The Obstacle Detection and Measurement Based on Machine Vision

Xitao Zheng[1], Shiming Wang[2], Yongwei Zhang[1]

[1]College of IT, Shanghai Ocean University, Shanghai 201306, China

[2]College of Engineering Science and Technology, Shanghai Ocean University, Shanghai 201306, China

xtzheng@shou.edu.cn, smwang@shou.edu.cn, zhangyongwei_108@163.com

Abstract - **To develop a quick obstacle detection and measurement algorithm for the image-based autonomous vehicle (AV) or computer assisted driving system, this paper utilize the previous work of object detection to get the position of an obstacle and refocus windows on the selected target. Further calculation based on single camera will give the detailed measurement of the object, like the height, the distance to the vehicle, and possibly the width. It adopts a two camera system with different pitch angles, which can perform real-time monitoring for the front area of the vehicle with different coverage. This paper assumes that the vehicle will move at an even speed on a flat road, cameras will sample images at a given rate and the images will be analyzed simultaneously. Focus will be on the virtual window area of the image which is proved to be related to the distance to the object and speed of the vehicle. Counting of the blackened virtual sub-area can quickly find the existence of an obstacle and the obstacle area will be cut to get the interested parameter measurements for the object evaluation.**

Index Terms - obstacle detection; object measurement, ALV, virtual window.

I. INTRODUCTION

Autonomous land vehicles (ALVs) are useful for many automation applications that serve for both indoor and outdoor environments. Vision or digital image based obstacle detection for ALV navigation in outdoor road environments is a difficult and challenging task because of the great variety of object and road conditions, like irregular and unstable features on objects, moving objects, changes of illumination, and even rain. Successful ALV navigation requires the integration of the many techniques, like the environmental sensing and learning, image processing and feature extraction, ALV location retrieving, travel path planning, vehicle wheel controlling, and so on.

Vehicle intelligent drive system is also an important part of the intelligent transport system. Based on our driving experience, more than 90% of our information is from our vision and hearing, so using computer vision technique to solve this problem is a challenging work for every researcher in this field. The lane keeping and distance measurement based on the computer vision technique are the key problems of the auxiliary drive system.

In reference [1], a real-time distance detection method is proposed to get the real-time distance between the camera car and the obstacle in the front. The paper uses the geometrical reasoning method to obtain the depth information. The shoot angle of the camera is the key factor that affects the accuracy of output. The work uses the two lane boundaries as a constraint to get the pitch angle. The problem is that the results are for static car tests, and algorithm needs to be combined with others to get all the required parameters.

Reference [2] presents an obstacle detection method for AGV under indoor or outdoor environment. Corner or turn points are detected and tracked through an image sequencing camera and grouped into ground region using a method that is called co-planarity checking algorithm. The algorithm is initialized by 5-point planar projective and contour cue which is added to segment the ground region. The method can be applied to some ideal environment and the contour matching can be a problem in a lot of environments.

Reference [3] is the previous work of this paper and it provides the quick detection of the existence of an obstacle, this is a key feature for ALV's moving where safety is the highest requirement. The frames can be analyzed in a fragment of milliseconds and alert can be sent if an obstacle is found. But sometime the vehicle does not need to stop to wait for a clear road, it need to further analyze the situation and the surrounding environment to make decision whether to slow down or to stop, this part of job is not included in [3].

Reference [4] proposes an effective approach to obstacle detection and avoidance for ALV navigation in outdoor road environments using computer vision and image sequence techniques. To judge whether an object newly appearing in the image is an obstacle or not, the object shape boundary is first extracted from the image. After the translation from the ALV location in the current cycle to that in the next cycle, the position of the object shape in the image is predicted, using coordinate transformation techniques based on the assumption that the height of the object is zero. The predicted object shape is then matched with the extracted shape of the object in the image of the next cycle to decide whether the object is an obstacle. A reasonable distance measure is used to compute the correlation measure between two shapes for shape matching. Finally, a safe navigation point is determined, and a turn angle is computed to

Author:Xitao Zheng, Professor. Research Field: Pattern Recognition, Artificial Intelligence, etc. Email:xtzheng@shou.edu.cn

Corresponding Author: Shiming Wang, Professor, Dean of Engineering Science and Technology College.

