

2D/3D 동영상 변환을 위한 그룹화된 객체별 깊이 정보의 차등 적용 기법

한성호¹, 홍영표², 이상훈^{3*}

¹광운대학교 대학원, ²한국국제대학교 병원관리학과, ³광운대학교 교양학부

Applying differential techniques for 2D/3D video conversion to the objects grouped by depth information

Sung-Ho Han¹, Yeong-Pyo Hong² and Sang-Hun Lee^{3*}

¹Kwangwoon University, Graduate School

²International University of Korea, Department of Hospital Management

³Kwangwoon University General Education

요 약 본 논문은 2D/3D 동영상 변환을 위한 그룹화된 객체별 깊이 정보의 차등 적용 기법에 관한 연구이다. 기존의 깊이 정보 획득 기법 중 움직임 정보를 깊이 정보로써 사용할 때 움직임이 존재하지 않는 객체의 경우 깊이 정보를 획득할 수 없어 해당 객체의 3D 효과를 얻을 수 없는 문제점이 있다. 본 논문에서는 이러한 문제점을 해결하고자 객체와 배경을 추출하고 객체에 움직임 정보를 이용한 깊이 정보를 할당하는 과정을 거친 후, 배경과 깊이 정보가 할당되지 못한 객체에 깊이 단서 중 하나인 상대적 높이 단서를 이용한 깊이 정보를 할당함으로써 모든 객체에서 3D 효과를 얻을 수 있는 기법을 제안한다. 제안한 기법으로 깊이맵을 생성하여 DIBR(Depth Image Based Rendering)으로 3D 영상을 생성하여 확인한 결과 움직임이 없는 객체에서도 3D 효과를 얻을 수 있음을 확인하였다.

Abstract In this paper, we propose applying differential techniques for 2D/3D video conversion to the objects grouped by depth information. One of the problems converting 2D images to 3D images using the technique tracking the motion of pixels is that objects not moving between adjacent frames do not give any depth information. This problem can be solved by applying relative height cue only to the objects which have no moving information between frames, after the process of splitting the background and objects and extracting depth information using motion vectors between objects. Using this technique all the background and object can have their own depth information. This proposed method is used to generate depth map to generate 3D images using DIBR(Depth Image Based Rendering) and verified that the objects which have no movement between frames also had depth information.

Key Words : 2D/3D Conversion, Cut Detection, Object Segmentation, Motion Estimation, Depth Information, Depth Image Based Rendering

1. 서론

3D 영상은 기존의 2D 영상보다 시청자에게 생동감 및 현실감을 느낄 수 있게 만드는 시각적 정보를 제공하고

실재감을 고조시킬 수 있다. 또한 최근 3D 영상 시청이 가능한 모니터, 핸드폰, TV 등 많은 디스플레이 기기의 제작이 활성화 되고 있다. 그러나 3D 영상의 수요 증가에 비해 공급량이 턱없이 부족한 문제점이 있어 사용자의

본 논문은 2011년도 광운대학교 교내 학술 연구비 지원에 의해 연구되었음.

본 논문은 중소기업청에서 지원하는 2011년도 산학연공동기술개발사업(No.00046375)의 연구수행으로 인한 결과물임을 밝힙니다.

*교신저자 : Sang-Hun Lee

Tel: +82-2-940-5287 e-mail: leesh58@kw.ac.kr

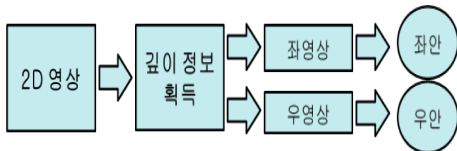
접수일 12년 01월 02일 수정일 (1차 12년 02월 06일, 2차 12년 02월 24일, 3차 12년 03월 02일) 게재확정일 12년 03월 08일

요구를 충족시킬 수 없는 것이 현실이다[1].

