Coupling 2D-Wavelet decomposition and Multivariate Image Analysis (2D WT-MIA)

Mario Li Vigni[1], José Manuel Prats-Montalban[2], Alberto Ferrer[2], Marina Cocchi[1]

[1]Department of Chemical and Geological Sciences, University of Modena & Reggio Emilia, Via G. Campi, 183, 41125 Modena, Italy

[2]Multivariate Statistical Engineering (GIEM), Departamento de Estadística e I.O. Aplicadas y Calidad, Universitat Politècnica de València, Cno. de Vera s/n, Edificio 7A, 46022 Valencia, Spain

ABSTRACT

The use of 2D Discrete Wavelet Transform in the Feature Enhancement phase of the Multivariate Image Analysis is discussed and implemented in a comparative way with respect to what already present in the literature. In the proposed approach, all of the resulting sub-images obtained by the Discrete Wavelet Transform decomposition are unfolded pixel-wise and mid-level datafused to a Feature Matrix which is used for the Feature Analysis phase. Congruent sub-images can be obtained either by reconstruction of each decomposition block to the original pixel dimensions, or by using the Stationary Wavelet Transform decomposition scheme. The main advantage is that all possible relationships among blocks, decomposition levels and channels are assessed in a single multivariate analysis step (Feature Analysis). This is particularly useful, in a monitoring context, when the aim is building multivariate control charts based on images. Moreover, the approach can be versatile to handle context where several images are analysed at a time as well as in the multispectral images analysis.

Both a set of simple artificial images will be used to highlight the details of the methodology and show how the wavelet transform allow extracting features that are informative of how strong, and in which direction, the texture of the image varies and a set of real images as representative of the on-line quality monitoring context.

Keywords: 2D Wavelet Transform, Multivariate Image Analysis, Multi resolution, Quality monitoring

1. INTRODUCTION

The use of Multivariate Analysis to evaluate images dates back to the mid-late 80's, with the work of Esbensen and Geladi [1] who introduced the use of Principal Component Analysis for the study of multi-channel images. Multivariate Image Analysis (MIA) has soon gained boost with the application in many contexts, typically those with images of such complexity that they could benefit of a multivariate analysis approach (e.g. remote sensing[2-4] and medical imaging[5-9]). In the 90's, the pioneering works of MacGregor and his group made the field of process industry accessible by the MIA approach[2, 10]: the possibility to acquire images for on-line process monitoring purposes and effectively analyse them represents a viable, PAT-like sensor to investigate process changes in time in an environment, that is the process line, where often the room for installing and interface new "traditional" sensors is poor, not to mention the fact that often a single image can acquire simultaneously several different potential sources of variability.

At present, several uses of MIA are reported in literature for different tasks[11-29], all of which are characterised by being well described by the information an image can carry, i.e. textural variability (based on the "two" dimension relationship structure of pixels) and optical variation properties (based on the "third" dimension, that is the channels acquired for each pixel). The latter aspect becomes the more relevant as the number of channels increases, moving from simple binary or gray-scale images (where not much information can be given more than texture homogeneity/non homogeneity), to RGB images (where changes in colour can be related to the presence of non homogeneous texture or underlying phenomena which alter the composition), to multi-channel images and spectral images (where chemical information can effectively be acquired for each pixel). Therefore, most image-based challenges which can be addressed with MIA approach represent the detection of product defects in quality control[11-18], the monitoring of changes in process behaviour and its feed-back control[16, 19-21], the prediction of product properties on the basis of the joint evaluation of texture and channel information (this latter aspect is in particular addressed to by the development of Multivariate Image Regression – MIR – methods[22-23]); or more recently the development of imaging biomarkers in cancer diagnosis[8-9].

As far as the core details of the MIA approach and its evolution in time, the MIA approach proposed by Bharati and MacGregor is based on a framework[11-12] which can be summarised in two main steps: a Feature Extraction (or Enhancement) phase, in which the image (pre-processed, if necessary) is treated so that the texture information carried out by the pixels is made clearer, and a Feature Reduction (or Analysis) phase, where a suitable Multivariate Analysis method is applied (e.g. Principal Component Analysis, Partial Least Squares Regression, Partial Least Squares Discriminant Analysis) on the feature matrix obtained after the first phase. The two phases are strictly connected to each other, since the first step can strongly influence the outcome of the following analysis in a way, which is not much different from the effect of data pre-processing in many other situations. However, a certain degree of freedom can be considered when choosing the feature enhancement method (whilst the feature analysis phase is more application driven): the fundamental aspect to be preserved in this case is that it is not only important to preserve the information given by the channels, for which a simple unfolding of the image structure so that each pixel becomes a sample could be sufficient, but to retain the correlation among neighbouring pixels (that is, the texture information) and, most of all, present it to the following analysis step in such a way that both sources of variability (texture and channel-based properties) can be easily evaluated. In the approach proposed by Bharati and MacGregor[2] the texture information is extracted by augmenting column-wise the unfolded pixel vector with a series of its copies, shifted row-wise so that each row of the generated matrix is formed by a pixel and all its surrounding neighbours. This corresponds to stacking copies of the original image shifted according to a given pace. The number of neighbours, hence of columns, of the feature matrix is $(2w + 1)^2$, governed by the window aperture parameter w, which indicates the dimension of the window, centred on the pixel, encompassing the neighbours to be considered (typically, $w = 1$ or 2). In Prats-Montalbán et al.[17] this augmentation is extended to each channel of the image, we will refer to this now on as colour-textural MIA (ct-MIA), while Facco et al.[14] proposed a method to reduce the computation cost when operating with a larger window size, $w > 2$. Other approaches have been proposed and discussed, among which the most common are based on the application of a transformation of the image, again for each channel, in a different domain, such as the Fourier domain (*via* the 2D Fourier Transform) or the wavelet domain[10-12, 18, 24-29]. The wavelet advantage with respect to Fourier is that it has both good frequency and spatial resolution.

