

## 일반적 총 변이와 가이드 깊이맵을 이용한 스테레오 정합

백으뜸, 호요성  
 광주과학기술원  
 {eutteum, hoyo}@gist.ac.kr

### 요 약

스테레오 정합은 컴퓨터 비전 분야에서 활발하게 연구되고 있는 연구 주제로 서로 다른 위치에서 획득된 두 영상을 정합하여 거리 정보를 얻는 방법이다. 이 방법은 초음파나 레이저를 광원으로 거리를 측정하는 것보다 실제 응용 환경의 제약을 적게 받아 다양한 분야에서 응용되고 있다. 하지만, 텍스처가 반복되거나 텍스처가 없는 영역 혹은 객체의 경계 부근에서 정확한 깊이 정보를 획득하지 못한다는 단점이 있다. 본 논문은 일반적 총 변이와 가이드 깊이맵을 사용하여 정합 비용을 정제 방법을 사용하여 정확한 깊이 정보 획득 방법을 제안한다. 실험 결과를 통해 제안한 방법이 기존의 색상 영상의 텍스처 복사 문제를 해결하였으며, 기존의 방법에 비해 bad pixel rates 측면에서 월등한 성능을 보이는 것을 확인하였다.

### 1. 서론

스테레오 정합은 컴퓨터 비전 분야에서 중요한 과정으로 3 차원 복원, 3 차원 지도 제작, 공장 자동화, 그리고 의료 영상과 같은 다양한 분야에서 사용되고 있다. 일반적으로 스테레오 정합은 4 단계로 이루어진다. 정합 비용 계산 단계, 정합 비용 정제 단계, 최적화/ 변위 획득 단계, 그리고 후처리 단계이다.

스테레오 정합 방법은 지역적 방법과 전역적 방법으로 나눌 수 있다. 전역적 방법은 정합 비용의 전체적인 에너지를 고려하여 깊이 정보를 예측하는 방법이다. 전역적 에너지 함수를 최적화하는 방법은 동적 프로그래밍, 그래프 컷, 상수 공간 신뢰 확산, 그리고 세미 글로벌 매칭 등이 사용된다. 지역적 방법은 화소 혹은 블록 사이의 상관 관계를 계산하여 최소의 비용을 찾는 방법이다. 일반적으로 지역적 방법은 거리 차 제곱 합계(SSD: sum of squared differences), 절대 거리 차 합계(SAD: sum of absolute differences), 평균 상관관계수(NCC: normalized cross correlation)과 센서스 변환(census transform) 등이 사용된다. 지역적 탐색 방법은 전역적 탐색 방법에 비해 빠르게 깊이 영상을 획득할 수 있는 장점이 있지만, 균일한 화소값을 갖는 평탄한 영역이나, 반복되는 패턴을 갖는 영역, 혹은 폐색 영역에서 정확한 대응점을 찾는 데 한계가 있다. 이러한 문제점을 해결하기 위해 정합 비용을 계산한 후에 비용을 정제하는 방법들도 많이 소개되고 있다. Yoon et al. 은 결합형 양방향 필터 [1]를 사용해 깊이를 정제하는 방법으로 정합 비용을 정제하는 데 효과적이지만, 시간이 오래 걸리고 텍스처가 복사되는 문제가 있다. He et al.은 가이드 이미지 필터 [2] 사용해 깊이를 정제하였는데, 결합형 양방향 필터보다 속도가 월등히 빠르고, 정확한 깊이를 획득하는 것을 알 수 있다. Zheng et al.가 최근에 제안한 교차 계층 구조를 이용하여 비용을 정제하는 방법 [3]으로 크기가 다른 볼륨끼리의 관계를 고려하여 정합 비용을 정제하는 방법으로, 텍스처가 없는 부분에서 정확한 깊이 정보를 획득하는 장점이 있지만, 텍스처가 반복되는 부분에서 부정확한 깊이 정보를 획득하는 단점이 있다.

위에서 제기한 문제를 해결하기 위해 본 논문은 일반적

총변이 (total generalized variation) [4]와 색상 영상과 깊이 맵으로 구한 텐서(tensor)로 에너지 함수로 비용을 정제하는 방법을 제안한다.

