

## 시점 간 일관성을 고려한 다시점 깊이 영상의 해상도 개선

최진욱 손광훈

연세대학교 전기전자공학과

khsohn@yonsei.ac.kr

## Multi-view Depth Image Resolution Enhancement based on View Coherence

Choi, Jin-Wook Sohn, Kwang-Hoon

Department of Electrical and Electronic Engineering

Yonsei University

## 요약

최근 들어, Time-of-Flight (TOF) 방식의 깊이 센서가 컴퓨터 비전이나 영상처리분야 등의 연구에서 다양하게 활용되고 있다. TOF 방식의 깊이 센서는 실시간으로 정확한 깊이 정보를 획득할 수 있는 장점이 있는 반면, 해상도가 낮고 센서의 광학적 특성으로 인해 노이즈가 발생한다는 단점이 있다. 본 논문에서는 이러한 깊이 센서의 단점을 극복하고 고품질의 다시점 영상 합성을 위해 기존 깊이 센서에 여러 대의 Charge-coupled Device (CCD) 카메라를 결합하여 시점 간 일관성 있는 고해상도의 깊이 영상을 획득하는 기술을 제안한다. 한정되어 있는 시점의 영상을 이용하여 다시점 영상을 합성할 때, 각 시점에 해당하는 깊이 영상의 정확도에 따라 합성되는 영상의 품질이 결정된다. 따라서 보유 시점에 해당하는 깊이 영상 간의 일관성이 유지되고 객체의 경계가 정확히 보존되어야만 생성되는 임의 시점 영상의 품질이 보장될 수 있다.

이에, 본 논문에서는 Joint Bilateral Upsampling (JBU)을 기반으로 시점 간 일관성을 고려한 깊이 영상의 해상도 개선 기술을 제안하고 이를 통해 다시점 영상 합성 결과의 품질을 향상시킨다.

## 1. 서론

영상 혹은 객체에 대한 깊이 정보를 획득하기 위한 방법은 크게 능동적 (Active) 방식과 수동적 (Passive) 방식으로 분류할 수 있다. 우선, 능동적 방식은 레이저를 이용하여 장면 혹은 물체를 스캔하는 3차원 레이저 스캔 방식과 적외선 파장을 이용하여 물체에 반사되어 돌아오는 시간을 거리로 환산하는 Time-of-Flight (TOF) 센서 방식 등을 예로 들 수 있다. 수동적 방식의 대표적인 예로는 두 대 이상의 카메라로부터 획득한 영상을 이용하는 스테레오 매칭 (Stereo Matching) 방식이 있다. 3차원 레이저 스캔 방식은 레이저를 이용하여 물체를 한 점 혹은 한 줄씩 정밀하게 읽어 들여서 3차원 형상을 복원하여 깊이 정보를 얻어내는 방식으로 매우 정교하지만 가격이 비싸고 실시간으로 활용이 불가능하다. 스테레오 매칭 방식은 여러 대의 카메라로부터 획득한 영상들을 이용하여 비용 함수 (Cost function)를 계산하고 최적화하여 번이 정보를 추정하고 이를 바탕으로 깊이 정보로 환산하는 방식으로 텍스처가 없거나 반복되는 영역, 가려진 영역 등에서는 정확도가 매우 떨어지고 성능이 뛰어난 알고리즘의 경우 시간이 오래 걸린다는 단점이 있다. TOF 센서는 낮은 해상도의 깊이 영상을 제공하고 반사가 심한 물체 혹은 광량이 풍부한 야외 등에서 노이즈가 발생한다. 그러나 비교적 정확한 깊이 영상을 실시간으로 획득할 수 있어서 컴퓨터 비전이나 영상처리분야 등의 연구에서 다양하게 활용되고 있다. 본 논문에서는 TOF 센서를 여러 대의 고해상도의 CCD 카메라

와 결합하여 해당 시점의 깊이 영상 해상도를 개선하고 각 시점 간의 일관성을 유지시킴으로써 다시점 영상 합성의 품질을 향상시키는 방법을 제안한다. 향후, 무안경식 다시점 3D 기술로 발전해감에 따라 제안 방법은 고품질의 3D 콘텐츠를 제작할 수 있는 기술로써 활용 가능할 것이다.

