

Multi-scale 3D Panorama Content Augmented System using Depth-map

Kim Cheeyong[†], Eung-Kon Kim^{**}, Jong-Chan Kim^{***}

ABSTRACT

With the development and spread of 3D display, users can easily experience an augmented reality with 3D features. Therefore, the demand for content of an augmented reality is exponentially growing in various fields. A traditional augmented reality environment was generally created by CG(Computer Graphics) modelling production tools. However, this method takes too much time and efforts to create an augmented environment. To create an augmented environment similar to the real world, everything in the real world should be measured, gone through modeling, and located in an augmented environment. But the time and efforts spent in the creation don't produce the same environment as the real world, making it hard for users to feel the sense of reality. In this study, multi-scale 3D panorama content augmented system is suggested by using a depth-map. By finding matching features from images to add 3D features to an augmented environment, a depth-map is derived and embodied as panorama, producing high-quality augmented content system with a sense of reality. With this study, limits of 2D panorama technologies will be overcome and a sense of reality and immersion will be provided to users with a natural navigation.

Key words: Multi-scale, 3D Panorama, Augmented Reality, Depth-map

1. INTRODUCTION

3D content system is a technology which helps users accept a VR(Virtual Reality) as reality through 3D visual information and feel as if they are in the field, which are utilized in various fields such as education, advertising, home shopping, military training, and mock labs and medical industry. This kind of computer display system significantly contributes to the development of user interface, real-time object recognition, cutting-edge systems for the disabled. Various studies on 3D content including information about a sense of 3D effect and immersion are carried out home and abroad, increasing fields utilizing them[1,2].

The typical technology is a image-based modeling and rendering which establish 3D augmented environment based on images. This method lacks a sense of reality since it provides realistic images of the environment or objects at arbitrary times based on information from images taken at actual environments. However, it has merits in that it doesn't need a traditional 3D modeling process, which reduces costs in the establishment and a production speed of printing images. To overcome these problems, studies on 3D image production through image synthesis and processing are carried out actively home and abroad[3-7].

In this study, a multi-scale 3D panorama content augmented system using a depth-map of images

※ Corresponding Author : Jong-Chan Kim, Address: (540-742) Jungangro 255, Suncheon, Jeonnam, Korea, TEL : +82-10-2741-7026, FAX : +82-61-755-7999, E-mail : seaghost@sunchon.ac.kr

Receipt date : Apr. 21, 2014, Approval date : May. 7, 2014

[†] Dept. of Visual Information Engineering, Dong-eui University
(E-mail : kimchee@deu.ac.kr)

^{**} Dept. of Computer Engineering, Sunchon National University
(E-mail : kek@sunchon.ac.kr)

^{***} Dept. of Computer Engineering, Sunchon National University

※ This work was supported by Dong-eui University Grant.(2013AA172)

is suggested. By finding matching features from images to add 3D features to an augmented environment, a depth-map is derived and embodied as panorama, producing high-quality augmented content system with a sense of reality. This system adds a high-quality panorama technology to user-friendly technology, providing 3D panorama contents which can be intuitively and easily manipulated, thereby improving a sense of immersion and reality of users.

2. RELATED WORKS

2.1 Panorama VR

Panorama VR is to show nearby sight at one's location. it gives an effect just like he feels in the place and it can reenact at certain days. To acquire this effect, panoramic technique is used. Images of real environment taken in 360° are called omni images or omnidirectional images. And images transformed them into cylinder forms, circular or rectangular forms are called panoramic images[8-10].

To establish panorama VR, two steps are needed. The first step is to acquire panoramic images and the second step is to display them by panorama VR viewer. When users designate a certain gaze direction after receiving panoramic images, the panorama VR viewer shows the image seen from the direction by producing it from panoramic images in real time. Since panorama VR is widely used in the Internet, various kinds of panorama VR viewers have appeared. The most currently used viewers are Apple's QuickTime Player, Macromedia's Shockwave Player, Anything 3D's viewer, iSeeMedia's viewer, Zoomify's viewer, iPIX's viewer, and MG System's Viewer. Fig 1. shows how to produce panoramic VR using a cylinder.

