

전경 깊이 지도와 배경 깊이 지도의 결합을 이용한 깊이 지도 생성

김진현, 백열민, *김희율
한양대학교

jhkim@vision.hanyang.ac.kr, ymbaek@vision.hanyang.ac.kr, *wykim@hanyang.ac.kr

Depth-Map Generation using Fusion of Foreground Depth Map and Background Depth Map

Jin-Hyun Kim, Yeul-Min Baek, *Whoi-Yul Kim
Hanyang University

요 약

본 논문에서 2D-3D 자동 영상 변환을 위하여 2D 상으로부터 깊이 지도(depth map)을 생성하는 방법을 제안한다. 제안하는 방법은 보다 정확한 깊이 지도 생성을 위해 영상의 전경 깊이 지도(foreground depth map)와 배경 깊이 지도(background depth map)를 각각 생성 한 후 결합함으로써 보다 정확한 깊이 지도를 생성한다. 먼저, 전경 깊이 지도를 생성하기 위해서 라플라시안 피라미드(laplacian pyramid)를 이용하여 포커스/디포커스 깊이 지도(focus/defocus depth map)를 생성한다. 그리고 블록정합(block matching)을 통해 획득한 움직임 시차(motion parallax)를 이용하여 움직임 시차 깊이 지도를 생성한다. 포커스/디포커스 깊이 지도는 평탄영역(homogeneous region)에서 깊이 정보를 추출하지 못하고, 움직임 시차 깊이 지도는 움직임 시차가 발생하지 않는 영상에서 깊이 정보를 추출하지 못한다. 이들 깊이 지도를 결합함으로써 각 깊이 지도가 가지는 문제점을 해결하였다. 선형 원근감(linear perspective)와 선 추적(line tracing) 방법을 적용하여 배경 깊이 지도를 생성한다. 이렇게 생성된 전경 깊이 지도와 배경 깊이 지도를 결합하여 보다 정확한 깊이 지도를 생성한다. 실험 결과, 제안하는 방법은 기존의 방법들에 비해 더 정확한 깊이 지도를 생성하는 것을 확인할 수 있었다.

1. 서론

최근 들어 3D-TV 와 3D 모니터 등의 3D 디스플레이 장치들이 발전함으로써 많은 3D 디스플레이 사용자들이 다양한 3D 콘텐츠를 요구한다. 그러나 사용자들의 요구를 충족시킬 충분한 3D 콘텐츠를 보유하고 있지 않다. 따라서 사용자의 요구를 충족시키기 위해서 충분한 3D 콘텐츠 제작이 매우 중요하다.

3D 콘텐츠 제작 방법은 크게 두 가지 방법으로 구분된다. 첫 번째 방법은 깊이 카메라(depth camera) 또는 스테레오 카메라(stereo camera)를 이용하여 3D 콘텐츠를 직접 제작한다. 이 방법은 실제 깊이 지도를 획득할 수 있지만, 별도의 센서(sensor) 또는 카메라가 필요하기 때문에 추가적인 비용이 발생한다. 또 다른 방법은 2D 영상을 3D 영상으로 변환하는 것이다. 이 방법은 수동(manual)으로 변환하는 방법, 반자동적(semi-automatic)으로 변환하는 방법, 그리고 자동적(automatic)으로 변환하는 방법으로 나누어 진다. 먼저, 수동으로 변환하는 방법은 사람이 직접 2D 영상 내의 모든 화소(pixel)를 수평 방향으로 이동시킨다. 이는 양질의 3D 콘텐츠로 얻을 수 있지만, 3D 콘텐츠를 제작하는 시간이 상당히 많이 걸리는 단점이 있다. 두 번째로, 반자동적으로 변환하는 방법은 사용자의 입력을 요구한다[1]. 이 방법은 수동으로 변환하는 방법에 비해서 시간을 줄일 수 있지만,

