The Dynamic Hand Gesture Recognition Method

A Anoop[1] H Chidananda[2]
[1,2]Department of Computer Science & Engineering
[1,2]RYMEC, Bellary, KA -583101

Abstract— The human computer interaction system is an very much interactive and very much challenging task in nature and the hand gesture recognition system is one of the major and current interesting topic in the research of computer vision field. So here as we are humans we can communicate with other people as the same way we the humans we can communicate with machines. The human machine interaction can be implemented by using these methods. And here the interaction between the humans and the machines is became the primary objective in the human computer interaction (HCI) research field. Here there is rapid development in the human computer interaction field by gesture recognition and speech recognition the humans can interact easily with machines and conveniently without any difficulties. Here in this paper I am exploring the one of the method of gesture recognition in real time of human hand using hidden Markov model method (HMM). And this HMM is being used in dynamic hand gesture recognition systems. Here the bare hand is being used and no usage of any other gloves for hand detection.

Key words: Human Computer Interaction(HCI), Dynamic Hand Gesture Recognition (HGR), Hidden Markov Model(HMM), Skin Colour Detection And Computer Vision Methods

I. INTRODUCTION

Earlier the machines were running in vacuum tubes and later semi conductivity came into existence which made an evolutionary development in the field of computers. Now present super computers are being used in the application of banking such as usage of super computers in the ATM machines. Earlier keyboard was the only device which is being used for communication between the humans and the computers but later the invention of mouse which made an rapid development and it made still easy and convenient for communication of computers by clicking the mouse end user can easily communicate with machines. Here speech recognition is one of the rapid development in the computer vision field that is the machines can be communicated and can be controlled using the speech by humans best example is robots and presently this implementation is made in smartphones that is text to speech recognition application which is being developed and interfaced in the android smart phone. This application is being interfaced in the windows operating systems also and still its not yet being released and the application by name "CORTANA" which is an speech recognition help assistance for the humans. This is the power of computer vison in the field of computer science.

And this is the main aid tool for the physically handicapped people and this can be used in sign languages to communicate with handicap peoples. Here there is gesture recognition can be implemented in two ways one way is in static nature and the other way is dynamic in nature. Both are challenging task in nature in the implementation mechanism. Here there are two methods again for hand recognition that is one by using sensory gloves or secondly by using the computer vision concepts. In the usage of sensor gloves it measures the suitable angles and the spatial positions of the hand and fingers of hand. And this can recognise 28 number of static gestures in one character/second time by using 6 accelerometers. This method or approach is not efficient in the real time application. And here the glove should be wear for detection of hand and its not natural so people wants natural detection without using any type of gloves in hand for hand recognition.

Here natural HCI should be convenient and more reliable and natural in nature. For daily usage the sensor gloves is not an good option. So in computer vision technique one or more number of cameras are being used for capturing the images for hand gesture recognition.

By using the backward references of the Haar transform method for gesture recognition, we came across an another proposed method or algorithm for gesture recognition is firstly it separates the hand region in the complex background images by measuring the entropy of the adjacent frames. These hand gestures then recognised by using the appropriate centroid profile. And the Misrecognitions are being caused by the hand recognition gestures with similar spatial features and angles therefore here the number of gestures recognised could be limited in nature.

