Image analysis by bidimensional empirical mode decomposition

J.C. Nunes*, Y. Bouaoune, E. Delechelle, O. Niang, Ph. Bunel

Laboratoire d’Etude et de Recherche en Instrumentation, Signaux et Systèmes (LERISS), Université Paris XII-Val de Marne, Bat P2-Piece 230 61 Avenue du Général de Gaulle, 94010 Créteil Cedex, France

Received 30 September 2002; received in revised form 18 April 2003; accepted 9 May 2003

Abstract

Recent developments in analysis methods on the non-linear and non-stationary data have received large attention by the image analysts. In 1998, Huang introduced the empirical mode decomposition (EMD) in signal processing. The EMD approach, fully unsupervised, proved reliable monodimensional (seismic and biomedical) signals. The main contribution of our approach is to apply the EMD to texture extraction and image filtering, which are widely recognized as a difficult and challenging computer vision problem. We developed an algorithm based on bidimensional empirical mode decomposition (BEMD) to extract features at multiple scales or spatial frequencies. These features, called intrinsic mode functions, are extracted by a sifting process. The bidimensional sifting process is realized using morphological operators to detect regional maxima and thanks to radial basis function for surface interpolation. The performance of the texture extraction algorithms, using BEMD method, is demonstrated in the experiment with both synthetic and natural images.

Keywords: Bidimensional empirical mode decomposition; Texture analysis; Unsupervised texture decomposition; Radial basis function; Surface interpolation

1. Introduction

The joint space-spatial frequency representations have received special attention in the fields of image processing, vision and pattern recognition. Huang [15] introduces a multiresolution decomposition technique: the empirical mode decomposition (EMD), which is adaptive and appears to be suitable for non-linear, non-stationary data analysis. We propose a new analysis method of texture images based on bidimensional empirical mode decomposition (BEMD), firstly presented in Ref. [24].

This paper is organized as follow. Section 2 presents state of the art of texture analysis methods. Section 3 presents an introduction to the EMD and its extension on bidimensional data. It describes too the implementation details of sifting process, including the extrema detection by morphological reconstruction and the radial basis function (RBF) for surface interpolation. In Section 4, experimental results on different texture images indicate the interest of this new multiresolution decomposition with statistical descriptors. Finally, a conclusion is presented in Section 5.

2. Texture analysis methods

Texture features are determined by the spatial relations between neighbouring pixels. The difficulty of texture analysis is demonstrated by the number of different texture definitions attempted by vision researchers [12,17,28].

Approaches to texture feature extraction and recognition span a wide range of methods. Several books and articles give overviews of the available methodology [9, 33]. There are four major issues in texture analysis [22]: feature extraction (local texture properties), texture discrimination (image partition corresponding to different textures), texture classification (classes definition in finite number, normal and pathological) and shape from texture (reconstruction of 3D surface geometry from texture information).

Four major method categories may be identified [20]: statistical with cooccurrence and autocorrelation features [12], geometrical with Voronoi tessellation [33] and structural features, model based with Markov random field [5,11] and fractal parameters [6,26,32], and multiscale features with AM–FM analysis [13], morphological operators [27], Wigner-Ville Distributions [14,31], Gabor filters [8,16], and wavelet transforms [21,23].
A few comparisons between texture feature extraction schemes have been presented in Refs. [22,29,30]. Comparative studies showed different conclusions. Different setups, test images, and filtering methods may be the reasons for the contradicting results. No single approach did perform best or very close to the best for all images. The multiresolution techniques, which based on human visual perception, intend to transform images into a representation in which both spatial and frequency information are present. The most commonly multiscale features used are Wigner distributions [14,31], Gabor functions [2,7,16] and wavelet transforms [18,19,21,34,36]. These multiscale features try to characterize textures by filter responses directly. However, one difficulty of the multiresolution analysis is its non-adaptive nature since it uses filtering schemes.

We have applied the EMD technique to texture images for two reasons. The first, the advantage of this technique, the EMD is a fully data driven method [25], does not use any pre-determined filter [8] or wavelet functions [19]. The second, we can easily implement a bidimensional extension of EMD. This paper, which is based on multiscale decomposition, examines the issue of designing texture decomposition by BEMD.

