

Epipolar geometry를 활용한 개선된 depth 평가 방법

김성민, 한종기

세종대학교

ksmksj5265@sju.ac.kr hjk@sejong.edu

Improved depth evaluation using Epipolar geometry

Seong-Min Kim, Jong-Ki Han

Sejong University

요 약

실재하는 물체나 장소를 디지털 카메라나 휴대폰 카메라로 여러 장 촬영하여 얻은 2차원 이미지 데이터셋으로부터 3차원 영상을 얻기 위해서 이미지를 이루는 각 pixel의 depth 정보를 얻는 것은 필수적인 과정이다. 주어진 이미지에서 depth 정보를 얻기 위해 Shuhan Shen은 PatchMatch 알고리즘을 활용하는 것을 제안하였다. 그 이후 PatchMatch 기반의 알고리즘은 널리 사용되며 우수한 성능을 보이고 있다. PatchMatch 기반의 알고리즘을 사용해 depth를 추정하는 과정에서 depth와 법선 벡터를 Zero-mean Normalized Cross Correlation(ZNCC)를 사용해 평가한다. 하지만, ZNCC는 depth를 평가하려는 pixel의 주변 pixel들의 밝기 값 혹은 색상 값의 분포를 사용하기 때문에 밝기 값이나 색상 값의 변화가 적은 texture-less region에서는 신뢰성이 떨어진다. 본 논문에서는 이 문제를 epipolar geometry를 활용한 기하학적 정보를 이용하여 개선하고자 한다.

1. 서론

Structure from Motion(SfM) 알고리즘[1]을 활용하면 여러 장의 2차원 이미지로부터 카메라의 위치 정보와 3차원 상의 저밀도의 point cloud를 얻을 수 있다. SfM을 통해 얻은 정보들을 활용해 PatchMatch 기반의 알고리즘에서 각 pixel의 depth 정보를 추정할 수 있다. 이 depth 정보를 활용하면 SfM에서 얻은 point cloud보다 더욱 밀도가 높은 point cloud를 얻을 수 있으며 밀도가 높은 point cloud를 얻을수록 더 실물과 비슷한 3차원 복원 영상을 얻을 수 있다.

Shuhan Shen은 depth 정보를 추정하기 위해 PatchMatch 알고리즘을 활용하는 것을 제안하였다.[2] 이후 PatchMatch 기반의 알고리즘은 좋은 성능을 자랑하며 널리 사용되고 있다. 하지만, 밝기 값 혹은 색상 값의 변화가 적은 texture-less region에서는 정확한 depth를 추정하지 못한다는 문제점을 가지고 있다.[3] 따라서 해당 부분은 각 이미지로부터 얻은 depth-map과 최종적으로 얻은 point cloud에서 표현되지 못해서 구멍이 뚫린 것처럼 보인다.

위 문제점의 근본적인 원인은 PatchMatch 기반의 알고리즘은 추정된 depth를 평가할 때 ZNCC를 활용한다는 것이다. ZNCC는 일반적인 경우에는 높은 신뢰도를 보이지만, texture-less region에서는 낮은 신뢰도를 보이기 때문에 추정된 depth를 올바르게 평가하기 어려워 결

과적으로 부정확한 depth를 추정하게 한다.

본 논문에서는 epipolar geometry를 활용해 depth 평가 수식에 기하학적인 정보를 추가하는 새로운 depth 평가 수식을 제안함으로써 위의 문제를 개선하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서 기존에 사용한 depth 평가 방법과 문제점을 살펴보고 3장에서 기존 방법의 문제점을 해결하기 위한 방법을 제안한다. 4장에서 실험 결과를 보이고 5장에서 결론짓는다.

