

# A Human Identity Recognition System Based on Kinect Skeletal Tracking and ANN Classifier

Xianhui Zeng<sup>1</sup>, Jinwen Deng<sup>1</sup>, Mujun Liu<sup>1</sup>, Yue Wang<sup>1</sup> and Xuejun Zhang<sup>1,2,\*</sup>

<sup>1</sup>School of Computer, Electronics and Information, Guangxi University, Nanning, Guangxi 530004, P. R. China

<sup>2</sup>Guangxi Key Laboratory of Multimedia Communications and Network Technology & Guangxi Colleges and Universities Key Laboratory of Multimedia Communications and Information Processing, Guangxi University, Nanning 530004, China

\*Corresponding author

**Abstract**—We proposed a novel method of people recognition based on the structure of bones, which is detected by Kinect sensor for skeletal tracking. By using 3D depth image achieved by Kinect, body's skeleton information is extracted as 20 joints with joint coordinates, from which body's skeletal features are calculated into different length ratio between each bone. Later these features are processed by a BP artificial neural network (ANN) classifier to differentiate several samples. The results show a high accuracy up to 98% test samples, which indicate that the body's skeletal features can provide the effective judgment in people recognition.

**Keywords**-human identity recognition; kinect; ANN classifier; skeletal identification

## I. THE BACKGROUND INFORMATION AND THE SIGNIFICANCE

With the development of contemporary society, the requirement of effective authentication and human identity recognition becomes higher and higher. Traditionally, passwords and ID cards were used to restrict access to secure systems but these methods can easily be breached and are unreliable. Biometric recognition[1][9] is developing in rise with the favorable security, validity and irreplaceability. In these systems, physiological traits, such as fingerprint[2], face recognition[4], iris recognition[7], retina scan, and behavioral traits, such as typing rhythm, key stroke, pattern and handwriting, have been intensively studied and developed to solve security problem inherent in traditional personal

recognition methods. In this paper, by using Kinect skeletal tracking technology[6] developed by Microsoft, we proposed a novel method for human identity recognition based on the structure of bones which has some advantages: more acceptable by the users since it is more convenient without any specific requirements; better anti-falsification due to the unique of human's bone organization, it is hard to be falsified or stolen; easier to carry on and with less risk of loss, because the bone organization is human's intrinsic properties. It can be applied to many other industry and fields, such as Social Security, AI Robots and Health Care, Entertainment and Gaming. Our system firstly utilizes Kinect skeletal tracking technique[11] to capture the human body movement, after evaluating and recording the properties of skeletal data from different in-stances, such as length of the arm, length of the leg, etc., we compile the skeletal data and select the distinguished properties and train the data using artificial neural network (ANN) classifier[5]. Finally the trained ANN classifier is used to detect skeletal features of individual data to achieve the recognition of the individual.

## II. METHODS

### A. Flowchart of Our Skeletal Tracking System

From the system structure as shown in Figure I, the central part is the bone identification hardware based on the Kinect skeletal tracking sensor[3] and the human recognition software based on ANN-BP classifier.

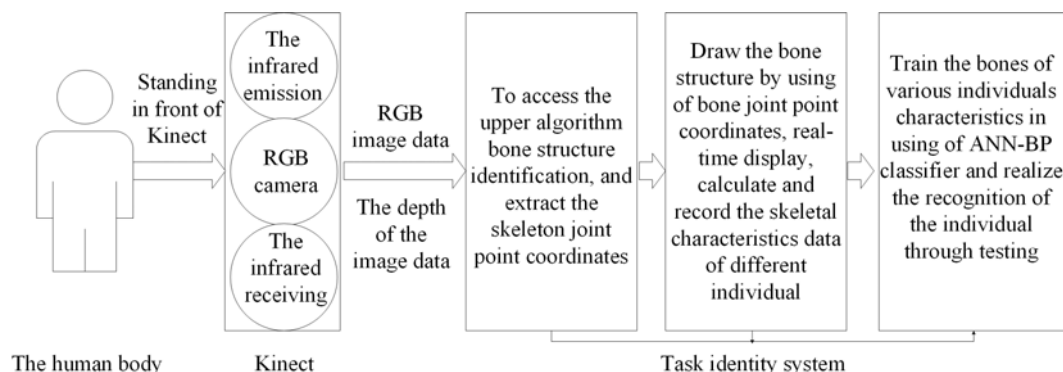


FIGURE I. THE STRUCTURE OF OUR SYSTEM

First of all, human skeletal structure will be identified using the Kinect skeletal tracking techniques. Second, skeletal properties will be calculated and its structure will be displayed based on the extracted joint point position data. Last, skeletal properties data need to be input into the ANN-BP classifier, which enable the system to conduct training on the data. The accuracy of human identity recognition will be test by detecting different instances.