3D 콘텐츠는 깊이 카메라, 입체 카메라, 컴퓨터 그래픽 툴, 2D/3D 변환 기법 등을 이용하여 생성하게 된다. 그러나 카메라를 이용하는 방법의 경우 실제 촬영이 불가능한 영상의 3D 콘텐츠를 획득할 수 없으며 고가인 문제점이 있고, 컴퓨터 그래픽 툴을 이용하는 방법의 경우 영상의 제작 시간이 많이 소요되는 문제점을 갖는다. 반면에 2D/3D 변환 기법은 비용이 적게 들며 제작시간이 짧고, 기존의 2D 콘텐츠를 활용할 수 있기 때문에 콘텐츠 부족 문제의 해결 기법으로 2D/3D 변환이 활발하게 연구되고 있다[2].

입체감을 일으키는 요인 중 가장 중요한 요인으로 양안시차가 있다. 이것은 사람의 양 눈의 간격이 약 6.5cm 정도 떨어져있기 때문에 발생하는 시차로써 이러한 양안시차로 인하여 좌안과 우안은 각각 서로 다른 2차원 영상을 보게 되고, 이것이 뇌에서 융합됨으로써 입체감을 얻을 수 있다.

따라서 이러한 양안시차를 갖는 좌영상과 우영상을 생성함으로써 3D 영상을 볼 수 있게 된다. 이러한 원리를 이용하여 Okino 그룹의 MTD(Modified Time Difference)를 이용한 기법[3], Garcia의 공간 시간 보간(spatial-temporal interpolation)을 이용한 기법[4], 영상의 깊이 정보를 이용한 2D/3D 변환 기법 등이 제안되었다[5].



[그림 1] 깊이 정보를 이용한 2D/3D 변환 원리
[Fig. 1] 2D/3D conversion principle using depth information

깊이 정보를 이용하는 변환 기법의 경우 깊이 정보가 입체 영상을 생성하는데 가장 큰 역할을 하게 되므로 깊이 정보의 획득이 매우 중요하다. 이러한 깊이 정보는 단일 영상 또는 다중 영상의 움직임, 물체의 가림, 그림자, 중첩, 상대적 높이, 상대적 크기 등과 같은 깊이 단서들을 분석함으로써 얻어질 수 있다. 그 중 움직임 추정 기법으로 Optical flow, 블록 매칭 등의 움직임 추정 알고리즘이 제안되었다[6][7]. 그러나 움직임 정보가 존재하지 않는 영역의 경우 깊이 정보를 획득할 수 없는 문제점이 있다.

본 논문에서는 움직임 추정 기법을 이용한 깊이 정보 생성에서 발생하는 문제점을 해결하기 위해 영상을 객체와 배경으로 분리하여 움직임 정보가 존재하는 객체와 존재하지 않는 객체를 나눠 그룹화 시키고, 배경과 움직

임 정보가 존재하지 않는 객체에 상대적 높이 단서를 이용한 깊이 정보를 할당함으로써 배경 및 모든 객체에서 깊이 정보를 갖도록 하여 문제점을 해결하고, 최종 입체 영상 생성 시 모든 객체에서 입체 효과를 얻을 수 있도록 하였다.

2. 장면 분석

2.1 장면전환점(컷) 검출

움직임 추정의 경우 다중 프레임을 이용하기 때문에 장면의 전환이 이루어지는 부분에서 오류가 발생할 수 있다. 따라서 이러한 오류를 줄이기 위해 MPEG-7의 EHD(Edge Histogram Descriptor)를 이용한 에지 히스토그램을 생성하고, 식 (1)을 이용하여 프레임간 히스토그램 분포의 차이 $H_i(j) - H_{i+1}(j)$ 를 허용오차값인 $T1$ 값과 비교하여 허용오차 이상일 경우 장면전환점, 즉 컷이라고 판단하여 장면전환점을 검출하게 된다[8].

$$\sum_{j=1}^N |H_i(j) - H_{i+1}(j)| > T1 \quad (1)$$

2.2 객체 및 배경 추출

객체 추출은 식 (2)와 (3)을 이용한 그래프 분할 알고리즘인 Normalized cut을 이용하여 추출하고[9], 배경은 원영상에서 추출된 객체를 제거하여 얻어진다.

