Morphology and EMD-Based Patch-Wise Image Fusion



Y. Pavan Kumar Reddy, Y. Sunanda, S. Charitha Adena, Sharmila Eruru, Naveen Chennamsetty, and Naveen Reddy Pothireddy

Abstract The process of integrating various source photos into a single image that is more informative than all the source images is known as image fusion. It is an efficient way of retrieving the information from the multiple sources into single image. The main purpose of image fusion is to not only decrease amount of data but also construct images that are more appropriate and comprehensible for human and machine perceptions. This paper describes morphology and empirical mode decomposition (EMD)-based image fusion strategy. The goal of this technique is to minimise the spatial distortions caused by noisy attributes of pixel-wise maps and to construct fusion images of high quality. Initially, we design a multi-channel, bidimensional empirical mode decomposition (EMD) algorithm that divides the image data into IMFs of different scales and a residue utilising morphological dilation as well as erosion filters. While retaining the decomposing quality, it further increases the computing efficiency of EMD. Additionally, we create a patch-based fusion method that merges the IMFs as well as the residue with intersecting partitions to reduce noisy attributes.

Keywords Empirical mode decomposition • Image fusion • Morphological filters • Patch-based fusion • Intrinsic mode functions

121

Y. Pavan Kumar Reddy · Y. Sunanda (\boxtimes) · S. Charitha Adena · S. Eruru · N. Chennamsetty · N. R. Pothireddy

Department of Electronics and Communication Engineering, Annamacharya Institute of Technology and Sciences, Rajampet, India e-mail: sunanda.bujji@gmail.com

[©] The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2023 A. Kumar et al. (eds.), *Proceedings of the 4th International Conference on Data Science, Machine Learning and Applications*, Lecture Notes in Electrical Engineering 1038, https://doi.org/10.1007/978-981-99-2058-7_14

1 Introduction

A wide range of image acquisition sensors are available as a result of technological advancements. The information collected with a single image acquisition sensor is insufficient, irrespective of the fact that each sensor offers characteristics that cannot be replaced in its optimum operating environment and range. Image fusion is a technique for creating a composite image from several different source images. It is an effective method of combining significant data from various sources into a single image [1]. Image fusion objects to generate images that remain more relevant and understandable for both human and machine perceptions, as well as lower the amount of data necessary to hold numerous images. Image fusion is widely cast off in computer apparition, medical imaging, remote sensing, military, etc. Various benefits of image fusion include image sharpening, feature extraction, replacement of defective data.

Different fusion techniques like (Intensity Hue Saturation), high-pass filtering, pyramid techniques, wavelet transform, discrete cosine transform are developed so far. Huang proposed a classical EMD algorithm that can decompose the onedimensional (1D) time series signal into different IMFs and residue using iterative sifting process. It is further developed for 2D images known as bidirectional empirical mode decomposition (EMD) which is applied to image fusion by many researchers to extract the feature and to overcome the distortion introduced by pre-defined functions based on transformation techniques like wavelet transform and Fourier transform [2]. Still, the known EMD methods have some drawbacks. They are highly time consuming with the increase in image size which in turn reduces computation efficiency. While merging IMFs, the spatial distortions are caused by noisy attributes of pixel-wise maps. So, these fusion methods will produce inappropriate outcomes.

To enhance the efficiency of EMD-centred fusion technique, this research paper proposes a morphology and EMD-based fusion technique. Initially, to enhance computation efficiency, we design a multi-channel bidirectional EMD algorithm using morphological dilation and erosion filters which can moulder the source images obsessed by IMFs of various scales in addition to a residue. Further, to reduce spatial distortion, we design a patch fusion technique with overlapping partitions, where maximum selection rule based on energy levels is developed to merge IMFs and residue, and the final output image is obtained by accumulating all the IMFs and also the residue collected.

The assistances in this paper include developing a morphological filter-based empirical decomposition algorithm for multi-channel images and patch-based fusion technique to fuse IMFs and the residue which can further minimise the decomposition time and maximise computation efficiency.