### B. Bone Identification

A skeleton is displayed by 20 joint points in Kinect[10] as shown in Figure II. When people walk into Kinect's vision field, Kinect can find the location of the object's' 20 joint points (standing position) marked by (x, y, z) coordinates.

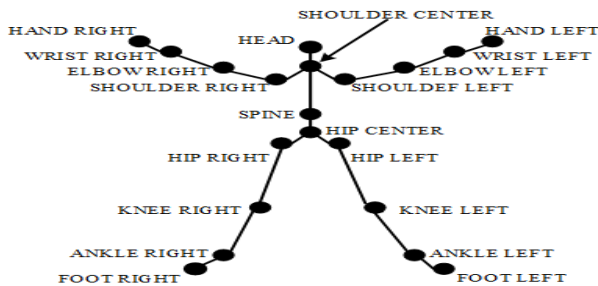


FIGURE II. THE HUMAN SKELETON DEFINED BY KINECT

### C. Calculation of Skeletal Characteristics

When extracting the human bone's 20 joint point coordinates, we can calculate a variety of characteristics of the skeleton, such as the length of arm and leg and so on. The skeletal characteristics data are stored in the TXT documents for further processing by ANN. Twenty joint points are defined as follows:

```
typedef enum _NUI_SKELETON_POSITION_INDEX
{
    1. NUI_SKELETON_POSITION_HIP_CENTER= 0
    2. NUI_SKELETON_POSITION_SPINE,
    .....
    20. NUI_SKELETON_POSITION_FOOT_RIGHT,
} NUI_SKELETON_POSITION_INDEX;
```

### D. Artificial Neural Network and Back Propagation (BP) Algorithm

Artificial neural network is an information solution system to simulate the brain structure and function. The processing units in the network are classified into three categories: input cell, export cell and hidden layer.

Back Propagation(BP) Network, a kind of multilayer feedforward network practiced by back propagation algorithm. The training algorithm of BP neural network repeats "Positive going calculation output - error back propagation" process until the error is reduced to an acceptable range, then the

training process of BP neural network comes to the end. Classical BP neural network is shown in Figure III .

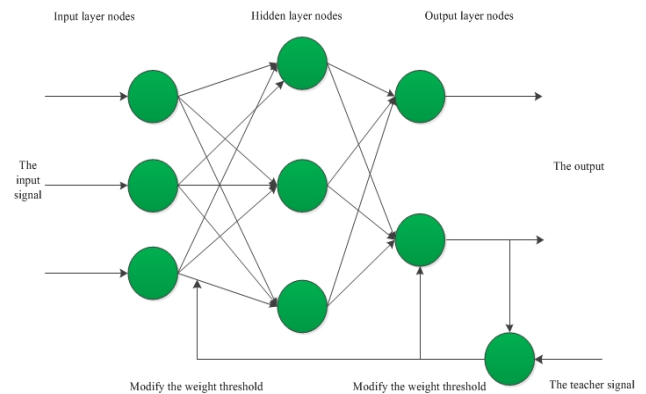


FIGURE III. CLASSICAL BP NEURAL NETWORK

BP classifier includes two steps: training and testing. In training stage, skeletal features calculated from individual data are selected as input unites, with a total of 12 samples and six input layer nodes. By setting the desired output, sample A's input value corresponds to the value of desired output to {0,0,0,1}. Sample B corresponds to {0,0,1,0}. C corresponds to {0,1,0,0}, and D corresponds to {1,0,0,0}. Neural network training start with four output layer node, eight hidden layer nodes, 0.01 as the maximum allowable error and the desired out-put that was set. When the system total error is less than 0.01, the training comes to the end. The weights and threshold of the system will be saved in TXT file for detection step. Taking a new image as input of ANN, such as a sample of character A is {0.595468, 0.489753, 0.789444}, our system read its weights and thresholds of the file which is saved after the training. Start test, if the result is {0,0,0,1} that character A, then the experiment is successful. In reality, the actual testing input and the training input do not necessarily coincide with those in another, and the output is not a standard binary value 0, 1, but a decimal number approached 0 or 1. As such we can do the detection by setting threshold value to achieve human identity recognition, when the output nodes greater than 0.7, then determine the output 1.

