Symmetric Trinocular Dense Disparity Estimation For Car Surrounding Camera Array

Yi-Min Tsai[a] and Yu-Lin Chang[b] and Liang-Gee Chen[b]

[a]Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan;
[b]DSP/IC Design Lab, Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan

ABSTRACT

This paper presented a novel dense disparity estimation method which is called as symmetric trinocular dense disparity estimation. Also a car surrounding camera array application is proposed to improve the driving safety by the proposed symmetric trinocular dense disparity estimation algorithm. The symmetric trinocular property is conducted to show the benefit of doing disparity estimation with three cameras. A 1D fast search algorithm is described to speed up the slowness of the original full search algorithms. And the 1D fast search algorithm utilizes the horizontal displacement property of the cameras to further check the correctness of the disparity vector. The experimental results show that the symmetric trinocular property improves the quality and smoothness of the disparity vector.

Keywords: Car Radar, Disparity Estimation, Trinocular, Camera Array

1. INTRODUCTION

Disparity estimation through binocular cameras has become popular recently since the estimated depth map is used on the 3D video, car safety and other real world things. It is a fundamental and advanced work in relating video processing and human life. For example, depth map generation and 3D scene construction are two growing demands for 3D-TV and vehicle-embedded application. Fig.1 shows a car radar system. It detects the surrounding environment and converts it into a depth map, then the depth map is shown from third-person viewpoint, just like the radar shown in the right part of Fig.1.

Mark et al proposed a dense stereo system for obstacle detection on a car.[1] The idea of using stereo camera to detect the distance of the obstacle instead of using millimeter wave radar was presented to make the car radar system cheaper and broader. Table 1 shows the comparison between the millimeter wave radar and stereo camera using on the vehicles. The detecting distance and relative velocity of the two devices are both acceptable, but the size and cost of stereo camera is much smaller and fewer than the millimeter wave radar. Since the size of the stereo camera is smaller, it would be possible to build up a car surrounding camera array to detect the obstacles in its panoramic environment with some more stereo cameras.

Table 1. The comparison between the millimeter wave radar and stereo camera

Parameters	Millimeter wave radar	Stereo camera
Detecting distance	150m	20cm∼100m
Relative velocity	-80∼240km/hr	±172.8km/hr
Hori. detecting angle	±8 degrees	±20 degrees
Scanning frequency	33Hz	30Hz
Power consumption	500mW∼1W	200mW∼400mW
Size	Large	Small
Cost	∼ USD400/radar	∼USD100/pair

Further author information: (Send correspondence to Liang-Gee Chen)
L.-G. Chen: E-mail: lgchen@video.ee.ntu.edu.tw, Telephone: +886-2-2363-5251 ext 344

Visual Communications and Image Processing 2007, edited by Chang Wen Chen, Dan Schonfeld, Jiebo Luo,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6508, 65081F, © 2007 SPIE-IS&T · 0277-786X/07/$18

