International Journal of Current Engineering and Technology
ISSN 2277 - 4106

Available at http://inpressco.com/category/ijcet

Research Article

A Hand Gesture Recognition using Feature Extraction

Ashis Pradhan[a]*, M.K.Ghose[a] and Mohan Pradhan[a]

[a]CSE Department, SMIT, Majhitar, East-Sikkim, India

Accepted 13 Nov. 2012, Available online 1Dec.2012, **Vol.2, No.4 (Dec. 2012)**

Abstract

Information communication between two peers can be done using various mediums. These mediums can be either linguistic or gestures. The development of procedure for realizing gestures into meaningful information plays a pivotal role in instances where linguistic feature cannot be taken as a basis and gestures can be used as the alternative for the conveying the same. This paper presents a very simple and efficient approach for recognizing the hand gesture that represents numbers from zero to nine. The work basically represents the active and in-active fingers with binary value 0 and 1 respectively, in different combination for representing different numbers. The method of representing the hand gesture in binary pattern contributes a lot for increasing the performance of classification process. The binary Support Vector Machine (SVM) is considered as a recognition tool.

Keywords: *Information Communication, linguistic feature, hand gesture, binary pattern, support vector machine.*

1. Introduction

Digital Cameras are now integrated into personal computers, mobile cellular devices and handheld systems. These devices usually include a powerful microprocessor, capable of performing millions of computations per second. The technology on digital cameras and microprocessors are advancing rapidly that it is possible to create a human computer interfaces using these resources for recognition of user gestures. The Gesture recognition interface acts as a communication channel between humans and machines. The human-machine interaction is similar to human-human interaction, in which, the valuable information are communicated using the human organs like hand gesture, head movement, face expression, voice communication and overall body posture.

The design of a gesture recognition system should be based on common hardware support such as web-cams or mobile-integrated cameras, to be applicable to current PCs, mobile devices, Digital Cameras, etc. While designing such systems, certain parameters have to be included, so that the system will be able to operate under complex or non-uniform background, i.e., different light intensity and noisy environment, etc.

2. Literature survey

There are several approaches that have been used to design a hand gesture recognition system. In all the approaches, the primary focus was given to feature extraction of the hand gesture and it was found that the better feature extraction step is performed, better will be the performance of classification. Ilan Steinberg et. al, proposed a method of recognizing hand gesture, by considering each pixel of the binary image as a feature. To reduce the complexity of computation for classification, the size of the images was reduced. Omlin and Vapnik et. al., have recognized user hand gesture using color histogram as their feature, in which, the whole image was divided into smaller blocks and plot the corresponding histogram, for classification. Mokhtar M. Hasan et. al., observed that the finger-tip can also be used as a feature for recognition.

Nasser H. Dardas in, used a bag of feature for classifying hand gesture. In this approach, all the key points of the training images were extracted and these key points are mapped to its corresponding histogram feature vector for classification.

Oleg Rumyantsev et. al introduced an efficient method for recognizing hand gesture, based on PCA method, in which, the test image and database images were represented using Eigen values and the hand gestures were correctly classified using the Euclidean distance of these Eigen values.

3. Proposed methodology

This project basically deals with the design of a system that acquires a user's hand gesture and classifies it based on the predefined hand gestures, stored in a database.
The design of a system is basically divided into parts, namely, pre-processing and classification phase.
The figures 1 shown below are the list of gestures that the system will recognize it correctly:

* Corresponding author: Ashis Pradhan

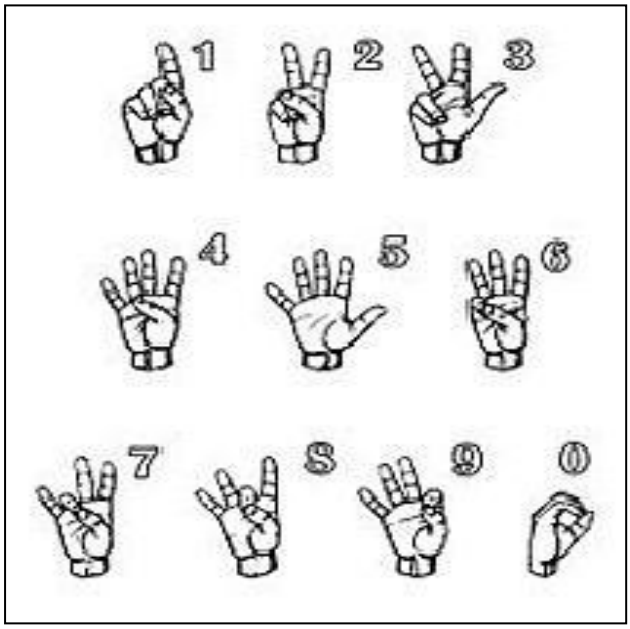


Fig. 1Gesture representing Numbers

The system will include low-resolution web cam for capturing the hand gestures and an algorithm that processes the acquired images and then classifies the hand gesture correctly. The work mainly emphasizes on the feature extraction from the hand gestures and use that features in the recognition algorithms. Initially, the system will contain a setup procedure, in which, the algorithm is trained on given training set, based on significant feature extracted for different hand gestures. Once the setup in completed successfully, the system will be able to classify the given hand gesture based on the knowledge acquired during the training phase.

