HUMAN COMPUTER INTERACTION USING HAND GESTURE

Nalluri Veda Kumar[1] Vidyadhar[2]
[1]Electronics& Communication Engineering, Gandhi Engineering College (GEC, Bhubaneswar), India
[2] Electrical & Electronics Engineering, Gandhi Engineering College (GEC, Bhubaneswar), India

Abstract

In this modern age the advancement in ubiquitous computing has made the use of natural user interface very much required. The presence of computers and making use of the facilities of human computer interaction in our societies will obviously bring and mark a positive impact on our societies. Either it was the day when the technologies had not been so advanced or todays when the technologies has been advanced so much that we spent most of our times to communicate, play, do our jobs with the machines and many more, even then human beings had used and are still using a broad range of gestures to communicate or interact with each other. Human gesture is a mode of non - verbal interaction medium and can provide the most intuitive, originative and natural way to interact with computers. Our main goal is to make the interaction between human and computer as natural as the interaction between humans. The objective of this paper is to recognize the static hand gesture images (i.e. frames) based on shapes and orientations of hand which is extracted from input video stream recorded in stable lighting and simple background conditions. We can use this vision based recognized gestures to control multimedia applications (like Windows Media Player, Windows Picture Manager, VLC Player etc.) running on computer using different gestural commands.

Keywords: Gestures recognition; Gesture technologies; Human-computer interaction; Static hand gesture; Vision-based gesture recognition.

1. Introduction

With the massive influx and advancement of technologies, a computer system has become a very powerful machine which has been designed to make the human beings' tasks easier. Due to which the HCI (human – computer interaction) has become an important part of our lives. Now-a-days, the progress and development in interaction with computing devices has increased so fast that as a human being even we could not remained left with the effect of this and it has become our primary thing. The technologies has so much surrounded us and has made a place in our lives that we use it to communicate, shop, work and even entertain ourselves[1]. There are many applications like media player, MS-office, Windows picture manager etc. which require natural and intuitive interface. Now-a-days most of the users uses keyboard, mouse, pen, Joysticks etc. to interact with computers, which are not enough for them. In the near future, these existing technologies which are available for the computing, communication and display will become a bottleneck and the advancement in these technologies will be required to make the system as natural as possible.

Nevertheless the invention of mouse and keyboards by the researchers and engineers has been a great progress, there are still some situations where interaction with computer with the help of keyboard and mouse will not be enough.

This is the case with the advancement in hand held devices like mobiles or i-pods or Tabs which are relatively very small in size. It's very difficult to interact with them due to their determined input spaces and small touch screen or keyboard. This is also the case of interacting 3D objects where these devices are incompatible for HCI.

One long-term goal in HCI has been to migrate "natural" means that human used it to interact with each other. With this goal human speech recognition was the area of research for a decade. It has made a tremendous progress in its field. However from the recent years there has been an increased extent in trying to introduce other human-to-human communication modalities in HCI. So human hand gestures can provide a natural and visceral alternative to some incompatible devices. We can use hand as a device to interact and communicate with computers as we do in our daily lives to interact with each other. We use our hand to point a person or an object, express or carry information about something, and to move, modify and transform an object. Exactly the same way we can use our hand to gesticulate while speaking to convey ideas.

It is required to provide a way to explore the use of gestures in HCI so that it can be interpreted by computers. The static and/or dynamic form of gestures of human arm, hand and even some other body parts require to be measurable by machine for the HCI interpretation. To facilitate and accomplish the advanced interaction between humans and computers, the designing of some special input devices has been found to be of great care in this area. The aggregation of traditional devices (i.e. keyboard, mouse etc.) with the new designed interaction devices such as face and gesture recognition, haptic sensors[2], and tracking devices provides flexibility in Virtual Reality (VR)[3], cars system control[4], Tele-operating[5], robot control[6], text editing[5], gesture recognition, video games[5], and multimedia interfaces[5].

