Analysis of Various Methodology of Hand Gesture Recognition System using MATLAB

Komal Hasija[1], Rajani Mehta[2]

Abstract— *Recognition is a very effective area of research in regard of security with the involvement of biometric analysis, human computer interface and digital image processing. Humans inherently use faces to recognize individuals and similarly gestures are used in non-verbal communication to efficiently and effectively express thoughts. So, in order to migrate the natural means of communication by gesture into computer can setup a good move in making systems more interactive.*

Keywords: Artificial Intelligence (AI), Gesture, Human Computer Interaction (HCI).

I. INTRODUCTION

Gesture detection is the process of finding and extracting features within images or videos and hence Digital Image Processing will be the dominating tool to implement it practically. This area itself is a branch of artificial intelligence (AI) whose objective is to emulate human intelligence. The area of image analysis is in between image processing and computer vision.

As a digital image is composed of a finite number of elements, each of which has a particular location and value, these elements are called picture elements or pixels and the field dealing with the processing of digital images by means of digital computer is Digital Image Processing.

Digital image processing stems from two principal application areas

1. Improvement of pictorial information for human interpretation.
2. Processing of image data for storage, transmission, and representation for autonomous machine perception.

Various traditional input devices are available for interaction with computer, such as keyboard, mouse, and joystick as well as touch screen; yet these are not considered natural interface.

To develop human hand gesture recognition system which must be able to recognize a few hand gesture as an image format and then after processing, launches the application associated with the gesture respectively, regardless the person hand sizes and other external causes. Hand gestures have been widely used in the deaf community as the major communication media called sign language. [7] Natural Human Computer Interaction (HCI) is the demand of today's world. Survey and Sign language study shows that from various gesture communications modality, the hand gesture is the most easy and natural way of communication. This paper organized as follows:

II. COLOR SPACES

RGB: RGB is a color space originated from CRT (or similar) display applications, when it was convenient to describe color as a combination of three colored rays (red, green and blue). It is one of the most widely used color spaces for processing and storing of digital image data. However, high correlation between channels, significant perceptual non-uniformity, mixing of chrominance and luminance data makes RGB not a very favorable choice for color analysis and color-based recognition algorithms. [2]

Normalized RGB: Normalized RGB is a representation, which is easily obtained from the RGB values by a simple normalization procedure:

$$r = \frac{R}{R+G+B} \qquad g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B}$$

HSI, HSV, HSL - Hue Saturation Intensity (Value, Lightness)

Hue-saturation based color spaces were introduced when there was a need for the user to specify color properties numerically. They describe color with intuitive values, based on the artist's idea of tint, saturation and tone. Hue defines the dominant color (such as red, green, purple and yellow) of an area; saturation measures the colorfulness of an area in proportion to its brightness. The "intensity", "lightness" or "value" is related to the color luminance. The intuitiveness of the color space components and explicit discrimination between luminance and chrominance

properties made these color spaces popular in the works on skin color segmentation [3] [2]

$$H = \arccos \frac{\frac{1}{2}((R-G) + (R-B))}{\sqrt{((R-G)^2 + (R-B)(G-B))}}$$

$$S = 1 - 3\frac{\min(R,G,B)}{R+G+B}$$

$$V = \frac{1}{3}(R+G+B)\ \frac{1}{2}((R-G) + (R-B))$$

YCbCr: It is an encoded nonlinear RGB signal, commonly used by European television studios and for image compression work. Color is represented by luma (which is luminance), constructed as a weighted sum of the RGB values, and two color difference values Cr and Cb that are formed by subtracting luma from RGB red and blue components.

$$Y = 0.299R + 0.587G + 0.114B$$
$$Cr = R - Y$$
$$Cb = B - Y$$

The transformation simplicity and explicit separation of luminance and chrominance components makes this color space attractive for skin color modeling.

III. SKIN MODELING

The goal of skin color modeling is to build a decision rule that will discriminate between skin and non-skin pixels. This is usually accomplished by introducing a metric, which measures distance (in general sense) of the pixel color to skin tone. The type of this metric is defined by the skin color modeling method.

Parametric skin distribution modeling: The most popular histogram-based non-parametric skin models require much storage space and their performance directly depends on the representativeness of the training images set. The need for more compact skin model representation for certain applications along with ability to generalize and interpolate the training data stimulates the development of parametric skin distribution models. [3]

Non-Parametric skin distribution modeling: The key idea of these skin modeling methods is to estimate skin color distribution from the training data without deriving an explicit model of the skin color. The result of these methods sometimes is referred to as construction of Skin Probability Map. [3]

