

그래프 컷 커널을 이용한 스테레오 대응

이용환*·김영섭**†

*원광대학교 디지털콘텐츠공학과, **단국대학교 전자전기공학과

Stereo Correspondence Using Graphs Cuts Kernel

Yong-Hwan Lee* and Youngseop Kim**†

*Dept. of Digital Contents, Wonkwang University

**† Dept. of Electronic and Electrical Engineering, Dankook University

ABSTRACT

Given two stereo images of a scene, it is possible to recover a 3D understanding of the scene. This is the primary way that the human visual system estimates depth. This process is useful in applications like robotics, where depth sensors may be expensive but a pair of cameras is relatively cheap. In this work, we combined our interests to implement a graph cut algorithm for stereo correspondence, and performed evaluation against a baseline algorithm using normalized cross correlation across a variety of metrics. Experimental trials revealed that the proposed descriptor exhibited a significant improvement, compared to the other existing methods.

Key Words : Stereo Correspondence, Graphs Cuts, Markov Random Field, Energy Function, Alpha-Expansion, Min-cut Graph

1. 서 론

두 장의 스테레오 장면 영상이 주어지면 3D 장면 영상으로 생성할 수 있다. 이러한 개념은 인간 시각 시스템이 깊이를 추정하는 주요 방법을 제공한다. 이러한 과정과 기술은 깊이 센서를 갖는 3D 카메라가 높은 비용을 가지기 때문에, 한 쌍의 카메라를 통해 비교적 저렴한 가격으로 3D 장면 영상을 생성하여 기계 또는 로봇에 적용하는 응용 분야에서 매우 유용하다.

스테레오 영상에서 깊이 맵을 생성하기 위해서는 스테레오 대응 문제를 해결해야 한다. 스테레오 대응 문제는 전통적으로 영상 처리 및 컴퓨터 비전 분야에서 많이 주목되는 주제 중에 하나이며, 최근까지도 활발하게 연구되고 있는 영역이기도 하다 [1]. Boykov는 그래프-컷(Graph-Cuts) 알고리즘을 통해 네트워크 흐름(Network flow) 문제에 대한 에너지 최소화(Energy minimization) 문제를 해결하는

프레임워크를 제안하였다 [2]. 컴퓨터 시각의 다양한 문제점들이 에너지 최소화 문제와 유사한 형태로 공식화될 수 있기 때문에, 그래프-컷 기법은 영상 복원(Image restoration), 스테레오 대응(Stereo correspondence) 및 영상 분할(Segmentation)과 같은 광범위한 범위에서 응용 알고리즘으로 적용되고 있다 [3-4].

본 논문에서는 스테레오 대응을 위한 그래프-컷 알고리즘을 구현하기 위한 매개변수를 실험적으로 분석하고, 다양한 매트릭스를 통해 정규화된 상호-상관 관계(NCC, Normalized Cross-correlation)를 사용하는 기법을 제안한다. 특히, 그래프-컷 알고리즘에 대한 불일치 레이블과 교합 처리의 효과를 실험적으로 평가한다. 실험에서는 사전에 정제된 공용의 데이터셋을 대상으로 수행하여 성능 측면에서 우월한 기법임을 검증한다.

본 논문의 구성은 다음과 같다. 2장에서 제안 알고리즘과 관련된 기술을 살펴보고, 3장에서 제안하는 스테레오 대응 알고리즘을 설명한다. 4장에서는 구현한 제안 알고리즘을 실험적으로 분석하고, 5장에서 결론과 향후 연구 방향으로 마무리한다.

