



Image Fusion and Enhancement via Empirical Mode Decomposition

Harishwaran Hariharan

Imaging, Robotics, and Intelligent Systems Lab.

*Department of Electrical and Computer Engineering, University of Tennessee
410 Science & Engineering Building, Knoxville, TN 37996, USA*

hari@utk.edu

Andrei Gribok

*Nuclear Engineering Department, University of Tennessee
315 Pasqua Engineering, Knoxville, Tennessee 37996, USA*

agribok@utk.edu

Mongi A. Abidi

Imaging, Robotics, and Intelligent Systems Lab.

*Department of Electrical and Computer Engineering, University of Tennessee
328 Ferris Hall, Knoxville, TN 37996, USA*

abidi@utk.edu

Andreas Koschan

Imaging, Robotics, and Intelligent Systems Lab.

*Department of Electrical and Computer Engineering, University of Tennessee
330 Ferris Hall, Knoxville, TN 37996, USA*

akoschan@utk.edu

Invited paper. Published January 20, 2006.

Abstract

In this paper, 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 (IMFs). In this method, we decompose images, rather than signals, from different imaging modalities into their IMFs. Fusion is performed at the decomposition level and the fused IMFs are reconstructed to realize the fused image. We have devised weighting schemes which emphasize features from both modalities by decreasing the mutual information between IMFs, thereby increasing the information and visual content of the fused image. We demonstrate how the proposed method improves the interpretive information of the input images, by comparing it with widely used fusion schemes. Apart from comparing our method with some advanced techniques, we have also evaluated our method against pixel-by-pixel averaging, a comparison, which incidentally, is not common in the literature.

Keywords: Data fusion, Empirical mode decomposition, Image fusion, Intrinsic mode function.

1. Introduction

Image fusion is the capacity to produce a single fused image from a set of input images. The fused image has enhanced information that is more understandable and decipherable for human perception and, preferably, for machine learning and computer vision. Images from different imaging sensors have been fused using various techniques in the literature [1-3]. The necessity for fusion techniques increased with the inception of new image acquisition devices. By fusing images, it is possible to discern the useful information from the input images. Apart from collating the contributive features, an image fusion scheme of a higher abstraction suppresses inconsistencies, artifacts and noise in the fused images, which are distractions in the parent images.

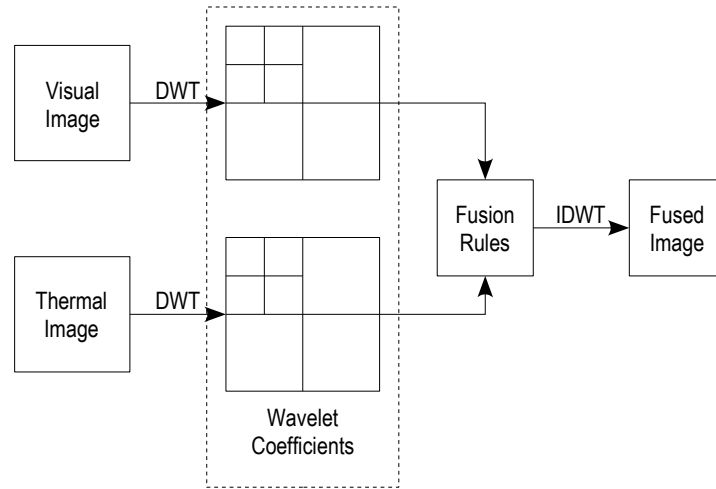


Fig. 1: General framework for wavelet-based image fusion techniques.

In this paper, we propose an image fusion technique wherein image decomposition is performed on input images by EMD. The basis functions are subjected to a weighting scheme that minimizes mutual information and enriches information from the input images. The weighting scheme emphasizes desired, and de-emphasizes undesired features, in the fused image.

In section 2, a review of the image fusion techniques that are widely used is presented. The review presented is limited to fusion methods that are relevant or similar to our method. The foundation of our method lies in the EMD algorithm, which is explained in section 3. Our fusion technique is described in detail in section 4. Image pairs with a high degree of alignment and spatial correspondence were used in our experiments. In section 5, the experimental results are shown and compared with the widely used fusion methods. Our conclusions and plans for the future in advancing our technology are presented in the last section.

2. Review of Existing Methods

The fusion of images has been performed using various techniques. The data used in these efforts stem from a variety of imaging sensors. Thermal and visual data were fused using the most widely-used fusion methods to compare with our results. The fusion methods that we have used as benchmarks have been explained herewith.

2.1 Wavelet-based Fusion

Wavelet based methods are common in image fusion. The wavelet transform is a data analysis tool that provides a multi-resolution decomposition of an image. The input image is decomposed into a set of wavelet decomposition levels. The basis functions are generated from one single basis function popularly referred to as the mother wavelet. The mother wavelet is shifted and scaled to obtain the basis functions. Wavelet decomposition can be administered on an image in many ways. The fusion of images using wavelets follows a standard procedure and is performed at the decomposition level. The standard operation of a wavelet fusion scheme is outlined in Fig. 1.

The input images are decomposed by a discrete wavelet transform and the wavelet coefficients are selected using a set of fusion rules and an inverse discrete wavelet transform is performed to reconstruct the fused image. Wavelet fusion methods differ mostly in the fusion rule used for selection of wavelet coefficients. Selection of the wavelet coefficients with the maximum value

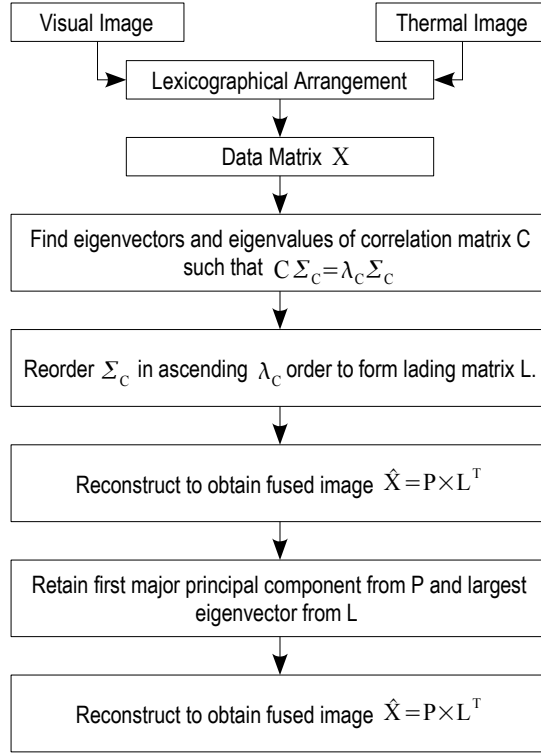


Fig. 2: Fusion of visual and thermal images using principal components.

is used as the fusion rule in [4]. Another popular fusion rule is an area-based selection criterion, where regional windows are compared and a consistency measure is used to select wavelet coefficients to reconstruct the fused image [5]. Nunez *et al.* developed a method using wavelet coefficient addition between high-resolution panchromatic image and low-resolution multi-spectral intensity images [6].

