

변이공간영상에서의 비용 함수의 결정

준회원 박 준 희*, 종신회원 이 병 옥*

Determination of Cost Function in Disparity Space Image

Jun-hee Park* Associate Member, Byung-uk Lee* Lifelong Member

요 약

변이공간영상(Disparity space image, DSI) 방법은 스테레오 영상간의 정합을 찾는 방법이다. 이 방법은 각 화소 단위로 변이값을 계산해 내는 장점이 있다. DSI 방법은 비용함수를 최소화시키는 정합을 찾는 방법이다. 이 비용함수에서 폐색영역비용 값과 정합 보상값을 경험적으로 정하여왔다. 본 논문에서는 변이공간영상 방법에서 폐색영역비용과 정합보상값이 영상 잡음과 물체와 배경간의 차이에 영향을 받는 것을 이론적으로 분석하였다.

Key Words : Disparity space image, Stereo matching, Occlusion, Disparity estimation

ABSTRACT

Disparity space image (DSI) technique is a method of establishing correspondence between a pair of images. It has a merit of generating a dense disparity map for each pixel. DSI has a cost function to be minimized, and it needs empirical weighting factors for occlusion penalty and match reward. This paper provides theoretical basis for the weighting factors, which depend on image noise and contrast between an object and background.

I. 서 론

스테레오 영상 인식은 두 개의 2차원 영상에서 3차원 깊이 정보를 찾아내는 방법이다. 영상의 밝기정보를 사용하여 두 개의 영상간의 대응점을 찾아낸다. 이 대응점들의 변이(disparity)는 카메라로부터의 깊이에 반비례하므로 이로부터 깊이를 측정할 수 있다. 스테레오 영상 기법은 3차원 정보의 복원^[1], 동영상 부호화^[2], 다시점 영상 합성^{[3][4]}에 사용된다. 스테레오 정합 알고리즘은 크게 전역방법과 국소 방법으로 나눌 수 있다^{[5][6]}. 전역방법은 스캔라인이나 전체 영상에 대해 최적의 대응점을 찾는 방법이다. 이 방법은 텍스처가 없는 부분이나, 한쪽 영상에서는 보이지만 다른 영상에서는 가려진 폐색영역(occlusion)에서도 잘 동작한다. 그러나 이 방법은 각 부분 비용 함수를

더하여 최적의 비용함수를 구하므로 계산 속도가 느리고 부분적인 오차가 전파 될 수 있는 단점이 있다. 전역방법은 계산량이 많으므로 분할 및 정복 기법(divide and conquer algorithm)^{[7][8]}이나 동적 계획법^[9]을 사용한다. 그에 비하여 국소 방법은 관심있는 특징점이나 화소의 주변에 대해서만 대응점을 찾으므로 효율적이나 폐색영역을 잘 찾지 못하는 단점이 있다.

변이공간영상^[1,7,8,9]은 스캔라인단위로 영상의 정합을 최적화시키는 전역방법의 하나이다. 영상잡음으로 인하여 폐색영역으로 잘못 분류하는 것을 막기 위하여 깊이 불연속에 대한 페널티 함수를 넣거나^[1], 폐색영역비용(occlusion penalty cost)으로 상수값을 넣는다^[7,8,9]. 그러나 이러한 폐색 영역 비용값을 경험적인 값으로 정하여 사용해 왔다. 본 논문에서는 이 폐색영역비용을 정하는 방법

※ 본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음.

* 이화여자대학교 정보통신학과 영상처리연구실 (corelpark@gmail.com, bulee@ewha.ac.kr)

논문번호 : KICS2005-07-292, 접수일자 : 2005년 7월 18일, 최종논문접수일자 : 2007년 5월 2일

을 이론적으로 분석하여 이 값이 영상 잡음과 배경과 물체간의 밝기 차이에 영향을 받는다는 것을 밝혔다. 여기에서는 Birchfield와 Tomasi^[9]가 제안한 비용함수를 채택하여 폐색영역비용과 정합 보상값(match reward)을 정하는 방법을 분석한다.

