

PAPER • OPEN ACCESS

A Research on Dynamic Target Tracking with Camera Based on Kalman Filter

To cite this article: Yimeng Ma 2020 *J. Phys.: Conf. Ser.* **1453** 012106

View the [article online](#) for updates and enhancements.

You may also like

- [Enhancing the dynamic range of ultrasound imaging velocimetry using interleaved imaging](#)
C Poelma and K H Fraser
- [Fast Aerial UAV Detection Using Improved Inter-frame Difference and SVM](#)
Li Xiaoping, Lei Songze, Zhang Boxing et al.
- [A Comparative Study of Plastic and Glass Optical Fibers for Reliable Home Networking](#)
Hussein K. Manea, Yousif N. Molood, Qussay Al-Jubouri et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

A Research on Dynamic Target Tracking with Camera Based on Kalman Filter

Yimeng Ma^{1*}

¹Woodward Academy, College Park 30337, United States

*Corresponding author's e-mail: 20yma@woodward.edu

Abstract. With the popularization of the camera, all kinds of video network communication technologies have developed towards civil use and become increasingly popular. For household application, a common PC is used to obtain the frame data from the home network, while the constant frame is set as the background. The update function is used to update background in real time, and thereby forming an inter-frame relation matrix. Based on the matrix, Kalman filter can match and track targets. This method can effectively complete the calculation on the PC, perform precise detection and deal with the complexity of the household application. Finally, with the test results, this paper further proves that the method is feasible and can be used in actual situations.

1. Introduction

With the popularization of the camera and the development of various video network communication technologies, the camera is not only used in the special fields such as public security, traffic, security and protection, and military, but also tends to be civil and popular. Both the camera and the video technology are also increasingly applied in civil and commercial fields, such as supermarkets, households, exhibitions, shopping malls, and other public and private places. The main purpose of the camera is to monitor and record video images as well as provide upper application functions. However, the analysis of the abnormal conditions in the video is undertaken by human beings, the efficiency will become very low, while the cost will be very high. Therefore, it is impossible to adopt this approach on a large scale.

One of the most fundamental tasks of video analysis is to track dynamic objects, generally known as the detection and tracking of moving objects, which can support the upper applications, such as abnormal behavior detection, target classification, target comparison and other upper application requirements. However, due to some factors in the application of cameras at home, including the complexity of application circumstance, large differences in light, and great dynamic changes of the detected and tracked targets, the target detection is more difficult and the tracking of moving targets may be occluded. At present, multiple sets of camera systems are used to deal with occlusion among targets, but generally, there is only one camera system in a household. In addition, EM algorithm is applied to track targets in accordance with the color information model, the defined area, the defined scope of the model, or the contour feature model ^[1-4]. However, these methods attach much attention to defining the scope of the model and scene, and require powerful computing ability of devices ^[5-7]. Accordingly, for the complexity of the household application, the above method is unavailable.

This paper aims to meet the daily needs of households, and make use of the existing PC, router, camera and online Cloud to detect and track targets with fixed cameras in order to achieve remote target tracking.



2. Solutions

In most cases, the camera is installed at a fixed position in a household. Family members, visitors, people who may appear in the range of the camera, and family pets are the main detection and tracking targets. For the above scene, this paper takes two perspectives into account. On the one hand, this paper designs an algorithm to detect and track targets. On the other hand, data can be shared to the Cloud through home computer and network devices, and then to the mobile phones of family members through the Cloud. As a result, remote monitoring of the family situation can be realized.

In view of the above scene, the key is to extract the current frame of the camera and take the preset fixed area as the background. At the same time, an update function is introduced to dynamically update the background in real time. Based on the time continuity among frames, the target relation matrix can be established, and then the target can be predicted by Kalman filter. By doing so, the target can be tracked with the continuous frames of the fixed camera even if the targets are occluded or the light is deficient. Since Kalman filter features easy implementation and low footprint, it is feasible to the home PC application and can effectively track the targets in real time.

