IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 26, 2019, accepted August 12, 2019, date of publication August 19, 2019, date of current version August 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2936030

Bidimensional Multivariate Empirical Mode Decomposition With Applications in Multi-Scale Image Fusion

YILI XIA[1], (Member, IEEE), BIN ZHANG[1], (Student Member, IEEE), WENJIANG PEI[1], AND DANILO P. MANDIC[2], (Fellow, IEEE)

[1]School of Information Science and Engineering, Southeast University, Nanjing 210096, China
[2]Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K.

Corresponding authors: Yili Xia (yili.xia06@gmail.com) and Wenjiang Pei (wjpei@seu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771124, and in part by the Zhi Shan Young Scholar Program of Southeast University.

ABSTRACT Empirical mode decomposition (EMD) is a fully data-driven technique designed for multi-scale decomposition of signals into their natural scale components, called intrinsic mode functions (IMFs). When EMD is directly applied to perform fusion of multivariate data from multiple and heterogeneous sources, the problem of uniqueness, that is, different numbers of decomposition levels for different sources, is likely to occur, due to the empirical nature of EMD. Although the multivariate EMD (MEMD) has been proposed for temporal data, which employs real-valued projections along multiple directions on a unit hypersphere in the n-dimensional space to calculate the envelope and the local mean of multivariate signals, in order to guarantee the uniqueness of the scales, its direct usefulness in 2D multi-scale image fusion is still limited, due to its inability to maintain the spatial information. To address this issue, we propose a novel bidimensional MEMD (BMEMD) which directly projects a bidimensional multivariate signal, which is composed of multiple images, on the unit hypersphere in the n-dimensional space. This is achieved via real-valued surface projections and the mean surface is estimated by interpolating the multivariate scatter data so as to extract common spatio-temporal scales across multiple images. Case studies involving texture analysis and multi-focus image fusion are presented to demonstrate the effectiveness of the proposed method.

INDEX TERMS Empirical mode decomposition (EMD), bidimensional multivariate EMD (BMEMD), real-valued surface projections, multi-scale image fusion.

I. INTRODUCTION

Image fusion is a process of gathering salient features from multiple images to produce a single "fused" image, which is especially important in situations where optical cameras, due to the limited depth of focus, cannot be focused simultaneously on all objects at different distances to gain a clear image [1]. In such cases, multi-scale images are obtained with each of them containing partial information of a scene of interest, which can be further merged to present the complete information in the fused image. To fuse multiple images, conventional solutions are established based on the assumptions that data exhibit some structure (linearity, sparsity) and on the subsequent applications of projections onto a set of predefined basis functions in spatial and/or frequency domain [2]–[6]. However, the static filter banks and/or fixed basis functions within the existing approaches prevent fusion of intrinsic and matched spatial frequency content among multi-scale input images.

Empirical mode decomposition (EMD) is a fully data-driven technique, which adaptively decomposes a multi-scale signal into a finite set of amplitude- and/or frequency-modulated (AM/FM) components, called intrinsic mode functions (IMFs), and a residual component [7]–[9]. Unlike Fourier or wavelet based methods, EMD does not impose *a priori* assumptions about the data while decomposing signals, and hence, it is particularly suitable for the time-frequency analysis of real-world nonlinear and nonstationary

The associate editor coordinating the review of this article and approving it for publication was Yingsong Li.

signals. Owing to the excellent characterization of intrinsic scales at local level by EMD, both its multivariate extensions and bidimensional ones [10]–[20] have been widely applied in heterogeneous image fusion [21]–[29], for which, a set of common frequency scales must be determined beforehand.

Multivariate extensions of EMD based data fusion schemes aim to address the problem of uniqueness within the original EMD by considering a multichannel signal as a whole and using multiple real-valued projections to find the local mean of the original signal, a key issue to find physically meaningful IMFs [16]–[20]. This is particularly important, given that univariate EMD processes multichannel signals component-wise, it cannot guarantee that decompositions of different data sources are matched, either in number or properties of local scales, making a multi-scale comparison often difficult. However, a prerequisite to implement these EMD extensions in 2D image processing is to vectorize multiple images by concatenating their columns/rows to form a one-dimensional multivariate signal [21]–[23]. This preprocessing, however, unavoidably breaks up the spatial correlation within the original images and results in artifacts in the scale images. On the other hand, although bidimensional EMD (BEMD) is able to directly decompose an image into several 2D IMFs via a 2D-sifting process, so as for better maintaining of the image spatial information [10], [11], the problem of uniqueness still remains due to the univariate nature of BEMD, which yields inconsistent numbers and scales of IMFs for different images.

To this end, we propose a bidimensional multivariate EMD (BMEMD) method, which possesses both the capability of multivariate EMD (MEMD) to address the problems of uniqueness and mode-mixing and the 2D processing nature of BEMD, by directly projecting a bidimensional multivariate signal, e.g., composed of multiple images, on the unit hypersphere in the multidimensional space via novel real-valued surface projections and estimating the mean surface by interpolating the multivariate scatter data to extract IMFs with matched scales across data channels. Simulations involving texture analysis and multi-focus image fusion demonstrate the effectiveness of the proposed method.

II. ORIGINAL MEMD

The key issue to implement EMD is the computation of local mean of the original signal, a step which critically depends on finding local maxima and minima. For multivariate signals, however, it is not straightforward to detect the local extrema and to estimate the mean envelope, since the fields of complex and hyper-complex numbers are not ordered. MEMD overcomes this difficulty by employing real-valued projections in an n-dimensional space, known as *n-dimensional real-valued projections*, where projection direction vectors start from the origin of n-dimensional coordinates and end at the points, which are uniformly distributed on the unit sphere, also known as the $(n-1)$-sphere, in the n-dimensional space [16]–[20]. Two approaches have been provided in [16] on how to generate uniform point sets on the $(n-1)$-spheres, which employ either uniform angular coordinates or low-discrepancy point sets stemming from quasi-Monte Carlo methods in order to obtain a suitable set of direction vectors. Once the projection matrix $\mathbf{U} = \left[\mathbf{u}^{\theta_1}, \mathbf{u}^{\theta_2}, \ldots, \mathbf{u}^{\theta_K}\right]$ is obtained for the original n-variate signal $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_n(t)]^T$, where $\mathbf{u}^{\theta_k} = \left[u_1^{\theta_k}, u_2^{\theta_k}, \ldots, u_n^{\theta_k}\right]^T$ is kth projection vector along the angle θ_k on the unit $(n-1)$-sphere for $k = 1, 2, \ldots, K$ and K is the total number of projection vectors, the n-dimensional real-valued projection of $\mathbf{s}(t)$ on $\mathbf{u}^{\theta_k}$ can be achieved as

$$p^{\theta_k}(t) = \mathbf{s}^T(t)\mathbf{u}^{\theta_k} = \sum_{l=1}^{n} s_l(t) u_l^{\theta_k}. \tag{1}$$

The iterative decomposition process of MEMD for general n-variate signals can be described in Algorithm 1.

