29

Biometrics: Face Recognition in Thermal Infrared

I. Pavlidis, P. Buddharaju, and C. Manohar
University of Houston

P. Tsiamyrtzis
University of Athens

29.1 Introduction

Biometrics has received a lot of attention in 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 emedded 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]. Efforts were directed into solving three complementary problems: face detection, feature extraction, and classification (see Figure 29.1) [6,7]. Face detection is a prerequisite step, since a face cannot be recognized unless first it is detected in the scene. Feature extraction

follows face detection and reduces the face to a succinct mathematical description — the feature vector. Classification operates upon the feature vector and matches the incoming face to one of the records kept in the database.

Many of the research efforts in thermal face recognition use the thermal infrared band as a way to see in the dark or reduce the deleterious effect of light variability [8]. Methodologically, they do not differ very much from face recognition algorithms in the visible band [9]. In this chapter, we will present a novel approach to the problem of thermal facial recognition that realizes the full potential of the thermal infrared band. It consists of a statistical face detection and physiological feature extraction algorithm tailored to thermal phenomenology. We will not elaborate on classification, since we do not have to add anything new in this part of the face recognition process. 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.

29.2 Face Detection in Thermal Infrared

We approach the face detection problem in thermal infrared from a statistical point of view. Due to its physiology, a typical human face consists of *hot* and *cold* parts. Hot parts correspond to tissue areas that are rich in vasculature (e.g., periorbital and forehead). Cold parts correspond to tissue areas that feature either sparse vasculature or multi-sided exposure to the environment (e.g., cheeks and nose). This casts the human face as a bimodal distribution entity. The background can also be described by a bimodal distribution. It typically consists of walls (*cold* objects) and the upper part of the subject's body dressed in clothes (*hot* objects). The bimodal distributions of the face and background vary over time, due to thermophysiological processes and environmental noise respectively. Therefore, the problem renders itself naturally to the Bayesian framework, since we have a priori knowledge of the bimodal nature of the scene, which can be updated over time by the incoming evidence (new thermal frames from the camera).

29.2.1 Facial Tissue Delineation Using a Bayesian Approach

We consider an indoor area that is being monitored with a thermal infrared camera. The objective is to detect and delineate a human face should this become available in the scene. The segmented face data feed the feature extraction and classification modules, which complete the face recognition process (see Figure 29.1). Initially, we consider the face detection problem at the pixel level and label every pixel as either facial skin s or background b pixel.

In more detail, we use a mixture of two Normal distributions to model facial skin temperature. The dominant mode is in the upper band of the values where usually about 60 to 70% of the probability mass resides. The secondary mode is in the lower band of the values. For subjects in climate controlled rooms

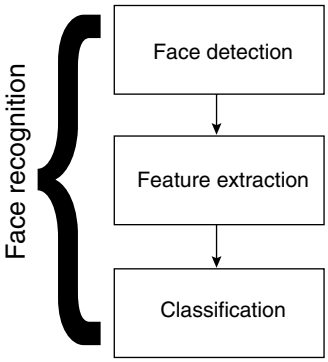


FIGURE 29.1 Architecture of a face recognition system. We will focus on face detection and feature extraction, where thermal infrared warrants a totally different approach from the one usually undertaken in the literature.

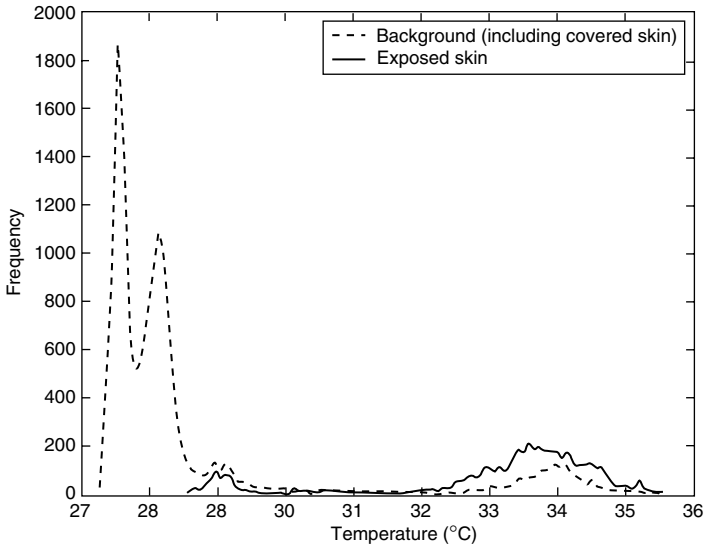


FIGURE 29.2 Temperature distributions of exposed skin and background from a set of facial thermal images in the University of Houston database. The bimodal nature of the distributions is evident.

a typical temperature range for the dominant skin mode is ∼(32–35)°C, while for the secondary mode is ∼(28–30)°C. The latter may overlap with the background distribution since areas of the face like the nose and ears have temperatures similar to the environment (see Figure 29.2).

Similarly, we use a mixture of two Normal distributions to model background temperature. The dominant mode is in the lower band of the values where about 80% of the background probability mass resides (typically, in the range ∼[27 to 29]°C). The secondary background mode has its mean somewhere between the two modes of the skin distribution and variance large enough to cover almost the entire band. Therefore, the secondary background distribution includes some relatively high temperature values. These are due to the fact that light clothes offer spots of high temperature (e.g., places where the clothes touch the skin), which mimic the skin distribution but are not skin (see Figure 29.2).

For each pixel we have some prior distribution (information) available of whether this particular pixel represents skin ($\pi(s)$) or background ($\pi(b) = 1 - \pi(s)$). Then, the incoming pixel value represents the data (likelihood) that will be used to update the prior to posterior distribution via the Bayes theorem. Based on the posterior distribution we will draw our inference of whether the specific pixel represents s or b. At the end, this posterior distribution will be used as the prior for the next incoming data point.

