Vision based Calculator for Speech and Hearing Impaired using Hand Gesture Recognition

Vishal Bhame
Student, Pune Institute of Computer Technology
Pune, India

R. Sreemathy
Prof., Pune Institute of Computer Technology
Pune, India

Hrushikesh Dhumal
Research Engineer,
Hyper-Ions,
Pune, India

***Abstract*— Even after more than two decades of development input devices such as data gloves, infrared cameras, many people still find the interaction with computers an uncomfortable experience. Efforts should be made to adapt computers to our natural means of communication: speech and body language. This paper is the proposal of a real time and fast command system through hand gesture recognition, using low cost sensors, like a simple personal computer and an USB web cam, so any user could make use of it in his industry or home. This paper describes the new methodology for vision based, fast and real time hand gesture recognition which can be used in many HCI applications. The proposed algorithm first detects and segments the hand region. Then using our novel approach, it locates the fingers and classifies the gesture. The proposed algorithm is invariant to hand position, orientation or distance from web cam. We have developed a gesture based mathematical tool (Calculator) as an application of proposed algorithm.**

***Keywords*— Features Extraction; Gesture Based Calculator (GBC); Hand Gesture Recognition (HGR); Human computer Interaction (HCI).**

I. INTRODUCTION

Human computer Interaction (HCI) is a branch of artificial intelligence, it is a scientific discipline that is concerned with the development of algorithms that take as input empirical data from sensors or databases, and yield patterns or predictions thought to be features of the underlying mechanism that generated the data. A major focus of HCI research is the design of algorithms that recognize complex patterns and make intelligent decisions based on input data. As the integration of digital cameras within personal computing devices becomes a major trend, a real opportunity exists to develop more natural Human-Computer Interfaces that rely on user gestures.

Vision-based automatic hand gesture recognition has been a very active research topic in recent years with motivating applications such as human computer interaction (HCI), robot control, and sign language interpretation. In particular, visual interpretation of hand gestures can help in achieving the ease and naturalness desired for HCI. HGR system confronts many challenges as addressed in [13], likeIllumination conditions, Background problem, Rotation problem, Scale problem and Translation problem.This has motivated a very active research area concerned with computer vision-based analysis and interpretation of hand gestures.

Accuracy and gesture-recognition speed depend on advanced software algorithms. Algorithm proposed in [1] uses wavelet transform and principal component analysis for face and hand gesture recognition on digital images. [3] proposed a hybrid algorithm which uses Gabor filter followed by Mel Scaling to get hand structure. In [5] CAMSHIFT algorithm is used to recognize alphabet characters (A-Z) in real-time from color image sequences.Mentioned algorithm gives effective performance but still they are computationally demanding. [7], [8] and [10] gives techniques to recognize the hand gesture in real time environment. It gives effective performance of face recognition and hand gesture recognition on digital images as well as image from video. It uses complex algorithms for recognition and hence lags in recognition speed.

In this paper we proposed a fast and simple algorithm for hand gesture recognition which can be used in real time HCI applications. We have also build a mathematical tool known as 'Gesture Based Calculator' (GBC) based on our proposed algorithm. GBC takes input from user in the form of sign language and produces the output. It is explained in detail under the section of 'HCI application' in this paper.

II. HAND GESTURE RECOGNITION

Like any recognition system, HGR uses collecting the input, preprocessing, feature extraction and finally the recognition algorithm in order to recognize the input gesture. Figure1 describes the flow chart of our proposed algorithm. Proposed method mainly consists of image capturing, image segmentation, Region of Interest Extraction followed by finger counting logic as recognition.

A) Image Capturing:

iBall C12.0 Webcam is used for capturing hand image with resolution 640x480. Image is captured in RGB Colorspace format and resized to 160x120. Resizing is necessary to reduce computational time. Resized image is stored in Tagged Image File Format ('.tiff'). TIFF files are larger than JPEG files, but they retain the full quality of the image and uses lossless compression scheme.

B) Image Segmentation:

After acquiring the image, the next phase of a tracking system involves separating potential hand pixels from non-hand pixels. Various methods are given in [4], [9] and [12] to achieve this task. Here we used a simple background subtraction scheme along with skin color mapping for detecting and tracking the hand. Before performing segmentation, we first convolve all captured images with a 5x5 Gaussian filter.

