

TRACKING OF HAND MOVEMENTS IN GESTURE RECOGNITION A REVIEW ON TECHNIQUES AND APPROACHES

Prerna Gupta

Computer Science and Engineering
NITTTR
Chandigarh

Prerna14gupta@gmail.com

Garima Joshi

Electronics and Communications
UIET
Chandigarh

Joshi_garima5@yahoo.com

Dr. Maitreyee Dutta

Computer Science & Engineering
NITTTR,
Chandigarh

d_maitreyee@yahoo.co.in

Abstract— *Gesture recognition is a vast topic of research in computer science. It deals with creating a system that can identify human gestures and use them for different applications. In gesture recognition, a camera is used to capture human body's motions which are sent to the computer for interpreting human body language. One of the major application areas of gesture recognition is automatic sign language recognition.*

In sign language very motion of human body parts like hands and head are very important to consider. Whenever trajectories and movements are considered, tracking plays a vital role and is essential too. Tracking helps in recognizing the motion trajectories. This paper presents literature review on research and findings by researchers on different techniques and approaches of tracking.

Keywords— *Review, Tracking, Sign Language, Mean shift, Kalman Filter, Condensation, Particle Filter*

I. INTRODUCTION

In Human Computer Interactions(HCI), hand gestures are considered a natural and convenient way as communication is done largely through hands/arm. However, facial expression, body postures, head and torso movements are also included in HCI but researcher Karam(2006) in his work reported that hand has major contribution in gesturing as compared to other body parts.

One of the major application areas of gesture communication is SIGN LANGUAGE. It varies from country to country but all have some common feature such as people use continuous motion of hands to form a gesture. However, capturing and tracking these hand motions is very challenging task.

Therefore, some approaches are used for HCI

- 1) Gloved Based method that uses some special sensors to directly measure the position and movement of hand of a signer who is wearing that glove but these gloves incur cost and are cumbersome.

- 2) Vision Based method can be either marker based that require reflective markers to be placed on target to ease tracking or markerless which more flexible and useful

Since, sign language is a boon to deaf people. The authors try to build a hand gesture recognition system for Indian Sign Language. Here, in this paper our focus is on gesture through hands only.

II. PREVIOUS WORK

In recent years large number of research has been done on tracking of hand motions in a sign language.

Ying Wu et al. [2005] presented an approach for tracking of hand in image sequences using Monte Carlo algorithm. For their convenience, they used divide and conquer approach for capturing hand poses and finger articulations. The tracking algorithm is based on sampling technique. The important part of sequential Monte Carlo is to generate samples at time(t+1) from samples at time (t) [8].Rezaei et al.(2008) also gave a novel algorithm for estimation of motion trajectory and orientation of hand. The algorithm presented by them is based on basic principle of control point extraction to calculate motion parameters [9].

Elmezain et al. [2009] proposed meanshift analysis and Kalman filter method for hand tracking in a complex environment using depth information. They introduced robust method of efficient tracking using 3D depth map as depth information solves the problem of overlapping between face and hands. Various video sequences with various hand shapes are tested. The images are taken by bumblebee camera system with 25 fps and 240X320 pixel resolutions. The meanshift iterations are applied

on both hands and result shows the proposed work to be robust for online tracking [5].

Akmeiliawati et al.[2009] introduced marker less hand gesture recognition for New Zealand sign language. Firstly they gave an introduction of that language and compared gloved based techniques as well as vision based techniques. The steps that they considered are detection of desired object (head and hands). Second is the segmentation and tracking using different color spaces RGB, ICrCb, HASV. Tracking algorithm implemented is mean shift. The novel idea introduced the work of image segmentation by finding peaks in image's histogram. Finally gesture modeling and interpretation is done. 13 gestures and one unknown gesture category is used in dataset which is limited [10].

Patwardhan and Roy [2007] designed the novel tracker capable of tracking objects based on position, shape, size and appearance. They worked for designing the tracker used to track gestures to do a particular task (for controlling an audio player). Algorithm is improvised predictive Eigen tracker that used particle filtering. In the experiments with representative gesture set 4 basic hand shapes, 64 gestures for training set and 16 additional are used and found that their proposed work is robust appearance based visual tracker [17].

Allen et al. [2006] extended the camshaft algorithm in combination with type of features. Comparison of effectiveness with other general purpose object tracking is done. Camshift is used by the authors because according to them it consumes lowest number of CPU cycles assuming only a single hue in colour space model. Weighted histograms is used. A synthetic video is taken as input that includes occlusion as well as orientation. The result shows that camshaft based tracker fails if weighted histograms are not used while if it is used then some noise such as snow and sleigh colors can be eliminated from target histograms [7].

Binh et al. [2005] introduced the vision based hand gesture recognition system. They employed the technique of Pseudo two dimension hidden markov models. For tracking Kalman filter is used and colour space model for detection. The experiment consists of 36 gestures of American Sign Language. Various papers and their approaches have been discussed. Tests are done at about 25 frames per second (fps) and for single hand recognition accuracy was found to be 98% [4].

