Hand Gesture Recognition using Kinect Sensor

Sai Rakshit S Harathas
sairakshit0517@gmail.com

Abstract— Previously hand gesture recognition was done using glove based techniques. But with the increase in the use of automation, humans preferred recognition methods which did not use any hardware even if it meant that the accuracy was compromised. The Human computer interaction with user is not limited to keyboard, mouse and pen due to evolution of ubiquitous computing. Applications of hand gesture recognition include Phobia therapy, military simulation and medical training. In this project, a two-stage HGR system is proposed to recognize hand using Kinetic sensor which includes segmentation and recognition using both colour and depth information. As most of the mobile cameras used have both depth and normal camera, thus this technique can also be implemented using an app. Here for the recognition part of the project I have used deep learning using autoencoders and verified the algorithm for different training sets

.Index Terms—Human computer Interaction, Human Gesture Recognition, Kinetic sensor

.Index Terms—Human computer Interaction, Human Gesture Recognition, Kinetic sensor

I. INTRODUCTION

The gesture message can be conveyed as posture of the body, gesture of the hand or it can be through movement of the body the hand coveys some messages. Gesture can be physical characteristic and they are different from traditional hand recognition. The hand gesture method is greatly different method when compared to conventional for Human computer interface..The hand gesture helps in application like sign language interpreters, virtual reality, medical service.
In real world situation, the interaction between user and computer has given more preference so that leads to hand, vision and all other sensory nodes gesture recognition have become a great attention in the research field. These innovation gesture technologies replace the convention used human machine interactive devices like keyboard, mouse etc.

A. HUMAN COMPUTER INTERFACE SYSTEM

Computer has become important working model for our regular basis life style. The computer model converts raw facts and features into meaningful information and labels. So special input and output devices are design to communicate between computer and human [1].The conventional interface is the general common message exchanger but new ideas and innovation allows computer more intelligent and that allows human to work in different complicated application with human computer interface. This is possible due to result oriented efforts made by computer professional creating innovative human computer interface [1].The modernized daily life in human style are turned into more complex and continues to grow more complicate application so, the need for complex programming skills and ability are needed for computer programmers to survive in a competitive environment.

The applications like monitoring system, medical application has developed with higher quality because of wide use research and developed tele operating robotic. Before software development process were avoiding complex programs as the focus was limited for configurable features .Now days it is observed that because of user friendly environment and better quality has made programmers to revisit the focus area[1]. Gestures are the nonverbally communication and play vital role in human computer interface researchers and it allows performing in more advance application at a time.

B. HAND GESTURE

In present research field, vision based hand gesture has become very advance in techniques in order to replace commonly used human machine interface devices conventional methods like keyboard, mouse etc. and it has become important segment in real time application. Use of hand gestures provides attractive and natural way to interact with computer interface and plays vital role with computer. People can easily convey information with computer in many ways using hand gesture. Hand gesture movements is non verbal communication can range from simple actions to more complex ones [2]. So, the primary goal of researcher is to develop a specific human gesture system to identify gesture and convey information or control device [3].
The gesture recognition system can be classified into preprocessing using kinetic sensor, Feature extraction method Training and classification or recognition as illustrated in fig 1.

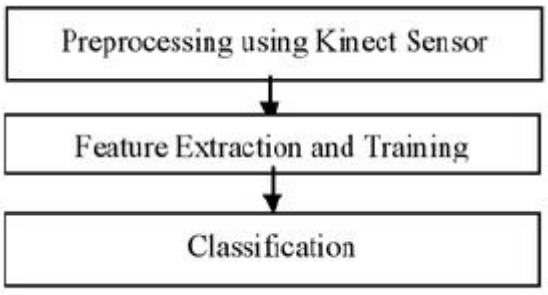

Fig 1. Hand Gesture recognition system steps

To perceive hand gesture, segmentation is the underlying procedure and afterward in this procedure input image is partitioned into locales isolated by limits [4]. segmentation procedure can be either powerful or static motion. On the off chance that it is motion is dynamic, at that point the hand gesture found and followed yet on the off chance that it is

Received: 15 July 2019/Accepted 14 September 2019.

static, at that point information picture is sectioned .The location of hand can be done using bounding box depending on the depth of skin color. Two main approaches for tracking the hand; Video is divided into frames and frames which are treated as gesture and segmented can be processed alone [5]. Kalman filter tools are used for tracking hand gesture by using information like shape and skin color [6].

