A Novel Method for Resolving Vehicle Occlusion in a Monocular Traffic-Image Sequence

Clement Chun Cheong Pang, William Wai Leung Lam, and Nelson Hon Ching Yung

***Abstract*—This paper presents a novel method for resolving the occlusion of vehicles seen in a sequence of traffic images taken from a single roadside mounted camera. Its concept is built upon a previously proposed vehicle-segmentation method, which is able to extract the vehicle shape out of the background accurately without the effect of shadows and other visual artifacts. Based on the segmented shape and that the shape can be represented by a simple cubical model, we propose a two-step method: first, detect the curvature of the shape contour to generate a data set of the vehicles occluded and, second, decompose it into individual vehicle models using a vanishing point in three dimensions and the set of curvature points of the composite model. The proposed method has been tested on a number of monocular traffic-image sequences and found that it detects the presence of occlusion correctly and resolves most of the occlusion cases involving two vehicles. It only fails when the occlusion was very severe. Further analysis of vehicle dimension also shows that the average estimation accuracy for vehicle width, length, and height are 94.78%, 94.09%, and 95.44%, respectively.**

***Index Terms*—Composite signature, curvature, monocular traffic image sequence, occlusion, signature decomposition, vanishing point.**

I. Introduction

AUTOMATED visual traffic surveillance (AVTS) has been actively investigated in the past decade, as it can potentially generate a large amount of useful information for other intelligent transportation system (ITS) applications [1]–[4]. As the name implies, AVTS allows the visualization of vehicles on the road by using a single camera (monocular vision) mounted in perspective view of the road segment that it is monitoring, thus enabling traffic-scene analysis, such as traffic-conditions assessment and travel-speed estimation, as well as queue-length measurement, in which traditional nonvisual surveillance systems could not do [5]. However, the performance of an AVTS system deteriorates when vehicles appear to occlude each other from the camera's point of view in a traffic-image sequence [6]. As a result, methods for occlusion detection and resolution must be adopted in order to produce meaningful results. These include using stereo vision [7]–[9], an overhead camera with a viewing axis perpendicular to the road surface [1], or nonvisual sensors such as an ultrasonic infrared or laser to aid the vision system [10], [11]. Insofar as the problem is concerned,

Manuscript received January 12, 2004; revised April 26, 2004 and May 9, 2004. The Associate Editor for this paper was L. Vlacic.

The authors are with the Laboratory for Intelligent Transportation Systems Research, Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong SAR, China (e-mail: ccpang@eee.hku.hk; wllam@eee.hku.hk; nyung@eee.hku.hk).

Digital Object Identifier 10.1109/TITS.2004.833769

Fig. 1. Vehicle occlusion. (a) Same lane and (b) adjacent lanes.

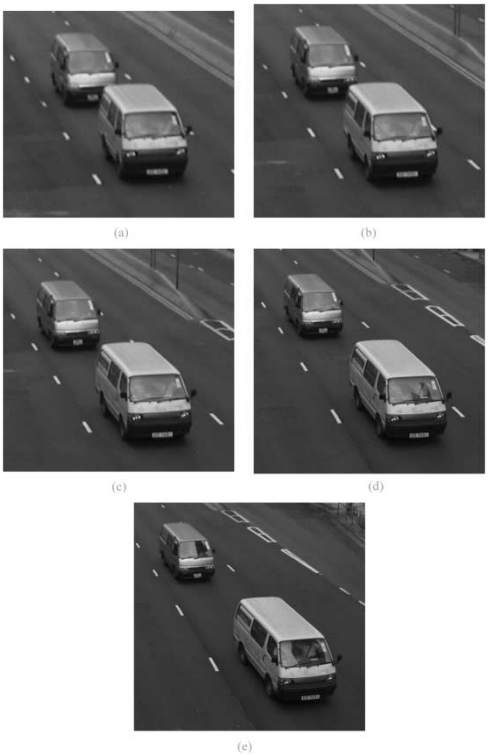


Fig. 2. Occlusion event (frame x–y, from left to right).

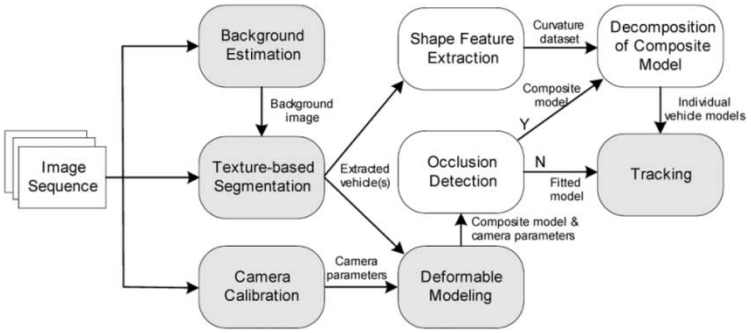


Fig. 3. Conceptual diagram of model-based vehicle modeling.

it is considered that stereo vision makes little difference from monocular vision in this case, while changing the camera configuration to overhead mounted reduces the variety of vehicle information obtainable. The added installation and signal-processing complexities of nonvisual sensors simply makes it prohibitive. Therefore, we are motivated to challenge the problem of resolving vehicle occlusion from an image sequence captured from a single roadside mounted camera.

This paper presents a novel method for resolving the occlusion of vehicles seen in a sequence of traffic images taken from a single roadside mounted camera. Its concept is built upon a previously proposed vehicle-segmentation method, which is able to accurately extract the vehicle shape out of the background without the effect of shadow and other visual artifacts. Based on the segmented shape and that the shape can be represented by a simple cubical model, we propose a two-step method: first, detect the curvature of the shape contour to generate a curvature data set of the occluded vehicles and, second, decompose it into individual vehicle models using a vanishing point in three dimensions and the set of curvature points of the composite model. The proposed method has been tested on a number of monocular traffic image sequences and found that it detects the presence of occlusion correctly and resolves most of the occlusion cases involving two vehicles. It only fails when the occlusion was very severe. Further analysis of vehicle dimension also shows that the average estimation accuracy for vehicle width, length, and height are 94.78%, 94.09%, and 95.44%, respectively.

The rest of this paper is organized as follows. The problem analysis is detailed in Section II. Following that, a literature review is given in Section III. The proposed method is presented in Section IV and experimental results and the discussion are depicted in Section V. The conclusion can be found in Section VI.

II. Problem Analysis

In this paper, occlusion refers to the overlapping of vehicles, visually in an image due to the perspective view of the camera (Fig. 1). Occlusion is a dynamic event changing over the duration, in which vehicles appear within the field of view. For instance, vehicles may occlude for a short time and then separate after that or vice versa (Fig. 2). Failing to detect and resolve the presence of occlusion may lead to surveillance errors, including incorrect vehicle count, incorrect tracking of individual vehicles, and incorrect classification of vehicle type on that road segment. However, occlusion detection and resolution are inherently complex, as they rely on *a priori* vehicle features that would indicate whether a particular moving object consists of one or more than one vehicle. If it is the latter case, then these features would have to provide a basis for differentiating which vehicle is which. The question is often regarding what kind of features would enable the detection and how to utilize what kind of features to resolve the occlusion. This paper intends to provide an answer to this question.

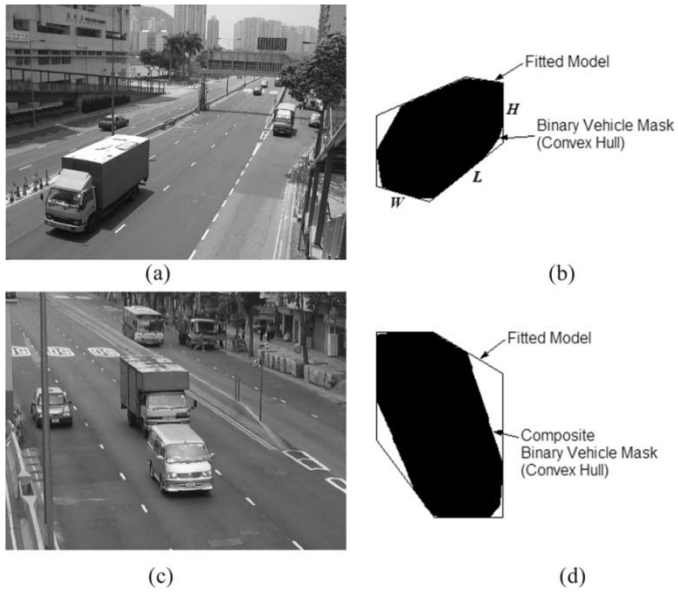


Fig. 4. Vehicle model of occluded and unoccluded cases (a) vehicle image, (b) deformable model, (c) two-vehicle occlusion, and (d) occluded model.

III. Related Works

Researchers have done an extensive amount of work on the problem of vehicle tracking in the past [12]–[35]. They have addressed the problem of occlusion detection and occlusion handling as the subset of their investigation on vehicle tracking. A brief review on some of the proposed methods for resolving occlusion follow.

