論文99-36S-10-13

에지 정보를 강조한 동적계획법에 의한 스테레오 정합

(Stereo Matching by Dynamic Programming with Edges Emphasized)

朱 在 欽 *, 吳 宗 奎 **, 薛 盛 旭 **, 李 哲 憲 ***, 南 基 坤 **

(Jae-Heum Joo, Jong-Kyu Oh, Sung-Wook Seol, Chul-Hun Lee, and Ki-Gon Nam)

요 약

본 논문에서는 에지를 강조하여 동적계획법을 적용한 스테레오 정합 알고리즘을 제안한다. 기존 알고리즘에서는 표면의 평활화 제약조건의 적용, 폐색 영역에서의 정합 화소 부재 등으로 인하여 대체로 불연속 지점에서의 무뎌짐 현상을 보이고 있다. 또한 밝기 변화가 없는 영역에서는 정합 정보 부족으로 인하여 정합 에러를 동반하게 된다. 본 논문에서는 좌·우 영상의 에지 부분은 에지 부분 사이에 정합이 이루어지고, 그 외의 영역에서는 그 외의 영역 사이에 정합이 이루어지게 유도함으로써 기존 알고리즘에서 야기되었던 문제점들을 보완하는 새로운 비용 함수를 정의하였다. 또한 영상에서 에지가 다량으로 발생할 가능성에 대비해 에지 정보사이의 정합은 비용 함수에 경로 거리와 비례하는 가중치를 추가하였다. 제안된 알고리즘을 평행 카메라 모델 뿐만아니라 수렴 카메라 모델로 획득한 다양한 형태의 영상에 적용한 결과, 기존의 알고리즘에 비해 폐색 영역에서의 처리와 정합 에러 측면에서 개선된 성능을 보였고, 특히 불연속 지점에서의 흐려짐 현상이 개선됨을 확인하였다.

Abstract

In this paper, we proposed stereo matching algorithm by dynamic programming with edges emphasized. Existing algorithms show blur generally at depth discontinuities owing to smoothness constraint and non-existence of matching pixel in occlusion regions. Also it accompanies matching error by lackness of matching information in the untextured regions. This paper defines new cost function to make up for the problems occurred to existing algorithms. It is possible through deriving matching of edges in left and right images to be carried out between edge regions and deriving that in the other regions to be performed between the other regions. In case of the possibility that edges can be produced in a large amount, matching between edge information adds weight to cost function in proportion to path distance. Proposed algorithm was applied to various images obtained by convergent camera model as well as parallel camera model. As the result, proposed algorithm showed improved performance in the aspect of matching error and processing in the occlusion regions compared to existing algorithms. Also it could improve blur especially in discontinuity regions.

* 正會員, 智山大學 電子計算科
(Dept. of Computer Science, Jisan College)
** 正會員, 釜山大學校 電子工學科
(Dept. of Electronics Eng., Pusan National University)
*** 正會員, 梁山大學 事務自動化科
(Dept. of Office Automation, Yangsan College)
接受日字:1999年5月28日, 수정완료일:1999年9月28日

Ⅰ. 서 론

최근 컴퓨터 공학의 발달과 더불어 인간의 능력을 컴퓨터에 접목시키고자 하는 연구와 노력이 각 분야에서 활발하게 진행되고 있다. 특히, 카메라를 이동하거나 또는 두 대의 카메라로부터 얻은 두 영상을 이용하는 스테레오 비젼은 공간상의 한 점이 두 영상에 투사되는 점의 위치 차이 즉 변이(disparity)를 이용하여 3차원 깊이 정보를 획득하는 것으로 인간의 시각 구조와 유사하며 환경의 특별한 제약이 없고 일반적인 자연환경을 대상으로 할 수 있는 장점 때문에 많이 응용되고 있다. 여기서 획득된 깊이 정보는 로봇의 정확한 제어, 자동 항법 시스템 등에 이용될 수 있을 뿐만 아니라 현재 많은 연구가 진행되고 있는 가상 현실 시스템 구축을 위한 기초 자료로서 활용될 수 있다.[1]

