

Hand Gesture Recognition Using PCA

Mandeep Kaur Ahuja

*CE Department
Punjabi University Main Campus,
Patiala, INDIA*

Dr. Amardeep Singh

*Professor, CE Department
Punjabi University Main Campus,
Patiala,INDIA*

Abstract: Interacting with physical world using expressive body movements is much easier and effective than just speaking. Gesture recognition turns up to be important field in the recent years. Communication through gestures has been used since early ages not only by physically challenged persons but nowadays for many other applications. As most predominantly hand is used to perform gestures, Hand Gesture Recognition have been widely accepted for numerous applications such as human computer interactions, robotics, sign language recognition, etc. This paper focuses on bare hand gesture recognition system by proposing a scheme using a database-driven hand gesture recognition based upon skin color model approach and thresholding approach along with an effective template matching which can be effectively used for human robotics applications and similar other applications.. Initially, hand region is segmented by applying skin color model in YCbCr color space. In the next stage otsu thresholding is applied to separate foreground and background. Finally, template based matching technique is developed using Principal Component Analysis (PCA) for recognition. The system is tested with the controlled and uncontrolled database and shows 100% accuracy with controlled database and 91.43% with low brightness images.

Keywords: hand, gesture, recognition, segmentation.

I. INTRODUCTION

Gestures are the movement of any body part used to convey the meaningful information. Communication through gestures has been widely used by humans to express their thoughts and feelings. Gestures recognition refers to the process of identifying gestures performed by human so that machine can perform the corresponding action. Gestures have been classified in two categories static and dynamic[16]. Static gestures refer to still body posture and dynamic refers to movement of body part. Gestures can be performed with any body part like head, face, arms, hands, etc. but most predominately we use hand to perform gesture like we wave hand to say 'good bye'. Hand gestures have been widely used for many applications[like human – computer interaction (HCI), robotics, sign language[19], human machine interaction, TV interaction[18] etc. With the advancement of technology, human robot interaction (HRI) has become an emerging field in recent years. Hand gestures can be effectively used to give commands to the robot which in turn can be employed in large number of applications. Now-a-days, human robot interaction using hand gestures has widely been used in medical sciences[15]. But still challenges regarding robustness and efficiency are to be considered.

Hand Gestures Recognition techniques have been divided into two categories[13]- Sensor based and Vision Based recognition.

Sensor based recognition collects the gesture data by using one or more different types of sensors. These sensors are attached to hand which record to get the position of the hand and then collected data is analyzed for gesture recognition. Data glove[1][2] is an example of sensor based gesture recognition. Other sensors used were Wii controller, EMG sensors, accelerometer sensors[26], etc. Sensor based recognition has certain limitations. First of all it requires a proper hardware setup which is very expensive. Secondly, it hinders the natural movement of the hand. So to overcome the limitation of sensor based recognition vision based techniques came into existence.

Vision based techniques [1]uses one or cameras to capture the hand images. Various type of cameras used for capturing image can be stereo cameras, monocular cameras, fish eye cameras, time-of-flight cameras, infrared cameras, etc. Vision based techniques uses various image processing algorithms to get hand posture and movement of hand. Some vision based techniques uses colored markers to get the position of hand. But the vision based recognition also has some limitations that it is affected by illumination changes and cluttered backgrounds. Vision based techniques are further divided into two categories[28] – 3D model based and Appearance base recognition.

Model based approaches uses 3D hand model to search kinematics parameters by comparing 2D projection of 3D hand image and input frame. 3D model are further divided into Volumetric 3D recognition [20] and skeleton 3D recognition[22]. Because of the complexity of 3D model it is not preferred. Appearance based techniques[24] are based on extracting features from the visual appearance of the image and compare it with already defined templates. Various features that can be extracted from the image can be shape based features[4] that can be geometric or non-geometric. Geometric features include- position of fingertips, location of palm, centroid[8], orientation[3], direction[7],etc. Non- geometric features includes color[9], silhouette and textures, contour, `edges, image moments, Fourier descriptors[10][23], Eigen vectors[21], etc. Some techniques uses skin color model[14] to extract skin colored pixels. Other techniques HOG features[5], SIFT features, etc. Appearance based technique is preferred over model based technique because of the complexity of the model based techniques.

