

Wide-scoped Surrounding Top-view Monitor for Advanced Driver Assistance Systems

Din-Chang Tseng^{1, a}, Yu-Chi Lin^{2, b}, and Tat-Wa Chao^{3, c}

¹Department of Computer Science & Information Engineering, National Central University, Taoyuan, Taiwan

²Department of Computer Science & Information Engineering, National Central University, Taoyuan, Taiwan

³Department of Computer Science & Information Engineering, National Central University, Taoyuan, Taiwan

^aemail: tsengdc@ip.csie.ncu.edu.tw, ^bemail:985402002@cc.ncu.edu.tw,
^cemail:985402003@cc.ncu.edu.tw

Keywords: Advanced Driver Assistance Systems, Parking Assistance, Around View Monitor, Image Processing

Abstract. To improve the scope range of the traditional around view monitor (AVM) systems for Advanced Driver Assistance Systems (ADAS), we propose a wide-scoped surrounding top-view monitor (STM) system with the same equipment as the traditional AVM system. In the proposed system, four wide-view cameras are mounted on front, rear, and the both sides of the vehicle to capture sequential images; then each four images are composed into a single seamless surrounding top-view image of the vehicle with the view radius more than 30 meters. The proposed system consists of eight stages: camera calibration, vignetting effect elimination, distortion calibration, top-view transformation, image registration, brightness uniformity, color blending, and bowl model mapping. The system performance and image quality are evaluated in experiment.

Introduction

One of the main reasons for traffic accidents is caused from the limited field of view for the drivers. The driver is unable to absolutely understand the surrounding area when paying attention in front and also turning his head, look through the side-view mirrors to watch the area behind the vehicle. Especially the large vehicles like trucks, busses or other public transportation. Because of the construction, it's a challenging mission to take control with such large blind spot area around the vehicle, as the blind spot area of general vehicle

In last decade, several surrounding monitor systems employed different devices were proposed. In 2005, Nissan announced a parking assistance system known as "Around View Monitor". The system uses four wide-angle cameras mounted on the front, sides, and rear of the vehicle to capture images. These images are then synthesized into an image like a bird's eye view of the vehicle with its surrounding scenery. However, the synthesis image has four clefts in which the objects are invisible.

Employing four fish-eye CCD cameras located in the front, rear and both side of the vehicle, Honda has completed the development in Multi-view camera system in 2007. For the system setup and the manner of display, Honda Multi-view camera system is similar as Nissan Around view monitor, but for the function, Multi-view camera system has increased other information like visibility support and tight driving support, which provides a reference for the distance between the vehicle and approaching obstacles.

Ehlgen et al. [1] provided a parking constructed with two catadioptric cameras. The cameras have been mounted on the left and right rear roof edge of the vehicle to create a bird-view of the surrounding area behind the vehicle as well as the area on the left and right hand side. For the sake of extending to a larger visible region, their study has stretched the bird-view image that provides a larger field of view.

Ehlgen and Pajdla [2] developed a bird-view surrounding monitor system that suitable for the large vehicles like buses, trucks and truck-trailer combinative vehicles. Their system is consists of four omnidirectional cameras mounted on a truck and trailer. The magnetic sensor is used to measure the kink angle between the truck and the trailer and depends on the direction of a magnetic field to correct the overlap of bird's-eye view image between the truck and trailer.

To increase the view scope, we propose a wide-scoped surrounding top-view monitor system. The proposed system consists of eight stages: camera calibration, vignetting effect elimination, distortion calibration, top-view transformation, image registration, brightness uniformity, color blending, and bowl model mapping as shown in Fig. 1.