There may be several advantages by moving to wavelet domain in terms of image compression, background removal and denoising. Moreover, the use of wavelet transform allows a better understanding of the pixel correlation structure at different scales, since at each level of decomposition, the coefficients carry both the information pertaining to the energy which characterise a frequency range (based on the selected filter) and an indication about the orientation in which varies (according to the type of decomposition block, being it Approximation or one of the three Details blocks, namely

Horizontal, Vertical and Diagonal). In this way, the features extracted by wavelet transform can be seen as a truly enhanced visualization of how strong, and in which direction, the texture of the image varies. Literature differs in the way these features could be expressed and handled: some authors have pointed the attention to the use of global indicators to synthesise the relevant information for a given decomposition level and block, by using, e.g. the Frobenhius norm (Energy), the entropy, statistical momenta or the standard deviation of the coefficients[18, 26-28], thus having a single variable which summarises the effect, while maintaining the orientation information by means of the level-block combination at which it is computed. This approach surely reduces the computational cost of the following analysis, but carries with itself the potential loss of interesting information, which goes together with an "averaging" procedure of a richer set of data. Also, since all the information of an image is compressed in a single vector of descriptors for each decomposition block and level, a somehow conspicuous set of images has to be considered to create a reference set, for example of Normal Operating Conditions (NOC), when moving to the following Feature Analysis phase. On the contrary, working at pixel level, that is considering each pixel as a sample, opens the possibility to reduce the requirements when building a reference set, often allowing the use of a single representative image only – being it a real one, or a combination of snapshots of NOC texture areas.

Liu and MacGregor[10] have proposed an approach where the wavelet transform is used for Feature Enhancement of images working at pixel level, i.e. the MultiResolutional Multivariate Image Analysis (MR-MIA). MR-MIA is conjugated in two frameworks that differ in the stage at which the wavelet transform is applied, i.e. before (MR-MIA I) or after (MR-MIA II) the Feature Analysis step (in this case PCA). In particular, in MR-

MIA I by applying the discrete wavelet transform (2D DWT) to each channel *ch* of the image, each decomposition block, at a given level, can be seen as an image itself with the same number of channels, but representing a different "resolution" and texture orientation. The Feature Analysis (e.g. PCA) step is then applied to each of these images, once unfolded pixel-wise, so as many latent variable models as number of blocks per decomposition level (L) are obtained. This approach relies on the orthogonality of the wavelet decomposition blocks, thus implying that there is no interest in evaluating correlations among blocks at different scales, and the possibility to evaluate texture – channel correlations is maintained. However, the complexity aspect of this approach can be a hindrance when considering how many multivariate models one should compute and the necessity of a high-level data fusion step where all the results are combined to create a decision rule in order to e.g. decide if a new product image has to be rejected when compared to the NOC modelled one. Recently, basing on similar considerations, Juneau *et al*.[25] proposed an approach where all sub images obtained by wavelet decomposition, once unfolded pixel-wise, are merged row-wise and analysed by a single PCA. However, they used the continuous wavelet transform (undecimated scheme, UWT) and in this way a rather large number of features is obtained, since scale and shift parameters vary continuously.

The MSMIA approach proposed by Reis[18] is similar to the MR-MIA I, although more images are considered at the same time as references NOC, but it differs in the way information from the Feature Analysis step (e.g. PCA) of each WT decomposition block is fused. In this approach, an index evaluating the distance to the scores distribution histogram of the reference NOC images for each decomposition block, at a given scale, is calculated in order to obtain a set of variables, which are then used for building a monitoring chart. In addition to this, multivariate control charts based on PCA of wavelet features (e.g. standard deviation of wavelet coefficients for each decomposition block), extracted for each decomposition block, at a given scale, are also considered. The approach is effective in compressing information and for on-line implementation, however defects location requires a further step. This step, similarly to MR-MIA II, consists in building a spatial shifting feature matrix (then analysed by PCA) for each of the sub-images contributing to “out of control” observations in the preceding step. Moreover, correlation structure among textural/colour pattern at different scales is only indirectly taken into account (the information from different scales is always merged at features, not pixels level).

Here we present an approach, which is named 2D WT-MIA, where the Feature Enhancement step, is similar to MR-MIA I, but as in Juneau[25] all of the resulting sub-images obtained by the 2D-DWT decomposition are unfolded pixel-wise and mid-level datafused to a Feature Matrix which is used for the Feature Analysis phase. In order to have congruent sub-images all decomposition blocks are reconstructed separately to the original pixel dimensions. This reconstruction step can be omitted, when the Stationary Wavelet Transform (2D-SWT) is used[30-31]. In this way, all possible relationships among blocks, decomposition levels and channels are assessed in a single multivariate analysis step (Feature Analysis). This is particularly useful, in a monitoring context, when the aim is building multivariate control charts based on NOC images. Thus, our proposed approach can be versatile to handle context where several images are analysed at a time as well as in the multispectral images analysis.

The rest of this paper is organised as follows: in Section 2: Methods, the proposed 2D WT-MIA approach is described into more details and compared to colour-textural MIA approach to highlight common and differing aspects. In Section 3: Materials, the dataset images which will be used will be presented, consisting in a set of simple artificial binary images, used to illustrate how texture is captured within the two approaches and a set of real images, and in Section 4: Results and Discussion, these images will be analysed according to the two-step MR-MIA framework, using Principal Component Analysis as Feature Analysis technique with the target of simulating a quality control task.

2. METHODS

The approach here described belongs to the more general framework of MultiResolution Multivariate Image Analysis, thus basing on a two-phase elaboration of the image: a first step of Feature Extraction (Enhancement) and a second step of Feature Reduction (Analysis) (Figure 1). In particular, the 2D WT-MIA (wavelet based feature enhancement) and the colour-textural MIA (spatial shifting based feature enhancement) approaches will be discussed and compared in terms of results in the present work.

2.1 Spatial Shifting Feature Enhancement

Colour-textural MIA[17] is summarised in Figure 2. This approach to Feature Enhancement consists in capturing, for each channel *ch*, the pixel proximity correlation by means of a spatial shifting of neighbouring pixels with respect to each pixel of the original image, according to a selected window aperture parameter, w. In practice, starting from a pixel element of the image $p_{i,j}$, a row vector is created by adding the intensity value of the channel corresponding to the closest surrounding pixels: if $w = 1$, the composition appears as reported in Figure 2. When this is done for all the pixels of a pixel-wise unfolded channel matrix, and the feature matrices for the different channels are then fused, a total Feature Matrix is obtained which has as many rows as the number of pixels $I = n_1 \times n_2$, and as many columns as the number of channels ch times the number of spatial shifted pixels, which is given by $(2w + 1)^2$. This means that, regardless of the number of channels, for a window parameter $w = 1$ (the closest neighbours) the Feature matrix is $9 \times ch$, and when moving to $w = 2$ (the closest neighbours and the next surrounding layer), the Feature matrix is $25 \times ch$. This implies a fast increase of the number of variables considered in the analysis, the higher is the number of channels.

