Physical rehabilitation assistant system based on Kinect

Ching-Tang Hsieh[a], Ruei-Chi Chung[a], Yeh-Kuang Wu[b], and Liung-Chun Chang[a]

Department of Electrical Engineering, Tamkang University, Tanshui, 25137, Taiwan [a]

Institute for Information Industry, Taipei, Taiwan [b]

E-mail: hsieh@ee.tku.edu.tw

Abstract —**In this paper, we set up a physical rehabilitation assistant system based on skeleton detection with Kinect produced by Microsoft. First of all, the users do not have to install the detectors on the exercise equipment. Secondly, they pay a little extra expensive to buy the rehabilitation equipments with Kinect using skeleton detection technique. In this study, we build a normalized three-dimensional Cartesian coordinates location of correct postures under OpenNI system. We find out 15 human skeleton joints with three dimensional coordinates and calculate the feature values, than we use support vector machine (SVM) as classifier to define the accuracy of posture. Finally, the system can judge the correct degree of user's postures.[1]**

Keyword: Kinect, rehabilitation, SVM.

I. INTRODUCTION

The application of the Image processing makes the development of the technology more life-oriented, and there are the wide varieties of the software to capture the images. Kinect is the popular device in these days, and it uses the infrared ray to build the Image depth information. Using the image in medical assistance is also a popular topic now. For example, Chen [1] mentioned that using face orientation in face images to help the vestibular rehabilitation. Yeh [2] provides three systems to operate the communication aid contain movements of their arms, open and close their mouths, and eye blinks to build the Human-Computer Interface. Those two theses are the good examples of the image technique apply to the medical.

Traditionally, the rehabilitation related equipments for physical are very expensive, and users have to put on the considerable numbers of sensor also they make quite uncomfortable for the users. However, the price of the Kinect is acceptable, and the skeleton detection technique is accurate. Most important of all, users do not have to put on the sensors on them. It is quite acceptable and comfortable for the users, and it also can promote this rehabilitation system to the home life. Therefore, we purpose to recognize human posture and correct user's action for rehabilitation process. So, how to seamlessly track and successfully is very important. Kinect is a good tool to tell the human acting, the character of skeleton tracking system is not affecting by illumination, and use depth image to individual object from environment also provide the information what we needed quickly for our work.

Huang [3] using the horizontal and vertical projection, star skeleton, neural network and similar feature process techniques to recognize the five human postures, its advantage is apply to different indoor environment and different distance between human and Kinect, also can stable and real-time to recognize different human physiques. Our method can recognize more detail human posture such as wave hand, raise foot. Wu [4] establishes a motion model of view-Invariant and training samples by K-means and calculating the three dimensional distance between fifteen points to matching samples, its advantage is high average recognize similarity and real-time, but only can recognize one motion. Wang [5] apply three-dimensional coordinates to classify under Support Vector Machine (SVM) [6], and its experiment result shows good performance, its average recognition rate is 95.6%, but they need to use very expensive equipment to achieve goal. We combine SVM with Kinect skeleton tracking system under OpenNI [7]. Therefore, not only could make the system equipment low cost but also improve the disadvantage of user restriction.

In this paper, we recognize the human action to build an assistance system for medical rehabilitation. During the rehabilitation process, the patients need to do the same motion repeated in a long period. After the patients are discharged from hospital, they cannot make sure to do the right medical rehabilitation postures without the help of the professional medical staff. Finally, our goal is to use simple equipment to help those patients who needs physical rehabilitation.

II. METHOD

We set up a system based on Kinect skeleton tracking under OpenNI. First at all, the user go into the region of Kincet and start to skeleton tracking, after generate real world three-dimensional Cartesian coordinates we normalize it to get a object coordinate system which use the torso center as the origin. So we can calculate the feature value of 45 human joint coordinates for SVM to classify. After SVM training we set up correct posture sample database for rehabilitation. According to the posture sample which trained by SVM before, we can recognize the user posture immediately and use human skeleton joint coordinates to correct posture. The flow chart of our system shown in Fig. 1.

[1] This work was supported in part by the NSC under Grant No. XX-0123-456-789 (sponsor acknowledgment goes here).

