

Tracking of an Object in Video Stream Using a Hybrid PSO-FCM and Pattern Matching

Sai Krishna Borra ¹ ; Suman Krishna Chaparala ²

1: K L University, Vijayawada,

2: K L University, Vijayawada,

Abstract— In this paper we propose a novel method for object tracking in video images. The method is based on image segmentation and pattern matching. All moving and still objects in video images can be detected accurately with the help of efficient image segmentation techniques. We propose a hybrid algorithm for image segmentation using the notion of Particle Swarm Optimization (PSO) and Fuzzy-C-Means (FCM) clustering techniques. The results obtained using segmentation of successive frames are exploited for pattern matching in a simple feature space. As a consequence, multiple moving and still objects in video images are tracked simultaneously. We perform simulation experiments on object tracking to validate the efficiency of our proposed algorithm. The algorithm outperforms the existing algorithm in context of accuracy and time complexity.

Keywords -Image segmentation, object tracking, pattern matching, motion estimation, Particle Swarm optimization, Fuzzy-C-Means clustering.

I. INTRODUCTION

Tracking is a significant and difficult problem that arouses interest among computer vision researchers. The objective of tracking is to establish correspondence of objects and object parts between consecutive frames of video. It is a significant task in most of the surveillance applications since it provides cohesive temporal data about moving objects which are used both to enhance lower level processing such as motion segmentation and to enable higher level data extraction such as activity analysis and behavior recognition.

There are two common approaches in tracking objects as a whole: one is based on correspondence matching and other one carries out explicit tracking by making use of position prediction or motion estimation. On the other hand, the methods that track parts of objects (generally humans) employ model-based schemes to locate and track body parts which allows you to see the footnotes. Moving object tracking is the process of locating a moving object in time using a camera. An

algorithm analyses the video frames and outputs the location of moving targets within the video frame. The main difficulty in video tracking is to associate target locations in consecutive video frames, especially when the objects are moving fast relative to the frame rate. Here, video tracking systems usually employ a motion model which describes how the image of the target might change for different possible motions of the object to track.

Examples of simple motion models are:

- To track planar objects, the motion model is a 2D transformation (affine transformation or homograph) of an image of the object.
- When the target is a rigid 3D object, the motion model defines its aspect depending on its 3D position and orientation
- For video compression, key frames are divided into macro blocks. The motion model is a disruption of a key frame, where each macro block is translated by a motion vector given by the motion parameters
- The image of deformable objects can be covered with a mesh, the motion of the object is defined by the position of the nodes of the mesh.

The role of the tracking algorithm is to analyze the video frames in order to estimate the motion parameters. These parameters characterize the location of the target.

II. METHODS OF APPROACH

A. Conventional Approach

There are two major components of a visual tracking system; Target Representation and Localization and Filtering and Data Association. Target Representation and Localization is mostly a bottom-up process. Typically the computational complexity for these algorithms is low. The following are some common Target Representation and Localization algorithm.

Blob tracking: Segmentation of object interior (blob detection, block-based correlation)

Kernel-based tracking (Mean-shift tracking): An iterative localization procedure based on the maximization of a similarity measure (Bhattacharyya coefficient).

Contour tracking: This is used for the Detection of object boundary.

Filtering and Data Association is mostly a top-down process, which involves incorporating prior information about the scene or object, dealing with object dynamics, and evaluation of different hypotheses. The computational complexity for these algorithms is usually much higher. The following are some common Filtering and Data Association algorithms:

Kalman filter: An optimal recursive Bayesian filter for linear functions subjected to Gaussian noise.

Particle filter : Useful for sampling the underlying state-space distribution of non-linear and non-Gaussian processes.

Object tracking is an important task within the field of computer vision. The proliferation of high-powered computers and the increasing need for automated video analysis has generated a great deal of interest in object tracking algorithms. In general it is used mostly in surveillance systems. There are three key steps in video surveillance: detection of interesting moving objects , tracking of such objects from frame to frame, and analysis of object tracks to recognize their behavior.

