
계위 공간을 이용한 고품질 3차원 비디오 생성 방법

다단계 계위공간 개념을 이용해 깊이맵의 경계영역을 정제하는 고화질 복합형 카메라 시스템과 고품질 3차원 스캐너를 결합하여 고품질 깊이맵을 생성하는 방법

High-quality 3-D Video Generation using Scale Space

이은경, Eun-Kyung Lee*, 정영기, Young-Ki Jung**, 호요성, Yo-Sung Ho***

요약 본 논문은 고화질(high definition, HD) 복합형 카메라 시스템과 고품질(high-quality) 3차원 스캐너를 결합하여 다시점 비디오와 그에 상응하는 다시점 깊이맵을 생성하는 시스템을 제안한다. 복합형 카메라 시스템과 3차원 스캐너를 이용해 3차원 비디오를 생성하기 위해서는, 우선 움직임이 없는 배경영역에 대한 깊이정보를 고품질 3차원 스캐너를 이용해 미리 획득하고, 동적으로 움직이는 전경영역에 대해서는 다시점 카메라와 깊이 카메라를 결합한 복합형 카메라 시스템을 이용해 다시점 비디오와 깊이맵을 획득한다. 그리고 3차원 스캐너와 깊이카메라를 통해 획득한 깊이정보를 이용해 3차원 워핑(warping)을 적용하여 각 다시점 카메라를 위한 초기 깊이정보를 예측한다. 초기 깊이정보를 이용해 다시점 깊이를 예측하는 것은 다시점 카메라의 각 시점에서의 초기 깊이맵을 계산하기 위한 것이다. 고화질의 다시점 깊이맵을 생성하기 위해서 belief propagation 방법을 이용하여 초기 깊이맵을 정제한다. 마지막으로, 전경영역의 경계선 영역의 불규칙적인 깊이맵을 정제하기 위해 전경영역의 외곽선 정보를 추출하여 생성된 깊이맵의 경계선 영역을 다시한번 정제한다. 제안한 3차원 스캐너와 복합형 카메라를 결합한 시스템은 기존의 깊이맵 예측 방법보다 정확한 다시점 깊이맵을 포함하는 3차원 비디오를 생성할 수 있었으며, 보다 자연스러운 3차원 영상을 생성할 수 있었다.

Abstract In this paper, we present a new camera system combining a high-quality 3-D scanner and hybrid camera system to generate a multiview video-plus-depth. In order to get the 3-D video using the hybrid camera system and 3-D scanner, we first obtain depth information for background region from the 3-D scanner. Then, we get the depth map for foreground area from the hybrid camera system. Initial depths of each view image are estimated by performing 3-D warping with the depth information. Thereafter, multiview depth estimation using the initial depths is carried out to get each view initial disparity map. We correct the initial disparity map using a belief propagation algorithm so that we can generate the high-quality multiview disparity map. Finally, we refine depths of the foreground boundary using extracted edge information. Experimental results show that the proposed depth maps generation method produces a 3-D video with more accurate multiview depths and supports more natural 3-D views than the previous works.

핵심어: 3-D video generation, depth camera, depth estimation, multiview camera

본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 정보통신대학 IT연구센터(ITRC)와 한국과학재단 특정기초연구(R01-2007-000-20330-0)의 지원에 의한 것입니다.

*이은경 : 광주과학기술원 정보통신공학과 박사과정 e-mail: eklee78@gist.ac.kr

**정영기 : 호남대학교 컴퓨터공학과 교수 e-mail: ykjung@honam.ac.kr

***호요성 : 광주과학기술원 정보통신공학과 교수; e-mail: hoyo@gist.ac.kr

1. 서론

3차원 멀티미디어 응용분야가 다양해짐에 따라 3차원 비디오에 대한 관심이 증가하고 있고, 이에 따라 다시점 비디오와 그에 상응하는 깊이맵을 획득하는 것이 중요한 연구 분야로 관심을 받고 있다. 최근에 MPEG에서도 이러한 중요성을 인지하고 깊이 정보를 포함한 다시점 비디오(multi-view video with depth)를 3차원 비디오로 간주하여 자유시점 TV (free-viewpoint TV, FTV)에 활용하려는 표준화 활동이 진행되고 있다. 뿐만 아니라 3차원 비디오의 표준화를 위한 여러 방법들이 제안되고 있다 [1].