Ikeda *et al.* [14] proposed a split and merging algorithm for handling occlusion in traffic surveillance. Their idea is to approximate rectangles onto the binary mask of the vehicle based on the geometric property of the extracted vehicle mask. To generate rectangles that have a one-to-one correspondence with the vehicles, adjacent rectangles are split and merged under speci-

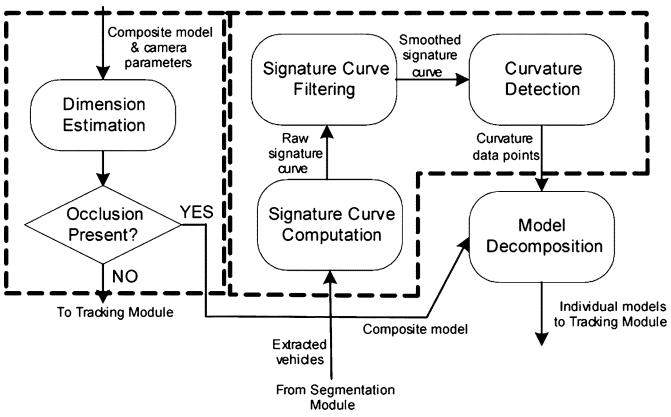


Fig. 5. Conceptual flow of the proposed method.

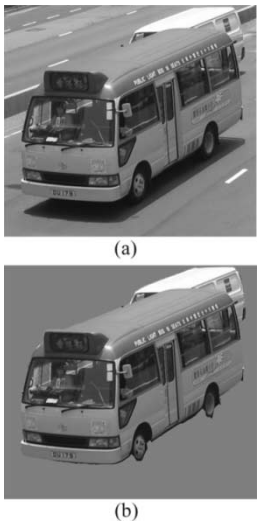


Fig. 6. Texture-based vehicle segmentation: (a) original image and (b) segmented image.

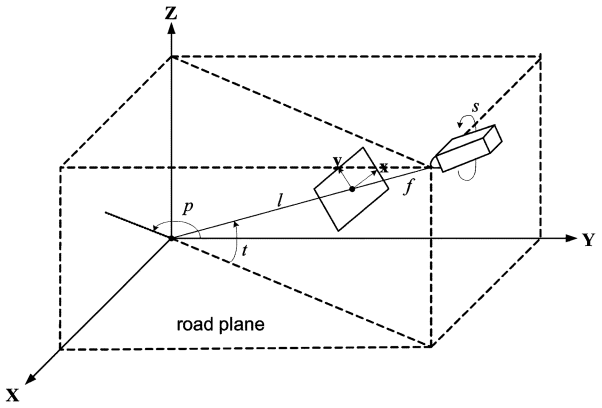


Fig. 7. Camera model.

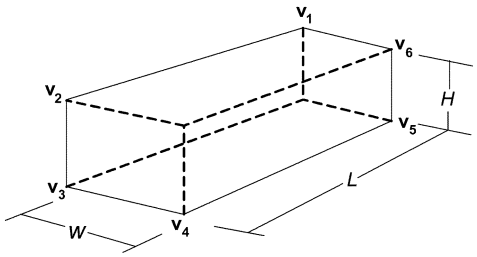


Fig. 8. Generalized deformable model.

fied rules so that redundant rectangles can be eliminated. It was claimed that the algorithm can achieve an accuracy of 94.8% for counting vehicles under occlusion. However, by using the rectangular models, only two dimensions of the vehicles can be identified (either front or back); thus, the algorithm cannot be used to classify the type of vehicles effectively.

Kamijo *et al.* [20] proposed a spatio-temporal Markov random model for segmenting vehicles under occlusion and confusion at an intersection. The basic idea is to divide the image taken at an intersection into small square blocks. The algorithm then determines to which vehicle each block should belong by computing the texture correlation of blocks between consecutive images, as well as the neighboring blocks within an image and segmenting the blocks into most likely objects by the stochastic process. The proposed algorithm has the advantage that it can be employed to monitor traffic at an intersection where there are extensive vehicle occlusions and confusions. The success rate of counting is claimed to be 95%. However, the major drawback of the proposed algorithm is that it can only count or track vehicles, but cannot estimate the dimension of the vehicles. Therefore, it cannot be employed for vehicle-type classification.

Masoud *et al.* [25] proposed a blob-based algorithm to track vehicles under confusion at weaving sections on highways. The blobs representing the position of the vehicle are first extracted from the image by a feature-extraction technique. A blob relationship graph representing the split and merge of the blobs between each subsequent frame is then computed. Vehicles are then associated to each blob with the aid of the blob graph. The idea is that if a vehicle is related to a blob in frame $(i-1)$ and if that blob is related to another blob in the ith frame (through split or merge), then the vehicle is also related to the latter blob. This algorithm was able to track and count vehicles as well as extracting the velocity and direction of each vehicle at weaving sections of highways at an average accuracy of 85%. The proposed algorithm has the advantage that it has addressed the problem of lane changing on highways, which frequently occurs in traffic images. However, similar to Kamijo's method, it was not able to estimate the dimensions of the vehicle and, thus, it cannot classify the type of vehicles using the road, which is a very important part of AVTS for traffic-volume measurement.

Lai [1] proposed a partitioning algorithm for handling occlusion in AVTS. His idea was to fit a six-vertex deformable model onto the binary masks of the vehicles based on the geometric property of the binary vehicle mask, as well as the direction of the road. In the case when there is occlusion, the binary vehicle masks are fused together and, thus, one model is fitted onto the occluded vehicles instead of several models. To generate models that have one-to-one correspondences with the vehicles, the fitted model is partitioned into equal portions of sub-models. The number and direction of partitions were determined by the width and length ratios between the occluded model and the nonoccluded model. This algorithm was able to improve the vehicle trajectories when tracking the individual vehicles. How-

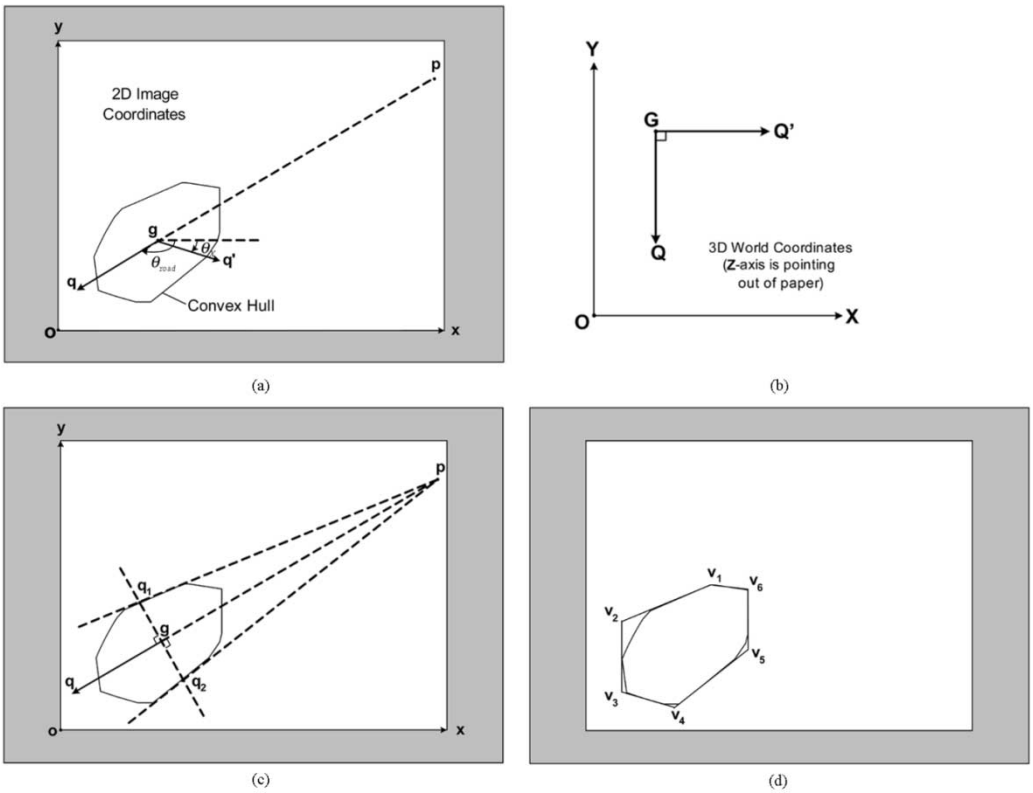


Fig. 9. Fitting of deformable model onto a vehicle. (a) Obtain $\overrightarrow{\mathbf{gq}'}$. (b) Obtain $\overrightarrow{\mathbf{GQ}'}$. (c) Obtain points $\mathbf{q}_1$ and $\mathbf{q}_2$. (d) Fitted model $\mathbf{v}$.

ever, the assumption of partitioning the occluded model into equal portions is questionable, since vehicles that involved in an occlusion might be of different sizes. Moreover, this method of model partitioning has not considered the severity of occlusion, which might eventually lead to a large dimension-estimation error in the partitioned model.

IV. Proposed Method

A. Philosophy and Concept

On the question of what kind of features would enable the detection of occlusion, we have adopted Yung and Lai's model [6], which uses a deformable cubic model to fit a segmented moving object. The simplicity of the model permits us to determine object dimension based on a set of two-dimensional (2-D)/three-dimensional (3-D) transformation equations derived from a calibrated camera model [36]. The object dimension indicates whether occlusion is present and what kind of occlusion it is (side-by-side or front-and-back). Fig. 3 depicts the flow of this concept and Fig. 4 illustrates the results of modeling and dimension estimation for the unoccluded and two-vehicle occlusion cases. In the latter, the length of the object dimension is much longer in relation to the width and height of the object.

On the second question of how to utilize what kind of features to resolve the occlusion, we propose to utilize the feature points (curvature) of the occluded object contour and decompose the occluded vehicles by model fitting individual vehicles using the geometrical property of the feature points. As in Yung and Lai's model, the fitting is performed in three dimensions that the model vectors are extended along the road direction, which converge to the vanishing point $\mathbf{p}$ due to the perspective effect. Fig. 5 depicts the flow of this concept.