Algorithm 1 The Original MEMD Algorithm

1. Calculate projections of $\mathbf{s}(t)$ along projection vectors $\mathbf{u}^{\theta_k}$, denoted by $p^{\theta_k}(t)$.
2. Extract time instants t^{θ_k} which correspond to maxima of the projection signal $p^{\theta_k}(t)$.
3. Interpolate $\left[t^{\theta_k}, \mathbf{s}(t^{\theta_k})\right]$ to obtain the multivariate envelope $\mathbf{e}^{\theta_k}(t)$, where $\mathbf{e}^{\theta_k}(t) = [e_1^{\theta_k}(t), e_2^{\theta_k}(t), \cdots, e_n^{\theta_k}(t)]$ and $e_l^{\theta_k}(t)$ refers to the envelope of $s_l(t)$ along the projection vector $\mathbf{u}^{\theta_k}$, for $l = 1, 2, \ldots, n$.
4. Repeat the above steps for all k from 1 to K.
5. Calculate the mean envelope of $\mathbf{s}(t)$, that is, $\mathbf{m}(t)$, as
$$\mathbf{m}(t) = \frac{1}{K}\sum_{k=1}^{K} \mathbf{e}^{\theta_k}(t). \tag{3}$$
6. Extract the detail $\mathbf{d}(t)$ using $\mathbf{d}(t) = \mathbf{s}(t) - \mathbf{m}(t)$. If $\mathbf{d}(t)$ fulfills the stoppage criterion for a multivariate IMF, apply the above procedure to $\mathbf{s}(t) - \mathbf{d}(t)$, otherwise, apply it to $\mathbf{d}(t)$.

Now, consider a 2D n-variate signal $\mathbf{I}$, which, for instance, consists of n heterogeneous images. Then, its lth channel, that is, $\mathbf{I}_l$, is given by

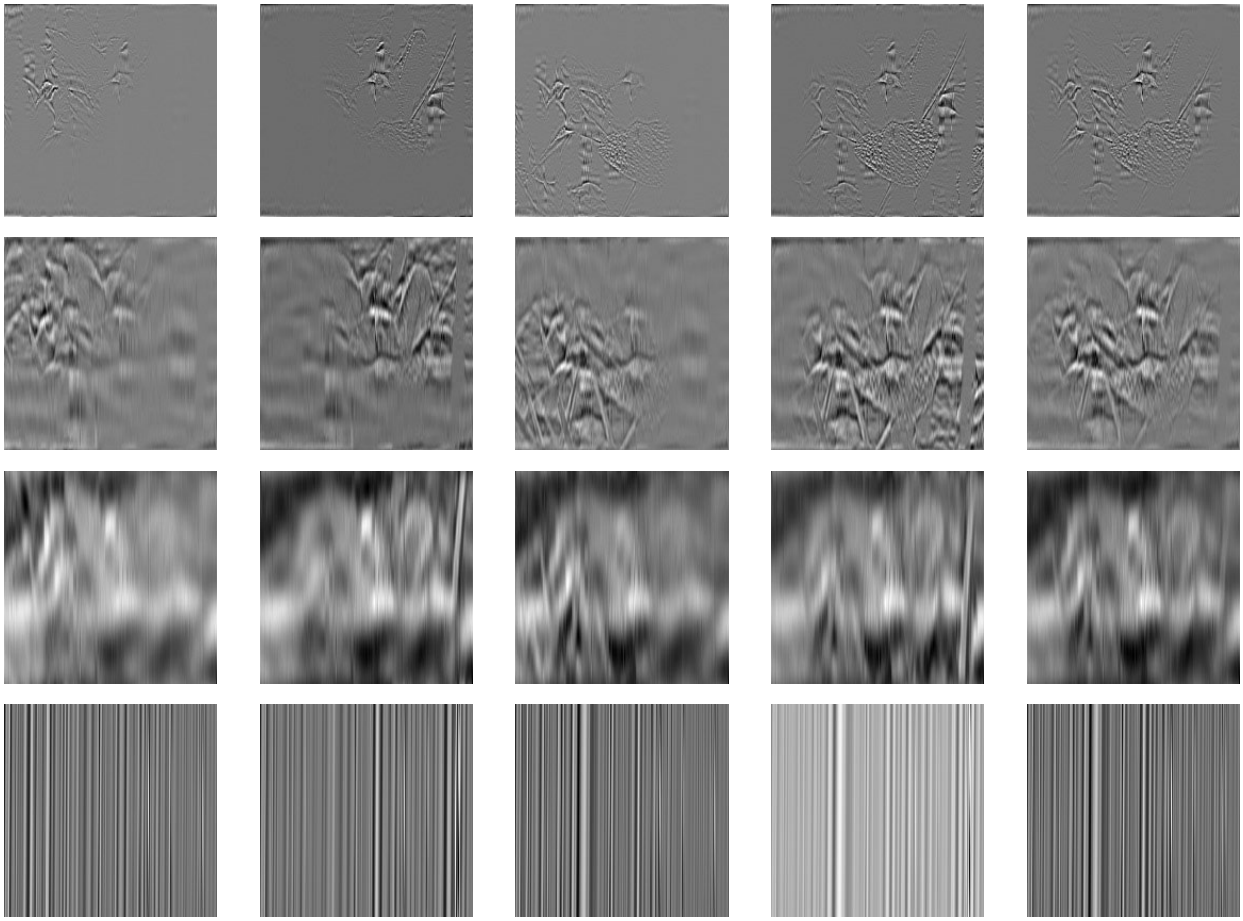
$$\mathbf{I}_l = \begin{bmatrix} I_l(1,1) & I_l(1,2) & \cdots \\ I_l(2,1) & I_l(2,2) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, \tag{2}$$

where $I_l(i,j)$ is the value of the pixel at the coordinate (i,j). When dealing with such a 2D n-variate signal, the original MEMD first vectorizes the lth image by concatenating its pixels by columns/rows to form $s_l(t)$ for $l = 1, 2, \ldots, n$ [23], [24]. After obtaining the 1D n-variate signal $\mathbf{s}(t)$, the iterative decomposition steps discussed above are subsequently applied on $\mathbf{s}(t)$ to obtain 1D n-variate IMFs; a set of common frequency scales prepared for multi-scale image fusion.

Although the simulations in [23], [24] illustrate the potentiality of MEMD in fusion of multi-focus images and pan-

(a) Image 1 (b) Image 2 (c) Image 3 (d) Image 4 (e) Image 5

FIGURE 1. Artificially generated multi-focus images of *Iris ensata*. The focus on the images is respectively in the upper left, the upper right, the lower left, the lower right, and the center of the image.