Since we need, at each pixel location, to use all the neighbouring pixels until a distance w, this implies that we lack of information for all those pixels in the borders with width w. Theferore, the solution commonly adopted is to diminish the size of the image from $n_1 \times n_2$ to $(n_1\text{-}2w) \times (n_2\text{-}2w)$

2.2 2D Wavelet-based Feature Enhancement

Figure 3 shows the general scheme of the Feature Enhancement step involving 2D – DWT application, through the fast Mallat algorithm[32-33], on an image. For each channel ch, the low-pass and high-pass filters (which are the same as in the 1D case) are first operated row-wise on the image and then, after downsampling of the coefficients, in each of the resulting blocks column-wise. In this way four decomposition blocks are obtained: Approximations (low-pass+low-pass), namely CA; Horizontal details (low+high), namely CH; Vertical details (high+low), namely CV, and Diagonal details (high+high), namely CD. The procedure is then iterated by applying it to the Approximations, i.e. increasing the decomposition level. Downsampling is skipped when the 2D - SWT scheme is used since, instead, the filters are up-sampled[26]. The maximum possible decomposition level, L, depends on the image size. The four decomposition blocks obtained from each level of decomposition (CA, CH, CV and CD) when 2D DWT is used are independently reconstructed by means of the inverse 2D – DWT so that their dimensions are the same of the original image, while they are already of the same size when 2D SWT is used (in fact, each block of coefficients at every level maintains the same size as the original image, and congruent images are obtained). This leads to a set of $4 \times L$ images for each channel *ch*, which can be unfolded and column-wise merged to obtain a total Feature Matrix which has as many rows as the number of pixels $I = n_1 \times n_2$, and as many columns as $4 \times L \times ch$. If we compare this column dimension to the one obtained with the Spatial Shifting approach, which is $(2w + 1)^2 \times ch$, it might appear that there is little benefit in terms of reduction of the Feature Matrix dimensions. However, two aspects have to be underlined:

i) in the spatial shifting approach the image is analysed by moving a $(2w + 1)x(2w + 1)$ pixels window by step of 1 in all image directions; on the other hand, with wavelet, a *filter length x filter length* pixels window is moved by step of 1 in all image directions, but using a larger filter does not increase the number of features, which remain always four;

ii) the two approaches lead to the same number of feature descriptors if $L = \text{round}[(w + ½)^2]$. This corresponds, for e.g. a window parameter of $w = 2$, to a decomposition level $L = 6$, which in terms of multiresolution means to have gone very deep in the analysis of coarse and smooth aspects of the image. In other words, such a decomposition level (if allowed by the nature of the chosen wavelet) usually leads to the possibility of evaluating correlations and high distance relationships among pixels to an extent, which is superior to the use of a moving window of fixed size.

When applying the wavelet transform, the selection of the most appropriate wavelet filter is considered a critical issue and a limiting step in implementation of routine applications (i.e. which wavelet family and which filter length, to analyse the specific characteristics of the images at hand). This issue has been dealt in literature by analysing the different properties of the decomposition filter in terms of texture description capability in order to propose general criteria[34] or focusing on goodness of image reconstruction[33], or proposing a design of experiments approach[36]. We recently proposed[37] a methodology based on N-way modelling to provide a range of possible wavelet choices (in terms of families, filters, and decomposition levels), for each image and problem at hand. Any of these strategies requires a preliminary analysis step to be conducted by experienced people in the field, although this step is only required once in model building. However, some considerations, based on our experience can be drawn: i) there is in general a relation between the decomposition level and the filter length, i.e. by using a larger filter a lower decomposition level is required to capture the different image aspects (coarse and smooth) and ii) taking into account the wavelet families characteristics, such as for degree of symmetry or regularity, number of vanishing moments[38-39] it is possible to focus on a small number of wavelet filters to test, by choosing a representative one for each type of property.

3. MATERIALS

3.1 Artificial Images datasets

These sets are used to illustrate how the colour-textural MIA and 2D WT-MIA approaches analyse texture and their capability to detect faulty pixels. These images are characterized by two main features: a particularly limited pixel size, so that computational time is not a relevant benchmark property at this stage, and a simple, yet well defined, pattern. Also, the differences between "normal", i.e. reference image, and "defective", i.e. image (or images) for which a perturbation of the pattern was created, are well controlled, in the sense that the number and position of pixels which have been changed is known, and the entity of the disturb is enough to obtain simulated test images which are not too similar to their reference image. In spite of the simplicity of this simulated case, the information which can be acquired from the analysis with both approaches is interesting to better understand how the two methods under comparison work, and the conclusions which can be drawn are helpful and can be extended, as shown in the next section where real images are presented and dealt with, to cases of higher complexity.

The set is composed by three binary images, as reported in Figure 4, of pixel size 32 × 32. Figure 4 "SimSetA" reports the "normal" (reference) image, on the basis of which an alternation pattern has been generated. In this case, the squares which alternate in both image directions to give a chequered pattern have a dimension of 8 × 8 pixels: starting from upper left corner and moving over columns dimension, a white (1's) 8 × 8 pixel square is alternated to a black (0's) 8 × 8 pixel square, and the same alternated pattern is repeated over the rows dimension. Starting from this image, two changes in pattern were produced, leading to two "defective" (test) images. Figure 4 "SimSetB" shows an overlying irregular shape which extends from the diagonal to the lower left part of the image: for this figure, a total of 55 pixels have been inverted in value (from 1 to 0 or from 0 to 1) over the total of 32 × 32 = 1024 pixels. Figure 4 "SimSetC" shows another change in the pattern, this time according to a regular shape which is applied on top of each of the 8 × 8 pixel squares: for each of these squares, starting from the second diagonal element, a single pixel every fourth has been modified both in the rows and in the columns, thus resulting in a change of four pixels for each of the squares. For this figure, a total of 64 pixels have been inverted in value (from 1 to 0 or from 0 to 1) over the total of 32 × 32 = 1024 pixels.

