

논문 2017-54-2-13

# 폐색 영역을 고려한 시간 축 스테레오 매칭

## ( Temporal Stereo Matching Using Occlusion Handling )

백 으 뜸\*, 호 요 성\*\*

( Eu-Tteum Baek<sup>Ⓒ</sup> and Yo-Sung Ho )

### 요 약

스테레오 정합은 두 영상의 색상과 공간 유사성이 최대가 되는 지점을 찾아 깊이 정보를 예측한다. 그런데, 두 시점 사이에 발생하는 폐색 영역으로 잘못된 깊이 정보를 얻게 되고, 폐색 영역을 보완하지 않은 깊이 정보는 시간 축 스테레오 매칭에서 잡음을 전파하는 문제를 일으킨다. 본 논문은 폐색 영역을 보완하여 시간 축 상에서 발생하는 잡음의 전파를 줄이고, 정확한 깊이 정보를 공유하여 스테레오 매칭의 정확성을 높이는 방법을 제안한다. 제안한 알고리즘을 수행하기 위해 색상과 공간의 유사성을 계산하는 함수를 정의하여 초기 깊이 정보를 예측하고, 세 가지 제약사항을 고려한 에너지 함수를 세워 (EM: expectation maximization) 으로 폐색 영역을 구한 뒤, 동적 프로그래밍 방법으로 예측된 폐색 영역을 보정한다. 끝으로, 이전 시점에 구해진 정확한 깊이 정보를 사용하여 시간 축 스테레오 매칭을 수행한다. 실험을 통해 제안한 알고리즘이 기존의 폐색 영역 예측 방법보다 우수한 성능을 가지는 것을 알 수 있었으며, 시간 축 정보를 고려하지 않은 스테레오 매칭 방법보다 정확한 결과를 얻는 것을 확인할 수 있었다.

### Abstract

Generally, stereo matching methods are used to estimate depth information based on color and spatial similarity. However, most depth estimation methods suffer from the occlusion region because occlusion regions cause inaccurate depth information. Moreover, they do not consider the temporal dimension when estimating the disparity. In this paper, we propose a temporal stereo matching method, considering occlusion and disregarding inaccurate temporal depth information. First, we apply a global stereo matching algorithm to estimate the depth information, we segment the image to occlusion and non-occlusion regions. After occlusion detection, we fill the occluded region with a reasonable disparity value that are obtained from neighboring pixels of the current pixel. Then, we apply a temporal disparity estimation method using the reliable information. Experimental results show that our method detects more accurate occlusion regions, compared to a conventional method. The proposed method increases the temporal consistency of estimated disparity maps and outperforms per-frame methods in noisy images.

**Keywords** : 폐색 영역, 스테레오 매칭, 그래프컷, EM, GMM, 동적 프로그래밍

### I. 서 론

스테레오 매칭은 컴퓨터 비전 분야에서 널리 연구되는 주제로서, 3차원 영화, 3차원 프린팅, 3차원 복원, 그리고 객체 인식 등과 같은 응용분야에서 사용된다. 깊이 정보는 카메라와 객체 사이의 거리 값을 나타내는데, 일반적으로, 능동형 깊이 카메라, 수동형 깊이 카메라, 그리고 혼합형 깊이 카메라를 통해 깊이 정보를 획득한다.

\* 학생회원, \*\* 평생회원 광주과학기술원 전기전자컴퓨터공학부 (Gwangju Institute of Science and Technology, School of Electrical Engineering and Computer Science)  
Ⓒ Corresponding Author (E-mail : hoyo@gist.ac.kr)

Received ; July 26, 2016

Revised ; December 19, 2016

Accepted ; January 24, 2017

능동형 깊이 카메라는 빛을 쏘아서 반사되어 오는 시간을 측정하여 거리를 계산하거나 적외선 점 패턴을 물체에 투영하여 점 패턴의 특성을 분석하여 깊이 정보를 획득한다<sup>[1]</sup>. 수동형 깊이 카메라는 두 대 혹은 여러 대의 색상 카메라로부터 획득한 영상의 상관관계를 계산하여 깊이를 획득한다<sup>[2]</sup>. 혼합형 깊이 카메라는 수동형 깊이 카메라와 능동형 깊이 카메라를 사용하여 깊이를 예측하는 방법이다<sup>[3]</sup>.

