Robust Hand Gesture Recognition by Using Depth Data and Skeletal Information with Kinect V2 Sensor

S.Chandrasekhar [#1], N.N.Mhala [*2]

[#] Ph.D. Scholar, Bapurao Deshmukh College of Engineering Seagram
Wardha, Maharashtra State-442102, India
[1] srikantamchandu@gmail.com

[*] Professor & HOD, Dept. of Electronics Engineering, Bapurao Deshmukh College of Engineering Seagram
Wardha, Maharashtra State-442102, India
[2] nitiket.nhale@gmail.com

Abstract—**Hand gesture based human-computer interaction is one of the most natural and intuitive ways to communicate between people and machines. Here we propose hard gesture recognition by using depth data and skeletal joint information based on gradient value. This procedure is divided into two ways they are hard shape features and hand arm movement. In this paper, we propose depth data and skeletal joint information based on gradient value instead of conventional 2D shape features and arm movement features based on angles between each joint.**

Human-Computer-Interaction, hand gesture recognition, Kinect sensor, depth sensor, HMM

I. INTRODUCTION

The hand gesture recognition is a powerful and general communication way it consists of two factors, they are shape of the hand and variation of arm in the way of movement. To give the instruction with good meaning to robot the gesture recognition is the best one [1].

The human hand is connected with number of parts and Joints. The hand movements are measured by using gloves and with different equipment. The hand movements are also measured by using cameras and also by using 3D cameras with good scanners and depth sensors like Kinect V2[2].But conventional methods are categorized into two major ways such as the best 3D- model and 2D conventional shape features .By using the 3D-model and glove-based devices very powerful ways to measure gesture recognition. But the Glove method is very cumbersome method it contains the wires, the gloves and their attached wires are still cumbersome for user to wear.

From the past few years, our researches have tried for gesture recognition methods using depth sensor. Which is in built with Kinect v2 .In our proposed method the gesture recognition are divided into two parts hand shape and hand arm movement. The human hand consists of many articulated object it containing many connected parts and joints .It is not possible from the method of 2D feature to calculate the ratio and counter based objects. From the method we have to go through the features, not only environmental changes but also we should keep on study on differences in hand sizes, the major distances between sensor and human body. In our procedure, the major coordinates are used to calculate the cantered of both hand and arm gestures.

In this paper, our aim is extraction of hand gesture and arm movements by using the advanced method of gradient value instead of conventional 2D procedure [3]. In our research or performance we evaluated performance of the proposed system is very well compare with other experiments such as 2Dfeatures.In this paper our aim is extraction of hand gesture and arm movements by using the advanced method of gradient value instead of conventional 2D procedure[3]. In our research or performance we evaluated performance of the proposed system is very well compare with other experiments such as 2Dfeatures.

Our approach involves an extraction of hand shape features based on gradient value instead of conventional 2D shape features, and arm movement features based on angles and angular velocity between each joints. Our approach involves an extraction of hand shape features based on gradient value instead of conventional 2D shape features, and arm movement features based on angles and angular velocity between each joints.
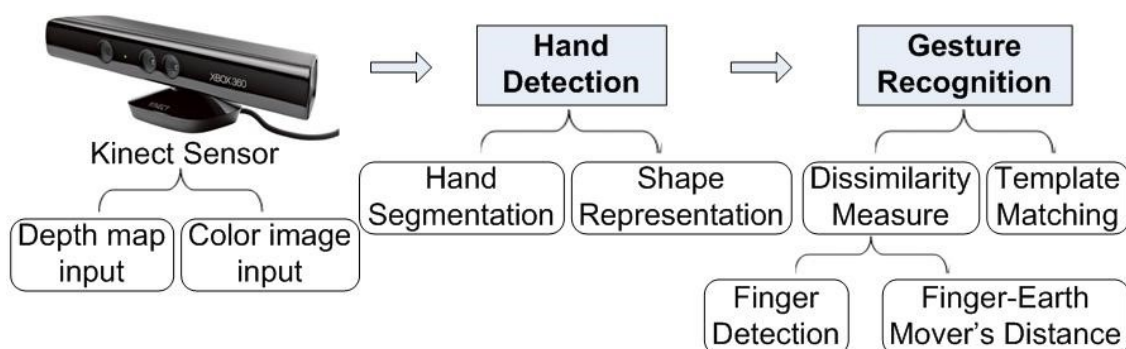
Figure 1: The framework of our robust hand gesture recognition system [4]

From the fig (1) shows the framework of complete hand gesture recognition system. Generally, it consists of the blocks of hand diction, gesture recognition. These are the main blocks of the system. From these blocks, we are again generalized as sub blocks of hand segmentation and shape representation. The depth map and colour image is feed from the input of the Kinect senor [4]. The gesture recognition also divided in to two sub parameters they are measuring of dissimilarity and template matching.

