A LOCAL FEATURE-BASED APPROACH FOR ACTIVITY 3-D SKELETON POSTURE RETRIEVAL

MD. ZIA UDDIN

Department of Computer Education, Sungkyunkwan University, Seoul, Republic of Korea
E-mail: ziauddin@skku.edu

Abstract- In this work, a novel method is proposed that utilizes depth information of human body to retrieve appropriate posture from a database. The retrieval method utilizes Enhanced Independent Component Analysis (EICA) of depth postures of different classes over which Linear Discriminant Analysis (LDA) is performed. Then, the postures of each class are ranked in sorted order based on the distance to their class center. To retrieve a testing posture from the training database, the posture is projected on the feature space based on which its class is determined by measuring the distances to all class centers. Finally, a binary search is done in the corresponding class to find the nearest distance of the corresponding posture and the one is chosen with the least distance.

Keywords- EICA, LDA, Depth, Posture.

I. INTRODUCTION

Lately, video cameras are very commonly used for various surveillance applications [1]. Among which, human body posture analysis from video data gets a considerable attention recently due to its potential use for human computer interfaces, which can contribute to a proactive system by responding to the user's conditions. The most representative methodology for posture analysis is extracting some features from postures and comparing with each trained features.

To represent the complete posture of human body, binary postures are most commonly employed from which useful features are derived [4]-[9]. In 1992, the authors proposed a binary posture-based human activity recognition system where the time-sequential postures were transformed into a feature vector sequence where binary pixel-based mesh features were extracted from every image [4]. The posture features were then utilized to recognize several tennis strokes. This approach showed poor performance since it was completely dependent upon the pixels and hence very sensitive to noise. In 2004 and 2005, the authors proposed Principal Component Analysis (PCA) on binary posture features to represent several view-invariant human activities [5], [6]. In 2006, the authors used binary postures to recover 3-D human body pose [7]. In 2007, the author proposed a binary posture lookup based approach to articulate the 3-D pose of human movements [8]. In 2008, the authors proposed local features from the binary postures via Independent Component Analysis (ICA) to represent human body in different activities [9]. In [9], the authors applied PCA on the body postures before applying ICA, which is called Enhanced ICA or EICA. The binary postures were used in human gait analysis too [10], [11]. However, binary postures are not efficient enough to describe human body properly due to its two-level limited pixel intensity distribution although they have been very extensively applied in this regard. From the binary postures, it is impossible to obtain the body part differences based on the distances of them to the camera. To overcome that, depth information-based body posture representation was utilized in some works [12] where the pixel values were set based on the distance to the camera which can provide better body information than the binary postures.

In this work, a depth posture retrieval system is proposed from a posture database. To extract the posture features, the EIC features are extracted from the depth postures and further processed by Linear Discriminant Analysis (LDA) to obtain strong posture features based on which the corresponding 3-D skeleton postures are ranked based on the distance to the specific class. Later on, from testing depth postures, after extracting EICA-LDA features, the nearest 3-D skeleton posture is obtained based on the distance to the features of the specific class first and then, to find the appropriate posture, the nearest distant feature is obtained based on binary search as each distance of the training features to the class mean was saved as sorted order.

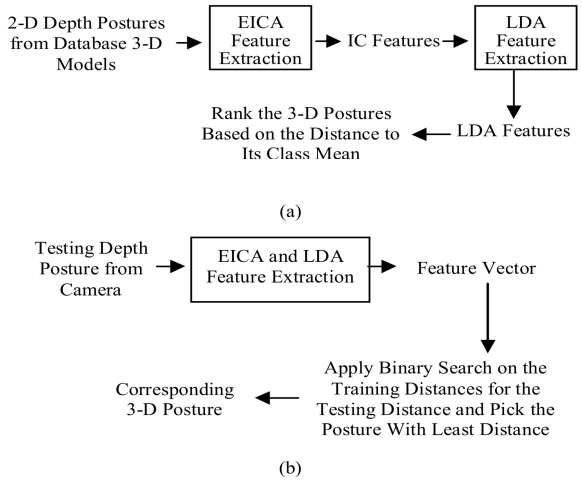


Fig. 1. Database Posture Ranking and (b) Testing Posture Retrieval.

II. PROPOSED METHODOLOGY

A strong feature space is generated via EICA and LDA to project all the feature vectors. The training posture features are then ranked based on the distance to their corresponding class means. To test a posture image, the testing feature's class is determined by measuring the distance to the mean features of all the classes. Finally, the corresponding posture from the specific class is chosen by binary searching the testing posture feature distance with the database posture feature distances to the same class. Fig. 1 shows the basic steps of ranking the database postures and retrieval of the corresponding posture for a testing posture.