(a) 1st channel of multivariate IMFs (b) 2nd channel of multivariate IMFs (c) 3rd channel of multivariate IMFs (d) 4th channel of multivariate IMFs (e) 5th channel of multivariate IMFs

FIGURE 2. Partial IMFs obtained by MEMD for the five multi-focus images shown in Fig. 1, where (a), (b), (c), (d), and (e) represent the five data channels, respectively. Originally, MEMD yielded 17 IMFs. For illustration purpose, several IMFs were accumulated together, and the first four sets of IMFs are presented. Images in the first row correspond to the sum of first 3 IMFs in different channels, images in the second row represent the sum of the 4th to the 6th IMFs, images in the third and fourth rows are the sum of the 7th to the 9th IMFs and that of the 10th to the 12th IMFs.

sharpening of multi-spectral images, an inherent obstacle encountered by applying a 1D algorithm on image processing lies in the loss of spatial information within original images. When columns of pixels are concatenated, spatial relations between pixels in neighboring rows are broken up and vertical artifacts are likely to appear in IMF images of one dimensional EMD based fusion methods [21]. This deficiency becomes clear when the original MEMD was applied to decompose the five multi-focus images of *Iris ensata* in Fig. 1. Its partial IMF components are illustrated in Fig. 2 in a descending order of the spatial frequency by rows. Obviously, the detail of flower leaves in the upper panels is difficult to recognize due to the contamination from vertical stripes. These artifacts arose because MEMD cannot maintain the spatial correlation within 2D images, and they became more pronounced in the low-frequency IMF components, as shown in the lower panels of Fig. 2. This is then expected that the undesirable decomposition deficiency of MEMD unavoidably deteriorates the quality of multi-scale image fusion.

III. PROPOSED BIDIMENSIONAL MEMD

Inspired by n-dimensional real-valued projections used in MEMD, the proposed bidimensional MEMD (BMEMD) first converts the bidimensional n-variate signal $\mathbf{I}$ into several bidimensional and univariate signals via novel *n-dimensional surface projections*. Recall that in (2), the lth channel of $\mathbf{I}$, that is, $\mathbf{I}_l$, can be regarded as a 1-dimensional surface, so that, $\mathbf{I}$ is an n-dimensional one. In this way, given a unit projection vector $\mathbf{u}^{\theta_k}$ along the angle θ_k and distributed on the unit $(n-1)$-sphere, the projection of $\mathbf{I}$ on $\mathbf{u}^{\theta_k}$, denoted by $\mathbf{P}^{\theta_k}$, can be described as

$$\mathbf{P}^{\theta_k} = \begin{bmatrix} \sum_{l=1}^{n} u_l^{\theta_k} I_l(1,1) & \sum_{l=1}^{n} u_l^{\theta_k} I_l(1,2) & \cdots \\ \sum_{l=1}^{n} u_l^{\theta_k} I_l(2,1) & \sum_{l=1}^{n} u_l^{\theta_k} I_l(2,2) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}. \tag{4}$$

Similar to the original MEMD, the mean surface of $\mathbf{I}$ can now be estimated by averaging all mean surfaces of its projections obtained by using (4), in which the uniformity of the direction set on the $(n-1)$-sphere plays a important role. As discussed in [16], and two solutions exist to generate a suitable set of direction vectors, based on either uniform angular coordinates or low-discrepancy point sets which stem from quasi-Monte Carlo simulations. It is shown that a set of direction vectors based on uniform sampling in the angular coordinate system is convenient to deal with, however, it yields non-uniformly distributed direction vectors, for which, a further calibration procedure is required. On the other hand, the approach based on low-discrepancy pointsets [30] provides a more uniform distribution of direction vectors [31], and hence, it is employed in the proposed BMEMD to yield projection vectors for more accurate local mean estimates on the $(n-1)$-sphere.

A convenient way to generate multidimensional low-discrepancy sequences is to use the Hammersley sequence, which is proven to show considerable improvement in terms of error bounds, over standard Monte Carlo methods. Moveover, the set of direction vectors generated by the Hammersley sequence also yields improved generalized discrepancy estimates as compared with other sampling methods, and hence, are uniformly distributed on a sphere; for more detail, we refer to [31]. In order to generate the Hammersley sequence, we first briefly refresh the concept of *radical inversion*. Let c be an integer greater than zero, its base-b representation is given by

$$c = \sum_{\lambda=0}^{L-1} a_\lambda b^\lambda. \tag{5}$$

Accordingly, the base-b radical inversion of c, denoted by g_c^b, is defined as [32]

$$g_c^b = \sum_{\lambda=0}^{L-1} a_\lambda b^{-\lambda-1}. \tag{6}$$

For example, we have $g_1^2 = 0.5$, $g_2^2 = 0.25$, and $g_3^2 = 0.75$. Next, since the total number of projection directions K can be predefined *a priori* and $K \geq n$, the lth sample, where $l = 1, 2, \ldots, n$, within the Hammersley sequence, denoted by $\mathbf{h}_l^r$, is calculated as

$$\mathbf{h}_l^r = \left[\frac{l}{K}, g_l^{b_1}, \ldots, g_l^{b_{K-1}}\right], \tag{7}$$

where the radicals $b_1, b_2, \ldots, b_{K-1}$ are typically the first $K-1$ prime numbers. In this way, the whole Hammersley sequence can be expressed in an $n \times K$ matrix $\mathbf{H}$, given by

$$\begin{aligned} \mathbf{H} &= \left[(\mathbf{h}_1^r)^T, (\mathbf{h}_2^r)^T, \ldots, (\mathbf{h}_n^r)^T\right]^T \\ &= \begin{bmatrix} \frac{1}{K} & g_1^{b_1} & \cdots & g_1^{b_{K-1}} \\ \frac{2}{K} & g_2^{b_1} & \cdots & g_2^{b_{K-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{n}{K} & g_n^{b_1} & \cdots & g_n^{b_{K-1}} \end{bmatrix} \end{aligned} \tag{8}$$

On the other hand, the Hammersley sequence matrix $\mathbf{H}$ in (8) can be viewed column by column as

$$\mathbf{H} = \left[\mathbf{h}_1^c, \mathbf{h}_2^c, \ldots, \mathbf{h}_K^c\right]. \tag{9}$$

Each column vector $\mathbf{h}_k^c$, where $k = 1, 2, \ldots, K$, contains the coordinates of a point in $\mathbb{R}^n$, and each coordinate is guaranteed to span over $(0, 1]$ [33]. Next, in order to obtain the projection vectors uniformly distributed on the unit $(n-1)$-sphere, the points generated by the Hammersley sequence should be further normalized as follows [33]. Firstly, let $\mathbf{h}_k^c = [x_1, x_2, \ldots, x_n]^T$, the corresponding angular coordinates, that is, $\boldsymbol{\theta}_k = [\varphi_1, \varphi_2, \ldots, \varphi_{n-1}]^T$, are calculated as

$$\begin{cases} \varphi_1 = \tan^{-1}\left(\dfrac{\sqrt{(2x_n-1)^2 + (2x_{n-1}-1)^2 + \cdots + (2x_2-1)^2}}{2x_1 - 1}\right) \\ \varphi_2 = \tan^{-1}\left(\dfrac{\sqrt{(2x_n-1)^2 + (2x_{n-1}-1)^2 + \cdots + (2x_2-1)^2}}{2x_2 - 1}\right) \\ \quad \cdots \\ \varphi_{n-2} = \tan^{-1}\left(\dfrac{\sqrt{(2x_n-1)^2 + (2x_{n-1}-1)^2}}{2x_{n-2} - 1}\right) \\ \varphi_{n-1} = \tan^{-1}\left(\dfrac{2x_n - 1}{2x_{n-1} - 1}\right) \end{cases} \tag{10}$$

where $\tan^{-1}(\cdot)$ is the inverse function of $\tan(\cdot)$.

