

# 영상 분할을 이용한 다이내믹 프로그래밍 기반의 스테레오 정합

준회원 서용석\*, 종신회원 유지상\*

## Dynamic Programming-based Stereo Matching Using Image Segmentation

Yong-seok Seo\* Associate Member, Ji-sang Yoo\* Lifelong Member

### 요약

본 논문에서는 영상 분할(image segmentation)을 이용한 다이내믹 프로그래밍(dynamic programming, DP) 기반의 스테레오 정합 (stereo matching) 기법을 제안한다. 다이내믹 프로그래밍은 스테레오 정합을 포함하는 여러 가지 컴퓨터 비전 문제들의 고전적이고 인기가 있는 최적화 방법이다. 그러나 스테레오 정합 문제에 적용될 때 스캔라인들 사이의 수직 상호 관계가 적절하게 고려되지 않기 때문에 기존의 DP의 성능은 만족스럽지 않다. 본 논문에서 제안하는 알고리즘에서는 영상을 분할하여 정확한 경계정보를 획득한 다음 경계 정보에 의거하여 변이의 불연속과 폐색영역을 고려한다. Middlebury 스테레오 영상에 적용한 실험 결과들은 제안된 알고리즘이 이전의 다이내믹 기반 알고리즘보다 더 좋은 성능을 보여주는 것을 입증해준다.

**Key Words** : Image Segmentation, Dynamic Programming, Stereo Matching

### ABSTRACT

In this paper, we present a dynamic programming(DP)-based stereo matching method using image segmentation algorithm. DP has been a classical and popular optimization method for various computer vision problems including stereo matching. However, the performance of conventional DP has not been satisfactory when it is applied to the stereo matching since the vertical correlation between scanned lines has not been properly considered. In the proposed algorithm, accurate edge information is first obtained from segmented image information then we considers the discontinuity of disparity and occlusions region based on the obtained edge information. The experimental results applied to the Middlebury stereo images demonstrate that the proposed algorithm has better performances in stereo matching than the previous DP based algorithms.

### I. 서론

최근 초고속 인터넷 망의 확대와 통신 위성의 개발 보급에 힘입어 정보의 디지털화가 가속되면서 방송과 통신기술이 융합되고 그 경계가 점차 모호해지고 있다. 이에 따라 다양한 멀티미디어에 대한

요구가 증가하고 있으며, 각종 방송 서비스에서 시청자의 참여도 매우 활발하다. 증가하는 시청자의 요구를 수용하고 보다 나은 서비스를 제공하기 위해 2012년말까지는 지상파 아날로그 방송이 모두 디지털화 될 예정이며, 그 이후에는 입체 3DTV 등 실감 방송이 서비스될 예정이다. 실감 방송은 다

※ 본 연구는 서울시 산학연 협력사업(NT080528) 지원으로 수행되었음.

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

논문번호 : KICS2010-03-114, 접수일자 : 2010년 3월 22일, 최종논문접수일자 : 2010년 8월 2일

차원 실감미디어를 이용하여 사용자에게 몰입감을 줄 수 있는 방송 서비스이다. 몰입감을 줄 수 있는 하나의 수단으로써 입체감을 들 수 있다. 3차원 입체감을 느끼게 하는 가장 일반적인 방법은 인간의 두 눈이 한 장면을 바라볼 때 생기는 양안 시차(binocular parallax)를 이용하는 스테레오 방식이다. 즉, 좌, 우 두 눈에 서로 다른 영상이 보이게 함으로써 우리가 실생활에서 보는 것과 같은 입체감을 느낄 수 있게 하는 것이다. 그러나 이러한 스테레오 방식은 시점이 고정되어 있기 때문에 여러 방향에서 관찰할 수 있는 자유도가 존재하지 않고, 무안경식의 경우에는 여러 사람이 동시 시청이 불가능하다. 또한 스테레오 영상에서 더욱 확장된 다시점 영상의 경우에도 데이터양의 증가와 시점간의 불연속성으로 인해 3차원 입체 영상 서비스 구현에 많은 어려움이 있다.