### 2. 일반적 총변이를 사용한 정합 비용 정제



그림 1. 제안한 방법의 흐름도

그림 1 은 제안한 방법의 전체적인 흐름도를 보여 준다. 먼저, 스테레오 영상으로 정합 비용을 계산하는데, 색상 차이와 변화량의 차이의 합으로 구한다.

$$C(p, d) = \lambda \cdot \min(T_c, C_{AD}(d)) + (1 - \lambda) \cdot \min(T_g, C_{GD}(p, d)) \quad (1)$$

식 (1)의 C(p,d)는 변위 d 에 따른 픽셀 p 의 위치에서의 비용 값이고, T<sub>c</sub>와 T<sub>g</sub> 절단 값을 나타낸다. 정합 비용을 계산한 후에 가이드 이미지 필터로 비용을 정제하여 가이드 깊이맵 d<sub>g</sub>를 만든다.

$$C_G(p, d) = W_p^G C(p, d) \quad (2)$$

식 (2)의 C<sub>G</sub>(p,d)는 정합된 정합 비용이고, W<sub>p</sub><sup>G</sup> 는 가이드 이미지 필터의 커널 비용을 나타낸다. 가이드 깊이맵 d<sub>g</sub> 와 색상 영상 I<sub>r</sub> 로 얻어진 텐서와 일반적 총변이 (total variation generalized: TGV)로 에너지 함수를 구한다. TGV 기반의 에너지 함수의 데이터 항은 각각의 비용 레벨의 형태를 유지하게 하고, 두 번째 항은 비용 레벨의 구조를 유지하면서 비용을 평탄화는 항이다.

$$\min_{u,v} \left\{ \int_{\Omega} |u - C(\rho, d)| dx + \alpha_1 \int_{\Omega} |w_d (\nabla u - v)| dx + \alpha_0 \int_{\Omega} |\nabla v| dx \right\} \quad (3)$$

식 (3)은 최종 에너지 함수로써,  $u$  는 최종결과이고,  $v$  는 색상 영상의 1 차 미분 벡터 필드,  $w_d$  는 텐서 합이다.  $\alpha_0$  와  $\alpha_1$  는 1 차 미분 항과 2 차 미분 항의 균형을 잡아주는 값이다.

$$w_d = \alpha_T TC_c + (1 - \alpha_T) TC_d \quad (4)$$

$TC_c$  와  $TC_d$  는 색상 영상과 깊이 맵의 텐서이고,  $\alpha_T$  는 가중치 값이다.

$$TC_c = \exp(-\beta |\nabla I_r|^{\gamma}) nn^T + n^{\perp} n^{\perp T} \quad (5)$$

$$TC_d = \exp(-\beta |\nabla I_g|^{\gamma}) nn^T + n^{\perp} n^{\perp T}$$

식 (5)의  $N$  은 영상 변화량의 정규화된 방향이고,  $\beta$  와  $\gamma$  는 각각 규모를 조정하고, 텐서의 가파름을 조절한다. 최종 에너지 함수를 구하기 위해 primal-dual 최적화 방법을 사용한다.

