Fusion of Visual and Thermal Signatures with Eyeglass Removal for Robust Face Recognition

Conference Paper · July 2004

DOI: 10.1109/CVPR.2004.77 · Source: IEEE Xplore

Besma Abidi
Phelps2020, Inc.

Mongi Abidi
University of Tennessee

Fusion of Visual and Thermal Signatures with Eyeglass Removal for Robust Face Recognition

Jingu Heo, Seong G. Kong, Besma R. Abidi, and Mongi A. Abidi

Imaging, Robotics, and Intelligent Systems Laboratory
Department of Electrical and Computer Engineering
The University of Tennessee
Knoxville, TN 37996-2100, U.S.A.

***Abstract* – This paper describes a fusion of visual and thermal infrared (IR) images for robust face recognition. Two types of fusion methods are discussed: *data fusion* and *decision fusion*. Data fusion produces an illumination-invariant face image by adaptively integrating registered visual and thermal face images. Decision fusion combines matching scores of individual face recognition modules. In the data fusion process, eyeglasses, which block thermal energy, are detected from thermal images and replaced with an eye template. Three fusion-based face recognition techniques are implemented and tested: Data fusion of visual and thermal images (Df), Decision fusion with highest matching score (Fh), and Decision fusion with average matching score (Fa). A commercial face recognition software FaceIt® is used as an individual recognition module. Comparison results show that fusion-based face recognition techniques outperformed individual visual and thermal face recognizers under illumination variations and facial expressions.**

I. INTRODUCTION

Despite a significant level of maturity and a few practical successes, face recognition is still a highly challenging task in pattern recognition and computer vision [1]. Face recognition based only on the visual spectrum has shown difficulties in performing consistently under uncontrolled operating conditions. Face recognition accuracy degrades quickly when the lighting is dim or when it does not uniformly illuminate the face [2]. Light reflected from human faces also varies depending on the skin color of people from different ethnic groups.

The use of thermal infrared (IR) images can improve the performance of face recognition under uncontrolled illumination conditions [3][4]. Thermal IR spectrum comprising mid-wave IR (3-5μm) and long-wave IR (8-12μm) bands has been suggested as an alternative source of information for detection and recognition of faces. Thermal IR sensors measure heat energy emitted, not reflected, from the objects. Hence thermal imaging has great advantages in face recognition in low illumination conditions or even in total darkness, where visual face recognition techniques fail. However, thermal imaging needs to solve several challenging problems. Thermal signatures are subject to change according to body temperatures caused by physical exercise or ambient temperatures. Eyeglasses may result in loss of useful information around the eyes in thermal face images since glass material blocks a large portion of thermal energy.

In this paper, the fusion of visual and thermal IR images is presented for enhancing robustness of face recognition. Fusion exploits synergistic integration of information obtained from multiple sources [5]. Two types of fusion-based face recognition techniques are developed and compared: *data fusion* and *decision fusion*. Data fusion refers to a sensor-level fusion of visual and thermal face images to produce a new face image that is invariant to illumination conditions. When eyeglasses are present, eyeglass regions are detected with an ellipse fitting method and replaced with template eye patterns to retain the details useful for face recognition. Experiments show that the data fusion method with eyeglass removal improves the recognition accuracy. Decision fusion combines the matching scores generated from the individual face recognition modules. The decision fusion with average matching score produced the highest recognition rate.

II. FUSION-BASED FACE RECOGNITION

A. Fusion of Visual and Thermal Face Recognition

Fusion techniques take advantage of the merits of multiple information sources to improve the overall recognition accuracy. Low-level data fusion integrates the data from different imaging modalities to produce a new data that contains more details. High-level decision fusion combines the decisions from multiple classification modules [6][7]. Decision fusion can be accomplished with majority voting, ranked-list combination [8], and the use of Dempster-Shafer theory. Several fusion methods have been attempted in face recognition. Biometric systems that integrate face and fingerprint data [9] and face and speech signals [10] improved the performance of personal identification. Fusion of local and global features in the face increased face recognition accuracy [11].

