

# Stereovision-Based Object Segmentation for Automotive Applications

**Yingping Huang**

*International Automotive Research Center, Warwick Manufacture Group, University of Warwick, Coventry CV4 7AL, UK  
Email: yingping.huang@warwick.ac.uk*

**Shan Fu**

*Applied Mathematics & Computing Group, School of Engineering, Cranfield University, Bedford MK43 0AL, UK  
Email: s.fu@cranfield.ac.uk*

**Chris Thompson**

*Applied Mathematics & Computing Group, School of Engineering, Cranfield University, Bedford MK43 0AL, UK  
Email: chris.thompson@cranfield.ac.uk*

*Received 14 January 2004; Revised 8 November 2004*

Obstacle detection and classification in a complex urban area are highly demanding, but desirable for pedestrian protection, stop & go, and enhanced parking aids. The most difficult task for the system is to segment objects from varied and complicated background. In this paper, a novel position-based object segmentation method has been proposed to solve this problem. According to the method proposed, object segmentation is performed in two steps: in depth map ( $X$ - $Z$  plane) and in layered images ( $X$ - $Y$  planes). The stereovision technique is used to reconstruct image points and generate the depth map. Objects are detected in the depth map. Afterwards, the original edge image is separated into different layers based on the distance of detected objects. Segmentation performed in these layered images can be easier and more reliable. It has been proved that the proposed method offers robust detection of potential obstacles and accurate measurement of their location and size.

**Keywords and phrases:** stereovision, segmentation, objects detection.

## 1. INTRODUCTION

Vision-based driver assistance system in complex urban area is highly demanding, but desirable for pedestrian protection, stop & go, and enhanced parking aids. The basic requirement for the system is the capability of detecting potential obstacles and providing complete three-dimensional (3D) information of the obstacles, that is, size, location. Furthermore, the capability of classifying obstacles is also essential for the system to better interpret the driving environment and give correct reaction.

Two basic techniques have been adopted for vision-based obstacle detection: optical flow and stereovision. The optical flow technique is equipped with monocular camera, and objects are segmented according to motion pattern (e.g., optical flow vectors) by analysing two or more consecutive images taken at different time instants [1, 2, 3, 4]. Viewing a scene from two different points allows us to extract 3D structure of the scene. Franke et al. [5, 6] applied stereovision technique to interpret urban traffic scene. Bohrer et al. [7] combined

stereovision technique with inverse-perspective method to warp the left and right images so that all ground plane points have zero disparity. Simple differencing or low correlation values between identically located images points correspond to an obstacle. Bertozzi et al. [8, 9] employed the similar method in their GOLD project to detect obstacle in high way traffic. In this paradigm, the perspective effect is removed from the stereo images, and the two images generated present the differences where the initial assumption of a flat road is not valid, thereby detecting the free space in front of the vehicle. These applications were not concentrated on object segmentation. In this paper, we proposed a different way to use stereovision technique for object segmentation. Moreover, a tailored area-based stereo matching has been designed to achieve quality disparity map.

Object segmentation is a major difficulty because of the complexity of urban environment where obstacles to be detected are various and mixed with buildings, trees, traffic signs, and road markings. Traditional image segmentation approaches such as boundary-detection and region-growing

methods are almost impossible to realise this goal without semantic knowledge about the scene. Some special segmentation strategies have been designed for automotive applications. Symmetry-based [9] and pattern-based [10, 11] segmentation methods were widely used for the tracking of a vehicle in highway scene. These segmentation methods make only use of 2-dimensional projection feature of object without consideration of its depth information. This paper presents a novel position-based segmentation method in virtue of stereovision technique. The idea is based on the fact that potential obstacles may be mixed in projection image, but must be separated in depth map in terms of their lateral and longitudinal positions. Position-based segmentation is therefore performed in two steps; in depth map ( $X$ - $Z$  plane) and in layered images ( $X$ - $Y$  planes). The depth map is generated by calculating 3D world coordinates of image points and transforming them into a scaled  $X$ - $Z$  plane. In the depth map, objects can be easily segmented by conventional object grouping method. The first segmentation serves to detect objects and provide precise location and preliminary size of the objects. The second step of the method is performed in layered images, which are obtained by separating the original image into different layers based on the distances of detected objects. Each image layer contains fewer or even only one object, therefore, object segmentation based on these images is easier and more reliable than doing that in the original image. By this step, refined object size is generated. Furthermore, separating the image by this method will also benefit further object classification. Image separation was also used by Fang and Masaki [12] to segment targets for intelligent vehicles. However, their method was based on an assumption that object distances had been detected by radar sensor and did not perform segmentation in the depth map.

