2004 IEEE Intelligent Transportation Systems Conference
Washington, D.C., USA, October 3-6, 2004

Real-Time Detection Of Crossing Pedestrians For Traffic-Adaptive Signal Control

Vignesh Bhuvaneshwar, *Member, IEEE,* and Pitu B. Mirchandani, *Senior Member, IEEE*

Abstract—**In this paper we propose an approach to detect and count pedestrians at an intersection, in real-time, using a fixed camera. After identifying moving objects in sequential images via motion segmentation, median filtering and erosion/dilation operations are performed to suppress noise. Connected component extraction is then employed to extract and label the moving objects in the image. The binary foreground region of each component is projected onto the axis perpendicular to its major axis to obtain a vertical projection histogram from which shadows can be detected, extracted and suppressed. Information about the size and coordinates of each component is then utilized to compute the number of people in the scene.**

Index Terms—**pedestrian detection, connected component extraction, projection histogram**

I. INTRODUCTION

Real-time pedestrian detection using video imagery has myriad applications ranging from traffic monitoring and incident detection to traffic signal control. Previous research on real-time traffic adaptive signal control systems has primarily focused on vehicle traffic for adaptive signal control. Very little research has been conducted on incorporating real-time pedestrian traffic data into adaptive signal control systems. The objective of this paper is to present an approach to detect and count pedestrians at an intersection, in real-time, using a fixed camera. This real-time data can then be incorporated into a system for traffic adaptive signal control. We envision that our approach could also be used to detect and track bicycles among other vehicles.

V. Bhuvaneshwar is a graduate student with the Systems & Industrial Engineering, and Electrical & Computer Engineering departments at the University of Arizona, P.O. Box 210020 Tucson, AZ 85721-0020, USA (E-mail: vignesh@ece.arizona.edu)

P. B. Mirchandani is a Professor, Systems & Industrial Engineering, Electrical & Computer Engineering, and Director ATLAS Research Center, University of Arizona (E-mail: pitu@sie.arizona.edu)

Fixed cameras can be mounted on poles or atop tall buildings overlooking intersections and output real-time imagery in the form of video or individual frames at a desired frequency, say, once frame per second. The background image for these cameras is fixed. Several background segmentation models (e.g., Stauffer [1], and Haritaoglu [2]) have been proposed in the literature for real-time implementation, and we assume the availability of an algorithm that predicts and continuously updates a background image.

The background image is subtracted from the first "real-time" frame to obtain the "difference image" that detects the moving objects in the frame. Median filtering and thresholding are applied to the difference image to generate a binary image. The moving objects are represented as blobs in this binary image using connected component extraction. Vehicles and other moving objects are differentiated from pedestrian traffic using information about the height and the area of each blob. A novel shadow removal algorithm (discussed in Section IV of the paper) has been developed that detects and remove shadows from the scene. Thus, the designed system outputs an approximate count of the number pedestrians in the image, and their coordinates. The above-mentioned steps are then applied to the next (successive frame) taken, say, one second after the first. Once pedestrians in consecutive frames have been detected, they are then matched and tracked using a maximum likelihood-matching algorithm (see [14] for more details). The pedestrian detection algorithm is summarized in the flow chart in Figure 1.

The rest of this paper is organized as follows. Section II reviews related work in the area of pedestrian detection through image processing. Section III describes the proposed approach for pedestrian detection. Section IV describes our shadow detection approach in detail and shows some implementation results. Conclusions are given in Section V.

II. RELATED WORK

The literature on pedestrian detection and tracking is vast, and a comprehensive survey is beyond the scope of this paper. We only mention some representative works.

