

동적계획법을 이용한 컬러 스테레오 정합

오종규, 이찬호, 김종구
현대중공업 기전연구소

Color Stereo Matching Using Dynamic Programming

Jong-Kyu Oh, Chan-Ho Lee, Jong-Koo Kim
Mechatronics Research Institute, Hyundai Heavy Industries Co.,Ltd.

Abstract - In this paper, we proposed color stereo matching algorithm using dynamic programming. The conventional gray stereo matching algorithms show blur at depth discontinuities and non-existence of matching pixel in occlusion regions. Also it accompanies matching error by lack of matching information in the untextured region. This paper defines new cost function makes up for the problems happening in conventional gray stereo matching algorithm. New cost function contain the following properties. i) Edge points are corresponded to edge points. ii) Non-edge points are corresponded to non-edge points. iii) In case of exiting the amount of edges, the cost function has some weight in proportion to path distance. Proposed algorithm was applied in various images obtained by parallel camera model. As the result, proposed algorithm showed improved performance in the aspect of matching error and processing in the occlusion region compared to conventional gray stereo matching algorithms.

1. 서 론

최근 인간의 감각능력을 컴퓨터에 접목시키려는 시도가 빈번한 가운데 인간의 시각 능력을 이용하는 컴퓨터 비전 분야는 많은 발전을 이루어 왔다. 비전 시스템은 고화질의 영상 장비와 고성능 H/W 및 고속 PC 개발로 인해 그 활용 범위가 한층 더 넓어졌으며, 특히 스테레오 영상에서의 일치점을 이용하여 거리정보를 획득하는 스테레오 비전은 로봇 제어, 자동 항법 시스템, 그리고 현재 많은 연구가 진행되고 있는 가상 현실 시스템에 이용되고 있다. 그러나 기존 흑백 영상에서의 스테레오 정합 알고리즘은 표면의 평활화 제약조건의 적용, 폐색 영역에서의 정합 화소 부재 등으로 인하여 불연속 지점에서 무뎌짐 현상을 보였고, 텍스처하지 않은 영역에서는 정합 정보 부족으로 정합 어려움을 동반하였다.[1][2] 따라서 본 논문에서는 기존의 흑백 스테레오 정합 알고리즘에서 야기되었던 문제점을 보완하기 위해서 컬러 정보를 사용하는 새로운 비용함수를 정의하여, 동적계획법에 의한 컬러 스테레오 정합 알고리즘을 제안하고자 한다.

2. 본 론

2.1 동적계획법에 의한 컬러 스테레오 정합

동적계획법은 어떤 문제의 해를 구하기 위해서 최적성의 원리에 기초하여 비용이 최소가 되는 경로를 찾음으로써 해를 구하는 알고리즘이다. 스테레오 정합에서 동

적계획법이 사용하는 특징점은 각 화소점이나 전처리 과정에서 추출된 에지 등이 될 수 있다. 본 연구에서는 모든 화소점들에 대한 색 정보의 상이 정도를 구하여 동적계획법을 수행하기 위한 정보로 사용하였다. 정합을 위한 시스템은 그림 1과 같다.

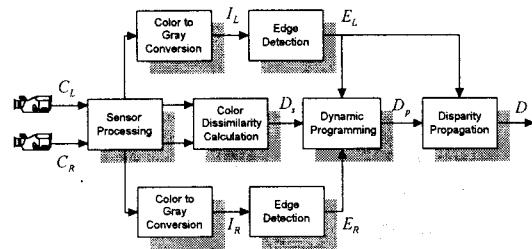


그림 1. System model for color stereo matching.

여기서 C_L , C_R 은 카메라에서부터 나오는 컬러 영상의 RGB 정보를 나타내고, I_L , I_R 은 컬러 정보를 흑백 정보로 바꾼 밝기 정보를 의미한다. E_L , E_R 은 추출된 에지 정보, D_S 는 상이 정보, 그리고 D_P , D 는 후처리 전·후의 변이값을 의미한다. 에지 검출은 Sobel 연산자를 사용하였다.

