International Journal of Computer Applications Technology and Research
Volume 4– Issue 5, 386 - 389, 2015, ISSN:- 2319–8656

Hand Gesture Recognition using Colour Based Segmentation and Hidden Markov Model

Kshitish Milind Deo
Computer Engineering Department, Pune Institute of Computer Technology, Pune, India.

Avanti Yashwant Kulkarni
Computer Engineering Department, Pune Institute of Computer Technology, Pune, India.

Tirtha Suresh Girolkar
Computer Engineering Department, Pune Institute of Computer Technology, Pune, India.

Abstract: Automatic gesture recognition is a key technology used in Human Computer Interaction. In this paper we introduce a hand gesture recognition system which consists of 4 modules, image segmentation, feature extraction, HMM training and gesture recognition. Image or video is divided into multiple frames and segmentation process which uses colour based detection of the path of the object is applied to each frame. Feature extraction process mainly considers the orientation of the state tracked. This is done using HSV[hue-saturation value] image and contour mapping of the image. The training part of the HMM model works on basis of LRB(Left-Right-Banded)topology and uses the BW (Baum Welch) algorithm. We have used Viterbi algorithm for mapping the state to a symbol i.e. recognition. HMM is used to predict the gesture and increase the tolerance of the system to incorporate human errors.

Keywords: Hidden Markov Model (HMM), Forward Backward Algorithm, Baum Welch Algorithm, Viterbi Algorithm, Colour Segmentation.

1. INTRODUCTION

The goal of gesture interpretation is to enhance the advanced human machine communication so as to make it more close to human-human interaction. Few of the models used for this purpose are Neural Networks, Fuzzy logic and HMMs. We are going to propose a system which is based on HMM model.

In this paper, HMM model is used for dynamic hand gesture recognition. HMMs can be successfully used for both speech and two-dimensional signs, because their state based nature enables them to capture variations in duration of signs, by remaining in same state for several time frames. Gesture recognition is a step by step process which has input as sequence of image frames and output as a symbol.

Here, a system is developed to recognize geometric shapes drawn with a blue coloured object. Colour is detected and pattern of hand movement is analyzed. This gesture is divided into multiple states. Output symbols are extracted from the gesture. These form parameters for the Hidden Markov Model (HMM).Colour detection technique has been used in our proposed system so as to track the path of the object using which the desired shape is being drawn. After the detection part the main issue is how to make the computer understand the gesture. Recent works can be said to use two methods: Data-glove based methods and vision based methods. The Data Glove method uses sensor devices for digitizing hand and finger motions for multi-parametric data. For Vision-based method, the only required equipment is a camera.

Challenges in vision based system are, it needs to be background invariant and lighting insensitive.

In the upcoming sections we will see how the segmentation part is being done using colour filtering. We will briefly talk about the process of feature extraction which requires contour detection and calculation of centroid of each contour in each frame. Section 3.3 will be encompassing the explanation of how HMM training and recognition works for our system.

2. HISTORY AND LITERATURE SURVEY

There exist many reported research projects related to learning and recognizing visual behaviour. However due to its recent introduction to the vision community, only a small number have been reported which use Hidden Markov Models. HMM has been traditionally used as tool for speech recognition tool, recent researches have begun relating the speech variations to visual gestures. Moni M. A. et al in their review paper [6] have analysed various techniques and approaches in gesture recognition for sign language recognition using HMM. They have provided an overview of HMM and its use in vision based applications, working in two stages that of image capturing and processing using cameras, and the second stage for identifying and learning models has eliminated the need of previously used sensor embedded equipment such as gloves for tracking of a gesture. T. E. Stanner have employed HMM in 1995 in identifying the American Sign language. On similar grounds authors Gaus Y. F. A. et al have successfully recognized the Malaysian Sign Language[5], skin segmentation procedure throughout frames and feature extraction by centroids, hand distances and orientation has been used, gesture paths define the hand trajectory. Kalaman filters have been used by researchers to identify overlaping hand-head and hand-hand regions. In [1] Elmezain M, Al-Hamadi A, MichaelisB,have quantized features form spatio-temporal trajectories into codewords. They have used a novel method of tracking the gesture by using 3D depth map along with colour information, this helps at separating the same colour at different surfaces in a complex background. In order to separate continuous gestures a special zero codeword is defined, using the start and end points of meaningful gestures the viterbi algorithm is employed or recognition. In [2] the authors have used the LRB topology along with forward algorithm to achieve the best performance. With a recognition rate of 95.87% arabic numbers have been identified. Shrivastav R[3] use OpenCV image processing library to perform the isolation of gesture frames, the entire process form per-processing to testing. In coordination with this processing, Baum-Welch algorithm and LRB topology with forward algorithm is applied for recognition.

