Shivganga Udhan, Prof. Pravin Futane / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 3, Issue 1, January-February 2013, pp.289-293

Alphabet Recognition Using Hand Motion Track

Shivganga Udhan[*] , Prof. Pravin Futane[**].
ME student Sinhgad College of Engineering Pune[*],
Professor and HOD Sinhgad College of Engineering Pune[**].

Abstract

This paper describes method for recognition of alphabets from hand motion trajectory. This method uses Hidden Markov Model (HMM). Gesture recognition for alphabets is done in three main stages; preprocessing, feature extraction and classification. In first stage, preprocessing hand is detected using color information. After detection of hand, motion trajectory which is also called as gesture path will be determined by tracking the hand. The second stage, feature extraction gives pure path by enhancing gesture path it also determines the orientation between the center of gravity and each point in gesture path. This orientation vector gives discrete vector that is used as input to HMM. In the final stage alphabet is recognized by gesture path. Our method will recognizes alphabets and we are expecting more than 90% of recognition rate.

Keywords—Hand Tracking, Hidden Markov Model, Human Computer Interaction, Gesture Recognition

I. INTRODUCTION

Human-machine interfaces are becoming more important as computer technology is growing. Now a days, hand gesture have been receiving a lot of attention in the area of HCI as an alternative to conventional interface devices. The hand provides a natural means for communication, which is highly expressive and simple to manipulate. Now Current filed of research is Hand gesture recognition and a lot of work is being done on it. There are so many applications have been designed on it. Computer version hand based tracking is used to interact with computers. method can take benefit by using the Hidden Markov Model(HMM).This method is introduced many days before, but it is very popular now, due to the effective use in the area of gesture recognition.Some kind of Artificial intelligence is required for recognising the things from the gesture path or motion trajectory.

Barbara Resch have given a gentle introduction to Markov models and hiddenMarkov models (HMMs) and relates them to their use in automatic speech recognition. [7].Nianjun Liu have given several ways to initialize and train hidden markov model (HMMs) for gesture recognition[8]. Ankit Gupta, Kumar Ashis Pati used hand traking and finger tip detection to create a simple user interface to browse and edit photos.[5]Cha-Sup Jeong,Dong-Seok Jeong,Designed method based on contour information and fourier descriptors [9]. Hongwei Ying, proposed a method for the tracking of fingertip in scenes with complex background [10]. Jonathan Alon, Quan Yuan,introduces a Unified Framework For simultaneously performing spatial segmentation, temporal segmentation and recognition[11].

II. LITERATURE SURVEY

In earlier work, the effects of varying weighting factors in learning from multiple observation sequences is studied. Multiple-sequence Baum-Welch, Ensemble methods, and Viterbi Path Counting methods are compared on Synthetic and real data. This study determined that Viterbi Path Counting was the best and fastest method. However the problem of choosing the initial model for the training (re-estimation) process was discussed by M Elmezain, A. Al-Hamagi,G. Krell, S. El-Etriby, B. Michaelis[1].

A method to determine an alphabet from a single hand motion using HMM is described by M Elmezain, A. Al-Hamagi,G. Krell, S. El-Etriby, B. Michaelis[1].System overview of their work is given as follows:

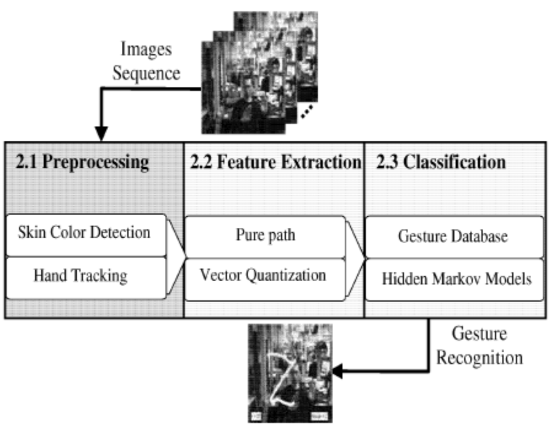

Fig 1. Recognition of Alphabet using HMM (Courtesy from [1]).

1. Preprocessing: point out hand location and track hand motion to generate its motion trajectory (gesture path).

2. Feature Extraction: enhance the gesture path to give pure path and then quantize the orientation to determine the discrete vector
3. Classification: the graphical gesture path is recognized using discrete vector and Left-Right banded model with six states.

