

확률기반의 윈도우 왜곡을 이용한 스테레오 정합

정회원 김재철*, 이호근**, 하영호**

Stereo Matching Using Probability-Based Window Warping

Jae-Chul Kim**, Ho-Keun Lee**, and Yeong-Ho Ha** *Regular Members*

요약

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

ABSTRACT

This paper proposes a new stereo matching algorithm using an adaptive window warping. Since the projective distortion regions can't be estimated with any fixed-size block matching algorithm, the window warping technique is used in block matching process. The position of a reference window to warping is adaptively obtained according to the degree of reliability calculated by estimates disparities in the previously. The initial disparity and reliability are obtained by applying a well known hierarchical strategy. The experimental results show that considerable improvements are obtained especially in projective distortion region.

1. 서론

인간의 시각 체계는 서로 다른 위치에서 획득된 두 영상을 적절히 정합함으로써 거리 정보를 얻는 것으로 알려져 왔다^[1,2]. 스테레오 정합은 인간 시각 체계의 거리 추출 능력을 자동화하기 위한 컴퓨터 시각 분야중 하나이다. 거리 정보를 얻기 위한 기본 단계는 영상 획득, 특징 추출, 스테레오 정합, 변위 추정, 변위로 부터의 거리계산 등으로 이루어지는데 여기에서 가장 중요한 인자는 특징으로 사용될 정합 요소의 선택과 그 요소에 적절한 정합 전략을 구하는 것이라 할 수 있다.

스테레오 정합은 정합 요소에 따라 크게 영역 기반(area-based)법과 특징 기반(feature-based)법으로 나눌 수 있다. 특징 기반에 사용되는 정합 요소

(matching primitive)는 영 교차점^[3], 경계선^[4], 모서리, 마루, 골, 원추 곡선^[5] 등이 있다. 이들은 정합 점이 정확하고 잡음에 강한 특징을 가지나, 정합되어지는 점들이 적으므로 전체 영상의 변위값을 구하기 위해서는 폐색 모델링(occlusion modeling)과 변위연속성(disparity continuity) 등을 포함하는 어려운 내삽(interpolation) 과정을 필요로 하게 된다. 반면에, 영역 기반법에서의 정합 요소는 밝기 정보의 변화가 평탄하거나 동일한 영역의 모양, 평균 밝기 및 면적 등을 이용한다.^[6] 이 정합 요소는 밝기 정보에 많이 의존하기 때문에 잡음에 약한 면이 있지만 영상의 전체 거리 정보를 얻을 수 있다는 장점이 있다. 그리고 정합과정이 전 영상에 대해서 이루어지기 때문에 영상 내에서의 특징정보를 모두 이용할 수 있으므로 특징 기반의 정합과정에서 발

* 한국전자통신연구원 컴퓨터소프트웨어연구소 (e-mail : kimjc@etri.re.kr)

** 경북대학교 전자전기공학부 (e-mail : yha@ee.knu.ac.kr)

논문번호: 00493-1229, 접수일자: 2000년 12월 29일

생하는 문제점을 극복하게 된다. 최근에는 색정보를 이용하는 정합법,^[7] 여러장의 스테레오 영상을 이용하는 정합법,^[8] 위상을 이용하는 정합법,^[9] 그리고 여러 가지 정합요소를 이용하는 복합적인 정합법^[10] 등이 제안되고 있다. 하지만 스테레오 영상의 기하학적인 특성으로 기인한 좌·우 영상에서의 투영적인 왜곡(projective distortion)현상은 극복하기가 어렵다. 이러한 현상은 그림 1에서 보여진다.

만약 물체의 표면(π)이 좌·우 각각의 카메라에 투영된다면 각 영상은 다른 시각적인 특성을 가지게 되고 정합을 수행하는 도중에 많은 오정합을 일으키게 된다. 최근에 이러한 투영적인 왜곡현상에 대한 연구가 이루어지고 있다^[11-12]. 이러한 문제를 해결하기 위해서 가변 윈도우를 이용한 방법과 확률에 기반한 이완처리 방법^[13]이 이용되었다. 하지만 기존의 방법들은 고정된 윈도우를 사용함으로써 근본적인 문제를 해결하지는 못했으며 결과도 만족스럽지 못하다.

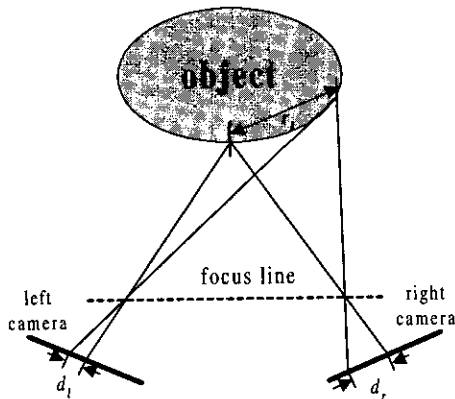


그림 1. 투영적 왜곡

본 논문은 이러한 왜곡현상을 해결하기 위해서 윈도우 와핑 방법에 근거한 새로운 스테레오 정합을 제안하였다. 본 논문은 이러한 투영적 왜곡현상이 현저하게 드러나는 근거리 실물 영상에 대해서 실험을 하였고, 좌·우영상의 정합 정도를 높이기 위해 영상내에서의 정합 신뢰성을 생성하였으며, 정합이 필요한 실물 영역내에서만 정합을 하기 위해서 객체를 추출하는 과정을 거치게 된다. 또한 영역 기반의 방법이 가질 수 있는 잡음의 영향을 줄이기 위해서 정합의 전 과정은 계층적인 과정을 가진다. 제안한 스테레오 정합 방법은 투영적으로 왜곡이 일어난 영역에 대해서 적응적인 윈도우를 사용함으

로 인해서 정합의 정확성을 높였으며, 정합의 신뢰도에 근거한 윈도우 와핑^[14]으로 인해 연산량이 감소하였다. 특히 기존의 정합 과정에서 필요했던 이완처리 과정이 없음에도 불구하고 오정합이 감소되었음을 알 수 있었다.

