2013 5th International Conference on Computational Intelligence and Communication Networks

Fuzzy L Membership Function Based Hand Gesture Recognition for Bharatanatyam Dance

Sriparna Saha, Lidia Ghosh, Amit Konar
Electronics & Telecommunication Engineering Dept.
Jadavpur University
Kolkata, India
sahasriparna@gmail.com, lidiaghosh.bits@gmail.com, konaramit@yahoo.co.in

Ramadoss Janarthanan
Computer Science & Engineering Dept.
TJS Engineering College
India
srmjana_73@yahoo.com

***Abstract*— This paper presents a method for automatic hand gesture recognition of 'Bharatanatyam' dance using Fuzzy L membership function based approach. Here, a 3-stage system has been designed. In the first stage, the hand of the dancer from background is isolated using Texture based segmentation and thus the contour of the hand is extracted by using Sobel edge detection technique. In the next stage, centre point of the boundary is located and based on this eight spatial distances are calculated. These distances are normalized by dividing the maximum distance value. In the final stage, fuzzy L Membership values are calculated for each distance and matching of an unknown hand gesture is done with the known hand gestures from the database based on L fuzzy membership function. The proposed algorithm gives overall an accuracy of 85.1% and timing complexity is 2.563 sec in an Intel Pentium Dual Core processor running Matlab R011b for each hand gesture.. This simple yet effective code is very useful for e-learning of 'Bharatanatyam' dance.**

Keywords- boundary extraction; fuzzy L membership; similarity function; spatial information; similarity function

I. INTRODUCTION

Gesturing provides an effective and powerful way of communication among human beings. It is so much implanted in our communication that we often use gesturing while speaking. Human use gestures as a complementary way to communicate for expressing their ideas. So, it is essential for computing devices to possess the ability to recognize meaningful gestures to achieve natural interaction with people. Hand gesture recognition has become an obvious way of making a suitable interface between machines and their users.

'Bharatanatyam' is the most popular Indian classical dance form, performed by male and female dancers all over the world. It is said to be originated in Thanjavoor of Tamil Nadu, South India. In Bharatnatyam, a dancer performs to a classical Carnatic music, creating a parallel kinetic poetry in movement, registering subtle expressions on the face and the entire body reacts to the emotions, evoking sentiments in the spectator for relish, the 'rasa'. This traditional dance comprises of elegant and graceful hand movements. Like normal hand gestures, 'Bharatanatyam' hand gestures are expressive and meaningful. Here we are concentrating on 'Asamyukta Hastas' (single hand) hand gestures. The Natyashastra mentions about 28 single hand 'mudras'. 'Katarimukham' hand gesture is used to indicate scissors, death, lightning etc, whereas 'Bramaram' denotes bee, parrot or some other birds.

In this paper we focus our attention to a simple system development for promoting e-learning of 'Bharatanatyam' dance. Besides the altruistic aspirations of e-learning, it is a flexible and cost effective way to learn in today's fast paced and fickle economic climate.

Another aim of our research is to establish more elegant way of human-computer interaction. In recent years, interactions between human and computer are performed through keyboard, mouse or many other Haptic devices. However, these tools are not flexible enough in some situations like interacting with physically challenged people or manipulating objects in a virtual environment. An automatic hand gesture recognition technology eliminates the need for physical contact between operator and machine, allowing the operation of complex machines using only a series of fingers and hand movements.

A system for Hand gesture recognition of 'Bharatanatyam' dance making use of orientation filter [1] and connectivity graph for 'Asamyukta Hastas' has been designed here. However, that algorithm has two major problems. To accentuate the hand gestures, the fingertips of the hands are coated with red dye, when 'Bharatanatyam' dance is actually represented. Since in [2], boundary of the hands is extracted by using skin color based segmentation, the red tips of the fingers are erased after segmentation. Second problem is that eight hand gestures are not identified by [2]. In this paper, we have proposed an algorithm dealing with these two problems. Here we have chosen an appropriate alternate of skin color segmentation i.e. texture based segmentation here to avoid the first problem. The problem is solved by taking eight spatial information bout centre of the hand gesture. Hand gesture recognition has been a research topic for many years. The authors in [3] prepared a survey on gesture recognition. There are several approaches of Hand gesture recognition based on hidden Markov model [4], neural network [5], particle filter [6], graph theory [7], etc.

