ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 1, Issue 7, October 2013

K One Hand Gesture Recognition

Ketki P.Kshirsagar, R.M.Sahu, S.M.Bankar, R.K.Moje

Asst. Professor of Electronics Department of P.D.E.A's College of Engg.,Manjari(BK),Pune, India.

Dr. D.D.Doye

Professor Electronics Department of SGGS Institute & Technology, Nanded, Maharashtra, India

Abstract - Sign language is the most natural and expressive way for the hearing impaired. Its most appealing application is the development of more effective and friendly interfaces for human-machine interaction. Gestures are a natural and powerful way of communication. A hand gesture recognition system can provide an opportunity for a mute person to communicate with normal people without the need of an interpreter.

We proposed use to the finite state machine with key frame selection facility and gesture trajectory features for one hand gesture recognition. Experimental results demonstrate the effectiveness of our proposed scheme for recognizing One Handed American Sign Language.

Index terms – Sign language, Finite state machine and Trajectory feature for one hand.

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[12][13]. 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 hand shape, hand position, static trajectory features and dynamic trajectory features. The basic recognition system diagram is as shown figures 1.These are main attributes of hand gesture recognition. In this system, finite state machine (FSM) algorithms are used for one as well as two hand gesture recognition. FSM based recognition (with trajectory feature and without feature) procedure includes key frame selection and shape similarity measured by Hausdorff distance (test frame sequence and standard frames) used recognition procedure includes key frame selection, pixel-to-pixel distance measured by lowest weight distance.

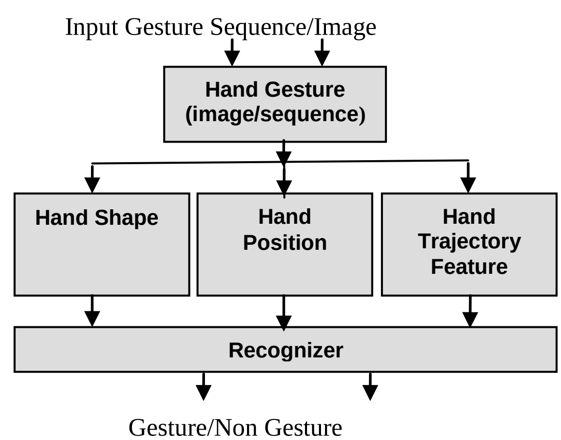


Figure 1: Basic recognition system

The main contributions in this paper are summarized as follows:

Firstly, we propose novel technique for one hand gesture recognition to select key frames using FSM (with trajectory features and without feature) method. The advantage of proposed method is that shape and pixel distance measurement for key frames are required instead of all frames of video sequences. The key frame based gesture representation is equally useful for quick gesture recognition and coding of video frames in compressed domain. Detection co-articulation or non gesturing phase is possible by extracting the key frames in video sequence. Secondly to validate our result we have used two set of databases for one and two hand recognition: first database consist of static alphabet (A to Z) signs, one handed of American Sign Language and second database consist of dynamic alphabet sequences one handed of American Sign Language. The result of proposed method found outstanding with compared to other existing methods.

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 1, Issue 7, October 2013

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

2. PROPOSED HAND GESTURE RECOGNITION SYSTEM

The proposed hand gesture recognition method is shown in figure 2.where the input is the video sequence or static image. The first step is converted video sequence into frame format (any size of frame format). Backgrounds in the used video sequences are uniform and non uniform. In hand gesture recognition first step is video object plane (VOP) generation.

2.1 VOP (Video Object Plane) 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 or static image are converted into gray scale. Canny operator is used for luminance edge mapping.

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

Difference edges (DE_n) is computed between two successive frames using equation that is the inter-frame change detection algorithm (2)

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

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

Input Video Sequence/Image

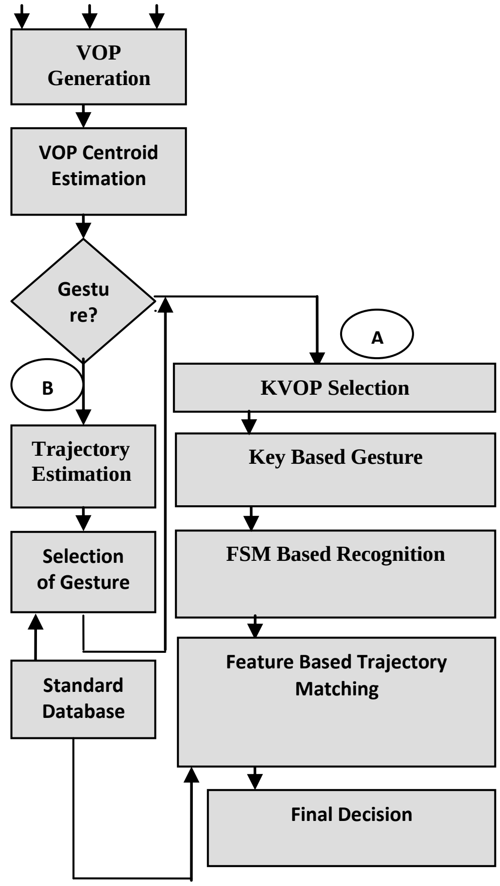


Figure 2: Proposed hand gesture recognition system
A -------- Local motion only.
B -------- Global only or local and global motion

