BIVARIATE EMD-BASED IMAGE FUSION

N. Rehman[1] and D. Looney[1] and T. M. Rutkowski[2] and D. P. Mandic[1]

[1]Imperial College London, UK
[2]RIKEN Brain Science Institute, Saitama, Japan
E-mail: {naveed.rehman07,david.looney06,d.mandic}@imperial.ac.uk, {tomek@brain.riken.jp}

ABSTRACT

The Empirical mode decomposition (EMD) algorithm is a fully data-driven method which is used to perform an adaptive decomposition of nonlinear and nonstationary signals. It has been recently illustrated that its complex extensions can be used to carry out fusion of multiple images. This is possible because the complex EMD allows comparison between common frequency scales, by aligning them within a single complex IMF. In this paper, complex extensions of EMD are proposed for the fusion of two images; the fusion methodologies are presented for both gray-level and RGB based color images. The potential of the proposed scheme is highlighted by showing its superiority to wavelet based fusion schemes, through simulations on real world multi-exposure images.

Index Terms— Empirical Mode Decomposition, complex signals, complex/bivariate EMD, image fusion, multi-exposure images.

1. INTRODUCTION

The fusion of multiple images is aimed at producing a single output image which carries salient features of all fused images [1]. Fusion techniques are particularly relevant in cases where it is difficult to obtain an image in which all relevant objects are 'in-focus', and multiple 'out-of-focus' images must be combined to to yield a single 'in-focus' image. Another important application is the fusion of multiple multi-exposure images from a static scene, which is required when a wide range of luminosity/irradiance inhibits the camera to capture all features of a scene in a single shot. In such cases, there will always be regions of an image which are either over-exposed or under-exposed. These over- and under-exposed regions usually carry less information as compared to regions which are properly exposed to light, and therefore need to be fused. For fusion of multiple images, one class of techniques perform 'local' fusion by obtaining local details from images by dividing them into several non-overlapping blocks. Subsequently, image block carrying greater information is selected in the fused image. Using this methodology, Goshtasby proposed entropy as an information measure for the fusion of multiple exposure images [2]. This technique, however, does not perform local fusion on decomposed images corresponding to multiple natural frequency scales of the input signal (multi-scale fusion), and is therefore sub-optimal.

Other established techniques for the fusion of multiple images are mostly based on combining multi-scale decompositions of input images using some information measure. The multi-resolution wavelet transform is commonly used to provide the required decomposition, after which the coefficients corresponding to decomposed signals are conveniently combined to yield coefficients of the fused image [3]. The inverse wavelet transform of the fused coefficients yields the well-exposed fused image. Fusion has also been achieved recently with the Discrete cosine transform (DCT) domain [4], and in the Fourier domain. However, these techniques fail to extract the details of an image at the 'local' level, and hence are not able to perform fusion 'locally', due to their non-adaptive nature. Furthermore, the Fourier analysis and DCT project the input signal on orthogonal sinusoids, regardless of the nature of input data; whereas wavelets are also based on projecting the input signal on some fixed a priori basis function which makes them unsuitable for processing nonstationary signals. As real world images are typically nonlinear and nonstationary, a fully adaptive approach is required for their fusion.

The empirical mode decomposition (EMD) algorithm, recently proposed for the processing of nonlinear and nonstationary signals, creates an adaptive decomposition of the signal in hand [5]. The fully adaptive and data driven nature of EMD algorithm makes it superior to other established signal processing tools (Fourier, wavelet analysis), which employ projections on fixed basis functions to decompose an input signal. In [6], real-valued EMD was used for the fusion visual and thermal images, and was shown to outperform wavelets and PCA based approaches. However, due to the empirical nature of EMD, it is not guaranteed to obtain the same number of IMFs from two sources, which makes 'scale by scale' comparison difficult. Also, even if the same number of IMFs are generated, it is not possible to align the similar frequency modes from different sources using the real-valued EMD. This problem of uniqueness was solved by Looney and Mandic using bivariate extensions of EMD to process two images simultaneously in the context of fusing multifocus images [7]. This way, not only the same number of IMFs are generated by the two sources, but the frequency scales from both images are aligned in single complex IMFs, which facilitates a better comparison between the relevant scale-images from input data.

