

# 깊이 카메라를 이용해 획득한 깊이 영상의 경계 오류 제거 방법

이천<sup>o</sup>, 호요성

광주과학기술원 정보통신공학부

leecheon@gist.ac.kr, hoyo@gist.ac.kr

## 요 약

최근 편안한 3차원 영상을 제공하기 위해 다시점 영상과 깊이 영상을 이용한 중간 시점의 영상을 생성하는 기술이 활발히 연구되고 있다. 이 과정에서 깊이 영상의 정확도는 합성 영상의 화질에 직접 영향을 주기 때문에, 정확하고 빠른 깊이 정보를 획득하기 위해 적외선 센서를 이용한 깊이 카메라와 컬러 카메라를 동시에 이용하는 혼합형 카메라 시스템이 이용되고 있다. 하지만 깊이 카메라는 객체의 경계 부근에서 깊이 값을 잘못 측정하여 발생하는 경계 오류가 자주 발생한다. 본 논문에서는 이러한 경계 오류를 보정하기 위해 동일 시점의 컬러 영상과 깊이 영상의 경계 정보를 이용해 경계 오류 화소를 검출하고, 거리 기반 혼합형 다변적 필터(distance-based joint multilateral filter)를 이용해 경계 오류의 깊이값을 재정의한다. 이 방법을 깊이 영상에 적용하여 깊이 영상의 경계를 더욱 뚜렷하게 복원했다.

## 1. 서론

3차원 비디오에 관한 연구 주제는 입체영상을 관람할 때 발생할 수 있는 불편함을 최대한 줄이고 깊이감을 효율적으로 제공하는 방법과 관련이 많다. 이러한 관점에서 DIBR(depth-image-based rendering)을 이용하는 다시점 영상 생성 방법이 많이 이용되고 있다 [1]. 깊이 영상은 각 화소의 3차원 거리 정보를 나타내므로, 이를 이용해 입의 시점의 영상을 생성할 수 있다. 이렇게 생성한 다시점 영상은 스테레오스코픽 재현 장치나 무안경식 다시점 재현 장치(auto-stereoscopic display)에 입력시켜 편안한 3차원 영상을 제공할 수 있다.

3차원 비디오 콘텐츠를 만드는 과정에서 정확한 깊이 정보를 얻는 것은 매우 중요하다. 깊이 영상을 획득하는 방법은 스테레오 정합과 같은 알고리즘을 이용해 획득하는 방법[2]과 깊이 카메라와 컬러 카메라를 혼용하는 방법[3]으로 구분할 수 있다. 스테레오 정합을 이용한 방법은 장면의 구성에 대한 제약이 적은 반면, 복잡도가 높고 정확도가 떨어진다. 반면, 깊이 카메라를 이용하는 방법은 정확도가 높은 깊이 정보를 실시간으로 획득할 수 있으나, 제공하는 해상도가 낮다. 이러한 단점을 보완하기 위해 깊이 영상을 고해상도로 변환하는 방법도 연구되고 있다 [4].

혼합형 카메라 시스템은 한 대의 깊이 카메라와 두 대 이상의 컬러 카메라를 이용해 각 컬러 카메라 위치의 깊이 영상을 획득하는 것이 목적이다. 이러한 시스템을 이용해 고해상도의 깊이 영상을 생성하기 위해서는 깊이 영상에 일반적인 업샘플링(upsampling) 방법을 적용하지 않고, 컬러 영상을 이용해 수행한다. 먼저 깊이 카메라를 이용해 획득한 깊이 영상을 각 컬러 카메라 위치로 시점 이동한 후, 컬러 영상과 시점 이동한 깊이 영상을 동시에

이용해 고해상도의 깊이 영상을 획득한다. Zhu[3]와 Lee[5]는 고해상도의 깊이 영상을 획득하기 위해 시점 이동한 깊이 영상을 스테레오 정합의 초기 조건으로 이용하는 방법을 제안했다. 하지만 이러한 방법들은 복잡도가 높아 실시간으로 깊이 영상을 획득하기 어렵다는 단점이 있다.

