S4/Venue B(Oral)/ Animation & Broadcasting /11-Jan-10, 14:00-16:00

Session Chair: Prof. Zen-Chung Shih, National Chiao Tung University

Authors	Paper Title
Purevtsogt NUGJGAR, Tadahiro FUJIMOTO ,Norishige CHIBA	Generating endless animation of water stream based on time-limited simulation
Liu Dongquan, Seah Hock Soon, Yu Jun, Vignesh Kavaserry Rajagopala	Stroke extraction and stylization for line drawing systems
Eiji Sugisaki, Fumihito Kyota ,Hock Soon Seah ,Masayuki Nakajima	In-between creation for anime character's hair
Akira Yutani, Masatoshi Kakiuchi, Atsuo Inomata, Kazutoshi Fujikawa, Keishi Kandori, Yoshitsugu Manabe, Kunihiro Chihara	A realization of the total solar eclipse live transmitting with liveliness by using 4k
Yun-Suk Kang, Eun-Kyung Lee, Yo-Sung Ho	Multi-depth camera system for 3d video generation
Eunjo Lee, Sungkwon Park, Phooi Yee Lau	Dynamic multicast DSID forwarding for IP-based switched digital video services in cable networks

GENERATING ENDLESS ANIMATION OF WATER STREAM BASED ON TIME-LIMITED SIMULATION

Purevtsogt NUGJGAR, Tadahiro FUJIMOTO and Norishige CHIBA
Graduate School of Engineering
Iwate University
Morioka, Japan
E-mail: {puugii@cg., fujimoto@, nchiba@}cis.iwate-u.ac.jp

ABSTRACT

Endlessly flowing fluids such as brooks, rivers, torrents, and fountains are some of the most visually spectacular natural phenomena. Most physically based simulation methods employed in computer graphics to depict flowing fluids are too time and memory consuming and are thus reserved for off-line simulations and small-domain real time simulations, especially in the case of fluids that flow endlessly. In this paper, we present a novel hybrid model that employs a tracer particle technique to generate endless animation of a water stream using a non-static velocity field that is obtained from physically based particle simulation. The simulation is thus performed in a limited time step. We introduce a new type of velocity field which we refer to as a Markov-type velocity field. This velocity field allows us to endlessly animate a water stream in real time by avoiding the time-consuming process of solving numerical equations.

Keywords: natural phenomenon, endless animation, tracer particle, water stream, stochastic-modeling

A REALIZATION OF THE TOTAL SOLAR ECLIPSE LIVE TRANSMITTING WITH LIVELINESS BY USING 4K

Akira Yutani, Masatoshi Kakiuchi*, Atsuo Inomata*, Kazutoshi Fujikawa*, Keishi Kandori**, Yoshitsugu Manabe*, Kunihiro Chihara**
* Nara Institute of Science and Technology
Nara, Japan
** Asahi Broadcasting Corporation
Osaka, Japan
E-mail: yuta@itc.naist.jp

ABSTRACT

The total solar eclipse of July 22, 2009 was the longest that of the 21st century, many astronomy shows were held in Asia. To transmit this valuable moment to a distant area with enough liveliness, we considered a mechanism how to transmit. Only that transmits the image of the corona or the diamond ring which the moon lines up with the sun like existing methods, it is difficult to say that we can transmit a total solar eclipse perfectly. In this paper, we describe the technique for transmitting in 4K ultra high definition image of the whole sky all with an atmosphere. Also we discuss about the effectiveness of 4K from results of experimentation, especially for liveliness and immersive.

Keywords: Total solar eclipse, 4K, the whole sky all

MULTI-DEPTH CAMERA SYSTEM FOR 3D VIDEO GENERATION

Yun-Suk Kang, Eun-Kyung Lee, and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro Buk-gu, Gwangju 500-712, Korea
E-mail: {yunsuk, eklee78, hoyo}@gist.ac.kr

ABSTRACT

In this paper, we describe a multi-depth camera system for 3D video contents generation. We combine five video cameras and five TOF depth cameras to capture a scene in real time. By taking advantages of both active and passive sensor based depth acquisition methods, we can estimate multi-view depth sequences accurately. After performing several steps of preprocessing, the depth sequences are warped to the corresponding video camera positions and used as initial depth values for refinement. By applying the stereo matching method, we can improve multi-view depth sequences. Intermediate view image sequences are then synthesized using both the multi-view color and depth sequences.

