ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

A Hand Gesture Recognition for Human Computer Interaction

Anuj Kumar[1], Prof.N.M. Wagdarikar[2]

PG Student, Dept. of E&TC, Smt. Kashibai Navle College of Engineering, Pune, India

Assistant Professor, Dept. of E&TC, Smt. Kashibai Navle College of Engineering, Pune, India

ABSTRACT: This paper presents a novel and real time system with an application of sign language recognition via hand gesture recognition. This system includes detecting & tracking of bare a hand in uniform & cluttered background using skin colour detection after RGB image is converted into YCbCr image. In training stage, after extracting the features for every training image using moment invariant method and feature vector is created. This feature vector is taken as input for multiclass artificial neural network classifier. In testing stage, a webcam is used to capture the image and out of seven, first four feature values are used for classification and recognition of hand gestures.

KEYWORDS: sign language recognition; RGB; YCbCr; training; testing, artificial neural network.

I. Introduction

Human and Computer Interface has become a major portion of our lives because of technology potion into our current way of life. This paper describes the design and execution of a vision-based hand gesture recognition (VHGC) for human computer interaction (HCI) applications and for man optional and augmentative message applications. Hand gestures and even whole body play an important role in human communication. Computer analysis of hand gesture has become an important area in hand gesture recognition.

Gesture recognition is a challenging task, especially hand gesture recognition. Hand Gesture has very high degree of freedom and can move freely in any direction. The shape of the hand is also complicated.

Gesture recognition mostly involves two approach. One is static hand gesture recognition [5] and another is to analyse dynamic hand movements. In this method hand movement is tracked and recognized in real time [1] [4].

Our system is different from the earlier system from the following ways.

- A high accuracy with different hand gestures.
- A low processing power is required because of highly efficient computational operation of artificial neural network.
- Light variation have very trivial effect on our system.
- Real time operation.

This paper presents vision based hand gesture recognition for sign language recognition. The advantage of vision based system is one would not has to use glove for hand segmentation process. Hand segmentation is based on skin colour detection [5] [7]. RGB image is converted into YCbCr image. Y component implies luminance of the image but Cb and Cr imply chrominance of the image.

II. Related works

There are normally two types of category for vision based hand gesture recognition. One is three dimensional based hand prototypical and another is two dimensional based mode [2] [3] or it can be said appearance based. The three-dimensional hand gesture recognition technique is captured by Kinect camera which also calculates the depth of the image.Appearance-based model technique extract the features of the image and relate those features with the extracted features computed from the video frames. This is our approach because of its effectiveness.

In [1], Feature extraction through Moment invariant is used for object recognition. While taking image, the gesture variations consists of rotation, scaling and translation and these variations can be avoided using set of features that are

ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

invariant to these operations. Moment invariant offers a set of features whose values are independent of these degree of freedom.

In [2], Scale Invariance Feature Transform (SIFT) is used to take out the key points of every training image. This feature is independent of the scaling of the image. The accuracy can be achieved up to 92.67%, as presented by the author using Support Vector Machine (SVM) classifier.

The authors of [3] obtained a contour of the hand to recognize the hand gestures. Contour of hand is obtained after segmentation of the hand images. Then the contour of the hand is represented by linear and non-linear localized contour sequence. Gesture likeness is determined by measuring the similarity between the localized contour sequences.

In [5], gesture analysis method was developed using a skin colour based method to extract the area representing the hand. Hole and the finger direction is taken as feature for classification and Hidden Markov Model (HMM) is used as a classifier

The author of [7] proposed pixel-based hierarchical-feature Ada Boosting (PBHFA), skin pixel based colour segmentation and cluttered background subtraction. This feature meaningfully diminishes the training time by a factor of at least thousand times compared to the conventional Haar-like feature. In this paper ,authors has tried to minimize the tracking time of hand gesture and training time and training image of hand gesture.

III. PROPOSED ALGORITHM

The system comprises a camera and gesture pre-processing unit. The webcam is used to capture the image with uniform and cluttered background which are then segmented using skin-colour, normalized using morphological operation and median filter and featured extracted using moment invariant for classification and sign language recognition. Matlab2015a is used during the course of project. The advantage of using Matlab is real-time data processing, classification and recognition. The artificial neural network is used as a classifier which constructs training stage and used in testing stage to recognize hand gesture which is captured from webcam. Below Fig. 1 summarizes the whole project.

