Object Extraction with Excessive Disparities in 3D Stereoscopic Images

Sang Hyun Kim, Jeong Yeop Kim, and Gil Ja So, *Member, IACSIT*

Abstract—In this paper, we propose the method to extract objects with excessive disparities in 3D stereoscopic images using cost function considering the intensity and the depth information. The traditional region segmentation method such as CLRG (Centroid Linkage Region Growing) used in the general 2D images is processed only based on intensity information. 3D stereoscopic images have the additional information that is depth. In the proposed method, first the excessive disparity candidate regions are decided using the pre-defined threshold in the disparity-map. Next the extracted excessive disparity candidate regions are labeled to object regions and the other regions are labeled to background regions. The regions labeled as object are set as region that will be segmented in 3D stereoscopic left image. The proposed segmentation method considering the intensity and the depth is applied to these regions in 3D stereoscopic left image. Finally to eliminate small regions, the morphological filter is used.

Index Terms—3D Stereoscopic images, binary, disparities, region-based segmentation, discomfort and fatigue, cost function.

I. INTRODUCTION

Recently, the demand for 3D stereoscopic images is growing sharply. Typically, the real 3D stereoscopic image contents are made by bilateral approach using two left and right cameras, in other words, each side of the camera to acquire images at the same time. If 3D stereoscopic images using left and right cameras are inputted left and right eyes. The human brain fuses the left and right images and, from retinal disparity. Because our eyes are horizontally separated, each eye has its own perspective of the world, and thus both eyes receive slightly different. Stereoscopic imaging is the perception of depth that is constructed based on the difference between these two retinal images. This approach should ensure uniformity of left and right images to provide the safe and comfortable viewing [1], [2].

Stereoscopic image viewing comfort is one of the main problems that should be solved before the mass market proliferation of stereoscopic 3D content services. The complaints of some viewers about visual discomfort such as eyestrain and associated symptoms including nausea and headaches present a barrier in the popularization of stereoscopic 3D technology. Some visual discomfort may occur due to excessive demand on the accommodation-vergence linkage, that is excessive

disparities, fast local and global motion, and various stereoscopic distortions [3].

The conventional stereo matching estimates the relation between corresponding pairs, calculates disparities and depth information. It is the method [4]-[6] to describe 3D surface to acquire depth information for all pixels in the image using interpolation.

The viewer discomfort issue is especially important for children as their vision systems have not grown mature and their inter-pupil distances are shorter than adults. It is need to prevent that the children will not be exposed to excessive disparities and thus have no additional eye fatigue due to 3D stereoscopic image [7].

In Yuan's method [7], the discomfort and fatigue in 3D stereoscopic images occur from the recognition error of depth information made by excessive positive and negative disparities. These the recognition error of depth information is detected by measurement of perceived depth on screen and fixed by depth tuning methods as the depth shift and the depth scaling. Here, if some perceived depths are outside the pre-defined threshold in histogram of disparity map, those perceived depths are regarded as the excessive disparities and fixed by depth tuning methods.

In detection of the excessive disparities based on the histogram, there are some problems. Because objects in 3D stereoscopic images have three-dimensional position and depth depending on the difference of lighting effects, such as smoking. The excessive disparities are also generated locally by particular objects. The histogram based method dose not correctly detects it because of inaccuracy in disparity detection made from occlusion or noise. The many small isolated excessive disparity regions may be occurred [8]-[10]. To reduce the problem, it is required segmentation method additionally.

In general, image segmentation methods divide an input image into disjoint sub-regions satisfying homogeneity and connectivity conditions: a region is called homogeneous if all its pixels are homogeneous. A region is called connected if there is a connected path between any two pixels. There are three major approaches to segmentation as follows: binary imaging with threshold, edge-based method, and region-based method [11], [12]. LoG operator based method in edge-based methods forms closed-loop and extracts contour of each objects [13]. This method segments the input image based on vertical length of the object and the variance of the input image.

In Kim's method [14], the method to detect the excessive disparities using segmentation based on LoG operator and binarization is proposed. Because this method is operated in disparity-map, if the adjacent objects have very similar disparities, it's impossible to distinguish these regions. It's required the method to separate objects using pixel intensities

S. H. Kim is with the Department of Computer Engineering of the University of Young-San at Yangsan, Kyungnam (e-mail: ksh50@ ysu.ac.kr).