본 논문에서는 깊이 카메라를 이용해 고해상도의 깊이 데이터를 실시간으로 획득하기 위한 보정 방법을 제안한다. 깊이 카메라는 적외선을 이용해 화소의 거리를 측정하는데, 객체의 경계 부근에서는 배경 객체에서 반사되어 오는 빛과 전경 객체에서 반사되어 오는 빛이 혼합되어 객체의 경계 부근에서 깊이값이 부정확하게 측정된다. 따라서 깊이 영상의 객체 경계가 컬러 영상의 객체 경계와 일치하지 않는 문제가 발생한다. 이러한 문제를 깊이 영상의 경계 오류라고 지칭하고, 이 문제를 해결하기 위한 방법을 제안한다.

## 2. 깊이 카메라를 이용한 깊이 영상 획득

### 2.1 혼합형 카메라 시스템

혼합형 카메라 시스템은 그림 1과 같이 다시점 영상을 생성을 위한 입력 데이터를 획득하기 위해 고해상도의 컬러 카메라와 저해상도의 깊이 카메라를 이용한다. 양쪽에 위치한 컬러 카메라는 고해상도(1280x960)의 컬러 영상을 획득한다. 그리고 중앙에 위치한 깊이 카메라는 상대적으로 낮은 해상도(176x144)의 깊이 영상을 획득한다. 혼합형 카메라 시스템의 목표는 기하학적 왜곡이 없는 두 시점의 컬러 영상과 동일 시점에서 동일 해상도의 깊이 영상을 획득하는 것이다. 이렇게 획득한 두 시점의 컬러 비디오와 깊이 비디오를 다시점 영상 생성의 입력 데이터로 이용한다.



그림 1. 혼합형 카메라 시스템

그림 2는 고해상도의 깊이 영상을 획득하기 위한 혼합형 카메라 시스템의 블록 다이어그램을 보인 것이다. 중간 시점에서 획득한 깊이 영상은 3차원 워핑(3D warping)을 이용해 좌/우 시점으로 각각 시점 이동한다. 시점 이동으로 발생하는 빈 공간은 홀 채움(hole filling) 기술을 이용해 채운다. 홀 채움을 수행한 깊이 영상은 객체의 경계가 컬러 영상의 객체 경계와 일치하지 않는 문제를 갖는다. 이 문제를 본 논문에서 제안하는 경계 오류 제거 방법을 이용해 해결한다. 이 단계까지 수행한 깊이 영상은 여전히 낮은 해상도이다. 이 깊이 영상을 컬러 영상의 해상도와 같이 만들기 위해서 깊이 영상을 업샘플링(upsampling)한다.

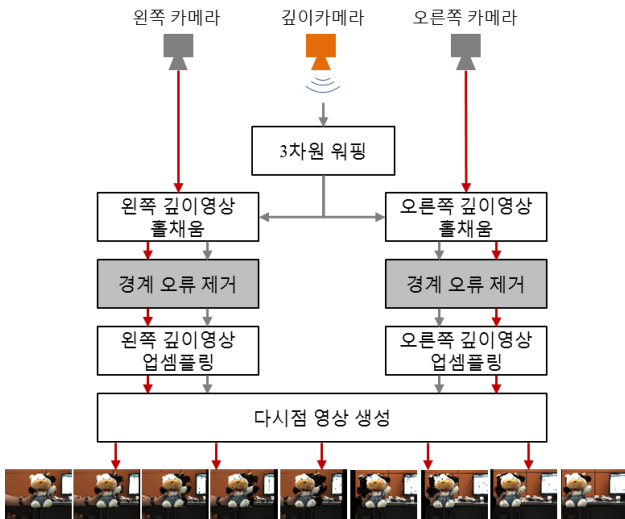


그림 2. 혼합형 카메라 시스템의 블록도

이렇게 획득한 데이터는 두 시점의 고해상도 컬러 영상과 두 시점의 깊이 영상이다. 이 데이터를 중간 시점 영상 생성 방법에 적용하면, 그림 2와 같이 다시점의 영상을 생성할 수 있다. 깊이 영상의 정확도가 높을수록 생성한 중간시점의 영상의 화질이 높아진다.

