8-5

MVA2009 IAPR Conference on Machine Vision Applications, May 20-22, 2009, Yokohama, JAPAN

Improvement of Particle Filtering for Intersection of Targets with Similar Patterns

Naoki Enda
Faculty of Engineering
Chubu University
Kasugai 487-8501, Japan
enda@cvl.cs.chubu.ac.jp

Shinji Fukui
Faculty of Education
Aichi University of Education
Kariya 448-8542, Japan
sfukui@auecc.aichi-edu.ac.jp

Yuji Iwahori
Faculty of Engineering
Chubu University
Kasugai 487-8501, Japan
iwahori@cs.chubu.ac.jp

Robert J. Woodham
Dept. of Computer Science
University of British Columbia
Vancouver, B. C. Canada V6T 1Z4
woodham@cs.ubc.ca

Abstract

The particle filtering, which is a probability technique, can be used for a robust tracking to noise and occlusion. It often uses characteristics based on the appearance information of a target object, such as color and edge, for calculating likelihood. As far as the appearance information is used, the particle filtering fails to track the target object when it intersects other objects which have similar patterns. This paper proposes a new method for addressing the problem. The method uses color and other characteristics such as velocity information of the tracking object and distance information between tracking objects except the appearance.

1 Introduction

The particle filtering is used for tracking moving objects[1]. It is robust to occlusion and noise. When a target object is tracked by the particle filtering, the likelihood is often calculated by appearance information, such as color and edge information of the object[1][2][3][4]. Methods using appearance information sometimes fail to track moving objects with similar patterns when they intersect with each other.

S. Khan *et al.* proposed a method to address the intersection of some targets by observing a scene from several points of view and integrating the information[5]. The method needs at least four cameras to track some moving objects with high performance. Using many cameras costs high expense and causes a problem that it is necessary to secure locations for installing them.

J. Sullivan *et al.* proposed a trajectory based method to solve the intersection problem in a soccer game[6]. The method solved the problem by analyzing the trajectory of each player which is acquired by background subtraction and labeling. The method does not perform the whole process in real time because it needs the trajectories of all players in the whole game.

L. Zhu *et al.* proposed a online sample based approach for multiple objects tracking [7]. The method uses local texture information and relative spatial features which are obtained from several frames. This method fails to track objects when they have similar patterns and intersect in a scene because it estimates the center position of the tracking object based on only the local texture information.

This paper proposes a new approach for tracking multiple objects with similar patterns in a video sequence taken by a single camera. The proposed method is based on the particle filtering. We improve the likelihood function in order to address the problem that tracking fails when target objects with similar patterns intersect. The likelihood is calculated by color histogram of the target object and other characteristics except the appearance of the object, such as velocity of the tracking object and distance information between tracking objects. Moreover, the method uses the result of background subtraction in order to track targets efficiently and accurately. Results are demonstrated by experiments using real video sequence.

2 Increasing Efficiency and Accuracy of Particle Filtering

The particle filtering is the approach that state variables are estimated by approximating the posterior probability distribution of them by many particles. Each particle consists of the state variables and a weight. The particle filtering estimates the state variables by following steps[1]. First, the state variables at time t are predicted by them at $t-1$. Next, the weight is calculated by the likelihood value. At last, particles are selected by Sequential Importance Sampling[8]. As a result, the discrete approximation of the posterior distribution of the state variables is obtained.

The particle filtering usually generates the particles at t by adding random values to the state variables of the particles at $t-1$. The previous methods scatter many particles into the region in which calculating the weights of the particles is useless. We proposes a method to increase efficiency and accuracy of the particle filtering in this section. The proposed method uses the velocity of the target object and the result of background subtraction. The method defines two types of the state variables and improves the state transition function to use velocity and distance information.

[1]See details about the particle filtering in [1]

2.1 Two Types of State Variable

Let the state variables be $\boldsymbol{x}_t$. In this paper, $\boldsymbol{x}_t$ is defined as follows:

$$\boldsymbol{x}_t = \begin{cases} [u_t, v_t, \dot{u}_t, \dot{v}_t, w_t, h_t]^T \\ \quad \text{(if the target object intersects with others)} \\ \\ [u_t, v_t, \dot{u}_t, \dot{v}_t]^T \quad \text{(otherwise)} \end{cases}$$

where the set of (u_t, v_t) represents the center position of the tracking object, the set of $(\dot{u}_t, \dot{v}_t)$ represents the velocity of the tracking object and w_t and h_t represent the width and height of the rectangle which involves the tracking object.