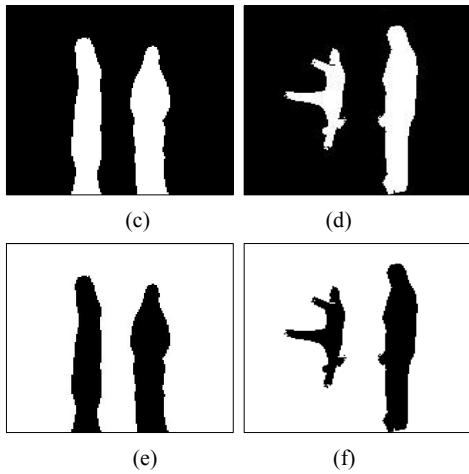
$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \quad (2)$$

$$w_{uv} = \exp \frac{-|F(u) - F(v)|^2}{\sigma_U^2} * \begin{cases} \exp \frac{-|X(u) - X(v)|^2}{\sigma_X^2}, & \text{if } \|X(u) - X(v)\| < r \\ 0, & \text{otherwise} \end{cases} \quad (3)$$



(a)

(b)



[그림 2] 입력 영상(a,b), 추출된 객체(c,d), 추출된 배경(e,f)
 [Fig. 2] Input image(a,b), Extracted object(c,d), Extracted background(e,f)

3. 깊이 정보 획득 기법

최종 깊이맵은 그레이스케일 영상으로 표현하기 위해 각 픽셀을 0~255 사이의 값을 갖도록 하고, 255에 근접할 수록 깊이 정보가 크며 가깝다는 것을 의미한다.

3.1 움직임 정보를 이용한 깊이 정보 획득

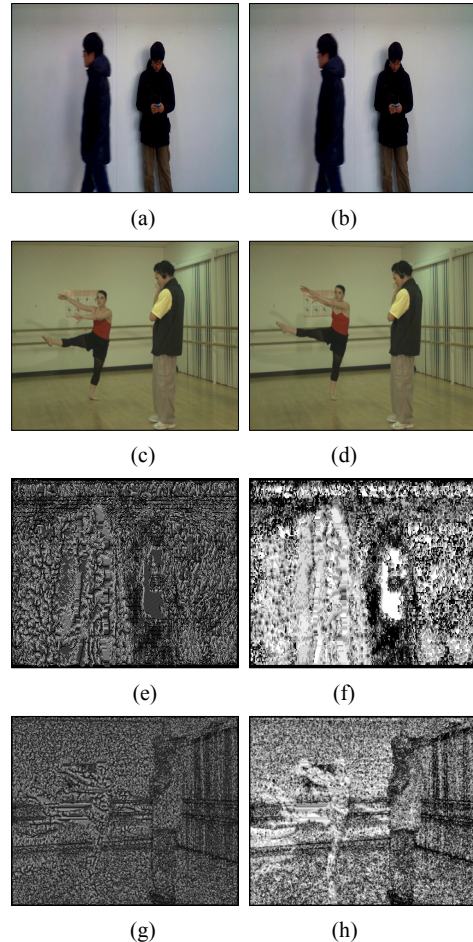
움직임 정보는 3x3 크기의 정합창을 이용하여, SAD(Sum of Absolute Difference) 값을 구해 가장 작은 SAD값을 갖는 픽셀을 구함으로써 획득되며, 다양한 방향에 대한 움직임 정보를 획득하기 위해 15x15의 블록을 설정하여 해당 블록의 중심으로부터 정합창의 시작위치까지의 거리를 픽셀 단위로 구하게 된다. 모든 픽셀에 대한 SAD값을 구하게 되면 연산 시간이 매우 커지기 때문에, 특정 픽셀 간격으로 값을 구한 후, 팽창연산을 통해 비어있는 값을 채움으로써 최종 움직임 정보를 획득한다. 블록 단위로 구하지 않고 픽셀 단위로 구함으로써 더 세밀한 움직임 정보를 얻을 수 있다.

$$C_{SAD}(x,y) = \sum_{i=0}^2 \sum_{j=0}^2 |F_n(x+i,y+j) - F_{n+1}(d_x+i,d_y+j)| \quad (4)$$

