PAPER

Iteration-Free Bi-Dimensional Empirical Mode Decomposition and Its Application

Taravichet TITIJAROONROJ\(^{a}\), Student Member and Kuntpong WORARATPANYA\(^{b}\), Nonmember

SUMMARY A bi-dimensional empirical mode decomposition (BEMD) is one of the powerful methods for decomposing non-linear and non-stationary signals without a prior function. It can be applied in many applications such as feature extraction, image compression, and image filtering. Although modified BEMDs are proposed in several approaches, computational cost and quality of their bi-dimensional intrinsic mode function (BIMF) still require an improvement. In this paper, an iteration-free computation method for bi-dimensional empirical mode decomposition, called iBEMD, is proposed. The locally partial correlation for principal component analysis (LPC-PCA) is a novel technique to extract BIMFs from an original signal without using extreme detection. This dramatically reduces the computation time. The LPC-PCA technique also enhances the quality of BIMFs by reducing artifacts. The experimental results, when compared with state-of-the-art methods, show that the proposed iBEMD method can achieve the faster computation of BIMF extraction and the higher quality of BIMF image. Furthermore, the iBEMD method can clearly remove an illumination component of nature scene images under illumination change, thereby improving the performance of text localization and recognition.

key words: iteration-free computation, locally partial correlation for principal component analysis (LPC-PCA), bi-dimensional empirical mode decomposition (BEMD), bi-intrinsic mode functions (BIMF), iteration-free bi-dimensional empirical mode decomposition (iBEMD), principal component analysis (PCA)

1. Introduction

An empirical mode decomposition (EMD) proposed by Huang et al.\([1]\) in 1996 is one of the powerful data adaptive analysis tools for non-linear and non-stationary signals. It has been successfully employed to analyze some non-stationary signals, mostly one-dimensional, such as signals from earthquake, Tsunami, and wind. This method can decompose an input signal into a set of intrinsic mode functions (IMFs) without using a priori function. Nonetheless, its decomposition process requires a high computational cost\([2]\). Shortly afterwards, Linderhed\([3]\) demonstrated a 2D empirical mode decomposition, called 2D-EMD, for image compression. That was the first extension of Huang’s EMD method to be used in two-dimensional signals. It requires a spline interpolation technique for estimating upper and lower surfaces to calculate mean surface, then the surface is used to extract IMF from bi-dimensional data. The extraction result is called bi-dimensional intrinsic mode function (BIMF). This paper focuses on a method for decomposing bi-dimensional signals that we called bi-dimensional empirical mode decomposition (BEMD). The terms “2D-EMD” and “BEMD” are interchangeable; hereafter, for ease of explanation, EMD and BEMD are used as the terms for one-dimensional and bi-dimensional signals, respectively.

In general, BEMD extracts BIMF by subtracting mean surface, estimated from upper and lower surfaces, from an original signal. Thus, the mean surface estimation plays an important role in decomposing BIMF. In this issue, Nunes et al.\([4]\) attempted to improve the interpolation of upper and lower surfaces by using radial basis function (RBF)\([5]\). It is known as BEMD-RBF. However, Nunes’s approach leads to a trade-off between BIMF quality and computational cost.

Since then, many papers\([6]\)–\([10]\) have proposed methods to reduce computation time. Most of them used an iteration technique to estimate mean surface. However, the iteration technique requires a lot of computation time, thus making it impractical in real world applications. For example, Lee et al.\([6]\) attempted to directly approximate mean envelop to reduce the number of iterations by half. Additionally, Damerval et al.\([7]\) and Bhuiyan et al.\([8]\) employed, respectively, piecewise cubic polynomial with Delaunay triangulation and order statistics filtering, respectively, instead of spline interpolation to speed up their algorithms. Furthermore, a minimal curvature of the curve technique (MC) and a maximum number of allowable iterations (MNAI) were used in\([7]\) and\([8]\), respectively, to limit the number of iterations.

Even though many modified BEMDs have been successful in speeding up their algorithms by using the iteration techniques, none of papers have proposed the algorithm that can decompose a BIMF in the first sifting result. In other words, no methods can instantaneously extract each BIMF from original signals with iteration-free computation. Therefore, this paper proposes an iteration-free technique that can drastically reduce the computational cost and to improve the quality of BIMFs. The performance of BEMDs is evaluated in terms of computation time and quality of BIMFs. Furthermore, the proposed method is successfully applied to Thai text localization and recognition under illumination effects on Thai text in natural scene images to demonstrate its effectiveness. The precision, recall, F-value, and accuracy rate are key indicators to reveal the achievement of the proposed method for resolving illumination change.
The rest of this paper is organized as follows. Section 2 describes the fundamentals and analysis of BEMD. The proposed method is described in Sect. 3. Experimental results are shown and discussed in Sect. 4. Finally, the conclusion is presented in Sect. 5.

2. Problem Formulation

This section describes the fundamentals of empirical mode decomposition (EMD) and bi-dimensional empirical mode decomposition (BEMD) methods and points out their limitations.

2.1 Empirical Mode Decomposition and Its Extension

EMD uses a sifting process to decompose an input signal into a set of intrinsic mode functions (IMFs). Huang et al. [1] has defined its important properties as follows.

