The Research of Long Time Occlusion Problem Based on Object Permanence

Xiang-Dan Hou[(✉)] and Ming-Jing Zhang

Department of Computer Science and Engineering,
Hebei University of Technology, Tianjin, China
hxd@scse.hebut.edu.cn

Abstract. In the video surveillance system, due to the complicated background, the targets in the movement process often appear some or full of occlusion. How to detect occlusions, handle of issues efficiently, especially in the event of long time occlusion. Accurate target identification and tracking is the key indicators to evaluate the robustness of a target tracking algorithm. This paper deals with long time occlusions based on the concept of "object permanence" in psychology. This paper proposes a method based on "object permanence" algorithm to solve the identification and tracking problems after long time occlusions. The experimental results show that this algorithm can effectively solve the occlusion problem.

Keywords: Object permanence · Long time occlusion · Target tracking

1 Introduction

In the research field of computer vision, target tracking technology combines advanced achievements of image processing, artificial intelligence and pattern recognition. It has important practical value and broad development prospects in military target tracking, traffic monitoring, industrial production monitoring and many other fields [1]. However, since the complexity and the diversity of the target features and external environments, target tracking has been a challenging task. A robust target tracking algorithm must be able to resolve all the difficulties encountered during tracking process, such as rotation, size changes, illumination changes, and occlusion is the most difficult one [2].

The occlusion problem is very common in the process of target tracking. General can be divided into the following three types [3]:

(1) According to the occlusion type: self-occlusion; other objects occlusion; occlusion between targets;
(2) According to the degree of occlusion: partial occlusion; total occlusion;
(3) According to the time of occlusion: short time occlusion; long time occlusion.

The occlusion process can be divided into the following three stages [4]:

(1) Target into the occlusion state, this process lost some or all of the target information;
(2) Target holds the occlusion state, target information keeps the state of loss;
(3) Target leaves the occlusion state, target information is gradual recovery.

The target information is not stable or even lost during tracking process caused by occlusion. The key to tracking algorithm is to get enough target information and determine the location of the target. When the target appeared two times, the method can accurately identify the target. Thus, when occlusion occurs, how can we use the remaining target information continue to track is an important basis for the evaluation of a target tracking algorithm [5]. This paper deals with long time occlusion of the target tracking based on the "object permanence" method. It has achieved good tracking effect for long time occlusion.

2   Tracking Algorithm in Occlusion

In order to solve the occlusion problem, domestic and foreign researchers have done a lot of research. They have put forward many effective algorithms, which achieved better results in certain situations. These algorithms can be divided into five categories: (1) Center weighted matching region; (2) sub-block matching; (3) the trajectory prediction; (4) Bayesian theory; (5) multi-algorithm fusion.

A typical center-weighted algorithm is mean-shift algorithm [6]. Mean-shift algorithm is a non-parametric pattern matching algorithm based on kernel density estimation. This algorithm describes the color of the target distribution through the establishment of weighted histogram. It requires large amount of calculation and delays for application of real-time monitoring system.

Matching algorithm based on sub-block main idea is that the entire target area is divided into several sub-blocks and tracked separately [7]. This algorithm requires accurate matching of target sub-blocks. But in the case of severe or total occlusion, this algorithm can't find enough sub-block information and may cause the target lost.

Trajectory prediction is the use of target motion information, such as position, velocity, acceleration, etc. to predict the target position in the next frame.

Trajectory prediction refers to uses the motion information, such as position, velocity, acceleration, to predict the target position in the next frame. Kalman filter algorithm [8] is a classic predictive estimation algorithm. It is often applied to target tracking, especially the trajectory prediction in occlusion. But this trajectory prediction is mainly applied to linear motion, does not apply to non-linear movement, or non-Gaussian assumption of dynamic systems.

Bayesian filter theory, models tracking problem into a Bayesian posterior probability distribution maximization problem. The core idea is to use some discrete random sampling points (particles) to approximate the probability density function of the state variables [9]. The algorithm can solve partial occlusion, short-term occlusion and analogue interference and other complex issues.

For occluded target, using a single tracking algorithm is often difficult to achieve perfect results, especially for the case of server occlusion. Therefore, the researchers consider integration of multiple algorithms to enhance the robustness of the algorithm. Paper [10] combines the Mean-shift algorithm, Sift algorithm and Kalman filter. It has a good tracking performance for the complex situation of distortion, occlusion, rotation, but the time complexity of the algorithm is high.

Aiming at the shortcomings of the above algorithms, this paper put forward the research of solving long time occlusion problem based on "object permanence". This algorithm does not need to know prior knowledge of the target's color, size, location and other characteristics, so that it has good flexibility and applicability.