The design of hand gesture recognition system is broadly divided into two phase. The first phase is the preprocessing phase and the second phase is the classification phase. The efficiency of the Classification

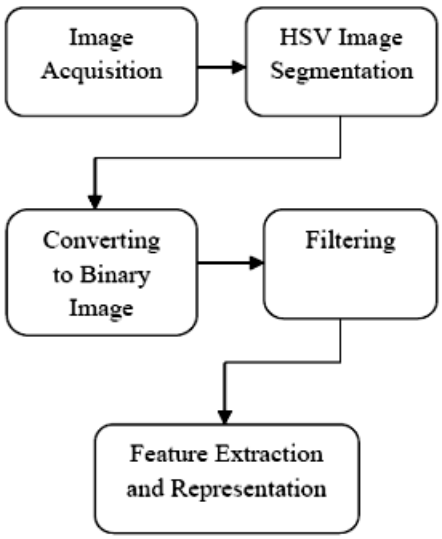


Fig. 2 Pre-processing Phase

phase entirely depends on the preprocessing phase i.e., better the task performed in pre-processing phase, better will be the performance of classification phase. So, all the tasks in pre-processing phase are to be carried out properly.

The figures 2 shown below are the list of task for pre-processing phase.

The main purpose of the pre-processing stage is to:

- Extract the only hand gesture from an image.
- Remove the noises (if present) and unwanted region.
- Process the extracted image to form a binary image and
- Extract the distinguishable significant features from the processed image, to form a feature set for classification.

Basically , in training phase, the database which consists of list of unique binary pattern for different hand gesture representing some specific number between 0 and 9 is created and is represented in some form.

The general methodology for a design of a system is shown in the figure 3 below:

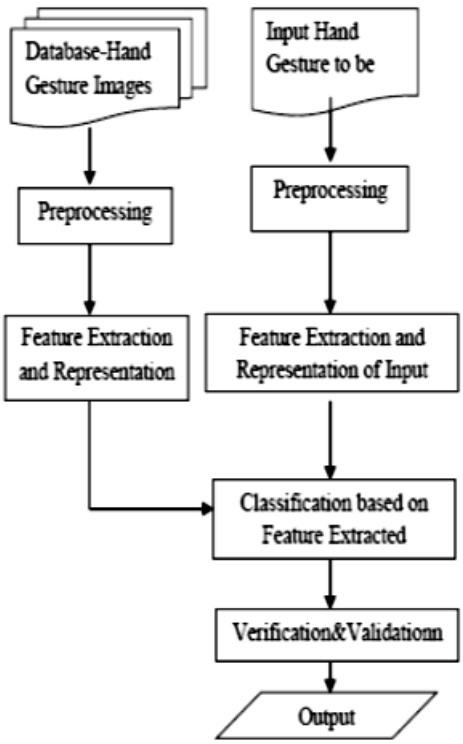


Fig. 3 General flow diagram of a System

Based on the literature survey, for classification, there are lots of efficient algorithms that are already used. They are Gradient, PCA (Principal Component Analysis) and SVM (Support Vector Machine). But, the paper mainly focuses on SVM, which is a machine learning algorithm. Initially, the traditional methods will be followed in preprocessing step, for preparing an image for feature extraction. Once a prepared image (i.e. noise free image), is successfully

achieved, the significant features representing a different hand gestures are extracted and represented to include in the classification algorithm. This is because the more features we include in the algorithm the more will be the accuracy of classification.

4. Experimental setup

As stated earlier, the set up procedure for a design of a system was basically divided into two parts, pre-processing and classification part. However, the pre-processing part plays an important role for classification. The first step in pre-processing part is the image acquisition step, in which, the scene containing hand portion as a major object was stored. During image acquisition, a particular constrain is included in a system, i.e., the images for training as well as testing were taken at equidistant level. Once an image containing hand portion only is stored, the second step in preprocessing was to extract the hand portion only from uniform or non-uniform background. This task for the extracting the hand portion was accomplished by considering the certain range of pixel values that represents skin color of hand. In this process an assumption is made that is, the color of background or the color of objects in background should not be a similar to that skin color.