The combined use of visual and thermal IR image data makes a viable means for improving the performance of face recognition techniques [12]. Face recognition algorithms applied to the fusion of visible and thermal IR images consistently demonstrated better performance than when applied to either visible or thermal IR imagery alone [13]. Wilder et al. [14] showed an improved recognition performance of the fusion of visual and thermal images at the decision level.

B. *Proposed Fusion Approach*

This paper implements and tests fusion-based face recognition methods. Figure 1 shows a schematic diagram of the face recognition approaches discussed in this paper. Data fusion (Df) produces illumination-invariant face images by adaptively integrating visual and thermal face images. Decision fusion schemes refine the classification based on the average matching score (Fa) or on the highest matching score (Fh). Although the concept of decision fusion can have a much broader interpretation, the decision fusion discussed combines the matching scores obtained from individual face recognition modules.

Registered visual and thermal images of the same size are normalized using the eye coordinates extracted from the visual image. When eyeglasses are present in the images, eyeglass regions are found by the use of ellipse fitting and replaced, in the thermal images, with an average eye template to enhance data fusion. FaceIt®, a commercial face recognition software package highly ranked in the face recognition vendor test (FRVT) [15][16], is used as an individual face recognition module for generating matching scores.

III. VISUAL AND THERMAL IMAGE FUSION

A. *Weighted Averaging for Data Fusion*

A simple data fusion can be represented as a weighted sum of pixel intensities from individual sensor data:

$$F(x,y) = a(x,y)V(x,y) + b(x,y)T(x,y) \quad (1)$$

where $F(x,y)$ denotes the fused output of a visual image $V(x,y)$ and a thermal image $T(x,y)$. The coefficients $a(x,y)$ and $b(x,y)$ represent the weights of each pixel ($a(x,y) + b(x,y) = 1$). Figure 2 shows the image fusion based on average intensity using both images ($a(x,y) = b(x,y) = 0.5$). In general, weight factors can be determined according to brightness intensity distributions. When a subject is measured in low-illumination conditions, the weight factors will be adjusted so that $a(x,y) < b(x,y)$. When the overall thermal contour of the face exceeds the average contour measured in a normal room temperature range, the weights will need to be $a(x,y) > b(x,y)$.

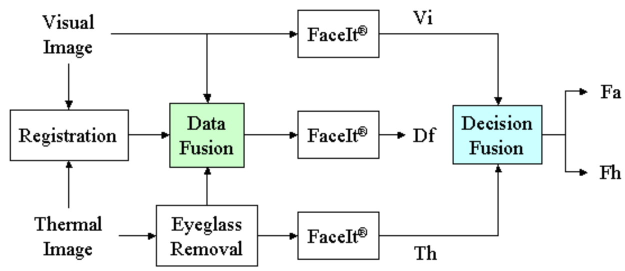
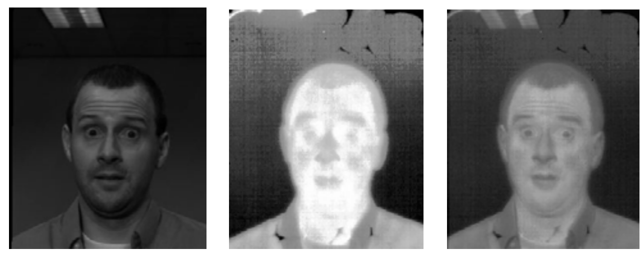


Figure 1: Visual and thermal face recognition techniques. Visual (Vi), thermal (Th), data fusion (Df), and decision fusion based on average matching score (Fa) and highest matching score (Fh).



(a) $V(x,y)$ (b) $T(x,y)$ (c) $F(x,y)$

Figure 2: Data fusion of visual and thermal images. (a) Visual image, (b) Thermal image, and (c) Data-fused image of (a) and (b) with $a(x,y) = b(x,y) = 0.5$.