본 논문은 다음과 같은 순서로 구성되어 있다. II장에서는 Birchfield와 Tomasi가 제안한 비용 함수에 대해서 살펴본다. III장에서는 변수 결정방법에 대해 설명한다. IV장에서는 실험과정 및 분석을 제시한다. 그리고 V장에서 연구의 결과 및 의의에 대해서 언급하였다.

II. 변이공간영상에서의 비용함수

Birchfield와 Tomasi가 제안한 화소간 스테레오 정합 알고리즘은 스캔라인에서 비용 함수를 최소화 하는 최적정합을 찾는다. 그림 1은 스캔라인에서 정합된 예를 보이고 있다. $I_L(x)$ 는 좌측영상의 밝기, $I_R(y)$ 는 우측영상의 밝기를 나타내며 x 와 y 는 각각 가로축을 표시한다. 뿔음 표시(○)는 폐색영역을 보인다. 위 그림에서 정합된 화소는 23개, 폐색영역의 개수는 6개이다.

비용 함수는 다음과 같이 정의하였다.

$$r(M) = N_{occ}k_{occ} - N_Mk_r + \sum_{i=1}^{N_M} d(x_i - y_i) \quad (1)$$

M : 대응점 집합

k_{occ} : 폐색영역비용, N_{occ} : 폐색영역 개수

k_r : 정합 보상, N_M : 정합된 화소의 개수

$d(x_i, y_i)$: x_i 와 y_i 가 대응 화소의 밝기일 때 화소 비유사성

여기에서 화소의 비유사성을 나타내는 함수 $d(x_i, y_i)$ 는

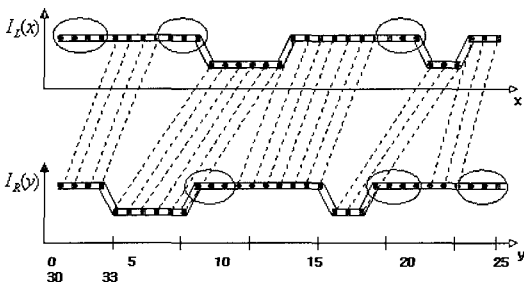


그림 1. 스캔라인과 매칭. 좌측 영상과 우측 영상의 정합을 나타내며 타원은 폐색영역을 나타낸다.

밝기 함수를 x_i 와 y_i 의 좌우 1/2 화소 범위에서 선형보간하여 $I_L(x)$ 와 $I_R(x)$ 의 차이가 제일 작은 경우의 값을 취한다. 이는 영상의 급격한 밝기 변화로 인하여 오차가 발생하는 것을 줄이기 위한 것이다^[10]. 화소 비유사성 함수가 스테레오 정합에서 중요한 역할을 하므로 이에 대한 연구가 계속되고 있다^[11]. 이 비용함수는 좌 우 영상의 밝기차이가 작을수록 함수값이 감소하며, 폐색 영역의 개수가 적고 정합된 화소가 많은 것을 선호한다. 여기에서 사용된 폐색영역 비용과 정합 보상 상수는 경험적인 값이다. 만일 입력된 영상에 잡음이 없고 물체와 배경의 밝기 차이가 크다면 정확한 정합에서 화소 비유사성 함수가 0이 되므로 폐색영역비용과 정합보상 상수를 0으로 주어도 오류 없이 영상 정합을 한다. 그러나 실제의 영상에서는 카메라 소자, 양자화 등으로 인한 여러 가지 종류의 잡음이 항상 존재한다. 따라서 영상에 텍스처의 존재 유무, 물체와 배경의 밝기 차이의 정도에 따라서 폐색영역 비용이 달라져야 한다. 본 논문에서는 여러 가지의 경우에 대하여 오류 없이 영상 정합을 할 수 있는 가중치를 정하기 위하여 각 경우를 분석한다.

