

# 그룹화된 객체별 깊이정보를 차등 적용한 2D/3D 동영상 변환 기법

한성호\*, 홍영표\*\*, 이종용\*, 이상훈\*

\*광운대학교,

\*\*한국국제대학교,

e-mail:han@kw.ac.kr

## 2D/3D Video Conversion Method using Differential Depth Information by Grouped Object

Sung-Ho Han\*, Yeong-Pyo Hong\*\*, Jong-Yong Lee\*, Sang-Hun Lee\*

\*Kwangwoon University,

\*\*International University of Korea

### 요 약

본 논문은 그룹화된 객체별 깊이정보를 차등 적용한 2D/3D 동영상 변환 기법에 관한 연구이다. 기존의 연구에서는 프레임의 움직임 정보와 기하학적 깊이 단서를 이용해 깊이정보를 추출하여 2D/3D 동영상 변환을 한다. 그러나 영상의 움직임 정보를 획득할 수 없는 영역의 경우 정확한 깊이정보를 얻을 수 없어 해당 영역의 3D 효과를 얻을 수 없는 문제점이 있다. 제안하는 기법에서는 객체 및 배경을 추출하고 움직임 정보와 기하학적 단서를 이용한 깊이정보를 그룹화된 객체별 차등 적용하여 움직임 정보가 없는 영역에서도 3D 효과를 얻을 수 있는 방법을 제안한다. 최종적으로 원본 영상과 생성된 깊이맵을 DIBR(Depth Image Based Rendering) 과정을 통해 3D 동영상을 생성한 결과 움직임 정보가 없는 영역에서도 3D 효과를 얻을 수 있었다.

### 1. 서론

3D 디스플레이에 표시되는 입체 영상은 기존의 2D 디스플레이보다 시청자에게 화려한 시각적 경험을 얻게 하고 실제감을 고조시킬 수 있으며 방송, 영화, 게임 등 많은 분야에서 유용하게 사용될 수 있다. 또한 입체 디스플레이 기술의 빠른 발전으로 최근 모니터, 핸드폰, TV 등 많은 디스플레이 장치에서 입체 영상을 볼 수 있도록 제작 되고 있다. 하지만 3D 콘텐츠가 부족하여 사용자의 요구를 충족시킬 수 없다[1].

3D 콘텐츠를 생성하는 방법 중 깊이 카메라 및 입체 카메라를 사용하는 경우 깊이정보를 갖는 동영상을 획득하여 쉽게 3D 동영상을 제작할 수 있지만 고가인 문제점을 갖으며 이미지 편집 툴을 이용한 경우에는 많은 시간이 소요되는 문제점을 갖는다. 반면에 2D/3D 변환 방법은 기존의 2D 콘텐츠를 이용하여 3D 콘텐츠를 생성하고, 제작시간이 짧아 이

를 활용하면 3D 콘텐츠 부족 문제를 해결할 수 있다[2].

사람이 입체 영상을 보는 원리는 좌, 우안에 맺히는 영상이 서로 차이가 나는 양안시차가 존재하기 때문이다. 이러한 양안시차를 생성하기 위해 깊이정보를 이용한 방법들이 많이 제안되었다[3].

기존의 2D 영상은 단일 렌즈 카메라로 촬영돼 깊이정보를 갖고 있지 않아 깊이정보를 추출하는 작업이 필요하다.

깊이정보를 획득하는 방법에는 Optical flow를 이용한 움직임 추정 방법, 블록 기반 또는 화소 단위 움직임 추정 방법, 소실점 및 소실선을 이용한 기하학적 방법, 움직임 추정 방식과 기하학적 방식을 같이 사용하는 방법 등이 있다[4].

이 중 움직임 추정 방법은 움직임 정보를 추출하는 방법에 따라 정확한 결과를 얻을 수 있지만 수행시간이 오래 걸리게 된다. 또한 움직임 정보가 없는 영역에서는 정확한 움직임 정보를 얻을 수 없는 단점이 있다. 반면에 기하학적 방법의 경우 수행시간이 짧지만 정확한 깊이정보를 얻을 수 없는 단점이 있다.

