregor, W. Tropper, Ind. Eng. Chem. Res. 42 (2003) 5345–5353.
- [9] H. Yu, J.F. MacGregor, G. Haarsma, W. Bourg, Ind. Eng. Chem. Res. 42 (2003) 3036–3044.
- [10] X. Ye, K. Sakai, A. Sasao, S.-i. Asada, Chemom. Intell. Lab. Syst., Lab. Inf. Manag. 90 (2008) 132–144.
- [11] J. Burger, P. Geladi, Analyst 131 (2006) 1152–1160.
- [12] M.H. Bharati, J.F. MacGregor, M. Champagne, Tappi J. 3 (5) (2004) 8–14.
- [13] L. Eriksson, S. Wold, J. Trygg, J. Chemom. 19 (2005) 393–403.
- [14] J.J. Liu, J.F. MacGregor, C. Duchesne, G. Bartolacci, Miner. Eng. 18 (2005) 65–76.
- [15] M. Bouydaï, J.F. Colom, J. Pladellourens, Tappi J. 82 (7) (1999) 153–158.



- [16] J. Scharcanski, *IEEE Trans. Instrum. Meas.* 55 (5) (2006) 1778–1785.
- [17] M.S. Reis, P.M. Saraiva, *Qual. Technol. Quant. Manag.* 3 (3) (2006) 263–282.
- [18] S. Mallat, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (7) (1989) 674–693.
- [19] M. Jansen, in: P. Bickel, P. Diggle, S. Fienberg, et al., (Eds.), *Noise Reduction by Wavelet Thresholding*, vol. 161, Springer, NY, 2001.
- [20] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, 1998.
- [21] M. Vetterli, J. Kovačević, *Wavelets and Subband Coding*, Prentice Hall, Upper Saddle River, NJ, 1995.
- [22] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, 1999.
- [23] J.E. Jackson, *A User's Guide to Principal Components*, Wiley, New York, 1991.
- [24] I.T. Jolliffe, *Principal Component Analysis*, Springer, New York, 2002.
- [25] J.E. Jackson, *Technometrics* 1 (4) (1959) 359–377.
- [26] J.E. Jackson, G.S. Mudholkar, *Technometrics* 21 (3) (1979) 341–349.
- [27] B.W. Wise, N.B. Gallagher, *J. Process Control* 6 (6) (1996) 329–348.
- [28] J.V. Kresta, J.F. MacGregor, T.E. Marlin, *Can. J. Chem. Eng.* 69 (1991) 35–47.