With the connection of the WIFI or the wired network to the home router, the home video monitor can be connected to the home PC. The home PC can support the target tracking algorithm and share the calculation results and video data to the Cloud. Users can connect to the Cloud through the mobile network, and thereby having access to the remote application of the algorithm and the remote monitoring through mobile phones.

3. Design of the Algorithm

In this paper, the design of the tracking algorithm follows two steps. First, dynamic targets need to be detected. Second, the moving state of the targets should be predicted to complete target tracking.

As the home camera is usually installed at a fixed position, the background is relatively fixed with less change. For this reason, the fixed background difference method can be used. With the static region image as the background, the moving target can be detected based on the difference of the current frame image matrix and the background image matrix.

The target tracking can be regarded as a model-based matching. The model consists of some relevant movement features of the targets in the continuous frame, including color, shape, speed, position, and so on. In an image, there may be a single target or multiple targets. For a single target, the tracking can be simplified as a trajectory tracking. However, multiple targets cannot be simply regarded as the accumulation of several single targets because the occlusion among those targets may lead to a temporary disappearance of the target, resulting in the failure of the target tracking model. With the ability to predict the parameters and the target position, Kalman filter can predict the occluded target, and thereby track multiple targets.

The motion of the targets can be considered as uniform motion and can be expressed as follows based on the kinetic formula, in which Δt refers to the interval of two frames.

$$\begin{aligned}x_t &= x_{t-1} + (\Delta t)v_{t-1} \\ v_t &= v_{t-1}\end{aligned}$$

As the system is a linear dynamic model, it can be expressed as follows according to Kalman filter:

$$\begin{aligned}X(t) &= AX(t-1) + W(t-1) \\ Y(t) &= CX(t) + V(t)\end{aligned}$$

In the equation, $X(t)$ and $X(t-1)$ respectively refer to the state vector at time t and $t-1$; $Y(t)$ is the observation vector at time t ; W is the system noise, which can be assumed to follow Gaussian distribution; V represents the observation noise with the mean value of zero, which also follows Gaussian distribution.

With the basis of the above formula, the algorithm can be divided into three steps.

- 1) Find the target in the current frame, and capture the target with a rectangular box, and label the target according to the sequence.
- 2) Calculate the intersection area of the target box, that is, the definitive overlapping area;
- 3) Based on the prediction of Kalman filter, the motion trajectory can be obtained by checking

whether the current frame and the previous frame can match with each other.

4. Design of the system

This section mainly focuses on the application of the designed algorithm for the home camera to the specific system. Taking the home environment and application feasibility into account, the network architecture composed of the home environment, the Cloud and the mobile phone terminal can be designed, which is shown in Figure 1.

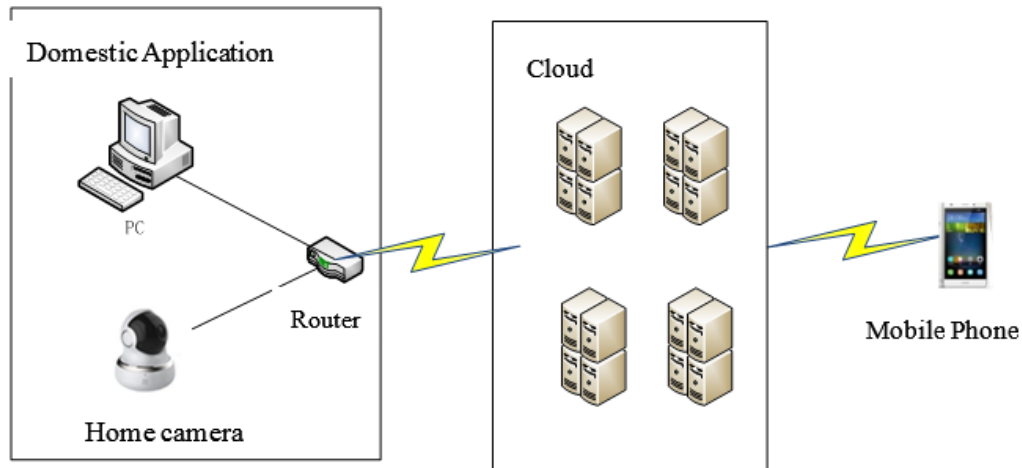


Figure 1. Network architecture of dynamic target tracking system with a home camera