ANN - BP classifier[8] algorithm is introduced so that it can input the same individual who is in different position with different input data of skeletal characteristics to the ANN. System can correctly identify the individual even if the individual standing in a different position.

## III. RESULTS AND DISCUSSIONS

### A. User Interface of Human Identity Recognition System

Figure IV shows the final function and interface of the human identity recognition system.



FIGURE IV. THE SYSTEM INTERFACE DIAGRAM

Systematic realizing program are shown as follow: Press “Show the Skeleton” button to show both the human photo and the skeleton framework when one is around in the screen, as shown in Figure V.

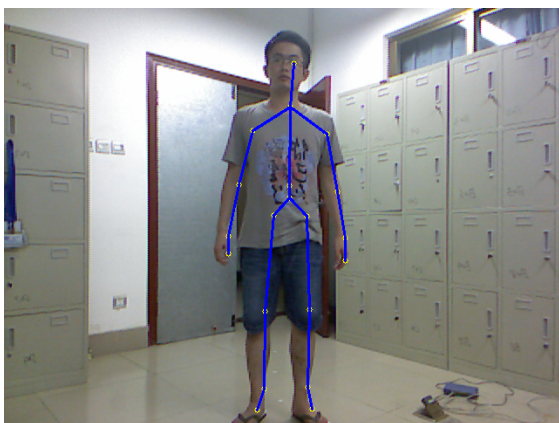


FIGURE V. SKELETON FRAMEWORK DISPLAYED IN HUMAN PHOTO

After the skeletal frameworks is established, the systematic procedure can calculate and draw the human skeleton framework, output and record the body's skeletal features in real time. The characteristic data of the bones is copied into the txt document when the user click “Record Data” after entering the name in the text box labeled User Name. Being able to use ANN artificial neural network classifier to train the data after processing the sampled body's skeletal features with the click of “Training” button. When experimenter walk into Kinect's visual field, click “People Recognition”, the experimenter himself name will be shown and pronounced in the “name” box if the recognition system is successful.

### B. Analysis of Experimental Result

The experiment results are contrasted with the feasibility of human identity recognition based on the structure of human skeleton, which is showed in the system demonstrations. Successful recognition from 4 body's skeletal features was obtained in accuracy of 100%. In order to test the ability of system and prove the reliability of the experimental results, we increase the number of samples and bone characteristics to check human identity recognition rates in various situations. When the number of features from body's skeletal keeps to 3

while increasing the sample number, the accuracy rates are changed as the following statistics shown in Figure VI.

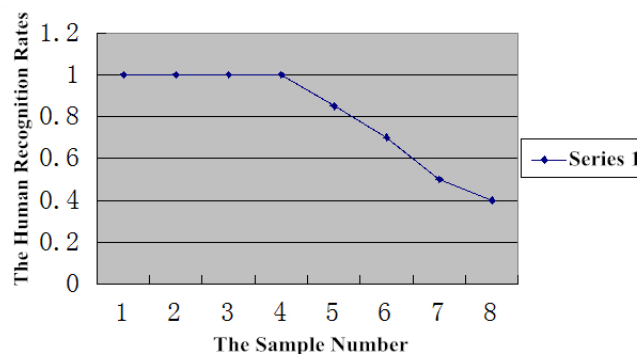


FIGURE VI. THE CHANGES OF THE HUMAN RECOGNITION RATES WITH A INCREASING SAMPLE NUMBER

If the sample number remains to constant 8 while changing the number of body's skeletal features, the statistics of accuracy changing curve is shown in Figure VII.

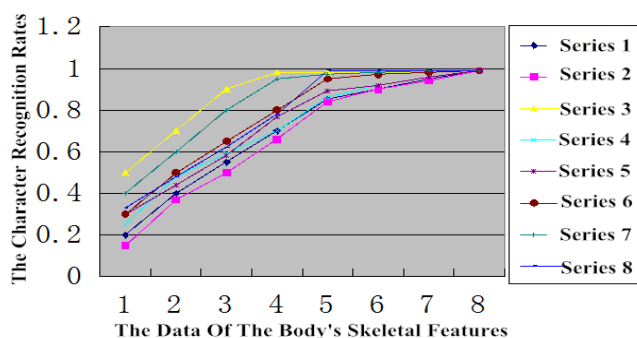


FIGURE VII. THE CHANGES OF THE HUMAN RECOGNITION RATES WITH THE DATA

The experiment results indicate that when the sample group increased to eight person, the human identity recognition rate is greatly reduced. Only 40% people were successful identified if each individual still use three bones features data for training and testing. The reason mainly due to the fact that an increase in the number of individual would made individual skeletal characteristics between simples similar and close in vector space, leading ANN to reduce the classification ability of among the individual and group. More features are need to add in training dataset to improve classification accuracy. If the bone features of each individual data increase to 8, the success rate of human identity recognition will be enhanced up to 98%.

