

## 패치매치 기반 및 분할 기반 조밀 깊이지도의 효율적인 결합 방법

임한신, 서정일

한국전자통신연구원

hslim@etri.re.kr seoji@etri.re.kr

### An Efficient Method to Combine PatchMatch-Based and Segmentation-Based Dense Depth Maps

Hanshin Lim, Jeongil Seo

Electronics and Telecommunications Research Institute

#### 요 약

본 논문에서는 패치매치 기법 및 분할 기법의 조밀 깊이지도들의 효율적인 결합을 통해 기존의 패치매치 기반의 방법들이 낮은 깊이값 추정 정확도를 보인 영역들인 텍스처가 부족한 영역과 기존의 분할 기반 방법들이 깊이값 추정에 한계를 보인 세밀한 영역에서의 깊이값 추정 정확도를 동시에 높이고 고품질의 조밀 깊이지도를 얻는 것을 목표로 한다. 이를 위해 제안한 방법에서는 신뢰지도를 바탕으로 패치매치 기법의 조밀 깊이지도, 조밀 노말지도와 분할 기법의 조밀 깊이지도, 조밀 노말지도의 초기 결합 깊이지도 및 초기 결합 노말지도를 생성한다. 이후 각 픽셀에서 원래 픽셀과 주변 픽셀에서의 깊이값, 노말값들로 업데이트를 위한 후보들을 만든다. 이후 각각의 후보들에 대해서 깊이값, 노말값, 컬러값들을 바탕으로 비용을 계산한다. 이후 가장 최적의 비용을 가지는 후보값으로 각 픽셀의 깊이값과 노말값을 업데이트한다. 이를 통해 패치매치 기법 및 분할 기법의 조밀 깊이지도들의 장점을 합친 결합 조밀 깊이지도를 생성한다.

#### 1. 서론

한 장면이나 물체를 다양한 시점에서 찍은 다시점 영상들로부터 원래의 장면이나 물체의 정확한 3D 구조를 복원하고 모델링하는 기술은 일반적으로 크게 특징점 추출 및 매칭, 영상의 카메라 파라미터 추정, 조밀 포인트 클라우드(dense point cloud) 생성, 3D 표면 복원으로 이루어진다. 이 중에서 조밀 포인트 클라우드 생성 단계에서는 다시점 영상과 structure-from-motion 등의 방법으로 구한 카메라 파라미터를 입력으로 하여 다시점 영상 안의 장면 및 객체들의 구조를 3D 포인트 클라우드 형태로 정확히 표현하는 것을 목표로 한다. 다시점 영상으로부터 조밀 포인트 클라우드를

생성하는 과정에서 가장 핵심적인 단계는 다시점 영상들로부터 고품질의 조밀 깊이지도를 얻어내는 과정이다.

다시점 영상들로부터 고품질 깊이지도를 생성하는 방법들 중 가장 많이 쓰이고 있는 방법 중의 하나가 패치매치(PatchMatch) 방법이다. 다시점 영상에서는 하나의 기준 영상과 주변의 여러영상들 간의 컬러값의 유사성 등으로부터 깊이값을 추정하는데 깊이값의 후보들을 최대한 정확하고 효율적으로 줄이는 것이 중요하다. 패치매치에서는 이러한 깊이값의 후보들의 모임을 효율적으로 줄여주어서 다시점 영상에서의 깊이지도 생성에 성공적으로 적용될 수 있었다. 하지만 패치매치 기반 방법들은 텍스처 정보가 부족한 영역과 같이 컬러값의 유사성 등으로 깊이값을 추정하기 어려운 영역에서 깊이값 추정의 정확도가 매우 떨어지는 것으로 알려져 있다.



그림 1. 제안하는 기술의 전체 순서

다시점 영상들로부터 고품질 깊이지도와 노말지도를 생성하는 또 다른 대표적인 방법은 영상을 메쉬 등의 분할 기법을 적용하여 분할된 영역 단위로 깊이값을 구하는 것이다. 이러한 방법은 패치매치 방법에 비해 텍스처 정보가 부족한 영역 등의 컬러값의 유사성 등으로 깊이값 추정이 어려운 영역에서 좀 더 안정적인 값을 얻을 수 있다. 하지만 표현상의 문제로 세밀한 영역에서는 깊이값 추정에 한계가 있다.

본 논문에서는 패치매치 및 분할 기반의 깊이지도의 효율적인 결합을 통하여 기존의 패치매치 기반의 방법들이 낮은 깊이값 추정 정확도를 보인 영역들인 텍스처가 부족한 영역과 기존의 분할 기반 방법들이 깊이값 추정에 한계를 보인 세밀한 영역에서의 깊이값 추정 정확도를 동시에 높임으로써 고품질의 조밀 깊이지도와 노말지도를 얻는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2 절에서는 제안 기술에 대해 설명하고 3 절에서는 제안한 기술의 성능 실험 결과를 보여준다. 4 절에서는 본 논문에 대한 결론을 맺는다.

