

논문 2010-47SP-1-14

# Belief Propagation 기반 스테레오 정합을 위한 정합 파라미터의 추정방식 제안

(Estimating the Regularizing Parameters for Belief Propagation Based Stereo Matching Algorithm)

오 광 희\*, 임 선 영\*, 한 희 일\*\*

(Kwang Hee Oh, Sun Young Lim, and Hee-il Hahn)

## 요 약

본 논문에서는 스테레오 이미지로부터 디스패리티 맵을 추출하기 위한 확률모델을 제시하고 이의 해를 구하는 과정은 에너지 기반 스테레오 정합과 일치함을 이론적으로 증명한다. 정합되는 화소 간의 차와 인근 화소에 해당되는 디스패리티의 차는 exponential 확률분포에 근사하다는 사실을 실험적으로 확인하고 이에 근거하여 이들의 정합 파라미터를 최적화하는 식을 유도하고 이를 실험적으로 구하는 방법을 제시한다. 에너지 기반 스테레오 정합 알고리즘의 성능은 기본적으로 정합 파라미터의 크기에 매우 민감하므로 이미지에 따라 적절한 값을 사전에 구하여 적용하여야 한다. 제안한 방식은 초기에 임의의 파라미터로 디스패리티 맵을 구한 후에 이의 통계적 특성을 이용하여 정합 파라미터를 추정하고 추정된 파라미터를 적용하여 디스패리티 맵을 재차 구하는 과정을 반복함으로써 최적의 파라미터에 적용적으로 수렴하도록 조정한다. 따라서, 이미지에 따라 사전에 정합 파라미터를 구하여야 하는 문제를 해결할 수 있다. Middlebury 웹사이트에서 제공한 다양한 스테레오 이미지를 이용하여 제안한 방식으로 구한 파라미터가 최적의 값으로 수렴하는지를 조사하고 이의 수렴 속도와 성능 개선 효과 등을 확인한다.

## Abstract

This paper defines the probability models for determining the disparity map given stereo images and derives the methods for solving the problem, which is proven to be equivalent to an energy-based stereo matching. Under the assumptions the difference between the pixel on the left image and the corresponding pixel on the right image and the difference between the disparities of the neighboring pixels are exponentially distributed, a recursive approach for estimating the MRF regularizing parameter is proposed. Usually energy-based stereo matching methods are so sensitive to the parameter that it should be carefully determined. The proposed method alternates between estimating the parameter with the intermediate disparity map and estimating the disparity map with the estimated parameter, after computing it with random initial parameter. It is shown that the parameter estimated by the proposed method converges to the optimum and its performance can be improved significantly by adjusting the parameter and modifying the energy term.

**Keywords :** Stereo matching, Belief propagation, Markov Random Fields, Parameter estimation

\* 학생회원, \*\* 평생회원, 한국의국어대학교  
정보통신공학과

(Department of Information and Communications  
Engineering, Hankuk University of Foreign  
Studies)

※ 이 논문은 2009년도 한국의국어대학교 교내 학술연  
구지원에 의하여 연구되었음

접수일자: 2009년10월13일, 수정완료일: 2009년12월30일

## I. 서 론

스테레오 정합은 컴퓨터 비전 분야에서 수십 년 동안  
연구가 진행되고 있는 고전 분야이지만 최근에 삼차원  
영상 합성, 이미지 기반 렌더링 등에 요소기술로 적용  
되면서 이에 대한 연구가 다시 활성화되고 있다. 스테

레오 정합은 기준 영상과 정합이 이루어지는 영상 사이에 실제 동일한 한 점을 나타내는 화소의 위치 차를 구하는 과정으로, 이 정보를 이용하여 이차원의 영상으로부터 삼차원 구조를 복원한다. 기본적으로 스테레오 정합은 공간 상에서 임의의 점은 좌, 우 이미지에서 동일한 화소 값을 갖는다는 가정에 기반하고 있다. 하지만 스테레오 정합은 다음과 같은 이유로 인해 정확한 추출이 매우 어려운 문제로 알려져 있다. 첫째, 두 카메라의 시점에 따라 반사되는 빛의 양이 변할 수 있고, 카메라의 센서 잡음 등으로 인하여 동일한 점이라도 좌우 카메라에 찍힌 해당 화소는 다른 값을 가질 수 있다. 둘째, 텍스처 없이 일정 영역이 유사한 화소 값을 갖게 되면, 그 영역 내의 화소에 정합되는 화소를 찾기가 매우 어려워 부정확하게 정보를 추출할 가능성이 높다. 이러한 문제를 해결하기 위해서 텍스처가 있는 영역에서 추출한 정보를 텍스처 없는 영역으로 확산하는 방법을 적용하기도 한다. 셋째, 기본적으로 디스패리티나 깊이가 정보는 영역 내에서는 거의 동일한 값을 갖지만 물체의 경계선 부분에서는 디스패리티가 변할 가능성이 매우 높으므로 이에 대한 별도의 처리가 요구되며, 한쪽 이미지에서 폐색 영역(occlusions)에 속하는 화소는 정합이 되지 않도록 주의해야 한다. 그런데 이들이 정확히 검출되지 않으면 그 영역에서의 정합은 오류로 나타나며 그 인근 영역에서의 정확한 정합을 어렵게 한다. 이와 같은 이유로 스테레오 정합은 해가 항상 존재하지는 않으므로 추론을 통해서 해를 찾는 수 밖에 없다.

