4

Face Recognition under the Skin
Pradeep Buddhuraja and Ioannis Pavlidis

4.1 Introduction

Face recognition stands as the most appealing biometric modality, since it is the natural mode of identification among humans and is basically inalienable. At the same time, however, it is one of the most challenging modality (Zhao et al. 2003). Several face recognition algorithms have been developed over years, mostly in the visible and a few in the infrared domain. A serious problem in visible face recognition is light variability, due to the real-time nature of visible light in this band. This can clearly be seen in Figure 4.1. The visible image of the same person in Figure 4.1(a) acquired in the presence of normal light appears totally different from that in Figure 4.1(b), which was acquired in low light.

Most of the research efforts in thermal face recognition were mainly aiming to combat the dark or reduce the detrimental effect of light variability (Zhao et al., 2003; Schonlau and Schonlau, 2000). Methodologically, such approaches did not differ very much from face recognition algorithms in the visible band, which can be classified into appearance-based (Zhu et al., 2002) and feature-based (Buddhuraja et al., 2006). Recently attempts have been made to fuse the visible and infrared modalities to improve the performance of face recognition (Buddhuraja and Schonlau, 2006; Weng et al., 2004; Chen et al., 2003; Kang et al., 2003).

The infrared have presented a physiological facial recognition method that promises a different way of thinking about face recognition in the thermal infrared (Buddhuraja et al., 2006; 2007; Buddhuraja and Pavlidis, 2008). This work has shown that facial physiological information, contained in the level of a superficial vascular network, can serve as a good and time-invariant feature vector for face recognition. However, the methodology in this pilot work had some weak points. The recognition performance reported in past experiments...
Figure 4.1. Example showing illumination effects on visible and thermal infrared images. All the input images were acquired from the same subject in the same time. Left: Visible image captured by a VIS/NIR camera in 800 nm. Right: Thermal infrared image captured by a thermal infrared imagers in 9-11 μm.

Buckland et al. 2004 can be substantially improved by curing these weaknesses. This chapter presents an advanced methodological framework that relies on maximum physiological face recognition from feasible to visible. Specifically, the main contributions in the chapter are:

- A new vessel segmentation post-processing algorithm that removes fake vascular contours detected by the top-hat vessel segmentation method.
- A new vascular network matching algorithm that is robust to nonlinear deformations due to facial pose and expression variations.
- Experiments and comparative experiments to evaluate the performance of the proposed methods with respect to previous methods.

The rest of the chapter is organized as follows. Section 4.2 presents an overview of the new methodology. Section 4.3 presents in detail the vessel segmentation post-processing algorithm. Section 4.4 discusses the new matching...
network matching algorithm. Section 4.5 presents the experimental results and
reaches a critical evaluation. The chapter concludes in Section 4.6.

4.2 Methodology

Figure 4.2 shows the methodological architecture. The method operates in the
following three models:

Offline Mode. The thermal facial images are captured by a mid-wave infrared
(MWIR) camera. For each subject to be registered in the database, a
feature extraction algorithm extracts the feature vectors from the facial images
and links it to the subject's record. The feature extraction algorithm has two
steps:

First, a Bayesian face representation separates facial tissue from background.
Second, face representation post-processing extracts four segmentation areas
which are due to occasional overlapping between portions of the tissue and
background disturbances. Third, a non-linear vessel segmentation algorithm
extracts the vascular network from the four segments after minimizing
different change of fuzzy edges, then to best deliver. These three steps have been
described in Buddhajiva et al. (2007) and are briefly presented in this section.
Fourth, a new vessel segmentation post-processing algorithm, which is one of
the chapter's contributions, occurs vessel segmentation. The vessel segment
occasionally is fooled by areas of skin contrast (i.e., hairline and skin edges)
and reports them as vascular contours. These fake vascular contours participate
in the matching process with deleterious effects.

