
실안개와 상대적 높이 단서 기반의 깊이 지도를 이용한 2D/3D 변환 기법

한성호*, 김요섭*, 이종용**, 이상훈***

2D/3D conversion method using depth map based on haze and relative height cue

Sung-Ho Han*, Yo-Sup Kim*, Jong-Yong Lee**, Sang-Hun Lee***

요약 본 논문은 단일영상의 실안개 정보와 상대적 높이 단서를 기반으로 깊이 지도를 생성하고, 이를 이용하여 2D/3D 변환을 하는 기법에 관한 연구이다. 기존의 실안개 정보만을 깊이 지도로써 이용하는 경우, 안개가 없는 영상에서 오류가 발생할 수 있다. 본 논문에서는 이러한 오류를 줄이기 위해, 상대적 높이 단서 기반의 깊이 지도를 생성하고, 실안개 정보와 결합하는 방법을 제안하였다. 또한 Mean Shift Segmentation을 이용한 gray scale 영상과 실안개 정보의 깊이 지도를 결합하여 객체의 경계를 선명화함으로써 3D 영상의 품질을 향상시킬 수 있도록 하였다. 입력영상과 최종 깊이 지도를 DIBR(Depth Image Based Rendering)을 통해 좌영상과 우영상의 시점영상을 생성하고, 적청영상의 형태로 결합함으로써 3D 영상을 생성하였고, 깊이 지도간의 PSNR을 측정하여 검증하였다.

주제어 : 실안개, 상대적 높이, Mean Shift Segmentation, 2D/3D 변환, 깊이 지도

Abstract This paper presents the 2D/3D conversion technique using depth map which is generated based on the haze and relative height cue. In cases that only the conventional haze information is used, errors in image without haze could be generated. To reduce this kind of errors, a new approach is proposed combining the haze information with depth map which is constructed based on the relative height cue. Also the gray scale image from Mean Shift Segmentation is combined with depth map of haze information to sharpen the object's contour lines, upgrading the quality of 3D image. Left and right view images are generated by DIBR(Depth Image Based Rendering) using input image and final depth map. The left and right images are used to generate red-cyan 3D image and the result is verified by measuring PSNR between the depth maps.

Key Words : haze, relative height, Mean Shift Segmentation, 2D/3D conversion, depth map

1. 서론

사람들의 3D 영상에 대한 관심 증가와 동시에 모니터, 핸드폰, TV 등 다양한 디스플레이에서 3D 영상을 지원하는 기기들이 등장하고 있다. 하지만 아직까지 3D 영상에 대한 콘텐츠의 양과 품질 측면에서 부족한 점이 많기 때문에 3D산업의 활성화가 더디게 진행되는 문제점이

있다.

3D 콘텐츠 생성은 스테레오 카메라, 컴퓨터 그래픽 툴, 2D/3D 변환 등을 통해 생성될 수 있지만, 스테레오 카메라와 그래픽 툴을 이용하는 경우, 실제로 촬영이 불가능한 영상의 3D 영상 생성 불가능, 시간과 비용의 많은 소비 발생과 같은 문제 등이 발생하게 된다. 반면에 2D/3D 변환의 경우, 기존 콘텐츠를 활용할 수 있고, 시간

본 논문은 2010년도 광운대학교 교내 학술 연구비 지원에 의해 연구되었음.

*광운대학교 대학원 Kwangwoon University Graduate School

**광운대학교 교양학부 Kwangwoon University, Department of General Education

*교신저자 : 이상훈(leesh58@kw.ac.kr)

논문접수: 2012년 9월 17일, 1차 수정을 거쳐, 심사완료: 2012년 10월 17일

과 비용이 적게 들기 때문에 최근 대다수의 3D 콘텐츠 생성에 관한 연구는 2D/3D 변환을 중심으로 다루어지고 있다. 아직까지는 수작업과 반자동 방식으로 생성되는 3D 영상과 품질을 비교했을 경우 많이 부족하지만, 3D 산업의 활성화를 촉진시키기 위해서 3D 콘텐츠 양의 확보는 중요한 문제이며 해결책으로 자동 2D/3D 변환이 손꼽히고 있다[3][5].