†E-mail: wangcho@dankook.ac.kr

2. 관련 연구

최근, 스테레오 대응(Stereo Correspondence) 문제를 해결하기 위한 다양한 기술과 방법이 연구되고 있다 [5]. 이러한 연구 방법은 크게 지역적 방법과 전역적 방법으로 분류된다. 지역적 방법은 각 대상 픽셀 주변에 윈도우 크기가 작은 윈도우를 고려한 문제 해결 방법이며, 전체 윈도우 영역에 대해 계산된 비용 함수에 특정 스무딩(Smoothing) 제약 조건을 부여하여 코딩한다. 비용함수를 최소화함으로써, 주어진 픽셀들을 대상으로, 가장 좋은 불일치 값을 선택한다. Adaptive Support Weight [6]와 Slanted Window [7]에서와 같이 다양한 최적화 기법이 제안되었지만, 지역적 방법은 매칭 과정에서 교합을 처리하는데 문제점을 가지고 있다. 정규화된 상호-상관 기법(Normalized Cross-correlation, NCC)은 지역적 방법의 대표적인 사례이다. 전역적 방법은 스무딩(Smoothness)과 교합(Occlusion) 과정에서 제한적 요소를 갖는다는 차이점이 있다. 이러한 제한 요소는 주어진 불일치 맵(Disparity map)의 전역 에너지 함수(Global energy function)를 직접 인코딩하기 때문에 발생한다. 전역 에너지 함수를 최소화하는 불일치 맵을 생성하기 위해 다양한 최적화 기법이 활용된다. 에너지 함수의 최적화는 NP-문제이지만, 스무딩 항목을 정확하게 선택함으로써, 동적 프로그래밍을 통해 최소화 문제를 효율적으로 계산할 수 있게 된다 [8]. 그래프-컷(Graph cut) 알고리즘이 전역적 방법들 중에 대표적인 방법이며, 일반적인 다른 기법에 비해, 에너지 함수를 최적화하는 데에 많이 적용되고 있다. 최근 LogCut [9]와 같은 성능 또는 런타임을 개선하기 위해 기존의 그래프-컷 알고리즘을 변형하는 다수의 논문이 발표되었다. 그래프-컷 알고리즘의 핵심 단계인 알파-확장(Alpha-Expansion) 기법이 최근 다양한 연구 논문에서 인용되고 있으며, 알고리즘에 대한 최적화 엔진으로 널리 사용되고 있다.

3. 제안 알고리즘

본 절에서는 그래프-컷 커널의 기본 흐름과 이를 이용하여 제안하는 알고리즘을 보다 상세하게 다룬다. 제안 알고리즘의 첫 번째 단계는 마코프 랜덤 필드(MRF, Markov Random Field)에서 에너지 최소화를 재구성하고, 최소화 컷을 사용하여 최적의 레이블링을 구성하여 알파-확장을 반복적으로 수행한다. 이러한 제안 알고리즘의 흐름은 Fig.1과 같다.

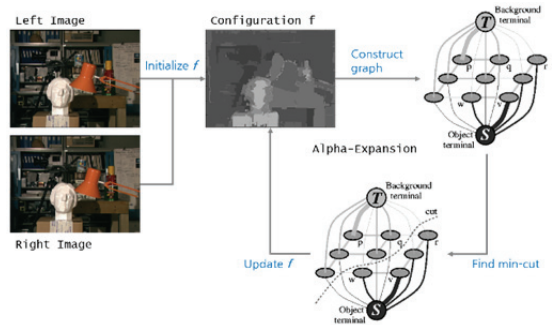


Fig. 1. Diagram of the proposed method.

3.1 마코프 랜덤 필드(Markov Random Field)

참조 영상의 각 픽셀을 마코프 랜덤 필드(MRF)에서 하나의 노드(Node)로 간주하면, 유일한 불일치 값(Disparity Value)을 레이블로 사용할 수 있다. MRF는 유한 개의 레이블 집합을 가지기 때문에, 불일치 값에 대한 이산 집합을 선택해야 한다. 제안 알고리즘에서는 입력으로 받는 좌영상과 우영상에 대한 GTS(Ground truth sets) 불일치 맵으로부터 불일치의 최대값(d_{max})과 최소값(d_{min})을 계산하고 $[d_{max}, d_{min}]$ 을 모든 불일치 값에 대한 거리로 선택한다.