Keywords: Depth generation, 3D video, TOF camera,
Depth camera, Multi-view video.

MULTI-DEPTH CAMERA SYSTEM FOR 3D VIDEO GENERATION

Yun-Suk Kang, Eun-Kyung Lee, and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro Buk-gu, Gwangju 500-712, Korea
E-mail: {yunsuk, eklee78, hoyo}@gist.ac.kr

ABSTRACT

In this paper, we describe a multi-depth camera system for 3D video contents generation. We combine five video cameras and five TOF depth cameras to capture a scene in real time. By taking advantages of both active and passive sensor based depth acquisition methods, we can estimate multi-view depth sequences accurately. After performing several steps of preprocessing, the depth sequences are warped to the corresponding video camera positions and used as initial depth values for refinement. By applying the stereo matching method, we can improve multi-view depth sequences. Intermediate view image sequences are then synthesized using both the multi-view color and depth sequences.

Keywords: Depth generation, 3D video, TOF camera, Depth camera, Multi-view video

1. INTRODUCTION

Since the first television (TV) broadcasting started in 1930s, a variety of technologies related to broadcasting such as image capturing and processing, data compression, communications, and display device manufacturing have been steadily studied and realized. As these technologies have been developed, the demand for multimedia services which can provide more immersive and realistic sense also has rapidly increased. The multi-view image or multi-view video is considered as the strongest candidate to satisfy this demand.

The multi-view video is a set of image sequences of the same scene captured by more than two cameras. We can select and change viewpoints within the available range, and then obtain images from multiple viewpoints. In addition, we can generate images representing depth information of the scene and also images at intermediate viewpoints. In recent years, much research for applications using the multi-view image such as free viewpoint TV (FTV) and three-dimensional TV (3DTV) has actively progressed [1-2].

There are a variety of methods to obtain depth information of a scene. Generally, they can be divided into two major categories. One is based on passive methods and the other is based on active methods. Among the passive methods, stereo matching is the most popular one which is widely used [3]. Although it is widely used in the fields, some difficult problems have been remained such as occlusion.

The most popular active method is time-of-flight (TOF) range sensor to obtain the depth information of the scene in real time. The principle of depth acquisition of TOF is based on measuring the arriving time of emitted infrared signal from the sensor. However, TOF has several problems to overcome such as low spatial resolution and noisy acquisition depending on the capturing environment.

There is research on a system that combines a TOF depth camera and video cameras [4-7]. In these approaches, they can take the advantages and discard the disadvantages of passive and active methods. The depth information from the passive method can be enhanced by using the depth obtained from the active method.

In this paper, we introduce a multi-depth camera system that has multiple video cameras and multiple depth cameras. They are arranged in two rows with parallel type. The captured color and depth image sequences with 30fps are preprocessed to increase the inter-view correlation and the computational accuracy. The preprocessed depth image sequences are then warped to the corresponding video cameras and used as initial depth values for stereo matching. After depth estimation, we obtain the 3D video that is composed of the multi-view color image sequences and the corresponding multi-view depth sequences. With this 3D video, we can reconstruct the intermediate view image sequences and watch the 3D scene through multi-view display devices.

2. SYSTEM CONFIGURATION

In the proposed system, we combine five video cameras and five depth cameras as shown in Fig. 1. The video camera model we use is Basler Pylon GigE [8] that can provide up to HD resolution. The depth camera model is Swiss Ranger SR4000 [9] that uses the TOF depth sensor. Two different kinds of cameras are mounted on a camera frame. There are two rows to amount the cameras, and they can hold and shift both types of the cameras. The fundamental setup is that the first row has five video cameras and the second row has five depth cameras. These two rows also move to the upper and lower sides. The specification of the camera frame is shown in Table 1.

For capturing a sequence simultaneously, we connect the video cameras to the PC for storing through a synchronizer CA-1000 by National Instrument [10]. To synchronize the depth cameras, we modified the software development kit (SDK) provided by Swiss Ranger [9] for synchronized capturing. However, there is no instrument for synchronization of two different types of cameras. Thus,

video capturing and depth capturing are respectively started with a small time difference.