스테레오 정합 방법은 지역적 방법과 전역적 방법으로 나눌 수 있다. 지역적 방법은 화소 혹은 블록 사이의 상관관계를 계산하여 최소의 에너지 값을 찾는 방법이다. 일반적으로 지역적 방법은 거리 차 제곱 합계(SSD : sum of squared differences), 절대 거리 차 합계

(SAD: sum of absolute differences), 평균 상관계수 (NCC: normalized cross correlation)과 센서스 변환 (census transform) 등이 사용된다. 지역적 탐색 방법은 전역적 탐색 방법에 비해 빠르게 깊이 영상을 획득할 수 있는 장점이 있지만, 균일한 화소값을 갖는 평탄한 영역이나, 반복되는 패턴을 갖는 영역, 혹은 폐색 영역에서 정확한 대응점을 찾는 데 한계가 있다. 이러한 문제점을 해결하기 위해 정합 비용을 계산한 후에 비용을 정제하는 방법들이 소개되었다. 결합형 양방향 필터<sup>[4]</sup>, 가이드 이미지 필터<sup>[5]</sup>, 그리고 교차 계층 구조를 이용한 정합 비용 정제<sup>[6]</sup> 방법 등이 사용되어 비용 불륨을 정제하여 정확한 깊이 정보를 획득한다. 전역적 방법은 지역적 방법과는 다르게 정합 비용의 전체적인 에너지를 고려하여 깊이 정보를 예측하는 방법이다. 전역적 에너지 함수를 최적화하는 방법은 동적 프로그래밍<sup>[7]</sup>, 그래프 컷<sup>[8]</sup>, 상수 공간 신뢰 확산<sup>[9]</sup> 등이 사용된다.

폐색 영역은 기준 영상의 시점에서 객체나 장면이 가려져서 목적 영상의 시점에서 볼 수 없는 영역을 말한다. 이로 인해, 스테레오 정합은 정확한 깊이 정보를 얻을 수 없다. 그림 1은 폐색 영역과 나타내는데, 객체에 의해 시야가 가려진 폐색 영역 부분은 점선으로 표시했고, 시야가 가려지지 않은 비 폐색 영역 부분은 빨간색으로 표시했다.

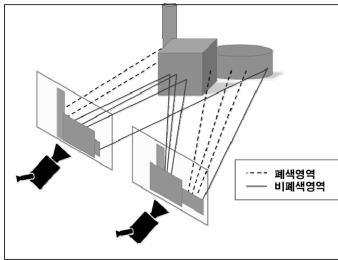


그림 1. 폐색 영역과 비 폐색 영역  
Fig. 1. Occlusion and non-occlusion.

폐색 영역을 검출하는 대부분 알고리즘은 순서 제약 사항과 유일성 제약사항을 사용하여 에너지 함수를 세우고 최적화하는 방법을 사용한다<sup>[10-11]</sup>. 일반적으로, 폐색 영역을 검사할 때, 대조조사 방법(cross-checking)을 사용한다. 하지만, 이전에 제안된 폐색 영역을 보완하는 방법은 작은 구멍이나 물체 혹은 기울어진 물체가 있는 부분에서 좋은 성능을 내지 못한다. 이러한 한계를 극복하기 위해 폐색 영역을 처리하는 다양한 알고리즘이 소개되고 있는데, Kolmogorov et al. 은 폐색 영역을

고려해 (MRF: Markov random field) 기반 에너지 함수를 만들어, 그래프 컷으로 최적화하는 방법을 제안했다<sup>[12]</sup>. 에너지 함수에 폐색 영역을 고려한 패널티 항을 추가하여 이전의 방법보다 좋은 결과를 얻었지만, 유일성 제약조건만을 고려한 식을 세웠다는 단점이 있다. 이러한 단점을 해결하기 위해 Jang et al. 은 초기 깊이 정보를 사용하여 폐색 영역을 검출하고, 주변 깊이 정보를 사용하여 폐색 영역을 채우는 방법을 소개하였다. 폐색 영역을 처리하는 데 좋은 성능을 보였지만 평활화 제약사항을 고려하지 않았다는 단점이 있다. 본 논문에서는 평활화 제약사항을 고려한 에너지 함수를 활용하여 폐색 영역으로 인하여 발생하는 깊이 정보의 부정확성을 해결하고, 시간 축 상에서 깊이 정보를 공유하여 더욱 정확한 깊이 정보를 획득하는 스테레오 정합 방법을 소개한다.