Paper titles should be written in uppercase and lowercase letters, not all uppercase. Full names of authors are preferred in the author field.

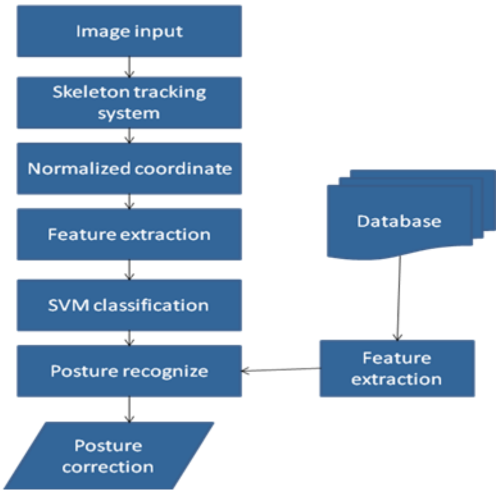

Fig. 1 The flow chart of the proposed system

A. *Skeleton tracking system*

The OpenNI defined 24 joints, but actually we can only use 15 joints to analyze skeleton by middleware. Fig.2 shows that we capture the information of skeleton by using User Generator, Depth Generator and Skeleton Capability, etc.

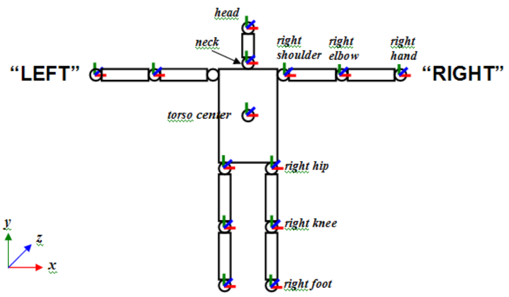

Fig. 2 The OpenNI defined the human skeleton and 15 joints from Hersy's Space

According to the definition of OpenNI, the user must stand in the view of Kinect and do PSI showed in Fig.3. Then we can not only locate the position of human body but also calculate the 3D coordinates (position) and the probable direction (orientation) of the 15 joints.

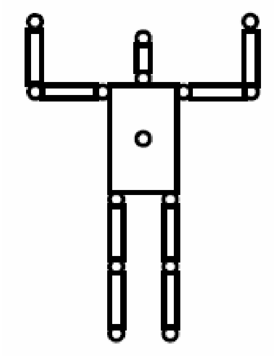

Fig. 3 PSI

B. *Normalization of coordinates*

According the define of OpenNI's real world coordinate. In front of the Kinect is the positive Z axis, above the Kinect is the positive Y axis and the right of the Kinect is the positive X axis. However, people can stand at any position in the real application. For classification, it will be a little trouble to set the value of feature. Therefore, we need to normalize the coordinates again to express the skeleton system by homogenous coordinate system [4] in the training process. So the system will calculate the translation of the geometry conversion in 3D Cartesian coordinates and define torso as origin. In computer graphics, it always applies the homogeneous coordinate system to simplify the operation of the matrix. It's convenient to record the translation, resize and rotation by the matrix. The homogeneous coordinate system use $d+1$ dimensional coordinates to represent the point on d-dimensional space. For example, three-dimensional coordinates point (x, y, z) is become (wx, wy, wz, w) in the homogeneous coordinate system. The forth coordinate value w is called homogenous coordinate and cannot equal to zero. Also w is expressed the distance parameter. The translation is to move a graph in the certain distance and direction. Also v is the position before translation; v' is the position after translation; dx is the translational component of x-axis; dy is the translational component of y-axis; dz is the translational component of z-axis. Furthermore, they can be represented by translation matrix. The 3D Cartesian coordinates is expressed by the homogeneous coordinate system. The position after translation is the dot product of the translation matrix and the original homogeneous coordinate matrix in 3D Cartesian coordinates. If we move the original 3D Cartesian coordinate position $(x_{torso}, y_{torso}, z_{torso})$ of the torso center as origin $(0,0,0)$ of the Cartesian coordinates, it can generate an object coordinate system as the origin of the torso center. The original 3D coordinate's position (x_i, y_i, z_i) of the other 14 joints take the original 3D Cartesian coordinate's position $(x_{torso}, y_{torso}, z_{torso})$ of the torso center as the reference point. They translate form the 3D Cartesian coordinate's position to the new coordinate (x_i', y_i', z_i'), this paper use the translation of geometry conversion in 3D Cartesian coordinates as Fig.4

