Hand-Shape Biometrics Combining the Visible and Short-Wave Infrared Bands

Article in IEEE Transactions on Information Forensics and Security · December 2011
DOI: 10.1109/TIFS.2011.2162948 · Source: DBLP

CITATIONS
13

READS
129

2 authors:

Miguel A. Ferrer
Universidad de Las Palmas de Gran Canaria
281 PUBLICATIONS 2,943 CITATIONS

Aythami Morales
Universidad Autónoma de Madrid
130 PUBLICATIONS 1,335 CITATIONS

Some of the authors of this publication are also working on these related projects:

- Handwriting evolution and learning, Handwriting synthesis. View project
- Signatures & robotics View project

All content following this page was uploaded by Aythami Morales on 10 August 2015.
The user has requested enhancement of the downloaded file.
Hand-Shape Biometrics combining the visible and Short Wave InfraRed Bands

Miguel A. Ferrer and Aythami Morales

Instituto Universitario para el Desarrollo Tecnológico y la Innovación en Comunicaciones,
Universidad de Las Palmas de Gran Canaria,
Campus de Tafira s/n, E35017 Las Palmas de Gran Canaria, Spain.
Phone: +34 928 451 269, Fax: + 34 928 451 243, Email: mferrer@dsc.ulpgc.es

Abstract

This paper proposes a hand shape biometric device with two sensors, respectively working in the visible and 1470nm bands. The inclusion of the 1470nm band sensor is to improve both security and performance. The security is improved by including a spoof detector and the performance by combining both bands. The spoof detector combines three skin detection indices obtained by comparing the reflectance of the hand image in the red, green and blue bands with that from the 1470nm band. The hand tissues reflect the visible radiation while absorbing the 1470nm radiation. The band combination is carried out at a score level which reduces the error rate because different images were obtained under different physical principles (reflection and absorption). The system performance has been evaluated with a database containing 10 acquisitions from each of a group of 100 users and 390 acquisitions from 62 imitated hands made of different materials. The experimental results confirm both security and performance improvement.

Index Terms: Biometrics, Hand-shape biometrics, aliveness detection, spoof detection, Biometric vulnerabilities.
I. INTRODUCTION

Personal identification based on the hand shape biometric has become a more and more popular technique for identity verification. This can be deduced from the number of recent reviews published [1 to 4]. It could be said that the three main research areas that cover the development of hand shape biometric schemes are the following: 1) research on hand acquisition devices, 2) research on hand characterization, and 3) research on feature classification. Obviously, there are other research areas such as evaluation and testing, normalization, etc.

Focusing our attention on the first research area mentioned above, most of the academic and industrial devices acquire the hand image in the visible band. These devices pay special attention to obtaining a good hand shape or hand contour in order to allow a precise measure of the hand shape and therefore to obtain a good model for the hand silhouette. For this reason, they frequently use backlighting [5] in order to increase the hand contrast. The same effect can be obtained using reflective material under the hand [6] or a uniformly colored material as background [7][8].

These techniques work well, but they can be physically spoofed, i.e., to can accept as genuine a fake or imitated hand. Even a simple, printed hand picture is able to deceive the system. To alleviate this weakness, spoof detection techniques are used, which are based on intrinsic properties of a living body (density, elasticity, capacitance, etc.) or involuntary signals from living tissue, such as pulse, blood pressure, etc. These countermeasures come into effect when the hand (genuine or imitated) touches the device [9] although recently [10] has proposed a touchless antispooing approach based on pulse oximetry. Other touchless spoof detectors have been developed in the visible band, which analyze the R-G-
B components of a skin picture. In [11] the authors develop a statistical model for human skin with a skin acceptance rate around 95% and false skin acceptance at approximately 30%. In [12], the RGB space is transformed into another color space as YIQ and compared to different classifiers with similar results. Solutions based on color eliminate the need for secondary sensors but provide lower antispooing accuracy levels although [13] presents a spoof detector for fingerprints based on multispectral imaging in the visible band where a rigorous test gives a true accept rate of 99.5% and a false accept rate of 0.9%

Spoof detectors based on color can be improved by looking for spectroscopy solutions. In [14] the reflectance spectroscopic curve between 300nm and 1100nm is used for both spoof detection and personal identification. Spoof is characterized by the area of the spectroscopic curve enclosing seven relevant bands, which correspond to absorption peaks of specific physiological components. In people search and rescue applications, [15] studied 81 bands from 800 to 1600nm proposing a normalized difference skin index (NDSI) obtained by combining the reflectance of 1100nm and 1400nm bands. The bands selected by the NDSI coincide with the results of spectroscopic studies of the human skin reflectance [16] which has a minimum around 1470nm due to the water (H₂O) content of the human body. This minimum is due to the vibrational absorption maximum of the OH molecule (OH-stretching Raman overtones) at 1470nm [17].

A. Our work

In this work we propose a hand shape biometric device with two sensors: a visible camera and a 1470nm camera which operates in the Short Wave InfraRed band (SWIR). The inclusion of the second sensor is justified because the resulting system improves both the security and the performance. The security is improved further by including a spoof detector and the performance is improved by combining the hand shape biometric schemes
for both bands. This reduces significantly the verification error rate. Fig. 1 shows a schematic view of the proposed device.

![Schematic view of the device]

The physical principles behind the improvements are simple. The human skin reflects the visible band radiation and absorbs the 1470nm wavelength because of the skin’s water content [17]. Therefore, the hand appears bright in the image of the visible band because of reflection and dark in the image of the 1470nm band because of skin absorption.

The underlying hypotheses of this work are: 1) The combination of the biometric system for each band will improve the overall performance because each is able to correct the errors from the other since they are each based on different physical principles (reflection and absorption); therefore, some noise present in the visible image, such as shading or shadowing, will not be present in the 1470nm image. 2) It is possible to design a skin detector by comparing the relationship between the pixel values of the hand image in the visible and the 1470nm bands.
Both hypotheses have been evaluated with a database containing 10 acquisitions for each band of a group of 100 users. The spoof detector system is verified with 400 hand images of the database users and 390 images of 62 imitated hands made of different materials.