A. Posture Acquisition

The major objective of posture feature extraction is to find efficient representation of body postures in an optimal feature space. Here, EICA is utilized to extract local features from the depth postures followed by LDA for further classification. The detailed description of the posture feature processing is given in the following sections. The RGB and depth images of different activities are acquired by Kinect, a commercial system that generates RGB and distance information for the objects captured by the camera. Figs. 2(a), 2(b), and 2(c) represent a sample RGB, corresponding binary, and depth images respectively in this regard. In the depth image, the higher pixel intensity indicates the near and the lower the far distance. For instance, in Fig. 2(c), the hand region is brighter than others as they are very near to the camera whereas the binary postures cannot even show the hands due to a flat pixel value distribution on all over the human body.



(a)



(b)



(c)

Fig. 2. (a) Sample grey image, (b) corresponding binary image, and (c) corresponding depth image of human posture.

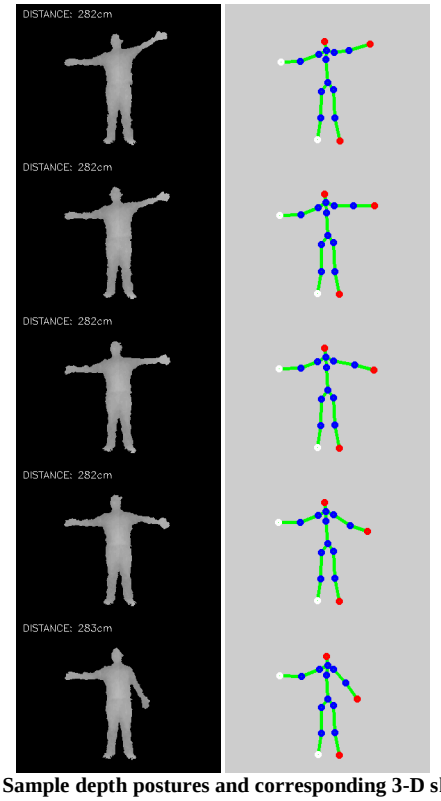


Fig. 3. Sample depth postures and corresponding 3-D skeleton models from hand movement activity.

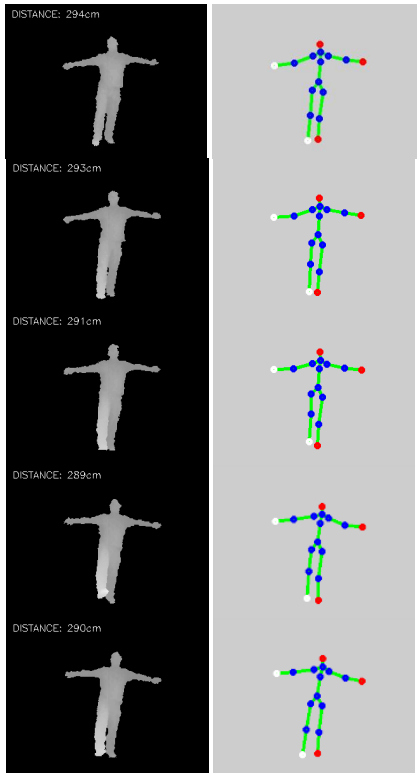
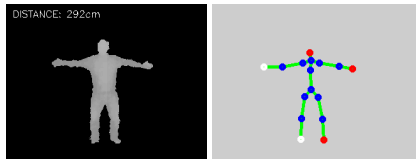


Fig. 4. Sample depth postures and corresponding 3-D skeleton models from leg movement activity.



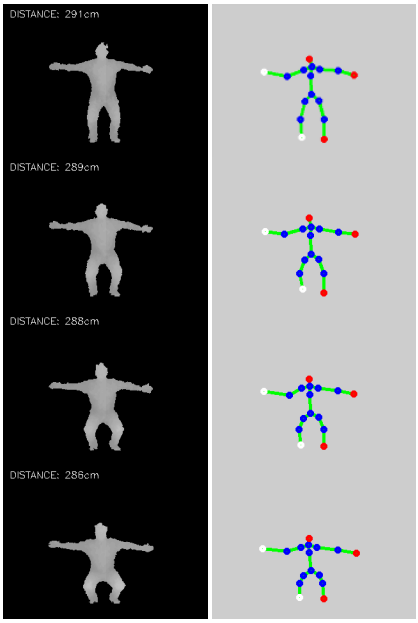


Fig. 5. Sample depth postures and corresponding 3-D skeleton models from sitting-standing movement activity.

