

## 얼굴 모델링을 위한 검색 기반 헤어 모델 증강 기법

이정우, 원소미, 박인규  
 인하대학교 정보통신공학과

leejw2807@gmail.com, yuanxiaoweichn@gmail.com, pik@inha.ac.kr

### Retrieval-Based Hair Model Augmentation for Face Modeling

Jung-Woo Lee, Xiaowei Yuan, and In Kyu Park

Department of Information and Communication Engineering, Inha University

#### 요 약

주어진 영상으로부터의 3 차원 얼굴 모델링은 얼굴 분석, 애니메이션, 생체 인식 등의 많은 컴퓨터비전 및 그래픽스 응용분야에서 중요한 역할을 하고 있다. 그 중에서도 헤어 영역은 얼굴에 비해 모양의 다양성과 모델의 복잡도가 현저히 높다. 기존의 연구는 주로 얼굴 영역에 한정된 3 차원 얼굴 모델링을 중심으로 이루어졌지만 헤어 모델링은 중요하게 다루지 않고 있는 경우가 많다. 본 논문에서는 심층인공신경망의 일종인 FCN (fully connected network)을 이용하여 인물 영상에서 헤어 부분을 영역화하고 그와 가장 유사한 3D 헤어 모델을 데이터베이스에서 검색하여 3 차원 얼굴 모델에 증강함으로써 완전한 얼굴 모델링을 수행하는 방법을 제안한다. 이는 FCN 을 이용하여 다양한 인물 영상에 대하여 네트워크 학습을 수행하는 과정과 3D 헤어 데이터베이스의 구축 과정을 포함한다. 실험 결과 적절한 수준의 헤어 모델이 3 차원 얼굴 모델링 결과물에 증강됨을 확인하였다.

#### 1. 서론

인체 3D 모델링은 3D 프린터, 증강현실과 가상현실 등의 많은 컴퓨터비전 및 그래픽스 응용분야에서 중요한 역할을 하고 있다. 3 차원 헤어 모델은 그 다양성과 복잡함 때문에 매개 변수 모델로 모델링하기가 매우 어렵다.

본 논문에서는 검색 기반의 3 차원 헤어 모델링 기법을 제안한다. 우선 입력된 인물 영상에 대해 FCN (fully convolutional network) [1]을 사용하여 헤어 부분의 영역화를 수행한다. 또한 영역화 된 헤어 부분 영상과 가장 일치하는 3D 헤어 모델 검색을 위해 헤어 모델들을 수집하여 데이터베이스를 구축한다. 입력 영상에서의 헤어 영역과 수집된 각각의 헤어 모델들의 비교를 위해 3 차원 헤어 모델의 정면도를 구하고, 정면도의 헤어 부분의 중횡비(aspect ratio)와 HoG (histogram of oriented gradients)를 구하여 각 3 차원 헤어 모델의 메타 데이터를 생성한다. 입력된 영상에서 영역화 된 헤어 영상에서도 동일하게 중횡비와 HoG 를 구하고, 입력된 영상의 중횡비와 일정 수준 이내로 가까운 데이터베이스 내의 3 차원 헤어 모델들 중 HoG 값이 가장 가까운 모델들을 검색 결과로서 사용자에게 제안하게 된다.

#### 2. 제안하는 기법

본 절에서는 영상 영역화에 주로 쓰이는 FCN 기반 학습을

이용하여 입력된 인물 영상에서 헤어 부분을 영역화 하는 단계와 입력 헤어 영역과 가장 일치하는 헤어 모델을 3 차원 헤어 모델 데이터셋에서 검색하여 생성하는 단계를 세부적으로 설명한다.

##### 2.1. 헤어 영역화를 위한 FCN 학습

본 논문에서는 입력된 인물 영상에서 헤어 부분을 영역화 하기 위해 FCN 을 사용한다. 본 논문에서는 그 중에서도 인물의 영역화에 집중한 VGG16 을 기반으로 만들어진 FCN[2]을 사용하였다.

