International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-2 Number-4 Issue-7 December-2012

Segmentation of One and Two Hand Gesture Recognition using Key Frame Selection

Ketki P. Kshirsagar
Department of Electronics Engineering Ph.D. student of S.G.G.S. Institute and Technology, Nanded
Working in P.D.E.A.College of Engineering, Manjari (BK) Pune, India

Abstract

The sign language recognition is the most popular research area involving computer vision, pattern recognition and image processing. It enhances communication capabilities of the mute person. In this paper, we present an object based key frame selection and skin colour segmentation uniform and no uniform background for one and two hand gesture recognition. Experimental results demonstrate the effectiveness of our proposed scheme for recognizing efficiency 99.50% for static and 98.09% for dynamic one handed American Sign Language and 99.00% for static and 97.03 % for dynamic two handed British sign language.

Keywords

Sign Language, Skin colour segmentation, Video object plane generation.

1. Introduction

Sign language is not universal; it changes from country to country. Every country has its unique interpreter. Recognition of sign language is to provide most important opportunity for deaf community. Sign language provides an opportunity for a mute person to communicate with normal or mute person without any interpreter.

The work on sign language recognition is reported by Starner and Pentland [1], [2], who developed a glove-environment system capable of recognizing a subset of the American Sign Language (ASL). Liang and Ouhyoung [3] used the hidden markov model (HMM) approach for recognition of continuous Taiwanese Sign Language with a vocabulary of 250 signs. Yang and Ahuja [4] used Time-Delay Neural Networks (TDNN) to classify motion patterns of ASL. Bhuyan, et.al. [5] developed recognition of Indian sign language with a vocabulary of 48 signs. For segmentation, Meier and Nagan [6] had developed hausdorff distance algorithm. Though different researchers proposed different methods for sign language recognition, none of these methods are successful to address all the problems encountered in sign language recognition.

In our system, recognition depends upon segmentation of hand in form of shape and contour. In this system Inter-change detection algorithms used for video object plane generation and segmenting the one as well as two hand static images as well as dynamic video sequences.

Rest of the paper is organized as follows. In section 2 we describe proposed segmentation of hand gesture recognition algorithm. Experimental results are given in section 3. Conclusions are given in section 4.

2. Proposed Algorithm

Recognition of sign languages is one of the major concerns for the international deaf community. Sign language recognition is a research area involving pattern recognition, computer vision, natural language processing and psychology. Sign language recognition is a comprehensive problem because of the complexity of the visual analysis of hand gestures and the highly structured nature of sign languages.

Initially hand image is captured by a video camera. The captured image is divided into two categories: method which use data gloves and method which relay on computer vision. Data gloves have sensors attached so that they can transmit electrical signals concerning the movement of hand. The hand portion is extracted from the image by segmentation method. Segmentation included, hand detection, motion analysis and filtering. Once the hand is detected, a complete hand gesture recognition system must be able to extract the hand motion and the spatial position and form ("gesture") of the hand. This information is then used to drive an application. Hand recognition system is explained by figure 1.

Hand gesture recognition used for many applications. It can be used for teleconferencing since it does not require any hardware. It can be applied to the interpretation and learning of sign languages. Recently research in hand gesture recognition aims at applying to sign language recognition, control of household electronic appliances human-computer interaction, and so on.

Hand region is achieved through color segmentation since color is one of the most distinctive clues to find objects and is generally input as RGB information via a CCD or a video camera.

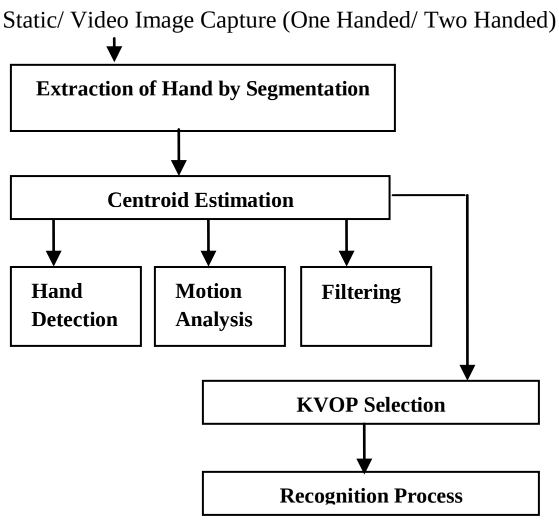


Figure 1: Proposed Segmentation Hand Gesture Recognition System

Colour is a robust feature that performs highly even if the conditions of the environment are not exactly the requisite. First of all, color is invariant to rotation and scaling as well as to geometric variations of the hand. Secondly and importantly it allows a simple and fast processing of the input image. On the other hand, skin color varies quite dramatically. Segmentation done by different color space methods. It is vulnerable to changing lighting conditions and it differs among people. In hand gesture recognition first step is video object plane (VOP) generation.