Algorithm 2 The Proposed Bidimensional MEMD (BMEMD) Algorithm

1.

1) Calculate the projections of $\mathbf{I}$ along the unit projection vector $\mathbf{u}^{\theta_k}$ by using (4), denoted by $\mathbf{P}^{\theta_k}$.
2) Extract the locations $\left(i^{\theta_k}_{max}, j^{\theta_k}_{max}\right)$ and $\left(i^{\theta_k}_{min}, j^{\theta_k}_{min}\right)$, which correspond to the local maxima and the local minima of $\mathbf{P}^{\theta_k}$, respectively.
3) Interpolate $\mathbf{I}(i^{\theta_k}_{max}, j^{\theta_k}_{max})$ and $\mathbf{I}(i^{\theta_k}_{min}, j^{\theta_k}_{min})$ to obtain the n-dimensional maximal and minimal surfaces, denoted by $\mathbf{e}^{\theta_k}_{max}$ and $\mathbf{e}^{\theta_k}_{min}$, respectively.
4) Repeat the above steps for all k from 1 to K.
5) Estimate the n-dimensional mean surface of $\mathbf{I}$, that is, $\mathbf{M}$, as

$$\mathbf{M} = \frac{1}{2K}\sum_{k=1}^{K}\left(\mathbf{e}^{\theta_k}_{max} + \mathbf{e}^{\theta_k}_{min}\right). \tag{12}$$

6) Extract the detail $\mathbf{D}$ using $\mathbf{D} = \mathbf{I} - \mathbf{M}$. If $\mathbf{D}$ fulfills the 2D stoppage criterion [10], [11] for a bidimensional n-variate IMF, apply the above procedure to $\mathbf{I} - \mathbf{D}$, otherwise apply it to $\mathbf{D}$.

Then, the n-dimensional projection vector $\mathbf{u}^{\theta_k}$ which corresponds to $\mathbf{h}^c_k$ is computed as

$$\begin{cases} u^{\theta_k}_1 = \cos(\varphi_1) \\ u^{\theta_k}_2 = \sin(\varphi_1)\cos(\varphi_2) \\ u^{\theta_k}_3 = \sin(\varphi_1)\sin(\varphi_2)\cos(\varphi_3) \\ \cdots \\ u^{\theta_k}_{n-1} = \sin(\varphi_1)\cdots\sin(\varphi_{n-2})\cos(\varphi_{n-1}) \\ u^{\theta_k}_n = \sin(\varphi_1)\cdots\sin(\varphi_{n-2})\sin(\varphi_{n-1}) \end{cases} \tag{11}$$

The original bidimensional multivariate signal $\mathbf{I}$ can now be converted into K bidimensional univariate signals, i.e., $\mathbf{P}^{\theta_k}$, for $k = 1, 2, \ldots, K$, by projecting $\mathbf{I}$ along $\mathbf{u}^{\theta_k}$ via (4). The neighboring window method is next employed to identify local maximum and minimum pixels of each $\mathbf{P}^{\theta_k}$ in the sense that a pixel is considered as a local maximum (minimum), if its value is strictly higher (lower) than all of its neighbors. This facilitates the construction of the maximal surface and the minimal one corresponding to each $\mathbf{u}^{\theta_k}$ by interpolating the extreme scatter data using Delaunay triangulation [34], [35] in order to compute the mean surface and extract bidimensional multivariate IMFs. The above operation can be considered as a multivariate extension of the 2D-sifting process within BEMD [10], [11]. The detailed iterative process of the proposed bidimensional MEMD (BMEMD) to extract IMFs is outlined in Algorithm 2.

IV. SIMULATIONS

Simulations were next conducted to investigate the potentiality of the proposed BMEMD on multi-scale multi-focus image fusion.[1] For all the images, the low-discrepancy Hammersley sequence was used to generate a set of $K = 16$ direction vectors for taking signal projections.

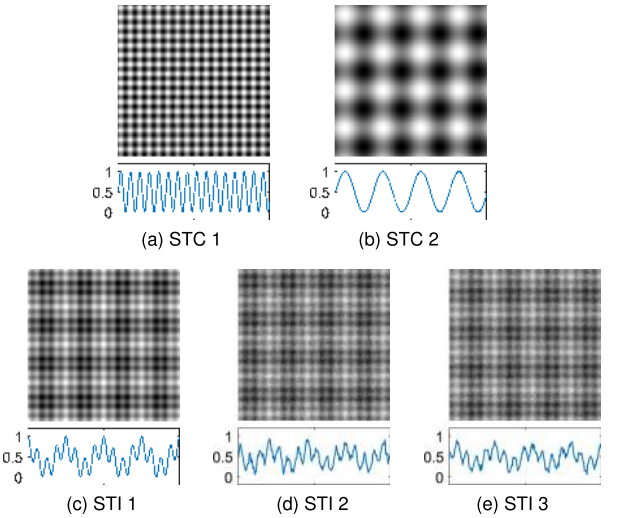


FIGURE 3. Two synthetic texture components (STCs) and three synthetic texture images (STIs), as well as their corresponding 1D diagonal intensity profiles (DIPs).

A. TEXTURE ANALYSIS

In most EMD based fusion applications, a major requirement lies in a careful synchronization and alignment of information among different data channels. However, applying univariate EMD methods separately to different data channels typically fails to align (synchronize) the respective common oscillatory modes and also suffers from the mode-mixing problems, represented by the mixing of different frequency scales within a single IMF and/or a single scale spread across different IMFs [16], [17], [21], since there exists a prerequisite that same-index IMFs contain the information pertaining to the same scale when using EMD based data fusion schemes.

In order to illustrate the problems discussed above, similar to the analysis in [36], [37], we considered two synthetic texture components (STCs), that is, STC 1 and STC 2, which were generated from horizontal and vertical sinusoidal waveforms with different spatial frequencies and are shown in the upper panels of Fig. 3(a) and (b), respectively. Their corresponding 1D diagonal intensity profiles (DIPs) are plotted in the lower panels, whose position started from the upper left corner of the image. Three synthetic texture images (STIs) were constructed as weighted combinations of the two STCs, where their horizontal and vertical sinusoidal waveforms were shifted by random phase values. Moreover, STI 2 and STI 3 were corrupted by different realizations of additive white Gaussian noise with the same statistics (mean, variance). The considered three STIs and their DIPs are shown in Fig. 3(c), (d) and (e), respectively. By applying

[1]The Matlab source code for BMEMD is downloadable at https://www.mathworks.com/matlabcentral/fileexchange/72343-bidimensional-multivariate-empirical-mode-decomposition.

