Vehicle Detection and Tracking for Freeway Traffic Monitoring

Brian Gloyer[1] Hamid K. Aghajan[2] Kai-Yeung Siu[1] Thomas Kailath[2]

[1] Department of Electrical and Computer Engineering
University of California, Irvine, CA 92717

[2] Information Systems Laboratory, Department of Electrical Engineering
Stanford University, Stanford, CA 94305

Abstract

In this paper an image processing and object tracking approach is proposed for the design of a video-based freeway traffic monitoring system. Proper estimation of the traffic speed in different lanes of a freeway allows for timely detection of possible congestions. The proposed method consists of a road modeling stage and a vehicle tracking stage. In the first stage, a three-dimensional model of the background road image is generated. In the tracking stage, each car in the scene is isolated and tracked over many frames. Experimental results on frame sequences taken from Santa Ana area freeways will be presented.

1 Introduction

Vision-based traffic monitoring systems have received considerable attention during the past several years. Monitoring the traffic on different lanes of a freeway helps to detect and mitigate possible congestions in a timely manner by allowing the traffic management centers to apply effective response strategies. A number of incident detection algorithms have been proposed in the last decade for freeway surveillance and control systems. Pattern recognition algorithms compare predetermined incident patterns with the received data to detect any such incident occurrences [1, 2, 3]. Time series approaches use statistical indicators to detect deviations from normal traffic conditions [4, 5, 6]. A comprehensive review of existing image processing systems for freeway monitoring appears in [7, 8].

In this paper, an image processing approach is proposed for designing a video-based freeway traffic monitoring system. The proposed method consists of a modeling and a tracking stage. In the modeling stage, a 3-D model of the background road image is generated. In the tracking stage, each car in the scene is isolated and tracked over many frames, and estimates of the vehicle speeds in different freeway lanes are obtained. Experimental results on frame sequences taken from Santa Ana area freeways will be presented.

2 Road Modeling

The developed technique uses an initialization algorithm to build a model of the road in order both to compare the frames with a background image, and to create a mapping between the detected vehicle locations on each frame to the three-dimensional space. This direct mapping makes the run-time algorithm more computationally efficient. In the later stage of tracking, images of moving cars are isolated in each frame and tracked over consecutive frames. Figure 1 shows two frames of a sequence taken from a Santa Ana area freeway. For making comparisons with each frame to detect the moving objects, a background image is generated from an initial set of frames. Figure 2 shows the flowchart of the initialization algorithm.

2.1 Background Image Generation

The task of generating the background from the frames is repeated from time to time to incorporate any changes in the illumination and general conditions of the freeway. This choice makes the vehicle detection procedure adaptive to the scene conditions. The background image is found by looking at different frames in the image sequence. In any one frame, parts of the road are covered by cars. As time goes on, the cars will move and reveal the covered road. If the sequence is long enough, a clear picture of the car-free road can be found. Car images have different intensities while

1058-6393/95 $4.00 © 1995 IEEE

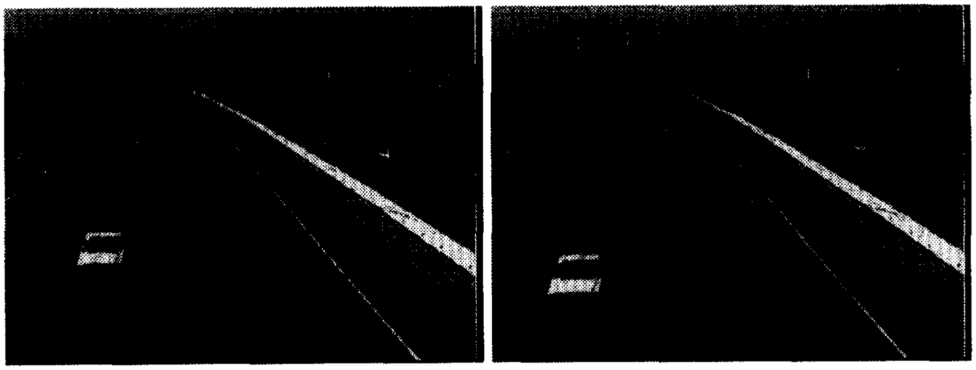

Figure 1: Two frames of the test sequence.

