Optimized Visual and Thermal Image Fusion for Efficient Face Recognition

Mohammad Hanif
COMSATS Institute of Information Technology,
Abbottabad Campus, Pakistan
hanif_bce@hotmail.com

Usman Ali
Muhammad Ali Jinnah University,
Islamabad Campus, Pakistan
usman_gikian_abbott@hotmail.com

Abstract – *Data fusion of thermal and visual images is a solution to overcome the drawbacks present in individual thermal and visual images. Data fusion using different approach is discussed and results are presented in this paper. Traditional fusion approaches don't produce useful results for face recognition. An optimized approach for face data fusion is developed which works for face data fusion equally well as for non-face images. This paper presents the implementation of Human face recognition system using proposed optimized data fusion of visual and thermal images. Gabor filtering technique, which extracts facial features, is used as a face recognition technique to test the effectiveness of the fusion techniques. It has been found that by using the proposed fusion technique Gabor filter can recognize face even with variable expressions and light intensities.*

Keywords: Gabor Filter, Data fusion, Face recognition.

I. INTRODUCTION

Human face recognition is a challenging task and its domain of applications is very vast, covering different areas like security systems, defense applications, and intelligent machines. It involves different image processing issues like face detection, recognition and feature extraction.

Face recognition using visual images is gaining acceptance as a superior biometric [1]. Images taken from visual band are formed due to reflectance. Therefore they depend on the external light source, which sometimes might be absent e.g. night time or when there are heavy clouds. Imagery is also difficult because it depend on the intensity of light and angle of incident of light.

Commonly used face recognition techniques are discussed in [2, 3, 4, 5, 6]. Recently, face recognition on thermal infrared spectrum has gained much popularity because thermal image is formed due to emission not reflection [7, 8]. They do not depend on the external light source and intensity of light and are also less dependent on angle of incident of light [9, 10]. Different methodologies of face recognition are performed on visual/thermal images and are compared in [11, 12, 13].

Image fusion can be divided into three categories namely, pixel level, feather level and decision level. Both feather

level and decision level image fusion may results in inaccurate and incomplete transfer of information [14, 15].

PCA based image fusion for a night vision application is discussed in [16]. A hybrid approach, advance wavelet transform (aDWT) method that incorporates principle component analysis (PCA) and morphological processing into a regular DWT, is proposed in [17].

An optimized image fusion approach comprising Structural Similarity, Principle Component Analysis and Discrete Wavelet is discussed in [18].

In [19] many multisensor data fusion architectures are presented to create a more informative fused image. Fusion of visual and thermal signatures for robust face recognition explained in [20] has outperformed the individual visual and thermal face recognizers.

For the detailed review on current advances in visual and thermal face recognition refer to [21]. Simple data fusion in spatial domain is discussed in [22], where faceIt recognition is used for testing the fusion of Equinox face database images. Both data fusion and decision fusion is employed in [23] to improve the accuracy of the face recognition system.

In this paper we present 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. Paper is organized as follow: In section 2 data fusion techniques are discussed. Sections 3 discuss, very briefly, the Gabor filter based face recognition technique and section 4 discuss results of image fusion for various spatial and wavelet domain data fusion techniques. Finally, the conclusions and future work is provided.

II. DATA FUSION TECHNIQUES

For accurate and effective face recognition we require more informative image. Image by one source (i.e. thermal) may lack some information which might be available in images by other source (i.e. visual). So if can combine the feathers of both the images i-e visual and thermal then an improved image, suitable for efficient and accurate face recognition, can be obtained. Spatial and DWT domain

Fusion techniques are discussed in the following subsections.

A. Spatial Domain Data Fusion

Data fusion technique provides an advantage of combining the information of both the sources (visual and thermal) to produce more informative image for efficient recognition. The Equinox facial database [24], the most extensive infrared facial database that is publicly available at the moment, was used for testing. The Equinox database has a good mix of subject images with accessories (e.g.,

Figure 1a

Figure 2b

Figure 3c

B. Wavelet Domain Data Fusion

1. Absolute maximum selection in DWT:

In this technique first decompose both the images (A & B) by discrete wavelet transformation (DWT) up to 'n' level, resulting in the 3n sub bands plus one approximate band. Denote the approximation wavelet coefficients for image A and B as A_A and B_A and detail coefficients as $A_{n_{vd}}$, $A_{n_{hd}}$, $A_{n_{dd}}$ and $B_{n_{vd}}$, $B_{n_{hd}}$, $B_{n_{dd}}$, for vertical, horizontal and diagonal details respectively. Then, at each level select the absolute maximum at every pixel from both the images. Perform this operation in all sub-bands to preserve more dominant feather at each level. Enhanced image is achieved by taking inverse DWT of the resultant sub-bands [19].