In this report we present a feature based object tracker which uses a pan-tilt (PT) camera to keep track of the target. Generally, the task is to keep the target at the center of the grabbed image. As the target starts moving in the real world, its position in the grabbed image is reported in subsequent frames through a feature based tracking algorithm, based on the work of Lukas, Kanade and Tomasi. The image position error is processed by a proportional-integral controller and the camera is re-positioned accordingly to place the target in the pre-specified image region. The task of video-object segmentation is to identify and separate the important objects in a video scene from the scene background. Clearly, to approach this problem, it is necessary to define what is exactly meant with important objects and how the correct object masks should look like. However, in practice, it turns out that even an unambiguous definition of video objects is a fundamental problem. In the following, the involved definition problems are addressed and grouped into physical problems, being a consequence of the image formation, and semantic problems. The physical problems involved are reflections, occlusions, translucent objects, foreground objects, small background objects, object status change.

B. Proposed Approach

Image segmentation and Pattern matching

The frames extracted from the video are segmented first, features of each object in the segmented image are extracted, pattern matching is done on the consecutive frames having the desired features in the hand, the motion vectors are calculated and mask is moved accordingly. The three techniques are:

- Histogram based segmentation and pattern matching .
- Otsu's global thresholding and pattern matching.
- Fuzzy C means clustering using Particle swarm optimization and pattern matching.

In this paper, we propose a novel method for object tracking. It consists of three stages,

- Image segmentation using hybrid algorithm exploiting the notion of Particle Swarm Optimization and Fuzzy-C-Means clustering techniques,
- Feature extraction for pattern matching,
- Motion vector determination and object tracking.

The segmentation is done using Fuzzy-C-Means clustering technique. The motion of PSO is exploited for assigning each pixel to a cluster. The efficiency of both global optimization techniques are hybridized here. each pixel to a cluster. The efficiency of both global optimization techniques are hybridized here. The segmented images of consecutive frames are used for pattern matching after feature extraction.

The motion of the object from



Input Image Segmented Image

Particle Swarm Optimization

Particle Swarm Optimization (PSO) algorithm is a kind of evolutionary computational technique developed by Kennedy and Eberhart in 1995. Like other evolutionary techniques, PSO also uses a population of potential solutions to search the explore space. In PSO, the population dynamics resembles the movement of a "birds' flock" searching for food, while social sharing of information takes place and individuals can gain from the discoveries and previous experience from all other companions. Thus, the companion (called particle) in the population (called swarm) is assumed to "fly" over the search space in order to find promising region of the landscape. Let, particle i of the swarm is represented by the D dimensional vector $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$ and the best particle of the swarm, is denoted by the index g . The best previous position of particle i is recorded and represented as $p_i = (p_{i1}, p_{i2}, \dots, p_{id})$. The position change (velocity) of particle i is $V_i = (V_{i1}, V_{i2}, \dots, V_{id})$. Particles update their velocity and position through tracing two kinds of 'best' value. One is its personal best (p best), which is the location of its highest fitness value. Another is the global best (g best), which is the location of overall best value, obtained by any particles population. Particles update their positions and velocities according to equations:

$$v_{id} = \chi(v_{id}(t) + c_1\phi_1(\rho_{id}(t) - \chi_{id}(t)) + c_2\phi_2(\rho_{gd}(t) - \chi_{id}(t)))$$

$$\chi_{id}(t+1) = \chi_{id}(t) + v_{id}(t+1)$$

Where, $V_{id}(t)$ is the velocity of the d th dimension of the i th particle in the t th iteration, $\chi_{id}(t)$ is the corresponding position, $P_{id}(t)$ and $P_{gd}(t)$ are the corresponding personal best and global best respectively, the variables w is the inertia weight, the variables ϕ_1 and ϕ_2 are the accelerate parameters.

Fuzzy C-Means Clustering

Fuzzy C Means (FCM) is one of the most commonly used fuzzy clustering techniques for different degree estimation problems. It provides a method that shows how to group data points that populate some multidimensional space into a specific number of different clusters which must be known a priori. FCM employs fuzzy partitioning such that a data point can belong to several groups with the degree of membership matrix U is constructed of elements that have value between 0 and 1. The aim of FCM is to find cluster centers that minimize a dissimilarity function.

Proposed concept for object tracking

The hybrid clustering approach to image segmentation starts by choosing the number of clusters and a random initial cluster center for each cluster. PSO plays its part in assigning each pixel to a cluster. It is done according to a probability which is inversely dependent to the distance (similarity) between the pixel and cluster centers.

