

DETECTION OF VEHICLE OCCLUSION USING A GENERALIZED DEFORMABLE MODEL

Nelson H. C. Yung and Andrew H. S. Lai

Department of Electrical & Electronic Engineering

The University of Hong Kong

Chow Yei Ching Building, Pokfulam Road, Hong Kong

Tel: (852) 2859-2685 Fax: (852) 2559-8739 Email: nyung@eee.hku.hk

ABSTRACT

This paper presents a vehicle occlusion detection algorithm based on a generalized deformable model. A 3D solid cuboid model with up to six vertices is employed to fit any vehicle images, by varying the vertices for a best fit. The advantage of using such a model is that the number of parameterized vertices is small which can be easily deformed. Occlusion is detected by recording the changes in the *Area Ratio* and the dimensions of the generalized deformable model. Our tests show that the new modeling algorithm is effective in detecting vehicle occlusion

1. INTRODUCTION

Visual traffic surveillance[1] (VTS) is an active research area as it is believed by many that it is going to have significant impacts to most of the traffic problems in big cities today. A common goal for many of these VTS systems is to be able to identify and track the movement of vehicles. In this, deformable models[2-8] are often used to model the vehicles, where vehicle types and orientation information may be extracted[2-6]. However, the complexity of deformable models is high and the success of matching the vehicle image with the model can be hard to come by if the vehicle image is incomplete or occluded by other vehicles or stationary objects. One possible solution to this particular problem is to use a simple generalized model[7] which may be more effective and robust. Evolved from the same line of argument, a simple generalized deformable model is proposed in this paper. It employs a 3D solid cuboid model with up to six vertices to fit any vehicle images, by varying the vertices for a best fit. The advantage of using such a model is that the number of parameterized vertices is small which can be easily calculated. This model was tried on a series of occlusion cases and was found to be able to discern the occluded vehicles and those that are not.

2. METHODOLOGY

The whole idea lies on the fitting of the model on the binary representation of the vehicle in 2D. Once a best fit is determined,

the resulting model is transformed into 3D, where parameters such as height, width and length of the vehicle are calculated. By considering the area ratio between the model and the binary representation of the vehicle, as well as it's width and height, vehicles that are occluded can be detected based on these parameters. Fig. 1 depicts the conceptual flow of the method employed. In this method, the binary representation of the vehicle is extracted using a background estimation algorithm[9]. The fitting of the generalized model is performed in 2D, as the vehicle image is also in 2D. This is done by transforming the 3D model into 2D, and calculate the lines between the vertices that best fit the vehicle image. Once the best fit model is determined, this 2D model is transformed back to 3D where 3D parameters are calculated. In general, if a vehicle is not occluded, the area of the model should be slightly larger than the area of the vehicle image. When occlusion occurs, this ratio changes. In addition, the width and height of the model also indicate the presence of occlusion.

3. BINARY REPRESENTATION OF A VEHICLE

In principle, the deformable model can be applied directly to a segmented version of the image. The problem with this is that if the segmentation algorithm is not ideal, the segmented result may not be faithfully representing the vehicle concerned. Besides, segmentation usually ends up with far more details than required, and can introduce unwanted effects when fitting the model[9]. Based on this, the concept of background estimation[10] is adopted here. By estimating the stationary background, the moving vehicle (object) can be deduced by simply subtracting the background from the original image, as depicted in Figs. 2(b) and 2(c). Once the vehicle is extracted, it's binary representation can be easily determined as given in Fig. 2(d). In this case, only the 2D outline of the vehicle is important for the model fitting, not any internal details.

