# Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction

**Archana S. Ghotkar** archana.ghotkar@gmail.com
*Department of Computer Engineering*
*Pune Institute of Computer Technology*
*University of Pune*
*Pune 411 043, India*

**Gajanan K. Kharate** gkkharate@yahoo.co.in
*Department of Electronics and Telecommunication Engineering*
*Matoshri College of Engineering and **R**esearch **C**entre*
*University of Pune*
*Nasik 422 105, India*

**Abstract**

This work is the part of vision based hand gesture recognition system for Natural Human Computer Interface. Hand tracking and segmentation are the primary steps for any hand gesture recognition system. The aim of this paper is to develop robust and efficient hand segmentation algorithm where three algorithms for hand segmentation using different color spaces with required morphological processing have were utilized. Hand tracking and segmentation algorithm (HTS) is found to be most efficient to handle the challenges of vision based system such as skin color detection, complex background removal and variable lighting condition. Noise may contain, sometime, in the segmented image due to dynamic background. An edge traversal algorithm was developed and applied on the segmented hand contour for removal of unwanted background noise.

**Keywords:** Human Computer Interface, Hand tracking and Segmentation, Hand Gesture Recognition, Edge Traversal Algorithm.

## 1. INTRODUCTION

Natural Human Computer Interaction (HCI) is the demand of today's world. Survey and Sign language study shows that from various gesture communications modality**,** the hand gesture is the most easy and natural way of communication. Real-time vision-based hand gesture recognition is considered to be more and more feasible for Human-Computer Interaction with the help of latest advances in the field of computer vision and pattern recognition [1].

There are various applications using Hand Gesture Recognition. Zhao *et al*. [2] explained virtual reality system based on hand gesture. Gaussian distribution was used for building complexion model, YCbCr color space was used for segmentation purpose and Fourier descriptor as a feature vector**.** BP neural network was utilized for recognition and all these resulted in an improved recognition rate. High light and shadow segmentation results were, however, found to be not perfect. Guan and Zheng [3] introduced a novel approach to pointing gesture recognition based on binocular stereo vision, in which user needs to wear special clothes or markers and was found to be suitable for both left and right handed users. Freeman and Weissman [4] explained television control application by hand gesture. Here**,** the user uses only one gesture: the open hand, facing the camera for controlling the television. Sepehri *et al.* [5] proposed algorithms and applications for using hand as an interface device in virtual and physical spaces.

In Real-time Vision based Hand Gesture recognition system**,** hand tracking and segmentation are most important and challenging steps towards gesture recognition. Uncontrolled environment,

lighting condition, skin color detection, rapid hand motion and self-occlusions are the challenges need to be considered while capturing and tracking the hand gesture [6]. Various researchers are working on hand tracking and segmentation to make it robust to achieve natural interface with machine. Bao *et al.* [7] introduced a new robust algorithm called Tower method for hand tracking module where**,** skin color was considered for hand gesture tracking and recognition.

Alon *et al*. [8] introduced unified framework for gesture recognition and spatiotemporal gesture segmentation applied to American Sign Language (ASL). To recognize manual gestures in video, it is required to do both spatial and temporal gesture segmentations. Spatial gesture segmentation is the problem of determining where the gesturing hand is located in each video frame. Temporal gesture segmentation is the problem of determining when gesture starts and ends.

Various gray-level segmentation techniques**,** such as use of single threshold value, adaptive thresholding, P-tile method, edge pixel method, iterative method and use of fuzzy set are available for object segmentation. Stergiopoulou and Papamarkos [9] proposed hand segmentation by color segmentation using YCbCr color map. Approximation of the hand morphology was accomplished by SGONG (self-growing and self organized neural gas) Network. Algorithm was able to identify finger, palm center, hand center, hand slope as well as finger features for recognition purpose. Probability based classification method was used for gesture classification. Skin color detection and complex background are major challenges in hand gesture recognition. Binary linked objects are groups of pixels that share the same label due to their connectivity in a binary image. Burande *et al.* [10] implemented Blobs analysis technique for skin color detection under complex background. Kalman filtering, HMM and Graph matching algorithm were used for gesture recognition. Howe *et al.* [11] had introduced fusion of skin color and motion 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 motion segmentation was used to find the difference between the moving foreground object and the stationery background.