II. 기존의 블록 정합 기반의 스테레오 정합

스테레오 영상은 두 개의 다른 카메라로부터 동일물체를 다른 각도에서 얻은 영상을 말하며, 그림 2는 스테레오 카메라의 기하학적인 구조를 나타낸다. 각 변수(L, f, d_l, d_r)는 스테레오 카메라 사이의 거리, 초점거리, 그리고 좌·우 영상의 중심으로부터의 거리를 가리킨다. 그리고 이차원적인 탐색영역이 극상선 제약조건(epipolar constraint)에 의해서 일차원으로 줄어든다면 연산량은 상당히 줄어들게 된다.^[15-16] 실험에 사용된 영상은 이러한 조건을 모두 만족하는 영상이므로 탐색영역은 수평방향으로 제한되었다.

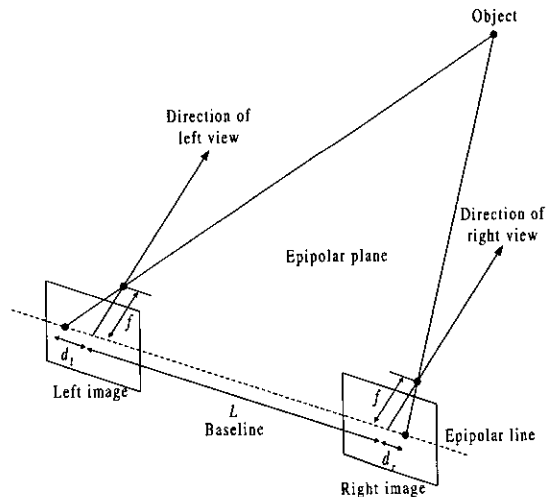


그림 2. 평행축을 가진 스테레오 시각의 기하학적 구조

2.1 블록 정합 알고리즘

블록 정합알고리즘은 기준 영상에서 블록내에서의 밝기값과 정합의 대상이 되는 영상간의 상관도를 구함으로써 정합을 수행하게 된다. 먼저 기준영상에서 블록내의 밝기 값을 추출하고 정합 대상이 되는 영상에서 모든 지점에 대해서 cross-correlation을 구해서 가장 상관도가 높은 지점을 정합점으로 취하게된다. 하지만 이러한 방법은 투영적으로 왜곡이 일어난 영역이나 조명의 변화에 따른 밝기값의 변화에 민감하여 오정합을 일으키기 쉽게 된다.

2.2 반복학습에 의한 이완처리 정합 알고리즘

반복학습에 의한 이완처리에 기반한 방법으로 신경회로망과 최근 유전자 알고리즘 및 개체군기반 증강학습을 이용한 방법, 그리고 이들 모두를 이용하는 방법이 사용되고 있으며 이완 처리를 동시에 수행하기 때문에 실영상이 아닌 인공 영상에 대해서는 좋은 결과를 가지지만 실 영상에서는 만족할 만한 결과를 얻지 못한다. 왜냐하면 정합의 초기 모델이 스테레오 영상의 특성 정보를 고려하지 않기 때문에 투영적으로 왜곡이 일어난 영역에서는 학습과 이완처리 과정을 거친 후에도 오정합의 결과를 계속 유지되게 된다.

이러한 기존의 방법들은 왜곡현상으로 인한 오정합을 줄이기 위해서 정합의 기준이 되는 영상만의 신호적 특성을 이용하여 탐색 윈도우의 크기를 다르게 하여 정합을 실시하거나 초기 정합의 결과 중에서 정합의 제약 조건에 가장 근접한 값을 선택하는 방법들이 존재했지만 반복 수행으로 인한 정합시간의 증가와 더불어 오정합의 결과가 넓은 영역에서 발생할 경우 오정합의 결과가 전파가 되는 결과를 얻게 된다. 하지만 본 논문이 제시한 방법에서는 영상의 변형 특성을 극복하기 위해서 정합의 기준 영상 정합 윈도우를 와핑시켜서 정합을 실시한다.

III. 윈도우 와핑

3.1 투영적 왜곡

스테레오 시각체계에 있어서 두 카메라에 투영된 영상은 카메라의 광학축에 대하여 기울어진다. 이것은 그림 3에서 보이는 것처럼 같은 영역에 대해서 좌·우 영상이 다른 형태를 가지게 되는 투영적인 왜곡현상이 일어난다.

만약 그림 3에 나타난 것처럼 3차원 공간상에 $L \times L$ 크기의 정사각형 모양이 존재한다고 가정했을 때 각 스테레오 영상은 Base line에 대해서 다른 각도로부터 영상이 얻어진다. P_1 과 P_2 를 각 영상에서의 중심점이라고 했을 때 L_{x1} 과 L_{x2} 와의 투영적 왜곡은 다음과 같은 식(1)으로 유도될 수 있다.^[15]

$$\frac{L_r}{L_l} = \frac{k_r \cdot f_r \cdot (\cos(\alpha - \beta) + k_l \cdot \sin(\alpha - \beta))}{k_l \cdot f_l (L \cdot \cos \alpha - k_l \cdot P_l \cdot \sin \alpha)} \cdot \frac{L \cdot (1 + k_r \cdot \sin \alpha)}{(1 + k_l \cdot \sin(\alpha - \beta))} \quad (1)$$

$$k_r = \frac{L}{f_r + Z_r}, \quad k_l = \frac{L}{f_l + Z_l}, \quad k_l = \frac{1 - \tan \beta \cdot \tan \gamma}{\tan \beta + \tan \gamma} \quad (2)$$