이러한 문제를 해결하는 데에는 확률개념에 근거한 베이시안(Bayesian) 방식이 비교적 만족스러운 결과를 제시하는 것으로 알려져 있다. 상호상관관계를 이용하는 지역적 방식은 정합 시 소요되는 비용함수를 계산하고 이를 합산한 후에 디스패리티를 바로 구하는데 비하여 베이시안 방식 같은 전역적 방식은 최적화할 전역함수가 존재하여 디스패리티를 구한 후에 이를 최적화하는 과정이 포함된다. 필드 내의 각 화소는 그 인근 화소와의 사이에 불연속점이 없으면 유사한 값을 갖는 동시에 그 화소에 해당하는 디스패리티는 인근 화소의 디스패리티와 거의 동일하다. 이러한 특성을 이용하여 많은 연구자들이 다음과 같은 에너지 함수를 정의하고 이를 최소화하는 값을 찾음으로써 디스패리티를 구하고 있다.<sup>[6~7]</sup>

$$E(d_i) = \sum_{i \in I} U(d_i) + \lambda \sum_{(i,j) \in N} V(d_i, d_j) \quad (1)$$

여기서,  $I$ 는 전 이미지 화소를 나타내고,  $N$ 는  $i$  인근의 4개 화소를 나타낸다.  $U(d_i)$ 는 화소  $i$ 에 디스패리티를  $d_i$ 로 정할 경우의 비용을 나타내고  $V(d_i, d_j)$ 는 인근의 화소  $i$ 와 화소  $j$ 에 디스패리티를 각각  $d_i, d_j$ 로 할 때의 비용을 나타낸다. 식(1)의  $E(d_i)$ 를 최소화한다는 의미는 곧, 정합되는 화소는 좌우 서로 유사한 값을 가지면서 동시에 그 인근 화소의 디스패리티와 유사한 값을 가져야 한다는 가정을 담고 있다. 또한, 한 화소의 디스패리티를 구하는데 있어서 그 화소에 대응되는 화소와 그 인근의 화소만을 고려한다는 면에서 이미지를 MRF(Markov Random Fields)로 가정하고 있다는 사실을 알 수 있다. 이와 같은 에너지 기반 정합 방식은 잡음에 취약한 화소기반 방식과 성감에 의한 부정확한 변위 값을 갖는 특징 기반 방식의 단점을 보완하는 특성을 지니고 있다. 식 (1)의 해를 구하기 위해 dynamic programming 기법을 적용할 수도 있으나 보다 더 정확한 디스패리티를 얻기 위해서는 MRF 알고리즘이 효과적이고 효율적이다.

지난 몇 년 동안 MRF 모델을 이용하여 스테레오, 옵티칼 플로우, 이미지 복원 등의 비전 문제를 해결하려는 연구 결과가 많이 발표되었다. 이 중에서도 특히, BP(belief propagation)는 글로벌 솔루션에 근사한 해를 제공하여 정확도 높은 결과를 제시해 준다는 점에서 주목을 받고 있다.<sup>[1]</sup> BP는 이미지가 MRF라고 가정하여 식(1)에 정의된 바와 같이 정합되는 화소 간의 차  $U(d_i)$ 와 인근 화소의 디스패리티 차  $V(d_i, d_j)$ 의 합으로 정의되는 에너지를 최소화시키는 디스패리티를 매우 정확하게 찾는 알고리즘으로 알려져 있다. Sun et al.<sup>[4]</sup>는 스테레오 정합을 세 개의 MRF 즉, 기준 이미지의 좌표에 해당하는 디스패리티 맵, 디스패리티의 불연속점, 폐색 영역으로 모델링하였다. 모델링 후에 BP에 적용하는 과정에서 불연속점과 폐색 영역에 대한 고려를 생략하는 문제를 노출하였으나 BP의 가능성을 충분히 보여주고 있다. Felzenszwalb와 Huttenlocher<sup>[6]</sup>는 멀티스케일로 이미지를 분할한 다음에 BP를 적용시킴으로써 메시지 전파속도를 개선하는 동시에 계산량을 크게 감소시키는 알고리즘을 제안하였다. 하지만, 아직까지 디스패리티의 불연속점이나 폐색 영역에 대한 차별화된 처리 없이 BP가 내부적으로 처리하는 능력에 의존하는 것이 대부분이다.

따라서, 본 논문에서는 불연속점과 폐색 영역은 경계

영역에서 주로 발생함에 착안하여 이들을 한 그룹으로 묶어 전파되는 메시지 량을 조절하는 기능을 추가함으로써 BP의 성능을 향상시키는 기술을 제안하였다. 또한, 식(1)의 에너지 함수를 이용하는 알고리즘들은 정합 파라미터  $\lambda$ 의 크기에 따라 성능이 매우 민감하게 변하므로 이미지가 주어지면 그 이미지에 적합한  $\lambda$ 를 별도로 미리 구해 놓아야 하는 문제가 있다. 이러한 문제로 인해 실제의 응용환경에서는 활용하기 어려운 한계를 지니고 있다. 이를 해결하기 위해 본 논문에서는 BP의 적용과정에서 확보한 중간 단계의 디스패리티 맵을 이용하여  $\lambda$ 를 확률모델에 근거해서 구하고 이렇게 구한  $\lambda$ 를 활용하여 재차 디스패리티 맵을 구하는 과정을 반복하는 방식을 제안하고 실험을 통하여 제안 방식의 적정성을 확인하였다. 즉, 본 논문의 목적은 스테레오 이미지에서 디스패리티 검출의 정확도 향상 보다는 입력되는 스테레오 이미지의 통계적 특성에 따라 정합 파라미터를 사전에 미리 구해야 할 필요 없이 적응적으로 파라미터를 조정하므로써 최적의 성능에 수렴하도록 하는 것이다.

