IEEE India Council Conference, INDICON 2009; Ahmedabad; 18 December 2009 through 20 December 2009; Category number CFP09598; Code 79708; Article number 5409361; DOI- http://dx.doi.org/10.1109/INDCON.2009.5409361

Object Tracking in Video Images Using Hybrid Segmentation Method and Pattern Matching

Dipti Patra[#1], Santosh Kumar K[#], Debarati Chakraborty[#]

[#] *Electrical Engineering Department, National Institute of Technology, Rourkela*

[1]dpatra@nitrkl.ac.in

***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

The moving object tracking in video images has attracted a great deal of interest in computer vision. For object recognition, navigation systems and surveillance systems etc. object tracking is an indispensable first-step [1, 2, 3, 4]. 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 prediction or motion estimation.

The conventional approach based on the difference, image can't simultaneously detect still objects. They can't be applied to the case of a moving camera also. The other approaches including the camera motion information have been proposed by many researchers. But, they still contain problems in separating the information from the background.

In this paper, we propose a novel method for object tracking. It consists of three stages, (i) image segmentation using hybrid algorithm exploiting the notion of Particle Swarm Optimization and Fuzzy-C-Means clustering techniques, (ii) feature extraction for pattern matching, (iii) 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.

The segmented images of consecutive frames are used for pattern matching after feature extraction. The motion of the object from reference frame to present frame is calculated in both X and Y directions to track the moving objects in video sequence.

II. PARTICLE SWARM OPTIMIZATION

Particle Swarm Optimization (PSO) algorithm is a kind of evolutionary computational technique developed by Kennedy and Eberhart in 1995 [5]. 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 $x_{i} = (x_{i1}, x_{i2}, \ldots, x_{id})$ 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 $p_i = (\rho_{i1}, \rho_{i2}, \ldots, \rho_{id})$. The position change (velocity) of particle i is $V_i = (V_{i1}, V_{i2}, \ldots, V_{id})$. Particles update their velocity and position through tracing two kinds of 'best' value. One is its personal best (*pbest)*, which is the location of its highest fitness value. Another is the global best (*gbest*), which is the location of overall best value, obtained by any particles in the population. Particles update their positions and velocities according to the following equations:

$$V_{id}(t+1) = \chi(wV_{id}(t) + c_1\phi_1(\rho_{id}(t) - x_{id}(t)) + c_2\phi_2(\rho_{gd}(t) - x_{id}(t)) \quad (1)$$

$$x_{id}(t+1) = x_{id}(t) + V_{id}(t+1) \quad (2)$$

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

III. FUZZY C-MEANS CLUSTERING

Fuzzy C Means (FCM) is one of the most commonly used fuzzy clustering techniques for different degree estimation problems [6,7]. It provides a method that shows how to group data points that populate some multidimensional space into a specific number of different

978-1-4244-4859-3/09/$25.00 ©2009

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 centres that minimize a dissimilarity function.

IV. PROPOSED CONCEPT FOR OBJECT TRACKING

The hybrid clustering approach to image segmentation starts by choosing the number of clusters and a random initial cluster centre 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 centres.

A. *Image Segmentation using hybrid PSO-FCM Clustering Technique*

In fuzzy clustering, a single particle represents a cluster centre vector . Let V_i is the *D*-dimensional vector of i^{th} cluster centre and can be represented as $\{V_{i1}, V_{i2}, \ldots, V_{iD}\}$.

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 centre and each iteration presents a solution giving a vector of cluster centres. We determine the position of vector V_l for every particle and update it. We then change the position of cluster centres based on these particles. For the purpose of this algorithm, the fitness of particles is easily measured as follows:

$$\rho_l^{(t)} = \sum_{K=1}^{n}\sum_{i=1}^{C} [A_i(X_K)]^m \left\| X_K - V_i^{(t)} \right\|^2 \qquad (3)$$

where *n* is the no. of data vector, C is no. of cluster centres, $V_l^{(t)}$ is the position of particle *l* in stage t, $Vel_l^{(t)}$ is the velocity of particle *l* in stage *t*, X_k is the vector of data and k=1,2,…,*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 *i*.

Proposed Segmentation Algorithm:

Step1: Let t=0, select initial parameters initial cluster centres *c*, initial position of particle χ, initial velocity of particles *w,* a real number m, a small positive number ε, and stopping criterion.

Step2: Calculate the $A_i^{(t)}(X_K)$ for all particles as follows where $i = 1, 2, .., C$; $k = 1, 2, .., n$. For each $X_K \subset X$, if $\|X_k - V_i^{(t)}\|^2 > 0$ for all i=1, 2, …, C; then define:

$$A_i^{(t+1)}(X_K) = \left[\sum_{j=1}^{C} \left(\frac{\left\| X_K - V_i^{(t)} \right\|^2}{\left\| X_K - V_j^{(t)} \right\|^2} \right)^{\frac{1}{m-1}} \right]^{-1}$$

If $\|X_k - V_i^{(t)}\|^2 = 0$; then define $A_i^{(t+1)}(X_K)$ by any nonnegative real numbers satisfying the following equation.

$$\sum A_i^{(t+1)}(X_K) = 1$$

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

Step4: Update the global best and local best position.