2. Window base absolute maximum selection:

As pixel by pixel absolute maximum selection does not work for face image so a window base absolute maximum selection scheme is proposed in [14]. In which, initially images are decomposed by DWT up to 'n' levels and absolute maximum value is selected with in a mask of size 3x3 at each sub-bands from both images A and B. Binary decision map is constructed equal to image size, where '0' represent that absolute maximum value at that pixel from image A is greater than absolute maximum value from image B and '1' represent that absolute maximum value from B is greater then from A. This binary map is subjected to consistency verification. In the transform domain, specially, if the center pixel value comes from image A, while the majority of the surrounding pixel values come from image B, the center pixel value is switched to that of image B. A majority filter [19] is applied to the binary decision map; the map is then negated, followed by the

glasses) as well as expressions of happiness, anger, and surprise, which account for pose variation. The visual and thermal image is combined by using equation as proposed in [15]:

$$F(x, y) = F_w T(x, y) + (F_w - 1)V(x, y)$$

Where $F(x,y)$ is the fused image $T(x,y)$ and $V(x,y)$ are the thermal and visual image respectively and F_w is the weight of fusion and its value is from 0 to 1. Fused image $F(x,y)$ is passed to Face recognition system as input. Fused image of images shown in figure 1a and 1b is shown in figure 1c.

application of a majority filter, and is negated again. A fused image is finally obtained based on the new binary decision map. This selection scheme ensures that most of the dominant feathers are incorporated into the fused image [19].

3. Simple averaging in DWT :

As both the pixel level absolute maximum selection and the window base absolute maximum selection approaches are not effective for the face images. A simple averaging in the wavelet transform domain is applied at every pixel for all sub-bands using the following equation

$$\text{Fused image} = 0.5 * \text{thermal} + 0.5 * \text{visual}.$$

By this blocky effects apparent in the previous techniques are removed significantly, although it causes a little blurry effect.

4. Optimized image fusion in DWT :

To curve resulting blurry effect, utilizes weighted average approach. As the thermal images are brighter then visual images, so assign weights accordingly as given below

$$\text{Fused image} = 0.3 * \text{thermal} + 0.7 * \text{visual}.$$

Next section would explain the Gabor filter face recognition technique, which would decide the effectiveness of the above proposed optimized fusion for accurate face recognition.

III. FACE RECOGNITION-GABOR FILTER

A. Feature point calculation

Physiological studies found simple cells in human visual cortex that are selectively tuned to orientation as well as to

special frequency, it was suggested that the response of a simple cell could be approximated by 2-D Gabor filters [25]. One of the most successful recognition methods is based on graph matching of coefficients, which have disadvantages due to there matching complexity.

Escobar and Javier [26] proposed a model in which they manually located the feature points and then calculated the Gabor jet manually which describes the behavior of the image around that point.

As for automated face recognition we can't locate feature points manually so to detect points automatically we have calculated the Gabor filter response at every point to see the behavior of the image around that point. A filter response at any point can be calculated by convolving the filter kernel with the image at that point. For point (X, Y), filter response denoted as R is defined as

$$R1 = x \cos(\theta) + y \sin(\theta)$$

$$R2 = -x \sin(\theta) + y \cos(\theta)$$

$$R(X, Y, \theta, \lambda) =$$

$$\sum_{x=-X}^{N-X-1} \sum_{y=-Y}^{M-Y-1} I(X+x, Y+y) f(x, y, \theta, \lambda)$$

$$f(x, y, \theta, \lambda, \sigma X, \sigma Y) = e^{-0.5 \left(\frac{R1^2}{\sigma X^2} + \frac{R2^2}{\sigma Y^2} \right)}$$

$$\theta = \lambda * k / n, k = 1, 2, \dots, n$$

Where σX and σY are the standard deviation of the Gaussian envelop along the x and y dimensions respectively. λ , θ and n are the wavelength, orientation and no of orientations respectively. $I(x, y)$ denotes NxM image.

