Research Article

Optical Flow and Principal Component Analysis-Based Motion Detection in Outdoor Videos

Kui Liu, Qian Du, He Yang, and Ben Ma

Department of Electrical and Computer Engineering, Mississippi State University, MS 39762, USA

Correspondence should be addressed to Qian Du, du@ece.msstate.edu

We propose a joint optical flow and principal component analysis (PCA) method for motion detection. PCA is used to analyze optical flows so that major optical flows corresponding to moving objects in a local window can be better extracted. This joint approach can efficiently detect moving objects and more successfully suppress small turbulence. It is particularly useful for motion detection from outdoor videos with low quality. It can also effectively delineate moving objects in both static and dynamic background. Experimental results demonstrate that this approach outperforms other existing methods by extracting the moving objects more completely with lower false alarms.

1. Introduction

The detection of moving objects is critical in many defense and security applications, where motion detection is usually performed in a preprocessing step, a key to the success in the following target tracking and recognition. Many videos used in defense and security applications are outdoor videos whose quality may be degraded by various noisy sources, such as atmospheric turbulence, and sensor platform scintillation. Meanwhile, moving objects may be very small occupying a few pixels only, which makes motion detection very challenging. Under this circumstance, existing approaches may generate significant amount of false alarms.

Motion detection has been extensively investigated [1–3]. Many research works are conducted for indoor videos with large objects. As one of the major techniques, optical flow-based approaches have been widely used for motion detection. There are two classic methods of optical flow computation in computer vision: Horn-Schunck (HS) method and Lucas-Kanade (LK) method [4–7]. Both of them are based on the two-frame differential algorithms. LK method may not perform well in dense flow field; on the other hand, HS method can detect minor motion of objects and provide a 100% flow field [7]. Thus, we focus on HS method for optical flow computation in our research. Considering

outdoor videos with low quality, special care needs to be taken in order to better extract features related to moving objects from optical flows while suppressing false alarms.

Principal component analysis (PCA) is a typical approach in multivariate analysis [8]. It is also named the discrete Karhunen-Loève transform (KLT) or the Hotelling transform [9]. PCA includes the eigen-decomposition of a data covariance matrix or singular value decomposition of a data matrix, usually after mean centering. It projects the original data onto an orthogonal subspace, where each direction is mutually decorrelated and major data information is present in the first several principal components (PCs). For optical flows in a local window, moving objects have consistent flows while pixels with only turbulence have random flows. Thus, if PCA is applied to the two-dimensional (2D) data of optical flows, the difference between desired motion pixels and random motion pixels may be magnified because their contributions to the two eigenvalues are very different; the contribution from random motion pixels can be very small, even to the second eigenvalue. Experimental results show that this approach actually is an effective way of analyzing outdoor videos; it can reduce false alarms for videos with either static or dynamic background, and it is also useful to delineate the size of moving objects.

This paper is organized as follows. Section 2 explains the proposed method based on optical flow and PCA. Section 3 presents experiments using ground-based and airborne videos. Section 4 draws the conclusion.

2. Proposed Method

HS method is a special approach of using global constraint of smoothness to express a brightness variation in certain areas of the frames in a video sequence. It is also a specially defined framework to lay out the smoothness of the flow field. Let $I(x, y, t)$ represent the brightness of a pixel at (x, y) coordinates and the tth frame. According to [4], the image constraint at $I(x, y, t)$ with Taylor series can be expressed as

$$\frac{\partial I}{\partial x}\partial x + \frac{\partial I}{\partial y}\partial y + \frac{\partial I}{\partial t}\partial t = 0, \tag{1}$$

which results in

$$I_x u + I_y v + I_t = 0, \tag{2}$$

where $u = \partial x/\partial t$ and $v = \partial y/\partial t$ are the x and y components of the velocity or optical flow of $I(x, y, t)$, respectively, and $I_x = \partial I/\partial x$, $I_y = \partial I/\partial y$, and $I_t = \partial I/\partial t$, are the derivatives of the image at (x, y, t) in the corresponding directions. A constrained minimization problem can be formulated to calculate optical flow vector (u^{k+1}, v^{k+1}) for the $(k + 1)$th frame:

$$u^{k+1} - \overline{u}^k = I_x \cdot \frac{I_x \overline{u}^k + I_y \overline{v}^k + I_t}{\alpha^2 + I_x^2 + I_y^2},$$

$$v^{k+1} - \overline{v}^k = I_y \cdot \frac{I_x \overline{u}^k + I_y \overline{v}^k + I_t}{\alpha^2 + I_x^2 + I_y^2}, \tag{3}$$

where $\overline{u}^k$ and $\overline{v}^k$ are the estimated local average optical flow velocities, and α is a weighting factor. A larger value of α results in a smoother flow; in our experiments using 8-bit videos, it is empirically set to be 30000. Based on the norm of an optical flow vector, one can determine if the motion exists or not, while the direction of this vector provides the motion orientation.

