

ToF 카메라의 특성과 그 한계

홍수민, 호요성
광주과학기술원
{sumin, hoyo}@gist.ac.kr

Specification and Limitation of ToF Cameras

Su-Min Hong, Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)

요 약

요즘 들어, 3차원 콘텐츠의 수요는 지속적으로 증가하고 있다. 3차원 콘텐츠의 품질은 해당 장면의 깊이 정보에 큰 영향을 받기 때문에 정확한 깊이 정보를 얻는 방법이 매우 중요하다. 깊이 정보를 얻는 방법은 크게 수동형 방식과 능동형 방식으로 나뉘는데, 수동형 방식은 계산 과정이 복잡하고 깊이맵의 품질이 보장되지 않는 단점을 갖기 때문에 능동형 방식이 많이 사용되고 있다. 능동형 방식은 깊이 카메라를 이용하여 직접적인 깊이 정보를 얻는 방식으로, 대개 ToF(Time-of-flight) 기술이 사용된다. 이 논문에서는 ToF 깊이 카메라로 촬영된 실제 깊이맵의 특성을 분석하기 위해 여러 가지 촬영 환경과 객체에 대해서 SR4000 깊이 카메라와 키넥트 v2 센서를 이용하여 깊이맵 품질을 비교했다. 실험 결과, 적외선이 제대로 반사되기 어려운 방사성 물질이나 표면, 경계 영역, 어두운 영역, 머리 영역 등에서 정확한 깊이 정보를 얻기 어려웠으며, 실외 환경에서 정확한 깊이 정보가 획득되지 않는 것을 확인할 수 있었다.

1. 서론

최근 들어, 사용자에게 깊이감과 몰입감을 줄 수 있는 3 차원 콘텐츠는 게임, 교육, 문화산업 등 많은 분야에서 각광 받고 있다[1]. 이러한 3 차원 콘텐츠의 품질은 깊이 정보에 큰 영향을 받기 때문에, 정확한 깊이 정보를 획득하고, 그 품질을 개선하는 과정이 매우 중요하다.

깊이 정보를 획득하는 방법은 크게 수동형 방식(passive sensors-based)과 능동형 방식(active sensor-based)으로 나눌 수 있다. 수동형 방식은 널리 알려진 스테레오 정합 방법과 같이 이미 촬영된 양안식 혹은 다시점 색상 영상을 이용하여 깊이 정보를 획득한다. 이러한 방법은 부가적인 장비 없이 색상 영상만을 이용해서 깊이 정보를 계산할 수 있다는 장점이 있지만, 정합 방법에 따라 계산 과정이 매우 복잡하며 품질이 보장되지 않는 단점을 갖는다. 이러한 문제 때문에 깊이 카메라 같은 능동 센서를 이용하는 직접적인 방법이 많이 사용되고 있다.

깊이 카메라는 단어 그대로 실제 환경의 깊이 정보를 획득하여 깊이 영상으로 출력하는 장치로써 깊이 센서(depth sensor)라고도 불린다. 깊이 영상은 각 화소의 깊이 카메라와 물체 간의 거리 정보를 나타내기 때문에, 기존의 수동형 방식과 달리 촬영하는 환경의 깊이 정보를 바로 얻을 수 있다. 이러한 깊이 카메라는 Time-of-flight(ToF)라고 불리는 카메라와 물체 사이의 거리를 적외선이나 광신호를 이용하여 측정하는 기술을 사용한다. 깊이 카메라를 이용한 깊이 정보 획득은 수동형 방식에 비해 정확한 깊이맵을 빠르게 얻을 수 있지만, 기술적 한계로 인해 획득한 깊이맵의 낮은 해상도나 짧은 촬영 범위 등이 문제로 남아 있다.

2. ToF 카메라의 특성

그림 1 은 깊이 카메라의 기본적인 원리를 나타낸다[2]. 그림과 같이 적외선 송출기(Infrared Emitter)에서 나온 적외선 파동(빨간색)은 깊이를 측정하고자 하는 객체에 발사되고 센서 부에서는 객체에 부딪혀 반사된 적외선 파동(파란색)을 감지한다. 이때 생겨나는 두 파동의 위상차를 통해 깊이 카메라는 객체와 카메라 사이의 거리정보를 계산한다.