B. Posture Feature Processing

In posture-based HAR, PCA produces the global features representing frequently moving parts of the human body in activities. Thus, after extracting the top m basis vectors (i.e., E_m) in the feature space, the posture vector representation in the feature space becomes $C_i = \tilde{X}_i E_m$ where C_i is the PCA projection of the i^{th} posture and $\tilde{X}_i$ zero-mean i^{th} posture vector to be applied for feature processing. Alternative to PCA, we can apply a higher order statistical approach called ICA to focus on the local feature extraction of the postures, which is called enhanced ICA or EICA. ICA is a blind source separation problem that decomposes a mixture of observed variables into a linear combination of some unknown statistically independent components and their mixing matrix. Let K and X be the collection of the basis and input posture vectors respectively. Thus, S and X can be modeled as

$$X = GK \quad (1)$$

where G is an unknown linear mixing matrix of full rank. Assuming the sources are independent of each other and the mixing matrix is invertible, the ICA algorithm tries to find the mixing matrix G or the separating matrix W such that

$$V = WX \quad (2)$$

where V represents estimated independent sources. In brief, the ICA algorithm tries to learn the weight matrix W to recover the set of independent basis images S. Fig. 6 shows 12 ICA basis images of the depth postures with the parts in high contrast representing the local body components used in the activities. The ICA algorithm on E_m^T produces $E_m^T = W^{-1}S$ and the reconstructed posture vectors can be represented as $X_r = UW^{-1}S$ where $U = XE_m$ is the projection of the postures X on E_m. Thus, the ICA projection J_i of a posture vector $\tilde{X}_i$ is as

$$J_i = \tilde{X}_i E_m W^{-1}. \quad (3)$$

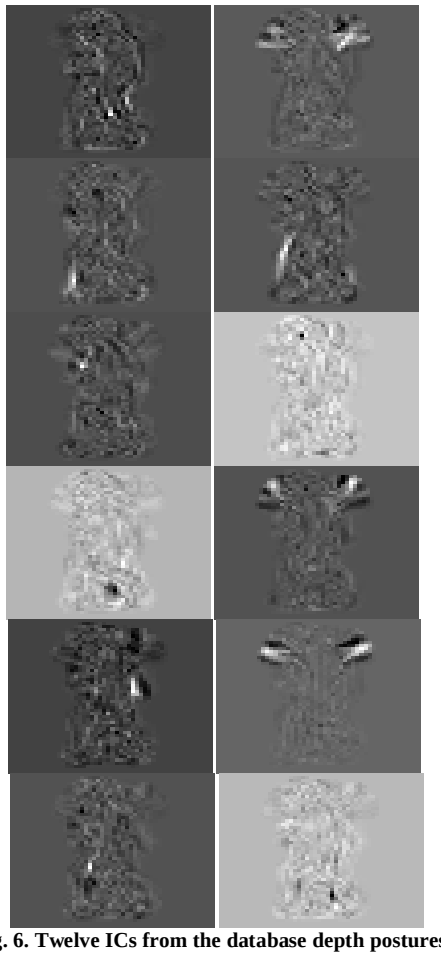


Fig. 6. Twelve ICs from the database depth postures.

Then, LDA is applied on EICA features for more robust feature extraction. LDA looks for the vectors in the underlying space that tries to best discriminates among the prototypes of different classes [14]. Thus, LDA can be applied on the IC or PC features of the postures of different activity classes to further classify them. Basically, the LDA algorithm tries to maximize B and minimize W so that their ratio can be maximum. Thus, the optimal discrimination matrix can be obtained as

$$O_{LDA} = \left|O^T BO\right| / \left|O^T WO\right|. \quad (4)$$

As a consequent, the LDA representation on the IC feature (i.e., F_i) for i^{th} mage vector can be defined as $F_i = J_i O_{LDA}^T$. Fig. 7 represents the 3-D plot of the

features after applying LDA on the EICA representations of the depth postures of different classes. After applying LDA, the class means are calculated and the postures of each class are ranked according to the ascending order of their distances to the class center. Usually, the raw depth activity images acquired by the camera consist of random noises. To reduce them, median filtering is performed on the depth images to make them smooth. Then, corresponding 3-D skeletons are obtained based on random forest-based body pixel labeling and 3-D skeleton joint position estimation. Figs 3, 4, and 5 show 12 depth postures as well as corresponding 3-D skeletons from hand movement, leg movement, and sitting-standing postures respectively.

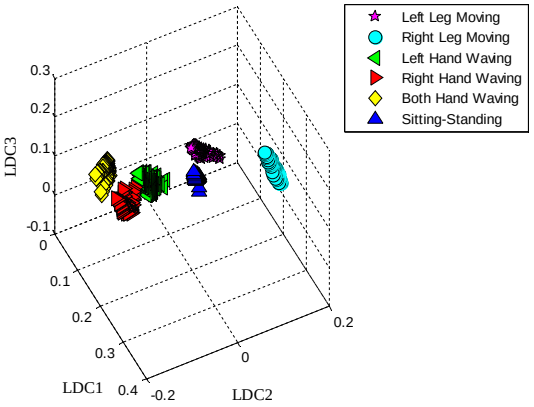


Fig. 7. 3-D plot of the features applying LDA on the IC features of the depth postures of the different classes.