J. Y. Kim is with the School of Undeclared Majors, University Collage of the University of Young-San at Busan (e-mail: neocopy@ ysu.ac.kr).

G. J. So is with the Department of Cyber & Police Science of the University of Young-San at Yangsan, Kyungnam (e-mail: kjso@ ysu.ac.kr).

in 3D stereoscopic left or right images.

In this paper, we propose the method to extract objects with excessive disparities in 3D stereoscopic images using cost function considering the pixel intensities and the depth information, disparities. The traditional region segmentation method such as CLRG (Centroid Linkage Region Growing) [15] used in the general images is processed only based on intensity information. 3D stereoscopic images have the additional information that is depth or disparities. In the proposed method, first the excessive disparities regions are extracted using the pre-defined threshold in the disparity-map. Next the extracted regions are labeled to object regions and the other regions are labeled to background regions. The regions labeled as object are set as region that will be segmented in 3D stereoscopic left image. The proposed segmentation method considering the intensity and the depth is applied in these regions. To eliminate small regions, the morphological filter is used.

II. THE PREVIOUS RESEARCHES

A. Yuan's Method [7]

This method proposes the depth tuning method. The overall algorithm flow is shown in Fig. 1. The method starts by estimating 2D image disparities between the left and right images. The viewer's preference, e.g., input through the TV remote, is also fed into the algorithm. A human visual model is then applied to predict a comfortable range of image disparities.

The comfortable disparity range is compared to the image disparities as well as viewer's preference, leading to the decision of how much depth tuning is needed. Finally a view synthesis algorithm is applied to generate newer versions of the left and right images.
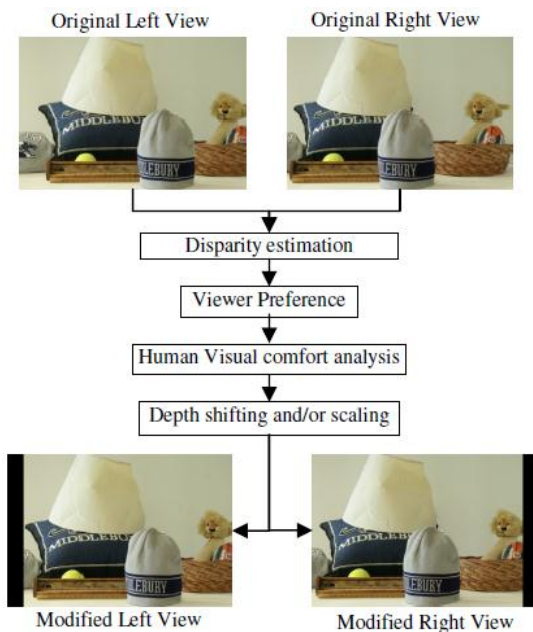

Fig. 1. Overview of the stereoscopic content tuning in Yuan's method.

To estimate disparities, a regularized block matching algorithm is used between left and right images. Human visual comfort analysis algorithm computes the comfortable range of disparities. The depth tuning method, namely "depth shifting," applies lateral shifting to the two images. The next depth tuning method, "depth scaling", produces a new disparity map with scaled disparities [16].

In Fig. 2, this method's depth tuned image result from "Monsters vs Aliens 3D". This method automatically detects that the image disparities (see Fig. 2(c)) are out of the comfort zone, and applies depth scaling method. In this image, the only big red ball region is located out of the comfort zone. The disparities of the big red ball are shifted and scaled.

(a) Original image pair from "Monsters vs Aliens".
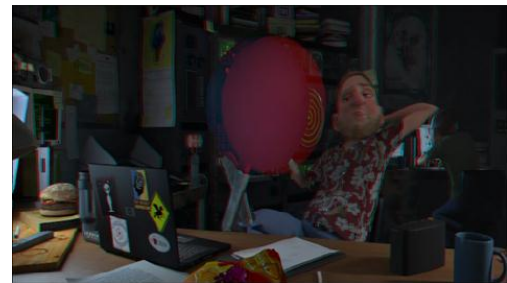
(b) Tuned image pair from "Monsters vs Aliens".
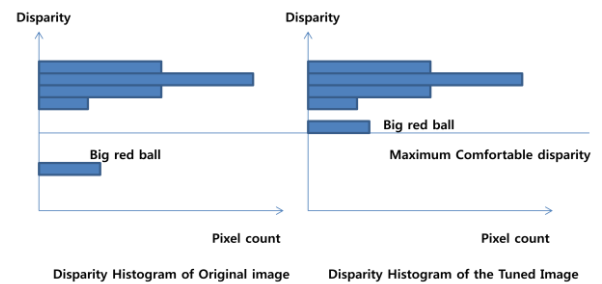
(c) Visual Comport analysis of the 3D image.

