

Naked-eye 3D Video Display using Adaptive Support-Weight Algorithm with Variable Support Window

Zhichao Yang¹, Guanglin Yang^{*1}, Haiyan Xie²

¹Laboratory of Advanced Optical Communication System and Network, School of Electronics Engineering & Computer Science, Peking University, Beijing, China, 100871

²China Science Patent Trademark Agents Ltd., Beijing, China, 100083

*ygl@pku.edu.cn

Abstract- A naked-eye three-dimensional (3D) video real-time display system is proposed. In this system, the cameras number of capturing 3D information were reduced from nine to two, and the naked-eye 3D images multi-view display was realized using the 2 cameras. The disparity image and the left image from stereo views were applied to generate the intermediate seven virtual viewpoint images along the horizontal baseline. The Adaptive Support-Weight (ASW) algorithm with Variable Support Window (VSW) was applied to obtain the disparity image. The depth-image-based rendering and image interlacing were adopted to get a multi-view image (i.e. 3D information). The screen of integral-imaging using lenticular lens was applied to display 3D images with continuous and wider viewing zone. The high speeding parallel programming architecture CUDA was applied to realize the high-quality naked-eye 3D video real-time display. Experimental results showed that the stereo-matching algorithm was suitable for the proposed system.

Keywords- Naked-eye 3D; Virtual Viewpoint Images; Image-based Rendering; Real-time Display

I. INTRODUCTION

In three-dimensional (3D) video systems, the viewers need to wear a pair of polarized glasses (or shutter, anaglyph glasses) to enjoy the 3D video, and that causes great inconvenience. Currently, the stereoscopic 3D displays include two methods, the first requires polarization glasses, only in two-view manner, and provides high-quality 3D images in a simple manner. It is widely used in movie theatres and some broadcasting TV. The second method is auto-stereoscopic 3D display, including two-view manner and multi-view manner; and it is more comfortable for viewing. Towards the multi-view manner, the most widely and currently used method is based on both the principle of parallax barrier and the one of lenticular lens.

In this research, a naked-eye 3D video real-time display system as shown in Fig. 1 was established. The system consists of three parts, responsible for 3D information collecting, transmission encoding/decoding, and depth-image-based rendering and image interlacing, respectively. For the 3D information collecting, two cameras were applied to get the left image and the right image, and the Adaptive Support-Weight (ASW) algorithm [5] with Variable Support Window (VSW) [6] was applied to produce the disparity images. Graphic Processing Unit (GPU) was employed to improve the disparity image computation speed, as shown in Fig. 1. The second part is responsible for transmission encoding/decoding, where H.264/AVC was used to compress the 3D information, as shown in Fig.1. H.264/AVC (Advanced Video Coding) is a video compression format that is one of the most commonly used video compression formats for recording, compression and transmission [22]. The last part is responsible for depth-image-based rendering and image interlacing. The depth-image-based rendering was adopted to generate the other seven virtual viewpoint images. One multi-view image (i.e. 3D information) was synthesized by the image interlacing from the nine viewpoint images (i.e. the 7 virtual viewpoint images plus the left image and right image) as shown in Fig.1.

In naked-eye 3D video systems, the major bottleneck is the generation of real-time high-quality dense disparity images [9-11]. In experiments, the stereo colour binocular camera has been employed to obtain stereo information. In order to satisfy the multi-view rendering and reduce the pixel data that will be encoded and transmitted, stereo matching algorithm was adopted to generate disparity images. In order to realize naked-eye 3D image real-time display, as shown in Fig. 1, the high speeding parallel programming architecture CUDA was applied to improve the naked-eye 3D video beam, and to display the naked-eye 3D video on the special screen [7, 8, 19-22].

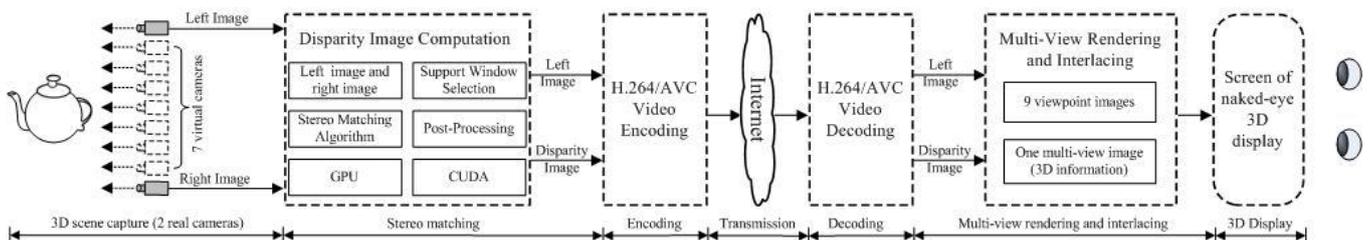


Fig. 1 A naked-eye 3D video real-time display system

In Section I, the research background of 3D displays and the system structure of the naked-eye 3D video real-time display system were introduced. In Section II, the reduction of the number of cameras from nine to two is detailed. In Section III, the Adaptive Support-Weight algorithm with Variable Support Window is described. In section IV, the multi-view rendering and interlacing algorithm are described. In section V, the naked-eye 3D display screen is described. In section VI, the experiment results and analysis are described. In section VII, the conclusions are made.

