A SURVEY ON GESTURE RECOGNITION TECHNOLOGY

Deeba Kazim[1] , Mohd Faisal[2]

[1]*MCA Student, Integral University, Lucknow (India)*

[2] *Assistant Professor, Integral University, Lucknow (india)*

ABSTRACT

Gesture recognition is a technology through which inputs are given to the devices through human gestures. It is a human-machine interaction. By implementing the gesture recognition technology a device can identify and respond to the various gestures of a human. Gestures are movements or positions of the person's hands, face, fingers, to convey any information. These are non-verbal communication. Hand gestures can be mainly used for expressing ones ideas. This hand recognition approach can be mainly divided into Data-Glove Based and Vision Based approaches. Facial Gesture Recognition uses the different face expressions to interact with the machine. Nose tip being the highest point on the face can be considered as an important feature to indicate the head pose.

Keywords*: *Facial gesture, Gestures, Hand gestures, Nose tips.

I. INTRODUCTION

Gestures are non-verbal means of communication which includes the movements or positions of the person's hands, face or fingers to convey any information. Its basic process involves user, sensor and device. The user inputs the gestures. The sensor detects the gesture, deduces it into signals and passes it to the device; the device performs the task as conveyed by the gesture and gives output. For example, a person placing his hand as shown in fig (1) stops a running process when the gesture is fed through a camera attached to the computer.

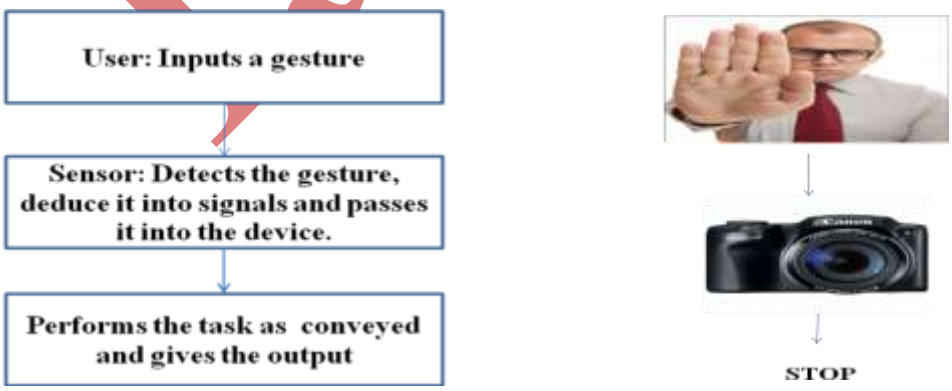

Fig 1: The process of gesture recognition.

1.1 Input Devices

These are the devices which take gestures (positions and movements) as inputs and convey the signals to the processing unit to deduce the information. In the beginning humans interacted with the machine through input devices such as mouse and keyboard. With the advent of technology there felt a need to interact with the machine as humans convey information to each other, through speech and gestures.

Various devices which make natural hand gestures to be taken as input by the machine are:

1.1.1 Wired Gloves

These have magnetic or internal tracking devices that provide the information about the position and rotation of the hands as the input to the computer. One such device is Data Glove which uses fiber optic cables running down the back of the hand as shown in the figure 2. It could detect the bending of fingers, hand position and movements.

Fig 2: Data Glove

1.1.2 Depth-aware Cameras

These are special cameras that can generate a depth map and use the data to generate 3D representation of what is being seen. A depth map is generated, which is basically a rough 3D representation of what the camera sees within a short range. The depth aware cameras can detect the gestures within short range.

1.1.3 Stereo Cameras

Stereo cameras have two lenses about the same distance apart as human eyes. It simulates the way humans actually see and creates 3D effect when viewed. Its functioning is very much similar to the human brain which takes the stereoscopic information from the eyes and decides how far or near any given object is located.

1.1.4 Thermal Cameras

These cameras convert the target object into electronic signals by detecting its radiation (temperature). Also known as thermographic or infrared camera forms images using the infrared radiations. Infrared cameras operate in wavelengths as long as 14,000 nm.