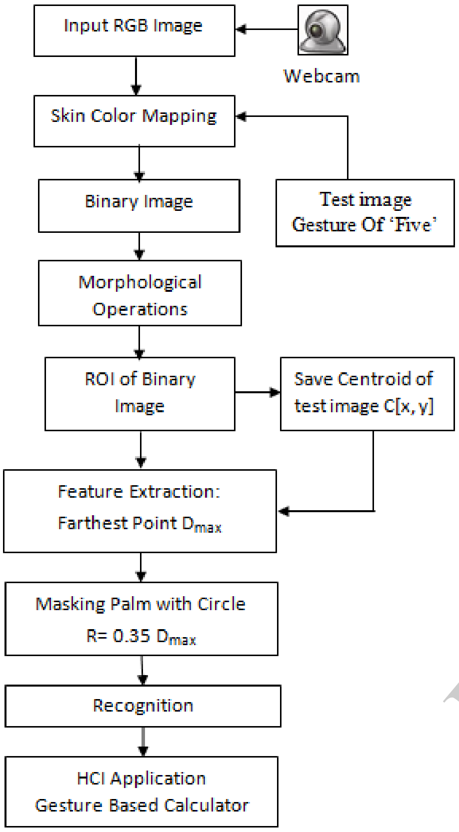


Figure1. Flowchart of proposed algorithm

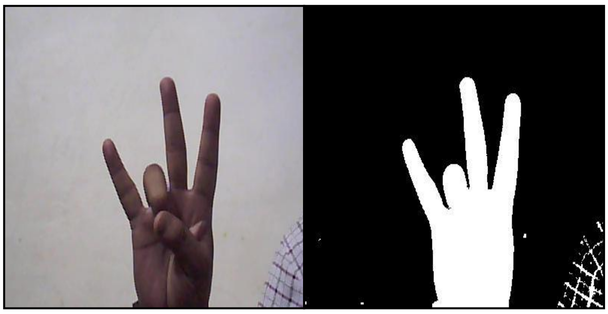


Figure2. (a) Input RGB image, (b) Segmented image

and then scale this filtered image by one half in each dimension in order to reduce noisy pixel data.

Background subtraction scheme segments any potential foreground hand information from the non-changing background scene. For each pixel in image I, we compute the foreground mask image I_F as in equation

$$I_F = \begin{cases} 255 & \text{if} |I_i - I_B| > \sigma_B \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Where, σ_B is a fixed threshold to differentiate foreground data from background data. Value of σ_B of 8 provides good result which is found after several experiments. Figure 2 shows captured RGB image and the result image after segmentation followed by morphological operations. Morphological operations like erosion and dilation are needed to reduce the noise in the image.

C) Region of Interest Extraction:

After segmentation, the binary image contains hand as well as non-hand region. White pixels represent the hand region while black represents background. The hand is extracted as biggest continuous blobfrom the binary image by using the bilinear interpolation. The image is once again resized to 80×60 and stored as blob.tiff for further processing. Figure 3(a) shows blob image of captured input image.

D) Recognition(Finger counting):

It is required to assign a meaning to image. In this case, it is a finger count ranging from 0 to 9. This stage explains the logic behind finger counting. It is carried out using following steps:

i) Calculating the Centroid C [x, y] of blob image: In this step Centroid point location C[x, y] of the blob image is calculated and is stored in workspace for future reference. Centroid point location is given by C[x, y] and is calculated as:

$$x = \frac{\sum_{i=0}^{k} X_i}{k} \ , \ y = \frac{\sum_{i=0}^{k} Y_i}{k} \quad (2)$$

where X_i and Y_i are x and y coordinates of the i^{th} pixel in the hand region, and K denotes the number of pixels in the region.

