Physiology-Based Face Recognition

P. Buddharaju
Dept. of Computer Science
Univ. of Houston
Houston, TX 77204
braju@cs.uh.edu

I. T. Pavlidis
Dept. of Computer Science
Univ. of Houston
Houston, TX 77204
ipavlidi@central.uh.edu

P. Tsiamyrtzis
Dept. of Statistics
Athens Univ.
Athens, Greece
pt@aueb.gr

Abstract

We present a novel approach for face recognition based on the physiological information extracted from thermal facial images. First, we delineate the human face from the background using a Bayesian method. Then, we extract the blood vessels present on the segmented facial tissue using image morphology. The extracted vascular network produces contour shapes that are unique for each individual. The branching points of the skeletonized vascular network are referred to as Thermal Minutia Points (TMPs). These are reminiscent of the minutia points produced in fingerprint recognition techniques. During the classification stage, local and global structures of TMPs extracted from test images are matched with those of database images. We have conducted experiments on a large database of thermal facial images collected in our lab. The good experimental results show that our proposed approach has merit and promise.

1. Introduction

Biometrics has received a lot of attention during the last few years both from the academic and business communities. It has emerged as a preferred alternative to traditional forms of identification, like card IDs, which are not embedded into one's physical characteristics. Research into several biometric modalities including face, fingerprint, iris, and retina recognition has produced varying degrees of success [1]. Face recognition stands as the most appealing modality, since it is the natural mode of identification among humans and totally unobtrusive. At the same time, however, it is one of the most challenging modalities [2]. Faces are 3D objects with rich details that vary with orientation, illumination, age, and artifacts (e.g., glasses). Research into face recognition has been biased towards the visible spectrum for a variety of reasons. Among those is the availability and low cost of visible band cameras and the undeniable fact that face recognition is one of the primary activities of the human visual system. Machine recognition of human faces, however, has proved more problematic than the seemingly effortless face recognition performed by humans. The major culprit is light variability, which is prevalent in the visible spectrum due to the reflective nature of incident light in this band. Secondary problems are associated to the difficulty of detecting facial disguises [3].

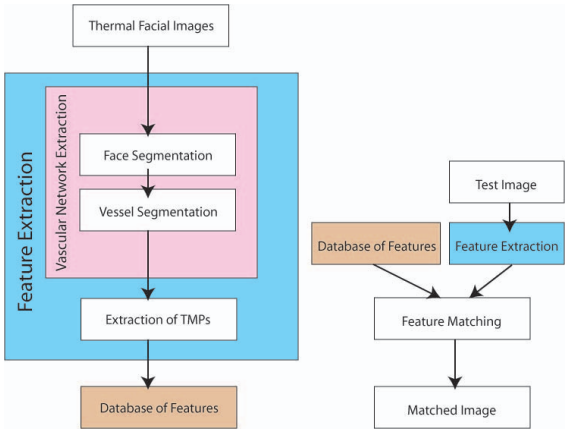
As a cure to the aforementioned problems, researchers have started investigating the use of thermal infrared for face recognition purposes [4, 5, 6]. However, many of these research efforts in thermal face recognition use the thermal infrared band only as a way to see in the dark or reduce the deleterious effect of light variability [7, 8]. Methodologically, they do not differ very much from face recognition algorithms in the visible band and can be classified as appearance-based [9, 10] or feature-based approaches [11, 12].

In this paper, we present a novel approach to the problem of facial recognition that realizes the full potential of the thermal infrared band. It consists of a statistical face segmentation and physiological feature extraction algorithms tailored to thermal phenomenology. Prokoski et al. anticipated the possibility of extracting the vascular network from thermal facial images and using it as a feature space for face recognition [13]. However, they did not present an algorithmic approach for achieving this. To the best of our knowledge, this is the first attempt to develop a face recognition system using physiological information on the face. Our goal is to promote a different way of thinking in areas where thermal infrared should be approached in a distinct manner with respect to other modalities.

Figure 1 shows the architecture of the proposed system. It operates in two phases:

- *Off-line phase:* The thermal facial images are captured by a thermal infrared camera. A two-step segmentation algorithm is applied on the input image to extract the vascular network on the face. TMPs are detected on the branching points of the vascular network and are stored in the database (see Figure 1(a)).