C. A Testing Posture Retrieval

To retrieve a posture for a testing posture, after extracting EICA and LDA features, corresponding class is determined based on the least distance. Then, the distance is used to find out the corresponding posture from that class. To search the posture, binary search is applied here on the distances.

Table 1 Retrieval results based on the features from the binary and depth postures.

Feature	Posture	Correct Retrieval Rate
PCA	Binary	47 %
ICA	Binary	53%
LDA on PCA	Binary	44%
LDA on EICA	Binary	64%
PCA	Depth	83%
ICA	Depth	89%
LDA on PCA	Depth	87%
LDA on EICA	Depth	96%

III. EXPERIMENTAL SETUPS AND RESULTS

Our depth posture-based database was built where there were a total of 5,000 frames. Furthermore, each generalized posture was with the size of 64x64 pixels. A total of 1000 postures were used for testing. After applying EICA, 100 features were taken into the feature space. Four different feature extraction methods (i.e., PCA, LDA on the PCA features, ICA, and LDA on the EICA features) were utilized to evaluate the retrieval performances on the binary and depth postures respectively. Table 1 shows the success results of the binary and depth posture-based experiments utilizing the different feature extraction approaches where LDA on the EIC-features of depth postures show the highest retrieval rate i.e., 96%. Moreover, using the depth information-based effective approach, we can overcome the limitations of the binary postures to describe human body postures properly.

IV. CONCLUDING REMARKS

In this paper, a novel approach has been proposed for retrieving 3-D skeleton human posture from posture database using useful features of depth postures. It has been investigated here the IC-based local features followed by LDA on the depth postures obtained superior results over the other approaches including binary posture-based approaches. The proposed system could be used for modeling a good human computer interaction system to be applied in smart applications such as smart home healthcare.

ACKNOWLEDGEMENT

This paper was supported by Faculty Research Fund, Sungkyunkwan University, 2013.

REFERENCES

[1] J. Kim, J. Park, K. Lee, K.-H. Baek, and S. Kim, "A Portable Surveillance Camera Architecture using One-bit Motion Detection," IEEE Transactions on Consumer Electronics, 53(4), 1063 – 1068 (2007).

[2] H. Kang, C.W. Lee, and K. Jung: 'Recognition-based gesture spotting in video games', Pattern Recognition Letters, 25, 1701-1714 (2004).

[3] FS. Chen, CM Fu., and CL. Huang, "Hand gesture recognition using a real-time tracking method and Hidden Markov Models", Image and Vision Computing, 21, 745-758 (2005).

[4] J. Yamato, J. Ohya, and K. Ishii, "Recognizing human action in time-sequential images using hidden markov model", in Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 379-385 (1992).

[5] F. Niu and M. Abdel-Mottaleb, "View-invariant human activity recognition based on shape and motion Features", in Proceedings of the IEEE Sixth International Symposium on Multimedia Software Engineering, 546-556 (2004).

[6] F. Niu and M. Abdel-Mottaleb, "HMM-based segmentation and recognition of human activities from video sequences", in Proceedings of IEEE International Conference on Multimedia & Expo., 804-807 (2005).

[7] A. Agarwal and B. Triggs, "Recovering 3D human pose from monocular images", IEEE Transactions on Pattern Analysis and Machine Intelligence, 28, 44-58 (2006).

[8] N. R. Howe, "Posture lookup for monocular 3D pose tracking", Image and Vision Computing, 25, 331-341 (2007).

[9] M. Z. Uddin, J. J. Lee, and T.-S. Kim, "Shape-Based Human Activity Recognition Using Independent Component Analysis and Hidden Markov Model," in Proceedings of the 21st international conference on Industrial, Engineering, and Other Applications of Applied Intelligent Systems: New Frontiers in Applied Artificial Intelligence, 245-254 (2008).

[10] A. Kale, A. Sundaresan, A. Rajagopalan, N. Cuntoor, A. R. Chowdhury, V. Kruger, and R. Chellappa, "Identification of humans using gait", IEEE Transactions on Image Processing, 13, 1163-1173 (2004).

[11] T. H. W. Lam, R. S. T. Lee, and D. Zhang, "Human gait recognition by the fusion of motion and static spatio-temporal templates", Pattern Recognition, 40, 2563-2573 (2007).

[12] R. Munoz-Salinas, R. Medina-Carnicer, F. J. Madrid-Cuevas, and A. Carmona-Poyato, "Depth postures for gesture recognition", Pattern Recognition Letters, 29, 319-329 (2008).

[13] Microsoft Corporation, Kinect for Xbox 360-Xbox.com, Available: http: //www.xbox.com/en-GB/kinect/, Accessed November 10 2014.

[14] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), 711–720 (1997).

★★★