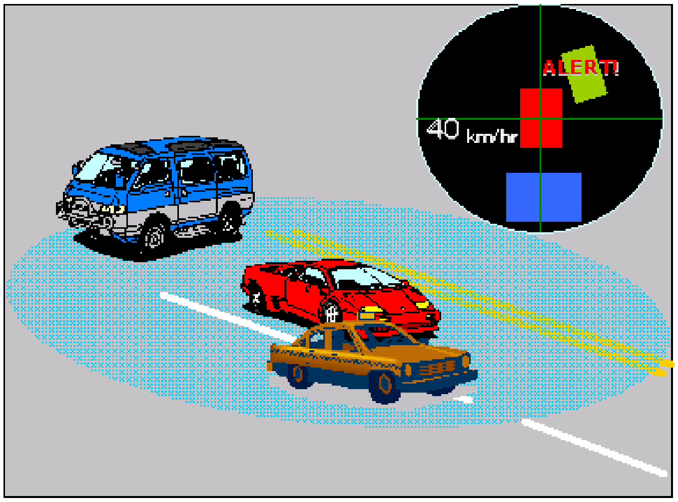
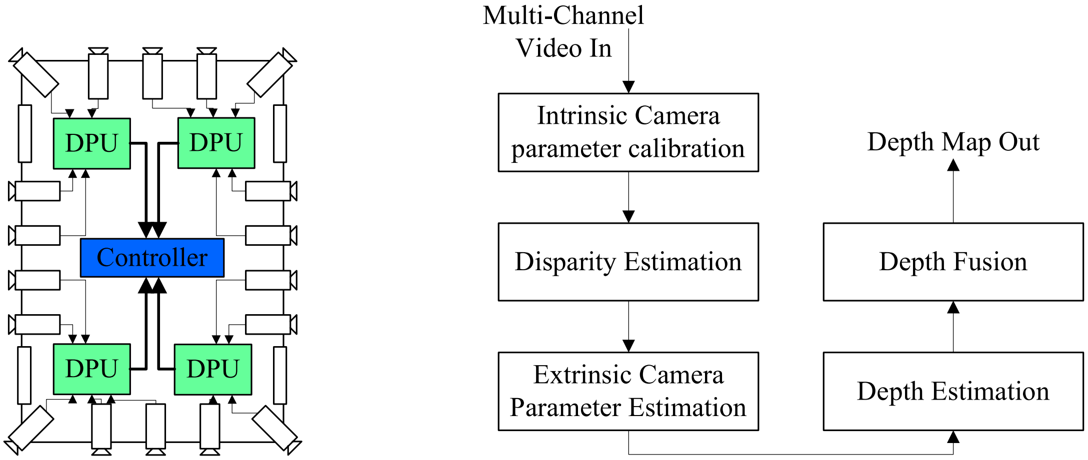


Figure 1. Car Radar System



(a) (b)

Figure 2. Car surrounding camera array (a) Overall architecture (b) DPU flow

The overall architecture of the car surrounding camera array is shown in Fig.2(a). Cameras are put around the car, and there are four dimension processing units(DPU) to calculate the disparity and depth information retrieved from the cameras. After the disparity/depth estimation, the central controller fuses the depth map from the depth processing units. Finally, the panoramic depth map is converted to third-person viewpoint to show on the car radar. The processing flow of the DPU is show in Fig.2(b). The intrinsic camera parameter calibration helps to retrieve the focal length and distortion of the devices. Then the disparity estimation calculates the dense disparity vector of each pixel. If needed, extrinsic camera parameter such as camera distances are estimated. Then, the disparity vectors are converted to the values on the depth map by the camera distance. Finally, the depth map is fused and sent out. The disparity estimation is the main part for the dimension processing unit. With the largest computation in the car surrounding camera array, the accuracy and the speed of the disparity estimation is a tradeoff.

2. PRIOR ART

Disparity estimation is divided into two classes: one is block-based disparity estimation, the other is dense disparity estimation. The block-based disparity estimation is usually adopted in the multi-view video coding system. The dense disparity estimation is also called per-pixel disparity estimation, since its disparity vector density is one disparity vector per pixel. The disparity estimation used by the car surround camera array should be as fine as possible to detect further objects. So dense disparity estimation is adopted.

There are two kinds of dense disparity estimation: area-based methosd and phase-based methods. Area-based methods work on pixel domain. They relate block correlation and find a matching disparity vector for each pixel. The phase-based methods work on frequency domain. The matching of the frequency response of the textures is the point for the phase-based methods. Fleet proposed a phase-based disparity estimation[2] method to extract the binocular disparity. The phase-difference and phase-correlation are considered. But the matching criterion is not enough since textures are not fully considered. Grammalidis et al took multiocular system into consideration[3] and proposed an area-based disparity estimation method for the coding of occlusion and disparity information. They concentrated on the image data reproduction through disparity and occlusion information. Ouali et al described a phase-based disparity estimation method by Gabor filter.[4] The Gabor filter helps to figure out the real texture for the matching. However, the tuning of the filters is another problem for the accuracy of the disparity vectors. Strecha presented a PDE-based Multi-view depth and disparity estimation method.[5] But the PDE-based method does not produce smooth results while processing natural images. An one dimensional dense disparity estimation is proposed by Oisel et al[6] by optical flow. But the obtained disparity vector density is too sparse to produce accurate depth map. Huang et al proposed a three-view dense disparity estimation method with Gabor filter[7].[8] The disparity estimation method with Gabor filter corrects most of the occlusion but costs too much computational power.

