

論文98-35S-9-12

거리영상과 밝기영상의 fusion을 이용한 영상분할

(Image Segmentation Based on Fusion of Range and Intensity Images)

張仁壽*, 朴來弘*

(In Su Chang and Rae-Hong Park)

요 약

본 논문에서는 거리영상과 밝기영상의 fusion을 이용한 영상분할을 제안한다. Bayes 이론을 기반으로 하여 Markov random field (MRF)로 선형적인 정보를 모델링한다. 거리영상과 밝기영상에서 추출한 특징을 이용하여 maximum a posteriori (MAP) 추정기를 구성한다. 거리영상에서 물체는 국부적인 평면으로 근사되어 각 점마다 평면 계수를 추정해 계수 공간을 구성한다. 밝기영상에서는 α 트림드 (α -trimmed) 분산이 밝기 특성을 구성한다. 각 공간상의 특징을 에지에 대한 likelihood를 설정하여 구성된 MAP 추정기를 최적화함으로써 영상을 분할한다. 모의실험을 통해 제안된 구조가 그림자, 잡음 그리고 광원의 blurring에 관계없이 영상을 잘 분할한 것을 보였다.

Abstract

This paper proposes an image segmentation algorithm based on fusion of range and intensity images. Based on the Bayesian theory, *a priori* knowledge is encoded by the Markov random field (MRF). A maximum a posteriori (MAP) estimator is constructed using the features extracted from range and intensity images. Objects are approximated by local planar surfaces in range images, and the parametric space is constructed with the surface parameters estimated pixelwise. In intensity images the α -trimmed variance constructs the intensity feature. An image is segmented by optimizing the MAP estimator that is constructed using a likelihood function based on edge information. Computer simulation results show that the proposed fusion algorithm effectively segments the images independentl of shadow, noise, and light-blurring.

I. 서 론

컴퓨터 시각은 자동화 시스템에서 물체의 구분 또는 인식 등을 위해 인간의 시각적 기능을 컴퓨터에게 부여한다. 기존의 자동화 시스템은 카메라로 필요한 목적에 맞는 영상을 취득하여 영상 처리를 통해 시각적

기능을 얻었다. 그러나 일반 CCD 카메라로 얻는 밝기영상은 취득 과정에서 생기는 왜곡이나 잡음의 영향, 그림자 등에 크게 영향을 받는다. 대기중에서 발생하는 빛의 굴절과 반사, 카메라 렌즈에 낀 먼지 등 여러 가지 잡음의 요소가 존재하고 물체의 밝기에 갑작스런 변화를 유발시키는 그림자는 물체 인식시 인식을 저하시킨다. 또한, 물체의 표면마다 각기 반사특성이 다르기 때문에 곡면 등에서 하이라이트나 확산현상이 일어난다. 그러므로 고정밀을 요구하는 자동화 시스템의 응용분야에 밝기영상을 사용하는 경우 원하는 결과를 얻기 어렵다. 이런 이유로 빛의 밝기 값이 아닌 3차원

* 正會員, 西江大學校 電子工學科

(Dept. of Electronic Eng., Sogang Univ.)

※ 본 연구는 제어계측 신기술연구센터 (ERC-ACI)의 연구비지원에 의해 이루어졌음

接受日字: 1997年1月27日, 수정완료일: 1998年7月4日

거리값을 나타내는 거리영상을 많이 이용한다. 거리영상은 능동적 접근 방법으로 얻는데, 3차원 물체의 표면에서 미리 정해진 기준면까지의 수직 거리로 화소값을 구성한다. 거리영상은 밝기영상과는 달리 그림자나 빛의 반사, 조명 환경 등에 영향을 받지 않아 밝기영상에서 발생하는 문제점들이 상당히 감소되고 3차원 물체의 기하학적 특징을 직접적으로 제공하는 이점이 있다. 따라서 밝기영상에서 얻을 수 없는 거리정보를 이용해서 고수준 (high-level)의 처리과정 없이 3차원 기하학적 특징을 이용한 3차원 모델링이나 3차원 물체 인식 등이 가능하다.^[1]

3차원 모델링을 위해 거리 영상을 얻을 경우 밝기영상에서 일반적으로 가정하는 가우시안 잡음뿐 아니라 registration 과정의 변환과정과 물체간의 겹쳐짐으로 인한 거리값의 누락으로 임펄스 잡음이 거리영상에 첨가된다. 이것은 직접적인 3차원 모델링에 오차를 발생시키는 단점이 있다. 이런 단점을 극복하기 위해 밝기영상과 거리영상의 정보를 같이 고려하여 각각의 영상에서 유효한 정보를 결합하여 원하는 결과를 얻는 방법이 소개되었는데 이런 종류의 연구들은 센서 fusion 분야에서 많이 이루어졌다.^[2-5]

Fusion 기법은 서로 다른 영상에서 나오는 정보를 결합하는 방법에서 크게 경험적인 방법과 Bayes의 정리에 근거한 방법으로 분류할 수 있다. 거리영상과 밝기영상의 fusion의 초기 연구 단계에서는 주로 경험적인 접근방식이 많이 연구되었는데 화소간의 논리 합에 근거한 Gil 등의 방법^[3] 등이 제안되었다. 그러나 경험적인 접근방식에는 센서의 불확실성을 통합하는 데 효율적인 규칙을 찾기 어렵기 때문에 영상간의 관계에서 생기는 장점에 대한 일반성이나 강건성이 결여되는 단점이 있다. 이에 대해 Bayes의 정리는 센서의 불확실성을 쉽게 통합할 수 있는 장점을 가지고 있다. Zhang과 Wallace^[4]의 방법과 Nadabar와 Jain의 방법^[5] 등은 Bayes의 정리에 근거한 접근방법으로 센서의 불확실성을 모델링하는 선택적인 정보와 관측치에 기반한 후험적인 정보를 결합하는 과정을 통해 원하는 결과를 얻는다.