2.2 Software method of production of panoramic images

Image mosaic is a method for matching several

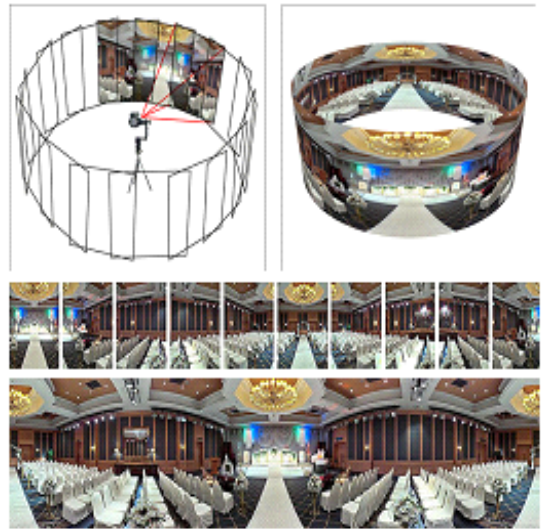


Fig. 1. Panorama VR process step.

images and connecting them by general camera. First, set the starting point of photo-taking. And then rotate horizontally by overlapping 10%-50% of images and take 12-24 pictures according to the degrees of overlapping. Connect the overlapping parts to make a one piece of panorama, which is called a stitching process. For images to configure a cylinder by matching with neighboring ones, each image will be bent. Overlapped pictures should be bent in the same angle to fit perfectly. Each picture image is a flat rectangular image first, but to configure a cylinder its corner should be bent. The degrees of bending is different according to a filed of view of a used lens, and the images which become the cylinder part are more distorted. And when playing them next, the images will become flat as original. A final source image will have different dimension and ratio according to each images filed of view. The bigger the filed of view, the higher the height becomes compared to area. In other words, the smaller the filed of view, the larger the panorama. Pictures completing stitching have different tops and bottoms in length since they seem to connect pictures cut long. And then the blending process which make them a panorama picture by connecting them smoothly is car-

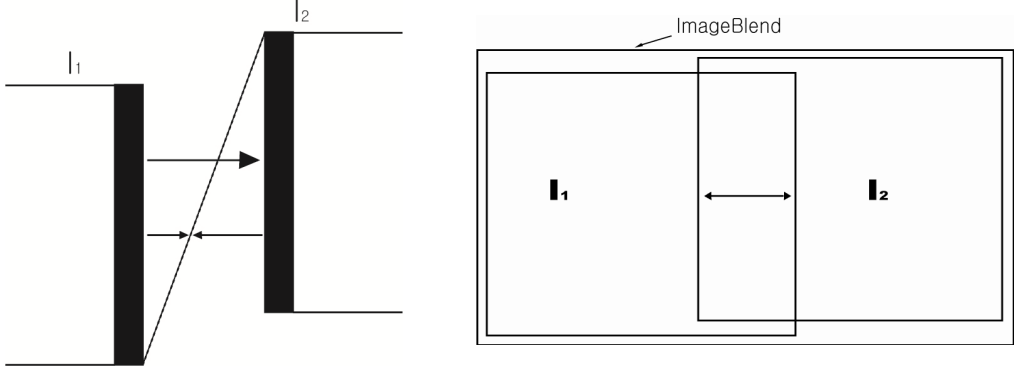


Fig. 2. Image blending step.

ried out. The process is shown in Fig 2.

To produce high-quality panoramic images effectively, the exact camera direction of each input image should be found out. To do so, after finding the relative camera rotations of neighboring images through image matching, the camera direction is calculated by adding them in order. This approach has demerits in that errors occurring in image matching are accumulated and final panoramic images have gaps or overlays[10-13].

3. 3D PANORAMA GENERATION ALGORITHM

Let $x = (x, y)$ which is a point from projection of $X = (X, Y, Z)$, a point in 3D(dimensional) space that is taken by lens distortion free camera, to 2D image, the relationship between X and x is modelled as follow[14].

$$u = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = K(RX + t) \quad (1)$$

$$= \begin{bmatrix} f & s & x_c \\ 0 & af & y_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}, \quad (2)$$

$$x = \frac{u}{v}, \quad y = \frac{v}{w}$$

u is a vector that represents x as Homogeneous Coordinates. K , as a Camera Calibration Matrix, shows the relationship between 3D point on camera coordinates and 2D point on image within which 3D point is projected. R and t , that represent

