

# 다중 정보 융합 비용 체적을 이용한 라이트 필드 카메라에서의 깊이지도 추정

김민정<sup>0</sup>, 오태현, 권인소  
한국과학기술원 전기및전자공학과  
{mjkim, thoo}@rcv.kaist.ac.kr, [iskweon77@kaist.ac.kr](mailto:iskweon77@kaist.ac.kr)

## 요약

상용 라이트 필드 카메라(Light field camera)가 등장함에 따라 한 번의 촬영으로 원하는 곳에 초점을 맞추고, 다양한 시점에 대응되는 영상을 생성하는 것이 가능해졌다. 이러한 기능 구현을 위해서는 라이트 필드 카메라로부터 깊이 정보를 추정하는 과정이 필수적이지만, 라이트 필드 카메라를 이용하여 촬영한 영상에 잡음이 많고 부분 영역 영상(Sub-aperture image)들의 베이스라인(Baseline)이 좁다는 특성 때문에 라이트 필드 카메라로부터 깊이 정보를 추정하는 것에는 어려움이 있다. 본 논문에서는, 비용 체적(Cost volume)으로부터 추정된 깊이 정보 중, 신뢰도가 높은 깊이 정보를 전파하는 방식을 통해 초기 깊이지도(Depth map)를 추정함으로써 비-볼록(Non-convex) 목적함수 최적화의 성능 개선에 도움을 주는 방식을 제안하였다. 실험 결과 이러한 방식은 라이트 필드 카메라로부터 깊이 지도를 안정적으로 구하는데 도움을 주었다.

## 1. 서론

상용 라이트 필드 카메라가 등장함에 따라 라이트 필드 카메라로부터 깊이 정보를 추정하는 연구가 많은 사람들의 주목을 받게 되었다. 하지만 라이트 필드 카메라의 최대 베이스라인이 렌즈의 구경 크기로 제한될 뿐 아니라 라이트 필드 카메라의 부분 영역 영상이 영상 잡음을 포함하고 있기 때문에, 라이트 필드 카메라로부터 깊이 정보를 추정하는 것에는 상당한 어려움이 있다.

검출된 직선상에 있는 시차 정보(Disparity)는 연속적으로 변화할 것이라는 가정을 사용한 연구[1]는 직선의 검출 결과에 의존하기 때문에 때로 부정확한 결과를 도출하였고, MRF 모델을 사용함으로써 최종적으로 연속적이지 않은 깊이 지도를 생성하였다. EPI(Epipolar image) 상에서 직선의 기울기를 검출하여 깊이 정보를 추정하는 연구[2]는 잡음이 많은 라이트 필드 카메라 영상에서 깊이 정보를 얻는 것에 쉽게 실패한다. EPI 상의 기울어진 직선이 수직이 되는 순간을 찾는 연구[3]는 초점 정보와 색상 정보를 모두 사용하여 각각의 깊이 지도를 생성하지만 이를 효율적으로 융합하지 못했을 뿐 아니라 절대적 깊이 정보가 아닌 상대적 깊

이 정보만을 계산할 수 있다.

본 논문에서는, 비용 체적을 이용하여 안정적인 깊이 지도를 추정하는 방법론을 제시하였다. 비용 체적 개념은 빠른 깊이 지도의 계산을 가능하게 한다. 특징 정보가 거의 없는 영역과 부분 영역 영상들의 좁은 베이스라인 때문에 발생하는 깊이 정보 모호성 문제를 해결하기 위하여, 믿을 수 있는 깊이 정보를 선발한 후 주변의 믿을 수 없는 깊이 정보를 추정하는 데 사용하였다. 믿을 수 있는 깊이 정보들로부터 신뢰도가 낮은 깊이 정보들을 계산하기 위해 유사한 색을 가진 인접한 픽셀들은 유사한 깊이 정보를 가질 것이라는 가정이 사용되었다. 이러한 가정은 비-볼록 최적화 과정의 성능을 상당히 개선시켜주었다. 색상-색상변화량 일관성을 나타내는 항과 인접 깊이 정보의 변화량을 나타내는 항으로 구성된 목적 함수를 최소화함으로써 최종 깊이 지도를 얻기 위해 이산-연속 최적화(Discrete-Continuous optimization) 기법이 채택되었다. 이산-연속 최적화 기법은 이산적인 비용 체적 구조와 연속적인 인접 깊이 유사성을 효율적으로 결합하는 데 매우 적합하다.