## II. 스테레오 매칭

스테레오 매칭 방법은 카메라와 물체 사이의 거리 정보를 얻는 데 사용되는 방법으로, 일반적으로 변위를 계산하기 위하여 정합 에너지 함수를 정의하고, 에너지 함수를 계산하여 그 값이 최소가 될 때의 변위를 선택하는 방법이다. 본 논문에서는 MRF-MAP 기반의 에너지 함수를 사용하여 변위 값을 획득하였다. MRF-MAP 기반의 에너지 함수는 두 가지 요소를 포함하고 있다. 첫 번째는 두 영상에서 변위만큼 떨어진 색상은 서로 유사성을 갖는다는 것이며, 두 번째는 주변의 픽셀들은 비슷한 변위를 가질 것이라 가정한다. 이 두 가지 가정은 정합 에너지 함수에서 데이터 항(Data)과 평활화 항(Smoothness)으로 표현되며 정합 에너지 함수는 다음 식과 같이 정의된다.

$$E(d) = \sum_s D_s(d_s) + \sum_{s,t \in N(s)} S_{s,t}(d_s, d_t) \quad (1)$$

$D_s()$ 는 데이터 항이고  $S_{s,t}()$ 는 평활화 항이고,  $d_s$ 는 변위를 나타낸다. 본 논문에서 데이터 항은 SAD를 사용하였는데, 다음과 같이 정의된다.

$$D_s(d_s) = \min(SAD(x, y, d), T_d) \quad (2)$$

$$SAD(x, y, d) = \sum_{i,j \in W} |I_r(x+i, y+j) - I_t(x+i+d, y+j)|$$

위 식에서  $I_r$ 과  $I_t$ 은 각각 기준 영상과 목표 영상을 나타내고  $T_d$ 은 절단 값을 의미한다.

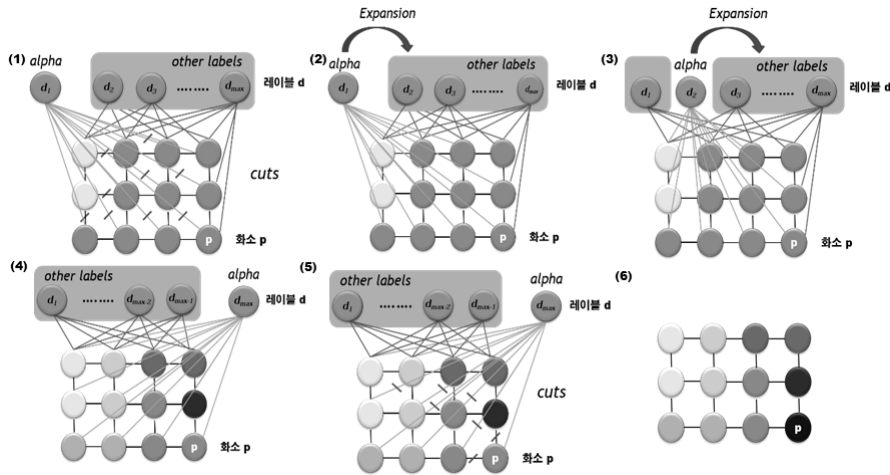


그림 2. alpha expansion 알고리즘. (1) 과 (5)는 min cuts이고 (2)와 (3)은 expansion이고 (6) 결과  
Fig. 2. alpha expansion algorithm. (1) and (5) min cuts, (2) and (3) expansion, (6) result.

평활화 항은 L2-norm을 사용하는데, 다음 식과 같다.

$$S_{s,t}(d_s, d_t) = \min(\lambda \|d_s - d_t\|, T_s) \quad (3)$$

$T_s$ 는 절단 값이고,  $\lambda$ 는 가중치 값을 의미한다.

본 논문에서는 alpha-expansion 알고리즘을 사용해 에너지 함수를 최적화하였다<sup>[8]</sup>. alpha-expansion 알고리즘은 min cut/max flow 알고리즘을 사용하여, multi labeling을 하는 최적화 방법이다. 그림 2는 alpha-expansion 알고리즘을 수행하는 방법을 보여준다. 먼저, alpha 노드와 나머지 노드를 정하고, min cut/max flow 알고리즘을 수행하여 alpha 노드에 속한 화소  $p$ 를 예측한다. 그리고, expansion 단계에서 alpha 노드와 나머지 노드 중 하나를 바꾸고 다시 min cut/max flow 알고리즘을 순차적으로 수행하여 최적의 레이블을 찾는 방법을 alpha-expansion 알고리즘이라고 한다.

### III. 폐색 영역 보정

#### 1. 폐색 영역 검출

폐색 영역은 정확한 일치점을 찾는 주요한 사항이다. 폐색 화소는 기준 영상에서는 보이지만 목표 영상에서는 보이지 않는 부분을 의미한다. 본 논문에서는 폐색 영역을 찾기 위해 순서 제약사항, 유일성 제약사항, 색상 유사성 제약사항을 사용하여 에너지 함수를 세운다. 다음은 폐색 영역 검출을 위한 에너지 함수를 나타낸다.

$$E_T(D^R) = \lambda_1 E_g(D^R) + \lambda_2 E_o(D^R, D^T) + \lambda_3 E_c(D^R) \quad (4)$$

$\lambda_1$ ,  $\lambda_2$ , 그리고  $\lambda_3$ 는 각각 가중치를 나타내고,  $D^R$ 와

$D^T$ 는 각각 기준 영상과 목표 영상의 변위를 나타낸다.