본 논문의 구성은 다음과 같다. II장에서는 BP가 어떻게 동작하는지를 알아보고 III장에서는 디스패리티 맵을 구하기 위한 확률 모델을 설명하고 MLE (Maximum likelihood estimation) 기준에서 최적의 해는 식(1)과 동일하다는 사실을 보여준다. 또한, 이를 이용하여  $\lambda$ 를 구하는 방식을 제안한다. 스테레오 정합에 BP를 적용했을 때의 결과와 함께 제안한 방식이 성능개선에 미치는 영향과 수렴속도 등을 실험에 근거하여 분석한 결과를 IV장에 제시한다. 마지막으로 V장에서는 결론을 맺고 향후 연구 진행방향에 대하여 논의한다.

## II. Belief Propagation

베이지안 네트워크는 여러 개의 노드와 이들의 상호 연결로 구성되는데 각 연결마다 조건부 확률이 주어진다. 이 네트워크는 최종적으로 모든 노드들의 joint probability를 구한 다음, 관심 있는 노드의 marginal probability를 최대화하는 점을 구함으로써 추론한다. 그런데, 베이지안 네트워크는 각 노드들 간에 방향성이 있고 폐루프를 형성하지 않으면서 각 노드에 대해서 통계적 연관성이 있는 노드의 수가 많지 않아야 적용 가능하다. 이러한 문제를 BP는 효율적이고 효과적으로 해결하고 있다. 예를 들어, 이미지로부터 주어진 정보  $y_i$

를 이용하여 원하는 정보  $x_i$ 를 추론한다고 가정한다. 여기서 인덱스  $i$ 는 화소나 영역의 위치 정보를 나타낸다. 또한, 각 위치  $i$ 에서  $x_i$ 와  $y_i$  간에는 통계적 연관성이 있다고 가정한다. 이 연관성을  $\Phi_i(x_i, y_i)$ 라고 표현하고, 원하는 정보  $x_i$ 와 그 인근의 정보  $x_j$  간의 연관 함수를  $\Psi_{ij}(x_i, x_j)$ 라고 정의하면, 이미지가 MRF라는 가정 하에 이미지 정보와 원하는 정보들의 joint probability는 다음과 같이 표현할 수 있다.

$$P(\{x\}, \{y\}) = \frac{1}{Z} \prod_{(i,j)} \Psi_{ij}(x_i, x_j) \prod_i \Phi_i(x_i, y_i) \quad (2)$$

위 식에서  $Z$ 는 정규화 상수이고,  $(i, j)$ 는  $i$ 를 중심으로 한 4개의 인근 화소를 나타낸다. 일반적으로  $\Phi_i(x_i, y_i)$ 와  $\Psi_{ij}(x_i, x_j)$ 는 각각 조건부 확률인  $P(x_i/y_i)$ ,  $P(x_i/x_j)$ 에 비례하는 값을 갖는다. 그런데, 이미지는 일반적으로 화소나 영역 간의 방향성이 없고 폐루프가 형성되므로 베이지안 네트워크 보다는 BP가 더 적절하다. 일반적으로  $y_i$ 는 화소 값과 같이 고정된 값을 가지므로  $\Phi_i(x_i, y_i)$ 를  $\Phi_i(x_i)$ 로 나타내면 위 식은 다음과 같이 표현할 수 있다.

$$P(\{x\}) = \frac{1}{Z} \prod_{(i,j)} \Psi_{ij}(x_i, x_j) \prod_i \Phi_i(x_i) \quad (3)$$

이와 같이,  $x_i$ 들의 joint probability가 구해지면 각  $P(\{x\})$ 에 대한 marginal probability는 다음과 같이  $x_i$ 를 제외한 나머지 변수들에 대하여 합을 계산함으로써 구할 수 있다.

$$P(x_i) = \sum_{x_1} \sum_{x_2} \dots \sum_{x_{i-1}} \sum_{x_{i+1}} \dots \sum_{x_N} P(x_1, x_2, x_3, \dots, x_N)$$

BP에서는 marginal probability  $P(x_i)$ 를 노드  $i$ 에서의 belief라고 부른다. 노드  $i$ 에서의 belief  $B_i(x_i)$ 는 다음과 같이 구할 수 있다.<sup>[1~2]</sup>

$$B_i(x_i) = k \Phi_i(x_i) \prod_{j \in N(i)} M_{ji}(x_i) \quad (4)$$

여기서  $k$ 는  $\sum_i B_i(x_i) = 1$ 이 되도록 정규화하는 상수이고  $j$ 는  $i$  인근의 4개 노드를 나타낸다. 그리고,  $M_{ji}(x_i)$ 는 노드  $j$ 에서 노드  $i$ 로 보내는 메시지인데,  $x_i$ 와 동일한 크기의 벡터이며 노드  $j$ 가 바라본 노드  $i$ 의

추정 값을 의미한다. 메시지  $M_{ji}(x_i)$ 는 다음과 같은 방법으로 업데이트된다.

$$M_{ji}(x_i) = \max_{x_j} \Phi_j(x_j) \Psi_{ji}(x_j, x_i) \prod_{k \in N(j) \setminus i} M_{kj}(x_j) \quad (5)$$

여기서  $N(j) \setminus i$ 는  $i$ 를 제외한  $j$  인근 세개의 노드를 나타낸다. 식(4)와 식(5)의 구현을 용이하게 하기 위해 양변에  $-\log$ 를 취하여 정리하면 다음과 같이 표현할 수 있다.

$$b_i(x_i) = k\phi_i(x_i) + \sum_{j \in N(i)} m_{ji}(x_i) \quad (6)$$

$$m_{ji}(x_i) = \min_{x_j} \left\{ \phi_j(x_j) + \psi_{ji}(x_j, x_i) + \sum_{k \in N(j) \setminus i} m_{kj}(x_j) \right\} \quad (7)$$

BP를 구동시키기 위해서는 매 단계마다 메시지를 업데이트시킨 다음 전 노드에 대해서 belief를 구하는데, 수렴할 때까지 이 과정을 충분히 반복 적용한다.

