International Journal of Electronics Signals and Systems

Volume 2 | Issue 2 Article 12

October 2012

GESTURE RECOGNITION SYSTEM

ANUBHAV SRIVASTAVA
Electronics and Communication Department, I.T.S. Engineering College, Greater Noida, Uttar Pradesh, India, swatiag2106@gmail.com

PRANSHI AGARWAL
Electronics and Communication Department, I.T.S. Engineering College, Greater Noida, Uttar Pradesh, India, its.pranshiagarwal@gmail.com

SWATI AGARWAL
Electronics and Communication Department, I.T.S. Engineering College, Greater Noida, Uttar Pradesh, India, anubhav.sani@gmail.com

USHA SHARMA
Electronics and Communication Department, I.T.S. Engineering College, Greater Noida, Uttar Pradesh, India, adorable_usha@rediffmail.com

Recommended Citation
SRIVASTAVA, ANUBHAV; AGARWAL, PRANSHI; AGARWAL, SWATI; and SHARMA, USHA (2012) "GESTURE RECOGNITION SYSTEM," *International Journal of Electronics Signals and Systems*: Vol. 2 : Iss. 2 , Article 12.
DOI: 10.47893/IJESS.2012.1080
Available at: https://www.interscience.in/ijess/vol2/iss2/12

GESTURE RECOGNITION SYSTEM

[1]ANUBHAV SRIVASTAVA, [2]PRANSHI AGARWAL, [3]SWATI AGARWAL & [4]USHA SHARMA

[1,2,3&4]Electronics and Communication Department, I.T.S. Engineering College, Greater Noida, Uttar Pradesh, India
Email: swatiag2106@gmail.com, , its.pranshiagarwal @gmail.com, anubhav.sani @gmail.com, adorable_usha@rediffmail.com

Abstract: In this paper, the hand gesture of a person is recognised and it identifies which hand of the person is raised. The skin colour is taken to recognise hands and face and the dark background is taken so that the skin detection may become easier. The hands and face are differentiated on the basis of area and centroid. Camera is the only input device used in this algorithm. No other input device is used to differentiate hands from the remaining body. This algorithm can be used both on the captured images and real time images.

Keyword: *Hand; area; centroid; skin detection; gesture recognition*

1. INTRODUCTION

Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures via mathematical algorithms. Gestures can originate from any bodily motion or state but commonly originate from the face or hand. These are a form of nonverbal communication in which visible bodily actions are used to communicate important messages, either in place of speech or together and in parallel with spoken words.

Gesture recognition can be seen as a way for computers to begin to understand human body language, this building a richer bridge between machines and humans than primitive text user interfaces or even GUIs (graphical user interfaces), which still limit the majority of input to keyboard and mouse.

In this paper, the hand gesture is recognised that identifies whether only a left or right hand is raised or both hands are raised. This can be used in the security purposes and in distance communication in which hand raising can be made a special sign and on the basis of which hand has been raised the particular action can take place.

During the implementation of the algorithm, it has been realised that the hand gesture recognition algorithm may produce errors in following conditions:

- Dynamic background: The dark background is best suited for this algorithm. In case, if the background will contain the different objects or the colours that are the same as that of skin colours, the errors may occur and it can be considered as noise.
- Lightning conditions: The variable lightning conditions may produce errors so it should have the adaptability to produce the correct results.
- Clothes colours and types: The colours of clothes should be different from that of the colour of skin. The different neck designs, sleeves designs are considered differently by the algorithm.
- Real-time interaction: As we move towards the real time images, the noise increases and hence the complexity. [2]

1.1 Some previous work related to hand gesture recognition:

Gesture recognition was first proposed by Myron W. Krueger as a new form of interaction between human and computer in the middle of seventies [3]. It has become a very important research area with the rapid development of computer hardware and vision systems in recent years. The important role of hand gesture for different applications has increasingly being acknowledged, although some approaches still focus only on static hand poses rather than dynamic use of more general types of gesture in context. Currently, there are several available techniques that are applicable for hand gesture recognition, which are either based on auxiliary devices or computer vision. A typical widespread device based example is VPL data glove, which is developed by Zimmerman in 1987 [7]. In this system, user wears a VPL data glove that is linked to the computer. The glove can measure the bending of fingers, the position and orientation of the hand in 3-D space. Data glove is able to capture the richness of a hand's gesture.