The basic step of hand gesture recognition is to localize and segment the hand from the image. Various techniques are available for hand segmentation. The most popular and simplest technique is skin color model[7][14] which is used to get the skin pixels in the image but it has some

limitations that skin color of different person can vary and background image can also contain the skin pixels. Other techniques are thresholding which divides the image into two regions foreground and background based on color, depth[6][10][12], etc. Some researcher uses background subtraction[9][11] for segmenting the hand. In our approach keeping in view the limitation and simplicity of skin color model we will combine it with thresholding for hand segmentation. Skin color segmentation[25] can be applied on any color space- RGB, HSV, YCbCr, YUV, etc. Every color space has its own benefits. We will use YCbCr color space for skin color segmentation. For gesture recognition HMM[3][8], SVM[18][8][23], Nearest Neighbour classifier[10], neural network[17][14], PCA[21], finite state machine (FSM)[22] etc.

In our approach vision based hand gesture recognition technique is proposed using a database-driven approach based upon skin color model and thresholding along with an effective template matching using PCA which will be used for controlling robotics hand in surgical applications and many other similar applications.

The rest paper is organized as follows: Section II describes the methodology, in section III results are discussed and finally conclusion and future scope in discussed.

II. METHODOLOGY

In this section we will discuss our proposed methodology step by step.

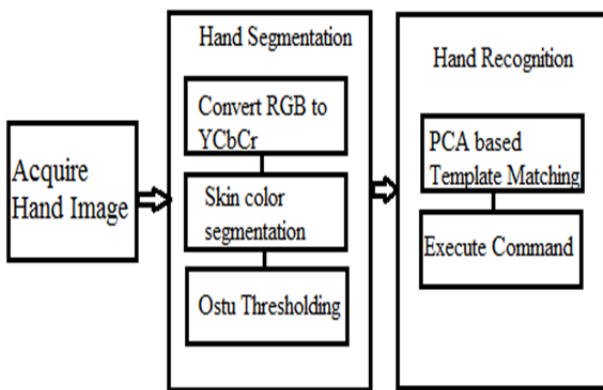


Fig. 1 Flow Diagram of proposed Methodology

A. *Image Acquisition*: Images are acquired using the 13 megapixel real-aperture camera in controlled background as well as by varying the lightning conditions.

B. *Hand Segmentation*: The main and basic step in hand gesture recognition is to segment the hand from the whole image so that it can be utilized for recognition. In our proposed color skin color segmentation is applied to segment the hand. As skin color of different person can vary and background image can also contain the skin pixels so after skin color model Otsu Thresholding is applied to remove the background.

- *Conversion from RGB to YCbCr*: The proposed skin color segmentation in applied to YCbCr color space. So first of all RGB color space is converted to YCbCr

color space. Y represents the luminance and Cb and Cr represents chrominance. The RGB color space is converted to YCbCr color space using the following equation:

$$Y = 0.299R + 0.587G + 0.114B$$

$$Cb = (B - Y) * 0.564 + 128$$

$$Cr = (R - Y) * 0.713 + 128$$

- *Skin Color Segmentation*: The skin color segmentation is used to classify the pixel as skin pixel or non-skin pixel. As or hand is connected component made of skin pixels we will get the hand after skin color segmentation. Steps for skin color segmentation:

1. The first step in skin color segmentation to specify the range for the skin pixels in YCbCr color space.

$$[R_{Cb}, R_{Cb}'] = [77, 127] \ \&$$

$$[R_{Cr}, R_{Cr}'] = [133, 173]$$

R_{Cb}, R_{Cb}' is lower and upper bound for Cb component.