The proposed method has been verified by a sample scenario in the paper. A prototype system based on the proposed method has been implemented and integrated into a Jaguar XKR-Type demonstrator. A dual-processor (2.4 GHz CPU) industrial PC was employed in the system. The system was capable of running in a real-time basis with the update rate of 12 Hz for a  $320 \times 240$  image size. Substantial experimental studies have proved that the method proposed in this paper offers robust object detection and accurate measurements on their location and size. The paper is composed of six sections. Section 2 introduces two schemes of stereo matching tailored to our application. Section 3 describes 3-dimensional reconstruction of image. Detailed two-step segmentations are presented in Sections 4 and 5. Finally, conclusions are given in Section 6.

## 2. STEREO MATCHING

It is known that stereovision obtains 3D information by establishing correspondence between the left and right images and applying triangulation. The core of a stereovision algorithm is correspondence searching (stereo matching) between the images, which accounts for most computational time and makes the most significant effect on system performance. Many matching approaches have been reported for

a variety of applications [13, 14, 15]. Area-based correlation method has been adopted in our system with specific considerations. The sum of absolute differences (SAD) over the two matching windows is used as correlation measure because it is a fast and effective measure. The SAD of the two windows is calculated from the following equation:

$$SAD = \sum_{(i,j) \in W(m,n)}^N [x_{i,j} - y_{i,j}], \quad (1)$$

where  $x_{i,j}$  and  $y_{i,j}$  are the grey levels within the rectangular window  $W(m, n)$ ,  $m$  and  $n$  are the sizes of the window. The calculated SAD indicates the similarity extent of intensity profile of the two rectangular windows. The area-based matching is conducted on the images filtered by the Laplacian of Gaussian (LOG) rather than on raw images so as to remove the effects of intensity variations between images due to difference of camera gains, ambient light, and perspective [15]. Moreover, LOG transform of raw image enhances the image features, and increases the signal-to-noise ratio of matching. In consideration that the area-based matching may generate false matching in nontextured region of image, a confidence measure based on edge energy is applied as matching control, which gives high confidence to regions that are textured in intensity. The deviation of the grey level within the matching windows can be a measure of the texture, and therefore is taken as the confidence measure. Since the stereo matching is based on the LOG-transformed images where the edges of the objects are enhanced and other flat areas are homogenized, the adopted confidence measure indicates the edge energy. This operation increases the reliability of the disparity image but reduces its density.

The area-based matching applies correspondence searching to all image points and generates a dense disparity, thereby leading to more expensive computation. Two constraints are used in our algorithm to improve computational efficiency. There exists a strict constraint of search position for binocular stereovision, that is, corresponding points between the left and right images must be on the corresponding epipolar lines. For a parallel-axis stereovision rig, the epipolar lines coincide with the horizontal scan lines. Since our cameras are aligned, the corresponding search can be only conducted on a single scan line. In addition, constraint on disparity search range, which corresponds to the depth range of interest, is also helpful to improve computational efficiency. Thus, the correspondence searching can be limited within a disparity range.

The procedure for the area-based matching is introduced as follows.

- (1) Apply LOG operator into the input images to obtain two transformed images.
- (2) Pick points one by one in the left transformed image as seed points, and find out all candidate points in the right transformed image, which are located on the left-hand side of the point with the same coordinates within a certain range along the same scan line.



FIGURE 1: Sample image indicating a typical urban traffic scenario.

- (3) Assign a low-aspect-ratio rectangular window to the seed and candidate points.
- (4) Calculate the *SAD* of the candidate point pairs. The output at this stage is an array of *SAD* values corresponding to the candidate point pairs.
- (5) Find out the minimum *SAD* value and check if it is less than a prespecified acceptance level. If it is, the point pair is regarded as the corresponding points. The disparity between the points is calculated.
- (6) The confidence measure is applied to filter the potential mismatches. In particular, the deviation of the grey level within the matching windows is calculated and checked with a predefined value. The area with the confidence measure below this value is filtered out.

There are two points worthy to be noted: size of matching window and disparity accuracy. The size of matching window is a compromise. Small window size is more likely to be similar in images with different viewpoints, while large size (not over large) increases the reliability. The shape of the window should have a low aspect ratio even tending to a line window. In addition, disparity detected must have a subpixel accuracy to obtain an accurate depth measurement since depth resolution per pixel disparity can be huge in a far distance. To achieve subpixel disparity, a quadratic fit process is then performed on the *SAD* value array obtained in step (4) to estimate the optimal position where the minimum *SAD* value occurs. This position has subpixel accuracy, and consequently we will obtain a very precise disparity map with subpixel accuracy.