Image Segmentation using hybrid PSO-FCM Clustering Technique

In fuzzy clustering, a single particle represents a cluster center vector. Let V_i is the D -dimensional vector of i th cluster center and can be represented as $\{V_{i1}, V_{i2}, \dots, V_{iD}\}$. Each point or data vector belongs to ever cluster by different membership function, thus a fuzzy membership is assigned to each data vector. Each cluster has a cluster center and each iteration presents a solution giving a vector of cluster centers. We determine the position of vector V_l for every particle and update it. We then change the position of cluster centers based on these particles. For the purpose of this algorithm, the fitness of particles is easily measured as follows:

$$\rho_l^{(t)} = \sum_{k=1}^n \sum_{i=1}^n [A_i(\chi_k)]^m \|\chi_k - v_i^{(t)}\|^2$$

where n is the no. of data vector, C is no. of cluster centers, $V_l(t)$ is the position of particle l in stage t , $V_{ll}(t)$ is the velocity of particle l in stage t , X_k is the vector of data and $k=1,2,\dots,n$, $\rho_l(t)$ best position funded by particle l in stage t , $\rho_g(t)$ is the best position funded by all particles in stage t , $p(t)$ is fuzzy pseudo partition in stage t and $A_i(t)(X_k)$ is the Membership function of data k vector in stage t into cluster

Proposed Segmentation Algorithm

Step1: Let $t=0$, select initial parameters initial cluster centers c , initial position of particle X , initial velocity of particles w , a real number m , a small positive number ξ , and stopping criterion.

Step2: Calculate the $A_i(t)(X_k)$ for all particles as follows where $i = 1, 2, \dots, C$; $k = 1, 2, \dots, n$.

$$A_i^{(t+1)}(\chi_k) = \left[\sum_{j=1}^c \frac{\|\chi_k - v_i^{(t)}\|^2}{\|\chi_k - v_j^{(t)}\|^2} \right]^{m-1} \Big]^{-1}$$

Step3: For each particle calculate the fitness using (3).

Step4: Update the global best and local best position.

Step5: Update $V_{ll}(t)$ and $V_l(t)$ for all particles using (1) and(2)

Step6: Update $p(t+1)$ by step2.

Step7: Compare $\rho(t)$ and $\rho(t+1)$. If $p(t+1) - p(t) < \xi$, then stop; Otherwise, increase t by one and continue from Step3

Feature extraction of segmented objects

In this subsection, we describe the extracted features of segmented objects. We can do this by the following calculations

1) *Area:* By counting the number of pixels included in segment I of the t -th frame, we calculate the area of the object $a_i(t)$.

2) *Width and Height:* We extract the positions of the pixel P_{xmax} (P_{xmin}) which has the maximum, minimum) X component by

$$P_{x,max} = (X_{max,x}, X_{max,y}), P_{x,min} = (X_{min,x}, X_{min,y})$$

Where $X_{max,x}$, $X_{max,y}$, $X_{min,x}$, and $X_{min,y}$ are the x - and y coordinates of the rightmost and leftmost boundary of segment i , respectively. In addition, we also extract

$P_{y,max} = (Y_{max,x}, Y_{max,y})$, $P_{y,min} = (Y_{min,x}, Y_{min,y})$
Then we calculate the width w and the height h of the objects as follows $W_i(t) = X_{max,x} - X_{min,y}$, $h_i(t) = Y_{max,x} - Y_{min,y}$

3) *Position:* We define the positions of each object in the frame as follows

$$X_i(t) = (X_{max,x} + X_{min,x})/2$$

$$Y_i(t) = (Y_{max,x} + Y_{min,x})/2$$

4) *Color:* Using the image data at P_{xmax} , P_{xmin} , P_{ymax} and P_{ymin} , we define the color feature of each object as in for the R (Red) component.

$$R_i(t) = \{ [R(P_{x,max}) + R(P_{x,min}) + R(P_{y,min}) + R(P_{y,min})] \} / 4$$

as well as by equivalent equations for the G and B components.

