

Dance Gesture Recognition: A Survey

Mampi Devi
Tezpur University

Sarat Saharia
Tezpur University

D.K.Bhattacharyya
Tezpur University

ABSTRACT

Gesture recognition means the identification of different expressions of human body parts to express the idea, thoughts and emotion. It is a multi-disciplinary research area. The application areas of gesture recognition have been spreading very rapidly in our real-life activities including dance gesture recognition. Dance gesture recognition means the recognition of meaningful expression from the different dance poses. Today, research on dance gesture recognition receives more and more attention throughout the world. The automated recognition of dance gestures has many applications. The motive behind this survey is to present a comprehensive survey on automated dance gesture recognition with emphasis on static hand gesture recognition. Instead of whole body movement, we consider human hands because human hands are the most flexible part of the body and can transfer the most meaning. A list of research issues and open challenges is also highlighted.

Keywords

Gesture, hand gesture, automated dance gesture, vision-based gestures, glove-based gestures.

1. INTRODUCTION

Gestures are powerful and natural ways of non-verbal communications in which visible body actions are used to communicate important messages. Gesture recognition means the identification of meaningful expressions of human motion; including recognition of fingers, arms, hands, head and body expressions [1].

To study gesture recognition, it is essential to understand the applications, definitions and nature of the gestures. Gesture recognition is a multi-disciplinary research area.

The application areas of gestures spread from our daily life activities to Artificial Intelligence through Electronics Engineering [2], Medical Science, Education, and Robotics [3] and also for entertainment [4] purpose.

We formally define the gesture and hand gesture as follows.

Definition1: Gesture: It is defined as an image expression that communicate a message.

Definition2: Hand gesture: It is defined as an hand image expression which convey meaning during natural conversation.

Hand gestures are the most important types of gestures because human hands are the most flexible parts of the body and can transfer most meaning. Culture specific hand gestures are able to convey different meanings in different culture.

In the recent past, several reviews on gesture recognition [1], skeleton gesture recognition [5], vision-based gesture recognition [6], and vision-based motion capture [7] have been published in literature. However, our survey differs from them in the following ways.

- Like [1][5], we also provide a taxonomy of gesture classification techniques and additionally, we discuss

about vision-based gesture recognition. Also, we include comparison of various glove-based and vision-based techniques.

- Unlike most of them, we highlight a list of issues and research challenges in this live research domain.

The main aspects of this survey is to provide information about existence work on automated gesture recognition in dance domain and to point out the issues and to provide ideas for future scope in this domain. Automated recognition of dance gestures can help in creating universal communication environment for a dance drama, independent of the language used in the associated song [8]. Therefore, such recognition system helps a viewer to understand the meaning of dance sequences irrespective of the language of the background song. Also, it includes applications such as dance self-assessment and e-learning of dances [9], [10]. Next, we discuss six major applications of gesture recognition.

1.1. Computer vision

Gesture recognition in computer vision helps improving the Human Computer Interaction (HCI) in the form of, say, providing support to young children or physically challenged people to interact more effectively with computer. Developing of various sign languages like American sign language[11], Arabic sign language(ArSL) [12], Japanese sign language (JSL) [13], [14], Korean sign language(KSL) [15], [16], Brazilian sign language[17] gives new life for hearing impaired and physically challenged people. System developed for hand gesture recognition in computer vision can easily recognize the gestures and use them for controlling the electronic devices.

1.2. Electronics Interface

The demand for touch and touch-less technologies are another major scope for gesture recognition research. The remote control device, multi-touch device, touch pad, iPhone, input devices like wired gloves, depth-aware cameras, stereo cameras all are directly or indirectly involved with gesture recognition research.

1.3. Artificial Intelligence

Gesture recognition mostly focused with the analysis of functionality of human being. It aims to gather and represent human gestures and use them to convey meaning. Recognition of sign languages and lie detection are other possibilities of gesture recognition.

1.4. Neurology

The main task of neurology is to develop and support human intelligence. Understanding human emotion and behaviour is the most important research area in present day world. Gestures are generally processed in the left inferior frontal gyrus (Broca's area) and the posterior middle temporal gyrus area. These areas of the brain support the pairing of gesture and meaning and it adapts in human evolution. It works for both co-speech and speechless gestures [18].

1.5. Medical Science

Monitoring the patient's emotional state and stress is another major application of gesture recognition. Hand gesture recognition systems can help doctors in surgical environment to manipulate digital images during medical treatment.

1.6. Social Significance

Today, gesture recognition also has an important role in our daily life from official works to entertainment activities. An important contribution of gesture recognition is e-learning. It makes easy to communicate in video conferencing. Gesture can work as a bridge between human and computer i.e., human-computer interaction (HCI) which makes it more user-friendly and more flexible. Expressions of human body are roughly classified into three types of gestures [1]: (a) Hand gestures, (b) Head and face gestures and (c) Body gestures.