To reduce the complexity of feature extraction for hand gestures, the output image of hand portion extraction process was converted into binary image using a thresholding technique. The image so formed may contain some noises. The appearance of such noises may be due to the atmospheric condition at which the images were taken and also the type of source that is used for capturing an image. Hence to remove the noise, the median filter was generally used. Also, depending upon the percentage of noise present in the binarized image, it was found that the morphological operator namely image erosion can also be used for remove small sharp unwanted details (i.e., noise) from an image. The extent of noise removal is directly proportional to the extent to which a system can be trained correctly and hence classifies the input hand gestures correctly.

The next immediate process after the removal of noise is the feature extraction process, in which, the different techniques are applied based on the type of feature to be extracted. There are various different kinds of distinguished features that can be extracted from the filtered image, but the paper focuses only on the active and in-active finger which is represented by 1 and 0 respectively. So to identify the active and in-active finger, the first task was to identify the finger tip. To identify the finger tip, the skeletonization technique was used to thin the finger portion of hand. In this technique, the thinned image was obtained that represents the finger portion only by singly connected pixels, the co-ordinate values for the finger tip was stored. The pixel is considered as a finger tip that has only one neighbor in 3x3 window mask. For such pixel, the co-ordinate values are stored for different active finger. After storing the co-ordinate values of different active finger tip, the next step is to find the centroid of a hand. The centroid of a hand is calculated as follows:

$$\overline{X} = \frac{\sum_{i=0}^{k} xi}{k}, \qquad \overline{Y} = \frac{\sum_{i=0}^{k} yi}{k}$$

Where $(\overline{X}, \overline{Y})$ represents the centroid of the hand, xi and yi are x and y coordinates of the ith pixel in the hand region and k denotes the number of pixels that represent only the hand portion. In the next step, the distance between the centroid and the finger tip was calculated. For distance, the following Euclidean distance was used:

$$\text{Distance} = \sqrt{(x2 - x1)^2 - (y2 - y1)^2}$$

Where (x1, x2) and (y1, y2) represent the two co-ordinate values.

To increase the efficiency of the system, a certain degree of deviation was included in the distance parameter, so that a system can recognize the hand gesture correctly even there is a small bents of finger in the gesture to be recognized. Through manual process, it was found that the distance from centroid to different finger tip were different. So, this distance can be used as a significant feature for identifying the active or in-active finger. Once the combination of different active finger are found based on distance from centroid, these active and in-active finger were represented with the help of binary code. That is, the active fingers were represented by 1 and the in-active finger was represented by 0. So, for different kinds of gesture representing different number, there will be a different combination of 0's and 1's. For example, the binary code for gesture 1 is will be 01000, in which, the second finger is the only active finger and rest of the fingers are in-active. In order to maintain consistency for binary representation of fingers, the numbering of finger is done form left to right.

Using this method of representing the hand gesture that is in the form of binary code, a repository is created for different gestures. The different hand gesture representing different number will have different combination of binary numbers. Hence, the repository will contain the distinguishable binary pattern. Finally, the system is trained based on these binary patterns.

For testing phase, the same steps are followed in order to extract the binary pattern for the input hand gesture to be tested. And the extracted binary pattern for each input hand gesture is compared with the content of repository created in the training phase. Finally, if there is match of the input binary pattern with any one of the binary pattern in repository, the index of the matched binary pattern from repository is displayed which is considered as the matched gesture that represents a number between 0 and 9 by the system. The Binary SVM is used for classifying, in which, the input image is classified as either correct or incorrect. Therefore, number of binary SVM comparison has to be made with repository. Wherever there is a match with any binary pattern in repository, the system will display its corresponding recognized gesture number.

5. Results and discussion

The result of the system is very simple, but the paper aims at the efficient way of representing data or information for comparison and focuses on the method to reduce the complexity during comparison for a hand gesture image to be tested and hence to increase the correctness of the system. The figure 5 shown below is the abstract view of result of the system.

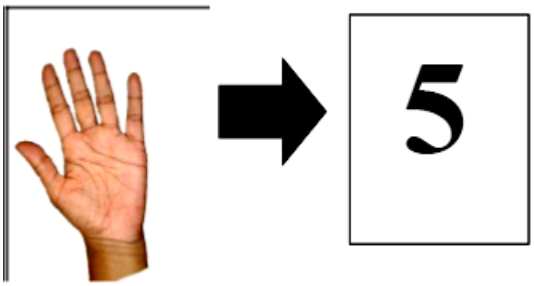
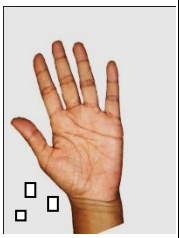
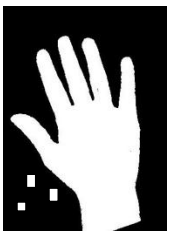
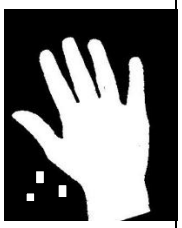


Fig. 5 Input Gesture and Output Gesture

The lots of task are carried out to design a system that will correctly classify and give the above sample output for its corresponding input.