In this work, we propose to use complex extensions of empirical mode decomposition algorithm to simultaneously achieve both 'multi-scale' and 'local' fusion of multiple images. For that purpose, we present separate fusion methodologies for both gray-level and RGB based colored images. The results of the proposed scheme are compared against the standard wavelet based fusion approach via simulations on real world multi-exposure images.

2. EMPIRICAL MODE DECOMPOSITION

Empirical mode decomposition is a fully adaptive method for decomposing nonlinear and nonstationary signals [5]. It uses the 'sifting process' to obtain a finite set of oscillatory components, also known as intrinsic mode functions (IMFs), from the input signal. These intrinsic mode functions are designed to ensure that each IMF carries a local, natural oscillatory mode embedded in the original signal, and that it exhibits well-behaved time-frequency spectra [5].

For an input real-valued signal $x(k)$, application of EMD algo-

rithm results in the decomposition of the signal into M oscillatory modes, IMFs, and a residue signal $r(k)$, given by

$$x(k) = \sum_{j=1}^{M} d_j(k) + r(k) \qquad (1)$$

The residual $r(k)$ is a monotonic function, and represents an overall trend within the signal. To extract the individual IMFs from the input signal, an iterative procedure known as sifting algorithm is used. The details of sifting algorithm for obtaining the first IMF from a signal $x'(k)$ are as follows:

1. Find the locations of all the extrema of $x'(k)$;

2. Interpolate (using spline interpolation) between all the minima (resp. maxima) to obtain the signal envelope passing through the minima, $e_{min}(k)$ (resp. $e_{max}(k)$);

3. Compute the local mean $m(k) = (e_{min}(k) + e_{max}(k))/2$;

4. Subtract the mean from the signal to obtain the "oscillating" signal $s(k) = x'(k) - m(k)$;

5. If the resulting signal $s(k)$ obeys the stopping criteria, $d(k) = s(k)$ becomes an IMF, otherwise set $x'(k) = s(k)$ and repeat the process from Step 1.

The stoppage criterion used in the final step can be taken as the normalized squared difference between two successive sifting iterates $s_n(k)$ and $s_{n-1}(k)$ [5].

3. COMPLEX EXTENSIONS OF EMPIRICAL MODE DECOMPOSITION

Recently, several extensions of EMD have been proposed for complex/bivariate signals. An extension of EMD which operates fully in the complex domain was first proposed in [8], and termed as rotation-invariant EMD (RI-EMD). The extrema of a complex/bivariate signal are chosen to be points where the angle of the derivative of the complex signal becomes zero, which is based on the change in the phase of the signal. Mathematically, this definition is equivalent to the extrema of the imaginary parts of the complex signal, as given for a complex signal $z(t)$ as:

$$\angle \dot{z}(t) = 0 \Rightarrow \angle\{\dot{x}(t) + \jmath\dot{y}(t)\} = 0$$
$$\Rightarrow \tan^{-1}\frac{\dot{y}(t)}{\dot{x}(t)} = 0 \Rightarrow \dot{y}(t) = 0. \qquad (2)$$

The envelopes are calculated using component-wise spline interpolation, corresponding to the maxima and minima points which are then averaged to obtain the local mean of the bivariate signal. The rest of the process is same as that of the real-valued EMD. The complex IMFs obtained from the sifting process are physically meaningful as the process operates directly in the complex domain.

