

통계적 신체 외형 데이터베이스를 활용한 실루엣으로부터의 3차원 인체 외형 예측

안다솔¹

박상일^{2*}

¹넥슨 게임즈 주식회사, ²세종대학교 소프트웨어학과

¹award09130@nexongames.co.kr, ²sipark@sejong.ac.kr

3D Human Shape Estimation from a Silhouette Image by using Statistical Human Shape Spaces

Dasol Ahn¹ Sang Il Park^{2*}

¹Nexon Games Co., Ltd, ²Department of Software, Sejong University

요 약

본 논문은 특정 자세를 취하고 있는 인간형 캐릭터의 단순한 실루엣 그림 등을 입력으로 받아 이로부터 3차원 신체 외형 정보를 복원하는 기술을 다룬다. 기하학적인 방법 만으로 실루엣에 담기지 않은 내부 형태를 추출하는 것은 매우 어려운 일이기 때문에 본 논문은 대규모 신체 3차원 형태 데이터베이스 및 이의 통계적 분석 데이터를 활용하여 인간형 템플릿 모델을 실루엣에 맞도록 정합 하는 것을 주된 아이디어로 한다. 본 기술은 다음의 세 단계로 이뤄진다. 먼저 주어진 실루엣 이미지로부터 이미지 분석을 통해 윤곽선 및 법선 벡터를 추출한다. 둘째로, 실루엣 이미지와 3차원 모델 간의 점 대 점 대응관계를 수립한다. 마지막으로 수치적 최적화를 통해 실루엣과 모델이 최대한 일치하도록 파라미터들을 정한다. 본 논문은 실루엣 정보로부터 3차원 신체 형태를 최적화하는 실용적 방법론을 제시한 것이 주된 기여이며, 기술의 유효성을 실험을 통해 검증하였다.

Abstract

In this paper, we present a method for estimating full 3D shapes from given 2D silhouette images of human bodies. Because the silhouette only consists of the partial information on the true shape, it is an ill-posed problem. To address the problem, we use the statistical human shape space obtained from the existing large 3D human shape database. The method consists of three steps. First, we extract the boundary pixels and their appropriate normal vectors from the input silhouette images. Then, we initialize the correspondences of each pixel to the vertex of the statistically-deformable 3D human model. Finally, we numerically optimize the parameters of the statistical model to fit best to the given silhouettes. The viability and the robustness of the method is demonstrated with various experiments.

키워드: 형태 모델링, 인간 체형 공간, 형상 최적화, 실루엣 영상

Keywords: Shape Modeling, Human body space, Shape optimization, Silhouette image

1. 서론

게임, 영화, 패션 등 다양한 분야에서 실제와 같은 인간형 캐리

터 모델은 많이 사용되고 있다. 하지만 개개인마다 체형이 다양하고 인체 특유의 해부학적 복잡성 때문에 이를 수작업으로 생성하는 것은 어렵고 시간이 오래 걸리는 작업이며, 따라서 보다 손쉽게 모델링을 하기 위해 많은 방법이 개발되어왔다. 기존의

*corresponding author: Sang Il Park / Department of Software, Sejong University (sipark@sejong.ac.kr)

방법 중 전문적인 3차원 스캐너를 이용해 인체를 모델링하는 것이 있으나 이는 전문 장비의 구비 유무와 스캔 데이터의 후처리 과정의 비용 등으로 상용 애플리케이션으로 널리 사용하기에는 무리가 있다. RGB-D 카메라 또는 다 시점(multi-view) 사진을 이용해 인체를 모델링하기 위한 시도 또한 활발하게 이루어졌는데 RGB-D 카메라의 스캔 데이터는 노이즈가 많이 발생하며, 다 시점 사진을 이용해 인체를 모델링할 경우 조명과 카메라 등의 장비 세팅의 번거로움과 비용적인 측면에서 비효율적일 수 있다. 본 연구는 이러한 전문가의 기술이나 고비용의 장비 혹은 번거로운 후처리과정을 생략하고도 사실적인 인체를 모델링할 수 있는 방법을 제안하려 한다.

본 연구에서는 간편한 입력으로 실루엣을 담은 이미지나 간단하게 그린 그림 한 장 또는 여러 장을 입력 받아 이로부터 3차원 인체 모델을 유추하는 방법을 제안한다. 이미지로부터 실루엣 정보만을 사용하기에 반드시 실제 사진일 필요도 없으며, 간단히 사용자가 그린 스케치 만으로도 충분히 활용될 수 있어 사용성을 높일 수 있는 장점이 있다. 이러한 방법론에서는 실루엣 정보에 담기지 않은 내부 정보를 유추하는 것이 핵심이 되며, 우리는 대규모 신체 외형 데이터를 분석하여 얻은 통계적 신체 외형 데이터를 활용하는 것으로 이 문제를 해결하려 한다. 이 경우 신체 외형 예측은 3차원 인체 기준 모델(템플릿 모델)을 주어진 실루엣과 일치하도록 통계적 방법으로 변형하여 일치시킴으로써 이뤄질 수 있다. 이 모든 과정은 사용자의 상호작용 없이 자동으로 이뤄지며, 다양한 체형에 대해 강건하게 사실적인 인체를 재구성할 수 있다.

본 논문의 주된 기여는 실루엣으로부터 3차원 신체 형태를 최적화하는 실용적이며 효율적인 방법론을 제시하였다는 것이며, 또한 다양한 실험을 통해 이의 유용성 및 유효성을 검증하여 실제 충분히 활용 가능한 점을 보인 것이다. 본 논문의 구성은 다음과 같다. 2장에서는 관련된 기존 연구를 살펴보고, 3장에서는 방법의 전체적인 개요를, 그리고 4장에서는 본 기술에 활용된 통계적 인체 모델에 대해 설명한다. 5장에서는 통계적 모델을 활용한 최적화 방법론에 대해 제시하고, 6장에서는 실험 결과 및 검증을 제시하며, 마지막으로 결론과 논의할 점들을 제시한다.

2. 관련 연구

이미지로부터 인간의 체형을 3차원으로 유추 및 복원하는 기존의 기술들은 인체에 대한 사전 정보 없이 주어진 기하학적 정보 만으로 유추해 내는 기하학적 기반 방법과, 인체 기준 모델을 변형함에 있어 일반적인 인체에 대한 통계적 정보를 활용하는 방

법론으로 나뉜다. 또한 인체 정보는 자세 정보와 체형 정보로 나뉘는데 본 연구에서는 다양한 자세의 3차원 체형을 유추하는 문제를 연구 대상으로 하지 않고 그림 1에서 보이는 것과 같은 특정한 자세(A 포즈)에서의 체형만을 대상으로 한다.