B. *Eyeglass Detection using Ellipse Fitting*

The eyeglass regions in thermal face images can be represented by ellipses. A thermal image, binarized with a threshold, provides data points for fitting with ellipses. After morphological filtering for noise reduction, the data points in the binarized image are connected using the Freeman chain coding with 8-connectivity [17]. A non-iterative ellipse-fitting algorithm [18] is applied to each set of connected components to produce an ellipse. Figure 3 shows an ellipse with the parameters used for eyeglass detection in thermal face images. The center of an *i*th ellipse is denoted by C_i, $2\alpha_i$ and $2\beta_i$ are the lengths of the major axis and the minor axis respectively, and θ_i indicates the orientation angle of the ellipse in the range of $-\pi/2 < \theta_i < \pi/2$.

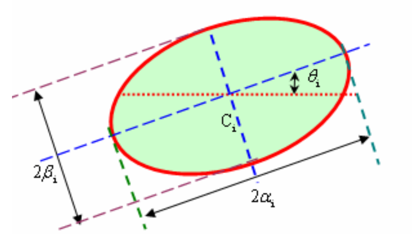
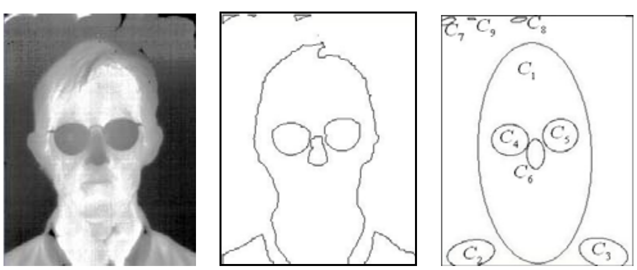


Figure 3: Ellipse parameters.

Similarities of the ellipses within the face region, or inside the biggest ellipse, are tested for possible eyeglass regions. Among all the candidate glasses, a pair of ellipses of similar shape and size is considered as eyeglasses in thermal images. In this paper, the similarity of *i*th and *j*th ellipsoids is defined as:

$$S_{ij} = \frac{\alpha_i \beta_i}{\alpha_j \beta_j} \left(\frac{1}{1+|\theta_{ij}|} \right) \left(\frac{1}{1+|\theta_i + \theta_j|} \right) \qquad (2)$$

where θ_{ij} represents the angle of the line segment that connects the centers of the two ellipses C_i and C_j. We assume that $\alpha_j\beta_j > \alpha_i\beta_i$ so the similarity measure S_{ij} is less than 1. For a shape constraint, the ellipses must have the ratio of major and minor axis (α/β) in the range of $0.5 < \alpha/\beta < 1.5$. For a size constraint, the ratio of major axis to the face height is $0.2 < \alpha/\alpha_F < 0.8$ and the ratio of minor axes to face height is $0.4 < \beta/\beta_F < 0.8$, where α_F and β_F indicate major and minor axes of the biggest ellipse. Two ellipses with the highest similarity measure of $S_{ij} > 0.7$ are considered as eyeglasses. Figure 4 illustrates an example of detecting eyeglasses in thermal images using the ellipse fitting. Among the ellipses generated from each connected component, the biggest ellipse (C_1) corresponds to the face. Ellipses outside the face region (C_2, C_3, C_7, C_8, and C_9) are not considered for similarity checking. For the three ellipses inside the face region, the similarities are calculated as $S_{45} = 0.96$, $S_{46} = 0.38$, and $S_{56} = 0.40$. As a result, the two ellipses C_4 and C_5 with the highest similarity are identified as eyeglasses.



(a) (b) (c)

Figure 4: Eyeglass detection example using ellipse fitting. (a) Original image, (b) Connected components of the binary image, (c) Eyeglass regions detected using the ellipse fitting method.

Figure 5 shows the performance of the glass detection method discussed above as a function of intensity in the range of [0,1] in terms of false acceptance rate (FAR) and false rejection rate (FRR). The threshold can be found where the false rejection rate reaches the minimum. In this paper, the threshold was selected to be 0.57.