We call θ the parameter of interest, which takes two possible values (s and b) with some probability. The prior distribution at time t consists of the probabilities of the two complementary events, s and b. Thus, we have:

$$\pi^{(t)}(\theta) = \begin{cases} \pi^{(t)}(s), & \text{when } \theta = s \\ \pi^{(t)}(b) = 1 - \pi^{(t)}(s), & \text{when } \theta = b \end{cases} \tag{29.1}$$

The incoming pixel value x_t has a conditional distribution $f(x_t|\theta)$, which depends on whether the particular pixel is skin (i.e., $\theta = s$) or background (i.e., $\theta = b$). Based on our bimodal view of the skin and background distributions we will have for the likelihood:

$$f(x_t|\theta) = \begin{cases} f(x_t|s) : \omega_s^{(t)} N(\mu_{s_1}^{(t)}, \sigma_{s_1}^{2(t)}) + (1 - \omega_s^{(t)}) N(\mu_{s_2}^{(t)}, \sigma_{s_2}^{2(t)}), & \text{when } \theta = s \\ f(x_t|b) : \omega_b^{(t)} N(\mu_{b_1}^{(t)}, \sigma_{b_1}^{2(t)}) + (1 - \omega_b^{(t)}) N(\mu_{b_2}^{(t)}, \sigma_{b_2}^{2(t)}), & \text{when } \theta = b \end{cases} \tag{29.2}$$

There are ten parameters that are involved in the conditional distribution of Equation (29.2), namely, 2 weights, 4 means, and 4 variances. We will discuss later on, how to initialize and update these parameters.

The prior distribution, $\pi^{(t)}(\theta)$, will be combined with the likelihood, $f(x_t \mid \theta)$, to provide (via the Bayes theorem) the posterior distribution $p^{(t)}(\theta \mid x_t)$. Thus, we will have:

$$p^{(t)}(\theta \mid x_t) = \begin{cases} p^{(t)}(s \mid x_t), & \text{when } \theta = s \\ p^{(t)}(b \mid x_t) = 1 - p^{(t)}(s \mid x_t), & \text{when } \theta = b \end{cases} \tag{29.3}$$

where according to the Bayes theorem:

$$p^{(t)}(s \mid x_t) = \frac{\pi^{(t)}(s) f(x_t \mid s)}{\pi^{(t)}(s) f(x_t \mid s) + \pi^{(t)}(b) f(x_t \mid b)} \tag{29.4}$$

In the Bayesian philosophy the posterior distribution of the parameter of interest represents how the prior information (distribution) is updated in light of new evidence (data). At every time point t our inference (on whether the particular pixel represents s or b) will be based on the posterior distribution, $p^{(t)}(\theta \mid x_t)$.

This posterior distribution will also be used to provide the prior distribution for the next stage. More precisely:

$$\pi^{(t+1)}(\theta) = \begin{cases} \pi^{(t+1)}(s) = p^{(t)}(s \mid x_t), & \text{when } \theta = s \\ \pi^{(t+1)}(b) = p^{(t)}(b \mid x_t) = 1 - \pi^{(t+1)}(s), & \text{when } \theta = b \end{cases} \tag{29.5}$$

29.2.1.1 Initialization

As part of the initialization, we need to specify two things: the prior distribution, $\pi^{(1)}(\theta)$ and the likelihood $f(x_1 \mid \theta)$. Absence of information on whether a pixel is skin or background, leads us to adopt a noninformative prior distribution where a priori each pixel is equally likely to be s or b. Thus, we have:

$$\pi^{(1)}(s) = \frac{1}{2} = \pi^{(1)}(b) \tag{29.6}$$

Regarding the likelihood, we need to calculate the initial values of the ten parameters involved in the likelihood Equation (29.2). For that, we select N facial frames (off-line) from a variety of subjects. This is the so-called training set. It is important for this set to be representative, that is, to include people of both sexes, different ages, and with different physical characteristics. We manually segment on all the N frames, skin and, background areas (see Figure 29.3). The segmentation needs to be representative. For example,

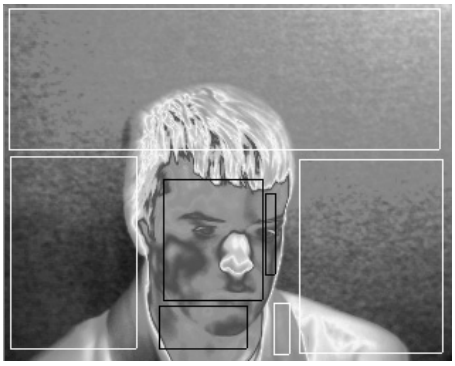


FIGURE 29.3 (See color insert following page xxx) Manual segmentation of skin (black rectangles) and background (white rectangles) areas for initialization purposes.

in the case of skin areas, we need to include eyes, ears, nose and the other facial areas, not parts thereof. Out of the N frames the segmentation will yield N_s skin and N_b background pixels.

For the skin distribution we have available at the initial state the pixels $x_1, x_2, \cdots, x_{N_s}$, which are assumed to be sampled from a mixture of two Normal distributions:

$$f(x_j \mid s) = \sum_{i=1}^{2} \omega_{s_i} N(\mu_{s_i}, \sigma_{s_i}^2) \tag{29.7}$$

where $\omega_{s_2} = 1 - \omega_{s_1}$. We estimate the mixture parameters ω_{s_i}, μ_{s_i}, and $\sigma_{s_i}^2$ using the N_s skin pixels of the training set via the EM algorithm. Initially, we provide the EM algorithm with some crude estimates of the parameters of interest: $\omega_{s_i}^{(0)}$, $\mu_{s_i}^{(0)}$, and $(\sigma_{s_i}^{(0)})^2$. Then, we apply the following loop: For $k = 0, 1, \ldots$ we calculate:

$$z_{ij}^{(k)} = \frac{\omega_{s_i}^{(k)} (\sigma_{s_i}^{(k)})^{-1} \exp\left\{-\frac{1}{2(\sigma_{s_i}^{(k)})^2}(x_j - \mu_{s_t}^{(k)})^2\right\}}{\sum_{t=1}^{2} \omega_{s_t}^{(k)} (\sigma_{s_t}^{(k)})^{-1} \exp\left\{-\frac{1}{2(\sigma_{s_t}^{(k)})^2}(x_j - \mu_{s_t}^{(k)})^2\right\}} \tag{29.8}$$

$$\omega_{s_i}^{(k+1)} = \frac{\sum_{j=1}^{N_s} z_{ij}^{(k)}}{N_s} \tag{29.9}$$

$$\mu_{s_i}^{(k+1)} = \frac{\sum_{j=1}^{N_s} z_{ij}^{(k)} x_j}{N_s \omega_{s_i}^{(k+1)}} \tag{29.10}$$

$$(\sigma_{s_i}^{(k+1)})^2 = \frac{\sum_{j=1}^{N_s} z_{ij}^{(k)} (x_j - \mu_{s_i}^{(k+1)})^2}{N_s \pi_{s_i}^{(k+1)}} \tag{29.11}$$

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

$$|\omega_{s_i}^{(k+1)} - \omega_{s_i}^{(k)}| < \varepsilon \quad i = 1, 2 \tag{29.12}$$

where ε is a small positive number $(10^{-3}, 10^{-4}, \ldots)$. We apply a similar EM process for determining the initial parameters of the background distributions.