Email:smwang@shou.edu.cn.

guide the ALV toward the navigation point for obstacle avoidance. This is one camera based obstacle detection system and the moving speed is relatively slow, the work on the corner sensing can be applied in this paper.

Reference [5] tries to establish restriction functions for lane lines tracks and reconstructs the lane line to get the necessary parameters. There is binocular camera to guide the vehicle moving and another camera on the roadside to trajectory comparison. Reference [6] proposes a function that can use a monocular camera to detect the lane edge and keep the vehicle in the lane. Reference [7] works on stereo color image with one camera and a lens. The job is based on the assumption that all the brake lights are working and can be clearly pictured. It is a good paper for car-following model which is sensitive to the emerging of the brake signal, i.e., when the traffic is slowing down. It will lose target if one brake light is out.

There are a lot of papers that involve the lane detection and obstacle avoidance for ALV navigation in indoor or outdoor road environments using computer vision and image sequence techniques [4, 5, 8]. These methods need to search the whole pictured areas and feature matching can be unreliable.

Reference [9] addresses the problem of visual object class recognition and localization in natural images. Field experiences show how histogram-based image descriptors can be combined with a boosting classifier to provide a state of the art object detector. Among the improvements the paper introduces a weak learner for multi-valued histogram features and shows how to overcome problems of limited training sets.

Reference [10] summarizes the past work on object detection of feature extraction and classi1cation, proposes that feature selection is an important problem in object detection and demonstrates that genetic algorithms (GAs) provide a simple, general, and powerful framework for selecting good subsets of features, leading to improved detection rates.

Reference [11] proposes an object detection approach using spatial histogram features. As spatial histograms consist of marginal distributions of an image over local patches, they can preserve texture and shape information of an object simultaneously. Their work employs Fisher criterion and mutual information to measure the discriminability and features correlation of spatial histogram features. They further train a hierarchical classifier by combining cascade histogram matching and support vector machine. The cascade histogram matching is trained via automatically selected discriminative features. A forward sequential selection method is presented to construct uncorrelated and discriminative feature sets for support vector machine classification.

Reference [12] establishes a system that can detect moving object in the 3-D space and propose a background reduction method. The camera can be repositioned to track the moving of a target.

As we know all these works are confined in a specific environment and the methods proposed above can be used in a certain circumstances.

In this paper we will utilize a virtual area-based object detection rule to detect the height of an obstacle and the combination of the two camera images will be able to confirm the movement of the object over the picture; then use the location to retrieve the object area for further analysis. So the work is correlated to reference [3] and some problems of reference [3] are resolved. The system can be very efficient because the object area can be very small compared with the whole image and the image center is nearly found when the obstacle detection is done. The further analysis will tell the driver whether to wind around or to apply brake action.

.

II.   PROPOSED OBJECT HEIGHT DETECTION METHOD
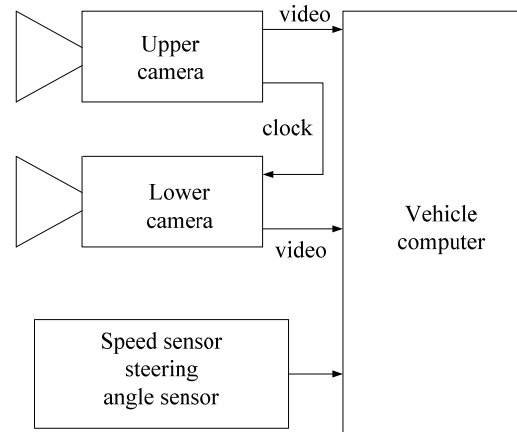
A.   System Architecture

Fig 1 The structure diagram of target detection

The system hardware architecture can be shown in Fig 1. The speed sensor is used to get the vehicle speed to do frame differentiation. The speed will be used to determine the height of the smaller object, like a cone or a deer. Lower camera can be viewed as collision prevention precaution and precise distance calculation camera; upper one can be used as lane detection, object pre-alerting, object height analysis, and observation camera.