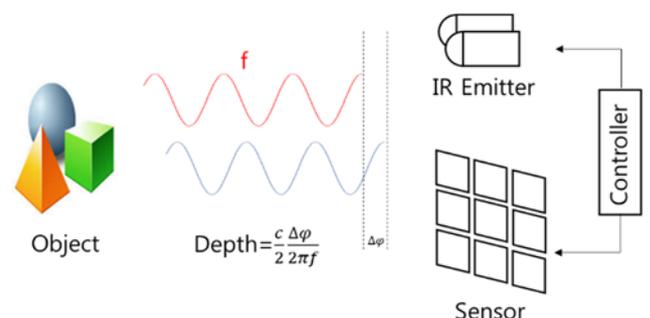


그림 1. 깊이 카메라 원리

두 파동의 위상차를 계산하기 위해서, 각각 90 도 차이의 지연과 서로 다른 전하량을 갖는 4 개의 신호를 사용하게 된다. 깊이 카메라는 이 신호의 집합을 유추함으로써, 두 파동의 위상차를 계산할 수 있다. 식 1 은 위상차 t_d 를 구하는 식을 보여준다.

$$t_d = \arctan\left(\frac{Q_3 - Q_4}{Q_1 - Q_2}\right) \quad (1)$$

여기서 $Q_1 \sim Q_4$ 는 90 도 차이의 지연을 갖는 신호 $C_1 \sim C_4$ 들의 각 전하량을 나타낸다. 최종적으로 깊이값 d 는 식 1 의 위상차 정보를 이용하여 계산되어 진다.

$$d = \frac{c}{2f} \frac{t_d}{2\pi} \quad (2)$$

여기서 c 는 빛의 속도, f 는 파동의 진동수를 나타내며 $\frac{c}{2f}$ 는 최대 측정 거리를 나타낸다. 그림 2 는 위상차를 이용한 깊이맵 계산의 원리를 보여준다.

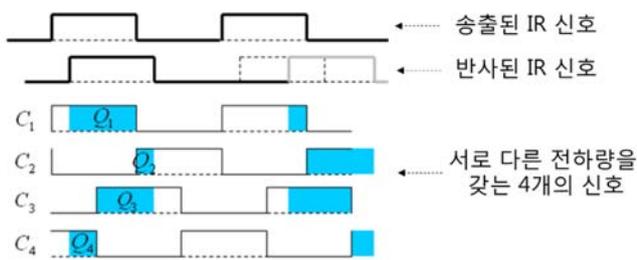


그림 2. 위상차를 이용한 깊이맵 계산

2.1. SR4000

SR4000 은 스위스의 Mesa Imaging 사에서 개발한 대표적인 깊이 카메라 모델이다[3]. 그림 3 은 SR4000 깊이 카메라이며, 크기는 $65 \times 65 \times 68\text{mm}$, 무게는 470g 이다.



그림 3. SR4000 깊이 카메라

SR4000 은 기존 ToF 카메라들에 비해 가볍고 크기가 작으며, CCD/CMOS 프로세스를 이용하여 광학 및 전기적 기능 블록들의 각각 독립적인 최적화를 구현하였다. 표 1 은 SR4000 깊이 카메라의 특성을 보여준다.

표 1. SR4000 의 특성

항목	상세내용
깊이 측정 범위	0.3~5.0m
깊이 정확도	1cm
깊이맵 해상도	176 X 144
최대 프레임률	초당 50프레임
시야각	43.6 X 34.6도
LED Wavelength	850nm
변조 주파수	29, 30, 31MHz

2.2. Kinect

키넥트(Kinect)는 마이크로소프트에서 XBOX-360 게임 디바이스로 출시한 센서이다. 키넥트는 '키네틱(Kinetic)'과 '커넥트(Connect)'의 합성어로 사람의 움직임을 인식하여 컴퓨터 시스템에 연결하는 센서를 의미한다. 2010 년 11 월 출시된 키넥트 v1 은 깊이 정보 획득을 위해 구조광 방식을 사용하였다. 구조광 방식이란, 깊이를 얻고자 하는 객체에 좁은 밴드의 빛을 비추고, 다른 시점에서 볼 때 왜곡으로 나타나는 광 패턴(light pattern)을 분석하여 표면 형태를 복원하는 기술을 말한다[4].

키넥트 카메라는 본래 게임을 위해 개발되었지만, 작은 크기로 인한 설치의 용이성, 저렴한 가격, 깊이맵의 품질 등 여러 가지 장점으로 인해 다양한 3 차원 콘텐츠 제작에 사용되고 있다. 또한 마이크로소프트는 기존의 키넥트 v1 의 구조광 방식을 버리고 ToF 기술을 도입한 키넥트 v2 를 2014 년 7 월 출시하였다[5]. 그림 4 는 키넥트 v1 과 v2 를 보여준다.