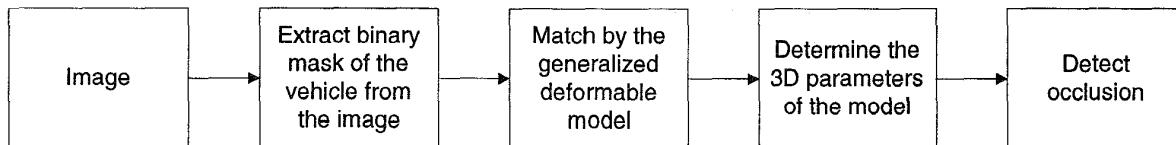


Figure 1 - Conceptual diagram of the proposed vehicle occlusion detection method

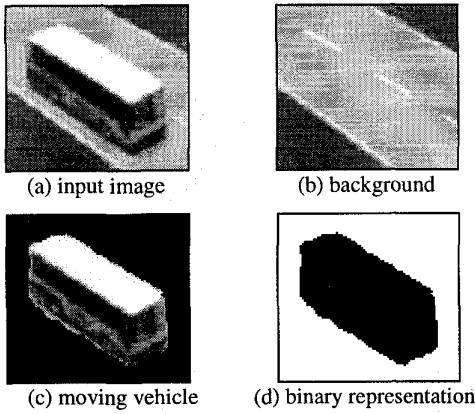


Figure 2 - Binary representation of a vehicle

4. GENERALIZED DEFORMABLE MODEL

The generalized deformable model, M , proposed is basically a solid 3D cuboid with 6 vertices, $M = \{v_k | k = 0, \dots, 5\}$. The vertices are parameterized such that their values can be calculated to fit a particular vehicle. A model of this type has several properties: 1) it can be used to model any type of vehicles; 2) easily deformed to fit any size of vehicles; 3) does not suffer from incomplete or occluded vehicle images. However, it should also be pointed out that because of the simplicity of the model, the type of vehicles it can differentiate would be limited. This is a compromise from the more complex deformable model with many more vertices.

In general, the model can be realized as a modified 3D wire frame projected on the a 2D plane as shown in Fig. 3. This is a modified version as the front facing three edges (dotted lines) are not considered as part of the wire frame. In other words, the 2D model consists of only the outlines defined by v_0 to v_5 . For this reason, this model has three pairs of parallel edges along three 2D orientations (\bar{x} , \bar{y} and \bar{z}) as depicted. The fitting is only confined to these six lines.

Once the model fitting in 2D is completed, the \bar{x} , \bar{y} and \bar{z} orientation components must be transformed to 3D orientations \bar{X} , \bar{Y} and \bar{Z} , if the 3D vehicle features are to be determined. This transformation between the 2D and 3D spaces is mainly determined by the camera model depicted in Fig. 4, where the camera is focusing at point O with focal length f . The

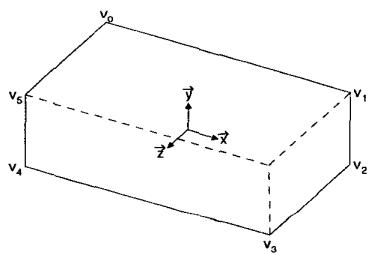


Figure 3 - Generalized deformable model.

camera height is h and the distance between the focusing point and the camera is d . A 3D point $P = (P_u, P_v, P_w)$ is thus projected onto the 2D camera image plane as $p = (p_x, p_y)$.

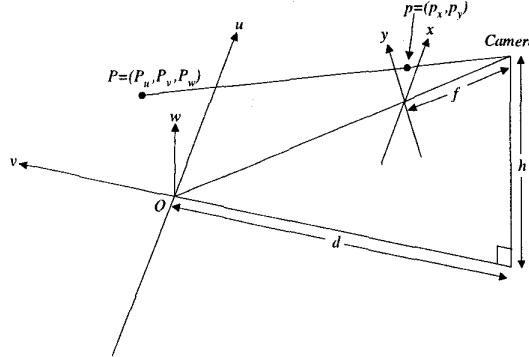


Figure 4 - Camera model.

From this camera model, the 3D-to-2D forward mapping function, Φ , and the 2D-to-3D inverse mapping function, Φ^{-1} , are computed by Eqs. (1) and (2) respectively.