3차원 깊이 정보를 획득하기 위한 스테레오 정합 방법은 크게 영역 기반(arca-based), 특징 기반(feature-based), 그리고 밝기 기반(intensity-based)으로 나눌 수 있고[2], 이러한 접근 방법을 바탕으로 다양한 방향으로의 연구가 진행되어 왔다. 그 중 베이시안 모델(Bayesian model)로부터 사후 에너지 비용 함수(a posteriori cost energy function)를 유도하여 반복법의 형태로 변이 정보를 구하는 밝기 기반 스테레오 정합 알고리즘에 대한 연구가 있었다.[2,3] 이러한 연구에서는 평활화 제약조건(smoothness constraint)으로 인해 대체로 불연속 지점에서 흐려짐(blurring) 현상이 발생함을 볼 수 있었다. 한편 스테레오 정합 알고리즘에 동적계획법(dynamic programming)을 적용한 연구도 계속적으로 진행되어 왔다. 동적계획법을 사용한 대표적인 논문으로 Baker & Binford[4]와 Ohta & kanade[5]가 있는데 이들은 기본적인 정합 특징으로 에지를 사용하였다. 그러나 에지를 특징점으로 사용할 경우 에지 영역에 대한 변이 값 만을 구할 수 있으므로 그 외의 영역에 대해서는 보간을 해야하는 어려움이 있다.[6] 이러한 문제점을 보완하기 위해 보간 과정없이 전 영역의 변이를 얻기 위한 방법으로 모든 화소를 정합 특징으로 사용한 연구들이 계속적으로 발표되었다.[6-9] 이러한 연구에서는 먼저 좌·우 영상에서 기준선(epipolar line) 상의 화소들의 밝기 차를 취하는 차이도를 구성한다. 차이도는 좌·우 영상의 일치점에서 작은 값을 가지며, 일치되지 않는 지점에서는 큰 값을 가지게 된다. 따라서 좌·우 영상의 일치점을 찾는 것은 차이도 상의 시작점에서부터 목표점까지 최소값을 갖는 경로를 찾는 것으로 해결할 수 있다. 이러한 형태의 동적계획법은 영상의 밝기 변화가 있는 부분에서는 대체로 정확한 일치점을 찾을 수 있으나 영상의 밝기 변화가 없는 부분에서는 이전의 경로를 그대로 유지하려는 성질이 있기 때문에 정확한 경로를 찾을 수 없는 경우가 발생한다. 또한 차이도는 획득된 영상을 디지털 데이터화 하는 과정에서 샘플링 문제를 야기할 수 있고 이런 문제를 보완하기 위해 상이도(dissimilarity map) 개념을 도입한 연구도 있었다.[10]

본 논문에서는 모든 화소의 밝기 정보를 특징점으로 사용하고, 기준선 상의 모든 화소들에 대한 상이도를 도입하여 동적계획법을 적용하였다. 특히, 에지 화소들은 상대 영상의 에지 화소들에 정합되게 하고, 특징이 없는 화소들은 특징이 없는 화소들간에 정합되기를 유도하기 위하여, 에지 영역에서는 그 비용 값을 줄이고 그 외의 영역에서는 비용 값을 상대적으로 크게 적용하는 비용 함수를 정의하였다. 이때 에지 정보가 너무 많은 영상의 경우에는 에지를 강조하여 정합하는 의미가 반감되므로 에지사이의 정합에서는 경로에 비례하는 가중치를 두어 부적절하게 폐색 영역으로의 경로 확장을 억제하게 하였다. 또한 구해진 변이 정보에 대해 기준선 사이의 상관 정보를 이용하여 보다 정확한 변이 정보를 획득하기 위하여 후처리 과정을 두었다.

제안된 알고리즘은 기존 발표된 여러 알고리즘들[7-9]과 그 결과를 비교하였다. 이때 객관적인 성능 비교를 위하여 스테레오 영상 획득시 가장 널리 사용되는 평행 카메라 모델로서 획득된 영상 뿐만아니라, 중심 거리(base line)를 길게하여 측정의 정밀도를 높일 수 있는 수렴 카메라 모델[11]로서 획득한 영상 등 다양한 영상에 대해 알고리즘을 적용하였다.

본 논문의 구성은 다음 같다. 먼저 Ⅱ장에서는 동적계획법에 의한 스테레오 정합의 전체 구성과 제안된 비용 함수의 적용에 대해 기술하고, Ⅲ장에서는 다양하게 획득된 영상에 대해 제안한 알고리즘의 적용 결과를 보이고, Ⅳ장에서 결론을 맺는다.

Ⅱ. 동적계획법에 의한 스테레오 정합

동적계획법은 어떤 문제의 해를 구하기 위하여 연속된 결정을 해야 하는 다단계 결정 과정을 최적성 원리

에 기초한 방법으로 최소 비용 경로를 찾음으로써 해를 구하는 알고리즘이다. 스테레오 정합에서 동적계획법이 사용하는 특징점은 각 화소점이나 전처리 과정에서 추출된 에지 등이 될 수 있다.[6] 본 연구에서는 모든 화소점들에 대한 밝기 상이 정도를 구하여 동적계획법을 수행하기 위한 정보로 사용하였다. 정합을 위한 시스템은 그림 1과 같다.

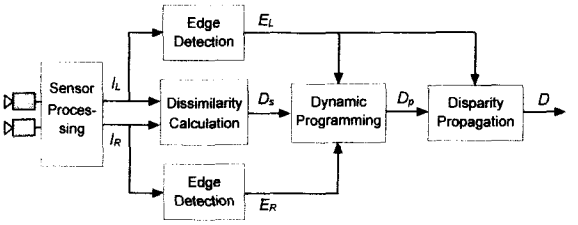


그림 1. 정합을 위한 시스템 모델
Fig. 1. System model for stereo matching.

여기서 IL, IR은 좌·우 영상의 밝기 정보이고, EL, ER은 추출된 에지 정보, Ds는 상이 정보, 그리고 Dp, D는 후처리 전·후의 변이값을 의미한다.

