Manuscript for Information Fusion in Biometrics

Using Multiple Models to Uncover Blood Vessel Patterns in Color Images for Forensic Analysis

Chaoying Tang[1,3], Hengyi Zhang[2,3] and Adams Wai-Kin Kong[2]

[1] Nanjing University of Aeronautics and Astronautics, 29 Yudao Street, Nanjing, P. R. China, 210016.
[2] Nanyang Technological University, 50 Nanyang Avenue, Singapore, 639798.
[3] Contributed equally, listed according to the alphabetical order of their surnames.
E-mails: cytang@nuaa.edu.cn, hengyi@ntu.edu.sg, adamskong@ntu.edu.sg

Corresponding Author:
Chaoying Tang
College of Automation Engineering,
Nanjing University of Aeronautics and Astronautics,
29 Yudao Street, Nanjing, 210016,
P.R.China.
Tel: +8613851606886
E-mails: cytang@nuaa.edu.cn

Abstract— With the proliferation of digital cameras, images of crimes, such as child sexual abuse images, are increasing dramatically. Both verification and identification of criminals and victims in these images are highly difficult and often impossible for the current biometric technology because their faces, tattoos, and distinctive skin mark patterns are not always observable. Superficial blood vessels under skin are a potential solution to compensate the weaknesses of the traditional biometric traits. However, blood vessels were neglected by law enforcement agencies because they are generally invisible in color images. To use blood vessel patterns in forensic analysis, this paper proposes three computational models to uncover hidden patterns, two optimization schemes to handle illumination variations and prevent over-relying on biophysical parameters measured in ideal medical conditions, a matching algorithm to automatically extract and compare noisy patterns, and two fusion rules to combine patterns from the three models for performance enhancement. The experimental results on 1900 color images and 1900 infrared images from 490 forearms and 460 thighs show that the matching performance of the blood vessel patterns from the color images is comparable with that from the infrared images. The proposed models are also applied to hands, arms, thighs, chests, breasts, and abdomens of men, women, and children in indoor and outdoor images collected from the Internet. Though these images were taken in uncontrolled environments and the subjects had different poses, the proposed models can uncover blood vessels. These results indicate that the potential of using blood vessel patterns in forensic analysis was underestimated.

Keywords **—**Forensics, child pornography, biometrics, skin marks, tattoos.

1. Introduction

Both verification and identification of criminals and victims are always critical tasks for law enforcement agencies. These tasks are becoming more important, because images of crimes are increasing exponentially. Let us take child sexual offenses as an example. In Canada alone, Cybertip.ca received over 21,000 tips about online child exploitation between 2002 and 2008 [1] and found 12,696 websites offering child sexual abuse images (also known as child pornography) in 2009 [2]. Australia is also facing an explosion in online child sexual abuse images [3]. Anyone who possesses, makes, prints, publishes, distributes, sells, or imports child sexual abuse images commits a criminal offense. Criminals always hide their faces and tattoos, which are regularly used by law enforcement agencies for criminal verification and identification. Thus, prosecuting them is highly challenging. U.S. attorneys declined to prosecute over 30% of child sexual abuse suspects because of weak or inadmissible evidence, which was the major reason for declination of prosecution [4]. In a child sexual exploitation case *United States v. Michael Joseph Pepe*, the author A.K., who was recruited by the U.S. Department of Justice as an expert witness, faced the same challenge [5]. He was requested to verify an arrested suspect and a criminal in a set of images, which showed the lower part of his body, but his face was not observable. In addition to this case, the author A.K. saw a large number of child sexual abuse materials in the Toronto Police Service and the U.S. Immigration and Customs Enforcement with the same characteristics. Though neither faces nor tattoos of the pedophiles were available for identification, the images were high resolution and close-up of their non-facial body sites.

In addition to child sexual abuse images, many cities such as London and Athens experienced riots. Many rioters, who threaten the law and order in many societies and damage others' property, always wear face masks to avoid verification and identification. Identifying masked terrorists and masked gunmen is also important in many counties.

To address these tough verification and identification problems, new biometric traits have to be developed. Though skin marks and androgenic hair have been considered, not all body sites, e.g., hands, have distinctive skin mark patterns and androgenic hair patterns [6-7]. The patterns formed by the blood vessels that lie between the skin and muscle are a potential solution because of their universality, permanence, and distinctiveness. Currently, vein recognition systems depend on infrared and laser imaging techniques to capture high-quality blood vessel patterns from palms, wrists, and fingers (where the skin is thin) for commercial applications, e.g., access control [8-10, 32-33]. Very limited works attempted to uncover blood vessel patterns hidden in color images for forensic analysis. The optical-based vein uncovering (OBVU) method, which was the first attempt, is very sensitive to illumination changes and heavily depends on the biophysical parameters measured in ideal medical conditions [11]. Skin is a complex structure. Different persons and different body sites have diverse skin properties. The OBVU method fails to handle skin diversity because it is a single-model approach. Only 32 color images were examined through manual comparison. To perform forensic identification, i.e., searching suspects in a given database and systematically evaluating blood vessel patterns from color images, automatic blood vessel extraction, representation, and matching algorithms are essential. Hand, wrist, and finger vein verification methods have been developed, but they were designed for high-quality near infrared (NIR) images captured in controlled environments from cooperative users. It is nearly impossible to guarantee the quality of blood vessel patterns extracted from color images because of subcutaneous fat and other uncontrollable physical parameters weakening the penetration of visible light. It is worth mentioning that the penetration capability of visible light in skin is much weaker than that of NIR in skin. Noisy blood vessel patterns are almost unavoidable. Without user cooperation, non-linear distortion is also inevitable. To address these problems and make verification and identification of criminal and victim based on blood vessel patterns hidden in color images possible, this paper proposes three optical skin models to uncover hidden patterns, two optimization schemes to handle illumination variation and avoid over-relying on biophysical parameters measured in ideal medical conditions, an identification algorithm to automatically extract and match noisy patterns, and two fusion rules to combine patterns from the three models to enhance the matching performance.

The preliminary version of this work presented in [12] was a single model approach. In this paper, two new skin optical models are presented to uncover more blood vessels in different imaging conditions. A new dissimilarity measure which takes orientation, distance, and magnitude of corresponding blood vessels into consideration and two fusion schemes which combine information from different models are also proposed. The preliminary method was evaluated on a database with 300 images from 150 right forearms. The proposed models, schemes, and algorithms are examined on a database with 3800 images from 490 forearms and 460 thighs. Furthermore, the proposed models are applied to hands, arms, thighs, chests, breasts, and abdomens of men, women, and children in indoor and outdoor images collected from the Internet.

The rest of this paper is organized as follows. Section 2 presents a physical model to compute skin reflectance. This model is used to develop the three optical models. Section 3 provides the three optical models with the two optimization schemes. Section 4 describes the proposed blood vessel extraction, representation, and matching algorithms. Section 5 reports the experimental results. Section 6 discusses the impacts of our findings.

2. A Physical Approach for Skin Reflectance Computation

Skin is a semi-transparent multilayer object. When light hits skin, some is absorbed and scattered by compounds in the skin, while some is reflected and captured by the sensor in a camera. By analyzing the reflected light, the internal structure of the skin can be revealed. A number of optical theories have been developed for studying radiation passing through a scattering medium. The Kubelka-Munk (K-M) theory is the simplest one. In this paper, the three optical models use the K-M theory and Reichman's solution to calculate skin reflectance [13-14]. Skin can be considered as an n-layered material and its total reflectance, $R_{12\ldots n}$, and transmittance, $T_{12\ldots n}$, can be computed by the recursive equations:

$$R_{12\ldots n} = R_{12\ldots n-1} + \frac{T_{12\ldots n-1}^2 R_n}{1 - R_{12\ldots n-1} R_n}, \tag{1}$$

$$T_{12\ldots n} = \frac{T_{12\ldots n-1} T_n}{1 - R_{12\ldots n-1} R_n}, \tag{2}$$

where R_i and T_i are respectively the reflectance and transmittance of the i^{th} layer [13]. Fig. 1 shows a three-layered skin model. The total reflectance and transmittance of an n-layered skin are controlled by the absorption and scattering coefficients and the thicknesses of different layers. To simplify the notations, R_T is used to represent $R_{12\ldots n}$.

It is worth mentioning that the K-M theory has been employed by medical scientists for dermatological research [13]. Their empirical findings are extensively used in our optical models. The K-M theory is only a rough solution to the radiative transfer equation, which quantitatively describes light transport in different materials. Its exact analytical solution has not been obtained for complex and multiple scattering mediums such as human skin [15].

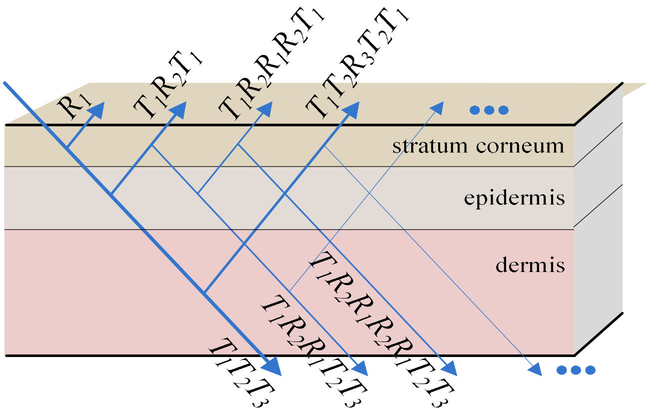


Fig. 1 Illustration of the recursive equations (Eqs. 1-2) applied to a layered skin model.

3. Three Optical Models for Uncovering Blood Vessel Patterns Hidden in Color Images

To uncover blood vessels hidden in color images, three optical models that simulate skin color formation are developed. Each model consists of a camera model, an illuminant model, a skin structure, and a method to compute the total reflectance from the skin structure. Though our previous method has a number of problems (see the introduction) [11], we still employed it as a baseline model because with the two optimization schemes presented in this section, its performance can be significantly improved. The baseline model and the two optimization schemes together are regarded as the first model. In the second and third models, Reichman's solution to the radiative transfer equation is used to replace the K-M theory in the first model [14]. In the first and second models, a three-layered skin structure constituted by the stratum corneum, the epidermis, and the dermis is used, while in the last model, the hypodermis is also included to form a four-layered skin structure. Subsection 3.1 presents the general idea based on the optical models for uncovering blood vessels. Subsection 3.2 summarizes the baseline model for completeness. Subsection 3.3 describes the two optimization schemes to overcome the weaknesses of the baseline model. Subsection 3.4 describes the other two optical models.

3.1 From Skin Color Formation to Skin Parameter Estimation

The proposed optical models estimate biophysical parameters in the skin. In the first two models, the spatial distributions of the volume fraction (%) of the epidermis occupied by melanosomes (organelles containing melanin), the volume fraction (%) of the dermis occupied by blood, and the depth of the dermis are considered. They are denoted as ϑ_m, ϑ_p, and d^{der}, respectively. Intuitively, blood vessels can be found in ϑ_p because they contain a higher concentration of blood than other skin components. In the last model, the depth of the blood vessel d^v and the depth of adipose tissue d^a are included. For the sake of mathematical convenience, d^v and d^a are combined and represented by one symbol, d^{hyp}, which is called the depth of the hypodermis and let $\Lambda_1 = [\vartheta_m, \vartheta_p, d^{der}]$, $\Lambda_2 = [\vartheta_m, \vartheta_p, d^{der}]$, and $\Lambda_3 = [\vartheta_m, \vartheta_p, d^{der}, d^{hyp}]$. The computational details of d^{hyp}, d^v, and d^a are given in Section 3.5.

In this section, three mathematical functions based on the three optical models are sought to estimate the biophysical parameters, Λ_i, where blood vessels are expected to be seen. These mathematical functions can be represented as

$$\Lambda_i = g_i(R, G, B, E_i), \tag{3}$$

where $i \in \{1,2,3\}$; R, G, and B represent the three color components in a given image; and E_i represents other information. E_i includes the prior knowledge from a camera, an illuminant, and a skin structure and its biophysical parameters. The subscript i is used to emphasize that different g_i uses different prior knowledge. The illuminant is characterized by its spectrum $Z(\lambda)$, and the camera is characterized by its R, G, and B spectral response functions, which are denoted as $S_R(\lambda)$, $S_G(\lambda)$, and $S_B(\lambda)$, respectively, where λ represents wavelengths. $Z(\lambda)$, $S_R(\lambda)$, $S_G(\lambda)$, and $S_B(\lambda)$ are regarded as known variables. Section 3.2 will discuss how to obtain these functions in legal cases. To establish g_i, the skin color formation in an image should be modeled. Mathematically, it can be represented as

$$[R, G, B] = f_i\big(Z(\lambda), \Lambda_i, S_R(\lambda), S_G(\lambda), S_B(\lambda)\big). \tag{4}$$

In Eq. 4, other prior knowledge, e.g., the molecular weight of eumelanin, is considered internal constants in f_i. If it was a bijective function, Λ_i could be computed from

$$[Z(\lambda), \Lambda_i, S_R(\lambda), S_G(\lambda), S_B(\lambda)] = f_i^{-1}(R, G, B). \tag{5}$$

Because $Z(\lambda)$, $S_R(\lambda)$, $S_G(\lambda)$, and $S_B(\lambda)$ are known, Eq. 5 can be simplified as $\Lambda_i = f_i^{-1}(R, G, B)$. In general, f_i is not a bijective function. A function g_i is thus sought to approximate f_i^{-1} such that

$$\Lambda_i \approx g_i(R, G, B). \tag{6}$$

E_i in Eq. 3 is omitted because it is regarded as a constant. Eq. 4 is called a forward model, which simulates the skin color formation, and Eq. 6 is called a backward model, which is used to estimate the skin parameters, where blood vessels can be seen.

Given a forward model f_i and the domain of Λ_i (i.e., the ranges of ϑ_m, ϑ_p, and d^{der} for the first two models and the ranges of ϑ_m, ϑ_p, d^{der}, and d^{hyp} for the third model), the corresponding Λ_i and $[R, G, B]$ can be generated. Using a regression method and the Λ_i and $[R, G, B]$ obtained from the forward model as training data, g_i, which estimates Λ_i based on R, G, and B values in color images, can be derived. A neural network with 5 neurons in the hidden layer is used to produce g_i for all the optical models. Fig. 2 illustrates the relationship between the forward and backward models.

