

영상의 컬러 정보를 이용한 계층적 스테레오 정합

준회원 김 태 준*, 정회원 유 지 상*

Hierarchical Stereo Matching with Color Information

Tae-june Kim* Associate Member, Ji-sang Yoo* Regular Member

요 약

본 논문에서는 컬러 정보를 이용한 계층적 스테레오 정합 기법을 제안한다. 특징기반의 스테레오 정합 방법을 이용하여 초기 변이지도를 생성하고, 계층적 스테레오 정합 기법으로 최종 변이지도를 획득한다. 영상을 R, G, B, white 4개의 색상 성분으로 분할하여 영상의 경계(edge)를 추출하고, 추출된 경계에서 정합 창을 이용하여 변이(disparity)를 추정한다. 추정된 변이는 각 색상 성분에서 자기상관도(autocorrelation)에 따라 주변 영역으로 확산되어 초기 변이지도(disparity map)를 생성한다. 초기 변이지도는 최종 변이지도를 생성하기 위한 변이 탐색의 초기 값으로 사용되고, 각 색상 성분에서 정합 창과 탐색 범위(search range)의 변화를 이용하여 최종 변이 지도를 생성시킨다. 본 논문에서는 Middlebury stereo vision의 4개의 실험 영상을 가지고 객관적 성능 평가를 하였다. 실험 결과 제안한 기법이 기존의 Graph-cuts와 Dynamic Programming 기법보다 우수한 성능을 보였다. 최종 변이지도의 부정확한 변이는 전체 영상에서 평균 11% 존재했고, 변이지도에서 불연속점의 경계가 뚜렷한 것을 확인하였다.

Key Words : stereo image, feature, hierarchical stereo matching, disparity, disparity map

ABSTRACT

In this paper, a hierarchical stereo matching with color information is proposed. To generate an initial disparity map, feature based stereo matching is carried out and to generate a final disparity map, hierarchical stereo matching is carried out. The boundary (edge) region is obtained by segmenting a given image into R, G, B and White components. From the obtained boundary, disparity is extracted. The initial disparity map is generated when the extracted disparity is spread to the surrounding regions by evaluating autocorrelation from each color region. The initial disparity map is used as an initial value for generating the final disparity map. The final disparity map is generated from each color region by changing the size of a block and the search range. 4 test images that are provided by Middlebury stereo vision are used to evaluate the performance of the proposed algorithm objectively. The experiment results show better performance compared to the Graph-cuts and Dynamic Programming methods. In the final disparity map, about 11% of the disparities for the entire image were inaccurate. It was verified that the boundary for the non-contiguous point was clear in the disparity map.

I. 서 론

최근 전 세계적으로 3D 입체 영상에 대한 연구가 활발히 이루어지고 있다. 이미 일본은 2008년 4월부터 BS11에서 3D 입체방송을 서비스하고 있고, MPEG

에서도 다시점 비디오 부호화(MVC:multiview video coding) 기술의 표준화가 완료 단계에 있으며, 자유 시점 텔레비전(FTV:freeviewpoint TV)에 대한 표준화 작업도 활발히 이루어지고 있다. 이와 같이 입체 영상에 관심을 가지는 이유는 미디어가 소비자의 요

※ 본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업(IITA-2009-(C1090-0902-0017))과 서울시 산학연 협력사업(NT080528) 지원으로 수행되었음.

* 광운대학교 전자공학과 디지털 미디어 연구실 (mirrine@kw.ac.kr, jsyoo@kw.ac.kr)

논문번호 : KICS2009-01-003, 접수일자 : 2009년 1월 15일, 최종논문접수일자 : 2009년 2월 20일

구에 따라 실감 미디어를 향해 발전하고 있기 때문이다. 영상에서의 실감은 사용자가 실제로 시청하고 있는 영상 안으로 들어와 있는 듯한 입체감을 주는 것이 목적이다. 사람이 입체감을 느끼는 방법에는 여러 가지가 있는데 대표적으로 알려져 있는 방법은 양안시차(binocular disparity)에 의한 방법이다. 사람의 눈은 평균적으로 좌, 우 6.5cm의 간격을 두고 존재하기 때문에 좌 안과 우 안에 각각 차이가 있는 다른 영상이 맺혀 물체와의 거리감을 인지할 수 있다. 이것을 양안 시차라 한다. 양안시차의 원리를 이용하여 입체감을 느끼기 위해서는 3D 모니터에서 사용이 가능한 스테레오 영상이 필요하다. 스테레오 영상이란 사람의 좌 안과 우 눈에 해당하는 좌, 우 영상을 지칭하며, 스테레오 영상을 3D 모니터를 통해 보게 되면 시청자는 양안 시차에 의한 입체감을 느낄 수 있게 된다.

스테레오 영상의 좌, 우 영상은 인접한 시점에서 획득되기 때문에 상관도가 매우 높다. 이러한 높은 상관도를 이용하여 대응점을 찾아내는 과정을 일반적으로 스테레오 정합(stereo matching)이라 한다. 두 영상간의 상관도에 따른 대응점을 찾아 상대적인 거리를 표시한 것이 변이지도(disparity map)이고, 변이지도는 두 영상간의 상관도를 나타내는 중요한 정보를 가지고 있다. 이러한 변이정보를 이용하면 중간시점 영상 합성(IVR:intermediate view reconstruction)을 할 수 있고, 획득된 영상의 카메라 정보를 이용하여 실제 깊이 정보 추출이 가능하다. 뿐만 아니라 변이 지도로부터 특정 영역을 분리하여 새로운 영상과의 합성도 가능하다. 이처럼 스테레오 영상에서 변이지도는 상당히 중요한 의미를 갖고 있다.

