

# Gesture Recognition: A Survey

Bharti Bansal  
Assistant Professor  
Amity University Haryana

## ABSTRACT

With increasing use of computers in our daily lives, lately there has been a rapid increase in the efforts to develop a better human computer interaction interface. The need of easy to use and advance types of human-computer interaction with natural interfaces is more than ever. In the present framework, the UI (User Interface) of a computer allows user to interact with electronic devices with graphical icons and visual indicators, which is still inconvenient and not suitable for working in virtual environments. An interface which allow user to communicate through gestures is the next step in the direction of advance human computer interface. In the present paper author explore different aspects of gesture recognition techniques.

## General Terms

Pattern Recognition, Gesture Recognition, Machine learning.

## Keywords

Gesture, HMM,

## 1. INTRODUCTION

Gesture recognition is an emerging field of computer science which focuses on the problem of segmentation and interpretation of the human gestures. It can be seen as a process in which receiver will recognizes the gesture made by a user. In that sense gesture recognition provide a natural mean for human computer interaction.



Figure 1 : A few static hand gestures

Gesture recognition can be seen as identifying the motion or posture of a body part for example hands, fingers, face expressions etc. Gesture can be static or dynamic or can have the elements of both. Figure 1 show few static hand gestures. The most familiar cultural specific gestures are called quotable gestures which are basically used as replacement of words like a hand wave can signify a "Hello" as well as "Good Bye". Some gestures (beat gestures) are used with conjunction of the speech to emphasis certain words. Some are full of more content like the shivering or the rubbing of hands will be interpreted as the one is feeling cold.

Gesture can be classified as follows:

- 1) Gestures using face and head: recognition of facial gestures and movement of head. Examples: i) use of eyes ii) Expression of different emotions like anger, fear, surprise etc using face iii) nodding of head.
- 2) Gestures using hands and arms: recognition of gestures involving movements of hands and arm [4]. Examples: i) sign language ii) hand poses.
- 3) Gestures using full body: recognition of gestures involving full body. Example: dance movements.

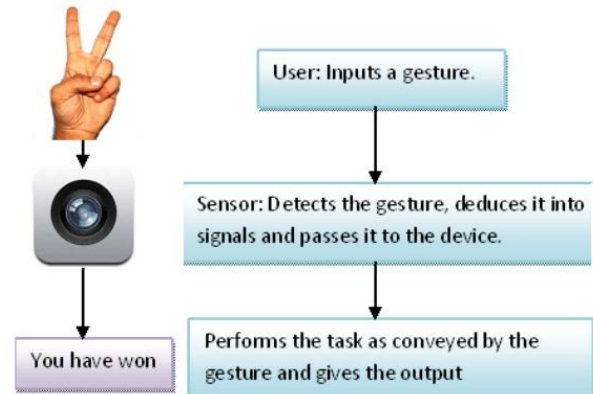


Figure 2 : Gesture Recognition Process

The process of gesture recognition can be seen as a three way process which involve identification of gesture and segmentation the gesture and matching of gesture. The gesture identification involves finding out the meaningful gesture in the input signal stream and then applying the techniques to segment that gesture. The complexity of the system lies in the time and shape variability of the gesture. From application point of view gesture recognition has found it's application in the field of Sign language recognition [2, 5], Immersive game Technology [8], Aid to physically challenged, Remote Control and Socially assistive robotics etc.

In this paper, author provides a survey of different aspects of gesture recognition. In section 2, author outlined the different sensing technology used to track the gesture. Then in section 3, author explored different algorithms used for gesture recognition briefly. In section 4, author explained one of the popular tools for gesture recognition. Author conclude in Section 5 with challenges involve and future scope in gesture recognition.

## 2. GESTURE SENSING TECHNOLOGY

Gesture tracking is based on identifying motion, posture and position of human body [6]. Technology being used in gesture recognition can be seen as of two types: contact based and non-contact based. In contact based sensing technology touch screen is used to track the movement. While non-contact based can be device based, vision based and electrical sensing based technology. In device based techniques a device equipped with variety of sensors is used like stylus, gloves etc to send the signals to the system. Vision based technology having two approaches: Model based technique, which try to create a 3D model and Image based methods, which detects the gesture by using pictures. In electric field sensing by sensing electric field propinquity of body part can be measured, which is used to measure the distance between human body and object.

There are various tools which are used as input devices in gesture recognition. These tools have the ability to track the

movement of human body. Following is a list of some of such tools.