II. RELATED WORKS

In last few years' hand gesture, recognition methods have been widely used. The major methods are 3D and 2D mode-based approaches. The features of 3D method is glove which optical sensors were added to find gesture reflections. But the gloves and their attached wires are still cumbersome for user to wear [5]. In current method, the gesture recognitions are based on tracking the bare hand and recognizing hand gestures without gloves. This shown in the following fig (2).

Figure 2: Data glove [7]

Generally extract the features from hand shape the 3D sensor can be used .These sensors are very costly for general use and also take much time to investigate he other hand toget the features from hand is multiple camera system the other method is vowel model for calculating hand finger gestures by using multiple cameras It is difficult to understand hand gestures with a complex background. It is difficult to understand hand gestures with a complex background [6].

Now days, to recognize the gestures only the most important method is depth sensor such as kinect sensors. By using this sensor, we can get hand and fingertip tracking using depth data but it is difficult to classify a hand shape by only using 2D hand fractures because the human hand is very complex to identify the gestures. The hand and arm consisting of many joints it shown in fig (3).From the studies of global hand gestures each finger joints of human hand motion has normally 280 of freedom [7]. Here environmental conditions are also most important,it has less robustness compare to the arm movements. Now consider the individual difference in human body size, hand size and length of the arm and arm movements and skeletal joints in the hand.

III. PROPOSED GESTURE RECOGNITION METHOD

Here we are using best procedure of HMM to recognize the gestures by using 3D coordinates of eight joints out of 20 joints and depth data from the Kinect sensors. By using the HMM we extract hand shape features and arm movement features. Finally, the gestures are classified by kusing HMM [8].

A. Input device

Here, we use Kinect sensor as input depth sensor with resolution of colour image is 640*480 pix, it embedded with IR projector, depth image and colour image for RGBdata.the resolution of depth image is 320-*240 pix.The major important thing is that the depth data and RGB image of the object should be obtained at the same time.
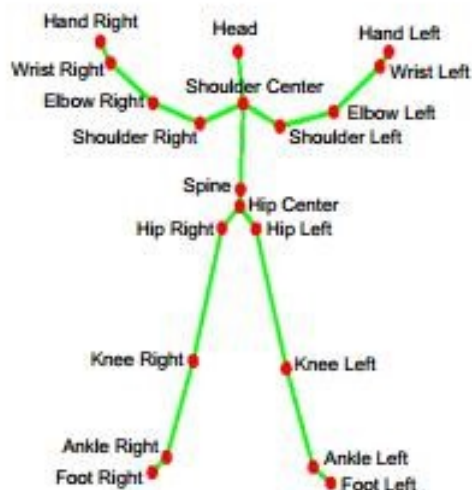

Figure 3: Kinect joints [5]

From fig (3) we use eight joints which are Hand Left, Hand Right, Wrist Left, Wrist Right, Shoulder Left, Shoulder Right and knee left and knee right obtained from skeleton model. In addition, the user image without background pixels is used for gesture recognition [9].

We use Kinect software development kit; it is a software development kit fig (4) shows RGB image, visualized depth data, depth image and skeleton model of 76%.
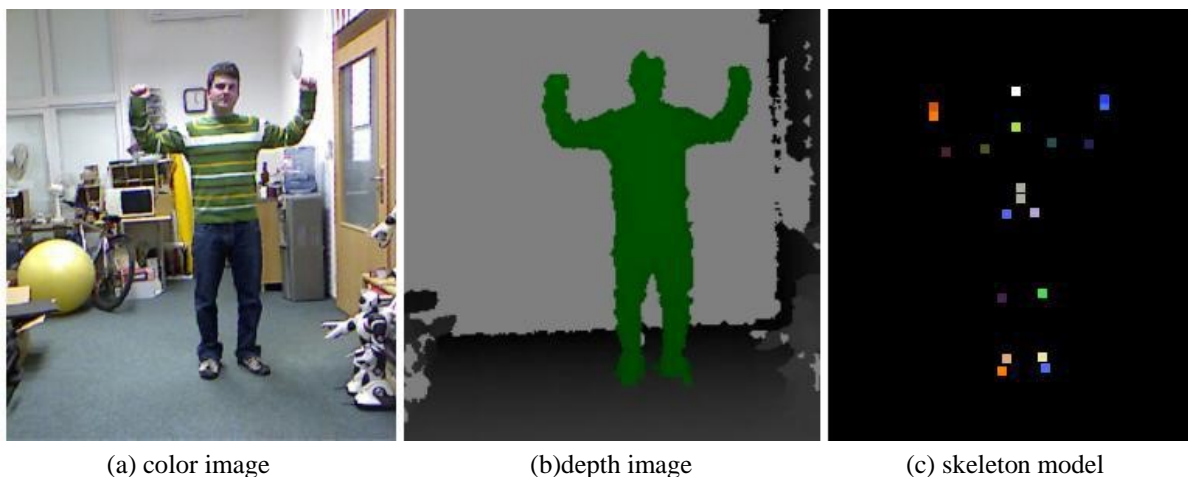
(a) color image   (b)depth image   (c) skeleton model

