Available online at http://www.journalajst.com

International Journal of Current Research
Vol. 5, Issue, 06, pp.1457-1462, June, 2013

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

RECOGNIZING SAMYUKTHA HAND GESTURES OF BHARATANATYAM USING SKELETON MATCHING AND GRADIENT ORIENTATION

***Dr. Srimani, P. K. and Kavitha, S.**

Department of Computer Science and Maths, B.U, Director, R and D, B.U.
Department of MCA, Dayananda Sagar College of Arts, Science and Commerce, Bangalore, India

ARTICLE INFO

Article History:
Received 14th March, 2013
Received in revised form
19th April, 2013
Accepted 08th May, 2013
Published online 15th June, 2013

Key words:
Gesture recognition,
Feature extraction,
Skeleton matching,
Orientation histogram,
Edge detection methods.

ABSTRACT

Hand gesture recognition system can be used for providing the interface between computer and human using hand gestures. The main objective of the present work is to develop algorithms for the recognition of twenty three samyuktha mudras of the bharatanatyam. By employing a pattern recognition technique in which the orientation histogram and silhouette of the different gestures is used as a feature vector for mudra classification and interpolation accurate results are obtained.

INTRODUCTION

Digital image processing is the use of computer algorithms to perform image processing on digital images. A digital image is a numeric representation of a two-dimensional image. The field of digital image processing has experienced continuous and significant expansion in recent years. The usefulness of this technology is apparent in many different disciplines covering medicine through remote sensing. The Application of Digital Image Processing includes medical applications, restorations and enhancements, digital cinema, image transmission and coding, pattern recognition, high-resolution display etc. A number of real-world problems from astronomy to consumer imaging find applications for image restoration algorithms. Plus, image restoration is an easily visualized example of a larger class of inverse problems that arise in all kinds of scientific, medical, industrial and theoretical problems. A new and simple two-level decision making system has been designed for performing scale-, translation- and rotation-invariant recognition of various double hand gestures (samyuktha) of bhartanatyam. The orientation filter is used at the first-level to generate a feature vector that is able to distinguish between several gestures. At the second-level the silhouette of the different gestures is extracted, followed by the generation of the corresponding edge detection and the evaluation of the gradients at its end points. These gradients constitute the second feature set, for recognizing those gestures which remain to be identified at the first-level. An application has been provided in the domain of double-hand gestures of bharatanatyam. Computer recognition of hand gestures may provide a more natural-computer interface, allowing people to point, or rotate a CAD model by rotating their hands. Hand gestures can be classified in two categories: static and dynamic. A static gesture is a particular hand configuration and pose, represented by a single image. A dynamic gesture is a moving gesture, represented by a sequence of images. We will focus on the recognition of double hand mudras.

***Corresponding author:** Dr. Srimani, P. K.
Department of Computer Science and Maths, B.U, Director, R and D, B.U.

Related Work

A thorough survey of the literature pertaining to the topic of research reveals that very sparse literature is available with regard to the subject and no work is available with regard to the present work. Hence, the present investigation is carried out to throw light on the subject. Some related works include [1] to [6]. The main objective of the present investigation is recognition of Bharatanatyam mudras by using the Image processing techniques. The codes are written and implemented through MATLAB. The work is first of its kind in the literature. The scope of this investigation is to create a method to recognize Samyukta Bharatanatyam gestures, based on a pattern recognition technique; employing histograms of local orientation. The orientation histogram will be used as a feature vector for gesture classification and interpolation.

Background

Research on hand gestures can be classified into three categories. The first category, glove based analysis, employs sensors (mechanical or optical) attached to a glove that transducers finger flexions into electrical signals for determining the hand posture. The second category, vision based analysis, is based on the way human beings perceive information about their surroundings. It is probably the most difficult to implement in a satisfactory way. The third category, analysis of drawing gestures, usually involves the use of a stylus as an input device. Analysis of drawing gestures can also lead to recognition of written text.

METHODOLOGY

The present investigation is carried out by using MATLAB [7]. MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. This allows us to solve many technical computing problems, especially those with matrix and vector formulations. In a fraction of the time it would take to write a program in a scalar non-interactive language such as C or Fortran.