Online Mode. Given a query image, its vascular network is presented using
the feature extraction algorithm outlined in the offline mode, and it is matched
against vascular networks stored in the database. The new matching algorithm, which is another of this chapter's contributions, has two stages:

First, after more conventional algorithms are used to detect the most salient features, a measure of correspondence is required to calculate the vascular network attenuation between input and database images. Second, the dual-source maximum likelihood matching algorithm matches the test and database vascular networks. The matching score between the two depends on the amount of overlapping.

4.2.1 Fine Segmentation

Because of its physiology, human face consists of "soft" parts that correspond to areas rich in vasculature and "hard" parts that correspond to areas rich with aponeurotic tissue. Most cases the human face is a bimodal mixture distribution entity, which can be modeled using a mixture of two Normal distributions. Similarly, the background can be described by a bimodal mixture distribution with walls being the "solid" object and the upper part of the object's body dressed in nothing being the "soft" object. The consistency of simplicity means algorithm and image transformation is identical. We approach the problem of detecting facial image from background using a Bayesian framework since we have a prior knowledge of the essential nature of the work.

We call it the parameter of intensity, which takes two possible values value 0 or 1 with the same probability. For each pixel $x$ in the image at time $t$, we know an indicator of whether it represents skin ($x \in A$) or background ($x \in B$). Let $y_0$ be the posterior distribution $P(y_0 | x_t)$ given.

$$P(y_0 | x_t) = \begin{cases} P(y_0 = 1 | x_t), & \text{when} \quad y = 0, \\ P(y_0 = 0 | x_t), & \text{when} \quad y = 1, \end{cases} \quad (4.1)$$

We develop the assumption only for skin and then the statistics for the background can easily be inferred from Eq. (4.1).

According to Bayes' theorem,

$$P(y_0 | x_t) = \frac{P(y_0 = 1 | x_t) P(x_t | y_0 = 1) + P(y_0 = 0 | x_t) P(x_t | y_0 = 0)}{P(x_t | y_0 = 1) + P(x_t | y_0 = 0)}, \quad (4.2)$$

Here $\pi^*(y_0)$ is the prior skin distribution and $f(x_t | y_0)$ is the likelihood for pixel $x_t$ representing skin at time $t$. In the first frame ($t = 1$) the prior distributions for skin and background are considered equiprobable:

$$\pi^*(y_0) = \frac{1}{2}, \quad \pi^*(y_0). \quad (4.3)$$
The likelihood of the model is defined as follows for time $t$.

$$L(\theta, \phi) = \sum_{i=1}^{n} \log p(x_i \mid \theta, \phi).$$

Next, we define a parameter $\alpha$ to represent the number of bins, and $n_i$ represents the number of elements in the $i$th bin.

4.3.2 Blood Vessel Segmentation

First, we define a model to describe the phenomenon of blood vessel segmentation in the field of image processing in the biological setting.

Next, we process the image to enhance noise and enhance the edges.

Applying morphological processing to localize the vessel-like structures.

This will yield objects that are diffused in the local space. The blood vessels in the image are then segmented by performing thinning and enhancement of object boundaries using morphological operators.

The mathematical expression that describes this process is

$$\frac{d^2x(t)}{dt^2} = \text{vector field}.$$
In eq. (4.1), it is the that an "image" refers to the spatial distribution of color and intensity, and Q(t) is called the diffusion term. The diffusion process is described by the diffusion equation, which is given by

\[ \frac{\partial I}{\partial t} = \frac{1}{4} \nabla^2 I + Q(t) \]

or

\[ \nabla^2 I - \frac{1}{4} \nabla^2 I = -Q(t) \]

The four constant coefficients and four variables in Eq. (4.1) are, respectively, the diffusion coefficient D in the x-direction, the diffusion coefficient D in the y-direction, the diffusion coefficient D in the z-direction, and the diffusion coefficient D in the t-direction, respectively. The corresponding problems are solved in the same manner. For example, the separation along the nth direction is calculated as follows

\[ c_n (r) = \exp \left( -\frac{r^2}{4D_n t_n} \right) \]  

where \( D_n = \frac{1}{4} \) (if \( n = 1, 2, 3 \)).