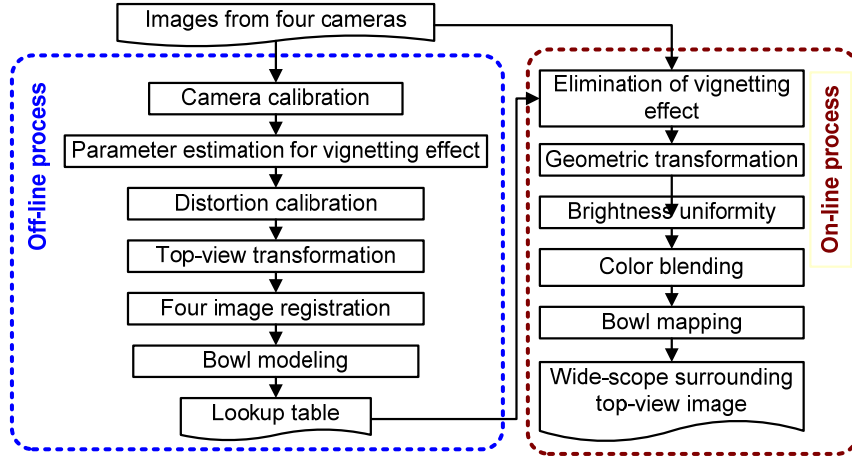


Fig. 1. The block diagram of wide-scoped STM system.

Seamless Top-view Monitoring

The tasks of the wide-scoped surrounding top-view monitor system are divided into off-line and on-line processes as shown in Fig. 2. The off-line processes: camera calibration, vignetting effect elimination, distortion calibration, top-view transformation, image registration and the on-line processes: bright uniformity, color blend are described in this section. These processes are also used to create a seamless surrounding top-view monitor (STM) system.

Camera's parameter calibration

The purpose of camera calibration is to obtain the corresponding relation between three-dimensional world coordinate system and the two-dimensional image coordinate system. Before the cameras are installed, we can calibrate the intrinsic parameters for getting the transformation between image coordinates and camera coordinates. After mounting the cameras on the vehicle, we calibrate the extrinsic parameters which relate world coordinates to each of cameras coordinates.

Zhang [3] proposed a flexible, robust and low cost method for camera calibration. We print the pattern and attach to a planar surface, and then only utilize the camera to observe a planar pattern shown in two or more different orientations. Either the camera or the planar pattern can be freely moved. The motion need not be known.

We utilize the algorithm and camera model to calculate the constraints on the camera's intrinsic and extrinsic parameters by estimating the homography between a planar model containing the calibration target and several images of this target.

In the Zhang's model [3], if the coordinates of a 3D point p_w is $[x_w \ y_w \ z_w]^T$ and its 2D image point q is $[u \ v \ 1]^T$, then

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A[R \ t] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \quad (1)$$

where s is a non-zero scale factor; R and t are 3×3 rotation matrix and 3×1 translation vector, respectively; A is camera intrinsic parameters matrix,

$$A = \begin{bmatrix} \alpha & \gamma & u_o \\ 0 & \beta & v_o \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

where (u_o, v_o) is the camera optical axis center on the image coordinate system; α and β are the focus lengths in image u and v axes, respectively; γ is the skew parameter between photosensitive element array (CCD) and lens structure.

There are totally twelve extrinsic parameters and five intrinsic parameters. Zhang has proposed a sequence computation to initially estimate the intrinsic parameters and then refine the extrinsic parameters by solving the nonlinear minimization problem with the Levenberg-Marquardt algorithm [3].

Wide-angle lens distortion correction

Based on the characteristic of wide-angle camera, we can achieve the surrounding monitor system with a few cameras; however, the wide-angle lens bright distorted images. The goal of distortion calibration is just to correct the lens distortion to get the mapping between the actual image plane and the perspective camera model. Many distortion models have been proposed such as the classical polynomial model [4], the division model [5], the rational model [6], stereographic projection [7], and the unified catadioptric model [8].