2. 기존 방법의 문제점

PatchMatch 알고리즘의 전체적인 과정은 다음과 같다.[3] 먼저, 각 pixel의 depth와 법선 벡터를 랜덤으로 초기화하고, 각 이미지에 대해 이웃 이미지를 선택한다. 그 후 모든 pixel에 대해 주변 pixel들의 depth와 법선 벡터를 자신의 것과 비교하여 더 좋은 depth와 법선 벡터를 선택하며 이 과정을 spatial propagation이라고 한다. 이 과정이 끝나면 모든 pixel에 대해 추정된 depth와 법선 벡터를 약간씩 움직여보며 더 나은 depth와 법선 벡터를 찾는 random refinement를 진행한다. spatial propagation과 random refinement는 여러 번 반복될 수 있다. 이 과정이 끝난 후에 각 이미지별로 구한 depth 정보를 비교하여 서로 일관성이 없는 depth는 걸러진다. 마지막으로 모든 이미지에 대해 depth map이 구해졌으므로 이를 활용하여 pixel을 공간상에 역투영하여 point cloud로 합친다.

1) 연락처: 한종기

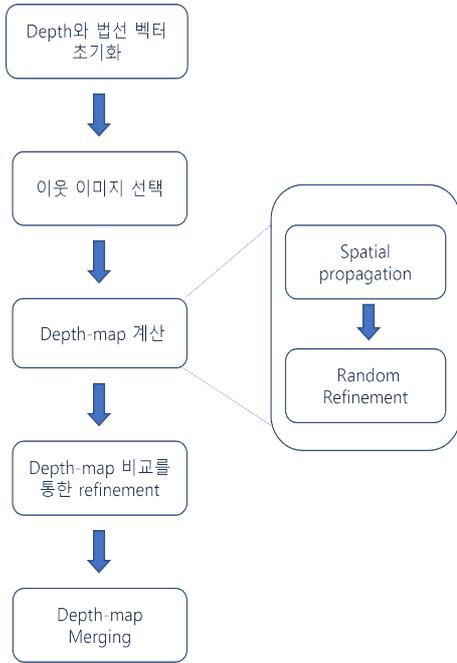


그림 1. PatchMatch 기반의 depth 추정 알고리즘의 전반적인 Flow chart

위의 과정을 보면 알 수 있듯이 PatchMatch 알고리즘에서는 추정된 depth와 법선 벡터가 얼마나 잘 추정되었는지에 대해 평가하는 것이 매우 중요하다. 기존에 사용하는 평가 기법은 다음과 같다. 현재 pixel에서 추정된 depth와 법선 벡터를 사용해 아래와 같이 호모그래피 행렬을 구한다.[2]

$$H_{ij,p} = K_j(R_j R_j^{-1} + \frac{R_j(C_i - C_j)n_{i,p}^T}{n_{i,p}^T X_{i,p}})K_j^{-1} \quad (1)$$

$$X_{i,p} = \lambda K_i^{-1} p \quad (2)$$

i 는 depth를 구하고자 하는 이미지의 인덱스이며, j 는 i 번째 이미지의 이웃 이미지의 인덱스이다. p 는 i 이미지 위의 pixel 좌표로 동차좌표계로 표현되며 현재 depth를 구하고자 하는 pixel이다. K, R, C 는 각 이미지에 대한 카메라 내부 파라미터, 카메라 회전 행렬, 카메라 중심 좌표이며 λ 는 p 에서 추정된 depth, $n_{i,p}$ 은 추정된 법선 벡터. $X_{i,p}$ 는 추정된 depth로 p 를 i 의 3차원 카메라 좌표계 상에서 표현한 좌표이다. $H_{ij,p}$ 는 pixel p 를 이미지 i 에서 j 로 warping 해주는 호모그래피 행렬이다. 즉, 호모그래피 행렬은 각 pixel마다 구해진다.