B. Preprocessing

1) Background Estimation and Texture-Based Segmentation: The background-estimation algorithm adopted in this paper is the "scoreboard algorithm," as described in [37], which has the advantage of being fast and accurate for estimating the stationary background. The idea of the algorithm is to estimate the background either by the running-mode or running-average method. The decision of using which method is based on the intensity difference of the pixels between the current frame and the previously estimated frame. If the difference is high, then the running-mode method is used. Otherwise, the running-average method is used. By selecting the method in this way, the speed of the algorithm is greatly improved and accuracy is not severely affected.

For vehicle segmentation, we have employed the texture-based vehicle-segmentation approach described in [38] to extract the moving vehicles. This algorithm is capable of extracting the vehicle body without being affected by the vehicle cast shadow. The segmentation process contains three steps. First, three likelihood maps, namely, the T-map, L-map, and C-map, are computed according to the differences in texture, luminance, and chrominance between the input and background frames, respectively. Then, a logical OR circuit operation is performed on the likelihood maps to produce the OR map. Finally, morphological operations are performed on the OR map to form the foreground mask. The vehicle is then extracted based on the shape of the foreground mask. This

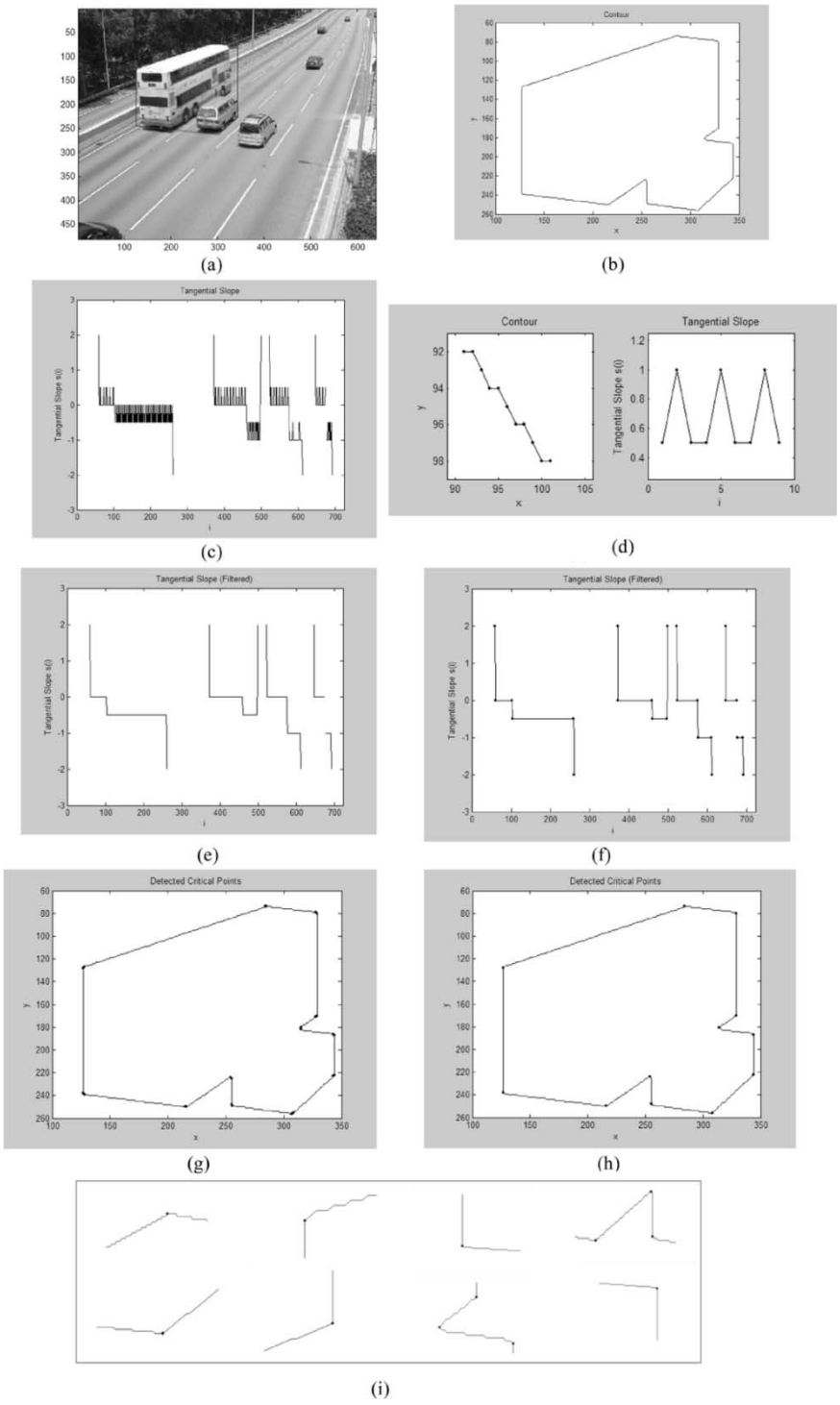


Fig. 10. Tangential slope curvature points extraction: (a) original image; (b) contour or shape of the segmented object; (c) signature curve of (b); (d) local variations on the signature curve; (e) filtered signature curve; (f) detected curvature points on curve; (g) curvature points in clusters; (h) final curvature points detected; and (i) closeup of the detected curvature points.

method proves to be very effective in extracting vehicle details without shadow and other visual artifacts (Fig. 6).

2) Camera Calibration: The camera model from [36] is depicted in Fig. 7. It helps to define the relationship between the 2-D image coordinates and the 3-D world coordinates in terms of the pan angle p, tilt angle t, swing angle s, focal length f, and camera distance l. By the process known as *camera calibration*, these parameters can be determined. The mapping between the 3-D and 2-D coordinates is defined in terms of the forward transformation Φ and the backward transformation Φ^{-1}, which is depicted as

$$\mathbf{g} = \Phi\{\mathbf{G}\} \tag{1}$$

$$\mathbf{G} = \Phi^{-1}\{\mathbf{g}\} \tag{2}$$

where $\mathbf{g}$ is a point in the 2-D image coordinates and $\mathbf{G}$ is the transformed point $\mathbf{g}$ in the 3-D world coordinates. A detailed

derivation of how these parameters are determined can be found in [36].

3) Deformable Modeling: From [22], the generalized deformable model is a wire frame with six vertices that can be parameterized to fit onto any vehicles (Fig. 8). The six vertices of the model can be represented as $\mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_6]^T$. This model is designed to fit onto the convex hull of the segmented vehicle. The steps of model fitting are summarized as follows (Fig. 9).

Step 1) Compute the center of gravity $\mathbf{g}$ of the convex hull of the segmented vehicle, which is defined as

$$\mathbf{g} = \frac{1}{N}\sum_{i=0}^{N-1} \mathbf{z}(i) \tag{3}$$

where $\mathbf{z}(i)$ are the sample points on the outline of the convex hull and N is the number of sample points on the outline [Fig. 9(a)].

Step 2) Draw the vector $\overrightarrow{\mathbf{gq}}$ along the road direction in two dimensions (θ_{road}). Vector $\overrightarrow{\mathbf{gq}}$ is defined as a vector with arbitrary length that points from the center of gravity $(\mathbf{g})$ to the road direction (θ_{road}) [Fig. 9(a)].

Step 3) Transform vector $\overrightarrow{\mathbf{gq}}$ from the 2-D to 3-D coordinates using (2). The transformed $\overrightarrow{\mathbf{gq}}$ is depicted as uppercase letters $\overrightarrow{\mathbf{GQ}}$ [Fig. 9(b)]. Rotate vector $\overrightarrow{\mathbf{GQ}}$ by 90° in the 3-D world coordinates. The rotated vector is depicted as $\overrightarrow{\mathbf{GQ}'}$. Transform $\overrightarrow{\mathbf{GQ}'}$ from 3-D coordinates back to 2-D coordinates by (1). The resultant vector is $\overrightarrow{\mathbf{gq}'}$, which is perpendicular to the road direction in the 2-D perspective view [Fig. 9(a)]. Traverse vector $\overrightarrow{\mathbf{gq}'}$ upward and downward to fit the upper and lower edges of the convex hull.

Step 4) Join points $\mathbf{g}$ and $\mathbf{p}$ to form line $\overline{\mathbf{gp}}$, where $\mathbf{p}$ is the vanishing point of the road. Search in the perpendicular direction of $\overline{\mathbf{gp}}$ (in the 2-D image coordinates) for points $\mathbf{q}_1$ and $\mathbf{q}_2$. Connect the vanishing point $\mathbf{p}$ with $\mathbf{q}_1$ and $\mathbf{q}_2$ [Fig. 9(c)].

Step 5) Draw vertical lines in 2-D to fit the two side edges.

Step 6) Find the intersections of the lines found in Steps 3), 4), and 5). The intersections define the six vertices of model $\mathbf{v}$ [Fig. 9(d)].

The above model has the advantage that it does not require any *a priori* knowledge of the vehicle shape for model fitting. Therefore, it is able to fit onto any type of vehicles, including complex-shaped vehicles such as cement trucks or even semitrailers. This advantage has greatly increased its applicability in AVTS.