Ongkittikul et al.[2008] employed the two hand recognition using skin color feature with particle filtering. As per the authors, they have proposed the system which can undergo rotations, fast movements and rigid deformations. The novel method is used for tracking in clutter background for cash machine simulation. YCbCr color space is used in the work. The K means technique is considered for hand merge and split. 10 signers with 10 gestures that pose as commands of cash machine are used. More than 45 videos at 20fps are tested and result indicates that proposed scheme deals with fast motion till the end and works for 83% of all videos [2].

Fahn et al. [2007] presented their work using particle filter technique. They considered particle filter to be a flexible simulation based method which is suited for non linear tracking. The novel approach designed by them includes efficient and robust hand motion tracking technique. Since, they used fewer particles. So, it reduces the computational cost and improves the system performance. As discussed in above mentioned works, hand is detected through skin colour model for which HSV colour model was used. After detection and segmentation connected component labelling method is applied for obtaining more accurate results.

Gupta et al. Designed a novel framework for tracking hand gestures when position and view of hand changes. Eigen tracking is used for such an application. They introduced predictive framework with condensation algorithm that according to them is absent in all related work with condensation algorithm. An eigen tracking simply updates the eigenspace and affine coefficients while predictive framework helps to speed up the tracking process [1].

III. OVERVIEW OF HAND GESTURE TRACKING TECHNIQUES

In today's scenario of sign language the tracking of movement is a crucial task as most of the gestures include movements. Different techniques have been proposed by many researchers. Some have used the traditional techniques and some have improvised the traditional techniques and formed the new one. As discussed in literature survey section, some of the techniques are discussed below.

A. Mean shift

Meanshift is a non-parametric feature space analysis technique for locating the maxima of a density function, a so-called mode seeking algorithm[11]. Application domains include cluster analysis in computer vision and image processing[12].

The idea originally came in 1975 by Fukunaga and Hostetler [13]. The camshift algorithm is based on the mean shift algorithm. The mean shift algorithm works well on static probability distributions [14] but not on dynamic ones as in a movie. Camshift is based on principles of the mean shift but also a facet to account for these dynamically changing distributions.

Camshift is able to handle dynamic distributions by readjusting the search window size for the next frame based on the current frames distribution. This allows the algorithm to anticipate object movement to quickly track the object in the next scene. Even during quick movements of an object, Camshift is still able to correctly track. The camshift algorithm is a variation of the mean shift algorithm. Camshift works by tracking the hue of an object, in this case, flesh color. The movie frames were all converted to HSV space before individual analysis.

B. Kalman Filter

The filter is named after Rudolf Emil Kalman and is a very powerful idea for controlling noisy systems. In words of G. Welch and G. Bishop it is an efficient tool that estimates states of a process in terms of mathematical equations. It can estimate past, present and future states of a process. According to a study the achievement of Apollo 8 (first human spaceflight from earth to an orbit around moon) was impossible without Kalman filter. Apart from tracking objects, Kalman filter has various application areas such as navigation, depth measurement, computer vision applications, economics and much more.

C. Particle Filter

It is utilized to track hand gestures in dense cluttered environment. The location of hand is modelled with set of particles. The disadvantage is many particles are required even for hands. So, for high dimensional models or other body parts large amount of particles are needed for representation. Hence researchers are working on reducing number of particles and modification of this algorithm is

1) *Sequential Monte Carlo*: This simulation based method is used for computing that calculates posterior density estimations. This method does not assume dynamics of state space (can be non linear also). The samples from distribution are represented through set of particles having weights. Weights assigned to particles represent probability of particles being sampled [15].

2) *ConDensation (Conditional Density Propagation)*: The algorithm is to track edges or contours of objects. It solves the problem of finding pixels that make up the contours of an object. It is a probability based algorithm. According to Isard and Blake (1998) who introduced this algorithm, it is an application of particle filter estimation technique which allows general representation of probabilities.

D. Eigen value based

Eigen values or vectors are part of linear transformations. Operations like stretching, compressing, flipping etc are used in this. Eigen vectors of covariance are found out. It denotes variation in data. More Eigen vector means better information obtained from linear transformations [16].

IV. RESEARCH METHODOLOGY

Current phase of our development also deals with image processing in which four basic steps are followed that are shown in Fig. 1

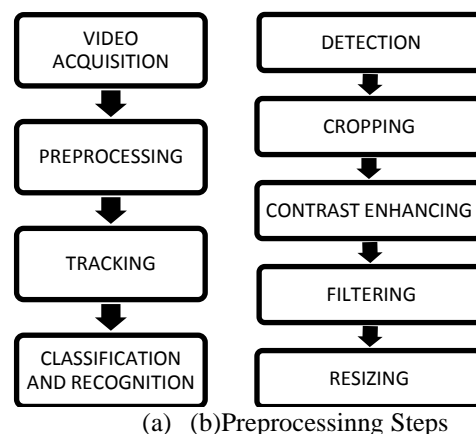


Fig.1 (a) showing basic steps of recognition system

Thus, review on tracking methods will contribute to the idea of how we are going to track the hand in image sequences and recognize it.