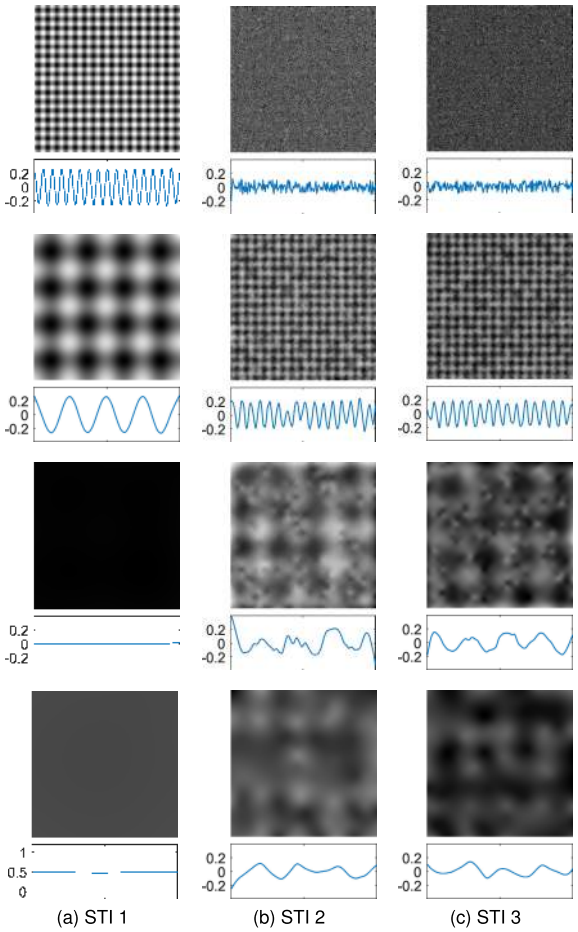


FIGURE 4. IMFs and their corresponding DIPs, obtained by applying BEMD to the three STIs.

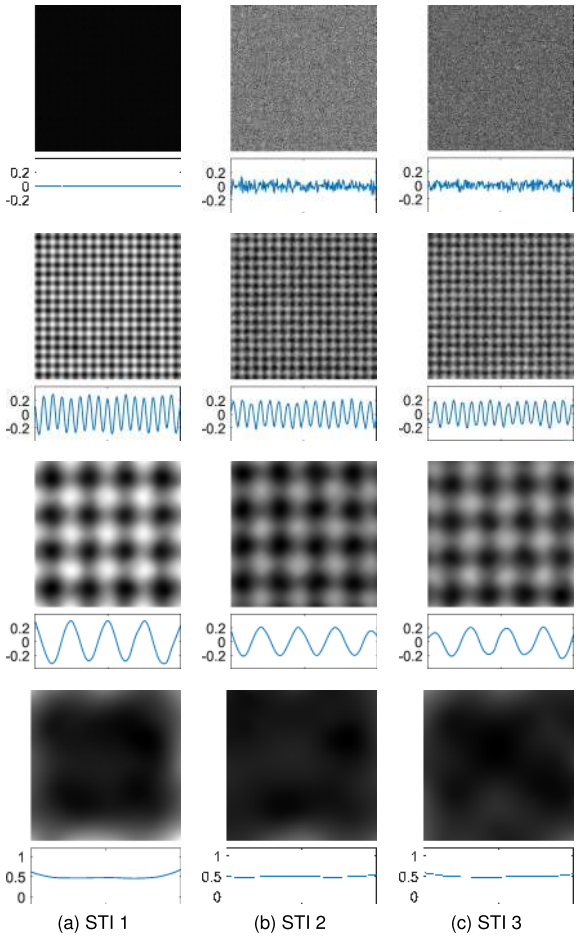


FIGURE 5. IMFs and their corresponding DIPs, obtained by the proposed BMEMD applied on the same STI set.

BEMD to STIs separately, different numbers of IMF images were obtained in each case, due to the nonuniqueness of the resulting decompositions. Fig. 4 shows all the three IMFs and a residue for STI 1 and the first four IMFs for STIs 2 and 3. Mode mixing is clearly visible between IMFs 3 and 4 in both STIs 2 and 3, where the low frequency texture information spread across. Moreover, the three sets of IMF images are not aligned. On the other hand, the proposed BMEMD is able to attenuate the above problems of mode mixing and mode misalignment within IMFs owing to the manner in which the 2D images are collectively processed in a multidimensional nature. This is supported by Fig. 5, where mode mixing has been largely eliminated within the resulting multivariate IMFs. In addition, IMFs from different image channels are also shown to be well aligned with respect to their respective scales. Both features together enable the proposed BMEMD to provide a meaningful comparison between scales and to form a robust basis for image fusion. Note that, although MEMD possesses similar mode alignment and mode matching capabilities, they make less sense in image fusion due to the inability of MEMD to maintain the spatial information of 2D data, as shown in Fig. 2.

