Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Engineering

Elixir Comp. Engg. 79 (2015) 30296-30299

ISSN: 2229-712X

Object tracking based on particle filter and feature adaptation based on SIFT

Yazdanfar Heidary and Farokh Koroupi

Department of Computer Management, Baft Branch, Islamic Azad University, Baft, Iran.

ARTICLE INFO

Article history:

Received: 19 December 2014;

Received in revised form:

20 January 2015;

Accepted: 5 February 2015;

Keywords

Particle filter, Moving object,
SIFT, Soccer ball,
Feature points.

ABSTRACT

In this paper Particle filter and SIFT algorithm are combined for moving objects tracking. SIFT key points constitute parts of particles to improve the distribution sample. The work, experiments of this study are performed using a movie from a soccer match. First, the ball is selected in the first frame by removing the background. Afterwards, key points are extracted via SIFT algorithm which are combined with particle filter. The particle filter algorithm tracks the ball till it is inside the scene.

Introduction

In the proposed method, the desired object (which might be a soccer ball, pedestrian, automobile and so on) is selected by removing the background in the first frame. Then, it is expected that the algorithm tracks the object till it exits the scene. In the proposed method 2D tracking of objects in a colored video stream is achieved. A colored video stream is a set of colored images which are dependent semantically and have close locations. The proposed algorithm in this study is a combination of particle filter and SIFT algorithm which is called SIFT particle filtering (SPF). The flowchart of the algorithm and its steps are shown before. It includes background modeling, removing shadow, identifying objects movement and so on. The proposed algorithm is utilized for object tracking. In the details of the first frame static background is extracted. Then, the difference between current frame and background is extracted to track the object. When the shadow is removed complete information of the object might be extracted. If the object is lost, the tracking model must be updated. The particle filter combines important particles with the feature points of SIFT in order to track the object. Using particle retention features the particle destruction could be reduced. The proposed algorithm improves identification precision and decreases computational complexity [1,2,3].

Methodology

The proposed methodology consists of several steps. Shadow and background are removed, moving objects are identified and then SPF is used to track the objects. More precisely, firstly, frames associated with video sequence are exploited to create static background. Afterwards, the difference between current frame and background image is used to identify information about moving object. When the shadows are removed the complete information about objects might be extracted. SIFT feature points are extracted for objects which are going to be tracked and using particle filter the object is tracked. In case of losing the object the tracking model is updated if necessary. Using combination of SIFT and particle filter reduces particle destruction as SIFT feature points are highly stable. Furthermore, the precision is increased while computational complexity is decreased.

Motion detection

There are various methods for detecting the object or modeling the moving object. One of the simplest ones is the sliding window method where a rectangular window in different sizes scans the image intensively. Each of these regions is considered as a favorable region for future decision making. It is one of the easiest methods for detecting objects with different sizes. In this paper the main idea of the study is mentioned. Then, details of implementation of proposed method including combination of SIFT algorithm and particle filter are explicated. The block diagram of the system is depicted here and its details are explained [1,2].

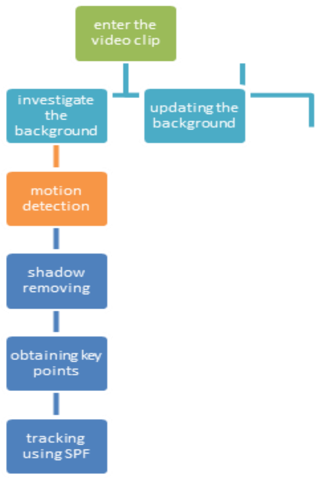

Figure 1. Block diagram of implementation of the study

Background subtraction

When the video is transformed to a sequence of frames, the object must be separated from background in the first frame. The common problems regarding machine vision is locating the specific object in the image. Background subtraction is a simple but effective technique for extracting foreground objects. In each frame the result which includes a preview of object motion is obtained by subtracting the background which was extracted previously. After image processing this image is transformed to a binary-colored space. The derived image may include shadows and small moving objects in addition to the desired one. To

Tele:
E-mail addresses: farokh.koroupi@gmail.com

address this problem and eliminate unwanted information, a set of morphologic operations are performed. In this study a background of video frame which excludes the moving object is utilized.

Key points initialization

In this step the key points must be extracted from desired motion area. Which is called feature extraction or key point extraction. These key points are fed to particle filter to be combined with its particles so that the new location of the object could be estimated using motion equation.

Feature extraction in SIFT algorithm

Feature extraction in SIFT algorithm is composed of three main stages: 1) scale space exterma detection, 2) eliminating unstable exterma, 3) orientation assignment to each feature. In the following these steps are explained in details.