The proposed method extracts moving object regions by background subtraction at first[2]. Then the likelihood of only the particles that the set (u_t, v_t) of each particle is the point in the extracted region is calculated. This process decreases the useless calculations. Furthermore, the number of the state variables can be reduced because w_t and h_t does not need to be calculated by the particle filtering when the size of the target object can be obtained by the result of background subtraction. This makes it possible that a moving object is tracked more steadily with less particles.

When the target object intersects with others, its width and height can not be obtained accurately from the result of background subtraction. In such a case, w_t and h_t are added as the state variables and calculated by the particle filtering. w_{t-1} and h_{t-1}, which are obtained by using the result of background subtraction at $t-1$, are set to w_t and h_t respectively. In addition, the number of particle is increased because the number of state variables increases. How to judge whether the target object intersects with others or not is described in [10].

2.1.1 Obtaining Rectangle Size from Result of Background Subtraction

The process of obtaining the rectangle size from the result of background subtraction is described. First, the position of the tracking object at t is estimated. Next, the region which is the nearest region to the estimated position is determined as the target object region. At last, the width and height of the region are obtained.

It is necessary to find the region of the target object tracked by the particles from the detected regions by background subtraction. The position of the tracking object at t is estimated by the position and the velocity which are obtained by particle filtering at $t-1$. The nearest region to the estimated position is determined as the target object region. The estimated position of the tracking object, (u'_t, v'_t), is calculated by Eq. (1).

$$u'_t = \hat{u}_{t-1} + \hat{\dot{u}}_{t-1}, \; v'_t = \hat{v}_{t-1} + \hat{\dot{v}}_{t-1} \tag{1}$$

where the set of $(\hat{u}_{t-1}, \hat{v}_{t-1})$ represents the center position of the tracking object at $t-1$ and the set of $(\hat{\dot{u}}_{t-1}, \hat{\dot{v}}_{t-1})$ represents the velocity of the tracking object. Both of them are estimated by particle filtering at $t-1$.

[2]The proposed method obtains background image by background update method[9] in order to extract moving objects stably even though background environment is changed.

The region closest to (u'_t, v'_t) is determined as the region of the target object. Then the width and height of the region are obtained. As a result, the rectangle involving the target object can be generated with high performance and it is used for tracking.

2.2 State Transition Function of Proposed Method

Let $\boldsymbol{x}_t^{(i)}$ be state variables of i-th particle. The state transition function of the proposed method is defined by Eq. (2).

$$\boldsymbol{x}_t^{(i)} = \boldsymbol{x}_{t-1}^{(j)} + \boldsymbol{V} + \boldsymbol{\omega}_t \tag{2}$$

where $\boldsymbol{\omega}_t$ means the Gaussian noise vector and $\boldsymbol{V}$ means velocity vector. When the target object intersects with others, $\boldsymbol{\omega}_t$ consists of six Gaussian noises with means 0 and standard deviations $\{\sigma_u, \sigma_v, \sigma_{\dot{u}}, \sigma_{\dot{v}}, \sigma_w, \sigma_h\}$, which mean the standard deviations of Gaussian noise for $u_t, v_t, \dot{u}_t, \dot{v}_t, w_t$ and h_t respectively, and $\boldsymbol{V} = [\dot{u}_{t-1}^{(j)}, \dot{v}_{t-1}^{(j)}, 0, 0, 0, 0]^T$. The set of $(\dot{u}_{t-1}^{(j)}, \dot{v}_{t-1}^{(j)})$ means $(\dot{u}_{t-1}, \dot{v}_{t-1})$ of j-th particle at $t-1$. Otherwise, $\boldsymbol{\omega}_t$ consists of four Gaussian noises with means 0 and standard deviations $\{\sigma_u, \sigma_v, \sigma_{\dot{u}}, \sigma_{\dot{v}}\}$ and $\boldsymbol{V} = [\dot{u}_{t-1}^{(j)}, \dot{v}_{t-1}^{(j)}, 0, 0]^T$.

In the case that speed of the target object does not change rapidly in a short time, the probability that the target object exists at (u'_t, v'_t) is very high. So, $(u_t^{(i)}, v_t^{(i)})$ is transited based on $(u_{t-1}^{(j)}, v_{t-1}^{(j)})$, $(\dot{u}_{t-1}^{(j)}, \dot{v}_{t-1}^{(j)})$ and Gaussian noises. On the other hand, the proposed method can track the target object well even in the opposite case because Gaussian noises are added in addition to the velocity information.

