A Two-Level Classification Scheme for Single-Hand Gestures of Sattriya Dance

Mampi Devi
Computer Science and Engineering Dept.
Tezpur University,
Tezpur, Sonitpur, Assam.
Email: mampi@tezu.ernet.in

Sarat Saharia
Computer Science and Engineering Dept.
Tezpur University,
Tezpur, Sonitpur, Assam.
Email: sarat@tezu.ernet.in

***Abstract*—The single-hand gestures of Indian classical dance are termed as 'Asamyukta Hastas' which is a combination of two Sanskrit words, asamyukta meaning 'single' and hastas meaning 'hand gestures'. This paper introduces a simple two-level classification method for asamyukta hastas of Sattriya dance which is an Indian classical dance form. In the first level, twenty nine classes of hastas are categorized into three groups based on their structural similarity. Then, in the next level hastas are individually recognized from the database within the group. Moreover, the proposed method extracts Medial Axis Transformation (MAT) from the captured images to identify the groups in the first level. One of the applications of the outcome of this research work can be in the e-learning and self learning of the dance hand gestures (mudras or hastas).**

***Index Terms*—Asamyukta hastas, hand gestures, Indian Classical Dance, Medial Axis Transformation**

I. Introduction

Sattriya dance was originated in the state of Assam. It is one of the popular dance forms among the eight Indian Classical Dance forms. This classical dance uses of several hand gestures, most of them are similar to other classical dances which is performed by both male and female dancers. It is mentioned in the book 'Sattriya Nrittyar Rup Darshan' by Karuna Borah [1] that about 29 single-hand gestures and 47 double hand gestures are used in performing Sattriya dance. The single-hand gestures are known as Asamyukta hastas and double hand gestures are divided into two parts: samyukta hastas and nrittya hastas.

This paper focuses on the classification of the single-hand gestures. The main aim of this paper is to propose a two-level classification method which can gives better recognition of the hand gestures. In the first level, Support Vector Machine (SVM) is used to classify an unknown hasta image into one of three groups and, in the second level Decision Tree classifier is used to recognize the hasta within the group.

The rest of paper is organized as: the next section summarizes the related works available in literature. The propose two-level recognition method described in section III, and section IV presents the experimented result. Finally, we conclude the paper with their improvement scope in section V.

978-1-5090-4291-3/16/31.00©2016$IEEE$

II. Literature Review

Research on gesture recognition has been going on since 1941. The dance gesture recognition is one of the emerging research areas in this domain. Surveys of dance gesture recognition has been done in [2]- [4]. Some of the significant works report in the literature are (i) Bharatnatyam [5], where the recognition based on two-level decision making system and works with single-hand gesture, (ii) Odissi [6], the gesture of whole body using kinect sensor. (iii)Bali Traditional Dance [7], which works on probabilistic grammar-based classifier, (iv) Ballet Dance [8] where multiple stage system is proposed to recognize the different dance posture, (v) Kazakh Traditional Dance [9] is basically concerned with the head gestures and many others. However, no research works on recognition of Sattriya dance gestures are reported in the literature.

III. The Proposed Method

The basic steps involve for any hand gesture recognition system are preprocessing, feature extraction and classification. The work flow diagram as shown in Fig. 1 depicts the basic framework of the whole system. The main purpose of this two-level classification system is to get better recognition accuracy. Each step of this diagram has been explained in the next sub-sections.

A. Preprocessing

This step is an important step to do any further processing. In this method, the preprocessing step is done in two sub steps: background removal and Gaussian filtering. Background removal is done by using GMM [10] on RGB images. Then, a Gaussian filter approach has been used to make the images smooth and noise free. Thereafter, the RGB images are converted into gray images by using MATLAB function and continue to the next section. The steps of preprocessing are shown in Fig. 2

B. Feature Extraction

The main aim of feature extraction is to transform the input image into set of numeric values which is also known as feature vector [11]. The extracted features are used to find out the meaning of gestures. In this paper, geometrical features like centroid, eccentricity, orientation, bounding box, major

