IMAGE FUSION USING MULTIVARIATE AND MULTIDIMENSIONAL EMD

Naveed ur Rehman[1] *M. Murtaza Khan*[2] *Ishaq Sohaib*[1] *M. Jehanzaib*[1] *Shoaib Ehsan*[3] *Klaus McDonald-Maier*[3]

[1] Department of Electrical Engineering, COMSATS Institute of Information Technology, Park Road, Islamabad, Pakistan
[2] School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad, Pakistan
[3] School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK

ABSTRACT

We present a novel methodology for the fusion of multiple (two or more) images using the multivariate extension of empirical mode decomposition (MEMD). Empirical mode decomposition (EMD) is a data-driven method which decomposes input data into its intrinsic oscillatory modes, known as intrinsic mode functions (IMFs), without making a priori assumptions regarding the data. We show that the multivariate and multidimensional extensions of EMD are suitable for image fusion purposes. We further demonstrate that while multidimensional extensions, by design, may seem more appropriate for tasks related to image processing, the proposed multivariate extension outperforms these in image fusion applications owing to its mode-alignment property for IMFs. Case studies involving multi-focus image fusion and pan-sharpening of multi-spectral images are presented to demonstrate the effectiveness of the proposed method.

Index Terms— Empirical mode decomposition (EMD), Multivariate EMD (MEMD), Bidimensional EMD, Multi-focus image fusion, Pan-sharpening.

1. INTRODUCTION

Image fusion is the process of combining multiple images to produce a single image which carries more information than any of the images used for blending [1]. Fusion techniques are useful for cases where the limitations of optical sensors and imaging conditions make it difficult to view multiple objects clearly in a single image. In such cases, multiple images are obtained with each containing partial information about a scene. Multi-focus and Multi-exposure images are two example classes in which the relevant objects may be obscured as a result of either being out of focus or not being properly exposed to the light source. These sets of images can be merged to present the complete information in a single image via multi-focus or multi-exposure image fusion.

Similarly, in remote sensing applications, we often require both high spatial and spectral information in a single image which is not physically possible to obtain via available sensors. In such cases, fusion of high spatial resolution panchromatic image (PAN) and low spatial resolution (but high spectral resolution) multispectral (MS) images is performed to obtain the desired high spatial and spectral resolution MS image. This process is also called Pan-sharpening [2].

Image fusion methods may be characterized as pixel-level fusion, multi-scale fusion and hybrid fusion techniques. The key steps of pixel level fusion include: i) generation of a quantitative map of information content for each image; ii) comparison of information content at pixel level; iii) assigning weights to individual pixels (or a set of pixels) based on information content; iv) and weighted recombination to obtain fused image. The advantages of such class of methods include their low computational cost and simplicity, while the main disadvantage is their susceptibility to noise. Multi-scale techniques, on the other hand, operate by first decomposing input images in terms of their frequency components which are then combined to obtain a single fused image. Here, the main steps include: i) converting input images into transform domain coefficients; ii) assigning weights to the coefficients based on information content; iii) selecting the relevant coefficients; and iv) taking the inverse transform. Typical examples are the methods based on Gaussian pyramids, Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT) [3], and Discrete Wavelet Transform (DWT) [4].

We propose a hybrid (multi-scale and pixel-level) and data-driven scheme for image fusion based on multivariate extensions of empirical mode decomposition (MEMD) algorithm [6]. We also compare our results with the standard bi-dimensional EMD (BDEMD) [7] based fusion approach. The EMD based fusion methods are employed since they are fully data adaptive, enable fusion of intrinsic scales at *local* level, and allow fusion of matched spatial frequency content between input images. Standard multiscale methods (based on Fourier and wavelet transform) employ static filter banks and predefined basis functions which hinder the fusion of matched spatial frequency content between input images. We demonstrate the potential of the proposed scheme in two application scenarios: a) multi-focus image fusion; and b) pan-sharpening of MS images. In both cases, the fusion results obtained from the proposed scheme outperforms the results obtained by BDEMD both qualitatively and quantitatively.