본 논문의 2장에서는 비용 체적을 생성하는 방법을 설명한다. 3장에서는 초기 깊이 지도 생성 방

법을 소개한다. 4장에서는 깊이지도 최적화를 위해 설계된 식을 소개한다. 5장에서는 실험 결과를 보인 후 이를 분석하고, 6장에서 논문을 마무리 짓도록 한다.

## 2. 비용 체적 생성

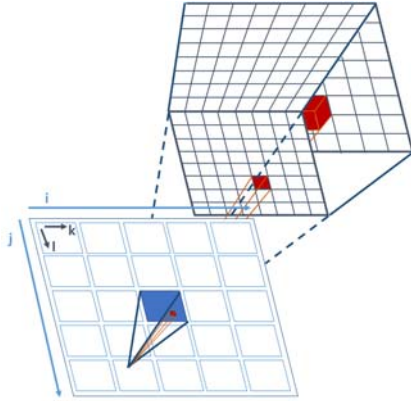


그림 1. 비용 체적 생성

만일 3차원 공간상에 존재하는 올바른 깊이 정보를 가진 복셀(Voxel)이 다수의 부분 영역 영상으로 투영(Projection)될 경우 그에 대응되는 화소들은 모두 같은 색상과 색상 변화량 정보를 갖게 된다. 그렇기 때문에, 만일 기준이 되는 부분 영역 영상 화소와 임의의 깊이 정보를 가정하고 투영했을 때 그에 대응되는 부분 영역 영상 화소들과의 색상-색상 변화량 차이의 합을 각 복셀 안에 저장할 경우 그 값이 최소가 되는 복셀의 깊이 정보가 가장 높은 확률로 올바른 깊이 정보에 해당하게 된다.

색상-색상 변화량을 표현하기 위해 수식 1 과 같은 특징 정보 벡터(Vector)가 제안되었다.

$$\alpha = [R \ G \ B \ \frac{dR}{dx} \ \frac{dG}{dx} \ \frac{dB}{dx} \ \frac{dR}{dy} \ \frac{dG}{dy} \ \frac{dB}{dy}] \quad (1)$$

제안된 특징 정보 벡터를 이용하여 각 복셀 내에 저장되는 값은 수식 2 와 같다.

$$C(p, d) = \frac{1}{N} \sum_{I(i', j') \in I_n} (\|\alpha(i, j, k, l) - \alpha(i', j', k', l')\|_1) \quad (2)$$

$I_{(i', j')}$ 은 ( $j'$ 번째 행,  $i'$ 번째 열)에 존재하는 부분 영역 영상을,  $I_n$ 은 복셀을 투영했을 때 대응되는 좌표가 영상 크기를 벗어나지 않는 부분 영역 영상의 집합을,  $N$ 은  $I_n$ 의 수를 의미한다. 수식 2 의 값을 저장하고 있는 복셀들로 구성된 체적을 비용

체적이라 하며 그림 1에서 비용 체적의 개념을 잘 보여주고 있다. 그림 1에서 파란색 네모 박스가 기준 부분 영역 영상을 나타내고, 파란색 네모 박스 안에 있는 조그만 빨간 네모 박스가 한 픽셀을 의미한다. 색이 칠해지지 않은 하늘색 네모 박스들은 기준 부분 영역 영상 외의 다른 부분 영역 영상들을 의미한다. 그림 상에서 부분 영역 영상들은 ( $i, j$ )의 좌표계로 표기되고, 각 부분 영역 영상 내에 존재하는 화소들의 좌표계는 ( $k, l$ )로 표기되었다. 공간상에 존재하는 육면체들은 색상-색상 변화량 차이의 합을 저장하는 복셀을 나타낸다. 복셀 투영을 위해서는 카메라 교정(Calibration) 과정이 필요한데, 이 때 연구 [4]에서 제안된 방법이 그대로 사용되었다.