In this proposed algorithm, we have designed a three level system. At the first level, the hand gesture is separated

978-0-7695-5069-5/13 $26.00 © 2013 IEEE
DOI 10.1109/CICN.2013.75

out from the background using hand texture. Then the boundary of the hand is detected using Sobel edge detection technique. At the next level, centre of the boundary is calculated and eight spatial distances starting with 00 and with 450 changes in slope are noted. The distances are normalized by dividing with the maximum distance value. At the final stage, unknown hand gestures are matched with the hand gestures from the database using the chain codes by similarity function. This algorithm is implemented to deal with translation invariant hand gestures. The average timing complexity of the proposed algorithm is 2.563 sec in an Intel Pentium Dual Core processor running Matlab R011b for each static image. We have achieved 85.1% overall accuracy.

In this paper, we provide a method for hand gesture recognition of 'Bharatanatyam' dance for e-learning and human computer interaction applications. Section II includes the preliminary tools of the proposed algorithm, while section III gives the block diagram. Section IV demonstrates result of study and analysis are elaborated. Finally, section V concludes with some idea of future work.

II. 3 – Stage System

A. First Stage

In the first stage, texture based segmentation and Sobel edge detection technique is implemented to extract the boundary of the hand gesture from the background.

- Texture based Segmentation

The primary step of gesture recognition is the identification the hands of the dancer and the segmentation of the corresponding image regions. The importance of this segmentation lies in the isolation of the required data from the image-background before accessing with other relevant stages. Here, texture based segmentation [8] [9] [10] has been utilized by entropy calculation for hand detection and elimination of unnecessary information about the background. This technique usually detects hand texture by employing knowledge acquired from the neighbourhood of each pixel in original image. The segmented result may contain a rather noisy boundary that would also contain huge topological errors like holes. For this reason, morphological filling of holes are done to capture actual patterns of the data, while leaving out noise.

- Sobel Edge Detection

For Sobel edge detection [11] [12] technique, a 3×3 neighbourhood is taken into account as explained in Fig. 2. Partial derivatives are calculated with constant c using the following two equations

$$s_x = (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6) \tag{1}$$

$$s_y = (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4) \tag{2}$$

Magnitude, M is measured by (3)

$$M = \sqrt{s_x^2 + s_y^2} \tag{3}$$

a_0	a_1	a_2
a_7	[i,j]	a_3
a_6	a_5	a_4

Figure 1. Labelling of neighbourhood pixels

After the Sobel edge detection [13], in some part of the boundary, noise creeps in. Various type of noise removal schemes are there, each having their merits and demerits. Some of these are based on frequency domain technique, while the others on spatial domain technique. Here, we have incorporated a spatial domain technique, to smooth the edges. This technique involves morphological shrink operation to one value. Fig. 2 explains the extraction of boundary of the 'Chaturam' hand gesture. The lines corresponding to the edges have become thicker in some places due to the increased smoothing of Sobel operator, so to reduce the threshold output of the Sobel operator to lines of a single pixel thickness, while preserving the full length of these lines; morphological shrink operation is carried out.

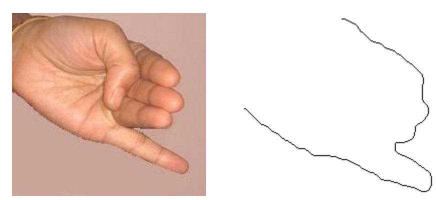


Figure 2. RGB and extracted boundary image of'Chaturam' hand gesture

B. Second Stage

The second stage comprises with centre point [14] [15] and the eight spatial information calculations.