II. REDUCING NUMBER OF CAMERAS

In naked-eye 3D video real-time display systems, one challenge is the difficulty to capture and transfer data of the video streams from multiple cameras. In order to solve this difficulty, the number of cameras was reduced from nine to seven, with two remaining cameras that can generate the intermediate seven virtual viewpoint images between the left viewpoint image and the right viewpoint image, as shown in Fig.2. In experiments, it was demonstrated that the nine actual cameras can be replaced by the 2 cameras plus the seven virtual viewpoint images [1-4].

In the proposed system, two digital cameras were adopted to get the left image and the right image, and then the Adaptive Support-Weight (ASW) algorithm [5] with Variable Support Window (VSW) [6] was adopted to generate the disparity images. The depth-image-based rendering was applied to generate the seven virtual viewpoint images. The seven virtual viewpoint images plus the left image and the right image equal the nine images, as shown in Fig. 2.

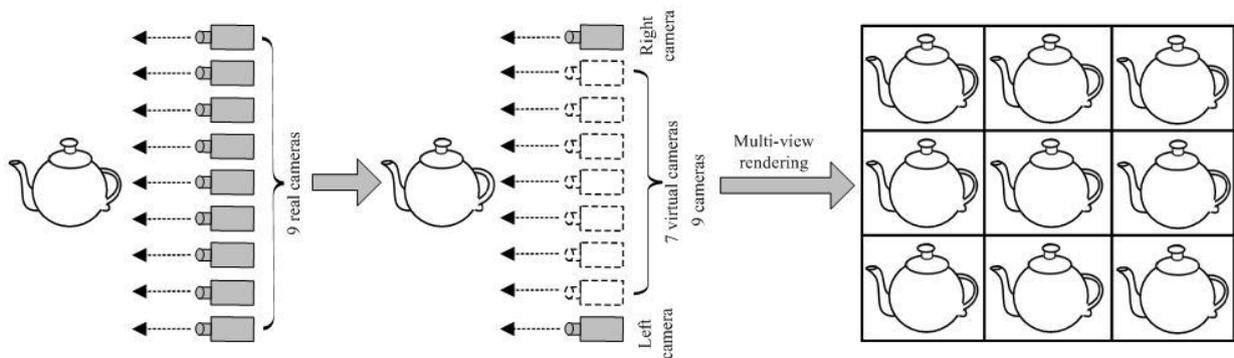


Fig. 2 Reduction of the nine cameras to capture 3D information to two cameras, and application of depth-image-based rendering to generate the seven virtual viewpoint images, to get the nine viewpoint images

III. ADAPTIVE SUPPORT-WEIGHT ALGORITHM WITH VARIABLE SUPPORT WINDOW

Considering the stereo matching speed, accuracy and structure complexity, ASW algorithm was applied to obtain the disparity images. However, in ASW, the matching window size is fixed. The mismatch is relatively larger in the discontinuous region of the disparity images. In this case, VSW was applied to judge the support pixels before stereo matching [6]. The mismatch rate of the discontinuous regions of the disparity images was reduced, as shown in Figs. 3 and 4.

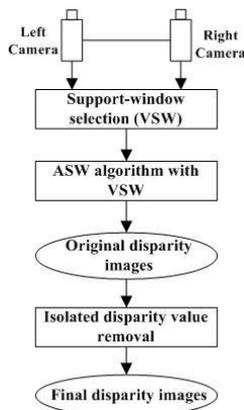


Fig. 3 Flowchart of ASW algorithm with VSW

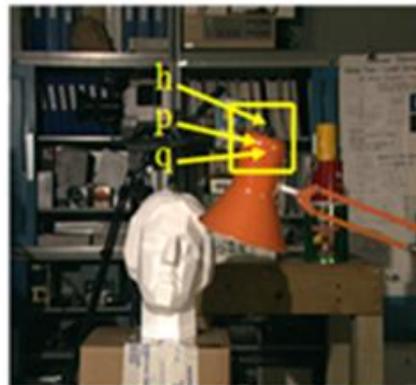


Fig. 4 Judgment of the support pixel q using colour similarity

In stereo matching, according to the Gestalt Psychology Theory [5, 12], the main visual perceptual principles includes both the colour similarity and the geometric proximity. In the scheme of this study, colour similarity was applied to judge the support pixel q , as shown in Fig. 4. When the pixel p is given, if the pixel q is similar to p , then q is the support pixel. On the other hand, if the pixel h is not similar to the pixel p , h is not the support pixel [13]. In calculation, $p_r(q_r)$ represents the Red component of the pixel $p(q)$, $p_g(q_g)$ represents the Green component of the pixel $p(q)$, and $p_b(q_b)$ represents the Blue component of the pixel $p(q)$, thereby the colour distance $f_c(p, q)$ between two pixels (i.e. p and q) was defined as follows [5]:

$$f_c(p, q) = |p_r - q_r| + |p_g - q_g| + |p_b - q_b| \quad (1)$$

Based on Eq. (1), Eqs. (2) and (3) can be obtained,

$$f_c(p_L, p_R) = |p_{Lr} - p_{Rr}| + |p_{Lg} - p_{Rg}| + |p_{Lb} - p_{Rb}| \quad (2)$$

$$f_c(q_L, q_R) = |q_{Lr} - q_{Rr}| + |q_{Lg} - q_{Rg}| + |q_{Lb} - q_{Rb}| \quad (3)$$

