Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Contact-free and pose-invariant hand-biometric-based personal identification system using RGB and depth data*

Can WANG, Hong LIU‡, Xing LIU

(*Engineering Laboratory on Intelligent Perception for Internet of Things (ELIP) and MOE Key Laboratory of Machine Perception, Shenzhen Graduate School, Peking University, Shenzhen 518055, China*)

E-mail: herowc@pku.edu.cn; hongliu@pku.edu.cn; liuxing@sz.pku.edu.cn

Received July 13, 2013; Revision accepted Mar. 24, 2014; Crosschecked May 6, 2014

Abstract: Hand-biometric-based personal identification is considered to be an effective method for automatic recognition. However, existing systems require strict constraints during data acquisition, such as costly devices, specified postures, simple background, and stable illumination. In this paper, a contactless personal identification system is proposed based on matching hand geometry features and color features. An inexpensive Kinect sensor is used to acquire depth and color images of the hand. During image acquisition, no pegs or surfaces are used to constrain hand position or posture. We segment the hand from the background through depth images through a process which is insensitive to illumination and background. Then finger orientations and landmark points, like finger tips or finger valleys, are obtained by geodesic hand contour analysis. Geometric features are extracted from depth images and palmprint features from intensity images. In previous systems, hand features like finger length and width are normalized, which results in the loss of the original geometric features. In our system, we transform 2D image points into real world coordinates, so that the geometric features remain invariant to distance and perspective effects. Extensive experiments demonstrate that the proposed hand-biometric-based personal identification system is effective and robust in various practical situations.

Key words: Hand biometric, Contact free, Pose invariant, Identification system, Multiple features
doi:10.1631/jzus.C1300190 **Document code:** A **CLC number:** TP391.4

1 Introduction

With rapid development in information technology, biometrics have emerged to provide a greater degree of security for personal identification systems. Among the various biometric systems, hand-biometric-based systems have received increasing interest because of their user acceptability for biometric traits (Kanhangad *et al.*, 2010; Ramalho *et al.*, 2011; Michael *et al.*, 2012).

Shape features extracted from the hand carried only limited information for discrimination. Over the years, various approaches have been proposed to address this problem. Some systems have been developed to eliminate the use of pegs and avoid the need for the user to make contact with a flat surface. These hand-biometric-based identification systems can be grouped into three categories based on the method of image acquisition, as described by Kanhangad *et al.* (2011a):

1. Constrained and contact based: This kind of system employs pegs and a flat surface to constrain the position of the user's hand. Many commercial systems and early research systems fall in this cat-

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. 61340046, 60875050, and 60675025), the National High-Tech R&D Program (863) of China (No. 2006AA04Z247), the Scientific and Technical Innovation Commission of Shenzhen Municipality (Nos. JCYJ20120614152234873, CXC201104210010A, JCYJ20130331144631730, and JCYJ20130331144716089), and the Specialized Research Fund for the Doctoral Program of Higher Education, China (No. 20130001110011)

egory (Sanchez-Reillo *et al.*, 2000; Sanchez-Reillo, 2000). Sanchez-Reillo *et al.* (2000) and Sanchez-Reillo (2000) used six pegs for hand geometry implementation. They measured 25 geometric sizes of a hand, obtaining an equal error rate (EER) of 4.9%. Although pegs provide consistent measuring positions, they also cause user inconvenience and raise public-health concerns.

2. Unconstrained and contact based: This kind of system also requires the user to put his/her hand on a flat surface (Woodard and Flynn, 2005; Methani and Namboodiri, 2009) or a digital scanner (Ribaric and Fratric, 2005; Xiong *et al.*, 2005). These systems do not constrain position or posture, which is more user friendly. Woodard and Flynn (2005) extracted 3D finger surface features from 3D range images of hands placed on a flat surface. Ribaric and Fratric (2005) scanned images of the hand using a low-cost scanner. However, contact based image acquisition still brings hygiene and public-health concerns.

3. Unconstrained and contact-free: This kind of system requires no peg or flat surface during hand image acquisition. It is believed to be more user-friendly and has therefore attracted much attention recently (Zhang *et al.*, 2003; Kanhangad *et al.*, 2010; 2011a; Dai *et al.*, 2012; Morales *et al.*, 2012). A typical peg-free hand geometry technique uses optical devices to capture the image of a hand. An expensive 3D digitizer is used to capture 3D hand images (Kanhangad *et al.*, 2010; 2011a; 2011b). Although the systems are unconstrained and contact-free, the expensive device makes these systems inapplicable in real environments. Moreover, clear and simple backgrounds are required for hand segmentation. Skin-like and bad illumination conditions pose difficulties for these systems. In this study, we focus on eliminating these illumination and background constraints by means of an inexpensive Kinect sensor (Microsoft Corporation, Kinect for Xbox 360).

Another problem to be addressed is the segmentation of hands. Many existing systems cannot segment a hand from the background accurately. Kanhangad *et al.* (2011b) used an expensive 3D scanning digitizer to acquire color and depth images of hands, but this method also requires a clean and simple background. Morales *et al.* (2008) modified a webcam to receive infrared emissions. Maximizing contrast and brightness with a low value achieves a very high contrast between the hand and the background, so the hand can be separated easily. In our system, an inexpensive Kinect sensor is used to acquire color and depths image of a hand. Hand segmentation is implemented based on the depth image of a hand, which is insensitive to changes in illumination and background distractions. So, unlike many other systems, our system can still work in poor illumination environments with cluttered backgrounds.

The main contribution of our work lies in two aspects. First, a novel geodesic contour analysis method is proposed to localize landmark points of the hand and extract its geometric features. Besides explicit features like finger length widely used in previous systems, implicit geometric features can also be obtained by analyzing a curvature matrix which contains details of all hand contour features. Second, 3D hand registration is conducted to ensure that features extracted from hand regions are kept rotation, pose, and depth invariant. This greatly reduces the intra-class difference when describing the same hand with different poses and distances. Extensive experiments were conducted which validate the effectiveness and robustness of the proposed framework for describing hand biometric features and its practicality for personal identification.

2 System description

2.1 Data acquisition

An illustration of the proposed hand-biometric-based identification system framework given in Fig. 1. The first step in our system is to acquire data from the hand. A Kinect sensor is used to capture color and depth images of the hand. During image acquisition, the user is requested to position his/her palm in front of the Kinect sensor. Fingers are slightly and casually stretched apart. There are no guide pegs or flat surfaces to constrain the user's hand. The user can put his/her hand naturally above the sensor in various postures. The optimal interaction region in our system is set as 60–100 cm from the Kinect sensor. The background and illumination conditions in our system are not strictly constrained.

2.2 Hand segmentation

The second step in our system is to segment the hand for subsequent feature extraction.