- Locate Centre Point

For calculation of centre point, we start to record all the boundary pixel co-ordinates from any end point. After tracking of all the co-ordinates, the minimum co-ordinate value along X axis (X_{min}) and maximum value (X_{max}) are noted. So the centre point (X_c) along X axis is obtained using (4).

$$X_c = X_{\min} + \frac{X_{\max} - X_{\min}}{2} \tag{4}$$

For calculation of centre point along Y axis, we are keeping in mind that Y axis for image is opposite to Y axis as shown in Fig. 3. Thus the minimum value (Y_{min}) and maximum value (Y_{max}) are reversed. The centre point along Y axis (Y_c) is calculates using (5) as shown in Fig. 3.

$$Y_c = Y_{\min} + \frac{Y_{\max} - Y_{\min}}{2} \tag{5}$$

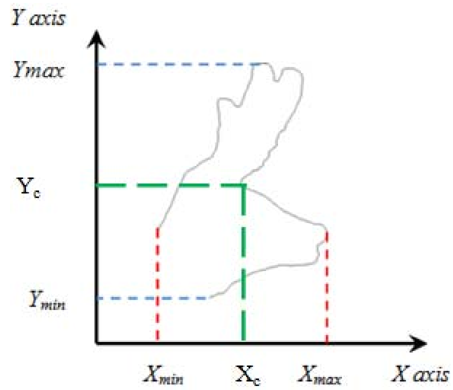


Figure 3. Centre point calculation for 'Hamsayam' hand gesture

- Calculation of Spatial Infomarmation

As shown in Fig. 4, if the co-ordinate of centre point is [i,j], then the 8 neighbor of that point along 0^0 angle is [i+1,j]. The axes i and j for image is also shown in Fig. 4. Now if we concentrate on point [i+1,j], then its corresponding neighbor along 0^0 angle is [i+2,j]. The loop goes on until we found the boundary pixel along 0^0 angle. Then that boundary pixel co-ordinate is noted and the Euclidean distance between boundary co-ordinate and centre of the image is taken as the spatial distance value along 0^0 angle.

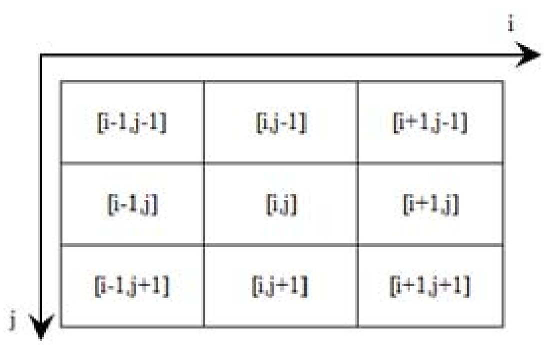


Figure 4. 8 neighbors

For obtaining spatial distances along 45^0, 90^0, 1350, 180^0, 225^0, 270^0, 315^0 angles, we are moving towards [i+1,j-1], [i,j-1], [i-1,j-1], [i-1,j], [i-1,j+1], [i,j+1] and [i+1,j+1] directions according to 8 eighbor principle. Let the boundary pixel along any direction is (x_b,y_b). So the Euclidean distance of the boundary pixel with respect to centre point (x_c,y_c) is obtained by (4). All the spatial distances are normalized by dividing them by the maximum distance value.

$$\text{Euclidean Distance} = \sqrt{(x_c - x_b)^2 + (y_c - y_b)^2} \tag{6}$$

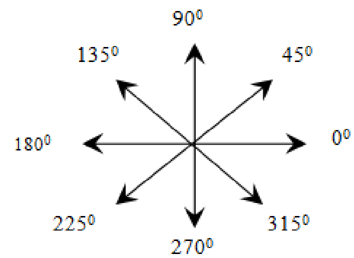


Figure 5. Eight spatial directions

C. *Third Stage*

Based on fuzzy L membership values, matching of unknown hand with respect to unknown 28 primitives are done in this stage.

