

키넥트를 이용한 색상 및 깊이 기반 영상 분할 기법

김영배, 장원동, 김창수
고려대학교

{youngbaekim, wdjang, cskim}@mcl.korea.ac.kr

요 약

본 논문에서는 색상 및 깊이 기반 영상 분할 기법을 제안한다. 계층화된 영상 분할을 수행하기 위해서 색상을 기준으로 영상을 과분할 한 후, 과분할 영역의 깊이를 기준으로 영역 병합을 수행한다. 적은 개수의 화소로 이루어진 병합 영역을 제거하기 위해서 인접한 분할 영역 중 화소 수가 많은 영역에 병합시키는 이상영역 처리 기법을 수행한다. 제안하는 영상 분할 기법을 기존의 데이터셋 및 키넥트 취득 영상에 적용하여 신뢰도 높은 객체 단위 영상 분할이 이루어짐을 확인한다.

1. 서론

최근 영상 취득 장비의 기술 향상에 힘입어 색상 정보와 깊이 정보를 동시에 사용한 영상 분할 기법의 연구가 활발하게 이루어지고 있다. 키넥트는 색상과 근거리 깊이 정보를 동시에 취득할 수 있는 대표적인 장비로, 가격이 저렴하고 다양한 기본 기능을 제공하여 영상 분할 및 안면 인식, 동작 인식 등 다양한 연구 분야에서 사용되고 있다.

영상 분할 기법은 감독 기법과 비감독 기법으로 나뉜다. 감독 기법은 사용자가 분할하고자 하는 영역을 입력으로 하여 사용자가 표시한 영역과 특징이 유사한 영역을 분할한다. 감독 기법은 주로 영상에서 사용자가 원하는 전경을 배경으로부터 정확하게 분할하고자 할 때 사용되며 대표적인 기법으로 Grabcut [1] 기법이 연구되었다. 비감독 기법은 사용자의 입력 없이 색상, 텍스처, 모서리 위치 등 영상의 특징만을 이용하여 영상을 분할하는 기법이다. 깊이 정보는 객체 단위 영상 분할을 위한 유용한 정보를 제공한다. 하나의 객체가 다양한 색상으로 구성된 경우, 깊이 유사성을 바탕으로 같은 영역으로 분할할 수 있다. 본 연구에서는 색상 및 깊이 정보를 기반으로 한 비감독 영상 분할 기법을 제안한다.

최근 영상 분할은 주로 계층화된 화소 병합 기법으로 수행된다. 우선 색상이 유사한 화소를 병합하는 과분할을 수행한 후, 특징이 유사한 과분할 영역을 병합하여 최종 분할 영상을 획득한다. 대표적인 과분할 기법으로 SLIC [2] 과 같은 슈퍼픽셀 분할 기법이 사용되고 있다. 과분할 기법을 통해 생성된 영역간 경계는 최종 분할 영역의 경계를 포함하고 있어야 하기 때문에, 적절한 과분할 기법을 적용하는 것이 매우 중요하다. 본 연구에서는 색상 기반 과분할을 수행한 후, 깊이 기반 영역 병합을 수행하는 계층화 기법을 사용한다.

본 논문의 구성은 다음과 같다. 2-1 절에서는 분할에 사용하는 색상 및 깊이 영상의 정규화 규칙을 정의한다. 2-2 절에서는 색상 기반 과분할 기법과 깊이 기반 영역 병합 기법을 살펴본 후, 2-3 절에서 이상영역을 처리하는 기법을 제안한다. 3 절에서는 제안한 기법의 성능을 실험을 통해서 확인한다. 마지막으로 4 절에서 본 논문에 대한 결론을 맺는다.

이 논문은 2014 년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구(No.2009-0083495) 및 2015 년도 에스케이텔레콤의 지원을 받아 수행된 연구임.

2-1. 색상 및 깊이 영상 정규화

컬러 및 깊이 기반 영상 분할을 수행하기 위해 취득한 색상 정보와 깊이 정보는 동일한 상황을 기술하고 있어야 한다. 동일한 상황을 기술하기 위해서 두 정보는 동일한 위치에서 같은 시점에 취득되어야 한다. 두 센서를 동기화 하여 시간적 동시성은 보장할 수 있는 반면, 같은 공간에 두 센서가 동시에 존재할 수 없으므로 취득 정보간 공간적 차이는 필연적으로 발생한다. 이에 따라 색상 영상과 깊이 영상간의 시점 차이가 발생하고, 이를 극복하기 위한 다양한 정합 방법이 연구되었다. 일반적으로 영상간 정합은 한 영상을 상대 영상의 시점에 맞추어 변환하는 방식으로 이루어진다. 시점 변환에는 각 센서의 고유 특성인 내부 파라미터와 두 센서의 위치 차이에 따라 결정되는 외부 파라미터가 사용된다. 키넥트의 경우, 각 센서의 파라미터를 내부적으로 제공하며, 두 센서간 위치 차이가 고정되어 있다. 따라서 본 연구에서는 키넥트가 제공하는 내장 파라미터를 사용하여 두 영상을 정합하는 방식을 사용하였다.

