This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2018.2873490, IEEE Sensors Journal

Two User Adaptation-Derived Features for Biometrical Classifications of User Identity in 3D-Sensor-Based Body Gesture Recognition Applications

Ing-Jr Ding and Zong-Gui Wu

***Abstract*—For human gesture recognition applications, 3D-sensor-based approaches have received considerable attention and are crucial for future applications of an advanced body sensor network (BSN). A practical gesture recognition system is to apply active gesture patterns of a gesture-making user for body action classifications. 3D-sensor-based gesture recognition, categorized as biometric recognition in BSN applications, is greatly lack of the extensible cognition ability due to substandard recognition accuracy on the gesture-making user identity. To overcome this problem, this study proposes an active gesture-based user identity recognition approach using a robust feature design, called user adaptation (UA) features, derived from a UA process. Two different UA features, namely *Eigen Centroid-UA* and *Eigen Transform-UA* features, were developed in this study to accurately represent the adaptive learning tendency of gesture recognition for a specific gesture-making user. Compared with traditional 3D sensor gesture-based identity recognition approaches that employ only the feature of fixed body skeleton information without any UA designs, the presented UA-feature can exhibit fine adaptation learning continuously to the specific action user and therefore, superior identity recognition accuracy will be constantly ensured. To demonstrate the efficiency and effectiveness of the developed robust UA features in this paper, experiments on gesture-making user identification by Gaussian mixture model (GMM) applying the proposed *Eigen Centroid-UA* feature and verification by support vector machine (SVM) applying the proposed *Eigen Transform-UA* feature were conducted.**

***Index Terms*—3D sensor, body skeleton, gesture recognition, identity recognition, eigenspace, user adaptation, feature**

I. Introduction

FOR IMPLEMENTING an advanced body sensor network (BSN)-based application in smart home or health care areas, contactless sensors can be alternatives to wearable sensors. Audio-based BSN applications with acoustic sensors, such as speech recognition and speaker recognition applications, have been developed for obtaining voice characteristics for further analysis [1]–[3]. Video-based BSN applications with RGB image sensors, such as face recognition and facial emotion recognition applications, have been frequently used for obtaining facial image characteristics for further analysis. Currently, with rapid developments in three-dimensional (3D) image sensors, the maturity of 3D sensing techniques has been improved and the categories of 3D-sensor-based BSN applications have been widely extended [4]–[6]. Kinect [4], Xtion Motion [5], and Leap Motion [6] sensors comprising both RGB and additional depth cameras (i.e., the well-known RGB-D camera) are prevalently used 3D sensor devices for detecting human body actions. These sensors can effectively resolve a real-time calculation problem, which is difficult for a conventional RGB sensor [7], [8]. The aforementioned 3D sensors may be efficiently employed to establish a BSN environment for biometric recognition. Currently, biometric recognition using an active person's gesture data acquired using a 3D sensor has become a new human computer interaction (HCI) scheme (i.e., 3D-sensor-based gesture recognition). However, the effectiveness of 3D gesture feature designs in developing gesture recognition systems with satisfactory recognition performance levels is still a problem that must be addressed. Among the three sensor devices used in the aforementioned 3D-sensor-based gesture recognition techniques, the use of the Kinect sensor for developing a human gesture recognition system (the so-called Kinect 3D sensor recognition) has attracted the most attention recently because of the sensor's relatively early introduction to the market.

Kinect 3D sensor recognition by biometric human gesture data have two main application fields, which are gesture command recognition (or shortly called gesture recognition) and gesture-based user recognition. Gesture-based user recognition can be further categorized into gesture-based user identification and user verification. However, conventional 3D gesture feature designs in gesture recognition or gesture-based user recognition systems do not take into considerations user adaptation (UA). Such feature without UA considerations is seriously lack of identical characteristics of the specific user's gesture actions and can cause dissatisfactory system recognition performances. For addressing this problem, this study will present two UA features, the *Eigen Centroid-UA* feature and the *Eigen Transform-UA* feature, which will be efficiently employed in gesture-based user identification and user verification, respectively.

The use of the Kinect sensor for gesture recognition has been seen in numerous relevant studies [9]–[18]. These Kinect-sensor-based gesture recognition techniques can be primarily divided into the following categories: hidden-Markov-model-based gesture recognition using Kinect 3D-sensing data features [9]–[11], gesture template matching

This research is partially supported by the Ministry of Science and Technology (MOST) in Taiwan under Grant No. MOST 106-2221-E-150-058.

Ing-Jr Ding is with the Department of Electrical Engineering, National Formosa University, Huwei Township, Yunlin County, Taiwan (e-mail: ingjr@nfu.edu.tw).

Zong-Gui Wu. was with the Department of Electrical Engineering, National Formosa University, Huwei Township, Yunlin County, Taiwan.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2018.2873490, IEEE Sensors Journal

and gesture feature designs using a dynamic time warping method to rapidly classify the different gesture actions of specific gesture-making active users [12]–[14], an eigenspace approach with principal component analysis (PCA) for gesture command recognition [15], and system constructions of Kinect-sensor-based applications for specific purposes [16]–[18]. Although several studies have investigated Kinect-sensor-based related gesture recognition, extremely few studies have performed UA to implement an initial user-independent gesture recognition model for a user-adaptive gesture recognition system. In fact, such UA estimations obtained by properly adjusting gesture recognition model parameters using active gesture data of the user can considerably increase gesture recognition performance.

Studies on 3D sensor gesture-based user recognition have also been found recently, and those studies have directly applied a user's active gesture represented by Kinect 3D coordinate (x, y, z) in user identification or verification [19]–[27]. Kinect 3D sensor data have also been used to monitor the active behavior of elderly people [19]–[23]. The primary aims of these studies reported on [19]–[23] were to analyze specific activities of elderly staff and then perform user identification by specific gait gestures (walking behavior) through Kinect sensor surveillance. Feature-based approaches have been employed to establish a system for gesture-based user identification [24], [25]. However, the approaches presented in [24] and [25] are essentially feature-based designs and still lack recognition classification models; these approaches presented in [24] and [25] cannot perform UA calculations for parameter adjustments in the gesture recognition model to further improve recognition accuracy. Zhang et al. [26] presented a personalized gesture interaction system with the Kinect sensor to provide user identification based on gesture password sets. Wu et al. [27] presented two-stream deep learning neural networks for learning user styles to implement gesture-based user verification and identification systems. Although the gesture recognition systems in [26] and [27] learn a user's gesture styles to perform gesture-based identification or verification, they still use traditional 3D (x, y, z) data without any UA information for finely determining the gesture feature. Studies have yet to explore a feature design primarily incorporating the UA information of a specific user estimated from the parameter adjustment data of the gesture recognition model.

