

폐색영역을 고려한 밝기 기반 쌍방향 스테레오 정합

신 홍철*, 주 재흠**, 이상욱***, 남 기곤*, 이 장명*
 *부산대학교 전자공학과, **지산 대학, ***경상 대학교
 부산광역시 금정구 장전동 산 30번지
 heshin@vision.ee.pusan.ac.kr

Intensity-Based Bidirectional Stereo Matching with Occlusions

Hong-Chul Shin*, Jae-Heum Joo**, Sang-wook Lee***, Ki-Gon Nam*, Jang-myung Lee*

*Dept. of Electronics Eng. Pusan National University, **Jisan College, ***Kyung Sang Univ.

요 약

본 논문에서는 스테레오 정합에서 불연속성과 폐색영역을 고려하고 다중 계층 구조를 사용한 relaxation 알고리즘을 이용하여 좌우 영상간의 최적의 정합점을 찾는 스테레오 정합을 구현하였다. 잡음을 제거하기 위해 베이지안 필터링 방법을 사용하였다. 좌우 영상에서 얻은 변이(disparity)를 이용하여 좌우 폐색영역(occlusion region)을 구한 후, 다시 최적의 변이를 얻기 위해 내삽과정을 통해 양쪽의 정보를 상호 보완하는 병렬적인 relaxation 방법을 구현한다. 구현한 방법은 다양한 영상에 적용하였다.

노와 local minima에 빠지지 않기 위해서 다중 계층 구조를 사용하였다.^[4] 또한 최적의 깊이정보를 얻기 위해 베이지안 방법으로 깊이정보의 모델화를 시도했다.^[5] 본 논문은 두 개의 종속적인 영상을 사용하므로 약간의 복잡성은 있지만, 잘못된 정합을 극복하기 위해 각각의 서로 다른 정보를 상호 보완적으로 이용하므로 잘못된 정합을 완화시키고 깊이정보의 정확성과 안정성을 증가시켰으며 이때, 정확한 폐색정보도 얻을 수 있었다.

I. 서론

스테레오 비전은 쌍안의 스테레오(binocular stereo)인 2-D 영상으로부터 일치점을 찾아 scene의 3차원 깊이(depth) 정보를 얻어내기 위한 방법이다. 이러한 스테레오 정합 방법은 크게 영역기반(Area-based), 특징기반(Feature-based), 밝기기반(Intensity-based) 세 가지 부분으로 대별할 수 있다.^[1,2] 특징기반은 특징부분만을 이용하여 정합하기 때문에 영상의 전체 영역에 대한 정확한 정보를 얻기 어려운 단점이 있다. 그러나 밝기기반은 스테레오 이미지 쌍의 서로 다른 밝기에 대처하기 때문에 특징기반보다는 알고리즘에 의해 밀집한 깊이정보를 얻을 수 있고 정합방법이 특징점에만 의존하지 않는 장점과 통계적인 최적치에 의해서 잡음 문제가 부분적으로 극복되는 장점 있다. 따라서, 본 논문에서는 밝기기반을 바탕으로 스테레오 정합을 구현한다. 스테레오 정합에 문제점은 폐색영역의 대해 알고리즘이 정확하게 다루어지지 않고 있다는 것이다.^[3] 따라서, 최적의 정합을 이루기 위해 cost 함수의 최소화 문제와 이완방법(relaxation method)을 통한 반복법(iterative method)을 구현하여, 특징기반과 폐색영역에서 나타나는 문제를 해결하고자 하였다. 빠른 수렴속

II. Bayesian model과 폐색

1. 폐색도(Occlusion Map)

스테레오 정합에서 폐색문제는 몇 가지의 기본 조건과 기하학이 폐색영역에서는 무용하기 때문에 중요한 역할을 한다. 폐색영역은 한 영상에서는 보이는데 다른 영상에서는 가려져있거나 보이지 않을 때 발생한다. 스테레오 정합을 위해 두 종류의 폐색영역은 반드시 분리하여 고려되어야 한다. 간단한 기하학 해석으로부터, 하나의 영상에 폐색영역은 다른 영상의 깊이정보의 불연속점에 대응하게 된다. 이런 두 폐색영역의 위치는 이전 폐색정보들에 의해 다음과 같이 표현되어질 수 있다.

$$O_s(x, y) = \begin{cases} 1 : (x, y) \notin s\text{-이미지폐색영역} \\ 0 : (x, y) \in s\text{-이미지폐색영역} \end{cases} \quad (1)$$

여기서, s는 좌측 혹은 우측을 나타낸다.