In our system, given depth data captured from

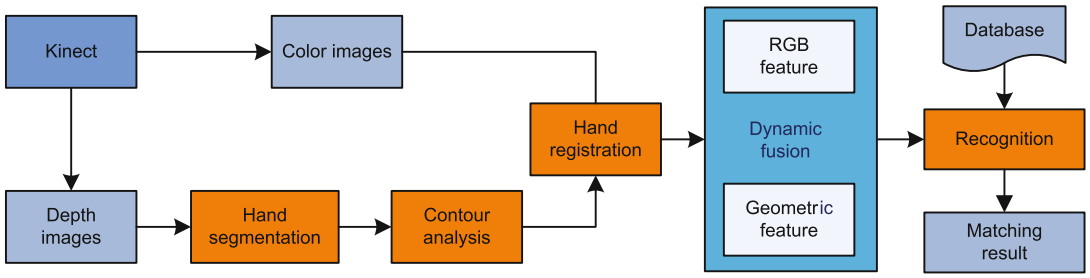

Fig. 1 A summary illustration of the proposed biometric-feature-based human hand identification system

the Kinect sensor, the hand region will be segmented and its contour will be extracted for the following analysis.

In this framework, a weighted-depth-histogram-based first-peak hand segmentation method is adopted to segment the hand in front of the depth sensor. Suppose the depth histogram of the depth map is $\boldsymbol{H}_{\mathrm{d}} = (h_{\mathrm{d}}^1, h_{\mathrm{d}}^2, \cdots, h_{\mathrm{d}}^{N_{\mathrm{d}}})$. The weighted depth histogram $\boldsymbol{H}_{\mathrm{w}}$ can be given as

$$\boldsymbol{H}_{\mathrm{w}} = (h_{\mathrm{w}}^1, h_{\mathrm{w}}^2, \cdots, h_{\mathrm{w}}^{N_{\mathrm{d}}}), \quad h_{\mathrm{w}}^k = \frac{\omega_{\mathrm{d}}}{k} h_{\mathrm{d}}^k, \tag{1}$$

where ω_{d} is a weight controlling parameter and h_{w}^k $(k = 1, 2, \cdots, N_{\mathrm{d}})$ indicates the histogram value in the kth bin. Based on the weighted histogram, the first peak is extracted, which is regarded as the hand region. A brief illustration is given in Fig. 2. Fig. 2b is the original depth histogram and Fig. 2c the weighted histogram. The histogram value around the hand region becomes larger after weighting, so the first peak is obvious and the hand region can be separated more accurately and robustly from noise.

After the hand is segmented on the 2D depth image, its contour is extracted based on Canny edge extraction, denoted as $\hat{\mathcal{C}}_{\mathrm{h}} = \{\hat{\boldsymbol{p}}_1, \hat{\boldsymbol{p}}_2, \cdots, \hat{\boldsymbol{p}}_N\}$, where $\hat{p}_k$ $(k = 1, 2, \cdots, N)$ is the kth contour point of the hand with 2D coordinates (u_k, v_k) on the depth image. In this framework, a depth-invariant transformation (DIT) is adopted to transform 2D points on the image plane into 3D camera coordinates, formulated as

$$x_{\mathrm{c}} = d \cdot \frac{u - u_0}{f_u}, \ y_{\mathrm{c}} = d \cdot \frac{v - v_0}{f_v}, \ z_{\mathrm{c}} = d, \tag{2}$$

where u_0, v_0, f_u, f_v are intrinsic parameters of the sensor and d is the depth value of pixel (u, v) on the depth map. DIT transforms a point (u, v) in the depth map with depth value d to a point $(x_{\mathrm{c}}, y_{\mathrm{c}}, z_{\mathrm{c}})$ in the 3D camera coordinates. Fig. 2d shows the segmented hand from Fig. 2a transformed to camera coordinates.

DIT is simple but important for subsequent geometric feature extraction in our framework. It transforms the hand contour to the same scale as the real world length metric, which keeps all geometric features of the hand real and makes them invariant to depth changes or perspective effects.

2.3 Geodesic contour analysis

Based on the hand segmentation described above, geodesic contour analysis is conducted to extract geometric features of the hand, which is based on our previous work in Wang *et al.* (2013). Let $\mathcal{R}_{\mathrm{h}}$ denote the segmented hand region and $\mathcal{C}_{\mathrm{h}}$ its 3D contour. Suppose there are N points on $\mathcal{C}_{\mathrm{h}}$. Let p_i denote a contour point. The contour $\mathcal{C}_{\mathrm{h}}$ can be represented by $\mathcal{C}_{\mathrm{h}} = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_N\}$.

Given any point $\boldsymbol{p}_i$ on the hand region contour $\mathcal{C}_{\mathrm{h}}$, a geodesic curve $\Gamma_i^{s_j}$ around $\boldsymbol{p}_i$ at step s_j is defined as a point set:

$$\Gamma_i^{s_j} = \chi(\boldsymbol{p}_i, s_j) = \{\boldsymbol{p}_k | i - s_j \le k \le i + s_j\}, \tag{3}$$

where step s_j determines the length span of local geodesic curve $\Gamma_i^{s_j}$.

For each contour point $\boldsymbol{p}_i$, its geodesic curve set S_i is defined, which consists of local geodesic curves in M steps, formulated as

$$S_i = \{\Gamma_i^{s_j} = \chi(\boldsymbol{p}_i, s_j) | 1 \le j \le M\}. \tag{4}$$

Given any local geodesic curve $\Gamma_i^{s_j}$ on the hand contour, we calculate its curvature for biometric feature representation. Let $\kappa_i^{s_j}$ denote the curvature of curve $\Gamma_i^{s_j}$, formulated as

$$\kappa_i^{s_j} = \delta \cdot \frac{l_{p_i}^{s_j}}{|\boldsymbol{p}_{i-s_j}, \boldsymbol{p}_{i+s_j}|}, \tag{5}$$