Figure 1 is the left image of a stereo image pair, indicating a typical scene of urban traffic. This image is used in this paper as a sample scenario to be analysed. The scenario was composed of two parked cars, pedestrian, moving truck, and roadside trees. Figure 2 shows its disparity histogram obtained from area-based stereo matching. The disparity histogram indicates that the closer the point, the greater the disparity. It can also be seen that the image had uniform disparity distribution in most of the regions and some noise patches caused by false-matching scattered randomly in the image. The noise patch pointed out in Figure 2 made an impact on detecting car 1, which will be discussed in Section 4.2.

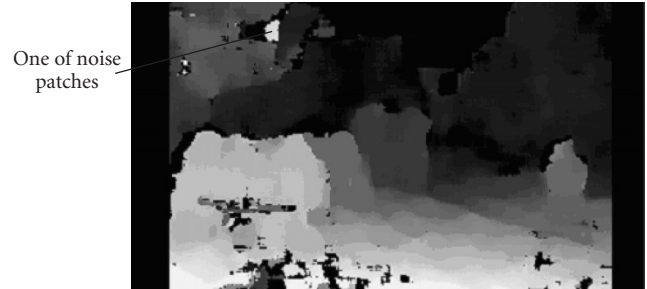


FIGURE 2: Disparity image generated by area-based matching.

### 3. 3D RECONSTRUCTION OF IMAGE

Figure 3 indicates the geometry of a parallel-axis stereo rig with a baseline  $b$ . In our system, the world coordinates were set to be coincided with the coordinates of the left camera. The coordinate origins were located in the centres of the images. For a given scene point  $P(X, Y, Z)$ , the projection points on the left and right images are  $p_l(x_l, y_l)$  and  $p_r(x_r, y_r)$ , respectively, and its world coordinates can be obtained from

$$X = \frac{b \cdot x_l}{d}, \quad y = \frac{b \cdot y_l}{d}, \quad Z = \frac{b \cdot f}{d}, \quad (2)$$

where  $f$  is the focus length of the lens,  $d$  is the disparity equalling to  $x_l - x_r$ . Based on these relations, all points in the image with a nonzero disparity can be reconstructed into real 3D world.

Not all points in the disparity map correspond to an obstacle. Some of them correspond to those lying on the road surface, for example, road boundaries, lane markings, shadows of objects. These points are not regarded as obstacles and should be discarded. Assuming a planar road surface, the camera height is  $H$ , and the tilt angle towards the road plane is  $\theta$ . The  $Y$ -axis coordinate ( $Y_g$ ) of those points lying on the road surface depends on distance ( $Z$ ):

$$Y_g = \frac{Z \cdot \sin(\theta) - H}{\cos(\theta)}. \quad (3)$$

For those points with  $Y$ -axis coordinate greater than  $Y_g$ , we classified them as obstacle points above the road.

## 4. OBJECT SEGMENTATION IN DEPTH MAP

### 4.1. Formation of depth map

As long as 3D information of image points has been obtained, a transform from  $X$ - $Y$  plane to  $X$ - $Z$  plane can be performed to generate depth map. A depth map is a bird's-eye view of a 3D scene, where the horizontal and vertical coordinates correspond to lateral and longitudinal distances, respectively. The formation procedure of depth map is introduced as follows.

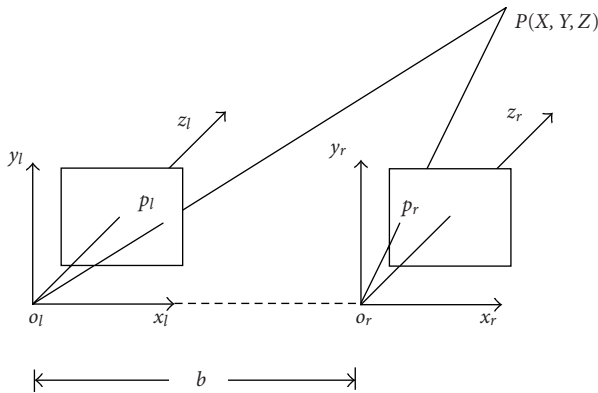


FIGURE 3: Geometry of parallel-axis stereo rig used in this study.

- (1) Scale the area of interest into a map with a predefined range resolution corresponding to a rectangular area in real scene.
- (2) Calculate 3D world coordinates of all image points according to (2), and discard the points on road surface according to (3).
- (3) Locate the left points in the map by transforming their world coordinates  $(X, Z)$  into the map coordinates. Count the points that fall into each rectangular area, that is, one pixel in the map.
- (4) Draw the depth map by using the number as the grey level of each corresponding pixel.