이러한 문제점을 보완하기 위해 두 시점의 영상으로부터 임의의 시점에 해당하는 영상을 생성하는 중간시점 영상 합성(IVR : intermediate view reconstruction) 기법에 관한 연구가 활발히 이루어지고 있다. 중간시점 영상 합성 기술을 통해 다시점 영상에서의 효율적인 데이터 전송 및 수신이 가능해지며 시점간의 불연속 문제를 해결할 수 있게 된다. 또한 스테레오 영상에서 시점에 대한 자유도를 제공할 수 있다. 중간시점 영상 합성 시 가장 중요한 과정은 좌, 우 영상에 해당하는 두 시점의 영상이 주어졌을 때 영상 내 정합점을 찾아내는 변이 추정(disparity estimation) 방식이다. 변이 추정 방식은 정합 방법에 따라 크게 국부적 방법(local method)과 전역적 방법(global method)으로 나눌 수 있다<sup>11,12)</sup>. 국부적 방법은 비교적 연산량이 적고 구현이 쉬운 반면 가려진 영역(occlusion)이나 텍스처 정보가 부족한 영역(textureless)에서 오정합 발생 확률이 매우 높다. 또한 정합창 사용으로 인한 경계 영역에서의 부정확한 변이가 할당되는 단점으로 인해 정확성 보다는 실시간이나 고속의 프로세스가 요구되는 분야에 사용되고 있다. 국부적 방법의 가장 대표적인 알고리즘은 한 영상의 작은 영역에 대하여 상관성이 가장 높은 영역이나 여러 값이 가장 낮은 영역을 참조 영상에서 탐색하는 블록매칭 방식이다. 또한 특징기반 매칭방식은 단순히 화소의 밝기값을 이용하지 않고 모서리(edge), 구석(coner), 골(cave) 등 영상이 가지는 특징을 이용한다. 이외에도 화소의 밝기값 변화와 잡음 등에 취약한 블록 매칭 방식의 단점을 보완한 영상의 경사도를 이용한 방식이 있

다. 이에 반해 전역적 방법은 텍스처 정보가 부족한 영역이나 가려진 영역에서의 오정합 발생을 줄이기 위해 비국부적 제약조건을 활용하는 방식이다. 이러한 전역적 방법은 정확한 변이를 추정할 수 있는 반면 계산 비용이 크다는 단점이 있으며 대표적인 방법으로는 그래프 컷(graph cut), 신뢰확산(belief propagation), 다이내믹 프로그래밍(dynamic programming) 기법 등이 있다. 그래프 컷은 컴퓨터 네트워크 흐름제어에서 출발하였으며 최대-흐름/최소-컷(max-flow/min-cut) 알고리즘을 이용하여 변이를 할당하는 방식으로 변이 추정에 가장 좋은 성능을 내는 방법 중 하나이다. 신뢰확산은 메시지 전달에 기반을 둔 반복 알고리즘으로 성능은 우수한 반면 완전히 수렴된 메시지를 얻기 위해서 많은 반복이 필요한 단점이 있다<sup>13)</sup>. 다이내믹 프로그래밍은 본래 어떠한 문제를 해결하는데 있어 최적화 문제에 대한 계산상의 복잡성을 줄이기 제안된 수학적 방법이다. 다이내믹 프로그래밍 기반 변이 추정 방식이란 기준영상과 참조영상 사이의 관계를 이용하여 생성한 변이 공간 영상(DSI : disparity space image)으로부터 최적의 변이 경로를 찾아가는 방식을 말한다<sup>14,15)</sup>. 이러한 다이내믹 프로그래밍 기반 변이추정 방식은 화소 단위의 정밀한 변이 맵을 얻을 수 있으며, 특징점 추출에 의한 오류를 막으면서 다른 전역적 방식들에 비해 수행 속도가 빠른 장점을 가지고 있다.

본 논문에서는 영상의 컬러 정보를 이용한 다이내믹 프로그래밍 기반의 스테레오 정합 기법을 제안한다. 다이내믹 프로그래밍의 단점인 스트리킹(streaking)을 해결하기 위해 영상 분할을 이용하여 최적의 변이값을 할당 받을 수 있도록 하였다. 먼저 정합점과 유사한 컬러 정보를 갖는 영역을 정합창으로 사용하여 DSI를 생성하고 영상의 경계를 추출하여 경계 정보를 획득한다. 다음 경계 지점에서 주변 화소값을 고려한 비용 행렬을 생성하고 여러 가지 제한 조건이 결합된 최적 변이 경로를 설정함으로써 변이 지도(disparity map)를 생성하였다. 제안하는 기법은 일반적인 다이내믹 프로그래밍이 가진 장점을 살리면서 생성된 변이 지도에 대해 적절한 후처리 과정으로 오류를 최소화함으로써 기존의 다이내믹 프로그래밍 기반의 스테레오 정합 알고리즘들보다 우수한 결과를 보여준다.

본 논문의 구성은 다음과 같다. 제 II장에서는 기존의 다이내믹 프로그래밍기반의 변이 추정 방식에 대하여 알아보고 제 III장에서는 제안된 변이 추정방식에 대하여 기술한다. 제 IV장에서는 실험 및

결과를 분석하고 제 V장의 결론을 끝으로 논문을 마친다.

## II. 다이내믹 프로그래밍

그림 1은 기본적인 다이내믹 프로그래밍 기반의 변이 추정 기법을 나타낸다. 영상의 각 라인에 대하여 좌·우 영상의 밝기값 차이를 나타내는 변이 공간 영상 생성 후, 적절한 비용 값을 이용하여 비용 행렬을 생성한다. 마지막으로 최적의 경로를 탐색함으로써 변이를 할당하게 된다<sup>6)</sup>.