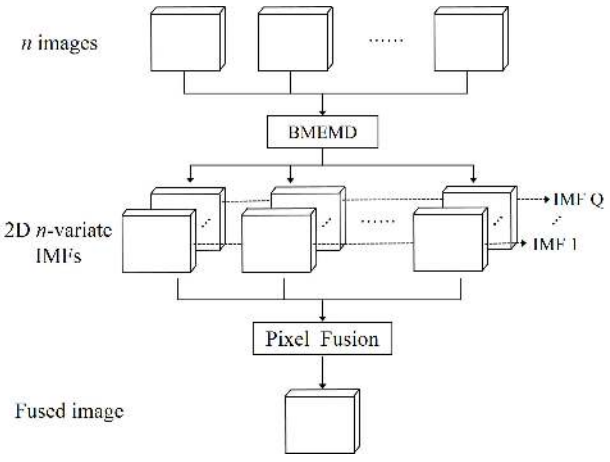
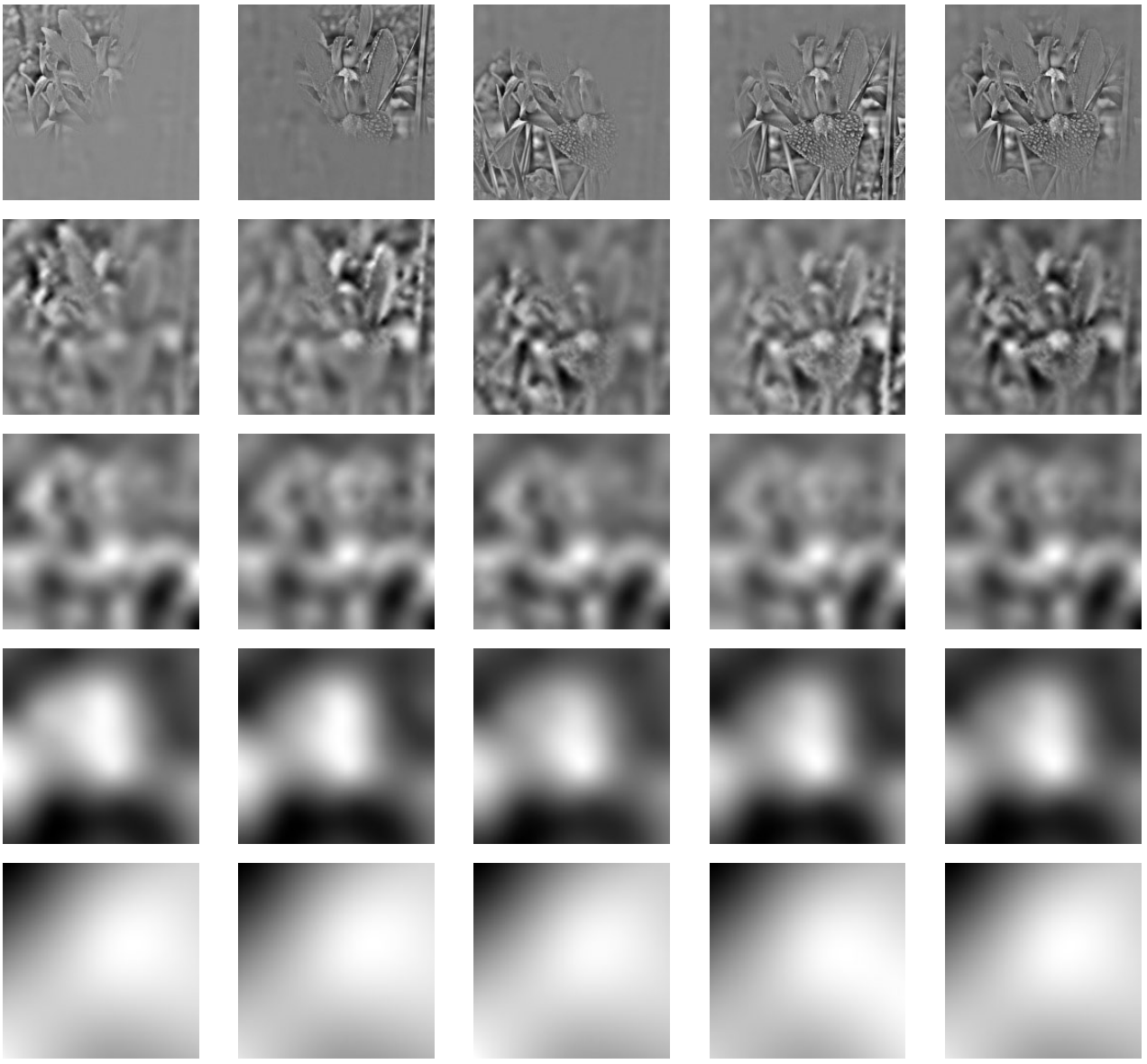


FIGURE 6. Proposed framework for pixel-level multi-scale image fusion based on BMEMD.

B. MULTI-SCALE IMAGE FUSION

The framework of the proposed BMEMD based multi-scale image fusion at the pixel level is illustrated in Fig. 6. A bidimensional multivariate signal $\mathbf{I}$, which is composed of n images, is first decomposed by the proposed BMEMD method to obtain a set of Q n-variate IMF images, denoted by $\mathbf{I}_l^q$, where $q = 1, 2, \ldots, Q$ and $l = 1, 2, \ldots, n$. For



(a) 1st channel of 2D multivariate IMFs (b) 2nd channel of 2D multivariate IMFs (c) 3rd channel of 2D multivariate IMFs (d) 4th channel of 2D multivariate IMFs (e) 5th channel of 2D multivariate IMFs

FIGURE 7. Decomposition results by the proposed BMEMD for the five multi-focus images in Fig. 1, where (a), (b), (c), (d), and (e) represent the five channels of the four 2D multivariate IMFs and one residue component obtained by BMEMD, respectively. The spatial frequencies of IMF images decrease gradually from top to bottom.

mathematical convenience, the residue is considered as the last IMF. In the next stage, an efficient windowed-based weighting averaging method for pixel fusion is adopted [21], [24]. At each pixel location (i, j) of the lth image channel with the qth IMF, that is, $\mathbf{I}_l^q$, the local variance of the pixels is computed within a sliding window of size $(2z+1)\times(2z+1)$ as

$$V_l^q(i,j) = \sum_{\alpha=-z}^{z}\sum_{\beta=-z}^{z}\left[I_l^q(i+\alpha, j+\beta) - \mu\right]^2, \quad (13)$$

where μ denotes the mean of all the pixel values inside the sliding window, given by

$$\mu = \frac{1}{(2z+1)^2}\sum_{\alpha=-z}^{z}\sum_{\beta=-z}^{z} I_l^q(i+\alpha, j+\beta). \quad (14)$$

Since the variance measure can be used to quantify the degree of local detail in input images that need to be transferred to the final fused image, all the pixels $\mathbf{I}_l^q(i,j)$ are assigned to local weighting factors, which are directly proportional to their local variance estimates. The idea behind this operation is that the IMF images which locally exhibit greater "activity of interest" are assigned larger weights than those exhibiting lower activity; this is in order to maximize their contribution to the fused image. The local weight factors, denoted by $W_l^q(i,j)$, are calculated through the following relations:

$$W_l^q(i,j) = \frac{V_l^q(i,j)}{\sum_{l=1}^{n} V_l^q(i,j)}, \quad (15)$$

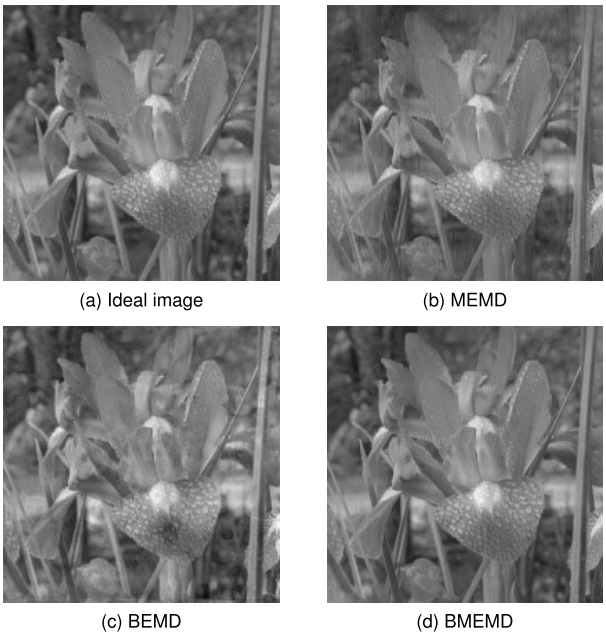


FIGURE 8. Ideal all-in-focus image and fused results obtained by MEMD, BEMD, and BMEMD.

and the qth IMF of the fused image, denoted by $\hat{\mathbf{I}}^q$, can be computed as

$$\hat{\mathbf{I}}^q = \sum_{l=1}^{n} \mathbf{W}_l^q \otimes \mathbf{I}_l^q, \tag{16}$$

where the symbol $\otimes$ denotes a point-wise multiplication operation between two matrices. This procedure is repeated for all Q IMFs to obtain a set of fused IMFs $\hat{\mathbf{I}}^q$, for $q = 1, 2, \ldots, Q$, which are added together to yield the fused image $\hat{\mathbf{I}}$, i.e., $\hat{\mathbf{I}} = \sum_{q=1}^{Q} \hat{\mathbf{I}}^q$.