The binary depth map of Figure 2 is shown in Figure 4, in which the lateral ranges from  $-8$  to  $8$  m, the longitudinal ranges from  $4$  to  $60$  m, and the range resolution is  $0.2 \times 0.4$  m. The binary depth map was achieved by applying a grey-level threshold into the depth map. This threshold also acts as a cutoff to discard some points in the depth map that contains less nonzero points in the disparity map. The points lying on the road surface such as lane marks have been discarded. The five dense clusters of points corresponded to the five objects. It is evident that these objects were separated in terms of their lateral and longitudinal positions. It should be noted that all points in the depth map are visible in the disparity map because they all come from it. The volume of the point cluster determines the width and thickness of the object. In fact, the thickness could be very rough since the object cannot be viewed through in the deep direction, and errors exist during the disparity analysis.

#### 4.2. Object grouping

Traditional region-growing method was used to group the points in the depth map into different object entities. All 8-orientation-connected points are grouped as one object. The procedure for the operation is as follow.

- (1) Pick a point as a seed point, group points connected from 8 directions as one object, and assign the object number to all grouped points.
- (2) Locate next seed point and repeat step (1) until all nonzero points have been labelled.



FIGURE 4: Depth map generated from Figure 2.

Another way of object grouping is to search for the centroid point of each point clusters by using a Gaussian-based kernel, and then a boundary condition can then be established to encapsulate the object. The shape of the kernel depends on the object for which we are looking. For example, a kernel size equivalent to  $1\text{m}$  by  $1\text{m}$  is good to isolate human objects. This method is more effective to segment specific objects with known aspect ratio.

Following the object grouping, the points in the original image within the objects are known. The lateral and longitudinal distances of the individual objects were then calculated from the average values of all points within the object. The width and height of the object were determined from the boundary points within the object. In addition, point within the objects was counted. A predefined threshold based on this number was applied to remove spurious objects, which contained very few remapped points or had negligible physical size. This threshold value was dynamic and inversely proportional to the object distance since the image size of an object is inversely proportional to the distance due to the perspective. Furthermore, all objects segmented were prioritized according to the longitudinal distance; the closer to the camera, the higher priority the object possesses.

Back to the sample image, the detected objects were outlined with rectangular boxes in the original left image, as shown in Figure 5. The rectangular boxes represented the maximum boundary of the objects. Five objects were detected, that is, two cars, pedestrian, truck, and tree. The measured location and size of the objects are listed in Table 1, where  $X$ ,  $Y$ , and  $Z$  represent the average coordinates of the object points. Please note that the origin of the coordinates is the centre of the image, as indicated in Figure 3. It can be seen that the boundary of car 1 was wrongly extended upright due to the noise patch pointed out in Figure 2. This noise patch generated some points with the same coordinates as car 1 in depth map, and therefore regarded as part of car 1 in depth map, and therefore regarded as a magnified car height. Careful observations on Figure 5 indicate that rectangular boundary of the objects was not precisely right, such as pedestrian. This was because the object contour might be enlarged or shrunk by transforming image points into depth map due to imperfect stereo matching.



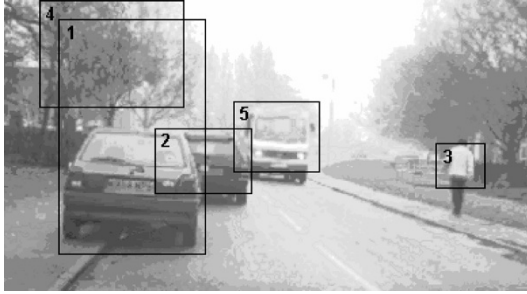


FIGURE 5: Object boundary obtained by first segmentation.

## 5. OBJECT SEGMENTATION IN LAYERED IMAGES

Object segmentation in depth map segments objects and tells their location and size information by making use of the  $X - Z$  information. However, the detected height and width of objects can be wrong or imperfect due to the impact of other objects located in the same position as the objects but with different height or noise points in the disparity image. To obtain the refined height and width, the best way is to do further segmentation in the original  $X - Y$  image. In addition, objectives of the system are not only detecting objects but also classifying them. For classification purpose,  $X - Y$  information of objects must be used since object classification is normally based on its geometry shape. From this point of view, we also need to go back to the original images. It is evident that the original image is a mixture of all objects; hence object segmentation and classification are extremely difficult if the operations are directly performed in this image. To solve these problems, we proposed that further object segmentation and classification could be operated in different image layers. Each layer of the image contains fewer objects or even only one object, thereby enabling easier and more reliable object segmentation and classification. That is, the original difficult task can be decomposed into multiple simpler tasks. The basis of layering the original image is the distances of the objects detected by the first segmentation step. Since object distance is determined from disparity of the points within the object, disparity image is firstly layered based on determined disparity ranges. Accordingly, the original image can be layered by using layered disparity images as index. In contrast to the layered disparity images, the layered greyscale images contain richer information, which is useful for further segmentation and classification.