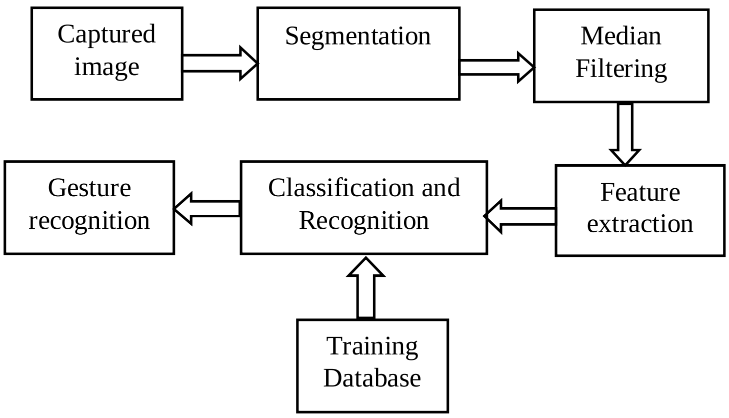


Fig. 1. Block diagram of hand gesture recognition for sign language recognition

A. Gesture Registration

The captured hand gesture needed to be processed before they can be processed by the computer because it contains unwanted component in background. In this project only hand image is taken without face.

B. Skin colour segmentation

The 'skin' pixel colour segmentation used in this project which comprises converting RGB image into YCbCr image format [1]. Then threshold value of Cb and Cr is applied to remove non-skin colour elements. The advantage of thismethodology is brightness can be removed during the transformation process of RGB into YCbCr format. So brightness does not matter after capturing images. So light compensation can be achieved through this approach.

Through 'imtool, command in matlab Y, Cb and Cr values of image can be seen at any particular coordinates. Y (luminosity) values is neglected. Chrominance component i.e. Cb and Cr values can be set by observing the values of Cb and Cr using 'imtool' command.

ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

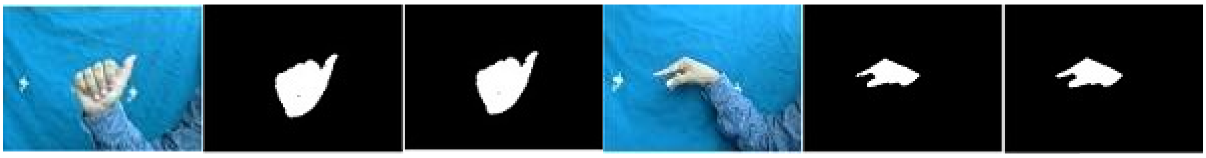
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

C. Median Filtering

After thresholding the YCbCr image, there are some noisy spot can be seen in the figure 2. These noisy spot affect the feature values more than expected deviation. It also affect recognition accuracy of the system at the subsequent stages.So these distortion can be through 'Median' filtering. In this research 5x5 identity matrix is used to remove the noisy spot.



Original Image Segmented Image Filtered Image Original Image Segmented Image Filtered Image

Fig. 2. Median filter Image output

The advantage of using Median Filter is 'skin-colour' region can be extracted from the captured image more accurately. So in conclusion, it can be said that skin colour segmentation and median filter is steadfastly robust against in uniform background and lighting conditions.

D. Feature Extraction

Hu's Movement Invariant algorithm is used as a feature extraction because of its effectiveness and usefulness in body identification. The algorithm has been broadly used in classification of airplanes, vessels, gesture classification and leaf recognition etc.