Rotation Matrix and Translation Vector each, indicate the transformation relationship between world coordinates and camera coordinates. As we call parameters included in vector K as Internal Parameter of camera, f is Focal Length, Q is Aspect Ratio of pixel, S is Skew and (X_c, Y_c) is Principal Point.

$$u_1 = KR_1X = KR_1R_0^{-1}K^{-1}u_0 = KR_{01}K^{-1}u_0 = H_{u_0} \quad (3)$$

Rotation Matrix and Translation Vector of camera before rotating are R_0 and t_0 respectively and those of camera after rotating are R_1 and t_1 . When conforming projection center to zero point of world coordinates for convenience's sake, it becomes $t_0 = t_1 = 0$. If homogeneous vector of a point that is from the projection of certain point X at 3 dimensional space to I_0 (original image) and I_1 (rotated image) is each u_0 and u_1 , then.

That is, as pixel correspondence relationship between two images that are from camera rotation, which is presented by a 3×3 matrix H , we call it Planar Homography. Considering that corresponding image point is the same even though multiplying any real number except 0 to Homogeneous Vector, H also can be defined regardless of its scale. Hence, Planar Homography is defined 8 parameters, not 9 parameters. Actually, as the number of parameter that is included in K from $H = KR_{01}K^{-1}$ is 5 and the number of parameter that is in R_{01} is 3, we can aware that independent pa-

Table 1. Camera condition of normalized model

condition) - As pixels are generally square pixel, pixel's aspect ratio satisfies $a = 1$. - As CCD cells that compose image are arranged on the square, Skew meets $s = 0$. - As an assumption that optical axis of camera go through array center is safe, Principal Point (x_c, y_c) accords with image center.

rameters that are composed of H is 8. This Homography is called as 8-Parameter Homography.

Given two images, finding H which is pixel correspondence relationship between the two images is Image Registration. In general, H can be calculated with more than 4 correspondent points but it takes caution to get more accurate H. Therefore, we assume that camera satisfies this normalized model. Table 1. shows the conditions of camera based on normalized model.

Therefore, the number of independent parameter included in H on normalized model is 4(focal point distance f and three rotating angles within R_{01}), which we call as 4-Parameter Homography. General rotation matrix R is divided as follow.

$$R = R(\psi)R(\theta)R(\varnothing) \tag{4}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\varnothing & \sin\varnothing & 0 \\ -\sin\varnothing & \cos\varnothing & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In formula(4), $R(\psi), R(\theta), R(\varnothing)$ are respective matrix that present rotation transformation from each X, Y, Z axis and $\psi, \theta, \varnothing$ are called Euler Angle. Given Planar Homography H in normalized model, calculating K and R are as shown below. Matrix K can be divided as follow.

$$K = K_c K_f = \begin{bmatrix} 1 & 0 & x_c \\ 0 & 1 & y_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

Substituting to formula(3)

$$K_f R K_f^{-1} = K_c^{-1} H K_c = \hat{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \tag{6}$$

Therefore,

$$R = \begin{bmatrix} h_{11} & h_{12} & h_{13}/f \\ h_{21} & h_{22} & h_{23}/f \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \tag{7}$$

As R is Orthogonal Matrix, it satisfies following formula. After calculating f from formula(8), it gets R by substituting to formula(7).

$$h_{11}^2 + h_{12}^2 + \frac{h_{13}^2}{f^2} = h_{31}^2 f^2 + h_{32}^2 f^2 + h_{33}^2 \tag{8}$$

Camera models described above do not consider distortion of its lens. As it utilizes wide angle lens to get 360 angle panorama video, it is necessary to consider lens distortion to apply algorithm to wider area. (x_d, y_d) is coordinates of pixel in original image(that is, image that contains lens distortion) and (x_u, y_u) is corresponding coordinates of pixel in image that removes lens distortion. As (x_c, y_c) is the center coordinates, given $(x_d, y_d), (x_u, y_u)$ is calculated as follow.

$$\begin{aligned} x_u &= (x_d - x_c)(1 + k_1 r_d^4 + k_2 r_d^4 + \dots) + x_c \\ y_u &= (y_d - y_c)(1 + k_1 r_d^4 + k_2 r_d^4 + \dots) + y_c \\ r_d &= \sqrt{(x_d - x_c)^2 + (y_d - y_c)^2} \end{aligned} \tag{9}$$