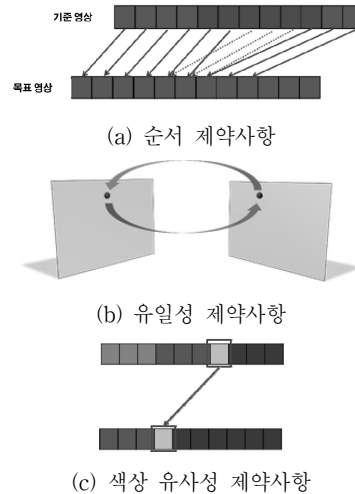


그림 3. 폐색 영역 제약사항

Fig. 3. Constraints for occlusion.

그림 3(a)에서 나타내듯 순서 제약사항은 기준 영상에서 목표 영상으로 위핑할 때, 단 하나의 화소만 일치점으로 나타내야 하고, 만약 두 개 이상의 화소가 겹치게 위핑되는 화소를 폐색 영역으로 여긴다.  $E_g(D^R)$ 는 순서 제약 사항 항으로 다음 식과 같다.

$$\begin{cases} E_g(D^R) = 1, & \text{if } D^R = \text{occlusion candidate} \\ E_g(D^R) = \alpha, & \text{else if } D^R = \text{largest value of candidate} \\ E_g(D^R) = 0, & \text{otherwise} \end{cases} \quad (5)$$

식 (5)의  $\alpha$ 는 1보다 작은 양수이다. 만약 임의의 화소가 폐색 영역 후보이면 에너지값은 1이고, 폐색 영역

후보 중에 변위 값이 가장 큰 값은  $\alpha$ 이다. 그림 3(b)는 유일성 제약사항을 나타내는데 특정한 화소는 유일한 일치점을 가져야 한다는 제약사항이다.  $E_o(D^R, D^T)$ 는 유일성 제약사항 항으로 다음 식과 같다.

$$\begin{cases} E_o(D^R, D^T) = 0, \text{ if } D^R(x_z) = D^T(x_z - D^R(x_z)) \\ E_o(D^R, D^T) = 1, \text{ otherwise} \end{cases} \quad (6)$$

그림 3(c)는 색상 유사성 제약사항을 나타내는데 일치점은 다른 화소에 비해 색상이 유사해야 한다는 제약사항이다.  $E_c(D^R)$ 는 색상 유사성 제약사항 항으로 다음 식과 같다.

$$E_c(C^R, C^T) = \frac{1}{W_C} \sqrt{\sum_{i=1}^3 C_i^R(x_z) - C_i^T(x_z - D^R(x_z))} \quad (7)$$

$C^R$ 와  $C^T$ 는 각각 기준 색상 영상과 목표 색상 영상을 나타내고,  $W_C$ 는 정규화 값이다.

페색 영역을 찾기 위해 식(4)으로 얻어진 에너지 값의 (GMM: Gaussian mixture model)을 구한다. GMM은 데이터들의 분포 특성을 분석하는데 적합한 확률 모델로, 가중치가 부여된  $k$ 개의 확률밀도함수의 선형 결합이다. 일반적으로 GMM을 구하기 위해 EM을 사용한다<sup>[13]</sup>. EM은 최대 우도를 찾는 파라메트릭 최적화 방법의 일종으로, 매개변수에 관한 추정값으로 로그가능도의 기댓값을 계산하는 expectation 단계와 이 기댓값을 최대화해주는 변수값을 찾는 maximization 단계를 반복적으로 수행하는 최적화 방법이다. 본 논문에서는 2개의 파라미터를 가지는 GMM을 찾는다.



그림 4. 두 종류의 페색 영역  
Fig. 4. Two kinds of occlusion.

## 2. 페색 영역 보정

본 절은 앞 절에서 예측한 페색 영역에 근접한 비 페색 영역의 변위 값을 사용하여 페색 영역 보정한다. Jang et al.이 제안한 논문에서는 페색 영역을 두 가지로 분류하는데, 일반적인 영역과 극좌의 영역이다<sup>[2]</sup>.