29.2.1.2 Inference

Once a data point x_t becomes available we are faced with the decision of whether the particular pixel represents skin or background. In a decision theory framework this can be cast as testing the statistical hypotheses for the parameter of interest θ:

$$\begin{Bmatrix} H_0\text{: } \theta = s \text{ (i.e., the pixel represents skin)} \\ H_1\text{: } \theta = b \text{ (i.e., the pixel represents background)} \end{Bmatrix}$$

By using Equations (29.2) and (29.4) we combine the data with the prior distributions to obtain the posterior distributions. Our inference will be based on these posterior distributions. We need to decide whether we will accept H_0 or H_1. The easiest way to do this within the Bayesian framework is to favor the hypothesis that has the highest a posteriori coverage. That simply means:

$$\begin{Bmatrix} \text{Accept } H_0 \text{ if: } p^{(t)}(s \mid x_t) > p^{(t)}(b \mid x_t) \Leftrightarrow p^{(t)}(s \mid x_t) > 0.5 \\ \text{Accept } H_1 \text{ if: } p^{(t)}(s \mid x_t) < p^{(t)}(b \mid x_t) \Leftrightarrow p^{(t)}(s \mid x_t) < 0.5 \end{Bmatrix}$$

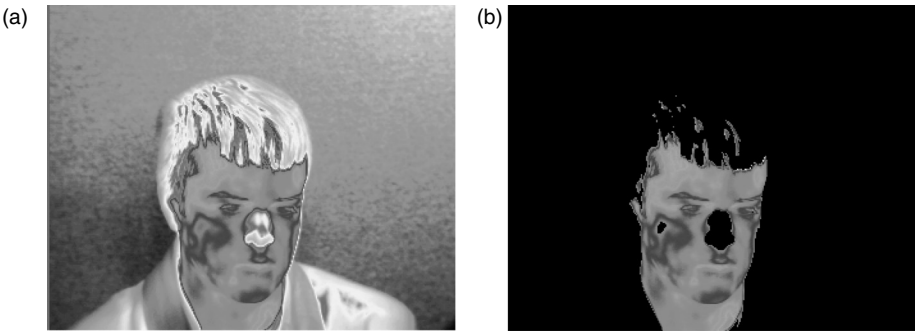


FIGURE 29.4 (See color insert) Visualization of Bayesian segmentation on a subject: (a) original image; (b) segmented image. The nose has been erroneously segmented as background and a couple of hair patches have been erroneously marked as facial skin. This is due to occasional overlapping between portions of the skin and background distributions. The isolated nature of these mislabeled patches makes them easily correctable through post-processing.

Figure 29.4 visualizes the results of Bayesian segmentation on a typical facial thermal image from the University of Houston database.

29.3 Facial Feature Extraction in Thermal Infrared

Once the face is delineated from the rest of the scene, one can extract the features necessary for classification. In contrast to the visible band, thermal infrared provides the capability to extract physiological features. In particular, blood that flows in the major superficial vessels creates a substantial convective heat effect that is manifested in thermal imagery. Segmentation of the vascular network can provide the basis of a unique feature vector. The topology and extent of the recovered vascular network depend on the genetic traits and physical characteristics of the individual (e.g., facial skin fat deposits). Therefore, facial vasculature is an identifying entity that endures through aging and superficial changes of appearance.

The problem for vessel segmentation has been solved in other imaging modalities using a variety of methods. To the best of our knowledge, it is the first time that vessel segmentation is reported in the thermal imaging modality. In our effort, we took into account methodologies used for vessel segmentation in modalities other than thermal imaging (e.g., ultrasound and MRI). We can broadly categorize these methodologies as follows:

- Center line detection approaches
- Ridge-based approaches
- Region growing approaches
- Differential geometry-based approaches
- Matched filter approaches
- Mathematical morphology approaches

In the center line extraction scheme, the vasculature is developed by traversing the vessels' center lines. The method employs thresholding and thinning of the image, followed by connected component analysis. Center line extraction techniques are used by References 10 and 11. They use graph search theory to find out vessel center lines and then curvature features to reconstruct the vasculature. In Reference 12, the authors detect center lines from images acquired from different angles. As a result, they are able to represent the 3D shape of the vessels.

Ridge-based approaches convert a 2D gray scale image into a 3D surface by mapping the intensity values along the z-axis. Once the surface is generated, the local ridge points are those where the directional gradient is the steepest. The ridges are invariant to affine transforms and this property is exploited in

medical registration [13,14]. The method of segmentation of vessels using the ridge-based approach is discussed in Reference 15 and more extensive details are available in its references.

One of the most widely used methods of vessel segmentation is the region growing method. Traditionally, the parameters used for growing the vessel region are pixel intensity and proximity. Recent improvements in this method include gradient and texture [16,17]. Such methods, although give good results, depend upon initialization. In Reference 18, based on the starting seed provided by the user the region grows upon similar spatial, temporal, and structural features. A similar method is also used in Reference 19.

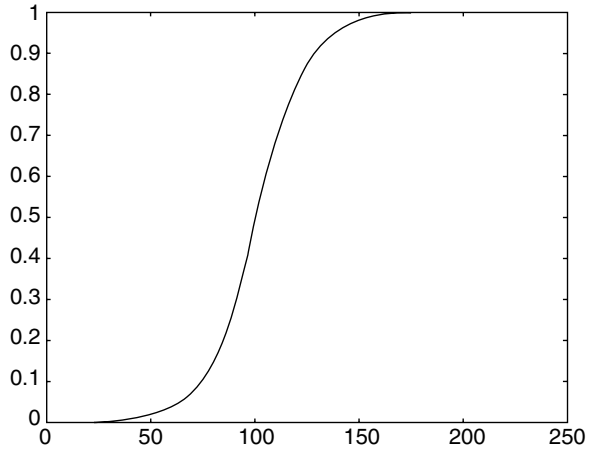
If a 2D image is mapped into a 3D surface, then the crest lines are the most salient features of that surface. This gives rise to the differential geometry based methods for vessel segmentation. The crest lines are obtained by linking the crest points in the image, which are the local maxima of the surface curvature. A prominent differential geometry-based method is the directional anisotropic diffusion (DAD), which is discussed in Reference 20. It uses Gaussian convolution to remove the image noise. This method is a generalized form of the method reported in Reference 21 and uses the gradient as well as the minimum and maximum curvature information to differentiate the diffusion equation. The method removes the noise without blurring and is thus, very useful in edge enhancement procedures.