폐색도(Occlusion map) $O_s(x, y)$ 는 같은 영상의 $D_s(x, y)$ 로부터는 구하기가 어렵지만 다른 영상의 정확한 변이도 $D_r(x, y)$ 로부터 직접적으로 구할 수 있다. 즉, $O_s(x, y)$ 는 $D_r(x, y)$ 로부터 다음과 같은 방법으로 구해진다.

$$[(x, y)|O_l(x, y) = 1, (x, y) \in B] \\ = [(x_l, y)|x_l = x + D_r(x, y), (x, y) \in B] \quad (2)$$

$$[(x, y)|O_r(x, y) = 1, (x, y) \in B] \\ = [(x_r, y)|x_r = x - D_l(x, y), (x, y) \in B] \quad (3)$$

여기서, B는 영상의 전체영역을 나타낸다. discrete 한 경우에는, 기하학으로부터 좌측 영상의 모든 점 (x,y)는 선형적인(a priori) 지식 $D_l(x, y)$ 을 이용하여 우측 영상의 sub-point (x_r, y) 로 재투영된다. (x_r, y) 의 이웃한 두 개의 폐색도 $O_r(\lfloor x_r \rfloor, y)$ 와 $O_r(\lfloor x_r \rfloor + 1, y)$ 는 factor라는 변수에 의해서 값들이 누적된다. 이것은 subpixel에서 다른 이웃 점들 사이의 거리를 뜻한다. 좌측 영상에 대해 위의 작업을 한 후에, 적절한 Threshold(임계값-0.65)을 설정하여 작은 폐색영역과 잡음을 제거한다. 이 같은 처리과정을 아래와 같은 변환(Transformation) OT로 나타낼 수 있다.

$$O_L(x, y) = OT_{LR}(D_l(x, y)) \quad (4)$$

$$O_R(x, y) = OT_{RL}(D_r(x, y)) \quad (5)$$

따라서, 정확한 깊이(변이)노출 얻은 경우, 폐색영역 또한 정확하게 얻어지게 된다.

2. 변이도의 prior model

MRF(Markov random filed)처리의 확률 분포는 유한한 이웃에 의존하는데, 이것은 어떠한 local 특징을 가지는 변이도를 모델화하는데 사용되어진다. 변이도의 선형적 확률 분포는 아래와 같이 얻을 수 있다.^[6]

$$P(D_s) = \frac{1}{C_s} \exp\left(-\frac{U(D_s)}{T}\right) \quad (6)$$

여기서, T는 상수, C_s 는 정규화 상수이다. 에너지 함수 $U(D_s)$ 는 local potential 함수의 합으로 아래와 같이 쓸 수가 있다.

$$U(D_s) = \sum_{(x,y) \in B} V_{x,y}(D_s) \quad (7)$$

여기서, B는 영상내에 모든 점들의 집합이다.

potential 함수 $U(D_s)$ 의 선택은 물체표면의 국부적 특징을 모델화하는 방법에 직접적으로 반영되어진다. 그래서, 물체표면의 전체적인 평활에 관한 조건을 위해 아래와 같은 표준 선형적 평활 모델의 간단한 quadratic 함수로서 다음과 같이 사용하였다.

$$V_c(D_s) = (D_s(x+1, y) - D_s(x, y))^2 \\ + (D_s(x, y+1) - D_s(x, y))^2 \quad (8)$$

3. Bayesian model - MAP estimation

스테레오 이미지 $G_l(x, y)$ 와 $G_r(x, y)$ 이 주어질 경우에, 최적의 변이를 얻어야 하는데 최적의 변이도는 모든 확률에 걸쳐 아래와 같은 최대의 귀납적(a

posteriori) 확률을 가져야 한다.

$$\max_{D_l, D_r} P(Z_s | G_l, G_r) \quad (9)$$

$G_l(x, y) - G_r(x_r, y)$ 로부터 G_l 과 D_l , O_l 이 주어지면 G_r 의 조건부 확률을 아래와 같이 구할 수가 있다.

$$P(D_l | G_l, G_r) = \frac{1}{C} \exp\left(-\frac{1}{2\sigma^2} \sum O_l(x, y)(G_l(x, y) - G_r(x_r, y))^2\right) \quad (10)$$