When we apply all Gabor filters at multiple frequencies and orientations at a specific point we thus get filter response for that point. We have chosen four orientations and a constant wavelength because feature points are relatively insensitive to the Gabor kernel wavelength, while vary significantly across different orientations [5]. We have constant $\lambda = 2 * 1.414$ and $\sigma X = \sigma Y = \lambda / 2$.

B. Feature Point selection

Mostly eyes, nose, mouth and corner of lips are taken as feature points. However, in our implementation we do not fix the feature points because of varying facial characteristics of different faces such as dimples, moles etc. Human mind also uses these characteristics for face recognition. We chose the point, in a particular window of size SxT around which the behavior or response of Gabor filter kernel is maximum, as feature point.

$$\text{Where } S = \frac{N}{\sqrt{W}} \text{ and } T = \frac{M}{\sqrt{W}}$$

Where N = no of columns and M = no of rows and W is the no of windows.

Feature point located at any point can be evaluated as

$$R_f(x_o, y_o) = \max_{(x, y) \in C} (R_j(x, y))$$

Where R_j is the response of the image to the j^{th} Gabor filter and C is any window. Window size is one of the important constraints of our implemented model. It should be small enough to capture all important facial feature points, but it should be large enough so that no redundancy occurs. Feature responses are obtained by applying above method on all windows.

C. Feature vector generation

Feature vectors are generated at feature points as discussed in previous sections. p^{th} feature vector of i^{th} reference face is defined as:

$$v_{i, p} = [x_p, y_p, R_{i, j}(x_p, y_p)]$$

Where j = no of responses. Feature vector contains response with location information.

D. Similarity calculation

The degree of similarity is calculated between input image and all the images from the database. Similarity between features of input image and any image from database is calculated using.

$$S_i(p, j) = \frac{\sum_l |v_{i, p}(l)| |v_{i, j}(l)|}{\sqrt{\sum_l |v_{i, p}(l)|^2 \sum_l |v_{i, j}(l)|^2}}$$

Where $S_i(p, j)$ represents the similarity of j^{th} feature vector of input face ($v_{i, j}$) to p^{th} feature vector of i^{th} reference face, ($v_{i, p}$), where l the no of vector elements.

We chose the greatest similarity value (i.e. nearest to one) of a feature vector of input image with all the feature vectors of any image from the database, as it determines the highest degree of similarity between two feature vectors.

$$D(T_i, I) = \min[|S(i, 1) - 1|, |S(i, 2) - 1|, \dots, |S(i, (z-1)) - 1|, |S(i, z) - 1|]$$

Where $D(T_i, I)$ is the difference of i^{th} feature vector of input image T with all the feature vectors of image I of the respective degree from the database where z = n x W.

Finally, we calculate the overall difference $D(T, I)$ between an input image and an image from data base by using following equation:

$$D(T, I) = \frac{1}{z} \sum_{i=1}^z D(T_i, I)$$

E. Architecture of system

The proposed architecture consists of four main processing modules a) Data fusion b) feature value calculation c) Feature vector selection d) Similarity

calculation .Figure 3 shows the block diagram of architecture of the system, there are few pre-processing and storage modules as well.

Figure 2: Data Fusion for Face Recognition Architecture

IV. RESULTS

A. Fusion Results

Fig 3 show the face and non-face images obtained from infrared and visual image sources. Fig 4 shows the resultant fused face images obtained using the data fusion techniques discussed in section 2. Fig 4(a) shows resultant fused image obtained from the pixel level absolute maximum selection technique [22]. Although, a brighter image is obtain but with blocky effects which make it difficult to extract the desire feathurs. Fig 4(b) shows the fused image by Window base absolute maximum selection approach [22]. Again the resultant image is not useful for the face recognition applications. Fig 4(c) shows the fused image by using simple averaging technique. Here the resultant image is significantly improve and is more informative. Simple averaging removes the blocky effects present in the previous approaches with a little blurring. Fig 4(d) shows

the resultant fused image obtained from the proposed weighted average approach. Here the improvement is evident; a more informative image is obtained without any blurry effect. This image can be use for the accurate and efficient face recognition. Fig 5 shows fused satellite images after application of the various data fusion techniques. Fig 5(a) shows fused image obtained from pixel level absolute maximum selection technique. Fig 5(b) shows fused image achieved from window base absolute maximum selection approach. Fig 5(c) shows fused image from simple averaging technique and fig 5(d) shows fused image from the weighted average technique. Note that all the technique produce better results for the satellite images. All the resultant fused images are more informative then the other images (i-e visual and thermal).