In this paper, a symmetric trinocular property is proposed to utilize the trinocular camera characteristic. By this property, the proposed area-based symmetric trinocular dense disparity estimation (STDDE) and 1D fast search algorithm are described in section 4. The experimental results of the synthetic images and real images are shown in section 5. At last, the conclusion remarks the proposed algorithms.

3. PROBLEM DEFINITION

Disparity estimation algorithms can roughly be divided into two kinds: optical flow and block matching method. The application of our design is going to be put on mobile devices, so hardware-oriented algorithms such as block matching methods are more suitable for this kind of application. The searching strategy of the block matching disparity estimation influences the computational speed most. The full search block matching algorithm is accurate, but it is time consuming under software. For this reason, we have to use fast search algorithm instead of full search. However, with fast search, using only binocular camera may result in high error rate on disparity vectors. For this reason, we have to find a new method for finding the correct disparity vector.

4. PROPOSED METHOD

Our goal is not only the disparity estimation method but also the application of it. Nowadays, more and more car accidents on road happen due to carelessness of the driver. The avoidance of car accidents becomes a serious work. In this reason, a car surrounding camera array combined with dense disparity estimation is needed. Car surrounding camera array detects the distance between the car and surroundings, it warns the driver while getting too close to the obstacles and then display the real-time result on the monitor for drivers. The following detection flow chart in Fig.3 shows the block diagram of our proposed system. Three input frames are captured in the same time, and we then use the proposed symmetric trinocular dense disparity estimation algorithm to generate disparity map. After the generation of disparity map, the depth map is converted out and drawn on the car radar display.

Firstly, the block-based and dense disparity estimation method is compared to obtain the method with better quality. For block-based DE, shown in Fig.4(a), the disparity vector in each block (16x16) is the same, and each block does not overlap with other block. In this way, block-based DE may result in information loss if the image structure is complicated. For dense DE, it means that disparity vectors are constructed in per pixel way but now

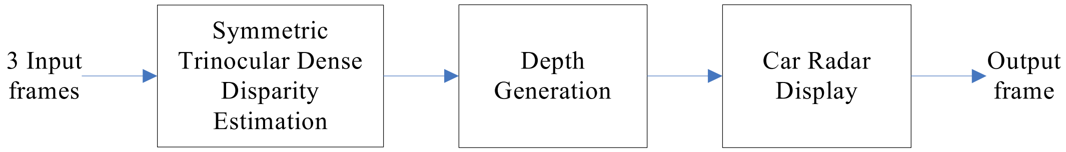
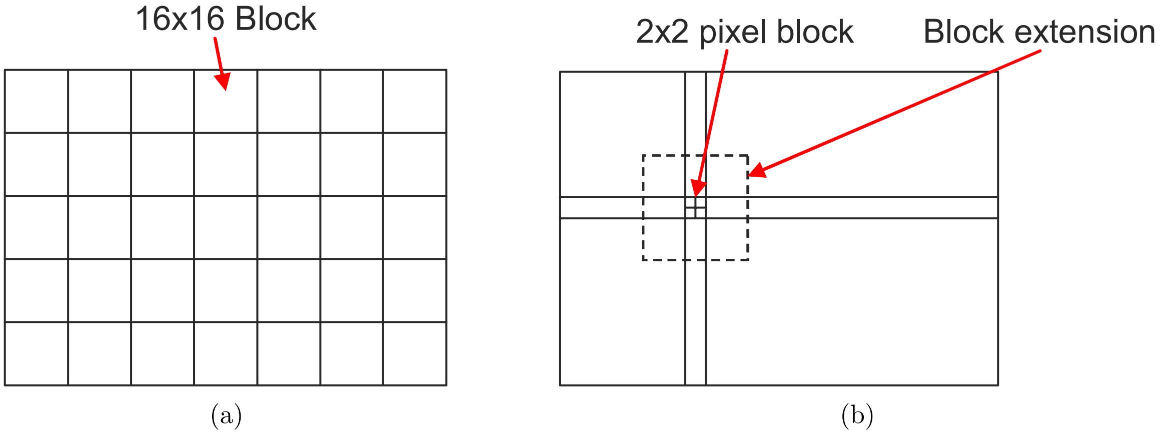