3D-to-2D mapping function:

$$p_x = P_u \cdot \sqrt{\frac{f^2 + p_y^2}{(d + P_v)^2 + (h - P_w)^2}} \quad \dots(1a)$$

$$p_y = f \cdot \frac{hP_v + dP_w}{d \cdot (d + P_v) + h \cdot (h - P_w)} \quad \dots(1b)$$

2D-to-3D mapping function:

$$P_u = p_x \cdot \sqrt{\frac{(d + P_v)^2 + (h - P_w)^2}{f^2 + p_y^2}} \quad \dots(2a)$$

$$P_v = \frac{p_y \cdot (d^2 + h \cdot (h - P_w)) - f \cdot d \cdot P_w}{f \cdot h - p_y \cdot d} \quad \dots(2b)$$

When the three 2D orientations \bar{x} , \bar{y} and \bar{z} are determined, model M is created along these orientations and its vertices are adjusted to fit the vehicle image, where these vertices must satisfy the following criteria:

$$|v_0 - v_1| = |v_3 - v_4| \text{ and } |v_1 - v_2| = |v_4 - v_5| \text{ and } |v_2 - v_3| = |v_5 - v_0| \quad \dots(3)$$

where $|a-b|$ denotes the distance between a and b . This ensures the model obtained is regular as the model does not make any assumption on the camera perspective. If we allow M to be irregular without knowing or considering the camera perspective, such irregularity may be due to the camera perspective or the shape of the vehicle. Using the above fitting approach and criteria, Fig. 5 illustrates an example of a vehicle fitted by the model. It can be seen that the model fitted is not perfect, but is close enough to represent the vehicle.

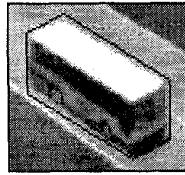


Figure 5 - Vehicle fitted with the deformable model

5. 3D MODEL PARAMETERS

To determine the 3D parameters of model M , it is necessary and appropriate to assume that V_2 , V_3 and V_4 are on the plane of the road surface, i.e. $P_w = 0$. Therefore, the *width* and *depth* of the model can be computed by Eqs. (4) and (5) respectively. Furthermore, as V_1 is directly above V_2 , by solving Eqt. (6), the *height* of the model can be determined.

$$\text{width} = |\Phi^{-1}(v_2, 0) - \Phi^{-1}(v_3, 0)| \quad \dots(4)$$

$$\text{depth} = |\Phi^{-1}(v_3, 0) - \Phi^{-1}(v_4, 0)| \quad \dots(5)$$

$$\begin{cases} \text{height} = V_1 - V_2 \\ \Phi^{-1}(v_1, \text{height}) = \Phi^{-1}(v_2, 0) + \begin{pmatrix} 0 \\ 0 \\ \text{height} \end{pmatrix} \end{cases} \quad \dots(6)$$

6. OCCLUSION DETECTION

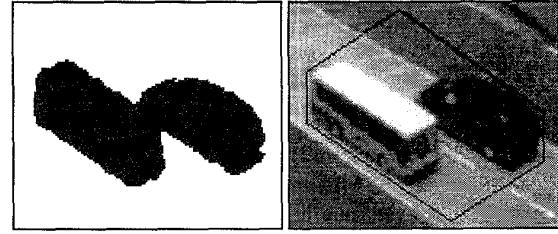
Essentially, occlusion is detected by monitoring the changes of the 3D model parameters and an extra parameter called *Area Ratio*, R_{area} , as given in Eqt. (7).

$$R_{area} = \frac{A_{mask}}{A_{model}} \quad \dots(7)$$

where A_{mask} is the area of the binary vehicle image and A_{model} is the area of model M . The *Area Ratio* basically describes how well the model fits the binary vehicle image. As A_{model} is always larger than A_{mask} , R_{area} is always less than one. An ideal fit is that R_{area} equals to one, whereas a good fit means R_{area} is close to one, and a poor fit means otherwise. The difference in R_{area} can also be used to discern different types of vehicles in a limited sense.