Images fused using the pixel maximum selection criteria are selected for comparison as the results obtained are noteworthy. For any decomposition level i , the maximum coefficients are selected to form the fused image $F_i(x, y)$ and x, y are pixel coordinates.

$$F_i(x, y) = \begin{cases} V_i(x, y) & \text{if } V_i(x, y) > T_i(x, y) \\ T_i(x, y) & \text{if } V_i(x, y) < T_i(x, y) \end{cases} \quad (1)$$

2.2 Principal Component (PC) Fusion

Image fusion via decomposition of a signal into its principal components has been performed in many different ways in the literature. Multiple video frames are used as inputs to the principal component analysis [7] procedure outlined in Fig. 2. Similarly in multi-spectral imaging applications, images from different spectral bands are used [8] as inputs to the PC analysis. Genetic algorithms have been used by Singh *et al.* [9] to select the principal components used for the reconstruction. In our implementation, we have used the channel information from the input images to form the data matrix. The pixels from the visual and thermal image channels are arranged in lexicographically arranged columns to form a data matrix X . An eigen analysis is performed on the correlation matrix C , to obtain the eigenvalues λ_c and eigenvectors Σ_c that

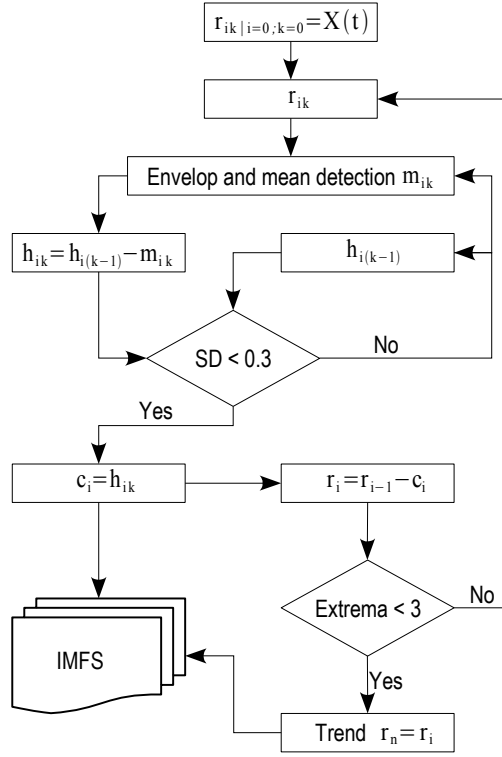


Fig. 3: A flow diagram of signal decomposition via EMD.

satisfy the following equation:

$$C = \Sigma_c^T \times \lambda_c \times \Sigma_c \quad (2)$$

where,

$$C = X^T \times X \quad (3)$$

The eigenvectors are arranged such that the eigenvector corresponding to the largest eigenvalue is the first column of the loading matrix L and the second column corresponds to the eigenvector corresponding to the second largest eigenvalue and so on. The principal components are obtained by multiplying the data matrix with the loading matrix.

$$P = X \times L \quad (4)$$

The reconstruction is performed by use of principal component synthesis, given by,

$$\hat{X} = P \times L^T \quad (5)$$

The first major principal component is multiplied with the largest eigenvector from the loading matrix to obtain the fused image. When the input images are fused using this method, many features from both the modalities are seen in the fused image. The principal component fusion scheme is outlined in Fig. 2.

2.3 Pixel-by-pixel Averaging

Many fusion schemes have been proposed with rigorous procedures to decompose the images using computationally intensive and mathematically demanding algorithms. A comparison against pixel-by-pixel averaging is a good measure of effectiveness for evaluating image fusion results. An algorithm which has little or no improvement over averaging is a futile exercise. Incidentally, comparisons with fusion-by-averaging are not very visible in the literature. We have compared images fused using our method with averaged images.

3. Empirical Mode Decomposition

In this paper, we harness the potential of a relatively recent method for analyzing nonlinear and non-stationary data sets developed by Huang et al [10]. The aspect of decomposing a signal into IMFs is employed in the fusion process. One is able to decompose any complicated data set into a finite set of IMFs that admit well-behaved Hilbert transforms. This decomposition method is data-driven and hence highly effective. The decomposition is based on the local characteristic time scale of the data, and hence extendable to nonlinear and non-stationary processes. With the Hilbert transform, the IMFs allow representation of instantaneous frequencies as functions of time. The main conceptual benefits are the decomposition of parent signal into IMFs and the visualization of time-frequency characteristics. We present the EMD algorithm and elaborate on the sifting process. The flow diagram of the EMD algorithm is shown in Fig. 3.

3.1 EMD Assumptions

Contrary to many of the former decomposition methods, EMD is intuitive and direct, with the basis functions based on and derived from the data. 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, whose IMFs can be extracted. Integration may be employed for reconstruction. The time between the successive extrema was used by Huang et al. [10] as it allowed the decomposition of signals that were all positive, all negative, or both. This implied that the data did not have to have a zero mean, as in the case of image data. This also allowed a finer resolution of the oscillatory modes.