Image smoothing is also realized on the image intensity. The intensity at each point is smoothed by linear smoothing with a kernel size of 3x3, which is equivalent to the smoothing in the spatial domain. The linear smoothing is applied to the intensity image to reduce noise and accentuate edges. For instance, in a specific case where the smoothing is applied to the edge of a polygon, it helps to enhance the sharpness of the edge-like structures corresponding to the boundary of the polygon. In practice, the original image is first upsampled and then the upsampled image is smoothed to remove noise.

\[ I_{NS} = G_{NS} \otimes I \\
I_{US} = \frac{1}{3} (I_{NS} + I_{US} + I_{US}) \]

where \( I_{NS} \) and \( I_{US} \) are the upsampled and original images, respectively, and the smoothing is applied on the upsampled image on the plane of pixels, respectively.

### 4.3 Visual Segmentation Post-Processing

As a result of the visual segmentation process, an image is divided into segments, each of which is characterized by specific features such as color, texture, and shape. The segmentation process is often used to identify and separate different objects or regions within an image. This step is crucial for various applications, including image analysis, object recognition, and computer vision tasks. The post-processing stage typically involves refining the segmentations to improve their accuracy and completeness, as well as enhancing the overall quality of the segmented images. Techniques such as merging small segments, applying morphological operations, and smoothing boundaries are commonly used in this phase to achieve better results.

---

[Image-related content and equations would be added here, but not transcribed due to the nature of the task.]
The top hat segmentation algorithm is successful in localizing vessels because it iterates ten times from hot to cold tub. In some instances, such iterations are not due to the presence of vessels. Examples include those between the legs or between the fingers, and an instance from Figure 4.5a.

It is crucial to detect and remove these variations from the vessel network before applying a matching algorithm. To study the properties of the outliers, the authors selected 25 representative subjects from the University of Houston dataset and analyzed the segmentation errors. Specifically, the authors identified the location of true, false, and vessels. Then, they drew measurement lines across each of the vessels and measured and plotted the corresponding temperature profiles. They noticed that the difference between the minimum and maximum temperature values was much larger in outliers than in real vessels. Indeed, the problem in outliers is quite severe because high C02 as it is formed between facial hair or glabella and tissues. By contrast, in real vessels, this gradient is quite small (only a few units of degrees C02 in 1 h). Therefore, the projection of the vessel's location and surrounding areas (Figure 4.4). Figure 5 shows the difference between minimum and maximum temperatures (Tmin - Tmax). Table 3 shows all the selected line profiles from the 25 representative subjects.

The new segmentation post-processing algorithm minimizes outliers based on the above findings. Specifically, it relies on the following steps:

**Step 1:** Sketch out the vascular network in one grid dimension.

**Step 2:** Use a normal pixel discontinuity across each decision point and gather all the points covered by this pixel discontinuity.
4.3 Vascular Network Matching

4.4 Vascular Network Matching

The matching method presented in [Boudhaya et al., 2017] could not cope with discontinuities in the deformation of the vascular network, because of variations...
at initial pose and coarse registration. This chapter presents a novel vascular network matching algorithm that is robust to nonlinear deformations.

### 4.4.1 Registration of Vascular Networks

The aim is to register the vascular network of the test image with that of the database image, so that they can be aligned. The iterative Closest Point (ICP) algorithm has appealing properties for point-based registration (Besl and McKay, 1992). ICP requires proper initialization, as different initializations of the algorithm can lead to different combinations of image poses, shape metrics, and transformation models. In Strojnik et al. (2003), developed a variant of the ICP algorithm called dual bootstrap ICP, that works well when initialization provides just a "best guess" on the correct estimate of the transformation and successfully registers vessel-like structures, such as vasculature. Specifically, they reported good results from registration of medical vascular images in the visible band. Since superficial vasculature exhibits distinct thermal infrared morphological resemblance to actual vasculature in the visible, the authors adopted the dual bootstrap algorithm for the registration task of vasculature.

