WIDE-SCOPED AROUND VIEW MONITOR

[1]YU-CHUNGKUO, [2]DIN-CHANG TSENG

[1,2] Institute of Computer Science and Information Engineering, National Central University,
Jhongli, Taoyuan City 32001, Taiwan
Email: [1]yuchung.kuo@gmail.com, [2]tsengdc@gmail.com

Abstract - In recent years, around-view monitoring systems are becoming a practical driving aid that help reducing collision hazards by eliminating blind spots. Many of such systems provide short range views surrounding the vehicle, limiting its application to parking and reversing. In this paper, we propose a practical system for creating a wide-scoped around-view imagery surround the vehicle. We use four cameras mounted on four sides of a vehicle to capture the surrounding images; these images are then processed and projected on a dual-camber model centered by the vehicle. The projected imagery gives drivers the freedom to change view-point to suit different driving needs.

Index terms - Advanced driver assistance system, around view monitor, 3D visualization, Wide-scoped surveillance

I. INTRODUCTION

Following the advancement of science and technology, improving drive safety has become of higher concern. Among the causes of traffic accidents, one main cause is due to limited field of view of the driver. The driver is unable to acknowledge the entire area surrounding the vehicle, pay attention of different directions at the same time. In order to monitor the vehicle's surrounding; cameras are mounted around the vehicle to capture images and provide the driver with a visual monitor. Nissan [1]developed a parking assistance system called "Around View Monitor" as shown in Fig.1. The system uses four wide-angle cameras mounted at front, sides, and rear of the vehicle. Capturing images from the surrounding area and provides the driver with a bird's eye view of the vehicle and its surrounding. Honda [2] has also developed a similar system called Multi-view camera system.

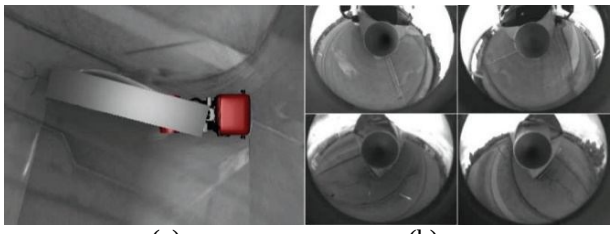
Fig.1.Nissan Around View Monitor system with top-view and rear view display.

Ehlgenet al. [3] provided a parking assistance system composed of two catadioptric cameras as shown in Fig.2 (a) and (b). The cameras have been mounted on the rear left and right top edge of the vehicle to create a bird's-eye view of the area behind the vehicle as well as the area of the sides. To be able to extend to a larger visible area, their study has stretched the bird's-eye view image that provides a larger field of view, as shown in Fig.2 (c) and (d).

(a) (b) (c) (d)

Fig.2.Omnidirectional cameras for backing-up aid. (a) Two cameras mounted on the test vehicle. (b) A catadioptric camera. (c) Standard bird's-eye view. (d) Extended view.

Ehlgen and Pajdla [4] developed a bird's-eye view surrounding monitoring system that is suitable for large vehicles such as buses, trucks and tractor-trailer combinative vehicles. Their system consists of four omnidirectional cameras mounted on the tractor and trailer. The magnetic sensor is used to measure the kink angle between tractor and trailer. Depending on the direction of magnetic field of the sensor, the overlap of bird's-eye view image between tractor and trailer is corrected, as shown in Fig.3.



(a) (b)

Fig.3.Bird's-eye view image. (a) The kink angle between tractor and trailer is -14°. (b) Raw images from the omnidirectional cameras.

There are many algorithms for detection with moving cameras. Bertozzi and Broggi [5] proposed a lane and object detection system using inverse projection transformation. It first assumes ground as a flat surface, through inverse projection transforms pair of input images into top-view images, then detect lanes within the top-view image. Bertozzi and Broggi combined the top-view images by subtraction; parts with high difference levels are where objects are located. Bertozziet al. [6]set up two cameras on a

truck and by using grid points on the ground to inverse project the images to top-view images, depicted in Fig.4.Then polar histogram is used to analyze the difference of two images to check for objects perpendicular to the ground.

