

확률모델에 기반한 스테레오 정합 및 객체추출

이상화¹, 노민호², 조남익¹, 박종일²

¹서울대학교 전기컴퓨터공학부

²한양대학교 전기컴퓨터공학부

Stereo Matching and Objects Extraction Using Stochastic Models

Sang Hwa Lee¹, Min Ho Roh², Nam Ik Cho¹, and Jong-Il Park²

¹School of Electrical and Computer Engineering, Seoul National University

²Division of Electrical and Computer Engineering, Hanyang University

E-mail: lsh@jpl.snu.ac.kr

Abstract

본 논문은 확률적 확산 기법 및 확률모델을 이용하여 스테레오 영상간의 대응점을 추정하고, 영상의 배경으로부터 객체를 추출해 내는 연구를 다루고 있다. 스테레오 영상의 정합 및 객체 추출을 위하여 시차, 세그먼트, 라인, 및 오클루전 필드를 Markov random field 모델로 정의하고, 확률적 에너지 최소화 방법을 이용하여 최적의 시차 필드 및 객체추출을 수행한다. 본 논문에서는 우선 이러한 다양한 필드간의 MRF 모델링 기법을 제안하고, 각 필드에 대한 에너지 함수를 정의한다. 그리고, 확률적 확산 기법을 이용하여 각 필드에 대하여 정의된 에너지 함수를 최소화함으로써, 최적의 시차필드 및 객체추출 결과를 구한다.

I. 서론

최근의 영상신호처리에 대한 연구는 영상에 존재하는 다양한 영역을 객체단위로 분할하고, 분할된 객체를 독립적으로 처리하는 방향으로 발전하고 있다. 이러한 연구의 흐름은 MPEG과 같은 영상표현 및 압축에서부터 영상감시, 영상합성, 비디오분석, 3차원 모델링, 가상현실, 등등, 다양한 분야에서 매우 중요한 기초기술로 요구되고 있다. 따라서, 주어진 영상조건으로부터 응용분야에 적합한 객체를 효과적으로 분할하는 것은 대단히 중요하며, 특히, 배경영역으로부터 객체를 구분하고 정

확히 추출하는 것은 많은 연구의 기본 주제가 되어 왔다. 분리된 객체는 하나의 독립적인 영상으로서, 형태, 색상, 위치, 움직임 또는 깊이 등의 정보를 포함하는 layer로 정의된다 [1]. 객체를 배경으로부터 구분하는 방법은, 배경과 객체간의 색상정보를 이용하거나, 비디오 신호로부터 객체의 움직임을 추정하여 독립적인 움직임을 갖는 영역을 객체로 분리하는 방법이 널리 활용된다. 그러나, 일반적으로 추정된 움직임 정보는 부정확하고, 색상정보에 대한 정확한 해석도 불가능하기 때문에, 보다 정확하고 다양한 정보를 활용하여 영상분할을 수행하는 것이 최근의 추세이다. 특히, 확률적 접근으로서 MRF 모델과 에너지 최소화 기법을 기반으로 다양한 영상정보를 동시에 고려하여 영상분할 및 신호추정에 대한 연구가 널리 진행되고 있다 [2].

본 논문에서는 확률적 모델을 기반으로 하여 스테레오 영상으로부터 시차(disparity), 세그먼트(segment), 라인(line), 오클루전(occlusion) 정보를 동시에 추정하고, 이를 이용하여 영상의 배경으로부터 객체영역을 추출하는 기법을 다루고자 한다. 본 연구에서는 시차, 라인[2,7], 오클루전[2,4]과 세그먼트(segment)[5,6] 필드를 하나의 확률적 추정이론에서 동시에 고려하고, Markov random field (MRF) 이론에 따라 각 필드 및 필드간의 상관성을 확률적으로 모델링 한다. 배경으로부터 분리되는 객체는 시차정보 및 색상정보를 기준으로 나뉘어지며, 이들 정보의 복잡도에 따라 단일 객체

또는 다중 객체의 형태로 분할된다. 이렇게 분할된 객체는 3 차원 공간에 대한 깊이정보(시차)를 포함하고 있기 때문에, 가상현실, 가상스튜디오와 같은 3 차원 영상합성에서 매우 유용한 영상정보가 되며, 3 차원 디스플레이 시스템에서 공간의 입체감을 표현하는데 효과적인 도구가 될 수 있다.