The reason for selecting the MATLAB tool[7] for the development of this paper is its toolboxes. Toolboxes allow you to learn and apply specialized technology. Toolboxes are comprehensive collections of MATLAB functions (M-files) that extend the MATLAB environment to solve particular classes of problems. It includes among other image processing and neural networks toolboxes. The objective of this research is to develop a suitable prototype for the recognition of the 24 Samyukta Hastas of Bharatanatyam in a two dimensional space, using image processing techniques [1]. We aim to make the computer act as a teacher to correct the dance gestures, for the purpose of promoting e-learning of the nuances of Bharatanatyam across the world.

System Overview

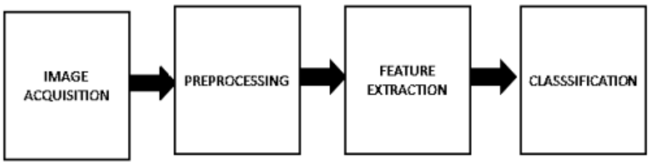

Fig 1. Block Diagram of hand gesture recognition system

Data Set Description

The term Indian Classical Dance comprises all the art forms of the Natyashastra written by the ancient musicologist Sage Bharata. The Natyashastra confers classical status to eight Indian dances, of which Bharatanatyam is one of the oldest. A popular interpretation of the name is BHAva (expression) + RAga (musical mode) + TAla (rhythm) + NATYAM (dance) = BHARATANATYAM. A distinctive feature of Bharatanatyam Dance is the use of expressive hand gestures as a way of communication. Hastas [3] refers to a variety of hand symbols that a dancer can use. Hastas can be broadly classified into two categories, viz., Asamyukta Hastas (Single hand gestures) and Samyukta Hastas (Double hand gestures). There exist 28 Asamyukta Hastas and 24 Samyukta Hastas.

Fig. 2. Samyukta Bharatanatyam Mudras

Anjali: This mudra is used to offer salutation to the God, Elders, Teachers and a Brahmin. The Anjali mudra is kept above the head to offer salutations to the God. It is kept in front of the face to offer respects to the Teachers and Elders. It is kept in front of the chest to offer pranams to the twice -born or the Brahmin's.

Karkata: This verse says that Karkata is used to denote: Arrival of people (Get together), Showing the Belly Blowing the council, Twisting and stretching of Limbs, Bending a Branch

Kapota: The shloka says that Kapota can be used to denote the following: Respectful salutations. Such a Mudra is held while conversation with teachers or Guru. Its a mark of acceptance or obedience. It shows the Vinayam (down to earth attitude) quality of an individual.

Swastika: The shloka says that swastika denotes a Crocodile: My imagination says it can be used, To say "No", To show a blocked road or a passage, To show imprisonment

Dola: The shloka says that Dola can be used: At the commencement of natya (Natyarambhe), It is a pose prescribed by the experts.

Pushpaputa: Thus according to the shloka, Pushpaputa is used to denote the following: Indicate waving of lights to gods (to perform Aarti), Acceptance of water or any fruits etc, To show offerings to Gods Offering floral tributes at the time of Mantra puja

Utsanga: The shloka says that Utsangha is used to denote: An embrace, To show Modesty, Shyness etc, Display of Armlets and other such ornaments, Coaching children, The shloka say that this Mudra indicates a Linga of lord Shiva. Linga means sign, mark or symbol. Thus a Shivlinga is a symbol that represents Hindu God lord Shiva, Linga is derived from the Sanskrit word lingam, which comes from li meaning, "to dissolve" and gam "to move on". This refers to the belief that one appears as a being in the world and then dissolves back into the universe. A Shivalinga is phallus-shaped and fixed on a base, which is shaped like a yoni. The structure symbolises the supreme creative energy. It represents the Shivashakti ie the union of Lord Shiva and Devi Parvati. It is usually made of stone.

Katakavardhana: The shloka says that katakavardhana can be used to denote: Coronation ceremony, Worship, Marriage

Karariswastika: According to the the shloka karariswastika help to show: Stems and branches of the trees, A huge tree, A hill top or its peak

Shanka: It says that the Shankha mudra is used to denote a council, Shankha was also blown in the battlefield as a sign to begin the war. The sound of a Shankha is considered sacred and is also blown in some of the Indian rituals. Shankha has a divine significance in Indian mythology. It says that this Mudra denotes a Wheel. In Sanskrit Chakra means a circle or a wheel. Very often in Dance this hasta is used to represent Lord Vishnu's Sudarshan chakra. He is dipicted holding it with his right forefinger. This chakra is a Divine weapon with sharp edges and is supposed to spin like a Disc.

