

# 효율적인 스테레오 정합을 위한 동적계획법의 역 추적 방법

\*박장호 \*최현준 \*\*서영호 \*김동욱

\*광운대학교 전자재료공학과

\*\*광운대학교 교양학부

\*jhchip@kw.ac.kr

## A back tracing in dynamic programming for efficient the stereo matching

\*Park, Jang-Ho \*Choi, Hyun-Jun \*\*Seo, Young-Ho \*Kim, Dong-Wook

\*Dept. of Electronic Materials Eng., Kwangwoon University,

\*\*Dept. of General Education., Kwangwoon University

### 요약

변이영상은 두 스테레오 영상의 시차에 의해 발생하는 각 화소의 변위를 수록한 영상이다. 이 영상은 깊이영상을 생성하여 시점 간 가상영상을 생성하는데 사용된다. 따라서 변이영상은 다시점 비디오 서비스와 직접적인 연관이 있다. 본 논문에서는 유일성(uniqueness) 제약과 순차성(ordering) 제약을 사용하여 기준영상과 참조영상 사이의 관계를 이용하여 생성한 변이 공간 영상(DSI : disparity space image)으로부터 비용 행렬을 계산하여 최적의 변이 경로를 찾아가는 다이내믹 프로그래밍을 분석하였다. 다이내믹 프로그래밍은 정밀한 변이 맵을 얻을 수 있고, 다른 방식들에 비해 연산 속도가 빠르다는 장점을 가지고 있지만, 영상의 화소값의 변화가 없는 영역에서 이전의 경로를 계속 유지하려는 성질에 때문에 발생 하는 오류확산과 가려진(occluded) 영역에 의한 오차로 인해 정확한 경로를 찾을 수 없는 경우가 빈번히 발생 하여 에러율이 높아지는 단점을 가지고 있다. 이러한 이론을 토대로 기존의 기법들에 비하여 정확도가 우수한 기법들을 제안하였다. 개선된 역 추적 과정을 이용하여 기존의 다이내믹 프로그래밍 기반의 스테레오 정합 기법들보다 우수성이 뛰어난 결과들을 나타내었다.

### 1. 서론

최근 컴퓨터 산업의 발전과 함께 그래픽 하드웨어 및 소프트웨어가 크게 발전하면서 더욱 실감나는 영상을 표현하여 실제처럼 느낄 수 있도록 추구하는 3차원 그래픽 렌더링 기술이 부각되고 있다. 이와 함께 컴퓨터 그래픽과 사람이 상호 작용을 할 수 있는 스테레오 비전 기술 등에 대한 연구가 활발히 진행되고 있다. 차세대 비디오 서비스로 주목되고 있는 다시점(multi-viewpoint) 및 자유 시점(free-viewpoint) 비디오 서비스는 배열된 카메라 시점에서 획득된 영상 이외에 사용자의 의도에 따라 임의의 시점에서 새롭게 만들어진 가상시점을 제공한다. 여기에 사용된 깊이정보는 공간상에서 카메라 시점을 기준으로 대상 물체까지의 거리로 정의되는데, 깊이정보 추출의 기본원리는 각각의 눈에서 따로 관찰되는 물체의 상을 대뇌에서 분석하여 물체의 원근을 인지하는 인각 시각 시스템(HVS : human visual system)에 근거를 두고 있다. 여기서 망막에 맺힌 각각 다른 물체의 상을 공학적으로 해석한 것이 스테레오 영상(stereoscopic image)이다. 그러나 깊이정보 추출 과정은 잡음, 가려진 영역(occluded region) 등으로 인한 깊이 불연속, 모호성 때문에 정확한 정보를 얻는 것은 사실상 불가능하다. 이러한 어려운 점들을 해결하기 위한 방법으로 영역 기반(area-based), 특징 기반(feature-based), 그리고 다이내믹 프로그래밍(dynamic programming) 등의 연구가 진행중이다.

