International Journal of Electronics Signals and Systems

Volume 3 | Issue 4 Article 3

April 2014

A COLORED FINGER TIP-BASED TRACKING METHOD FOR CONTINUOUS HAND GESTURE RECOGNITION

DHARANI MAZUMDAR
Dept. of Electronics & Communication Engineering, Gauhati University, Guwahati-14, Assam, India,
dharani_mazumdar@yahoo.com

ANJAN KUMAR TALUKDAR
Dept. of Electronics & Communication Engineering, Gauhati University, Guwahati-14, Assam,
anjan.nov@gmail.com

Kandarpa Kumar Sarma
Department of Electronics and Communication Technology, Gauhati University, Guwahati, Assam-781014,
kandarpaks@gmail.com

Recommended Citation

MAZUMDAR, DHARANI; TALUKDAR, ANJAN KUMAR; and Sarma, Kandarpa Kumar (2014) "A COLORED FINGER TIP-BASED TRACKING METHOD FOR CONTINUOUS HAND GESTURE RECOGNITION," *International Journal of Electronics Signals and Systems*: Vol. 3 : Iss. 4 , Article 3.
DOI: 10.47893/IJESS.2014.1176
Available at: https://www.interscience.in/ijess/vol3/iss4/3

A COLORED FINGER TIP-BASED TRACKING METHOD FOR CONTINUOUS HAND GESTURE RECOGNITION

DHARANI MAZUMDAR[1], ANJAN KUMAR TALUKDAR[2] & KANDARPA KUMAR SARMA[3]

[1,2&3]Dept. of Electronics & Communication Engineering, Gauhati University, Guwahati-14, Assam, India
E-mail: dharani_mazumdar@yahoo.com, anjan.nov@gmail.com, kandarpaks@gmail.com

Abstract- Hand gesture recognition system can be used for human-computer interaction (HCI). Proper hand segmentation from the background and other body parts of the video is the primary requirement for the design of a hand-gesture based application. These video frames can be captured from a low cost webcam (camera) for use in a vision based gesture recognition technique. This paper discusses about the continuous hand gesture recognition. The aim of this paper is to report a robust and efficient hand segmentation algorithm where a new method, wearing glove on the hand is utilized. After that a new idea called "Finger-Pen", is developed by segmenting only one finger from the hand for proper tracking. In this technique only a finger tip is segmented in spite of the full hand part. Hence this technique allows the hand (excepting the segmented finger tip) to move freely during the tracking time also. Problems such as skin colour detection, complexity from large numbers of people in front of the camera, complex background removal and variable lighting condition are found to be efficiently handled by the system. Noise present in the segmented image due to dynamic background can be removed with the help of this adaptive technique which is found to be effective for the application conceived.

Keywords*- Hand gesture recognition, Finger-pen, Segmentation, Tracking, Complex background.*

I. INTRODUCTION

Gestures recognition system increasingly becoming significant part of human-computer interaction. Gestures can originate from any bodily motion or state but commonly originate from the face and/or hand. A gesture is a movement of the body parts that contain information and/or feelings [1][2]. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed. Many researchers have tried to define gestures but their actual meaning is still arbitrary. S. Mitra and T. Acharya stated gestures as expressive, meaningful body movement involving physical movements of the fingers, hands, arms, head, face, or body with the intent of 1) conveying meaningful information or 2) interacting with the environment [3]. According to V. I. Pavlovic, R. Sharma and T. S. Huang gestures originate as a mental concept, possibly in conjunction with speech. They are expressed through the motion of arms and hands, the same way speech is produced by air stream modulation through the human vocal tract. Also, observers perceive gestures as streams of visual images which they interpret using the knowledge they possess about those gestures [4].

Gesture recognition is a topic pursued with the goal of interpreting human motions via mathematical algorithms. Gesture recognition can be seen as a way for a computer to begin to understand human body language, thus building a richer bridge between machines and humans than primitive text interfaces or even GUIs (graphical user interfaces), which still limit the majority of input to keyboard and mouse.

Hence, gesture recognition is a process by which the system will know, what is going to be performed by the gesturer. Gesture recognition can be conducted with techniques from computer vision and image processing. Gesture recognition enables humans to interface with the machine and interact naturally without any mechanical devices [5]. Using the concept of gesture recognition, it is possible to point a finger at the computer screen so that the cursor will move accordingly. This could potentially make conventional input devices such as mouse, keyboards and even touch-screens redundant.