레이블 집합을 정의하는 것뿐만 아니라, MRF에 대한 인접 시스템을 선택해야 하며, 제안 알고리즘에서는 4개의 인접 시스템을 적용한다. 스테레오 대응(Stereo correspondence)을 위한 마코프 랜덤 필드를 정의하기 위해, 특정 에너지 함수를 최소화하는 마코프 랜덤 필드의 최적화된 레이블 설정을 정의한다.

3.2 에너지 함수(Energy Function)

레이블 설정으로 가능한 집합은 수식 (1)과 같이 나타낼 수 있다.

$$A = \{(p, q) | p_y = q_y, 0 \leq q_x - p_x < k\} \quad (1)$$

영상 전체에서 유일한 불일치 설정 f 에 대한 에너지 함수는 수식(2)과 같이 정의된다[2].

$$E(f) = E_{data}(f) + E_{occ}(f) + E_{sm}(f) \quad (2)$$

E_{data} 는 대응 픽셀들간의 강도(Intensity) 차이비용을 나타내며, E_{occ} 는 폐색된 픽셀에 대한 일정 패널티 비용을 의미한다. E_{sm} 은 이웃 픽셀들이 유사한 불일치치를 갖게 만드는 제약조건으로 추가된 항목이다. 특히, 하나의 레이블 설정이 구성되고 동일한 불일치 값을 갖는 인접 픽셀의 설정이 구성상에 존재하지 않는 경우에는 일정한 비용이 발생하게 된다.

3.3 알파-확장(Alpha-Expansion)

다음 단계의 업데이트를 위해, 먼저 가능한 모든 할당 집합으로 알파-확장을 정의한다. 제약조건은 이전 단계에서 활성화된 집합 및 할당에 대한 불일치 알파가 결합된 형태로 나타난다. 즉, 반복할 때마다 활성 할당을 취소하거나 새로운 불일치 알파를 할당해야 한다. 제안 알고리즘에서는 새로운 할당이 에너지 함수보다 낮을 때까지 반복적으로 이를 수행한다.

유일한 설정을 초기화 한 다음, 전역 에너지 함수를 최소화하기 위해 알파-확장을 수행하여 설정을 반복하여 수정한다. 이때, 모든 알파 값에 대해 알파-확장을 수행한다. 모델이 더 이상 변하지 않을 때까지 알파 값을 계속적으로 순환시키며, 지역 최적 값에 도달하면 알고리즘을 종료한다. 이러한 과정을 수행하는 데 있어, 알파-확장을 수행하는 동안에 에너지 함수를 최소화하는 설정을 찾는 것이 문제가 될 수 있다. 이는 다음 절에 기술하는 최소컷(Min-Cut) 그래프를 통해 해결한다.

3.4 최소-컷(Min-Cut) 그래프

반복되는 모든 단계에서 Table 1과 같이 표기법을 정의한다.

Table 1. Cost model terminology

A^0	현재 설정에서 활성 할당 집합
A^∞	불일치 알파를 가진 잠재적 할당 집합
$D(a)$	할당에 따른 데이터 비용
$D_{occ}(a)$	할당에 따른 교합 비용
$D_{sm}(a)$	할당에 따른 혼합의 부드러움 비용
V	혼합의 부드러움 상수
C_p	교합 상수

최소-컷 그래프를 생성하는 과정은 다음과 같이 기술된다.

- (1) 소스와 싱크 노드를 생성;
- (2) 노드로 설정된 각 대응 할당을 그래프에 삽입;
- (3) Table 2에 명시된 경계선 가중치에 따라 두 노드 사이의 모서리와 무게를 그래프에 삽입;

이 그래프에서 최소-컷이 어떻게 에너지를 최소화하는 최적의 설정과 동일하게 되는지는 [3]을 참조한다.