3.2 Camera and Illuminant Models

To realize the forward model f_i, three components—the illuminant $Z(\lambda)$, the total reflectance of the skin $R_T(\lambda)$, and the spectral response functions of the camera $S_R(\lambda)$, $S_G(\lambda)$, and $S_B(\lambda)$—are necessary. Once these three components are known, the corresponding R, G, and B values can be computed by

$$\tau = \int_0^{\infty} Z(\lambda) R_T(\lambda) S_\tau(\lambda) d\lambda, \tag{7}$$

where $\tau \in \{R, G, B\}$. Though the integration range in Eq. 7 is from 0 to ∞, spectral response functions of consumer cameras only have non-zero responses at visible light wavelengths between 400 nm and 700 nm. Other wavelengths such as infrared are blocked by the RGB filter in front of the sensor. It is worth mentioning that to capture good-quality blood vessel images, NIR in the range between 800 nm and 930 nm is recommended [16]. In forensic investigation, the spectral response functions can be known through camera models, which can be found in the EXIF headers of evidence images. If the headers are removed or tampered with, camera identification methods can be applied to identify the camera model [17]. The use of these methods is outside the scope of this paper.

Three types of light sources are generally encountered in daily life: daylight, incandescent lamps, and fluorescent lamps. They correspond respectively to the standard illuminants D65, A, and F, defined by the International Commission on Illumination [18]. In addition to the standard illuminants, forensic scientists may be able to reconstruct crime scenes so that a more accurate $Z(\lambda)$ can be measured directly. In our experiments, the standard illuminant D65 was used.

Now, we know how to obtain $Z(\lambda)$, $S_R(\lambda)$, $S_G(\lambda)$, and $S_B(\lambda)$. Once the total reflectance of the skin $R_T(\lambda)$ is known, the forward model f_i can be established. Subsections 3.3-3.5 present three optical models based on different optical theories and skin structures to approximate $R_T(\lambda)$.

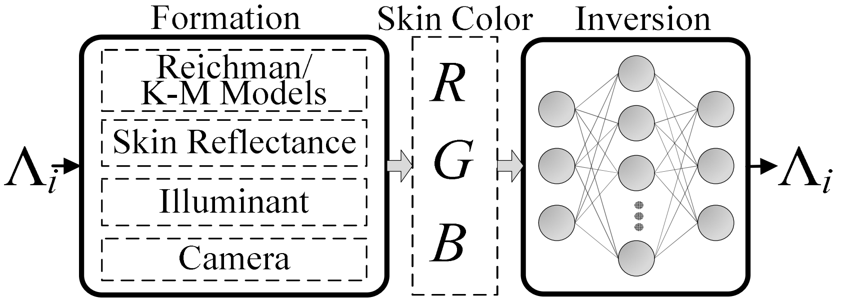


Fig. 2 A schematic diagram illustrating the blood vessel uncovering process based on the optical models. The left block represents the skin color formation model, and the right block represents the inversion of the skin color formation to obtain Λ_i from $[R, G, B]$. They are, respectively, the graphical representation of f_i and g_i in Eqs. 4 and 6.

3.3 The Baseline Model

The baseline model uses a three-layered skin structure composed of the stratum corneum, the epidermis, and the dermis. A fraction of the incident light is reflected by the stratum corneum, and the rest penetrates into the epidermis and the dermis (Fig. 1). In the epidermal and dermal layers, the light is scattered multiple times and absorbed by bilirubin, melanin, β-carotene, and hemoglobin [19]. In this model, the K-M theory and the recursive equations (Eqs. 1-2) are used to calculate $R_T(\lambda)$. According to the K-M theory, $R_T(\lambda)$ is determined by the absorption and scattering coefficients and the depths of the stratum corneum, the epidermis, and the dermis. The details of these coefficients can be found in [11].

3.4 Two Optimization Schemes

3.4.1 Skin Parameter Range Optimization

The baseline model takes the skin parameter ranges directly from medical literature to be the domain of ϑ_m, ϑ_p, and d^{der} in the training. These ranges may not be suitable for forensic analysis because they were obtained in ideal medical settings from a particular race (e.g., Caucasian) and a particular body site (e.g., hand). The three optical models use an optimization scheme to determine the domain of Λ_i. Our image database mainly consists of Asian skin, so we selected a typical skin image from the thigh of an Asian subject and used its average RGB values as a target to compute the corresponding optimal Λ_i, which is denoted as Λ_i^*. We then linearly discretized the domain formed by $[0.5\Lambda_i^* \;\; 2\Lambda_i^*]$ and used the forward model to calculate the corresponding RGB values. The trained neural network approximates the inversion of the skin color formation. To uncover blood vessel patterns in a testing image, its RGB values are input to the trained neural network and the outputs are collected to form the spatial distribution maps of the biophysical parameters.

We did not compute the optimal Λ_i for each pixel in an image because the optimization process is very time-consuming. Using MATLAB on a PC embedded with an Intel® Core™2 Quad processor (3.0 GHz), it takes approximately 50 hours to process an image with a size of 200×300 pixels. However, using neural network mapping, it takes only one second.

3.4.2 An Automatic Adjustment Scheme for Illumination Intensity Variation

Illuminants significantly influence skin color in images. Though several standard illuminants can be used as the spectrum of $Z(\lambda)$, its power is determined by the power of the illuminant source and the distance between the source and the skin, which can vary greatly. To avoid manually estimating the power, an automatic adjustment scheme is proposed.

Eq. 7 indicates that the influence from the power variation can be eliminated by adjusting the image intensity. Let $Z(\lambda)$ and $aZ(\lambda)$ be two illuminants with the same spectrum but different power. According to Eq. 7, their pixel values are $\tau = \int_0^\infty Z(\lambda)R_T(\lambda)S_\tau(\lambda)d\lambda$ and $a\tau = \int_0^\infty aZ(\lambda)R_T(\lambda)S_\tau(\lambda)d\lambda$, respectively, where $\tau \in \{R, G, B\}$. Their color values differ by a factor of a, and therefore, we can eliminate the power variation by adjusting the image intensity.

Though all spatial distribution maps of the parameters show blood vessel patterns, the distribution maps of d^{der} and d^{hyp} are the clearest. They are used in the intensity adjustment and also later in automatic matching. An intensity adjustment ratio a is introduced in the computation, i.e.,

$$d^{layer} = g_{layer}(C/a), \tag{8}$$

where d^{layer} represents either d^{der} or d^{hyp}, C is a color skin image and g_{layer} represents the neural network function mapping RGB values to d^{der} or d^{hyp}. The average local variance of the d^{layer} map is used as an objective function to determine the optimal adjustment ratio a^*:

$$a^* = \underset{a \in A}{\operatorname{argmax}} \left\{ \frac{1}{N} \sum_{i=1}^{N} \underset{(x,y) \in M_{b_i}}{\operatorname{var}} \left[d_{b_i}^{layer}(x, y) \right] \right\}, \tag{9}$$

where A is a set of different intensity adjustment ratios, $d_{b_i}^{layer}$ is the i^{th} 5×5 block in d^{layer}, N is the total number of blocks, (x, y) is the position of a pixel, and M_{b_i} is a set of skin pixels in $d_{b_i}^{layer}$. Note that only skin pixels are considered in this calculation.

3.5 The Other Two Models

3.5.1 The Second Optical Model

The first optical model utilizes the K-M theory to approximate the transport of light in human skin. A comparative study based on the Monte Carlo simulation of human skin showed that the K-M theory is very sensitive to the K-M coefficients, which are very difficult to measure accurately because of the complexity of the skin structure [14]. To overcome this problem, the solution to the radiative transfer equation based on the Schuster-Schartzchild approximation derived by Reichman is employed in the second and third optical models [14]. When the incidence is normal, the reflectance and transmittance of this solution can be computed by:

$$R = \frac{\omega_0}{2} \frac{3\Gamma + (1 - \Gamma^2)\exp[-(1+\alpha)\gamma d] - 1 - \Gamma(3 - \Gamma)\exp(-2\alpha\gamma d)}{(1 - \alpha^2)[1 - \Gamma^2 \exp(-2\alpha\gamma d)]}, \tag{10}$$

$$T = \frac{\omega_0}{2} \frac{\Theta}{(1 - \alpha^2)[1 - \Gamma^2 \exp(-2\alpha\gamma d)]} + \exp(-\gamma d), \tag{11}$$

where $\Theta = 3(1 - \Gamma^2)\exp(-\alpha\gamma d) - (3 - \Gamma)\exp(-\gamma d) + \Gamma(3\Gamma - 1)\exp[-(1 + 2\alpha)\gamma d]$, $\gamma = \mu_a + \mu_s, \omega_0 = \mu_a/(\mu_a + \mu_s)$, $\alpha = 2\sqrt{1 - \omega_0}$, $\Gamma = (2 - \alpha)/(2 + \alpha)$, $\mu_a(\lambda)$ and $\mu_s(\lambda)$ are the absorption and scattering coefficients of a layer at a particular wavelength, and d is the thickness of the layer. The advantage of this solution over the K-M theory is that it does not rely on the K-M coefficients [20].

The second optical model is still based on the three-layered skin structure composed of the stratum corneum, the epidermis, and the dermis. For each of these three layers, the reflectance and transmittance are calculated based on Reichman's equations (Eqs. 10-11). The computational processes of the first model and the second model are the same, except that the K-M theory in the first model is replaced with Reichman's equations. The two optimization schemes proposed in Subsection 3.4 are also adopted in this model.

3.5.2 The Third Optical Model

In the first and second optical models, we assume that the optical properties of human skin are determined by three layers – the stratum corneum, the epidermis, and the dermis, and blood vessels are located in the dermis. However, skin anatomy shows that on many occasions, blood vessels are located under the dermis and inside the hypodermis [21]. In our third optical model, one more skin layer, the hypodermis consisting of adipose tissue and blood vessels, is added. Thus, the new skin structure has four layers in total – the stratum corneum, the epidermis, the dermis, and the hypodermis. Using the recursive equations (Eqs. 1-2) and the four-layered skin structure, the total reflectance $R_{1234}(\lambda)$ and transmittance $T_{1234}(\lambda)$ of the new skin structure can be computed.

The variables $R_i(\lambda)$ and $T_i(\lambda)$ $(i = 1,2,3,4)$ required in the recursive equations are the reflectance and transmittance, respectively, of the i^{th} layer of the skin. The 1st-4th layers are the stratum corneum, the epidermis, the dermis, and the hypodermis, respectively. Reichman's solution is still utilized to calculate the reflectance and transmittance of each layer. The calculations of the absorption and scattering coefficients of the stratum corneum, the epidermis, and the dermis are the same as the second optical model. This model assumes that the hypodermis is composed of adipose tissue and blood vessels, and blood vessels are located just under the dermis. The absorption and scattering coefficients of the adipose tissue are obtained from the experimental data [22].

With the same illuminant and sensor models, the skin color formation based on the four-layered model can be simulated. In addition to ϑ_m, ϑ_p, and d^{der}, two more parameters, d^v and d^a, are added to represent the depth of blood vessels and adipose tissue, respectively. They are combined and represented by one symbol, d^{hyp}, which is called the depth of the hypodermis. Let the range of d^v be $[0, \alpha]$ and the range of d^a be $[0, \beta]$. The range of d^{hpy} is set to be $[-\alpha, \beta]$. When d^{hpy} is less than zero, its absolute value representing d^v is used to calculate the reflectance and transmittance of blood vessels; otherwise, it represents d^a and can be used directly to calculate the reflectance and transmittance of adipose tissue. This model assumes that blood vessels and adipose tissue cannot coexist in the same column in the fourth layer. ϑ_m, ϑ_p, d^{der}, and d^{hyp} are sampled exhaustively in the domain, and the corresponding RGB values can be obtained. The two optimization schemes are also adopted in this model.

4. Automatic Blood Vessel Matching

Low-quality blood vessel images[1] from the optical models are unavoidable because of the weak penetration capability of visible light and the high concentration of fat in some body sites. Blood vessel patterns extracted from these low-quality images can be noisy and partially overlapping. The proposed automatic extraction and matching algorithms are specially designed to address these challenges. Subsection 4.1 presents the automatic blood vessel extraction algorithm, which utilizes directional groups to enhance patterns in low-quality images. Subsection 4.2 describes the matching algorithm with explicit noise and outlier removal

---

[1] In this section, blood vessel images refer to the spatial distribution maps of the biophysical parameters.

steps. Subsection 4.3 gives two fusion schemes to utilize information from different optical models for enhancing the matching performance.

4.1 Blood Vessel Extraction based on Directional Groups

The proposed extraction algorithm includes four components: preprocessing, local information estimation, enhancement and representation. In preprocessing, contrast-limited adaptive histogram equalization (CLAHE) is first used to normalize the contrast of input images [23]. This method can avoid amplifying noise in homogeneous regions and greatly improve the overall contrasts of the input images, especially for NIR images. Examples are given in Fig. 3b. Then a filter bank composed of the real parts of 16 Gabor filters with different scales and orientations are applied to NIR images and distribution maps of d^{layer} to capture their local information. A blood vessel segment can be regarded as a dark ridge, and therefore, only the real parts are used. The direct current (DC) components of the Gabor filters are removed to enhance their robustness against brightness variation, and their powers are normalized for accurately estimating local orientation. Let G_{DC} be a normalized zero DC Gabor filter and $Img(x,y)$ be an NIR image or a distribution map of d^{layer} from one of the three optical models. $F_{\lambda_{mk},\theta_k,\sigma_m,\gamma}$ denotes a filtered $Img(x,y)$, which can be obtained from

$$F_{\lambda_{m,k},\theta_k,\sigma_m,\gamma}(x,y) = -G_{DC}(x,y,\lambda_{mk},\theta_k,\sigma_m,\gamma) * Img(x,y), \tag{12}$$

where * represents an operation of a two-dimensional convolution.