Segmentation process is a vital play role in recognition process and based on specific application, the different methods are used to extract feature vector from segmented image. Hand contour and silhouette [6] are the different methods are used to get feature vectors from shape of the hand.

There are various popular Gesture classification algorithms are used to recognize the gesture. Some of the algorithm like Edge detection or contour operators [7] cannot be utilized for gesture recognition because after generation could create misclassification. Some of the tools like Euchdean distance metric are applied to classify gestures by using suitable distance measure, statistically based models are used for gesture classification, Temporal Gesture which characterize movement of hand can be recognized by using HMM tool, Finite State Machine (FSM), Learning Vector Quantization and classical competitive learning by Principal Component Analysis (PCA)[8,9,10].Hand shape gesture extraction can be applied using Neural network classifier [12] and also NN can be used for hand gesture recognition [11]. Fuzzy C means clustering (FCM) [5] can be applied to recognize by using clustering method in which one feature data belong to two or more clusters and Genetic Algorithms GAs method is biological evolution [12].

II. RELATED WORK

C. Keskin [1] has introduced a framework which distinguishes the hand gestures and captures the gestures of user having glove with shading. For noise elimination this framework utilizes 3D kalman filters and two color cameras for 3D reconstruction. Hidden Markov Model is used for 3D dynamic gesture recognition. This HMM is utilized for diminishing the spatio-temporal variability. The spatio temporal variability means the duration and shape of every gesture differs when capturing the image. This framework has two section First is the division .Using HMM it is difficult to find the begin and end focuses so they have remade the model with Baum-Welch calculation. The subsequent part is the gesture spotting and is executed by Adaptive Threshold model. The precision of 98.75% was accomplished with 160 trails of 8 defined gestures. Starner et al. [13] used Hidden Markov Model using 2D feature vector input to recognize American sign language .The precision was limited to 92% and able recognize 40 signs. In order to get stability, the user used inexpensive colored gloves and used single colored camera to track orientation and shape of the hand. Huang and Huang [14] used kinetic sensor along with SVM classification to achieve a recognition rate of 97% with a 55 samples. The author implemented to recognize American sign language based on arm movement The approach failed to provide facial and finger movement information and also author have used fixed time slot duration for ten frames so that restrict the speed of which sign can be recognized.Biswas and Basu[15] used images from kinetic sensor for hand recognition like clapping and waving .The author presented lesser compute intensive approach and that makes accuracy to be improved further considering skin color information from RGB feature data.

III. ISSUES OF TECHNIQUES OF HAND GESTURE

After acquiring the input image from camera(s), videos or even data glove instrumented device, the researchers can characterize gesture recognition framework functionality into Extraction Method, features estimation and extraction, and classification or recognition

The Extraction method of hand gesture recognition can be done by utilizing division [4] and if it is static gesture then the input image must be divided and in the event that it is dynamic gesture then the hand gesture should be found and followed [5].Bounding box is utilized to locate the hand based on skin shading and secondly, the hand must be followed. There are two fundamental methodologies for following hand the video is isolated into edges and each edge must be handled alone, for this situation the hand edge is treated as a Gesture and sectioned [4], or utilizing some following data, for example, shape, skin shading utilizing a few apparatuses, for example, Kalman filter [5].