그림 4는 두 종류의 페색 영역을 보여주는데, 노란색으로 표시된 영역을 극좌의 페색 영역이라고 하며, 좌측 영상에서는 보이지만 오른쪽 영상에는 보이지 않는 좌측 맨 끝 부분을 의미하고, 나머지 부분을 일반적인

페색 영역이라고 하며, 객체로 가려져서 보이지 않는 영역을 의미한다. 두 영역의 특성을 고려하여 극좌의 페색 영역은 오른쪽에서 값을 채워 나가고 일반적인 페색 영역은 전체적인 변위 값을 고려하여 페색 영역을 보정한다. 본 논문에서 페색 영역으로 검출된 부분을 보정하기 위해 MRF-MAP 기반의 에너지 함수를 사용하는데, 식은 아래와 같다.

$$E_{occ}(d) = E_d(d) + \lambda_{occ} \cdot E_s(d) \quad (8)$$

$E_d(d)$ 은 데이터 항으로  $E_s(d)$ 은 평활화 항이다. 데이터 항은 색상과 거리를 고려하였는데, 다음 식과 같다.

$$E_d(s, d) = \frac{1}{dist(s, t)} \exp\left(-\frac{dif_{s,t}}{\sigma^2}\right) \text{ s.t. } t \in \text{non-occ} \quad (9)$$

$$\text{where } dif_{s,t} = \sum_{c \in \{R, G, B\}} |I_c(s) - I_c(t)|$$

$dist(s, t)$ 는 기준점  $s$ 에서  $t$ 까지의 유클리디안 거리 값을 나타내고  $dif_{s,t}$ 는 색상의 유사성을 나타내는 식이다. 평활화 항을 사용하기 위해 식(3)과 같은 L2-norm을 사용하였다.

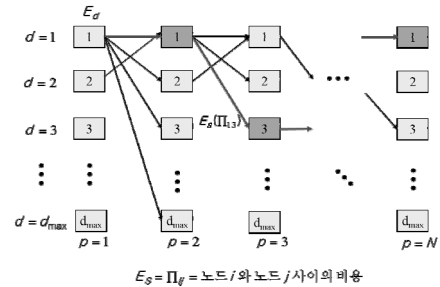


그림 5. 동적 프로그래밍  
Fig. 5. Dynamic programming.

본 논문에서는 식(8)을 최적화하여 페색 영역을 보정하기 위하여 동적 프로그래밍을 사용한다. 페색 영역의 특성상 단방향의 에너지 정보를 사용해야 더 좋은 성능을 보이기 때문에 동적 프로그래밍을 사용한다<sup>[2]</sup>. 그림 5는 동적 프로그래밍을 사용하여 에너지의 최솟값을 찾는 과정을 보여준다. 그림 5에서  $p$ 는 화소의 노드를 나타내며  $d$ 는 변위를 나타낸다. 먼저 현재 화소의 데이터 항의 값  $E_d$ 를 구하고, 평활화 값  $E_s$ 를 합하여 그 값이 최소가 되는 변위를 찾는다. 마지막으로 가이드 이미지 필터 방법을 사용하여 나머지 잡음을 보정한다<sup>[14]</sup>.

## IV. 시간 축 스테레오 매칭

### 1. 움직임 예측

시간 축 상에서 이전 시점의 정확한 정보를 사용하기 위해서는 두 시점 사이의 움직임을 예측해야 한다. 본 논문에서는 정확한 움직임 정보를 얻기 위해 optical flow 방법으로 각 화소의 움직임을 예측하였다.

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (10)$$

$I(x, y, t)$ 는 시간  $t$ 에서의 색상 값이고,  $(\Delta x, \Delta y)$ 는 속도를 나타낸다. 위의 식을 테일러 급수로 표현하면 아래와 같다.

$$\begin{aligned} I(x + \Delta x, y + \Delta y, t + \Delta t) \\ = I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t + \text{Higher orders} \end{aligned} \quad (11)$$

식 (10)과 식 (11)를 아래의 식과 같이 속도에 관한

$$\begin{aligned} \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0 \\ \frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} = 0 \end{aligned} \quad (12)$$

식으로 바꾸어 사용할 수 있다.

$V_x$ 와  $V_y$ 는 각각  $x$ 와  $y$ 에 대한 optical flow이고  $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t}$ 는 편미분을 나타낸다. 각각의 편미분을  $I_x, I_y, I_t$ 으로 바꾸어 식을 세울 수 있다.

$$I_x V_x + I_y V_y = -I_t, \quad \nabla I^T \cdot \vec{V} = -I_t \quad (13)$$

본 논문은 식(13)을 풀기 위해 pyramidal Lucas-Kanade flow를 사용하였다<sup>[15]</sup>.

## 2. 시간 축 스테레오 매칭

본 절에서는 움직임 정보와 이전 시점의 깊이 정보를 활용하여 시간 축 스테레오 매칭을 하는 법을 소개한다. 본 논문에서는 이전 시점의 깊이 정보 폐색 영역 처리로 잡음이 제거되었고, 움직임 정보를 알고 있으므로 이전 시점의 깊이 정보가 정확하다고 가정하였다.