2D/3D 변환은 입체감을 일으키는 가장 중요한 요인인 양안시차를 대부분 이용한다. 이것은 사람의 양 눈의 간격이 약 6.5cm 정도 떨어져있기 때문에 발생하는 것으로써 이로 인해 좌안과 우안은 서로 다른 영상을 보게 되고, 두 영상이 뇌에서 융합됨으로써 입체감을 느끼게 된다. 이러한 양안시차를 바탕으로 좌안과 우안에 해당하는 두 영상을 생성 및 결합함으로써 3D 영상이 생성된다. 따라서, 2D/3D 변환에 관한 연구는 얼마나 정확하게 좌안과 우안에 해당하는 영상을 생성할 수 있는지에 초점이 맞춰진다. 이러한 생성 방법으로 Okino 그룹의 MTD (Modified Time Difference)를 이용한 기법, Garcia의 공간 시간 보간(spatial-temporal interpolation)을 이용한 기법, 영상의 깊이 정보를 이용한 2D/3D 변환 기법 등이 제안되었다. 다양한 기법이 제안됐지만, 원영상과 해당 픽셀의 깊이 정보를 갖는 깊이 지도를 통해 가상의 시점 영상을 생성하는 DIBR(Depth Image Based Rendering) 기법이 제안 및 연구되면서 최근의 변환 연구는 깊이 정보를 이용한 기법에 초점이 맞춰져있다. DIBR을 이용하는 변환 기법의 경우, 깊이 지도가 영상 내의 사물들 사이에 대한 거리 차이를 얼마나 정확하게 나타냈는지에 따라 품질이 결정된다. 그러므로 2D/3D 변환 연구의 핵심은 깊이 지도의 생성에 있으며, 세부적으로는 단일영상 및 다중영상의 다양한 깊이 단서들을 이용한 깊이 정보 획득 및 깊이 지도 생성에 있다[10].

단일영상에서 획득할 수 있는 깊이 단서는 단안단서라고 불리며, 결 기울기, 중첩, 상대적 높이 및 크기, 채도, 디포커스, 선형 원근감, 기하학적 구조 등이 있다. 이러한 단서들을 이용하여 깊이 정보를 획득하고 깊이 지도를 생성하는 많은 연구가 이루어졌다. 하지만 하나의 단서만 이용하여 깊이 지도를 생성하는 경우 깊이 지각의 어려움을 느낄 수 있기 때문에 최근의 연구는 다양한 단서들을 결합하는 형태로 이루어지고 있다[4].

본 논문에서는 단일영상의 단안단서 중 채도 및 기하학적 단서가 될 수 있는 실안개 정보와 상대적 높이 단서

를 이용한 깊이 지도를 생성하고, 객체 경계의 선명화를 위해, Mean Shift Segmentation의 gray scale 영상을 실안개 정보와 결합하고, 최종 깊이 지도를 생성하는 기법을 제안한다. 입력영상과 최종 깊이 지도를 DIBR을 통해 좌영상 및 우영상의 시점영상을 생성하고, 적청영상의 형태로 결합함으로써 3D 영상을 생성 및 검증하였다.

2. 제안하는 깊이 지도 생성 기법

2.1. 실안개 정보 생성

실안개 정보란 영상의 안개가 낀 정도를 의미하며, 안개의 양에 따라 채도변화가 이루어지기 때문에, 단안단서 중 채도 단서로 구분할 수 있으며, 따라서 실안개 정보를 깊이 단서로써 활용한다[1][2].

실안개 정보는 안개 모델링을 나타내는 식 (1)과 Dark Channel Prior를 나타내는 식 (2)를 이용하게 된다.

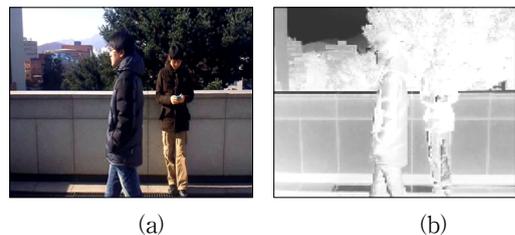
$$I(x) = J(x)t(x) + A(1 - t(x)) \quad (1)$$

$$J^{dark}(x) = \min_{c \in r, g, b} \left(\min_{y \in \Omega(x)} (J^c(y)) \right) \quad (2)$$

식 (1)의 t(x)는 전달량으로써 영상의 안개 값의 비율을 나타내고, 식 (2)는 일정 구간 내 한 픽셀 중 어느 한 채널은 대부분 0의 값을 갖는다는 관측을 나타내는 Dark channel prior를 나타낸다.

$$t(x) = 1 - \min_c \left(\min_{y \in \Omega(x)} \left(\frac{I^c(y)}{A^c} \right) \right) \quad (3)$$

식 (1)과 (2)를 정리하여 식 (3)을 얻게 되고 이것을 실안개 정보로써 사용하게 된다. 본 논문의 Dark channel prior의 일정 구간은 5x5로 설정하였다.



