이기종 카메라를 이용한 3 차원 재구성 기반 다시점 깊이 영상 생성

신동원O, 호요성
광주과학기술원 전자전기컴퓨터공학부
dongwonshin@gist.ac.kr, hoyo@gist.ac.kr

요 약

본 논문에서는 이기종 카메라를 이용한 3 차원 재구성을 기반의 다시점 깊이 생성 방법을 소개한다. 이 연구의 주요 목표는 서로 다른 위치에 배치된 깊이 카메라를 사용하여 컬러 카메라의 각 시점에서 정확한 깊이 영상을 생성하는 것이다. 기존의 필터 기반 다시점 깊이 영상 생성 방법은 잘린 깊이 영역과 혼합 된 깊이 값과 같은 중요한 문제를 가지고 있다. 이 문제는 깊이 영상의 품질 뿐만 아니라 합성된 중간 시점 영상의 질을 저하시킨다. 제안하는 카메라 시스템은 하부 층에 4 개의 색상 카메라를 포함하고 상부 층에 2 개의 깊이 카메라를 포함하는 2 개의 카메라 층으로 구성된다. 먼저 사전 과정으로 카메라 캘리브레이션을 사용하여 올바른 카메라 매개 변수를 예측한 뒤 본 과정에서는 이기종 다시점 카메라 시스템을 이용하여 동기화된 색상 및 깊이 영상을 촬영한다. 다음으로 2 차원 깊이 영상으로부터 3 차원 점군을 생성하고 반복적 점군 정합 방법으로 최적의 카메라 외부 매개 변수를 예측하여 통합 된 3 차원 점군 모델을 얻는다. 그 후, 절단 된 부호 거리 함수로 희소한 형태의 3 차원 점군 모델로부터 체적 표면 모델을 만든다. 마지막으로 체적 표면 모델을 투영하여 각 색상 카메라 시점에서의 깊이 영상을 예측한다. 실험 결과 장에서는 제안하는 방법이 앞서 언급된 문제를 해결하며 기존 방법에 비해 가지는 장점에 대한 내용을 설명한다.

1. 서론

최근 우리는 일상생활에서 3 차원 TV, 헤드 마운트 디스플레이 및 스마트 폰과 같은 다양한 장치를 통해 실감나는 3 차원 콘텐츠를 쉽게 접할 수 있다. 이 맥락에서 정확한 3 차원 깊이 정보는 실제 장면의 완벽한 표현을 생성하고 사용자에게 몰입감을 주는 경험을 제공하는 데 반드시 필요하다.

이러한 3 차원 콘텐츠는 스테레오 카메라, 다시점 카메라 및 깊이 카메라와 같은 다양한 카메라 시스템을 이용해 생성 될 수 있다. 일반적으로 이러한 깊이 획득 시스템은 수동 센서 기반 방법과 능동 센서 기반 방법으로 분류 할 수 있다. 수동 센서 기반의 방법에서는 양안식 카메라를 이용한 스테레오 매칭 방법을 이용하여 깊이 영상을 계산한다. 능동 센서 기반 방법의 경우에는 적외선이나 레이저 광선을 사용하는 깊이 카메라를 활용하여 카메라와 대상 간의 거리를 측정한다.

색상과 깊이 카메라를 포함하는 이기종 카메라 시스템을 이용하여 3 차원 콘텐츠를 생성할 경우 각 색상 카메라 시점에서 정확한 깊이 영상을 만드는 것이 무엇보다 중요하다. 그러나 기존의 다시점 깊이 생성 방법에서는 3 차원 영상 워핑 및 양방향 필터를 사용하여 절단 및 혼합 깊이 값과 같은 문제를 유발한다. 따라서 본 논문에서는 이러한 문제를 해결하기 위해 3 차원 재구성을 기반으로 하는 다시점 깊이 영상 생성 방법을 제안한다.

2. 기존의 다시점 깊이 영상 생성 방법

기존의 다시점 깊이 영상 생성 방법에서는 8 개의 컬러 카메라와 3 개의 깊이 카메라를 포함하는 이기종 다시점 카메라 시스템을 사용했다. 이 방법의 목적은 색상 카메라와 동일한 해상도로 각 시점에서 해당 깊이 영상을 생성하는 것이다. 이 목표를 달성하기 위해 카메라 캘리브레이션, 기하학적 영상 보정, 3 차원 영상 워핑 및 양방향 필터와 같은 다양한 컴퓨터 비전 기술이 사용되었다. 그림 1 은 기존 다시점 깊이 영상 획득 방법의 전체 과정을 나타낸다 [1].