To capture local orientation and estimate the information quality of a blood vessel segment, $F_{\lambda_{mk},\theta_k,\sigma_m,\gamma}$ is further processed. The orientation of a blood vessel segment can be estimated from

$$O(x,y) = \arg_{\theta_k} \max_{m,k} F_{\lambda_{mk},\theta_k,\sigma_m,\gamma}(x,y)/P_m(x,y), \tag{13}$$

which is called the orientation map, and the quality of the local information can be estimated from

$$R(x,y) = \max_{m,k} F_{\lambda_{mk},\theta_k,\sigma_m,\gamma}(x,y)/P_m(x,y), \tag{14}$$

which is called the response map. P_m is the local image power around the point (x, y) for further minimizing the influence of image contrast. Note that P_m depends on the scale of the filter. Figs. 3(c) and (d) give orientation and response maps of four images.

To enhance blood vessel patterns and suppress noise, the information in orientation maps is operated in groups. First, a labeled map $L(x, y) \in \{1, \cdots, n\}$ with the same size as $R(x, y)$ is produced by a connected component labeling method [24]. Pixels in $L(x, y)$ with the same label i $(i = 1, \cdots, n)$ are connected and share the same orientation $o(i)$. They are called a component. If $R(x, y)$ is greater than a threshold, the corresponding pixel is classified as a potential blood vessel pixel, and the corresponding component is then denoted as a potential blood vessel component. Blood vessel components are further connected based on their orientation difference and spatial distance. For every component pair, if their angular distance is shorter than or equal to $\pi/8$ and their spatial distance is shorter than a threshold, they are connected. The threshold is determined by experiments. A large threshold connects more components together and is more effective for restoring broken blood vessels, while a small threshold links fewer components and is more robust to noise (e.g., hair). There is a trade-off in setting this threshold. All elements in the connected component pairs are searched iteratively and assigned the same label if they share the same component. This iterative process generates a map of grouped labels. A weighting function is defined based on the map [12] and the pointwise multiplication is applied to the weighting function and $R(x, y)$ to enhance blood vessel patterns (Fig. 3f).

To retain the geometric relationship among blood vessel segments and keep a simple structure, point representation is employed. We do not use minutiae features, which have been considered for recognizing infrared hand vein images [8], because blood vessels of other body sites from color images are noisy and the bifurcation and ending points are unreliable. Otsu's method [25] is used to binarize the enhanced blood vessel images and a skeletonization method is applied to obtain their structures, which are then uniformly sampled and represented by point sets. In our experiment, each blood vessel pattern is finally represented by a set of points with equal distance. The average number of points is 227. Forearm boundaries are also extracted and sampled in the same manner to improve matching performance, while thigh boundaries are not used because their exposed skin regions vary considerably (Fig. 6). Four sampled blood vessel patterns are given in Fig. 3g.

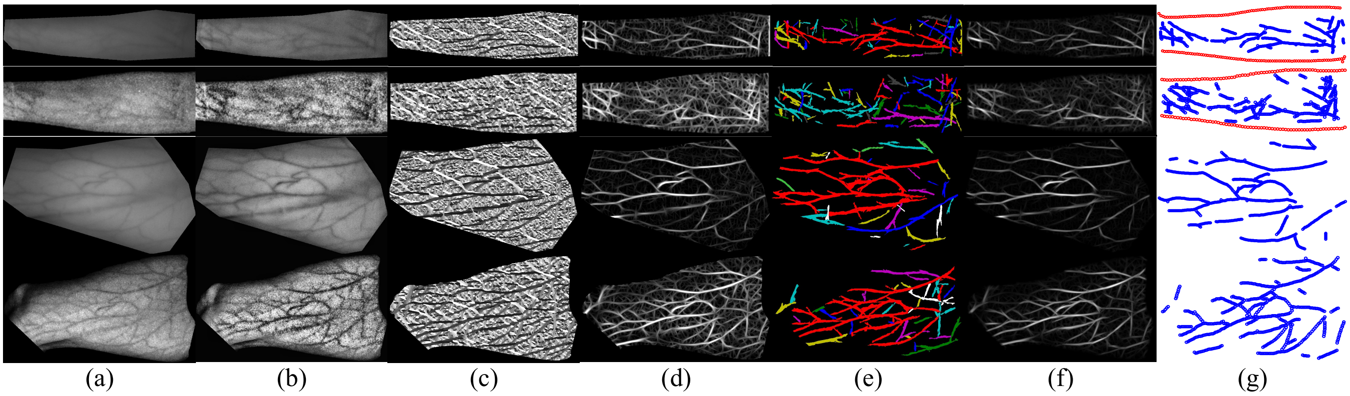


Fig. 3 The blood vessel pattern extraction processes. The first and third rows are NIR images and the second and fourth rows are uncovered results from color images. Column (a) is the input images; column (b) is the results after the local brightness adjustment; column (c) is the orientation maps generated by the Gabor filters; column (d) is the response maps; column (e) is the maps of grouped blood vessel components (different colors indicating different groups); column (f) is the enhanced response maps; and column (g) is the point sets for matching. (a color figure)

4.2 Blood Vessel Pattern Matching with a Noise Removal Scheme

Blood vessel pattern matching based on the point set representation can be regarded as a point registration problem, which has been extensively studied. Among all point registration methods, the Iterative Closest Point (ICP) [26] is the best known one, and it has many variants, such as LM-ICP [27]. ICP and most of its variants are designed for rigid point set registration. However, they are not suitable for our problem because of the non-linear distortion in the blood vessel patterns. Non-rigid point registration methods such as Robust Point Matching (RPM) [28] and Coherent Point Drift (CPD) [29] are more desirable. Directly applying existing methods to match blood vessel patterns is not an optimal approach because it neglects characteristics of the problem, including partial overlap, noisy patterns, missing blood vessel segments, viewpoint differences, and non-linear distortion. In preliminary experiments, we tested some existing hand vein matching methods, but no encouraging results were obtained. Thus, a new blood vessel matching algorithm is developed.

A set of notations is given first for clear presentation. Let M and D be two point sets from one of the optical models. M and D are called model and data patterns, respectively. Each point in M, e.g., $(x, y) \in M$, is associated with a response value $R_M(x, y)$ and an orientation value $O_M(x, y)$, where R_M and O_M are the response and orientation maps of M. For forearms, the point sets representing their boundaries are denoted as M_{BD} and D_{BD}. Transformations of the point sets are denoted as $T(\bullet, \tau)$, controlled by a parametric vector τ.

In general, a point registration algorithm determines the parametric vector τ in a transformation T, which is applied to the model set, i.e., $T(M, \tau)$, so that the two point sets are aligned optimally and their dissimilarity can be measured. The proposed matching algorithm utilizes explicit schemes to remove outliers and noise, a non-rigid point registration method to align remaining points, and their response and orientation information to determine the dissimilarity between two patterns.

When matching forearms, the boundaries M_{BD} and D_{BD} are first used to remove points outside the overlapping area of M and D. A rigid transformation T_R is used to describe the deformation between M_{BD} and D_{BD}. The rigid Coherent Point Drift (CPD) method is used to determine a parametric vector τ_1. Fig. 4(c) shows two boundaries and their alignment result. The point set M is transformed into $M_{t1} = T_R(M, \tau_1)$. Then, the overlapping area B of M_{t1} and D is determined. Points that are located outside B are regarded as outliers and removed to increase robustness. The pruned point sets are denoted as M_1 and D_1. If the numbers of points in these two sets are less than a threshold, they are assumed to be from two different blood vessel patterns. Fig. 4(d) shows two point sets after outliner removal. Because thigh boundaries in our database are unreliable, we set $M_1 = M$ and $D_1 = D$ when matching them.

Then, the matching algorithm removes outliers and noise inside M_1 and D_1. The rigid transformation T_R is used again to roughly align M_1 and D_1. Note that M_1 is regarded as a model set in this alignment. Using the rigid CPD method and the resultant parametric vector τ_2, the correspondence c_2 between M_1 and D_1 can be obtained. The point set $M_{t2} = T_R(\mathrm{M}_1, \tau_2)$ is roughly aligned with D_1, and the distances between the corresponding points in M_{t2} and D_1 are calculated. If the distances are very long, the corresponding points in M_{t2} are considered as outliers. All points in M_{t2} are sorted in ascending order according to their distances, and the top 80% of the points in M_{t2} are retained. The remaining points are denoted as M_2. The same process is repeated to remove outliers in D_1. The rigid CPD method is applied to align D_1 and M_2. In this alignment, D_1 is considered as a model set, and a transformed point set $D_{t2} = T_R(D_1, \tau_2')$ is computed. As with the previous case, only 80% of the points in D_{t2} are retained, and distances between corresponding points in D_{t2} and M_2 are used as a selection criterion. The remaining points in D_{t2} are denoted as D_2.

The final step is to match M_2 and D_2 with a non-rigid transformation. The non-rigid CPD method with a Gaussian kernel as a regularizer is used to match M_2 and D_2. With a non-rigid transformation T_N and corresponding parametric vector τ_3 generated by the CPD method, $M_{t3} = T_N(M_2, \tau_3)$, and D_2 are aligned optimally. As with the previous steps, a correspondence vector c_3 is obtained. Note that multiple points in M_2 can correspond to the same point in D_2 and that data points with no correspondence are recognized as outliers.

The proposed dissimilarity function is based on the distance between the two aligned point sets and the corresponding response and orientation differences. Let $m^i = (m_x^i, m_y^i)$ be the i^{th} point in M_{t3}; its corresponding response and orientation values be $M_{t3}^R(m^i)$ and $M_{t3}^O(m^i)$, respectively; the corresponding point in D_2 be $d^{c_3(i)} = (d_x^{c_3(i)}, d_y^{c_3(i)})$; and its corresponding response and orientation values be $M_2^R(d^{c_3(i)})$ and $D_2^O(d^{c_3(i)})$. The dissimilarity function is defined as

$$s(M,D) = |I|^{-1} \times \sum_{i\in I} \left\| m^i - d^{c_3(i)} \right\| \times \left| M_{t3}^R(m^i) - M_2^R(d^{c_3(i)}) \right| \times \Omega\left(M_{t3}^O(m^i), D_2^O(d^{c_3(i)})\right), \qquad (15)$$

where I is an index set containing points with dissimilarity values defined as $\left\| m^i - d^{c_3(i)} \right\| \times \left| M_{t3}^R(m^i) - M_2^R(d^{c_3(i)}) \right| \times \Omega\left(M_{t3}^O(m^i), D_2^O(d^{c_3(i)})\right)$ in the 80th percentile.

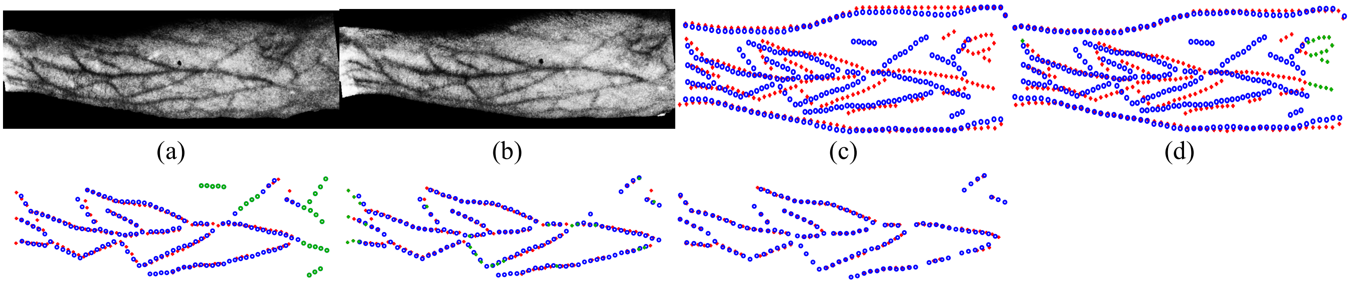

(a) (b) (c) (d)
(e) (f) (g)

Fig. 4 Illustration of blood vessel matching. (a) and (b) are resultant images from an optical model. (c) shows the points extracted from (a) and (b). Red and blue dots represent model and data points. (d) illustrates M_{t1} and D_1. Outliers outside the overlapping area are marked in green. (e) illustrates M_{t2} and D_1. Green points are regarded as noise. (f) shows the alignment result of M_2 and D_{t2}, and (g) is the final non-rigid alignment result. (a color figure)

4.3 Fusion of Blood Vessel Patterns from the Three Models

Section 3 presents three optical models based on two skin structures and two approximate solutions to the radiative transfer equation for simulating light passing through skin. To utilize information provided by different optical models simultaneously for improving matching performance, a feature-level fusion scheme and a score-level fusion scheme are proposed.

4.3.1 A Feature-Level Fusion Scheme based on the Noise Removal Scheme

The optical models provide similar major blood vessels, but different minor blood vessels and noise (Fig. 8). If their point sets are fused directly through a union operator, the fused patterns will be very noisy. Before applying the union operator, the noise removal scheme in the matching algorithm is used to suppress noise and retain blood vessel patterns.

Let C_i be an input color image and its corresponding point sets from the three optical models be $M^{i1}, M^{i2},$ and M^{i3}, and let C_j be another color image and its corresponding point sets from the three optical models be $D^{j1}, D^{j2},$ and D^{j3}. The proposed algorithm is used to match M^{ik} and D^{ik}, where $k \in \{1,2,3\}$, to obtain the noise-suppressed point sets M_2^{ik} and D_2^{jk}, which are denoted as M_2 and D_2, respectively, in the previous section. The proposed fusion scheme applies the union operator to these point sets. Mathematically, this fusion scheme is defined as

$$F_i = \bigcup_{k=1,2,3} M_2^{ik}. \tag{16}$$

The fused point set F_j is obtained in the same way. Then, the scheme performs the non-rigid CPD matching on the fused point sets F_i and F_j and computes their dissimilarity.

