Multi-Spectral Face Recognition - Fusion of Visual Imagery with Physiological Information

Pradeep Buddharaju and Ioannis Pavlidis

Dept. of Computer Science, University of Houston
4800 Callhoun Road, Houston, TX 77204
braju@cs.uh.edu, ipavlidi@central.uh.edu

Summary. We present a novel multi-spectral approach for face recognition using visual imagery as well as the physiological information extracted from thermal facial imagery. The main point of this line of research is that physiological information available only in thermal infrared, can improve the performance and enhance the capabilities of standard visual face recognition methods. For each subject in the database, we store facial images collected simultaneously in the visual and thermal bands. For each of the thermal images, we first delineate the human face from the background using the Bayesian framework. We then extract the blood vessels present on the segmented facial tissue using image morphology. The extracted vascular network produces contour shapes that are unique to each individual. The branching points of the skeletonized vascular network, referred to as thermal minutia points (TMPs), are an effective feature abstraction. During the classification stage, we match the local and global structures of TMPs extracted from the test image with those of the corresponding images in the database. We fuse the recognition results of our thermal imaging algorithm with those of a popular visual imaging algorithm. We have conducted experiments on a large database of co-registered visual and thermal facial images. The good experimental results show that the proposed fusion 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 is totally unobtrusive. At the same time, however, it is one of the most challenging modalities [2]. Research into face recognition has been biased towards the visual spectrum for a variety of reasons. Among those is the availability

and low cost of visual 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 proven more problematic than the seemingly effortless face recognition performed by humans. The major culprit is light variability, which is prevalent in the visual spectrum due to the reflective nature of incident light in this band. Secondary problems are associated with the difficulty of detecting facial disguises [3].

As a solution 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 visual band, which can be classified as appearance-based [9, 10] and feature-based approaches [11, 12].

Recently, attempts have been made to fuse the visual and thermal infrared modalities to increase the performance of face recognition [13, 14, 15, 16, 17, 18]. However, almost all these approaches use similar algorithms for extracting features from both visual and thermal infrared images. In this chapter, we present a novel approach to the problem of thermal facial recognition that realizes the full potential of the thermal infrared band. Our goal is to promote a different way of thinking in the area of face recognition in thermal infrared, which can be approached in a distinct manner when compared with other modalities. It consists of a statistical face segmentation and a physiological feature extraction algorithm tailored to thermal phenomenology. The use of vessel structure for human identification has been studied during recent years using traits such as hand vessel patterns [19, 20] and finger vessel patterns [21, 22]. 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 [23]. However, they did not present an algorithmic approach for achieving this. We present a full methodology to extract and match the vascular network from thermal facial imagery [24].

Figure 1 depicts the essence of the proposed multi-spectral face recognition methodology. The goal of face recognition is to match a query face image against a database of facial images to establish the identity of an individual. We collect both thermal and visual facial images of the subject whose identity needs to be tested. We extract the thermal minutia points (TMPs) from the thermal facial image and match them against TMPs of subjects already stored in the database. We then extract the principal components (eiganfaces) from the visual face image and project it to the face space constructed from visual database images. The eigenspace match score is fused with the TMP match score to produce the final match score.

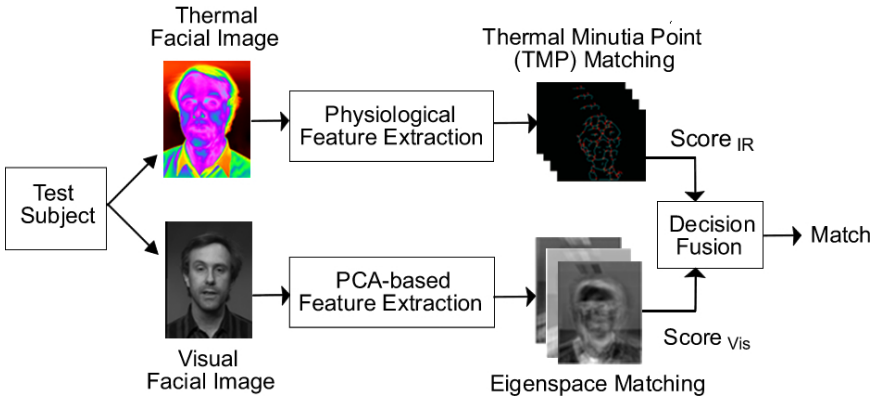


Fig. 1. Multi-spectral face recognition methodology.

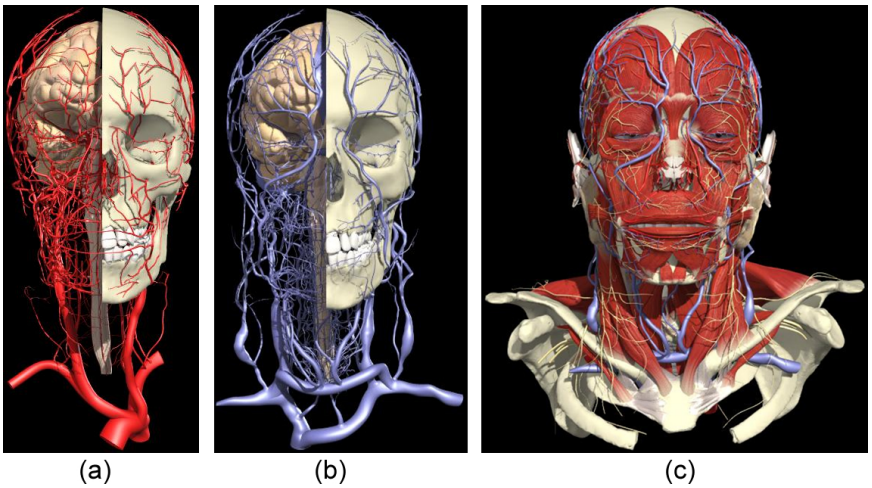


Fig. 2. Generic maps of the superficial blood vessels on the face - courtesy of Primal Pictures [25]: (a) Overview of arterial network. (b) Overview of venous network. (c) Arteries and veins together under the facial surface.

