Based on Auto Comfortable Disparity Adjustment Real-Time 3-D Generation System

Hong-Peng Lee, Yi-Ming Chung, and *Pei-Jun Lee

Department of Electrical Engineering, National Chi Nan University

HongPeng617@gmail.com

ShoPeter@gmail.com

*Advisor

Abstract— **This paper use colour image and its depth information to generate virtual images in different viewpoint. There is a distance sensor to detect the distance form viewer to displayer which is one of parameters for generating virtual images. The comfort of watching 3-D image could be enhanced by changing the parameters. All the system is implemented on FPGA.**

***Keywords*— FPGA; DIBR; View distance; Watching comfort; Disparity adjustment**

I. INTRODUCTION

The sense of watching 3-D is combing two images captured from left-eye and right-eye in brain. In other word, let left-eye and right-eye watch relative image would have the sense of 3-D. The scene watching in left-eye and right-eye is different in horizontal and the difference is according to the distance form viewer to objects. The disparity of objects in left-eye and right-eye is bigger when the objects are closer to viewer while the disparity is smaller when the objects are farer from viewer. It is easy to figure out when close right-eye and left-eye in turns to watch.

Placing two cameras in parallel could generate simulated images as human visual system. However, it cost more resources. Viewer will not always watch image at same position. When the position of viewer and cameras is different the sense of 3-D will not match real situation. Therefore, utilize an image and its depth information which is called depth map to generate virtual views in different viewpoints is a popular method in recent years.

This paper use FPGA to fulfil the generation for multi-view virtual images from an image and its depth map. The interface of input and output is DVI. Place original image in left-hand side and its depth map in right-hand side as the input of DVI. There is a distance sensor to detect the distance form viewer to displayer. Furthermore, adjust the quality of virtual images. The disparity in virtual images would be increased when viewer stand close to displayer and the disparity would be decreased when the viewer is far from displayer.

II. RELATIVE WORK

Objects have horizontal disparity in left-eye and right-eye. Objects capturing in left-eye is closer to right-hand side than right-eye. Moreover, the shift will be more obvious when objects are more close to viewer. 3-D virtual image could be generated by utilizing this phenomenon.

A. Depth Map

The distance from viewer to objects is presented in grey-scale value. The range of grey-scale value is from 0 to 255. Smaller value presents longer distance and bigger value presents shorter distance. It could be seen in Fig. 1.

Fig.1 A colour image and its corresponding depth map

B. The Generation of Virtual Image

As mentioned in section I, the sense of 3-D is due to the combination of images captured from left-eye and right-eye in brain. Depth Image Based Rendering (DIBR) is a technique combing single-view image and its depth map to generate multi-view images which is suitable for human visual system.

C. Multi-View Images

People will not always watch video at the same position. When the position of viewer and displayer is different comparing with cameras and the scene which generate video in the displayer, the sense of 3-D is not match real-situation. That is the motivation for several different virtual images in different viewpoint.

D. DVI Signal Protocol

The signals needing process in DVI protocol are as below:

- HS: This signal stays in high-level until a line of a frame is finished transition. Low-level is the condition to wait synchronization.
- VS: This signal stays in high-level until a frame is finished transition. Low-level is the condition to wait synchronization.

- Data: It is the information of a pixel in a frame containing red, green and blue data. Each colour is delivered in parallel at the same time and presented in 12 bits.
- Clock: An oscillator for data synchronization.
- Data Enable (DE): Data could be only transited or delivered when DE is high-level.

E. Distance Sensor

The sensor is used to obtain the distance between the viewer and displayer as shown in Fig. 2. Utilize the distant parameter to adjust the quality of 3-D image. The sensor sends a pulse signal and stop when receive the echo. According to the interval between sending pulse and receiving echo the distance between viewer and displayer could be calculated.