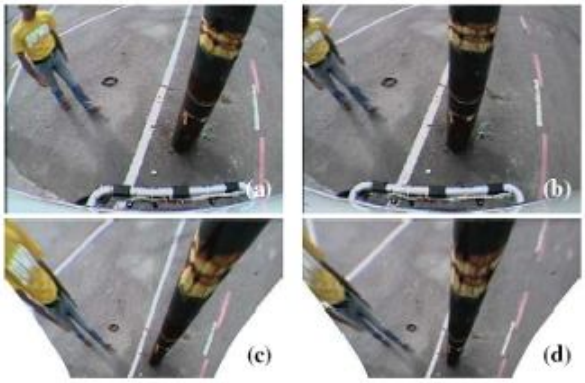


Fig.4.Top-view image. (a) Left camera image. (b) Right camera image. (c) Left camera top-view image. (d) Right camera top-view image.

In this study, we propose a driving assistance system. The proposed system employs four wide-view cameras mounted at front, rear, and both sides of the vehicle to capture images of vehicle's surrounding. After a sequential transformation and combination, the system provides a bowl-shaped wide-scale view of vehicle's surroundings, giving driver a full view around the vehicle. The system is comprised of three parts: offline calibration, combined calibration, and online wide-scoped surrounding top-view monitoring. Offline calibration process includes the calibration of the camera parameters, image and distortion parameters. The combined calibration includes top-view transformation, image registration, and projection to bowl shaped model. With the offline calibration and transformation process, we create a mapping that maps input images to the wide-scoped view on the bowl shaped model. Because the cameras' positions are fixed on the vehicle, the geometric relationship between cameras and the vehicle does not change, the mapping created at offline calibration process can be used repeatedly during the online phase. Wide-scoped around view monitor (WAVM) takes the input images and projects them onto the bowl shaped model according to the mapping produced during offline process. The projection includes interpolation and color blending. The remaining sections are arranged as follows. In Section II, we describe the transformation and creation of the mapping from input to bowl shaped model. Experiments and results are presented in Section III, with conclusions and future works in Section IV.

II. WIDE-SCOPED AROUND-VIEW MONITOR

A top-view monitoring system provides the driver with visual assistant to area surrounding the vehicle, reducing the chance of possible collisions. The integration of such technology on vehicles has been proliferating. Mainstream surrounding top-view monitoring systems [7, 8] creates a bird's-eye view of the vehicle's surrounding. Such systems have the drawback of limited viewing distance, for driver can only observe near stationary obstructions near the vehicle rather than identifying on-coming objects that moves relatively to the vehicle. A wide-scoped surrounding top-view monitoring system provides a broader visibility, giving the driver more information about his or her surroundings. We here describe the offline transformation process used to construct a wide-scoped surrounding top-view monitoring system that can take advantage of graphics hardware which is increasingly used in vehicles [9, 10].

A. Dual-camber modeling

The wide-scope surrounding top-view imagery is created by projecting input frames onto a 3-D semispherical model. Because curved surface distorts the ground texture which could influence driver's judgment, the 3-D model is constructed as a flat bottom dual-chamber model. The dual-camber model is defined as a function M of radius from z-axis r and angle θ from positive x-axis

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = M(r,\theta), \quad (1)$$

With x, y, and z components defined as

$$x = M_x(r, \theta) = r\cos\theta, \quad (2)$$

$$y = M_y(r,\theta) = \frac{L}{w} r\sin\theta, \quad (3)$$

and

$$z = M_z(r,\theta) = \begin{cases} 0, & \text{if } r^2 < r_i^2 \\ -(1-\cos(r^2 - r_i^2))\, r_0, & \text{if } r_i^2 < r^2 < (r_i + r_0)^2 \end{cases}, \quad (4)$$

where r_i is radius of inner circle, and r_o is the radius of the curved surface. The curved surface has the same gradient as inner circle at $r = r_i$, resulting a smooth transition. Fig.5 shows the cross section of the model.

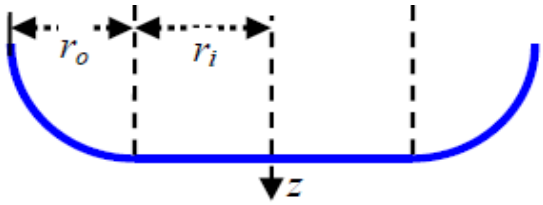


Fig.5. Cross section of dual-camber model.