### 2.2 깊이 영상의 경계 오류

앞서 언급한 것과 같이, 깊이 카메라는 객체의 경계에서 부정확한 깊이값을 가질 수 있다. 깊이 카메라는 적외선을 주사하고 객체에 반사되어 돌아오는 광선의 지연 시간을 이용해 객체의 거리를 계산한다. 이때 객체의 경계 부근에서는 배경 객체에서 반사된 광선과 전경 객체에서 반사된 광선이 혼합

되어 부정확한 깊이값이 측정된다. 특히, 객체의 표면이 Lambertian 표면이 아닌 난반사가 일어나는 표면일 경우에 이러한 경계 오류가 두드러지게 나타난다. 그림 3은 이러한 경계 오류를 단적으로 보여준다. 그림 3(b)을 잘 살펴 보면, 붉은색 사각형으로 표시한 물체의 경계에서 불안정한 값을 갖는 것을 확인할 수 있다.



(a) 컬러 영상 (b) 획득한 깊이 영상  
그림 3. 깊이 영상의 경계 오류

### 2.3 기존의 깊이 영상 필터링 방법

깊이 영상의 경계를 컬러 영상의 경계와 일치하도록 보정하는 방법은 여러 가지가 있다. 그 중 대표적으로 JBF(joint bilateral filter)[6]와 JMF(joint multilateral filter)[7]를 들 수 있다. JBF는 식 (1)과 같이 보정할 화소  $p$ 를 중심으로 하는 커널  $\Omega$  내의 참조 화소  $D(q)$ 에 각각 가중치를 주어 새로운 값  $D^{new}$ 를 결정한다.

$$D^{new}(p) = \frac{\sum_{q \in \Omega} \omega_{JBF}(p, q) \cdot D(q)}{\sum_{q \in \Omega} \omega_{JBF}(p, q)} \quad (1)$$

여기서  $\omega_{JBF}$ 는  $p$ 와  $q$ 의 화소간 거리와 컬러값을 이용해 결정하는 가중치 함수이며, 이는 다음과 같다.

$$\omega_{JBF}(p, q) = \exp\left(-\frac{\|I(p) - I(q)\|^2}{2\sigma_c^2}\right) \cdot \exp\left(-\frac{\|p - q\|^2}{2\sigma_r^2}\right) \quad (2)$$

여기서  $I(p)$ 와  $I(q)$ 는  $p$ 와  $q$ 의 컬러 화소값을 나타낸다. 그리고  $\sigma_c$ 와  $\sigma_r$ 은 각각 컬러의 유사도와 거리 향에 사용되는 표준편차를 나타낸다. JBF를 사용하면 깊이 영상에 포함되어 있는 잡음을 효과적으로 제거하고 객체의 경계를 더욱 선명하게 보정한다. 이 방법의 한계는 깊이 영상의 값을 컬러 화소의 유사도를 위주로 결정하기 때문에 깊이 영상의 객체 경계가 제대로 보정되지 못한다.

이러한 문제를 보완하는 방법으로 JMF는 식 (3)과 같이 JBF에서 깊이 영상의 유사도를 추가적으로 고려한다.

$$\omega_{JMF}(p, q) = \exp\left(-\frac{\|D(p) - D(q)\|^2}{2\sigma_d^2}\right) \cdot \exp\left(-\frac{\|I(p) - I(q)\|^2}{2\sigma_c^2}\right) \cdot \exp\left(-\frac{\|p - q\|^2}{2\sigma_r^2}\right) \quad (3)$$

여기서  $D(p)$ 와  $D(q)$ 는  $p$ 와  $q$ 의 깊이값을 나타내고,  $\sigma_d$ 는 깊이값 유사도 향의 표준편차를 나타낸다. 이

JMF를 이용하면 깊이 영상의 객체 경계를 컬러 영상의 객체 경계에 맞추어 깊이 영상을 보정할 수 있다. 하지만 이 방법도 객체 경계의 불일치가 클 경우 불안정한 깊이값을 갖는다.

### 3. 깊이 영상 경계 오류 제거 방법

깊이 카메라를 이용해 획득한 깊이 영상은 경계 오류를 포함한다. 이러한 경계 오류를 제거하기 위해 그림 4와 같은 경계 오류 제거 방법을 제안한다.