A. Preprocessing

First stage in our method is preprocessing which will be divided into two steps. The first step is skin color detection and second step is hand localization and tracking.

B. Feature extraction

The feature extraction is a very important part in this method to recognize the alphabet gesture path. There are three basic features as location, orientation and velocity. The previous researches showed that the using of orientation feature is the best in terms of results. So, it is more reliable. A gesture path is spatio-temporal pattern which consists of centroid points (Xhand, Yhand). The gesture path is having some stable (i.e. non movable) points possibly starting and end point of gesture path, so enhance the gesture path to obtain a pure path as follows (Fig2.):

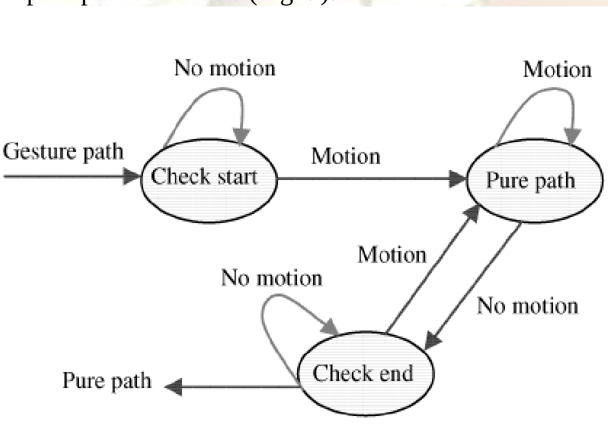

Fig2. Gesture Path Enhancement (Courtesy from [1]).

Firstly, the gesture path is entered to a check start state to see, whether there is no motion if yes, then it takes the next point else the pure path is begin generated. Then the pure path state continuously generates the pure path while the points of gesture path input move. When the point does not move, check end state is called. Finally, the pure path is generated from check end state when input gesture path is ended. Also while there are points in gesture path, this state perform a check to see if the point not moves then delete it and takes the next point, else return to the pure path state.

In contrast to the enhancement gesture path to obtain a pure path, our method is based on the angle(orientation) as a basic feature. Therefore, the orientation is based on (Xc,Yc) where Xc and Yc are the center of gravity of pure path and are determined by Eq. 1 and Eq. 2. Since the location of pure path for the same gesture according to the start point is different, we calculate the orientation between any point in pure path and center of gravity (fig.3).

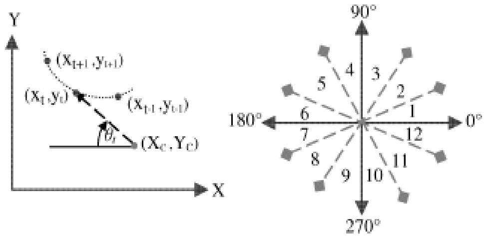

Fig. 3: Orientation and vector quantization range (Courtesy from [1]).

$$X_c = \frac{1}{n}\sum_{t=1}^{n} X_t \qquad (1)$$

$$Y_c = \frac{1}{n}\sum_{t=1}^{n} Y_t \qquad (2)$$

$$\theta_t = \arctan\left(\frac{Y_c - y_t}{X_c - x_t}\right) \qquad ; t = 1,2,\dots,n \qquad (3)$$

Where n is the length of pure path. In addition, the orientation θ_t determined according to eq. 3 where the orientation is divided by 30 in order to quantize the value of it from 1 to 12. Thereby, the discrete vector is obtained which is used as input to HMM.

C. Classification

Classification is the last stage in this method. Throughout this step, the pure path of hand graphical is recognized by using Left-Right Banded model with 6 states and building gesture database. According to this stage, Baum-Welch algorithm (BW) is used for training the initialized parameters of HMM to provide the trained parameters. Then the trained parameters and discrete vector are used as input to Viterbi algorithm in order to obtain the best path. By this best path and gesture database, the pure path is recognized. The following subsections describe this stage in detail [1].

1) Hidden Markov Models: Markov model is a mathematical model process where these processes generate a random sequence of outcomes according to certain probabilities. An HMM is a triple A = (A, B, Π) as follows:

- The set of states S {s1, s2, , SN} where N is a number of states.
- An initial probability for each state Π_i, i=l, 2,....N such that $\Pi_i = P(s_i)$ at the initial step.
- An N-by-N transition matrix A = $\{a_{ij}\}$ where a_{ij} is the probability of a transition from state Si to Sj; 1 <=i, j<=N and the sum of the entries in each row of matrix A must be 1

because this is the sum of the probabilities of making a transition from a given state to each of the other states.

