

복합형 카메라 시스템을 이용한 고화질 3차원 비디오 생성 방법

이은경⁰, 호요성

광주과학기술원 정보통신공학과

eklee78@gist.ac.kr, hoyo@gist.ac.kr

요약

본 논문은 다시점 비디오와 그에 상응하는 다시점 깊이 영상을 생성하기 위해 다시점 비디오 카메라와 깊이 카메라를 결합한 복합형 카메라 시스템을 제안한다. 제안한 복합형 카메라 시스템을 이용해 3차원 비디오를 생성하기 위해서는, 우선 다시점 카메라와 깊이 카메라를 이용해 다시점 비디오와 깊이 영상을 획득하고, 획득한 깊이 영상을 이용해 3차원 워핑(warping)을 적용하여 각 다시점 카메라를 위한 초기 깊이 정보를 예측한다. 3차원 워핑을 이용해 초기 깊이 정보를 예측하는 것은 다시점 카메라의 각 시점에서의 초기 깊이 영상을 계산하기 위한 것이다. 본 논문에서는 물체의 외곽선 영역의 깊이 영상 정확도를 높이기 위해 비디오 앞 뒤 영상의 색상 차분 정보를 이용해 움직임 영역을 추출한다. 마지막으로, 고화질의 다시점 깊이 영상을 생성하기 위해서 제안한 비용 함수를 이용해 깊이 영상을 정제한다. 제안한 복합형 카메라 시스템은 기존의 깊이 영상 예측 방법보다 정확한 다시점 깊이 영상을 포함하는 3차원 비디오를 생성할 수 있었으며, 보다 자연스러운 3차원 영상을 생성할 수 있었다.

1. 서론

3차원 멀티미디어 응용분야에 대한 관심이 증가하면서 3차원 비디오를 생성하고자 하는 노력이 많은 분야에서 진행되고 있다. 3차원 TV와 같은 차세대 3차원 응용에서 고화질의 3차원 비디오를 지원하기 위해서는 정확한 깊이 정보를 포함하는 깊이 영상을 효과적으로 생성해야 한다 [1]. 최근, 국제 표준화 그룹인 Moving Picture Experts Group (MPEG)에서는 이러한 자유시점 TV와 3차원 TV에 대한 중요성을 인식하고, 3차원 비디오 부호화 기술의 표준화 작업을 활발히 진행하고 있다 [2].

머지않은 미래에 시청자들은 많은 응용 분야를 통해 3차원 깊이를 체험할 수 있을 뿐만 아니라, 3차원 비디오를 이용해 시청자가 원하는 위치의 시점을 선택하여 볼 수도 있을 것이다. 이미 유럽의 ATTEST 프로젝트는 깊이 카메라를 이용한 3차원 TV 시스템을 선보였으며, 광주과학기술원 실감방송연구센터에서도 깊이 카메라를 이용하여 3차원 방송 콘텐츠를 제작했다. 이와 같이 3차원 비디오를 이용해 임의 시점에서 3차원 장면을 생성하기 위해서는 정확한 깊이 정보를 획득하는 것이 무엇보다 중요하다. 오래 전부터 컴퓨터 비전 분야에서는 스테레오 정합 기술을 이용한 깊이 영상을 생성하는 방법이 연구되어 왔다 [3]. 그러나 신뢰할 만한 깊이 영상을 생성하는 것은 여전히 해결해야 할 문제점으로 남아 있다. 깊이 영상을

생성하는 방법은 크게 능동적 깊이 센서 방식(active depth sensors)과 수동적 깊이 센서 방식(passive depth sensors)으로 나누어진다. 능동적 깊이 센서 방식은 레이저 센서, 적외선 센서, 패턴 센서와 같은 물리적인 센서 장치를 이용해 3차원 공간상의 깊이 정보를 직접 획득하는 방식이다 [4]. 능동적 깊이 센서 방식은 정확한 깊이 정보를 실시간으로 획득할 수 있지만, 저해상의 깊이 영상만을 제공할 뿐만 아니라 큰 비용이 든다는 단점이 있다.