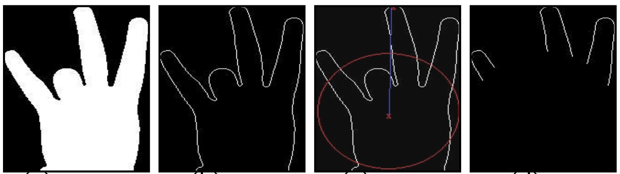
ii) Calculating the farthest point distance $D_{max}$from Centroid point:

In this step we calculate the distance from Centroid point to farthest pixel point Q[q1, q2] on the counter of the hand region using Euclidean distance. To get the counter image we applied canny's edge detection algorithm on blob image. Figure 3 (a) and (b) shows the blob image and its edge image. Euclidean distance is calculated by following equation:

$$d_{(C,Q)} = \sqrt{(x - q_1)^2 + (y - q_2)^2} \quad (3)$$

Where C(x, y) and Q (q1, q2) are two points on image and Dmax (C, Q) is the distance between these two points.

iii) Constructing a circle centered at C [x, y] that intersects all the fingers that are active in the count: We draw a circle whose radius is 0.35 of the farthest distance D_{max} from the Centroid C[x, y].Such a circle is likely to intersect all the fingers active in a particular gesture or 'count'. We masked the circle region with pixel value 0 (Non hand region).The resulting image is saved in workspace as Masked_image.tiff.



(a) (b) (c) (d)

Figure 3.Logic for Finger counting (a)Blob image; (b) Edge image; (c) Circle intersecting fingers; (d) Masked image

iv) *Recognition:*

It is the stage in which we apply a meaning to the image. In previous step we got the binary masked image which contains isolated fingers shown in figure 3(d). These fingers (white objects) are counted and stored in workspace or passed to HCI application GBC.

III. HCI APPLICATION

In this paper, we developed a mathematical tool (calculator) for hearing and speech impaired people based on our proposed algorithm.

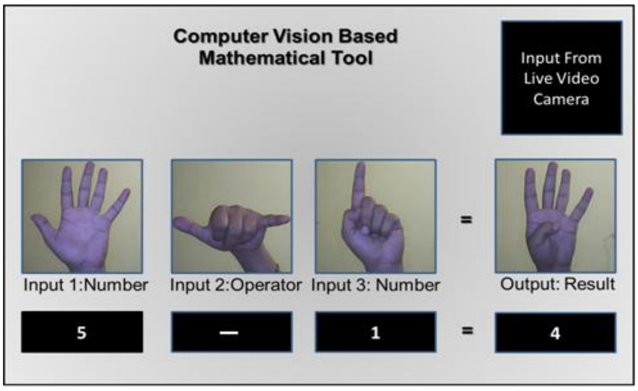


Figure 4.HCI application: Gesture Based Calculator

The graphical user interface (GUI) of the same is shown in Figure 4. The system takes three inputs from user.

Input 1, Input3: Numbers

It is a single handed gesture of 0 to 9 numbers made by user in front of the webcam. Camera captures the image and passes it to our proposed algorithm for further recognition. Figure 5 shows typical single hand gestures to count the numbers from 0 to 5.Similer gesture is made by remaining hand to count the numbers from 5 to 10.

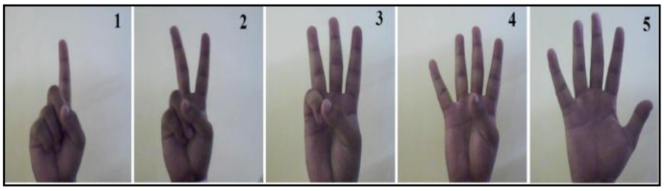


Figure 5: Single Hand Gestures as input numbers

Input2: Operators

It is a two handed gesture of arithmetic operations (plus, minus, multiplication, division) as shown in figure 11.to recognize these gestures we used Gesture Recognition Algorithm Based on Wavelet transforms and Principal Component Analysis proposed by [1].

Figure 6: Dual Handed Gesture of Mathematical operation

Output:

The recognized outcomes of input gestures are passed to 'domath()' function. It does the arithmetic operations depending upon input arguments.

```
domath (input1, operator, input2)
{
if operator==plus;
output=input1+input2;
if operator==minus;
output=|input1   input2|;
if operator==multiply;
output=input1   input2;
if operator==divide;
output=input1   input2;
}
```

Here input 1 and input 2 are the 0 to 9 numbers and operator is either +, -, x or / recognized by HGR algorithm. After evaluating the output, it is displayed in the form of the number as well as corresponding gesture image.