Hidden Markov models are used prominently and successfully in speech recognition and, more recently, in handwriting recognition. Consequently, they seem ideal for visual recognition of complex, structured hand gestures as are found in sign languages. Explicit segmentation on the word level is not necessary for either training or recognition. Language and context models can be applied on several different levels, and much related development of this technology has been done by the speech recognition community [9]. The perseus system developed by R. E. Kahn [4] applies a variety of techniques (e.g. motion, color, edge detection) to segment person's hands, it can be used to recognize the pointing gesture. This system requires a static background, and relies on off-board computation which causes delays in gesture recognition. The Hidden Markov Models (HMM) have intrinsic properties that make them very attractive for dynamic gesture recognition, T. Starner [5] developed a gesture recognition system for sign language recognition, HMM is used to learn repeatable patterns of human gestures with a recognition rate of 92%. Thad Starner and Alex Pentland presented two real-time hidden Markov model-based systems for recognizing sentence-level continuous American Sign Language (ASL) using a single camera to track the user's unadorned hands. In that system, the first system observes the user from a desk mounted camera and achieves 92 percent word accuracy. The second system mounts the camera in a cap worn by the user and achieves 98 percent accuracy (97 percent with an unrestricted grammar). Both experiments use a 40-word lexicon [6].

Multimodal and voice analysis can also help to infer intent via prosodic patterns, even when ignoring the content of speech. Robotic recognition of a small number of distinct prosodic patterns used by adults that communicate praise, prohibition, attention, and comfort to preverbal infants has been employed as feedback to the robot's `affective' state and behavioural expression, allowing for the emergence of interesting social interaction with humans (Breazeal and Aryananda, 2002). Hidden Markov Models (HMMs) have been used to classifying limited numbers of gestural patterns (such as letter shapes) and also to generate trajectories by a humanoid robot matching those demonstrated by a human (Billard et al., 2004). Multimodal speech and gesture recognition using HMMs has been implemented for giving commands via pointing, one-, and two-handed gestural commands together with voice for intention extraction into a structured symbolic data stream for use in controlling and programming a vacuuming cleaning robot (Iba et al., 2002) [1].

Recently, in the work of Tomoichi Takahashi and Fumio Kishino, they used hand shape input device Data Glove to recognise hand gesture [8]. So if a person will not wear that glove, their method may not work properly. Here, a method is presented takes a single camera as its input, no extra device is required to identify hands.

2. OVERVIEW OF HAND GESTUE RECOGNITION SCHEME

In this paper, the hand gesture is recognised and it explains which hand of the person is raised. It can be used as a sign language. The right hand raised condition can be used as the sign of moving right, left hand raised can be used as the sign of moving left, both hands raised means move straight and no hand raised means stop.

After taking the image, the skin filtering is done. Then, right and left hands are distinguished on the basis of the concept of area and centroid. The concept of segmentation was implemented first but that was very time taking algorithm. Then the concept of recognising the vertical axis of hand, i.e. the length from the tip of middle finger to end of palm was implemented but that increased the computational complexity. Hence, the concept of area and centroid is implemented which is very simple and give results equally good.

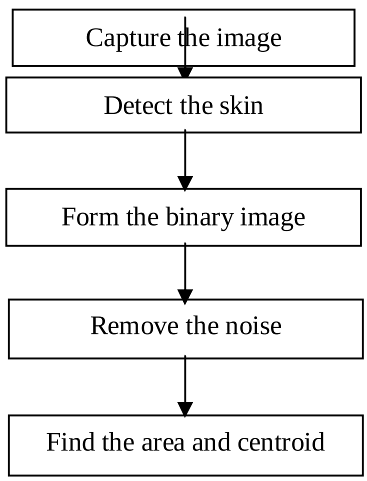
The platform used for the programming is MATLAB. Initially, we used R2007a version but it raised problem in executing the command "imtool" so we moved to R2007b but that raised problem in executing the commands of video capturing and taking snapshots from running video. Then finally we moved to R2010a version which supports both of these commands and it runs in 64-bit computers.