식 (4)의 F_n 은 n번째 프레임이며 F_{n+1} 은 n+1번째 프레임을 의미한다. d_x 와 d_y 는 설정한 블록의 중심점으로부터 이동된 x방향과 y방향의 움직임 크기를 나타낸다.

$$D(x,y) = d_x + d_y \quad (5)$$

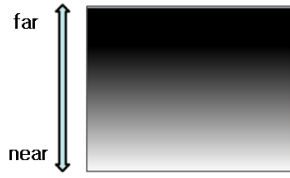
획득한 움직임 정보를 식 (5)를 이용하여 깊이정보로 변환한다. 여기서 $D(x,y)$ 는 (x,y) 좌표의 깊이 정보를 나타낸다.



[그림 3] n번째 영상(a,c), n+1번째 영상(b,d), 움직임 정보(e,g), 깊이 정보(f,h)
 [Fig. 3] n frame(a,c), n+1 frame(b,d), Motion information(e,g), Depth information(f,h)

3.2 상대적 높이 단서를 이용한 깊이 정보 획득

상대적 높이 단서란 물체가 영상의 하단에 있을수록 크게 인식되는 것으로써 이것을 이용하여 최하단의 깊이 정보를 255로 설정하여 상단으로 갈수록 0에 가까워지는 형태의 깊이 정보를 획득한다.



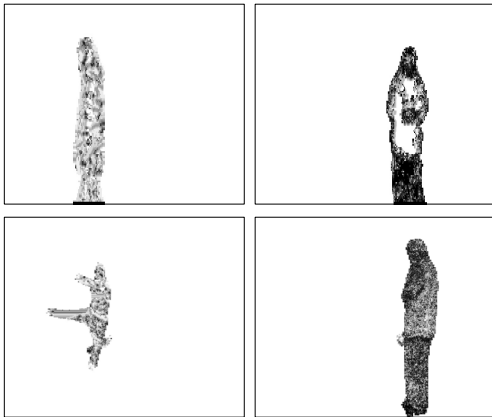
[그림 4] 상대적 높이 단서를 이용한 깊이 정보
[Fig. 4] Depth information using relative height cue

4. 그룹화된 객체별 깊이 정보 할당 및 보정

4.1 그룹화

획득한 두 가지 깊이 정보를 움직임 정보가 있는 객체와 없는 객체에 차등 적용하기 위해 그룹화 과정을 거치게 된다[10].

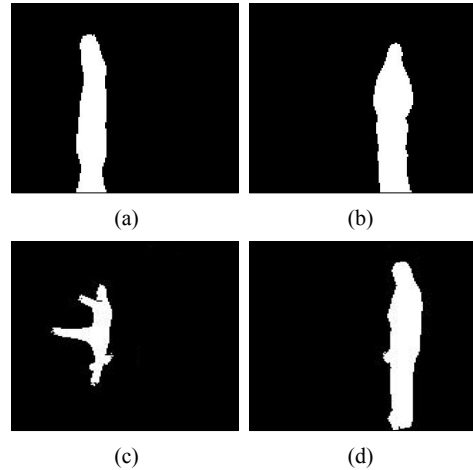
우선 움직임 정보를 이용하여 생성한 깊이 정보를 추출된 객체들에 할당하고, 식 (6)을 이용하여 각 객체별 움직임 정보가 존재하는지 판단하게 된다.



[그림 5] 움직임 정보를 이용한 깊이 정보 할당
[Fig. 5] Depth information allocation using motion information

$$B_{obj}/P_{obj} < T2 \quad (6)$$

식 (6)에서 B_{obj}/P_{obj} 는 0의 값을 갖는 픽셀의 비율을 의미하고, 이 값을 특정 임계값 $T2$ 와 비교하여 임계값 이하이면 움직임 정보가 존재하는 객체로 판단하고, 그렇지 않은 경우 움직임 정보가 존재하지 않는 객체로 판단하여 각각 그룹화 하게 된다.