R_{Cr}, R_{Cr}' is lower and upper bound for Cr component

2. Find the pixels (p) that are in the range defined above:

$$R_{Cb} \leq \text{Pixel value}(Cb(i, j)) \leq R_{Cb}'$$

$$R_{Cr} \leq \text{Pixel value}(Cr(i, j)) \leq R_{Cr}'$$

3. Summation of all the pixels in the above step belongs to Region of interest i.e hand.

$$ROI = \sum p(i, j)$$

After Skin color segmentation we will the hand but may be some other pixels in the background also. To remove that background pixels we will use Otsu Thresholding.

- *Otsu Thresholding*: Thresholding is used to separate the object from its background by assigning pixel to either background or foreground based on threshold value. In our proposed system hand is in foreground. Otsu threshold is a global thresholding method which chooses threshold that minimizes within class variance.

1. Calculating threshold value:

In MATLAB there is a function Graythresh(I) which calculate global threshold value using Otsu Threshold.

$$TH = \text{graythresh}(I)$$

2. Convert Image pixel values into binary value according to THR.

Then

$$g(i, j) = \begin{cases} 1 & \text{if } p(i, j) \geq T \\ 0, & \text{otherwise} \end{cases}$$

C. *Gesture Recognition*

One of the important technique of recognition is template matching in which a template to recognize is available and is compared with already stored template. In our approach PCA method for feature extraction and matching is used.

Principal Component Analysis: PCA is used to reduce the dimensionality of the image while preserving much of the information. It is the powerful tool for analyzing the data by identifying patterns in the dataset and reduces the

dimensions of the dataset such that maximum variance in the original data is visible in reduced data. PCA was invented by Karl Pearson in 1901. It works by converting set of correlated variables to linearly uncorrelated variable called principal components. Principal components are calculated by computing Eigen vectors of covariance matrix obtained from the group of hand images. The highest M eigenvectors contains the maximum variance in the original data. These principal components are orthogonal to each other and the first component is in the direction of greatest variance.

Mathematical Model for PCA:

The PCA approach has 2 stages: Training and Testing stage. In the training stage the Eigen space in established using training images of hand gestures and these images are mapped to the Eigen space. In the testing stage the image to be tested is mapped to same Eigen space and is classified using distance classifier.

Algo. for PCA:

Training Stage: Calculation of Eigen vectors:

1. Obtain the database containing N training images of dimensions $M \times M$: $I_1, I_2, I_3 \dots \dots \dots, I_N$.
2. Convert these N images into vectors Z_i , $1 \leq i \leq N$ of dimension M^2 .

3. Obtain mean image vector Ψ Ψ

$$= \frac{1}{N} \sum_{i=1}^N Z_i$$

4. Obtain the difference image by subtracting the mean image vector from the training image.

$$\phi_i = Z_i - \Psi$$

5. Obtain the covariance Matrix C having dimensions $M^2 \times M^2$.

$$C = \frac{1}{N} \sum_{n=1}^N \phi_n \phi_n^T = AA^T$$

$A = [\phi_1, \phi_2, \phi_3, \dots \dots \dots \phi_n]$ with dimension $M^2 \times N$.

6. Compute the Eigen vectors v_i of AA^T

As the dimensions of $AA^T [M^2 \times M^2]$ are very large so computation of eigenvectors is impractical.

7. Obtain Eigen vectors u_i of $A^T A$ [dimensions $N \times N$].

AA^T has M^2 Eigen vectors and Eigen values.

$A^T A$ has N Eigen vectors and Eigen values.

8. Obtain the best N eigenvectors of AA^T by following equation.

$$v_i = Au_i$$

9. Take only V Eigen vectors corresponding to V largest Eigen values.

Representation of Training Database using Eigen Vectors:

Weight of each training image is calculated as:

$$w_j = v_j^T \cdot (Z_i - \Psi),$$

where $j=1,2,3,\dots,N$.