카메라로부터 획득된 좌·우 영상으로부터 상이도를 구하고 이들 정보를 특징점으로 동적계획법을 적용하여 최소 비용 경로를 구한다. 여기서 에지 영역의 화소들은 상대 영상의 에지 영역의 화소들에 정합되도록 유도하기 위하여, 에지 영역에서는 그 비용 값을 줄이고 그 외의 영역에서는 비용값을 상대적으로 크게 적용한다. 에지 검출은 Sobel 연산자를 사용하였다. 구해진 최소 비용 경로를 역으로 추적함으로써 변이값을 구할 수 있고, 구해진 변이로부터 에지 정보를 이용한 후처리 과정을 통하여 보다 명확한 변이 정보를 획득한다.

1. 영상 획득 카메라 모델

스테레오 영상 획득시 가장 널리 사용되는 카메라 배열 방법은 평행 카메라 모델로서 중심 거리가 일정하게 되게 두 대의 카메라를 평행 이동시킨 후 물체를 촬영하는 형태이다. 이때 두 카메라의 광축은 서로 평행하게 된다. 이 경우 기준선 제약조건이 정확히 적용될 수 있고, 두 영상의 대응점을 보다 쉽게 검색할 수 있을 뿐만 아니라 대응점에 대한 기하학적 해석도 용이하다. 그러나 물체와 카메라와의 거리에 비해 중심 거리가 상대적으로 클 경우 한쪽 영상에서는 보이는 물체가 다른 영상에서는 보이지 않는 현상이 발생될 수 있다. 즉, 중심 거리가 크지면 공통 시야가 줄어들게 된다. 한편 공통 시야의 범위를 증가시키기 위한 방법으로는 수렴 카메라 모델(convergent camera model)이 있다. 이 방법은 공통 시야의 범위를 넓히기 위해 두 카메라에 의한 광축이 평행되지 않고 물체 방향으로 수렴되게 구성하여 영상을 획득하는 방법이다. 이 경우 중심 거리를 길게 해도 일정한 공통 시야를 확보할 수 있고 이로 인해 삼각법에 의한 거리 측정의 정밀도를 높일 수 있다. 이러한 이유로 본 연구에서는 제안된 알고리즘의 내구성(robustness)을 입증하기 위해 평행 카메라 모델로 획득된 영상 뿐만 아니라 수렴 카메라 모델로서 획득된 영상에 대해서도 제안된 알고리즘을 적용하였다.

2. 상이값 계산

동적계획법을 사용한 스테레오 정합에서 특징점으로 모든 화소점들을 사용할 경우, 좌·우 영상에서 대응하는 두 기준선 사이의 모든 화소점들에 대한 차이값을 주로 특징점으로 사용 하였다. 이 경우 영상 획득시 센서에 의한 영향을 받게 된다. 즉, 에지 점들 근처의 지점에서는 아주 큰 차이값을 유발시킬 가능성이 있다. 이것은 획득된 영상을 디지털 데이터화 하는 과정에서 발생하는 샘플링의 문제로서 이러한 문제를 완화시키기 위하여 상이값을 도입하여 동적계획법의 특징점으로 사용한다.

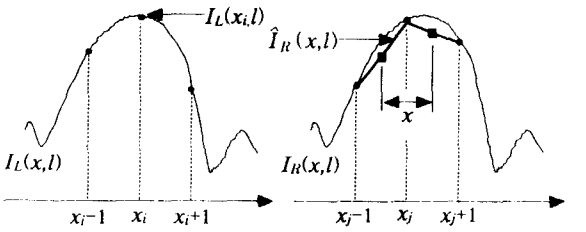


그림 2. 상이값의 설명도
Fig. 2. Illustration for dissimilarity.

그림 2는 기준선 l에 대한 좌·우 영상의 xi와 xj 화소 사이의 상이값을 구하는 경우를 설명하고 있다. 여기서 IL(x,l)과 IR(x,l)은 대응하는 기준선 상의 밝기 값들이고, 표시된 점들은 영상 획득시 샘플링된 이웃한 화소들을 나타낸다. 또한, $\hat{I}_R$ 은 우측 영상에서 인근 두 점 사이의 선형 함수값을 의미한다. 여기서 IL(xi,l)과 $\hat{I}_R$(x,l)사이의 최소 차이값 d를 다음과 같이 정의한다

$$d(x_i, I_L; x_j, I_R) = \min_{x_j-\frac{1}{2} \le x \le x_j+\frac{1}{2}} |I_L(x_i, l) - I_R(x, l)| \quad (1)$$

이때 두 화소 x_i와 x_j 사이의 상이값은 다음과 같이 정의한다.

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

3. 제안된 동적계획법

식 (2)의 상이값으로 그림 3의 (a)와 같은 상이도를 구성할 수 있다. 이것은 최대 변이가 ±3인 경우에 대한 상이도이고, 이 경우 흰색의 영역을 검색 영역으로 제한하였다. 여기서 x_i와 x_j는 좌·우 영상의 기준선의 x 좌표를 나타낸다. 그림 (b)는 (a)를 y 축이 수평이 되게 세로축으로 열의 수 만큼 위로 이동시킨 것이다. 여기서 가로축은 우측 영상을 기준으로 한 기준선 상의 화소 위치이고, 세로축은 해당 화소에 대한 변이를 나타낸다.