In this study, an user recognition approach based on active gesture data of the user was developed. The developed user recognition system including both user identification and verification employs an eigenspace-based 3D sensor gesture recognition approach to derive the UA feature information for identity classifications. The Kinect sensor gesture-based user recognition approach proposed in the current study is essentially based on the natural user experience (NUX) thought-line. Kinect sensor gesture-based human computer interface designs with a fine interface that considers NUX have been primarily studied in [28] and [29]. The key aspect of the gesture-based user recognition method in this study is the NUX-like UA feature design, where user-adaptive estimation can be obtained from eigenspace-based gesture model adjustment parameters and variation information between the initial user-independent and user-adaptive eigenspaces of Kinect 3D gesture data. Compared with Kinect 3D sensor recognition studies by active gesture data, this study has the following advantages:

- This study proposes a gesture command recognition system involving both user gesture command recognition and user identity recognition, in contrast to those used in [9]–[18] that perform only gesture recognition.
- This study presents a robust feature design that excellently estimates user-adaptive information from initial and user-adaptive gesture recognition models that can considerably increase the recognition rate with an increase in the UA time, in contrast to conventional Kinect-sensor-based active gesture recognition systems with limited recognition accuracy [19]–[25] that consider only Kinect 3D data features and disregard any UA information in feature designs.
- The study presents a UA feature design with improved recognition performance in user identification or verification, in contrast to the less robust feature design in [26] and [27] that applies only traditional 3D (x, y, z) location data for gesture-based user classification.

II. EIGENSPACE-BASED 3D SENSOR BODY GESTURE RECOGNITION USING KINECT 3D (x, y, z) FEATURES (*EIGEN3DGESTURE* APPROACH)

The author's previous study [15] presented a human gesture recognition system applying the *Eigen3Dgesture* approach based on the eigenspace scheme and PCA, with Kinect 3D (x, y, z) space location data serving as the features. Different to robust feature designs incorporated with UA schemes in this paper, the presented *Eigen3Dgesture* approach in the previous work of [15] adopts the feature for gesture classifications by just conventional 3D (x, y, z)-skeleton joint location information. The following sections introduce conventional 3D gesture recognition systems that use the Kinect sensor to acquire 3D sensing features for rapid gesture recognition, as well as the *Eigen3Dgesture* approach that applies the Kinect sensor to obtain 3D sensing features.

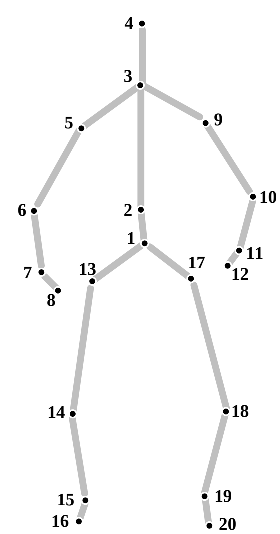

Fig. 1. 3D-sensing data obtained from the Kinect sensor to reveal the location information of the kernel joint in a human body skeleton

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2018.2873490, IEEE Sensors Journal

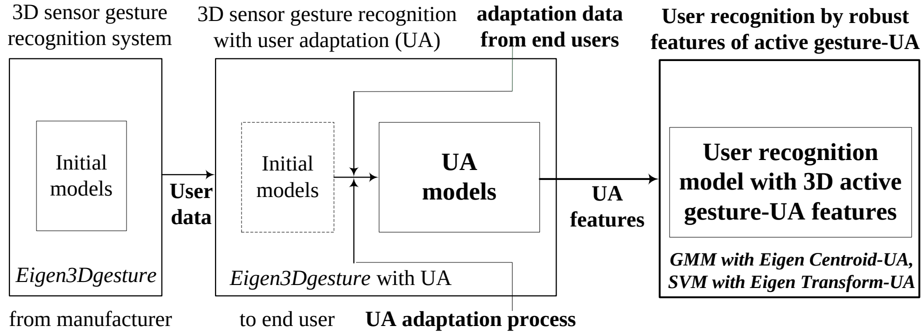

Fig. 2 The presented architecture for active gesture-based user identity recognition using the developed robust user adaptation (UA) features, *Eigen Centroid-UA* and *Eigen Transform-UA*

A. Utilization of the Kinect 3D Sensor to Extract 3D Sensing Data for Rapid Body Gesture Recognition

Fig. 1 presents a human skeleton with numerous kernel joints obtained using the Kinect sensor. From the acquired human skeleton, the 3D (*x*, *y*, *z*) information of the primary kernel joints of the human body in the person's active space can be rapidly estimated using the Kinect software development kit from Microsoft. Such 3D (*x*, *y*, *z*) data of the human body joints obtained from RGB and depth (RGB-D) sensors are also termed Kinect 3D data. Kinect 3D data enable easily implementing a gesture recognition system. As illustrated in Fig. 1, variations of 20 relative positions of joints in 3D (*x*, *y*, *z*) coordinate data can be obtained when the user makes a specific gesture for a certain continuous period [e.g., 90 frames, each containing 20 (*x*, *y*, *z*)-coordinate information acquired in 3 s].

