

Vision Based Hand Gesture Recognition Using Skin Detection and Morphological Operations

Shreyasi Sarkar¹, Ashwini Gade²Student, Department of Electronics and Telecommunication, MPSTME, NMIMS, Mumbai, India¹Assistant Professor, Department of Electronics and Telecommunication, MPSTME, NMIMS, Mumbai, India²

ABSTRACT: Gesture recognition technology deals with the interpretation of gestures using various mathematical algorithms to develop a user-friendly human computer interaction (HCI) interface. Recognition of gestures aims at creation of a system that can effectively translate specific human gestures to convey meaningful information for applications such as home automation, sign language interpretation, and device control etc. In this paper, we propose a fast algorithm for recognition of a limited set of gestures from hand images of different skin tones. Our approach contains steps for applying pre-processing techniques, segmenting the hand region and finally classifying the gesture using MATLAB. The system detects skin regions based on a set of skin detection criteria which helps to filter out the non-skin regions and extract the hand shape using morphological operations.

KEYWORDS: Gesture, Gesture recognition, Skin detection, Morphological operations, Human computer interface, MATLAB.

I. INTRODUCTION

In today's world, hand gesture recognition that provides a natural form of human machine interaction is one of the growing fields of research. Gestures are actions, which a human being uses to express meaningful information across other individuals without having to say anything. Gesture recognition technique is a tool to manipulate and exploit this meaningful information conveyed through gestures, using mathematical algorithms. For efficient human computer interaction, hand gestures are a spontaneous and a powerful mode of communication.

Traditionally, keyboard, mouse, touch screen and other attachments are used as input devices to communicate with computer. However, these provide an artificial interface and always require a physical contact to be maintained. [1] Interaction between humans comes from different sensory modes like gesture, speech, facial and body expressions [2]. The advantages of using hand gestures are that hands are the most flexible part of the body and can show different information to interact with computer from a distance as a non-contact human computer input modality. This makes the man-machine interface more intuitive in nature. Interfaces based on successful recognition of gestures provides a natural form of communication between man and machine giving rise to a number of applications for example remote free systems for centralizing the control of electronic appliances, interpretation of sign language and many more. [3] Gesture recognition systems can be one of the three states, namely, glove based, vision based and drawing based. [4] The glove-based system depends on sensors and wires and is considered costly. A number of issues are faced while using glove based and drawing based input devices for collection of raw postures of hand and gesture recognition data, which requires to be connected to the computer and the user to wear gloves or draw circles on the fingertips. [5] Usage of these devices also restricts the movement of the hand and cause inconvenience to the users in a similar manner to the traditional interaction methods. The most promising and a good alternative for glove-based system is the vision-based system. [6]

By using concepts like image segmentation and image enhancement, a vision based interaction system for gesture recognition and processing can be established.[7] Vision based methods of hand gesture recognition needs only a camera and the gesture recognition algorithm responsible for translating the human gestures into a command to be carried out by the machine [8]. Thus, without the use of any sensors and wires it realizes a natural human computer

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

interaction. To achieve real time performance, this system needs to be background invariant, skin tone invariant, light insensitive, person and camera position independent.

Gesture recognition solutions can be divided into two types on the basis of the type of gesture used for establishing a human computer interface. [4][9] Static gestures can be considered as a change of the hand shape, which means systems, based on such gestures a single pose for single command. Dynamic gestures can be considered as change of the hand position in a particular time interval with a given velocity and the systems utilizing these gestures require a sequence of gestures for a single command.

The paper is organized as follows. Section II contains a discussion of the work previously done in the field of gesture recognition. Section III presents the proposed vision based hand gesture recognition system. The block diagram and the flowchart of the proposed system and the description of the algorithm are provided in Section IV. Section V contains the experimental results obtained on the tested gesture at each step of the system. Section VI presents conclusion.

II. RELATED WORK

In the past years, a lot of work has been done in the field of vision based gesture recognition solutions aiming to develop fast intelligent human interaction systems to successfully translate gestures into speech or text.