2. EMD AND ITS MULTIVARIATE AND MULTIDIMENSIONAL EXTENSIONS

2.1. Standard EMD

Empirical mode decomposition (EMD) [5] is a data-driven method which decomposes an arbitrary signal $x(k)$ into a set of multiple oscillatory components called the intrinsic mode functions (IMFs) via an iterative process known as sifting algorithm [6]. The IMFs represent the intrinsic temporal modes (scales) that are present in the input data which when added together reproduce the input $x(k)$, as shown in eq. (1) below:

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

The residual $r(k)$ does not contain any oscillations and represents a trend within the signal.

The recursive sifting algorithm operates by defining the upper and lower envelopes of an input signal by interpolating its extrema. The local mean $m(k)$ is then estimated by averaging these envelopes, which is subsequently subtracted from the input signal $x(k)$ to obtain the fast oscillating signal $d(k) = x(k) - m(k)$. Next, $d(k)$ is checked for an IMF condition; if it is not satisfied, the process is repeated until the condition for IMF is satisfied and we obtain an IMF. The sifting process stops when $d(k)$ has inadequate extrema.

2.2. Bi-dimensional EMD (BDEMD)

Bi-dimensional EMD (BDEMD) [7] is a generic extension of EMD for images. Various algorithms for computing BDEMD decomposition exist which mainly differ in the way the extrema are interpolated to obtain upper and lower envelopes. Radial basis functions (tensor product) or B-splines are commonly used methods for interpolation [7], whereas the method by Linderhed [8] uses thin-plate splines for the interpolation of the extrema.

2.3. Multivariate EMD (MEMD)

Multivariate EMD (MEMD) algorithm extends the functionality of EMD to signals containing multiple channels [6]. The rationale behind the MEMD is to separate inherent rotations (rather than oscillations) within a signal. This is achieved by estimating the local mean of a multivariate signal in multidimensional spaces where the signal resides. For multivariate signals, however, the concept of extrema cannot be defined in clear terms and therefore envelopes cannot be obtained as a trivial extension of univariate case. To address this issue, MEMD operates by projecting an input multivariate signal in V uniformly spaced directions on a unit p-sphere; the extrema of the so projected signals are then interpolated to obtain multiple envelopes which are subsequently averaged to obtain the local mean.

3. MEMD VS BDEMD: MODE ALIGNMENT

Fig. 1 shows correlations of normalized IMFs for two multi-focus images (Fig. 3(a) and Fig. 3(c)), obtained using the
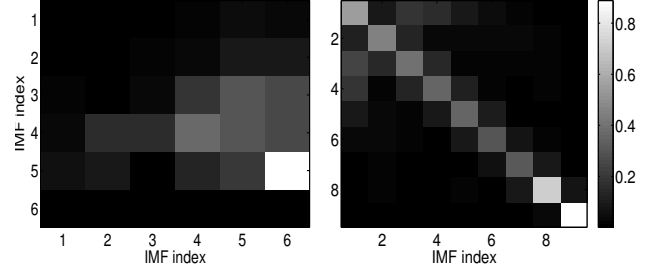
**Fig. 1**: Cross-correlation of normalised IMFs for multi-focus images (left) BDEMD (right) MEMD.
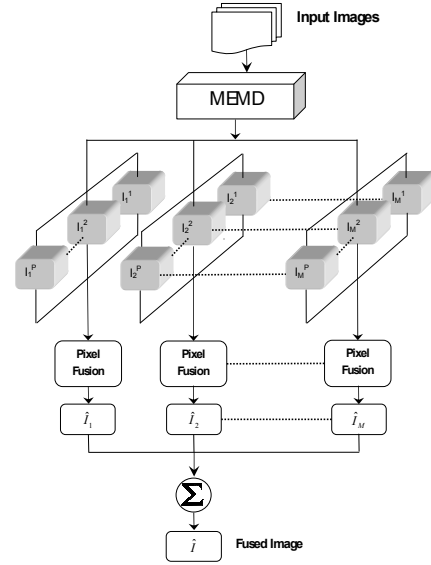
Fig. 2: The proposed scheme illustrating the local fusion of P arbitrary images to yield a single fused image $\hat{I}$ using the MEMD algorithm.