Table 1: Specification of the camera frame

Width	120cm
Height	50cm to 160cm
Min. distance between two adjacent cameras	Horizontal direction: 6.5cm
	Vertical direction: 4cm

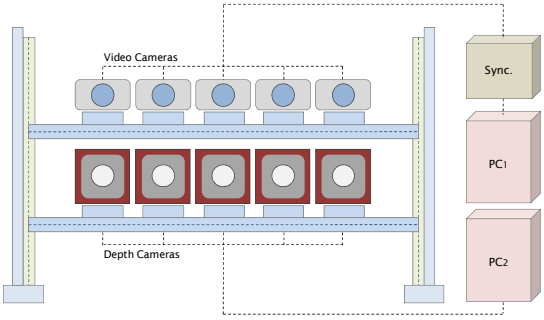


Fig. 1: The proposed multi-depth camera system

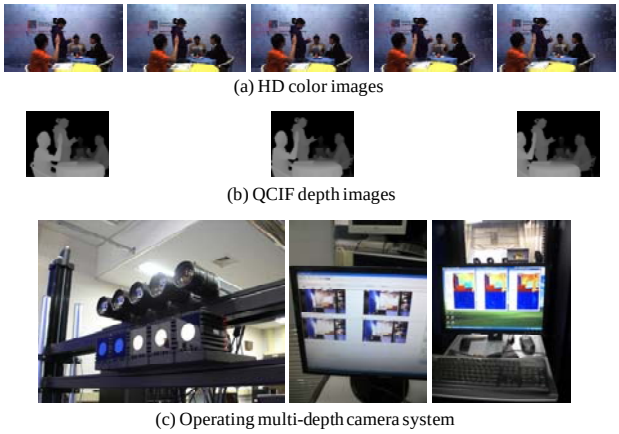
3. DESCRIPTION OF MODULES

In this section, we explain several principal modules for 3D video generation using the proposed multi-depth camera system. Figure 2 shows the whole procedure to obtain the 3D video.

3.1 Data Acquisition

We can acquire not only the multi-view video sequence but also its corresponding multi-view depth video sequence by using the proposed camera system. Five video cameras capture a scene with 30fps. The maximum resolution is HD and the lower resolutions are available by setting the ROI from the original HD resolution images. Figure 3(a) is the captured five images by the video cameras.

Five depth cameras also obtain the 3D information of the same scene with 30fps. These depth cameras have two types of output images which are depth and intensity images. The depth image represents the 3D information of the scene and the intensity image is considered as the grey images of the scene. Despite the depth camera has a small output resolution of QCIF, we can acquire the 3D information of the scene in real time. Figure 3(b) shows the depth image and the intensity image of the scene.



(a) HD color images

(b) QCIF depth images

(c) Operating multi-depth camera system

Fig. 3: Captured images by the proposed camera system

However, one problem of the depth camera is that simultaneous capturing is allowed up to three cameras. It is the problem of modulation frequencies. To avoid interference among radiated signals, each SR4000 uses different modulation frequencies. Since they support three different frequencies, we can simultaneously operate up to three cameras. Figure 3(c) shows the operating multi-depth camera system.