### III. 제안 알고리즘

#### 1. 디스패리티 맵을 구하기 위한 확률 모델

스테레오 이미지  $I$ 와  $J$ 가 주어졌을 때 스테레오 정합을 두개의 MRF인 디스패리티 맵  $D$ 와 디스패리티의 불연속점  $S$ 로 모델링한다. 베이시안 범칙에 따라 주어진 이미지  $I$ 와  $J$ 에 대하여  $D$ 와  $S$ 에 대한 조건부 확률은 다음과 같이 표현할 수 있다.

$$P(D, S/I, J) = \frac{P(I, J, D/S)P(S)}{P(I, J)} \quad (8)$$

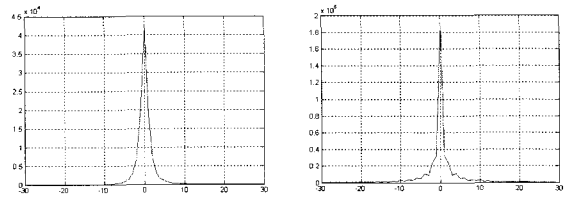
이 식에서,  $P(I, J)$ 와  $P(S)$ 는 기본적으로 장면에 따라 일정하여  $P(D, S/I, J) \propto P(I, J, D/S)$ 로 볼 수 있으므로  $P(I, J, D/S)$ 를 최대화하는  $D$ 를 구함으로써 원하는 디스패리티 맵을 구할 수 있다.  $D$ 와  $S$ 가 주어졌을 때 이미지  $I$ 와  $J$ 의 생성 확률모델  $P(I, J/D, S)$ 는 각 화소가 서로 독립이라는 가정 하에 식(9)와 같이 표현할 수 있고,  $S$ 가 주어졌을 때  $D$ 의 확률  $P(D/S)$ 는 MRF라고 가정하면 식(10)과 같이 모델링할 수 있다.

$$P(I, J/D, S) \propto \prod_i P(e(d_i)/d_i, s) \quad (9)$$

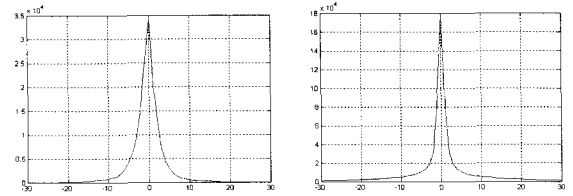
$$P(D/S) \propto \prod_{(i,j) \in N} P(\nabla d_{ij}/s) \quad (10)$$

여기서,  $e(d_i)$ 는 정합되는 두 화소 간의 차이를 나타내고,  $\nabla d_{ij}$ 는 인근 두 화소의 디스패리티 차이를 나타낸다. 그런데,  $P(I, J, D/S) = P(I, J/D, S)P(D/S)$ 이므로  $P(D, S/I, J)$ 는 다음과 같이 정리할 수 있다.

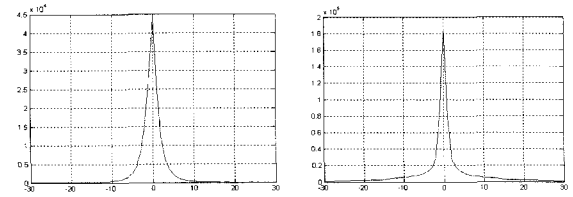
$$P(D, S/I, J) \propto \prod_i P(e(d_i)/d_i, s) \prod_{(i,j) \in N} P(\nabla d_{ij}/s) \quad (11)$$



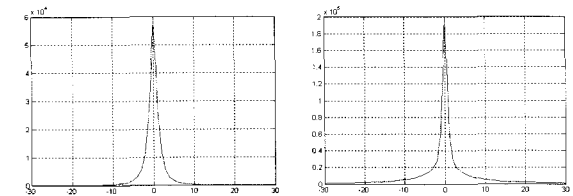
(a) Tsukuba



(b) Cones



(c) Teddy



(d) Venus

그림 1. Middlebury 이미지에서 구한 대응 화소 간의 오차  $P(e(d_i)/d_i)$ 와 인근 디스패리티의 차  $P(\nabla d_{ij}/s)$ 에 대한 히스토그램

Fig. 1. Histograms of errors between the corresponding pixels in two images  $P(e(d_i)/d_i)$  and of the neighboring disparity difference  $P(\nabla d_{ij}/s)$ .