Table 1 Sample Result for Number 5

Process	Input	Output
RGB to Binary		
Noise Removal		
Skeleton-ization		
Repository Creation		**11111**
Pattern Comparison	**11111**	**5**

The table 1 shown above lists of task performed serially for a sample image that represents number five and classifies it correctly.

The first task after image acquisition was to extract only the hand portion from any complex or uniform background, which was accomplished with the help of skin color of hand and convert the resultant image to binary image. The thresholding technique was used to convert image into binary image. While working in real time, there exists some noise in an image. This appearance of noise may be due to low resolution cameras that it used during image acquisition step or due to the environmental factor. Hence, the image in which the hand portion is extracted also contain some noises. So, the median filtering technique, followed by erosion morphological operation is used on the segmented image to remove noise and to remove un-necessary small sharp detail in an image, respectively. The outcome of these operations is the noise free regularized image, in which, the maximum area in an image is covered with the object of interest only. Finally, the resultant image is processed to determine the finger tips and store its co-ordinate values. The Skeletonization technique is used to determine the finger tip. In the next step, the centroid of a hand region is calculated. The significance for calculating the centroid is to determine the actively raised finger. The actively raised fingers were determined by calculating the distance between the centroid and the finger tip. Visually, it can be determined that the distance between centroid and the different fingers of a hand are different. Hence, this concept is used to create a binary pattern for actively raised finger in an image. The deviation of 10% (approxly) is included in the system for distance to determine actively raised finger. The active raised finger is marked 1 and the in-active fingers is marked 0. These different combinations of 0's and 1's represent any one number between 0 and 9. Thus, a database is created for different hand gesture that represents number.

The following lookup table represents the binary pattern for different hand gestures.

Table 2 Lookup table for Number Pattern

Binary Code	Equivalent Number
00000	0
01000	1
01100	2
11100	3
01111	4
11111	5
01110	6
01101	7
01011	8
00111	9

Initially, the system is trained by extracting binary pattern for different hand gestures representing numbers and thus creating a repository that is used later in recognition phase. After the creation of database, the testing phase starts. In testing phase, any input image is taken that represent number. Same tasks are carried out for the input image, as described above, to extract the binary pattern. Finally, this binary pattern is matched with the repository to recognize it as correct gesture and hence display the result.

Conclusion

The hand gesture recognition system that is designed is tested with different gestures and is able to classify it correctly. The input image to be classified during testing phase has to be taken at same distance as that of training phase. The system is able to successfully classify the hand gesture representing number and the system can be further extended to recognize alphabets, expressions, etc. Furthermore, the search procedure can be enhanced to increase the performance of the system. The designed system is able to classify only the static images which can be extended more to recognize hand gesture in videos as well.

References

A.Malima, E.Ozgur, M.Cetin (2006), A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control, *IEEE 14th Conference*, ISBN: 1-4244-0238-7, Date: 17-4-‘06.

Aryuanto Soetedjo (2008),SkinColor Segmentation Using Coarse-to-Fine Re-gion on Normalized RGB Chromaticity Diagram for Face Detection., *BIEICE Trans. Inf. & Syst.*, Vol.E91-D, No.10, October.

C.W. Omlin, M. Glaser (2009), Vision Based Static Hand Gesture Recognition, *Art and Literature category*, Date of posting: 12/16/2009.

E. Sanchez-Nielsen, et.al (2003), ,Hand gesture recognition for human machine interactionl, *Journal of WSCG*, Vol.12, No.1-3 (February).

Ilan Steinberg (2010), Tomer M. London, Dotan Di Castro,, Hand Gesture Recognition in Images and Video,, *Technion – IIT*, January 5.

Mokhtar M. Hasan , Pramod K. Mishra (2012),Real Time Fingers and Palm Locating using Dynamic Circle Templates, *International Journal of Computer Applications*, (0975 – 8887), Vol: 41– No.6, March.

Nasser H. Dardas and Nicolas D. Georganas (2011) ,Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and SVM, *IEEE Transactions On Instrumentation And Measurement*,Vol.60,No.11.

Oleg Rumyantsev, Matt Merati, Vasant Ramachandran (2012),Hand Sign Recognition through Palm Gesture and Movement, *Image Processing*, EE 368, Spring.

Olivier Chapelle, Patrick Haffner and Vladimir Vapnik (1999),SVM for Histogram-BasedImage Classification, *IEEE*, Vol. 10, No. 5, September.