3.5 갱신 설정

앞 절에서 설명한 그래프를 설정한 다음, 네트워크 흐름 루틴을 호출하여 그래프의 최소-컷을 계산한다. 산출된

Table 2. Edge weights

Edge	Weight	for
(s, a)	$D_{occ}(a)$	$a \in A^0$
(a, t)	$D_{occ}(a)$	$a \in A^\infty$
(a, t)	$D(a) + D_{sm}(a)$	$a \in A^0$
(s, a)	$D(a)$	$a \in A^\infty$
$(a1, a2)(a2, a1)$	$V_{a1, a2}$	$\{a1, a2\} \in N,$ $a1, a2 \in \bar{A}$
$(a1, a2)$	∞	$p \in P, a1 \in A^0,$ $a2 \in A^\infty,$ $a1, a2 \in N_p(\bar{f})$
$(a2, a1)$	C_p	$p \in P, a1 \in A^0,$ $a2 \in A^\infty,$ $a1, a2 \in N_p(\bar{f})$

최소-컷으로부터, 새로운 활성 할당을 검색하고, 이에 따라 새롭게 갱신되는 정확한 설정으로 업데이트한다.

4. 실험

제안 알고리즘의 성능을 평가하기 위해, 이미지와 GTS (Ground Truth Sets) 불일치 값을 포함하는 실험 데이터셋은 Middlebury Stereo Vision 홈페이지(<http://vision.middlebury.edu/stereo/>)에서 다운받았다. GT 불일치 맵 상에서의 교합은 제공되지 않으며, 물리적 개념에 따라 계산하였다. 다른 영상의 경계 외부에 있는 픽셀에 대응하는 픽셀은 교합되는 레이블로 지정한다. 다른 불일치(임계값 2)를 갖는 픽셀에 대응하는 픽셀은 교합 (Occlusion)으로 레이블을 설정한다. 본 절에서 GT 불일치 맵을 사용하여 제안하는 스테레오 대응 알고리즘의 성능을 객관적으로 평가한다. 제안 알고리즘에 대한 실험 결과는 Fig. 2와 같다.

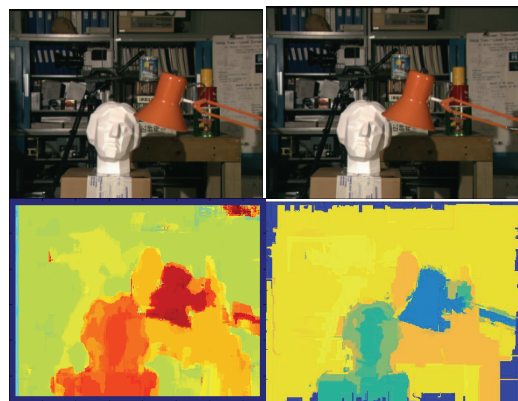


Fig. 2(a). Disparity algorithm results, upper left: input left image, upper right: input right image, lower left: NCC algorithm, lower right: proposed method.

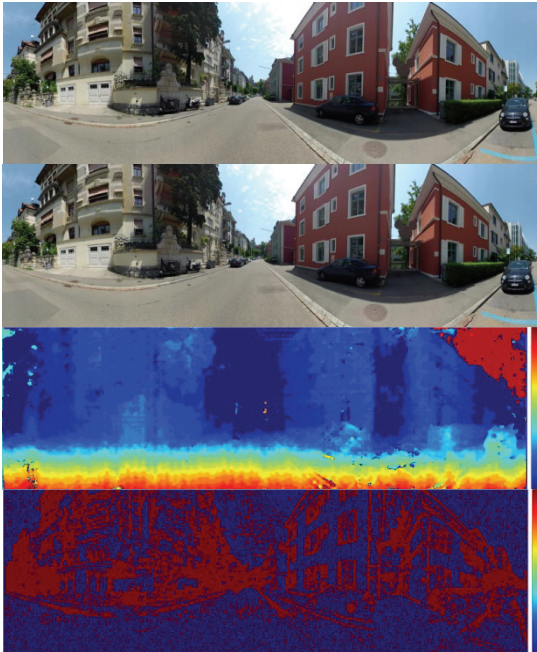


Fig. 2(b). Disparity algorithm results, first: input left image, second: input right image, third: NCC algorithm, fourth: proposed method.