Fig. 2. 3D depth tuning and analysis results on the "Monsters vs Aliens 3D" image [7].

B. Kim's Method [14]

Kim's method extracts the excessive disparity regions using the adaptive LoG operator and binarization. Contour extraction using LoG operator is shown in Fig. 3. In Fig. 3, Double-differential $\nabla^2 G$ is applied to the input image. The zero-crossing point is extracted as a contour. Here, I(m, n) is a input image, E(m, n) is a contour image made by zero-crossing points in the double-differential image, h(m, n).
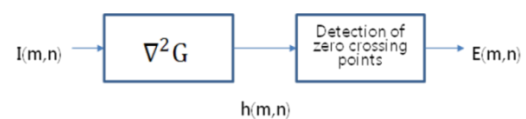

Fig.3. Block diagram of contour extraction using LoG operator.

In Fig. 3, the equation for the procedure of contour extraction using LoG operator with the double-differential $\nabla^2 G$ is as follows:

$$h(m,n) = \nabla^2 G(m,n) \times I(m,n) \qquad (1)$$

Here, LoG operator with the double-differential $\nabla^2 G$ is as follows.

$$\nabla^2 G(m,n) = \frac{1}{2\pi\sigma^4}(2 - \frac{m^2+n^2}{\sigma^2})\exp(-\frac{m^2+n^2}{2\sigma^2}) \qquad (2)$$

If contour extraction using LoG operator is applied to the input images, then the contour images with the various resolutions, depending on the value of the standard deviation σ can be obtained. We can obtain the variety of contour images using LoG operator by adjusting the standard deviationσ.

If the standard deviation is smaller, the effect of smoothing by the Gaussian filter is smaller. It can be obtain the detail contour image but many small regions are also occurred. In contrast, if the standard deviation is greater, the distinct contour can be detected but the objects and background can be merged. In contour extraction using LoG operator, it is important to decide the proper standard deviation.

In [13], to solve these problems, the method to automatically set the standard deviation based on the vertical position and the size informed by user, and the variance of the image, was proposed.

The 3D stereoscopic images have the disparity-map with depth information between the left and right images. The distribution in this disparity map is generated by the distribution of objects in 3D space. We can set the standard deviation proportional to the distribution as the variance in this disparity map.

The block diagram of the proposed method is as follows. In Fig. 1, the first the proposed method generates the disparity-map based on left image. The histogram made from the disparity-map and the standard deviation of LoG operator is decided from this. The next, region segmentation in disparity-map is performed and binariztion using threshold is also processed. In the segmented regions, the overlapped regions with the binarization results are detected as the excessive disparity regions.

A block matching algorithm is used to calculate disparities between left and right images. The standard deviation in LoG operator is estimated using the constant of proportionality with the standard deviation of disparity-map. It is also adjusted by user preference manually. The threshold is decided in histogram. A pixel in disparity-map which has value is over the pre-defined threshold, classified into the excessive disparity pixel. In the segmented regions using LoG operator, the overlapped regions with the binary results over the pre-defined threshold are detected as the excessive disparity regions.

In general, binary images made by threshold have the rough boundaries and the many small regions. In the proposed method, because region segmentation using LoG operator has the smooth boundary and small region elimination function, this region segmentation using LoG operator can make up for these disadvantages occurred from binary results.