학습 데이터는 [2]에서 제공하는 1,466 개의 인물 사진을 Grab-Cut 영역화 기법[3]을 사용한 반자동 수작업으로 헤어 영역을 분리하고 데이터 증강을 위하여 좌우 반전 후 이진화 하여 생성한 총 3 천여개의 영상으로 구성하였다.

##### 2.2. 3 차원 헤어 모델 데이터 베이스 생성 및 검색

본 논문에서는 여러 3 차원 헤어 모델을 수집하여 데이터 베이스를 구축하고 입력된 인물 영상의 구분된 헤어 영역과 가장 유사한 헤어 모델을 검색하는 방법을 제안한다.

데이터 수집은 다음과 같이 진행된다. 온라인 게임 커뮤니티에서 [4]에서 게임 유저들에게 공유되고 있는 서로 다른 187 개의 헤어 모델을 수집하고 영역화 된 헤어와 비교하기 위해 각각의 헤어 모델의 정면도를 구한다. 정면도는 3D 헤어 모델의 뒷부분 절반을 삭제한 후 정면으로 투영한 영상을 이진화 하여 구하며, 그 후 각각의 정면도에 대해서 HoG 와 중횡비를 구한다. 즉, HoG 는 투영된 헤어 모델의

전경과 배경 경계의 에지 방향을 히스토그램으로 저장한 것이다. 중형비는 영상에서 실제 머리가 존재하는 부분의 세로 화소 수를 가로 화소 수로 나눈 수치이다.

본 논문에서 제안하는 검색 기법은 다음과 같다. 입력 얼굴 영상에서 FCN 을 통해 분리된 헤어 영역에 대해 HoG 와 중형비를 구한다. 이후 3 차원 헤어 모델 정면도의 헤어 영역이 갖는 중형비의  $\pm 20\%$ 이내의 중형비를 갖는 데이터베이스 내의 정면도들을 후보로서 우선 검출한다. 마지막으로 후보 정면도 중 HoG 벡터가 입력 헤어 영상의 HoG 벡터와 L2 거리가 가장 일치하는 순으로 사용자에게 제안한다.

### 3. 실험 결과

본 논문의 실험은 Intel i7-7700 3.6GHz CPU 와 NVIDIA GeForce GTX 1080 Ti 를 장착한 컴퓨터에서 수행하였다. FCN 네트워크에서는 stochastic gradient descent (SGD) solver 와 softmax 손실 함수를 사용하였으며 학습율은  $10^{-4}$ 이 적합함을 경험적으로 찾아냈다. 600x800 RGB 영상에 대해서 영역화는 0.5 초, 헤어 모델 검색에는 1 초의 실행시간이 소요되었다.

그림 1 에 영역화 결과와 3 차원 헤어 모델 검색 결과를 제시하였다. 입력 영상과 중형비와 HoG 값이 가까운 상위 3 개의 모델을 도시하였다. 상대적으로 헤어스타일의 다양성이 적은 남성의 정확도가 여성의 정확도보다 높음을 알 수 있다.

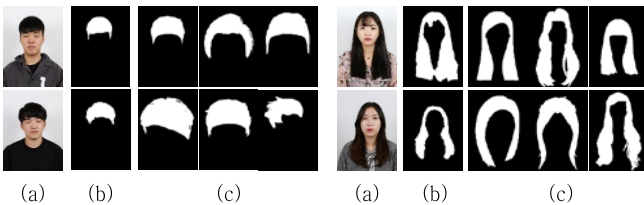


그림 1. 헤어 모델 검색 실험결과. (a) 입력 영상. (b) 헤어 영역화 결과. (c) 검색된 상위 3 개의 3 차원 헤어 모델의 정면도.