정규화된 상호-상관 알고리즘의 경우, 알고리즘을 통해 자동 생성된 불일치 맵을 GTS와 시각적으로 대략 비교를 하면, NCC는 장면의 일반적인 느낌을 캡처하지만, 많은 양의 잡음을 포함하고 있다는 사실을 알려준다. 템플릿 매칭할 때 발생하는 잡음은 대응을 트리거 하기에는 너무 신호가 약하다. Fig 2에 보여진 두 번째 영상 집합을 살펴보면, 매끄러운 표면에서 특히 많은 잡음이 발생하였다. 실험의 첫 번째 영상 집합을 살펴보면, 반복적인 패턴은 영상의 여러 위치에서 대응이 생성되며, 관찰되는 단순 대응을 유도하게 된다. 그러나 알고리즘의 실행 간편성을 고려할 때, 수용할 수준의 성능을 보였다.

그래프-컷 차이를 살펴보면, 3가지 주요 특성을 알 수 있다. 그래프-컷 알고리즘은 대략적인 값을 계산하지만, 불일치 값의 급격한 증가를 나타내며, 가려진 픽셀의 가는 선을 보여준다. 경계선에 대해서는 대응하는 픽셀들이 서로 대응해야 한다는 제약을 갖는다. 이러한 일정한 불균형 영역의 경계에서 오른쪽 또는 왼쪽 중에 하나는 정합된다는 것을 알 수 있다.

Table 3에서 실험 데이터셋에 대한 표준과 표준편차의 실험 결과를 비교하였다.

Table 3. Benchmark results

	NCC	GC	GC Ref
Gross Errors	72 ± 5%	16 ± 12%	1.9%
False Negative Rate	68 ± 12%	26 ± 17%	42.6%
False Positive Rate	3.7 ± 0.8%	4.8 ± 3.7%	1.1%
Bias	0.8 ± 3.8	-1.0 ± 1.8	-
R ²	0.25 ± 0.53	0.76 ± 0.31	-

NCC 알고리즘의 긍정적인 편향은 높은 불균형으로 값을 표시한다. 이는 실험 영상을 육안으로 검사하여 지원된다. 그래프-컷 알고리즘의 네거티브 편향은 높은 불일치를 갖는 값 또는 장면의 전경을 검출하지 못한다. 전경 요소는 급격하게 변하는 경향을 보이며 알고리즘을 부드럽게 변화되는 부분에서 높은 불일치가 낮은 값으로 클리핑된다.

그래프-컷 알고리즘은 서로 다른 3개의 비용함수인 픽셀 매칭, 교합, 스무딩에 영향을 받는다. 이러한 3개 함수의 상대적 중요성은 2개의 변수(C_p, I)에 영향을 받는다. C_p 는 픽셀에 교합되는 비용을, I 는 불일치 값에서 불연속성에 대한 비용을 나타낸다. 2개 매개변수는 자연스러운 영상 통계에 따라 달라지는 픽셀 매칭 비용 함수로 조정된다.

본 실험에서는 교합의 오류 레이블과 불일치의 오류 레이블간의 최적 균형을 맞추기 위해, C_p 에 200, I 에 200를 적용하였다. NCC 알고리즘에는 윈도우 크기가 영향을 미친다. 윈도우 크기가 클수록 정규화가 보다 세밀하게 적용되어 영상은 더 매끄러워진다. 그러나, 윈도우 크기가 크면 템플릿의 다른 영역이 매칭 과정에서 일치되지 않기 때문에 날카로운 모서리 검출이 어렵게 된다. 실험에서 윈도우 크기는 9픽셀일 때 가장 좋은 성능을 보였다.