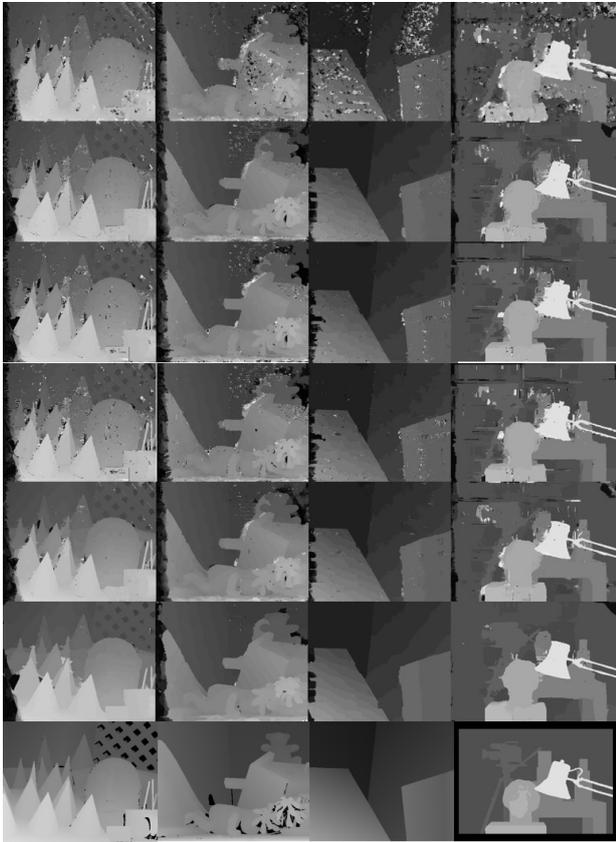


그림 2. 제안한 방법의 흐름도. 첫째 줄은 box filter 를 사용한 결과, 둘째 줄은 bilateral filter, 셋째 줄은 non-local filter, 넷째 줄은 segmented tree aggregation, 다섯 째 줄은 guided filter, 여섯 번째 줄은 제안한 방법, 그리고 마지막 줄은 ground truth 영상이다.

### 3. 실험 결과

제안한 방법을 평가하기 위해 본 논문에서는 Middlebury 에서 제공하는 데이터 셋을 사용하였다 [2]. 그림 2 는 실험한

결과를 보여주는데 기존의 비용 정제 방법과 제안한 비용 정제 방법의 결과를 보여주고 있으며, 마지막 줄은 ground truth 영상이다. 실험의 공정한 평가를 위해 스테레오 매칭 이외의 후처리는 사용하지 않았다. 그림 2 에서 나타나듯이, 제안한 방법이 텍스처 영역과 비 텍스처 영역에서 기존의 방법에 비해 잡음이 적고, 정확한 깊이 정보를 갖는 것을 알 수 있다. 또한, 본 논문에서는 BPR 을 사용하여 정성적 평가를 하였는데, 기존의 방법에 비해 제안한 방법이 좋은 성능을 얻은 것을 보여준다.

표 1. BPR 결과 비교

	tsukuba		venus		cones		teddy	
	non occ	all	non occ	all	non occ	all	non occ	all
box	14.54	16.47	16.26	17.61	11.2	21.26	15.69	23.95
bilateral [1]	6.08	7.11	2.03	2.82	7.14	15.47	9.37	16.75
non-local [5]	5.46	6.54	2.58	3.38	7.19	16.4	8.21	15.6
segmented [6]	6.31	7.53	3.18	4.07	8.22	17.56	9.43	16.9
guided [2]	5.77	7.06	2.03	3.1	<b>4.38</b>	13.06	7.61	15.4
proposed	<b>3.23</b>	<b>4.08</b>	<b>1.02</b>	<b>1.9</b>	5.32	<b>12.88</b>	<b>7.01</b>	<b>13.5</b>

### 4. 결론

본 논문은 일반적 총 변이와 가이드 깊이맵의 정보로 정합 비용을 정제하여 정확한 깊이 정보를 획득하는 스테레오 매칭 방법을 제안한다. 가이드 깊이맵과 색상 영상을 활용하여 텐서 값으로 일반적 총변이 기반의 에너지 함수를 만들고, primal dual 최적화 방법으로 정합 비용을 정제하였다. 제안한 스테레오 매칭 방법을 통해 기존의 방법보다 정확한 깊이 맵을 생성하였으며, BPR 측면에서도 월등한 성능을 보이는 것을 확인하였다.

### 감사의 글

본 논문은 민 · 군기술협력사업 (Civil-Military Technology Cooperation Program) 으로부터 지원을 받아 수행된 연구임

### 참고문헌

- [1] Yoon K. J., Kweon I. S.: 'Adaptive support-weight approach for correspondence search', Trans. Pattern Analysis and Machine Intelligence, 28(4), 2006, pp. 650-656.
- [2] He, K., Sun, J., Tang, X.: 'Guided image filtering', In European conference on computer vision, 2010, pp. 1-14.
- [3] Zheng K., Fang Y., Min D., Sun L., Yang S., Yan S., Tian Q.: 'Cross-scale cost aggregation for stereo matching', Conf. Computer Vision and Pattern Recognition, 2014, pp. 1590-1597.
- [4] Bredies K., Kunisch K., Pock T.: 'Total generalized variation', SIAM J. Imaging Sci., 2010, 3(3), pp. 492-526.
- [5] Yang Q.: 'A non-local cost aggregation method for stereo matching', Conf. Computer Vision and Pattern Recognition, 2012, pp. 1402-1409.
- [6] Mei, X., Sun, X., Dong, W., Wang, H., Zhang, X.: 'Segment-tree based cost aggregation for stereo matching', Conf. Computer Vision and Pattern Recognition, 2013, pp. 313-320.