2 Literature Survey

Huang proposed a classical 1D EMD algorithm [3] for processing non-stationary and nonlinear 1D time signals. Through an iterative sifting process, any complex set of data is divided into a pre-defined and indeed small number of intrinsic mode functions and residue. It is confined to only 1D signals. Later, Nunes and their fellows developed it for 2D images and proposed image [4] analysis by bidirectional empirical mode decomposition which lacks stability. In order to improve the dissolution by averaging algorithms of all noise-added images, Wang [5] created a BEEMD approach, that also consists of highly costly time expenses. By using automatic vehicle-selected selective restoration and improved fast empirical mode decomposition, Trusiak's [6] advanced computation of optical fringe patterns was proposed. While using segmentation method instead of order statistics filters, this will increase the efficiency of the algorithm of envelops, but it is only applicable to single-channel images.

A wide range of fusion techniques have been developed in spatial, frequency transform, deep learning, and neural network-based domains. The decomposition algorithms are mainly suited for transform-domain fusion techniques. The input source images' transformed coefficients, which were gathered using transform-domain techniques, are integrated, and restoration step with a conforming inverse transform results in the creation of the fused image [7, 8]. Certainly, in these techniques, the choice of the transform domain is crucial. The Laplacian pyramid, empirical mode decomposition (EMD), multi-scale geometric analysis, wavelet transform, fast Fourier transform, and other transformations have all been used till now to conduct image fusion. Unlike traditional transform techniques that rely on pre-explained basis functions. EMD is completely flexible, and data dependent. Qin [9] created the decomposition. The extreme selection criteria dependent on two saliency parameters and a pixel-based algorithm were used to combine the residual and all IMFs, which may have caused some distortions. The multivariate 1D EMD is cast off to dissect source images in order to equalise the quantity and properties of the decompositions of various source images [10]. A variance-based weighted averaging method can then be used to aggregate each component pixel by pixel. In order to gain the multiscale breakdown, Xia [11] used the MBEMD based on surface projection, which may progress the fusion excellence of the multivariate one-dimensional empirical mode decomposition-based fusion technique. To handle Zhu's [12] ground-breaking fusion technique, sparse representation (SR) and bivariate bidimensional empirical mode decomposition (B-BEMD) are used, by properly combining the common and novel characteristics of two patterns of pictures. In order to successfully keep the fine qualities of the source pictures, the high-frequency components are combined using the "max-absolute" method as the activity level measurement. Then, in order to emphasise the common features and reserve the innovation features, the common and innovative features among low-frequency components are extracted by the deftly devised SR-based approach and fused, respectively, by the appropriate fusion rules. The fused picture is then rebuilt using the inverse B-BEMD procedure. Sufyan [13] proposed a new MMAI fusion method constructed on structure extraction and contrast which eliminates distortions from source images and then fuses the images based on local contrast and salient structure [14]. This paper presents a morphological filter-based empirical mode decomposition algorithm for multi-channel images and then the extracted IMFs and residue are fused with the help of patch-based fusion technique, where a maximum selection rule based on energy levels is employed [15]. This method increases the computation efficiency and minimises the decomposition time [16].

3 Methodology

3.1 Existing Method

IHS (Intensity Hue Saturation) Transform

The three characteristics of a colour—intensity, hue, and saturation—provide a regulated visual representation of an image. The IHS transform method is the most traditional picture fusion technique. Because the IHS space carries the majority of the spectral information, hue and saturation need to be carefully regulated [17].

High-pass filtering (HPF)

High-pass filtering is used to create high-resolution multispectral photographs. The high-frequency data from the high-resolution panchromatic image and the low-resolution multispectral image are combined to create the final image [18]. Either a high-pass filter is employed to filter the high-resolution panchromatic image or the original HRPI is used and the LRPI is removed from it.

Wavelet Transform

The wavelet transform is an alternative to the rapid Fourier transforms. The Fourier transform only provides the proper resolution in the frequency domain, but this method supplies it in both the time domain and the frequency domain [19]. In contrast to the wavelet transform, which scales and shifts versions of the mother wavelet or function, the Fourier transform separates the signal into sine waves of various frequencies.

Discrete Cosine Transform

It has become important for the MPEG, JVT, and other compressed picture formats. The spatial domain image is transformed into frequency-domain image using the discrete cosine transform [20]. Low frequency, medium frequency, and high frequency are three categories used to divide the images. The DC value reflects average illumination, whereas the AC values are the high-frequency coefficients. The RGB picture is divided into 8×8 pixel blocks for segmentation. The picture

is then turned into a greyscale image after being separated into groups based on the red, green, and blue matrices [21].

