

Research Article

Moving Target Positioning Based on a Distributed Camera Network

Long Zhao,^{1,2} Zhen Liu,¹ Tiejun Li,¹ Baoqi Huang,² and Lihua Xie²

¹ School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China

² School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798

Correspondence should be addressed to Long Zhao; buaa_dnc@buaa.edu.cn

Received 14 February 2014; Accepted 18 April 2014; Published 13 May 2014

Academic Editor: Guoqiang Hu

Copyright © 2014 Long Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a systematic framework for moving target positioning based on a distributed camera network. In the proposed framework, low-cost static cameras are deployed to cover a large region, moving targets are detected and then tracked using corresponding algorithms, target positions are estimated by making use of the geometrical relationships among those cameras after calibrating those cameras, and finally, for each target, its position estimates obtained from different cameras are unified into the world coordinate system. This system can function as complementary positioning information sources to realize moving target positioning in indoor or outdoor environments when global navigation satellite system (GNSS) signals are unavailable. The experiments are carried out using practical indoor and outdoor environment data, and the experimental results show that the systematic framework and inclusive algorithms are both effective and efficient.

1. Introduction

The theory of navigation and positioning has been used in various application fields such as positioning equipment to monitor its working state and positioning a car or people to guide them to a certain place. In these applications, it is required that targets are positioned and tracked, which can be realized using the global navigation satellite system (GNSS) or GNSS aided inertial navigation system (INS). However, the GNSS system is subject to various limitations, the most critical one of which is jamming. GNSS signals are not always available due to the blockage of high buildings, canyons, and forests, among others. For this reason, a number of alternative technologies, including optical [1], radio [2–4], RFID [5], and acoustic [6], have been proposed for indoor and outdoor positioning systems over the years. Most efforts were focused on WiFi based localization which takes the advantage of WiFi infrastructures. By using the user's smartphone to measure the signal strength from multiple WiFi access points, the user's location can be constrained within a relatively small region in a large indoor environment. However, these systems rely on prepared infrastructures installed beforehand, and their accuracy critically depends on the number of available

access points, which results in certain restrictions [5]. In addition, many results related to the application of the camera based positioning systems have been reported in the last few years [1], including simultaneous localization and mapping (SLAM) [7] and visual odometry [8]. Although SLAM is becoming a standard technique for indoor robotic applications, it is still challenging to apply SLAM in large outdoor environments.

In recent years, the size of camera networks grows quickly with the development of building safe and smart city. These cameras can provide complementary positioning information for moving target positioning in both indoor and outdoor environments when GNSS signals are unavailable. In this paper, we will deal with the moving target positioning based on a distributed camera network in a three-dimensional space. In existing camera networks, static cameras are generally deployed to cover a large region, and moving targets can only be detected and tracked via certain algorithms running in a central supervisor unit, but their positions are not determined [9–12]. In some practical applications, the total number of cameras is usually restricted by various factors such as the cost and placement of cameras. To address this problem, multiple pan, tilt, zoom (PTZ)

cameras or the combination of PTZ cameras and static cameras can be deployed to fulfill some practical tasks [13–18].

The target detection, target tracking, and camera calibration are key to the moving target positioning process. To extract moving targets from a video frame of a static camera, background subtraction is the most widely used approach [19, 20]. When the camera is stationary, the background scene is unchanging, such that it is convenient to construct a background model [21, 22]. The capability of efficiently and accurately estimating background images is critical for any robust background subtraction algorithm. A well-known method presented by Stauffer and Grimson [21] uses an adaptive strategy for modeling background. Therein, each pixel is modeled using a separate Gaussian mixture, which is continuously learnt based on online approximations. Target detection at the current frame is then performed at pixel level by comparing its value against the most probable background Gaussians. However, the adaptive Gaussian mixture algorithm suffers from a low convergence speed in the learning process, especially in complicated environments. For this reason, an improved adaptive Gaussian mixture learning algorithm was introduced in [23].

Moving target tracking is an important component in the field of computer vision and has been widely used in many applications, such as video surveillance [16], intelligent transportation [24], and multiagent systems tracking and control [25]. Target tracking aims to estimate the position and the shape of a target or a region in subsequent frames. During target tracking, a target is continuously tracked by correctly associating a target detected in subsequent frames with the same identified track. These methods and their variations commonly make use of the one-to-one assumption in the sense that a target can only generate at most one measurement in each frame and a measurement can only originate from at most one target. However, the one-to-one assumption rarely holds in practical applications due to the splitting and merging processes as well as multitargets existing in a common scene. In order to overcome these shortcomings, several approaches have been proposed for multitargets tracking in recent years [26–29].

Camera calibration is an essential procedure in distributed multitarget positioning and determines the mapping between 3D world coordinates and 2D image coordinates in practical applications. The basic task of camera calibration is to compute the camera extrinsic and intrinsic parameters which determine the imaging model and the relationship between multiple camera coordinates. With respect to different applications, the corresponding calibration algorithms include the direct linear transformation (DLT) algorithm [9], Tsai algorithm [10], vanishing point algorithm [11], and Zhang algorithm [12]. These algorithms have respective advantages and disadvantages in various practical applications. In this paper, we will focus on a fast calibration algorithm based on the vanishing point theory, which overcomes the defects of traditional measurements.

The paper is organized as follows. The system framework is presented in Section 2. Section 3 focuses on the target detection and tracking. The fast calibration algorithm is

presented in Section 4. The test results of target positioning based on a distributed camera network are reported in Section 5. Finally, we draw some conclusions and shed light on future work in Section 6.

2. Systematic Framework and Problem

The work presented in this paper originates from a research project of moving target tracking and positioning in the Digital Navigation Center (DNC) at Beihang University. The primary goal of the project is to develop a target positioning platform to realize monitoring and positioning targets in a large region. The systematic framework of the moving target positioning based on a distributed camera network is shown in Figure 1. Due to the field of view and price limitations, a mass of static cameras are installed in a practical application environment. In order to realize moving target positioning in a large region, it is necessary that the system supports targets detection and tracking. Since a target can no longer be detected, because either it leaves the field of view, it stops and becomes static, or it can no longer be distinguished from the background, it is reasonable to take target splitting and merging into account or detect multiple targets. Therefore, the performance of the target detection, tracking, and association algorithms will influence the reliability of the target positioning, and it is necessary that the target positioning results between cameras are fused into a world coordinate system $X_w Y_w Z_w$.