본 논문의 II 절에서는 다양한 필드를 확률적으로 모델링하고 이를 에너지함수로 정의하여 최소화하는 구체적인 기법에 대하여 소개한다. III 절에서는 실제 스테레오 영상을 이용하여 실험한 다양한 결과를 제시한다.

II. 확률적 확산 기법

3.1 MRF-MAP 추정과 에너지 최소화 기법

시차(D), 세그먼트(S), 라인(L), 그리고, 오클루전(O) 필드를 동시에 추정하는 확률적 방법은 *maximum a posteriori* (MAP) 추정기법에 따라 다음과 같은 조건부 확률이 최대가 되도록 하는 각 필드를 찾는 것이다.

$$p(\mathbf{D}, \mathbf{S}, \mathbf{L}, \mathbf{O} | \mathbf{I}_l, \mathbf{I}_r) = p(\mathbf{I}_l | \mathbf{D}, \mathbf{S}, \mathbf{O}, \mathbf{L}, \mathbf{I}_r) \cdot p(\mathbf{D} | \mathbf{S}, \mathbf{O}, \mathbf{L}, \mathbf{I}_r) \cdot p(\mathbf{S} | \mathbf{O}, \mathbf{L}, \mathbf{I}_r) \cdot p(\mathbf{O} | \mathbf{L}, \mathbf{I}_r) \cdot p(\mathbf{L} | \mathbf{I}_r) / p(\mathbf{I}_l | \mathbf{I}_r)$$

그러나, 일반적으로 확률적으로 정의된 위의 식은, MRF 모델 및 Gibbs random field (GRF) 이론에 따라 에너지함수로 변환될 수 있다. 즉, MRF 모델에 따라 인접필드 (Neighborhood field)를 정의하고, 인접필드간의 상호작용을 에너지함수로 표현하여 Gibbs 확률분포함수를 적용하면, 확률공간의 문제는 에너지공간의 문제로 대응시킬 수 있다.

$$p(x) = \frac{1}{Z} \exp \left\{ -\frac{1}{T} \sum_{c \in C} V_c(x, y | y \in c) \right\}$$

위의 Gibbs 확률분포에서, 필드 x 에 대하여 인접하는 모든 필드 y 에 대한 상호작용을 $V_x()$ 라는 에너지 함수 (clique potential)로 표현하고 있다. 이 상호작용 에너지함수의 전체 합을 최소화 하도록 각 필드를 추정하면, 결과적으로 조건부 확률을 최대화하는 필드를 구하게 된다.

$$U(\mathbf{D}, \mathbf{S}, \mathbf{L}, \mathbf{O} | \mathbf{I}_l, \mathbf{I}_r) = U(\mathbf{I}_l | \mathbf{D}, \mathbf{S}, \mathbf{O}, \mathbf{L}, \mathbf{I}_r) + U(\mathbf{D} | \mathbf{S}, \mathbf{O}, \mathbf{L}, \mathbf{I}_r) + U(\mathbf{S} | \mathbf{O}, \mathbf{L}, \mathbf{I}_r) + U(\mathbf{O} | \mathbf{L}, \mathbf{I}_r) + U(\mathbf{L} | \mathbf{I}_r) - U(\mathbf{I}_l | \mathbf{I}_r)$$

문제는 각 필드 및 필드간의 에너지함수를 정의하는 방법과 그 에너지 함수를 최소화하는 기법이다. 이와 같이 MRF-MAP에 기반한 추정에서 에너지함수를 최소화하는 방법은 컴퓨터비전의 영역에서 다양한 방법들이

제안되어 왔다. 가장 일반적인 기법은 simulated annealing으로서, 필드를 선택하고 결정하는 방식에 따라 Metropolis 기법, Iterated Conditional mode(ICM), Gibbs sampler 등으로 구분할 수 있다 [1]. Bayesian network, mean field 이론 등도 에너지 최소화 기법이며, 최근에는 그래프 개념을 적용하여 더욱 효과적으로 최적의 필드를 추정하는 graph cut 기법이 소개되었다 [3]. 본 논문에서는 이미 저자들이 제안한 확률적 확산 기법 (stochastic diffusion [7]) 이라는 새로운 방법을 적용하여 에너지함수를 최소화한다.