본 논문의 구성은 다음과 같다. 먼저 Bayes 추정기에서 선택적인 내용을 모델링하는 Markov random field (MRF)에 대해 II장에서 고찰한다. III장에서는 fusion을 이용한 제안한 영상 분할에 대해 기술하고 IV장에는 실험영상에 대한 실험 결과를 보인다. 마지

막으로 V장에서 결론을 내렸다.

II. MRF

MRF는 많은 응용 분야에서 영상 모델링에 사용되어 왔다.^[2-8] 이 절에서는 fusion 과정에서 선택적인 정보를 모델링하는 MRF의 특성에 대해 알아보고 선택적인 정보의 모델링에 대해 알아본다.^[9]

1. MRF와 Gibbs 확률 분포

$F = \{F_1, \dots, F_m\}$ 가 집합 S 에서 정의된 랜덤 변수의 집합일 때 F_i 는 대응 함수 g 의 지역의 한 원소 f_i 를 그 값으로 갖는다. 이런 성질을 갖는 F 를 랜덤 필드라 한다. F_i 가 f_i 의 값을 취하는 사건을 $F_i = f_i$ 라 표기하고 F_i 의 joint 사건을 $(F_1 = f_1, \dots, F_m = f_m)$ 으로 표시하면 근접 시스템 N 과 집합 S 에 대해 다음의 두 조건을 만족하는 F 를 MRF라 한다.

- 1) 확실성: $P(f) > 0, \forall f \in F$
- 2) Markovianity: $P(f_i | f_{S-(i)}) = P(f_i | f_N)$ (1)

여기서 $S-(i)$ 는 주어진 집합에서 위치 i 를 제외한 집합을 의미하고 f_N 는 위치 i 의 근접 시스템의 위치들의 라벨의 집합을 의미한다. 이 정의에 의해 라벨은 근접한 라벨들에 의해서만 직접적인 상호 작용을 한다. Hammersley-Clifford 정리^[8]에 의해 MRF F 는 Gibbs 확률 분포를 갖는다. Gibbs 확률분포의 에너지 함수로부터 MRF의 joint 확률분포를 구할 수 있다. 확률 분포가

$$P(f) = \frac{1}{Z} \exp\left\{-\frac{1}{T} U(f)\right\} \quad (2)$$

의 형태를 보일 때 이것을 Gibbs 확률분포라 한다. 여기서 T 는 온도 상수이고 보통 1로 가정되고, $U(f)$ 는 에너지 함수로

$$U(f) = \sum_{c \in C} V_c(f) \quad (3)$$

로 정의되는데 여기서 $V_c(f)$ 와 Z 는 각각

$$V_c(f) = \text{clique } c \text{에서의 potential}$$

$$Z = \sum_{f \in F} \exp\left\{-\frac{1}{T} U(f)\right\} \quad (4)$$

와 같이 정의된다.

MRF를 사용하여 영상을 모델링하기 위한 MRF의 조건부 확률밀도함수를 추정하는 것은 Gibbs 확률분포함수에 의해 만들어지는 MAP (maximum a posteriori) 추정기를 최적화하는 방법을 이용한다.

2. MAP 추정기의 최적화 기법

Geman과 Geman^[6]은 MAP 추정에 대한 근사화를 위하여 모든 경우를 고려하는 simulated annealing (SA) 방법을 제안했으나 이 방법에는 많은 계산시간을 요구하는 단점이 있다. SA의 많은 계산량을 줄인 방법인 iterative conditional modes (ICM) 방법^[7]은 위치의 라벨을 반복적으로 갱신하는 방법이다. ICM은 SA보다 계산량이 적지만 국부 최소 (local minima)에 쉽게 빠지는 단점이 있다.

Chou와 Brown^[8]은 효율성, 예측성 그리고 강건성의 세가지 특성에 근거하여 highest confidence first (HCF)라고 알려진 최적화 알고리즘을 제안하였다. 효율성은 반복적인 과정이 결정론적이고 각 단계가 최대로 개선되어 계산비용이 절감되는 것을 뜻한다. 예측성의 측면에서는 최종적인 추정결과는 입력과 선택된 선택적인 정보에 의존하므로 초기화와 같은 성능에 영향을 줄 수 있는 파라미터의 설정과 무관해야 한다. 강건성은 잡음이나 모델링 오차가 증가해도 최종 추정 결과에는 영향이 없어야 하는 것이다.