The matched filtering approach uses a series of Gaussian kernels of different sizes and orientations. In Reference 22 the orientation of the Gaussian filters is chosen using the Hessian matrix. A similar approach is used in Reference 23 to enhance and detect vessels in real time. These methods are similar to the Gabor filters, which are extensively used in texture analysis.

29.3.1 Morphological Reconstruction of Superficial Blood Vessels

In our method, first, we smooth the image to remove the unwanted noise and then, we apply the morphological operators. In thermal imagery of human tissue the major blood vessels do not have strong edges. Due to the thermal diffusion process the edges that we get have a sigmoid temperature profile. A sigmoid function can be written as $y = 1/(1 + e^{-x})$ (see Figure 29.5).

The method of anisotropic diffusion has proved very effective in handling sigmoid edges [24,25]. Nonlinear anisotropic diffusion filters are iterative filters introduced by Perona et al. [21] Greig et al. used such filters to enhance MR images [26]. Sapiro et al. used a similar technique to perform edge preserving smoothing of MR images [27]. Others have shown that diffusion filters can be used to enhance and detect object edges within images [21,28].



AQ: Please provide 'x' and 'y' co-ordinate for Figure 29.5.

FIGURE 29.5 Sigmoid function plot.

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:

$$\frac{\partial I(\bar{x}, t)}{\partial t} = \nabla(c(\bar{x}, t)\nabla I(\bar{x}, t)) \tag{29.13}$$

In our case $I(\bar{x}, t)$ is the thermal image, $\bar{x}$ refers to the spatial dimensions, and t to time. The function $c(\bar{x}, t)$ is a monotonically decreasing function of the image gradient magnitude and is called the *diffusion function*

$$c(\bar{x}, t) = f(|\nabla I(\bar{x}, t)|) \tag{29.14}$$

This allows locally adaptive diffusion strengths. Edges are thus, selectively smoothed or enhanced on the basis of the evaluated diffusion function.

Perona and Malik have suggested the following two formulations for the diffusion function:

$$c_1(\bar{x}, t) = \exp\left(-\left(\frac{|\nabla I(\bar{x}, t)|}{k}\right)^2\right) \tag{29.15}$$

$$c_2(\bar{x}, t) = \exp\left(\frac{1}{1 + \left(\frac{|\nabla I(\bar{x}, t)|}{k}\right)^{1+\alpha}}\right) \tag{29.16}$$

Typical responses of the thermal image of the wrist to the Perona–Malik filters with diffusion functions c_1 and c_2 respectively are shown in Figure 29.6. One can observe that for the same image gradient and value of the k parameter a steeper slope is obtained for c_1 as compared to c_2.

The discrete version of the anisotropic diffusion filter of Equation (29.13) is as follows:

$$\begin{aligned} I_{t+1}(x, y) = I_t + \frac{1}{4} \times [&c_{N,t}(x, y)\nabla I_{N,t}(x, y) + c_{S,t}(x, y)\nabla I_{S,t}(x, y) \\ &+ c_{E,t}(x, y)\nabla I_{E,t}(x, y) + c_{W,t}\nabla I_{W,t}(x, y)] \end{aligned} \tag{29.17}$$

The four diffusion coefficients and gradients in Equation (29.17) correspond to four directions (i.e., North, South, East, and West) with respect to 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:

$$c_{N,t}(x, y) = \exp\left(\frac{-\nabla I_{N,t}^2(x, y)}{k^2}\right) \tag{29.18}$$

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

Figure 29.7(a) and Figure 29.7(b) show the original thermal image of skin tissue (wrist in this case) and its temperature surface plot respectively. One can observe in the surface plot the noisy ridges formed due to the hair. Figure 29.7(c) shows the filtered skin image. The hair has been removed from the surface, resulting into smoother ridges and heightened peaks in the temperature surface plot (Figure 29.7[d]).

As these figures testify anisotropic diffusion is highly beneficial in improving the contrast in the image and removing the noise. This is in preparation of vessel segmentation through morphological methods.

29.3.1.1 Image Morphology

Image morphology is a way of analyzing imagery based on shapes. It is rooted on set theory, the Minkowski operators, and DeMorgan's Laws. The Minkowski operators are usually applied to binary images where it

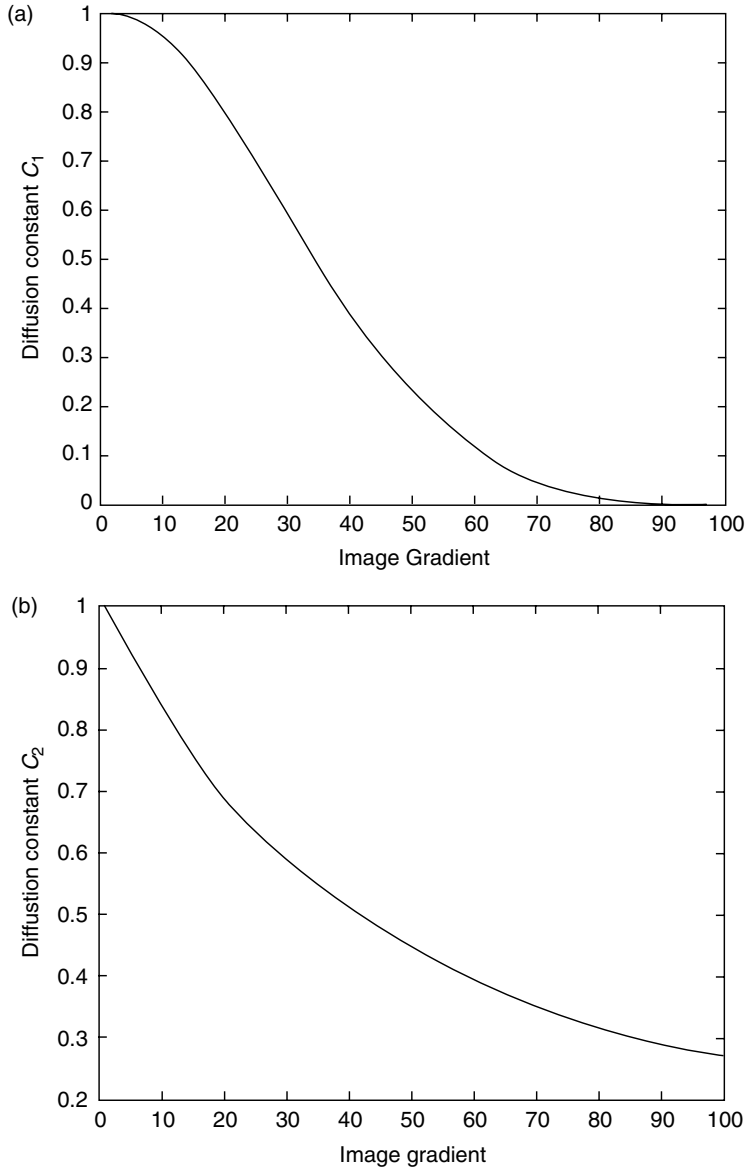


FIGURE 29.6 (a) Plot of diffusion function c_1 against the image gradient. (b) Plot of diffusion function c_2 against the image gradient.

is easy to perform OR and AND operations. The same operators can be applied to gray scale images with small modifications.