- Fuzzy L Membership Value

Fuzzy L membership [16] [17] curve follows by Fig. 6. The membership value (µunknown) is 1 for 0 to α and it started decreasing to 0 for the next β values. Here α and β values are taken 0.07 and 1.00 empirically. A variable flag is calculated based on the absolute difference between normalized distance of unknown ($N_{unknown}$) and known (N_{known}).

$$\alpha = 0.07 \tag{7}$$

$$\beta = 1.00 \tag{8}$$

$$flag = abs(N_{unknown} - N_{known}) \tag{9}$$

where abs returns the absolute value.

$$\mu_{unknown} = \begin{cases} 1 & if \quad flag < \alpha \\ \left(\dfrac{\beta - flag}{\beta - \alpha}\right) & if \quad \alpha \le flag \le \beta \\ 0 & if \quad flag > \beta \end{cases} \tag{10}$$

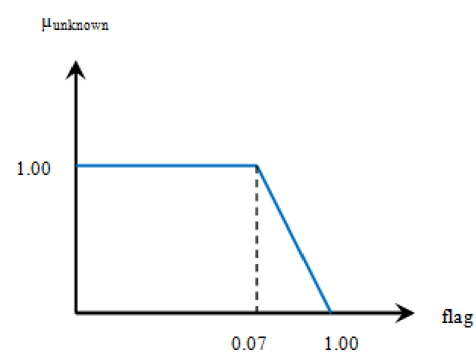


Figure 6. Fuzzy L membership curve

- Matching of Hand Gestures

For matching of unknown hand gesture with respect to 28 hand primitives, we obtain 28 set of different membership values ($\mu_{unknown}$) when comparing normalized distances of unknown and known hand gestures using (10). For matching with each known primitive eight membership values are produced. These eight membership values are summed and total 28 summations of membership values are obtained. The result is that known hand gesture for which maximum membership value summation is obtained using (12).

$$Sum_{unknown} = \sum_{1}^{8} \mu_{unknown} \tag{11}$$

$$SF = \arg\max \underset{p=28}{\forall} \left[Sum_{unknown} \right] \tag{12}$$

where SF is the similarity function, which returns argument based on summation of membership values. So the proposed algorithm easily identifies any unknown gesture.

III. Block Diagram of the Proposed Work

Fig. 7 shows each step of the proposed algorithm elaborately.