그런데  $e(d_i)$ 는 세그멘테이션 정보  $s$ 에 무관하므로  $P(e(d_i)/d_i, s) = P(e(d_i)/d_i)$ 로 표현할 수 있다.

ground\_truth 디스패리티 맵이 주어진 4쌍의 Middlebury 스테레오 이미지를 이용하여  $P(e(d_i)/d_i)$ 와  $P(\nabla d_{ij}/s)$ 를 실험적으로 구한 결과는 그림 1에 제시하였는데, 이들 모두 exponential 분포에 근사함을 알 수 있다. 따라서, 본 논문에서는  $P(e(d_i)/d_i)$ 와  $P(\nabla d_{ij}/s)$ 를 다음과 같은 확률분포로 모델링하였다.

$$\begin{aligned} P(e(d_i)/d_i) &= \rho e^{-\mu|e(d_i)|} \\ P(\nabla d_{ij}/s) &= \gamma e^{-\delta|\nabla d_{ij}|} \end{aligned} \quad (12)$$

결국, 디스패리티를 구하는 문제는 다음 식의 해를 구하는 문제로 귀결된다.

$$\max P(D, S/I, J) \propto \max_i \prod_i \rho e^{-\mu|e(d_i)|} \max_{(i,j) \in N} \prod_{(i,j) \in N} \gamma e^{-\delta|\nabla d_{ij}|} \quad (13)$$

이 식의 근을 구하는데 있어서 계산량을 줄이기 위해 다음과 같이 양 변에  $-\log$ 를 취하면 다음과 같이 나타낼 수 있다.

$$\begin{aligned} \min(-\log P(D, S/I, J)) &\propto \\ \min \sum_i (-\log(\rho e^{-\mu|e(d_i)|})) &+ \min \sum_{(i,j) \in N} (-\log(\gamma e^{-\delta|\nabla d_{ij}|})) \end{aligned} \quad (14)$$

위 식을 정리하면 다음과 같이 나타낼 수 있다.

$$\begin{aligned} \min(-\log P(D, S/I, J)) &\propto \\ \min \sum_i (\mu|e(d_i)| - \log \rho) &+ \min \sum_{(i,j) \in N} (\delta|\nabla d_{ij}| - \log \gamma) \end{aligned} \quad (15)$$

여기서,  $\log \rho$ 와  $\log \gamma$ 는 상수이어서 이를 제거해도 결과에는 영향을 주지 않는다. 또한, 양변을  $\mu$ 로 나누면 식(15)는 다음과 같이 간단히 정리할 수 있다.

$$\begin{aligned} \min(-\log P(D, S/I, J)) &\propto \\ \min(\sum_i |e(d_i)| + \lambda \sum_{(i,j) \in N} |\nabla d_{ij}|) &\quad (16) \end{aligned}$$

여기서,  $\lambda = \frac{\delta}{\mu}$ 이다. 위 식을 최소화하는  $d_i$ 를 찾아냄으로써 디스패리티를 구할 수 있다.  $U(d_i)$ 와  $V(d_i, d_j)$ 를 각각  $|e(d_i)|$ ,  $|\nabla d_{ij}|$ 로 정의하면 위 식은 식(1)과 동일함을 알 수 있다.

## 2. 파라미터 $\lambda$ 의 유도

파라미터  $\lambda$ 를 계산하기 위해서는 확률분포 파라미터인  $\mu$ 와  $\delta$ 를 먼저 구하여야 한다. 그런데, BP를 구동하는 과정에서 매 단계마다 메시지 벡터를 업데이트하고 belief를 구하게 되는데, 이를 최대화하는 디스패리티를 전 화소에서 구할 수 있다. 이와 같이 디스패리티 맵이 주어지면  $P(e(d_i)/d_i)$ 와  $P(\nabla d_{ij}/s)$ 를 실험적으로 구할 수 있는데, 이를 통하여 식(17)과 같이  $\mu$ ,  $\delta$ 를 구한 다음  $\lambda = \frac{\delta}{\mu}$ 를 계산할 수 있다. 기본적으로 BP를 적용하여 디스패리티를 구하려면 주어진 이미지에 따라 파라미터  $\lambda$ 를 별도로 구해서 적용하여야 하는 어려움이 있다. 하지만, 본 논문에서 제안한 방법으로 파라미터  $\lambda$ 를 구하여 적용하면 이미지의 특성에 구애 받지 않으면서도 보다 정확한 디스패리티를 구할 수 있는 장점이 있다.

$$\begin{aligned} \mu &= -\log \left\{ \frac{P(e(d_i) = n+1/d_i)}{P(e(d_i) = n/d_i)} \right\} \\ \delta &= -\log \left\{ \frac{P(\nabla(d_{ij}) = n+1/s)}{P(\nabla(d_{ij}) = n/s)} \right\} \end{aligned} \quad (17)$$

## 3. 에너지 함수의 개선

식 (1)의 에너지 함수를 재해석해 보면  $U(d_i)$ 는 화소  $i$ 에 디스패리티를  $d_i$ 로 정할 경우의 비용이고  $V(d_i, d_j)$ 는 인근의 두 화소  $i$ 와  $j$ 에 디스패리티를 각각  $d_i$ ,  $d_j$ 로 할당할 때의 비용을 나타낸다. 에너지 함수가 최소 값을 가지려면 정합되는 두 화소가 동일한 값을 가지면서 동시에 인근 화소는 디스패리티가 동일하여야 한다. 에너지 함수를 최소화시키는 디스패리티 맵을 구하여 살펴보면, 에너지 함수의 특성으로 인하여 영역 간의 경계 부분에서 블러링 현상이 발생하는 것을 피할 수 없다. 즉, 화소  $i$ 와 화소  $j$ 가 서로 다른 영역에 속해 있으면  $d_i$ 와  $d_j$ 는 서로 다른 값을 가질 확률이 높는데 이들이 동일한 값을 갖도록 한다면 오히려 오류를 증폭시킬 수 있어 경계 면에서 블러링이 발생하는 요인이 된다. 이러한 문제를 해결하기 위하여, 본 논문에서는  $V(d_i, d_j)$ 를 계산하는데 있어서 line process의 개념<sup>[3]</sup>을 도입하여 화소  $i$ 와 화소  $j$  간에 값의 차이가 커서 이들이 다른 영역에 속해 있을 가능성이 있으면  $V(d_i, d_j) = 0$ 으로 정함으로써 이를 배제하였다. 즉, 위의 식(1)을 다

