Chapter 5

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

From the experimental results presented in the previous chapter, it is observed that the classification accuracy of the state of the art classifier using the geometric feature is poor. To improve the performance of single-hand gestures recognition system of Sattriya dance, a two-level classification approach has been proposed which can also help to minimize the search time for recognition of the hand gestures. In the first level, twenty nine of hastas are categorized into three groups based on their structural similarity. Then, in the next level individual hastas are recognized using a decision tree classifier. The proposed method uses Medial Axis Transformation (MAT) as shape descriptor to classify the hastas at 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 motivation of the propose method. The propose two-level recognition method is described in section 5.1, and the experimental results are presented in section 5.2. Finally, we conclude the chapter with a discussion on the limitation and scope for further improvement of the proposed method in section 5.3.

5.1 The Proposed Method

The main purpose of this two-level classification system is to reduce the search space for recognition. This methods has two phases: training and testing. The work flow diagram given in Figure. 5-1 depicts the basic framework of the whole system.

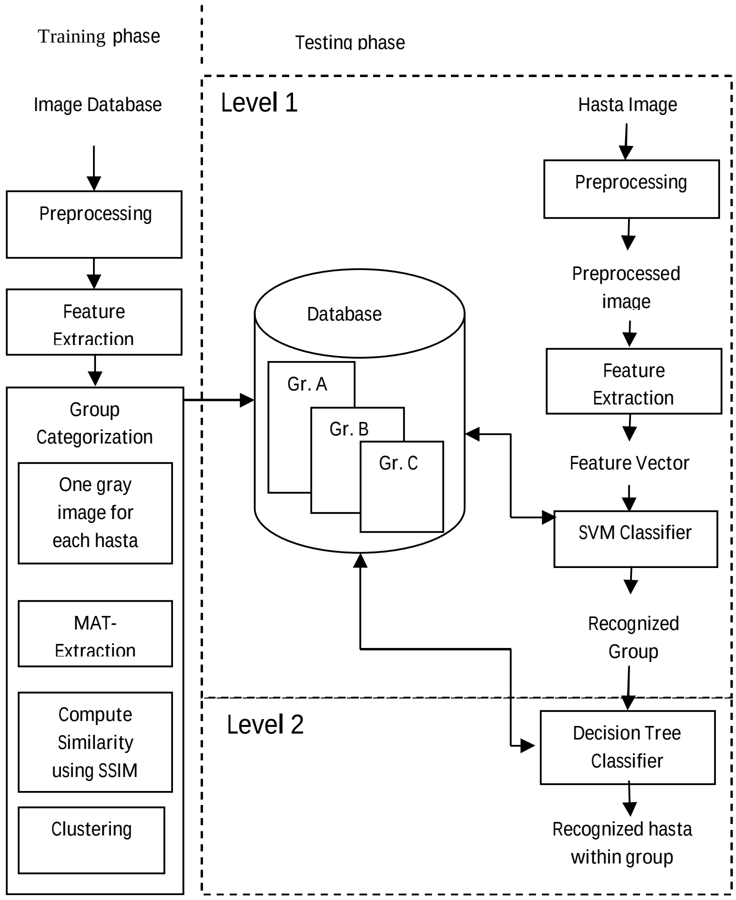

Figure 5-1: Workflow Diagram of Two-Level Classification Method

5.1.1 Training Phase

In training phase, the twenty nine hastas are categorized into three groups. The three basic steps involved in training phase are preprocessing, feature extraction and grouping. Each step of this diagram has been explained in the next sub-sections.

5.1.1.1 Preprocessing

This step is an important step to do any further processing. The hasta images of the dataset are pre-processed as described in chapter 3.

5.1.1.2 Medial Axis Transformation(MAT)

This section explains how the MAT image and extracted from the gray image dataset. 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. Algorithm 1 given below is used to extract the MAT image. The

---

Algorithm 1: ConvertRGBtoMAT

Input: RGB Image
Output: Medial Axis Transformation (MAT) Image

1. Convert input RGB image to gray image
2. Creates a two-dimensional rotationally symmetric Gaussian low-pass filter of size 25 with standard deviation sigma of 15.
3. Filter the Gray image using the Gaussian low-pass filter
4. Convert the Gray image into Binary image by multiplying automatic gray threshold value with 1.6
5. Apply morphological operation skel on the Binary image.
6. Apply spur operation repeatedly until all the spur are removed in the binary image

---

steps are carried out in order to extract Medial Axis Transformation (MAT). 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 25×25 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 outputs of a hasta image at different steps of the algorithm are shown in Figure 5-2.