위에서 구한 호모그래피 행렬을 사용해 이웃 이미지로 warping 시킨다. 이 때, 추정된 depth와 법선 벡터가 정확하다면 올바른 위치로 warping 된다. warping이 잘 되었는지를 평가하기 위해 현재 pixel의 주변에 있는 pixel의 밝기 분포와 warping 된 위치 주변에 있는 pixel의 밝기 분포를 비교하는데, 이 때, ZNCC를 활용해 $1 - ZNCC$ 를 손실 함수로 사용한다.

$$ZNCC = \frac{\sum_{q \in B} (q - \bar{q})(H_{ij,p}(q) - \overline{H_{ij,p}(q)})}{\sqrt{\sum_{q \in B} (q - \bar{q})^2 \sum_{q \in B} (H_{ij,p}(q) - \overline{H_{ij,p}(q)})^2}} \quad (3)$$

B 는 pixel p 주변의 $n \times n$ box를 의미하며 q 는 pixel p 의 밝기 값

을 의미, \bar{q} 는 B 내부 pixel들의 밝기 값의 평균이며 $H_{ij,p}(q)$ 는 호모그래피 행렬로 pixel p 를 이웃 이미지로 warping한 좌표에서의 밝기를 의미한다.

ZNCC는 [그림 2]와 같은 경우에 신뢰도가 높다. [그림 2-(a)]의 노란색 box의 중심 pixel에서 depth와 법선 벡터가 잘 추정되었다면 [그림 2-(b)]의 파란색 box로 올바르게 warping되며 ZNCC 값이 높다. 반면 depth와 법선 벡터가 잘 추정되지 않는다면 빨간색 box처럼 이상한 곳으로 warping되며 ZNCC 값도 낮다. 따라서, ZNCC값을 통해 depth와 법선 벡터를 잘 구했는지 알 수 있다. 하지만, [그림 3]과 같은 경우에는, [그림 3-(a)]의 노란색 box가 [그림 3-(b)]의 파란색 box로 warping된 경우가 노란색 box의 중심 pixel의 depth와 법선 벡터가 잘 추정된 경우이지만 depth와 법선 벡터가 잘못 추정되어 빨간색 box와 같이 잘못된 위치로 warping되더라도 ZNCC 값이 파란색 box와 빨간색 box가 비슷하기 때문에 ZNCC 값으로는 depth와 법선 벡터가 잘 추정되었는지 알 수 없다. 이는 [그림 3]의 두 이미지에 표현된 box들과 같은 texture-less region은 넓은 영역에 걸쳐 pixel의 밝기 값이 균등하여, 비슷한 분포를 가지게 되기 때문이다. 즉, texture-less region에서는 ZNCC의 신뢰도가 떨어진다고 말할 수 있다.

위의 이유로 인해 PatchMatch 기반의 알고리즘은 texture-less region에서 depth를 잘 추정하지 못한다는 문제점을 가지고 있다. 따라서, depth map에서 해당 부분이 구멍이 뚫린 것처럼 보이며 마지막으로 얻게 되는 point cloud에도 해당 부분이 촘촘하게 구성되지 못하고 구멍이 뚫리게 된다.