III. 비용 함수의 상수 결정방법

이번 장에서는 정확한 정합을 찾을 수 있는 비용 함수의 폐색영역비용과 정합보상 값에 대한 분석을 한다. 정합에 오류가 발생하는 경우를 배경을 폐색영역으로 잘못 판단하는 경우, 물체와 배경의 경계선에서 오류가 발생하는 경우, 배경을 물체로 판단하는 경우로 나누었다. 각 경우에서 오류가 발생하지 않을 조건을 찾고 이 조건들이 영상 잡음과 물체와 배경의 밝기 차이에 영향을 받는 것을 밝혔다.

3.1 배경에서 정합이 잘못된 경우정합 보상 결정방법

그림 2에서 위쪽 점들이 왼쪽 영상의 스캔라인이고 아래쪽 점들이 오른쪽 영상의 극상선을 나타낸다. 실선으로 표시한 것이 정답일 때의 정합

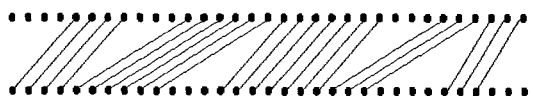


그림 2. 정합이 잘 된 경우



그림 3. 배경을 가려짐 영역으로 잘못 찾은 경우

M_1 을 보여준다. 이때의 가려짐 영역의 수는 N_{occ1} , 정합의 수는 N_{M1} 이다. 비용 함수는 다음과 같다.

$$r(M_1) = N_{occ1}k_{occ} - N_{M1}k_r + \sum_{i=1}^{N_{M1}} d_1(x_i - y_i) \quad (2)$$

그림 3은 배경에 텍스처가 없을 때 배경을 가려짐 영역으로 잘못 찾은 경우의 오차 정합 M_2 를 그린 것이다. 가려짐 영역의 개수는 N_{occ2} , 그리고 정합된 화소의 수는 N_{M2} 이다. 이때 비용 함수는 다음과 같다.

$$r(M_2) = N_{occ2}k_{occ} - N_{M2}k_r + \sum_{i=1}^{N_{M2}} d_2(x_i - y_i) \quad (3)$$

이때의 가려짐 영역의 개수는 정확한 정합의 개수와 같지만, 정합된 화소의 개수는 줄어든다. 그러므로 N_{occ1} , N_{occ2} , N_{M1} 과 N_{M2} 는 다음과 같은 관계식을 갖게 된다.

$$N_{occ1} = N_{occ2}, N_{M1} > N_{M2}$$

비용 함수는 정합이 잘 될 경우 (M_1)가 정합이 잘못된 경우 (M_2)보다 작아야 정답을 찾기 때문에 다음과 같은 식을 만족시켜야 한다.

$$\begin{aligned} r(M_1) - r(M_2) \\ = -(N_{M1} - N_{M2})k_r + \sum_{i=1}^{N_{M1}} d_1(x_i - y_i) - \sum_{i=1}^{N_{M2}} d_2(x_i - y_i) < 0 \end{aligned} \quad (4)$$

배경의 오정합은 화소 비유사성으로부터 많은 대가를 유발하지 않기 때문에 일어날 가능성이 많다. 최악의 경우는 배경과 물체가 모두 texture가 없을 때이다. 이 경우에 화소 비유사성의 차이를 분산이 인 zero mean Gaussian 분포로 가정한다 [11,12]. 영상잡음을 independent, identical distribution (i.i.d.) 가우시안 분포로 가정하면 분산 σ_1 은 영상잡음 σ 의 $\sqrt{2}$ 의 배가 된다.

$$\sum_{i=1}^{N_{M1}} d_1(x_i - y_i) - \sum_{i=1}^{N_{M2}} d_2(x_i - y_i) \sim N(0, \sqrt{2}\sigma)$$