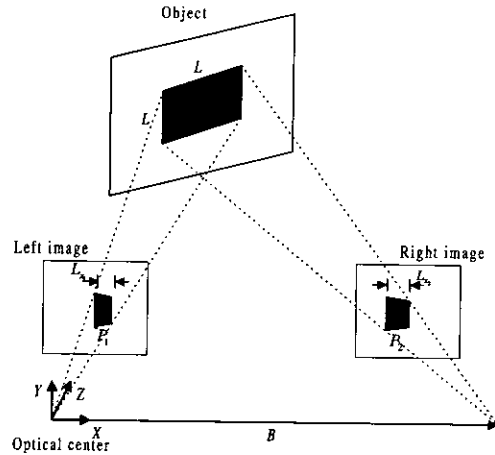


그림 3. 평면 정사각형의 투영적 왜곡현상

β 는 base line과 카메라 사이의 회전각을 의미한다. f_l 과 f_r 은 초점거리를 의미하고 α 는 3차원 평면과 기준 카메라와의 평면과의 각을 의미한다. 그리고 α 는 물체가 카메라 축에 대해서 움직인 회전각을 의미한다. Z_l 와 Z_r 는 각 카메라와 3차원 평면과의 거리를 의미한다. 부록 A은 위 식의 유도과정을 상세히 설명해 준다. 식 (1)에서 알 수 있듯이 각 α 가 커질수록 왜곡 현상은 증가하며 이러한 현상은 실제 3차원 물체가 정면에 대해서 기울어지거나 곡면을 가질 때 같은 현상을 가져온다. 기존의 방법들이 이러한 왜곡현상을 고려하지 않고 정합을 실시함으로써 해서 오정합이 증가되었다.

3.2 윈도우 와핑 알고리즘

그림 4는 기준 영상의 윈도우가 정합을 수행하는 과정에서 대상이 되는 영상의 윈도우와 정합하기 위한 윈도우 와핑 과정을 설명하고 있다. 그림에서 보이는 것과 같이 투영적인 왜곡이 일어난 영역에서는 윈도우를 와핑시킨 후보 윈도우와 상관도를 구해서 정합점을 찾게 된다. 여기에서 신뢰성이 낮은 영역을 찾는 문제는 다음 장에서 다룬다.

기존의 영상처리 기법에서 사용된 재표본화 기법은 여러 단계를 거치기 때문에 정합후보점이 많은 경우에 연산량이 상당히 증가하게 된다. 제안한 알고리즘에서는 연산량을 감소시키기 위해서 하드웨어적으로도 구현이 간단한 Fant 알고리즘을 이용하였다. Fant 알고리즘은 입력과 출력의 공간적인 사상에 따라 결정되는 비율에 의해서 생성과 소모가 이루어지는 열로 간주한다. 따라서 단단계 수행과정이 필요없어서 연산량이 감소하고 수행속도가 증가한다.

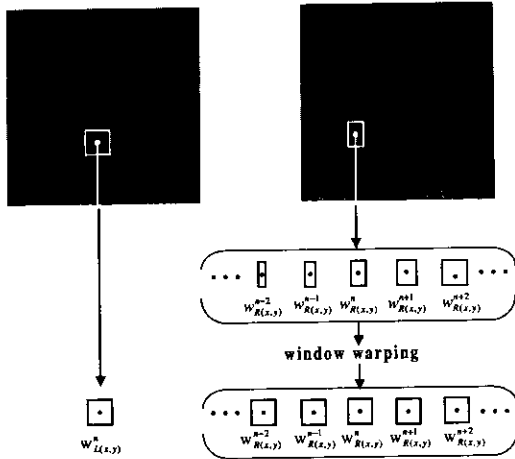


그림 4. Fant's 알고리즘을 이용한 윈도우 와핑.

정합과정을 수행하기 위한 전처리과정(외핑과정)은 세 단계를 가진다. 첫째, 기준 윈도우를 왼쪽 기준 영상에서 추출한다. 둘째, 다른 크기의 표본들을 우측 영상에서 추출한다. 셋째, 추출된 표본들을 기준영상에서 추출된 윈도우와 정합을 수행하기 위해서 외핑을 시킨다. 외핑된 후보 윈도우중에서 최적의 윈도우의 선택은 식(3)과 같이 정의 된다.

$$w_{\phi} = \arg \min_{k \in \Omega} |w_{L(x,y)}^n - w_{R(x,y)}^{n+k}| \quad (3)$$

식(3)에서 Ω 는 윈도우 와핑 크기를 나타내고 $w_{L,R(x,y)}^n$ 는 블록 정합의 국부 윈도우를 가리키며 n 은 후보 윈도우를 의미한다. 이렇게 선택된 윈도우를 정합시에 사용할 경우에는 정합영역의 정합신뢰성을 근거로해서 선택적으로 정합을 실시한다.

정합의 신뢰성이 낮은 영역에서는 표1에서와 같이 윈도우를 와핑시켜 정합을 실시하는데 기존의 정합방법을 사용했을 때보다 정합오차가 줄어들었음을 확인할 수 있다. 기존의 가변 윈도우기법은 정합영역 정보의 감소 또는 증가로 인해서 정합오차가 제안한 방법에 비해서 상대적으로 증가하며, 이러한 현상은 근거리에서 얻은 스테레오 영상일수록 증가하며 그로 인한 오차도 증가하게 된다. 이러한 결과는 표1에 나타난 결과와 같이 영역의 변화가 심해서 정합의 신뢰성이 떨어지는 지점에서 그 차이가 크다.