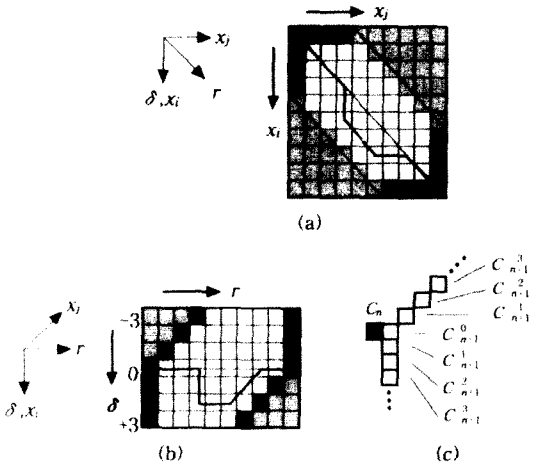


그림 3. 비용 경로의 검색을 위한 개념도
(a) 상이도 (b) 압축된 상이도
(c) 가능한 검색 경로

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

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

동적계획법에서 비용 함수는 통상 베이시안 이론(Bayesian formulation)으로부터 최대 사후 비용 함수(maximum a posteriori cost function)를 유도하여 적용한다.[12]

본 연구에서는 경험적인 접근을 바탕으로 비용 함수를 정의한다. 즉, 이동 가능한 구간의 경로에서는 기본적으로 상이값을 누적해서 더하고, 폐색 영역의 경우는 폐색 패널티 상수(constant occlusion penalty), Pocc를 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 에지가 아닌 영역의 패널티 상수(constant non-edge penalty), Pne를 추가하여 비용값을 상대적으로 크게 한다. 여기서, 시작점부터 목표점까지의 전체 경로중 m번째 임의의 경로에 대한 비용 함수, Cm은 식 (3)과 같이 표현할 수 있다.

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

여기서, N_{occ}는 폐색 영역의 화소의 수, n은 정합 화소의 수, N_e는 에지 화소의 수, i는 정합쌍의 순서를 의미한다.

동적계획법을 스테레오 정합에 보다 효율적으로 적용시키기 위해 경로 검색을 위한 다양한 방향의 연구가 있었다.[7-9,12] 본 연구에서는 그림 3의 (c)와 같은 방법을 사용하였다. 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.

$$C^{k}_{n+1} = C_n + D_s{}^{k}_{n+1} + P^{k}_{n+1},$$
$$\{k | \cdots -3, -2, -1, 0, 1, 2, 3 \cdots\}$$
$$P^{k}_{n+1} = \begin{cases} P_{ne} & \text{if } \hat{E}^{k}_{n+1} = 0 \\ d_n \cdot P_{occ} & \text{if } \hat{E}^{k}_{n+1} = 1 \end{cases} \quad (4)$$
$$\hat{E}^{k}_{n+1} = \begin{cases} E_R{}^{k}_{n+1} & \text{if } k \ge 0 \\ E_L{}^{k}_{n+1} & \textit{otherwise} \end{cases}$$

여기서 C_n은 현재까지의 누적 비용을 나타내고, C^{k}_{n+1}은 임의의 다음 위치까지의 누적 비용, $D_s{}^{k}_{n+1}$는 다음 위치의 상이값, d_n은 C_n와 C^{k}_{n+1}사이의 거리, $E_R{}^{k}_{n+1}$, $E_L{}^{k}_{n+1}$은 좌·우 영상의 에지 함수를 의미한다. 그리고 k는 임의의 다음 경로를 의미하는 것으로 실제 알고리즘에서는 대응점 검색상 충분한 값 $(-14 \le k \le 14)$으로 두었다. 여기서 에지 영역과 그 외의 영역에 대해 차별적으로 정합이 가능하게 하기위해 $P_*(=1000) \gg P_{occ}$인 조건이 만족되게 두 상수값을 정하였다. 그리고 에지가 너무 많은 영상의 경우 부적절하게 폐색 영역으로의 경로 확장을 억제하기 위해 P_{occ}는 $dn \cdot P_{occ}$와 같이 이동 경로의 거리에 비례하는 값으로 가변 적용한다. 한편 여기서 얻어진 C^{k}_{n+1}의 값은 이전의 검색에서 n+1의 위치에 채워진 비용값들 보다 작을 경우 새로운 $n+1$위치의 비용값으로 결정된다.

한편, 알고리즘을 간편화하고 보다 명확한 정합을 위해 기준선 제약조건(epipolar constraint), 유일성 제약조건(unique constraint), 순서성 제약조건(ordering constraint)을 사용하였다.

4. 변이 정보의 보정

주사선(scanline) 사이의 화소의 밝기값들은 서로 무관한 것이 아니다. 마찬가지로 스테레오 영상에서의 실제 변이는 각 주사선 사이에 독립적인 것은 아니다. 그러나 지금까지의 처리는 주사선 별로 독립적인 처리가 이루어 졌다. 이들 주사선 사이의 관계를 규명하고 이들 정보를 이용하여 이미 얻어진 변이 값들을 더욱 타당한 값들로 수정하고자 한다.

여기서 변이들 사이의 신뢰도 개념을 도입한다. 통상 정합 에러는 산발적으로 일어날 가능성이 높기 때문에 변이 값이 자주 바뀌는 것은 오류일 가능성이 높다. 이러한 이유로 같은 변이가 연속되는 횟수를 해당되는 변이값들의 신뢰도로 정의한다. 이때 동적계획법을 적용함으로서 구해진 변이값 D_p에 대해 수평 방향으로의 신뢰도 R_h와 수직 방향으로의 신뢰도 R_v는 다음 식 (5), (6)과 같이 표현할 수 있다.

