International Journal of Electronics Communication and Computer Engineering
Volume 4, Issue (6) NCRTCST -2013, ISSN 2249–071X

National Conference on Recent Trends in Computer Science and Technology (NCRTCST)-2013

A Survey on Different Fusion Techniques of Visual and Thermal Images for Human Face Recognition

Tumpa Dey
Department of IT,
Dasaratha Deb Memorial College, Tripura, India
Email: tumpadey91@gmail.com

Abstract – **In this paper we do a survey on different fusion techniques of visual and thermal images for human face recognition. Image fusion constructs a single image by combining information from a set of source images together using different techniques. It is quite simpler to extract and locate facial features in visual images. Another advantage of using visual image is that it works well under controlled illumination condition. Infrared (IR) imaging is especially useful to determine the location and shape of objects in darkness and provides information about the degree of heating of individual segment of complex surfaces. Now a days, researchers have explored the use of fusion of thermal infrared and visual face images for person identification to combine the good aspects of visual and thermal images and to overcome the negative aspects of individual thermal and visual images. The aim of fusion is to produce a fused result or image that offers the more detailed and reliable information than the source images. In our paper we discussed the different fusion techniques which is used to combine visual and thermal images for human face recognition.**

Keywords - **Image Fusion, Visual Image, Thermal Image, Infrared Imaging, Face Recognition.**

I. Introduction

In our daily life, face is the most important part for conveying identity and emotions, as it is the prime centre of interest. Right from birth, all our social interactions depend on our face and its various emotions. Face recognition means the ability to distinguish and identify people by their facial features. It is one of the most dynamic and extensively used techniques because of its reliability in the process of recognizing and verifying a person's identity. There are various images of different modalities (e.g. visual, thermal, sketch etc.) which are used for performing the face recognition task. Optical or visual image is an intellectual image that is similar to a visual insight or can be referred to as a percept that arises from the eyes or as an image in the visual system i.e. the image which we can see with our normal eyes. Most of the researchers have focused on visible spectrum imaging for research in biometric field. This has improved the development of better recognition algorithms in case of working with visual images. It is quite simpler to extract and locate facial features in visual images. Another advantage of using visual image is that it works well under controlled illumination condition. But, there are some disadvantages of visual images and thermal images have the capability to overcome some of these problems. The disadvantage of using visual image is that it results in poor performance with illumination variations (such as in indoor and outdoor lighting conditions); again it is not efficient to distinguish different facial expressions, it is difficult to segment out faces from cluttered scene, visual images are useless in very low lighting, and unable to detect disguise [1].To solve these problems of visual images, researchers have started investigating the use of thermal infrared images for face recognition purposes. However, many of these research efforts in thermal face recognition use the thermal infrared band only as a way to see in the dark or trim down the detrimental effect of light variability [2] [3]. Infrared (IR) imaging is especially useful to determine the location and shape of objects in darkness and provides information about the degree of heating of individual segment of complex surfaces. It also helps to know about internal structure of bodies that are opaque in visible light. Every heated body emits thermal radiation whose intensity and spectrum depend on the body's properties and temperature. Infrared imaging is mostly used in medicine and technical diagnosis, navigation, geological exploration, meteorology, flaw detection, research on thermal processes, and military affairs. Image fusion constructs a single image by combining information from a set of source images together using different techniques. Pixel-level Image fusion implies fusion at the lowest processing level which refers to the integration of measured physical parameters. The image after the pixel-level image fusion contains much richer and more accurate information content, which is advantageous to the analysis and processing of image signal. It makes human observation easier and more suitable for computers in detection processing. Advantage of image fusion at pixel level includes minimum loss of information, but it has the larger amount of information to be processed, and so, slower processing speed, and a higher demand for equipment.