확률적 확산 기법은 기존의 MRF 모델을 배경으로 하는 에너지함수 최소화 기법과는 달리, 인접하는 필드의 확률분포함수를 적용하여 인접하고 있는 필드간의 상관성을 확률적으로 모델링하는 기법이다. 예를 들어서 시차필드의 경우, 인접하는 시차필드의 확률분포를 확률적 기대치를 계산하는 과정에 도입하면, 현재위치의 시차필드의 확률분포를 계산할 수 있는데, 이러한 형태를 반복적으로 계산하여 인접하는 필드의 확률분포를 현재위치의 필드값에 반영하도록 한다.

$$p(d_{ij}) = \sum_{all \mathbf{d}_N} p(d_{ij} | \mathbf{d}_N) p(\mathbf{d}_N)$$

위의 식에서처럼 인접필드의 확률분포는 현재 필드의 확률분포에 영향을 주게 되며, 필드간의 상호작용 에너지함수를 정의하는 과정에서 일종의 가중치함수로서 작용하게 된다. 위와 같은 계산을 반복하게 되면, 각 화소의 필드는 인접하고 있는 필드로부터 확률적 상호작용을 받게 되며, 이러한 상호작용은 더욱 넓은 인접영역으로 전파된다. 따라서, 상호작용을 주고 받는 인접영역은 결국에는 임의의 형태 및 크기를 갖게 되며, 에너지공간을 확률적으로 확산시켜 안정적인 상태로 수렴시키는 역할을 수행하게 된다 [7].

3.2 확률 모델 및 상호작용 에너지함수

필드간의 상호작용 에너지함수는 MRF 모델에 따라 정의할 수 있으며, 각 필드 또는 필드간의 MRF 모델은 인접하고 있는 필드값들의 연속성과, 경계영역에서의 불연속성을 동시에 보장하는 방향으로 설계되어야 한다. 본 연구에서는 시차 및 세그먼트 필드에 대해서는 화소간에 연속성을 보장하고, 라인 및 오클루전 필드에 대해서는 다른 필드들의 불연속성을 유지시켜 주는 방향으로 에너지 함수를 정의하고 구현한다. 그리고, 오클루전 필드의 경우에는 시차필드의 불연속성뿐만 아니라, 확률공간에서 오클루전이 발생할 가능성이 높

은 영역을 관측하여 이를 오클루전 필드 모델에 반영하는 독창적인 확률모델을 제시한다.

3.2.1 정합오류 모델 (likelihood model)

정합오류 모델은 추정된 시차 및 오클루전 필드가 스테레오 영상간에 얼마나 잘 부합되는지를 평가하는 에너지 함수로서, 일반적으로 블록정합의 개념과 동일하다. 잡음과 색상정보의 부정확성을 보정하기 위해, 국소평균(m)도 고려하였다. 그리고, 오클루전에 대해서는 정합오류를 판단할 수 없기 때문에, 이를 보정하는 문턱에너지(T_0)를 설정하였다.

$$\{f_i(\mathbf{I}'_{i,j} - \mathbf{I}'_{i,j-d_j}) + (m'_{i,j} - m^r_{i,j-d_j})^2\}(1 - o_{ij}) + o_{ij}T_0$$

3.2.2 시차모델

시차필드는 확률변수이므로, 앞에서 설명한 확률적 확산의 개념을 적용할 수 있다. 또한, 인접하고 있는 시차필드간에는 일반적으로 유사한 값을 갖는 경향이 강하므로, 인접하는 두 시차필드간의 차이를 고려하여 에너지함수($V(d_{ij}, \mathbf{d}_n)$)를 계산하였다. 그러나, 라인필드가 두 필드 사이에 존재하는 경우에는 불연속성을 보장하기 위하여 두 시차필드간의 상호작용을 차단하는 형태로 에너지함수를 유도하였다.

$$-\ln \sum_{N \in S} p(\mathbf{d}_N | \mathbf{S}, \mathbf{L}) \prod_{n \in N} e^{-V_d(\mathbf{d}_{ij}, \mathbf{d}_n) \delta(s_{ij}, s_n) [1 - L(ij, n)]}$$

3.2.3 세그먼트 모델

세그먼트 필드는 인접하는 두 화소가 동일한 영역인가를 판단해 주는 모델로서, 객체추출의 결과이기도 하다. 세그먼트 필드도 시차필드와 같은 개념으로 에너지함수를 정의할 수 있다. 단, 동일한 세그먼트 영역이라 하더라도 시차필드는 각 화소마다 조금씩 다를 수 있기 때문에, 이를 보정해 주는 에너지항($V_d(\hat{\mathbf{d}}_{ij}, \mathbf{d}_s)$)이 필요하다.

