International Journal of Pure and Applied Mathematics
Volume 114 No. 7 2017, 305-313
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue
ijpam.eu

Human Posture Recognition Based on Tracking Data of Kinect Skeleton

P.Tirumala Rao[1], V.Sindhu Bhargavi[2] , Dr.Rudra Pratap Das[3], A.Sampath Dakshina Murthy[4], I. Krishna Rao[5] and Dr.S.Koteswara Rao [6]

[1,2,3,4,5]Department of ECE. Vignan's institute of information technology Visakhapatnam , India

[6]Department of ECE, K L University Vaddeswaram, India

[1]pechettytirumalarao@gmail.com,
[2]sindu.bhar@gmail.com,
[3]drrudrapratapdas@yahoo.com,
[4]sampathdakshinamurthy@gmail.com,
[5]krishnaraomvpippili@gmail.com,
[6]rao.sk9@gmail.com

April 14-15, 2017

Abstract

Recognition of human posture is very important in the areas of computer vision because of prospective application in human computer recognition, surveillance and personal health. In this paper, a method has been proposed for posture recognition by kinect of Microsoft. Some salient joints of the human body are estimated by this and experiments are carried out on a predetermined dataset. Results show that supervised tracking is superior to the unsupervised counterpart. An estimate of precision/recall is of the order of 70.6%..

KeyWords: Human posture, Tracking, KINECT, Computer

1 Introduction

Our human actions can be depicted by evolution of different body posture based on spat ion (space related) - temporal (time related) evolutions. Salient parts such as arms or legs are studied to make meaningful configuration of joint positions. ECG signals are utilized to detect emotions coupled with posture recognition. This method is relatively simple and less accurate [1]. In this paper, it is suggested that KINECT should be attached to a small computer which can evaluate the levels of obtrusiveness.

In the proposed technique KINECT is used as a motion sensor due to good user tracking. The output parameter is evolved from the inputs of general actions and event recognition system (ERS). The ERS analyzes various information of sensing mechanism.

Basing on a dense matrix of spatial-temporal parameters, a relatively efficient system has been proposed. Use of time invariant features is required for recognizing action at various speeds. It also leads to error when speed is used to discover various movements.

The disadvantage is that a single object needs to be active before a background which has to be steady. This constraint is overcome by using advanced devices which can manage visual and depth parameters [2].

The problem of active recognition by using intrusive sensors like the wearable variety has been previously discussed. The unobtrusive sensors like video sensors are preferred to the wearable sensors. In the context of the above discussion, it is pertinent to take up kinect as the primary sensor for collection of user's movements.

The problem of action recognized using intrusive sensors like the wearable variety has been discussed. The unobtrusive sensors like video sensors are preferred to the wearable sensors. In the context of the above discussions, it is pertinent to take up kinect as the primary sensor for collectors of users movements.

2 Description of Equipment

There are essentially two cameras in Microsoft Kinect. One is the RGB camera and the other is the IR camera. The resolution is

640x480 and IR projector is suitable for shifting infrared rays for the environment.

The principle of operation is that the degree of distortion of each ray projected from the field of action is eminently suitable for estimating map of depth in which each pixel can represent the distance of a specific point in 3D space with respect to the sensor [3]. In a specific case human bodies have been modeled a set of movable joints and the actions which are defined as the interaction among the joint subsets.

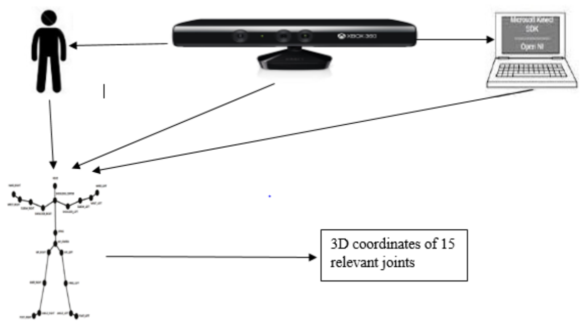

Fig: 1 images captured by the Kinect are processed to detect a set of joints, which are subsequently normalized with respect to scale and position.