In this moment, k_1, k_2 are radial distortion coefficients. In common,

$$\begin{aligned} x_u &= (x_d - x_c)(1 + k r_d^2) + x_c \\ y_u &= (y_d - y_c)(1 + k r_d^2) + y_c \end{aligned} \tag{10}$$

To the contrary when (x_u, y_u) is given, (x_d, y_d) is as follow.

$$x_d = \frac{x_u - x_c}{1 + k r_d^2} + x_c, y_d = \frac{y_u - y_c}{1 + k r_d^2} + y_c \tag{11}$$

In formula (12),

$$r_d^2 = h(k) + \frac{1}{9h(k)k^2} - \frac{2}{3k} \tag{12}$$

$$h(k) = \left(\sqrt{\left(\frac{1}{27k^3} + \frac{r_u^2}{2k^2} \right)^2} - \frac{1}{729k^6} + \frac{1}{27k^3} + \frac{r_u^2}{2k^2} \right)^{\frac{1}{3}} \tag{13}$$

$$r_d^2 = (x_u - x_c)^2 + (y_u - y_c)^2 \tag{14}$$

4. MULTI-SCALE 3D PANORAMA CONTENT AUGMENTED SYSTEM USING DEPTH-MAP

Multi-scale 3D panorama content augmented system using depth map produces one panoramic image by analyzing input images of multi-scale images. The image stitching technique is a method for making images taken in certain angles as a stretched out form without joints. Multiple images acquired from more than one cameras are produced as one panoramic image without joints through a image synthesis process. Since cameras used to acquire a panoramic image are set in different directions and locations, overlapping parts of images should be found out and geometrical adjustments should be carried out to produce one panoramic image without joints.

Multi-scale 3D panorama augmented reality content system produces a depth-map between images internally to produce panoramic images. To derive a depth-map, median filters are applied to remove noises from images. Edges are detected through the derivation of canny edges to input images through preprocessing, and vanishing lines are detected through Hough transformation. Through the intersection points of the detected vanishing lines, vanishing points are designated. Since the depth step of the vanishing points should be different depending on the locations of images, the range of locations of corresponding vanishing points between images and vanishing points are defined. By setting the defined range as standard, a standard depth step on a base level is produced. By linear interpolation of neighboring vanishing points, a depth-map is produced.

The production of a depth-map comprises three steps. The first step, the detection of edge, reduces noise as a preprocessing for input images by applying median filters and derives geometric features by applying canny edge. The second step, the production of vanishing lines and points, derives

straight lines through Hough transformation, calculates the intersection points of the straight lines, and estimates the locations of vanishing points through the range of locations of defined vanishing points. The final step, the estimation of vanishing point locations and production of a depth-map, estimates the locations of vanishing points through the range of locations of vanishing points, sets the steps of depth, and finally produces a depth-map based on the locations of vanishing points.

To produce multi-scale panorama image, the relationships between images are set internally. Multi-scale panorama image is produced through the blending process which makes input images into panoramic pictures. According to the change of image scale, the exact corresponding locations with neighboring images are automatically calculated, securing the exactness within sub pixels and stitching the images. Fig 3. shows a general configuration map of multi-scale 3D panorama augmented system using a depth-map suggested in this study. Fig 4. shows the steps of stitching images derived from a depth-map, producing depth-map panorama, and embodying multi-scale 3D panorama augmented system.

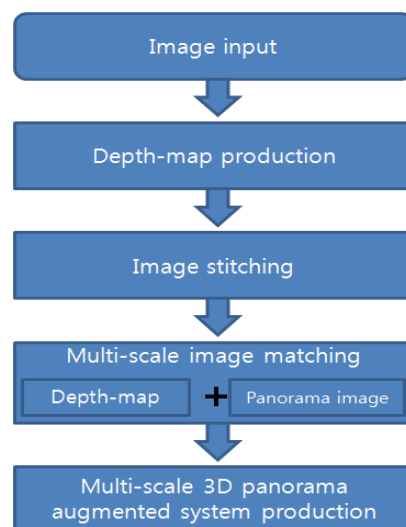


Fig. 3. Multi-scale 3D panorama augmented system configuration map.