$$ME_n^{change} = \left\{ e \in E_n \,\middle|\, \min_{x \in DE_n} \|e - x\| \le 1 \right\} \tag{3}$$

$$ME_n^{still} = \left\{ e \in E_n \,\middle|\, \min_{x \in ME_{n-1}} \|e - x\| \le 1 \right\} \tag{4}$$

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

$$ME_n = ME_n^{change} \cup ME_n^{still} \tag{5}$$

$E_n = \{e_1, e_2, e_3 \ldots e_n\}$ Current frame edges

Now extract the VOP using contour mapping [8], [9] in figure 3.

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 1, Issue 7, October 2013

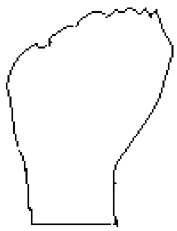
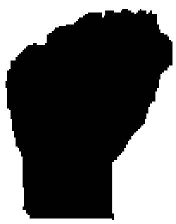


Image (Alphbet S) Contour



Binary Alpha Plane
Figure 3: VOP Generated

Centroid of palm region decides local motion (static) or global motion (dynamic) [5]. 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. VOP range is shown in figure 4.

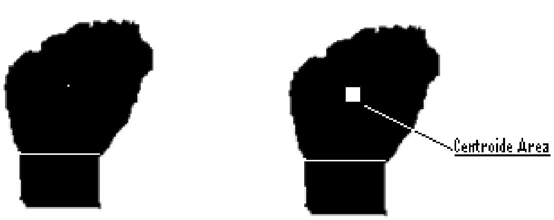


Figure 4: centroid region

As shown in Figure 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 and selects the object key frames as well as binary alpha plans that are the segmentation.

2.2 Local motion

Finite state machine and Dynamic time warping methods are used for the local motion detection.

2.2.1 Finite State Machine

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).

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

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

Hausdorff distance is used to measure the similarity between the test frames sequence and standard frames using equation (7).

Contour points (randomly selected) of Standard frame $\left(S = \{s_1, s_2, s_3, \dots s_n\}\right)$ and Test frame $\left(T = \{t_1, t_2, t_3, \dots, t_m\}\right)$.

$$\begin{aligned} H(S,T) &= \max\{h(S,T), h(T,S)\} \\ \textit{where} & \\ h(S,T) &= \max s \in S \min t \in T \|s-t\| \\ \textit{and} & \\ h(T,S) &= \max_{t \in T} \min_{s \in S} \|t-s\| \end{aligned} \qquad (7)$$

In FSM based recognition every key frame similarity is checked by hausdorff distance and key frame duration criteria [10], [11]. If it satisfies both criteria, then recognition is successful otherwise it is co-articulation or non gesture phase.

Next important step for FSM method in one hand as well as two hand gesture recognitions is the selection of suitable features. Selecting good feature is crucial for gesture recognition, because it is totally depend on their shape and motion. For trajectory matching consider both static and dynamic features. Dynamic features used for global motion only. Static features are low level features and dynamic features are high level feature. Static features correspond to shape of hand trajectory and dynamic features correspond to motion of hand trajectory [15]. For gesture recognition both static

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING
Vol. 1, Issue 7, October 2013

and dynamic features are equally important. Key trajectory point selection is the features used for recognition.

Key point selection is merging of adjacent approximation interval of the estimated trajectory. These key points best represent the prominent locations of the hand in the gesture trajectory. Trajectory length calculated using equation (12).

$$D = \sum \left\{ \left(S_i - S_{i+1} \right)^2 + \left(T_i - T_{i+1} \right)^2 \right\}^{1/2} \qquad (12)$$

2.3 Global motion

Palm region centroid decided the global motion [5]. In global motion most important part is the trajectory estimation. For trajectory estimation first find out the centroid of two successive frames and these are the C_i and C_{i-1}. In global motion where centroid C_i is obtained by translating C_{i-1} by respective motion vector in figure 6. The final centroid is taken as the average of these two this nullifed the effect of slight shape changes in successive VOPs [5]. After nullifying effect it behaves as local motion and follows the same path for recognition.