Hu's moment is defined seven functions, calculated from central moments of orders two, which were invariant with respect to translation, rotation and scaling of object. Hu's moment invariants is calculated as below-

$$\Omega_1 = \eta_{20} + \eta_{02} \quad (1)$$

$$\Omega_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}{}^2 \quad (2)$$

$$\Omega_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (3)$$

$$\Omega_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (4)$$

$$\Omega_5 = (\eta_{30} - 3\eta_{12})\,(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})$$
$$[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (5)$$

$$\Omega_6 = (\eta_{20} + \eta_{02})\,[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] +$$
$$4\eta_{11}(\eta_{30} + \eta_{12})^2(\eta_{21} + \eta_{03}) \quad (6)$$

$$\Omega_7 = (3\eta_{21} - \eta_{03})\,(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2\text{-}3(\eta_{21} + \eta_{03})^2](3\eta_{12} - \eta_{30})\,(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2\text{-}(\eta_{21} + \eta_{03})^2 (7)$$

Where η_{ab} is the stabilized central moment of order (a + b), which is defined as

$$\eta_{ab} = \frac{\varphi_{ab}}{\varphi_{00}{}^{\gamma}},$$

$$\Psi = \frac{a+b}{2} + 1$$

$\varphi_{ab}$is the parallel central moment, defined as

$$\varphi_{ab} = \sum_x \sum_y (p - \overline{q})^a (q - \overline{q})^b f(p, q),$$

$$\overline{p} = \frac{m_{10}}{m_{00}} \text{ And } \overline{q} = \frac{m_{01}}{m_{00}}.$$

m_{ab} is two dimensional moment of order (a + b), well-defined as

ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

International Journal of Innovative Research in Computer and Communication Engineering

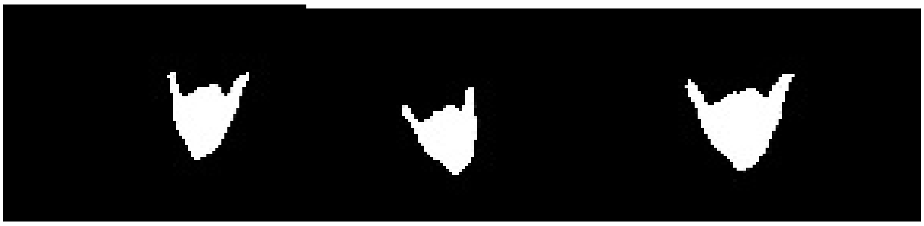
(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

$$m_{ab} = \sum_p \sum_q p^a q^b f(p,q).$$

Ω_1 - Ω_6 are moment invariants with respect to revolution and reflection, while Ω_7 fluctuates sign under reflection.

Example of Moment invariants



Normal Image Rotated Image Scaled Image

Fig. 3. Normal, Rotated and Scaled image of a hand gesture

Figure 3 shows three hand images out of which first image captured at normal distance and straight in line. Second image is captured while hand is slightly tilted and distant from the webcam and third image is captured nearer to the webcam.

From Table I it can be seen that first four moment value i.e. $\Omega_1, \Omega_2, \Omega_3$ and Ω_4 is being changed slightly approx. 2% variations from normal image while next three can be discarded for object recognition. So as a result feature extraction using moment invariant can be progressed rapidly and efficiently.

TABLE I: Moment Invariant calculated using formula (1)-(7)

f_i	Normal Image	Rotated Image	Scaled Image
f_1	0.2000	.2006	0.2006
f_2	5.739e-04	6.047e-04	2.835e-04
f_3	.00300	.0036000	.0043000
f_4	1.549e-04	1.917e-04	1.889e-04
f_5	-3.31e-08	-1.50e-08	3.947e-08
f_6	3.204e-06	4.826e-06	1.461e-06
f_7	-1.02e-07	1.113e-07	-2.64e-08

E. Gesture Classification

The Gesture classification in this project is done via the artificial neural network.For the above method training set is required which contain a number of sample images. In this project approx. 100 images is taken for training set for each alphabet recognition. In feature pulling out phase, each set of taster images that represent the same gesture produces approximately similar range of $\Omega_1, \Omega_2, \Omega_3$ and Ω_4.

These ranges of deviations are used as a preset values to classify and recognize hand gesture. According to variations of background as feature changes, 2% of deviations is used for feature classification.

The Fig.4. Shows the classification accuracy for the each hand gesture. The early artificial neural network (ANN) is trained with 9000 of images representing 26 hand images for each alphabet and each image is taken under uniform background. Each gesture approximately has 400 different images.

Input is single feature vector which is given to ANN. For example, detecting 'A' feature vector is [1 0 0 0 0 0 0 0 0 0], for detecting 'B' feature vector is [0 1 0 0 0 0 0 0 0 0] and so on.