A. VOP generation

Inter-frame change detection algorithm is used for extracting the VOP using contour mapping. It is one of most feasible solution to object tracking. For removing the background, skin color segmentation is used [7] as shown in equation (1).

In this process, sequence frames are converted into gray scale. Canny operator is used for luminance edge mapping.

$$\begin{bmatrix} (R > 95) \quad \& \quad (G > 40) \quad \& \quad (B > 20) \\ \& \quad (max\{R,G,B\} - min\{R,G,B\} > 15) \\ \& \quad (|R - G| > 15) \quad \& \quad (R > G) \quad \& \quad (R > B) \end{bmatrix} \quad (1)$$

Difference edges (DE_n) is computed between two successive frames using equation (2)

$$DE_n = \phi(|F_{n-1} - F_n|) \quad (2)$$

Moving change edges (ME_n^{change}) are calculated using difference edges (DE_n) and current frame edges (E_n) using equation (3). Moving still edges (ME_n^{still}) are calculated using moving edges of previous moving edges of previous frame (ME_{n-1}) and current frame edges (E_n) using equation (4).

$$ME_n^{change} = \left\{ e \in E_n \,\middle|\, \min_{x \in DE_n} \|e - x\| \leq 1 \right\} \quad (3)$$

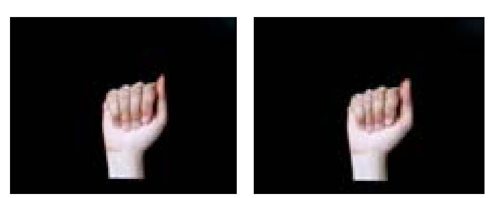
$$ME_n^{still} = \left\{ e \in E_n \,\middle|\, \min_{x \in ME_{n-1}} \|e - x\| \leq 1 \right\} \quad (4)$$

Using ME_n^{change} and ME_n^{still} moving edges (ME_n) are calculated using equation (5).

$$ME_n = ME_n^{change} \cup ME_n^{still} \quad (5)$$

$E_n = \{e_1, e_2, e_3 ... e_n\}$ Current Image edges now extract the VOP using contour mapping [8], [9].

In VOP extraction uses logical-OR operation on horizontal VOP region and vertical VOP region, where horizontal region are declared if they are inside the first and last edge points for each row and vertical region for each column. Then take intersection of both region and false part erased by threshold method for improve the performance of VOP extraction. After false part erasing morphological closing operation is used in post-processing step, and finally extracted VOP [9][10]. Figure 2 shows the results.



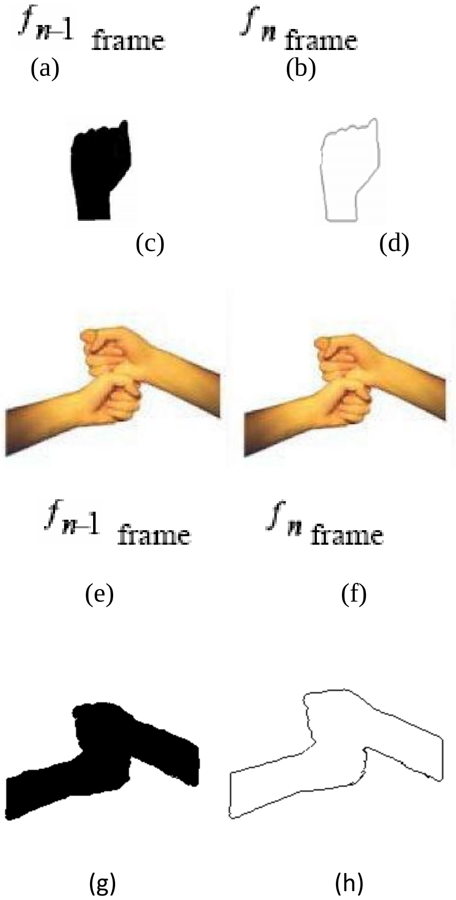


Figure 2: Conture Mapping
(a,b,c,d- One Handed Result)
(e,f,g,h- Two Handed Result)

Centroid of palm region decides local motion or global motion. There is generally no movement of VOP centroids in case of gestures having only local motions, except for very small movement due to hand trembling. On the other hand, in case of gestures having global motions, VOP centroids will move by large amount from one key VOP to the next.

