A method of perspective normalization for video images based on map data

Bingxian Lin, Changlu Xu, Xin Lan and Liangchen Zhou

Key Laboratory of Virtual Geographic Environment, Nanjing Normal University, Ministry of Education, Nanjing, China; State Key Laboratory Cultivation Base of Geographical Environment Evolution (Jiangsu Province), Nanjing Normal University, Nanjing, China; Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing Normal University, Nanjing, China

ABSTRACT
Objective perspective distortion is a problem that needs to be solved by video surveillance analysis. Compared with the street scene method, which depends on prior knowledge of the scene or 3D scene of the dedicated hardware recovery scene, the commonly used perspective distortion correction method is based on the linear relationship to monitor a video image in perspective normalization. However, the distortion caused by perspective imaging is nonlinear, and the linear perspective normalization model cannot guarantee the accuracy of the correction in the scene where the perspective phenomenon is evident. An image normalization method based on map data is proposed to solve this problem. A nonlinear perspective correction model is introduced by establishing a single relation between video image space and map space. With selected control points between image and map, we can calculate homography matrix in order to build the perspective correction model, which is computed to know the single pixel real size in map. The proposed perspective correction model is applied to the moving target detection. The results of the linear correction model and the proposed nonlinear correction model prove the validity and practicability of the method.

1. Introduction
In recent years, video surveillance, which is an important application technology in the field of public security, attracted the attention of scholars in computer vision and video GIS (Milosavljević, Dimitrijević, and Rančić 2010). The main contents of video surveillance include human detection and track of population density estimation and other analytical methods (Ianasi et al. 2005). These studies aim to extract and analyse various types of feature information from video images (Dalal, Triggs, and Schmid 2006).

However, the objective deformation perspective has a serious impact on the detection accuracy in all kinds of video monitoring analysis. In the image, when the same object is near the camera, the visual expression of Angle distortion is that it occupies a large pixel area, and away from the camera position, which accounted for a smaller area of the pixels. Evidently, this position has brought the interference to the characteristic information extraction and analysis based on the video image; thus, it affected the precision of the video monitoring analysis.

Through perspective normalization, the interference of perspective distortion is eliminated and the accuracy of monitoring video analysis is improved. Perspective normalization means that video images can eliminate the same object caused by distance measurement difference in perspective imaging through various types of transformation. Existing research can be divided into three categories. The image normalization method by scene linear relation, the target normalization method, and the 3D reconstruction method.

Some researchers use the linear relation to normalize the image. For example, Chan et al. calculated the linear change weight map of near and far pixels, and extracted the image features of the dense population based on the obtained weight map to improve the accuracy of population density detection (Chan, Liang, and Vasconcelos 2008). Panlong et al. corrected the optical flow field in the image according to the normalized compensation method of near and far distance scaling, and applied it to the pedestrian detection in infrared image scene with small range and low angle of view (Panlong and Yu-ming 2008). Qinglong et al. used the image vertical coordinate and the pedestrian size linear fitting method for the video image normalization after high-density population estimation. This method is simple, feasible, and has been widely used (Qinglong, Hongsha, and Ning 2014). However, the experiment shows the same object in the monitor video image in different locations of the

CONTACT Liangchen Zhou · zhoulch@njnu.edu.cn · Key Laboratory of Virtual Geographic Environment, Nanjing Normal University, Ministry of Education, Nanjing, China
© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Nanjing Normal University. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
scale changes, and the change in its distance to the camera is not linear (Figure 1). Therefore, this method reduces the non-linear problem in the linear problem processing; a large error exists in the large-scale monitoring with the evident perspective phenomenon.

Unlike the method that global normalizes the video surveillance image, several researchers attempted to normalize the local image feature vector of the target when recognizing and classifying the monitoring target. D. Hoiem performed segmentation of video surveillance images according to the ground, buildings, and the spatial relationship between the sky to obtain the level of viewpoint position as a basis for local window normalization and to improve the accuracy of pedestrian detection (Hoiem, Efros, and Hebert 2008). Similar studies include that of B. Leibe, Z. Lin, and others (Ess et al. 2010; Lin, Lin, and Weng et al. 2011). Compared with global normalization, the local normalization method can effectively improve the accuracy of image feature detection and analysis in the video; it is useful for population density estimation, pedestrian detection, and tracking applications. However, these methods must rely on a priori knowledge and the inherent clues of the scene to restore the three-dimensional structure of the scene, which has no structural information does not apply to more blocks and irregular structure (Figure 2).