- Wired Gloves:** Also called the data gloves [12] are devices which can track the position and the rotation of the hands. Base on the property of tracking devices, Data gloves can be used to provide a varied degree of accuracy.
- Controller based gesture:** These controller works as an extension of human body. Whenever gesture is performed, motion can be captured by the software. For example: AudioCubes, A collection of wireless light emitting object capable of detecting the gestures.
- Camera:** A standard camera is a suitable tool for gesture recognition in an environment which is not convenient for other form of image based recognition. Several companies are embedding the software for gesture recognition using 2D cameras for detecting the gestures, hand sign etc.
- Stereo Camera:** A camera having two lenses, with a known physical relationship. It will take two pictures at a time. By using the output of two cameras whose relationship is known, a 3D representation can be generated.
- Depth aware camera:** These are suitable for short range that's why very effective for detection of hand gestures. One can generate the depth map of the image taken by the specialized cameras such as structured light cameras.
- Thermal cameras:** Infrared radiation emitted by the target object is detected by the camera and convert those radiation in electronic signal to generate a thermal picture or to make temperature calculations. This temperature can be quantified.

### 3. APPROACHES AND TOOLS FOR GESTURE RECOGNITION

The problem of gesture recognition is not specific to a particular domain. Researchers from different domain have tried to tackle the problem using the tools and techniques available from different domains. In past approaches from computer vision, pattern recognition, signal processing, statistical modeling and particle filtering has been applied to solve gesture recognition problem. Tools available from different domains are PCA, HMMs [9, 7], particle filtering, Time-delay neural networks (TDNN) to name a few. From modeling point of view, Different approaches for gesture recognition can be classified into following classes: 3D model-based approaches and Appearance-based approaches. We discuss both approaches in this section followed by popular HMM model for gesture recognition.

#### 3.1 Approaches for gesture recognition

The movement of arms and hands in 3d space is observed as gesture. In order to interpret the gesture we have to classify movements of body according to common properties and the message they may express. For example in sign languages every word is represented using a gesture [3]. There are different approaches for interpreting the gesture depending upon the type of input data. Two different approaches are used in gesture recognition: 3D based model and appearance based model [6].

- 3D based model:** In 3D based model the key elements of body parts are represented as 3D information and this information is used to interpret the several parameters

like position of palms. It can use volumetric model, skeletal model or both. Complicated 3D surfaces are used to create model in volumetric model approach. The simplicity of the parameters for these objects is the key advantage of this approach but is requires an intensive computation which is treated as a drawback. In skeletal model the parts of the human body are mapped with a virtual skeleton to certain segments. Using the orientation and position of these segments the analysis is performed. The key advantages of skeletal models are:

- Due to the use of template database, pattern matching becomes easy.
  - Only key parameters are analyzed as a result algorithm becomes faster.
- Appearance based model:** Instead of spatial representation of the body Images or video is used for interpretation in appearance based model. A template database is created for the parts of human body using images or video. In a different approach in appearance model image sequences are used as a gesture template.

#### 3.2 HMM for gesture recognition

If in a system the conditional probability density of a particular event depends only on a fixed number of past events, say  $N$ , then system is said to be following Markov model. If  $N$  is 1 then we call the system first order Markov model. There are a number of variant of Markov model, like Hidden Markov model (HMM), which are being used in different applications pertaining to temporal pattern recognition such as speech, handwriting and bioinformatics. A typical Hidden Markov model involve

- A set of states, Denoted by be  $S$ .
- Initial and Final states, denoted by be  $S_i, S_f$ .
- A matrix representing probability of transition from one state to another state.
- An output probability matrix which capture the probability of getting a particular output.

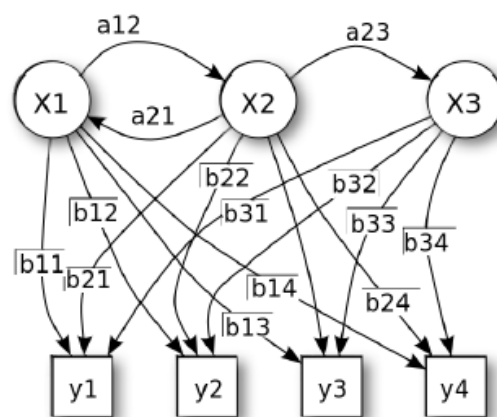


Figure 3 : General Hidden Markov Model

The Figure 2 shows an example of HMM. For detailed introduction on HMM we refer reader to [11].