### 2.1 변이 공간 영상(Disparity Space Image)

변이 공간 영상 생성은 다이내믹 프로그래밍 기반의 변이 추정 시 가장 먼저 수행되는 과정으로 이후의 비용행렬 생성과 최적경로 탐색에 영향을 주게 된다. 변이 공간 영상은 좌·우 영상의 n번째 스캔라인에 대한 정합 오류를 나타내는 것으로 그림 2에 좌·우영상의 100번째 스캔라인에 대해 생성된 변이공간 영상을 나타내었다.

생성된 변이 공간 영상은 변이를 구하고자 하는 영상과 참조되는 영상의 n번째 스캔라인을 두 축으로, 대응되는 좌표의 정합오류를 나타내며 영상내의

한 점은 식 (1)과 같이 나타낼 수 있다.

$$DSI_n(x_L, x_R) = \frac{1}{N} \sum_{i,j=-\frac{W}{2}}^{\frac{W}{2}} |I_L(x_L+i, n+j) - I_R(x_R+i, n+j)| \quad (1)$$

여기서  $I_L$ 은 좌영상,  $I_R$ 은 우영상을 나타내며  $W$ 는 정합창의 크기,  $N$ 은 정합시 사용된 블록내 화소 수를 나타낸다.  $x_L$ 과  $x_R$ 은 각각 n번째 스캔라인에서 좌영상, 우영상의 좌표를 나타낸다.

식 (1)을 통하여 변이 공간 영상 내에 어두운 부분은 정합 오류가 적고 밝은 부분일수록 정합 오류가 크다는 것을 알 수 있다. 변이 공간 영상은 그 특징에 따라 다음과 같이 3개의 영역으로 나뉜다.

#### 2.1.1 수직, 수평경계성분(Vertical, horizontal edge)

변이 공간 영상 내에 수직 경계성분은 좌영상의 경계성분에 해당되는 것으로써 경계 영역에서 화소값의 차이가 크다는 특성이 반영된 것이다. 마찬가지로 수평 경계 성분은 우영상의 경계성분에 해당되며. 이러한 수직, 수평 경계성분은 최적 경로를 탐색하는데 영향을 주게 된다.



그림 1. 다이내믹 프로그래밍 기반의 변이 추정  
Fig. 1. Disparity estimation by dynamic programming

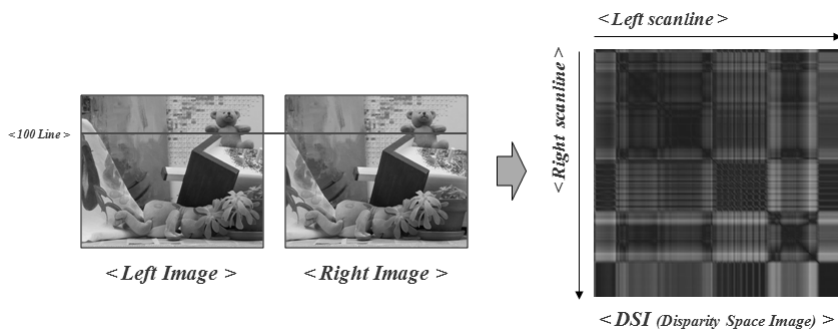


그림 2. 변이 공간 영상의 생성  
Fig. 2. Generation of DSI(disparity space image)

2.1.2 질감 영역 성분 (Texture region)

좌영상과 우영상의 화소변화가 큰 영역이 서로 대응될 때 발생하는 영역으로 변이가 대응되는 좌표를 중심으로 화소의 변화가 크게 나타나기 때문에 변이를 쉽게 찾을 수 있는 영역이다.

2.1.3 평활 영역 성분 (Plain region)

평활 영역은 좌영상과 우영상의 화소변화가 거의 없고 평활한 영역이 서로 대응될 때 발생한다. 대응되는 변이를 중심으로 화소의 차이가 거의 존재하지 않으므로 변이를 찾는 데 어려움이 존재한다.

2.2 비용행렬(Cost Matrix) 생성

비용행렬은 앞서 얻어진 변이 공간 영상을 바탕으로 수직, 수평 경계성분을 고려한 비용값과 변이의 연속성을 고려한 비용값을 이용하여 생성하게 된다. 단순한 변이 공간 영상에서 최적의 경로를 찾게 될 경우 일반적인 화소기반의 변이 추정 방식과 크게 다르지 않게 되므로 비용행렬의 생성은 다이내믹 프로그래밍 기반의 변이 추정 방식에 특별성을 부여한다. 비용행렬 생성 시 고려되는 3가지 비용값은 식 (2)와 같다.

$$\begin{aligned}
 C_{plane}(x, y) &= DSI(x, y) + Cost(x-1, y-1) \\
 C_{horizontal}(x, y) &= P_h + Cost(x-1, y) \\
 C_{vertical}(x, y) &= P_v + Cost(x, y-1)
 \end{aligned}
 \tag{2}$$