2.1 다시점 및 기하학 기반 3차원 인체 재구성

다 시점(multi-view) 사진이나 RGB-D 카메라를 통해 3차원 물체 재구성하는 방법은 기존에 많이 연구되어 왔다[3, 7, 8, 9, 12, 18]. 그러나 이러한 방법들은 여러 시점 간의 카메라 보정(Camera Calibration)을 통해 복잡한 계산 과정을 거쳐야 하며 특수한 센서와 조명이 설치된 환경을 조성해야 한다는 번거로움이 뒤따른다. 이러한 문제점을 해결하기 위해 한 장 내지 두 장의 사진만을 이용한 3차원 재구성 기법들도 활발하게 연구되고 있다. 2009년에 Kraevoy 등은 입력 받은 실루엣에 맞게 기준 인체 모델을 변형시킴으로써 물체를 3차원으로 재구성하였다[4]. 이 방법은 반복적 최단 거리 정점(iterative closest point) 방법을 이용해 실루엣과 템플릿 모델을 매칭시켜 둘 사이의 기하학적 오차를 최소화 방법이다. 다양한 물체에 대해서 3차원 재구성을 수행할 수 있다는 장점이 있지만, 템플릿 모델의 초기 정렬을 해주어야 하고 추가적인 사용자 제약조건이 필요한 경우가 자주 발생한다는 단점이 있다. 2010년에는 Zhou 등이 키, 몸무게와 같은 직관적인 특성에 대한 매개변수를 조작하여 사진 속 사람의 체형을 수정하는 방법을 제안했다[6]. 이 방법은 입력 받은 이미지와 인체의 직관적 특성에 대한 정보를 가진 매개변수 인체 모델을 매칭한 다음, 매개변수에 따라 변형시킨 후 다시 이미지로 렌더링 함으로써 결과를 도출해냈다. 본 연구의 목표 또한 실루엣을 기반으로 3차원 모델을 재구성하는 것으로 이와 유사하지만 이와 다르게 하나의 템플릿 모델을 이용하는 것이 아니라 통계적 모델을 통해 획득한 변형 가능한 3차원 인체 통계 모델을 이용하며, 또한 윤곽선 내부는 고려하지 않는 차이점이 있다.

2.2 통계적 방법론에 기반을 둔 3차원 인체 재구성

실루엣으로부터의 체형 유추 기술은 체형과 자세에 대한 매개변수 값에 따라서 3차원 모델을 변형시킬 수 있는 변형 가능한 3차원 체형 모델을 기반으로 발전되어왔다. 2003년에 Allen 등은 특정 자세를 취하고 있는 실제 인간에 대해 범위 스캔(range scan)을 수행하여 얻은 3차원 인체 메쉬 모델들의 주성분 분석(Principal Component Analysis, PCA)을 통해 매개변수 변형 모델을 구축했다[1]. 이 모델은 매개변수 값에 적절한 변화를 주어 현실적인 인체의 모습을 효율적으로 표현할 수 있다. 또한 2005년에 Anguelov 등의 SCAPE 모델은 인간의 자세와 체형에 대해서 변형 가능한 3차원 모델을 구축했다[2]. 먼저 한 사람의 자세에

대한 모델을 별도로 세운 후 여러 사람의 체형에 대한 모델을 결합하는 방법으로 3차원 모델을 구축하였다. 이 방법과 이후의 비슷한 연구에서는 인체 모델의 변형을 메쉬의 삼각형을 변형하는 방식으로 모델을 구축했다. 이로 인해 정점 좌표로 표현되는 인체 메쉬 데이터를 삼각형 좌표로 표현되는 변형 모델로 변환하여 표현해야 했고, 이는 전체적인 계산을 까다롭게 하였다. 이를 해결하기 위해 단순화된 버전의 SCAPE(simplified-SCAPE, S-SCAPE) 모델이 제안되었다[5]. 이 방법은 정점 좌표를 직접적으로 변형하는 모델이며 자세에 대한 변형은 골격 기반(skeleton-based)의 스킨닝(skinning) 방법을 통해 해결하였다. 2017년에는 MPI 연구소의 인체 프로젝트에서 CAESAR[11] 모델을 기반으로 이를 PCA로 분석하고 그 통계 데이터를 MPII Human Shape라는 이름으로 공개하였다. 그 동안 발전되어온 3차원 인체 모델 처리 기술을 적용하여 더 정확한 3차원 정합 모델을 구성하였으며, 4000여명 이상의 방대한 실측 모델이 되었다[10]. 본 논문에서도 이 MPII Human Shape 데이터를 활용하여 인간의 체형 추정 문제를 통계적 최적화를 통해 수행한다. 그러나 일반적인 3차원 절점이나 측량 데이터들을 입력으로 받는 것이 아닌 2D 실루엣 정보만을 받아 최적화를 수행하는 것이 기존의 방법들과 차이점이다. 위 모델들이 성별, 인종과 상관없는 정보였다면, Choi 등은 한국인 여성 2200여명을 기반으로 한 인체 분석 데이터를 공개하였다[19].

2.3 딥러닝 기반 3차원 인체 재구성

최근 딥러닝을 이용하여 이미지 한 장으로부터 3차원 신체를 재구성하는 연구들이 활발하게 진행되고 있다. 이러한 방법들은 위에 기술한 통계적 모델에 기반한다기 보다는 하나의 이미지에서 깊이정보 또는 3차원 기하를 추출하는 방법의 연속성상에서 발전되어 왔다. Varol 등은 이미지에서 인체의 볼륨 정보로 재구성하는 방법을 제시했으며[14], Zheng 등은 이후 이보다 더 고해상도의 볼륨을 재구성하는 방법을 제안하였다[15]. 최근 Saito 등은 Pixel-Aligned Implicit Function(PIFu) 방법론에 기반을 둔 인체 추출 방법들을 제안했다[16,17]. 이러한 딥러닝 기반 방법들은 인체 사진으로부터 3차원 볼륨 또는 기하를 얻어내는 데 성공적으로 적용되고 있다. 그러나, 본 연구는 딥러닝에 기반하지 않고 통계학적 모델에 기반하는데, 이는 활용면에서 실루엣 정보만을 활용할 수 있어 실제 사진이 아닌 일반적인 스케치나 드로잉에 대해서도 적용 가능하다는 차이점이 있다.

