

3차원 얼굴 복원을 위한 그래프 컷 기반의 전경 물체 추출 방법

박안진^o 홍광진 정기철

송실대학교 IT대학 미디어학과

{anjin^o, hongmsz, kcjung}@kiss.or.kr

Better Foreground Segmentation for 3D Face Reconstruction using Graph Cuts

Anjin Park^o Kwangjin Hong Keechul Jung
School of Media, College of IT, Soongsil University

요 약

영상기반의 3차원 복원(reconstruction)에 대한 연구가 컴퓨터 성능의 발전과 다양한 영상기반의 복원 알고리즘의 연구로 인해 최근 좋은 결과를 보이고 있으나, 이는 얼굴영역과 같은 목적이 되는 영역이 각 입력영상으로부터 미리 정확하게 추출되어있다고 가정하기 때문이다. 일반적으로 목적이 되는 영역을 추출하기 위해 차영상이 많이 이용되고 있지만, 차영상은 잡음과 구멍(hole)과 같은 오 추출된 영역이 발생하기 때문에 목적이 되는 영역을 3차원으로 복원을 할 때 심각한 오류를 초래할 수 있다. 전경 물체(목적이 되는 영역)를 정확하게 추출하기 위해 최근 그래프 컷(graph cut)을 이용한 방법이 다양하게 시도되고 있다. 그래프 컷은 데이터 항(data term)과 스무드 항(smooth term)으로 구성된 에너지 함수를 전역적으로 최소화하는 방법으로 여러 공학적 문제에서 좋은 결과를 보이고 있지만, 에너지 함수의 데이터 항을 설정할 때 필요한 사전정보를 자동으로 얻기가 어렵다. 스테레오 비전의 깊이 정보가 최근 전경 물체 추출을 위한 사전정보로 많이 이용되고 있고 그들의 실험환경에서는 좋은 결과를 보이지만, 3차원 얼굴 복원에서 얼굴의 대부분이 동질의 영역을 가지고 있기 때문에 깊이 정보를 구하기 어려워 정확한 사전정보를 구하기가 어렵다. 본 논문에서는 3차원 얼굴 복원을 효과적으로 하기 위한 그래프 컷 기반의 전경 물체 추출 방법을 제안한다. 에너지 함수의 데이터 항을 설정하기 위해 전경 물체에 대한 사전정보를 추출해야 하며, 이를 위해 차영상을 이용하여 대략적인 전경 물체 추출하고, 사전 정보에 대한 오류를 줄이기 위해 잡음과 그림자 영역을 제거한다. 잡음과 그림자 영역을 제거하면 구멍이 발생하거나 실루엣이 손상되는 문제가 발생한다. 손상된 정보는 근접한 픽셀이 유사하지 않을 때 낮은 비용을 할당하는 에너지 함수의 스무드(smooth) 항에 의해 에지 정보를 기반으로 채워진다. 결론적으로 제안된 방법은 스무드 항과 대략적으로 설정된 데이터 항으로 구성된 에너지 함수를 그래프 컷으로 전역적으로 최소화함으로써 더욱 정확하게 목적이 되는 영역을 추출할 수 있다.

1. 서 론

우리가 살고 있는 세상이 3차원 공간이기 때문에, 현실성을 느끼고 3차원 물체에 대한 정확한 정보를 얻을 수 있다. 이런 이유 때문에 최근 다양한 물체나 사람의 몸, 얼굴영역에 대한 3차원 복원(reconstruction)이 컴퓨터 비전이나 그래픽스 분야에서 각광받고 있다. 특히 여러 대의 카메라에서 획득한 멀티 영상을 기반으로 3차원 물체 등을 복원하는 영상기반의 3차원 복원에 대한 연구가 컴퓨터의 성능의 발전과 다양한 복원 알고리즘의 연구로 인해 최근 좋은 결과를 보이고 있다[1].

영상기반의 3차원 복원이 좋은 결과를 보이고 있지만, 이는 얼굴영역과 같은 목적이 되는 영역이 각 입력 영상으로부터 미리 정확하게 추출되어있다고 가정하기 때문이며, 일반적으로 차영상이 많이 이용되고 있다. 차영상은 첫 영상을 이용하는 정적인 방법과 연속 영상을 계속해서 갱신하는 동적인 방법으로 배경을 모델링하고,

모델링된 영상과 입력 영상을 픽셀단위로 비교하여 전경물체(목적이 되는 영역)를 추출한다[2]. 차영상은 간단한 연산 때문에 많이 이용되고 있지만, 카메라에서 발생하는 잡음이나 대비(contrast)가 낮은 배경 영역으로 인해 잘못된 전경물체가 추출되는 경우 또는 구멍과 같이 추출되지 못하는 경우가 생길 수 있다[3].