오정합의 발생확률을 0.5% 이하로 줄이려면

$(N_{M1} - N_{M2})k_r$ 가 표준 편차의 세배 이상이 되어야 한다. 즉 $3\sigma_1 < (N_{M1} - N_{M2})k_r$ 식을 만족시켜야 한다. 그러므로 k_r 의 최소값은 다음 식에 의하여 결정된다.

$$k_r > \frac{3\sigma_1}{N_{M1} - N_{M2}} = \frac{3\sqrt{2}\sigma}{N_{M1} - N_{M2}} \quad (5)$$

위의 식에서 $N_{M1} - N_{M2}$ 가 1보다 큰 자연수 이므로 $k_r > 3\sqrt{2}\sigma$ 인 영역은 위의 (5)식의 충분조건이다.

3.2 물체와 배경의 경계에서 잘못 찾아진 경우

그림 4는 물체와 배경의 경계에서 정합점을 잘못 찾을 경우를 나타낸 것이다. 물체와 배경의 경계에서 대응점을 잘못 찾은 경우를 굵은 선으로 표시하였다. 잘못 찾아진 경우의 정합을 M_3 , 가려짐 영역의 개수를 N_{occ3} , 정합의 개수를 N_{M3} 이라고 하면 비용 함수는 다음과 같이 나타낼 수 있다.

$$r(M_3) = N_{occ3}k_{occ} - N_{M3}k_r + \sum_{i=1}^{N_{M3}} d_3(x_i - y_i) \quad (6)$$

이때의 정합된 화소의 수는 정확한 정합 (M_1)의 개수와 같지만 가려짐 영역의 개수는 늘어난다. 그러므로 N_{occ1} , N_{occ3} , N_{M1} 과 N_{M3} 은 다음의 관계를 갖는다.

$$N_{occ1} < N_{occ3}, N_{M1} = N_{M3}$$

M_3 의 비용 함수는 M_1 의 값보다 커야 하므로 다음의 식을 유도할 수 있다.

$$\begin{aligned} r(M_1) - r(M_3) \\ = -(N_{occ1} - N_{occ3})k_{occ} + \sum_{i=1}^{N_{M1}} (d_1(x_i - y_i) - d_3(x_i - y_i)) < 0 \end{aligned} \quad (7)$$

이때의 화소 비유사성도 식 (4)와 같이 zero-mean Gaussian 분포로 볼 수 있으므로

$$\sum_{i=1}^{N_{M1}} d_1(x_i - y_i) - \sum_{i=1}^{N_{M3}} d_3(x_i - y_i) \sim N(0, \sqrt{2}\sigma)$$

앞의 경우와 마찬가지로 k_{occ} 의 최소값을 다음 식에 의해 찾을 수 있다.

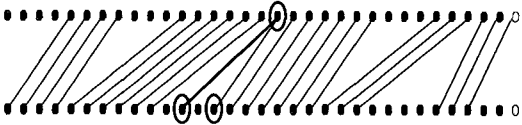


그림 4. 경계에서 잘못 찾아진 경우. 굵은 선은 잘못된 정합을 나타낸다.

$$k_{occ} > \frac{3\sigma_1}{N_{occ3} - N_{occ1}} = \frac{3\sqrt{2}\sigma}{N_{occ3} - N_{occ1}} \quad (8)$$

$N_{occ3} - N_{occ1}$ 은 물체와 배경의 경계에서 잘못 찾아진 개수를 의미하며 1보다 크다. 따라서 $k_{occ} > 3\sqrt{2}\sigma$ 의 영역은 위의 식 (8)을 항상 만족시킨다.