In Fig. 5, this method's excessive disparity extraction result from "Art" in Middlebury University dataset [17]. Fig. 5(a) shows the original image. There are some objects on the table as crayon, plaster head, paper cup, and small plaster

manipulation. Fig. 5(b) shows the disparity-map's histogram of "Art" ground image given in Middlebury dataset. For the real world image which is not animation image such as "Monsters vs Aliens 3D", the specific object cannot be distinguished as object out of the comfort zone. In Fig. 5(b), the red circle area is out of the comfort zone. Fig. 5(c) shows the contour image extracted as excessive disparity region, region out of the comfort zone. In Fig. 5(c), the extracted excessive disparity region includes a small plaster manipulation and the front part of table. To adjust depth, the depth tuning algorithm can adjust the depth a small plaster manipulation but cannot adjust the front part of table linked to another part. In disparity-map, it is impossible distinguish a small plaster manipulation and the front part of table have similar disparities. Therefore to distinguish objects with similar disparities and to do the effective depth tuning, in the actual 3D stereoscopic image as well as in disparity-maps. It is necessary segmentation and combination in 3D stereoscopic images.
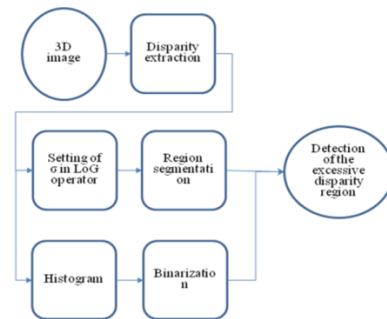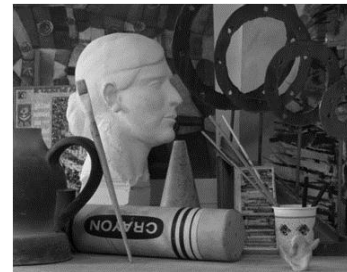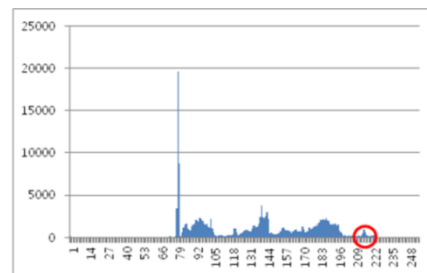

Fig. 4. Block diagram of Kim's method.

(a) Original image from "Art".

(b) Histogram of disparity-map of "Art".

(b) Contour image of the excessive disparity region from "Art".
Fig. 5. Excessive disparity region extraction results using Kim's method on the "Art" image [14].

III. THE PROPOSED METHOD

We propose the method to extract objects with excessive disparities in 3D stereoscopic images using cost function considering the pixel intensities and the depth information, disparities. The conventional segmentation method considers the only pixel intensities in 2D images. Because 3D stereoscopic images have the additional information that is depth or disparities, the proposed method considers the both pixel intensities and the disparities.

The block diagram of the proposed method is as follows: In Fig. 6, the first the proposed method generates the disparity-map based on left image. The histogram is calculated from the disparity-map. The excessive disparity candidate regions are decided using the pre-defined threshold in the disparity-map. The extracted excessive disparity candidate regions are labeled to object regions and the other regions are labeled to background regions. The regions labeled as object are set as region that will be segmented in 3D stereoscopic left image. The proposed segmentation method considering the intensity and the depth information is applied to these regions in 3D stereoscopic left image. Finally to eliminate the protruding regions or the small regions, the morphological filter by means of the post processing is used.

Disparity calculation: A block matching algorithm is used to calculate disparities between left and right images. In our algorithm, we cannot adopt the additive occlusion detection and processing algorithm. To reduce the impression of simulation results come from the disparity calculation error, we used the ground images which are provided by Middlebury dataset by means of disparity-map.

Binarization based on histogram: A pixel in disparity-map which has value is over the pre-defined threshold, classified into the excessive disparity candidate pixel. The threshold can be decided by user preference or recommended based on human visual comfort analysis algorithm [7].
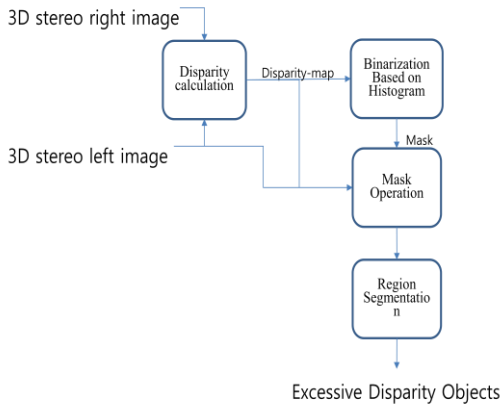

Fig. 6. Block diagram of the proposed method.