Similar data processing procedures have been adopted by EMD based image fusion [21]–[27]. However, for those univariate ones, they need to be applied on all the input images, separately. In this sense, the problem of uniqueness, that is, different numbers of IMFs for different images, is likely to occur, due to the empirical nature of EMD. Although this problem can be partially addressed by forcing the iterative process of EMD to stop once a predefined number of IMFs are obtained, the residue may still maintain physically meaningful AM/FM components which would violate the fully data-driven property of EMD.

The proposed BMEMD method was applied to decompose the five multi-focus images shown in Fig. 1. The so-produced four 2D multivariate IMFs and a residue are shown in Fig. 7, displayed in the descending order of the spatial frequency. As compared with their counterparts obtained by the original MEMD in Fig. 2, they were much more meaningful for the image spatial analysis in the sense that the image edge information was well contained in the high frequency scales while the illumination information was better preserved in the low frequency components and without artifacts, owing to the 2D data processing nature of BMEMD. This is evidenced by the high-frequency IMF images in both the first and second rows, where the essential differences among different IMF channels, corresponding to different focus areas of the original images in Fig. 1, can be clearly observed. On the other hand, as expected, the low-frequency IMF images in the third and fourth rows and the residue image had very similar spatial structures in different image channels.

TABLE 1. Quantitative fusion results.

	Entropy	SF	RMSE	Corr.	SSIM
Ideal image	7.1203	0.0035	-	-	-
BEMD	6.9250	0.0032	0.0495	0.9433	0.8760
MEMD	6.8752	0.0024	0.0920	0.9454	0.9442
BMEMD	**7.1129**	**0.0036**	**0.0205**	**0.9932**	**0.9889**

We next performed fusion of the multi-focus images in Fig. 1 using BMEMD and compared the results with those obtained from BEMD and MEMD based fusion approaches. The ideal all-in-focus image and fused results are shown in Fig. 8. Observe that in Fig. 8(b), MEMD still produced several vertical artificial stripes, similar to its IMFs shown in Fig. 2, due to the loss of spatial information, although it was not that obvious after fusion. Moreover, the detail of leaves of *Iris ensata* and the image background in upper corners was unclear compared with the ideal image. In case of BEMD, there are several deficiencies occurring on the stem of *Iris ensata* on the right-hand side of Fig. 8(c). As desired, the fusion image of the proposed BMEMD in Fig. 8(d) retained almost all the detail of the ideal image and any spurious fusion artifacts were kept to a minimum. This is further supported by Table 1, where several quantitative metrics covering both the spatial and spectral quality of the fused image were considered, including entropy, spatial frequency (SF), root mean square error (RMSE), correlation (Corr.), and structural similarity (SSIM) [1], [38], [39]. The closest match obtained by BMEMD to the ideal image can be observed.

V. CONCLUSION

A bidimensional multivariate EMD (BMEMD) has been proposed for multi-scale image analysis. The critical step within BMEMD lies in n-dimensional surface projections, which directly convert a bidimensional n-variate signal into several bidimensional and univariate signals along multiple directions in the n-dimensional space, so that the spatial information within images is well-preserved. It has been shown that the proposed method has the ability to extract common spatio-temporal scales across multiple images. This mode alignment property together with the 2D processing nature suggests BMEMD as a better candidate for multi-scale image fusion than the conventional BEMD and MEMD. Simulations on synthetic texture sets and multi-focus images support the analysis.

REFERENCES

[1] T. Stathaki, *Image Fusion: Algorithms and Applications*. New York, NY, USA: Academic, 2008.

[2] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1391–1402, Jun. 2005.

[3] U. Patil and U. Mudengudi, "Image fusion using hierarchical PCA," in *Proc. Int. Conf. Image Inf. Process. (ICIIP)*, Nov. 2011, pp. 1–6.

[4] G. Pajares and J. M. de la Cruz, "A wavelet-based image fusion tutorial," *Pattern Recognit.*, vol. 37, no. 9, pp. 1855–1872, 2004.

[5] J. Tang, "A contrast based image fusion technique in the DCT domain," *Digit. Signal Process.*, vol. 14, no. 3, pp. 218–226, 2004.

[6] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.

[7] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[8] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 112–114, Feb. 2004.

[9] G. Rilling and P. Flandrin, "One or two frequencies? The empirical mode decomposition answers," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 85–95, Jan. 2008.

[10] J. C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, and P. Bunel, "Image analysis by bidimensional empirical mode decomposition," *Image Vis. Comput.*, vol. 21, no. 12, pp. 1019–1026, 2003.

[11] J. C. Nunes, S. Guyot, and E. Deléchelle, "Texture analysis based on local analysis of the bidimensional empirical mode decomposition," *Mach. Vis. Appl.*, vol. 16, no. 3, pp. 177–188, May 2005.

[12] M. U. B. Altaf, T. Gautama, T. Tanaka, and D. P. Mandic, "Rotation invariant complex empirical mode decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2007, pp. 1009–1012.

[13] T. Tanaka and D. P. Mandic, "Complex empirical mode decomposition," *IEEE Signal Process. Lett.*, vol. 14, no. 2, pp. 101–104, Feb. 2007.

[14] G. Rilling, P. Flandrin, P. Gonçalves, and J. M. Lilly, "Bivariate empirical mode decomposition," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 936–939, Dec. 2007.

[15] N. ur Rehman and D. P. Mandic, "Empirical mode decomposition for trivariate signals," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1059–1068, Mar. 2010.

[16] N. U. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 466, no. 2117, pp. 1291–1302, 2009.

[17] D. Looney, A. Hemakom, and D. P. Mandic, "Intrinsic multi-scale analysis: A multi-variate empirical mode decomposition framework," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 471, no. 2173, pp. 1–26, 2015.

[18] A. Hemakom, V. Goverdovsky, D. Looney, and D. P. Mandic, "Adaptive-projection intrinsically transformed multivariate empirical mode decomposition in cooperative brain–computer interface applications," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 374, no. 2065, pp. 1–15, 2016.