We here utilized the *FOV* model proposed by Deverney and Faugeras [9] to calibrate the lens distortion. The larger incident angle between 3-D point and optical axis is, the larger distance between image point and image center is. In the *FOV* model, the change of incident angle is proportional to the change of distance,

$$r_d = \frac{1}{\omega} \tan^{-1}(2r_u \tan \frac{\omega}{2}), \quad (3)$$

where ω is the distortion parameter associated to field of view, and the inverse function is

$$r_u = \frac{\tan(r_d \omega)}{2 \tan \frac{\omega}{2}} \quad (4)$$

One example result of lens distortion calibration is shown in Fig. 2.

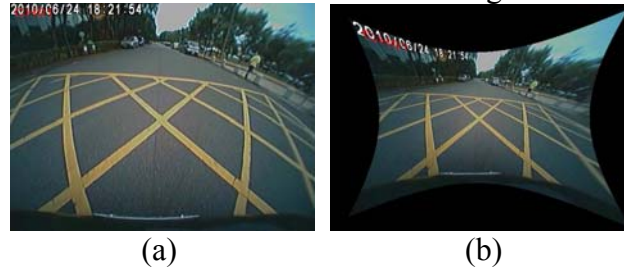


Fig. 2. The result of distortion correction. (a) Original image. (b) After distortion correction.

Elimination for vignetting effect

Vignetting is caused by partial obstruction of light from the object space to image space. We divide the vignetting effect into two reasons: natural vignetting and mechanical vignetting. We observe if the object is a plane of uniform brightness exactly perpendicular to the optical axis, the luminance of image fall off with the distance away from the image center. The luminance at the off-axis image point is described as

$$I(\theta) = I_0 \cos^4 \theta, \quad (5)$$

where I_0 is original luminance, $I(\theta)$ is the luminance, and θ is the angle of the off-axis at the entrance angle [10]. If the effective focal length f of camera lens, we can take the distance r of the point to the center in the image as

$$I(r) = I_0 \left(\frac{f}{\sqrt{f^2 + r^2}} \right)^4 = I_0 \frac{1}{(1 + (r/f)^2)^2}. \quad (6)$$

The effect resulted from the mechanical vignetting is modeled by a parameter α ,

$$I'(r) = (1 - \alpha r) I(r). \quad (7)$$

We combine Eqs.(6) and (7) as

$$I'(r) = I_0 \frac{1 - \alpha r}{\left(1 + (r/f)^2\right)^2}. \quad (8)$$

To estimate parameter α , we calculate the square error between the points in the image and their actual illuminance by capturing the object with uniform illuminance,

$$E_{total} = \sum_{i=1}^n E(p_i), \quad (9)$$

where $E(p_i) = \|I'(r_u) - I_0\|^2$ is the error for a point. r_u is calculated by inverse function of distortion model for the distance r . One result of elimination for vignetting effect is shown in Fig. 3.

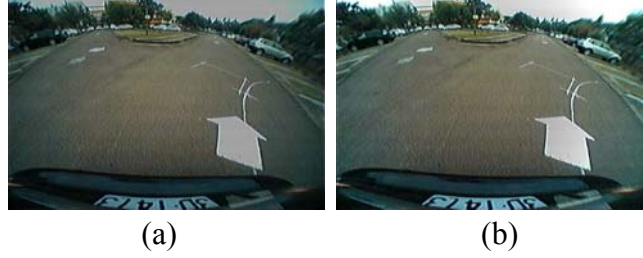


Fig. 3. Elimination for vignetting effect. (a) Original image. (b) Enhanced brightness.

Top-view transformation

The top-view transformation is to back project images onto the ground surface. The task can be achieved by a homographic matrix. If $q_i = [u_i \ v_i]^T$ and $q_j = [u_j \ v_j]^T$ are projected from one 3D point on planes I_i and I_j , respectively. The point-to-point relation can be expressed by

$$s q_j = H q_i, \quad (10)$$

where H is a homography matrix.