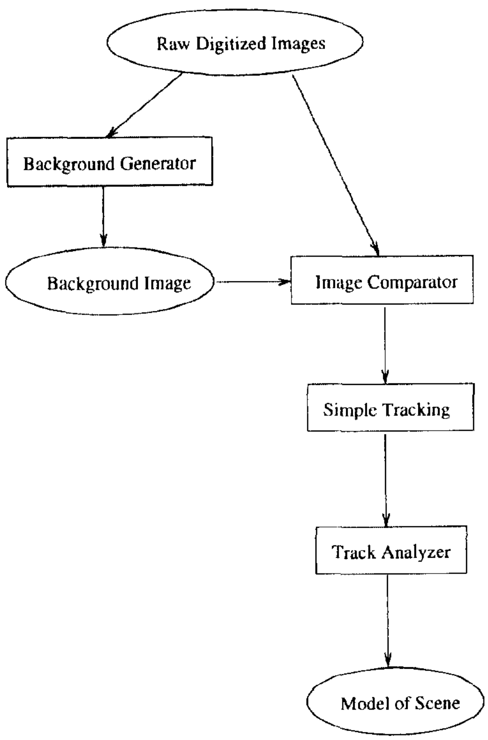

Figure 2: The road modeling stage.

the intensity of the pixels representing points on the road remains the same. The background image is generated pixel by pixel. The intensity of each point is observed in several initialization frames. The intensity value that occurred most often is chosen to be the background value at that point. This procedure avoids filtering and blurring of the background image. Figure 3 shows the result of background generation for the sequence of Figure 1.

2.2 Road Model Generation

A three dimensional geometrical model for the road is produced by tracking cars in a set of frames and extracting their spatial coordinates. This model will greatly reduce the computation costs during the run-time by allowing direct 2-D image mapping.

Assume a general setting for a camera imaging a rigid object (a car) located at the point $P = (X, Y, Z)^T$ in the coordinate system (x, y, z) attached to the camera, and moving with the translational velocity vector $V = (\dot{X}, \dot{Y}, \dot{Z})^T = (v_x, v_y, v_z)^T$. The center of projection is assumed to be at the origin, and the image plane is parallel to the $x - y$ plane at a distance f from the origin, where f is the focal length of the camera. Figure 4 shows such a setting.

We will consider the parameters of the moving point in this coordinate system. In practice the camera can be slightly rotated around the y-axis by a known angle ϕ to mitigate the horizon effect and to include a larger portion of the freeway scene in the image. The point P is imaged on the camera's image plane at the image point (x', y'), which is related to the 3-D coordinates of P via

$$x' = \frac{X}{Z} f \quad (1)$$

$$y' = \frac{Y}{Z} f \quad (2)$$

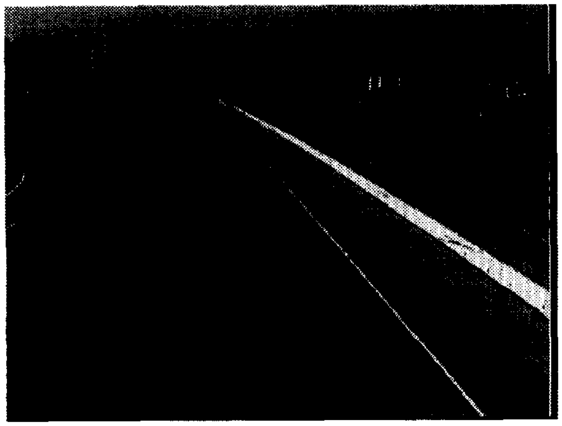

Figure 3: Background image generated from the sequence.

We assume constant velocities for the cars over a few frames, and take the derivatives of the 3-D coordinates X, Y, and Z. For the X coordinate, we will have

$$X_n - X_{n+1} = X_{n+1} - X_{n+2} \tag{3}$$

or

$$X_{n+2} = 2X_{n+1} - X_n \tag{4}$$

and similarly:

$$Y_{n+2} = 2Y_{n+1} - Y_n \tag{5}$$

$$Z_{n+2} = 2Z_{n+1} - Z_n \tag{6}$$

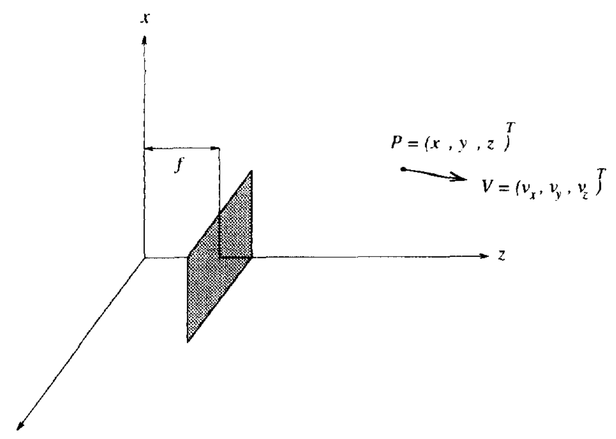

Figure 4: Imaging geometry.