스테레오 정합 방법에는 대표적으로 영역기반(window based), 특징기반(feature based), 에너지기반(energy based) 등이 있다^{14,15,16}. 영역기반이란 특정 크기의 정합 창을 이용하여 정합 창 내부에 존재하는 화소와의 오차가 적은 다른 시점의 정합 창을 찾아내는 방법이다. 가장 간단하면서도 효율성이 높은 방법이지만, 폐색영역(occlusion)과 균일한 색상을 가지는 영역(homogeneous region) 등에서 좋지 않은 결과를 발생시키고, 정합 창의 크기에 따른 결과가 다르다는 단점이 있다. 특징기반은 영역기반과 달리 전체 영상에서 변이를 찾는 것이 아니라 특징 점에서만 변이를 찾는다. 특징 점은 영상에서의 모서리, 경계선과 같은 부분이 해당된다. 특징 점 이외의 변이가 존재하지 않는 영역은 특징 점에서 획득된 변이를 내삽법(interpolation)과 외삽법(extrapolation)을 이용

하여 채운다. 내삽법과 외삽법에 의한 에러 전파는 생성되는 변이지도의 정확성에 영향을 미치는 단점이 있다. 최근에 가장 우수한 정합 방법들 대부분은 신뢰확산(belief propagation)과 같은 에너지 기반의 알고리즘을 사용한다. 주변의 화소값이 가지는 변이 값을 고려하여 최적의 확률을 선택하여 변이지도를 생성하는 신뢰도 높은 방법이지만, 복잡도가 높아 수행시간이 오래 걸리는 단점이 있다.

본 논문에서 제안하는 영상의 컬러 정보를 이용한 계층적 스테레오 정합 기법은 영상의 컬러 정보를 이용하여 영상을 분할하고 각 독립된 분할 영역에서 변이를 추정하여 특징기반의 스테레오 정합과 계층적 스테레오 정합이 내포하고 있는 에러 전파의 단점을 보완하는 신뢰도 높은 기법이다. 계층적 스테레오 정합은 매 계층마다 영상의 크기, 정합 창 및 탐색범위 등을 제한하며 스테레오 정합을 하는 방법이다¹¹. 매 계층은 이전 계층에서 생성된 변이 정보를 이용하여 스테레오 정합을 수행하기 때문에 이전 계층에서 발생하는 에러가 전파가 된다. 따라서, 이러한 에러 전파를 방지하기 위하여 영상을 R, G, B, white의 4가지 색상 성분으로 분할하고, 각 색상 성분에서 독립적으로 변이를 추정한다. 초기 변이지도를 생성하기 위하여 분할된 각 색상 성분에서 특징기반의 스테레오 정합을 수행하고, 오 정합을 검출하여 확산에 의한 에러 전파를 방지하는 cross checking 기법을 이용한다². 생성된 초기 변이지도는 최종 변이지도를 생성하기 위하여 계층적 스테레오 정합의 초기값으로 이용된다. 계층적 스테레오 정합의 매 계층에서는 이전 계층의 변이 값을 이용하여 탐색범위의 중심점을 정하고, 변이에 따라 참조영상을 분할한다. 변이에 따라 분할된 각 영역은 R, G, B, white의 색상 성분으로 분할되어 하나의 정합 창 역할을 하여 계층적 스테레오 정합을 수행한다. 또한, 매 계층에서 폐색영역과 cross-checking을 이용한 오 정합 검출을 통해 신뢰도 높은 변이지도를 생성한다.

본 논문의 II장에서는 제안하는 영상의 컬러 정보를 이용한 계층적 스테레오 정합 기법을 설명한다. III장에서는 본 논문에서 제안하는 알고리즘의 실험 결과를 비교하고, IV장에서 결론을 맺는다.

II. 영상의 컬러 정보를 이용한 계층적 스테레오 정합 기법

본 논문에서 제안하는 방법은 크게 두 부분으로 이루어져 있다. 초기 변이지도를 생성하는 부분과



그림 1. 제안하는 스테레오 정합 과정
Fig. 1. Proposed algorithm for stereo matching

최종 변이 지도를 생성하는 부분이다. 그림 1은 본 논문에서 제안하는 스테레오 정합 과정의 전체 구조를 나타낸다. 초기 변이 지도는 참조영상의 특징 점에서 정합 창을 이용하여 추출된 변이가 자기 상관도에 의해 확산됨으로써 생성되고, 최종 변이 지도는 페색영역과 오 정합 등을 고려하여 계층적 스테레오 정합에서 발생할 수 있는 에러 전파를 최대한 방지하여 생성된다.

2.1 초기 변이 지도 생성

2.1.1 영역 분할

초기 변이 지도를 생성하기 위하여 먼저 참조영상을 그림 2와 같이 각 색상 성분으로 영역분할을 하고, 각 분할된 영역에서 경계 변이 지도를 생성한다. 각 분할된 영역에서 초기 변이 지도를 생성하는 독립적인 과정을 수행하게 되면 인접 유사 화소 간의 간섭 효과가 낮아지기 때문에 확산을 통한 신뢰도 높은 초기 변이 지도를 생성할 수 있다.

그림 2는 참조영상을 4가지 색 성분으로 영역 분할하는 원리를 나타낸다. RGB 영상에서 white축을 중심

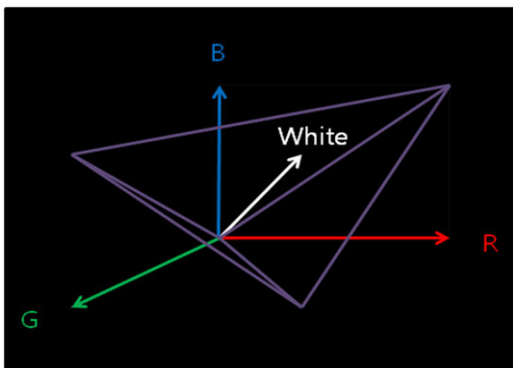


그림 2. R, G, B, White 영역 분할
Fig. 2. Segmentation of R, G, B and White

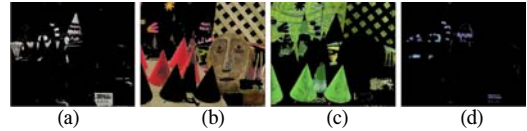


그림 3. 참조 영상의 성분의 색 영역 분할 (a) white, (b) red, (c) green, (d) blue
Fig. 3. Segmentation of a reference image (a) white, (b) red, (c) green, (d) blue

으로 일정한 거리 안에 존재하는 값은 white로 변환하고, 나머지 영역은 각 색상 성분의 값을 비교해 비율이 많은 값을 차지 하는 색상 성분으로 판단하게 된다.

