

# 배경이 제거된 마스크를 활용한 깊이지도 업샘플링 방법

김선호, 호요성  
광주과학기술원 전기전자컴퓨터공학부  
sunhokim@gist.ac.kr, hoyo@gist.ac.kr

## Depth Map Upsampling Method Using Subtracted Background Mask

Sunho Kim, Yo-Sung Ho  
School of Electrical Engineering and Computer Science  
Gwangju Institute of Science and Technology (GIST)

### 요 약

하나의 깊이 카메라로 촬영한 깊이지도는 워핑 (Warping) 및 업샘플링 (Upsampling) 과정을 통해 다양한 색상 카메라 시점에 대응되는 깊이지도로 변환될 수 있다. 그러나 이 과정에서 만들어지는 깊이지도는 경계 부분이 모호하게 출력되는 단점이 있다. 본 논문에서는 업샘플링된 깊이지도에 배경이 제거된 마스크를 활용해 모호한 경계 부분을 명확하게 하고, 필요한 객체의 깊이지도만을 생성하는 방법을 제안한다. 배경 제거 알고리즘에선 색상 공간의 변환과 SAD (Sum of absolute differences) 연산, 그리고 객체의 형태학적 팽창 (Dilation)을 활용한 연산을 수행한다. 깊이지도 업샘플링 알고리즘에선 3차원 워핑과 보간법을 통해 다시점 깊이지도를 생성하고, 이전 과정에서 만든 마스크를 이용하여 경계 부분을 명확하게 만든다.

### 1. 서 론

3차원 콘텐츠는 사용자에게 보다 실감나는 영상을 제공한다. 따라서 3차원 콘텐츠는 많은 사람이 접할 수 있는 교육, 게임, 대중매체와 같은 분야에서 활용되고 있다 [1].

이러한 3차원 콘텐츠를 제작하기 위해서는 기존의 2차원 콘텐츠에서 사용했던 색상 정보 외에 깊이에 대한 정보가 필요하다. 깊이 정보를 얻는 방법에는 깊이 카메라를 사용하는 방법과 스테레오 정합을 사용하는 방법이 있다. 스테레오 정합은 별도의 깊이 카메라를 사용하지 않고 서로 다른 두 시점의 색상 카메라로부터 촬영한 색상 영상들을 바탕으로 깊이 값을 예측하는 방법이다. 이 방법은 별도의 장비를 요구하지 않지만, 다소 복잡한 연산을 수행해야 한다는 단점이 있다.

반면에 깊이 카메라를 이용하면 빠르고 쉽게 원하는 깊이지도를 얻을 수 있다. 측정 가능 거리가 다소 제한적이며 객체의 색상과 환경적 요인에 따라 측정에 영향을 받는 문제가 있으나, 스테레오 정합과는 달리 별도의 복잡한 연산이 요구되지 않는 장점이 있다. 일반적으로 이 방법은 적외선 파동을 발사해 분출한 파동과 반사된 파동 간의 위상차를 통해 객체와 카메라 사이의 거리 정보를 계산하고, 이를 바탕으로 깊이지도를 생성한다 [2].

그러나 깊이 카메라를 통해 촬영한 깊이지도는 색상 카메라로 촬영한 색상 영상과 비교했을 때, 해상도와 초점 거리, 카메라의 위치, 주점 등 다양한 점에 대한 차이가 있다. 따라서 색상 영상의 이러한 조건들과 일치하는 깊이지도를 얻기 위해서는 카메라 캘리브레이션 (Calibration)

과정을 통해 얻은 카메라 인자를 활용해 3차원 워핑과 업샘플링을 수행해야 한다.



(a) 원본 영상



(b) 추출된 경계

(그림 1) Canny 경계 검출법을 이용한 경계 추출

깊이 카메라로 촬영한 깊이지도로 다른 시점의 깊이 영상을 만들 때, 단순한 업샘플링 방법만을 사용하면 깊이 영상 속 객체의 경계 부분이 모호하거나 빈 공간이 채워지지 않는다. 이러한 문제점을 개선하기 위해 그림 1과 같이 색상 영상의 경계 정보를 추출하고 그 정보를 이용해 업샘플링하는 방법도 제안되었다 [3, 4]. 그러나 그림 1(a)처럼 배경에 다양한 무늬가 있으면 경계 추출 과정에서 배경 속 무늬의 경계까지 검출된다. 따라서 이 경우엔 객체와 배경을 정확히 분리하기 힘들고 결과적으로 부정확한 업샘플링을 수행하게 된다.

