Hand Gesture Recogniton

Neha Choudhary

Sri Sai College of engineering and Technology Pathankot ,Pb, India

ABSTRACT

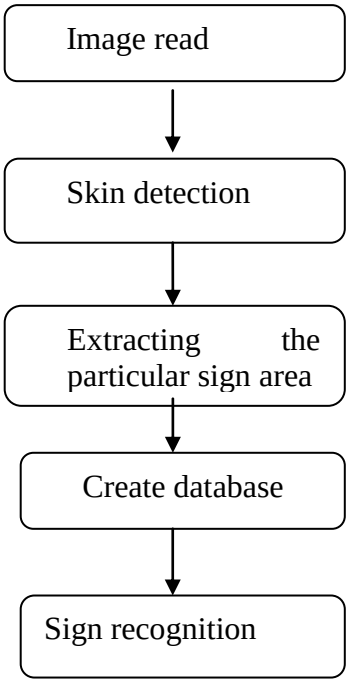
Hand gestures recognition provides a easy way to Interact and communicate with machines of different types. A hand gesture is classified into two categories: static and dynamic. A static gesture is a particular hand configuration or pose which is represented by a single image. A dynamic gesture is a moving pose, represented by a sequence of images. In this paper we will focus on the recognition of static images. The first step is skin detection. Skin detection basically helps to identify the presence of a human face, hand or any other body part by identifying and marking skin-like pixels in a given image .It is based on gray world algorithm. The second step is to extract the particular sign area or the area to be recognized. Converting the input image into the set of features is called feature extraction. The aim of this phase is to find and extract features that can be further used to determine the meaning of a given gesture. Hand gestures recognition system has been applied for different applications on different areas like sign language translation and recognition, virtual environments, smart surveillance, robot control, medical systems etc.

Keywords: Sign recognition, hand gesture, Gray world Algorithm.

I. INTRODUCTION

Hand gestures recognition provides a natural way to interact and communicate with computers. As Compared to the currently used human-machine-interfaces (HMI) such as a keyboard or a remote control, static hand gesture recognition does without any additional devices which are used to give instructions to a machine. In a system, which is generally known as static hand gesture recognition, a person instructs the machine using his bare hands, whereas images of the persons hand gestures are taken and analysed in order to determine the exact meaning of the hand gesture.[1] Basically hand gestures can be classified in two categories: static and dynamic. A static gesture is a particular hand pose which is represented by a single image. A dynamic gesture is a moving gesture which is represented by a sequence of images. In this paper we will focus on the recognition of static images. The main problem to deal with is how, by use of image processing and given set of pictures, to study the pictures and to find out the user's hand in it. Suppose the user will not always put his hand in the same position, the solution could not be based on certain targeted area.. Although a static hand gesture is only theoretically possible, normally only a limited set of well defined postures are considered to be used in this system. Generally gesture recognition is considered as a very challenging field because some natural environments is unsuitable for gesture recognition due to bad illumination, non uniform backgrounds, and so on. The recent publications of the few years show that static hand gesture recognition is still a field of active research area.[2] Sign languages are the most simple and natural form of languages.

II. BLOCK DIAGRAM OF SYSTEM



Working of block diagram of this proposed system is very easy and efficient. First of all we have to read image. Images are read into the Matlab environment using the function imread. [3] Skin detection is the process of finding skin like pixels and regions in an image or a video. This process is basically used as a pre-processing step to find raw regions that contain human faces and limbs in images. Different computer vision approaches have been developed for skin detection. [4] A skin detector typically converts a given pixel into an appropriate color space and then use a skin classifier to mark the pixel whether it is a skin or a non-skin pixel. Good segmentation of the image process leads to perfect features extraction and latter play an important role in a complete recognition process. Features vector of the segmented image can be used in different ways according to particular application. Various methods have been applied for representing how features can be extracted. Some of the methods use the shape of the hand such as hand contour while others use fingertips position and palm center[5]. Creation of the database is used to store the images and other information regarding images. Database is created in the form of rows and columns.