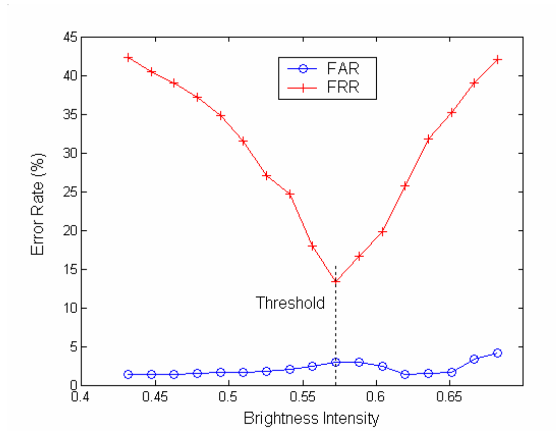


Figure 5: Performance of eyeglass detection.

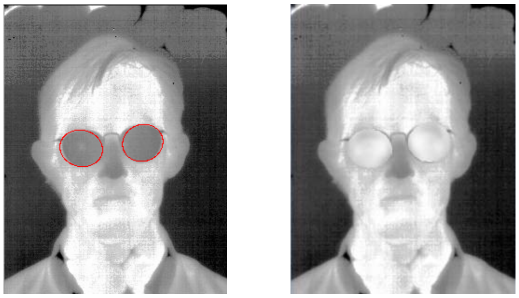
Table 1 summarizes the performance of eyeglass detection algorithm with the ellipse fitting method. Correct detection rate was 86.6% for the subjects wearing eyeglasses. For the face images with no eyeglasses, 97.1% true negative accuracy was achieved. False positive and false negative errors were 2.9% and 13.4%, respectively. The database used in this experiment is comprised of thermal images from the database developed by the National Institute of Standards and Technology (NIST) and Equinox Corporation [19].

Table 1: Performance of eyeglass detection

Detection Types	Matched /Total	Rate (%)
Eyeglass → Eyeglass (True Positive)	445/ 514	86.6
No Eyeglass → No Eyeglass (True Negative)	1096/1129	97.1
No Eyeglass → Eyeglass (False Positive)	33/1129	2.9
Eyeglass → No eyeglass (False Negative)	69/514	13.4

C. Data Fusion with Eyeglass Removal

Detected eyeglass regions are replaced with an average eye template in the thermal images to enhance visual quality around the eyes in data-fused images. Template eye regions are obtained from the average of all thermal face images without glasses. Figure 6 shows the result of eyeglass replacement with a template eye pattern. A geometrical transformation of eye templates is performed to fit the templates to the eyeglass regions detected by the use of ellipse fitting. The eye templates for the left and the right eyeglasses superimpose eyeglass regions after rotating and resizing.

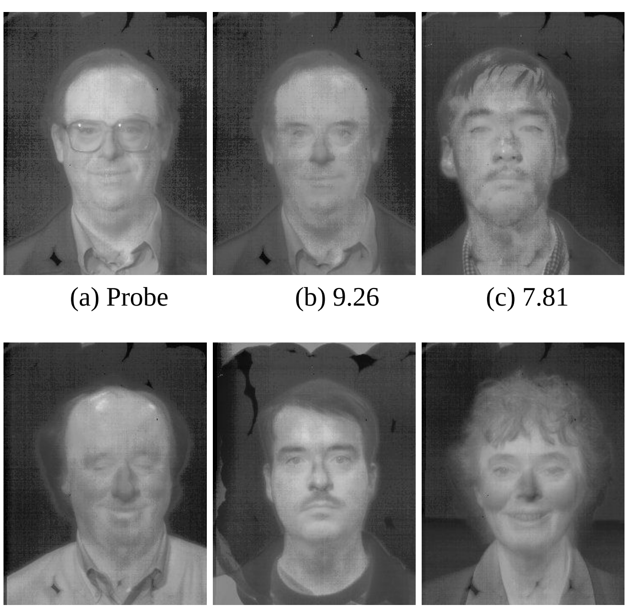


(a) (b)

Figure 6: Eyeglass removal (a) Eyeglasses detected, (b) Eyeglasses replaced by eye templates after rotation and resizing.

