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

SCHOLARONE™
Manuscripts

1 **Coupling 2D-Wavelet decomposition and Multivariate Image Analysis (2D WT-**

2 **MIA)**

3

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

5 [1]Department of Chemical and Geological Sciences, University of Modena & Reggio

6 Emilia, Via G. Campi, 183, 41125 Modena, Italy

7 [2]Multivariate Statistical Engineering (GIEM), Departamento de Estadística e I.O.

8 Aplicadas y Calidad, Universitat Politècnica de València, Cno. de Vera s/n, Edificio 7A,

9 46022 Valencia, Spain

10

ABSTRACT

11

12 The use of 2D Discrete Wavelet Transform in the Feature Enhancement phase of

13 Multivariate Image Analysis is discussed and implemented in a comparative way with

14 respect to previous publications. In the proposed approach, all the resulting sub-images

15 obtained by Discrete Wavelet Transform decomposition are unfolded pixel-wise and

16 mid-level datafused to a Feature Matrix which is used for the Feature Analysis phase.

17 Congruent sub-images can be obtained either by reconstruction of each decomposition

18 block to the original pixel dimensions, or by using the Stationary Wavelet Transform

19 decomposition scheme. The main advantage is that all possible relationships among

20 blocks, decomposition levels and channels are assessed in a single multivariate analysis

21 step (Feature Analysis). This is particularly useful in a monitoring context where the

22 aim is to build multivariate control charts based on images. Moreover, the approach can

23 be versatile for contexts where several images are analysed at a time as well as in the

24 multispectral images analysis.

25   Both a set of simple artificial images and a set of real images, representative of the on-

26   line quality monitoring context, will be used to highlight the details of the methodology

27   and show how the wavelet transform allows extracting features which are informative of

28   how strong the texture of the image is and in which direction it varies.

29

30   **Keywords:** 2D Wavelet Transform, Multivariate Image Analysis, Multi resolution,

31   Quality monitoring

32   **1. INTRODUCTION**

33   The use of Multivariate Analysis to evaluate images dates to the mid-late 80's, with the

34   work of Esbensen and Geladi[1] who introduced the use of Principal Component Analysis

35   for the study of multi-channel images. Multivariate Image Analysis (MIA) has soon

36   gained boost with the application in many contexts, typically those with images of such

37   complexity that they could benefit of a multivariate analysis approach (e.g. remote

38   sensing[2-4] and medical imaging[5-9]). In the 90's, the pioneering works of MacGregor and

39   his group made the field of process industry accessible by the MIA approach[2, 10]. The

40   possibility to acquire images for on-line process monitoring purposes and effectively

41   analyse them represents a viable, PAT-like sensor to investigate process changes in time

42   in an environment, the process line, where often the room for installing new

43   "traditional" sensors is poor, not to mention the fact that often a single image can

44   acquire simultaneously several different potential sources of variability.

45   At present, several uses of MIA are reported in literature for different tasks[11-29], all of

46   which are characterised by being well described by the two main sources of information

47   an image can carry. Textural variability, which can be gathered by analysing the "two"

48   dimensions relationship structure of pixels and "spectral" properties variability, which is

49   based on the "third" dimension, that is the channels acquired for each pixel. The latter

50   aspect becomes the more relevant as the number of channels increases, moving from

51   simple binary or grey-scale images (where not much information can be given more

52   than texture homogeneity/non homogeneity), to RGB images (where changes in colour

53   can be related to the presence of non homogeneous texture or underlying phenomena

54   which alter the composition), to multi-channel images and spectral images (where

55   chemical information can effectively be acquired for each pixel). Therefore, most

56　image-based challenges, which can be addressed with MIA approach, represent the

57　detection of product defects in quality control[11-18], the monitoring of changes in process

58　behaviour and its feed-back control[16, 19-21], the prediction of product properties on the

59　basis of the joint evaluation of texture and channel information (in particular addressed

60　to by the development of Multivariate Image Regression – MIR – methods[22-23]); or

61　more recently the development of imaging biomarkers in cancer diagnosis[8-9].

62　As far as the core details of the MIA approach and its evolution in time are concerned,

63　the MIA approach proposed by Bharati and MacGregor is based on a framework[11-12]

64　which can be summarised in two main steps: a Feature Extraction (or Enhancement)

65　phase, and a Feature Reduction (or Analysis) phase. In the Enhancement phase the

66　image (pre-processed, if necessary) is treated so that texture information carried out by

67　the pixels is made clearer. In the Analysis phase, a suitable Multivariate Analysis

68　method is applied (e.g. Principal Component Analysis, Partial Least Squares

69　Regression, Partial Least Squares Discriminant Analysis) on the feature matrix obtained

70　after the first phase. The two phases are strictly connected to each other, since the first

71　step can strongly influence the outcome of the following analysis in a way which is not

72　much different from the effect of data pre-processing in many other situations.

73　However, a certain degree of freedom can be considered when choosing the feature

74　enhancement method (whilst the feature analysis phase is more application driven). The

75　fundamental aspect to be considered in this case is that it is not only important to

76　preserve the information given by the channels, for which a simple unfolding of the

77　image structure so that each pixel becomes a sample could be sufficient, but to retain the

78　correlation among neighbouring pixels (that is, the texture information) and, most of all,

79　to present it to the following analysis step in such a way that both sources of variability

80 (texture and channel-based properties) can be easily evaluated. In the approach

81 proposed by Bharati and MacGregor[2] the texture information is extracted by

82 augmenting column-wise the unfolded pixel vector with a series of its copies, shifted

83 row-wise so that each row of the generated matrix is formed by a pixel and all its

84 surrounding neighbours. This corresponds to stacking copies of the original image

85 shifted according to a given step. The number of neighbours, hence of columns, of the

86 feature matrix is $(2w + 1)^2$, governed by the window aperture parameter w, which

87 indicates the dimension of the window, centred on the pixel, encompassing the

88 neighbours to be considered (typically, $w = 1$ or 2). In Prats-Montalbán et al.[17] this

89 augmentation is extended to each channel of the image, and will be referred to from

90 now on as colour-textural MIA (ct-MIA). Facco et al.[14] proposed a method to reduce the

91 computational cost when operating with a larger window size, $w > 2$. Other approaches

92 have been proposed and discussed, among which the most common are based on the

93 application of a transformation of the image, again for each channel, in a different

94 domain, such as the Fourier domain (*via* the 2D Fourier Transform) or the wavelet

95 domain[10-12, 18, 24-29]. The wavelet advantage with respect to Fourier is that it has both

96 good frequency and spatial resolution.

97 There may be several advantages by moving to wavelet domain in terms of image

98 compression, background removal and denoising. Moreover, the use of wavelet

99 transform allows a better understanding of the pixels correlation structure at different

100 scales. At each level of decomposition, the coefficients carry both the information

101 pertaining to the energy which characterise a frequency range (based on the selected

102 filter) and an indication about the orientation in which varies (according to the type of

103 decomposition block, being it Approximation or one of the three Details blocks, namely

104   Horizontal, Vertical and Diagonal). In this way, the features extracted by wavelet

105   transform are a truly enhanced visualization of how strong, and in which direction, the

106   texture of the image varies. Literature differs in the way these features could be

107   expressed and handled. Some authors have pointed the attention to the use of global

108   indicators to synthesise the relevant information for a given decomposition level and

109   block, by using, e.g. the Frobenhius norm (Energy), the entropy, statistical momenta or

110   the standard deviation of the coefficients[18, 26-28]. In this way, a single variable

111   summarises the effect, while the orientation information is maintained by means of the

112   level-block combination at which it is computed. This approach surely reduces the

113   computational cost of the following analysis, but carries with itself the potential loss of

114   interesting information, which goes together with an "averaging" procedure of a richer

115   set of data. Also, since all the information of an image is compressed in a single vector

116   of descriptors for each decomposition block and level, a somehow conspicuous set of

117   images must be considered to create a reference set, for example of Normal Operating

118   Conditions (NOC), when moving to the following Feature Analysis phase. On the

119   contrary, working at pixel level, that is considering each pixel as a sample, opens the

120   possibility to reduce the requirements when building a reference set, often allowing the

121   use of a single representative image, being it a real one, or a combination of snapshots

122   of NOC texture areas.