3. DESIGNAND ANALYSIS OF SYSTEM

3.1 Segmentation:

We use in our implementation, a colour based segmentation approach to extract the object used. Gesture video is captured using a generic web cam. For each frame in the video, contour of object (blue colour) is tracked. Minimum threshold area is given so as to put a constraint on the size of the object to be tracked, this avoids the tracking of the accidental blue colour that appears in the background. After identifying this contour we calculate the centroid of the area.

3.2 Extraction.

Selecting good features to recognize the hand gesture path play significant role in system performance. There are three basic features; location, orientation and velocity. The previous research showed that the orientation feature is the best in terms of accuracy results. Therefore, we will rely upon it as a main feature in our system. We will use the calculated centroid co-ordinates of each frame as a measure to deduce the orientation feature. Orientation is defined as the angel of the vector made by the centroid of two consecutive frames (refer Figure 1). As observation symbols for HMM this orientation is normalized. For normalization purpose we divide the 360 degree angles into 18 parts, 20 degrees each. Codewords are calculated after this normalization, to be further used(refer Figure 2.).

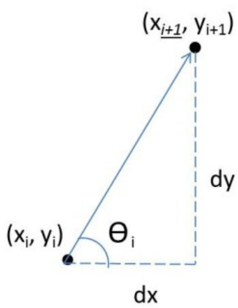

Figure 1. Orientation Calculation

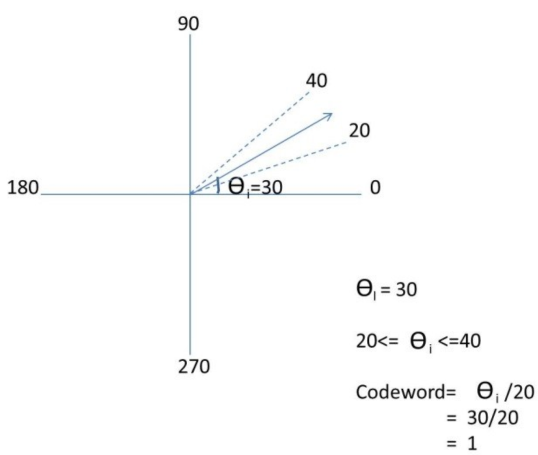

Figure 2.Codeword calculation

3.3 Recognition

HMM is a mathematical model of stochastic process. Evaluation, Decoding and Training are the main problems of HMM and they can be solved by using Forward-Backward, Viterbi and BW algorithms respectively. Also, HMM has three topologies; Fully Connected (i.e. Ergodic model) where any state can be reached from other states, LR model such that each state can go back to itself or to the following states and LRB model in which each state can go back to itself or the next state only.

3.3.1 Hidden Markov Model:

$HMM = (\pi;A;B)$ where π represents initial vector, A is the transition probability matrix and B refers to emission probability matrix.

