Volume 4, Issue 3, March 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Hand Gesture Recognition through Rule Based Classification

Veerapalli Lathasree
Undergraduate Scholar, RGUKT
RK Valley, India 516329

Shaik Noor e Karishma
Undergraduate Scholar, RGUKT
RK Valley, India 516329

Ramakant Yadav
Faculty, ECE, RGUKT
RK Valley, India 516329

Abstract-The advancements in science and technology will have a real meaning if they can give an opportunity for the disabled to lead a normal life in our society. Vision for the blind, voice for the dumb and audibility for the deaf through man-machine interface would aid the less fortunate people to communicate with the outside world normally. This paper is a part of the project whose aim is to design a real-time system that would understand sign language accurately and give text output for the corresponding hand gestures of the dumb. Instead of traditional sensor based gloves, camera is used to capture the gesture information. A rule based approach is proposed to recognize the gestures.

Keywords: Human computer interaction, Vision based system, Hand gestures, Sign recognition.

I. INTRODUCTION

Sign language is mostly used a vision language by the deaf, dumb or people with any other kind of disabilities. Nevertheless, the society normally communicates with these groups of people with gesture or signs that are unofficial, causing communication breakdown and low self-esteem to these special needs people. A Human Computer Interface (HCI) system that can understand the sign language accurately aids in the communication of hearing impaired with the normal world. The system can further be enhanced to a sign-to-speech translator.

The aim of sign recognition system is the interpretation of the semantics that the hand(s) location, posture, or gesture conveys. Gloves based method is the traditional approach to Gesture Modeling. The person signs by wearing gloves and the system functions through sensing apparatus like wires, antenna and computer. This is a 3d model that gives a direct measurement of hand position, joint angle and orientation. The gloves are too expensive and are not user-friendly. With the advancements in Computer Vision and Pattern Recognition, Vision based solutions which are simple, easy and affordable to implement in real-time are practiced. The generalized block diagram of Vision based approach is as shown in the figure 1.

Hand Gesture Acquisition can be done by using a camera to grab images or video frame sequences of the signing person. The hand is cropped till the wrist to obtain hand gestures. Hand Detection includes Segmentation and Edge Detection. In hand gesture segmentation stage, regions which represent the hand gestures are to be distinguished from the background. Segmentation handles the challenges of vision based system such as skin color detection, complex background removal and variable lighting condition. Efficient segmentation is the key of success towards any gesture recognition.

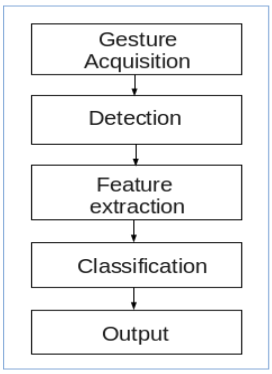

Figure 1: Hand Gesture Recognition System

Feature Extraction is used to calculate particular dimensions that capture the bulk of variation in the image data. Features that do not contribute towards predicting the response are discarded. Classifier identifies the hand gesture from the alphabets of the sign language. Typically, the larger the vocabulary is, the harder the recognition task becomes.

II. LITERATURE REVIEW

Various methods were proposed by researches to detect various hand gestures. These are broadly classified as segmentation, feature extraction, gesture recognition (classification) approaches. The process of extracting the desired gesture image from the rest of visual image is called segmentation. In literature, it is also termed as image localization. Segmentation can be done based on color, shape, motion and anatomical models of hand. Color based segmentation techniques were used at the initial stages. Most of the color segmentation techniques rely on Histogram Matching. Several color spaces like HSV, RGB, normalized RGB, YCrCb and YUV were provided, in order to approximate the "chromaticity" of skin (or, in essence, its absorption spectrum) rather than its apparent color value. Hand extraction is much easier in HSV (hue-saturation-value) color space as the hue component of hand can be easily threshold. The major drawback of color based segmentation is color of human skin varies greatly across human races or even between individuals of the same race. This is sensitive to quickly changing or mixed lighting conditions. In general, color segmentation can be confused by background objects that have a color distribution similar to human skin. Background subtraction can be done to overcome this problem. However, background subtraction is typically based on the assumption that the camera system does not move with respect to a static background. Sometimes we need preprocessing to distinguish between hands and faces in the case where faces are visible in the camera field of view [7].