Figure. 4. The obtained data from Kinect sensor [11]

B. Hand Detection by using sensor

In the proposed system, the method used in cantered coordinates of both the hands first we obtained the hand regions which are cropped on the basis is shown in the fig (4)a the hand cropping is based on the distance between hand subjected and a sensor, from this we can dynamically determine a window size of hand cropping. The horizontal angle of the field view is 430 and vertical angle of the field view is 570[10].Here kindest sensor. The range that each pixel is express is measured by

$$pworld = 0.00354 \times zworld \quad (1)$$

Where pw in "mm" is the depth value from the sensor and Qw is the length of each pixel expresses in real world. In real –world the hand size is as 250*250[mm]. The window size of hand region is defined by

$$RW=250/PW \quad (2)$$

The above fig (4) shows that in-hand elements which are body and head are included in cropped region.
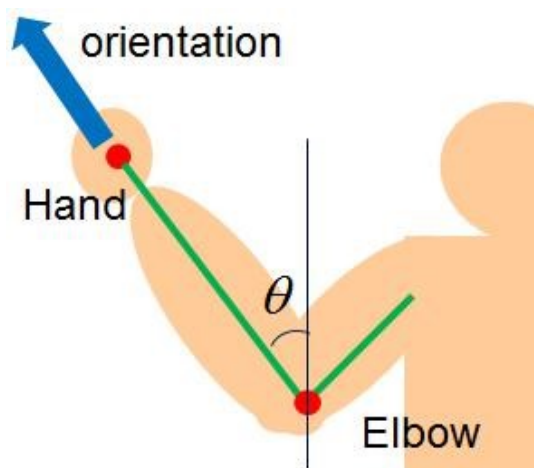
C. Feature extraction of gesture

Figure 4 (a) Hand orientation [12]

The features for gesture recognition are divided into two parts one is hand shape and another one is arm movement. These positions are depends on the velocity of the hand movement on the other hand, the ration (or) contour of the hand has been used as hand shape features. This shown fig (4) (a) This paper we propose a novel.

1) *Arm movement features*: Angles between selected eight joints (hand, elbow, wrist and shoulders) are defined as two angles between shoulder and elbow. In this case, elbow is defined as origin of fracture space shown in the fig (5).
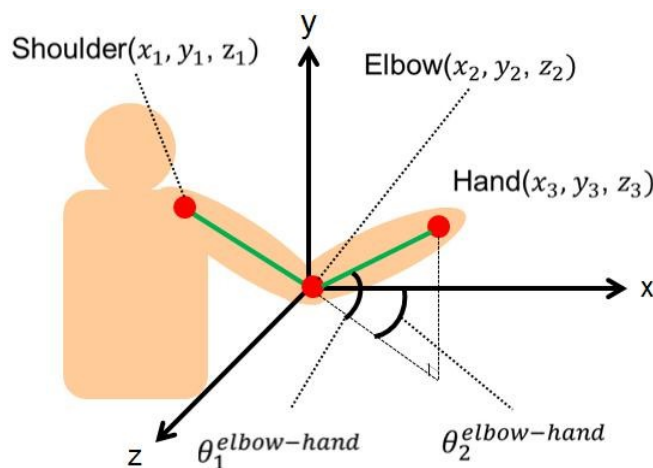

Fig. 5. The detail of proposed arm angle features [9]

b) Before

a) After

Figure. 5(a). Example of hand detection by Otsu threshold [13]

2) *Hand shape features:* Generally, the distance between hand and sensor will give the scale transform of obtained hand shape. A movement of shoulder joint causes the rotation transform of captured hand shape. This feature extraction method by utilizing the concept of speed up robust feature which is known as local robust feature descriptor in various environments [11].This SRUF algorithm is used to estimate the orientation is performed in order to obtain the angle between hand and elbow as orientation it is shown in fig (5a).