더욱 효과적인 전경 물체 추출을 위해 최근 그래프 컷(graph cut)을 이용한 다양한 방법들이 시도되고 있다[4-10]. 그래프 컷은 데이터 항(data term)과 스무드 항(smooth term)으로 구성된 에너지 함수를 전역적으로 최소화하는 방법으로, 에너지 함수로 표현할 때 데이터 항의 값을 설정하기 위한 전경, 배경에 대한 사전정보가 주어지면 전경 물체 추출에서 높은 정확도를 보인다. Li 등[4], Boykov와 Funka-Lea[5], Rother 등[6]은 그래프 컷을 이용하여 전경물체를 추출하는 인터랙티브 툴(interactive tool)을 제안하였다. Li 등[4]과 Boykov와 Funka-Lea[5]가 제안한 툴에서는 마우스 버튼을 누른

상태에서 드래그하여 전경 물체와 배경 물체에 대한 사전 정보를 얻었으며, 왼쪽 마우스 버튼은 전경을 오른쪽 마우스 버튼은 배경을 획득하기 위해 이용되었다. Rother 등[6]이 제안한 툴에서는 전경 물체의 주위에 사각형으로 표시함으로써 기존 방법보다 쉽게 전경과 배경에 대한 사전 정보를 획득했다. 위에서 언급한 방법들[4-6]은 그래프 컷을 이용하여 전경 물체를 추출하기 때문에 높은 정확도를 보이지만, 전경, 배경을 위한 사전 정보를 사용자가 직접 입력해야 하는 단점을 가지고 있다.

최근 전경물체와 배경에 대한 사전 정보를 자동으로 추출하기 위해 스테레오 비전 기술을 이용하였다[7,8]. 이 방법은 특정 범위내의 깊이를 가진 영역이 전경물체라고 가정하여 전경물체와 배경을 위한 사전 정보를 자동으로 추출하였다. 이 방법은 그들의 실험 환경에서 좋은 결과를 보였지만, 3차원 얼굴 복원에 적용할 때 얼굴 영역의 대부분이 유사한 색상이나 텍스처를 가지기 때문에 깊이 정보를 정확하게 구하기가 힘들며, 이로 인해 그래프 컷을 위한 정확한 사전 정보를 얻기가 어렵다. 그래프 컷을 이용한 또 다른 전경 물체 추출 방법으로 Howe와 Deschamps[10]는 데이터 항을 설정하기 위해 배경 영상과 입력 영상의 차이 값을 이용하였으며, 전경 물체를 위한 사전 정보는 주어진 매개변수에 의해 설정되었다. 이 방법은 차이 값을 사전 정보로 사용하기 때문에 잡음에 민감하며, 전경물체에 대한 부정확한 사전 정보 때문에 추출된 전경 물체에서 많은 구멍이 발생한다. Sun 등[11]은 배경에 대한 사전 정보를 얻기 위해 Gaussian mixture models(GMM)을 이용하였으며, 이는 잡음에 강인한 정보를 얻을 수 있는 장점을 가지고 있다. 하지만, 전경 물체가 균일 분포(uniform distribution)에 가진다는 가정하에, 전경 물체에 대한 사전 정보를 매개변수로 할당하였으며, 이는 실제 응용분야에서 제약사항이 된다.

본 논문에서는 영상기반의 3차원 얼굴 복원을 위해 스테레오 비전 기술대신 최근 많이 이용되고 있는 볼륨 데이터 교차 영역 추출(Volume Intersection)[1]과 3차원 영역 수정(Object Carving)[9]을 이용한다. Volume Intersection은 3차원 얼굴 영역을 대략적으로 복원하며, 실루엣 정보를 이용하기 때문에 전경 물체의 정확한 실루엣이 필요하다. Object Carving은 정교한 복원을 위해 사용되며, 홀이나 잡음 등이 없는 정확하게 추출된 전경 물체가 필요하다. 본 논문에서는 3차원 얼굴 복원을 효과적으로 하기 위한 그래프 컷 기반의 정교한 전경 물체 추출 방법을 제안한다. 전경 물체 추출에 대한 정확도를 향상시키기 위해 매개 변수가 아닌 보다 정확한 전경 물체에 대한 사전 정보 추출해야 한다. 이를 위해 차영상을 이용하여 대략적인 전경 물체 추출하고, 사전 정보에 대한 오류를 줄이기 위해 잡음과 그림자 영역을 제거한다. 잡음과 그림자 영역을 제거하면 구멍이 발생하거나 실루엣이 손상되는 문제가 발생한다. 손상된 정