From Fig 2, we can see that each camera has different coverage and they will overlap in a considerable range. In Fig 3, we can see that the same virtual area will be in different position if we merge the two pictures. This will offer ID verification and approaching information of the object.
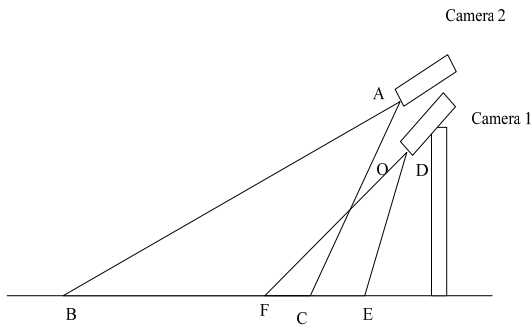
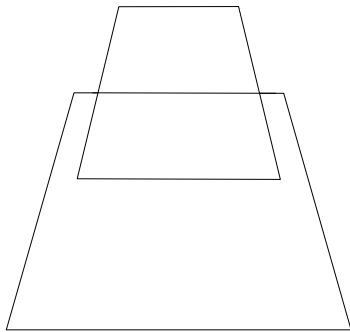Fig 2 The covering ranges of the two cameras

Fig 3 Overlapping of virtual area between two cameras

B.   Virtual Window(s)

As we discussed above, the virtual windows are the only areas that we will be interested, which can be seen in Fig 9-11 as the gridded area. These windows can be dynamically added or deleted depending on the detected road condition. In this model, we will assign the area from 6-11 meters in front of the vehicle as the virtual grid, which is usually used for low-speed drive. Usually multiple windows should be assigned to lower window for smaller object detection and lane detection and larger object detection windows should be assigned to the upper camera. Grid shown in Fig 4 can be an example for slow country road driving window setting.
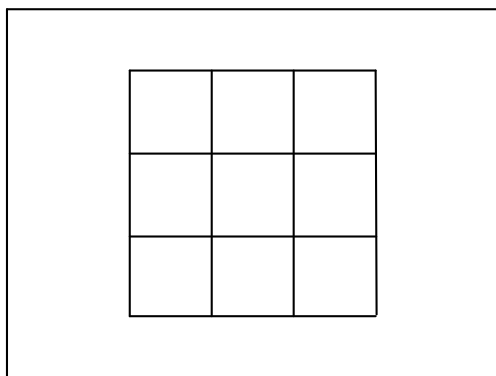

Fig 4 Small object detection: multiple windows

Virtual windows can be placed at different positions to serve as the information window, attention window, and action window. Here for the obstacle height detection purpose, we will use the same number and size of windows for both cameras. The accurate mapping and

quick resolution of the virtual windows on the picture re resolved by our previous work.
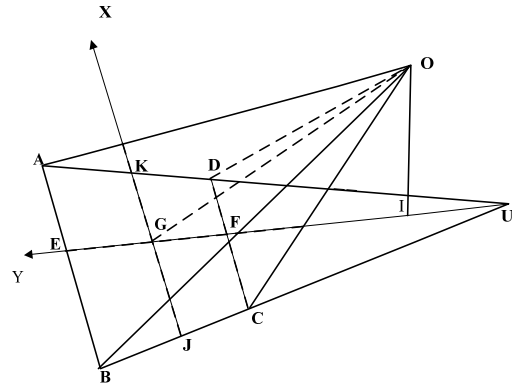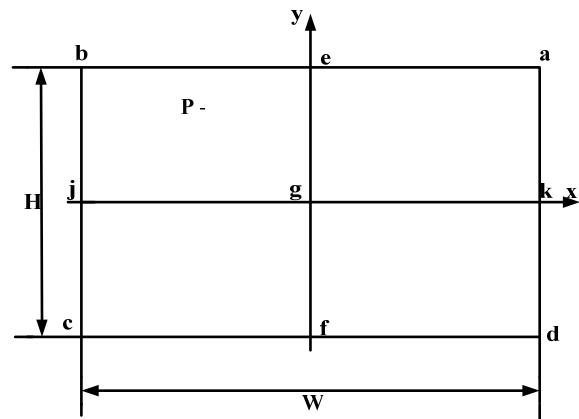