Great segmentation procedure leads to perfect features extraction process and the latter play an vital role in a successful recognition process [5]. The shape of the hand ,for example, hand form and silhouette and used fingertips position, palm focus, etc. are the methods to recognize hand by using feature vector data from segmented image.

The next section therefore discusses about certain standard for Hand Gesture, which were introduced in past decade and still considered to be adopted by various researchers.

IV. PROPOSED SYSTEM DESIGN

The prime reason for the proposed framework is to propose a hand motion displaying utilizing deep neural system. The prime explanation behind actualizing deep neural system in our demonstrating is essentially to diminish the reaction time and increment the precision. The schematic architecture of proposed system is shown in Fig.2

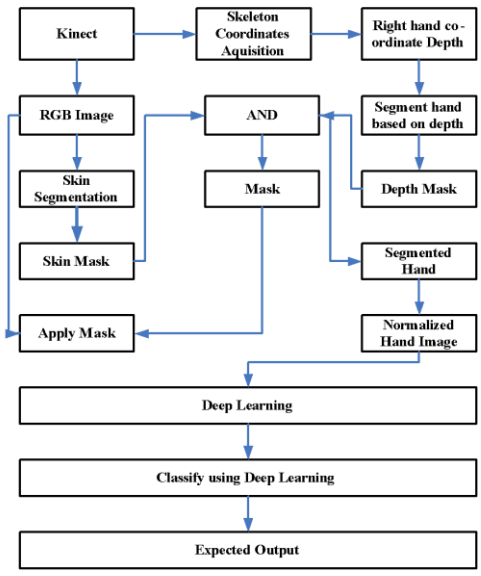

Fig. 2 Proposed system block diagram

Received: 15 July 2019/Accepted 14 September 2019.

The figure shows the proposed system. The kinetic camera and an RGB camera to allow hands overlapping with the face, however, calibration is necessary for a mapping from depth data to RGB image. The method uses various Kinect to follow the hand and a HD shading sensor to section the hand for stable gesture recognition. That, be that as it may, requires the administrator wears shaded gloves just as all sensors to be adjusted. In this, depth information is first used to locate the front most object from the sensor, which is viewed as hand and arm part; the user should wear dark belt so as to separate hand from the arm.

V RESULT ANALYSIS

In this paper a code is designed to check whether it can detect hand gestures correctly. Video clips from signer of hand postures are captured.

To all the almost certain recognize and track hands, Kinect is used to rough data streams and skeletal information in our structure. It incorporates a RGB camera and a depth sensor, which captures full-body development.

Kinect SDK supports for default skeletal framework and seated skeletal system, and we use the last mode, where ten chest region joints (shoulders, elbows, wrists, arms and head, as appeared in Figure 2) are followed rather than twenty joints of the whole body. This is satisfactory for our system, since communication via gestures just contains hand movements.

In this, the Kinect camera is associated with PC and coded by utilizing MATLAB. From, the kinect camera both the depth and shading pictures of the human body is caught. The figure for the equivalent is appeared in Fig 3. From the gained shading picture, utilizing the YCbCr model the skeleton focuses were followed and the left hand directions of the hand were gathered

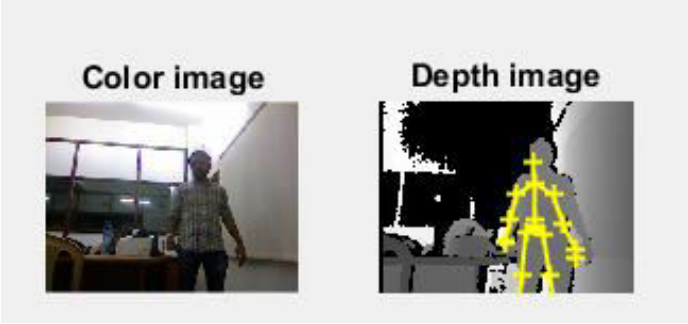

Fig 3: Capture from Kinect Sensor

The depth value is obtained from the depth image using the depth sensor and the complete palm image is obtained using the equation as mentioned

$$Depth_{palm} - \Delta d_{back} < img(x,y) < Depth_{palm} - \Delta d_{front}$$

Where $Depth_{palm}$ is obtained from the kinect sensor and Δd_{front} and Δd_{back} is decided by trial and error method. The remaining regions are blacked out