3. 연구의 개요

본 연구는 입력 받은 한 장의 인체 실루엣과 실제 인체의 통계에

바탕을 둔 변형 가능한 3차원 인체 모델을 이용해 입력 사진 속의 인체를 3차원으로 재구성하는 것을 목표로 하며 다음과 같은 3 단계의 과정으로 이뤄진다.

- 1) 입력 실루엣 이미지의 분석을 통해 윤곽선에서의 픽셀들과 이들의 법선 방향을 추출한다.
- 2) 인체 템플릿 모델의 실루엣 대응 정점 후보 집합을 구하고 이 정점들과 입력 사진 속 실루엣 픽셀들의 초기 대응관계를 설정한다.
- 3) 인체 템플릿 모델을 실루엣 대응점과의 기하학적 오차가 최소화되며 형태 변형의 통계적 가능성도(likelihood)가 최대화될 수 있도록 최적화를 수행하여 3차원 인체 모델을 유추한다.

4. 통계적 3차원 인체 변형 모델

본 장에서는 입력 사진을 3차원으로 재구성하는데 필요한 3차원 통계적 인체 변형 모델에 대해 설명한다. 이 모델은 매개변수 값에 따라 인체의 모습을 사실적으로 변형시킬 수 있으며, 다변량 분석의 한 방법인 주성분 분석(Principal Component Analysis, PCA)를 통해 만들어진다. 본 연구에서는 MPI[10]에서 공개한 PCA 인체 모델을 사용한다. 이 모델은 4000명 이상의 CAESAR[11] 실측 모델에 대해 [1,5]와 같은 발전되어온 3차원 인체 모델 기술을 이용해 보다 더 사실적이고 효율적인 인체 모델을 재구성했다. 따라서 MPI에서 제공하는 모델은 현존하는 공개 데이터 중 가장 폭 넓고 활용성이 높은 모델이다. 4000여 개의 CAESAR 모델은 각각 정점의 개수와 정점의 연결구조가 다른 삼각형 메쉬(triangular mesh) 구조이며, MPI에서는 표준화된 템플릿 모델을 각각의 CAESAR 모델에 정합시켜 동일한 정점의 개수와 연결 구조를 갖는 메쉬 모델로 변형시켰다. 따라서 동일한 인덱스에서의 정점은 대상 모델과 상관없이 모두 동일한 의미를 갖는다.

4.1 주성분 분석을 통한 3차원 인체 모델 구성

MPI 인체 모델은 이렇게 구성된 각각의 모델들을 가지고 정점들 간의 상호 상관관계를 파악하기 위해 주성분 분석을 수행하여 인체 통계 모델을 제공하고 있다. 주성분 분석은 고차원의 데이터에 대해 표본 간의 공분산이 최대가 되도록 상호 직교하는 독립된 축을 계산하여 표본의 차이를 가장 잘 나타내는 성분들로 분해하는 통계적 방법이다. PCA를 통해 얻은 주성분은 표본

간 공분산 행렬의 고유벡터(eigen vector)이며 고유값(eigen value)의 크기에 따라 내림차순으로 그 중요도가 정해진다. 또한 고유값은 데이터 분포의 표준편차 의미로도 해석될 수 있어 변형하여 얻은 모델의 우도(likelihood)를 예측하는데 유용하다.

MPI 인체 모델은 4300개의 샘플 모델(각 모델은 20000여개의 정점으로 구성)에 대해 주성분 분석을 수행하여 고유 벡터와 고유 값, 그리고 전체 샘플의 평균 모델을 제공하고 있다. 주성분 분석은 모델의 정점을 표현하는데 필요한 자유도 (즉 정점의 개수 $N \times 3 = 60000$)의 개수와 같은 수의 고유벡터를 얻게 되나, 본 데이터의 경우 샘플의 숫자가 자유도 보다 적기에 고유값이 0이 아닌 의미 있는 주성분은 샘플 수와 같은 총 4300여개를 얻게 된다. 그림 1은 이렇게 구한 PCA의 상위 4개의 고유 벡터로 재구성한 모델을 보여준다. 또한 본 연구에서 사용하는 모델은 그림1의 중앙에서 보여지는 평균 모델을 템플릿 모델로 정하고, 이 모델에 고유 벡터를 매개변수의 값만큼 적용함으로써 변형시킬 수 있는 PCA 인체 모델을 사용한다. 이 모델은 다음과 같은 식으로 표현된다.

$$\tilde{V} = \bar{V} + wE. \quad (1)$$

매개변수의 값에 따라 변형이 가능한 PCA 인체 모델 \tilde{V} 는 템플릿 모델 \bar{V} 와 매개변수의 값만큼 곱해진 고유 벡터 wE 의 합으로 나타낼 수 있다. 본 절 이후로 논문에서는 위와 같이 표현되는 인체 변형 모델을 템플릿 모델이라고 칭한다.

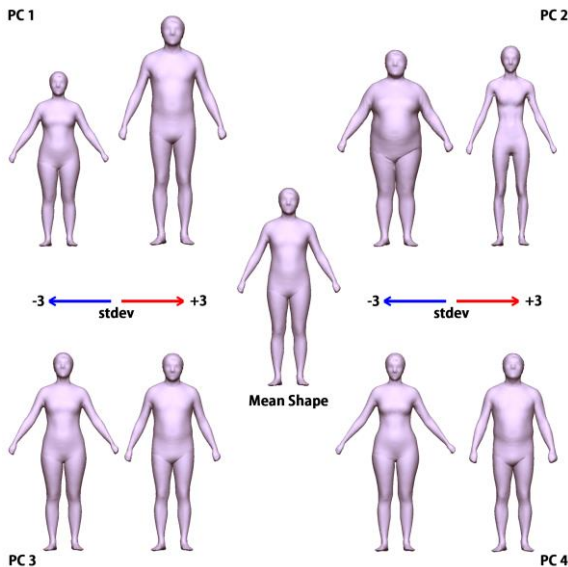


Figure 1. The mean shape of the database(center) and its deformed body shapes using the first four principal components of the MPII Human Spaces dataset [10].

5. 실루엣으로부터의 3차원 인체 재구성

본 장에서는 입력 받은 사진 속 윤곽선에서의 정점 좌표와 법선 벡터를 추출하고, 이를 바탕으로 4장에서 기술한 템플릿 모델과 매칭하여 최종적으로 입력 실루엣에 대한 3차원 인체를 재구성하는 방법을 설명한다. 입력 받은 실루엣 이미지는 배경이 단색 또는 투명으로 되어있으며, 바라보는 방향이 주어진다고 가정한다. 좌표축은 그림 2(a)과 같이 바라보는 방향을 -z축으로 위쪽 방향을 y축으로 (y-up), 그리고, 캐릭터의 오른쪽 방향을 x축으로 되어있다고 가정한다. 따라서 입력 실루엣은 xy 평면에 투사한 것이고, 투영 방향은 평면의 법선 벡터로 입력 받으며 z축 방향이 된다.