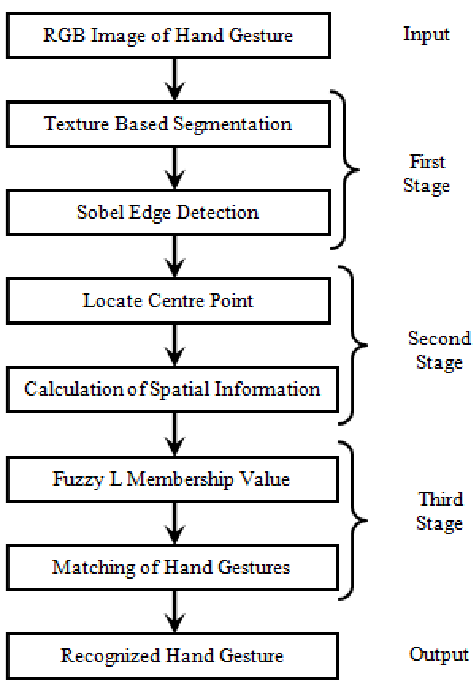


Figure 7. Block Diagram of the proposed algorithm

IV. Result of Study and Analysis

In order to achieve the human-computer interaction, the detection of human gestures or emotions in an image, becomes a crucial pre-processing of system design. The segmentation problem classifies an input image in two groups: skin and non-skin pixels. If the pixel values in a region are within a specific range, then the region is considered as skin region and the others are rejected as non-skin pixels. In general, when a 'Bharatanatyam' dancer actually performs the dance, his or her fingertips and middle of the palm are designed with red dye. Now, if skin color segmentation is used to isolate hand gesture from the back ground eliminating the unnecessary information as described in [2], it can be shown that the portion of the hand, coated with red color is not taken as the segmented output. Fig 7(ii) illustrates the skin color segmentation of 'Shukathundam' hand gesture and Fig. 7(iii) shows that after applying the edge detection technique, proper boundary is not obtained. The texture based segmentation on the other hand, deals not only with one pixel but also with the positional relation among pixels. Its basic principle is based on entropy calculation. If texture based segmentation operation is performed on the same input RGB image, depicted in Fig. 7(i) with threshold value 0.7, it extracts the exact boundary of the hand gesture as shown in Fig. 7 (vii). Thus, the texture based approach reduces the percentage of misclassification of the skin pixels with respect to the skin color based approach and thus gives a better result. The following Fig.7 explains the above mentioned process.

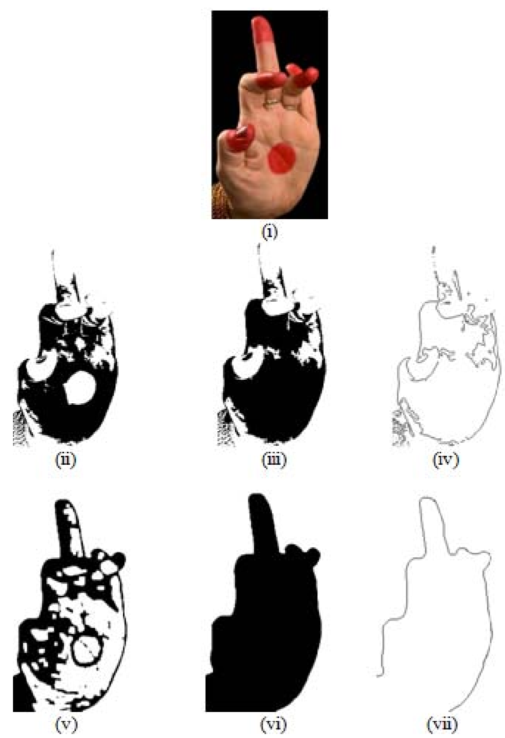
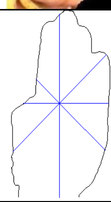
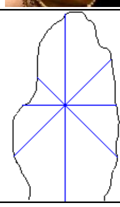


Figure 8. Comparison between skin color and texture based segmentation

Table I elaborately explains the matching of 'Pathakam' hand gesture for known and unknown images. First row gives the RGB image for unknown and known images, while the second row shows the corresponding extracted boundary images with eight spatial directions. Actual eight spatial distances and the normalized values (i.e., dividing by maximum distance value) is given in the next two rows. Using (9), flag value is calculated and based on the flag value fuzzy L membership value using (10) are produced. Summation of all membership value is 4.0398. This summation value is maximum for 'Pathakam' hand gesture, thus correct hand gesture is recognized using our proposed work.

TABLE I. COMPARISON OF KNOWN AND UNKNOWN HAND GESTURES FOR 'PATHKAM' HAND GESTURE

Features	Unknown Image	Known Image
RGB Image		
Boundary with Spatial Information		
Actual Distances	30 39 53 16 25 43 52 37	35 43 61 23 26 47 65 40
Normalized Distances	0.5660 0.7358 1.0000 0.3019 0.4717 0.8113 0.9811 0.6981	0.5385 0.6615 0.9385 0.3538 0.4000 0.7231 1.0000 0.6154
flag	0.0276 0.0743 0.0615 0.0520 0.0717 0.0882 0.0189 0.0827	
Membership Function Values	1.0000 0.9954 1.0000 1.0000 0.9982 0.9804 1.0000 0.9863	
Summation of Membership Values	7.9602	
SF	19	

V. CONCLUSION AND FUTURE DIRECTIONS

The accuracy rate achieved by this method is much better than that provided by the previous work [2]. Here, the proposed algorithm gives us 85.1% accuracy, which is also very much adoptable in e-learning purpose. Moreover, an average computation time of 2.563 sec also estimating the complexity of the dance hand gestures is highly efficacious for each input RGB image in an Intel Pentium Dual Core processor running Matlab R011. Again, the above mentioned procedure is very much cost effective as the input images required for this purpose can be taken only using a single static camera.