### 5.1. Formation of image layers

Assuming  $N$  objects have been detected by previous processing, the relationship between distance and disparity generates  $N$  corresponding groups of points in the disparity image and in the original image. In order to ensure that all image points within the objects will be included in the image layer, a range of disparity is used to sort the image points in disparity map. This range should cover the span of disparity of all points within the object and can be represented as

$$\left( \frac{b \cdot f}{Z} - \Delta d, \frac{b \cdot f}{Z} + \Delta d \right). \quad (4)$$

TABLE 1: Detected objects, location and size.

Objects	Car 1	Car 2	Pedestrian	Tree	Truck
Z (m)	10.6	19.6	20.0	29.2	36.8
X (m)	-1.65	-1.02	4.14	-5.32	-0.23
Y (m)	-0.33	-0.29	-0.51	4.12	0.30
Width (m)	1.98	1.64	1.32	4.49	3.51
Height (m)	2.92	1.45	1.22	4.78	3.02
Disparity	11.03	5.96	5.87	4.25	3.18

The formation of image layers can be described as follows.

- (1) Determine the number of image layer according to the number of detected objects, and the disparity range for each image layer.
- (2) Sort the points in the disparity image according to the disparity ranges. Use the sorted points as index to obtain corresponding points in the original image. The image layer can be drawn by using these points.

For the sample image, the disparity values corresponding to the five objects are also listed in Table 1. Ignoring the trees and in consideration that car 2 and the pedestrian were located at similar distance, three disparity values 11.03, 5.9, and 3.18 were taken as the center of the disparity range. Disparity ranges were determined as (9.5, 12.5), (4.9, 6.9), and (2.68, 3.68). Three image layers were generated and displayed in Figures 6a, 6b, and 6c. In contrast to the Figure 1, four obstacles have been separated into three individual images. Whether object segmentation or classification based on these images would be easier because each image only contains one or two objects.

### 5.2. Object boundary determination

The way of determining object contour can be various in terms of applications and requirements. In this study, we implemented an edge-linking method to determine the object contour. The method is introduced as follows.

- (1) Extract the edge layers of the raw image. Actually, the method of layering the raw image described above can be also applied to the LOG-transformed image. Figure 7a shows the binary edge image which was generated by applying a cutoff threshold to the LOG-transformed image. This edge image was separated into three layers corresponding to three layers of the raw image, as shown in Figures 7b, 7c, and 7d.
- (2) Apply a morphological “opening” operation into the edge image to smooth the object contour. Firstly, the image is eroded to eliminate some scattered noise points adjacent to the edge. A small element consisting only two vertical pixels was selected as erosion element. The dilation operation is followed to connect some separated edge points so that the contour is enhanced and the length is increased. By these operations, the object should be outlined with a smoother edge contour, and some noise points have been filtered out.