보는 근접한 픽셀이 유사하지 않을 때 낮은 비용을 할당하는 에너지 함수의 스무드(smooth) 항에 의해 에지 정보를 기반으로 채워진다. 결론적으로 제안된 방법은 대략적으로 추출된 물체를 기반으로 에너지 함수의 데이터 항을 보다 정확하게 설정하고 스무드 항을 기반으로 오 추출된 영역으로 채움으로써 정확도가 높은 전경 물체를 추출할 수 있다.

본 논문의 구성은 다음과 같다. 제 2장에서 본 논문에서 제안한 그래프 컷을 이용한 전경 물체 추출에 대해 기술하며, 제 3장에서 Volume Intersection과 Object Carving을 이용한 3차원 얼굴 복원 방법을 기술한다. 제 4장에서 실험 및 결과에 대해 언급하고, 제 5장에서 결론 및 향후 연구 방향을 기술한다.

2. 전경 물체 추출

본 논문에서는 보다 정확한 전경물체를 추출하기 위해 그래프 컷을 이용한다. 일반적으로 그래프 컷을 컴퓨터 비전 문제에 적용할 때, 주어진 문제를 어떻게 에너지 함수로 표현하며, 데이터 항을 위한 사전 정보는 어떻게 구해야 하는지를 고려해야 한다. 2.1에서 전경 물체 추출 문제를 어떻게 에너지 함수로 표현하는지에 대해 기술하며, 2.2에서는 제안된 방법으로 전경 물체의 사전 정보를 어떻게 자동으로 추출할 것인지를 기술한다.

2.1. 전경 물체 추출을 위한 에너지 함수

본 논문에서는 전경 물체 추출 문제를 레이블링 (labeling) 문제로 고려한다. 레이블링 문제는 주어진 영상 P 의 각 픽셀에 레이블 집합 중 하나를 할당하는 것으로, 전경 물체 추출에서 레이블 집합은 $\{ 'fg', 'bg' \}$ 두 개이며, 여기서 fg 는 전경, bg 는 배경을 의미한다. 레이블링 문제는 $F = \{ f_1, f_2, \dots, f_{Dn} \}$ 표현하며 여기서 Dn 은 주어진 영상의 픽셀 수이며 F 의 각 요소들은 레이블 집합 중에 하나이다. 각 픽셀에 특정 레이블을 할당하기 위해 먼저 전경 물체 추출 문제를 에너지 함수(식 1)로 표현해야 한다.

$$E(F) = \sum_{p \in P} D_p(f_p) + \lambda \sum_{(p,q) \in N} V_{p,q}(f_p, f_q) \quad (1)$$

여기서 $D_p(f_p)$ 는 픽셀 p 가 f_p 에 얼마나 적합한지를 나타내는 데이터 항이다. 다르게 표현해서 픽셀 p 가 각 레이블에 대한 사전 정보와 유사하면 $D_p(f_p)$ 에 낮은 비용(cost)를 할당하며, 유사하지 않으면 높은 비용을 할당한다. $V_{p,q}(f_p, f_q)$ 는 두 근접한 픽셀 p, q 사이가 유사한지를 나타내는 스무드 항이다. 다르게 표현해서 $V_{p,q}(f_p, f_q)$ 는 두 근접한 픽셀이 유사하면 높은 비용을 유사하지 않으면 낮은 비용을 할당한다. λ 는 두 항의 상대적인 중요성을 나타내며, N 은 입력영상의 근접한 이웃에서 픽셀 쌍의 집합으로 $n-links$ (neighborhood links)라 불린

다.