BDEMD (left) and MEMD (right). Note that the MEMD produced diagonally dominant correlograms of IMFs as compared to the BDEMD, proving that the same-indexed IMFs generated from MEMD are highly correlated, a major requirement in most fusion applications. This mode alignment property of MEMD is a result of direct processing of input images within MEMD, whereas the lack of it in BDEMD is due to the fact that it processes multiple input images separately.

4. MEMD- AND BDEMD-BASED IMAGE FUSION

The proposed algorithm based on MEMD operates by first converting the P input images into a vector form by concatenating their rows/columns. The resulting vectors are then put together to form a multivariate signal containing P number of data channels. MEMD is next applied to the resulting signal yielding M number of IMFs for each channel; let us denote the $m-$th IMF of the $p-$th channel (input image) by $I_m^p(a, b)$, where $m = 1 \ldots M$; $n = 1 \ldots P$; and a and b represent the spatial coordinates. To perform fusion at the pixel level,

(a) Image 1   (b) Image 3   (c) Image 7   (d) BDEMD   (e) MEMD
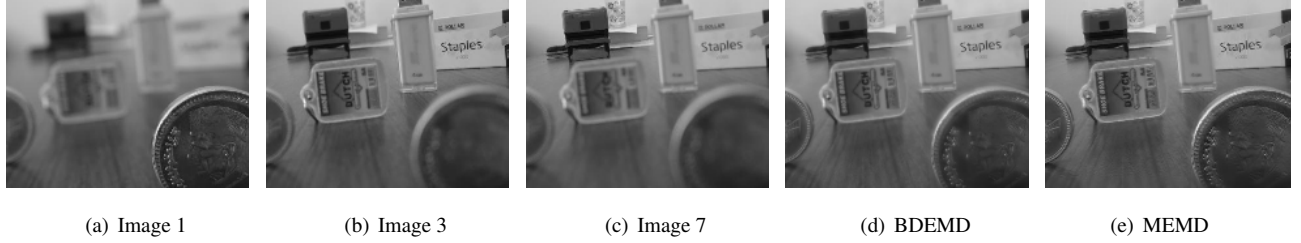
Fig. 3: Multi-focus image fusion results for data set 1.

sub-images $I_m^p(a,b)$ are divided into small windows of size $N \times N$ and their variances $\xi_m^p(a,b)$ are computed. Next, the sub-images $I_m^p(a,b)$ are assigned local weights, $WF_m^p(a,b)$, based on $\xi_m^p(a,b)$ by using the following relation:

$$WF_m^p(a,b) = \frac{\xi_m^p(a,b)}{\sum_{p=1}^{P} \xi_m^p(a,b)} \qquad (2)$$

This means that the IMFs exhibiting greater variance are assigned higher weights $WF_m^p(a,b)$ than those exhibiting lower variances, thereby maximising their contribution to the fused image. To obtain the $m-$th IMF of the fused image, the IMFs of all P input images $\{I_m^p(a,b)\}_{p=1}^{P}$ are multiplied by their respective weight factors $\{WF_m^p(a,b)\}_{p=1}^{P}$ and added together to obtain:

$$\hat{I}_m(a,b) = \sum_{p=1}^{P} WF_m^p(a,b) \times I_m^p(a,b) \qquad (3)$$

This procedure is repeated for all M IMFs to obtain a set of fused IMFs $\{\hat{I}_m(a,b)\}_{m=1}^{M}$ which are added together to yield the fused image $\hat{I}$.

$$\hat{I}(a,b) = \sum_{m=1}^{M} \hat{I}_m(a,b) \qquad (4)$$

The block diagram of the proposed multivariate EMD based fusion algorithm is shown in Fig. 2.

The BDEMD based fusion algorithm operates similarly to the MEMD algorithm illustrated above. The only difference is that instead of a single operation of MEMD on all input images, BDEMD is applied separately on P input images to obtain M_p IMFs, where $p = 1 \ldots M$. Note that owing to the empirical nature of the EMD algorithm, typically different number of IMFs are obtained for multiple input images resulting in mismatched IMFs, thus hindering the fusion process.

5. CASE STUDY 1: MULTI-FOCUS IMAGE FUSION

As a case study, we first performed fusion of multi-focus images using MEMD and compared the results with those obtained from BDEMD based fusion approach. For this purpose, multiple images of 30 different scenes were used; seven images were taken of each scene with different parts of the scene out-of-focus in each image.