In this paper, we tackle several problems including the targets detection, tracking, and association, as well as the fast camera calibration and target positioning in the moving target positioning system based on a distributed camera network in practical applications.

3. Target Detection and Tracking

3.1. Target Detection. Target detection is the basis of target tracking, target positioning, target recognition, action recognition, and so forth. There are some common algorithms such as the optical flow algorithm [30], the frame difference algorithm [31], and the background subtraction algorithm [32] in practical applications. The most well-known and the most widely used one is background subtraction for static cameras because it is convenient to construct a background model and extract moving targets.

The background modeling techniques can be divided into two categories: the parametric techniques that use a parametric model for each pixel location and the samples-based techniques that build their model by aggregating previously observed values for each pixel location [33]. The most popular parametric technique is based on the Gaussian mixture model (GMM) presented by Stauffer and Grimson [21]. This algorithm relies on the principle that the pixel value for the same location in the image sequences satisfies a Gaussian distribution, as illustrated in Figure 2.

While updating the background model, each pixel of a scene image is independently modelled by a mixture of at most K Gaussian distributions and employs an adaptive

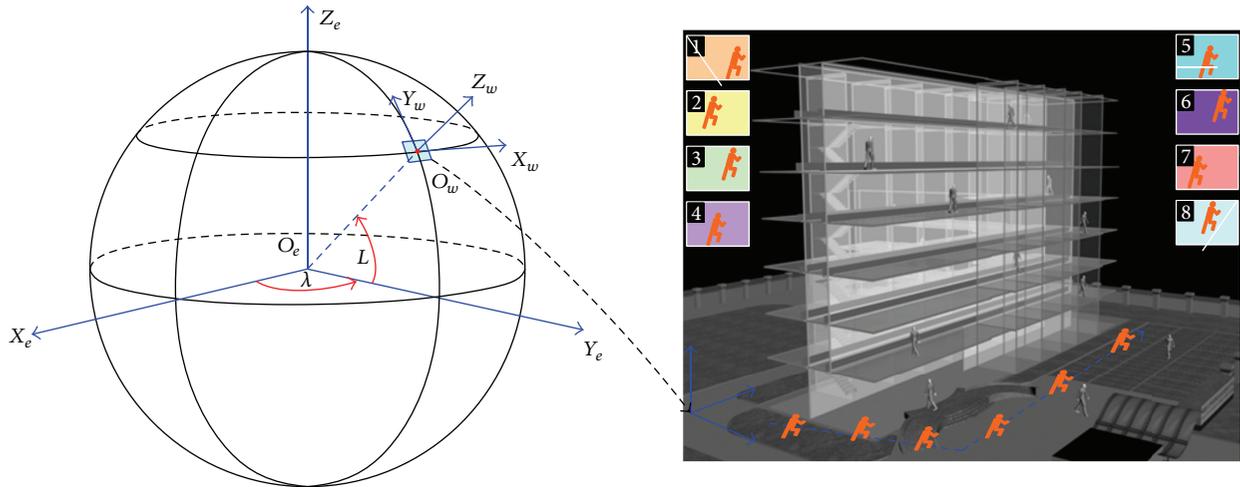


FIGURE 1: The systematic framework of the moving target positioning based on a distributed camera network.

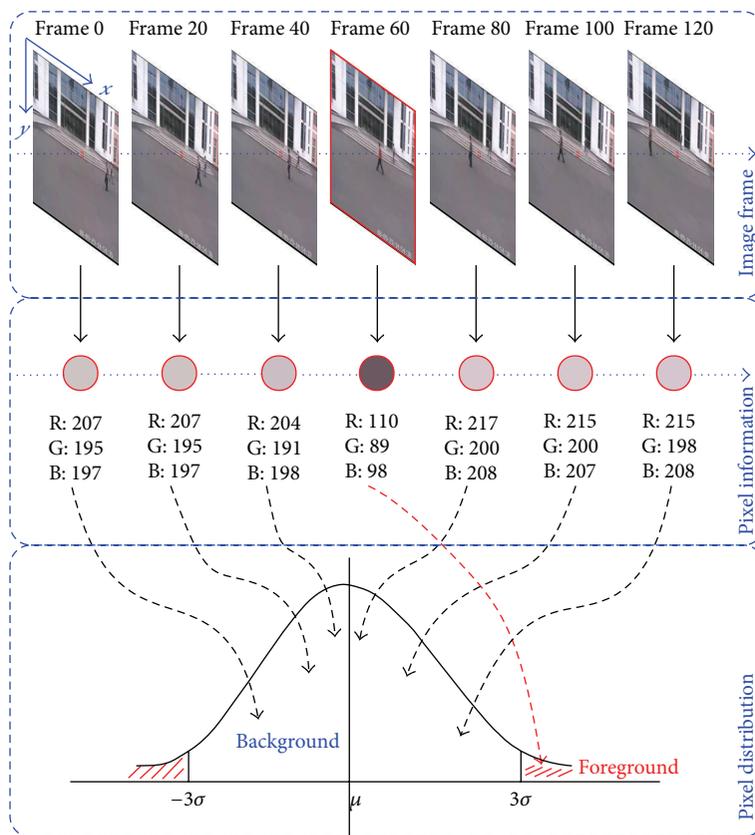


FIGURE 2: The model of pixel distribution.

strategy, with the result that the algorithm is adaptive and able to deal with multimodal backgrounds in a dynamic environment (e.g., changing time of day, clouds, swaying tree leaves, and etc.). However, since its sensitivity cannot be properly tuned, its ability to successfully handle high- and low-frequency changes in the background is debatable. To overcome these shortages, samples-based techniques [34]

circumvent a part of the parameter estimation step by building their models from observed pixel values and enhance their robustness to noises. They provide fast responses to high-frequency events in the background by directly including newly observed values in their pixel models. However, since they update their pixel models in a first-in first-out manner, their ability to successfully handle concomitant