0-7803-8500-4/04/$20.00 ©2004 IEEE

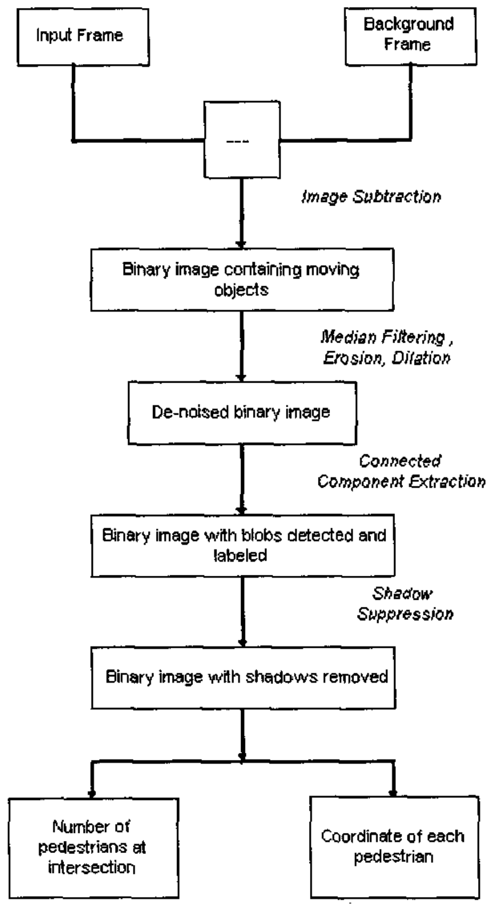

Figure 1. Flow chart summarizing the methodology for pedestrian detection.

The W4 system [2] is a real-time visual surveillance system that has been designed to detect, track, and monitor people in an outdoor environment. W4 uses a statistical background model (bimodal distribution of intensity at each pixel) to isolate the moving objects in the image, and then connected component extraction and silhouette shape analysis to locate people, classify their body parts (head, arms, torso, etc.), and detect body postures (standing, sitting, lying, and bending). A motion prediction and template-matching model is then employed to track the detected foreground objects in subsequent frames.

In [3] the authors propose an approach to detect pedestrians and vehicles at an intersection for incident prevention applications. The system uses the motion segmentation approach of [1] to separate foreground regions from the background. Connected component analysis is employed to label individual blobs. The blobs are tracked by constructing an undirected bi-partite graph whose vertices are comprised of the blobs detected in the previous and the current frames. In [4] the authors propose an approach similar to the one proposed above for tracking pedestrians in real-time using a single camera.

Bertozzi et al. [5] proposes a vision-based system that uses the vertical symmetry of pedestrians and the presence of strong edges to calculate a set of candidate pedestrian regions. The bounding boxes that span these regions are then refined using a Kalman-filter estimator. Qui et al. [6] detects pedestrians and bicycles in urban traffic scenes using a feature extraction algorithm. After background subtraction, edge and corner detection algorithms are employed to extract feature points from the moving objects in the frame. These feature points are grouped into objects using a "grouper" and are tracked in subsequent frames using a prediction algorithm. Other approaches for pedestrian detection and tracking include neural-network based techniques [9][10], and stereo-based algorithms [7][8] (see [6] for more details).

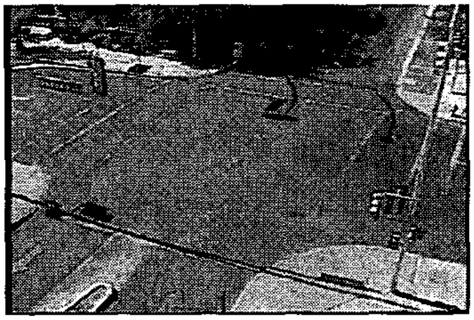

Figure 2. The background frame.

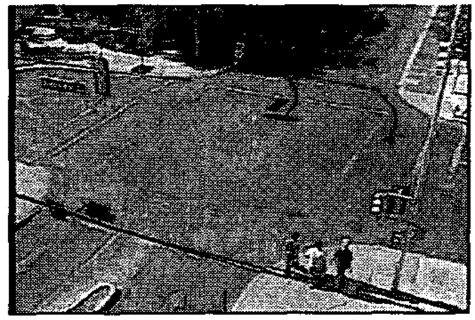

Figure 3. The current frame.

Figure 4. The thresholded difference image.

III. PROPOSED APPROACH FOR PEDESTRIAN DETECTION AND COUNTING

Motion segmentation is the primary step in detecting the moving objects in a scene. As mentioned earlier, we assume the presence of a robust background subtraction model similar to the approach of Stauffer [1]. For our analysis in this paper, the background image has been extracted directly from a video sequence captured by the fixed camera mounted on a tall pole, overlooking an intersection in Tucson, Arizona. The background and the current frame are shown in Figures 2 and 3 respectively.

The background image is subtracted from the current frame to obtain the "difference image" (see Figure 4) that identifies the moving objects in the frame. Median filtering and thresholding are applied to the difference image to obtain a binary image. Morphological operators erosion and dilation are applied to this binary image to generate a de-noised binary image (see Figure 5).