III. SKIN DETECTION AND GRAY WORLD ALGORITHM

a) Introduction

Skin detection is basically the process of finding skin-colored pixels and regions in an image or a video. This process is typically used as a pre-processing step in skin detection. Several approaches have been developed for skin detection. A skin detector typically converts a given pixel into a particular color space and then use a skin classifier to mark the pixel whether it is a skin or a non-skin pixel [9] A skin classifier defines a particular boundary of the skin color class in the color space based on the stored database of skin-colored pixels .Skin color and textures are important cues. Skin color and texture represent race, health, age, wealth, beauty of the person. However, such conditions vary across cultures and across the history. [6]In images and videos, skin color is a representation of the existence of humans in such media. Therefore, in the past few years extensive research have focused on skin detection in images. Skin detection basically means detecting image pixels and regions that contain skin-tone color. Most the research in this area just focus on detecting skin pixels and regions based on their color. Very few also use texture information to classify and mark skin pixel. Detecting skin pixels is basically easy task and can be done very efficiently in many different ways. For example, in earlier applications, detecting skin-colored regions was used to identify nude pictures on the internet for the sake of content filtering and security.

b.) Skin detection and color spaces

The human skin color has a restricted range of hues and saturation. The skin pixels are formed by a combination of blood (red) and melanin (brown, yellow). Therefore, the human skin color does not fall in a particular color space, but clustered at some area in the color space. But it is not always the same for all the color spaces. Varieties of color spaces have been used in skin detection to finding a color space where the skin color is invariant to illumination conditions. Color spaces affect the shape of the skin class, which affects the detection process. [7]

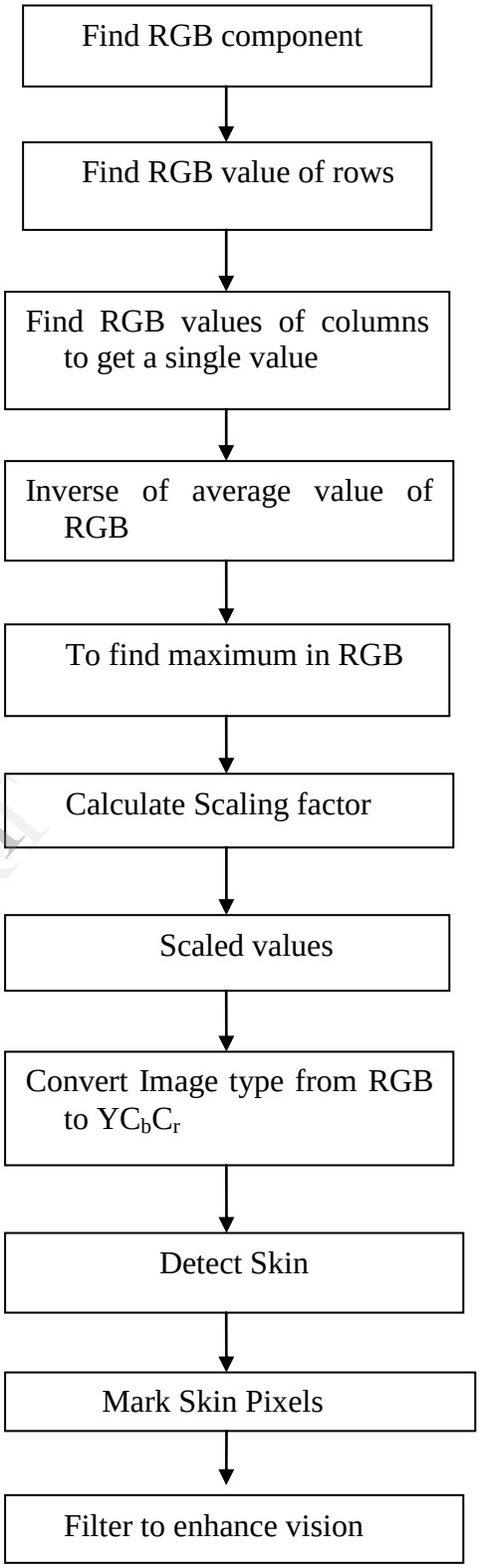
c) Gray World Algorithm

Gray World algorithm is based on the theory that given an image with sufficient amount of color variations, the average value of the R, G, and B components of the image should average to a common gray scale. This assumption is proved very well in a real world image and basically it is true that there are a lot of different color variations.[8] The variations in the color pattern are random and independent, the average should converge to the mean value as well as gray value, by given an enough amount of samples. Color balancing algorithms can be applied on this assumption by forcing its images to have a common gray value for its R, G, and B components. In the case when an image is taken by a digital camera under a particular lighting environment, the effect of the special lighting conditions can be removed by using the gray world assumption on the image. Result based on approximation is the color of the image is much closer to the original scene.

The main role of this color balance algorithm is to try to correct for the various factors that might have affected the color intensity values of the image pixels. The most

simple color balance algorithm uses the gray-world assumption. The gray-world algorithm states that given an image with proper amount of color variations, the average value of the red, green and blue components of the image should average to a common gray value or scale. This fact in general is valid for many cases. In any given real world scene we find lots of different color variations. Basically the variations in color are random and independent, it would be true to say that given a large amount of samples, and the average should always tend to converge to the mean value in the color space. [10].

