

특집논문 (Special Paper)

방송공학회논문지 제20권 제3호, 2015년 5월 (JBE Vol. 20, No. 3, May 2015)

<http://dx.doi.org/10.5909/JBE.2015.20.3.368>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

관심맵과 에지 모델링을 이용한 2D 영상의 3D 변환

김만배^{a)†}

Generation of Stereoscopic Image from 2D Image based on Saliency and Edge Modeling

Manbae Kim^{a)†}

요 약

2D영상의 3D변환 기술은 3D 디스플레이 및 3DTV에 기본적으로 장착된 기술로 꾸준히 연구 및 상업화가 진행된 기술이다. 이 기술은 3D 입체영상 콘텐츠 부족을 해결할 수 있다는 장점이 있다. 3D변환은 정지영상으로부터 다양한 깊이단서를 이용하여 깊이맵을 추출한 후에, DIBR(Depth Image Based Rendering)로 입체영상을 생성한다. 특정 영상외에는 영상에서 신뢰성 있는 단서가 있는 경우는 많지 않다. 따라서 3D변환 기술은 일반 영상에서도 우수하고, 일관된 입체영상이 생성하는 것이 중요하다. 이러한 관점에서 본 논문에서는 상기 조건을 만족할 수 있는 3D변환 방법을 제안한다. 주 기술로 최근 다양한 분야에서 활용되는 관심맵과 에지를 활용한다. 깊이맵을 생성하기 위해서 기하적 투영, 근접 모델 및 바이노믹 필터를 활용한다. 실험에서는 제안한 방법을 24개의 2D 비디오 콘텐츠에 적용하였고, 입체감 및 시각적 피로도 등의 주관적 평가를 통해 3D 콘텐츠의 우수한 만족도를 확인하였다.

Abstract

3D conversion technology has been studied over past decades and integrated to commercial 3D displays and 3DTVs. The 3D conversion plays an important role in the augmented functionality of three-dimensional television (3DTV), because it can easily provide 3D contents. Generally, depth cues extracted from a static image is used for generating a depth map followed by DIBR (Depth Image Based Rendering) rendering for producing a stereoscopic image. However except some particular images, the existence of depth cues is rare so that the consistent quality of a depth map cannot be accordingly guaranteed. Therefore, it is imperative to make a 3D conversion method that produces satisfactory and consistent 3D for diverse video contents. From this viewpoint, this paper proposes a novel method with applicability to general types of image. For this, saliency as well as edge is utilized. To generate a depth map, geometric perspective, affinity model and binomic filter are used. In the experiments, the proposed method was performed on 24 video clips with a variety of contents. From a subjective test for 3D perception and visual fatigue, satisfactory and comfortable viewing of 3D contents was validated.

Keyword : 3D conversion, depth map, saliency, edge modeling, affinity model

a) 강원대학교 컴퓨터정보통신공학과(Dept. of Computer and Communications Engineering, Kangwon National University)

† Corresponding Author : 김만배(Manbae Kim)

E-mail: manbae@kangwon.ac.kr

Tel: +82-33-250-6395

ORCID:<http://orcid.org/0000-0002-4702-8276>

※ This research was supported by Kangwon National University.

※ 이 논문의 연구결과 중 일부는 "IWAIT 2015"에서 발표한 바 있음.

· Manuscript received January 12, 2015; revised February 24, 2015; accepted February 24, 2015.

I. Introduction

The generation of a stereoscopic image from a 2D image has been investigated over the past decades due to the success of 3D TVs and displays^[1-8]. Most of conversion methods derive a depth map of each frame and then use DIBR (Depth Image Based Rendering) to synthesize the stereoscopic view.

To generate a depth map from a given 2D image, diverse methods have been proposed according to the principle of a human visual system. Some of them are depth from motion^[6], depth from defocus^[8], depth from geometrical linear perspective and gradient plane assignment^[4], depth from shadow^[7] and so forth. Most depth estimation algorithms make use of the combination of aforementioned multiple monocular depth cues. Therefore, they are expected to work well for some particular images that contain suitable cues. In other words, if the depth cues do not deliver sufficient information, the algorithms might fail, yielding the production of uncomfortable 3D images. Furthermore, the accurate detection of a depth cue type is also needed and is somehow a difficult task. In case of incorrect type classification, a wrong depth map would be also obtained. This observation requires the necessity of the design of a global conversion method that produces a satisfactory 3D perception regardless of image contents as well as depth cues.

The proposed method is composed of four main components; (1) visual saliency estimation, (2) affinity model and binomic filter, (3) edge modeling, and (4) depth generation. The overall block diagram is shown in Fig. 1. Given an RGB image, an edge map is obtained from a grayscale image. As well, a saliency map is extracted from an RGB image. Due to the lack of distance information in the saliency map, we incorporate a geometrical perspective cue to the saliency map. Then, a binomic filter as well as an affinity model is applied on the saliency map to reduce the saliency discontinuities present at neighboring pixels. This re-

sult is combined with an edge map and the transformed edge map are binomic-filtered with the saliency map. Finally a depth map is generated and subsequently left and right images are constructed by a DIBR method.