여기서  $C_{plane}$ 은 현재 화소의 변이가 이전 화소와 연속적인 변이값을 가질 때를 고려한 비용이며  $C_{horizontal}$ 과  $C_{vertical}$ 은 각각 변이 공간 영상 내에 수평 경계성분 및 수직 경계성분을 고려한 비용이다.  $P_h$ 와  $P_v$ 는 수평, 수직 경계성분을 고려한 오프셋 값으로 가려진 영역을 판별하는데 영향을 미치게 된다. 위 3가지 값의 비교를 통한 최종적인 비용행렬 생성은 식 (3)과 같다.

$$Cost(x, y) = \min(C_{plane}, C_{horizontal}, C_{vertical})
 \tag{3}$$

2.3 최적 경로 탐색(Optimal Path-Search)

최적 경로 탐색은 각각의 라인에 대해 생성된 비용행렬을 이용하여 최소 비용이 소요되는 경로를 찾는 과정이다. 목표점에서부터 시작점까지의 최소 비용 경로를 탐색하면서 최종적으로 변이를 할당하게 된다. 그림 3은 최적 경로 탐색을 통하여 변이

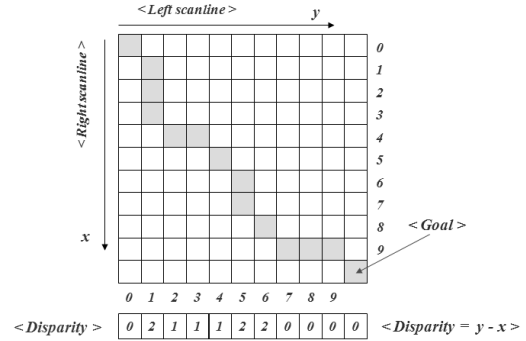


그림 3. 최적 경로 탐색 및 변이 할당  
Fig. 3. Optimal path-search and disparity allocation

가 할당되는 과정을 나타낸다.

III. 제안하는 변이 추정 기법

스테레오 정합 방법 중에서 영역기반의 정합 방법으로 다단계 결정과정에서 최적성의 원리를 이용하고 폐색 영역을 찾는 데 유리한 동적 계획법 (dynamic programming)을 많이 적용하고 있다<sup>[7]</sup>. 다른 정합 방법들과 비교했을 때 유일성과 순차성 제약 때문에 계산상의 복잡도가 적은 장점이 있지만 좁은 폐색 영역을 갖는 영상에 대해서는 비교적 좋은 결과를 얻지 못하는 단점이 있다<sup>[8]</sup>. 본 논문에서는 비용행렬 생성 시, 스트리킹(streaking)의 원인이 되는 연속적인 경로 탐색에 대한 신뢰성을 검증하기 위한 알고리즘을 제안하였다. 이전에도 신뢰성을 검증하기 위한 방법은 제안되었으나, 본 논문에서는 보다 효율성을 높이기 위하여 영상의 특징 정보를 이용하여 연속적인 경로 탐색에 대한 신뢰성을 확인하여 최적의 변이 값을 할당 받을 수 있도록 하였다. 영상의 특징 정보로는 영상의 경계를 사용하였다. 그림 4는 제안하는 스테레오 정합 과정의 전체 구조를 나타낸다<sup>[9]</sup>.

영상은 일반적으로 모서리, 경계 등에서 특징 점이 존재하며, 이러한 특징 점을 추출하기 위한 방법은 다양하다. 본 논문에서는 Mean-Shift 알고리즘을 이용하여 영상을 분할하고 경계 정보를 획득하여 이를 특징 점으로 이용한다. Mean-Shift 알고리즘은 어느 공간상에 존재하는 점들의 밀도가 높은 곳에서 군집화가 이뤄지도록 하는 방법이다. 즉, 점들의 경사도(gradient)가 큰 쪽으로 이동을 해 가면서 국부최대가 되는 지점(그 점에서의 경사도는 0이다)에서 군집화를 이루게 되는 것이다.

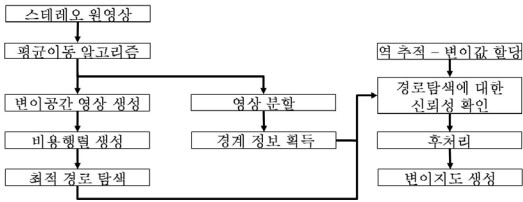


그림 4. 제안하는 스테레오 정합 알고리즘  
Fig. 4. Proposed algorithm for stereo matching

$$m_h(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \quad (4)$$

여기서, n개의 데이터 점  $x_i(i=1,\dots,n)$ 이 되고, d-차원 공간  $R^d$ 이다. 그리고  $h$ 는 윈도우의 반지름이고  $x$ 는 그 윈도우의 중심이다. 식 (4)를 Mean-Shift 벡터라고 부르며, 항상 밀도가 최대로 증가하는 방향을 가리키고 있다. 이 벡터가 가리키는 방향으로 윈도우를 이동하면 국부 최대 밀도에 수렴해 가게 된다. 그림 5-(a)의 “Teddy” 영상에 Mean-Shift 알고리즘을 적용한 그림 5-(b)의 결과 영상을 보면 영상의 경계는 보존되면서 공간 영역에서의 화소값을 균일하게 하는 효과를 확인할 수 있다.