II. RELATED WORK

The user's cultural background can be utilized for analyzing the gesture expressivity in a model based on cultural dimensions [1]. Users make different gestures for the same purpose differently according to their cultural background. Thus in future gesture recognition technology can be used to predict ones culture through his gestures.

The most important features of human face are eyes and nose. The detection of eyes and nose is important for various facial gesture recognition algorithms. Eyes and nose can be detected computationally. One of its approaches is by using template based methods.

In the template based methods, a generic eye model, based on the eye shape, was designed at first [2]. Template matching is then used to search the image for the eyes. While these methods can detect eyes accurately, they are normally time-consuming.

Latest work related to gesture recognition is **'Sixth Sense'**. It is a wearable gestural interface that converts the physical world into digital information. It bridges the gap digital information and hand gestures. Using it we can access, change and data very easily, without the use of keyboard and mouse.

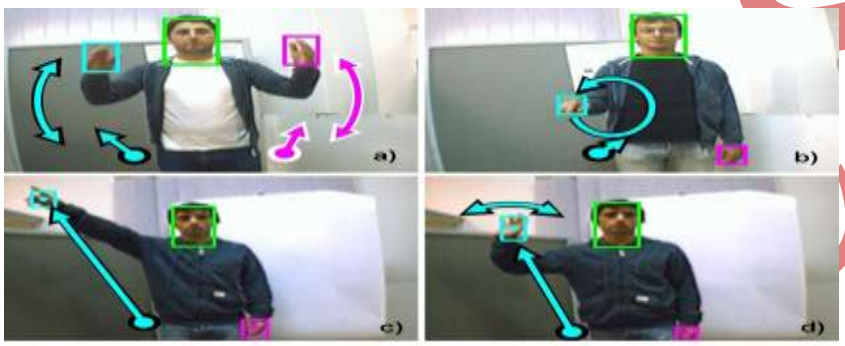

Fig 3: a, b , c and d shows different hand movements.

2.1 Hidden Markov Model

This method [3] deals with the dynamic aspects of gestures. Gestures are extracted from a sequence of video images by tracking the skin-color blobs corresponding to the hand into a body– face space centered on the face of the user. Hidden Markov Models (HMMs) are popular tools for modeling a large time series data and for generating sequences of dependent observations.

2.2 YUV Colour Space and CAMSHIFT ALGORITHM

This method deals with recognition of hand gestures [4]. It is done in the following five steps:

a. First, a digital camera records a video stream of hand gestures.

b. All the frames are taken into consideration and then using YUV colour space skin colour based segmentation is performed. The YUV colour system is employed for separating chrominance and intensity. The symbol Y indicates intensity while UV specifies chrominance components.

c. Now the hand is separated using CAMSHIFT algorithm .Since the hand is the largest connected region, we can segment the hand from the body.

d. After this is done, the position of the hand centroid is calculated in each frame. This is done by first calculating the zeroth and first moments and then using this information the centroid is calculated.

e. Now the different centroid points are joint to form a trajectory .This trajectory shows the path of the hand movement and thus the hand tracking procedure is determined.

2.3 Vision Based Hand Gesture Recognition

In vision based hand gesture recognition system [5], the movement of the hand is recorded by video camera(s). This input video is decomposed into a set of features taking individual frames into account.

III NOSE DETECTION AND RECOGNITION

3.1 Using Hough Transform

In this method (Hough Transform [6]), the face vertexes in proximity of the found sphere are extracted, and magnifying the extracted cloud of a factor of 10, in order to center the sphere on the nose. The mean value of normal vectors of each face vertex is evaluated with the purpose of obtaining information about face orientation in the three-dimensional space.

3.2 Effective Energy Calculation

Since the Nose is the protruding area on the facial front, this method uses Edge Detection techniques [7] to identify potential nose tip candidates by investigating the neighboring pixels. The nose tip is then identified from the potential candidates by their effective energy, mean and variance values obtained from Principal Component Analysis Technique. This method has the ability to locate the nose tip in Frontal as well as Non-Frontal Faces.

