

## 3차원 복원을 위한 세그먼트 기반의 전경물체 추출

## Segment-based Foreground Extraction Dedicated to 3D Reconstruction

김정환\*, 박안진\*\*, 정기철\*\*\*

**요약** 영상기반의 3차원 복원(reconstruction)에 대한 연구가 최근 좋은 결과를 많이 내고 있지만, 이는 복원의 목적이 되는 영역이 각 입력영상으로부터 미리 정확하게 추출되어 있다고 가정하기 때문이다. 본 논문에서는 다시점 영상에서 *세그먼트(segment)*라 불리는 초기 분할된 영역을 기반으로 전경물체 추출과 3차원 복원을 EM 형식으로 반복적으로 수행하는 정교한 전경물체 추출방법을 제안한다. 본 논문에서 세그먼트는 수행시간을 최소화하기 위해 사용된 3차원 복원방법인 visual hull에 의해 발생한 외곽선 오류를 보상해야 하며, 사용자의 입력을 최소화하기 위해 적은 수의 집합으로 구성되어야 한다. 이를 위해 데이터 항과 스무드 항으로 구성된 에너지 함수를 최적화할 수 있는 그래프 컷 방법을 이용하는 영상분할을 초기단계로써 수행하며, 전경물체 추출과 3차원 복원은 에너지 함수가 최적화될 때까지 반복 수행한다. 실험에서 간단한 3차원 복원 방법을 이용함에도 불구하고 전경물체의 외곽선에서 정확한 결과를 보였다.

**Abstract** Researches of image-based 3D reconstruction have recently produced a number of good results, but they assumed that the accurate foreground to be reconstructed is already extracted from each input image. This paper proposes a novel approach to extract more accurate foregrounds by iteratively performing foreground extraction and 3D reconstruction in a manner similar to an EM algorithm on regions segmented in an initial stage, called *segments*. Here, the segments should preserve foreground boundaries to compensate for the boundary errors generated by visual hull, simple 3D reconstruction to minimize the computational time, and should also be composed of the small number of sets to minimize the user input. Therefore, we utilize image segmentation using the graph-cuts method, which minimizes energy function composed of data and smoothness terms, and the two methods are iteratively performed until the energy function is optimized. In the experiments, more accurate results of the foreground, especially in boundaries, were obtained, although the proposed method used a simple 3D reconstruction method.

**핵심어:** 3D 복원(Reconstruction), 전경 물체추출(Foreground Extraction), 영상 분할(Image Segmentation), 그래프 컷(Graph Cuts)

본 논문은 서울시 산학연 협력사업(10581)지원에 의하여 연구되었음.

\*김정환 : 숭실대학교 미디어학과 e-mail: kim598@ssu.ac.kr

\*\*박안진 : 숭실대학교 미디어학과 e-mail: anjin@ssu.ac.kr

\*\*\*정기철 : 숭실대학교 미디어학과 교수; e-mail: jungkeechul@gmail.com

## 1. 서론

우리가 살고 있는 세상이 3 차원 공간이기 때문에, 현실성을 느끼고 3 차원 물체에 대한 정확한 정보를 얻을 수 있다. 이런 이유 때문에 최근 다양한 물체나 사람의 몸, 얼굴영역에 대한 3 차원 복원(reconstruction)이 컴퓨터 비전이나 그래픽스 분야에서 각광받고 있다[1].