We can use the intrinsic and extrinsic parameters of the cameras to construct the homography matrix. Assume the ground is the plane of WCS which means $Z = 0$. The top-view image is generated from a virtual camera with image plane being parallel to the ground. Then the relation between a 3D point $[X \ Y \ 0]^T$ on the ground and the image point $[u \ v]^T$ is just

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A[R | t] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}. \quad (11)$$

The projection is a one-to-one transformation; thus

$$s \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = H^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \quad (12)$$

Assume the relation between the virtual image plane and the WCS is

$$s \begin{bmatrix} u_v \\ v_v \\ 1 \end{bmatrix} = H_v \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \quad (13)$$

where (u_v, v_v) is the point of the virtual image plane. The transformation from the image point $[u \ v]^T$ to a point on the virtual image is

$$s \begin{bmatrix} u_v \\ v_v \\ 1 \end{bmatrix} = H_v H^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \quad (14)$$

The 3×3 homography matrix $H_v H^{-1}$ can also be estimated by a least-squares estimation method by taking

$$s \begin{bmatrix} u_v \\ v_v \\ 1 \end{bmatrix} = H \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \quad (15)$$

One example through the distortion calibration and top-view transformation is shown in Fig. 4.

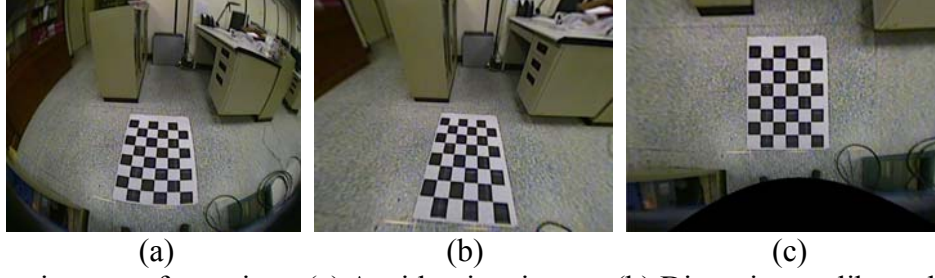


Fig. 4. Top-view transformation.. (a) A wide-view image. (b) Distortion- calibrated image. (c) Top-view transformed image.

Image registration

Theoretically, the four top-view images can be registered by the camera extrinsic parameters; however, there is still a little error; thus the four top-view images need further registered by an extra process. The registration only needs to consider the rotation and translation problem; thus a similarity transformation,

$$\begin{aligned} \begin{bmatrix} u' \\ v' \end{bmatrix} &= \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} \\ &= \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix}, \end{aligned} \quad (16)$$

with the least squares estimated parameters (a , b , Δu , Δv).

Brightness uniformity

Environmental light source and the different levels of exposure of the four cameras make the image's brightness significantly different. We here compute the brightness mean and standard deviation of each image, then adjust those images brightness into a similar illumination by

$$I_{\text{result}}(p_i) = \frac{\sigma_{\text{reference}}^I}{\sigma_{\text{source}}^I} (I_{\text{source}}(p_i) - \mu_{\text{source}}^I) + \mu_{\text{reference}}^I, \quad (17)$$

where

$$\mu^I = \frac{\sum_{i=1}^n I(p_i)}{n}, \quad \sigma^I = \sqrt{\frac{1}{n} \sum_{i=1}^n (I(p_i) - \mu^I)^2},$$

$I(p_i)$ is the brightness of pixel p_i , μ^I is the brightness mean, σ^I is the brightness standard deviation. One example of the brightness adjustment is shown in Fig. 5.