The motivation behind using HMM for gesture recognition is the applicability of HMM in applications related to speech processing. HMM is an efficient tool for modeling process

involving transition or motion so HMM is a natural choice for gesture recognition. A typical HMM based gesture recognition approach can be seen as consisting of following stages:

- First a set of gesture need to be defined. These are the gestures one expects to find in the input data.
- The next step is to represent these gestures in term of HMM. A typical representation would contain the states and transition information of gestures.
- Then data for training need to be collected.
- The training data is then passed through the HMM which will compute a set of parameters. These parameters, generally, depends on the input data.
- Then the gestures are evaluated using the above trained models.

In past a number of different has been proposed based on HMM for gesture recognition. The work by Yamato et al. [10] where HMM was used to recognize different classes of tennis strokes is among then earliest work in the field of gesture recognition using HMM. Complexity of HMMs is high because of the time needed to train system. A mixture of HMMs and FSMs was proposed to increase the efficiency of gesture recognition systems. There are many other tools available for gesture recognition like FSMs, FACS, contour models, wavelets, optical flow, Connectionist Approach to name a few. The introduction to all approaches is out of scope of this paper. Author refers interested reader to [3].

#### **4. CONCLUSION AND FUTURE SCOPE**

In this paper author provided a brief overview of different aspects involved in gesture recognition. The application of gesture recognition varies from sign language, facial expression matching and lip movement tracking to virtual gaming etc.

As the gesture is culture specific a gesture can be interpreted differently. Gesture varies from user to user. The interpretations of same gesture can be different for different users. Due to which identification of a gesture is a hard task. Background and environment conditions like lighting conditions, noise etc plays a vital role in gesture tracking. Due to these factors gesture recognition system do not track the gesture properly. The storage and retrieval of heterogeneous multimedia data is also very important. Due to the increasing pictorial information we need compression techniques and analyzing software for the multimedia database.

#### **5. REFERENCES**

- [1] Manuel Caputo, Klaus Denker, Benjamin Dums, and Georg Umlauf. 3d hand gesture recognition based on sensor fusion of commodity hardware. In Harald Reiterer and Oliver Deussen, editors, *Mensch Computer*, pages 293{302. Oldenbourg Verlag, 2012.
- [2] D. Kumarage, S. Fernando, P. Fernando, D. Madushanka, and R. Samarasinghe. Real-time sign language gesture recognition using still-image comparison amp; motion recognition. In *Industrial and Information Systems (ICIIS)*, 2011 6th IEEE International Conference on, pages 169{174, August 2011.
- [3] S. Mitra and T. Acharya. Gesture recognition: A survey. *Trans. Sys. Man Cyber Part C*, 37(3):311{324, May 2007.
- [4] G. R. S. Murthy and R. S. Jadon. A review of vision based hand gestures recognition. *International Journal Of Information Technology and Knowledge*, pages 2-2009.
- [5] M. Panwar. Hand gesture recognition based on shape parameters. In *Computing, Communication and Applications (ICCCA)*, 2012 International Conference on, pages 1-6, Feb 2012.
- [6] Vladimir I. Pavlovic, Rajeev Sharma, and Thomas S. Huang. Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:677-695, 1997.
- [7] L. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257-286, Feb 1989.
- [8] S.S. Rautaray and A. Agrawal. Interaction with virtual game through hand gesture recognition. In *Multimedia, Signal Processing and communication Technologies (IMPACT)*, 2011 International Conference on, pages 244-247, Dec 2011.
- [9] Ferdinando Samaria and Steve Young. Hmm-based architecture for face identification. *Image and vision computing*, 12(8):537{543, 1994.
- [10] J. Yamato, Jun Ohya, and K. Ishii. Recognizing human action in time-sequential images using hidden markov model. In *Computer Vision and Pattern Recognition*, 1992. *Proceedings CVPR'92*, 1992 IEEE Computer Society Conference on, pages 379-385, Jun 1992.
- [11] Jie Yang and Yangsheng Xu. Hidden markov model for gesture recognition. Technical Report CMU-RI-TR-94-10, Robotics Institute, Pittsburgh, PA, May 1994.
- [12] Thomas G. Zimmerman, Jaron Lanier, Chuck Blanchard, Steve Bryson, and Young Harvill. A hand gesture interface device. *SIGCHI Bull.*, 18(4):189-192, May 1986.