3.2 Proposed Method

The bidirectional EMD method, which is based on morphological dilation and erosion filters, first divides the contribution source images hooked on numerous IMFs and also a residue. Second, it uses an overlapping patch-based fusion technique to fuse the residue and IMFs separately. A maximum selection method based on energy levels is constructed about the fusion of the IMFs, along with two separate rules which are built for the fusing of the residue based on the key information collected by IMFs from the input images. The intrinsic mode functions and also fused residue are ultimately used to reconstruct the required fused image.

The block diagram below depicts the entire structure of EMD algorithm (Fig. 1).

Morphological Filter Based EMD Algorithm

In the proposed morphological filter-based multi-channel bidirectional empirical mode decomposition, using morphological dilation as well as erosion filters which have the same window size for every channel, which retrieve the very same spatial extent from every channel image at the moment of decomposition, envelope edges for the inter image are produced. The lower (upper) envelope $D = (D_1, ..., D_n)$ ($(U = (U_1, ..., U_n))$ for a multi-channel image $I = (I_1, ..., I_n)$ with window size



Fig. 1 Block diagram of morphology and EMD-based patch-wise image fusion

 $W \times H$ can be generated by

$$U_{k}(x, y)|_{k=1,...,n} = (I_{k} \oplus b)(x, y) = \max_{(s,t) \in Z_{xy}} I_{k}(s, t)$$
$$D_{k}(x, y)|_{k=1,...,n} = (I_{k} \oplus b)(x, y) = \min_{(s,t) \in Z_{xy}} I_{k}(s, t)$$
(1)

where \oplus represents the morphological dilation filter, \ominus represents the morphological erosion filter, b characterises a binary group pointer function on Z_{xy} , and Z_{xy} denotes the group of pixels centred on the pixel (x, y) in the window w × w. To obtain significantly smoother envelopes, the average filter is utilised.

$$U_{k}'(x, y)\Big|_{k=1,...,n} = \frac{1}{\omega \times \omega} \sum_{(s,t) \in Z_{xy}} U_{k}(s, t)$$
$$D_{k}'(x, y)\Big|_{k=1,...,n} = \frac{1}{\omega \times \omega} \sum_{(s,t) \in Z_{xy}} D_{k}(s, t)$$
(2)

The window size w in Eqs. (1) and (2) is set to the smallest average extreme distance of all image channels in order to evaluate feature abstraction for all statistics channels of the source images.

$$\omega = \min\{\omega 1, ..., \omega n\}$$
(3)

where average extreme space of *k*th channel picture I_k is represented by w_k (k = 1, ..., n) and is determined by

$$\omega_k = \sqrt{\frac{w \times h}{N_k}} \tag{4}$$

where N_k represents the average value of all I_k 's local maxima and minima. In each iteration, this compares the values of each pixel and neighbourhood pixels in 3×3 window centred on it to locate all local maxima (minima) of I_k . This could iteratively extract IMFs of various scales using a sifting technique based on the envelope calculation technique described above, until residue is a monotonic function or a constant or the required number of IMFs is obtained.

EMD-Based Patch-Wise Image Fusion

The two source images I_1 and I_2 are combined to form a two-channel image $I = (I_1, I_2)$ that is decomposed by Algorithm 1 into K IMFs and a residue.

$$I = \sum_{l=1}^{K} F_l + R_k \tag{5}$$

where $F_i = (F_{i1}, F_{i2})(i = 1, ..., k)$ is the *i*th IMF and $R_k = (R_{K1}, R_{K2})$ is the associated residue. All intrinsic mode functions and residues are divided into several patches of size $M \times M$, with N overlapping columns/rows. This overlapping patch technique is developed to minimise the distortions that occur around partition boundary while using patch-based fusion techniques.