**Definition 1** An intrinsic mode function (IMF) is defined as a function that satisfies two conditions: (i) in the whole dataset, the number of extrema and the number of zero-crossing must either be equal or differ at most by one; and (ii) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

Mathematically, a collection of IMFs and a residue that are completely decomposed from an input signal, \( s \), can be expressed as

\[
s = \sum_{i=1}^{n} \delta_i + r
\]  

where \( \delta_i \), \( r \), and \( n \) are a IMF component, a residue, and a total number of IMFs, respectively. From Definition 1 and Eq. (1), let \( g \), \( r \), and \( h \) be an input signal of sifting process of each iteration, an original input signal from the source, and an intermediate subtraction result, respectively. The sifting algorithm generated by Huang’s method [1] can be described as follows.

**Step 1:** Set an original signal to \( g_i \), for \( i = 1 \) and set \( j = 1 \).

**Step 2:** If \( i \) is equal to 1 then set \( gt_j = g_i \); otherwise set \( g_{j+1} = r_i \); \( gt_j = g_i \) and \( i = i + 1 \).

**Step 2.1:** Determine all maxima and minima points of \( gt_j \).

**Step 2.2:** Estimate upper \( \epsilon_{\text{upper}} \) and lower \( \epsilon_{\text{lower}} \) envelopes to obtain the mean envelope \( \epsilon_{\text{mean}} \) by using Eq. (2).

\[
\epsilon_{\text{mean}} = \frac{\epsilon_{\text{upper}} + \epsilon_{\text{lower}}}{2}
\]  

**Step 2.3:** Calculate \( h_j = g_i - \epsilon_{\text{mean}} \).

**Step 2.4:** Use standard deviation (SD) between \( h_j \) and \( h_{j+1} \), calculated from Eq. (3), as a stopping criterion instead of IMF conditions as reported in [11]. If SD is less than a threshold, equivalent to \( h_j \) approaches to no extrema points, then \( \delta_i = h_j \) and \( r_i = g_i - \delta_i \); else, set \( gt_{j+1} = h_j \), \( j = j + 1 \) and go to Step 2.1.

\[
SD = \sum_{j=1}^{T} \frac{|h_j - h_{j+1}|}{h_j^2}
\]  

where \( j \) and \( T \) denote an index of the element and the total number of elements in \( h_j \), respectively.

**Step 3:** Repeat Step 2 until the number of extrema of \( r_i \) is less than three points.

The sifting algorithm generates a set of IMFs in a descending order of frequencies such that the first IMF and residual components contain high and low frequencies, respectively. For extending EMD in 2002 and later, many papers have proposed to employ EMD to various types of bi-dimensional signal analysis such as image compression [3], texture extraction [4], image filtering, and image decomposition [8]. In 2005, Damerval et al. [7] formally defined bi-intrinsic mode function (BIMF), extended from the Definition 1, for two dimensional signals. A BIMF must satisfy the following two conditions in Definition 2.

**Definition 2** An image is a bi-dimensional IMF (BIMF): if it has a zero mean, if the maxima are positive and the minima are negative and if the number of maxima equals to the number of minima.

In this definition, the second condition proposed by Huang et al. [1] is replaced with a condition on the number of extrema, i.e., the number of maxima should be equal to the number of minima. To support the new condition in this definition, Damerval et al. showed that the number of maxima and minima are close to each other in practice even though they might not be exactly the same, i.e., the sifting result is a BIMF when the number of maxima and minima are nearly equal. This definition was successfully used by Xie [10].

2.2 Analysis of Bi-Dimensional Empirical Mode Decompositions

Most of bi-dimensional empirical mode decompositions are implemented in several approaches to speed up their algorithms. This subsection points out some shortcomings of those BEMD methods in the following paragraphs.

Several BEMD approaches [6]–[10] extract BIMFs in two stages: extrema detection and surface interpolation. A set of IMFs are decomposed by means of the iteration technique as explained in [1]. However, this technique consumes more computing time. Many papers have proposed to attempt to resolve this problem. They can be viewed as two attempts.

The first attempt was to make use of direct mean surface interpolation to reduce computation time. Lee et al. [6] used a moving average technique to interpolate in their
BEMD method called modified empirical mode decomposition (MEMD). Although Lee’s method can reduce the number of iterations by half for one-dimensional signal, it cannot be used directly with bi-dimensional signal. Furthermore, when MEMD is applied to an image vector, the interstice discontinuity appears in the resulting BIMFs as investigated in [11].

The last attempt was to limit the number of iterations. For example, Bhuiyan et al. [8] proposed a fast and adaptive bi-dimensional empirical mode decomposition (FaBEMD) method that was implemented by limiting the number of iterations to specify a maximum number of allowable iterations (MNAI) and using order statistics filter (OSF) to reduce computation time and artifacts in images. However, it has one major drawback; the optimum value of MNAI as well as those of the lowest distance (LD) and highest distance (HD) utilized in OSF for a particular problem are very difficult to estimate, thus making it impractical to apply in some situations. For another example, Damerval et al. [7] proposed a fast algorithm for bi-dimensional EMD (FBEMD) to reduce time-consuming by using Delaunay triangulation with piecewise cubic polynomial interpolation and minimal curvature of the curve technique (MC). Nonetheless, Damerval’s method is still computationally expensive as reported in [11]. Later, Xie [10] extended FBEMB to a logarithmic BEMD (LBEMD) by employing a logarithmic scale with Lambertian reflectance assumption [12]. This method focuses on the illumination preprocessing in face recognition, a different kind of problem tackled by the other investigations above, but it still inherits the high computational cost and artifact problems from FBEMD.

All of those methods can be summarized in Table 1. Many papers [6]–[8] made an effort to speed up their algorithms; nevertheless, none of papers can greatly reduce computation time and artifacts in BIMFs. In order to overcome these problems, this paper proposes a novel BEMD method as presented in Sect. 3.