그림 4. 키넥트 v1, v2

키넥트 v2 는 기존 버전에 비해 반응속도와 정확도가 빨라졌으며, 기존 버전보다 향상된 기능의 SDK 도 제공된다. 표 2 는 키넥트 v2 의 특성을 보여준다.

표 2. 키넥트 v2 의 특성

항목	상세내용
깊이 측정 범위	0.4~4.5m
색상영상 해상도	1920 X 1080
깊이맵 해상도	512 X 424
최대 프레임률	초당 30프레임
시야각	70 X 60도

3. ToF 카메라의 한계

ToF 카메라는 깊이 정보를 획득하기 위해 ToF 기술을 사용한다. 이 기술은 적외선을 사용하는 기술이기 때문에, 깊이맵을 획득할 때, 빛에 민감한 특성을 보인다. 우선 투명한 유리컵과 같은 빛을 투과하는 물체는 깊이 정보를 정확히 측정하지 못한다. 또한, 객체의 경계나 뾰족한 물체 같은 경우 빛이 제대로 반사되지 못하기 때문에 정확한 깊이 정보를 기대하기 어렵다. 방사성 물질, 광택표면, 머리카락, 가루 등과 같은 물체들도 표면 질감이나 색상 때문에 빛의 왜곡이 일어나 정확한 깊이 정보를 획득하지 못한다.



그림 5. 깊이정보 획득이 어려운 물체

보다 정확한 깊이정보를 획득하기 위해서는 깊이 카메라의 위치도 중요하다. 그림 6 은 깊이 카메라가 잘못된 위치에 설치된 경우를 보여준다. ToF 카메라가 바닥에 너무 가까이 위치하거나, 벽에 가까이 위치한 경우, 카메라에서 나가는 적외선 신호가 바닥이나 벽에 반사되어 잘못된 깊이값으로 인식될 수 있다[5].

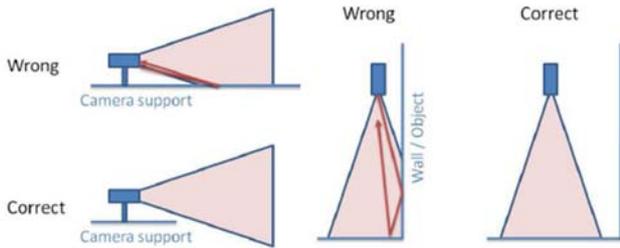


그림 6. ToF 카메라 배치 방법

4. 실험결과

이 논문에서는 ToF 깊이 카메라로 촬영된 실제 깊이맵의 특성을 분석하기 위해 여러 가지 촬영 환경에서 깊이맵 품질을 비교해 보았다. 해당 실험을 위해 실내, 실외 촬영 실험에는 Microsoft 의 키넥트 v2 센서를 실제 객체를 이용한 촬영에는 Mesa Imaging 사의 SR4000 깊이 카메라를 사용하였다.

4.1. 실내, 실외 촬영

ToF 깊이 카메라를 이용하여 깊이정보를 획득할 때, 실외의 촬영 환경에서는 적외선의 원활한 송수신이 불가능하기 때문에 높은 품질의 깊이 정보를 기대하기 어렵다.



그림 7. 실내 촬영 깊이맵

그림 7 은 실내에서 촬영된 색상 영상과 그에 대응하는 깊이맵을 보여준다. 그림에서 볼 수 있듯이, 객체의 경계나 유리 같은 영역에 대해서는 약간의 오류가 있지만, 일정 거리 내에서 깊이 정보가 안정적으로 획득된 것을 확인할 수 있다. 그림 6 은 거리가 먼 경우의 깊이맵을 보여준다.



그림 8. 실내 촬영 깊이맵 (거리가 먼 경우)

그림 8 에서 볼 수 있듯이, 카메라와 촬영하는 영역의 거리가 장비의 촬영 범위를 넘어갈 경우에 깊이 정보가 획득되지 않는 것을 확인할 수 있다. SR4000 깊이 카메라의 경우, 촬영하려는 장면이 촬영 범위를 넘어갈 경우에 그 깊이 정보가 다시 0~5m 의 깊이 정보로 표현된다고 적혀있지만, 실제 촬영 결과 5m 이상의 깊이맵은 사용할 수 없을 정도의 품질을 보였다.