[그림 6] 움직임 정보를 갖는 객체(a,c), 움직임 정보를 갖지 않는 객체(b,d)
[Fig. 6] objects which have moving information(a,c), objects which have no moving information(b,d)

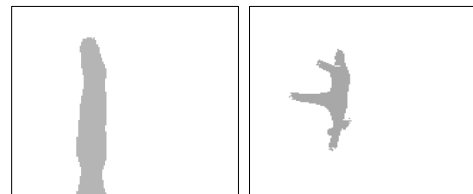
4.2 객체별 깊이 정보 차등 적용 및 보정

입체 영상을 생성할 때 객체가 동일한 깊이 정보를 갖지 않으면 양안시차에 맞지 않는 입체 영상이 생성되어 시청자의 눈의 피로감을 가중시킬 수 있다. 따라서 한 객체는 동일한 깊이 정보를 갖는 것을 목적으로 깊이 정보의 할당 및 보정 과정을 거치게 된다.

할당 및 보정 과정은 총 세 단계로 이루어진다. 첫 번째 단계에서는 움직임 정보가 존재하는 그룹의 각 객체별로 깊이 정보의 평균값을 구하여 객체별 깊이 정보의 보정을 하게 된다.

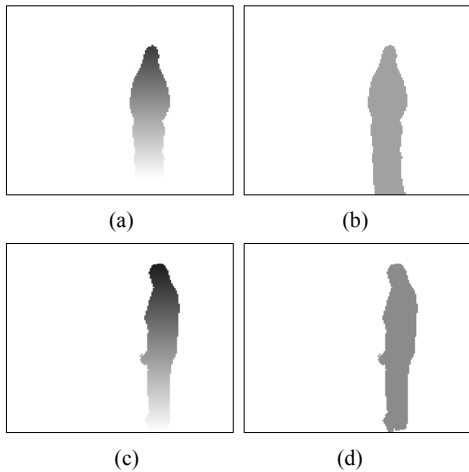
$$D_{avg} = (Pv_{total} - Pv_{back})/P_{obj} \quad (7)$$

식 (7)의 D_{avg} 는 깊이 정보의 평균값을 의미하고, Pv_{total} 과 Pv_{back} 은 각각 영상의 모든 픽셀의 합과 배경의 픽셀의 합을 의미하며 P_{obj} 는 객체의 픽셀수이다.



[그림 7] 움직임 정보를 이용한 깊이 정보 보정
[Fig. 7] Depth information correction using motion information

두 번째 단계는 움직임 정보가 존재하지 않는 그룹의 각 객체별로 상대적 높이 단서를 이용하여 생성한 깊이 정보를 할당하고, 각 객체별로 깊이 정보의 평균값을 구하여 객체별 깊이 정보의 보정을 하게 된다. 만약 그룹화 과정에서 모든 객체가 움직임 정보를 갖고 있다고 판단 되면 두 번째 단계를 거치지 않고 세 번째 단계로 넘어가게 된다.

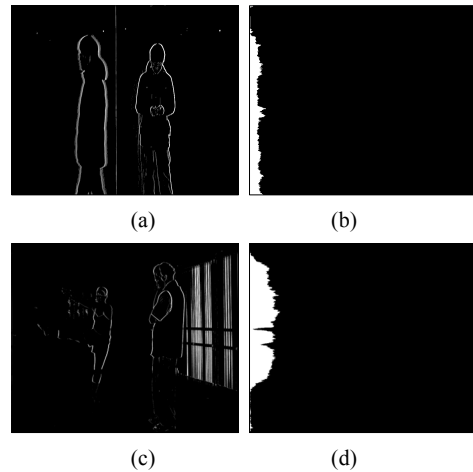


[그림 8] 상대적 높이 단서를 이용한 깊이 정보 할당(a,c), 상대적 높이 단서를 이용한 깊이 정보 보정(b,d)
 [Fig. 8] Depth information allocation using relative height cue(a,c), Depth information correction using relative height cue(b,d)