### Table 1: A summary of state-of-the-art methods.

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Purpose</th>
<th>Approach</th>
<th>Computation-intensive</th>
<th>Artification</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D-EMD [3]</td>
<td>To propose 2D-EMD for image compression.</td>
<td>The extrema points are defined by nearest neighbors. The upper and lower surfaces are computed by using a thin-plate spline interpolation.</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>BEMD-RBF [4]</td>
<td>To develop EMD for texture extraction and image filtering.</td>
<td>A morphology operator is used to define extrema points. The upper and lower surfaces are computed from the radial basis function.</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>MEMD [6]</td>
<td>To extract features from an iris image.</td>
<td>The extrema points are defined by nearest neighbors. The mean envelope is computed from the moving averaging.</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>FBEMD [7]</td>
<td>To speed up the conventional BEMD.</td>
<td>The extrema points are defined by nearest neighbors. The upper and lower surfaces are computed by piecewise cubic polynomial interpolation. The minimal curvature of the curve technique is used to limit the number of iterations.</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>FaBEMD [8]</td>
<td>To reduce computation time and artifacts in BIMFs of the conventional BEMD.</td>
<td>The extrema points are defined by nearest neighbors. The upper and lower surfaces are computed by using order statistics filters. The maximum number of allowable iterations is set to limit the number of iterations.</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>LBEMD [10]</td>
<td>To restore a frontal illuminated face image based on BEMD.</td>
<td>The BIMF is extracted from FBEMD in logarithm scale.</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Yes means problem is found whereas No means no problem is found.

### 3. Proposed Method

As stated in Sect. 2, decreasing computation time of BIMF extraction is key to improving the performance of existing BEMD methods. Therefore, this paper proposes an iteration-free computation method based on locally partial correlation for principal component analysis (LPC-PCA) to directly estimate mean surface from bi-dimensional signals without using iteration technique for BIMF extraction. More details of the conventional PCA and LPC-PCA are described as follows.

#### 3.1 Locally Partial Correlation for Principal Component Analysis

Principal component analysis (PCA) is one of the linear schemes for optimal approximation to stochastic data and of the low computational cost techniques [13]. Consequently, it is a powerful tool for several purposes, such as data compression [14], dimensionality reduction [15], image analysis [16], and feature extraction [17]–[19]. A conventional PCA is performed in two steps: computing covariance matrix and constructing eigenvectors and eigenvalues. The covariance matrix is determined from correlations between attributes. Eigenvectors and eigenvalues are then constructed from that covariance matrix. The eigenvectors are sorted according to the descending order of their corresponding eigenvalues. The lowest-order eigenvector is used for construction of a mean surface by using an image reconstruction technique.
PCA estimates the mean surface by using coefficients, the straight line in Fig. 3 (a). Second, the conventional estimate from the global mean that is a single mean of the whole signal, the covariance matrix could not be accurately obtained from the first BIMF, since the data zero mean for computing the BIMFs is estimating an accurate mean surface instead of a smooth one with low entropy value. These errors occurred from the following two causes. First, the low frequencies could not be removed from the first BIMF, since the data zero mean for computing the covariance matrix could not be accurately obtained from the global mean that is a single mean of the whole signal, the straight line in Fig. 3 (a). Second, the conventional PCA estimates the mean surface by using coefficients of the covariance matrix to determine the correlation between two column vectors within a given image, but this estimation is inadequate because such coefficients cannot represent the structure of the given image; hence, artifacts appear in the residual component as shown in Fig. 2 (a). To overcome these drawbacks, this paper proposes an LPC-PCA method that uses partial correlation matrix instead of covariance matrix and local means instead of a global mean.

LPC-PCA is derived from the conventional PCA by modifying two steps. In the first modification, a covariance matrix is calculated from an original image, $A$, with resolution $H \times W$ and local means, $M$, representing the individual mean in each small partition of the whole image. Consequently, a subtracted result only contains the high frequencies. In other words, the covariance matrix obtained by local mean subtraction is the low-frequency free. The local mean vector, $m_k$, is defined in Eq. (4).

$$m_k = \begin{cases} 
\frac{1}{w+1} \sum_{i=k}^{k+w} a_i & \text{if } (k-w) \leq 0 \\
\frac{1}{w+1} \sum_{i=k-w}^{k} a_i & \text{if } (k+w) > j \\
\frac{1}{2w+1} \sum_{i=k-w}^{k+w} a_i & \text{otherwise}
\end{cases} \quad (4)
$$

where $w$ is a width of nearest vector neighborhood, $j$ is the last index of input data, $k$ is an index of column vector of input data, and $a_i$ is a column vector of the given image $A$. Statistically, the covariance matrix can be rewritten as Eq. (5).

$$\Gamma = \frac{1}{N-1} (A - M)^T (A - M) \quad (5)$$

where $N$ is the number of column vectors in $A$ and $M$ is a local mean matrix derived from the local mean vector as defined by Eq. (6).

$$M = [m_1, m_2, m_3, \ldots, m_j] \quad (6)$$

In the second modification, the covariance matrix in Eq. (5) is used to generate a partial correlation matrix so that the eigenvectors and eigenvalues are calculated from it. This modification provides the more accurate mean surface estimation for reducing artifact in the residual component. The partial correlation matrix, $\rho$, can be written as Eq. (7).

$$\rho = \begin{pmatrix}
\rho_{1,1} & \rho_{1,2} & \cdots & \rho_{1,j} \\
\rho_{2,1} & \rho_{2,2} & \cdots & \rho_{2,j} \\
\vdots & \vdots & \ddots & \vdots \\
\rho_{j,1} & \rho_{j,2} & \cdots & \rho_{j,j}
\end{pmatrix} \quad (7)$$

where $\rho_{x,y}$ is an element of the partial correlation matrix in $(x, y)$ coordinates. It is computed by Eq. (8).

$$\rho_{x,y} = \sum_{x=1}^{j} \sum_{y=1}^{j} \frac{-\gamma^{-1}_{x,y}}{\sqrt{\gamma^{-1}_{y,y} \gamma^{-1}_{x,x}}} \quad (8)$$

![Fig. 1](image1.png) A comparison of the first BIMF in Fourier domain, (a) from PCA and (b) from LPC-PCA.