B. Eigenspace-Based Approach Involving Principal Component Analysis for Body Gesture Recognition Through Traditional 3D Sensor-(x, y, z) Features (Eigen3Dgesture approach)

The previously presented *Eigen3Dgesture* approach [15] for 3D-gesture recognition contains two primary phases, namely PCA applied to the aforementioned 3D (*x*, *y*, *z*) information of the Kinect 3D data and training and recognition test of the eigenspace *Eigen3Dgesture*. The following eigen-decomposition procedure is performed to establish the *Eigen3Dgesture* space after PCA operations:

$$C \cdot \xi = \lambda \cdot \xi \, , \tag{1}$$

where ξ and λ are the eigenvector and the corresponding eigenvalue, respectively. To achieve the set threshold of the error rate, only *K* eigenvectors in Eq. (1) are employed to establish the 3D gesture eigenspace. These *K* eigenvectors (i.e., *K* supervectors) form the "*K-space*" of the 3D gestures, as indicated by the following transformation matrix:

$$A = \begin{bmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1M} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{K1} & \xi_{K2} & \cdots & \xi_{KM} \end{bmatrix}_{KM} . \tag{2}$$

The original large-sized gesture data X_{MN} of $M \times N$, where *N* represents the training gesture data associated with a specific active gesture and *M* represents the data dimension, are transformed to relatively small-sized data with a dimension of *K* by projecting the transformation matrix *A*:

$$G = A \cdot \left(X_j - u \right), \quad j = 1,2,3,\ldots,N \, , \tag{3}$$

where u, X_j, and *G* represent the mean vector of X_{MN}, the *j*th eigenvector, and the new eigenspace of the 3D gesture with the dimension *K* after matrix projection, respectively. The corresponding *Q* centroids of a gesture recognition system with *Q* gesture commands are distributed in the *Eigen3Dgesture* space for recognizing a test gesture pattern. In the recognition phase, the *Eigen3Dgesture*-based approach employs the preestablished *Q* centroids, each of which has *K* dimensions— $Centroid_{qk}$, *q* = 1, 2,…, *Q*, and *k* = 1, 2,…, *K*—for classifying the test gesture command. A recognition technique for determining the minimum physical distance between the test gesture pattern and all preestablished *Q* gesture command centroids is adopted.

III. User Identity Authentication in Eigenspace-Based 3D Sensor Body Gesture Recognition Using the Proposed Gesture-UA Features with User-Adaptive Information

Fig. 2 illustrates a schematic for the efficient utilization of variations in the eigenspace generated using 3D human gesture recognition for obtaining the desired active gesture-based user identity recognition models. As shown in Fig. 2, this study primarily adopted the eigenspace approach involving PCA for human gesture recognition; the initial model represents the user-independent gesture command recognition model without any recognition system parameter adjustments using a specific user's adaptation gesture data (i.e., the initial model does not involve adaptive system learning). After the appropriate UA of the initial eigenspace gesture recognition model is determined, the initial model gradually becomes a user model that can sufficiently indicate identical gesture characteristics of the specific user. Assume that a total of *P* users perform active gestures in the gesture recognition system comprising *Q*-specified gesture commands; the corresponding *P* gesture UA models are established, each of which particularly denotes the identical gesture characteristic of the specified user from the *P* users among these *Q* available gesture commands. The

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2018.2873490, IEEE Sensors Journal

Algorithm to compute the *Eigen Centroid-UA* feature
Set adaptation mode (*r* values) of incremental UA-Eigen3Dgesture model with the *r*-th adaption.
For (*i* = 1 to *r*)
/*Acquisitions of UA model with the *r*-th adaptation*/
Input the *i*-th active gesture pattern obtained from the user.
3D-sensor data feature extraction on the *i*-th active gesture pattern.
Perform gesture recognition to the *i*-th active gesture pattern by initial *Eigen3Dgesture* model.
Check the class label (*q* value in *Q* specified gesture commands) of the recognition result.
Update the centroid value of the *q*-th gesture command class in the *Eigen3Dgesture* space.
End For
/*Estimations of the *r*-th *Eigen Centroid-UA* feature */
Calculate the difference of updated and original *Eigen3Dgesture* spaces of the user (see Eq. (4)).
Output the estimated UA feature, *Eigen Centroid-UA*.

Fig. 3. A pseudo-code explanation for detailed describing acquisitions of the designed UA feature, *Eigen Centroid-UA*

UA feature in Fig. 2 implies the model variation degree between "the UA model with the incremental adaptation of the *r*th system learning iteration using *r* times of adaptation data" and "the initial model without any UA." In this study, two types of UA features, *Eigen Centroid-UA* and *Eigen Transform-UA*, were developed. These features were effectively employed to develop the Gaussian mixture model (GMM)- and support vector machine (SVM)-based user verification systems, respectively.

A. GMM User Identification Using Active Gesture UA Features, Called Eigen Centroid-UA Features, Derived From Centroid Variations of Eigenspace Models

As mentioned, if the system comprises *P* users performing active gesture commands in the established gesture recognition system composed of *Q*-specified gesture commands, a total *P* UA eignespace models are available (i.e., overall *P* user gesture spaces of *Eigen3Dgesture*) for feature extraction of UA feature parameters. For the *p*th model of all these *P* *Eigen3Dgesture* models with UA in *r*th incremental learning using the adaptation data from the p^{th} user, the UA features of *Eigen Centroid-UA* for the pattern features to establish the user identity recognition system of the *p*-th user will be calculated using Eq. (4) as follows,

$$\begin{aligned} &Eigen\,Centroid - UA(p_r) = \\ &Centroid_{qk}(p_r) - Centroid_{qk}(initial), \\ &q = 1,2,...,Q,\, k = 1,2,...,K,\; p = 1,2,...,P. \end{aligned} \tag{4}$$

As shown in (4), for a specific gesture-making user *p*, the presented UA feature of *Eigen Centroid-UA* is based on the degree of variation between the set of updated *Q* gesture command centroids of UA models with *r*th adaptation and the set of original *Q* gesture command centroids of initial models without any system parameter adjustments of UA. The dimension of the developed *Eigen Centroid-UA* feature for establishing the user identification model of each user is $Q \times K$, where each *Q* gesture command centroid is obtained from the estimated *Eigen3Dgesture* space of the corresponding specified gesture command. The gesture command centroid in the *Eigen3Dgesture* space is essentially a vector with *K* dimensions (*K*-space), as mentioned in Section II-B.

The GMM [30] with the acquired UA feature, the aforementioned *Eigen Centroid-UA*, is then employed to identify these gesture-making users. Assume that *P* users make active gesture commands in the gesture recognition system; a total of *P* GMM gesture-making user models must be developed using the *Eigen Centroid-UA* feature, each of which represents the corresponding *P* user. In the user identity recognition phase executed using these *P* appropriately trained GMM gesture-making user models, the following equation provides the recognition decision:

$$\begin{aligned} \hat{p} &= \max_{p=\{1,2,...,P\}} Prob(\lambda_p \mid UA) \\ &= \max_{p=\{1,2,...,P\}} \frac{f(UA \mid \lambda_p)}{f(UA)} \cdot Prob(\lambda_p) \end{aligned}, \tag{5}$$

where UA denotes the active gesture pattern of the test user, which comprises the estimated *Eigen Centroid-UA* features with a time interval covering a certain number of UA feature vectors of $Q \times K$ dimensions; λ_p denotes the GMM of the *p*th user for all *P* gesture-making users in the system; and $Prob(\lambda_p)$ denotes the prior probability of GMM user models of the *p*th user. The category of the test user's identity can be determined by maximizing the a *posteriori* probability $Prob(\lambda_p \mid UA)$; the gesture action test user is finally categorized as one of the *P* user classes according to the class label indicated by $\hat{p}$.

