International Journal of Engineering Research and General Science Volume 3, Issue 4, Part-2, July-August, 2015
ISSN 2091-2730

SKIN COLOR DETECTION AND BACKGROUND SUBTRACTION FUSION FOR HAND GESTURE SEGMENTATION

Yogesh B.Dongare, Prof.Ramesh Patole
G.H.Raisoni College of Engineering & Management,Chass,Ahmednagar, India.
yogesh.dongare05@gmail.com
rameshpatole@gmail.com

ABSTRACT

Hand gestures recognition play a important role in Human Computer Interaction(HCI). They serve as primary interaction tools for gesture based computer control. The present work is a part of vision based hand gesture recognition system for Human Computer Interaction. Here we have proposed an algorithm with the fusion of skin color model and background subtraction that gives robust output in the presence of drastic illumination changes. In this we paper we compare different methodologies of hand segmentation approaches for gesture recognition systems. This study is a first step towards development of a reliable, efficient and robust gesture recognition system with high detection rate for better performance of system on Android mobile device.

Keywords: hand gesture detection, appearance based segmentation, skin color detection, background subtraction;

INTRODUCTION

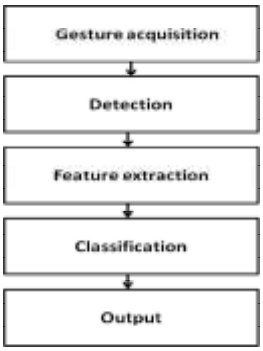
As computers become more usefull in society, facilitates natural human-computer interaction (HCI) have a very positive impact on their use [1].By using various improvements in the image acquisition and image processing technology, hand gestures become a significant,growing and popular tool in human computer interaction (HCI) systems [2].This hand gesture recognition is a field of research with growing applications in different area of research. The applications where we can use this techniques are Sign language recognition, Virtual reality, Robotics, Gesture-to speech, Television control, Smart room interactions and medical interactions [3]. The goal of gesture recognition system is the rendering of the semantics that the hand(s) location, posture of hand, or gesture information conveys.

Segmentation of gestures is the separation of hand gestures from a continuous image sequences containing gestures [2]. Hand segmentation is the key and major task to the analysis of gesture and detection of a gesture. A static gestures is a specially hand configuration and pose and this is represented by a single image. A dynamic gesture are the moving gestures, represented by a sequence of images [3]. Quality of the gesture segmentation in real time affects the rate of recognition of gestures. Effective use of various information of images such as color, motion, geometric information are the key of the study [2].For the detection of static gestures (i.e. postures), a general classifier or a template-matcher are used. Hand detection an be done in various ways such as pixel based [4], shape based [5], 3d model based [6], motion based [7] parameters. However, dynamic hand gestures have a temporal characteristcs and requires techniques that can handle this dimension like Dynamic Time Warping, Time Delay Neural Network, Finite State Machines, and Advanced Hidden Markov Models (HMM) [1]. Hand tracking can be done by using template based, optimal estimation, particle filtering and camshift algorithms [1].

Segmentation can be mainly done through two major mostly used technologies for human computer interaction are contact based devices and vision based devices. Contact based device method is the old traditional technique for modeling of Gesture. The user signs by using gloves and the system will perform functions through sensing apparatus like wires, accelerometers and multi touch based detectors [8]. This is a 3d model which gives a direct measurement of hand position, joint angle and orientations [3]. Limitations are the system can detect and recognize only single hand gestures. Experiencing all these, the contact based devices will not provide much acceptability in hand gesture recognition, so that vision based devices have been introduced for taking images for gesture recognition in human computer interaction [1].

Vision based techniques are simple to use, easiest and affordable to implement human computer interaction in real-time. Contact based devices are user cooperative for user,precise and flexible to use and configure, on the other hand ,Vision based devices are flexible to use and healthy [1]. The two major types of vision based hand gesture representation are 3D model based methods and appearance based methods. The 3D model based method is advantageous that it updates the different model parameters while checking the matches of transition in temporal model, which leads to proper hand gesture recognition and representation, though making it computationally important with requirement of dedicated hardware [9]. Appearance based hand gesture representation has

color based model, silhouette geometry model and motion based model. Appearance based hand gesture representation methods are classified into two subcategories,2D static model based methods and motion based methods [3]. The generalized block diagram of appearance based static approach is as shown in the fig 1.



.

Fig 1: Appearance based Approach

The device camera is used for the purpose of collecting images or video frame for the process of image or gesture acquisition. Hand Detection contains Segmentation and Edge Detection steps. After performing segmentation of the hand gestures, an edge traversal algorithm is used on the contour of segmented hand for removing noise from unwanted background. Feature Extraction is used to calculate particular features because of that is variation in the image data. Features that are not useful are discarded. Classifier is used to identify the hand gesture from the alphabets of the sign language used [2]. Generally, if there is large vocabulary then recognition task becomes complex.

