

# Low-textured 영역에 강한 세그먼트 기반의 다시점 스테레오

\*박해솔 \*\*이경무 \*\*\*이상욱

서울대학교 전기공학부 자동화시스템공동연구소

\*hspark@diehard.snu.ac.kr \*\*kyoungmu@snu.ac.kr \*\*\*sanguk@ipl.snu.ac.kr

## A Segment-based Multiview Stereo Robust to the Existence of Low-textured Region

\*Park, Haesol \*\*Lee, Kyoungmu \*\*\*Lee, Sanguk

Dept. of EECS, ASRI, Seoul National University

### 요약

이 논문에서 우리는 텍스처 정보가 적은 영역이 존재하는 입력 영상들에 대해서도 안정적인 복원을 도출하는 새로운 다시점 스테레오 방법을 제시한다. 제안된 방법에서는 입력 영상들을 인접한 픽셀간의 색 유사성을 이용하여 세그먼트이션 후, 세그먼트 단위로 다시점 스테레오를 수행한다. 특히 그 과정에서 한 영상 내의 이웃한 세그먼트들의 깊이 값 유사성, 그리고 서로 다른 시점에서 상응하는 세그먼트 간의 깊이 값 일관성을 가정하여, 텍스처 정보가 적은 영역에 대해서도 안정적으로 3D 점들을 생성해준다. 생성된 3D 점들은 그래프 컷 기반의 복원 알고리즘을 통해 일관된 3D 표면으로 복원 되었고, 복원 결과는 제안된 방법이 실제로 기존의 다른 다시점 스테레오에 비해 보다 안정적으로 깊이 정보를 추출할 수 있음을 보여준다. 결과적으로 제안된 방법은 보다 일반적이고 실생활에 가까운 입력 영상들에 대해서도 3D 복원을 수행할 수 있는 방향을 제시한다.

## 1. 서론

대상 물체의 3D 모형을 복원하는 것은 컴퓨터 비전의 중요한 연구 분야 중 하나였고 그 동안 이 문제를 해결하기 위한 많은 접근 방법이 제안되어 왔다. 특히 다시점 스테레오를 이용한 복원 방법은 그 뛰어난 범용성과 넓은 적용가능성, 그리고 시스템 구현의 간편성 등의 이유로 매우 활발하게 연구되어 왔다.

특히 Seitz 외의 연구 논문[1]이 발표된 이후 다시점 스테레오의 정량적 평가 기준이 확립됨에 따라 연구가 가속화되었고, 그 결과 laser scanning에 의한 복원에 근접한 정도의 정확도를 갖는 양질의 결과를 도출하는 기법들이 제안되었다[2][3].

하지만 기존의 기법들은 실생활에 가까운 환경에서 얻은 입력 영상들에 대해서는 통제된 상황에서의와 같은 정도의 정확도를 보여주지 못한다. 대부분의 방법들이 photo-consistency를 metric으로 하는 매칭에 기반하여 3D points들을 생성하거나 mesh를 변형하고 있어, 실생활 환경에서 자주 볼 수 있는 단색 옷을 입은 사람이나 단색에 가까운 색을 갖는 개체 등을 정확하게 복원하기 힘들기 때문이다.

이 논문에서 우리는 텍스처가 부족한 영역의 존재에 강한 복원 기법을 제안한다. 제안된 기법은 먼저 (1)텍스처 정보가 부족한 영역에 대해서도 보다 안정적으로 작동하는 세그먼트 기반의 다시점 스테레오를 이용해 깊이 정보를 추출하고 (2)생성된 깊이 정보를 그래프 컷 기반의 전역적 최적화 방법을 이용해 하나의 단일한 표면으로 변환한다. 실험은 일반적인 다시점 입력 영상에 대해 제안된 기법을 적용하여 진행되었으며, 다른 state-of-the-art 기법과의 결과비교는 목표 상황에 대해 제안된 기법이 의미 있는 결과를 만들어낸다는 것을 보여준다.