Image morphology is a simple but effective tool for shape analysis and segmentation. In retinotherapy it has shown great results in localization of blood vessels in the retina. Leandro et al. [29] have used morphological operators for vessel delineation in the retina where the background intensity was very close to that of the blood vessels. In our case, we have the blood vessels, which have a relatively low contrast compared to that of the surrounding tissue. As per our hypothesis the blood vessel is a tubule like structure running either along the length of the forearm or the face. Thus, our problem is to segment tubule structures from the image. We employ for this purpose a top-hat segmentation method, which is a combination of erosion and dilation operations.

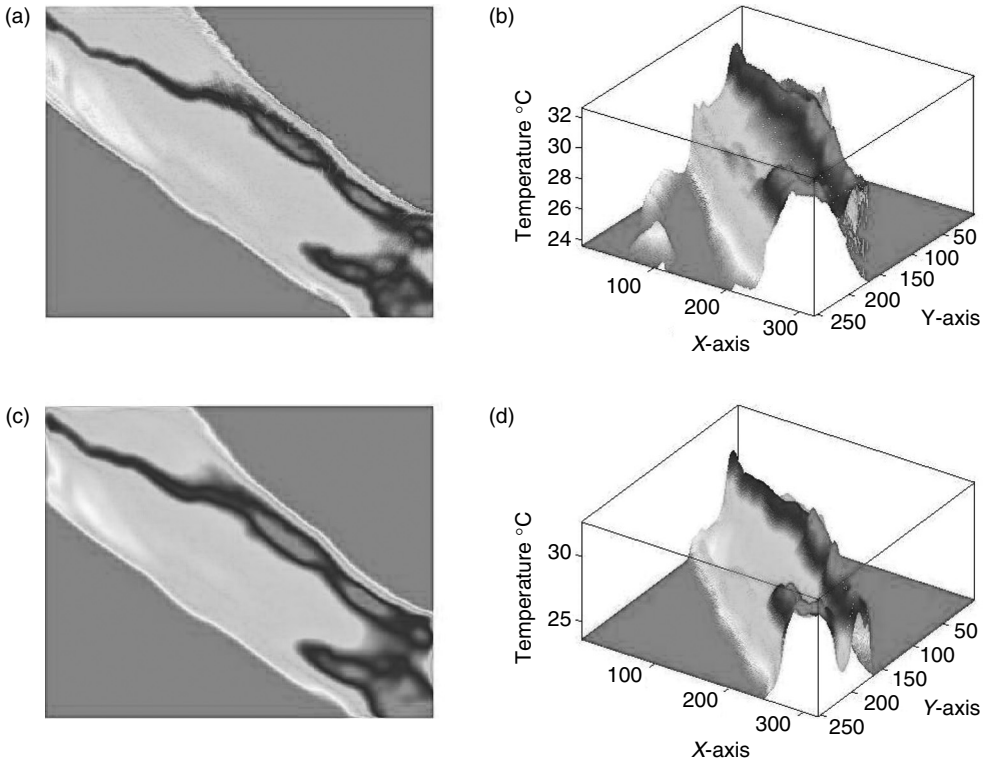


FIGURE 29.7 (See color insert) (a) Original thermal image of a wrist. (b) Temperature surface plot of the original image. (c) Diffused thermal image of the wrist. (d) Temperature surface plot of the diffused image.

1	1	1
1	1	1
1	1	1

A 3×3 structuring element

0	0	1	0	0
0	1	1	1	0
1	1	1	1	1
0	1	1	1	0
0	0	1	0	0

A 5×5 structuring element

FIGURE 29.8 Structuring elements.

Erosion and dilation are the two most basic operations in mathematical morphology. Both of these operations take two pieces of data as input: an image to be eroded or dilated and a structuring element. The structuring element is similar to what we know as a kernel in the convolution operation. There are a variety of structuring elements available but a simple 3×3 square matrix is used more often. Figure 29.8 shows some commonly used structuring elements.

The combination of erosion and dilation results into more advanced morphological operations such as:

- Opening
- Closing
- Skeletonization

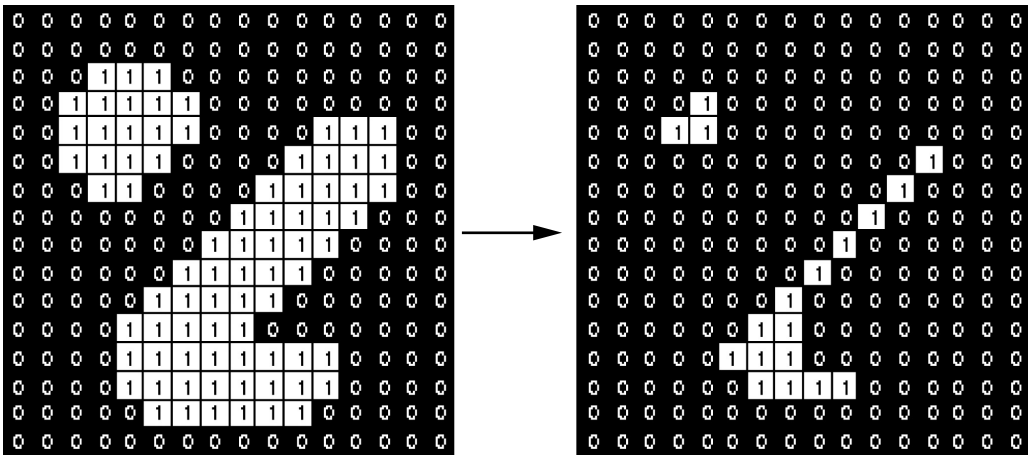


FIGURE 29.9 Binary erosion by a 3 × 3 flat structuring element.