5.1 실루엣 분석을 통한 윤곽선 및 법선 방향 추출

입력 실루엣은 인체의 정면 모습을 담은 이미지이다. 본 절에서는 이러한 입력 실루엣을 분석하여 템플릿 모델과의 매칭에 필요한 윤곽선에서의 정점 좌표와 법선 벡터를 구하는 방법에 대해 설명한다. 입력 실루엣 이미지는 배경이 단색이거나 투명하다는 가정을 통해 쉽게 이진화 할 수 있다. 그림 2(b)는 입력 실루엣 이미지의 이진 영상이다. 외각선을 추출하기 위해서는 캐니 엣지 디텍터(Canny Edge Detector) 등을 이용할 수도 있으나, “이미지 내부에 속하면서 3x3 주변에 외부에 포함되는 점이 존재하는 점”이라는 조건을 사용하여 쉽게 윤곽선에 해당하는 픽셀들만 추출할 수 있다(그림 2(c)).

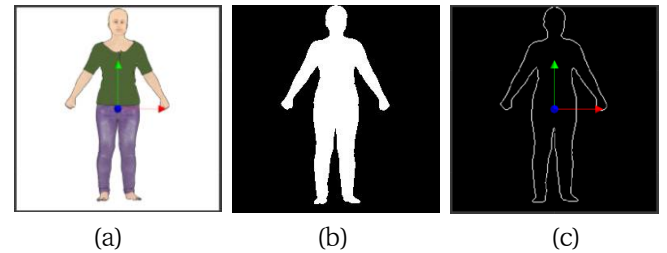


Figure 2. Input and preprocessed silhouette images: (a) an input silhouette; (b) binary image from (a); (c) extracted silhouette pixels from (b).

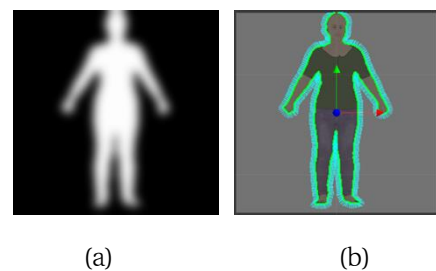


Figure 3. (a) the blurred binary image; (b) Computed normal vectors from the negative gradient directions.

이에 구해진 외각선의 각 pixel들은 향후 템플릿 모델을 정합 할 때 맞춰줘야 할 목적 점들로 사용하게 된다. 이때 템플릿 모델의 정점들과의 대응점을 찾는 과정이 필요하며 단순히 점의 위치만을 고려하게 되면, 그림과 템플릿모델의 초기 크기 및 위치에 따라 잘못된 대응점들을 찾게 될 수 있다. 이러한 문제를 해결하기 위해 추출된 외각선 각 픽셀들에 대해 표면의 법선 방향을 추가로 설정해 준다. 이미지에서 인체 내부에서 외부를 바라보는 방향을 법선 벡터로 설정할 수 있으며, 이는 이진화된 이미지에서 이미지 그래디언트 방향을 구함으로써 쉽게 설정할 수 있다. 이미지 그래디언트는 다음 식과 같이 주어진다.

$$\nabla I(x, y) = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right). \quad (2)$$

이 때, 윤곽선에서의 이미지 그래디언트가 부드럽게 변하도록 하기 위해 이진 영상에 가우시안 블러(Gaussian Blur)처리를 해 준다. 이미지의 인체의 내부에서 외부로 향하는 법선 벡터는 그래디언트의 반대방향이 되며 그림 3은 이렇게 얻은 법선 벡터를 보여준다.

5.2 입력 실루엣과 템플릿 모델 매칭

입력 받은 실루엣에 맞게 인체 변형 모델(템플릿 모델)을 변형 시키기 위해선 둘 사이의 대응관계가 필요하다. 이를 구하기 위해 5.1절에서는 입력 실루엣 이미지에 대해서 윤곽선에서의 정점 위치와 법선의 방향에 대한 정보를 얻었고, 본 절에서는 템플릿 모델의 정점들로부터 윤곽선에 해당되는 점들을 추출하는 방법을 설명한다. 나아가 입력 실루엣과 템플릿 모델 각각의 윤곽선 정보를 바탕으로 상호 매칭하는 방법에 대해 설명한다.

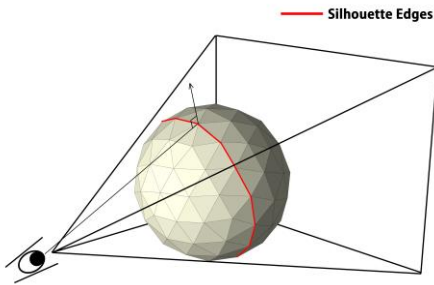


Figure 4. Extracted silhouette information from 3D

입력 받은 실루엣 이미지의 투영 방향과 템플릿 모델의 정점 법선 벡터(vertex normal)가 서로 수직인 경우라면 해당 정점은 템플릿 모델의 실루엣으로 간주할 수 있다. 이는 두 벡터의 내적의 값을 확인함으로써 쉽게 이뤄진다. 그림 4는 이러한 실루엣 엣지 및 정점들의 개념을 보여준다. 본 연구에서는 두 벡터 내적의 절대 값이 0.1 이하가 되는 정점들에 대해서만 실루엣 후보군으로 선정하였다.