After successful registration of the test and database vascular images, the matching score is computed based on the number of overlapping vessel pixels. If \( I_{\text{test}} \) represents the test vascular image with \( N_{\text{test}} \) vessel pixels, and \( I_{\text{d}} \) represents the database vascular image with \( N_{\text{d}} \) vessel pixels, the matching score is

\[
\text{Score} = \frac{N_{\text{overlap}}}{\min(N_{\text{test}}, N_{\text{d}})} \times 100
\]

where \( N_{\text{overlap}} \) represents the number of vessel pixels in \( I_{\text{test}} \) with a corresponding vessel pixel in \( I_{\text{d}} \) within a certain distance. Figure 4.8 shows examples of the performance of the dual bootstrap ICP algorithm in registering vascular images in the thermal infrared. The example at the bottom row of the figure features substantial pose variation between...
the test and database images. The larger the pose variation, the more difficult it becomes for the dual invariant KLT to cope successfully. Thus, an approach for determining the pose of the test and database images and setting accordingly the threshold values (T10) need to accept or reject the test image. This paper presents an algorithm that was developed for this purpose in presented in the next section.

### 4.4.2 Face Pose Estimation

A neutral pose (0°) is the pose in the center of the face, that is, the position of the face at halfway between the left and right ends of the face. From the face segmentation algorithm presented in Budilingun et al. (2000), one can find the left and right ends of the face. Hence, if one localizes the nose, he or she can estimate the facial pose.

In a thermal infrared image, the nose is typically at a gradient with its surroundings, as shown in Figure 4.7. This is because of the nose's shape (hollowed out), its temperature (cooler), and the forced circulation of air due to breathing. The combination of all these creates a nasal thermal signature, but considerably thinner than that of the surrounding soft-tissue.
Figure 4.1: Nose edge detection from thermal facial image. The nose edges were detected using the Canny edge detection algorithm. (d) color map.

By using an edge detection algorithm, one can extract the nose area from the facial image. The next step is to search for the nose edge model in the edge detection map. The authors used a Tomosynthesis-based matching algorithm to localize the nose model in the face edge image. (Hummelschäfer et al. 1999)

Figure 4.1 shows performance examples of the face model estimation algorithm.

4.5 Experiments

The authors conducted several experiments to validate the performance of the new physiological face recognition method. This section presents the experimental setup and results in detail.

Figure 4.2: Nose area edge performance example (thermal facial image). (a) nose detection using headless head matching algorithm; (b) nose extension.
4.5.1 Experiments on the University of Houston Database

The authors collected a supplemental thermal frontal dataset for the purposes of this evaluation. This set, known as the University of Houston (UH) database, has thermal facial images of varying expressions and poses from 400 subjects. The images were captured using a high-quality mid-wave infrared (MWIR) camera.

4.5.1.1 Facial Expression Dataset

To test the performance of the new method in the presence of varying facial expressions between gallery and probe images, the authors created the facial expression dataset (FEDS) out of the UH database as follows: From each of the 400 subjects, one frontal thermal image at 0° pose and neutral expression was used as a gallery image. Then five different facial images at ±45° pose but with varying facial expressions were used as probe images for each subject. Hence, FEDS has a total of 400 gallery images and 1500 probe images from 400 subjects. Figure 4.9 shows a sample subject set from FEDS.
Figure 4.10. Experimental results on the FERID dataset using dual-bounce ICP versus TMI matching (a) CMC curve (b) ROC curve.

Figures 4.10(a) and 4.10(b) show, respectively, the Cumulative Match Characteristics (CMC) and Receiver Operating Characteristic (ROC) curves of the FERID experiments using the dual bounce point (TMI) matching algorithm reported in Badshah et al. (2005) (a) long running, (b) variety versus real running, (c) dual matching ICP matching algorithm.

The results demonstrate that the dual bounce ICP outperforms the TMI matching algorithm. In the case of TMI matching, the CMC curve shows that rank 1 recognition is 81% whereas for the dual bounce ICP is 99%. Also, the dual bounce ICP matching method achieves a high matching precision at very low FAR and the matching rates, as shown in Figure 4.10(b). This indicates that the ICP matching algorithm is highly robust to deformations caused in facial expression variations.