## 3. 초기 깊이지도 생성

비용 체적 내에 저장된 값을 최소로 만드는 깊이 정보를 취하더라도 특징 정보가 없는 구간에서는 깊이 정보 추정에 실패하게 된다. 그렇기 때문에 본 논문에서는 신뢰도가 높은 깊이 정보로부터 신뢰도가 낮은 깊이 정보를 추정하는 방식을 제안하였다.

### 3.1 기준 화소 선택

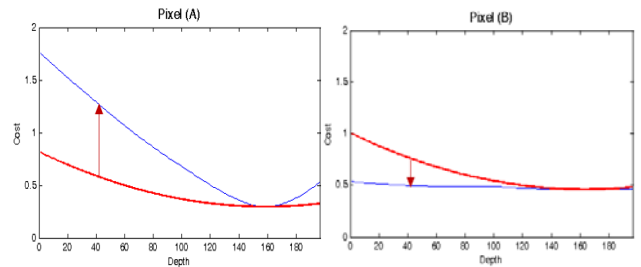
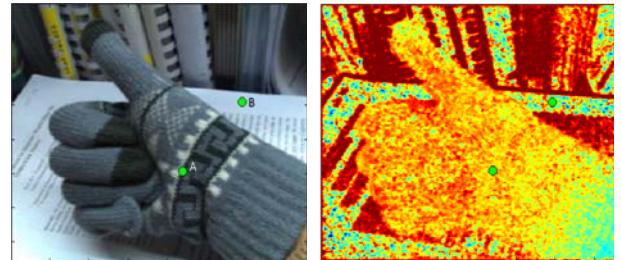


그림 2. 비용 그래프

신뢰도가 높은 깊이 정보로부터 신뢰도가 낮은 깊이 정보를 추정하기 위해서는 신뢰도가 높은 기준 픽셀을 선택하는 과정이 필요하다. 보통 신뢰도

높은 깊이 정보를 가지는 화소의 비용 그래프는 기울기가 크고 최저 비용과 최고 비용 값의 차이가 크다. 그림 2에서 이러한 특징을 잘 보여주고 있다. 화소 A는 특징 정보가 뚜렷한 화소이고 화소 B는 특징 정보가 모호하여 깊이 정보 추정이 쉽지 않은 화소이다. 이 때 화소 A의 경우 비용 그래프의 기울기 값이 크고 화소 B의 경우 비용 그래프의 기울기 값이 작으며 최대 비용과 최소 비용의 차이가 크지 않음을 확인할 수 있다. 이러한 특징으로부터 신뢰도 높은 깊이 정보를 가지는 화소가 정의되었다. 최대 최소 비용의 차이가 크지 않으면서, 사전에 정의된 2차 곡선(그림 2에서 빨간색 그래프)보다 2차 계수 값이 작은 비용 그래프는 신뢰도가 낮은 화소로 판단하여 걸러내졌다.

3.2 깊이 정보 전파

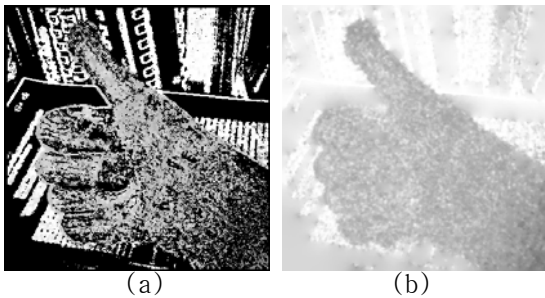


그림 3. (a) 신뢰도 높은 깊이 정보만 남은 깊이지도 (b) 신뢰도 높은 깊이 정보로부터 추정된 초기 깊이지도

3.1 절에서 선택된 신뢰도가 높은 깊이 정보를 전파하기 위해 색상 정보가 유사한 곳에서 깊이 정보 또한 유사하다는 가정이 사용되었다. 이러한 가정을 만족하는 깊이 지도를 구하기 위해, 다음 식이 정의되었다.