3.3 물체 사이의 배경 사이에서 잘못 찾아질 경우

그림 5는 물체와 물체 사이에 배경이 있을 때, 배경이 물체로 잘못 찾아진 경우를 굵은 선으로 표시하였고 잘 찾아진 경우를 가는 선으로 함께 표시한 것이다. 잘못 찾아진 경우의 정합을 M_4 , 가려짐 영역의 개수를 N_{occ4} , 정합의 개수를 N_{M4} 이라고 하면 비용 함수는 다음과 같이 쓸 수 있다.

$$r(M_4) = N_{occ4}k_{occ} - N_{M4}k_r + \sum_{i=1}^{N_{M4}} d_1(x_i - y_i) \quad (9)$$

가려짐 영역이 배경으로 정합되었기 때문에 가려짐 영역의 개수는 줄고 정합된 화소의 수는 증가한다. 그러므로 N_{occ1} , N_{occ4} , N_{M1} 과 N_{M4} 는 다음의 관계를 갖는다.

$$N_{occ1} < N_{occ4}, N_{M1} = N_{M4}$$

비용 함수는 정합이 올바른 경우가 정합이 잘못된 경우 (M_4) 보다 작아야 하기 때문에 다음의 식을 얻을 수 있다.

$$\begin{aligned} r(M_1) - r(M_4) &= (N_{occ1} - N_{occ4})k_{occ} + (N_{M1} - N_{M4})k_r \\ &\quad + \sum_{i=1}^{N_{M1}} d_1(x_i - y_i) - \sum_{i=1}^{N_{M4}} d_1(x_i - y_i) < 0 \end{aligned} \quad (10)$$

이와 같은 경우에는 가려짐 영역이 두 개 없어진 경우이므로 $N_{occ1} - N_{occ4} = 2$ 가 되고, $-(N_{M1} - N_{M4}) \equiv \alpha$ 는

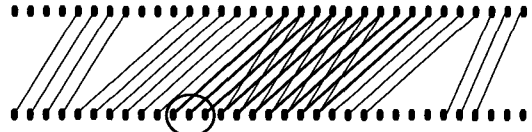


그림 5. 물체와 배경 사이에서 잘못 찾아질 경우

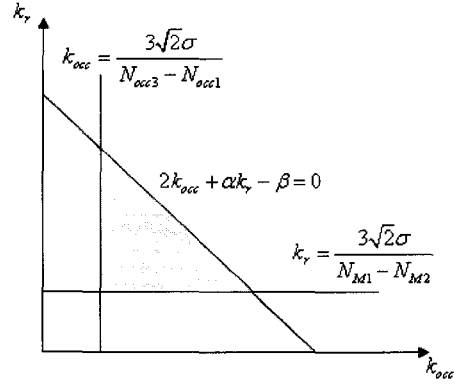


그림 6. 올바른 정합을 위한 폐색영역 비용과 정합 보상값의 영역

정합된 화소 수의 오차가 된다. 그리고 정합이 잘된 경우와 잘못된 경우의 화소 비유사성의 차이는 다음의 식과 같이 β 로 정의하였다.

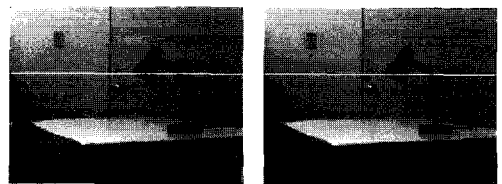
$$2k_{occ} - \alpha k_r + \beta < 0 \quad (11)$$

위 식 (11) 이 k_{occ} 와 k_r 의 범위를 나타낸다.

이상의 세 경우를 모두 만족시키는 폐색영역 비용과 정합 보상값의 영역은 그림 6과 같다. 폐색영역 비용과 정합 보상의 최소값은 영상 잡음에 비례한다. 만일 영상 잡음이 없는 경우 최소값은 0이 된다.

IV. 실험 및 분석

합성 영상과 실제 영상을 이용해 제안한 방법으로 폐색영역 비용의 범위를 실험해 보았다. 본 논문에서는 Birchfield와 Tomasi가 사용한 실험 영상 [9]을 예로 들어서 실험한 결과값을 제시한다.