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

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

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

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

여기서 N_x와 N_y는 영상 폭과 영상 높이를 나타낸다. 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, 이들 값들은 다양한 영상에 대한 적용 후 설정된 값들이다.

한편, 높은 신뢰도로부터 낮은 신뢰도로의 변이값의 전파는 우선 수직 방향의 모든 열에 대해 적용하고 그 결과에 대해 수평 방향의 모든 행에 대해 적용한다. 한편 이러한 전파 방법은 불연속 지점과 같이 변이 값들이 완전한 독립인 경우에도 전파될 가능성이 있다. 그러나 이 단계에서 불연속 지점을 명확하게 지정할 수 없기 때문에 불연속 구간이 될 가능성이 있는 에지 구간에서는 아무리 높은 신뢰도의 변이값이라 할지라도 전파되지 않게 하였다.

Ⅲ. 실험 결과 및 검토

본 논문에서 제안한 에지를 강조한 동적계획법을 이용한 스테레오 정합 알고리즘은 256×256의 다양한 스테레오 영상에 적용하여 기존 동적계획법을 채택한 알고리즘들과 결과를 비교하였다. 실험에 사용된 스테레오 영상은 표 I과 같은 환경으로 획득하여 그림 4~7에 나타내었다.

표 1. 영상 획득 환경

Table 1. Environments for Image Capture.

Env. / Image	Baseline [mm]	Distance from Cameras[m]				Image Property	Camera Model
		Object1	Object2	Object3	Background		
Test1	10	2	2.5	3	4	Textured	parallel Camera Model
Test2	10	2	3	·	10	Untextured	〃
Test3	30	2	3	·	4	Textured	convergent Camera Model
Test4	30	2	3	·	10	Untextured	〃

그림 4, 5의 영상은 평행 카메라 모델로서 획득된 영상이고, 그림 6, 7은 수렴카메라 모델로서 획득된 영상이기 때문에 주어진 환경에서 중심 거리를 30mm까지 길게 둘 수 있었다. 이때 광축의 수렴각은 약 0.9°로 하였고 이것은 현재의 영상 획득 환경에서 공통시야를 최대로 확보할 수 있을 때의 각도이다. 그리고, 그림 4, 6은 배경을 텍스쳐하게, 그림 5, 7은 밝기 변화가 없는 형태로 구성한 영상이다. 또한 모든 영상은 2~3개의 물체와 배경을 카메라로부터 다양한 거리에 두었다.

그림 4~7에서 (a), (b)는 획득된 스테레오 원 영상이고, (c)는 종래의 알고리즘[7], (d)는 Cox 알고리즘[8], (e)는 Birchfield 알고리즘[9], (f)는 제안된 알고리즘을 적용했을 때의 결과 변이도이다. 여기서 Cox 알고리즘은 최대 유사도 비용함수를 이용하고, 경로에 대한 가중치를 베이시안 확률을 이용하여 결정하는 알고리즘이다. Birchfield 알고리즘은 경험에 의한 새로운 비용함수를 정의하였고, 먼저 대략적으로 변이정보를 구한 뒤, 후처리 과정에서 보다 명확하게 변이를 보정하는 전략을 취하는 알고리즘이다. Test1 영상에 대한 종래의 알고리즘과 Cox 알고리즘의 결과는 3단계로 놓여 있는 물체의 변이값을 정확하게 표현하고 있으나, 물체 내부 곳곳에 에러가 발생함을 볼 수 있고, Birchfield 알고리즘과 제안된 알고리즘은 보다 개선된 결과를 보여주고 있다. Test2 영상에서는 밝기 변화가 없는 배경으로 인하여 종래의 알고리즘과 Cox 알고리즘에서 많은 에러가 발생함을 보여 준다. Test3, Test4 영상은 수렴카메라 모델로서 획득된 영상이기 때문에 카메라 사이의 거리를 크게 하더라도 공통시야를 일정하게 유지할 수 있고, 이로 인하여 삼각법에 의한 깊이 정보 추출 시 거리 측정의 정확성을 높일 수 있다. 이렇게 획득된 경우의 영상은 음의 변이값을 가질 수도 있다. 이와 같은 경우, 변이 영상의 구성은 우선 전체 변이 값들에 최소의 변이 값의 크기만큼 더하고, 이렇게 수정된 변이값들을 8비트 그레이 레벨(0~255)로 정규화한 데이터를 사용하였다. 이 두 영상에서는 특히 Birchfield 알고리즘에서 문제점이 발견되는데, 이 알고리즘은 수렴카메라 모델로서 획득된 영상에 대한 고려가 없기 때문이다. 이 두 영상에서도 제안된 알고리즘의 성능이 개선되었음을 확인할 수 있었다.