Samputa: It says that Samputa is used to denote: Concealment of objects. It can be used to show secrets or something that is kept safely. It can also denote something that is hidden.

Pasha: The shloka says that the pasha hasta is used to denote: Internal fights or Quarrels, A String, A Chain

Kilaka: The shloka says that this kilaka hasta is used to denote: A feeling of Love and affection, Funny talks, hilarious conversation etc.

Preprocessing

Preprocessing is very much required task to be done in hand gesture recognition system. We have taken prima database[1] which is standard database in gesture e recognition. We have taken total 25 signs each sign with 40 images. Preprocessing is applied to images before we can extract features from hand images. Preprocessing consist of two steps

- Segmentation
- Morphological filtering

A. Segmentation is done to convert gray scale image into binary image so that we can have only two object in image one is hand and other is background. Otsu algorithm [2] is used for segmentation purpose and gray scale images are converted into binary image consisting hand or background. After converting gray scale image into binary image we have to make sure that there is no noise in image so we use morphological filter technique. Morphological techniques consist of four operations: dilation, erosion, opening and closing. The preprocessing of the hand gesture consists of detecting the skin color in the image, and cropping the hand region in order to avoid unnecessary details in the background. The images are then resized to 240 x 240 and converted to gray scale. The next step is the feature extraction procedure. A suitable feature vector which is invariant to translation, rotation, scaling and reflection needs to be chosen for the purpose of distinguishing between the different gestures.

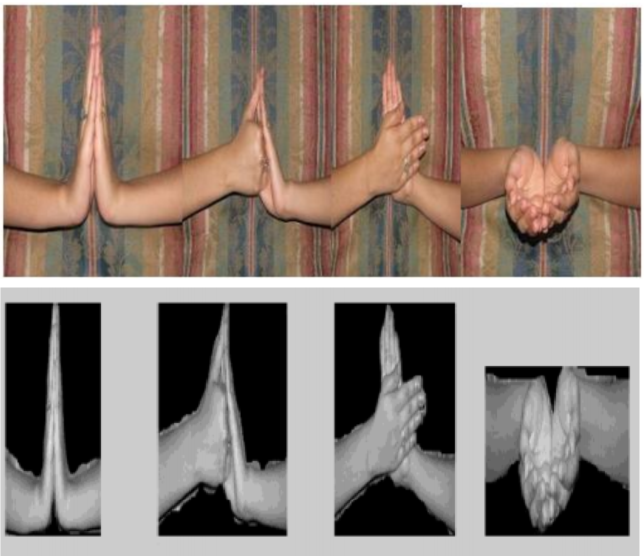

Fig.3. Sample gestures (a) Before preprocessing (b) After preprocessing

Train set

There are twenty three training sets of images, each one containing three images. Each set originates from a single image for testing. Three operations were carried out in all of the images. They were converted to grayscale, the background was made uniform and applied spatial transformations (Rotation and Scaling) (Fig. 4.). A spatial transformation (also known as a geometric operation) modifies the spatial relationship between pixels in an image, mapping pixel locations in an input image to new locations in an output image. The toolbox includes functions that perform certain specialized spatial transformations, such as resizing and rotating an image. In addition, the toolbox includes functions that you can use to perform many types of 2-D and N-D spatial transformations, including custom transformations: Rotation and Scaling.

The final form of database would be:

Fig.4. Spatial transformations for sample mudra (i) Original image Gyanmudra (ii) Rotated image (iii) Resized image

A major difficulty is associated with the rotation and scaling involved. For example, a gesture image rotated to any degree or scaled to any level should represent the same gesture. Orientation filters [2] have been used in various image processing and vision tasks, by applying filters of arbitrary orientation and phase. Typically a few filters, corresponding to a few angles, are employed and the intermediate responses are interpolated. With a correct filter set and interpolation rule, it becomes possible to evaluate a filter of any arbitrary orientation.

Feature Extraction

In this paper we use the orientation filter at the first-level to generate a feature vector for distinguishing between different gestures. At the second-level the silhouette of the different gestures is extracted, followed by the generation of the corresponding skeleton and the evaluation of the gradients at its end points. This constitutes the second feature set, for recognizing those gestures which remain to be identified at the first-level.