123   Liu and MacGregor[10] have proposed an approach where the wavelet transform is used

124   for Feature Enhancement of images working at pixel level, i.e. the MultiResolutional

125   Multivariate Image Analysis (MR-MIA). MR-MIA is conjugated in two frameworks

126   that differ in the stage at which the wavelet transform is applied, i.e. before (MR-MIA I)

127   or after (MR-MIA II) the Feature Analysis step (in this case PCA). In particular, in MR-

128   MIA I by applying the discrete wavelet transform (2D DWT) to each channel of the

129   image, each decomposition block, at a given level, can be seen as an image itself with

130   the same number of channels, but representing a different "resolution" and texture

131   orientation. The Feature Analysis (e.g. PCA) step is then applied to each of these

132   images, once unfolded pixel-wise, so that as many latent variable models as number of

133   blocks per decomposition level (L) are obtained. This approach relies on the

134   orthogonality of the wavelet decomposition blocks, thus implying that there is no

135   interest in evaluating correlations among blocks at different scales, and the possibility to

136   evaluate texture – channel correlation is maintained. However, the complexity aspect of

137   this approach can be a hindrance when considering how many multivariate models one

138   should compute and the necessity of a high-level data fusion step where all the results

139   are combined to create a decision rule in order to e.g. decide if a new product image has

140   to be rejected when compared to the NOC modelled one. Recently, basing on similar

141   considerations, Juneau *et al.*[25] proposed an approach where all sub images obtained by

142   wavelet decomposition, once unfolded pixel-wise, are merged row-wise and analysed

143   by a single PCA. However, they used the continuous wavelet transform (undecimated

144   scheme, UWT) and in this way a rather large number of features is obtained, since scale

145   and shift parameters vary continuously.

146   The MSMIA approach proposed by Reis[18] is similar to the MR-MIA I, although more

147   images are considered at the same time as references NOC, but it differs in the way

148   information from the Feature Analysis step (e.g. PCA) of each WT decomposition block

149   is fused. In this approach, an index evaluating the distance to the scores distribution

150   histogram of the reference NOC images for each decomposition block, at a given scale,

151   is calculated in order to obtain a set of variables, which are then used for building a

152   monitoring chart. In addition to this, multivariate control charts based on PCA of

153   wavelet features (e.g. standard deviation of wavelet coefficients for each decomposition

154   block), extracted for each decomposition block, at a given scale, are also considered.

155   The approach is effective in compressing information and for on-line implementation,

156   however defects location requires a further step. This step, similarly to MR-MIA II,

157   consists in building a spatial shifting feature matrix (then analysed by PCA) for each of

158   the sub-images contributing to "out of control" observations in the preceding step.

159   Moreover, correlation structure among textural/colour pattern at different scales is only

160   indirectly taken into account (the information from different scales is always merged at

161   features, not pixels level).

162   Here we present an approach, which is named 2D WT-MIA, where the Feature

163   Enhancement step is similar to MR-MIA I, but as in Juneau[25] all of the resulting sub-

164   images obtained by the 2D-DWT decomposition are unfolded pixel-wise and mid-level

165   datafused to a Feature Matrix which is used for the Feature Analysis phase. In order to

166   have congruent sub-images all decomposition blocks are reconstructed separately to the

167   original pixel dimensions. This reconstruction step can be omitted, when the Stationary

168   Wavelet Transform (2D-SWT) is used[30-31]. In this way, all possible relationships among

169   blocks, decomposition levels and channels are assessed in a single multivariate analysis

170   step (Feature Analysis). This is particularly useful, in a monitoring context, when the

171   aim is building multivariate control charts based on NOC images. Thus, our proposed

172   approach can be versatile to handle contexts where several images are analysed at a time

173   as well as in the multispectral images analysis.

174   The rest of this paper is organised as follows: in Section 2: Methods, the proposed 2D

175   WT-MIA approach is described into more details and compared to colour-textural MIA

176   approach to highlight common and differing aspects; in Section 3: Materials, the images

177   datasets will be presented, consisting in a set of simple artificial binary images, used to

178   illustrate how texture is captured within the two approaches and a set of real images,

179   and in Section 4: Results and Discussion, these images will be analysed according to the

180   two-step MR-MIA framework, using Principal Component Analysis as Feature Analysis

181   technique with the target of simulating a quality control task.

182   **2. METHODS**

183   The approach here described belongs to the more general framework of MultiResolution

184   Multivariate Image Analysis, thus basing on a two-phase elaboration of the image: a

185   first step of Feature Extraction (Enhancement) and a second step of Feature Reduction

186   (Analysis) (Figure 1). In particular, the 2D WT-MIA (wavelet based feature

187   enhancement) and the colour-textural MIA (spatial shifting based feature enhancement)

188   approaches will be discussed and compared in terms of results in the present work.

189   2.1 Spatial Shifting Feature Enhancement

190   Colour-textural MIA[17] is summarised in Figure 2. This approach to Feature

191   Enhancement consists in capturing, for each channel ch, the pixel proximity correlation

192   by means of a spatial shifting of neighbouring pixels with respect to each pixel of the

193   original image, according to a selected window aperture parameter, w. In practice,

194   starting from a pixel element of the image $p_{i,j}$, a row vector is created by adding the

195   intensity value of the channel corresponding to the closest surrounding pixels: if $w = 1$,

196   the composition appears as reported in Figure 2. When this is done for all the pixels of a

197   pixel-wise unfolded channel matrix, and the feature matrices for the different channels

198   are then fused, a total Feature Matrix is obtained which has as many rows as the number

199　of pixels $I = n_1 \times n_2$, and as many columns as the number of channels *ch* times the

200　number of spatial shifted pixels, which is given by $(2w + 1)^2$. This means that,

201　regardless of the number of channels, for a window parameter $w = 1$ (the closest

202　neighbours) the Feature matrix is $9 \times ch$, and when moving to $w = 2$ (the closest

203　neighbours and the next surrounding layer), the Feature matrix is $25 \times ch$. This implies

204　a fast increase of the number of variables considered in the analysis, the higher is the

205　number of channels.

206　Since we need, at each pixel location, to use all the neighbouring pixels up to a distance

207　*w*, this implies that we lack information for all those pixels in the borders with width *w*.

208　Therefore, the solution commonly adopted is to diminish the size of the image from

209　$n_1 \times n_2$ to $(n_1\text{-}2w) \times (n_2\text{-}2w)$

210　2.2 2D Wavelet-based Feature Enhancement

211　Figure 3 shows the general scheme of the Feature Enhancement step involving 2D –

212　DWT application, through the fast Mallat algorithm[32-33], on an image. For each channel

213　*ch*, the low-pass and high-pass filters (which are the same as in the 1D case) are first

214　operated row-wise on the image and then, after downsampling of the coefficients, in

215　each of the resulting blocks column-wise. In this way four decomposition blocks are

216　obtained: Approximations (low-pass + low-pass), namely CA; Horizontal details (low +

217　high), namely CH; Vertical details (high + low), namely CV, and Diagonal details (high

218　+ high), namely CD. The procedure is then iterated by applying it to the

219　Approximations, i.e. increasing the decomposition level. Downsampling is skipped

220　when the 2D - SWT scheme is used since, instead, the filters are up-sampled[26]. The

221　maximum possible decomposition level, L, depends on the image size. The four

222   decomposition blocks obtained from each level of decomposition (CA, CH, CV and

223   CD) when 2D DWT is used are independently reconstructed by means of the inverse 2D

224   – DWT so that their dimensions are the same of the original image, while they are

225   already of the same size when 2D SWT is used (in fact, each block of coefficients at

226   every level maintains the same size as the original image, and congruent images are

227   obtained). This leads to a set of $4 \times L$ images for each channel ch, which can be

228   unfolded and column-wise merged to obtain a total Feature Matrix which has as many

229   rows as the number of pixels $I = n_1 \times n_2$, and as many columns as $4 \times L \times ch$. If we

230   compare this column dimension to the one obtained with the Spatial Shifting approach,

231   which is $(2w + 1)^2 \times ch$, it might appear that there is little benefit in terms of reduction

232   of the Feature Matrix dimensions. However, two aspects have to be underlined:

233   i) in the spatial shifting approach the image is analysed by moving a $(2w + 1)x(2w + 1)$

234   pixels window by step of 1 in all image directions; on the other hand, with wavelet, a

235   *filter length x filter length* pixels window is moved by step of 1 in all image directions,

236   but using a larger filter does not increase the number of features, which remain always

237   four;

238   ii) the two approaches lead to the same number of feature descriptors if $L = \text{round}[(w +$

239    $\frac{1}{2})^2]$. This corresponds, for e.g. a window parameter of $w = 2$, to a decomposition level

240    $L = 6$, which in terms of multiresolution means to have gone very deep in the analysis of

241   coarse and smooth aspects of the image. In other words, such a decomposition level (if

242   allowed by the nature of the chosen wavelet) usually leads to the possibility of

243   evaluating correlations and high distance relationships among pixels to an extent, which

244   is superior to the use of a moving window of fixed size.