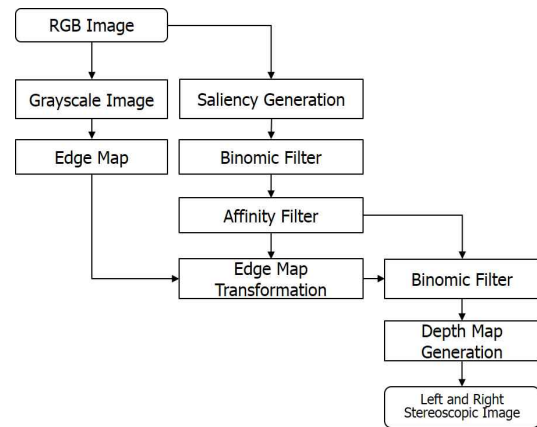


그림 1. 제안방법의 블록 다이어그램
Fig. 1. Block diagram of the proposed method

II. Saliency Map Generation

The saliency generation has gained much interest over past decades in the diverse fields^[9-16]. Contrast is an important factor that affects visual attention in static images. Whether an object is perceived as salient or not depends greatly on the distinctiveness between itself and the background. Color is one of main features for saliency detection. For colors, red/green and blue/yellow are two strong contrast color pairs. Recently, 2D-to-3D conversion researchers have applied the saliency to depth map generation^[11-16]. In this paper, a global contrast-based method proposed by Zhai and Shah^[9] is adopted because it is suitable to real-time processing unlike other complex methods. This method is simple to implement, but yet efficient for a baseline saliency production. Moreover, the performance is comparative to other methods.

The important fact is that only saliency containing dis



그림 2. 각 영상별 관심객체의 다른 위치를 보여주는 관심객체의 예들 (수작업으로 지정). 개인마다 다른 관심객체를 선택할 수 있다.
 Fig. 2. Examples of images showing the salient objects (manually marked in yellow box). Individual human can select different salient objects. The different positions of the saliency objects are observed.

tance (or depth) information is useful data. As observed in Fig. 2, natural scenes contain the arbitrary locations and unknown relative distances of salient objects in the image. For instance, the saliency objects (manually marked in yellow box) can be located at the top or bottom as well as in the left, middle and right. Note that the location of saliency objects depends on human decision especially in complex scenes. Therefore an additional procedure is needed to compensate for the lack of the geometric information shortage of the saliency map.

Most of natural images have general properties that a top area is far from a camera and that a bottom region is close to the camera due to inherent geometric perspective. The examples are also shown in Fig. 2. The upper regions of the four example images have a larger distance compared with lower regions. Such characteristic will be utilized in the saliency construction. Furthermore, the location of salient objects imposes the uncertainty on background regions.

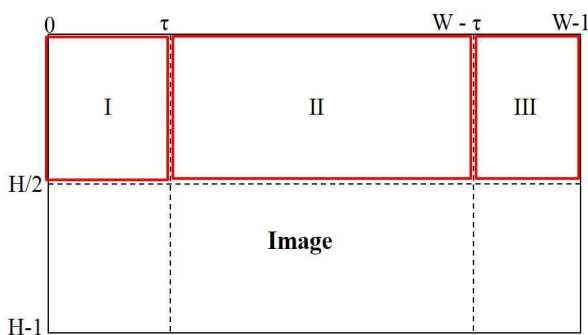


그림 3. 관심객체의 가능한 위치를 고려한 3개의 영역
 Fig. 3. Regions I, II, III are used to consider possible locations of salient objects. Region I = [0,0]×[τ,H/2], Region II = [τ,0]×[W-τ,H/2], and Region III = [W-τ,0]×[W-1,H/2]

For instance, the yellow boxes considered as salient objects in each image are in the left, middle, middle, and right, respectively. Therefore to deal with all possible cases, the three regions I, II, and III are combined to compensate for the locational uncertainty of salient objects as illustrated in Fig. 3. Given a W x H image, three regions I, II, and III are configured in the upper region.

The weights of the three regions are illustrated in Fig. 4, where an weight w[i] is assigned to ith region. The mathematical formula are defined by

$$\begin{aligned}
 w[1] &= -\frac{j}{W-1} + 1.0 \\
 w[2] &= \begin{cases} \frac{j}{(W/2)} & \text{if } j \leq W/2 \\ -\frac{j}{(W/2)} + 2 & \text{if } j > W/2 \end{cases} \\
 w[3] &= \frac{j}{W-1}
 \end{aligned} \tag{1}$$