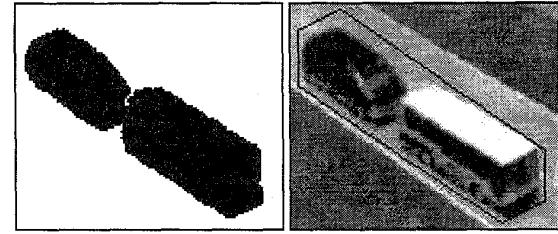
The argument is that when occlusion occurs, R_{area} and the dimension of the model change abruptly. Figs. 6 and 7 illustrate two possible occlusion cases. In Fig. 6, the vehicle on the right is occluded by the bus in the adjacent lane. The 3D model fitted as a result is much wider than expected and does not match with the usual dimensions of typical vehicles. It also reduces the R_{area} value substantially. Fig. 7 is similar except that the model lengthens. The reduction in R_{area} is not as severe as Fig. 6.

Therefore, the monitoring of occlusion is carried out by comparing the 3D dimensions and R_{area} of the model in the current image, n , with the last image without occlusion, m , according to Eqt. (8).



(a) Binary representation (b) Model fitted image

Figure 6 - Occlusion of vehicles in adjacent lane



(a) Binary representation (b) Model fitted image

Figure 7 - Occlusion of vehicles in the same lane

$$|R_{area}(n) - R_{area}(m)| > \varepsilon_1 \text{ or}$$

$$|R_{width}(n) - R_{width}(m)| > \varepsilon_2 \text{ or } |R_{height}(n) - R_{height}(m)| > \varepsilon_2 \quad \dots(8)$$

where R_{width} and R_{height} are the normalized *width* and *height* of the vehicle model. The values ε_1 and ε_2 are thresholds to permit small changes in the values. The values of ε_1 and ε_2 used in our evaluations are both 0.05. When there is no occlusion, R_{area} should be relatively constant, close to one. This is also true for the other model parameters too. However, when occlusion occurs, R_{area} decreases and some of the other parameters increase.

To prove the proposed occlusion detection method, two model cars moving in a laboratory environment are set up, with a camera mounted on a tripod at about 0.3m above a model road. An image sequence was acquired and subject to the method proposed in Fig. 1, where for each image, the vehicles' binary representations are generated, 2D model fitted, and 3D model and parameters determined, as well as R_{area} calculated. Fig. 8 depicts a sample set of the image sequence with the 3D model overlaid on the images, and Table 1 shows the parameters obtained for the 'bus' and the detection results.

Image	R_{area}	R_{width}	R_{height}	Occlusion ?
1	0.88	0.42	0.57	✗
2	0.87	0.44	0.56	✗
3	0.60	0.62	0.32	✓
4	0.79	0.82	0.60	✓
5	0.80	0.37	0.84	✓
6	0.61	0.22	0.64	✓
7	0.85	0.40	0.55	✗
8	0.86	0.38	0.54	✗

Table 1 - Simulation results for 'bus'