1) Scale space exterma detection: the only kernel which is appropriate for scale space (different scale sigmas) is Gaussian.

$$g(x,y,z) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)} / 2\sigma^2 \quad (1)$$

For a 2D image I[x,y], $L(X,Y,\sigma)$ is scale space which is calculated as follows:

$$L(X,Y,\sigma) = (X,Y,\sigma) * I(X,Y) \quad (2)$$

$G(X,Y,\sigma)$ is variable space of the Gaussian function which is obtained experimentally and Laplacian matrix of Gaussian (LOG) is $\sigma^2 \nabla^2 G$. Then, exterma is selected according to figure2.

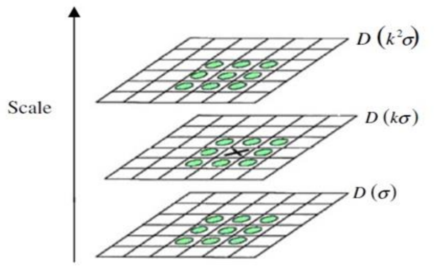

Figure 2: Exterma selection

In figure 2.the best concept of scale is shown.

(LOG)

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \quad \mathbf{(3)}$$

In this study σ is 0.5 and it should be calculated for all DOG images. To describe the feature, we need to subtract two images. Afterwards, exterma must be selected in 3*3*3 neighboring (depicted in figure 2).

2) Eliminating edge unstable points

Consider Tailor series (4).

$$D(\vec{X}) = D + \frac{\partial D^T}{\partial \vec{X}} \vec{X} + \frac{1}{2} \vec{X}^T \frac{\partial^2 D^2}{\partial \vec{X}^2} \vec{X} \quad (4)$$

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X}, where\ D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^{-1}}{DX} \hat{X} \quad (5)$$

Unfounded points are eliminated via $|D(\hat{X})| 0.03$ rule.

Moreover, rejected points with strong edge response in the same direction as follows

$$\frac{(Tr(H))^2}{Det(H)} \prec \frac{(r+1)^2}{r} \quad (6)$$

If the following is met:

$$where\ \mathrm{Tr(H)} = D_{xx} + D_{yy}, Det(H) = D_{xx} D_{yy} - (D_{xy})^2, r = 10 \quad (7)$$

3) Scale orientation assignment is performed based on points which are derived from correct image according to equation 7-2; thus we will have:

$$L(x,y) = G(x,y,\sigma) * I(x,y) \quad (8)$$

Then, gradient and its orientation are obtained from 9.

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (9)$$

Figure 3. The results of SIFT algorithm adaptation.

Figure 3 demonstrates matching of automobile image points which were received before and the moving automobile.

$$\theta(x,y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1))}{(L(x+1,y) - L(x-1,y))}\right) \quad (10)$$

4) Obtaining matching points: the SIFT descriptor error of moving object exterma is achieved from 11.

$$error = (\hat{F}_1 - \hat{F}_2)^2 \quad (11)$$

Figure 3 demonstrates SIFT feature points and matching of a stored image of the object and its image during motion. As can be seen the SIFT feature points achieve acceptable results in detection and matching of the object.

Object tracking using combination of particle filter and SIFT (SPF tracking)

Particle filter takes a set of samples and assigns a number of particles to them during formatting. Then a recursive Bayesian is applied to them and the accuracy of this estimation is achieved by a specific weighting criterion. Particle filters use a probability distribution function to estimate the closest state to the reality. Particles are state vectors which are accumulated in a location and utilize the mentioned probability distribution. A weight is assigned to each particle which is originated from the similarity between particle location and main location. Particle filter algorithm uses sampling to create N particles. Then, comparing the samples (tracking region) determines important weights. Resampling step is the most prominent factor in each run of particle filter because in conventional particle filter the variance of particle's weight rapidly increases. The main idea of resampling is to convert particles to normalized value by eliminating particles with large or small weight. As SIFT key points are utilized the number of particles which should be sampled decreases. It reduces computational burden and avoids useless distribution of weights. The problem of tracking is the lack of natural synchronization between iterations. In tracking new iterations are dependent on previous ones. Therefore in this study the state space for first frame is initialized using matching key points obtained by SIFT in order to realize an active tracking.