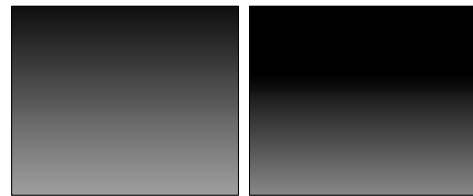
마지막으로 세 번째 단계에서 배경에 상대적 높이 단서를 이용한 깊이 정보를 할당함으로써 최종 깊이맵을 생성하게 된다. 이 때 배경이 객체 뒤에 있다고 인식될 수 있도록 깊이 정보의 추가 작업을 거치게 된다. 객체 중 깊이 정보의 최소값을 구하고, 그 값을 시작으로 상대적 높이 단서를 이용한 깊이 정보를 생성하게 된다. 이 때 깊이 정보의 간격은 영상의 수직 에지 정보를 구해, 행간 에지 분포 차이 $Edge_j - Edge_{j-1}$ 를 특정 임계값 $T3$ 과 비교하여 임계값을 넘는 경우와 넘지 않는 경우 차등적으로 적용한다.

$$D_{back_{j-1}} = \begin{cases} D_{back_j} - (D_{min}/256), & \text{if } |Edge_j - Edge_{j-1}| < T3 \\ D_{back_j} - (D_{min}/128), & \text{otherwise} \end{cases} \quad (8)$$

식 (8)의 j는 영상의 높이, $D_{back_{j-1}}$ 는 j-1행의 깊이 정보이다.



[그림 9] 추출된 수직 에지(a,c), 에지 분포(b,d)
 [Fig. 9] Extracted horizontal edge(a,c), Edge distribution (b,d)



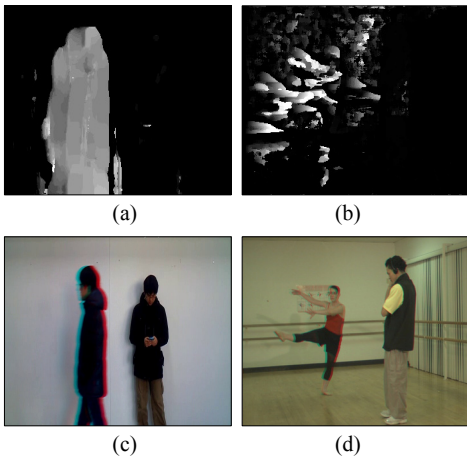
[그림 10] 상대적 높이 단서를 이용한 깊이 정보(배경)
 [Fig. 10] Depth information using relative height cue (background)

5. 실험 및 고찰

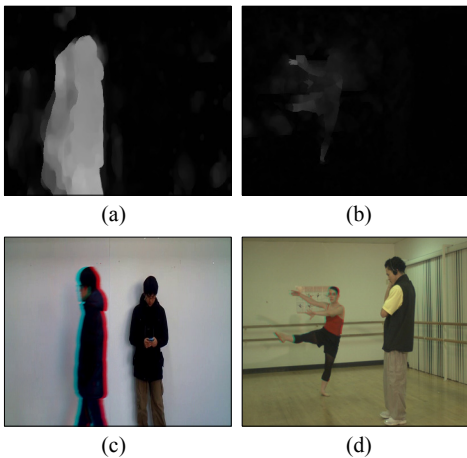
원본 영상과 블록 매칭, Optical flow, 제안하는 기법으로 생성된 각각의 깊이맵을 이용하여 원본 영상의 좌, 또는 우영상을 생성하는 DIBR(Depth Image Based Rendering)을 통해 양안시차에 해당하는 영상을 생성하고[11], 최종 입체 영상은 눈으로 쉽게 식별할 수 있도록 적청 영상의 형태로 생성하였다.