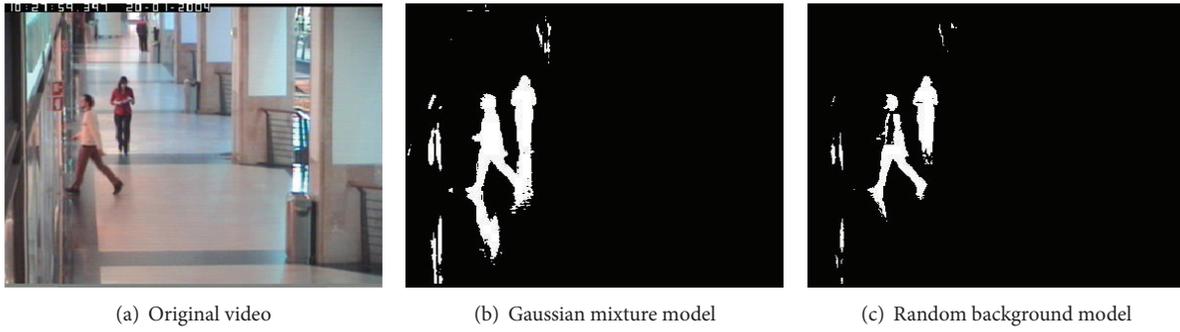


FIGURE 3: The target detection results of video data set 1.

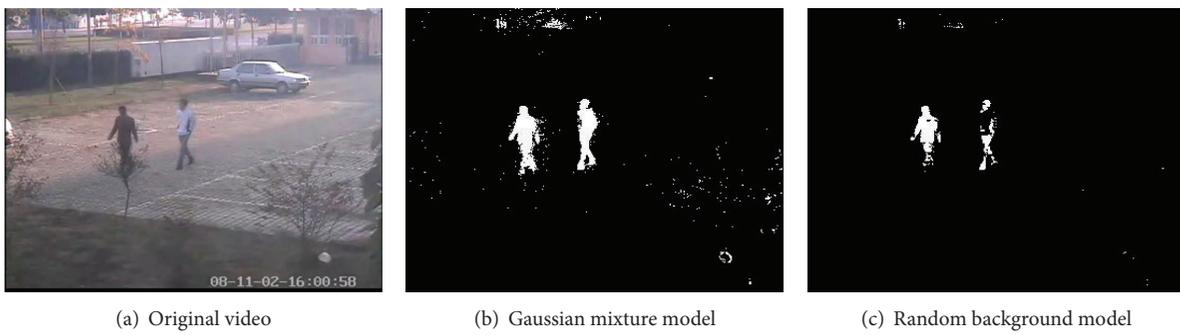


FIGURE 4: The target detection results of video data set 2.

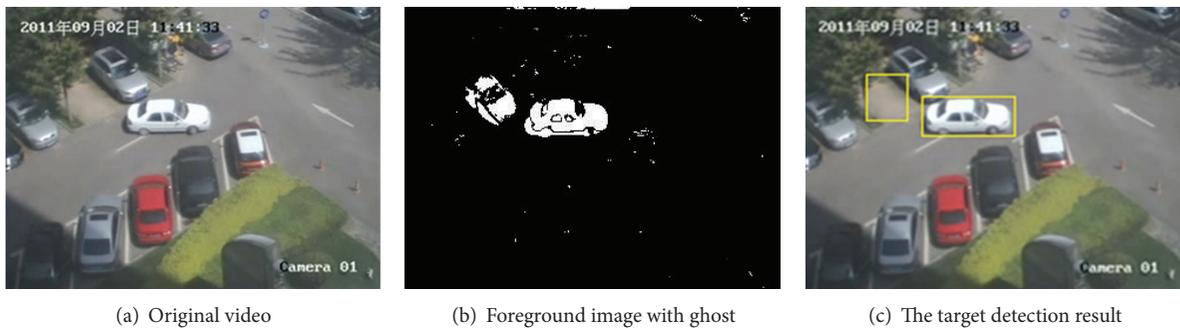


FIGURE 5: The target detection results included the ghost.

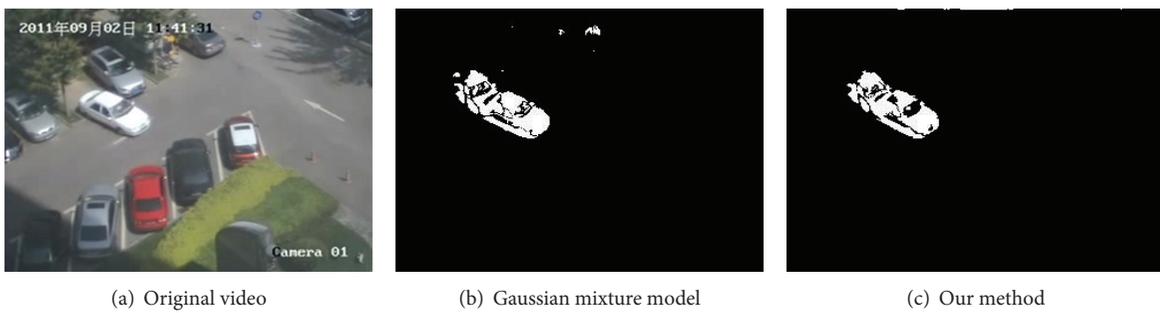


FIGURE 6: The target detection results for frame 80.

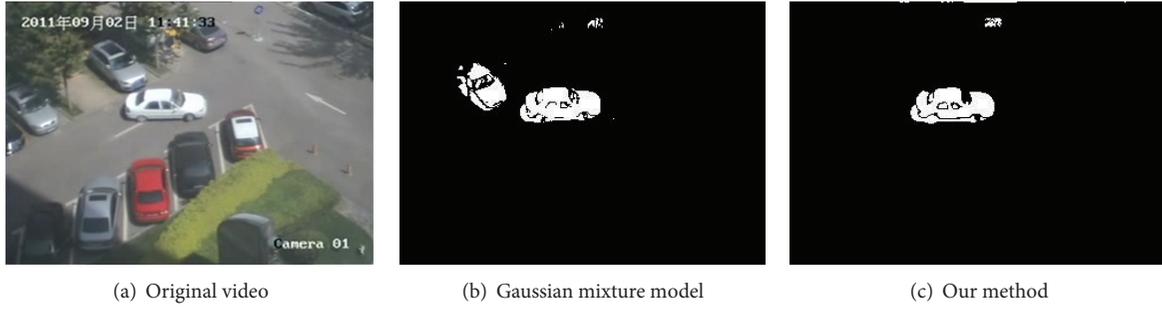


FIGURE 7: The target detection results for frame 105.