2.1   The Principle of Object Permanence

Object permanence refers to the understanding of the objective world [11, 12] "Even if the eyes can't see, the objects also remain". Simple said is the object still exists after leaving the eye-sight. Applied to the occlusion problem refers to a complete occluded object will appear near the occluder.

2.2   Object Modeling

In the process of target tracking, effectively extract prospects information is the key to successful tracking. But the reality is, the complex background, constantly changing illumination, silhouette interference bring a great deal of difficulties for foreground object extraction [13]. To solve this problem, we need a way to dynamically establish and maintain a simple, common object models [14], and timely updates. Therefore, we use the foreground object extraction method based on background subtraction Gaussian mixture model, and use the ellipses approximate target contours.

For later convenience, definitions are as follows:

Contour ellipse $e = (c_x, c_y, \alpha, \beta, \theta)$. The (c_x, c_y) is the coordinates of the center point of the ellipse, which is the coordinates of the central position of the object. α, β, respectively, represent a major axis and a minor axis, θ represents the angle of the ellipse in the two-dimensional plane.

Foreground pixel $p = (x, y)$. The (x, y) represents the coordinates of the current pixel.

The space distance between the foreground pixels and the ellipse:

$$D(p,e) = \sqrt{\vec{v} * \vec{v}} \tag{1}$$

$$\vec{v} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \left(\frac{x - c_x}{\alpha}, \quad \frac{y - c_y}{\beta} \right)^T \tag{2}$$

3 The Application of Object Permanence for Long Time Occlusion

In the experiments, we assume that the camera is stationary. First, input video image sequence and use background subtraction method to separate foreground and background pixels. We can get different foreground pixels blobs. Then, obtain the relationship between the targets and foreground pixels blobs. Next, analysis the occlusion relationship based on "object permanence". Finally, update the object models and predict the targets positions.

3.1 The Relationship Between the Foreground Pixel and Object

The position relationship between the targets maybe independent of each other or exist occlusion. When each target motion is independent, it's easy to mark. But when interactions occur between moving targets, the problem of foreground pixels belonging to needs a reasonable allocation. Therefore, the foreground pixels are marked as follows:

Target pixels set inside the ellipse:

$$I(e) = \{p | D(p,e) \leq 1\} \tag{3}$$

Assume that in a period of time, obtaining M foreground pixel blobs $b_j (1 \leq j \leq M)$, tracking to N targets $e_i (1 \leq i \leq N)$. In case of occlusion, especially when full occlusion, a foreground blob may contain more than one target, but a target will not be recognized as several blobs, so M $\leq$ N. At this time, a foreground blob may relate to multiple targets. The relationship between the blobs and the foreground pixels can be found in the following four conditions, as shown in Fig. 1.
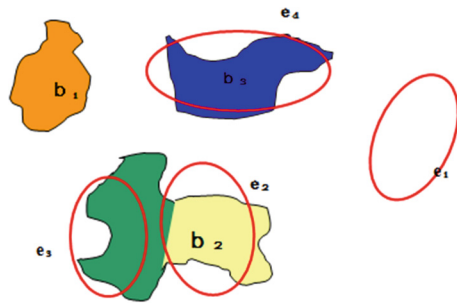

Fig. 1. The relationship between foreground pixel blobs and targets

(1) A foreground pixel blob doesn't relate to any target

$$\forall e_i, b \cap I(e_i) = \Phi \tag{4}$$

The formula (4) shows that the existing targets are not relate to the blob. Therefore, it is a new target appears in the video screen. Just like b_1.

(2) A target object is not relate to any foreground pixel blobs

$$\left(\cup_{j=1}^{M} b_j\right) \cap I(e) = \Phi \tag{5}$$

The formula (5) shows that the targets sets don't contain any detected pixels inside the blob. Therefore the target has disappeared, like e_1.

(3) A target associates with only one blob. As target e_4 and blob b_3.

(4) A blob associates with multi-targets. As target e_2 and blob e_3. They both scramble for the pixels in blob b_2 at the same time.

3.2   The Judgment of Occlusion Types Based on Object Permanence

In the occlusion problem, even in a simple partial occlusion situation, the relationship between targets and blobs is not just one to one. In order to solve the problem of missing and error tracking caused by occlusion, this paper based on "Object permanence" analyses the occlusion-ship between targets.