[그림 11] DIBR 구조
 [Fig. 11] DIBR Architecture

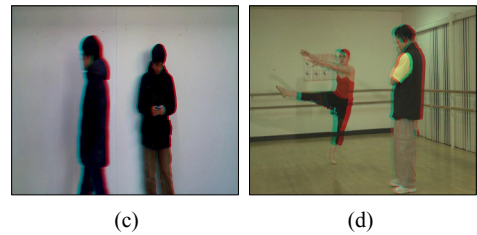
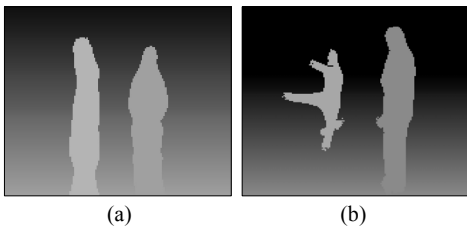


[그림 12] 블록 매칭을 이용한 깊이맵(a, b) 및 해당 입체 영상(c, d)
 [Fig. 12] Depth map using block matching and corresponding stereopsis



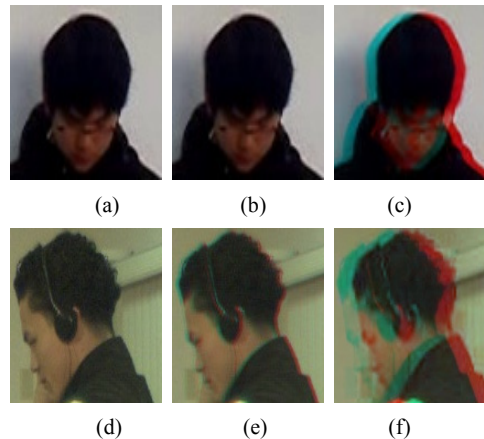
[그림 13] Optical flow를 이용한 깊이맵(a, b) 및 해당 입체 영상(c, d)
 [Fig. 13] Depth map using optical flow and corresponding stereopsis

기존 기법을 이용한 그림 12, 그림 13의 입체 영상에서 움직임이 없는 객체에서 해당 영역의 입체효과를 얻을 수 없는 문제점을 확인할 수 있다.



[그림 14] 제안하는 기법을 이용한 깊이맵(a, b) 및 해당 입체 영상(c, d)
 [Fig. 14] Depth map using proposed method and corresponding stereopsis

제안한 기법을 이용하여 움직임이 없는 객체의 입체효과를 그림 14의 입체 영상을 통해 확인하였고, 따라서 기존 기법에 대한 문제점을 해결할 수 있다.



[그림 15] 움직임 정보를 갖지 않는 객체의 입체 영상, 블록 매칭(a, d), Optical flow(b, e), 제안하는 기법(c, f)
 [Fig. 15] Stereopsis of objects which have no moving information, block matching(a, d), Optical flow(b, e), proposed method(c, f)

비교실험은 기준영상을 1로 잡고, 오차가 클수록 0에 가까운 수치를 갖게 되는 형태의 영상 품질 측정 방법인 SSIM(Structural SIMilarity index)을 이용하여 진행하였다.

[표 1] SSIM 결과
 [Table 1] Result of SSIM

영상	SSIM		
	블록 매칭	Optical flow	제안하는 기법
Men	0.692	0.726	0.869
Ballet	0.685	0.723	0.813

제안한 기법이 기존 기법에 비해 SSIM수치가 높아 보다 오차가 없음을 알 수 있다.

6. 결론

본 논문에서는 2D 동영상을 3D 동영상으로 변환하기 위해 객체를 추출해 그룹화 과정을 거치고, 움직임 정보를 이용한 깊이 정보와 상대적 높이 단서를 이용한 깊이 정보를 획득하여, 객체별 깊이 정보를 할당 및 보정함으로써 최종 깊이맵을 생성하는 기법을 제안하였다. 제안하는 기법을 통해 3D 영상을 생성한 결과 움직임 정보가 없는 객체에서도 깊이 정보를 획득할 수 있어 모든 객체에서 입체 효과를 구현할 수 있었으며 실험을 통해 제안한 기법이 기존 기법에 비해 오차가 적은 입체영상을 생성할 수 있음을 확인하였다.

향후 연구에서는 보다 정확한 객체 추출 및 움직임 정보의 획득 기법과 객체별 유기적인 깊이 정보를 획득할 수 있도록 추가적인 깊이 단서 및 깊이 정보를 이용하는 연구가 필요하다.