그림 2 는 검색된 헤어 모델을 3 차원 얼굴 모델링 결과에 증강하여 전체적인 머리 모델링을 수행한 결과이다. 실험에 사용된 영상은 직접 취득한 인물의 정면 영상이며, 얼굴 모델링은 [5]의 기법을 사용하였다. 헤어 모델은 사전에 동일한 위치로 설정을 하여 미세한 상하 조정으로 얼굴 모델과의 정렬이 가능하다. 위에서 언급하였듯이 남자의 경우 비교적 정확한 모델링이 이루어진 반면 여성의 모델링은 첫번째 여성의 경우 전체적인 양상은 비슷하지만 세부적인 모양이 다르거나 두번째 여성의 경우처럼 장발이 가려져 단발로 모델링이 된 것을 볼 수 있다. 이것은 데이터베이스내의 여성 헤어스타일이 다양하지 않기 때문에 발생하는 문제이며, 이는 데이터베이스의 규모를 늘림으로써 해결할 수 있다.

### 4. 결론

본 논문에서는 FCN 기반의 학습을 이용한 인물 영상에서 헤어 부분 영역화 기법과 검색 기반 3 차원 헤어 모델링 기법을 제안하였다. 제안한 방법은 1466 쌍의 입력 영상과 그의 헤어 영역 영상으로 학습된 FCN 을 통해 헤어 부분 영역화를 수행한다. 영역화 결과를 바탕으로 187 개의 3 차원 헤어

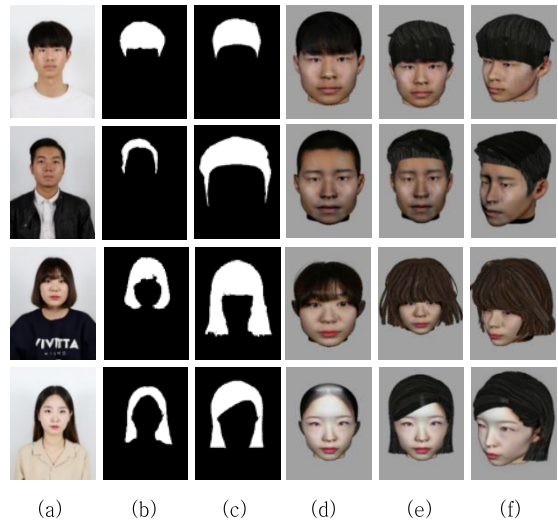


그림 2. 복원된 헤어 모델이 증강된 얼굴 모델링 결과. (a) 입력 영상. (b) 헤어 영역화 결과. (c) 최상위 유사도를 갖는 헤어 모델 검색 결과. (d) 얼굴 모델링 결과. (e) (f) 복원된 3 차원 헤어 모델이 증강된 머리(얼굴 및 헤어) 모델.

모델로 구성된 데이터베이스에서 가장 유사한 헤어 모델을 검색하여 3 차원 얼굴 모델링의 결과에 증강한다. 실험 결과 다양한 인물의 입력 영상에 대해서 유사한 헤어 모델이 생성되어 3 차원 얼굴 모델링의 현실성을 높임을 확인하였다. 향후 더 큰 규모의 헤어 데이터베이스의 구축을 통해 실용적인 3 차원 인물 모델링 시스템으로 확장할 수 있다.

### 감사의 글

본 논문은 CoAsia Holdings 의 지원을 받아 수행된 연구입니다.

### 참고문헌

- [1] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *Proc. IEEE CVPR*, pp. 3431-3440, June 2015.
- [2] X. Shen, A. Hertzmann, J. Jia, S. Paris, B. Price, E. Shechtman, and I. Sachs, "Automatic portrait segmentation for image stylization," *Computer Graphics Forum*, vol. 35, no. 2, pp. 93-102, May 2016.
- [3] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: interactive foreground extraction using iterated graph cuts," *ACM Trans. on Graphics*, vol. 23, no. 3, pp. 309-314, July 2004
- [4] *Maysims*, <https://www.maysims.com/>
- [5] Xaiwei Yuan, Myung Rai Cho, and In Kyu Park, "3D head modeling from uncalibrated three images," *Proc. IPIU*, February 2018.