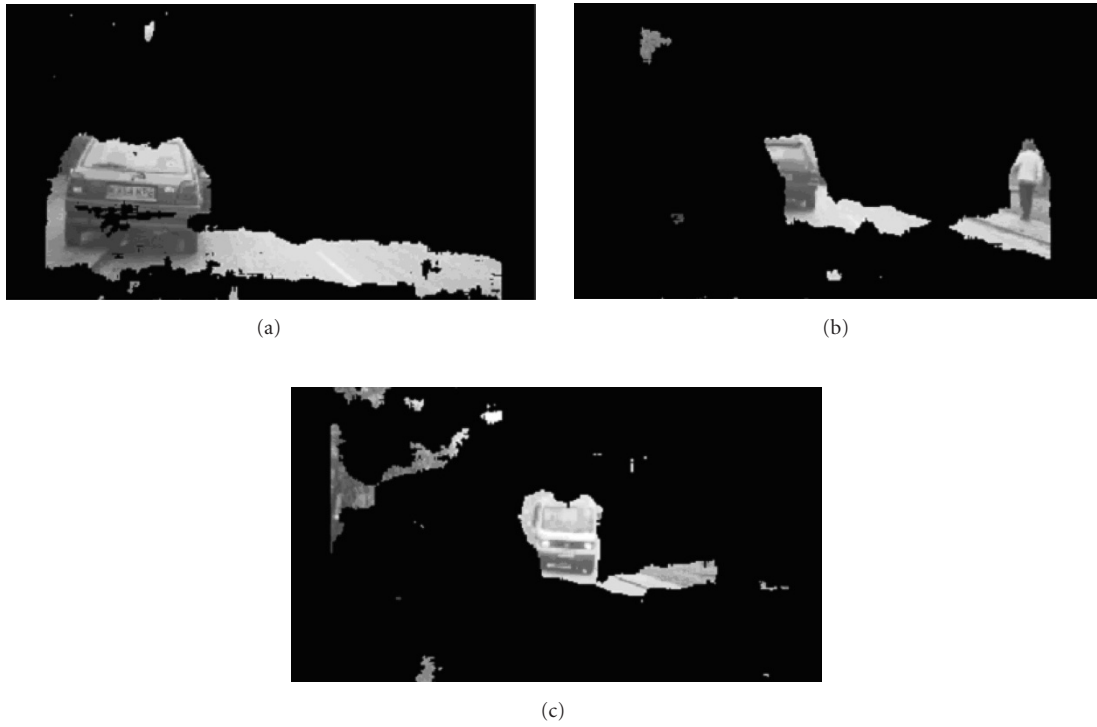


FIGURE 6: Layered images. (a) Image layer 1 containing car 1. (b) Image layer 2 containing car 2 and pedestrian. (c) Image layer 3 containing truck.

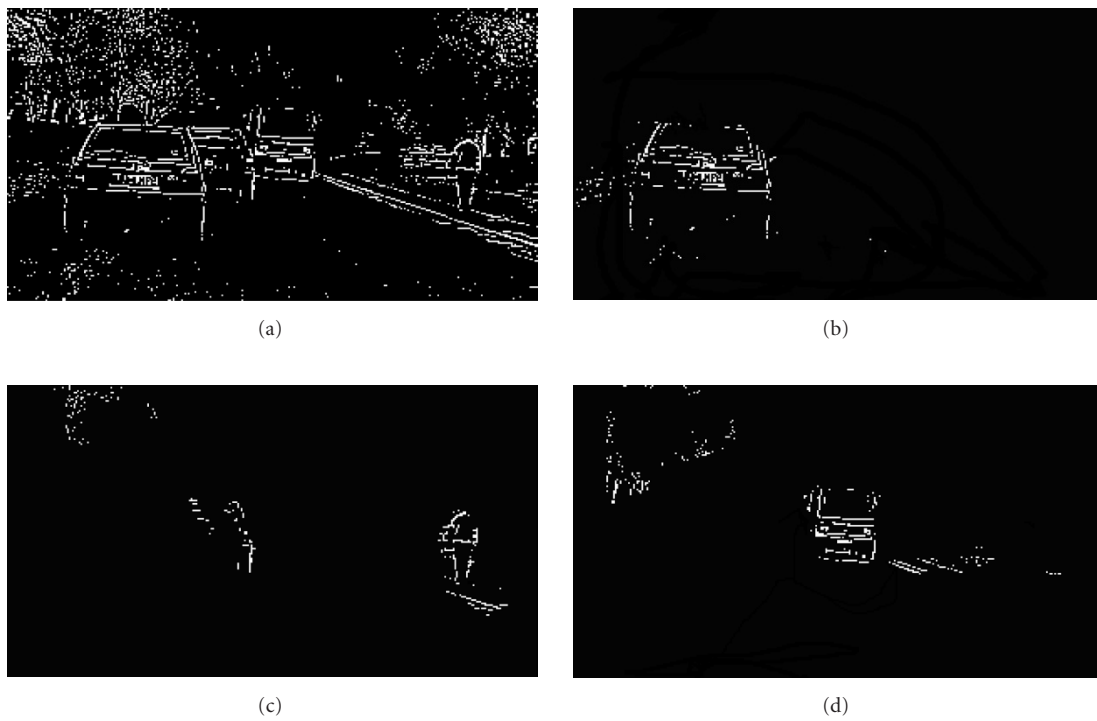


FIGURE 7: Layered edge images. (a) Edge image. (b) Edge layer 1 containing car 1. (c) Edge layer 2 containing car 2 and pedestrian. (d) Edge layer 3 containing truck.