Fig. 3 shows pseudocode explaining the detailed derivation of the designed *Eigen Centroid-UA* feature for a specific gesture-making user in the eigenspace of *Eigen3Dgesture* containing *Q*-specified gesture commands.

B. SVM User Verification Through Active Gesture UA Features, Called Eigen Transform-UA Features, Estimated From Transformation Matrix Variations of Eignespace Models

As mentioned in Section II, the "*K*-space" can be used to preserve most information of the 3D sensor gesture features from a gesture-making person. Therefore, such *K*-space composed of *K* supervectors provides considerable benefit for deriving the UA features from the 3D eigensapce-based gesture recognition system. According to (3) in Section II-B, the transformation matrix *A* is projected for dimension reduction in order to transform the large-sized original gesture data set X_{MN} with $M \times N$ dimensions into a relatively small-sized eigenspace gesture data set G_{KN} with $K \times N$ dimensions, which is composed of these *K* supervectors. Therefore, the transformation matrix *A* with *K* supervectors can sufficiently provide the identical 3D gesture action characteristics from the specific user and can be directly employed to develop UA features for user identity recognition. As described in Section III-A, for all *P* gesture-making users in the gesture recognition system, for the corresponding

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2018.2873490, IEEE Sensors Journal

Eigen3Dgesture gesture recognition model with UA of the *p*th user, the UA model can be established using the *r*th incremental learning from the adaptation data of the *p*th user; moreover, the UA features of *Eigen Transform-UA* can be used as the pattern features to demonstrate the variation degree between the UA gesture model of the *r*th UA and the initial model without UA for constructing the user identity recognition model, which can be determined by (6):

$$Eigen\ Transform-UA(p_r)=A_{km}(p_r)-A_{km}(initial),\quad k=1,2,...,K,\ m=1,2,...,M,\ p=1,2,...,P. \tag{6}$$

In (6), variations between the updated transformation matrix of the UA *Eigen3Dgesture* model with the *r*th adaptation and the original transformation matrix of the initial *Eigen3Dgesture* model without UA can be completely represented in the proposed UA feature of *Eigen Transform-UA*. Fig. 4 illustrates the flowchart of the acquisition process of the designed *Eigen Transform-UA* feature.

The designed *Eigen Transform-UA* feature acquired using (6) can be efficiently used for SVM-based [31] user verification by the user's active gesture patterns, where *Eigen Transform-UA* serves as the pattern feature for the training and testing of the SVM classification model. All users in SVM identity recognition are categorized as follows: the valid user and the impostor. In SVM training, the *Eigen Transform-UA* feature parameters calculated using the active gesture data of the valid user and impostor groups are used to establish a separation hyperplane, which is as shown in (7):

$$y_{km}(w\cdot \mathrm{UA}+b)\geq 1-\xi_{km},\quad k=1,2,...,K,\quad m=1,2,...,M, \tag{7}$$

where UA represents the test pattern with the calculated *Eigen Transform-UA* feature parameters in a certain time interval to be verified. This study adopted a nonlinear SVM design because of misclassifications of the UA parameter, and a proper nonzero value was set for the corresponding ξ_i to appropriately satisfy the inequality of SVM classifications. As shown in (7), the trained SVM separation hyperplane with the parameter pair (w , b) can be used to verify the user in the test phase of the user verification procedure, and the classification label y_{km} ($y_{km}\in\{-1,1\}$) for $k=1,2,...,K, m=1,2,...,M$, can finally be used to assign the test user to the valid or impostor class.

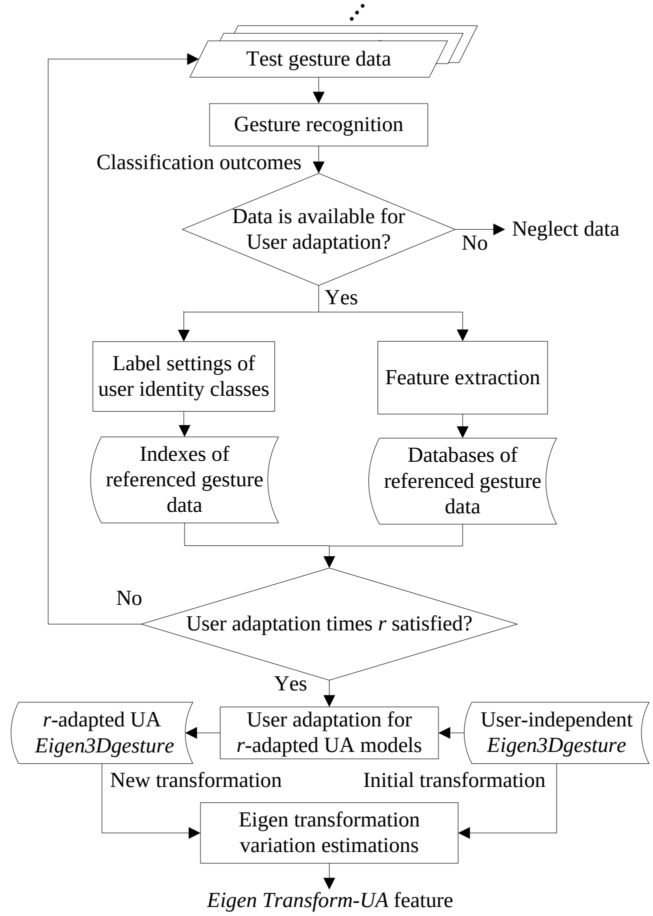

Fig. 4. A flow chart illustration for acquisitions of the designed UA feature, *Eigen Transform-UA* (from the *r*-adapted UA *Eigen3Dgesture*)

IV. EXPERIMENTS

In this study, active gesture-based user recognition experiments were performed in a laboratory environment. To make all the specified gesture actions, a total of eleven active gesture-making users, including eight male (subjects S1, S2,…, and S8) and three female (subjects S9, S10, and S11) users, were recruited. Seven specified gesture actions were designed in this study to be used as gesture commands in an HCI gesture recognition application. These gesture actions were stored in the gesture action database and are outlined as follows:

"Gesture-1: holding the right hand flat";
"Gesture-2: swinging the right hand in the clockwise direction";
"Gesture-3: swinging the right hand in the counterclockwise direction";
"Gesture-4: holding the left hand flat";
"Gesture-5: swinging the left hand in the clockwise direction";
"Gesture-6: swinging the left hand in the counterclockwise direction"; and
"Gesture-7: the standing pose."