2.1.1 상이값 계산

동적계획법을 사용한 흑백 스테레오 정합에서는 모든 화소점들에 대한 밝기 차이값을 주로 특징점으로 사용하였다. 이 경우 획득된 영상을 디지털화 하는 과정에서 샘플링의 문제가 발생한다. 이러한 문제를 완화시키기 위하여 상이값을 도입하여 동적계획법의 특징점으로 사용한다.[3]

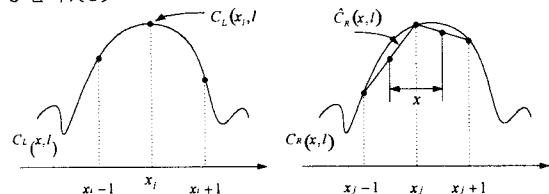


그림 2. Illustration for dissimilarity.

그림 2는 기준열 l 에 대한 좌·우 영상의 x_i 와 x_j 화소 사이의 상이값을 구하는 경우를 설명하고 있다. 여기서 $C_L(x, l)$ 과 $C_R(x, l)$ 은 대응하는 기준열 상의 색 정보인 RGB값이고, 표시된 점들은 영상 획득시 샘플링된 이웃한 화소들을 나타낸다. 또한, \hat{C}_R 은 우측 영상에서 인근 두 점 사이의 RGB 각각의 선형 합수값을 의미한다. 여기서 $C_L(x_j, l)$ 화소와 우측 영상에서 $x_j - \frac{1}{2} \leq x \leq x_j + \frac{1}{2}$ 구간에서의 $\hat{C}_R(x, l)$ 사이의 최소 차이값 d 를 다음과 같이 정의한다.

$$d(x_i, C_L; x_j, C_R) = \min(|R_L(x_i, l) - R_R(x_j, l)| + |G_L(x_i, l) - G_R(x_j, l)| + |B_L(x_i, l) - B_R(x_j, l)|) \quad (1)$$

for $\{x|x_i - \frac{1}{2} \leq x \leq x_j + \frac{1}{2}\}$

여기서, $R_L(x, y)$, $G_L(x, y)$, $B_L(x, y)$ 은 원쪽영상의 (x, y) 에서의 각각의 RGB 값이고 R_R , G_R , B_R 는 우측 영상에서 인근 두 점 사이의 RGB 각각의 선형 함수를 나타낸다. 이때 두 화소 x_i 와 x_j 사이의 상이값은 다음과 같이 정의한다.

$$D_s(x_i, x_j) = \min\{d(x_i, C_L; x_j, C_R), d(x_j, C_R; x_i, C_L)\} \quad (2)$$

2.1.2 제안된 동적계획법

식 (2)의 상이값으로 그림 3의 (a)와 같은 상이도를 구성할 수 있다. 이것은 최대 변이가 ± 3 인 경우에 대한 상이도이고, 이 경우 흰색 영역을 검색 영역으로 제한하였다.

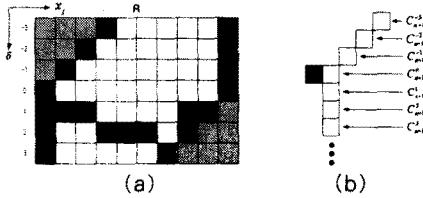


그림 3. Concept for searching cost path.
(a) Dissimilarity Map.
(b) Possible Searching Path.

그림 3의 (a)에서 검은 색으로 나타낸 지점은 동적계획법에 의한 경로를 검색하기 위한 시작점과 목표점을 나타낸 것이다. 검은색 사선은 최적의 비용 경로에 대한 예이다. 즉, 동적계획법은 시작점과 목표점을 잇는 최소 비용 경로를 찾는 것이다.

본 연구에서는 경험적인 접근을 바탕으로 비용 함수를 정의한다. 즉, 이동 가능한 구간의 경로에서는 기본적으로 상이값을 누적해서 더하고, 폐색 영역의 경우는 폐색 폐널티 상수(constant occlusion penalty), P_{occ} 를 이동 경로에 비례하는 값으로 가중적으로 추가한다. 이것은 부적절하게 폐색 영역으로의 경로 확장을 억제하게 한다. 또한, 에지 영역의 화소들은 상대 영상의 에지 영역의 화소들과 정합되기를 유도하기 위하여, 에지 영역에서는 그 비용 값을 줄이고 그 외의 영역에서는 비에지 영역 폐널티 상수(constant non-edge penalty), P_{ne} 를 추가하여 비용값을 상대적으로 크게 한다. 여기서, 시작점부터 목표점까지의 전체 경로 중 임의의 경로에 대한 n 번째 정합 화소의 비용 함수, T_n 은 식 (3)과 같이 표현할 수 있다.

$$T_n = N_{occ} \cdot P_{occ} + (n - N_e) \cdot P_{ne} + \sum_{i=1}^n D_s(r_i, \delta_i) \quad (3)$$

여기서, N_{occ} 는 폐색 영역의 화소의 수, n 은 정합 화소의 수, N_e 는 에지 화소의 수, i 는 정합쌍의 순서를 의미한다.