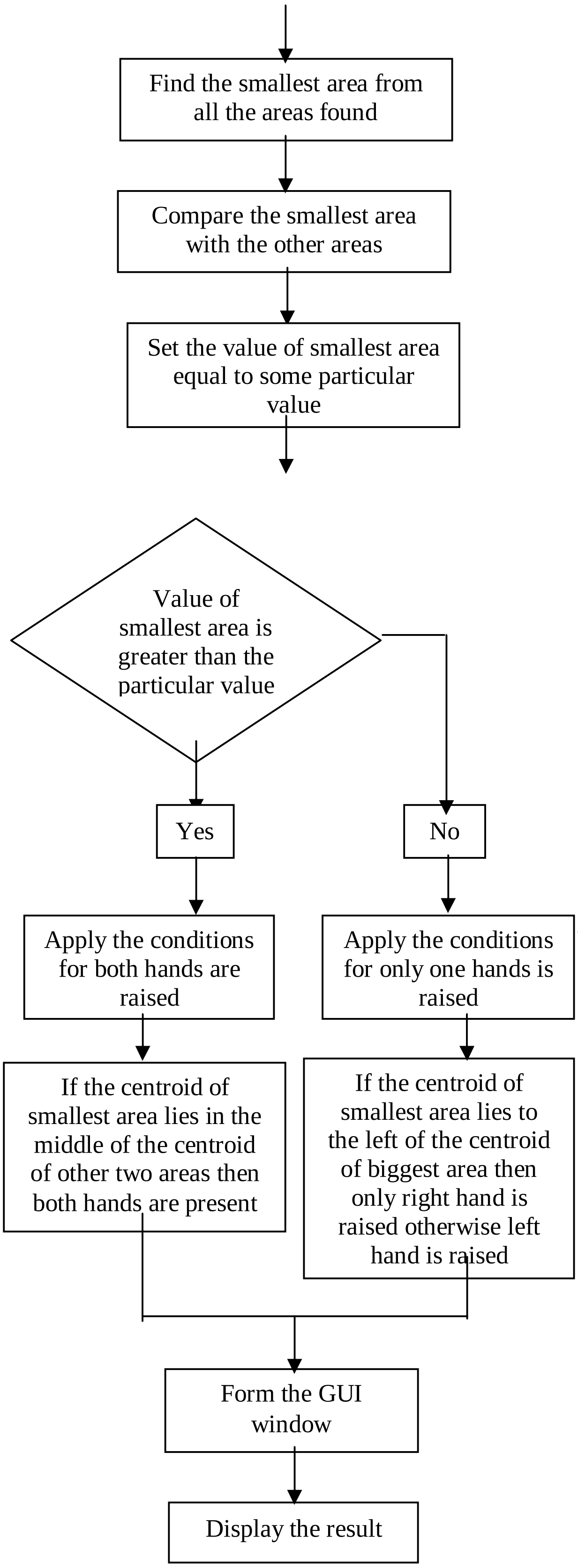
After recognising the hands as left or right, the result is finally displayed in GUI window.

3. ALGORITHM USED

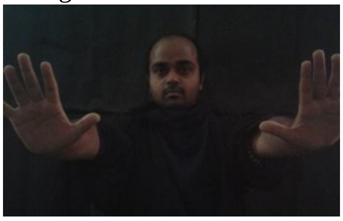
This algorithm is based on the concept of area and centroid, therefore we have named this algorithm as "AC ALGORITHM".

Following are the steps that are followed in order to recognise hands using area and centroid concept:





a. Capture the image: Firstly we take an image already present in the database or we can capture the image in MATLAB itself but it will create a lot of noises and exact results may not be produced. So preferably we take an image from the database. Fig. 1 (a), (b), (c) shows the original images taken from database.