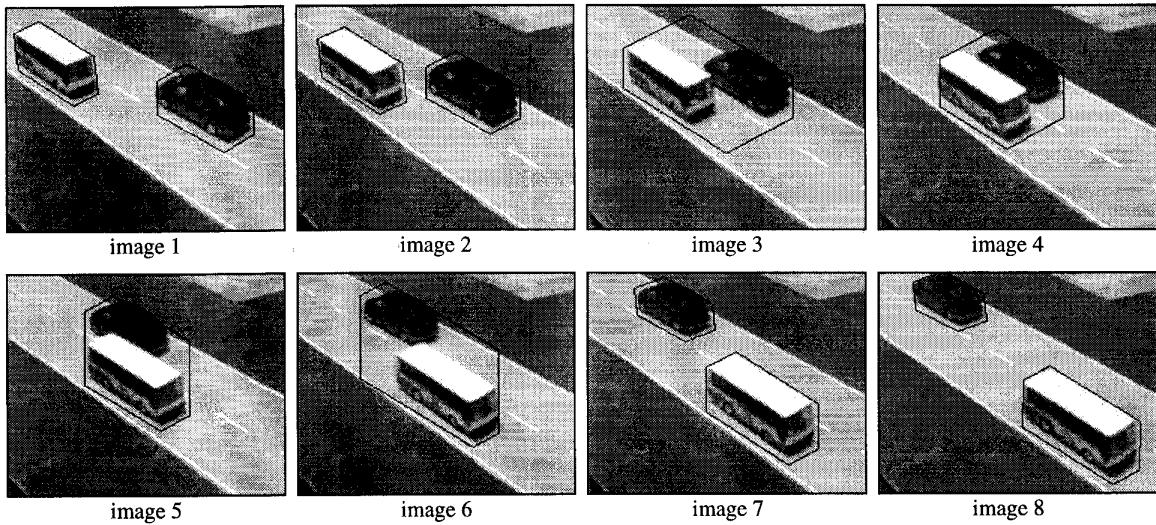


Figure 8 - Simulation results for 'bus'

As can be seen from the table, the non-occluded cases have R_{area} in the range of 0.85 or above, and the model width and height vary within a small range. On the other hand, the occluded cases have small R_{area} and both R_{width} and R_{height} different from the typical figures by a large margin.

7. CONCLUSION

In conclusion, the occlusion detection method based on a generalized deformable model as presented in this paper is an effective way for detecting vehicle occlusion. Central to the method is a generalized deformable model, which is characterized by its small number of vertices and the transformation between the 2D and 3D spaces for model fitting. According to our preliminary tests, the model is able to fit different types and sizes of vehicles. By monitoring the changes in the area ratio, model width and height with respect to previous images, vehicle occlusion can be detected successfully. This result will have positive impact to the subsequent segmentation and tracking of vehicles. Further development of this method will be focused on the evaluation on road traffic scenes and on the study of how the generalized model can be used for a wider class of vehicles.

8. REFERENCES

- [1] Neil Hoose, "Computer Image Processing in Traffic Engineering", *Research Studies Press Ltd.*, London, 1991.
- [2] Marie-Pierre Dubisson Jolly, Sridhar Lakshmanan, and Anil K. Jain, "Vehicle Segmentation and Classification Using Deformable Templates", *IEEE Trans. on PAMI*, vol. 18, no. 3, pp. 293-308, 1996.
- [3] G. D. Sullivan, K. D. Baker, A. D. Worrall, C. I. Attwood and P. R. Remagnino, "Model-Based Vehicle Detection and Classification Using Orthographic Approximations", *Proc. of 7th British Machine Vision Conference*, 9th September, 1996.
- [4] D. Koller, K. Daniilidis, T. Thorhallson, and H.-H. Nagel, "Model-Based Object Tracking in Traffic Scenes", *Proc. 2nd European Conference on Computer Vision ECCV'92*, pp. 437-452, May, 1992.
- [5] Warren F. Gardner and Daryl T. Lawton, "Interactive Model-Based Vehicle Tracking", *IEEE Trans. on PAMI*, vol. 18, no. 11, pp. 1115-1121, 1996.
- [6] Thomas Frank, Michael Haag, Henner Kolling, and Hans-Helmut Nagel, "Tracking of Occluded Vehicles in Traffic Scenes", *4th European Conference on Computer Vision ECCV'96*, pp. 487-490, April, 1996.
- [7] Dieter Koller, Joseph Weber, and Jitendra Malik, "Robust Multiple Car Tracking with Occlusion Reasoning", *UCB Technical Report UCB/CSD-93-780*, University of California, October, 1993.
- [8] Sylvia Gil, Ruggero Milanese and Thierry Pun, "Combining Multiple Motion Estimates for Vehicle Tracking", *Computer Vision - ECCV'96*, Volume II, Springer, Germany, pp. 307-320, 1996.
- [9] Nelson H. C. Yung & Andrew H. S. Lai, "Segmentation of Color Images Based on the Gravitational Clustering Concept", *to appear in SPIE Journal - Optical Engineering*, vol. 37, no. 3, March, 1998.
- [10] Andrew H. S. Lai & Nelson H. C. Yung, "A Fast and Accurate Scoreboard Algorithm for Estimating Stationary Backgrounds in an Image Sequence", *to appear in Proc. ISCAS'98*.