Hand gestures are spontaneous and powerful communication modality for HCI. Various traditional input devices are available for interaction with computer**,** such as keyboard, mouse, joystick as well as touch screen**;** yet these are not considered natural interface. Proposed system will consist of desktop/laptop interface, with the help of hand gesture where users need not to wear any data glove**,** along with the web camera for capturing the hand image. The primary step toward any hand gesture recognition (HGR) is hand tracking and segmentation. In the present work**,** three techniques for hand segmentation were explored. The objective of this work is to overcome the vision-based challenges**,** such as dynamic background removal, skin color detection for natural human computer interface and variable lighting condition.

The organization of the paper is as follows. In section 2, system model of hand gesture recognition for desktop/laptop handling operations is described. Color model used in different algorithms are discussed in section 3. In section 4, techniques of hand detection and segmentation with anticipated dataset are introduced. In section 5, experimented results are shown. The results of experimented techniques are compared and discussed in section 6. Conclusion and future work is discussed in section 7.

## 2. HAND GESTURE RECOGNITION SYSTEM MODEL

Let S be a hand gesture recognition system that recognizes hand gesture.

S = { I,G,M,F,O }

where, I is a set of input hand gestures;

G represents a set of single-handed anticipated static gestures;

M represents mouse operation such as left click, right click;

F represents feature vector for G;

O represents output with application interface (A);

$G = \{G_1, G_2, ..G_7\}$

$I = \{I_1, I_2 \ldots I_5\}$

$G \subseteq I$

M = {center position (x,y), movement of (x,y) position, left click, right click};

$G \subseteq M$

$F = \{f1, f2 \ldots f5\}$

$O = \{A_1, A_2 \ldots A_7\}$

**Success of the system will be depend upon when**

(i) $I_i = F_j$ where $I_i \in I$

$F_j \in F$ where $1 <= j <= 5$

**Failure of the system when**

(ii) For a gesture(I) no feature vector is found

$I_i \neq F_j$ where $I_i \in I$

$F_j \in F$ where $1 <= j <= 5$

(iii) For two different gesture (I) same feature vector found

$I_i = F_j$ and $I_k = F_j$ where $I_i, I_k \in I$

$F_j \in F$ where $1 <= j <= 5$ and

## 3. COLOR MODELS

The aim of the proposed project is to overcome the challenge of skin color detection for natural interface between user and machine. So to detect the skin color under dynamic background the study of various color models was done for pixel based skin detection [12][13][14]. In this paper, three color spaces has been chosen which are commonly used in computer vision applications.

- **RGB:** Three primary colors red(R), green(G), and blue(B) are used. The main advantage of this color space is simplicity. However, it is not perceptually uniform. It does not separate luminance and chrominance, and the R, G, and B components are highly correlated.
- **HSV (Hue, Saturation, Value):** It express Hue with dominant color (such as red, green, purple and yellow)of an area. Saturation measures the colorfulness of an area in proportion to its brightness. The "intensity", "lightness", or "Values" is related to the color luminance. This model discriminates luminance from chrominance. This is a more intuitive method for describing colors, and because the intensity is independent of the color information this is very useful model for computer vision. This model gives poor result where the brightness is very low. Other similar color spaces are HSI and HSL(HLS).
- **CIE –Lab:** It defined by the International Commission on Illumination. It separates a luminance variable L from two perceptually uniform chromaticity variable(a,b)

## 4. HAND SEGMENTATION

Since, efficient hand tracking and segmentation is the key of success towards any gesture recognition, due to challenges of vision based methods, such as varying lighting condition, complex background and skin color detection**;** variation in human skin color complexion required the robust development of algorithm for natural interface. Color is very powerful descriptor for

object detection. So for the segmentation purpose color information was used, which is invariant to rotation and geometric variation of the hand [18]. Human perceives characteristics of color component such as brightness, saturation and hue component than the percentage of primary color red, green, and blue. Color models are useful for to specify a particular color in standard way. It is space-coordinated system within which any specified color represented by single point. Here, three techniques were introduced using different color spaces for robust hand detection and segmentation. Hand tracking and segmentation (HTS) technique using HSV color space is identified for the preprocessing of HGR system.

