

Research on the Collaborative Tracking system for Moving Target in multi-camera surveillance system based on space-time video fusion

Gang Zhao¹, YaLi Yu¹, Liqin Yu^{2,*}, Qi Wang¹

¹School of Educational Information Technology, Huazhong Normal University, Wuhan, Hubei, China. 430079

² National Engineering Research Center for E-Learning, Huazhong Normal University, Wuhan, Hubei, China. 430079

^aemail: zhaogang@mail.ccnu.edu.cn, ^bemail: 13007102880@163.com,

^cemail: yu_liqin_hope@163.com, ^demail: 1191879849@qq.com

*Corresponding author Email: yu_liqin_hope@163.com

Keywords: Collaborative Tracking; space-time video fusion; multi-camera surveillance system

Abstract. With the increasing demand of social security, intelligent moving target tracking technology play a more important role in video surveillance system. However, current moving target tracking technology works well for single camera situation but always lost targets in multi-camera environments due to the difficulty of object matching. In this paper, we proposed a new collaborative tracking system for moving target based on space-time video fusion. The idea of this system is firstly to register and fuse real-time surveillance video to 3D virtual model with situation awareness, then the space relativity between different cameras and the time relativity of tracking target moving across multi-cameras are extracted. With the help of this space-time information, we can accurately match the moving targets between consecutive cameras and coordinate adjacent cameras to collaboratively track the target by using moving target tracking for single camera. Test results show that our approach can realize continuous tracking for moving objects between multi-cameras with good robustness.

1. Introduction

With the increasing demand of social security, video surveillance systems are widely deployed. Intelligent technologies such as intelligent moving target tracking technology play a more important role in video surveillance system. As a part of intelligent video surveillance systems, moving object tracking technology includes single camera tracking and multi-camera tracking. Moving object tracking technology in single camera situation has been well research. Zhou S.K. proposed an adaptive particle filters method for visual object tracking with more accuracy [1]. Comaneci. D used meanshift methods to realize moving object tracking[2]. Although many methods works well for single camera but always lost targets in multi-camera environments due to the difficulty of object matching.

The key problem of moving object tracking in multi-camera environments is the object matching between different cameras. However, it is largely affected by the complex out-door situations such as environment illumination, camera pose, occlusion and other noises. The imagery of the same object in different cameras will vary largely. So finding strong image features to realize robust object matching between different cameras is studied. A method based region SIFT feature extraction has been proposed to solve the object matching problem with shooting angle limitation[3]. Zhou Q integrated several image features such as color feature, contour feature to realized more robust object matching with computational complex cost[4]. Guo Y used several image features to realize vehicle tracking in multi-camera system[5]. Huanxi Liu proposed an adaptive feature integrated method for object matching including color histogram, UV chrominance, spectrum and SIFT[6].

In general, strong image features or integrated feature for object matching are still not found for

all situations. To solve the above problem, we proposed a new collaborative tracking system for moving target based on space-time video fusion. The idea of this method is firstly to register and fuse real-time surveillance video to 3D virtual model with situation awareness, then the space relativity between different cameras and the time relativity of tracking target moving across multi-cameras are extracted. With the help of this space-time information, we can accurately match the moving targets between consecutive cameras and coordinate adjacent cameras to collaboratively track the target by using moving target tracking for single camera. Our approach can realize continuous tracking for moving objects between multi-cameras with good robustness.

2. Collaborative Tracking system for multi-camera Surveillance Based Video fusion

A. Video Fusion for 3D panoramic surveillance video

Traditional video surveillance system independently displays surveillance video images on several monitors. However such system cannot provide enough situation information of every camera in real-time and flexible way. The operators of traditional video surveillance system cannot link the video content and video situation to analyze and estimate the security state in time. This enormously limited the ability and efficiency of video surveillance system especially for real-time use cases such as suspicious target tracking.

With the development of computer technology such as augmented reality, the fusion and interactive between virtual world and reality world are available. Virtual 3D models of the city, buildings and regions give us a virtual observation method of real world, meanwhile provide actual situation and measurement information. Integrating real surveillance video with virtual 3D models can solve the space-time coordination problem of traditional video Surveillance System. In this paper, we use Surf-based image matching and projective texture mapping method to register and fuse SURF [7] to 3D virtual model in order to form a broader monitoring screen than a single camera sight. This method consists of three steps:

1) SURF feature extraction and matching

We can firstly use equation (1) to find the situation of SURF feature point[7] both for real surveillance video image and texture image of the 3D model. Based on the SURF feature points, Harr wavelet is used to get SURF feature description vectors.