V. DISCUSSIONS

Most important factor in tracking is accuracy. A good tracking algorithm should have most accurate results which can be measured by calculating in how many frames the region of interest is tracked properly. Also time of tracking, speed and system performance should be taken in consideration.

Some assumptions are also made in signer's location and background as image sequences are captured by a Logitech digital camera.

VI. CONCLUSION

In this paper, some work has been discussed related to hand gesture tracking and recognition. Generally in all related work skin colour modelling is used for detection of hand region and then some processing is done. For tracking part various techniques are applied.

Orientation and position of hands are other different areas that can be touched in future.

REFERENCES

- [1]N. Gupta, P. Mittal, S. Dutta Roy, S.Chaudhury, and S. Banerjee, "Developing a gesture-based interface," *Journal of the Institution of Electronics and Telecommunication Engineers* 48, Vol. 3 ,2002.
- [2]S. Ongkittikul, S. Worrall, & A. Konoz, "Two hand tracking using colour statistical model with the K-means embedded particle filter for hand gesture recognition," in *Proc. CISIM'08. 7th, 2008* , IEEE , pp. 201-206.
- [3] Toni Benussi, Darko Jurić, "A Robust Hand Detection and Tracking Algorithm with Application to Natural User Interface," in *MIPRO, Proceedings of the 35th International Convention*, IEEE , 2012., pp. 1768-1774.
- [4]N.D. Binh, E. Shuichi, T. Ejima, "Real-Time Hand Tracking and Gesture Recognition System," in *GVIP 05 Conference, CICC, Cairo, Egypt, 2005*.
- [5] Elmezain, Mahmoud, A. Al-Hamadi, R. Niese, and B. Michaelis, "A robust method for hand tracking using mean-shift algorithm and kalman filter in stereo color image sequences," in *WASET*, 2010, pp. 131-135.
- [6] S. S. Rautaray, A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey", in *Springer Science+Business Media Dordrecht , Artificial Intelligence Review* , 2012, pp. 1-54
- [7] Allen, GJohn., Y. X. Richard, and S. Jin Jesse, "Object tracking using camshift algorithm and multiple quantized feature spaces." in *Proc. of the Pan-Sydney area workshop on Visual information processing, Australian Computer Society, Inc., 2004*, pp. 3-7.
- [8]Wu, Ying, J. Lin, and T. S. Huang, "Analyzing and capturing articulated hand motion in image sequences," in *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 27.12 (2005), pp. 1910-1922.
- [9] Rezaei, Ahmadreza, M. Vafadoost, S. Rezaei, and Y. Shekofteh, "A feature based method for tracking the 3-D trajectory and the orientation of a signer's hand," in *Computer and Communication Engineering, 2008. ICCCE* pp. 346-351.
- [10]Akmeliawati, R., F. Dadgostar, S. Demidenko, N. Gamage, Y. C. Kuang, C. Messom, M. Ooi, A. Sarrafzadeh, and G. SenGupta, "Towards real-time sign language analysis via markerless gesture tracking," in *Instrumentation and Measurement Technology Conference, 2009. I2MTC'09. IEEE*, pp. 1200-1204.
- [11] Cheng, Yizong, "Mean Shift, Mode Seeking, and Clustering," in *Transactions on Pattern Analysis and Machine Intelligence (IEEE)* , August 1995, pp. 790-799.
- [12]Comaniciu, Dorin, P.Meer , "Mean Shift: A Robust Approach Toward Feature Space Analysis," in *Transactions on Pattern Analysis and Machine Intelligence, IEEE*, May 2002, pp. 603-619.
- [13]Fukunaga, Keinosuke; Larry D. Hostetler "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition," in *Transactions on Information Theory (IEEE)*, (January 1975)., pp. 32-40.
- [14] Feng, Yingxue, J. Yang, P. Yi, and P. Guo. "Video Based Tracking & Recognition via (Wireless) Network Transmission," 2007
- [15] Doucet, A, Johansen, A.M., "A tutorial on particle filtering and smoothing: fifteen years later," Technical report (Department of Statistics, University of British Columbia), December 2008.
- [16]J. Singha, and K. Das. "Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique," *International Journal of Advanced Computer Science and Applications*, Vol. 4, No. 2, 2013.
- [17] K.S.Patwardhan and S.D. Roy, "Modelling and recognising spatio-temporal hand gestures with an uncalibrated camera," in *Proc. First international conference on Expressive Interactions for Sustainability and Empowerment, British Computer Society, 2009*, pp. 6.): 237-244.