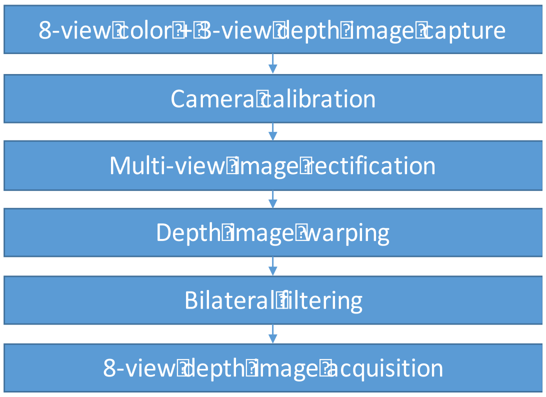


그림 1. 기존의 다시점 깊이 영상 획득 방법

2.1 절단된 깊이 영상 문제

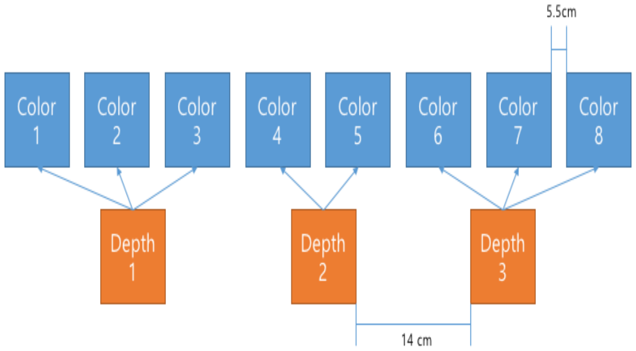


그림 2. 기존 방법의 카메라 시스템

기존의 방법에서는 깊이 카메라 3 대 중 가운데 카메라를 색상 카메라의 중간 위치에 고정했고 다른 두 개의 깊이 카메라를 중간에 고정된 카메라로부터 각각 좌우로 14cm 떨어져 배치했다. 그림 2 는 기존의 이기종 다시점 카메라 시스템을 도식화한 그림이다.

그림 2 에서 깊이 2 카메라는 색상 4 및 색상 5 카메라에 대한 깊이 영상 예측을 담당하고 깊이 1 과 깊이 3 카메라는 각각 인접하는 3 개의 색상 카메라에 할당되어 해당하는 깊이 영상을 예측한다. 색상과 깊이 카메라 사이의 거리가 너무 멀면 잘못된 깊이 결과가 발생하기 때문에 이러한 구성이 불가피하다. 따라서 이러한 구성 때문에 기존의 방법에서는 절단된 깊이 영역 문제가 발생한다. 즉, 깊이 카메라의 깊이 값이 모든 색상 카메라로 전송되지 않기 때문에 각 깊이 카메라의 제한된 시야에 의해 깊이 영상이 절단되어 예측된다.

2.2 혼합된 깊이 값 문제

기존의 다시점 카메라 시스템에서 색상 카메라 시점의 깊이 데이터는 깊이 카메라에 의해 획득된 깊이 값에 기초하여 생성된다. 이 때, 깊이 영상과 색상 영상의 해상도를 일치시키기 위해 원래의 깊이 카메라 매개 변수와 보정된 색상 카메라 매개 변수를 사용하여 깊이 영상에 3 차원 영상 워핑을 적용한다.

영상 워핑을 이용해 색상 카메라의 각 시점에서 8 개의 깊이 영상이 생성되지만 이 깊이 영상에는 색상 카메라와 깊이 카메라 간의 해상도 차이로 인해 빈 공간이 포함된다. 기존의 방법에서는 그러한 빈 영역을 채우기 위해 양방향 필터를 이용했다 [2]. 이 단계에서 워핑된 깊이 영상의 깊이 값은 양방향 필터 기반 홀 채우기 방법의 특성에 의해 오염되고 혼합 될 수 있다. 이는 희박한 형태로 워핑된 깊이 값이 서로 다른 위치의 깊이 값과 혼합되기 때문이다.

3. 제안하는 다시점 깊이 영상 생성 방법

제안하는 이기종의 다시점 카메라 시스템은 그림 3 과 같이 하부 층에 4 개의 색상 카메라가 있고 상부 층에 2 개의 깊이 카메라가 있는 2 개의 층으로 구성된다.