그래프 컷은 에너지 함수를 적역적으로 최소화하기 위해 사용되며, 그래프 컷을 사용하기 위해 에너지 함수를 그래프 $G=\langle v, \epsilon \rangle$ 로 변환해야 한다. v 는 영상의 각 픽셀과 터미널이라 불리는 두 개의 추가적인 노드로 구성되며, 터미널은 두 레이블을 표현하기 위해 사용된다. ϵ 는 근접한 이웃 픽셀을 연결하는 n -links와 각 픽셀과 각 터미널 사이를 연결하는 t -links(terminal links)로 구성된다. 그래서 그래프의 v, ϵ 는 $v = P \cup \{S, T\}$ 와

$$\epsilon = N \cup_{p \in P} \{(p, S), (p, T)\}$$

로 변환할 때 에너지 함수의 각항이 t -links와 n -links에 설정된다. 터미널과 픽셀을 연결하는 t -links에는 픽셀과 레이블 사이의 적합도를 표현하는 데이터 항이, 근접한 이웃 픽셀을 연결하는 n -links에는 두 근접한 픽셀 사이의 유사도를 표현하는 스무드 항이 그래프의 가중치로 할당된다. 그래프 G 가 완벽하게 설정된 후에 그래프 컷은 가중치의 합이 최소가 되는 컷(cut)을 찾으며, 컷된 두 터미널 S 와 T 에 각각 연결된 각 픽셀에 각 레이블을 할당한다.

그래프 컷은 모든 컷 중에서 최소의 가중치를 가진 컷을 찾기 위해 이용되며, 최소 컷 문제는 Ford와 Fulkerson[13]의 방법을 이용하여 s 에서 t 로의 최대 흐름(maximum flow)을 찾음으로써 해결할 수 있다.

2.2. 사전 정보를 위한 대략적인 전경 물체 추출

에너지 함수의 데이터 항(t -links)에 값을 할당하기 위해 각 레이블에 대한 사전 정보가 필요하며, 초기 그래프 컷을 이용한 방법은 사용자로부터 수동으로 사전 정보를 입력 받았다.

전경 물체 추출에서 자동으로 에너지 함수를 정의하기 위해 배경 정보를 기반으로 한 방법들이 제안되었다. Howe와 Deschamps[10]는 데이터 항을 설정하기 위해 모델링된 배경 영상과 입력 영상의 차이 값인 δ_p 을 이용하였으며, 데이터 항은 다음과 같다.

$$D_p(f_p) = \begin{cases} \delta_p & \text{if } f_p = 'bg' \\ 2\tau - \delta_p & \text{if } f_p = 'fg' \end{cases}$$

이 데이터 항은 배경을 위한 사전 정보로 차이값을 이용하기 때문에 잡음에 민감하게 반응하며, 전경 물체를 위한 사전 정보는 주어진 매개변수인 τ 에 의해 설정되기 때문에 전경 물체에 대한 정확한 사전 정보를 구하지 못한다. 뿐만 아니라 스무드 항(식 2)으로 주어진 매개변수 α 와 τ 에 의해 설정되는 Pott model을 이용하기 때문에 근접한 두 픽셀 사이의 유사성을 표현하지 못하는 단점을 가지고 있다.

$$V_{p,q}(f_p, f_q) = \begin{cases} \tau\alpha & \text{if } f_p \neq f_q \\ 0 & \text{if } f_p = f_q \end{cases} \quad (2)$$

Sun 등[11]은 배경에 대한 사전 정보를 얻기 위해 Gaussian mixture models(GMM)을 이용하였으며, 이는 잡음에 강인한 정보를 얻을 수 있는 장점을 가지고 있다. 하지만, 전경 물체가 균일 분포(uniform distribution)에 가진다는 가정하에, 전경 물체에 대한 사전 정보 const(상수)으로 할당하였으며, 이는 실제 응용분야에서 제약사항이 된다. Sun 등[11]이 제안한 데이터 항은 다음과 같다.

$$D_p(f_p) = \begin{cases} \arg \min_k \left\| (I_p - \mu_k) \Sigma_k^{-1} \right\|^2 & \text{if } f_p = 'bg' \\ const & \text{if } f_p = 'fg' \end{cases}$$

여기서 k 는 가우시안 모델의 개수, μ_k 와 Σ_k 는 k 번째 가우시안 모델의 평균과 공분산, I_p 는 p 번째 픽셀의 정보를 나타낸다. 스무드 항으로 영상 분할에서 에지 정보에 강인하다고 알려져 있는 함수[14]를 이용하였으며, 다음과 같다.

$$V_{p,q}(f_p, f_q) = \begin{cases} 1 - \frac{\tan^{-1}(|\delta_{pq}|)}{\pi/2} & \text{if } f_p \neq f_q \\ 0 & \text{if } f_p = f_q \end{cases} \quad (3)$$

여기서, δ_{pq} 는 이웃하는 픽셀 p 와 q 사이의 차이값이다.