245   When applying the wavelet transform, the selection of the most appropriate wavelet

246   filter is considered a critical issue and a limiting step in the implementation of routine

247   applications (i.e. which wavelet family and which filter length, to analyse the specific

248   characteristics of the images at hand). This issue has been dealt in literature by

249   analysing the different properties of the decomposition filter in terms of texture

250   description capability in order to propose general criteria[34] or focusing on goodness of

251   image reconstruction[33], or proposing a design of experiments approach[36]. We recently

252   proposed[37] a methodology based on N-way modelling to provide a range of possible

253   wavelet choices (in terms of families, filters, and decomposition levels), for each image

254   and problem at hand. Any of these strategies require a preliminary analysis step to be

255   conducted by experienced people in the field, although this step is only required once in

256   model building. However, some considerations, based on our experience can be drawn:

257   i) there is in general a relationship between the decomposition level and the filter length,

258   i.e. by using a larger filter a lower decomposition level is required to capture the

259   different image aspects (coarse and smooth) and ii) taking into account the wavelet

260   families characteristics, such as degree of symmetry or regularity or number of

261   vanishing moments[38-39] it is possible to focus on a small number of wavelet filters to

262   test, by choosing a representative one for each type of property.

263

264   **3. MATERIALS**

265   3.1 Artificial Images datasets

266   These sets are used to illustrate how the colour-textural MIA and 2D WT-MIA

267   approaches analyse texture and their capability to detect faulty pixels. These images are

268 characterized by two main features: a particularly limited pixel size, so that

269 computational time is not a relevant benchmark property at this stage, and a simple, yet

270 well defined, pattern. Also, the differences between "normal", i.e. reference image, and

271 "defective", i.e. image (or images) for which a perturbation of the pattern was created,

272 are well controlled, in the sense that the number and position of pixels which have been

273 changed is known, and the entity of the disturb is enough to obtain simulated test

274 images which are not too similar to their reference image. In spite of the simplicity of

275 this simulated case, the information which can be acquired from the analysis with both

276 approaches is interesting to better understand how the two methods under comparison

277 work, and the conclusions which can be drawn are helpful and can be extended, as

278 shown in the next section where real images are presented and dealt with, to cases of

279 higher complexity.

280 The set is composed by three binary images, as reported in Figure 4, of size 32 × 32

281 pixels. Figure 4 "SimSetA" reports the "normal" (reference) image, on the basis of

282 which an alternation pattern has been generated. In this case, the squares which alternate

283 in both image directions to give a chequered pattern have a dimension of 8 × 8 pixels:

284 starting from upper left corner and moving over columns dimension, a white (1's) 8 × 8

285 pixel square is alternated to a black (0's) 8 × 8 pixel square, and the same alternated

286 pattern is repeated over the rows dimension. Starting from this image, two changes in

287 pattern were produced, leading to two "defective" (test) images. Figure 4 "SimSetB"

288 shows an overlying irregular shape which extends from the diagonal to the lower left

289 part of the image: for this figure, a total of 55 pixels have been inverted in value (from 1

290 to 0 or from 0 to 1) over the total of 32 × 32 = 1024 pixels. Figure 4 "SimSetC" shows

291 another change in the pattern, this time according to a regular shape which is applied on

292 top of each of the 8 × 8 pixel squares: for each of these squares, starting from the second

293 diagonal element, a single pixel every fourth has been modified both in the rows and in

294 the columns, thus resulting in a change of four pixels for each of the squares. For this

295 figure, a total of 64 pixels have been inverted in value (from 1 to 0 or from 0 to 1) over

296 the total of $32 \times 32 = 1024$ pixels.

297 3.2 Real Images datasets

298 To further explore the performance of the method proposed in this work and compare it

299 to the results of the ct-MIA approach, additional datasets have been taken into account,

300 belonging to different applicative contexts, tiles and bread production, respectively. In

301 both cases, the control of the final product undergoes visual inspection, while the

302 datasets differ as for image dimensions and number of channels.

303 3.2.1 Tiles

304 These datasets come from a production of tiles of marble-like materials for surface

305 coverage: all the cases share a common issue, that is presenting product samples which

306 do not comply to a strict definition of "normal" images, characterized, for instance, by a

307 precise colour shade or by the absence of defects such as spots and scratches. Therefore,

308 it is necessary to develop a method, complementary or alternative to visual inspection,

309 which is able to: a) recognize the presence of a defectiveness when a new tile is

310 compared to the reference one(s); b) indicate the kind of defectiveness (e.g. colour

311 shade and/or presence of unwanted changes in surface pattern); c) locate on the surface

312 the position of the defect in order to obtain an enhanced perception of the same, so that

313 its visualization and recognition by the operator is made easier. Samples from two

314 different products were considered with different degree of irregularity of the pattern in

315   the defective tiles. They both consists of RGB images of dimensions 256 × 256 pixels

316   (Figure 5). Figure 5a reports dataset 1: Blanco Zeus, from now on referred to as

317   BZdataset, which is composed by three reference images (BZN01, BZN02 and BZN03),

318   and three images of tiles showing defects (BZD01, BZD02, and BZD03). This kind of

319   tile shows a mostly homogeneous shade of grey all over its surface, so that defects (as

320   for instance white or dark spots and scratches) do not usually present particularly high

321   difficulty of detection also by visual inspection. Figure 5b reports dataset 2: Blanco

322   Norte, from now on referred to as BNdataset, which is composed by three reference

323   images (BNN01, BNN02 and BNN03), and three images of tiles showing defects

324   (BND01, BND02, and BND03). In this case, the tile main colour is grey, but the surface

325   is characterized by an inhomogeneous distribution of darker spots, in a grainy structure,

326   which makes quite difficult to detect the presence of defectiveness, both when

327   represented by darker and paler areas.

328   3.2.2 Bread

329   This data set comes from an industrial production of bun bread, where a digital scanner

330   is already used to automatically assess defects concerning mainly bun dimensions,

331   while surface defectiveness, such as dark spots, blisters, and pale areas is still evaluated

332   by visual inspection of expert personnel. These defects arise from different causes, some

333   of which not perfectly known, and are also often difficult to be detected by RGB online

334   cameras. Thus, a feasibility study has been undertaken[29] by acquiring offline

335   multispectral images, covering the UV-visible range (from 430 to 700 nm, 10 channels)

336   and the short-wavelength NIR range (from 850 to 970 nm, 8 channels), which can

337   improve the acquisition of information on bread quality, combining spectral (NIR may

338   represent also a "chemical signature") and textural information. The whole data set has

339 been analysed by WT-MIA (DWT scheme) approach and described in detail in ref. 33

340 while here a subset of images has been analysed in order to discuss comparatively the

341 performance of WT-MIA (DWT and SWT) and ct-MIA.

342 The raw images were cropped to remove the distortion effect of the round bun shape,

343 background and noise were removed *via* preliminary wavelet analysis[29], finally giving

344 images of dimensions of about 387 x 420 pixels for 18 channels. Here two non-

345 defective images (N01, used as reference, and N02) and two defective ones (D04 and

346 D07) are analysed, shown in Figure 6.

347

348 **4. RESULTS AND DISCUSSION**

349 4.1 Artificial Images datasets

350 All the three images (SimSetA, SimSetB and SimSetC) have been treated according to

351 the same Feature Enhancement step, by considering:

352 - Spatial Shifting, colour-textural MIA (ct-MIA) with window size parameter w =

353 1

354 - Wavelet Decomposition (WT-MIA) by using a Daubechies 1 (db1) at

355 decomposition level L = 1, both DWT and SWT.

356 The Feature Enhancement step gave a Feature Matrix of dimensions I = 900 rows

357 (reduction from 32 x 32 to 30 x 30 is necessary to cope with borders) and 9 columns for

358 the ct-MIA approach and I = 1024 rows x 4 columns for the WT-MIA approaches.