![Fig. 2](image2.png) A comparison of the residual BIMF in spatial domain, (a) from PCA (local entropy $= 3.92$) and (b) from LPC-PCA (local entropy $= 3.54$).

![Fig. 3](image3.png) A comparison of the difference between (a) global and (b) local means as indicated in red straight lines.
where $\gamma_{x,y}$ is a covariance coefficient between $a_x$ and $a_y$ and $\gamma_{x,y}^{-1}$ is an inverse covariance coefficient of $\gamma_{x,y}$ determined by Eq. (9).

\[
\Gamma = \begin{bmatrix}
\gamma_{1,1} & \gamma_{1,2} & \cdots & \gamma_{1,j} \\
\gamma_{2,1} & \gamma_{2,2} & \cdots & \gamma_{2,j} \\
\vdots & \vdots & \ddots & \vdots \\
\gamma_{j,1} & \gamma_{j,2} & \cdots & \gamma_{j,j}
\end{bmatrix}
\] (9)

In this way, LPC-PCA can fulfill two conditions for decomposing BIMFs effectively. First, low frequency-free condition in the first BIMF is achieved as shown in Fig. 1 (b). Second, rough surface reduction in the residual component is improved as shown in Fig. 2 (b). In other words, the first BIMF in Fourier domain can clearly be removed low frequency better than the first BIMF in Fig. 1 (a) whereas the residual component carried out by LPC-PCA is smoother than Fig. 2 (a) in spatial domain. Moreover, when evaluated by local entropy, the entropy value of Fig. 2 (b) is lower than Fig. 2 (a) in spatial domain. Moreover, when evaluated by local entropy, the entropy value of Fig. 2 (b) is lower than Fig. 2 (a) in spatial domain.

3.2 Iteration-Free Computation for Bi-Dimensional Empirical Mode Decomposition

As described in Sect. 2, the heart of the conventional BEMD methods is a mean surface construction through extrema detection and upper-and-lower surface interpolation. Both of which are extremely time-consuming. In order to overcome this problem, this paper proposes an iteration-free bi-dimensional empirical mode decomposition (iBEMD) method by using LPC-PCA that can construct a direct mean surface; i.e., extrema detection and upper-and-lower surface interpolation is no longer needed. This can reduce the computational cost of the conventional BEMD methods. It can be done by an additional modification defined by Definition 3.

**Definition 3** Let $m_i$ be a mean surface of an image $g_i$ and $r_i$ be a residual component. Each first sifting result of subtracting $m_i$ from $g_i$ is a bi-directional intrinsic mode function (BIMF): (i) if it has a zero mean, (ii) if the maxima are positive and the minima are negative, and (iii) if the number of maxima equals to the number of minima.

Based on Definition 3, iBEMD can instantaneously extract a BIMF from an image. Its algorithm can be described as follows.

**Step 1:** Set an original image to $g_i$ and $i = 1$, where $i$ is an index of BIMF.

**Step 2:** Estimate a mean surface $m_i$ by using LPC-PCA.

**Step 3:** Calculate $h_i = g_i - m_i$.

**Step 4:** Set $\omega_i = h_i$, where $\omega_i$ is a BIMF, $r_i = g_i - \omega_i$, $g_{i+1} = r_i$ and $i = i + 1$.

To estimate the mean surface $m_i$ from the original image $g_i$, LPC-PCA is calculated by two main steps: (i) extracting eigenvector and eigenvalue and (ii) constructing mean surface. In the former step, let $\Gamma_c$ and $\Gamma_r$ be $W \times W$ column and $H \times H$ row covariance matrices defined by Eqs. (10) and (11), respectively.

\[
\Gamma_c = \frac{1}{W-1} (A - M_c)(A - M_c)^T
\] (10)

\[
\Gamma_r = \frac{1}{H-1} (A^T - M_r)(A^T - M_r)^T
\] (11)

where $M_c$ and $M_r$ are $H \times W$ column and $W \times H$ row directions of local means calculated by Eq. (6) with $A$ and $A^T$, respectively.

After that, $\Gamma_c$ and $\Gamma_r$ are employed to compute $\rho_c$ and $\rho_r$ defined by Eq. (7). The column eigenvector $\psi_c$, and eigenvalue $\Lambda_c$ and the row eigenvector $\psi_r$, and eigenvalue $\Lambda_r$ are determined by Eqs. (12) and (13), respectively.

\[
\rho_c \psi_c = \Psi_c \Lambda_c
\] (12)

\[
\rho_r \psi_r = \Psi_r \Lambda_r
\] (13)

where $\rho_c$, $\psi_c$, and $\Lambda_c$ are $W \times W$ matrices while $\rho_r$, $\psi_r$, and $\Lambda_r$ are $H \times H$ matrices. In this paper, QR transformation was used for computing each eigenvalue [21], [22]. This transformation produces eigenvalues by using an iterative function. To extract eigenvalue and eigenvector without using iterative operations, this work sets the maximum round of QR transformation to “1”; that is, we accept a small error in eigenvalue calculation, but the eigenvalue is still sufficient for mean surface interpolation. Moreover, the normal complexity of the conventional PCA is $O(n^3)$ for $n \times n$ matrix while the complexity of the proposed LPC-PCA method based on QR transformation is $O(n^2)$ as reported in [23], [24].