In Eqs. (2) and (3), $p_L(q_L)$ represents the $p(q)$ pixel in the left image, $p_R(q_R)$ represents the $p(q)$ pixel in the right image; $p_{Lr}(q_{Lr})$ represents the Red component of pixel $p_L(q_L)$, $p_{Lg}(q_{Lg})$ represents the Green component of pixel $p_L(q_L)$, $p_{Lb}(q_{Lb})$ represents the Blue component of pixel $p_L(q_L)$; $p_{Rr}(q_{Rr})$ represents the Red component of pixel $p_R(q_R)$, $p_{Rg}(q_{Rg})$ represents the Green component of pixel $p_R(q_R)$, $p_{Rb}(q_{Rb})$ represents the Blue component of pixel $p_R(q_R)$; $f_c(p_L, p_R)$ represents the colour distance between p_L and p_R , and $f_c(q_L, q_R)$ represents the colour distance between q_L and q_R .

According to the Gestalt Psychology Theory [5], Eq. (4) can be obtained from Eqs. (2) and (3),

$$\Delta f_c = |f_c(p_L, p_R) - f_c(q_L, q_R)| \quad (4)$$

In Eq. (4), Δf_c represents the colour similarity between the colour distance $f_c(p_L, p_R)$ and the colour distance $f_c(q_L, q_R)$, and was used to judge whether pixel q is the support pixel of the given pixel p . In order to achieve the final disparity, the truncated value T was chosen to compare with Δf_c . If $\Delta f_c \leq T$, pixel q_L is the support pixel of the given pixel p_L , and pixel q_R is the support pixel of the given pixel p_R . If $\Delta f_c > T$, pixel q_L is not the support pixel of the given pixel p_L , and pixel q_R is not the support pixel of the given pixel p_R .

After the support pixels are selected, the support pixels can be discontinuous, and the support-window obtained will be variable in its shape. In the ASW algorithm, both the reference support window N_{pd} and the target support window N_p are square, as shown in Fig. 5(a). In the proposed method, both the reference support window N_{pd} and the target support window N_p are variable in shape, as shown in Fig. 5(b).

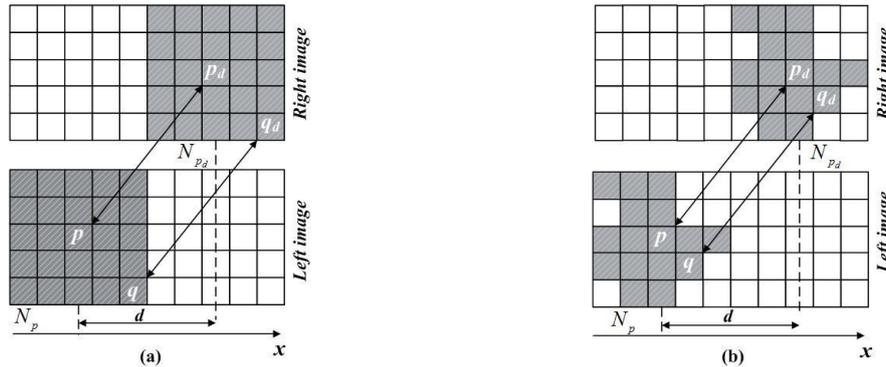


Fig. 5 (a) Reference support window N_{pd} and target support window N_p in ASW algorithm, (b) Reference support window N_{pd} and target support window N_p in the ASW algorithm with VSW

According to the ASW algorithm [5], the support-weights were calculated by Eqs. (5) and (6),

$$w(p, q) = \exp\left(-\left(\frac{\Delta c_{pq}}{\gamma_c} + \frac{\Delta g_{pq}}{\gamma_p}\right)\right) \quad (5)$$

$$w(p_d, q_d) = \exp\left(-\left(\frac{\Delta c_{p_d q_d}}{\gamma_c} + \frac{\Delta g_{p_d q_d}}{\gamma_p}\right)\right) \quad (6)$$

In the following equations, Δc_{pq} (or $\Delta c_{p_d q_d}$) represents the Euclidean distance of the colour similarity between the support pixel q (or q_d) and the given pixel p (or p_d), and Δg_{pq} (or $\Delta g_{p_d q_d}$) is the Euclidean distance of the geometry proximity between the support pixel q (or q_d) and the given pixel p (or p_d). The values of γ_c and γ_p are empirical factors that are used to modify the colour and space similarity, respectively, and γ_p is half of the common window size, $p(i_p, j_p)$, $q(i_q, j_q)$ [5].