### 4.1 Anticipated Static Gesture Set

Static gesture is a specific posture assigned with meaning. Following are the static gesture set specified for the proposed system with the specific meaning. Application interface will be provided after recognition of specified posture for action. Simplicity and user frendliness were taken into consideration for the design of anticipated posture set. For the mouse cursor movement the center of the hand gesture window was passed as a mouse cursor. Figure 1. shows anticipated static gestures set with defined task.

**FIGURE 1:** a)Open notepad (b) Open paint (c) Log off (d) Open media player (e) Picture browsing (f) mouse left click (g) mouse right click

### 4.2 Hand Segmentation Using HSV Color Space and Sampled Storage Approach [16]

A Novel Approach for Image segmentation algorithm has been developed and tested for green color glove. In this approach, color based segmentation was attempted using HSV color space. The H, S and V separation was done using following equations.

$$V=\max\{R,G,B\}$$
$$\delta= V-\min\{R,G,B\}$$
$$S= \delta/V$$

To obtain value for hue following are the cases

(i) if $R=V$ then $H=1/6(G-B)/\delta$
(ii) if $G=V$ then $H=1/6(2+(B-R)/\delta)$
(iii) if $B=V$ then $H=1/6(4+(R-G)/\delta)$

The input image of green color samples was passed to the algorithm and from H-S histogram the H_range = [0.4 0.55 0.6 0.6] and S_range = [0.2 1.0] were experimented for segmentation. Algorithm could able to subtract dynamic background. Skin color samples needed to be passed to the algorithm for skin color detection. The drawback of this algorithm was training samples of the color need to be stored. It was sensitive to little variation in color brightness.

### 4.3 Hand Segmentation Using Lab Color Space (HSL)

The Input captured RGB image was converted to lab color space. In CIE L* a* b* co-ordinates, where L* defines lightness, a* represent red/green value and b* denotes the blue/yellow color value. a* axis and +a direction shift towards red while along the b* axis +b movement shift toward yellow. Once the image gets converted into a* and b* planes, thresholding was done. Convolution operation was applied on binary images for the segmentation. Morphological processing was done to get the superior hand shape. This algorithm was found to work for skin color detection but it was sensitive for complex background. Figure 2 shows the output with intermediate steps.

#### 4.3.1 HSL Algorithm

i) Capture the Image

ii) Read the input image

iii) Convert RGB image into lab color space

iv) Convert the color values in I into color structure specified in cform

v) Compute the threshold value.

vi) Convert Intensity image into binary image

vii) Performing morphological operations such as erosion.