The motivation behind this research is to make an interaction between human and computer using various applications running on computer by aiming basic shapes made by hand. Our hand movements have an important role while interacting with other people, as they convey very rich information in many ways. According to this thought hand gestures would be an ideal option for expressing the feelings, or controlling the dynamic applications of computers through easier hand gesture.

In compare to other body parts, human hand which has been considered as a natural means for human to human interaction, has been used widely for gesturing and can be best suitable for communication between human and computer[1]. There are several typical applications of hand gesture recognitions such as virtual game controller[7], sign language recognition[8], Directional indication through pointing, making young children to interact with computer, human computer interaction[8], robot control[8], lie detection[8] etc. The increasing interest in this field has made the researchers to do a large number of research which has been endured in a number of surveys given in[1]. These surveys are directly or indirectly related with hand gesture recognition.

We have organized this paper by breaking it into main components as follows: Section 2 provides some of the important related works done in this field. Section 3 discusses the methodology we have proposed for the gesture recognition system. Section 4 presents details about the experimental setups. The results and discussion for the attempt we have taken has been presented in Section 5. Finally the conclusion and future scope of the present approach has been submitted in the Section 6.

2. Related Work

In computer science and language technology, gesture recognition is an important topic which interpret human gesture through computer vision algorithms. There are various bodily motion which can originate gesture but the common form of gesture origination comes from the face and hands. The entire procedure of tracking gesture to their representation and converting them to some purposeful command is known as gesture recognition[1]. Various technologies has been used for the design and implementation of such kind of devices, but contact based and vision based technologies are two main types of technologies used for robust, accurate and reliable hand gesture recognition systems. Contact based devices like accelerometers[7], multi-touch screen, data glove [9] etc. based on physical interaction of user who will be required to learn their usages. Whereas vision based devices like cameras has to deal with the prominent variety of gestures. Gesture recognition involves to handle degrees of freedom[4, 10] (DOF), variable 2D appearances, different silhouette scales (i.e. spatial resolution) and temporal dimension (i.e. gesture speed variability). Vision based gesture recognition further classified into two main categories, which are 3D model based methods and appearance based methods[1]. 3D based hand models[4] describes the hand shapes and are the main choice of hand gesture modeling in which volumetric analysis is done. In appearance based models[4], the appearance of the arm and hand movements are directly linked from visual images to specific gestures. A large number of models belong to this group. We have followed one of these models i.e. silhouette geometry based models to recognize the gesture in our project.

In[6], a fast, simple and effective gesture recognition algorithm for robot application has been presented which automatically recognizes a limited set of gestures. However, the segmentation process should be robust and required to be deal with temporal tracking, occlusion and 3D modelling of hand. The author of[7] has used multi-stream Hidden Markov Models (HMMs) consisting of EMG sensors and 3D accelerometer (ACC) to provide user friendly environment for HCI. However, there are some problems or limitations in ACC-based techniques and EMG measurement. In[11], a method has been proposed which firstly store the human hand gesture into the disk, convert them into binary image by extracting frame from each video one by one and then creates 3D Euclidian space for binary image, for recognizing vision-based hand gesture. They have used back propagation algorithm and supervised feed-forward neural network based training for classification. However it is suitable for only simple kind of gesture against the simple background. In[12], a method for detecting finger from the detected hand, can be used as a non-contact mouse, has been proposed. They have used skin color technique for segmentation and contour as the feature to locate the fingertip in hand. The authors in[13] have used bag-of-features and multiclass SVM to detect and track bare hand, and to control an application using command generated by a grammar in a complex background, via skin detection and contour comparison algorithm. They

have also used K-means clustering algorithm and scale invariance feature transform (SIFT) to extract the main features from the trained images. However, the segmentation and localization method is unclear for the system and there is no rigorous geometric information of the object components. In[14], the author has used Lucas KanadePyramidical Optical Flow algorithm to detect moving hand and K-means algorithm to find center of moving hand. Here Principal Component Analysis (PCA) was used to extract features and then the extracted features were matched using K-nearest neighborhood. However, PCA made whole system slower and required more memory. In[15], a comparative analysis of different segmentation techniques and how to select an appropriate segmentation method for the system have been presented. It has also described Gaussian Model Classifier along with some other classification techniques.