$$\Delta c_{pq} = \sqrt{(p_r - q_r)^2 + (p_g - q_g)^2 + (p_b - q_b)^2} \quad (7)$$

$$\Delta c_{p_d q_d} = \sqrt{(p_{dr} - q_{dr})^2 + (p_{dg} - q_{dg})^2 + (p_{db} - q_{db})^2} \quad (8)$$

$$\Delta g_{pq} = \sqrt{(i_p - i_q)^2 + (j_p - j_q)^2} \quad (9)$$

$$\Delta g_{p_d q_d} = \sqrt{(i_{p_d} - i_{q_d})^2 + (j_{p_d} - j_{q_d})^2} \quad (10)$$

Through computation, the support window was selected, and then the matching cost $E(p, p_d)$ between pixel p and pixel p_d were derived by Eq. (11) [5],

$$E(p, p_d) = \frac{\sum_{q \in N_p, q_d \in N_{p_d}} w(p, q)w(p_d, q_d)e(q, q_d)R_s(p, q)}{\sum_{q \in N_p, q_d \in N_{p_d}} w(p, q)w(p_d, q_d)} \quad (11)$$

where $e(q, q_d)$ is expressed as

$$e(q, q_d) = f_c(q, q_d) = |q_r - q_{dr}| + |q_g - q_{dg}| + |q_b - q_{db}| \quad (12)$$

Through the matching cost computation, the disparity of each pixel was selected by WTA (Winner-Takes-All) method [5] as below,

$$d_p = \arg \min_{d \in D} E(p, p_d) \quad (13)$$

where $D = \{d_{min}, \dots, d_{max}\}$ is the set of all possible disparity values.

After all pixels' original disparities were calculated, a filtering method was adopted to reduce the isolated disparity values with wrong correspondences. The details of the filtering method are as follows: Four neighbouring pixels around the given pixel p were considered, including the left pixel, the right pixel, the upper pixel and the lower pixel. If the disparity value of the left pixel was equal to that of the right pixel, the disparity value of the given pixel p was equal to that of the left pixel. If the disparity value of the upper pixel was equal to that of the lower pixel, the disparity value of the given pixel p was equal to that of the upper pixel. In other cases, the disparity value of the given pixel p remained unchanged. After the isolated disparity values were reduced, the final disparity image that can be encoded/decoded using H.264/AVC could be obtained.

IV. MULTI-VIEW RENDERING AND INTERLACING

Depth-image-based rendering and image interlacing are the key technologies to realize naked-eye 3D video display [14-16]. In the proposed naked-eye 3D video display system, depth-image-based rendering was adopted to produce the other seven virtual viewpoint images, and one multi-view image (i.e., 3D information) was synthesized using interlacing from the nine viewpoint images, which are the 7 virtual viewpoint images plus the left image and the right image.

As shown in Fig. 6, c_1, c_2, \dots, c_n are the positions of the virtual cameras, which can be created at a specific baseline distance. The total distance between the virtual cameras is t . Moreover, P is a point of interest in the scene. The focal length of cameras is f , and the depth value is Z . x_1, x_2, x_3, \dots and x_n are the horizontal coordinates of the view positions, respectively [14].

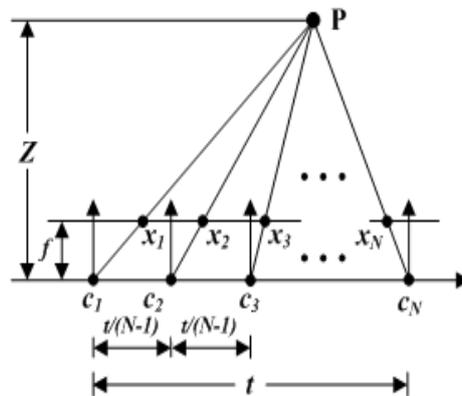


Fig. 6 Top-view of camera configuration for generating the virtual viewpoints

According to the geometrical relation of the camera configuration, the equation (14) can be got as follows:

$$x_m = x_{m-1} + \left(\frac{t}{N-1} \cdot \frac{f}{Z} \right), m = 2, 3, \dots, N \quad (14)$$

From Eq. (14), Eqs. (15) ~ (17) were derived,

$$x_{m-1} = x_{m-2} + \left(\frac{t}{N-1} \cdot \frac{f}{Z} \right), m = 3, 4, \dots, N \quad (15)$$

$$x_{m-2} = x_{m-3} + \left(\frac{t}{N-1} \cdot \frac{f}{Z} \right), m = 4, 5, \dots, N \quad (16)$$

$$x_2 = x_1 + \left(\frac{t}{N-1} \cdot \frac{f}{Z} \right) \quad (17)$$

From Eqs. (14) ~ (17), Eq. (18) was derived,

$$x_m = x_1 + (m-1) \left(\frac{t}{N-1} \cdot \frac{f}{Z} \right), m = 2, 3, \dots, N \quad (18)$$

For disparity $d=f \times t/Z$, Eq. (19) was derived,

$$x_m = x_1 + (m-1) \frac{d}{N-1}, m = 2, 3, \dots, N \quad (19)$$

Based on Eq. (19), the new viewpoint images were obtained. However, some holes existed in the disparity images due to occlusions. The hole-filling algorithm was adopted to fill the holes in order to get high-quality viewpoint images [17] that can improve the naked-eye 3D image display quality.