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix} \quad (1)$$

In which, $X = (x, y)$ is a point of Image I, σ is the scale of Image I, $L_{xx}(X, \sigma)$ represents the convolution of two order partial derivative of Gaussian function at point X and Image I.

Then we use SURF feature description vectors to initial a matching procedure between the real video images and texture image and find the geometric transformation model between the two images.

2) Camera calibration

In this paper we use linear imaging model as equation (2) for camera calibration. Self-calibration method by using Matlab Tollbox and Surf feature based matching in step 1) is executed to compute the intrinsic parameters and extrinsic parameters of the surveillance camera respectively.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (2)$$

In which, $(u, v)^T$ is the point coordination of image plane, $(x_w, y_w, z_w)^T$ is the point coordination in the real world coordinate system. M_1 and M_2 are the intrinsic parameters matrix and extrinsic parameters matrix of the camera respectively.

3) Projective texture mapping

After getting the intrinsic parameters matrix and extrinsic parameters matrix of the camera, we can then use equation (3) to projecting the image point of the real video to the 3D model as texture point[8].

$$\begin{bmatrix} s \\ t \\ r \\ 1 \end{bmatrix} = \begin{bmatrix} Normal \\ Matrix \end{bmatrix} * M_1 * M_2 * \begin{bmatrix} ModelView \\ Matrix \end{bmatrix} * \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \quad (3)$$

B. Collaborative Tracking system for multi-camera Surveillance

Based on the video fusion technology for 3D panoramic surveillance video presented previously, we proposed a collaborative tracking system for multi-camera surveillance as illustrated in Figure 1

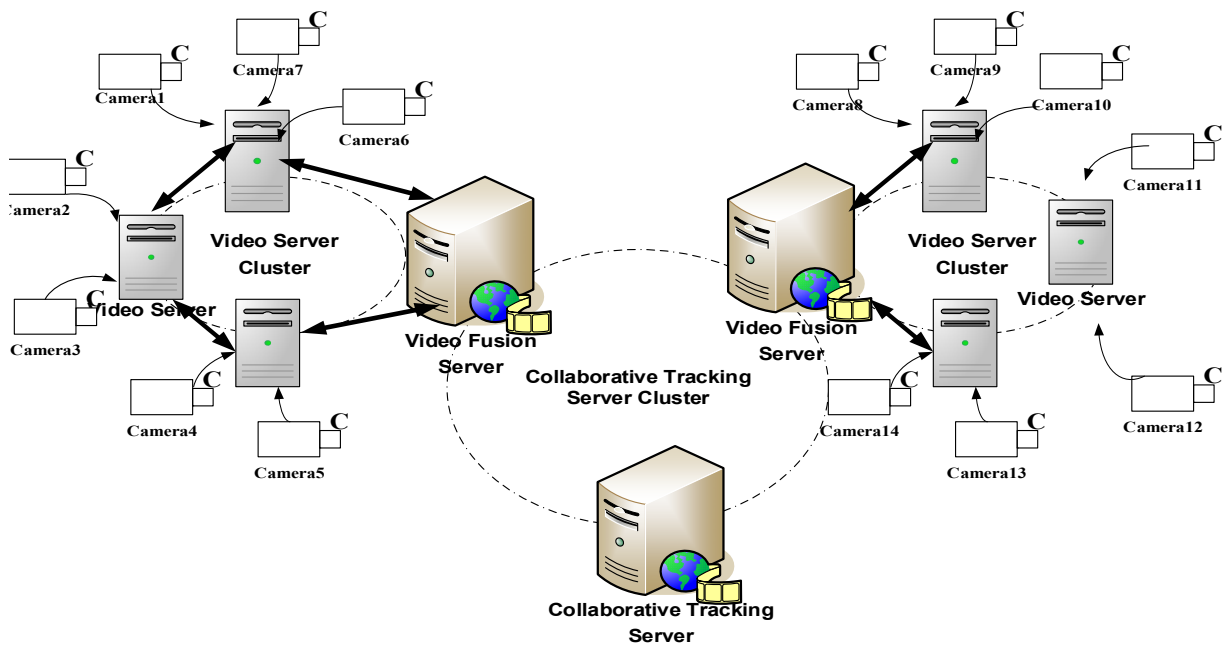


Figure 1. The system architecture of collaborative tracking system for multi-camera surveillance

This system consists of server parts: video server cluster, video fusion server and collaborative tracking server. The video server cluster may include several video server. Each video server responds to the encoding, decoding and management for a number of surveillance videos. A video fusion server is designed to fusion the videos with the whole or part of region 3D model covered by the video server cluster. Considering the complexity and overload of video fusion computation, several video fusion servers maybe form a video fusion cluster to realize the video fusion of a whole surveillance region. Collaborative tracking server can receive related information from the video fusion servers and initial a suspicious target tracking if necessary, which will be presented in Chapter III in detail.