Below discussion will explain the Appearance based segmentation, current difference and skin color fusion algorithm, comparison with previous algorithms and Conclusion.

APPEARANCE BASED SEGMENTATION

In Gesturer Localization, the person who is performing the gestures is extracted from the rest of the visual image scene [7]. Appearance based static method includes finding target region from the intensity image that includes data descriptive of a hand. These methods utilize several types of visual features like skin color, shape, motion and anatomical models of hands for detection of human hand motion and gestures[10] 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 hand segmentation[11]. Thresholding technique is applicable for simple hand images in the static, uniform backgrounds.

Clustering technique is also used at initial stages [5]. Initially this algorithm locates k clusters in the image. Each pixel in the image is grouped to the nearest cluster; clusters are moved to the average of their class values. This process is repeated until the stopping condition is met [5]. The time complexity of this algorithm is very less but false detection rate is high.

Color based segmentation generally rely on histogram matching, look up table approach and skin pixel data training through various color space [7]. Several color spaces have been proposed including RGB, normalized RGB [2], HSV [12], YCbCr [2], YUV [13], etc. Color spaces that separate luminance component from the chrominance component are preferred. This is due to the fact that by employing chromaticity-dependent components of color only, some degree of robustness to illumination changes and shadows can be achieved [1].

Burande et al. [14] implemented Blobs analysis technique for skin color detection under complex background. In this technique, several skin colored blobs are formed by making connected components and hand region is detected.

The major drawback of color based segmentation is color of human skin varies greatly across human races or even between individuals of the same race. In general, color segmentation can be confused by background objects that have a color distribution similar to human skin. Background subtraction can be done to overcome this problem. However, background subtraction is typically based on the assumption that the camera system does not move with respect to a static background. The difference in luminance of pixels from two successive images is close to zero for pixels of the background [15].

Segmentation handles the challenges of vision based system such as skin color detection, complex background removal and variable lighting condition. Efficient segmentation is the key of success towards any gesture recognition.

CURRENT APPROACH

Our current approach is a fusion of skin color and background 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 background subtraction was used to find the difference between the hand gesture image and the background object.

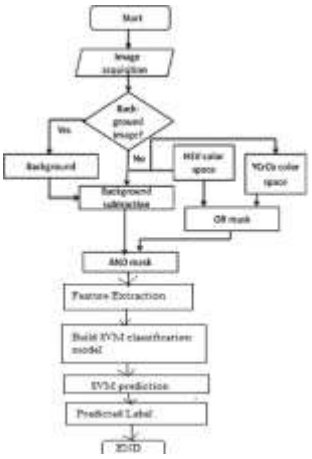


Fig 2: Flowchart of proposed algorithm

I. Background subtraction:

In gesture making only position of the hands of gesturer will change, where as background and other body parts remain almost static. In image acquisition, background image (bgr(x,y)) without hand gesture is taken initially. The new image taken is considered as foreground image (fgr(x,y)). To isolate the gesture (gst(x,y)) from image, difference principle is applied.

$$gst_i(x,y) = fgr_i(x,y) - bgr(x,y) \quad (1)$$

Difference foreground gesture obtained is converted into binary image setting appropriate threshold. As the background is not perfectly static, noise part is added. To obtain clean hand image, this technique is combined with skin detection. To remove this noise, connected component analysis is applied, and to fill the holes in hand if any region filling is applied and to obtain clear edges morphological processing is applied.

II.Skin detection in HSV and YCrCb color space:

Every color image is composed of three planes namely hue (H), saturation (S) and value (V). To extract the hand from the image, foreground image is decomposed into H , S and V planes. The following threshold is imposed on each plane to extract skin regions like hand, head and the rest body parts.

$$0.05<H<0.17$$
$$0.1<S<0.3$$
$$0.09<V<0.15 \quad (2)$$

In YCbCr color space, every color image is decomposed into Yellow (Y), Chromium blue (Cb) and Chromium red (Cr) planes. Threshold is applied in YCbCr color space as shown below.

$$76<Cb<128$$
$$132<Cr<172 \quad (3)$$

Results obtained from two color spaces are converted into binary images and added to maximize the skin color detection. Finally results obtained from background subtraction and skin color models are multiplied to eliminate the body parts other than hand. If the

background is in the range of skin color, false detection is eliminated to a considerable manner. Region filling and morphological processing are done to enhance the gesture image.

III. Experimental Results:

In accordance with the established background subtraction of equation (1), the binary gesture image is obtained. The background and foreground images are shown in the fig 3(a) and (b).

Fig 3 a) Foreground image a) Difference image b) Color space segmented image c) Hand gesture image

Hand gesture area detected contained some interference region caused due to clothes. Biggest blob analysis is implemented to obtain hand region as shown in the fig 3(c).