2 Physiological Feature Extraction from Thermal Images

A thermal infrared camera with good sensitivity (NEDT $> 0.025^0C$) provides the ability to directly image superficial blood vessels on the human face [26]. The pattern of the underlying blood vessels (see Figure 2) is characteristic to each individual, and the extraction of this vascular network can provide the basis for a feature vector. Figure 3 outlines the architecture of the feature extraction algorithm.

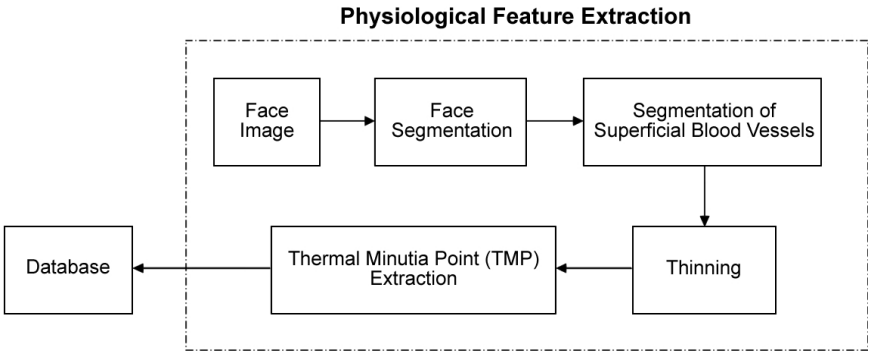


Fig. 3. Architecture of physiological feature extraction algorithm.

2.1 Face Segmentation

Due to its physiology, a human face consists of 'hot' parts 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 temperature distribution entity, which can be modeled using a mixture of two Normal distributions. Similarly, the background can be described by a bimodal temperature distribution with walls being the 'cold' objects and the upper part of the subject's body dressed in cloths being the 'hot' object. Figure 4(b) shows the temperature distributions of the facial skin and the background from a typical thermal facial image. We approach the problem of delineating facial tissue from the background using a Bayesian framework [24, 27], because we have apriori knowledge of the bimodal nature of the scene.

We call θ the parameter of interest, which takes two possible values (skin s or background b) with some probability. For each pixel x in the image at time t, 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:

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

We develop the statistics only for skin and then the statistics for the background can easily be inferred from Equation (1).

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 skin distribution and $f(x_t|s)$ is the likelihood for pixel x representing skin at time t. In the first frame ($t = 1$) the prior distributions for skin and background are considered equiprobable:

$$\pi^{(1)}(s) = \frac{1}{2} = \pi^{(1)}(b). \tag{3}$$

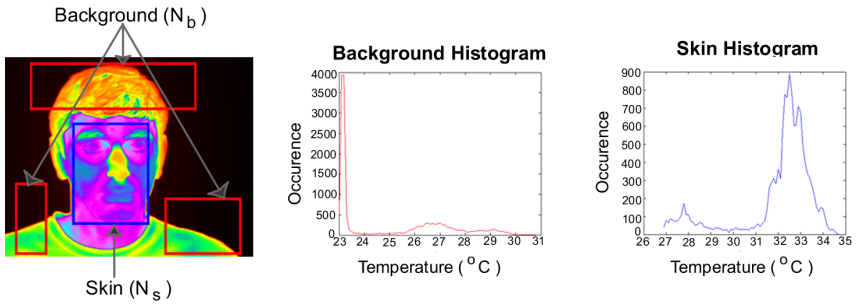
For $t > 1$, the prior skin distribution $\pi^{(t)}(s)$ at time t is equal to the posterior skin distribution at time $t-1$:

$$\pi^{(t)}(s) = p^{(t-1)}(s|x_{t-1}). \tag{4}$$

The likelihood $f(x_t|s)$ of pixel x representing skin at time $t \geq 1$ is given by:

$$f(x_t|s) = \sum_{i=1}^{2} w_{s_i}^{(t)} N(\mu_{s_i}^{(t)}, \sigma_{s_i}^{2(t)}), \tag{5}$$

where the mixture parameters $w_{s_i}(\text{weight}), \mu_{s_i}(\text{mean}), \sigma_{s_i}^2(\text{variance}) : \ i = 1, 2$ and $w_{s_2} = 1 - w_{s_1}$ of the bi-modal skin distribution can be initialized and updated using the EM algorithm. For that, we select N representative facial frames (off-line) from a variety of subjects that we call the training-set. Then, we manually segment, for each of the N frames, skin (and background) areas, which yields N_s skin (and N_b background) pixels as shown in Figure 4(a).



(a) (b) (c)

Fig. 4. Skin and Background: (a) Selection of samples for EM algorithm. (b) Corresponding bi-modal temperature distribution of background region. (c) Corresponding bi-modal temperature distribution of skin region.

To estimate the mixture parameters for the skin, we initially provide the EM algorithm with some crude estimates of the parameters of interest: $w_{s_0}, \mu_{s_0}, \sigma_{s_0}^2$. Then, we apply the following loop for $k = 0, 1, ..$:

$$z_{ij}^{(k)} = \frac{w_{s_i}^{(k)} (\sigma_{s_i}^{(k)})^{-1} \exp\left\{ -\frac{1}{2(\sigma_{s_i}^{(k)})^2} (x_j - \mu_{s_i}^{(k)})^2 \right\}}{\sum_{t=1}^{2} w_{s_t}^{(k)} (\sigma_{s_t}^{(k)})^{-1} \exp\left\{ -\frac{1}{2(\sigma_{s_t}^{(k)})^2} (x_j - \mu_{s_t}^{(k)})^2 \right\}},$$