HCF는 ICM과 같이 결정론적이고 반복적인 갱신을 이용하는데 ICM이 무작위로 방문하는 위치를 선택하는 반면 HCF는 가장 안정도가 낮은 위치부터 선택하는 방법으로 방문순서에 따른 계산 효율을 증가시킨다. 라벨링에 대한 후험적 지식은 Gibbs 분포에 의해 표현된다.^[8-9] O 를 관측치라고 하면 라벨 f 의 후험적 확률은 clique의 집합 C 에서

$$P(f|O) = \frac{1}{Z} \exp\left\{-\frac{1}{T}\left(\sum_{c \in C} V_c(f) - T \sum_{i \in S} \ln[P(O_i|f_i)]\right)\right\} \quad (5)$$

로 나타낼 수 있다. 여기서 $V_c(f)$ 는 potential 함수이고 이것은 현재 근접 시스템의 구성 (configuration)에 대한 연속성을 보장해준다. O_i 와 f_i 는 위치 i 에서 관측치와 라벨을 의미한다.

HCF 방법에서는 모든 위치들은 라벨이 없는 상태 (l_0)로 초기화된다. HCF의 기본 개념은 초기화된 구성에서 F 의 측면에서 에너지 $U(f)$ 가 최소가 되는 구성을 찾는 것이다.

후험적 확률로부터 위치 i 의 후험적 에너지 함수를 유도하면

$$E_i(f) = \sum_{c \in C} V_c'(l) - T \ln P(O_i|f) \quad (6)$$

와 같다. 여기서 $V_c'(l)$ 은

$$V_c'(l) = \begin{cases} 0, & \text{for } l = f \text{ or } l = l_0 \\ V_c(l), & \text{otherwise} \end{cases} \quad (7)$$

이다. 후험적 국부 에너지는 현재의 근접 시스템의 구성에 대해 라벨을 보장해주는 척도가 된다. 여기서 라벨이 없는 상태의 위치는 근접 시스템의 구성에 직접적으로 영향을 주지는 못하지만 근접된 위치의 라벨링에는 고려가 된다.

갱신 순서는 가장 안정성이 낮은 위치를 우선으로 한다. 여기서 안정성에 대한 척도는

$$G_i(f) = \begin{cases} \min_{k \in L, k \neq f} \Delta E_i(k, f_i), & \text{if } f_i \in L \\ -\min_{k \in L, k \neq j} \Delta E_i(k, j), & \text{if } f_i = l_0 \end{cases} \quad (8)$$

으로 정의되고 여기서 $\Delta E_i(a, b) = E_i(a) - E_i(b)$, $E_i(j) = \min_{k \in L} E_i(k)$ 이며 이것은 다른 라벨의 상태에 대한 현 상태의 선택적 지식과 관측치의 혼합된 척도이다. G 의 음의 값은 그 위치에서 다른 라벨의 상태로 발생하는 구성이 더욱 안정하다는 것을 나타낸다. 식 (8)에서 라벨이 없는 상태의 위치의 안정화 척도는 항상 음의 값을 갖는다. G 값의 크기는 라벨의 변화에 의한 에너지의 손실 또는 이득의 정도를 보이므로 G 의 음의 값이 클수록 라벨이 변화될 확률이 커진다. 초기에 라벨이 없는 상태에서는 관측치에 의한 likelihood에 의해서만 안정화 척도가 구성되므로 최대 likelihood 추정과 동치가 된다.

HCF는 G 의 값에 따라 방문 순서가 결정되는 heap 구조로 구현할 수 있다. 이것은 가장 작은 G 값을 갖는 위치를 heap의 최상의 자리에 놓는 방법이다. 가장 작은 G 를 갖는 위치 i 의 상태를 바꾸고 G 를 갱신하여 heap을 다시 조정하고 위치 i 의 근접시스템의 G 를 갱신한다. 이런 과정의 반복에서 안정화 척도 G 의 최소값이 양의 값이 되면 모든 위치의 상태가 안정화 된 것이므로 수행을 끝낸다. 라벨 변화의 과정은 위치 i 의 라벨을

$$l = \begin{cases} \min_i \Delta E_i(l, f_i), & \text{if } f_i \in L, \\ \min_i E_i(l), & \text{if } f_i = l_0 \end{cases} \quad (9)$$

의 식을 만족하는 라벨 l 로 바꾸는 과정이다. 이 $Change_label$ 이 일어나는 위치에서 안정도는 양의 값을 갖게 되고 라벨의 변화에 의해 그 위치의 근접 시스템은 새로운 구성을 갖는다.

III. 거리영상과 밝기영상의 fusion을 이용한 제한한 영상분할

1. 특징 추출

1) 밝기영상에서의 특징추출

밝기영상에는 밝기의 변화에 대한 정보가 존재한다. 밝기 정보는 3차원 물체에 대한 각종 에지 정보를 가지고 있는데 밝기 정보는 물체간에 존재하는 계단형 에지와 면 사이에 존재하는 지붕형 에지, 그리고 물체간의 가림 등에서 발생하는 그림자 에지와 반사광의 각도 차이와 같은 albedo 에지 정보 등을 제공한다. 밝기영상에서 얻을 수 있는 많은 에지 정보중에서 실제로 3차원 물체의 인식에서 필요한 에지정보는 계단형 에지와 지붕형 에지이다. 그림자 에지와 albedo 에지는 3차원 물체의 분할과 인식을 방해하는 요소이다. 따라서 본 논문에서는 기존의 방법과 같이 에지를 여러 부류로 분류하지 않고 {EDGE, NO_EDGE}의 라벨 집합을 설정한다.

