Content Based Image Retrieval using the Generalized Gamma Density to model BEMD’s IMFs.

Aziza Benkuider
LESSI, Faculty of sciences Dhar el mahraz P.Box 1796
Department of physic, Fez, Morocco.
Email: abekuider@hotmail.com

Abdellah Aarab
LESSI, Faculty of sciences Dhar el mahraz P.Box 1796
Department of physic, Fez, Morocco.
Email: azbekuider@yahoo.fr

Abstract—In this paper, we present a texture-image retrieval approach, which is based on the idea of to characterize images without extracting local features, by using global information extracted from the image Bidimensional Empirical Mode Decomposition (BEMD) together with the Generalized Gamma (GG) Density. The BEMD method decompose image into a set of functions named Intrinsic Mode Function (IMF) and residue. The Generalized Gamma (GG) Density is used to represent the coefficients derived from each IMF and the Kullback-Leibler Distance (KLD) compute the similarity between Gamma Generalized function’s coefficients. The experimental results indicate that our approach can achieve higher retrieval rates.

Index Terms—Content Based Image Retrieval (CBIR), Empirical Mode Decomposition (EMD), Bidimensional Empirical Mode Decomposition (BEMD), Fast and Adaptive Bidimensional Empirical Mode Decomposition (FABEMD), Intrinsic Modes Functions (IMF), Generalized Gamma (GG), Kullback-Leibler (KL).

I. INTRODUCTION

Nowadays, we see the advancement in image acquisition technologies and storage systems which always encourages us to design a sophisticated system to retrieve the images effectively. Fast access to such huge database requires efficient indexing algorithms. CBIR system is one of possible solutions to effectively manage image databases [1,2]. The purpose is to retrieve desired images from database using only the numerical content of images. Typical CBIR systems is decomposed in two steps: the first is feature extraction (FE), where a set of features, called image signatures, is generated to accurately represent the content of each image in the database and stored in a feature database. The second is similarity measurement (SM), where a distance between the query image and each image in the database using their signatures is computed. For definition and extraction of visual features, many classical approaches in the field of texture-image retrieval leave the spatial domain including image characterization using a transform domain such as wavelet transform and Gabor filter [3,4].

Recently, a novel multiresolution decomposition method, The Empirical Mode Decomposition (EMD) originally proposed by Huang [5] for one dimensional signal and extended specifically to image processing two years after (BEMD) [6]. The BEMD is a very powerful tool in image retrieval [7,8]. It is based on the characterization of the image through its decomposition in IMFs where the image can be decomposed into a redundant set of signals and a residue, adding all the IMF’s together with the residue reconstructs the original image without information loss or distortion.

Some researchers used BEMD for image processing. Nunes and al. [6] applied BEMD to texture extraction and image filtering. Linderhed used BEMD for image compression [9]. Liu and al. presented the definition and framework of Directional Empirical Mode Decomposition (DEMD) and applied DEMD to texture segmentation [10]. The BEMD process cause massive computation time and complexity. In this context, Sharif, M and all. suggested a new simple, efficient and fast method of envelope estimation that replaces the surface interpolation (Fast and Adaptive BEMD: FABEMD) [11].

In this work, we propose a new texture-retrieval method which takes advantage of image FABEMD. Roland Kwitt and all. [4,12] used Generalized Gamma density functions (GG) to represent texture images in the wavelet domain. In this paper we propose a new
algorithm, based on Roland Kwitt work, using FABEMD decomposition instead of wavelet transform. In order to evaluate the proposed FABEMD using the Generalized Gamma (FABEMD-GG), we presented a comparison study with the Gabor Wavelet transform (GWT), the classical pyramidal Discrete Wavelet Transform (DWT) or the Dual-Tree Complex Wavelet Transform (DT-CWT). The results showed that FABEMD-GG texture can achieve encouraging retrieval precision and recall.

The remainder of this paper is structured as follows: In the first section, we introduce the BEMD with FABEMD. In the next section, we introduce the Generalized Gamma density functions (GG) model of the IMFs. Section 3 then briefly review the connection between the KLD and the principle of maximum-likelihood selection. In section 4, we present the approach and provide closed-form solutions for the corresponding KLD. Section 5 and 6 presents a comparison of the experimental retrieval results. Finally, concludes the paper with a brief summary and an outlook on further research.