3.2 Real Images datasets

To further explore the performance of the method proposed in this work and compare it to the results of the ct-MIA approach, additional datasets have been taken into account, belonging to different applicative contexts, tiles and bread production, respectively. In both cases, the control of the final product undergoes visual inspection, while the datasets differ as for image dimensions and number of channels.

3.2.1 Tiles

These datasets come from a production of tiles of marble-like materials for surface coverage: all the cases share a common issue, that is presenting product samples which do not comply to a strict definition of "normal" images, characterized, for instance, by a precise colour shade or by the absence of defects such as spots and scratches. Therefore, it is necessary to develop a method, complementary or alternative to visual inspection, which is able to: a) recognize the presence of a defectiveness when a new tile is compared to the reference one(s); b) indicate the kind of defectiveness (e.g. colour shade and/or presence of unwanted changes in surface pattern); c) locate on the surface the position of the defect in order to obtain an enhanced perception of the same, so that its visualization and recognition by the operator is made easier. Samples from two different products were considered with different degree of irregularity of the pattern in the defective tiles. They both consists of RGB images of dimensions are 256 × 256 pixels (Figure 5). Figure 5a reports dataset 1: Blanco Zeus, from now on referred to as BZdataset, which is composed by three reference images (BZN01, BZN02 and BZN03), and three images of tiles showing defects (BZD01, BZD02, and BZD03). This kind of tile shows a mostly homogeneous shade of gray all over its surface, so that defects (as for instance white or dark spots and scratches) do not usually present particularly high difficulty of detection also by visual inspection. Figure 5b reports dataset 2: Blanco Norte, from now on referred to as BNdataset, which is composed by three reference images (BNN01, BNN02 and BNN03), and three images of tiles showing defects (BND01, BND02, and BND03). In this case, the tile main colour is gray, but the surface is characterized by an inhomogeneous distribution of darker spots, in a grainy structure, which makes quite difficult to detect the presence of defectiveness, both when represented by darker and paler areas.

3.2.2 Bread

This data set comes from industrial production of bun bread, where a digital scanner is already used to automatically assess defects concerning mainly bun dimensions, while surface defectiveness, such as dark spots, blisters, and pale areas is still evaluated by visual inspection of expert personnel. These defects arise from different causes, some of which not perfectly known, and are also often difficult to be detected by RGB online cameras. Thus, a feasibility study has been undertaken[29] by acquiring offline multispectral images, covering the UV-visible range (from 430 to 700 nm, 10 channels) and the short-wavelength NIR range (from 850 to 970 nm, 8 channels), which can improve the acquisition of information on bread quality, combining spectral (NIR may represent also a "chemical signature") and textural information. The whole data set has been analysed by WT-MIA (DWT scheme) approach and described in detail in ref. 33 while here a subset of images has been analysed in order to discuss comparatively the performance of WT-MIA (DWT and SWT) and ct-MIA .

The raw images were cropped to remove the distortion effect of the round bun shape, and background and noise were removed via preliminary wavelet analysis[29], finally giving images of dimensions of about 387 x 420 pixels for 18 channels. Here two non-defective images (N01, used as reference, and N02) and two defective ones (D04 and D07) are analysed, shown in Figure 6.

4. RESULTS AND DISCUSSION

4.1 Artificial Images datasets

All the three images (SimSetA, SimSetB and SimSetC) have been treated according to the same Feature Enhancement step, by considering:

- Spatial Shifting, colour-textural MIA (ct-MIA) with window size parameter w = 1
- Wavelet Decomposition (WT-MIA) by using a Daubechies 1 (db1) at decomposition level L = 1, both DWT and SWT.

The Feature Enhancement step gave a Feature Matrix of dimensions I = 900 rows (reduction from 32 x 32 to 30 x 30 is necessary to cope with borders) and 9 columns for the ct-MIA approach and I = 1024 rows x 4 columns for the WT-MIA approaches.

SimSetA was used as the reference set, upon which for the ct-MIA approach a Principal Component Analysis model was obtained after mean centring of the Feature Matrix.

Figure 7 reports the PCA results from left to right in the order scores image SimSetA, loadings, projected scores images SimSetB and SimSetC (in the order PC1 to PC4 from top to bottom). PC1 captures, for the reference ("normal") image both the difference in grey intensity value (colour) and the variation in pattern (texture) when passing from the pixels having zero value to pixels with value one, i.e. it shows the change of value when moving along the borders from one square to another, where pixel values invert, leading to a "blurring" effect of the borders. This is the expected effect since ct-MIA window of one (which is actually 3×3 pixels) moves pixel by pixel on the image structure, which is made of 16 squares of dimensions 8×8. All features contribute similarly to the PC1 loadings since the pattern change takes place in all directions. Thus, PC1 works as an average grey scale image, which in fact extracts out the spectral information (we have no other source of spectral info that a single gray scale channel). The following PCs capture only the borders effects, i.e. only the frame around the squares are visible in the scores images, and by inspecting the loadings it is possible to understand the directions of the pattern variation, e.g. to PC4 the features accounting for diagonal shift do not contribute.

When the Feature Matrix corresponding to SimSetB and SimSetC are projected onto this model, the same chequered pattern is correctly reproduced (Figure 7), but the changes in pixel correlations due to the small scale modifications of its regularity produce a large blurred area, which roughly encompasses the whole shape of the differences but extends further with respect to the faulty pixels, i.e. an area of about 3×3 pixels around each defective pixel as it is detailed in the following text. This is due to the fact that the perturbation, although being well defined (in particular for SimSetC) to a small number of pixels, influences the neighbouring correlation structure of all the pixels, which are contained by the moving window.

In a monitoring context the defective images with respect to the reference one/s can be identified by the Hotelling-T^2 and squared residuals (RSS, SPE or Q) multivariate control charts by using the percentage of pixels beyond control limits[16]. However, in this case being the images binary simply the pixel by pixel difference of the residuals sum of squares (RSS) of the test images with respect to the reference image (NOC) can be used. Both the RSS from a one or a four components PCA model are suitable to depict the faulty pixels for SimSetB and SimSetC, but also an area of about 3 x 3 around each faulty pixel will show up differing in RSS values with respect to NOC (Figure S2, supplementary material). This can be expected on the basis of the considerations made above on the neighbouring pixels correlation structure.