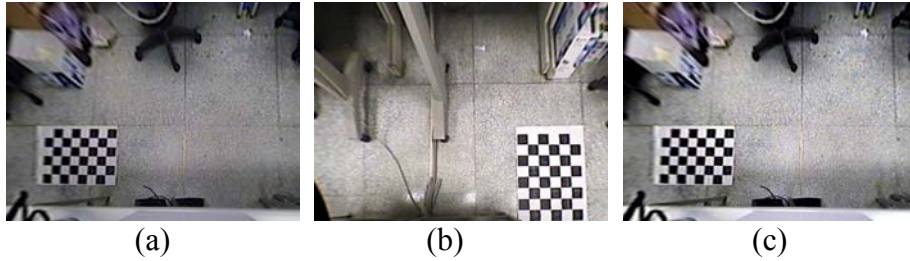


Fig. 5. Example of brightness adjustment. (a) Original image. (b) Reference image. (c) Brightness adjustment image.

Wide-Scoped Top-view Monitoring

To generate a wide-scoped top-view surrounding image, the bowl modeling and bowl mapping are combined into the seamless *STM* system and will be proposed in this section.

Actually, the top-view transformation can be extended to a larger-area surrounding image as one example shown in Fig. 6 (a). However, the far-distance region is heavily distorted; thus the key task is how to de-distort the image as shown in Fig. 6 (b).

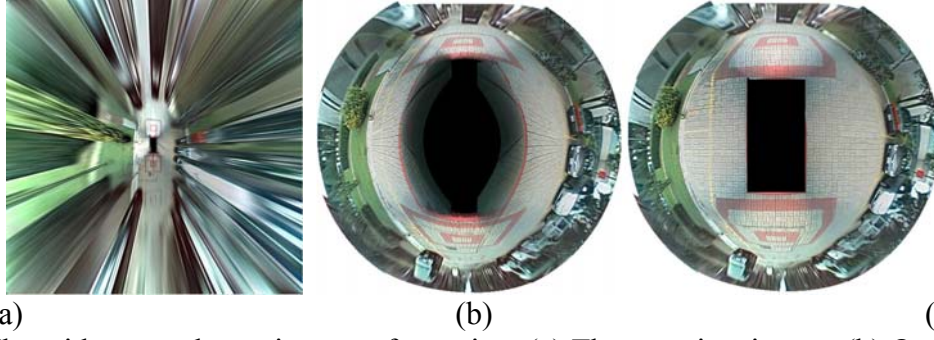


Fig. 6. The wide-scoped warping transformation. (a) The top-view image. (b) One warping transformation. (c) Two warping transformations.

Here we utilize the *sine* function to warp the synthesis image into a curve surface. The new position is defined by *sine* inverse function

$$\begin{cases} x_n = \frac{(x_o - x_c) r}{(S - \sin^{-1}(r/R)) R} \\ y_n = \frac{(y_o - y_c) r}{(S - \sin^{-1}(r/R)) R} \end{cases}, \quad (18)$$

where (x_n, y_n) is the new coordinates after warping transformation, (x_o, y_o) are the coordinates of top-view image, (x_c, y_c) is image center, R is the maximum radius after transformation, and $r = \sqrt{(x_o - x_c)^2 + (y_o - y_c)^2}$. After transformation, the original square image will become a circle image. To maintain the center area un-deformed, the center area must be transformed by another formula. Thus,

$$x_n = \begin{cases} x_e = \frac{(x_o - x_c) r}{(S - \sin^{-1}(r/R)) R} \\ x_i = \frac{(x_o - x_c) \sin^{-1}(r/R)}{r} cR \end{cases} \quad (19)$$

is used instead of the original x_n , where x_n is the new coordinates location after warping transformation, x_e is the external-range coordinate, x_i is the internal-range coordinate, c is a constant usually set at 1.50~1.70. One example of the final warning image is shown in Fig. 6 (c).

Experiments

We employed four low-cost wide-angle 1/4 inch *CMOS* cameras for experiments. The cameras providing a horizontal view angle about 136 degrees and 115 degrees vertical angle. The algorithms were implemented in C++ programming language and Microsoft Foundation Class (*MFC*) Library, and all experiment were excepted on a general *PC* with *Intel® Pentium® Core2 Duo 2.66GHz* and *1.99GB RAM*, Microsoft® Windows *XP* professional operation system.