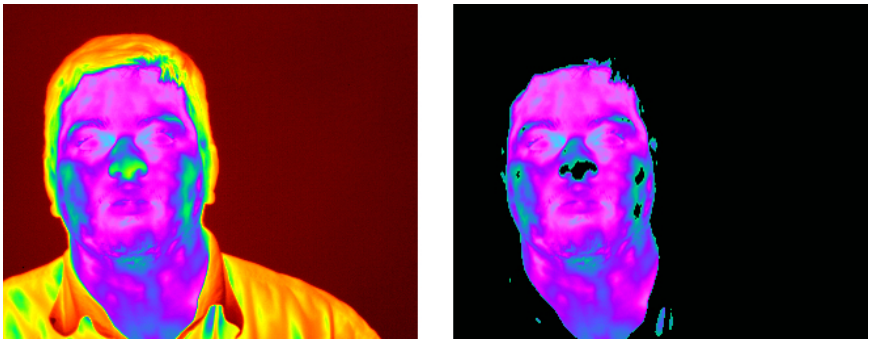
$$w_{s_i}^{(k+1)} = \frac{\sum_{j=1}^{N_s} z_{ij}^{(k)}}{N_s},$$

$$\mu_{s_i}^{(k+1)} = \frac{\sum_{j=1}^{N_s} z_{ij}^{(k)} x_j}{N_s w_{s_i}^{(k+1)}},$$

$$(\sigma_{s_i}^{(k+1)})^2 = \frac{\sum_{j=1}^{N_s} z_{ij}^{(k)} (x_j - \mu_{s_i}^{(k+1)})^2}{N_s w_{s_i}^{(k+1)}},$$

where $i = 1, 2$ and $j = 1, ..., N_s$. Then, we set $k = k + 1$ and repeat the loop. The condition for terminating the loop is:

$$|w_{s_i}^{(k+1)} - w_{s_i}^{(k)}| < \epsilon, \ i = 1, 2. \tag{6}$$



(a) (b)

Fig. 5. Segmentation of facial skin region: (a) Original thermal facial image. (b) Result of Bayesian segmentation.

We apply a similar EM process for determining the initial parameters of the background distributions. 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 5(b) depicts the visualization of Bayesian segmentation on the subject shown in Figure 5(a). Part of the subject's nose has been erroneously classified as background and a couple of cloth patches from the subject's shirt have been erroneously marked as facial skin. This is due to occasional overlapping between portions of the skin and background distributions. The isolated nature of these mislabeled patches makes them easily correctable through post-processing. We apply our three-step post-processing algorithm on the binary segmented image. Using foreground (and background) correction, we find the mislabeled pixels in foreground (and background) and remove them. The specific algorithm that achieves this is the following:

i. Label all the regions in the foreground and background using a simple flood-fill or connected component labeling algorithm [28]. Let the foreground regions be $R_f(i), i = 1, \ldots, N_f$, where N_f represents the number of foreground regions, and let the background regions be $R_b(j), j = 1, \ldots, N_b$, where N_b represents the number of background regions.
ii. Compute the number of pixels in each of the foreground and background regions. Find the maximum foreground (R_f^{max}) and background (R_b^{max}) areas:

$$R_f^{max} = \max\{R_f(i), i = 1, \ldots, N_f\},$$
$$R_b^{max} = \max\{R_b(i), i = 1, \ldots, N_b\}.$$

iii. Change all foreground regions that satisfy the condition $R_f(i) < R_f^{max}/4$ to background. Similarly, change all background regions that satisfy the condition $R_b(i) < R_b^{max}/4$ to foreground. We found experimentally that outliers tend to have an area smaller than one fourth of the maximum area, and hence can be corrected with the above conditions. Figure 6 shows the result of our post-processing algorithm.

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 [26, 27]:

i. The image is processed to reduce noise and enhance edges.
ii. Morphological operations are applied 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:

$$\frac{\partial I(\bar{x}, t)}{\partial t} = \nabla(c(\bar{x}, t)\nabla I(\bar{x}, t)). \tag{7}$$

In our case $\mathrm{I}(\bar{x}, t)$ is the thermal infrared image, $\bar{x}$ refers to the spatial dimensions, and t to time. $\mathrm{c}(\bar{x}, t)$ is called the diffusion function. The discrete version of the anisotropic diffusion filter of Equation (7) is as follows:

$$\begin{aligned} I_{t+1}(x, y) = I_t + \frac{1}{4} * [&c_{N,t}(x, y)\nabla I_{N,t}(x, y) \\ &+ c_{S,t}(x, y)\nabla I_{S,t}(x, y) + c_{E,t}(x, y)\nabla I_{E,t}(x, y) \\ &+ c_{W,t}(x, y)\nabla I_{W,t}(x, y)]. \end{aligned} \tag{8}$$