Figure 3. Detection flow



(a) (b)

Figure 4. Disparity Estimation (a) Block-based Disparity Estimation (b) Dense Disparity Estimation

a 2x2 pixel block is regarded as a calculating unit since the depth map resolution is usually half than the original picture in the width and height. In calculating disparity, the 2x2 small block is extended to a 16x16 block, and the SAD value of the 16x16 block is calculated. After finding the disparity value, it is filled into this 2x2 small block shown in Fig.4(b). The dense DE is much finer than the block-based DE, so we decide to use dense DE.

Secondly, the binocular camera DE is compared with trinocular camera DE. Binocular camera DE is the traditional way, however, our new DE algorithm is based on multi-camera. The three cameras are used to do the tirnocular DE. With one more camera information, the disparity vectors are more smooth, and the occlusion areas are no more existed.

4.1. Symmetric trinocular dense disparity estimation (STDDE)

The traditional dense disparity estimation methods often suffer from large computation and slow operating speed. But if the dense disparity estimation is going to be adopted on the car surrounding camera array, the operating speed should be kept as fast as possible. For our algorithm, the input video format is CIF with 30FPS. So the operating speed must be kept as 30FPS, too. Common area-based dense disparity estimation uses full search algorithm as their searching strategy. The full search algorithm consumes much time to calculate the disparity vectors. Here, an 1D fast search algorithm is proposed to speed up the calculation. Along with the proposed symmetric trinocular property, the proposed dense disparity estimation does not suffer from accuracy degradation. The proposed smmetric trinocular dense disparity estimation is defined below.

4.2. 1D fast search algorithm

In order to speed up the disparity estimation, the fast search algorithm is used instead of full search algorithm. Due to the horizontal displacement of trinocular camera, the disparity tends to be in the horizontal direction. Because of this characteristic, applying 1D fast search is reasonable.

The 1D fast search speeds up the DE by reducing the disparity candidate, and in each step it searches only horizontal and vertical direction once. Fig.5 shows the two 1D directions searched by 1D fast search algorithm.

The three step fast search algorithm and the 1D fast search algorithm are also compared in Fig.6. The first step is to find a best matching candidate in the horizontal direction. Then from the best matching position, the