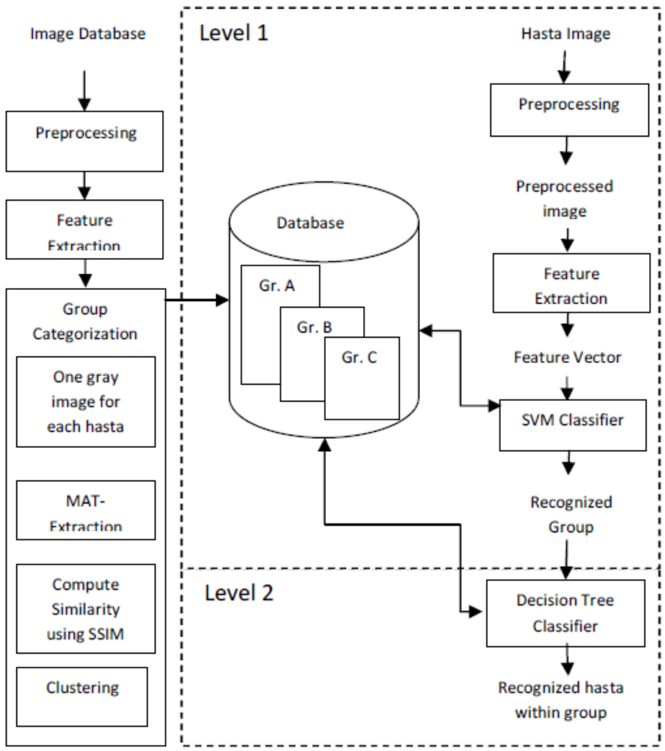

Fig. 1. Workflow diagram for Sattriya dance hand gesture recognition

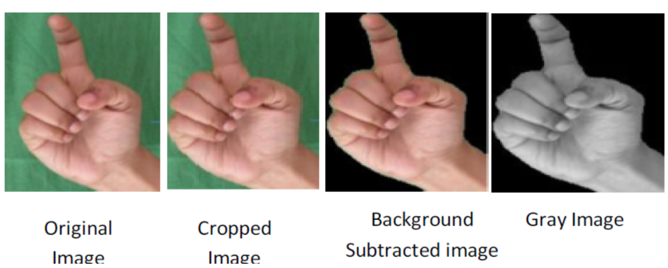

Fig. 2. Preprocessing steps for a hasta (Ankusha hasta)

axis length, minor axis length, aspect ratio and perimeter [12] are used. These features are found out by using feature vector region props techniques. The features are extracted for all the 1015 images.The above mentioned features are defined as follows [12]:

- Centroid: Centroid of the image can be defined as center of mass of the object. In two dimensional image, it represents by two co-ordinates(x,y), where the first element denotes the horizontal co-ordinates of the center of mass and the second element denotes the vertical center of mass.
- Eccentricity: The eccentricity feature of the image can be defined as the ratio of the distance between the foci of the object to its major axis length.
- Orientation: The orientation feature vector represents the angle (in degrees) between the x-axis and the major axis of the object.
- Bounding Box: The bounding box feature returns the smallest rectangle that can contain the object.
- Major Axis Length: The major axis length feature returns the longest diameter of the object.
- Minor Axis Length: The minor axis length feature returns the shortest diameter of the object.
- Aspect Ratio: The aspect ratio can be measured by finding the ratio of the width to the height of the object.
- Perimeter: The perimeter feature can be calculated by finding the distance between each adjoining pair of pixels around the border of the region.

C. Grouping Similar Hastas

This section mainly focuses on the group categorization. To classify the hasta images in the database into three groups, the following steps are performed.

- Take one image for each type of hasta.
- Extract the Medial Axis Transformation (MAT) for each image.
- Apply Structural Similarity Index Method with window size 11X11.
- Create 29X29 similarity matrix and covert into distance matrix.
- Apply hierarchical agglomerative clustering algorithm on distance matrix.
- Draw a complete linkage dendrogram for clustering.
- Cut at threshold point as per requirement of number of group.