이제 템플릿 모델의 실루엣 정점과 입력 실루엣에서 픽셀과의 대응관계를 찾는다. 본 연구에서는 이를 위해 정점 사이의 거리와 법선 방향 두 가지를 고려한다. 템플릿 모델의 실루엣 후보 정점 v , 입력 실루엣 이미지의 각 픽셀을 p 라고 할 때, 두 정점 집합의 거리는 입력 받은 실루엣 이미지의 평면에서 계산되며, 실루엣 이미지를 입력 받을 때 투영방향을 $(0, 0, 1)$ 로, 투영평면을 xy 로 가정하였으므로, 다음과 같이 표현할 수 있다.

$$dist_p = (p_x - v_x)^2 + (p_y - v_y)^2. \quad (3)$$

또한, 템플릿 모델의 표면 방향과 실루엣의 법선 방향을 고려하기 위해 두 방향이 어느 정도 같은 방향을 향하고 있는지를 측정한다. 이를 위해, 두 방향 벡터 간의 내적을 이용하여 얼마나 비슷한 방향을 향하는 지 측정한다. 이 때 실루엣 이미지 정점의 법선 벡터 n_p 와 템플릿 모델의 정점 법선 벡터 n_v 의 내적은 실루엣 이미지 평면에서 계산되며, 다음과 같이 표현된다.

$$d_n = n_p \cdot n_v \quad (4)$$

따라서 본 연구에서는 $d_n > 0$, 즉 같은 방향을 바라보는 두 정점들에 대해서 $dist_p$ 가 최소인 점들을 매칭하였다. 이렇게 세워진 대응관계는 실루엣 이미지의 모든 윤곽선 정점들이 하나의 템플릿 모델 실루엣 상의 대응 정점을 갖도록 하였다. 더 나은 대응관계를 계산하기 위해 Kraevoy 등 [4]이 사용한 템플릿 모델과 입력 실루엣의 곡선 연속성을 고려할 수도 있지만, 본 연구에서는 [4]의 방법과는 다르게 외부의 윤곽선만을 이용하였으며, 템플릿 모델이 점진적으로 실루엣에 맞게 변형됨에 따라 대응관계를 다시 계산하기 때문에 빠른 계산이 중요한 본 연구의 특성상 정점의 연속성은 고려하지 않았다.

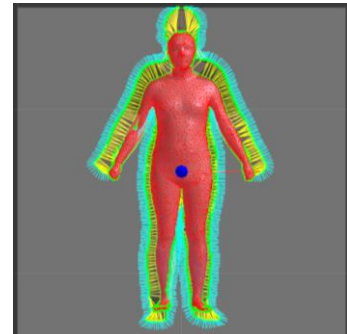


Figure 5. Initial Correspondence between input silhouette pixels and the vertices of the template model

5.3 PCA 인체 모델 기반 템플릿 변형을 통한 최적화

본 절에서 입력 실루엣과 매칭된 템플릿 모델 사이의 기하학적

오차를 최소화하는 템플릿 모델의 변형을 구하고 최종적으로 입력 실루엣으로부터 3차원 인체를 유추 및 재구성하는 방법에 대해 설명한다. 이 때 초기에 주어지는 템플릿 모델과 입력 그림은 비슷한 자세를 취하고 있더라도 전체적인 위치와 크기가 다르다. 따라서 PCA 모델만으로 최적화하는 것은 자연스럽지 못한 결과를 초래할 수 있다. 그러므로 이런 문제가 발생하지 않도록 템플릿 모델의 위치와 크기도 함께 고려하여 PCA 고유 벡터의 매개변수 값과 함께 최적화를 수행한다. 따라서 최적화를 수행하는 매개변수의 차원은 다음 표 1과 같다.

Table 1. Degree of freedom of the optimization parameters

매개변수	자유도
전역 크기(global scale)	1
전역 이동(global translation)	3
PCA 고유 벡터	30~200

최적화 문제는 표 1의 매개변수에 대해 실루엣 이미지와 템플릿 간의 거리에서 발생하는 오차 E_d 와 PCA 매개변수의 우도 (likelihood) 오차 E_l 를 고려한 오차 합 E 를 최소화하는 문제로 정의할 수 있다.

먼저 템플릿 모델 V 와 입력 이미지 실루엣 정점 집합 Q 와의 거리를 최소화하는 오차함수는 다음과 같다:

$$E_d = \sum_{i=1}^{N_d} (\text{deform}(\tilde{V}_i) - Q_i)^2, (N_d : Q \text{의 픽셀갯수}) \quad (5)$$

$$\text{deform}(\tilde{V}_i) = S \cdot (\tilde{V} + \sum_{j=1}^{N_e} w_j \cdot e_{j,i}) + T. \quad (6)$$

식 (5)는 $\text{deform}(\tilde{V}_i)$ 와 입력 사진의 윤곽선 정점과의 최소 거리를 계산하는 식이며, $\text{deform}(\tilde{V}_i)$ 는 식 (6)에서 볼 수 있듯 확대/축소 S , 이동 T 및 PCA 고유벡터에 의해 변형된 템플릿 모델을 의미한다. 식 (5)만을 이용하면 템플릿 모델 V 의 통계를 반영한 형태를 유지하기 힘들다. 즉 PCA의 고유 벡터의 중요도는 각각 다르며 각각의 고유 값의 가중치에 반비례하는 만큼의 우도 (likelihood)를 고려해줘야 한다(식 (7)).

$$E_l = \sum_{k=1}^{N_e} \left\| \frac{w_k}{\epsilon_k} \right\|^2, (N_e : \text{사용되는 주성분 갯수}) \quad (7)$$

w_k 는 각각의 k번째 고유 벡터의 가중치를 의미하며, ϵ_k 는 그 고유값을 의미한다. 최적화는 비선형 오차 함수이므로 그래디언트 값을 주어 반복적인 수치 최적화를 수행하였으며, 매 반복마다 모델과 그림사이의 대응 관계가 변할 수 있기 때문에 앞 절에서 설명한 입력 받은 실루엣과 템플릿 사이의 매칭 과정을 갱신

해줘 변화된 모델을 반영한 거리 계산이 가능하도록 하였다. 그림 6은 그림 2에 주어진 입력 이미지로부터 획득한 3차원 신체 예측 결과이다.

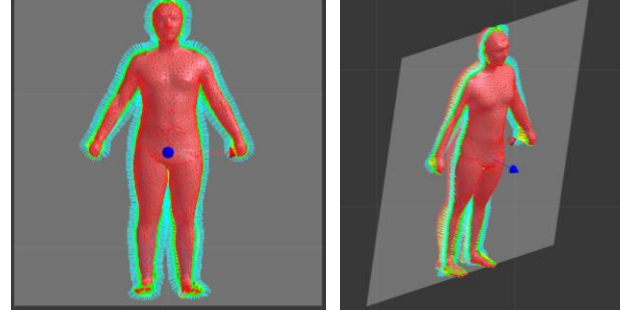


Figure 6. Resulting fitted template model to the input image given in Figure 2.