In the latter step, we propose a bi-directional image reconstruction technique to estimate the mean surface with the eigenvector and eigenvalue in both directions. The mean surface $\bar{A}$ in iBEMD is defined by Eq. (14).

\[
\bar{A} = \frac{1}{2} (\Psi_c \psi_c^T + (\Psi_r \psi_r^T)
\] (14)

where $\Psi_c$ and $\Psi_r$ are the column and row directions of the first eigenvector that is selected from the highest eigenvalue. The column and row directions of coefficient matrices, $C_c$ and $C_r$, with size $1 \times H$ and $1 \times W$ can be determined by Eqs. (15) and (16), respectively, such that matrix $A$ is projected onto $\Psi_c$ and $\Psi_r$.

\[
C_c = \Psi_c^T A
\] (15)

\[
C_r = \Psi_r^T A
\] (16)

The results of using either column or row eigenvectors in PCA image reconstruction technique to decompose the first BIMF from the image are illustrated in Figs. 4 (a) and 4 (b). It is found that the column and row eigenvectors provide clear edges of the wood texture only in either the vertical or horizontal directions [25], [26]. On the other hand, the
result of using both eigenvectors in the bi-directional image reconstruction technique; Eq. (14), illustrated in Fig. 4 (c), shows clear edges in both directions. Therefore, for obtaining complete edges of wood texture, the bi-directional image reconstruction technique proves more suitable for estimating the mean surface. The first BIMF in the Fourier domain and the last BIMF in the spatial domain achieved by the proposed iBEMD method are illustrated in Figs. 1 (b) and 2 (b), respectively. For another BIMF, if it is required, Step 2 to 4 are repeated until the number of extrema of \( r_l \) is less than three points.

As described before, the iBEMD method is able to reduce the artifacts in each BIMF and also the computation time. More importantly, researches in bidimensional empirical mode decomposition (BEMD) are based on empirical background. In other words, there has been no theoretical background that directly supports the conventional BEMD method as reported in [27]. However, in 2005, Damerval et al. [7] presented their BIMFs and showed that their BIMFs satisfied certain conditions by exhibiting the number of maxima and minima points in each BIMF for three cases. It was a strong experimental proof. Therefore, this paper follows this prove-by-experiment method as introduced by Damerval [7]. In order to prove that the proposed method is iteration-free computation and our BIMF satisfies Definition 3, the numbers of maxima and minima in three cases, including Gaussian white noise, Lenna image, and Lenna with Gaussian white noise as recommended in [7] are determined. Typically, the numbers of maxima and minima will be nearly equal if the first sifting result already yields the BIMF as mentioned in Sect. 2.2. Therefore, in Fig. 5, the tendency that the numbers of maxima (red line) and minima (blue line) in the first sifting result in a BIMF for all cases is almost similar to those from Damerval’s approach [7]. This demonstrates that the first sifting result of the iBEMD method satisfied Definition 3. In other words, the proposed method is iteration-free computation.

4. Experiments and Discussion

In this section, the performance of the proposed method, iBEMD, is tested and evaluated in terms of (i) the image quality, (ii) computational cost, and (iii) Thai text localization and recognition under illumination effects. Three experiments are set up on each of those aspects as follows.

4.1 Quality of BIMFs

In the first experiment, the quality of BIMFs provided by iBEMD is tested with synthesized images and the progression of BIMFs from the first to the last components is tested with a real image. The results of these two tests are compared against those of the state-of-the-art methods including BEMD-RBF [4], FBEMD [7], FaBEMD [8], and MEMD [6].

4.1.1 Testing with Synthesized Images

Two test images synthesized from a combination of two clearly different contrived frequencies are used to evaluate the capability of decomposition methods. Synthesized image I and II as depicted in Figs. 6 (a) and 7 (a) are constructed from high and low frequency components as demonstrated in Figs. 6 (b) and 6 (c); Figs. 7 (b) and 7 (c) in Fourier domain, and then transformed to spatial domain in Figs. 8 and 9, respectively, since BIMFs are typically compared in spatial domain. Figures 8 (a), 8 (b), and 8 (c) are images of Synthesized Image-I and the high and low frequency components in spatial domain. In the same way, Figs. 9 (a), 9 (b), and 9 (c) are images of Synthesized Image-II and the high and low frequency components in spatial domain. These components are used to evaluate the BIMFs decomposed by all methods in terms of quality index (Qi) value proposed by Wang et al. [28]. The range of Qi values is in interval \([0, 1]\) where ‘1’ implies the best quality, while ‘0’ represents the worst quality. The Qi values can reflect to
Fig. 6 Synthesized Image-I on Fourier domain, (a) original synthesized image, (b) high, and (c) low frequency components.

Fig. 7 Synthesized Image-II on Fourier domain, (a) original synthesized image, (b) high, and (c) low frequency components.

Fig. 8 Synthesized Image-I on spatial domain, (a) original synthesized image, (b) high, and (c) low frequency components.

Fig. 9 Synthesized Image-II on spatial domain, (a) original synthesized image, (b) high, and (c) low frequency components.