위에서 언급한 방법들을 간단히 설명하자면, 밝기 정보의 변화가

평탄하거나 동일한 영역의 모양, 평균 밝기 및 면적등을 정합 요소로 사용하는 방법을 영역 기반이라고 한다. 밝기 정보에 많이 의존하기 때문에 잡음에 약한 면이 있지만, 영상의 전체 거리 정보를 얻을 수 있다는 강점이 있다. 특징 기반은 교차점, 경계선, 모서리, 마루, 골, 원추 곡선등 영상의 특징이 되는 요소들을 사용하는 방법이라 정의할 수 있다. 이 방법은 정합점이 정확하고 잡음에 강한 강점을 가지고 있으나, 정합 되어지는 점들이 소수이므로 전체 영상의 변이 정보를 구하기 위해서는 가려진 영역과 변이 연속성을 고려하여 복잡한 보간 과정을 필요로 하게 되므로 연산량의 증가를 초래하게 되는 단점이 있다. 이러한 어려움을 극복하기 위하여 영상 내 화소의 유사도에 관한 평가함수에 국한하지 않고 변이정보의 분포 통계에 근거한 확률모델을 이용하여 이를 해결하는 기법이나 신경망(neural network)이론을 도입하여 연산 중 출력되는 정보를 정규화 하고 학습효과를 이용하여 변이정보를 판단하는 기법이 연구되었다. 또한 영상의 단일 영역에 국한하지 않고 영상의 주사선(scan-line)에 대한 화소들의 대응관계를 다이내믹 프로그래밍(dynamic programming)을 이용하여 풀어나가는 기법 또한 연구되고 있다.

본 논문에서 사용하는 다이내믹 프로그래밍 기반의 변이 추정 방식은 유일성(uniqueness) 제약과 순차성(ordering) 제약을 사용하여 기준영상과 참조영상 사이의 관계를 이용하여 생성한 변이 공간 영상(DSI : disparity space image)으로부터 비용 행렬(cost matrix)을 계산하여 최적의 변이 경로를 찾아가는 방식을 말한다. 다이내믹 프로그

래밍은 화소 단위의 정밀한 변이 맵을 얻을 수 있어서 다른 방식들에 비해 연산 속도가 빠르다는 장점을 가지고 있다. 그러나 영상의 화소값의 변화가 없는 영역에서 이전의 경로를 계속 유지하려는 성질 때문에 발생 하는 오류확산과 가려진 영역에 의한 오차로 인해 정확한 경로를 찾을 수 없는 경우가 빈번히 발생 하여 에러율이 높게 된다. 이러한 현상을 스트리킹(streaking)이라고 한다.

본 논문에서는 이전의 경로를 그대로 유지함으로써 발생하는 변이 정보의 오류 확산을 줄이기 위해 영상의 특징점을 이용하여 보다 정확한 변이 정보를 획득하기 위해 경로를 재탐색하는 알고리즘을 제안한다.

## 2. 동적계획법 기반의 스테레오 정합 기법

동적계획법은 수학자 리처드 벨만에 의해 제시된 다이내믹 프로그래밍이라는 최적화 기법을 적용한 스테레오 정합 방법 중 하나이다. 동적계획법은 최소 비용 경로를 찾음으로써 정합을 할 수 있는 알고리즘으로 알려져 있다. 스테레오 정합에서의 다이내믹 프로그래밍은 좌·우 두 영상의 주사선의 화소값을 비교하여 차이도를 만들고, 차이도 상에서 시작점부터 목표점에 이르는 최소 비용 경로(path)를 구함으로써 좌·우 영상을 정합(matching) 시키는 것이다. 전체 순서도는 그림 1과 같다.

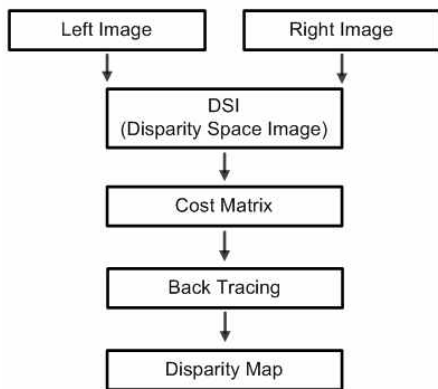


그림 1. 동적계획법의 전체 순서도

### 가. 변이 공간 영상 (DSI : disparity space image)

동적계획법은 좌·우영상의 같은 스캔라인(scanline)의 화소값 차이를 이용하여 그림 2처럼 우측 영상에서 좌측 영상의 탐색구간에서 유사성이 높은 화소를 탐색하는 기법으로 변이 공간 영상을 생성한다. 동적계획법은 WTA(winner-take-all) 방법과 달리, 탐색 구간에서 가장 유사성이 높은 화소를 택하는 것이 아니라, 탐색 구간에 있는 모든 화소들의 비용값을 저장해 둔다. 변이 공간 영상을 생성할 때, 일반적으로 SSD(sum of squared differences), SAD(sum of absolute differences), MAD(mean of absolute differences)등의 비용 함수를 사용한다. 본 논문에서는 수식 (1)과 같은 SAD 비용 함수를 사용하였다. 생성된 변이 공간 영상을 이용하여 비용 행렬을 생성하게 된다. 그림 3은 좌·우영상과 변이 공간 영상이다. 어두운 부분은 정합 오류가 적고 밝은 부분일수록 정합 오류가 크다는 것을 나타낸다. 변이 공간 영상은 그 특성에 따라 수직(horizontal), 수평(vertical), 대각선(diagonal)등의 3개 영역으로 나눌 수 있다. 이 특성들을 이용하여 비용 행렬을 계산한다.  $u, v$ 는 좌·우 영상의 화소를 말하며,  $i, j$ 는 정합창의 크기를 말한다.