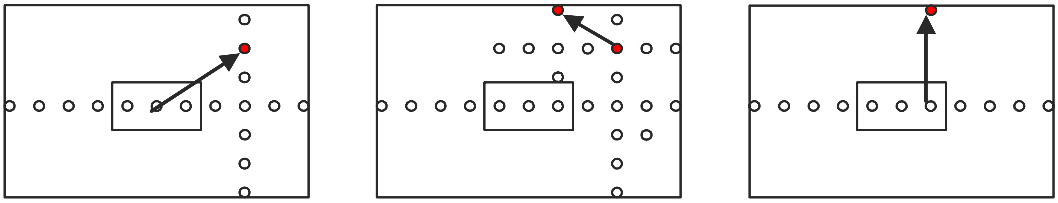
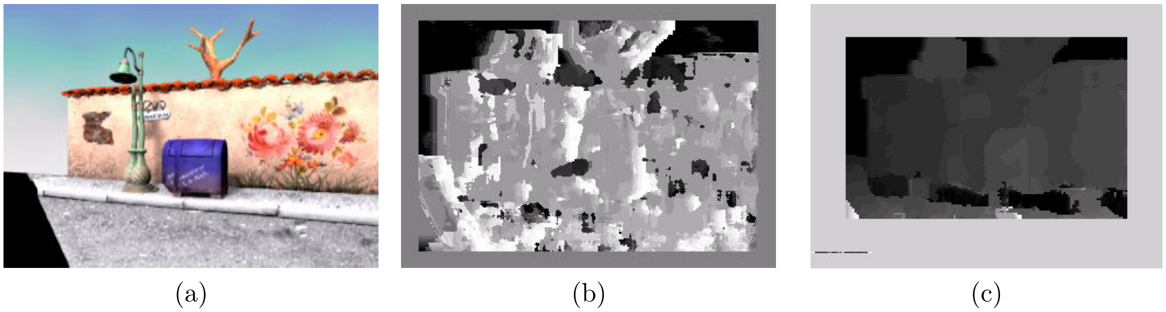


Figure 5. 1D Fast Search for trinocular dense disparity estimation



(a) (b) (c)

Figure 6. 3-step fast search and 1D fast search algorithms (a) Original (b) 3-step (c) 1D fast search

best matching candidate in the vertical direction is found. The second step is to find the best matching candidate in the horizontal and vertical direction again. Finally, the best matching criterion over the whole search window is found.

4.3. Trinocular Camera

The Trinocular camera setup is to put the three lenses horizontally separated with 20cm. Based on the 1D fast search algorithm, two of the three cameras are chosen to get a general direction in calculating disparity. The middle camera and left camera (or right camera) are used to select local disparity candidate. Then our proposed trinocular camera calculating algorithm follows up the previous calculation. The middle image provides the reference image. Three images are captured in the same time. If the distance between right and middle camera and the distance between middle and left camera are the same, the object disparity vector of back image and middle image and that of middle image and front image must have the same magnitude but inversed direction as shown in Fig.7. The property is called as symmetric trinocular property. As for the matching criterion of the disparity estimation methods, the Sum of Absolute Difference (SAD) is used for this area-based method. It is expressed as the following equation:

$$SAD_m = \sum_{i,j \in MB} (|f(i - MV_{kx}, j - MV_{ky}, n-1) - f(i,j,n)| + |f(i + MV_{kx}, j + MV_{ky}, n+1) - f(i,j,n)|) \quad (1)$$

The function f(i, j, n) denotes the intensity of the input image. i and j are the horizontal and vertical axes, and n describes the camera number. The first SAD calculation describes the relationship between the left and middle camera, and the second SAD calculation describes that between the middle and right camera. The symmetric property of the two SAD calculation is called as the symmetric trinocular property.

In this way, the symmetric trinocular property is combined with the 1D fast search to refine disparity vector. the result of disparity map generation which uses only binocular camera and which uses our proposed algorithm, STDDE are shown in Fig.8, the STDDE outperforms the original binocular method.