하지만 이는 복원된 대상이 각 입력영상으로부터 미리 정확하게 추출되어있다고 가정하기 때문이며, 일반적으로 차영상(background subtraction)이 많이 이용되고 있다. 차영상 방법은 먼저 한 장의 고정된 영상을 이용하는 정적인 방법과 순차적으로 입력되는 영상을 계속해서 갱신하는 동적인 방법으로 배경 정보를 모델링하고, 모델링된 영상과 입력영상을 픽셀단위로 비교하여 전경물체를 추출하는 방법이다[2]. 이 방법은 연산이 간단하기 때문에 많이 이용되고 있지만, 잡음이나 대비(contrast)가 낮은 배경 영역에서 잘못된 전경이 추출되는 경우 또는 구멍(hole)과 같이 추출이 되지 않는 경우가 생길 수 있다[3]. 이로 인해 다시점 영상에 전경물체를 추출하는 방법들이 최근 다양하게 제안되고 있다[4-7]. Goldlücke 와 Magnor[4]는 입력된 여러 장의 영상에서 전경물체를 추출하는 방법을 제안했지만, 3D 물체를 복원하는 대신 영상 평면(plane)에 대응하는 깊이 정보를 계산하였다. 뿐만 아니라 전경물체를 추출하기 위해 각 뷰에 대응하는 배경영상을 가지고 있어야만 전경물체를 추출할 수 있기 때문에 실제환경에서 많은 제약이 가지게 되는 단점이 있다. 각 뷰에 대응하는 배경영상의 필요성을 없애기 위해 Sormann 등[5]은 지능적인 그래픽 사용자 인터페이스를 제안했다. 이 인터페이스는 다시점의 영상을 3D 큐브(cube) 형식을 디자인하여 전경과 배경을 쉽게 지정할 수 있을 뿐만 아니라 모든 영상을 동시에 접근할 수 있게 설계하였다. Tsai 등[6]은 그래프 컷을 이용한 전경물체 추출과 3D 복원 방법을 결합한 다시점 영상에서의 새로운 전경물체 추출방법을 제안하였으며, 두 방법은 더 좋은 전경물체 추출을 위해 반복적으로 수행되었다. 이 방법은 초기 대략적으로 추출된 전경의 집합에서 사용자가 정교하게 추출된 결과를 선택하며, 선택된 결과는 3D 물체를 복원하기 위해 이용되었다. 그리고 복원된 3D 물체는 다시 다시점의 각 영상으로 재투영되어 추출된 전경물체의 정확도를 향상시키며, 이 단계를 사용자가 만족할 때까지 반복 수행한다. 위의 두 방법[5,6]은 전경영역 추출과 3D 복원에서 좋은 결과를 보였지만, 많은 양의 정보가 필요하다는 단점이 있다. 예를 들어, Sormann 등[5]이 제안한 방법은 정교한 전경, 배경에 대한 사전정보를 위해 많은 수의 사용자입력이 필요하며, Tsai 등[6]이 제안한 방법에서는 더 정확한 전경물체를 매 반복단계에서 사용자가 선택해야만 한다.

Campbell 등[7]은 다시점 영상에서 전경물체를 자동으로 추출하기 위한 방법을 제안하였으며, 사용자의 입력대신 '추출될 대상은 입력영상의 중심에 있다'는 가정을 이용하였다. 이 방법은 입력영상 중심의 특정범위

내에 포함되어있는 색상값을 GMM 으로 학습하고, 이 결과를 그래프 컷의 사전정보로 이용하여 전경물체를 추출하였으며, 추출된 전경물체는 3D 물체 복원을 위한 정보로 이용되었다. 이 역시 두 단계를 반복적으로 수행하였으며, 3D 물체 복원의 재 투영 결과를 기반으로 학습할 색상값을 확대해 나가고 이로 인해 더 좋은 전경물체를 추출할 수 있게 되었다. 그러나 이 방법은 초기 특정범위 내에 추출될 전경물체의 색상정보가 대략적으로 포함되어있어야 하기 때문에 초기값에 영향을 많이 받는다. 뿐만 아니라 전경물체 추출 결과를 정교하게 하기 위한 그래프 컷을 이용한 정교한 3D 물체 복원을 수행하기 때문에 많은 수행시간이 필요하다.

이와 같이 위에서 언급한 방법들[4-7]은 다음에서 정리된 것과 같이 4 가지 조건을 만족하지 못하는 단점을 가지고 있다.

- 1) 각 뷰에 대한 배경정보를 요구하지 않아야 한다[4].
- 2) 전경물체의 정확도를 향상시키면서 사용자의 입력을 최소화해야 한다[5,6].
- 3) 전경물체의 정확도를 향상시키기 위해 이용되는 3D 복원의 수행시간을 최소화해야 한다[6,7].
- 4) 초기값에 영향을 받지 않아야 한다[7].