After the seven virtual viewpoint images were generated and the hole-filling was implemented, image interlacing was adopted, and the pixels of the nine viewpoint images were rearranged to fit the optical structure of 3D display screen in the 9 directions. In the proposed naked-eye 3D video display system, the sub-pixels of the interlaced multi-view images were selected from the RGB sub-pixels of the corresponding nine viewpoint images [14], as shown in Fig. 7(a).

V. SCREEN OF NAKED-EYE 3D DISPLAY

A naked-eye 3D display screen that has the nine optimum viewpoint perspectives to display the depth information, was adopted. Therefore, each pixel of the naked-eye 3D display screen in the multi-viewpoint image has the information of nine directions. The screen mechanism of the naked-eye 3D video display for one pixel in one direction is shown in Fig. 7(b). The nine viewpoint information of the pixel are represented as the 1st, 2nd, ... 9th elements, and the RGB (Red, Green, Blue) components distribution of the naked-eye 3D display screen's pixel is shown in Fig. 7(a).

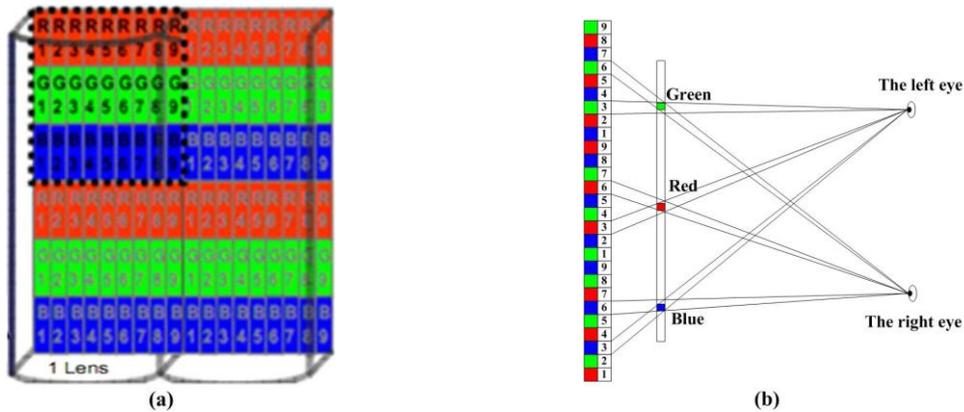


Fig. 7(a) RGB components distribution for one pixel in one direction, (b) The screen principle of naked-eye 3D video display for one pixel

In one direction, the right eye and the left eye of a man will see two different viewpoints' RGB information of the pixel at the same time, thereby he can enjoy the impression of depth. In the other eight different directions, he will also experience different impressions of depth. In the proposed naked-eye 3D video real-time display system, in order to get better 3D display effect, the integral imaging auto-stereoscopic 3D Display technique was applied. The technique provides a wider viewing zone and a shorter viewing distance. As shown in Fig. 8(a), the number of views is more than nine, and the viewing zone is continuous and wider. As seen in Fig. 8(c), the number of views is more than two, and the viewing zone is discrete viewpoint (not zone). There is no "flipping" in Fig. 8(a), and the natural visibility is smooth motion parallax, while in Fig. 8(c), the flipping makes user feel tiresome and the natural visibility is unsmooth motion parallax. Moreover, due to 3D effect, the 3D resolution in Fig. 8(a) decreases, while the 3D resolution in Fig. 8(c) does not change. Fig. 8(b) shows the multi-view (parallax barrier); the "L"s of the LCD represent the left image pixels, and the "R"s represent the right image pixels.

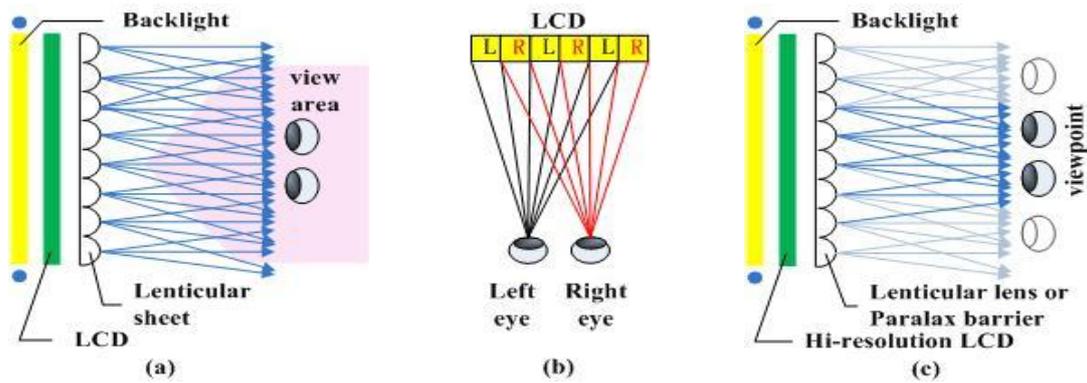


Fig. 8 (a) Integral imaging 3D display using lenticular lens, (b) Multi-view 3D display using parallax barrier, (c) Multi-view 3D display using lenticular lens