For quantitative evaluation of the fusion results, we have employed Entropy (E) [11], objective image fusion ($QABF$) [10] and the spatial frequency (SF) [9] performance measures. Collectively, these measures served to quantify the spatial distortion as well as the information present in the fused image. The parameters used in MEMD algorithm were: number of direction vectors $V = 8$ and the square window length of $N = 9$. The standard BDEMD algorithm given in [7] was implemented for fusion purposes. Fig. 3 shows a subset of input multi-focus images and the
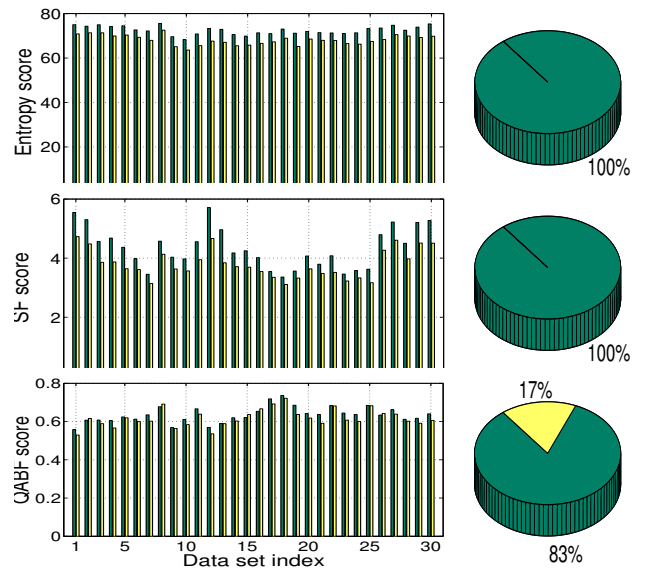
Fig. 4: Quantitative comparison of the proposed fusion schemes on 30 input multi-focus images. (Left column) Bar graphs of the values of quantitative measures, including %Entropy, %SF, and $QABF$ shown respectively from top to bottom, obtained for the MEMD- (dark green), and BDEMD-(yellow) based fusion methods. (Right column) Pie charts of the quantitative measures highlighting the relative performance of the MEMD (dark green), and BDEMD (yellow).

fused images obtained from the two methods; only three out of seven input images are shown due to the space restrictions. Note from Figs. 3(a-c) that each input image has some specific objects within focus: Image 1, for instance, focuses on the nearest objects such as the coin, whereas the Image 2 and Image 3 focus on the middle and the farthest objects respec-

| (a) MS | (b) PAN | (c) AWT | (d) BDEMD | (e) MEMD |

Fig. 5: Fusion results for Pleiades Tolouse Image.

tively. The proposed MEMD-based fusion algorithm resulted in a single output image, shown in Fig. 3(d), which has all objects within focus. Similarly, the BDEMD-based fused image also yielded an improved image, though not as sharp as the one obtained via MEMD: Please observe the improved sharpness of the word 'DUTCH' written on the key chain in the MEMD-fused image. Similar trend was observed in most of the 30 input multi-focus data sets used in our experiments, with MEMD outperforming BDEMD-based fusion. This is evident from Fig. 4 (left) which shows the bar graphs of the values of the quantitative performance measures (E, SF, and $QABF$) for all data sets. To complement that, Fig. 4 (right) shows pie-charts highlighting the number of cases in terms of percentage where each method performed best. It can be observed that for E and SF measures, the MEMD yielded superior results for approximately all input data sets, whereas for $QABF$, the MEMD produced better results for 83% of the input data sets.

6. CASE STUDY 2: PAN-SHARPENING

We next performed experiments for Pan-sharpening of multispectral (MS) images using the existing BDEMD-based fusion and the proposed MEMD-based fusion algorithms. Their performance was compared against a multiscale Pan-sharpening algorithm called àtrous wavelet transform (AWT) method [12]. The simulated Pleiades data set consisting of 1024×1024 pixels of i) Strasbourg and ii) Tolouse was used in our experiments. All the input MS images contained four bands i.e. blue (B), green (G), red (R) and near-infrared (NIR). The ground truth for both data sets was also available (not shown in Fig. 5 due to space restrictions) and was used for the quantitative analysis of the fusion results.