II. THE BIDIMENSIONAL EMPIRICAL DECOMPOSITION.

The EMD method is an adaptive decomposition and suitable for analysis of non-linear and non-stationary processes. This decomposition technique extracts a finite number of oscillatory components or (well-behaved) AM-FM functions, called IMF, directly from the data. IMFs are monocomponent functions that have well defined instantaneous frequencies. EMD does not use any pre-determined filter or basis functions; which is quite different from Gabor analysis [13] and Wavelet analysis [14]. This technique extended to analyse two-dimensional data is known as Bidimensional EMD (BEMD), the process being called 2D sifting process. The procedure of decomposition is iterative. The 2D sifting process is performed in two steps: extrema detection by neighboring windows or morphological operators and surface interpolation by radial basis functions [15] or multigrid B-spline. Its principle is to decompose adaptively an image into its IMFs. This technique has been applied successfully on real data in areas such that the Oceanography and the Climatically phenomenon study [16, 17], the Seismology [18], the Control No-Destructive [19, 20], the Acoustic submarine [21, 22], or Biology [23, 24].

A. The BEMD steps

1. Algorithm

Given an image I, the sifting process of BEMD can be defined as follows [6]:

- Step 1) Fixed, i, I ← 1
- Step 2) r_{i-1} ← I (residue)
- Step 3) Extract the i^{th} IMF:

a) h_{i,j-1} ← r_{i-1}, j ← I (j, iteration of the loop of sifting).

b) Extract local maxima and minima of h_{i,j-1}.

c) Compute upper envelope and lower envelope functions U_{i,j-1} and L_{i,j-1} by interpolating, respectively, local minima and local maxima of h_{i,j-1}.

d) Compute local mean surface:

\[ m_{i,j-1} = \frac{(U_{i,j-1} + L_{i,j-1})}{2} \]

f) Calculate stopping criterion: (Standard Deviation):

\[ SD(j) = \sum_{i=1}^{T} \left| h_{i,j} - h_{i,j+1} \right| \]

\[ SD(j) = \sum_{i=1}^{T} \left| h_{i,j} - h_{i,j+1} \right|^2 \]  \tag{1}

\[ SD(j) = \sum_{i=1}^{T} \left( \frac{U_{i,j-1} + L_{i,j-1}}{2} \right) \]

T: the number of iterations.

When the decomposition is complete, the original image may be reconstruct by adding all the IMFs and the last residue (2)

\[ I_{reconstruct} = \sum_{i=1}^{k} C_i \]  \tag{2}

where, C_i is the i^{th} IMF or residue and k is the number of IMFs.

Extraction of each IMF requires several iterations Because the surface interpolation method itself fits a surface in an iterative optimization approach, it makes the BEMD process complex and excessively time consuming, so to overcome this problem we can accelerate the process and minimize the computational time by the use of the New Bidimensional Empirical Mode Decomposition (NBEMD) [25], the Olympics Bidimensional Empirical Mode Decomposition (OBEMD) [26], the Lapped Block Bidimensional Empirical Mode Decomposition (LBBEMD) [27] and FABEMD [11].

2. Image extrema extraction
The first step of the sifting process in the BEMD is to locate the local extrema points (maxima and minima) of the image intensity. The simple way is to use the pixel neighbouring information in the image; so a point is maxima (respectively minima) if its value is strictly higher (respectively strictly lower) than all its 8-connected closer neighbours.