3. Texture analysis based on the empirical mode decomposition

3.1. The 1D empirical mode decomposition

In this paper, we use the EMD, first introduced by Huang et al. [15]. Different applications as medical and seismic signals analysis have showed the effectiveness of this method. This method permits to analyse non-linear and non-stationary data. Its principle is to decompose adaptively a given signal into frequency components, called intrinsic mode functions (IMF). These components are obtained from the signal by means of an algorithm called sifting process. This algorithm extracts locally for each mode the highest frequency oscillations out of original signal.

3.1.1. Sifting procedure

The sifting procedure decomposes a sampled signal $s(k)$ by means of the EMD. The sifting procedure is based on two constraints:

- each IMF has the same number of zero crossings and extrema;
- each IMF is symmetric with respect to the local mean.

Furthermore, it assumes that $s$ has at least two extrema.

The EMD represent adaptively non-stationary signals as sums of zero-mean AM–FM components [10], i.e. an IMF is an AM–FM component. AM–FM analysis [13] has been used successfully in a variety of applications including non-stationary analysis, edge detection, image enhancement, recovery of 3D shapes from texture, computational stereopsis, texture segmentation and classification.

The sifting algorithm for $s \in \mathcal{L}^2(\mathbb{R})$ reads as follows:

1. Initialise: $r_0 = s$ (the residual) and $j = 1$ (index number of IMF),
2. Extract the $j$th IMF:
3. (a) Initialise $h_0 = r_{j-1}, i = 1$,
   (b) Extract local minima/maxima of $h_{i-1}$,
   (c) Compute upper envelope and lower envelope functions $x_{i-1}$ and $y_{i-1}$ by interpolating, respectively, local minima and local maxima of $h_{i-1}$,
   (d) Compute $m_{i-1} = (x_{i-1} + y_{i-1})/2$ (mean envelope),
   (e) Update $h_i := h_{i-1} - m_{i-1}$ and $i := i + 1$,
   (f) Calculate stopping criterion (standard deviation $SD_{hi}$)
   (g) Repeat steps (b) to (f) until $SD_{hi} \leq SD_{MAX}$ and put then $s_j = h_i$ (jth IMF)
4. Update residual $r_j = r_{j-1} - s_j$.
5. Repeat steps 2–4 with $j := j + 1$ until the number of extrema in $r_j$ is less than 2.

3.1.2. Finding all the IMFs

1. Once the first set of ‘sifting’ results in an IMF, define $c_1 = h_{1i}$. This first component contains the finest spatial scale in the signal.
2. Generate the residue, $r_1$, of the image by subtracting out $c_1, r_1 = I - c_1$. The residue now contains information about larger scales.
3. Resift to find additional components $r_2 = r_1 - c_2, \ldots, r_n = r_{n-1} - c_n$

The superposition of all the IMF reconstructs the data: $I = \sum_{i=1}^{n} (c_i) + r_n$.

We have to determine a criterion for the sifting process to stop. This can be accomplished by limiting the size of the SD, computed from the two consecutive sifting results as:

$$SD_j^2 = \sum_{k=1}^{K} \left[ \frac{|h_{ji(k)}(k) - h_{ji}(k)|^2}{h_{ji-1}^2(k)} \right]$$

3.2. The bidimensional empirical mode decomposition

Texture analysis is considered as a challenging task. The ability to effectively classify and segment images based on textural features is of key importance in scene analysis, medical image analysis, remote sensing and many other application areas. Feature extraction is the first stage of image texture analysis. To extract the 2D IMF during the sifting process, we have used morphological reconstruction.
Determine the local mean
Generate the 2D ‘envelope’ by connecting maxima
and competitive with the best analysis results [13,25].

We define a bidimensional sifting process [24]:

- Identify the extrema (both maxima and minima) of the image $I$ by morphological reconstruction based on geodesic operators;
- Generate the 2D ‘envelope’ by connecting maxima points (respectively, minima points) with a RBF;
- Determine the local mean $m_1$, by averaging the two envelopes;
- Since IMF should have zero local mean, subtract out the mean from the image: $I - m_1 = h_1$;
- repeat as $h_1$ is an IMF.

As described above, the process is indeed like sifting to separate the finest local mode from the image first based only on the characteristic multiscale. The sifting process, however, has two effects to eliminate riding waves and to smooth uneven amplitudes.

3.2.1. Extrema detection

Morphological reconstruction is a very useful operator provided by mathematical morphology [1]. Its use in hierarchical segmentation proves its efficiency in all the steps of the process, from extrema detection to hierarchical image construction. The image extrema have been detected by using morphological reconstruction based on geodesic operators. We define the geodesic reconstruction as follows.