Rest of this paper is organized as follows Section 2 provides a generic view of dance gesture recognition which is the prime focus of this work. Section 3 provides taxonomy of gestures and in section 4 covers the work review on dance gesture recognition domain followed by study on glove-based as well as vision-based approaches. Section 5 gives a short description of the existing approaches and methods and Section 6 gives the future direction of the study. Finally, section 7 concludes the summary of this survey paper.

2. DANCE GESTURE RECOGNITION

Research on dance gesture recognition has been receiving an increasing attention day-to-day. Automated dance gesture recognition is a special task of general gesture recognition. Dance gesture recognition means the recognition of meaningful expression from different dance poses. It expresses the meaning of dance drama and help to create universal communication environment. It can also help for self assessment and e-learning of dances [8].

Next, we provide few major definitions to provide a foundation on dance gesture recognition.

Definition 3: Gesture Feature: Gesture feature f_i of a gesture type g_i is defined as a characteristic of g_i .

Definition 4: Gesture Feature Vector: It is defined as a subset of features f_i 's to represent a gesture type.

Definition 5: Gesture Class: It is defined as a group of similar gesture images.

Definition 6: Gesture Recognition: It is a process of recognizing a type of gesture

The procedure for dance gesture recognition can be divided into four major phases; viz., (a) image acquisition (b) preprocessing, which can be further divided into two major tasks, i.e., (i) segmentation and (ii) boundary detection, (c) feature extraction and (d) classification and recognition

We now describe each of these phases, in brief.

2.1. Image Acquisition

The capturing of image is generally performed by a clean environment with uniform background using single camera. If we use two cameras with approximately orthogonal direction then we can easily reduce the ambiguities compared to single view approach. Kinect sensor camera [19] or depth sensor camera [9] are used for capturing 3-D images. The images can be directly skeletonized and also the position of the skeleton can be estimated by using both these cameras.

2.2. Preprocessing

The main task of this phase is to process the images those are captured in the previous phase without losing their salient features. The images are processed by cropping, resizing, filtering, transforming (using say Gaussian filter or Gabor Wavelet filter) and extracting the object from the background.

The following steps are mainly involved for preprocessing the images.

2.2.1 Segmentation

Segmentation is the process by which objects of different shapes can be extracted from the background. The segmentation algorithms are dependent on different types of features, skin color based [20], texture based [21], saliency object based. Hybrid Saliency algorithm [8] is a very popular segmentation algorithm which is used on images with complex background to highlight the object from background and to find out the salient feature of the images. Using this algorithm the salient regions can be differentiated from the background details easily.

2.2.2 Boundary Detection

The image segmentation always not gives proper result (boundary). So, for detecting the boundary edges, many morphological operations are used. The general morphological operations of image processing are erosion, dilation, thinning, filtering and hole filling operations. For example, skin color-based segmentation algorithm contains many irregularities [10]. Using opening operations (erosion preceded (dilation operation), it is possible to remove these irregularities. The chain code generation algorithm is used for detecting the unbounded edges. Resultant boundary of texture-based segmentation contains topological error like hole. To fill up these holes different filling operations, Sobel edge detection and many others techniques are used [22].

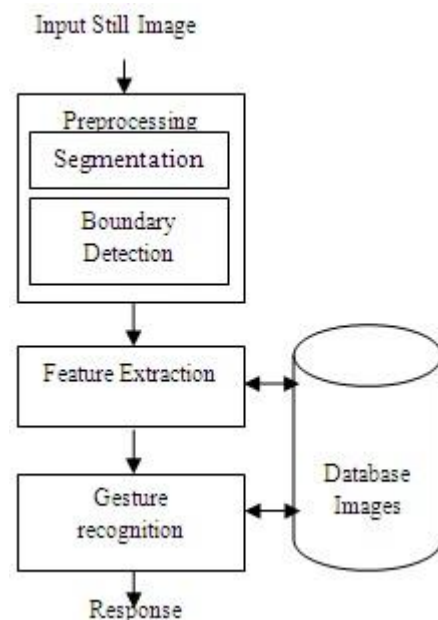


Fig 1. Generic view of Dance Gesture Recognition System

2.3. Feature Extraction

After preprocessing, the next step of recognition is feature selection. It is a process by which we can transform the input data into a set of features. The selection of feature(s) is/are not an easy task, it is very important to choose proper feature(s) vector to classify the images. Orientation histogram [10], [23],

hybrid saliency techniques [8], Steer filter [10] are some existing feature extraction techniques used to extract types of features like (e.g., area, major axis length, minor axis length, centroid and eccentricity) of mudra images can be extracted. A major issue is to find out an optimal and relevant feature vector that can help classify the images in any complex background with high recognition accuracy.