(a) 좌측 영상 (b) 우측 영상
그림 7. 실험 영상

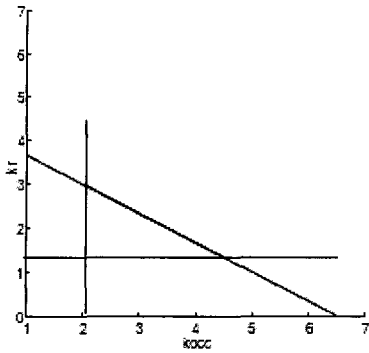


그림 8. 실험영상에서의 k_r 과 k_{occ} 의 범위

그림 7은 실험 영상을 보이고 흰 선은 disparity map에 사용된 스캔라인을 표시한다. 본 논문에서는 제안된 정합 보상과 폐색영역비용 계산 방법을 적용하였다. 그림 8은 실험영상에서 식 (5), (8), (11)으로 둘러싸인 영역을 보여준다.

실험영상에서 기존의 폐색영역비용과 정합보상을 사용하여 구한 변이도 (disparity map) 는 다음과 같다 ($k_{occ} = 25, k_r = 5$).

실험영상에서 본 논문에서 제안한 방식으로 구한 폐색영역비용으로 구한 변이도는 그림 10과 같다 ($k_{occ} = 2, k_r = 4$).

변이는 물체까지의 거리에 반비례한다. 배경은 거리가 멀어서 변이가 작고 물체는 변이가 크다. 그림 9와 그림 10을 비교해 보면, 기존의 k_r 과 k_{occ} 로 구한 변이는 실제의 변이값과 크게 다른 것을 알 수 있다. 그림 10에서는 스탠드의 갯 부분과 받침대 부분이 배경보다 가까운 부분이어서 변이가 크게 나온다. 그러나 그림 9에서보면 스탠드의 갯 부분을 인식하지 못하였고, 오른쪽 배경을 물체로 판단하고 있다.

실험영상에서 k_r 이 계산된 최소값보다 작을 때 변이도를 그려보면 다음과 같다 ($k_r = 2, k_{occ} = 1$). 스탠드의 갯과 받침대 사이의 부분의 배경이 잘못정 합되어 변이가 작게 되었고 이것은 이 부분이 다른 배경보다 더 먼 것처럼 보이게 한다.

실험영상에서 k_{occ} 가 식 (8)의 최소값보다 작을 때 변이도를 그려보면 다음과 같다 ($k_r = 2, k_{occ} = 1$). 이 그림에서는 물체와 배경부분에 오류가 발생한 것을 알 수 있다.

실험영상에서 k_r, k_{occ} 가 식 (11)의 상한보다 클 경우의 변이도를 그려보면 그림 13과 같다.

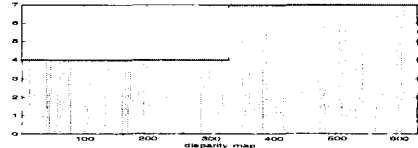


그림 9. 실험영상 1에서 기존의 k_r 과 k_{occ} 을 적용하여 구한 변이도

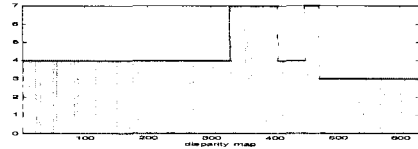


그림 10. 실험영상 1에서 제안한 방식으로 구한 k_r 과 k_{occ} 을 적용하여 구한 변이도



그림 11. 실험영상 1에서 k_r 이 최소값보다 작을 때의 변이도



그림 12. 실험영상 1에서 k_{occ} 가 최소값보다 작을 때의 변이도

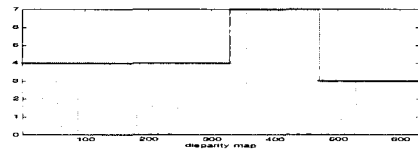


그림 13. 실험영상 1에서 k_r, k_{occ} 가 상한보다 클 때의 변이도

($k_r = 4, k_{occ} = 4$). 이 그림에서 스탠드의 갯과 받침대 사이의 배경을 물체로 잘못 판단한 것을 알 수 있다.