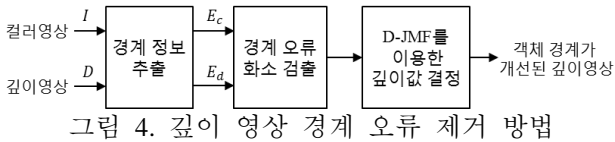


그림 4. 깊이 영상 경계 오류 제거 방법

#### 3.1 경계 영상 추출

본 논문에서 제안하는 경계 오류 제거 방법의 처음 과정으로 컬러 영상과 깊이 영상의 경계 정보를 추출한다. 깊이 영상의 경계 오류는 주로 객체의 경계에 위치하기 때문에 객체 경계 주변의 화소만 관심 화소로 지정한다. 제안하는 방법에서는 Canny 경계 검출 알고리즘을 이용한다. 경계 정보를 정확하게 추출하기 위해서는 각 영상에 최적화된 변수를 이용해야 하지만, 깊이값의 변화가 작은 영역도 경계 정보를 검출할 수 있는 변수값을 이용한다.

#### 3.2 경계 오류 탐색

경계 오류가 있는 화소를 탐색하기 위해 앞에서 획득한 경계 정보를 이용한다. 그림 5는 경계 정보를 이용하여 경계 오류를 검출하는 방법을 설명한 것이다. 컬러 영상의 경계 화소를 중심 화소  $p$ 로 하고, 이 화소를 중심으로 일정한 크기의 커널 내의 깊이 영상의 경계 정보를 탐색한다. 그 중  $p$ 와 가장 가까운 경계 화소  $q$ 를 선택한다. 그리고 두 화소  $p$ 와  $q$ 의 위치가 동일할 경우는 경계 정보가 일치하므로 경계 오류 화소가 검출되지 않고,  $p$ 와  $q$ 가 일치하지 않을 경우엔 그 중간의 화소를 경계 오류 화소  $m$ 이라고 결정한다.

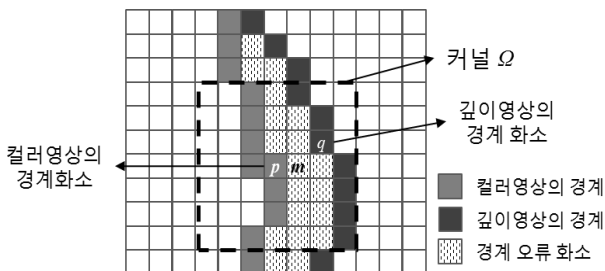


그림 5. 경계 정보를 이용한 경계 오류 탐색

이런 방법으로 컬러 영상의 모든 경계 화소에 대해서 오류 화소  $m$ 을 검출한다. 만약 컬러 영상의

경계 화소 주변에 경계 화소가 없을 경우엔 오류 화소가 없는 것으로 간주한다. 검출한 경계 오류 화소를 깊이 영상에서 제거하고, 다음에 제안하는 방법으로 깊이값을 다시 결정한다.

### 3.3 D-JMF 를 이용한 경계 보정

앞에서 제거한 경계 오류 화소의 깊이값을 재정의하기 위해 식 (4)와 같이 경계에서 가까울수록 낮은 가중치를 할당하고, 거리에서 멀수록 높은 가중치를 할당하는 거리 정보를 이용한다. 경계 오류가 깊이 영상의 경계에 가까울수록 자주 발생하기 때문에 깊이 영상의 신뢰도를 반영한다.

$$\mu_{p,q} = \begin{cases} 0 & \text{if } q = m \\ \frac{\|p - q\|}{K} & \text{otherwise} \end{cases} \quad (4)$$

여기서  $K$ 는  $p$ 와  $q$ 가 가장 멀리 떨어질 경우의 상수값으로 커널의 크기에 따라 결정된다.

경계 오류 화소의 깊이값 결정은 식 (5)와 같이 거리기반 가중치 정보를 이용한다.

$$D^{new}(p) = \begin{cases} \sum_{q \in \Omega} \omega_{MP}(p, q) \cdot D(q) / \sum_{q \in \Omega} \omega_{MP}(p, q) & \text{if } p = m \\ D(p) & \text{otherwise} \end{cases} \quad (5)$$

여기서 깊이값의 재정의는 앞에서 결정한 경계 오류 화소에 대해서만 수행한다. 각 참조 화소  $D(q)$ 에 부여하는 가중치는 다음의 식 (6)과 같이 거리기반 가중치를 이용한다.