In the process of occlusion, target tracking relies on the spatial distribution information and appearance model information of targets. When two targets belong to different blobs, two objects began to scramble for the same blob of pixels, following the gradual process of occlusion. As shelter target, the original pixels have no obvious changes, still be marked for the objects. But the occluded target's number of pixels is gradually reduced. On account of the reduction of pixels is accompanied by the occlusion occurring. Therefore, introducing "occlusion rate" is to the quantitative characterization of the changes.

$$R_i = \frac{A_i}{A'_i} \tag{6}$$

A_i represents the current area of the target, A'_i is the area of a target before the occlusion. Occlusion rate is a standard when object e_i scrambles for the same blob of pixels with other objects. When occlusion rate is small, it shows that this target area is less than the area observed before. The occlusion has occurred. When occlusion rate is less than a certain threshold value, the target is completely occluded. At this time, the object e_i has disappeared.

$$R_i \leq T \tag{7}$$

Through the experimental tests, the threshold set to 35 % (T), can achieve good effect of distinction. When T is close to 0 %, which means a contour color of mis-classification will make the occluded target appearances again. It is a false judgment.

In fact, occlusion process is not only to determine whether the occlusion occurred, but also need to judge the shelter. The occlusion of two objects is evident. But for multiple objects occlusion, shelter judgment is very complicated. When multiple objects take place occlusion, the shelter should be located near the occluded target, and just occupy a partial of occluded target.

Among the many shelters, for each possible one, which occupies the largest part of occluded pixels, is considered to be the ultimate shelter. On the principle of "object permanence", before occluded targets appearing, we assume that the shelter completely covered the occlude targets, and they move together. When the occluded targets appear again near its shelter, the two objects separated gradually. The occluded pixels will reconstruct the target and displayed on the monitor screen.

3.3   Linear Prediction

The above mentioned method is based on the spatial distribution relationship between objects and blobs. It also can predict trajectories according to the position of target contour ellipse linear equation [15]. The judgments of targets need to mark the contour of different colors and central algorithm tags. Therefore, accurate tracking of a target need to ensure the same contour color, labels, accurate positioning in the entire process of tracking.

4   The Results and Analysis

An ideal multi-target tracking algorithm evaluation should include the following four aspects:

(1) It can accurately detect the position of each target;
(2) It can keep continuous tracking of the target;
(3) Each target corresponds to a unique ID;
(4) It has certain robustness to occlusion.

In order to verify the effectiveness of the proposed algorithm, we test respectively for outdoor, indoor, single pedestrians, multi pedestrians and standard video library CAVIAR pedestrian motion video. In this experiment, we use the windows 7 operating system as the platform, use the C++ language coding and testing in the visual studio 2008 editor. The hardware platform is the Intel core i5 CPU, 2.9 GHz, 2 GB memory.

Figure 2 is a single outdoor movement. It is the results for the pedestrian was full occlude by a tree after a period of time. The target disappeared for 102 frames. Figure 3 is the outdoor single movement process. It is tracking results of the target after a long time occlusion by a large vehicle. Figure 3(b) shows the target re-emerges after

disappearing of 241 frames. This program can accurately identify the original target after long time occlusion.

Figure 4 shows the tracking results of opposite interior movement of totally occlusion after for some time. Figure 4(b) frame 498: two targets in the first encounter.
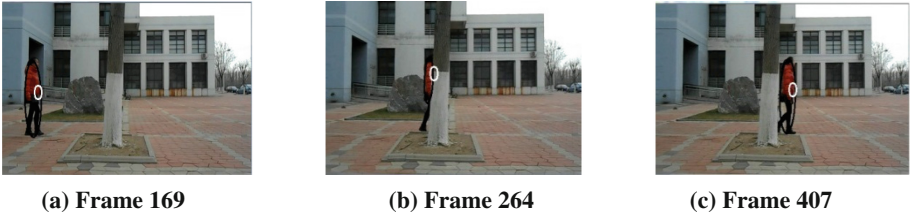

| (a) Frame 169 | (b) Frame 264 | (c) Frame 407 |

Fig. 2. Pedestrian occlude by a tree

| (a) Frame 154 | (b) Frame 498 | (c) Frame 537 |

Fig. 3. Pedestrian occlude by a large vehicle

| (a) Frame 154 | (b) Frame 498 | (c) Frame 537 |
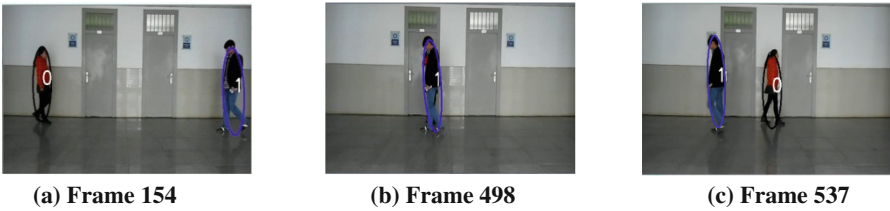
Fig. 4. The opposite interior movement

| (a) Frame 426 | (b) Frame 457 | (c) Frame 469 |
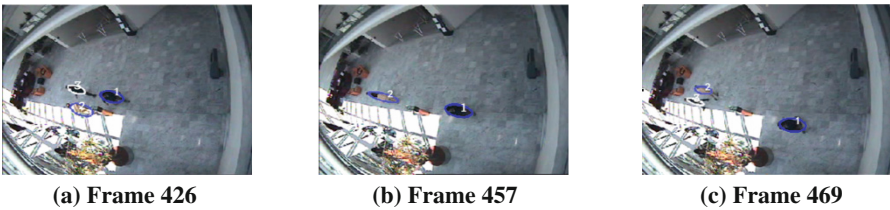
Fig. 5. Meet_Walk_Split