The grayscale reconstruction $I_{rec}(J)$ of $I$ from $J$ is obtained by iterating grayscale geodesic dilations $\partial_I^J$ of $J$ under $I$ until a stability is reached, i.e. $I_{rec} = \bigvee_{n=1}^N \partial_I^J(J)$.

A well-known use of the morphological reconstruction is the extraction of the extrema (minima and maxima). If we take $J = I - 1$ (subtract one gray level to every pixel of original image) and if we perform the reconstruction $I_{rec}$ (by geodesic dilation) of $J$ by $I$, the difference $I - I_{rec}$ corresponds to the indicator function of the maxima of $I$ (Fig. 1).

Conversely, the difference between $I_{rec}$ (reconstruction by geodesic erosion) and $I$ (original image) produces the indicator function of the minima of $I$. This extrema detection method is described in Ref. [35].

3.2.2. Surface interpolation by radial basis function

In Ref. [15], Huang proposed to use cubic spline interpolation on non-equidistant sampled data. We have choice to use the RBF rather than the bicubic spline for different reasons developed in Ref. [4]. One of the technical problems is that the cubic spline fitting creates distortions near the end points. RBFs are presented as a practical solution to the problem of interpolating incomplete surfaces derived from three-dimensional (3D) medical graphics. A RBF is a function of the form:

$$s(x) = p_m(x) + \sum_{i=1}^{N} \lambda_i \Phi(\|x - x_i\|), \quad x \in \mathbb{R}^d, \quad \lambda_i \in \mathbb{R},$$

where

- $s$ is the radial basis function,
- $p_m$ is a low degree polynomial, typically linear or quadratic, a member of $m$th degree polynomials in $d$ variables,
- $\|\|$ denotes the Euclidian norm,
- the $\lambda_i$’s are the RBF coefficients,
- $\Phi$ is a real valued function called the basic function,
- the $x_i$’s are the RBF centres.

The radial basis approximation method offers several advantages over piecewise polynomial interpolants. The geometry of the known points is not restricted to a regular grid. Also, the resulting system of linear equations is guaranteed to be invertible under very mild conditions. Finally, polyharmonic RBFs have variational characterizations, which make them eminently suited to interpolation of scattered data, even with large data-free regions. These applications include geodesy, geophysics, signal processing, and hydrology. RBFs have also been successful employed for medical imaging and morphing of surfaces in three dimensions.

Experimental results demonstrate that high-fidelity reconstruction is possible from a selected set of sparse and irregular samples. RBF are introduced to compute a continuous surface through a set of irregularly spaced points (extrema) during the sifting process.

4. Results and discussion

Real-world texture frame, as well synthetic texture frames, has been used to test and validate the proposed approach. The decomposition approach is applied to both synthetic and natural images, created by composing textures selected from Brodatz [3]. A study is performed to show the efficiency and performance of the texture extraction. Then, a set of texture images is selected to illustrate the application procedure of the EMD method.
4.1. Texture extraction

The 2D decomposition by sifting process of an image provides a representation that is easy to interpret. Every mode (IMF) contains information of a specific scale, which is conveniently separated. Spatial information is retained within the mode. Our algorithm is able to decompose a broad range of texture elements. Examples using images (200 × 200 in grey scale) from the Brodatz texture album, or using synthetic and real images are shown in Figs. 3, 6–12.

To stop the sifting process, we used the standard deviation (SD). We have used SDMAX between 0.05 and 0.75.

In a first step, we applied our algorithm on a synthetic image (Fig. 2a). It is an image of sinusoidal components, built with a sum of three horizontal ($h_1 = 40$, $h_2 = 14$, $h_3 = 0.2$) and three vertical ($v_1 = 60$, $v_2 = 17$, $v_3 = 0.3$) frequencies with respective amplitudes ($a_{h1} = 40$, $a_{h2} = 30$, $a_{h3} = 190$, $a_{v1} = 40$, $a_{v2} = 50$, $a_{v3} = 190$). All stages of results are shown in Figs. 2 and 3.