Figure 1 shows a system which can automatically infer the actions performed as per known set of postures there are three results of the system 1) Feature detection 2) Recognition System 3) Distance calculation.

3 Modus Operand

The movements of the body are to be tracked first .The features of interest must be initially identified. In the human body there are many interactions such that no work can be there in isolation. The musco-skeletal system supports the human body enabling the movements as per the stimulus provided [4] [5]. The significant parts of the skeleton are head, neck, arms, hands, legs & feet. All these need to be focused. Such parts can be modeled as segments connected to each other through nodes or joints. These joints in turn restrict the movements of the parts in the three dimensional space.

The Kinect device can be tested to be an unobtrusive sensor for real-time detection in 3-D for almost all the body joints.

An evolutionary algorithm is used to select feasible joints without noise and other peculiarities of the human body. The algorithm selects the optimal subset of a specific training data. The joints selection becomes centered on the data. So any variations in activity cause the choice of different types of joint subsets.

Some fundamental tests are necessary to evaluate the significance of the act of joints by adjusting the IR sensor. When two segments overlap like the fingers touching body parts, the detection is not perfect due to the presence of objects between the user and action. For this purpose the recognition rates needs to be evaluated.

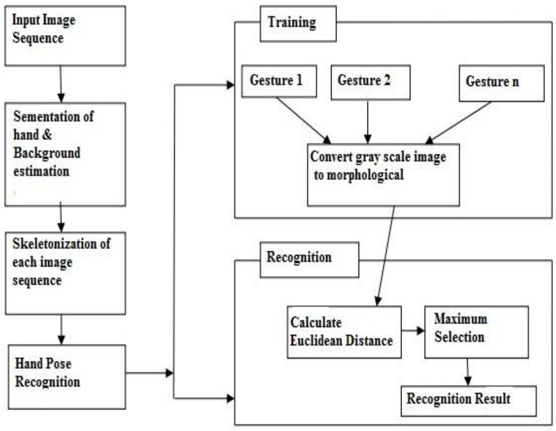

Fig: 2 Block diagram for recognition system

Figure 2 shows the recognition process. The training phase involves
a) Detection of features of interest.
b) Organization of features space
c) Refining the detection by Euclidean distance
For recognition it is required to
a) Detect the main points of the feature
b) Classify the activity
c) Test the posture sequence against all stored images
d) Select for maximum probability
e) Compare against threshold.

4 Distance Calculation

There are three distance functions: -
a) Chess board distance
b) City block distance
c) Euclidean distance
Basically distance calculation is based on transformation which converts a digital binary image to a grayscale image in which pixel value is the minimum distance from the back ground to that pixel set by a predetermined function.

The signed Euclidean transformation is quite popular in application such as smoothing and finding convex skulls.

By morphological erosion, the minimum value from the combination of an image and the predefined weighted element are suitably determined [6] [7].The decomposition of morphology is useful for parallel computing. The boundaries of the objects are encoded as chains.

Before the Euclidean analysis is performed, input source data (feature class) is converted to a raster. Normally the resolution is smaller than height/width (smaller of these values). It is divided by 250. The resolution can be set with output cell parameter.

The Euclidean distance is measured from the centre of source cells to the centre batch. For each cell the distance is calculated to each source cell and true distance is computed to each cell in the distance functions. (Hypotenuse is measured from x-axis and y-axis).

The shortest distance [8][9] to source is determined by comparing with a specified maximum distance. Finally the values are assigned to the cell action on the output raster. The raster of Euclidean distance gives us data regarding how close each cell is to the nearest source cell and the raster defines the source and corresponding value [10].