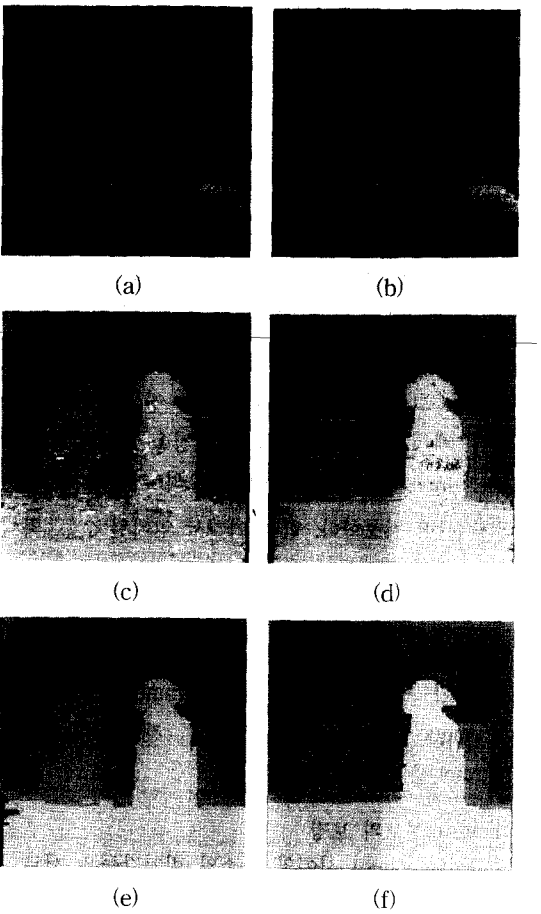
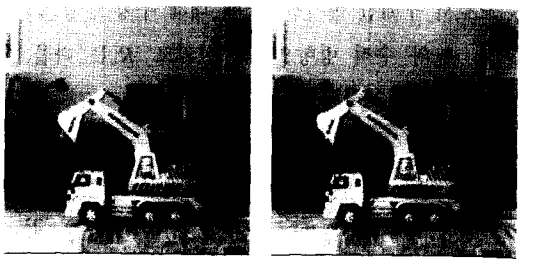
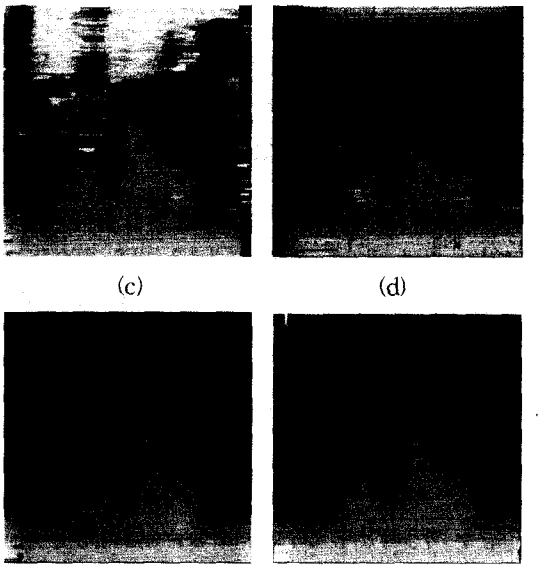


그림 4. 텍스쳐한 배경에서 평행 카메라 모델로 획득한 Test1 영상
(a), (b) 좌·우 원 영상
(c), (d), (e) 종래의 방법, Cox, Birchfield 방법에 의한 결과
(f) 제안된 방법에 의한 결과

Fig. 4. Test1 stereo images with textured background and parallel camera model.
(a), (b) Left and Right images.
(c), (d), (e) Results of convertional, Cox and Birchfield method.
(f) Result of proposed method.



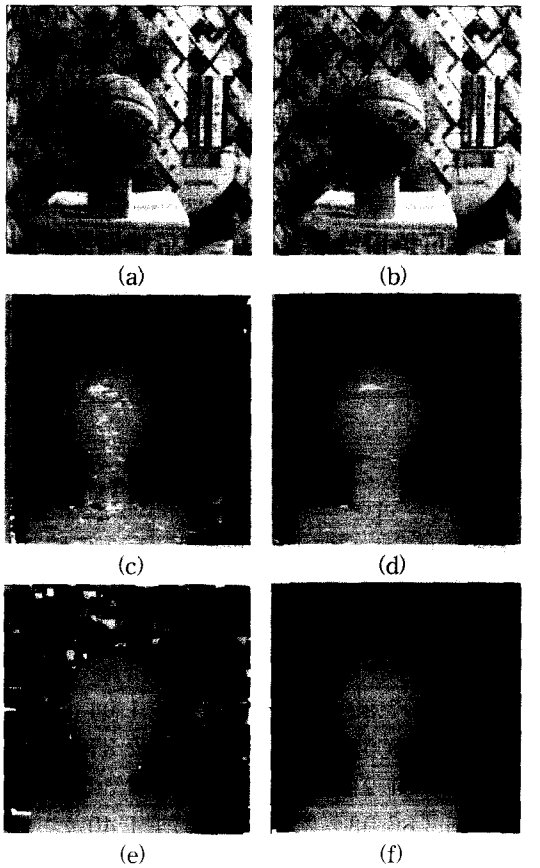


(c) (d)

(e) (f)

그림 5. 밝기 변화가 없는 배경에서 평행 카메라 모델로 획득한 Test2 영상

Fig. 5. Test2 stereo images with untextured background and parallel camera model.