## 2. 방법

### 가. 세그먼트 기반의 깊이 정보 추출

앞의 논의에서 언급했던 것처럼, 기존의 다시점 스테레오 기법들은 텍스처 정보가 적은 영역에 대해서는 안정적으로 3D 점구름을 생성하지 못한다. 이러한 문제가 발생하는 이유는 대부분의 다시점 스테레오 기법이 픽셀 단위의 window 기반 매칭을 계산하기 때문이다. Window 기반 매칭에서는 window의 크기가 일정 이상으로 커지는 경우 fattening effect 등으로 인해 복원의 전체적인 정확도가 크게 떨어지기 때문에 window 크기를 크게 함으로써 더 넓은 영역의 텍스처 정보를 활용하는 것에 한계가 있다.

이러한 문제점을 극복하기 위해 우리는 세그먼트 기반의 스테레오를 수행하는 Zitnick의 방법[4]을 수정하여 받아들였다. 원래 view interpolation을 위해 사용되었던 Zitnick의 방법은 입력 영상들을 인접한 픽셀간의 색 유사성을 이용하여 세그먼트이션 후, 세그먼트 단위로 다시점 스테레오를 수행한다. 따라서 텍스처 정보가 적은 영역에 대해 보다 큰 크기의 세그먼트를 활용할 수 있다는 이점을 갖게 되는데, 그에 더하여 여러 단계의 smoothness를 고려함으로써 매우 안정적인 깊이 정보를 생성한다.

원래 기존의 Zitnick의 방법[4]은 선형적인 카메라 배열 시스템을 가정하고 있다. 따라서 초기의 깊이 공간 분포(DSD)를 구할 때나 후에 각 세그먼트간 혹은 각 입력 영상 간의 smoothness와 일관성을 고려하여 깊이 값을 세밀하게 조정하는 과정에서도 각 입력 영상에서 대응되는 픽셀 혹은 세그먼트가 같은 깊이 값 혹은 disparity를 가지는 것으로 가정하고 수식을 전개하고 있다. 우리는 원래의 논문에 언급된 수

식들의 변수가 가지는 의미를 부분적으로 재정의함으로써 보다 일반적인 다시점 스테레오 문제를 다룰 수 있도록 기존의 알고리즘을 수정, 보완하고자한다.

먼저, 기존의 수식에서의 모든 disparity를 깊이 값으로 수정하고 이에 따라 기존의 방법에서 이웃한 세그먼트간의 smoothness를 고려하기 위해 고안되었던 식에서  $d_{il}$ 은 이웃한 세그먼트의 DSD가 최대가 되는 깊이를 의미하게 된다. 그 외의 변수들은 기존 방법에서의 동일한 의미를 가진다.

한편, 이웃한 영상 간의 깊이 정보 일관성을 확보하기 위한 수식을 계산하기 위한 과정은 보다 복잡해지는데, 이제 카메라의 배열이 아니게 되었으므로, 인접한 카메라들이 같은 3D 점에 대해 동일한 깊이 값을 갖는다는 가정이 더 이상 성립하지 않기 때문이다.

이에 따라 원래 방법에서의  $p_{ijk}^t(d)$ 와  $o_{ijk}^t(d)$  식이 아래와 같이 변형되어야 한다.

$$p_{ijk}^t(d) = \frac{1}{C_{ij}} \sum_{x \in S_{ij}} p_{\pi(k,x)}^t(d_{proj})$$

$$o_{ijk}^t(d) = 1.0 - \frac{1}{C_j} \sum_{x \in S_{ij}} p_{\pi(k,x)}^t(d_{proj}) h(d - d^{kl} + \lambda)$$

위식에서  $d$ ,  $d_{proj}$ 을 제외한 나머지 모든 변수, 함수들의 정의는 원래 방법에 제시되었던 것과 같고,  $d$ 는 현재 계산하고자하는 픽셀이 가질 수 있는 후보 깊이 값을,  $d_{proj}$ 은 그 픽셀이 원래의 이미지에서 깊이  $d$ 를 이용해 생성하는 3D 점을 이웃한 카메라  $k$ 에서 바라보았을 때의 깊이 값이다.

## 나. 그래프 컷 기반의 표면 추출 방법

3D 점구름으로부터 일관된 3D 표면을 추출하는 방법에는 많은 종류의 방법이 있지만, 여기에서는 특히 노이즈와 외부자들에 대해 강인하게 동작하면서 매끄러운 3D 표면을 복원할 수 있는 그래프 컷 기반의 방법을 이용하기로 한다. 이와 같은 방법의 사용은 세그먼트 기반으로 계산된 깊이 정보에 존재하는 평면 가정 error등에 효과적이다.