여전히 사람의 도움을 필요로 하는 단점이 있다. 마지막으로 자동으로 변환하는 방법은 2D 영상 내의 깊이 단서(depth cue)들로부터 추출된 깊이 지도를 기반으로 좌, 우 영상을 생성함으로써 3D 콘텐츠를 제작한다. 이 방법은 기존의 방대한 2D 콘텐츠를 3D 콘텐츠로 변환할 수 있는 적합한 방법이다. 따라서 본 논문에서는 자동적으로 2D 콘텐츠를 3D 콘텐츠로 변환하기 위해 깊이 정보들로부터 깊이 지도를 생성하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2 장에서는 2D 영상을 3D 영상으로 변환하기 위해 깊이 지도를 생성하는 관련 연구에 대해 살펴본 후, 3 장에서 본 논문에서 제안하는 방법을 설명한다. 4 장에서는 실험 결과를 보여주고, 5 장에서는 본 논문에 대한 결론을 기술한다.

2. 관련 연구

2D 영상 내의 깊이 단서들로부터 깊이 지도 생성 방법은 오랜 시간 연구되어 왔다. 깊이 지도 생성 방법은 크게 단안 깊이 단서(monocular depth cue) 방법과 양안 깊이 단서(binocular depth cue) 방법으로 나누어 진다. 본 논문은 자동으로 2D 영상을 3D 영상으로 변환하는 것을 목적으로 하기 때문에 단안 깊이 단서 방법에 대해서만 언급한다. 그림 1 은

단안 깊이 단서의 분류를 보여준다. 이 분류가 모든 단안 깊이 단서를 포함하는 것은 아니다. 단안 깊이 단서는 화상 단서(pictorial cue)와 움직임 단서(motion cue)로 분류된다. 먼저, 포커스/디포커스(focus/defocus) 방법은 가까운 객체가 선명도(sharpness)가 높다는 가정을 기반으로 각 화소 또는 블록(block)의 블러(blur)된 정도를 추정하여 깊이 지도를 생성한다[2]. 선형 원근감 방법은 소실점(vanishing point)이 카메라에서 가장 먼 곳이라는 정보로부터 깊이 지도를 생성한다[3]. 상대적 높이(relative height) 방법은 영상의 아래 부분이 더 가깝다는 가정을 이용하여 깊이 지도를 생성한다[4]. 움직임 시차(motion parallax) 방법은 카메라에 움직임이 발생하였을 때 가까운 물체가 더 많은 시차를 갖는다는 가정을 기반으로 수평과 수직의 움직임 시차 양을 이용하여 깊이 지도를 생성한다[5].

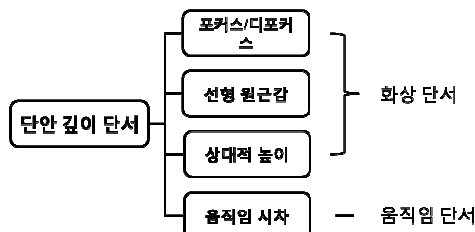


그림 1. 단안 깊이 단서의 분류

3. 제안하는 방법

본 논문에서 제안하는 방법은 전경 깊이 지도와 배경 깊이 지도를 생성하여 결합함으로써 보다 정확한 깊이 지도를 생성한다. 먼저, 전경 깊이 지도를 생성하기 위해서 앞 절에서 설명한 포커스/디포커스, 움직임 시차 방법을 결합함으로써 각 깊이 단서가 가지는 문제점을 보완한다. 배경 깊이 지도를 생성하기 위해서 선형 원근감 깊이 단서를 이용한다. 마지막으로 전경 깊이 지도와 배경 깊이 지도를 결합함으로써 보다 정확한 깊이 지도를 생성한다. 그림 2 는 제안하는 방법의 흐름을 보여준다.