A gesture recognition system generally consists of 4 main segments, namely collection of input, image preprocessing, feature extraction and recognition algorithm. [1] Preprocessing operations are performed at the lowest level of abstraction to make the image suitable for further processing. Applications to recognize static hand gesture have used HSV color space [1][4][9], black and white color space [10], RGB color space [11] and YCbCr color space [12] for preprocessing of the input image. Background suppression procedure in the HSV color space was used to model the scene discarding illumination variations to focus on areas containing human skin color.[1] With predefined values of H, S, and V to apply segment filling strategy, the segment was filled before applying the edge detection method. This helped to vanish all the black spots caused by the segmentation operation and led to better edges. [9] Conversion of the input images from RGB into binary image is an efficient yet simple preprocessing operation because in binary images the skin region is shown with white pixels whereas the background is represented by black pixels. [10] Pixel count algorithm that simply counts the skin pixels to give the count of fingers raised was found to be simple but a size variant algorithm. [10] Intensity-based position marking method was used in order to detect the red color marker placed at fingertip of the white glove worn by the user to track the user's hand. [11] The hand gesture was singled out from the image using skin-based thresholding and features were extracted from the thresholded image based on the distance between the centroid and key contour points on the hand. [12] [13] The contour of the hand was obtained through serial morphological operations and then the fingertips were marked out through the computation of Fingertip-Angle. [13] To overcome the problems including skin-like color interference, complex environmental background and overlapping, depth map segmentation with K-means algorithm was used to get the contour of human body and determine the center location of face or hands based on the detected skin blobs. [14] Apart from these, various techniques for data collection, algorithms for recognizing hand postures and gestures and methods for hand detection and segmentation along with recognition models have been analyzed can be found in [5].

Colour is a powerful descriptor for skin that can be exploited to segregate skin and non-skin regions. An efficient and effective algorithm can be developed to recognize gestures of hand by exploiting the current advancements in the field of image processing.

III. PROPOSED SYSTEM

This paper presents a system for recognition of static hand gestures by using skin detection in conjunction with morphological operations. To reduce the complexity and make the system more natural and user-friendly, the user doesn't need to wear hand gloves or any other sensor device. These images are subjected to MATLAB based algorithms, which preprocesses the image, detects the skin region, extracts the gesture and finds an approximate match with the database. The database consists of a set of gestures which are used to train the computer. The database images as well as the input images undergo the same algorithms for extracting the gesture components. The aim of this project is to implement a user independent, convenient and robust vision based gesture recognition system for static gestures.

IV. METHODOLOGY

The vision based gesture recognition system is divided into two subsystems: creation of training database and recognition of gestures. Both these subsystems are interdependent. Fig. 1. gives the block diagram of the proposed system. The first subsystem carries out the task to the train the system with a set of gestures that are to be recognized. The other subsystem deals with the recognition of the gestures.

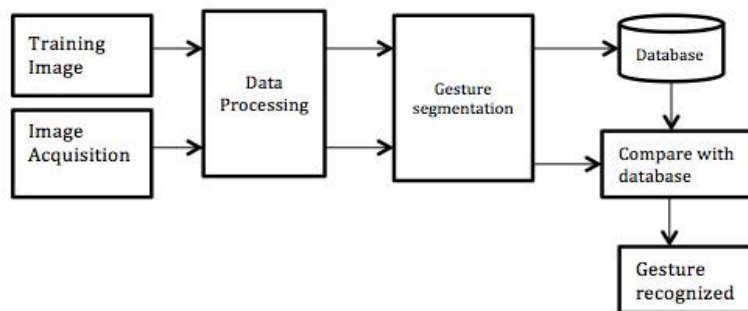


Fig. 1. Block diagram of vision based hand gesture recognition system

The system designed to recognize hand gestures consists of 5 steps. Fig. 2 shows the flow of the steps involved to recognize the set of hand gestures. A live frame is grabbed spontaneously as the input image through the webcam or the externally connected camera. After image acquisition, conversion to YCbCr is done as a preprocessing step. Then skin detection is carried out and gesture segmentation is done by performing filling of holes and opening operation. After performing all these steps a binary image is obtained in which the white pixels represent the gesture and black pixels represent the background.

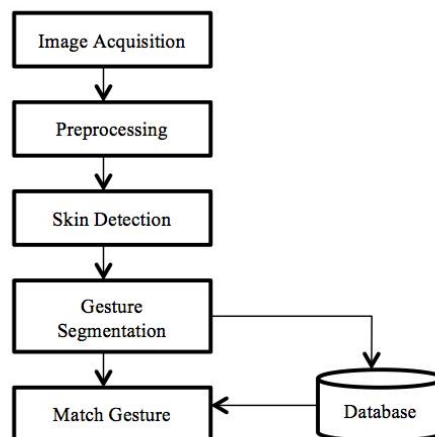


Fig. 2. Flowchart

A. Data Collection and Training

To train the vision-based solution with a set of specific gestures and subsequently recognize the hand gestures with high accuracy, four components play a major role. First is the placement of camera. It is critically important to achieve maximum visibility of the hand for fast and robust recognition. The second important component is the placement of the hand. It will provide ease of visibility to the camera and facilitate simpler extraction of the gesture components. The third component is the extraction of features from a stream of raw data on which the recognition

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

algorithms are to be applied. The fourth component is the recognition algorithm applied on the extracted features. [5] The first step is the collection of images for creating a database, which will comprise of the images of the various hand postures. All images were real time photographed through a digital camera of 8MP in a background that was devoid of any objects. The images are originally present in RGB format and are resized to 400x400 to reduce the size and increase the computational speed of the proposed system.