음과 같이 수정하여 적용하였다.

$$E(d_i) = \sum_{i \in I} U(d_i) + \lambda \sum_{(i,j) \in N} V(d_i, d_j) l_{ij} \quad (18)$$

여기서,  $l_{ij}$ 는 다음과 같이 결정된다.

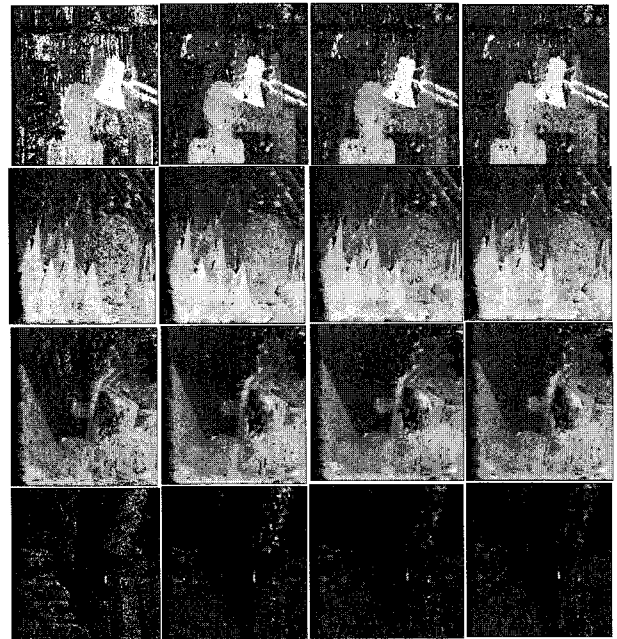
$$l_{ij} = \begin{cases} 1 & |I(i) - I(j)| \leq \theta \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

참조 이미지  $I$ 의 인근 화소  $I(i)$ 와  $I(j)$  간의 차이가 작아 이들이 동일 영역에 속해 있을 가능성이 높을 때에는  $V(d_i, d_j)$ 를 활성화시키지만, 이들의 차이가 크면 두 화소 사이에 에지가 있어서 서로 다른 영역에 속해 있을 확률이 높다. 이러한 경우에는 두 화소가 서로 다른 디스패리티를 가질 확률이 비교적 높으므로 인근 두 화소는 유사한 디스패리티를 가져야 한다는 제약조건을 약화시킴으로써 디스패리티 맵에서 영역 간의 블러링을 억제하는 효과를 얻을 수 있다. 본 논문에서는  $\theta$ 를 50으로 실험적으로 정하였는데, 이에 대한 별도의 지속적인 연구가 필요하다.

#### IV. 실험 결과

본 논문에서는 Middlebury 웹사이트<sup>[8]</sup>가 제공한 것으로 ground\_truth가 주어진 여러 스테레오 이미지를 이용하여 제안한 알고리즘의 동작을 확인하고 성능을 분석하였다. 성능분석은 Scharstein<sup>[5]</sup>이 제안한 방식으로 다음과 같이 nonocclusion 영역에서의 정합오율( $B_0$ )을 계산하고 또한 Middlebury 웹사이트에서 제공하는 평가방식을 이용하여 수행하였다. 이미지에 잡음이 존재하면 전처리를 통하여 잡음을 제거하거나 또는  $e(d_i)$ 와  $\nabla d_{ij}$ 를 각각  $\min(|\mathcal{I}(x, y) - \mathcal{I}(x, y - d_i)|, \tau)$ ,  $\min(|d_i - d_j|, \theta)$  등으로 모델링하여 임펄스 잡음에 대처해야 할 것이다. 이런 경우에는 조정해야 할 파라미터의 수가 증가하게 되어 알고리즘이 매우 복잡해지는 문제가 있어서 본 논문에서는 구현을 용이하게 하고 복잡도를 줄이기 위하여  $e(d_i)$ 와  $\nabla d_{ij}$ 를 각각  $|\mathcal{I}(x, y) - \mathcal{I}(x, y - d_i)|$ ,  $|d_i - d_j|$ 으로 정의하였다. 그림 2는 Middlebury 이미지에 대하여 식(1)에서 파라미터  $\lambda$ 를 1로 고정하여 BP를 반복 수행하였을 때의 디스패리티 맵을 보여주고 있다.