4.3.2 A Score-Level Fusion Scheme based on the Weighted Sum

The other fusion scheme uses the weighted sum to combine the dissimilarity values from the optical models. A new dissimilarity based on this fusion scheme is defined as

$$s_{fuse}(i,j) = \frac{\sum_{k=1}^{3} w^k * S(M^{ik}, D^{jk})}{\sum_{k=1}^{3} w^k}, \quad (17)$$

where w^k is $0.8 \times |M_2^{ik}|$. Note that $|M_2^{ik}|$ is the cardinality of M_2^{ik} and that w^k is the size of the index set in Eq. 15. The fusion schemes presented in these two subsections are also applicable to match a color image with an infrared image. We simply replace D^{j1}, D^{j2}, and D^{j3} with the point set from an infrared image. Thus, the two fusion schemes can be applied without any modification.

5. Experimental Results

To evaluate the proposed optical models, the fusion schemes, and the matching algorithm, NIR and color images were collected from the inner forearms and thighs of 328 subjects[2], mainly Chinese, Indians, and Malays in Singapore. 301 of them are males. The NIR images are used for comparisons and validation. The proposed models and algorithms do not depend on them. Two data collection sessions were carried out 11 days apart on average. The maximum time gap between the two sessions was 81 days. During each session, two color images and two NIR images were taken from each body site. In total, 16 color images and 16 NIR images were collected from each subject. Two consumer cameras (Canon 500D and Nikon D70s) and one NIR camera (JAI-AD080CL) were used. The maximum resolutions of the two consumer cameras are 4752 by 3168 and 3008 by 2000 pixels, respectively, and the maximum resolution of the NIR camera is 1024 by 768 pixels. Though the subjects were expected to be present in both sessions, 29 subjects were present only in one session. At each body site, approximately 60 subjects have large tattoos and/or very dense androgenic hair covering their blood vessels. Even in NIR images, their blood vessels are not observable. It is well-known that tattoos are regularly used by law enforcement agencies for criminal verification and identification and tattoo retrieval methods have been developed [30]. Recently, a study shows that androgenic hair patterns can be used for personal identification [7]. Thus, for these images, blood vessel recognition is not necessary. In addition to tattoo-obscured and hairy images, some unusable images, such as

---

[2] A database will be shared with the biometric community for this research direction. Download URL will be provided before Mar 31, 2015.

incorrect body sites and images with motion blur or incorrect infrared illumination were also removed. Images collected from wrong poses were removed because images from the same person in the gallery set and the probe set do not form a pair for matching. Color images whose corresponding NIR images were captured under wrong illumination were removed because we cannot validate our uncovering results through the NIR images. In this study, motion blur is not considered because we target on still images, e.g., child sexual abuse images. Fig. 5 shows some of the removed images. In real cases, law enforcement agencies may not have more than one image of a suspects or a criminal for matching. Thus, if a subject had more than one color (NIR) image from a particular body site in any of the sessions, one of them was randomly selected to form the testing database. The testing database contained 1900 NIR images and 1900 color images from 250 right forearms, 240 left forearms, 230 right thighs, and 230 left thighs. Although some images were removed, many challenging images collected under different illumination conditions, viewpoints, and poses still remained in the database. Fig. 6 shows some typical images in the testing database.

The color images in the testing database were processed semi-automatically. A rectangular region containing the body site was manually selected, and then the skin region was segmented automatically based on skin color. The NIR images were manually segmented. The segmented regions were finally rotated and resized. Eigenvalues computed from segmented skin regions were used to perform rotation. The average sizes of the processed color (NIR) forearm and thigh images were 990 by 383 pixels (643 by 259 pixels) and 1118 by 636 pixels (687 by 449 pixels). Fig. 8 shows some of the processed images. After applying the models presented in Section 3, the color forearm and thigh images were downsampled. Consequently, the average sizes of the distribution maps from the forearms and the thighs for matching were 495 and 191 pixels and 559 and 318 pixels, respectively. In the experiments, blood vessel patterns were extracted from the distribution maps of d^{layer} generated from the optical models.

In the rest of this section, four sets of experimental results are reported. The blood vessel patterns from the NIR images were matched to evaluate the proposed matching algorithm (Subsection 5.1). The blood vessel patterns from the color images were matched to examine the proposed optical models and the fusion schemes

(Subsection 5.2). The blood vessel patterns from the color images were also matched with those from the NIR images to evaluate their similarity (Subsection 5.3). One of the proposed optical models was applied to hands, arms, thighs, chests, breasts, and abdomens of men, women, and children in indoor and outdoor images collected from the Internet to evaluate their effectiveness on images taken from completely uncontrolled environments (Subsection 5.4). In the first three experiments, the images collected during the first session were regarded as a gallery set, and the images collected during the second session were regarded as a probe set. Cumulative match curves (CMCs) were used as a performance index in these experiments.

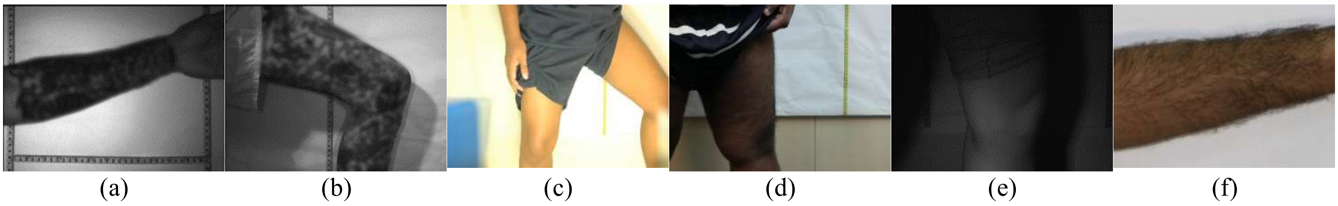


Fig. 5 Examples of the removed images. (a)-(b) large tattoos, (c) blur and wrong pose, (d) blur and incorrect illumination, (e) incorrect NIR illumination, and (f) dense androgenic hair.

Fig. 6 Typical images in our testing database. The first and second rows show color images collected during the first and second sessions, respectively. The third row shows corresponding NIR images from the first session. The images in each column were collected from the same subject.

5.1 Matching Blood Vessel Patterns from NIR Images

To evaluate the proposed blood vessel matching algorithm, we compared it with the LM-ICP method [27], the original CPD method [29], and our preliminary matching method [12]. NIR images were selected for this comparison because NIR has a higher penetration capability and is expected to offer the best-quality blood vessel patterns. The original CPD method with a non-rigid transformation was selected for this comparison because it is the state-of-the-art point matching method with the ability to handle non-rigid deformation and also because the proposed matching algorithm depends on it. In the experiments, the original CPD method and the proposed matching algorithm used the same parameters. When matching thigh images, the preliminary matching algorithm did not use their boundaries. The LM-ICP method was also included in this comparison because of its popularity.

The first column of Fig. 7 shows the CMCs from different matching methods and different body sites. Their rank-1 and rank-10 accuracy is given in Table 1. The term "rank-10 identification accuracy" refers to the percentage of input blood vessel patterns whose corresponding patterns can be found in the database within the top 10% of the patterns given by a matching algorithm. The proposed matching algorithm achieved rank-1 accuracy of 66.8%, 80.4%, 57.9%, and 56.9% for the left forearms, the right forearms, the left thighs, and the right thighs, respectively. It clearly outperforms the original CPD, the LM-ICP, and the preliminary matching method for all body sites. These results demonstrate that the proposed outlier and noise removal schemes are effective, that blood vessel patterns have serious non-rigid distortion, and that the proposed new dissimilarity function and other techniques for improving the preliminary matching method are useful. The proposed matching algorithm was thus used in the rest of the experiments.

5.2 Matching Blood Vessel Patterns from Color Images

In this experiment, blood vessel patterns extracted from color images were matched. Before offering CMCs as an objective evaluation, blood vessel patterns generated from different optical models are given in Fig. 8 for visual comparison. The first and second columns are color and NIR images, respectively. The third column contains the results from the OBVU method. The fourth, fifth, and sixth columns are the results from the first, second, and third optical models, respectively. This figure shows that the proposed optical models can effectively uncover blood vessel patterns hidden in color images and perform significantly better than the OBVU method. The visual quality of the blood vessel patterns from the proposed optical models is comparable with that from the NIR images. Fig. 8 shows that the major blood vessels from different optical models are the same, but the minor blood vessels can differ. It is difficult to visually conclude which is the best optical model.

Table 1 Rank-1 and Rank-10 accuracies from matching blood vessel patterns from NIR images (measured in percentages).

		Left Arm	**Right Arm**	**Left Thigh**	**Right Thigh**
Rank-1 accuracy	LM-ICP	10.0	10.8	5.7	2.2
	Original-CPD	51.9	62.4	46.1	50.2
	Preliminary matching method	62.2	76.0	53.1	53.3
	Proposed matching algorithm	66.8	80.4	57.9	56.9
Rank-10 accuracy	LM-ICP	28.6	32.8	21.9	18.7
	Original-CPD	61.8	67.6	62.3	59.1
	Preliminary matching method	72.2	82.8	64.5	66.2
	Proposed matching algorithm	76.8	86.0	70.6	70.7

The underlining indicates the best matching results.

To objectively compare the optical models, the fusion schemes, and the OBVU method, their blood vessel patterns from the color images were matched. In this experiment, we only matched blood vessel patterns from the same optical model or the same fusion scheme. No cross model matching was performed. The corresponding CMCs are given in the second column of Fig. 7, and the corresponding rank-1 and rank-10 accuracy is given in Table 2. The NIR and red-channel images were also included for comparison. NIR images are expected to provide blood vessel patterns with the highest quality and to be an upper performance limit. Red-channel images are always available in color images and red light is close to NIR in terms of wavelength. The CMCs in the second column of Fig. 7 show that the three proposed optical models outperform the red channel and the OBVU method with significant margins. Because both the first optical model and the OBVU method use the same skin structure and the K-M theory, these experimental results indicate that the proposed optimization schemes can effectively handle illumination variation and avoid over-relying on the biophysical parameters measured in ideal medical conditions. Among the three optical models, the first and second optical models perform similarly for all the body sites. The third optical model performs slightly better than the other two models for the forearms, but slightly worse for the thighs. The CMCs also demonstrate the effectiveness of the fusion schemes. The feature-level and score-level fusion schemes both outperform any individual optical model. In terms of rank-1 accuracy, the feature-level fusion scheme outperforms the score-level fusion scheme for all body sites. In terms of rank-10 accuracy, the feature-level fusion scheme outperforms the score-level fusion scheme for all body sites, except for the right forearms. The CMCs of the fusion schemes are comparable with those of NIR images, which are regarded as an upper performance limit. In terms of rank-1 and rank-10 accuracy, the feature-level fusion scheme performs better than the NIR images for the right forearms. For the other body sites, the average differences of rank-1 and rank-10 accuracy between the feature-level fusion scheme and the NIR images are 2.8% and 1.8%, respectively. The results obtained from the forearms are better than those from the thighs because the skin of forearms is thinner and light has a higher chance to penetrate it. Some ROIs of the left forearms are smaller than those of the right forearms, because of watches and accessories on the left forearms. Thus, the right forearms perform better.

Table 2 A summary of the matching accuracy (measured in %)

		Matching the same type of images (e.g., NIR vs. NIR and color images vs. color images)				Matching different types of images (color images vs. NIR images)			
		Left Arm	Right Arm	Left Thigh	Right Thigh	Left Arm	Right Arm	Left Thigh	Right Thigh
Rank-1 accuracy	NIR	66.8	80.4	57.9	56.9				
	Model OBVU	29.7	37.6	13.6	14.5	23.1	26.4	4.3	2.3
	Model 1	52.3	68.0	46.5	46.5	35.5	49.6	13.9	14.0
	Model 2	52.7	68.0	47.8	46.9	37.2	47.6	16.5	10.7
	Model 3	52.7	70.4	39.9	43.4	35.0	42.0	16.1	10.7
	Red channel	44.8	38.4	26.3	31.6	33.8	28.4	7.8	7.0
	Feature fusion	63.6	*81.2	54.8	54.8	44.4	54.4	16.5	13.0
	Score fusion	57.7	76.0	51.8	48.2	35.9	48.4	13.9	10.2
Rank-10 accuracy	NIR	76.8	86.0	70.6	70.7				
	Model OBVU	46.0	54.8	30.7	30.3	37.6	33.2	14.8	10.7
	Model 1	66.9	81.6	59.6	61.0	50.4	60.8	28.3	27.9
	Model 2	67.8	80.4	59.6	59.6	51.7	58.0	30.0	26.5
	Model 3	68.6	83.2	51.3	55.3	52.1	59.2	28.7	24.2
	Red channel	61.9	59.6	47.4	49.6	46.6	43.6	20.4	22.3
	Feature fusion	76.2	*86.8	69.3	67.1	54.7	63.6	30.0	28.4
	Score fusion	74.5	*89.6	66.7	62.7	55.6	60.8	32.2	25.1

* denotes that the results are better than the corresponding results from NIR images. The underlining indicates the best results from matching color images with color images and matching color images with NIR images.

5.3 Matching Blood Vessel Patterns from Color Images with those from NIR Images

The previous experimental results have demonstrated that the proposed optical models can uncover blood vessel patterns hidden in color images and that these patterns can be matched for personal identification. Fig. 8 shows that blood vessel patterns from color and NIR images are not exactly the same. In this experiment, blood vessel patterns from color images were matched with those from NIR images. The last column of Fig.

7 gives the corresponding CMCs, and Table 2 lists the corresponding rank-1 and rank-10 accuracy. As with the previous results, the three optical models outperform the OBVU method and the red channel. The feature-level fusion scheme provides the highest rank-1 accuracy for all body sites, except for the right thighs. In terms of rank-10 accuracy, the two fusion schemes perform similarly. The CMCs indicate that the blood vessel pattern similarity between the color images and the NIR images is higher than that between the red channels and the NIR images. Comparing with the corresponding CMCs in the second column of Fig. 7, significant performance drops are noted. They indicate that the blood vessel patterns from NIR images and those from optical models have some differences. In fact, these differences are also observable in Fig. 8. The optical models tend to provide more blood vessels than NIR images, especially for the thighs. Some may expect that this cross spectral matching requires some mechanisms, e.g., domain adaption. Because blood vessels appear in both the NIR images and the uncovered results, direct matching can be performed.