Mean-Shift 알고리즘을 적용한 영상에서 정합점과 유사한 컬러 정보를 갖는 영역을 정합창으로 사용하여 DSI를 생성한다. 일반적으로 다이내믹 프로그래밍에서 사용되는 그림 6-(a)와 같은 3x3, 5x5 등의 일정 크기의 정합창 대신 그림 6-(b)처럼 정합창 내의 중심이 되는 정합점과 유사한 컬러 정보를 갖는 영역을 정합창으로 사용한다.

정합점을 찾을 때 정합창의 크기는 매우 중요하다. 형식적으로 3x3, 5x5, 7x7, 9x9를 주로 사용하며 창 크기가 너무 작으면 잡음에 민감하고 반면

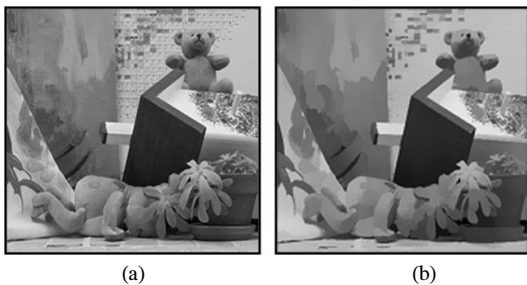


그림 5. (a) “Teddy” 영상, (b) Mean-Shift Algorithm을 적용한 결과 영상  
Fig. 5. (a) “Teddy” image (b) Result image using Mean-Shift Algorithm

에 너무 크면 일종의 저역 필터(low-pass filter)의 역할을 하고 깊이 불연속점을 쉽게 놓치게 된다. 정합창의 크기가 커질수록 변이 지도는 세밀하지만 연산 속도가 증가하고, 정합창의 크기가 작아질수록 변이 지도는 세밀하지 못하지만 연산 속도가 감소하므로 적당한 정합창을 설정하는 것이 중요하다. 실험을 하여 결과를 비교했을 때 5x5의 정합창 성능이 가장 우수하다. 제안한 정합창을 사용한 SAD (Sum of Absolute Difference)의 비용은 아래 다음과 같은 식 (5)가 된다.

$$C_{SD}(x_r, x_l) = \sum_{m=-W}^W \sum_{n=-W}^W |I_r(x_r+m, s+n) - I_l(x_l+m, s+n)|, (x_r, s) \in R \quad (5)$$

식 (5)에서  $x$ 는  $x$ 축에서의 위치,  $s$ 는  $s$ 번째 스캔라인을 나타내고,  $I(x, s)$ 는 그 위치에서의 화소값을 나타내며, 아래첨자  $l, r$ 은 각각 좌 영상, 우 영상을 나타낸다.  $W$ 은 사용된 정합창의 크기,  $R$ 은 분할된 영역을 나타낸다. 그림 6-(a),(b)의 검은색 점을 정합점이라고 했을 때 이 점과 근접한 위치에 있는 유사한 컬러 정보를 가지는 영역을 정합창 크기  $W$  안에서 찾고 거기에 해당하는 영역에 대해서만 SAD 비용 값을 구하게 된다. 정합창 전체에 대해 이루어지는 기존의 계산법에서 나타날 수 있는 정합창의 크기 변화에 의해 생기는 영향을 줄일 수 있다.

DSI를 생성한 후 경계 지점에서 주변 화소값을 고려한 비용 행렬을 생성하기 위하여 영상을 분할하여 경계 정보를 획득한다. 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

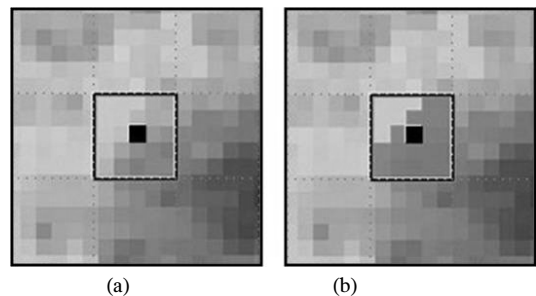


그림 6. (a) 5x5 정합창, (b) 제안한 정합창  
Fig. 6. (a) 5x5 window, (b) Proposed window

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 사용하기에 충분한 신뢰성을 가진다고 할 수 있다. 앞에서 정합점과 유사한 컬러 정보를 갖는 영역을 정합창으로 정하는 방법과 같이 영상 전체를 R,G,B 값을 고려하여 정해진 범위만큼 분할하며 이때 한 번 분할된 영역은 또 다시 분할되지 않도록 제외시킨다. 그림 7-(a)는 “Teddy” 영상에서 곱인형 부분을 나타내며 그림 7-(b)는 그 영상을 유사한 컬러 정보를 갖는 영역별로 분할하고 랜덤 함수를 사용하여 화소값을 할당한 결과이다. 이때 부정확한 영상 분할이 일어났을 경우에는 신뢰도 높은 변이를 예측할 수 없다.