The paper is organized as follows. Section 2 proposes a hand shape biometric device acquiring hand images in the visible and 1470nm bands. The hand shape based verifiers and spoof detector are described in section 3 and 4 respectively. Section 5 is devoted to the evaluation and section 6 closes the paper with the conclusions.

II. SYSTEM DESIGN

The experimental device designed consists of two cameras, a flat plate and illumination. The image in the 1470nm band is acquired by a XENICS camera XEVA 1.7-320 with an InGaAs sensor, sensitive from 900 to 1700nm, with a band pass filter lens centered on 1470nm and bandwidth of 250nm. Its resolution is $N \times M$ equal to 320x256 pixels. The image in the visible band is acquired with a color webcam, the quickcam E2500, with a resolution of 640x480 pixels. For illumination, we used an incandescent bulb with a tungsten filament which radiates from 400nm to 2500nm, covering both cameras bands.

The individual whose identity is going to be verified has to place his/her right hand freely over the flat plate with the fingers spread. We do not use pegs, templates or any other user constraining method, to capture the hand image. The illumination and the video cameras are positioned directly above the hand position to reduce the shadowing effect.

The minimum spatial resolution required for hand geometry biometry with a reasonable performance is $R = 45 \ dpi$. To ensure such a spatial resolution in the hand image, the cameras are located 30cm over the flat plate.
This distance has been obtained as follows. The XEVA 1.7-320 camera contains a sensor of size $L \times P$ equal to 0.96x0.768 cm. The focal length of its lens is $f = 1.6\, cm$ and the relative aperture can be selected from 1.4 to 16. So, its horizontal field of view is $FOV_h = 2.4 \cdot N/R = 17.07\, cm$ and its vertical field of view is $FOV_v = 2.4 \cdot M/R = 13.65\, cm$. This gives a magnification parameter equal to $m = L/FOV_h = P/FOV_v = 0.0563$. As the horizontal and vertical angles of view are $\alpha_h = 2 \cdot \text{atan}(L/(2 \cdot f \cdot (1 + m))) = 31.71^\circ$ and $\alpha_v = 2 \cdot \text{atan}(P/(2 \cdot f \cdot (1 + m))) = 25.60^\circ$ respectively, the distance between the lens and the object to cover the field of view should be:

$$p = \frac{FOV_h/2}{\tan(\alpha_h/2)} \cong \frac{FOV_v/2}{\tan(\alpha_v/2)} \cong 30\, cm.$$  

As the hand width ranges from 2 to 4 $\, cm$, the depth of field should be in the approximate range $\pm 2\, cm$. So the relative aperture is selected to be 7.1.

The webcam contains a sensor of 640x480 pixels which has a size equal to 0.418x0.364 cm. and its lens focal length is $f = 0.6\, cm$. The distance $p = 30\, cm$ implies a $FOV_h = 24\, cm$, $FOV_v = 18\, cm$ and a spatial resolution of $R = 64$ dpi. The camera is focused manually.

The plate size will take the minimum $FOV_h$ and $FOV_v$ values, which are 17 by 13 cm. An example of image captured by the proposed device is shown in Fig. 2.

![Fig 2. Images of the hand in the 1470nm (SWIR) band (left) and in the visible band (right)](image-url)
III. HAND BIOMETRIC VERIFIERS

In order to be able of verify the first hypothesis in different scenarios, we will use three different hand shape biometric approaches, one based on geometrical measures and two based on global hand shape: ICA architecture 2 for hand silhouette and Hausdorff distance for hand contour data. All hand biometric verifiers used will be applied to the images acquired in both bands in the same way, except in the image preprocessing step.

A. Hand detection and image preprocessing

The cameras of our hand biometric device are continuously acquiring images. The hand is automatically detected: when the absolute difference between two consecutive frames of the XEVA 1.7-320 camera is greater than a threshold, it is assumed that a hand is being put on the plate. When the movement finishes, it is assumed that the hand has been placed and it is still. Then both images are acquired: the image in the visible band $I_{vc}(x,y,z), 0 \leq I_{vc}(x,y,z) \leq 255, 1 \leq x \leq 640, 1 \leq y \leq 480, 1 \leq z \leq 3$ where the $z$ coordinate is the red, yellow and green colour components respectively, and the image in the SWIR band $I_{SWIR}(x,y), 0 \leq I_{SWIR}(x,y) \leq 255, 1 \leq x \leq 320, 1 \leq y \leq 240$. Note that the last 15 columns of the SWIR images have been discarded in order to keep a size proportional relationship between $I_{vc}(x,y,z)$ and $I_{SWIR}(x,y)$.

Due to the 99% pixel operability of the XEVA 1.7-320, the contents of the 298 erroneous pixels located are interpolated as the mean of the surrounding pixels and the stripped aspect of the image is alleviated by subtracting an acquired black background image.

The visible image is converted from color to gray by means of: $I_v(x,y) = 1.2 \cdot I_{vc}(x,y,1) - 0.3 \cdot (I_{vc}(x,y,2) + I_{vc}(x,y,3))$. The Otsu’s threshold [18] is used to binarize $I_v(x,y)$ and $I_{SWIR}(x,y)$.
B. Fingers tips and valley location

Let \( \{x_c(i), y_c(i)\}, \ 1 \leq i \leq L \) be the Cartesian coordinates of consecutive contour pixels of the hand binary image acquired by one of the cameras. To locate the tips and valleys between the fingers, the coordinates of the contour are converted to polar coordinates \( (r_c(i) \text{ and } \phi_c(i), 1 \leq i \leq L \) for radius and angle respectively) considering as the coordinate’s origin the center of the first column, which corresponds to the wrist side (see Fig. 2). The indices \( i_p^f, 1 \leq f \leq 4 \) of \( r_c(i) \) maxima are considered to be the finger tips, corresponding to \( f = 1,2,3 \) and 4 the little, ring, medium and index fingers. The indices of the valleys between fingers are obtained as:

\[
i_v^f = \arg\min_{i_p^f \leq i \leq i_p^{f+1}} \{r_c(i)\} \quad 1 < f < 3.
\]

The exterior bases of the index and little fingers are obtained as the nearest pixel of the exterior contour to the valley between the index and middle fingers and the valley between the ring and little fingers, respectively, i.e.:

\[
i_v^4 = i_{\text{index}} = \arg\min_{i_v^4 \leq i \leq i_v^3} d((x_c(i), y_c(i)), (x_c(i_v^3), y_c(i_v^3)))
\]

\[
i_v^0 = i_{\text{little}} = \arg\min_{1 \leq i \leq i_v^1} d((x_c(i), y_c(i)), (x_c(i_v^1), y_c(i_v^1)))
\]

where \( d(\cdot, \cdot) \) is the Euclidean distance. Fig 3 illustrates the \( i_p^f \) and \( i_v^f \) location.