0-7803-9385-6/05/$20.00 ©2005 IEEE



(a) (b)

Figure 1: Block diagram of our face recognition system: (a) off-line phase; (b) on-line phase.

- *On-line phase:* Given a query image, TMPs of its vascular network are extracted and are matched against those that are stored in the database (see Figure 1(b)).

In the following sections, we will describe our face recognition system in detail. In Section 2, we present the vascular network extraction algorithm. In Section 3, we discuss about vascular network matching. In Section 4, we present the experimental results and attempt a critical evaluation. We conclude our paper in Section 5.

2. Vascular Network Extraction

A thermal infrared camera with reasonable sensitivity provides the ability to directly image superficial blood vessels on the human face. The pattern of the underlying blood vessels is unique to each individual, and the extraction of this vascular network can provide the basis for a unique feature vector.

2.1. Face Segmentation

Due to its physiology, a typical human face consists of hot parts (dominant mode) that correspond to tissue areas that are rich in vasculature and cold parts that correspond to tissue areas with sparse vasculature. This casts the human face as a bimodal distribution entity, which can be modelled using a mixture of two Normal distributions. Similarly, the background can be described by a bimodal distribution with walls being the cold objects (dominant mode) and parts of the subject's body dressed in cloths being the hot objects. We approach the problem of delineating facial tissue from background using a Bayesian framework [14] since we have apriori knowledge of the bimodal nature of the scene.

We call θ the parameter of interest, which takes two possible values (s and b) with some probability. For each pixel x in the image, we draw our inference of whether it represents skin (i.e., $\theta = s$) or background (i.e., $\theta = b$) based on the posterior distribution $p^{(t)}(\theta|x_t)$ given by:

$$p^{(t)}(\theta|x_t) = \begin{cases} p^{(t)}(s|x_t), \ \theta = s \\ p^{(t)}(b|x_t) = 1 - p^{(t)}(s|x_t), \ \theta = b \end{cases} \quad (1)$$

where according to the Bayes' theorem:

$$p^{(t)}(s|x_t) = \frac{\pi^{(t)}(s) f(x_t|s)}{\pi^{(t)}(s) f(x_t|s) + \pi^{(t)}(b) f(x_t|b)}. \quad (2)$$

Here, $\pi^{(t)}(s)$ is the prior distribution at time t, which is equal to the posterior distribution at time $t-1$.

In the first frame ($t = 0$), the priors for skin and background are considered equiprobable:

$$\pi^{(0)}(s) = \frac{1}{2} = \pi^{(0)}(b), \quad (3)$$

for $t \geq 1$:

$$\pi^{(t)}(\theta) = \begin{cases} \pi^{(t)}(s) = p^{(t-1)}(s|x_{t-1}), \ \theta = s \\ \pi^{(t)}(b) = p^{(t-1)}(b|x_{t-1}), \ \theta = b \end{cases} \quad (4)$$

Based on the bimodal view of the skin and background distributions we will have for the likelihood for $t \geq 0$:

$$f(x_t|\theta) = \begin{cases} f(x_t|s) = \sum_{i=1}^{2} w_{s_i}^{(t)} N(\mu_{s_i}^{(t)}, \sigma_{s_i}^{2(t)}), \ \theta = s \\ f(x_t|b) = \sum_{i=1}^{2} w_{b_i}^{(t)} N(\mu_{b_i}^{(t)}, \sigma_{b_i}^{2(t)}), \ \theta = b \end{cases} \quad (5)$$

where $w_{s_2} = 1 - w_{s_1}$ and $w_{b_2} = 1 - w_{b_1}$. The mixture parameters $w_{s_i}, \mu_{s_i}, \sigma_{s_i}^2, w_{b_i}, \mu_{b_i}, \sigma_{b_i}^2, \ i = 1, 2$ can be initialized and updated using the EM algorithm. For that, we select N representative facial frames (off-line) from a variety of subjects. This will serve as our training set. Then, we manually segment on all the N frames, skin and background areas, which will yield N_s skin and N_b background pixels.

Once a data point x_t becomes available, we decide that it represents skin if the posterior distribution for the skin, $p^{(t)}(s|x_t) > 0.5$ and that it represents background otherwise. Figure 2 depicts the Bayesian segmentation of a subject's face. On one hand, small parts of the subject's nose and left cheek have been erroneously labelled as background (white). On the other hand, a couple of small cloth patches around the subject's neck have been erroneously marked as facial skin. These are due to occasional overlapping between portions of the skin and background distributions. The isolated nature of these mislabelled patches makes them easily correctable through post-processing (morphological opening and closing).