Substituting Eqs. (4) and (6) into Eq. (1) results in

$$2Z_{n+1} - Z_n = \frac{f(2X_{n+1} - X_n)}{x'_{n+2}} \tag{7}$$

Using Eq. (1) again, we obtain

$$2Z_{n+1} - Z_n = \frac{2Z_{n+1}x'_{n+1} - Z_n x'_n}{x'_{n+2}} \tag{8}$$

Finally,

$$Z_{n+1} = Z_n \frac{x'_{n+2} - x'_n}{2(x'_{n+2} - x'_{n+1})} \tag{9}$$

With this recurrence, the relative distances the vehicles travel can be found. If the real distances are to be found, the real position of some point in the scene must be known. To accomplish this, it is assumed that the real position of the bottom row in the image plane in known. This is measured when the camera is installed. These positions can be used as Z_0 in the recurrence equations.

Figure 5: Three dimensional model of the road.

In practice, x' and y' can be labeled as the coordinates of the center of gravity of the moving object's image. The tracking stage yields estimates of x' and y', and in turn can take advantage of the estimated coordinates for tracking the cars in the following frame.

By observing all of the tracks, a three-dimensional map of the road can be built. The real-world coordinates of any point in the image plane can be found by using this map. Figure 5 shows the 3-D model of the road for our images.

3 The Tracking Stage

The road model and the background image are used in the tracking algorithm to detect and map the vehicles onto the real 3-D scene. Figure 6 shows the flowchart of the tracking algorithm. Once an accurate

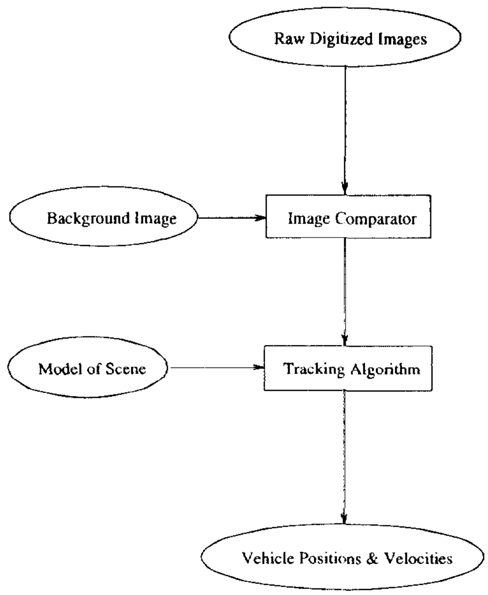

Figure 6: The tracking stage.

background image has been found, comparing each frame to the background image results in a difference map that is thresholded to form a binary image including images of moving patterns. Figure 7 shows the result of this comparison. In the difference image, the cars appear as large solid blobs, whereas noise and camera jitter produce spurious non-zero difference values that generally span over only a few pixels. These extra points are removed at the low level processing step. Connected component methods are used to connect the blobs in each frame. For each blob, the position, the area, and an estimate of its density is computed. The blobs that are small or have a low density are considered to be noise. The position of the blobs in each frame is used by the tracking algorithm to follow each vehicle through the sequence. The tracking algorithm considers each frame in a three dimensional space defined by stacking several frames, and locates nearest blobs in the neighboring frames. Initial appearance of cars at the far distant in the sequence is handled by setting a threshold in the search for nearest neighbors.

The tracking algorithm maintains track of all the vehicles on the road, and also makes an estimate of the expected position of the blobs in the next frame.

The image coordinates of the detected and tracked blobs are mapped onto the generated 3-D model of the road to obtain the three-dimensional car coordinates in the scene. This in turn leads to an estimate of the car velocity. Figure 8 presents the result of the tracking stage on the whole sequence, and in Figure 9 the detected vehicle locations are superimposed on the road image.