The moving objects represented as blobs in this binary image are extracted using connected component extraction (see Figure 6). Vehicles and other moving objects are differentiated from pedestrian traffic using information about the height and the area of each blob. A shadow removal algorithm (discussed in the next section) is then employed to detect and remove shadows from the image. Figure 7 shows the binary image after shadows have been removed. The pedestrian detection algorithm gives an approximate count of the number pedestrians in the image and their coordinates. The above-mentioned steps are then applied to the next frame taken, say, one second after the first. Once pedestrians in consecutive frames have been detected, they are matched and tracked using a maximum likelihood-matching algorithm (see [14] for more details).

IV. SHADOW SUPPRESSION

One of the primary concerns surrounding background subtraction and thresholding is that shadows are also detected as part of the moving object, as shown in Figure 4. Since our approach uses the size of each component to estimate the number of pedestrians in the scene, the presence of shadows could result in a significant amount of error in the detected number of pedestrians. We propose a novel technique for shadow suppression using the vertical projection histogram of each detected component in the current image (for other techniques for shadow detection and removal see [11][12][13]).

Figure 5. The denoised binary image.

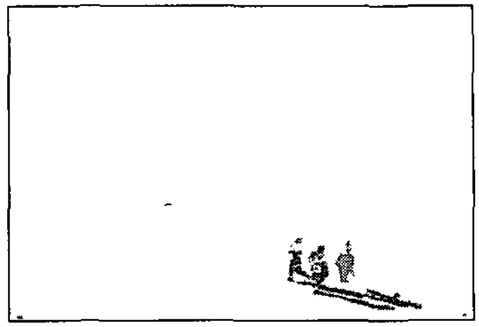

Figure 6. The image after connected component labeling.

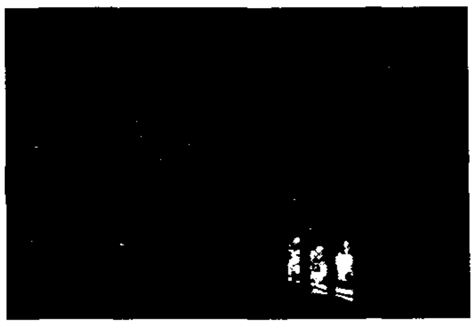

Figure 7. The image after shadow removal.

The vertical histogram projection of a component is simply the plot of the number of white pixels of that particular component versus the column index. The vertical projection histogram of the current frame is depicted in Figure 8. As seen in this figure, the shadow pixels are much lower in number than the pixels representing pedestrians. Therefore, by suppressing those columns of pixels that are below a predefined histogram threshold, it is possible to remove most of the shadow pixels as shown in Figure 7. The number of pedestrians in the scene can be estimated using the area of each component. Bounding boxes have been drawn around detected pedestrians in Figure 9.

We are also currently testing a second approach for estimating the number of pedestrians directly from the histogram. As seen in Figure 8, the three peaks of the histogram correspond to three pedestrians in the scene, each pedestrian being about 50-70 pixels wide. Our algorithm for counting pedestrians is given below.

- ***Step 1:*** Find the points of maxima (peaks) in the projection histogram.
- ***Step 2***: Sort the peaks in the descending order of height.
- ***Step3:*** Starting from the highest peak P_1 find the distance D to the next highest peak P_2. If D is lower than a threshold (in this case 50 pixels), peak P_2 belongs to the same object, and is rejected. If D is higher than the threshold, peak P_2 represents another pedestrian.
- ***Step4:*** Repeat step 3, this time, finding the distance to the next highest peak from all previously detected pedestrian peaks (e.g., P_1 , $P_2 \ldots P_n$), and continue until all peaks have been considered.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we propose an efficient approach for pedestrian detection and counting at an intersection, in real-time, using a single, fixed camera. After the moving objects in the image have been detected using background subtraction and connected component extraction, vehicles are differentiated from pedestrians using information about the size and height of each component. A projection histogram of each blob (component) on an axis perpendicular to the major axis of the blob (vertical projection histogram) is then employed to suppress shadows.