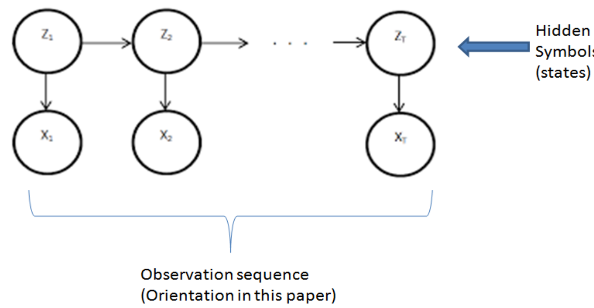

Figure 3. Trellis Diagram

In above trellis diagram, z is the hidden states and the x is the observation symbol. There is transition from z_1 to z_2 and so on to z_n. z_1 gives the observation symbol x_1, z_2 gives x_2 and so on. We can find the transition probability and the emission probability from the given trellis diagram. Transition matrix is the matrix of probabilities of the transitions of states to other states and the emission matrix is the matrix of the probability of states to emit observation symbols.

We can write the equation of HMM as

$$P(x_1,x_2,.....x_n,z_1,z_2,.......,z_n) = P(z_1)\ P(x_1 \mid z_1) \prod_{k=2}^{n} P(z_k \mid z_{k-1}) . P(x_k \mid z_k)$$

Let us write $P(x_1,x_2,.....x_n,z_1,z_2,.......,z_n)$ as P(X,Z).

Where,

$P(z_1)\ P(x_1 \mid z_1)$ is the probability of x_1 given z_1. It is the initial state(π). We have added it as it doesnot have any previous state.

$P(z_K \mid z_{k-1})$ is the probability of z_K given z_{k-1} .This represents the transition state. Let us denote it as A.

$P(x_K \mid z_K)$ is probability of x_K given z_k .This represents the emission state. Let us denote it as B.

So the HMM equation is

$$P(X,Z)=\pi(i).B_{z1}(x_i)\prod_{k=2}^{n} A(z_{k-1}, z_k) . B_{zk}(x_k)$$

With Hidden Markov Model, we can solve following problems

1. Match most likely system to sequence of observation(using forward algorithm)
2. Determine hidden sequence generated by sequence of observations(using Viterbi algorithm)
3. To model parameters which might have generated sequence of observations(Using forward backward algorithm)

In the paper we have used, Viterbi algorithm for recognition and Baum Welch (BW) algorithm for training purpose. Forward backward algorithm is used for evaluation purpose. In forward backward algorithm, probability of z_k given x is found ie. $P(z_k|x)$.

International Journal of Computer Applications Technology and Research
Volume 4– Issue 5, 386 - 389, 2015, ISSN:- 2319–8656

Assumption is that HMM parameters π(initial state), transition probability and the emission probability is known. For this purpose we use forward algorithm and backward algorithm.
In forward algorithm we calculate probability of z_k given $x_{1:k}$ ie. $P(z_k|x_{1:k})$. Note that when we write $x_{1:k}$ it means $x_1,x_2\ldots\ldots,x_k$.

In backward algorithm we calculate probability of $x_{k+1:n}$ given x_k ie $P(x_{k+1:n}|x_k)$.
Thus the forward backward algorithm is the multiplication of the two probabilities generated from the forward algorithm and backward algorithm.
$P(z_k,x)=P(x_{k+1:n}|z_k,x_{1:k}).P(z_k,x_{1:k})$
In Viterbi algorithm, we find the maximum likelihood of the given sequence to the trained model. Thus the goal is to find $Z^*=\mathrm{argmax}P(z|x)$.

4. EXPERIMENTATION ANALYSIS

4.1 Segmentation

Our experiment consists of detecting the motion of blue coloured object. Now one issue is that different shades of blue from the background can get detected and create disturbance. To avoid this we have specified a particular range of blue intensity that is to be considered for detection. Range values used in our code are min[95,50,70,0] and max[145,255,255,0].
Another issue which we deal with is the size of blue object. To avoid detection of unnecessary blue objects we have put a constraint of area of the object to be detected i.e. the area should be greater than 1000{units}. Each frame is then passed to a filter where contours of the object are drawn.