3. Surface interpolation

The interpolation is regarded as a delicate stage in the sifting process and has a remarkable influence on the results of the decomposition. A natural extension for the interpolation method used in the EMD for one dimensional is the cubic spline interpolation for the bidimensional signals. Unfortunately, this method decomposes the two dimensional down to residue with a many number of extrema points and it creates distortions near the end points. All this is because the data we have to interpolate is often described as scattered. Radial Basis Functions provide well-established tools for solving interpolation problems of the above form in any space dimension. An RBF is a function of the form:

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

where \( s \) is the RBF, \( p_m \) is a low-degree polynomial, typically linear or quadratic, a member of the \( m \)th degree polynomials in \( d \) variables, \( \| \cdot \| \) denotes the Euclidean norm, The \( \lambda_i \)'s are the RBF coefficients, \( \Phi \) is the basis function and \( x_i \) are the RBF centers. For this interpolation method we should fix in advance the basis function and since there are many functions that can be used as a basis function we can consider the Thin-Plate [28] as an extension of the cubic spline function as in the case of one dimension:

\[ \Phi(r) = r^2 \log(r) \]  

For the BEMD interpolation, the RBF offers several advantages over piecewise polynomial interpolates because the geometry of the known points is not restricted and the resulting system of linear equations is guaranteed to be invertible under very mild conditions.

4. Stopping criteria for the sifting process

The sifting process in the BEMD decomposition consist on decomposing an input signal into set of functions defined by the signal itself, these functions are called IMFs. An IMF is characterized by some specific properties [6, 29]:

- Each IMF is expected to have the following properties
  - The number of zero crossings and the number of extrema points is equal or differs by only one,
  - The envelopes defined by the local maxima and minima, respectively, are locally symmetric around the envelope mean.

The sifting process stops when the resulting image satisfied the characteristics of an IMF, as it is described above. In other term, it stops when the envelope mean signal is close enough to zero. After an IMF is found, we define the residue as the result of subtracting this IMF from the input image and then we iterate on the residue.

The BEMD is completed when the residue, ideally, does not contain any extrema points. This is if we suppose that the sifting process will converge, things that never been rigorously demonstrated. So, 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:

In practice, we have used SD between 0.02 and 0.3 and this stop criterion gives satisfying results.

5. Stopping criteria for the decomposition

Generally, the decomposition stops when the number of extrema is less than 2; that’s mean that there is no more oscillations to extract. In some case we can have a high number of IMFs without meeting this condition, so we should set a maximal number of IMF to extract. We can stop the decomposition depending on the need; for example, in image denoising we will need only the first IMF [7].

B. FABEMD steps

The decomposition FABEMD is a new simple, efficient and fast method of envelope estimation that replaces the surface interpolation step by a direct envelope estimation method precisely we used a filter. The FABEMD has the same steps as the BEMD except that the construction of the mean envelope is different from that of the BEMD.

Construction of the mean envelope by FABEMD:

- calculate a max-min filter.
- calculate the upper and lower envelope after smoothing.
- Determine the mean envelope.

To determine the filtering window, FABEMD uses an adaptive method to each IMF [11]. Fig. 1 shows an example of decomposition of an image by FABEMD.

![Example of decomposition an image of the vistex base by FABEMD.](image)

III. MODELLING WITH GENERALISED GAMMA DISTRIBUTION.

In this paper we suggest to characterize globally images by generating a numerical signature of image, based on IMFs. Roland Kwitt and all [4] show that, for textures images, the distribution of coefficients in the
subbands of the Dual-Tree Complex Wavelet Transform (DT-CWT) modelled by a generalized gamma law. Empirically, we noticed that the histogram derived from each IMF seems also to follow a generalized gamma law as illustrated below.

![IMFs coefficients distribution for image of Fig. 1.](image)

The probability density function of the GG distribution is:

\[ f(y; \alpha, \tau, \lambda) = \frac{\tau}{\lambda^\alpha} y^{\alpha-1} e^{-\left(\frac{\lambda}{\tau}y\right)^\alpha}, \quad y \geq 0, \quad \alpha, \tau, \lambda > 0 \]

where \( \Gamma(\cdot) \) is the Gamma function, \( \alpha \) and \( \tau \) are shape parameters, and \( \lambda \) a scale parameter.