C. Occlusion Detection Based on Estimated Dimensions

From the vertices of the fitted model, the dimensions of the vehicles can be estimated by transforming the model vertices $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_6]^T$ from the 2-D image coordinates to the 3-D world coordinates through (2) [22]. The model vertices after the transformation are denoted by the capital letter $\mathbf{V}$; that is, $\mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_6]^T$. The dimensions of the model are then estimated as (see Fig. 8)

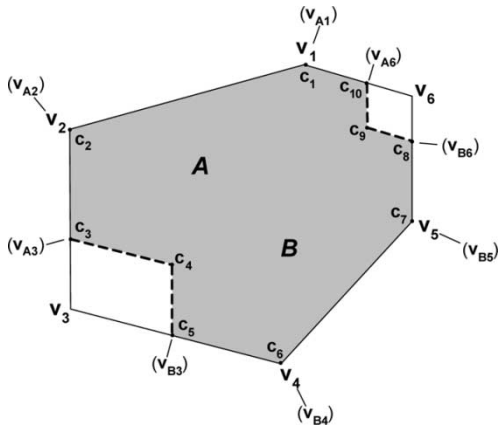


Fig. 11. Model fitted onto two occluded vehicles ($\boldsymbol{A}$ and $\boldsymbol{B}$).

$$W = |\mathbf{V}_3 - \mathbf{V}_4| \tag{4}$$

$$L = |\mathbf{V}_4 - \mathbf{V}_5| \tag{5}$$

$$H = |\mathbf{V}_2 - \mathbf{V}_3|. \tag{6}$$

When an occlusion occurs in the image sequence, the ratios between the dimensions of the fitted model are substantially different from the norm [6]. Besides, if occlusion starts or ends while the vehicles are within the field of view, the estimated dimensions change abruptly. Moreover, the area ratio (R_{area}) of the model is substantially lower than the norm, too, where

$$R_{\text{area}} = \frac{A_{\text{mask}}}{A_{\text{model}}} \tag{7}$$

and A_{mask} is the area of the segmented object and A_{model} is the area of the model fitted onto the segmented object. Occlusion is detected if either of the following criteria is met:

$$|R_{\text{area}}(i) - R_{\text{area}}(i-1)| > \varepsilon_1$$

or

$$|W(i) - W(i-1)| > \varepsilon_2$$

or

$$|L(i) - L(i-1)| > \varepsilon_2$$

or

$$|H(i) - H(i-1)| > \varepsilon_2 \tag{8}$$

where i is the current frame and $i-1$ is the previous frame. The values ε_1 and ε_2 are thresholds to permit small changes in the values. When there is no occlusion, R_{area} should be close to 1 and the differences of W, L, and H between frames should be relatively constant.

D. Shape Feature Extraction

The purpose of this approach is to extract feature points that can be used to fit individual models to the composite segmented object later on. The feature points chosen are those that represent a change in direction along the extracted shape. Some of these

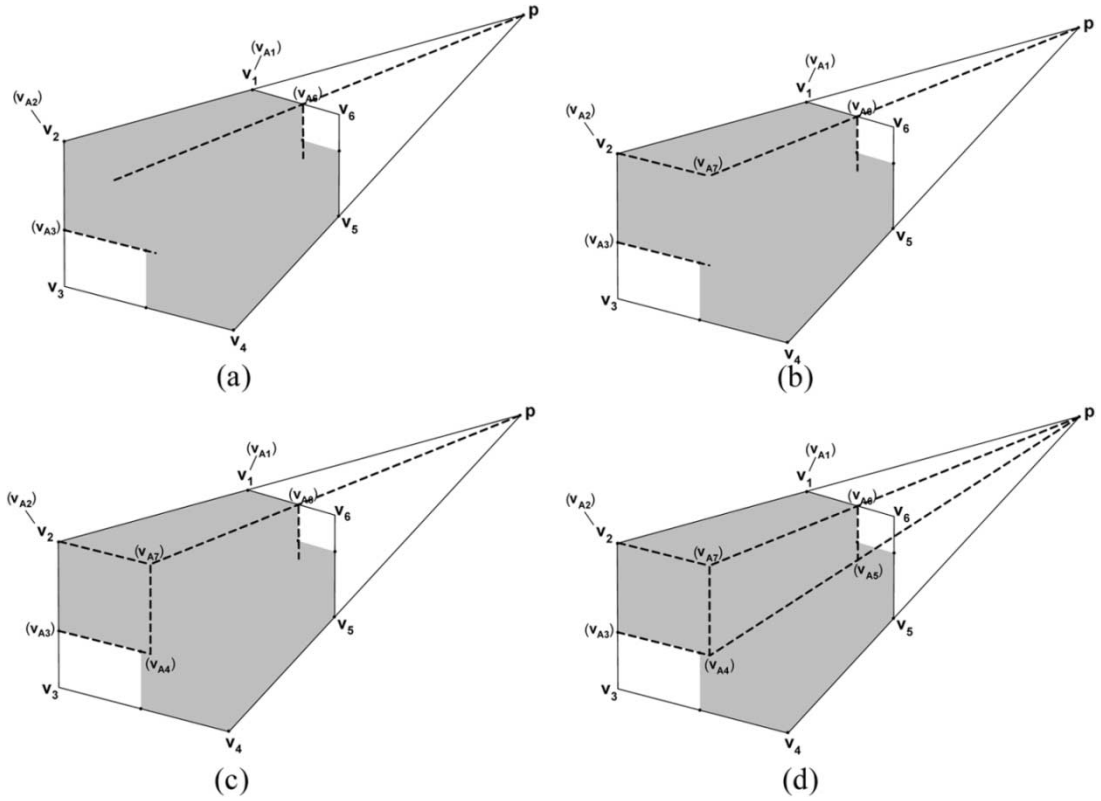


Fig. 12. Model partition of vehicle $\boldsymbol{A}$: (a) Step 2); (b) Step 3); (c) Step 4); and (d) Step 5).

points are clear corners with sharp curvature. On the other hand, some of these points have a more gradual curvature and are less definite as a corner point. In order to detect both types of points, we propose the following approach.

1) Computation of Signature Curve: In the case of occlusion, the segmented object would consist of more than one vehicle. The contour or shape that describes the segmented object is a composite of the occluded vehicles [Fig. 10(a) and (b)]. To prepare for the subsequent fitting of the individual vehicle model, we first compute the tangential slope of every data sample on the contour of the segmented object [Fig. 10(c)]. For simplicity, the tangential slope of a data sample i is defined as the vector that joins samples $i-1$ and $i+1$. The signature curve of the contour, which is a set of tangential slopes, can be written mathematically as

$$s_i(x_i, y_i) = \frac{y_{i+1} - y_{i-1}}{x_{i+1} - x_{i-1}} \tag{9}$$

for $i = 2, \ldots, N-1$, where N is the number of sample points on the contour and x_i and y_i are the x and y components of the sample point i, respectively. The merit of this method is that it is straightforward and retains points of sharp and gradual curvatures. However, this also implies that the signature curve also contains a considerable amount of local variations or noise [Fig. 10(d)] that required suppression in the next stage.

2) Signature Curve Filtering: These local variations or noise are due to quantization in the digitizing process of the contour and must be filtered before the curvature points can be detected from the signature curve. From a detailed study of the signature curve, it is found that most of these local variations have a duration of not more than two samples in length. As such, the filter can simply eliminate any spikes that satisfy this condition. In other words, the filter changes $[a, b, a]$ or $[a, b, b, a]$ to $[a, a, a]$ or $[a, a, a, a]$, respectively, where a and b are two different tangential slope values. The resulting filtered signature curve is depicted in Fig. 10(e). Comparing this with Fig. 10(c), all the local variations have been removed.

3) Curvature Detection: First, start and end points of each disjointed segment on the filtered signature curve are considered curvature points, as they represent a substantial change in tangential slopes. Second, curvature points are also detected at the transition point where the filtered signature curve changes state.

The detected curvature points are shown in Fig. 10(f) and (g). We can observe from Fig. 10(g) that some of the detected points are clustered. To alleviate this problem, we simply select one point with the highest curvature out of a cluster of detected points. The set of final curvature points is depicted in Fig. 10(h). Fig. 10(i) depicts the closeup of the detected curvature points and we can observe that the proposed algorithm has accurately detected the curvature points. The set of detected curvature points can be written as $\mathbf{c} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K]^T$, where K is the number of detected curvature points.