The RI-EMD algorithm uses only the extrema of the imaginary part of the complex signal to calculate the local mean, which results in envelopes corresponding to only two directions. To estimate an accurate value of the local mean of the complex signal, the envelopes corresponding to the multiple directions must be considered in complex domain. Bivariate EMD [9] achieves this by projecting the complex signal in multiple directions in 2D spaces and calculating the envelope corresponding to each direction. The envelope
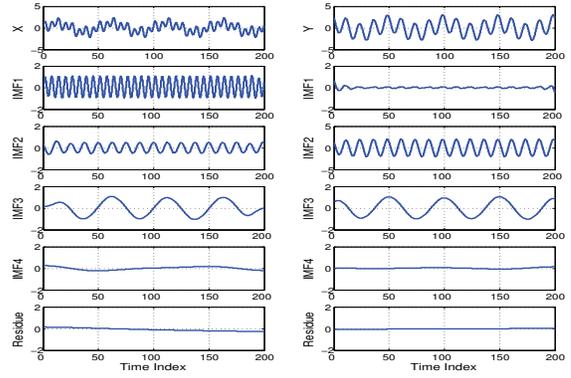
Fig. 1. Decomposition of a synthetic complex signal, with multiple frequency modes, via the bivariate EMD algorithm.

curves are formed by interpolating between the extrema locations of the projected signal component-wise. The resulting multiple envelopes are averaged to obtain the local mean signal. It can be noticed that RI-EMD can be treated as a special case of bivariate EMD, with two projected directions.

As bivariate EMD uses multiple directions for taking projections, it estimates multiple envelopes for calculating the local mean. The resulting estimate of local mean is therefore more accurate than that obtained by RI-EMD. For this paper, we performed all simulations using bivariate EMD with eight directions for taking projections.

4. BIVARIATE EXTENSIONS FOR FUSION OF MULTI-EXPOSURE IMGAES

We propose to use complex extensions of EMD for the fusion of two images, performing both 'local' and multi-scale fusion simultaneously. In this section, we present two separate methodologies to achieve fusion of both gray scale and RGB based color images using a simple local variance-based fusion algorithm, which operates on each scale.

4.1. Alignment of Common Frequency Modes

Firstly, it is necessary to illustrate how complex-valued IMFs align 'common scales' present within the input data [7], as this is vital for multi-scale fusion of multiple images. For this purpose, a complex signal was constructed from a set of three sinusoidal signals. The real and imaginary components of a complex signal are shown in the top row of Figure 1 (denoted by X and Y). Common to both real and imaginary parts are two sinusoidal frequency components, while a high frequency sinusoid is added to the real part only. The bivariate EMD algorithm was applied to the resulting complex signal yielding multiple complex-valued IMFs, the components of which are shown in Figure 1. Observe that the sinusoids common to both real and imaginary parts are aligned in the second and third IMFs, while the remaining high frequency mode is present in the real part of the first IMF only. Common frequency modes are therefore aligned in corresponding IMFs, which allows a meaningful comparison between real and imaginary components, thus facilitating fusion at local level.

2009 IEEE/SP 15th Workshop on Statistical Signal Processing
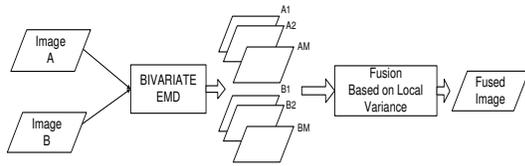
Fig. 2. Multiple gray-scale image fusion methodology using the bivariate EMD.

4.2. Complex EMD based Local Image Fusion

In this section, we present an approach for the fusion of two images to a single improved image using complex extensions of EMD. The methodology employed is similar to the one used in [7] for the fusion of two out-of-focus images. We propose a new fusion rule for multiscale image fusion, which is more suited for the fusion of exposure images. Also, a method to fuse two color images using multiple applications of the bivariate EMD is presented.
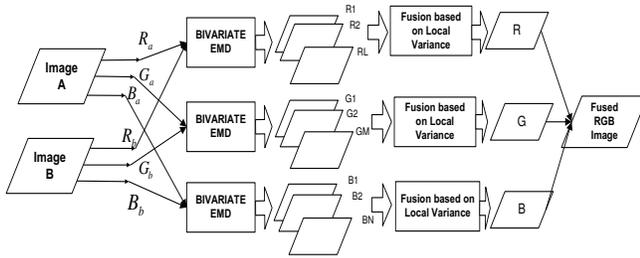
Fig. 3. RGB color image fusion methodology using the bivariate EMD.