$$SAD(u, v, d) = \sum_{(i,j)} |I_L(u+i, v+j) - I_R(u+i+d, v+j)| \quad (1)$$

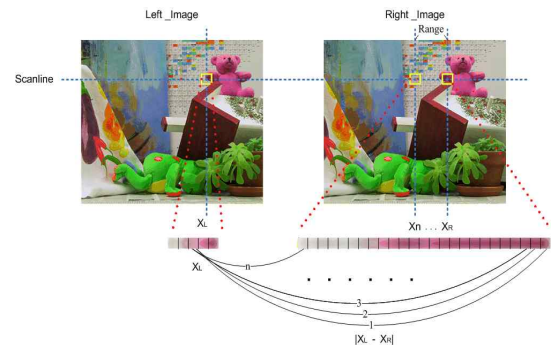


그림 2. 최소화 비용함수를 이용한 유사성 판단

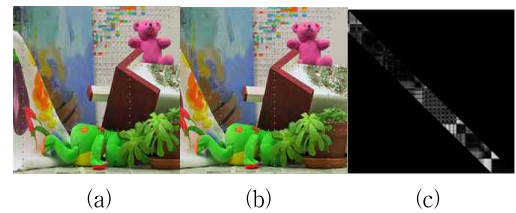


그림 3. (a) 좌영상 (b) 우영상 (c) 변이 공간 영상

### 나. 비용 행렬

비용 행렬은 앞서 미리 구한 변이 공간 영상을 기본으로 수직, 수평, 경계성분을 고려한 비용값과 변이의 연속성을 고려한 비용값을 이용하여 생성하게 된다. 다이내믹 프로그래밍 기반의 스테레오 정합 기법은 이 성분들을 고려하여 다음 화소가 가질 수 있는 변이값을 예측하여 보다 안정되고, 빠르게 변이값을 추출할 수 있도록 한다. 비용 행렬 생성 시 고려되는 3가지 비용값은 식 (2)와 같다.

$$\begin{aligned} C_b(i, j) &= Cost(i-1, j) + P_{occlusion} \\ C_c(i, j) &= Cost(i-1, j-1) + DSI(i, j) \\ C_a(i, j) &= Cost(i-1, j) + P_{occlusion} \end{aligned} \quad (2)$$

그림 4의 (a)는 현재 화소 P의 비용값을 구할 때 고려하는 세 가지 성분 (a,b,c)를 나타내고, 비용값의 진행 경로 방향을 보여주며, (b) 생성된 비용 행렬을 보여준다. Left\_scanline과 Right\_scanline은 좌영상과 우영상의 기준열을 뜻한다.

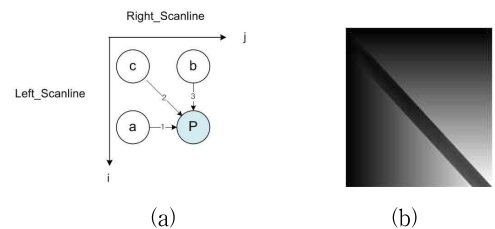


그림 4. (a) 비용 행렬의 경로 (b) 생성된 비용 행렬

식 (2)에서  $C_c$ 는 현재 화소의 변이값이 이전 화소와 연속적인 변이값 즉, 같은 변이값을 가질 때를 고려한 비용이며  $C_a$ 와  $C_b$ 는 변이 공간 영상 내에 수평 경계 성분 및 수직 경계 성분을 고려한 비용이다.  $P_{occlusion}$ 은 가려진 영역을 판별하는데 영향을 미치는 일종의 페널티 (penalty)이다. 위의 세 가지 값을 비교하여 수식 (3)를 통해 가장 최소



하여 그림 9와 같이 a,b,c의 실제 비용값을 비교하여 최소 비용을 갖는 방향으로 경로로 이동한다.

$$DSI(i,j) = Penalty \times DSI(i,j) \quad (4)$$

$$Penalty = \begin{cases} 0 & : |I(i,j) - I(i-1,j)| = 0 \\ 1 & : |I(i,j) - I(i-1,j)| < threshold \\ k & : |I(i,j) - I(i-1,j)| > threshold \end{cases} \quad (5)$$

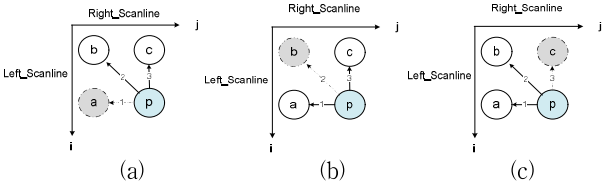


그림 9. (a) 1번경로 (b) 2번경로 (c) 3번경로

#### 4. 실험 방법 및 결과

##### 가. 실험 방법

본 논문에서는 실험영상은 Middlebury 사이트에서 제공하는 4개의 실험영상을 사용하였다. 탐색범위를 제외한 파라미터는 모든 실험영상에 동일하게 적용하였다. 실험영상은 그림 10에 나타내었다.