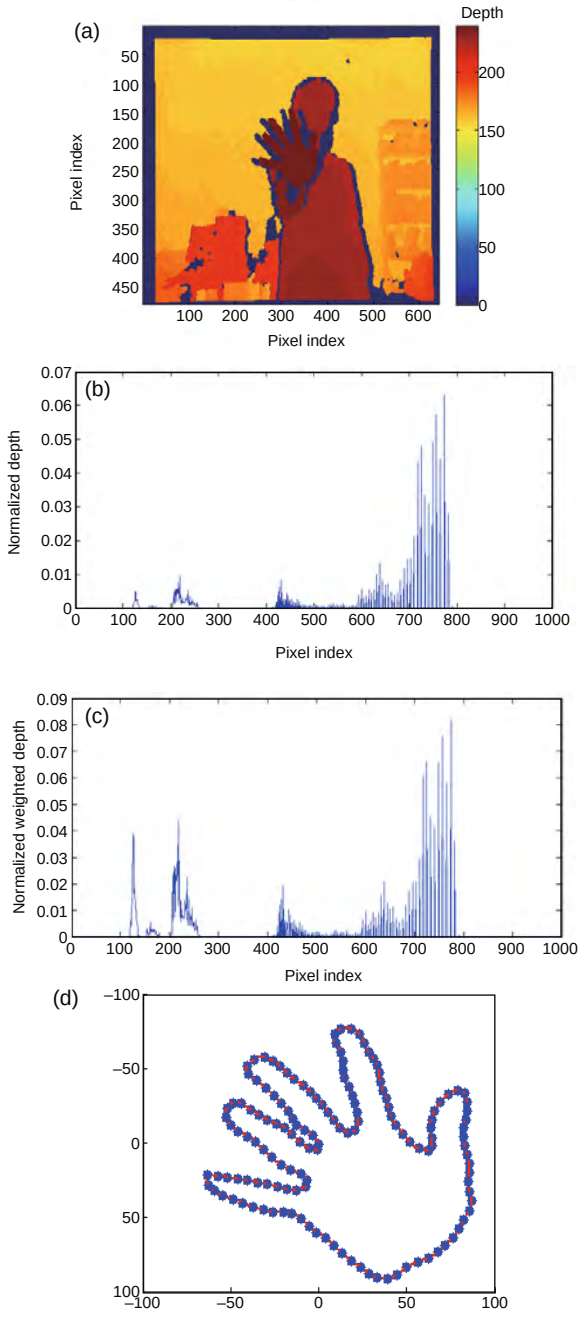

Fig. 2 An illustration of the weighted-depth-histogram based first-peak hand segmentation method: (a) original depth map; (b) its depth histogram; (c) weighted depth histogram; (d) 3D hand contour transformed to the camera coordinates

where $|\boldsymbol{p}_i, \boldsymbol{p}_j|$ is defined as the Euclidean distance between any two points $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$, and $l_i^{s_j}$ denotes the geodesic length of curve $\Gamma_i^{s_j}$:

$$l_{p_i}^{s_j} = \sum_{k=i-s_j}^{i+s_j-1} |\boldsymbol{p}_k, \boldsymbol{p}_{k+1}|. \tag{6}$$

To avoid repeated calculation of the Euclidean distance between two points in Eqs. (5) and (6), a Euclidean distance look-up table is adopted in our implementation, which ensures that all Euclidean distances between each two points used in curvature calculation are calculated only once.

In Eq. (5) δ is defined to indicate whether local curve $\Gamma_i^{s_j}$ is concave or convex:

$$\delta = \begin{cases} 1, & [\boldsymbol{p}_{i-s_j}, \boldsymbol{p}_{i+s_j}] \in \mathcal{R}_{\mathrm{h}}, \\ -1, & \text{otherwise}, \end{cases} \tag{7}$$

where $[\boldsymbol{p}_i, \boldsymbol{p}_j]$ indicates a line segment between points $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$. The term $[\boldsymbol{p}_i, \boldsymbol{p}_j] \in \mathcal{R}_{\mathrm{h}}$ indicates that the line segment $[\boldsymbol{p}_i, \boldsymbol{p}_j]$ is inside the hand region $\mathcal{R}_{\mathrm{h}}$. In our framework, for simplicity, the middle point $\boldsymbol{p}_{\mathrm{m}}^{iS_j}$ is used to judge whether the line segment between points $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$ is in the hand region. Thus, Eq. (7) can be reformulated as

$$\delta = \begin{cases} 1, & \boldsymbol{p}_{\mathrm{m}}^{iS_j} \in \mathcal{R}_{\mathrm{h}}, \\ -1, & \text{otherwise}. \end{cases} \tag{8}$$

Based on Eqs. (3) and (5), given any point $\boldsymbol{p}_i$ on the hand contour $\mathcal{C}_h$, at a certain step s_i, the corresponding geodesic curve is denoted as $\Gamma_i^{s_j}$ with curvature value $\kappa_i^{s_j}$. For its geodesic curve set S_i in Eq. (4), let K_i denote its corresponding curvature set as S_i, formulated as

$$K_i = \{\kappa_i^{s_j} | 1 \le s_j \le M\}. \tag{9}$$

For N contour points $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_N\}$, their curvature set $\{K_1, K_2, \cdots, K_N\}$ constitutes a 2D curvature matrix with M rows and N columns, denoted as $\boldsymbol{\mathcal{M}}_{\mathrm{c}}$. The element on the jth row and ith column of $\boldsymbol{\mathcal{M}}_{\mathrm{c}}$ is actually the curvature $\kappa_i^{s_j}$ of the local geodesic curve of the ith contour point $\boldsymbol{p}_i$, denoted as $\mathcal{M}_{\mathrm{c}}^{ji}$. A curvature map with pseudo color is given in Fig. 3. The values vary due to different curvatures of contour points at different curvature steps.

Then, the 2D curvature matrix $\boldsymbol{\mathcal{M}}_{\mathrm{c}}$ is analyzed to find contour points on finger tips and finger valleys. A curvature histogram H is constructed to give statistics of curvatures on various steps for every contour point:

$$\begin{aligned} H &= \{h_1, h_2, \cdots, h_N\}, \\ h_i &= \sum_{j=1}^{N} \mathcal{M}_{\mathrm{c}}^{ij} = \sum_{j=1}^{N} \kappa_i^{s_j}, \end{aligned} \tag{10}$$

where h_i is the mean value of the curvatures of contour point $\boldsymbol{p}_i$ at various steps. A curvature histogram is shown in Fig. 3b.

Given the curvature histogram H, a 1D Gaussian kernel is first adopted to smooth it, and then the candidate points on finger tips and valleys are selected according to the following criteria:

$$\hat{P}_{\text{tip}} = \{\boldsymbol{p}_i | h_i > h_{i-1} \text{ and } h_{i+1} > h_i\},$$
$$\hat{P}_{\text{val}} = \{\boldsymbol{p}_i | h_i < h_{i-1} \text{ and } h_{i+1} < h_i\}. \quad (11)$$