본 논문에서는 색상 공간 변환을 이용한 배경 제거 기법을 통해 객체의 마스크를 추출하고, 추출된 마스크를 바탕으로 업샘플링된 깊이지도에서 더 명확하게 객체의 깊

이 정보를 얻는 방법을 제안한다.

$$T = \alpha T_{Huc} + \beta T_{Saturation} \quad (2)$$

## 2. 색상 영상의 배경 제거

제안하는 방법에서는 업샘플링된 객체의 깊이도를 색상 영상의 객체와 일치시키기 위해 배경이 제거된 마스크를 사용한다. 객체의 마스크를 얻기 위해 본 논문에서는 HSL 색상 공간으로의 변환과 SAD 연산, 그리고 형태학적 객체 팽창 연산을 활용한다.

### 2.1 색상 공간 변환을 이용한 색상 및 채도의 차이 구하기

일반적인 색상 영상은 RGB 색상 공간을 기반으로 하는 화소 값을 가진다. 본 논문에서는 영상의 색상 값과 채도 값을 얻기 위해 RGB 색상 공간 기반의 영상을 HSL 색상 공간 기반의 영상으로 변환해 배경 제거를 수행한다.

영상의 각 화소에 빨강, 초록, 파랑에 대한 값을 갖는 RGB 색상 공간과는 달리 HSL 색상 공간은 영상의 각 화소에 색상, 채도, 명도 값을 가지고 있다. 따라서 이 색상 공간은 RGB 색상 공간보다 직관적으로 화소의 색상 값에 접근할 수 있다 [4]. 식 (1)은 RGB 색상 공간을 HSL 색상 공간으로 변환하는 공식이다.

$$A_{max} = \max\left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255}\right)$$

$$A_{min} = \min\left(\frac{R}{255}, \frac{G}{255}, \frac{B}{255}\right)$$

$$H = \begin{cases} 60^\circ \times \left(\frac{G-B}{255(A_{max}-A_{min})} \bmod 6\right) & (A_{max} = \frac{R}{255}) \\ 60^\circ \times \left(\frac{B-R}{255(A_{max}-A_{min})} + 2\right) & (A_{max} = \frac{G}{255}) \\ 60^\circ \times \left(\frac{R-G}{255(A_{max}-A_{min})} + 4\right) & (A_{max} = \frac{B}{255}) \end{cases} \quad (1)$$

$$S = \begin{cases} 0 & (A_{max} = 0) \\ \frac{A_{max}-A_{min}}{A_{max}} & (A_{max} \neq 0) \end{cases}$$

$$V = A_{max}$$

식 (1)에서  $A_{Max}$ 는 R, G, B 값의 최댓값,  $A_{Min}$ 은 최솟값을 나타낸다. 최댓값과 최솟값을 바탕으로 영상의 색상 H 값을 구하는데, 이 값은 최댓값이 R, G, B 중 어느 값이 가장 큰지에 따라 다른 수식을 통해 구한다. 채도 S는 R, G, B의 최댓값과 최솟값의 차이를 최댓값으로 나눈 값으로 나타내는데, 최댓값이 0인 경우 채도는 0이 된다. R, G, B의 최댓값은 명도 V의 값이 된다.

제안하는 배경 제거 방법에선 RGB 영상을 HSL 색상 공간으로 변환시킨 후, 배경 영상과 대상 영상 속 각 화소의 색상 및 채도 차이를 구한다. 이때 차이 값이 임계값보다 크면 해당 화소는 객체의 일부로 간주한다. 식 (2)는 임계값을 구하는 공식이다.

식 (2)에서  $\alpha$ 와  $\beta$ 는 색상과 채도의 임계값에 대한 가중치를 나타내고,  $T$ 는 최종 임계값을 나타낸다. 최종 임계값을 계산하기 위해 색상 임계값  $T_{Huc}$ 와 채도 임계값  $T_{Saturation}$ 을 설정하고, 이에 대한 가중치를 부여한다. 이때  $\alpha$ 는 0과 1 사이의 값을 갖고,  $\beta$ 는 1보다 큰 값을 가진다.

색상과 채도에 대한 임계값과 가중치는 고정된 값이 아니며 영상마다 다른 값을 가진다. 따라서 촬영 환경이나 배경의 복잡도, 카메라의 차이 등을 고려해 임계값과 가중치를 설정해야만 한다.