2.4. Classification and Recognition

This is the final phase of gesture recognition. The main job of this phase is to understand the images. To understand the images, it has to cross two phases: image classification and image recognition. Image classification means which class it belongs to. For example it is human hand i.e., find out the position of certain object. Image recognition means which gesture it is, i.e., recognize the pattern. The classification and recognition may be of supervised or unsupervised. To classify any images, the gesture dataset are divided into training set and testing set [8]. A major portion of the dataset are use as training set and the enduring set are used to classify or test the data [19]. The different classification techniques[24] for dance gestures are artificial neural network (ANN), decision tree (DT), support vector machine (SVM)[25], fuzzy classifier, fuzzy set theory-based techniques [25], hidden markov model [26], back propagation neural network [27], particle filtering and graph theory-based [1], fuzzy L-membership [23] and KNN(K-nearest neighbour) algorithm.

3. GESTURES TAXONOMY

Gestures are broadly classified into two types: static gestures and dynamic gestures (Table I). Apart from this broad classification, several other classifications are provided by various authors. Some of these classifications are depicted in Fig. 2 and described in the remaining part of this section. Gestures have variously been classified by different people [28], [29] in different period of time.

TABLE 1: BASIC TYPES OF GESTURES [1]

Gestures	Meaning	Example
Static gestures	Static gestures are those where the user assumes certain pose or configuration.	Stop 
Dynamic gestures	Dynamic gestures are those gesture where the three phases prestroke, stroke and post stroke are considered	Goodbye 

3.1. Efron’s Classification

Efron is considered to be the pioneer of classifying gestures. His classification is based on the linguistic aspects of the gesture behaviour and expression [28]. According to him, gestures can be of two types: 'objective' gestures which are speechless and 'logical discursive' which describes along with speech. Objective gestures are farther categorized into three types of gestures i.e., deictic, physiographic and symbolic gestures. Logical discursive gestures are sub classified into Batons gestures and Ideographic gestures

3.2. Kendon Classification

In the year 1988, Kendon classified the gestures based on the variety of requirement and speech dependency[28]. These gestures are classified into five types as i) Gesticulation gestures which are used in general conversation with utterance, ii) Speech-framed gestures are the part of the sentence itself. iii) Emblems are recognized in absence of speech, iv) Pantomime, these are used to narrate the story in absence of speech and v) Signs gestures are mostly used in sign languages. These types of gestures have their own linguistic structures, grammatical patterns, vocabulary and morphological patterns.

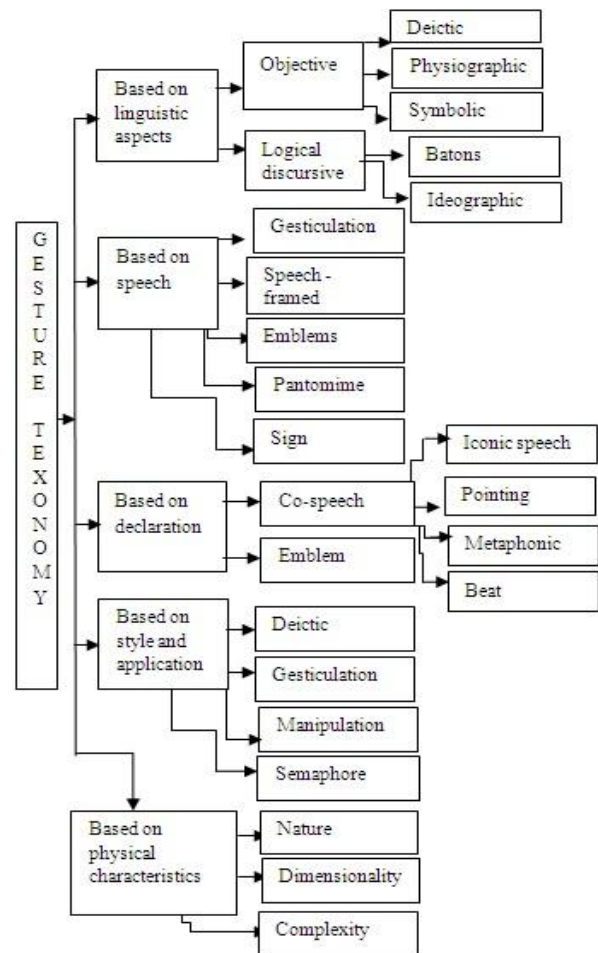


Fig 2: Gesture Classification: A Taxonomy

3.3. David McNeil Classification

According to David McNeil, gestures are classified into Co-speech gestures and Emblems gestures [29].

3.3.1 Co-speech gestures

Co-speech gestures are the type of gestures which are used along with word. These types of gestures are categorized into the following types:

3.3.1.1. Iconic

These gestures are used along with speech. Whenever we express some physical matter we use our hands to explain it more clearly. For example, if we describe about some real item like how big or small it is, then we paint our hand along with the word. Iconic gestures are different from the rest of gestures in the sense that it is used to describe real and existing matter.