Fig 3. Example of finger tips, valleys and width measures in the 1470nm band hand image.
The fingertip and valley locations vary with the hand pose on the plate or the camera view. So, the position of the tip of each finger is finely adjusted as follows:

1. Four equally spaced points are selected on each finger side.

2. The lines that minimize the square error with the selected point of each finger side are calculated, where \( y = m_r^f \cdot x + b_r^f \) is the line for the right side and \( y = m_l^f \cdot x + b_l^f \) for the left side. See Fig 4.

3. The finger axis is defined as \( y = m_a^f \cdot x + b_a^f \) where \( m_a^f = (m_r^f + m_l^f)/2 \) and \( b_a^f = (b_r^f + b_l^f)/2 \).

4. The new finger tips are the points where the finger axis and the finger contour intersect.

\[
i_p^f = \arg\min_{i_{v-1} \leq i \leq i_{v+1}} \left\{ d_{y=m_a^f \cdot x+b_a^f}([x_c(i), y_c(i)]) \right\}
\]

where \( d_{y=m_a^f \cdot x+b_a^f}([x(i), y(i)]) \) is the Euclidean distance of \{x(i), y(i)\} to the line \( y = m_a^f \cdot x + b_a^f \). In our test, this tip adjustment improves the EER of the hand shape biometric system by 44%.

Fig. 4. Fingertip adjustment. Contour line approximation of each finger side \( y = m_r^f \cdot x + b_r^f \) and \( y = m_l^f \cdot x + b_l^f \), finger axis \( y = m_a^f \cdot x + b_a^f \), and finger tip: first obtained as radius maximum, newly obtained as intersection of finger axis and finger contour.
C. Hand shape characterization

Many methods have been proposed for hand shape characterization [1]. The two main approaches are the geometrical and the global shape features. The first focuses on characteristics such as finger widths, length, etc. while the second is based either on the comparison of the hand silhouettes or the distance between vectors that model the hand shape. In order to check that the combination of $I_v(x, y)$ and $I_{SWIR}(x, y)$ improves a hand shape based biometric system, we will test both feature approaches.

C.1. Geometrical based verifier

Geometrical measures: The geometric features are obtained by measuring the widths of each finger. This is conducted as follows: The center of each finger bottom $\{x_{fb}^f, y_{fb}^f\}$ is defined as the point where the finger axis $y = m_a^f \cdot x + b_a^f$ intersects the finger base line between $\{x(i_v^f-1), y(i_v^f-1)\}$ and $\{x(i_v^f), y(i_v^f)\}$ (see Fig. 4). Let $\{x_s^f(k), y_s^f(k)\}$ be 12 equally spaced points in the line going from $\{x_{fb}^f, y_{fb}^f\}$ to $\{x_p(i_p^f), y_p(i_p^f)\}$. Supposing that the perpendicular line to the finger axis at these points is $y = m_{pa}^f(k) \cdot x + b_{pa}^f(k)$, which intersects with the right and left finger sides at the points $i_{cr}^f(k)$ and $i_{cl}^f(k)$ respectively, the finger widths at these points are obtained as:

$$d_{w}^f(k) = d(\{x_c(i_{cr}^f(k)), y_c(i_{cr}^f(k))\}, \{x_c(i_{cl}^f(k)), y_c(i_{cl}^f(k))\})$$

The feature vector is obtained by concatenating the 12 width vectors of index, middle, ring and little fingers: $d_{w}^f(k)$, $1 \leq f \leq 4, 1 \leq k \leq 12$ and the 4 finger lengths obtained as the distance between the finger bottom $\{x_{fb}^f, y_{fb}^f\}$ and finger tip $\{x_p(i_p^f), y_p(i_p^f)\}$. An example can be seen in Fig. 3. Therefore, the feature vector comprises $12 \cdot 4 + 4 = 52$ features.
Verifier: As verifier we have used a Least Squares Support Vector Machine (LS-SVM) to model each user hand in each band. The LS-SVMs are reformulations to standard SVMs which improve the robustness, sparseness, and weightings [19]. The toolbox used can be freely downloaded from [20].

The meta-parameters of the LS-SVM model are the width of the Gaussian kernels $\sigma$ and the regularization factor $\gamma$. This is taken as $\gamma = 20$ and is identical for all the LS-SVM models we use. The Gaussian width $\sigma$ parameter is optimized for each user as follows: the training sequence is randomly partitioned into two equal subsets $P_i, 1 \leq i \leq 2$. The LS-SVM is trained $L = 30$ times with the first subset $P_1, \gamma = 20$ and Gaussian width equal to $L$ logarithmically equally spaced values between $10^1$ and $10^4 \sigma_l, 1 \leq l \leq L$. Each one of the $L$ LS-SVM models is tested with the second subset $P_2$, obtaining $L$ Equal Error Rate $EER_l, 1 \leq l \leq L$ measures. The Gaussian width $\sigma$ of the signature model is selected as $\sigma = \sigma_j$ where $j = \arg \min_{1 \leq l \leq L}\{EER_l\}$. Finally, the signature model is obtained by training the LS-SVM with the whole training sequence.

Given a questioned hand image, its score is obtained by using the LS-SVM model of the identity claimed and the hand geometrical features. If the score is greater than a given threshold, the questioned hand image is accepted as genuine.