특히 장주용 외에 의해 제안된 방법[5]은 기존의 그래프 컷 기반 3D표면 추출의 한계인 regular sampling 효과를 surfel 개념의 도입을 통해 해결하여 높은 정확도와 안정성을 보인다. 이 방법은 깊이 정보 지도로부터 얻은 3D 점들을 이용해 3D 공간 상에 surfel을 생성하고 각 surfel의 위치와 방위를 photo-consistency 하게 세부 조정된 후에 이들을 그래프 구조와 결합하고 photo-consistency 값을 그래프의 각 edge의 weight로 정의하여 그래프 컷을 시행함으로써 안정적인 3D 표면을 복원한다.

## 3. 실험 결과 및 결론

실험에 사용된 dataset은 장당 1280×720의 해상도를 가진 총 20장의 영상으로 이루어져 있다. 이 dataset의 복원 대상인 사람의 경우 대부분의 영역에 대해 텍스처 정보가 부족해 기존의 방법들 중 하나를 이용하여 일관된 표면을 복원하는 것이 상당히 어렵다.

그림 1.은 해당 dataset에 대해 제안된 방법을 적용한 실험 결과와 State-of-the-art 기법 중 하나인 Furukawa의 방법을 적용하여 얻은 결과를 각각 아래 행 왼쪽과 오른쪽에 도시했다. 왼쪽 결과의 빨간 원으로 표시된 부분들은 Furukawa의 방법에서 복원이 제대로 이루어지

지 않아 구멍이 생긴 부분으로, 그림 1.의 위행에 제시된 입력 영상에서 확인할 수 있듯이 해당 영역들은 텍스처 정보가 부족한 부분에 대응된다. 한편, 제안된 방법(오른쪽 아래)을 이용한 경우는 해당 영역들이 비교적 잘 복원되어 있음을 확인할 수 있다.

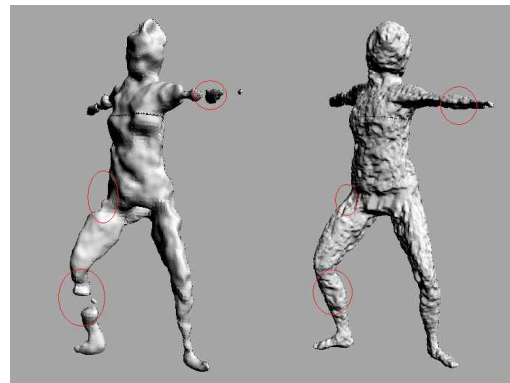


그림 1 위행은 입력 영상을 아래 영상은 복원 결과를 각각 나타내며, 아래 영상의 왼쪽이 Furukawa의 결과, 오른쪽이 제안된 방법의 결과이다.

제안된 방법은 텍스처 정보가 부족한 영역에 대해서도 안정적으로 깊이 정보를 추출하여 3D 복원을 수행할 수 있다. 하지만 3D 점들의 방위 정보가 정확하게 계산되지 않았고, 세그먼트 단위로 이루어진 깊이 정보 추출의 정확도가 높지 않아 그림 1.의 오른쪽에서 보이는 것과 같이 복원 결과의 표면이 매끄럽지 못하다. 남은 과제는 세그먼트 기반으로 수행된 깊이 정보를 픽셀 수준에서 더 정밀하게 추가 계산하고 그를 기반으로 보다 정확한 표면을 복원할 수 있는 기법을 고려하는 것이다.

## 참고문헌

- [1] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. "A comparison and evaluation of multi-view stereo reconstruction algorithms". In CVPR, 2006.
- [2] Y. Furukawa and J. Ponce. "Accurate, dense, and robust multiview stereopsis". In CVPR, 2007.
- [3] C. Zach, "Fast and high quality fusion of depth maps", International Symposium on 3D Data Processing, Visualization and Transmission 2008.
- [4] C. Lawrence Zitnick, Sing Bing Kang, Matthew Uyttendaele, Simon Winder, Richard Szeliski, "High-quality video view interpolation using a layered representation", Proceedings of ACM SIGGRAPH 2004.
- [5] 장주용, 박해솔, 박인규, 이경무, 이상욱, "표면요소 표현과 그래프컷 기반의 다시점 스테레오", In IPIU, 2009.