When calculating the features of bone we can find that even the same individual, in different distance to Kinect, a feature will not be the same, and sometimes even with big differences. This is one of the reasons the ANN - BP classifier algorithm is introduced so that it can input the same individual who is in different position with different input data of skeletal characteristics to the ANN. After training the ANN classifier, our human identity recognition system can correctly identify the individual even if the individual standing in a different position.

This recognition system introduces BP -ANN algorithms based on the difference of bone organization in human. For time limit, the inadequacies still exist which could be improved and the further research as follows:

First, fewer body's skeletal features data collection can not meet the requirement which is strict to the performance of the ANN classifier, resulting in deviation on statistical results. In order to improve the system performance, more data have to be collected for building a robust model.

Second, the limitations of human identity recognition of this system still exist which need further study: data deviation of the system will be occurred if the experimenter is not standing toward the Kinect properly when the data acquisition is taken. This sideways data could be corrected by some deformation algorithm.

#### IV. CONCLUSIONS

Our studies involved two aspects: (1) Identification of the structure of human skeletons and calculation of features of skeletal characteristics by using Kinect skeletal tracking technology. (2) Introduction of the BP -ANN classifier to construct the human identity recognition applications. It's feasible to identify people through the body's bone structure. With the recognition group's individual number increasing, the skeleton characteristic is also needed to increase. The further work includes selection of efficient skeleton features with optimized number of inputs to ANN; application of more training samples to built a real system; and adding face features to achieve higher accuracy. Our system will be of a potential real-world application to human identity recognition by combination of face recognition with bone identification methods.

#### ACKNOWLEDGMENT

This work was supported in part by two research support from the National Natural Science Foundation of China (Nos.61262027&45627390).

#### REFERENCES

- [1] Tatepamulwar, C.B., Pawar, V.P., Fadewarr, H.S. "Biometric Recognition", NCI2TM: 2014. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.
- [2] Galar, M., Derrac, J., Peralta, D., et al., 2015a. A survey of fingerprint classification part I: taxonomies on feature extraction methods and learning models. *Knowl.-Based Syst.*, 81:76-97. K. Elissa, "Title of paper if known," unpublished.
- [3] GU, R., HE, Y., JIN, L. Design and implementation of Kinect skeletal tracking technology-based intelligent teaching control system [J]. *Computer Application and Software*, 2013, 30( 6) : 276 - 280.
- [4] Li B Y L, Mian A, Liu W, et al. Using Kinect for face recognition under varying poses, expressions, illumination and disguise[C]// *Proc of IEEE Workshop on Applications of Computer Vision*. Washington DC: IEEE Computer Society, 2013: 186-192.
- [5] Nong, Zhang. Real-time identification of vehicle motion modes using neural networks. *Mechanical Systems and Signal Processing*, 2015(05).
- [6] Open Kinect Organization. Imaging information for Kinect [EB/OL]. [2011- 07 - 10 ]. [http://open.kinect.org/wiki/Imaging\\_Information](http://open.kinect.org/wiki/Imaging_Information).
- [7] R. T. Al-Zubi, D. I. Abu-Al-Nadi. Automated personal identification system based on human iris analysis [J]. *Pattern Analysis and Applications*. 2007 (2).
- [8] Seman N, Bakar Z A, Bakar NA. The optimization of Artificial Neural Networks connection weights using genetic algorithms for isolated spoken Malay parliamentary speeches. *2010 International Conference on Computer and Information Application (ICCIA)*. IEEE. 2010. 162-166.
- [9] Shweta Gaur, V.A. Shah, Manish Thakker, "Bio-metric Recognition Techniques", *International Journal of Advanced Electrical, Electronics and Instrumentation Engineering*, Vol. 1, issue 4, October 2012.
- [10] Tilak D. Evaluation of the Kinect TM sensor for 3- D kinematic measurement in the workplace [J]. *Applied Ergonomics*, 2012, 43( 4) : 1 - 5.
- [11] Liang, Wang, Tieniu, T. Hu zhong, N. Weiming, H. Silhouette, Analysis-based Gait Recognition for Human Identification [J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2003, 25 ( 12 ) : 1505 - 1518.