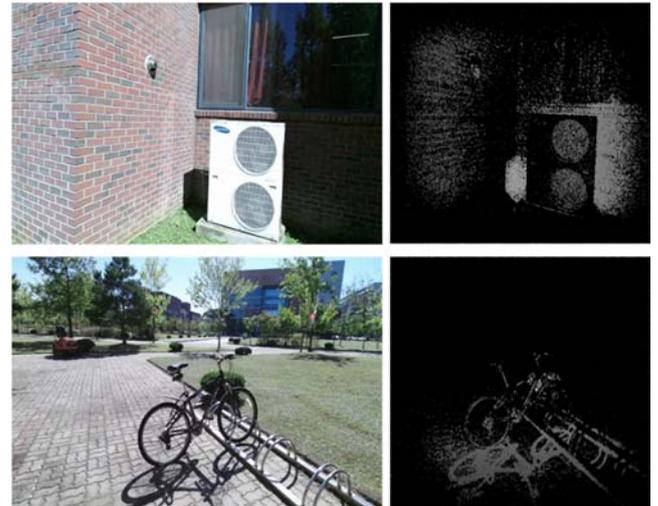


그림 9. 실외 촬영 깊이맵

그림 9 는 실외에서 촬영된 색상 영상과 그에 대응하는 깊이맵을 보여준다. 그림에서 볼 수 있듯이, 강력한 햇빛 아래에서는 적외선 센서가 객체의 표면에 제대로 닿거나 반사되기 어렵기 때문에 깊이 정보를 획득하지 못하는 문제가 발생하였다.

4.2. 실제 객체를 이용한 촬영

또한 실제 객체에 대한 ToF 깊이 정보의 특징을 살펴보기 위해, 실내에서 곰 인형과 사람에 대한 깊이맵을 획득하였다. 해당 실험을 위해, 동일한 장면을 1초 간격으로 10장씩 촬영하였으며, 1초마다 촬영된 깊이맵 10 장의 특정 위치의 깊이값을 그래프로 표현하였다.

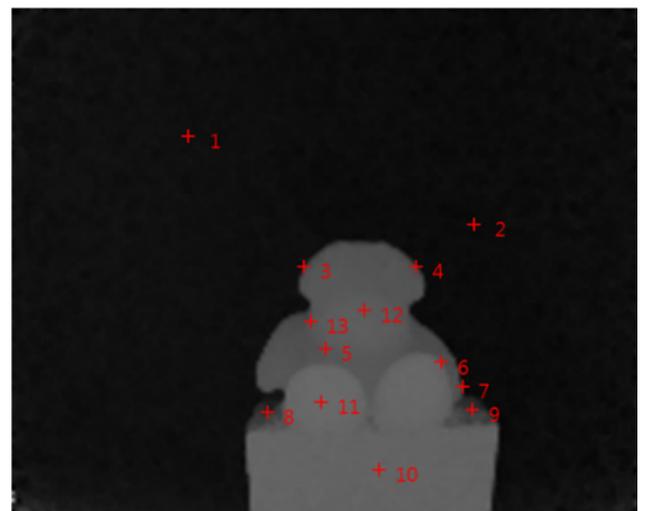


그림 10. 곰 인형 깊이맵

그림 10 은 곰 인형을 촬영한 깊이맵과 임의의 특징점을 보여준다. 그림 11 은 1 초 간격을 두고 촬영된 10 장의 깊이맵에서 임의의 특징점들의 깊이값을 그래프로 표현한 것이다.

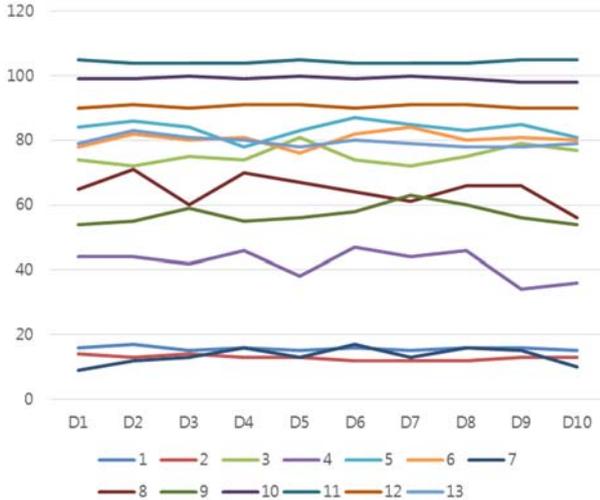


그림 11. 곰 인형의 특징점 깊이값 (10 장)