3.2 The Sifting Process

As per the definition of an IMF, the decomposition method can simply employ the envelopes defined by the local maxima and minima individually. The extrema are identified and all local maxima are connected by a cubic spline to form the upper envelope. This process is repeated for the local minima and the lower envelope is constructed. While interpolating, care is taken that the upper and lower envelopes cover all the data between them. The point-wise mean of the envelopes is called m_1 , and is subtracted from the data r_0 for the first component h_1 . For the first iteration, $X(t)$ is the used as the data,

$$r_0 = X(t) \quad (6)$$

$$h_1 = r_0 - m_1 \quad (7)$$

As per mathematical definitions, h_1 should be considered as one of the IMF, as h_1 seems to satisfy all the requirements of an IMF. But since we are interpolating the extrema with numerical schemes, overshoots and undershoots are bound to occur. These generate new maxima and minima, and distort the magnitude and phase of the existing extrema. These effects will not

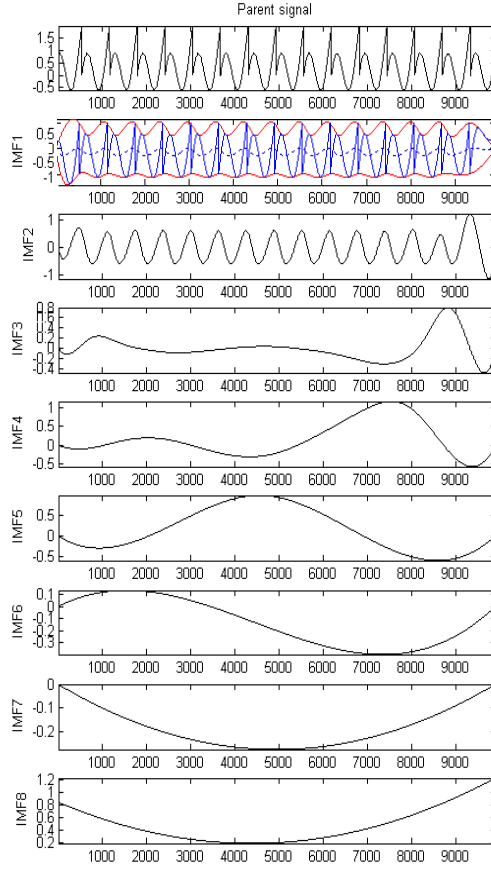


Fig. 4: Empirical mode decomposition of a synthetic data set. An example of envelope detection is also shown. The envelopes are shown by red lines and the envelope mean is shown with a dashed blue line.

affect the process directly as it is the mean of these envelopes that pass on to the next stages of the algorithm and not the envelopes themselves. The formation of false extrema cannot be avoided easily and an interesting offshoot is that this procedure inherently recovers the proper modes lost in the initial examination and recovers low-amplitude riding waves on repeated sifting.

The envelope means may be different from true local mean and consequently some asymmetric waveforms may occur but they can be ignored as their effects in the final reconstruction are very minimal. Apart from a few theoretical difficulties, in practice, a ringing effect at the ends of the data array can occur. But even with these effects, the sifting process still extracts the essential scales from the data set. The sifting process eliminates riding waves and makes the signal symmetrical. In the second sifting process, h_1 is considered as the data where m_{11} is the mean of the h_1 envelopes.

$$h_{11} = h_1 - m_{11} \quad (8)$$

The sifting is continued k times until the first IMF is obtained.

$$h_{1k} = h_{1(k-1)} - m_{1k} \quad (9)$$

We designate c_1 as the first IMF,

$$c_1 = h_{1k} . \quad (10)$$

3.3 Stopping Criteria

In sifting, the finest oscillatory modes are separated from the data, analogous to separating fine particles through a set of fine to coarse sieves. As can be expected of such a process, uneven amplitudes will be smoothed. But if performed too long, the sifting process becomes invasive and destroys the physical meaning of the amplitude fluctuations. On sifting too long, we get IMFs that are frequency modulated signals with constant amplitude. To retain the physical meanings of an IMF, in terms of amplitude and frequency modulation, a standard deviation based stopping criterion is used. The standard deviation, SD, computed from two consecutive sifting results, is used as one of the stopping criteria.

$$SD = \sum \left[\frac{|h_{i(k-1)} - h_{i(k)}|^2}{h_{i(k-1)}^2} \right] \quad (11)$$

Sifting is stopped if SD falls below a threshold. The isolated intrinsic mode function, c_1 contains the finest scale of the signal and we separate c_1 from the data.

$$r_1 = r_0 - c_1 \quad (12)$$

The new signal called the residue, r_1 , 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 IMFs.

$$r_1 - c_2 = r_2 , \dots , r_{n-1} - c_n = r_n \quad (13)$$

The sifting through 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 more IMFs. It is not expected, to always have a residue with zero mean, because even for data with zero mean, the final residue can still be different from zero. The final residue is the trend of the data. Reconstruction of the signal is performed using the relation,

$$\hat{X}(t) = \sum_{i=1}^n c_i + r_n \quad (14)$$

Thus, the data is decomposed into n-empirical modes, and a residue, r_n , which can be either the mean trend or a DC shift. In Fig. 4, a decomposition of a synthetic data set by EMD is shown. The formation of envelopes has been shown for the first IMF. The envelopes are shown in red and the envelope mean is shown in a dashed blue line.

4. Intrinsic Mode Image Fusion

The key contributions in our algorithm are the use of EMD to decompose the input images and,

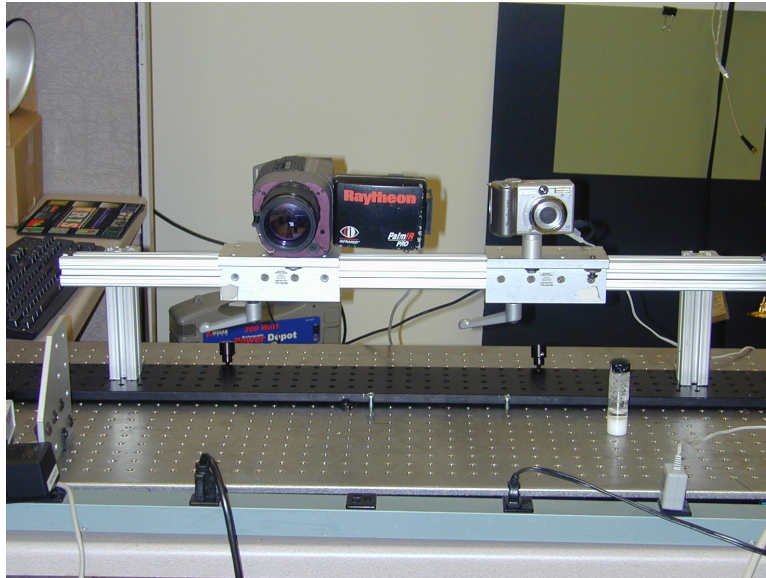


Fig. 5: Data acquisition setup used to obtain well-aligned images from thermal and visual imaging modalities.