따라서 본 논문에서는 위의 4 가지 조건을 만족하는 3 차원 복원을 위한 새로운 전경물체 추출 방법을 제안한다.

본 논문의 구조는 다음과 같다. 제 2 장에서 제안된 방법을 개략적으로 설명하며, 제 3 장에서 초기단계에서 세그먼트를 구하기 위한 그래프 컷 기반의 영상분할 방법을 설명한다. 제 4 장에서는 세그먼트와 3D 복원을 기반으로 한 제안된 전경물체 추출 방법을 기술한다. 마지막으로 제 5 장에서 실험 및 결과에 대해 언급하고, 제 6 장에서 결론 및 향후 연구 방향을 제시하면서 논문을 마무리한다.

## 2. 제안된 방법의 개요 (overview)

제안된 방법의 목적은 다시점 영상에서 적은 수의 입력을 기반으로 정확도를 향상시키는 것이다. 이를 위해 우리는 EM 알고리즘과 유사한 방식으로 3 차원 복원과 전경물체 추출을 반복수행하며, 복원된 3 차원 물체는 이전단계에서 추출된 전경물체에 재 투영함으로써 더욱더 향상된 전경물체를 추출하고, 추출된 전경물체를 기반으로 향상된 3 차원 물체를 복원한다. 첫 번째 단계는 본 논문에서 *세그먼트*라 불리는 분할된 영역을 기반으로 전경물체를 추출하는 것으로, 처음에는 사용자의 입력을 기반으로 대략적으로 전경물체를 추출하며, 첫 번째 이후의 전경물체는 복원된 3 차원 물체의 재 투영에 의해 추출된다. 이 단계가 3 차원 물체를 고정된 상태에서 최적의 전경물체를 추출하기 때문에, EM 단계의 "E" 단계와 유사하다. 두 번째 단계는 첫 번째 단계에서 추출된 전경물체를 기반으로 3 차원 물체를 구축하는 것으로, 3 차원 물체는 다시점의 모든 영상의 광선 원뿔(ray-cone) 교차영역(intersection)을 기반으로 복원되기 때문에 실제 3D 물체의 외곽에 위치한 잡음을 제거할 수 있다. 이

단계는 전경이 고정된 상태에서 3D 를 복원하기 때문에 “M”단계를 유사하다.

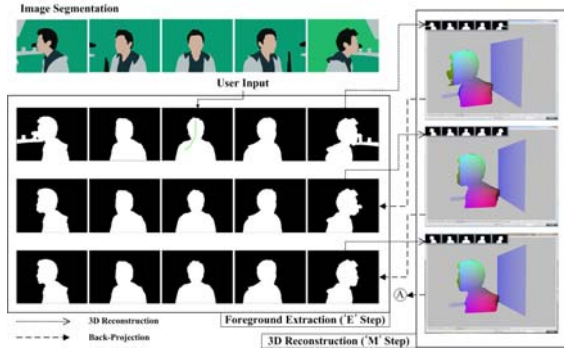


그림 1. 제안된 방법의 흐름도.

그림 1 은 제안된 방법의 흐름도를 예를 들어 설명한다. 다시점 영상이 주어지면, 제안된 방법은 먼저 전경물체의 외곽선에 특히 강인한 세그먼트를 생성하는 자동 영상분할 방법을 수행한다. 본 논문에서는 그래프 컷 기반의 방법으로 영상분할을 수행하며, 이는 영상을 분할하기 위해 데이터 항에 의해 특징공간에서뿐만 아니라 스무드 항에 의해 영상공간에서 특징값을 분석할 수 있는 에너지 함수를 최소화할 수 있는 장점 때문에 이용되었다. 데이터 항을 자동으로 설정하기 위해 임의로 선택된 초기 mean 으로부터 밀도가 높은 지역인 모드로 이동하는 mean 들의 집합을 각 클래스(세그먼트)를 위한 사전정보로 이용하였으며, MeanShift 분석 방법은 다양한 클래스의 모양을 보다 정확하게 분석할 수 있는 장점 때문에 1 장에서 기술한 3 번째 조건을 만족한다. 스무드 항은 영상공간에서 이웃하는 픽셀간의 불연속성을 보존하는 Potts 모델을 이용하였으며, 다중 레이블링을 위해 이용된  $\alpha$ -expansion 방법에 최적화된 모델이다. 에너지 최소화에 의해 분할된 결과는 입력영상을 전체를 채우는 세그먼트로 구성된다. 초기 영상분할 후 그림 1 에서 보는 바와 같이 하나의 입력영상에서 실제 전경과 일치하는 세그먼트를 선택하여 초기 전경물체를 추출하며, 이로 인해 두 번째 조건에서 언급한 사용자의 입력을 최소화할 수 있다.