3 Likelihood Function of Proposed Method

The function for the calculation of the likelihood value is very important factor for performance of the particle filtering. Many methods calculate the value based on the appearance information of the target object. However, they fail to track the target object when it intersects others with similar patterns. The likelihood function of the proposed method is improved so as to use velocity of the target object and distance information between tracking objects.

The Likelihood function L is given by Eq. (3).

$$L = L_c L_V L_d \tag{3}$$

where L_c represents the likelihood of color histogram which is the horizontal split histogram for the color histogram[10], L_V represents the likelihood of the velocity of the target object and L_d represents the likelihood of the distance information between tracking objects.

L_V and L_d are described in the following.

3.1 Likelihood of Velocity Information of Tracking Object

The proposed method uses velocity information of the target object. Suppose that velocity of the target object does not change rapidly in a very short time. In this case, the probability that the target object exists at (u'_t, v'_t) is very high. L_V should be defined so that it may become larger when (u_t, v_t) of a particle is closer to (u'_t, v'_t).

L_V is calculated by Eq. (4).

$$L_V = \exp\left(-\frac{\sqrt{(u_t^{(i)} - u'_t)^2 + (v_t^{(i)} - v'_t)^2}}{2\sigma_V^2}\right) \quad (4)$$

where the set of $(u_t^{(i)}, v_t^{(i)})$ means (u_t, v_t) of i-th particle and σ_V means the spreading factor. When $(u_t^{(i)}, v_t^{(i)})$ equals to (u'_t, v'_t), L_V becomes 1. And it becomes smaller when the distance between $(u_t^{(i)}, u_v^{(i)})$ and (u'_t, v'_t) becomes larger.

σ_V is given by Eq. (5).

$$\sigma_V = \sqrt{\hat{\dot{u}}_{t-1}^2 + \hat{\dot{v}}_{t-1}^2} + 1 \quad (5)$$

When the velocity of the target object is lower, σ_V should be smaller because $u_t^{(i)}$ and $v_t^{(i)}$ include less errors. In the contrast, it should be larger in the opposite case because they may include larger errors. σ_V is defined so that it may be changed according to the velocity values of the tracking object.

3.2 Likelihood of Distance Information of Tracking Objects

When the target object is close to other target objects which have the similar patterns, the object might be tracked by two or more particle groups because likelihood of particles which track other target object becomes large. Or no particle group might track the target object in such a case because likelihood of tracking particles becomes large on the other objects. The proposed method defines the likelihood function so that it may decrease the likelihood of particles more when they are closer to other objects.

When one or more objects are detected near the target object by the method described in [10], L_d is calculated by Eq. (6) and Eq. (7).

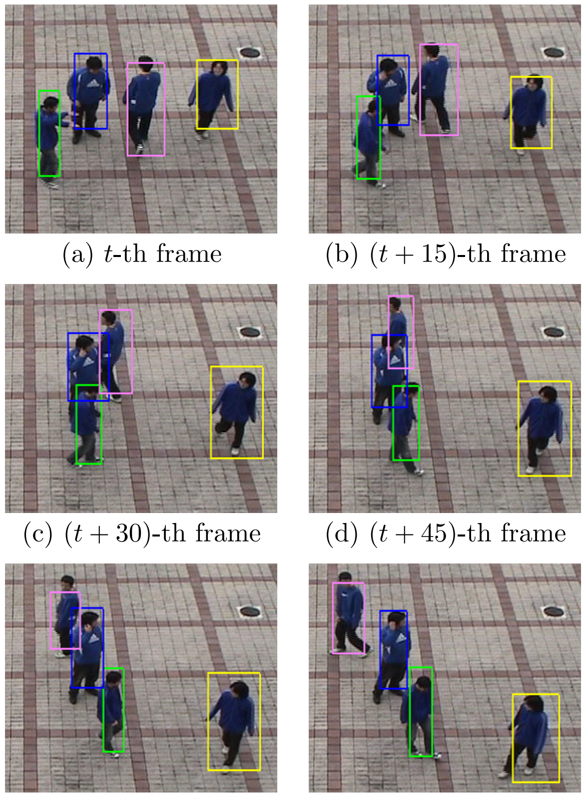
$$L_d = \prod_{k=1}^{M} L_d^k \quad (6)$$

$$L_d^k = 1.0 - \exp\left(-\frac{(u_t^{(i)} - u_t'^k)^2 + (v_t^{(i)} - v_t'^k)^2}{2\sigma_d^{k\,2}}\right) \quad (7)$$

where M means the number of objects which are detected near the target object, $(u_t'^k, v_t'^k)$ means the predicted center position of k-th moving object and σ_d^k means the spread factor for k-th object. On the other hand, when no object is detected, L_d is set to 1.