Hand gesture recognition system can be used for interfacing between computer and human using hand gesture. Proper hand segmentation is the primary requirement from the background and other body parts of the image frame for the design of a hand-gesture based application. These image frames can be captured from a low cost webcam (camera) for use in a vision based gesture recognition technique. This paper discusses about the continuous hand gesture recognition. The aim of this paper is to report a robust and efficient hand segmentation algorithm where a new method, wearing glove on the hand is utilized. After that a new idea called "Finger-Pen", is developed by segmenting only one finger tip from the hand for proper tracking. In this technique, only the tip of the finger is segmented instead of whole hand (palm). Problems such as skin color detection, complexity from large numbers of people in front of the camera, complex background removal and variable lighting conditions are found to be efficiently handled by the system. Noise present in the segmented image due to dynamic background can be removed with the help of this adaptive technique which is found to be effective for the application conceived.

This paper is organized as follows: basic theoretical considerations are presented in Section II. Section III covers our proposed model. Experimental results and discussions and conclusion are described in Sections IV and V respectively.

II. BASIC THEORITICAL CONSIDERATIONS

A generalized block diagram for gesture recognition process is shown in fig 1.

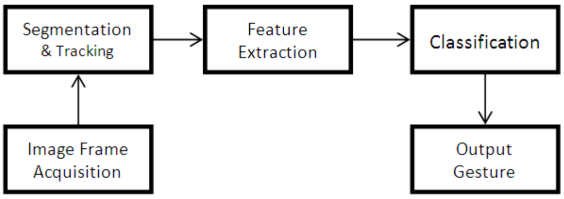


Figure 1: Block diagram for gesture recognition system

Hand segmentation is the pre-requisite to track the movement of the hand. There are various types of commonly used hand segmentation techniques as described below-

- Segmentation based on skin color- In this technique, depending upon the skin color, the rest of the background is removed. By doing this, the hand and face parts are segmented as they have almost same color. After this, the face region in the frame is removed by using a face detection algorithm, since we are interested only on the hand motion. Then the original color plane is converted to two separate HSV and YCbCr color planes, and a threshold is applied to chrominance component of both color spaces. RGB plane is converted to HSV because hue component gives the average of the R, G and B components. Saturation represents the amount to which that respective color is mixed with white and value represents the amount to which that respective color is mixed with black. A logical 'AND' operation is done between them to get the most probable skin region (hand). Noise is minimized using morphological operations like erosion, dilation etc. [6].
- Segmentation based on image intensity- 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 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 [7].
- Background Subtraction- Apart from the above two process, hand segmentation can be done with the help of background subtraction. In this method first, the image of our working background (without gesturer) is stored. Now the image frame (with gesturer) is subtracted from the previously stored background image plane. This gives the image of gesturer's body parts. Now we required to perform the operation for detecting face and segmented the moving hand. 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 [8][9].

III. PROPOSED MODEL FOR HAND SEGMENTATION AND TRACKING

All the above techniques (section II) are very sensitive to varying lighting conditions and dynamic background situation. Apart from these drawbacks, there are some limitations related to gesturer as below [10]-

1. Gesturer should always wear full sleeve to easily detect the palm part from rest of the body parts.
2. It is difficult to detect the correct gesturer when more than one gesturers are there. (Fig.2)
3. When the hand comes in front of the face, then it is difficult to separate it from the face. (Fig. 2)

(a) (b)

Figure 2: Common problems of the conventional system

To overcome the various limitations as described in the above, we propose a new technique based on the hue components parts of a HSV color plane. For this model, we need a low cost hand glove, which is easily available. There are no such limitations for our technique as discussed above.

Figure 3 shows the process logic of our proposed model. HSV color plane is the most suitable color plane for color based image segmentation. So, we have converted the color space of original image frame of the camera, i.e. from RGB to HSV plane. Now setting the hue value ranges from the lower threshold to upper threshold for the particular color of hand gloves, we can easily eliminate all the other parts of the image frames. Finally, this output is converted to binary form and certain morphological operation likes erosion, dilation etc. are carried out. It results in a noiseless segmented hand or region of interest (ROI) that can be used subsequently.

The track path found by this segmented hand is better than the other conventional system output. But, due to change of only one pixel of the segmented region, track path deformed. So, this method creates some limitations on the degrees of freedom of the hand. At the tracking time, if by mistake the orientation or shape of the hand is changed, number of pixel may change. Hence, the position of the centroid changes. So, we have modified our idea by segmenting a very small region from the hand. For this we wear a cap of the same color as glove on the finger as seen in figure 4. This cap works as ink of a pen for tracking purpose. Hence, the method is named as 'Finger-Pen'.