깊이맵을 생성하는 방법은 크게 능동적 깊이센서 방식(active depth sensors)과 수동적 깊이센서 방식(passive depth sensors)으로 나눈다. 능동적 깊이센서 방식은 레이저 센서, 적외선 센서, 패턴 센서를 이용하여 3차원 공간의 깊이 정보를 직접 획득하는 방법이다. 능동적 깊이센서 방식은 정확한 깊이 정보를 획득할 수 있는 반면, 저해상도의 깊이맵을 제공하고 큰 비용이 든다는 단점이 있다.

반면에, 수동적 깊이센서 방식은 다시점 및 스테레오 영상으로부터 스테레오 정합과 같은 간접적인 방법으로 3차원 깊이 정보를 획득하는 방법이다. 수동적 깊이센서 방식은 폐색 영역과 변화가 적은 단색 색상 영역에서 부정확한 깊이 정보를 제공하는 단점이 있지만, 고해상도의 깊이맵을 제공하고, 능동적 깊이센서 방식에 비해 적은 비용으로 깊이맵을 생성할 수 있다 [2, 3].

좀 더 정확한 3차원 깊이 정보를 획득하고자 능동적 깊이센서 방식과 수동적 깊이센서 방식의 장점을 결합한 복합형 방법이 연구되었다. 한국전자통신원(ETRI)에서는 3차원 스캐너 대신 깊이 카메라를 이용한 복합형 방법을 소개하였다. 깊이 카메라는 실시간으로 실물의 깊이 정보를 제공한다. 그러나 이와 같은 복합형 방법은 깊이 카메라를 이용하여 고품질의 깊이맵을 생성할 수 있지만, 깊이 카메라에 의존하여 깊이 정보를 측정하였기 때문에, 깊이 카메라가 제공하는 저해상도의 깊이맵을 생성하였다 [4].

Zhu는 다시점 깊이맵을 생성하기 위해 확률 분산 함수를 적용하여 깊이 카메라에서 획득한 깊이정보를 보정하는 방법을 제안하였다. 그러나 깊이맵 생성을 정지된 물체에 대해서만 적용하였다 [5]. 미래의 3차원 응용분야에서는 고품질의 3차원 비디오를 요구하기 때문에, 우리는 고화질의 다시점 3차원 비디오를 생성할 필요가 있다.

본 논문에서는 고화질 (high definition, HD) 복합형 카메라 시스템과 고품질 (high-quality) 3차원 스캐너를 결합하여 다시점 비디오와 그에 상응하는 다시점 깊이맵을 생성하는 시스템을 제안하여, 고해상도의 다시점 영상과 그에 상응

하는 깊이맵을 포함하는 3차원 비디오를 생성하는 새로운 방법을 제안한다. 제안하는 방법은 고화질 3차원 스캐너와 깊이 카메라에서 획득한 깊이정보를 이용하여 다시점 카메라를 위한 고품질 3차원 비디오를 생성한다. 또한 제안한 방법은 깊이 카메라가 아닌 고화질 다시점 카메라에 의존하기 때문에, 생성된 3차원 비디오의 해상도는 고해상도이다.

2. 제안하는 3차원 비디오 생성 방법

2.1 3차원 비디오 생성을 위한 시스템 구성

본 논문은 제안한 복합형 카메라 시스템과 고품질 3차원 스캐너를 이용해 다시점 카메라의 고화질 깊이맵을 생성한다. 그림 1은 복합형 카메라 시스템의 구성을 보여준다 [1]. 복합형 카메라 시스템은 기본적으로 고화질 다시점 카메라와 표준화질 깊이 카메라로 구성된다. 또한, 모든 카메라에 동기화 신호를 연속적으로 보내기 위해 동기화 신호 발생기가 각 카메라에 연결된다.