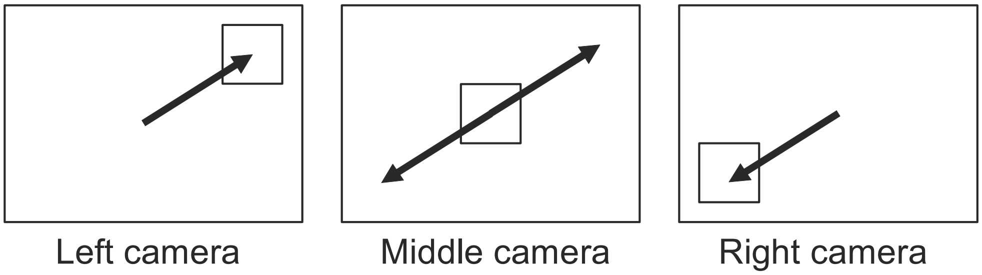


Figure 7. Symmetric trinocular property

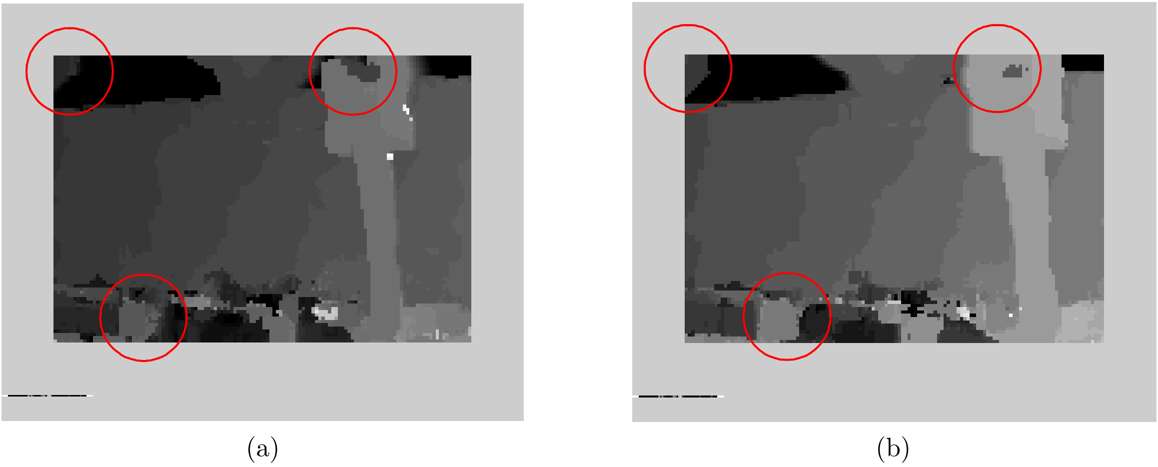

(a)
(b)

Figure 8. Binocular DE v.s. STDDE (a) 1D fast search with binocular camera (b) 1D fast search with trinocular camera

4.4. Depth Generation

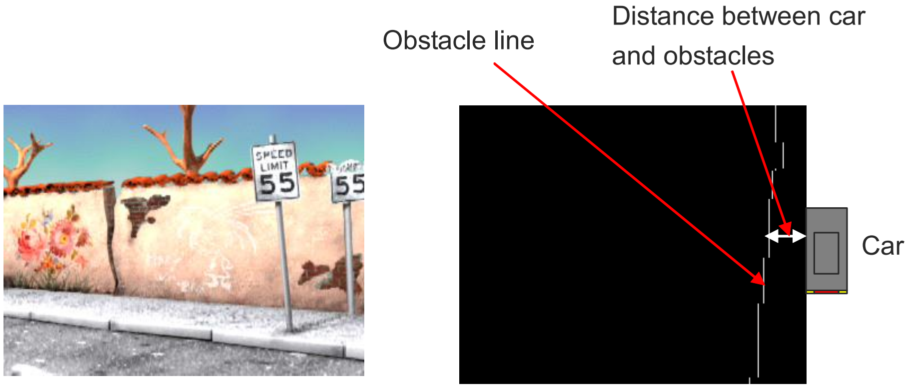
After the per-pixel disparity vectors are generated, we convert the disparity vectors to depth map by applying the formula:

$$Z = \frac{fB}{|x_1 - x_2|} \tag{2}$$

in which B means baseline, f denotes the focal length, Z is the depth, and —x2-x1— is the disparity. So the depth map is easily made.

5. EXPERIMENTAL RESULTS

Fig.9 shows the result of our algorithm by synthetic images. The right most edge of the radar display is the position of a car and the white line with several segments is the obstacle line which represents our surroundings, thus, it is the wall in the original frame. A real trinocular camera is setup as Fig.10. The three cameras are Logitech QuickCam Pro 4000 and separated 20cm with each other. The captured images from the real trinocular camera is shown in Fig.11. After the STDDE computation and the depth map generation, the calculted depth map is shown in Fig.12. The depth map of the relation between the people and the background are presented clearly. This also proves that the proposed algorithms work well even when the illumination varies on the each camera. The symmetric trinocular property helps to consider more on the correspondense rather than the intensity. The computational speed achieves 30FPS on a Pentium-4 2.4GHz computer since the 1D fast search really reduces the search candidates and find correct disparity vectors.