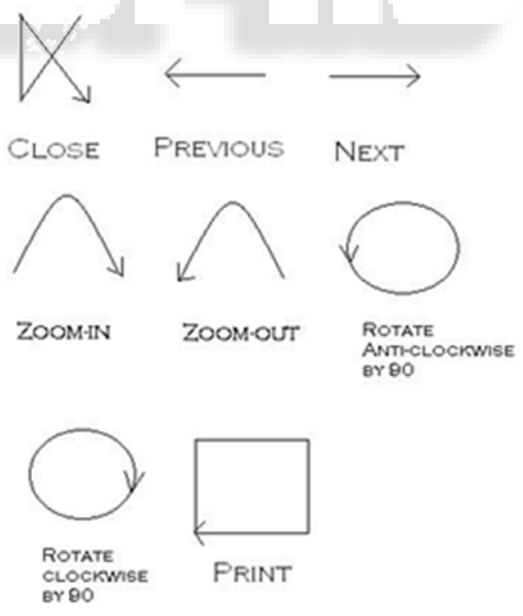

Diagram 1: Various Gestures Used in the System

The above diagram represents the various gesture movements of hand. By referencing the above concept called Haar transformation for detecting the hand gestures we can come across an another approach for an robust and effective computer vision system. Here the proposed system is the skin colour detection technique. Here the hands were detected by the skin colour detection mechanism. The

problem which causes the hand orientation in the image can also be overcome and solved by using the idea of axis of elongation approach. This approach helped in immense by keeping the data bases smaller by standardizing the hand gestures in the recognition using the fixed orientation in nature. Here to make the facility of the searching process the code word scheme mechanism technique is being used here. This experimentation results shows the good hit rate of scores in the recognition of the hand gestures.

Here actually the different and actual process in the hand gesture recognition is the hidden Markov model (HMM). In this paper the graphic editor tool which recognises twelve dynamic gestures and five static gestures was being developed here. Gesture recognition was being facilitated by structural analysis for static gestures and for dynamic gestures hidden Markov model (HMM) is being implemented.

Here this HMM models have intrinsic properties which will make attract for an option of gesture recognition and also it will explicit the segmentation is not all necessary for training the model or recognition of the model.

I referenced the thesis proposed by the jinly and the tianding "hand gesture trajectory based on HMM" theses this main theme of this thesis is that we can model an spatio-temporal information in an natural way. In order to differentiate the unidentified gestures an modified threshold model was proposed. The hand is being separated by the or from the complex background by using the technique called skin colour detection method by firstly converting the RGB based colour pixels into YCbCr colour model and being defining the suitable range for skin colour detection.

In this paper I am proposing the Hidden Markov model (HMM) method for recognition of dynamic hand gestures against the static background method. The technique is being in addition to the HMM model, here the representation of hand gestures which is being represented by using the adjacency matrix and also I am using the idea of the principal axis through the centroid of the hand for standardizing the hand gestures by reducing the image sizes of the data set. I am here will be working with the HSV colouring model.

Here the section 2 gives the system overview and the section 3 determines the segmentation based on skin colour detection and section 4 the usage of HMM method for hand gestures dynamically and section 5 the implementation and section 6 represents the conclusion.

II. SYSTEM OVERVIEW

The following block diagram represented below represents the approach we will be implementing in our system

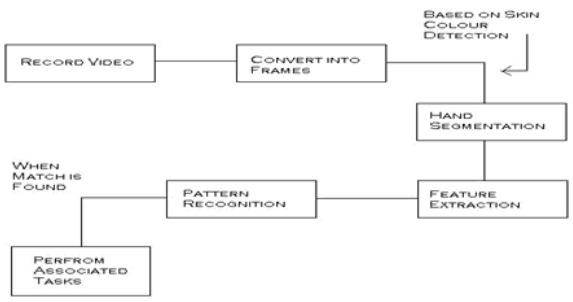

Diagram 2 : Block Diagram of the System

The state transition diagram of the system is being represented below here the operations of each state is being shown below