E. Model Decomposition

Given the detected curvature points $\mathbf{c} = [\mathbf{c_1}, \mathbf{c_2}, \ldots, \mathbf{c_K}]^T$ and the vanishing point $\mathbf{p}$, assume that model $\mathbf{v} = [\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_6}]^T$ is fitted geometrically onto the segmented object of two occluded vehicles, namely, vehicle $\boldsymbol{A}$ and $\boldsymbol{B}$ (Fig. 11). It is observed that, in an ideal case, the expected number of K is 10 when there are two vehicles occluded as seen in the image. Based on this set of curvature points, the

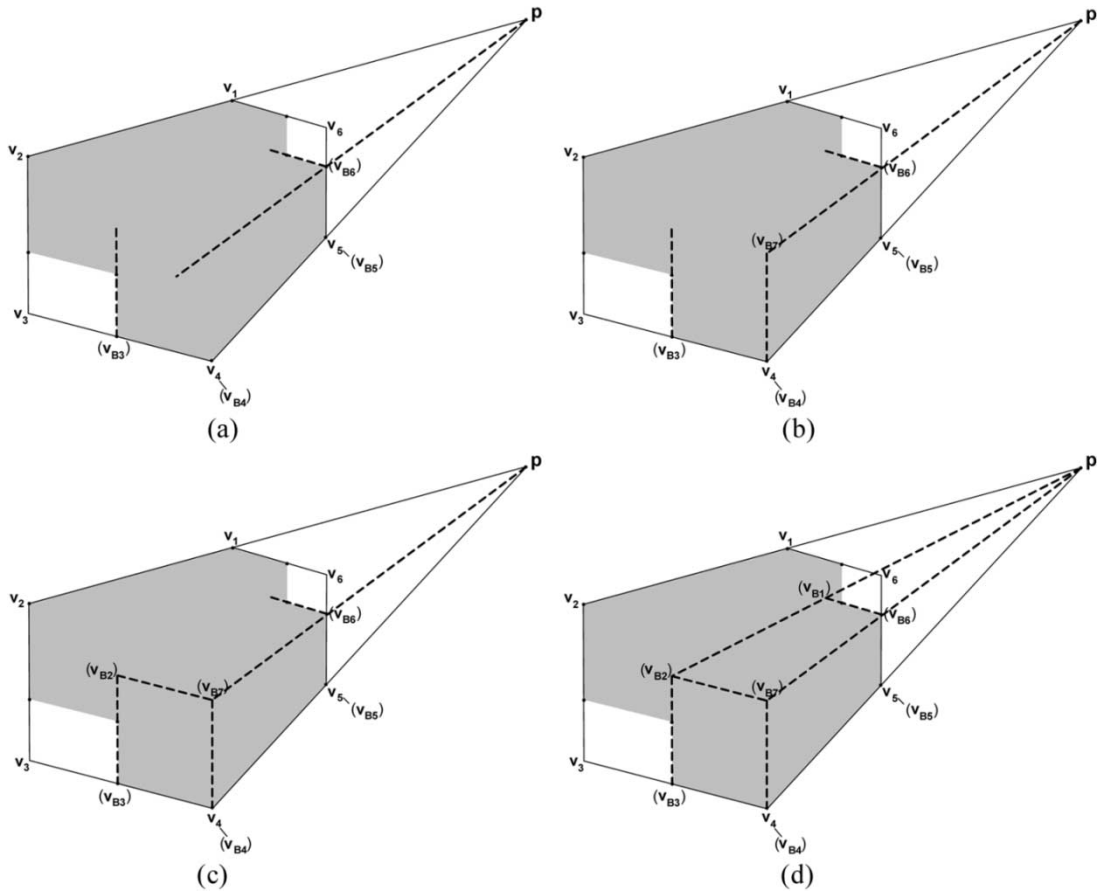


Fig. 13. Model partition of vehicle $\boldsymbol{B}$: (a) Step 2); (b) Step 3); (c) Step 4); and (d) Step 5).

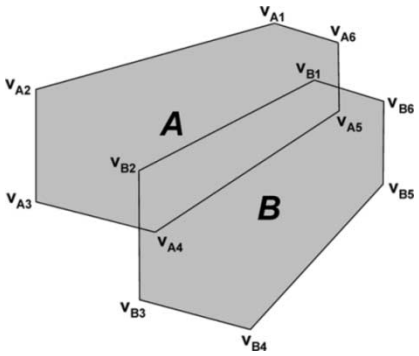


Fig. 14. Partitioned models $\mathbf{v_A}$ and $\mathbf{v_B}$.

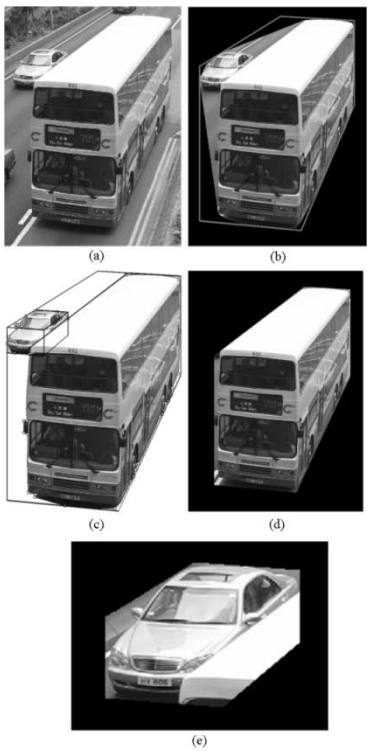


Fig. 15. Sedan occluded by a double-decker bus: (a) image; (b) fitted model; (c) model partitioning; (d) segmented bus; and (e) segmented sedan.

the occluded model $\mathbf{v}$ can be decomposed into two separate models, namely, $\mathbf{v_A} = [\mathbf{v_{A1}}, \mathbf{v_{A2}}, \ldots, \mathbf{v_{A6}}]^T$ and $\mathbf{v_B} = [\mathbf{v_{B1}}, \mathbf{v_{B2}}, \ldots, \mathbf{v_{B6}}]^T$, so that model $\mathbf{v_A}$ and $\mathbf{v_B}$ represent vehicle $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively.

From Fig. 11, it can be seen that point $\mathbf{v_{A3}}$ can be determined by traversing line $\overline{\mathbf{v_3 v_4}}$ upward until it aligns with curvature points $\mathbf{c_3}$ and $\mathbf{c_4}$ and point $\mathbf{v_{A6}}$ can be determined by traversing line $\overline{\mathbf{v_5 v_6}}$ leftward until it aligns with curvature points $\mathbf{c_9}$ and $\mathbf{c_{10}}$. Similarly, points $\mathbf{v_{B3}}$ can be determined by traversing line $\overline{\mathbf{v_2 v_3}}$ rightward until it aligns with $\mathbf{c_4}$ and $\mathbf{c_5}$ and $\mathbf{v_{B6}}$ can be determined by traversing line $\overline{\mathbf{v_1 v_6}}$ downward until it aligns with $\mathbf{c_8}$ and $\mathbf{c_9}$.

Now, points $\{\mathbf{v_{A1}}, \mathbf{v_{A2}}, \mathbf{v_{A3}}, \mathbf{v_{A6}}\}$ of model $\mathbf{v_A}$ and points $\{\mathbf{v_{B3}}, \mathbf{v_{B4}}, \mathbf{v_{B5}}, \mathbf{v_{B6}}\}$ of model $\mathbf{v_B}$ are already known. Only points $\{\mathbf{v_{A4}}, \mathbf{v_{A5}}\}$ of model $\boldsymbol{A}$ and $\{\mathbf{v_{B1}}, \mathbf{v_{B2}}\}$ of model $\boldsymbol{B}$

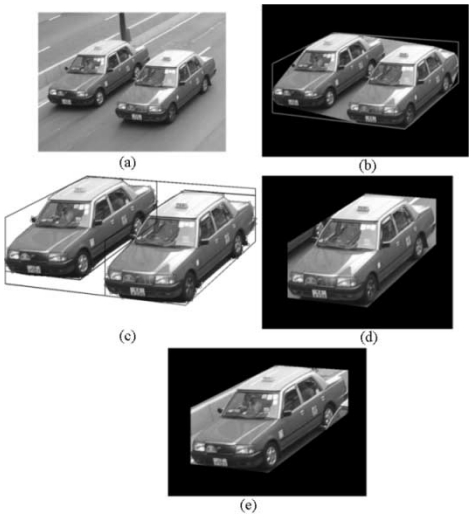


Fig. 16. Taxi occluded by another taxi: (a) image; (b) fitted model; (c) model partitioning; (d) segmented taxi (right); and (e) segmented minivan (left).

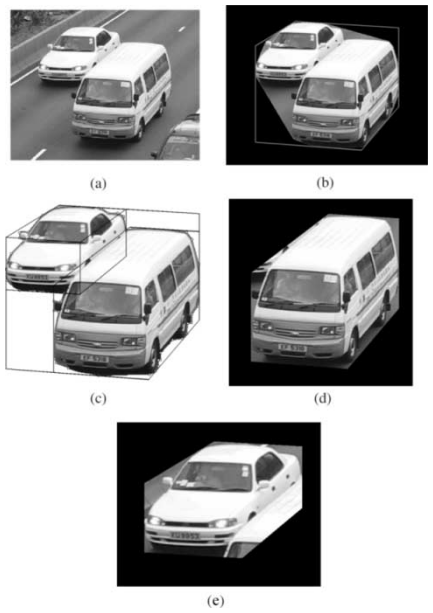


Fig. 17. Sedan occluded by a minivan: (a) image; (b) fitted model; (c) model partitioning; (d) segmented minivan; and (e) segmented sedan.

are missing. These points can be determined by exploiting the information given by the curvature points, as well as the vanishing point $\mathbf{p}$. The steps are depicted as follows.

Vehicle A (Fig. 12):

Step 1) Extend line $\overline{\mathbf{v_2v_1}}$ and line $\overline{\mathbf{v_4v_5}}$ until they meet at point $\mathbf{p}$.

Step 2) Connect points $\mathbf{p}$ and $\mathbf{v_{A6}}$ and extend the line [Fig. 12(a)].

Step 3) Traverse line $\overline{\mathbf{v_{A1}v_{A6}}}$ along plane $\overline{\mathbf{pv_{A1}v_{A6}}}$ until it intersects with point $\mathbf{v_{A2}}$. The new corner on the extended plane is now labeled as $\mathbf{v_{A7}}$ [Fig. 12(b)].

Step 4) Traverse line $\overline{\mathbf{v_{A2}v_{A3}}}$ along the plane $\overline{\mathbf{v_{A2}v_{A3}v_{A7}}}$ until it intersects with point $\mathbf{v_{A7}}$. The intersection between the new line and the horizontal line passing through $\mathbf{v_{A3}}$ is now labeled as $\mathbf{v_{A4}}$ [Fig. 12(c)].