(a) (b) (c)

Figure 7: Adaptive data fusion with eyeglass removal. (a) Original image, (b) Direct fusion of visual and thermal images without eyeglass removal, and (c) Fused image after eyeglass removal.



(a) Probe (b) 9.26 (c) 7.81

(d) 7.09 (e) 7.01 (f) 6.74

Figure 8: Face recognition with data fusion (Df) with eyeglass removal. (a) Probe, (b)-(f) First five matches.

Figure 7 shows an example adaptive data fusion result by eyeglass detection and replacement with the eye template in thermal images. Eyeglass removal enhances the visual quality of data-fused images. Figure 8 shows an example of face recognition using data-fused visual and thermal images (Df). The five classification result faces are in descending order of matching scores.

D. Decision Fusion

Decision fusion produces a new ranked list by combining confidence measures from individual face recognition modules. In this paper, matching scores are used to determine ranked lists. Matching scores generated by FaceIt® measure the degree to which the probe image and the gallery image are similar. The matching score (M_F) of decision fusion can be derived using the individual scores of visual recognition module (M_V) and of thermal recognition module (M_T).

Decision fusion with average matching score (Fa) determines the matching score as a weighted sum of M_V and M_T:

$$M_F = w_V M_V + w_T M_T \tag{3}$$

Decision fusion with highest matching score (Fh) takes the largest matching score of the two:

$$M_F = \max(M_V, M_T) \tag{4}$$

where w_V and w_T denote weight factors for the matching scores of visual and thermal face recognition modules. In this paper, $w_V = w_T = 0.5$.

IV. PERFORMANCE EVALUATION

A. Database

The National Institute of Standards and Technology and Equinox Corporation built an extensive database of face images using registered broadband-visible/IR camera sensors for experimentation and statistical performance evaluations [19]. The NIST/Equinox database used for evaluation of fusion-based face recognition performances consists of visual and thermal IR images of 3,244 (1,622 per modality) faces from 90 individuals. One image for each face taken with a frontal lighting condition is used for the gallery. Probe images are divided according to different conditions. Original 12-bits gray level thermal images were converted into 8 bits and histogram equalized. Table 2 describes the NIST/Equinox databases of visual and thermal IR face images used in the experiments.

Table 2: The NIST/Equinox database of visual and thermal IR face images

Dataset	Visual (Thermal)	Eyeglass	Lighting	Expression
Gallery	90(90)	Off	Frontal	Neutral
Probe 1	283(283)	Off	Frontal	Vary
Probe 2	370(370)	Off	Left	Vary
Probe 3	365(365)	Off	Right	Vary
Probe 4	177(177)	On	Frontal	Vary
Probe 5	172(172)	On	Left	Vary
Probe 6	165(165)	On	Right	Vary

B. *Performance Comparison*

The recognition results of fusion-based techniques (Df, Fa, and Fh) are compared with the single modality cases (Vi and Th) at various lighting directions. The three probe sets 1, 2, and 3 contain 1,018 images in total with no eyeglasses. Figure 9 demonstrates the first 10 best matches of different recognition methods in terms of matching scores. Visual face recognition relatively under-performed due to illumination variations. Fusion-based methods yield reliable recognition results.