Given certain conditions, some researchers use the depth camera or camera parameters in 3D reconstruction scene according to pedestrian depth information or 3D posture of the image-normalized operation. Nevatia et al. obtained the internal and external parameters of the camera directly through the PTZ camera interface, thereby recovering the 3D information of pedestrians in the image and eliminating the perspective distortion (Yuan, Bo, and Nevatia 2008). Wang et al. used a depth-based camera to obtain the depth information to achieve the perspective of the normalization of the scene. However, this method must obtain relevant and

Figure 1. Nonlinear relationship between the pixel size and the ordinate.

Figure 2. Lack of structural information on the video scene.
real-time parameters of the camera through the relevant interface of the video monitoring device; thus, it is difficult to be widely used in practical application.

Obviously, compared with the practical application, the linear transformation based on the normalization method has more restrictions, simple operation and strong practical value. The normalization method based on linear transformation can eliminate the perspective deformation of the monitored video image in a small range of the scene, but it will produce a large error in the evident scene of the fluoroscopic phenomenon. As a result, normalization is difficult to achieve.

One possible solution is to fully utilize the 2D map data that corresponds to the monitoring scene and normalize the video image accurately. Map is a 2D representation of the geographical space and the orthographic expression of geographical scenes; the moving objects on the map have the same projection area characteristics which can be represented by 2D area both in map and image.

If the map projection area is pixel-by-pixel, the corresponding actual area must be obtained. Only in this way can the non-linear relation between the precise normalization be reflected.

This paper presents a perspective normalization method based on map data. First, we obtain the homography matrix from the same point of the video image space and the 2D map space. Then, we establish the mapping relationship between the video image space and the map space and obtain the accurate non-linear perspective normalized weight graph. Finally, we use the perspective normalized weight map and the linear perspective normalized weight graph of the existing literature, which are used in the post-processing of the moving target detection, to verify the validity of the method. Then, the effectiveness of the two methods is evaluated.

The following sections of the paper are organized as follows: Section 2 describes the basic idea of this method. In Section 3, the concrete steps of this method are introduced in detail. In Section 4, we introduce the verification scheme of this method and verify the environment, data, and the results. Finally, the methods of this paper are discussed in Section 5 and the conclusion is summarized in Section 6.

2. The basic idea

The basic idea of this paper is to establish the relationship between the video image space and the 2D map space, calculate the map area that corresponds to each pixel in the image, and obtain the weight map that corresponds to the video image. In this method, the neighbouring pixels have smaller weight, whereas the distant pixels have a larger weight. The value of the distance pixel weight is used to achieve the normalized perspective and to eliminate the deformation perspective of the subsequent video feature extraction and interference analysis.

The following basic flow of this method is shown in Figure 3. Step 1: We created a set of video image pairs with the same name in the map space using a manual marker. Step 2: The homography matrix of the video image space and the map space is calculated by the point with the same name to establish the mapping relationship between the video image space and the map space. Step 3: The corresponding geographic area of each pixel in the video image is calculated based on the homography matrix, and the corresponding weight map is obtained to realize the perspective normalization of the video image.

3. Methodology

3.1. Video image space and map space definition of a single relationship

The core of this method is to construct a one-to-one mapping relationship between the video image space and the 2D map space by using the homography matrix. (Figure 4). This mapping is called homography (Criminisi 2002). This relationship can convert each point in one plane to another plane. The map area that corresponds to the midpoint of the video image can be calculated by mapping each point in a video image. Then, a weight map is created to eliminate the perspective distortion.

If the p point in the image plane transitions to the p’ point in the map plane, then we have the following definition:

\[
\begin{align*}
p &= [x, y, 1]^T \\
p' &= [x', y', 1]^T
\end{align*}
\] (1)

The homography relationship between the image plane and the map plane can be simply expressed as:

\[p = Hp'\] (2)

H is a homogeneous matrix, which can be expressed as a 2D matrix of \(3 \times 3\):

\[
\begin{bmatrix}
x \\
y \\
1
\end{bmatrix}
= \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}
\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}
\] (3)