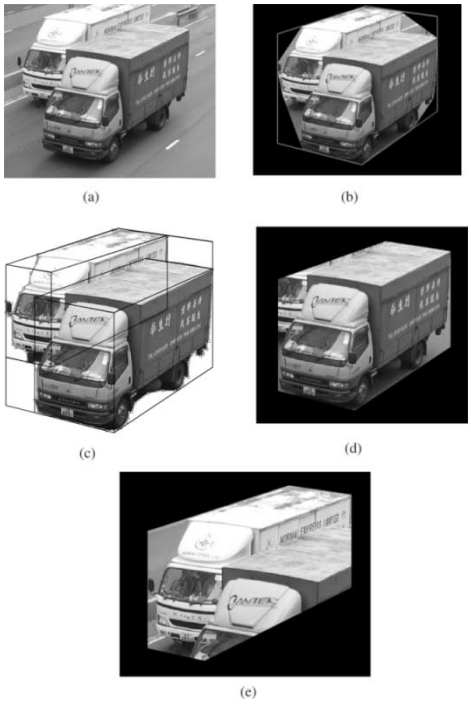


Fig. 18. Truck occluded by another truck: (a) image; (b) fitted model; (c) model partitioning; (d) segmented truck (left); and (e) segmented truck (right).

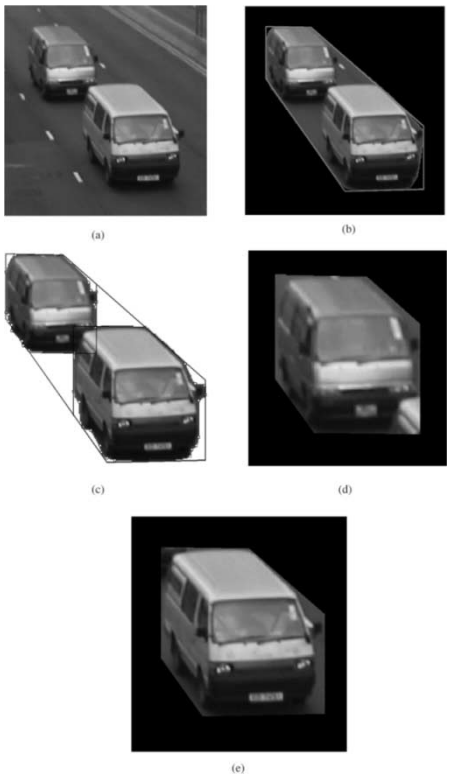


Fig. 19. Minivan occluded by another minivan: (a) image; (b) fitted model; (c) model partitioning; (d) segmented minivan (back); and (e) segmented minivan (front).

Step 5) Connect point $\mathbf{p}$ and $\mathbf{v_{A4}}$. An intersection point $\mathbf{v_{A5}}$ is formed [Fig. 12(d)].

Vehicle B (Fig. 13):

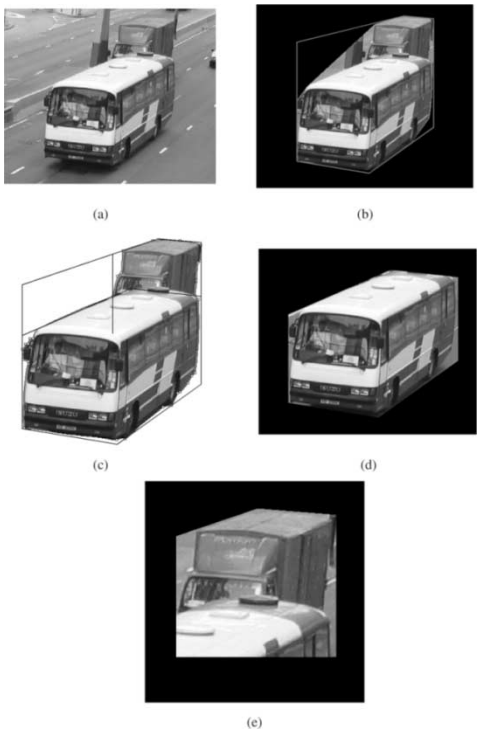


Fig. 20. Truck occluded by a tour bus: (a) image; (b) fitted model; (c) model partitioning; (d) segmented tour bus; (e) segmented truck.

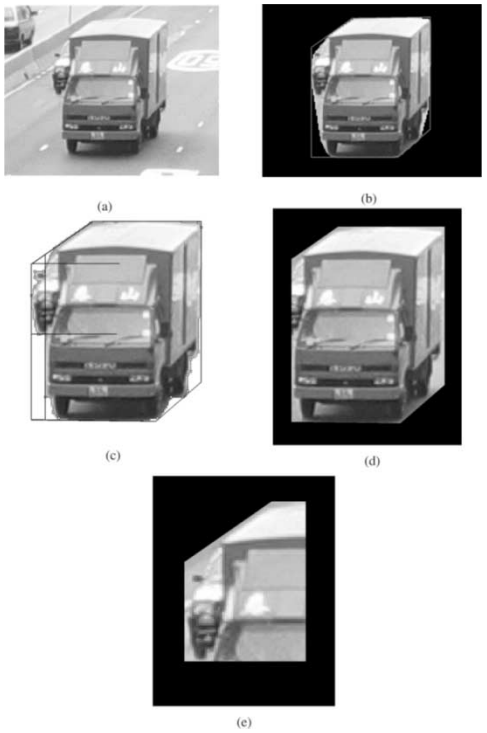


Fig. 21. Taxi occluded by a truck: (a) image; (b) fitted model; (c) model partitioning; (d) segmented truck; and (e) segmented taxi.

Step 1) Extend lines $\overline{\mathbf{v_2v_1}}$ and $\overline{\mathbf{v_4v_5}}$ until they intersect at point $\mathbf{p}$.

Step 2) Connect points $\mathbf{p}$ and $\mathbf{v_{B6}}$ and extend the line [Fig. 13(a)].

Step 3) Traverse line $\overline{\mathbf{v_{B5}v_{B6}}}$ along plane $\overline{\mathbf{pv_{B5}v_{B6}}}$ until it intersects with point $\mathbf{v_{B4}}$. The new corner on the extended plane is now labeled as $\mathbf{v_{B7}}$ [Fig. 13(b)].

Step 4) Traverse line $\overline{\mathbf{v_{B3}v_{B4}}}$ upward until it intersects with point $\mathbf{v_{B7}}$. The intersection between the new line and the vertical line passing through $\mathbf{v_{B3}}$ is now labeled as $\mathbf{v_{B2}}$ [Fig. 13(c)].

Step 5) Connect point $\mathbf{p}$ and $\mathbf{v_{B2}}$. An intersection point $\mathbf{v_{B1}}$ is formed [Fig. 13(d)].

From the above steps, the model can be partitioned into two separate but complete models $\mathbf{v_A}$ and $\mathbf{v_B}$, representing vehicles $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively (Fig. 14). Dimensions of the individual vehicle can be estimated by the same procedure as introduced in Section IV-C.

V. Results and Discussions

The proposed method is evaluated by 50 sets of monocular traffic-image sequences that are taken on a busy highway where occlusions occur frequently. All the image sequences are acquired during daytime, when shadows are prominent. These results are subjectively analyzed for each case to see if the proposed method performed in a satisfactory manner (Figs. 15–21). To objectively evaluate the accuracy of the proposed method, the dimensions of the vehicles segmented from the test image sequences were calculated from the partitioned model. The estimated values are then compared with the actual dimensions published by the manufacturers and the percentage accuracies of the estimated dimensions are plotted in Fig. 22. The first image sequence contains a sedan that is partially occluded by a double-decker bus, as depicted in Fig. 15(a). The convex hull of the two vehicles as well as the fitted model is depicted in Fig. 15(b). The segmented vehicles as well as the decomposed models are shown in Fig. 15(c). The resolved vehicles are shown in Fig. 15(d) and (e). In Fig. 15(d), the decomposed model fits the double-decker bus very well, except that the width is slightly wider due to the effect of the side mirror, which causes the left edge of the decomposed model to move to the left. It can be observed from Fig. 15(e) that the decomposed model fits the sedan well and that the side mirror in this case has no effect to the modeling.

The second image sequence contains a taxi that is partially occluded by another taxi, as depicted in Fig. 16(a). The decomposed models are shown in Fig. 16(d) and (e). It can be observed that both decomposed models fit the taxis very well.

The third image sequence contains a sedan that is partially occluded by a minivan, as depicted in Fig. 17(a). The decomposed model fits the width and height of the minivan accurately. However, the fitting of the top edge of the minivan is slightly deviated from the vehicle [Fig. 17(d)]. This error is due to the fact that the traveling direction of the vehicle is slightly misaligned from the road direction. Therefore, when the vehicle is aligned with the vanishing point, such deviation occurred. On the other hand, in Fig. 17(e), the decomposed model fits the sedan well.

The fourth image sequence contains a truck that is partially occluded by another truck [Fig. 18(a)]. Similar to the minivan in the previous case, the fitting of the top edge is slightly deviated from the truck roof for the same reason [Fig. 18(d)]. For the left truck, the decomposed model is slightly larger due to the side mirror.