Fig. 5 illustrates all these designed seven different gesture command actions, which are denoted as Gesture-1, Gesture-2,…, and Gesture-7, for command control applications to recognize the gesture-making user identity. The Kinect sensor produced by Microsoft has a frame rate of 30 in 3D image acquisition. Fig. 6 is an example to record the human body skeleton variations acquired by the Kinect 3D sensor when the gesture action "Gesture-1"is made by the user. Observed from Fig. 6, when the user performs the gesture command "Gesture-1," for conventional feature designs considering only 3D-(x, y, z) location values, just three joints P9, P10, and P11 (see Fig. 1 mentioned in Sec. 1) in the body skeleton vary significantly. However, conventional 3D-(x, y, z) features with such joint location variations are not proper for gesture-based user recognition. Fig. 7 depicts variations of 3D-(x, y, z) values of joints P9, P10, and P11 at the period of about one second when the gesture command action "Gesture-1" is performed by the user of subject S1. It is seen in Fig. 7, only slight variations at z-coordinate values for these three joint points appear, which explains the weakness of conventional 3D-(x, y, z) features on gesture-based user

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2018.2873490, IEEE Sensors Journal

recognition. The presented two robust UA features in this work will thoroughly tackle this problem.

Table I shows the user characteristics of the recruited eleven users for gesture-based user recognition evaluations, including mainly height, weight, age, upper-arm, and lower-arm lengths. It is noted that for upper-arm and lower-arm lengths in Table I, such arm length information is measured by the Kinect 3D sensor directly and therefore also scaled by 3D sensor acquisitions. Fig. 8 shows the effect of different distances between the user and the 3D sensor on acquired body skeleton information. It can be found from Fig. 8 that the obtained body skeleton information is uncertain and will greatly vary with the distance of the user and 3D sensor. Therefore, it can be seen in Table I that uncertain values of upper-arm and lower-arm lengths of each recruited user are acquired by the 3D sensor. It can be observed that due to distance uncertainty, conventional features composed of only 3D-(x, y, z) or additional user body characteristics (height and limb sizes) derived by skeleton will seriously lack of robustness and therefore can't be adaptive in a practical gesture-based user recognition environment.

The gesture database was determined to contain 1540 (11 × 7 × 20) active gestures, where the seven specified gesture commands were recorded 20 times for all eleven gesture-making users. The Kienct 3D image sensor was employed to obtain each active gesture-making user's body skeleton with all 3D (x, y, z) joint position data (i.e., Kinect 3D data, as mentioned in Section II) for rapidly establishing gesture recognition and active gesture-based user recognition systems. Table II is averaged recognition performances of gesture recognition by previous proposed *Eigen3Dgesture* approach [15] among these 11 users making 7 different specific gestures, which achieves 91.8%. In gesture-based user recognition experiments, half of the data in the database were used as the training data for establishing user classification recognition models, and the other half of data were used as the adaptation and test data for evaluating user identification and verification performances of the proposed UA feature scheme.

Table III shows the average recognition accuracy achieved using only conventional features of Kinect 3D data without any UA design for SVM-based user verification and GMM-based user identification among the 11 recruited users making the 7 gestures. SVM-based user verification by conventional Kinect 3D data achieved a recognition accuracy of 94.6%, and GMM-based user identification yielded a recognition accuracy of 92.9% (Table III). Although the conventional user verification or identification processes without UA features exhibits acceptable recognition accuracy levels, some gesture commands, such as Gesture-2, Gesture-6, and Gesture-7 in the SVM-based user verification process and Gesture-4 and Gesture-7 in the GMM-based user identification process, were still associated with dissatisfactory classification errors, which could not be improved due to the lack of UA.

Table IV presents the average recognition performance of SVM-based user verification using the presented *Eigen Transform-UA feature* for the eleven users making the seven gestures. The average user verification performance of the SVM classifier improved distinctively when the presented *Eigen Transform-UA feature* was derived with an increase in the number of adaptations (Table IV). At the beginning of SVM-based verification with the *Eigen Transform-UA* feature with two adaptations, the recognition accuracy was only a little substandard, which is 92.5%. For SVM-based verification using the *Eigen Transform-UA* feature with four adaptations, the recognition rate intensely increases by 3.1%. When using the presented UA feature with six adaptations, a satisfactory recognition accuracy of 96.9% was reached, after which, the performance was invariably maintained. For each specified gesture command action, SVM-based user verification using the *Eigen Transform-UA* feature exhibited significant improvement. After continuous adaptations, the *Eigen Transform-UA* feature will lead the recognition system to the state of recognition accuracy saturation. Fig. 9 shows the adaptation learning curve of SVM-based user verification performances by the developed *Eigen Transform-UA* feature.

Table V shows the average recognition accuracy of GMM-based user identification using the presented *Eigen Centroid-UA* feature for the eleven users making the seven specified gestures. Fig. 10 illustrates the adaptive learning curve of GMM-based user identification performances using the proposed *Eigen Centroid-UA* feature. The average recognition rate gradually increased with the amount of UA information incorporated with the *Eigen Centroid-UA* feature (Fig. 10). As can be seen in Fig. 10, on four adaptations (i.e. *r* = 4), the presented *Eigen Centroid-UA* feature performs user adaptation and outperforms the conventional feature using 3D-(x, y, z) information by 1.3%. On the state of recognition accuracy saturation, the GMM applied the *Eigen Centroid-UA* feature will achieve 95.5% recognition accuracy, indicating excellent user adaptation effects. For the effect of each gesture command on user identification performances, the adaptive learning curve of the proposed *Eigen Centroid-UA* feature was determined to be very significant, an average recognition accuracy improvement of 4.9% (from 90.6% to 95.5%).