The system, illustrated above, allows mobile phone terminals and home cameras to share data on the Internet. In households, users can arrange the home webcam in any place that needs to be monitored, and then the camera can transmit the monitoring video frame to the home PC through the home router. The video frame can be analyzed by the video analysis software installed by the users on the PC, and

then the targets will be tracked by the designed algorithm module. In addition, according to the threshold preset by the users, the algorithm will analyze the abnormal video and generate abnormal messages. The abnormal messages will be first sent to the video data sharing system on the Cloud through the home router, and then forwarded to the designated user's mobile phone terminal by the Cloud. Consequentially, the user can check the message on the mobile terminal and monitor abnormal behaviors based on the video clip shared on the Cloud.

In conformity with the design of the above network architecture and the physical partition of each communication module, a target tracking application software should be installed on the PC, and the function of receiving abnormal messages and checking abnormal video clips should be added on the mobile phone. Therefore, two software modules, consisting of a PC-side application and a mobile application, should be designed.

The design of the PC-side application can be divided into three layers: access layer, algorithm layer, and application layer from bottom to top, which is shown in the following figure:

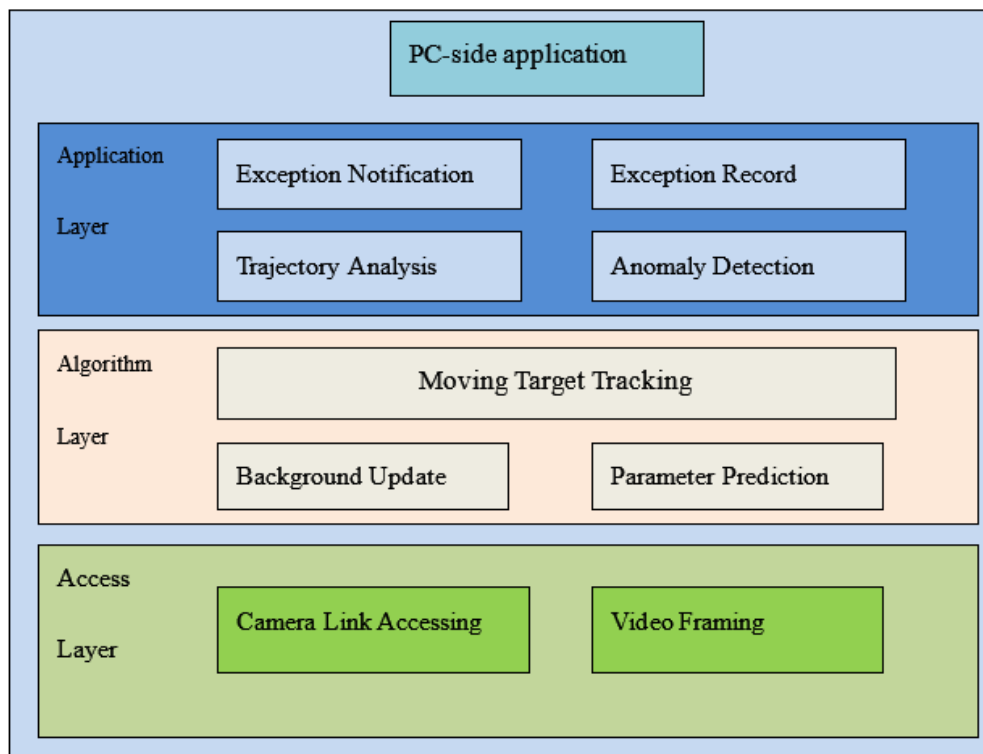


Figure 2. Framework design of the PC-side application software

The access layer is mainly responsible for the TCP/IP of the camera and the video coding stream, which is completed through the camera link accessing module. After that, the video framing module is used to decode video data stream and split video according to the frame.