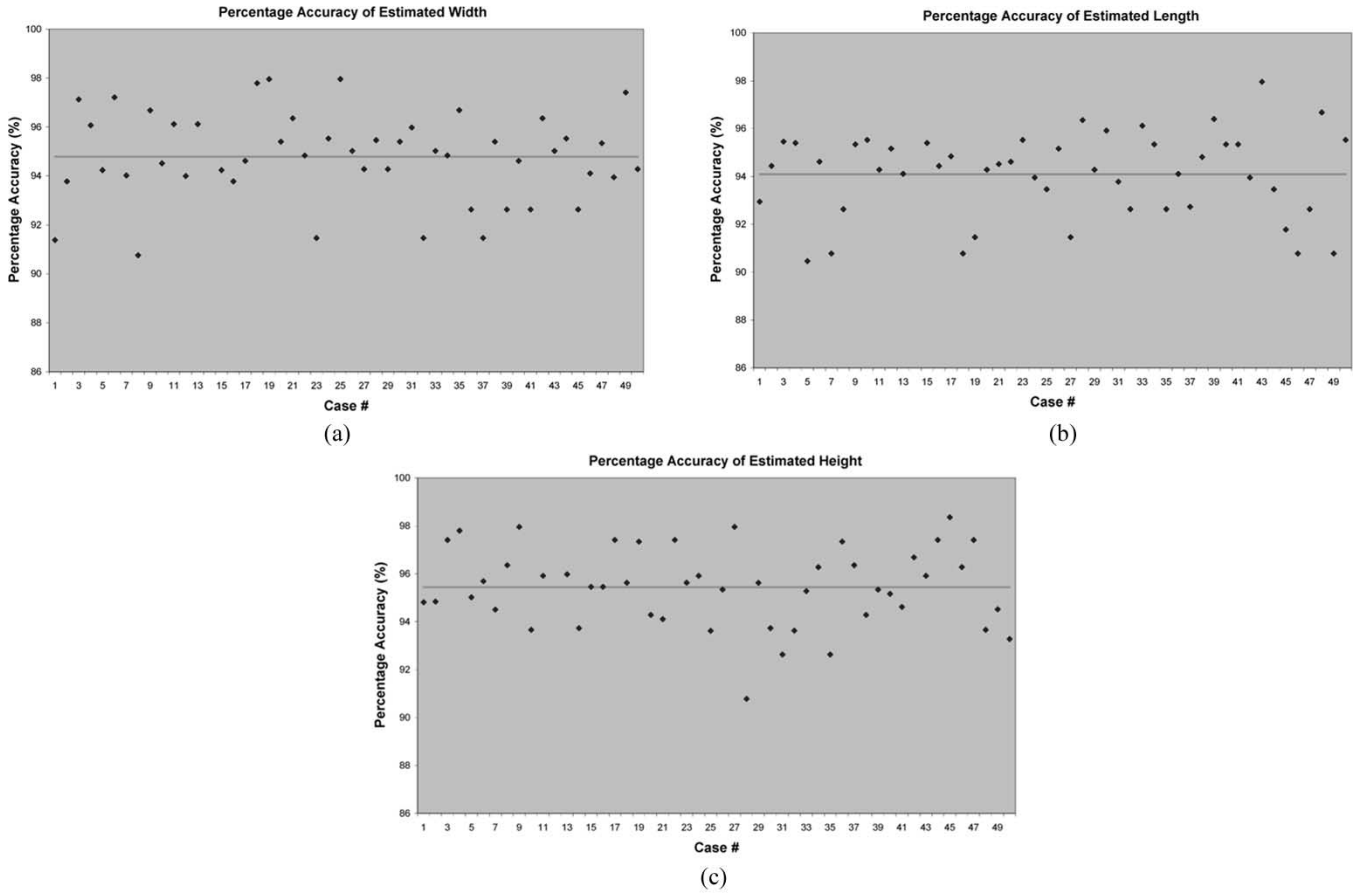


Fig. 22. Percentage accuracy of the dimension estimation of the individual vehicles: (a) percentage accuracy of the width estimation; (b) percentage accuracy of the length estimation; and (c) percentage accuracy of the height estimation.

The fifth image sequence contains a minivan that is partially occluded by another minivan on the same lane [Fig. 19(a)]. The decomposed model fits the van at the back reasonably well with the top right side edge slightly too far to the right [Fig. 19(d)]. For the front van, the model is slightly larger due to the side mirror.

The sixth image sequence contains a truck that is partially occluded by a tour bus [Fig. 20(a)]. The truck is so severely occluded by the tour bus that the proposed method fails to determine its height. As a result, the decomposed model for the truck [Fig. 20(e)] is far too large, whereas the decomposed model of the tour bus is satisfactory [Fig. 20(d)]. As a matter of fact, it would be equally difficult for human beings to estimate the height of the truck in this case too.

The seventh image sequence contains a taxi that is being severely occluded by a truck, as depicted in Fig. 21(a). In this case, the proposed method fails to find the width and length of the model of the taxi and, as a result, calculated an erratic model, as depicted in Fig. 21(e). Similar to the sixth case, it would be challenging for human beings to estimate the height and length without applying our knowledge of what vehicle it is (as in recognition).

The accuracy of the proposed algorithm is being evaluated qualitatively by determining the dimensions of the vehicles in the test-image sequences from the decomposed model. The estimated values are then compared with the actual dimensions published by the manufacturers and the percentage accuracies of the estimated dimensions are plotted in Fig. 22. The average accuracy for width estimation is 94.78% and the average accuracy for length and height estimation is 94.09% and 95.44%, respectively. In all three cases, the worst case accuracy is around 91% while the best case accuracy is around 98%. The errors for width, length, and height estimation are 5.22%, 5.91%, and 4.56% respectively. These figures again show the potential of the proposed method.

TABLE I
TIME REQUIREMENT OF THE PROPOSED METHODOLOGY

	Time required
Signature curve computation	0.31 sec
Signature curve filtering	0.55 sec
Model decomposition	0.12 sec

The computational speed values for the three major steps of the proposed methodology are depicted in Table I. The proposed algorithms were implemented in MATLAB on a Pentium III 800-MHz platform. We could see that the signature curve-filtering step required the longest time to execute, while the model decomposition step required the shortest execution time. The total time delay for performing all three tasks is less than 1 s. It should be noted that MATLAB implementation general is not very efficient and is much slower than dedicated C++ or C# implementation. As such, the proposed algorithms are computationally feasible.

VI. Conclusion

This paper proposed a novel method for detecting and resolving occlusion between vehicles, as seen in a monocular image sequence. By adopting the texture-based segmentation and deformable modeling method, it performed the curvature detection reasonably accurate, which then enables the decomposition of the composite model into individual vehicle models using a geometric approach. The proposed method has been evaluated on a number of real-world traffic image sequences and the average estimation accuracy in estimating vehicle width is 94.78%, length is 94.09%, and height is 95.44%. It should be noted that the errors are due to: 1) the side mirror (left or right) that protrudes out of the vehicle body and, being part of the vehicle, it is detected as part of the shape of the vehicle and 2) the vehicle is not completely aligned with the road direction, which causes slight deviations in aligning lines through the vanishing point. Both drawbacks do not introduce large errors to the method. Further, the proposed method in its current form is not able to handle severely occluded cases, as shown in the sixth and seventh cases. This is perhaps not a deficiency of the proposed method, but is a problem that inherently require more complex approaches, such as vehicle recognition, to deal with it. Although we have not considered the more complex scenarios, such as occlusion, that involve more than two vehicles and a long queue of vehicles in congesting situations, the proposed method does indeed present a way forward for further development in this direction.

References

[1] A. H. S. Lai, "An effective methodology for visual traffic surveillance," Ph.D. dissertation, Univ. Hong Kong, China, 2000.

[2] J. Lou, T. Tan, and W. Hu, "Visual vehicle tracking algorithm," *Electron. Lett.*, vol. 38, no. 18, pp. 1024–1025, 2002.

[3] J. Lou, Q. Liu, T. Tan, and W. Hu, "Semantic interpretation of object activities in a surveillance system," in *Proc. IEEE 16th Int. Conf. Pattern Recognition*, vol. 3, 2002, pp. 777–780.

[4] Z. Kim and J. Malik, "Fast vehicle detection with probabilistic feature grouping and its application to vehicle tracking," in *Proc. IEEE 9th Int. Conf. Computer Vision*, 2003, pp. 524–531.

[5] D. Beymer and K. Malik, "Tracking vehicles in congested traffic," in *Proc. IEEE Intelligent Vehicles Symp.*, 1996, pp. 130–135.

[6] N. H. C. Yung and A. H. S. Lai, "Detection of vehicle occlusion using a generalized deformable model," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 4, 1998, pp. 154–157.

[7] T. Rabie, A. Shalaby, B. Abdulhai, and A. El-Rabbany, "Mobile vision-based vehicle tracking and traffic control," in *Proc. IEEE 5th Int. Conf. Intelligent Transportation Systems*, 2002, pp. 13–18.

[8] M. Bertozzi, A. Broggi, A. Fascioli, and S. Nichele, "Stereo vision-based vehicle detection," in *Proc. IEEE Intelligent Vehicles Symp.*, 2000, pp. 39–44.

[9] T. Hasegawa, K. Imamura, H. Zen, and A. El-Rabbany, "Observation of moving vehicles by the plural cameras established freely," in *Proc. IEEE Intelligent Transportation Systems Conf.*, 2000, pp. 328–333.

[10] B. Steux, C. Laurgeau, L. Salesse, and D. Wautier, "Fade: A vehicle detection and tracking system featuring monocular color vision and radar data fusion," in *Proc. IEEE Intelligent Vehicle Symp.*, vol. 2, 2002, pp. 632–639.

[11] R. P. Ramachandran, G. Arr, C. Sun, and S. G. Ritchie, "A pattern recognition and feature fusion formulation for vehicle reidentification in intelligent transportation systems," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 4, 2002, pp. 3840–3843.