본 연구에서는 그림 3의 (b)는 검색 가능한 경로를 나타낸다. 여기서 검은색 화소는 현재의 위치를 나타내고, 흰색 화소들은 가능한 다음의 경로를 표시한다. 이때 비용 함수의 적용은 그림에서와 같이 현재의 위치를 기준으로 가능한 다음 위치들에 대한 비용의 누적 형태를 취하고 이것은 식 (4)와 같이 표현할 수 있다.

$$\begin{aligned} T_{n+1}^k &= T_n + D_s \cdot \delta_{n+1} + P_{n+1}^k, \\ \{k\} &\cdots -3, -2, -1, 0, 1, 2, 3 \cdots \\ P_{n+1}^k &= \begin{cases} P_{ne} & \text{if } E_{n+1}^k = 0 \\ d_n \cdot P_{occ} & \text{if } E_{n+1}^k = 1 \end{cases} \\ E_{n+1}^k &= \begin{cases} E_R \cdot \delta_{n+1} & \text{if } k \geq 0 \\ E_L \cdot \delta_{n+1} & \text{otherwise} \end{cases} \end{aligned} \quad (4)$$

여기서 T_n 은 현재까지의 누적 비용을 나타내고, T_{n+1}^k 은 임의의 다음 위치까지의 누적 비용, $D_s \cdot \delta_{n+1}$ 는 다음 위치의 상이값, d_n 은 T_n 와 T_{n+1}^k 사이의 거리, $E_R \cdot \delta_{n+1}$ 은 좌·우 영상의 에지 함수를 의미한다. 그리고 k 는 임의의 다음 경로를 의미하는 것으로 실제 알고리즘에서는 변이 검색의 충분한 값인 $-15 \leq k \leq 15$ 로 두었다. 여기서 에지 영역과 그 외의 영역에 대해 차별적으로 정합이 가능하게 하기 위해 $P_{ne} (= 1000) \gg P_{occ}$ 인 조건이 만족되게 두 상수값을 정하였다. 그리고 에지가 너무 많은 영상의 경우 부적절하게 폐색 영역으로의 경로 확장을 억제하기 위해 P_{occ} 는 $d_n \cdot P_{occ}$ 와 같이 이동 경로의 거리에 비례하는 값으로 가변 적용한다. 한편 여기서 얻어진 T_{n+1}^k 의 값은 이전의 검색에서 $n+1$ 의 위치에 채워진 비용값들 보다 작을 경우 새로운 $n+1$ 위치의 비용값으로 결정된다.

2.2 변이 정보의 보정

변이들 사이의 신뢰도 개념을 도입하여 스캔라인 사이의 관계를 규명하고 이를 정보를 이용하여 이미 얻어진 변이값들을 더욱 타당한 값들로 수정하고자 한다. 통상 정합에러는 산발적으로 일어날 가능성이 높기 때문에 변이 값이 자주 바뀌는 것은 오류일 가능성성이 높다. 이러한 이유로 같은 변이가 연속되는 횟수를 해당되는 변이값들의 신뢰도로 정의한다. 이때 동적계획법을 적용함으로서 구해진 변이값 D_p 에 대해 수평 방향으로의 신뢰도 R_h 와 수직 방향으로의 신뢰도 R_v 는 다음 식 (5), (6)과 같이 표현할 수 있다.

$$R_h(k, j) = \sum_{i=k}^N 1, \text{ for } D_p(i, j) = D_p(k, j) \quad (5)$$

$$R_h(k+m_1, j) = R_h(k, j), \quad \{m_1|1, 2, \dots, R_h(k, j)-1\}$$

$$R_v(i, k) = \sum_{j=k}^N 1, \text{ for } D_p(i, j) = D_p(i, k) \quad (6)$$

$$R_v(i, k+m_2) = R_v(i, k), \quad \{m_2|1, 2, \dots, R_v(i, k)-1\}$$

구해진 신뢰도를 이용하여 높은 신뢰도를 갖는 변이 정보를 낮은 신뢰도를 갖는 변이값들에 전파함으로써 산발적으로 발생하는 에러는 보정할 수 있다. 이때, 높은 신뢰도는 같은 변이가 25회 이상 반복되는 경우로 하였고, 낮은 신뢰도는 12회 이상 반복되는 경우로 하였다. 같은 변이가 13~24회 반복되는 경우는 신뢰도의 높고 낮음의 중간적인 수치이므로 전파 또는 피전파 과정에서 제외하였고, 이를 값들은 다양한 영상에 대한 적용 후 설정된 값들이다.