Mask operation: The extracted excessive disparity candidate pixels are labeled to object pixels and the other pixels are labeled to background pixels. In 3D stereoscopic left image, the overlapped pixels with the pixels which are labeled to object in disparity-map preserve the original pixel value. The other pixels except the overlapped pixels are reset to '0'.

Region Segmentation: The masked 3D stereoscopic left image will be segmented using image segmentation method. The goal of image segmentation is to partition an input image into disjoint sub-regions satisfying homogeneity and connectivity conditions: a region is called homogeneous if all its pixels are homogeneous. A region is called connected if there is a connected path between any two pixels.

Precisely, a segmentation of an image R is a partition of R into disjoint sub-regions R_1, R_2, ..., R_N satisfying the following conditions:

1) $\bigcup_{i=1}^{N} R_i = R$
2) $C(R_i)$ = TRUE for i=1, 2, …, N and C: connectivity
3) $H(R_i)$ = TRUE for i=1, 2, …, N and H: homogeneity predicate
4) $H(R_i \cup R_j)$ = FALSE for any adjacent R_i and R_j

There are three major approaches to segmentation as follows: binary imaging with threshold, edge-based method, and region-based method.

In the various conventional segmentation methods, region-based method is robust and to the image noise better than the others. CLRG (Centroid Linkage Region Growing) [15] is the region-based method which is the simple and the most frequently used. This method's structure is as follows:
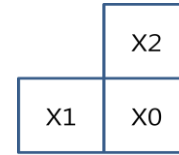

Fig. 7. Region growing geometry of centroid linkage method.

Step 1: Scan the image in the raster scanning order.

Step 2: Compare the pixel X_0 in scanning to the mean of a neighboring region to which each neighboring pixel X_i, i = 1, 2, belongs.

Step 3: If the pixel X_0 and the mean of the closest region is close enough, then add the pixel to the region and if the merged region is close enough with another neighboring region then merge the two regions. Else establish a new region having the pixel X_0. The cost function $C(\cdot)$ is used in comparison with a pixel X_0 and neighboring region R_m can be represented as follows:

$$C(R_m, X_0) = \sqrt{(\overline{R_m} - I(X_0))^2} \tag{3}$$

Here, $\overline{R_m}$ is the intensity mean of the neighboring m-th region R_m, and $I(X_0)$ is intensity value of a pixel X_0.

Step 4: Move to the next pixel and then go to step 2.

This segmentation method as CLRG based on the only pixel intensity have problem cannot distinguish the neighboring regions have the similar intensities at region boundary partially. In the entire image the specific target object detection is not easy. Although in our method, the mask operation improve the performance of the target object segmentation to reduce range of segmentation, the additional method is required. 3D stereoscopic images have the additional information that is depth or disparities. We can modify the equation (3) to cost function considering the additional depth information as follows:

$$C(R_m, X_0) = \sqrt{(\overline{R_m} - I(X_0))^2 + \alpha(\overline{D_m} - d(X_0))^2} \tag{4}$$

Here, $\overline{D_m}$ is the disparity mean of the neighboring m-th region R_m, and $d(X_0)$ is disparity value of a pixel X_0, α is the

constant.

IV. SIMULATION RESULTS

"Art" image of 3D stereoscopic images in Middlebury University dataset were used in this paper's simulation. In Middlebury University dataset, 3D stereoscopic images are composed of images with the full and half size generated by 3 types of lighting and exposure. The original "Art" images are color images. It is converted into gray image. Fig. 5(a) shows the original "Art" image. Fig. 8 shows the disparity-map. In Fig. 8, the more pixel intensity is bright, the more its disparity value is strong. A small plaster manipulation and the front part of table are the brightest in this disparity map.

Fig. 8. Disparity-map of art image.

Fig. 9 shows the binary image made by using threshold in disparity-map Fig. 5(b). In Fig. 9, the extracted excessive disparity candidate region, the white region, includes a small plaster manipulation and the front part of table. To adjust depth, the depth tuning algorithm can adjust the depth a small plaster manipulation but cannot adjust the front part of table. It is linked to another part. In disparity-map, it is impossible distinguish a small plaster manipulation and the front part of table have similar disparities. Therefore to distinguish objects with similar disparities and to do the effective depth tuning, in the actual 3D stereoscopic image as well as in disparity-maps, it is necessary to do segmentation and combine the results.