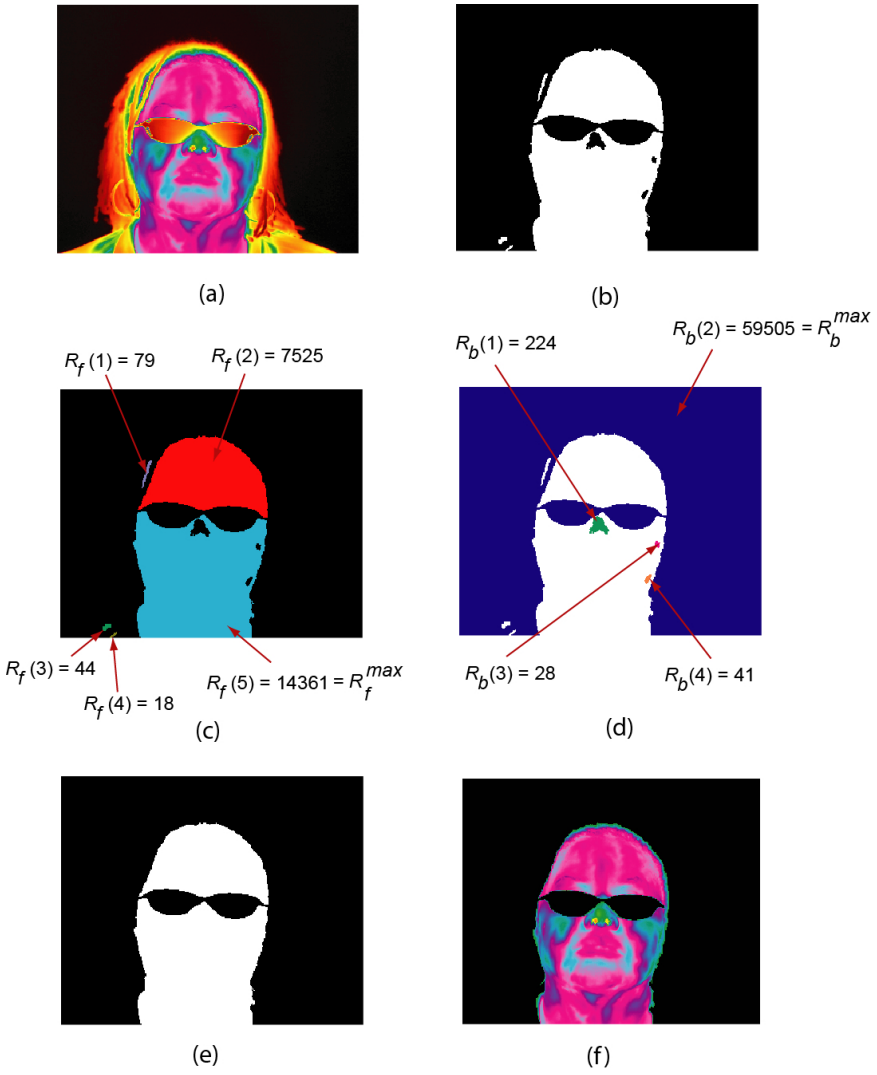


Fig. 6. Segmentation of facial skin region: (a) Original thermal facial image. (b) Binary segmented image. (c) Foreground regions each represented in different color. (d) Background regions each represented in different color. (e) Binary mask after foreground and background corrections. (f) Final segmentation result after post-processing.

The four diffusion coefficients and four gradients in Equation (8) correspond to four directions (i.e., North, South, East & West) with respect to the location (x,y). Each diffusion coefficient and the corresponding gradient are calculated in the same manner. For example, the coefficient along the north direction is calculated as follows:

$$c_{N,t}(x,y) = exp(\frac{-\nabla I^2_{N,t}(x,y)}{k^2}), \tag{9}$$

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 at a relatively low contrast compared to that of 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 takes two forms. First form is the white top hat segmentation that enhances the bright objects in the image, while the second one is the black top hat segmentation that enhances dark objects. In our case, we are interested in the white top hat segmentation because it helps with enhancing the bright ('hot') ridge like structures corresponding to the blood vessels. In this method the original image is first opened and then this opened image is subtracted from the original image as follows:

$$\begin{aligned} I_{open} &= (I \ominus S) \oplus S, \\ I_{top} &= I - I_{open}, \end{aligned} \tag{10}$$

where I, I_{open}, I_{top} are the original, opened, and white top hat segmented images respectively, S is the structuring element, and $\ominus$, $\oplus$ are morphological erosion and dilation operations respectively. Figure 7(a) depicts the result of applying anisotropic diffusion to the segmented facial tissue shown in Figure 4(b), and the Figure 7(b) shows the corresponding blood vessels extracted using white top hat segmentation.

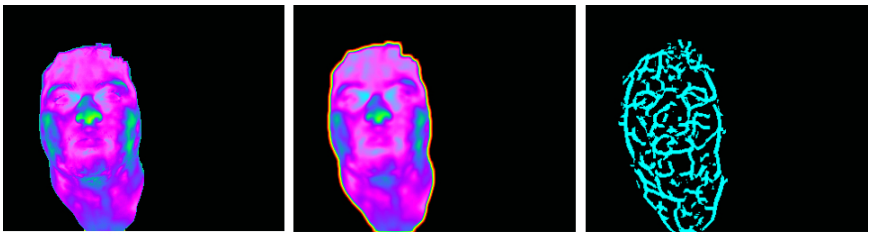


Fig. 7. Vascular network extraction: (a) Original segmented image. (b) Anisotropically diffused image. (c) Blood vessels extracted using white top hat segmentation.

2.3 Extraction of TMPs

The extracted blood vessels exhibit different contour shapes between subjects. We call the branching points of the blood vessels Thermal Minutia Points (TMPs). TMPs can be extracted from the blood vessel network in ways similar to those used for fingerprint minutia extraction. A number of methods have been proposed [29] for robust and efficient extraction of minutia from fingerprint images. Most of these approaches describe each minutia point by at least three attributes, including its type, its location in the fingerprint image, and the local vessel orientation. We adopt a similar approach for extracting TMPs from vascular networks, which is outlined in the following steps:

i. The local orientation of the vascular network is estimated.
ii. The vascular network is skeletonized.
iii. The TMPs are extracted from the thinned vascular network.
iv. The spurious TMPs are removed.

Local orientation $\Psi(x, y)$ is the angle formed at (x, y) between the blood vessel and the horizontal axis. Estimating the orientation field at each pixel provides the basis for capturing the overall pattern of the vascular network. We use the approach proposed in [30] for computing the orientation image because it provides pixel-wise accuracy.

Next, the vascular network is thinned to one-pixel thickness [31]. Each pixel in the thinned map contains 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 $(\sum_{i=0}^{7} N_i) > 2$ (see Figure 8).

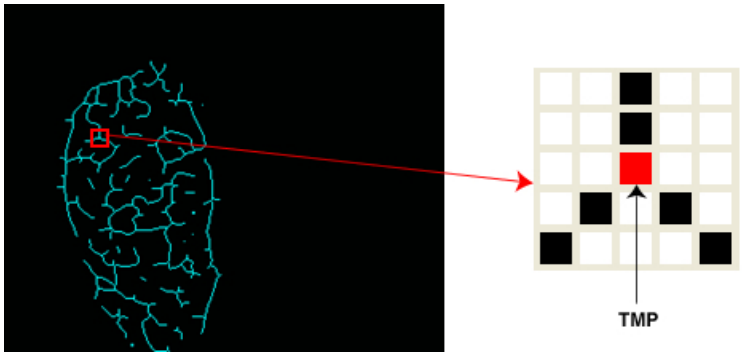


Fig. 8. Thermal Minutia Point (TMP) extracted from the thinned vascular network.

It is desirable that the TMP extraction algorithm does not leave any spurious TMPs since this will adversely affect the matching performance. Removal of clustered TMPs (see Figure 9(a)) and spikes (see Figure 9(b)) helps to reduce the number of spurious TMPs in the thinned vascular network.