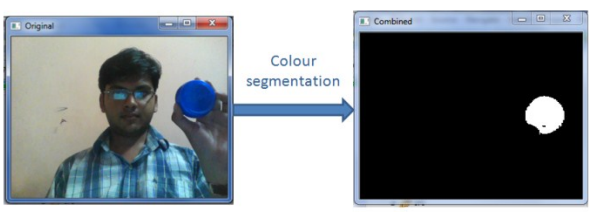

Figure 4. Image Segmentation

4.2 Feature Extraction

Here, we used the Orientation feature for extraction. We first calculated centroids of contour of each frame. Thus we got the position of the blue spot in each frame. Using position of blue spot in consecutive frames we calculated angle of orientation.
For convenience we have normalised the angle by forming groups of 20 degrees each as follows.

```
angle = atan2(y2 - y1, x2 - x1));
                angle -=180;
           if(angle < 0){
                  angle += 360;

             }        Where x1,x2,y1,y2 are the coordinates of the centroid.
```

Figure 5.Code snippet for angle calculation

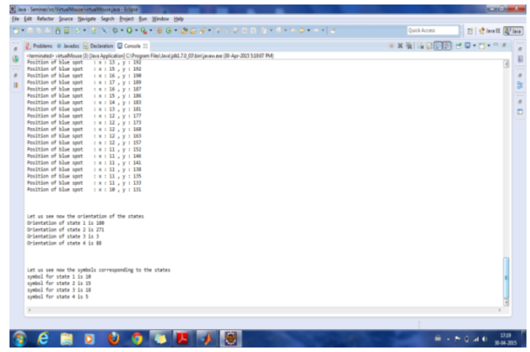

Figure 6. Centroid co-ordinates of each frame and output states and observation symbols

After normalising the orientation angles we get an associated code word for each state as follows.

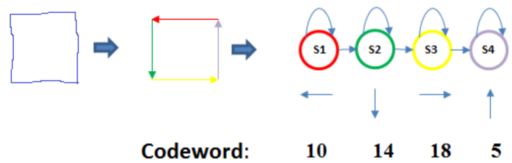

Figure 7. State Symbol Mapping for Anticlockwise square.

In above figure, s1 is first state with horizontal line from right to left. The approximate angle is 180°. After normalization we get symbol as 10. Similarly for s2 we get symbol as 14, for s3 we get 18 and for s4 we get symbol 5.

4.3 Hidden Markov model analysis and recognition

We have got hidden states and observation symbols for analysis of HMM. From previously trained samples we have emission and transition probability matrix. This sequence of symbols is given to the viterbi algorithm for checking the likelihood of the model with the trained model. The threshold is fixed to 80% likelihood. Thus if the model matches with the given trained model, then the emission and transition matrices are modified accordingly.
However if there is no match with the given model, then next model is taken for matching.

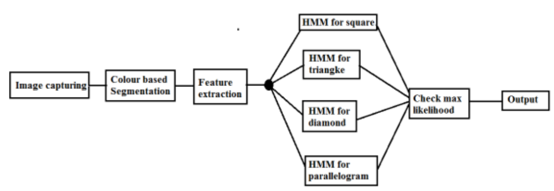

Figure 8. HMM flow diagram

4.4 RESULTS:

In this paper we analysed the algorithm for four shapes which are square, rhombus, parallelogram and triangle. The analysis includes test cases from four different users represented in the Figure 9.

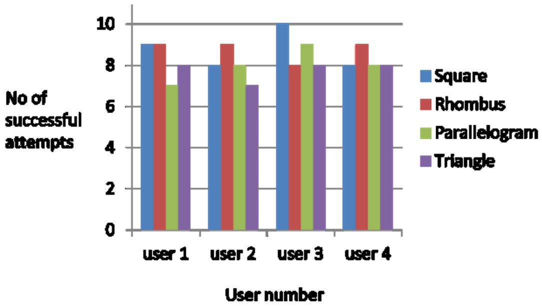

Figure 9. Analysis of Test Cases

The results obtained in the various test cases are summarized in the form of percentage accuracy for each shape in the following table 1.