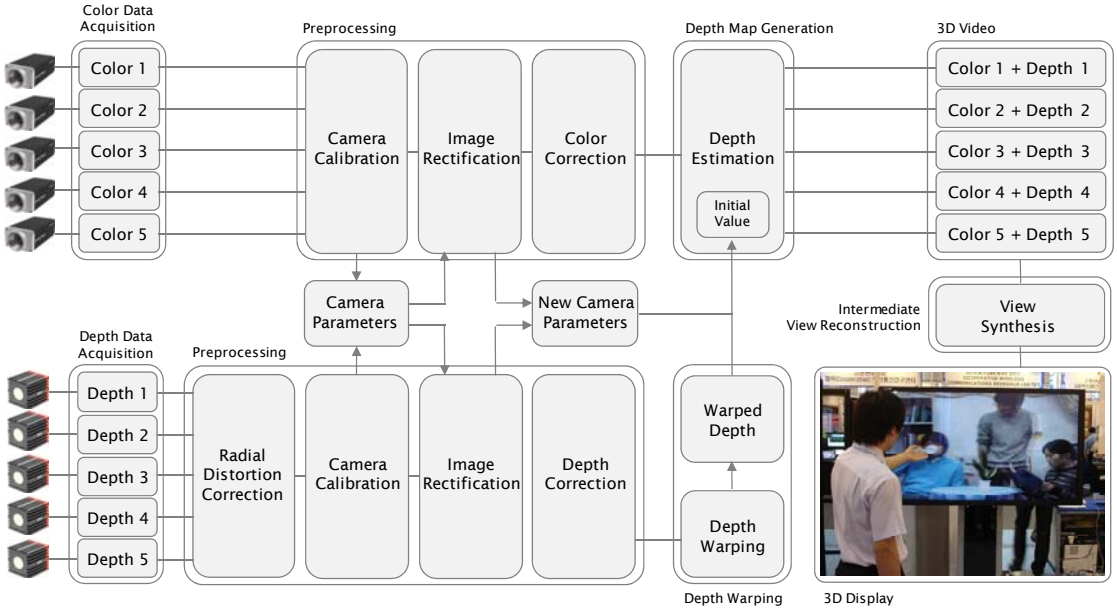


Fig. 2: Procedure of the proposed system

3.2 Preprocessing

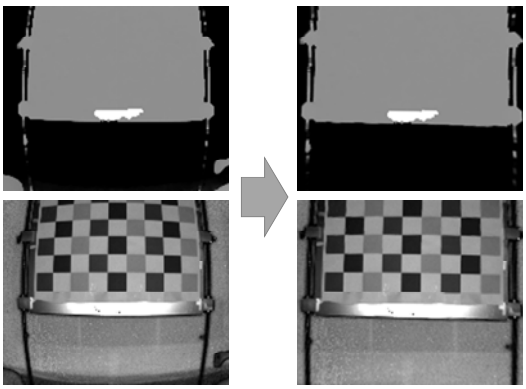
Captured color and depth video sequences require several preprocessing steps to obtain metric information and to correct the mechanical defects. In this subsection, we briefly introduce each preprocessing step.

3.2.1 Radial Distortion Correction for Depth Images

Depth images captured by SR4000 have a large amount of lens radial distortion. There are two types lens distortion which are barrel distortion and pincushion distortion. In this case, the barrel distortion is occurred by the intrinsic problem of the depth camera. This distortion causes not only the shape mismatch between the color image and the corresponding depth image but also the errors in the results of some feature point based processing such as camera calibration.

In order to avoid that situation, we have to perform radial distortion correction to the obtained depth images. In general, there are two main categories of radial distortion correction. Methods in the first category use the point correspondences between two or more views. The second category also has lots of approaches which are based on the distorted straight line components in the image.

In the proposed multi-depth camera system, we use one of the second approaches to correct the radial distortion in the depth images [11]. After finding the curved straight line component in the captured image, we estimate the distortion center and the distortion parameter. With the distortion information, we can reconstruct the image from the distorted image. Figure 4 shows the depth and intensity images before and after the correction.



Original Corrected

Fig. 4: Results of radial distortion correction

3.2.2 Camera Calibration

Camera calibration is the process to extract metric information of the camera from captured images. To calibrate multiple cameras relatively, we need several images of a planar check-patterned board from each camera. Then, by extracting a number of regular feature points from each image with the real distance between the neighboring two points, we can find the camera intrinsic and extrinsic parameters [12].

In our approach, we execute camera calibration of the depth cameras after radial distortion correction. Because the radial distortion interferes the extraction of feature points, the depth images we use are the corrected images. We employ the Camera Calibration Toolbox for MATLAB provided by Caltech [13]. For camera calibration, we captured 11 planar check-patterned board images for each camera. As shown in Fig. 5, the board has 81 small squares and the distance between two adjacent points is 90mm.

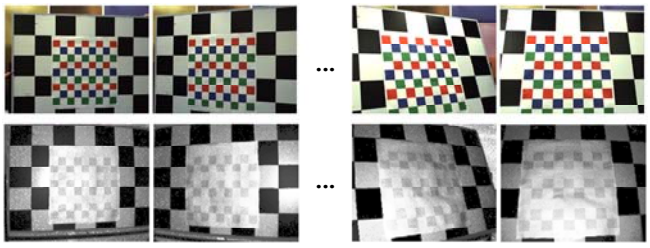


Fig. 5: Check-patterned board images

3.2.3 Multi-view Image Rectification

To capture the multi-view video and multi-view depth video to generate the 3D video, our proposed camera system has the video cameras and depth cameras on one-dimensional parallel arrays. However, the cameras are manually mounted on the camera frame. Therefore, there exist geometric errors that mean the non-ideal parallel camera arrangement. These errors are represented as the vertical pixel mismatches and the irregular horizontal disparities between the inter-view corresponding points as shown in Fig. 6(a).