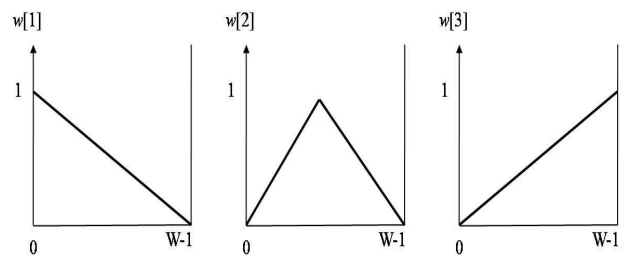


그림 4. 3개의 영역에 대응하는 가중치 함수들
 Fig. 4. Weight functions w[1], w[2] and w[3] are associated with regions I, II, and III, respectively

One of saliency methods is to directly use Red, Green, Blue channels. In [9], a mean value is computed from an

entire image. On the contrary, in our method, the mean values of R, G and B channels are computed only from upper three regions.

$$\begin{aligned}\overline{R}_1 &= \sum_{(i,j) \in I} R(i,j), \overline{G}_1 = \sum_{(i,j) \in I} G(i,j), \overline{B}_1 = \sum_{(i,j) \in I} B(i,j) \\ \overline{R}_2 &= \sum_{(i,j) \in II} R(i,j), \overline{G}_2 = \sum_{(i,j) \in II} G(i,j), \overline{B}_2 = \sum_{(i,j) \in II} B(i,j) \quad (2) \\ \overline{R}_3 &= \sum_{(i,j) \in III} R(i,j), \overline{G}_3 = \sum_{(i,j) \in III} G(i,j), \overline{B}_3 = \sum_{(i,j) \in III} B(i,j)\end{aligned}$$

where (\overline{x}) is a mean of x.

The purpose of computing a mean value only from an upper region is to obtain higher saliency in the lower region. Then weighted saliency maps of the three channels are computed by

$$\begin{aligned}S_R &= w[1] \cdot |R - \overline{R}_1| + w[2] \cdot |R - \overline{R}_2| + w[3] \cdot |R - \overline{R}_3| \\ S_G &= w[1] \cdot |G - \overline{G}_1| + w[2] \cdot |G - \overline{G}_2| + w[3] \cdot |G - \overline{G}_3| \quad (3) \\ S_B &= w[1] \cdot |B - \overline{B}_1| + w[2] \cdot |B - \overline{B}_2| + w[3] \cdot |B - \overline{B}_3|\end{aligned}$$

The first saliency map S_1 can be made by either the maximum of three saliency maps or their average.

$$S_1 = \text{MAX}(S_R, S_G, S_B) \text{ or } S_1 = \frac{(S_R + S_G + S_B)}{3} \quad (4)$$

In the first method, R, G and B are directly utilized. The usage of a single saliency is not enough for getting satisfactory outcome. To solve this, we adopt another approach. We utilize the variation of RGB, from which a different saliency can be produced. Using Eq. (5), the transformed colors a, b, c are derived from the RGB.

$$\begin{aligned}a &= |R - \frac{G+B}{2}|, \quad b = |\frac{|B+G|}{2} - \frac{|B-G|}{2} - R|, \\ c &= |B - \frac{R+G}{2}| \quad (5)\end{aligned}$$

Similar to Eq. (2), the averages of a, b, c are derived for the three regions.

$$\begin{aligned}\overline{a}_1 &= \sum_{(i,j) \in I} a(i,j), \overline{b}_1 = \sum_{(i,j) \in I} b(i,j), \overline{c}_1 = \sum_{(i,j) \in I} c(i,j) \\ \overline{a}_2 &= \sum_{(i,j) \in II} a(i,j), \overline{b}_2 = \sum_{(i,j) \in II} b(i,j), \overline{c}_2 = \sum_{(i,j) \in II} c(i,j) \quad (6) \\ \overline{a}_3 &= \sum_{(i,j) \in III} a(i,j), \overline{b}_3 = \sum_{(i,j) \in III} b(i,j), \overline{c}_3 = \sum_{(i,j) \in III} c(i,j)\end{aligned}$$

For each channel, the saliency map is made by an weighted average of transformed colors.

$$\begin{aligned}A &= w[1] \cdot |a - \overline{a}_1| + w[2] \cdot |a - \overline{a}_2| + w[3] \cdot |a - \overline{a}_3| \\ B &= w[1] \cdot |b - \overline{b}_1| + w[2] \cdot |b - \overline{b}_2| + w[3] \cdot |b - \overline{b}_3| \quad (7) \\ C &= w[1] \cdot |c - \overline{c}_1| + w[2] \cdot |c - \overline{c}_2| + w[3] \cdot |c - \overline{c}_3|\end{aligned}$$

The second saliency map S_2 is obtained by either the maximum of A, B, C or their average.

$$S_2 = \text{MAX}(A, B, C) \text{ or } S_2 = \frac{(A + B + C)}{3} \quad (8)$$

To obtain the best saliency map, we tested a variety of combinations of S_1 and S_2 . From this experiment, we found that the following relation outperforms other combinations. A final saliency map S is obtained from S_1 multiplied by normalized S_2 and is expressed by

$$S(i,j) = S_1(i,j) \times \frac{S_2(i,j)}{S_2^{\max}} \quad (9)$$

where S_2^{\max} is a maximum value of S_2 .