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

본 논문에서는 객체 및 배경을 추출하고 화소 단위의 움직임 정보와 기하학적 단서를 이용하여 깊이 정보를 획득한 뒤 그룹화된 객체별 차등 적용 및 보정하여 최종 깊이맵을 생성함으로써 움직임 정보가 없는 영역에서도 3D 효과를 얻을 수 있는 2D/3D 동영상 변환 방법을 제안한다.

## 2. 장면 분석

### 2.1 장면 전환점 검출

순차적으로 프레임별 움직임 정보를 획득하면 장면 전환점이 발생하는 구간에서 정확한 움직임 정보를 획득할 수 없게 된다. 따라서 이러한 문제점을 해결하기 위해 장면 전환점의 검출 방법으로 영상의 에지 정보를 효과적으로 표현하는 MPEG-7의 에지 히스토그램 서술자를 이용한다[5].

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

식 (1)을 이용하여 히스토그램 분포의 차이  $|H_i(j) - H_{i+1}(j)|$  가 허용오차 T 이상이라면 장면 전환점이라고 판단한다.

### 2.2 객체 추출 및 배경 분리

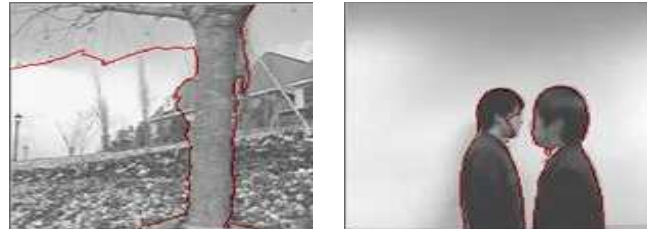
Normalized cut 알고리즘을 이용하여 객체 추출을 한다. 이것은 그래프 분할 알고리즘으로써 분할된 그래프와 영상을 대응시킨다[6].

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

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

식 (2)를 이용하여  $Ncut(A, B)$ 가 가장 작은 값을 갖는  $A, B$ 를 찾아 분할하여 영상 내의 객체를 추출할 수 있다.

식 (3)은 정확한 객체 추출을 위한 가중치에 해당하고 배경 분리는 획득한 객체 부분을 원영상에서 제거함으로써 얻을 수 있다.



(a) 객체 추출



(b) 배경 분리

[그림 1] 객체 추출 및 배경 분리

## 3. 깊이맵 생성

### 3.1 움직임 정보를 이용한 깊이맵 생성

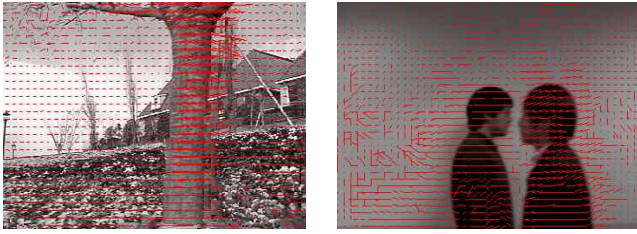
움직임 정보는 식 (4)의 SAD(sum of absolute difference)값을 이용한다. 7x7 크기를 갖는 정합창의 화소값을 비교하여 획득되며 장면 전환점의 전영상과, 영상의 마지막 프레임의 경우 역탐색을 통해 움직임 정보를 획득한다. i번째 프레임과 (i+1)번째 프레임에서 움직임 정보의 픽셀수가 전체 픽셀수의 10% 이하로 존재하는 경우 움직임 정보가 없다고 판단하고, i번째 영상과 (i+2)번째 영상과 연산을 수행하며, 이와 같은 방식으로 다음 프레임과 연산이 진행되고 역탐색의 경우에도 이와 동일한 방식으로 진행된다. 또한 카메라의 움직임이 발생 하는 경우, 영상의 모든 움직임 값과 방향의 평균값을 구하고 이전에 얻은 움직임 값에서 상쇄함으로써 객체에 대한 움직임 정보를 획득할 수 있다.

