GESTURE RECOGNITION USING HIDDEN MARKOV MODEL

M. K. S. S. Jayasinghe[1], T. W. A. S. L Jayasinghe[2], A. Jeevithan[3], M. B Dissanayake[4]

Department of Electrical and Electronic Engineering, Faculty of Engineering, University of Peradeniya, Sri Lanka.
Email: [1]suwanijayasinghe@gmail.com,[2]amaljayasinghe1@gmail.com, [3]ejeevi10@gmail.com and
[4]maheshi@ee.pdn.ac.lk

ABSTRACT

Hand gestures are used by many applications, such as mobile phone, robots, television and gaming purposes. In order to employ hand gestures as input for a control application, it requires recognition of gesture with a high precision rate, with minimum probability of error. In the research presented, the signs are executed with one hand as it is sufficient for many everyday applications. The gestures are captured by using general purpose web camera. (HMM) is employed for gestures recognition as this technique has been applied successfully in speech and symbol recognition with low latency. Recognition is performed by employing the HMM based analysis. HMM was implemented using MATLAB HMM toolbox.

***Key words*:** Hidden Markov Model, Hand Gestures, Gesture recognition

1. INTRODUCTION

Hand gesture or similar procedure is used for communication in many applications even in devices such as mobile phone camera operations, gaming consoles, robot controlling and etc. In order to implement hand gesture as inputs for control applications, it requires precision with minimum probability of error. Hidden Markov has being used for similar purposes such as voice recognition. Most of the gesture recognition techniques are based upon recognition of static gestures [1].

There are two basic approaches used.

1. The top-down approach, where a previously created model of collected information about hand configurations is rendered to some feature in the image co-ordinates. Comparing the likelihood of the rendered image with the real gesture image is then used to develop whether the gesture of the real image corresponds to the rendered one.
2. The bottom-up approach, which extracts features from an input image and uses them to query images from a database, where the result is based on a similarity measurement of the database image features and the input feature.

The disadvantage of the first approach is that it seems to use a high computational effort in order to achieve robust recognition. The second approach however requires an adequate preprocessing in order to achieve a reliable segmentation.

This report mainly keeps the focus on the latter approach stick with Hidden Markov model.
It is very challenging to recognize dynamic gestures [2]. Hidden Markov Model would rise proficiently giving resolution for dynamic gestures.

In Hidden Markov Model the state of the object observed is not visible [3]. But the observation is always available. Emission matrix and the transition matrix are essential to recognize using Hidden Markov Model. Apart from that the initial state probability has to be known. To calculate the set of states using the observation equations such as Viterbi can be used [4].

Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture recognition, part of speech tagging, musical score following, partial discharges and bioinformatics.

2. METHODOLOGY

2.1. Hidden Markov Model

In this research, the signs are executed with one hand as it is sufficient for many applications. Four dynamic gestures are used for the results presented in this paper, yet the model presented can be explained with simple modifications. Model using MATLAB hmm toolbox and

achieve in order to feed it into the control application.

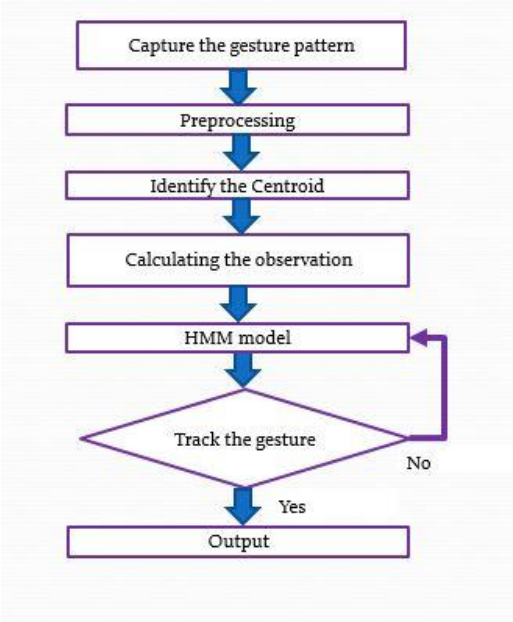

Figure 1: High level architecture of dynamic gesture recognizing

Hidden Markov Model dynamic gesture matching system presented in this paper uses several steps as shown in the Figure 1. In this system, a green background is used to capture the video for the simplicity of the implementation [5]. First, input to the system is captured as RGB video from the camera. The input sequence must be of a minimum of 3 seconds. Centroid of the hand which is used as a parameter for estimation. In order to identify the centroid of hand as in figure 2, red mark was placed on the palm of the hand and red color [6] was detected by setting a specific threshold value.

Figure 2: Detected centroid in a frame

Next the video was converted into frames and then each frame(image) into binary mage with read mark as white. Blob analysis is used to find the centroid of the hand in each frame and the (X,Y) coordinates of the centroid of the hand were extracted

Each frame after every 14th frame is selected as a state. Then angle between two successive frames was calculated which was taken as observation.