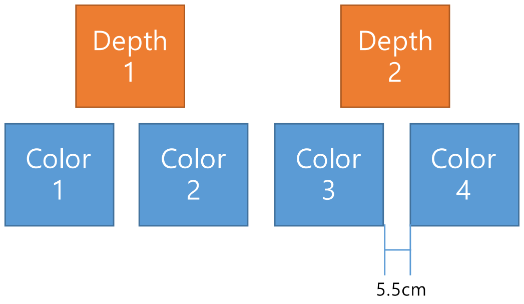


그림 3. 제안하는 방법의 카메라 시스템

카메라 캘리브레이션 단계는 사전 과정에서 올바른 카메라 매개 변수를 얻는 데 반드시 필요하다. 본 과정에서는 다시점 카메라 시스템에서 동기화된 색상 및 깊이 영상을 획득한 다음 깊이 영상의 잡음을 줄이기 위해 양방향 필터를 수행한다. 2 차원 영상 좌표의 깊이 화소를 3 차원 세계 좌표의 정점으로 만들고, 반복적 점군 정합 방법을 사용하여 점군 모델을 정합하면 통합 된 3 차원 점군 모델을 만들 수 있다. 조밀한 3 차원 표면 모델은 희소한 형태의 3 차원 점군 모델에 절단된 부호 거리 함수 방법을 수행하여 구성 할 수 있다 [3]. 마지막으로, 표면 모델을 각 색상 카메라 시점으로 투사하여 각 색상 카메라에서 깊이 영상을 얻을 수 있다. 그림 4 는 제안하는 방법의 전체적인 순서도를 보여준다.

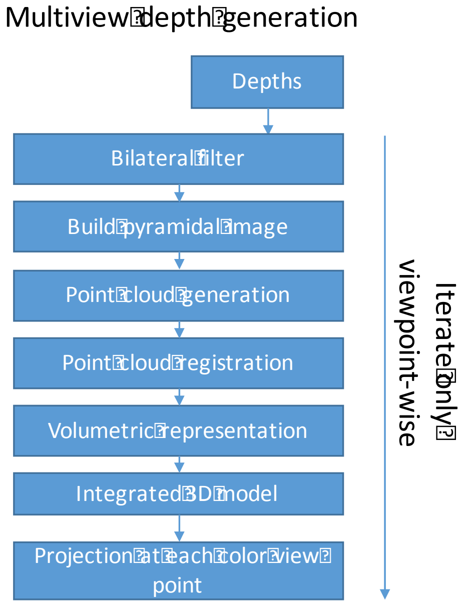


그림 4. 제안하는 방법의 순서도

3.1 3 차원 점군 생성

TOF 카메라로부터 깊이 영상을 획득한 다음, 카메라 내부 매개변수를 이용하여 3 차원의 점군을 생성할 수 있다.

$$\mathbf{M}_{\mathrm{W}} = \mathbf{A}_{\mathrm{i}}{}^{-1}\tilde{\mathbf{m}} \quad where\ \boldsymbol{A} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

행렬 **A** 는 카메라 내부 매개변수를 나타내고 $\tilde{m}$ 는 2 차원의 영상 좌표를 동차좌표계에서 표현한 것이다. M_W 는 세계 좌표계에서 3 차원의 점 좌표를 나타낸다.

3.2 점군 정합

깊이 영상으로부터 3 차원 점군 모델을 획득하면 통합된 하나의 좌표계에 각 깊이 카메라로부터의 점 구름을 정합해야 한다. 반복적 점군 정합 방법은 2 개의 점군 모델 사이의 6 차원의 자유도 변환을 계산하는 점군 정합 기법이다 [4]. 이 방법은 대응 설정, 법선 벡터 계산 및 반복 형태 최적화와 같은 주요 단계로 구성된다.

첫번째로 첫번째 점군 모델에서 두번째 점군 모델 사이의 대응 관계가 확립되어야 한다. 본 논문에서는 비교적 정확한 대응 관계를 확립하기 위해 점군 사영을 사용하였다 [5]. 정확한 대응 관계는 점군 정합의 정확성을 결정하므로 매우 중요하다.

다음으로, 대응간의 거리를 측정하는데 이 때, 각 3 차원 점에서의 법선 벡터는 대응 간의 거리를 측정하기 위해 중요한 요소이다. 본 논문에서 대응 간 거리는 두 인접 벡터 사이의 단순 외적에 의해 계산되고 최적화 단계에서 그림 5 와 같이 시작점 s_i 에서 목적점 d_i 의 접평면까지의 수직 거리 l_i 를 측정하기 위해 사용된다. 여기에서 행렬 M 은 회전 행렬 및 이동 벡터를 포함하는 3 차원의 강체 변환을 나타낸다. 따라서 이러한 대응 간의 거리를 측정하려면 법선 벡터 n_i 가 필요하다 [5].