A simple method of gray-world assumption is to find the average values for each of the R, G, and B color components in the image and use these average values to calculate an overall gray value for the image. Each color component is then scaled according to the total amount of its deviation from this gray value. We can also obtain the scale factors for each color component in the image by simply dividing the average gray-value for that component by the average gray-value across all color components. This scale factor is then to be divided by each color component individually. Therefore if an image under the normal white lighting satisfies the gray-world assumption, putting it under color filtered lighting would surely disturb this behavior. By applying the gray-world assumption on the image again we will find that it will remove the colored lighting to reacquire the true colors of the original image. In order to perform skin detection first of all we need to find a possible range of values for skin pixels. Basically when we load an image the pixel values of the image correspond to their RGB coordinates. As discussed earlier it is not the ideal space when we are working with the color. [11] First of all we need to convert the images to the YC_bC_r space. YC_bC_r is basically used in JPEG image compression and MPEG video compression.



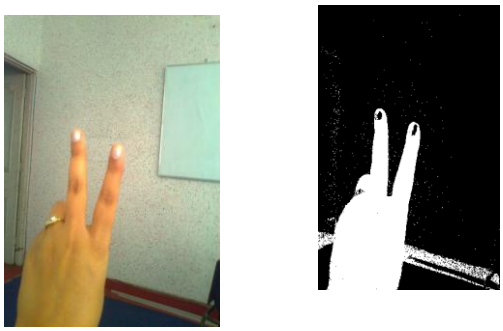


Figure 1.(a) Original image (b) image after skin detection without use of filters

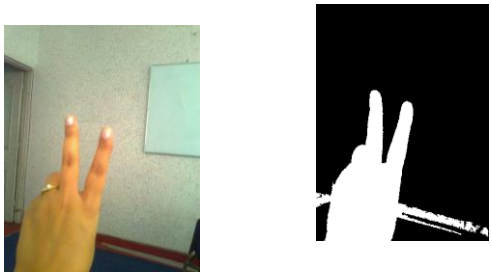


Figure 2. (a)Original image (b) image after skin detection with use of filter

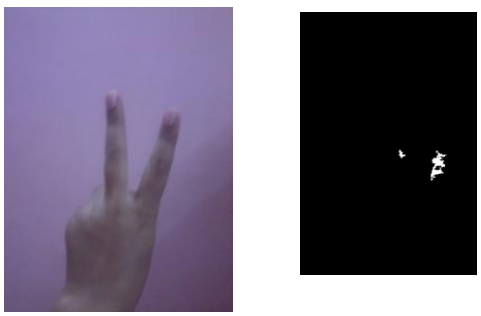


Figure 3. (a) Image in bad condition (b) image after skin detection.

IV. EXTRACTING THE PARTICULAR SIGN AREA

Good segmentation process leads to efficient features extraction process and latter it plays an important role in a successful gesture recognition process. Feature vector of the segmented image can be extracted in many ways according to its application. Various methods have been used for representing how the features can be extracted. Some methods used the shape of the hand such as hand contour and others used the fingertips position, palm centre, etc. By creating 13 parameters as a feature vector the first parameters represents the ratio of the boundary box of the hand and the rest of the 12 parameters are mean values of brightness pixels in the image. By using the Self-Growing and Self-Organized Neural Gas (SGONG) neural algorithm we capture the shape of the hand to obtain palm region, palm centre, and hand slope. Hand region divides the segmented image into various blocks size and each of its block represents the brightness measurements in the image. Many experiments were performed to decide the perfect block size that can achieve good recognition rate.

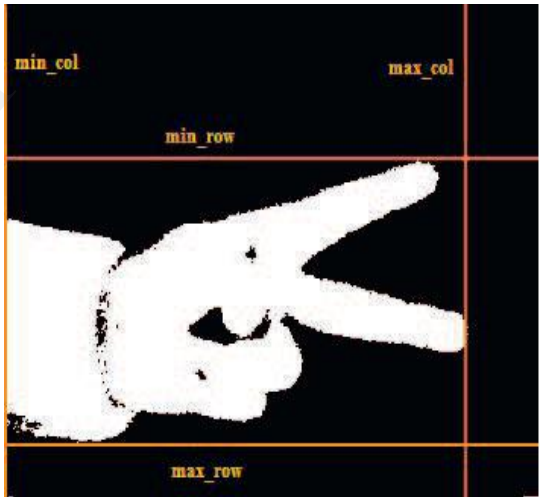


Figure 3. Local maxima and minima.