We performed the saliency generation methods on diverse images, whose scene complexity varies from low to high and the results are shown in Fig. 5.

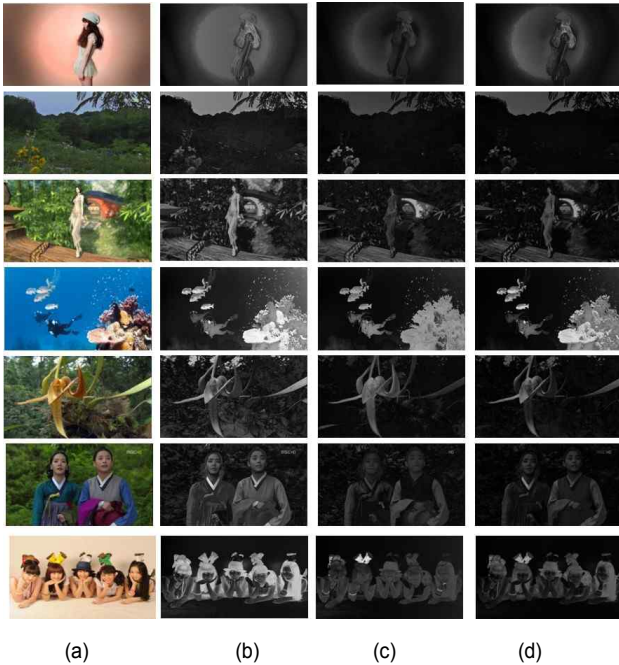


그림 5. 제안 방법으로 얻어진 관심도맵. (a) RGB 입력, (b)는 식 (4), (c)는 식 (8)에서 얻어진 관심도맵, 및 (d)는 식 (9)로부터 얻어진 관심도맵
 Fig. 5. Diverse saliency maps obtained by the proposed method. (a) is RGB input image. (b) and (c) are obtained from Eqs. (4) and (8), respectively, and (d) is a final saliency made by Eq. (9)

III. Binomic Filter and Affinity Model

As observed in Fig. 5, the lack of consistent saliencies are present in most of images. For instance, different values in the inner area of a woman in the first row of Fig. 5(a) are noticed and thus an additional processing is needed. We summarize the problems as follows: (1) an identical object

has different saliency values. Especially the inconsistent values at the boundary are apparent, (2) the saliency of the inner region of a foreground object is not consistent within its boundary, and (3) a background is relatively homogeneous except some particular regions, but still needs constant values. Such problems prevent the saliency maps from being used as a depth. To solve this, we employ a binomic filter^[17] and an affinity model^[19-21]. The binomic filter is performed on the saliency map to solve the inconsistency between the inner region and the boundary. The aim of the binomic filter is to fill in the inner region using the boundary values. As well, the affinity model is to soothe the discontinuities of the close-by pixels.

A) Binomic Filter

Since the saliency uses a color, different saliency values might spread over an identical object. To alleviate such problem, one of effective methods is a binomic filter^[17], where its elements are binomic numbers which are created as a sum of the corresponding two numbers in Pascal's triangle. The effect of the binomic filter is illustrated in Fig. 6. Suppose that the distribution of S along x axis is like Fig. 6(a). Then its distribution is changed to Fig. 6(b) after the filter is applied. It is observed that the large variation of the pixels values is much reduced.

We expand this filter to an image as follows; Based on an $N \times N$ pixel block, then we convolve $S(i,j)$ with its scaled image $S^T(i,j)$.

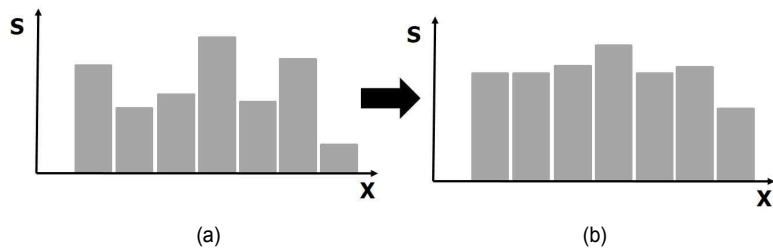


그림 6. 주변 픽셀의 차이를 줄여주는 바이노믹 필터의 특성

Fig. 6. The binomic filter lessens the difference between neighboring pixels

$$S_{bino}(i,j) = S(i,j) \otimes S^\tau(i,j) \\ = \frac{1}{N^2} \sum_{m=-N/2}^{m=N/2} \sum_{n=-N/2}^{n=N/2} S(i,j) \cdot S^\tau(i-m,j-n) \quad (10)$$

where S^τ is the scaled version. For the sake of clarity, the scaling used here is the value scaling. For instance of 1-D, if $S = [120 \ 140 \ 80]$, S^τ becomes $[30/3, 140/3, 80/3]$ in the scale of $\tau = 1/3$. τ varies at $[0, 1]$. The larger it is, the more the output is saturated. The result is shown in Fig. 7. It is observed that saliency discontinuities of Fig. 7(a) are much alleviated in Fig. 7(b)

A) Affinity Model

Pixels with same RGB have identical saliency values. Therefore, two neighboring pixels with different colors, but belonging to the same object might have different saliency

values that produce the inconsistent depth. To solve this problem, we employ an affinity model in order to alleviate the boundary discontinuities at the identical object region.