정합된 영상 중에서 깊이 영상은 피사체와의 거리에 따라 값의 범위가 다양하게 나타날 수 있다. 따라서 깊이 영상은 [0,1] 범위에 맞추어 정규화 함으로써 알고리즘이 입력 영상의 특징에 무관하게 안정적으로 작동할 수 있도록 보장한다. 일반적으로 16 비트 깊이 영상의 폐색 영역은 한 화소가 표현할 수 있는 최소값인 0 혹은 최대값인 $2^{16}-1$ 로 저장된다. 본 연구에서 폐색 영역은 최소값인 0 으로 표현되었다. 이때, 깊이 영상을 단순히 최대값과 최소값에 대해 정규화할 경우 폐색 영역의 값인 0 과 실제 최소 깊이간의 깊이 값은 사용할 수 없기 때문에, 정규화 영상의 동적 영역이 줄어드는 문제가 발생한다. 본 연구에서는 폐색 영역의 값을 실제 최소 깊이 보다 1 만큼 작도록 깊이 영상을 가공하여 정규화 깊이 D 를 다음과 같이 계산한다.

$$D = \frac{D^0 - (D^{\min} - 1)}{D^{\max} - (D^{\min} - 1)} \quad (1)$$

여기에서 D^0 는 입력 깊이, D^{\min} 은 실제 최소 깊이, D^{\max} 는 최대 깊이를 나타낸다. 제안하는 정규화 깊이를 사용하면 폐색 영역의 값을 정규화 전과 동일하게 0 으로 나타내면서 정규화 이전에 비해 넓은 동적 영역을 표현할 수 있는 장점이 있다.

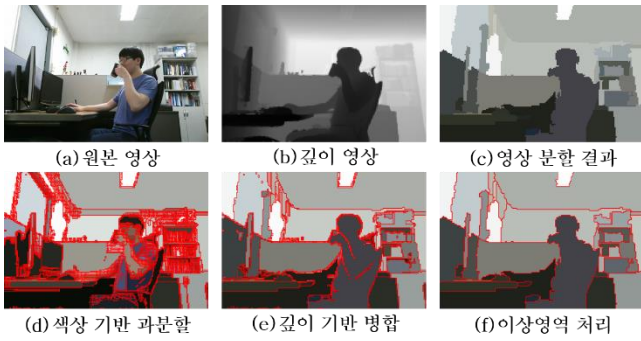


Figure 1. 영상 분할 흐름도. (a), (b), (c)는 각각 원본 영상과 깊이 영상 및 분할 결과 영상을 나타낸다. (d), (e), (f)는 순서대로 색상 기반 과분할 결과, 깊이 기반 병합, 이상영역 처리 간 경계를 붉게 처리하여 나타낸다.

2-2. 색상 및 깊이 기반 영상 분할

색상 정보에 추가적으로 깊이 정보가 주어지면 영상의 효과적인 객체 단위 분할이 가능하다. 객체가 다양한 색상으로 구성되어 있더라도 깊이가 유사하다면 영역을 병합하여 하나의 객체로 분할할 수 있다. 따라서, 본 연구에서는 색상을 기반으로 과분할 영역을 생성한 후, 깊이가 유사한 영역을 병합하였다. 본 연구에서는 각 채널이 8 비트로 표현된 RGB 영상을 원본 영상으로 사용한다. 이때, 인접한 i 번째 화소와 j 번째 화소의 색상 거리 d_{ij}^c 는 다음과 같이 RGB 유클리드 거리의 제곱으로 계산한다.

$$d_{ij}^c = |R_i - R_j|^2 + |G_i - G_j|^2 + |B_i - B_j|^2 \quad (2)$$

색상 거리가 일정값 보다 작은 인접 화소는 병합하여 과분할 영역을 생성한다.