3. Methodology

The system we have proposed and designed for vision-based hand gesture recognition system contained various stages which we have explained through an algorithm. The working flow-chart of gesture recognition system has also shown in Fig. 1.

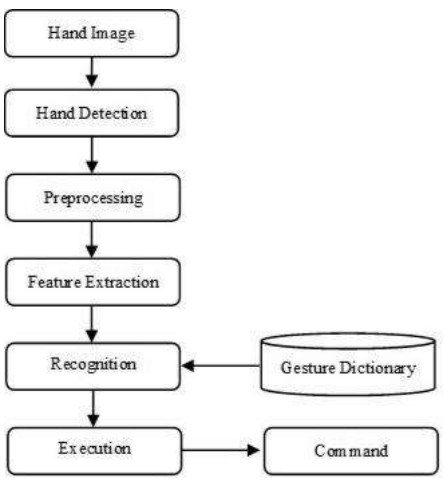

Fig. 1. Flowchart of hand gesture recognition.

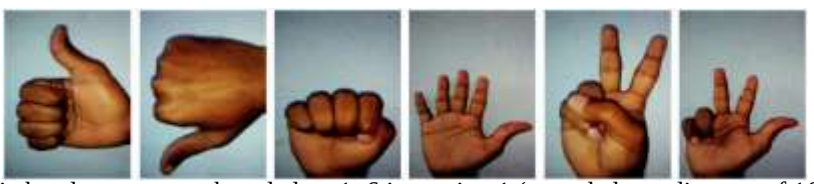

Fig. 2. Static hand gesture numbered class 1–6 in session 1 (recorded at a distance of 16 cm approx.)

Fig. 3. Static hand gesture numbered class 1–6 in session 2 (recorded at a distance of 21 cm. approx.)

Algorithm

1. Extract a frame (i.e. hand image) from recorded video stream.
2. Extracted frame is transformed from RGB color space to YCbCr[15] color space model and then hand is detected in the image using skin color based detection techniques[15].
3. After detection of hand we have converted the image into black & white i.e. marked the skin pixels as white and non-skin pixels (i.e. background) as black and then we have applied some preprocessing techniques like image filling, morphological erosion using 15 × 15 structuring elements etc. to increase the quality of image and to remove some noise.
4. For the feature extraction centroid, equivalent diameter, area, perimeter and orientation of detected objects is found out in the frame. With the help of centroid and diameter a circle is drawn same as background color pixels as shown in Fig. 5. The radius of the circle is calculated as shown in equation 1. All the features have been used until we have got the non-conflicting output.

$$Rf = (\text{EquivDiameter}/2) + \sigma \quad (1)$$

σ is some threshold value.

5. Gesture is recognized by counting the number of white objects in the image and orientation of image. Finally a command is passed to the applications running on the computer corresponding to the recognized gesture.

4. Experimental Setups

For the implementation of hand gesture recognition (HGR) system for multimedia interaction we have built our own

dataset. We have recorded each video stream of duration time approximately 10 seconds at the rate of 30 frames per seconds and at the resolution of 1280 × 720 using digital camera of 8 megapixel. We have performed the experiments in three different sessions. These sessions are classified based on the images (i.e. frames) extracted from the recorded video streams at different distance and positions. Each session consists of 300 images of 6 different classes where each class having 50 images. Some samples of images used in different sessions are shown in the Figs. 2, 3 and 4 based on which we will calculate the hand gesture recognition accuracies.

The whole system is designed using image processing and computer vision techniques implemented in Matlab-2013 under Windows 8 operating system. We have used a 64-bit computer with 2.40 GHz processor and 2 GB RAM as the hardware requirement for processing the hand gesture recognition system.

Fig. 4.Static hand gesture numbered class 1–6 in session 3 (recorded at a distance of 26 cm. approx.)