The geometric errors in the multi-view image are obstacles to processing time and accuracy in 3D video generation and application. Also, they deteriorate not only the correlation among views but also the visual quality of the 3D video. Thus, it is essential to minimize the geometric errors in the multi-view image [14].

Multi-view image rectification is the transformation to minimize the geometric errors in the multi-view image. Based on the original camera parameters of each camera, we estimate the new camera parameters which have the same intrinsic characteristics and the smallest geometric errors. By using these two sets of camera parameters of each camera, we compute the rectifying transformations and then apply them to the original images. As depicted in Fig. 6(b), rectified multi-view image has the same vertical coordinates and regular disparities between the corresponding points. All intrinsic parameters have the same characteristics after rectification [15].

(a) Original (b) Rectified

Fig. 6: Results of multi-view image rectification

In the proposed system, we apply the multi-view image rectification to both color and depth images. As a result, we obtain the rectified multi-view color and depth images and also regular camera parameters.

3.2.4 Multi-view Color Correction

Although we capture a scene using the same kind of multiple cameras in the consistent condition, there exists color mismatch problem among views. It is caused by different color characteristics of the cameras. Because this problem can decrease the efficiency and quality of some processes such as depth generation and intermediate view synthesis, we have to adjust each image to reduce the color mismatch.

There are three main approaches to correct the color mismatch problem in the multi-view image. The first approach usually sets one view image as a reference and then adjusts the other view images. The advantage of this approach is the simplicity. However, the disadvantage is the high dependency on the input image characteristics since the occlusion and disocclusion regions are not considered [16].
The second approach is to use the standardized color chart called Macbeth chart shown in Fig. 7. This approach is independent on the input images and provides high accuracy. However, we cannot apply the color correction to the image that has no color chart information [17].

Fig. 7: Captured image of the color chart

Recently, research on the third approach that handles the color mismatch problem based on the corresponding points has been actively studied. To regularize the different color characteristics, we find a number of sets of the corresponding points among views, and then we extract the color information. By using this information, we can reduce the color mismatch in the multi-view image [18].

In the proposed camera system, we apply one method that belongs to the second approach to the color images. The original images and the color corrected images are indicated in Fig. 8(a) and (b), respectively.

(a) Original

(b) Color corrected

Fig. 8: Result of multi-view color correction

3.2.5 Depth Value Correction

Depth information obtained from the depth camera has nonlinear characteristic between the real depth and its depth value. Moreover, when we use multiple depth cameras, there is irregularity in depth values of multiple depth images due to the camera intrinsic features. Since these characteristics give wrong 3D information of the captured scene, we have to correct depth values before applying 3D warping.

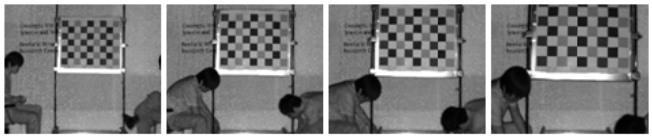
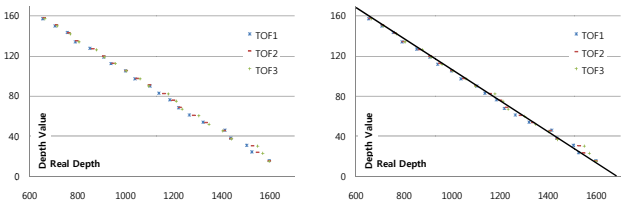


Fig. 9: Image capturing according to regular interval

To correct multiple depth images, we obtain still images of the check-patterned board with moving it from the farthest plane to the front with regular interval as shown in Fig. 9. From these images, we can calculate real depth values Z at each captured point by using Eq. 1, where i indicates the i-th image and f, B, and d mean the focal length, baseline, and disparity, respectively.

$$Z_i = \frac{f \cdot B}{d_i} \tag{1}$$

The relationship between depth value for the check-pattern and real depth is shown in Fig. 10(a). It describes that there are irregular and nonlinear characteristics in multiple depth images. Therefore, we find a line as a function shown in Fig. 10(b) that minimizes the squared distance from each point to the line to correct depth values.