Saengsri [1] in his paper Thai Sign Language Recognition used 5DT Data Glove 14 Ultra data glove which was attached with 14 sensors- 10 sensors on fingers and rest 4 sensors between the fingers which measures flexures and abductions respectively. The accuracy rate was 94%. Weissmann [2] used Cyber glove which measured features like thumb rotation, angle made between the neighboring fingers and wrist pitch. Limitations were that the system could recognize only single hand gestures. Limitation of glove based approach is it is unpractical; it limits the free motion of hand [2]. Shape of the hands can be used to extract them from the image. This is based on the edge detection techniques. Edges of an image are considered as a type of crucial information that can be extracted by applying detectors with different methodology. This technique is robust to the illumination and the variability of skin color. Edge detection results in large number of edges belonging to hands and also background. To eliminate the background edges it is combined with skin color technique. Motion based detection technique assumes only hand in the image is moving. The difference in luminance of pixels from two successive images is close to zero for pixels of the background. By choosing and maintaining an appropriate threshold, moving objects are detected within a static scene [7].

Orientations of hand, centroid of hand, angle between fingers are some features of hand. There have been many models for feature extraction like Linear discriminant analysis, PCA (Principal component analysis), wavelet packet decomposition (WPD). In linear discriminant analysis, p-dimensional space is projected on to a line. Lamar [3] in his research used PCA to obtain features like position of finger, shape of finger and direction of the image. The accuracy was about 93% and this system could recognize only single hand gestures. Generic Fourier descriptor and Generic Cosine Descriptor is used [5] for feature extraction as it is rotation invariant hand image leads to shifting of hand in polar space. Rotation invariance is obtained by only considering the magnitude of the Fourier coefficient. While using centroid as the origin translational invariance is achieved and finally ratio of magnitude to area scale invariance is obtained [5].

After feature extraction, proper classifiers were used to recognize gestures. Classification approaches can be classified into two categories namely, [4] machine learning based and rule based. Machine based approaches use set of examples to infer model of gestures. Different researchers used different approaches for classification like Support Vector machines (SVM), Artificial Neural Network (ANN), Fuzzy Logic, Euclidean distance and Hidden Markov model (HMM). Neural Network (ANN) was used as classifier to recognize a gesture [6]. Accuracy rate obtained was 97.4%. Jyothi Singha and Karen Das[6] used Support Vector Machines (SVM) as classifier. Recognition rate obtained was 95.2%. Rule based approaches use manually encoded rules to classify gestures.

III. PROPOSED ALGORITHM

1. Gesture Acquisition and Detection

Segmentation contributes to the major part of the Gesture Acquisition. Our current approach is a fusion of skin color and background segmentation. Face and hand of signer were successfully detected by using skin color segmentation. False detection of skin region in the uncontrolled background also occurs due to light variation. So background subtraction was used to find the difference between the hand gesture image and the background object[9].

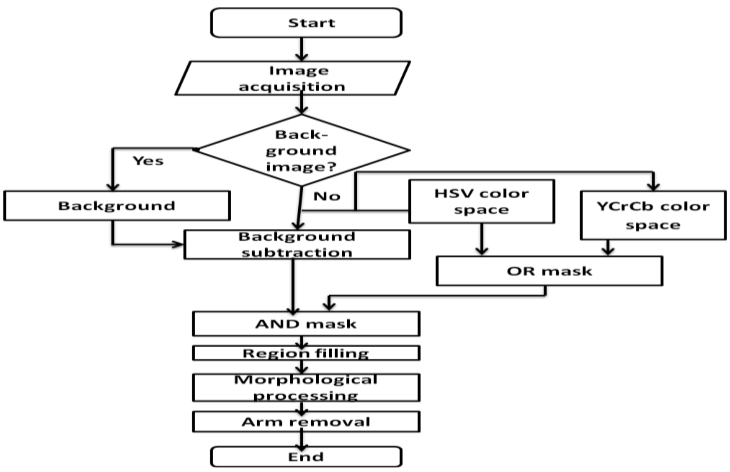

Figure 2: Flowchart of proposed segmentation algorithm

Segmentation process has direct impact on balancing accuracy-performance-usefulness trade-off of recognition systems. In accordance with the established background subtraction, the binary gesture image is obtained. Hand gesture area detected contained some interference region caused due to clothes. Biggest blob analysis is implemented to obtain hand region. The skin color detection method in HSV and YCrCb color spaces is used to make skin color segmentation on the foreground image. The hand region is highlighted after skin color detection. The two images of background subtraction and skin segmentation were multiplied. Region filling and morphological operations were performed to enhance the image. Then the hand gesture image is obtained as shown in fig 3(c).