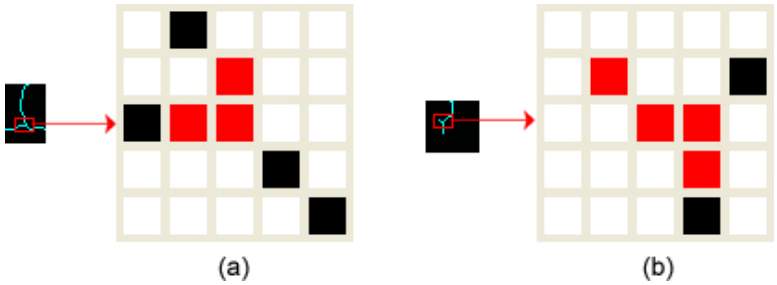


Fig. 9. Spurious TMPs: (a) Clustered TMPs. (b) Spike formed due to a short branch.

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 10 shows the results of each stage of the feature extraction algorithm on a thermal facial image.

2.4 Matching of TMPs

Numerous methods have been proposed for matching fingerprint minutiae, most of which try to simulate the way forensic experts compare fingerprints [29]. 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. Global minutiae matching algorithms feature high distinctiveness. A few hybrid approaches [32, 33] have been proposed where the advantages of both local and global methods are exploited. We use a hybrid method [32] to perform TMP matching.

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

$$L_M = \{\{d_1, \varphi_1, \vartheta_1\}, \{d_2, \varphi_2, \vartheta_2\}, ..., \{d_N, \varphi_N, \vartheta_N\}, \Psi_i\} \tag{11}$$

where

$$\begin{aligned} d_n &= \sqrt{(x_n - x_i)^2 + (y_n - y_i)^2} \\ \varphi_n &= diff(\Psi_n, \Psi_i),\ n = 1, 2, ..., N \\ \vartheta_n &= diff\left(\arctan\left(\frac{y_n - y_i}{x_n - x_i}\right), \Psi_i\right) \end{aligned} \tag{12}$$

The function $diff()$ calculates the difference of two angles and scales the result within the range $[0, 2\pi)$ [33]. Given a test image $\mathbf{I_t}$, the feature vector of each of its TMP is compared with the feature vector of each 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

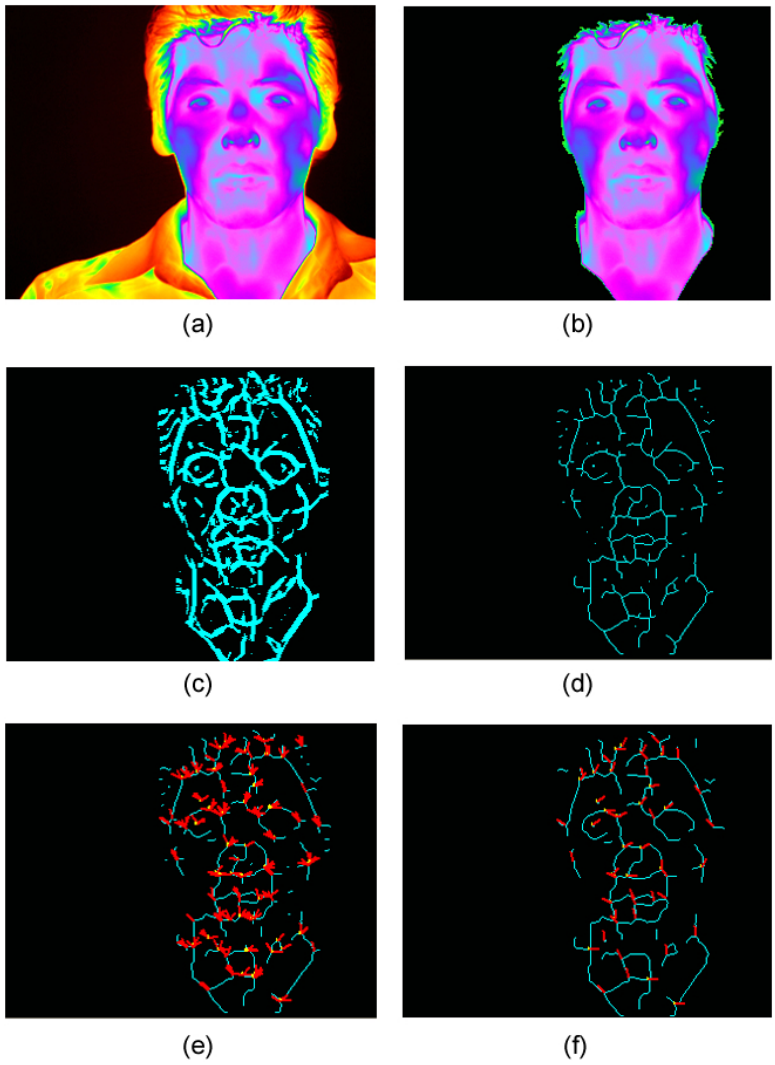


Fig. 10. Visualization of the various stages of the feature extraction algorithm: (a) A typical thermal facial image. (b) Facial tissue delineated from the background. (c) Vascular network extracted from thermal facial image. (e) Thinned vessel map. (f) Extracted TMPs from branching points. (g) Spurious TMPs removed.

$\{\delta_d, \delta_\varphi, \delta_\vartheta, \delta_\Psi\}$. The threshold values should be chosen in such a way that they accommodate linear deformations and translations. The final matching score between the test image and a database image is given by:

$$Score = \frac{NUM_{match}}{\max(NUM_{test}, NUM_{database})} \tag{13}$$

where NUM_{match} represents number of matched TMP pairs, and NUM_{test}, $NUM_{database}$ represent number of TMPs in test and database images respectively.

3 PCA-based Feature Extraction from Visual Images

Principal Component Analysis (PCA) is a well known approach for dimensionality reduction of the feature space. It has been successfully applied in face recognition [9, 34]. The main idea is to decompose face images into a small set of feature images called eigenfaces, which can be considered as points in a linear subspace called "face space" or "eigenspace". Recognition is performed by projecting a new face image into this eigenspace and then comparing its position with those of known faces.

Fig. 11. Eigenfaces extracted from our training set that correspond to decreasing order of eigenvalues.