실험을 통해, 교합 비용이 너무 낮으면 전체 영상이 교합 전으로 표시되어 여러 유형의 오류가 발생한다는 것을 알 수 있다. 교합을 회피하기 위해 너무 강조하게 될 경우, 교합된 픽셀의 레이블을 지정하고 오류가 발생할 수도 있다. 이와 마찬가지로, 스무딩이 너무 적게 적용되면, 불일치 레이블에 잡음이 발생한다. 스무딩이 너무 크게 적용되면 결과가 줄어들고 시각적 효과가 많이 낮아짐을 실험결과로 알 수 있었다.

5. 결론

2장의 좌우 영상을 입력 받아 3D 영상을 생성하는 다양한 연구가 진행되고 있으며, 이러한 스테레오 영상에서

깊이 맵을 생성하기 위한 스테레오 대응 문제를 해결하기 위한 방법이 연구되고 있다. 본 논문에서는 스테레오 대응을 위한 그래프-컷 알고리즘을 구현하기 위한 매개 변수를 실험적으로 분석하고, 다양한 매트릭스를 통해 NCC 관계를 사용하는 기법을 제안하였다. 실험 결과, 제안 알고리즘이 기존의 알고리즘보다 성능적으로 개선된 효과를 얻었음을 보였다.

감사의 글

본 연구는 2016년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(과제번호: 2015R1D1A1A01061152).

참고문헌

1. D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two Frame Stereo Correspondence Algorithms," *International Journal of Computer Vision*, 2002.
2. Y. Boykov and V. Kolmogorov, "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision", *IEEE transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 26, no. 9, pp. 1124-1137, 2004.
3. V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts", *International Conference Computer Vision*, volume II, 508-515, 2001.
4. Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Nešić, N., Wang, X., and Westling, P. "High Resolution Stereo Datasets with Subpixel-Accurate Ground Truth", *Lecture Notes in Computer Science Pattern Recognition*, 3142.
5. M. Bleyer and C. Breiteneder, "Stereo Matching - State-of-the-Art and Research Challenges," *Advanced Topics in Computer Vision*. Springer, pp. 143-179, 2013.
6. K.-J. Yoon and I.-S. Kweon, "Locally Adaptive Support-Weight Approach for Visual Correspondence Search", *Computer Vision and Pattern Recognition*, pp. 924-931, 2005.
7. M. Bleyer, C. Rhemann, and C. Rother, "Patch match Stereo Matching with Slanted Support Windows," *Proceeding on British Machine Vision*, 2011.
8. I. J. Cox, S. L. Hingorani, S. B. Rao, and B. M. Maggs, "A maximum likelihood stereo algorithm", *Computer Vision and Image Understanding*, 63(3): 542-567, 1996.
9. Victor Lempitsky, Carsten Rother, and Andrew Blake, "Logcut-efficient graph cut optimization for markov random fields," *International Conference on Computer Vision*, 2007.
10. J. Zbontar and Y. LeCun, "Stereo Matching by Training a Convolutional Neural Network to Compare Image Patches", *Computer Vision and Pattern Recognition*, 1510.05970, 2015.
11. Mikhail G. Mozerov and Joost van de Weijer, "Accurate Stereo Matching by Two-Step Energy Minimization," *IEEE Transactions on Image Processing*, vol. 3, issue. 3, 2015.
12. H. Hirschmuller and D. Scharstein. "Evaluation of cost functions for stereo matching." *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007.
13. S. Birchfield and C. Tomasi, "A Pixel Dissimilarity Measure That Is Insensitive to Image Sampling", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 401-406, 1998.

접수일: 2017년 6월 14일, 심사일: 2017년 6월 21일,
 게재확정일: 2017년 6월 21일