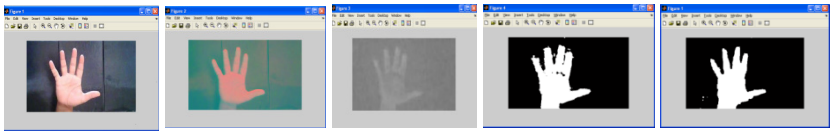

**FIGURE 2:** (a) Input Image (b) rgb2clb (c) Gray Scale (d) Black and white (e) Image after erosion

### 4.4 Hand Tracking and Segmentation (HTS) Algorithm

The prime objective of this algorithm was robust skin color detection and removal of complex background. The limitation of previous two techniques was overcome with the hand tracking algorithm. In this approach**,** hand detection and segmentation were attempted. Hand tracking was done using mean shift algorithm [17]. Odd frame has been considered for fast processing. For robust performance user's skin color sample was passed and HSV histogram was created. Experimented threshold value has been used for the segmentation. Algorithm was found to be working for dynamic color tracking under complex background and able to segment required hand image on which image morphological operations were applied to get the contour. CamShift function (a variation on mean shift algorithm) within the OpenCV library is used for tracking and detection. Edge traversal algorithm was applied for getting fine contour of the hand shape. As dynamic background was considered**,** while capturing the user's gesture after edge detection**,** there was a possibility to detect unwanted edges from the background. In an attempt to only identify the boundary of user's hand edge**,** traversal algorithm was devised. Figure 3 explains the flow chart of HTS algorithm and Figure 4(a) shows that at the run time, skin color sample is passed to the algorithm. Figure 4(b) shows the color of the hand is tracked, circle is drawn for the extraction purpose and Figure 4(c) shows the segmented result after applying edge traversal algorithm. This algorithm was found to work for skin color tracking including for any color glove hand tracking.

#### 4.4.1 HTS Algorithm

i) Capture the image frames from camera.

ii) Process odd frames, tracked the hand using CamShift function by providing skin color samples at the run time.

iii) HSV histogram is created and the experimented threshold value is passed to the CamShift function for tracking required hand portion.

iv) Segment the required hand portion from Image

v) Find the edges by using ***Canny*** edge detection.

vi) Dilate the image.

vii) Erode the image.

viii) Apply edge traversal algorithm to get final contour.

ix) Stop

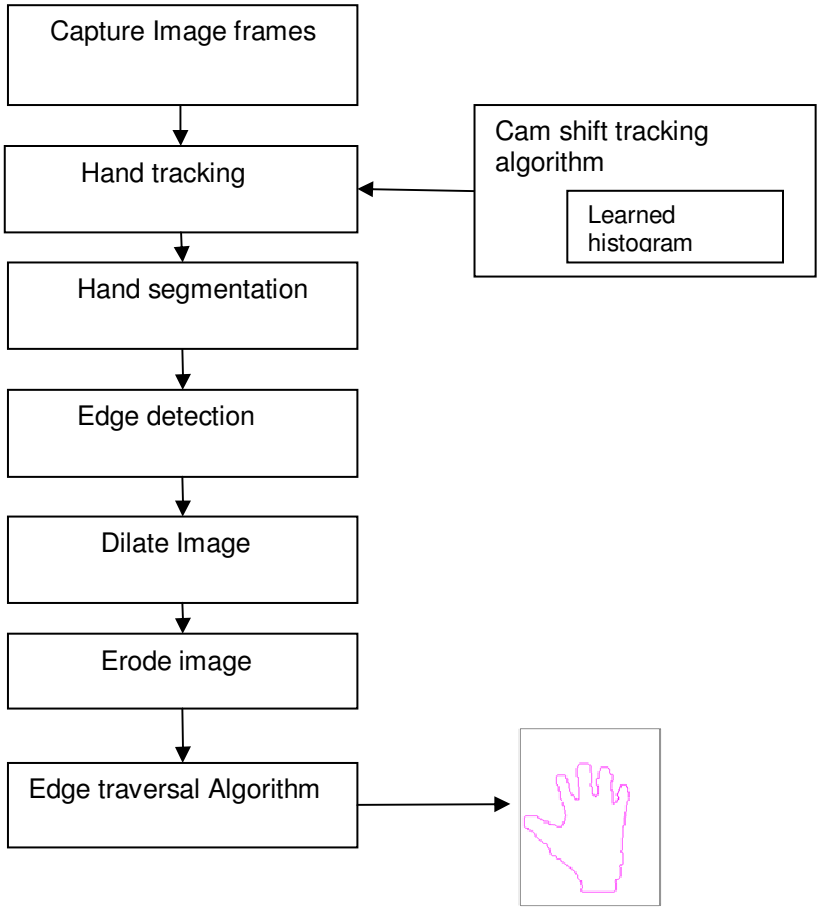


**FIGURE 3:** Hand tracking and segmentation (HTS)

### 4.4.2 Edge Traversal Algorithm

i) Take the center of the window reference

ii) Traverse to the left of the image till we get the first white pixel and record this pixel.