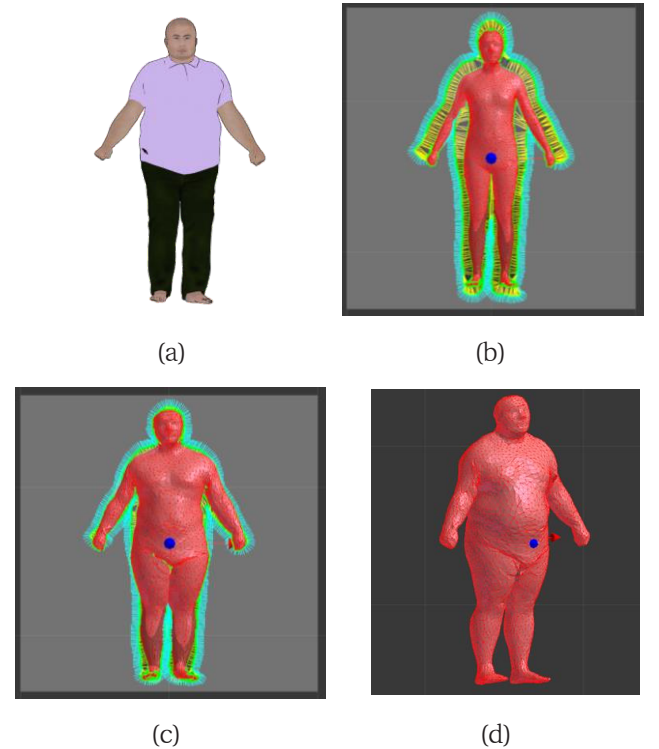


Figure 7. Experimental results for a different body type from the mean shape: (a) given input image; (b) the initial correspondences; (c) the resulting fitted template model; (d) the final result.

6. 실험 결과

본 장에서는 본 연구의 방법으로 수행된 결과를 보여준다. 제안된 방법의 강건성을 보여주기 위해 그림 2에서 주어진 입력 이미지에 대한 실험 외에도 그림 7에서 보이는 것과 같이 평균 모델과는 조금 더 차이나는 인체에 대해서도 실험을 수행하였으며 잘 동작하는 것을 확인하였다.

본 방법은 기본적으로 실루엣 이미지 한 장에 대해서 동작하지만 여러 장에 대해서도 큰 방법의 변화 없이 적용될 수 있다. 아래 그림 8(a)에서 보이는 것과 같이 두 장의 투영 평면이 직교하는 이미지(그림 8(b))로부터 초기 대응점을 설정하고(그림 8(c)), 이로부터 최적화 과정을 거쳐 최종 3차원 신체 변형 이미지를 구해낸다(그림 8(d)와 (e)).

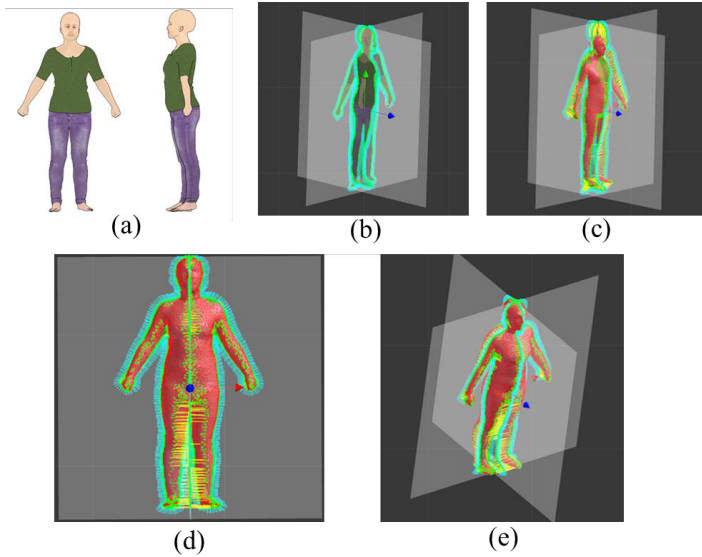


Figure 8. Experimental results for multiple input images: (a) given input two input image; (b) the images assume to have orthogonal projection planes to each other; (c) the initial correspondences; (d) and (e) the resulting fitted template model.

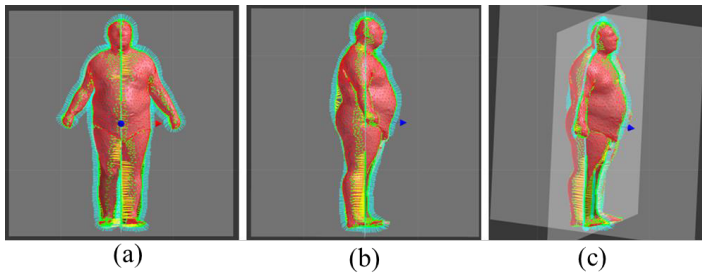


Figure 9. Experimental results for multiple input images for the input image shown in Figure 7: (a) the front view (a) and the side view (b) of the resulting model (c).

6.1 방법의 정량적 평가

본 논문에서 제안하는 방법을 정량적으로 평가하기 위해 기존에 존재하는 3차원 인체 모델(그림 10)의 실루엣을 입력 받아 본 연구의 방법으로 재구성한 3차원 인체 모델과 원본 모델의 정점 간의 유클리디언 거리(Euclidean distance)를 측정하였다. 두 모델의 정점 개수와 연결 정보가 다르기 때문에 정합된 모델의 각

정점에 대해 원본으로부터 가장 가까운 정점을 찾아 정확도를 측정했다. 그림 11은 결과 모델을, 그림 12는 오차의 특정 임계값 미만의 정점 비율과 각 정점에서의 원본과의 오차를 가시적으로 보여준다. (특히, 그림 12(a)는 특정 error 값에 해당되는 누적분포를 y축으로 표기한 것이다.)

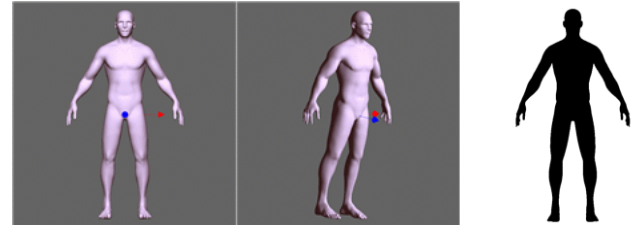


Figure 10. Reference 3D model for evaluating the fitting accuracy of our method. The right image shows the input for the method.

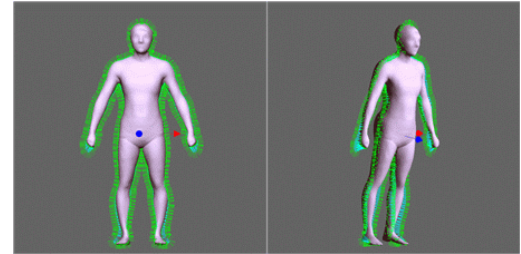


Figure 11. Fitting result for the reference 3D model from Figure 10.