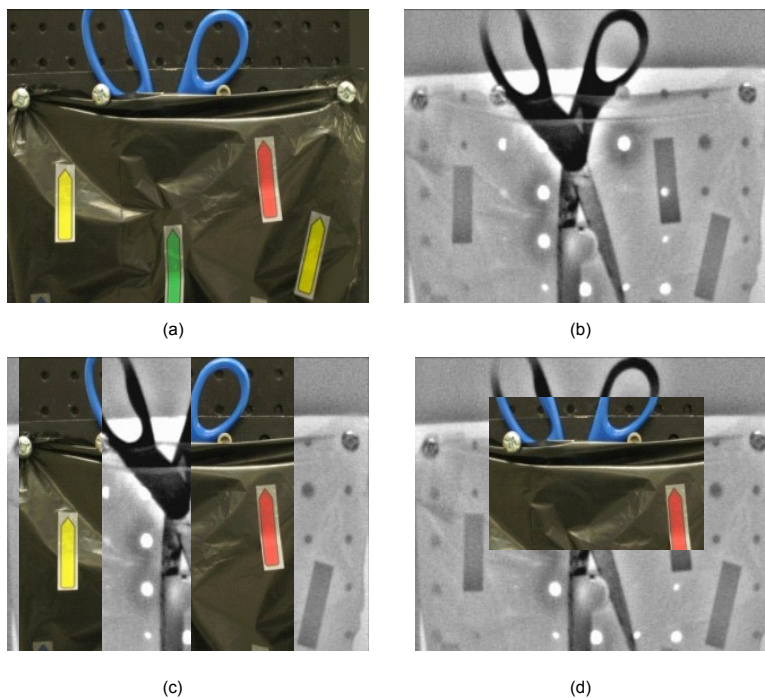


Fig. 6: A pair of test images used for experimentation, (a) visual image, (b) thermal image, (c) and (d), demonstration of the degree of registration of the test images. For example, the shape of a pair of scissors can be perceived.

establishing a weighting scheme which decreases the mutual information between IMFs. In the first sub-section, the nature of the data used in our experiments has been described. Empirical mode decomposition has been extended to images in our method and the process employed is described in the second subsection. The IMFs showed consistent behavior when visualized as

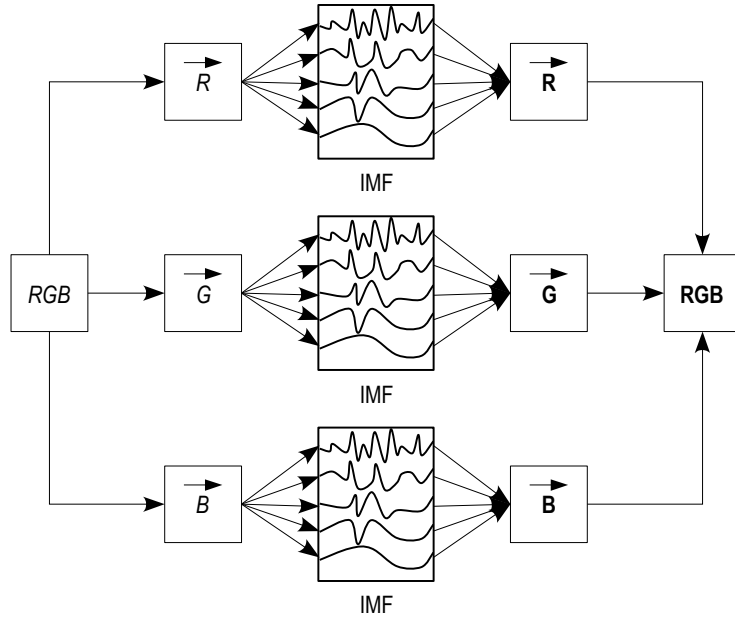


Fig. 7: Extension of one dimensional EMD to images via channel vectorization. The input image and channels are shown in italics and the reconstructed channels and image are shown in bold face.

images and the nature of the IMFs is shown in the second subsection. The manner in which mutual information assists in weight selection is described in the third subsection.

4.1 Data Collection

One of the major requirements for successful image fusion is that the images from the different modalities have to be aligned correctly. Image registration techniques are found extensively in the literature and the assumption made is that the images used in our experiments are registered. A hardware setup was constructed in such a manner that the optical axes of both the imaging sensors could be collimated. The setup allowed movement of the cameras on an axis, perpendicular to their optical axes. To acquire a registered image pair, an image is captured with the thermal camera first and by laterally displacing the visual camera, the principal axes are collimated and the visual image is obtained. The data acquisition setup is shown in Fig. 5. The images were visually examined to ensure a high degree of spatial correspondence between the multi-sensor images. A Raytheon PALM IR Pro® camera was used to acquire thermal images and a Canon PowerShot A-80 was used for visual images.

The registration performed was manual and refined by inspection. An example pair is shown in Fig. 6. In Fig. 6(a) and (b), the input images from different modalities are shown. In Fig. 6(c) and 6(d), sections of the thermal image are replaced with sections of the visual image. One is able to discern the shape of a pair of scissors and the degree of registration can be seen at the extremities of the replaced sections.