Suppose a face image consists of N pixels, so it can be represented by a vector Γ of dimension N. Let $\{\Gamma_i | i = 1, \ldots, M\}$ be the training set of face images. The average face of these M images is given by

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i. \tag{14}$$

Then, each face Γ_i differs from the average face Ψ by Φ_i:

$$\Phi_i = \Gamma_i - \Psi; i = 1, \ldots, M. \tag{15}$$

A covariance matrix of the training images can be constructed as follows:

$$C = AA^T, \tag{16}$$

where $A = [\Phi_1, \ldots, \Phi_M]$. The top M' eigenvectors $U = [u_1, \ldots, u_{M'}]$ of the covariance matrix A, called eigenfaces, constitute the eiganspace. Figure 11 shows the top six eigenfaces extracted from our training set in decreasing order. Given a test image, Γ_{test}, it is projected to the eigenspace and an Ω_{test} vector is obtained as follows:

$$\Omega_{test} = U^T(\Gamma_{test} - \Psi). \tag{17}$$

The distances between this vector and the projected vectors from the training images are used as a measure to find the best match in the database. Any standard distance measure such as Euclidean distance, Mahalanobis distance, or MahCosine measure can be used to compare the vectors [17].

4 Experimental Results and Discussion

We used the Equinox Corporation's database in our experiments. It is a large database of both infrared (short-, mid-, and long-wave) and visual band images available for public download at *equinoxsensors.com/products/HID.html*. Image frame sequences were acquired at 10 frames/sec while the subject was reciting the vowel sequence "a,e,i,o,u". The database also consists of subject images wearing glasses and with expressions of happiness, anger and surprise, which were used to account for variation in poses. In order to induce variability in visual band images, three different illumination conditions were used during acquisition - frontal, frontal-left, and frontal-right. For each subject in the database subset we used, images were acquired simultaneously from visual and mid-wave infrared cameras.

We used a total of 4552 co-registered visual and mid-wave infrared images for our experiments, which comprised of 45 different subjects. For each subject, we used 3 training images (one per each illumination condition). If the subject was wearing glasses, we included images with and without glasses in the training set. Figure 12 shows the training examples of two subjects from the database. For each test image, we applied the physiological face recognition algorithm on the thermal image and the PCA algorithm on its visual counterpart. Finally, we applied decision fusion by combining the scores from

Fig. 12. Sample training images of two subjects (one subject per row) in the database.

the visual and thermal infrared recognition algorithms. Specifically, the fusion was performed by combining the individual scores from each of the algorithms. We found from our experiments that the physiological algorithm on thermal images performed slightly better than the PCA algorithm on visual imagery. The rank 1 recognition rate for thermal images was 97.74%, whereas that of visible images was 96.19%. Since the mismatches in each of these experiments were disjoint, fusion yielded an increase in performance with a rank 1 recognition rate of 98.79%. Figure 13 shows the CMC curves for the visual, thermal, and fusion algorithms.

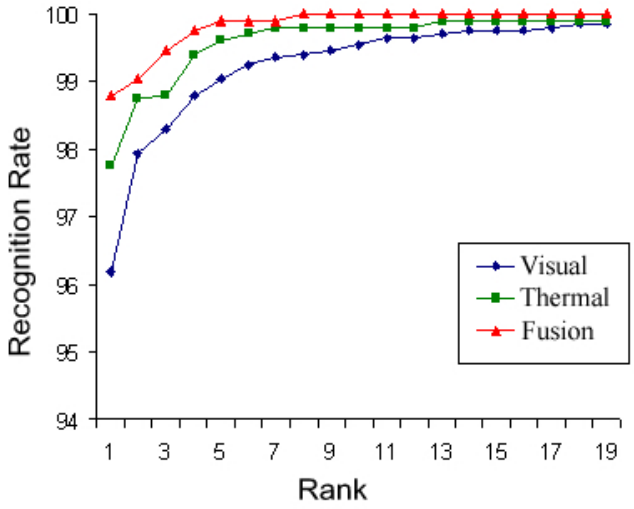


Fig. 13. CMC curves of the visual, thermal, and fusion face recognition algorithms.

We noticed that the recognition performance from the physiological algorithm on thermal images can be improved by estimating and eliminating the incorrect TMPs as well as non-linear deformations in the extracted vascular network caused due to large facial expressions and non-linear pose transformations. Figure 14 shows an example of non-linear deformations caused in the vascular network between gallery and probe images of the same subject due to pose and facial expression changes. Even though the matching algorithm described in Section 2.4 works fine with linear transformations in the vascular network, it affords small latitude in the case of non-linear transformations. Our future work is directed towards addressing this issue.

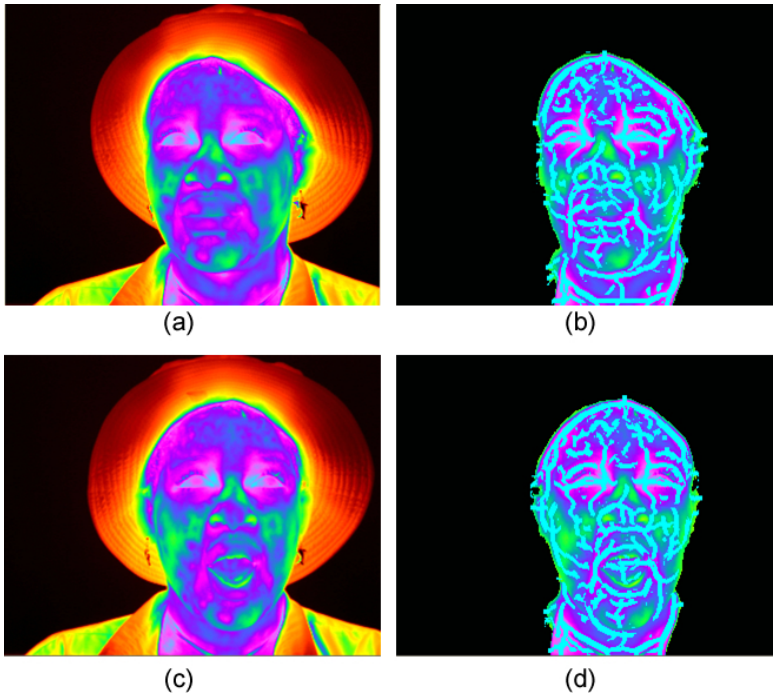


Fig. 14. (a) Training image and (b) corresponding vascular network (overlaid over the segmented image). (c) Test image of same subject exhibiting large facial expression and (d) corresponding vascular network (overlaid over the segmented image).