Object Tracking And Motion Estimation

The proposed algorithm for object tracking exploits pattern matching with the above features and makes use of the minimum distance search in the feature space. Using the image segmentation result of the object i in the t -th frame, we first extract the features of the object (t, i) . Here, the notation (t, i) stands for the objects i in the t -th frame. Then we perform the

minimum distance search in the feature space between (t, i) and $(t - 1, j)$ for all objects j in the preceding frame. Finally, the object (t, i) is identified with the object in the preceding frame which has the minimum distance from (t, i) . Repeating this matching procedure for all segments in the current frame, we can identify all objects one by one and can keep track of the objects between frames. A few comments on further refinements of the proposed algorithm are in order. 1) In calculation of the distance between (t, i) and $(t - 1, j)$ in position space, it is more appropriate to take account of motion determination and use estimated positions $(x(t) \& y(t))$.

$$\begin{aligned} \chi_i'(t + 1) &= \chi_i(t) + m_{x,i}(t) \\ y_i'(t + 1) &= y_i(t) + m_{y,i}(t) \\ m_{x,j}(t - 1) &= \chi_j(t - 1) - \chi_j(t - 2) \\ m_{y,j}(t - 1) &= y_j(t - 1) - y_j(t - 2) \end{aligned}$$

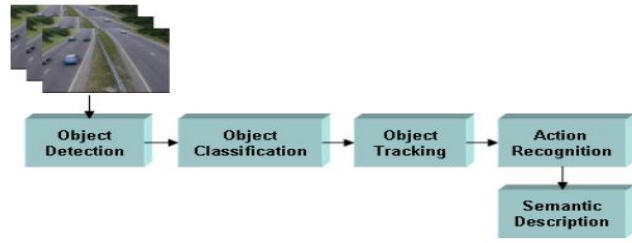
Instead of raw positions $x_j(t - 1)$ and $y_j(t - 1)$. The quantities $m_{x,j}(t - 1)$ and $m_{y,j}(t - 1)$ correspond to the motion vector of x - and y -directions of the object j . These replacements are available and used from the third frame onwards. The Euclidean distance DE and Manhattan distance DM is already sufficient for object tracking purposes. These two distances between vectors (x_1, \dots, x_n) and (y_1, \dots, y_n) are defined as

$$\begin{aligned} D_E &= \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} \\ D_M &= |x_1 - y_1| + \dots + |x_n - y_n| \end{aligned}$$

In order to treat all object features with equal weights, it is necessary to normalize the features. One possible way is dividing them by their maximum values. Dividing by 2^n , where the integer n is determined for each features so that approximately equal weights results, is another possibility. Pattern Matching in the Feature Space If $(t == 1)$ then

- a) Perform feature-extraction for segments.
 - b) go to (image segmentation of the next frame).
- If $(t >= 2)$ then
- a) Perform feature-extraction for segment i .
 - b) Calculation of distances $D(t, i; t - 1, j), \forall j$.
 - c) Search for the minimum distance among the distances
 - d) Identify (t, i) with $(t-1, k)$ and remove $(t-1, D(t, i; t - 1, k) = \min D(t, i; t - 1, j), \forall jk)$ from reference data.
 - e) Estimation of the positions of the segment i in the next Frame using equations mentioned above.
 - f) Repeat the matching procedure [from b) to e)] for all segments in the t -th frame.
 - g) Go to (image segmentation of the next frame).

It is determined for each feature so that approximately equal weights results, is another possibility. The second possibility has the advantage that the division can be realized by a shifting operation in a hardware realization.



III. SIMULATION AND RESULTS

For simulations, a tennis video sequence frame from 21-23 are considered. Each image of this video sequence is of size (352×220) . 2 depict the object tracking results obtained after implementing conventional adaptive thresholding image segmentation method followed by pattern matching. Fig. 3 shows the results obtained taking the same video using our proposed hybrid FCM-PSO image segmentation and pattern matching. It is obtained that Here two techniques have been repeated in tennis video sequence. Each image of this video sequence is of size 352×220 . In the tennis video sequence the frames from 21 to 23 are taken. The ROI (region of interested) is 16×16 for this video sequence. First we have shown the results obtained using thresholding method and then using FCMPSO method. In this video the ball and the hand both are moving. And we are going to track both of them. In each frame the mask regions are 24×24 for ball and 35×35 for hand. It is observed in Fig. 2 that the moving hand in the video sequence is little bit out of track by implementation of conventional methods. In Fig. 3, which is tracked by proposed hybrid PSO-FCM method, the tracking of moving hand is accurate. All parameters for Fig. 1 are presented in Table I and Table II. Similarly the extracted features and Distance measurements for Fig. 3 are presented in Table III and Table IV respectively.