IV. LITERATURE REVIEW

Hand gestures have been widely used in the deaf community as the major communication media called sign language. [7] Natural Human Computer Interaction (HCI) is the demand of today's world. Survey and Sign language study shows that from various gesture communications modality, the hand gesture is the most easy and natural way of communication. Gesture-based interaction was firstly proposed by M. W. Krueger as a new form of human computer interaction in the middle of the seventies Krueger, 1991. Real-time vision-based hand gesture recognition is considered to be more and more feasible for Human-Computer Interaction. In Real-time Vision based Hand Gesture recognition system, hand tracking and segmentation are most important and challenging steps towards gesture recognition. [8]

The keyboard and mouse are currently the main interfaces between man and computer. In other areas where 3D information is required, such as computer games, robotics and design, other mechanical devices such as roller-balls, joysticks and data-gloves are used. Humans communicate mainly by vision and sound, therefore, a man-machine interface would be more intuitive if it made greater use of vision and audio recognition. Another advantage is that the user not only can communicate from a distance, but need have no physical contact with the computer. However, unlike audio commands, a visual system would be preferable in noisy environments or in situations where sound would cause a disturbance. The visual system chosen was the recognition of hand gestures. The amount of computation required to process hand gestures is much greater than that of the mechanical devices.

A general problem in computer vision is building robust systems, where higher-level modules can tolerate inaccuracies and ambiguous results from lower-level modules. This problem arises naturally in the context of gesture recognition: existing recognition methods typically require as input, either the location of the gesturing hands or the start and end frame of each gesture. Requiring such input to be available is often unrealistic, thus making it difficult to deploy gesture recognition systems in many real-world scenarios. [9]

We organize the discussion into the following main components based on the general view of a gesture recognition system as [10]:

- Gesture Modeling: The approach used for modeling plays a pivotal role in the nature and performance of gesture interpretation.
- Gesture Analysis: The goal is to estimate the parameters of the gesture model using measurements from the video frames or static images of a human operator engaged in HCI. An analysis stage is used to

compute the model parameters from the image features that are extracted from single or multiple video input streams.

- Gesture Recognition: Here, the parameters are classified and are interpreted in the light of the accepted model and perhaps the rules imposed by some grammar.
- Gesture-Based Systems and Applications: CD Player Control Panel, Computer Game Control, Robot manipulator control, and Hand sign recognition are some of the realistic systems using hand gesture recognition.

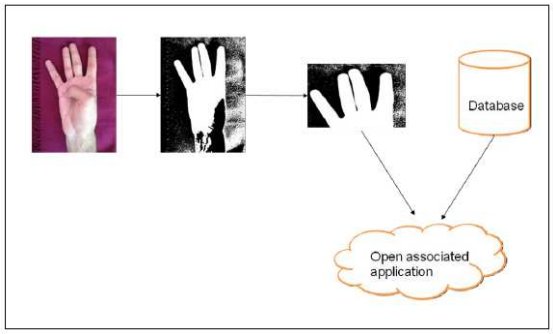
After having a view point on the domain the major problems which are faced in developing such system are as:

- Involvement of other factors with the input hand image i.e. additional objects in background, lightning conditions etc.
- Efficiency of algorithm depends on hardware i.e. camera quality.
- It is difficult to formulate an effective recognizing and matching package.
- Efficiency of algorithm also depends on the distance between acquisition device and human hand used for gesture.

V. WORKING METHODOLOGY

The image processing pipeline used for hand gesture recognition system involved following:

1) Hand image acquisition: The picture is taken by camera, digitized and prepared for further processing. It is usually done with the help of frame grabber. A frame grabber can be described as a GUI or interface which allows user to take snapshots by initializing the camera hardware and trigger manually for taking snaps.
2) Hand image segmentation: In this step those area in the image which represents the skin part of hand are separated from the background by applying different approaches using different color models. Using different color models helps in comparing the efficiency of segmentation on basis of models respectively.
3) Feature extraction: Aim of this step is to derive smallest possible amount of features out of the segmented hand region, in order to differentiate the different gestures. As in this we use output of previous step and for effective utilization, result image is converted into binary and then parameters as pixel intensities & quantity helps in forming base for feature extraction followed by cropping and standardization.

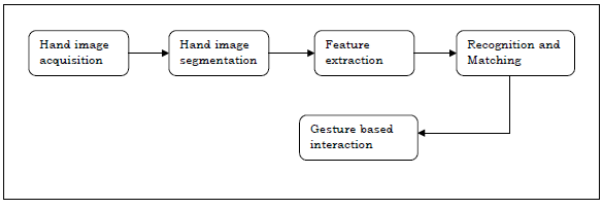


4) Recognition and Matching: By this step we conclude the exact gesture and by applying suitable matching technique results the recognition of gesture based on database images. For efficient matching algorithm we can implement neural networks, skeleton matching, chain coding, or can define algorithm which is capable of fetching best match for the input one.

Here basically we applied the matching/recognition phase on the basis of presence of number of white pixels which are of the binary image obtained from image segmentation operation on input image. Then this magnitude of number of pixels is matched with those previously stored images in database.