IV. EXPERIMENTAL RESULTS

The system performance is validated in still images and can be applied to real video sequences. In average the system recognized static gestures in cluttered background with great accuracy and less computational time. Recognition is invariant to hand position, orientation and distance from web cam.

V. CONCLUSION

In this study a fast and simple algorithm for hand gesture recognition is proposed. Algorithm segments the hand region and then recognizes the input gesture. For gesture recognition centroid distance features are used and a high recognition rate can be achieved with improved computational time. Also this paper presented a Gesture Based Calculator system able to interpret dynamic and static gestures from a user with the goal of real-time human-computer interaction.

The system only uses 2D gesture paths for dynamic gestures, although as future work it is our intention to test and include not only the possibility of 3D dynamic gestures but also to work with several cameras to thereby obtain a full 3D environment and achieve view-independent recognition.

REFERENCES

[1] V.G. Spitsyn and N.H. Phan,"Face and Hand Gesture Recognition Algorithm Based on Wavelet transforms and Principal Component Analysis," *FifthIEEE International Conference on Computational Intelligence, Communication Systems and Networks , Aug.* 2013

[2] Prashanth Suresh and NirajVasudevan, " Computer-aided Interpreter for Hearing and Speech Impaired," *Fourth International Conference on Computational Intelligence, Communication Systems and Networks,* June2012.

[3] Shweta. K. Yewale, Pankaj. K. Bharne, "Hybrid Algorithm for Hand Gesture Recognition," in *Proc International Conference on Computer & Information Science (ICCIS),* Aug.2012.

[4] Rohit Kumar Gupta, “A Comparative Analysis of Segmentation Algorithms for Hand Gesture Recognition,” in *ProcThird International Conference on Computational Intelligence*, *Communication Systems and Networks*, June 2011.

[5] S.D. Sarwarkar, A.D.Gawande, “Hand Gesture Recognition using CAMSHIFT Algorithm” in *ProcThird International Conference on Emerging Trends in Engineering and Technology,IEEE* 2010.

[6] Husheng Li and Zhu Han, “Hand Gesture Recognition System Using Standard Fuzzy C-Means Algorithm for Recognizing Hand Gesture with Angle Variations for Unsupervised Users” in *Proc.IEEE Transactions on Wireless Communications*, vol. 9, no. 11, November 2010.

[7] Hang Zhou, QiuqiRuan, “A Real-time Gesture Recognition Algorithm on Video Surveillance,”in *ProcInternational Conference on Signal Processing, June 2006.*

[8] BurakOzer, Tiehan Lu, and Wayne Wolf, “Design of a Real-Time Gesture Recognition System,” in *ProcIEEE Signal Processing Magazine [57]* **,** *May 2005.*

[9] Paulraj M.P, Yaacob, S., Desa, H, "Extraction of head and hand gesture features for recognition of sign language," *Electronic Design, 2008. ICED 2008. International Conference on* , vol., no., pp.1-6, 1-3 Dec. 2008.

[10] Zhu, Hong-Min, Pun, Chi-Mun, "Movement Tracking in Real-Time Hand Gesture Recognition" *Information Technology: Computer and Information Science 2010. IEEE/ACIS 2010. International Conference*, Aug 2010.

[11] Soontranon, N, Aramvith, S.Chalidabhongse, T.H, "Improved face and hand tracking for sign language recognition," *Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference* on , vol.2, no., pp. 141- 146 Vol. 2, 4-6 April 2005.

[12] Baoyun Zhang, Ruwei Yun Digital Entertainment Research Center, Nanjing Normal University, Nanjing, China “Gesture Recognition System Based on Distance Distribution Feature and Skin-Color Segmentation”, 978-1-4244-5858-5/10 2010 IEEE.

[13] M. M. Hasan& P. K. Mishra, (2011) “Performance Evaluation of Modified Segmentation on Multi Block For Gesture Recognition System”, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 4, No. 4, pp 17-28.