II. Previous Work on Fusion Methology

In search of an enhanced and robust face recognition system, researchers started to use fusion technique over the visual and thermal images. G. Bebis et al. [4] focused on the study of sensitivity of thermal IR imagery to facial occlusions caused by eyeglasses. Specifically, their experimental results illustrated that recognition performance in the IR spectrum degrades seriously when eyeglasses are present in the probe image but not in the gallery image and vice versa.V. Jyothi et al. [5] discuss that they have compared the regular image fusion techniques with the genetic algorithm (GA) based techniques and observe that from their experiment GA based techniques are having much better results compared to conventional techniques. T. Zaveri et al.[6] applied discrete wavelet transform using high boost filtering on large number of dataset of each category and simulation results are found with superior visual quality compared to other earlier reported pixel and region based image fusion method. Here, T. Zaveri et al. investigated on two different fusion rules are applied on broad range of images. In their proposed method MMS fusion rule is introduced to select desired regions from multimodality or multisensor images and SF based rule is used for single sensor based multifocus images. Proposed algorithm is compared with standard reference based and nonreference based image fusion parameters and from simulation results and this proposed algorithm preserves more details in fused image M. Hanif et al. [7] said data fusion of thermal and visual images is a solution to overcome the drawbacks present in individual thermal and visual images. They presented an optimized and efficient wavelet domain data fusion of thermal and visual face images to achieve better face recognition system. The proposed fusion technique has proven to be effective even for variable expression and light condition.Z. Shu-Long [8] describes data fusion and decision fusion of registered visual and thermal infrared (IR) images for robust face recognition. In data fusion, eyeglasses are detected from thermal images and replaced with an eye template. They implemented three techniques: Data fusion of visual and thermal images (Df), Decision fusion with highest matching score (Fh), and Decision fusion with average matching score (Fa). Their comparative study show that fusion-based techniques outperformed individual visual and thermal face recognizers under illumination variations and facial expressions. D. R. Kisku et al. [9] presented a face recognition technique based on the extraction and matching of SIFT features related to independent face areas. Both a global and local (as recognition from parts) matching strategy is proposed. The local strategy is based on matching individual salient facial SIFT features as connected to facial landmarks such as the eyes and the mouth. As for the global matching strategy, all SIFT features are combined together to form a single feature. S. Singh et al. [10] proposed fusing IR with visible images, exploiting the relatively lower sensitivity of visible imagery to occlusions caused by eyeglasses. Two different fusion schemes have been investigated in this study: (1) image based fusion performed in the wavelet domain and, (2) feature-based fusion performed in the eigenspace domain. In both cases, they employed Genetic Algorithms (GAs) to find an optimum strategy to perform the fusion. R. Singh et al. [11] presented an integrated image fusion and match score fusion of multispectral face images. The fusion of visible and long wave infrared face images is performed using 2ν Granular SVM which uses multiple SVMs to learn both the local and global properties of the multispectral face images at different granularity levels and resolution. The 2ν GSVM performs accurate classification which is subsequently used to dynamically compute the weights of visible and infrared images for generating a fused face image. 2D log polar Gabor transform and local binary pattern feature extraction algorithms are applied to the fused face image to extract global and local facial features respectively.

III. Advances in Fusion and New Technique

During the past two decades, several fusion techniques have been proposed. Among the hundreds of variations of image fusion techniques, the widely used methods include, intensity-hue-saturation (IHS), high pass filtering, High-pass modulation, different arithmetic combination(e.g. Brovey transform), multi-resolution analysis-based methods (e.g. pyramid algorithm, wavelet transform), Generic Multiresolution Fusion Scheme, Ehlers transform, etc. The chapter will provide a general introduction of these selected methods.

A. Intencity hue-saturation

The IHS technique is a standard procedure in image fusion. Originally, it was based on the RGB true colorspace [26]. RGB color space, a pixel is identified by the intensity of red, green, and blue primary colors. It offers the advantage that the separate channels outline certain color properties, namely intensity (I), hue (H), and saturation (S). Geometrically the IHS colour system can be represented as a cylinder (Figure 3 (a)) or a sphere (Figure 3 (b)). In cylindrical coordinates, the colour space is defined by two vectors and one angle. Intensity, which represents the brightness of a colour defines the axis, saturation, which represents the purity of a colour, defines the radius while hue, the average wavelength of the colour, defines the circumferential angle of the cylinder. In spherical coordinates, the colour space is defined by one vector and two angles. Intensity defines the vertical axis, saturation the co-latitude and hue the circumferential angle of the sphere. In the cylindrical system the definition of saturation implies that a constant distance between any RGB vector and intensity axis is maintained with varying intensity, while in the spherical system, the absolute distance with the intensity axis is proportional to intensity. Because saturation is defined as an angle, the ratio between this distance and intensity defines its magnitude [12].