IV. 제안한 적응적 윈도우 와핑을 이용한 계층적인 스테레오 정합

제안한 방법은 네 단계로 이루어져 있으며 전처리 단계로 첫 번째 단계인 객체추출의 단계를 가진다. 두 번째 단계부터 마지막단계는 변형된 계층적인 정합의 모든 과정을 포함하고 있다. 그림 5는 제안한 알고리즘의 블록도를 나타낸다. 두 번째 단계에서는 계층적인 정합의 최상위 계층의 영상으로부터 블록정합을 실시한다. 이 단계의 결과는 다음 단계의 결과를 얻는 곳에 사용이 된다. 본 논문에서는 모두 세 개의 계층을 정합에 사용하였으며, 기존의 방법이 오정합의 결과를 하위계층으로 전파를 시킬 수 있으므로 새로운 계층적인 정합법이 제안되었다. 제안된 방법의 정합 방법은 양방향 정합을 실시하여 신뢰성이 작은 정합점은 제거하고 나머지 부분에 대해서 내삽을 실시하여 정합도를 구함으로써 하위로 값을 전파시킨다. 이러한 방법은 계층적인 정합의 신뢰성을 더욱 증가시킨다. 계층적인 정합의 세 번째 단계에서는 두 번째 단계에서 얻은 결과로부터 신뢰성 정보를 획득하게 되는데, 세부적으로는 양 방향 검사와 국부 변위 분산을 구하는 과정이 포함되어 있다. 정합도에서 정합의 결과에 대한 신

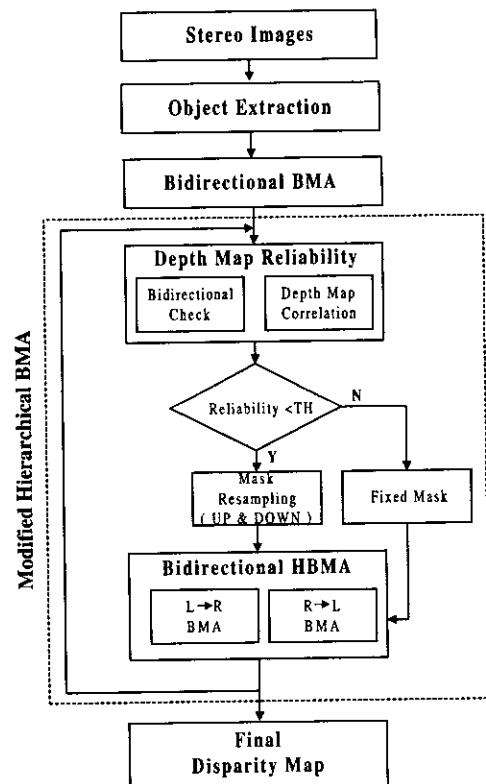


그림 5. 제안한 스테레오 정합과정

되도의 확률값을 이 과정에서 구하게 되며 마지막 단계에서 사용이 된다. 세 번째 단계에서 사용된 두 가지 방법은 스테레오 정합의 제약 조건이 사용되며, 이를 위반한 정합값은 오정합일 가능성이 상당히 크게 되므로 적응적 윈도우 왜곡의 대상이 된다. 마지막 단계에서는 신뢰성에 근거한 정합을 실시하게 되는데 신뢰성이 높은 영역에서는 고정된 크기의 윈도우를 사용하고 신뢰성이 낮은 영역에서는 왜곡 윈도우를 이용한 정합을 실시한다.

4.1 객체 추출 과정

스테레오 영상의 객체는 정합의 대상이 아닌 배경과 정합의 대상이 되는 객체로 구분될 수 있으며 배경 부분을 제거하여 정합의 대상에서 제외시킨다면 정합의 속도뿐만 아니라 정합의 경계에서 정확한 값을 가지게 된다. 따라서 본 논문에서는 정합의 전처리단계로 객체 추출의 단계를 가지며 모두 여섯 단계를 가진다.

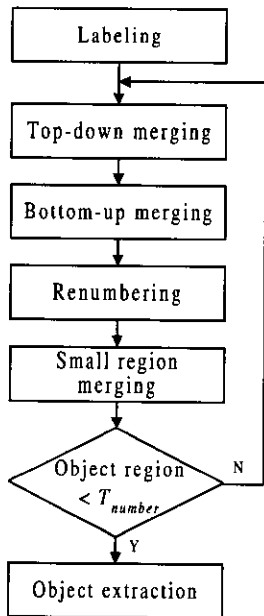


그림 6. 객체추출 과정

첫 번째 단계에서는 모든 점들에 대한 레이블링을 실시한다. 두 번째 단계에서는 top-down 순서로 밝기값의 임계치 차이내에 있는 값들을 하나의 영역으로 병합을 하게 된다. 다음 단계로는 이렇게 분류된 영역을 다시 bottom-up 순서로 다시 병합을 한다. 다음 단계로는 이렇게 병합이 된 영역을 다시 레이블링을 하고 임계영역의 수 안에 들어 올 때까지 반복

수행의 과정을 실시한다. 이렇게 merging이 된 영역에 대해서 다시 영역의 크기가 크지 않은 영역은 밝기 값이 가장 가까운 영역으로 merging이 된다. 다음 단계에서는 분류가 된 영역의 수가 임계치를 넘지 않을 때까지 반복 수행을 한다. 반복 수행과정을 마치면 마지막 단계에서 객체를 얻게 된다.

4.2 계층적 블록정합 알고리즘

스테레오 영상의 정합의 과정에서 정합의 오류를 일으키는 요인 중에서 주요한 요인인 잡음과 광원의 영향을 줄이기 위해서 본 논문은 계층적인 구조로 이루어져 있다. 각 계층에서 사용되는 블록의 크기를 설정할 때 계층별로 다른 크기의 블록을 이용하는 방법과 같은 크기의 블록을 사용하는 방법이 있을 수 있다. 두 방법에 대한 비교를 보면 계층별로 같은 크기의 블록을 사용하는 것이 더 나은 결과를 보여주었다.¹⁷⁾ 일반적인 계층적인 구조는 상위계층의 오차가 하위 계층으로 전파가 쉽게 이루어지기 때문에 본 논문에서는 새로운 계층적인 정합법을 제안하며 그림 7에 자세하게 나타나 있다. 실험에서는 모두 세 개의 계층을 사용하여 실험을 하였고 더 많은 계층을 사용하더라도 더 좋은 결과는 얻지 못했다.