영상을 분할하여 획득한 경계정보를 이용하여 경계 지점에서 주변 화소값을 고려한 비용 행렬을 생성한 후, 최적 경로 탐색을 하여 변이값을 할당한다. 이때 부정확한 변이를 검출하기 위해서 획득된 변이를 참조영상에 적용해 원영상과의 오차를 판단하는 역 추적(cross-checking)방법을 이용한다. 식 (6)은 역 추적을 이용하여 부정확한 변이를 검출하는 식을 나타낸다<sup>[10]</sup>.

$$CrossChecking = |R(x + Dis(x, y), y) - L(x, y)|$$

$$Dis(x, y) = \begin{cases} Dis(x, y), & CrossChecking < Tolerance \\ 0, & else \end{cases} \quad (6)$$

여기서  $R(x, y)$ 는 우영상의 화소값,  $L(x, y)$ 는 좌영상의 화소값을 나타내고,  $Dis(x, y)$ 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 영

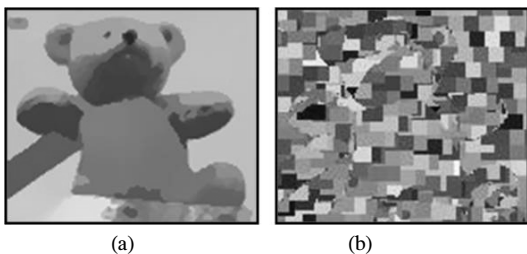


그림 7. (a) “Teddy” 영상, (b) 영상 분할  
Fig. 7. (a) “Teddy” image, (b) image segmentation

상이 획득될 때 조명 등의 주변 환경과 카메라의 특성에 의해 차이 값은 ‘0’이 되지 않는다. 이러한 오차의 영향을 줄이기 위해 허용오차를 설정한다. 역 추적 값이 허용오차 이내에 있다면 신뢰도 높은 변이라 판단하고, 그렇지 않다면 부정확한 변이라 판단하여 변이를 ‘0’으로 초기화 시켜 오정합이라 판단한다<sup>[11]</sup>.

역 추적과정에서 허용오차가 정해 놓은 문턱치 이상인 경우, 그 위치의 화소들은 다른 객체간의 화소라고 인정하고 식 (7)과 (8)을 이용하여 올바른 경로에 대한 신뢰성을 확인하여 변이 정보의 탐색을 실시한다. 현재 지점이 연속적인 경로에 해당할 때, 화소 위치를 확인하여 그 지점이 경계에 해당하는지를 점검한다. 만일 그 화소의 위치가 경계 지점이라면 연속적인 경로를 인정하고, 경계 지점이 아니라면 연속적인 경로에 대해 재탐색을 한다.  $I(i, j)$ 는 원영상의 화소값을 말하며,  $k$ 는 정해놓은 문턱치 이상이면 부여되는 패널티 값이다.

$$DSI(i, j) = Penalty \times DSI(i, j) \quad (7)$$

$$Penalty = \begin{cases} 0 & : |I(i, j) - I(i-1, j)| = 0 \\ 1 & : |I(i, j) - I(i-1, j)| < threshold \\ k & : |I(i, j) - I(i-1, j)| > threshold \end{cases} \quad (8)$$

신뢰도 높은 변이 지도를 생성하기 위해서는 오정합 변이 검출이 중요하다. 폐색 영역에서 오정합 변이가 많이 나타나므로 연속적으로 오정합 변이가 할당되는 것을 방지하기 위해 폐색 영역 검출이 필요하다. 동적 계획법은 보통 유일성(uniqueness) 과 순차성(ordering) 제약을 사용하기 때문에 잘못된 시차는 연속적인 에러를 일으키게 된다. 폐색 영역에서 오정합 변이가 많이 나타나는 이유는 기준 영상에는 존재하지만 원 영상에는 존재하지 않는 유사한 영역을 찾아 변이를 폐색 영역에 할당하기 때문이다. 오정합 변이 검출을 위해서 획득한 변이 지도를 분할된 영역별로 평가한다. 각 분할된 영역은 비슷한 변이를 가지며, 오직 영역의 경계에서만 변이의 변화가 커지므로 변이 지도를 분할된 영역에 따라 평가하기 위해서 우리는 분할된 영역에 대한 분산 값  $var$ 을 계산하고 일정한 임계값  $\alpha$ 에 따라 그 영역의 옳고 그름을 판단한다.

$$\begin{cases} valid & var < \alpha \\ invalid & var \geq \alpha \end{cases} \quad (9)$$