반면에, 수동적 깊이 센서 방식은 스테레오 및 다시점 카메라에서 획득한 영상을 이용해 스테레오 정합을 수행하여 3차원 정보를 획득하는 방법이다 [5]. 수동적 깊이 센서 방식은 고해상도의 깊이 영상을 생성할 수 있지만, 능동적 깊이 센서 방식에 비해 부정확한 깊이 영상을 생성한다.

이러한 두 방식의 단점을 해결하고 좀 더 정확한 3차원 깊이 정보를 생성하기 위하여 능동적 깊이 센서 방식과 수동적 깊이 센서 방식을 결합한 복합형 카메라 시스템을 이용한 방법들이 제안되었다[6][7]. 한국전자통신원(ETRI)에서는 깊이 카메라를 3시점 카메라 시스템을 결합한 복합형 카메라 시스템을 소개하였다. 그러나 기존의 복합형 카메라 시스템은 깊이 카메라를 이용하여 고품질의 깊이 영상을 생성할 수 있지만, 저해상도의 깊이 카메라를 기준으로 다시점 카메라의 해상도를 낮추어 깊이 영상을 생성하였기 때문에 저해상도의 3차원 비디오를 생성하였다 [8].

최근 정확한 깊이 영상을 생성하기 위해 깊이

카메라에서 획득한 깊이 정보를 보정하는 방법들이 제안되었다. 그러나 이 방법들은 모두 정지된 장면에 대한 깊이 영상을 생성하였고, 저해상도의 깊이 영상만 생성할 수 있다. 미래의 3차원 응용분야에서는 고품질의 3차원 비디오를 요구하기 때문에, 우리는 고품질의 다시점 3차원 비디오를 생성할 필요가 있다.

본 논문에서는 고훈상도의 다시점 카메라와 저해상도 깊이 카메라를 결합하여 고훈상도의 다시점 영상과 그에 상응하는 깊이 영상을 포함하는 3차원 비디오를 생성하는 방법을 제안한다. 제안한 방법은 깊이 카메라에서 획득한 깊이 정보를 이용해 다시점 카메라를 위한 고훈상도 3차원 비디오를 생성한다.

2. 복합형 카메라 시스템

그림 1은 제안하는 복합형 카메라 시스템과 그 구성을 보여준다. 제안하는 복합형 카메라 시스템은 5대의 고훈상도 다시점 카메라와 1대의 저해상도의 깊이 카메라로 이루어진다. 각 카메라는 비디오 저장 보드가 갖추어진 PC 와 연결된다. 또한, 동기화 신호 재생기가 모든 카메라에 연결되어 동기화 신호를 연속적으로 보낸다. 제안하는 복합형 카메라 시스템은 프레임마다 다시점 영상과 깊이 영상을 동시에 획득할 수 있다. 표 1은 복합형 카메라 시스템의 구성을 보여준다. 복합형 카메라 시스템은 다시점 카메라, 깊이 카메라, 동기화 세 부분으로 구성된다. 제안하는 복합형 카메라 시스템은 매 프레임마다 다시점 카메라에서 획득한 다시점 영상과 깊이 카메라로부터 획득한 깊이 영상을 제공한다.

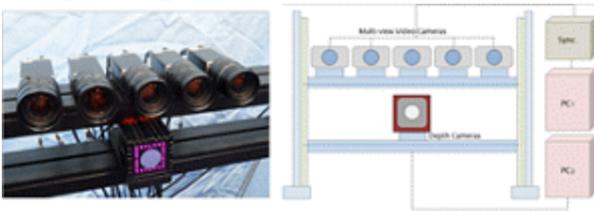


그림 1. 제안한 복합형 카메라 시스템

표 1. 복합형 카메라 시스템 구성

장치	상세정보	
다시점 카메라	출력형식	NTSC / PAL (16:9)
	해상도	1920 (h) × 1080 (v)
깊이 카메라	깊이범위	0.5 ~ 5.0 m
	시야각	43.6° (h) x 34.6°(v)
	출력양식	NTSC / PAL (4:3)
	해상도	176 (h) × 144 (v)
동기화 신호	출력형식	SD / HD 비디오 생성