However, this proposed algorithm also has some limitations. This algorithm is suitable for an image possessing simple background, whether it should be much more improved for an image having complex background. Here as we have used simple background for hand gesture recognition, an accuracy rate of 85.1% is achieved easily. Another problem lies in the use of cropped images, whether in reality, most of the images of dancers taken by a static camera contain the whole body as an input. These shortcomings can be overcome by using practical filtering approach for detection of hand region in a video. Thus, modeling of a fully automatic hand posture recognition system still remains a challenging work. Our final goal is to build such a system.

REFERENCES

[1] W. T. Freeman and M. Roth, "Orientation histograms for hand gesture recognition," in *International Workshop on Automatic Face and Gesture Recognition*, 1995, vol. 12, pp. 296–301.

[2] D. Hariharan, T. Acharya, and S. Mitra, "Recognizing hand gestures of a dancer," in *Pattern Recognition and Machine Intelligence*, Springer, 2011, pp. 186–192.

[3] S. Mitra and T. Acharya, "Gesture recognition: A survey," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 37, no. 3, pp. 311–324, 2007.

[4] H.-S. Yoon, J. Soh, Y. J. Bae, and H. Seung Yang, "Hand gesture recognition using combined features of location, angle and velocity," *Pattern Recognition*, vol. 34, no. 7, pp. 1491–1501, 2001.

[5] M.-H. Yang and N. Ahuja, "Recognizing hand gestures using motion trajectories," in *Face Detection and Gesture Recognition for Human-Computer Interaction*, Springer, 2001, pp. 53–81.

[6] O. Rashid, A. Al-Hamadi, and B. Michaelis, "Integration of gesture and posture recognition systems for interpreting dynamic meanings using particle filter," in *Soft Computing and Pattern Recognition (SoCPaR), 2010 International Conference of*, 2010, pp. 47–50.

[7] P. P. Kumar, P. Vadakkepat, and L. A. Poh, "Graph matching based hand posture recognition using neuro-biologically inspired features," in *Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on*, 2010, pp. 1151–1156.

[8] C. Sagiv, N. A. Sochen, and Y. Y. Zeevi, "Integrated active contours for texture segmentation," *Image Processing, IEEE Transactions on*, vol. 15, no. 6, pp. 1633–1646, 2006.

[9] X. Liu and D. Wang, "Image and texture segmentation using local spectral histograms," *Image Processing, IEEE Transactions on*, vol. 15, no. 10, pp. 3066–3077, 2006.

[10] D. Dunn and W. E. Higgins, "Optimal Gabor filters for texture segmentation," *Image Processing, IEEE Transactions on*, vol. 4, no. 7, pp. 947–964, 1995.

[11] R. Maini and H. Aggarwal, "Study and comparison of various image edge detection techniques," *International Journal of Image Processing (IJIP)*, vol. 3, no. 1, pp. 1–11, 2009.

[12] O. R. Vincent and O. Folorunso, "A descriptive algorithm for sobel image edge detection," in *Proceedings of Informing Science & IT Education Conference (InSITE)*, 2009, pp. 97–107.

[13] R. Jain, R. Kasturi, and B. G. Schunck, *Machine vision*, vol. 5. McGraw-Hill New York, 1995, pages 147-148.

[14] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, 1999, vol. 2, pp. 1150–1157.

[15] V. Goetcherian, "From binary to grey tone image processing using fuzzy logic concepts," *Pattern recognition*, vol. 12, no. 1, pp. 7–15, 1980.

[16] A. Konar, *Computational intelligence: principles, techniques and applications*. Springer, 2005, page no 45.

[17] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.