추출된 전경물체를 기반으로 3 차원 물체를 복원하며, 우리는 평면기반의 교차영역 방법(plane-based intersection method)을 이용하였다. 평면기반의 방법은 visual hull 을 만드는 방법들 중에 하나이며, 복셀(voxel)기반의 방법들보다 빠른 수행시간을 보이는 장점을 가지고 있다. 뿐만 아니라 이 방법은 각 평면의 교차영역을 이용하여 3D 물체를 복원하기 때문에 교차영역 밖에 있는 잡음들을 제거한다. 하지만 visual hull 은 광선 원뿔을 기반으로 교차영역을 구하기 때문에 3 차원 물체를 대략적으로 복원하며, 이를 2 차원 영상에 재 투영할 경우 전경물체의 외곽선이 정확하지 않을 문제점을 가지고 있다. 이를 해결하기 위해 우리는 영상공간에서 불연속성을 보존하는 세그먼트를 이용한다.

제안된 방법의 기본적인 아이디어는 완성된 퍼즐에서 조각을 제거해가면서 원하는 형태의 그림을 도출하는

것이다. 본 논문에서 세그먼트가 조각과 같은 역할을 하며, 세그먼트를 제거하면서 실제 전경물체보다 큰 초기에 추출된 전경물체를 정확한 결과로 도출한다. 세그먼트를 제거하는 건 3 차원 물체를 각 세그먼트에 투영했을 때의 픽셀 수와 일치하는 세그먼트의 총 픽셀수의 비율을 기반으로 하며, 이로 인해 전경물체의 외곽선에서 정확하게 추출할 수 있다. 뿐만 아니라 제안된 방법은 3D 복원에 필요한 카메라 교정(camera calibration)에서 발생하는 오류의 범위가 크지 않다면 정확한 전경물체를 추출할 수 있는 장점을 가지고 있다.

### 3. 초기 영상분할

에너지 함수가 전경물체 추출과 3D 복원의 반복을 종료하기 위한 조건으로 사용되고 세그먼트가 대략적으로 복원된 3D 물체에 의해 발생할 수 있는 전경물체 외곽선 오류를 제거할 수 있기 때문에, 초기 영상분할은 제안된 방법에서 중요한 역할을 한다.

본 논문에서 영상분할을 레이블링(labeling)을 위해 MRF 모델로 고려하였다. 레이블링 문제는 주어진 영상  $\mathbf{P}$  의 각 픽셀  $p$  에 레이블 집합 중 하나를 할당하는 것으로  $\mathbf{F} = \{f_1, f_2, \dots, f_n\}$  으로 표시하며,  $\mathbf{F}$  의 각 요소는 레이블의 집합 중 하나가 할당되고  $n$  은 픽셀의 수를 나타낸다. 입력영상의 각 픽셀에 적합한 레이블을 할당하기 위해, 영상분할 문제를 먼저 에너지 함수로 표현하며 다음과 같다.

$$E(\mathbf{F}) = \sum_{p \in \mathbf{P}} D_p(f_p) + \lambda \sum_{\{p, q\} \in \mathbf{EN}} V_{p, q}(f_p, f_q)$$