Based on point sets P_{tip} and P_{val}, five points with the largest histogram values are selected as the final finger tip points (P_{tip}), and four points with the smallest histogram values are selected as finger valley points (P_{val}) (Fig. 3). The five finger tip points and four finger valley points can be ordered by their indexes on the contour. The finger length $L_{\text{f}} = \{l_{\text{f}}^{(1)}, l_{\text{f}}^{(2)}, \cdots, l_f^{(5)}\}$ can be obtained by calculating the geodesic distance from finger tips to finger valleys:

$$P_{\text{tip}} = \{\boldsymbol{p}_{\text{tip}}^{(1)}, \boldsymbol{p}_{\text{tip}}^{(2)}, \boldsymbol{p}_{\text{tip}}^{(3)}, \boldsymbol{p}_{\text{tip}}^{(4)}, \boldsymbol{p}_{\text{tip}}^{(5)}\},$$
$$P_{\text{val}} = \{\boldsymbol{p}_{\text{val}}^{(1)}, \boldsymbol{p}_{\text{val}}^{(2)}, \boldsymbol{p}_{\text{val}}^{(3)}, \boldsymbol{p}_{\text{val}}^{(4)}\},$$
$$l_{\text{f}}^{(1)} = |\boldsymbol{p}_{\text{tip}}^{(1)} - \boldsymbol{p}_{\text{val}}^{(1)}|,\ l_{\text{f}}^{(5)} = |\boldsymbol{p}_{\text{tip}}^{(5)} - \boldsymbol{p}_{\text{val}}^{(4)}|,$$
$$l_{\text{f}}^{(k)} = \min\{|\boldsymbol{p}_{\text{tip}}^{(k)} - \boldsymbol{p}_{\text{val}}^{(k-1)}|, |\boldsymbol{p}_{\text{tip}}^{(k)} - \boldsymbol{p}_{\text{val}}^{(k)}|\}, k = 2, 3, 4. \quad (12)$$

The order $\{\boldsymbol{p}_{\text{tip}}^{(1)}, \boldsymbol{p}_{\text{tip}}^{(2)}, \boldsymbol{p}_{\text{tip}}^{(3)}, \boldsymbol{p}_{\text{tip}}^{(4)}, \boldsymbol{p}_{\text{tip}}^{(5)}\}$ is little finger, ring finger, middle finger, forefinger, and thumb.

Based on this order, the curvature matrix $\boldsymbol{\mathcal{M}}_{\text{h}}$ of the hand contour can be re-formulated with the last column corresponding to the thumb's tip point $\boldsymbol{p}_{\text{tip}}^{(5)}$. After the five finger tips are determined (Fig. 3d), the curvature matrix in Fig. 3a is re-formulated using $\boldsymbol{p}_{\text{tip}}^{(5)}$ as reference. Thus, all hand contours can be represented by this geodesic curve curvature matrix $\boldsymbol{\mathcal{M}}_{\text{h}}$, which contains details of the hand's biometric contour features. Geometric features are extracted by analyzing the hand's curvature matrix in our framework, which will be described in detail in Section 2.5.1.

2.4 Three-dimensional hand registration

2.4.1 Hand rotation invariance

To achieve hand rotation invariance which is a prerequisite for the subsequent hand feature extraction, the orientation of the hand must be calculated

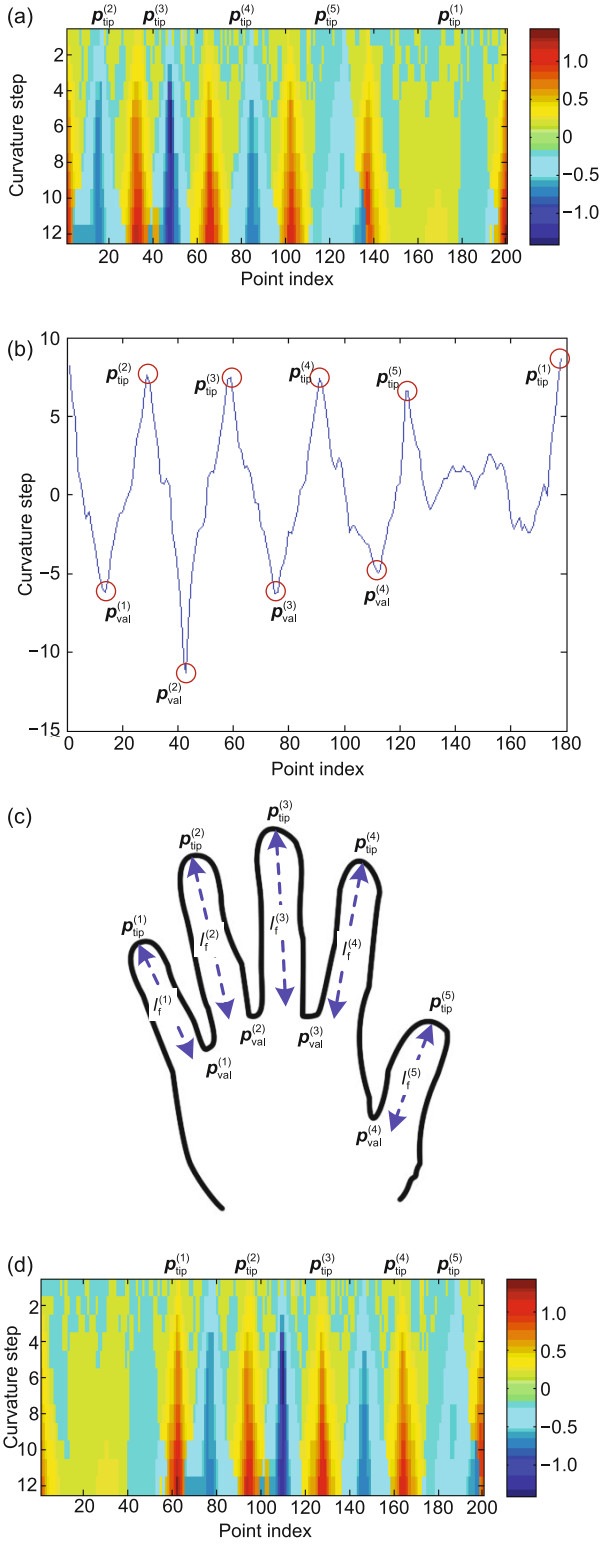

Fig. 3 (a) Original curvature matrix; (b) Curvature histogram with five peak points and four valley points marked by circles; (c) Illustration of the tip and valley points found along the length of the finger; (d) Re-formulated curvature matrix

first. Given any point $\boldsymbol{p}_i$ on boundary $\mathcal{C}_\mathrm{h}$, and supposing $\boldsymbol{p}_m^{s_j}$ is the middle point between points $\boldsymbol{p}_{i-s_j}$ and $\boldsymbol{p}_{i+s_j}$, the orientation of point $\boldsymbol{p}_i$ is calculated by averaging local orientations of a contour point at various steps:

$$\boldsymbol{o}_i = \frac{1}{M}\sum_{j=1}^{M}\frac{\boldsymbol{p}_i - \boldsymbol{p}_m^{s_j}}{|\boldsymbol{p}_i - \boldsymbol{p}_m^{s_j}|}. \tag{13}$$

In our framework, the orientation of point $\boldsymbol{p}_{\mathrm{tip}}^{(3)}$ on the middle finger tip is defined as the principal orientation of the hand, denoted as $\boldsymbol{o}_\mathrm{h}$ (Fig. 4a).