The above steps are carried out in order to classify the twenty nine images into three groups. Initially, we take one image for each hasta and extract its Medial Axis Transformation(MAT). The Medial Axis Transformation (MAT) finds out the closest boundary points for each point in an object and finally gives the skeletal of the images. To extract MAT, at first we convert the gray image to binary using threshold value determined by multiplying 1.6 with automatic gray threshold value. In the next step, we apply 25x25 Gaussian filter with average mask sigma=15 to make the images smooth. Then 'skel' operation is used to extract the skeletal from their corresponding image and to reduce the spurs of skeletal 'spur' operation is used repeatedly until its value becomes approximately equal to half of sigma value. The skeletal images are extracted as the gray scale images do not give the ideal output for the next step.
The different steps for MAT is shown in Fig. 3. The next step of this process is implementation of Structural Similarity Index (SSIM) Method. The SSIM works with a square window that moves pixel by pixel over the entire image and calculates the local statistics like mean intensity and standard deviation for each step of movement [13]. The SSIM method is experimented on both gray image dataset and MAT image dataset of the Gaussian weighting function with varying window size of 8x8, 9x9, 10x10 and 11x11. However, in case of gray image dataset, while trying to group the images, it is found that the groups to which an image falls vary with varying window sizes

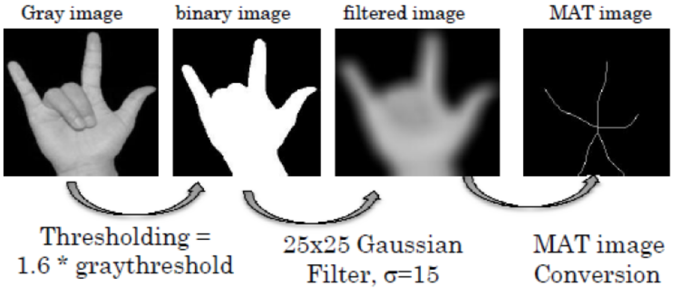

Fig. 3. Steps for MAT extraction from image

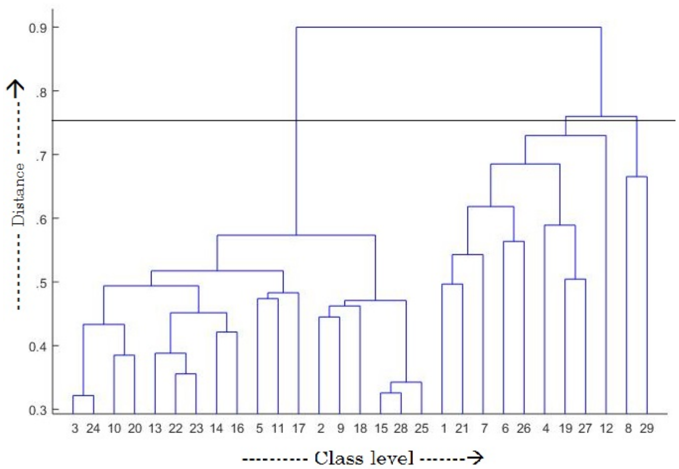

Fig. 4. Creation of Groups on MAT- image dataset

which is not so in case of MAT image dataset where the group does not vary with varying window sizes. Finally, from this experiment we come to conclusion that MAT image dataset is more robust than gray image dataset for group identification. Hence, we use MAT-image dataset and standard window size 11x11 throughout this experiment. Outcome of this experiment is 29x29 similarity matrix with values varying from 0-1. Then, the similarity matrix is converted to a distance matrix by subtracting each value of this matrix from 1. The hierarchical clustering using complete linkage method [14] is applied on the distance matrix and the corresponding dendrogram is obtained as shown in Fig. 4. In this experiment, it was observed that the complete linkage gives the best clustering result. The horizontal axis of the dendrogram represents the class of hastas and vertical axis denotes the distance. The numerical values are used to represent the classes. Their corresponding class names are given in Table I. Finally, to determine the groups, an optimal threshold point of 0.75 is chosen. At this threshold point, three groups are determined from the dendogram. After categorizing the twenty nine hastas into three groups, the datasets are trained for three groups classification.