C.2. Global shape based schemes

The silhouettes contain much richer information than geometrical measures of the hand. For example, the roundness of fingertips, the shape of the thumb, the sharpness of finger valleys, etc., which are not necessarily incorporated in the geometric measurements. The global shape methods are based on direct silhouette alignment [21] or distance of between
vectors that model the hand shape [22]. In this paper both approaches are used to evaluate the improvement of combining visible and SWIR band biometric schemes.

For the second approach, we used the scheme suggested in [22] which is based on independent component analysis ICA architecture 2. As the authors provide freely the software to develop such an approach [23], we have used their programs codes in order to be sure that our procedure is the same as that described in [22].

The first approach, based on the contour mean alignment error, has been developed as follows. Let be \( \{x_f^A(i), y_f^A(i)\} \) and \( \{x_f^B(i), y_f^B(i)\} \) be 50 equal spaced samples of the finger contour \( f \) of the hands A and B respectively. The finger of hand A is adjusted to the same finger of hand B by a linear transformation that includes translation and rotation with no shape deformation. The transformation is \( (u(i) \ v(i))^T = S_f * (x_f^A(i) \ y_f^A(i) \ 1)^T \) where \( (u(i), v(i)) \) are the coordinates of \( f^{th} \) finger of hand A adjusted to \( f^{th} \) finger of hand B, and the matrixes \( S_f \) are obtained as:

\[
S_f = \begin{pmatrix}
    x_f^B(1) & \ldots & x_f^B(50) \\
y_f^B(1) & \ldots & y_f^B(50)
\end{pmatrix} \ast pinv \begin{pmatrix}
    x_f^A(1) & \ldots & x_f^A(50) \\
y_f^A(1) & \ldots & y_f^A(50) \ 1 & \ldots & 1
\end{pmatrix}
\]

The distance \( d_f(A, B) \) between \( f^{th} \) finger of hand A and B is calculated as the modified Hausdorff distance defined in [3] which measures how far \( \{x_f^B(i), y_f^B(i)\} \) and \( \{u(i), v(i)\} \) are from each other using a Euclidean metric. The distance between hand A and B is defined as \( d(A, B) = \sum_{f=1}^{4} d_f(A, B) \).

To verify whether an input hand image corresponds to the claimed user, the distances \( d(A, B_i) \) of the input hand A to each one of the claimed user reference hands \( B_i, 1 \leq i \leq T \)
are worked out. The final score is 
\[ s = \sum_{i=1}^{T} \min_{1\leq i \leq T} \{d_f(A, B_i)\}. \]
If the score is lower than a given threshold, the input hand is accepted as belonging to the claimed user.

\section*{D. Visible and SWIR Fusion}

This section describes the procedure used to combine the Visible and SWIR based hand shape biometric devices. Before combining them, the scores of each band are transformed to a distribution with mean of 0 and standard deviation of 1 with Z-score normalization, using the mean and standard deviation of the training sequence.

Let \( s_v \) and \( s_{SWIR} \) be the normalized scores obtained with the image acquired in the visible and SWIR bands respectively and \( s_{comb} \) the combined score. Several linear and non-linear methods have been evaluated to combine the biometric scores:

1) The simple score sum (SS), \( s_{comb} = s_v + s_{SWIR} \).

2) The Minimum Score (MIS), \( s_{comb} = \min\{s_v, s_{SWIR}\} \).

3) The Maximum Score (MAS), \( s_{comb} = \max\{s_v, s_{SWIR}\} \).

4) The Matcher Weighting (FLD), \( s_{comb} = w_1 s_v + w_2 s_{SWIR} = \bar{w}^T \bar{s} \), where \( \bar{s} = \{s_v, s_{SWIR}\}^T \) and the weight vector \( \bar{w} = \{w_1, w_2\}^T \) is the Fisher linear discriminant vector [24] obtained with the training subset of the database.

5) The kernel Fisher discriminant vector (KFD), \( s_{comb} = \bar{w}^T \Phi(\bar{s}) \) where \( \bar{w} \) is a weight and \( \Phi \) is a kernel function [25]. The experiments with the KFD have been carried out with the following functions [26]:

a) KFDtanh, a hyperbolic tangent function: \( \Phi(\bar{s}) = \tanh(\bar{s}) \).

b) KFDexp, an exponential function: \( \Phi(\bar{s}) = \exp(\bar{s}) \).

The weighting vector \( \bar{w} \) is obtained as the fisher linear discriminant of the \( \Phi(\bar{s}) \) scores.
6) Another non-linear combination is the score multiplication. As the scores have different signs, we transform the scores to positive numbers by means of an exponential function, which is the weighted multiplication combination (WMC): \( s_{comb} = \exp(w^T s) \) [26].

IV. REDUCING SCHEME VULNERABILITY

A. Spoof detection: physical principles

A hand based biometric device could be easily physically spoofed if an adversary imitates a genuine hand to generate a fake sample and tries to overcome the identity verification stage. As the proposed hand biometric device acquires hand images in both visible (red: 650nm, green: 520nm and blue: 470nm bands) and 1470nm bands, a spectroscopic measure using the available bands can be used to discriminate the skin from other materials and reduce such vulnerability.

Fig. 5 shows the measured skin reflectance at different wavelengths [15]. As can be seen, the spectrographic curve presents a minimum around 1470nm due to the vibrational absorption of the HO molecule present in the water content of the skin [17]. The minimum magnitude depends on the HO molecule concentration. Additionally, the skin reflectance contains a maximum around 650 nm (red color) and minima around 470 and 520nm (blue and green color) due mainly to oxyhemoglobin and deoxyhemoglobin [16]. The reflectance of Dark skin displays neither maxima nor minima because of the dominant absorption of the melanosome present in the darker skin [16].