B. Image registration

For a surround top-view to produce a continuous imagery from four cameras, the relative positions of each camera in world space must be known. The camera position and orientation is calculated from corresponding feature points in world space and image. Once camera position and orientation is known the images from cameras can be transformed to from a seamless view. In this section we describe

the process to automatically perform image registration. For wide-scoped surrounding top-view monitoring, we define a right-handed coordinate system as world coordinate system. The coordinate system is defined that center of vehicle is located at origin of x-y plane with $z_w = 0$ being the ground plane. x-axis is to the right, and y-axis is to the back of the vehicle. Since the cameras are fixed to the vehicle, they're at a fixed location and orientation in the world space. We define a grid on ground surface that is of a checkered board pattern with grid coordinates in the same direction as WCS. At each side around the vehicle, there is one red block in the place of a black block. The red blocks' world coordinates and block's width are measured prior to calibration. In order to project the input frames of cameras onto the 3D model, we need to calculate the cameras' positions and orientations within WCS. The cameras' positions and orientations can be calculated using extrinsic parameters by

$$R_i c_w^i + t_i = c_c^i \;\Rightarrow\; c_w^i = R_i^{-1}(c_c^i - t_i), \qquad (5)$$

wherei is the camera index, c_w^i is the position of camera i in WCS, $c_c^i = [0\ 0\ 0]^T$ is the position of camera i in CCS. Substituting c_c^i in Eq.(5) with $[0\ 0\ 1]^T$ and $[0\ -1\ 0]^T$ gives us the optical axis vector l_w^i and up vector u_w^i of the camera. Right vector r_w^i is calculated as $r_w^i = l_w^i \times u_w^i$.

C. Model construction

We construct the dual-camber model as 3-D mesh, and input images from the cameras are used as textures. Each vertex in the model has five components.

i. Vertex coordinates
ii. Texture coordinates for camera A
iii. Texture coordinates for camera B
iv. Blending factor for camera A
v. Blending factor foe camera B

The dual-camber model is divided into four meshes, separated by x-z plane and y-z plane, mesh for each quadrant as shown in Fig.6 slightly separated here for viewing. Dividing it this way produces four meshes, each only overlapped by two cameras, reduces the texture coordinates needed for the vertices. The texture sampling method is set to clamp and border's alpha set to 0, meaning that any coordinates outside the texture boundary will be transparent. The model can also be split in other ways depending on the position of cameras, but the goal is to reduce the number of textures used in each vertex.

Chou [11] generated the wide-scoped surround view by first transform input images to a single image then projects it to the dual-camber model. To produce higher resolution output, the transformed image would need to have a much higher resolution, and still the two-step transformation inherently produce more rounding errors. Our method projects directly onto the dual-camber model, which eliminates the drawbacks mentioned above.

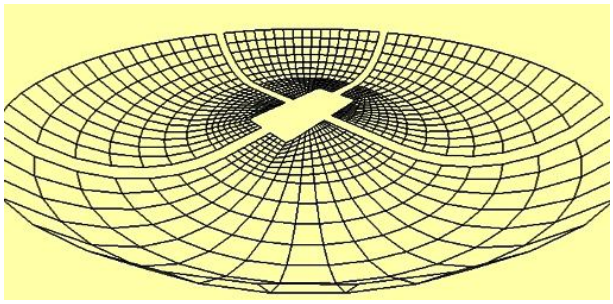


Fig.6.Dual-camber model divided into four meshes, one for each quadrant.

The dual-camber model's vertices are defined by slices around the z-axis, ring segments in base and ring segment in bowl shape as shown in Figure 6. Number of slices and segments can be adjusted for quality or performance.

The position of each vertex in the dual-camber model is calculated by following steps.

Step 1. For each ring segment in dual-camber model
Step 2. For each slice around z-axis
Step 3. Calculate r in Eq.(1)
Step 4. Generate vertex coordinate with Eq.(1)

The texture coordinates are generated per quadrant, because each quadrant uses different cameras as input. Texture coordinates are derived from vertex coordinates in following steps.