Fusion of IMFs

On measuring the energy levels of two related patches, the fused patch G_i^j is generated using a maximum selection rule based on energy levels, for the *j*th patch $F_i^j = (F_{i1}^j, F_{i2}^j)$ of *i*th IMF F_i

$$G_{i}^{j} = \begin{cases} F_{i1}^{j}, E(F_{i1}^{j}) \ge E(F_{i2}^{j}) \\ F_{i2}^{j}, E(F_{i1}^{j}) < E(F_{i2}^{j}) \end{cases}$$
(6)

Equation (6) computes the energy of each patch by

$$E(F_{ip}^{j}) = \sum_{(s,t)\in Z_{j}} F_{ip}^{j}(s,t)^{2}, p = 1, 2,$$
(7)

where Z_j stands for the pixels group in *j*th patch. The above formula is utilised to obtain the significant features from the source pictures.

Fusion of Residue

As for *j*th patch $R_K^j = (R_{K1}^j, R_{K2}^j)$ of residue R_K , two different methods are designed centred on the statistics collected by the intrinsic mode functions to extract fusion residue of patch H_k^j in accordance with the image types.

The first one combines multi-focus images using a maximum selection method that is energy-based. In the first IMF, the energy of two identical patches is compared as

$$H_{k}^{j} = \begin{cases} R_{K1}^{j}, E\left(F_{11}^{j}\right) \ge E\left(F_{12}^{j}\right) \\ R_{K2}^{j}, E\left(F_{11}^{j}\right) < E\left(F_{12}^{j}\right) \end{cases}$$
(8)

where $E(F_{ip}^{j})$ (p = 1, 2) represents the energy of initial intrinsic mode functions, and this could acquire the features up to finest scales. The above fusion method accurately describes the focused area of multi-focus pictures.

The second combines multi-modal images using an energy-based algorithm. The knowledge area retrieved by IMFs is used to merge the actually imply area of the residue patch, and the mean of a residue serves is helpful to fuse the illumination of each multi-modal image. The fusion equation can be obtained from

$$H_{k}^{j} = \sum_{p=1}^{2} a_{p}^{j} (R_{Kp}^{j} - \mu_{Kp}^{j}) + \sum_{p=1}^{2} b_{p}^{j} \mu_{Kp}^{j}$$

$$a_{p}^{j} = \left| \sum_{i=1}^{K} E(F_{ip}^{j}) \right|^{l} / \left(\left| \sum_{i=1}^{K} E(F_{i1}^{j}) \right|^{l} + \left| \sum_{i=1}^{K} E(F_{i2}^{j}) \right|^{l} \right)$$

$$b_{p}^{j} = \frac{\left| \mu_{Kp}^{j} \right|^{m}}{\left| \mu_{K1}^{j} \right|^{m} + \left| \mu_{K2}^{j} \right|^{m}}$$
(9)

where μ_{kp}^{j} is mean of the *j*th patch of residue R_{kp}^{j} . The two non-negative exponent parameters for controlling feature guidance intensity as well as brightness fusion intensity, respectively, are p = 1, 2, 1, and *m*. If *l* and *m* are both set to zero, the result is just an average of the leftover information from two multi-modal images. The knowledge area retrieved by IMFs is used as the guidance to merge the brightness of each modal image if the virtues *l* and *m* have been set greater than 0, and the mean of the residue patch is used to fuse the mean divided section of the residue of *each* modal image if the results of *l* and *m* are set larger than 0. A higher value of *l* denotes that the merged result contains much stronger features, while a greater value of *m* indicates that more brilliant targets are included in the fusion result. Both the values of *l* and *m* are set to 6 in trials, which can produce successful outcomes.

Image Reconstruction

The value at each pixel (x, y) of the fused IMFs and residue is determined by averaging the values of the pixel (x, y) in all overlapping patches after all IMFs and residue patches have been fused.

$$G'_{i}(x, y) = \frac{1}{s(x, y)} \sum_{j} G^{j}_{i}(x, y)$$
$$H'_{K}(x, y) = \frac{1}{s(x, y)} \sum_{j} H^{j}_{K}(x, y)$$
(10)

where S(x, y) represents the overlapping patch number at the pixel (x, y), and the resultant fused image I' is created by combining the fused IMFs and the fused residue.

$$I'(x, y) = \sum_{i=1}^{k} G'_i(x, y) + H'_K(x, y)$$
(11)