The algorithm layer mainly consists of a background update module and a parameter prediction module. According to the algorithm designed above, the background update module can analyze the background with the method of fixed background difference, and continuously update the background based on the binary noise reduction. After that, the parameter prediction module can predict the parameters of the moving targets with Kalman filter. Finally, the moving target tracking module can generate the target relation matrix, and thereby tracking the moving targets.

The application layer provides an interactive UI interface for users and carries out track analysis, anomaly detection, exception notification, and exception records.

The design of the mobile application terminal mainly aims to make users receive abnormal messages and access to the data and video. To be specific, the design consists of four modules, as Figure 3 shows.

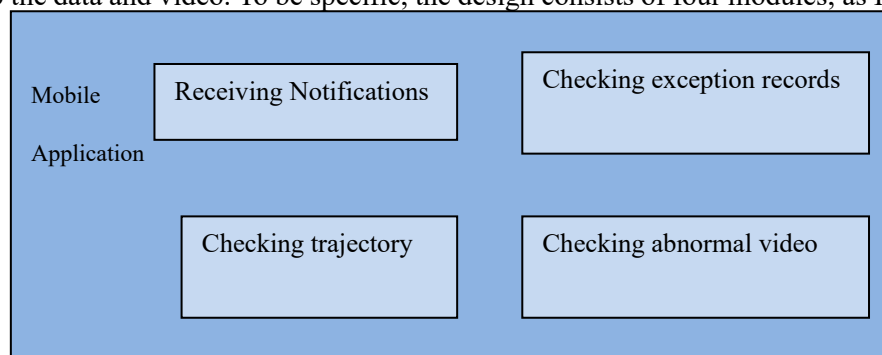


Figure 3. Framework design of the mobile terminal application

The receiving notification module is designed for regularly receiving messages from the Cloud, while the exception records are used to check whether there is any abnormal case. Besides, by checking the

trajectory, the moving trajectory of the target can be acquired. Additionally, users can also check the video clip with exceptions. With the above modules, users can remotely check the video of abnormal cases without too much data flow when they are out.

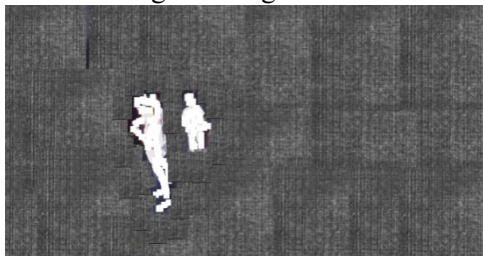
5. System Test

In terms of the PC running environment, the system test is carried out on an Intel i3 processor with 4-core, 1.5GHz main frequency, 4G memory, 500G hard disk drive, 100M network interface card, and integrated graphics. Such the test environment can be feasible for most home PCs. At the same time, the camera with HD 1080P resolution and both WIFI and cable network can shoot 30 frames per second. Apart from that, for the router, a 100M optic modem is adopted in order to support locally real-time monitoring.

As shown in Figure 4, the photo (a) is the original image of the 238th frame, the photo (b) is the binarization of the original 238th frame with the method of difference, and the photo (c) is the target detection result of the 238th frame.



(a) The original image of the 238th frame



(b) The binarization of the 238th frame with the method of difference



(c) The target detection result of the 238th frame

Figure 4. Background update test

Since the position of the camera is fixed, the background will also not change. Under this circumstance, the binarization based on the method of difference can quickly mark the targets. With the unchanged target background and the real-time updated non-target region, the error will be minimized. Hence, the accuracy of the moving target detection will also be promoted.

When the targets in the image occlude with each other, just as the figure below shows, the system will first identify targets, and then mark them in accordance with the sequence. Figure 5 (a) shows the image of the 213th frame, in which the three targets are respectively marked with the number 1, 2, and 3. In the image of the 238th frame shown in Figure 5 (b), two moving targets overlap with each other, resulting in occlusion. For this case, the identification box will be marked with the smallest number--1. In Figure 5 (c) showing the image of the 257th frame, the target 2 moves out of the monitoring region,

while only the target 1 and the target 3 are still there, indicating that the target can also be tracked after occlusion.