In this hand rotation invariant module, we consider only 2D in-plane rotation. Suppose the 2D orientation of the hand can be formulated as $\boldsymbol{o}_h = (o_x, o_y)$. The rotation matrix can be given as

$$\boldsymbol{R}_\theta = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \theta = \arctan(\frac{o_x}{o_y}). \tag{14}$$

The rotation matrix $\boldsymbol{R}_\theta$ is then used to correct the rotation of the hand on the intensity images. Suppose the original intensity data in the 2D image plane can be represented as

$$\boldsymbol{I}_{2\mathrm{D}} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \\ i_1 & i_2 & \cdots & i_n \end{bmatrix}, \tag{15}$$

where i_k $(k = 1, 2, \cdots, n)$ indicates the intensity value of pixel (u_k, v_k) of intensity image $\boldsymbol{I}_{2\mathrm{D}}$, and (x_k, y_k) the corresponding coordinates in the camera coordinates after DIT (Eq. (2)). The rotation invariant hand data can be given by

$$\hat{\boldsymbol{I}}_{2\mathrm{D}} = \boldsymbol{R}_\theta \boldsymbol{I}_{2\mathrm{D}}. \tag{16}$$

2.4.2 Hand pose invariance

To achieve an invariant description of a hand from all view points, besides 2D in-plane rotation invariance, 3D out-of-plane rotation caused by different hand poses must be corrected. In our framework, hand pose invariance is achieved by rotating 3D hand points according to the normal vector of the palm plane. Four valley points $\{\boldsymbol{p}_{\mathrm{val}}^{(1)}, \boldsymbol{p}_{\mathrm{val}}^{(2)}, \boldsymbol{p}_{\mathrm{val}}^{(3)}, \boldsymbol{p}_{\mathrm{val}}^{(4)}\}$ are used for palm plane fitting (Fig. 4). The fitting results and normal vector of the palm plane are shown (Fig. 4c).

Let $\boldsymbol{n} = (n_x, n_y, n_z)$ denote the normal vector of the palm plane. The 3D rotation matrix can be

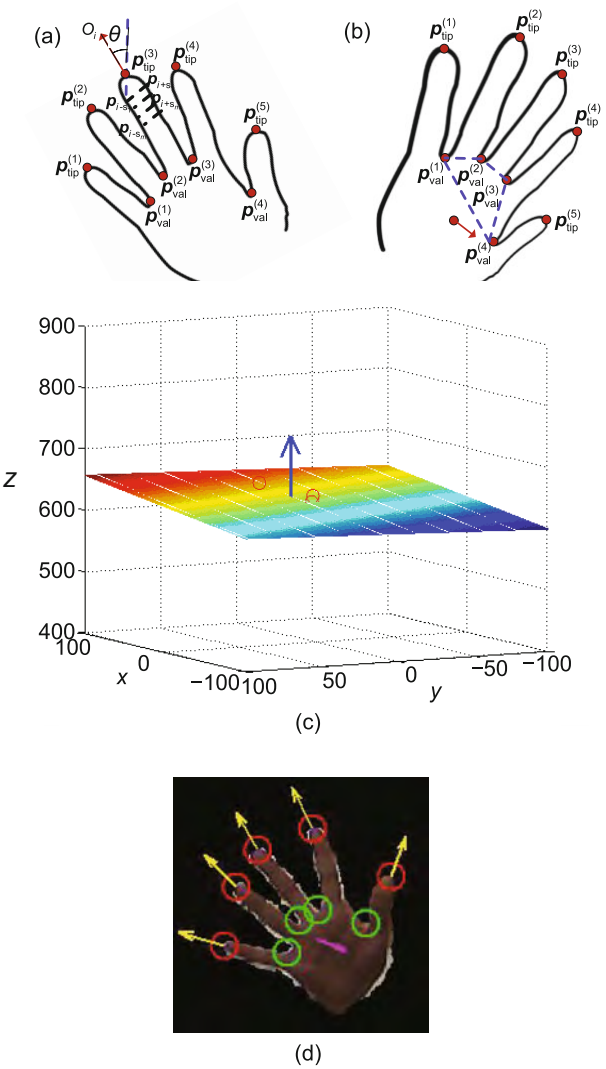

Fig. 4 (a) Hand 2D orientation; (b) Hand 3D orientation and the normal vector of the palm plane; (c) Plane fitting for the palm; (d) The final results of the segmented hand with finger tips (red circles), finger valleys (green circles), finger orientations (yellow arrows) and the palm's 3D orientation (purple arrow). References to color refer to the online version of this figure

formulated as

$$\boldsymbol{R}_{\boldsymbol{n}} = \begin{pmatrix} \cos\theta_y & 0 & \sin\theta & 0 \\ \sin\theta_x \sin\theta_y & \cos\theta_x & -\sin\theta_x \cos\theta_y & 0 \\ -\cos\theta_x \sin\theta_y & \sin\theta_x & \cos\theta_x \cos\theta_y & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{17}$$

where

$$\theta_x = -\arctan\left(\frac{n_y}{n_z}\right), \quad \theta_y = \arctan\left(\frac{n_x}{n_z}\right). \tag{18}$$

Suppose the 3D hand points with intensity

values can be represented as

$$\boldsymbol{I}_{3\mathrm{D}}=\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \\ z_1 & z_2 & \cdots & z_n \\ i_1 & i_2 & \cdots & i_n \end{bmatrix}, \tag{19}$$

where (x_k, y_k, z_k) $(k=1,2,\cdots,n)$ are the 3D coordinates of any point in hand regions and i_k its corresponding intensity value. Then, hand pose invariance can be achieved by the following 3D rotation transformation:

$$\hat{\boldsymbol{I}}_{3\mathrm{D}}=\boldsymbol{R_n}\boldsymbol{I}_{3\mathrm{D}}. \tag{20}$$

Thus, through 2D in-plane rotation and 3D out-of-plane rotation, hands with various poses are registered to the same view. The hand registration results of our framework will be evaluated in Section 3.2. The next step is to extract hand biometric features from the registered hand.

2.5 Biometric feature extraction

2.5.1 Geometric features

Geometric features can be applied in recognition as described by Malassiotis *et al.* (2006) and Kumar and Zhang (2007). As mentioned in Section 2.3, a hand contour can be represented by the geodesic curve curvature matrix $\boldsymbol{\mathcal{M}}_{\mathrm{h}}$, which contains details of the hand's geometric contour features. In this framework, we extract the main geometric features of the hand by conducting singular value decomposition (SVD), formulated as follows:

$$\boldsymbol{\mathcal{M}}_{\mathrm{h}}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{\Delta} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{V}^{\mathrm{H}}, \tag{21}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}^{\mathrm{H}}$ are unitary matrices and $\boldsymbol{\Delta}$ is a diagonal matrix with singular values of $\boldsymbol{\mathcal{M}}_{\mathrm{h}}$, denoted as

$$\boldsymbol{\Delta}=\mathrm{diag}\{\delta_1,\delta_2,\cdots,\delta_r\},\ \ \delta_i=\sqrt{\lambda_i}, i=1,2,\cdots,r. \tag{22}$$

Here r is the rank of $\boldsymbol{\mathcal{M}}_{\mathrm{h}}$ and λ_i is one of its singular values.