The coordinate relationship between p and p’ is further derived by:

\[
\begin{align*}
x &= h_{11}x' + h_{12}y' + h_{13} \\
y &= h_{21}x' + h_{22}y' + h_{23}
\end{align*}
\] (4)
Finally, it is converted to the following form:

\[
\begin{bmatrix}
  x' & y' & 1 & 0 & 0 & 0 & -xx' & -xy' & -x \\
  0 & 0 & 0 & x' & y' & 1 & -yx' & -yy' & -y \\
\end{bmatrix}
\begin{bmatrix}
  h_{11} \\
  h_{12} \\
  h_{13} \\
  h_{21} \\
  h_{22} \\
  h_{23} \\
  h_{31} \\
  h_{32} \\
  h_{33}
\end{bmatrix}
= \begin{bmatrix}
  0 \\
  0
\end{bmatrix}
\]

(5)

In Formula 5, to normalize \( h_{33} \) to 1, eight unknown equations are required to solve eight unknowns. Four groups of points or more points must be obtained to solve the homography matrix (Criminisi 2002).

### 3.2. The same point based on interactive mode selection

The derivation shows that it is necessary to select four or more homonymic points to establish the mapping relationship between video space and image space by using the homography matrix. We use artificial interaction to select the same point in the video image and 2D map (Figure 5). The same point should be evenly distributed on the image surface, covering most of the monitoring area to avoid the distortion caused by the camera lens. We also need to acquire distinctive features, such as buildings, roads, etc. Therefore, the mapping is accurate. The location on a 2D map is identified.
3.3. Calculation of perspective weights based on the map area

After determining the mapping relationship between the image plane and the 2D map plane, the map area that corresponds to each pixel in the video image can be calculated to construct the perspective weight map. Weight map, which is the same size of the video image, and each pixel assigned by the corners size of area in the map. In the camera imaging model, the pixel expresses information of a rectangular area with a certain length and width, and each pixel corresponds to a quadrilateral area on the 2D map space. The image coordinates of the pixel corners are converted to...
cartesian coordinates in the map space. Then, we can solve the map space under the corresponding quadrilateral map area. In order to simplify the calculation and ignore the anisotropy of the pixels, we assume that the corresponding quadrilateral pixel region is a square. \((X + 0.5, x + 0.5), (x + 0.5, x + 0.5), (x + 0.5, x + 0.5), (x + 0.5), (x + 0.5) (X + 0.5, x + 0.5)\), as shown in Figure 6. The coordinates of the corners of the pixel converter correspond to the map coordinates based on the homography matrix, and the map area is calculated by using the map coordinates to obtain the perspective map.

3.4. Moving target detection based on a perspective weighted graph

The validity of the perspective correction effect must be applied to specific video analysis methods. The obtained nonlinear perspective weights and the linearized perspective weights in the existing literature are applied to the post-processing stage of the moving target detection, and the results are compared and verified. Moving object detection is a method that distinguishes moving objects from background information in video sequence images; it is the basis of various video analysis algorithms and video compression algorithms (Kim 2003; Kim and Hang 2003). In video surveillance applications, usually within a certain period, the background of the video does not change; thus, more backgrounds are subtracted based on the moving target detection. A variety of background subtraction algorithms have been incorporated into the open-source BGS Library. We select five background subtraction algorithms that are widely used to separate the front/back scene of video images. For the extracted foreground image, two different perspective weights were used to denoise, and the actual motion pedestrian was retained by using the digital morphological method. The process is shown in Figure 7.

To evaluate the effectiveness of the method, further calculations are performed on TP (actual class), TN (actual negative class), FP (false positive class), and FN (false negative class) to obtain Precision (precision), Recall (recall), and F1-measure (Lipton, Elkan, and Naryanaswamy 2014).

TP represents the actual statistic value, that is, the number of detected actual foreground pixels; TN represents the statistic value of the actual negative sample, that is, the pixel number of the actual background detected; FP represents the statistical value of false positive. FN is the statistical value of the false negative sample, that is, the number of pixels mistakenly recognized as the foreground of the background.
Background subtraction method is commonly used in the two metrics of recall rate (Recall) and precision (Precision); their corresponding formula is as follows:

\[
Recall = \frac{TP}{TP + FN}
\]

\[
Precision = \frac{TP}{TP + FP}
\]

When the recall rate and accuracy are high, the performance of the algorithm is better. Moreover, obtaining one-sided results is easy if the accuracy and performance of the background subtraction method are evaluated only by the recall rate and accuracy. For example, when all the pixels in the image are detected as foreground, the recall rate is 100%, which is evidently incorrect. F1-measure is used as an integrated measure of the experiment, which represents the average of recall and precision:

\[
F1 - measure = 2 \times \frac{Recall \times Precision}{Recall + Precision}
\]

If the F1-measure of the background subtraction method is closer to 1, then the performance of the algorithm is better. When the value is closer to 0, the performance of the algorithm is worse.