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

Step6: Update $\rho^{(t+1)}$ by step2.

Step7: Compare $\rho^{(t)}$ and $\rho^{(t+1)}$. If $p^{(t+1)} - p^{(t)} \le \varepsilon$, then stop; otherwise, increase t by one and continue from Step3.

B. *Feature extraction of segmented objects:*

In this subsection, we describe the extracted features of segmented objects. Figure. 1 shows an example of a segment for explanation purposes.

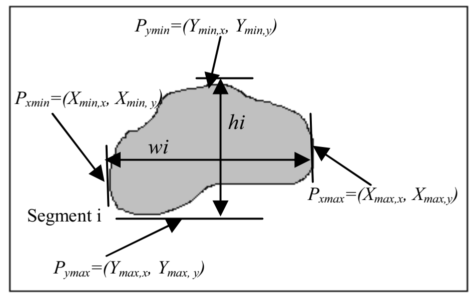

Fig.1

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

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

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

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

$$X_i(t) = \frac{X_{\max,x} + X_{\min,x}}{2}$$
$$Y_i(t) = \frac{Y_{\max,x} + Y_{\min,x}}{2}$$

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

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

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

V. 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 ($x(t)$ & $y(t)$)

$$x_i^{'}(t+1) = x_i(t) + m_{x,i}(t) \quad (4)$$

$$y_i^{'}(t+1) = y_i(t) + m_{y,i}(t) \quad (5)$$

$$m_{x,j}(t-1) = x_j(t-1) - x_j(t-2)$$
$$m_{y,j}(t-1) = y_j(t-1) - y_j(t-2)$$

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.

2) The Euclidean distance D_E and Manhattan distance D_M is already sufficient for object tracking purposes. These two distances between vectors $(x_1, \cdots, x_n)$ and $(y_1, \cdots, y_n)$ are defined as

$$D_E = \sqrt{(x_1 - y_1)^2 + ... + (x_n - y_n)^2}$$

$$D_M = |x_1 - y_1| + ... + |x_n - y_n|$$

3) 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 ($t \geq 2$) 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 $D(t, i; t-1, k) \equiv \min D(t, i; t-1, j)$, $\forall j$.
d) Identify (t, i) with $(t-1, k)$ and remove $(t-1, k)$ from reference data.
e) Estimation of the positions of the segment i in the next frame using equations (4) and (5).
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.

V. SIMULATION AND RESULTS

For simulations, a tennis video sequence frame from 21-23 are considered. Each image of this video sequence is of size (352X220). Fig.2 depicts 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 X 220. In the tennis video sequence the frames from 21 to 23 are taken. The ROI (region of interested) is 16 X 16 for this video sequence. First we have shown the results obtained using thresholding method and then using FCM-PSO 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 24X24 for ball and 35X35 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.

Fig2.

TABLE I
EXTRACTED FEATURES FOR TENNIS VIDEO SEQUENCE

(Frame, Object)	Area	Width	Height	Cenx	Ceny	mvx	mvy
(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)	190.6253	**32.0624**	142.3938
(22,3)	70.2353	115.2085	**1**

Fig. 3

TABLE III

THE EXTRACTED FEATURES USING FCM-PSO

(Frame, Object)	Area	Width	Height	Cenx	Ceny	mvx	mvy
(22,1)	16497	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**	60.8353
(22,3)	185.7148	158.3824	**52.9528**

VI. CONCLUSIONS AND DISCUSSION

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 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.

REFERENCES

[1] G.L. Foresti, "A real-time system for video surveillance of unattended outdoor environments," *IEEE Trans*. Circuits and Systems for Video Technologies, vol. 8, no. 6, pp. 697–704, 1998.

[2] T. Morimoto, O. Kiriyama, Y. Harada, "Object Tracking in Video Pictures based on Image Segmentation and Pattern Matching", *IEEE International Symposium on"Circuit and System"*, vol. 04, pp. 3215-3218, 2005

[3] Junqiu Wang, Yasushi Yagi, "Integrating colour and shape texture features for adaptive real time object tracking", *IEEE Transactions on Image Processing,* vol . 17,no. 2, pp. 235-240,2007

[4] Wen-Yan Chang, Chu-Song Chen, and Yong-Dian Jian," Visual Tracking in High-Dimensional State Space by Appearance-Guided Particle Filtering", *IEEE Transactions on Image Processing*, vol. 17, No. 7, pp. 1154-1168, July 2008.

[5] J. Kennedy, R. C. Eberhart, "Particle Swarm Optimization", *in Proceeding of IEEE International Conference on Neural Networks,* IEEE Press, Perth, WA, vol. 4, pp. 1942-1948,1995.

[6] E.. Mahdizadeh, R.. Tavakkoli-moghaddam, "A hybrid Fuzzy clustering PSO algorithm for a clustering supplier problem", *Proceedings of IEEE IEEM 2007,* vol 7, pp 1466-1470, 2007.

[7] Z. Xiang Cheng, S. Quan-tai, L. Li Mei, ""New two dimensional fuzzy C means clustering algorithm for image segmentation", *Springerlink*, vol. 15, pp 882-887, 2008.