TABLE I
LABEL OF SATTRIYA DANCE HASTAS

Sl No	Class name
1	Alpadma
2	Ankusha
3	Ardhachandra
4	Ardhasuchi
5	Ban
6	Bhromora
7	Chatura
8	Dhanu
9	Granika
10	Hangsamukha
11	Kartarimukha
12	Kopittha
13	Krishnamukha
14	khatkhamukha
15	Mukula

Sl No	Class name
16	Mustika
17	Padmokosha
18	Pataka
19	Sandangsha
20	Sarpasira
21	Sasaka
22	Sikhara
23	Singhamukha
24	Suchimukha
25	Tamrachura
26	Tantrimukha
27	Tripataka
28	Trishula
29	Urnanava

D. *Classification*

In this phase, the two level classification methods is described as shown in Fig 1. Here, the input image is first preprocessed and then sent to the feature extraction step. The output of this step gives the feature vector centroid, eccentricity, orientation, bounding box, major axis length, minor axis length, aspect ratio and perimeter. These values of feature vector are then compared with those in the database. The group with which the input image matches the most is returned as the output of this step. Support Vector Machine [15] is used for this first level classification. The SVM classifier is used as it gives better result compared to other classifiers. Similarly, decision tree classifier [16] is used at the second level as best results are observed with this classifier at this level. In this level, the classification is narrowed down to the group to which the image matches the most as identified in the first level classification. And with the help of decision tree classifier the image is classified within its relevant group.

IV. EXPERIMENTAL RESULTS

The experimental setup and the results for the proposed method on single-hand gestures of Sattriya dance dataset are discussed in the following sub-sections.

A. *Experimental Setup*

The experiment is carried out on a machine with following configuration: Windows10 OS(64 bits), 4GB RAM, 500 GB Hard-disk and MATLAB 2015. Additionally, the open source machine learning tool 'Weka-3-6-13' is used to test the classification accuracy.

B. *Dataset Description*

The dataset of asamyukta hastas used for this work is our own dataset and it consists of a total of 1015 images from four volunteer dancers. The images are captured using a digital camera of 13 megapixels. Out of four dancers, from three we have collected 870 images and from one, collected 145 images. So, finally 1015 (29types x 10 variation x 3 individuals)+ (29 types x 5 variation x 1 individual) images of asamyukta hastas

TABLE II
STATISTICS FOR FIRST LEVEL CLASSIFICATION

Classifier	Total no. of Instances	Correctly classified instance	Accuracy (%)
K-nn(n=5)	1015	704	71.68
Bayesian Network	1015	639	62.95
Decision Tree	1015	818	80.59
Support Vector Machine	1015	987	97.24

TABLE III
CONFUSION MATRIX FOR FIRST LEVEL CLASSIFICATION WITH SVM

	a	b	c
a	428	0	2
b	15	192	7
c	4	0	367

TABLE IV
STATISTICS FOR SECOND LEVEL CLASSIFICATION

Group	Classifier	Total no. of Instances	Correctly classified instances	Accuracy (%)
GroupA	Decision Tree	430	307	71.54
GroupB	Decision Tree	214	161	75.29
GroupC	Decision Tree	371	295	79.51

are used for our experiment. Since, it is a vision based analysis; there are two limitations that have been taken into account. The limitations are that the images should have a uniform background and there should be a fixed distance between the camera and dancers. The resolution of captured images is too large, so to reduce the size all the captured images are cropped and re-sized into 200 x 200 resolutions.

C. Results Discussion

The overall classification accuracy achieved for asamyukta hastas of Sattriya dance for first level classification is shown in Table II. Among all the classifiers, SVM gives the best results with 97.24% accuracy using RBF kernel with 10 fold cross validation based on extracted feature sets. The kernel parameter C=15 and gamma=0.09 is obtain by varying C and gamma value within a range. So, we have chosen SVM classifier with RBF kernel at the first level classification. Confusion matrix for first level classification using SVM is shown in Table III.