We calculate the F1-measure index of the moving target detection to evaluate the effectiveness of the perspective correction in three cases (non-linear perspective weighting chart is adopted without correcting the perspective, which is corrected by the linear perspective weight map).

4. Case study

4.1. Verification scheme design

The video images which are captured in square and corridor, meanwhile the map is covered by the same place. This system is based on VC++ 2013. It uses OpenCV 3.0 to load and display video images, ArcEngine 10.0 to load and display 2D maps, and BGS Library (Sobral and Vacavant 2014) Background separation. The system operating environment is Windows 8.1 operating system. The processor is Intel Core i7 clocked at 3.5 GHz with a memory of 8GB. The Verification system interface is shown in Figure 8, including the following functions:

1. The video images which are captured in square and corridor, meanwhile the map is covered by the same place.

   The mapping relation between the video image space and the map space is established based on the homonymy point pair.

   (1) The map area value that corresponds to each pixel in the video image is calculated according to the homography matrix, and the perspective weight map is obtained;

   (2) Separation of the front/back scene of the video image;

   (3) The linearized perspective weights obtained by the method proposed by Chan (Chan, Liang, and Vasconcelos 2008) and the non-linear perspective weight map obtained in this paper are used to post-process the separated foreground and

---

![Figure 8. Prototype verification system.](image-url)
To verify the data, two sections of the 640 * 480 resolution outdoor monitoring video (Figure 9) are selected. The first video is a shooting scene for the high downstairs channel, and the second video is a shooting scene for a bus station next to the road; the moving targets detected are pedestrians. Usually, the pedestrian height is 1 to 2 metres, and the shoulder width is 0.3 to 1 metre; thus, the maximum pedestrian footprint area is \(T_{\text{max}}\), which is set as two square metres, and the minimum map area is \(T_{\text{min}} \) 0.3 square metres.

### 4.2. Verify the results

For the two scenes of Videos 1 and 2, we first calculate the perspective weighting map based on the previous method (Figure 10), and then the background subtraction algorithm is used to separate the foreground and background. The perspective weighted map is used for the foreground image processing operation. The experimental results are shown in Tables 1 and 2.

In the test video 1, the shooting scene is in the building between the channels. The building glass reflection light changes the ground of the light and shadow constantly; thus, the movement of the target detection caused interference. The experimental results indicate that the traditional method is based on the linear hypothesis for normalization and cannot effectively suppress the interference information that is caused by the change of light and shadow. Therefore, various kinds of moving target detection algorithms that are closer to the camera will retain more pseudo-moving targets. After adopting this method, the detection results of various motion target detection algorithms have been improved evidently by the nonlinear perspective normalization based on the map data.

In the test video 2, the shooting scene is located along the roadside. Therefore, various kinds of moving target detection algorithms that are closer to the camera will retain more pseudo-moving targets. The experimental results show that the traditional method, which is based on the linear hypothesis to normalize the processing, cannot effectively interfere with the tree shake caused by interference information. The detection results retained more targets that are pseudo-moving. After adopting this method, the detection results of various motion target detection algorithms have been improved evidently by the nonlinear perspective normalization based on the map data.

### 4.3. Accuracy analysis

We obtain the TP (actual class), TN (actual negative class), FP (false positive class), and FN (false negative class) of the foreground and background, which are generated and normalized by the moving object extraction algorithm to further describe the accuracy of the detection and obtain Precision, Recall, and F1-measure (Lipton, Elkan, and Naryanaswamy 2014). The statistical results are presented in Tables 3 and 4.

The accuracy of the statistical detection shows that when the same post-processing method is used to

---

**Figure 9.** Verification data.
process the foreground/background images that are separated by the BGS algorithm library, compared with the linear perspective weight map, the proposed non-linear perspective weighting method has an evident improvement in precision.

Compared with the F1-measure, the accuracy of the inter-frame difference method is improved, and the accuracy of the improved multi-layer background modelling method is less. The result is due to the low complexity of the inter-frame difference algorithm. The extraction of foreground pixels contains more noise; thus, the effect of this algorithm is more evident. The multi-layer background modelling algorithm uses a complex algorithm with high accuracy. The extracted foreground contains less noise; thus, the precision is improved.