Figure 3: a) Background Image b)Foreground Image c) Subtracted Image d)Segmented Image

After the segmentation of hand from the image, region filling is applied to remove the black spots in the image. Image cleaning is done in order to avoid the wrong calculations in the further steps.

2. **Feature Extraction**

Features used in the algorithm: Centroid, Orientation of the hand, Thumb position, Tip position of active fingers.

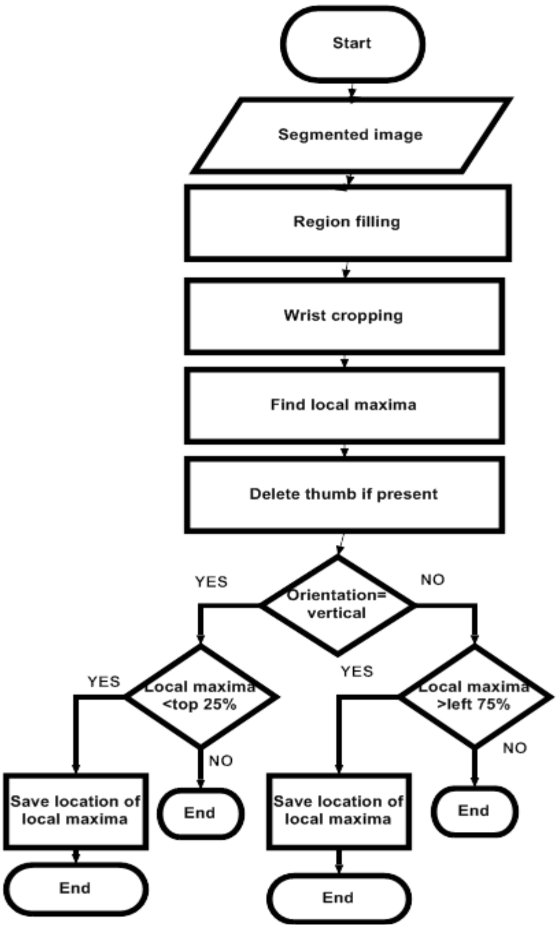

Figure 4: Flowchart of recognition algorithm

a. Centroid:

Centroid is the geometric center of the image it is calculated using the weighted average of the pixels. If the image is uniformly distributed centroid becomes center of mass.

$$M_{ij}=\sum\sum x^i y^j I(x,y)$$

Where M_{ij} is is the image moment.

I(x,y) is the intensity of pixels at the coordinate (x,y).

$$(X_C, Y_C) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}}\right)$$

Where $\{x_c,y_c\}$ are the coordinates of the centroid.

b. Orientation:

To find out the orientation of hand find the minimum and maximum row value of the image containing the white pixel and similarly find the minimum and maximum column of the image containing the white pixel .

Height = row_max – row_min= height of the image containing hand part.

Width = col_max – col_min = Width of the image containing hand part.

$$s = \left[\frac{Height}{Width}\right]$$

If S>=1 ↔ vertical image.

If S<1 ↔ horizontal image.

Wrist cropping is done to avoid the redundant information. The principle employed for wrist cropping is that the minimum number of white pixels is available in the row (column) of the vertical (horizontal) image at the wrist.

c. Thumb Detection:

The above algorithm is designed for the closed left hand gestures so thumb can be at left side and at the bottom of the image. Thumb is a significant feature to recognize some gestures in the gesture recognitition system. To detect thumb for the vertical image take the 10% at both the sides. Calculate the number of white pixels at both the sides. If the number of white pixels at left side is less than 5.5% of the total white pixels present in the image then there is a thumb. If we get both sides less than 5.5% then it is consider that there is no thumb. Similarly for the horizontal image calculate the number of white pixels at top and bottom 10%. If number of white pixels at bottom is less than 5.5% then there is thumb. If both sides it is less than 5.5% then it is consider that there is no thumb.

d. Detection of Fingertips:

To compute the finger tips, find the global maxima in the image. If local maxima lie below the centroid it is not a finger. If the image is vertical then check whether the local maxima lies below the 25% of the image height then the fingertip is consider to be active. If the image is horizontal and finger tip lies above the 75% of the image width then it is considered to be an active finger. From the above procedure we can find the number of fingers raised.

3. Classification

After detecting the count of active fingers, the position of the active fingers is to be determined[10]. The thumb region is omitted. The recognition of the hand gestures is done in the following sequence.

The palm region with four fingers is our region of interest. This region of interest is divided into four parts to find the position of active finger and the corresponding finger name. The locations of the local maxima saved during detection of finger tips are considered as 'tips' in the following approach.

a. Vertical image

Consider step size Δ = 0.25*(width of the image).