There are two major operational limitations in the current physiological feature extraction method:

i. Glasses are opaque in the thermal infrared spectrum and hence block important vascular information around eyes. Also, facial hair curtails the radiation emitted from the covered surface of the skin, and may cause

facial segmentation to break down. Figure 15 shows examples of failed face segmentation when glasses and facial hair are present.

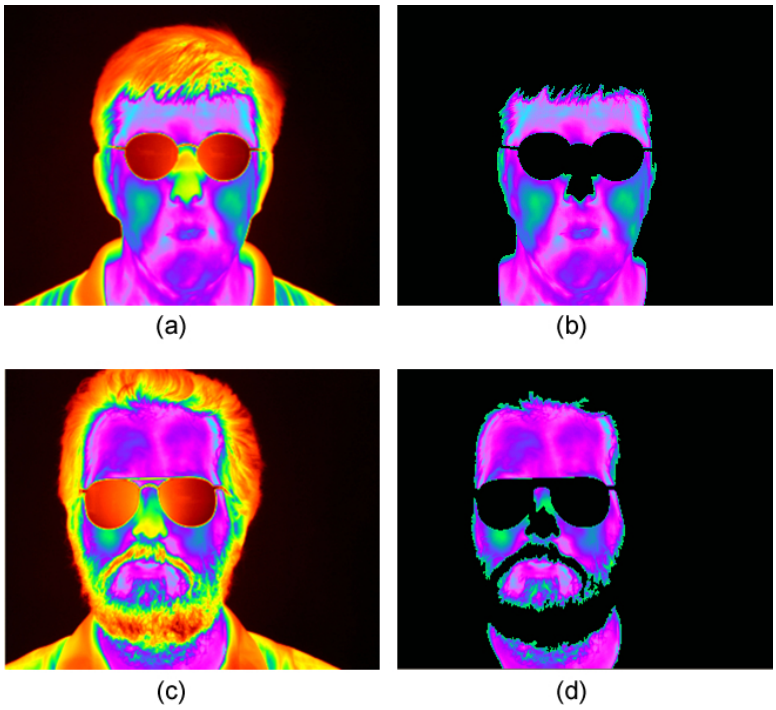


Fig. 15. (a) Thermal facial image with glasses and (b) result of segmentation. (c) Thermal facial image with facial hair and glasses and (d) result of segmentation.

ii. The robustness of the method degrades when there is substantial perspiration. This results in a highly non-linear shift of the thermal map that alters radically the radiation profile of the face. A practical scenario where such a case may arise is when a subject is imaged after a strenuous exercise that lasted several minutes. Another such scenario may arise when a heavily dressed subject is imaged in a very hot environment.
We have performed an experiment whereby a subject is imaged at the following instances:
- In a baseline condition (Figure 16, image 1a)
- After 1 min of rigorous walking (Figure 16, image 2a)
- After 5 min of rigorous walking (Figure 16, image 3a)
- After 5 min of rigorous jogging (Figure 16, image 4a)

Column b of Figure 16 shows the corresponding vessel extraction results. In the case of image 2a, the metabolic rate of the subject shifted to higher gear, but perspiration is still not a major problem. One can find evidence

of the higher metabolic rate by looking at the left temporal area, where the region around the rich vasculature has become deeper cyan (hotter) in image 2a with respect to image 1a. This is an example of a positive linear shift (warming up), which the vessel extraction algorithm handles quite well (see image 2b versus image 1b). As the exercise become more strenuous and lasts longer, perspiration increases and introduces a negative non-linear shift (cooling down) in the thermal map. This is especially pronounced in the forehead where most of the perspiration pores are. Due to this, some unwanted noise starts creeping in image 3b, which becomes more dramatic in image 4b. The performance of the vessel extraction algorithm deteriorates but not uniformly. For example, the vessel extraction algorithm continues to perform quite well in the cheeks where perspiration pores are sparse and the cooling down effect is not heavily non-linear. In contrast, performance is a lot worse in the forehead area, where some spurious vessel contours are introduced due to severe non-linearity in the thermal map shift.

5 Conclusions

We have outlined a novel multi-spectral approach to the problem of face recognition by the fusion of thermal infrared and visual band images. The cornerstone of the approach is the use of unique and time invariant physiological information as feature space for recognition in thermal imagery. 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 method although young, performed well on a nontrivial database. We also applied a PCA-based (eigenface) recognition approach on concomitant visual imagery. We have shown that the recognition performance in the thermal domain is slightly better than the visual domain, and that fusion of the two modalities/methodologies is better than either one of them. In a nutshell, this research demonstrated that standard visual face recognition methods can gain in performance if they are combined with physiological information, uniquely extracted in thermal infrared. The most notable outcome besides performance increase is the striking complimentarily of the two modalities/methodolgogies as it is revealed in the experimental results. It is the latter that renders fusion a natural strategy that fits the problem.