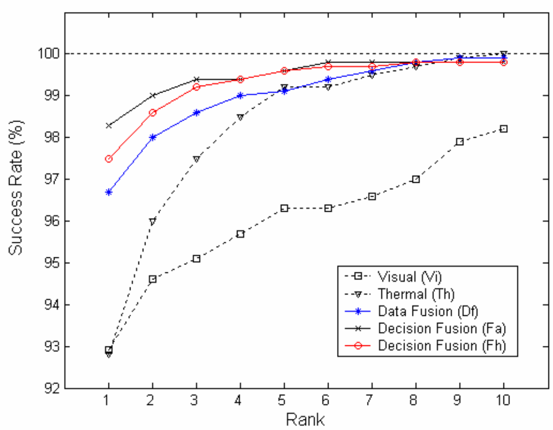
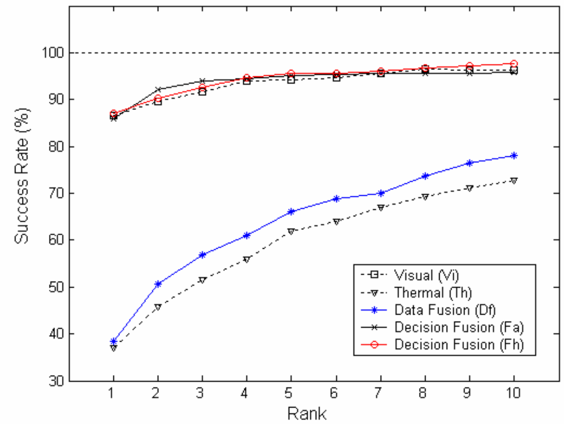
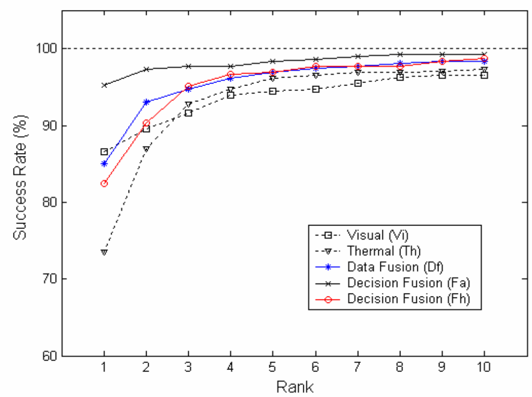


Figure 9: Performance evaluation of fusion-based face recognition when no eyeglasses are present (probes 1, 2, 3).

Figure 10 compares the performances of the five face recognition techniques when the subjects wear eyeglasses (probes 4, 5, and 6). There are total 514 images in the set. In Figure 10(a), thermal face recognition and data fusion without eyeglass removal show unsatisfactory results due to the energy blocking effect of eyeglasses. Eyeglasses slightly affect the performance of visual face recognition while affecting that of thermal face recognition significantly. Figure 10(b) demonstrates that eyeglass removal greatly improves the recognition performance in thermal and data fusion techniques. Decision fusion with average matching score gives the best performance.



(a)



(b)

Figure 10: Performance evaluation of fusion-based face recognition when eyeglasses are present (probe 4, 5, 6) (a) Without eyeglass removal, (b) With eyeglasses replaced with templates.

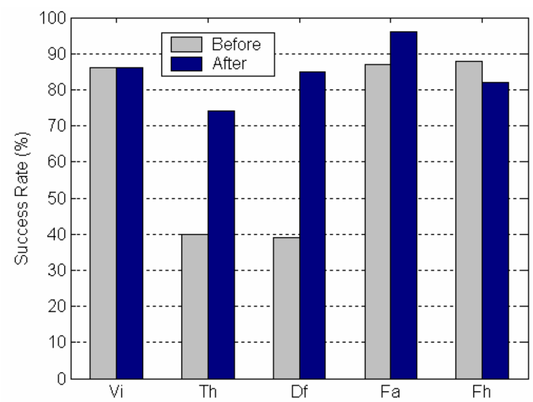


Figure 11: Performance comparison in terms of the first match success rates for the face images with eyeglasses before and after eyeglass removal.

Figure 11 summarizes the performance comparison of fusion-based face recognition methods before and after eyeglass removal in terms of first match success rates. The probe sets 4, 5, and 6 are used. First match is important especially for verification tasks. Decision fusion with average matching score shows higher performance than any other methods. Eyeglass removal greatly improved the performance for thermal and data fusion face recognition methods. Decision fusion techniques demonstrate consistent recognition accuracies.