5.4 Uncovering Blood Vessels from Internet Images

In addition to uncovering blood vessels from images in the database, the optical models were applied to images collected from the Internet. They were taken under uncontrolled environments with large lighting, pose and viewpoint variations. The subjects include male and female, adult and children. Figs. 9 and 10 show the uncovering results from the first optical model. The faces were pixelated to protect the people's privacy. Because some of these images have blocking artifacts, the parameter a in Eq. 8 was selected manually. To give more results from different images, results from the second and third models are not given. Though these images were taken in uncontrolled environments, blood vessels are still uncovered by the proposed model. In addition to forearms and thighs, blood vessels in other body sites, such as neck, chest, upper arms, and legs become visible in the resultant images. These results demonstrate clearly the feasibility of using blood vessels uncovered by the proposed optical models for criminal and victim authentication.

5.5 Uncovering Blood Vessels from Facial Images

To our knowledge, no research has been done on facial vein pattern identification, since there are fewer veins on face than in other body parts. Most of current research utilizes facial marks and other prominent features for identification. Fig.11 shows an NIR facial image obtained from the Internet. We can find that compared with shoulder and chest, face has much fewer veins. The first optical model was applied to two facial images. Fig. 12 shows the uncovering results. For privacy protection, partial faces were blurred. It can be seen that the optical model visualizes some veins.

6. Discussion

Identifying criminals and victims in evidence images is challenging and sometimes is not possible with current technology because faces, tattoos, and distinctive skin mark patterns are not observable. Traditionally, blood vessel patterns could not be used for criminal and victim verification because they are nearly invisible in color images. This paper presents three optical models, two optimization schemes, blood vessel extraction and matching algorithms, and two fusion schemes to eliminate this traditional barrier. The proposed models, algorithms, and schemes were examined on a database containing 1900 color images and 1900 NIR images from 328 subjects' forearms and thighs. The experimental results demonstrate the effectiveness of the approach, and most importantly, they show that blood vessel patterns hidden in color images can be uncovered for personal identification. In addition to images in this database, the proposed models were also applied to hands, arms, thighs, chests, breasts, and abdomens of men, women, and children in indoor and outdoor images collected from the Internet. They were also applied to facial images. The results are very encouraging.

However, we admit that the proposed optical models are not yet perfect. They cannot uncover blood vessel patterns from some color images, especially when blood vessels in the corresponding NIR images are unclear or even unobservable. The proposed models are based on blood vessel information hidden in color images. If color images do not carry any of this information, no model will work. However, we cannot conclude now whether the color images have no blood vessel information or the proposed models cannot make use of it to visualize the blood vessels. Compared with the accuracy of other well-developed biometric techniques, e.g., fingerprint, iris and face identification, which have been studied for more than two decades, the recognition accuracy of blood vessel patterns is low. We target at forensic identification where no fingerprint, iris or face is available. In forensic applications, quality of fingerprints and face images can also be very low and the state-of-the-art algorithms cannot always handle them well. Because uncovering blood vessels from color images is an extremely new research direction, we believe that significant progress will be made in the coming years.

In the experiments, tattoo-obscured and hairy images were removed. In legal cases, once tattoos are available, the identification process becomes straightforward. Though dense androgenic hair can cover other biometric traits, such as skin marks and blood vessels [6], androgenic hair patterns can in fact be used for personal identification directly [7]. Identification and verification using skin marks, blood vessels, hair, and tattoos all have some weaknesses. These biometric traits should be used simultaneously. Fusion of different biometrics is an effective solution. In addition to the challenges from the availability of biometric traits, image quality is always an issue in forensic verification and identification. More research should be performed on uncovering blood vessels from images collected in low-light environments and larger pose and viewpoint variations. Skin marks and blood vessels requiring middle- to high-resolution images are more suitable for child sexual abuse images and images of other sexual offenses because close-up images are expected in these types of legal cases. However, identifying masked criminals based on exposed skin in low-resolution images still remains a challenge [7]. In addition to resolution, evidence images and videos are always compressed by the JPEG and MPEG methods [31]. How to restore the skin features from these low-quality images and videos is also an important problem. Algorithms, especially for restoring veins from JPEG images are also demanded. Furthermore, a huge database with long time-lapse body image pairs is also needed for the further research.

In this paper, we have demonstrated that blood vessel patterns can be uncovered from color images for personal identification and verification. However, significant research is still demanded to address all these problems for forensic verification and identification.

Acknowledgements

Hengyi Zhang and Adams Wai-Kin Kong are supported by the Ministry of Education, Singapore, through Academic Research Fund Tier 2, MOE2012-T2-1-024, and Chaoying Tang is supported by the National Natural Science Foundation of China (61403196), the Natural Science Foundation of Jiangsu Province (BK20140837), and the Specialized Research Fund for the Doctoral Program of Higher Education of China (20133218120018).

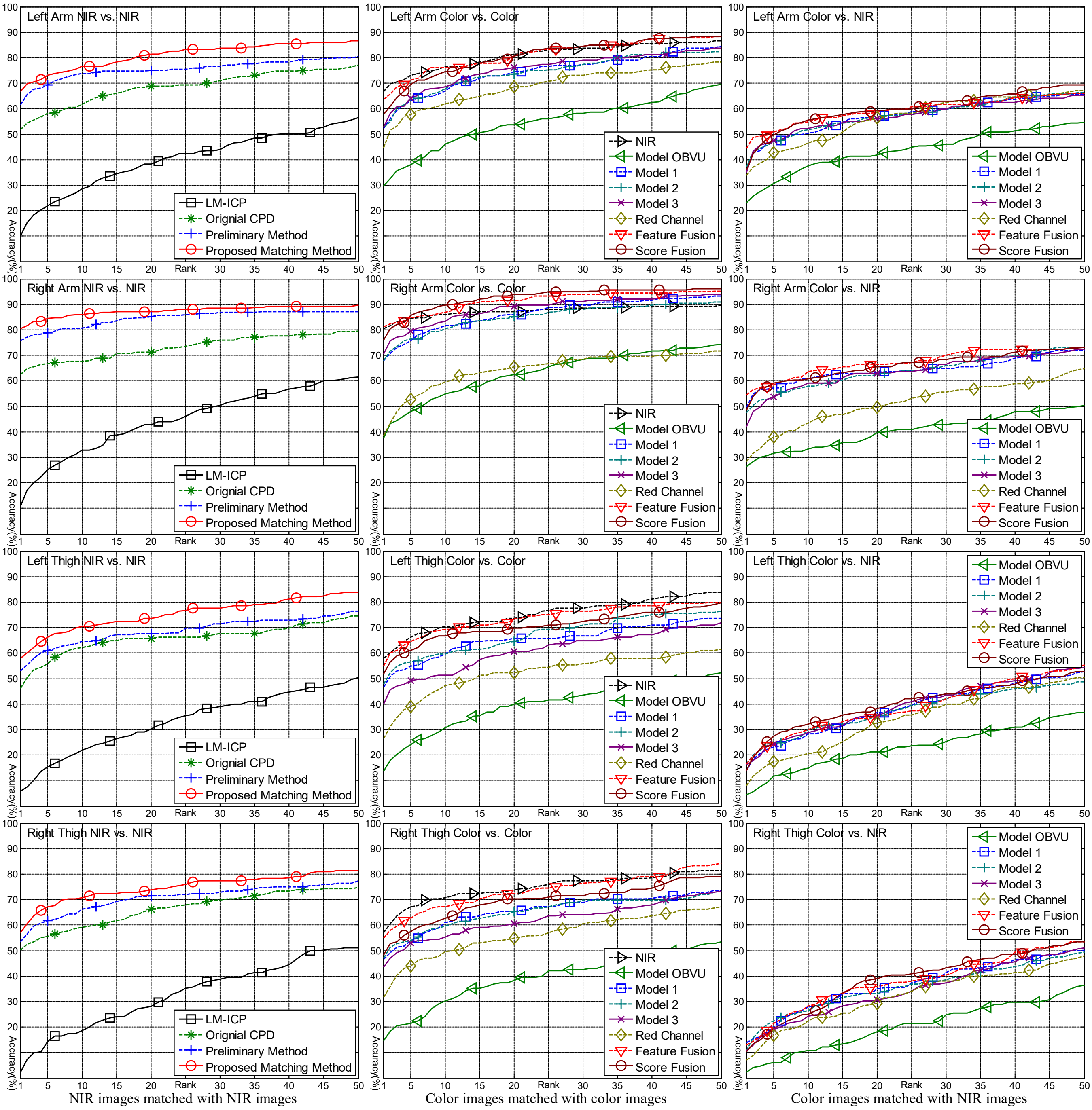


Fig. 7 CMCs generated from matching results. The first column compares the proposed matching algorithm with other matching methods based on NIR images. The second column compares different optical models and fusion schemes based on color images. The third column shows the CMCs from matching color images with NIR images. (a color figure)

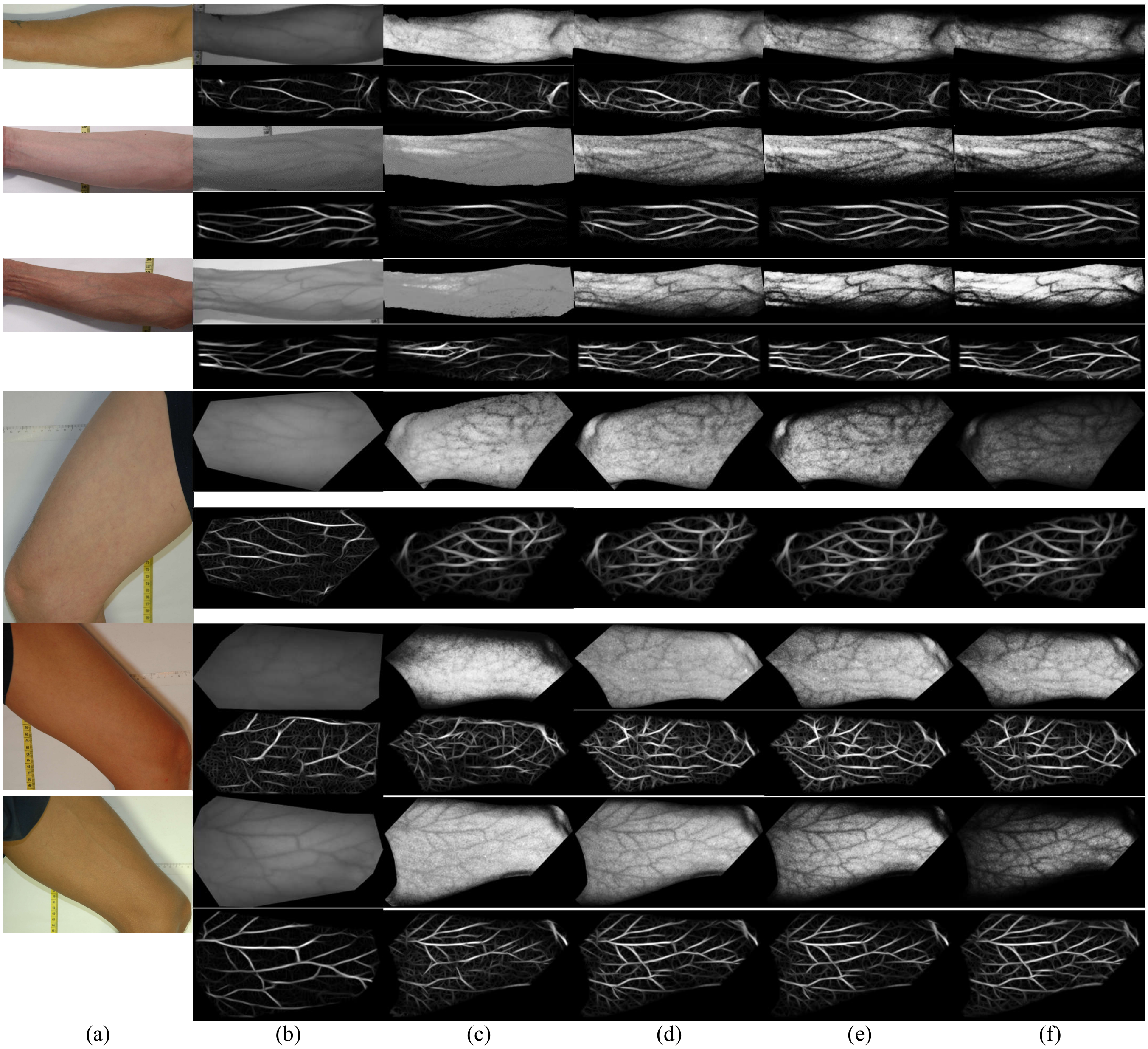


Fig. 8 Resultant images from the optical models and the corresponding enhanced response maps. Column (a) contains color images; column (b) contains NIR images; columns (c)-(f) contain results from the OBVU and the first, second, and third optical models, respectively. The enhanced response maps are shown under the processed images.