The measure of KLD between two Generalized Gamma as follows

\[
\text{KL}_{\text{GGamma}}(p//p_0) = \log\frac{\phi_p}{\phi_{p_0}} - \log\frac{\Gamma(\alpha)}{\Gamma(\alpha_0)} - \alpha + \mu(\alpha, \phi_\tau, \phi_\lambda) + \mu(\alpha, -\alpha, \alpha_0) \lambda(\alpha, \phi_\tau, \phi_\lambda)
\]

with \( p = \text{GG}(\alpha, \tau, \lambda) \) and \( p_0 = \text{GG}_\theta(\alpha_0, \tau_0, \lambda_0) \).

\[ \phi_\tau = \frac{\tau}{\tau_0}, \quad \phi_\lambda = \left(\frac{\lambda}{\lambda_0}\right)^{\tau_0}/\tau \quad \mu(\alpha, \tau, \lambda) = \frac{\lambda^{\alpha} + 1}{\tau^\alpha} \]

\( \mu \): The moment of order 1 of the gamma generalized.

\( \nu(\alpha, \tau, \lambda) = \log \lambda + \frac{1}{\tau} \nu(\alpha) \)

\( \nu \): The geometric mean of gamma generalized.

We propose to characterize images by using the coefficients \( (\alpha, \tau, \lambda) \) of the distribution law for all IMFs of the BEMD decomposition.

IV. PROBABILISTE IMAGE RETRIEVAL

Let’s assume that we have \( N \) images \( I_i, 1 \leq i \leq N \) in our database. Each image is represented by a data vector \( x_i = \{x_{i1}, \ldots, x_{in}\} \), which is an element of some feature space \( X \subset \mathbb{R}^n \) and is obtained by feature extraction (FE). The retrieval task is to search the \( K \) most similar images to a given query image \( I_q \). We assume the query image is represented by data vector \( x_q \). From the probabilistic point of view, each data vector contains \( n \) realizations i.i.d. random variables \( x_1, \ldots, x_n \) which follow a parametric distribution with probability density function (PDF) \( p(x_i^\prime) \), \( i \in \mathbb{R}^n \). Given that we have a consistent estimator \( \hat{\theta} \) for the parameter vector \( \theta \), we can use \( \hat{\theta} \) without limitations. Under these premises, it is natural to select the most similar image \( I_i \) to \( I_q \) as the one, whose parameter vector \( \hat{\theta}_i \) leads to a maximization of the likelihood/log-likelihood function, i.e.

\[
\text{r} = \arg\max_i \frac{1}{n} \log(p(x_i/\theta_i))
\]

Note, that the additional factor \( \frac{1}{n} \) does not affect the maximization result. By applying the weak law of large numbers to (7) as \( n \to \infty \) (asymptotic case), we obtain

\[
\text{r} = \arg\max_i E_{p(x/\theta_i)} \log(p(x/\theta_i))
\]

and

\[
\text{r} = \arg\max E_{p(x/\theta_i)} \log(p(x/\theta_i))
\]

where the term \( E_{p(x/\theta_i)} \) denotes the expectation with respect to \( p(x/\theta_i) \) and \( D \) denotes the domain of \( p(x/\theta_i) \).

By observing that \( p(x/\theta_i) \) is an independent term for maximization, we can rewrite Eq.(8) as the following minimization problem:

\[
\text{r} = \arg\min_i \{ -\int_D p(x/\theta_i) \log(p(x/\theta_i))dx \}
\]

We note that the term \( \int_D p(x/\theta_i) \log(p(x/\theta_i))dx \) of (11) is simply the measure of KL distance \( \text{KL}(p_i/p_\theta) \) between \( p(x/\theta_i) \) and \( p(x/\theta_i) \) with \( p_i = p(x/\theta_i) \). So, the problem of maximizing the log-likelihood function can be seen as a minimization of the KL measure.

A reasonable similarity measure between images is one of the essential parts of every image retrieval and classification system. To compute the distance between two signatures, we estimate the KLD of the coefficient distribution law.

V. PROPOSED METHOD

We focus on the problem of texture-image retrieval, where we have an arbitrary query image and we want to
obtain the K most similar image from a given image database, according to some similarity criterion. Extraction of each IMF in the BEMD requires several iterations. Because the surface interpolation method itself fits a surface in an iterative optimization approach, it makes the BEMD process complex and excessively time consuming. For this, we used the FABEMD.