## 2. 제안 기술

### 1. 초기 조밀 깊이지도 및 조밀 노말지도 생성

다시점 영상 및 각 영상의 카메라 파라미터가 주어졌을 때 먼저 패치매치 기법을 통해 조밀 깊이지도, 조밀 노말지도를 생성 및 분할 기법을 통해 조밀 깊이지도, 조밀 노말지도를 생성한다.

### 2. 조밀 신뢰지도 생성

패치매치 기법을 통해 생성된 조밀 깊이지도에서의 각 픽셀에서의 깊이값들의 신뢰도의 정도를 나타내는 조밀 신뢰지도를 생성한다. 깊이값의 신뢰도는 다양한 방법으로 측정될 수 있다. 본 논문에서는 각 픽셀에서의 깊이값에 대한 주변 영상들과의 컬러일관성에 기반한 비용들을 계산하고 이를 평균한 값을 신뢰도라고 판단하였다.

### 3. 초기 결합 깊이지도 및 결합 노말지도 생성

패치매치 기법과 분할 기법으로 각각 생성된 조밀 깊이지도와 조밀 노말지도를 조밀 신뢰지도를 바탕으로 결합하여 초기 결합 깊이지도 및 초기 결합 노말지도를 생성한다. 조밀 신뢰지도의 각 픽셀에서 신뢰값이 일정 수준 이상이면 패치매치

기법을 통해 생성된 조밀 깊이지도와 조밀 노말지도에서의 동일 위치에서의 깊이값과 노말값을 초기 결합 깊이지도 및 결합 노말지도의 값으로 할당한다. 만일 조밀 신뢰지도의 신뢰값이 일정 수준 아래이면 분할 기법을 통해 생성된 조밀 깊이지도와 조밀 노말지도에서의 동일 위치에서의 깊이값과 노말값을 초기 결합 깊이지도 및 결합 노말지도의 값으로 할당한다.

### 4. 각 픽셀에서의 깊이값 및 노말값 업데이트

각 픽셀에서 현재 픽셀의 깊이값, 노말값과 이전 픽셀의 깊이값, 노말값들로 업데이트할 깊이값과 노말값의 후보쌍을 만든다. 본 논문에서의 후보쌍들은 아래와 같다.

$$\{(d_{cur}, n_{cur}), (d_{cur}, n_{prev}), (d_{prev}, n_{cur}), (d_{prev}, n_{prev}), (d_{cur}, n_{rand}), (d_{prev}, n_{rand}), \}$$

여기서  $d_{cur}, n_{cur}$  은 각각 현재 픽셀에서의 깊이값, 노말값을 나타낸다. 그리고  $d_{prev}, n_{prev}$  은 각각 이전 픽셀에서의 깊이값, 노말값을 나타낸다.  $n_{rand}$  은 랜덤하게 선택된 노말값을 나타낸다.

각 깊이값과 노말값의 후보쌍들로부터 주변 영상들과의 컬러일관성에 기반한 비용(cost)을 재계산한다. 여기서 후보쌍의 깊이값이 현재 깊이값인지의 여부와 후보쌍의 노말값이 랜덤인지의 여부에 따라 컬러일관성에 기반한 원래의 비용에 추가적인 값을 더한 아래와 같은 업데이트된 비용을 구한다.

$$\begin{aligned} Cost(d_{cur}, n_{cur})' &= Cost(d_{cur}, n_{cur}) + \alpha w_c + \alpha w_d \\ Cost(d_{cur}, n_{prev})' &= Cost(d_{cur}, n_{prev}) + \alpha w_c + \alpha w_d \\ Cost(d_{prev}, n_{cur})' &= Cost(d_{prev}, n_{cur}) + \alpha w_d \\ Cost(d_{prev}, n_{prev})' &= Cost(d_{prev}, n_{prev}) + \alpha w_d \end{aligned}$$

여기서  $Cost(d, n)$ 는 깊이값  $d$ , 노말값  $n$  일 때 컬러일관성에 기반한 원래의 비용이고  $\alpha$  는 상수이다.  $w_c$  와  $w_d$  는 각각 현재 픽셀과 이전 픽셀의 컬러값 및 깊이값의 변화정도를 나타낼 수 있다. 본 논문에서  $w_c$  와  $w_d$  는 각각 아래와 같다.

$$\begin{aligned} w_c &= \exp(-||C_{cur} - C_{prev}||^2 / \sigma_c) \\ w_d &= \exp(-||d_{cur} - d_{prev}||^2 / \sigma_d) \end{aligned}$$

여기서  $C_{cur}$  와  $C_{prev}$  는 각각 현재 픽셀과 이전 픽셀에서의 컬러값이다.  $\sigma_c$  와  $\sigma_d$  는 각각 상수이다.