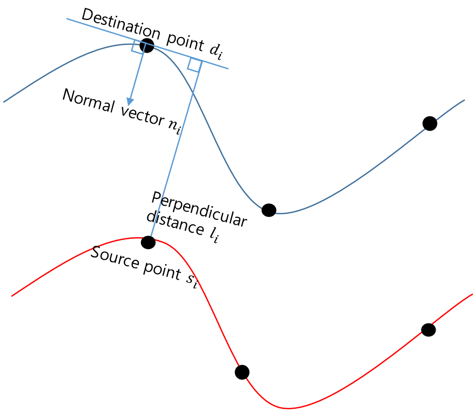


그림 5. 대응 간 거리 측정

3.2 복셀을 이용한 체적 모델 표현

통합된 3 차원 점군 모델은 희소한 형태로부터 부드럽게 보간된 형태를 얻기 위해 표면 복셀을 이용한 체적 표현으로 변환된다. 이 단계에서 절단된 부호 거리 함수 방법이 변환에 사용된다 [6]. 절단된 부호 거리 함수 방법의 주요 아이디어는 카메라의 중심에서 특정 복셀까지의 거리가 해당 정점과 관련된 깊이 값보다 큰 경우 복셀에 대한 함수 값이 음수가 되고 그 반대의 경우는 양수가 되는 것이다.

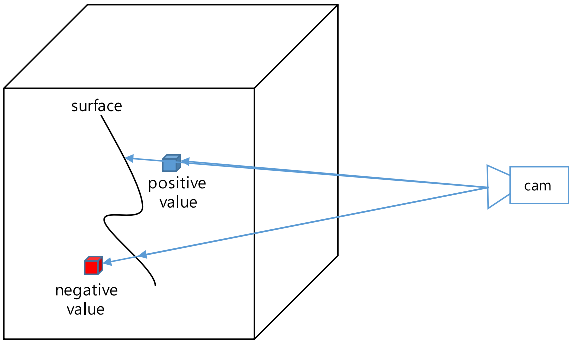


그림 6. 절단된 부호 거리 함수 방법

3.3 색상 카메라 시점으로의 사영

제안하는 방법의 궁극적인 목표는 동일한 해상도로 색상 카메라의 시점에 맞게 정렬된 깊이 영상을 생성하는 것이다. 통합된 체적 모델을 각 색상 카메라 뷰에 투영함으로써 깊이 영상을 생성할 수 있다. 이 단계에서 사전에 얻은 카메라 매개 변수가 활용된다.

먼저 체적 모델이 첫 번째 깊이 카메라에 의존하는 위치에서 생성되기 때문에 모든 카메라 외부 매개 변수는 첫 번째 깊이 카메라를 3 차원 공간에서의 기준($[0,0,0]^{\mathrm{T}}$) 위치로 고려하여 조정된다. 카메라 매개 변수를 조정 한 후에는 각 복셀을 영상 평면에 투영하여 카메라에서 해당 복셀까지의 거리를 기록 할 수 있다.

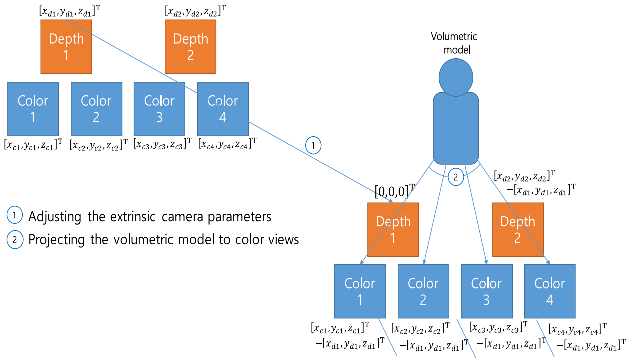


그림 7. 사영을 통한 깊이 영상 생성

4. 실험 결과

앞서 3 절에서 언급한 바와 같이 촬영 시스템의 상부 층에 2 개의 깊이 카메라를 사용하고 하위 층에는 4 개의 색상 카메라를 사용했다. 카메라 모델은 512 × 424 깊이 영상 해상도를 가진 Kinect V2 깊이 카메라와 1280 × 720 색상 해상도를 가진 Basler 색상 카메라를 사용했다. 본 논문에서는 기존 방법과의 호환성을 위해 그림 8 과 같이 깊이 카메라 1 을 색상 카메라 1 과 2 에, 깊이 카메라 2 를 컬러 카메라 3 과 4 에 할당했다.