ISSN(Online): 2320 - 9801
ISSN (Print) : 2320 - 9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 6, June 2016

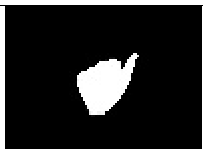
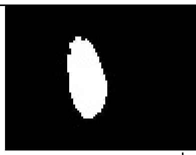
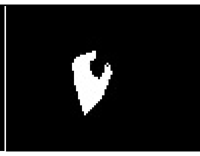
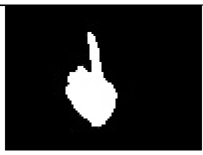
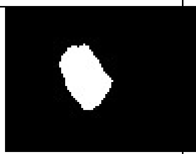
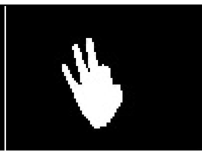
Hand Gesture			
Classification accuracy	95%	85%	83%
Hand Gesture			
Classification accuracy	90%	92%	82%

Fig. 4. Classification Accuracy

IV. CONCLUSION

The main theme of this system is to improve robust hand gesture recognition with unintended and irregular hand gestures such as children's unintended movement of hand. The recognition of hand gesture is independent of the size of the palm. Rotation and Scaled gesture can be recognized reliably using this system. In this work a set of different gestures of Indian Sign Language database which is distinct from each other is created. Some of the gestures have been taken and classified successfully with good accuracy rate up to 90%. Artificial Neural Network Classifier reduces the computational work up to great extent. So Power and Time consumption can be reduced. In future all the hand gesture of Indian sign language can be recognized with more accuracy by using more efficient feature extraction such as aspect ratio, distance between centroid of the palm and fingertips, Scale Invariant Feature Transform (SIFT) etc.

REFERENCES

1. Q. Nguyen and P. Premaratne, "Consumer electronics control system based on hand gesture moment invariants," IET computer vision, , April 2007, pp. 35-41.
2. Nicolas D. Georganas and Nasser H. Dardas, "Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques," IEEE Transaction on Instrumentation and Measurement, vol. 60, no. 11, Nov 2011, pp 3592-3607.
3. Suwei Ma and Lalit Gupta, "Gesture based interaction and Communication: Automated Classifier of Hand Gesture Contours," IEEE Tran. On system, mans and cybernetics, Vol 31, no. 1, Feb 2001, pp. 114-120.
4. T.Y. Kim and C. Nyirarugira, "Adaptive Differential Evolution Algorithm for Real Time Object Tracking," IEEE Transactions on Consumer Electronics, Vol. 59, No. 4, November 2013, pp. 833-838.
5. Lin Xueyin and Ma Gengyu, "Operational Gesture Segmentation and Recognition," Xsinghua science and technology," Volume 8, Number 2, April 2003, pp. l69 – 173.
6. Yun-Fu Liu and Jing-Ming Guo, "Improved Hand Tracking System," IEEE Transaction on circuits and systems for video technology," vol. 22, no. 5, May 2012, pp. 693-701.
7. Tom Chau and Natasha Alves, "Vision-Based Segmentation of Continuous Mechanomyographic Grasping Sequences," IEEE Transaction on biomedical engineering, vol. 55, no. 2, Feb. 2008, pp. 765-773.
8. A. El-Sawah, N. Georganas, and E. Petriu, "A prototype for 3-D hand tracking and gesture estimation," IEEE Trans. Instrum. Meas., vol. 57, no. 8, pp. 1627–1636, Aug. 2008.
9. A. Argyros and M. Lourakis, "Vision-based interpretation of hand gestures for remote control of a computer mouse," in Proc. Workshop Compute. Human Interact. pp. 40–51, 2006.
10. Q. Chen, N. Georganas, and E. Petriu, "Real-time vision-based hand gesture recognition using Haar-like features," in Proc. IEEE IMTC, pp. 1–6, 2007.
11. P. Viola and M. Jones, "Robust real-time object detection," Int. Jour. Compute. Vis., vol. 2, no. 57, pp. 137–154, 2004.
12. D. G. Lowe, "Distinctive image features from scale-invariant key points,"Int. J. Compute. Vis., vol. 60, no. 2, pp. 91–110, Nov. 2004.