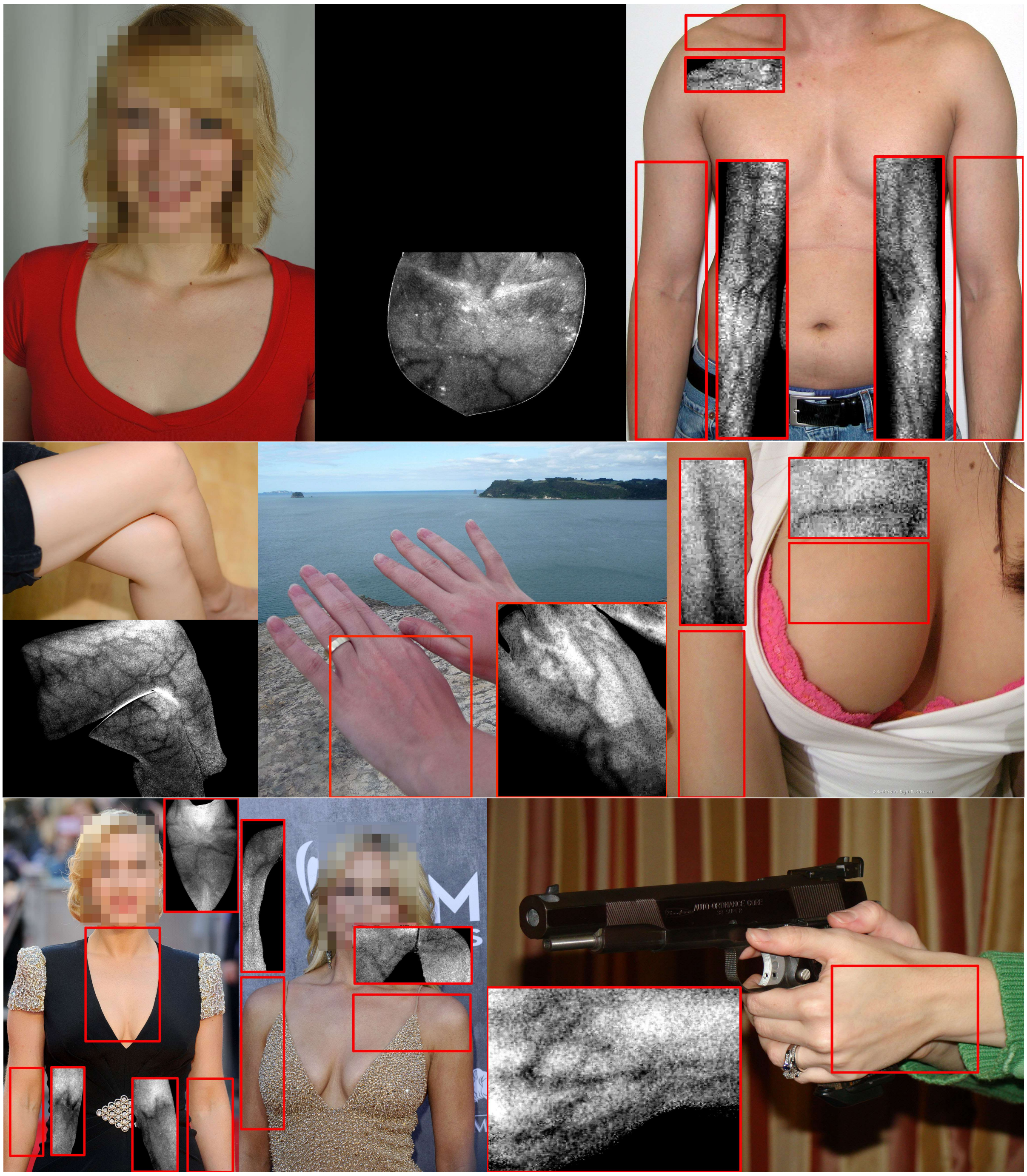


Fig. 9 Results from Internet images. Each pair of red rectangles indicates a skin region and the corresponding result from the first optical model. (To see clearly the images, please read the electronic version.)

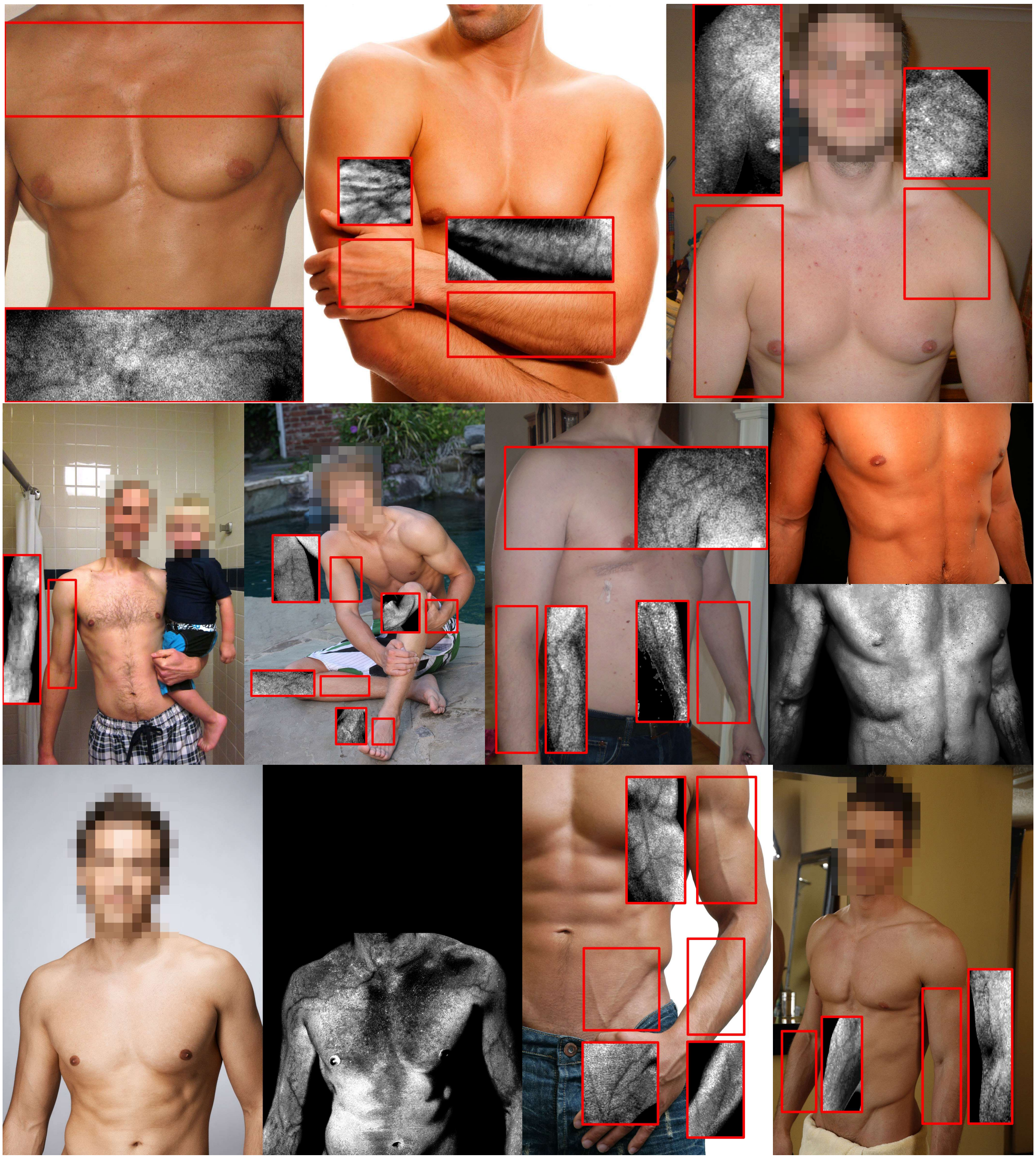
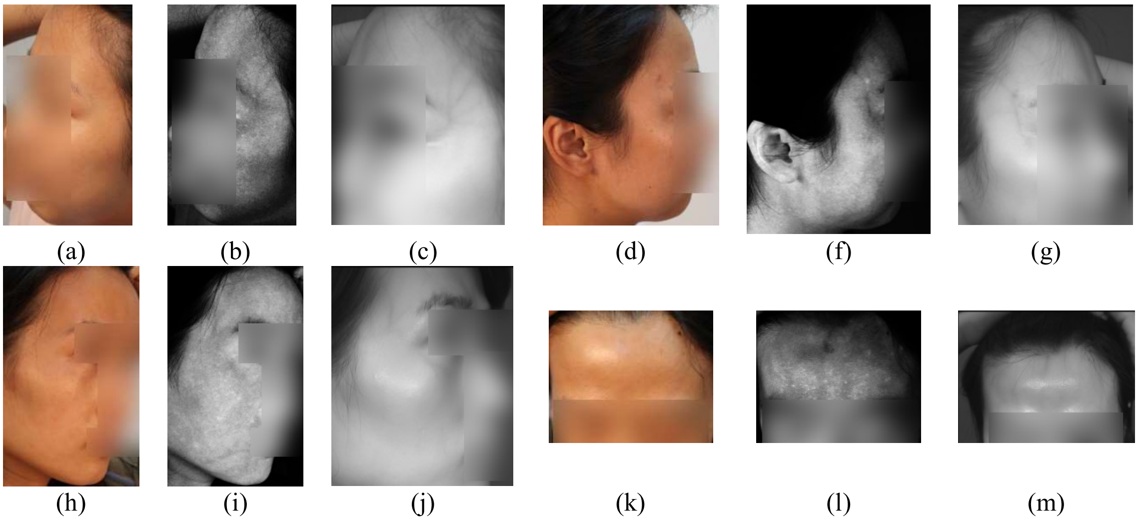


Fig. 10 Results from Internet images. Each pair of red rectangles indicates a skin region and the corresponding result from the first optical model. (To see clearly the images, please read the electronic version.)

Fig. 11 An NIR facial image collected from the Internet.



(a) (b) (c) (d) (f) (g)

(h) (i) (j) (k) (l) (m)

Fig. 12 Uncovering results obtained from some facial images. (a), (d), (h) and (k) are color facial images; (b), (f), (i) and (l) are their corresponding uncovered results obtained from optical model 1; (c), (g), (j) and (m) are the NIR images. For privacy protection, partial faces are blurred.

References

[1] J. Loughlin and A. Taylor-Butts, Child luring through the Internet, Juristat, 29 (2009).

[2] K. Bunzeluk, Child sexual abuse images–analysis of websites by Cybertip. ca, ed: Canadian Centre for Child Protection, 2009.

[3] AFP, Australia sees explosion in online child porn, Asiaone News, ed, Feb, 08, 2012.

[4] M. Motivans and T. Kyckelhahn, Federal prosecution of child sex exploitation offenders, 2006, Washington, DC: US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, 2007.

[5] The U.S. Attorney's Office, Central District of California, Ex-marine Guilty of Using Drugs and Force to Have Sex with Young Girls in Cambodia, 29 May 2008. Available: http://www.justice.gov/usao/cac/Pressroom/pr2008/074.html

[6] A. Nurhudatiana, A.W.K. Kong, K. Matinpour, D. Chon, L. Altieri, S.Y. Cho and N. Craft, The individuality of relatively permanent pigmented or vascular skin marks (RPPVSM) in independently and uniformly distributed patterns, TIFS, 8 (2013), 998-1012.

[7] H. Su and A.W.K. Kong, A study on low resolution androgenic hair patterns for criminal and victim identification, TIFS, 9 (2014), 666-680.

[8] L. Wang, G. Leedham, and S. Y. Cho, Minutiae feature analysis for infrared hand vein pattern biometrics, Pattern Recognition, 41 (2008), 920-929.

[9] Fujitsu-Laboratories-Ltd. (2003, March). Fujitsu laboratories develops technology for world's first contactless palm vein pattern biometric authentication system. Available: http://pr.fujitsu.com/en/news/2003/03/ 31.html.

[10] J.G. Wang, W.Y. Yau, A. Suwandy and E. Sung, "Person recognition by fusing palmprint and plam vein images based on "Laplacianpalm" representation", 41 (2007), 1514-1527.

[11] C. Tang, A. Kong, and N. Craft, Uncovering vein patterns from color skin images for forensic analysis, CVPR (2011), 665-672.

[12] H. Zhang, C. Tang, A. Kong, and N. Craft, Matching vein patterns from color images for forensic investigation, BTAS (2012), 77-84.

[13] E. Claridge, S. Cotton, P. Hall, and M. Moncrieff, From colour to tissue histology: Physics-based interpretation of images of pigmented skin lesions, Medical Image Analysis, 7 (2003), 489-502.

[14] J. Reichman, Determination of absorption and scattering coefficients for nonhomogeneous media. 1: Theory, Applied Optics, 12 (1973), 1811-1815.

[15] A. Sawetprawichkul, A study of the transient radiative transfer within participating media, Florida Institute of Technology, 2002.

[16] V. P. Zharov, S. Ferguson, J. F. Eidt, P. C. Howard, L. M. Fink, and M. Waner, Infrared imaging of subcutaneous veins, Lasers in Surgery and Medicine, 34 (2004), 56-61.

[17] T. Van Lanh, K. S. Chong, S. Emmanuel, and M. S. Kankanhalli, A survey on digital camera image forensic methods, ICME (2007), 16-19.

[18] G. Wyszecki and W. S. Stiles, Color Science: Wiley New York, 1982.

[19] C. Donner, T. Weyrich, E. d'Eon, R. Ramamoorthi, and S. Rusinkiewicz, A layered, heterogeneous reflectance model for acquiring and rendering human skin, Siggraph Asia, 2008, article no. 140.

[20] S. Avrillier, E. Tinet, and E. Delettre, Monte Carlo simulation of collimated beam transmission through turbid media, Journal de Physique*,* 51 (1990), 2521-2542.

[21] A. Michelle, J. Hopkins, C. W. McLaughlin, S. Johnson, M. Q. Warner, D. LaHart*,* and J.D. Wright, Human biology and health, Englewood Cliffs, New Jersey, USA: Prentice Hall, 1993.

[22] A. Bashkatov, E. Genina, V. Kochubey, and V. Tuchin, Optical properties of the subcutaneous adipose tissue in the spectral range 400–2500 nm, Optics and spectroscopy*,* 99 (2005), 836-842.

[23] K. Zuiderveld, Contrast limited adaptive histogram equalization, Graphics gems, IV (1994), 474-485.

[24] R. M. Haralick and L. G. Shapiro, Computer and robot vision, vol. I: Addison-Wesley, 1992.

[25] N.Otsu, , A Threshold Selection Method from Gray-Level Histograms, TSMC, 9 (1979), 62-66.

[26] P. J. Besl and N. D. McKay, A method for registration of 3-D shapes, TPAMI*,* 14 (1992), 239-256.

[27] A. Fitzgibbon, Robust registration of 2D and 3D point sets, Image and Vision Computing, 21 (2003), 1145-1153.

[28] H. Chui and A. Rangarajan, A new point matching algorithm for non-rigid registration, Computer Vision and Image Understanding*,* 89 (2003), 114-141.

[29] A. Myronenko and S. Xubo, Point set registration: coherent point drift, TPAMI, 32 (2010), 2262-2275.

[30] J.E. Lee, R. Jin, A.K. Jain and W. Tong, Image retrieval in forensics: tattoo image database application, IEEE Multimedia (2012), 40-49.

[31] C. Tang, A. Kong, and N. Craft, Using a knowledge-based approach to remove blocking artifacts in skin images for forensic analysis, TIFS, 6 (2011), 38-1049.