Fig 5 The projected plane X-Y

Fig 6 The original rectangle on the road

In reference [1], the relationship between the actual road coordinate and the projected coordinate, which is the picture, can be shown in Fig 5 and 6, the conversion formula between the two coordinate can be seen in formula (1, 2). Detailed definitions of the subscripts can be found in [1].

$$
\begin{cases}
Y_p = h \cdot k_1 \cdot y_p \cdot \dfrac{1 + k_2{}^2}{1 - k_2 \cdot k_1 \cdot y_\mu} \\[2mm]
X_P = \dfrac{UG + Y_P}{UG \cdot k_3 \cdot x_p \cdot k_4} \\[2mm]
y_p = \dfrac{Y_p/k_1}{h + h \cdot k_2{}^2 + Y_P \cdot k_2} \\[2mm]
x_p = \dfrac{UG \cdot X_P}{k_3 \cdot k_4 \cdot (UG + Y_P)}
\end{cases}
\quad (1)
$$

$$
\begin{cases}
k_1 = 2tg/H(\alpha_0) \\
k_2 = tg(\gamma_0) \\
k_3 = h/\cos(\gamma_0) \\
k_4 = 2tg(\beta_0)/W \\
UG = \dfrac{h \cdot (tg(y_0) - tg(y_0 - \alpha))\cos(\gamma_0 - \alpha_0)}{\cos(\gamma_0 - \alpha_0) - \cos(\gamma_0)}
\end{cases}
\quad (2)
$$

Using formula (1) and (2), we can convert a pre-designated area in front of the car to a virtual area in the picture and vise versa.

C. Object Height and Width Analysis

Each virtual window will be counted for the pixels occupied, if the occupied percentage is more than 50%, this window will be assigned black, otherwise white. The black and white windows will be compared against the other camera, the current frame situation will compared again with the following frame, the one with changing black spots will be the obstacle and the numbers of the windows will be the picture height, which can be converted to real height using formula (1) and (2). As seen in Fig 7, the object with a height FS is projected to the ground segment St, ST can be further projected to image size as shown in Fig 8. OE is the height of the camera.

In order to implement our detection algorithm, we find that most road image areas have similar RBG values, so the pixels with this characteristics and the area with even color range are reset to white color. This road information needs to be inputted if different road base color is present. The painting and signs on the road will not affect road color selection. In our samples the asphalt road and stone road are tested and they are a good match to the RGB rule.
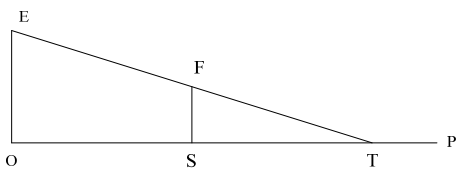

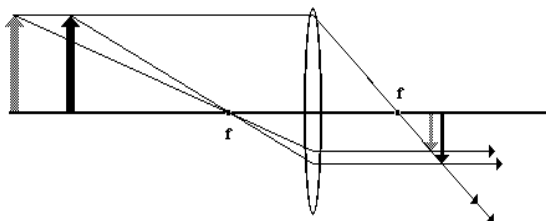Fig 7 Projection of image length ST to obstacle height FS
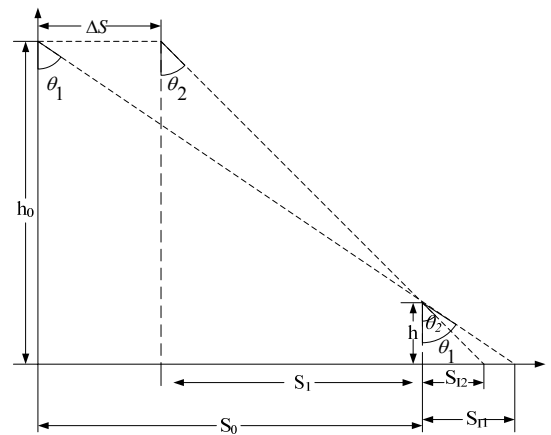