(a) (b) (c) (d)

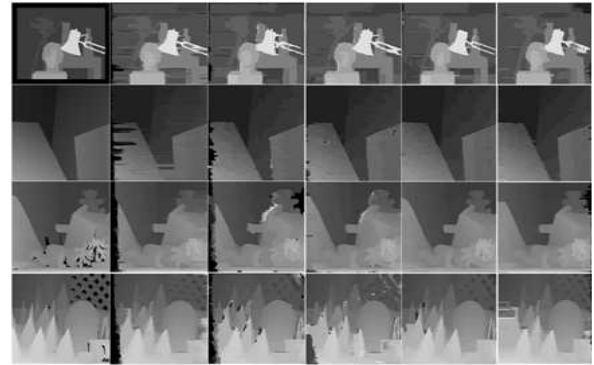
그림 10. 실험 영상; (a) Tsukuba, (b) Venus, (c) Teddy, (d) Cones

##### 나. 실험 결과

본 논문에서는 기존의 다이나믹 알고리즘인 DP(dynamic programming)<sup>1)</sup>, Tree DP(tree dynamic programming)<sup>2)</sup>, Reliability DP(reliability dynamic programming)<sup>3)</sup>, Region Tree DP(region tree dynamic programming)<sup>4)</sup>와 비교하였을 때, 그림 11과 같이 우수한 결과를 보여주고 있다. 표 1은 성능 비교를 나타낸다. 표를 보면 알 수 있듯이, Venus 가 가장 낮은 에러율을 나타내며 Cones 과 Teddy 영상이 높은 에러율을 보이고 있다. 그 이유는 Venus 영상은 복잡도가 다른 테스트 영상들보다 낮지만, Cones 같이 반복적이면서 복잡한 영상은 에러율이 높게 나오는 경향을 보인다. 미들버리에서 정의해 놓은 에러율은 가려진 영역을 제외한 에러율(nonocc)과 불연속지점에서의 에러율(disc), 그리고 앞의 두 가지를 모두 포함한 전체 에러율(all)로 분류할 수 있다. 본 논문에서는 전체 에러율만 나타내었다.

표 1 기존의 방법들과의 성능 비교

Algorithms	Matching Error (%)			
	Tsukuba	Venus	Teddy	Cones
Proposed	4.34	1.79	13.7	15.4
Region Tree DP	1.64	0.57	11.9	11.9
Reliability DP	3.39	3.43	16.9	19.9
Tree DP	2.84	2.1	23.9	18.3
DP	5.04	11	21.6	19.1



(a) (b) (c) (d) (e) (f)

그림 11. 결과 영상들; (a) Ground truth (b) DP (c) Tree DP (d) Reliability DP (e) Region Tree DP (f) Proposed

#### 5. 결론

본 논문에서는 이러한 오류를 줄이기 위해 이전화소와 현재 화소와의 화소값 차이를 분석한 후, 영상의 특징점을 이용하여 역추적과정에서 변이값을 할당 시, 연속적인 변이정보에 대한 신뢰성을 점검 한다. 그리하여 보다 신뢰성이 높은 변이맵을 추출할 수 있었다. Region Tree DP, Reliability DP, Tree DP등의 기존의 다이나믹 프로그래밍 기반의 기법인보다 6~7%의 에러율을 감소하였다.

#### 감사의 글

본 논문은 서울시 신기술 연구개발 지원사업(NT080528)의 지원을 받아 수행된 연구임.

#### 참고문헌

- [1] D. Scharstein, R. Szeliski. "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," International Journal of Computer Vision Springer Netherlands, Vol. 47, no. 1-3 pp. 0920-5691, April 1989.
- [2] Olga Veksler, "Stereo correspondence by Dynamic programming on a Tree", IEEE Computer Society Conference, Vol. no. 2, pp. 20-25, June 2005.
- [3] Minglun Gong, Yee-Hong Yang, "Near real-time reliable stereo matching using programmable graphics hardware", IEEE Computer Society Conference, Vol. 1, pp. 924-931, June 2005.
- [4] Federico Tombari, Stefano Mattocchia, and Luigi Di Stefano, "Segmentation-Based Adaptive Support for AccurateStereo Correspondence", D. Mery and L. Rueda (Eds.): PSIVT 2007, LNCS 4872, pp.427-438, 2007.