In the usage of saliency data, defining an affinity model gained by integrating local grouping local cues such as saliency and boundary is of importance. As mentioned, a single object can be represented by multiple different saliency data, resulting in different depth values. The affinity model used in the segmentation can solve the aforementioned problem^[19-21]. Close-by pixels with similar saliency values likely belong to the same segment. The color-based affinity model is defined by an exponential function.

$$\psi_c(i,j) = \exp\left(-\frac{\|x_i - x_j\|}{\sigma_c}\right) \\ \psi_s(i,j) = \exp\left(-\frac{\|S_i - S_j\|}{\sigma_s}\right) \quad (11)$$

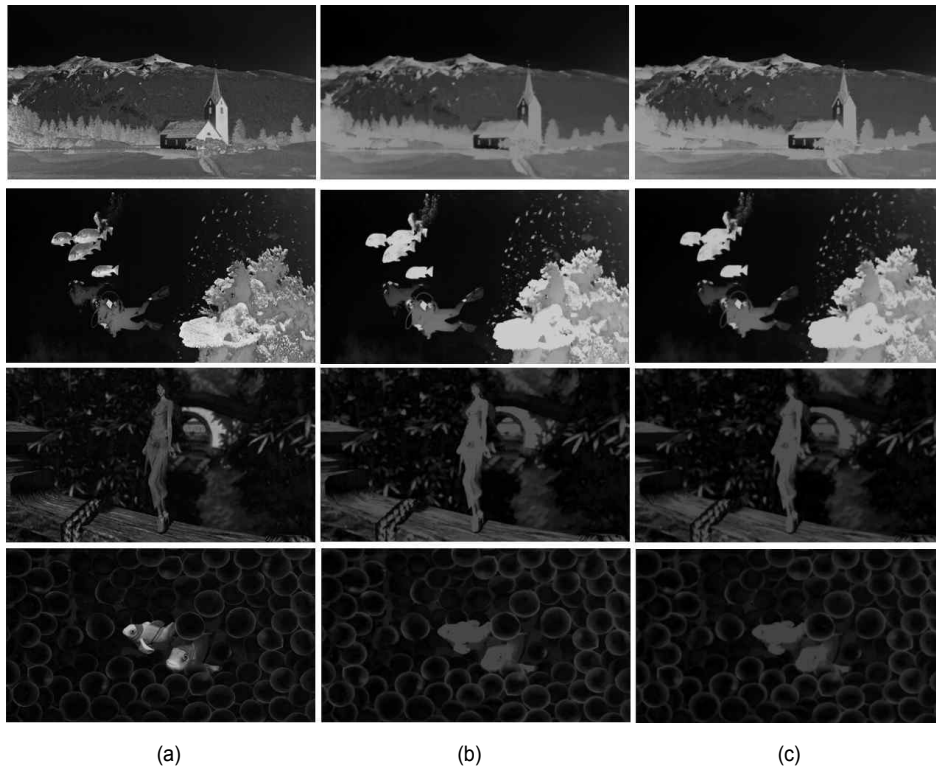


그림 7. 바이노믹 필터의 결과. (a) 입력 영상, (b) 바이노믹 필터를 적용한 출력 영상, 및 (c) 근점 모델을 적용한 출력 영상
 Fig. 7. The results obtained after a binomic filter. (a) input image, (b) image after binomic filtering ($\tau=0.5$) and (c) image after affinity modeling ($N=4$).

where x_i and s_i denote the position and saliency values of pixel i respectively and, α_c and α_s control the weights of the two factors.

The combined model is used to design a better affinity model. Two models can be simply combined with a parameter to produce a combined model ψ_m

$$\psi_m = \alpha\psi_c + (1-\alpha)\psi_s \text{ 또는 } \psi_m = \psi_c \cdot \psi_s \quad (12)$$

where α is a weight of $[0, 1]$.

We apply the affinity model to the binomic-filtered image using the convolution.

$$\begin{aligned} S_{aff}(i,j) &= \psi_m(i,j) \otimes S_{bino}(i,j) \\ &= \frac{1}{N^2} \sum_{m=-N/2}^{m=N/2} \sum_{n=-N/2}^{n=N/2} S_{bino}(i-m,j-n) \cdot \psi_m(m,n) \end{aligned} \quad (13)$$

The resulting image is shown in Fig. 7(c). As observed, the filtered output shows more smoothed outputs. Two close-up images are shown in Fig. 8, where the surface of resulting images after the affinity model become more smooth.

IV. Edge Map Transformation

Edge plays an important role in the proposed method^[22].