D. Classification of gesture by using HMM

Here we use HMM with missing of Gaussians output as a classifier in order to recognize hand gesture. The HMM is the statistical model which is in the process with un observed (hidden) states [12]. The following fig showed the left-right Marco model which consist of states, transaction probabilities and output probabilities this shown in fig (6).
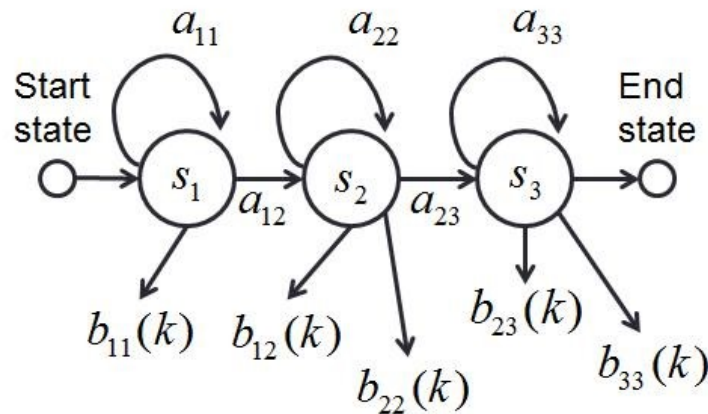

Fig. 6. Hidden Markova model [11]

The arm movement features are defined by:

$$\theta_1 elbow - hand = \tan^{-1} \frac{y_3 - y_2}{x_3 - x_2} \qquad (3)$$

$$\theta_2 elbow - hand = \tan^{-1} \frac{z_3 - z_2}{x_3 - x_2} \qquad (4)$$

In the same way, angles between elbow and shoulder are defined by:

$$\theta_3 shoulder - elbow = \tan^{-1} \frac{y_2 - y_1}{x_2 - x_1} \qquad (5)$$

$$\theta_4 shoulder - elbow = \tan^{-1} \frac{z_2 - z_1}{x_2 - x_1} \qquad (6)$$

It is thought conventionally that the movement velocity of hand is important to recognize hand gestures. Thus, angular velocities are employed for arm movement features.

$$\theta_i velocity = \theta_i t - \theta_i t - 1 \; (i = 1,...,4) \qquad (7)$$

Where t is frame number. Finally, we obtain 16 dimensional feature vectors in both arms.

IV. EXPERIMENT AND DISCUSSION

The good effectiveness of the proposed method, it is compared with conventional method using depth sensor. The performance of proposed hand gesture recognition method gives the good results and also improves the robustness of the conventional 2D method..

The test data is getting from five objects as a restriction, a distance between sensor and subject's hand is set to about 3000mm, 3500mm and 4000mm.Inthe cade1,as a training and test distance dataset used. In the case2, 3500mm gesture data captured at red point is used for tr5aining data. the different distance data of 3000mm and 3500mm is used for test data ,in order to evaluate robustness for environmental changing.

The parameter of mixtures is set to 16 Gaussions, and states of HMM is set to 14experimentally.The number of dimensions in feature vector of conventional method is 62and the number of dimensions in feature vector of proposed method is 42. As a result, we show the recognition accuracies of proposed method and conventional method in table1.The conventional method uses 2D hand fractures such as ratio and ellipse approximation of extracted hand shape.

TABLE I. EXAMPLE OF MISRECOGNITION [5]

(%)	Case 1	Case 2
Proposed method	81.8	84..9
Conventional method	77.6	68.4

From the following table the conventional method in case2 decreased from case1 the accuracy of proposed of method is very high in both methods. Therefore, we conformed that effectiveness and robustness of the proposed method.

V. CONCLUSIONS

In this paper, we proposed the novel hand gesture recognition system by using depth data, which is robust for environmental changing. In this approach an extraction fo hand shape features based on gradient value instead of conventional 2D shape features and arm movement features based on angles between each joints.