To measure the spectroscopic relationship among these maxima and minima, we use an extended version of the normalized difference skin index (NDSI) used for search and rescue [15] called by us Modified NDSI and defined as:
\[ MNDSI(\lambda_1, \lambda_2) = \frac{J(\lambda_1) - J(\lambda_2)}{J(\lambda_1) + J(\lambda_2)} \]

where \( J(\lambda) \) are the intensities of the reflected radiation at wavelength \( \lambda \). As defined in [15], \( NDSI = MNDSI(1100,1400) \) is the best choice of \( \lambda_1 \) and \( \lambda_2 \) for skin detection. In our case, and limited by the available bands, we use \( MNDSI(650,1470) \), \( MNDSI(520,1470) \), and \( MNDSI(470,1470) \).

An expected value of these indices for skin can be estimated using the simple Kubelka and Munk model [16] which defines the reflected radiation as \( J(\lambda) = R(\lambda)I(\lambda) \), where \( I(\lambda) \) is the intensity of the incident radiation and \( R(\lambda) \) the reflectance. Fig 5 shows the measured reflectance of the human skin and the radiation curve of the light bulb used. The index can be estimated for Caucasian skin as:

\[ MNDSI(650,1470) \approx \frac{0.62 \cdot 0.9 - 0.06 \cdot 1.85}{0.62 \cdot 0.9 + 0.06 \cdot 1.85} = 0.67 \]

and \( MNDSI(520,1470) \approx MNDSI(470,1470) \approx 0.44 \). For black skin the values are approximately equal to 0.23, -0.15 and -0.15 respectively.

To circumvent the proposed spoof detector we need an imitated hand made of material with a similar concentration of HO molecules to skin and with a skin like color.

Fig 5 Spectral reflectance of black and Caucasian skin [15] and radiation curve of the lamp used in the proposed hand biometric device
B. Spoof detection procedure

The spoof detection is based on comparing the skin tissue response in the visible and in the SWIR imaging. The correspondence between pixels of both band images is not linear because they have been recorded with cameras located in different positions. A procedure to correct such a projection distortion between them is to map the coordinates of both images with a second order polynomial [27]. The polynomial coefficients are obtained by minimizing the mean square error over the nine fingertips and valleys of each image as landmarks. Therefore, the spoof detector has been as follows:

1. Let $\{x_i, y_i\}_{i=1}^{9}$ be the ordered tip and valley coordinates of the hand in the $I_{SWIR}(x,y)$ image.

2. Let $I_{rr}(x,y)$ be the $I_{vc}(x,y,1)$ image reduced to 320x240 pixels and $\{x_{ri}, y_{ri}\}_{i=1}^{9}$ the ordered tip and valley coordinates.

3. The $\{x_i, y_i\}_{i=1}^{9}$ coordinates are mapped onto $\{x_{ri}, y_{ri}\}_{i=1}^{9}$ with a two dimensional least squares polynomial as follows:

$$
\begin{pmatrix}
    x_{ri} \\
    y_{ri}
\end{pmatrix}
= Q \cdot \begin{pmatrix}
    1 & x_i & x_i^2 & y_i & y_i x_i & y_i^2 & y_i^2 x_i & y_i^2 x_i^2
\end{pmatrix}^T
$$

where $Q$ is the 2 column by 9 row matrix that minimizes the mean square error. This is obtained via the pseudoinverse ($\text{pinv}$) as:

$$
Q = \begin{pmatrix}
    x_{r1} & x_{r2} & \cdots & x_{r9} \\
    y_{r1} & y_{r2} & \cdots & y_{r9}
\end{pmatrix} \cdot \text{pinv}\left(\begin{pmatrix}
    1 & x_1 & x_1^2 & y_1 & y_1 x_1 & y_1^2 & y_1^2 x_1 & y_1^2 x_1^2 \\
    1 & x_9 & x_9^2 & y_9 & y_9 x_9 & y_9^2 & y_9^2 x_9 & y_9^2 x_9^2
\end{pmatrix}^T\right)
$$

4. As the user identity is characterized by the finger widths or finger shape, we work out the $MNDSI(650,1470)$ values for each pixel inside the finger area as:
\begin{align*}
D_r(x, y) &= \frac{I_{rr}(x, y) - I_{1470}(x_r, y_r)}{I_{rr}(x, y) + I_{1470}(x_r, y_r)} \quad \forall \{x, y\} \in \text{inside finger area}
\end{align*}

where \{x_r, y_r\} are the coordinates of \{x, y\} mapped by the Q matrix. The number of pixels inside the four finger areas range from 4000 to 1000 approximately, in our database.

5. The spoof detection score \( s_r \) is obtained calculating the 10\(^{th} \) percentile of the \( D_r(x, y) \) distribution. This percentile is chosen because any other measure such as mean or mode of \( D_r(x, y) \) can be spoofed if the forger changes just a part of the finger, e.g. by adding silicone on the tip to make it longer, or on one side to make it wider. A lower percentile makes the system too sensitive to noise, i.e. salt and pepper noise.

6. The steps 4 and 5 are repeated for reduced images \( I_{vc}(x, y, 2) \) and \( I_{vc}(x, y, 3) \) to obtain \( s_g \) and \( s_b \) respectively.

7. The spoof detection is performed with a linear discriminator based on Rosenblatt’s perceptron [24] the coefficients of which are obtained by minimizing the square error over the training sequence, that is to say, the input sample is considered skin if

\[ s_{\text{spooof}} = w_0 + w_1s_b + w_2s_g + w_3s_r > 0, \]

otherwise is considered an imitation.

V. SWIR HAND BIOMETRIC EVALUATION

A. Hands Database

Our hands database consists of 10 times 2 acquisitions (visible, and 1470nm bands) from 100 people. The 2000 images were taken from the users’ right hand. Most of the users are between 23 to 40 years old. Approximately half of the database volunteers are male.

Additionally, we have another set of 30 people by 10 repetitions called the negative training database.
As our *imitated hand database*, we have made 62 imitated hands of different materials. We have chosen classical synthetic materials such as silicone or gelatine not because of their malleability for imitate fingers [13][28] but because they contain HO molecules in their composition which absorbs the 1470nm radiation, as do genuine hands. Additionally we have made hands with natural materials such as leaves or fruits in which the water content is similar to the human body. We acquired 5 images of each imitated hand per session. For synthetic materials, we took just one session and for natural material we took several sessions during which the samples experienced a natural drying procedure which results in a loss of water content. This can be useful for measuring the time validity of the natural samples. The rest of materials evaluated, such as the clay, paper, etc. with different colours and textures, were chosen just to check for any possible anomalies in our results.