To utilize the space relativity between different cameras and the time relativity of tracking target moving across multi-cameras, the related information received by the collaborative tracking server should include not only the real stream of the surveillance video but also the position of every surveillance camera, the pose of every surveillance camera, the speed and direction of suspicious target in single camera which can be extracted by traditional single-camera tracking methods.

To store the above related information, server databases are designed such as video fusion parameter database, space information database of the surveillance camera and moving information database of moving objects not shown in Figure 1.

3. Collaborative Tracking method for Moving Target based on space-time video fusion

In the collaborative tracking system for multi-camera surveillance, a new tracking method for moving target based on space-time video fusion has been investigated. This method is illustrated in two adjacent cameras with overlap area as example.

At first, extracting the background images of two cameras respectively. Then, detecting the matching points by using SIFT algorithm and Compute the homographic matrix of two background images. Homographic matrix is the coordinate's transformation of the same point on two images. At the same time, according to the homographic matrix can draw overlapping field of views between two cameras. Then, estimating whether the target enter into the public overlapping field of views. If not ,Continue to detecting and tracking in the camera 1;If the target enter into the public overlapping field of views, camera1 and camera 2 finish the target handover and continue to continue to track the target in the camera 2.The detailed tracking processes are illustrated in Figure 2.

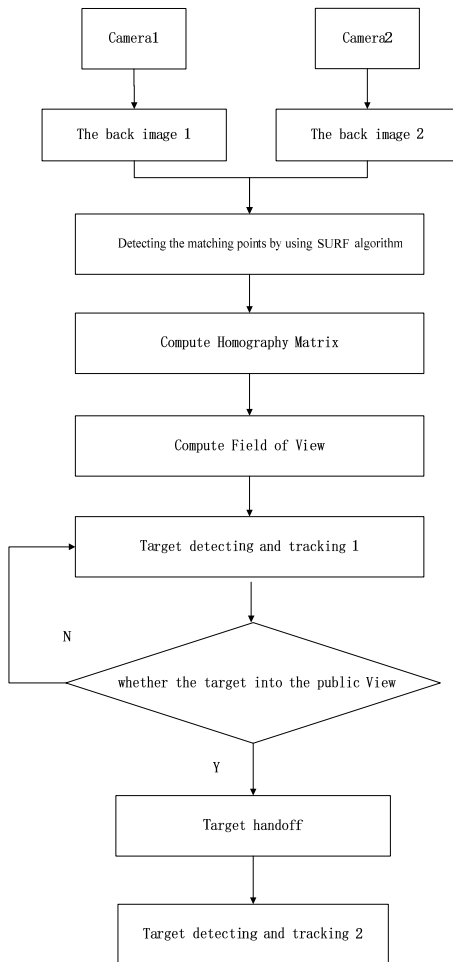


Figure 2. The process of target tracking under multiple cameras with Overlap area

As shown in the figure 3, an example of collaborative tracking is presented. The public overlapping field of views is marked by red line. There are three moving targets in the public view. When targets completely enter into the overlapping area of the camera, we can compute the coordinates of tracking box $P_1^1 (238, 190)$, calculate the projective location of point in another camera by using homographic matrix $\overline{P}_1^1 (14, 215)$. At the same time, there are three moving targets detected in another camera and marked by green box. The coordinates of moving targets are $P_1^2 (4, 217)$, $P_2^2 (68, 157)$, $P_3^2 (75, 207)$. Then, computing the Euclidean distance of two point to implement the target handover. the Euclidean distance $\|\overline{P}_1^1 P_1^2\| = 14.866$, $\|\overline{P}_1^1 P_2^2\| = 77.077$, $\|\overline{P}_1^1 P_3^2\|$

=61.400. The shortest distance is $\|\overline{P_1^1} P_1^2\|$. The best match is $P_1^1 \Leftrightarrow \overline{P_1^1} \Leftrightarrow P_1^2$, P_1^1 and P_1^2 is the same target. Then track the P_1^2 in the another camera.