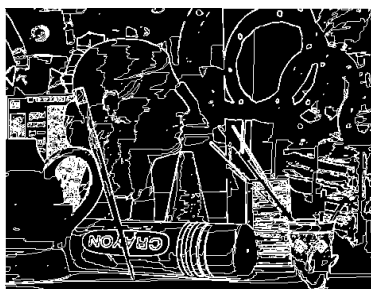Fig. 9. Binary image using the threshold in disparity-map.

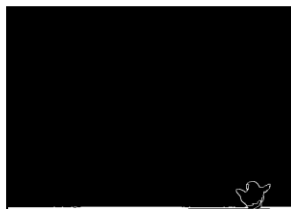
Fig.10. Region segmentation result using CLRG.

Fig. 10 shows the region segmentation result using CLRG in the "Art" 3D stereoscopic left image. In Fig. 10, the specific objects cannot be distinguished. To extract the specific objects, it is necessary to define the region of interest (ROI) previously and to reduce range of segmentation. In the proposed method, the binary image can be used by means of ROI.

Fig. 11 shows the segmentation results in the masked 3D stereoscopic left image using the binary image in Fig. 10. Fig. 11(a) is the result of the conventional method using the only intensity. Fig. 11(b) is the result of the proposed method using the intensity and depth information. In Fig. 11(a), a small plaster manipulation is partitioned to a number of small regions but the front part of table is made to a region linked to small plaster manipulation. It's impossible to separate the small plaster manipulation from the front part of table using the post processing. On the other hand, in Fig. 11(b), the proposed method's result, the front part of table is partitioned to a number of small regions because its disparities are spread. The small plaster manipulation is made to almost a region. The small plaster manipulation can be extracted by using the post processing as morphological filter to eliminate the protruding regions or the small regions. Fig. 12 shows the extracted object as the excessive disparity region in the proposed method.

(a) Region segmentation using intensity in left image.

(b) Region segmentation using intensity and disparity in left image
Fig. 11. Region segmentation results in the masked left image.

Fig. 12. The extracted object as the excessive disparity region.

V. CONCLUSION

Stereoscopic image viewing comfort is one of the main problems that should be solved before the mass market proliferation of stereoscopic 3D content services. The discomfort and fatigue in 3D stereoscopic images occur from the recognition error of depth information made by excessive positive and negative disparities.

In this paper, we proposed the method to extract objects with excessive disparities in 3D stereoscopic images using cost function considering the intensity and the depth

information. In the proposed method, first the excessive disparity candidate regions are decided using the pre-defined threshold in the disparity-map. Next the extracted excessive disparity candidate regions are labeled to object regions and the other regions are labeled to background regions. The regions labeled as object are set as region that will be segmented in 3D stereoscopic left image. The proposed segmentation method considering the intensity and the depth is applied to these regions in 3D stereoscopic left image. Finally to eliminate small regions, the morphological filter is used. In simulation results, we extracted region with the greatest disparities as the excessive disparity region effectively. Therefore to distinguish the target objects from and the other regions with similar disparities and to do the effective depth tuning, it is necessary to upgrade the proposed segmentation method.

REFERENCES

[1] C. W. Choi, " HUD factor classification of online FPS Game for 3D stereoscopic application," M.S. thesis, Dept. of Media Broadcasting, Graduate School of Information Sciences, Soongsil Univ., Seoul, Korea, 2011.

[2] S. Choi, "Technical status and prospect of 3D stereoscopic images," *Journal of Korea Information Processing*, vol. 17, no. 4, pp. 4-11, 2010.

[3] M. Lambooji, W. A. IJsselsteijn, and I. Heynderickx, "Visual discomfort of 3D TV: Assessment methods and modeling," *Displays*, vol. 32, Issue 4, Oct. 2011.

[4] J. S. Kim, "A study on a balance and registration of non-rectified stereo images," Ph.D. Dissertation, Chung-Ang Univ., Seoul, Korea, 2009.

[5] S. Kim, G. So, and J. Kim, "Intensity correction of 3D stereoscopic images using region segmentation," in *Proc. 34th Conf. Korea Information Processing Society*, Seoul, 2010.

[6] J. S. Lee, S. D. Park, S. H. Han, and C. Y. Kim, "Implementation of 3-dimension stereoscope using 3D graphics," in *Proc. Conf. Korea Multimedia Society*, vol. 13, no. 2, 2010, 11, pp. 184-187.