Later, post-processing is performed on these segmented depth and color image. All images are converted to gray scale and resized to a resolution of 32x32 pixels. Then later the image is converted to binary image where the foundation is dark in shading and the frontal area is white in colour. An model is appeared in Fig 4

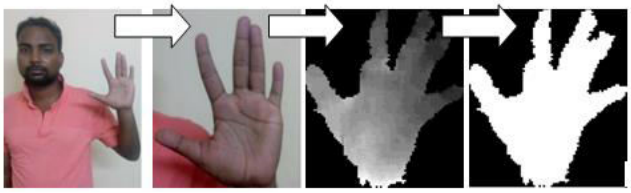

Fig 4: Pre-Processing of Hand Gesture

A total of 225 images are used for 11 different and hand gestures namely for gestures 1, 2, 3, 4, 5, A, B, C, I, K, L The input database obtained from the kinect sensor. In order to check the robustness of the algorithm the images are taken in different background ,skin and lighting conditions Also the hand size is varied for different images along with the rotational effect of the hand is also taken into consideration

After obtaining the hand gesture database, based on different training and testing sets we have checked the performance of the algorithm . These images are trained using autoencoders and the result obtained are as follows are shown in Table I

Training Data %	*Training Nos Based on %*	*Testing Nos Based on %*	*Best Accuracy*	*Avg Accuracy*	Worst Accuracy
88.88	199	26	100	100	99
80	180	45	100	99	97
70	157	68	98	95	93
60	135	90	96	91	86
50	112	113	90	85	79
40	90	135	85	79	69
30	67	158	80	73	57
20	45	180	72	61	45
10	22	203	65	50	38

Table 1: Results for different training sets

Thus we can infer from the table that even with low training sampling (about 60%) we can get 85% average accuracy and 79% worst accuracy. Thus this project can be extended to get more advanced results like using varieties gestures

Received: 15 July 2019/Accepted 14 September 2019.

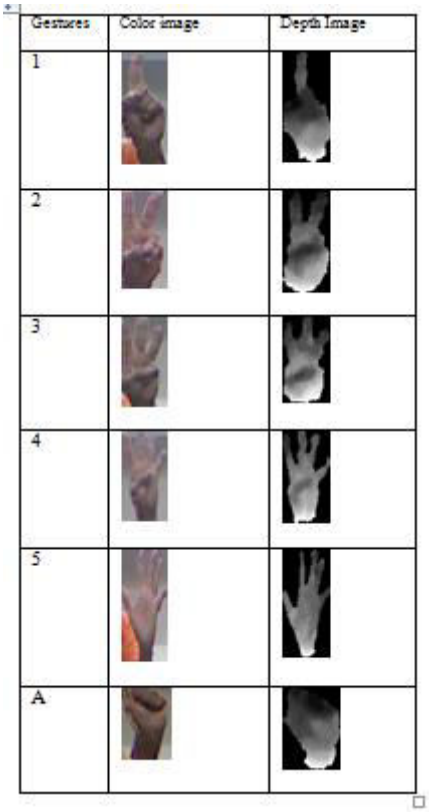

Fig.5. Images for different Hand Gestures

IV. CONCLUSION

In this section, different kinds of hand motion acknowledgment strategies and its application in programmed communication via gestures acknowledgment is featured. Different squares of hand motion acknowledgment framework are talked about with related research works. As the hand signal acknowledgment is experienced foundation conditions, a couple of looks about significant works have been considered. A couple of difficult issues like development epenthesis and co-explanation which keep from appropriate constant communication via gestures acknowledgment are additionally secured. It is seen that in spite of the fact that glove-based hand signal acknowledgment systems are precise, vision based methods are ideal on account of the instinctive nature of the hand development and simplicity of operation. As future work the number of gestures can be increased and also the performance of the used algorithm can also be checked by comparing it with other algorithm. Also this application can also be extended for controlling vehicles and for other real life suitations.