(a) (b)

Figure 2: Segmentation of facial skin region: (a) result of Bayesian segmentation; (b) correction of mislabelled patches through post-processing.

(a) (b)

Figure 3: Vascular network extraction: (a) anisotropically diffused image; (b) blood vessels extracted using white top hat segmentation.

2.2. Segmentation of Superficial Blood Vessels

Once a face is delineated from the rest of the scene, the segmentation of superficial blood vessels from the facial tissue is carried out in the following two steps [14, 15]:

1. Smooth the image to remove the unwanted noise.
2. Apply morphological operations to localize the superficial vasculature.

In thermal imagery of human tissue the major blood vessels have weak sigmoid edges, which can be handled effectively using anisotropic diffusion. The anisotropic diffusion filter is formulated as a process that enhances object boundaries by performing intra-region as opposed to inter-region smoothing. The mathematical equation for the process is:

$$\frac{\partial I(\bar{x},t)}{\partial t} = \nabla(c(\bar{x},t)\nabla I(\bar{x},t)). \tag{6}$$

In our case I($\bar{x},t$) is the thermal infrared image, $\bar{x}$ refers to the spatial dimensions, and t to time. c($\bar{x},t$) is called the diffusion function. The discrete version of the anisotropic diffusion filter of Equation (6) is as follows:

$$\begin{aligned} I_{t+1}(x,y) = I_t + \frac{1}{4} * [&c_{N,t}(x,y)\nabla I_{N,t}(x,y) \\ &+ c_{S,t}(x,y)\nabla I_{S,t}(x,y) + c_{E,t}(x,y)\nabla I_{E,t}(x,y) \\ &+ c_{W,t}(x,y)\nabla I_{W,t}(x,y)]. \end{aligned} \tag{7}$$

The four diffusion coefficients and four gradients in Equation (7) correspond to four directions (i.e., North, South, East, and West) with respect to the location (x,y). Each diffusion coefficient and the corresponding gradient are calculated in a similar manner. For example, the coefficient along the North direction is calculated as:

$$c_{N,t}(x,y) = exp(\frac{-\nabla I^2_{N,t}(x,y)}{k^2}), \tag{8}$$

where $I_{N,t} = I_t(x,y+1) - I_t(x,y)$.

Image morphology is then applied on the diffused image to extract the blood vessels that are in low contrast compared with respect to the surrounding tissue. We employ for this purpose a top hat segmentation method, which is a combination of erosion and dilation operations. Top hat segmentation has two versions: white top hat segmentation that enhances bright objects and black top hat segmentation that enhances dark objects in the image. In our case, we are interested in white top hat segmentation [14] since it helps to enhance the bright ('hot') ridge like structures corresponding to the blood vessels. Figure 3(a) depicts the result of applying anisotropic diffusion to the segmented facial tissue shown in Figure 2(b). Figure 3(b) shows the corresponding blood vessels extracted using white top hat segmentation.

3. Vascular Network Matching

The branching points of the extracted blood vessels are localized as Thermal Minutia Points (TMPs). Different features of each TMP are considered for matching an input facial image with a database image.

3.1. Extraction of Thermal Minutia Points

Although, various methods have been proposed for robust and efficient extraction of minutia from fingerprint images [16], we are not aware of any for facial images. Most of the fingerprint approaches describe each minutia by at-least three attributes, including its location, orientation, and type. We adopt a similar approach for extracting TMPs from the facial vascular network. Specifically, we undertake the following steps:

1. Estimate the local orientation of the vascular network.
2. Skeletonize the vascular network.
3. Extract TMPs from the thinned vascular network.
4. Remove spurious TMPs.