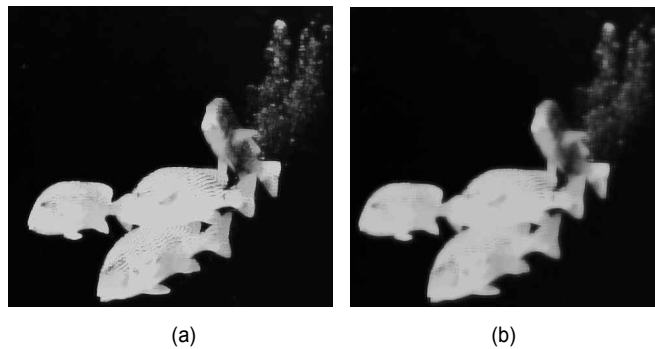


그림 8. 그림 7의 두 번째 영상의 확대영상. (a) 입력영상 및 (b) 픽셀값 변화를 감소시키는 근접모델
Fig. 8 Close-up of the second image in Fig. 7. (a) input image and (b) affinity model reduces the pixel value variation

The presence of connected edge boundary in the object provides useful information. On the other hand, note that the edge does not provide any depth information. Based on this fact, the edge processing helpful to depth generation focuses on the smooth edge preservation as well as the edge adaptation to saliency map.

The edge processing procedure is shown in Fig. 9. Given an edge map, we decompose it into multiple sub-images in the vertical direction. Then, the Bezier surface modeling is applied to entire image. Considering the edge ratio of a subimage, we adapt the edge map to the surface model and then a transformed edge map is derived.

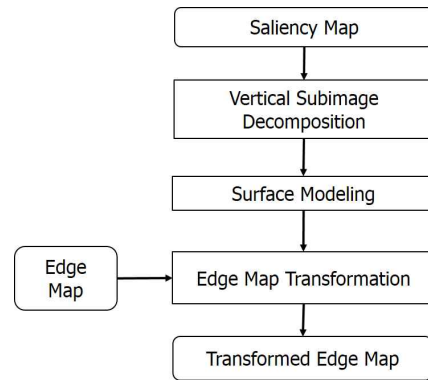


그림 9. 에지맵 변환 방법
Fig. 9. Edge map transformation

An image is decomposed into K subimages in the verti-

cal direction as in Fig. 10. Then a maximum saliency value is computed for each kth subimage, SB^k .

$$S_{\max}^k = \text{MAX}_{(i,j) \in SB^k} S_{\text{aff}}(i,j) \quad (14)$$

Then, an edge ratio E_R is derived from each subimage.

| | | |
|------------|-------------------------------|-----|
| | 0 | W-1 |
| Subimage 1 | $SB^1 \rightarrow S_{\max}^1$ | |
| Subimage 2 | $SB^2 \rightarrow S_{\max}^2$ | |
| | · | |
| | · | |
| | · | |
| Subimage K | $SB^K \rightarrow S_{\max}^K$ | |
| H-1 | | |

그림 10. 수직방향으로 K개의 서브영상으로 분할.
 Fig. 10. An image is decomposed into K subimages in the vertical direction

Since edge contains no depth information, the 3D depth is diminished if we use edge values. Therefore, if an edge ratio is large, we decrease the saliency value and vice versa. For each subimage, the edge ratio E_R^k is computed by

$$E_R^k = \frac{E^k}{N^k} \quad (15)$$

where N^k is the number of pixels and E^k is the number of edge pixels. E_R^k is at $[0,1]$.

For each subimage, we compute a saliency maximum value Q^k that is a maximum saliency value multiplied by $(1-E_R^k)$.