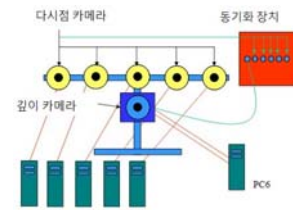


그림 1. 복합형 카메라 시스템

제안하는 3차원 비디오 생성 방법은 먼저 움직임에 없는 배경영역에 대한 깊이정보를 3차원 스캐너를 이용해 획득한다. 일반적으로 3차원 스캐너는 고품질의 깊이정보를 획득할 수 있지만, 움직이는 물체에 대한 깊이정보를 획득할 수 없는 단점이 있다. 그리고 움직임이 많은 전경영역은 다시점 카메라와 깊이 카메라를 이용해 색상정보와 깊이정보를 획득한다. 깊이 카메라의 경우, 동적인 영역에 대한 깊이정보를 실시간으로 획득할 수 있지만, 그 획득영역에 제한이 있어서 배경영역과 전경영역을 동시에 획득할 수 없는 단점을 가진다. 일반적인 깊이 카메라의 깊이정보 측정거리는 약 0.5~7m이다. 그러나 실제 응용에서 측정 가능한 깊이 정보 측정거리는 보통 약 2~4m 정도이다. 더욱이, 깊이정보 측정거리를 넓게 하면 할수록 깊이정보의 정확도가 떨어진다.

본 논문에서는 고품질의 깊이맵 생성을 위해 두 방법의 단점을 보완하면서 서로의 장점을 결합하기 위해 3차원 스캐너를 이용하여 배경영역의 깊이맵을 먼저 획득하고, 움직임이 많은 전경영역의 깊이맵을 깊이카메라를 이용해 획득하여 두 깊이정보를 결합하는 방법을 제안한다.

2.2 3차원 워핑을 통한 배경영역 깊이맵 생성

3차원 스캐너로 획득한 깊이 정보를 다시점 카메라의 배경영역의 깊이 정보로 간주한다. 다시점 카메라의 색상 정보와 상응하는 배경영역의 깊이 정보를 정합하기 위해, 다시점 카메라와 3차원 스캐너를 독립적으로 카메라 보정(camera calibration)한다. 결과적으로 다시점 카메라와 3차원 스캐너의 투영 행렬인 P_m 과 P_s 을 식 (1)과 같이 얻는다.

$$P_s = K_s [R_s | t_s] = \begin{bmatrix} K_{s_x} & 0 & P_{s_x} \\ 0 & K_{s_y} & P_{s_y} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{s_{00}} & R_{s_{01}} & R_{s_{02}} & t_{s_x} \\ R_{s_{10}} & R_{s_{11}} & R_{s_{12}} & t_{s_y} \\ R_{s_{20}} & R_{s_{21}} & R_{s_{22}} & t_{s_z} \end{bmatrix} \quad (1)$$

$$P_l = K_l [R_l | t_l] = \begin{bmatrix} K_{l_x} & 0 & P_{l_x} \\ 0 & K_{l_y} & P_{l_y} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{l_0} & R_{l_1} & R_{l_2} & t_{l_x} \\ R_{l_3} & R_{l_4} & R_{l_5} & t_{l_y} \\ R_{l_6} & R_{l_7} & R_{l_8} & t_{l_z} \end{bmatrix}$$

여기서, K_s , R_s , t_s 는 3차원 스캐너의 내부 및 외부 인수 행렬이고, K_m , R_m , t_m 은 다시점 카메라의 내부 및 외부 인수 행렬이다. 3차원 스캐너와 다시점 카메라의 상대적인 위치 관계를 계산하기 위해, 3차원 스캐너의 위치를 기준으로 다시점 카메라의 위치를 결정한다. 먼저, 식 (2)를 이용하여, 3차원 스캐너의 회전 행렬 R_s 를 단위 행렬 I 로, 이동 행렬 t_s 를 영행렬 O 로 변환한다.

$$R'_{ori} = R_s \cdot R_s^{-1} = I \quad (2)$$

$$t'_{ori} = t_s - t_s = O$$

그런 다음, 식 (3)과 같이, 3차원 스캐너의 회전 행렬 R_s 의 역행렬 R_s^{-1} 를 다시점 카메라의 회전행렬 R_m 에 곱하고, 3차원 스캐너의 이동 행렬 t_m 를 다시점 카메라의 이동 행렬 t_m 에서 빼준다. 3차원 스캐너에 대한 다시점 카메라의 상대적 위치인 회전 행렬 R'_m 과 이동 행렬 t'_m 를 결정한다.

$$R'_m = R_m \cdot R_s^{-1} \quad (3)$$

$$t'_m = t_m - t_s$$

끝으로, 3차원 스캐너의 색상 영상의 화소 위치 (p_{sx}, p_{sy}) 와 상응하는 깊이맵의 화소값이 $D_s(p_{sx}, p_{sy})$ 일 때, 3차원 화소 $p_s = (p_{sx}, p_{sy}, D_s(p_{sx}, p_{sy}))$ 를 식 (4)에서 계산한 투영 행렬 P'_m 을 이용하여 식 (5)와 같이 3차원 워핑을 수행한다.