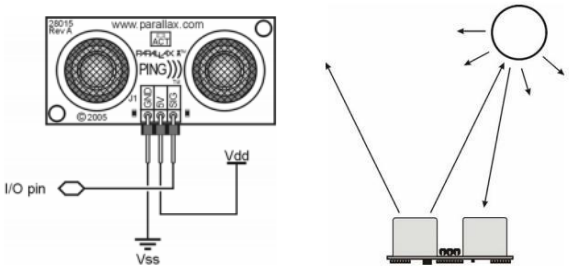

Fig. 2 The chip of distant sensor and its working method

III. Formula Deriving And Simulation

A warping formula is proposed in this paper to achieve more comfortable 3-D watching. The 3-D image would have positive disparity and negative disparity by applying this formula. 3-D visual system with negative disparity means the objects would project in front of the displayer in 3-D display while positive disparity means the objects projecting behind the displayer as shown in Fig. 3.

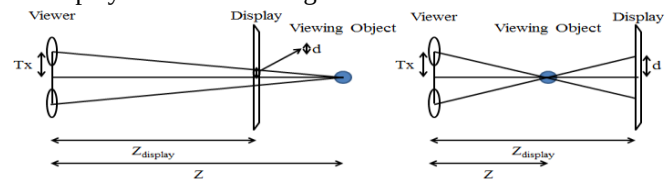

Fig. 3 The positive disparity shown in left-hand side and negative disparity shown in right-hand side

The formula refers to the distance from viewer to displayer is shown in formula (1).

$$\begin{cases} \text{right view}: x_r = x + d \\ \text{left view}: x_l = x + d \end{cases}, \quad d = \frac{Tx}{2}\frac{\left(Z - Z_{display}\right)}{Z} \tag{1}$$

Where x is the x-axis of the center pixel in central view, x_r and x_l is the corresponding x-axis of the pixel in the right-view and left-view respectively. T_x is a baseline distance or the distance between cameras. $Z_{display}$ is the distance from viewer to displayer. For the better watching comfort, this paper uses formula (2) to obtain Z value which presents the distance from viewer to the objects.

$$Z = \frac{1}{\frac{depth}{255}\left(Z_{near}^{-1} - Z_{far}^{-1}\right) + Z_{far}^{-1}} \tag{2}$$

Where Z_{near} is the nearest distance and Z_{far} is the farthest distance presenting in the depth map respectively.

The proposed system utilizes Look up Table (LUT) for depth to disparity table instead of calculating warping formula to reduce computing time. Calculate the correspond disparity of each depth and write into LUT.

The reference pixel is shifted to target position by the corresponding disparity. To avoid overwrite the foreground objects by background objects when doing warping. Before warping the reference pixel, the depth value in Z buffer is compared to prevent the foreground pixels filling by background pixels. If the depth value of current pixel is larger than warped depth value, the depth value will be updated and the address will pass to row buffer to warping the pixel. Otherwise, the depth value and color pixel pair is abandoned.

The correlation of pixel's disparity and the distance from viewer to displayer is close. Fix $Z_{display}$ will cause uncomfortable watching when viewer is changing viewing distance. Therefore, adjust $Z_{display}$ in real-time and update the disparity to enhance to viewing comfort is proposed.

There are numerous holes after warping which are lost information. Holes are needed to fill by using the information from neighbor pixels. Consider hardware structure the holes are filled by the average of left-pixel and right-pixel which are not holes.

IV. Implement on Hardware

The real-time 2-D to 3-D comfort generation in this paper is utilizing DVI as the interface for input image and output image. Each module in this system is processing according to the synchronization signal of DVI.

The system could divide into "distance detection and LUT updating" and "3-D generation" two main modules. Distance detection and LUT updating using ultrasound to detect the distance and adjust $Z_{display}$ in formula (2). Finally re-calculate the disparity from its corresponding depth to update the LUT for 3-D generation. 3-D generation module utilizing DIBR combing colour image and depth map to generate side-by-side 3-D image with the help of LUT.

A. Distance Detection and LUT Updating

Coordinate the synchronization signal of DVI LUT updating optimal after the end of a frame. Use two LUTs to avoid read and write at the same time. One LUT for read and the other for write. Change LUTs after a frame finished. Timing diagram is shown as Fig. 4.