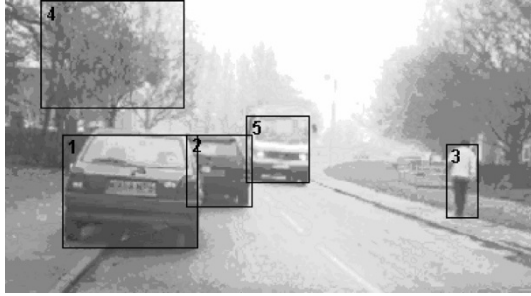


FIGURE 8: Refined object boundary obtained by second segmentation.

- (3) Link the connected points within a local neighborhood from 8 directions. The size of the neighbourhood is a compromise. A larger scale of the neighborhood is beneficial for overcoming the gaps of the contour, but may cause a false link to the noise points. In the meantime, count the number of the connected points as the contour length. By these operations, a long continuous object edge contour is generated. Note that the linked edge contour would also contain some inner points.
- (4) Locate next seed point and repeat step (3) until all points have been linked as a contour.
- (5) Apply a constraint to the length of the detected contours. This constraint is also configured by considering the perspective relationship since the distance has been measured. The short edge contours are more likely caused by the noise points and discarded. The boundary points of the edge contour of the confirmed objects form a rectangular bounding box which gives the width and height of the objects.

Note that the operations addressed above were only applied to a local area rather than global image to save the computation time. The local area was determined from the distribution range of the points, which had been grouped into the object in depth map and remapped back into the original image. In addition, objects can be scaled to a standard size by using perspective principle because the distance of the objects has been measured, which would also be very helpful to further object classification.

Back to the example, the boundary of the four objects was detected by applying above operations into the three images, respectively. The object boundaries detected are shown in Figure 8, the measured width and height of the objects are listed in Table 2. In contrast to Figure 5 and Table 1, the detected boundaries were more precise, and measured object sizes were more accurate.

Apart from the example presented here, the proposed method has been validated by a number of scenarios. The statistic experimental results indicate that the success rate on object detection reaches 95% within a detection range of 4–50 m under reasonable illumination conditions, and the relative errors for distance and size measurements are less than 5% and 10%, respectively.

TABLE 2: Refined object location and size.

Objects	Car 1	Car 2	Pedestrian	Truck
Width (m)	1.76	1.28	0.61	2.98
Height (m)	1.49	1.44	1.65	3.02

Precise measurement on object size is very useful for object classification. Preliminary classification based on aspect ratio has currently been available in our system. More precisely classification based on point-distributed model and statistic criteria is in progress in virtue of this image separating method.

## 6. CONCLUSIONS

Object segmentation is a major difficulty for vision-based driver assistance system due to the complexity of urban environment where obstacles to be detected are various and mixed with buildings, trees, traffic signs, and road markings. This paper proposes a novel stereovision-based object segmentation method for this application. Stereovision technique is employed to reconstruct image points into 3-dimensional world coordinates. Object segmentation makes use of 3-dimensional information by splitting it into two steps. The first segmentation step is performed in depth map ( $X - Z$  plane) and provides accurate location and preliminary size information. The second segmentation step is operated in layered images ( $X - Y$  planes) and generates refined size information. It has been proved that the method proposed in this paper offers robust detection on potential obstacles and accurate measurement on their location and size. In details, a tailored area-based scheme has been designed to produce dense and high-quality disparity image, which delivers rich information for further processing. Accordingly, the depth map is generated by transforming image points into a scaled  $X - Z$  plane. In the depth map, objects can be easily segmented by using conventional object grouping methods. Based on the distance of detected objects, the original complicated image can be separated into different layers. Each layer of the image contains fewer objects or even only one object, and object segmentation based on these images is easier and more reliable. Furthermore, separating the image by this method will also greatly benefit further object classification.

## ACKNOWLEDGMENT

The authors would like to thank the Engineering and Physical Sciences Research Council (UK), the Department of Trade and Industrial (UK), and Jaguar Cars Limited for their sponsorship and enormous supports throughout the project.

## REFERENCES

- [1] T. Suzuki and T. Kanade, "Measurement of vehicle motion and orientation using optical flow," in *Proc. IEEE/IEE/JSAI International Conference on Intelligent Transportation Systems (ITSC '99)*, pp. 25–30, Tokyo, Japan, October 1999.