4.2 Empirical Mode Decomposition for Images

The EMD theory was originally proposed for one dimensional data. It has been extended for two-dimensional data in the literature towards image compression [12, 13] and texture analysis [14]. Though seemingly apt for image compression, these methods require sophisticated fitting schemes employing surface fitting by means of radial basis functions, bi-cubic splines etc. The

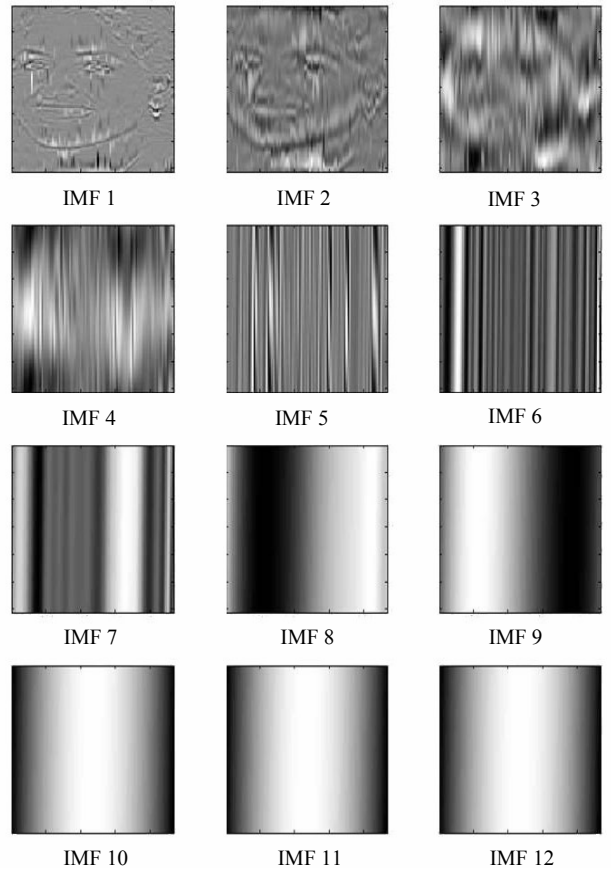


Fig. 8: An example of the decomposition of an image into edges, details and illumination.

mathematical soundness was unclear from the literature and it was supposed that vectorization of data was a more promising experiment. The input images were vectorized in lexicographical order. EMD was performed on each channel vector separately and converted to the reconstructed channel matrix. In the process of testing and understanding the decomposition process, the intrinsic mode functions were reconstructed and viewed as images. This led to an interesting and beneficial observation, upon which, one of the contributions of this paper is centered. It was observed that the lesser IMFs were those pertaining to the edge information of the image and the higher IMFs receded into details and illumination of the image. It was also observed that the EMD produced a smaller set of IMFs. A schematic of the flow of operations of the extension of empirical mode decomposition to images via vectorization is shown in Fig. 7. An example for one visual image channel is presented in Fig. 8 where the nature of the IMFs is exhibited.

4.3 Weighting Scheme Inspired by the Nature of IMF

Based on the nature of the IMFs, experiments were conducted to utilize this knowledge towards image fusion. The versatility of image fusion using weighted IMFs was high for highlighting features from both modalities. In principal component fusion, the number of principal components to be employed for reconstruction is either an arbitrary or a statistically-influenced choice. The nature of the application is also another influence in the number of principal components selected. But in the case of the EMD, the number of IMFs does not constitute an

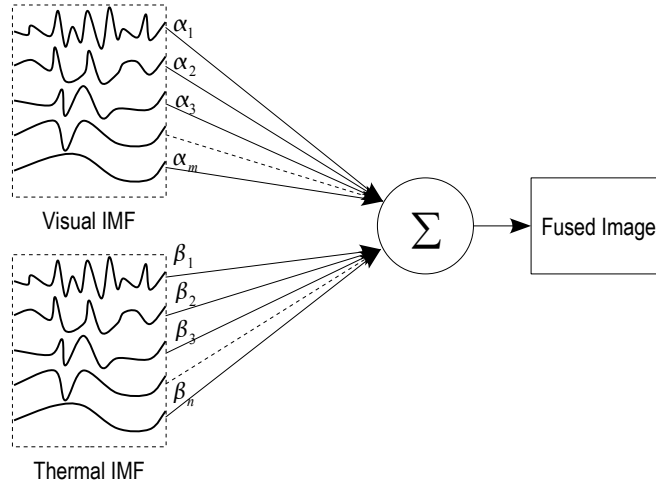


Fig. 9: A conceptual schematic of the new fusion scheme. This schematic shows the process of utilizing IMF and the weighting scheme towards image fusion.

issue. The decomposition is performed on all the channels of the input images. The RGB channels and the thermal image channel were decomposed into intrinsic mode images. At this pre-reconstruction stage, the IMFs are multiplied by a set of weights that decrease the mutual information between them. A flow diagram which elucidates our fusion scheme is shown in Fig. 9. The central idea behind our choice of weights in this manner is that image fusion is more meaningful if the components being fused are independent from each other. We conducted experiments with different weights and used mutual information and inspection to study the effects of the weighting scheme on the resultant image. It was observed that with the decrease in mutual information the features from both modalities were emphasized well and the resultant image was richer in thermal and visual features. The mutual information was calculated for different weighting schemes (shown in Table 1) and it is seen in Fig. 11 (a-c) that for decrease in mutual information the visual quality of the fused image improves.

Table 1: Mutual information calculated for various weighting schemes.

Case	Weighting Scheme 1	Weighting Scheme 2	Weighting Scheme 3	Weighting Scheme 4	Weighting Scheme 5
Mutual Information	0.6034	0.4184	0.4154	0.4116	0.3856

The mutual information (MI) is calculated using [15]

$$I(x, y) = \sum_{x \in V} \sum_{y \in T} p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right), \quad (15)$$

where $I(x, y)$ is the mutual information of the image pair, $p(x, y)$ is the joint probability mass function of the co-occurrence of pixel pairs (x, y) in an image pair, $p(x)$ is the probability mass function of the occurrence of pixel x in the visual image V , $p(y)$ is the probability mass function of the occurrence of pixel y in the thermal image T .

An empirical understanding of the nature of the IMFs is used to initialize the weights. The



Fig. 10: Examples of increase in visual and thermal content of fused image with decrease in mutual information (MI).

initial weights were chosen to emphasize or de-emphasize a particular feature. Based on the decrease in mutual information and increase in visual information, the weights are changed to arrive at better fused images. The compact form of the weighting scheme used in our method is given by

$$F(x, y) = \sum_{i=1}^3 \sum_{j=1}^k [\alpha_{ij} V_{ij} + \beta_{ij} T_{ij}], \quad (16)$$

where $F(x, y)$ 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, T_{ij} is the j^{th} thermal IMF corresponding to the i^{th} channel.