각 픽셀의 후보쌍들에 대해서 업데이트된 비용을 구하고 가장 비용이 큰 값을 가진 후보쌍을 결합 깊이지도 및 결합 노말지도의 해당되는 픽셀에서의 깊이값 및 노말값으로 업데이트

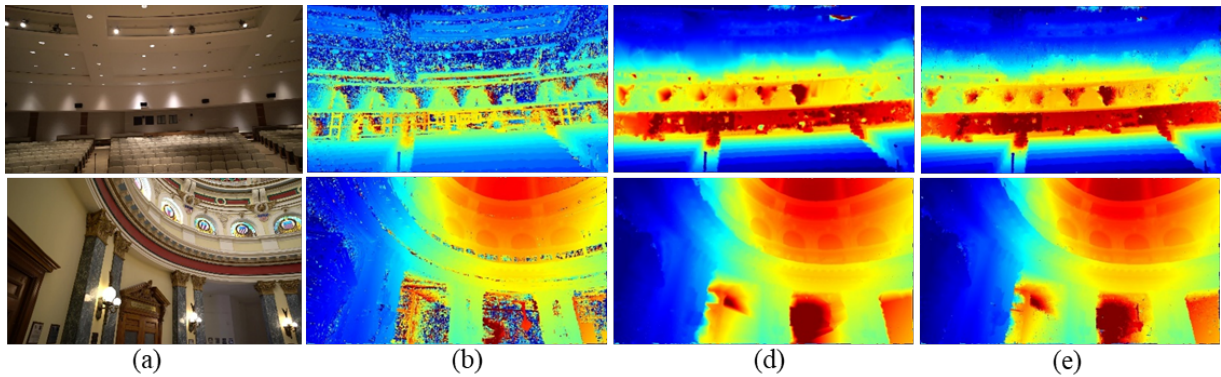


그림 2. Tanks and Temples Auditorium, Museum dataset 에서 (a)원본 영상과 (b)패치매치 기법을 통해 생성된 조밀 깊이지도, (c)분할 기법을 통해 생성된 조밀 깊이지도 (d) 제안하는 조밀 깊이지도의 결합 방법을 통해 생성된 조밀 깊이지도

표 1. Tanks and Temples advanced dataset 에서 기존 방법들과 제안한 방법의 성능(F-score) 비교

	Auditorium	Ballroom	Courtroom	Museum	Palace	Temple	Mean
COLMAP[1]	16.02	25.23	34.70	41.51	18.05	27.94	25.83
EPP-MVSNet[2]	21.28	39.74	35.34	49.21	30.0	38.75	35.72
Altizure-HKUST[3]	24.04	44.52	36.64	49.51	30.23	39.09	37.34
Proposed	27.68	40.15	42.69	53.76	23.19	37.60	37.51

한다. 이러한 과정을 모든 픽셀에 대해서 수행한다.

효율적으로 할당하는 작업을 수행하여 텍스처 정보가 부족한 영역 및 세밀한 영역에서의 깊이값 추정 정확도를 향상시켰다.

### 3. 실험 결과

실험에 사용한 영상은 Tanks and Temples public dataset 중 advanced testset(1920x1080)이었다. 카메라 파라미터는 COLMAP[1]을 통해 구하였다. 그림 2 는 패치매치 기법을 통해 생성된 조밀 깊이지도와 분할 기법을 통해 생성된 조밀 깊이지도로부터 정제된 조밀 깊이지도를 생성한 예시이다. 표 1은 Tanks and Temples advanced dataset 에서 기존 방법들과 제안한 방법의 성능(F-score)을 비교한 것이다. 제안한 방법은 기존 기술들의 문제점들인 텍스처가 부족한 영역 및 세밀한 영역에서의 깊이값 추정을 모두 효과적으로 할 수 있으므로 기존보다 고품질의 깊이지도를 구할 수 있다.

### 4. 결론

본 논문에서는 두 가지 방법으로 구한 조밀 깊이지도에서 각 영역별로 신뢰성이 더욱 높다고 판단되는 깊이값을 각 픽셀에

### Acknowledgement

본 연구는 한국전자통신연구원 연구운영비지원사업의 일환으로 수행되었음. [22ZH1200, 초실감 입체공간 미디어.콘텐츠 원천기술 연구]

### 참고문헌

[1] Schonberger, Lutz, J., Jan-Michael, F., "Structure-from-motion revisited," Conference on Computer Vision and Pattern Recognition, 2016.

[2] Ma, X., Gong, Y., Wang, Q., Huang, J., Chen, L., Yu, F., "Epp-mvsnet: Epipolar-assembling based depth prediction for multi-view stereo," International Conference on Computer Vision, 2021.

[3] Yao, Y., Luo, Z., Shen, T., Li, S., Fang, T., Quan, L., "Recurrent mvsnet for high-resolution multi-view stereo depth inference," Conference on Computer Vision and Pattern Recognition, 2019.