359 SimSetA was used as the reference set, upon which for the ct-MIA approach a Principal

360 Component Analysis model was obtained after mean centring of the Feature Matrix.

361   Figure 7 reports the PCA results, from left to right in the order are shown: scores image

362   SimSetA, loadings, projected scores images SimSetB and SimSetC (in the order PC1 to

363   PC4 from top to bottom). PC1 captures, for the reference ("normal") image both the

364   difference in grey intensity value (colour) and the variation in pattern (texture) when

365   passing from the pixels having zero value to pixels with value one, i.e. it shows the

366   change of value when moving along the borders from one square to another, where pixel

367   values invert, leading to a "blurring" effect of the borders. This is the expected effect

368   since ct-MIA window of one (which is actually 3×3 pixels) moves pixel by pixel on

369   the image structure, which is made of 16 squares of dimensions 8×8.   All features

370   contribute similarly to the PC1 loadings since the pattern change takes place in all

371   directions. Thus, PC1 works as an average grey scale image, which in fact extracts out

372   the spectral information (we have no other source of spectral information than a single

373   grey scale channel). The following PCs capture only the borders effects, i.e. only the

374   frames around the squares are visible in the scores images, and by inspecting the

375   loadings it is possible to understand the directions of the pattern variation, e.g. to PC4

376   the features accounting for diagonal shift do not contribute.

377   When the Feature Matrix corresponding to SimSetB and SimSetC are projected onto

378   this model, the same chequered pattern is correctly reproduced (Figure 7), but the

379   changes in pixel correlations due to the small scale modifications of its regularity

380   produce a large blurred area, which roughly encompasses the whole shape of the

381   differences but extends further with respect to the faulty pixels, i.e. an area of about 3×3

382   pixels around each defective pixel as it is detailed in the following text. This is due to

383   the fact that the perturbation, although being well defined (in particular for SimSetC) to

384   a small number of pixels, influences the neighbouring correlation structure of all the

385   pixels, which are contained by the moving window.

386   In a monitoring context the defective images with respect to the reference one/s can be

387   identified by the Hotelling-T^2 and squared residuals (RSS, SPE or Q) multivariate

388   control charts by using the percentage of pixels beyond control limits[16]. However, in

389   this case being the images binary, simply the pixel by pixel difference of the residuals

390   sum of squares (RSS) of the test images with respect to the reference image (NOC) can

391   be used. Both the RSS from a one or a four components PCA model are suitable to

392   depict the faulty pixels for SimSetB and SimSetC, but also an area of about 3 x 3

393   around each faulty pixel will show up differing in RSS values with respect to NOC

394   (Figure S2, supplementary material). This can be expected on the basis of the

395   considerations made above on the neighbouring pixels correlation structure.

396   In the WT-MIA both decomposition schemes, DWT and SWT, have been applied,

397   considering the simulated pattern, i.e. inversion of the binary value of some not

398   consecutive pixels, db1 seems an appropriate filter. The feature matrix, holding the four

399   decomposition blocks CA, CH, CV and CD (reconstructed only in DWT case), already

400   captures the texture pattern, as highlighted in Figure 8 (DWT) and Figure 9 (SWT)

401   where the sub-images corresponding to each block of the DWT and SWT

402   decomposition of SimSetA and SimSetB are reported. This is a first difference with

403   respect to ct-MIA approach where the feature matrix holds just the shifted version of the

404   raw image in all possible neighbouring direction (as shown in Figure S1, supplementary

405   material) and thus PCA (or, more in general, a multivariate decomposition technique) is

406   needed to reveal the texture pattern. Further, in this case with only one reference image

407 and one channel the feature analysis step by PCA is not needed at all (the decomposition

408 blocks are orthogonal).

409 In the DWT the Approximation Block is the only one which carries the structural

410 information of image SimSetA (Figure 8, top). This is explainable by considering that

411 the db1 filter is of length two, thus operates like a window of pixel size 2×2 which

412 moves at steps of one, and due to the down-sampling scheme of DWT only one

413 coefficient every two is retained. Thus, since the binary values change every 8 x 8

414 pixels, there is no blurring effect at the squares edges; also the coefficients in all the

415 other decomposition blocks are zeros (Figure 8, top). On the contrary, the presence of

416 deviations in the two test images, related to "sharper" structures i.e. alternating by one

417 pixel, is well captured by all decomposition blocks (Figure 8, middle). In particular

418 Approximation (CA) shows both intensity change and texture, while Details blocks

419 capture horizontal (CH), vertical (CV) and diagonal (CD) neighbouring pixels

420 alternation of binary values. Thinking of a monitoring context, in this case the defects

421 can be depicted by the difference between the decomposition sub-images of the

422 reference and test image, as shown in Figure 8, bottom; to this aim, considering the

423 specific pattern of the defects in SimSetB, CA and CD are the most suitable blocks.

424 Figure 9 shows the results of the SWT decomposition. Similar considerations can be

425 drawn. The only difference is that now the effect of the variation in the binary values of

426 the pixels at the edges of the 8 x 8 pixels squares are visible (similarly to ct-MIA). This

427 is well explainable by the fact that in SWT down-sampling of the coefficients is not

428 operated (it is worth mentioning that the window size and the moving step remain the

429 same). Analogous considerations hold for SimSetC decomposition (figure not shown for

430 sake of brevity).

431   The behaviour of a larger filter, i.e. Daubechies 2 (db2) of length 4, has been also

432   inspected by using the SWT scheme on SimSetB (Figure S3, supplementary materials).

433   The db2 operates as a 4×4 window moving at steps of one pixel: the result is similar to

434   the one obtained by ct-MIA, leading to a blurring of the borders among squares and

435   around the area (of wideness about 3 x 3) which is interested by the defect.

436   Finally in Figure 10 the performance of ct-MIA and WT-MIA (SWT, db1) are compared

437   in terms of capability of detection and localization of the faulty pixels for SimSetB

438   (Figure 10a) and SimSetC (Figure 10b), respectively. The detection is good in both

439   approaches being all the faulty pixels correctly identified, the difference is in the

440   blurring area, which is strictly connected to the wideness of the analysing window, i.e. 3

441   x 3 for ct-MIA and 2 x 2 for db1. This is a general known advantage of WT of being

442   more efficient for feature enhancement because of the availability of several filter

443   shapes and length compared to spatial shifting approach.

444

445   4.2 Real Images datasets

446   4.2.1 Tiles

447   Several wavelet filters, belonging to Daubechies (filter length from 1 to 5), Symlet

448   (filter length from 1 to 5), Coiflet (filter length form1 to 3) and biorthogonal (1.3 and

449   1.5) families were tested (decomposition levels from 1 to maximum), by using an

450   approach as described in ref. 30. For both BZdataset and BNdataset Daubechies filter

451   length 1 (db1 or Haar) resulted among the best performing wavelet filters and we report

452   results relative to this filter, at decomposition level $L = 3$. While for the ct-MIA

453   approach window size 1 and 2 were considered, better performance was obtained with w

454　= 1 for BZdataset and w = 2 for BNdataset. This led, considering the three RGB

455　channels, to an unfolded feature matrix of size 65536 × 27 (w = 1), or 65536 x 75 (w =

456　2) in the ct-MIA case, and 65536 × 36 in the WT-MIA case.

457　In both datasets, a single reference image has been used to calibrate the PCA models

458　and build the Hotelling's T^2 and Q statistics (control charts). Autoscaling pretreatment

459　gave for both datasets and approaches the best results.

460　The choice of model dimensionality, i.e. number of principal components, in this

461　context cannot be automated, i.e. assessed on the basis of a priori fixed criterion, since it

462　is problem dependent[40]. General guidelines that we adopted in this work consist in: i)

463　inspecting how spatial features of the image are accounted for in scores images and ii)

464　scree-plot to ensure the systematic variation is modelled. Further, when enough defects

465　images are available, to preserve some for model validation, few can be used to see

466　which are the components that maximize detection capacity. It is worth noticing that

467　minimizing the squared prediction error in cross validation, as most used in PCA

468　modelling, is not appropriate in this context, because it is not necessarily related to the

469　capability of fault detection which is the objective pursued in process monitoring.

470　Image BZN01 has been used as reference NOC image for BZdataset. The PCA model

471　dimensionalities were 4 PC's for both approaches ct-MIA (captured variance 77%) and

472　WT-MIA (captured variance 44%), which correspond to a number of components each

473　explaining more than 1% variance (ct-MIA) and to the first minimum in the scree-plot,

474　i.e. number of components vs. eigenvalues plot, (WT-MIA), respectively. All the

475　remaining images of the dataset were projected onto the models and distances were

476　calculated. Table 1 reports the results in terms of percentage of pixels scoring above the

477 critical limits, which were chosen on the basis of the reference image by obtaining the

478 99th percentile values of its distances distributions. Both models are able to accept as

479 normal behaving images such as BZN02 and BZN03, which are actually defectless, and

480 indicate, especially in terms of T^2 distance, the presence of anomalies on all of the three

481 defective tiles, BZD01, BZD02 and BZD03; albeit the results are quite similar, a higher

482 percentage of pixels above the critical limits is detected by the WT-MIA approach. Both

483 approaches show similar results, although the WT-MIA identification of defects appears

484 better defined, especially for image BZD02 where more clusters of pixels are identified,

485 which are in particular connected to the presence of darker spots on the surface of the

486 tile, especially when using SWT (Figure S4, supplementary Material). SWT monitoring

487 results are also shown on Table 1 and are very close to DWT ones.