Bayesian 방법에 의하여, 식(6)과 (7)으로부터 posteriori 확률을 아래와 같이 구할 수가 있다.

$$P(D_l | G_l, G_r) = \frac{P(G_l | D_l, G_r)}{P(G_l | G)} \\ = \frac{1}{C_b} \exp\left(-\frac{U_{bl}(D_l)}{T}\right) \quad (11)$$

여기서 C_{bl} 은 normalization 상수이고 귀납적 에너지 함수 $U_{bl}(D_l)$ 는 다음과 같이 주어진다.

$$U_{bl}(D_l) = \lambda \sum_{(x,y) \in B} O_l(x, y)(G_l(x, y) - G_r(x_r, y))^2 \\ + \sum_{(x,y) \in B} V_{x,y}(D_l) \quad (12)$$

또 다른 하나의 귀납적 에너지 함수 $U_{br}(D_r)$ 는 같은 방법으로 다음과 같이 구할 수가 있다.

$$U_{br}(D_r) = \lambda \sum_{(x,y) \in B} O_r(x, y)(G_l(x_l, y) - G_r(x, y))^2 \\ + \sum_{(x,y) \in B} V_{x,y}(D_r) \quad (13)$$

MAP estimation을 다른 방법인 귀납적 에너지 비용 함수를 최소화하는 문제로 변환할 수가 있다. 즉, 아래와 같은 최소화 과정을 해결함으로써

$$\min_{D_l, D_r} (U_{bl}(D_l) + U_{br}(D_r)) \quad (14)$$

폐색도를 고려한 최적의 변이도를 얻을 수 있다.

III. Relaxation 알고리즘

식(8)에 관한 평활 조건하에 비용 함수를 최소화하고 좋은 초기값을 선택하여 전체적 최적값(global optimal)을 찾기 위해서 relaxation 방법을 사용하였다. relaxation 과정은 식(14)을 Euler 방정식의 discrete 근사화함으로써 아래와 같은 식을 얻을 수가 있다.

$$D_s^{n+1}(x, y) = \overline{D_s^n} + \lambda O_l^n(G_l(x, y) - (G_r(x_r, y))G_{xx}(x, y)) \quad (15)$$

$$D_s^{n+1}(x, y) = \overline{D_s^n} + \lambda O_r^n(G_l(x_l, y) - (G_r(x, y))G_{xx}(x, y)) \quad (16)$$

$$\text{여기서, } x_r = x - \overline{D_l^n}, \quad x_l = x + \overline{D_r^n}, \quad G_{xx} = \frac{\partial^2 G_s(x, y)}{\partial x^2}$$

이며 $\overline{D_s^n}(x, y)$ 는 $D_s^n(x, y)$ 의 국부적 평균치로써 아래와 같이 표현할 수가 있다.

$$\overline{D_s^n}(x, y) = 1/9 (D_s^n(x-1, y-1) - D_s^n(x-1, y) + D_s^n(x, y) \\ + D_s^n(x-1, y+1) + D_s^n(x, y-1) + D_s^n(x, y+1) \\ + D_s^n(x+1, y-1) + D_s^n(x+1, y) + D_s^n(x+1, y+1)) \quad (17)$$

또한, 알고리즘의 안정성을 향상시키기 위해 식(11)

과 (12)로부터 아래와 같이 수정할 수가 있다.

$$D_{s,t}^{n+1}(x, y) = \overline{D_{s,t}^n} + \lambda O_s^n(G_t(x, y) \cdot (G_r(x_s^+, y) - G_r(x_s^-, y))) \quad (18)$$

$$D_{s,t}^{n+1}(x, y) = \overline{D_{s,t}^n} + \lambda O_s^n(G_t(x_s^+, y) - (G_r(x, y) - G_r(x_s^-, y))) \quad (19)$$

여기서, $x_s^+ = x - \overline{D_{s,t}^n}$, $x_s^- = x + \overline{D_{s,t}^n}$

$$O_s^n(x, y) = OT_{RL}(\overline{D_{s,t}^n}(x, y)), \quad O_r^n(x, y) = OT_{LR}(\overline{D_{s,t}^n}(x, y)) \quad (20)$$

$$\overline{D_{s,t}^{n+1}}(x, y) = \frac{\overline{D_{s,t}^n}(x, y) + D_{s,t}^n(x_s^+, y)O_s^n(x_s^+, y)}{1 + O_s^n(x_s^+, y)} \quad (21)$$

$$\overline{D_{s,t}^{n+1}}(x, y) = \frac{\overline{D_{s,t}^n}(x, y) + D_{s,t}^n(x_s^-, y)O_r^n(x_s^-, y)}{1 + O_r^n(x_s^-, y)} \quad (22)$$

위의 방정식들은 폐색도를 이용한 최적의 변이도를 얻어내는 relaxation 알고리즘의 완전한 순환식이다.

