

윈도우 와핑 기법을 이용한 계층적인 스테레오 정합

김재철, 이호근, 권순규, 김희수, 하영호
경북대학교 전자전기공학부

Hierarchical Stereo Matching Using Window Warping Technique

Jae Chul Kim, Ho Keun Lee, Sun Kyu Kwon, Hee Soo Kim, and Yeong Ho Ha
School of Electronic and Electrical Eng., Kyungpook National University
E-mail : yha@ee.knu.ac.kr

요약

본 논문에서는 스테레오 정합시에 발생하는 투영적인 왜곡을 해결하기 위해 윈도우 와핑(warping)을 이용한 새로운 스테레오 정합방법을 제안한다. 투영된 영역의 왜곡은 고정된 윈도우를 사용해서는 변위값을 추정하기 어렵기 때문에 본 논문에서는 블록정합 시에 윈도우 와핑기법을 사용한다. 와핑이 이루어지는 지점은 이전에 구해진 변위값의 신뢰성에 근거해서 구해진다. 초기의 신뢰성과 변위값은 계층적인 정합에서 구해진다. 실험의 결과는 제안한 방법이 정합이 향상되었음을 보여준다.

1. 서론

인간의 시각체계는 서로 다른 위치에서 획득된 두 영상을 적절히 정합함으로써 깊이 정보를 얻는 것으로 알려져 있다[1,2]. 깊이 정보는 카메라 변수와 두 영상의 변위값으로부터 구해진다. 만약 이차원적인 탐색영역이 극상선 제약조건에 의해서 일차원으로 줄어든다면 연산량은 상당히 줄어들게 된다. 이러한 정합지점을 찾는 것을 스테레오 정합이라고 한다. 그림1에서 각 변수(L, f, dl, dr)는 스테레오 카메라 사이의 거리와 초점거리 그리고 좌·우 영상의 중심으로부터의 거리를 가리킨다.

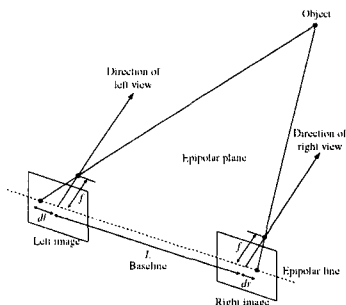


그림 1. 평행축을 가진 스테레오 시각의 기하학적 구조

스테레오 영상은 다른 두 개의 시각체계에 얻어지며, 각각의 영상은 그림2에서처럼 투영적인 왜곡을 갖는다.

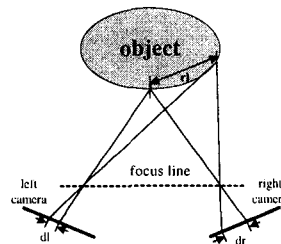


그림 2. 투영적 왜곡

만약 물체의 표면(rl)이 좌·우 각각의 카메라에 투영된다면 각 영상은 다른 시각적인 특성을 가지게 될 것이다. 최근에 이러한 투영적인 왜곡현상에 대한 연구가 이루어지고 있다. Kanade는 이러한 왜곡현상을 줄이기 위해서 적응적 윈도우 방법을 제안하였다[3]. 이 방법에서는 윈도우 안의 변위값에 대한 통계적인 방법을 이용하는데, 변위값의 국부적인 분산을 구해서 각 윈도우의 형태를 결정하게 된다. Wang과 Ohishi는 이러한 문제를 해결하기 위해서 3D-to-3D 방법을 제안하였다. 이 방법에서는 통계적인 방법보다는 기하학적인 방법을 이용하였다[4]. 하지만 이제까지의 방법들 중에서 윈도우 와핑방법을 이용하여 투영적 왜곡현상에 접근하지는 않았다. 그리고 이러한 왜곡현상은 고정된 윈도우를 사용해서는 해결하는 것은 상당히 어렵다. 본 논문은 이러한 왜곡현상을 해결하기 위해서 윈도우 와핑방법에 근거한 새로운 스테레오 정합을 제안한다.

2. Window Warping algorithm

2.1 투영적 왜곡

스테레오 시각체계에 있어서 두 카메라에 투영된 영상은 카메라의 광학축에 대하여 기울어진다. 이것은 그

림3에서 보이는 것처럼 같은 영역에 대해서 좌·우 영상이 다른 형태를 가지게 되는 투영적인 왜곡현상이 일어난다. 그림3은 좌·우 각 영상의 밝기 종단선을 보여준다.