### 2.2 SAD (Sum of Absolute Differences)

대부분의 불필요한 배경 정보는 색상 값과 채도 값의 차이를 이용한 방법으로 제거할 수 있다. 그런데도 아직 제거되지 않은 배경 정보가 여전히 존재한다. 제안하는 방법은 남아있는 배경 정보를 배제하고 객체 정보에 보다 초점을 맞추기 위해 SAD 연산을 사용한다. 식 (3)은 SAD 공식을 나타낸다.

$$|R_{back} - R_{fore}| + |G_{back} - G_{fore}| + |B_{back} - B_{fore}| > \lambda \quad (3)$$

식 (3)을 사용하기 위해 HSL로 변환된 색상 공간을 다시 RGB 색상 공간으로 바꾼다. 남아있는 영역 속 각 화소에 대해 전경 영상과 배경 영상의 R, G, B 차이의 절댓값을 구한다. 전경과 배경 화소의 R, G, B 차이의 절댓값을 더한 후 그 값이 일정 기준 값보다 클 경우, 해당 화소를 객체에 대한 정보로 간주한다. 반대의 경우, 배경에 대한 정보로 간주한다.

### 2.3 형태학적 연산을 활용한 홀 채움

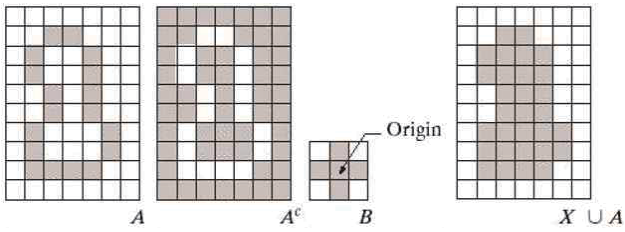
영상의 채도와 색상 값, 그리고 SAD 공식을 통해 만들어진 마스크 객체 내에는 아직 빈 공간이 일부 존재한다. 홀 채움에는 다양한 방법이 있는데, 본 논문에서는 형태학적 연산 기반의 홀 채움 방법을 사용한다.

형태학적 연산은 기본적으로 대상 객체의 팽창과 침식 두 연산을 활용한다. 그중에서 팽창 연산을 활용해 빈 공간을 채워나간다. 식 (4)와 그림 2는 이 과정을 식과 그림을 통해 표현한 것이다.

$$X_k = (X_{k+1} \oplus B) \cap A^c \quad (k=1,2,3,\dots)$$

$$A = X_k \cup A \quad (4)$$

식 (4)에서 A는 이미 채워진 객체 부분이고, 채워지지 않은 부분은 이의 여집합으로 표현했다. 채워진 홀에 대한 정보를 담은 별도의 집합은 X로, 모든 홀이 채워질 때까지 단계적으로 추가된다. B는 과정 수행을 위해 사용될 구조적 요소를 뜻한다.



(그림 2) 팽창 연산을 응용한 홀 채움 연산

객체 부분의 임의의 화소 값 하나를 시작으로 구조적 요소를 통해 A의 여집합과 비교를 한다. 일치하는 부분이 있으면 X에 저장한다. 저장된 X에 대해서 또다시 구조적 요소를 통해 A의 여집합과 비교를 하고, X를 포함한 일치하는 모든 화소 값들을 다음 X에 저장한다. 객체 내에 있는 모든 홀에 대해서 이 연산을 수행할 때까지 과정을 반복한다.



(a) 원본 영상 (b) 객체 마스크

(그림 3) 배경 제거 결과

그림 3은 배경을 제거하고 객체만을 추출한 결과 영상이다. 원본 영상에 대해 배경 제거 연산을 수행하고, 이를 통해 객체 마스크를 생성하고, 이를 이용하면 원본 영상에 대해서도 객체만을 갖는 색상 영상도 만들어낼 수 있다.

### 3. 깊이지도 업샘플링

#### 3.1 3차원 워핑

깊이 카메라를 통해 촬영한 깊이지도는 색상 카메라에서 촬영한 색상 영상과 초점 거리나 카메라의 위치 등의 부분에서 차이를 보인다. 3차원 워핑을 사용하면 깊이 카메라에서 촬영한 깊이지도를 다양한 색상 카메라의 시점에 대한 깊이지도로 변환할 수 있다. 식 (4)는 3차원 워핑에 대한 전반적인 공식을 나타낸다.

$$\begin{aligned} m_s &= A_s \cdot R_s \cdot M_W + A_s \cdot t_s \\ M_W &= R_s^{-1} \cdot A_s^{-1} \cdot m_s - R_s^{-1} \cdot t_s \\ m_t &= A_t \cdot R_t \cdot M_W + A_t \cdot t_t \end{aligned} \quad (5)$$

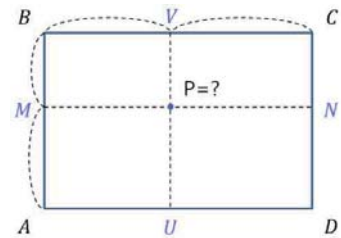
3차원 워핑을 수행하기 위해선 각 카메라의 내부 및 외부 파라미터의 값을 알아야만 한다. 내부 파라미터는 카메라의 초점 거리와 주점을, 외부 파라미터는 카메라의 회전 및 위치 변형에 대한 정보를 가지고 있다. 식 (5)에서 행렬 A는 내부 파라미터 정보를, R은 회전, t는 위치 변형에 관한 정보를 가지고 있는 행렬이다.