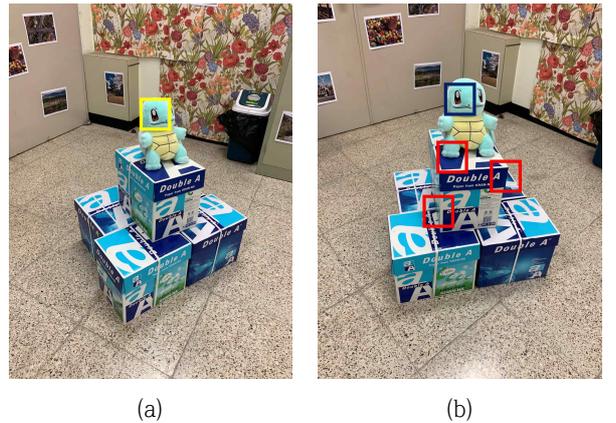


그림 2. ZNCC의 신뢰도가 높은 경우

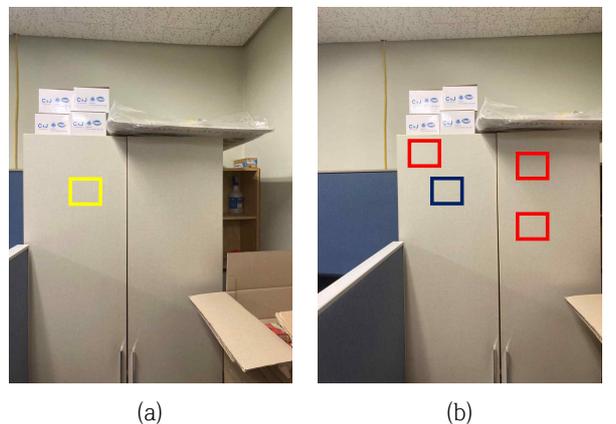


그림 3. ZNCC의 신뢰도가 낮은 경우

3. 제안하는 방법

제안하는 방법은 기존의 ZNCC에 epipolar geometry를 활용해 기하학적인 정보를 추가해주는 것이다. epipolar geometry는 같은 장면을 다른 각도에서 찍은 두 장의 이미지 i, j 가 있을 때, 두 이미지 간에 서로 매칭되는 pixel들 사이의 기하학적 관계를 설명한다.[4] 공간상의 점 X 가 이미지 i 에서 p 에 나타나고, 이에 매칭되는 pixel p' 이 이미지 j 에 존재할 때, depth 정보를 모르므로 p' 이 정확히 어떤 위치에 있는지는 알 수 없지만 어떤 직선 l 위에 존재하게 된다. 이 직선 l 을 epipolar line이라고 한다. epipolar line은 Fundamental 행렬을 이용하면 쉽게 구할 수 있다.

$$l = Fp \quad (4)$$

F 는 Fundamental 행렬이고 p 는 이미지 i 위의 pixel이며 동차좌표계로 표현된다. 만약, p 에서 구한 depth와 법선 벡터를 활용해 호모그래피를 구한 후 이미지 j 로 warping한 위치가 l 과의 거리가 멀수록 부정확한 warping이 일어난 것이며 depth와 법선 벡터가 부정확하게 구해졌다는 것을 의미한다.

이를 활용하여 기존의 주변 pixel의 밝기 값 분포를 활용한 ZNCC에 epipolar line l 과 warping 위치의 거리 정보를 더한 새로운 depth 평가 방법을 제안한다.

$$ZNCC + \omega e^{-distance'} \quad (5)$$

$$distance' = \frac{distance \times 100}{\sqrt{width^2 + height^2}} \quad (6)$$

$$distance = \frac{|l \cdot p'|}{\sqrt{l_1^2 + l_2^2}} \quad (7)$$

ω 는 가중치 파라미터이며 이 값이 크면 클수록 epipolar geometry의 기하학적인 정보가 더 많이 반영된다. $distance'$ 은 정규화된 거리 값이다. 이웃 이미지들의 크기가 제각각인 경우를 고려하며 100을 곱하는 이유는 해당 항의 값이 너무 작아지는 것을 방지하기 위함이다. $width$ 와 $height$ 는 이미지 j 의 너비와 높이이며 l_1 와 l_2 는 각각 l 의 첫 번째 원소와 두 번째 원소이다. p' 은 이미지 j 로 warping된 위치이며 동차좌표계로 표현된다.

4. 실험

본 실험은 OpenMVS[5]의 결과와 제안하는 방법을 비교하였다. 더 정확한 비교를 위해 제안하는 평가 수식의 결과 값도 ZNCC처럼 -1에서 1사이로 출력되도록 아래와 같이 사용하였다.

$$\frac{ZNCC + \omega e^{-distance'}}{0.5\omega + 1} - 1 \quad (8)$$

이는 depth 추정 뒤에서 일어나는 refinement 단계와 point cloud를 생성하는 단계에서 depth 평가 값을 기반으로 진행되는 과정에 일관성을 주기 위함이다. 또한 ω 는 1.5를 사용하였다.