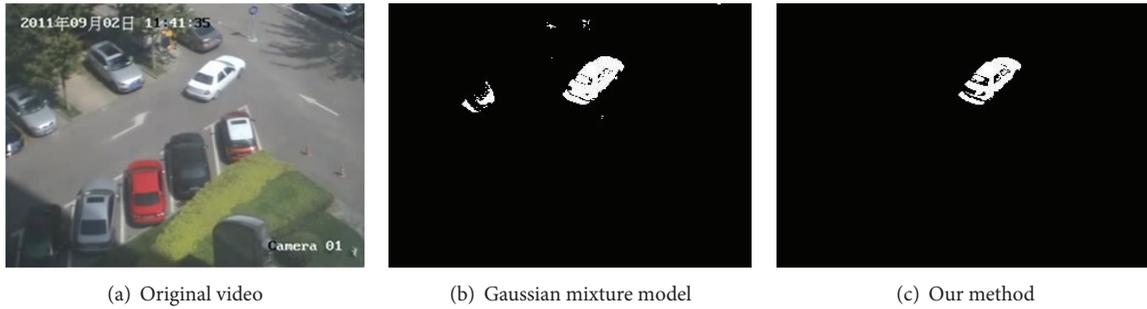


FIGURE 8: The target detection results for frame 145.

events evolving at various speeds is limited, similarly to the limitation in its adaptive ability of dealing with the concurrent events with different frequency. In order to address this issue, random background modeling that is intuitively an improved samples-based algorithm is found in [33]. This algorithm assumes $p_t(x)$ to be the value of the pixel x at time t and imposes the constraint that the influence of a value on the polychromatic space is restricted within the local neighborhood. Then, a set of sample values is used as a pixel model to classify a value $p_t(x)$ to be either a background or a foreground pixel value.

An experiment was carried out using video data and compared with the Gaussian mixture model presented by Stauffer and Grimson [21], in which the video data set 1 comes from the evaluating data from Performance Evaluation of Tracking and Surveillance (PETS) database with the video image resolution of 768×576 pixels and the frame rate of 25 frames per second (f/s); the video data set 2 comes from a practical surveillance system data in the DNC of Beihang University, with the video image resolution of 352×288 pixels and the frame rate of 25 f/s. The experimental results are shown in Figures 3 and 4. As can be seen from Figures 3 and 4, since trees were swinging in the wind, such movements were classified as foreground motions by the Gaussian mixture model, while the random background model effectively detected the trees as the background. However, both algorithms do not take into account the shadow of the target and thus are severely damaged in terms of the reliability and robustness of the target detection and tracking, as illustrated in Figure 5, in which the used video data set 3 comes from a practical surveillance system in the DNC of

Beihang University and is with the video image resolution of 352×288 pixels and the frame rate of 25 f/s.

In order to remove the damage from the shadow on target detection and tracking, we propose an algorithm by combining the random background model and the frame difference algorithm, and the mathematical model is described as follows:

$$\text{Mask}(x, y) = \begin{cases} 1 & |\text{Dilate}(D(x, y)) - \text{Erode}(D(x, y))| = 1 \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where $D(x, y)$ denotes the mask image of the background differencing; $\text{Dilate}(\ast)$ denotes the dilation operation of the target region block; $\text{Erode}(\ast)$ denotes the erosion operation of the target region; $\text{Mask}(x, y)$ denotes the mask image of the difference between the dilated and eroded image operations.

Suppose that the number of the pixels with their values equal to 1 in $\text{Mask}(x, y)$ is N_1 and the number of the pixels that are detected as foreground from differencing image and the values of which at (x, y) in the template $\text{Mask}(x, y)$ equal to 1 is N_2 . If $N_1/N_2 > T$, where T denotes a threshold, then the target region block is the foreground target; otherwise, it is the shadow of the target.

An experiment is carried out using the video data set 3 and compared with the GMM. The experimental results are shown in Figures 6, 7, and 8. As can be seen, the algorithm presented in this paper is effective to remove the shadow of the target.



FIGURE 9: The multitarget tracking results for video data set 4.



FIGURE 10: The target tracking results across multicamera for video data set 3.

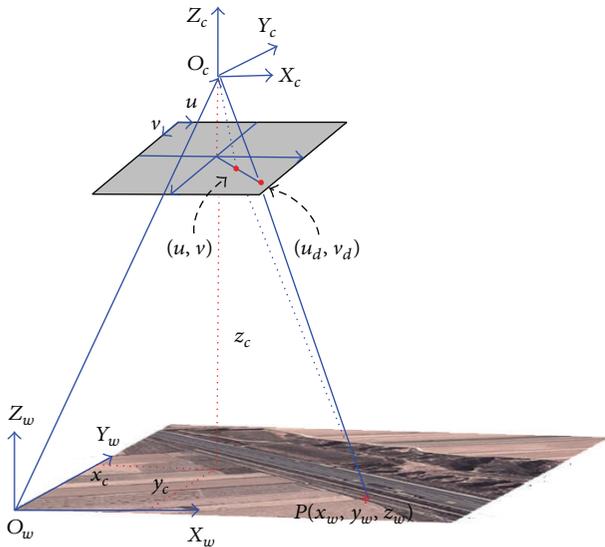


FIGURE 11: The relationship between the world coordinates and pixel coordinates.