$$E_{init} = \sum_p (D(p) - \sum_{n \in N(p)} w_{RGB} D(n)) \quad (3)$$

$N(p)$ 는  $p$ 의 인접 화소들,  $w_{RGB}$ 는  $p$ 의 인접 화소 색상에 의해 결정되는 가중 함수이다.  $w_{RGB}$ 는 다음과 같이 정의된다.

$$w_{RGB} = e^{-\frac{(R(p)-R(n))^2 + (G(p)-G(n))^2 + (B(p)-B(n))^2}{2\sigma^2}} \quad (4)$$

신뢰도가 높은 깊이 정보는 그대로 두고, 그 이외

의 화소에서만 깊이 정보를 수식 3에 기반 하여 추정하면 매트릭스 역함수에 의해 깊이 지도를 계산할 수 있다. 그림 3에서 이렇게 추정된 깊이 지도를 보여준다.

4. 최적화

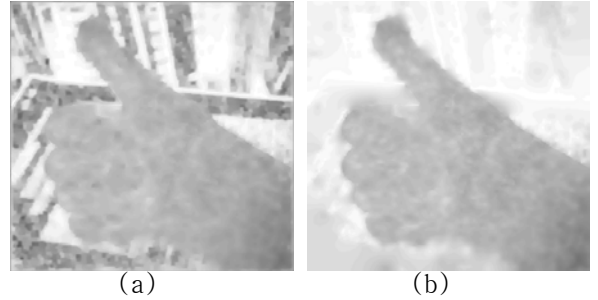


그림 4. (a) 비용 체적에서 구한 초기 깊이 지도를 사용하여 최적화 한 경우 (b) 3장에서 추정된 깊이 지도를 사용하여 최적화 한 경우

최종 깊이 지도를 구하기 위해 인접 화소의 깊이 정보 차이와 비용 체적 내에 저장된 비용 값을 동시에 최소화하는 수식이 제안되었다. 제안된 수식은 다음과 같다.

$$E_{opt} = \sum_p (\lambda_s w_{grad} \|\nabla d(p)\|_1 + C(p, d)) \quad (5)$$

$\lambda_s$ 는 인접 화소 깊이 정보 유사성의 비중을 결정하는 매개변수,  $w_{grad}$ 는 인접 화소 색상 정보 유사성을 나타내는 가중 함수,  $d(p)$ 는  $p$ 에서의 깊이 정보,  $C(p, d)$ 는  $p$  화소가 임의의 깊이  $d$ 에서 갖는 비용을 의미한다.  $w_{grad}$ 는 다음과 같이 정의 된다.

$$w_{grad} = e^{-\frac{(\|\nabla R(p)\|_2^2 + \|\nabla G(p)\|_2^2 + \|\nabla B(p)\|_2^2)}{\sigma_{grad}}} \quad (6)$$

수식 5는 비-볼록 함수이기 때문에 벌칙 항 (Penalty term)을 추가함으로써 식을 두 개로 나누고, 이 두 식의 해가 수렴할 때까지 곱으로써 해를 얻었다. 벌칙 항에 의해 정의된 두 식은 다음과 같다.

$$d^* = \arg \min_d \sum_p w_r C(p, d) + \frac{1}{2\theta} \|d - z\|_2^2 \quad (7)$$

$$z^* = \arg \min_z \sum_p \lambda_s w_{grad} \|\nabla z(p)\|_1 + \frac{1}{2\theta} \|d - z\|_2^2 \quad (8)$$

수식 7에서는 모든 깊이 후보 중 식의 값이 가장 작은 값을 갖는 깊이 정보가 선택되었다. 수식 8을 풀기 위해서는 Goldstein과 Osher에 의해 소개된 split Bregman anisotropic TV denoising 기법[5]이 사용되었다.