여기서  $D_p(f_p)$ 는 픽셀  $p$  가  $f_p$  에 얼마나 적합한지를 나타내는 데이터 항으로, 픽셀  $p$  가 레이블  $f_p$  에 대한 사전정보와 유사하면  $D_p(f_p)$ 에 높은 비용(cost)을 할당하며, 유사하지 않으면 낮은 비용을 할당한다.  $V_{p, q}(f_p, f_q)$ 는 스무드 항으로 영상공간에서 이웃하는 픽셀과의 불연속성을 나타내며, 이웃하는 두 픽셀이 유사하면 높은 비용(cost)을 할당한다. 우리는 스무드 항을 Potts 모델로 표현하기 위해  $V_{p, q}(f_p, f_q)$ 을  $V_{p, q} \cdot \delta(f_p, f_q)$ 로 대체한다. 여기서  $\delta(f_p, f_q)$ 는 델타 함수로  $f_p$  와  $f_q$  가 같은 레이블을 가지면 0, 아니면 1 의 값을 가지며, 두 픽셀이 다른 레이블을 가질 때의 비용만을 고려하기 위해 사용한다.  $N$ 은 입력영상에서 이웃하는 픽셀 쌍의 집합으로  $n$ -links(neighborhood links)라 불린다.

에너지 함수에서 데이터 항의 비용을 할당하기 위해 각 레이블을 위한 사전정보가 필요하며, 본 논문에서는 특징공간에서 임의로 선택된 초기 mean 으로부터 각 모드로 이동하는 mean 들의 경로를 사전정보로 이용한다. Mean shift 의 초기단계에서 임의로 선택된 데이터 집합을 mean 의 초기값으로 설정하며, 각각의 mean  $ms_i$  는  $mean_{sk}(ms_i)$ 를 이용하여 갱신한다

여기서  $mean_{s_k}(s) = \frac{\sum_{j=1}^n s_j \cdot sk(\|s-s_j\|^2)w(s_j)}{\sum_{j=1}^n sk(\|s-s_j\|^2)w(s_j)}$  이며,  $w(s_j)$ 는 가중치 함수,  $sk(\cdot)$ 는 커널 함수의 섀도우(shadow)이다. 다음단계로  $ms_i$ 는 모드에 도착할 때까지 반복해서 갱신되며,  $k$  번째 모드로 향하는 연속적인 mean 의 집합을 레이블  $k$  를 위한 사전정보로 이용한다. 이렇게 획득한 각 레이블의 집합은 수행시간을 줄이기 위해 히스토그램으로 표현하며, 데이터 항은 다음과 같이 정의한다.

$$\begin{cases} D_p(f_p = 1) = \infty, & D_p(f_p = 2) = 0, & \forall p \in ms^1, \\ D_p(f_p = 1) = 0, & D_p(f_p = 2) = \infty, & \forall p \in ms^2, \\ D_p(f_p = 1) = \ln P(p|H_1), & D_p(f_p = 2) = \ln P(p|H_2), & \forall p \in ms^U. \end{cases}$$

여기서  $ms^U$ 는  $ms^1$ 과  $ms^2$ 에 속하지 않는 픽셀들의 집합을 의미하며,  $\ln P(p|H)$ 는 집합  $ms^i$ 의 히스토그램  $H_i$ 가 주어졌을 때  $p$ 의 확률값을 나타낸다.

스무드 항의 비용은 각 픽셀 사이의 불연속 보전을 위해 할당되고 이를 위해 Potts 모델을 사용하였으며  $V_{p,q}$ 는 다음과 같이 정의된다.

$$V_{p,q} = dis(p, q)^{-1} \exp(-\beta \|p - q\|^2).$$

여기서  $dis(\cdot)$ 는 영상공간에서 픽셀 사이의 거리를 나타내며, 우리는 유클리디언 거리를 이용하였다. 에너지 함수를 최소화하기 위해 우리는 그래프 컷을 사용하며 이를 위해 먼저 픽셀과 동일한 정점을 가진 그래프  $G=\langle v, \varepsilon \rangle$ 를 생성한다. 그래프에 레이블 정보를 반영하기 위해 두 개의 추가적인 정점  $source(S)$ 와  $sink(T)$ 가 필요하며 각 정점에  $t-links$ (terminal links)라 불리는 에지  $\{p, S\}$ 와  $\{p, T\}$ 를 추가하게 된다. 결론적으로 그래프의  $v$ 와  $\varepsilon$ 는 다음과 같이 나타낼 수 있다.

$$v = P \cup \{S, T\}, \varepsilon = N \cup_{p \in P} \{\{p, S\}, \{p, T\}\}.$$

그래프의 가중치는  $t-links$ 와  $n-links$ 에 할당되며, 터미널과 각 정점을 연결하는  $t-links$ 는 픽셀과 레이블 사이의 유사도를 나타내는 데이터 항이, 근접한 이웃 정점을 연결하는  $n-links$ 에는 두 근접한 픽셀 사이의 연속성을 나타내는 스무드 항이 그래프의 가중치로 각각 할당된다.