밝기영상에서 에지를 검출하기 위해 일반적으로 gradient를 이용한다.^[10] 그러나 gradient는 잡음이 존재할 경우 그 값에 영향을 많이 받기 때문에 잡음에 강건한 특징을 필요로 하는 방법에서는 많이 사용되지 않는다. 본 논문에서는 각 화소를 중심으로 밝기값의 분산을 특징값으로 한다. 가우시안 잡음을 가정하면 밝기의 불연속이 존재하는 지역에서 분산은 큰 값을 보이는 반면 균일한 지역에서는 작은 값을 갖는다. 그러나 임펄스 잡음을 가정한다면 임펄스 잡음이 있는 지역에서 분산은 큰 값을 갖는다. 큰 제공오차와 작은 제공오차를 제외한 α 트림드 분산 (α -trimmed variance)은 임펄스 잡음에 강건한 특성을 갖는다. α 트림드 분산은 robust 통계학에서 사용되는 α 트림드 방법^[11]을 이용하여

$$\sigma_a^2 = \frac{1}{(1-2\alpha)n} \sum_{i=an}^{(1-\alpha)n} (I - \bar{I})_{i,n}^2 \quad (10)$$

와 같이 큰 오차와 작은 오차 부분을 제외한 나머지 부분으로 분산을 계산하는 방법이다. 여기서 \bar{I} 는 α

트림드 평균으로

$$\bar{I} = \frac{1}{(1-2\alpha)n} \sum_{i=an}^{(1-\alpha)n} I_{i,n} \quad (11)$$

와 같이 정의되고 α 는 0과 0.5 사이의 값이며 $I_{1:n} \leq I_{2:n} \leq \dots \leq I_{i:n}$ 는 오름정렬된 밝기값의 관측치이고 $(I - \bar{I})_{1:n}^2 \leq (I - \bar{I})_{2:n}^2 \leq \dots \leq (I - \bar{I})_{i:n}^2$ 은 오름정렬된 제곱차이다. α 가 작으면 잡음에 영향을 받기 쉬운 반면 α 가 크면 계산에 사용되는 데이터의 수가 작게 되어 분산값의 오차가 커진다.

2) 거리영상에서의 특징추출

거리영상에서 지붕형 에지는 일반적으로 Laplacian을 취하여 검출하지만 잡음이 첨가된 경우 잘못 검출하는 경향이 있다. 이런 이유로 거리영상의 거리값을 측정치로 국부적인 평면계수를 추정하여 평면계수를 특징값으로 하는 방법이 제안되었다.^[12] 이 방법에서는 국부 영역을 평면으로 가정하여 평면의 방정식의 계수를 추정한다.

평면계수를 추정하는 기법은 robust 통계학에서 제안된 LTS^[13] 방법을 사용한다. LTS 방법은

$$\min \sum_{i=1}^h (r^2)_{i,n} \quad (12)$$

와 같이 오차의 일부의 합을 최소화하는 비선형적인 식이다. 여기서 $(r^2)_{1:n} \leq \dots \leq (r^2)_{n,n}$ 은 오름정렬된 제곱오차이고 h 는 1과 n 사이의 값이고 h 가 $n/2$ 일 때 breakdown point는 0.5이다. 여기서 breakdown point는 전체 데이터량에서 추정을 왜곡시키는 데이터량의 비를 뜻한다. LTS 방법은 높은 breakdown point를 갖는 반면 물체의 모서리 부분을 잡음의 영향으로 판단하기 때문에 모서리의 일부가 유실되는 경우가 있다. 추정영역의 크기가 크면 추정 성능은 좋아지나 계산시간이 증가하고 추정영역의 크기가 작으면 추정 성능이 떨어진다. 제한한 방법에서는 9×9 영역에 대해 화소 단위로 LTS 방법을 적용하여 평면계수를 추출하여 이것을 거리영상의 특징값으로 한다.

LTS 기법을 이용하여 평면계수를 추정하면 두 개의 평면이 만난 부분에서 상대적으로 큰 평면의 계수가 추정된다. 추정된 평면계수로 다른 평면을 fitting하면 추정오차가 증가하기 때문에 에지의 존재 가능성으로 추정오차를 사용할 수 있다. 화소 (x, y) 에서 제곱 추정오차는

$$\varepsilon^2(x, y) = \sum_{i=-4}^4 \sum_{j=-4}^4 \{z(x+i, y+j) - \hat{z}(x+i, y+j)\}^2 \quad (13)$$

와 같다.

2. 영상분할

1) Likelihood

본 논문에서는 II장에서 언급한 MAP 추정기의 최적화를 통한 에지추출을 제안한다. 에지는 영상에서 특징값의 불연속을 의미한다.

MRF를 모델링하는 데 있어 중요한 요소가 likelihood의 정의이다. 본 논문에서는 밝기영상과 거리영상에서 얻을 수 있는 특징값을 토대로 likelihood를 정의하였다.

라벨 위치 집합은 EDGE와 NON-EDGE의 원소만 가정하여 $L = \{0, 1\}$ 로 설정한다. 여기서 '1'은 EDGE 위치를 의미하고 '0'은 NON-EDGE 위치를 의미한다. 밝기영상에서 likelihood $P(O_I|f)$ 는 α 트립드 분산 $\sigma_{o_i}^2$ 을 이용하여 정의한다. 밝기영상에서 에지가 존재하면 그 분산값이 커지고 에지가 없는 영역은 분산이 작으므로 likelihood를

$$P(O_I|f=0) = \frac{1}{Z_1} \exp(-\sigma_{o_i}^2) \quad (14)$$

$$P(O_I|f=1) = \frac{1}{Z_2} \exp\left(-\frac{1}{\sigma_{o_i}^2}\right) \quad (15)$$

과 같이 정의하였다. 여기서 Z_1, Z_2 는 likelihood에 대한 정규화상수이다.