iii) Consider the 3*3 matrix, with the white pixel at the center of the matrix.

iv) Find the previous black pixel**,** which helped to locate the first white pixel.

v) Starting from the previous black pixel, traverse the matrix in anti-clock wise direction till the next white pixel is obtained.

vi) Compare this white pixel with first white pixel recorded.

vii) If new white pixel is same as the first white pixel then go to step (viii)

viii) else record this white pixel and go to step (iii)

ix) Stop

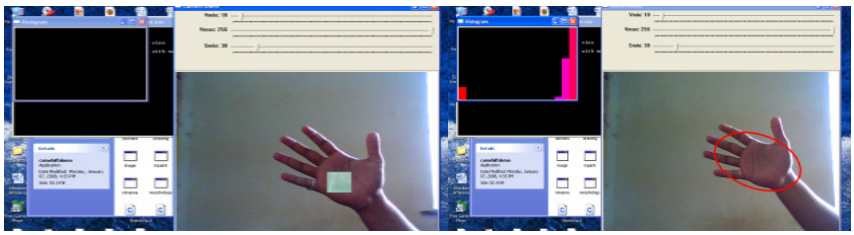

**FIGURE 4:** (a) Learning color histogram (b) Hand tracking

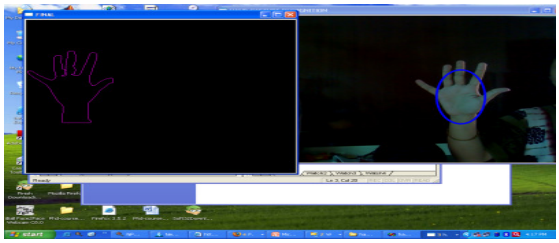

**FIGURE 4:** (c) Hand tracking and final contour

### 4.5 Mouse Cursor Movement

Movement of the mouse cursor is done on the logic where center point of the segmented region is calculated and tracked with respect to x,y position.

i) Capture frames from webcam

ii) Perform hand segmentation(using HTS algorithm)

iii) Centre point detection and tracking

iv) After tracking the centre point, pass the centre point coordinates to the SetCursorPos() Function.

v) Move the cursor position according to the centre point.

vi) In order to perform any operations from mouse, give the appropriate gestures for right and left click.

vii) Stop

## 5. RESULTS

HTS algorithm was tested on four users for five gestures. Instances of three to five gestures on same users were tested. Algorithm was tested on green color glove and skin color on 70 samples. It was observed that using green color glove segmentation process is faster than using skin color due the varying lighting condition. Figure 5 shows segmented results using green color gloves and Figure 6 shows segmented results on skin color.

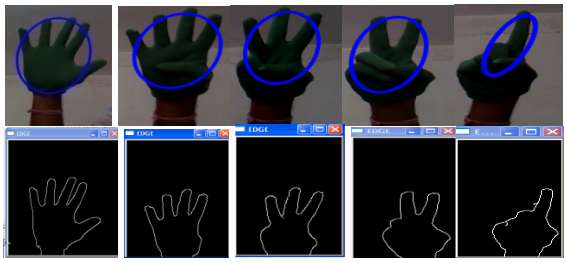

**FIGURE 5:** Segmentation results on green color glove

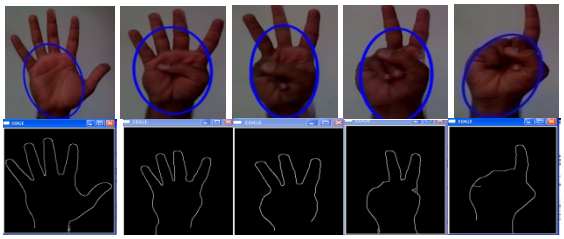

**FIGURE 6:** Segmentation results on skin color

### 5.1 Software/Hardware Used

Sample storage **(**using HSV) and HSL (using Lab) algorithms**,** was done using MATLAB 2010Ra and HTS algorithm was implemented using OpenCV. It was found that for real time interfacing and object tracking**,** OpenCV was a good choice. All algorithms were found to be working on low-resolution web camera (2.0 M pixel and more). For constant lighting source night vision**,** 8 M pixel camera was used for desired results.