488 Interpretation of the features enhancement step can be gathered by loadings analysis. ct-

489 MIA loadings are shown in Figure 11 (left) both as bar plot (top left) and refolded

490 (bottom left) in the corresponding position of neighbours window (the central pixel is

491 the pixel itself). As usual, PC1 is gathering an (approximately) average colour effect (all

492 loadings have the same sign). Moreover, it can be observed that colour intensity varies

493 left to right for red and blue channels, while green is more uniform; similarly does PC1

494 of WT-MIA (bar plot, top right, and decomposition sub-images, bottom right), to which

495 the Approximations of all levels and channels contribute (Approximations in fact act as

496 an averaging tool at each decomposition level, hence extracting out the same

497 phenomenon as ct-MIA). Also the WT-MIA Approximations sub-images (Figure 11,

498 bottom right) highlight the varying intensity from left to right, especially for

499 decomposition levels two and three (the green channel is uniform at level 1). This effect

500   may be due to illumination and eventually (not the aim here) it could be easily removed

501   in WT domain, e.g. by suppressing level 2 or 3 approximations as data pre-treatment[29].

502   PC2 and PC3 show the main contrast in horizontal and vertical directions, respectively

503   both for ct-MIA and WT-MIA (for PC2, the Horizontal details of level 3 for all channels

504   are the most relevant, and for PC3 the vertical details, level 1 opposite to level 3). PC4

505   shows a mixed pattern, loadings sign and values vary in all directions for ct-MIA and

506   for PC4 in WT-MIA the vertical details of level 2 are the most relevant.

507   It can be noticed that the possibility to analyse the images at different resolution (the

508   different decomposition levels) enhances the colour-textural pattern recovery, with

509   respect to ct-MIA where only the neighbouring window size can be varied (that in WT-

510   MIA can roughly corresponds to the filter length/family).

511   *Table 1 to be inserted about here*

512   As reference NOC image for BNdataset, image BNN01 has been used, the PCA model

513   dimensionalities were 2 PC's for both approaches ct-MIA (variance captured 39%) and

514   WT-MIA (variance captured 26%), which correspond to the first minimum in the scree-

515   plot. All the remaining images of the dataset were projected onto the models and

516   distances were calculated. Table 2 reports the results in terms of percentage of pixels

517   scoring above the critical limits, which were chosen on the basis of the reference image

518   by obtaining the 95th percentile values of its distances distributions.

519   *Table 2 to be inserted about here*

520   Neither of the models appear particularly satisfactory, since the normal behaving images

521   BNN02 and BNN03, which are defectless, appear to have Q distances higher than 5%.

522   The defective tiles, BND01 and BND03 appear above limits for both models, according

523   to T^2 distance statistic. On the contrary, defective BND02 is only detected by WT-MIA,

524   T^2 distance, albeit close to the limit.

525   By considering the T^2 distance values reshaped in the original pixel domain it is

526   possible to identify the groups of pixels which have distances higher than the critical

527   values. Figure 12a) and 12b) shows the comparison of defective images and the normal

528   images, with the corresponding distance images for ct-MIA and WT-MIA (DWT). The

529   WT-MIA identification of defects appears better defined, while ct-MIA seems to find

530   fewer clusters of pixels and more darker, well separated, spots all over the surface.

531   In a monitoring context, the results of Table 2 would indicate products BNN02 and

532   BNN03 as defective (false negatives) and shed doubt on rejecting or not product

533   BND02. On the other hand the possibility to look at above limits T^2 distance images (or

534   in general to the images corresponding to the above limit statistic) may clarify if defects

535   are present or not. In particular, this is a case were the defects are mainly due to a non

536   uniform distribution of pixels with a given colour content and texture that if normally

537   distributed on the image, as in the case of BNN01, BNN02 and BNN03, would be

538   acceptable. In this situation, it may be useful to calculate and represent the local

539   entropy[41] of the scores images, were the defective area is a region of low entropy

540   encircled by high entropy values, as shown in Figure 13.

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

542

543   4.2.2 Bread

544   The Daubechies 2 (db2) wavelet filter was used up to decomposition level 5 and for

545   both DWT and SWT decomposition schemes. The feature data matrix results of

546 dimension I$_{pixels}$ x 360 (4 blocks x 5 levels x 18 channels). In ct-MIA both a window

547 size of 1 (162 features) and 2 (450 features) were tested. Since results were similar we

548 will discuss the ones corresponding to $w = 1$, which gave a better defects localization.

549 The reference PCA model for non-defective image has been calculated by considering

550 as feature matrix the one obtained for N01 image (Figure 6). The PCA model refers to

551 mean centred data and model dimensionalities were 6 PC's for both approaches ct-MIA

552 (variance captured 66%) and WT-MIA (variance captured 52%), which correspond to

553 reaching the plateau in the scree-plot. We tested also a model made on two NOC images

554 but the results were analogous. Q and T^2 statistics were computed, and the critical limits

555 for each of the two statistics were computed on the basis of the 99th percentile. The

556 total percentage of pixels exceeding the critical limits is reported in Table 3. For all

557 approaches a clear detection of the two defective images can be obtained, with relevant

558 percentages of pixels above the critical limits for both distances, as well as N02 being

559 defectless.

560 However, when the Q and T^2 values above the reference limits are refolded to the

561 original pixel x pixel domain to locate the defective areas on the image (Figure 14),

562 differences among the approaches emerge. ct-MIA is less efficient to detect the

563 defective area for D07 and for D04, it is also worth noticing that ct-MIA provides these

564 results when applied on the pretreated images, i.e. after denoising and background

565 removal with WT; otherwise it detects as faulty only pixels on the borders of the image.

566 DWT seems more efficient than SWT to locate the faulty pixels, notwithstanding the

567 fact that the same wavelet filter and resolution have been used.

568 Now focusing on the WT-MIA DWT results, it is worth noticing that not only the

569 stains, which are also easy to detect visually, but also the blisters and tiny scratches can

570 be detected. Moreover, we can assess which features are responsible of the defects by

571 inspecting the T^2-contributions, which can be interpreted in terms of the spectral

572 channels. In particular, Figure 15 shows the T^2-contributions for some of the blisters. To

573 make the representation clearer distinct plots are made for each decomposition block,

574 and each decomposition level is represented as a distinct line, so that the x-axis reports

575 just the channels (wavelengths): the main contributions are from approximations

576 decomposition levels 1-3. Interestingly, besides the visible channels, some of the NIR

577 ones (11th to 18th corresponding to the range from 850 to 970 nm at 20 nm resolution)

578 contribute, which point to carbohydrate, fat and water bands. This may indicate a

579 segregation of some of the ingredients on surface spots where blisters appear.

580

581 **CONCLUSIONS**

582 The artificial image datasets allowed highlighting the distinct way in which textural

583 information can be recovered by the ct-MIA and WT-MIA approaches: both are

584 efficient in depicting the salient pattern of the images and the area were the defects are

585 located. The main distinctive characteristics of the two methods are:

586 i) the feature matrix obtained by ct-MIA just holds the shifted version of the raw image

587 thus always requires coupling to a multivariate decomposition technique to highlight

588 textural patterns while the feature matrix obtained by WT-MIA already captures it;

589 ii) in general WT-MIA is more efficient for feature enhancement because of the

590 availability of several filter shapes and length compared to the spatial shifting approach

591 where only the window size can be varied.

592   The analysis of the tiles data sets reveals a similar behaviour of the two considered

593   approaches although identification of defects appears better defined with the WT-MIA

594   approach. Also both decomposition schemes DWT and SWT show similar performance.

595   In a monitoring context it is worth noticing that when the defects are due to a non

596   uniform distribution of pixels, whose colour content and texture if normally distributed

597   on the image would be instead acceptable, further image analysis tools (e.g. local

598   entropy or any other to assess homogeneity or heterogeneity of pixels distribution), on

599   the beyond Q or T^2 limits images, are required to avoid false negative to be detected.