시간 축 정보를 고려한 스테레오 매칭 에너지 함수는 아래의 식과 같다.

$$\begin{aligned} E(d) = \sum_s g(x) D_s(d_s) + \sum_{s,t \in N(s)} S_{s,t}(d_s, d_t) \\ g(x) = 1 - \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(x-d_{t-1})^2}{2\sigma^2} \right) \end{aligned} \quad (14)$$

식 (14)에서  $g(x)$ 는 가우시안 가중치 함수로 이전 시점의 깊이 정보에 가중치를 주는 함수이고,  $d_{t-1}$ 는 이전 시점의 깊이 값을 나타낸다. 그림 8은 가우시안 가중치 함수의 값을 그래프로 나타내는데 이전 프레임의 깊이 값에 가까울수록 에너지가 낮아진다. 에너지 함수의 데

터 항과 평활화 항은 각각 식 (2)와 식 (3)을 사용하였고, alpha-expansion 알고리즘을 사용하여 최적화를 하였다.

## V. 실험 결과

정확한 실험 평가를 위해 폐색 영역 보정과 시간 축 스테레오 매칭을 각각 나누어 평가하였다. 폐색 영역 보정의 결과를 평가하기 위해 본 논문에서는 Middlebury에서 제공하는 ground truth 깊이 정보로 얻어진 폐색 영역과 ground truth 폐색 영역을 비교하였다<sup>[16]</sup>. 정량적 평가를 위해 본 논문에서는 (BPR: bad pixel rate)을 사용하였는데, 표 1과 그림 6에서 나타나듯이 제안한 방법이 Jang's 방법보다 폐색 영역을 더욱 정확하게 찾는 것을 알 수 있다.

표 1. 폐색 영역 평가  
Table1. Evaluation for occlusion map.

	Teddy	Cone	Venus
Jang's <sup>[2]</sup>	4.75	6.78	1.16
proposed	1.55	1.70	3.57

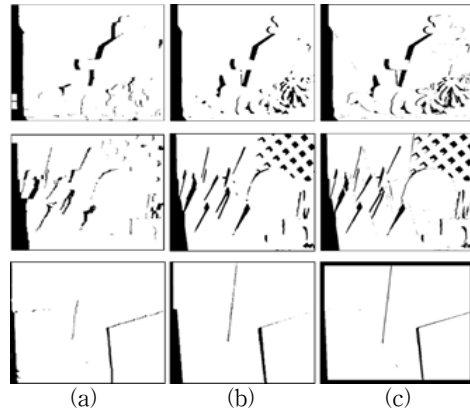


그림 6. 폐색 영역 검출: (a) Jang's method<sup>[2]</sup>, (b) 제안한 방법, (c) ground truth

Fig. 6. Occlusion detection: (a) Jang's method<sup>[2]</sup>, (b) Proposed, and (c) ground truth.

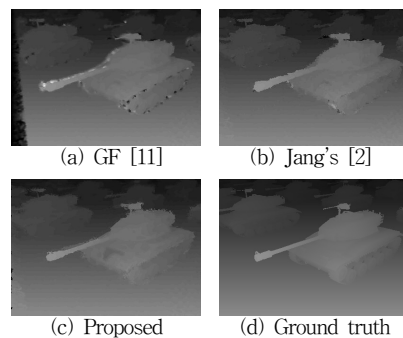


그림 7. 제안한 방법과 기존 방법 비교  
Fig. 7. Comparison of proposed and other methods.

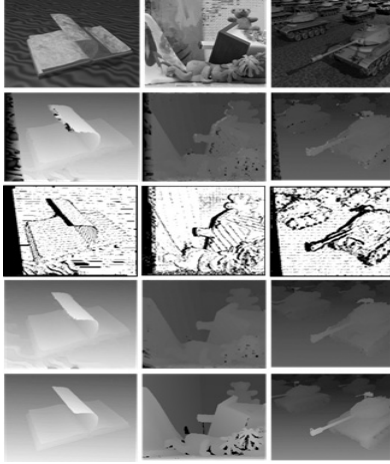


그림 8. 깊이 정보 예측과 폐색 영역 보완  
Fig. 8. Disparity estimation and occlusion refinement.

그림 7은 기존의 방법과 제안한 방법의 스테레오 정합 결과를 비교한 그림을 보여준다. 가이드 이미지 필터를 사용한 방법은 폐색 영역을 처리하지 않았기 때문에 객체 주변에 잡음이 발생한다. 반면 Jang's 방법은 폐색 영역을 개선하였다. 제안한 방법은 폐색 영역을 개선하며, 에지 부분을 잘 보존한다는 장점이 있다.