$$P'_m = K_m [R'_m | t'_m] \quad (4)$$

$$p'_m = P'_m \cdot P_s^{-1} \cdot p_s \quad (5)$$

여기서, $p_m = (p_{mx}, p_{my}, 1)$ 은 p_s 에 상응하는 다시점 영상의 위치 정보 (p_{lx}, p_{ly}) 를 포함한다. 또한, p_m 에서의 깊이 정보 $D_m(p_{mx}, p_{my})$ 는 식 (6)과 같다. 그림 4는 3차원 워핑을 통해 얻어진 배경영역의 초기 깊이맵을 보여준다.

$$D_m(p_{mx}, p_{my}) = (t_{m_z} - t_{s_z}) + D_s(p_{sx}, p_{sy}) \quad (6)$$

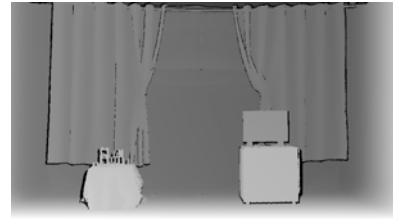


그림 2. 배경영역을 위한 초기 깊이맵

2.3 색상 분할에 기반으로 한 스테레오 정합

본 논문은 스테레오 정합을 수행하는 동안 세그먼트 단위로 3차원 워핑을 하고, 생성한 초기 깊이정보를 고려하여 다시점 영상의 정확한 깊이맵을 생성한다. 본 논문에서는 mean-shift 기반의 색상 분할 방법을 이용하여 다시점 영상을 색상 분할한다. 제안한 방법은 3차원 워핑된 세그먼트를 기반으로 한 스테레오 정합을 수행하여 깊이맵을 생성한다. 그림 2는 색상 분할된 세그먼트를 3차원 워핑을 통해 시점 변화에 따른 형태를 예측하고, 예측된 세그먼트를 이용해 주변영역을 탐색하는 방법을 그림으로 보여준다.

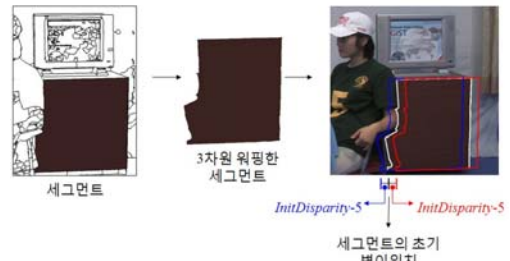


그림 3. 색상 분할 기반의 스테레오 정합

2.4 깊이맵 정제

색상분할에 기반한 스테레오 정합 방법으로 깊이 정보를 획득할 경우, 하나의 세그먼트는 하나의 변이값을 갖는다. 그러나 실제 환경은 연속적인 깊이정보로 이루어져 있다. 이러한 문제점을 해결하기 위해 본 논문은 *belief propagation* 방법[6]을 이용해 깊이맵을 정제한다. 깊이 카메라를 통해 획득한 깊이정보를 이용해 검색영역을 결정하고 *belief propagation* 방법을 이용해 깊이맵을 정제한다.

스테레오 정합과 같은 영역 기반 정합 방법으로 깊이 정보를 획득할 경우, 폐색영역과 같은 불연속적인 영역에서 부정확한 깊이 정보를 예측한다. 본 논문은 계위공간(scale space)을 이용해 다시점 영상의 경계 정보를 생성하고, 이 경계 정보를 정합하여 3차원 경계 조각을 추출한다. 그런 다음, 경계 조각을 특징 기반 정합 방법을 적용하여 깊이맵 경계 영역의 깊이 정보를 다시 한번 정제한다. 계위공간의 원리는 영상을 점진적으로 흐리게 하는 처리과정을 말한다. 그리고 계위공간은 영상 특징의 계층화 개념을 가지고 있으며, 화소와 관련이 있는 표현부터 의미 있는 영상 영역까지 묘

사하는 중요한 단계들로 구성된다.