Four states were used as only four hand gestures are recognized. The only parameter used as observation is the directional angle. That is the angle which is made by position of the centroid of the hand in the current frame with the previous frame. The angle is calculated and it is rounded off to the nearest integer 0, 20, 40, to 180. Therefore only possible set of observations are 1, 2, to 18 denoting the full 360 degrees.

We assumed the first portion of each hand gesture to be the first state. Then the transition matrix and observation matrix were estimated using the available gestures. Recognition is performed by adapting the Hidden Markov. The Emission matrix of HMM represents the possible observations. Using the observations initial transition matrix (TRANS_GUESS) and the emission matrix (EMIS_GUESS) were assumed. After acquiring those two separate matrices, they were fed into hmm toolbox which would able to provide estimated transition (TRANS_EST) and estimated emission matrix (EMIS_EST) precisely with the aid of actual observations using eq.(01).

$$[TRANS_{EST}, EMIS_{EST}] = hmmtrain(seq, TRANS_{GUESS}, EMIS_{GUESS}) \quad (01)$$

In order to recognize real time gesture hmmviterbi was used. It captures the observation parameters of the real time gesture and was able to provide the most likely path which could have been taken by the gesture with the aid of estimated transmission matrix and estimated emission matrix.

3. RESULTS

Throughout the test four different gestures were used such as letter ‘L’, letter ‘O’, letter ‘V’ and a combination of letter ‘L’ and ‘V‘ as mentioned in table 01.

The success percentage of this model is calculated using eq(02).

$$\textbf{Success percentage} = \frac{Success\ matches}{Input\ Videos} \times 100\% \quad (02)$$

The same four gestures were obtained by different participants and it is applied to the

system. Then the success percentage is calculated using eq(02).

Table 1: Gesture and its corresponding id

Gesture ID	Gesture
one	L
two	O
three	V
four	

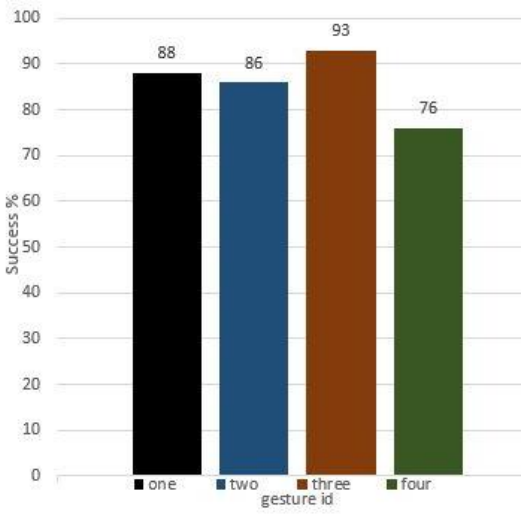

Figure 3: Success percentage for each gesture

According to the results in Figure 3, the gesture id 4th gets the lowest successe rate. That's because gesture id 1st and gesture id 2nd are close enough to gesture id 4th. Since that it has comparatively low likelihood to give higher success rate compared to others.

4. CONCLUSION

The research paper presents a mechanism that can be used to recognize dynamic hand gestures using Hidden Markov Model and image processing techniques. The target of this project is to develop a prototype system using above mentioned mechanism that will help to access or control modern gadgets in a convenient manner. The system implemented is capable of analyzing real time gestures in a more accurate way if number of training videos for each gestures is increased and the number of observation parameters is increased. However with improvement of the accuracy of the recognition system complexity will be increased. So it will be the trade-off between accuracy and the latency.

5. REFERENCES

[1] H. Zhou, D. Lin, and T. Huang. *"Static hand gesture recognition based on local orientation histogram feature distribution model*", In CVPRW '04: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition workshop (CVPRW'04), Vol.10, pp. 161, Washington, DC, USA, IEEE Computer Society, 2004.

[2] A. D. Wilson, and A. F. Bobick, "*Hidden Markov Models for Modeling and Recognizing Gesture under Variation*", International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), Vol.15, no.1, pp.123-160, 2001.

[3] M. Elmezain, A.H. Ayoub, S. Sadek, B. Michaelis, *"Robust methods for hand gesture spotting and recognition using hidden Markov model and conditional random fields"*, IEEE Proc. International symposium, pp. 131-136, 2010.

[4] L. R. Rabinar, "*A tutorialon Hidden Markov Models and Selected Applications in Speech Recognition",* Proceedings of the IEEE vol. 77, no. 2, pp.257- 286, 1989.

[5] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. *"Pfinder: Real–time tracking of the human body",* IEEE"Transactions on Pattern Analysis and Machine Intelligence, vol. 19 of *7*, pp. 780–785, 1997.

[6] M. Dissanayake, H. Herath, W. Kumari and W. Senevirathne, *"Image processing based Sinhala sign language recognition system"* Proceedings of the IESL Transactions of Technical Papers, vol.1-part B, IESL, 2013.