- [2] Z. Hu and K. Uchimura, "Tracking cycle: a new concept for simultaneous tracking of multiple moving objects in a typical traffic scene," in *Proc. IEEE Intelligent Vehicles Symposium (IV '00)*, pp. 233–239, Dearborn, Mich, USA, October 2000.
- [3] N. Ancona, "A fast obstacle detection method based on optical flow," in *Proc. European Conference on Computer Vision (ECCV '92)*, pp. 267–271, Santa Margherita Ligure, Italy, May 1992.
- [4] D. Willersinn and W. Enkelmann, "Robust obstacle detection and tracking by motion analysis," in *Proc. IEEE Conference on Intelligent Transportation System (ITSC '97)*, pp. 717–722, Boston, Mass, USA, November 1997.
- [5] U. Franke and A. Joos, "Real-time stereo vision for urban traffic scene understanding," in *Proc. IEEE Intelligent Vehicles Symposium (IV '00)*, pp. 273–278, Dearborn, Mich, USA, October 2000.
- [6] U. Franke and S. Heinrich, "Fast obstacle detection for urban traffic situations," *IEEE Trans. Intell. Transport. Syst.*, vol. 3, no. 3, pp. 173–181, 2002.
- [7] S. Bohrer, T. Zielke, and V. Freiburg, "An integrated obstacle detection framework for intelligent cruise control on motorways," in *Proc. IEEE Intelligent Vehicles Symposium (IV '95)*, pp. 276–281, Detroit, Mich, USA, September 1995.
- [8] M. Bertozzi and A. Broggi, "GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. Image Processing*, vol. 7, no. 1, pp. 62–81, 1998.
- [9] M. Bertozzi, A. Broggi, and A. Fascioli, "Stereo vision-based vehicle detection," in *Proc. IEEE Intelligent Vehicles Symposium (IV '00)*, pp. 39–44, Dearborn, Mich, USA, October 2000.
- [10] M. Betke, E. Haritaoglu, and L. S. Davis, "Multiple vehicle detection and tracking in hard realtime," in *Proc. IEEE Intelligent Vehicles Symposium (IV '96)*, pp. 351–356, Tokyo, Japan, September 1996.
- [11] T. Kalinke, C. Tzomakas, and W. V. Seelen, "A Texture-based object detection and an adaptive model-based classification," in *Proc. IEEE Intelligent Vehicles Symposium (IV '98)*, vol. 1, pp. 143–148, Stuttgart, Germany, October 1998.
- [12] Y. Fang and I. Masaki, "Depth-based target segmentation for intelligent vehicles: fusion of radar and binocular stereo," *IEEE Trans. Intell. Transport. Syst.*, vol. 3, no. 3, pp. 196–202, 2002.
- [13] W. E. L. Grimson, "Computational experiments with a feature based stereo algorithm," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 7, no. 1, pp. 17–34, 1985.
- [14] T. Kanade and M. Okutomi, "A stereo matching algorithm with an adaptive window: theory and experiment," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, no. 9, pp. 920–932, 1994.
- [15] T. Kanade, A. Yoshida, K. Oda, H. Kano, and M. Tanaka, "A stereo machine for video-rate dense depth mapping and its new applications," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '96)*, pp. 196–202, San Francisco, Calif, USA, June 1996.

Ericsson Mobile Communication (UK) Ltd. During 2002–2004, he worked as a Research Fellow for the Centre for Computer Vision and Image Processing, Cranfield University. Currently, he is a Research Scientist at the International Automotive Research Centre, the University of Warwick. His research interests include image processing for intelligent vehicles, vision-based instrumentation, and artificial intelligence for automobile diagnostics.

**Shan Fu** is the Director of the Image Processing Laboratory, School of Engineering, Cranfield University in UK. He received a B.Eng. degree in electronic engineering from Northwestern Polytechnic University and a Ph.D. degree from Heriot-Watt University. All his research relates to computer vision, image processing, and related system development.

**Chris Thompson** is a Professor of applied computation currently and the Head of the Applied Mathematics and Computing Group, Cranfield University. After leaving Oxford University with a B.A. degree in mathematics and an M.S. degree in numerical analysis, he joined the Applied Maths Group, AEA Technology, Harwell. He has also worked at the Advanced Computer Research Facility, Argonne National Laboratory, Chicago, and at Bergen Scientific Centre, IBM. He holds a D.Phil. degree from the University of Bergen. He joined Cranfield University in 1994. His current research interests include image processing for automotive applications; computational fluid dynamics, and the application of parallel computers and the development of software tools for large-scale problems. He has written over sixty papers in these areas.

**Yingping Huang** received the Ph.D. degree in electronics engineering from the University of Greenwich in 2001, the M.S. degree in precision instrumentation engineering from Shanghai University of Science and Technology in 1990, and the B.S. degree from Tianjing University in 1987. From September 1990 to December 1996, he was working as an R&D Engineer for East China Electric Power Test and Research Institute. During 2000–2001, he was an R&D Engineer for