참조영상을 white 성분을 제외한 R, G, B 3개의 색상 성분으로 변환을 하면, 각 색상 성분간의 경계가 모호함으로 인해 변이 생성에 있어 오히려 낮은 신뢰도의 결과를 도출한다. 그러나 white축을 추가 함으로써, 이러한 모호한 경계를 개선시키는 효과가 있어 신뢰도 높은 변이를 얻을 수 있다. White 영역은 발생할 영역의 범위를 제한함으로써 증가 혹은 감소시킬 수 있다. 그림 3은 초기 변이 지도를 생성하기 위해 참조영상을 각 색상 성분으로 분할한 영상을 나타낸다.

2.1.2 경계 변이 지도 생성

영상은 일반적으로 모서리, 경계 등에서 특징 점이 존재하며, 이러한 특징 점을 추출하기 위한 방법은 다양하다. 본 논문에서는 소벨(Sobel) 마스크를 이용하여 각 색상 성분에서 참조영상의 경계를 추출하고, 이를 특징 점으로 이용한다. 그림 4는 수평, 수직의 소벨 마스크와 이를 이용하여 획득된 각 색상 성분에서의 경계 영상을 나타낸다.

경계 변이 지도를 생성하기 위하여 3×3, 5×5, 7×7, 9×9의 정합창을 사용하여 실험을 하여 결과를 비교해 본 결과, 7×7의 정합창 성능이 가장 우수하여 영상의 경계에서 7×7 정합 창을 이용하고, 비용 함수는 식 (1)의 SSD(sum of square difference)를 이용한다.

$$d(x, y) = \arg \min_{d_{\min} \leq d \leq d_{\max}} \sum_{(i,j) \in N(x,y)} (R(i+d, j) - L(i, j))^2 \quad (1)$$

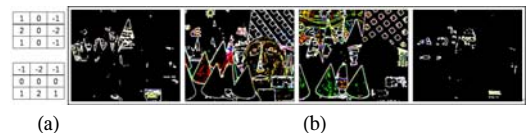


그림 4. (a) 수직, 수평 소벨(Sobel) 마스크, (b) 각 색상 성분에서의 경계영상
Fig. 4. (a) Vertical and Horizontal Sobel mask, (b) Edge image on each color components

여기서 $N(x,y)$ 은 (x,y) 를 중심으로한 7×7 정합 창을 나타낸다. (i,j) 는 정합 창 내의 화소 위치를 나타낸다. $R(x,y)$ 와 $L(x,y)$ 는 좌, 우 영상의 화소값을 나타내며, d 는 범위내에 존재하는 변이를 나타내고, $d(x,y)$ 는 (x,y) 에서 얻어진 최종 변이 값이다.

영상의 균일한 영역은 영상의 경계 부분보다 변이를 예측할 수 있는 정보가 상대적으로 적으며 따라서 경계 부근에서 변이를 예측하게 되면, 보다 신뢰도 높은 변이를 예측할 수 있다. 하지만, 영상의 경계는 전체 영상에서 일부분에 지나지 않으므로, 경계에서 예측된 변이는 주변 영역의 정보를 반영하도록 하여야 한다. 일반적으로 변이의 불연속점은 대부분 영상의 경계에서 발생하고, 영상의 균일한 영역에서는 변이가 유사하다는 특징이 있기 때문에 영상의 불연속점인 특징 점에서 예측된 변이가 주변 영역으로 확산되어 생성되는 초기 변이지도는 최종 변이를 추정하기 위한 초기값으로 사용하기에 충분한 신뢰성을 가진다고 할 수 있다. 하지만 부정확한 변이가 경계 변이지도에서 발생하면, 확산이 되는 과정에서 영상 전체로 여러 값이 확산된다. 본 연구에서는 이러한 부정확한 변이를 검출하기 위해서 획득된 변이를 참조영상에 적용해 원영상과의 오차를 판단하는 cross-checking 방법을 이용한다^[2]. 식 (2)는 cross-checking을 이용하여 부정확한 변이를 검출하는 식을 나타낸다.

$$\begin{aligned}
 \text{CrossChecking} &= |R(x + \text{Dis}(x, y), y) - L(x, y)| \\
 \text{Dis}(x, y) &= \begin{cases} \text{Dis}(x, y), & \text{CrossChecking} < \text{Tolerance} \\ 0, & \text{else} \end{cases} \quad (2)
 \end{aligned}$$

여기서 $R(x,y)$ 는 우영상의 화소값, $L(x,y)$ 는 좌영상의 화소값을 나타내고, $\text{Dis}(x,y)$ 는 $R(x,y)$ 에서의 변이 값을 나타낸다. 이상적인 영상이 획득 되었을 때 참조영상 화소에 대응되는 원영상 화소의 차이 값은 '0'이 되어야 한다. 하지만 일반적으로 영상이 획득 될 때 조명 등의 주변 환경과 카메라의 특성에 의해 차이 값은 '0'이 되지 않는다. 이러한 오차의 영향을 줄이기 위해 허용오차를 설정한다. Cross-checking 값이 허용오차 이내에 있다면 신뢰도 높은 변이라 판단하고, 그렇지 않다면 부정확한 변이라 판단하여 변이를 '0'으로 초기화 시켜 오 정합이라 판단한다.