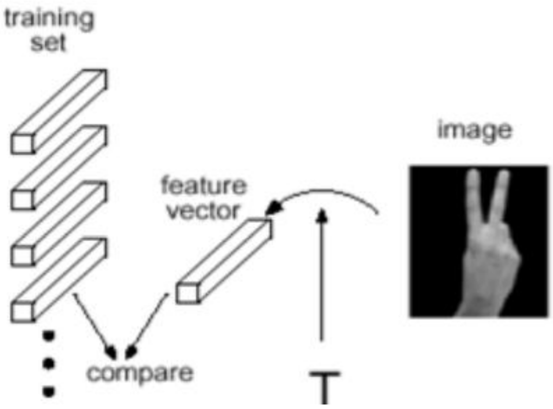

Fig. 5. Pattern Recognition System

The pattern recognition system that will be used can be seen in Fig. 5. Some transformation T, converts an image into a feature vector, which will be then compared with feature vectors of a training set of gestures. Histogram orientation has the advantage of being robust in lighting change conditions. If we follow the pixel -intensities approach certain problems can arise for varying illumination. Taking a pixel-by-pixel difference of the same photo under different lighting conditions would show a large distance between these two identical gestures. Orientation analysis should give robustness in illumination changes while histogramming will offer translational invariance. This method will work if examples of the same gesture map to similar orientation histograms, and different gestures map to substantially different histograms[4] and [5].

Edge Orientation Histogram

Considering the aspects of translation and scaling invariance, the orientation histogram [2] was found to be a useful feature component. Here these constitute histograms of local orientation of the hand gesture. The orientation histograms are robust to illumination changes, and are simple and fast to compute. Local orientation is obtained by the use of steerable filters, in which a filter of arbitrary orientation is synthesized as a linear combination of a set of "basis filters". A set of 36 one-dimensional Gaussian steerable filters and their first order derivatives have been used for extracting the local edge orientation properties of the hand gesture for every 10°, ranging from 0° to 350°. The gesture image I, filtered at an arbitrary orientation and convolved with the filters, can be synthesized for the oriented filter response

$$R_1^{\theta} = (G_1^{0^\circ} * I)\cos\theta + (G_1^{90^\circ} * I)\sin\theta,$$

where * represents the convolution operator, and G^0_1 is considered at an arbitrary orientation 0.

In order to enhance the ability of the edge orientation histogram in recognizing the rotated gestures, the direction of maximum local orientation of every gesture is found. Accordingly the original gesture image is rotated in such a manner that the maximum orientation is obtained at 180°. The edges of the hand gesture are extracted using the Laplacian of Gaussians (LOG) [1], and the image is multiplied with the filter response for every value of 0. Fig. 6 illustrates the polar plots of the edge orientation histogram corresponding to the edges from the gestures "Anjali", "Shanka", "Chakra" and "Pushpaputa".

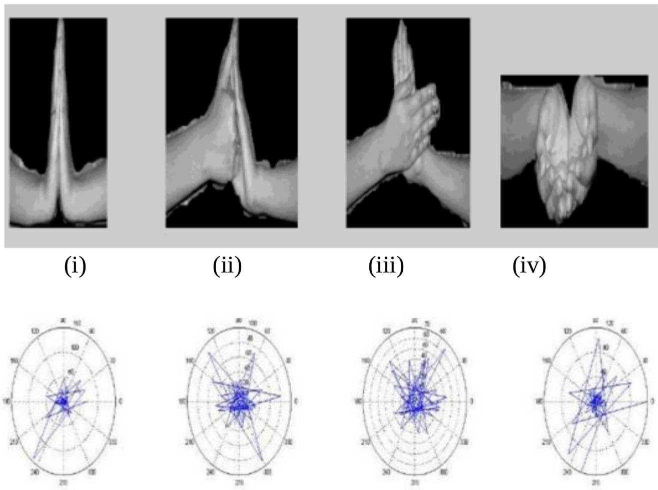

Fig. 6. Double-hand Bharatanatyam gestures. (i)Anjali (ii) Shanka (iii) Chakra (iv) Pushpaputa.

Polar plots of corresponding edge orientation histograms of gestures Anjali, Shanka, Chakra and Pushpaputa

Gradients at Corner Points of Skeleton

Those gestures that cannot be recognized at the first-level, by the use of edge orientation histograms, are processed further at a second-level. It is known that the end points of a skeleton correspond to a change of curvature. Refined categorization is next made, using the gradients at the extremities of the skeleton as a new set of features. A given gray scale image is initially binarized in a uniform manner over hand gestures of different people. The boundary is extracted from the binary image, followed by the flood-fill operation to generate a uniform silhouette. This is depicted in Fig. 7(i) for a sample dance gesture.