5. Results

We present the results of our fusion technique juxtaposed with the other fusion techniques. In Fig. 12(a), we see an example of an object partially hidden in the visual spectrum. The lower portion of a pair of scissors is completely hidden from view. In Fig. 12(b), the thermal signature of the whole pair of scissors is seen. The edges are not very distinct, especially the round holes on the metal surface in the back ground and there is no color information present in the thermal image. In Fig. 12(c), fusion by pixel-by-pixel averaging is shown. There is significant loss in the strength of edges. Averaging is akin to low pass filtering and the surface reflectance of the object obstructing the scissors is not seen in the fused image. In Fig. 12(d), wavelet based fusion of the input images is shown where the contrast in the fused image is high the fusion undermines one of the modalities. As a consequence, images acquired under extremely hot or over-illuminated conditions tend to look over-exposed. In Fig. 12(e), principal component fusion is similar to averaging but without any color information. In Fig. 12(f), an example of image fusion using EMD is demonstrated. Color information is well preserved in our fused image. Features from both the modalities are preserved well and it gives the visual effect of being able to see through the visual obstruction. The surface reflectance of the obstructing material is seen well in the images fused by our method. Edges are introduced from the input images thereby making object recognition a simpler task. For instance, information from both modalities can be easily observed, if the fused image is presented for human inspection.

In Fig. 13(a), we are not able to see a hand concealed in the pocket whereas in Fig. 13(b), the concealed hand is visible. However in Fig. 13(b), the edges are very feeble. The edges of the objects in the background are not prominent. In Fig. 13(c), fusion by pixel-by-pixel averaging is shown and the fused image looks more like the thermal image, due to the poor illumination of the

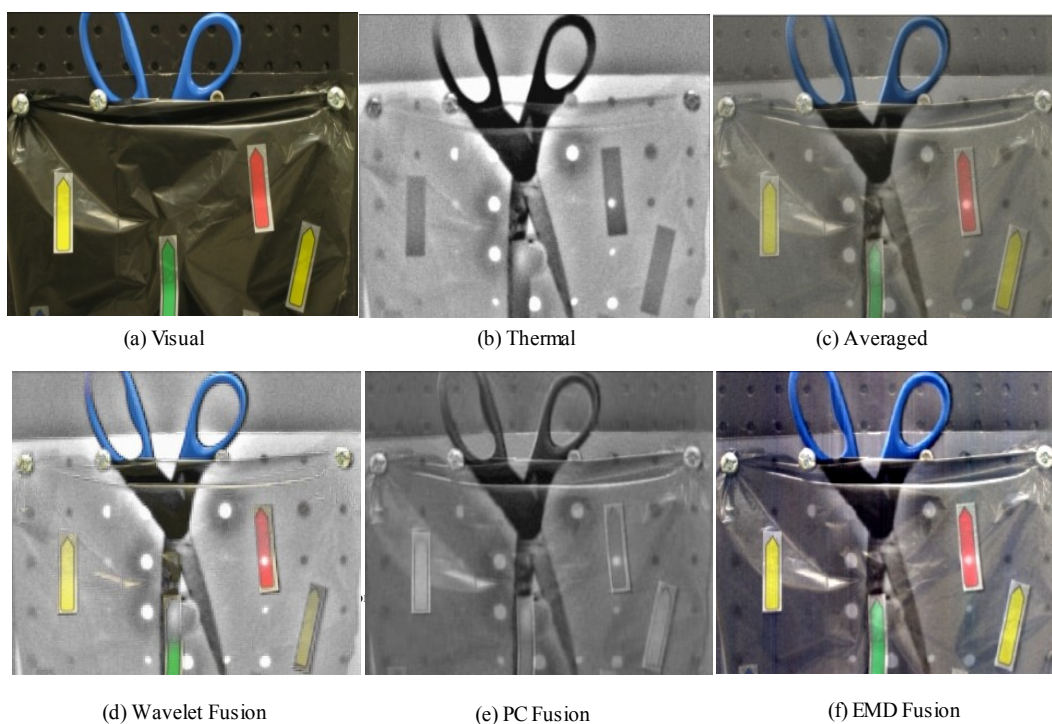


Fig. 11: Comparison of different fusion schemes, (a) the visual image, (b) the thermal image, (c) pixel-by-pixel averaging, (d) wavelet based fusion, (e) principal component fusion and, (f) EMD fusion.

visual image. This is also the case in PC fusion, as shown in Fig. 13(e). Due to the high intensity of the thermal signatures, the wavelet fusion is completely devoid of visual information as shown in Fig. 13(d). Image fusion via EMD increases the contrast in such a manner that an outline of hidden hand in the visual spectrum is seen in the fused image. The edges of the object in the background and the fabric texture become visible in the fused image as shown in Fig. 13(f). This experiment was conducted to establish feasibility in security applications. Hidden metal objects like fire-arms and blades can be detected without any invasive physical contact. Similarly in Fig. 14(f), we see that a hidden human face is more discernible than in the isolated input images, and also than in the other fused images. This can be used as an asset in face related imaging procedures where the heat signature of a human face can be used to increase feature reliability. The eyes of the individual are seen clearly in EMD fused image as opposed to the other fused images. Face detection and human tracking are some of the possible areas where the potential of the fused images can be harnessed.

The advantages of image fusion are not limited to detection of hidden objects. In Fig. 15(a), an image fused by pixel-by-pixel averaging is shown along with the edges detected from the same in Fig. 15(b). The edges from the thermal image are not very well emphasized in Fig. 15(b). In Fig. 15(c), an image fused by wavelet based fusion is shown. The emphasis is on the thermal edges in this case, as seen in Fig. 15(d). In Fig. 15(e), an image fused using principal components is shown. The edges that are detected are similar to the edges of the pixel-by-pixel averaging, as in Fig. 15(f). Color information is lost in the principal component based fusion. In Fig. 15(g), an image fused using EMD is shown. It is evident that the strengths of the edges of the fused images are stronger than any of the other fused images. Further, the edges from both modalities are introduced in the fused image. It is noteworthy that the eyes are seen clearly and, this is a