Fig 3a: Face Thermal Image

Fig 3b: Face Thermal Image

Fig 3c: Satellite Thermal Image

Fig 3d: Satellite Visual Image

Fig 4a: Pixel level absolute

Fig 4b: Window base absolute

Fig 4c: Simple averaging

Fig 4d: Weighted average

Fig 5a: Pixel level absolute

Fig 5b: Window base absolute

Fig 5c: Simple averaging

Fig 5d: Weighted average

B. Face Recognition Accuracy Results

Experiments were performed on total of 17 images of 7 different candidates from Equinox facial database [14] with variable illuminations. We have performed face recognition on the fused image achieved from the fusion techniques discussed in section II. Accuracy results achieved for different data fusion techniques employed are provided in Table 1.

Table 1: Accuracy results of Gabor Filter technique

Image Fusion Technique	Accuracy for 60 X 60 resolution images
Simple Spatial Fusion	91.00%
Abs max selection in DWT	90.31%
Window base absolute maximum selection:	90.31%
Optimized Image Fusion	95.84%

The above table shows that the accuracy of recognition system utilizing the fused images is higher for proposed simple averaging DWT domain fusion than any other image fusion technique.

V. CONCLUSIONS

In this paper various image fusion techniques are discussed and Gabor filter face recognition technique is utilized to measure the effectiveness of the fusion technique employed. We have studied both spatial and wavelet domain data fusion technique and have proposed a wavelet based data fusion, which works equally good for face images as it works for all other types of images. Our

experiments have shown that the proposed optimized wavelet domain fusion technique can handle light and expression variations. The fused images were passed to face recognition system and the accuracy results were calculated and compared. It has been concluded that recognition accuracy for proposed fusion technique is greater than all other fusion techniques.

VI. FUTURE WORK

We are currently working on the extension of the proposed work for the video based real time object tracking/face recognition. Where, both thermal and visual video frame would be fused using the proposed technique to improve the efficiency of the object tracking and video based face recognition.

ACKNOWLEDGEMENT

We are thankful to COMSATS Institute of IT, HEC and PTCL Pakistan, for providing us a research environment, development facilities and travel grant support to attend conferences.

REFERENCES

- [1] Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: *Face recognition: A literature survey*. ACM Computing Surveys (CSUR) 35 (2003) 399–458
- [2] Turk, M., Pentland, A.: *Eigenfaces for Recognition*, *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 1991, pp. 71-86.
- [3] Bartlett, M., Movellan, J., Sejnowski, T.: *Face Recognition by Independent Component Analysis*,