The integral imaging auto-stereoscopic 3D display method in Fig. 8(a) has several advantages. Firstly, it provides continuous and wider viewing zone (32°), while other multi-view methods have discrete viewpoints (not viewing zone). Secondly, as for the natural visibility, there is no “flipping” in the viewing zone and the motion parallax is smooth, while other multi-view methods have the “flipping” phenomenon, which makes the viewer feel tiresome and produces unsmooth motion parallax. The integral imaging auto-stereoscopic 3D display method is based on the controlling viewing zone method, and its display screen material employs the Low-Temperature Poly Silicon for high resolution. Its viewing zone is wider and the viewing distance is shorter, and its viewing distance is influenced by the shape of the viewing zone and can be controlled by software. Viewers can easily see 3D effect with wider and continuous viewing zone.

VI. EXPERIMENTAL RESULTS

In the experiments, the naked-eye 3D video real-time display system was implemented using two digital cameras, a workstation with CPU 2.5 GHz with 32 GB RAM and GPU (Graphic Processing Unit), and the screen of integral imaging 3D display using lenticular lens. The viewing distance was about 60 cm in front of the screen of 3D display. The naked-eye 3D video real-time display can be appreciated in wider and continuous viewing zone (32°) with nine viewpoints of high resolution.

A. Experiments of Disparity Image Computation

The Middlebury School Database Cones [23], Tsukuba and Venus [18, 24] were adopted to analyse the proposed method using ASW algorithm, as shown in Fig. 9. It was found that the mismatch was relatively little in the discontinuous region of the disparity images. In the proposed naked-eye 3D video system, the ASW algorithm with VSW was adopted to produce better disparity images. The experimental parameters are shown in Table 1 and Fig. 9. In order to compute the disparity images in high speed, the support window of 15×15 was adopted.

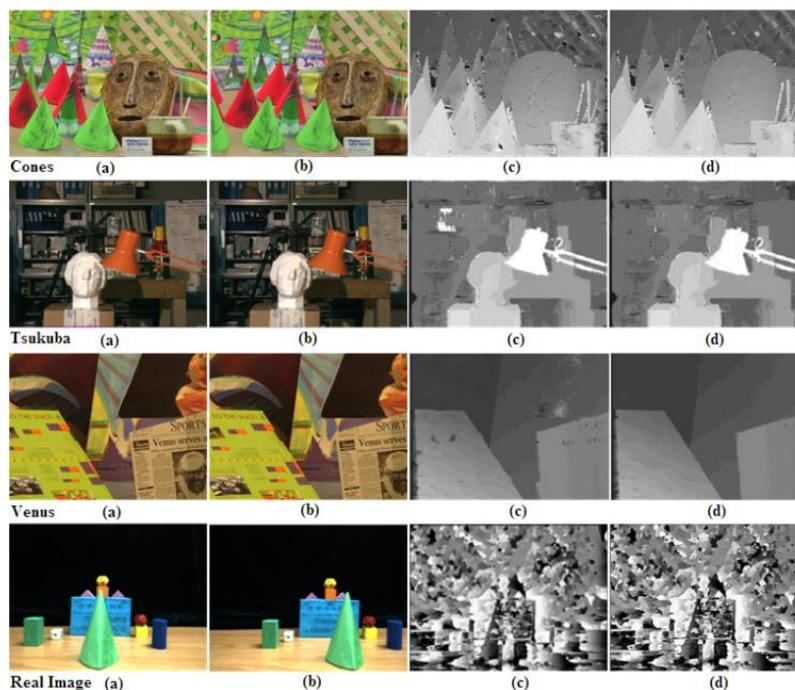


Fig. 9 Disparity image computation (a) Left image, (b) Right image, (c) Disparity image using ASW, (d) Disparity image using ASW with VSW

The bad pixels between the disparity image using ASW algorithm and the disparity image using ASW algorithm with VSW were tested. Comparison of the matched points is shown in Table 2. The data in Table 2 represents the percentage (%) of bad pixels (i.e. pixels that each has absolute disparity error of greater than "1") in the non-occluded pixels (nonocc.), the pixels in untextured areas (untext.), and the pixels near depth discontinuities (disc.). As shown in Table 2, for the pixels near depth discontinuities, the performance of the proposed method is better than that of the ASW algorithm. The 1st, 3rd, 5th, and 7th images of the seven virtual viewpoint images rendered by the depth-image-based rendering are shown in Fig. 10.