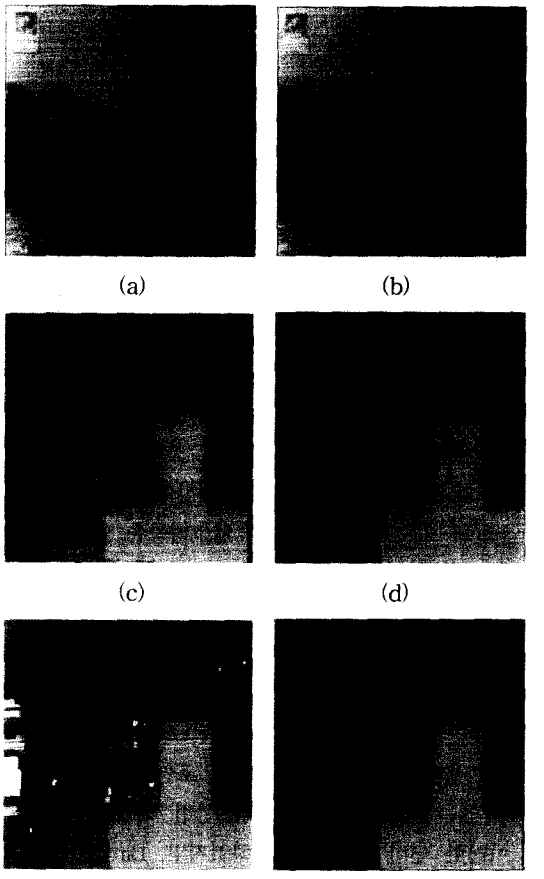
(a) (b)

(c) (d)

(e) (f)

그림 6. 텍스쳐한 배경에서 수렴 카메라 모델로 획득한 Test3 영상

Fig. 6. Test3 stereo images with textured background and convergent camera model.



(a) (b)

(c) (d)

(e) (f)

그림 7. 밝기 변화가 없는 배경에서 수렴 카메라 모델로 획득한 Test4 영상

Fig. 7. Test4 stereo images with untextured background and convergent camera model.

실험 결과를 보다 객관적으로 비교하기 위하여 정합률과 오차율을 도입하였다.[3] 정합률은

$$m=\frac{1}{N_x\times N_y}\sum_{y=0}^{N_y-1}\sum_{x=0}^{N_x-1}k(x,y)$$

$$where,\ k(x,y)=\begin{cases}1 & \text{if } |I_R(x,y)-I_L(x+D(x,y),y)|\le\theta\\ 0 & otherwise\end{cases} \tag{7}$$

과 같이 정의하였다. 여기서 θ는 최대 밝기차의 5%, 즉 12로 설정하였다. 이것은 정합률 m이 전체 영상에

대해 좌·우 정합 화소의 밝기차가 12 이내인 정합쌍의 비율을 의미한다.

또한, 오차율은

$$e=\frac{1}{N_x\times N_y}\sum_{y=0}^{N_y-1}\sum_{x=0}^{N_x-1}\frac{|I_R(x,y)-I_L(x+D(x,y),y)|}{I_{\max}} \tag{8}$$

와 같이 정의하였다. 오차율 e는 좌·우 정합 화소의 평균적인 밝기차를 나타낸다.

표Ⅱ는 4가지의 시험 영상에 대한 각 알고리즘별 정합률과 오차율을 나타낸 것이다. 표Ⅱ에 나타낸 바와 같이 제안된 알고리즘이 기존의 알고리즘에 비해 정합률과 오차율에서 우수함을 알 수 있었다.

한편 각 알고리즘은 Pentium PC-350 MHz에서 수행하였을 때의 시간을 표시하였다. 종래의 알고리즘과 Cox 알고리즘의 대략적인 계산량은 $n\cdot M^2$이고, Birchfield 알고리즘과 제안된 알고리즘은 $n\cdot M\cdot\log M$이다. 여기서 n은 주사선에 대한 화소의 수, M은 누적 비용 계산시 한번에 이동 가능한 최대 검색 경로를 나타낸다. 그리고 검색 영역의 제한에 대한 사항은 포함하지 않은 상태이다. Birchfield 알고리즘은 각 화소에 대한 누적 비용 계산시 중복된 경로에 대한 비용 계산은 생략하는 기법을 도입함으로써 계산시간을 줄일 수 있는 것으로 알려져 있다. 한편, 제안된 알고리즘은 이러한 연산 기법을 도입하였지만, Birchfield 알고리즘에 비해 검색 영역이 2배이고, 누적 비용 계산시 가중치를 부여했기 때문에 Birchfield 알고리즘보다는 다소 수행속도가 느리다. 그러나 기존의 동적계획법을 이용한 알고리즘에 비해서는 수행속도가 상대적으로 빠름을 알 수 있었다.