3.3.1.2. *Pointing:*

The pointing gestures refer to the way people show the knowledge by pointing. One can indicate a place or perhaps a thing moving from one place to another using this gesture. For example, if we point someone across the room, then, naturally we make point our finger in the appropriate direction.

3.3.1.3. *Metaphoric*

This type of gestures can express an idea in a general way. For example, when we are describing something complex, also we wave our hand in the air. These kinds of gestures also add emotion and spice when something being said. Sometime it is more or less dramatic. Making our hand into heart shape and placing in our chest to show the affection to a loved one.

3.3.1.4. *Beat gestures*

Beat gestures are related to rhythmic beating of hand expressions. Sometime they may be described by a single beat or may continue the whole duration as long as explaining some particular point. These gestures can vary depending on situation. [29]

3.3.2. *Emblems gestures*

Emblem gestures indicate those are independent of speech, it is a specific type of gesture with specific meaning that are intentionally used and unintentionally understood. These gestures are used as a substitute of words which are similar to sign language. For example we can recognize different alphabet and numbers by using our hand expressions.

3.4. David Karam's classification

Karams classification of gestures are based on representation of gesture, application domain and input output technology[28]. These gestures are of four types: Deictic, Semaphores, Gesticulation and Manipulation

3.4.1. *Deictic*

Deictic gestures refer to pointing to identity or spatial location of an object.

3.4.2. *Semaphore*

This kind of gestures can be expressed via body part or other objects and electronic devices, such as a mouse.

3.4.3. *Gesticulation*

Gesticulations gestures are co-speech multimodal gesture, consisting of hand movement, this kind of gestures are not pre-planned.

3.4.4. *Manipulation*

This kind of gestures is also referred to as physical gestures. This class of gestures gives importance to the relation between hand movement and the object being manipulated.

3.5. Ruiz classification for 3D -motion gestures

These classifications of gestures are used in 3D-motion gestures that are applied on mobile device (smart phone). The gestures are classified based on physical characteristics [28]. The physical characters are kinetic impulse, dimensionality and complexity. These gestures are categorized as body part, handedness, hand shape and the range of motion. The four types of dimensionality for the gestures are used as i) Single-axis which are used around the single axis. ii) Double-axis gestures are used on plane i.e., on 2D surface area. iii) Tri-axis gestures are used in 3D space and works either on translational motion or rotational motion and iv) Six-axis gestures, are also used in 3D space but difference is the it

works on both types of motion, translational as well as rotational motion. The Complexity of these types of gesture may be of simple or complex, depending on the situation.

4. LITERATURE REVIEW

Research on dance gesture is still on in fast stage, however within this short period it has received a good momentum. Probably, Efron [30] is the pioneer who worked on gesture analysis in psychology. The performance of hand gestures are divided into three steps: preparation, stroke and retraction. He proved gestures as a main block for human motion in the early part of 1940. Later, Kendon continued Efron's work [31] and converted how gesture related to words in language. He also showed the first segment of human gestures and recognition. Since then, gesture analysis in dance field has come into existence.

The current research on dance field can be categorized as (i) Bharatnatyam [10], where recognition based on two level decision making system and works with single hand gesture, (ii) Odissi [25], where the gesture of whole body using kinetic sensor. Their work is carried with only eleven co-ordinates out of twenty different joints of skeleton,(iii)Bali Traditional Dance [9], which works on probabilistic grammar-based classifier, (iv) Ballet Dance [22] where multiple stage system is proposed to recognize the different dance posture, (v) Kazakh Traditional Dance [26] is basically concerned with the head gestures and many others. The details of these gesture recognition methods are described in the later part of this survey. Probably, the most important and the oldest algorithm for dance gesture segmentation is the hierarchical activity segmentation [32] algorithm. This algorithm is based on an understanding of human body instruction. It represents all the 22 physical segmentation of human anatomy and each of which has independent movement. Here, all parent segments inherit the combined characteristics of all child segments. The algorithm considered the variation of gestures from one choreographer to others. But this algorithm is limited to a finite set of pose states. Another popular algorithm found in this domain based on torso frame for human joint is known as Skeletal Tracking Algorithm (STA) [33]. The algorithm represents all the angular skeleton and mapping the skeleton motion data into a smaller set of features. The torso frame is computed as a basis of co-ordinate of the other skeleton joint. Research on classical dance ontology-based approach for designing a cultural heritage repository for digitizing the record of music and performance of Indian classical dance [34] also has been carried out. Since dance and music are stored as multimedia formats so they use multimedia web ontology language to encode the domain knowledge. They proposed an architectural framework to construct the ontology with a training data and use that for automatically annotate new instances of digital heritage artifacts. Also they develop a browsing application for semantic access of traditional collection of Indian dances. The author gives emphasis on Odissi dance, which is a subclass form of Indian classical dance. Also the MOWL language has been used which is an extension of OWL language used to ensure the compatibility with W3C standards. The author collect ICD dance video which are contributed for research. A joint team [10] also has been found worked on gesture recognition of Bharatnatyam hastas. Bharatnatyam is the most oldest dance form of Indian classical dance. This group aims to develop one prototype to recognize the 28 asamyukta hastas of Bharatnatyam in two dimensional form. They work for making self-learning facilities for dancers and to promote e-learning of Bharatnatyam across the world. In [9], emphasis has been given on Bali traditional dance. Their aim to build a robust