The evaluated materials and their colours are listed in Table I. The table also shows the $s_{spoof}$ scores obtained with each material (35 for plastic, 35 for paper, 5 for picture, and so on) and the Mahalanobis distance between these and the 980 $s_{spoof}$ scores obtained using the Caucasian hands. Images of several imitated hands can be seen in Fig 6.

![Fig 6. Some examples of images acquired at 1470nm $l_{1470}(x_r, y_r)$ and in red $l_r(x, y)$ bands. Upper row: red band; Lower row: 1470nm band. From left to right: genuine hand, clay, cork, plaster, black plastic and silicone imitated hands.](image-url)
### Table I. Hand imitiation materials used for spoof detector test

<table>
<thead>
<tr>
<th>Imitated hands</th>
<th>Observations</th>
<th>Images</th>
<th>$\Delta m^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plastic</td>
<td>Seven colors(^1)</td>
<td>35</td>
<td>25.00</td>
</tr>
<tr>
<td>Paper</td>
<td>Seven colors(^1)</td>
<td>35</td>
<td>20.42</td>
</tr>
<tr>
<td>Paper</td>
<td>Hand picture scanned at 600 dpi</td>
<td>5</td>
<td>18.34</td>
</tr>
<tr>
<td>Cardboard</td>
<td>Seven colors(^1)</td>
<td>35</td>
<td>20.56</td>
</tr>
<tr>
<td>Play-doh</td>
<td>Seven colors(^1)</td>
<td>35</td>
<td>22.34</td>
</tr>
<tr>
<td>Clay</td>
<td>Brown</td>
<td>5</td>
<td>22.32</td>
</tr>
<tr>
<td>Wood</td>
<td>Beech, cherry, riga, pine and wood cement</td>
<td>25</td>
<td>19.32</td>
</tr>
<tr>
<td>Cork</td>
<td>White color</td>
<td>5</td>
<td>31.73</td>
</tr>
<tr>
<td>Plaster</td>
<td>White color</td>
<td>5</td>
<td>10.63</td>
</tr>
<tr>
<td>Metal</td>
<td>Gray stainless steel</td>
<td>5</td>
<td>15.31</td>
</tr>
<tr>
<td>Leather</td>
<td>Brown color</td>
<td>5</td>
<td>16.67</td>
</tr>
<tr>
<td>Clothes</td>
<td>Red, brown and beige gloves</td>
<td>15</td>
<td>14.09</td>
</tr>
<tr>
<td>Wax</td>
<td>Natural and mixed with red tempera powder</td>
<td>10</td>
<td>10.34</td>
</tr>
<tr>
<td>Latex</td>
<td>Glove</td>
<td>5</td>
<td>6.86</td>
</tr>
<tr>
<td>Rubber</td>
<td>Brown color</td>
<td>5</td>
<td>10.84</td>
</tr>
<tr>
<td>Glue</td>
<td>Wood glue</td>
<td>5</td>
<td>3.22</td>
</tr>
<tr>
<td>Chewing gum</td>
<td>Red and pink color</td>
<td>10</td>
<td>3.44</td>
</tr>
<tr>
<td>Gelatin</td>
<td>Made with red and pink powder</td>
<td>10</td>
<td>2.05</td>
</tr>
<tr>
<td>Silicone</td>
<td></td>
<td>5</td>
<td>2.82</td>
</tr>
<tr>
<td>Tree leaves</td>
<td>Eucalyptus, orange &amp; laurel trees</td>
<td>15</td>
<td>1.85</td>
</tr>
<tr>
<td>Tree leaves</td>
<td>As above, dried during a week</td>
<td>15</td>
<td>2.91</td>
</tr>
<tr>
<td>Tree leaves</td>
<td>As above, dried during a fortnight</td>
<td>15</td>
<td>3.65</td>
</tr>
<tr>
<td>Fruit</td>
<td>Orange, banana, red apple</td>
<td>15</td>
<td>1.23</td>
</tr>
<tr>
<td>Fruit</td>
<td>As above, dried during one day</td>
<td>15</td>
<td>1.92</td>
</tr>
<tr>
<td>Fruit</td>
<td>As above, dried during two days</td>
<td>15</td>
<td>2.04</td>
</tr>
<tr>
<td>Natural latex</td>
<td>Harvested from common fig and Euphorbia Canariensis</td>
<td>10</td>
<td>1.34</td>
</tr>
<tr>
<td>Natural latex</td>
<td>As above, dried during a week</td>
<td>10</td>
<td>1.97</td>
</tr>
<tr>
<td>Natural Resin</td>
<td>Harvested from Pine</td>
<td>5</td>
<td>1.72</td>
</tr>
<tr>
<td>Natural Resin</td>
<td>As above, dried during a week</td>
<td>5</td>
<td>2.45</td>
</tr>
<tr>
<td>Ham</td>
<td>Fresh</td>
<td>5</td>
<td>1.05</td>
</tr>
<tr>
<td>Ham</td>
<td>As above, dried during a day</td>
<td>5</td>
<td>1.78</td>
</tr>
<tr>
<td>Genuine hands</td>
<td>Caucasian (50 male and 48 female)</td>
<td>980</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Black (2 female)</td>
<td>20</td>
<td>0.45</td>
</tr>
<tr>
<td>Negative Training</td>
<td>Caucasian (15 male and 15 female)</td>
<td>300</td>
<td>0.12</td>
</tr>
</tbody>
</table>

\(^1\)White, yellow, orange, pink, red, brown and violet color

\(^2\)Mahalanobis distance between each material and hand tissue.