Weight vector is represented as:

$$\mu = [w_1, w_2, w_3, \dots \dots \dots, w_N]^T;$$

Every image in training database is represented by weight vector:

$$\mu_i = [w_1^i, w_2^i, w_3^i \dots \dots \dots, w_N^i]^T;$$

Testing Stage:

Let the

image to be tested is q , its weight (w_i) is calculated by multiplying Eigen vector v_i with the difference image.

$$w_i = v_i^T \cdot (q - \Psi)$$

Weight vector of unknown image is calculated as:

$$\mu = [w_1, w_2, w_3, \dots \dots \dots, w_N]^T;$$

Compute:

$$E_j = \min || \mu - \mu_j ||; j = 1, 2, 3 \dots N$$

N is no. of training images.

So, q is recognized as jth hand gesture from training database.

III. EXPERIMENTAL RESULTS

The following systems were implemented using MATLAB 2013a and tested on an Intel Core i3 with 4GB of RAM running Windows 7. The datasets has been obtained in both controlled and non-controlled forms.

Controlled Dataset: The controlled dataset has been obtained on the similar background which can provide the maximum accuracy. The database collected in the ideal conditions has proved to be the most efficient database in terms of accuracy. The controlled database has been collected from the various objects (persons). A few images from the dataset are shown below:

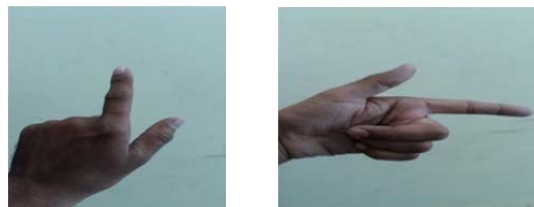


Figure 2: The controlled hand gestured obtained using the 13 megapixel real-aperture camera

The result obtained of testing the system is shown in the following tables.

Table 1: The accuracy tables of the controlled database, (a) The accuracy test on the controlled database, (b) The statistical analysis of the proposed model results

Condition of Image	Total tested	Successful detection	Failures	Accuracy
Controlled Background	30	30	0	100%

(a)

PROPERTY	VALUE
True Positive	30
True Negative	0
False Positive	0
False Negative	0
Sensitivity (Recall)	100%
Precision (Positive Predictive)	100%
Prevalence	100%

(b)

Royalty Free Image Hand Gesture Dataset: The database has also been obtained from the internet sources by collecting the royalty free images of hand gestures. This database has been used to test the flexibility and adaptability of the proposed model in the case of different situations than the controlled conditions. Uncontrolled database include images of low brightness and normal or below than normal light conditions.

Table 5.2: Table 5.1: The accuracy tables of the uncontrolled database, (a) The accuracy test on the uncontrolled database, (b) The statistical analysis of the proposed model results

Condition of Image	Total tested	Successful detection	Failures	Accuracy
Low Brightness images	35	32	3	91.43 %

(a)

PROPERTY	VALUE
True Positive	32
True Negative	0
False Positive	2
False Negative	1
Sensitivity (Recall)	96.97%
Precision	94.12%
Prevalence	94.29%

(b)

Formulas for calculating Recall, Precision and F-measure

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Prevalence} = \frac{\text{True positive} + \text{False Negative}}{\text{Total number of Gestures}}$$