If 0 < tip < Δ ↔ little finger

*If Δ< tip < 2*Δ ↔ ring finger*

*If 2*Δ< tip < 3*Δ ↔middle finger*

*If 3*Δ< tip < 4*Δ ↔ index finger*

An array is created with [index finger, middle finger, ring finger, little finger] as variables and is updated with '1' if the finger is detected, else '0'.

b. Horizontal Image

Consider step size Δ = 0.25*(height of the image).

If 0 < tip < Δ ↔ little finger

*If Δ< tip < 2*Δ ↔ ring finger*

*If 2*Δ< tip < 3*Δ ↔ middle finger*

*If 3*Δ< tip < 4*Δ ↔ index finger*

An array is created with [little finger, ring finger, middle finger, index finger] as variables and is updated with '1' if the finger is detected, else '0'.

4. Output

For our experiment, we chose a set of 16 static hand gestures. These 16 gestures comprised of binary open/close configuration of the fingers excluding the thumb.

Figure 5: Sixteen static hand gestures

The binary code is assigned to each gesture with ‘1’ for the active finger and ‘0’ for the inactive finger.

Gesture Index	Segmented Image	Code	Recognition rate
0		0000	100%
1		0001	100%
2		0010	100%
3		0011	98%
4		0100	98%
5		0101	100%
6		0110	98%
7		0111	97%
8		1000	100%

9		1001	100%
10		1010	98%
11		1011	96%
12		1100	96%
13		1101	96%
14		1110	96%
15		1111	100%

Table 1: Recognition rate of Segmented 16 gestures

IV. DISCUSSION:

The average recognition time for the hand gestures is 1.6 seconds. Though the classification algorithm works fast, the segmentation of image from a complex background makes the recognition system. This method is adaptable to illumination changes and complex background, which makes it efficient over other algorithms. The next part of the project is detecting alphabets and thus, implementing our rule based classification system to the Indian Sign Language. In this project, the segmentation is done for both right and left hand gestures, but only left handed gestures are classified. We will back our classification algorithm adaptive to both the hands. We will also be focusing on the computation time in the future work.

V. CONCLUSION:

Gesture recognition system should balance accuracy-performance-usefulness trade-off according to the type of application, the cost of the solution and several criteria such as real-time performance, robustness, scalability and user-independence. Analysis of different research papers and journals indicates that the techniques implemented so far are not completely robust and are often sensitive to poor resolution, drastic illumination conditions, cluttered backgrounds and several other problems. There is a wide scope of future research to overcome the disadvantages with current state of systems and realize efficient gesture recognition systems. In coming future, our focus would be on implementing above rule based algorithm for alphabets of sign language with high detection rate and minimum computation time.

ACKNOWLEDGMENT:

The authors are very thankful to Sri Ramakant Yadav, Faculty, RGUKT for his inspiring guidance and constructive criticism.

REFERENCES:

[1] S. Saengsri, V. Niennattrakul, and C.A. Ratanamahatana, “TFRS: Thai Finger-Spelling Sign Language Recognition System” , IEEE, 2012, pp.457-462.
[2] J. Weissmann and R. Salomon, “Gesture Recognition for Virtual Reality Applications Using Data Gloves and Neural Networks”, IEEE, 1999, pp. 2043-2046.
[3] M. V. Lamar, S. Bhuyan, and A. Iwata, “Hand Alphabet Recognition Using Morphological PCA and Neural Networks”, IEEE, 1999, pp. 2839-2844.
[4] A. Birdal, R. Hassanpour, “Region based hand gesture recognition”.
[5] M. Pahlevanzadeh, M. Vafadoost, and M. Shahnazi, “Sign Language Recognition”, IEEE, 2007.
[6] Jyothi singha and Karen Das, “Hand Gesture Recognition Based on Karhunen-Loeve Transform”.
[7] X. Zabulis, H. Baltzakis and A. Argyros, “Vision Based Hand Gesture Recognition for Human-Computer Interaction”.
[8] M. Pahlevanzadeh, M. Vafadoost, and M. Shahnazi, “Sign Language Recognition”, IEEE, 2007.
[9] L. Howe, F. Wong and A. Chekima, “Comparison of Hand Segmentation Methodologies for Hand Gesture Recognition”, IEEE-978-4244-2328-6, 2008.
[10] M. Panwar and P. S. Mehra, “Hand Gesture Recognition for Human Computer Interaction”, Proceedings of IEEE, International Conference on Image Information Processing(ICIIP 2011), Wakhanghat, India, November 2011.