IV. 다중 계층 구조(Multilevel structure)

다중 계층 구조는 컴퓨터 비전 하위 단계에서 다양한 분해에 대해 넓게 사용된다. 특히, 국부 특징 때문에 발생하는 오차를 줄이고 비용 경비의 계산 절약과 함께 알고리즘의 수렴속도에 아주 효과적이다. 양쪽의 스테레오 영상에 Average pyramids를 생성하는데, 이것은 $0 \leq t \leq M(4)$ 범위를 가지며, 아래와 같이 나타낸다.

$$D_{s,t}(x, y) = (D_{s,t-1}(2x, 2y) + D_{s,t-1}(2x+1, 2y) + D_{s,t-1}(2x, 2y+1) + D_{s,t-1}(2x+1, 2y+1))/4 \quad (23)$$

최상위(coarsest) 단계에서 변이도 $D_{s,t}$ 와 폐색도 $O_{s,t}$ 를 0의 값으로 초기화한다. 이때, 알고리즘은 최상위 단계($t=4$)에서 최하위(finest) 단계($t=0$)까지 수행한다. 각 단계 t 에서 일어난 변이도 $D_{s,t}(x, y)$ 와 폐색도 $O_{s,t}(x, y)$ 의 최적값은 아래와 같은 내삽과정을 통해서 다음 하위 단계에 초기값으로 전달되어진다.

$$D_{s,t}(x, y) = D_{s,t+1}(x/2, y/2) \times 2 \quad (24)$$

$$O_{s,t}(x, y) = O_{s,t+1}(x/2, y/2) \times 2 \quad (25)$$

따라서, 본 논문에서는 최적의 변이도와 폐색도를 얻기 위해 각 단계마다 식(17)을 사용하여 식(18)부터 (22)까지 이용하였다.

V. 실험 및 결과

다중 계층 구조를 가지는 relaxation 알고리즘을 이용한 스테레오 정합 실험은 256×256 스테레오 영상을 이용했다. 최상위 단계의 변이도와 폐색도를 0으로 초기화하였으며, 이때 얻은 변이도와 폐색도를 다음 finer한 단계의 초기값으로 전달하는 형태로 알고리즘을 실행하였다. 실험은 disparity만을 이용한 수행 결과와 논문에서 강조하는 occlusion map을 도입했을 때의

수행 결과를 비교하였다. λ 는 0.0003의 값으로 주었고, 반복 횟수는 각 단계마다 50번으로 고정하였다. 알고리즘 수행은 펜티엄 150MHz의 시스템에서 하였고, 수행시 소요된 시간은 2분 35초였다.

그림 1.과 2에 나타난 것처럼 폐색영역을 도입했을 때와 하지 않았을 때의 결과는 잡음제거 뿐만 아니라, 표 1.에서 나타난 것처럼 정합률과 밝기차의 에러 값에서도 좋은 결과를 얻을 수가 있었다. 여기서, 정합률은 좌 우 영상의 밝기차가 전체밝기의 5%이내에 존재하는 점들을 정합점이라고 가정하여 비로 나타낸 것이며, 에러률은 각 점에 대한 밝기차의 절대값을 모두 합하여 비로 나타낸 것이다.

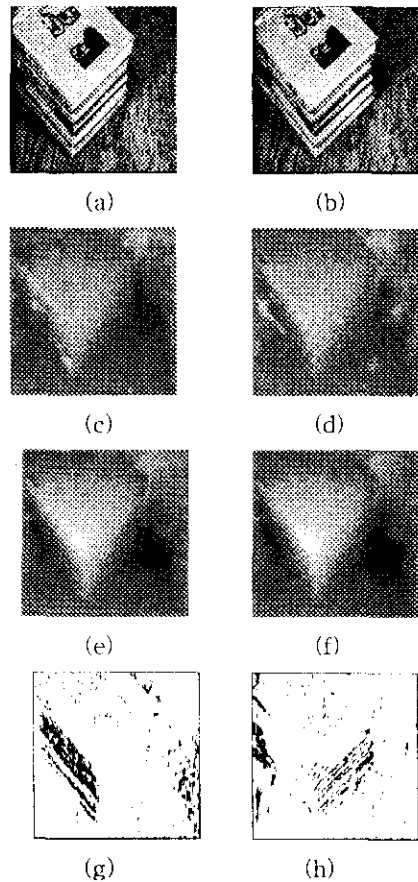


그림 1. Books 영상에 대한 결과 영상
(a) 좌측 원영상 (b)우측 원영상,
(c)(d) 변이만을 이용한 좌 우 결과 영상
(e)(f) 폐색영역을 고려한 좌 우 결과 영상
(g)(h) 좌 우 폐색 영상