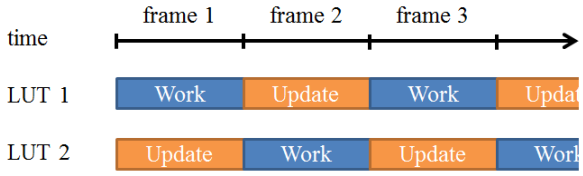

Fig. 4 Timing diagram of LUT

The sensor will detect the distance to determine Z_{zero} before updating LUT and then calculate each depth's corresponding disparity to write into LUT shown as Fig. 5.

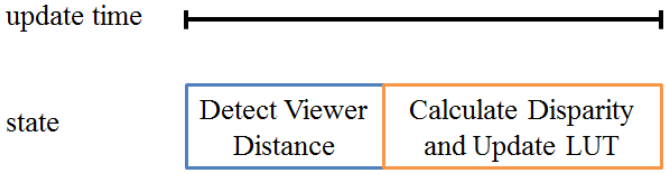

Fig. 5 The processing of LUT

B. Distance Detection

The sensor used for distance detection is PARRALLEX Pin ultrasound distance sensor. Distance detection could be done by a single pin shown as Fig. 6. The sensor will send a pulse to detect objects and generate a square wave at the same time. The pin will stay high-level after send a pulse and drop to low-level when receive echo.

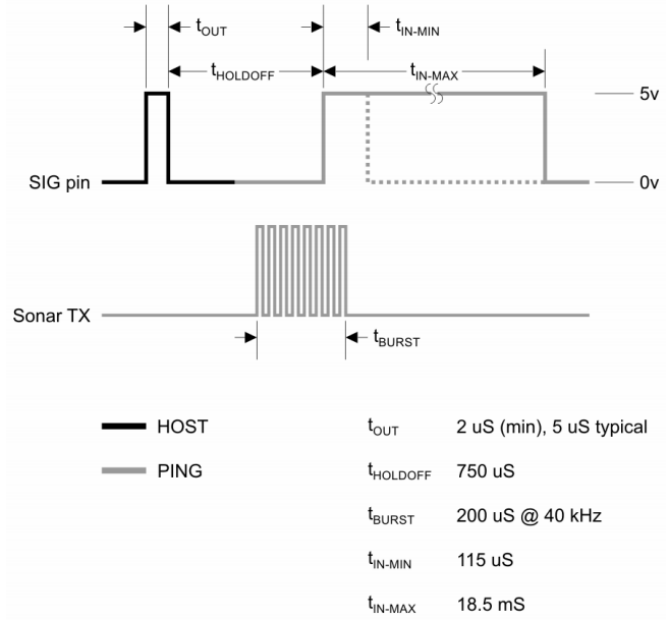

Fig. 6 The pin for sending pulse and detecting distance

The position could be calculated by using the formula of sonic speed shown as formula (3).

$$d = t_{in} \times c \tag{3}$$

$$c = 331 + 0.6T \tag{4}$$

Where d is the distance from viewer to sensor, T is the degree in Celsius and t_{in} is the time between sending pulse and receiving echo presented in micro second. A unit of distance is set as 10 centimetres.

To reduce the computing time it could be found that every 10 centimetres takes 0.288 micro-second in sonic speed. Then calculate the positive edges with a cycle of 0.288 micro-second. Multiply the times of positive edge and 10 centimetres could obtain the viewer distance $Z_{display}$ as shown in Fig. 7.

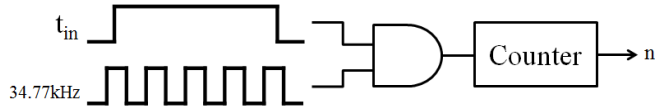

Fig. 7 Calculating each 10 centimetres

In the last, update LUT from warping formula. First refer the correspond Z of depth value from depth-to Z LUT then calculate the disparity from formula (1) and write into the LUT of depth-to-disparity.