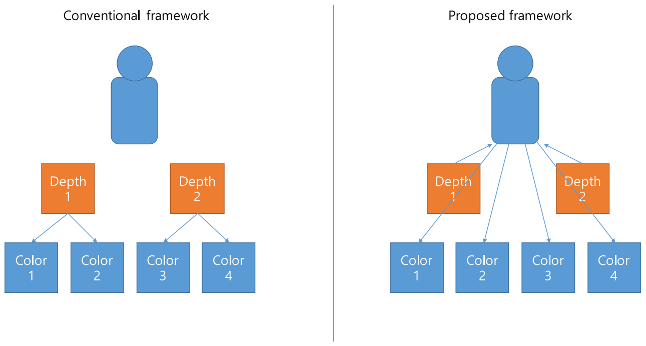
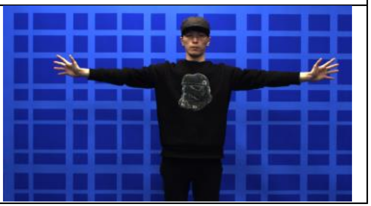


그림 8. 실험에서 사용된 카메라 시스템

표 1 은 촬영된 시퀀스를 보여준다

표 1. 촬영된 시퀀스

Name	Description	Still shot
Person	A single person, 175 frames	

4.1 깊이 영상의 시각적 비교

깊이 영상의 품질을 검증하기 위해, 기존의 방법과 제안하는 방법을 이용하여 깊이 영상을 생성했다. 그런 다음 점군 모델을 생성하고 3 차원 공간에서 시각화했다. 표 2 는 생성된 점군 모델의 시각화를 보여준다.

표 2. 점군 모델의 시각화

Name	Conventional framework	Proposed framework
Person		

기존 방법의 결과는 잡음 또는 이상치 때문에 벽과 사람 사이에서 상당히 지저분한 형태를 보여준다. 이러한 오차는 2 장에서 설명했듯이 3 차원 영상 워핑 이후의 양방향 필터로 인해 발생한다.

그러나 제안 된 프레임 워크는 체적 방식으로 3 차원 공간에서 대부분의 단계를 처리하기 때문에 기존의 방법보다 상대적으로 명확하고 깔끔한 결과를 보여준다.

5. 결론

본 논문에서는 이기종 카메라를 이용하여 3 차원 재구성 시스템을 기반으로 한 다시점 깊이 생성 방법을 제안하였다. 기존 방법에서는 절단된 깊이 영역 문제와 혼합된 깊이 값 문제가 발생하여 정확한 깊이 영상을 계산하지 못했다. 하지만 제안하는 방법에서는 이러한 문제를 해결했을 뿐만 아니라 더 깔끔한 형태의 3 차원 모델을 획득할 수 있었다. 또한 제안하는 방법은 가상 현실(VR) 헤드 마운트 디스플레이, 홀로그래픽 3 차원 디스플레이 등과 같은 최첨단 3 차원 디스플레이에서 생성된 모델을 표현하는데 유리하다.

재구성된 3 차원 모델을 각 색상 카메라 시점으로 투영하는 절차는 간단하지만 임의의 시점에서 깊이 영상을 생성하는 데 이점을 가진다. 제안하는 방법이 가상 현실 및 증강 현실과 같은 최첨단 기술로 사실적인 3 차원 컨텐츠를 획득하는 데 널리 사용되기를 바란다.

감사의 글

본 연구는 미래창조과학부 '범부처 Giga KOREA 사업'의 일환으로 수행하였음. [GK16C0100, 기가급 대용량 양방향 실감 콘텐츠 기술 개발]

참고문헌

[1] Y. Song, D. W. Shin, E. Ko, and Y. S. Ho, "Real-time depth map generation using hybrid multi-view cameras," in *APSIPA*, 2014, pp. 1-4.

[2] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele, "Joint Bilateral Upsampling," in *ACM SIGGRAPH*, 2007, pp. 1-5.

[3] S. Izadi, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, *et al.*, "KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera," in *ACM symposium on User interface software and technology*, 2011, pp. 559-568.

[4] K. L. Low, "Linear least-squares optimization for point-to-plane icp surface registration," 2004.

[5] S. Rusinkiewicz and M. Levoy, "Efficient Variants of the ICP Algorithm," in *International Conference on 3D Digital Imaging and Modeling* 2001, pp. 145-152.

[6] B. Curless and M. Levoy, "A Volumetric Method for Building Complex Models from Range Images," in *ACM SIGGRAPH*, 1996, pp. 303-312.