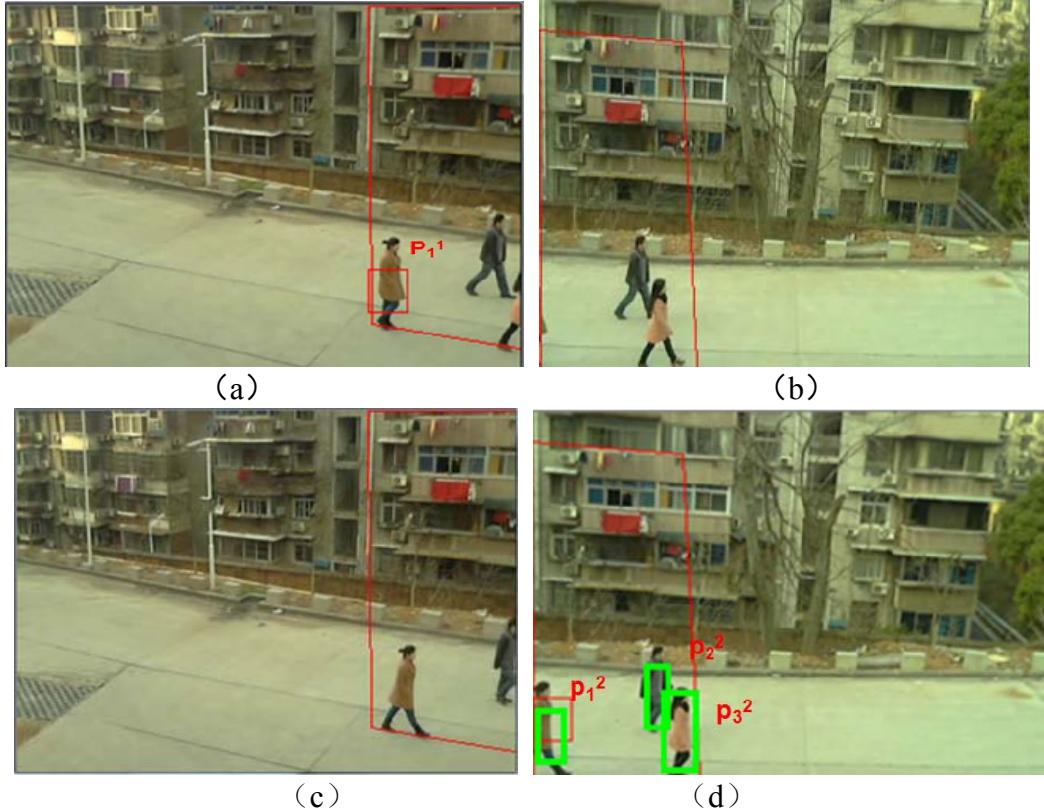


Figure 3 an example of collaborative tracking

4. Conclusion

In this paper, we proposed a new collaborative tracking system for moving target based on space-time video fusion. Based on registering and fusing real-time surveillance video to 3D virtual model with situation awareness, this system provides space-time information for object matching and a new collaborative tracking method for moving target in adjacent cameras with little overlap area has been studied. However collaborative tracking method for moving target in adjacent cameras with no overlap area remain further research.

Acknowledgement

This work is supported by the project “Research on Key technology of 3D Full-View Intelligent Surveillance System based on Space-Time video fusion and analysis” (No. 2014010101010025) granted by the Applied Basic Research Plan for Science and Technology of Wuhan City. This work is also financially supported by self-determined research funds of CCNU from the colleges’ basic research and operation of MOE(No.CCNU15A02020).

References

- [1] Zhou S K, Chellappa R, Moghaddam B. Visual tracking and recognition using appearance-adaptive models in particle filters[J]. Image Processing, IEEE Transactions on, 2004.
- [2] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2003, 25(5): 564-577.
- [3] Ming An-Long, Ma Hua-Dong, Region-SIFT Descriptor Based Correspondence Between

Multiple Cameras, Chinese Journal of Computer, 2008, 31(4): 650-661.

[4] Zhou Q, Aggarwal J K. Object tracking in an outdoor environment using fusion of features and cameras[J]. Image and Vision Computing, 2006, 24(11): 1244-1255.

[5] Guo Y, Hsu S, Sawhney H S, et al. Robust object matching for persistent tracking with heterogeneous features[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2007, 29(5): 824-839.

[6] Huanxi Liu, Xiaowei Lv. An Adaptive Feature-fusion Method for Object Matching over Non-overlapped Scenes[J]. IEEE CS. 2014.

[7] Herbert Bay, Tinne Tuytelaars, Luc Van Gool. SURF: speeded up robust features [A]. Proceedings of the 9th European Conference on Computer Vision[C]. 2006, 404-417.

[8] Mark Segal, P. Haeberli. Fast Shadows and Lighting Effects Using Texture Mapping[J]. Computer Graphics (SIGGRAPH'92 Proceedings), 1992, 249-252.