위의 두 방법은 전경 물체에 대해 사전 정보를 제대로 구하지 못하는 단점을 가지고 있기 때문에, 실제 응용 분야에 적용할 때 정확도가 떨어지는 결과를 보인다. 따라서 본 논문에서는 전경 물체 문제에서 보다 효과적인 데이터 항을 설정하기 위해 전경 물체를 대략적으로 추출하며, 기본적으로 차영상을 이용한다.

차영상을 이용하여 사전 정보를 추출하면, 잡음이나 그림자와 같은 잘못 추출된 영역 때문에 에너지 함수를 제대로 설정하지 못하는 문제가 생기며, 이 때문에 심각한 문제를 야기할 수 있다. 그래서 전경 물체의 사전 정보를 추출하기 전에 잘못 추출된 영역을 제거해야 한다. 본 논문에서는 잘못 추출된 영역으로 카메라에서 생긴 잡음과 그림자 영역을 고려하였으며, 잡음은 간단한 모폴로지(morphology) 연산을 이용하여 제거한다. 모폴로지 연산을 이용하여 잡음을 제거할 경우 가장 큰 문제는 실루엣이 손상되는 문제가 발생한다. 이 문제는 영상의 에지를 기반으로 다른 레이블이 할당되고 근접한 두 픽셀이 유사할 때 낮은 비용을 할당하는 스무드 항(식 3)으로 인해 손상된 실루엣을 회복할 수 있다. 둘째 차영상을 이용할 때 그림자 영역이 추출되지만 대부분이 그림자 영역은 배경이기 때문에 제거해야 한다. 식 4와 같이, 우리는 그림자 영역은 배경과 비교해서 유사한 색상(hue)을 가지지만 명도(brightness)가 어둡다는 가정하에 그림자를 제거한다.

$$Shadow_{xy} = \begin{cases} 1 & \text{if } |Hue_{xy}^{fg} - Hue_{xy}^{bg}| < \tau_H \text{ and } Brt_{xy}^{fg} - Brt_{xy}^{bg} < 0 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$D_p(f_p) = \begin{cases} \arg \min_k \left| \alpha_k^{bg} \cdot \frac{1}{2\pi|\Sigma_k^{bg}|} \exp \left[-\frac{1}{2} (I_p - u_k^{bg})^T \Sigma_k^{bg} (I_p - u_k^{bg}) \right] \right| & \text{if } f_p = 'bg' \\ \arg \min_k \left| \alpha_k^{fg} \cdot \frac{1}{2\pi|\Sigma_k^{fg}|} \exp \left[-\frac{1}{2} (I_p - u_k^{fg})^T \Sigma_k^{fg} (I_p - u_k^{fg}) \right] \right| & \text{if } f_p = 'fg' \end{cases} \quad (5)$$

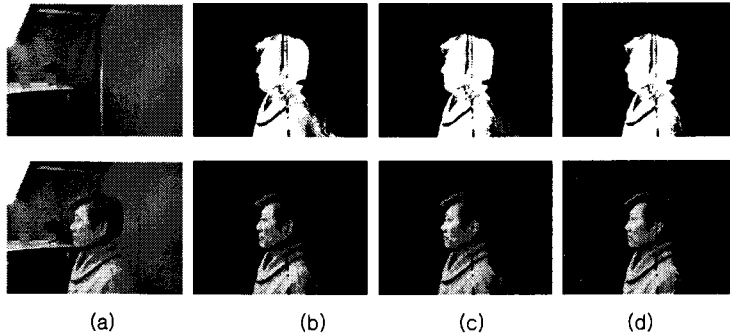


그림 2. 대략적으로 추출된 전경물체의 예: (a) 입력영상과 배경영상, (b) 차영상, (c) 그림자영역을 회색으로 표시한 영상, (d) 잡음과 그림자가 제거된 결과영상.

식 4의 Hue_{xy}^{fg} 는 입력영상에서 x, y 좌표의 색상 값이며,

Brt_{xy}^{bg} 는 배경 영상에서 x, y 좌표의 색상 값이다. 그림 2

는 대략적인 전경 물체를 보여준다. 그림 2(b)는 차영상을 보여주며, 그림 2(c)는 그림자 영역을 회색으로 표시한 영사이며, 그림 2(d)는 잡음이 제거된 대략적인 전경 물체를 보여준다.