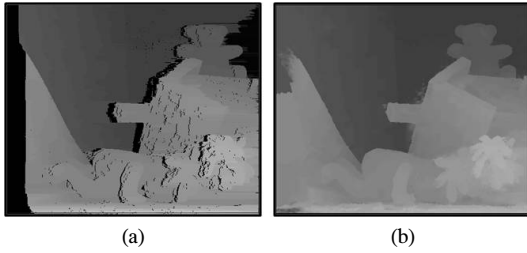


그림 8. (a) 폐색 영역, (b) 최종 변이 지도  
Fig. 8. (a) Occlusion region, (b) Final disparity map

일반적으로, 비슷한 변이를 갖는 영역의 분산 값은 작을 것이고, 그렇지 않은 영역, 즉 변이의 변화가 큰 영역은 큰 분산 값을 가질 것이다. 식 (9)은 영역별로 이루어지며 이 과정을 통해 우리는 올바른 정합영역과 잘못된 영역을 임계값  $\alpha$ 에 따라 정합 영역과 오정합 영역으로 구분하여 오정합 영역으로 판단된 영역은 주변 정보와 경계 정보를 고려하여 가장 신뢰도 높은 변이를 할당한다. 그림 8(a)의 검은색 영역은 폐색 영역을 나타내고 그림 8(b)는 후처리한 최종 변이 지도이다.

#### IV. 실험 결과 및 고찰

스테레오 정합의 성능을 평가하기 위하여 본 실험에 사용된 데이터는 Middlebury stereo vision (<http://vision.middlebury.edu/stereo/>)의 4개의 실험 영상을 사용하였다. “Teddy”, “Cones”, “Venus”, “Tsukuba”로써 9개의 카메라로부터 획득된 다시점 영상이다. 실험에 사용된 파라미터 값들은 모든 영상에 공통적으로 적용하여 객관적인 결과가 도출되도록 하였다. 각 영상의 특성은 표 1과 같다.

그림 9는 실험영상과 본 연구를 통해 얻은 각 실험 영상에 대한 최종 변이지도와 기존의 다이내믹 기반의 스테레오 정합 알고리즘들의 최종 변이지도이다.

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

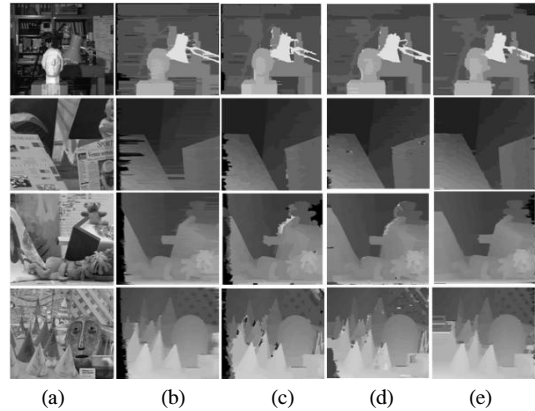


그림 9. (a) 스테레오 영상, (b) DP, (c) Tree DP, (d) Reliability DP, (e) Proposed  
Fig. 9. (a) Stereo image, (b) DP, (c) Tree DP, (d) Reliability DP, (e) Proposed

표 2는 Middlebury stereo vision에서 ground truth map과 비교한 에러율이다. Middlebury stereo vision에서 정의해 놓은 에러율은 가려진 영역을 제외한 에러율(nonocc)과 불연속지점에서의 에러율(disc), 그리고 앞의 두 가지를 모두 포함한 전체 에러율(all)로 분류되어 있다.

표 2를 보면 4개의 실험 영상에 대한 평균 에러율이 각각 DP 기법은 14.18%, Tree DP 기법<sup>[12]</sup>은 11.67%, Reliability DP 기법<sup>[13]</sup>은 10.72%이며 제안한 기법이 10.39%로 가장 적다. 다이내믹 프로그래

표 2. 변이 지도들의 에러율  
Table 2. Error rates(%) of the disparity maps

Algorithm	Tsukuba			Venus			Teddy			Cones		
	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc
DP	4.12	5.04	12.0	10.1	11.0	21.0	14.0	21.6	20.6	10.5	19.1	21.1
Tree DP	1.99	2.84	9.96	1.41	2.10	7.74	15.9	23.9	27.1	10.0	18.3	18.9
Reliability DP	1.36	3.39	7.25	2.35	3.48	12.2	9.82	16.9	19.5	12.9	19.9	19.7
Proposed	1.69	3.63	8.84	2.35	3.19	10.36	8.92	15.7	19.4	12.02	19.4	19.2

밍은 가려진 영역이 발생하는 영상의 경계에서 변이가 전파되는 현상 때문에 변이의 불연속점 구별이 뚜렷하지 않다. 반면 제안하는 기법에 의한 결과는 변이의 불연속점에서 신뢰도 높은 변이를 제공한다. 제안하는 기법은 일반적인 다이나믹 프로그래밍이 가진 장점을 살리면서 생성된 변이 지도에 대해 적절한 후처리 과정으로 오류를 최소화함으로써 기존의 다이나믹 프로그래밍 기반의 스테레오 정합 알고리즘들보다 우수한 결과를 보여준다.