B. Skin Detection

In the proposed approach, after image acquisition skin filtering is carried out. Here, the RGB image is converted into YCbCr because this model helps to accommodate most of the skin tone variations in our experience. Y is the luminance, and Cb and Cr are the chrominance of the pixel. The main goal is to segment out the hand in the image from the background which is achieved by finding the pixels, which are more likely to belong to the skin region in the scene. Two sets of upper and lower thresholds are used, each one individually for Cb and Cr, to successfully extract the hand-like intensities in the image of the gesture. The two thresholds are selected in such a way that it effectively helps in identifying the location of the hand. These thresholds define the range for Cb and Cr to mark the skin pixels within the image of interest. This range is found empirically and it varies with the skin complexion of ethnic groups and lighting condition. Therefore, it is required to select a wide range for Cb and Cr, which covers almost all skin tones to make the system user independent.

C. Morphological Operations

Morphological processing is carried out to get the superior hand shape. They apply a structuring element, a matrix consisting of 0's and 1's with any arbitrary shape, to an input image for creating an output image of the same size. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. [15] To find the count of fingers the image is converted from YCbCr to binary image using Otsu's algorithm after skin detection is performed. Otsu's algorithm performs clustering based thresholding to reduce the image to binary. Then morphological operations of filling the holes in the image and opening is performed to eliminate the effect of noise which was added during the process of image capturing or during the process of conversion of the image from RGB format to YCbCr and from YCbCr to binary. Opening of an image by a structuring element is erosion followed by dilation of that image. This results in the final image which shows the given gesture in white on a black background.

D. Gesture Recognition

The process of skin detection and morphological operation is repeated for each and every image of the collection of gestures that is used for training the system. After going through the algorithm, the results of each image are stored serially in the rows of a two dimensional matrix. This matrix is referred to as the database that trains the vision-based solution and acts as a reference with which the input test image is compared to match the gestures. A video input interface is created to grab live frames spontaneously through MATLAB using a webcam or the camera which is connected to the computer. This image is subjected to the same algorithms which were carried on the database images to filter and mark the skin region out of the scene, and extract the count of fingers from the superior hand shape obtained after performing morphological operations. The resultant pixels are stored in a matrix and this matrix is subtracted from the database matrix. The gesture, which gives the minimum difference with input image matrix, is considered to be the closest match of the gesture.

V. EXPERIMENTAL RESULTS

The following figures show the results of the proposed vision based gesture recognition system for an input image using skin detection in conjunction with morphological operations. In Figure 3 and 4, (a) shows the original image, (b) is the image obtained by converting the original image form RGB format to YCbCr, (c) is the resultant image after applying detection criteria to mark the skin pixels on the original image, (d) is the image after converting into binary format and (e) is the resultant image after applying morphological operations of filling the holes and opening on the binary image.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