3. 3차원 비디오 생성

3.1 세그먼트 기반 스테레오 정합

본 논문은 스테레오 정합을 수행하는 동안 세그먼트 단위로 3차원 워핑을 하고, 생성한 초기 깊이정보를 고려하여 다시점 영상의 정확한 깊이 영상을 생성한다. 본 논문에서는 mean-shift 기반의 색상 분할 방법을 이용하여 다시점 영상을 색상 분할한다. 제안한 방법은 3차원 워핑된 세그먼트를 기반으로 한 스테레오 정합을 수행하여 깊이 영상을 생성한다. 이때 세그먼트의 크기가 너무 클 경우, 하나의 세그먼트가 하나의 깊이 값으로 정의되기 때문에 본 논문에서는 최소 128 화소 이상, 최대 256 화소 이내의 크기로 세그먼트 크기를 다시 정의한다. 그림 2는 세그먼트를 다시 정의하는 방법을 보인 것이다.

모든 영상을 먼저 16×16 블록으로 나눈다. 그리고 하나의 블록 안에 각 세그먼트를 하나의 세그먼트 단위로 정회한다. 그 다음 각 세그먼트의 크기가 앞에서 정의한 최소 세그먼트 크기보다 작은 경우는 주변 블록을 탐색하여 블록을 정합한다. 정합을 위한 블록 탐색 순서는 대각선 방향을 포함하여 오른쪽, 아래, 왼쪽, 위쪽 방향으로 진행한다. 왜냐하면 아직 정합되지 않을 영역을 먼저 정합하고 이후 정합이 이루어 지지 않은 이전 블록을 탐색하기 위해서 이다. 이러한 과정을 앞서 정의한 최대 블록 크기를 만족할 때까지 수행한다.

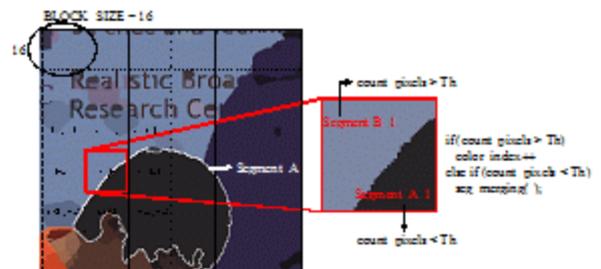


그림 2. 세그먼트 재정의

색상 분할에 기반으로 한 스테레오 정합에서 일반 자연 영상은 비슷한 변이를 가지는 영역으로 분할되고, 세그먼트는 하나의 변이를 가진다고 가정한다. 각 세그먼트에 초기 변이를 설정하기 위해서, 변이와 깊이 변환 식을 이용해 초기 깊이 영상의 깊이 정보를 변환한다.

각 세그먼트의 변이를 결정하기 위해, 우리는 초기 깊이 값을 위한 탐색 영역을 설정한다. 그림 3에서 보는 것처럼 3차원 워핑을 통해 얻은 초기 깊이 영상은 실제 색상 영상과 경계영역에서 잘 일치하지 않는다. 이 때문에 경우에 따라서는 깊이 값이 잘못된 영역에 위치하는 문제도 발생한다. 이렇게 깊이 값이 잘못 설정될 경우 잘못된 정합

결과를 가져올 수 있기 때문에 본 논문에서는 초기 깊이 값 설정을 위한 영역을 그림 3처럼 가로, 세로 방향으로 확대하여 설정한다. 그리고 설정한 영역 안의 초기 깊이 값을 이용해 재정의한 세그먼트를 기반으로 스테레오 정합을 수행한다. 초기 깊이 값을 기반으로 MAD 값이 가장 작은 영역을 정합 위치로 결정하여 초기 깊이 영상을 생성한다.