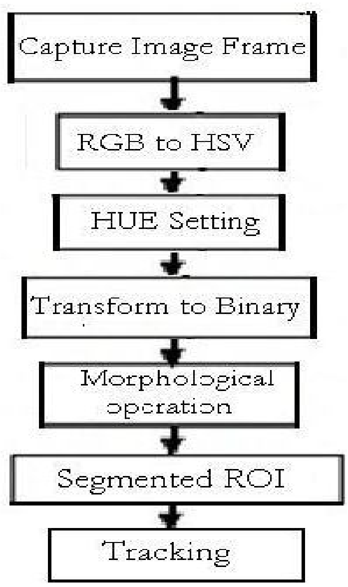


Fig 3: Process of our proposed model

Fig 4: Development of finger-pen from naked hand

Figure 4a shows a naked hand used for conventional segmented technique. Figure 4b shows a gloved hand, which is used in our proposed model. Figure 4c shows a 'Finger-Pen' developed from the gloved hand by wearing a cap on the finger tip.

In our technique, 'hue' setting is the main controlling factor. 'Saturation' and 'value' may be varying according to the lighting condition of that environment. So, these two factors do not create any major problem for good segmented result. Hue value of red color generally in the range of 160-179. We found that the range of 172-179 is perfect for the range of hue values of our red glove.

The proposed model can be summarized as below:-

Step 1- The input image frame is first converted to HSV color plane from RGB plane.

Step 2- Hue value is then set properly for the color of the hand glove or the cap of the finger.

Step 3- HSV plane is converted to binary plane with proper thesholding.

Step 4- Certain morphological operations such as erosion-dilation on the binary plane gives a proper segmented result and the motion of the segmented ROI gives a track path.

IV. RESULTS AND DISCUSSIONS

A few samples are recorded by a webcam and indoor-outdoor sets are created for use with the proposed system. The samples consider illumination and background variation. A glove is used to make the system more robust.

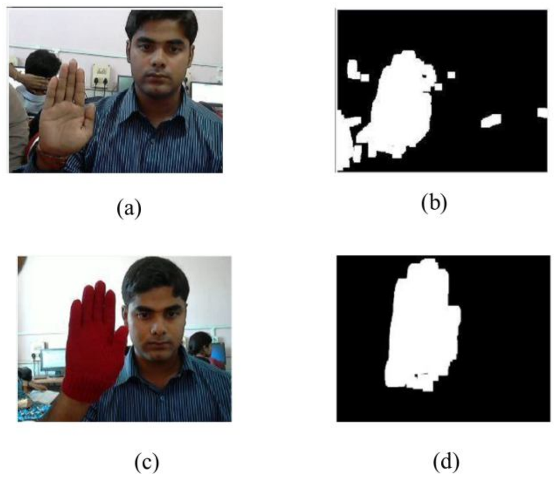


Fig 5: Segmentation result at static background condition

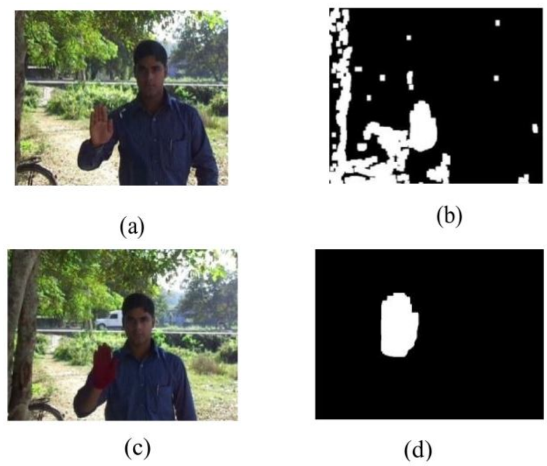


Fig 6: Segmentation result at dynamic background condition

Figure 5a is the input to any of the ordinary segmentation technique at static background with poor light condition, which gives a noisy output as shown in fig 5b. All the three ordinary segmentation techniques give almost same output. So, we have shown here only one result. But the same samples when are applied as input as shown in fig 5c, to our proposed model gives a better segmented output as shown by figure 5d. We have performed a number of trials by changing the lighting conditions, but our proposed model was robust to all these variable conditions and gave the same output. This is because the proposed system shows no effect to the hue value of the glove.

Again, figure 6 shows some of the test results at dynamic background conditions. To know the robustness of the system, various environments are considered. Figure 6b is the output from one of the commonly used segmentation technique for the input as shown in figure 6a. This gives a noisy segmented result. Figure 6c and 6d shows the input and output of our proposed model respectively. Hence we can conclude that this proposed model is robust to lighting and complex background conditions and it does not depend on any situation.

We observed the segmented output of various segmentation techniques. Table I shows the result of these technique at static and dynamic background conditions. For static background condition, we have observed the situation like normal, complex background and dynamic lighting. Among these dynamic lighting condition is very sensitive than the situation basically for conventional segmentation techniques We have found that at any conditions our system is robust and effective.