## 6. DISCUSSION

There are various approaches are found in survey for hand tracking and segmentation. Different researchers are using different color spaces such as YCbCr[9], HSI[10], YIQ[15] and various techniques, such as gray level segmentation techniques, for getting hand contour [18]. Basically there are two methods for skin detection i) Pixel-Based Segmentation and ii) Region Based Segmentation. Pixel based segmentation uses color based methods. The choice of the color model should consider separability between skin and non skin color and should decrease separability among skin tones. Color distribution has modeled into non parametric and parametric model. However Non parametric Model uses Histogram and Bays classifier for skin color distribution and parametric model derives Gaussian model from training set. In the present study, pixel based non parametric methods were used. Three techniques were implemented and compared. HTS algorithm was giving best result under complex background and skin color detection. It was robust than first two techniques. Using first technique, with HSV color space, the color samples needed to be stored and then threshold value was calculated for segmentation purpose. The limitation of this technique was for the slight variation of the hand color, desired segmentation result could not be achieved. We could segment hand on various skin tone using Lab approach but sometimes result was varying for complex background. Table 1 shows the comparison of discussed techniques with respect to challenging factor.

From the comparison between three techniques it was observed that hand tracking with learned histogram is giving better result than other two for skin color detection under complex background. Algorithms were tested for skin as well as non skin color (with color glove).

| Techniques | Skin color detection | Dynamic background subtraction | Different lighting condition |
|---|---|---|---|
| Storage sampled based technique using HSV | working | Working but if background is having the same hand/glove color then fail | Sensitive |
| Lab based technique | working | sensitive | Sensitive |
| Object Tracking with HSV | working | working | More robust than previous two |

**TABLE 1:** Comparisons of three techniques.

Since, for the detection of the skin color various methods were presented by various researchers such as Bayesian classification with the histogram techniques using various color spaces[12][13][19] ,Bayesian theorem and Kernel Density estimation[20], Using parametric color modeling using Gaussian Mixture Model [21] ,and using unsupervised learning such as K-mean clustering[22]. However for all the classification method, correct choice of color space is very important. Any real time HCI application requires robustness, accuracy and fast processing, for skin color detection, as color varies with illumination the choice of the color space and classification method is really challenge.

## 7. CONCLUSION AND FUTURE WORK

This is an ongoing project and first phase of the project where we tried to implement and tested various techniques for efficient and robust segmentation. Better results on skin color detection using HTS algorithm were achieved**,** where for fast hand tracking, odd frames for processing were considered. The next phase of the project is creation of multiple feature vectors for desired classification and recognition accuracy. Geometric features**,** such as line identification using Hough transform , Fourier descriptor, and Image Hu moments will be used due to the properties of rotation, scale and translation invariant. Genetic algorithm will be used for gesture recognition. Window APIs (Application Program Interface) will be used for specified action after Gesture recognition. Further the system will be extended to manual alphabet Indian sign language Interpretation.

## ACKNOWLEDGEMENT

The authors are thankful to the Department of Science and Technology, Ministry of Science and Technology, New Delhi for supporting this non-commercial research under the Fast Track Scheme for Young Scientists (SR/FTP/ETA-18/08).