Table 1. Percentage Accuracy

	Square	Rhombus	Parallelogram	Triangle
% accuracy	80	87.5	80	77.5

5. FUTURE SCOPE

The proposed system can be further developed to include different sign language gestures. This will become an interactive aid for people unable to speak. Thus they'll be able to communicate with other human beings, who are unaware of the sign language, as if they themselves were speaking. Hence our system will act as an interface between the sign language gestures and English words.

6. CONCLUSION

This paper proposes an automatic recognition system that can recognise geometric figures. The proposed system uses HMM for recognising the gestures. Further experiments would focus on larger array of geometric shapes, number and alphabets.

7. ACKNOWLEDGEMENT

We would like to acknowledge Mrs. Archana Ghotkar, Professor Pune Institute of Computer Technology, Pune for providing us this idea to work upon.

8. REFERENCE:

[1] Elmezain, M. ; Al-Hamadi, A. ; Michaelis, B, *Hand trajectory-based gesture spotting and recognition using HMM*, Image Processing (ICIP), 2009 16th IEEE International Conference on

[2] Elmezain, M. ; Al-Hamadi, A. ; Michaelis, B ,*A Hidden Markov Model-based continuous gesture recognition system for hand motion trajectory,* Pattern Recognition, 2008. ICPR 2008. 19th International Conference on

[3] Shrivastava, R. ; Dept. of Electron. & Commun. Eng., Maulana Azad Nat. Inst. of Technology.,Bhopal, India, *A hidden Markov model based dynamic hand gesture recognition system using OpenCV,* Advance Computing Conference (IACC), 2013 IEEE 3rd International

[4] Gaus, Y.F.A. ; Sch. of Eng. & Inf. Technol., Univ. Malaysia Sabah, Kota Kinabalu, Malaysia; Wong, F., *Hidden Markov Model-Based Gesture Recognition with Overlapping Hand-Head/Hand-Hand Estimated Using Kalman Filter*, Intelligent Systems, Modelling and Simulation (ISMS), 2012 Third International Conference on

[5] Moni, M.A. ; Dept. of Comput. Sci. & Eng., Jatiya Kabi Kazi Nazrul Islam Univ.,Bangladesh ; Ali, A.B.M.S., *HMM based hand gesture recognition: A review on techniques andapproaches,* Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference on

[6] M Elmezain, A Al-Hamadi, B Michaelis,*Hand gesture recognition based on combined feature extraction ,* Int. J. Inf. Math. Sci, 2010 [7]. TE Starner ,*Visual Recognition of American Sign Language Using Hidden Markov Models,* DTIC Document-1995

[7] R Shrivastava, *A hidden Markov model based dynamic hand gesture recognition system using OpenCV*, Advance Computing Conference (IACC), 2013 IEEE 3rd International Conference on

[8] J Yamato, J Ohya, K Ishii, *Recognizing human action in time-sequential images using hidden markov model*, Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92., 1992 IEEE Computer Society Conference on

[9] Bregler, C. ; Div. of Comput. Sci., California Univ., Berkeley, CA, USA,*Learning and recognizing human dynamics in video sequences*, Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on

[10] AD Wilson, AF Bobick*, Hidden Markov models for modeling and recognizing gesture under variation***,** International Journal of Pattern Recognition and Artificial Intelligence,Volume 15, Issue 01, February 2001

[11] Wilson, A.D. ; Media Lab., MIT, Cambridge, MA, USA ; Bobick, A.F.,*Parametric hidden Markov models for gesture recognition*, Pattern Analysis and Machine Intelligence, IEEE Transactions on (Volume:21 , Issue: 9)

[12] Eickeler, S. ; Fac. of Electr. Eng., Gerhard-Mercator-Univ. Duisburg, Germany ; Kosmala,A. ; Rigoll, G.,*Hidden Markov model based continuous online gesture recognition,* Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on (Volume:2)