워핑 대상이 될 깊이지도를 촬영했던 깊이 카메라의 파

라미터 값들을 활용해 깊이지도 내의 화소 값들을 식 (5)를 통해 모두 3차원 월드 좌표계로 이동시켜준다. 식 (5)의  $m_s$ 는 깊이지도의 좌표 값을,  $M_W$ 는 월드 좌표계의 좌표 값을 나타낸다. 월드 좌표계로 이동한 모든 화소 값들은 색상 카메라의 파라미터 값들을 활용해서 이에 대응되는 2차원 영상 평면으로 워핑을 시켜준다. 식 (5)를 통해 최종적으로 얻게 될  $m_t$ 가 이를 뜻한다.

#### 3.2 이중선형보간법

워핑된 깊이 값 화소 사이에 있는 빈 공간들에 들어갈 값을 추정하기 위해 이중선형보간법을 활용한다. 보간법은 데이터의 값을 알고 있는 특정 두 대상 사이의 값을 추정하기 위해 사용되는 방법이다. 2차원 영상의 보간을 위해서는 이중선형보간법을 사용한다. 식 (6)과 그림 4는 이중선형보간법에 대한 설명을 나타낸다.



(그림 4) 이중선형보간법

$$\begin{aligned} P = & \frac{C-V}{C-B} \left( \frac{B-M}{B-A} \times A + \frac{M-A}{B-A} \times B \right) \\ & + \frac{V-B}{C-B} \left( \frac{B-M}{B-A} \times D + \frac{M-A}{B-A} \times C \right) \end{aligned} \quad (6)$$

점 P의 값을 알기 위해 그 주변의 값들을 이용한다. 값에 대한 정보를 가지고 있는 화소들을 A, B, C, D라 할 때, 점 P에서의 데이터 값은 식 (6)과 같은 공식을 통해 구한다. 이 공식의 원리는 A와 B를 선형 보간법을 이용해 두 점 사이의 중간 값을 구하고, 맞은편의 C와 D에 대해 선형 보간법을 적용해 그 중간 값을 구하는 순서로 진행된다. 그림 4의 M, N, U, V는 좌표들 사이의 중간 값을 뜻한다. 두 중간 값을 알게 되면 이를 통해 다시 한 번 선형 보간법을 이용한다. 두 중간 값의 선형보간이 끝나면 최종적으로 우리가 구하고자 하는 P의 값을 알 수 있다.

#### 3.3 평균값 필터를 이용한 경계 보정

이중선형보간법을 통해 채워진 깊이지도는 최종적으로 필터를 적용해 빈 공간을 채우거나 경계 부분을 다듬어준다. 본 논문에서는 이를 위해 3x3 크기의 평균값 필터를 활용해 이중선형보간법을 수행했다.

식 (7)은 본 논문에서 적용한 평균값 필터다. 0 값을 갖는 부분을 계산 범위에 포함하게 되면 구하게 될 평균값에 영향을 끼쳐 주변 값들과 비교하면 부정확한 결과가 나온다. 따라서 이 과정에서 값이 비어있는 부분은 연산

과정에서 제외했다. 3x3 크기의 범위 내에서 0이 아닌 값을 갖는 화소의 수를 m, 화소 값의 전체 합을 S에 저장한다. 계산된 S값을 m으로 나눠 화소의 최종 값을 얻는다.

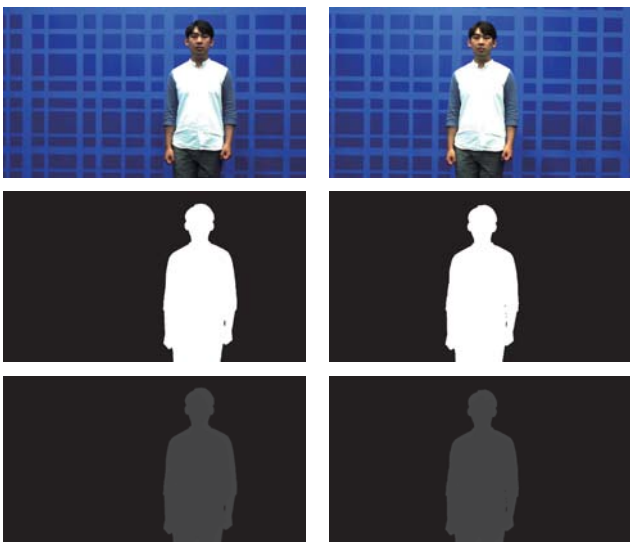
$$\begin{cases} m = \text{count}(p_{i,j} \neq 0) \\ S = \sum_{i=-1}^1 \sum_{j=-1}^1 p_{i,j} \\ p'_{i,j} = \frac{S}{m} \end{cases} \quad (7)$$