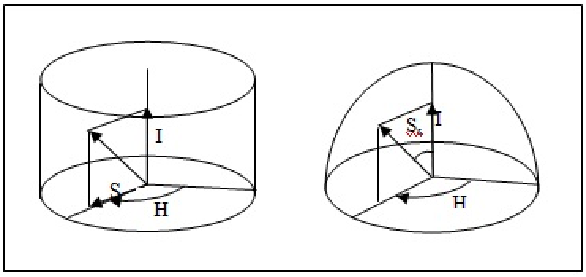

Fig.1. Cordinate system for defining the IHS transform; (a) cylindrical (b) spherical.

B. Brovey Trasform

The Brovey Transform is based on the chromaticity transform [13]. Before know the chromaticity transform we have to know the chromaticity space. chromaticity space, is two dimensions of the normalized RGB space. Chromaticity space, has no intensity information. Normalized RGB space, a color is represented by the proportion of red, green, and blue in the color, rather than by the intensity of each.A color (R,G,B) where R, G, B are intensity of Red, Green and Blue respectively. This can be converted to color (r,g) where r, g imply the proportion of red and green in the original color is called chromaticity transform. The Brovey transform is based on the mathematical combination of the multispectral images and high resolution Panchromatic image. Panchromatic images is an image collected in the broad visual wavelength range but rendered in black and white and multispectral images are images optically acquired in more than one spectral or wavelength interval. Each individual image is usually of the same physical area and scale but of a different spectral band. Each Each multispectral image is normalized based on the other spectral bands and multiplied by the pan image to add the spatial information to the output image. Its purpose is to normalize the three multispectral bands used for RGB display and to multiply the result by any other desired data to add the intensity or brightness component to the image.

C. High-pass filtering

The principle of HPF is to add the high-frequency information from the High Resolution Panachromatic Images to the Low Resolution Multispectrum Images to get the High Resolution Multispectrum Images. The high spatial resolution image is filtered with a small high pass-filter resulting in the high frequency part of the data which is related to the spatial information. This is pixel wise added to the low resolution bands. It can be of advantage to use subtraction of pre-processed data (e.g., HPF filtered imagery) from the original data in order to enhance lines and edges. The principle of HPF is to add the high-frequency information from the High Resolution Panachromatic Image to the Low Resolution Multi-spectrum Images to get the High Resolution Multispectrum Images. The high-frequency information is computed by filtering the High Resolution Panachromatic Images with a high-pass filter or taking the original High Resolution Panachromatic Images and subtracting the Low Resolution Panachromatic Images, which is the low-pass filtered High Resolution Panchromatic Images. This method preserves a high percentage of the spectral characteristics, since the spatial information is associated with the high-frequency information of the High Resolution Mulispectrum Images, which is from the High Resolution Panachromatic Images, and the spectral information is associated with the low-frequency information of the HRMIs, which is from the Low Resolution Multispectrum Images[14][15].

D. Enhler Transform

The basic idea behind this method is to modify the input Pan image to look more like the intensity component. In the first step in order to manipulating, three low resolution multispectral RGB bands are selected and transformed into the IHS domain. Then, the intensity component and the panchromatic image are transformed into the spectral domain via a two-dimensional Fast Fourier Transform (FFT). Low pass (LP) and high pass (HP) filter were directly performed in the frequency domain on the intensity component and the high resolution panchromatic image respectively. The idea is to replace the high frequency part of the intensity component with that from the Pan image. To return both components back into the spatial domain an inverse FFT transform was used. Then the high pass filtered panchromatic band and low pass filtered intensity are added and matched to the original intensity histogram. Finally, an inverse synthesis is the normalized average error of each band[13][16]