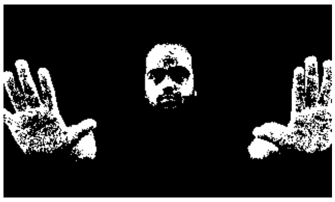
(a)

(b)

(c)

Fig.1 Original image from database

b. Skin detection: Then the skin is filtered from the image captured/taken. For this we use 'imtool' command that gives the pixel value of the skin portion of the image. Fig.2 (a), (b), (c) shows the images after skin detection.



(a)

(b)

(c)

Fig. 2 Image obtained after skin detection

c. Remove noise: The result found after the skin filtering contains a lot of noise, so it is removed using ‘strel’, ‘imclose’, ‘imfill’ and ‘imopen’ command. Fig.3 (a), (b), (c) shows the image obtained after removing noise.

(a)

(b)

(c)

Fig.3 Image obtained after removing noise

d. Binary image: Up to this step, the binary image is obtained that shows the skin area with white colour and rest of the portion by black colour.
e. Find area and centroid: Now the area and centroid of the white portion of the image is obtained. For this firstly label the image and take its properties in a certain variable. Now using the command stats.area and stats.centroid find the area and centroid of the image.
f. Remove small areas: As this image still contains some noise elements, so remove the smaller areas as the area of face and hands cannot be smaller than a particular value.
g. Final image to be processed: The image thus obtained is the final image to be processed, fig.4 (a), (b), (c) shows this.

(a)

(b)

(c)

Fig.4 Final image to be processed

h. Again find the area and centroid and store it: The area and centroid of this final image is found again. The areas and centroids are stored in different variables. The x-coordinate of the centroid is also stored separately.
i. Find the smallest area: Now the value of the smallest area is found by comparing all the areas’ value with one another.
j. Compare the smallest area with a particular value: Based on the results, set the particular value of the smallest area, so that if the value of smallest area value is more than that value than the condition can be of both hands raised otherwise only one hand is raised.
k. Apply the specific conditions: After comparing the value of the smallest area, compare the value of the three areas with one another one-by-one.
l. Compare the x-coordinate of centroid: Now compare the values of the x-coordinate of centroid.
m. Define the position of hands: If the x-coordinate of the smallest area is the smallest then only left hand is raised, if the x-coordinate of the smallest

area is largest then only right hand is raised and if it lies in the middle then both hands are raised.

n. Display the particular result: Display this result using the command 'disp' in command window.

o. Form the GUI window: There is one more impressive way of displaying the result, i.e. forming the GUI window. To form it open a blank GUI and write whatever you want to display in it and then save it. It will automatically generate an m-file.

p. Display results in GUI window: Write the name of the saved file in program wherever the result needs to be displayed; it will open the GUI window automatically. Fig.5 shows the final result displayed in GUI window

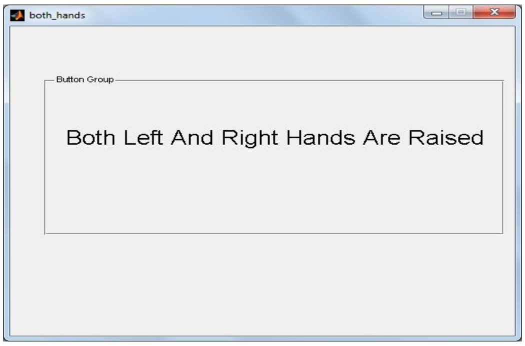


Fig.5 Final result displayed in GUI window

q. Particular message using sign language: A particular message can also be given using this gesture recognition system. For example, right hand raised can mean moving in right direction, left hand raised can mean moving in left direction, both hands raised may mean moving in both the directions and no hands raised may mean stop moving.