그래프 컷은 그래프  $G$ 가 설정된 후에 가중치의 합이 최소가 되는 컷(cut)을 찾는 방법으로 컷된 두 터미널  $S$ 와  $T$ 에 각각에 레이블을 할당한다[8]. 최소 컷은 Ford와 Fulkerson의 이론[9]을 기반으로  $S$ 에서  $T$ 로 흐르는 max flow를 찾는 것으로 구한다. maxflow는 정점을 두 개의 중복되지 않는 영역으로 나누는 그래프의 에지들을 산출하며, 그 에지의 가중치 합이 최소 컷의 비용과 같다. 따라서 그래프에서 max flow가  $t-links$ 와  $n-links$ 가 할당된 가중치의 합이 최소가 되는 에지를 산출하기 때문에, 본 논문에서 max flow를 에너지 함수를 전역으로 최소화하기 위해 이용한다.

#### 4. 복원된 3D 물체를 기반으로 한 전경물체 추출

전경물체는 첫 번째 단계를 제외하고 복원된 3D 물체를 다시점 각 영상의 세그먼트에 투영함으로써 추출한다. 따라서 이 장에서는 대략적으로 추출된 전경물체를 기반으로 3D를 어떻게 복원하는지 대략적으로 소개하고, 복원된 3D를 기반으로 어떻게 전경물체를 좀더 정교하게 추출하는지를 기술한다.

##### 4.1 3D 복원

실루엣 기반으로 한 볼륨공간상의 교차영역 방법(silhouette volume intersection)은 다시점 영상에서 추출된 전경물체들을 기반으로 3D를 구축하는 가장 유명한 방법 중에 하나이다[1]. 이 방법은 3차원 볼륨공간에 2차원 전경물체를 투영하여 3D 물체를 복원하는 방법이며, 다시점의 전경물체와 함께 3D 공간에서 교차영역을 기반으로 3D 물체를 대략적으로 추출할 수 있다. 이를 일반적으로 visual hull 이라고 부른다[1].

일반적인 교차영역기반의 방법들은 3차원 공간을 작은 복셀(voxel)로 분할하며, 각 복셀은 전경물체 내에 속하지는 아닌지를 평가하기 위해 각각의 영상공간으로 투영된다. 이 복셀 기반의 원근 투영(perspective projection)방법의 가장 큰 단점은 계산량이 많이 필요하다는 것이다. 계산량을 줄이기 위해 본 논문에서는 평면기반의 교차영역 방법을 이용하였다. 3차원 복셀공간은  $z$  축을 기반으로 평행의 평면으로 분할되고 각 평면에 투영된 전경물체는 다음의 3 단계를 이용하여 3D로 복원된다. 첫째, 다시점 각각의 영상으로부터 추출된 전경을  $z$  축의 가장 아래에 있는 평면에 투영하고, 둘째  $z$  축으로 한 평면씩 올라가면서 추출된 전경을 투영하며, 마지막으로 각 평면에 투영된 모든 전경의 2차원 교차영역을 구한다. 이렇게 구한 2차원 교차영역의  $z$  축 집합이 3차원 복원의 결과이다.