REFERENCES

[1] C. Keskin, A. Erkan, L. Akarun .Real time hand tracking and 3D gesture recognition for interactive interfaces using HMM‖ .IEEE Transaction J. U. Duncombe,

[2] Hong Li, Michael Greenspan. ―Model based Segmentation and recognition of dynamic gestures in continuos video streams. Springer publications in Patternrecognition,44(2011)1614-1628

[3] Fabio Dominio, Mauro Donadeo, Pietro Zanuttigh Combining multiple depth-based descriptors for hand gesture recognition, Elsevier, Pattern recognition Letters 2013.

[4] Sheng-Yu Peng, Kanoksak Wattanachote, Hwei-Jen Lin, Kuan-Ching Li, ―A Real-Time Hand Gesture Recognition System for Daily Information Retrieval from Internet, IEEE Fourth International Conference on Ubi-Media Computing, 978-0-7695-4493-9/11 © 2011.

[5] M.M.Gharasuie, H.Seyedarabi, Real-time Dynamic Hand Gesture Recognition using Hidden Markov Models‖ , 8th Iranian Conference on Machine Vision and Image Processing (MVIP), 2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP), 978-1-4673-6184-2/13

[6] Zhou Ren, Junsong Yuan, JingjingMeng, Zhengyou Zhang. ―Robust Part-Based Hand Gesture Recognition Using Kinect Sensor. IEEE Transactions on Multimedia, Vol.15, No.5, Augus

[7] Nick C. Tang, Chiou-Ting Hsu, Chih-Wen Su, Timothy K. Shih, and Hong-Yuan Mark Liao.Video Inpainting on Digitized Vintage Films via Maintaining Spatiotemporal Continuity‖ published at IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 13, NO. 4, AUGUST 2011.

[8] N.Neelima , M.Arulvan, B. S Abdur, ―Object Removal by Region Based Filling Inpainting‖ . IEEE transaction 978-1-4673-5301/2OI3 IEEE.

[9] Jan Herling, and Wolfgang Broll, High - Quality Real -Time Video In painting with Pix Mix, IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS,VOL.20, NO. 6, JUNE 2014.

[10] Qing-Ge Ji ,Zhi-DangFang ,Zhen-HuaXie , Zhe-Ming Lu Video abstraction based on the visual attention model and online clustering . Published at Elsevier, Signal Processing: Image Communication 28 (2013)241–253 in 2013.

[11] Massimo Camplani , Carlos R. del Blanco , Luis Salgado , Fernando Jaureguizar , Narciso García, ―Multisensor background subtraction by fusing multiple region-based probabilistic classifiers‖, published at Elsevier Pattern Recognition Letters 2013.

[12] Massimo Camplani , Luis Salgado, ―Background foreground segmentation with RGB-D Kinect data: An efficient combination of classifiers‖. published at Elsevier J. Vis. Commun. Image R. 25 (2014) 122–136.

[13] Blanca Miriam Lee-Cosioa, Carlos Delgado-Mataa, Jesus Ibanezb. ―ANN for Gesture Recognition using Accelerometer Data. Elsevier Publications, Procedia Technology 3 (2012) 109 – 120.

[14] Jino Lee, Dong-Kyu Lee, and Rae-Hong Park, A Robust Exemplar-Based Inpainting Algorithm Using Region Segmentation‖. IEEE Transactions on Consumer Electronics, Vol. 58, No. 2, May 2011

Received: 15 July 2019/Accepted 14 September 2019.