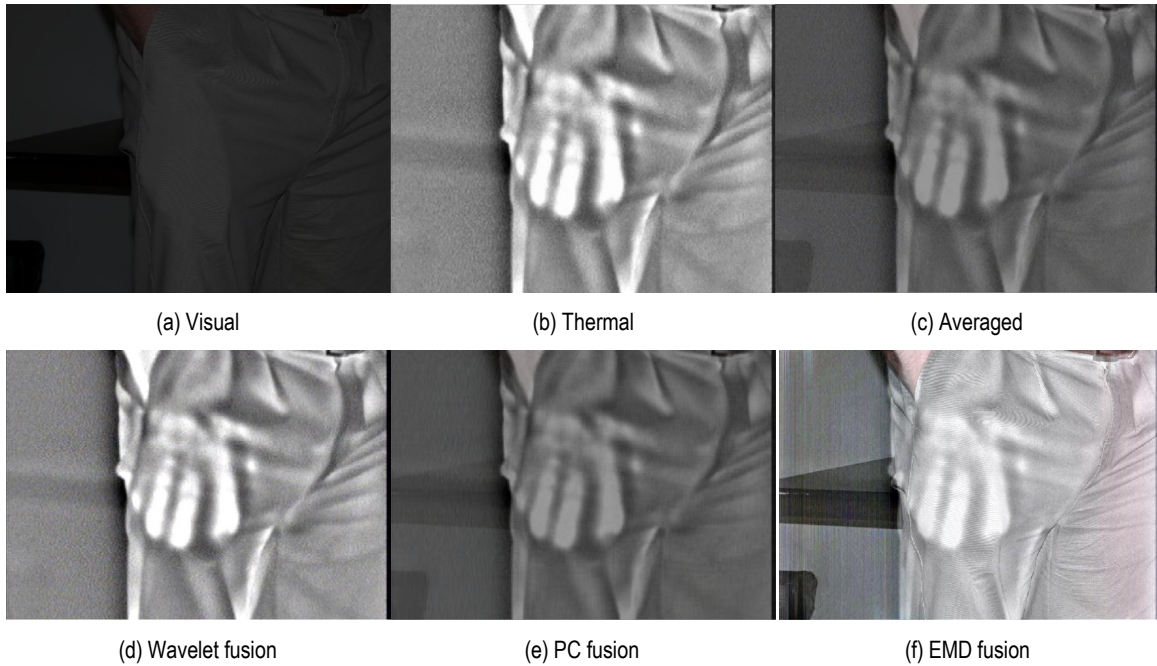


Fig. 12: Comparison of different fusion schemes, (a) the visual image, (b) the thermal image, (c) pixel-by-pixel averaging, (d) wavelet based fusion, (e) principal component fusion and, (f) EMD fusion.

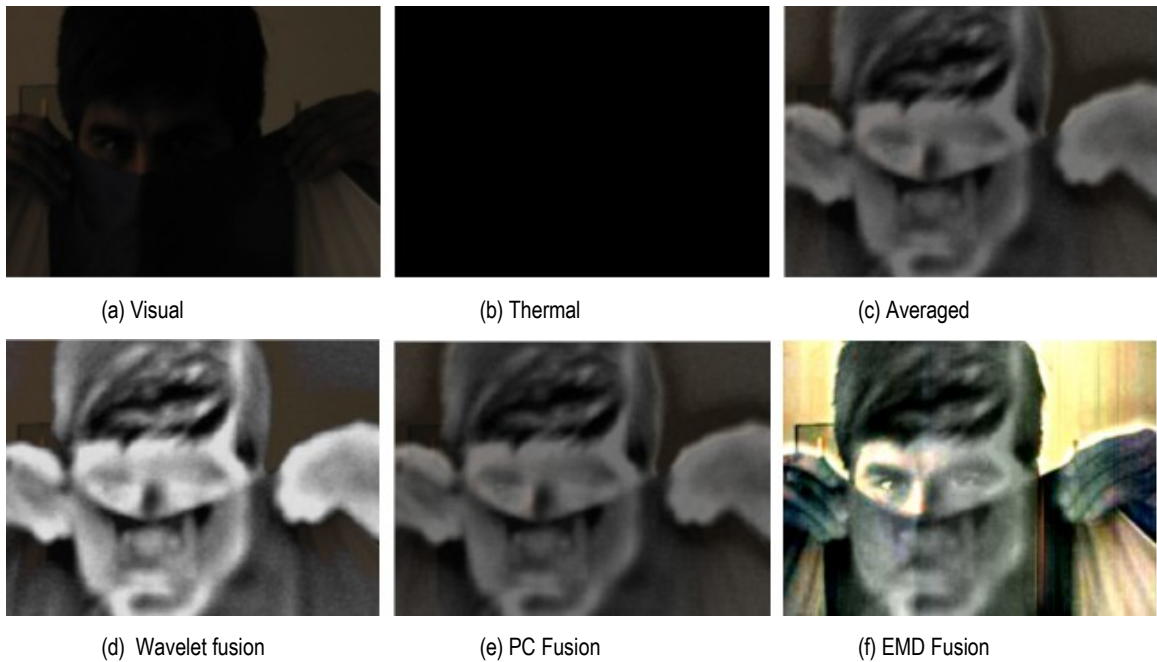


Fig. 13: Comparison of different fusion schemes, (a) the visual image, (b) the thermal image, (c) pixel-by-pixel averaging, (d) wavelet based fusion, (e) principal component fusion and, (f) EMD fusion.

significant factor from the view point of a face recognition application. This demonstrates a confluence of information from both the imaging modalities.