거리영상에서 likelihood $P(O_R|f)$ 는 국부적으로 추정된 평면계수 벡터를 이용하여 정의한다. 두 개의 평면이 만난 영역의 추정오차는 크고 단일 평면의 영역에서는 추정오차가 작으므로 likelihood를

$$P(O_R|f=0) = \frac{1}{Z_3} \exp(-\varepsilon^2) \quad (16)$$

$$P(O_R|f=1) = \frac{1}{Z_4} \exp\left(-\frac{1}{\varepsilon^2}\right) \quad (17)$$

과 같이 정의하였다. 이 경우 평면계수 벡터의 크기는 1로 정규화된다. 여기서 ε^2 은 식 (13)의 평면계수 제곱 추정 오차이고 Z_3 과 Z_4 는 likelihood에 대한 정규화상수이다.

밝기값은 물체의 색과 반사특성, 조명환경에 의해 큰 영향을 받고 거리값은 기준위치와의 거리와 3차원

기하학적인 위치에 의해 결정된다. 밝기영상에서 그림자의 경우 광원의 위치와 물체의 형상에 의해 결정되나 본 논문에서는 그림자에 의해 발생한 밝기영상의 에지는 잘못된 에지라 판단하므로 밝기값과 거리값은 독립적인 관측치로 가정한다. 밝기값과 거리값이 서로 통계적으로 독립이라고 가정하면 구하고자 하는 likelihood는 밝기영상의 likelihood와 거리값의 likelihood의 곱으로 표현할 수 있다.

$$P(O|f) = P(O_I|f) \cdot P(O_R|f) \quad (18)$$

(18)식을 (6)식에 대입하여 MAP 추정기를 구성하고 HCF 방법으로 추정기를 최적화한다.

2) Clique potential 함수

Potential 함수는 에지 위치의 공간적인 관계에 대해 segment의 연속성의 보장 및 그 위치가 잘못 선택되는 것을 막아 주는 효과가 있다. 기존의 방법에서는 potential 값을 실험적으로 최적의 값을 찾아서 사용하였다. 본 논문에서는 다음과 같은 clique potential 함수를 제안한다.

$$V_c = -\lambda_0(1 - l_k)\|a_k - a_j\|^2 - \lambda_1 l_k \quad (19)$$

여기서 λ_0 과 λ_1 은 가중치 상수이고 l_k 는 화소 k 와 j 간의 에지 상태로서 에지일 때 $l_k=1$, 에지가 아닐 때 $l_k=0$ 을 의미한다. c 는 clique의 집합으로 $i, j \in c$ 이고 a_k, a_j 는 각각 화소 k 와 j 에서의 평면계수 벡터이다. 여기서 clique은 1차와 2차의 경우만 고려한다. 에지가 존재하지 않는 clique에서는 $\lambda_0\|a_k - a_j\|^2$ 의 potential 값이 감소되고 에지가 존재하는 clique에는 λ_1 만큼 potential 값이 감소된다. 가중치 λ_0 은 평활화의 양을 조절한다. λ_0 가 크면 평활화의 효과가 커져서 두 개의 독립된 영역이 하나의 영역으로 잘못 나오고 λ_0 가 작으면 하나의 영역이 두 개의 영역으로 분할된다. λ_1 은 선형적인 정보로 에지의 상태가 발생하는 정도를 나타낸다.

그림 1은 제안한 방법에 대한 흐름도를 보인다. 각각의 영상에서 잡음에 강건한 특징으로 likelihood를 구성하고 선형적인 내용을 MRF로 모델링한다. Clique에 대한 potential 함수를 설정하고 MAP 추정기를 구성하여 최적화를 통해 에지를 구한다. 에지 지도를 근거로 하여 영상분할을 한다.

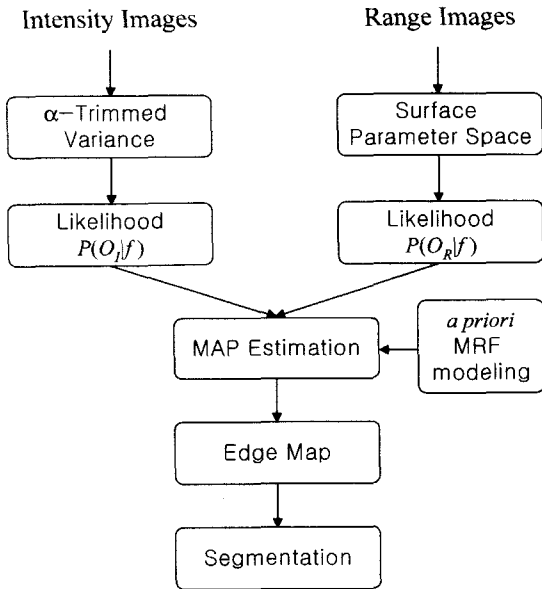


그림 1. 제안한 알고리즘의 흐름도
Fig. 1. Flowchart of the proposed algorithm.