recognizer based on linguistic motivated method. The author use the Alergia algorithm with Symbolic Aggregation Approximation (SAX) discretization method and found 92% accuracy. The work based on important joint features like left/right foot and left /right elbow of Bali traditional dance. The authors of [25] jointly propose gesture recognition algorithm for Indian classical dance style using sensor. They made one device which generates the skeleton of human body from which twenty different junction 3-dimensional coordinates are obtained. They use a unique system and extract the features to distinguish anger, fear, happiness, sadness and relaxation. They calculate the distance between different parts of the upper human body and generate velocity, acceleration along with the angle between different angles. On the basis of that they extract twenty three features. The performance of their method is almost 86.8%. In [8], the authors recognize the mudra sequence using image processing and pattern recognition techniques and apply the result to the different expressions of classical dance. The recognition of mudra sequence can thus create language independent universal communication environment for the dance drama. Their system consists of two major components: training and testing. They use hybrid saliency technique in image to highlight the object from background and to find out the salient features of double hand mudra image. They use hypercomplex representations for static mudra representation and k-nearest neighbour algorithm for classification. The features of the different mudra are extracted and the values of these features are compared with the feature values for each mudra in the database. They apply the k nearest neighbour algorithm for classification during testing. Output of this technique is emotional description for the recognized mudra images. In [26], the author works on Kazakh traditional dance gesture recognition for detecting the head movements. They use MS kinetic camera to collect human skeleton and depth information. Tree structured Bayesian network and expectation maximization (EM) algorithm with K-means clustering was applied to calculate and classifying the poses. Finally, Hidden Markov Model (HMM) is used for recognition As per the literature review, the dance gesture recognition approach has been divided into two approaches:

(i) Glove-based approach and

(ii) Vision-based approach. We discuss each of these approaches next.

4.1. Glove-based Approach

In this approach [2], sensor devices and hand gloves are used in the image acquisition phase. It provides the co-ordinate points of skeleton and orientation. In this approach, user can directly connect with computer. However, two major limitations of this approach are: quite expensive and inefficient for working in virtual reality. The related works on dance gesture recognition in glove-based approach is summarized in Table 2. In glove-based approach, segmentation algorithm is not required. The sensor use in image acquisition gives skeletonized image automatically.

4.2. Vision-based Approach

In this approach [6], the images are captured by camera. It is a very simple approach deals with the simple image characteristics like color, texture and intensity values. A limitation of this approach is that it is not able to provide the appropriate result in some circumstances like complex processing techniques which have been used to handle background, lighting variation, indistinct color combination The related works on dance gesture recognition in vision-

based approach is summarized in Table 3.

5. EXISTING METHODS

Various methods based on soft computing, statistical modeling, computer vision, pattern recognition and image processing techniques have been used to handle the gesture recognition problem.

5.1 Statistical Methods

The statistical methods are [35] [36] broadly applied for audio signal analysis, data mining, bioinformatics and many other subjects related to science and engineering branches. It covers the basic topics like dimensionality reduction, linear classification, regression as well as other modern topics such as Support Vector Machine (SVM), Hidden Markov Model (HMM), Principal Component Analysis and Linear Discriminate Analysis. This approach usually follows the supervised learning method but sometime it may show unsupervised characteristics. In this approach, one can easily apply probability theory and decision theory. These statistical approaches crucially depend on the choice of features.

5.2. Methods Based on Finite State Machine

A finite state machine (FSM) is a machine that makes transition with a finite number of state based on some input parameter. This machine is design for both computer program as well as sequential logic circuit. One common example of finite state machine is traffic light. The traffic light has three states red, green and yellow. If the light is initially red then after certain time it becomes green or yellow based on some input signal. Considering the input signal t , the transition diagram for traffic signal will be as shown in fig:3. Here the transition is based n two input parameter 1 and 0. If $t = 1$ then it jumps to the next state or continue in the same state for $t = 0$. FSM can model a huge no. of problems like automata, gesture recognition, networking design and artificial intelligence. The recognition of gesture can be performed using online trained FSM. When feature vectors are supplied as a input to the recognizer, if it reaches the final state then the gesture has been recognized [1].