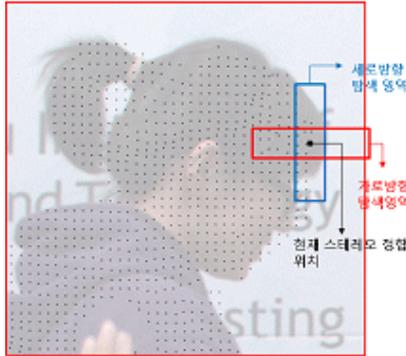


그림 3. 탐색 영역 설정

2.2 움직임 영역 분리

일반적으로 스테레오 정합을 통해 생성된 초기 깊이 영상은 전경영역과 배경영역의 경계부분에서 많은 불일치를 가진다. 본 논문에서는 깊이 영상의 정확도를 높이기 위해서 움직임 정보 찾고 이를 이용해 전경 영역을 추출한다. 그림 4는 이전 프레임과 현재 프레임의 색상 차이를 이용해 현재 프레임에서의 움직임 영역을 추출한 결과이다. 움직임이 없는 영역은 이전 프레임과 현재 프레임이 차이가 없는 영역으로 이전 프레임에서 생성한 깊이 정보를 참조할 수 있다. 차분 영상은 RGB 차분의 합이 임계값 이상일 경우 움직임 영역으로 판단한다. 본 논문에서는 10이상인 경우를 움직임 영역으로 설정하였다.



그림 4. 움직임 영역 추출

2.3 깊이 영상 정제

색상분할에 기반한 스테레오 정합 방법으로 깊이 정보를 획득할 경우, 하나의 세그먼트는 하나의 변이 값을 갖는다. 그러나 실제 환경은 연속적인 깊이정보로 이루어져 있다. 이러한 문제점을 해결하기 위해 본 논문은 깊이 영상 정제 함수를 제안한다. 깊이 카메라에서 획득한 깊이

정보, 스테레오 정합을 통해 생성한 변이 정보, 그리고 이전 프레임에서 생성한 깊이 정보를 이용해 식 (1)과 같은 비용함수를 정의한다.

$$E(x, y, d) = \begin{cases} w_s f_s(x, y, d_s(x, y)) + w_d f_d(x, y, d_d(x, y)) & \text{if } obj_mov(x, y) = 1 \\ w_s f_s(x, y, d_s(x, y)) + w_d f_d(x, y, d_d(x, y)) + w_f f_f(x, y, d_f(x, y)) & \text{if } obj_mov(x, y) = 0 \end{cases} \quad (1)$$

여기서 f_s 는 깊이 영상의 자연스러운 깊이 변화를 주기 위해 깊이 카메라에서 획득한 깊이 영상에서 현재 위치를 기준으로 주변 화소 간의 편미분 값을 적용해 자연스러운 깊이 영상 생성을 고려한 값이고, f_d 는 세그먼트 기반에서 한번 정제한 깊이 정보를 다시 화소 단위로 스테레오 정합을 수행하여 최소 MAD 를 갖는 위치를 찾은 결과이다.

마지막으로 f_f 는 이전 프레임에서 생성한 깊이 정보로 이전 단계에서 계산한 움직임 영역결과에서 움직임 없는 영역 즉, obj_mov 값이 0인 위치에서만 이전 프레임의 깊이 정보를 참조한다. 이 세 가지 정보를 하나의 함수로 정의하여 비용함수 값이 가장 큰 값을 최상의 깊이 값으로 결정한다. 이 비용함수의 경우 이전의 정제함수에서 항상 존재하는 반복적인 계산이 없기 때문에 깊이 영상을 생성하는 시간을 단축시킬 수 있는 장점이 있다.

4. 실험 결과 및 분석

제안한 깊이 영상 생성 방법을 평가하기 위해, 5대의 다시점 카메라와 깊이 카메라를 이용하여 복합형 카메라 시스템을 구성하였다. 그림 5는 제안한 복합형 카메라 시스템을 이용해 획득한 입력영상이다. 5 개의 다시점 색상영상과 1장의 깊이 영상을 획득하였다.



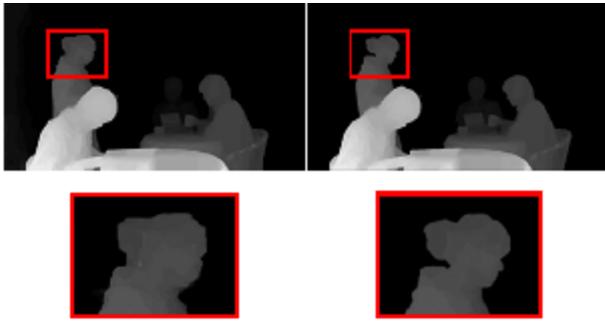
그림 5. 입력 영상

그림 6은 제안한 방법으로 생성한 깊이 영상과 기존의 알고리즘을 이용해 생성한 깊이 영상 결과를 비교한 결과이다. 기존 방법은 DERS 프로그램을 이용했다. 결과에서 확인할 수 있듯이 기존의 방법은 몇몇 영역에서 잘못된 정합 결과를 포함하고 있으나, 제안한 방법으로 생성한 깊이 영상은 머리카락과 같이 복잡하고 미세한 외곽선 정보를 가지는 영역과 같이 비슷한 색상정보가 반복적으로 존재하는 영역에서도 좋은 결과를