$$-\ln \sum_{N \in S} p(s_N | \mathbf{L}) \prod_{n \in N} e^{-V_s(s_{ij}, s_n) [1 - L(ij, n)]} + \alpha V_d(\hat{\mathbf{d}}_{ij}, \mathbf{d}_s)$$

3.2.4 오클루전 모델

오클루전 필드는 좌우 스테레오 영상에서 부분적으로 가리워져서 정확한 대응점을 추정할 수 없는 영역을 나타낸다. 따라서, 오클루전에 대해서는 스테레오 영상간에 얼마나 정합되는지를 구하는 것은 의미가 없으므로, 일정한 정합에러로 구현된다. 그런데, 확률적으로 시차필드를 구하는 경우, 각 화소에서 최고의 확률로 선택된 시차의 확률은 오클루전 영역에서는 상대적으로 작거나 불안정한 값을 지닌다는 특징을 확인할 수 있다.

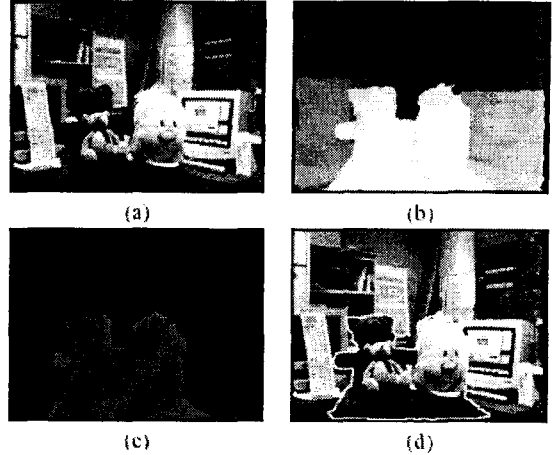


그림 1. 시차추정 및 객체추출 (a) 스테레오영상, (b) 시차, (c)와 (d) 추출된 객체 영역

이러한 확률적 불안정성을 활용하면 오클루전 검출에 유용하며, 각 화소에서 추정된 시차의 확률값에 반비례하도록 오클루전의 에너지를 정의한다.

$$U_0(\mathbf{I}_l | \mathbf{d}_{ij}, o_{ij}, \mathbf{I}_r)(2.0 - p(d_{ij})) + \sum_{n \in N} \frac{1}{p(d_{ij})} \delta(o_{ij}, o_n)$$

3.2.5 라인모델

라인필드는 시차 및 세그먼트 필드의 불연속성을 보장하기 위한 것으로서, 색상정보의 변화율을 계산하여, 에지가 존재할 만한 부분에 대해서는 두 필드의 불연속성을 강화시켜 준다.

$$\sum_{n \in N} L(ij, n) \left\{ \frac{C}{\nabla \mathbf{I}_{l+1}(ij, n)} + T_l \right\}$$

3.2.6 반복적인 확률적 확산

각 필드모델에 대한 에너지함수는 반복적으로 계산되어 다른 필드에 영향을 주며, 이러한 반복적인 과정을 통하여 에너지공간이 확률적으로 확산된다. 즉, 시차필드는 시차필드가 관여된 에너지 함수가 최소가 되도록 추정되며, 이 추정된 시차필드는 세그먼트 및 오클루전 필드의 에너지함수 계산에 이용된다. 최종적으로 에너지함수가 안정되면, 각 화소에서 추정된 필드는 에너지함수를 최적화하는 형태로 추정된 필드가 된다.