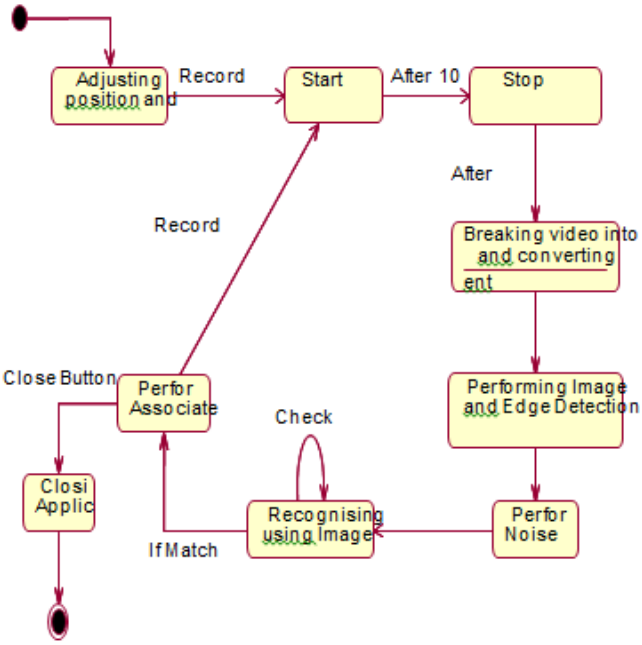

Diagram 3: State Transition Diagram of the System

III. HAND DETECTION USING SKIN COLOUR DETECTION

Here the RGB colour model which it contains both the colouring and brightness properties which is vulnerable in nature for changes in the background illumination and environment. For ensuring that this doesn't affect in the detection of skin colour we will be converting this RGB model into HSV colouring model. Because in this the latter is more similar to the human skin colour perception. The skin colour in channel H characterised by the values in between 0 to 50 and in the channel S it is in between 0.23 to 0.68 for Asian skin colour.

Diagram 5: Original Image

After changing or converting from RGB to HSV colouring model we get the following image shown below in diagram 6

Diagram 6: Image Converted into HSV

The next image we will be obtaining is that its an intermediate image obtained by immediately after setting all the pixels which are being fallen into our skin colour range that is 255 white pixel values and non-pixel is being set to 0 that is black pixel value. And although this image as some certain noise in it and it should be filtered or noise should be removed. Here some pixels which are not the part of the skin but also will be fallen in the given range and this is to be eliminated here. So to eliminate these pixels morphological filters or operations are being implemented and performed here.

The final image obtained using morphological operations performed shown below in diagram

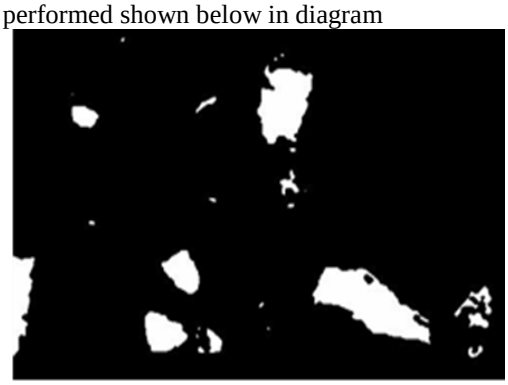

Diagram 7: Noise Removed for the Original Diagram

Here in the above diagram only the skin pixels are being represented with white pixels. To convert from RGB to HSV colouring model (normalized RGB values to be assumed) first find the maximum and minimum values of RGB triplet. Saturation S is given by the equation below

S = (Max – Min) / Max (eq:1) and the value of 'V' is given V = Max , the hue H is being Calculated as follows that is first calculate the values of 'R''G''B' v'R' = (Max-R) / (Max – Min)

'G' = (Max-G) / (Max – Min)

'B" = (Max-B) / (Max – Min) (eq :2)

If the saturation S value is 0(zero) then hue value is undefined (the colour is not having any hue hence it is monochrome in nature) otherwise it is

If R = Max and G = Min then H = 5+B' (eq : 3)

Else if

If R = Max and G6 = Min then H =1 – G' (eq : 4)

Else if

If G = Max and B = Min then H = R' + 1 (eq : 5)

Else if

If G = Max and B6 = Min then H = 3 – B' (eq : 6)

Else if

If R = Max then H = 3 + G' (eq : 7)

Otherwise

H = 5 – R' (eq : 8)

after this the hue H values is being converted into degrees by multiplying 60 by giving the HSV values with S and V in between 0 and 1 and H value is given between 0 and 360.