Fig 8 Image formation on camera

Fig 9 The view of the observer from different positions

$$
\begin{cases}
\Delta S = S_0 - S_1 \\
h_0 \cdot \tan\theta_1 = S_0 + S_{I_1} \\
h_0 \cdot \tan\theta_2 = S_1 + S_{I_2} \\
h = S_{I_1} \cdot ctg\,\theta_1 \\
h = S_{I_2} \cdot ctg\,\theta_2
\end{cases}
\quad (3)
$$

Applying Fig 7 to a moving condition, Fig 9 can be drew to show the height of an object in the observer's eyes, i.e., the shadow length. In formula (3), ΔS is the distance moved in the given time interval, S_0 is the initial distance to the object, S_1 is the distance to the object at the next time interval, h_0 is the height of the cameras, θ_1 and θ_2 are the pitch angles at the two time intervals, h is the height of the object, S_{I1}, S_{I2} are the viewed object lengths, need to be converted to the picture length using Fig 8, which is very typical for optical camera imaging system.

When an obstacle is found and so an object is located, formula (3) can be used to retrieve the interested height data. There are five variables in formula (3), not include θ_1 and θ_2. Where ΔS is known if the speed and the time interval is known, and the height of the camera, which is h_0, is also known. So formula will be reduced to an equation of θ_1 and θ_2, which are unknown because of our moving positions. Note that we have two cameras that are positioned to the same level and symmetrically to the center line, so all the parameters in formula (3) are shared for two cameras, this triangle will give another equation for θ_1 and θ_2 to get their solution, therefore the solution of all the parameters. Note that we can only get pixels from the picture so the image height and width will be counted from black white small picture to get their pixels, which

can be converted to actual shadow distances S_{I1}, S_{I2} using Fig8. We have data redundancy so we can cross verify the effectiveness of image edge detection algorithms.

III. EXPERIMENT

The experiment is done with a vehicle moving at 25km/h, the two cameras are horizontally positioned at the front of the vehicle and closely together, so in our model calculation the two cameras are considered at the same vertical location, the upper camera (which has a bigger pitch angle, or have a farther view, not actually higher in position) serves as the information and alert window and the lower one, which has a smaller view area but better focused, serves as the verification window. Both cameras are sampling at 10frame/s, calculation of 5 sample frames pictures are presented here from Fig 10 to Fig 19. Each frame is only processed against its sister frame and the following one or two frame. Calculations are focused on virtual windows and the half processed windows are shown in Fig 20-23, the grids are added to show the virtual window arrangement.

As we already know, the virtual window will be positioned to interested area, where a possible object will appear. If the object is confirmed, which is done by the counting of blackened boxes (data shown in Tab 1) and compared with the previous frames, an object window will appear to cut the object off from the original picture and to do image processing, like the image format conversion and edge detection, pixel data can be found and compared with previous frame for stableness evaluation. Tab 2 is the data for this experiment.
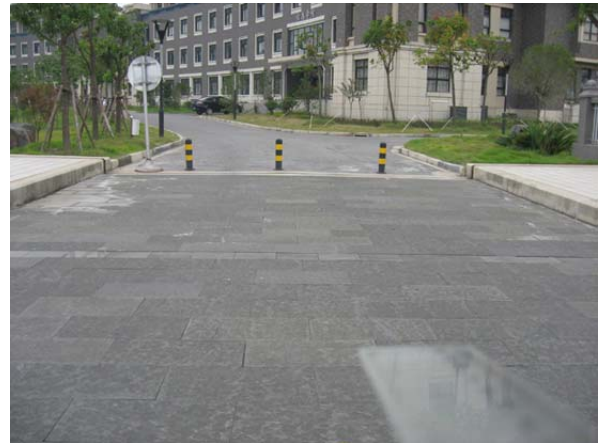