Step 1. For each quadrant Q in dual-camber model
Step 2. For each vertex v_w in Q
Step 3. For each cam c^i associated with Q
Step 4. Let p be $\overline{c^i v_w}$
Step 5. Calculate angle θ between p and l_w^i
Step 6. If θ> half the FOV angle of c^i, rotate p along the plane containing p and l_w^i so that θ= half the FOV angle.
Step 7. Intersect line c^i + xpwith c^i's imaging plane at p' where focal length f = 1
Step 8. Distort and transform p' to ICS as p" with intrinsic parameters A and distortion parameter ω.
Step 9. Calculate texture coordinates

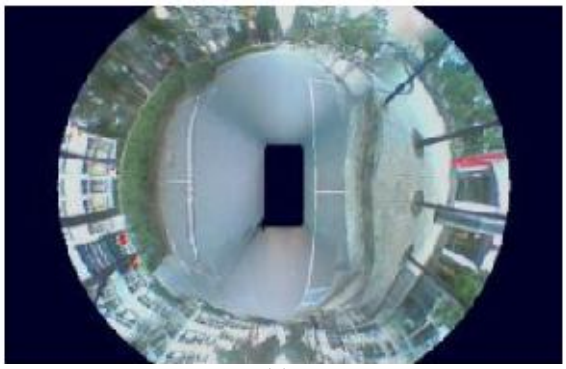
$$[u\ \ v]^T = \left[\frac{p''_x}{\text{image width}}\ \ \frac{p''_y}{\text{image height}}\right]^T.$$

To create a seamless surround view on the dual-camber model, the overlapped regions between cameras is blended. The blending factor is calculated for each camera in camera group $\{C^A, C^B\}$ for each vertex v in the model. The blending factor b^A for v in camera C^A is calculated as

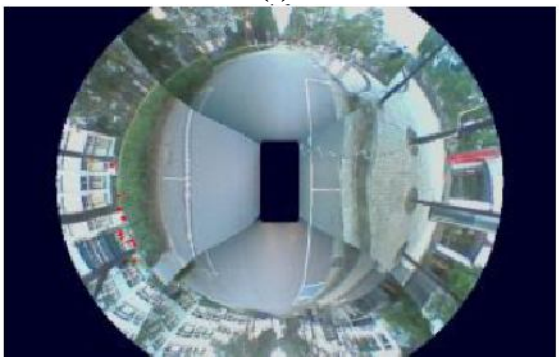
$$b^A = B(t^A, t^B, d^A, d^B) = \begin{cases} 0, \text{ if } t^A \notin T \\ 1, \text{ if } t^B \in T \text{ and } t^A \notin T \\ d^A/(d^A + d^B), \text{ otherwise} \end{cases} \qquad (6)$$

wheret ^{A} and t ^{B} are texture coordinates for each camera, d^A and d^B are minimum distance to the edge of texture edge bounded by $T = [T_x T_y]^T$, where $0 \leq T_x \leq 1$, $0 \leq T_y \leq 1$ The result of blending is shown in Fig.7.

Although color blending creates a seamless view, objects captured by cameras at different angles will only coincide if the object is located on the surface of the model. Objects at other distances would create a ghosting effect with blending as shown in Fig.8, which may influence the driver's understanding of surroundings. On the other hand, above-ground objects may be cropped as images of the object on the cameras are projected on to each other as shown in Fig.9. Because we perform obstacle detection and overlays detected region on the image, color blended image can provide a better visual result as the detected obstacle is visible rather than cropped.



(a)



(b)

Fig.7.Color blending. (a) With blending. (b) Without blending.

Fig.8. Ghosting effect from blending of two camera inputs.

Fig.9. Object cropped without color blending.

Besides the dual-camber model described in previous section, other models were tested as well. Dual-camber shape provides a generic model that can be viewed with little distortion. In certain driving situations, other shapes can be more informative depending on the area of interest. For example, when parking the car, the flat region can be extended so the imagery near the vehicle is not distorted by the curvature of dual-camber model. Another use of different model shape is for checking the correctness of the calibration process visually. By rendering images as-is from the cameras and position those images in the correct position in world space as shown in Fig.10, we can quickly determine the errors made during calibration. Figure 11 shows a model that use elliptical rings instead of circular rings which provide more flat base at front and back of the vehicle. This matches the vehicle better as most vehicles are longer than they are wide.