그림 5의 (a)는 원영상, (b)는 소벨(Sobel) 연산자를 이용하여 각 색상 성분에서의 경계영상을 취합한 그림이고, (c)는 cross-checking을 한 후 영상의 각 색상성분의 경계에서 정합 창을 이용하여 획득

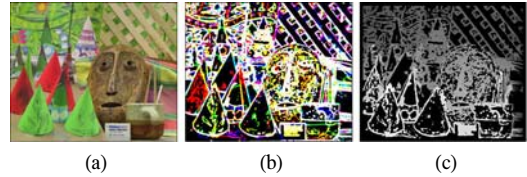


그림 5. (a) 원본영상, (b) 경계영상, (c) 경계 변이지도
Fig. 5. (a) Original image, (b) Edge image, (c) Boundary disparity map

한 경계 변이지도를 취합한 결과를 나타낸다.

2.1.3 확산

본 연구에서는 영상의 경계에서 획득한 신뢰도 높은 변이를 주변 유사한 영역으로 확산되도록 한다. 이때 색상의 유사도를 가지고 화소의 유사성을 판단하게 되며 경계변이에 해당하는 참조영상의 화소값과 주변의 유사한 화소값을 spiral search를 통해 찾는다^[3]. 탐색은 참조영상에서만 수행하고, 얻어지는 좌표 정보를 이용하여 경계 변이지도에서 확산작업을 수행한다. 그림 6은 spiral full search 방법을 나타낸 그림이다.

참조영상의 탐색범위 내에서 중심점의 화소값과 유사하다고 판단된 화소의 좌표를 이용하여 경계 변이지도에서 확산을 수행한다. 유사한 화소의 판단 기준은 식(3)을 이용한다. R, G, B의 각 색상 성분에 대한 비용함수 합이 허용오차보다 작으면 유사한 화소라 판단하고, 그렇지 않으면 불확실한 화소라 판단하여 확산을 수행하지 않는다.

$$\begin{cases} \text{Valid} & , |R_{r,g,b}(x + \Delta x, y + \Delta y) - R_{r,g,b}(x, y)| < \text{Tolerance} \\ & (-\text{SearchRange} \leq \Delta x, \Delta y \leq \text{SearchRange}) \\ \text{Invalid, else} & \end{cases} \quad (3)$$

여기서 $R(x,y)$ 는 참조영상을 나타내고, Δx 와 Δy 는 x 와 y 에서의 상대적인 거리를 나타낸다. r, g, b는 RGB영상의 각 red, green, blue 성분을 나타낸

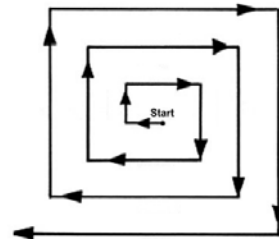


그림 6. Spiral full search
Fig. 6. Spiral full search



그림 7. 초기 변이지도
Fig. 7. Initial disparity map

다. 확산에서 사용되는 영상은 RGB의 원본영상을 사용하기 때문에 white 색상 정보는 무의미 하므로, white 색상 정보는 이용하지 않는다.

2.1.4 초기 변이지도 생성

각 색상 성분에서의 경계 변이 확산은 각 색상 성분을 고려하지 않았을 때보다 높은 신뢰성을 보장한다. 변이지도는 일반적으로 균일한 색상 성분에서 유사한 변이를 가진다. 따라서 균일한 색상 성분에서 획득된 불연속적인 변이를 평활화하기 위하여 평균필터를 적용하여 초기 변이지도를 생성하게 된다. 그림 7은 위의 과정을 거쳐 획득된 초기 변이지도를 나타낸다.

2.2 최종 변이지도 생성

2.2.1 폐색영역 검출

신뢰도 높은 변이지도를 생성하기 위해서는 폐색영역과 오 정합 검출이 중요하다. 폐색영역 검출이 필요한 이유는 폐색영역에서 부정확한 변이가 많이 나타나고, 폐색 영역에 부정확한 변이가 존재 했을 경우 계층적 스테레오 정합의 다음 계층에서 폐색영역에 부정확한 변이가 할당되는 것을 방지하기 위해서이다. 폐색영역에서 부정확한 변이가 많이 나타나는 이유는 참조영상에는 존재하지만 원 영상에는 존재하지 않는 유사한 영역을 찾아 변이를 폐색영역에 할당하기 때문이다. 따라서 폐색영역을 검출하고 계층이 증가 할 때마다 이전 계층의 폐색영역을 제외한 변이가 존재하는 영역에 대해서만 계층적 스테레오 정합을 한다.

그림 8에서 x축은 영상의 스캔라인에 해당하고, d축은 변이를 나타낸다. Segment B는 그림 8(a)의 큰 변이를 가지는 영역에 해당하고, Segment A,

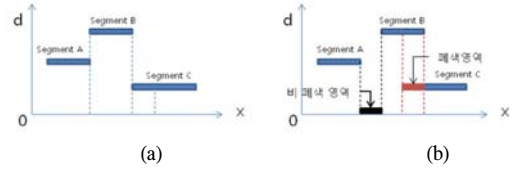


그림 8. 폐색영역 판단 과정 (a) 참조영상의 변이지도, (b) 원 영상에 대응되는 변이지도
Fig. 8. Process to determine an occluded region
(a) Disparity map for the reference image, (b) Disparity map corresponding to the original image

Segment C는 상대적으로 작은 변이를 가지는 영역에 해당한다. 스캔라인을 따라 변이 값을 탐색 할 때 작은 변이에서 큰 변이로 변할 때에는 그림 8(b)에서 보듯이 겹쳐지는 영역에 해당하는 비 폐색영역(disocclusion)이 나타나고, 큰 변이에서 작은 변이로 변할 때에는 겹쳐지는 영역에 해당하는 폐색영역이 나타난다. 비 폐색영역은 참조영상에 존재하지 않는 영역이므로 변이 지도를 생성할 때 오차를 발생할 수 있는 영역이 아니다. 따라서 변이지도를 생성시에는 폐색영역만을 고려한다. 일반적으로 폐색영역은 큰 변이와 작은 변이 사이의 불연속점에서 발생한다. 식(4)는 폐색영역을 검출하기 위한 식이다.

$$Length = SegmentB - SegmentC \quad (4)$$

여기서 Segment B는 현재 위치에서의 변이이고, Segment C는 이웃한 변이이다. 그리고 이 두 값의 차이가 폐색영역의 길이가 된다. Length가 음수가 발생 했을 때는 비 폐색영역이 발생 했을 때 이므로 폐색영역이라 판단하지 않고, Length가 양수일 때만 폐색영역이라 판단하고 변이 지도상에 '0'으로 초기화 시킨다. 그림 9는 위 과정을 거친 후 검출된 폐색영역을 나타낸 것이다.