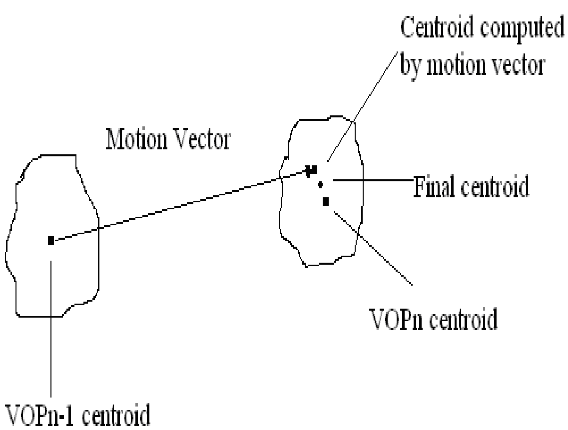


Figure 6: Estimate centroid in VOPs

3. EXPERIMENTAL RESULT

For recognition we have used two set of databases: first database consist of static alphabet (A to Z) signs, one handed of American Sign Language and two handed of British Sign Language and second database consist of dynamic alphabet sequences one handed of American Sign Language [16] . For recognition of one handed static hand gesture 6500 alphabet signs (Few static alphabet sign shown in figure 7) and for dynamic hand gesture 130 sequences are used of the five different persons(Result is shown in Table I and II). All the signs are one handed. Table III and IV show the average recognition efficiency of our proposed sign language recognizer.

Table I
Gesture Recognition efficiency in % (without feature)

	Method	Persons	FSM
One Handed	Static	1	68.23
		2	82.77
		3	63.08
		4	60.85
		5	55.85
	Dynamic	1	63.84
		2	62.15
		3	60.61
		4	60.07
		5	53.30

Table II
Gesture Recognition efficiency in % (with feature)

	Method	Persons	FSM
One Handed	Static	1	70.75
		2	68.07
		3	67.61
		4	63.92
		5	56.36
	Dynamic	1	66.38
		2	60.84
		3	64.38
		4	77.15
		5	59.76

Table III
Average Gesture Recognition efficiency (without feature)

	Method	**FSM**
One Handed	**Static**	**62.49%**
	Dynamic	**61.330%**

Table IV
Average Gesture Recognition efficiency (with feature)

	Method	FSM
One Handed	Static	65.34%
	Dynamic	65.702%

Result of proposed FSM (with feature) method is compared with existing FSM (without feature). The recognition efficiency of one handed has been improved up to 65.34% for static database and 65.702% for dynamic database with proposed method

4. CONCLUSIONS

The proposed hand gesture recognition system can used different type gesture signs. In this system key frame selection is done by three methods. Proposed system is suitable for static as well as dynamic one 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. And trajectory features improve recognition efficiency.

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.

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 roboust moving object segmentation in video sequence", IEEE International Conference on Image Processing 2004.

[9] Changick Kim and Jenq-Neng Hwang, "Fast and Automatic video object segmentation and tracking for content –based application", IEEE transactions on circuits and systems for video technology, VOL.12, NO. 2, Feb 2002.

[10] M.K. Bhuyan, D. Ghosh and P.K. Bora, "Finite state representation of hand gesture using key video object plane", Proc. IEEE Region 10 Asia-Pacific Conference, vol. A, pp. 579-582, 2004.

[11] M.K. Bhuyan, D. Ghosh and P.K. Bora, "Key video object plane selection by MPEG-7 visual shape descriptor for summarization and recognition of hand gestures", Proc. Fourth Indian Conference on Computer Vision, Graphics and Image Processing , pp. 638-643, 2004

[12] Vinod Nair and James clark (2002), "Automatic video surveillance using hidden markov model", Proc. Of 15th vision interface conference, pp. 88-92.

[13] Lawrence Rabiner, "A tutorial on hidden markov model and selected application in speech recognition", Proc. IEEE, VOL. 77, No. 2, Feb 1989.

[14] D.D.Doye ,T.R. Sontakke and Smita Nagtode, "The nonlinear time alignment model for speech recognition system", IETE Journal of research, Vol 47, No. 6 pp271-275, Nov-Dec 2001.

[15] M.K. Bhuyan, D. Ghosh and P.K. Bora, "Feature extraction from 2D gesture trajectory in dynamic hand gesture recognition", cybernetics and intelligent systems, IEEE-2006.

[16] www.lifeprint.com.

[17] www.deafblind.com.

[18] www.british-sign.co.uk/fingerspelling_alphabet.php.