Ⅳ. 결 론

본 논문에서는 스테레오 영상으로부터 깊이 정보를 추출하기 위해 에지를 강조한 새로운 비용 함수를 정의하여 동적계획법을 적용시키는 스테레오 정합 알고리즘을 제안하였다. 제안된 알고리즘의 성능 평가를 위해 동적 계획법을 채택한 기존 발표된 알고리즘들과 정합 결과를 비교하였다. 이때 알고리즘의 내구성을 평가하기 위해 평행 카메라 모델뿐만 아니라 수렴 카메라 모델로 획득된 다양한 스테레오 영상에 대해 각 알고리즘을 적용하였다. 그 결과 제안된 알고리즘은 기존 알고리즘에서 발생한 불연속 지점에서의 무뎌짐 현상을 크게 보완하였고, 밝기 변화가 없는 영역에서 발생하는 다수의 정합 에러 등을 개선함을 볼 수 있었다. 또한 광축이 평행하지 않고 물체에 수렴하는 카메라 모델로 획득한 영상에 대해서도 평행 카메라 모델로의 좌표 변환 없이 제안된 알고리즘을 적용한 결과, 비교 알고리즘들 보다 우수한 결과를 보였다. 객관적 성능 비교를 위해 정합률과 오차율을 도입하여 수치적으로 비교한 결과에서도 제안된 알고리즘의 개선된 성능을 확인할 수 있었다.

본 논문에서 제안한 알고리즘은 변이값을 정수로 획득하게 되어있다. 변이값을 실수 값으로 획득하게 하면 더욱 정밀한 변이 정보를 획득할 수 있을 것이다. 이러한 데이터는 3-D 영상 재구성 작업에 보다 유용한 자료로 쓰여질 수 있을 것이고, 향후 과제이다.

표 2. 정합률(m)과 오차율(e)에 의한 알고리즘의 성능 비교
Table 2. Performance of Algorithms for Matching and Error Rate[%].

Algorithm / Image	Conventional Method[7]		Cox's Algorithm[8]		Birchfield's Algorithm[9]		Proposed Algorithm	
	m	*e*	*m*	*e*	*m*	*e*	*m*	*e*
Test1	95.89	1.25	96.86	1.08	97.99	0.92	97.51	1.04
Test2	85.93	2.96	86.79	3.23	89.59	2.51	89.98	2.37
Test3	94.22	1.41	94.59	1.45	61.21	6.31	95.70	1.30
Test4	93.85	1.40	93.85	1.55	86.03	2.62	94.67	1.40
Time[sec]	28.5		29.4		3.5		6.3	

참 고 문 헌

[1] R. Koch, "3-D Surface reconstruction from stereoscopic image sequences," *in Proc. ICCV'95*, Cambridge, MA., USA, Jun. 1995.

[2] A. Ruo and H. Burkhardt. "An intensity-based cooperative bidirectional stereo matching with simultaneous detection of discontinuities and occlusions," *IJCV*, 15, pp. 171-188, Jul. 1995.

[3] 주재흠, 신홍철, 강창순, 이상욱, 남기곤, "폐색 영역을 고려한 밝기 기반 쌍방향 스테레오 정합," 전자공학회 논문지, *제 36권 S편, 제 5호*, 1999

[4] H. H. Baker and T. O. Binford, "Depth from edge and intensity based stereo," *in Proc. IJCAI'81*, Vancouver, B. C., Canada, pp. 961-986, Aug. 1981.

[5] Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search," *IEEE Trans. on PAMI*, 7(3), pp. 139-154, Mar. 1985.

[6] 강창순, 에지변이를 기반으로 하는 동적계획법에 의한 스테레오 정합, 박사학위 논문, 부산대학교, 1997

[7] 김종덕, 강창순, 이상욱, 남기곤, "영역 분할에 의한 동적계획법을 이용한 스테레오 정합," 전자공학회 논문지, *제 34권 S편, 제 6호*, pp. 11-20, 1997

[8] I. J. Cox, S. L. Hingorani and S. B. Rao, "A Maximum likelihood Stereo Algorithm," *Computer Vision and Image Understanding*, 64(3), pp. 542-567, May, 1996.

[9] S. Birchfield and C. Tomasi, "Depth discontinuities by pixel-to-pixel stereo," *in Proc. ICCV'98*, Bombay, India, pp. 1073-1080, Jan. 1998.

[10] S. Birchfield and C. Tomasi, "A pixel dissimilarity measure that is insensitive to image sampling," *IEEE Trans. on PAMI*, 20(4), pp. 401-406, Apr. 1998.

[11] N. Alvertos, D. Brzakovic and R. C. Gonzalez, "Camera geometries for image matching in 3-D machine vision," *IEEE Trans. on PAMI*, 11(9), pp. 897-915, Sep. 1989.

[12] D. Geiger, B. Ladendorf and A. Yuille, "Occlusions and binocular stereo," *IJCV*, 14, pp. 211-226, Apr. 1995.

저 자 소 개

朱 在 欽(正會員) 第 36卷 S編 第 5號, 參照
현재 지산대학 전자계산과 조교수.

吳 宗 奎(正會員)
1998년 2월 부산대학교 전자공학과 학사. 1999년 9월~부산대학교 전자공학과 석사과정. 주관심분야는 영상처리 및 컴퓨터 비젼, 가상현실, 멀티미디어

薛 盛 旭(正會員) 第 36卷 S編 第 3號, 參照

李 哲 憲(正會員) 第 36卷 S編 第 3號, 參照
현재 양산대학 사무자동화과 전임강사

南 基 坤(正會員) 第 36卷 S編 第 3號, 參照
현재 부산대학교 전자공학과 교수. 부산대학교 컴퓨터 및 정보통신 연구소 연구원