[32] W. Song, T. Kim, H.C. Kim, J. H. Choi, H. J. Kong and S. R. Lee, A finger-vein verification system using mean curvature, Pattern Recognition Letters, 32 (2011), 1541-1547.

[33] J. Yang and Y. Shi, Finger-vein ROI localization and vein ridge enhancement, Pattern Recognition Letters, 33 (2012), 1569-1579.

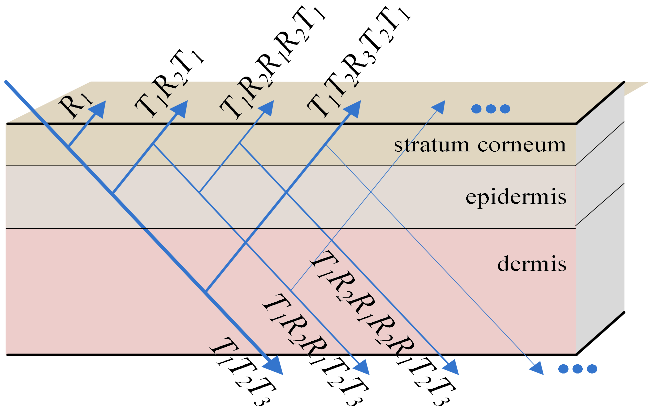


Fig. 1 Illustration of the recursive equations (Eqs. 1-2) applied to a layered skin model.

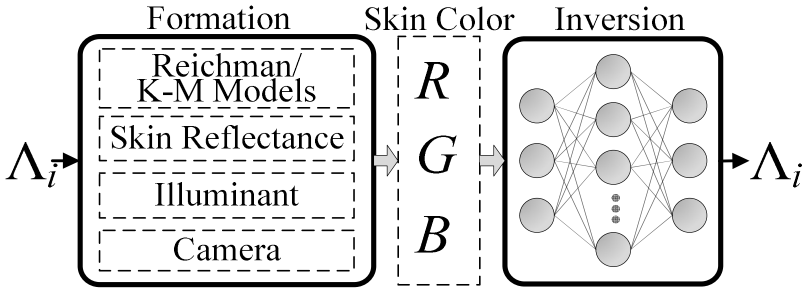
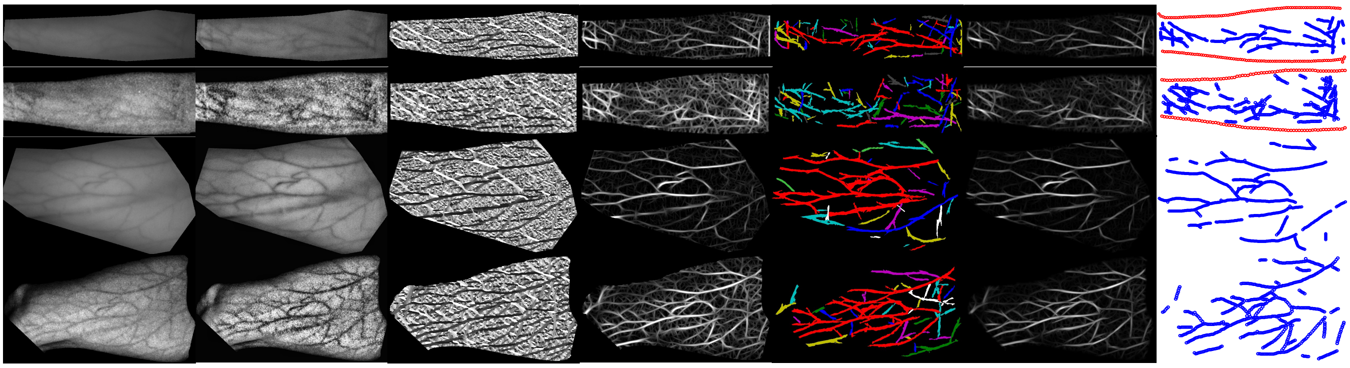
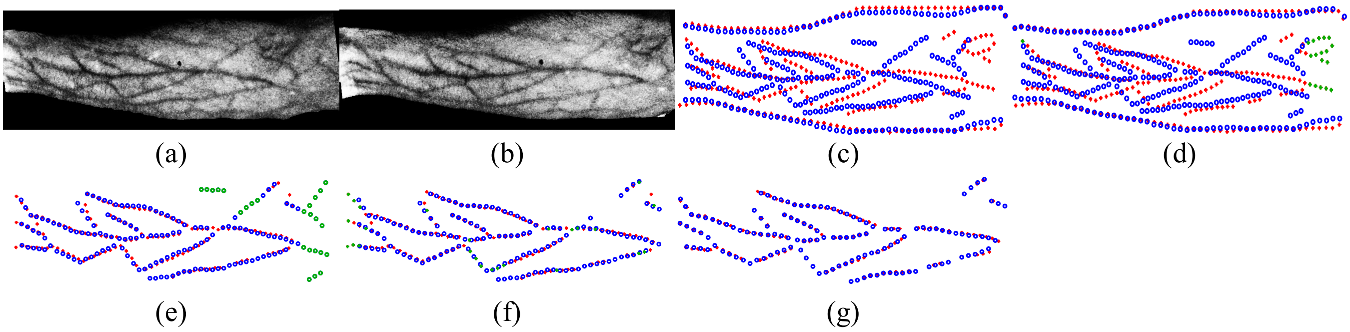


Fig. 2 A schematic diagram illustrating the blood vessel uncovering process based on the optical models. The left block represents the skin color formation model, and the right block represents the inversion of the skin color formation to obtain Λ_i from $[R, G, B]$. They are, respectively, the graphical representation of f_i and g_i in Eqs. 4 and 6.



(a) (b) (c) (d) (e) (f) (g)

Fig. 3 The blood vessel pattern extraction processes. The first and third rows are NIR images and the second and fourth rows are uncovered results from color images. Column (a) is the input images; column (b) is the results after the local brightness adjustment; column (c) is the orientation maps generated by the Gabor filters; column (d) is the response maps; column (e) is the maps of grouped blood vessel components (different colors indicating different groups); column (f) is the enhanced response maps; and column (g) is the point sets for matching. (a color figure)


(a) (b) (c) (d)

(e) (f) (g)

Fig. 4 Illustration of blood vessel matching. (a) and (b) are resultant images from an optical model. (c) shows the points extracted from (a) and (b). Red and blue dots represent model and data points. (d) illustrates M_{t1} and D_1. Outliers outside the overlapping area are marked in green. (e) illustrates M_{t2} and D_1. Green points are regarded as noise. (f) shows the alignment result of M_2 and D_{t2}, and (g) is the final non-rigid alignment result. (a color figure)

(a) (b) (c) (d) (e) (f)

Fig. 5 Examples of the removed images. (a)-(b) large tattoos, (c) blur and wrong pose, (d) blur and incorrect illumination, (e) incorrect NIR illumination, and (f) dense androgenic hair.

Fig. 6 Typical images in our testing database. The first and second rows show color images collected during the first and second sessions, respectively. The third row shows corresponding NIR images from the first session. The images in each column were collected from the same subject.

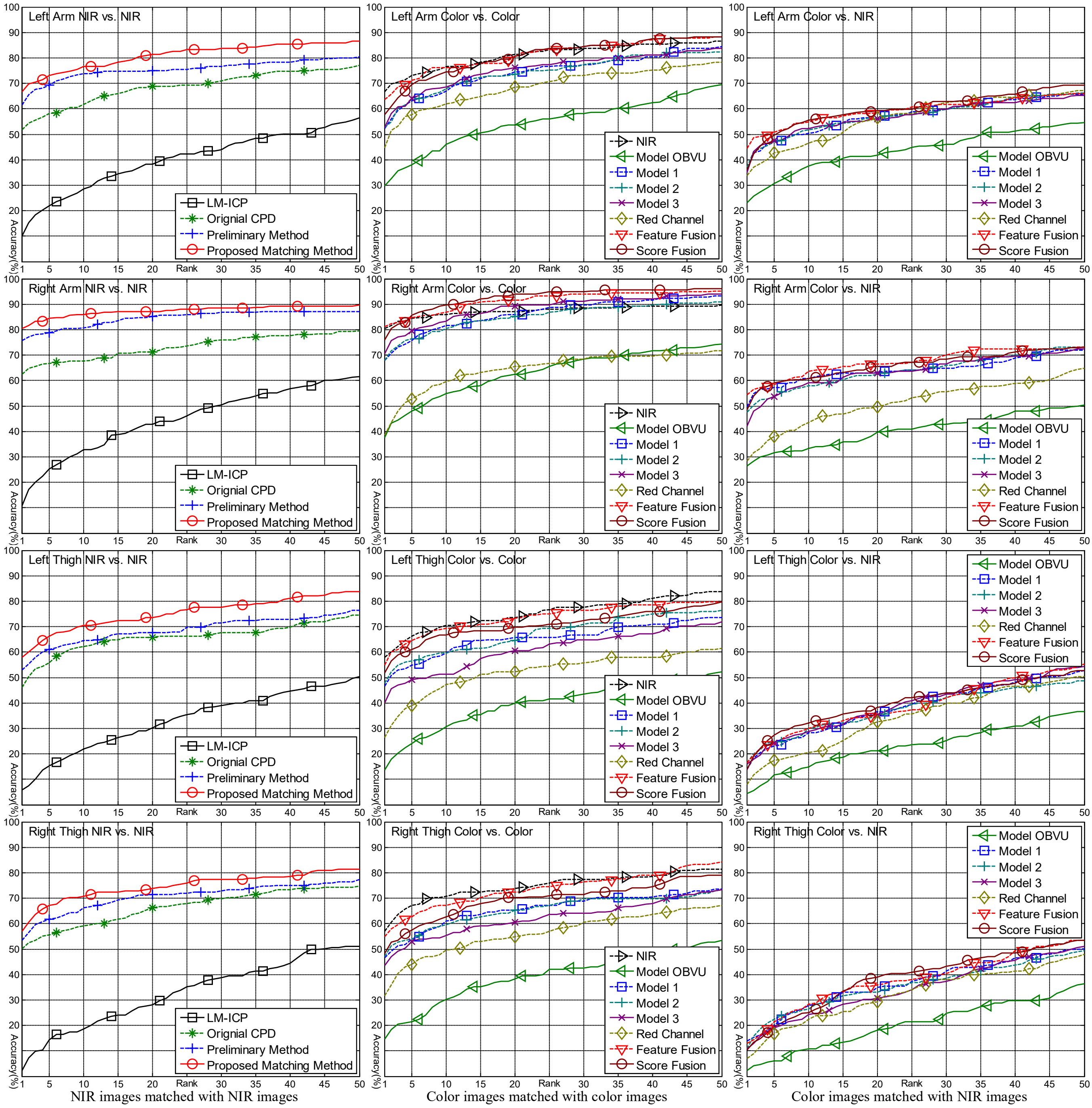


Fig. 7 CMCs generated from matching results. The first column compares the proposed matching algorithm with other matching methods based on NIR images. The second column compares different optical models and fusion schemes based on color images. The third column shows the CMCs from matching color images with NIR images. (a color figure)

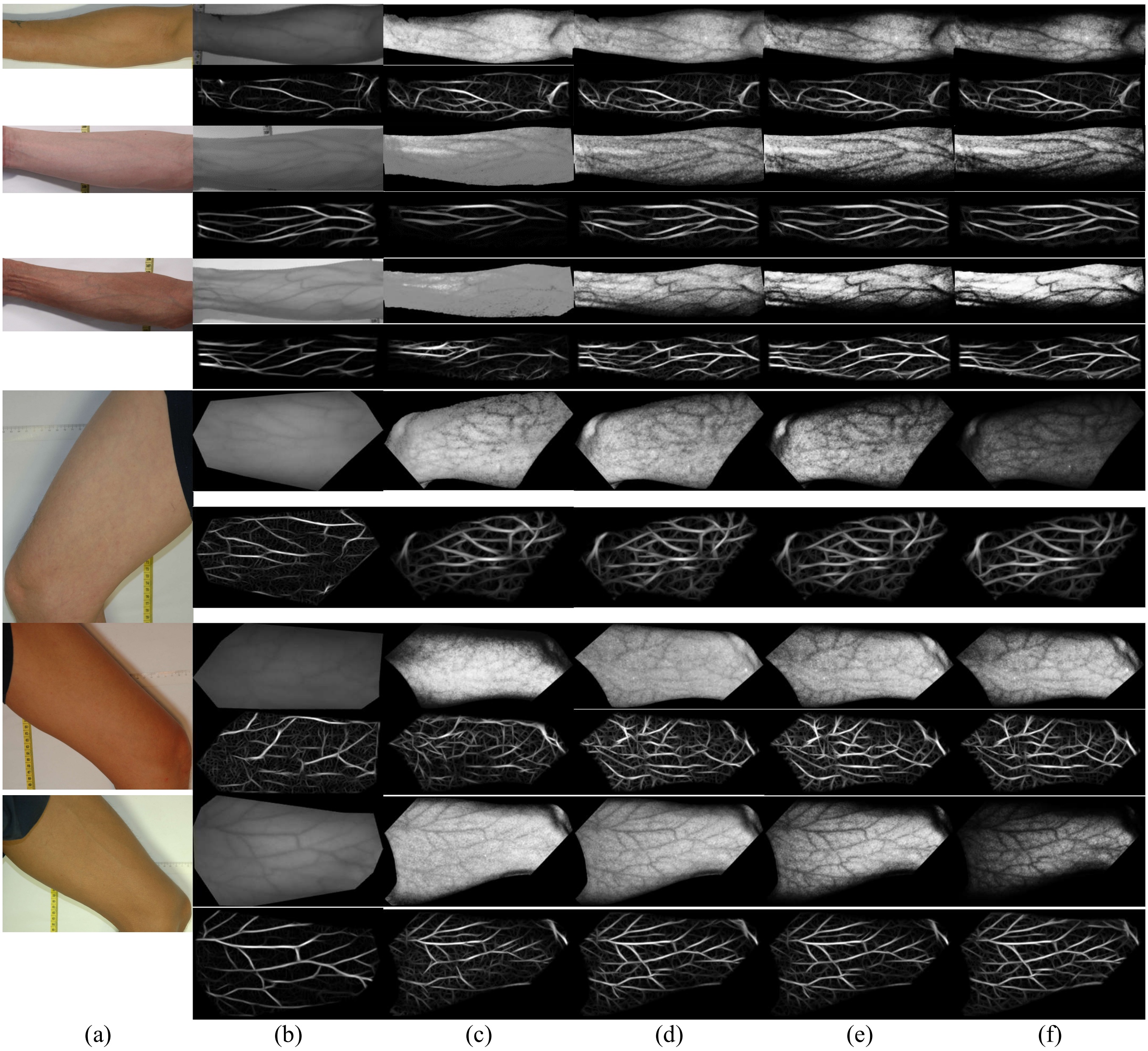


Fig. 8 Resultant images from the optical models and the corresponding enhanced response maps. Column (a) contains color images; column (b) contains NIR images; columns (c)-(f) contain results from the OBVU and the first, second, and third optical models, respectively. The enhanced response maps are shown under the processed images.