In the WT-MIA both decomposition scheme DWT and SWT have been applied, considering the simulated pattern, i.e. inversion of the binary value of some not consecutive pixels, db1 seems an appropriate filter. The feature matrix, holding the four decomposition blocks CA, CH, CV and CD (reconstructed only in DWT case), already captures the texture pattern, as highlighted in Figure 8 (DWT) and Figure 9 (SWT) where the sub-images corresponding to each block of the DWT and SWT decomposition of SimSetA and SimSetB are reported. This is a first difference with respect to ct-MIA approach where the feature matrix holds just the shifted version of the raw image in all possible neighbouring direction (as shown in Figure S1, supplementary material) and thus PCA (in general a multivariate decomposition technique) is needed to reveal the texture pattern. Further, in this case with only one reference image and one channel the feature analysis step by PCA is not needed at all (the decomposition blocks are orthogonal).

In the DWT the Approximation Block is the only one which carries the structural information of image SimSetA (Figure 8, top). This is explainable by considering that the db1 filter is of length two, thus operates like a window of pixel size 2 × 2 which moves at steps of one, and due to the down-sampling scheme of DWT only one coefficient every two is retained. Thus, since the binary values change every 8 x 8 pixels, there is no blurring effect at the squares edges; also the coefficients in all the other decomposition blocks are zeros (Figure 8, top). On the contrary, the presence of deviations in the two test images, related to "sharper" structures i.e. alternating by one pixel, is well captured by all decomposition blocks (Figure 8, middle). In particular Approximation (CA) shows both intensity change and texture, while Detail blocks capture horizontal (CH), vertical (CV) and diagonal (CD) neighbouring pixels alternation of binary values. Thinking of a monitoring context, in this case the defects can be depicted by the difference between the decomposition sub-images of the reference and test image, as shown in Figure 8, bottom; to this aim, considering the specific pattern of the defects in SimSetB CA and CD are the most suitable blocks. Figure 9 shows the results of the SWT decomposition. Similar considerations can be drawn. The only difference is that now the effect of the variation in the binary values of the pixels at the edges of the 8 x 8 pixels squares are visible (similarly to ct-MIA). This is well explainable by the fact that in SWT down-sampling of the coefficients is not operated (to be noticed that the window size and the moving step remain the same). Analogous considerations hold for SimSetC decomposition (figure not shown for sake of brevity).

The behaviour of a larger filter, i.e. Daubechies 2 (db2) of length 4, has been also inspected by using the SWT scheme on SimSetB (Figure S3, supplementary materials). The db2 operates as a 4 × 4 window moving at steps of one pixel: the result is similar to the one obtained by ct-MIA, leading to a blurring of the borders among squares and around the area (of wideness about 3 x 3) which is interested by the defect.

Finally in Figure 10 the performance of ct-MIA and WT-MIA (SWT, db1) are compared in terms of capability of detection and localization of the faulty pixels for SimSetB (Figure 10a) and SimSetC (Figure 10b), respectively. The detection is good in both approaches being all the faulty pixels correctly identified, the difference is in the blurring area, which is strictly connected to the wideness of the analysing window, i.e. 3 x 3 for ct-MIA and 2 x 2 for db1. This is a general known advantage of WT of being more efficient for feature enhancement because of the availability of several filter shapes and length compared to spatial shifting approach.

4.2 Real Images datasets

4.2.1 Tiles

Several wavelet filters, belonging to Daubechies (filter length from 1 to 5), Symlet (filter length from 1 to 5), Coiflet (filter length form1 to 3) and biorthogonal (1.3 and 1.5) families were tested (decomposition levels from 1 to maximum), by using an approach as described in ref. 30. For both BZdataset and BNdataset Daubechies filter length 1 (db1 or Haar) resulted among the best performing wavelet filter and we report results relative to this filter, at decomposition level $L = 3$. While for the ct-MIA approach window size 1 and 2 were considered, better performance was obtained with w

= 1 for BZdataset and w = 2 for BNdataset. This lead, considering the three RGB channels, to an unfolded feature matrix of size 65536 × 27 (w = 1), or 65536 x 75 (w = 2) in the ct-MIA case, and 65536 × 36 in the WT-MIA case.

In both datasets, a single reference image has been used to calibrate the PCA models and build the Hotelling's T^2 and Q statistics (control charts). Autoscaling pretreatment gave for both datasets and approaches the best results.

The choice of model dimensionality, i.e. number of principal components, in this context cannot be automated, i.e. assessed on the basis of a priori fixed criterion, since it is problem dependent[40]. General guidelines that we adopted in this work, is to inspect how spatial features of the image are accounted for in scores images, and to scree-plot to ensure the systematic variation is modelled. Further, when enough defects images are available, to preserve some for model validation, few can be used to see which are the components that maximize detection capacity. It is worth noticing that minimizing the squared prediction error in cross validation, as most used in PCA modelling, is not appropriate in this context, because it is not necessarily related to the capability of fault detection which is the objective pursued in process monitoring.

Image BZN01 has been used as reference NOC image for BZdataset. The PCA model dimensionalities were 4 PC's for both approaches ct-MIA (captured variance 77%) and WT-MIA (captured variance 44%), which correspond to a number of components each explaining more than 1% variance (ct-MIA) and to the first minimum in the scree-plot, i.e. number of components vs. eigenvalues plot, (WT-MIA), respectively. All the remaining images of the dataset were projected onto the models and distances were calculated. Table 1 reports the results in terms of percentage of pixels scoring above the critical limits, which were chosen on the basis of the reference image by obtaining the 99th percentile values of its distances distributions. Both models are able to accept as normal behaving images BZN02 and BZN03, which are actually defectless, and indicate, especially in terms of T^2 distance, the presence of anomalies on all of the three defective tiles, BZD01, BZD02 and BZD03; albeit the results are quite similar, a higher percentage of pixels above the critical limits is detected by the WT-MIA approach. Both approaches show similar results, although the WT-MIA identification of defects appears better defined, especially for image BZD02 where more clusters of pixels are identified, which are in particular connected to the presence of darker spots on the surface of the tile, especially when using SWT (Figure S4, supplementary Material). SWT monitoring results are also shown on Table 1 and are very close to DWT one.