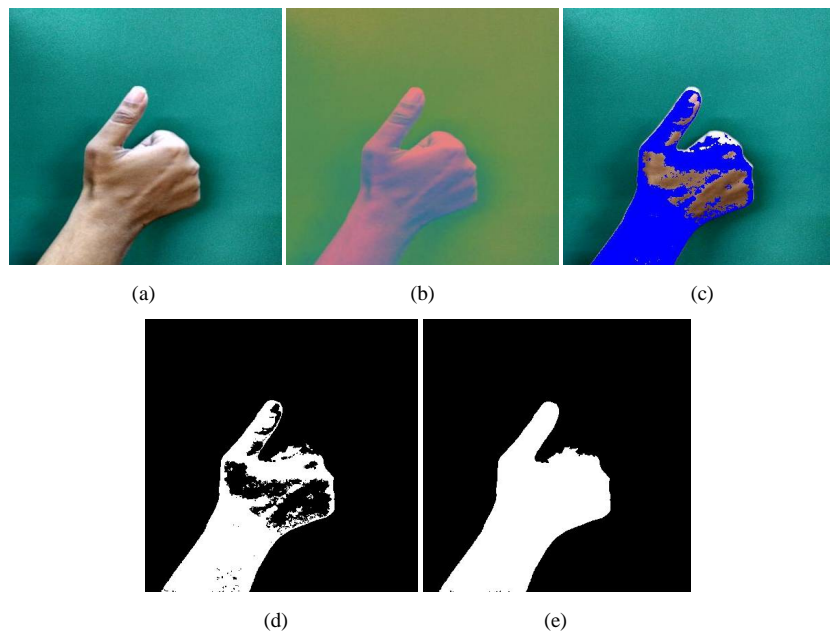


Fig. 3. Gesture Recognition (a) Original image (b) Image after converting into YCbCr (c) Image after applying skin detection criteria and marking the skin pixels on the original image (d) Image after converting into binary form (e) Image after filling the holes and applying opening operation.

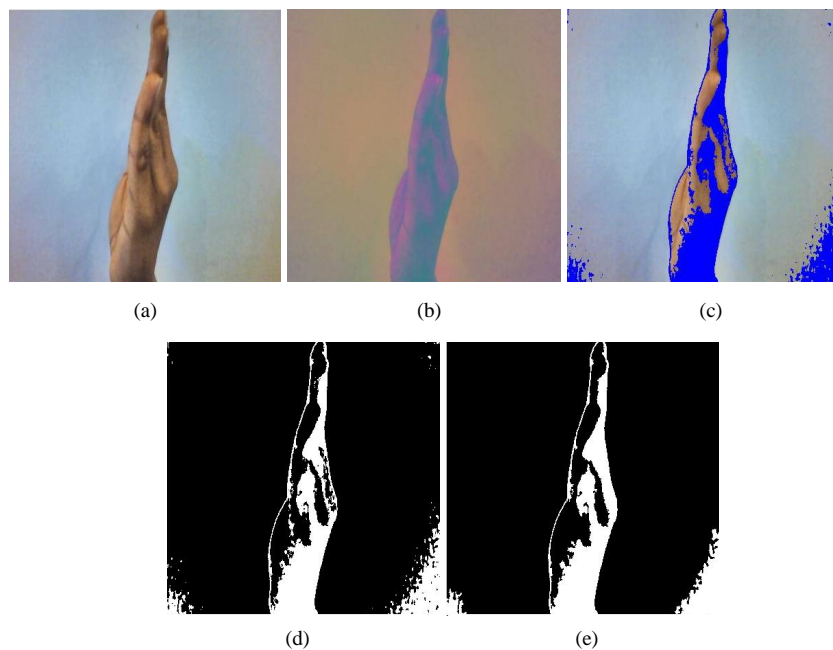


Fig. 4. Gesture Recognition (a) Original image (b) Image after converting into YCbCr (c) Image after applying skin detection criteria and marking the skin pixels on the original image (d) Image after converting into binary form (e) Image after filling the holes and applying opening operation.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

Images with green background showed better results when compared to the images with other colour backgrounds. This can be distinctly observed in fig.3 (e). Variations in light intensities added noise to the image which can be seen prominently in fig.4 (e). This did not affect the recognition performance of the system to a high degree because the closest match is considered to be the gesture that has minimum difference with the reference data. Therefore, images with green background were used to train the system and create the reference database.

Table 1 summarizes the results of the number of gestures recognised by the vision based static hand gesture recognition system which has been trained once against a set of 7 gestures. 10 images of each gesture are gathered from 5 users of different skin complexions. A sample set of 50(5*10) images for each gesture is used for testing the system. The system yields a maximum accuracy of 86 % per gesture and an average accuracy of 84 % per gesture. Also, it yields a maximum accuracy of 87.1 % per user and an average accuracy of 84 % per user.

| Gesture | User 1 | User 2 | User 3 | User 4 | User 5 | Accuracy per gesture |
|----------------------|--------|--------|--------|--------|--------|----------------------|
| One | 8 | 7 | 8 | 9 | 8 | 80 |
| Two | 8 | 8 | 8 | 9 | 9 | 84 |
| Three | 9 | 8 | 9 | 8 | 8 | 84 |
| Four | 8 | 8 | 8 | 9 | 9 | 84 |
| Five | 9 | 9 | 8 | 9 | 8 | 86 |
| Thumbs up | 9 | 7 | 9 | 8 | 9 | 84 |
| Hand raised sideways | 9 | 8 | 8 | 9 | 9 | 86 |
| Accuracy per user | 85.7% | 78.6% | 82.9% | 87.1% | 85.7% | -- |