$$Q^k = S_{\max}^k \cdot (1 - E_R^k) \quad (16)$$

Q^k will act as a control point for surface modeling. From

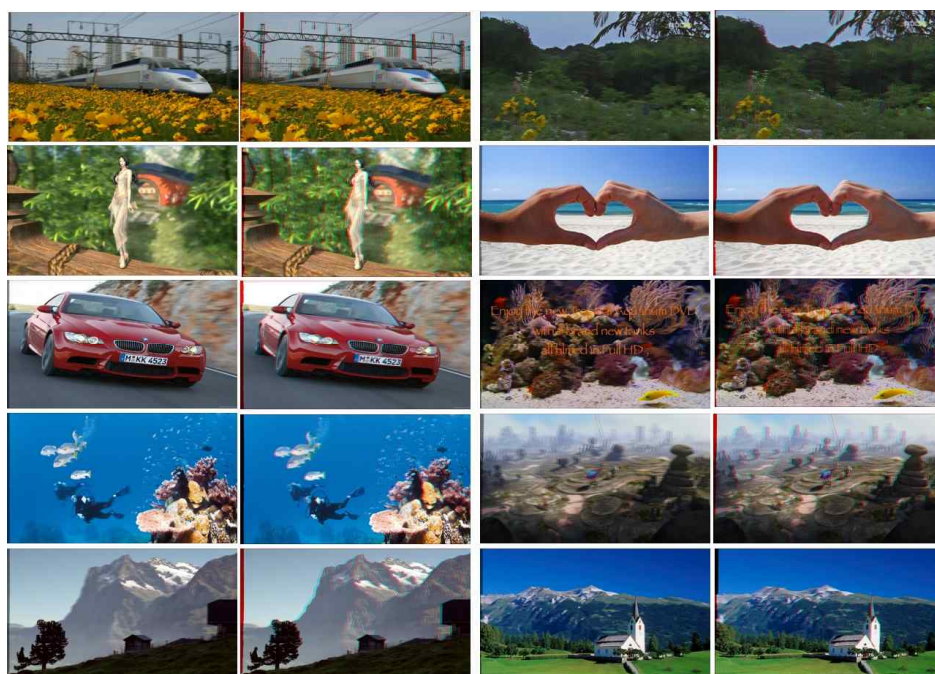


그림 11. 생성된 입체영상. (좌) 인터레이스 영상 및 (우) 아나글리프 영상
 Fig. 11. Output stereoscopic images. (left) interlaced images and (right) anaglyph images

this, it is apparent that the saliency is more dominant at the dense edge subimage and less at the sparse region. As verified in the experiment, this is expected to add more 3D perception. The surface is modeled by Bezier curve or surface using K control points and then continuous surface SB is generated. Finally, a binomic filter is applied to saliency and edge map and a final depth map is constructed.

V. Experimental Results

The proposed method has been tested on twenty four video clips as in Table 1. The stereoscopic images of some test images from a video are shown in Fig. 11. 3D formats are interlaced and anaglyph. The resolution of all the sequences is FHD (Full High Definition) 1920 x 1080. The duration of video clips is 300 ~ 10,000 frames.

To measure 3D subjective tests, we examined 3D perception grade as well as visual fatigue. Twenty subjects participated in the experiment. They are 3D experts from industrial and academic fields who have much experience in viewing and evaluating 3D contents. The test videos were captured from commercial movies, TV dramas, sports, animation movies to demonstrate that the evaluation is independent from contents types. The duration of viewing time of each sequence is proportional to the number of frames. The viewing distance is 3 meter from display monitor.

SSCQS (Single Stimulus Continuous Quality Scale) subjective test was performed. Human subjects observed the stereoscopic videos on LG FHD 40" 3DTV and evaluated the 3D perception and visual fatigue^[23]. The scale of the grade is [1, 5]. For the 3D perception, the grade of 5 is very good and 1 is bad. The grade of 5 is no fatigue and 1 is severe fatigue for the visual fatigue. As shown in Table 1, the 3D grade is 3.68 and visual fatigue is 3.49 on the average. 3.68 of the 3D perception grade is between mild

and good 3D. Regarding the performance limitation of automatic 3D conversion, the grade is satisfactory. As well, 3.49 of the visual fatigue is at the range of mild and less fatigues. One of the functionalities of 3D conversion is the ability to control a depth range. Therefore, if viewers feel any visual discomfort, they can control the strength of depth by adjusting a maximum parallax.

표 1. 3D 주관적 평가 결과

Table 1. 3D subjective evaluation results

| Video Contents | 3D Perception | Visual Fatigue | No. of Frames |
|------------------|---------------|----------------|---------------|
| dongyi | 4.2 | 3.8 | 500 |
| Girls generation | 3.0 | 3.4 | 1,000 |
| acua | 4.1 | 3.0 | 380 |
| airplane | 3.0 | 3.8 | 150 |
| ani | 3.8 | 3.7 | 8,000 |
| Alps | 3.6 | 3.3 | 10,000 |
| avatar | 3.9 | 3.2 | 12,000 |
| anicar | 4.2 | 3.8 | 400 |
| fan | 3.4 | 3.8 | 200 |
| football | 2.8 | 3.2 | 5,000 |
| gumi | 3.5 | 2.7 | 10,000 |
| IU | 4.2 | 3.7 | 4,000 |
| Kara | 4.3 | 2.5 | 450 |
| Konan | 3.7 | 3.2 | 300 |
| Kpop | 3.8 | 3.6 | 3,000 |
| natural | 4.1 | 3.8 | 1,200 |
| oriFrance | 4.3 | 3.3 | 5,000 |
| penguin | 3.7 | 3.8 | 350 |
| plant | 3.6 | 4.1 | 800 |
| Rainbow | 4.4 | 3.2 | 500 |
| rio | 2.7 | 4.0 | 5,000 |
| secret | 2.9 | 3.7 | 1,200 |
| starwars | 3.6 | 3.6 | 5,000 |
| Wonder girls | 3.6 | 3.5 | 5,000 |
| Average | 3.68 | 3.49 | |

VI. Conclusion

In this paper, a novel 3D conversion method was proposed. The proposed method stems from the fact that depth cues are not enough except some particular images,

thus motivating the necessity of a general conversion method. Our method is designed to meet such requirement using saliency and edge modeling. For this, two saliency maps are utilized and are fused into a single saliency map that is a baseline for depth map generation. Additional geometric perspective is integrated into the saliency map in order to follow general natural scenes. Edge modeling is based on the saliency map as well surface representation. By combining the saliency and edge surface, satisfactory depth map could be obtained, leading to high 3D effect and low visual fatigue. An important aspect of all conversion methods is to provide stable 3D perception with reduced visual discomfort that was verified in our intensive video testing.