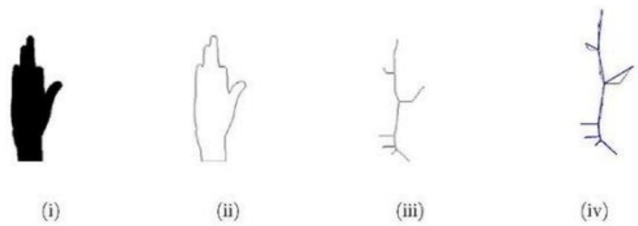

Fig. 7. Sample gesture (i) Silhouette, (ii) Boundary, (iii) Skeleton and (iv) Connectivity graph.

The skeleton [Fig. 7(iii)] is extracted by the morphological operation of thinning. Typically it consists of a number of branch points and end points, connected by curve segments. These points are considered to be the nodes of a graph, such that the skeletal curve segments form the edges. This graph [Fig. 7(iv)] is called the connectivity graph of the skeleton. It provides topological information about the hand gesture. Matching of the connectivity graphs, based on their topologies and geometric features, provides a distance measure for determining the similarity (or dissimilarity) between the different shapes. The adjacency matrix of the connectivity graph is constructed. The degree of every node in the graph is computed and assigned as its weight. Subsequently, a depth-first traversal sequence is constructed for the connectivity graph starting from one end point (typically, the leftmost point). Preference is provided to a node with a lower weight during the traversal. Thereby, the number of backtracking sequences gets reduced.

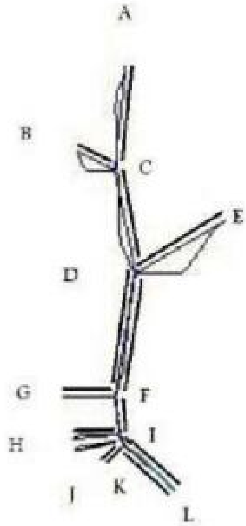

Fig. 8 . Depth-first traversal in edge connectivity graph of sample gesture

The thick black lines in Fig. 8 indicate the depth-first traversal sequence of the connectivity graph of Fig. 7(iv). The sequence obtained is expressed as

G -> F -> I -> H -> I -> K -> J -> K -> L -> K -> I -> F -> D -> E -> D -> C -> B -> C -> A.

Repetitions of nodes correspond to the backtracking of edges during traversal of the graph.

For every end point in the skeleton, the nearest boundary point is obtained. Without loss of generality, it can be inferred that this nearest boundary point will be either a finger tip or a point in the hand gesture where the curvature change is large. Hence the gradient at these points will significantly vary over the different gestures. For the branch points, this value is assigned to be zero in order to avoid unnecessary weights due to backtracking sequences. The values of gradients are then substituted for the nodes, in the depth-first traversal sequence. This sequence of gradients and zeros forms the feature vector for the second-level of recognition.

Edge Detection

The edge is a set of those pixels whose grey have the step change and rooftop change, and it exists between object and background object and object region and region between element and element. When image is acquired the factors as projection, mix, aberrance and noise are produced. Above mention factors bring on image features blur and distortion, due to this it is difficult to detect edge. In process of noising we use a “Gaussian Noise removal of image on the local feature” after then we apply different operators’ e.g Binary Morphology operators, canny operator, log operator, and differential operator for edge detection [11]. The image can be affected by noise inevitably in the process of saving and transmission and noise causes the negative effect on the image processing and analysis. For removing these effects, it is necessary to remove or decrease the noise, at the same time conserve the image information as much as possible, such as edge and the texture.

Image De-noising

Gaussian noise model has a very significant feature, it does not matter how much the variance and histogram of the original image is it will always follows the Gauss distribution. In Gaussian method firstly according to the feature that in the image the local neighbourhood pixel in the same object are smooth, we estimate whether the pixel point is on the image edge, the noise point or the edge texture point [4]. Then according to the local continuity of the image edge and the texture feature, using the continuity of the image. And then locate the noise points. Lastly for the noise which is not on the edge or the texture. Using the mean value of the non-noise points in the adaptive neighbourhood to eliminate the noise, and for the noise on the edge and texture region just using the pixel points of the neighbourhood edge and texture to smooth. With the help of this method we can remove the Gaussian noise in the image well and the number of the residual noise points decreases sharply.