식(17)의  $\mu$ 와  $\delta$ 를 이용하여  $\lambda$ 를 변경시키면서 구한 디스패리티 맵은 그림 3에 제시하였다. 여기서,  $\lambda$ 를 구



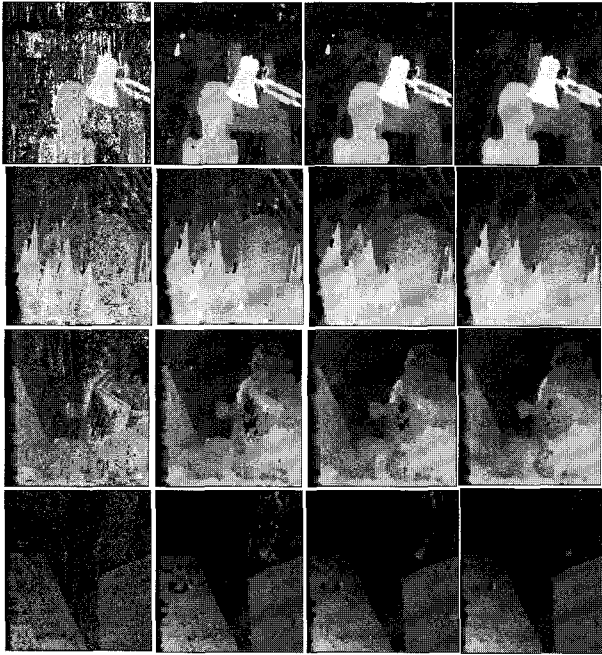
(a) 1회 (b) 10회 (c) 30회 (d) 60회

그림 2. Middlebury 이미지에 대하여  $\lambda = 1$ 로 고정된 상태에서 BP를 반복 적용하여 구한 디스패리티 맵

Fig. 2. The results of the BP-based stereo matching algorithm on the Middlebury data set when  $\lambda$  is fixed with 1. (a) Iteration (1). (b) Iteration (10). (c) Iteration (30). (d) Iteration (60).

하려면  $\mu$ 와  $\delta$ 를 동시에 구하여야 하는데, 본 논문에서는 BP를 적용하여 구한 디스패리티 맵으로부터  $P(e(d_i)/d_i)$ 와  $P(\nabla d_{ij}/s)$ 를 실험적으로 구한 다음, 이들이 식(12)에 나타난 바와 같이 exponential 확률분포를 갖는다는 가정을 이용하여  $\mu$ ,  $\delta$ 에 대한 측정치를 식(17)과 같이 구하고 이를 이용하여  $\lambda = \frac{\delta}{\mu}$ 를 구하였다.

그림 4는 Tsukuba 이미지에 대하여  $\lambda$ 를 변경하면서 구한, 반복 수행 회수에 따른 nonocclusion 영역에서의 정합오율을 그래프로 나타낸 것이다. Tsukuba 이미지에서는  $\lambda = 70$ 일 때 최적의 성능을 보여주고 있는데 제안 알고리즘은  $\lambda = 66$ 으로 수렴하고 있으며, 이때의 정합오율은 3.4%를 나타내고 있다. 또한 Venus, Cones, Teddy 등의 다른 이미지에 대해서도 각각 최적의 성능에 근접하여 수렴함을 확인하였다. 또한, 제안 알고리즘은  $\lambda$ 를 임의의 값으로 초기화하였을 때에도 회수가 반복됨에 따라 최적에 근사한 값으로 수렴함을 확인하였다. 그림 5는 Tsukuba 이미지에 대하여  $\lambda$ 의 초기값을 1, 50, 70, 100 등으로 변화를 주었을 때 반복회수에 따



(a) 1회 (b) 10회 (c) 30회 (d) 60회

그림 3. Middlebury 이미지에 대하여 식(17)에 제시한 방법으로  $\lambda$ 를 조정하여 구한 디스패리티 맵

Fig. 3. The results of the proposed method when the regularizing parameter  $\lambda$  is adjusted.

(a) Iteration (1). (b) Iteration (10). (c) Iteration (30). (d) Iteration (60).

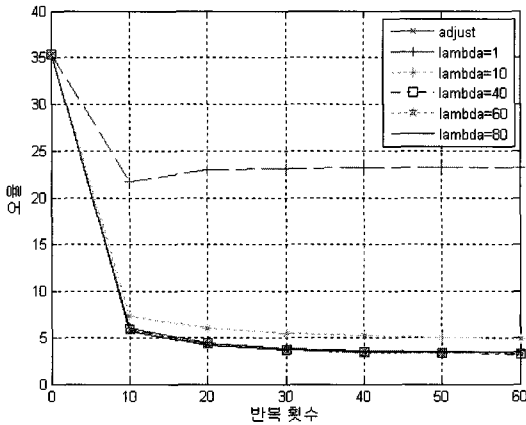


그림 4. 정합 파라미터  $\lambda$ 에 따른 반복 회수 별 nonocclusion 영역에서의 정합오율

Fig. 4. The error rates with respect to the ground truth depend on the regularizing weight  $\lambda$ . This graph shows the estimation of  $\lambda$  converges to the optimum.

큰  $\lambda$ 의 수렴과정을 보여주고 있다.

디스패리티 맵의 정확한 추정을 위해서는 이에 대한 적절한 확률모델과 함께 식(12)에 나타낸 바와 같은 좌우 이미지의 해당 화소 간의 차인  $e(d_i)$ 와  $\nabla d_{ij}$ 의 정확

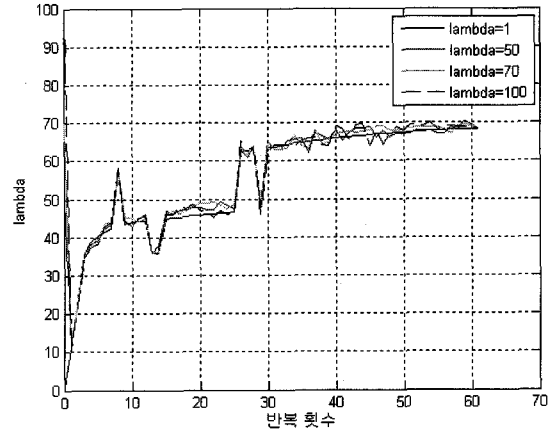


그림 5. 정합 파라미터  $\lambda$ 의 초기값의 변화에 대한 반복 회수에 따른  $\lambda$ 의 수렴과정 (Tsukuba 이미지)

Fig. 5. The proposed method converges the regularizing parameter  $\lambda$  to the specific value depending on the statistics of the input images irrespective of the initial value of  $\lambda$ .