3.2. Target Tracking. Once moving targets are detected, the track initialization event is triggered, such that the moving targets can be continuously tracked by the tracking algorithm in the living period of a track (which starts from its initialization to its termination [35]). The termination of a track occurs when a target can no longer be detected because it leaves the field of view, it stops and becomes static, or it can

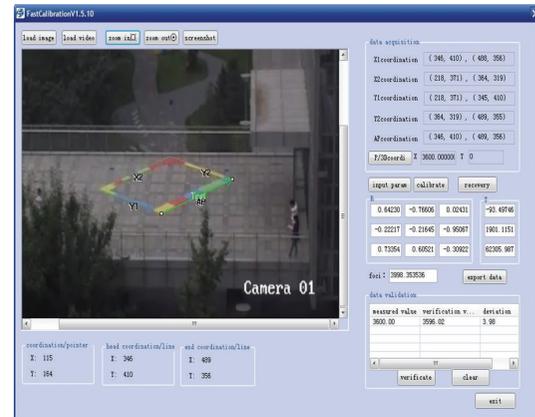


FIGURE 12: The fast calibrating camera tool software.

no longer be distinguished from the background. Detected targets are not confirmed to be true moving targets until they have been consistently tracked for a period of time before their target tracks are initialized. We create a dynamic list of potential tracks using all detected targets. Associations will be established between targets detected in a new image frame and potential tracking targets. When a potential target is tracked in several continuous frames, it is recognized as a true moving target and a track will be initialized.

Compared to a single target tracking, the multitarget problem poses additional difficulties: data association needs to be solved; that is, it has to be decided which observation

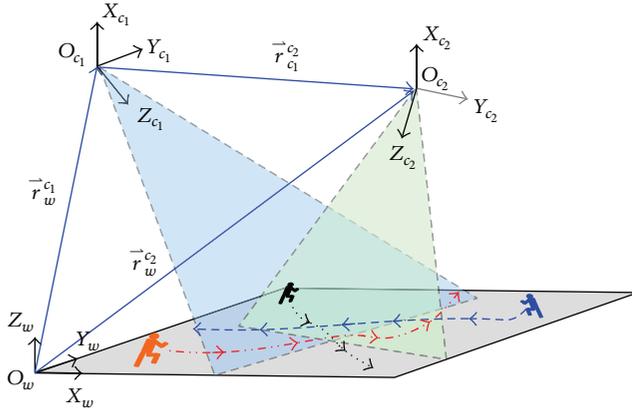


FIGURE 13: The schematic diagram of the target positioning.

corresponds to which target; constraints between targets need to be taken into account. Multitarget tracking algorithms can be roughly divided into two categories: the recursive algorithms and the nonrecursive algorithms. The recursive algorithms base their estimate only on the state of the previous frame such as Kalman filtering [36] and particle filtering [27], in which different strategy is used to obtain an optimal solution over multiple frames and can thus better cope with ambiguous, multimodal distributions. The nonrecursive algorithms seek optimality over an extended period of time [37, 38].

In practical applications, target tracking takes target splitting and merging into account because of the factors such as illumination changes and occlusion. Since the one-to-one tracking assumption rarely holds, multitarget tracking problem is still challenging. In order to realize multitarget tracking, we propose a solution by combining the pyramid Lucas-Kanade feature tracker [39] and Kalman filter. The mathematical model of Kalman filtering is described as follows:

$$\begin{aligned} X(k+1) &= A(k+1, k)X(k) + W(k), \\ Z(k) &= H(k)X(k) + V(k), \end{aligned} \quad (2)$$

where $X(k) = [x(k), y(k), w(k), h(k), v_x(k), v_y(k), v_w(k), v_h(k)]^T$ denotes the state; $A(k+1, k)$ denotes the state transition matrix; $W(k)$ denotes the system noise; $H(k)$ denotes the measurement matrix; $Z(k) = [x(k), y(k), w(k), h(k)]^T$ denotes the measurement value; $V(k)$ denotes the measurement noise; $x(k)$ and $y(k)$ denote the horizontal and vertical ordinates of the target centroid; $w(k)$ and $h(k)$ denote the width and height of the target envelope rectangle; $v_x(k)$, $v_y(k)$, $v_w(k)$, and $v_h(k)$ denote the speeds of the $x(k)$, $y(k)$, $w(k)$, and $h(k)$, respectively.

To realize multitarget tracking across multiple cameras, we construct a similarity function of target matching to

realize target association and target tracking across multiple cameras. The similarity function is described as follows:

$$\begin{aligned} f(a, b) &= \alpha M_d(a, b) + \beta M_s(a, b) + \gamma M_h(a, b), \\ M_d(a, b) &= \frac{(W_a + W_b)}{(W_a + W_b) + D_x} \cdot \frac{(H_a + H_b)}{(H_a + H_b) + D_y}, \\ M_s(a, b) &= \frac{2W_a H_a W_b H_b}{(W_a H_a)^2 + (W_b H_b)^2}, \\ M_h(a, b) &= \frac{2h_a h_b}{h_a^2 + h_b^2}, \end{aligned} \quad (3)$$

where a and b denote targets which will be matched; $M_d(a, b) \in [0, 1]$ denotes their position similarity (the larger the $M_d(a, b)$ is, the closer their positions are); $M_s(a, b) \in [0, 1]$ denotes the similarity of their sizes (the larger the $M_s(a, b)$ is, the closer their sizes are); $M_h(a, b) \in [0, 1]$ denotes the similarity of their heights (the larger the $M_h(a, b)$ is, the closer their heights are); $\alpha \in [0, 1]$, $\beta \in [0, 1]$, and $\gamma \in [0, 1]$ denote the weight coefficients and satisfy $\alpha + \beta + \gamma = 1$; X_a and Y_a denote the horizontal and vertical ordinates of the target a in the world coordinate system; X_b and Y_b denote the horizontal and vertical ordinates of the target b in the world coordinate system; W_a and W_b denote half of the widths of the targets a and b , respectively; H_a and H_b denote half of their heights, respectively; $D_x = \text{fabs}(X_a - X_b)$ denotes the absolute difference between X_a and X_b ; $D_y = \text{fabs}(Y_a - Y_b)$ denotes the absolute difference between Y_a and Y_b ; h_a and h_b denote their heights in the world coordinate system; $f(a, b) \in [0, 1]$ (the larger the $f(a, b)$ is, the higher their matching similarity is), and vice versa. In practical applications, α , β , and γ can be adjusted according to the accuracy of the $M_d(a, b)$, $M_s(a, b)$, and $M_h(a, b)$.

An experiment is carried out by using the video data sets 3 and 4 which come from a practical surveillance system in the DNC of Beihang University, with the video image resolution of 352×288 pixels and the frame rate of 25 f/s. The experimental results are shown in Figures 9 and 10, in which the rectangle with the dotted blue lines denotes the overlapping area between two cameras. As can be seen, the target tracking algorithm presented in this paper is effective and can track multiple targets across the multiple cameras.