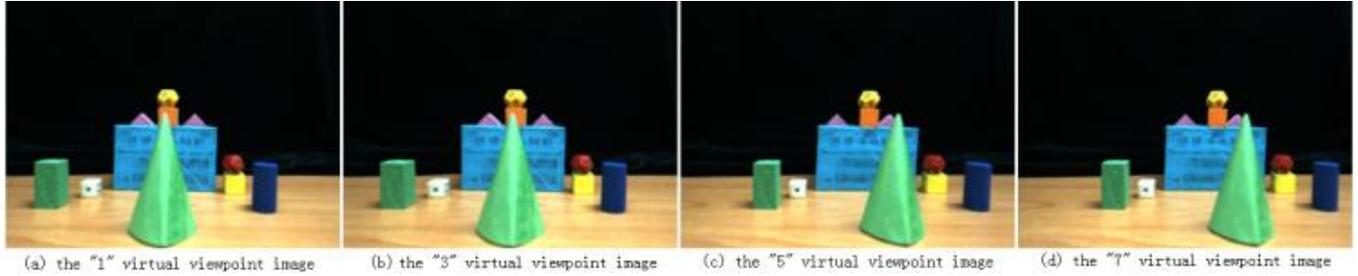


Fig. 10 The 1st, 3rd, 5th, and 7th images of the 7 virtual viewpoint images

TABLE 1 EXPERIMENTAL PARAMETERS OF THE ASW WITH VSW

Parameter	Image pixels	γ_c	γ_p	T	support window	disparity rang
Value	384×288	5	7.5	350	15×15	0~50

TABLE 2 COMPARISONS OF MATCHED POINTS WITH THE PRESENCE AND ABSENCE OF THE VARIABLE WINDOW

	Cones			Tsukuba			Venus		
	nonocc.	untext.	disc.	nonocc.	untext.	disc.	nonocc.	untext.	disc.
ASW algorithm	3.72	1.78	10.2	1.29	0.61	6.72	0.99	0.89	6.66
ASW algorithm with VSW	3.38	1.25	9.57	1.26	0.59	6.61	0.98	0.86	6.58

Through analysis of the experimental results, it was found that the proposed method not only could get the higher accuracy disparity images, but also cost less time than the ASW algorithm. The reason is that the support pixels (i.e. the support pixel q of the pixel p as shown in Fig. 4) were felicitously selected before the disparity calculation, and the support-window obtained was variable in shape and size as shown in Fig. 5, thus reducing the data redundancy and eliminating the unnecessary pixels.

B. Experiments of Disparity Image Parallel Computation

In order to reduce computation time, a GPU was adopted to improve the computation of disparity images, the depth-image-based rendering and the image interlacing. To compare the computation time of CPU with that of GPU in disparity, rendering and interlacing, system time was analysed to obtain the time consumed by computation, as shown in Table 3.

TABLE 3 TIME CONSUMPTION OF DISPARITY IMAGE COMPUTATION

Window size	CPU	GPU
	Time consuming	Time consuming
5×5	20.73 s	0.52 s
15×15	102.71 s	2.53 s
31×31	1892.56 s	44.89 s

Comparing with CPU, the GPU computation improved the time consumption by about 97%. In the experiments, as shown in Fig. 9, for the real images, the disparity image is not so satisfied. The main reason may be that the left and right cameras were calibrated not precisely enough. The GPU architecture was faster than the CPU. If the parallel threads of calculations could be changed correctly, disparity computation with higher speed may be realized, and this work will be completed in the future [7].

VII. CONCLUSIONS

A naked-eye 3D video real-time display system was established. In this system, the number of cameras for capturing 3D information was reduced from nine to two. The two cameras were applied to get the left image and the right image in order to reduce the complexity of the system structure, the transmittance time and the cost. ASW algorithm with VSW was applied to obtain the disparity images to satisfy the multi-view rendering and reduce the pixel data in the support and target windows.

Depth-image-based rendering was adopted to generate the other seven virtual viewpoint images, and the interlacing nine viewpoint images were adopted to synthesize one multi-view image (i.e. 3D information). The special screen of integral imaging auto-stereoscopic using lenticular lens was adopted to realize the naked-eye 3D video display with a wider and continuous viewing zone. Finally, in order to realize real-time 3D display, GPU was adopted to improve the computation speed of the disparity images, the multi-view rendering and the image interlacing. The experimental results showed that the ASW algorithm with VSW and the designing structure were suitable for the proposed system. This system can be applied in the entertainment industry.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (61271310).