##### 4.2 전경물체 추출

정교한 전경물체를 추출하기 위한 본 논문의 아이디어는 여러 개의 조각으로 구성된 완성된 퍼즐로부터 조각(본 논문에서는 세그먼트)을 제거하면서 최종적으로 원하는 형태를 만드는 것이며, 그림 2는 아이디어에 대한 예제를 보여준다. 초기 단계에서 선택된 실제 전경과 일치하는 조각을 기반으로 다시점의 모든 영상에서 대략적으로 전경을 추출(그림 2(b))한 후 나머지 조각들이 정교한 전경물체를 추출하기 위해 제거되어야 하는지 아닌지는 3D 물체로부터 투영된 결과와 세그먼트 사이의 비율을 기반으로 결정되며, 비율은 다음과 같이 계산된다.

$$Portion_i = \frac{\# \text{ of pixels projected from 3D object into } i^{\text{th}} \text{ segment}}{\# \text{ of pixels in } i^{\text{th}} \text{ segment}}$$

여기서  $i$  번째 세그먼트에 대한 비율(Portion)이 주어진 문턱치(threshold) 값보다 작으면  $i$  번째 세그먼트는 제거된다. 그림 3(a)와 (c)는 그림 2(b)와 (c)를 포함하는 다시점의 모든 영상을 기반으로 복원된 3D 물체로부터 재투영된 결과이며, 그림 3(b)와 (d)는 그림 2(b)와 (c)에 접목된 재투영 결과를 보여준다. 그림 2(c)와 3(b)에서 보는 바와 같이 세그먼트 중 하나가 위의 식의 비율을 기반으로 제거된다.

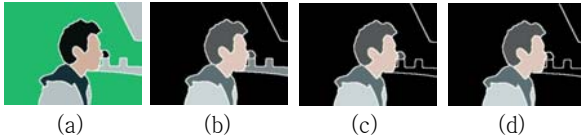


그림 2. 제안된 방법의 예: (a) 여러 개의 조각으로 구성되어 있는 완성된 퍼즐로부터 (d) 최종결과를 도출하는 과정.

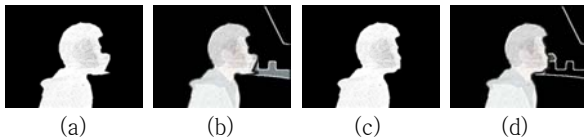


그림 3. 세그먼트와 3D 물체로부터 재투영된 결과의 관계: (a,c) 재투영된 결과, (b,d) 세그먼트에 접목된 재투영 결과.

### 5. 실험 및 결과

모든 실험은 2.66 GHz 펜티엄 4 에서 실행되었으며, 모든 구현은 오픈 소스(예를 들어 그래프 컷과 에너지 최소화 라이브러리(library)[8])를 기반으로 하였다. 그림 1 은 본 논문에서 사용한 6 개의 테스트 세트를 보여주며, 그림 1 을 포함해서 총 7 개의 세트를 제안된 방법을 테스트하기 위해 이용하였다. 그림 5 는 그림 4 의 영상에 대한 초기 영상분할에 대한 결과이며, 실험에서 사용한 모든 영상의 크기는  $800 \times 600$  이다.



그림 4. 실험에서 사용한 테스트 영상의 집합.

그림 5 의 3 번째 행에서 사용자가 전경에 해당하는 몇 개의 mean 들을 선택하면, 그 외의 영상에 대한 전경물체는 선택된 세그먼트를 기반으로 대략적으로 추출된다. 다음단계로 실제 전경물체가 다른 성분보다

크다는 가정하에 추출된 후의 각 성분의 크기를 이용하여 일차적으로 전경물체 추출에 대한 오류를 제거한다. 그림 6 은 그림 5 의 첫 번째 행에 대한 대략적인 결과를 보여주며, 그림 6(b)의 2 번째, 3 번째, 5 번째 결과는 실제 전경과 같다. 이는 제안된 방법은 전경물체 추출과 3D 복원을 한번만 반복해도 최종결과를 추출할 수 있음을 의미하며, 때문에 수행시간을 최소화할 수 있다.



그림 5. 그림 4 에 대한 초기 영상분할 결과.



(a)



(b)

그림 6. 초기단계에서 추출된 대략적인 전경: (a) 사용자 입력에 의한 결과, (b) 성분의 크기를 기반으로 약간의 오류가 제거된 결과.