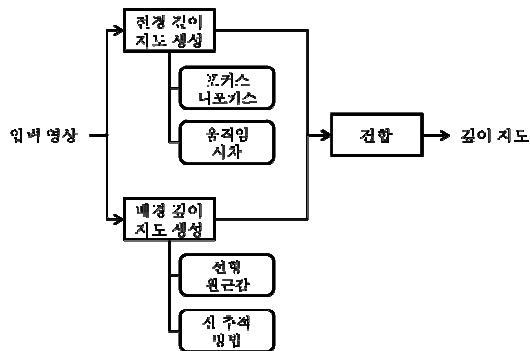


그림 2. 제안하는 방법의 흐름

3.1 전경 깊이 지도 생성

3.1.1 포커스/디포커스 깊이 지도

제안하는 방법에서 각 화소의 선명도를 추정하기 위해서 라플라시안 피라미드(laplacian pyramid)를 이용한다. 그림 3 은 포커스/디포커스 깊이 지도 생성 흐름을 보여준다. 먼저 가우시안 피라미드(Gaussian pyramid)로부터 라플라시안 피라미드를 생성하고, 라플라시안 피라미드 최상위층을 제외한 각 층별 영상을 이진화한 후 팽창연산(dilation operation)을 수행한다. 그리고 각 층별 영상을 병합하고 조인트 바이레터럴 필터(joint bilateral filter) [6]를 이용하여 후처리를 한 후 최종 포커스/디포커스 깊이 지도 $depth_{focus}$ 을 얻는다.

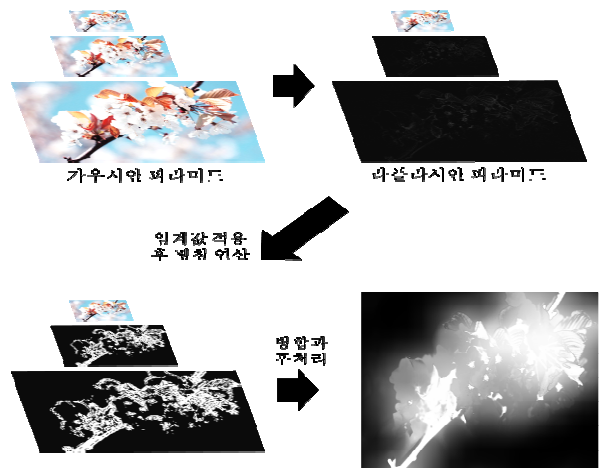


그림 3. 포커스/디포커스 깊이 지도 생성 흐름

3.1.2 움직임 시차 깊이 지도

그림 4 는 본 논문에서 제안하는 움직임 시차를 깊이 지도를 생성하는 흐름을 보여준다. 본 논문에서는 빠르게 움직임 벡터를 얻기 위해서 Cheung 의 블록정합 방법을 사용한다 [7]. Cheung 의 블록정합 방법 통해 획득한 움직임 벡터의 정확도를 높이기 위하여 식(1)을 이용하여 움직임 벡터의 정확도를 개선한다. 식(1)의 MV_{B_i} 는 주변 블록의 움직임 벡터를 나타내고, $E(MV_{B_i})$ 는 주변 블록의 움직임 벡터를 이용하여 구한 블록정합 방법의 오차를 나타낸다. 그림 5 는 주변 블록 B_i 의 위치를 보여주었고, 그림 6 은 움직임 벡터 개선 결과를 보여준다.



그림 4. 움직임 시차 깊이 지도 생성 흐름

$$Index = \arg \min_i |E(MV_{B_i}) - E(MV_{B_5})|$$

$$MV_{B_5} = MV_{B_{Index}}$$
(1)

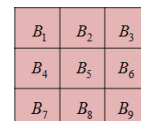


그림 5. 현재 블록과 주변 블록

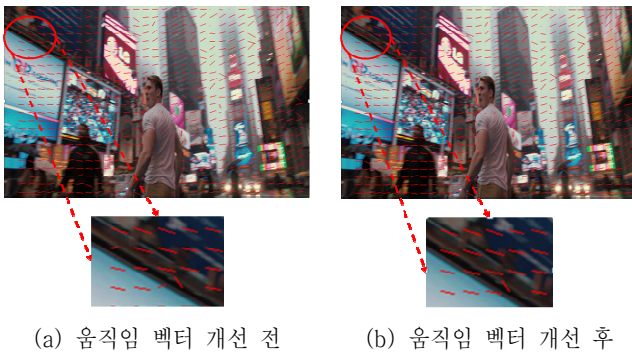


그림 6. 움직임 벡터 개선 결과

또한, 카메라의 움직임으로 발생하는 전역 움직임(global motion)을 보상하여 영상 내 객체만을 분할하는 움직임 시차 깊이 지도를 생성하기 위해, 영상 내의 객체는 영상의 중앙에 위치한다고 가정하고 그림 7 과 같이 영상 외각(음영 처리된 부분)의 움직임 벡터 히스토그램(histogram) [9]을 이용하여 전역 움직임을 추정하고, 식(2)을 이용하여 전역 움직임을 보상한다. 그림 8 은 전역 움직임 보상 후에 결과를 보여준다.