2.2.2 계층적 스테레오 정합

이전 계층의 변이를 이용하여 다른 변이 값에 따라 참조영상을 분할하면 그림 10과 같이 된다. 그림

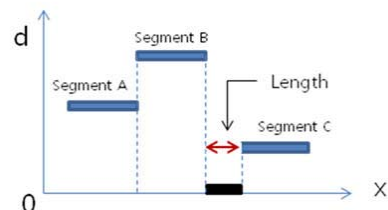


그림 9. 검출된 폐색영역
Fig. 9. Extracted occlusion region

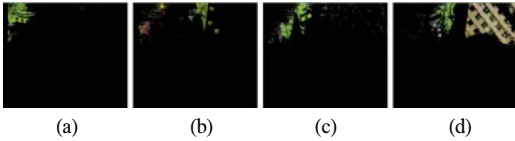


그림 10. 각 변이에 따른 참조영상의 영역 분할
 (a) 변이값이 18인 경우, (b) 변이값이 19인 경우, (c) 변이값이 20인 경우, (d) 변이값이 21인 경우
 Fig. 10. Segmentation of reference image with disparity
 (a) disparity=18, (b) disparity=19, (c) disparity=20, (d) disparity=21

10의 (a), (b), (c), (d)는 각각 변이 값이 18, 19, 20, 21일 때 해당되는 참조영상을 보여준다. 그림 10의 각 영상들을 다시 각각의 색상 성분으로 영역을 분할하고, 분할된 영역을 정합 창으로 이용하여 변이를 추정한다. 분할된 영역은 계층적 스테레오 정합의 매 계층에서 획득되는 변이지도에 따라 그 영역을 달리하여, 신뢰도 높은 변이지도를 생성한다. 색영역분할(color segmentation)과 유사한 방법이지만, 이전계층의 변이와 색상 성분의 비율로서 영상을 분할한다는 점에서 보다 간단하고 효율적인 방법이다.

본 논문에서는 초기 변이지도로부터 최종 변이 지도를 생성하기 위해서 fine-to-fine의 계층적 스테레오 정합과 유사한 기법을 적용한다. Fine-to-fine 방법은 계층이 증가할 때 마다 정합 창의 크기를 줄이고, 탐색범위를 제한하면서 변이를 추정하는 방법이다¹¹.

기존의 fine-to-fine 방법은 계층이 증가 할수록 사각형 모양의 정합 창 크기를 줄였지만 제안하는 방법에서는 분할 영역을 정합 창으로 사용하고, 계층적으로 분할영역의 크기를 줄여가며 정합하는 방법을 이용한다. 분할 영역은 경계가 일정하지 않기 때문에 영상의 경계를 충분히 표현할 수 있는 장점이 있다. 각 분할영역의 크기를 매 계층에서 반으로 줄이면서 변이를 추정하게 된다. 분할 영역의 크기가 7×7 보다 작게 되면 더 이상 분할을 하지 않는다. 이전 계층의 변이를 중심으로 제한된 범위까지 탐색을 시작하고 계층이 증가할수록 탐색 범위를 이전 계층의 반으로 줄여 나간다.

2.2.3 후처리

계층적 스테레오 정합은 폐색영역에 대해서는 수행을 하지 않는다. 따라서 계층적 스테레오 정합을 하고 나면 폐색영역이 남아있게 된다. 본 연구에서는 폐색 영역은 배경과 물체의 경계에서 발생하기 때문에 그림 11과 같이 폐색영역이라 판단된 영역은 배경과 유사한 변이를 가진다는 가정하에 배경의 변이를 가지고 채우는 방법을 적용한다.

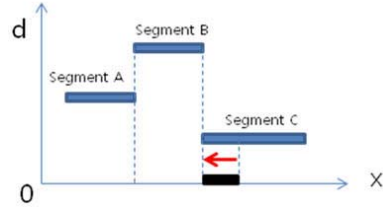


그림 11. 폐색영역 처리
 Fig. 11. Occlusion processing

폐색영역을 처리한 변이 지도는 식 (2)의 cross-checking을 이용하여 신뢰도를 확인하고, 불확실한 변이는 오정합이라 판단한다. 오정합이라 판단된 영역은 식 (3)의 확산 과정을 통하여 변이지도를 재생성한다. 폐색영역 검출, 계층적 스테레오 정합, 후처리의 3 가지 과정은 하나의 계층에서 이루어진다. 최종 계층까지 3 가지 과정을 반복적으로 수행함으로써 신뢰도를 향상시킬 수 있다.

2.2.4 최종 변이지도 생성

계층적 스테레오 정합이 완료되면 최종 변이 지도를 생성한다. 이 과정은 같은 색상 영역에서의 변이 값을 평활화하는 작업이다. 일반적인 2차원 평균 필터를 적용하게 되면 마스크내부에 변이의 경계가 포함되기 때문에 유사한 변이를 가지는 영역뿐만 아니라 변이의 경계 또한 흐려진다. 따라서 본 연구에서는 영상의 경계가 아닌 유사한 변이를 가지는 영역에 1차원 평균필터를 적용한다. 이때 1x5, 5x1의 수평, 수직 마스크가 사용되었다.

III. 실험 결과 및 고찰

본 연구에서 사용한 실험영상은 Middlebury stereo vision(<http://vision.middlebury.edu/stereo/>)의 4개의 실험 영상을 사용하였다. 변이 탐색범위를 제외한 파라미터는 모든 실험 영상에 동일하게 적용하였으며 각 영상의 특성은 표 1과 같다.