4. Fast Camera Calibration and Target Positioning

4.1. Fast Camera Calibration. Camera calibration is a key technology in determining the mapping between 3D world coordinates and 2D image coordinates for various computer vision applications. A schematic diagram describing the mapping between 3D world coordinates and 2D image coordinates is shown in Figure 11. (u_d, v_d) is the image coordinate of (x_w, y_w, z_w) if a perfect pinhole camera model is used. (u, v) is the actual image coordinate which deviates from (u_d, v_d) due to lens distortion. The distance between them is termed as the radial distortion. Therefore, the mathematical



FIGURE 14: The experimental results in indoor environment.

model from 3D world coordinates to 2D image coordinates is expressed by [40]

$$\begin{aligned} u + \delta u &= f_x \frac{r_1(x_w - x_c) + r_2(y_w - y_c) + r_3(z_w - z_c)}{r_7(x_w - x_c) + r_8(y_w - y_c) + r_9(z_w - z_c)} \\ v + \delta v &= f_y \frac{r_4(x_w - x_c) + r_5(y_w - y_c) + r_6(z_w - z_c)}{r_7(x_w - x_c) + r_8(y_w - y_c) + r_9(z_w - z_c)}, \end{aligned} \quad (4)$$

where r_i ($i = 1, 2, \dots, 9$) denotes the element of the rotation matrix from the world coordinate frame to the pixel coordinate frame; f_x and f_y denote the focal lengths of the camera in the x and y directions; δu and δv denote the photogrammetric distortions; $O_c(x_c, y_c, z_c)$ denotes the coordinate of the camera in the world coordinate frame; $P(x_w, y_w, z_w)$ denotes the coordinate of the target in the world coordinate frame.

Traditional calibration algorithms, for example, DLT algorithm [9] and Tsai algorithm [10], utilize a series of mathematical transformations and algorithms to obtain parameters of the camera model. They have been widely used because of their simple mathematical model and theory. However, these algorithms require a large amount of work to record and check calibration points during calibration and thus are inefficient in practical applications. For this reason, a fast calibration algorithm based on the vanishing point theory was introduced in [11, 41], in which the photogrammetric distortions δu and δv are unconsidered.

According to the mathematical model of the fast camera calibration presented in [11], we develop software to calibrate cameras fast and can promptly check the accuracy of camera parameters by (4). The calibration process and results of a practical camera are shown in Figure 12. As can be seen, the focal lengths of the camera are 3998.3535 pixels, the calibration error of the line segment AP is 3.98 mm, and

the rotation matrix R and translation vector T (unit: mm) are as the following:

$$R = \begin{bmatrix} 0.642 & -0.766 & 0.024 \\ -0.222 & -0.216 & -0.951 \\ 0.734 & 0.605 & -0.309 \end{bmatrix}, \quad T = \begin{bmatrix} -93.497 \\ 1901.11 \\ 62305.937 \end{bmatrix}. \quad (5)$$

4.2. Target Positioning. Once targets are continuously tracked, the space coordinates of the targets in the camera coordinate frame can be computed by imaging model with camera parameters as follows:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = Z_c \begin{bmatrix} \frac{d_x}{f_x} & 0 & 0 \\ 0 & \frac{d_y}{f_y} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} + \begin{bmatrix} \frac{u_0 Z_c d_x}{f_x} \\ \frac{v_0 Z_c d_y}{f_y} \\ 0 \end{bmatrix}, \quad (6)$$

where u_0 and v_0 denote the principle point in the pixel frame and d_x and d_y denote the pixel sizes of the camera in the x and y directions.

When the targets are across multiple cameras, the space coordinates of the targets in each camera coordinate frame are computed, respectively, by (6) and then are unified into the world coordinated system, as illustrated in Figure 13. The mathematical model is described as follows:

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = R_i \left(\begin{bmatrix} X_{c_i} \\ Y_{c_i} \\ Z_{c_i} \end{bmatrix} - T_i \right), \quad i = 1, 2, \dots, N, \quad (7)$$

where i denotes the i th camera and N denotes the number of cameras.

5. Experiment Test

The target positioning system presented in this paper is tested in indoor and outdoor environments, in which all the low-cost static cameras are calibrated, and the coordinates of

