Multi-modal Face Image Fusion using Empirical Mode Decomposition

Harishwaran Hariharan, Andrei Gribok, Besma Abidi and Mongi Abidi
Imaging Robotics and Intelligent Systems Laboratory
The University of Tennessee, Knoxville, Tennessee, 37996-2100
{hari,agribok,besma,abidi}@utk.edu

1. Introduction

Using multiple modalities for face recognition has proven to increase the recognition rates as opposed to conventional single modality systems, especially in challenging illumination conditions and in the case of disguised individuals. Images from the multiple modalities are fused to increase the information content in a resultant fused image. The fused image would have enhanced information which is more understandable and decipherable for face recognition applications. In this effort, we describe a novel technique for image fusion and enhancement, using Empirical Mode Decomposition (EMD). EMD is a non-parametric data-driven analysis tool that decomposes non-linear non-stationary signals into Intrinsic Mode Functions (IMF). In this method, we decompose images from different imaging modalities (visible-color and thermal) into their IMF. Fusion is performed at the decomposition level and the fused IMFs are reconstructed to realize the fused image. Based on an empirical understanding of the nature of the IMF we have devised weighting schemes which emphasize features from both modalities, thereby increasing the information and visual content of the fused image.

2. Image Fusion using Empirical Mode Decomposition

2.1 Empirical Mode Decomposition

In this effort, we harness the potential of a relatively new method for analyzing nonlinear and non-stationary datasets developed by Huang *et al* [1]. The assumptions for this method are (1) the signal has at least a pair of extrema; (2) the characteristic time scale is defined by the time between the successive extrema; and (3) if there are no extrema, and only inflection points, then the signal can be differentiated to realize the extrema, where all IMF can be extracted. As per the IMF definition, the decomposition method employs the envelopes defined by the local maxima and minima individually. The extrema envelopes are identified and the point-wise mean of the envelopes (m_1), is subtracted from the data r_0 for the first component h_1. For the first sift, r_0 is the original data X(t). In the second sifting, h_1 is considered as the data and the second component h_{11} is found using $h_{11} = h_1 - m_{11}$, where m_{11} is the mean of the h_1 extrema envelopes. While sifting, the finest oscillatory modes are separated from the data, analogous to separating fine particles through a set of fine to coarse sieves. But if performed too long, the sifting process becomes invasive and destroys the physical meaning of the amplitude fluctuations. The k^{th} sifted component is given by $h_{1k} = h_{1(k-1)} - m_{1k}$. The standard deviation, SD, computed from two consecutive sifting results, is used as criteria to stop sifting, given by, $SD = \sum_{t=0}^{T}\left[\frac{\left|h_{1(k-1)} - h_{1(k)}\right|^2}{h_{1(k-1)}^2}\right]$. Sifting is ceased if SD falls below a threshold. We designate c_1 as the first of the IMF, $c_1 = h_{1k}$. The isolated intrinsic mode function, c_1 contains the finest scale of the signal and we separate c_1 from the data, to obtain the new signal called the residue, by $r_1 = r_0 - c_1$, which still holds lower frequency information. In the next iteration, the residue r_1 is treated as the new data in place of r_0 and subjected to the sifting process. This procedure is repeated on all the subsequent residues r_j's, to realize a set of IMF, where $r_1 - c_2 = r_2, \ldots.., r_{n-1} - c_n = r_n$. The iterations on the residuals can be stopped by any of the following stopping criteria; if the residue becomes too small to be of any practical importance, or when the residue becomes a monotonic function containing no IMF. Thus, the data is decomposed into a set of n-empirical modes, c_i's, and the final residue, r_n, which can be either the mean trend or a DC shift. The complete set of IMF, $\hat{X}(t)$, is notated by,

$$\hat{X}(t) = \sum_{i=1}^{n} c_i + r_n. \tag{1}$$

2.2 Image Fusion Using EMD

The EMD theory was originally proposed for one dimensional data. It has been extended for two-dimensional data in the literature towards image compression [2, 3], and in texture analysis [4]. The input images were vectorized in lexicographical order. In our implementation, EMD was performed on each channel vector separately and the facet of decomposing a signal into IMF is employed in the fusion process. The key contributions in our algorithm are the use of EMD to decompose co-registered input images, and establishing fusion rules at the decomposition level. The decomposed IMF are multiplied by a set of weights that are chosen in a manner to minimize the mutual information between the corresponding IMF of the visual and thermal modalities. An empirical understanding of the intrinsic mode images and subjective inspection is used to guide the fusion process. The compact form of the fusion scheme used in our method is given by,

$$\Im(x,y)=\sum_{i=1}^{3}\sum_{j=1}^{k}\left[\alpha_{ij}V_{ij}+\beta_{ij}T_{ij}\right]. \tag{2}$$

where $\Im(x,y)$ is the fused image, α_{ij}, is the weight by which the j^{th} visual IMF is multiplied, V_{ij} is the j^{th} visual IMF corresponding to the i^{th} channel, β_{ij} is the weight by which the j^{th} thermal IMF is multiplied and, T_{ij} is the j^{th} thermal IMF corresponding to the i^{th} channel.