Table 2 A comparison between the quality index (Qi) of each BIMF obtained by the proposed method and those obtained by the state-of-the-art methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Synthesized Image-I</th>
<th>Synthesized Image-II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High Freq.</td>
<td>Low Freq.</td>
</tr>
<tr>
<td>iBEMD</td>
<td>0.90</td>
<td>0.89</td>
</tr>
<tr>
<td>BEMD-RBF [4]</td>
<td>0.67</td>
<td>0.03</td>
</tr>
<tr>
<td>FBEMD [7]</td>
<td>0.64</td>
<td>0.00</td>
</tr>
<tr>
<td>FaBEMD [8]</td>
<td>0.80</td>
<td>0.83</td>
</tr>
<tr>
<td>MEMD [6]</td>
<td>0.81</td>
<td>0.70</td>
</tr>
</tbody>
</table>

The purpose of this experiment is to see whether the progression of blurriness of BIMFs decomposed by proposed iBEMD method agreed with the constraint in the clause “Each first sifting result of subtracting $m_i$ from $g_i$ is a bi-directional intrinsic mode function (BIMF)” in Definition 3 that we extend from Definition 2 in order to make our method iteration-free. A wood texture image with a pixel resolution of $128 \times 128$ from Brodatz dataset [29] is used as a test image because it was successfully used for testing BIMFs by BEMD-RBF method [4]. According to the BIMF behavior, the progression of BIMFs of any BEMD methods should be similar to the output of the image successively filtered by a filter bank from high to low frequencies in frequency domain or from fine to coarse scales in spatial domain. Figures 14 to 18 show the progression of all of these methods. It can be seen clearly that the first BIMF of every method is the sharpest; the subsequent BIMFs are successively blurrier; and the last BIMF is almost homogeneous with no trace of the wood texture, strictly following the BIMF behavior. Thus, the results of these experiments demonstrate that the set of BIMFs decomposed by the proposed iBEMD method derived from Definition 3 is valid according to the BIMF behavior. Moreover, each of the successive BIMFs decomposed by iBEMD still retained some resemblance to the wood texture whereas those decomposed by FBEMD and MEMD as shown in Figs. 15 and 17 do not—the shapes in them are clearly artifacts—and while the BIMFs decomposed by BEMD-RBF also retain some

evident image quality discrimination as the human-eye perception can detect. Furthermore, the Qi metric is not responsive to error sensitivity, since it employs structural distortion measurement instead of error measurement as described in [28]. The hypothesis of this experiment is that the efficient method is able to decompose BIMFs from the synthesized images with the high image quality when evaluated by quality index.

Table 2 shows the Qi values of BIMFs decomposed by all of the tested methods. It can be seen that the BIMFs decomposed by the proposed method achieve higher Qi values than those decomposed by the state-of-the-art methods: BEMD-RBF, FBEMD, FaBEMD, and MEMD. Especially, our BIMFs of high and low frequency components exhibit the highest Qi values at 0.90 and 0.89 for Synthesized Image-I as well as 0.84 and 0.93 for Synthesized Image-II, respectively. The higher quality of our BIMFs is visually confirmed by their greater similarity to the original components, as can be observed in Figs. 10 (b) to 13 (b). On the other hand, BIMFs of the other methods shown in Figs. 10 (c-f) to 13 (c-f) are clearly degraded. The degradation can be seen easily in the BIMFs of low frequency component of Synthesized Image-I and -II that are decomposed by BEMD-RBF and FBEMD methods. This degradation is a side effect of the interpolation methods. Because of the more true-to-the-original quality of the BIMFs decomposed by the proposed method, it is deemed a better method than the other tested methods for this purpose.
Fig. 10  An illustration of high frequency component of Synthesized Image-I, (a) original synthesized image and decomposed images (BIMFs) obtained from (b) iBEMD, (c) BEMD-RBF, (d) FBEMD, (e) FaBEMD, and (f) MEMD.

Fig. 11  An illustration of low frequency component of Synthesized Image-I, (a) original synthesized image and decomposed images (BIMFs) obtained from (b) iBEMD, (c) BEMD-RBF, (d) FBEMD, (e) FaBEMD, and (f) MEMD.

Fig. 12  An illustration of high frequency component of Synthesized Image-II, (a) original synthesized image and decomposed images (BIMFs) obtained from (b) iBEMD, (c) BEMD-RBF, (d) FBEMD, (e) FaBEMD, and (f) MEMD.

Fig. 13  An illustration of low frequency component of Synthesized Image-II, (a) original synthesized image and decomposed images (BIMFs) obtained from (b) iBEMD, (c) BEMD-RBF, (d) FBEMD, (e) FaBEMD, and (f) MEMD.

Fig. 14  A set of BIMFs decomposed by BEMD-RBF with the wood texture image.
resemblance to the original texture, it is much more computationally expensive than iBEMD, as explained in the next subsection.

4.2 Computational Cost

This experiment aims to evaluate the computational cost of the proposed method, iBEMD, and compare it to those of the state-of-the-art methods: BEMD-RBF [4], MEMD [6], FBEMD [7], and FaBEMD [8]. Six images of different sizes are used to test the performance: 32×32, 64×64, 128×128, 256×256, 512×512, and 1024×1024 pixels. As illustrated in Fig. 19, the larger the size of the image, the more time it takes to decompose the image. In all cases, iBEMD takes much less computation time than all of the other methods. For example, the computation times that FBEMD takes to decompose the 32×32, 64×64, 128×128, 256×256, 512×512, and 1024×1024 pixel images are 57, 167, 472, 817, 1,202, 6,860, and 16,320 times longer than the computation times that iBEMD takes, respectively. While the computation times that FaBEMD, which is the closest in computational cost performance to iBEMD, takes for the case of the largest image size, the proposed method achieves the decomposition 105, 104, and 103 times faster than FBEMD, FaBEMD and MEMD methods do, respectively, while BEMD-RBF cannot complete it in a reasonable time at all, as explained in [8].