Original frame

Radar display

Figure 9. The result and the test sequence for the car surrounding camera array

Figure 10. The trinocular camera combination

Figure 11. The captured images from the real trinocular camera

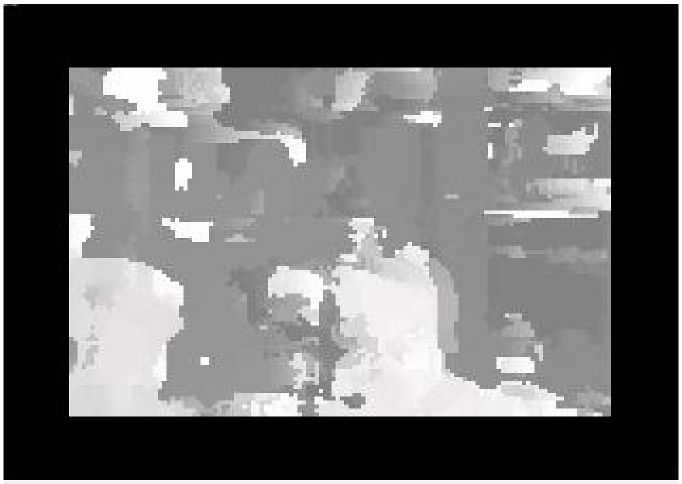


Figure 12. The calculated depth from the real trinocular camera

6. CONCLUSIONS

The proposed symmetric trinocular dense disparity estimation (STDDE) proves a much more accurate and faster algorithm than the traditional dense disparity algorithm. It is also reliable and robust enough. The experimental results show that the STDDE derives a useful application such as car surrounding camera array which detects danger on road.

ACKNOWLEDGMENTS

Acknowledgements to NSC foundation for the support of the experiment devices.

REFERENCES

1. W. van der Mark and D. M. Gavrila, "Real-time dense stereo for intelligent vehicles," *IEEE Transactions on Intelligent Transportation Systems* **7**, Mar. 2006.
2. D. J. Fleet, "Disparity from local weighted phase-correlation," *IEEE International Conference on Systems, Man, and Cybernetics, 1994. 'Humans, Information and Technology'* **1**, pp. 48–54, Oct. 1994.
3. N. Grammalidis and M. G. Strintzis, "Disparity and occlusion estimation in multiocular systems and their coding for the communication of multiview image sequences," *IEEE Transactions on Circuits and Systems for Video Technology* **8**, pp. 328–344, Jun. 1998.
4. M. H. Ouali, D. Ziou, and C. Laurgeau, "Dense disparity estimation using gabor filters and image derivatives," *Proceedings. Second International Conference on 3-D Digital Imaging and Modeling, 1999.* , pp. 483–489, Sep. 1999.
5. C. Strecha and L. V. Gool, "Pde-based multi-view depth estimation," *Proceedings of the First International Symposium on 3D Data Processing Visualization and Transmission* , 2002.
6. L. Oisel, E. Memin, L. Morin, and F. Galpin, "One-dimensional dense disparity estimation for three-dimensional reconstruction," *IEEE Transactions on Image Processing* **12**, pp. 1107–1119, Sep. 2003.
7. X. Huang and E. Dubois, "Three-view dense disparity estimation with occlusion detection," *International Conference on Image Processing* , 2005.
8. X. Huang and E. Dubois, "Two-dimensional regularized disparity estimation based on the gabor transform," in *Visual Communications and Image Processing 2006*, J. G. Apostolopoulos and A. Said, eds., *Proc. SPIE* **6077**, 2006.