Taking P (d_1,u_1) and Q (d_2,u_2) the travelly distance

$$=\sqrt{(d_1 - u_1)^2 + (d_2 - u_2)^2}$$

5 Results

User defined dataset

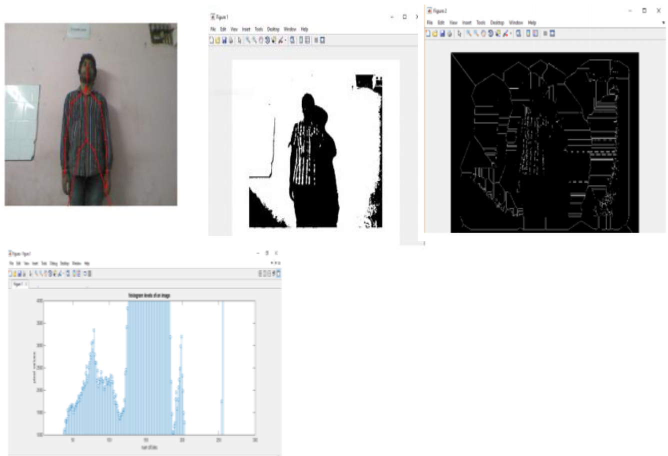

Fig: 3 a) Standing position b) gray level image c) morphological image d) histogram level of an image

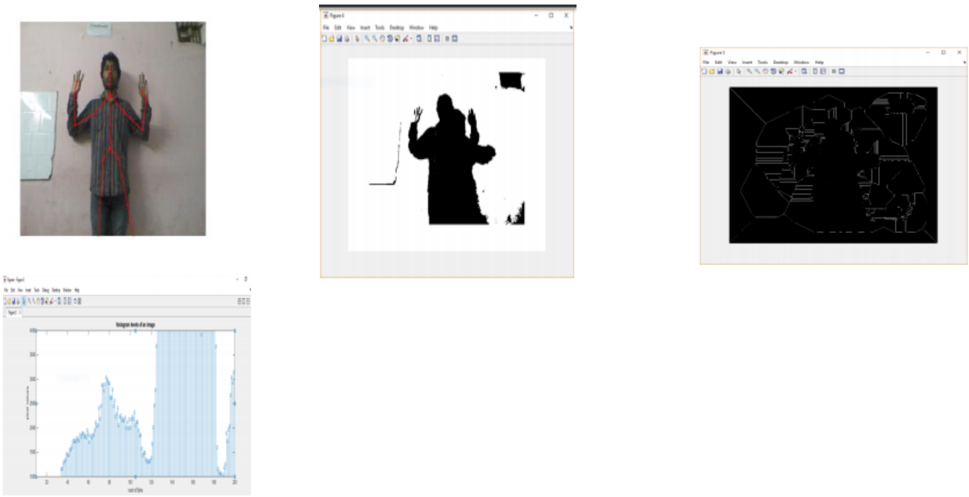

Fig: 4 a) Two hands up position b) gray level of an image c) morphology of image d) histogram level of an image

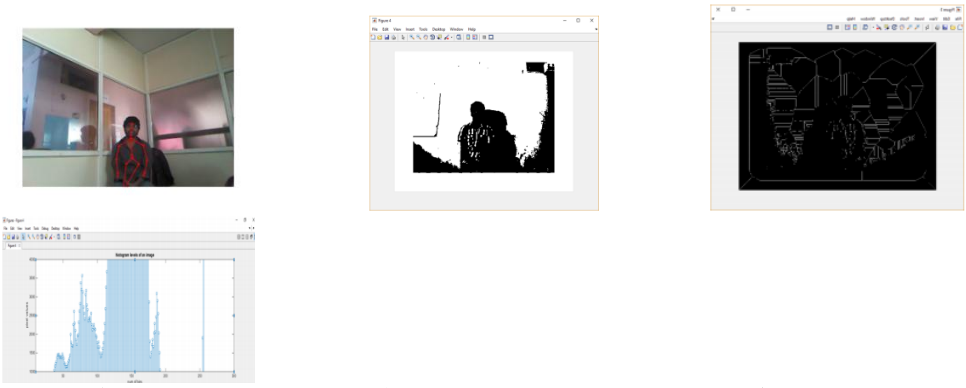

Fig:5 a) Sitting position b) gray level of an image c) morphology of image d)histogram level of an image

Fig:6 a) Stretch position b) gray level of an image c) morphological image d) histogram level of an image

Comparison one posture with the above postures we get the resultant posture and the position.