추출된 전경물체에서 사전 정보를 얻기 위해 잡음에 강인한 정보를 얻을 수 있는 장점을 가지고 있는 Gaussian mixture models(GMM)을 이용하였으며, 위의 식 5는 제안된 방법의 데이터 항을 보여준다.

식 5의 μ_k^{fg} 와 Σ_k^{fg} 는 전경 물체를 위한 k번째 가우시안 모델의 평균과 공분산이다. 그래프를 완벽하게 설정하기 위해 데이터 항과 스무드 항이 그래프의 t -links와 n -links에 각각 할당되어야 하며, 이때 $D_p(f_p='fg')$ 는 $\{p, T\}$ 에 $D_p(f_p='bg')$ 는 $\{p, S\}$ 에 할당된다¹.

3. 3차원 얼굴 복원

앞장에서 설명 한 전경 물체 추출 과정을 거쳐 얻어진 결과를 이용하여, 3차원 볼륨 데이터를 생성한다. 3차원 볼륨 데이터 생성 과정은 볼륨 데이터 교차 영역

추출과 3차원 영역 수정의 두 단계로 이루어진다.

볼륨 데이터 교차 영역 추출 단계에서 우리는 다시점 입력 영상을 이용하여 3차원 voxel을 생성할 때 가장 일반적으로 사용되는 Silhouette Volume Intersection(SVI) 알고리즘을 사용한다. 이 방법은 서로 대응하는 카메라 시점으로부터 voxel 공간 상의 각 층에 변환된 실루엣 영상들을 투영시키고, 영상들 사이의 교차 영역을 계산한다(그림 3).

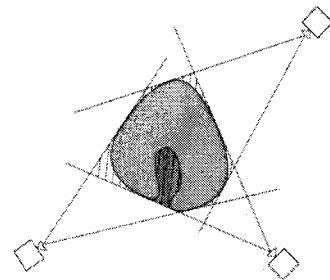


그림 3. 입력 영상의 교차영역 추출을 통한 3차원 볼륨 데이터 생성 개념도.

그림 4는 SVI 알고리즘의 세부 단계인 voxel 공간 각 층에 해당하는 plane 영상을 생성하는 단계와 생성된 plane 영상의 교차 영역을 계산하는 단계를 보여준다.

¹ 이 경우에 S 와 T 는 각각 전경과 배경을 표현한다.

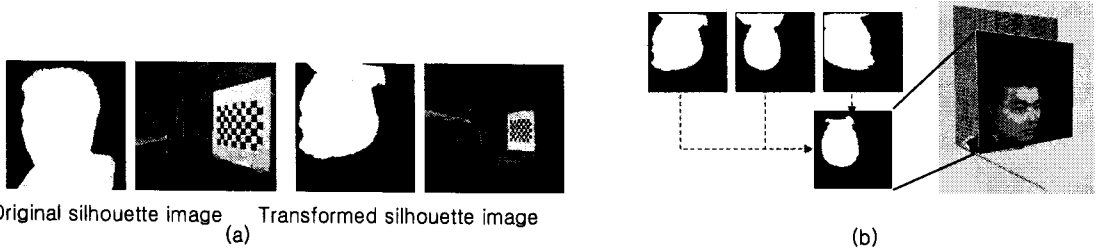


그림 4. SVI 알고리즘의 세부 단계: (a) plane 영상 생성 단계, (b) plane 영상 교차 영역 계산 단계.

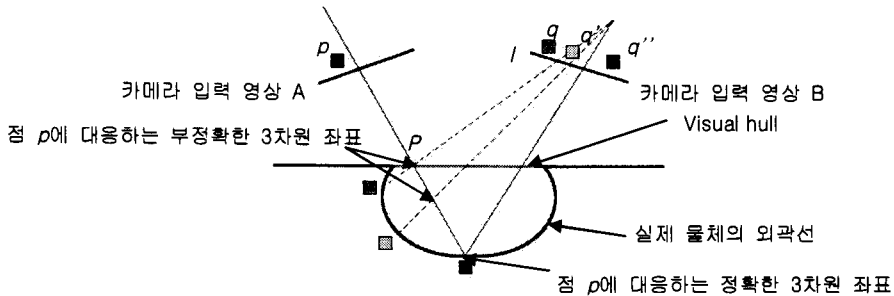


그림 5. 3 차원 데이터 수정 단계.