Based on the SVD of $\boldsymbol{\mathcal{M}}_{\mathrm{h}}$, an implicit geometric feature vector $\boldsymbol{v}_{\mathrm{g}}$ is formulated as a normalized vector:

$$\boldsymbol{v}_{\mathrm{g}}=\left(\hat{\delta}_1,\hat{\delta}_2,\cdots,\hat{\delta}_r\right),\ \ \hat{\delta}_i=\frac{\delta_i}{\sum_{k=1}^{r}\delta_k}. \tag{23}$$

Besides the implicit geometric feature vector $\boldsymbol{v}_{\mathrm{g}}$, an explicit geometric feature vector $\boldsymbol{v}_{\mathrm{f}}$ considering finger length is used in our framework:

$$\boldsymbol{v}_{\mathrm{f}}=\left(\hat{l}_{\mathrm{f}}^{(1)},\hat{l}_{\mathrm{f}}^{(2)},\cdots,\hat{l}_{\mathrm{f}}^{(n)}\right),\ \ \hat{l}_{\mathrm{f}}^{(i)}=\frac{l_{\mathrm{f}}^{(i)}}{\sum_{k=1}^{r}l_{\mathrm{f}}^{(k)}}, \tag{24}$$

where $l_{\mathrm{f}}^{(k)}$ is the finger length as given in Eq. (12).

The final geometric feature of a hand can be described as a concatenated vector obtained by weighting two normalized feature vectors $\boldsymbol{v}_{\mathrm{g}}$ and $\boldsymbol{v}_{\mathrm{f}}$, formulated as

$$\boldsymbol{v}_{\mathrm{gf}}=\omega_{\mathrm{g}}\boldsymbol{v}_{\mathrm{g}}+\omega_{\mathrm{f}}\boldsymbol{v}_{\mathrm{f}},\ \ \omega_{\mathrm{g}}{+}\omega_{\mathrm{f}}{=}1. \tag{25}$$

2.5.2 Intensity features

To increase the robustness of our system, similar to many previous studies (Choras and Choras, 2006; Kanhangad *et al.*, 2011b; Michael *et al.*, 2012), we combine hand geometry features with palmprint features for hand biometric feature description.

In this framework, geometric features are extracted through geodesic contour analysis using depth data. Through the hand registration module, depth data and intensity data in the hand region are registered with rotation and pose invariance. The next step is to extract palmprint features from intensity data after hand registration, which ensures the palmprint features remain invariant to various hand poses and orientations. There is another problem worthy of consideration, which is the scale problem: due to perspective effects, the projection scale of an object on the 2D image plane changes when its depth changes. To eliminate the scale problem, the hand region is often normalized to the same scale. However, this leads unavoidably to a loss of biometric features in the hand description, as different hands inherently have different sizes.

In our framework, to handle the scale problem for better hand feature extraction, rather than simply normalizing all hands to the same scale, we convert hand regions on the image plane to depth invariant image coordinates which have the same scale as the real world coordinates. For any point (u_i, v_i) on the intensity image plane, the transformation is conducted by

$$\begin{aligned} u_{\mathrm{DI}}^{i}&=\alpha'\left(W_{\mathrm{off}}+d_i(u_i-u_0)\frac{1}{f_u}\right),\\ v_{\mathrm{DI}}^{i}&=\alpha'\left(H_{\mathrm{off}}+d_i(v_i-v_0)\frac{1}{f_v}\right), \end{aligned} \tag{26}$$

where (u_0, v_0, f_u, f_v) are the camera's intrinsic parameters and α' is the scale factor. Here, W_{off} and H_{off} are set to make sure the minimum values of u^i_{DI} and v^i_{DI} are 1. d_i is the corresponding depth of (u_i, v_i). This is similar to the DIT in Eq. (2) which transforms the point (u_i, v_i) on the image plane to the 3D camera coordinates. Actually, this transformation can be regarded as projecting the 3D points with intensity data of the hand region onto a 2D image plane, without any perspective effects. That is to say, the size of hand projection on the transformed image plane depends only on its real size in the real world, which is depth-invariant.

Based on this useful attribute, palmprint features are extracted following the method proposed by Kong and Zhang (2004). This method uses a competitive coding scheme for palmprint verification. The competitive coding scheme uses multiple 2D Gabor filters to extract orientation information from palm lines. This information is then stored in a feature vector called the 'competitive code'. The angular matching with an effective implementation is then defined for comparing the proposed codes in nearly real time.

In our framework, feature vectors for hands of the same person can be simply trained and tested using the K-nearest-neighbors (KNN) method, for both geometric and palmprint features. As this module is not the main contribution of our work, details will not be described, but the results of experiments using geometric and palmprint features will be given in the experiments section.

3 Experiments and discussion

To demonstrate the effectiveness of our proposed method, extensive experiments were conducted on a PC connected to a Kinect sensor. The algorithm was processed on a Pentium i3-2100 3.10 GHz CPU with 2 GB RAM.

As discussed in Section 2.1, since there is no publicly available 3D hand database where hand images are obtained in a contactless and posture-unconstrained manner, we have developed our database using an inexpensive Kinect sensor. Participants in the data collection process conducted in our experiments included 30 students in our laboratory who volunteered to have their hand biometric features analyzed by our system. Left and right hand images were acquired from these 30 students. To test our approach under different pose variations, students were instructed to present their left hands in various poses.

For all experiments, parameters were set as follows: in Eq. (1) $\omega_{\text{d}} = 1000$, in Eq. (9) $M = N/16$, in Eq. (25) $\omega_{\text{g}} = \omega_{\text{f}} = 0.5$, in Eq. (26) $\alpha' = 0.05$, and in Eq. (27) (Section 3.3), $N_1 = 30$. Size of the depth image was 640×480, and of the color image was 1280×960.

3.1 Orientation and landmark point detection

The first group of experiments evaluated the performances of finger orientation and landmark point detection. As discussed in Section 2.3, geodesic contour analysis is conducted before extracting hand geometry features. Landmark points and finger orientations were obtained through geodesic contour analysis for hand registration. No strict illumination or background conditions are required in our personal identification system (Fig. 5). The complex and challenging conditions in our testing sequences included unstable illumination, cluttered background, and various 2D rotation and out-of-plane poses.

Finger orientations and landmark points were detected in depth images and then plotted on the color images in Fig. 5. Finger orientations and landmark points were well detected even in poorly illuminated and cluttered backgrounds with hands rotated. As in the last sequence in Fig. 5, the user's finger orientations and landmark points were detected outdoors with poor illumination. This is a tough problem for many state-of-the-art systems lacking depth information. With finger orientations and landmark points accurately obtained, hand registration and geometric feature extraction can be well conducted.