IV. 실험 결과 및 토의

이 장에서는 기존의 fusion을 이용한 에지추출 방법과 제안한 방법의 성능을 비교하기 위해 수행한 컴퓨터 모의실험에 대한 결과에 대해 논한다.

1. 실험영상

본 논문에서 제안한 알고리즘의 타당성을 실험하기 위해 그림 2의 실험 영상 set를 사용하였다. 사용한 실험 영상은 미시간 주립대학의 Pattern Recognition and Image Processing (PRIP) 연구실에서 제작된 240×240 크기의 실제 영상으로 전처리를 통해 각 화소에 대한 좌표 위치는 서로 대응되어 있다 (ftp.cps.msu.edu/pub/prip/data/range_images).

실제 영상을 취득할 때 몇 개의 화소에 대해 정보가 손실되어 원영상에는 hole이 존재한다. 따라서 이 hole을 임펄스 잡음으로 설정하는데 기존의 방법들은 임펄스 잡음에 강건하지 못하므로 전처리 과정을 통하여 hole을 제거하였다. 밝기영상에 대해서는 3×3 메디안 필터를 사용하여 hole을 제거하였고 거리영상에 대해서는 hole에 대해 평면 파라미터를 이용한 복원기법^[14]을 적용하였다.

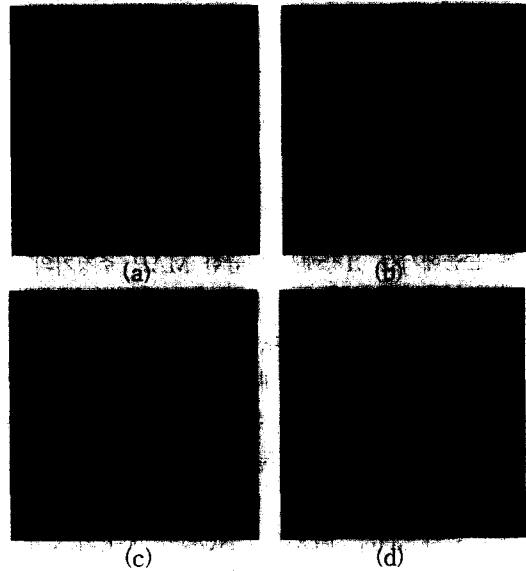


그림 2. 실험에 사용한 밝기영상과 거리영상(a) GB 밝기영상 (b) GB 거리영상 (c) Block+Column 밝기영상 (d) Block+Column 거리영상

Fig. 2. Intensity and range images used in experiments.

(a) GB intensity image (b) GB range image (c) Block+Column intensity image (d) Block+Column range image

2. 에지추출

그림 2의 밝기영상과 거리영상에 각각 $\sigma=3.0$ 의 가우시안 잡음과 5%의 임펄스 잡음을 첨가하여 실험영상으로 사용하였다. 기존의 방법인 Gil 등의 방법, Zhang과 Wallace의 방법, 그리고 Nadabar와 Jain의 방법과 제안한 방법의 성능을 비교하였다. 기존의 방법은 임펄스 잡음에 대해 강건성이 없기 때문에 메디안 필터를 통해 임펄스 잡음을 제거하였고 제안한 방법에서는 전처리를 하지 않았다.

그림 3은 GB영상에 대한 likelihood를 200배 확대하여 그림으로 나타낸 것이다. 그림에서 밝은 부분이 확률이 높은 영역이고 어두운 부분이 확률이 낮은 영역이다. α 트림드 분산은 3×3 영역에 대해 가장자리 4개를 제외한 가운데 5개의 화소에 대한 부분 ($\alpha=0.22$)을 취해 계산하였고 9×9 영역에 대해 평면계수를 추정하였다.

그림 4(a), 5(a)는 Gil 등의 방법을 이용하여 에지를 추출한 것을 보이고 있다. Kirsch 연산자로 밝기영상에서 에지를 추출하고 거리영상에서는 법선 벡터 간 사이각의 불연속점을 검출하여 거리영상의 에지를

구하였다. 그러나 거리영상에서 에지를 구하는 방법은 잡음에 상당히 민감하고 밝기영상의 에지 검출기 역시 성능이 뛰어나지 않기 때문에 지붕형 에지가 모두 나타나지 않았고 잡음의 영향이 나타나 있다.

