An Adaptive Algorithm for Hand Segmentation and Tracking for Continuous Hand Posture Recognition

Madhurjya Kumar Nayak
Dept. of Electronics and Communication Engineering,Gauhati University
Guwahati-14, Assam, India.

Anjan Kumar Talukdar
Dept. of Electronics and Communication Engineering,Gauhati University Guwahati-14,
Assam, India

Kandarpa Kumar Sarma
Dept. of Electronics and Communication Engineering,Gauhati University
Guwahati-14, Assam, India

ABSTRACT
This work reports the design of a continuous hand posture recognition system. Hand tracking and segmentation are the primary steps for any hand gesture recognition system. The aim of this paper is to report a robust and efficient hand segmentation algorithm where a new method for hand segmentation using different colour space models with required morphological processing are utilized. Problems such as skin colour detection, complex background removal and variable lighting condition are found to be efficiently handled with this system. Noise present in the segmented image due to dynamic background can be removed with the help of this adaptive technique. The proposed approach is found to be effective for a range of conditions.

General Terms
Gesture Recognition, Hand Segmentation and Tracking.

Keywords
Hand Tracking and Segmentation, Hand Gesture Recognition, Colour based Segmentation, Background Subtraction, Mixture Model.

1. INTRODUCTION
Gesture can be said to be some specific motions of body parts that represents some meaningful data. Gesture recognition is a mechanism through which a machine can understand the meaning of any gesture. Hand Gesture can be subdivided into two types, firstly global motion where the entire hand moves whereas in the second one i.e. local motion (or posture) only the fingers move [1] [2] [3].

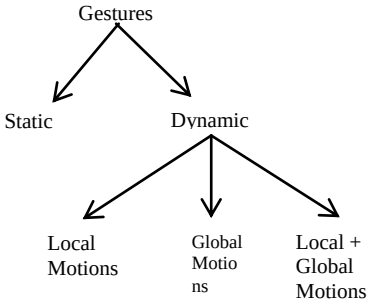

Figure1: Hand Gesture Types

Today's world has many applications of gesture recognition systems such as human computer interface (HCI), robotic arm control, gaming consoles and television control mechanisms, sign language etc [4]. With such widespread applications, it is imperative for us to study and to make this system as user friendly as possible. Hand segmentation is one of the most important process in gesture recognition system as if we get better segmented image of region of interest i.e. hand than better detection rates can be achieved. This paper reports an efficient algorithm for hand segmentation using different spaces employing required morphological processing for filtering. The proposed method tries to minimize the problems that come in gesture recognition systems namely skin color detection, complex background removal and dynamic lighting conditions. Moreover, it has the capability of reducing noise present in segmented image due to complex background.

The rest of the paper is organized as follows. Section II provides a brief review of the design aspects required for the work. Section III describes the proposed model. Section IV contains the experimental results. Section V concludes the work.

2. DESIGN NOTIONS
A generic hand posture recognition system can be implemented using the block diagram as shown below –

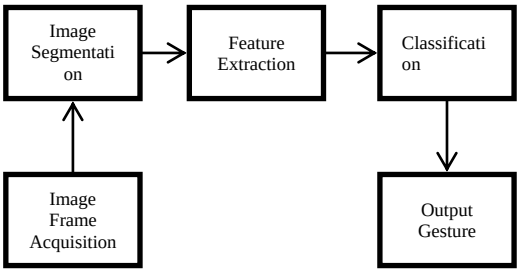

Figure 2: Block Diagram for Gesture Recognition

2.1 Image Acquisation-

The image frame is taken from the input video feed and then processed for the system to get it as a gesture.

2.2 Image Segmentation and Tracking-

Image segmentation is one of the most important steps in this system as without proper segmentation we cannot get a proper gesture input to the system.