Local orientation Ψ_{xy} is the angle formed by the blood vessel at pixel (x,y) with the horizontal axis. Estimating the orientation field at each pixel provides the basis for capturing the overall pattern of the vascular network. We use

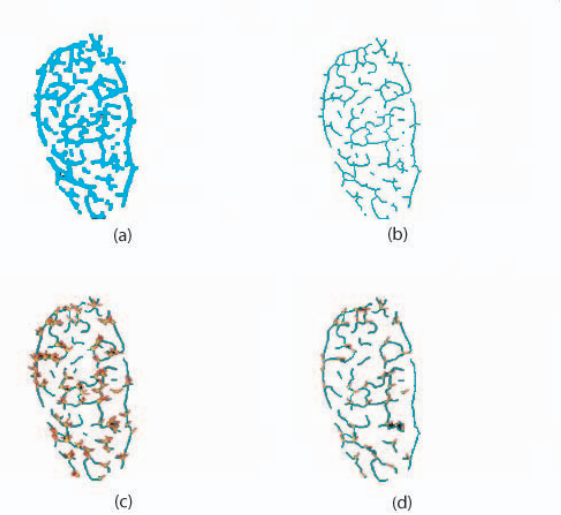


Figure 4: Results of TMP extraction algorithm: (a) vascular network extracted from thermal facial image; (b) thinned vessel map; (c) extracted TMPs from branching points; (d) spurious TMPs removed.

the approach proposed in [17] for computing the orientation image since it provides pixel-wise accuracy and also can be used in both binary and gray-scale images.

Next, the vascular network is thinned to one-pixel thickness [18]. Each pixel in the thinned map has a value of 1 if it is on the vessel and 0 if it is not. Considering 8-neighborhood $(N_0, N_1, ..., N_7)$ around each pixel, a pixel (x, y) represents a TMP if $(\sum_{i=0}^{7} N_i) > 2$. Removal of spikes and clustered TMPs reduces spurious TMPs. The vascular network of a typical facial image contains around 50-80 genuine TMPs whose location (x, y) and orientation (Ψ) are stored in the database. Figure 4 shows the results of each stage of the TMP extraction algorithm on a segmented vascular network.

3.2. Matching of Thermal Minutia Points

Numerous methods have been proposed for matching fingerprint minutiae, most of which try to simulate the way forensic experts compare fingerprints [16]. Popular techniques are alignment-based point pattern matching, local structure matching, and global structure matching. Local minutiae matching algorithms are fast, simple, and more tolerant to distortions whereas global minutiae matching algorithms have high distinctiveness. A few hybrid approaches [19, 20] have been proposed where the advantages of both methods are exploited. Our method of TMP extraction is modelled after these hybrid approaches.

For each TMP $M_{(x_i, y_i, \Psi_i)}$ that is extracted from the vascular network, consider its N nearest neighbor TMPs $M_{(x_n, y_n, \Psi_n)}$, $n = 1, ..., N$. The TMP M can then be defined by a new feature vector:

$$L_M = \{\{d_1, \varphi_1, \vartheta_1\}, \{d_2, \varphi_2, \vartheta_2\}, ..., \{d_N, \varphi_N, \vartheta_N\}, \Psi_i\} \tag{9}$$

where

$$\begin{aligned} d_n &= \sqrt{(x_n - x_i)^2 + (y_n - y_i)^2} \\ \varphi_n &= diff(\Psi_n, \Psi_i),\ n = 1, 2, ..., N \\ \vartheta_n &= diff\left(\arctan\left(\frac{y_n - y_i}{x_n - x_i}\right), \Psi_i\right) \end{aligned} \tag{10}$$

The function $diff()$ calculates the difference of two angles and bounds the result within the range $[0, 2\pi)$ [20]. Given a test image M, the feature vector of each of its TMPs is compared with the feature vector of every TMP of a database image. Two TMPs M and M' are marked to be a matched pair if the absolute difference between corresponding features is less than specific threshold values $\{\delta_d, \delta_\varphi, \delta_\vartheta, \delta_\Psi\}$. The threshold values should be chosen in such a way that they accommodate for linear deformations and translations. The final matching score between the test image and a database image is given by:

$$Score = \frac{NUM_{match}}{\max(NUM_{test}, NUM_{database})} \tag{11}$$

where NUM_{match} represents number of matched TMP pairs, and NUM_{test}, $NUM_{database}$ represent number of TMPs in test and database images respectively. If the highest matching score between the test and database images is greater than a specific threshold, we decide that the corresponding database image is the match. If the highest score is less than the threshold value, we decide that the test image does not exist in the database.

4. Experimental Results

We collected a large database of thermal facial images in our lab from volunteers representing different sex, race, and age groups. The images were captured using a high quality mid-wave infrared camera (Phoenix by FLIR Systems).