Fig11 Frame 1 image from camera 2

Fig 12 Frame 2 image from camera 1

Fig13 Frame 2 image from camera 2

Fig 10 Frame 1 image from camera 1

Fig 14 Grid image of frame1 from camera 1

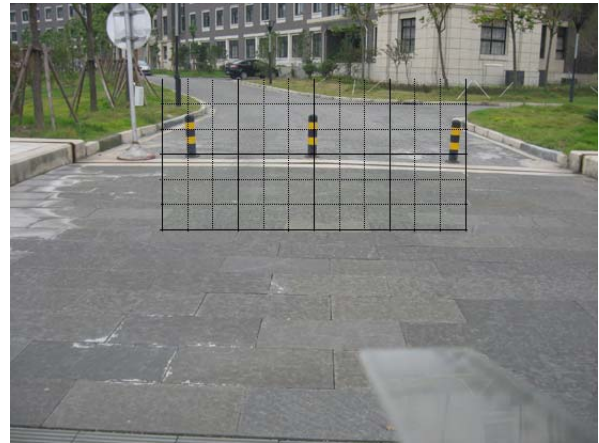Fig 15 Grid image of frame1 from camera 2
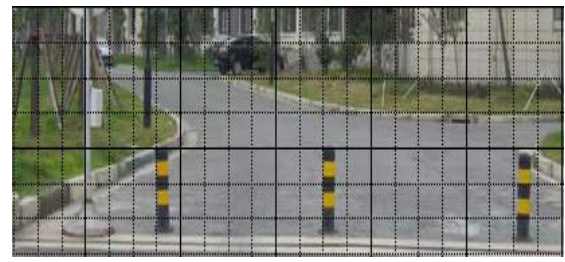

Fig 16 Grid image of frame2 from camera 1

Fig 17 Grid image of frame2 from camera 2

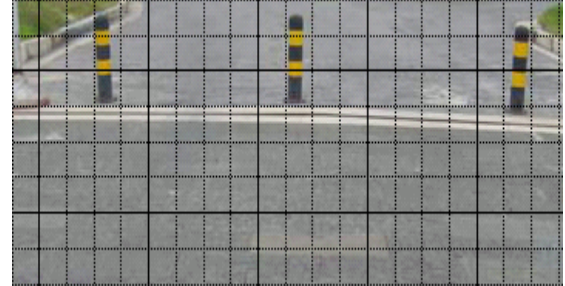Fig 18 Virtual windows isolated of frame1 (camera1)

Fig 19 Virtual windows isolated of frame1 (camera2)

It can be seen from Fig 10 to Fig 17 that the upper camera (camera 1) views farther and the lower one(camera 2) views closer. As the vehicle approaches the obstacles become bigger in the pictures. Figure 14 to Figure 17 are applied to intelligent grid. Fig 18 and Fig 19 are the separated virtual windows.
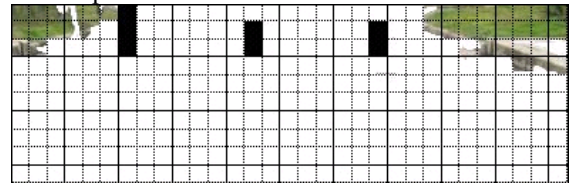

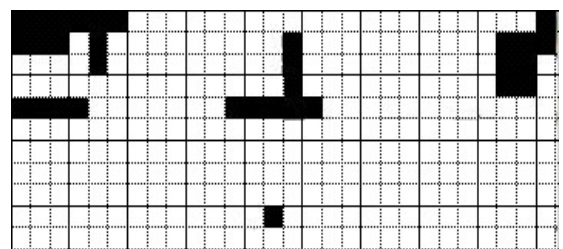Fig 20 Processed windows isolated of frame1 (camera1)