- White top hat segmentation
- Black top hat segmentation

In our application, we are interested only in image opening and top hat segmentation.

1. *Erosion:* The name suggests that this operation erodes the boundary pixels of an image and thus, the resultant image has a shrunk boundary. Mathematically, it is defined as follows:

$$A \oplus B = \{z \mid (A \cap B)_z \subseteq B\} \tag{29.19}$$

Therefore, erosion removes elements smaller than the structuring element. Figure 29.9 shows a simple erosion operation performed on a binary image. Gray scale erosion with a flat structuring element generally darkens the image. Bright regions surrounded by dark regions shrink in size and dark regions surrounded by bright regions grow in size. Small bright spots in images disappear, as they are eroded away down to the surrounding intensity value. In contrast, small dark spots grow. The effect is most pronounced at places in the image where the intensity changes rapidly. Regions of fairly uniform intensity are left more or less unchanged except at their edges.

2. *Dilation:* The name suggests that this operation gradually expands the boundary pixels of an image and thus, the resultant image will have an enlarged boundary. Dilation results in fusing small holes in the boundary area, by enlarging the boundary pixels. This is the equivalent of a smoothing function. Mathematically, it is defined as follows:

$$A \oplus B = \{z \mid (\cap B)_z \cap A \neq \emptyset\} \tag{29.20}$$

Therefore, dilation fills in gaps smaller than the structuring element. Figure 29.10 shows a simple dilation operation performed on a binary image. Gray scale dilation with a flat structuring element generally brightens the image. Bright regions surrounded by dark regions grow in size and dark regions surrounded by bright regions shrink in size. Small dark spots in images disappear as they are "filled in" to the surrounding intensity values. In contrast, small bright spots will grow in size. The effect is most pronounced at places in the image where the intensity changes rapidly. Regions of fairly uniform intensity will be largely unchanged except at their edges.

3. *Opening:* The basic effect of an opening operation is reminiscent of erosion, since it tends to remove some of the foreground (bright) pixels at the edges. However, it is less destructive than erosion. The size and shape of the structuring element plays an important role in performing opening. The operation

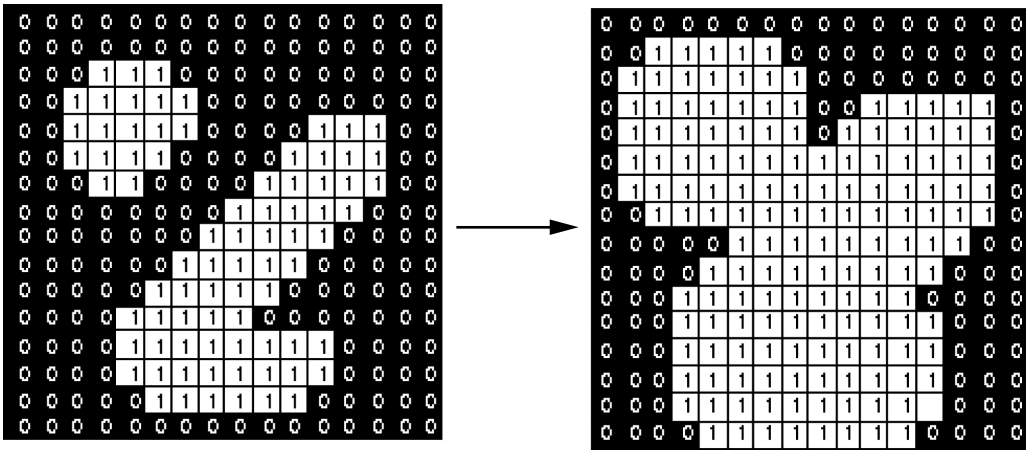


FIGURE 29.10 Binary dilation by a 3 × 3 flat structuring element.

preserves foreground regions that have a shape similar to its structuring element while erodes all other regions of foreground pixels. In mathematical terms opening can be written as:

$$A \circ B = (A \odot B) \oplus B \quad \text{or}$$
$$A \circ B = \cup[(B)_z \mid (B)_z \subseteq A] \tag{29.21}$$

While erosion can be used to eliminate small clumps of undesirable foreground pixels (e.g., "salt noise") quite effectively, it has the disadvantage that it affects all foreground regions indiscriminately. Opening gets around this by performing both erosion and dilation on the image. The effect of opening can be visualized quite easily. Imagine taking the structuring element and sliding it inside each foreground region without changing its orientation. All foreground pixels that can be covered by the structuring element with the structuring element being entirely within the foreground region will be preserved. However, all foreground pixels which cannot be reached by the structuring element without parts of it moving out of the foreground region will be eroded away. After the opening has been carried out, the new boundaries of foreground regions will all be such that the structuring element fits inside them. Therefore, further openings with the same element have no effect, a property known as idempotence. The effect of an opening on a binary image using a 3 × 3 flat structuring element is illustrated in Figure 29.11.

4. *White Top Hat Segmentation:* Many times gray scale images feature poor contrast. For example, in our case thermal imagery of human tissue has poor contrast around the vessels due to the thermal diffusion process. As a result, image thresholding yields very poor results. Top-hat segmentation is a morphological operation that corrects this problem. Top hat segmentation has two forms:

- White top hat segmentation
- Black top hat segmentation

The white top-hat segmentation process enhances the bright objects in the image, while the black top-hat segmentation enhances the dark objects. In our case, we are interested in enhancing the bright (hot) ridge like structures corresponding to the blood vessels. Therefore, we are interested only in the white top-hat segmentation process. Two methods have been introduced for performing the white top hat segmentation. The first one proposed in Reference 30 is based on image opening using a flat structuring element, while the second one proposed in Reference 31 uses H-dome transformation. We have adopted the first method where the image is first opened and then this opened image is subtracted from the original image. This gives only the peaks in the image and thus enhances the maxima. The step by step evolution of the original

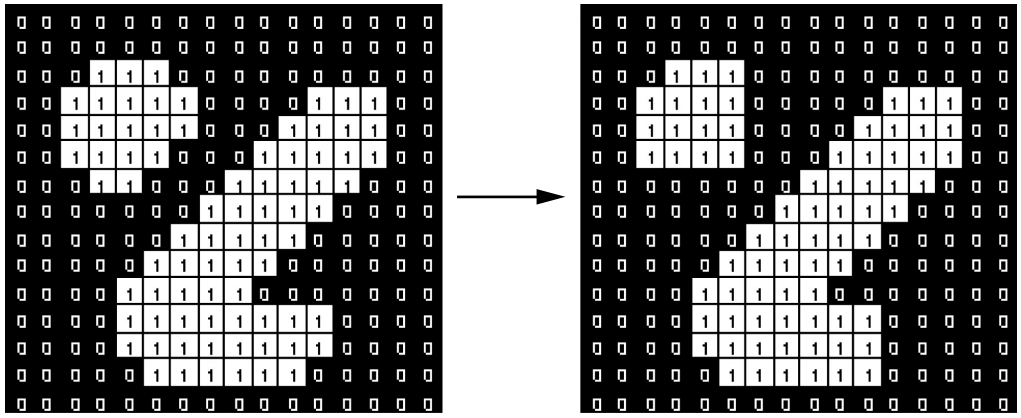


FIGURE 29.11 Binary opening by a 3 × 3 flat structuring element.