As the match is found out then user is supplied with all in between processed images along with the execution of application associated with gestures.

5) Gesture based interaction: Finally after obtaining the suitable match with respect to the gesture, associates the gesture with the assigned functionality i.e. launching the application.



VI. WORKING OF SYSTEM

A gesture recognition system is one which is capable of perceiving gesture by a means and then makes the system to act accordingly. So, here introducing a system which will recognize the gesture on the basis of image and after

processing on image the associated application is executed. The image on which processing is being carried out is a static part of human hand gesture in the form of signals by fingers.

The following steps are being carried out:

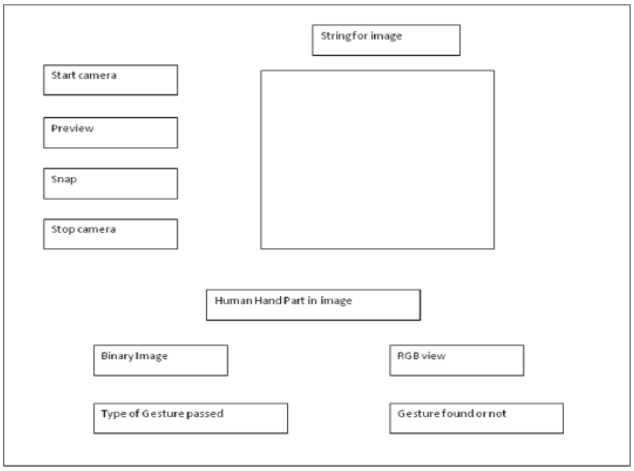
• Firstly the image is being capture/acquired with the help of webcam or external acquisition device, the condition applied is the background must be uniform.

• Then after the image is processed and the human hand is being separated out from the background i.e. segmentation is done.

• Then in order to match the supplied gesture with that in the database certain features are to be extracted out. Here we are concerned about the number of white pixels as the outcome from the previous step we get a binary image.

• So, on the basis of number of white pixels the particular gesture is recognized to that of images present in database.

• As soon as the gesture is recognized the associated application is being executed.



VII. CONCLUSION

The objectives and goal of this project are achieved successfully. We successfully implemented the image segmentation and recognition part in MATLAB. Where the images are static and provides interface which is user friendly. And the most important thing is that there is no extra hardware to perform all these tasks.

But taking the dynamic environment in mind this project is not enough robust and safe to guarantee result. Here we only considered limited set of gestures but if we take into account more number of gestures then we can facilitate more commands and the system will become more interactive.

As the system can serve in vital ways so the system should not involve any other external part/hardware except computer system equipped with webcam. This helps in keeping the cost minimum and everyone able to use this application also able to own.

REFERENCES

[1] Rafael C. Gonzalez and Richard E.Woods. 'Digital Image Processing', India, Published by Dorling Kindersley (India) Pvt. Ltd, 2011, Print.

[2] Vladimir Vezhnevets, Vassili Sazonov Alla Andreeva. A Survey on Pixel-Based Skin Color Detection Techniques.2003, Print.

[3] Zarit, B. D., Super, B. J., and Quek, F. K. H. 1999. Comparison of five color models in skin pixel classification. In ICCV'99 Int'l Workshop on recognition, analysis and tracking of faces and gestures in Real-Time systems, 58–63.

[4] J.R.Parker, "Algorithms for Image Processing and Computer Vision", Edition 2, 2011, Print.

[5] Rafael C. Gonzalez and Richard E.Woods. 'Digital Image Processing – Image Segmentation', India, Published by Dorling Kindersley (India) Pvt. Ltd, 2011, Print.

[6] Web link – www.mathworks.in/products/matlab/

[7] Chung-Lin Huang*, Ming-Shan Wu, Sheng-Hung Jeng. Gesture recognition using the multi-PDM method and hidden Markov model, Elsevier, 2000, Print.

[8] Archana S. Ghotkar & Gajanan K. Kharate. 'Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction', International Journal of Human Computer Interaction (IJHCI), Volume (3): Issue (1): 2012, Print.

[9] Jonathan Alon, Vassilis Athitsos, Quan Yuan, and Stan Sclaroff. 'A Unified Framework for Gesture Recognition and Spatiotemporal Gesture Segmentation', Pre-print, IEEE Transactions of Pattern Analysis and Machine Intelligence (PAMI), September 2009, Print.

[10] Vladimir I. Pavlovic, Rajeev Sharma, and Thomas S. Huang. 'Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review', IEEE TRANSACTIONS On Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, July 1997, Print.

[11] Xiaoming Yin and Ming Xie. 'Hand Posture Segmentation, Recognition and Application for Human-Robot Interaction', Human-Robot Interaction, Book edited by Nilanjan Sarkar, ISBN 978-3-902613-

13-4, pp.522, September 2007, Itech Education and Publishing, Vienna, Austria.