Four experiments with four sequences of synchronous images are presented as shown in Figs.7 - 10.

The performance of the proposed system is dependent on the speed of on-line processes: geometric transformation of images, brightness uniformity, and color blending. Currently, we can generate the *STM* images with 15 - 20 frames per second.

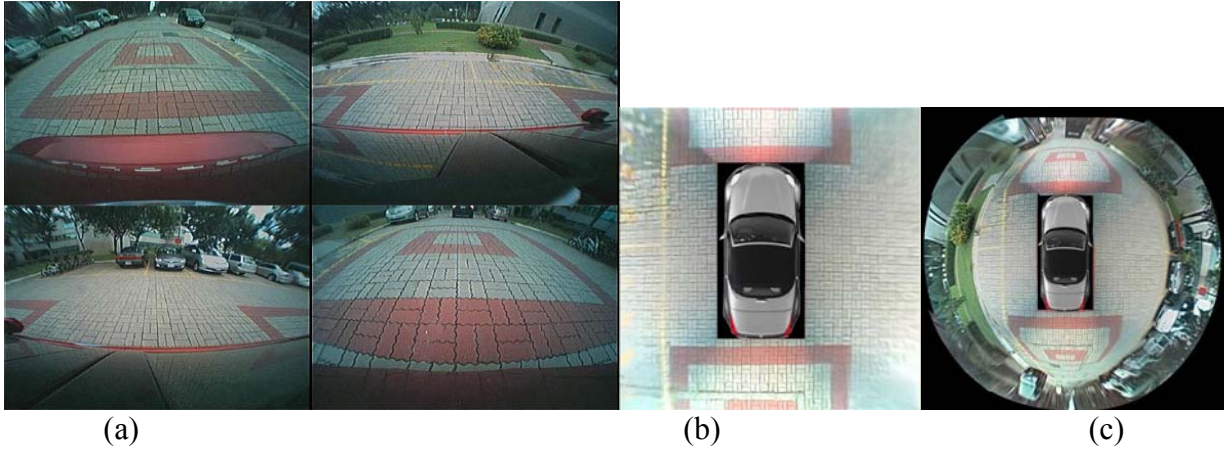


Fig. 7. The first wide-scoped *STM* images. (a) Four original images. (b) The general *STM* image. (c) The wide-scoped *STM* image.

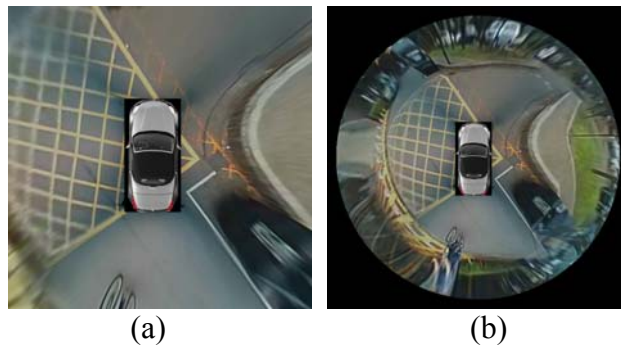


Fig. 8. The second wide-scoped *STM* images. (a) The general *STM* image. (b) The wide-scoped *STM* image.

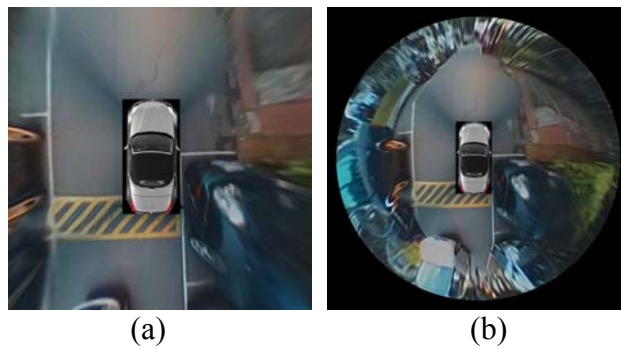


Fig. 9. The third wide-scoped *STM* images. (a) The general *STM* image. (b) The wide-scoped *STM* image.