REFERENCES

- [1] G. Ramachandran and M. Rupp, "Multiview synthesis from stereo views," the 19th IEEE International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 341-345, 2012.
- [2] J. Arai, "Three-dimensional television system based on spatial imaging method using integral photography," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5449-5452, 2012.
- [3] H. Urey, K. V. Chellappan, E. Erden, and P. Surman, "State of the Art in Stereoscopic and Autostereoscopic Displays," Proceedings of the IEEE, vol. 99, no. 4, pp. 540-555, 2011.
- [4] Yepin Lu and Suping Li, "A review on developing status of stereo video technology," International Conference on Computer Science and Information Processing (CSIP), pp. 715-718, 2012.
- [5] K. J. Yoon and I. S. Kweon, "Adaptive support-weight approach for correspondence Search," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, iss. 4, pp. 650-656, 2006.
- [6] O. Veksler, "Fast variable window for stereo correspondence using integral images," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 556-561, 2003.
- [7] Luat Do, G. Bravo, and S. Zinger, "GPU-accelerated real-time free-viewpoint DIBR for 3DTV," IEEE Transactions on Consumer Electronics, vol. 58, iss. 2, pp. 633-640, 2012.
- [8] A. Kubota, A. Smolic, M. Magnor, M. Tanimoto, T. Chen, and C. Zhang, "Multiview Imaging and 3DTV," IEEE Signal processing Magazine, vol. 24, iss. 6, pp. 10-21, 2007.
- [9] J. Congote, I. Barandiaran, J. Barandiaran, T. Montserrat, J. Quelen, C. Ferran, P. J. Mindan, O. Mur, F. Tarres, and O. Ruiz, "Real-time depth map generation architecture for 3D video conferencing," The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), pp. 1-4, 2010.
- [10] I. Feldmann, W. Waizenegger, N. Atzpadin, and O. Scheer, "Real-time depth estimation for immersive 3D videoconferencing," The True Vision - Capture, Transmission and Display of 3D Video(3DTV-CON), pp. 1-4, 2010.
- [11] P. Kauff, N. Atzpadin, C. Fehn, M. Müller, O. Schreer, A. Smolic, and R. Tanger, "Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability," Signal Processing: Image Communication, vol. 22, iss. 2, pp. 217-234, 2007.
- [12] David G. Lowe, Perceptual Organization and Visual Recognition (Kluwer Academic Publishers), 1985.
- [13] Zhengang Zhai, Yao Lu, and Hong Zhao, "Stereo Matching with Adaptive Arbitrary Support-Pixel Set and Adaptive Support-Weight," International Conference on Computational Intelligence and Security (CIS), vol. 1, pp. 598-602, 2009.
- [14] Hsin-Jung Chen, Feng-Hsiang Lo, Fu-Chiang Jan, and Sheng-Dong Wu, "Real-time multi-view rendering architecture for auto stereoscopic displays," Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1165-1168, 2010.
- [15] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV," in Stereoscopic Displays and Virtual Reality Systems XI., pp. 93-104, 2004.
- [16] Liang Zhang and W. J. Tam, "Stereoscopic Image Generation Based on Depth Images for 3D TV," IEEE Transactions on Broadcasting, vol. 51, iss. 2, pp. 191-199, 2005.
- [17] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image in painting," IEEE Transactions on Image Processing, vol. 13, iss. 9, pp. 1200-1212, 2004.
- [18] D. Scharstein and R. Szeliski, Middlebury Stereo Vision Page, <http://vision.middlebury.edu/stereo/eval/>, 2008.
- [19] Guanglin Yang and Zhichao Yang, Stereo Display Sampling Device System with Adaptive Support-Weight Stereo Matching Algorithm, CN Patent 201210023079, Peking University, 2012.
- [20] Guanglin Yang and Fuming Liu, Stereo Display Device with a Virtual Viewpoint Synthesis Algorithm of Gradient Optical Flow Method, CN Patent 201210023081, Peking University, 2012.
- [21] Guanglin Yang and Shun Li, a Naked-eye 3D Video Generation Method, CN Patent 2001310149142, Peking University, 2013.
- [22] ITU-T Rec.H.264 | ISO/IEC 11496-10 AVC[S]. Document JVT-G050, 7th Meeting: Pattaya, Thailand, 2003.
- [23] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003), volume 1, pages 195-202, Madison, WI, June 2003.
- [24] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision, 47(1/2/3):7-42, April-June 2002.



Zhichao Yang received his B. Eng. degree from School of Information and Electronics, Beijing Institute of Technology in July 2004. His M. Eng. degree from Dept. of electronics, school of electronics engineering & computer science, Peking University in July 2011. His research interests include image processing.



Guanglin Yang is an associate professor in Laboratory of image processing, Dept. of electronics, school of electronics engineering & computer science, Peking University. He received his Dr. Eng. degree from Faculty of Engineering, Osaka City University of Japan in Sept. 2001, and his M. Eng. degree from Dept. of optoelectronics engineering, Huazhong University of science and technology of China in July 1993. He became a lecturer in Faculty of automatic control engineering, Beijing University of Aeronautics and Astronautics of China in Sept. 1994. His research interests mainly include the optical information processing, digital holography, image processing, artificial neural network, intelligent infrared detecting, and digital control technique. He is a member of IEEE and OSA, respectively.



Haiyan Xie is a patent attorney in China Science Patent Trademark Agents Ltd., she respectively received her M. S. and Ph. D. degrees from School of Science, Osaka City University of Japan in March 2002 and April 2005. Moreover, she respectively received her B. and M. Eng. degrees from Dept. of Optoelectronic & Scientific Instrumentation Engineering, Zhejiang University of China in July 1990 and Dec. 1992. She became a lecturer in Faculty of Automatic Control Engineering, Beijing University of Aeronautics and Astronautics of China in 1995. Her interests mainly include the thin film optics, the semiconductor material and display screen.