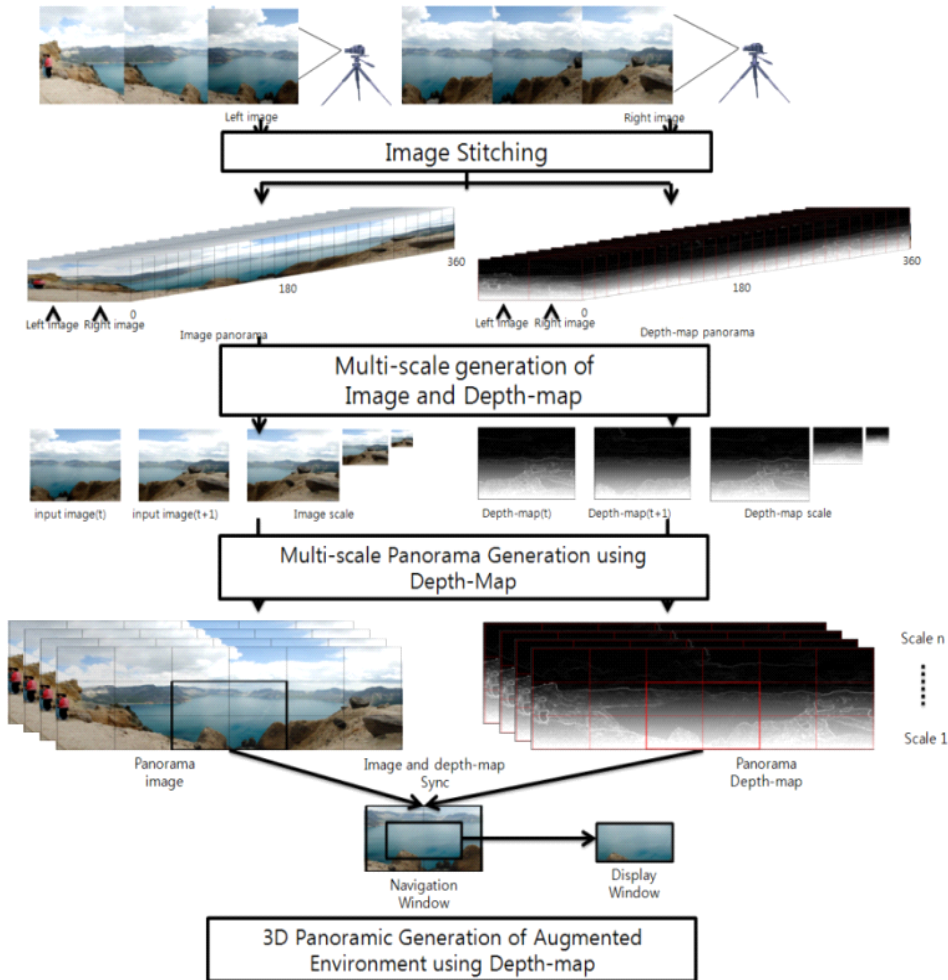


Fig. 4. Embodiment of multi-scale 3D panorama augmented system using depth-map.

5. CONCLUSION AND FUTURE WORKS

In this study, features matching images found out to add a sense of 3D effect to the augmented environment and a depth-map is used to embody 3D panorama, producing the augmented system with a sense of reality. This system will overcome the limits of 2D panorama technologies and provide users with a natural navigation, providing an improved sense of reality and immersion. Multi-scale 3D panorama augmented system using depth-map makes users feel as if they are at the location which photo-taking is taken and seeing the surroundings, in other words of a sense of 3D effect. Multi-scale

3D panorama augmented system provides high-quality 3D images with a sense of reality which has a free vision of close objects and distant backgrounds. This suggests a possibility in that it can be utilized in establishing augmented reality system based on images, attracting a attention from many people.

Further studies will be carried out based on computer vision technologies including image processing and interpretation technologies and computer graphics configuring media images with a sense of reality from the interpreted images to commercialize tourist attractions tours with a sense of reality for interaction with users and em-

bodied images and high-quality digital content of public relations office.