4 Conclusions

An algorithm has been proposed for monitoring speed of cars in different lanes of a freeway. The method consists of a road modeling stage and a tracking stage. Proper estimation of the traffic speed in different freeway lanes allows for timely application of effective response strategies to possible congestions. Experimental results on frame sequences taken from Santa Ana area freeways indicate promising performance efficiency for the proposed algorithm.

Acknowledgments

This work was supported in part by the NSF Young Investigator Awards Programs under Grant MIP-9357553 and by the Institute of Transportation Studies of the University of California, Irvine. The work of H. Aghajan and T. Kailath was supported in part by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Air Force Office of Scientific Research under contract F49620-93-1-0085.

References

[1] B. N. Persaud, F. L. Hall, and L. M. Hall, Congestion identification aspects of the McMaster incident detection algorithm, *Transportation Research Record*, **1287**:151–166, 1990.

[2] P. H. Masters, J. K. Lam, and K. Wong, Incident detection algorithms for COMPASS-an advanced traffic management system, In *Proc. Vehicle Navigation & Information Systems Conference*, pages 295–310, Dearborn, MI, 1991.

[3] D. Bullock, J. Garrett, and C. Hendrickson, A neural network for image-based vehicle detection, *Transportation Research - C*, **1**(3):235–247, 1993.

[4] S. A. Ahmed and A.R. Cook, Discrete dynamic models for freeway incident detection systems, *Transportation Planning and Technology*, **7**:231-242, 1982.

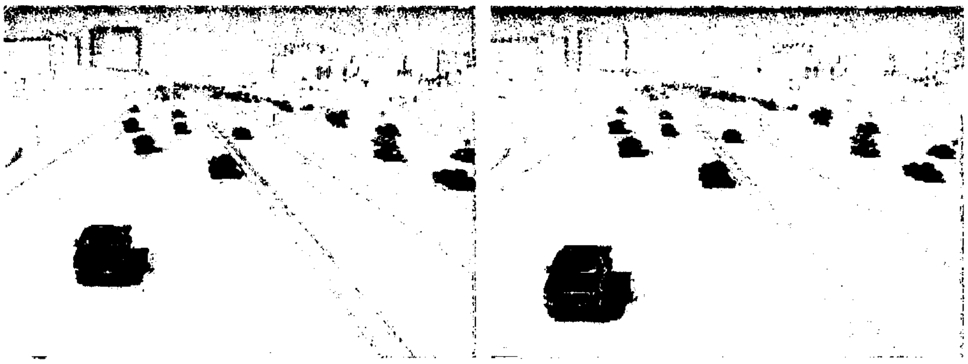

Figure 7: Moving objects detected by comparison with the background.

[5] S. G. Ritchie and R. L. Cheu, Simulation of freeway incident detection using artificla neural networks, *Transportation Research - C*, **1**(3):203–217, 1993.

[6] Y. J. Stephanedes and A. P. Chassiakos, Freeway incident detection through filtering, *Transportation Research - C*, **1**(3):219–233, 1993.

[7] S. Hockaday, Evaluation of image processing technology for application in highway operation, Technical Report TR 91-2, California Polutechnic State University, San Luis Obispo, CA, 1991.

[8] S. Hockaday, A. Chatziioanou, R. Nodder, and S. Kuhtenschmidt, Evaluation and comparison of video image processing systems for traffic detection, Preprint # 920744, Transportation Research Board, 1992.

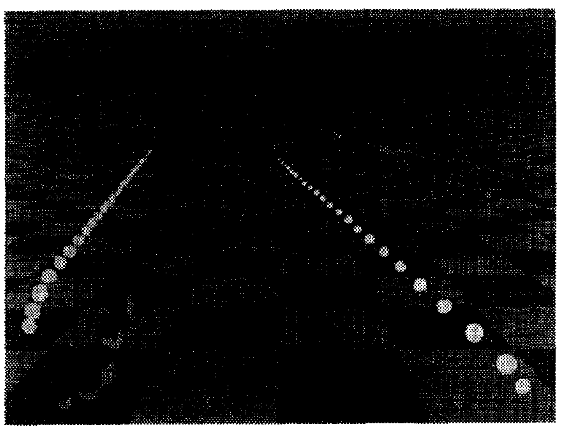

Figure 8: Track of moving vehicles on the 3-D road model.

Figure 9: Estimated car tracks superimposed on the road image.