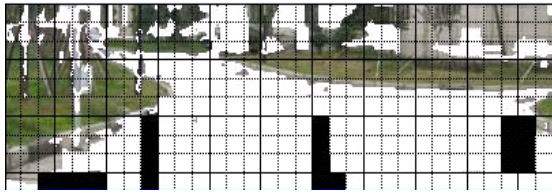Fig 21 Processed windows isolated of frame3 (camera1)

Fig 22 Processed windows isolated of frame3 (camera2)

Fig 20 to Fig 22 are the processed virtual windows to count black boxes for obstacle detection. Fig 23-25 are the height and width analysis figure where the width and height of the object are measured.
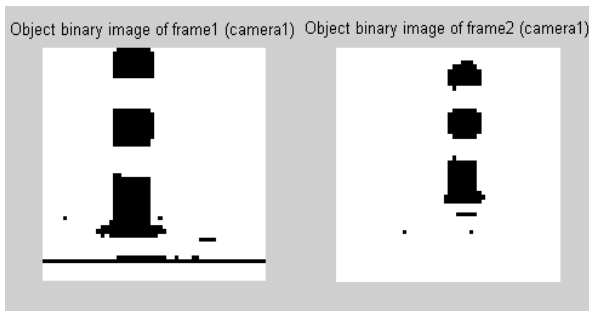

Fig 23 Middle object binary images of frame1 and frame2 (camera1)

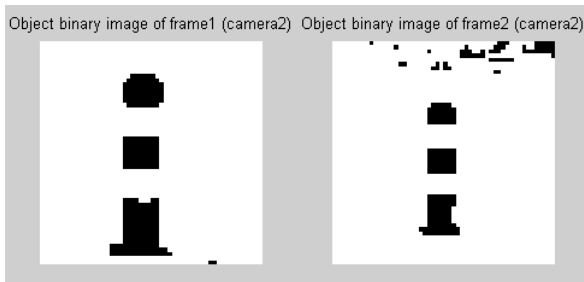Fig 24 Difference image of middle object of frame1 and frame2 (camera1)

Fig 25 Middle object binary images of frame1 and frame2 (camera2)

We can see the difference of the two objects in Fig 23-25, their height, width, the pixels occupied are all very

different, this is another verification of the effectiveness of obstacle detection algorithm.

So by counting the vertical continuous black spots, the difference prompts an obstacle, and by counting the pixels to the top and bottom of the object, we get the height of the object. If we apply the counting the left and right edge of the object window, we get the width. Applying formula (3), we have Tab 2. We will not elaborate the process to get the pixels length of height and width, because this is the known way to process the digital image.

This quick estimation and multiple verification mechanism will avoid the miscalculation and ensure the driving safety.

Tab 1 is the result of obstacle detection, where the black box numbers are presented for each frames and for each cameras. Tab 2 is the data from each object windows, and then converted by formula (3). The test run proves the two step method is effective. We can also see that the closer image intends to give bigger width which is still within our error tolerance level.

The current method can not resolve the case when the object has similar color to the road. It also proved that when the weather condition is bad, like in the rain or in the night, the system will not be effective. This reason is that we do not apply a good background reduction algorithm, and the complicated illumination condition will considerably affect the detection and analysis. If possible, we will try to research these issues in our later project.