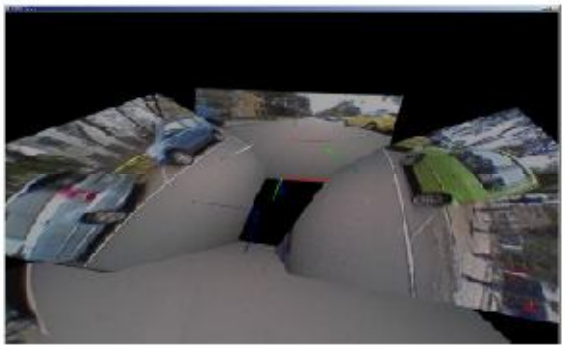


Fig.10. Rendering camera input images as-is in world space.

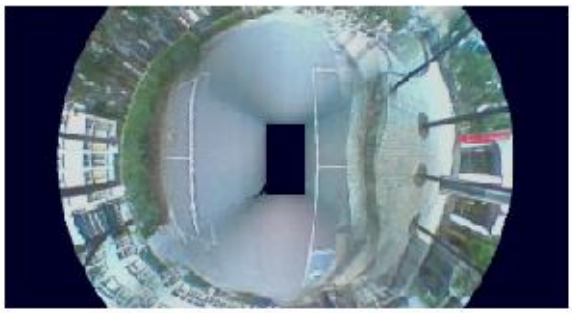


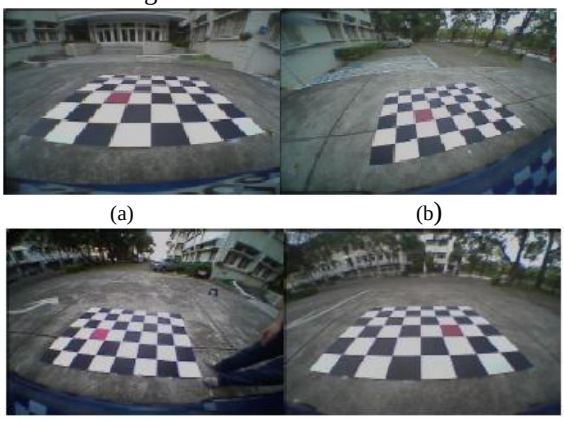
Fig.11.Ellipsoid dual-camber model matching vehicle's dimension.

III.EXPERIMENTS

Experiments using actual vehicle was performed. The experiments were set to test the detection performances on different types of road surfaces. Tests were also performed at different speeds to see how speed affects the accuracy of detection process. **b**In the experiment, the test vehicle is a 2007 Suzuki Swift 1.5L GLX. The dimension of the vehicle is 3.755m long and 1.69m wide. The vehicle has a small 4.6m turning radius, which is useful when testing ego-motion while performing turns on road-side parking scenario. Four 137° field-of-view cameras are mounted at each side of the vehicle. Front camera is mounted above the number plate; rear camera is mounted on the flat surface above the rear bumper; and side cameras are mounted under the side mirrors. Such setup is similar to various vehicle models with surround view technologies that are on the market. All cameras are connected to an industrial computer with a 4-input video capture card located at rear seat. The entire system is powered with a 12VDC lead-acid battery.

A. Combined calibration

Four calibration boards with 25cm wide blocks are places at pre-measured locations around the vehicle as shown in Fig.12. The cameras' extrinsic parameters as well as world positions and orientations are then estimated using least-squares estimation method. The estimated positions of cameras are recorded. Measurements were made to the closest centimeter with tape measurer for height and horizontal distance measured by distance from the ground point contact of plumb hanging from the camera. A top-view image is generated during combined calibration for manual verification of the process as shown in Fig.13. An optional world space x-y plane, or ground surface, can also be overlaid on input image for further verification. A 50cm grid overlaid onto input image of right camera is shown in Fig.14. Errors can be seen in Fig.14 (a)in which the grid does not match the calibration board, and horizon is above the true horizon in image.