위에서 실험한 결과와 같이 k_r 과 k_{occ} 의 범위 밖의 값을 적용하여 정합을 하는 경우 정합오류가 발생하는 것을 확인할 수 있다.

V. 결론

본 논문에서는 폐색영역을 고려한 스테레오 정합 알고리즘 중에서 비교적 계산량이 적고 변이를 잘

찾을 수 있는 변이공간영상 방법에서 사용하는 비용함수에 대하여 심층적으로 분석하였다. 기존의 논문에서는 비용함수에 사용되는 상수인 폐색영역비용과 정합보상값을 경험적으로 정하여왔으나 본 논문에서는 그 상수가 영상 잡음과 물체와 배경의 밝기 차이에 의존한다는 것을 밝혔다. 이를 사용하면 스테레오 정합의 정확도를 향상시킬 수 있다.

참 고 문 헌

[1] L. Oisel, E. Memin, L. Morin, and F. Galpin, "One-Dimensional Disparity Estimation for Three-Dimensional Reconstruction," *IEEE Trans. on Image Processing*, Vol. 12, No. 9, pp. 1107-1119, Sep. 2003.

[2] 박찬희, 손광훈, "변이-움직임 벡터의 상관관계를 이용한 객체 기반 스테레오 동영상 부호화," *방송공학회논문지*, 제 7권, 제 3호, pp. 238-247, 2002.

[3] 박창주, 고정환, 김은수, "스테레오 영상 보정 알고리즘에 기반한 새로운 중간시점 영상합성 기법," *한국통신학회 논문지*, Vol. 29, No. 5 C, pp. 632-641, May 2004.

[4] 장세훈, 한충신, 배진우, 유지상, "깊이 정보를 이용한 실시간 다시점 스테레오 영상 합성," *한국통신학회 논문지*, Vol. 30, No. 4C, pp. 239-246, Apr. 2005.

[5] M.Z. Brown, D. Burschka, and G.D. Hager, "Advances in Computational Stereo," *IEEE Trans. on PAMI*, Vol. 25, No. 8, pp. 993-1008, August 2003.

[6] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, 2nd Ed., PWS Publishing, pp. 441-507, 1998.

[7] 김철환, 이호근, 하영호, "화소기반 변이공간 영상에서의 스테레오 정합," *한국통신학회 논문지*, Vol. 29, No. 6C, pp. 848-856, Jun. 2004.

[8] C.-J. Tsai and A.K. Katsaggelos, "Dense Disparity Estimation with a Divide and Conquer Disparity Space Image Technique," *IEEE Trans. on Multimedia*, Vol. 1, No. 1, pp. 18-29, Mar. 1999.

[9] S. Birchfield and C. Tomasi, "Depth Discontinuities by Pixel-to-pixel Stereo," *Proceedings of the Sixth IEEE International Conference on Computer Vision*, Mumbai, India, pp. 1073-1080, January 1998.

[10] Stan Birchfield and Carlo Tomasi, "A Pixel Dissimilarity Measure that is Insensitive To Image Sampling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 4, pp. 401-406, Apr. 1998.

[11] R. Szeliski, and D. Scharstein, "Sampling the Disparity Space Image," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 3, pp. 419-425, Mar. 2004.

[12] T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window-Theory and Experiment," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.16, No.9, pp.920-932, Sep. 1994.

박 준 희 (Jun Hee Park) 준회원
 2001년 2월 : 이화여자대학교 정보통신학과 졸업
 2003년 2월 : 이화여자대학교 정보통신학과 석사
 2007년 3월~현재 : 이화여자대학교 전자정보통신공학
 과 박사과정
 <관심분야> 영상인식, 영상처리

이 병 욱 (Byung-Uk Lee) 중심회원
 한국통신학회 논문지 제 22권 6호 참조
 <관심분야> 영상인식, 영상처리