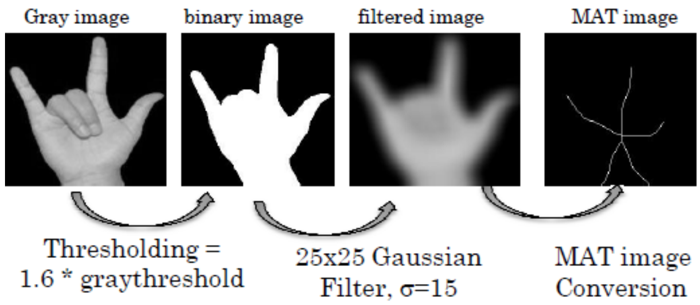

Figure 5-2: Outputs of Different Steps of MAT Extraction of an Hasta Image

5.1.1.3 Feature Extraction

Geometrical features like centroid, eccentricity, orientation, bounding box, major axis length, minor axis length, aspect ratio and perimeter [58] described in the previous Chapter 3 are used in this method. The features are extracted for all of the 1015 images.

5.1.1.4 Grouping

The Algorithm 2, is used to classify the hasta images into three groups. The input to the algorithm are 2D hasta images consisting one image for each hasta. Medial Axis Transformation (MAT) of these images are extracted using Algorithm 1. Then a similarity matrix is computed using Structural

---

Algorithm 2: Grouping

Input: Hasta Image
Output: Recognized Group

1. Take one image for each type of hasta.
2. Extract the Medial Axis Transformation (MAT) for each image using algorithm 1.
3. Apply Structural Similarity Index Method with window size 11×11.
4. Create 29×29 similarity matrix and covert into distance matrix.
5. Apply hierarchical agglomerative clustering algorithm on distance matrix.
6. Draw a complete dendrogram for clustering.
7. Cut at threshold point as per requirement of number of group.

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 [71]. The SSIM method is experimented on both gray image dataset

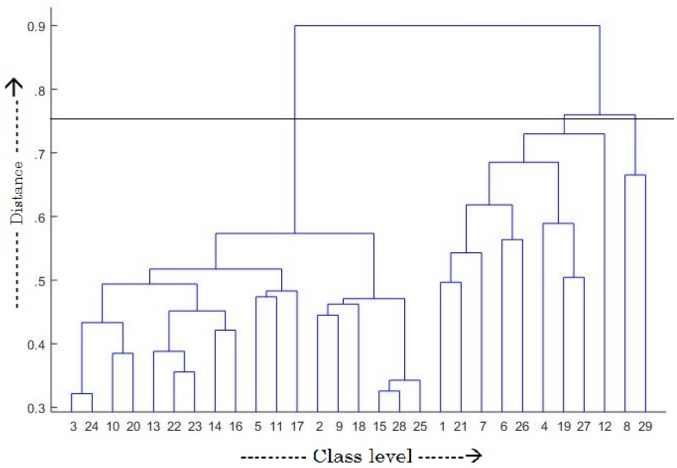

Figure 5-3: Creation of Groups on MAT Image Dataset

and MAT image dataset of the Gaussian weighting function with varying sizes of 8×8, 9×9, 10×10 and 11×11. It is observed in this experiment that the grouping of hastas do not vary with window size of SSID of MAT image whereas the grouping vary for gray images. Therefore, MAT images and a window size of 11×11 pixels are used in this method. The similarity matrix obtained is then converted to a distance matrix by subtracting each value of this matrix from 1. The hierarchical clustering method [26] is applied using the distance matrix and the corresponding dendrogram obtained is shown in Figure 5-3. The horizontal axis of the dendrogram represents the class of hastas and vertical axis denotes the distance. The serial numeric label used in horizontal axis are given in Table 3.2. Finally, to determine the groups, a threshold distance of 0.75 is chosen. At this threshold three groups are determined from the dendogram. After grouping of the 29 hastas in three groups, these group labels are used to train the classifier to classify an unknown hasta image into one of the three group at the level 1.

5.1.2 Testing Phase

The testing phase involve three sub-phases: preprocessing, feature extraction and classification. The classification in the testing phase has been completed in two level. In this phase, the input image is first preprocessed and then sent to the 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 output in the first

Level Classification. Support Vector Machine [24] is used for this first level classification. SVM has been chosen as it gives the best results compared to other classifiers. At Second Level classification, decision tree classifier is chosen as the group is already known at the first level classification and also it gives better results. And with the help of decision tree classifier [56] the input hasta image is recognized as one of the hastas in the group.

5.2 Experimental Results

The experimental description and the obtained outcome for the proposed method on Sattriya dance single-hand gestures dataset are discussed in the following sub-sections.

5.2.1 Dataset Description

In this experiment, we use MAT image dataset to reduce the space complexity which were extracted from our SSHG dataset explained in Chapter 3. This MAT image dataset contains 1015 test instances. Examples of the MAT image dataset are shown in Figure 5-4. The geometric features set consists of eight feature which were extracted from the MAT image dataset. This extracted features were used as a dataset throughout the experiment.