(a) (b)

(c) (d)

Fig.12.Calibration boards for each camera. (a) Front. (b) Right. (c) Left. (d) Rear

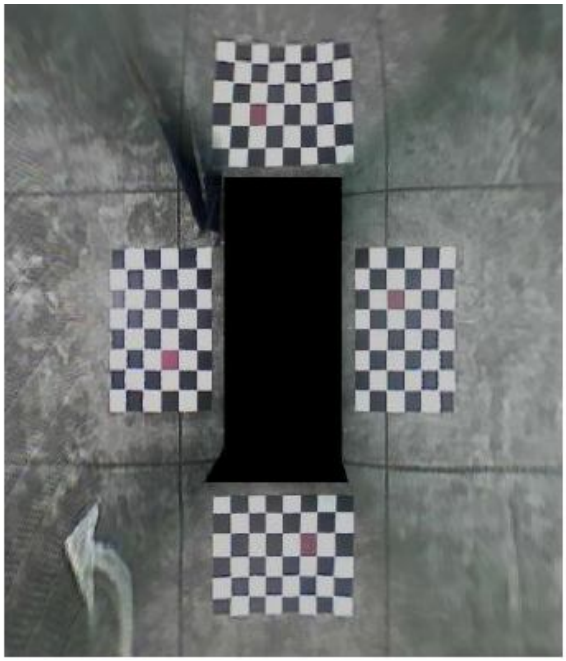
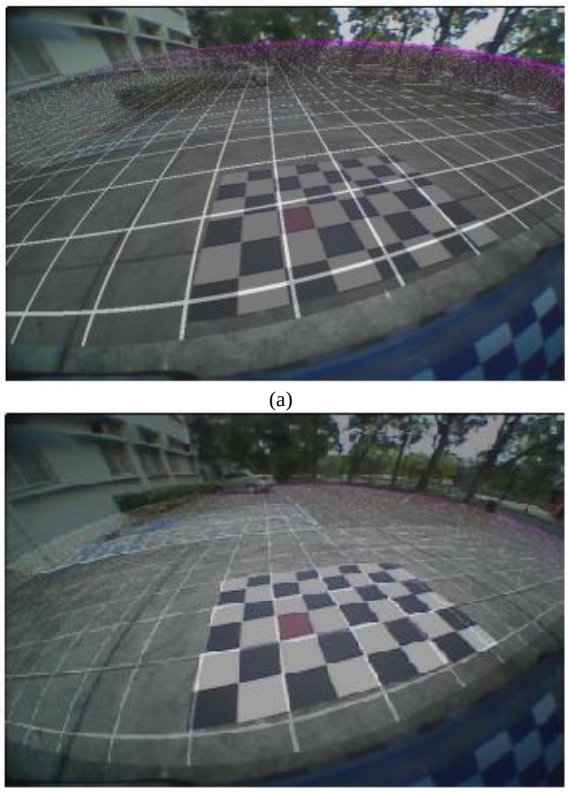


Fig.13. Top-view image generated for manual verification of combined calibration.



(a)

(b)

Fig.14.World space 50cm grid on ground plane overlaid on input image of right camera with calibration board of 25cm blocks. (a) Error in extrinsic parameters. (b) Correct extrinsic parameters

The calibration process is somewhat similar to the process used in HCE-C500 surround view by Alpine [12]. Alpine's system has the advantage of calibration boards is placed in relation to the vehicle; ours have the vehicle in relation to the boards. By performing

the calibration process, we discovered that Alpine's method would be affected by the deformation of the vehicle which can be caused by accidents or simply low strength body material.

B. Optimizations

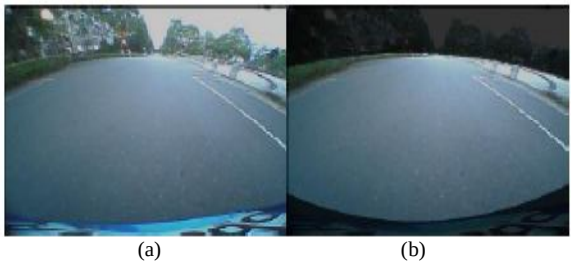
Because of the transformations from image coordinates to world coordinates and back involves many matrix operations and they are done repeatedly during online process, lookup tables were added to trade memory for performance.

- **Bird's-eye view map**

Bird's-eye view (BEV) maps are computed after combined calibration for each camera. BEV maps are 1-to-1 mapping tables that maps image coordinates into x-y coordinates of ground plane of world coordinate system. Areas above horizon and within the boundary of the vehicle are ignored. Fig.15 shows an input image and the area that has a mapping from ICS to WCS.