E. Imgae pyramid

Image pyramids have been initially described for multiresolution image analysis and as a model for the binocular fusion in human vision. A generic image pyramid is a sequence of images where each image is constructed by low pass filtering and sub sampling from its predecessor. Due to sampling, the image size is halved in both spatial directions at each level of the decomposition process, thus leading to a multi resolution signal representation. The difference between the input image and the filtered image is necessary to allow an exact reconstruction from the pyramidal representation. The image pyramid approach thus leads to a signal representation with two pyramids: The smoothing pyramid containing the averaged pixel values, and the difference pyramid containing the pixel differences, i.e. the edges. So the difference pyramid can be viewed as a multiresolution edge representation of the input image. The actual fusion process can be described by a generic multiresolution fusion scheme which is applicable both to image pyramids and the wavelet approach. There are several modifications of this generic pyramid construction method described above. Some authors propose the computation of nonlinear pyramids, such as the ratio and contrast pyramid, where the multistage edge representation is computed by a pixel-by-pixel division of neighbouring resolutions [17].

F. Pixel based fusion technique

In the process of image fusion where pixel data of 70% of visual image and 30% of thermal image of same class or same image is brought together into a common operating image or now commonly referred to as a Common Relevant Operating Picture (CROP) [45]. This implies an additional degree of filtering and intelligence applied to the pixel streams to present pertinent information to the user. So image pixel fusion has the capacity to enable seamless working in a heterogeneous work environment with more complex data. They assumed that each face is represented by a pair of images, one in the IR spectrum and one in the visible spectrum. Both images were combined prior to fusion to ensure similar ranges of values [18].

G. Wavelet transform

A signal analysis method similar to image pyramids is the discrete wavelet transform. The main difference is that while image pyramids lead to an over complete set of transform coefficients, the wavelet transform results in a non redundant image representation. The discrete 2-dimentional wavelet transform is computed by the recursive application of low pass and high pass filters in each direction of the input image (i.e. rows and columns) followed by sub sampling. Wavelet transform is a tool that provides a variety of channels representing the image feature by different frequency sub-bands at multi-scale. It is a famous technique in analyzing signals. When decomposition is performed, the approximation and detail component can be separated [18]

IV. Application Area of Fusion Technique

(a) Concealed weapon detection:

Concealed weapon detection is an increasingly important topic in the general area of law enforcement. Since no single sensor technology can provide acceptable performance in CWD applications, image fusion has been identified as a successful technology to achieve improved CWD procedures [20].

(b) Medical imaging:

Several diagnostic cases require integration of complementary information for better analysis. Fusion of multimodal medical images can provide a single composite image that is dependable for improved analysis and diagnosis[20].

(c) Flood monitoring:

In the field of the management of natural hazards and flood monitoring using multisensor VIR (visible–infrared)/SAR (synthetic aperture radar) images plays an important role. For the representation of the pre-flood situation the optical data provides a good basis. The VIR image represents the land use and the water bodies before flooding [19]. Then, SAR data acquisition at the time of the flood can be used to identify flood extent and damage.

(d) Topographic mapping and map updating:

Image fusion used as tool for topographic mapping and map updating has its importance in the provision of up-to-date information. Areas that are not covered by one sensor might be contained in another. In the field of topographic mapping or map updating often combinations of VIR and SAR are used [19].

(e) Land use, agriculture and forestry:

The defection of single data source can be solved with fusion between TM and SAR images. It provides a better solution for land surface monitoring in the mining areas. Multi-temporal SAR is a valuable data source in countries with frequent cloud cover and successfully used in crop monitoring. Optical and microwave image fusion is also well known for the purpose of identifying and mapping forest cover and other types. The combined optical and microwave data provide a unique combination that allows more accurate identification, as compared to the results obtained with the individual sensors [19].

(f) Ice and snow monitoring:

The fusion of data in the field of ice monitoring provides results with higher reliability and more detail. Regarding the use of SAR from different orbits for snow monitoring the amount of distorted areas due to layover, shadow and foreshortening can be reduced significantly [19].

References

[1] S. G. Kong, J. Heo, B. R. Abidi, J. Paik, and M. A. Abidi, "Recent advances in visual and infrared face recognition—a review", Computer Vision and Image Understanding 97 (2005) 103–135.

[2] D. A. Socolinsky, L. B. Wolff, J. D. Neuheisel, and C. K. Eveland, "Illumination invariant face recognition using thermal infrared imagery", Proc. of the IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR 2001), vol. 1, pp. 527-534, 2001.