600   In the analysis of multispectral images (bread data set) the WT-MIA approach

601   performed better and it was possible to highlight the full benefit of the proposed

602   approach from both the correct defects identification/location and interpretation in terms

603   of spectral features point of view.

604   A further remark is that the proposed WT-MIA approach is rather straightforward

605   requiring only the Feature Extraction (Enhancement) and Reduction (Analysis) steps, as

606   in ct-MIA; one or more NOC images can be analysed at the same time and assembled in

607   the same WT features matrix which is organized pixels wise, thus allowing defect

608   localization directly. Images denoising and background removal can be as well

609   accomplished at WT decomposition stage.

610   The proposed WT-MIA approach can be as well applied to hyperspectral images, the

611   bread data set being an example limited to eighteen channels. However, the

612   computational costs will be a limiting factor and further strategies could be considered

613   to render it more efficient, work is in progress in this direction.

614

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

711   Chemistry, (ed). Elsevier Science B. V.: Amsterdam, 2000: 77-84.

712 39. Misiti M., Wavelet Toolbox User's Guide, MathWorks: Massachusetts, 1996.

713 40. Camacho J., Ferrer A., Cross-validation in PCA models with the element wise k-fold (ekf)
714   algorithm: practical aspects. Chemom Intell Lab Syst, 2014; 131: 37–50.

715 41. Gonzalez R.C., Woods R.E., Eddins S.L., Digital Image Processing Using MATLAB, New
716   Jersey, Prentice Hall, 2003.

717

718   **Table 1**. BZDataset. Percentage of pixels above Hotelling's T^2 and Residuals Q

719   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%

720

721

722 **Table 2**. BNDataset. Percentage of pixels above Hotelling's T^2 and Residuals Q

723 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%

724

725

726

727

728   **Table 3**.   Bread Dataset. Percentage of pixels above Hotelling's T^2 and Residuals Q

729   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%

730

731

732

733
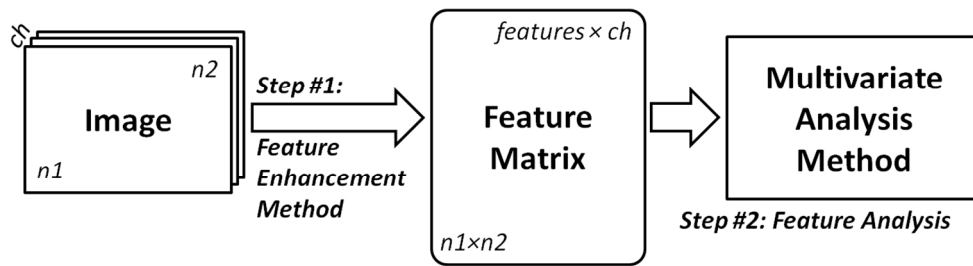
Figure 1: Feature Extraction (Enhancement) and Reduction (Analysis) steps in Multivariate Image Analysis (MIA).

381x102mm (96 x 96 DPI)
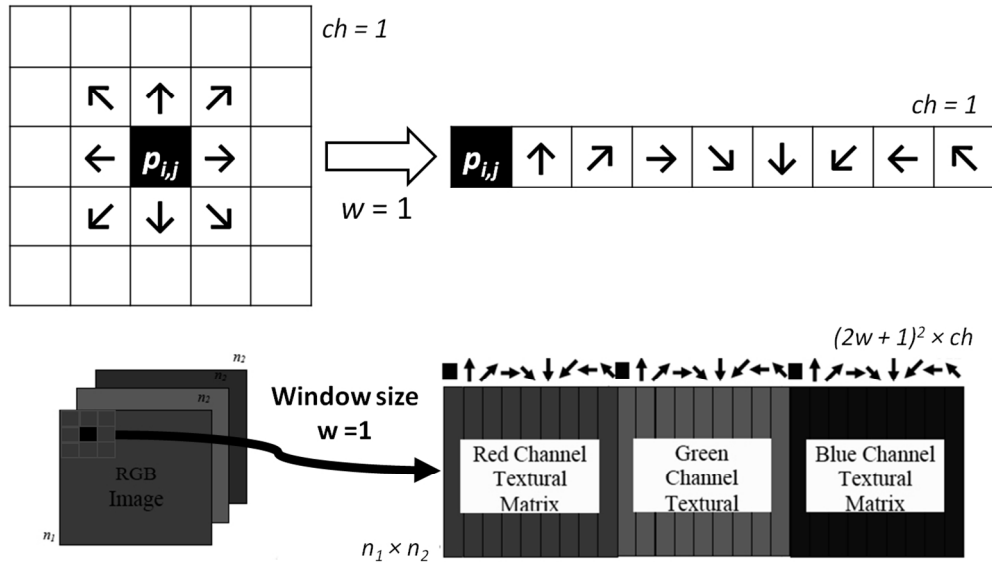
Figure 2: The Colour-textural MIA approach (ct-MIA)

389x222mm (96 x 96 DPI)
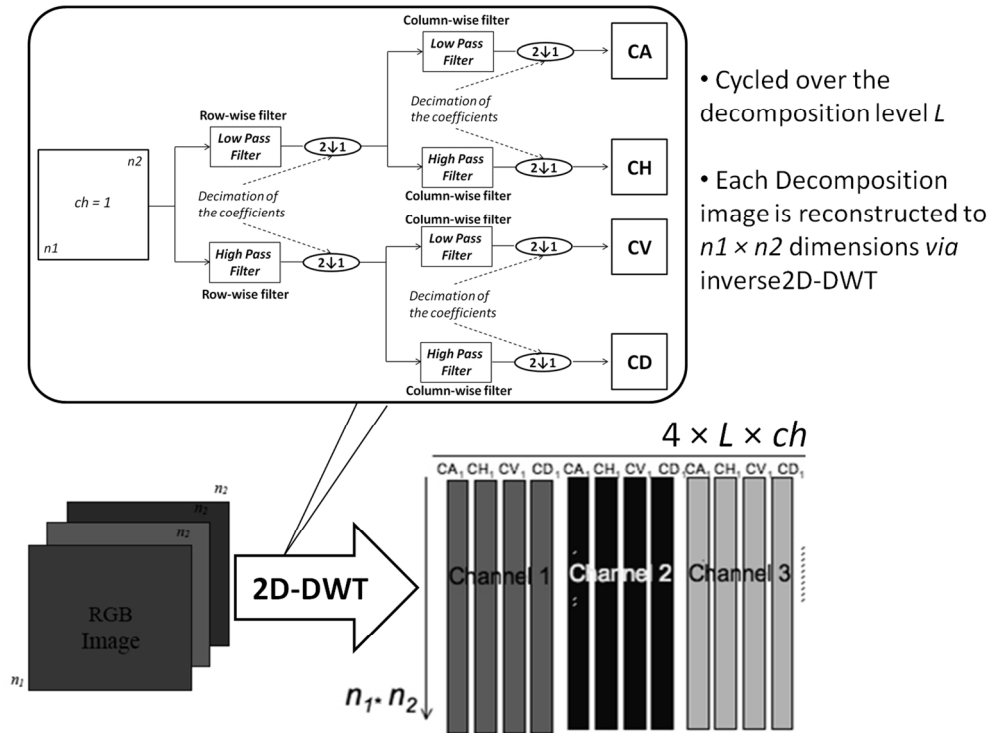
Figure 3: The Feature Extraction (Enhancement) step in the 2D WT-MIA approach, illustrated for an RGB image. The insert on the top of the figure shows the 2D DWT decomposition scheme at the first level of decomposition. CA, CH, CV and CD stand for Approximation, Horizontal details, Vertical details and Diagonal details coefficients, respectively.

395x290mm (96 x 96 DPI)

Figure 4: The SimSetA, SimSetB and SimSetC images.

376x127mm (96 x 96 DPI)

Figure 5: top (5a): Blanco Zeus (BZdataset) images; bottom (5b) Blanco Norte (BNdataset) images.

361x270mm (72 x 72 DPI)

Figure 6: The Bread dataset images

395x296mm (72 x 72 DPI)

Figure 7: The features reduction (Analysis) step for ct-MIA approach. PCA model built on SimSetA from left to right in the order: scores image, loadings, projected scores images from SimSetB and SimSetC, respectively.