입력된 영상은 먼저 저역 통과과정을 거친 다음 성긴 표본화를 한 영상을 정합하였다. 최상위 계층에서의 오정합 결과가 하위 계층으로 전파되는 것을 최소화 시키기 위해서 양방향 검사를 통해서 일치하는 영역만 남긴 후에 내삽과정을 거쳐 하위 계층으로 변위 값을 전파시킨다. 이러한 값들은 다음 계층으로 전파가 되고 다음 단계에서는 확률에 기반한 신뢰성 생성과정을 거쳐 구한 확률값으로부터 정합을 실시한다. 정합 영역에 대한 확률값은 다음 장에 상세히 기술되어 있다.

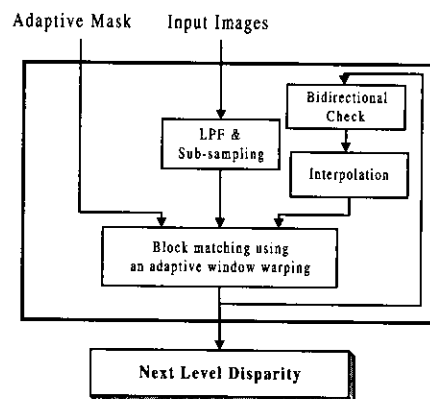


그림 7. 계층적 정합 블록도

4.2.1 블록 정합 알고리즘

첫 번째 수행과정은 세 번째 과정에서 수행될 계층과 같은 크기의 영상을 이용해서 블록 정합을 실시한다. 이 단계의 결과는 두 번째 단계의 입력값으로 사용 되고 양방향으로 정합이 이루어진다. 거리 측정 방법에는 여러 가지가 있을 수 있지만 본 논문에서는 MAD(mean absolute distance)를 사용하고 변위값은 좌·우 영상에서 MAD의 값을 계산해서 최소의 값을 가지는 점을 정합점으로 선택한다.

$$D(x, y, d) = \arg \min_{d(x, y)} \left[\frac{1}{mn} \sum_{y'=-\frac{n}{2}}^{\frac{n}{2}} \sum_{x'=-\frac{m}{2}}^{\frac{m}{2}} |I_R(x+x', y+y') - I_L(x+d+x', y+y')| \right] \quad (4)$$

여기서 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.

4.3 신뢰성 생성

스테레오 영상에서의 정합 결과는 실영상의 경우에 평탄한 특성과 하나의 물체에서 기인한 정합점의 유일성을 내포하고 있다. 따라서 본 논문에서는 이러한 특성으로부터 신뢰성값을 확률에 기반한 값으로 변환하여 윈도우 좌평 영역을 적응적으로 선택하게 된다. 그리고 제약성을 위반하는 변위값에서는 신뢰성이 낮아지게 되고 충족하는 지점에서는 신뢰성이 높아지게 된다. 이러한 신뢰성은 스테레오 영상의 배경영역과 투영적 왜곡이 일어나 영역에서는 신뢰성이 매우 떨어지게 된다. 스테레오 정합의 제약 조건은 여러 가지가 제기되고 있지만 실제 정합에서 가장 주요한 결정력을 가진 유일성과 평탄 조건을 사용함으로써 해서 논문이 제안하고 있는 확률값을 생성하게된다.

4.3.1 유일성 조건

기준 영상에서의 한 점은 다른 영상에서 두 개의 지점에 정합이 될 수가 없다 이러한 특성은 그림 8에서 보여지며 기준영상을 좌영상과 우영상으로 두어서 만든 두 개의 변위도를 가지고 제약 조건을 검사하게 된다. 하지만 양방향 검사에서 각 변위값들은 그림 6에서 보이는 것처럼 작은 차이를 보인다. 이상적인 경우에는 이러한 값들은 영이 되어야 하지만 얻어진 영상이 질감정보나 윤곽선 정보를

가지지 않을 경우에 올바른 정합점을 찾지 못해서 유일성의 조건을 만족시키지 못하게 된다. 이러한 유일성의 조건을 검사하기 위해서는 기준영상을 좌측과 우측일 때를 모두 고려하여야 하며 각 변위값의 차에 대해서 절대치를 구한 값이 오차가된다. 이로부터 첫 번째 확률값을 구하게된다. 식(5)는 양방향 검사의 정의를 나타낸다.

$$\delta = |D_L(x, y) - D_L(x + D_R(x, y), y)| \quad (5)$$

이러한 유일성 제약조건은 각 표본 지점(x, y)에서 이루어지고 편차(deviation) δ 는 이러한 조건의 만족척도에 이용된다.

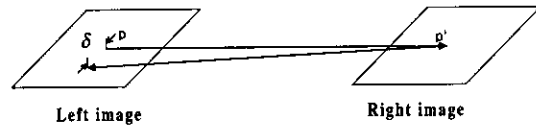


그림 8. 양방향 검사

입력 영상에 대한 정합의 결과가 어느 정도 유일성 제약조건을 만족하는가를 정의하기 위해서 다음의 신뢰성 함수를 정의한다. 이러한 정의는 정해진 임계치에 대해서 그 오차가 증가된 정도를 나타내며 실험의 결과는 계층적인 구조로 인해서 하위계층으로 갈수록 그 오차가 감소됨을 알 수 있었다. 식(6)은 제안된 알고리즘에서 사용될 첫 번째 확률값이 다음 단계에서 구해질 확률값과 함께 사용이 된다.

$$f_{R1} = \frac{T_{bc} - \delta}{T_{bc}} \quad \left\{ \begin{array}{l} T_{bc} : \text{threshold} \\ \delta : \text{bidirectional check distance} \end{array} \right. \quad (6)$$

f_{R1} 은 유일성 신뢰도를 나타내고 δ 는 양방향 검사의 오류를 나타내고 T_{bc} 는 최소 임계치를 가리킨다. 여기에서 사용된 임계치는 표 2에서와 같다.