The skin color detection method in HSV and YCrCb color spaces is used to make skin color segmentation on the foreground image in correspond to equations (2) and (3). The hand region is highlighted after skin color detection as shown in fig 3(d). The two images of background subtraction and skin segmentation were multiplied. Region filling and morphological operations were performed to enhance the image. Then the hand gesture image is obtained as shown in fig 3(e).

COMPARISON WITH PREVIOUS APPROACHES:

All the above discussed Appearance based approaches are implemented on different platform and are compared with our technique. The results of segmentation methods implemented are shown in the table (1). Compared with the previous approaches, our proposed algorithm is illumination invariant, skin color and shadow insensitive. It is applicable to even skin color backgrounds to obtain data descriptive of hand to a considerable extent.

Approach	**Implementation**	**Characteristics**
Dynamic thresholding(gray level segmentation)[11]	 Fig 4: a)Input Image (from web) b)Segmented Image	Single posture, uniform background, non -real time.
Clustering technique(LAB color space) [5]	 Fig 5: a)Input Image (from web) b)Segmented Image	Single posture, uniform background, skin color sensitive, non-real time.
Skin modeling (YIQ and YUV color spaces)[13]	 Fig 6: a)Input Image b)Segmented Image	Illumination invariant and applicable to complex background, but sensitive to skin color.

Blob analysis(YCBCR color space [2,14]	 Fig 7: a)Input Image b)Segmented Image	Applicable to complex background, but depth dependent. False detection rate is high.
Background subtraction[15]	 Fig 8: a)Background Image b)Foreground Image c)Segmented Image	Applicable to static complex background , but illumination variant and shadow sensitive.
Proposed approach	 Fig 9: a)Background Image b)Foreground Image c)Segmented Image	Illumination invariant, skin color insensitive, low false detection rate. Applicable to complex but static backgrounds.

Table 1: Implementation of Segmentation Approaches

CONCLUSION

Hand gesture segmentation is the most important step in the gesture recognition systems. Segmentation process has direct impact on balancing accuracy-performance-usefulness trade-off of recognition systems. The algorithm we implemented is robust with respect to drastic illumination changes and cluttered backgrounds. The proposed algorithm fails if the hand region overlaps the face. In future the focus would be on improving the algorithm to avoid false detections if the hand is overlapped with face. This is merely the first step towards implementation of effective gesture recognition system. Further the project will be extended to recognize the detected gestures.

REFERENCES

[1]
S Rautaray, A Agarwal. "Vision Based Hand Gesture Recognition for Human Computer Interaction: A Survey". Springer Science & Business Media Dordrecht 2012.

[2]
makant Yadav. "Fusion of Information from Data-Gloves and a Camera for Hand Gesture Recognition", M.Tech thesis, 2011.

[3]
I Pavlovic, R Sharma, T S Huang. "Visual Interpretation of Hand Gestures for Human Computer Interaction: A Review", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, 1997.
Viola, M Jones. "Robust Real-Time Object Detection", IEEE Workshop on Statistical and Computational Theories of Vision, Vancouver, 2001.

[4]
Panwar, P S Mehra. "Hand Gesture
Recognition for Human Computer Interaction", Proceedings of IEEE International Conference on Image Information Processing (ICIIP 2011), Waknaghat, India,November 2011.

[5]
Rehg, T Kanade. "Visual Tracking of High DoF Articulated Structures: An Application to Human Hand Tracking," European Conference on Computer Vision and Image Understanding: 35-46, 1994.

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

[7]
Karam. "A Framework for Research and Design of Gesture-Based Human Computer Interactions", PhD Thesis, University of Southampton, 2006.

[8]
Bourke, J O' Brien , G Lyons. "Evaluation of a Threshold-Based Tri-Axial Accelerometer Fall Detection Algorithm", Gait & Posture 26(2):194–199, 2007.

[9]
A Ibraheem, R Z Khan. "Vision Based Gesture Recognition Using Neural Networks Approaches: A Review", International Journal of Human Computer Interaction (IJHCI), Malaysia, Vol. 3(1), 2012.

[10]
Stergiopoulou, N Papamarkos. "Hand Gesture Recognition using a Neural Network Shape Fitting Technique". Elsevier Engineering Applications of Artificial Intelligence 22, 11411158, 2009.

[11]
M Hasan, P K Mishra. "HSV brightness factor matching for gesture recognition system". International Journal of Image Processing (IJIP), vol. 4(5), 2010.

[12] .
Metaxas. "Sign Language and Human Activity Recognition". CVPR Workshop on Gesture Recognition, June 2011.

[13]
Burande, R Tugnayat, N Choudhary. "Advanced Recognition Techniques for Human Computer Interaction", IEEE, Vol 2 pp., 480483, 2010.

X Zabulis, H Baltzakis and A Argyros. "Vision Based Hand Gesture Recognition for HumanComputer Interaction"