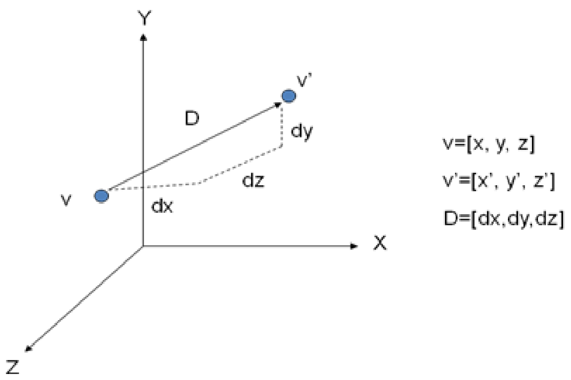

Fig. 4 Translation of geometry conversion

We perform the three-dimensional coordinates' translation of geometry conversion as:

$$
\begin{aligned}
v' &= v + d \\
&= T(dx, dy, dz) \bullet [x, y, z, 1]^T \\
&= \begin{bmatrix} 1 & 0 & 0 & dx \\ 0 & 1 & 0 & dy \\ 0 & 0 & 0 & dz \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \\
&= [x + dx, y + dy, z + dz, 1]^T \qquad (1)
\end{aligned}
$$

Fig. 5 shows the coordinate before normalizing with the torso center and the left shoulder. And Fig. 6 shows the after normalizing.

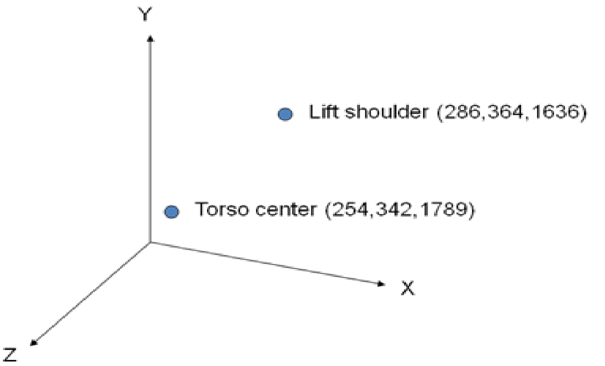

Fig. 5 The illustration before normalizing

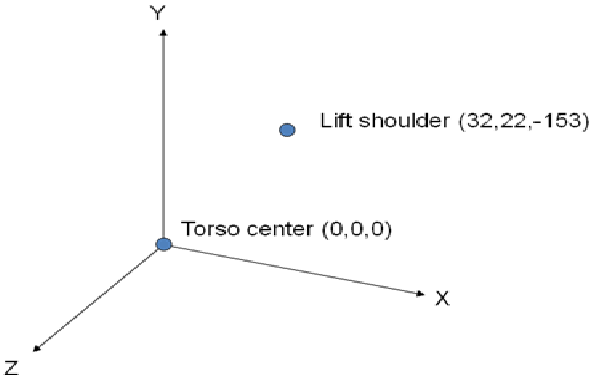

Fig. 6 The illustration after normalizing

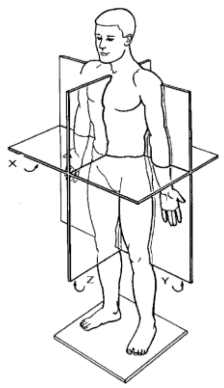

Fig. 7 The object (body) coordinates system as the origin of torso from Jacob & Francone 1970

C. Feature extraction

Based on the above method, we can get a normalized coordinates. According to normalized coordinate, each joint to torso center defined that as feature vector, we consider the 45 3D coordinate values (x', y', z') as our input data (feature points).

D. SVM classification

In the training procedure, the user must do the rehabilitation posture we want before training. Then record the training data points in every frame for training and building the samples. We take the 45 coordinate values as the feature points to do the SVM training in every frame for determining the posture. Every posture sample we use 100 frames for training. The parameter used in SVM are c=1, gamma=1, kernel=poly, and degree=1. The choice of SVM parameters is according to experimental experience.