plane 영상 생성 단계에서는 각 카메라 마다 미리 계산된 캘리브레이션 행렬을 이용하여 모든 카메라 입력 영상을 하나의 축이 기준이 되게 변환하며, 축은 z-축을 기준으로 기본 축에서 각 축에 투영되는 영상이 없을 때까지 이동한다. plane 영상 교차 영역 계산 단계에서는 각 축마다 투영된 모든 plane 영상의 공통 영역을 계산한다. 일반적으로 visual hull이라 부르는 공통 영역은 해당 물체의 단층 영상을 의미한다. 카메라 수가 많으면 많을수록 visual hull의 정확도는 향상되지만, 그림 3에서 보는 것과 같이 SVI 알고리즘을 이용하는 방법으로는 물체의 오목한 부분(진한 빗금 영역)을 표현하지 못한다.

이러한 문제점을 해결하기 위해, 우리는 3차원 데이터 수정 단계를 사용한다. 이 단계에서 우리는 서로 이웃한 입력 영상 내의 픽셀 일치 여부를 visual hull 수정의 기준으로 사용한다. 그림 5에서 보는 것처럼, visual hull 상의 한 점 P 의 수정 여부를 확인하기 위해 카메라 입력 영상 A에서 P 에 대응하는 점 p 를 찾고, 이웃한 입력 영상 B에서 점 q 를 찾는다. 그리고 점 q 를 지나고 점 P 에 대응하는 epipolar line l 을 찾는다. 점 p 와 q 가 일치하지 않으면 점 P 를 물체 외부에 존재하는 점으로 인식하여 삭제하고, 점 p 를 이웃한 점 q' 와 다시 비교하여 일치할 때까지 visual hull 상의 점을 삭제한다.

그림 6은 3차원 얼굴 복원 과정을 거친 결과를 보여준다. 그림 6(a)는 볼륨 데이터 교차 영역 추출 후 텍스처 매핑을 한 결과이고, 그림 6(b)는 6(a)의 볼륨 데이터를 3차원 데이터 수정 단계를 거쳐 물체의 오목한 부분을 표현한 결과이다.

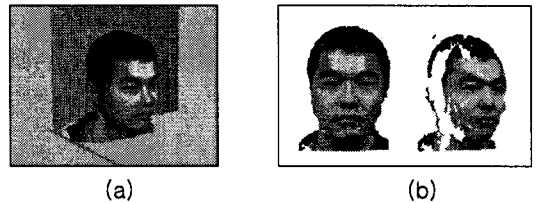


그림 6. 3차원 얼굴 생성 결과: (a) 볼륨 데이터 교차 영역 추출 결과, (b) 3차원 데이터 수정 결과.

4. 실험 및 결과

실험에서 사용한 컴퓨터는 intel Core2Quad Q6600 CPU와 nVidia GeForce 8800 GTX 그래픽 카드로 구성되어 있으며, 카메라는 6개의 Olympus E-500으로 구성되어 있다. 개발 툴로서 Microsoft C++ 6.0과 Olympus 카메라 SDK를 이용하였으며, 입력 영상의 크기는 800X600이다.

제안된 방법의 효능을 평가하기 위해 기존의 두 방법 차영상, Howe 등이 제안한 방법[10]과 제안된 방법을 비교하였으며, [10]에서 제안한 CCE(connected component error)(식 4)를 이용하여 평가하였다. 표 1은 각 방법에 대한 CCE를 보여준다.

표 1. 전경 물체 검출 결과 비교

Background Subtraction	Howe's	Ours
0.287	0.161	0.115

그림 7은 제안된 방법의 결과를 보여준다. 그림 7(a,c,e)는 입력 영상을, 그림 7(b,d,f)는 결과 영상을 보여준다. 그림 8은 제안된 방법의 결과 중 잘못된 추출된 결과를 보여주며, 이는 전경 물체에 생기는 그림자가 전경 물체를 위한 사전 정보에서 빠지기 때문에 생기는 문제점이다.

5. 결론

본 논문에서는 그래프 컷을 이용하여 3 차원 얼굴 복원을 위한 보다 효과적인 전경 물체 추출 방법을 제안하였다. 일반적으로 스테레오 비전 기술이 3 차원 복원을 위해 사용되면, 깊이 정보가 에너지 함수의 데이터 항을 위한 사전정보로도 이용된다. 하지만 3 차원 얼굴 복원에 적용할 때 얼굴 영역의 대부분이 유사한 색상이나 텍스처를 가지기 때문에 깊이 정보를 정확하게 구하기가 힘들며, 이로 인해 그래프 컷을 위한 정확한 사전정보를 얻기가 어렵다.