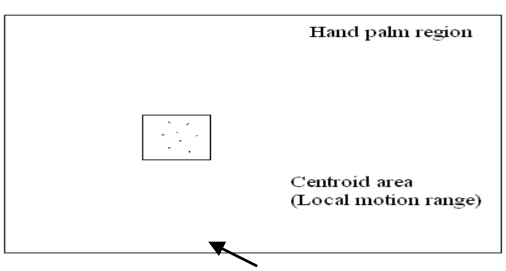


Figure 3: 2D centroid region

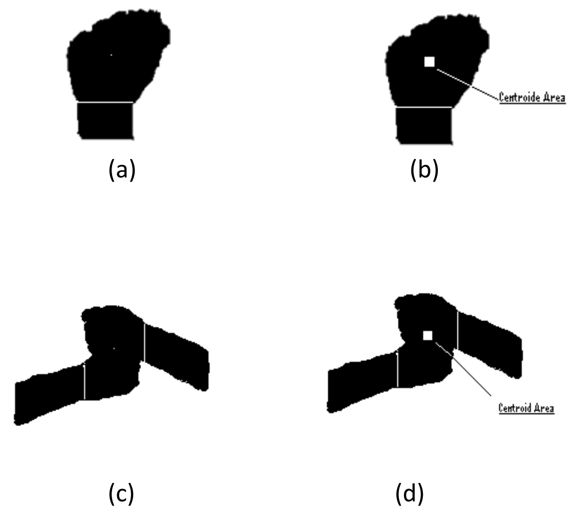


Figure 4: Centroid region for One Handed(a),(b) and Two Handed Images(c),(d)

As shown in Figure 3 and 4, square region indicate local motion range (centroid area). If frame centroid is within the square region it indicates the local motion, and centroid beyond that range indicates global motion. In our method, square region is considered around the centroid of the first key VOP and the square space represents the allowable movement of the VOP centroid. The square region is calculated during frames checking.

That centroid analyse local motion or global motion. After VOP extraction, binary alpha plane is generated. For the key VOP selection, the first VOP of video sequence is considered as the key VOP. If the canny edge difference between two successive frames is greater than the adaptive threshold, then next frame is selected as next key VOP. The threshold calculated using equation (6).

$$T = (Canny\ \ edge\ \ point\, F_n) - (Canny\ \ edge\ \ point\, F_{n-1}) \quad (6)$$

F_n - n^{th} frame $\qquad$ F_{n-1} - $n\text{-}1^{th}$ frame

3. Experimental Result

For recognition we have used two set of databases: first database consist of static alphabet (A to Z) signs and second database consist of dynamic alphabet sequences of American Sign Language for one handed and first database consist of static alphabet (A to Z) signs and second database consist of dynamic alphabet sequences of British Sign Language for Two handed. These signs are available in life print fingurespell library [11][12][13]. For recognition of static hand gesture 6500 alphabet signs (Few static alphabet sign shown in figure 5) and for dynamic hand gesture 130 sequences are used of the five different persons. All the signs are single handed gestures as well as two handed gestures and video sequences. Table I shows the average recognition efficiency of our proposed segmentation algorithm.