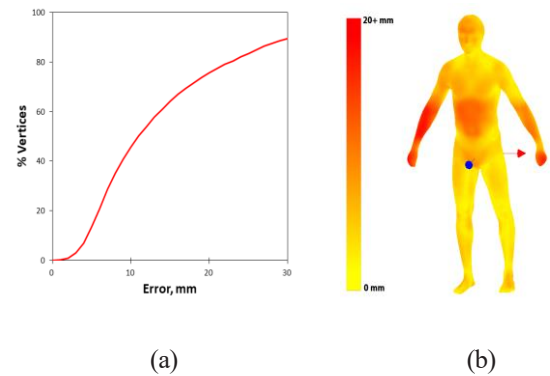


Figure 12. Fitting error for the reference model using our method; (a) the accumulated proportion [%] of vertices of which fitting error is below the threshold; (b) visualization of the fitting error on the surface.

그림 12(b)에서 볼 수 있듯 배 부분과 팔, 손 부분에서 오차가 크게 나타나는 것을 확인할 수 있다. 이는 템플릿 모델과 원본 모델이 서로 손모양이 다르고 원본 모델의 팔꿈치 아래의 팔 부분이 템플릿과는 다르게 조금 굽혀져 있기 때문이다. 템플릿 모델의 정합은 정면의 모습을 이용하였기 때문에, 이 오차는 템플릿과 원본의 팔의 z축 위치가 다르기 때문에 발생하는 오차이다. 또한 본 연구의 방법은 실루엣의 외형만을 이용하여 인체를 재구성

하기 때문에 실루엣 내부의 세부적인 신체 모습을 재구성할 수 없다. 따라서 본 연구의 방법으로는 원본 모델의 배 부분의 세부 모습을 복원할 수 없었고, 그로 인해 배 부분에서 오차가 나타남을 확인할 수 있다.

6.2 PCA 차원수와 최적화 횟수에 따른 결과 분석

사용한 PCA 차원수, 즉 최적화의 매개변수에 따른 결과모델의 정확도와 연산 시간이 달라지게 된다. 앞서 실험결과에서는 모두 공히 50개의 PCA 차원수를 이용한 것이다. 인체의 사실적인 재구성과 빠른 수행시간 사이의 균형을 이루기 위해 적절한 실험 환경을 조성하고자 최적화 매개변수의 자유도, 즉 PCA의 차원수에 따라 결과가 어떻게 변하는지 분석하였다.

그림 13는 PCA의 사용 차원수와 오차율에 대한 분석 결과를 보여준다. 이 그래프는 오차의 값에 따른 해당 점점들의 누적 분포를 보여주며 각기 색이 다른 곡선은 서로 다른 PCA 차원수를 의미한다. PCA의 차원수가 증가함에 따라 원본 모델과 정합 모델 사이의 큰 오차를 갖는 점점들의 분포는 점점 감소한다. 그러나 사실적인 결과를 위해 무작정 PCA 차원수를 늘리고 그에 따라 최적화를 많이 수행하는 것은 더 많은 수행시간을 필요로 한다. 표 2는 사용된 PCA 컴포넌트 개수에 따른 최적화 계산 소요 시간을 보여준다. 이 수치는 Intel i5-7200U 2.71GHz와 NVIDIA GeForce 940 MX를 장착한 개인용 노트북 컴퓨터에서 실험을 수행한 결과이다. 위 두 결과를 종합하여 인체의 경우 20개 정도의 PCA를 사용했을 때 충분한 정확도 및 수행속도 보여주는 것을 확인할 수 있다.

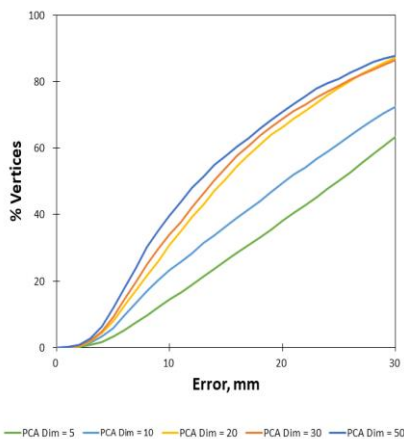


Figure 13. Relation between fitting error and the PCA components. The x-axis represents the error in mm and the y-axis is the percentage of the vertices with that error. Each colored curve represents the different numbers of PCA components use. As expected, the more components used result in the smaller error.

Table 2. Runtime according to the PCA dimension

PCA 차원수	수행 시간(초)
5	1.225
10	1.406
20	3.766
30	4.844
50	5.797
100	7.703

6.3 몸통과 팔의 각도가 45° 이상인 체형에 대한 결과

본 연구의 결과는 사용한 데이터베이스의 모델에 큰 영향을 받는다. 결과에 제일 영향을 주는 것으로는 MPII Human Spaces 데이터에서 강제한 특정 팔의 각도의 A-포즈를 취하고 있는 데이터이다. 이 특정 팔의 각도에서 벗어 날수록 재구성 에러는 증가하거나 아예 적용이 안될 수 있다. 이를 확인하기 위해 그림 14(a)와 같이 팔의 각도를 조금 더 T-포즈에 가깝게 변형한 입력 이미지를 넣을 경우 그림 14(b)(c)와 같은 결과를 얻을 수 있었다. 결론적으로 팔의 각도가 몸통으로부터 크게 벌어진 체형의 실루엣에 대한 3차원 정합은 어렵다고 할 수 있는데 이는 PCA가 갖고 있는 선형조합의 한계로부터 비롯된다. 즉 어깨의 회전은 대표적인 비선형적 변형이다.

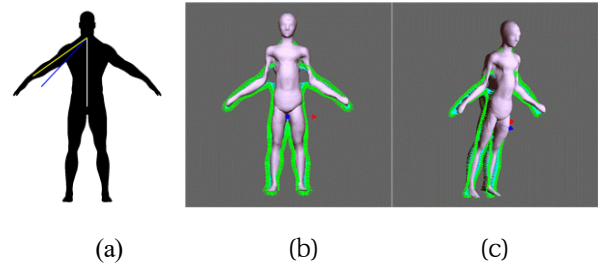


Figure 14. Input testing silhouette image with larger arm angle and its result: (a) input silhouette image: (b) and (c) erroneous fitting result.

7. 결론 및 향후 계획

본 연구는 특정 자세를 취하고 있는 인간형 캐릭터의 실루엣을 입력으로 받아 캐릭터의 3차원 외형을 유추하는 방법을 제안하고 있으며 이는 다음 세 가지 기술 요소를 포함한다.