$$C_{SAD}(x, y) = \sum_{i=0}^6 \sum_{j=0}^6 |Y_n(x+i, y+j) - Y_{n+1}(d_x+i, d_y+j)| \quad (4)$$

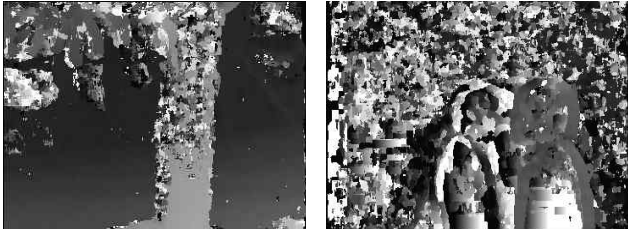
획득한 움직임 정보는 식 (5)을 통해 깊이 정보로 변환된다.

$$D(x, y) = Motion_x + Motion_y \quad (5)$$

$D(x, y)$ 는  $(x, y)$ 좌표에서의 깊이값을 나타내고,  $Motion_x$ 와  $Motion_y$ 는 획득한 움직임 정보의 x와 y 방향으로의 움직임 크기를 나타낸다.



(a) 움직임 정보



(b) 깊이맵

[그림 2] 움직임 정보와 깊이맵

### 3.2 기하학적 단서를 이용한 깊이맵 생성

허프 변환을 통해 라인을 찾은 이후 식 (6)을 이용해 소실점을 찾는다. 소실점이 발견되면 가우시안 분포를 이용하여 깊이맵을 생성하고, 소실점이 발견되지 않은 경우 상대적 높이 단서를 이용한 깊이맵을 생성하게 된다[7].

$$\min_{x_0, y_0} \sum_{i=1}^N W_i (\rho_i - x_0 \cos \theta_i - y_0 \sin \theta_i)^2 \quad (6)$$



(a) 원본 영상

(b) 소실점

[그림 3] 원본 영상 및 소실점



[그림 4] 상대적 높이를 이용한 깊이맵

## 4. 그룹화된 객체별 깊이정보의 차등 적용 기법

### 4.1 그룹화

움직임 정보가 있는 객체와 없는 객체는 서로 다른 깊이정보를 사용한다. 두 객체 중 한쪽의 객체에 깊이정보가 할당되고, 다른 쪽의 깊이정보를 할당할 때 이미 깊이정보가 할당되어 있는 객체에는 깊이정보가 새롭게 할당되는 일이 발생해서는 안 된다.

따라서, 움직임 정보가 있는 객체와 움직임 정보가 없는 객체는 그룹화를 통해 구별되어야 한다.

### 4.2 깊이정보의 차등 적용 깊이맵 생성 및 보정

한 객체에 동일한 깊이정보가 할당되지 않는 경우 정확한 3D 효과가 생성되지 않는다. 따라서 한 객체에 동일한 깊이정보가 할당되어야 하고, 획득한 두 깊이정보는 유기적으로 구성되어야 한다.

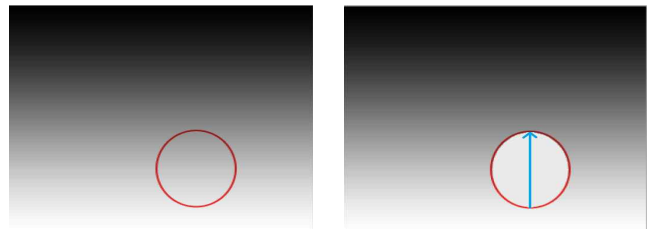
최종 깊이맵의 생성 및 보정은 총 세 단계로 이루어진다. 첫 번째는 추출된 객체에 움직임 정보를 이용한 깊이정보를 할당하고 [그림 5]와 같이 객체 내 깊이정보의 평균값을 할당함으로써 동일한 깊이정보를 갖도록 한다. 두 번째는 첫 번째 단계에서 깊이정보를 할당하지 않은 객체에 기하학적 단서를 이용한 깊이정보를 할당하고 [그림 6]과 같이 객체의 최하단 깊이정보를 할당함으로써 동일한 깊이정보를 갖도록 한다. 세 번째는 보정된 객체들을 상대적 높이 단서를 이용하여 선후관계를 분석하고, 깊이정보가 유기적으로 구성될 수 있도록 한다. 객체의 보정이 끝나고 배경에 기하학적 깊이정보를 할당하여 최종 깊이맵을 생성한다.