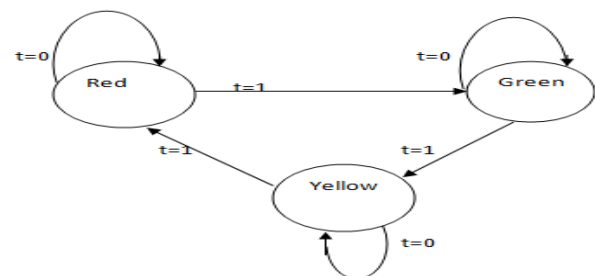


Fig 3. Finite state machine for traffic signal

5.3 Soft Computing Methods

Human mind is the main functional model of soft computing because of its uncertainty. Soft computing methods are used in the problem of computer science whose solutions are not predictable, uncertain and between 0 and 1. Soft computing problems are used when we do not have enough information about the problem itself. The main tools and techniques of soft computing method are fuzzy logic(FL), neural network (NN), support vector machine(SVM), evolutionary computing (EC), machine learning (ML), probabilistic reasoning (PR), genetic algorithm (GA) and rough sets. These techniques are very popular for gesture recognition domain. Generally, fuzzy set and rough sets are used to measure the uncertainty level of gesture. Other popular method like fuzzy rule based [17] and

NN [37] are also used for glove based hand gesture recognition and vision based hand gesture recognition respectively.

5.4 Hybrid Methods

The aims of hybrid methods are to combine the advantages of different approaches within a single system. The hybrid classification methods are very useful for gesture recognition. St Japan Rajko and Gang Qian [38] use the elements of both HMM and dynamic programming alignment (DPA) methods like edit distance, sequence alignment and dynamic time wrapping to propose a hybrid method HMM/PDA for gesture recognition. This hybrid method takes the robustness and effectiveness property of HMM and the simplicity property of DPA approaches.

Other than the above approaches, some of the most popular and frequently used models for dance gestures classification are described as follows:

5.5. Method using Hidden Markov Model

Hidden Markov Model (HMM) is a very power tool for gesture recognition. It is composed of number of state like finite state machine. The two basic state of HMM's hidden state and visible state. In HMM, the states are not directly visible, but each hidden state emits one visible state and output depends on this visible state. Each hidden state randomly generates one of n observations O_1, O_2, \dots, O_n . An HMM is represented as $M = (M_1, M_2, \pi)$. where $M_1 = (m_{ij})$, $m_{ij} = P(s_i / s_j)$ denotes to specify a matrix of transition probabilities, $M_2 = (P(Om/Si))$ represent the matrix of observation probabilities and $\pi = (P(s_i))$ is the initial probabilities[39]. The central issues of HMM is learning, if the machine is not trained properly then the recognition output never give the satisfactory result.

5.6. Method using Back Propagation Neural Network

Back Propagation Neural Network (BPNN) is one of the most popular training methods for neural networks. It is a very simple method even for complex problems which have huge amount of parameters. Generally, ANN are classified as single layer neural networks and multi layer neural networks. BPNN is under the category of multilayer networks. It has more than one or more hidden layers exist between the input and output units. Multilayer perception neural network is very important tool for classification purpose. Back propagation learning algorithm for static hand gesture recognition gives accuracy up to 86.38% [27]. A major advantage of this method is that it is based on learning rather than programming. So this method takes less time compared to programming method and no need to specify the exact behavior of the model. It is also very robust in nature i.e., flexible in changing environment. It takes some more time in case of drastic changes but it has good adjusting power for constantly changing environment.

6. FUTURE RESEARCH DIRECTION

Based on this literature survey, we have identified the following research issues and challenges.

- 1) Developing faster segmentation techniques for object in a complex background is still a challenging task.
- 2) Image of a particular posture may have wide variations due to position and angle of view of the camera. Therefore, developing a hand gesture recognition system capable of handling different types of transformations is an important research issue.
- 3) Dance poses may vary from dancer to dancer, so accurate

detection of gestures in presence of these variations is a challenging task.

- 4) There are no proper classification techniques available in literature, which can identify all the relevant groups of hand gestures as well as the subset of closely matched gestures within a subgroup.

7. CONCLUSION

Dances express human love, emotions, devotion and narrate stories of religious scriptures and are integral parts of the celebration of life. In this survey paper, the overview of gestures recognition with special emphasis on hand and dance gestures are discussed. This paper also provides a comprehensive study on vision based and glove based gesture recognition, gesture taxonomy. Finally, this survey paper concludes with existing approaches and methods to implement and scope for future work.