### B. Evaluation methodology

The evaluation of the security of the hand shape biometric device proposed has been conducted on the basis of the guidance provided in [29] and [30]. It has been performed in two steps, each evaluating one of the two hypotheses we set out above in section 1A.
To evaluate the first hypothesis, we make measurements of the statistical error rates in order to establish as reliable a figure as possible for the verification and identification performance of the device. The statistical measures provided for verification are the False Accept Rate (FAR) and the False Reject Rate (FRR) along with the Detection Error Trade (DET) curve and Equal Error Rate (EER) value. For identification we provide the Cumulative Match Curve (CMC).

In the case of global shape hand based verifiers, four randomly chosen hand images are used as user reference images. In the case of the geometrical based hand verifier, the LS-SVM of each user is trained with the same four reference hand images as positive samples. Two training case scenarios are considered:

- Case 1, the negative samples were the reference samples of the other users: 99x4=396.
- Case 2, the negative samples were the 300 samples of the negative training database.

This second case attempts to ascertain more realistic results because it did not use information about the imposter users in the training models and could be used as a fair comparison with the global shape hand verifier.

To evaluate the devices, we use the remaining 100x6 samples of the database not used for training. The EER values given are the average and variance of the results obtained by repeating the experiment 10 times. As a result, we have $N_g = 100 \times 6 \times 10 = 6000$ scores for genuine users and $N_f = 100 \times 99 \times 6 \times 10 = 594000$ scores for impostors.

Turning to the evaluation methodology for the second hypothesis, we assess whether the biometric system will accept an artifact as a valid biometric sample. The evaluation is conducted by supposing that the spoof detector has not been trained with the evaluated material. Therefore we divide the materials in two disjointed subsets $M_1$ and $M_2$ containing
half of the synthetic and half of the natural materials. The performance measures are obtained by concatenating the results of training with subset $M_1$ and testing with $M_2$ plus training with subset $M_2$ and testing with $M_1$. As for the skin class we have trained with two acquisitions of the 100 hand database users and tested with another two acquisitions. The black hand samples are used one for training and the other one for testing.

The aggregated results are given as FAR and FRR curves. The FAR and FRR curves do not adequately represent the performance if a sample material is able to deceive repeatedly the spoof detector [13]. In order to gain a better idea of the discriminative ability of the spoof detector, the Mahalanobis distance [24] between the score distributions of hands and different materials is given in Table I. A high Mahalanobis distance means a strong difference between genuine hands and materials used for imitations.

C. Biometric verifiers experimental results

To establish a baseline, the biometric verifiers developed in both bands have been evaluated separately. Table II shows the EER obtained from the experiments. As to global shape features, it can be seen that the results with ICA2 are very similar to those obtained in [22] and a little better than those obtained with alignment. Contrary to what we expected, the geometrical features work better than global shape features. Although is not the aim of this paper to compare them, this improvement is attributed to the classifier used. In our experiments, using a KNN classifier [24], the geometrical features performs at an EER around 4.5%, using SVM light [31] the EER decrease to 1.5% and using the above LS-SVM the error is below 0.6%. We have not been able to obtain these improvements by applying the SVM classifier to alignment or to ICA2 features.
Returning to the visible vs. 1470nm performance comparison, it can be seen that the best results are obtained with the SWIR band in all the cases, despite its lower resolution than in the visible images. This improvement could be attributed to the fact that the image acquisition based on hand radiation absorption gives a more precise hand contour, without the possible contour bias due to hand shade that appears in the visible band images.

**TABLE II. EERs (%) of the biometric verifiers in visible and SWIR bands**

<table>
<thead>
<tr>
<th>Camera band</th>
<th>Global shape features</th>
<th>Geometrical features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alignment</td>
<td>ICA2</td>
</tr>
<tr>
<td>Visible</td>
<td>3.33±0.40</td>
<td>3.36±0.95</td>
</tr>
<tr>
<td>SWIR</td>
<td>2.85±0.83</td>
<td>1.93±0.90</td>
</tr>
</tbody>
</table>

**D. Combining scores from visible and SWIR bands**

This section checks whether the hypothesis that the score combination of the visible and SWIR based hand shape biometric devices we have evaluated improves the system performance. Fig 7 represents the visible-SWIR joint distribution of the genuine and impostor scores for the geometrical features. It is easy to see that the genuine and impostor distribution are easier to separate taking into account the scores of both bands together than by raking each band on its own.

![Image](image-url)

Fig 7. visible-SWIR joint distribution of the genuine and impostor scores for the geometrical features and training case 2.
Table III shows the results of combining both bands at the score level for the different fusion methods evaluated. As can be seen, the combination improves significantly the scheme performance, the FLD and non-linear methods being the best strategies, providing similar results. We follow our discussion above about using the $s_{comb}$ obtained with the FLD method and both experimental methodologies (case 1 and case 2) because this is the simplest one to calculate. The improvement using this combination is confirmed in Fig. 8 which shows the DET curves. They compare the result of using $s_v$, $s_{SWIR}$, and $s_{comb}$ for the geometrical features (case 2 for a fair comparison with global shape methods), silhouette alignment and ICA2. For the case of identification, the CMC curves are provided in Fig 9.

This improvement experimentally confirms that the combination of hand shape biometric devices operating in visible and SWIR band improves the performance. As the improvement remains when different features are used, we consider that it is due to the images acquired in each band being complementary, that is to say, the image in the SWIR band is robust to the noises and distortion that affects the visible image and vice versa.

### TABLE III

| EERs (%) RESULTS OBTAINED COMBINING VISIBLE AND SWIR BAND AT SCORE LEVEL |
|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Biometric Fusion | Sum  | MIS  | MAS  | FLD  | KFDtanh | KFDexp | WMC |
| Global shape features | Alignment | 2.39 | 2.74 | 2.67 | 2.36 | 2.37 | 2.34 | 2.36 |
| ICA2 | 1.79 | 2.74 | 1.77 | 1.68 | 1.68 | 1.70 | 1.69 |
| Geometrical features | Case 1 | 0.014 | 0.15 | 0.015 | 0.004 | 0.004 | 0.005 | 0.004 |
| Case 2 | 0.32 | 0.51 | 0.32 | 0.25 | 0.25 | 0.25 | 0.25 |

Fig 8. From left to right DET curves which compare the result of using $s_v$ (red dotted line), $s_{1470}$ (blue dashed-dotted line) and $s_{comb}$ (green continuous line) for the silhouette alignment, ICA2 and geometrical features.
Fig. 9. From left to right CMC curves which compare the result of using $s_v$ (dotted line), $s_{1470}$ (dashed line) and $s_{comb}$ (continuous line) for the silhouette alignment, ICA2 and geometrical features.