Based on experimental assumption, iBEMD can speed up faster than the state-of-the-art methods due to two reasons: (i) the proposed method can directly estimate the mean surface. This greatly reduces the computation time of mean surface estimation while the almost state-of-the-art meth-
ods estimate both upper and lower surfaces. (ii) The proposed method excludes the phase of finding extrema points. It means that the computational cost of extrema detection of the proposed method is none. On the other hand, the state-of-the-art methods require extrema detection for mean surface estimation. Ordinarily, the time-consuming of this phase depends on image sizes; i.e., the extrema detection requires the high computational cost if the image size is large.

As mentioned in Sects. 4.1 and 4.2, it is summarized that the proposed method, iBEMD, realistically achieves the higher Qi value and extremely reduces the computational cost when compared with the state-of-the-art methods.

4.3 Thai Text Localization and Recognition

The last experiment aims to evaluate the performance of the proposed method applied to Thai text natural scene dataset [30], [31] for overcoming the illumination problems, such as illumination change, shadow, and reflection. These problems have a great impact on performance of Thai text localization and recognition. Hence, this experiment illustrates the accomplishment of iBEMD versus the state-of-the-art methods when applied for such problem images. Fortunately, it is a fact that the illumination component is always included in the last BIMF [10]; therefore, illumination removal is easily done by excluding it from the original image. For this reason, illumination removal is added to the framework of experimental design. The experimental framework of Thai text localization and recognition consists of four procedures: illumination preprocessing, text-background decomposition, Thai text localization, and Thai character recognition as shown in Fig. 20. Based on this scheme, the illumination component in natural scene images is extracted by all BEMD methods in the preprocessing phase for the better text localization and recognition.

4.3.1 Illumination Preprocessing

This procedure is to resolving illumination problems prior to Thai text localization and recognition procedures. Therefore, iBEMD is embedded to the illumination preprocessing and its algorithm is described as follows:

**Step 1:** Transform a given image \( I \) to a YCbCr color model image.

**Step 2:** Select the Y-component \( Y_c \) from the YCbCr color model image to serve as an input of iBEMD algorithm.

**Step 3:** Construct a new Y-component \( Y'_c \) by using the iBEMD algorithm as described in Sect. 3.2 from all BIMFs, except the residual component. It is calculated by Eq. (17).

\[
Y'_c = \frac{U-1}{\sum_{u=1}^{U-1} \omega_u} \cdot \sum_{u=1}^{U-1} \omega_u
\]  

where \( \omega_u \) is the BIMF and \( U \) is the total number of elements in BIMF. The output of this step is the new Y-component \( Y'_c \) whose illumination effects are eliminated by excluding the residual component.

Based on the survey experiments, the suitable number of BIMFs for using to eliminate illumination effects varies depending on the given image. Additionally, the retrieving rate of Thai character under illumination effects relates to the number of BIMFs. In order to discover the appropriate number of BIMFs, the new stopping criterion of iBEMD is proposed by means of variability analysis of an intensity image. Such a criterion is implemented by using the coefficient of variation (CV) defined by Eq. (18).

\[
CV = \frac{\sigma}{\mu}
\]

where \( \mu \) and \( \sigma \) are mean and variance of \( Y_c \), respectively. With this criterion, an iBEMD algorithm is terminated if CV is less than the threshold \( T_{cv} \).

\[
CV < T_{cv}
\]

In this experiment, the width of a nearest vector neighborhood \( w \) in Eq. (4) and coefficient of variation threshold \( T_{cv} \) in Eq. (19) are set to 2.0 and 0.2, respectively, for the best result. Similarly, all parameters of FBEMD, FaBEMD, and MEMD are set to recommended values as reported in [6]–[8].

4.3.2 Text-Background Decomposition

The aim of second procedure is to decrease the number of
noisy or unwanted objects. Adaptive boundary clustering (ABC) [30] is implemented to achieve this objective. Two types of features, text color and text position, are used. The outcomes of this procedure are the number of layers (k-layers) of boundary objects.

4.3.3 Thai Text Localization

Thai text localization is a procedure to identify an individual Thai character location on a natural scene image. To do this, modified connected component analysis (MCCA) proposed by Woraratpanya et al. [31] is used. The MCCA method
is designed for working with the Thai language structure. This procedure provides the set of detected characters. The performance of Thai text localization is evaluated by using precision $P$, recall $R$, and $f$-value $F$ as defined by Eqs. (20), (21), and (22), respectively.

$$P = \frac{tp}{tp + fp} \tag{20}$$

$$R = \frac{tp}{tp + fn} \tag{21}$$

$$F = \frac{2 \times P \times R}{P + R} \tag{22}$$

where $tp$ and $fp$ are a number of correctly detected characters and non-characters, respectively, and $fn$ is a number of incorrectly detected characters.

4.3.4 Thai Character Recognition

Transforming a Thai character image to editable characters is a goal of Thai character recognition. Commonly, this procedure has two main steps: feature extraction and classification. Here, an adaptive histogram of oriented gradient (AHOG) proposed by Woraratpanya et al. [32] is used for feature extraction whereas Euclidean distance is used for classification. The recognition rate, $acc$, is measured by Eq. (23).

$$acc = \frac{c_1}{c_2} \times 100 \tag{23}$$

where $c_1$ is a number of correctly classified characters and $c_2$ is a total number of input characters.