4.5.1.2 Facial Pose Database

To test the performance of the new method in the presence of varying pose between gallery and probe images, the authors created the facial pose dataset (FPDS) out of the UM dataset as follows. From each of the 210 subjects, 20 frontal facial images at five gaze and neutral expression were used as a gallery image. Then four different frontal images, at neutral facial expression and at varying poses between -90° and 90° were used as probe images. Hence, FPDS has a total of 840 gallery images and 1200 probe images from 210 subjects. Figure 4.11 shows a sample subject set from FPDS.

Figures 4.11(a) and 4.11(b) show the results of the FPDS experiments using the dual bounce ICP matching algorithm. The results demonstrate that the algorithm copes well with facial pose variations between gallery and probe images. Specifically, the CMC curve shows that rank 1 recognition is 99% and the ROC curve shows that it requires a false acceptance rate over 36% such a positive acceptance rate above the 90% range. One can notice that the false acceptance rate on FERID experiments is a bit higher than on FPDS.
experiments. This can be expected, as variations in pose typically cause some resilience deformations in the vascular network that are caused by variations in head exposure.

4.5.2 Experiments on the University of Notre Dame Database

A major challenge associated with thermal face recognition is recognition performance over time [11] [20]. In this section, we will present experimental results on the NTU database using the data Predictive Biometric Identification (PBI) system [11].

![Graphs showing results](image)

**Figure 4.12:** Experimental results on the NTU database using the data Predictive Biometric Identification (PBI) system. (a) S.E. curve.
they change depending on the physical condition of the person, which is to say, not to mention similar face and the same person over time. Previous face recognition methods in the neural network used visual facial configurations such as performance over time [Chen et al., 2018].

Therefore, physiological information, like using a camera to capture the features, and not having pitfalls regarding facial variance...

Since a vast majority of the subjects in the P3H database were collected during the winter season, it is statistically significant that the experiment was possible. For this reason, the subjects shown below are being used for the experiment that was conducted at the University of New York.

The database is divided into head of facial gallery and locations: P3H (left), K-A (right), and L-A (left). Each of these galleries (e.g., P3H-A) are used against the other three galleries (e.g., P3H-L-A, K-A, and L-A). This results in different possible pairs that are used for testing. The accuracy of the new method using the matching algorithm was compared to that of k-NN matching (k = 1), RST matching, and WTA matching [Chen et al., 2018].

The results demonstrate that a new matching algorithm yields better recognition results even in the presence of time and temperature variations, thus making the THF method more powerful than the traditional face recognition algorithms.

4.5.3 Experiments on the University of Arizona Database

The authors have also used data from stress experiments carried out at the University of Arizona (UA). The subject of the experiment acquired information from a few real subjects, thus demonstrating the power of the facial thermal map in times of stress. In fact, in a few minutes, much more accurate and precise information is gathered from these maps [Chen et al., 2018], which can be used to attempt the face recognition experiment in a similar way. In essence, face recognition, the collection of images, and the calculation of the facial features from the information thus obtained...
4.6 Conclusions

The results presented here indicate that substantially improved performance is achievable with the current network training method. The learning algorithm, based on the backpropagation method, is capable of producing networks with better generalization capabilities than those obtained through traditional training techniques. The improved performance is evident in both the training and testing phases, with lower error rates and faster convergence. This advancement opens new possibilities for the application of neural networks in various domains, including pattern recognition, function approximation, and data classification.
Figure 4.10. Sample subject from UA dataset: (a) front face image captured in a laboratory environment; (b) three images shown in a sequence; (c) thermal images at 17 seconds, 27 seconds, 75 seconds, and 17 minutes 45 seconds of the interview, respectively. (d) Thermal map improvement for shortening.

Figure 4.11. Result of improvements of the IA method using the proposed method with THP matching: (a) CMC curves; (b) ROC curves.
Acknowledgments

The authors would like to thank the University of Notre Dame for kindly providing a facial database. They also acknowledge Professor Hoon Park at the University of Arkansas for designing and conducting the visual experiments as part of a collaboration with the Defense Academy for Credibility Assessment (DACA). The views expressed in this chapter do not necessarily represent the views of the funding agencies.

References

[References are not provided in the image.]