Now there are various models that we follow to get the segmented output would be-

- Color Segmentation- In this model we first take the input image and then convert it into HSI or YCbCr color space as color intensity in RGB has to be controlled individually but in YCbCr color space, y controls the intensity. Then then the intensity is adjusted to match the color required (which in our case would be the color of our hand. Then thesholding is done on the image and is converted into a binary image. Noise is minimized using morphological operations like erosion, dilation etc. Disadvantage of this system is that if the background has any object having the same color as the hand, noise will be very high [5].
- Background Subtraction-In this model we first take the image of the background and store it. Now, when we get an image frame then the image is subtracted from the previously stored background. This gives only the moving or dynamic parts which in our case would be the body parts. Disadvantage of this system is that if the lighting conditions change abruptly then there is a change in pixel value where the light intensity changed and additive noise contributes to the output[6] [7].
- Object tracking with HSV color scheme- In this model along with color based segmentation, a hand tracking mechanism is used. Now through this mechanism we first track the hand and then segment it out using color segmentation. This method is more advantageous over both the above mechanisms in tracking. But this model also has a disadvantage that when the motion of the hand which is the region of interest (ROI) in our case moves too fast then the system fails to track it thus add noise to the output of this process [8].

3. PROPOSED MODEL FOR HAND POSTURE SEGMENTATION

Here, we describe the proposed model for hand posture segmentation. At first image frame is captured using a camera than this image is fed to the mixture model which does two processes, first the subtraction and next color segmentation. The subtraction i.e. background subtraction minimizes the noise due to the background containing objects having the same color as that of the hand. It removes all parts except the moving parts. This subtraction output is then color segmented, so that the background will now not have any objects having same color as ROI, so output will contain less noise.

Then face detection and removal is done as the color of face is same as that of ROI. Now this output is then fed to the adaptive thresholding mechanism, where the thresholding is done pixel to pixel. This minimizes the error of the system due to dynamic lighting conditions. Thus we get the segmented hand output from this system. This output is again converted to a binary format so that we can form the contour i.e. the outline of the hand so as to feed in the subsequent parts of the posture recognition system.

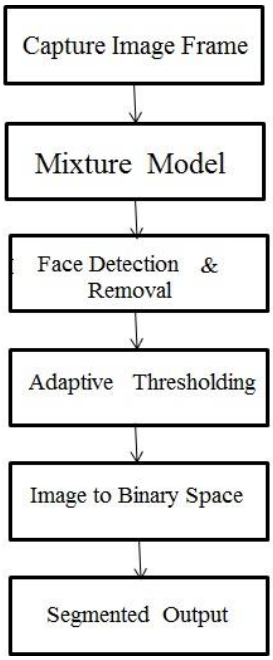

Figure 3: Flow Chart for Proposed model

The proposed mixture can be summarized as below:-

Step 1- The input is first dynamically stored in an array. Then the frame is compared with the current frame. It gives us the moving parts only as output.

Step 2- This image is then converted to HSV color scheme.

Step 3- In this color scheme we then set the range of intensity for a color similar to ROI.

Step 4- This will then highlight the parts having colors near to that of ROI.

Along with this there is a proposed Adaptive thresholding mechanism takes the input from the face detection and removal block. The input will then contain the hand and some noise due to dynamic lighting conditions, if any, in the image. This thresholding uses an adaptive local filtering mechanism to minimize this noise by working on the image pixel by pixel.

Thus the output of the system gives a segmented hand which has noise much less than the outputs of previous systems.