[그림 1] 입력영상(a), 실안개 정보 영상(b)

2.2. 상대적 높이 단서 기반의 깊이 지도 생성

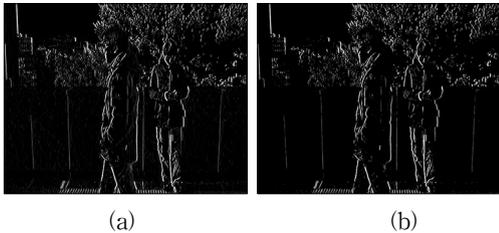
영상의 기하학적 구조로 인해 대부분의 영상의 물체는 영상의 하단부에서 상단부로 갈수록 근거리에서 원거리로 배치된다. 이것을 상대적 높이 단서라고 부르며, 본 논문에서는 이 단서를 이용하여 깊이 지도를 생성한다[7].

우선, 3x3 크기의 Sobel filter를 통해 수직 에지만을 검출한다.



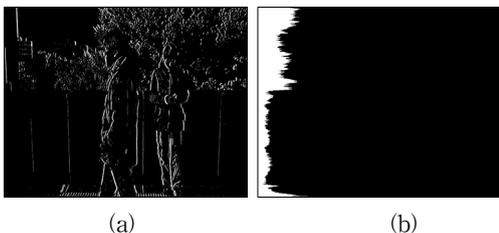
[그림 2] 입력영상(a), 수직 에지 영상(b)

다음, 잡음에 강건하고 신뢰성 있는 에지만을 남기기 위해, 특정 임계값 이하의 값을 갖는 에지를 제거한다. 본 논문에서는 에지 제거를 위한 임계값을 50으로 설정하였다.



[그림 3] 수직 에지 영상(a), 임계값을 통한 수직 에지 제거 영상(b)

이후 행별 에지 개수를 누적시킨 히스토그램을 생성하고, 히스토그램의 최하단부를 시작으로, 바로 윗행과의 에지 개수의 차이를 최상단부까지 반복적으로 검출하여 에지 개수의 행별 평균 차이값 E_{row_avg} 을 구한다.



[그림 4] 임계값을 통한 수직 에지 제거 영상(a), 에지 개수 히스토그램 영상(b)

다시, 앞서 구한 행별 평균 차이값 E_{row_avg} 을 넘는 값을 구하고, 255값을 E_c 값으로 나눔으로써 깊이 지도의 픽셀 간격 D_d 을 구하게 된다.

$$E_c = \sum_{j=height}^0 \begin{cases} 1, & \text{if } |(EH(j) - EH(j-1))| > E_{row_avg} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$D_d = \frac{1}{E_c} \times 255 \quad (5)$$

$$D_{rh}(y) = \begin{cases} 255 - D_d \times n & , \text{if } |(EH(y) - EH(y-1))| > E_{row_avg} \\ 255 - D_d \times (n-1) & , \text{otherwise} \end{cases} \quad (6)$$

$y = height, height-1 \dots 1, n = 1, 2, 3 \dots E_c$

$height$ 는 영상의 높이, EH 는 에지 개수 히스토그램, D_{rh} 는 상대적 높이 단서 기반의 깊이 지도를 의미하고, $D_{rh}(y)$ 는 0~255 사이의 값을 갖는다.

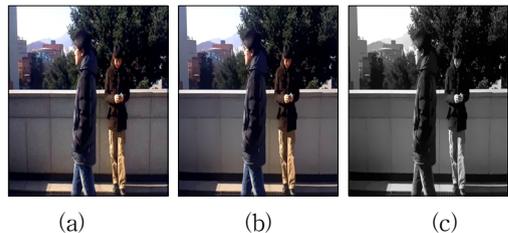


[그림 5] 입력영상(a), 상대적 높이 단서 기반의 깊이 지도(b)

2.3. 객체 경계 선명화

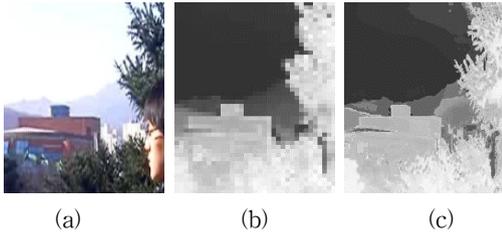
깊이 지도에서 객체와 배경의 경계가 명확하지 않은 경우 3D 영상의 품질이 저하될 수 있다. 따라서, 본 논문에서는 깊이 지도의 객체 경계를 선명화하기 위해, 색상 분할 알고리즘 중 하나인 Mean Shift Segmentation (MSS) 알고리즘과 2.1절에서 획득한 실안개 정보를 이용한다[6][8].