[19] X. Lang, Q. Zheng, Z. Zhang, S. Lu, L. Xie, A. Horch, and H. Su, "Fast multivariate empirical mode decomposition," *IEEE Access*, vol. 6, pp. 65521–65538, 2018.

[20] S. Tao, N. Li, S. Wei, and Y. Chai, "Weighted window sliding multivariate empirical mode decomposition for online multichannel filtering," *IEEE Access*, vol. 6, pp. 43170–43178, 2018.

[21] D. Looney and D. P. Mandic, "Multiscale image fusion using complex extensions of EMD," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1626–1630, Apr. 2009.

[22] N. Rehman, D. Looney, T. M. Rutkowski, and D. P. Mandic, "Bivariate EMD-based image fusion," in *Proc. IEEE Workshop Stat. Signal Process. (SSP)*, Aug./Sep. 2009, pp. 57–60.

[23] N. Rehman, M. M. Khan, M. I. Sohaib, M. Jehanzaib, S. Ehsan, and K. McDonald-Maier, "Image fusion using multivariate and multidimensional EMD," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2015, pp. 5112–5116.

[24] S. M. U. Abdullah, N. U. Rehman, M. M. Khan, and D. P. Mandic, "A multivariate empirical mode decompositionbased approach to pansharpening," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3974–3984, Jul. 2015.

[25] X. Xu, H. Li, and A. N. Wang, "The application of BEMD to multispectral image fusion," in *Proc. Int. Conf. Wavelet Anal. Pattern Recognit. (ICWAPR)*, Nov. 2007, pp. 448–452.

[26] X. Zhang, Y. Liu, and J. Chen, "Fusion of the infrared and color visible images using bidimensional EMD," in *Proc. Int. Conf. MultiMedia Inf. Technol. (MMIT)*, Dec. 2008, pp. 257–260.

[27] D. Zhang, J. Li, Z. Chen, and J. Zhang, "Fusion of polarization image using bidimensional empirical mode decomposition," in *Proc. Int. Conf. Comput. Sci., Mech. Automat. (CSMA)*, Oct. 2016, pp. 206–209.

[28] A. Linderhed, "2D empirical mode decompositions in the spirit of image compression," *Proc. SPIE*, vol. 4738, no. 7, pp. 1–8, 2002.

[29] J. Pan and Y. Tang, "A mean approximation based bidimensional empirical mode decomposition with application to image fusion," *Digit. Signal Process.*, vol. 50, pp. 61–71, Mar. 2016.

[30] H. Niederreiter, "Random number generation and quasi-Monte Carlo methods," *SIAM Rev.*, vol. 35, no. 4, pp. 680–681, 1993.

[31] J. Cui and W. Freeden, "Equidistribution on the sphere," *SIAM J. Sci. Comput.*, vol. 18, no. 2, pp. 595–609, 1997.

[32] J. Dick and F. Pillichshammer, *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[33] K. Suffern, *Ray Tracing from the Ground Up*. Natick, MA, USA: A K Peters, 2007.

[34] O. Kreylos and B. Hamann, "On simulated annealing and the construction of linear spline approximations for scattered data," *IEEE Trans. Vis. Comput. Graphics*, vol. 7, no. 1, pp. 17–31, Jan. 2001.

[35] L. P. Chew, "Constrained Delaunay triangulations," *Algorithmica*, vol. 4, no. 1, pp. 97–108, 1989.

[36] S. M. A. Bhuiyan, R. R. Adhami, and J. F. Khan, "Fast and adaptive bidimensional empirical mode decomposition using order-statistics filter based envelope estimation," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, Dec. 2008, Art. no. 728356.

[37] S. M. A. Bhuiyan, R. R. Adhami, and J. F. Khan, "A novel approach of fast and adaptive bidimensional empirical mode decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar./Apr. 2008, pp. 1313–1316.

[38] V. Tsagaris, "Objective evaluation of color image fusion methods," *Proc. SPIE*, vol. 48, no. 6, 2009, Art. no. 066201.

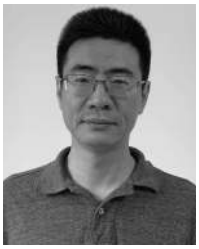
[39] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

YILI XIA (M'11) received the B.Eng. degree in information engineering from Southeast University, Nanjing, China, in 2006, the M.Sc. degree (Hons.) in communications and signal processing from the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K., in 2007, and the Ph.D. degree in adaptive signal processing from Imperial College London, in 2011.

Since 2013, he has been an Associate Professor in signal processing with the School of Information Science and Engineering, Southeast University, Nanjing, China, where he is currently a Professor and the Deputy Head of the Department of Information and Signal Processing Engineering. His research interests include complex and hypercomplex statistical analysis, detection and estimation, linear and nonlinear adaptive filters, and their applications on communications, power systems, and images. He was a recipient of the Education Innovation Award at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), in 2019. He is also an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

BIN ZHANG (S'18) received the B.S. degree in communication engineering from Hohai University, Changzhou, China, in 2014. He is currently pursuing the M.Sc. degree in electronics and communication engineering with Southeast University, Nanjing, China.

His research interests include multivariate signal processing and image processing.



WENJIANG PEI received the M.S. and Ph.D. degrees in instrumentation and measurement from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1995 and 1997, respectively.

He is currently a Professor in signal processing with the School of Information Science and Engineering, Southeast University, Nanjing. His research interests include statistical signal processing and hardware instrumentation.

DANILO P. MANDIC (M'99–SM'03–F'12) received the Ph.D. degree in nonlinear adaptive signal processing from Imperial College London, London, U.K., in 1999, where he is currently a Professor in signal processing. He has been working in the area of nonlinear adaptive signal processing, multivariate data analysis, and nonlinear dynamics. He has been a Guest Professor with Katholieke Universiteit Leuven, Leuven, Belgium, Tokyo University of Agriculture and Technology, Tokyo, Japan, and a Frontier Researcher with RIKEN, Japan. His publication record includes two research monographs entitled *Recurrent Neural Networks for Prediction*: *Learning Algorithms*, *Architectures and Stability* (1st ed., August 2001), and *Complex Valued Nonlinear Adaptive Filters*: *Noncircularity*, *Widely Linear and Neural Models* (1st ed., Wiley, April 2009), an edited book titled *Signal Processing Techniques for Knowledge Extraction and Information Fusion* (Springer, 2008), and a two-volume research monograph *Tensor Networks for Dimensionality Reduction and Large Scale Optimization* (Now Publishers, 2016 and 2017). Prof. Mandic has received the Denis Gabor Award for Outstanding Achievements in Neural Engineering, given by the International Neural Networks Society, and is a winner of the 2018 Best Paper Award in the *IEEE Signal Processing Magazine*, for his article on tensor decompositions. He has received the President Award for Excellence in Postgraduate Supervision at Imperial College London. He has been an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the *IEEE Signal Processing Magazine*, and the IEEE TRANSACTIONS ON INFORMATION AND SIGNAL PROCESSING OVER NETWORKS.