수식 5는 비-블록 함수이기 때문에 초기 깊이 지도에 최종 최적화 결과 값이 많은 영향을 받는다. 3장에서 구해진 초기 깊이 지도가 수식이 잘못된 해로 수렴하는 것을 막아주었다. 그림 4 에서 보이는 바와 같이, 3장에서 추정된 신뢰도 높은 깊이 지도는 결과 값이 잘못된 해로 수렴하지 않도록 도움을 주었다.

5. 실험 결과 및 분석

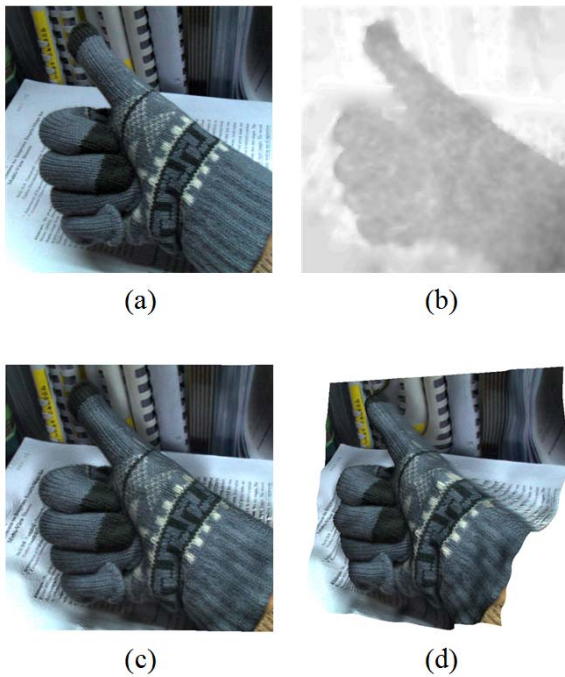


그림 5. 실험 결과 - 장갑

그림 5에서 제안된 방식에 의해 구해진 깊이 지도와 실험 결과를 보이고 있다. (a)는 라이트 필드 카메라로 촬영된 장면을, (b)는 제안된 방법을 통해 얻어진 깊이 지도를, (c)는 계산된 깊이 지도로 얻은 3차원 복원 결과를, (d)는 시점을 바꾸어 바라본 (c)의 결과를 보여주고 있다.

6. 결론

본 연구에서는 상용 라이트 필드 카메라로부터 깊이 지도를 안정적으로 추정하는 방법론을 제시하

였다. 비용 그래프 분석을 통해 신뢰도 높은 깊이 정보를 가지는 화소를 구분해내었으며 그로부터 신뢰도 높은 깊이 지도를 추정함으로써 목적함수의 최적화 성능을 높였다. 실험 결과 라이트 필드 카메라로부터 올바른 깊이 정보를 추정하는 것에 성공하였다.

감사의 글

이 논문은 2013년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2010-0028680).

참고문헌

- [1] Z. Yu, X. Guo, and J. Yu, "Line assisted light field triangulation and stereo matching," in ICCV 2013.
- [2] Sven Wanner and Bastian Goldluecke, "Globally consistent depth labeling of 4d light fields," in Computer Vision and Pattern Recognition(CVPR), IEEE, 2012, pp.41-48.
- [3] M. W. Tao, S. Hadap, J. Malik, and R. Ramamoorthi, "Depth from combining defocus and correspondence using light-field cameras," in ICCV, 2013.
- [4] Donald G Dansereau, Oscar Pizarro, and Stefan B Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," in CVPR 2013.
- [5] Tom Goldstein and Stanley Osher, "The split bregman for l1-regularized problems," SIAM Journal on Imaging Sciences, vol.2, no. 2, p323-343, 2009.