(a) Relationship between the depth value and real depth

(b) Function for depth value correction

Fig. 10. Depth value correction

To correct the multi-view depth image, we calculate the real depth of each pixel and then obtain corrected depth value by using the function. The calculation of the real depth Z of pixel (i, j) is

$$Z(i,j) = Z_{min} + \{D_{max} - D(i,j)\} \times \frac{Z_{max} - Z_{min}}{D_{max} - D_{min}} \tag{2}$$

where D_{max}, D_{min}, $D(i, j)$, Z_{max}, and Z_{min} indicate the maximum and minimum depth values, depth value of pixel (i, j), and maximum and minimum real depth values of the scene that are computed based on the minimum and maximum disparity values, respectively.

3.3 Multi-view Depth Map Generation

In this subsection, we introduce how to generate the multi-view depth map using the preprocessed data. Basically, the final depth maps are generated by using stereo matching. However, we can obtain an initial value for stereo matching from the captured depth information. Then the final depth map for each view is produced after the refinement.

3.3.1 Initial Depth Value Generation

For the accurate stereo matching, we can use the initial disparity values. We regard depth information obtained by the depth cameras as the initial disparity of multi-view image. In order to match the depth images that have the different resolution compared to the color images and captured by lower position than the video camera, we backproject the depth images to the world coordinates using 3D warping.

Then, the depth information in the 3D space is reprojected onto each image plane of the video camera. This reprojected depth information is changed to the form of disparity by using Eq. 1.

Figure 11(a) shows the original depth image and the warped depth image. The warped depth image has the same resolution as the color image. The warped depth image and the color image are almost matched as shown in Fig. 11(b). Because not only there is shape mismatch between the original depth image and the color image, but also the resolution difference is too large.

(a) Warped depth image

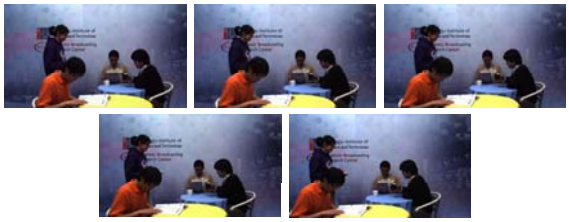
(b) Matched to color image

Fig. 11: Depth image warping

3.3.2 Multi-view Depth Estimation

In the proposed system, we obtain the depth map of each view based on the initial depth information that is from the depth cameras. After applying the mean-shift color segmentation to color images, the initial depth of each pixel is represented as initial disparity value using Eq. 1. At this point, we use the baseline and focal length of video cameras. Also, we assumed that each segment has one disparity value that is the average. Then, the disparity map of each view is obtained based on segments with their disparity values that minimize the sum of absolute difference (SAD) between the current and reference segments.

In order to increase the quality of the initial disparity map, we refine that using belief propagation method. Figure 12 shows the generated 3D video that is composed of the multi-view color and multi-view depth sequences.



(a) Multiview color sequence

(b) Multiview depth sequence

Fig. 12: Generated 3D video

3.4 Intermediate View Reconstruction

With the generated multi-view disparity or depth map, we can reconstruct intermediate view images between two neighboring views. These intermediate view images are considered as images that are captured by the virtual cameras. They can decrease the disparity between the two adjacent views, which helps users to feel the 3D sense. Because the excessive disparity makes user's eyes exhaust and causes flickering at the viewpoint change.

In order to reconstruct the intermediate viewpoint images, the generated disparity or depth images are used. We set the viewpoint we reconstruct, and then obtain two projected images from the left and right adjacent original images. By blending two images according to the distance, we can reconstruct the intermediate viewpoint images as shown in Fig. 13.

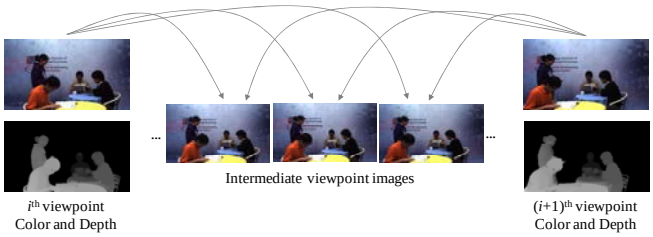


Fig. 13: Intermediate view reconstruction

3.5 3D Display

The generated multi-view video sequence can be displayed at various types of 3D display. For stereoscopic displays, adjacent two views are selected and displayed with one viewpoint. In the case of multi-view displays, there is a built-

in viewpoint limitation. For example, 9-view monitors have nine input video sequences. Users also can be provided nine viewpoints, and feel a 3D sense within the viewpoints. Figure 14 shows the multi-view displays showing the multi-view sequences.