III. 모의실험결과

본 논문에서는 다양한 스테레오 영상을 이용하여 시차, 객체추출, 오클루전 검출 등의 다양한 실험을 수

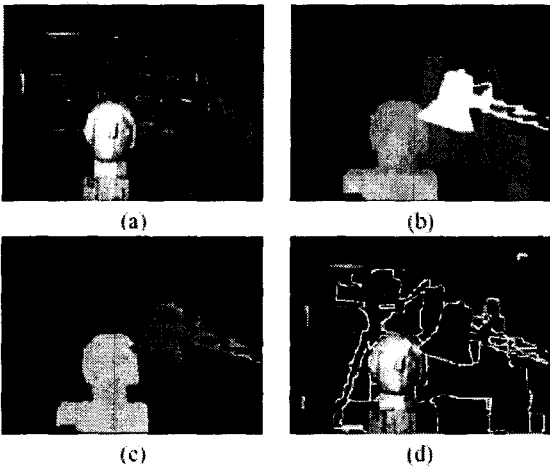


그림 2. 시차추정 및 다중객체 추출, (a) 스테레오영상, (b) 시차, (c)와 (d) 추출된 객체 영역

행하였다. 앞에서 설명한 확률적 모델과 확산의 개념을 구현하여 실험을 수행하였다. 그림 1 은 복잡한 배경으로부터 시차추정 및 전경에 있는 객체를 효과적으로 추출한 결과를 보여주고 있으며, 그림 2 는 다중객체를 추출한 결과이다. 그림 3 은 시차 및 오클루전 필드에 대한 결과를 제시하고 있다. 본 논문에서 제안한 모델에 따라 실험한 결과, 시차 추정 및 객체추출에서 우수한 성능을 보여 주었다. 오클루전 및 라인필드는 시차 및 세그먼트 필드의 불연속성을 보장하면서, 더욱 정확하게 객체의 경계를 추출해 내는데, 긍정적인 역할을 수행하였다. 그리고, 각 필드는 다른 필드의 추정에 긍정적으로 피드백 작용을 함으로써, 전체적인 필드의 추정 성능을 향상시키고 있다.

VI. 결론

본 논문은 확률적 확산 기법 및 확률모델을 이용하여 스테레오 영상간의 시차정보를 추정하고, 영상의 배경으로부터 객체를 추출해 내는 기법을 다루고 있다. 본 논문에서는 시차, 세그먼트, 라인, 및 오클루전 필드를 Markov random field (MRF) 모델로 정의하고, 각 필드간의 에너지함수를 제안하였으며, 에너지 최소화 기법으로서 확률적 확산 기법을 이용하여 최적의 시차 필드 및 객체추출을 수행하였다. 다양한 스테레오 영상을 이용하여 실험한 결과에 의하면, 본 논문에서 제안된 확률모델, 및 에너지 함수의 최적화 기법은 안정적

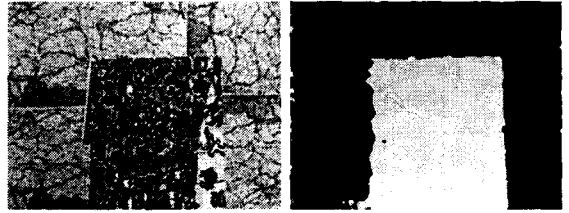


그림 3. 시차추정 및 오클루전 검출,

이고 우수한 시차추정과 객체추출 성능을 보여 주었다. 본 논문에서 추출된 객체는 3 차원 공간에서의 깊이(시차)정보를 포함하고 있으므로, 가상현실에서의 공간적 영상합성 및 모델링 등에서 매우 유용한 기법이다.

V. 감사의 글

이 논문은 2002 년도 한국학술진흥재단의 지원에 의하여 연구되었음 (KRF-2002-003-D00345).

참고문헌

- [1] J.Y.A. Wang and E.H. Adelson, "Representing Moving Images with Layers," *IEEE Trans. on IP*, Vol.3, No.5, pp.625-638, Sept. 1994.
- [2] A. Redert, E. Hendricks, and J. Biemond, "Correspondence estimation in image pairs," *IEEE Signal Processing Magazine*, Vol.16, No.3, pp. 29-46, 1999.
- [3] Y.Boykov, O.Veksler, and R.Zahih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. on PAMI*, Vol.23, No.11, pp.1222-1239, Nov. 2001.
- [4] K.P.Lim, A.Das, and M.N. Chong, "Estimation of occlusion and dense motion fields in a bidirectional Bayesian framework," *IEEE Trans. on PAMI*, Vol.24, No.5, pp.712-718, 2002.
- [5] Y.Tsaig and A. Averbuch, "A Region-Based MRF Model for Unsupervised Segmentation of Moving Objects in Image Sequences," *Proc. of CVPR*, 2001.
- [6] I.Patras, E.Hendriks, and R. Lagendijk, "Video segmentation by MAP Labeling of Watershed Segments," *IEEE Trans. on PAMI*, Vol.23, No.3, pp.326-332, March, 2001.
- [7] Sang Hwa Lee, Yasuaki Kanatsugu, and Jong-II Park, "Stochastic diffusion for stereo matching and line fields estimation," *IJCV*, vol. 47 (1/2/3) pp. 195-218, 2002