그림 12는 본 연구를 통해 얻은 각 실험 영상에 대한 최종 변이지도와 동적 계획법(Dynamic Programming) 및 그래프-컷(Graph-cuts) 알고리즘을 이용하여 획득된 변이 지도를 나타낸 그림이다. 생성된 최종 변이 지도는 블록 매칭(Block Matching), 동적 계획법과 같은 방법에 비하여 변이 지도의 경계가 뚜렷하다는 특징이 있다.

블록 매칭의 경우 정합 창의 크기에 따라 다른 결과가 발생하고, 많은 영역에서 부정확성을 나타

표 1. 실험 영상의 특성
Table 1. Characteristics of test images

| Test Image | Size | Disparity Range |
|------------|---------|-----------------|
| Cones | 450x375 | 0-59 |
| Teddy | 450x375 | 0-59 |
| Tsukuba | 384x288 | 0-15 |
| Venus | 434x384 | 0-19 |

낸다. 동적 계획법은 가려진 영역이 발생하는 영상의 경계에서 변이가 전파되는 현상 때문에 변이의 불연속점구별이 뚜렷하지 않다. 반면 제안한 방법에 의한 결과는 변이의 불연속점에서 신뢰도 높은 변이를 제공한다. 대부분의 디지털 영상은 경계가 부정확하다. 변이지도의 불연속점이 부정확한 변이가 나타나는 이유는 이러한 영상의 부정확한 경계에 기인한다. 본 논문에서는 정확한 변이의 불연속점을 찾기 위해 각 색상 성분으로 영상을 분할하여 계층

적 스테레오 정합 기법으로 최종변이를 생성하였다. 또한 정확한 폐색 영역 검출을 통해 변이의 불연속점에서 발생할 수 있는 부정확한 변이의 전파를 차단하여 신뢰도 높은 변이 지도를 생성하였다.

부정확한 변이는 4개의 실험 영상에서 평균 11%를 나타내었고, 평균 부정확한 변이에 대한 에러율은 성능평가표에 함께 제공된다. 제안한 방법에 의한 변이지도와 Ground truth의 비교에서 평균적으로 3~5의 적은 값의 절대오차가 발생하였다. 또한 Dynamic programming, Graph-cuts 및 PhaseBased보다 우수한 성능을 나타내었다^{4,5}. 반면, 현재 가장 우수하게 평가 받는 신뢰확산에 기반한 방법들에 비해 낮은 성능을 나타내었다. 이러한 방법들은 복잡한 색 영역 분할 방법과 MRF(markov random fields)에 기반한 확률적 접근을 통해 정교한 변이를 획득한다⁶. 이러한 방법들은 결과가 정교한 반면, 변이를 획득하는 과정의 복잡도가 높다는 단점이

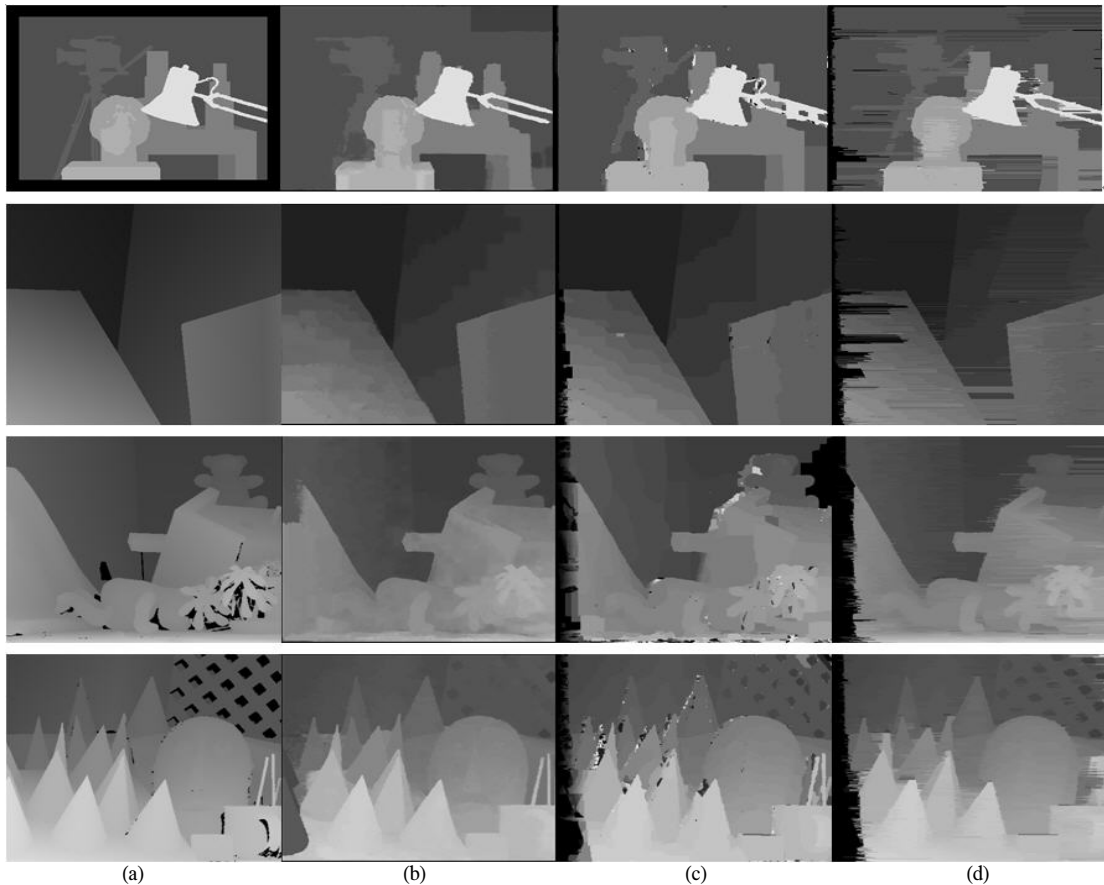


그림 12. 각 실험 영상의 최종 변이지도 (a) Ground Truth, (b) Proposed method, (c) Graph-cuts^[4], (d) Dynamic programming^[4]
Fig. 12. Final disparity map for each experimental image (a) Ground Truth, (b) Proposed method, (c) Graph-cuts^[4], (d) Dynamic programming^[4]