IV. HIDDEN MARKOV MODEL (HMM)

Consider an human being who is being sitting inside one room and imagine that he has three coins with him. And he is being tossing those three coins by sequence which that sequence is only known to him. And we are being positioned outside the room and are being showing by means of an display (which is also being placed outside the room also) the outcomes of the person of flipping the coins is "HTTHTHTTHHTHHTHH" this is being called as observation sequence. Here actually we don't know how the coins is being tossed with what sequence and also we don't know the bias of the person coins. To know the impact on significance of the bias of coins on the outcome let us imagine that the third coin is being highly biased to generate tails. Here now if all the coins are being tossed with equal probability then it would be naturally expectation made that the output will be having more tails than heads. So consider further that the probability of moving from first or second coin (state) to third coin (state) is zero. Here now if we started to start tossing the coins from the first and second coins then the output sequence will be generating more sequence of tails because of the transition probability between the coins/states and also the initial state. These three sets namely the set of individual bias of the three coins , the set of transition probability from one coin to the next coin and the set of initial probabilities characterizes what is being called as the Hidden Markov Model(HMM).

This HMM Model is being specified as shown below

The set of States S = { S1, S2,…….SN's} (corresponding to N possible gesture condition

Above)

And the set of parameters : λ = { Π, A, B}

Transition probabilities : A = { aij = P(qj) at t+1 | qi at t)},

Where P(a | b) is the conditional probability of a given b, t = 1,2,…..T, is time and Qi in Q.

Informally A is the probability that the next state is 'qj' that the given current state is 'qi'

Observation symbols O = {Ok} , k = 1,2,….M.

Emission probabilities B is B = {bik = bi(Ok) = P(Ok) / qi)} where Ok in O

Informally B is probability that the output Ok is given that the current state is qi.

Initial state probabilities are : Π = { pi = (p(qi) at T =1)}

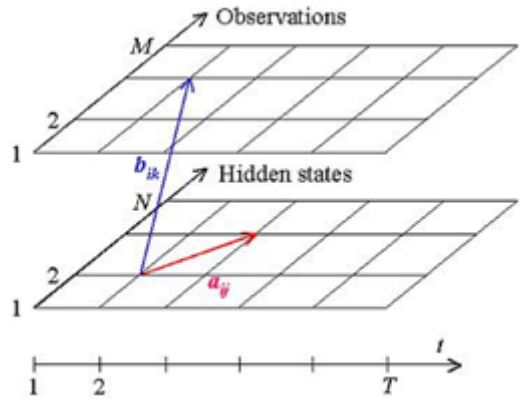

Diagram 9: Observations and Hidden Sequences

This HMM allows the transitions from any emitting state to any other emitting state and it is being called as an "Ergodic HMM" . The other extremes where the HMM only the transition will go from one state to itself or to an unique follower is called "Left-Right HMM".

There are important three canonical problems that o be solved by HMM are

1) The model parameters are being given the computation of the probability of that particular

output sequence. This problem is being solved by using the backward and forward algorithms.

2) The model parameters are being given find the most likely sequence of the hidden states which could have been generated in the given output sequence. This can be solved by the Viterbi algorithm and also usage of posterior decoding technique.
3) The output sequence is being given find most likely set of state transitions and the output probabilities. This can be solved by the Baum – Welch algorithm.

A. Forward Algorithm:

Let at(i) be the probability of the partial observation sequence
Ot = { O1,O2……,O(t) to be produced by the possible state sequences that end at the ith state at(i) = P(O(1),O(2),……O(t) | q(t) = q(i) (eq : 9)

Then the unconditional probabilities of the partial observation sequence is the sum of at(i) for all N states. This forward algorithm is the recursive algorithm for calculation of the at(i) for the observation sequence of the increasing length 't'. firstly the probabilities of the single symbol sequences is being calculated as the product of the initial i-th state probability and emission probability of the given symbol O(1) in the i-th state. Then here the recursive formula is being applied. Assume here we are being calculated the value of at(i) for some t. to calculate the value of a (t+1) (j), we will multiply every at(i) by the corresponding transition probabilities from the i-th state to the j-th state. Here the summing of all the products over all the states and then multiply the result by the emission probability of the symbol O(t+1). Iterating the procedure we can calculate the value of aT(i). Summing them over all the states we will be getting the required probability.