$$\begin{aligned} MV_x &= MV_x - MV_{global_x} \\ MV_y &= MV_y - MV_{global_y} \end{aligned} \quad (2)$$

MV_x , MV_y 는 각 화소의 수평, 수직 방향의 움직임 시차를 나타내고, MV_{global_x} , MV_{global_y} 는 수평, 수직 전역 움직임을 나타낸다.



그림 7. 전역 움직임 시차 추정 방법



그림 8. 전역 움직임 보상 결과

전역 움직임이 보상된 움직임 벡터로부터 각 화소에 식(3)을 이용하여 깊이 값(depth value)을 할당하고,

조인트 바이레터럴 필터[6]를 이용하여 후처리를 하여 최종 움직임 시차 깊이 지도를 생성한다.

$$depth_{motion}(x, y) = \omega \times \sqrt{MV_x^2 + MV_y^2} \quad (3)$$

x, y 는 2D 영상의 좌표를 나타낸다. ω 는 가중치 매개변수(weight parameter)이다.

3.1.3 포커스/디포커스 깊이 지도와 움직임 시차 지도의 결합

본 논문에서는 식(4)를 이용하여 포커스/디포커스 깊이 지도와 움직임 시차 지도를 결합함으로써 전경 깊이 지도를 생성한다.

$$depth_{foreground} = \omega_1 \times depth_{focus} + \omega_2 \times depth_{motion} \quad (4)$$

3.2 배경 깊이 지도 생성

그림 9 는 배경 깊이 지도 생성 방법의 전체 흐름을 보여준다. 먼저 허프 변환(hough transform) [8]을 이용하여 정확한 소실점의 위치를 찾는다. 그리고 나서 소실점의 위치를 중심으로 가우시안 분포(gaussian distribution) 형태로 퍼져나가는 초기 선형 원근감 깊이 지도를 생성한다 [3]. 초기 선형 원근감 깊이 지도를 초기모델로 하는 변형된 Jung 의 방법[4]을 적용하고, 조인트 바이레터럴 필터[6]를 이용하여 후처리를 하여 최종 배경 깊이 지도 $depth_{background}$ 을 생성한다.

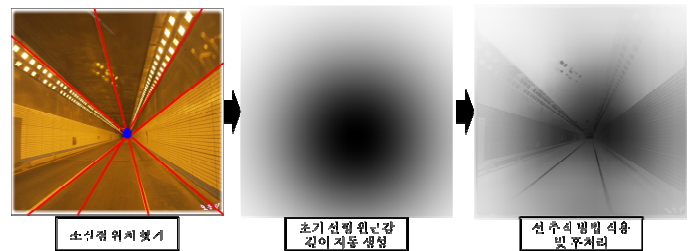


그림 9. 배경 깊이 지도 생성 흐름

3.3 전경 깊이 지도와 배경 깊이 지도의 결합

마지막으로 식 (5)와 같이 앞서 생성한 전경 깊이 지도와 배경 깊이 지도를 결합 함으로써 최종적인 깊이 지도를 생성한다.

$$depth = \omega_3 \times depth_{foreground} + \omega_4 \times depth_{background} \quad (5)$$

4. 실험 결과

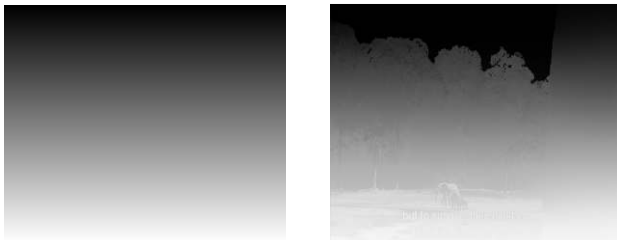
그림 10 은 포커스/디포커스 깊이 지도와 움직임 시차 깊이 지도를 결합함으로써 각 깊이 지도가 가지는 문제점이

보완되어 보다 정확한 전경 깊이 지도가 생성된 것을 보여준다. 그림 11 은 배경 깊이 지도 실험 결과를 보여준다. 그림 12 는 전경 깊이 지도와 배경 깊이 지도의 결합을 통해 생성된 깊이 지도를 보여준다.