We used a subset of the dataset for evaluating the performance of the proposed face recognition algorithm. The dataset consists of 1518 thermal facial images from 138 different subjects (11 images per subject) with varying pose and facial expressions. One image from each subject is used for training, the TMPs of which are extracted and stored in the database, and the remaining ten images are used for testing. A major challenge associated with thermal face recognition is the performance over time [21]. Facial thermograms may change depending on the physical condition and environmental conditions. A few approaches that use direct temperature data for recognition reported degraded performance over time [10]. However, our approach attempts to

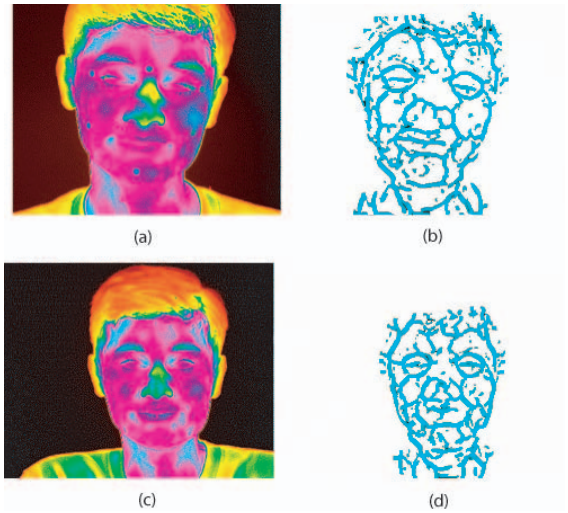


Figure 5: (a) Thermal facial image of a subject acquired on 10-17-2003 and (b) corresponding vascular network; (c) Thermal facial image of the same subject acquired on 04-29-2004 and (d) corresponding vascular network.

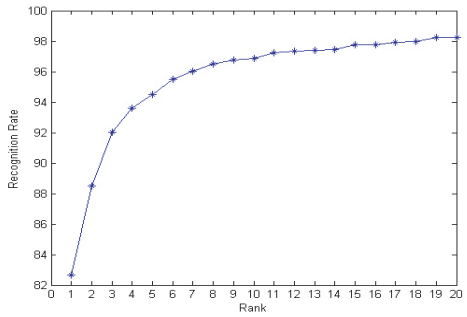


Figure 6: CMC curve of proposed method on our dataset.

solve this problem by using facial anatomical information as its feature space, which is unique to each person and at the same time is invariable to physical and environmental conditions as shown in Figure 5. The vascular network extracted from the same person with a time gap of about six months exhibits a similar pattern.

Figure 6 shows the Cumulative Math Characteristic (CMC) curve and Figure 7 shows the ROC based on various threshold values for the matching score discussed in Section 3.2. The results demonstrate the promise as well as some problems with our proposed approach. Specifically, CMC shows that rank 1 recognition is over 82% and rank 5 recognition is over 92%. This performance puts a brand new approach very close to the performance of mature visible band recognition methods. In contrast, ROC reveals a weakness of the current method, as it requires false acceptance rate over 20% to reach positive acceptance rate above the 80% range. To address this problem we believe we need to estimate and eliminate the non-linear deformations in the extracted vascular network. This will accommodate safely, large variations in pose and facial expression during the enrollment and testing phases.

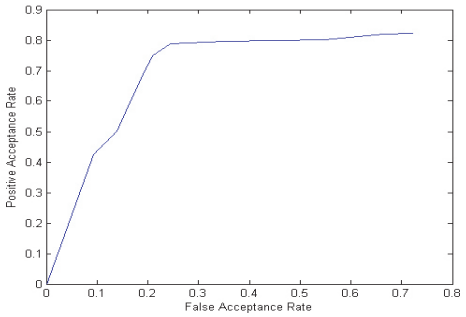


Figure 7: ROC of proposed method on our dataset.

5. Conclusions

We have outlined a novel approach to the problem of face recognition in thermal infrared - one of the fastest growing biometrics. The cornerstone of the approach is the use of unique and time invariant physiological information as feature space. The facial tissue is first separated from the background using a Bayesian segmentation method. The vascular network on the surface of the skin is then extracted based on a white top-hat segmentation preceded by anisotropic diffusion. Thermal Minutia Points (TMPs) are extracted from the vascular network and are used as features for matching test to database images. The experimental results demonstrate that our approach is very promising.