## REFERENCES

[1] V. Pavlovis,R. Sharma and T. huang, “Visual Interpretation of Hand Gesture for Human-Computer Interaction: A Review”, *IEEE Transaction on pattern Analysis and Machine Intelligence, Vol. 19, No.7, Jul 1997.pp.,677-695.*

[2] S. Zhao, W. Tan, C. Wu, L. Wen, “A Novel Interactive Method of Virtual Reality System Based on Hand Gesture Recognition”, *IEEE-978-1-4244-2723-9/09,pp., 5879-5882,2009.*

[3] Y. Guan, M. Zheng, “Real-time 3D pointing gesture recognition for natural HCI”, *Proceedings of the world congress on Intilligent Control and Automation, China, pp, .2433-2436.,2008.*

[4] W. Freeman, C. Weissman, "Television control by hand gesture", *IEEE international workshop on Automatic Face and Gesture Recognition, Zurich, 1995.*

[5] A. Sepehri, Y. Yacoob, L. Davis, "Employing the Hand as an Interface Device", *Journal of Multimedia, Vol. 1, No.7, pp., 18-2, 2006.*

[6] A. Erol, G. Bebis, M. Nicolescu, R.Boyle and X.Twombly, "Vision-based hand pose estimation: A review", *Science Direct, Computer Vision and Image Understanding 108,pp., 52-73,2007.*

[7] P. Bao, N. Binh, T. Khoa, "A new Approach To Hand Tracking and Gesture Recognition By A New Feature Type And HMM", *International Conference on Fuzzy Systems and Knowledge Discovery, IEEE Computer Society, 2009.*

[8] J. Alon, V. Athitsos, Q. Yuan, S. Sclaroff, "A Unified Framework for Gesture Recognition and Spatiotemporal Gesture Segmentation", *IEEE Transaction of Pattern Analysis and Machine Intelligence,2008.*

[9] E. Stergiopoulou, N. Papamarkos, "A New Technique for Hand Gesture Recognition",*IEEE-ICIP,pp., 2657-2660,2006.*

[10] C. Burande,R. Tugnayat, N. Choudhary, "Advanced Recognition Techniques for Human Computer Interaction", *IEEE, Vol 2 pp., 480-483,2010*

[11] L. Howe, F. Wong,A. Chekima, "Comparison of Hand Segmentation Methodologies for Hand Gesture Recognition",*IEEE-978-4244-2328-6,2008*

[12] V. Vezhnevets, V. Sazonov. and A. Andreeva, " A Survey on Pixel-Based Skin color Detection Techniques"

[13] S. Phung, A. Bouzerdoum and D. Chai, " Shin Segmentation Using Color Pixel Classification: Analysis and Comparision" *IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.27, No. 1,pp., 148-154,2005.*

[14] A. Elgammal, C. Muang and D. Hu, "Skin Detection – a Short Tutorial", Encyclopedia of Biometrics , *Springer-Verlag Berlin Heidelberg, 2009*.

[15] R. Rokade, D. ,Kokare," Gesture Recognition by Thinning Method", *International Conference on Digital Image Processing, IEEE Computer Society, pp., 284-287,2009.*

[16] A. Ghotkar, "A Novel Approach for Image Segmentation in Real Time Hand Gesture Recognition for HCI", *International Conference(ICSCI-2011),Pentagram Research, Hyderabad ,Jan-2011.*

[17] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift" *Computer Vision and Pattern Recognition, 2000.Proceddings. IEEE Conference on, 2:142-149 vol.2,2000.*

[18] A. Martin and S. Tosunoglu, "Image Processing Techniques for Machine Vision"

[19] C.Jung, C.Kim, S.Chae, and S. Oh, "Unsupervised Segmentation of Overlapped Nuclei Using Bayesian Classification", *IEEE Transaction on Biomedical Engineering, Vol. 57, No.12,2010.*

[20] H. Rahimizadeh,, M. Marhaban, R. Kamil, and N. Ismail, "Color Image Segmentation Based on Bayesian Theorem and Kernel Density Estimation", *European Journal of Scientific Research, ISSN 1450-216 vol.26 No.3, pp., 430-436,2009.*

[21] R. Hassanpour, A. Shahbahrami, and S. Wong, "Adaptive Guassian Mixture Model for Skin Color Segmentation", *World Academy of science, Engineering and Technology 41, 2008.*

[22] A. Chitade, S. Katiyar, "Color Based Image Segmentation Using K-means Clustering", *International Journal of Engineering Science and Technology, Vol.2(10), pp., 5319-5325,2010.*