361x270mm (72 x 72 DPI)
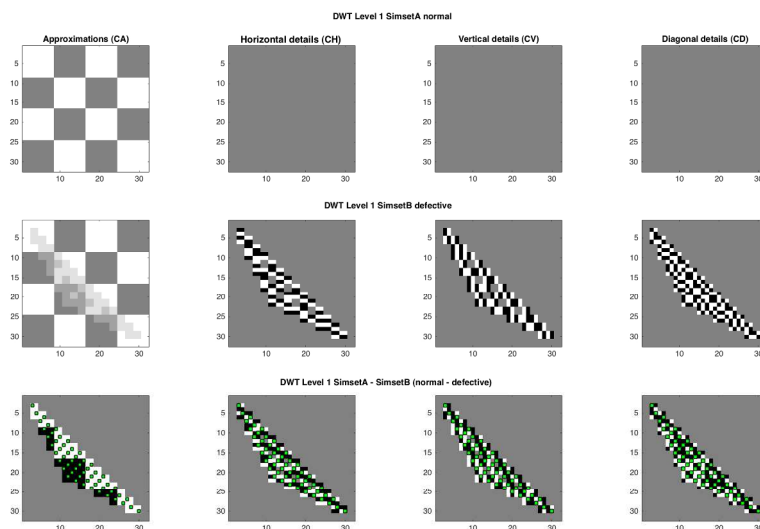
Figure 8: 2D WT-MIA approach. Decomposition blocks for level 1 by using Daubechies 1 (db1) wavelet filter and DWT scheme. In the order, from left to right, Approximations (CA), Horizontal (CH), Vertical (CV) and Diagonal (CD) details, respectively. Top) SimSetA; middle) SimSetB and bottom) difference (pixel by pixel) between SimSetA and SimSetB; small squares (green on on-line version) highlight location of the defective pixels.

903x528mm (72 x 72 DPI)
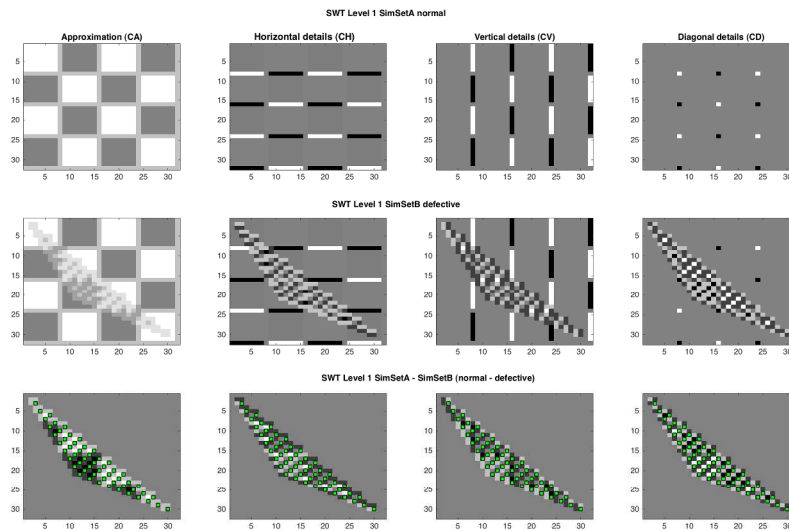
Figure 9: Figure 8: 2D WT-MIA approach. Decomposition blocks for level 1 by using Daubechies 1 (db1) wavelet filter and SWT scheme. In the order, from left to right, Approximations (CA), Horizontal (CH), Vertical (CV) and Diagonal (CD) details, respectively. Top) SimSetA; middle) SimSetB and bottom) difference (pixel by pixel) between SimSetA and SimSetB; small squares (green on on-line version) highlight location of the defective pixels.

903x528mm (72 x 72 DPI)
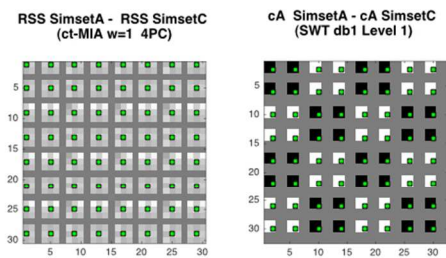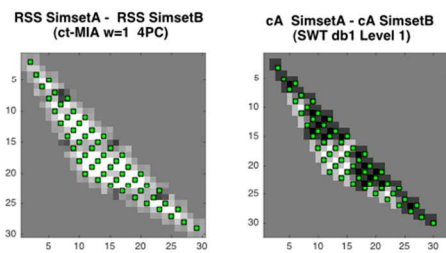
Figure 10: Differences (normal minus defective) among residuals sum of squares (RSS) images for ct-MIA approach (left) and differences among Approximations images from WT-MIA approach (right). Figure 10a (top). SimSetA (normal) minus SimSetB (defective). Figure 10b (bottom) SimSetA minus SimSetC.
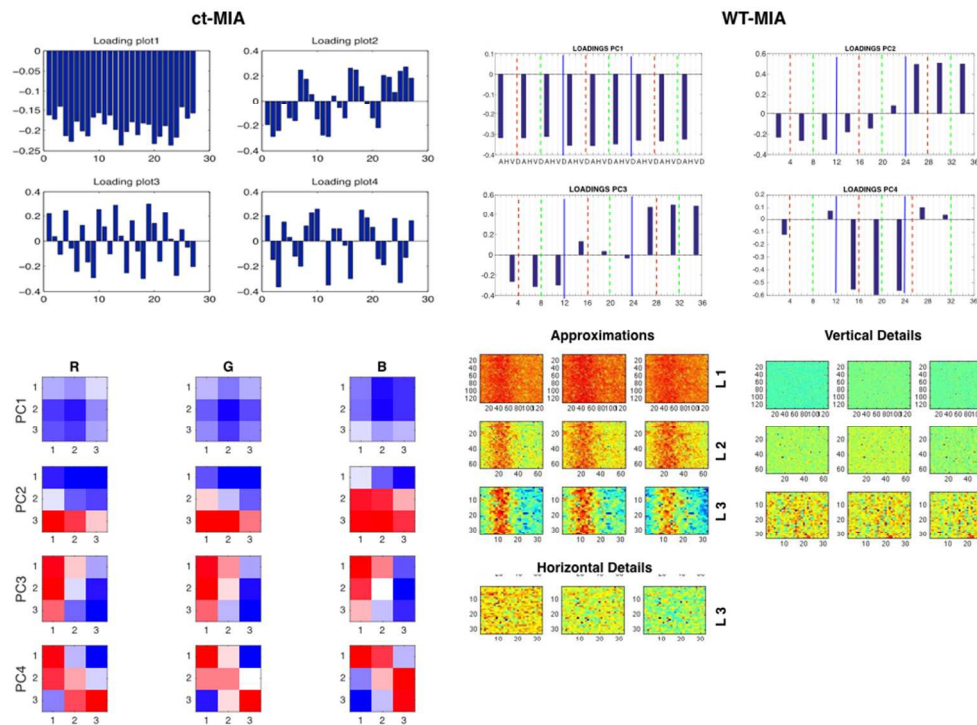
361x270mm (72 x 72 DPI)

Figure 11: On the right: ct-MIA loadings for BNdataset, as bar plot (top) and as image plot (bottom); on the left WT-MIA loadings (top) and WT images for each decomposition block that it is most contributing to the loadings: Approximations from level 1 to 3 contributing to PC1; Horizontal details, level 3, contributing to PC2; Vertical details level 1 and 3 contributing to PC3 and Vertical details level 2 contributing to PC4.