우선 입력영상에 공간 윈도우와 색공간 윈도우의 반경 각각 20의 크기를 갖는 MSS 알고리즘을 적용하고, gray scale 영상으로 변환한다.



[그림 6] 입력영상(a), MSS 영상(b), MSS gray scale 영상(c)

이후, gray scale 영상의 0~255 각 픽셀값에 해당하는 위치와 대응하는 실안개 정보 영상의 픽셀값들의 평균값을 획득한다. 마지막으로, 획득한 각 평균값들을 다시 0~255 각 픽셀값에 대응시켜 할당함으로써 객체 경계가 선명화 된 깊이 지도를 생성하게 된다.



[그림 7] 입력영상(a), 실안개 정보 영상(b), 객체 경계 선명화 영상(c)

2.4. 최종 깊이 지도 생성

2.2절과 2.3 절에서 생성한 두 깊이 지도를 식 (8)을 이용하여 결합한다. 본 논문에서 α 값은 0.7로 설정하였다.

$$D_c(x,y) = (1-\alpha)t(x,y) + \alpha D_{rn}(x,y) \quad (7)$$

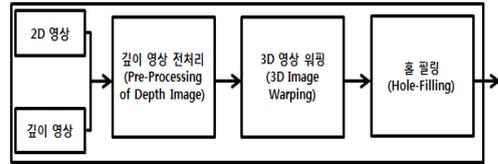


[그림 8] 입력영상(a), 최종 깊이 지도(b)

3. 실험 및 고찰

본 논문에서는 실험을 위해 안개가 포함되지 않은 영상인 Microsoft Research에서 제공하는 “Ballet”, “Breakdancing” 영상의 시퀀스 및 깊이 지도 영상을 사용하였다. 입력영상과 최종 깊이 지도를 Depth Image Based Rendering(DIBR)을 통해 좌영상과 우영상의 시점 영상을 생성하고, 적청영상의 형태로 합성함으로써 3D 영상을 생성하였다[9]. 3D 영상의 품질을 객관적으로 측정하기에는 어려움이 따르기 때문에, 실제 3D 영상의 품질을 결정하는 깊이 지도에 초점을 맞추어 제공된 깊

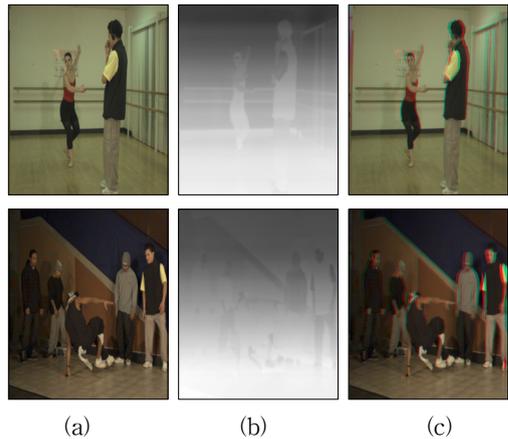
이 지도와 실안개 정보만을 사용한 깊이 지도 및 제안된 깊이 지도의 PSNR(Peak Signal to Noise Ratio) 측정을 통해 제안하는 기법을 검증하였다. PSNR이란 두 영상에 대한 차이를 수치로 나타내는 것으로써 db이 높을수록 비교영상과의 차이가 적다는 것을 알 수 있다.



[그림 9] DIBR 구조

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right)[db] \quad (8)$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [Gd(i,j) - d(i,j)]^2 \quad (9)$$



[그림 10] 입력영상(a), 깊이 지도(b), 3D 영상(적청)(b)

<표 1> PSNR 측정 결과

영상	PSNR(db)	
	실안개 깊이 지도	제안하는 깊이 지도
Ballet	7.87	10.12
Breakdancing	6.36	13.38

4. 결론

본 논문에서는 기존의 실안개 정보만을 깊이 지도로써 사용하는 경우, 안개가 없는 영상에서 오류가 발생할

수 있는 문제점을 줄이기 위해, 단안단서인 상대적 높이 단서를 깊이 정보로써 결합하는 깊이 지도 생성 방법을 제안하였다. 또한 3D 영상 품질 향상을 위해, 객체 경계를 선명화 할 수 있도록 Mean Shift Segmentation 알고리즘을 실안개 깊이 지도와 함께 사용하였다. PSNR 측정 결과에서 기존 기법에 비해 제안하는 기법의 db이 높음으로써 개선됨을 확인하였다.