표 2. Middlebury Stereo 성능 평가표
Table 2. Middlebury Stereo Evaluation

| Algorithm | Avg. | Tsukuba sound truth | | | Venus sound truth | | | Teddy sound truth | | | Cones sound truth | | |
|--------------------|------|------------------------|--------|------|----------------------|--------|------|----------------------|--------|------|----------------------|--------|------|
| | | Rank | nonocc | all | disc | nonocc | all | disc | nonocc | all | disc | nonocc | all |
| GenModel [20] | 30.5 | 2.57 | 4.74 | 13.0 | 1.72 | 3.08 | 16.9 | 6.86 | 15.0 | 19.2 | 4.64 | 14.9 | 11.4 |
| TensorVoting [9] | 31.9 | 3.73 | 4.79 | 8.86 | 1.23 | 1.88 | 11.5 | 9.76 | 17.0 | 24.0 | 4.36 | 11.4 | 12.2 |
| RealTimeGPU [14] | 31.5 | 2.05 | 4.22 | 10.6 | 1.92 | 2.98 | 20.3 | 7.23 | 14.4 | 17.6 | 6.41 | 13.7 | 16.5 |
| CostRelax [11] | 32.9 | 4.76 | 6.68 | 20.3 | 1.41 | 2.48 | 18.5 | 8.16 | 16.9 | 23.8 | 3.91 | 10.2 | 11.8 |
| ReliabilityCP [13] | 33.9 | 1.26 | 3.39 | 7.25 | 2.35 | 3.48 | 12.2 | 9.82 | 16.9 | 19.5 | 32.9 | 19.9 | 19.7 |
| TypeCP [8] | 36.1 | 1.92 | 2.84 | 9.96 | 1.41 | 2.19 | 7.74 | 16.9 | 23.9 | 27.1 | 10.0 | 10.3 | 19.9 |
| YOUR METHOD | 36.4 | 2.02 | 2.75 | 10.5 | 1.75 | 2.32 | 9.97 | 15.8 | 22.4 | 31.0 | 10.1 | 16.3 | 17.4 |
| GC [16] | 36.0 | 1.94 | 4.12 | 9.39 | 1.79 | 3.44 | 8.75 | 16.5 | 26.0 | 24.9 | 7.70 | 18.2 | 16.3 |
| BP+M.H [40] | 36.3 | 4.12 | 6.34 | 14.6 | 1.96 | 3.31 | 16.8 | 10.2 | 18.9 | 24.0 | 4.93 | 16.5 | 12.3 |
| DF [18] | 40.6 | 4.12 | 5.94 | 12.0 | 10.1 | 11.0 | 21.0 | 14.0 | 21.6 | 20.6 | 10.5 | 19.1 | 21.1 |
| PhaseBased [31] | 42.5 | 4.26 | 6.53 | 15.4 | 6.71 | 8.16 | 26.4 | 14.5 | 23.1 | 25.5 | 10.8 | 20.5 | 21.2 |
| RegionalSup [38] | 43.8 | 3.59 | 6.05 | 14.2 | 8.14 | 9.68 | 36.8 | 19.3 | 26.7 | 32.1 | 2.16 | 19.3 | 19.9 |
| SSD+MF [1a] | 43.0 | 5.22 | 7.07 | 24.1 | 3.24 | 5.16 | 11.9 | 19.5 | 24.8 | 32.9 | 10.6 | 19.8 | 26.3 |
| STICA [15a] | 44.3 | 7.70 | 9.63 | 27.8 | 8.19 | 9.58 | 40.3 | 15.8 | 23.2 | 37.7 | 9.80 | 17.8 | 28.7 |
| SO [16a] | 44.9 | 5.08 | 7.22 | 12.2 | 9.44 | 10.9 | 21.9 | 19.8 | 28.2 | 26.3 | 13.0 | 22.8 | 22.3 |
| PhaseDiff [23] | 45.8 | 4.83 | 7.11 | 16.3 | 8.34 | 9.76 | 26.0 | 20.0 | 28.0 | 29.0 | 19.8 | 28.5 | 27.5 |
| Infection [10] | 46.0 | 7.95 | 9.54 | 28.9 | 4.41 | 5.53 | 31.7 | 17.7 | 26.1 | 44.4 | 14.3 | 21.3 | 38.0 |

있다. 본 논문에서 제안한 방법은 이러한 신뢰확산에 기반한 방법들보다 성능이 낮지만 복잡도가 낮은 장점이 있고, 기존의 계층적 스테레오 정합 방법보다도 성능이 우수하다는 장점이 있다¹¹⁾.

표 2는 Middlebury stereo 평가 프로그램을 이용해서 성능을 평가한 자료이다. 본 논문에서 제안한 방법은 ‘YOUR METHOD’로 표시되어 있다. nonocc는 폐색영역을 제외한 영역에 대한 어려움을 나타내고, disc는 변이지도의 불연속점인 경계부분에 대한 어려움을 나타낸다. all은 모든 영역에 대한 어려움을 나타낸다. 제안한 방법의 결과 중 ‘Venus’영상의 어려움이 가장 낮은 것을 알 수 있다. 4가지 실험영상 결과의 순위만을 비교하였을 때는 ‘Tsukuba’영상의 전체적인 성능평가 순위가 가장 높았고, 변이의 불연속점에서의 ‘Teddy’영상이 가장 낮은 순위를 기록하였다. ‘Teddy’영상이 성능이 낮게 나온 이유는 영상이 복잡하여 초기 변이지도에서 발생한 부정확한 변이가 매 계층마다 전파되어 나온 결과라고 분석한다.