E. Spoof detection assessment

Fig. 10 shows the joint distribution of $s_g$ and $s_r$ for genuine and imitated hands. To make the illustration clearer, we have not drawn the 3D graphic. It can be seen that most of the scores lies around the above calculated Caucasian skin center. The skin scores distribution moves toward the black skin center with increasing skin melanosome level. Referring to imitated hands, the nearest scores to the linear discriminator border belong to natural material.

To illustrate further the spoof detector operating mode, Fig 11 shows the FAR and FRR curves of the scores $s_{spooof}$ obtained from genuine and imitated hands. As the linear discriminator that combines the red, green and blue scores has been designed for a threshold equal to zero, it can be seen that there are no errors when using this database. The nearest materials to the threshold are the natural ones while the nearest skin to the threshold is black skin.
The above aggregated results do not show the ability of the spoof detector to discriminate between the individual materials we used for imitation hands. To illustrate this ability, Table I show the Mahalanobis distance between the $s_{\text{spoof}}$ for the Caucasians hands and each material. The distance from the hands to the nearest material, which is the fresh ham, is greater than the distance between the hands themselves, including the black hands. The distance of the hands to the natural material increases when the natural samples are dried, confirming the importance of water contents on the spoofing material response. The nearest synthetic material is the gelatin.

As an impostor can change his/her finger shape by adding a piece of silicone or another material to their fingers, we should assess the spoof detector robustness against finger modifications. The relevant experiment, the result of which can be seen in Fig. 12, consists of adding more and more silicone to a finger until the spoof detector rejects the sample. In this experiment, the material added is silicone with pink powder, which is one of the nearest synthetic materials to the hand tissue. The score $s_{\text{spoof}}$ has been obtained by adding from 1% to 30% silicone to the finger area. In the ranges 1% to 10% and 10% to 30% the increasing steps are approximately 1% and 5% respectively. As $s_{\text{spoof}}$ is the combination of each band score obtained at $10^{\text{th}}$ percentile of MNDSI distributions, we have also varied this percentile from $2^{\text{nd}}$ to $26^{\text{th}}$ percentile. It can be seen that using the $2^{\text{nd}}$ or $4^{\text{th}}$ percentile, the image is always rejected as skin because of the noise or the non-perfect matching between the visible and SWIR images. At the $10^{\text{th}}$ percentile, the modified finger can deceive the spoof detector until approximately 7% of the finger area has been added which is equivalent to adding half a centimeter to the finger length. This means that an impersonator with a smaller hand may try to add to each of his fingers up to 7% of its
projected area in silicone to imitate someone else’s hand. To calculate this possibility, Fig. 13 shows the distribution of the percentage area differences between different human subjects. In our database, the probability of two subjects having a four finger area difference below 7% is 4.26% although a false positive in the aliveness detection algorithm does not imply a false acceptance during verification procedure.

\[ s_{\text{spoof}} \]

**F. Computational load**

In this section, Table IV show the estimated computational time for the feature extraction, matching and spoof detector stages using Matlab on a Pentium Dual-Core 1.66GHz with 4Gb RAM. The geometrical feature extraction includes the fingertip and valley detection plus the measures of the 4 finger widths while the matching is the time taken by the LSSVM. In the case of alignment, the feature extraction consists of obtaining the fingertips and valleys while the matching uses the Hausdorff distance measure between fingers. For ICA2, the feature extraction consists of obtaining the shape feature vector and the matching is the calculation of the cosine similarity measure. As can be seen in Table IV, except for ICA2 feature extraction, the verification and spoof detector takes about 0.5 seconds. The spoof detector takes about 60% of the computing time.
<table>
<thead>
<tr>
<th></th>
<th>geometrical</th>
<th>alignment</th>
<th>ICA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature extraction</td>
<td>0.233917</td>
<td>0.194861</td>
<td>5.21</td>
</tr>
<tr>
<td>Matching</td>
<td>0.000768</td>
<td>0.026920</td>
<td>0.000045</td>
</tr>
<tr>
<td>Spoof detector</td>
<td></td>
<td>0.329357</td>
<td></td>
</tr>
<tr>
<td>Total time</td>
<td>0.564042</td>
<td>0.551133</td>
<td>5.530045</td>
</tr>
</tbody>
</table>

Therefore, except in the case of ICA2, the hand based biometric identifier using one camera will take approximately half a second to accept or reject the user and a little bit more than a second using the proposed bimodal scheme.

VI. CONCLUSIONS

We have examined different strategies for combining visible and 1470nm band cameras for hand shape based biometric devices in order to increase the performance and the security of the biometric. The experiments have been designed to assess the two initial hypotheses: First, combining the visible and SWIR band scores will improve the overall performance of the hand shape biometric device because the images are acquired using different physical properties. Second, it is possible to design a reliable skin detector by comparing the pixel values of the hand image in the visible and 1470nm bands.

The first hypothesis has been validated using different hand shape features based on geometrical and global shape models. In all the tested cases, the score combination improves the overall performance. Among the different score combination strategies, the linear score normalization with weighted linear fusion and with \textit{a priori} estimation of the weight factor has shown to be the best.

To validate the second hypothesis, we have proposed a combined score obtained from the 10th percentile of the Modified Normalized Difference Skin Index distributions for the
red, green and blue bands. This has been tested with 100 people and 390 hand images of 62 imitated hands made of different synthetic and natural materials.

The result of the experiments show that, when working with hand shape biometric devices, to add a second camera working at 1470nm wavelength is a promising option for reliable spoof detection, especially when the same sensor is used to obtain the additional advantage of improving the overall scheme performance.

ACKNOWLEDGMENT
This work has been funded by Spanish government MCINN TEC2009-14123-C04 research project.

REFERENCES