Interpretation of the features enhancement step can be gathered by loadings analysis:

ct-MIA loadings are shown in Figure 11 (left) both as bar plot (top left) and refolded (bottom left) in the corresponding position of neighbours window (the central pixel is the pixel itself). As usual PC1 is gathering an (approximately) average colour effect (all loadings have the same sign). Moreover, it can be observed that colour intensity varies left to right for red and blue channels, while green is more uniform; similarly does PC1 of WT-MIA (bar plot, top right, and decomposition sub-images, bottom right), to which contribute the Approximations of all levels and channels (Approximations in fact act as an averaging tool at each decomposition level, hence extracting out the same phenomenon as ct-MIA). From the WT-MIA Approximations sub-images (Figure 11, bottom right) it is also evident the varying intensity left to right, especially for decomposition levels two and three (the green channel is uniform at level 1). This effect may be due to illumination and eventually (but was not the aim here) it could be easily removed in WT domain, e.g. by suppressing level 2 or 3 approximations as data pre-treatment[29].

PC2 and PC3 show the main contrast in horizontal and vertical directions, respectively both for ct-MIA and WT-MIA (for PC2, the Horizontal details of level 3 for all channels are the most relevant, and for PC3 the vertical details, level 1 opposite to level 3). PC4 shows a mixed pattern, loadings sign and values vary in all direction for ct-MIA and for PC4 in WT-MIA the vertical detail of level 2 is the most relevant.

It can be noticed that the possibility to analyse the images at different resolution (the different decomposition levels) enhances the color-textural pattern recovery, with respect to ct-MIA where only the neighbouring window size can be varied (that in WT-MIA can roughly corresponds to the filter length/family)

Table 1 to be inserted about here

As reference NOC image for BNdataset, image BNN01 has been used, the PCA model dimensionalities were 2 PC's for both approaches ct-MIA (variance captured 39%) and WT-MIA (variance captured 26%), which correspond to the first minimum in the scree-plot. All the remaining images of the dataset were projected onto the models and distances were calculated. Table 2 reports the results in terms of percentage of pixels scoring above the critical limits, which were chosen on the basis of the reference image by obtaining the 95th percentile values of its distances distributions.

Table 2 to be inserted about here

Neither of the models does not appear particularly satisfactory, since the normal behaving images BNN02 and BNN03, which are defectless, appear to have Q distances higher than 5%. The defective tiles, BND01 and BND03 appear above limits for both models, according to T^2 distance statistic. On the contrary, defective BND02 is only detected by WT-MIA, T^2 distance, albeit close to the limit.

By considering the T^2 distance values reshaped in the original pixel domain it is possible to identify the groups of pixels which correspond to the passing of the critical values. Figure 12a) and 12b) shows the comparison of defective images and the normal images, with the corresponding distance images for ct-MIA and WT-MIA (DWT). The WT-MIA identification of defects appears better defined, while ct-MIA seems to find less clusters of pixels and more darker, well separated, spots all over the surface.

In a monitoring context, the results of Table 2 would indicate products BNN02 and BNN03 as defective (false negatives) and shed doubt on rejecting or not product BND02. On the other hand the possibility to look at above limits T^2 distance images (or in general to the images corresponding to the above limit statistic) may clarify if defects are present or not. In particular, this is a case were the defects are mainly due to a non uniform distribution of pixels with a given colour content and texture that if normally distributed on the image, as in the case of BNN01, BNN02 and BNN03, would be acceptable. In this situation, it may be useful to calculate and represent the local entropy[41] of the scores images, were the defective area are region of low entropy encircled by high entropy values, as shown in Figure 13.

WT-MIA model based on SWT in this case yielded lower performance.

4.2.2 Bread

The Daubechies 2 (db2) wavelet filter was used up to decomposition level 5 and both DWT and SWT decomposition schemes. The feature data matrix results of dimension

I_{pixels} x 360 (4 blocks x 5 levels x 18 channels) is obtained. In ct-MIA both a window size of 1 (162 features) and 2 (450 features) were tested. Since results were similar we will discuss the ones corresponding to $w = 1$, which gave a better defects localization. The reference PCA model for non-defective image has been calculated by considering as feature matrix the one obtained for N01 image (Figure 6). The PCA model refers to mean centred data and model dimensionalities were 6 PC's for both approaches ct-MIA (variance captured 66%) and WT-MIA (variance captured 52%), which correspond to reaching the plateau in the scree-plot. We tested also a model made on two NOC images but the results were analogous. Then, Q and T^2 statistics were computed, and the critical limits for each of the two statistics were computed on the basis of the 99th percentile. The total percentage of pixels exceeding the critical limits is reported in Table 3. For all approaches a clear detection of the two defective images can be obtained, with relevant percentages of pixels above the critical limits for both distances, as well as N02 being defectless.

However, when the Q and T^2 values above the reference limits are refolded to the original pixel x pixel domain to locate the defective areas on the image (Figure 14), differences among the approaches emerge. ct-MIA is less efficient to detect the defective area for D07 and for D04, it is also worth noticing that ct-MIA provides these results when applied on the pretreated images, i.e. after denoising and background removal with WT; otherwise it detects as faulty only pixels on the borders of the image. DWT seems more efficient than SWT to locate the faulty pixels, notwithstanding the same wavelet filter and resolution has been used.

Now focusing on the WT-MIA DWT results, it is worth noticing that both stains, which are also easy to detect visually, but as well blisters and tiny scratches could be detected.

Moreover, we can assess which features are responsible of the defects by inspecting the T^2-contributions, which can be interpreted in terms of the spectral channels. In particular, Figure 15 shows the T^2-contributions for some of the blisters. To make the representation clearer distinct plots are made for each decomposition block, and each decomposition level is represented as a distinct line, so that the x-axis reports just the channels (wavelengths): the main contributions are from approximations decomposition levels 1-3, interestingly besides the visible channels some of the NIR ones (11^{th} to 18^{th} corresponding to the range from 850 to 970 nm at 20 nm resolution) contribute, which point to carbohydrate, fat and water bands. This may indicate a segregation of some of the ingredients on surface spots where blisters appear.