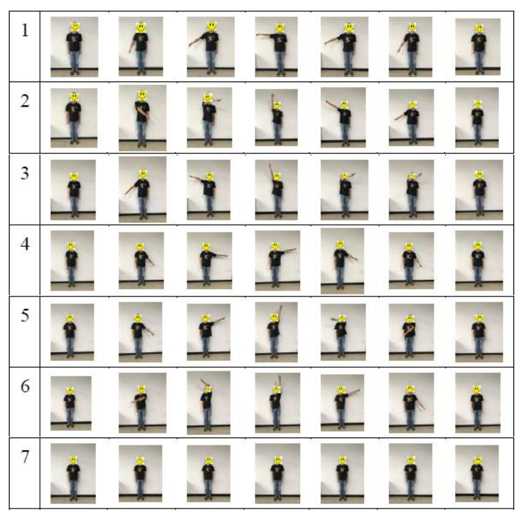

Fig. 5 Designed 7 different gesture actions (Gesture-1, Gesture-2,…, and Gesture-7) for command control applications to recognize the user identity

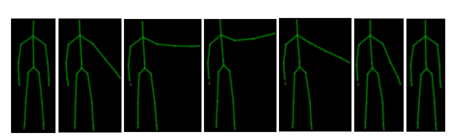

Fig. 6 Illustrations of variations of body skeleton information when the action "Gesture-1" is performed (also see Fig. 5)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2018.2873490, IEEE Sensors Journal

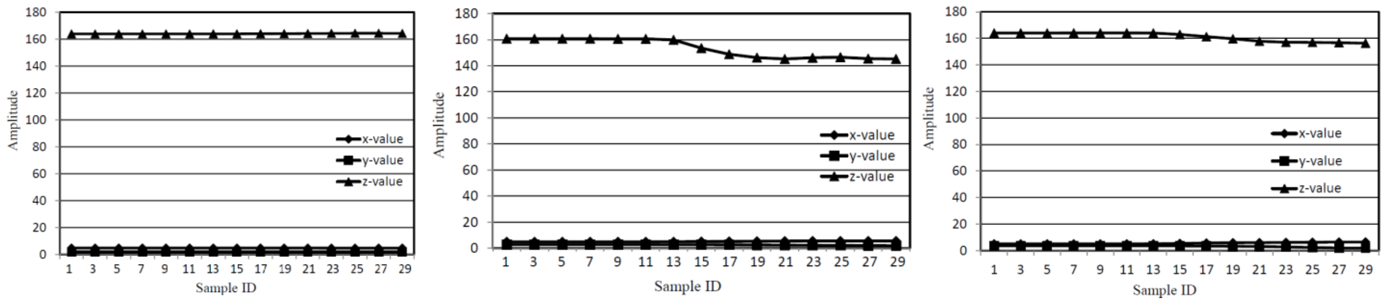

Fig. 7 Variations of 3D-(x, y, z) values of joints P9, P10, and P11(from right to left) at the period of about one second (sampling rates of Kinect 3D sensor = 30) when the gesture action “holding the right hand flat” (Gesture-1) is performed by the subject S1

TABLE I USER CHARACTERISTICS OF EACH OF THE COLLECTED USERS FOR GESTURE-BASED USER RECOGNITION EXPERIMENTS, INCLUDING HEIGHT, WEIGHT, AGE, UPPER-ARM, AND LOWER-ARM LENGTHS (UPPER-ARM AND LOWER-ARM INFORMATION WITH SCALES MEASURED DIRECTLY BY KINECT 3D SENSOR)

User characteristics of recruited users to perform seven gesture command actions for gesture-based user recognition											
Sex	Male								Female		
Subject (user)	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11
Height (cm)	178	173	168	178	172	175	170	170	158	158	160
Weight (kg)	70	63	61	85	50	78	60	65	45	63	52
Age	21	21	23	23	22	22	23	23	21	21	22
Upper-arm (P9 to P10)	119	137	98	131	163	130	193	138	83	146	116
Lower-arm (P10 to P11)	407	269	316	300	333	353	310	379	443	296	228

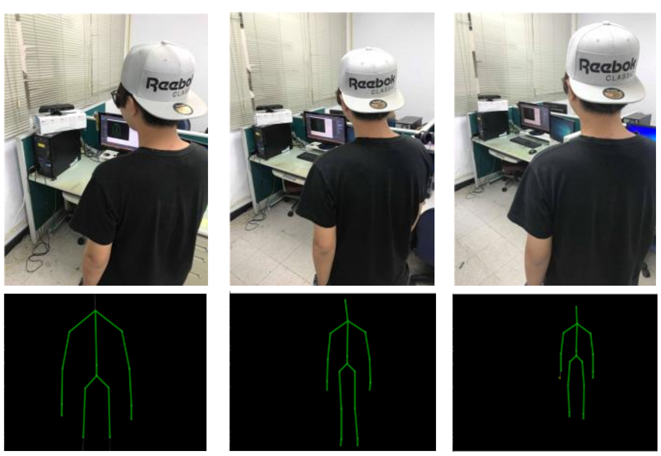

Fig. 8 Effects of distances between the user and the 3D sensor on acquired body skeleton information (from left to right: “too close,” “very proper,” and “too far”)

TABLE II AVERAGED RECOGNITION PERFORMANCES OF GESTURE RECOGNITION USING PREVIOUS PROPOSED *EIGEN3DGESTURE* APPROACH [15] AMONG 11 GESTURE-MAKING USERS THAT MAKE 7 DIFFERENT SPECIFIC GESTURES

Gesture commands	Recognition rates of gesture recognition (%) (gesture recognition using the previous proposed *Eigen3Dgesture* approach in [15])
Gesture-1	91.8
Gesture-2	93.6
Gesture-3	93.6
Gesture-4	90.9
Gesture-5	94.5
Gesture-6	95.5
Gesture-7	82.7
Averaged	91.8

TABLE III AVERAGED RECOGNITION PERFORMANCES OF CONVENTIONAL FEATURES OF KINECT 3D DATA WITHOUT ANY UA USED ON SVM USER VERIFICATION AND GMM USER IDENTIFICATION AMONG 11 GESTURE-MAKING USERS THAT MAKE 7 DIFFERENT SPECIFIC GESTURES