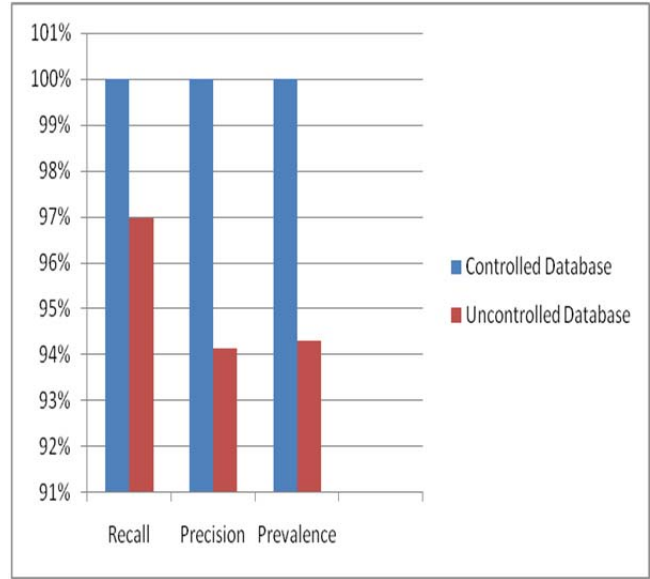


Figure 3: The comparison table of recall, precision and F-measure

IV. CONCLUSIONS

In this paper the hand gesture recognition system is developed using skin color model, Otsu thresholding and PCA. The system is tested in controlled background and in different lightning conditions. The database collected in the ideal conditions has proved to be the most efficient database in terms of accuracy and gives 100% accuracy and when the lightning conditions are changed the accuracy decreases as compare to the previous one. The system shows 91.43% with low brightness images .The hand images have been obtained for the purpose of human-computer interactions for the operation theatre robots, which must understand the hand language in order to take the actions. Our research empowers the medical experts to pass the instruction to the robotic hands remotely to add the accuracy in the operations. But the proposed model is not capable of working with the images containing hands of other than skin color. The proposed model does not evaluate the images clicked in other light colors where the hand gestures has been clicked and the model work only with static gesture .In future the system can be upgraded to support dynamic gestures and an application for controlling medical operations can be developed using the system.

REFERENCES

- [1] Haitham Hasan , Sameem Abdul-Kareem, "Human-computer interaction using vision-based hand gesture recognition systems: a survey", *Neural Comput & Applic*,2013.
- [2] Fahn, C.S,Sun, H, "Development of a fingertip glove equipped with magnetic tracking sensors", *Sensors 2010*,vol. 10, pp. 1119-1140, 2010
- [3] Chang-Yi Kao,Chin-Shyurng Fahn, "A Human-Machine Interaction Technique: Hand Gesture Recognition Based on Hidden Markov Models with Trajectory of Hand Motion", *Procedia Engineering 15*, pp. 3739 - 3743,2011.
- [4] Amit Gupta, Vijay Kumar Sehrawat, Mamta Khosla, "FPGA Based Real Time Human Hand Gesture Recognition System", *Procedia Technology 6* ,pp. 98 - 107, 2012
- [5] Jing Lin, Yingchun Ding, "A temporal hand gesture recognition system based on hog and motion trajectory", *Optik 124* pp.6795-6798, 2013.