Table I: Comparisons of segmented output among the systems

Segmentation Models	Normal	Complex background	Dynamic Lighting	Dynamic Background
Skin Color based.	Working	Sensitive	Working	Noisy output
Background Subtraction	Working	Working	Sensitive	Very Noisy output
Image Intensity based	Working	Robust	Sensitive	Sensitive
Proposed Model	Working	Robust	Robust	Robust

We have calculated the PSNR value of the output results of the various techniques as shown in Table II. The PSNR is defined as-

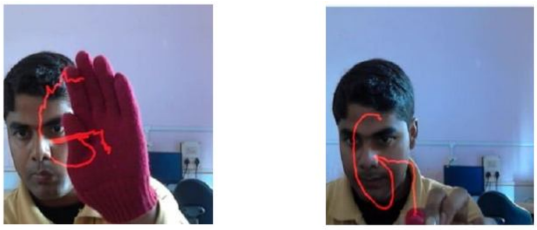
$$PSNR = 10\log_{10}\frac{255^2(A.B)}{\sum_{i,j}(D(i,j)-F(i,j))^2}$$

where, *D*□ is the output image, *F*_ is the input image and *(A X B)*is the size of the image; i,j are the pixels values.

The PSNR values further establish the advantages of the proposed method.

Models	Normal (PSNR)	Complex Background (PSNR)	Dynamic Lighting (PSNR)	Dynamic Background (PSNR)
Input (PSNR)	27.5342	27.5342	27.5342	27.5342
Color Based	28.1697	28.2862	27.8681	27.9334
Intensity Based	28.8935	29.7852	29.2145	27.7313
Proposed Model	**30.1923**	**30.2398**	**30.9188**	**31.0234**

We have tested a few number of meaningful gestures with gloved hand and finger-pen by tracking. The tracking result found is independent to background condition. So, we have shown here a single result of tracking by the gloved hand and ‘finger-Pen’ as seen in figure 7.



(a) (b)

Fig 7: Tracking result with gloved hand and finger-pen

Figure 7a and 7b shows the tracking result by the gloved hand and by the ‘Finger-pen’ respectively for a gesture ‘G’. It is clear from the figure that tracking found from the ‘Finger-Pen’ is better than the gloved hand. During the tracking if at least one pixel value of the segmented region change, then the COG (center of Gravity) of the segmented region changes. Hence, the path deforms. So, the orientation or shape of hand should not be change during tracking. It creates some limits to the degrees of freedom to the naked or gloved hand for tracking purpose. But the ‘Finger-Pen’ offers higher degrees of freedom to the other parts of the hand excepting the caped finger tip.

V. CONCLUSION

As compared to the other techniques, our proposed model is not dependent on any background and lighting conditions. On a hot day, there is no necessity to put on full sleeve for gesture recognition process using our technique, because it is not sensitive to body colors (for this color of the glove or finger cap should not be same with body color). In continuation with the proposed algorithm we can easily find out the back side (top part) of the gesturer hand by wearing glove of two different colors. Real time recognition process with no time delay is one of the features offered by our model. We can use it for real time purpose at any environment. This method allows moving freely all the fingers excepting the caped one, during the tracking time also. So, this tracking method is more useful than the tracking by a gloveless or gloved hand tracking.

REFERENCES

[1] F. Florez, J. M. Garcia, J. García and A. Hernandez, "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.318-323, 2002.

[2] 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. 583-588, 2004.

[3] S. Mitra and T. Acharya, "Gesture Recognition: A Survey", IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, vol. 37, No. 3, pp. 311-324, 2007.

[4] V. I. Pavlovic, R. Sharma and T. S. Huang, "Visual Interpretation of Hand Gesturesfor Human-Computer Interaction: A Review", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 677-695, 1997.

[5] E. S. Nielsen, L. A. Canalís and M. H. Tejera, "Hand Gesture Recognition for Human-Machine Interaction", Journal of WSCG, vol. 12, no.1-3, 2004.

[6] 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.

[7] L. Gu and D. Bone, "Skin Colour Region Detection in MPEG Video Sequences", in Proc. IEEE International Conferences on Image Analysis and Processing, Italy, pp. 898-903, 1999.

[8] Y. Wang, K. F. Loe and J. K. Wu, "A Dynamic Conditional Random Field Model for Foreground and Shadow Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 2, pp. 279-286, 2006 .

[9] M. H. Yang, N. Ahuja, M. Tabb, "Extraction of 2D Motion Trajectories and its Application to Hand Gesture Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 8, pp. 1061-1074, 2002.

[10] G. R. S. Murthy , R. S. Jadon," A Review Of Vision Based Hand Gestures Recognition", International Journal of Information Technology and Knowledge Management, vol. 2, no. 2, pp. 405-410, 2009.