- 1) 실루엣 분석을 통한 윤곽선 정점좌표 및 법선 벡터 추출
- 2) 3D 모델로부터 실루엣 추출 및 입력 실루엣과의 매칭
- 3) 통계적 인체 변형 모델 기반 템플릿 모델의 최적화

본 기술은 기존에 존재하는 인체 모델을 대상으로 하는 통계 모델[10]을 사용하기 때문에 장점과 단점이 극명하게 나타난다. 우선 PCA의 주성분을 이용해 인체를 표현하는 방식을 사용하기 때문에 보다 적은 변수로 다양하고 사실적인 인체를 효과적으로 표현할 수 있다는 장점이 있다. 그러나 본 연구에서 가정한 특정 자세(A-포즈)와 다른 자세, 가령 몸과 팔의 각도가 특정 각도 이상으로 벌어지면 정합이 제대로 이루어지지 않는다는 것 또한 확인하였다. 이러한 문제는 다양한 자세의 데이터를 추가함으로써 완화될 수 있겠지만, PCA라는 선형 방법의 한계로 인해 관절의 회전 등이 야기하는 비선형적 변형에 대해서는 재구성에 근본적인 한계가 있는 것도 사실이다. 따라서, 관절체(articulated body)에 대한 가정을 통한 내부 뼈대 및 스키닝 방법들을 활용한 자세에 강건한 재구성 법 등에 대한 추가 연구가 필요하다. 또한 최근 많이 연구되는 딥러닝을 통한 이미지 한 장에서의 깊이값 등의 3차원 정보 추출을 실루엣 또는 스케치 이미지로부터의 인체 재구성으로 확장 적용해 보는 것도 향후 연구 방향이 될 수 있다.

감사의 글

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 출연금 등으로 수행하고 있는 한국전자통신연구원의 문화기술(CT) 연구개발지원사업의 위탁연구과제 (R2018030391)를 통해 지원받았습니다.

References

- [1] B. ALLEN, B. CURLESS et al, The space of human body shapes: reconstruction and parameterization from range scans, SIGGRAPH 2003, pages 587-594
- [2] D. ANGUELOV, P. SRINIVASAN et al, SCAPE: shape completion and animation of people, SIGGRAPH 2005, pages 408-416
- [3] D. VLASIC, P. PEERS et al, Dynamic shape capture using multi-view photometric stereo, SIGGRAPH Asia 2009, Article No. 174
- [4] V. KREAVOY, A. SHEFFER, M. PANNE, Modeling from contour drawings, SBIM 2009, pages 37-44
- [5] A. JAIN, T. THORMHLEN et al, MovieReshape: tracking and reshaping of humans in videos, SIGGRAPH Asia 2010, Article No. 148
- [6] S. ZHOU, H. FU et al, Parametric reshaping of human bodies in images, SIGGRAPH 2010, Article No. 126
- [7] D. CREMERS, K. KOLEV, Multiview Stereo and Silhouette Consistency via Convex Functionals over Convex Domains, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol 33 Issue 6, pages 1161-1174, 2011
- [8] M. ZOLLHÖFER, M. NIEBNER et al, Real-time non-rigid reconstruction using an RGB-D camera, ACM Trans. on Graphics Vol. 33 Issue 4, Article No. 156, 2014
- [9] A. COLLET, M. CHUANG et al, High-quality streamable free-viewpoint video, ACM Trans. on Graphics Vol 34 Issue 4, Article No. 69, 2015
- [10] L. PISHCHULIN, S. WUHRER et al, Building statistical shape spaces for 3D human modeling, Pattern Recognition Vol 67, pages 276-286, 2017
- [11] K. M. ROBINETTE, H. DAANEN et al, The caesar project: a 3-D surface anthropometry survey, 3DIM'99 Proceedings of the 2nd international conference on 3-D digital imaging and modeling, pages 380-386, 1999
- [12] R. A. NEWCOMBE, D. FOX, S. M. SEITZ, DynamicFusion: Reconstruction and tracking of non-rigid scenes in real-time, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015
- [13] Desbrun, Anisotropic polygonal remeshing, ACM Transactions on Graphics, pp. 485-493, 2003.
- [14] G. VAROL, D. CEYLAN, B. RUSSEL, J. YANG, E. YUMER, I. LAPTEV, C. SCHMID, BodyNet: Volumetric Inference of 3D Human Body Shapes, Proceedings of the European Conference on Computer Vision (ECCV), pp. 20-36, 2018
- [15] Z. ZHENG, T. YU, Y. WEI, Q. DAI, Y. LIU, DeepHuman: 3D Human Reconstruction from a Single Image. The IEEE International Conference on Computer Vision (ICCV), 2019
- [16] S. SAITO, Z. HUANG, R. NATSUME, S. MORISHITA, A. KANAZAWA, H. LI, PIFu: Pixel-Aligned Implicit Function for High-Resolution Clothed Human Digitization, The IEEE International Conference on Computer Vision (ICCV), 2019
- [17] S. SAITO, T. SIMON, J. SARAGIH, H. JOO, PIFuHD: Multi-Level Pixel-Aligned Implicit Function for High-Resolution 3D Human Digitization, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020
- [18] 현대은, 윤승현, 김명수, 3차원 스캔 데이터로부터의 인체 팔, 다리 형상 복원, 컴퓨터그래픽스학회논문지, 8(4), 1-7, 2002
- [19] 최별이, 이성희, K-SMPL: 한국인 체형 데이터 기반의 매개화된 모델, 컴퓨터그래픽스학회논문지, 28(4), 1-11, 2022

〈 저 자 소 개 〉



안 다 술

- 2017년: 세종대학교 디지털콘텐츠학과 학사
- 2019년: 세종대학교 일반대학원
디지털콘텐츠학과 석사
- 2019년~현재: 벅슨게임즈 재직
- 관심분야: 컴퓨터 그래픽스, 컴퓨터 애니메이션,
기하처리 모델링 등
- <https://orcid.org/0000-0001-9741-6309>



박 상 일

- 1997년: 연세대학교 기계공학과 학사
- 1999년: KAIST 기계공학과 석사
- 2004년: KAIST 전산학과 박사
- 2005년~2007년: 미국 Carnegie Mellon
University 박사후 연구원
- 2007년~현재: 세종대학교 소프트웨어학과 교수
- 관심분야: 컴퓨터그래픽스, 캐릭터 애니메이션,
기하처리 및 모델링
- <https://orcid.org/0000-0003-1561-8148>