## 2. 본론

CCD 카메라는 일반적으로 고해상도의 영상을 제공하므로 상대적으로 낮은 해상도를 지원하는 TOF 방식의 깊이 센서의 단점을 보완할 수 있다. 따라서 깊이 영상의 품질을 향상시키기 위하여 이중 센서를 결합하여 사용하는 연구가 활발하게 진행되고 있다. 스테레오 카메라와 깊이 센서를 결합하여 수동적 방식으로 추정된 깊이 정보의 정확도를 높이기 위해 능동적 방식인 TOF 센서로부터 획득한 깊이 정보를 부가적으로 활용하는 방법이 제안되었다 [1]. 또한, 초해상도 개념을 이용하여 깊이 영상의 해상도를 향상시키는 방법도 제안되었다 [2]. 이러한 방식과 달리 깊이 센서에서 획득한 정보를 주 정보로 활용하여 한 쌍의 CCD 카메라와 깊이 센서를 매우 가까이 붙여 놓아 호모그래피 행렬을 통해 깊이 정보를 이동시켜 해상도를 개선하는 방법도 제안되었다 [3]. 본 논문에서는 [3]과 같이 깊이 센서의 정보를 주 정보로 활용하여 해상도를 개선하지만 여러 대의 CCD 카메라를 활용하여 카메라 캘리브레이션을 통해 각 센서의 좌표계를 통합하고 각 시점 간의

일관성을 유지시킬 수 있는 다시점 깊이 영상의 해상도 개선 기술을 제안한다.

우선, 컬러 영상 간의 색상 불일치를 해결하기 위해 히스토그램 매칭을 수행한다. 색상 불일치가 심할 경우, 임의의 시점 영상 합성이나 깊이 영상의 해상도 상황변환 결과의 품질에 악영향을 미치므로 타겟 영상의 히스토그램을 기준 영상의 히스토그램으로 매칭시키는 과정을 전처리로 수행한다. 각 센서의 위치 정보를 계산하기 위해 카메라 캘리브레이션(Camera Calibration)을 수행하고 계산된 왜곡 보정 계수를 통해 영상의 원주 방향 왜곡을 보정한다 [4]. 보정된 각 센서의 좌표계를 하나의 통합된 좌표계로 정보를 통합시킨다. 이러한 과정을 통해 깊이 센서의 정보를 각각의 CCD 카메라로 이동시킬 수 있고, 이동된 깊이 정보와 컬러 영상을 기반으로 다음의 Joint Bilateral Upsampling (JBU) [5] 방식을 활용하여 해상도를 개선할 수 있다.

$$R_p = \frac{1}{k_p} \sum_{q \downarrow \in k} R_{p_i} f(\|p \downarrow - q \downarrow\|) g(\|\bar{I}_p - \bar{I}_q\|) \quad (1)$$

식 (1) 과 같이 고해상도 영상  $\bar{I}_p$  와 저해상도 영상  $I_{p_i}$  이 주어졌을 때 저해상도 영상에서의 공간 함수인  $f(\|p \downarrow - q \downarrow\|)$  (Domain kernel) 와 고해상도 영상에서 각 픽셀의 화소 값 차이를 나타내는  $g(\|\bar{I}_p - \bar{I}_q\|)$  함수 (Range kernel)를 이용하여 저해상도에서의 정보  $R_{p_i}$  를 고해상도의 정보  $R_p$  로 상황변환해 주는 것이 JBU 이다. 각 함수는 주로 가우시안의 형태를 사용한다. JBU를 반복 수행하게 되면 경계부분은 더욱 확실히 부각되고 평활한 영역은 매우 부드럽게 필터링된다. 그러나 그림 1에서 보듯이 반복 수행의 결과로 동일한 깊이 영역에서 컬러 영상의 세부 텍스처가 표현됨으로써 깊이 영상의 품질을 저해하고 시점 간 일관성이 유지되지 않는다는 문제가 발생한다. 이를 해결하기 위해 다음 식 (2)와 같이 픽셀별 선택적 반복 수행을 적용함으로써 시점 간 일관성이 유지되는 고품질 깊이 영상을 획득할 수 있다.

$$R_p^{(t+1)} = \frac{1}{k_p} \sum_{q \downarrow \in k} R_{p_i}^{(t)} f(\|p \downarrow - q \downarrow\|) (w_1 g(\|\bar{I}_{L,p} - \bar{I}_{L,q}\|) + w_2 g(\|\bar{I}_{R,p} - \bar{I}_{R,q}\|)) \quad (2)$$

$$\begin{cases} R_p^{(t+1)} = R_p^{(t)} & , |R_p^{(t+1)} - R_p^{(t)}| \geq \text{threshold} \\ R_p^{(t+1)} = R_p^{(t+1)} & , |R_p^{(t+1)} - R_p^{(t)}| \leq \text{threshold} \end{cases}$$

식 (2) 에서 Range kernel 은 각 시점에서의 컬러 영상 정보를 모두 포함한다. 본 논문에서는 두 개의 시점을 사용하므로 좌, 우영상으로 Range kernel을 구성하였다. 사용가능한 컬러 영상 정보를 모두 활용하여 시점 간 일관성을 유지시킬 수 있다. 변수  $t$ 는 반복 수행 횟수를 의미한다. 반복 수행한 결과와 이전 결과 값의 차이가 한계치를 넘는다면 그 픽셀에서의 값은 잘못 필터링 된 값, 즉 세부 텍스처가 표현되는 영역으로 판단하여 이전 결과의 값을 다음 반복 수행에서 그대로 사용한다. 한계치를 넘지 않는다면, 경계를 보존하고 평활한 영역을 더욱 부드럽게 하는 JBU 의 성능에 맞게 필터링 된 값이라고 판단하여 다음 반복 수행에서 필터링 값을 사용한다. 이러한 픽셀별 선택적 반복 수행을 통해 각 시점 간의 일관성이 유지되는 고해상도 깊이 영상을 획득할 수 있다.