도에 따라 크게 좌우된다.

이미지의 통계적 특성에 따른 적절한 모델을 유도하고 앞으로 최적의 성능을 얻기 위한  $P(e(d_i)/d_i)$ ,  $P(\nabla d_{ij}/s)$ 의 결정에 관한 이론적인 연구가 절실히 요구된다.

### V. 결 론

본 논문에서는 스테레오 이미지에서 디스패리티를 구하기 위한 확률모델을 제시하고 이에 의거하여 디스패리티를 구하기 위한 수식은 식(1)에 주어진 바와 같은 에너지 함수를 최소화하는 수식과 동일함을 유도하였다. 또한, 정합되는 화소 간의 차와 인근 화소의 디스패리티 차이가 exponential에 근사한 확률분포를 갖는다는 사실을 실험적으로 확인하였으며 이에 근거하여 식(1)의 파라미터  $\lambda$ 를 유도하고 이를 실험적으로 구하는 방법을 제시하였다. 실험을 통하여 제안 알고리즘은 반복적으로 수행됨에 따라 최적의  $\lambda$ 에 수렴함을 보여 주었다. 본 논문에서는 구현의 편의를 위하여  $e(d_i)$ 와  $\nabla d_{ij}$ 를 각각  $|\mathcal{I}(x, y) - \mathcal{I}(x, y - d_i)|$ ,  $|d_i - d_j|$ 으로 정의하였으나 이는 스테레오 이미지에 잡음이 거의 존재하지 않음을 가정한 것이다. 이미지에 잡음이 존재하면 전처리를 통하여 잡음을 제거하거나 또는  $e(d_i)$ 와  $\nabla d_{ij}$ 를 각각  $\min(|\mathcal{I}(x, y) - \mathcal{I}(x, y - d_i)|, \tau)$ ,  $\min(|d_i - d_j|, \theta)$  등으로 모델링하여 임펄스 잡음에 대처해야 할 것이다. 향후에,

제안 알고리즘의 성능을 보다 더 개선하기 위해서는  $e(d_i)$ 와  $\nabla d_{ij}$ 에 대한 보다 적절한 모델과 함께 이들의 확률분포를 보다 정확히 정의할 수 있어야 하며, 이를 기반으로 관련 파라미터를 최적화하는 알고리즘을 유도해야 할 것이다. 또한, 폐색영역에 속하는 화소에 대해서는 정합이 이루어지지 않도록 함과 동시에 디스패리티가 서로 다른 불연속 지역에서는 메시지 전파를 억제하는 방안을 보완함으로써 BP의 내재적인 문제점인 경계영역에서의 블러링 현상을 극복해야 한다. 이와 함께 belief propagation을 구현하는데 필요한 엄청난 계산량을 줄여 효율적으로 처리하는 방식에 대한 연구가 진행되어야 할 것으로 기대된다.

참 고 문 헌

[1] J. S. Yedidia, W. T. Freeman and Y. Weiss, "Understanding Belief Propagation and Its Generalizations," TR-2001-22, Jan. 2002.

[2] F. R. Kschischang, B. J. Frey, H. Loeliger, "Factor Graphs and the Sum-Product Algorithm," IEEE Transactions on Information Theory, vol.47, no.2, pp. 498-519, Feb. 2001.

[3] M. J. Black and A. Rangardjan, "On the Unification of Line Processes, Outlier Rejection, and Robust Statistics with Applications in Early Vision," International Journal of Computer Vision, 1995.

[4] J. Sun, N. N. Aheng, and H. Y. Shum, "Stereo Matching Using Belief Propagation," IEEE Transactions on Pattern Analysis and Machine Intelligence ,vol.25, no.7, pp.787-800, July 2003.

[5] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," International Journal of Computer Vision, pp.7-42, 2002.

[6] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient Belief Propagation for Early Vision," Proc. Conf. Computer Vision and Pattern Recognition, pp.261-268, 2004.

[7] L. Zhang and S. M. Seitz, "Estimating Optimal Parameters for MRF Stereo from a Single Image Pair," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.29, no.2, pp.331-342, Feb. 2007.

[8] "Middlebury Stereo Vision Page," <http://vision.middlebury.edu/stereo>.

저 자 소 개



오 광 희(학생회원)  
 2007년 한국외국어대학교 정보통신공학과 학사 졸업.  
 2010년 한국외국어대학교 정보통신공학과 석사 졸업예정.  
 <주관심분야 : 신호처리, 영상처리>



임 선 영(학생회원)  
 2007년 한국외국어대학교 정보통신공학과 학사 졸업.  
 2010년 한국외국어대학교 정보통신공학과 석사 졸업예정.  
 <주관심분야 : 신호처리, 영상처리>



한 희 일(평생회원)  
 1984년 서울대학교 제어계측공학과 학사 졸업.  
 1988년 서울대학교 제어계측공학과 석사 졸업.  
 1995년 University of Arizona 전기및컴퓨터공학과 박사 졸업.

2010년 현재 한국외국어대학교 정보통신공학과 교수.  
 <주관심분야 : 영상처리, 컴퓨터비전, 패턴인식>