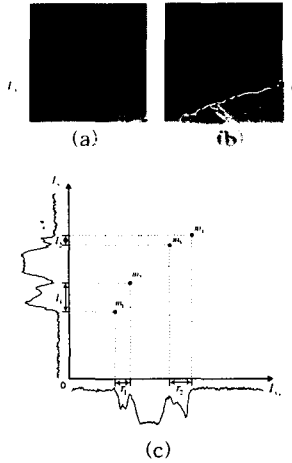


그림 3. 'man'의 원영상과 투영적인 왜곡현상결과. (a) 좌영상의 종단선. (b)우영상의 종단선. (c) 각 종단선의 비교

그림 3은 l_1 의 길이가 r_1 의 길이와 같지 않음을 보여주며 l_2 와 r_2 에서도 이와 같다. 따라서 고정크기의 블록 정합 알고리즘을 이용해서는 이러한 영역에 대해서 올바른 정합점을 찾지 못한다.

2.2 Fant's 와핑 알고리즘

계층적인 정합을 위해서는 영상을 서브 샘플링해서 구성해야 하는데 Fant's 알고리즘을 사용하였다.

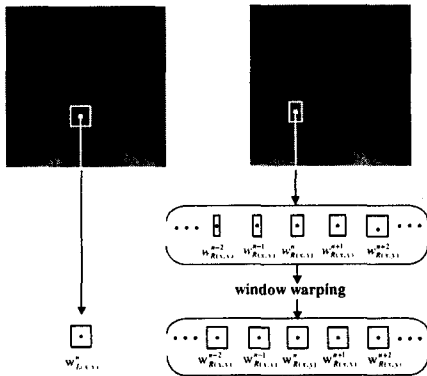


그림 4. Fant's 알고리즘을 이용한 윈도우 와핑

Fant는 입력과 출력을 공간적인 매핑에 따라 결정되는 비율에 의해서 생성과 소모가 이루어지는 열로 간주한다.

윈도우 와핑된 후보 윈도우중에서 최적의 윈도우의 선택은 식1과 같이 정의 된다.

$$w_p = \arg \min_{k \in \Omega} |w_{L(x,y)}^n - w_{R(x,y)}^{n+k}| \quad (1)$$

식1에서 Ω 는 윈도우 와핑 크기를 나타내고 w 는 블록 정합의 국부 윈도우를 카리킨다.

3. 적응적 윈도우 와핑을 이용한 스테레오 정합

그림 5는 제안한 알고리즘의 블록도 이다. 제안한 방법은 네 단계로 이루어져 있다. 첫 번째 단계에서는 계층적인 정합의 최상위 계층의 영상으로부터 블록정합을 실시한다. 이번 단계의 결과는 다음 단계의 결과를 얻는 곳에 사용이 된다. 두 번째 단계에서는 첫 번째 단계에서 얻은 결과로부터 신뢰성 정보를 획득하는데, 세부적으로는 양 방향 검사와 국부 변위 분산을 구하는 과정이 포함되어 있다. 세 번째 단계에서는 신뢰성에 근거한 계층적인 정합을 실시하게 된다. 이러한 계층적인 정합은 정합이 국부 minima에 빠지는 것을 피하고 밝기값의 영향을 줄이는데 효과가 있다. 신뢰성이 높은 영역에서는 고정된 크기의 윈도우를 사용하고 신뢰성이 낮은 영역에서는 와핑 윈도우를 이용한 정합을 실시한다. Fant의 와핑방법을 사용하여 저역 필터링과 서브 샘플링의 과정을 거치는 기존의 방법을 사용하는 것보다 연산량과 수행시간이 많이 줄어들게 된다.

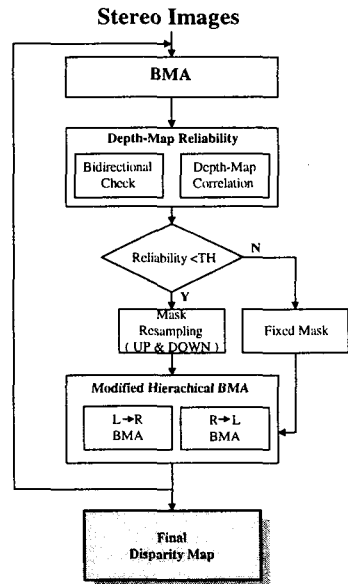


그림 5. 제안한 스테레오 정합과정

3.1 블록 정합 알고리즘

첫 번째 수행과정은 세 번째 과정에서 수행될 계층과 같은 크기의 영상을 이용해서 블록 정합을 실시한다. 이번 단계의 결과는 두 번째 단계의 입력값으로 이용이 되고 양방향으로 정합이 이루어진다.