The accuracy of the two video experiments are different. Improvement comparison of the two algorithms before and after using F1-measure is shown in Figures 11 & 12. In the second video scene, the precision of the algorithm is improved. Analysis of the surveillance video that corresponds to the scene structure shows that the depth of the scene greatly affected the detection accuracy. The perspective distortion of the video that corresponds to a scene with less depth is not evident, whereas that of the video two is otherwise. Therefore, the proposed algorithm is more effective for videos. In the case of large-scale outdoor real-time video surveillance, the distortion caused by perspective interference can be effectively eliminated by using this method to obtain a non-linear perspective weight matrix.

4.4. Multiplane processing strategy analysis

The key step of this algorithm is to create the relationship between the image and the map obtain the normalized weighting map; thus, it can realize the high precision normalization when the image scene is a single plane. However, when many planes are in the scene, the mapping relationship of the single plane cannot guarantee the accuracy of the normalized weight map calculation. It
to establish a mapping relationship between different images and map regions to obtain normalized weights of different plane regions. (Figure 13).

### 4.5. Application scalability analysis

This method can not only realize the improvement on moving target detection algorithm, but also can be extended to the population statistical process through simple modification. The method is as follows: We create a mapping relationship between the image and the mapping plane, obtain the normalized weight map, select the image area for statistics, and normalize the image area according to the normalized weight map. Taking the number of texture population as an example, the normalized population is obtained by performing an algorithmic analysis on the normalized corrected image.

### 5. Conclusions

Eliminating the interference of perspective distortion on video feature extraction and analysis is an important problem that needs to be addressed in many video analysis methods. This paper presents a perspective normalization
method based on map data. First, we select the pair of same name points in the video image space and the two-dimensional map space, establish the first-order relationship between the two, and then calculate the corresponding relation of each pixel in the video image. The map area that is a non-linear perspective weighting graphics obtained, and the perspective of the video image is finally normalized. The obtained non-linear perspective
## Table 2. Second video test results comparison.

| Algorithm name                  | Video image | Linear normalization | Non-linear normalization |
|---------------------------------|-------------|----------------------|-------------------------|
| Interframe difference method    | ![Frame](image1) | ![Frame](image2) | ![Frame](image3) |
| Average Gaussian modeling       | ![Frame](image4) | ![Frame](image5) | ![Frame](image6) |
| Mixed gaussian modeling         | ![Frame](image7) | ![Frame](image8) | ![Frame](image9) |
| Adaptive pixel segmentation modeling | ![Frame](image10) | ![Frame](image11) | ![Frame](image12) |
| Multi-layer modeling            | ![Frame](image13) | ![Frame](image14) | ![Frame](image15) |

## Table 3. Video-1 precision information statistics.

| Algorithm                              | Precision | Recall | F1-measure | Linear normalization | Non-linear normalization |
|----------------------------------------|-----------|--------|------------|----------------------|-------------------------|
| Interframe difference method           | 0.853     | 0.773  | 0.811      | 0.911                | 0.823                   | 0.865                   |
| Adaptive hybrid Gaussian modelling     | 0.912     | 0.785  | 0.844      | 0.925                | 0.801                   | 0.859                   |
| Mixed gaussian modelling               | 0.902     | 0.783  | 0.838      | 0.914                | 0.807                   | 0.857                   |
| Adaptive pixel segmentation modelling  | 0.913     | 0.792  | 0.848      | 0.922                | 0.802                   | 0.858                   |
| Multi-layer modelling                  | 0.916     | 0.842  | 0.877      | 0.926                | 0.855                   | 0.889                   |

## Table 4. Video-2 precision information statistics.

| Algorithm                              | Precision | Recall | F1-measure | Linear normalization | Non-linear normalization |
|----------------------------------------|-----------|--------|------------|----------------------|-------------------------|
| Interframe difference method           | 0.821     | 0.724  | 0.769      | 0.901                | 0.814                   | 0.855                   |
| Adaptive hybrid Gaussian modelling     | 0.867     | 0.745  | 0.801      | 0.905                | 0.807                   | 0.853                   |
| Mixed gaussian modelling               | 0.854     | 0.743  | 0.794      | 0.911                | 0.814                   | 0.859                   |
| Adaptive pixel segmentation modelling  | 0.859     | 0.739  | 0.794      | 0.927                | 0.806                   | 0.862                   |
| Multi-layer modelling                  | 0.87      | 0.772  | 0.818      | 0.92                 | 0.847                   | 0.881                   |
weights are compared with the linear perspective weights of the existing literature in post-processing of moving object detection. The results show that the method can eliminate the influence of perspective distortion on the accuracy of video analysis more effectively. Especially if the scene depth is large and the perspective deformation of the scene is evident, then the effectiveness of this method is particularly evident. Therefore, the perspective normalization method proposed in this paper can effectively eliminate the image caused by perspective deformation, and can be better applied to video surveillance and analysis under large-scale scenes.