Pan-sharpening of MS images via MEMD and BDEMD was performed as follows: the intensity plane I of the input MS image was first obtained by averaging the bands of the MS image. The so obtained intensity image I was then fused with the high resolution panchromatic image to obtain the 'detailed' intensity image $\hat{I}$. The details added by the fusion process were extracted by subtracting the original intensity plane I from $\hat{I}$, which were then separately added to the B, G, R, and NIR components of the MS image to obtain the Pan-sharpened MS image.

Figs. 5(a-b) show the source MS and PAN images of Strasbourg city taken from the Pleiades sensor. The fused images obtained using AWT, BDEMD and MEMD algorithms are shown in Figs. 5(c-e) respectively. It can be noticed that the spectral performance of the images obtained from the proposed MEMD-based method, shown in Fig. 5(c), matches the result of the state-of-the-art AWT technique. The proposed method, however, showed much improved spatial performance as compared to both AWT and BDEMD based fusion methods.

Table 1: Quantitative Results of Pan-sharpening

	Pleiades Strasbourgh			Pleiades Tolouse		
	AWT	BD..	MEMD	AWT	BD..	MEMD
S	2.91	4.96	**2.90**	**4.70**	5.23	5.01
ER	4.247	6.52	**2.981**	5.74	5.72	**3.52**
$Q4$	89.31	54.36	**93.89**	94.2	59.22	**94.98**

The improved performance of MEMD can be further validated by the quantitative results of pan-sharpening on both data sets. We employed the following set of performance metrics for this purpose: i) Relative dimensionless global error in synthesis (ER, ideally 0), ii) Spectral Angle Mapper (S, ideally 0), and iii) Quaternion Index ($Q4$, ideally 100%) [2]. The results of the quantitative analysis are presented in Table I with the best value for each quality measure are labeled in bold. Please observe that in both data sets the proposed scheme performed better than AWT and BDEMD fusion methods for all performance metrics with an exception to S value for the Tolouse image where AWT performed better. Superiority of MEMD over BDEMD can be attributed to data adaptive and local nature of its decomposition which manifested in improved spatial performance in both case studies.

7. CONCLUSIONS

We have presented a method for the fusion of multiple images using multivariate empirical mode decomposition (MEMD) algorithm. The superiority of the method has been demonstrated on a large data set for two applications: i) multi-focus fusion, and ii) pan-sharpening of multi-spectral images. In addition to the qualitative analysis, we have also employed a wide range of quantitative performance measures to compare the fusion results obtained from the two approaches.

8. REFERENCES

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

[2] 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 Transactions on Geoscience and Remote Sensing,* vol. 45, no. 10, pp. 3012–3021, 2007.

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

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

[5] N. E. Huang, Z. Shen, S. R. Long, M. L. Wu, H. H. Shih, Z. Quanan, 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", *Proceedings of the Royal Society A*, vol. 454, pp. 903–995, 1998.

[6] N. Rehman and D. P. Mandic, "Multivariate Empirical Mode Decomposition", *Proceedings of the Royal Society of Engineering A,* vol. 466, no. 2117, pp. 1291–1302, 2009.

[7] J. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, and Ph. Bunel, "Image analysis by bi-dimensional empirical mode decomposition", *Image and Visual Computing*, vol. 21, no. 12, pp. 10191026, 2003.

[8] A. Linderhed, "Image EMD: A new tool for image processing", *Advances in Adaptive Data Analysis*, vol. 1, no. 2, pp. 265294, 2009.

[9] Y. Zheng, E. A. Essock, B. C. Hansen and A. M. Haun, "A new metric based on extended spatial frequency and its application to DWT based fusion algorithms", *Information Fusion* vol. 8, no. 2, pp. 177–192, 2005.

[10] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure", *IEEE Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.

[11] C. Schmid, and R. Mohr, and C. Bauckhage, "Evaluation of Interest Point Detectors", *International Journal of Computer Vision,* vol. 37, no. 2, pp. 151-172, 2000.

[12] J. Nunez, X. Otazu, O. Aldrich, A. Prades, V. Pala, R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition", *IEEE Transactions on Geoscience and Remote Sensing*, vol.37, no.3, pp. 1204-1211, 1999.