REFERENCES

- [1] D. Burschka, D. Cobzas, Z. Dodds, G. Hager, M. Jagersand, and K. Yerec, "Recent Methods for Image-Based Modeling and Rendering," *IEEE Virtual Reality 2003 Tutorial 1*, pp. 1-85, 2003.
- [2] H.K. Shum, S.B. Kang, and S.C. Chan, "Survey of Image-based Representations and Compression Techniques," *IEEE Transaction On Circuits and Systems for Video Technology*, Vol. 13, No. 11, pp. 1020-1037, 2003.
- [3] B.E. Shelton and N.R. Hedley, "Using Augmented Reality for Teaching Earth-Sun Relationships to Undergraduate Geography Students," *Proceeding of First IEEE International Augmented Reality Toolkit Workshop*, pp. 1-8, 2002.
- [4] Z. Yan, X. Ma, and S. Tao, "The Virtual Display Study on 3D Panorama in Tourist Areas-Take Shilin World Geopark as an Example," *Proceeding of the International Symposium on Information Processing*, pp. 229-232, 2009.
- [5] S.M. Chun, J.Y. Choi, S.H. Kim, Y.C. Cho, and K.S. Park, "A Study on Depth Map Quantization for Multiview Image Generation," *Journal of Korean Society For Computer Game*, Vol. 26, No. 4, pp. 219-226, 2013.
- [6] S. Curti, D. Sirtori, and F. Vella, "3D Effect Generation from Monocular View," *IEEE Proceeding of the 3D Data Processing Visualization and Transmission*, pp. 550-553, 2002.
- [7] T. Hassner and R. Basri, "Example Based 3D Reconstruction from Single 2D Images," *Proceeding of the Conference on Computer Vision and Pattern Recognition Workshop*, pp. 15-22, 2006.
- [8] L. Juan and O. Gwun, "SURF Applied in Panorama Image Stitching," *Proceeding of 2nd International Conference on Image Processing Theory Tools and Applications*, pp. 495-499, 2010.
- [9] M. Brown and D.G. Lowe, "Automatic Panoramic Image Stitching using Invariant Features," *International Journal of Computer Vision*, Vol. 74, No. 1, pp. 59-73, 2007.
- [10] R. Szeliski, "Video Mosaics for Virtual Environments," *IEEE Computer Graphics and Applications*, Vol. 16, No. 2, pp. 22-30, 1996.
- [11] R. Szeliski and H.Y. Shum, "Creating Full View Panoramic Image Mosaics and Environment Maps," *Proceeding of the Conference on Computer Graphics and Interactive Techniques*, pp. 251-258, 1997.
- [12] N. Chiba, M. Minoh, and H. Kano, "Feature-Based Lens Distortion Correction for Image Mosaicing," *Proceeding of the IAPR Workshop on Machine Vision Applications*, pp. 607-610, 2000.
- [13] J. Luo, S.S. Shin, H. J. Park, and O. B. Gwun, "Stitching for Panorama based on SURF and Multi-band Blending," *Journal of Korea Multimedia Society*, Vol. 14, No. 2, pp. 201-209, 2011.
- [14] S.I. Cho, K.J. Kim, K.J. Ban, K.W. Park, C.Y. Kim, and E.K. Kim, "3D Panorama Generation Using Depth-Map Stitching," *Journal of Information and Communication Convergence Engineering*, Vol. 9, No. 6, pp. 780-784, 2011.



KIM Cheeyong

He received the B.S. degree, M.S. degree and Ph.D. degree from Inje University, Republic of Korea in 1991, 1994 and 2000 respectively. He is visiting professor at Oxford University in 2007, and visiting professor of

the Digital Clothing Center at Seoul National University in 2012. Currently, he is a professor of Department of Visual Information Engineering at Dong-eui University. Prof. Kim held private exhibitions four times both in Seoul and China and joined over 210 international group exhibitions. With his deep interest in 3D Animation, Fractal & Chaos Design, Computational Simulation, and 3D virtual fashion fitting system.



Jong-Chan Kim

He received his B.S. degree from Suncheon National University in 2000, his M.S degree from department of computer science, Suncheon National University in 2002, his Ph.D. degree from department of computer science,

Suncheon National University in 2007. His current research interests are image processing, computer graphics, AR, games and HCI.



Eung-Kon Kim

He received B.E. degree Chosun University in 1980, his M.E. degree from department of computer engineering, Hanyang University in 1986, his Ph.D degree form department of computer engineering, Chosun Uni-

versity in 1994, In 2011, he joined department of computer science, the Suncheon National University, where he is a professor. His current research interests are image processing, AR, computer graphics, multimedia and HCI.