그림 8은 Middlebury 제공하는 스테레오 영상<sup>[16]</sup>과 캠브리지 대학의 컴퓨터 비전 랩에서 제공하는 데이터 셋<sup>[17]</sup>을 사용해 깊이 맵을 획득하여 폐색 영역을 검출하고, 폐색 영역을 개선한 결과를 보여준다. 두 번째 줄은 초기 깊이 맵이고, 세 번째 줄은 초기 깊이 영상을 이용하여 폐색 영역을 예측한 영상이다. 네 번째 줄은 폐색 영역을 보정한 깊이 영상이고, 마지막 줄은 ground truth 영상이다. 그림 8을 통해 제안한 알고리즘이 폐색 영역과 잘 검출할 뿐 아니라 잘못 접합된 깊이 정보도 잘 검출 하는 것을 확인 할 수 있다. 또한, 예측된 폐색 영역 보정 알고리즘을 통해 정확한 깊이 맵을 획득한다.

표 2는 기존의 방법과 제안한 방법의 오차율 결과를 보여 주는데, 기존의 방법에 비해 제안한 방법의 BPR이 더 낮은 것을 확인 할 수 있다. 표 3은 캠브리지 대학의 컴퓨터 비전 랩에서 제공하는 데이터 셋을 활용하여 시간축 깊이 정보의 (MAD: Mean absolute difference)를 계산한 결과를 보여 준다<sup>[1, 7, 9]</sup>. 표 3을 통해 기존 방법보다 제안한 스테레오 매칭 방법이 정확한 깊이 정보를 생성하는 알 수 있다.

그림 9는 시간 축 스테레오 매칭의 MAD를 계산한 결과를 그래프로 표현한 것으로, 'book'과 'tanks' 데이

표 2. 결과 영상의 오차율 분석(BPR)  
Table2. Error rate comparison.

Algorithm		CSBF <sup>[6]</sup>	Jang's <sup>[2]</sup>	Proposed	GC+occ <sup>[8]</sup>
Tsukuba	nonocc	2.00	1.423	3.38	1.19
	all	4.17	2.30	3.43	2.01
	disc	10.50	7.94	16.6	6.24
Venus	nonocc	1.48	0.91	1.19	1.64
	all	3.11	1.54	1.93	2.19
	disc	17.70	12.71	11.62	6.75
Cones	nonocc	5.98	4.96	5.70	5.36
	all	16.50	12.70	12.58	12.40
	disc	16.00	14.44	16.42	13.00
Teddy	nonocc	11.10	6.34	7.81	11.20
	all	20.20	13.62	13.40	17.40
	disc	27.50	17.59	22.75	19.80

표 3. 시간 축 MAD 비교  
Table3. Comparison of Mean Absolute Difference of Sequence (Tanks)

No	Proposed	GF <sup>[5]</sup>	Jang's <sup>[2]</sup>	SG <sup>[9]</sup>
1	2.717	3.644	2.717	11.643
2	2.700	3.645	2.728	11.652
3	2.722	3.640	2.746	11.6491
4	2.766	3.657	2.753	11.6675
5	2.782	3.681	2.776	11.7343
6	2.779	3.718	2.805	11.7661
7	2.773	3.750	2.839	11.9887
8	2.809	3.806	2.885	12.1494
9	2.874	3.865	2.942	12.3425
10	2.907	3.910	2.989	12.6568
11	2.987	4.003	3.057	12.961
12	3.081	4.087	3.160	13.3963
13	3.182	4.206	3.253	13.794
14	3.250	4.279	3.331	14.0393
15	3.357	4.363	3.432	14.3933
16	3.410	4.432	3.495	14.7737
17	3.463	4.504	3.556	15.0365
18	3.447	4.520	3.599	15.3484

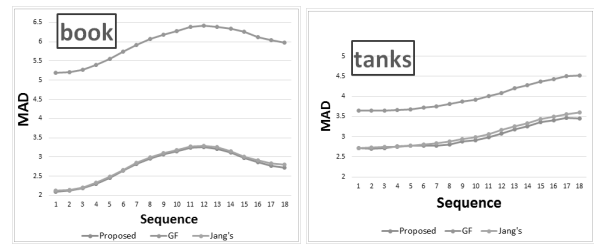


그림 9. Mean absolute difference (MAD)  
Fig. 9. Mean absolute difference.

터 셋을 사용하였다. 그래프에서도 나타나듯이 제안한 알고리즘이 기존의 방법보다 더 낮은 MAD를 보여준다.