표 1 은 1 회 수행시간을 보여준다. 만약 반복회수가 2 회 이상이면, 재투영에 대한 수행시간이 거의 소모되지 않기 때문에 3D 복원에 대한 수행시간만을 총 수행시간에 추가하면 된다. 3D 복원은 간단한 알고리즘을 사용했기 때문에 약 15 초의 수행시간을 보인다. [6,7]에서 사용한 그래프 컷 기반의 3D 복원 방법은 매우 높은 정확도를 보이지만, 3 차원 그리드(grid)의 크기가  $300^3$  일 때 약 45 분 정도가 소모된다. 기존의 방법과 비교해서 제안된 방법은 visual hull 을 이용하기 때문에 빠른 수행시간에 비해 3D 복원의 외곽선에서 정교하지 못한 결과를 보이지만 세그먼트를 이용하기 때문에 전경물체 추출에서는 정교한 외곽선을 추출할 수 있다.

표 1. 제안된 방법의 수행시간(sec.)

	그래프 컷	영상분할	3D 복원
수행시간	8.0	1.5	15.0

그림 7 은 그림 4 에 대한 최종 전경물체 추출결과이다.



그림 7. 제안된 방법의 전경물체 추출결과.

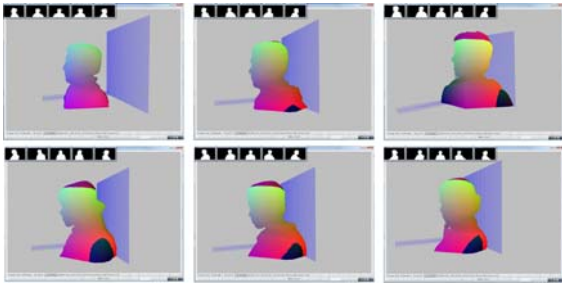


그림 8. 그림 7을 기반으로 복원한 3 차원 물체

## 6. 결론

본 논문에서는 다시점 영상에서 정교한 전경물체 추출방법을 제안하며, 하나의 영상에서 사용자가 실제 전경과 일치하는 세그먼트를 간단하게 선택함으로써 사용자의 입력을 최소화하였다. 제안된 방법은 초기 단계에서 분할된 세그먼트의 집합에서 EM 알고리즘과 유사하게 전경물체 추출과 3D 복원을 반복적으로 수행함으로써 정교하게 전경물체를 추출할 수 있다. 초기에 그래프 컷을 이용하여 분할된 세그먼트는 전경물체의 외곽선에 강건하며, 이는 수행시간을 최소화하기 위해 사용된 간단한 3D 복원에 의해 발생하는 외곽선 오류를 제거할 수 있으며, 기존 방법과 비교해서 높은 정확도와 빠른 수행시간을 보이는 장점을 가지고 있다.

## 참고 문헌

- [1] A. Laurentini, "The Visual Hull Concept for Silhouette-based Image Understanding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, Issue 2, pp. 150-162, 1994.
- [2] C. Stauffer and W.E.L. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 246-252, 1999.
- [3] Image Processing Toolbox, Chapter 9, *Morphological Operations*, The Mathworks, 2001.

- [4] B. Goldlücke and M.A. Magnor, "Joint 3D Reconstruction and Background Separation in Multiple Views using Graph Cuts," *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 683-688, 2003.

- [5] M. Sormann, C. Zach, and K. Karner, "Graph Cut based Multiple View Segmentation for 3D Reconstruction," *Proceedings of IEEE International Symposium on 3D Data Processing, Visualization, and Transmission*, pp. 1085-1092, 2006.

- [6] Y-P. Tsia, C-H. Ko, Y-P. Huang, and Z-C. Shih, "Background Removal of Multiview Images by Learning Shape Priors," *IEEE Transactions on Image Processing*, Vol. 16, No. 10, pp. 2607-2616, 2007.

- [7] N. Campbell, G. Vogiatzis, C. Hernandez, and R. Cipolla, "Automatic 3D Object Segmentation in Multiple Views using Volumetric Graph-Cuts," *Proceedings of British Machine Vision Conference*, Vol. 1, pp. 530-539, 2007.

- [8] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximation Energy Minimization via Graph Cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 11, pp. 122-1239, 2001.

- [9] L. Ford and D. Fulkerson, *Flows in Networks*, Princeton University Press, 1962.