Table 1. User specific results for 10 input images per gesture by each user

VI. CONCLUSION

Thus, we have proposed a fast, feasible, skin tone invariant and cost effective system using an algorithm based on skin detection and morphological operations to recognize static gestures. The need to wear additional hardware such as gloves, sensors and wires is eliminated to make it more user-friendly. The user must train the computer with different gestures in a clear background to translate the other real time gestures that are going to be fed to the computer as an input. An interface to acquire a live frame from the video input of camera is set up. Given the image of hand, the algorithm detects the hand region and extracts it. The gesture is compared with the existing database and returns the database gesture that gives the minimum difference with the input image as the closest successful match. The output of this robust gesture recognition system can be suitably used in applications such as interpretation of sign language, centralized control of home appliances for smart living, and device control by physically disabled and/or blind people. Under future work, improvements such as making the system background invariant, usage of dynamic gestures as input and rejection of unknown gestures may be implemented to increase the capabilities and functionality of the system.

REFERENCES

- [1] Rautaray S. and Agrawal A., " Real Time Hand Gesture Recognition System For Dynamic Applications" , International Journal of UbiComp (IJU), Vol.3 , No.1, pp. 21-31, 2012
- [2] Ismail, N. A. and O'Brien E. A., " Enabling Multimodal Interaction in Web-Based Personal Digital Photo Browsing", Proc. of the International Conference on Computer and Communication Engineering, IEEE , pp. 907-910, 2008

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

- [3] Ducloux J., Petrashin P., Lancioni W. and Toledo L. ,” Remote Control with Accelerometer-Based Hand Gesture Recognition for Interaction in Digital TV”, Proc. of the Argentine Conference on Micro-Nanoelectronics, Technology and Applications (EAMTA), IEEE , pp. 29-34, 2014
- [4] Hasan M. and Mishra P., “HSV Brightness Factor Matching for Gesture Recognition System”, International Journal of Image Processing, Vol. 4, Issue 5, pp. 456-467, 2010
- [5] Pradipa R. and Kavitha S., ”Hand Gesture Recognition - Analysis of Various Techniques, Methods and Their Algorithms” , International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, Special Issue 3, pp. 2003-2010, 2014
- [6] Erol A., Bebis G., Nicolescu M., Boyle R. D. and Twombly X., “Vision-Based Hand Pose Estimation: A Review”, Computer Vision and Image Understanding, Vol. 108, Issues 1–2, pp. 52–73, 2007
- [7] Panwar, M., “Hand Gesture Recognition based on Shape Parameters”, Proc. of the International Conference on Computing, Communication and Applications (ICCCA), IEEE, pp. 1-6, 2012
- [8] Hasan M. and Mishra P., “Gesture Recognition using Modified HSV Segmentation “, Proc. of the International Conference on Communication Systems and Network Technologies (CSNT), IEEE , pp. 328 – 332, 2011
- [9] Solanki U. and Desai N. “Hand gesture based remote control for home appliances : Handmote “ , Proc. of World Congress on Information and Communication Technologies (WICT), IEEE , pp. 419- 423, 2011
- [10] Chowdary P. R. V., Babu M. N. , Subbareddy T. V., Reddy B. M. and Elamaran V., “Image Processing Algorithms for Gesture Recognition using MATLAB “ , Proc. of the International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), IEEE, pp.1511-1514, 2014
- [11] Nancy and Sekhon G.S. “An Analysis Of Hand Gesture Recognition Technique Using Finger Movement Detection Based On Color Marker”, International Journal of Computer Science & Communication (IJCSC), Vol. 3, No. 1, pp. 129-133 , 2012
- [12] Tariq M., Iqbal A., Zahid A., Iqbal Z. and Akhtar J.,” Sign Language Localization: Learning to Eliminate Language Dialects”, Proc. of the 15th International Multitopic Conference (INMIC), IEEE, pp.17-22, 2012
- [13] Wen X. and Niu Y., “A Method for Hand Gesture Recognition Based on Morphology and Fingertip Angle ”, Proc. of the 2nd International Conference on Computer and Automation Engineering (ICCAE), IEEE, Vol.1, pp. 688 – 691, 2010
- [14] Xu D., Chen Y.L., Wu X., Ou Y. and Xu Y., “Integrated Approach of Skin-color Detection and Depth Information for Hand and Face Localization”, Proc. of the International Conference on Robotics and Biomimetics , IEEE, pp. 952-956, 2011
- [15] Konwar A. S. , Borah B. S. and Tuithung C. T. ,”An American Sign Language Detection System using HSV Color Model and Edge Detection” Proc. of the International Conference on Communication and Signal Processing (ICCSP), IEEE, pp. 743-747, 2014