Fig. 5. Intermediate results of HGR system (generalized for all sessions).

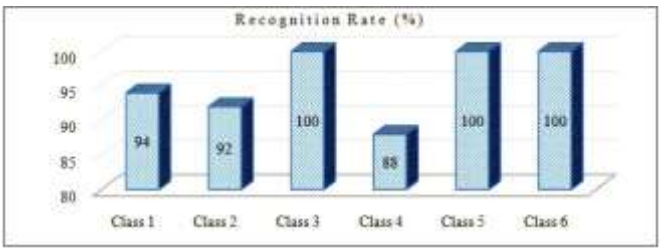

Fig. 6. Hand gesture accuracy in session 1.

This experiment has been performed in a good and stable lighting condition and under the simple background which does not contain any other skin like objects. The generalized intermediate results of the experiment for all the sessions for designing the hand gesture recognition system is shown in Fig. 5. It portrays the feature extraction of the hand by extracting the number of objects and orientation in the object of interest. Based on the recognized gesture a command can be passed to any multimedia applications running on computer for controlling its various operations.

5. Results and Discussion

After performing the practical experiments several times it has been seen that the system performance is very well in good lighting conditions and in a simple background environment which does not contain any skin like objects. The algorithm is not much robust, because in the complex background it is very difficult for it to detect the hand. The system performance also depends on the threshold value which we have taken to calculate the radius by doing the several practical approximations. Never the less the system is little bit fast responsive as compared to the other system which have been developed earlier as it does not require any training phase for gesture recognition.

We have executed the HGR system for each session and recorded the accuracy of the system which has shown in Figs. 6, 7 and 8.

After performing the experiments we can see that the overall accuracy of the system is 95.44% which is a very good result. The minimum accuracy we have achieved by class 3 gesture in session 3 due to the gesture shapes and positions. In session 1 the results were also not satisfactory for some of the classes due to the same reason. The threshold value (i.e. σ) which we have chosen plays an important role for recognizing the gestures. For increasing the performance of the system we can select a good threshold value after experimenting the system on some more number of images.

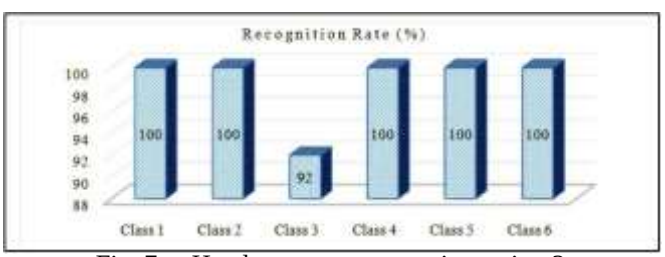

Fig. 7. Hand gesture accuracy in session 2.

Fig. 8. Hand gesture accuracy in session 3.

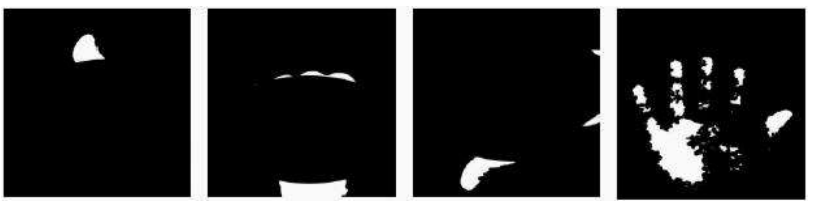

Fig. 9. Static hand gesture images not recognized by the system numbered 1–4.

We have compared our system with some other system and find that our system recognition rate is somewhat better than the others. The good thing about our system is that it does not require any training or classifications which is a very much time consuming tasks. There were certain cases where our system found troubled recognizing some of the gestures which is shown in Fig. 9.

From the above figure we can see that how these gestures had troubled to be recognized by our system. The first three gestures were detected by the system properly but not recognized by the system. The number 4 gesture is not detected by the system due to the bad lighting condition.