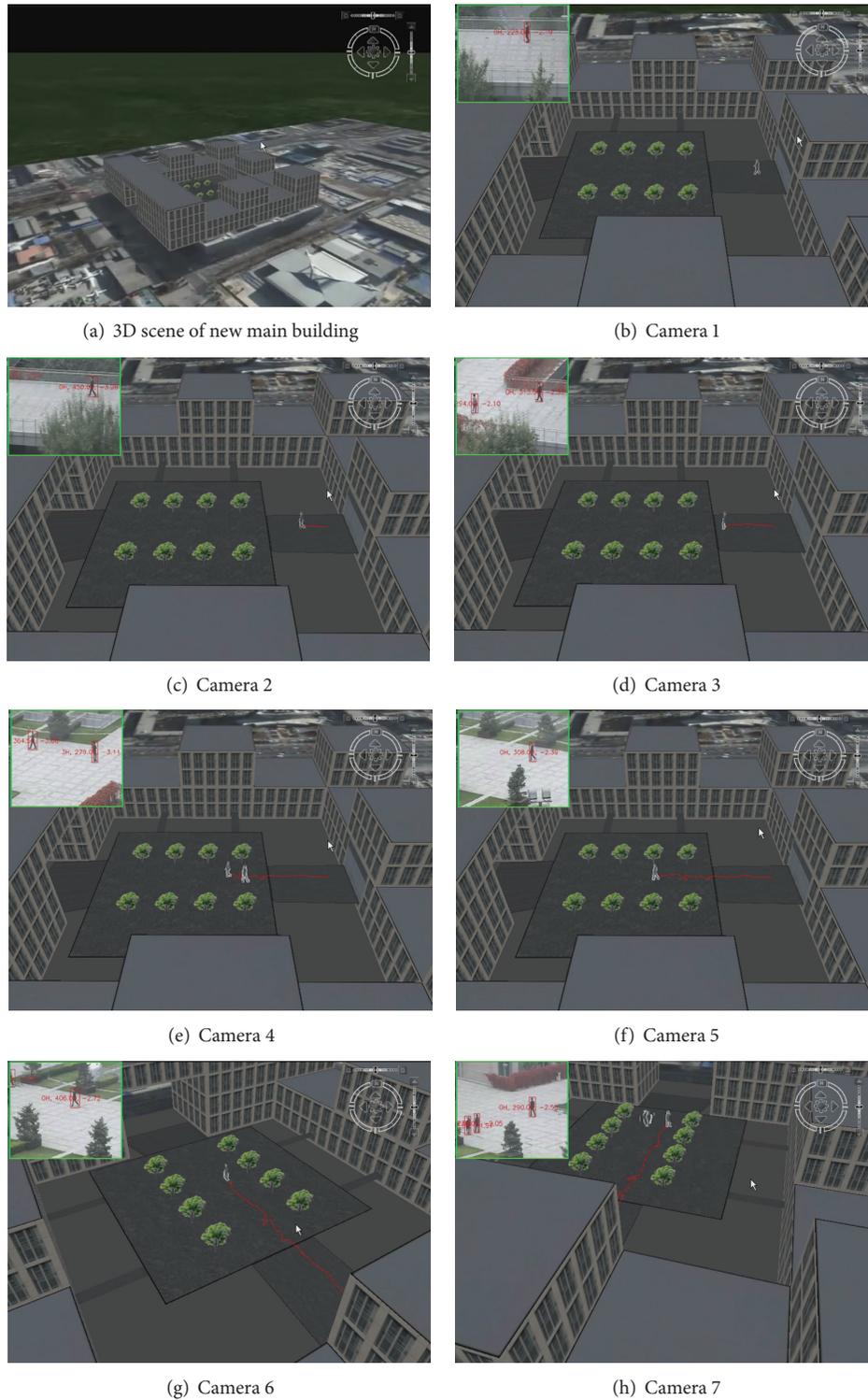


FIGURE 15: The experimental results in outdoor environment.

cameras are unified into the world coordinate system. Targets detection and tracking are done by the target detection and tracking algorithm. As a result, the targets are positioned by imaging model and camera parameters in real time, and their trajectories are displayed in a three-dimensional scene.

The test results are, respectively, shown in Figures 14 and 15 for indoor and outdoor environments. As can be seen from Figure 14, when a target is continuously moving within an indoor corridor, the positioning system consisting of six distributed cameras is able to position this target in real

time and display its trajectory in a three-dimensional space model. Likewise, as can be seen from Figure 15, when a target is continuously moving outdoors, the positioning system consisting of seven distributed cameras is able to position this target and display its trajectory in a three-dimensional space model in real time as well. The experimental results confirm that the systematic framework and inclusive algorithms presented in this paper are both effective and efficient.

In this paper, we assume that the ground is flat, which rarely holds in practical applications given a large region. In order to solve this problem, it is necessary to use digital elevation model (DEM) to describe topographic relief in large regions.

6. Conclusion and Future Work

This paper presented the comprehensive design and implementation of a moving target positioning system based on a distributed camera network. The system is composed of low-cost static cameras, which provide complementary positioning information for moving target positioning in indoor and outdoor environments when GNSS signals are unavailable. In this system, static cameras can cover a large region, moving targets are detected and then tracked using corresponding algorithms, target positions are estimated by making use of the geometrical relationships among those cameras after calibrating those cameras, and finally, for each target, its position estimates obtained from different cameras are unified into the world coordinate system. The experimental results of the target detection, tracking, and positioning system were reported based on real video data.

Targets positioning and tracking with multiple static cameras were verified in both indoor and outdoor environments. However, the reliability and accuracy of target tracking and positioning suffer from several environment factors. Hence it is necessary to fuse information from various sensors, such as radar, infrared camera, inertial measure unit (IMU), and wireless location system. Regarding future work, it is meaningful to develop and test these algorithms in the practical applications.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This project is supported by the Key Program of the National Natural Science Foundation of China (Grant no. 61039003), the National Natural Science Foundation of China (Grant no. 41274038), the Aeronautical Science Foundation of China (Grant no. 2013ZC51027), the Aerospace Innovation Foundation of China (CASC201102), and the Fundamental Research Funds for the Central Universities.

References

- [1] R. Mautz and S. Tilch, "Survey of optical indoor positioning systems," in *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN '11)*, Guimarães, Portugal, September 2011.
- [2] G. Retscher, E. Moser, D. Vredeveld et al., "Performance and accuracy test of a WiFi indoor positioning system," *Journal of Applied Geodesy*, vol. 1, no. 2, pp. 103–110, 2007.
- [3] S. J. Ingram, D. Harmer, and M. Quinlan, "Ultra wide band indoor positioning systems and their use in emergencies," in *Proceedings of the Position Location and Navigation Symposium (PLANS '04)*, pp. 706–715, Monterey, Calif, USA, April 2004.
- [4] G. Anastasi, R. Bandelloni, M. Conti et al., "Experimenting an indoor bluetooth-based positioning service," in *Proceedings of the 23rd International Conference on Distributed Computing Systems Workshops*, pp. 480–483, Providence, RI, USA, 2003.
- [5] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man and Cybernetics C: Applications and Reviews*, vol. 37, no. 6, pp. 1067–1080, 2007.
- [6] H.-S. Kim and J.-S. Choi, "Advanced indoor localization using ultrasonic sensor and digital compass," in *Proceedings of the International Conference on Control, Automation and Systems (ICCAS '08)*, pp. 223–226, Seoul, Korea, October 2008.
- [7] R. Karlsson, T. B. Schön, D. Törnqvist, G. Conte, and F. Gustafsson, "Utilizing model structure for efficient simultaneous localization and mapping for a UAV application," in *Proceedings of the IEEE Aerospace Conference Digital Object Identifier (AC '08)*, Rome, Italy, March 2008.
- [8] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, pp. 652–659, Washington, DC, USA, July 2004.
- [9] Y. I. Abdel-Aziz and H. M. Karara, "Direct linear transformation into object space coordinates in close-range photogrammetry," in *Proceedings of the Symposium on Close-Range Photogrammetry*, pp. 1–18, Urbana, Ill, USA, 1971.
- [10] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D Machine vision metrology using off-the-shelf TV camera and lenses," *IEEE Journal of Robotics and Automation*, vol. RA-3, no. 4, pp. 323–344, 1987.
- [11] J. Long, X. Zhang, and L. Zhao, "A fast calibration algorithm based on vanishing point for scene camera," *Applied Mechanics and Materials*, vol. 58–60, pp. 1148–1153, 2011.
- [12] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [13] P. D. Z. Varcheie and G.-A. Bilodeau, "People tracking using a network-based PTZ camera," *Machine Vision and Applications*, vol. 22, no. 4, pp. 671–690, 2011.
- [14] C.-H. Chen, Y. Yao, D. Page, B. Abidi, A. Koschan, and M. Abidi, "Heterogeneous fusion of omnidirectional and PTZ cameras for multiple object tracking," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 8, pp. 1052–1063, 2008.
- [15] X. Clady, F. Collange, F. Jurie, and P. Martinet, "Object tracking with a Pan-tilt-zoom camera: application to car driving assistance," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '01)*, pp. 1653–1658, Seoul, Korea, May 2001.
- [16] N. Bellotto, E. Sommerlade, B. Benfold et al., "A distributed camera system for multi-resolution surveillance," in *Proceedings*