Generic structure of SPF tracking system

Color distribution model

Color is the most prominent and popular feature of the image for modeling the object in tracking applications. It is constant in case of variations in size, orientation and perspective. To extract color feature the type of color space must be determined. Colored histogram of an object provides numerous advantages for its tracking. In particle filters the target is tracked based on comparison between its colored histogram and histogram of sample locations using Bhattacharyya distance. There are diverse methods for employing color information among which the most common is color histogram model [10, 11, 12]. Instead of using whole image, color histogram vector (hist) corresponding to inside of a part of object (e.g. the rectangular whose location is determined by xk state) is extracted. Weighted histogram is a simple statistical model of number of image pixels from each color component.

Favorable motion area

In this study the favorable motion area depends on the moving object which might be a soccer ball, pedestrian or moving automobile. The color distribution inside tracking area is derived from the following equation.

Distribution of color q in desired area is calculated as follows:

$$q = \{q(u)\}_n \tag{12}$$

Color distribution probability is:

$$q = f \sum_{i=1}^{i} k\left(\frac{\|xi\|}{a}\right)\delta(h(x_i) - u) \tag{13}$$

Where I is the number of pixels in particles region; u is discretization components in the color; **δ** is Kronecker delta remained in the function; a is the size of particle region; k is a weight function which assigns smaller weights to regions which are farther from circle origin; f is the normal coefficient and h is histogram function which is achieved in xi location of the color component. S includes combination of SIFT key points with current motion area

$$S_0 = \{sift(X_0^{(i)}), 1/N\}_{i=1}^{N} \tag{14}$$

Due to replaceable selective method, N samples are extracted from St-1 according to w_{t-1}^{i}

Calculating normal weight of accumulation:

$$c_{t-1}^{0} = 0, c_{t-1}^{i} = c_{t-1}^{i-1} + w_{t-1}^{i} \tag{15}$$

$$c_{t-1}^{'i} = c_{t-1}^{i} / c_{t-1}^{N} \tag{16}$$

For generating uniform random number r from [0,1] and searching for minimum K in 17

$$c_{t-1}^{'k} \geq r \tag{17}$$

Substituting in

$$S_{t-1}^{'i} = S_{t-1}^{k} \tag{18}$$

Dynamic equation

$$S_t = AS_{t-1} + R\,.\,rank \tag{19}$$

Where A is state variation matrix and R is rank which is a Gaussian random matrix.

For center of moving object:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad R = \alpha \begin{bmatrix} \sigma^3/3 & \sigma^2/2 \\ \sigma^2/2 & \sigma^3/3 \end{bmatrix}$$

$$\sigma = 3, \alpha = 0.35$$

Ssift1 might be obtained via calculating SIFT key points in at time t to St. In other words, SIFT algorithm firstly extracts key points and provides them for particle filter. Then, particle filter algorithm generates N particles. Specific weights are assigned to the particles according to SIFT key points. Afterwards, the location of moving object is estimated using probability function and based on significant weights.

Calculating Bhattacharyya p[pt,q] between candidate sample and target mask

The distance between histogram of reference image, q, and volunteer image histogram, hist, is stated by Bhattacharyya distance. To assign weights to samples the Bhattacharyya coefficient between target histogram and histogram of the area shown by each sample should be calculated. The region of each sample is identified by its state vector s(n) [1, 4, 6]. Since it is desired to have samples whose color distribution is more similar to target color distribution, larger weight is assigned to samples with smaller Bhattacharyya distance. Thus, samples' weights are calculated as follows. A famous magnitude between two distributions, p(u) and q(u), is Bhattacharyya coefficient such that for two matched normalized distribution we have p=1 which demonstrates complete matching between two distributions.

Calculating Bhattacharyya

$$p[p,q] = \sum_{u=1}^{m} \sqrt{p^u q^u} \tag{20}$$

So the distance between two distributions is:

$$d = \sqrt{1 - \rho[p,q]} \tag{21}$$

Resampling:

If the target is lost resampling must be performed using equation

$$w_t^i = \frac{1}{\sqrt{2\pi\sigma}} \exp(-(1-\rho[P_t, q]) / 2\sigma^2)\,.\,W_{t-1}^i \text{ of } S_{\text{sift},\,t} \tag{22}$$

6) The central location of moving object at t might be calculated by deriving average of weight sample.

$$E(S_t) = \sum_{i=1}^{N} W_t^i S_{sift,t}^i / \sum_{i=1}^{N} W_t^i \tag{23}$$

Combining particles and SIFT feature points the sample distribution is updated and included in particle filter process. Therefore, the precision of particle description is improved. Retaining definite number of independent particles with complete and strong characteristics, the particle destruction is reduced because the points which are selected as key points have special characteristics (such as object edges). These key point help motion and tracking precision improvement.