4. RESULTS

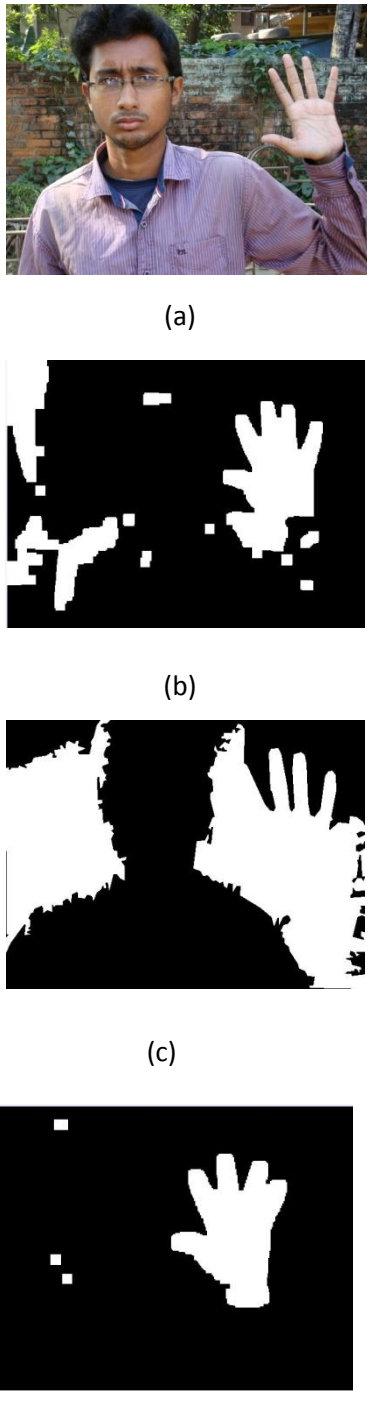

Figure 4: (a) Original Image with dynamic lighting conditions (b) Segmented image using color segmentation. (c) Segmented image using background subtraction. (d) Output from proposed model.

The results obtained from the experiments are shown in figure 4. Figure 4(a) shows the input frame to the system. The figures 4 (b) and 4 (c) shows the outputs of color segmentation method and background subtraction method respectively. Figure 4 (d) shows the output from the proposed system. We can clearly see that noise is much less in the output of the proposed.

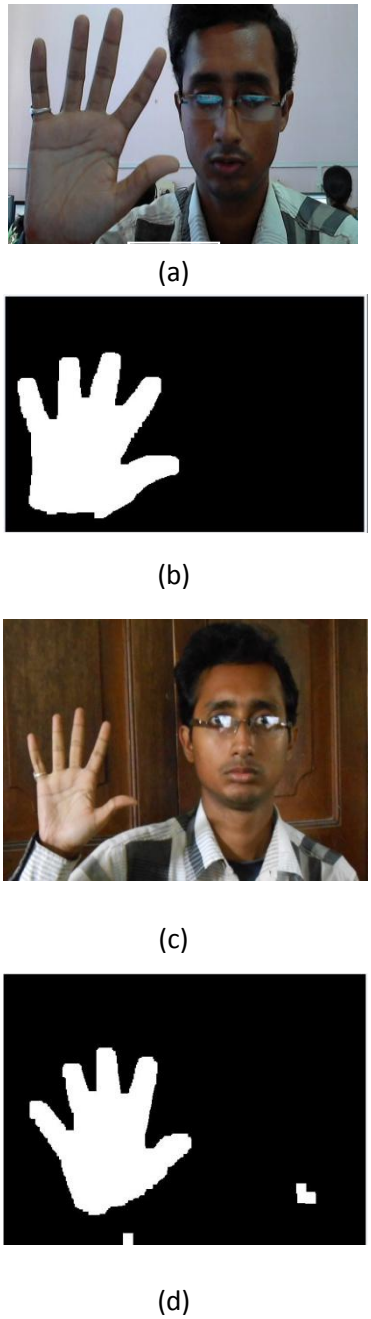

Figure 5: (a) & (b) Input and output from the proposed model in dim light condition. (c) & (d) Input and output from the proposed when complex background is considered.

Again, Figure 5 shows some of the testing output. To determine the robustness of the system, various environments are considered. Figures 5 (a) & 5 (b) shows the input and output of the system in dim light condition and 5 (c) & 5 (d) shows the input and output of the proposed model when we consider a complex background i.e. a background containing objects having color almost similar to that of the ROI. Thus from figure 4 & 5 we can say that this proposed model is very much robust to lighting and complex background conditions.

Table 1: Comparison among outputs obtained from different approaches.