- of the 3rd ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC '09), Como, Italy, September 2009.
- [17] C. Ding, B. Song, A. Morye et al., "Collaborative sensing in a distributed ptz camera network," *IEEE Transactions on Image Processing*, vol. 21, no. 7, pp. 3282–3295, 2012.
- [18] E. B. Ermis, P. Clarot, P.-M. Jodoin, and V. Saligrama, "Activity based matching in distributed camera networks," *IEEE Transactions on Image Processing*, vol. 19, no. 10, pp. 2595–2613, 2010.
- [19] D.-M. Tsai and S.-C. Lai, "Independent component analysis-based background subtraction for indoor surveillance," *IEEE Transactions on Image Processing*, vol. 18, no. 1, pp. 158–167, 2009.
- [20] J. Migdal and W. E. L. Grimson, "Background subtraction using Markov thresholds," in *Proceedings of the IEEE Workshop on Motion and Video Computing (MOTION '05)*, pp. 58–65, Breckenridge, Colo, USA, January 2005.
- [21] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, 2000.
- [22] D.-S. Lee, "Effective Gaussian mixture learning for video background subtraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 827–832, 2005.
- [23] L. Zhao and X. He, "Adaptive Gaussian mixture learning for moving object detection," in *Proceedings of the 3rd IEEE International Conference on Broadband Network and Multimedia Technology (IC-BNMT '10)*, pp. 1176–1180, Beijing, China, October 2010.
- [24] S. Sivaraman and M. M. Trivedi, "Looking at vehicles on the road: a survey of vision-based vehicle detection, tracking, and behavior analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1773–1795, 2013.
- [25] G. H. Wen, Z. S. Duan, G. R. Chen, and W. W. Yu, "Consensus tracking of multi-agent systems with lipschitz-type node dynamics and switching topologies," *IEEE Transactions on Circuits and Systems*, vol. 60, no. 9, pp. 1–13, 2013.
- [26] S. Oh, S. Russell, and S. Sastry, "Markov chain Monte Carlo data association for multi-target tracking," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 481–497, 2009.
- [27] H. Zhang and L. Zhao, "Integral channel features for particle filter based object tracking," in *Proceedings of the 5th International Conference on Intelligent Human-Machine Systems and Cybernetics*, pp. 190–193, Hangzhou, China, 2013.
- [28] M. Taj and A. Cavallaro, "Distributed and decentralized multi-camera tracking," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 46–58, 2011.
- [29] N. Krahnstoeber, T. Yu, S. Lim et al., "Collaborative real-time control of active cameras in large scale surveillance systems," in *Proceedings of the Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications*, pp. 1–12, Marseille, France, 2008.
- [30] C. Brailon, C. Pradalier, J. L. Crowley, and C. Laugier, "Real-time moving obstacle detection using optical flow models," in *Proceedings of the IEEE Intelligent Vehicles Symposium (IV '06)*, pp. 466–471, Tokyo, Japan, June 2006.
- [31] C. Kim and J. N. Hwang, "Fast and automatic video object segmentation and tracking for content-based applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 2, pp. 122–129, 2002.
- [32] R. Jain and H. H. Nagel, "On the analysis of accumulative difference pictures from image sequences of real world scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, no. 2, pp. 206–214, 1978.
- [33] O. Barnich and M. Van Droogenbroeck, "ViBE: a powerful random technique to estimate the background in video sequences," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '09)*, pp. 945–948, Taipei, Taiwan, April 2009.
- [34] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," in *Proceedings of the 6th European Conference on Computer Vision-Part II*, pp. 751–767, London, UK, 2000.
- [35] H. W. Mao, C. H. Yang, G. P. Abousleman, and J. Si, "Automated multiple target detection and tracking in UAV videos," in *Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications VII*, vol. 7668 of *Proceedings of SPIE*, Orlando, Fla, USA, 2010.
- [36] H. Medeiros, J. Park, and A. C. Kak, "Distributed object tracking using a cluster-based Kalman filter in wireless camera networks," *IEEE Journal on Selected Topics in Signal Processing*, vol. 2, no. 4, pp. 448–463, 2008.
- [37] A. Andriyenko and K. Schindler, "Multi-target tracking by continuous energy minimization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 1265–1272, Colorado Springs, Colo, USA, June 2011.
- [38] A. Andriyenko and K. Schindler, "Globally optimal multi target tracking on a hexagonal lattice," in *Proceedings of the 11th European Conference on Computer Vision*, pp. 466–479, Crete, Greece, 2010.
- [39] J. Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, pp. 1–9, 2001.
- [40] Y. Wang, G. Hu, and Z. Chen, "Calibration of CCD camera in image matching experimental equipment," in *Proceedings of the 2nd International Symposium on Instrumentation Science and Technology*, p. 3/146, Jinan, China, August 2002.
- [41] B. Caprile and V. Torre, "Using vanishing points for camera calibration," *International Journal of Computer Vision*, vol. 4, no. 2, pp. 127–139, 1990.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