The results of soccer ball tracking experiment

In conventional particle filter algorithm the particle is selected based on dynamic equation in a stochastic manner and in complicated situations it leads to incorrect results. The results of our proposed tracking method (figure 4-1) are compared to the results of conventional particle filter (figure 4-2). The tracking is performed in a video record of a soccer match. The algorithm is used in MATLAB environment. There are 300 initial particles in this experiment. As shown by results when the scale of soccer ball changes dramatically, the particles remain in tracking area using our algorithm; whereas, in conventional particle filter method incorrect tracking results are obtained. Additionally, the execution time of two methods are compared, the difference is 15%s which could be neglected in actual applications.

Figure 5. Conventional particle filter (PF)

Figure 6. Combination of particle filter and SIFT (SPF)

Tracking in different and complex environments

Figure 7 illustrates the video sequence of a few frames of soccer match where soccer ball might be covered by players. It is also worsen due to similar color of ball and players' sportswear. Moreover, variations occur when the light of stadium changes and uniform light is disturbed. Furthermore, when the ball goes farther from camera its scale changes. Since SIFT key points are invariant and do not change by scale, brightness and so on, our proposed method is robust against brightness, scaling and coverage changes. It is also confirmed by our experimental results. In case of long durations of being

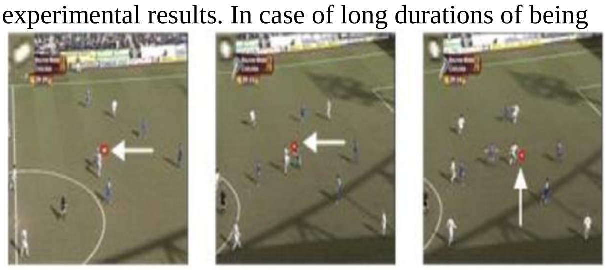

Figure 7. Soccer ball tracking in different and complex environment

The diagram of comparison between conventional particle filter and proposed method (SPF)

It can be seen that in diagram 1 tracking using SPF method is much more efficient than conventional particle filter. The red line shows conventional particle filter while the black one corresponds to SPF method.

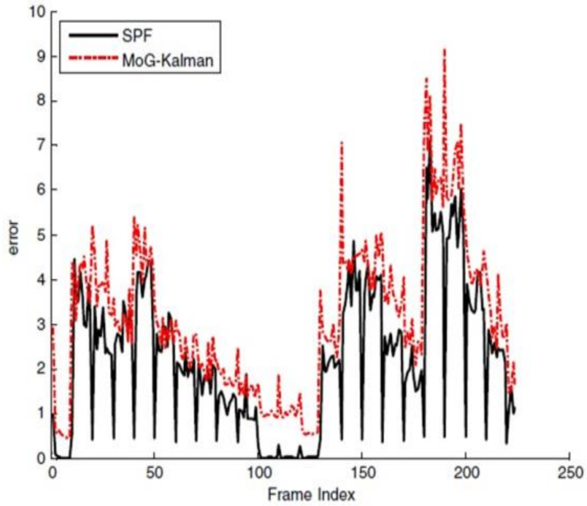

Figure 8. the diagram of comparison between conventional particle filter and proposed method (SPF)

References

[1] Arulampalam S, Maskell S, Gordon N, Clapp T (2002) A tutorial on particle filters for on-line nonlinear/

[2] nongaussian bayesian tracking. IEEE Trans Signal Process 50(2):174–188

[3] Ba HX, Zhao ZG, Yang F, Cao L (2004) JPDA in multisensor multitarget tracking. J Syst Simul 16

[4] Babu RV, Perez P, Bouthemy P (2007) Robust tracking with motion estimation and local kernel-based color modeling. Image Vis Comput 25(8):1205–1216

[5] Belcher C, Du Y (2009) Region-based SIFT approach to iris recognition. Opt Lasers Eng 47**147–139:(1)**

[6] Betke M, Haritaoglu E, Davis LS (2000) Real-time multiple object detection and tracking from a moving object. Machine Vision and Applications 2(2):69–83

[7] Breitenstein MD, Reichlin F, Leibe B, Koller-Meier E, Gool LV (2009) Robust tracking-by-detection using a detector confidence particle filter. In: Proceedings of IEEE International Conference on Computer Vision, Kyoto, Japan, 1515–1522

[8] Cai YZ, Freitas N, Little J (2006) Robust visual tracking for multiple targets. In: Proc. the 9th European Conference on Computer Vision (ECCV), Graz, Austria, vol.4, pp. 107–118

[9] Comaniciu D, Meer P (2002) Mean Shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal Mach Intell 24(5):603–19