Experimental Results

Fig.15 shows the Pop-up menu to select the gesture for recognition either from test setor from trainset.

Fig. 14. Pop-Up menu to select Test and Train set for Recognition.

The gesture vocabulary was built by capturing ten images of each of the 23 double hand gestures of Bharatanatyam from the hands of three different people, using a five megapixel camera. These ten images were translated, rotated, scaled and reflected versions of a single gesture. Hence every image is unique. Out of the ten images, eight were randomly selected for training the system, while the remaining two were kept for testing. This amounted to a total of 224 training images. The test images were identified based on the closest match to the learned examples (prototypes) in terms of Euclidean distance. In the first-level, orientation histogram was used for feature extraction. In the second-level a shape-based skeleton matching was used, with the gradients at the corner points being used as a new set of features.

Fig. 15. Pop-up menu

Fig.15. shows the pop-menu for the user to select the particular gesture to display the hand gesture.

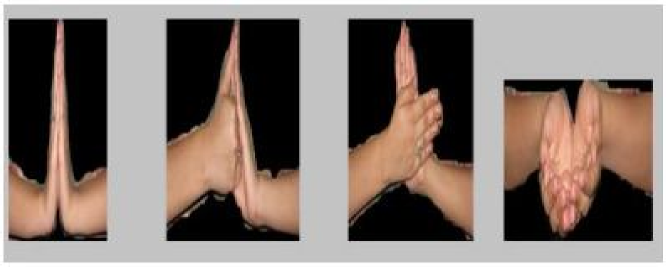

Fig. 16. Distinctly identifiable gestures at the first-level. (i) Anjali (ii) Shanka (iii) Chakra (iv) Pushpaputa.

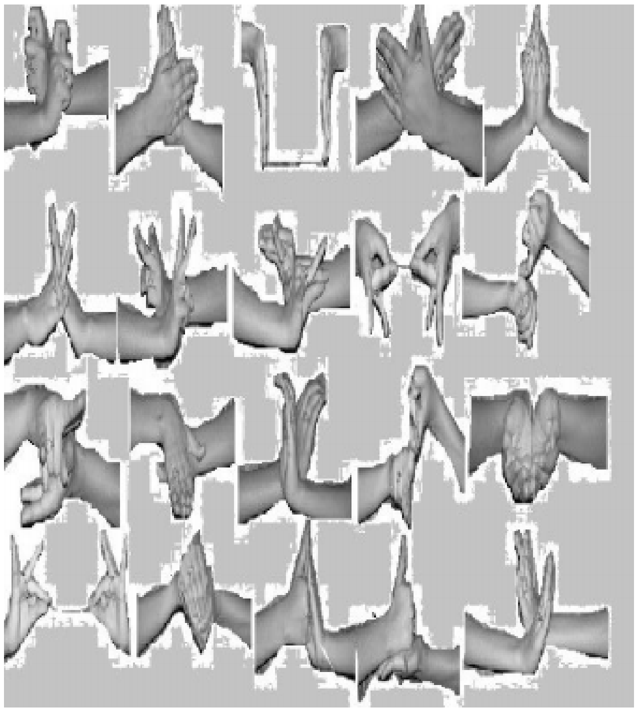

Fig.17. (i)Anjali (ii) Kapota (iii) Karkata (iv) Swastika (v) Dolahasta (vi) Pushpaputa (vii) Utsanga (viii) Shivalinga (viii) Katakarardhana (ix) KarthariSwastika (x) Shakata (xi) Shanka (xii) Chakra (xiii) Shamputa (xiv) Pasha (xv) Kilata (xvi) Matsya (xvii) Kurmat (xviii) Varaha (xix) Garuda (xx) Nagabandha (xxi) Khatwa (xxii) Bherunda (xxiii) Avahitta