8. REFERENCES

- [1] Mitra, S. and Acharya, T. 2007. Gesture recognition: A survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 37(3):311–324.
- [2] Sturman, D.J. and Zeltzer, D. 1994. A survey of glove-based input. Computer Graphics and Applications, IEEE, 14(1):30–39.
- [3] Fong, T., Nourbakhsh, I. and Dautenhahn, K. 2003. A survey of socially interactive robots. Robotics and autonomous systems, 42(3):143–166.
- [4] Freeman, W.T., Tanaka, K., Ohta, J. and Kyuma, K. 1996. Computer vision for computer games. In Automatic Face and Gesture Recognition, Proceedings of the Second International Conference on, pages 100–105. IEEE.
- [5] Pradeep kumar, B.P., Santhosh, S.Y., Manjuatha, M. B. 2014. Survey on skeleton gesture recognition provided by Kinect. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 3.
- [6] Moeslund, T.B. and Granum, E. 2001. A survey of computer vision-based human motion capture. Computer Vision and Image understanding, 81(3):231–268.
- [7] Wu, Y. and Huang, T. S. 1999. Vision-based gesture recognition: A review. In Gesture-Based Communication in Human-Computer Interaction, volume 1739, pages 103–115. Springer Berlin Heidelberg.
- [8] Mozarkar, S. and Warnekar, C.S. 2013 Recognizing Bharatnatyam Mudra using Principles of Gesture Recognition. International Journal of Computer Science and Network, 2(2):46–52.
- [9] Heryadi, Y., Fanany, M.I. and Arymurthy, A. M. 2012. Grammar of dance gesture from Bali traditional dance. International Journal of Computer Science Issues (IJCSI), 9(6).
- [10] Hariharan, D., Acharya, T., and Mitra, S. 2011. Recognizing hand gestures of a dancer. In Pattern recognition and machine intelligence, pages 186–192. Springer.
- [11] Starner, T. and Pentland, A. 1997. Real-time American sign language recognition from video using hidden Markov models. In Motion-Based Recognition, pages

- 227–243. Springer, 1997.
- [12] Maraqa, M. and Abu-Zaite,R. 2008. Recognition of arabic sign language (arsl) using recurrent neural networks. In Applications of Digital Information and Web Technologies, 2008. ICADIWT 2008. First International Conference on the, pages 478–481.IEEE.
- [13] Murakami, K. and Taguchi, H. 1991. Gesture recognition using recurrent neural networks. In Proceedings of the SIGCHI conference on Human factors in computing systems, pages 237–242. ACM.
- [14] Starner, T, Weaver, J. and Pentland, A. 1998. Real-time american sign language recognition using desk and wearable computer based video. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 20(12):1371–1375,.
- [15] Kim, J., Jang, w. and Bien, Z. 1996. A dynamic gesture recognition system for the korean sign language (ksl). Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 26(2):354–359.
- [16] Cho. M.G 2006 A new gesture recognition algorithm and segmentation method of korean scripts for gesture-allowed ink editor. Information Sciences, 176(9):1290–1303.
- [17] Bedregal, B. C., C. R. Costa, A. and Dimuro, G. P. 2006 Fuzzy rule-based hand gesture recognition. In Artificial Intelligence in Theory and Practice, pages 285–294, Springer.
- [18] Campbell, R., LANDIS, T., and REGARD, M. 1986 Face recognition and lipreading a neurological dissociation. Brain, 109(3):509–521.
- [19] Saha, S, Ghosh, S. Konar, A. and Janarthanan, R. 2013. Identification of odissi dance video using kinect sensor. In Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on, pages 1837–1842. IEEE.
- [20] Phung, S.L., Bouzerdoum, A. and Chai Sr, D.2005. Skin segmentation using color pixel classification: analysis and comparison. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 27(1):148–154.
- [21] Belongie, S., Carson, C., Greenspan, H. and Malik, J. 1998 Color-and texture-based image segmentation using em and its application to content-based image retrieval. In Computer Vision, Sixth International Conference on, pages 675–682, University of California, IEEE.
- [22] Saha, S., Banerjee, A. Basu, S., Konar, A. and Atulya, K. N. 2013. Fuzzy image matching for posture recognition in ballet dance. In Fuzzy Systems (FUZZ), 2013 .IEEE International Conference on, pages 1–8.
- [23] Saha, S., Ghosh, L., Konar, A. and Janarthanan, R. Fuzzy l membership function based hand gesture recognition for bharatanatyam dance. In Computational Intelligence and Communication Networks (CICN), 2013 5th International Conference on, pages 331–335. IEEE, 2013.
- [24] Kamavisdar, P., Saluja, S. and Agrawal, S. 2013. A survey on image classification approaches and techniques. International Journal of Advanced Research in Computer and Communication Engineering, 2(1).
- [25] Konar, A. Atulya, K. N., Saha, S. Shreya Ghosh, S. 2013. Gesture recognition from indian classical dance using kinect sensor. In Fifth International Conference on Computational Intelligence, Communication Systems and Networks
- [26] Nussipbekov, A.K., Amirgaliyev, E.N and Hahn, M. 2014. Kazakh traditional dance gesture recognition. In Journal of Physics: Conference Series, volume 495, page 012036. IOP Publishing.
- [27] Hasan, H. and Abdul-Kareem, S.2014. Static hand gesture recognition using neural networks. Artificial Intelligence Review, 41(2):147–181.
- [28] Vafaei F.2013. Taxonomy of Gestures in Human Computer Interaction. PhD thesis, North Dakota State University.
- [29] McNeill, D. 1992. Hand and mind: What gestures reveal about thought. University of Chicago Press.
- [30] Efron.,D 1941 Gesture and environment: A tentative study of some of the spatio-temporal and” linguistic” aspects of the gestural behavior of eastern Jews and southern Italians in New York city, living under similar as well as different environmental conditions. King’s crown Press, New York.
- [31] Canada Kendon A., Potyados, F. 1988. How gestures can become like words. cross cultural perspectives in nonverbal communication. pages 131–141.
- [32] Kahol,K.,Tripathi, P. and Panchanathan, S. 2004. Automated gesture segmentation from dance sequences. In Automatic Face and Gesture Recognition, In Automatic Face and Gesture Recognition. Proceedings. Sixth IEEE International Conference on, pages 883–888. IEEE.
- [33] Raptis, M., Kirovski, D. and Hoppe, H. 2011. Realtime classification of dance gestures from skeleton animation. In Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pages 147–156. ACM.
- [34] Mallik, A., Chaudhury, S. and Ghosh, H. 2011. Nriyakosha: Preserving the intangible heritage of indian classical dance. Journal on Computing and Cultural Heritage (JOCCH), 4(3):11.
- [35] Minka, T. 2005. A statistical learning/pattern recognition glossary, Retrieved June, 29:2008.
- [36] Hong, P., Turk, M. And Huang, T. S 2000. Gesture modeling and recognition using finite state machines. In Automatic Face and Gesture Recognition, 2000. Proceedings Fourth IEEE International Conference on, pages 410–415. IEEE.
- [37] Ibraheem, N.A. and Khan, R.Z. 2012. Vision based gesture recognition using neural networks approaches: A review. International Journal of human Computer Interaction (IJHCI), 3(1):1–14.
- [38] Rajko,S. and Qian, G. 2005. A hybrid HMM/DPA adaptive gesture recognition method. In Advances in Visual Computing, pages 227–234. Springer.
- [39] Sharma, A. Recognising bharatanatyam dance sequences using rgb-d data. Master’s thesis, IIT, Kanpur.