- The set of possible emission (an observation) 0={o1,o2, . . on} where T is the length of pure path.
- The set of discrete symbols V {vl, v2, ..., vI} M represents the number of discrete symbols.
- An N-by-M observation matrix B ={b_{im}} where b_{im} gives the probability of emitting symbol Vm from state s1 and the sum of entries in each row of matrix B must be 1 because this is the sum of the probabilities of making a transition from a given state to each of the other states.

There are three main problems for HMM: Evaluation, Decoding and Training that can be solved by using Forward- Backward algorithm, Viterbi algorithm and Baum-Welch algorithm respectively[1]. Also, HMM has a three topology: Fully Connected (Ergodic model) where any state in it can be reached from any other state, Left-Right model such that each state can go back to itself or to the following states and Left-Right Banded (LRB) model that also each state can go back to itself or the following state only (Fig 4.).

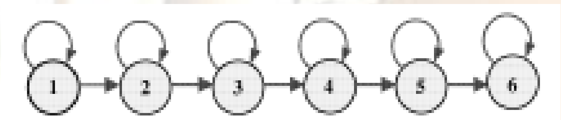

Fig 4. : Left-Right Banded model with 6 states (Courtesy from [1]).

2) Initializing parameters for LRB model: It will be more convenient to explain, why we use left-Right Banded model with 6 states before describing the initialization of HMM parameters. Since each state in fully connected model has many transitions rather than LRB model, the structure data can be losing easily. Moreover, LRB model is restricted and simple for training data that will be able to match the data to the model. In addition, the discrete vector contains a single sequence of codehook from 1 to 2 in this method. For that reason the LRB model is preferred than left-right model. About the number of steps considers the number of segmented part that is contained in graphical pattern when we represent it. For example, "L"graphical pattern contains two segmented parts, thus we need only 2 state for it, while "G" and "E" patterns need 5 and 6 states respectively. Therefore the no of states is 6 nearly for all alphabets. For this reason we selected the left-right banded model with 6 states. There is no doubt that, a good parameters initialization for HMM (A, B, Π) gives a better results. The matrix A is determined by Eq. 5 and it depends on the duration time d of states for each alphabet such that d is defined as;

$$d = \frac{T}{N}$$

Where T is the length of pure path and N represents the number of states that has a value 6 in this method.

$$A = \begin{pmatrix} a_{ii} & 1-a_{ii} & 0 & 0 & 0 & 0 \\ 0 & a_{ii} & 1-a_{ii} & 0 & 0 & 0 \\ 0 & 0 & a_{ii} & 1-a_{ii} & 0 & 0 \\ 0 & 0 & 0 & a_{ii} & 1-a_{ii} & 0 \\ 0 & 0 & 0 & 0 & a_{ii} & 1-a_{ii} \\ 0 & a_{ii} & 1-a_{ii} & 0 & 0 & 1 \end{pmatrix}$$

Such that;

$$a_{ii} = 1 - \frac{1}{d}$$

The second important parameter is a matrix B that is determined by Eq. 7. Since HMM states are discrete, all elements of matrix B can be initialized with the same value for all different states.

$$B = \{b_{im}\}; \quad b_{im} = \frac{1}{M} \qquad (7)$$

Where i, m run over the number of states and the number of discrete symbols respectively. The third parameter in the HMM is the initial vector Π which takes value;

$$\pi = (1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0)^T$$

That is because we use 6 states as the maximum numbers of the segmented graphical alphabet and in order to guarantee that it begin from the first state as shown in Fig. 6.

3) **Baum-Welch and Viterbi Algorithm.**

After the HMM parameters are initialized, M Elmezain, A. Al-Hamagi,G. Krell, S. El-Etriby, B. Michaelis[1], used Baum-Welch algorithm to perform the training for initialized parameters where the inputs of this algorithm are discrete vector i.e. obtained from feature extraction stage and initialize parameter. This algorithm give us a new parameter estimation of vector Π , matrix A and matrix B. In the next step the viterbi algorithm takes the discrete vector, new matrix A and new matrix B as its input and gives us the best path. For doing that the initial state determined by product initial vector Π with associated observation probability bit. After this the best result route of the next step (t+1) is determined by taking the minimum probability that is derived from the product of previous state observation probability with its transition. Finally, by backtracking through the trellis, the best path is obtained by selecting the maximum probability state at time T as shown in fig 7. After the best path is determined, we call the database gesture for this path to recognize it from A to Z by using higher priority for comparing and choosing. The following steps demonstrate how Viterbi algorithm works:

1) Initialization:For $1 \le i \le N$,
   a) $\delta_1(i) = \pi_i . b_i(O_1)$,
   b) $\emptyset_i(i) = 0$,

2) Recursion: For $2 \le t \le T,\ 1 \le j \le N$,
   a) $\delta_t(j) = \max_i[\delta_{t-1}(i). u_{ij}]. b_j(v_t)$
   b) $\emptyset_t(j) = arg\max_i[\delta_{t-1}(i). a_{ij}]$

3) Termination:
   a) $P^* = \max_i[\delta_T(i)]$
   b) $q_T^* = arg\max_i[\delta_T(i)]$

4) Reconstruction: For t=T-1, T-2, …, 1.
   $q_T^* = \emptyset_{t+1}(q_{t+1}^*)$

III. SUGGESTED METHOD

Alphabets recognition process mainly consists of three stages as given in fig. 5 below:

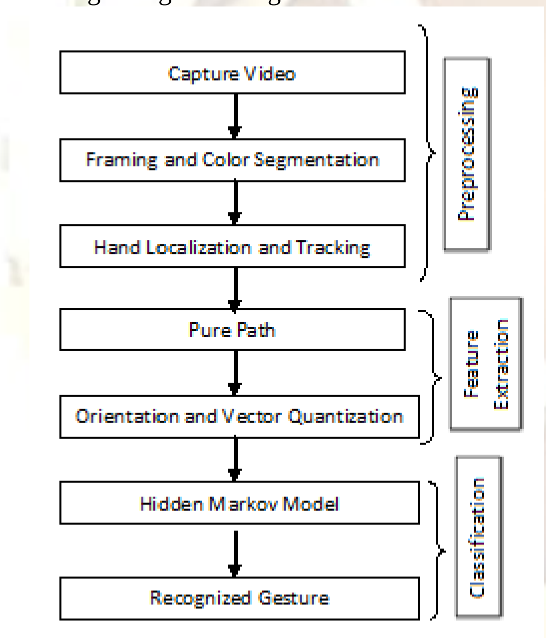

Fig. 5: System Overview.

The proposed system is self contained product this is basically a means of communication between physically disable people with computer or it can be replacement for input devices of computer systems. Main goal of this system is to accept the video of hand motion and to display appropriate alphabet identified from hand motion trajectory. In this method the gesture path is generated from hand motion trajectory by using HMM. This gesture path will recognise respective alphabets.

In our method we are hiding head of user form camera to reduce the complexity of the Segmentation and Localization module. Then we find the centroid if actual skin part (i.e. users hand) for each frame of input video. Then these centroids are joined to get gesture path (i.e. motion trajectory). After that we use Left-Right Banded model on gesture path to get pure path and thus build the gesture database.

IV. RESULTS AND DISCUSSIONS

These are some results which we got till now are shown in fig.6 below. Here are first and last frames i.e. starting and end points of gesture path for alphabet 'A', and centroids of hand in respective frames.

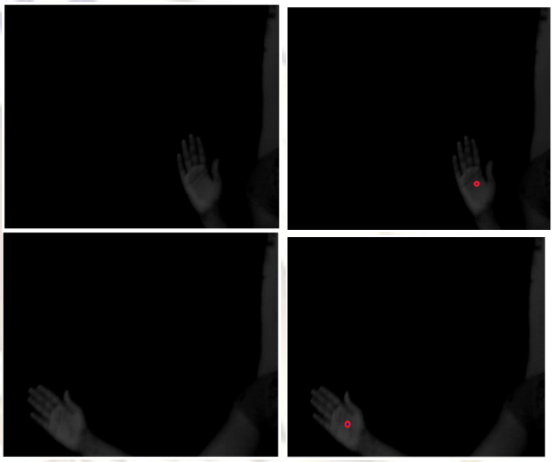

Fig. 6: Start and end point for 'A' and their centroids.

Possible gesture paths (Motion trajectory) for all 26 alphabets are as shown in fig.7 below. That means we are not considering the letter as it is but we can skip some part without any problem to reduce complexity of overlapping of gesture path. For example we can skip middle horizontal line from 'A' as shown in figure.