References

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

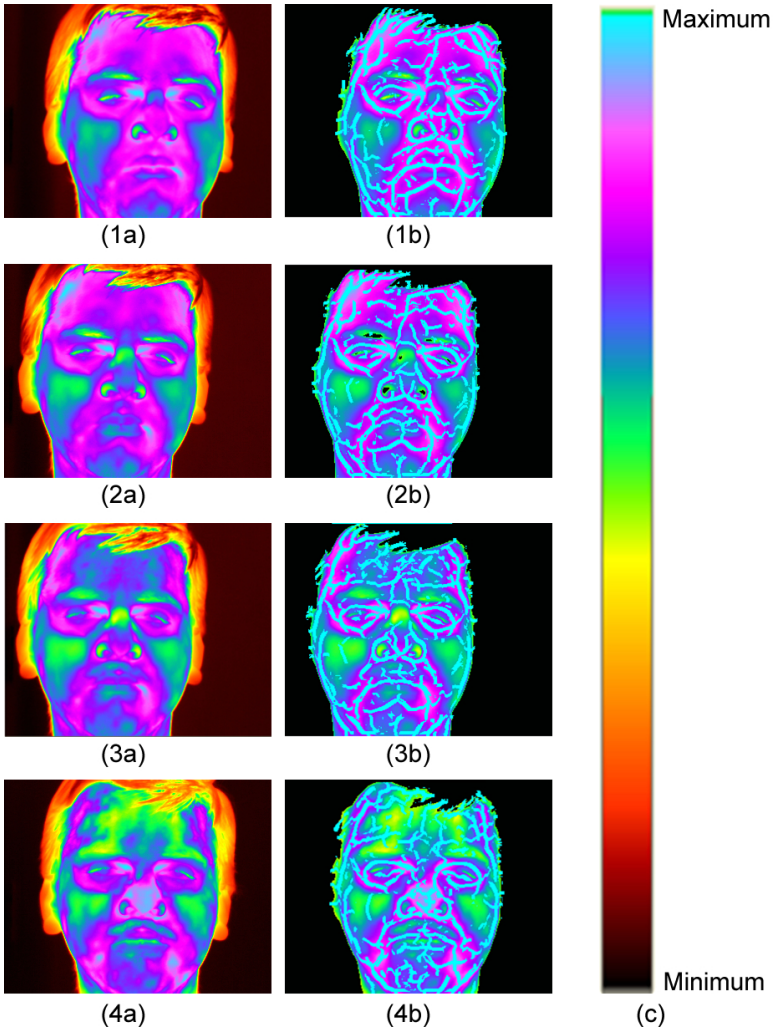


Fig. 16. Effect of perspiration on feature extraction. Thermal facial image of a subject (1a) at rest, (2a) after 1 minute of rigorous walking (3a) after 5 min of rigorous walking, (4a) after 5 min of rigorous jogging, and (1b,2b,3b,4b) corresponding vascular network maps, and (c) color map used to visualize temperature values.

2. Zhao W., Chellappa R., Phillips P. J., and Rosenfeld A. Face recognition: A literature survey. *ACM Computing Surveys (CSUR)*, 35(4):399–458, December 2003.
3. Pavlidis I. and Symosek P. 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. Prokoski F. 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. Socolinsky D.A. and Selinger A. 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. Wilder J., Phillips P.J., Jiang C., and Wiener S. 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. Socolinsky D.A., Wolff L.B., Neuheisel J.D., and Eveland C.K. 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. Selinger A. and Socolinsky D.A. 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. Cutler R. Face recognition using infrared images and eigenfaces. cs.umd.edu/rgc/face/face.htm, 1996.
10. Chen X., Flynn P.J., and Bowyer K.W. 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. Srivastava A. and Liu X. Statistical hypothesis pruning for recognizing faces from infrared images. *Journal of Image and Vision Computing*, 21(7):651–661, 2003.
12. Buddharaju P., Pavlidis I., and Kakadiaris I. 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. Heo J., Kong S.G., Abidi B.R., and Abidi M.A. Fusion of visual and thermal signatures with eyeglass removal for robust face recognition. In *Proceedings of the Joint IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum*, Washington D.C., June 2004.
14. Gyaourova A., Bebis G., and Pavlidis I. Fusion of infrared and visible images for face recognition. In *Proceedings of the 8th European Conference on Computer Vision*, Prague, Czech Republic, May 2004.
15. Socolinsky D.A. and Selinger A. Thermal face recognition in an operational scenario. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1012–1019, Washington, D.C, June 2004.

16. Wang J.G., Sung E., and Venkateswarlu R. Registration of infrared and visible-spectrum imagery for face recognition. In *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, pages 638–644, Seoul, Korea, May 2004.
17. Chen X., Flynn P., and Bowyer K. Ir and visible light face recognition. *Computer Vision and Image Understanding*, 99(3):332–358, September 2005.
18. Kong S.G., Heo J., Abidi B.R., Paik J., and Abidi M.A. Recent advances in visual and infrared face recognition–a review. *Computer Vision and Image Understanding*, 97(1):103–135, January 2005.
19. Lin C.L. and Fan K.C. Biometric verification using thermal images of palm-dorsa vein patterns. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(2):199–213, February 2004.
20. Im S.K., Choi H.S., and Kim S.W. A direction-based vascular pattern extraction algorithm for hand vascular pattern verification. *ETRI Journal*, 25(2):101–108, April 2003.
21. Shimooka T. and Shimizu K. Artificial immune system for personal identification with finger vein pattern. In *Proceedings of the 8th International Conference on Knowledge-Based Intelligent Information and Engineering Systems, Lecture Notes in Computer Science*, volume 3214, pages 511–518, September 2004.
22. Miura N., Nagasaka A., and Miyatake T. Feature extraction of finger vein patterns based on iterative line tracking and its application to personal identification. *Systems and Computers in Japan*, 35(7):61–71, April 2004.
23. Prokoski F.J. and Riedel R. *BIOMETRICS: Personal Identification in Networked Society*, chapter 9 Infrared Identification of Faces and Body Parts. Kluwer Academic Publishers, 1998.
24. Buddharaju P., Pavlidis I.T., and Tsiamyrtzis P. Physiology-based face recognition. In *Proceedings of the IEEE Advanced Video and Signal based Surveillance*, Como, Italy, September 2005.
25. B. J. Moxham, C. Kirsh, B. Berkovitz, G. Alusi, and T. Cheeseman. *Interactive Head and Neck (CD-ROM)*. Primal Pictures, December 2002.
26. Manohar C. Extraction of superficial vasculature in thermal imaging. Master's thesis, University of Houston, Houston, TX, December 2004.
27. Pavlidis I., Tsiamyrtzis P., Manohar C., and Buddharaju P. *Biomedical Engineering Handbook*, chapter Biometrics: Face Recognition in Thermal Infrared. CRC Press, February 2006.
28. Di Stefano L. and Bulgarelli A. Simple and efficient connected components labeling algorithm. In *Proceedings of the International Conference on Image Analysis and Processing*, pages 322–327, Venice, Italy, September 1999.
29. Maltoni D., Maio D., Jain A.K., and Prabhakar S. *Handbook of Fingerprint Recognition*. Springer Verlag, June 2003.
30. Oliveira M.A. and Leite N.J. 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.
31. Jang B.K. and Chin R.T. One-pass parallel thinning: analysis, properties, and quantitative evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1129–1140, November 1992.
32. Yang S. and Verbauwhede I.M. A secure fingerprint matching technique. In *Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications*, pages 89–94, Berkley, California, 2003.

33. Jiang X. and Yau W.Y. 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.
34. Turk M. and Pentland A.P. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1):71–86, 1991.

Index