CONCLUSIONS

The artificial image datasets allowed highlighting the distinct way in which textural information can be recovered by the ct-MIA and WT-MIA approaches, both are efficient in depicting the salient pattern of the images and the area were the defects are located. The main distinctive characteristics of the two methods are:

i) the feature matrix obtained by ct-MIA holds just the shifted version of the raw image thus always requires coupling to a multivariate decomposition technique to highlight textural patterns while the feature matrix obtained by WT-MIA already capture it;

ii) in general WT-MIA is more efficient for feature enhancement because of the availability of several filter shapes and length compared to the spatial shifting approach where only the window size can be varied.

The analysis of the tiles data sets reveals a similar behaviour of the two considered approaches although identification of defects appears better defined with the WT-MIA approach. Also both decomposition schemes DWT and SWT show similar performance.

In a monitoring context it is worth noticing that when the defects are due to a non uniform distribution of pixels, whose colour content and texture if normally distributed on the image would be instead acceptable, further image analysis tools (e.g. local entropy or any other to assess homogeneity or heterogeneity of pixels distribution), on the beyond Q or T^2 limits images, are required to avoid false negative to be detected.

In the analysis of multispectral images (bread data set) the WT-MIA approach performed better and it was possible to highlight the full benefit of the proposed approach from both points of view of correct defects identification/location and interpretation in terms of spectral features.

A further remark is that the proposed WT-MIA approach is rather straightforward requiring only the Feature Extraction (Enhancement) and Reduction (Analysis) steps, as in ct-MIA; one or more NOC images can be analysed at the same time and assembled in the same WT features matrix which is organized pixels wise, thus allowing defect localization directly. Images denoising and background removal can be as well accomplished at WT decomposition stage.

The proposed WT-MIA approach can be as well applied to hyperspectral images, the bread data set is an example limited to eighteen channels. However, the computational costs will be a limiting factor and further strategies could be considered to render it more efficient, work is in progress in this direction.

REFERENCES

1. Esbensen K.H., Geladi P., Strategy of Multivariate Image Analysis (MIA), Chemom. Intell. Lab. Syst. 1989; 7: 67– 86.
2. Bharati M. H., Liu J. J., MacGregor J. F., Image texture analysis: methods and comparisons, Chemom. Intell. Lab. Syst. 2004; 72: 57-71.
3. Bo S., Ding L., Li H., Di F., Zhu C., Mean shift-based clustering analysis of multispectral remote sensing imagery, International Journal of Remote Sensing 2009; 30: 817-827.
4. Villa A., Benediktsson J. A., Chanussot J., Jutten C., Hyperspectral Image Classification With Independent Component Discriminant Analysis, IEEE Transactions on Geoscience and Remote Sensing, 2011; 49: 4865-4876.
5. JohnN. M., Kabuka M. R., Ibrahim M. O., Multivariate statistical model for 3D image segmentation with application to medical images, Journal of Digital Imaging, 2003; 16: 365-377.
6. Nattkemper T. W., Multivariate image analysis in biomedicine, Journal of Biomedical Informatics, 2004; 37: 380-391.
7. Hackmack K., Paul F., Weygandt M., Allefeld C., Haynes J. D., Multi-scale classification of disease using structural MRI and wavelet transform, NeuroImage 2012; 62: 48-58.
8. Sanz-Requena, R; Prats-Montalbán, J.M.; Martí-Bonmatí, L.; Ferrer, A.; Alberich-Bayarri, A.; García-Martí, G.; Pérez, R. Automatic Individual Arterial Input Functions Calculated From PCA Outperform Manual and Population-Averaged Approaches for the Pharmacokinetic Modeling of DCE-MR Images. Journal of Magnetic Resonance Imaging, 2015; 42: 477–487.
9. Prats-Montalbán J.M., Aguado E.,Ferrer A.. Chapter 16: Multivariate Curve Resolution for Magnetic Resonance Image analysis: applications in prostate cancer biomarkers development. In Ruckebusch C ed. "Resolving Spectral Mixtures, with application from ultrafast spectroscopy to super-resolution imaging", Data Handling in Science and Technology 30, Elsevier, 2017: 519–550.
10. Liu J., MacGregor J., On the extraction of spectral and spatial information from images, Chemom. Intell. Lab. Syst. 2007; 85: 119-130.
11. Duchesne C., Liu J. J., MacGregor J. F., Multivariate image analysis in the process industries: A review, Chemom. Intell. Lab. Syst. 2012; 117: 116-128.
12. Prats-Montalban J. M., de Juan A., Ferrer A., Multivariate image analysis: A review with applications, Chemom. Intell. Lab. Syst. 2011; 107: -23.
13. Reis M.S., Multivariate image analysis. In: Granato D., Ares G. ed. Mathematical and Statistical Methods in Food Science and Technology. Chichester: Wiley-Blackwell; 2014: 201-218. ISBN: 978-1-118-43368-3.
14. Facco P., Masiero A., Beghi A., Advances on multivariate image analysis for product quality monitoring, Journal of Process Control, 2013; 23: 89-98.
15. Antonelli A., Cocchi M., Fava P., et al., Automated evaluation of food colour by means of multivariate image analysis coupled to a wavelet-based classification algorithm, Anal. Chim. Acta, 2004; 515: 3-13.
16. Prats-Montalban J.M., Ferrer A., Statistical process control based on Multivariate Image Analysis: A new proposal for monitoring and defect detection, Computers and Chemical Engineering 2014; 74: 501-511.
17. Prats-Montalban J.M., Ferrer A., Integration of colour and textural information in multivariate image analysis: defect detection and classification issues, J. Chemom. 2007; 21: 10-23.