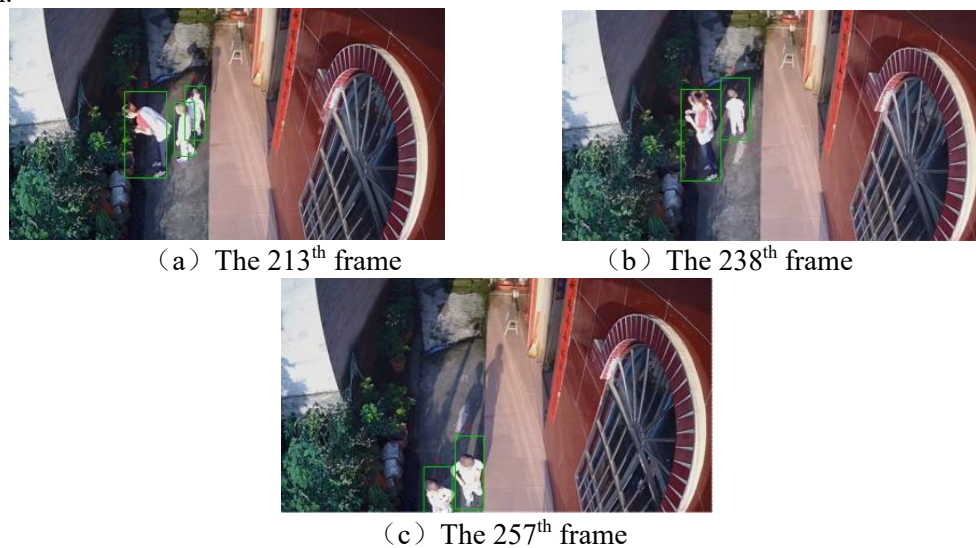


Figure 5. Occlusion Detection Test

6. Conclusion

Focusing on the application of cameras at home, this paper designs a dynamic target tracking system for detecting abnormal targets with the home camera. For household application, a common PC is used to obtain the frame data from the home network, while the constant frame is set as a background. The update function is used to update background in real time, and thereby forming an inter-frame relation matrix. Based on the matrix, Kalman filter can match and track targets. This method can effectively complete the calculation on the PC, perform precise detection and deal with the complexity of the household application. What's more, with this method, even if the target is occluded and moves out of the monitoring rejoin, its trajectory can still be predicted effectively. Moreover, when the target appears in the monitoring area at an abnormal moment, the system will push the warning message and share video clips through the Cloud so as to detect the abnormal cases with the home camera. Finally, this paper illustrates the test results, further proving the feasibility of this method for actual applications.

References

- [1] 1-Point RANSAC for extended Kalman filtering: Application to real-time structure from motion and visual odometry [J]. JavierCivera,Oscar G.Grasa,Andrew J.Davison,J. M. M.Montiel. J. Field Robotics. 2010 (5)
- [2] Drift-Free Real-Time Sequential Mosaicing [J]. Javier Civera, Andrew J. Davison, Juan A. Magallón,J. M. M. Montiel. International Journal of Computer Vision. 2009 (2)
- [3] Automatic scene structure and camera motion using a catadioptric system [J]. Maxime Lhuillier. Computer Vision and Image Understanding. 2007 (2)
- [4] Vision-Based SLAM: Stereo and Monocular Approaches [J]. Thomas Lemaire,Cyrille Berger,Il-Kyun Jung,Simon Lacroix. International Journal of Computer Vision . 2007 (3)
- [5] Research on Global Localization of Home Robot Based on Single Camera [D]. Ju Yujiang. Harbin Institute of Technology, 2006.
- [6] A Visual Slam Algorithm Based on Monocular Vision [J]. Wen Feng, Chai Xiaojie, Zhu Zhiping, Dong Xiaoming, Zou Wei, Yuan Kui. Journal of Systems Science and Mathematical Sciences. 2010(06).

- [7] Review on the Achievements in Simultaneous Localization and Mapping for Mobile Robot Based on Vision Sensor [J]. Sun Fengchi, Huang Yalou, Kang Yewei. Control Theory and Application. 2010(04).