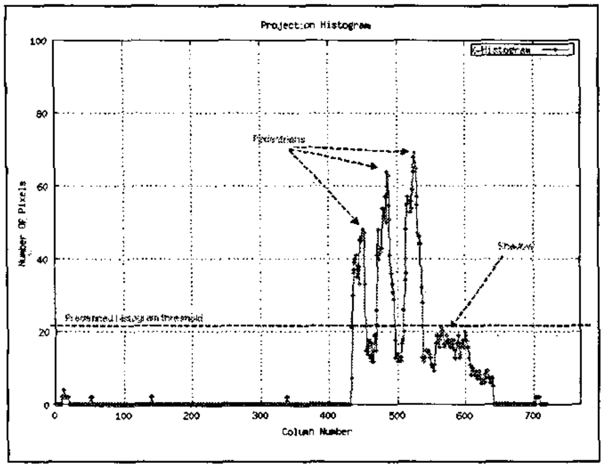

Figure 8. The vertical projection histogram of the current image.

Figure 9. Pedestrians detected

Our preliminary results are encouraging and we are confident that a system can be built using the approach and system architecture described in this paper. One area of future work would be to rigorously test our approach over a wide variety of images. For traffic adaptive signal control, only a good approximation on the number of pedestrians at a crosswalk is needed since a pedestrian-signal phase can accommodate several pedestrians at the same time and one needs only to provide a phase duration that accommodates the pedestrian cross time.

REFERENCES

[1] C. Stauffer and W.E.L. Grimson. "Adaptive Background Mixture Models for Real-Time Tracking". Proceedings of *The Computer Vision and Pattern Recognition Conference (CVPR)*, June 1999.

[2] I. Haritaoglu, D. Harwood, and L.S. Davis. "W4:Real-Time Surveillance of People and Their Activities". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, Aug 2000.

[3] H. Veeraraghavan, O. Masoud, N. Papanikolopoulos."Vision-Based Monitoring of Intersections". Proceedings of the *IEEE 5th International Conference on Intelligent Transportation Systems*, pp.7—12, Sep. 2002.

[4] O. Masoud, N. Papanikolopoulos. "A Novel Method for Tracking and Counting Pedestrians in Real-Time Using a Single Camera", *IEEE Transactions on Vehicular Technology*, Vol. 50, No. 5, pp. 1267—1278, Sep. 2001.

[5] M. Bertozzi, A. Broggi, R. Chapuis, F. Chausse, A. Fascioli, and A. Tibaldi. "Shape-Based Pedestrian Detection and Localization". Proceedings of the *IEEE Conference on Intelligent Transportation Systems*, Vol. 1, pp. 323—333, Oct. 2003.

[6] Z. Qiu, D. Yao, Y. Zhang, D. Ma, and X. Liu. "The Study of the Detection of Pedestrian and Bicycle Using Image Processing". Proceedings of the *IEEE Conference on Intelligent Transportation Systems*, Vol. 1, pp. 340--345, Oct. 2003.

[7] L. Zhao and C. Thorpe. "Stereo and Neural Network-Based Pedestrian Detection". *IEEE Transactions on Intelligent Transportation Systems*, Vol. 1, pp. 148--154, Sep. 2000.

[8] K. Fujimoto, H. Muro, N, Shimomura, T. Oki, Y. K. K. Maeda, and M. Hagino. "A Study on Pedestrian Detection Technology Using Stereo Images". *JSAE Review*, Vol. 23, pp. 383--385, Aug. 2002.

[9] C. Wohler, J. K. Aulaf, T. Portner, and U. Franke. "A Time Delay Neural Network Algorithm for Real-Time Pedestrian Detection". Proceedings of the *IEEE Intelligent Vehicles Symposium*, pp. 247-251, Oct. 1998.

[10] C. Wohler, U. Krebel, and J. K. Anlauf. "Pedestrian Recognition by Classification of Image Sequences – Global Approaches vs. Local Spatio-Temporal Recognition", Proceedings of the *IEEE International Conference on Pattern Recognition*, Sep. 2000.

[11] A. Bevilacqua. "Effective Shadow Detection in Traffic Monitoring Applications", *Journal of WSCG*, Vol.11, No.1, pp. 56—64, Feb 2003.

[12] A. Bevilacqua M. Roffilli. "Robust denoising and moving shadows detection in traffic scenes", Proceedings of the *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1—4, Dec. 2001.

[13] A. Prati, I. Mikic, M. M. Trivedi, R. Cucchiara, "Detecting Moving Shadows: Algorithms and Evaluation", *IEEE Transactions on pattern Analysis and Machine Intelligence*, Vol. 25, No. 7, pp. 918—923, Jul. 2003.

[14] A. Angel, M. Hickman, D. Chandnani, and P. Mirchandani. "Methods of traffic data collection using aerial video", Proceedings of the *IEEE 5th International Conference on Intelligent Transportation Systems*, September 2002.