3.3 Template Based Method

In this approach the convex shape of the nose is used to identify the nose region. It involves calculating the local search area which was used as a template. The template was scanned through the image and the pixel and one which had maximum correlation was identified. The position of best match was refined using Evident Based Convolution Filter [8].

3.4 Colour Data Model

The human facial features such nose tip and the nose ridge were identified using a shading technique on the colour data of the image. But, the highly sensitive nature of the nose region to light made the system disadvantageous.

IV CHALLENGES

There are many challenges related to gesture recognition software. The usefulness and accuracy are the main challenges associated. In order to capture human gestures by visual sensors, robust computer vision methods are also required.

4.1 Latency: Gesture recognition software can be slow creating unacceptable latency for video games and other similar applications. In real life, humans do very fast movements of hands and facial parts which is a big challenge for the software to grasp accurately and efficiently.

4.2 Robustness: The input devices do not read accurately. Different people make different signs and gestures for the same thing differently. Thus it is very difficult for the software to read and act accordingly.

4.3 Performance: Gesture recognition is resource intensive and may be difficult to run on resource constrained devices. The gesture recognition software requires complex devices and algorithms to perform according to the input.

4.4 Difficult to develop**:** The system for recognizing gestures are complex, thus difficult to develop.

4.5 High cost of implementation: The implementation of the software is expensive.

IV CONCLUSION

This paper provides information of the gesture recognition technology. This technology can be used in various spheres of life. It can be used in Sign language recognition, Remote controllers, aid to physically challenge and in gaming. Gesture recognition is a developed technology which identifies human position and actions.

V FUTURE SCOPE

In the near future we will be seeing the use of gestures in almost all the interactions with the system. It will play an increasingly important role in human-computer interaction in the future. Reduced complexity will provide less computation time so that system works in real time. Gesture recognition can be a boon for the physically challenged people, who by using simple finger gestures can control their surroundings. Various fields where we can extensively apply gesture recognition technology in the near future are:

1) Robotics
2) Directional indication through pointing
3) Sign Language
4) Games
5) Desktop and Tablet PC Applications
6) Monitoring automobile driver's alertness/drowsiness levels etc.

REFERENCES

[1] Matthias Rehm, Nikolaus Bee, Elisabeth André "Wave Like and Egyptian: Accelerometer Based Gesture Recognition for Culture Specific Interactions" Journal of science & Technology, Sep 2006.

[2] Kun Peng et.al, "A robust algorithm for eye detection on gray intensity face without spectacles", Journal of Computer Science & Technology, Oct, 2005

[3] Chih-Ming Fu et.al, "Hand gesture recognition using a real-time tracking method and hidden Markov models", Science Direct – Image and Vision Computing, Volume 21, Issue 8, 1 August 2003, pp.745-758

[4] Harshith.C, Karthik.R.Shastry, Manoj Ravindran,.S Srikanth, Naveen Lakshmikhanth, "Survey on Various Gesture RecognitionTechniques for interfacing machines based on ambient intelligence", International Journal of Computer Science & Engineering survey (IJCSES) Vol.1, No.2, November 2010

[5] Pragati Garg et.al, "Vision Based Hand Gesture Recognition", World Academy of Science, Engineering and Technology, pp.1-6 (2009)

[6] Vitoantonio B et.al, "Extending Hough Transform to a Points' Cloud for 3D-Face Nose-Tip Detection", Advanced Intelligent Computing Theories and Applications with Aspects of Artificial Intelligence Lecture Notes in Computer Science, 2008, Volume 5227/2008, pp. 1200-1209.

[7] Wei Jen Chew et.al, "Nose Tip Detection on a Three-Dimensional Face Range Image Invariant to Head Pose", International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18 - 20, 2009, Hong Kong.

[8] Jan-Mark Geusebroek et.al,"Fast Anisotropic Gauss Filtering", Lecture Notes in Computer Science, 2002, Volume 2350/2002,pp.99-112.

Biographical Notes

Ms Deeba Kazim is presently pursuing MCA final year in Department of Computer Application, Integral University, India.

Mr. Mohammad Faisal is working as an Assistant Professor in Department of Computer Application, Integral University, India.