(a) (b) (c)

그림 10. 전경 깊이 지도 실험 결과. (a) 포커스/디포커스 깊이 지도 (b) 움직임 시차 깊이 지도 (c) 전경 깊이 지도



(a) (b)

그림 11. 배경 깊이 지도 실험 결과. (a) 초기 선형 원근감 깊이 지도 (b) 선 추적 알고리즘 수행 후 결과



(a) 원본 이미지 (b) 최종 깊이 지도

그림 12. 최종 깊이 지도

5. 결론

본 논문에서는 전경 깊이 지도와 배경 깊이 지도를 결합함으로써 2D-3D 자동 변환을 위한 깊이 지도 생성 방법을 제안한다. 전경 깊이 지도를 생성하기 위해서 포커스/디포커스 깊이 지도와 움직임 시차 깊이 지도를 결합한다. 이들의 결합을 통하여 평탄영역에서 발생하는 포커스/디포커스 깊이 지도의 문제점과 움직임 시차가 없는 영상에서 발생하는 움직임 시차 깊이 지도의 문제점을 서로 보완한다. 배경 깊이 지도를 생성하기 위해서 선형 원근감 깊이 지도와 선 추적 방법을 결합하였다. 실험 결과 포커스/디포커스 깊이 지도와 움직임 시차 깊이 지도를 결합함으로써 보다

정확한 전경 깊이 지도가 생성된 것을 보여준다. 또한, 각각 생성된 전경 깊이 지도와 배경 깊이 지도를 결합함으로써 2D-3D 자동 변환을 위한 보다 정확한 깊이 지도가 생성된 것을 보여준다.

감사의 글

본 연구는 삼성전자와의 산학협력과제의 연구결과로 수행되었습니다. (No. 20110000001711)

참고 문헌

- [1] N. E. Yang and R. H. Park, "Depth map generation from a single image using local depth hypothesis," Proc. of IEEE Int. Conf. Consumer Electronics, pp. 311-312, 2012.
- [2] G. Guo, N. Zhang, L. Huo, and W. Gao, "2D to 3D conversion based on edge defocus and segment," Proc. of IEEE Int. Conf. Acoustics, Speech and Signal Processing, pp. 2181-2184, 2008.
- [3] K. Han and K. Hong, "Geometric and texture cue based depth-map estimation for 2D to 3D image conversion," Proc. of IEEE Int. Conf. Consumer Electronics, pp. 651-652, 2011.
- [4] Y. J. Jung, A. Baik, J. Kim, and D. Park, "A novel 2D-to-3D conversion technique based on relative height depth cue," Proc. of SPIE Conf. Stereoscopic Displays and Applications XX, Vol. 7237, 72371U, 2009.
- [5] L. Ideses, L. P. Yaroslavsky, and B. Fishbain, "Real-time 2D to 3D video conversion," Jour. of Real-Time Image Processing, Vol. 2, No. 1, pp. 3-9, 2007.
- [6] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," Proc. of European Conf. Computer Vision, Vol. 4, pp. 568-580, 2006.
- [7] C. K. Cheung and L. M. Po, "Normalized partial distortion search algorithm for block motion estimation," IEEE Trans. Circuits and Systems for Video Technology, Vol. 10, No. 3, pp. 417-422, 2000.
- [8] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Third edition, NJ: Person Education Inc., Upper Saddle River, 2008.
- [9] M. T. Pourazad, P. Nasiopoulos, and R. K. Ward, "Generating the depth map from the motion information of h.264-encoded 2D video sequence," Jour. of Image and Video Processing, Vol. 2010, No. 4, 2010.