4.2.1. Gray scale image fusion scheme

Two input gray scale images, each considered as a 2D array and denoted by A and B, are first converted into row vectors by concatenating their rows. Two such row vectors are combined to form the real and imaginary part of a complex signal. Bivariate EMD is applied to the complex signal resulting in M complex IMFs. Real and imaginary parts of each complex IMF are then separated, and reconverted back to 2D images, resulting in M scale images for each input image, denoted by A_i and B_i for $i = 1, ..., M$, as shown in Figure 2. Scale images are then combined locally, based on values of coefficients calculated from local variance estimates at each spatial point, to give a fused image F given by

$$F(i,j) = \sum_{n=1}^{M} \alpha_n(i,j)A_i(i,j) + \beta_n(i,j)B_i(i,j) \quad (3)$$

where $\alpha_i(i,j)$ and $\beta_i(i,j)$ are coefficients determined at each location (i,j), based on the the relative values of the local variance for each scale, and are given by

$$\alpha_n(i,j) = \frac{\text{var}[A_n(i,j)]}{\text{var}[A_n(i,j)] + \text{var}[B_n(i,j)]} \quad (4)$$

$$\beta_n(i,j) = \frac{\text{var}[B_n(i,j)]}{\text{var}[A_n(i,j)] + \text{var}[B_n(i,j)]} \quad (5)$$
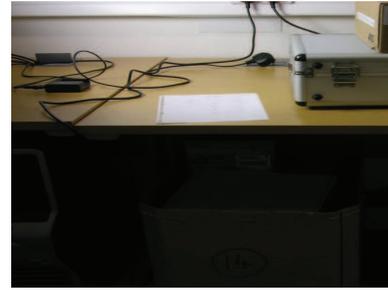
Fig. 4. Images of a work table at different exposures. In the image shown at the top, the lower region of the image is severely underexposed. whereas in the second image, details on the table are hidden due to overexposure to light.


**Fig. 5**. A fused image obtained using the discrete wavelet transform.

where $\text{var}[A_n(i,j)]$ represents the local variance of A_n calculated across a small block around (i,j).

4.2.2. RGB based color image fusion scheme

In order to extend the fusion methodology to RGB based colored images, we propose to apply the methodology presented for gray scale fusion to three channels (red, green, and blue) of a color image separately. That is, the red, green, and blue channels from two input images are therefore processed by three separate applications of bivariate EMD algorithm, as shown in Figure 3. As a result, three sets of complex valued IMFs are obtained which correspond to the red, green, and blue channels of the two input images. These are denoted by R_l, G_m, and B_n in Figure 3. Note that although the three instances of bivariate EMD generally yield different numbers of complex IMFs, each set is processed separately by the fusion algorithm to yield the fused red, green and blue channel. The fused channels are finally combined to form a fused RGB based color image.
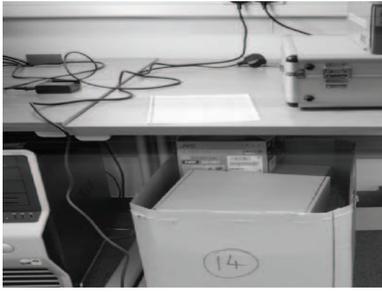
Fig. 6. A fused image obtained using complex extensions of EMD.
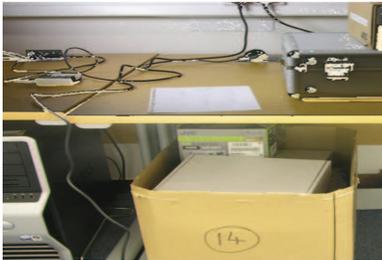
(a)

(b)

Fig. 7. (a) The fused color image obtained from 'local' fusion on input images. (b) The fused color image obtained from complex extensions of EMD.

5. RESULTS

The performance of our proposed method is evaluated on real-world multi-exposure images shown in Figure 4; these are images of a work table obtained at multiple exposures. Observe that the in the first input image, the lower region is severely underexposed, whereas the upper part of the image, showing details on the table itself, is well-exposed. In the second input image, the details on the table are hidden due to over-exposure, but the area below the table is well-exposed.