2.3 실험결과

본 논문에서 제안한 동적계획법을 이용한 컬러 스테레오 정합 알고리즘을 320×240 의 다양한 스테레오 영상에 적용하여 기존의 흑백 정합 알고리즘들과 결과를 비교하였다. 그림 4, 5의 (a), (b) 영상은 평행 카메라 모델로 획득된 영상으로 그림 4는 배경을 텍스처하게,

그림 5는 텍스처하지 않게 구성한 영상이다. 또한 모든 영상은 2~3개의 물체와 배경을 카메라로부터 다양한 거리에 두었다.

그림 4,5에서 (c)는 SSD[1], (d)는 Cox 알고리즘[2], (e)는 Birchfield 알고리즘[4], (f)는 제안된 알고리즘을 적용했을 때의 결과 변이도이다. (g)는 3차원으로 표면을 재구성한 와이어 프레임이다.[5]



(a) Left Image



(b) Right Image



(c) SSD



(d) Cox's MLMH+V



(e) Birchfield Method



(f) Proposed Method

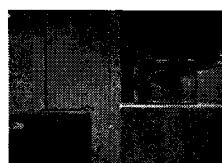


(g) Surface Reconstruction by Wire Frame

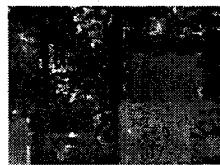
그림 4. Textured Image & Result



(a) Left Image



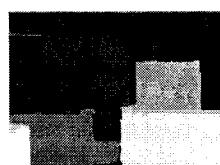
(b) Right Image



(c) SSD



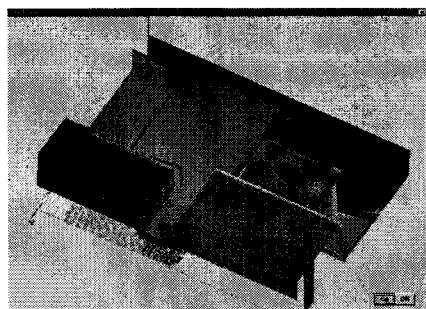
(d) Cox's MLMH+V



(e) Birchfield Method



(f) Proposed Method



(g) Surface Reconstruction by Wire Frame

그림 5. Untextured Image & Result

3. 결 론

본 논문에서는 컬러 스테레오 영상으로부터 깊이 정보를 추출하기 위해 새로운 비용 함수를 정의하여 동적 계획법을 적용시키는 스테레오 정합 알고리즘을 제안하였다. 제안된 알고리즘의 성능평가를 위해 기존의 흑백 스테레오 정합 알고리즘들과 정합 결과를 비교하였다. 제안된 알고리즘을 평행 카메라 모델로 획득한 다양한 영상에 적용한 결과 기존의 흑백 스테레오 정합 알고리즘에 비해 폐색영역에서의 처리와 불연속 지점에서의 무더짐 현상을 보완하였으며, 와이어 프레임으로도 제안된 알고리즘의 성능이 개선되었음을 확인하였다.

(참 고 문 헌)

- [1] G. A. Wood, "Realities of automatic", Photogrammetric Engineering and Remote Sensing, vol. 49, pp.537-538, Apr. 1983.
- [2] I. J. Cox, S. L. Hingorani and S. B. Rao, "A Maximum Likelihood Stereo Algorithm", Computer Vision and Understanding, 64(3), pp.542-567, May, 1996.
- [3] S. Birchfield and C. Tomasi, "A pixel dissimilarity measure that is insensitive to image sampling", IEEE Trans. PAMI, 20(4), pp.401-406, Apr. 1998.
- [4] S. Birchfield and C. Tomasi, "Depth discontinuity pixel-to-pixel stereo", in Proc. ICCV'98, pp.1073-1081, 1998.
- [5] S. Malassiotis and M. G. Strintzis, "Model-Based Motion and Structure Estimation from Stereo Images", Computer Vision and Image Understanding, Vol. 65 No. 1, pp.79-94, January, 1997.