The proposed method is implemented by setting the initial value of fused IMFs at each pixel (x, y), the fused residue, then with the help of overlapping patch number by $G'_i(x, y)|_{i=1,...,k} = 0$, $H'_K(x, y) = 0$, and S(x, y) = 0. Each time a patch is combined, and fused values are updated using

$$G'_{i}(x, y) \to G'_{i}(x, y) + G^{j}_{i}(x, y)$$

 $H'_{K}(x, y) \to H'_{K}(x, y) + H^{j}_{K}(x, y)$ (12)

and at the *j*th patch, each pixel's overlapping patch value is modified by

$$S(x, y) \to S(x, y) + 1 \tag{13}$$

After fusing all the patches of K IMFs and residue, on every pixel, the resultant joined IMFs and residue are retrieved.

$$G'_{i}(x, y) \rightarrow \frac{1}{S(x, y)}G'_{i}(x, y)$$

$$H'_{K}(x, y) \rightarrow \frac{1}{S(x, y)}H'_{K}(x, y)$$
(14)

In pixel-based fusion, the noisy features of pixel-wise maps are reduced with the help of overlapping patch partition.

4 Experimental Investigations

4.1 Selection of Key Parameters

K Decomposition level. The most effective selection principle based on the energy thresholds is used to combine all IMFs and extract the most important information from the images. However, K is chosen over 2 for multi-modal images so because top two IMFs of method 2 are where the majority of the input images' information is concentrated in the tests performed multi-modal images.

Overlapping number *N* **of rows/columns** Most of the time, decreasing distortions while also increasing computing costs can be accomplished by increasing the number of rows and columns that overlap *N* in the patch split. The block sizes are M = 2 for multi-modal images and M = 6 for multi-focus images. More tests have shown that these decisions can generate desired fusion outcomes.

Block size M of the division. Combining all IMFs, the most efficient selection premise using the energy threshold values is used to get the most important information from the images. The very same decomposition threshold K of sample b is set to 1 in sequence for the initial IMF to signify that the focused province of multi-focus images is good. The top two IMFs of step 2 represent that the large number of input



Fig. 2 Multi-focus colour image fusion using proposed method



Fig. 3 Multi-focus greyscale image fusion using proposed method

tiny pictures' information is focused in the tests conducted with multi-modal image sets, so K is chosen over 2 in this case.

4.2 Results

The MATLAB R2017b software was used to create all the experimental results shown in this work on a laptop with an Intel Core i5 processor with Windows 11 operating system and RAM size is 16 GB.

Using morphology and EMD-based patch-wise image fusion, Figs. 2, 3, 4, and 5 show the fusion of multifocal (colour) photos, greyscale images, multi-modal (medical) images, and infrared images, respectively. Given that the maximum selection rule based on energy levels for the fusion of each IMF can extract more significant information, the patch-based fusion technique can enhance the fusion quality of each EMD method in visualisation while also reducing the distortions caused by pixel-wise fusion method. The structure of multi-modal images can also be better represented by the extracted IMFs' energy-based weighted averaging method, and the focused area of multi-focus images can be captured more effectively by the first IMF's activity level. It is clearly observed that the essential details present in the output image but absent in either of the source images.



Fig. 4 Multi-modal medical image fusion using proposed method



Fig. 5 Multi-modal infrared image fusion using proposed method

5 Conclusion

To obtain good quality images, this paper describes a morphology and an EMD-based patch-wise image fusion. First of all, a morphological filter-based bidirectional EMD algorithm is developed for multi-channel images which uses dilation and erosion filters to calculate lower and upper envelopes of source images. This algorithm breaks down the input images into intrinsic mode functions of various sizes and a residue. This will gradually improve the computation efficiency. The IMFs and residue are then merged using a patch-based fusion method with overlapping partitions. With the aid of maximum selection rule based on energy levels, the IMFs are fused, and the residue is combined using the key information they have collected. The fused image is finally produced by combining the fused IMFs with all of the fused residues. The fused images for different sets of source images are displayed in Figs. 2, 3, 4, 5.