361x270mm (72 x 72 DPI)
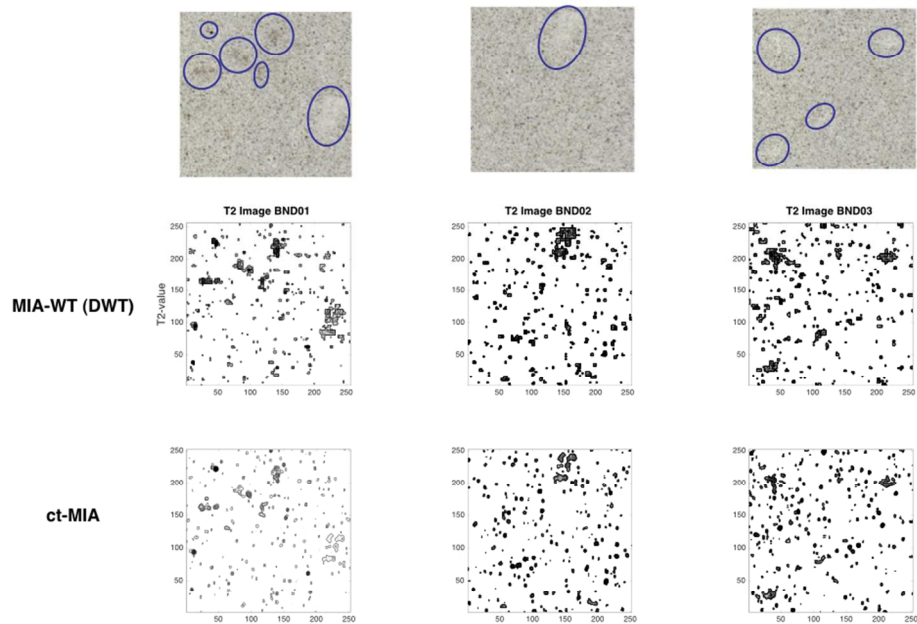
Figure 12: (a) Defects detection for BNdataset, from left to right the three defective images BND01, BND02 and BND03, respectively. On top defective images with defect encircled. Middle T2-chart obtained by WT-MIA (daubechies 1, level 3, DWT scheme). Bottom T2-chart obtained by ct-MIA (w = 2).
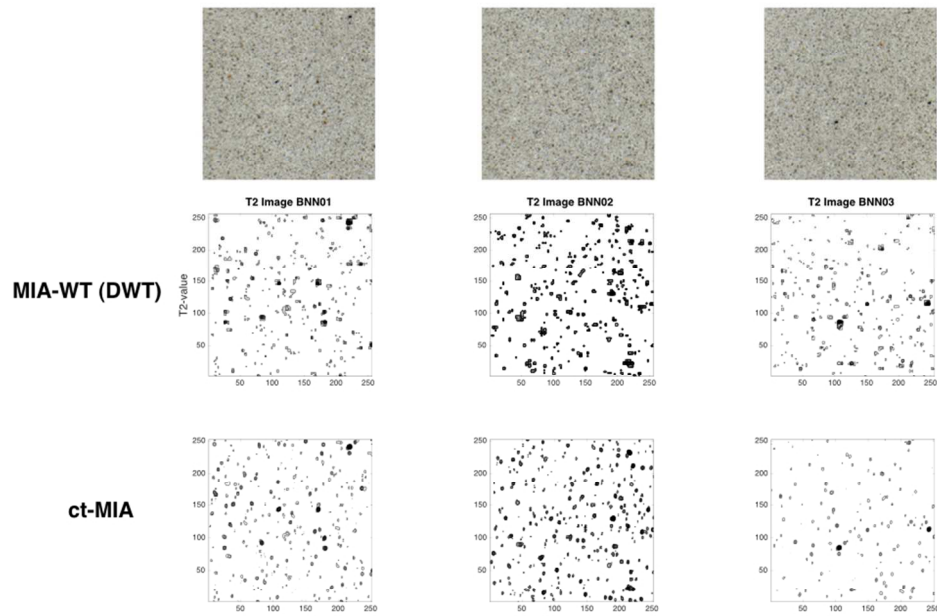
361x270mm (72 x 72 DPI)

Figure 12: (b) On top normal images; Middle T2-chart obtained by WT-MIA (Daubechies 1, level 3, DWT scheme). Bottom T2-chart obtained by ct-MIA (w = 2).‖ ┼

361x270mm (72 x 72 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
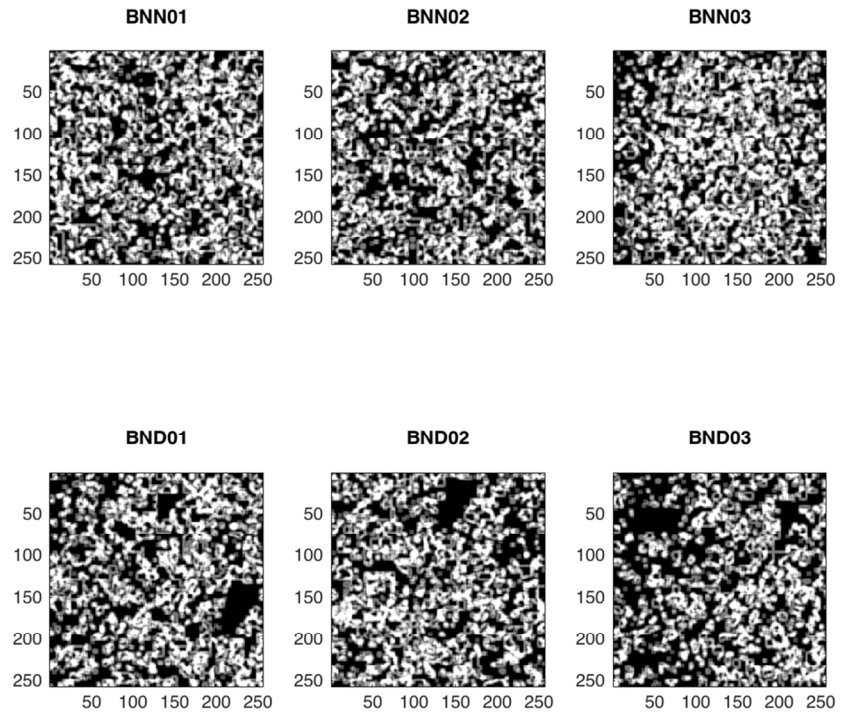43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 13: Results of local entropy analysis on scores images of normal and defective images for BNdataset.
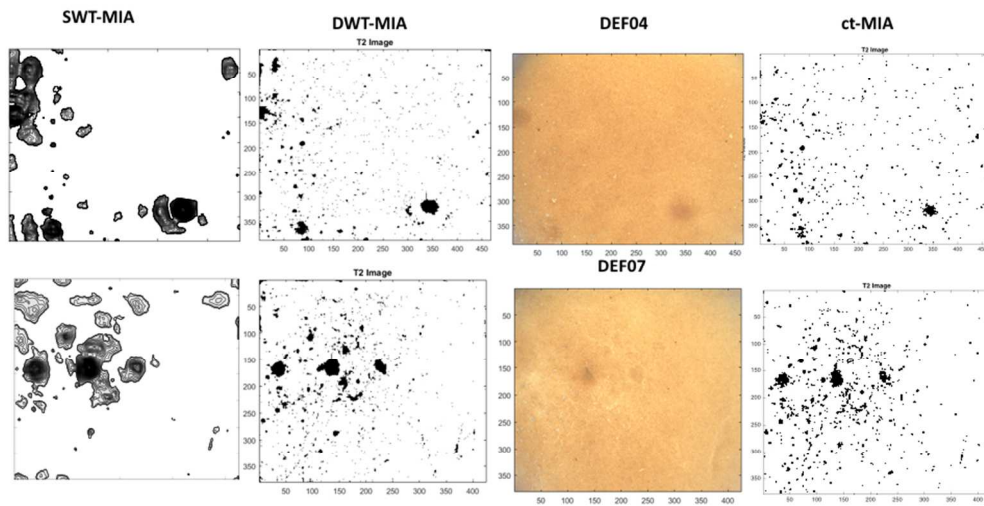
398x354mm (72 x 72 DPI)

Figure 14: Bread datas set: T2-chart images (from left to right SWT, DWT and ct-MIA) for defective images compared with raw images.
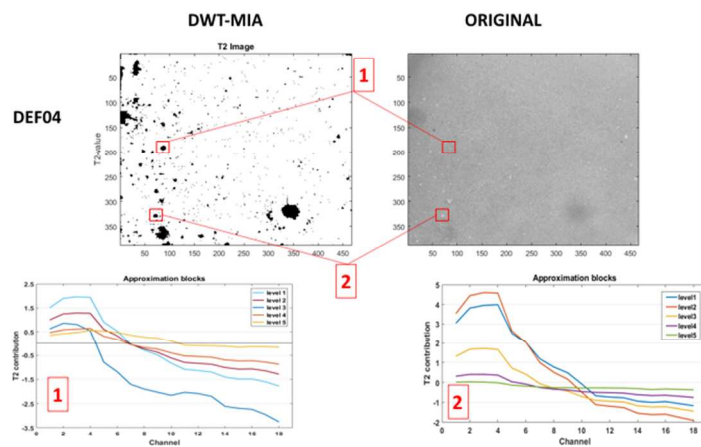
338x190mm (72 x 72 DPI)

Figure 15: Bread data set, D04 image. T2-chart and contribution plot for some of the pixels highlighted as defective (also shown on the raw image), for clarity only the Approximation block, whose T2-contribution is the most relevant, is shown (each curve represent a decomposition level) on the x-axis are reported the spectral channels.

338x190mm (72 x 72 DPI)