REFERENCES

[1] G. R. S. Murthy and R. S. Jadon, "A Review of Vision Based Hand Gestures Recognition," International Journal of Information Technology and Knowledge Management, vol. 2, no. 2, pp. 405-410, 2009
[2] M. Turk, "Gesture Recognition in Handbook of Virtual Environment Technology," Lawrence Erlbaum Associates, 2001.
[3] K. Hoshino and T. Tanimoto, "Realtime hand posture estimation with self-organization map for stable robot control," IEICE Trans. Inf. and Syst., vol. E89-D, no. 6, pp. 1813-1819, 2006.
[4] L. Dipietro, A. M. Sabatini, and P. Dario, "Survey of Glove-Based Systems and their applications," IEEE Trans. on systems, Man and Cybernetics, vol. 38, no. 4, pp. 461-482, 2008.
[5] A. Sato, K. Shinoda and S. Furui, "Sign Language Recognition Using Time-of-Flight Camera (in Japanese)," in Proc. of Meeting on Image Recognition and Understanding, IS3-44, pp. 1861-1868, 2010.
[6] L. Dipietro, A. M. Sabatini, and P. Dario, "A Survey of Glove-Based Systems and Their Applications," IEEE Trans. on Systems, Man and Cybernetics, vol. 38, no. 4, pp. 461-482, 2008.
[7] K. Higashiyama, S. Ono, Y. Wang, and S. Nakayama, "Finger character recognition using 3-dimensional template matching (in Japanese)," IEEJ Trans. EIS, vol. 125, no. 9, pp. 1444-1454, 2005.
[8] H. Takimoto, S. Yoshimori, Y. Mitsukura, and M. Fukumi, "Hand Posture Recognition Robust for Posture Changing in Complex Background," Journal of Signal Processing, vol. 14, no. 6, pp. 483-490, 2010.
[9] E. Ueda, Y. Matsumoto, M. Imai, and T. Ogasawara, "Hand pose estimation using multi-viewpoint silhouette images," in Proc. 2001 IEEE/RSJ Int. Conf. Intelligent Robots and Syst., pp. 1989-1996, 2001.
[10] H. Takimoto, T. Furukawa, M. Kishihara, and K. Okubo, "Robust Fingertip Tracking for Constructing an Intelligent Room", Proc. of 21th IEEE International Symposium in RO-MAN2012, pp. 759-763, 2012.
[11] Y. Wu and T. S. Huang, "Hand modeling analysis and recognition for vision-based human computer interaction," IEEE Signal Processing Magazine, Special Issue on Immersive Interactive Technology, vol. 18, no. 3, pp. 51-60, 2001.
[12] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-Time Human Pose Recognition in Parts from Single Depth Images," in Proc. of 24th IEEE Conference on Computer Vision and Pattern Recognition, pp. 1297-1304, 2011.
[13] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," IEEE Trans. on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62-66, 1979.
[14] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "SURF: SpeedUp Robust Features," Computer vision and Image Understanding(CVIU), vol. 110, no. 3, pp. 346-359, 2008.

Author Profile

S.Chandra Sekhar has Received B.E in Electronics and communications Engineering from Sri Dharamastla Manjunatheswra College of Engineering ,Dharawad in 2000, and M.Tech in Digital Electronics and Computer Systems. He perusing his PhD Degree from Nagapur University ,in 2016. He is life member of MISTE. He is dedicated to teaching field from the last 16 years. His research areas are Antennas and Digital Image Processing. He Published THREE International Journals and He Participated TWO International conferences. At present he is working as Professor in RISE Krishana sai Gandhi group of institutions: Ongole Andhra Pradesh.

Dr Nitiket N Mhala did his 10 and 12 std from Amravati with Merit. He had been awarded National Merit Scholarship. He qualified his BE in Electronics Engineering from Govt College of Engineering Amravati and his Post graduation as ME from Amravati University. He performed his Doctorate (PhD) in Electronics Engineering, RTM University ,Nagpur. Currently he is working as Senior UGC Approved Professor in Electronics & Telecommunication Enginering, BDCE, Sevagram. He was Head of the Department for last Seven years. He had substantial Full time teaching UGC experience of more than 23 years and Two yrs of Industrial Experience. Dr Nitiket Mhala has a expertise superiority in Wireless Adhoc Networks, Mobile Adhoc Networks, Computer Networking and Data communication ,Wireless Sensor Networks and New Emergent Haptics Technology He has a Specialized skills Techniques in Sense of Touch, Zigbee and REVEL Technology. His Research Domain is Remarkable. He had published more than 50 Research papers in a reputed Journals. 20 Research papers in International Conference. He is Working as a Principal Investigator for Major RPS Project and bagged a research grants of from AICTE, New Delhi. He published one book as a first author. He is nominated as Editorial board Member on reputed 12 International Journals. He worked as reviewer for reputed international journals.He is a Phd supervisor in three universities.