[7] C. Yuan, H. Pan, and S. Daly, "61.3: Stereoscopic 3D content depth tuning guided by human visual models," in *Proc. SID Symposium Digest of Technical Papers,* vol. 42, issue 1, June 2011, pp. 916–919.

[8] Y. J. Zhang, "Improving the accuracy of direct histogram specification," *Electronics Letters*, vol. 28, issue. 3, pp. 213-214, 1992.

[9] H. Tuo, L. Zhang, and Y. Liu, "Multisensory aerial image registration using direct histogram specification," in *Proc. IEEE Int'l Conf. Sensing and Control,* vol. 2, 2004, pp. 807-812.

[10] A. Mancini, "Disparity estimation and intermediate view reconstruction for noble application in stereoscopic video," M.S. thesis, McGill University, 1998.

[11] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Reading, MA: Addison-Wesley, 1992.

[12] A. K. Jain, "Fundamental of digital image processing," Printice-Hall, NJ: Englewood Cliffs, 1989.

[13] M. Jung, "3D organ segmentation in ultrasound images using LoG operator," M.S. thesis, Kyung-Pook National Univ., Daegu, Korea, 2002.

[14] S. H. Kim, J. Y. Kim, and G. J. So, "Segmentation of 3D stereoscopic images using LoG operator for the detection of excessive disparities," *International Journal of Computer and Communication Engineering*, vol. 2, no. 3, pp. 314-318, May, 2013.

[15] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer Vision, Graphic and Image Processing 29*, pp. 100-132, 1985.

[16] H. Pan, C. Yuan, and S. Daly, "3D video disparity scaling for preference and prevention of discomfort," in *Proc. SPIE 7863, Stereoscopic Displays and Applications XXII*, 786306, Feb. 2011.

[17] Middlebury. [Online]. Available: http://vision.middlebury.edu/vision/

Sang Hyun Kim received the diploma in electronics engineering from the Kyung-Pook National University of Daegu, South Korea in 1991, and MS and PhD degrees both in image processing of electronics engineering from the same university, in 1993 and 1998, respectively. He was with the Samsung Medison Co. Ltd, in Korea as a senior researcher from 1998 to 2001. From September 2001, he is at the Department of Computer Engineering of the University of Young-San at Yangsan, Kyungnam as an associate professor. His doctoral dissertation work focused on the development of fractal image coding method and transmission for image processing, image communication and computer vision problems. The topics he has worked on include virtual reality, ultrasound image processing, medical image processing, image feature extraction and classification, and 3D stereoscopic images. His current research is on region-based intensity correction in 3D stereoscopic images. He is a member of the KMMS (Korea Multimedia Society) and KIPS (Korea Information Processing Society).
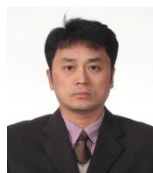
Jeong Yeop Kim received the diploma in electronics engineering from Kyung-Pook National University of Daegu, South Korea in 1990, and MS and PhD degrees both in image processing of electronics engineering, from the same university in 1992 and 2000, respectively. He was with the Samsung Advanced Institute of Technology in Korea as a senior research staff from 1994 to 2000. From March 2001, he is with the School of Undeclared Majors, University Collage of the University of Young-San at Busan as an associate professor. His doctoral dissertation work focused on detecting and adjusting the white point from single color image. The topics he has worked on include color printing, error diffusion, color correction and 3D stereoscopic images. His current research is on correction of color data in 3D stereoscopic images. He is a member of the KMMS (Korea Multimedia Society) and KIPS (Korea Information Processing Society).

Gil Ja So received the diploma in computer science from the Dong-Eui University of Pusan, South Korea in 1994, and MS and PhD degrees both in natural language processing of computer engineering, from the Pusan National university, South Korea in 1997 and 2012, respectively. From September 2001, she is with the Department of Cyber & Police Science of the University of Young-San at Yangsan, Kyungnam, as a lecturer. Her doctoral dissertation work focused on the Korean Grammar Checker for parsing Korean documents and correcting context-sensitive words in documents. The topics she has worked on include virtual reality, game development and artificial intelligence of non-players of game, and 3D stereoscopic images. Her current research is on region-based intensity correction in 3D stereoscopic images. She is a member of the KMMS (Korea Multimedia Society) and KIPS (Korea Information Processing Society).