The method although young, performed well in a non-trivial database. Our ongoing work is directed towards improving the sophistication of the method regarding variable pose and facial expression and testing it comparatively in larger databases.

Acknowledgments

This research was supported in part by NSF grant # DUE-0313880. The views expressed in this paper do not necessarily represent the views of the funding Agencies.

References

[1] A. Jain, R. Bolle, and S. Pankanti. *Biometrics: Personal Identification in Networked Society*. Kluwer Academic Publishers, 1st edition, 1999.

[2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys (CSUR)*, 35(4):399–458, December 2003.

[3] I. Pavlidis and P. Symosek. The imaging issue in an automatic face/disguise detection system. In *Proceedings of IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, pages 15–24, Hilton Head Island, South Carolina, USA, June 2000.

[4] F. Prokoski. History, current status, and future of infrared identification. In *Proceedings of IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, pages 5–14, Hilton Head Island, South Carolina, USA, June 2000.

[5] D.A. Socolinsky and A. Selinger. A comparative analysis of face recognition performance with visible and thermal infrared imagery. In *Proceedings of 16th International Conference on Pattern Recognition*, volume 4, pages 217–222, Quebec, Canada, 2002.

[6] J. Wilder, P.J. Phillips, C. Jiang, and S. Wiener. Comparison of visible and infrared imagery for face recognition. In *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, pages 182–187, Killington, Vermont, October 1996.

[7] D.A. Socolinsky, L.B. Wolff, J.D. Neuheisel, and C.K. Eveland. Illumination invariant face recognition using thermal infrared imagery. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR 2001)*, volume 1, pages 527–534, Kauai, Hawaii, United States, 2001.

[8] A. Selinger and D.A. Socolinsky. Face recognition in the dark. In *Proceedings of the Joint IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum*, Washington D.C., June 2004.

[9] R. Cutler. Face recognition using infrared images and eigenfaces. cs.umd.edu/rgc/face/face.htm, 1996.

[10] X. Chen, P.J. Flynn, and K.W. Bowyer. PCA-based face recognition in infrared imagery: Baseline and comparative studies. In *Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, pages 127–134, Nice, France, October 17 2003.

[11] A. Srivastava and X. Liu. Statistical hypothesis pruning for recognizing faces from infrared images. *Journal of Image and Vision Computing*, 21(7):651–661, 2003.

[12] P. Buddharaju, I. Pavlidis, and I. Kakadiaris. Face recognition in the thermal infrared spectrum. In *Proceedings of the Joint IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum*, Washington D.C., June 2004.

[13] F.J. Prokoski and R. Riedel. *BIOMETRICS: Personal Identification in Networked Society*, chapter 9 Infrared Identification of Faces and Body Parts. Kluwer Academic Publishers, 1998.

[14] I. Pavlidis, P. Tsiamyrtzis, C. Manohar, and P. Buddharaju. *Biomedical Engineering Handbook*, chapter Biometrics: Face Recognition in Thermal Infrared. CRC Press, November 2005.

[15] C. Manohar. Extraction of superficial vasculature in thermal imaging. Master's thesis, University of Houston, Houston, TX, December 2004.

[16] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer Verlag, June 2003.

[17] M.D. Almeida Oliveira and N.J. Leite. Reconnection of fingerprint ridges based on morphological operators and multiscale directional information. In *17th Brazilian Symposium on Computer Graphics and Image Processing*, pages 122–129, Curitiba, PR, BRAZIL, October 2004.

[18] B.K. Jang and R.T. Chin. One-pass parallel thinning: analysis, properties, and quantitative evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1129–1140, November 1992.

[19] S. Yang and I.M. Verbauwhede. A secure fingerprint matching technique. In *Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications*, pages 89–94, Berkley, California, 2003.

[20] Xudong Jiang; Wei-Yun Yau. Fingerprint minutiae matching based on the local and global structures. In *Proceedings of the 15th International Conference on Pattern Recognition*, volume 2, pages 1038–1041, Barcelona, Catalonia, Spain, September 2000.

[21] D.A. Socolinsky and A. Selinger. Thermal face recognition over time. In *Proceedings of the 17th International Conference on Pattern Recognition*, volume 4, pages 23–26, August 2004.