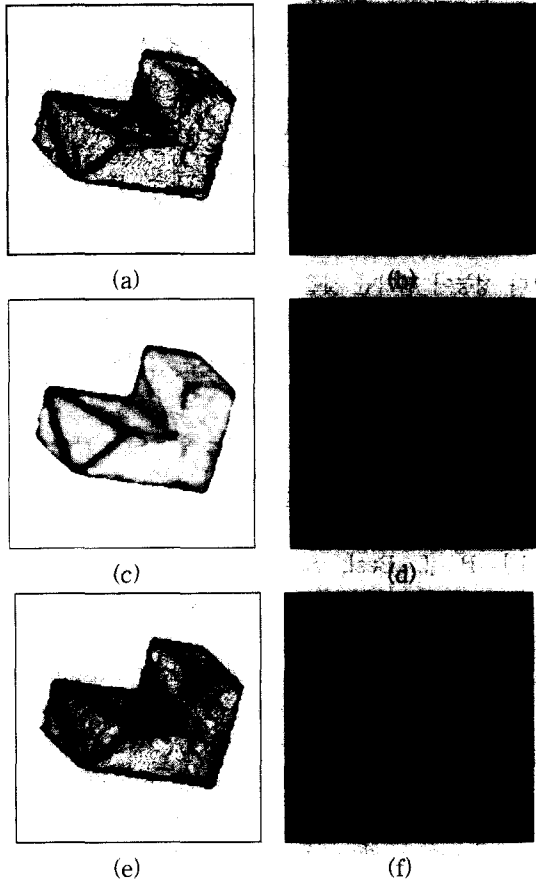


그림 3. GB영상에 대한 likelihood
 (a) 밝기영상의 $P(O_I|f=0)$ (b) 밝기영상의 $P(O_I|f=1)$ (c) 거리영상의 $P(O_R|f=0)$ (d) 거리영상의 $P(O_R|f=1)$ (e) $P(O_I|f=0)$ (f) $P(O_I|f=1)$

Fig. 3. Likelihood probability for GB images.
 (a) $P(O_I|f=0)$ (intensity image) (b) $P(O_I|f=1)$ (intensity image) (c) $P(O_R|f=0)$ (range image) (d) $P(O_R|f=1)$ (range image) (e) $P(O_I|f=0)$ (f) $P(O_I|f=1)$

그림 4(b), 5(b)는 Zhang과 Wallace의 방법을 이용하여 에지를 구한 결과이고 그림 4(c)와 5(c)는 Nadabar와 Jain의 방법을 이용하여 에지를 구한 결과이다. 이들의 방법에서는 각각의 에지 형태에 따라 라벨링했지만 결과는 {EDGE, NON_EDGE}의 2개의 라벨로 통합해서 보였다. 이 방법은 가우시안 잡음만

을 가정하였기 때문에 영상 취득시 생긴 비가우시안 잡음에는 성능이 저하된다. 그림 4(d)와 5(d)는 제안한 방법을 이용한 결과이다. λ_0 의 값은 15.0이고 λ_1 의 값은 20.0으로 실험적으로 정하였다. 제안한 방법에서는 잡음의 영향이 많이 제거된 것을 볼 수 있다. 거리영상의 특징인 평면 계수를 구하는 데 LTS를 사용했기 때문에 물체의 뾰족한 모서리 부분이 둥글게 나오는 현상을 볼 수 있다. 이 방법은 평면 계수를 구하는 데 다해상도 기법^[14]을 적용하면 모서리를 보존할 수 있는데 다해상도 기법은 계산시간이 커지는 단점이 있다.

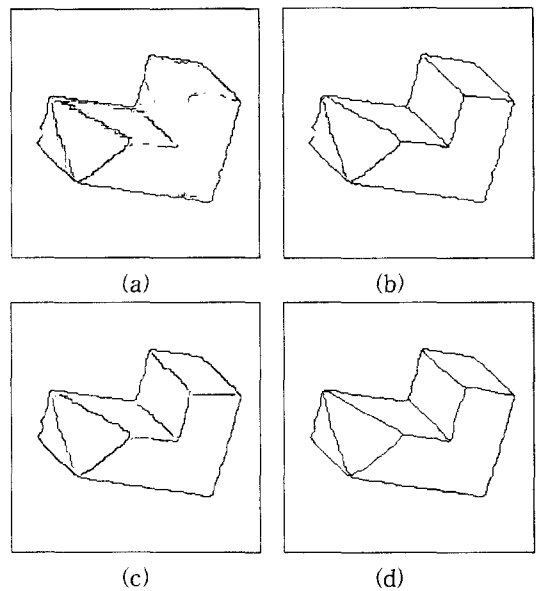
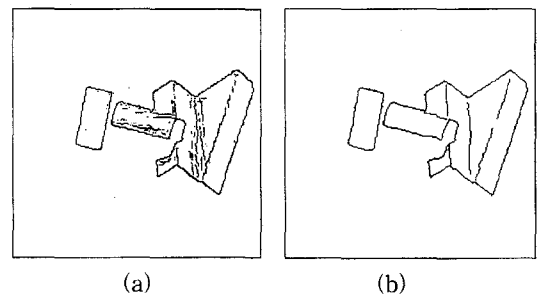


그림 4. GB 영상에 대한 에지추출 방법의 성능 비교
 (a) Gil 등의 방법 (b) Zhang과 Wallace의 방법 (c) Nadabar와 Jain의 방법 (d) 제안한 방법

Fig. 4. Performance comparison of edge detection methods for the GB images.
 (a) Gil et al.'s method (b) Zhang and Wallace's method (c) Nadabar and Jain's method (d) proposed method



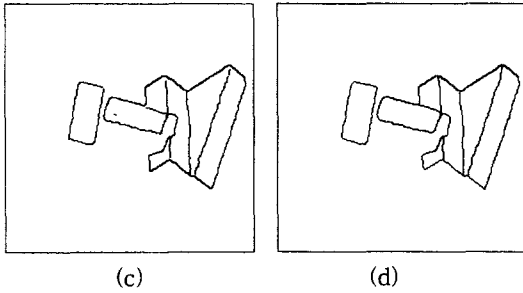


그림 5. Block+Column 영상에 대한 에지추출 방법의 성능 비교

(a) Gil 등의 방법 (b) Zhang과 Wallace의 방법 (c) Nadabar와 Jain의 방법 (d) 제안한 방법

Fig. 5. Performance comparison of edge detection methods for the Block+Column images.