Table 1 Camera 1 and Camera 2 frames

90^0 pitch angle	Fr1	Fr2	Fr3	Fr4	Fr5
camera height(m)	1.035	1.035	1.035	1.035	1.035
Black Boxes	2	4	4	5	7
77.8^0 pitch angle	Fr1	Fr2	Fr3	Fr4	Fr5
camera height(m)	1.035	1.035	1.035	1.035	1.035
Black Boxes	2	4	4	5	8

Table 2 Object Height and Width

	Fr1 pitch angle (87.4^0)	Fr2 pitch angle(86.9^0)
camera height(m)	1.035	1.035
Obj dist(m)	11.83	9.75
Obj Ht (pxl)	33	44
Obj Ht(m)	0.48	0.50
Obj Width(pxl)	8	11
Obj Width(m)	0.12	0.13

IV. CONCLUSION

This method presents a method to detector locate an obstacle, retrieve the window that cover the obstacle and then use digital image processing technique to get the detailed measurement of the object. The measurement project the object to a shadow on the background and

then project the shadow to the those in the camera. This is an important effort to facilitate the ALV for its decision making process. Usually the two cameras have the same pitch angle and are symmetric to the center, which can not benefit from our obstacle detection with far view and close view. The structure and algorithm need to be improved in later work for easier projection calculation and enable full object detection, not just the two dimensions in this paper. Pattern recognition and neural network need to be introduced to help the learning process. But because of the hardware architecture relative feature of the system, the learning experience and knowledge base may not be able to be shared like these geometrical algorithms. Also, most of the current works are in the day light, the algorithm needs to be tested modified on different weather conditions to ensure the robustness for actual applications.

REFERENCES

[1]   Lei Guo, Youchun Xu, Keqiang Li, Xiaomin Lian. "Study on Real-time Distance Detection Based on Monocular Vision Technique". *Journal of Image and Graphics*. Vol 11, No.1, Jan, 2006.(in Chinese)

[2]   Yong Zhou, Chao Chen, Qingtai Ye. "Obstacle Detection for AGVS Based on Computer Vision". *Machine Design and Research*. Vol.21, No. 5, Oct, 2005.(in Chinese)

[3]   Xitao Zheng, Yongwei Zhang. "An Obstacle Detection Method Based on Machine Vision". *The Chinese Symposium on Information Science and Technology*. 2010, page993-996.

[4]   Kuang-Hsiung Chen, Wen-Hsiang Tsai. "Vision based obstacle detection and avoidance for autonomous land vehicle navigation in outdoor roads". *Automation in Construction*, Volume 10, Issue 1, November 2000, Pages 1-25.

[5]   H.-H. Nagel, F. Heimes, K. Fleischer, M. Haag, H. Leuck, S. Noltemeier. "Quantitative comparison between trajectory estimates obtained from a binocular camera setup within a moving road vehicle and from the outside by a stationary monocular camera". *Image and Vision Computing*, Volume 18, Issue 5, April 2000, Pages 435-444.

[6]   Xin Zhou, Xiyue Huang, Yu Li. "Lane Keeping and Distance Measurement Based on Monocular Vision". *Journal of Image and Graphics*. Vol. 8(A), No.5. May, 2003. (in Chinese)

[7]   Joaquín Ferruz, Aníbal Ollero. "Integrated real-time vision system for vehicle control in non-structured environments". *Engineering Applications of Artificial Intelligence*, Volume 13, Issue 3, 1 June 2000, pages 215-236.

[8]   T Kato and Y Ninomiya. "An approach to vehicle recognition using supervised learning". *Proc. of Machine Vision Applications* [C]. 1998, 77～80.

[9]   Ivan Laptev. "Improving object detection with boosted histograms". *Image and Vision Computing* . 27 (2009), 535–544.

[10] Zehang Sun, George Bebis, Ronald Miller. "Object detection using feature subset selection". *Pattern Recognition*. 37 (2004), 2165 – 2176.

[11] Hongming Zhang, Wen Gao, Xilin Chen, Debin Zhao. "Object detection using spatial histogram features". *Image and Vision Computing*. 24 (2006) 327–341.

[12] Kun Zhu , Tangwen Yang , Qiuqi Ruan , Hongbo Wang , Jianda Han. "Real-Time Tracking and Measuring of Moving Objects Based OR Binocular Vision". *Robot*. Vol, 31. No. 4. July, 2009. (in chinese)