Gesture commands	Recognition rates of user recognition (%)	
	User recognition by conventional features of Kinect 3D data without any UA	
	SVM user verification	GMM user identification
Gesture-1	98.2	97.2
Gesture-2	90.9	99.1
Gesture-3	96.4	95.5
Gesture-4	98.2	83.6
Gesture-5	96.4	97.2
Gesture-6	92.7	95.5
Gesture-7	89.1	81.8
Averaged	94.6	92.9

TABLE IV AVERAGED RECOGNITION PERFORMANCES OF SVM USER VERIFICATION USING THE PRESENTED UA FEATURE, EIGEN TRANSFORM-UA, AMONG 11 GESTURE-MAKING USERS THAT MAKE 7 DIFFERENT SPECIFIC GESTURES

Gesture commands	Recognition rates of SVM user verification with UA features (%)				
	Adaptation times (r-values) on the *Eigen Transform-UA* feature				
	2	4	6	8	10
Gesture-1	93.6	95.5	98.2	98.2	98.2
Gesture-2	90.9	94.5	96.4	96.4	96.4
Gesture-3	96.4	98.2	98.2	98.2	98.2
Gesture-4	90.9	95.5	98.2	98.2	98.2
Gesture-5	95.5	98.2	98.2	98.2	98.2
Gesture-6	90.9	95.5	96.4	96.4	96.4
Gesture-7	89.1	91.8	92.7	92.7	92.7
Averaged	92.5	95.6	96.9	96.9	96.9

TABLE V AVERAGED RECOGNITION PERFORMANCES OF GMM USER IDENTIFICATION USING THE PRESENTED UA FEATURE, EIGEN CENTROID-UA, AMONG 11 GESTURE-MAKING USERS THAT MAKE 7 DIFFERENT SPECIFIC GESTURES

Gesture commands	Recognition rates of GMM user identification with UA features (%)				
	Adaptation times (*r*-values) on the *Eigen Centroid-UA* feature				
	2	4	6	8	10
Gesture-1	97.2	98.2	98.2	98.2	98.2
Gesture-2	91.8	96.4	99.1	99.1	99.1
Gesture-3	93.6	95.5	96.4	96.4	96.4
Gesture-4	83.6	89.1	90.9	90.9	90.9
Gesture-5	97.2	98.2	98.2	98.2	98.2
Gesture-6	89.1	94.5	96.4	96.4	96.4
Gesture-7	81.8	87.2	89.1	89.1	89.1
Averaged	90.6	94.2	95.5	95.5	95.5

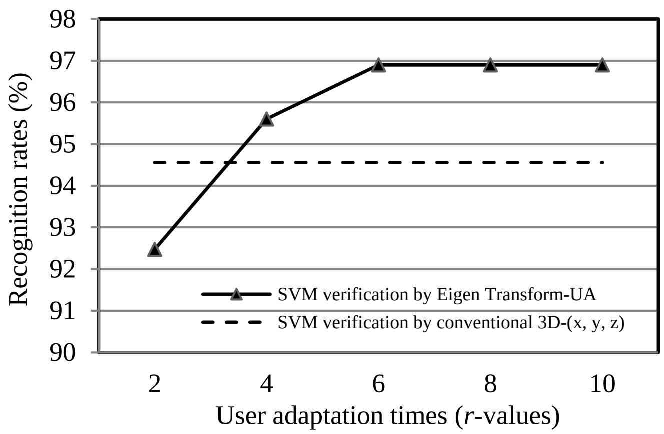

Fig. 9 The adaptation learning curve of SVM user verification using the presented UA feature of *Eigen Transform-UA*

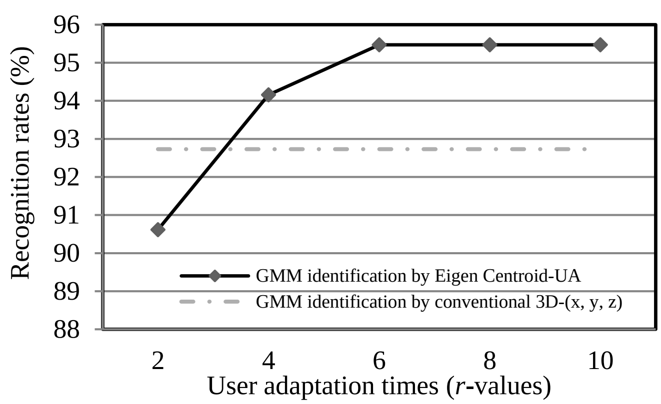

Fig. 10 The adaptation learning curve of GMM user identification using the presented UA feature of *Eigen Centroid-UA*

V. CONCLUSION

In this study, two robust UA features for active gesture-based user recognition, the *Eigen Centroid-UA* feature and *Eigen Transform-UA* feature, were developed on the basis of an eigenspace-based gesture recognition system using Kinect 3D data. The effectiveness and efficiency of the presented *Eigen Centroid-UA* and *Eigen Transform-UA* features in recognizing the identity of users making active gesture commands were demonstrated using GMM and SVM classification models for performing user identification and user verification tasks, respectively. Both active gesture-based user recognition approaches with UA feature designs were determined to have fine adaptive learning curves corresponding to recognition accuracy, indicating that the recognition approaches are evidently superior to conventional recognition using features composed of only Kinect 3D data.

ACKNOWLEDGEMENTS

This research is partially supported by the Ministry of Science and Technology (MOST) in Taiwan under Grant No. MOST 106-2221-E-150-058.

REFERENCES

[1] I.-J. Ding and S.-K. Lin, “Performance improvement of Kinect software development kit-constructed speech recognition using a client-server sensor fusion strategy for smart human-computer interface control applications,” *IEEE Access*, vol. 5, pp. 4154–4162, 2017.

[2] I.-J. Ding and J.-T. Liu, “Three-layered hierarchical scheme with a Kinect sensor microphone array for audio-based human behavior recognition,” *Computers & Electrical Engineering*, vol. 49, pp. 173–183, 2016.

[3] I.-J. Ding and C.-T. Yen, “Enhancing GMM speaker identification by incorporating SVM speaker verification for intelligent web-based speech applications,” *Multimedia Tools and Applications*, vol. 74, no. 14, pp. 5131–5140, 2015.