V. CONCLUSION

This paper presents and compares fusion-based face recognition techniques. Data fusion produces an integrated image from a pair of registered visual and thermal IR images. The data-fused image is invariant to illumination directions and is robust under low lighting conditions. When a subject is wearing eyeglasses, eyeglass regions are detected in thermal images using an ellipse fitting method and replaced with a template eye pattern to enhance data fusion. In decision fusion, a new rank list is generated by the use of the maximum matching score or the average score of the individual face recognition modules.

Experiment results show that thermal face recognition performed better than visual face recognition under various illumination and facial expression conditions in case of no eyeglasses (Figure 9). When eyeglasses are present in the image, thermal face recognition and a direct data fusion without eyeglass removal produced lower performance. Eyeglass removal significantly improved the recognition accuracy of the data fusion. Decision fusion with average matching score consistently demonstrated superior recognition accuracies (Figure 10 and Figure 11).

ACKNOWLEDGMENT

This research was supported by the Office of Naval Research under grant # N000143010022.

REFERENCES

[1] I. Craw, N. Costen, T. Kato, and S. Akamatsu, “How should we represent faces for automatic recognition?,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 8, pp.725-736, 1999.

[2] Y. Adini, Y. Moses, and S. Ullman, “Face recognition: The problem of compensating for changes in illumination direction,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp.721-732, 1997.

[3] F. Prokoski, “History, current status, and future of infrared identification,” *Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, pp.5-14, 2000.

[4] D. A. Socolinsky, A. Selinger, and J. D. Neuheisel, “Face recognition with visible and thermal infrared imagery,” *Computer Vision and Image Understanding*, Vol. 91, pp.72-114, 2003.

[5] D. L. Hall and J. Llinas, *Handbook of Multisensor Data Fusion*, CRC Press, 2001.

[6] A. Ross and A. Jain, “Information fusion in biometrics,” *Pattern Recognition Letters*, Vol. 24, No. 13, pp.2115-2125, 2003.

[7] T. K. Ho, J. J. Hull, and S. N. Srihari, “Decision combination in multiple classifier systems,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 1, pp.66-75, 1994.

[8] C. Sanderson and K. K. Paliwal, “Information fusion and person verification using speech and face information,” *IDIAP Research Report 02-33*, Martigny, Switzerland, 2002.

[9] L. Hong and A. Jain, “Integrating faces and fingerprints for personal identification,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 12, pp.1295-1307, 1998.

[10] S. Ben-Yacoub, Y. Abdeljaoued, and E. Mayoraz, “Fusion of face and speech data for person identity verification,” *IEEE Trans. on Neural Networks*, Vol. 10, No. 5, pp.1065-1074, 1999.

[11] Y. Fang, T. Tan, and Y. Wang, “Fusion of global and local features for face verification,” *Proc. of Int. Conf. on Pattern Recognition*, pp.382-385, 2002.

[12] 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, in print*, 2004.

[13] X. Chen, P. Flynn, and K. Bowyer, “Visible-light and infrared face recognition,” *Proc. of Workshop on Multimodal User Authentication*, pp.48-55, 2003.

[14] J. Wilder, P. J. Phillips, C. Jiang, and S. Wiener, “Comparison of visible and infrared imagery for face recognition,” *Proc. Int. Conf. Automatic Face and Gesture Recognition*, pp.182-187, 1996.

[15] D. M. Blackburn, M. Bone, and P. J. Phillips, *Facial Recognition Vendor Test 2000*: *Evaluation Report*, 2001.

[16] P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E.Tabassi, and M. Bone, *Face Recognition Vendor Test 2002*: *Evaluation Report*, 2003.

[17] H. Freeman and L. S. Davis, “A corner-finding algorithm for chain-coded curves,” *IEEE Trans. on Computers*, Vol. 26, pp.297-303, 1977.

[18] A. Fitzgibbon, M. Pilu, and R. B. Fisher, “Direct least square fitting of ellipses,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 5, pp.476-480, 1999.

[19] http://www.equinoxsensors.com/products/HID.html