Two optical flow images can be constructed by pixel optical flow vector (u, v). A mask of size $n \times n$ slides through these u and v images. At location (i, j), a two-dimensional (2D) data matrix $\mathbf{X}$ can be constructed, which includes all the 2D vectors covered by the mask. The covariance matrix can be calculated as

$$\mathbf{\Sigma} = \overline{\mathbf{X}}^T \overline{\mathbf{X}}, \tag{4}$$

where $\overline{\mathbf{X}}$ is the optical flow matrix after mean removal. After eigen-decomposition, two eigenvalues (λ_1, λ_2) are assigned to the central pixel of the mask. Motion detection is accomplished by analyzing or thresholding the eigenvalue(s). Since λ_1 is the major flow component and λ_2 is the minor flow component, it may be more effective to considering (λ_1, λ_2) than the values in the original (u, v) space.

Intuitively, only λ_1 needs to be considered because it corresponds to the major flow component and λ_2 corresponds to the minor flow component or even turbulence. An appropriate threshold can be determined by using the Ostu's method on the λ_1 histogram [10]. However, in practice, λ_2 should be considered as well since pixels inside object boundaries usually have quite large λ_2 but not λ_1. Thus, thresholding may need to be taken on the λ_2 histogram; a pixel is claimed to have motion if either λ_1 or λ_2 are above the corresponding thresholds.

Thus, the motion detection algorithm can be described as follows.

(1) Calculate optical flows between two adjacent frames (after registration as needed).

(2) For each pixel in the 2D optical flow data, perform PCA for a local mask (of size 3×3 in the experiment), and two eigenvalues are assigned to the central pixel.

(3) Apply the Ostu's thresholding to the eigenvalues of all the pixels (λ_2 in the experiment).

Figure 1 illustrates the framework of the proposed method with a 3×3 mask and resulting 2×9 data matrices.

It is noteworthy that some variants exist when implementing the proposed method differently.

(1) In Step (1), we may use the optical flow data from multiple frames. For instance, optical flow data from Frames 1 and 2 can be combined with optical flow data from Frames 2 and 3; this may help to emphasize the desired optical flows of moving objects and to emphasize the randomness of turbulence.

(2) In Step (2), masks with different sizes can be used. Intuitively, for a large moving object, mask size should be large.

(3) In Step (3), thresholding can take place on either λ_1 or λ_2, depending upon the object size and the features of turbulence.

In the experiments, we use two adjacent frames, a 3×3 mask, and only λ_2 for thresholding. It is to show that such simplest implementation is sufficient to provide better performance than other widely used techniques.

3. Experiments

In the experiments, videos with both static and dynamic backgrounds were analyzed. They were taken by a commercial Sony Camcorder. We compared our proposed method with the original optical flow method, the motion detection methods based on Kalman filtering [11], background modeling using Gaussian mixture model [12], difference-based spatial temporal entropy image (DSTEI) [13], and forward-backward motion history images (MHI) [14]. They were chosen for comparison because they are either typical methods or designed specifically for more complicated videos (e.g., those with dynamic background).
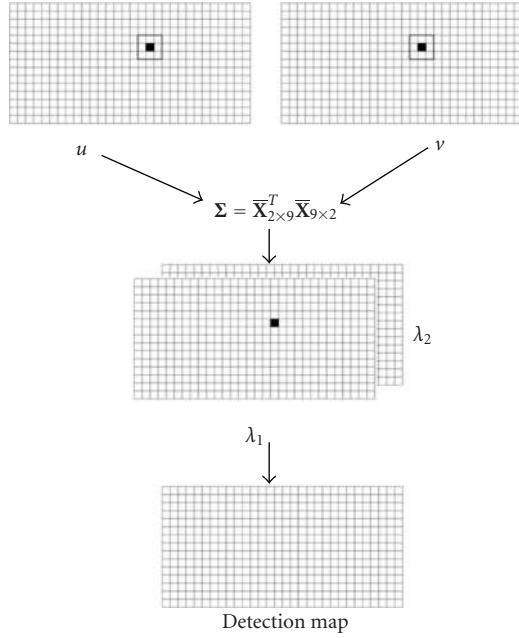
$$\Sigma = \overline{\mathbf{X}}_{2\times 9}^{T}\overline{\mathbf{X}}_{9\times 2}$$

FIGURE 1: The framework of the proposed method.