After knowing the group at the first level classification, Decision tree classifier is used to recognize the Hasta at the second level. The recognition accuracies obtained at this level are: 71.62% for Group A, 75.39% for group B and 79.15% for group C. The average accuracy obtained at the second level classification is 75.45%. The details of the recognition accuracy at the second level classification are shown in Table IV.

V. CONCLUSION AND FUTURE DIRECTION

In this paper, a two-level classification method for recognition of single-hand gestures (asamyukta hastas/mudras) of Sattriya dance is proposed. At the first level, an unknown hasta of Sattriya dance is classified into one of the three groups using multiple support vector machines. At this level, 98% accuracy is achieved. At the next level, the individual hastas are classified within the group using decision tree classifier. At this level, the observed recognition accuracies are 71.54%, 75.29% and 79.51% for group A, Group B and Group C respectively. The average recognition accuracy obtained at the second level is 75.45%. The recognition accuracy at second level is not very encouraging. The reason may be that most of the asamyukta hastas are very similar to each other and therefore are chances of misclassification is very high. We will focus more on the identification of some more distinguishing features from these Hastas to improve the recognition accuracy.

REFERENCES

[1] K. Borah, "Sattriya Nrityar Rup darshan," *Grantha-Sanskriti Tarazan, Jorhat*, pp. 113-135, 2009

[2] S. Mitra and T. Acharya," Gesture recognition: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C, Applications and Reviews*, vol. 37(3), pp. 311-324, 2007

[3] M.Hasan and P. K. Mishra, "Hand gesture modeling and recognition using geometric features:A review," *Canadian Journal on Image Processing and Computer Vision*, vol. 3(1), pp. 12-26, 2012

[4] M. Devi, S. Saharia and D.K. Bhattacharyya, "Dance Gesture Recognition: A Survey," *International Journal of Computer Applications*, vol. 122(5), pp. 19-26, 2015

[5] D. Hariharan, T. Acharya, and S. Mitra, "Recognizing hand gestures of a dancer," *In Pattern recognition and machine intelligence,Springer*, pp. 186-192, 2011

[6] S. Saha, S. Ghosh, A. Konar and A.K. Nagar, "Gesture recognition from indian classical dance using kinect sensor," *In Fifth International Conference on Computational Intelligence, Communication Systems and Networks*, pp. 3-8, 2013

[7] Y. Heryadi, M.I. Fanany, and A. M. Arymurthy,"Grammar of dance gesture from bali traditional dance," *International Journal of Computer Science Issues (IJCSI)*, vol. 9(6), 2012

[8] S. Saha, A. Banerjee, S. Basu, A. Konar, and K.N. Atulya, "Fuzzy image matching for posture recognition in ballet dance," *In Fuzzy Systems, IEEE International Conference on*, pp. 1–8, 2013

[9] A.K. Nussipbekov, E. N. Amirgaliyev, and M. Hahn, "Kazakh traditional dance gesture recognition," *In Journal of Physics.: Conference Series, IOP Publishing*, vol. 495, pp. 012-036, 2014

[10] S. Mukherjee and K. Das, "An adaptive gmm approach to background subtraction for application in real time surveillance," *ArXiv preprint arXiv 1307.5800*, vol. 2(1) 2013

[11] N. A. Ibraheem, "Finger Identification and Gesture Recognition Using Gaussian Classifier Model," *International Journal of Applied Engineering Research*, vol. 11(10), pp. 6924-6931, 2016

[12] P. Rajathi and S. Jothilakshmi, "A Static Tamil Sign Language Recognition System," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. (24), 2013

[13] H. R. Sheikh, and A. C. Bovik, " Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15(2), pp. 430-444, 2006

[14] A. K. Jain, M. Narasimha and J. Patrick "Data clustering: a review," *ACM computing surveys (CSUR)*, vol. 31(3), pp. 264-323, 1999

[15] C. Hsu, C. Chang and C. J. Lin, "A practical guide to support vector classification," pp. 1-16, 2003

[16] R, J. Quinlan, "Induction of decision trees," *Machine learning*, vol. 1(1) pp. 81-106, 1986