Models	Normal (PSNR)	Low light (PSNR)	Complex Background (PSNR)	Dynamic Lighting (PSNR)
Input (PSNR)	27.6362	27.6362	27.4347	27.6362
Color Based	28.7693	30.2892	28.2881	28.0433
Background Subtraction	27.9975	29.8822	30.2125	27.8314
Proposed Model	**31.0983**	**32.2328**	**31.4188**	**32.0660**

Table I shows the comparative output results for model as given in table. All the systems are tested in various conditions and the outputs are given in a qualitative manner.

The PSNR is defined as

$$PSNR = 10\log_{10}\frac{255^2(A.B)}{\sum_{i,j}(D(i,j)-F(i,j))^2}$$

where, D is the output image, F is the input image and $(A \times B)$ is the size of the image; i,j are the pixels values.

Table 2: Comparative noise reduction performance of the proposed model.

Models	Normal Conditions	Complex Background	Dynamic Lighting.	Tracking fast hand movements.
Color based	Working	Sensitive	Working	Working
Background Subtraction	Working	Working	Sensitive	Sensitive
Object Tracking based	Working	Robust	Robust	Sensitive
Proposed Model	**Working**	**Robust**	**More Robust then earlier**	**Robust**

Table II shows us the comparative noise reduction performance of the proposed model.

Earlier models such as color segmentation had problems when the background contained any object which had color almost similar to ROI. Also background subtraction failed to give suitable output when dynamic lighting conditions are taken into consideration. Also sometimes the input gesture maybe very fast, this makes detection through object based tracking process to contain large amount of noise. The proposed algorithm can deal with these problems very efficiently.

In continuation with the proposed algorithm a new model for tracking the hand for different viewing angles is being thought off. This new tracking algorithm will be able to track even when the gesturer tilts the hand by some degree or gives very fast gesture input. This would allow the gesturer more degree of freedom and also make the system much more user friendly.

5. CONCLUSION

Here, we proposed an adaptive algorithm for hand segmentation and tracking of continuous hand posture recognition. We have tested the system the system under illumination and background variations. The results establish the effectiveness of the system. But like other known system of similar type dynamic background puts restriction on the system performance. Subsequent work shall focus on this aspect.

6. ACKNOWLEDGMENTS

Our heartfelt gratitude to all the teachers and fellow students for their help and suggestions regarding this project.

7. REFERENCES

[1] T. Ishihara, N. Otsu, "Gesture Recognition Using Auto-Regressive Coefficients of Higher-Order Local Auto-Correlation Features," Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition,pp 583-584, 2004.

[2] F. Flórez, J. M.I. García, J. García and A. Hernández, "Hand Gesture Recognition Following the Dynamics of a Topology-Preserving Network", Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition,pp.318-319,Washington, 2002.

[3] H. K. Lee and J. H. Kim, "An HMM-Based Threshold Model Approach for Gesture Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 10, pp. 961-963, 1999.

[4] H. D. Yang,S. Sclaroff and S. W. Lee, "Sign Language Spotting with a Threshold Model Based on Conditional Random Fields", IEEE Transactions On Pattern Analysis And Machine Intelligence, vol. 31, no. 7. pp 1264-1265, 2009.

[5] S. L. Phung, A. Bouzerdoum and D. Chai, "Skin Segmentation Using Color Pixel Classification: Analysis and Comparison," IEEE Transactions On Pattern Analysis And Machine Intelligence, vol. 27, no. 1, pp 148-151, 2005.

[6] C. Manresa, J. Varona, R. Mas and F. J. Perales, "Real –Time Hand Tracking and Gesture Recognition for Human-Computer Interaction," Computer Vision

Center Universitat Autonoma de Barcelona, Barcelona,.pp-1-3, , Spain.2000.

[7] Y. C. Lu, "Background Substraction Based Segmentation Using Motion Feedback," First International Conference on Robot Vision and Signal Processing(RVSP), pp 224-227, Kaohsiung ,2011.

[8] E. Emami, M. Fathy, "Object Tracking Using Improved CAMShift Algorithm Combined with Motion Segmentation", 7th Iranian Machine Vision and Image Processing(MVIP), pp 1-4,Tehran, 2011.