Table I: Segmentation algorithm efficiency

	Segmentation One Handed	Segmentation Two Handed
Static	99.50%	99.00%
Dynamic	98.09%	97.03%

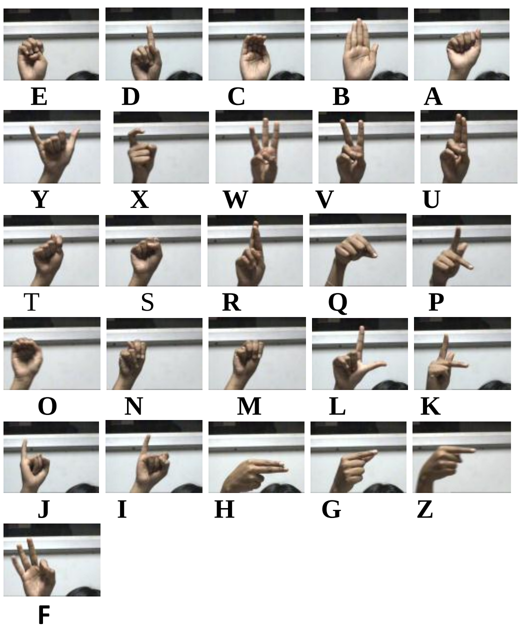


Figure 5: Static alphabet signs (A to Z)

4. Conclusions

The proposed hand gesture segmentation recognition system can used different type gesture signs. In this system key frame selection is done by both the method. Proposed system is suitable for static as well as dynamic hand gesture recognition. Advantage of proposed system is the instead of checking all frames in sequence only key frames are checked. Key frame based gesture recognition is more beneficial for fast recognition.

In future work, we would like to develop complete sign language recognition system by using other body parts i.e., head, arm, facial expression etc. The recognition rate of the proposed system is very much promising for future research in this area.

Acknowledgment

Authors would like to thank AICTE, New Delhi, for providing financial support to carry out this project under RPS sanctioned vide F.No.: 8023/BOR/RPS-67/2006-07, dated 26.02.2007.

References

[1] T. Starner and A. Pentland, "Real-time american sign language recognition from video using hidden markov models", Technical Report, M.I.T Media Laboratory Perceptual Computing Section, Technical Report No.375, 1995.

[2] T. Starner, J. Weaver and A. Pentland, "Real-time american sign language recognition using desk and wearable computer based video", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 12,pp. 1371 - 1375,1998.

[3] R. Liang and M. Ouhyoung, "Real-time continuous gesture recognition system for sign language", Proc. IEEE International Conf: on Automatic Face and Gesture Recognition, 1998.

[4] M.K. Bhuyan, D. Ghosh and P.K. Bora, “Estimation of 2D motion trajectories from video object planes and its application to hand gesture recognition”, Lecture Notes in Computer Science (Springer-Verlag) Pattern Recognition and Machine Intelligence, vol. 3776, pp. 509-514, 2005.

[5] M.K. Bhuyan, D. Ghoah, and P.K. Bora, “A Framework for Hand Gesture Recognition with Applications to Sign Language”, Annual Indian conference, IEEE-2006.

[6] Thomas Meier and King N. Ngan, “Automatic video sequence segmentation using object traking”, Speech and image technologies for computing and telecommunication. TECON IEEE-1997.

[7] Hebert Luchetti Ribeiro, Adilson Gonzaga, “Hand image segmentation in video sequence by GMM:a comparative analysis” SIBGRAPI 2006: 357-364.

[8] Changick Kim and Jenq-Neng Hwang, “A fast and robust moving object segmentation in video sequence”, IEEE International Conference on Image Processing 2004.

[9] R. Mech and M. Wollborn, “A noise robust method for segmentation of moving objects in video sequences”, ICASSP97**,** Vo1.4, pp.2657-2660, April 1997.

[10] Juan Wuchs, Helmun Stern, Yael Edun," Parameter search for an image processing fuzzy cmeansh and gesture recognition system" , 0-7803-7750-8/03/$I7.00 02003 IEEE .
[11] www.lifeprint.com.
[12] www.british-sign.co.uk/fingerspelling_alphabet.php.
[13] www.deafblind.com.

Prof. Ketki P. Kshirsagar completed her BE in 2005 from WIT Solapur and MTech 2008 from SGGS institute nanded Write now as Ph.D.student in SGGS institute Nanded. She is working in PDEA's College of Engineering, Manjari BK, and Pune. She was published papers at ICDIP2009 (07-08 March 2009) International Conference on Digital Image Processing Bangkok, Thailand, ARTCOM 2010 (15-16 Sept 2010) International conference on advances in recent technologies in communication & computing, Kottyam, Kerala and 2012 4th International Conference on Electronics Computer Technology (ICECT 2012), April 6 - 8, 2012, Kanyakumari, India.

.