A. Signature

The algorithm for construction of the signature is (The diagram of this construction is presented in Fig. 3):

- Decompose the image by FABEMD.
- Calculate the histogram of each IMF.
- Modeling each histogram by a Generalized Gamma density law.
- Constructed the signature with the coefficients the gamma generalization function of all IMFs.

B. Distance

The distance is a weighted sum of the KLD between corresponding IMF’s signature of two images.

\[ D(I_1, I_2) = \sum_{k=1}^{N} \gamma_k \text{KL}(p(y/\alpha, \tau, \lambda), p(y/\alpha_k, \tau_k, \lambda_k)) \]  

where \( \gamma_k \) the adjustment weights and \( N \) are is the number of IMFs.

VI. SYSTEMATIC EVALUATION

A retrieved image is considered a match if it belongs to the same category of the query image. To evaluate the performance of these texture-image retrieval algorithm we have used two well know parameters precision and recall derived from all four possible outcomes of CBIR experiment as show in Fig. 4.

Precision: precision is the fractal of the relevant images which has been retrieved (from all retrieval)

\[ \text{Precision} = \frac{A}{A+B} \]

where \( A \) is relevant retrieved and \( A+B \) is all retrieved images.

Recall: recall is the fraction of the relevant images which has been retrieval (from all relevant).

\[ \text{Recall} = \frac{A}{A+D} \]

where \( A \) is relevant retrieved and \( A+D \) is all relevant images.

We used the evaluation criteria which is the mean precision at 16, which is the ratio between the number of pertinent images recalled (same class than the query) and the total images.

Each image in the database is used as a query image.

- The algorithms find the 16 first images of the database closest to the query image.
- Precision is computed for this query.
- Finally, we compute the mean precision.

VII. RESULTS AND COMPARATIVE STUDY

We aim to compare our approach with those proposed in [3], [4] and [12]. For this, we kept the same experimental conditions; using the database Vistex consisting of 40 classes with size 512x512 pixels. They were downloaded from MIT vision textures database [30]. We take the K-nearest images to our query image and calculate how many they are really in the same class.

We used our algorithm for constructed image signatures.

From TABLE I we can see that a method based adapted FABEMD-GG achieved the highest rates as the GWT, DWT and DTCWT.

We used our algorithm for constructed image signatures.

From TABLE I we can see that a method based adapted FABEMD-GG achieved the highest rates as the GWT, DWT and DTCWT.

To see the differences in retrieval rates for 40 textures of our experiments, we additionally provide the texture-specifics rates in TABLE II. For 90% of the texture, the FABEMD-GG approach achieves higher retrieval rates.
In case of the gamma assumption, the results look similar, although we cannot provide a full listing of the results here, due to space limitations.

<table>
<thead>
<tr>
<th>Table II. Texture Specific Retrieval Results.</th>
</tr>
</thead>
<tbody>
<tr>
<td>textures</td>
</tr>
<tr>
<td>Bark 00</td>
</tr>
<tr>
<td>Bark 06</td>
</tr>
<tr>
<td>Bark 08</td>
</tr>
<tr>
<td>Bark 09</td>
</tr>
<tr>
<td>Brick 01</td>
</tr>
<tr>
<td>Brick 04</td>
</tr>
<tr>
<td>Brick 05</td>
</tr>
<tr>
<td>Buildings 09</td>
</tr>
<tr>
<td>Fabric 00</td>
</tr>
</tbody>
</table>

Finally, using the Directional Bidimensional Empirical Mode Decomposition.

**REFERENCES**


© 2011 ACADEMY PUBLISHER


Aziza Benkuider received her Master’s degree in 2006 from the faculty of sciences and techniques (FST) Fes Morocco. She is pursuing her Ph.D. degree. Her doctoral study focused on the application of the BEMD for image retrieval.

Abdellah Aarab is Professor at the Faculty of Sciences, Fes Morocco. He received his “Doctorat de 3ième cycle” degree from the University of Pau, France in 1982 and “These d’Etat” degree in 1999 from university Sidi Mohamed Ben Abdellah Fes Morocco. His research areas include, image processing, information retrieval and industrial applications.