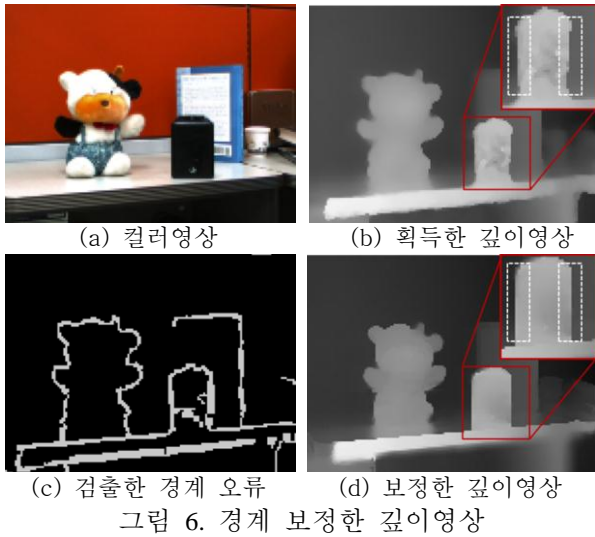
$$\omega_{MP}(p, q) = \exp\left(-\frac{\|p - q\|^2}{2\sigma_r^2}\right) \cdot \exp\left(-\frac{\|I(p) - I(q)\|^2}{2\sigma_c^2}\right) \cdot \mu_{p,q} \exp\left(-\frac{\|D(p) - D(q)\|^2}{2\sigma_d^2}\right) \quad (6)$$

여기서 가중치  $\mu_{p,q}$ 가 1에 가까울수록 신뢰도가 높으므로 세번째 항의 중요도가 높아진다. 동시에 깊이 영상 차이값의 크기가 작아지므로  $\omega_{MP}$ 의 값이 1에 가까워진다. 결국 이 참조 화소  $D(q)$ 에 가중치를 크게 할당하게 된다.

## 4. 실험 결과

본 논문에서 제안한 방법의 성능을 평가하기 위해 혼합형 카메라 시스템을 통해 획득한 깊이 영상과 스테레오 정합기술 개발에 이용되는 테스트 데이터를 이용했다. 이 과정에서 커널의 크기는 11x11이고,  $\sigma_r$ 는 0.5,  $\sigma_c$ 와  $\sigma_d$ 는 각각 0.1로 설정했다.

그림 6은 혼합형 카메라 시스템을 이용해 획득한 깊이 영상에 제안한 방법을 적용한 결과를 보인 것이다. 그림 6(a)와 그림 6(b)는 그림 3에 보인 것과 같은 영상이고, 그림 6(c)는 제안한 방법을 이용해 검출한 경계 오류 화소를 표시한 것이다. 그림

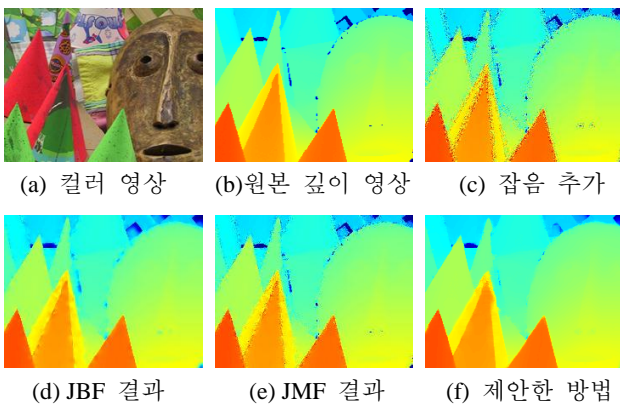


(a) 컬러영상 (b) 획득한 깊이영상  
(c) 검출한 경계 오류 (d) 보정한 깊이영상  
그림 6. 경계 보정한 깊이영상

6(d)는 경계 오류 정보와 D-JMF를 이용해 보정한 깊이 영상을 보인 것이다. 그림 6(b)와 그림 6(d)를 비교했을 때 붉은색 사각형으로 표시한 영역의 경계 정보가 정확하게 보정되었고, 인형의 경계 정보도 더욱 선명해졌음을 알 수 있다.