[12] R. Cucchiara, M. Piccardi, and P. Mello, "Image analysis and rule-based reasoning for a traffic monitoring system," *IEEE Trans. Intell. Transport. Syst.*, vol. 1, pp. 119–130, June 2000.

[13] S. Gupte, O. Masoud, R. F. K. Martin, and N. P. Papanikolopoulos, "Detection and classification of vehicles," *IEEE Trans. Intell. Transport. Syst.*, vol. 3, pp. 37–47, Mar. 2002.

[14] T. Ikeda, S. Ohnaka, and M. Mizoguchi, "Traffic measurement with a roadside vision system-individual tracking of overlapped vehicles," in *Proc. IEEE 13th Int. Conf. Pattern Recognition*, vol. 3, 1996, pp. 859–864.

[15] G. A. Jones and P. R. Giaccone, "Hierarchical tracking of motion in multiple images," in *Proc. IEE Colloq. Multiresolution Modeling Analysis Image Processing and Computer Vision*, vol. 7, 1995, pp. 1–6.

[16] Y. K. Jung, K. W. Lee, and Y. S. Ho, "Content-based event retrieval using semantic scene interpretation for automated traffic surveillance," *IEEE Trans. Intell. Transport. Syst.*, vol. 2, pp. 151–163, Sept. 2001.

[17] Y. K. Jung and Y. S. Ho, "Traffic parameter extraction using video-based vehicle tracking," in *Proc. IEEE Int. Conf. Intell. Transport. Syst.*, 1999, pp. 764–769.

[18] T. Kato, Y. Ninomiya, and I. Masaki, "Preceding vehicle recognition based on learning from sample images," *IEEE Trans. Intell. Transport. Syst.*, vol. 3, pp. 252–260, Dec. 2002.

[19] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi, "Traffic monitoring and accident detection at intersections," *IEEE Trans. Intell. Transport. Syst.*, vol. 1, pp. 108–118, June 2000.

[20] ——, "Occlusion robust tracking utilizing spatio-temporal markov random field model," in *Proc. IEEE 15th Int. Conf. Pattern Recognition*, vol. 1, 2000, pp. 140–144.

[21] J. B. Kim, H. S. Park, M. H. Park, and H. J. Kim, "Unsupervised moving object segmentation and recognition using clustering and a neural network," in *Proc. IEEE Int. Joint Conf. Neural Networks*, vol. 2, 2002, pp. 1240–1245.

[22] A. H. S. Lai, G. S. K. Fung, and N. H. C. Yung, "Vehicle type classification from visual-based dimension estimation," in *Proc. IEEE Intelligent Transportation Syststems Conf.*, 2001, pp. 201–206.

[23] I. Lee, H. Ko, and D. K. Han, "Multiple vehicle tracking based on regional estimation in nighttime CCD images," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 4, 2002, pp. 3712–3715.

[24] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, "Moving target classification and tracking from real-time video," in *Proc. IEEE 4th Workshop Applications Computer Vision*, 1998, pp. 8–14.

[25] O. Masoud, P. P. Nikilaos, and E. Kwon, "The use of computer vision in monitoring weaving sections," *IEEE Trans. Intell. Transport. Syst.*, vol. 2, pp. 18–25, Mar. 2001.

[26] T. Nakanishi and K. Ishii, "Automatic vehicle image extraction based on spatio-temporal image analysis," in *Proc. IEEE 11th IAPR Int. Conf. Pattern Recognition*, vol. 1, 1992, pp. 500–504.

[27] F. Oberti, S. Calcagno, M. Zara, and C. S. Regazzoni, "Robust tracking of humans and vehicles in cluttered scenes with occlusions," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, 2002, pp. 629–632.

[28] H. Sakaino, "Nonlinear robust velocity estimation of vehicles from a snowfall traffic scene," in *Proc. IEEE 16th Int. Conf. Pattern Recognition*, vol. 4, 2002, pp. 60–63.

[29] M. Schmid, "An approach to model-based 3-D recognition of vehicles in real time by machine vision," in *Proc. IEEE Int. Conf. Intelligent Robots Systems*, vol. 3, 1994, pp. 2064–2071.

[30] D. Sinclair, B. Boufama, and R. Mohr, "Independent motion segmentation and collision prediction for road vehicles," in *Proc. IEEE Computer Society Conf. Computer Vision Pattern Recognition*, 1994, pp. 958–961.

[31] Z. Tang and U. Ozguner, "An adaptive motion correspondence algorithm for multiple vehicle tracking from an air-borne platform," in *Proc. IEEE 5th Int. Conf. Intelligent Transportation Systems*, 2002, pp. 49–54.

[32] M. K. Teal and T. J. Ellis, "Occlusion prediction based on target dynamics and spatial reasoning," in *Proc. IEEE 6th Int. Conf. Image Processing Its Applications*, vol. 1, 1997, pp. 219–223.

[33] Z. Tang and U. Ozguner, "An adaptive motion correspondence algorithm for multiple vehicle tracking from an air-borne platform," in *Proc. IEEE 5th Int. Conf. Intelligent Transportation Systems*, 2002, pp. 49–54.

[34] Y. Won, P. D. Gader, and P. C. Coffield, "Morphological shared-weight networks with applications to automatic target recognition," *IEEE Trans. Neural Networks*, vol. 8, pp. 1195–1203, Nov. 1997.

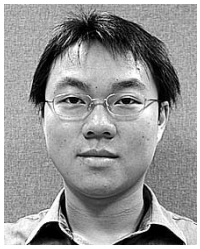
[35] H. Yang, J. Lou, H. Sun, W. Hu, and T. Tan, "Efficient and robust vehicle localization," in *Proc. IEEE Int. Conf. Image Processing*, vol. 2, 2001, pp. 355–358.

[36] G. S. K. Fung, N. H. C. Yung, and G. K. H. Pang, "Camera calibration from road lane markings," *Opt. Eng.— J. SPIE*, vol. 42, no. 10, pp. 2967–2977, 2003.

[37] A. H. S. Lai and N. H. C. Yung, "A fast and accurate scoreboard algorithm for estimating stationary backgrounds in an image sequence," in *Proc. IEEE Int. Symp. Circuits Systems*, vol. 4, 1998, pp. 241–244.

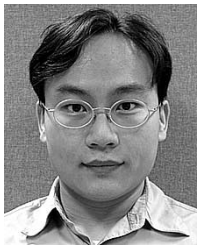
[38] W. W. L. Lam, C. C. C. Pang, and N. H. C. Yung, "A highly accurate texture-based vehicle segmentation method," *Opt. Eng.— J. SPIE*, vol. 43, no. 3, pp. 591–603, 2004.

[39] C. C. C. Pang, W. W. L. Lam, and N. H. C. Yung, "A novel method for handling vehicle occlusion in visual traffic surveillance," in *Proc. SPIE Conf. Electronic Imaging*, vol. 5014, 2003, pp. 437–447.



Clement Chun Cheong Pang received the B.Eng. and M.Eng. degrees in electrical engineering from McMaster University, Hamilton, ON, Canada, in 1999 and 2001, respectively. He is currently working toward the Ph.D. degree in the Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong, China.

His research interests include automated visual traffic surveillance and ECG signal analysis, as well as the detection of brain abnormalities from EEG signals.



William Wai Leung Lam received the B.Eng. degree in electrical engineering from McMaster University, Hamilton, ON, Canada, and the M.Phil. degree in electrical and electronic engineering from the Hong Kong University of Science and Technology, Hong Kong, China. He is currently working toward the Ph.D. degree in the Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong, China.

His research interests include digital image processing, texture synthesis, intelligent transportation systems, and computer networks.

Nelson Hon Ching Yung received the B.Sc. and Ph.D. degrees from the University of Newcastle-Upon-Tyne, Newcastle-Upon-Tyne, U.K.

He was Lecturer with the University of Newcastle-Upon-Tyne from 1985 to 1990. He was a Senior Research Scientist in the Department of Defence, Canberra, Australia from 1990 to 1993. He joined the University of Hong Kong (HKU), Hong Kong, China, as Associate Professor in 1993. He currently leads a research team in digital image processing and intelligent transportation systems. He is the Founding Director of the Laboratory for Intelligent Transportation Systems Research, HKU, and Deputy Director of HKU's Institute of Transport Studies. He also is Chairman of the Technology Advisory Board, mCommerce Online, Ltd., a university spinoff company. He has coauthored a computer vision book and has published over 100 journal and conference papers in the areas of digital image processing, parallel algorithms, automated visual traffic surveillance, and autonomous vehicle navigation and learning algorithms.

Dr. Yung is a Reviewer for the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, and IEEE TRANSACTIONS ON SIGNAL PROCESSING, as well as *IEE Part G, SPIE Optical Engineering, International Journal of Vehicle Mechanics and Mobility, HKIE Proceedings, Microprocessors and Microsystems,* and *Robotics and Autonomous Systems Journal*. He was a Member of the Advisory Panel of the ITS Strategy Review, Transport Department, HKSAR, a Council Member of ITS-HK, and Chair of the Computer Division, International Institute for Critical Infrastructures. He is a Chartered Electrical Engineer and a Member of the HKIE and IEE. He is listed in *Who's Who in the World* (Marquis, OH) since 1998. He was a Croucher Scholar and his team won the Silver Award from the Hong Kong Electronic Industry Association for Outstanding Innovation and Technology Product, 2000 for the Mobile and online vending enabler (MOVER) solution.