[4] I. Tashev, “Kinect development kit: A toolkit for gesture- and speech based human-machine interaction,” *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 129–131, 2013.

[5] A. S. Kundu, O. Mazumder, A. Dhar and S. Bhaumik “Occupancy grid map generation using 360° scanning xtion pro live for indoor mobile robot navigation,” in *Proc. 2016 IEEE First International Conference on Control, Measurement and Instrumentation*, Kolkata, India, 2016.

[6] W. Lu, Z. Tong and J. Chu, “Dynamic hand gesture recognition with Leap Motion controller,” *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1188–1192, 2016.

[7] R. Min, N. Kose and J.-L. Dugelay, “KinectFaceDB: a Kinect database for face recognition,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 11, pp. 1534–1548, 2014.

[8] Z. Zhang, L. Cui, X. Liu and T. Zhu, “Emotion detection using Kinect 3D facial points,” in *Proc. 2016 IEEE/WIC/ACM International Conference on Web Intelligence*, Omaha, NE, USA, 2016.

[9] J. Han, L. Shao, D. Xu and J. Shotton, “Enhanced computer vision with Microsoft Kinect sensor: a review, ” *IEEE Transactions on Cybernetics*, vol. 43, no. 5, pp. 2168–2267, 2013.

[10] A. Dubois and F. Charpillet, “Human activities recognition with RGB-Depth camera using HMM,” in *Proc. IEEE International Conference on Engineering in Medicine and Biology Society (EMBC)*, pp. 4666–4669, 2013.

[11] W. Xu and E. J. Lee, “Continuous gesture recognition system using improved HMM algorithm based on 2D and 3D Space,” *International Journal of Multimedia and Ubiquitous Engineering*, vol. 7, no. 2, pp. 335–340, 2012.

[12] I.-J. Ding and C.-W. Chang, “Feature design scheme for Kinect-based DTW human gesture recognition,” *Multimedia Tools and Applications*, vol. 75, no. 16, pp. 9669–9684, 2016.

[13] P. Saade, P. Joly and A. Awada, “Simulating actions for learning,” in *Proc. IEEE International Workshop of Electronics, Control, Measurement, Signals and their application to Mechatronics*, pp. 1–6, 2013.

[14] C. J. Su, “Personal rehabilitation exercise assistant with Kinect and dynamic time warping,” *International Journal of Information and Education Technology*, vol. 3, no. 4, pp. 448–454, 2013.

[15] I.-J. Ding and C.-W. Chang, “An eigenspace-based method with a user adaptation scheme for human gesture recognition by using Kinect 3D

data," *Applied Mathematical Modelling*, vol. 39, no. 19, pp. 5769–5777, 2015.

[16] K. Armin, Z. Mehrana and D. Fatemeh, "Using Kinect in teaching children with hearing and visual impairment," in *Proc. International Conference on E-Learning and E-Teaching*, pp. 86–90, 2013.

[17] A. Yargic and M. Dogan, "A lip reading application on MS Kinect camera," in *Proc. International Symposium on Innovations in Intelligent Systems and Application*, pp. 1–5, 2013.

[18] I.-J. Ding and Y.-J. Chang, "On the use of Kinect sensors to design a sport instructor robot for rehabilitation and exercise training of the elderly," *Sensors and Materials*, vol. 28, no. 5, pp. 463–476, 2016.

[19] Aniruddha Sinha and Kingshuk Chakravarty, "Pose based person identification using Kinect," in *Proc. IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 497–503, 2013.

[20] M. Machado-Molina, I. Bönninger, M. K. Dutta, T. Kutzner and C. M. Travieso, "Gait-based recognition of humans using Kinect camera," *Recent Advances in Computer Engineering, Communications and Information Technology*, pp. 63–71, 2014.

[21] L. Igual, À. Lapedriza and R. Borràs, "Robust gait-based gender classification using depth cameras," *EURASIP Journal on Image and Video Processing*, 2013.

[22] J. Preis, M. Kessel, M. Werner and C. Linnhoff-Popien, "Gait recognition with Kinect," in *Proc. the 1st International Workshop on Kinect in Pervasive Computing*, 2012.

[23] E. Hossain and G. Chetty, "Person identification in surveillance video using gait biometric cues", in *Proc. the 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pp. 1877–1881, 2012.

[24] M. Ferrara, A. Franco and D. Maio, "On the use of the Kinect sensor for human identification in smart environments," *Journal of Ambient Intelligence and Smart Environments*, vol. 6, no. 4, pp. 435–446, 2014.

[25] J. Wu, J. Konrad and P. Ishwar, "Dynamic time warping for gesture-based user identification and authentication with Kinect," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2371–2375, 2013.

[26] H. Zhang, W. Wu and Y. Lou, "A personalized gesture interaction system with user identification using Kinect," in *Proc. Pacific Rim International Conference on Artificial Intelligence (PRICAI)*, pp. 614–626, 2014.

[27] J. Wu, P. Ishwar and J. Konrad, "Two-stream CNNs for gesture-based verification and identification: learning user style," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2016.

[28] W. Xu and E. J. Lee, "A new NUI method for hand tracking and gesture recognition based on user experience," *International Journal of Security and Its Applications*, vol. 7, no. 2, pp.149–158, 2013.

[29] J. Pan and Y. P. Guan, "Human-computer interaction using pointing gesture based on an adaptive virtual touch screen," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 6, no. 4, pp. 81–91, 2013.

[30] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 1, pp. 72–83, 1995.

[31] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

Ing-Jr Ding received the B.S. degree from Chang-Gung University in 1999, M.S. degree from National Central University in 2001, and Ph.D. degree from National Chiao-Tung University in 2008. From March 2009 to July 2009, Dr. Ding served as a project assistant professor in the Graduate Institute of Automation and Control (GIAC), National Taiwan University of Science and Technology (NTUST). Dr. Ding was an assistant professor in the Department of Electrical Engineering (EE), National Formosa University (NFU) during the period of August 2009 to July 2012. Since August 2012, he has been an associate professor in the Department of EE, NFU. His research interests include signal process and recognition, artificial intelligence, and multimedia techniques.

Zong-Gui Wu received the B.S. degree and M.S. degree from the Department of Electrical Engineering, National Formosa University in 2013 and 2015, respectively. He is interested in 3D sensing techniques and pattern recognition.