To achieve a real time 2D to 3D conversion system based on DIBR technique, the proposed architecture is designed on FPGA. DIBR technique includes warping and hole-filling processing. To improve the memory utilization, row buffer is used for I/O processing. There are two set of modules for odd row and even row. Each row is processed by data receiving module (RX), warping module (W), and hole filling module (HF) then the result will output synthesis images. The hardware structure is shown in Fig. 8. The working timing schedule of each module is shown in Fig. 9.

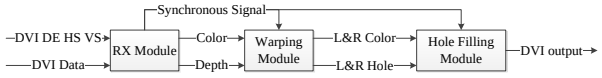

Fig. 8 Hardware structure of DIBR

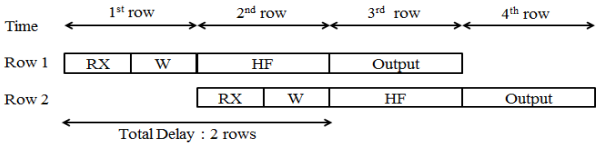

Fig. 9 Timing diagram of each DIBR module

C. Row Buffer Access Designed for RX Module

In this paper, RX module is designed for I/O timing synchronization, data buffering and repackage the data to other modules. All the data of a frame is transmitted pixel-by-pixel with Z-scan by DVI I/O interface as shown in Fig. 10. After I/O timing synchronization, RX module is able to control the timing of other module as shown in Fig. 11. RX module stores the color information and is combined with its corresponding depth value, then passes the data as a color and depth pair to warping module.

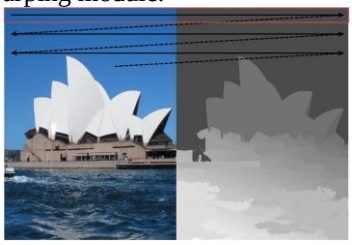

Fig. 10 Data transition in DVI I/O interface with Z-scan

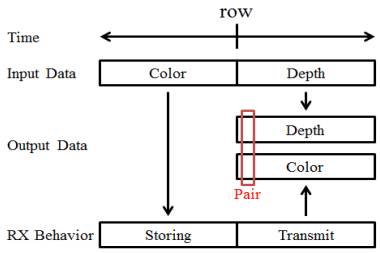

Fig. 11 Timing diagram of RX module

A pixel would be compared the warped pixel's depth value in Z buffer to avoid background object overwrite foreground object as shown in Fig. 12.

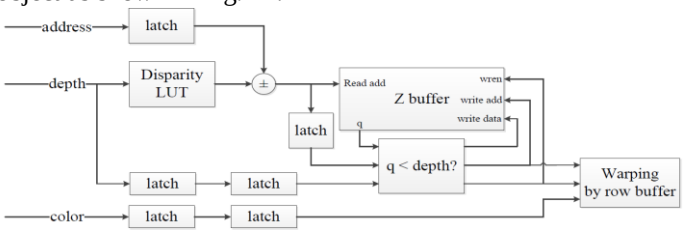

Fig. 12 Determine a pixel should warp or not

D. Proposed Hole-filling Module for Virtual Views Images

After rendering the virtual views, there is occlusion area, the hole, in the virtual views at the discontinuous area of depth. To improve the virtual view quality, the proposed architecture uses the average of the two non-hole pixel values at the neighbour holes to fill the hole.

The proposed architecture uses a Left Side Value (LSV) buffer to store the left side pixel value of holes by a forward scan register as Fig. 13 (a). And then the right side pixel value can took to average with LSV values to fill the hole by a backward scan, we latch the right side pixel value to take average with the left side pixel value in the buffer to fill the hole a backward scan register as Fig. 13(b).