4.3.2 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.

$$w_s(x, y) = \begin{cases} 0 & , \delta > T_{bc} \\ 1 & , \delta < T_{bc} \end{cases} \quad (7)$$

다음으로 추출된 변위도에서의 분산은 다음과 같이 정의된다.

$$\sigma_D^2(x, y) = \left[\frac{1}{N_V} \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} \{w_s(i+x, j+y) \cdot I(i+x, j+y)\}^2 \right] - \left[\frac{1}{N_V} \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} w_s(i+x, j+y) \cdot I(i+x, j+y) \right]^2 \quad (8)$$

그리고 N_V 는 다음과 같이 정의된다.

$$N_V = \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{w}{2}}^{\frac{w}{2}} w_s(i+x, j+y). \quad (9)$$

위와 같은 관계로부터 평탄 신뢰성을 다음과 같이 정의한다.

$$f_{R2} = e^{-\xi \sigma_D^2}, \quad (10)$$

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

4.3.3 신뢰도 함수

신뢰도 함수는 f_{R1} 과 f_{R2} 의 선형조합으로 구성이 되며 이러한 조합은 평탄성의 조건과 유일성의 조건을 모두를 포함하는 결과를 가져오며 가중치 계수를 조절함으로써 신뢰성 정보의 특성을 조절할 수 있다. 신뢰도 함수의 정의는 아래와 같다.

$$f_{total} = \lambda_1 f_{R1} + \lambda_2 f_{R2}, \quad (11)$$

식(11)에서 λ_1 과 λ_2 는 $\lambda_1 + \lambda_2 = 1$ 의 관계를 만족시키는 가중치 계수이다.

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.

$$D'(x, y, d) = \arg \min_{d(x, y)} \left[\frac{1}{mn} \sum_{y'=-\frac{w}{2}}^{\frac{w}{2}} \sum_{x'=-\frac{w}{2}}^{\frac{w}{2}} |I'_s(x+x', y+y') - I'_s(x+d+x', y+y')| \right] \quad (12)$$

$$S'(x, y, d) = \frac{1}{9} \sum_{y'=-1}^1 \sum_{x'=-1}^1 D'^{-1}(x', y') - D', \quad (13)$$

식(12)는 초기에 구한 변위값의 신뢰도에 대한 확률값으로부터 구한 와핑된 윈도우와 기준영상의 윈도우와의 MAD를 정의한 식이고, 식(13)은 상위 계층 9개의 변위벡터와 하위계층의 벡터와의 차로부터 구한 평탄성을 정의한 식이다.

식 (12)와 식(13)의 결과로부터 식(14)가 정의되며 각각의 계수는 영상에서의 특성을 고려한 계수이다. 실험에서는 모두 0.5를 사용하였으며 값을 크게 변화시키지 않는한 결과는 크게 변하지 않았다.

$$E^i = \gamma_1 D^i + \gamma_2 S^i, \quad (14)$$

식(14)에서 γ_1 과 γ_2 는 $\gamma_1 + \gamma_2 = 1$ 의 관계를 만족시키는 가중치 계수이다.

V. 실험 및 고찰

실험에서 사용된 'man'스테레오 영상에 대한 변수는 표 2에 나타난 것과 같다.

표 2. 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

실험에 사용된 영상은 384×384 크기를 가지는 'man' 스테레오 영상과 720×288크기의 'claude'영상을 사용하였다. 실험에 사용된 스테레오 영상은 극상선 조건을 만족시키므로 일차원 탐색만을 수행하여 변이를 추정하였다. 'man'영상은 질감정보가 작고 근거리에서 획득된 영상이므로 투영적인 왜곡 현상이 뚜렷하다. 그리고 하나의 조명만 사용되었기 때문에 광원의 영향이 오정합을 일으키는 요소로 작용한다. 'claude'영상은 'man'영상에 비해서 변이가 적고 스테레오 카메라의 초점 중심이 물체의 내부에 존재하기 때문에 배경이 더 큰 변이를 가지게 된다. 그림10에서는 객체추출 이전 단계인 제안한 방법에 대해서 기존의 블록 기반의 방법과 개체군 증가 학습법에 의한 방법의 비교가 나타나 있다. 그림에서 보이는 것처럼 기존의 방법에 비해서 질감 정보가

작은 머리 부분과 볼부분에 있어서 정합의 결과가 부드러움을 알 수가 있다. 기존의 방법이 오정합을 줄이기 위해서 이완처리의 단계를 가지지만 제안된 알고리즘에서는 이완처리의 단계를 가지지 않고도 인접 정합의 결과가 반영되어 오정합이 감소되었다. 그림 12의 'claude'영상에서도 제안된 방법이 기존의 방법에 비해서 정합의 결과가 우수함을 나타내고 있다. 실험에서 사용된 스테레오 영상의 경우에 정합의 대상에서 제외되어야 할 배경 부분을 가지므로 4장에서 설명된 방법에 의해서 객체를 추출 후 정합을 실시하였다. 객체 추출의 결과는 그림 13와 그림 14에 나타나 있으며 객체 추출 후의 정합 결과는 그림 15에 나타나 있다. 실험의 결과에서 보듯이 객체 내부에서는 오정합이 거의 발생되지 않았으며 객체 추출과정을 가져서 정합의 속도가 빨라졌음을 실험을 통하여 확인할 수 있었다.