기존의 방법은 흐림과정(smoothing)을 여러 번 반복 적용하면 영상이 흐려진다. 이때 기존의 계위공간 방법들과는 다르게 잡음을 제거하는 과정에서 생기는 에지의 흐림 효과를 줄이기 위해서 관심없는 영역만 흐리게 처리를 하고, 경계영역이라고 판단되는 부분에 있어서는 흐림과정을 적게 한다. 이런 계위공간의 개념과 경계의 흐림효과를 줄이면서 영상 내에 존재하는 잡음을 제거하는 특징을 이용한다. 그러나, 기존의 방법도 반복해서 적용하는 횟수가 늘어날수록 경계영역에 대한 흐림효과가 심해지기 때문에 낮은 계위에서 보다는 높은 계위에서 경계영역의 정보 손실이 심해진다.

3. 실험결과

본 논문은 복합형 카메라 시스템과 고품질 3차원 스캐너를 이용해 고화질 깊이맵을 생성하는 방법을 제안했다. 깊이맵의 정확도를 평가를 위해 기존의 깊이맵 생성 방법으로 깊이맵을 생성하여 그림 4와 같이 결과를 비교하였으며, 생성한 깊이맵을 이용해 3차원 장면을 생성하고 중간시점 영상을 그림 5와 그림 6과 같이 생성하였다.

본 논문은 기존의 깊이맵 생성 방법이 폐색영역과 비슷한 색상정보를 가지는 영역에 대해 잘못된 깊이정보를 생성하는 문제점을 해결하였고, 고해상도 3차원 스캐너와 저해상도 깊이 카메라를 이용해 고해상도 다시점 깊이맵을 생성하는 방법을 제시하였다.

4. 결론

본 논문은 복합형 카메라 시스템과 고품질 3차원 스캐너를 이용해 고화질 깊이맵을 생성하는 방법을 제안했다. 깊이맵의 정확도를 평가를 위해 기존의 깊이맵 생성 방법으로 깊이맵을 생성하여 그 결과를 비교하였으며, 생성한 깊이맵을 이용해 3차원 장면을 생성하고 중간시점 영상을 생성하였다. 본 논문은 기존의 깊이맵 생성 방법이 폐색영역과 비슷한 색상정보를 가지는 영역에 대해 잘못된 깊이정보를 생성하는 문제점을 해결하였고, 고해상도 3차원 스캐너와 저해상도 깊이 카메라를 이용해 고해상도 다시점 깊이맵을 생성하는 방법을 제시하였다.



그림 4. 다시점 영상과 생성한 고화질 깊이맵



그림 5. 색상영상과 깊이맵을 이용한 3차원 장면생성

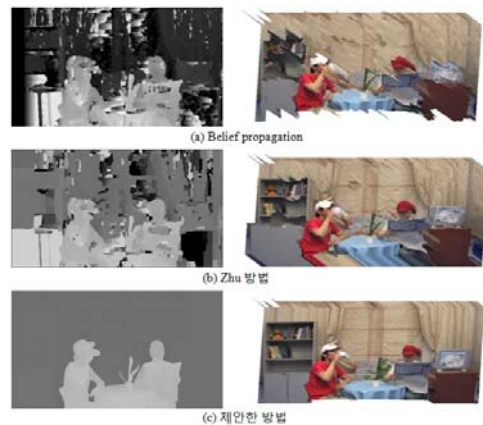


그림 6. 색상영상과 깊이맵을 이용한 3차원 장면재현

참고문헌

- [1]ISO/IEC JTC1/SC29/WG11 N8944, "Preliminary FTV Model and Requirements," 2007
- [2]M. Okutimi and T. Kanade, "A multiple-baseline stereo," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 4, pp. 353-363, 1993
- [3]C. Zitnick, S. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-quality video view interpolation using a layered representation," Proc. of ACM

SIGGRAPH, pp. 600–608, 2004

- [4]G. Um, K. Kim, C. Ahn, and K. Lee, “Three-dimensional scene reconstruction using multiview images and depth camera,” 3D Digital Imaging and Modeling, pp. 271–280, 2005
- [5]J. Zhu, L. Wang, R. Yang, and J. Davis, “Fusion of time-of-flight depth and stereo for high accuracy

depth maps,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 231–236, 2008.

- [6]F. Pedro and Felzenszwalb, P. Daniel, “Efficient belief propagation for early vision,” International Journal of Computer Vision vol.70, no 1, pp. 41–54, 2006.