References

- [1] L. Zhang, C. Vázquez, and S. Knorr, "3D-TV Content Creation: Automatic 2D-to-3D Video Conversion," (invited paper) IEEE Trans. on Broadcasting, Special Issue on 3D-TV, 2011.
- [2] Hyeon Ho Han, Gang Seong Lee, Sang Hun Lee, "A Study on 2D/3D image Conversion Method using Create Depth Map", Journal of The Korea Academia-Industrial cooperation Society, Vol 12, no. 4 pp. 1897-1903, 2011.
- [3] T. Okino, H. Murata, "New Television with 2D/3D image conversion technologies", SPIE Vol. 2653, pp. 96-103, 1996
- [4] Baxter J. Garcia "Approaches to stereoscopic video based on spatio-temporal interpolation," SPIE Vol. 2653, pp.85-95, 1990
- [5] C. Fehn, "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV", Proceeding of the SPIE Stereoscopic Displays and Virtual Reality Systems XI, Vol. 5291, pp. 93-104, 2004.
- [6] Deqing Sun, Roth, S., Black, M. J., "Secret of Optical Flow Estimation and Their Principles", IEEE Conference on Computer Vision and Pattern Recognition (CVPR),

pp. 2432-2439, 2010.

- [7] L. Zhang, B. Lawrence, D. Wang, and A. Vincent, "Comparison study on feature matching and block matching for automatic 2D to 3D video conversion," in The Second IEE European Conference on Visual Media Production (CVMP 2005), London, UK, pp. 122 - 129, 2005,
- [8] Sudong Kim, Jisang Yoo, "3D conversion of 2D video using depth layer partition", Journal of Broadcast Engineering, Vol. 16 No. 1 pp. 44-53, 2011
- [9] Jianbo Shi and Jitendra Malik, "Normalized Cuts and Image Segmentation", IEEE transactions on pattern analysis and machine intelligence, Vol. 22, pp. 888-905, 2000.
- [10] Sung Ho Han, Yeong Pyo Hong, Jong Youn Lee, Sang Hun Lee, "2D/3D Video Conversion Method using Differential Depth Information by Grouped Object", Proceedings of the KAIS Fall conference, Vol 12, no 2 pp. 445-448, 2011.
- [11] L. Zhang, W. J. Tam, "Stereoscopic image generation based on depth images for 3D TV", IEEE Trans. on Broadcasting, Vol. 51, no. 2 pp. 191-199, 2005.

한 성 호(Sung-Ho Han)

[준회원]



- 2011년 2월 : 광운대학교 컴퓨터 공학과 졸업(공학사)
- 2011년 3월 ~ 현재 : 광운대학교 대학원 재학(석사과정)

<관심분야>

Computer Vision, 3D영상처리, 2D/3D 변환

홍 영 표(Yeong-Pyo Hong)

[정회원]



- 1981년 2월 : 광운대학교 전자계산학과 졸업(이학사)
- 1985년 2월 : 광운대학교 대학원 졸업(공학석사)
- 2003년 2월 : 국립경상대학교 대학원 컴퓨터공학과 졸업(공학박사)
- 2009년 2월 : 가톨릭대학교 보건대학원 졸업(보건학 석사)
- 1985년 3월 ~ 2003년 2월 : 진주전문대학 부교수
- 2003년 3월 ~ 2008년 2월 : 진주국제대학교 교수
- 2008년 3월 ~ 현재 : 한국국제대학교 교수

<관심분야>

의료정보학, 기초의학, 보건학, 건강증진학

이 상 훈(Sang-Hun Lee)

[종신회원]



- 1983년 2월 : 광운대학교 응용전자 공학과 졸업(공학사)
- 1987년 8월 : 광운대학교 대학원 전자공학과 졸업(공학석사)
- 1992년 2월 : 광운대학교 대학원 전자공학과 졸업(공학박사)
- 1990년 9월 ~ 현재 : 광운대학교 정교수
- 2010년 3월 ~ 2012년 1월 : 광운대학교 교양학부장
- 2012년 2월 ~ 현재 : 광운대학교 정보통신처장

<관심분야>

무선인터넷, 무선네트워크, USN, 영상인식, 3D영상처리