18. Reis M.S., An integrated multiscale and multivariate image analysis framework for process monitoring of colour random textures: MSMIA. Chemom. Intell. Lab. Syst. 2015; 142: 36-48.
19. Kourti T., Process Analytical Technology Beyond Real-Time Analyzers: The Role of Multivariate Analysis, Critical Reviews in Analytical Chemistry, 2006; 36: 257-278.
20. Bharati M., MacGregor J. F., Softwood lumber grading through on line multivariate image analysis, Industrial and Engineering Chemistry Research, 2003; 42: 5345-5353.
21. Boldrini B., Kessler W., Rebner K., Kessler R. W., Hyperspectral imaging: a review of best practice, performance and pitfalls for in-line and on-line applications J. Near Infrared Spectroscopy, 2012; 20: 483-508.
22. Liu J. J., MacGregor J. F., Estimation and monitoring of product aesthetics: application to manufacturing of “engineered stone” countertops Machine Vision and Applications, 2006; 16: 374-383.
23. Elmasry G., Kamruzzaman M., Sun D. W., Allen P., Principles and Applications of Hyperspectral Imaging in Quality Evaluation of Agro-Food Products: A Review, Critical Reviews in Food Science and Nutrition, 2012; 52: 999-1023.
24. Ganesan R., Das T. K., Venkataraman V., Wavelet-based multiscale statistical process monitoring: A literature review, IIE Transactions, 2004; 36: 787-806.
25. Juneau P., Garnier A., Duchesne C., The undecimated wavelet transform – multivariate image analysis (UWT-MIA) for simultaneous extraction of spectral and spatial information, Chemom. Intell. Lab. Syst., 2015; 142: 304-318.
26. Reis M. S., Bauer A., Wavelet texture analysis of on-line acquired images for paper formation assessment and monitoring, Chemom. Intell. Lab. Syst., 2009; 95: 129-137.
27. Facco P., Tomba E., Roso M., et al., Automatic characterization of nanofiber assemblies by image texture analysis, Chemome. Intell. Lab. Syst. 2010; 103: 66–75.
28. Facco P., Santomaso A.C., Barolo M., Artificial vision system for particle size characterization from bulk materials, Chemical Engineering Science 2017; 164: 246–257.
29. Li Vigni M., Cocchi M., Chpt. 13 Multiresolution Analysis and Chemometrics for Pattern Enhancement and Resolution in Spectral Signals and Images, In Ruckebusch C ed. "Resolving Spectral Mixtures, with application from ultrafast spectroscopy to super-resolution imaging", Data Handling in Science and Technology 30, Elsevier, 2017:.
30. Pesquet J.C., Krim H., Carfantan H. Time invariant orthonormal wavelet representation. IEEE Transactions on Signal Processing, 1996; 44: 1964–1970.
31. Coifman R., Donoho D. Translation-invariant de-noising. In Wavelets and Statistics, (ed). Springer Verlag, Lecture Notes in Statistics: New York, 1995, 125–130.
32. Mallat S. A theory for multi-resolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell. 1989; 11: 674–693.
33. Cohen A., Daubechies I., Jawerth B., Vial P., Multiresolution analysis, wavelets and fast wavelet transform on an interval. CRAS Paris, Ser. A, 1993; 316: 417–421.
34. Mojsilovic A., Popovic V., Rackov D. M., On the selection for an optimal wavelet basis for texture characterization. IEEE Trans. Image Process. 2000; 9: 2043–2050.
35. Villasenor J. D., Belzer B., Liao J., Wavelet filter evaluation for image compression. IEEE Trans. Image Process., 1995; 4: 1053–1060.
36. Svensson O., Abrahamsson K., Engelbrektsson J., et al., An evaluation of 2D-wavelet filters for estimation of differences in textures of pharmaceutical tablets, Chemom. Intell. Lab. Syst., 2006; 84: 3–8.
37. Prats-Montalban J. M., Cocchi M., Ferrer A., N-way modeling for wavelet filter determination in multivariate image analysis. J Chemometr. 2015; 29: 379–88.
38. Mallet Y., de Vel O., Coomans D., Fundamentals of wavelet transforms. In Wavelets in

Chemistry, (ed). Elsevier Science B. V.: Amsterdam, 2000: 77-84.
39. Misiti M., Wavelet Toolbox User's Guide, MathWorks: Massachusetts, 1996.
40. Camacho J., Ferrer A., Cross-validation in PCA models with the element wise k-fold (ekf) algorithm: practical aspects. Chemom Intell Lab Syst, 2014; 131: 37–50.
41. Gonzalez R.C., Woods R.E., Eddins S.L., Digital Image Processing Using MATLAB, New Jersey, Prentice Hall, 2003.

Table 1. BZDataset. Percentage of pixels above Hotelling's T^2 and Residuals Q distances critical limits based on normal image BZN01 99th percentile

	ct-MIA *w* = 1, 4 PCs		**WT-MIA (DWT)** *Daubechies* 1, level = 3 4 PCs		**WT-MIA (SWT)** *Daubechies* 1, level = 3 4 PCs	
	T^2 distance	**Q distance**	**T^2 distance**	**Q distance**	**T^2 distance**	**Q distance**
BZN01	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%
BZN02	0.6%	0.6%	0.6%	0.7%	0.7%	0.6%
BZN03	0.9%	0.9%	0.8%	0.7%	0.8%	0.7%
BZD01	***3.6%***	***1.9%***	***5.1%***	***2.8%***	***4.1%***	***3.0%***
BZD02	***1.6%***	0.7%	***2.5%***	0.7%	***1.9%***	0.9%
BZD03	***1.9%***	0.9%	***2.4%***	0.7%	***2.1%***	0.8%

Table 2. BNDataset. Percentage of pixels above Hotelling's T^2 and Residuals Q distances critical limits based on normal image BNN01 95th percentile

	ct-MIA *w* = 2, 2 PCs		WT-MIA (DWT) *Daubechies* 1*,* level = 3 2 PCs	
	T^2 distance	**Q distance**	**T^2 distance**	**Q distance**
BNN01	5.0%	5.0%	5.0%	5.0%
BNN02	5.0%	6.0%	4.9%	5.7%
BNN03	4.9%	6.6%	4.4%	6.5%
BND01	6.2%	3.8%	6.6%	4.1%
BND02	4.5%	2.5%	5.3%	2.5%
BND03	5.2%	4.3%	6.3%	4.7%

Table 3. Bread Dataset. Percentage of pixels above Hotelling's T^2 and Residuals Q distances critical limits based on normal image N01 99th percentile

	ct-MIA w = 1, 6 PCs		**WT-MIA (DWT)** *Daubechies* 2, level = 5 6 PCs		**WT-MIA (SWT)** *Daubechies* 2, level = 5 6 PCs	
	T^2 distance	**Q distance**	**T^2 distance**	**Q distance**	**T^2 distance**	**Q distance**
N01	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%
N02	0.6%	1.0%	0.5%	0.7%	0.9%	0.2%
D04	2.7%	3.4%	3.3%	3.2%	11.9%	11.4%
D07	4.3%	3.0%	4.6%	4.3%	15.7%	14.3%