- IEEE Trans. on Neural Networks*", Vol. 13, No. 6, November 2002, pp. 1450-1464.
- [4] Belhumeur, P., P.Hespanha, J., Kriegman, D.: *Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection*, 4th European Conference on Computer Vision, ECCV'96, 15-18 April 1996, Cambridge, UK, pp. 45-58.
 - [5] Wiskott, L., Fellous, J.-M., Kuiger, N. von der Malsburg, C.: *Face Recognition by Elastic Bunch Graph Matching*, Chapter 11 in *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, eds. L.C. Jain et al., CRC Press, 1999, pp. 355-396.
 - [6] Creed F. Jones III, *Color Face Recognition using Quaternionic Gabor Filters*, PhD Dissertation, 15, January 2003.
 - [7] Prokoshi, F., Riedel, R.: *Infrared Identification of Faces and Body Part.*, In *BIOMETRICS Personal Identification in Networked Society*, Kluwer Academic Publishers (1998).
 - [8] Prokoshi, F.: *History, current status, and future of infrared identification*, In the Proceedings of IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, Hilton Head Island, South Carolina, USA (2000) 5-14
 - [9] Socolinsky, D., Wolff, L., Neuheisel, J., Eveland, C.: *Illumination invariant face recognition using thermal infrared imagery*. In the Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001). Volume 1., Kauai, Hawaii, United States (2001) 527-534
 - [10] Pavlidis, I., Symosek, P.: *The imaging issue in an automatic face/disguise detection system*, In the Proc. of IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, Hilton Head Island, South Carolina, USA (2000) 15-24
 - [11] Wilder, J., Phillips, P., Jiang, C., Wiener, S.: *Comparison of visible and infrared imagery for face recognition*, In the Proc. of the Second International Conference on Automatic Face and Gesture Recognition, Killington, Vermont (1996) 182-187
 - [12] Socolinsky, D., Selinger, A.: *A comparative analysis of face recognition performance with visible and thermal infrared imager*, In the Proc. of 16th International Conference on Pattern Recognition. Volume 4., Quebec, Canada (2002) 217-222
 - [13] Chen, X., Flynn, P., Bowyer, K.: *PCA-based face recognition in infrared imagery: Baseline and comparative studies*, In the Proc. of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, Nice, France (2003) 127-134
 - [14] Piella, G.: *A general framework for multiresolution image fusion: from pixels to regions*. Information Fusion, 9:259-280, December 2003
 - [15] Klein, L. A.: *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*, SPIE Vol. PM138-H, Li, B. image processing, Barcelona, Spain, 2003
 - [16] Das, S., Zhang, Y.-L., Krebs.: *Color night vision for navigation and surveillance*, In the Proceeding of the fifth joint conference on information science, Atlantic City, NJ, February 2000.
 - [17] Li, H., Manjunath, S., and Mitra, K.: *Multisensor image fusion using the wavelet transform*, Graphical Models and Image Processing, vol.57, no.3, pp.235-245, 1995
 - [18] Huq, A., Mirza, A. M., Qamar, Sajid: *An optimize image fusion algorithm for night time surveillance and navigation*, in the proceedings of IEEE International Conference on Emerging Technologies, Islamabad, Pakistan, 17-18 September 2005.
 - [19] H. Li, B. S. Manjunath and S. K. Mitra, "Multisensor Image Fusion Using the Wavelet Transform," Proc. first international conference on image processing, ICIP 94, Austin, Texas, Vol. I, pp. 51-55, Nov 1994
 - [20] Fay, D., Waxman, A., Aguilar, M., Ireland, D., Racamato, J.: *Fusion of Multi-Sensor Imagery for Night Vision: Color Visualization, Target Learning and Search*, In the proc. of the 3rd International Conference on Information Fusion, Vol 1, pp. TuD3-3 -TuD3-10, July, 2000.
 - [21] Kong, S., Heo, J., Abidi, B., Paik, J., and Abidi, M.: *Recent Advances in Visual and Infrared Face Recognition - A Review*, the Journal of Computer Vision and Image Understanding, Vol. 97, No. 1, pp. 103-135, June 2005.
 - [22] Taj, J. A., Ali, U., Qureshi, R. J., Khan, S.A.: "Fusion of Thermal and Visual Images for efficient Face Recognition using Gabor Filter" Accepted for publication in Proc. of The 4th ACS/IEEE International Conference on Computer Systems and Applications, March 8-11, 2006, Dubasi/Sharjah, UAE.
 - [23] Ali, U., Taj, J. A. "Gabor Filter based Efficient Thermal and Visual Face Recognition Fusion Architectures" accepted for publication in The fourth International Conference on Active Media Technology, June 7-9 2006 Brisbane, Australia.
 - [24] Equinox: Face database (2004). www.equinoxsensors.com/products/HID.html
 - [25] Heo, J., Kong, S., Abidi, B., and Abidi, M.: *Fusion of Visual and Thermal Signatures with Eyeglass Removal for Robust Face Recognition*, IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum in conjunction with CVPR 2004, pp. 94-99, Washington, D.C., July 2004.
 - [26] Fan, W., Wang, Y., Liu, W., Tan, T., *Combining Null Space-based Gabor Features for Face Recognition*, In the proc. of 17th International Conference on Pattern Recognition.