- [6] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust part-based hand gesture recognition using Kinect sensor," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1110–1120, Aug. 2013.
- [7] Jos'e Manuel Palacios, Carlos Sagues, Eduardo Montijano and Sergio Llorente, "Human-Computer Interaction Based on Hand Gestures Using RGB-D Sensors", *Sensors 2013*, vol.13, pp. 11842-11860, 2013.
- [8] Paulo Trigueiros, Fernando Ribeiro, Luis Paulo Reis "Generic System for Human-Computer Gesture Interaction", *IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), Espinho, Portugal*, pp. 175-180, May 2015.
- [9] Chetan Dhule, Trupti Nagrare, "Computer Vision Based Human-Computer Interaction Using Color Detection Techniques", *Fourth International Conference on Communication Systems and Network Technologies*, pp.934-938, 2014.
- [10] Stergios Poularakis and Ioannis Katsavounidis, "Finger Detection And Hand Posture Recognition Based On Depth Information", *IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, pp. 4329-4333, 2014.
- [11] Dong-Luong Dinha, Jeong Tai Kimb, and Tae-Seong Kimc, "Hand Gesture Recognition and Interface via a Depth Imaging Sensor for Smart Home Appliances", *6th International Conference on Sustainability in Energy and Buildings, SEB-14*, pp. 576 – 582, 2014.
- [12] Chong Wang, Zhong Liu, and Shing-Chow Chan, "Superpixel-Based Hand Gesture Recognition With Kinect Depth Camera", *IEEE Transactions On Multimedia*, vol. 17, no.1, pp. 29-39, January 2015.
- [13] Arpita Ray Sarkar, G. Sanyal, and S. Majumder, "Hand Gesture Recognition Systems: A Survey", *International Journal of Computer Applications*, vol. 71, no.15, pp. 0975 – 8887, May 2013.
- [14] E. Stergiopoulou, N. Papamarkos, "Hand gesture recognition using a neural network shape fitting technique", *Engineering Applications of Artificial Intelligence* 22, pp. 1141–1158, 2009.
- [15] Rong Wena, Wei-Liang Tayb, Binh P. Nguyena, Chin-Boon Chnga, Chee-Kong Chui, "Hand gesture guided robot-assisted surgery based on a direct augmented reality interface", *Computer methods and programs in biomedicine* 116, pp. 68–80, 2014.
- [16] Sushmita Mitra, and Tinku Acharya, "Gesture Recognition: A Survey", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 37, no. 3, pp. 311-324, May 2007.
- [17] Yoichi Sato, Makiko Saito and Hideki Koike, "Real-Time Input of 3D Pose and Gestures of a User's Hand and Its Applications for HCI", *Proceedings of the Virtual Reality Conference*, 2001
- [18] Sang-Heon Lee, Myoung-Kyu Sohn, Dong-Ju Kim, Byungmin Kim, and Hyunduk Kim, "Smart TV Interaction System Using Face and Hand Gesture Recognition", *IEEE International Conference on Consumer Electronics*, pp. 173 -174, 2013.
- [19] Archana S. Ghotkar and Dr. Gajanan K. Kharatem, "Study of Vision Based Hand Gesture Recognition Using Indian Sign Language", *International Journal On Smart Sensing And Intelligent Systems*, vol. 7, no. 1, pp. 96-115, March 2014.
- [20] B. Stenger, P. R. S. Mendonça, and R. Cipolla, "Model-based 3D tracking of an articulated hand", *Proc. CVPR*, Kauai, HI, USA, pp. 310–315, 2001
- [21] Krishnakant C. Mule & Anilkumar N. Holambe, "Hand Gesture Recognition Using PCA and Histogram Projection", *International Journal on Advanced Computer Theory and Engineering (IJACTE)* 2319 – 2526, Volume-2, Issue-2, 2013
- [22] X. Bai and L. J. Latecki, "Path similarity skeleton graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 7, pp. 1282–1292, Jul. 2008.
- [23] Kuan-Yu Chen, Cheng-Chin Chien, Wen-Lung Chang, Jyh-Tong Teng, "An Integrated Color and Hand Gesture Recognition Approach for an Autonomous Mobile Robot", *3rd International Congress on Image and Signal Processing*, pp. 2496-2500, 2010
- [24] G. Simion, V. Gui, and M. Oteteanu, "Vision Based Hand Gesture Recognition: A Review", *International Journal Of Circuits, Systems And Signal Processing*, Issue 4, Volume 6, 275-282, 2012.
- [25] Zaher Hamid Al-Tairi, Rahmita Wirza Rahmat, M. Iqbal Saripan, and Puteri Suhaiza Sulaiman, "Skin Segmentation Using YUV and RGB Color Spaces", *J Inf Process Syst*, Vol.10, No.2, pp.283–299, June 2014.
- [26] Xu Zhang, Xiang Chen, Yun Li, Vuokko Lantz, Kongqiao Wang, and Jihai Yang, "A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors", *IEEE Transactions On Systems, Man, And Cybernetics—Part A: Systems And Humans*, Vol. 41, No. 6, 1064-1076, November 2011