3.2 Performance of hand registration

The second group of experiments evaluated the performance of hand registration with 2D rotation and various poses. To make the hand geometry extraction more accurate, hand registration was conducted before feature extraction. For various hand rotations in a plane, we rotated the hand back to a position with the hand pointing upwards. An example of hand registration is illustrated in Fig. 6.

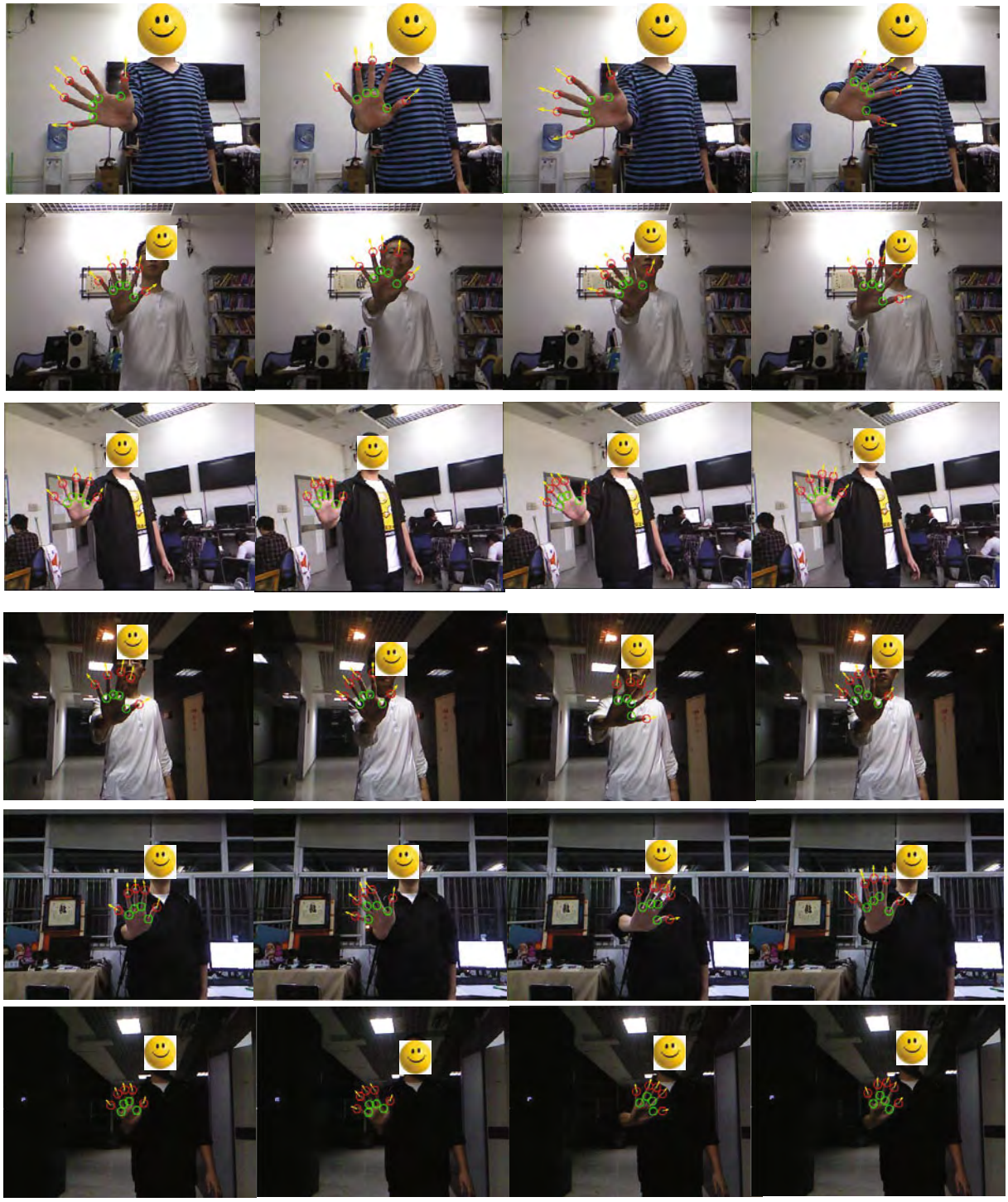

Fig. 5 Finger orientation and landmark point detection in six real and challenging environments. Orientations of fingers are plotted with a yellow arrow. Landmark points like finger tips and finger valleys are plotted as red and yellow circles, respectively. References to color refer to the online version of this figure

3.3 Hand geometry features

The third group of experiments evaluated the performance of hand geometry feature extraction. The geometric features extracted in our system are the length of the fingers and the SVD of the hand curvature matrix. To evaluate the accuracy of finger length calculation, we define an accuracy rate R as

$$R=\frac{1}{N_1}\sum_{i=1}^{N_1}\left(1-\frac{|l_i^{\mathrm{r}}-l_i^{\mathrm{c}}|}{l_i^{\mathrm{r}}}\right), \tag{27}$$

where N_1 is the number of users in our experiments, l_i^{c} is the calculated length of a finger of the ith user, and l_i^{r} is the real length of the finger.

To evaluate the effectiveness of the proposed approach, verification experiments were performed on the acquired database. In the first set of experiments, we evaluated the improvement of hand pose registration. Hand geometry features and palmprint features were evaluated before and after registration.

Table 1 gives the accuracy rates of four fingers of

Fig. 6 Hand registration based on finger orientation. Twelve real registration examples are given; hands are rotated to face the camera and point upwards. The subfigures on the left are the original color images and those on the right are the color images being rotated with the hand pointing upwards. To show clearly the performance of 2D rotation registration, we rotated the whole color image to illustrate the hand registration

the left hand, including the little finger, ring finger, middle finger, and forefinger. To evaluate the depth invariant performance of our system, we determined to test the accuracy rates with different distances between the hand and the Kinect sensor. The distance did not affect the performance of the finger length calculation (Table 1). The reason is that we have projected the coordinates of points on the hand to the real world, so the length calculation uses real-world coordinates instead of camera coordinates.

Table 1 Accuracy rates R of hand finger length calculations at various distances between the hand and the camera

Distance (cm)	Accuracy rate (%)			
	Little finger	Ring finger	Middle finger	Forefinger
60	95.1	94.5	92.7	96.2
70	96.2	95.4	94.6	93.8
80	94.7	93.6	92.9	97.1
90	95.5	94.7	94.4	93.6
100	93.2	95.2	95.8	95.7

To test the discrimination ability of geometric features, the identification rates for six challenging sequences are given in Table 2. The identification rate is an average identification rate in a sequence, calculated as $p = n_c/n$, where n_c is the number of correct identifications and n is the number of tests. From Table 2, we find that fusion of features can increase robustness, as concluded by Michael *et al.* (2012).