Table 2: Study on Glove-based Dance Gesture Recognition

Dance	Features/Parameter	Classifier	Accuracy (percentage)	Limitation
Oddissi[9]	Vertices of 4-side polygon, edge length and angle between different vertices.	multi-class SVM	92.7	Dancer should stand within a range.
Oddissi[25]	A total of 23 features have been extracted based on distance, velocity, acceleration, and angle measurement.	SVM	86.8	Dancer should stand within a range.
Bali traditional dance[9]	Skeleton features are measured by (Θ, Φ) where Θ denotes inclination and Φ denotes azimuth. Features with highest Cophenet Correlation Coefficient (CCC) value is considered.	Probabilistic grammar based classifiers.	92	Tested with limited set of samples and poses.
Kazakh Traditional Dance[26]	Skeleton features: inclination and azimuth angle values of particular joint vectors are considered. Joints are- Right Hand (RH), Left Hand (LH), Right Elbow (RE), Left Elbow (LE) and Head joint.	Bayesian network, hidden markov model, EM-approximation algorithm, k-means clustering.	90.82	Focus only on head movement.

Table 3: Study on Vision-based Gesture Recognition

Dance	Segmentation Techniques	Features/Parameter	Classifier	Accuracy (percentage)	Limitation
Bharat Natyam [10]	1. Skin color based 2. Morphological operation (thinning).	Geometric features histogram orientation, gradient at corner point of skeleton and cost function.	Two-level decision making system	No mentioned	1. Miss classification due to irregularities in skeleton structures (different people have different textural properties) 2. Fingertips of hands are coated with red dye during dance performed. The red tips of the fingers are erased after segmentation.
Bharat Natyam[8]	Saliency detection techniques	Area, major axis length, minor axis length and ccentricity of each mudra image.	K-nn classifier.	85.29	Miss classification chances are more
Ballet Dance [22]	Skin color based segmentation and morphological operation are used	Distance of Straight line, Fuzzy membership values	Fuzzy T-norm	82.35	1. Skin color segmentation algorithm produce sub optimal results when dressing and skin color matched. 2. Performance of the algorithm drops in case of postures containing bounded regions.
Bharat Natyam[23]	Texture based segmentation and sobel edge detection technique	Fuzzy L Membership values, Euclidean distance	Summation of fuzzy membership value and similarity function.	85.1	Algorithm is suitable for simple background images only.