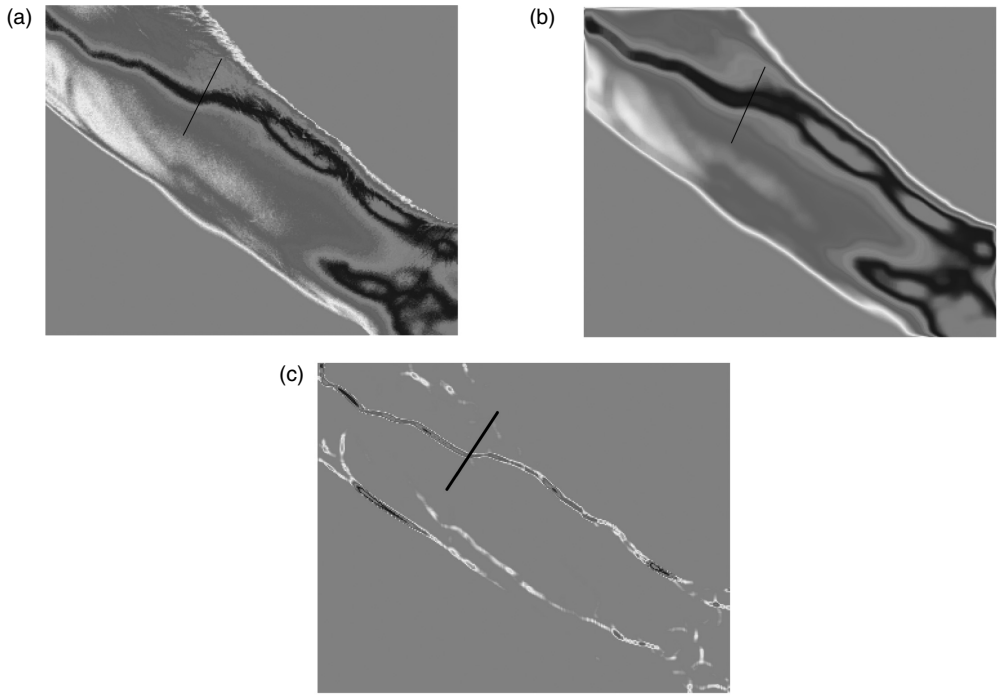


FIGURE 29.12 (See color insert) Thermal image of a wrist: (a) original, (b) opened, and (c) top hat segmented.

image toward the top-hat segmented image is shown in Figure 29.12. The simple functioning of top-hat transformation can be understood from the line profile plots in Figure 29.13.

29.4 Conclusion

We have outlined a novel approach to the problem of face recognition in thermal infrared. The cornerstones of the approach are a Bayesian face detection method followed by a physiological feature extractor. The face detector capitalizes upon the bimodal temperature distribution of human skin and typical indoor backgrounds. The physiological feature extractor delineates the facial vascular network based on a white

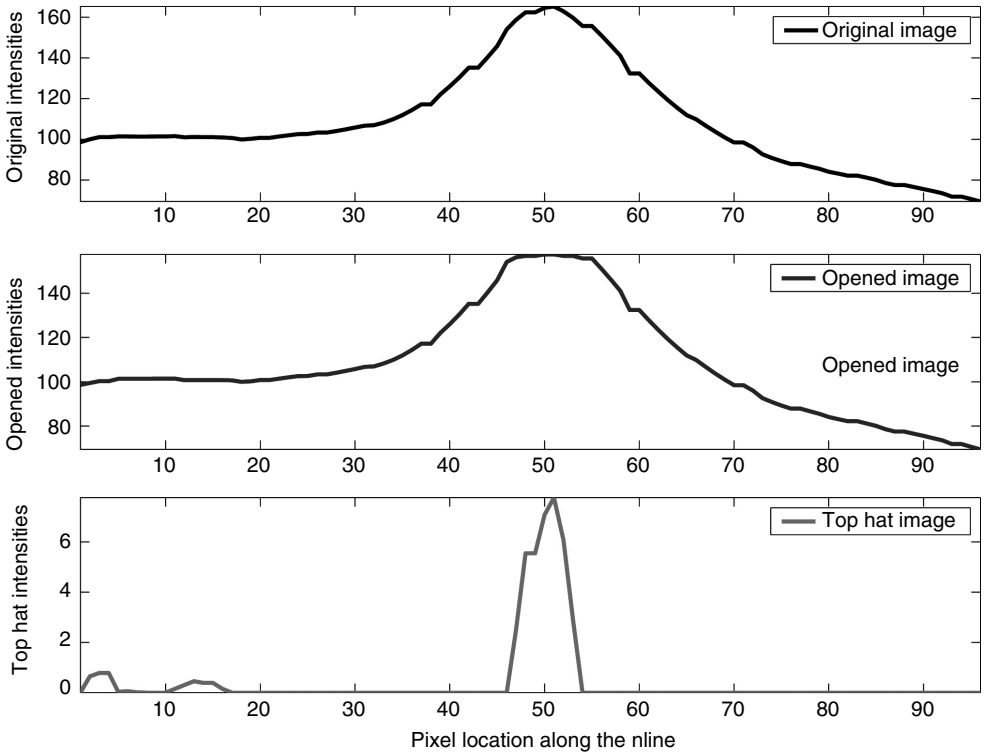


FIGURE 29.13 Image line profiles for (a) original, (b) opened, and (c) top-hat segmented image. The profiles were taken along the line shown in Figure 29.8.