In the similar manner we can implement the symmetric backward variable βt(i) as the conditional probability of the partial observed sequence from O(t+1) to the end to be produced by all the state sequences that start at i-th state
Bt(i) = P(O(t+1), O(t+2), …..O(T) | q(t) = q(i)) (eq : 10

This backward algorithm calculates backward variables recursively in nature. And it will be going backward along the observation sequence. Here both forward and backward algorithms are being frequently used in finding the optical state sequence and estimating the HMM parameters.

B. Viterbi Algorithm:

It chooses the best state sequence that maximise the likelihood of the state sequence for the given observation sequence.

Let δt(i) be the maximum probability of the state sequences of the length t and the end in sate i and producing the t first observations for the given image
δt (i) = Max{P(q1),(q2),….q(t-1) ; o(1),o(2),…..o(t) | q(t)= q(i)} (eq : 11)

This Viterbi algorithm is an dynamic programming algorithm that uses the same type of scheme as the forward algorithm except e value

1) It uses the maximization in place of summation at the recursion and the termination steps.
2) It will be keep tracking the arguments that maximizes δt (i) values for each t and i storing them in N by T matrix ψ. This matrix is being used to retrieval of the optimal states sequences at the backtracking step.
3) Baum – welch Algorithm: This algorithm is an very important in the HMM model for decoding the observation sequence in such a way that if this observation sequence having many charecterstics which is to be in similar one it should be encountered later and it should be identify first. There are two methods where identification can be made they are
   - The segmental K-means algorithm using
   - The Baum – Welch re-estimation formulas usage

Here we are being used the second method for identification and here it can be shown below
ξ t(i,j) = the joint probability of being in state qi at time t and state qj at time t=t+1,
given the observed sequence below
ξ t(i,j) = P(q(t) = q(i) ,q(t+1) = q(j) | O, Λ (eq:12)
Therefore we get = ξ t(i,j) =

$$\xi_t(i,j) = \frac{\alpha_t(i)\, a_{i,j}\, b_j(o(t+1))\, \beta_{t+1}(j)}{P(O \mid \Lambda)} \quad \text{(EQ : 13)}$$

Then the probability of the output sequence is being represented as below

$$P(O \mid \Lambda) = \sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t(i)\, a_{i,j}\, b_j(o(t+1))\, \beta_{t+1}(j) = \sum_{i=1}^{N} \alpha_t(i)\, \beta_t(i) \quad \text{(EQ : 14)}$$

The probability of state in being in state qi at t time is shown below

$$\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i,j) = \frac{\alpha_t(i)\, \beta_t(i)}{P(O \mid \Lambda)} \quad \text{(EQ : 15)}$$

V. CONCLUSION

This hidden Markov model is one of the important tool which is being used for dynamic hand gestures recognition in the real time. And also the idea of standardizing the axis through the centroid will reduce the size of the data base. The accuracy will be high in our proposed system.

VI. FUTURE WORKS

This methodology can be taken into consideration by taking into account of the effect of the speed of hand movement and the system can be implemented for the both hands can be able to recognize both the hand gestures .

VII. DISADVANTAGES

1) Misrecognition in case that the background as the elements that has to reassemble the human skin.
2) The other factors such as the velocity of the hand movement and the orientation and low background illumination will make an low score rate of the systems accuracy.

REFERENCE

[1] Wing Kwong Chung, Xinyu Wu, Yangsheng Xu, "A realtime hand gesture recognition based on Haar wavelet representation", Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference, pp. 336 – 341, 22-25 Feb. 2009.
[2] Byung-Woo Min, Ho-Sub Yoon, Jung Soh, Yun-Mo Yangc', Toskiaki Ejima, "Hand Gesture Recognition Using Hidden Markov Model", pp. 305-333.
[3] Jinli Zhao and Tianding Chen, "An Approach to Dynamic Gesture Recognition for Real-time Interaction", ISNN 2009, pp. 369-377.