본 논문에서는 데이터 항을 설정하기 위해 차영상을 이용하여 대략적인 전경 물체 추출하고, 사전 정보에 대한 오류를 줄이기 위해 잡음과 그림자 영역을 제거하였다. 잡음과 그림자 영역을 제거하면 구멍이 발생하거나 실루엣이 손상되는 문제가 발생하지만, 손상된 정보는 근접한 픽셀이 유사하지 않을 때 낮은 비용을 할당하는 에너지 함수의 스무드(smooth) 항에 의해 에지 정보를 기반으로 채워진다.

하지만 만약 그림자 영역이 전경물체에 위치하면 최종 결과에서 전경물체의 일부가 제거되는 경우가 생기며, 우리는 차후 전경물체에 강건한 그림자 영역 추출에 대해 연구할 것이다.

References

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. P. Wayne Power, J. Schoonees, "Understanding Background Mixture Models for Foreground Segmentation," Proceedings of Image and Vision Computing, New Zealand, pp. 267-271, 2002.
3. Image Processing Toolbox, Chapter 9, "Morphological Operations," The Mathworks, 2001.
4. Y. Li, J. Sun, C-K Tang, H-Y Shum, "Lazy Snapping," ACM Transactions on Graphics, Vol. 23, Issue 3, pp. 303-308, 2004.
5. Y. Boykov, G. Funka-Lea, "Graph Cuts and Efficient N-D Image Segmentation," International Journal of Computer Vision, Vol.70, No. 2, pp. 109-131, 2006.
6. C. Rother, V. Kolmogorov, A. Blake, "GrabCut - Interactive Foreground Extraction using Iterated Graph Cuts," ACM Transactions on Graphics, Vol. 23, Issue 3, pp. 309-314, 2004.
7. J-H. Ahn, K. Kim, H. Byun, "Robust Object Segmentation using Graph Cut with Object and Background Seed Estimation," Proceedings of International Conference on Pattern Recognition, Vol. 2, pp. 361-364, 2006.
8. V. Kolmogorov, A. Criminisi, A. Blake, G. Cross, C. Rother, "Probabilistic Fusion of Stereo with Color and Contrast for Bilayer Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 9, pp. 1480-1492, 2006.
9. K. N. Kutulakos, S. M. Seitz, "A Theory of Shape by Space Carving," International Journal of Computer Vision, Vol. 38, No. 3, pp. 199-218, 2000.
10. N. R. Howe, A. Deschamps, "Better Foreground Segmentation through Graph Cuts," Technical Report, <http://arxiv.org/abs/cs.CV/0401017>, 2004.
11. Y. Sun, B. Li, B. Yuan, Z. Miao, C. Wan, "Better Foreground Segmentation for Static Camera via New Energy Form and Dynamic Graph-cut," Proceedings of International Conference on Pattern Recognition, Vol. 2, pp. 49-52, 2006.
12. Y. Boykov, O. Veksler, R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 11, pp. 1222-1239, 2001.
13. L. Ford, D. Fulkerson, "Flows in Networks," Princeton University Press, 1962.
14. M. P. Kumar, P. H. S. Torr, a. Zisserman, "Obj Cut," Proceedings of Computer Vision and Pattern Recognition, Vol. 1, pp. 18-25, 2005.

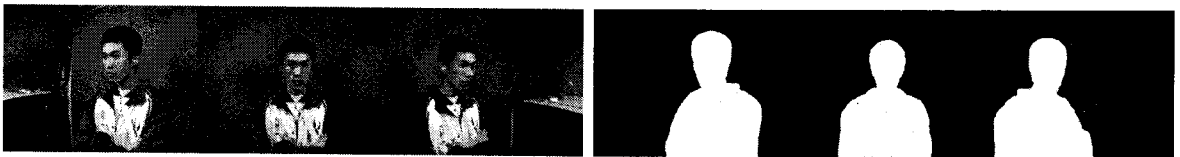


그림 7. 제안된 방법의 결과 영상: (a) 입력 영상, (b) 결과 영상.



그림 8. 오추출 결과: (a) 입력 영상, (b) 차영상, (c) 결과 영상.