추가적으로 객관적인 성능 평가를 위해 스테레오 정합에 이용되는 테스트 영상을 이용했다. 그림 7(b)와 같은 원본 깊이 영상의 경계에 인위적으로 Gaussian 잡음을 추가하여 그림 7(c)와 같이 객체의 경계가 왜곡된 깊이 영상을 획득했다. 그리고 기존의 방법과 제안한 방법을 왜곡된 깊이 영상에 적용하여 객체의 경계가 얼마나 잘 복원되는지 살펴보았다. 그림 7(d)와 그림 7(e)는 JBF와 JMF를 이용한 복원 결과이고, 그림 7(f)는 제안한 방법을 이용해 복원한 결과이다. 그림 7에 보인 것처럼, 제안한 방법이 경계를 보다 선명하게 복원했다.

객관적인 성능 평가를 위해 총 12개의 테스트 영상을 이용해 복원한 깊이 영상을 원본 깊이 영상과 비교하여 불량 화소 발생률(bad pixel rate)을 살펴 보았다. 그림 8에 보인 것과 같이, 본 논문에서 제안한 방법은 기존의 방법보다 불량 화소 발생률이 낮게 나타났다. 이로서 제안한 방법이 깊이 영상의 경계를 효과적으로 복원함을 알 수 있었다.



(a) 컬러 영상 (b)원본 깊이 영상 (c) 잡음 추가  
(d) JBF 결과 (e) JMF 결과 (f) 제안한 방법  
그림 7. 깊이 영상 보정 결과 비교

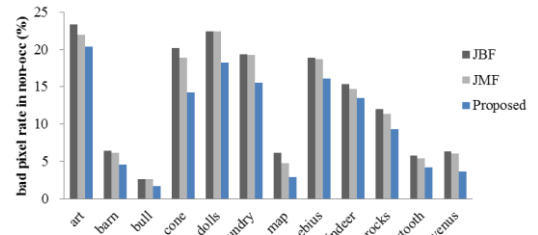


그림 8. 보정한 깊이 영상의 불량 화소 발생률 비교

### 5. 결론

본 논문에서는 깊이 카메라를 통해 획득한 깊이 영상의 객체 경계에 경계 오류를 보정 방법을 제안했다. 컬러 영상과 깊이 영상의 경계 정보를 이용해 경계 오류 화소를 검출한다. 그리고 거리 기반으로 정의한 신뢰도가 높은 깊이값을 이용하므로 보다 정확한 깊이값을 결정할 수 있었다. 이 방법을 이용하면 컬러 영상과 깊이 영상의 객체의 불일치 문제를 효과적으로 해결할 수 있었다.

### 감사의 글

본 연구는 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NO. 2012-0009228).

### 참고문헌

- [1] C. Fehn, "Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV," *Stereoscopic Displays and Virtual Reality Systems XI*, pp. 93-104, Jan. 2004.
- [2] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-frame Stereo Correspondence Algorithms," *International Journal of Computer Vision*, vol. 47, pp. 7-42, April 2002.
- [3] J. Zhu, L. Wang, R. Yang, J.E. Davis, and Z. Pan, "Reliability fusion of time-of-flight depth and stereo geometry for high quality depth maps," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 1400-1414, 2011.
- [4] C. Lee, H. Song, B. Choi, and Y.S. Ho, "3D Scene Capturing using Stereoscopic Cameras and a Time-of-Flight Camera," *IEEE Transactions on Consumer Electronics*, vol. 57, pp. 1370-1376, Aug. 2011.
- [5] E.K. Lee and Y.S. Ho, "Generation of Multi-view Video using a Fusion Camera System for 3D Displays," *IEEE Transactions on Consumer Electronics*, vol. 56, pp. 2797-2805, Nov. 2010.
- [6] O.P. Gangwal and R.P. Berretty, "Depth Map Post-processing for 3D-TV," *IEEE International Conference on Consumer Electronics*, Jan. 2009.
- [7] A.K. Riemens, O.P. Gangwal, B. Barenbrug, and R.P. Berretty, "Multi-step Joint Bilateral Depth Upsampling," *Visual Communications and Image Processing*, pp. 1-12, Jan. 2009.