참 고 문 헌 (References)

- [1] S. Battiato, A. Carpa, S. Curti and M. la Cascia, "3D Stereoscopic Image Pairs by Depth-Map Generation," Proceedings of 3DPVT, 2004.
- [2] W. Tam and L. Zhang, "3D-TV Content Generation: 2D-To-3D Conversion," Proc. of IEEE ICME, 2006.
- [3] L. Zhang and W. Tam, "Stereoscopic image generation based on depth images for 3DTV," IEEE Trans. On Broadcasting, Vol. 51, Issue 2, June 2005
- [4] S. Kim and J. Yoo, "3D conversion of 2D video using depth layer partition," Journal of Broadcast Engineering, Vol. 15, No. 2, Jan. 2011.
- [5] I. Ideses, L. Yaroslavsky, B. Fishbain, "Real-time 2D to 3D video conversion," Journal of Real-Time Image Processing, vol. 2(1), pp. 2-9, 2007
- [6] F. Xu, G. Fr, X. Xie, and Q. Dai, "2D-to-3D Conversion Based on Motion and Color Mergence," 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video, pp. 205-208, May 2008.
- [7] J. Jung, J. Lee, I Shin, J. Moon and Y. Ho, "Improved depth perception of single view images", ECTI Transactions on Electrical Engineering, Electronics and Communications, Vol. 8, No. 2, Aug. 2010.
- [8] G. Surya and M. Subbaro, "Depth from defocus by changing camera aperture: a spatial domain approach," IEEE CVPR, pp. 61-67, 171993.
- [9] Y. Zhai, and M. Shah, "Visual attention detection in video sequences using spatiotemporal cues," Proceedings of the 14th annual ACM Int' Conf. on Multimedia, pp. 815-824, 2006.
- [10] Achanta, S. Hemami, F. Estrada and S. Susstrunk, "Frequency-tuned Salient Region Detection," IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1597-1604, 2009.
- [11] Y. Zhang, G. Jiang, M. Yu, and K. Chen, "Stereoscopic visual attention model for 3D video", Advances in Multimedia Modeling, 2010.
- [12] J. Kim, A. Baik, Y. Jung and D. Park, "2D-to-3D image/video conversion by using visual attention analysis," ICIP, 2009.
- [13] C. Chamaret, S. Godeffroy, P. Lopez, and O. Le Meur, "Adaptive 3D rendering based on region-of-interest", in Proceedings of SPIE, vol.7524, 2010.
- [14] N. Ouerhani and H. Hugli, "Computing visual attention from scene depth", IEEE International Conference on Pattern Recognition, 2000.
- [15] E. Potapova, M. Zillich, and M. Vincze, "Learning what matters: combining probabilistic models of 2D and 3D saliency cues," Computer Vision Systems, pp. 132-142, 2011.
- [16] J. Wang, M. Perreira, D. Silva, P. Le Callet, and V. Ricordel, "Computational Model of Stereoscopic 3D Visual Saliency," IEEE Transactions on Image Processing, 22(6): 2151-2165, 2013.
- [17] M. Sonka, V. Hlavac, and R. Boyle, Image Processing, Analysis and Machine Vision, 3rd ed., Thomson, 2008.
- [19] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient Graph-Based Image Segmentation," Int' J. Computer Vision, vol. 59, no. 2, pp. 167-181, 2004.
- [20] Y. Boykov and G. Funka-Lea, "Graph Cuts and Efficient N-D Image Segmentation," Int' J. Computer Vision, Vol. 70, Nno. 2, pp. 109-131, 2006.
- [21] T. Kim, K. Lee and S. Lee, "Learning full pairwise affinities for spectral segmentation," IEEE Trans. PAMI, Vol 35, No. 7, July 2013.
- [22] G. Robinson, "Edge detection by compass gradient masks," Computer Graphics and Image Processing, Vol. 6, pp. 492-501, 1977.
- [23] M.Tanimoto, T. Fujii, and K. Suzuki, "View Synthesis Algorithm in View Synthesis Reference Software 2.0 (VSRS2.0)", ISO/IEC JTC1/SC29/WG11 M16090, Lausanne, Switzerland, February 2008.

저 자 소 개



김 만 배

- 1983년 : 한양대학교 전자공학과 학사
- 1986년 : University of Washington, Seattle 전기공학과 공학석사
- 1992년 : University of Washington, Seattle 전기공학과 공학박사
- 1992년 ~ 1998년 : 삼성종합기술원 수석연구원
- 1998년 ~ 현재 : 강원대학교 컴퓨터정보통신공학과 교수
- orcid : <http://orcid.org/0000-0002-4702-8276>
- 주관심분야 : 3D영상처리, 깊이맵처리, 입체변환