5.2.2 Results Discussion

The overall classification accuracy achieved for single-hand gestures of Sattriya dance for first level classification is shown in Table 5.1. Among all the classifiers, SVM gives the best results with 97.24% accuracy using RBF kernel. The result is validated with 10 fold cross validation based on extracted feature. The kernel parameter C=15 and gamma=0.09 is obtain by varying C and gamma value with in a range. Confusion matrix for first level classification with SVM is shown in Table 5.2. As mentioned earlier, decesion tree classifier is used at the second level. The accuracy obtain at this level for Group A is 71.62%, for group B is 75.39% and for group C is 79.15%. The average accuracy obtain at the second level classification is 75.45%. The details of the second level classification are shown in Table 5.3.

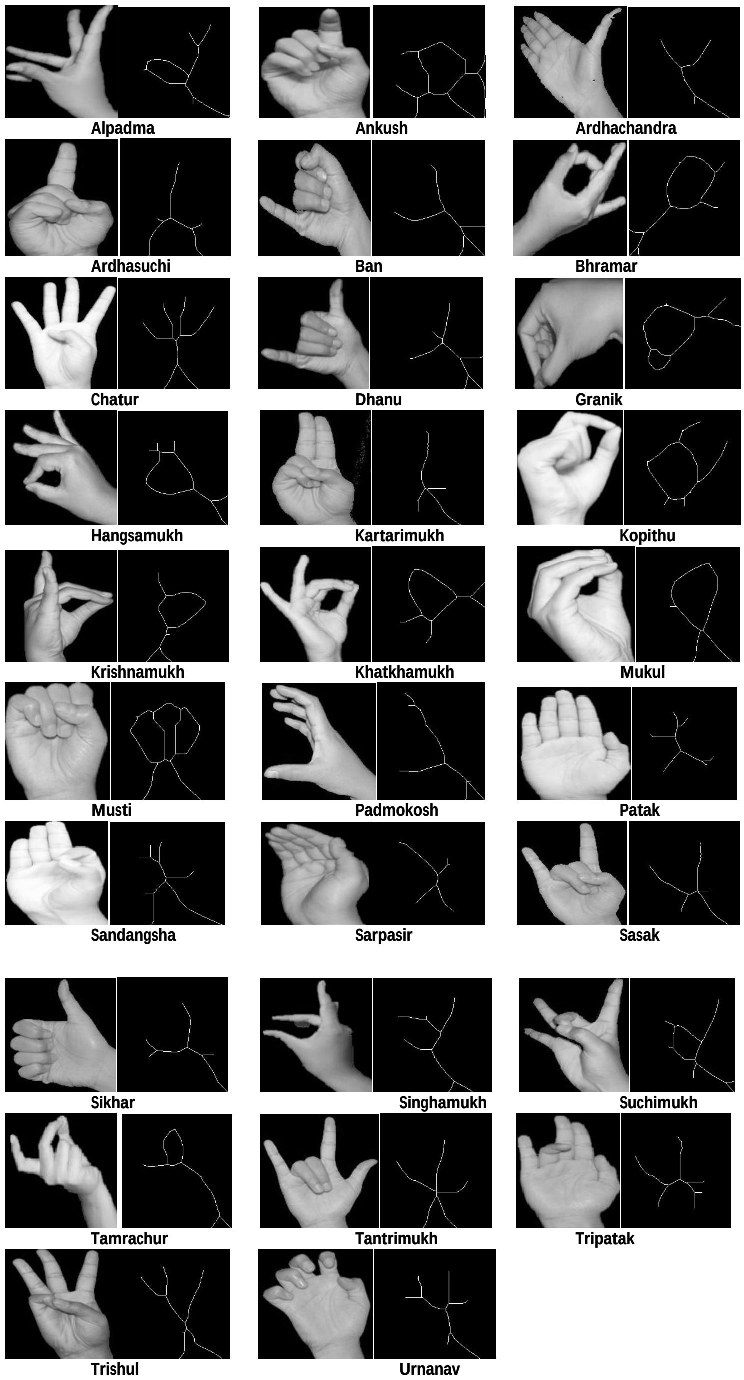

Figure 5-4: Sample of MAT Extracted Image Dataset

5.3 Conclusion and Future direction

In this chapter, a two-level classification method for single-hand gestures (Asamyukta hastas/mudras) recognition of Sattriya dance is proposed. At the first level, an unknown hasta image is classified into one of three group using support vector machine. At this level, 98% accuracy is achieved. At the next level, decesion tree classifiers is used to recognized the hasta image as

Table 5.1: 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 5.2: Confusion Matrix for First Level Classification With SVM

	GroupA	GroupB	GroupC
GroupA	428	0	2
GroupB	15	192	7
GroupC	4	0	367

Table 5.3: 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

one of the hastas of the group identified at level 1. At this level, the accuracy obtained for group A is 71.54%, for Group B is 75.29% and for Group C is 79.51%. The average accuracy for second level classification is 75.45%. The classification accuracy for second level is not very good. The reason may be that most of the asamyukta hastas are very similar to each other. Therefore, the next part of this research of more discriminative features from these images so that a better accuracy can be achieved. In order to improve the recognition accuracy, some more discriminative features are explored and a hierarchical classification method is proposed are presented in the next Chapter.