Application of the edge orientation histogram resulted in the correct identification of four gestures, which were found to be distinctly unique from the others. These are Anjali, Shanka, Chakra and Pushpaputa , as shown in Fig. 16. The remaining 19 gestures could be grouped into three classes, as indicated by the three rows of Fig. 17.
These 19 images were used for subsequent processing at the second-level, as described below. Shape-based skeleton matching was next used, with the set of gradients at the corner points in the skeleton serving as the new set of features for the remaining 192 (= 224 – 32) training images. The skeleton of the test image of a hand gesture was first obtained, and its connectivity graph generated. Then a depth-first traversal sequence with the gradient values was computed, for use as the feature vector. The left-most point of a skeleton image was generally considered as the starting point of the depth-first traversal for every gesture. Since all gestures are rotated to give a maximum orientation at 180°, therefore it is safe to assume that for the same gesture the starting point will be the same. The matching of two shapes now reduces to the problem of matching two feature sequences. Let V and U be two such sequences of lengths n and m respectively. Using concepts from dynamic programming, a cost function Cost(i,j) is defined as the cost of matching the i^{th} element V_i of the first sequence with the j^{th} element U_j of the second sequence. It is defined as

Cost(i,j)=min{Cost(I, j-1),Cost(i-1, j), Cost(i-1, j-1)}+Weight(Vi) *Vi-weight(Uj)*Uj (2)

In order to match the sequences V and U, one needs to compute the cost over their entire length. It is evident that the value of the cost function determines the similarity (or dissimilarity) between the shapes of different hand gestures. There are, however, possibilities of error due to the irregularities in the skeleton structure from the hands of different people with different textural properties.

Conclusion

This paper would be a very practical, simple and cost-effective mode of imparting training in the nuances of traditional dance across the globe, thereby assimilating the cultural divide between the Orient and the West. We aim to make the computer act as a teacher to correct the dance gestures for the purpose of promoting classical Indian dance across the world. A simple and new two-level decision making system has been designed for recognizing the samyuktha gestures of Bharatanatyam, a well-known Indian classical dance form. It has been shown to be scale-, translation- and rotation-invariant while recognizing various double-hand gestures of a dancer.

REFERENCES

1. William T. Freeman, Michael Roth, “Orientation Histograms for Hand Gesture Recognition” IEEE Intl. Wkshp. On Automatic Face and Gesture Recognition, Zurich, June, 1995.
2. Maria Petrou, Panagiota Bosdogianni, “Image Processing, The Fundamentals”,Wiley Vladimir I. Pavlovic, Rajeev Sharma, Thomas S Huang, “ Visual Interpretation of Hand Gestures for Human-Computer Interaction : A review” IEEE Transactions of pattern analysis and machine intelligence, Vol 19, NO 7, July 1997.
3. Srinivas Gutta, Ibraham F. Imam, Harry Wechsler, “ Hand gesture Recognition Using Ensembles of Radial Basis Functions (RBF) Networks and Decision Trees” International Journal of Pattern Recognition and Artificial Intelligence, Vol 11 No.6 1997.
4. Lalit Gupta and Suwei Ma “Gesture-Based Interaction and Communication: Automated Classification of Hand Gesture Contours”, IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 31, no. 1, February 2001.
5. Srimani P.K. and Kavitha S “Hand Gesture Recognition using Gradient Orientation”, International Journal of Current Research Vol. 4, Issue, 02, pp.191-195, February, 2012.
6. Srimani P.K. and Kavitha S, “A Comparative Study of Different Deblurring Methods Using Filters”, CP1414, 2nd International Conference On Methods And Models In Science And Technology (ICM2ST-11) © 2011, American Institute of Physics 978-0-7354-0879-1/11/$30.00.
7. Duane Hanselman, Bruce Littlefield, “Mastering MATLAB, A comprehensive tutorial and reference”, Prentice Hall, 2007
8. Srimani P.K. and Kavitha S, “Yoga Mudras Recognition using Gradient Orientation and Siloutte”, International Journal of Current Research Vol. 4, Issue, 09, pp.173-178, September, 2012.

[9 Matthew A. Turk and Alex P. Pentland, “ Face recognition using eigenfaces”, IEEE Society Conference on Computer Vision and Pattern Recognition, pages 586–591, Lahaina, Maui, Hawaii, June 3-6 1991.

10 J. Edward Jackson, “A Users Guide to Principal Components”, Wiley Series in Probability and Mathematical Statistics, A Wiley-Interscience Publication, 1st edition, 1991.

11 Canny J F., "A computational approach to edge detection [J]". IEEE Trans on PAMI, 1985, 8(6): 679 -698.

12 Heung-Soo Kim, Jong-Hwan Kim., "A Two-step detection algorithm from the intersecting chords". Pattern Recognition Letters. 2001, 22:787-798.

13 Lei Lizhen, "Discussion of digital image edge detection method", mapping aviso, 2006, 3:40-42.

****