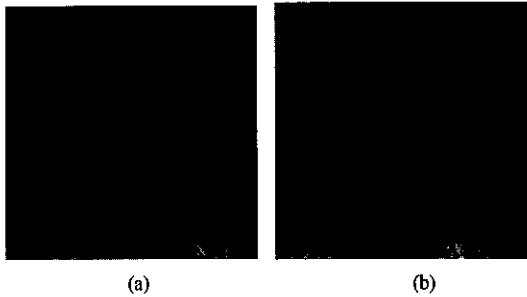


그림 9. 'man' 스테레오 영상. (a) 좌영상 (b) 우영상

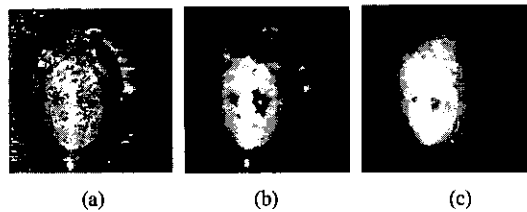


그림 10. 기존의 방법과 제안한 방법에 의한 변위도
(a) BMA의 변위도 (b)PBIL방법에 의한 변위도
(c) 객체추출을 하지 않은 제안한 방법에 의한 변위도

그림 10은 객체추출을 하지 않은 변위도(제안한 방법)와 일반적인 BMA (block matching algorithm)와 개체군기반 증가 학습방법(PBIL: population based incremental learning)에 의한 방법을 비교한 그림이다. 그림10의 (a)의 경우에 배경부분에서는 정합정보가 없기 때문에 잡음이 심하고 주위의 변위값을 고려하지 않았기 때문에 객체부분에서도 잡음이 상당히 심하다. 또한 투영적 왜곡이 일어나는

부분과 폐색영역에 대해서는 오정합이 일어났음을 확인할 수 있다. 그림 10의 (b)에서는 변위값 자체가 주위 변위벡터의 값을 고려하는 알고리즘을 포함하고 있기 때문에 변위도에서 잡음적인 결과는 없지만 오정합이 많이 일어난 영역이나 폐색영역에서는 많은 오정합 변위값들이 모여있기 때문에 오차가 더 증가함을 발견할 수 있다. 하지만 그림 10의 (c)에서는 그림(a)나 그림(b)에서와 같은 잡음은 상당히 줄었음을 발견할 수 있다. 하지만 배경과 객체의 경계부분에서는 오정합이 발생했음을 발견할 수 있다. 하지만 이러한 문제는 4.1절에서 다른 객체추출의 단계를 거침으로 인해서 정합의 신뢰성을 증가시키게 된다.



그림 11. 'claude' 스테레오 영상. (a) 좌영상 (b) 우영상

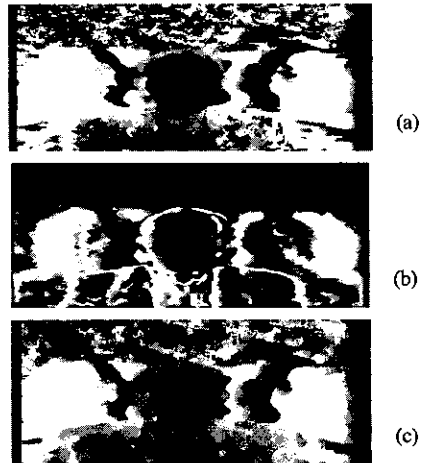


그림 12. 기존의 방법과 제안한 방법에 의한 변위도
(a)BMA의 변위도 (b)PBIL에 의한 변위도 (c)객체추출을 하지 않은 제안한 방법에 의한 변위도

그림 12의 결과도 그림10의 결과와 마찬가지로 (a)에서는 배경과 객체내에서 많은 오정합의 결과를 보여주며 그림(b)에서는 오정합의 결과가 인접 변위 벡터로 전파가 이루어져서 객체내에서 정합의 신뢰성이 떨어짐을 알 수 있다. 그림(c)에서는 정합정보가 존재하지 않는 배경을 제외하고는 객체 내에서는 정합의 신뢰성이 높음을 확인할 수 있다. 'claude'영상의 경우에는 스테레오 카메라의 초점이 인체의 내부에 존재하기 때문에 객체의 변위값이

양(+)¹의 값과 음(-)¹의 값을 동시에 가지며 배경이 더 큰 변위를 가지게 된다. 따라서 객체의 변위값은 배경에 비해서 작은 값을 가진다. 실험에 사용된 'claude'영상은 첫 번째 프레임의 4개의 다중 화면 중에서 왼쪽 끝과 오른쪽 끝의 영상이다. 중간 영상은 왜곡현상이 작고 폐색영역이 작기 때문에 실험에서는 제외되었다.

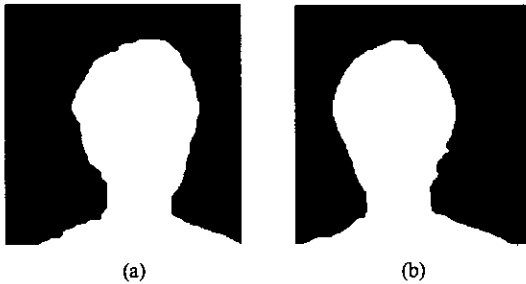


그림 13. 'man' 스테레오 영상의 객체 추출의 전처리 결과
(a) 좌영상의 객체추출 결과
(b) 우영상의 객체추출 결과



그림 14. 'claude' 스테레오 영상의 객체 추출의 전처리 결과
(a) 좌영상의 객체추출 결과
(b) 우영상의 객체추출 결과

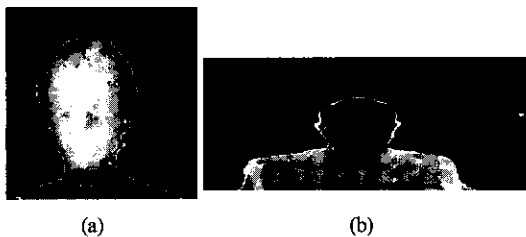


그림 15. 제안된 알고리즘에 의한 정합 결과
(a) 'man' 영상의 정합도
(b) 'claude' 영상의 정합도

그림 15는 객체 추출의 과정을 가진 정합의 결과를 나타내고 있다. 정합의 대상정보가 없는 배경을 제거한 뒤 객체 내에서만 정합을 실시한 결과이며 이상의 결과와 비교해 볼 때 투영적 왜곡이 일어난 영역과 광원의 영향으로 인한 오정합을 벗어난 결과를 보여주고 있다.