The first step is to convert selected frame into black-and-white image using global thresholding concept. Second step is to examine and extract the object of interest from the frame. In our case the object of interest is the part of human hand which is showing a gesture. [12] For the extra part other than the hand is cropped out so that the pattern matching can give more accurate and efficient results. For cropping extra parts row and column number is determined first from where the actual object of interest appears. This is done by calculating from each side of binary image and moving forward until the white pixels appears more than the offset value. Experimental results shows that the offset value set to 1% of

the total width gives better result for noise adjustment.

V. ADVANTAGES OF HAND GESTURE RECOGNITION

a. High speed and sufficient reliable for recognition process.
b. Efficient performance system with complex background.
c. Proper shape of the hand obtained led to a good feature extraction and recognition process.
d. Efficient and powerful results from the proposed algorithm.

VI. DISADAVANTAGES OF HAND GESTURE RECOGNITION

a. Rotation problem arises when the hand region is rotated in any direction in the scene.
b. Background problem is related to the complex background where there are many other objects in the scene with the hand objects and these objects might contain skin like color which would produce classification problem.
c. Scale problem appears when the hand gesture has different sizes in the image.
d. Translation problem is the variation of hand positions in different images which also leads to numerous representations of the features.
e. Variation of illumination conditions affects badly on the extracted hand skin region and extraction process.

VII. CONCLUSION

In this paper we have presented the work done on the static image. Hand gesture recognition is basically divided into two main steps. First is skin detection and extraction of particular sign area. For skin detection we use gray world algorithm. Good segmentation process leads to efficient features extraction process and latter it plays an important role in detection process. The recognition process suffers from many problems like rotation problem, background problem, scale problem etc. This work can be further extended to dynamic hand gesture recognition by using the sequence of images.

REFERENCES

[1.] Rafael C. Gonzalez, Richard E. Eoods, Steven L. Eddins," Digital Image Processing Using Matlab".
[2.] Rafiqul Z. Khan, Noor A. Ibrahim, (2012). "Survey on Gesture Recognition for Hand Image Postures", International Journal of Computer And Information Science, Vol. 5(3), Doi: 10.5539/cis.v5n3p110
[3.] N. Ibrahim, M. Hasan, R. Khan, P. Mishra, (2012). "comparative study of skin color based segmentation techniques", Aligarh Muslim University, A.M.U., Aligarh, India.
[4.] Mokhtar M. Hasan, Pramoud K. Misra, (2011). "Brightness Factor Matching For Gesture Recognition System Using Scaled Normalization", International Journal of Computer Science & Information Technology (IJCSIT), Vol. 3(2)
[5.] M. M. Hasan, P. K. Mishra, (2011). "HSV Brightness Factor Matching for Gesture Recognition System", International Journal of Image Processing (IJIP), Vol. 4(5).
[6.] G. R. S. Murthy, R. S. Jadon. (2009). "A Review of Vision Based Hand Gestures Recognition,"International Journal of Information Technology and Knowledge Management, vol. 2(2), pp. 405- 410.
[7.] P. Garg, N. Aggarwal and S. Sofat. (2009). "Vision Based Hand Gesture Recognition," World Academy of Science, Engineering and Technology, Vol. 49, pp. 972-977.
[8.] Jag dish Raheja, Radhey Shyam and Umesh Kumar, Hand Gesture Capture and Recognition Technique for Real-time Video Stream, In The 13th IASTED International Conference on Artificial Intelligence and Soft Computing (ASC 2009), September 7 - 9, 2009 Palma de Mallorca, Spain.
[9.] Reza Hassanpour, Stephan Wong, Asadollah Shahbahrami, Vision Based Hand Gesture Recognition for Human Computer Interaction: A Review, IADIS International Conference on Interfaces and Human computer Interaction. 25 - 27 July 2008 Amsterdam, Netherlands.
[10.] Mahmoud E., Ayoub A., J¨org A., and Bernd M., (2008). "Hidden Markov Model-Based Isolated and Meaningful Hand Gesture Recognition", World Academy of Science, Engineering and Technology 41.

[11.] Thomas B. Moeslund and Erik Granum, (2001). "A Survey of Computer Vision-Based Human Motion Capture," Elsevier, Computer Vision and Image Understanding, Vol. 81, pp.

[12.] Jones, M.J., Rehg, J.M.: Statistical color models with application to skin detection. international Journal of Computer Vision (IJCV) 46(1) (2002) 81–96.