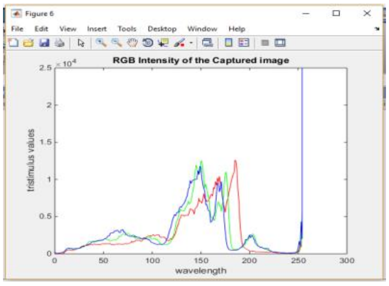

Fig:7 It display RGB intensity of an image   Fig8: Posture recognition: standing position

Table 1.1: Accuracy values for comparison images

Dataset Posture	Accuracy values
Standing	0.0712
Stretch	0.1210
Two hands up	0.1520
Sitting	0.0738

The simulation has been done on MATLAB using PC with dual core CPU .For all the dataset the position was done at nearly 2 meters from the object .Comparison was made with datasheet using threshold value and distance with respect to image.

6 Conclusion

It has been found that this technique is capable of giving a general model of the postures .This model s effectively independent of the user, speed of action .Moreover; these can be scaled upwards and expanded to future action. Similar activities done different time frames and durations can also be recognized.

The short timing the system is linked to the inability of having a stable video pattern .Post estimation process needs to be further developed to circumvent this problem .Overall performance of nearly 70.6% for precision & recall justify the application.

References

[1] J. Yamato, J. Ohya, and K. Ishii, "Recognizing human action in time-sequential images using hidden Markov model," in Proc. IEEEComput. Soc. Conf. Comput. Vision Pattern Recog. Proc., 1992, pp.379-385

[2] M. P. V. Kellokumpu and J. Heikkila, "Human activity recognition using sequences of postures," in Proc. IAPR Conf. Mach. Vision Appl., 2005, pp. 570–573

[3] B. Scholkopf and A. J. Smola, Learning with Kernels: Support VectorMachines, Regularization, Optimization, and Beyond. Cambridge, MA,USA: MIT Press, 2001.

[4] A. Oikonomopoulos, I. Patras, and M. Pantic, “Spatiotemporal salient points for visual recognition of human actions,” IEEE Trans. Syst., Man,Cybern. B, Cybern., vol. 36, no. 3, pp. 710–719, Jun. 2005.

[5] G. Willems, T. Tuytelaars, and L. Gool, “An efficient dense and scale-invariant spatio-temporal interest point detector,” in Proc.

[6] Wahabi, Saeid, Shahrzad Pouryayevali, and Dimitrios Hatzinakos. "Posture-invariant ECG recognition with posture detection." In Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on, pp. 1812-1816. IEEE, 2015.

[7] Rao, P. Tirumala, S. Koteswarao Rao, G. Manikanta, and S. Ravi Kumar. "Distinguishing normal and abnormal ECG signal." Indian Journal of Science and Technology 9, no. 10 (2016)

[8] Murthy, A. Sampath Dakshina, S. Koteswara Rao, G. Thiagarajan, and V. Suresh. "Noise Cancellation in Monte Carlo Simulation." Indian Journal of Science and Technology 9, no. 31 (2016).

[9] Murthy, A. Sampath Dakshina, S. Koteswara Rao, A. Naga Jyothi, and Rudra Pratap Das. "Analysis of effect of Ballistic coefficient in the formulations and performance of EKF with emphasis on air drag." Indian Journal of Science and Technology 8, no. 31 (2015).

[10] Zhang, Zhengyou. "Microsoft kinect sensor and its effect." IEEE multimedia 19, no. 2 (2012): 4-10.