We observe the images of the minima (Fig. 2b), the maxima (Fig. 2c) detection, the mean envelope (Fig. 2d) and the subtraction from mean envelope (Fig. 2e). Fig. 3 shows the performed decomposition in two modes (Fig. 3b and c) and the residue image (Fig. 3d). To prove the effectiveness of our method, we compare the density profiles of the original image and the various frequencies composing it. In Fig. 4, we traced the profile of density of sinusoidal images (corner in top on the left towards the corner in bottom in right-hand side): the sum of the three frequencies (Fig. 4a), the high frequency (Fig. 4b), the medium frequency (Fig. 4c) and the low frequency (Fig. 4d).

In Fig. 5, we can observe the density profile of decomposition, the sum of three frequencies (Fig. 5a), the first mode (Fig. 5b), the second mode (Fig. 5c) and the image residue (Fig. 5d). We can compare these different profiles. We observe that the first mode corresponds to the sinusoidal image with the highest spatial frequency, the second to the medium and the residue to the lowest. These profiles are very similar. However, we can watch in residue image (Fig. 5d) an irregularity, which is due to interpolation distortions near the end points.

In Fig. 6, the decomposition is performed in two modes. This fabric texture is quite regular, mainly diagonal and horizontal structures, making it well suited to the proposed approach.

![Fig. 2. Photographic texture fabric.](image)

![Fig. 3. Synthetic texture.](image)

![Fig. 4. Synthetic texture (frequency components).](image)
The first mode corresponds to the woven structure, the second to the pattern and the residue to the horizontal stripes (black and white). In Fig. 7, the decomposition is performed in three modes. It is interesting to observe that the first mode corresponds to the woven structure, the second to the finest pattern, the third to the medium pattern and the residue to largest pattern.

In Fig. 8, the decomposition is performed in three modes. In Fig. 9, the decomposition of MRI is performed in three modes. We observe the three modes and the residue, which contain the pattern structures from finest to coarsest.

We can observe that the extraction quality of a mode depend on the quality of the previous modes. Therefore, the stop criteria (SD) of sifting process is important.

4.2. Image filtering

Since sifting process extracts firstly the highest frequency, the first modes correspond generally to the noise. The information of the noise is contained in the very first modes and in the residual image. Conversely, the image tendency is contained only in the residual image.
The tendency can be represented by a polynomial of order relatively low \( (0, 1, 2) \). The results in Fig. 10 show the possibility of low level processing (filtering or denoising) with this technique. After having applied the decomposition, this filtering would be easily carried out by the subtraction with the original image of one or several modes. The residue represents the filtered image.

**4.3. Extraction of inhomogeneous illumination**

Since sifting process extracts firstly the highest frequency, the first modes correspond generally to the noise. Conversely, the image tendency is contained in the latest mode or more generally in residual image resulting to the sifting process (Fig. 11).

After the decomposition, we subtract residue image from original image (Fig. 12c) to perform inhomogeneous illumination correction.

**4.4. Perspectives**

In the bidimensional case, the regional extrema are not always well defined. The saddle points (or more generally, the pattern of ridges and valleys) should be taken into account. Thus, we propose the straightforward extension of the 1D EMD to the 2D case by using morphological operators. To detect lines peaks (respectively, the line...
connecting the lowest points) corresponding to the maxima (respectively, minima), we use the watershed, a powerful tool for image segmentation. More details of the BEMD from these lines (the saddle points) will be presented in future paper.

We can observe in Fig. 12 the BEMD from these lines and can compare these results with BEMD from regional extrema in Fig. 9. For the BEMD from the regional extrema (Fig. 9), we used SD\textsubscript{MAX} = 0.75. For the BEMD from the saddle points (Fig. 12), we used SD\textsubscript{MAX} = 0.47. By comparing Figs. 9 and 12, we can observe that the details corresponding to the modes are finer for the BEMD from the saddle points. Consequently, BEMD based on saddle points seem to produced more IMF than the preceding version.

5. Conclusion

In this paper, we present a new modulation domain feature-based approach for discriminating textured images. We have applied the EMD to texture image analysis. The BEMD permits to extract spatial frequency components or
different spatial scales, i.e. structures from finest to coarsest scales. This method, derived from the image data and fully unsupervised, permits to analyse non-linear and non-stationary data as texture images. We have shown experimental results for both natural and synthetic textures. By having access to these representations of scenes or objects, we can concentrate on only one or several modes (one individual or several spatial frequency components) rather than the image entirety. Clearly the 2D EMD offers a new and promising way to decompose and extract texture features without parameter.

References