Table 2 The identification rates based on different geometric features for six challenging sequences

Sequence	Identification rate (%)		
	Finger length	Hand curvature	Fused
1	81.2	80.4	81.5
2	79.4	82.7	82.4
3	81.1	82.7	83.8
4	84.5	83.8	84.2
5	83.5	84.6	84.7
6	82.3	81.4	82.9

3.4 Performance of the identification system

The fourth group of experiments evaluated the performance of our hand-biometric-based identification system. The correct identification rates p for our hand geometry and color features based identification system are given in Fig. 7. All experiments followed a leave-one-out strategy. We performed a full matching test on all the hand images, and each experiment was conducted five times.

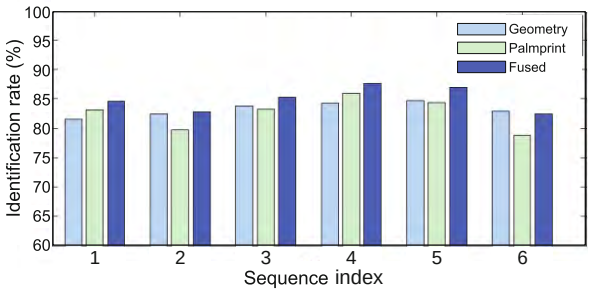

Fig. 7 Performance of the hand-biometric-based personal identification system. Geometric features, RGB features, and fused features were applied to identify people in six sequences

We evaluated the performance of geometric features based, color features based, and fused features based identification systems. A system based on the fusion of features is more robust than systems based on geometric or color features (Fig. 7). Geometric features give better results than color features, as geometric features are not sensitive to illumination or background conditions.

4 Conclusions and future work

In this paper, a contactless personal identification system is proposed based on matching hand geometry features and color features. An inexpensive Kinect sensor is used to acquire depth and color images of hands. Previous hand-biometric-based personal identification systems were severely limited by factors such as distance, background, illumination, and viewpoint. In our method, these limitations can be overcome. Extensive experiments have shown that hands can be registered and normalized based on finger and palm orientations. Real hand geometry data can be obtained, unaffected by the distance or viewpoint of the camera. Hand geometry features and color features used in our system show good discrimination ability. In future work, we will focus on developing a more natural and robust hand-biometric-based personal identification system for real environments. More robust features will be extracted to discriminate hands better.

References

Choras, R.S., Choras, M., 2006. Hand shape geometry and palmprint features for the personal identification. 6th

Int. Conf. on Intelligent Systems Design and Applications, p.1085-1090. [doi:10.1109/ISDA.2006.253763]

Dai, J., Feng, J., Zhou, J., 2012. Robust and efficient ridge-based palmprint matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, **34**(8):1618-1632. [doi:10.1109/TPAMI.2011.237]

Kanhangad, V., Kumar, A., Zhang, D., 2010. Human hand identification with 3D hand pose variations. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops, p.17-21. [doi:10.1109/CVPRW.2010.5543236]

Kanhangad, V., Kumar, A., Zhang, D., 2011a. Contactless and pose invariant biometric identification using hand surface. *IEEE Trans. Image Process.*, **20**(5):1415-1424. [doi:10.1109/TIP.2010.2090888]

Kanhangad, V., Kumar, A., Zhang, D., 2011b. A unified framework for contactless hand verification. *IEEE Trans. Inform. Forens. Secur.*, **6**(3):1014-1027. [doi:10.1109/TIFS.2011.2121062]

Kong, A., Zhang, D., 2004. Competitive coding scheme for palmprint verification. Proc. 17th Int. Conf. on Pattern Recognition, p.520-523. [doi:10.1109/ICPR.2004.1334184]

Kumar, A., Zhang, D., 2007. Hand geometry recognition using entropy-based discretization. *IEEE Trans. Inform. Forens. Secur.*, **2**(2):181-187. [doi:10.1109/TIFS.2007.896915]

Malassiotis, S., Aifanti, N., Strintzis, M.G., 2006. Personal authentication using 3-D finger geometry. *IEEE Trans. Inform. Forens. Secur.*, **1**(1):12-21. [doi:10.1109/TIFS.2005.863508]

Methani, C., Namboodiri, A.M., 2009. Pose invariant palmprint recognition. *LNCS*, **5558**:577-586. [doi:10.1007/978-3-642-01793-3_59]

Michael, G.K.O., Connie, T., Teoh, A.B.J., 2012. A contactless biometric system using multiple hand features. *J. Vis. Commun. Image Represent.*, **23**(7):1068-1084. [doi:10.1016/j.jvcir.2012.07.004]

Morales, A., Ferrer, M.A., Diaz, F., *et al.*, 2008. Contact-free hand biometric system for real environments. 16th European Signal Processing Conf., p.1-5.

Morales, A., Ferrer, M.A., Travieso, C.M., *et al.*, 2012. Multisampling approach applied to contactless hand biometrics. IEEE Int. Carnahan Conf. on Security Technology, p.224-229. [doi:10.1109/CCST.2012.6393563]

Ramalho, M., Correia, P., Soares, L., 2011. Distributed source coding for securing a hand-based biometric recognition system. 18th IEEE Int. Conf. on Image Processing, p.1825-1828. [doi:10.1109/ICIP.2011.6115820]

Ribaric, S., Fratric, I., 2005. A biometric identification system based on eigenpalm and eigenfinger features. *IEEE Trans. Pattern Anal. Mach. Intell.*, **27**(11):1698-1709. [doi:10.1109/TPAMI.2005.209]

Sanchez-Reillo, R., 2000. Hand geometry pattern recognition through Gaussian mixture modelling. Proc. 15th Int. Conf. on Pattern Recognition, p.937-940. [doi:10.1109/ICPR.2000.906228]

Sanchez-Reillo, R., Sanchez-Avila, C., Gonzalez-Macros, A., 2000. Biometric identification through hand geometry measurements. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22**(10):1168-1171. [doi:10.1109/34.879796]

Wang, C., Liu, H., Liu, X., 2013. Maximally stable curvature regions for 3D hand tracking. IEEE Int. Conf. on Image Processing, p.3895-3899. [doi:10.1109/ICIP.2013.6738802]

Woodard, D., Flynn, P., 2005. Finger surface as a biometric identifier. *Comput. Vis. Image Understand.*, **100**(3):357-384. [doi:10.1016/j.cviu.2005.06.003]

Xiong, W., Toh, K., Yau, W., *et al.*, 2005. Model-guided deformable hand shape recognition without positioning aids. *Pattern Recogn.*, **38**(10):1651-1664. [doi:10.1016/j.patcog.2004.07.008]

Zhang, D., Kong, W., You, J., *et al.*, 2003. Online palmprint identification. *IEEE Trans. Pattern Anal. Mach. Intell.*, **25**(9):1041-1050. [doi:10.1109/TPAMI.2003.1227981]