6. Conclusion and Future Works

In this modern world, where technologies is at the peak, there are many facilities available for offering input to any applications running on the computer systems, some of the inputs can be offered using physical touch and some of them without using physical touch (like speech, hand gestures, head gestures etc.). Using hand gestures many users can handle applications from distance without even touching it. But there are many applications which cannot be controlled using hand gestures as an input. This technique can be very helpful for physically challenged people because they can define the gesture according to their need.

The present system which we have implemented although seems to be user friendly as compared to modern device or command based system but it is less robust in detection and recognition as we have seen in the previous step. We need to improve our system and try to build more robust algorithm for both recognition and detection even in the cluttered background and a normal lighting condition. We also need to extend the system for some more class of gestures as we have implemented it for only 6 classes. However we can use this system to control applications like power point presentation, games, media player, windows picture manager etc.

References

[1] S. S. Rautaray and A. Agrawal, Vision Based Hand Gesture Recognition for Human Computer Interaction: A survey, Springer Transaction on Artificial Intelligence Review, pp. 1–54, (2012).

[2] P. Payeur, C. Pasca, A. Cretu and E. M. Petriu, Intelligent Haptic Sensor System for Robotic Manipulation, IEEE Transaction on Instrumentation and Measurement, vol. 54(4), pp. 1583–1592, (2005).

[3] S. Meena, A Study on Hand Gesture Recognition Technique, Master Thesis, Department of Electronics and Communication Engineering, National Institute of Technology, India, (2011).

[4] M. M. Hasan and P. K. Mishra, Hand Gesture Modeling and Recognition using Geometric Features: A Review, Canadian Journal on Image Processing and Computer Vision, vol. 3(1), pp. 12–26, (2012).

[5] B. A. Myers, A Brief History of Human Computer Interaction Technology, ACM Interactions, vol. 5(2), pp. 44–54, (1998).

[6] Malima, E. Ozgur¨ and M. C¸ etin, A Fast Algorithm for Vision-Based Hand Gesture Recognition For Robot Control, IEEE Signal Processing and Communications Applications, pp. 1–4, (2006).

[7] Z. Xu, et al., Hand Gesture Recognition and Virtual Game Control Based on 3D Accelerometer and EMG Sensors, In Proceedings of IUI'09, pp. 401–406, (2009).

[8] S. Mitra and T. Acharya, Gesture Recognition: A Survey, IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews, vol. 37(3), pp. 311–324, (2007).

[9] N. X. Tran, Wireless Data Glove for Gesture-Based Robotic Control, 13t h International Conference on HCI, vol. 5611, pp. 271–280, (2009).

[10] J. M. Rehg and T. Kanade, Visual Tracking of High DOF Articulated Structures: An Application to Human Hand Tracking, 3r d European Conference on Computer Vision, pp. 35–46, (1994).

[11] Murthy and Jadon, Hand Gesture Recognition using Neural Networks, In 2nd IEEE International Advance Computing Conference (IACC), pp. 134–138, (2010).

[12] S. K. Kang, M. Y. Nam and P. K. Rhee, Color Based Hand and Finger Detection Technology for user Interaction, IEEE International Conference on Convergence and Hybrid Information Technology, pp. 229–236, (2008).

[13] N. H. Dardas and N. D. Georganas, Real-Time Hand Gesture Detection and Recognition using Bag-of-Features and Support Vector Machine Techniques, IEEE Transaction on Instrumentations and Measurement, vol. 60(11), pp. 3592–3607, November (2011).

[14] S. S. Rautaray and A. Agrawal, A Novel Human Computer Interface Based on Hand Gesture Recognition using Computer Vision Techniques, In Proceedings of ACM IITM'10, pp. 292–296, (2010).

[15] N. A. Ibraheem, R. Z. Khan and M. M. Hasan, Comparative Study of Skin Color Based Segmentation Techniques, International Journal of Applied Information Systems (IJAIS), vol. 5(10), pp. 24–38, (2013).