L_d^k is defined so that it may become less when $(u_t^{(i)}, v_t^{(i)})$ is closer to $(u_t'^k, v_t'^k)$ because (u'_t, v'_t) never equals to $(u_t'^k, v_t'^k)$ and $(u_t^{(i)}, v_t^{(i)})$ is farther from (u'_t, v'_t) in such a case. When $(u_t^{(i)}, v_t^{(i)})$ equals to $(u_t'^k, v_t'^k)$, L_d^k becomes 0.

σ_d^k should be changed according to the distance between the target object and k-th object because L_d^k may become small even if $(u_t^{(i)}, v_t^{(i)})$ is on the target object when σ_d^k is large. The proposed method calculates Euclidean distance between (u'_t, v'_t) and $(u_t'^k, v_t'^k)$, and sets it to σ_d^k.



(a) t-th frame (b) $(t+15)$-th frame

(c) $(t+30)$-th frame (d) $(t+45)$-th frame

(e) $(t+60)$-th frame (f) $(t+75)$-th frame

Figure 1: Results of Proposed Method

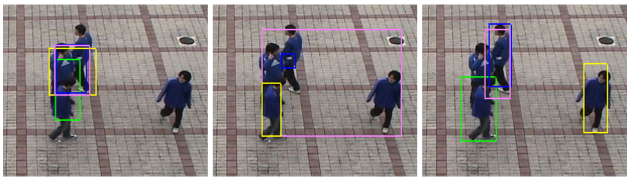
4 Experiments

Experiments using real video sequences were done for confirming the effectiveness of the proposed method. The size of each frame is 720 × 480 pixels and each pixel has 8bit RGB color values. A PC with Core 2 Quad Q6600 and 2G Byte main memory was used for calculation. $\sigma_u, \sigma_v, \sigma_{\dot{u}}, \sigma_{\dot{v}}, \sigma_w$ and σ_h were set to 15, 15, 1, 1, 2 and 2, respectively. The number of particles for each target was set to be 1000. It was set to be 2000 while the target object intersects with other targets. The rectangle including the target object is divided into five for the horizontal split histogram.

4.1 Experiment for Intersection of Moving Objects

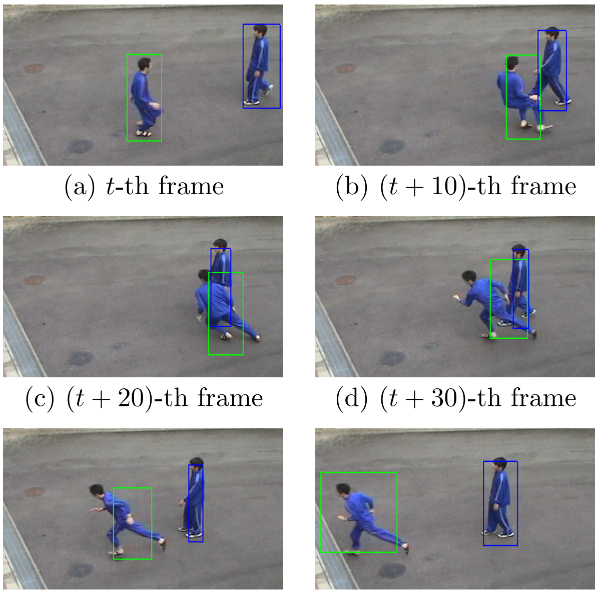
First, the experiment was done for an outdoor scene. Four people with similar clothes intersect in the scene. Fig. 1 shows a part of the experimental results. Fig. 1 shows that the proposed method can track multiple target objects with high performance even though they have similar patterns because it uses not only color information but also velocity and distance information. The success rate of the proposed method is 100.0% on this experiment.