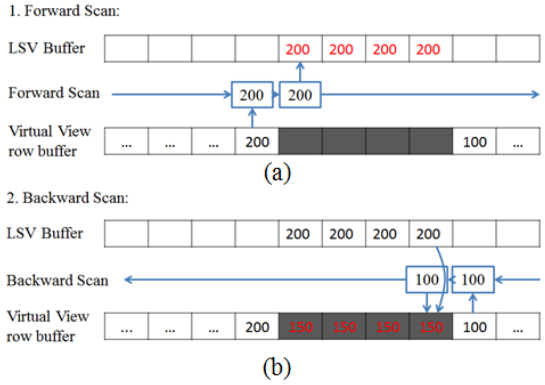

Fig. 13 (a) Search holes in forward scan (b) Search holes in backward scan

V. Implement Result

This paper uses Altera DE3-260 FPGA with a DVI I/O daughter board as hardware implementation platform as shown in Fig. 14 (a). Fig. 14 (b) and Fig. 14 (c) shows the input image and output side-by-side 2-views 3D virtual rendered images, respectively. Fig. 15 shows the hardware usage. The proposed system has 17% logic utilization of FPGA broad. Form the results, the proposed system can achieves real time 2D to 3D conversion. The simulation result of other pattern is shown in Fig. 16.

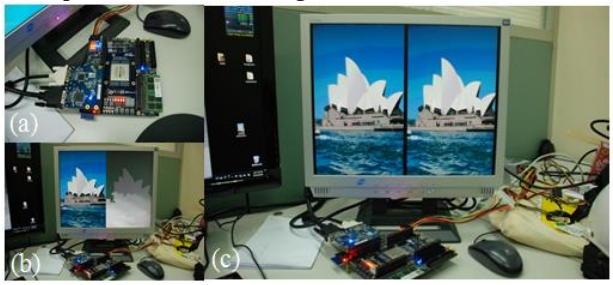

Fig. 14 (a) DVI daughter board (b) Input of DVI (c) Simulation result

Flow Summary	
Flow Status	Successful - Fri May 17 01:18:57 2013
Quartus II 64-Bit Version	10.0 Build 262 08/18/2010 SP 1 SJ Full Version
Revision Name	DVI_Demo
Top-level Entity Name	DVI_Demo
Family	Stratix III
Device	EP3SE260F1152C2
Timing Models	Final
Met timing requirements	N/A
Logic utilization	17 %
Combinational ALUTs	17,259 / 203,520 (8 %)
Memory ALUTs	0 / 101,760 (0 %)
Dedicated logic registers	24,720 / 203,520 (12 %)
Total registers	24720
Total pins	90 / 744 (12 %)
Total virtual pins	0
Total block memory bits	981,716 / 15,040,512 (7 %)
DSP block 18-bit elements	0 / 768 (0 %)
Total PLLs	2 / 8 (25 %)
Total DLLs	0 / 4 (0 %)

Fig. 15 The used resources of FPGA

Fig. 16 Simulation result of dog pattern

VI. Conclusion

The comfort of watching 3-D image is an important concern for the popular of 3-D visual system. It is not acceptable if the 3-D effect is not obvious or comfortable. That is the reason that an adaptive 3-D visual system is proposed in this paper.

A 3-D image is adjusted real-time according to the position of the viewer to enhance the 3-D watching comfort in different viewpoints.

The use of LUT reduced lot of computing time to reach real-time 3-D generation and the use of registers.

REFERENCES

[1] C. Fehn, “Depth-Image-Based-Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV,” in Proc. International Society for Optical Engineering Conf. Stereoscopic Displays and Virtual Reality Systems XI, vol. 5291, pp. 93–104, CA, U.S.A. Jan. 2004.

[2] I.Andorko, P.Corcoran and P.Bigioi,” Hardware implementation of a real-time 3D video acquisition system,” International IEEE Consumer Electronics Society's Games Innovations Conference, ICE-GIC 2009, pp.239 – 245, 2009.

[3] Yu-Cheng Fan, Yu-Ting Kung and Bing-Lian Lin, ”Three-Dimensional Auto-Stereoscopic Image Recording, Mapping and Synthesis System for Multiview 3D Display” IEEE Transactions on Magnetics, pp.683-686, 2011.

[4] Hsueh-Ming Hang and Du-Hsiu Li, “Sparse Multi-Camera Virtual View Synthesis Using Forward and Backward Depth Warping Algorithms”, thesis of National Chiao Tung University, 2012.