- **FOV lookup table**

Although BEV maps eliminate the calculation in ICS-to-WCS transformations, the transformation from WCS to ICS could not be done in the same manner without limiting the resolution and distance in world space. An FOV lookup table mapping undistorted r_u to distorted r_d where $r_u < 2.0$ is built to speed up distortion calculation. The use of FOV lookup table reduces average runtime by about 15%.



(a) (b)

Fig.15.BEV map. (a) Input image. (b) BEV map region shown highlighted.

CONCLUSIONS

Compared to current surround view systems, our system gives the driver immediate view of vehicle's surrounding environment and the freedom to change view point with current region of interest. The visual detection system provides additional hint of above-ground objects in the environment, allow faster focus on potential obstacles. The direct projection of input image onto the dual-camber model also increases the quality of output imagery with less memory requirement. The visual method used for detection removes the need for additional sensors, reduces the system cost. The reduced cost is transferred onto the processing power of the processor, which can be more easily integrated into vehicle's other systems, leading to better integration.

The system developed produces promising results on vision-based detection for vehicles. It also demonstrated a cost-effective method into building a vehicle safety system based on low-cost cameras with no radar and ultrasonic sensors. With recent advances in vehicle electronics such as NVIDIA's Jetson platform [10],our detection system can be integrated with it, taking advantage of massively parallel computation hardware. Modifying ego-motion estimation and detection to use parallel platforms CUDA [13] or OpenCL [14] would improve the performance by several times. Once integrated with vehicle's on-board computer, standard on-board information like speed sensor can also improve the ego-motion estimation, providing better detection result. The calibration method for building wide-scoped surrounding top-view monitoring can be extended to include more than four cameras to accommodate longer vehicles such as buses and trucks. Because the detection system uses only the Y component of YCbCr color space, the system may be able to incorporate infrared cameras and infrared lighting for detection to work at darker environment such as night.

ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Science and Technology, Taiwan under the grant of the research project MOST 104-2221-E-008-029-MY2.

REFERENCES

[1] Nissan, Around View Monitor, http://www.nissan-global. com/EN/TECHNOLOGY/INTRODUCTION/DETAILS/AVM/

[2] Honda, Multi-view Camera System, http://world. honda. com/news/2008/4080918Multi-View-Camera-System/

[3] T. Ehlgen, M. Thom, and M. Glaser, "Omnidirectional cameras as backing-up aid," in Proc. of IEEE 11th Int. Conf. on Computer Vision, Rio de Janeiro, Brazil, Oct.14-21, 2007, pp.1-5.

[4] T. Ehlgen and T. Pajdla, "Monitoring surrounding areas of truck-trailer combinations," in Proc. of 5th Int. Conf. on Computer Vision Systems, Bielefeld, Germany, Mar.21-24, 2007, pp.207-218.

[5] M. Bertozzi and A. Broggi, "GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection," IEEE Trans. on Image Processing, vol.7, no.1, pp.62-81, Jan. 1998.

[6] M. Bertozzi, A. Broggi, P. Medici, and P. P. Porta, "Stereo vision-based start-inhibit for heavy goods vehicles," in Proc. of IEEE Intelligent Vehicles Symp., Tokyo, Japan, June 13-15, 2006, pp.350-355.

[7] Luxgen, "Luxgen advanced technologies", http://www.luxgen-motor.com/technology-Advanced.html

[8] Spillard, "Spillard Optronics 360", June 2012, http://www.spillard.com/solutions/vision/optronics/optronics360.html

[9] Fujitsu, "Fujitsu MB86R24 automotive graphics processing LSI", May 2013, http://jp.fujitsu.com/group/fsl/downloads/en/release/20130516e.pdf

[10] NVIDIA, "Jetson automotive development platform", Apr. 2013, http://www.nvidia.com/object/jetson-automotive-developement-platform.html
[11] T.-H. Chou, "Wide-scoped top-view monitoring and image-based parking guiding," Master Thesis, Dept. of Computer Science and Information Engineering, National Central University, Chung-li, Taiwan, 2010.
[12] Alpine, "Alpine HCE-C500 Topview camera system", http://www.alpine-europe.com/p/Products/camera58/hce-c500
[13] NVIDIA, "CUDA developer zone", https://developer.nvidia.com/category/zone/cuda-zone
[14] Khronos, "OpenCL", http://www.khronos.org/opencl/

★★★