Table 1. Statistics results

Tracking algorithm	Target acquisition	Accurate rate		Error rate		
		No occlusion	With occlusion	Loss rate	Misjudgment rate	False matching rate
Mean shift	Manual	62 %	10 %	13 %	15 %	7 %
Object permanence	Auto	79 %	75 %	2 %	4 %	2 %

Target 0 is full occluded by target 1. Then target 0 disappears for 85 frames. Figure 4(c) shows the correct tracking results of the two targets after separation.

Figure 5 is the tracking results of "Meet_Walk_Split" in video standard video library CAVIAR. The results indicate that the algorithm can track the three people accurately after they met. Figure 5(b) shows the 457th frame, target 2 fully occludes target 3. They walked together for 10 frames and then separated. The recognition effect of this algorithm is shown in Fig. 5(c).

Table 1 shows the statistical average track of time, accuracy rate, error rate and other information of the results. Experimental results show that the algorithm is simple, high efficiency, and realizes real-time tracking multiple targets. In the target tracking process, this algorithm has advantages such as lower error matching rate. It has better solution of long time occlusion, error tracking of similar targets and illumination mutation problem.

5 Conclusion

This paper presents an algorithm based on "object permanence" to solve the occlusion in a longer time and larger space scope. The method can accurately track targets in and out of the monitor screen. In the process of tracking, it can dynamically build the appearance and motion model of the targets. The algorithm successfully uses object permanence to solve occlusion. The experimental results have shown the superiority of the proposed algorithm in solving the problem of target matching after long time occlusion. It can be used in indoor, outdoor surveillance, video capture, and has a wide range of application.

References

1. Zhang, H., Hu, S., Yang, G.: Video object tracking based on appearance models learning. J. Comput. Res. Dev. **52**, 177–190 (2015)
2. Peng, B.: Research on Occlusion in Image Sequence Tracking. Xidian University, Xidian (2014)

3. Zheng, K., Huang, W., Zhang, Z.: Occlusion handling in moving object tracking system. Comput. Eng. Des. **11**, 2816–2818 (2009)
4. Xue, C., Zhu, M., Liu, C.-X.: Review of tracking algorithms under occlusions. Chin. Opt. **2**, 388–393 (2008)
5. Ouyang, N., You, J.-H., Mo, J.-W., Zhang, T.: Tracking multiple objects in occlusions. Appl. Res. Comput. **27**, 1984–1986 (2010)
6. Ning, J., Zhang, L., Zhang, D., et al.: Scale and orientation adaptive mean shift tracking. Comput. Vis. IET **6**, 52–61 (2012)
7. Sun, Y.-Y., He, X.-H., Song, H.-Y.: A block-matching image registration algorithm for video super-resolution reconstruction. Acta Autom. Sin. **37**, 37–43 (2011)
8. Du, C., Liu, W.-N., Liu, L.: Target tracking algorithm based on kalman filter and particle filter. Chin. J. Liq. Cryst. Disp. **26**, 384–389 (2011)
9. Xia, S.-Z., Liu, H.-W., Jiu, B.: A method of relay of tracking based on bayesian theory. J. Electron. Inf. Technol. **33**, 652–658 (2011)
10. Chen, A.H., Meng, B., Zhu, M.: Multi-pattern fusion algorithm for target tracking. Opt. Precis. Eng. **17**, 185–190 (2009)
11. Zhang, C-F., Lin, Y-H.: Discussion of piaget object permanence. Psychol. Explor. **4**, 11–15 (1988)
12. Zeng, Q., Dong, Q., Tao, S.: The research object of the permanent mechanism for the development of infant. Acta Psychol. Sin. **4**, 394–399 (1997)
13. Lu, H., Li, H.-S., Fei, S.-J., Li, W.-C.: Adaptive method for tracking mobile object in occlusion. Comput. Eng. Des. **33**, 2343–2346 (2012)
14. Li, J.-Y., Xu, L.-Y.: A multi-target tracking algorithm combined with occlusion segmentation. Telecommun. Eng. **2**, 172–176 (2013)
15. Chen, X.-H, Zhu, S.-H.: Moving objects detection algorithm based on visual attention mechanism and gaussian mixture model. In: 2011 International Conference on Signal, Image Processing and Applications, vol. 21 (2011)