For comparison, a multi-resolution wavelet fusion scheme is also used to for multi-exposure image fusion. For wavelet fusion, the discrete wavelet transform (DWT) is first applied to both input images to obtain their multi-scale decompositions. The wavelet coefficients corresponding to the same decomposition level of two images, are combined to obtain fused multi-scale coefficients. Once the fused multi-scale coefficients are obtained, they are converted back to the fused image using the inverse wavelet transform. In our simulations, we choose the largest coefficient which correspond to sharper brightness changes in the image, such as edges, lines etc. The selection of the largest wavelet coefficient is also consistent with the local vari-

ance based fusion algorithm chosen in our EMD-based approach, since both criteria select variations in intensity as salient features in an image.

Figure 5 shows the results of the fusion of images in Figure 4 using the discrete wavelet transform (DWT). In the fused image, distortions are clearly evident, especially in the upper part of the image where details on the table are shown; observe the artifacts around the cable and the paper. It can also be observed from the wavelet fused image that the container on the right side of the table is still overexposed. The fused image from the EMD-based fusion scheme is shown in Figure 6; it is clear that EMD-based fusion outperforms wavelet fusion, as less distortions are present.

The distortions around the wire and paper in wavelet-based fusion arose due to the fact that the wavelet fusion is not performed at the 'local' level in the spatial domain; the coefficient merging occurs in the transform domain. As a result, the wavelet-based fusion scheme cannot accurately align, compare and process high frequency scales. On the other hand, the adaptive and data driven nature of EMD facilitates fusion at the local level, yielding better results even at high frequency scales.

To highlight the importance of 'multi-scale' processing in obtaining robust fusion results, we use our proposed scheme for RGB based color image fusion (Figure 3). The results are compared with those obtained from a fusion scheme which is 'local' but not multi-scale, that is the fused image (Figure 7(a)) retains salient features based on the local variance of the input images. Figure 7(b) shows the results obtained from applying our proposed scheme; it can be noticed that it clearly gives superior results to local fusion scheme, as it performs both 'local' and 'multi-scale' fusion simultaneously.

6. CONCLUSIONS

We have used complex extensions of Empirical Mode Decomposition for the fusion of multiple images. Furthermore, novel methodologies for the fusion of both gray scale images and color images, using the complex extensions of EMD, are presented. The proposed method outperforms multi-resolution wavelet based fusion algorithm and the local variance-based algorithm in simulations on real world multi-exposure images. This is achieved because the proposed methodology facilitates multi-scale and local fusion simultaneously, which is not possible using other existing fusion algorithms.

7. REFERENCES

[1] T. Stathaki, *Image Fusion: Algorithms and Applications*, Academic Press, 2008.

[2] A. A. Goshtasby, "Fusion of multi-exposure images," *Image and Vision Computing*, vol. 23, pp. 611–618, 2005.

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

[4] I. Zafar, E.A. Edirisinghe, and H.E. Bez, "Multi-exposure and multi-focus image fusion in transform domain," in *in Proceedings of IET International Conference on Visual Information Engineering*, 2006.

[5] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, and H. Liu, "The empirical mode decomposition and Hilbert spectrum for non-linear and non-stationary time series analysis," *Proceedings of the Royal Society A*, vol. 454, pp. 903–995, 1998.

[6] H. Hariharan, A. Gribok, M. A. Abidi, and A. Koschan, "Image fusion and enhancement via empirical mode decomposition," *Journal of Pattern Recognition Research*, vol. 1, pp. 16–32, 2006.

[7] D. Looney and D. Mandic, "Multi-scale image fusion using complex extensions of EMD," *IEEE Transactions in Signal Processing*, vol. 57, no. 4, pp. 1626–1630, 2009

[8] M. U. Altaf, T. Gautama, T. Tanaka, and D.P. Mandic, "Rotation in-variant complex empirical mode decomposition," in *Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Processing*, 2007, vol. 3.

[9] G. Rilling, P. Flandrin, P. Goncalves, and J. M. Lilly, "Bivariate empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 14, pp. 936–939, 2007.