VI. 결 론

본 논문에서는 적응적 윈도우 외평을 이용한 새로운 계층적 스테레오 정합을 제안하였다. 제안한 알고리즘은 변위값에 대한 확률적 모델에 근거해서 적응적으로 윈도우를 선택해서 계층적인 정합을 한다. 기존의 방법이 스테레오 카메라의 기하학적인 구조로 인한 투영적인 왜곡현상에 대해서 고려를 하지 않고 획득된 영상만으로 정합을 했기 때문에 오정합을 피할 수가 없었지만 제안된 방법에서는 왜곡현상을 줄이기 위해서 윈도우를 적응적으로 외평시키면서 계층적인 정합을 실시했다. 정합의 시간을 줄이고 오정합을 감소시키기 위해서 전처리과정으로 객체추출의 과정을 거치며 이로 인해서 정합 대상이 감소하여 속도가 개선되고 신뢰성이 증가되었다. 윈도우를 외평시킬 때 정합의 신뢰도를 이용함으로써 해서 왜곡이 일어난 영역과 일어나지 않은 영역을 분리해서 정합을 함으로써 정합의 시간 뿐만 아니라 정합의 신뢰성도 증가 시켰다. 제안된 알고리즘이 계층적인 정합을 수행하기 때문에 영상의 잡음에 영향과 변이 탐색 영역의 감소로 인한 오정합의 가능성을 감소시켰다. 실험의 결과에서는 기존의 정합 방법에 비해서 질감정보가 적고 투영적 왜곡이 일어나는 영역에 대해서 제안한 알고리즘이 정확한 정합을 수행함을 보여주었다.

대다수의 스테레오 카메라가 큰 base-line을 가지지만, 스테레오 영상이 원거리에서 촬영되거나 base-line이 좁을 경우에서 제안된 방법을 사용하더라도 영상에서 이러한 영역이 작고 현저하지 않기 때문에 성능의 향상은 크지 않을 것이다. 이러한 점을 극복하기 위해서 영상의 특징 정보를 윈도우 외평의 비율을 구하는데 사용하여 하이브리드 방식을 채택함으로써 개선할 수 있을 것이다.

참 고 문 헌

- [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] W. Eric L Grimson, "Computational Experiments with a Feature Based Stereo Algorithm," *IEEE Trans. Patt. Anal. Machin Intell.*, vol. PAMI-7, no. 1, pp.17-34, 1985.
- [4] Gerard Medioni and Ramakant Nevatia,

"Segment-based stereo matching," *Computer Vision, Graphics, and Image Processing*, vol. 31, pp. 2-18, 1985.

[5] Song De Ma, "Conics-based stereo, motion estimation, and pose determination," *Intern. J. Comp. Vision*, vol. 10, no. 1, pp. 7-25, 1993.

[6] Dimitrios S. Kalivas and Alexander A. Sawchuk, "A region matching motion estimation algorithm," *CVGIP: Image Understanding*, vol. 54, no. 2, pp. 275-288, 1991.

[7] John P. Jordan III and A. C. Bovik, "Using chromatic information in edge-based stereo correspondence," *CVGIP: Image Understanding*, vol. 54, no. 1, pp. 98-118, 1991.

[8] Masatoshi Okutomi and Takeo Kanade, "A Multiple-baseline stereo," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 15, no. 4, pp. 353-363, 1993.

[9] John (Juyang) Weng, "Image matching using the windowed Fourier phase," *Intern. J. Comp. Vision*, vol. 15, no. 4, pp. 211-236, 1993.

[10] Han, Kyu-Phil, Tae-Min Bae, and Yeong-Ho Ha, "Hybrid Stereo Matching with a New Relaxation Scheme of Preserving Disparity Discontinuity," *Pattern Recognition*, vol. 33, no. 5, pp. 767-785, May 2000.

[11] 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.

[12] Z. F. Wang and N. Ohnish, "Intensity-based Stereo Vision: from 3-D to 3-D," *SPIE*, vol. 2, pp. 434-443, Nov. 1994.

[13] 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.

[14] George Wolberg, *Digital Image Warping*, pp. 153-160 *IEEE Computer Society*, 1994

[15] R. Y. Tsai, "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision", *Proce, CVPR*, pp.364-374, 1986

[16] R. Y. Tsai, "A Versatile Camera Calibration Technique for High-accuracy 3D Machine

Vision Metrology Using Off-the-Shelf TV Camera and Lenses." *IEEE J. Robotics and Automation*, vol. RA-3, no. 4, pp. 323-344, 1987.

김 재 철(Jae-Chul Kim)

정회원



1999년 2월 : 경북대학교
전자공학과 졸업(공학사)
2001년 2월 : 경북대학교 대학원
전자공학과 졸업
(공학석사)
2001년 1월~현재 : 한국전자
통신연구원 연구원

<주관심 분야> 영상처리

이 호 근(Ho-Keun Lee) 정회원



1997년 8월 : 경북대학교
전자공학과 졸업(공학사)
2000년 2월 : 경북대학교 대학원
전자공학과 졸업
(공학석사)
2000년 3월~현재 : 경북대학교
대학원 전자공학과
박사과정

<주관심 분야> 스테레오 비전, 3D 모델링, 가상현
실

하 영 호(Yeong-Ho Ha) 정회원

제24권 제12호 참조