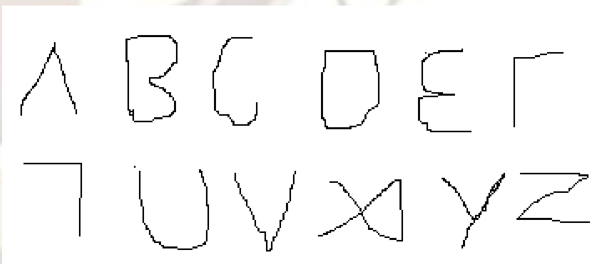

Fig. 7: Letter Gesture for A,B,C,D,E,F,T,U,V,X,Y,Z.

V. SUMMERY AND CONCLUSION

This paper presents a method for vision-based recognition of alphabets from A to Z by using Hand Motion Trajectory. It uses HMM for recognizing Alphabet. This method mainly consists of 3 main stages. The first stage is the preprocessing where hand is localized and tracked to produce gesture path. In the second stage i.e. feature Extraction, where the quantization and orientation is used to get the discrete vector. The final stage is classification, in this stage actual recognition of alphabet is done.

ACKNOWLEDGEMENT

The research in this paper was carried out at Sinhgad College of Engineering, Pune. So, special thanks to Head of Computer Department, Principal and Management of Sinhgad College of Engineering.

REFERENCES

[1] M Elmezain, A. Al-Hamagi,G. Krell, S. El-Etriby, B. Michaelis, "Gesture Recognition for alphabet from Hand Motion Trajectory Using Hidden Markov Models". 2007 IEEE International pp:1192-1197.

[2] N. Liu, B. C. Lovell, P. J. Kootsookos, R. I. A. Davis, W. 2004 "Model Structure Selection and Training Algorithm for a HMM Gesture Recognition System", International Workshop in Frontiers of Handwriting Recognition, pp.-100-106, October 2004.

[3] Lawrence R. Rabiner, Fellow, "A Tutorial on Hidden Markov Model and Selected Applications in Speech Recognition" Proceedings of the IEEE, vol. 77, No. 2, Feb 1989. Pp 257-286.

[4] Cha-Sup Jeong, Dong-Seok Jeong "Hand-written Digit Recognition Using Fourier Discriptors and Contour Information" 1999 IEEE TENCON.

[5] Ankit Gupta, Kumar Ashis Pati "Finger Tips Detection and Gesture Recognition" Indian Institute of Technology, Kanpur,India. November 12, 2009.

[6] Ming-Hsuan Yang and Narendra Ahuja "Recognizing Hand Gesture Using Motion Trajectories" Computer Vision and Pattern Recognition, IEEE Computer Society Conference on Jun 2002

[7] Barbara Resch , Erhard and Car line Rank "Hidden Markov Model". Signal Processing and Speech Communication Laboratory.

[8] Nianjun Liu, Richard I.A Davis,Brian C.Lovell and Peter J.Kootsookos "Effect of initial HMM Choice In multiple sequence training for gesture recognition" IEEE.

[9] Cha-Sup Jeong, Dong-Seok Jeong, 1999 "Hand-Written digit recognition using fourier descriptors and Contour information", IEEE TENCON.

[10] Hongwei Ying,Jiatao Song,Xiaobo Ren and Wei Wang "fingertip detection and traking using 2D and 3D information", 2008 IEEE Pp1149-1152

[11] Jonathan Alon, Quan Yuan "A Unified Framework for Gesture Recognition and Spatiotemporal Gesture Segmentation",2009 IEEE Computer So ciety Pp1685-1699.

[12] Ming-Hsuan Yang and Narendra Ahuja "Recognizing Hand Gesture Using Motion Trajectories" Computer Vision and Pattern Recognition, IEEE Computer Society Conference on Jun 2002.

[13] Nianjun Liu Brian et.al "Evaluation of HMM Training Algorithms for Letter Hand Gesture Recognition" Intalligent Real-Time Imaging and Sensing (IRIS) Group.Pp 648-651.

AUTHORS

First Author – Ms. Shivganga Udhan, ME Computer Networks Student, Sinhgad College of Engineering, Pune,

Second Author – Prof. P. R. Futane, Head of Computer Department, Sinhgad College of Engineering, Pune,