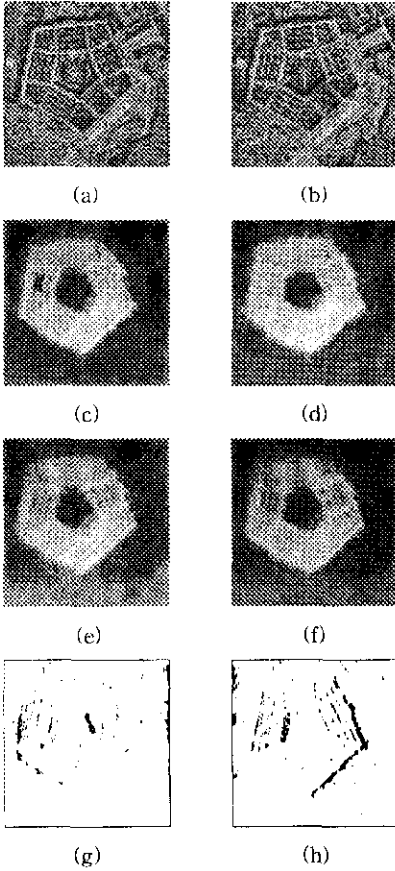


그림 2. Pentagon 영상에 대한 결과 영상

- (a) 좌측 원영상 (b) 우측 원영상,
- (c)(d) 변이만을 이용한 좌 우 결과 영상
- (e)(f) 폐색영역을 고려한 좌 우 결과 영상
- (g)(h) 좌 우 폐색 영상

표 1. 정합률과 에러율 비교

| 내용 \ 이미지 | Pentagon | | Books | |
|----------|----------|-------|-------|-------|
| | 정합률 | 에러율 | 정합률 | 에러율 |
| 폐색영역 도입 | 75.1% | 10.5 | 51.7% | 25.65 |
| 변이만을 이용 | 68.3% | 12.42 | 44.9% | 29.75 |

VI. 결론

스테레오 정합에서 문제화 되는 불연속성과 폐색성을 고려하여 최적의 변이율 얻기 위한 변이도의 확률적인 모델을 설정함으로써 비용 함수의 최소화 문제와 이완방법을 제안하였다. 그리하여 다중 계층 구조를 활용한 상호 보완적인 쌍방향 스테레오 정합 알고리즘을 구현하여 폐색영역을 얻을 수 있었고 이 정보를 활용하여 더 나은 변이 정보를 얻었다. 그 결과 폐색영역을 고려한 parallel relaxation 방법은 고려하지 않았을 경우의 결과보다 정합률과 에러율 측면에서 더 개선되었다. 향후, 에지부분에서의 블러링에 대한 연구가 필요하리라고 본다.

참고 문헌

- [1] 강장순, 남기곤, "변이율 정합 적도 분포를 이용한 선소의 정합", 전자공학회 논문지, 제34권 S편 제3호, pp. 74-83, 1997년 3월.
- [2] M.A.Gennert, "Brightness-based stereo matching", Proc. of 2nd Int. Conf. on computer Vision, Tampa, FL, 1988, pp. 139-143
- [3] D. Geiger, B.Ladendorf and A Yuille, "Occlusion and binocular stereo", Proc. of 2nd Europe Conf. on Computer Vision, 1992, pp. 425-33
- [4] C.Chang and S.Chatterjee, "Multiresolution stereo-a Bayesian approach", Proc. of 10th Int. Conf. on Pattern Recognition, Atlantic City, NJ, 1990, pp. 908-912
- [5] P.J.Burt, "The pyramid as a structure for efficient computation", in the book[22], 1984, pp. 6-35
- [6] S.Geman and D.Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images", IEEE Trans Pattern Anal. Mach. Intelligence, Vol.6, 1984, pp. 721-741