본 논문에서 제안하는 방법을 활용한 결과 [그림 4-6]에서 볼 수 있듯이 개선된 point cloud를 얻을 수 있고, [그림 7-8]에서 볼 수 있듯이 개선된 depth map을 얻을 수 있었다.



(a)



(b)

그림 4. (a) 기존 OpenMVS로 얻은 point cloud (b) 제안하는 방법으로 얻은 point cloud



(a)



(b)

그림 5. (a) 기존 OpenMVS로 얻은 point cloud (b) 제안하는 방법으로 얻은 point cloud

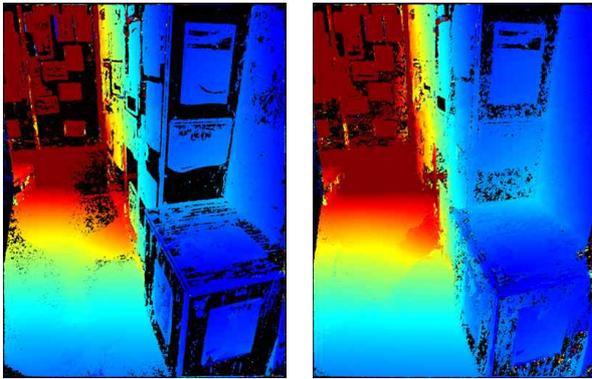


(a)



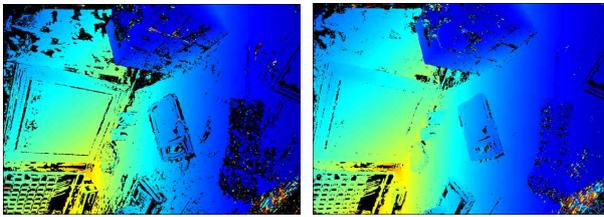
(b)

그림 6. (a) 기존 OpenMVS로 얻은 point cloud (b) 제안하는 방법으로 얻은 point cloud



(a) (b)

그림 7. (a) 기존 OpenMVS로 얻은 depth map (b) 제안하는 방법으로 얻은 depth map



(a) (b)

그림 8. (a) 기존 OpenMVS로 얻은 depth map (b) 제안하는 방법으로 얻은 depth map

5. 결론

본 논문은 PatchMatch 기반의 depth 추정 알고리즘의 과정 중에서 epipolar geometry를 활용하여 depth를 평가하는 방법을 제안한다. 제안하는 방법은 texture-less region에서의 depth 추정에 도움을 주어 기존의 방법보다 더 밀도가 높은 point cloud와 hole이 더 적은 depth map을 구할 수 있다.

하지만, feature가 너무 적은 이미지 세트에 대해서는 잘 작용하지 않는 모습을 보여 개선이 필요하다.

Acknowledgment

This work was partially supported by the National Research Foundation of Korea (NRF) under Grant 2022R1F1A1071513 and partially by the Institute for Information & Communications Technology Promotion (IITP) under Grant 2017-0-00486 funded by the Korea government through the Ministry of Science and ICT (MSIT).

참 고 문 헌 (References)

- [1] P.Moulon, P.Monasse, and R.Marlet, "Adaptive structure from motion with a contrario model estimation," In Asian Conf. Comput. Vision pp.257-270, Springer, 2013.

Available: <https://github.com/openMVG/openMVG/>.

- [2] Shuhan Shen, "Accurate Multiple View 3D Reconstruction Using Patch-Based Stereo for Large-Scale Scenes," IEEE transaction on image processing, 22(5):1901-1914, 2013
- [3] Shang Sun, Dan Xu, Hao Wu, Haocong Ying, Yurui Mou, "Multi-view Stereo for large-scale scene reconstruction with MRF-based depth inference", In Computers&Graphics 106 (2022) pp.248-258, 2022
- [4] Richard Hartley, Andrew Zisserman, "Multiple View Geometry in Computer Vision", 2nd edition, pp.239-261, 2004
- [5] OpenMVS. Accessed: 2019. [Online]. Available : <https://github.com/cdcseacave/openMVS>