3. Results and Conclusion

A novel and efficient technique for image fusion using empirical mode decomposition has been proposed. Our fusion technique augments information from both input images used for fusion experiments. As input, we use registered thermal and visible color images. EMD is used to obtain the decomposed IMF of the various channels of the visible and thermal image. Fusion is performed at the IMF level. Weighting schemes that reduce the mutual information between the IMF are used to emphasize features or, to discourage distracting features from the corresponding modalities. We are presently working on a data-driven scheme to select optimal weights for image fusion to advance this technique. Tests on face databases using various illuminators are also planned.

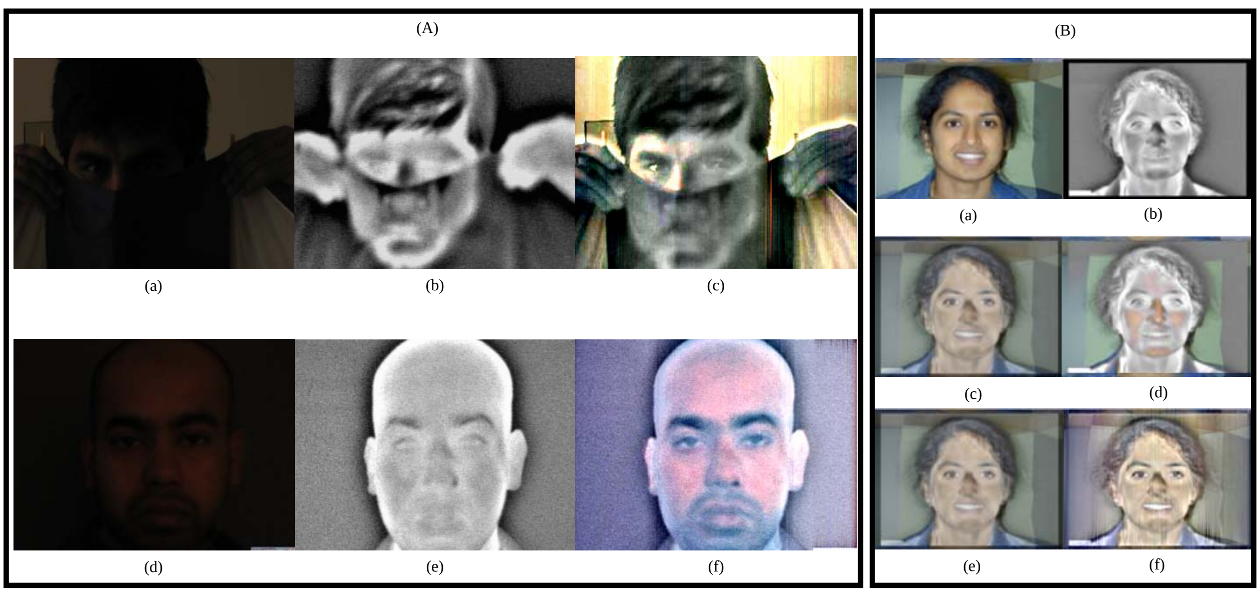

Fig. 1. **Panel A**: Some results of EMD fusion; (a,d) Low-illumination visible images; (b,e) Thermal images; (c,f) Output images, fused using our method. **Panel B**: Comparison with some other widely used fusion methods; (a) Visible input image; (b) Thermal input image; (c) Pixel averaged image; (d) Image fused using wavelet based fusion method; (e) Image fused via fusion of principal components; (f) Image fused using EMD.

4. References

[1] 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. R. Soc. London. A.*, vol. 454, 1998, pp. 903-995.

[2] A. Linderhed, “Image compression based on empirical mode decomposition,” *Proc. of SSAB Symposium Image Analysis*, Uppsala, March 2004, pp. 110-113.

[3] A. Linderhed, "2D empirical mode decompositions in the spirit of image compression", Wavelet and Independent Component Analysis Applications IX , SPIE Proceedings Vol. 4738, April 2002, pp: 1-8.

[4] Z. Yang, D. Qi, and L. Yang, “Signal period analysis based on Hilbert-Huang transform and its application to texture analysis,” *Intl. Conf. Image and Graphic, Hong Kong, 2004.*