There are certain observations during the analysis of result (shown in Table I) which contains some variables that means as follows:

Ra1: denotes the area of the first white portion encountered by MATLAB during execution

Ra2: denotes the area of the second white portion encountered by MATLAB during execution

Ra3: denotes the area of the third white portion encountered by MATLAB during execution

Rc1: denotes the centroid of the first white portion encountered by MATLAB during execution

Rc2: denotes the centroid of the second white portion encountered by MATLAB during execution

Rc3: denotes the centroid of the third white portion encountered by MATLAB during execution

S: denotes smallest area among all the three areas

Bx: denotes the x-coordinate of the biggest area

Mx: denotes the x-coordinate of the second biggest area

Cx: denotes the x-coordinate of the smallest area

A.

Database		**(a)**	**(b)**	**(c)**
Properties	**Ra1**	154899	169783	156007
	Ra2	78038	118878	1078
	Ra3	151432	197445	7569
	Rc1	238.0394 918.1101	0.3195 1.4454	0.3301 1.819
	Rc2	973.8653 661.0845	0.9413 1.1440	0.0492 2.5533
	Rc3	1.7439 0.9112	1.6120 1.4222	0.2992 2.5506
	S	78038	118878	1078
	Bx	238.0394	1.6120e+003	330.1336
	Mx	1.7439e+003	319.4933	299.1633
	Cx	973.8653	941.3275	49.2022

Table I. Observartions for 'both hands raised'

Similarly, we obtained results for the other conditions, i.e. only left hand is raised and only right hand is raised.

4. CONCLUSION

The gesture of the person is identified using the area and centroid concept. It requires less computation time. The concept used is easy to understand and the implementation is also not complex. It gives more accurate results. It requires less mathematical calculations.

5. APPLICATIONS

As we are recognising the hand gesture of a person, thus this can be used in security system checks, for example, a particular gesture can be made a password and if the person makes that gesture than only the password is accepted. The gestures can also be used as a sign language. The motion of a person can also be recognised using this algorithm. If the person is moving in the right direction then by knowing only the movement of hands and face, the movement of the person can be recognised. Replacement of mouse, gaming and entertainment industry are other areas in which gesture application is already in use.

6. FUTURE SCOPE

Gesture recognition has become widely used in so many applications; it is hard to think about it going away. For further development, work could be done to remove more noise as it is the major constraint in the project. We also have taken the dark background, dark coloured clothes and proper lightning conditions to implement the code and to remove as much noise as possible, so work can be done to make it more versatile so that the code can be implemented in any background colour. The sleeves and neck designs may not become constraints for its implementation. Although the code is taking very less time for execution, still the execution time can be reduced by putting some more efforts in it. At present we are taking the image from the running video and then the gesture is being recognised on that video. The future work may include recognising the gesture of the person in the running video and the output should also be shown in the video itself rather than on any other GUI window.

REFERENCES

[1]. Chrystopher L. Nehaniv, Classifying Types of Gesture and Inferring Intent

[2]. Dong,Guo,Yonghua Yan and M. Xie, Vision-Based Hand Gesture Recognition for Human-Vehicle Interaction

[3]. Myron W.Krueger, Artificial Reality II, Addison- Wesley, Reading, 1991.

[4]. Roger E. Kahn, Peter N. Prokopowicz, and R. James Firby, "Gesture Recognition Using the Perseus Architecture," IEEE Conference on Computer Vision and Pattern Recognition, San Francisco California, 1996.

[5]. T. Starner, Visual Recognition of American Sign Language Using Hidden Markov Models, Master's Thesis, MIT, Feb 1995.

[6]. Thad starner and Alex Pentland, Real-Time American Sign Language Recognition Using Desk and Wearable Computer Based Video, IEEE transactions on pattern analysis and machine intelligence, vol.20, no.12, December 1998

[7]. Thomas G. Zimmerman and Jaron Lanier, "A Hand Gesture Interface Device, " ACM SIGCHI/GI, pages 189-192, 1987. Tomoichi Takahashi and Fumio Kishino , A hand gesture recognition method and its application, Systems and Computers in Japan Volume 23, Issue 3, 1992, Pages: 38–48, Article first published online : 21 Mar 2007, DOI: 10.1002/scj.4690230304

[8]. X.D. Huang, Y. Ariki, and M.A. Jack, Hidden Markov Models for Speech Recognition, Edinburgh Univ. Press, 1990.

◈◈◈