Fig 1

Table I: Extracted Features For Tennis Video Sequence

(Fram, Object)	Area	Width	Height	Cen x	Ceny	Mv x	m vy
(22,1)	16497	47	351	217	177	2	1
(22,2)	304	19	16	44	140	14	0
(22,3)	1550	25	62	155	210	0	0

Table II : The Euclidian Distance Between Successive Frames

(Frame no, Object)	(23,1)	(23,2)	(23,3)
(22,1)	0	158.9025	71.196
(22,2)	19.6253	32.0624	142.3938
(22,3)	70.2353	115.2085	1



Fig 2

Table III: The Extracted Features Using FCM-PSO

(Frame, Object)	Area	Width	Height	Cenx	Ceny	Mvx	mvy
(22,1)	1649 7	47	351	217	177	0	0
(22,2)	304	19	16	44	140	10	12
(22,3)	1550	25	62	155	210	26	55

Table IV :The Euclidian Distance Between Successive Frames

(Frame no, Object)	(23,1)	(23,2)	(23,3)
(22,1)	1	175.9346	69.3542
(22,2)	185.4562	29.2373	60.8553
(22,3)	185.7148	158.3824	52.9528

IV. APPLICATIONS

Applications:

- Medical Imaging :
 - Locate tumors and other pathologies
 - Measure tissue volumes
 - Computer-guided surgery
 - Diagnosis
 - Treatment planning
- Study of anatomical structure
- Locate objects in satellite images (roads, forests, etc.)
- Face recognition

V. CONCLUSION

We have proposed an object tracking algorithm for video images using hybrid FCM-PSO image segmentation method and pattern matching of the segmented objects between frames. Simulation results for Tennis video frame sequences verify the suitability of the algorithm for reliable moving object tracking. The gray value at the centre pixel of an object is used in order to extract color features of segmented objects. The proposed FCM-PSO segmentation method could be found to be efficient to represent the object's color features for the tracking purpose. It is observed that if mistracking occurred at some frame, the proposed algorithm could recover correct tracking. This stability characteristic of the proposed method

results from the fact that object matching is performed in feature space between all objects in successive frames. As the proposed algorithm performs faster and with better accuracy than the existing techniques, implementation with a few moving objects in real time mode may be extended for future work.

VI. FUTURE WORK

The relative simplicity of this tracking algorithm promises that an FPGA implementation is possible and already sufficient for real time applications with a few moving objects. Thus, VLSI implementation of the algorithm is possible by using our developed architectures for image segmentation and a fully parallel associative memory with high-speed minimum Manhattan distance search, both of which have been already realized as VLSI circuits.

VII. REFERENCES

- [1] Yiwei Wang and John F. Doherty, *Moving Object Tracking in Video*, *IEEE Trans. Circuits Syst.*
- [2] Yoav Rosenberg and Michael Werman *Real-Time Object Tracking from a Moving Video Camera: A Software Approach on a PC*, *IEEE Trans. Circuits Syst.*
- [3] *Multi-Object Tracking in Video* by Johnson I Agbinya and David Rees *CSIRO Telecommunication and Industrial Physics, Image & Signal Processing Group, 76, Epping, NSW, 2121, Australia*
- [4] PATEUX S. "Tracking of video objects using a backward projection technique". *Proceedings of VCIP'2000 (Visual Communication and Image Processing)*. Perth, Australia. June 2000.
- [5]. J. Kennedy, R. C. Eberhart, and Y. Shi, —Swarm intelligence|| , San Francisco, Morgan Kaufmann Publishers, 2001.
- [6].N. Otsu, —A threshold selection method from gray-level histogram,|| *IEEE Transactions on System Man Cybernetics*, Vol. SMC-9, No. 1, pp. 62-66, 1979.
- [7] C. Stauffer and W.E.L. Grimson, —Learning patterns of activity using realtime tracking,|| *IEEE Trans. PAMIL*, Vol. 22, No. 8, pp. 747–757, 2000.
- [8] J. Kennedy, R. C. Eberhart, and Y. Shi, —Swarm intelligence|| , San Francisco, Morgan Kaufmann Publishers, 2001.
- [9]. Lipton, H. Fujiyoshi, and R. Patil, —Moving Target Detection and Classification from Real-Time Video|| , In *Proceedings of IEEE Workshop on Application of Computer Vision*, 1998,
- [10]. J. Kennedy, R. C. Eberhart, Particle Swarm Optimization, in *Proceeding of IEEE International Conference*