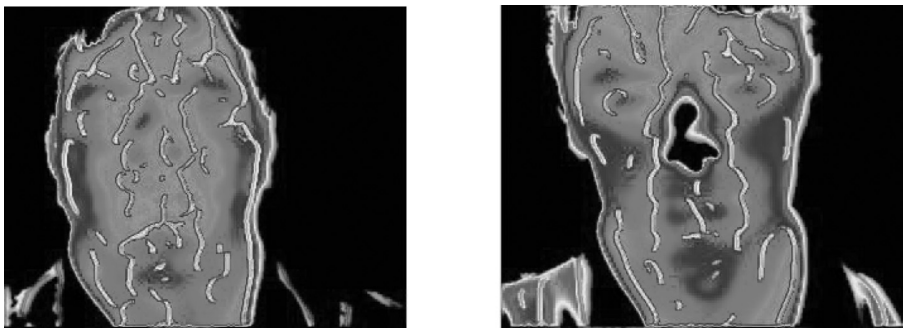


FIGURE 29.14 (See color insert) Segmented facial images annotated with the facial vascular network (yellow lines) per the approach detailed in this chapter.

top hat segmentation preceded by anisotropic diffusion (see Figure 29.14). These novel tools can be used in combination with traditional classification methods to exploit the full potential of thermal infrared for face recognition — one of the fastest growing biometrics.

Acknowledgments

Research related to the content of this chapter was supported by NSF grant #0313880 and ONR/DARPA grant #N00014-03-1-0622. The views expressed in this chapter do not necessarily represent the views of the funding Agencies.

References

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

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

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

[4] Wilder, J., Phillips, P., 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*, Killington, Vermont, 1996, 182.

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

[6] Srivastava, A., Liu, X., Thomasson, B., and Hesher, C., Spectral probability models for IR images with applications to IR face recognition, in *Proceedings of the IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, Kauai, Hawaii, 2001.

[7] 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., 2004.

[8] Socolinsky, D., Wolff, L., Neuheiser, J., and Evelenad, C., Illumination invariant face recognition using thermal infrared imagery, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1, Kauai, Hawaii, 2001, 527.

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

[10] Niki, N., Kawata, Y., Satoh, H., and Kumazaki, T., 3D imaging of blood vessels using x-ray rotational angiographic system, in *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference*, 3, San Francisco, California, 1993, 1873.

[11] Kawata, Y., Niki, N., and Kumazaki, T., Characteristics measurement for blood vessels diseases detection based on cone-beam CT images, in *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference*, 3, 1995, 1660.

[12] Parker, D.L., Wu, J., and van Bree, R.E., Three-dimensional vascular reconstruction from projections: a theoretical review, in *Proceedings of the IEEE Engineering in Medicine and Biology Society Conference*, 1, New Orleans, 1998, 399.

[13] Aylward, S., Pizer, S., Bullit, E., and Eberl, D., Intensity ridge and widths for tubular object segmentation and registration, in *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*, 1996, 131.

[14] Aylward, S. and Bullitt, E., Analysis of parameter space of a metric for registering 3D vascular images, *MICCAI*, 2001.

[15] Bullitt, E. and Aylward, S.R., Analysis of time-varying images using 3D vascular models, *Applied Imagery Pattern Recognition Works*, 2001, 9.

[16] Jones, T. and Metaxas, D., Image segmentation based on the integration of pixel affinity and deformable models, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Santa Barbara, California, 1998, 330.

[17] Pednekar, A.S. and Kakadiaris, I.A., Applications of virtual reality in surgery, in *Proceedings of the Indian Conference on Computer Vision, Graphics, and Image Processing*, Bangalore, India, 2000, 215.

[18] O'Brien, J.F. and Exquerra, N.F., Automated segmentation of coronary vessels in angiographic image sequences utilizing temporal, spatial, and structural constraints, in *Proceedings of the SPIE Visualisation Conference in Biomedical Computing*, 1994.

[19] Schmitt, H., Grass, M., Rasche, V., Schramm, O., Haehnel, S., and Sartor, K., An x-ray based method for determination of the contrast agent propagation ub in 3D vessel structures, *IEEE Transactions on Medical Imaging*, 21, 251, 2002.

[20] Krissian, K., Malandain, G., and Ayache, N., Directional Anisotropic Diffusion Applied to Segmentation of Vessels in 3D Images, Technical Report 3064, INRIA, 1996.

[21] Perona, P. and Malik, J., Scale-space and edge detection using anisotropic diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, 629, 2003.

[22] Sato, Y., Nakjima, S., Shiraga, N., Atsumi, H., Yoshida, S., Kikker, T., Greig, G., and Kikinis, R., 3D multiscale line filter for segmentation and visualization of curvilinear structures in medical images, *IEEE Medical Image Analysis*, 2, 143, 1998.

[23] Poli, R. and Valli, G., An algorithm for real-time vessel enhancement and detection, *Computer Methods and Programming in Biomedicine*, 52, 1, 1997.

[24] Acton, S.T., Bovik, A.C., and Crawford, M.M., Anisotropic diffusion pyramids for image segmentation, in *Proceedings of the IEEE International Conference on Image Processing*, 3, Austin, Texas, 1994, 478.

[25] Acton, S.T., Edge enhancement of infrared imagery by way of the anisotropic diffusion pyramid, in *Proceedings of the IEEE International Conference on Image Processing*, 1, Laussane, Switzerland, 1996, 865.

[26] Gerig, G., Kubler, O., Kikinis, R., and Jolesz, F.A., Nonlinear anisotropic filtering of MRI data, *IEEE Transactions on Medical Imaging*, 11, 221, 1992.

[27] Sapiro, G. and Tannenbaum, A., Edge Preserving Geometric Smoothing of MRI Data, Technical Report, University of Minnesota, Department of Electrical Engineering, April 1994.

[28] Nordstrom, N., Biased anisotropic diffusion — a unified generalization and diffusion approach to edge detection, *Image and Vision Computing*, 8, 318, 1990.

[29] Leandro, J., Cesar, R.M. Jr., and Jelinek, H., Blood vessel segmentation in retina: preliminary assessment of the mathematical morphology and wavelet transform techniques, *14th Brazilian Symposium on Computer Graphics and Image Processing*, 2001, 84.

[30] Dougherty, E.R., Euclidean gray-scale granulometries: representation and umbra inducement, *Journal of Mathematical Imaging and Vision*, 1, 7, 1992.

[31] Vincent, L., Morphological grayscale reconstruction in image analysis: application and efficient algorithms, *IEEE Transactions on Image Processing*, 2, 176, 1993.