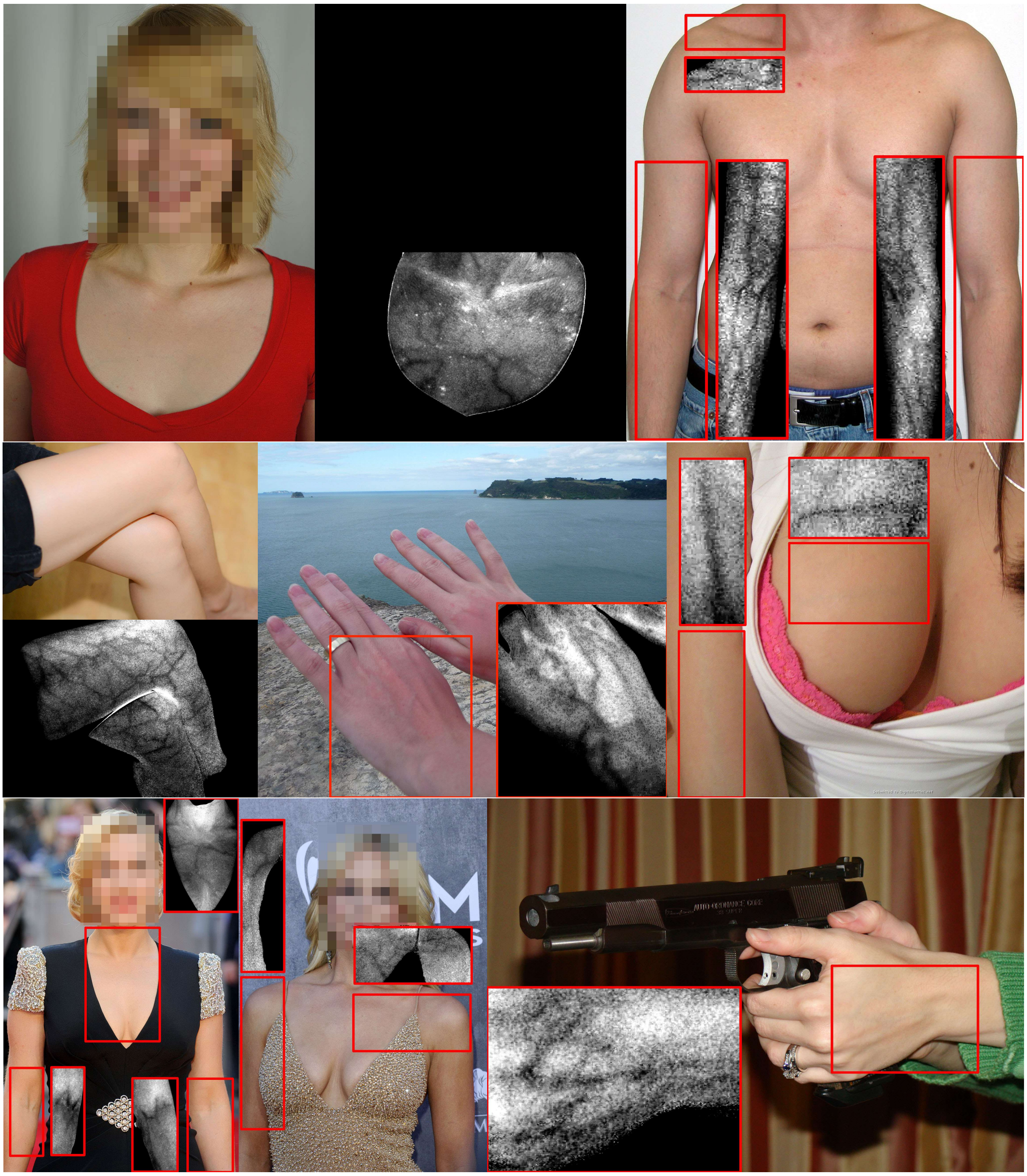


Fig. 9 Results from Internet images. Each pair of red rectangles indicates a skin region and the corresponding result from the first optical model. (To see clearly the images, please read the electronic version.)

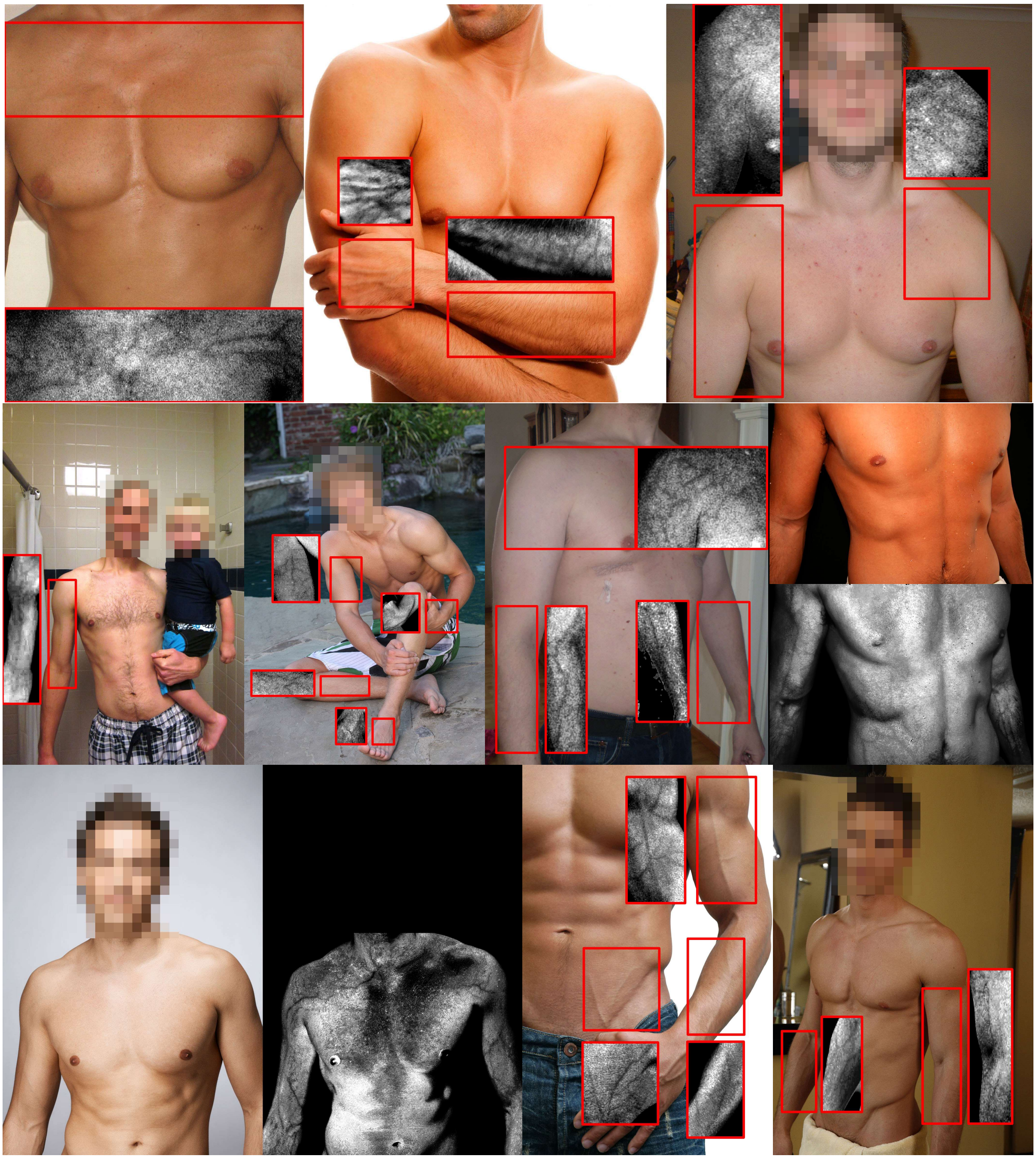
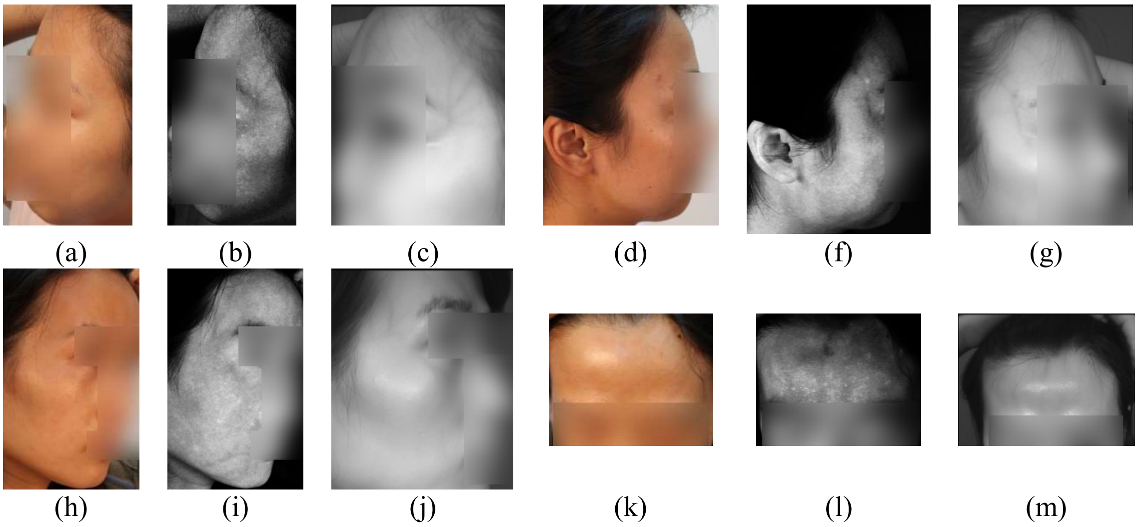


Fig. 10 Results from Internet images. Each pair of red rectangles indicates a skin region and the corresponding result from the first optical model. (To see clearly the images, please read the electronic version.)

Fig. 11 An NIR facial image collected from the Internet.



(a) (b) (c) (d) (f) (g)

(h) (i) (j) (k) (l) (m)

Fig. 12 Uncovering results obtained from some facial images. (a), (d), (h) and (k) are color facial images; (b), (f), (i) and (l) are their corresponding uncovered results obtained from optical model 1; (c), (g), (j) and (m) are the NIR images. For privacy protection, partial faces are blurred.

Table 1 Rank-1 and Rank-10 accuracies from matching blood vessel patterns from NIR images (measured in percentages).

		Left Arm	**Right Arm**	**Left Thigh**	**Right Thigh**
Rank-1 accuracy	LM-ICP	10.0	10.8	5.7	2.2
	Original-CPD	51.9	62.4	46.1	50.2
	Preliminary matching method	62.2	76.0	53.1	53.3
	Proposed matching algorithm	<u>66.8</u>	<u>80.4</u>	<u>57.9</u>	<u>56.9</u>
Rank-10 accuracy	LM-ICP	28.6	32.8	21.9	18.7
	Original-CPD	61.8	67.6	62.3	59.1
	Preliminary matching method	72.2	82.8	64.5	66.2
	Proposed matching algorithm	<u>76.8</u>	<u>86.0</u>	<u>70.6</u>	<u>70.7</u>

The underlining indicates the best matching results.

Table 2 A summary of the matching accuracy (measured in %)

		Matching the same type of images (e.g., NIR vs. NIR and color images vs. color images)				Matching different types of images (color images vs. NIR images)			
		Left Arm	Right Arm	Left Thigh	Right Thigh	Left Arm	Right Arm	Left Thigh	Right Thigh
Rank-1 accuracy	NIR	66.8	80.4	57.9	56.9				
	Model OBVU	29.7	37.6	13.6	14.5	23.1	26.4	4.3	2.3
	Model 1	52.3	68.0	46.5	46.5	35.5	49.6	13.9	14.0
	Model 2	52.7	68.0	47.8	46.9	37.2	47.6	16.5	10.7
	Model 3	52.7	70.4	39.9	43.4	35.0	42.0	16.1	10.7
	Red channel	44.8	38.4	26.3	31.6	33.8	28.4	7.8	7.0
	Feature fusion	63.6	*81.2	54.8	54.8	44.4	54.4	16.5	13.0
	Score fusion	57.7	76.0	51.8	48.2	35.9	48.4	13.9	10.2
Rank-10 accuracy	NIR	76.8	86.0	70.6	70.7				
	Model OBVU	46.0	54.8	30.7	30.3	37.6	33.2	14.8	10.7
	Model 1	66.9	81.6	59.6	61.0	50.4	60.8	28.3	27.9
	Model 2	67.8	80.4	59.6	59.6	51.7	58.0	30.0	26.5
	Model 3	68.6	83.2	51.3	55.3	52.1	59.2	28.7	24.2
	Red channel	61.9	59.6	47.4	49.6	46.6	43.6	20.4	22.3
	Feature fusion	76.2	*86.8	69.3	67.1	54.7	63.6	30.0	28.4
	Score fusion	74.5	*89.6	66.7	62.7	55.6	60.8	32.2	25.1

* denotes that the results are better than the corresponding results from NIR images. The underlining indicates the best results from matching color images with color images and matching color images with NIR images.