(a) frame 1

(b) frame 2

FIGURE 2: The two input frames of a helicopter video.

3.1. Experiment 1: Ground-Based Video with Relatively Large Object. In this experiment, a video with static background in a small regional airport was studied, which was taken when the camcorder was mounted on a tripod. As shown in Figure 2, a Hughes Cayuse helicopter was the moving object. Since the video was taken during a humid summer afternoon, there were significant atmospheric turbulence effects, which were visible around the vehicle, runway, and tree profiles.

FIGURE 3: The result from optical flow method.

FIGURE 4: The result from Kalman filtering.

FIGURE 5: The result from background modeling.

Figure 3 shows the detection result using optical flow only, where detected pixels were highlighted in red. It contained many false alarm pixels in runway and tree profiles. Figures 4, 5, 6, and 7 are the detection results using Kalman filtering, background modeling, DSTEI, and MHI methods, respectively. We can see that they all could detect the helicopter but with some regions missing and a few false alarm background pixels. The background modeling method could detect the largest areas of the helicopter; however, there were erroneously detected pixels scattered in the scene (even in the sky area). This method relies on an accurate background model, generally requiring complicated computations.

Figure 8 is the result of the proposed method, where almost all the false alarm pixels were removed (only two pixels in the vehicles were left) and major regions in the helicopter were detected. Compared to Figure 3, introducing PCA can significantly improve the performance of optical flow-based detection. Compared to the results in Figures 4–7, the proposed method can reduce false alarm while detecting larger regions in the moving object.

Figure 6: The result from DSTEI method.

(a) frame 1

Figure 7: The result from MHI method.

(b) frame 2

Figure 9: The two input frame about an airborne video.

Figure 8: The result from the joint optical flow and PCA method.

Figure 10: The result from optical flow method.

3.2. Experiment 2: Airborne Videos with Small Objects. The second experiment used an airborne video with low quality. It was taken by the camcorder mounted on the helicopter in the video shown in Experiment 1. In addition to atmospheric turbulence, scintillation from the airborne platform (i.e., the small helicopter) further degraded the video quality. As shown in Figure 9, there were three moving vehicles on the highway, highlighted in yellow circles. They consisted of only a few pixels. The two frames were pre-registered using the method in [15].

Figure 10 shows the detection result using optical flow only, where three vehicles on the highway were completely detected and the shape of the vehicles were outlined compactly. Figures 11, 12, 13, and 14 are the results for comparison, where the three vehicles were detected but not well delineated. For instance, the detected vehicle sizes were too small when using Kalman filtering and background modeling, and too big when using DSTEI and MHI. More false alarm pixels were contained in these results. Figure 15 is the result using optical flow and PCA, which could further reduce false alarm and the vehicle sizes seemed to be more reasonable. Although the proposed

method provided the best result, there were still several false alarmed pixels, mainly located around the edges of buildings.

We found out that such false alarms in airborne videos with small moving objects can be better removed by corner-based detection [16]. Harris corners were detected from two difference images, and many false alarm pixels around buildings could be removed; false alarms were further reduced through local tracking of detected corners in several consecutive frames. The drawback is that the detected result contains only object corners. In conjunction with the proposed method, the complete regions of moving objects can be segmented for the corner-based detection while the false alarm can be reduced in the proposed method. As shown in Figure 16(a), the corner-based method can accurately detect the three vehicles without false alarms; however, it detects only a corner corresponding to an object as detailed in Figure 16(b). Figure 16(c) shows the extracted vehicles using the MHI method, where the object sizes were slightly magnified. Figure 16(d) is the extracted vehicles using the proposed method, where the object sizes were reasonably reduced and pruned.

Figure 11: The result from Kalman filtering.

Figure 12: The result from background modeling.

Figure 13: The result from DSETI method.

Figure 14: The result from MHI method.

Figure 15: The result from the joint optical flow and PCA method.
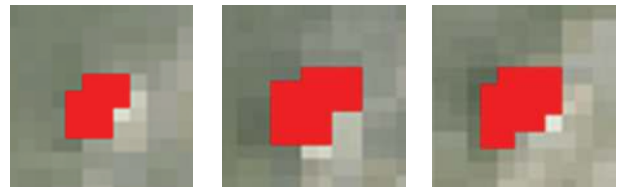

(a) detected vehicles based on corner detection

(b) the three vehicles in (a)

(c) extracted entire region using MHI method

(d) extracted entire region using our method

Figure 16: The result by combining the corner detection and the propose method.