보였다. 표 2는 제안한 방법과 기존 방법을 통해 생성한 깊이 영상을 이용해 중간 시점 영상을 생성하고 원본 영상과의 PSNR 을 구한 결과이다. 결과에서도 확일 할 수 있듯이 제안한 방법으로 생성한 중간영상이 더 정확하게 생성할 수 있음을 확인했다.

표 2. PSNR 비교

	기존방법	제안한 방법
PSNR	34.89	35.18



(a)DEIS를 이용해 생성한 깊이 영상 (b)제안한 시스템을 이용해 생성한 깊이 영상

그림 6. 깊이 영상 생성결과 비교

그림 7은 카메라 3에서의 생성한 깊이 영상을 30프레임마다 보여준 결과이다. 그림 4는 카메라 2과 카메라 4의 색상영상과 그에 상응하는 깊이 영상을 이용해 중간시점 영상을 생성한 결과이다. 그림 4에 보인 것처럼, 깊이 영상을 이용해 생성한 중간영상은 자연스러운 3차원 장면 영상을 생성했다.



그림 7. 깊이 비디오 생성 결과



그림 8. 중간영상 합성 결과

5. 결론

본 논문은 복합형 카메라 시스템을 이용해 고화질 깊이 영상을 생성하는 방법을 제안했다. 영상 내에 움직이는 물체를 구분하여 깊이 영상 경계 영역에서도 기존 방법과 비교했을 때 정확한 깊이 영상을 생성할 수 있었다. 또한 깊이 영상 정제를 위해 획득한 깊이 영상, 스테레오 정합 그리고 이전 프레임의 깊이 정보를 이용해 새로운 비용함수를 정의하였고, 이를 이용해 고화질의 다시점 깊이 영상을 생성 할 수 있었다. 이를 통해 실험 결과에서도 확인할 수 있듯이, 기존의 깊이 영상 생성 방법보다 향상된 화질의 깊이 영상을 얻었다. 최종적으로 복합형 카메라 시스템과 제안한 깊이 정제함수를 이용해 고해상도 3차원 비디오를 생성하는 방법을 제시했다.

감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행됨. (NIPA-2011-(C1090-1011-0003)).

참고문헌

- [1] P. Kauff, N. Atzpadin, C. Fehn, M. Müller, O. Schreer, A. Smolic, and R. Tanger, "Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability," *Signal Processing Image Communication*, vol. 22, no. 2, pp. 217-234, Feb. 2007.
- [2] ISO/IEC JTC1/SC29/WG11 N8944, "Preliminary FTV model and requirements," April 2007.
- [3] A. Smolic and D. McCutchen, "3DA V exploration of video-based rendering technology in MPEG", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, no. 3, .pp. 348-356, March 2004.
- [4] S. Nayar and Y. Nakagawa, "Shape from Focus," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 16, no. 8, pp. 824-831, Aug. 1994.
- [5] G. Iddan and G. Yahav, "3D imaging in the studio and elsewhere," *Proc. of SPIE Vidometrics and Optical Methods for 3D Shape Measurements*, pp. 48-55, Jan. 2001.
- [6] E. Lee, Y. Kang, Y. Jung, Y. Ho, "3-D video generation using hybrid camera system," *Proc. of International Conference on Immersive Telecommunications (IMMERSCOM)* , pp. T5(1-6), June 2009.
- [7] G. Um, K. Kim, C. Ahn, and K. Lee, "Three-dimensional scene reconstruction using multi-view images and depth camera," *Proc. of SPIE Stereoscopic Displays and Virtual Reality Systems XII*, vol. 5664, pp. 271-280, Jan. 2005.
- [8] 호요성, 이은경, 강윤석, "다시점 깊이 카메라를 이용한 3 차원 입체영상의 정보 획득 방법," *방송공학회지*, 제 25 권, 제 2호, pp. 88-100, 2010.