그림 1. 좌, 우 영상에 해당되는 시점 간 일관성이 고려되지 않은 해상도 상황변환 된 깊이 영상

### 3. 구현 및 결과

본 논문에서는 그림 2와 같이 2 대의 CCD 카메라와 1 대의 깊이 센서를 사용하여 3D 콘텐츠 획득 시스템을 구성하였다. 사용된 CCD 카메라는 Point Grey Research Inc. 의 1024X768 해상도를 지원하는 Flea 모델을 사용하였고 깊이 센서는 MesaImaging 사의 176X144 해상도를 지원하는 SR4000 모델을 사용하였다. 그림 3과 4는 시점 간 일관성을 고려한 고품질 깊이 영상과 그 정보를 바탕으로 중간 시점 영상 합성한 결과를 나타낸다.

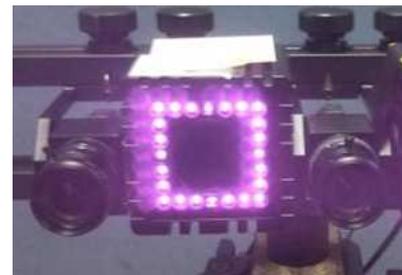


그림 2. 3D 콘텐츠 획득 시스템의 구성

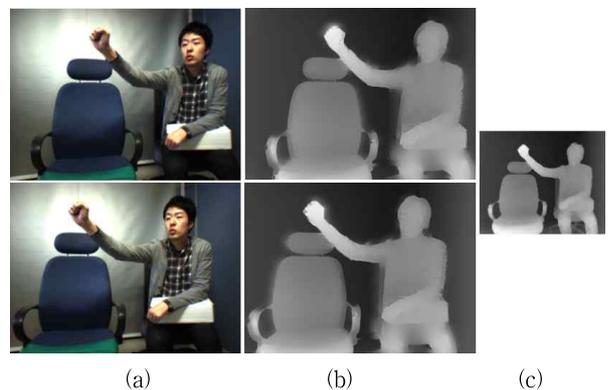


그림 3. (a) Flea 카메라로 획득한 좌,우 컬러 영상, (b) 시점 간 일관성을 고려한 해상도 상황변환 결과, (c) 깊이 센서로 획득한 초기 깊이 정보



그림 4. 해상도 상향 변환 된 깊이 영상을 이용한 중간 시점 합성 결과

#### 4. 결론 및 향후 연구 방향

영상신호처리, 컴퓨터 비전 등 다양한 연구 분야에서 활용되고 있는 TOF 방식의 깊이 센서는 3D 콘텐츠로 사용하기에는 해상도가 낮다는 치명적인 단점이 있다. 제안 방식을 통해 깊이 영상의 해상도를 개선하였고 더 나아가 시점 간 깊이 정보의 일관성을 유지시킴으로써 다시점 영상 합성 시 성능을 향상시켰으며 이는 무안경식 3D 콘텐츠 제작 등에서 적용될 수 있다. 향후에는 센서의 수를 늘려 다시점-다깊이 구조에서의 깊이 정보 신뢰도 판단 기술과 고품질 임의 시점 영상 합성 기술에 대한 연구를 통해 고품질 3D 콘텐츠 획득에 대한 연구를 진행할 수 있을 것이다.

#### 참고문헌

- [1] J. Zhu, L. Wang, R. Yang, J. Davis, "Fusion of Time-of-Flight Depth and Stereo for High Accuracy Depth Maps," *IEEE Proc. CVPR*, 2008.
- [2] S. Schuon, C. Theobalt, J. Davis, and S. Thrun. "High-quality scanning using time-offlight depth superresolution," *IEEE CVPR Workshop on Time-of-Flight Computer Vision*, 2008.
- [3] J. Choi, D. Min, B. Ham and K. Sohn, "Spatial and temporal up-conversion technique for depth video," *in Proc. IEEE Conf Image Processing*, pp. 3525-3528, 2009.
- [4] <http://www.vision.caltech.edu/bouguetj>, Camera Calibration Toolbox for MATLAB.
- [5] J. Kopf, MF. Cohen, D. Lischinski and M. Uyttendaele, "Joint bilateral upsampling," *ACM SIGGRAPH*, 2007.