The result using another airborne video is shown in Figure 17, which further demonstrated that our method can better extract object sizes.

4. Conclusion

In this paper, we propose a joint optical flow and PCA approach for motion detection. Instead of considering the original optical flow, the two eigenvalues of the covariance matrix of local optical flows are analyzed. Since the first eigenvalue represents the major motion component and the second eigenvalue represents the minor motion component or turbulence, they are more useful to detect true motions while more successfully suppressing false alarms. The proposed method is also effective in extracting the actual size of moving objects.
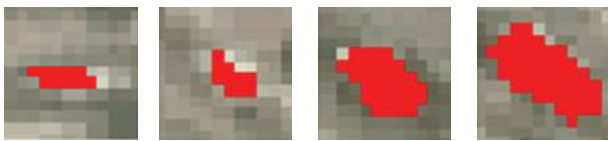
The computational complexity involved in PCA includes the calculation of covariance matrix of local optical flow and its eigen-decomposition. For a mask of size $n \times n$, the number of multiplications in calculating the covariance matrix of

(a) detected vehicles based on corner detection

(b) the four vehicles in (a)

(c) extracted entire region using MHI method

(d) extracted entire region using our method

FIGURE 17: The result by combining the corner detection and the propose method in another airborne video.

size 2×2 is $(2n)^2$, and complexity of eigen-decomposition is generally $O(2^3)$. For an image frame with m pixels, the total computational complexity is $O(m(2n)^2 + m2^3)$. It can be reduced to $O(\beta m(2n)^2)$ if using iterative PCA (IPCA) as discussed in [17], where β is a small integer. As the future work, we will investigate the performance when using IPCA to expedite motion detection.

Acknowledgment

References

[1] A. Mitiche and P. Bouthemy, "Computation and analysis of image motion: a synopsis of current problems and methods," *International Journal of Computer Vision*, vol. 19, no. 1, pp. 29–55, 1996.

[2] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on Systems, Man and Cybernetics Part C*, vol. 34, no. 3, pp. 334–352, 2004.

[3] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: a survey," *ACM Computing Surveys*, vol. 38, no. 4, pp. 1–45, 2006.

[4] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1–3, pp. 185–203, 1981.

[5] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 59, no. 1-2, pp. 81–87, 1993.

[6] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, vol. 2, pp. 674–679, 1981.

[7] A. Bruhn, J. Weickert, and C. Schnorr, "Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods," *International Journal of Computer Vision*, vol. 61, no. 3, pp. 211–231, 2005.

[8] K. Pearson, "On lines and planes of closest fit to systems of points in space," *Philosophical Magazine*, vol. 2, no. 6, pp. 559–572, 1901.

[9] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, vol. 24, no. 6, pp. 417–441, 1933.

[10] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[11] K.-P. Karmann and A. V. Brandt, "Moving object recognition using an adaptive background memory," in *Time-Varing Image Processing and Moving Object Recognition*, V. Cappellini, Ed., vol. 2, pp. 297–307, Elsevier, Amsterdam, The Netherlands, 1990.

[12] P. Kaewtrakulpong and R. Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection," in *Proceedings of 2nd European Workshop on Advanced Video-Based Surveillance Systems(AVBS '01)*, September 2001.

[13] G. Jing, C. E. Siong, and D. Rajan, "Foreground motion detection by difference-based spatial temporal entropy image," in *Proceedings of IEEE TENCON Conference*, pp. A379–A382, 2004.

[14] Z. Yin and R. Collins, "Moving object localization in thermal imagery by forward-backward MHI," in *Proceedings of IEEE Computer Vision and Pattern Recognition Workshops (CVPRW '06)*, 2006.

[15] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," in *Proceedings of the 2nd European Conference on Computer Vision*, vol. 588, pp. 237–252, Springer, 1992.

[16] H. Yang, B. Ma, and Q. Du, "Very small moving object detection from airborne videos using corners in differential images," in *Proceedings of IEEE International Conference on Image Processing*, 2010.

[17] Q. Du and J. E. Fowler, "Low-complexity principal component analysis for hyperspectral image compression," *International Journal of High Performance Computing Applications*, vol. 22, no. 4, pp. 438–448, 2008.