(a) Gil et al.'s method (b) Zhang and Wallace's method (c) Nadabar and Jain's method (d) proposed method

각 방법의 수행시간은 다음과 같다. MIPS-5000 시스템에서 Gil 등의 방법은 20초가 걸리고 Zhang과 Wallace의 방법은 6분, Nadabar와 Jain의 방법은 8분이 걸렸다. 제안한 방법은 15분이 걸린다. 이것은 평면계수를 추정하는 과정에 약 10분의 많은 계산 시간이 필요하기 때문이다.

3. 영상분할

구해진 에지를 근거로 하여 각 실험영상을 분할하여 그림 6에 나타내었다. 잡음은 $\sigma=3.0$ 의 가우시안 잡음과 5%의 임펄스 잡음을 설정하였다. 평면으로 이루어진 부분과 곡면으로 이루어진 부분이 올바르게 분할되었다. 거리영상을 1차 다항식으로 근사화하여 곡면에서는 오차가 생기지만 밝기영상의 특징으로 오차의 영향이 감소되어 곡면도 하나의 부분으로 분할되었다.

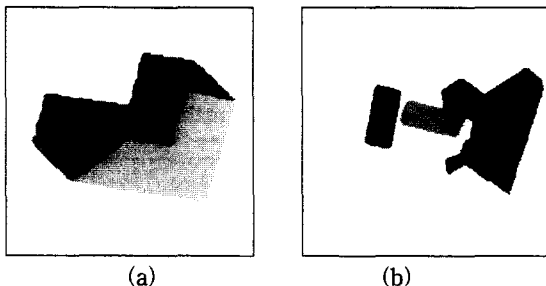


그림 6. 제안한 방법의 영상분할 결과

(a) GB 영상 (b) Block+Column 영상

Fig. 6. Segmentation results of the proposed method. (a) GB image (b) Block+Column image

V. 결 론

본 논문에서는 3차원 영상분할을 위해 기존의 거리영상분할이 아닌 밝기영상을 함께 고려하는 새로운 방법을 제시하였다. 밝기영상과 거리영상을 고려하는 방법으로 영상을 MRF로 모델링하여 조건부 확률을 최대화하였고 이것을 MAP 추정기로 구현하였다. 각 영상에서 관측치에 의한 likelihood를 위해 밝기영상에서 α 트림드 분산을 이용하였고 거리영상에서는 면계수를 특징으로 이용하였고 MAP 추정기의 최적화를 통해 3차원 물체를 분할한 결과를 보였고 기존의 방법보다 성능이 뛰어난 것을 보였다.

앞으로 거리영상에서 평면계수를 구하는 알고리즘의 속도 개선을 위한 연구와 곡면 등을 고차식으로 적용하는 연구가 필요하다.

참 고 문 헌

- [1] P. J. Besl, *Surfaces in Range Image Understanding*. Springer-Verlag, NY, 1988.
- [2] J. Clark and A. Yulle, *Data Fusion for Sensory Information Processing System*. Kluwer, Boston, MA, 1990.
- [3] B. Gil, A. Mitiche, and J. K. Aggarwal, "Experiments in combining intensity and range edge maps," *Computer Vision, Graphics, and Image Processing*, vol. 21, no. 1, pp. 395-411, Apr. 1983.
- [4] G. Zhang and A. Wallace, "Physical modeling and combination of range and intensity edge data," *CVGIP: Image Understanding*, vol. 58, no. 2, pp. 191-220, Sept. 1993.
- [5] S. G. Nadabar and A. K. Jain, "Fusion of range and intensity images on a connection machine (CM-2)," *Pattern Recognition*, vol. 28, no. 1, pp. 11-26, Jan. 1995.
- [6] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, no. 6, pp. 721-741, Nov. 1984.
- [7] J. Besag, "On the statistical analysis of

- dirty pictures," *J. Roy. Statist. Soc. B*, vol. 48, no. 3, pp. 259-302, Sept. 1986.
- [8] P. B. Chou and C. M. Brown, "The theory and practice of Bayesian image labeling," *Int. J. Computer Vision*, vol. 4, no. 3, pp. 185-210, June 1990.
- [9] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. Springer-Verlag, NY, 1995.
- [10] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679-698, Nov. 1986.
- [11] J. B. Bennar and T. L. Watt, "Alpha-trimmed means and their relationship to median filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no. 1 pp. 145-153, Feb. 1984.
- [12] A. Leonardis, A. Gupta, and R. Bajcsy, "Segmentation of range images as the search for geometric parametric models," *Int. J. Computer Vision*, vol. 14, no. 3, pp. 253-277, Apr. 1995.
- [13] P. J. Rousseeuw and A. M. Leory, *Robust Regression and Outlier Detection*. John Wiley & Sons, NY, 1987.
- [14] I. S. Chang, D.-G. Sim, and R.-H. Park, "Multiresolution surface parameter estimation for range images," in *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, vol. I, pp. 37-40, Sept. 1996.

 著者 소개

張仁壽(正會員) 第 34卷 S編 第 6號 参照

朴來弘(正會員) 第 35卷 S編 第 2號 参照

현재 서강대학교 전자공학과 교수