**Author Contributions**

B.L. proposed the original idea, set up the case study, performed most of the method, and wrote the draft of the manuscript. C.X., X.L. and L.Z. principally conceived the idea. L. Z. provided financial support.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This research was funded by the National Natural Science Foundation of China, grant number 41571381 and the National Key Research and Development Program of China, grant number 2017YFB0503503.

**References**

Chan, A. B., Z. S. J. Liang, and N. Vasconcelos. 2008. “Privacy Preserving Crowd Monitoring: Counting People without People Models or Tracking.” In Conference on computer vision and pattern recognition, IEEE CVP, 1–7. Anchorage, AK.

Criminisi, A. 2002. “Single-View Metrology: Algorithms and Applications (Invited Paper).” Dagm Symposium on Pattern Recognition, 224–239, Zurich, Switzerland.

Dalal, N., B. Triggs, and C. Schmid. 2006. Human Detection Using Oriented Histograms of Flow and Appearance, 428–441. Graz, Austria: ECCV.

Ess, A., K. Schindler, B. Leibe, and L. Van Gool. 2010. “Object Detection and Tracking for Autonomous Navigation in Dynamic Environments.” The International Journal of Robotics Research 29 (14): 1707–1725. doi:10.1177/0278364910365417.

Hoiem, D., A. A. Efros, and M. Hebert. 2008. “Putting Objects in Perspective.” International Journal of Computer Vision 80 (1): 3–15. doi:10.1007/s11263-008-0137-5.

Janjić, V., V. Gui, C. I. Toma, and D. Pescaru. 2005. “A Fast Algorithm for Background Tracking in Video Surveillance Using Nonparametric Kernel Density Estimation.” Facta Universitatis (Niš) 18 (1): 127–144.

Kim, J. B., and J. K. Hang. 2003. “Efficient Region-based Motion Segmentation for a Video Monitoring System.” Pattern Recognition Letters 24 (1): 113–128. doi:10.1016/S0167-8655(02)00194-0.

Kim, K. J. 2003. “Apparatus and Method for Motion Detection of Image in Digital Video Recording System Using MPEG Video Compression.” US 20030123551 A1.

Lin, T. Y., Y. Y. Lin, M. F. Weng. 2011. “Cross Camera People Counting with Perspective Estimation and Occlusion Handling.” In IEEE International Workshop on Information Forensics & Security, 1–6. IguacuFalls, Brazil: IEEE Computer Society.

Lipton, Z. C., E. Elkan, and B. Naryanaswamy. 2014. “Optimal Thresholding of Classifiers to Maximize F1 Measure[C].” In Joint European Conference on Machine learning and Knowledge Discovery in Databases, 225–239. Nancy, France.

Milosavljević, A., A. Dimitrijević, and D. Rančić. 2010. “GIS-augmented Video Surveillance.” International Journal of Geographical Information Science 24 (9): 1415–1433. doi:10.1080/13658811003792213.

Panlong, S., and Z. Yu-ming. 2008. “Optical Flow Analysis for Infrared Pedestrian Monitoring in Low View.” Microcomputer Applications 24 (7): 46–49.

Qinglong, C., X. Hongsha, and L. Ning. 2014. “A Method for People Counting in Complex Scenes Based on Normalized Foreground and Corner Information.” Journal of Electronics & Information Technology 36 (2): 312–317.

Sobral, A., and A. Vacavant. 2014. “A Comprehensive Review of Background Subtraction Algorithms Evaluated with Synthetic and Real Videos.” Computer Vision & Image Understanding 122 May: 4–21. doi:10.1016/j.cviu.2013.12.005.

Yuan, L., W. Bo, and R. Nevatia. 2008. “Human Detection by Searching in 3D Space Using Camera and Scene Knowledge.” In International Conference on Pattern Recognition, 1–5. Tampa,FL, USA.