E. Posture recognize

The SVM will normalize all the data to [-1, 1] for each sample dataset. It means we can find the maximum-margin hyperplane that divides the datasets, than we can find the nearest training posture sample for each frame in the time.

F. Posture correction

Based on the joints obtained in the same time, we can calculate the Euclidean distance with the sample and display it to help the user knowing the difference. And set the threshold make sure that the action is in the certain area.

III. EXPERIMENTAL RESULTS

The input RGB and skeleton images size are 640x480 pixels, the hardware used are Microsoft Kinect, CPU Intel® Core(TM)2 Duo CPU 3.00GHz,and RAM 3GB. And software used are Microsoft Visual Studio 2010, Microsoft Kinect for OpenNI, openCV2.3, and LIBSVM. The experimental results as follow: Fig.8 is hands held flat sample and Fig.9 is raised right hand sample.

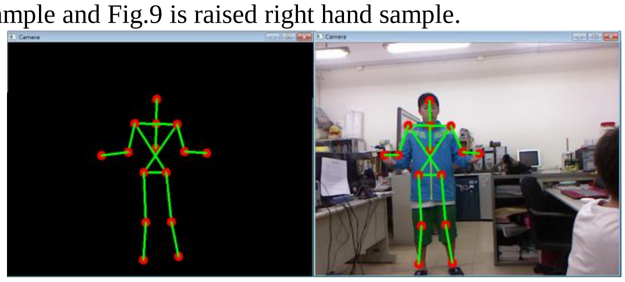

Fig. 8 Hands held flat sample

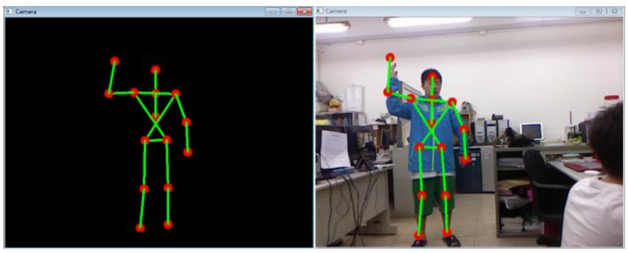

Fig. 9 Raise right hand sample

The left window in Fig. 8 and Fig.9 are the dataset we trained before, and right window are Kinect skeleton images. Our work process speed is 20FPS.

IV. CONCLUSION

At first, we use the SVM to training the samples of the correct postures and use it to build a correct sample database. After training, our system for classify we take the feature value of every frame and use the SVM to match with the samples, and it can tell the accuracy of those postures. According this system to the rehabilitation assistances, the user can make the correct postures which are required by rehabilitation program through this simple convenient equipment. The user do not need accompany with the medical personnel to do the rehabilitation program, it can reduce some waste of the medical human resource.

REFERENCES

[1] Bo-Ru Chen, "Detection of Face Orientation and its Applications in Vestibular Rehabilitation and Human-Computer Interface," Department of Computer Science Information Engineering, National Central University, 2011.
[2] ChinYeh Yeh, "An Implementation of an Eye-blink-based Communication Aid for people with Severe Disabilities,"pp351-356, ICALIP, 2008.
[3] Shih-fu Huang, "Human's posture recognition by using Kinect," Department of Electrical Engineering, National Central University, 2011.
[4] Kuei-Kang Wu, "Using Human Skeleton to Recognizing Human Motion by Kinect's Camera," Computer Science and Information Engineering, National Taipei University of Technology, 2011.
[5] Jung-Ying Wang, "Recognition of Human Actions Using Motion Capture Data and Support Vector Machine," pp234-238, World Congress on Software Engineering, 2009.
[6] Chih-Chung Chang and Chih-Jen Lin, "LIBSVM -- A Library for Support Vector Machines", *Network: http://www.csie.ntu.edu.tw/*.
[7] OpenNI , *Netowork : http://openni.org/*
[8] Heresy's Space, "OpenNI/Kinect", *http://kheresy.wordpress.com/index_of_openni_and_kinect/*