이번 단계에서 사용된 거리측정은 MAD(mean absolute distance)를 사용하고 변위값은 좌·우 영상에서 MAD의 값을 계산해서 최소의 값을 가지는 점을 정합점으로 선택한다.

$$D(x, y, d) = \arg \min_{d \in \{1, \dots, W\}} \left[\frac{1}{mn} \sum_{x'=-\frac{w}{2}}^{\frac{w}{2}} \sum_{y'=-\frac{w}{2}}^{\frac{w}{2}} |I_R(x+x', y+y') - I_L(x+d+x', y+y')| \right] \quad (2)$$

d 는 우영상에서의 이동을 나타내고 m 과 n 은 윈도우의 크기를 나타낸다. 그리고 I_L 과 I_R 은 각각 좌·우영상의 밝기값을 가리킨다. 극상선 제약조건을 가정했으므로 2차원적인 탐색을 1차원의 문제로 전환하여 연산량이 감소된다. 최적의 d 값은 MAD값을 최소화 시키는 값이다.

3.2 신뢰성 생성

정합지점의 신뢰성을 추정하기 위해서 유일성 조건과 평탄 조건을 이용하였다. 따라서 이러한 제약성을 위반하는 변위값에서는 신뢰성이 낮아지게 되고 충족하는 지점에서는 신뢰성이 높아지게 된다.

3.2.1 유일성 조건

기준 영상에서의 한 점은 다른 영상에서 두 개의 지점에 정합이 될 수가 없다. 하지만 양방향 검사는 각 변위값들은 그림6에서 보이는 것처럼 작은 차이를 보인다. 그리고 양방향 검사의 정의는 다음과 같다.

$$\delta = |D_L(x, y) - D_L(x + D_R(x, y), y)|. \quad (3)$$

유일성 제약조건은 각 표본 지점(x, y)에서 이루어지고 일탈(deviation) δ 는 이러한 조건의 변위의 척도에 이용된다.



그림 6. 양방향 검사

입력 영상에 대한 정합의 결과가 어느 정도 유일성 제약조건을 만족하는가를 정의하기 위해서 다음의 신뢰성 함수를 정의한다.

$$f_{R1} = \frac{T_{bc} - \delta}{T_{bc}}, \quad \begin{cases} T_{bc} : \text{threshold} \\ \delta : \text{bidirectional check distance} \end{cases} \quad (4)$$

f_{R1} 은 유일성 신뢰도를 나타내고 δ 는 양방향 검사의 오류를 나타내고 T_{bc} 는 최소 임계치를 가리킨다.

3.2.2 평탄 조건

각 정합점은 유사한 밝기값을 가지거나 윈도우 내의 밝기값은 상당한 상관성을 가진다. 이러한 특성을 고려해서 유일성 제약조건을 만족하는 값들에 대해서 분산을 구하게 된다. 그리고 그 값들은 가중함수로부터 추출되어진다.

$$w_s(x, y) = \begin{cases} 0 & , \delta > T_{bc} \\ 1 & , \delta < T_{bc} \end{cases}. \quad (5)$$

다음으로 추출된 변위 지도에서의 분산은 다음과 같이 정의된다.

$$\sigma_D^2(x, y) = \left[\frac{1}{N_V} \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} \{w_s(i+x, j+y) \cdot I(i+x, j+y)\}^2 \right] - \left[\frac{1}{N_V} \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} w_s(i+x, j+y) \cdot I(i+x, j+y) \right]^2 \quad (6)$$

그리고 N_V 는 다음과 같이 정의된다.

$$N_V = \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} w_s(i+x, j+y).$$

위와 같은 관계로부터 평탄 신뢰성을 다음과 같이 정의한다.

$$f_{R2} = e^{-\xi \sigma_D^2}, \quad (7)$$

ξ 는 f_{R2} 의 감쇄속도를 줄이기 위한 감쇄계수이다. 만약 σ_D^2 의 값이 작아지면 f_{R2} 의 값은 커지게 된다.

3.2.3 신뢰도 함수

신뢰도 함수는 f_{R1} 과 f_{R2} 의 선형조합으로 구성이 된다.

$$f_{total} = \lambda_1 f_{R1} + \lambda_2 f_{R2}, \quad (8)$$

λ_1 과 λ_2 는 $\lambda_1 + \lambda_2 = 1$ 의 관계를 만족시키는 가중치 계수이다.

3.3 계층적 블록정합 알고리즘

잡음에 대한 민감도를 줄이고 시간적 효율성을 증대시키기 위하여 계층적인 정합을 실시한다. 그림 7에 나타난 것처럼 본 논문에서는 기존의 계층적인 정합과는 다른 계층적인 정합과정이다. 더 자세하게 나타나 있다.