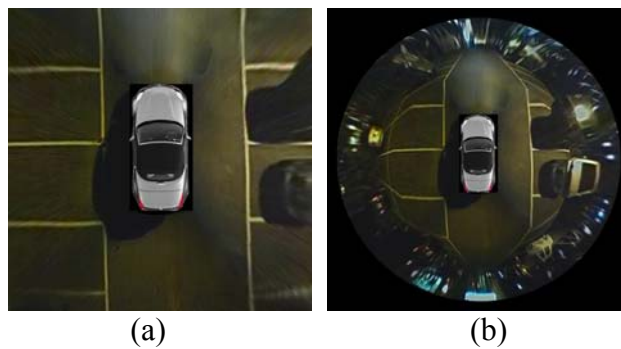


Fig. 10. The fourth wide-scoped *STM* images. (a) The general *STM* image. (b) The wide-scoped *STM* image.

Conclusion

In this study, we employed four wide-view cameras mounted on each side of the vehicle to monitor the surrounding traffic situation for driver to avoid the possible collision. Different from the current parking assistance systems, the proposed system provides a wide-scoped surrounding top-view monitoring area with the same equipment for the traditional *STM* system.

The surrounding top-view monitor system is useful for driving in the narrow street or parking. The larger view area helps the driver focus on understanding the parking size and the collision distance from the vehicle, it is especially useful in slower speed.

The wide-scoped surrounding top-view monitor system is also useful for driving in the urban area under 40 *km/hr* speed. It is sufficient in time to avoid possible collision around the host vehicle. The wide-scoped top-view monitoring system can also be used as a blind spot detection for lane change assistance on the highway.

The result of the wide-scoped top-view monitoring system is affected by the camera view angle and the image resolution. We are able to improve the quality of synthesis images by employ the wider-angle camera and higher resolution video devices. Furthermore, the system is just migrated into an embedded system.

Acknowledgement

This work is partially funded by the National Science Council, Taiwan under the grant of the research project NSC 100-2221-E-008 -115-MY3.

References

- [1] T. Ehlgen, M. Thorn, and M. Glaser, "Omnidirectional cameras as backing-up aid," in Proc. of IEEE Int. Conf. on Computer Vision., Rio de Janeiro, Brazil, Oct.14-21, 2007, pp.1-5.
- [2] 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, CD.
- [3] Z. Zhang, "A flexible new technique for camera calibration," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.22, no.11, pp.1330-1334, 2000.
- [4] C. C. Slama, Edit., Manual of Photogrammetry, 4th edition, American Society of Photogrammetry and Remote Sensing, Falls Church, Virginia, 1980.
- [5] A. W. Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," in Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, Hawaii, Dec.11-13, 2001, vol.1, pp.125-132.
- [6] D. Claus and A. W. Fitzgibbon, "A rational function lens distortion model for general cameras," in Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, San Diego, CA, Jun.20-25, 2005, vol.1, pp.213-219.
- [7] M. M. Fleck, Perspective Projection: The Wrong Imaging Model, Technical Report TR 95-01, Computer Science, University of Iowa, 1995.
- [8] C. Geyer and K. Daniilidis, "Catadioptric projective geometry," International Journal of Computer Vision, vol.45, no.3, pp.223-243, 2001.
- [9] F. Devernay and O. Faugeras, "Straight lines have to be straight," Machine Vision and Application, vol.13, no.1, pp.14-24, 2001.
- [10] S. B. Kang and R. Weiss, "Can we calibrate a camera using an image of a flat, textureless lambertian surface?," in Proc. European Conf. on Computer Vision, Dublin, Ireland, Jun.26 – Jul.1, 2000, pp.640-653.