(a) 보정 전

(b) 보정 후

[그림 5] 움직임 정보의 깊이정보 보정



(a) 보정 전

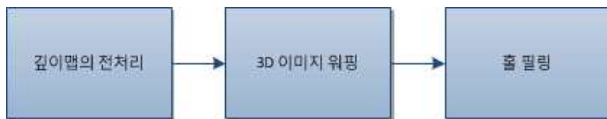
(b) 보정 후

[그림 6] 기하학적 깊이정보 보정

5. 결과 및 고찰

참고문헌

[그림 7]은 기본적인 Depth Image Based Rendering (DIBR) 방법의 구조를 보여준다. 원본 영상의 좌영상을 기준으로, 원본 영상과 해당 깊이맵을 DIBR 방법을 이용하여 원본 영상의 우영상을 생성하게 된다.



[그림 7] DIBR



(a) 원본 영상



(b) 최종 영상

[그림 8] 상대적 높이를 적용한 최종 입체 영상

6. 결론

본 논문에서는 그룹화된 객체별 깊이정보를 차등 적용 및 보정한 2D/3D 동영상 변환 기법을 제안하였다. 장면 전환점에서 발생할 수 있는 움직임 정보의 오류를 줄이기 위한 장면 전환점 검출 과정과 그룹화를 위한 객체정보를 획득하기 위해 객체 및 배경 추출 과정을 거치고, 화소 단위의 움직임 정보와 기하학적 단서를 이용한 깊이정보를 획득하여, 그룹화된 객체별 차등 적용 및 보정함으로써 최종 깊이맵을 생성하였다. 원본 영상과 생성된 깊이맵을 DIBR 방법을 이용하여 최종 3D 영상을 생성하여 확인한 결과 움직임 정보가 없는 영역의 경우에도 3D 효과가 생성되는 것을 확인할 수 있었다.

[1] Liang Zhang, Carlos Vázquez, Sebastian Knorr, “3D-TV Content Creation: Automatic 2D-to-3D Video Conversion”, Broadcasting, IEEE Transactions on, VOL. 57, NO. 2, JUNE 2011

[2] Jae-Il Jung, Yo-Sung Ho, “Depth map estimation from single-view image using object classification based on Bayesian learning”, 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), pp. 1-4, June 2010

[3] 한현호, 홍영표, 김진수, 이상훈, “연산량을 감소한 객체분할과 깊이정보 생성을 이용한 2D/3D 동영상 변환 연구”, 한국산학기술학회 추계 학술발표논문집, Vol. 11, No. 2, pp. 92-95, 2010

[4] Li Sisi, Wang Fei, Liu Wei, “The Overview of 2D to 3D Conversion System”, Computer-Aided Industrial Design & Conceptual Design (CAIDCD), 2010 IEEE 11th International Conference on, pp. 1388-1392, Nov 2010

[5] Chee Sun Won, Dong Kwon Park, Soo Jun Park, “Efficient Use of MPEG-7 Edge Histogram Descriptor”, ETRI Journal, vol. 24, no. 1, pp. 23-30, Feb 2002

[6] 한현호, 이강성, 이상훈, “2D/3D 변환을 위한 깊이 정보 생성 기법에 관한 연구”, 한국산학기술학회 논문지, Vol. 12, No. 4 pp. 1897-1903, 2011

[7] K. Han and K. Hong, “Geometric and texture cue based depth-map estimation for 2D to 3D image conversion,” Consumer Electronics (ICCE), 2011 IEEE International Conference on, pp. 651 - 652, Jan 2011