현재 가장 우수하게 평가 받는 신뢰확산에 기반한 방법들에 비해 낮은 성능을 나타내지만 이러한 방법들은 복잡한 색 영역 분할 방법과 MRF(Markov random fields)에 기반한 확률적 접근을 통해 정교한 변이를 획득한다. 이러한 방법들은 결과가 정교한 반면, 변이를 획득하는 과정의 복잡도가 높다는 단점이 있다. 이러한 신뢰확산에 기반한 방법들보다는 성능이 낮지만 복잡도가 낮은 장점이 있고, 기존의 다이나믹 프로그래밍 기반의 스테레오 정합 방법보다도 성능이 우수하다는 장점이 있다

## V. 결 론

본 논문에서는 영상 분할을 이용하여 최적의 변이값을 할당 받을 수 있도록 하였다. Mean-Shift 알고리즘을 적용한 균일화된 영상을 분할하여 경계정보를 획득하고, 정합점과 유사한 컬러 정보를 갖는 영역을 정합창으로 사용하여 DSI를 생성한 다음 경계 지점에서 주변 화소값을 고려한 비용 행렬을 생성하고 여러 가지 제한 조건이 결합된 최적 변이 경로를 설정함으로써 일반적인 스캔라인 사이의 불일치 문제를 해결하였고, 분할된 영역을 이용하여 정확한 변이를 찾아 폐색영역에서의 오류를 최소화하여 향상된 결과를 얻었다. 특히, 배경이나 유사한 화소값이 넓게 분포 되어있는 텍스트 정보가 부족한 영역에서 안정된 변이를 얻어 비교적 적은 연산을 가지고 정확한 결과를 이끌어 낼 수 있었다. 앞으로 이런 복합적인 기법에 대한 연구를 통해 좀더 정확하면서 빠른 결과를 얻어낼 수 있는 방법에 대한 연구가 계속되어야겠다.

## 참 고 문 헌

[1] Myron Z. Brown, Darius Burschka and Gregory

D. Hager, "Advances in Computational Stereo," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.25, No.8, pp.993-1008, August, 2003.

[2] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *IJCV* 47(1), pp.7-42, 2002.

[3] 김봉겸, 임재권, "신뢰확산 알고리즘을 이용한 선 그룹화 기반 스테레오 정합," *전자공학회 논문지* 제42권 SP 제3호, 2005.

[4] P.H.S Torr and A. Criminisi, "Dense Stereo Using Pivoted Dynamic Programming," *Microsoft Research*, MSR-TR-2002-51, 2002.

[5] Ingemar J. Cox L. Hingorani, Satish B.Rao "A Maximum Likelihood Stereo Algorithm," *CVIU*, 63:3, pp.542-567, 1996.

[6] 윤정환, 배병규, 박세환, 송혁, 김동욱, 유지상, "시공간 중복성을 고려한 다이나믹 프로그래밍 기반의 고속 변이 추정 기법," *전자통신학회 논문지* 제33권, 제10호, 2008.

[7] 강창순, 김종덕, 이상욱, 남기근, "영역 분할에 의한 동적계획법을 이용한 스테레오 정합," *대한전자공학회 논문지* 제34권, 제6호, 1997.

[8] 동원표, 정창성, "동적 계획법과 이미지 세그먼트를 이용한 스테레오 정합," *한국컴퓨터종합학술대회* 제32권, 제1호, 2005.

[9] 서용석, 김태준, 유지환, 김동욱, 유지상, "동적 계획법과 영상 분할을 이용한 중간지점 영상 합성," *신호처리합동학술대회 논문지* 제22권, 제1호, 2009.

[10] 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.

[11] 김태준, 유지상, "영상의 컬러 정보를 이용한 계층적 스테레오 정합," *전자통신학회 논문지* 제34권, 제3호, 2009.

[12] O. Veksler. "Stereo correspondence by dynamic programming on a tree," *CVPR*, 2005.

[13] M. Gong and Y.-H. Yang. "Near real-time reliable stereo matching using programmable graphics hardware," *CVPR*, 2005.



서 용 석 (Yong-seok Seo)

준회원



2009년 2월 광운대학교 전자공학과 학사

2010년 3월~현재 광운대학교 전자공학과 석사과정

<관심분야> Stereo Matching, Computer Vision, 3D 입체 영상처리

유 지 상 (Ji-sang Yoo)

중신회원



1985년 2월 서울대학교 전자공학과(공학사)

1987년 2월 서울대학교 전자공학과(공학석사)

1993년 5월 Purdue 전기공학과 (Ph.D.)

1993년 9월~1994년 8월 현대 전자산업(주) 산전연구소 선임연구원

1994년 9월~1997년 8월 한림대학교 전자공학과 조교수

1997년 9월~현재 광운대학교 전자공학과 교수

<관심분야> 3D 입체영상처리/압축, 웨이블릿 기반 영상처리, 비선형 디지털신호처리