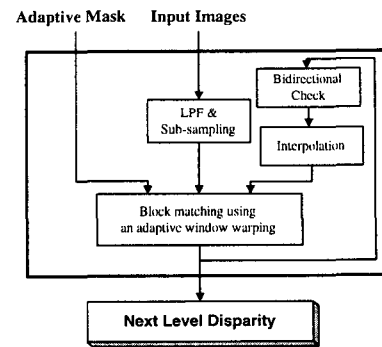


그림 7. 계층적 정합 블록도

Step 1 : 저역통과 필터링과 서브 샘플링

: 입력영상이 저역통과하고 서브 샘플된다.

Step 2 : 블록 정합 과정

: 스테레오 정합이 이루어진다.

Step 3 : 양방향 검사

: Step 2에서 구한 변위값에 대해서 양방향
: 검사를 실시한다.

Step 4 : 보간

: Step 3의 조건을 만족시키는 변위값에 대
: 해서 선형 보간을 한다.

국부 영역에서의 평탄조건을 고려하기 위해서 상위 계층의 변위벡터 9개와 현재의 변위벡터 4개를 이용한다. 이러한 계층적인 정합은 MAD와 상위계층의 평탄성을 함께 고려한 새로운 비용함수로 다음과 같이 정의할 수 있다.

$$D^i(x, y, d) = \arg \min_{d \in \mathbb{Z}} \left[\frac{1}{mn} \sum_{x'=x-\frac{m}{2}}^{x+\frac{m}{2}} \sum_{y'=y-\frac{n}{2}}^{y+\frac{n}{2}} |I'_k(x+x', y+y') - I'_l(x+d+x', y+y')| \right] \quad (9)$$

$$S^i(x, y, d) = \frac{1}{9} \sum_{y'=-1}^1 \sum_{x'=-1}^1 D^{i-1}(x', y') - D^i, \quad (10)$$

식 (9)와 식(10)의 결과로부터 식(11)이 유도된다.

$$E^i = \gamma_1 D^i + \gamma_2 S^i, \quad (11)$$

4. 실험 및 결과

실험에서 사용된 'man' 스테레오 영상에 대한 변수는 표1에 나타난 것과 같다.

표 1. 실험에서 사용된 변수값

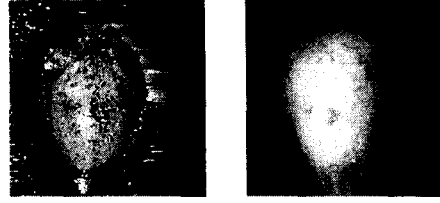
Window Size(w)	T_{bc}	λ_1	λ_2	γ_1	γ_2
7x5	3	0.5	0.5	0.4	0.6

실험의 결과는 본 논문이 제안한 방법이 질감정보가 적고 투영적 왜곡이 일어나는 영역에서 기존의 방법과 비교해서 좋은 결과를 나타낸다.



(a)

(b)



(c)

(d)

그림8. 'man'의 원영상과 제안한 방법에 의한 변위도. (a)좌영상. (b)우영상. (c)BMA의 변위도. (d)제안한 방법에 의한 변위도.

5. 결론

본 논문에서는 적응적 윈도우 와핑을 이용한 새로운 계층적 스테레오 정합을 제안하였다. 제안한 알고리즘은 변위값에 대한 확률적 모델에 근거해서 적응적으로 윈도우를 선택해서 정합의 신뢰성을 갖는다. 실험의 결과는 기존의 정합 방법에 비해서 투영적 왜곡이 일어나는 영역에 대해서 제안한 알고리즘이 정확한 정합을 수행함을 보여주었다.

참고문헌

- [1] M. W. Matlin, *Sensation and perception*, pp. 184-229. Allyn and Bacon, 1988.
- [2] D. Marr and T. Poggio, "Computational Theory of Human Stereo Vision," *Proc. of Roy. Soc. London B*, vol. 204, pp. 301-308, 1979.
- [3] Kanade and Okutomi, "A Stereo Matching Algorithm with an Adaptive Window," *IEEE Trans. on Patt. Anal. Machine Intell.*, vol. 16, no. 9, pp. 920-932, Sep. 1994.
- [4] Z. F. Wang and N. Ohnishi, "Intensity-based Stereo Vision: from 3-D to 3-D," *SPIE*, vol. 2354, pp. 434-443, Nov. 1994.
- [5] P. N. Belhumeur and D. Mumford, "A Bayesian Treatment of the Stereo Correspondence Problem Using Half-occlusion Regions," *IEEE Conf. on Computer Vision and Pattern Recognition*, pp.506-512, 1992.