References

- 1. Karimullah S, Vishnuvardhan D, Basha SJ (2020) Floorplanning for placement of Modulesin VLSI physical design using harmony search technique. ICDSMLA 2019, Lecture notes in electrical engineering 601, Springer Nature Singapore Pte Ltd.
- 2. Jiang Y, Wang M (2014) Image fusion with morphological component analysis. Inf Fusion 18(7):107–118
- Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen NC, Tung CC, Liu HH (1998) The empirical mode decom- position and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc Math Phys Eng Sci 454(1971):903–995
- 4. Nunes JC, Bouaoune Y, Delechelle E, Niang O, Bunel P (2003) Image analysis by bidimensional empirical mode decomposition. Image Vis Comput 21(12):1019–1026
- Wang X, Hu J, Guo L, Zhang D, Hong Q, Hao A (2018) Featurepreserving, mesh-free empirical mode decomposition for point clouds and its applications. Comput. Aided Geomet. Design 59:1–16
- Trusiak M, Wielgus M, Patorski K (2014) Advanced processing of optical fringe patterns by automated selective reconstruction and enhanced fast empirical mode decomposition. Opt Lasers Eng 52:230–240
- Liu Y, Chen X, Wang Z, Wang ZJ, Ward RK, Wang X (2018) Deep learning for pixel-level image fusion: recent advances and future prospects. Inf Fus 42:158–173
- Yu L, Lei W, Juan C, Chang L, Xun C (2020) Multi-focus image fusion: a survey of the state of the art. Inf Fus 64:71–91
- Qin X, Zheng J, Hu G, Wang J (2017) Multi-focus image fusion based on window empirical mode decomposition. Infrared Phys Technol 85:251–260
- Reddy YPK, Vaishnavi A, Devi MS, Prasad MS, Reddy BS (2023) Multimodal medical image fusion approach using PCNN model and shearlet transforms via max flat FIR filter. In: Kumar A, Senatore S, Gunjan VK (eds) ICDSMLA 2021. Lecture Notes in Electrical Engineering, vol 947. Springer, Singapore. https://doi.org/10.1007/978-981-19-5936-3_73
- Xia Y, Zhang B, Pei W, Mandic DP (2019) Bidimensional multivariate empirical mode decomposition with applications in multi-scale image fusion. IEEE Access 7:114261–114270
- 12. Zhu P, Liu L, Zhou X (2021) Infrared polarization and intensity image fusion based on bivariate BEMD and sparse representation. Multimed Tools Appl 80:4455–4471
- Sufyan A, Imran M, Shah SA, Shahwani H, Wadood AA (2022) A novel multimodality anatomical image fusion method based on contrast and structure extraction. Int J Imag Syst Technol 32(1):324–342
- Karimullah S, Vishnuvardhan D (2020) Experimental analysis of optimization techniques for placement and routing in Asic design. ICDSMLA 2019, Lecture notes in electrical engineering 601, Springer Nature Singapore Pte Ltd.
- Shaik F, Sharma AK, Ahmed SM (2016) Hybrid model for analysis of abnormalities in diabetic cardiomyopathy and diabetic retinopathy related images. Springerplus 5:507. https://doi.org/ 10.1186/s40064-016-2152-2
- Karimullah S, Vishnuvardhan D (2020) Iterative analysis of optimization algorithms for placement and routing in Asic design. ICDSMLA 2019, Lecture notes in electrical engineering 601, Springer Nature Singapore Pte Ltd.
- 17. Karimullah S, Vishnuvardhan D (2022) Pin density technique for congestion estimation and reduction of optimized design during placement and routing. Applied nanoscience
- Nagaraju CH, Sharma AK, Subramanyam MV (2018) Reduction of PAPR In MIMO-OFDM using adaptive SLM And PTS technique. International Journal of Pure and Applied Mathematics, Special issue 118(17):355–373. ISSN: 1311-8080 (printed version); ISSN: 1314-3395
- Karimullah S, Basha SJ, Guruvyshnavi P, Sathish Kumar Reddy K, Navyatha B (2020) A genetic algorithm with fixed open approach for placements and routings. ICCCE, Springer, pp 599–610

- Karimullah S, Vishnuvardhan D, Arif M, Gunjan VK, Shaik F, Siddiquee KN (2022) An improved harmony search approach for block placement for VLSI design automation. Wireless Communications and Mobile Computing, vol 2022, Article ID 3016709, 10 pages. https://doi. org/10.1155/2022/3016709
- Li H, Wang Y, Yang Z et al (2020) Discriminative dictionary learning-based multiplecomponent decomposition for detail-preserving noisy image fusion. IEEE Trans Instrum Meas 69(4):1082– 1102