그림 10 과 그림 11 에서 볼 수 있듯이, 배경을 나타내는 특징점 1, 2 와 같은 경우 깊이값이 안정적이게 획득된 것을 확인할 수 있었다. 또한, 곰 인형의 내부에서의 평탄한 지역인 특징점 10, 11, 12 의 경우에도 깊이값이 안정적이게 획득되었다. 그러나 객체와 배경의 경계를 나타내는 특징점 3, 4, 7 과 객체 내부에서의 경계를 나타내는 5, 6, 13 의 경우 깊이값이 불안정하게 획득되었다. 그리고 특징점 8, 9 와 같이, 상대적으로 적외선이 표면에 반사되기 어려운 곳에 위치해 있던 선반 위는 깊이값이 매우 불안정한 것을 확인할 수 있었다.



그림 12. 사람 깊이맵

그림 12 와 13 은 사람을 촬영한 깊이맵과 그 특징점, 그리고 1 초 간격을 두고 촬영한 10 장의 깊이맵에서 특징점들의 깊이값을 보여준다.

그림에서 볼 수 있듯이, 배경을 나타내는 특징점 1 은 깊이값이 안정적으로 획득되었다. 또한, 객체 내부의 평탄한 지역인 사람의 불과 몸을 나타내는 특징점 3 과 6 의 경우 깊이값이 안

정적으로 획득되는 것을 확인할 수 있었다. 그러나 적외선이 표면에 반사되기 어려운 객체와 배경의 경계 영역인 특징점 9 와 그림자나 어두운색 영역인 특징점 4, 5, 7 의 경우 깊이값이 불안정하게 획득되었다. 또한, 머리 영역을 나타내는 특징점 2 의 경우 깊이값이 매우 불안정하게 획득되었다.

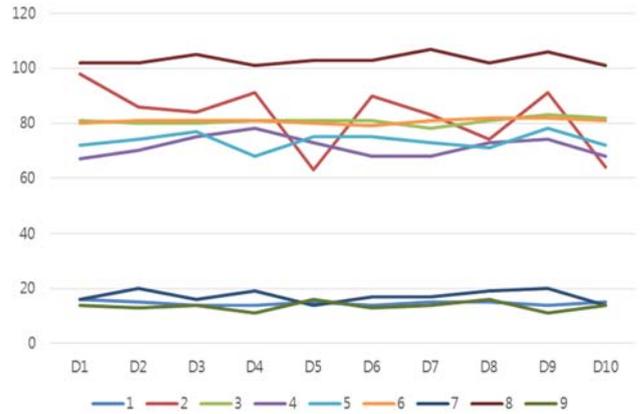


그림 13. 사람의 특징점 깊이값 (10 장)

5. 결론

요즘 들어, 3 차원 콘텐츠의 수요 증가로 인해 여러 관련 분야에서 정확한 깊이 정보의 중요성이 부각되고 있다. 깊이 정보를 획득하는 방법의 하나인 ToF 방식의 깊이 카메라는 기기의 크기가 작고 가격이 저렴해 사용이 증가하고 있지만, 촬영 환경이나 촬영하는 객체의 표면 질감 등에 인해 정확한 깊이 정보를 기대하기 어렵다는 단점을 갖는다. 이 논문에서는 ToF 카메라를 이용하여 실내, 실외의 촬영 환경과 다른 표면 질감을 갖는 객체의 깊이 정보의 품질을 비교하였다. 실험 결과, 실외 환경에서 정확한 깊이 정보를 얻기 어려웠으며, 적외선이 제대로 반사되기 어려운 방사성 물질이나 표면, 경계 영역, 어두운 영역, 머리 영역 등에서 정확한 깊이 정보가 획득되지 않는 것을 확인할 수 있었다.

감사의 글

이 논문은 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No.2011-0030079)

참고 문헌

- [1] C. Fehn, R. Barre, and S. Pastoor, "Interactive 3DTV-concepts and key technologies," Proc. of IEEE, vol. 94, no. 3, pp. 524-538, Mar. 2006.
- [2] M. Hansard, S. Lee, O. Choi, and R. Horaud, Time-of-Flight Cameras: Principles, Methods and Applications. Springer, 2013.
- [3] <http://www.msesa-imaging.ch>
- [4] P. Fechteler and P. Eisert, "Adaptive Color Classification for Structured Light Systems," IET Image Processing, vol. 3, no. 2, pp. 49-59, 2009.
- [5] <https://developer.microsoft.com/enus/windows/kinect>