향후 연구에서는 더 정확한 깊이 지도를 생성할 수 있도록 하는 단안 및 양안단서를 이용한 깊이 정보 획득과 기존의 깊이 지도와의 조합 및 결합 방법의 연구가 필요하다.

참고 문헌

[1] 김진환 · 김창수 (2010). Dark Channel Prior를 이용한 계층적 영상 안개 제거 알고리즘. 대한전기학회논문지, 59(2), 457-464.

[2] 소용석 · 현대영 · 이상욱 (2010). 실안개를 이용한 단일 영상으로부터의 깊이정보 획득 및 뷰 생성 알고리즘. 한국방송공학회 학술발표대회 논문집, 242-244.

[3] 한성호 · 홍영표 · 이상훈 (2012). 2D/3D 동영상 변환을 위한 그룹화된 객체별 깊이 정보의 차등 적용 기법. 한국산학기술학회논문지, 13(3), 1302-1309.

[4] 한종원 · 조진수 · 이일병 (2006). 단안단서를 이용한 자연영상의 상대적 깊이지도 생성. 한국정보과학회 학술발표 논문집, 367-369.

[5] 한현호 · 이강성 · 이상훈 (2012). Optical flow의 레벨 간소화 및 노이즈 제거와 에지 정보를 이용한 2D/3D 변환 기법. 한국산학기술학회논문지, 13(2), 827-833.

[6] Achanta, R., Hemami, S., Estrada, F., & Susstrunk, S. (2009). Frequency-tuned Salient Region Detection. IEEE Conference on Computer Vision and Pattern Recognition, 1597-1604.

[7] Chao-Chung Cheng, Chung-Te Li, & Liang-Gee Chen (2010). An Ultra Low Cost 2D-to-3D Video Conversion System. SID Symposium Digest of Technical Papers, 41(1), 766-769.

[8] ChaoLiu & LaurenChristopher (2012). Depth Map Estimation from motion for 2D to 3D Conversion. IEEE International Conference on Electro/Information Technology, 1-4.

[9] CHRISTOPH FEHN (2003). A 3D-TV Approach Using Depth-Image-Based Rendering (DIBR). Proceedings of the Visualization, Imaging, and Image Processing.

[10] Smolic, A., Kauff, P., Knorr, S., Hornung, A., Kunter, M., Muller, M. & Lang, M. (2011). Three-Dimensional Video Postproduction and Processing. Proceedings of the IEEE, 99(4), 607-625.

한 성 호



- 2011년 2월 : 광운대학교 컴퓨터공학과 졸업(공학사)
- 2011년 ~ 현재 : 광운대학교 대학원 재학(석사과정)
- 관심분야 : Computer Vision, 3D 영상처리, 2D/3D 변환
- E-Mail : han@kw.ac.kr

김 요 섭



- 2009년 : 광운대학교 정보콘텐츠 대학원 유비쿼터스 컴퓨팅공학과 졸업(공학석사)
- 2010년 ~ 현재 : 광운대학교 대학원 재학(박사과정)
- 관심분야 : 영상처리, 정보보안
- E-Mail : ysnewage@gmail.com

이 종 용



- 1983년 2월 : 한양대학교 원자력공학과 졸업(공학사)
- 1988년 2월 : 광운대학교 대학원 전자공학과 졸업(공학석사)
- 1993년 8월 : 광운대학교 대학원 전자공학과 졸업(공학박사)
- 1991년 3월 ~ 2004년 12월 : 광운대학교 정보과학교육원 대우교수
- 2005년 3월 ~ 현재 : 광운대학교 교양학부 부교수
- 관심분야 : 자동제어, 로보틱스, 영상인식
- E-Mail : jyonglee@kw.ac.kr

이 상 훈



- 1983년 2월 : 광운대학교 응용전자공학과 졸업(공학사)
- 1987년 8월 : 광운대학교 대학원 전자공학과 졸업(공학석사)
- 1992년 2월 : 광운대학교 대학원 전자공학과 졸업(공학박사)
- 1990년 9월 ~ 현재 : 광운대학교 정교수
- 2001년 ~ 2007년 : 세계기능경기대회(심사위원)
- 2006년 ~ 2007년 : 서울특별시 기능경기위원회(기술위원장)
- 2010년 3월 ~ 2012년 1월 : 광운대학교 교양학부장
- 2012년 2월 ~ 현재 : 광운대학교 정보통신처장
- 관심분야 : 무선인터넷, 무선네트워크, USN, 영상인식, 3D 영상처리, 수치해석
- E-Mail : leesh58@kw.ac.kr