According to [30], [31] discussed on an illumination problem on natural scene image, the number of detected Thai characters retrieved from Thai text localization framework is always low due to the illumination effect. In this work, the performance of all BEMD methods in reducing this effect is measured in terms of precision, recall, and $f$-value. When a scene image has illumination effects as shown in Fig. 21, the main metric used for evaluating the performance of each BEMD method is recall value, which is always low. In the performance test, the state-of-the-art methods and the proposed method are applied to Thai text localization and recognition framework as schematically displayed in Fig. 20, and the results are shown in Figs. 21 (a), 21 (b), 21 (c), 21 (d), and 21 (e). As shown in Table 3, three BEMD methods can reduce the illumination effect from the natural scene image, but to a different extent: FBEMD, FaBEMD, and iBEMD methods achieve a recall value of 55.64%, 59.79%, and 73.76%, respectively. (recall value is 51.67% without effect elimination by a BEMD). Even though the precision values achieved by FBEMD, FaBEMD and iBEMD methods are 5.22%, 1.02%, and 7.28%, respectively, lower than the value achieved by the baseline method without effect elimination, the $f$-values that indicate the average Thai text localization performance actually increased by 0.46%, 4.66%, and 8.50%. Overall, the performance of Thai text localization increased when any of the three BEMD methods are applied. Especially, the proposed method can greatly improve the performance of Thai text localization when compared to the state-of-the-art methods in terms of recall and $f$-value. The result in Fig. 21 (e) shows that the number of detected characters obtained from the proposed method is more than that of the state-of-the-art methods as displayed in Figs. 21 (b), 21 (c), and 21 (d).

Lastly, in Thai character recognition, a number of correctly detected characters and correctly classified characters are measured and evaluated in terms of recognition rate. In Table 3, the experimental results show that three BEMD methods: FBEMD, FaBEMD and iBEMD improve from the baseline of 78.06% (without using BEMD) to 78.15%, 79.90%, and 80.97%, respectively. Our proposed method’s improvement of 2.91% in recognition rate may look slightly different, but the achievement of iBEMD in making 461 more characters recognizable is very impressive (as far as character recognition goes). This improvement mainly comes from the illumination component removal. It is evident that the number of correctly detected characters greatly accomplishes when the BEMD methods are embedded to the illumination preprocessing procedure. Especially, the proposed method achieves 2,063 correctly detected characters—an increase of 580 more detectable characters than the baseline. The recall also proves that the higher value, the more characters are correctly detected. However, the MEMD method does not achieve in this work, since it is proposed for one-dimensional data, not for bi-dimensional data.

Table 3 A comparison of the performance in Thai text localization and recognition between the state-of-the-art and the proposed methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Text localization</th>
<th>Character recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision (%)</td>
<td>Recall (%)</td>
</tr>
<tr>
<td>FBEMD [4]</td>
<td>65.50</td>
<td>55.64</td>
</tr>
<tr>
<td>FaBEMD [8]</td>
<td>69.70</td>
<td>59.79</td>
</tr>
<tr>
<td>MEMD [6]</td>
<td>29.50</td>
<td>15.64</td>
</tr>
<tr>
<td>iBEMD</td>
<td>63.44</td>
<td>73.76</td>
</tr>
</tbody>
</table>
5. Conclusions

In this paper, an iteration-free bi-dimensional empirical mode decomposition (iBEMD) has been proposed for reducing computational cost and artifacts occurred from BIMF extraction. The novel contribution of the proposed method is the use of locally partial correlation for principal component analysis (LPC-PCA) to decompose a set of BIMFs in the first sifting result. Based on experiments, the proposed method achieves both computational cost and artifact reductions. It is due to the following reasons: (i) LPC-PCA inherited from original PCA is a low computational cost method and efficiently estimates a more accurate mean surface for BIMF extraction without using extrema detection. (ii) LPC-PCA can also well transform correlated variables into a set of linearly uncorrelated variables. This leads to artifact reduction on BIMFs. Additionally, when iBEMD is applied to Thai text localization and recognition to solve illumination effects on natural scene, its performance greatly improves, since it can clearly decompose illumination component on the last BIMF. This is an effective way to handle illumination problem.

In future work, the proposed method will be implemented in conjunction with Hilbert transform to make many benefits, for instance, identifying the possible text area on natural scene for text localization, and adaptive weighting feature area to discriminate the similar characters for recognition.

Acknowledgements

This research was supported by a research grant No.2559-06-002 from Faculty of Information Technology, King Mongkut’s Institute of Technology Ladkrabang, Thailand.

References


Taravichet Titijaroonroj received the B.S. and M.S. degrees in Information Technology from King Mongkut’s Institute of Technology Ladkrabang, Bangkok, Thailand in 2013 and 2014, respectively. He is currently working for Ph.D. degree in Information Technology from King Mongkut’s Institute of Technology Ladkrabang, Bangkok, Thailand. His research interests include signal analysis, machine learning, computer vision, pattern recognition and image processing.

Kuntpong Woraratpanya received his B.Ind.Tech. in computer technology, M.Eng in Computer Engineering, and D.Eng in Electrical Engineering, all from King Mongkut’s Institute of Technology Ladkrabang, Bangkok, in 1992, 1996, and 2005, respectively. He is currently an assistant professor at Faculty of Information Technology, King Mongkut’s Institute of Technology Ladkrabang, Bangkok, Thailand. His research interests include stereoscopic acquisition and compression, multimedia coding and processing, digital signal processing, speech recognition and processing, pattern recognition and image processing, computer vision, and machine learning.