IV. 결론

본 논문에서는 영상의 컬러 정보를 이용한 계층적 스테레오 정합 기법을 제안하였다. 영상의 특징점에서 정합 창을 이용하여 초기 변이지도를 생성하였고, 생성된 초기 변이지도를 이용하여 계층적 스테레오 정합을 통해 변이지도를 생성하였다. 계층적 스테레오 정합 기법에서 존재할 수 있는 오정합에 의한 에러전파를 사전에 차단하고, 매 계층에서 정확한 폐색영역 검출로 폐색영역에 의한 부정확한 변이의 확산을 방지하여 신뢰도 높은 변이지도를 생성하였다. 특히, 일반적인 변이지도에서 부

정확한 변이를 나타내는 변이의 불연속점에서 신뢰도 높은 변이지도를 생성하여 객관적인 평가를 하였다. 평가 결과 부정확한 변이는 4개의 실험 영상에서 평균 11%를 나타내었고, Dynamic programming, Graph-cuts 및 PhaseBased 보다 우수한 성능을 나타내었다^{4,5)}. 하지만 영상의 색이 균일한 영역에서 변이가 불분명 되는 현상이 나타났다. 이는 계층적 스테레오 정합 기법에 의해 나타나는 현상으로, 앞으로 개선해야 할 문제점이다. 최종변이를 예측할 때 주변의 신뢰도 높은 변이와 변이의 분포 패턴분석을 통해 보다 신뢰도를 향상시킨 알고리즘으로 해결이 가능할 것이다.

제안한 방법에 의한 신뢰도 높은 변이 지도를 이용하면 중간 시점 영상을 생성하여 보다 자연스러운 입체영상을 다시점 모니터에서 시청 할 수 있으며, 로봇연구에 활용되어 관심영역을 획득하여 실제 거리를 측정하여 정밀한 로봇의 움직임을 제어할 수 있는 연구에 활용이 가능하다. 뿐만 아니라, 의료, 보안, 안보등의 사회 전반적인 부분에 활용이 될 것으로 기대된다.

참고 문헌

- [1] Jong-Il Park, Seiki Inoue, “Hierarchical Depth Mapping from Multiple Cameras,” Proceedings of the 9th International Conference on Image Analysis and Processing, Florence, vol. 1310, pp. 685-692, 1997.
- [2] Cochran, S.D., Medioni, G., “3-D surface description from binocular stereo,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol. 14, Issue 10, Page(s):981 - 994, Oct. 1992.
- [3] Y. Fisher, “Fractal Image Compression,” Springer-verlage, New York, Inc. pp. 251 ~263, 1995.
- [4] Scharstein, D., Szeliski, R., Zabih, R., “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” Stereo and Multi-Baseline Vision, 2001. (SMBV 2001). Proceedings. IEEE Workshop on 9-10 Dec. 2001 Page(s):131 - 140.
- [5] S. El-Etriby, A. Al-Hamadi, and B. Michaelis., “Dense stereo correspondence with slanted surface using phase-based algorithm,” Industrial

Electronics, 2007. IEEE International Symposium on 4-7 June 2007 Page(s):1807-1813.

[6] A. Klaus, M. Sormann, K. Karner, "Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure," ICPR 2006. 18th International Conference on Volume3, 0-00 Page(s):15-18.

[7] Michael Bleyer and Margrit Gelautz, "A Layered Stereo Algorithm using Segmentation and Global Visibility Constraints," ICIP '04. Page(s):2997 - 3000 Vol. 5.

[8] Brown, M.Z., Burschka, D., Hager, G.D., "Advances in computational stereo," Pattern Analysis and Machine Intelligence, IEEE Transactions on Vol. 25, Issue 8, Page(s):993 - 1008, Aug. 2003.

[9] Fusiello, A., Roberto, V., Trucco, E., "Efficient stereo with multiple windowing," Computer Vision and Pattern Recognition, 1997. Proceedings, 1997 IEEE Computer Society Conference on 17-19 June 1997 Page(s):858 - 863.

[10] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 11, pp. 1222-1239, Nov. 2001.

[11] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient Belief Propagation for Early Vision," Proc. IEEE Conf. Computer Vision and Pattern Recognition, June 2004.

[12] Kuk-Jin Yoon, In-So Kweon, "Locally adaptive support-weight approach for visual correspondence search," Computer Vision and Pattern Recognition, 2005. IEEE Computer Society Conference on Volume2, Page(s):924 - 931, 20-25 June 2005.

[13] O. Veksler, "Fast Variable Window for Stereo Correspondence using Integral Images," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 556 - 561, 2003.

[14] R. C. Bolles, H. H. Baker, and M. J. Hannah, "The JISCT stereo evaluation," In DARPA Image Understanding Workshop, pages 263-274, 1993

[15] J. F. Canny, "A Computational approach to

edge detection," IEEE TPAMI, 8(6):34-43, 1986

[16] D. Terzopoulos, "Regularization of inverse visual problems involving discontinuities," IEEE TPAMI, 8(4):413-424, 1986

김 태 준 (Tae-june Kim)

준회원



2008년 2월 광운대학교 전자공학과 졸업
2008년 3월~현재 광운대학교 전자공학과 석사과정
<관심분야> Stereo Matching, Computer Vision, 동영상 코덱

유 지 상 (Ji-sang Yoo)

정회원



1985년 2월 서울대학교 전자공학과 졸업(공학사)
1987년 2월 서울대학교 대학원 전자공학과 졸업(공학석사)
1993년 5월 Purdue 대학교 전기공학과 졸업(Ph.D.)
1993년 9월~1994년 8월 현대 전자산업(주) 산전연구소 선임연구원
1994년 9월~1997년 8월 한림대학교 전자공학과 조교수
1997년 9월~2001년 8월 광운대학교 전자공학과 조교수
2001년 9월~2006년 8월 광운대학교 전자공학과 부교수
2006년 9월~현재 광운대학교 전자공학과 정교수
<관심분야> 웨이블릿 기반 영상처리, 영상압축, 영상인식, 비선형 신호처리