[3] A. Selinger, and D. A. Socolinsky, "Face Recognition in the Dark ",Computer Vision and Pattern Recognition Workshop, vol.8, pp. 129-134, June 2004.

[4] K. Edwards, and P. A. Davis, "The use of Intensity-Hue-Saturation transformation for producing color shaded-relief images," Photogramm. Eng. Remote Sens., vol. 60, no. 11, pp. 1369–1374, 1994.

[5] E. M. Schetselaar, "Fusion by the IHS transform: Should we use cylindrical or spherical coordinates?," Int. J. Remote Sens., vol. 19, no. 4, pp.759–765, 1998.

[6] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," Int. J. Remote Sens., vol. 21, no. 18, pp. 3461–3472, 2000.

[7] T. M. Tu, S. C. Su, H. C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," Inf. Fusion, vol. 2, no. 3, pp. 177–186, 2001.

[8] A. R. Gillespie, A. B. Kahle, and R. E. Walker, "Color enhancement of highly correlated images—II. Channel ratio and 'chromaticity' transformation techniques," Remote Sens. Environ., vol. 22, pp. 343–365,1987.

[9] D. R. Kisku, M. Tistarelli, J. K. Sing, and P. Gupta, "Face Recognition by Fusion of Local and Global Matching Scores using DS Theory: An Evaluation with Uni-classifier and Multi-

classifier Paradigm," IEEE Computer Vision and Pattern Recognition Workshop on Biometrics, Feb. 2010.

[10] S. Singh, A. Gyaourova, G. Bebis, and I. Pavlidis, "Infrared and Visible Image Fusion for Face Recognition," "Biometric Technology for Human Identification," Edited by A. K. Jain; N. K. Ratha, Proc. of SPIE, Vol. 5404, pp. 585-596, 2004.

[11] R. Singh, M. Vatsa, and A. Noore, "Integrated Multilevel Image Fusion and Match Score Fusion of Visible and Infrared Face Images for Robust Face Recognition," In Pattern Recognition Journal, Elsevier Science Inc., New York, USA Vol. 41 Issue 3, March 2008.

[12] Y. Z. Goh, A. B. J. Teoh, and M. K. O. Gog, "Wavelet Based Illumination Invariant Preprocessing in Face Recognition", Proc. of the 2008 Congress on Image and Signal Processing, IEEE Computer Society, Vol. 3, pp. 421 – 425.

[13] M. Turk, and A. Pentland, "Face recognition using eigenfaces," Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '91). pp. 586–591, 1991.

[14] T. Paul. "Affine coherent statesand the radial Schrodinger equation, Radial harmonic oscillator and hydrogen atom".

[15] T. Bourlai, N. Kalka, A. Ross, B. Cukic, and L. Hornak, "Cross-spectral Face Verification in the Short Wave Infrared (SWIR) Band," Proc. of Int. Conf. on Pattern Recognition (ICPR), Istanbul, Turkey, pp. 1343–1347, August 2010.

[16] J. Choi, S. Hu, S. S. Young, and L. S. Davis, "Thermal to Visible Face Recognition," (Accepted) Proc. of Biometric Technology for Human Identification IX, to be published by SPIE Defense, Security and Sensing, April 23, 2012.

[17] P. Buddharaju, I. Pavlidis, and I. Kakadiaris, "Face Recognition in the Thermal Infrared spectrum," Proc. of the 2004 IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition workshops (CVPRW'04), 2004.

[18] D. A. Socolinsky, L. B. Wolff, J. D. Neuheisel, and C. K. Eveland, "Illumination Invariant Face Recognition Using Thermal Imagery," *Proc. of the IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR'01)*, Hawaii, 2001.

[19] C. Pohl, and J. L. V. Genderen, "Multisensor image fusion in remote sensing: concepts, methods and applications," Review article, int. j. remote sensing, vol. 19, no. 5, 823- 854, 1998.

[20] Z. Xue, R. S. Blum, and Y. Li, "Fusion of Visual and IR Images for Concealed Weapon Detection", ISIF, 2002