#### IV. 결 론

본 논문은 폐색 영역과 시간 축 깊이 정보를 고려하여 스테레오 매칭의 정확성을 높이는 방법을 제안한다. 실험 결과를 통해 기존의 방법에 비해 제안한 방법이 폐색 영역을 잘 예측하며 폐색 영역의 홀 채움 역시 좋은 성능을 보이는 것을 확인하였다. 또한, 제안한 방법이 시간 축에서의 움직임을 고려하여 스테레오 정합을

하였기 때문에, 움직임을 고려하지 않은 스테레오 정합 방법보다 정확한 깊이 맵을 생성하는 것을 확인 할 수 있다. 제안한 방법을 다양한 3D 컴퓨터 비전 응용분야에 사용할 수 있다.

## REFERENCES

- [1] A. Frick, F. Kellner, B. Bartczak and R. Koch, "Generation of 3D-TV LDV-content with time of flight camera," Conf. 3DTV, pp. 45-48, 2009.
- [2] W. S. Jang, Y. S. Ho, "Efficient disparity map estimation using occlusion handling for various 3D multimedia applications," Trans. Consumer Electronics, vol. 57, no. 4, pp. 1937-1943, 2011.
- [3] E. K. Lee, Y. S. Ho, "Generation of high-quality depth maps using hybrid camera system for 3-D video," J. Visual Comm. Image Represent, vol. 22 no. 1, pp. 73-84, 2011.
- [4] K. J. Yoon and I. S. Kweon. "Adaptive support-weight approach for correspondence search," PAMI, vol. 28, no. 4, pp. 650-656, 2006.
- [5] A. Hosni, M. Bleyer, C. Rhemann, M. Gelautz, and C. Rother, "Real-time local stereo matching using guided image filtering," ICME, pp. 1-6, 2011.
- [6] K. Zheng, Y. Fang, D. Min, L. Sun, S. Yang, S. Yan, and Q. Tian, "Cross-scale cost aggregation for stereo matching," CVPR, pp. 1590-1597, 2014.
- [7] S. Birchfield and C. Tomasi, "Depth discontinuities by pixel-to-pixel stereo," J. Computer Vision, vol. 35, no. 3, pp. 269-293, 1999.
- [8] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," Conf. Computer Vision, pp. 508-515, 2001.
- [9] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Understanding belief propagation and its generalizations," Exploring Artificial Intelligence in the New Millenium, pp. 239-269, 2003.
- [10] A. F. Bobick and S. S. Intille, "Large occlusion stereo," J. Computer Vision, vol. 33 no. 3 pp. 1-20, 1999.
- [11] H. Ishikawa and D. Geiger, "Occlusions, discontinuities, and epipolar lines in stereo," Conf. Computer Vision, pp. 425-433, 1998.
- [12] T. Liu, P. Zhang, L. Luo, "Dense stereo correspondence with contrast context histogram, segmentation-based two-pass aggregation and occlusion handling," Advances in Image and Video Technology, pp. 449-461, 2009.
- [13] J. Bilmes., "A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models," Technical Report ICSI-TR-97-021, Univ. of Berkley, 1988.
- [14] K. He, J. Sun, and X. Tang, "Guided image filtering," Conf. Computer Vision, pp. 1-14, 2010.
- [15] J. Bouguet, "Pyramidal Implementation of the Lucas Kanade Feature Tracker: Description of the Algorithm," OpenCV Document, Intel, Microprocessor Research Labs, 2000.
- [16] D. Scharstein, R. Szeliski, and R. Zabih, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," Int. J. Computer Vision, vol. 47, no. 1, pp. 7-42, 2002.
- [17] C. Richardt, D. Orr, I. Davies, A. Criminisi, and N. A. Dodgson. "Real-time spatiotemporal stereo matching using the dual-cross-bilateral grid," Conf. Computer Vision, pp. 6311-6316, 2010.

## 저 자 소 개



백 은 뚝(정회원)  
2012년 전북대학교 컴퓨터공학과  
학사 졸업.  
2013년~2015년 광주과학기술원 정보  
통신공학부 석사.  
2015년~현재 광주과학기술원 전기  
전자컴퓨터공학부 박사 과정.

<주관심분야: 디지털 영상처리, 3D 영상합성>



호 요 성(평생회원)  
1981년 서울대학교 전자공학과 학사  
1983년 서울대학교 전자공학과 석사  
1989년 University of California,  
Santa Barbara, Department  
of Electrical and Computer  
Engineering 박사

1983년~1995년 한국전자통신연구소 선임연구원  
1990년~1993년 미국 Phillips 연구소, Senior Research  
Member

1995년~현재 광주과학기술원 전기전자컴퓨터공학부  
교수.

<주관심분야: 디지털 신호처리, 영상신호 처리 및  
압축, 멀티미디어 시스템, 디지털 TV와 고선명  
TV, MPEG 표준, 3차원 TV, 실감방송>