동일한 객체를 구성하는 영역의 경우 영역 경계에서 연속적인 깊이 변화를 갖는다. 따라서 영역 경계의 깊이 차이가 일정값 이하일 경우 두 영역을 합침으로써 객체를 분할할 수 있다. 그러나 영역 경계의 깊이만 비교하면, 서로 다른 두 객체가 접촉할 경우 두 객체를 병합하는 문제가 발생할 수 있다. 따라서, 본 연구에서는 영역 경계의 깊이 비교와 더불어 영역간 평균 깊이를 비교하였다.

2-3. 이상영역 처리

색상 및 깊이 기반 병합 이후, 충분한 크기를 갖는 영역은 정상적으로 분할된 영역으로 간주한다. 그러나 영상 취득 과정에 포함된 잡음에 의하여 동일 객체를 구성하지만 색상 및 깊이가 주변과 다른 영역이 발생할 수 있다. 이러한 영역들은 주변 영역에 병합되지 않고 적은 수의 화소를 포함한다. 본 연구에서는 일정값 이하의 화소를 갖는 영역을 이상영역으로 정의하고, 이상영역을 인접한 정상영역에 병합시키는 기법을 적용하였다. 인접한 정상영역이 둘 이상일 경우, 가장 큰 깊이를 갖는 영역으로 병합되도록 하여 깊이가 작은 객체일수록 윤곽이 선명한 영역으로 분할되도록 하였다. 주요 객체는 취득 장비에 가까운 곳에 위치하기 때문에 제안하는 이상영역 처리를 통해 객체를 정확히 분할한다. 이상영역은 점진적으로 가장 적합한 깊이를 갖는 정상 영역에 병합되도록 한다. Figure 1 은 알고리즘의 전체 흐름을 나타낸다.

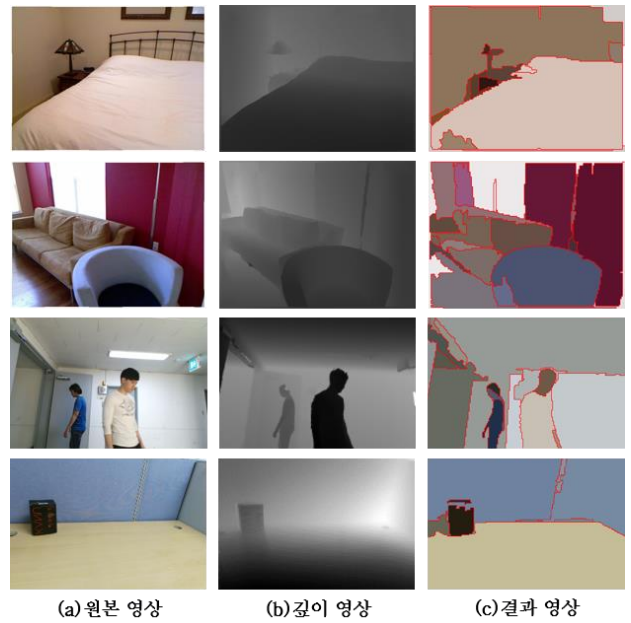


Figure 2. 실험 결과. NYU_v2 데이터셋 (위 두 줄)과 취득 영상 (아래 두 줄)에 대해서 순서대로 (a)원본 영상, (b)깊이 영상, (c) 결과 영상을 나타낸다.

3. 실험 결과

본 연구에서는 NYU_v2 데이터셋[3]과 키넥트를 사용하여 취득한 영상을 사용하여 제안하는 영상 분할 기법의 성능을 확인하였다. Figure 2 는 NYU_v2 와 취득 영상에 대한 분할 결과를 (a)원본 영상, (b)깊이 영상, (c)결과 영상 순서로 나타낸다. 실험을 통해 제안하는 알고리즘이 복잡한 영상에서도 객체 단위 분할을 정확히 수행함을 확인할 수 있다.

4. 결론

본 연구에서는 색상 및 깊이 기반 분할 기법을 제안하였다. 먼저, 색상 차이를 기반으로 영상 과분할을 수행한 후, 과분할 영역을 깊이를 기준으로 병합하였다. 화소 수가 적은 이상영역은 탐색 영역을 넓혀가며 인접한 정상 분할 영역에 병합시켜 최종적인 영상 분할결과를 획득하였다. 기존 데이터셋과 취득 데이터를 사용하여 본 알고리즘이 신뢰도 높은 객체 단위 영상 분할을 수행함을 확인하였다.

5. 참고논문

[1] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: interactive foreground extraction using iterated graph cuts," *ACM Trans. Graphics*, vol. 23, no. 3, pp. 309–314, 2004.

[2] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels," *EPFL Technical Report*, 2010.

[3] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *ECCV*, pp. 746–760, 2012.