Fig. 14: Multi-view 3D displays

4. CONCLUSION

In this paper, we introduced a multi-depth camera system for 3D video generation. The proposed camera system captures the multi-view video and the depth image sequences in real time by using five video cameras and five TOF depth cameras. The proposed system has an advantage on fast and accurate multi-view depth sequence generation. With the generated 3D video, we can reconstruct intermediate viewpoint images and feel a 3D sense from them through multi-view displays. However, the modulation frequency problem that limits the number of simultaneous working depth camera is still remained. The resolution difference between the color and depth images is also difficult to solve. In the future, we would like to improve each module in detail with hope that the mechanical problems are solved.

ACKNOWLEDGEMENT

This research was supported in part by the MKE (Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) [NIPA-2009-(C1090-0902-0017)].

REFERENCES

[1] A. Smolic and P. Kauff, "Interactive 3D Video Representation and Coding Technologies," Proceedings of the IEEE, Spatial Issue on Advances in Video Coding and Delivery, vol. 93, no. 1, pp. 99-110, 2005.

[2] ISO/IEC JTC1/SC29/WG11 M8595, "FTV-Free Viewpoint Television," 2002.

[3] J. Sun, N. N. Zheng, and H. Y. Shum, "Stereo Matching Using Belief Propagation," IEEE Transactions of Pattern Analysis and Machine Intelligence, vol. 25, no. 5, pp. 787-800, 2003.

[4] K.D. Kuhnert and M. Stommel, "Fusion of Stereo-Camera and PMD-Camera Data for Real-time Suited Precise 3D Environment Reconstruction," In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.4780-4785, 2006.

[5] J. Zhu, L. Wang, R. Yang, and J. Davis, "Fusion of Time-of-Flight Depth and Stereo for High Accuracy Depth Maps," in Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 231-236, 2008.

[6] G. Um, K. Kim, C. Ahn, and K. Lee, "Three-dimensional Scene Reconstruction using Multi-view Images and Depth Camera," in Proc. of 3D Digital Imaging and Modeling, pp. 271-280, 2005.

[7] S.Y. Kim and Y.S. Ho, "Generation of ROI Enhanced Depth Maps using Stereoscopic Cameras and a Depth Camera," IEEE Transaction on Broadcasting, vol. 54, no. 4, pp. 732-740, 2008.

[8] Basler Pylon GigE Camera piA I1900-32gm/gc, http://www.baslerweb.com

[9] Mesa Imaging SR 4000, http://www.mesa-imaging.ch

[10] National Instruments CA-1000, http://sine.ni.com

[11] A. Wang, T. Qiu, and L. Shao, "A Simple Method of Radial Distortion Correction with Centre of Distortion Estimation," Journal of Mathematical Imaging and Vision, vol. 35, no. 3, pp. 165-172, 2009.

[12] Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 22, no. 11, pp. 1330-1334, 2000.

[13] Camera Calibration Toolbox for MATLAB, http://www.vision.caltech.edu/bouguetj

[14] ISO/IEC JTC1/SC29/WG11 M15379, "Adjusting Method for Multi View Image; Color and Geometry Correction for MPEG-FTV Test Sequences," 2008.

[15] Y.S. Kang and Y.S. Ho, "Geometrical Compensation for Multi-view Video in Multiple Camera Array," in Proc. of International Symposium on Electronics and Marine (ELMAR), pp. 83-86, 2008.

[16] U. Fecker, M. Barkowsky, and A. Kaup, "Improving the Prediction Efficiency for Multi-View Video Coding Using Histogram Matching," in Proc. of Picture Coding Symposium, pp. 2-16, 2006.

[17] N. Joshi, B. Wilburn, V. Vaish, M. Levoy, and M. Horowitz, "Automatic Color Calibration for Large Camera Arrays," in UCSD CSE Technical Report CS2005-0821, 2005.

[18] K. Yamamoto, M. Kitahara, H. Kimata, T. Yendo, T. Fujii, M. Tanimoto, S. Shimizu, K. Kamikura, and Y. Yashima, "Multi-view Video Coding Using View Interpolation and Color Correction," IEEE Transactions on Circuits and Systems for Video, vol.17, no.11, pp. 1436-1449, 2007.

[19] D. Comaniciu and P. Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 24, no. 4, pp. 603-619, 2002.