필터를 적용하는 과정까지 마친 깊이지도는 촬영 당시 객체의 거리에 따라 배경의 깊이 값과 굉장히 가까운 값을 갖게 될 수도 있다. 값의 차이는 적으나, 객체로 간주하여 활용하고자 하면 이러한 부분을 고려해야 한다. 배경이 제거된 객체 마스크를 활용하면 배경의 깊이 값과 차이가 적은 객체 부분에 대해서도 명확하게 포함해서 객체의 깊이 값으로써 활용할 수 있다. 최종적으로 깊이지도와 앞선 과정에서 추출한 객체 마스크가 일치하는 부분만을 추출하여 객체의 깊이도를 획득할 수 있다.

#### 4. 실험 결과



(a) 깊이지도 원본



(b) 왼쪽 시점 업샘플링 (c) 오른쪽 시점 업샘플링

(그림 5) 다른 시점에 대한 실험 결과

본 논문에서는 SR-4000 깊이 카메라를 통해 그림 5(a)와 같은 깊이지도의 원본을 촬영했고, 여러 시점의 위치로 설치된 색상 카메라들로 색상 영상들을 촬영하였다. 촬영

된 영상들을 바탕으로 배경 제거 및 깊이지도 업샘플링 알고리즘을 수행하였다. 그림 5(b)는 두 시점에 대한 깊이지도 업샘플링 후 객체 마스크 이미지를 통해 객체의 깊이 정보를 중점적으로 추출한 결과다.

3차원 워핑은 카메라의 위치나 속성 등을 고려해서 수행한 연산이다. 따라서 이를 통해 만들어진 깊이 업샘플링 영상은 실제 촬영한 색상 영상 내의 객체 위치와 일치함을 보여준다. 이를 통해 제안된 방법이 객체의 깊이지도를 만드는 데 유용함을 알 수 있다.

단, 배경 제거 알고리즘은 색상과 채도를 이용하므로 배경의 색이나 반사되는 빛 등의 요인에 따라 그 결과가 달라질 수 있다. 따라서 배경 제거 알고리즘은 여러 환경을 고려해 색상과 채도의 가중치를 다르게 부여해야 한다.

#### 5. 결론

본 논문에서는 배경이 제거된 객체 마스크를 활용해 객체의 깊이지도 업샘플링 결과를 보다 명확하게 얻는 방법을 제안했다. Canny 경계 검출법을 활용해 객체의 깊이지도 업샘플링 결과를 명확하게 얻을 수도 있지만, Canny 경계 검출법은 복잡한 배경에 대해서는 불필요한 경계도 추출하는 등의 문제점을 가지고 있다. 반면, 배경이 제거된 객체 마스크를 활용하면 명확하게 필요한 객체에 대해서만 추출하는 데 도움이 되고, 불필요한 배경 정보를 제거할 수 있었다.

#### 감사의 글

본 논문은 민·군기술협력사업(Civil-Military Technology Cooperation Program)으로부터 지원을 받아 수행된 연구임.

#### 참고 문헌

- [1] C. Fehn, R. de la Barre and S. Pastoor, "Interactive 3-DTV-Concepts and Key Technologies." Proceedings of the IEEE, Vol. 94, No. 3, pp.524-538, 2006.
- [2] M. Hansard, S. Lee, O. Choi, and R. Horaud, "Time-of-Flight Cameras: Principles, Methods and Applications," Springer, pp.2-3, 2012.
- [3] J. Canny, "A Computational Approach to Edge Detection." IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, No. 6, pp.679-698, 1986.
- [4] S.M. Hong, Y.S. Ho, "Fast Joint Bilateral Upsampling using Color Edge Information," International Conference on Embedded Systems and Intelligent Technology (ICESIT), pp.101-102, 2014.
- [5] J. D. Foley, A. Van Dam, S. K. Feiner, and J. F. Hughes, "Computer Graphics: Principles and Practice," 3rd ed, Addition-Wesley, pp.776-777, 2013.