The experimental results of the previous methods [4][7][10] are shown in Fig. 2. Fig. 2-(a) shows the result of the method described in [4]. The method fails to track the target objects when tracking object intersects other object with similar patterns because it uses only color information. Fig. 2-(b) shows the result of the objects tracking after segmenting object regions by the method described in [7]. The method fails to segment



(a) (b) (c)

Figure 2: Results of Previous Methods at $(t+30)$-th frame: (a) Results of [4], (b) Results of [7] and (c) Results of [10]



(a) t-th frame (b) $(t+10)$-th frame

(c) $(t+20)$-th frame (d) $(t+30)$-th frame

(e) $(t+40)$-th frame (f) $(t+50)$-th frame

Figure 3: Experimental Results for Rapid Turn of Target Object

moving object regions into the region of each object. It can not distinguish between the target object and similar pattern objects because it uses color and texture features. Fig. 2-(c) shows the result of the method described in [10]. It can track the target object when one object with similar pattern exists near the target object. However, it fails to track when two or more objects exist near the target. The proposed method has better performance than the previous approaches [4][7][10] when target objects with similar patterns intersect.

4.2 Experiment for Rapid Turn of Moving Object

Next, another experiment was done. The scene that a running man from the left of the image turned rapidly to the opposite direction is used for the experiment. Fig.3 shows a part of the experimental results. The results show that the proposed method can track target objects correctly even when velocity of the target object changes rapidly in a short time.

5 Conclusion

This paper proposed a new method which is robust to the intersection of moving objects with similar patterns. The proposed method is based on particle filtering. The likelihood of each particle is calculated by not only color histogram of target object but also the velocity of the object and distance between the tracking objects. Furthermore, the method uses the result of background subtraction. The number of particles of which likelihood is calculated and the number of state variables can be reduced by using it. As a result, efficiency and accuracy increased in the proposed method.

The proposed method assumes that the velocity of the target object does not change rapidly. However, it is shown that the method can track the object well even when it changes rapidly.

The previous methods using only appearance information fail to track objects with similar patterns when they intersect. The proposed method can track each object steadily in such a case because the likelihood function is defined so as to decrease the likelihood value of particles of which the center points in the state variables exist in the other object region.

Future work includes addressing objects whose patterns change in the middle of the scene.

Acknowledgment

Iwahori's research is supported by JSPS Grant–in–Aid for Scientific Research (C)(20500168), Chubu University Grant. Woodham's research is supported by the Natural Sciences and Engineering Research Council (NSERC).

References

[1] Michael Isard and Andrew Blacke: "CONDENSATION - conditional density propagation for visual tracking -," *Intl. J. Computer Vision*, Vol.29, No.1, pp.5-28, 1998.

[2] J. MacCormick and M. Isard: "Partitioned Sampling, Articulated Objects, and Interface-quality Hand Tracking,", *Proc. ECCV 2000*, pp.II-3-19, 2000.

[3] J. J. Sherrah and S. Gong: "Fusion of Perceptual Cues for Robust Tracking of Head Pose and Position," *Pattern Recognition*, Vol.34, No.8, 2001.

[4] P. Perez, C. Hue, J. Vermaak and M. Gangnet: "Color-Based Probabilistic Tracking," *Proc. ECCV 2002*, Part I, PP.661-675, 2002.

[5] S. Khan and M. Shah: "A Multiview Approach to Tracking People in Crowded Scenes Using a Planar Homography Constraint," *Proc. 9th ECCV*, Vol.4, pp.133-146, 2006.

[6] J. Sullivan and S. Carlsson: "Tracking and Labeling of Interacting Multiple Targets," *Proc. 9th ECCV*, Vol.1, pp.301-312, 2006.

[7] L. Zhu, J. Zhou, J. Song: "Tracking Multiple Objects Through Occlusion with Online Sampling and Position Estimation," *Pattern Recognition*, Vol.41, pp.2447-2460, 2008.

[8] A. Doucet, S. Godsill and C. Andrieu: "On Sequential Monte Carlo Sampling Methods for Bayesian Filtering," *Statistics and Computing*, Vol.10, No.3, pp.197-208, 2000.

[9] A. Shimada, N. Tsuruta, R. Taniguchi: "Dynamic Control of Adaptive Mixture-of-Gaussians Background Model," *CD -ROM Proc. of IEEE International Conference on Advanced Video and Signal Based Surveillance 2006*, 2006.

[10] Y. Iwahori, S. Okada, H. Kawanaka, S. Fukui, R. J. Woodham: "Particle Filter Based Tracking for Crossing of Targets with Similar Pattern," *IAPR MVA2007*, pp.307-310, 2007