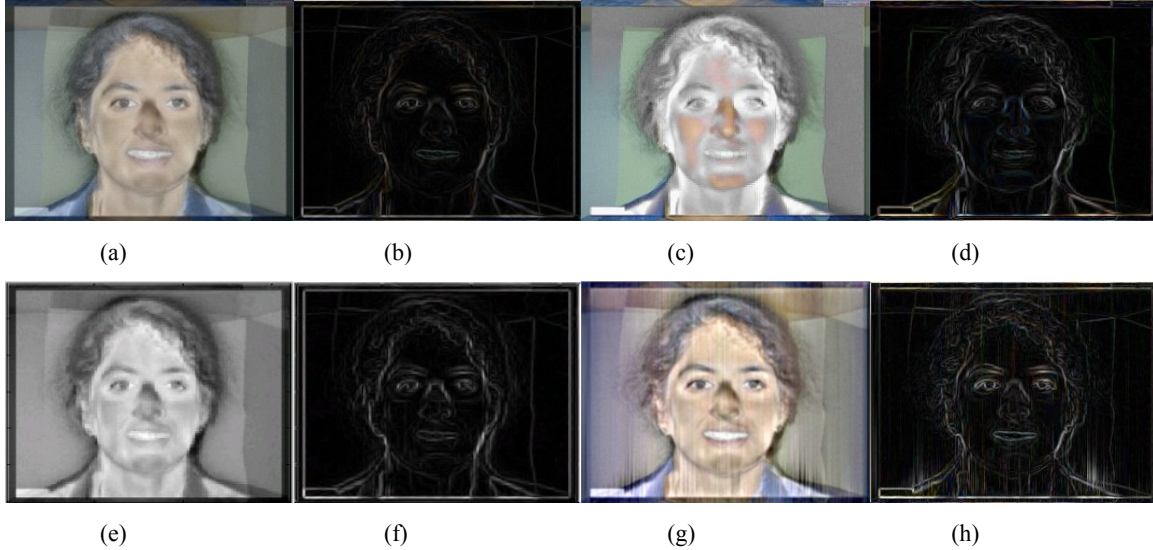


Fig. 14: An comparison of the strengths of the edge information obtained by different fusion schemes., (a) pixel-by-pixel averaged image , (b) edges corresponding to pixel-by-pixel averaged image , (c) wavelet based fused image, (d) edges corresponding to wavelet based fused image, (e) PC fused image, (f) edges corresponding to PC fused image, (g) EMD fused image, (h) edges corresponding to EMD fused image.

6. Conclusions And Future Work

A novel and efficient technique for image fusion using empirical mode decomposition has been proposed. Our fusion technique preserves information from both the input images. As input, we use registered thermal and visual images. Empirical mode decomposition is used to obtain the decomposed IMFs of the various channels of the visual and thermal image. Fusion is performed at the IMF level where a weighting scheme is used to emphasize features or, to discourage distracting features from one, or both, of the modalities by minimizing the mutual information between the IMFs. The output is a fused image containing enhanced visual and thermal information. The automatic minimization of mutual information is the next step in advancing this technology. Publications in the area of biometric authentication [16-18] indicate that the fusion of thermal and visual data increases the reliability of face data and hence our method is an interesting avenue for exploration for face recognition applications. The effects on our fusion scheme in increasing face recognition rates will be studied as part of our future work.

Acknowledgments

The authors wish to thank Dr. David Page, Dr. Besma Abidi and Bradley Grinstead for their valuable suggestions and personal communications. This support extended by the Office of Naval Research under Grant N00014-03-1-0022, Feature-Based Face Recognition with Thermal Signatures towards this effort is gratefully acknowledged.

References

- [1] L. P. Yaroslavsky, B. Fishbain, A. Shteinman, and S. Gepshtein, "Processing and fusion of thermal and video sequences for terrestrial long range observation systems," in *Proc. 7th Annu. Int. Conf. Information Fusion*, Sweden, 2004, pp. 848-855.
- [2] D. A. Fay, A. M. Waxman, M. Aguilar, D. B. Ireland, J. P. Racamato, W. D. Ross, W. W. Streilein, and M. I. Braun, "Fusion of multi-sensor imagery for night vision: Color visualization, target learning and search," in *Proc. 3rd Int. Conf. Information Fusion*, Paris, 2000, pp. 215-219.
- [3] D. A. Socolinsky and L. B. Wolff, "Multispectral image visualization through first-order fusion," *IEEE Trans. Image Processing*, vol. 11, no. 8, pp. 923-931, Aug. 2002.

- [4] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, no. 3, pp. 235-245, May 1995.
- [5] E. F. Canga, "Image fusion," Project report, Dept. Electronic and Electrical Eng., Univ. of Bath, Bath, UK, 2002.
- [6] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arniol, "Multiresolution-based image fusion with additive wavelet decomposition," *IEEE Trans. Geosciences and Remote Sensing*, vol. 37, no. 3, pp.1205-1211, May 1999.
- [7] O. Rockinger and T. Fletcher, "Pixel-level image fusion: The case of image sequences," *Proc. SPIE*, vol. 3374, pp. 378-388, 1998.
- [8] K. Vani, "Fusion of ASTER image data for enhanced mapping of landcover features," M.S. thesis, Inst. of Remote Sensing, Anna Univ., 1999.
- [9] S. Singh, A. V. Gyaourova, G. Bebis, and I. Pavlidis, "Face recognition using fusion," in *Proc. SPIE Defense and Security Symposium*, Orlando, FL, Apr. 2004, pp. 12-16.
- [10] 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, pp. 903-995, 1998.
- [11] A. Linderhed, "Variable sampling of the empirical mode decomposition of two-dimensional signals," *Int. Journal of Wavelets, Multiresolution and Information Processing*, vol. 3, no. 3, pp. 435-452, Sept. 2005.
- [12] A. Linderhed, "Image compression based on empirical mode decomposition," in *Proc. Intl. Conf. Image and Graphics*, Hong Kong, 2004, pp. 430-433.
- [13] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," *IEEE Trans. Electron Devices*, vol. 11, pp. 34–39, Jan. 1959.
- [14] Z. Yang, D. Qi, and L. Yang, "Signal period analysis based on Hilbert-Huang transform and its application to texture analysis," in *Proc. Intl. Conf. Image and Graphic*, Hong Kong, 2004.
- [15] S. Haykin, *Neural Networks: a Comprehensive Foundation*. Upper Saddle River, NJ: Prentice Hall, 1999.
- [16] S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi, "Recent advances in visual and infrared face recognition - A review," *Computer Vision and Image Understanding*, vol. 97, no. 1, pp. 103-135, Jan. 2005.
- [17] B. Abidi, S. Huq, and M. Abidi, "Fusion of visual, thermal, and range as a solution to illumination and pose restrictions in face recognition," *IEEE Carnahan Conf. Security Technology*, Albuquerque, NM, Oct. 2004, pp. 325-330.
- [18] J. Heo, S. Kong, B. Abidi, and M. Abidi, "Fusion of visual and thermal signatures with eyeglass removal for robust face recognition," *IEEE Workshop Object Tracking and Classification Beyond the Visible Spectrum* in conjunction with *CVPR 2004*, Washington, D.C., July 2004, pp. 94-99.