

스마트폰에서의 3 차원 인체 모델링

*한지수, *박인규

*인하대학교 정보통신공학과

{jshan0799@gmail.com, pik@inha.ac.kr}

3D Human Body Modeling on Smartphone

*Ji Soo Han, *In Kyu Park

* Department of Information and Communication Engineering, Inha University

요 약

본 논문에서는 스마트폰을 이용한 3 차원 인체 복원 시스템을 제안한다. 모바일에서의 인체 복원은 디바이스상의 한계로 인해 기기와 서버 간에 인체 모델링 과정을 분할하여 처리한다. 분할된 인체 모델링 과정은 복원 과정과 출력 과정으로 나뉘게 되며 안드로이드 기반의 스마트폰을 사용하여 취득한 영상을 서버로 전송하고 3 차원 인체 복원을 수행하게 된다. 3 차원 복원 알고리즘은 인체 변형 모델을 기반으로 3 차원 모델로부터 2 차원 영상으로의 투영을 통해 최적의 자세 및 체형 매개변수를 추정하고 복원된 결과는 모바일로 전송되어 복원된 모델링 결과를 출력한다.

1. 서론

3 차원 모델링 알고리즘은 2 차원 영상에 비해 필요한 데이터 처리량이 높아 모바일 기기에서 처리하는데 한계가 있다. 이러한 한계를 극복하기 위해 모바일기기와 서버로 인체 모델링 과정을 분할함으로써 처리 속도와 메모리 문제를 해결하는 것이 가능하다.

본 논문에서는 카메라가 탑재된 스마트폰을 이용하여 취득한 2 차원 인체 영상에 대해 3 차원 복원을 수행하고 모바일 기기에서 결과를 출력하는 시스템을 제안한다. 사전에 취득한 영상 및 카메라를 이용하여 실시간으로 촬영한 영상을 서버로 전송하여 복원을 수행하게 된다. 3 차원 인체 복원은 자세와 체형을 갖는 인체 데이터의 학습에 의해 만들어진 매개변수 기반의 인체 변형 모델을 이용하여 수행된다[1]. 모바일 기기로부터 전송된 영상에 대해 딥러닝을 이용하여 인체의 관절 위치와 윤곽 정보를 추정하여 자세 및 체형 정보를 취득한다. 측정된 인체 정보는 3 차원 변형모델로부터 2 차원 영상으로 투영된 모델과의 정합을 통해 최적의 매개변수 값을 찾게 된다[2]. 복원된 결과는 매개변수 혹은 오브젝트 형식으로 전송되며 모바일 상에서 구현된 3 차원 인체 모델을 변형한다. 이러한 접근은 디바이스의 한계로 인한 3 차원 인체 모델링의 한계를 극복하여 다양한 분야에서 활용 가능하도록 한다. 본 논문의 구성은 다음과 같다. 2 절에서는 모바일 기반의 3 차원 인체 모델링 방법에 대해 설명하고 3 절에서는 실험을 통해 제안한 시스템의 결과를 확인한다. 마지막으로 4 절에서는 본 논문에 대한 결론을 맺는다.

2. 모바일 기반의 3 차원 인체 모델링

모바일 기반의 3 차원 인체 모델링은 디바이스의 한계를 극복하기 위해 모바일 상의 입출력 과정과 서버 상의 3 차원 인체 복원 과정으로 분할 된다. 그림 1은 모바일 기반의 3 차원 인체 모델링의 전체 흐름도를 나타낸다.

2.1. 인체 모델 입출력

모바일 기기는 3 차원 인체 복원을 위한 입력 영상을 탑재된 카메라를 이용하여 쉽게 취득 가능하다. 안드로이드 기반의 어플리케이션은 액티비티 등의 전환이 일어날 때

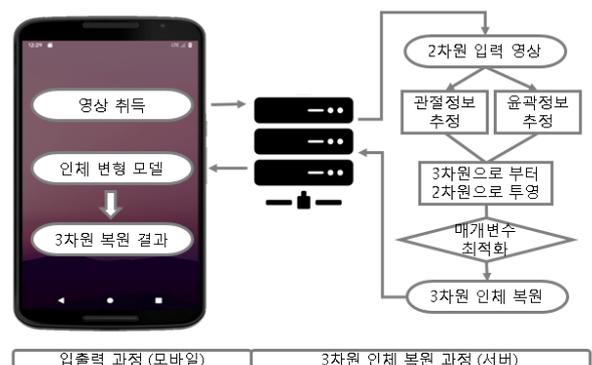


그림 1. 모바일 기반의 3 차원 인체 모델링의 전체 흐름도

호출이나 메시지를 전달하는 매개체인 Intent 를 사용하여 설정된 액션을 실행할 수 있다. 인체 모델링의 입력 값인 2 차원 영상은 접근 권한이 부여된 갤러리와 카메라를 통해 사전에 저장된 입력 영상을 불러오거나 실시간으로 영상을 취득하여 사용된다. 입력 영상에 대한 복원은 디바이스의 속도 및 메모리 한계로 인하여 데이터 통신을 통해 서버로 전송되어 처리된다. 서버는 리눅스와 윈도우 운영체제에서 apache 에 의해 구축되며 post 를 통해 데이터를 전송하고 get 을 통해 데이터를 전송받는 자바 스크립트를 url 형식으로 호출하여 통신한다. 3 차원 인체 모델은 매개변수 기반의 3 차원 인체 변형 모델로 6890 개의 정점과 23 개의 관절로 이루어진 SMPL 을 사용하며 OpenGLES 라이브러리를 이용하여 구현된다. 이러한 과정들은 셰이더라는 서브루틴을 이용하며 정점의 좌표를 기반으로 삼각형 집합을 취해 모델링 한다. 다양한 자세를 취하고 있는 데이터셋으로부터 회귀 모델을 통해 학습된 자세 매개변수와 다양한 형상의 데이터셋으로부터 principal component analysis(PCA)를 통해 학습된 체형 매개변수를 기반으로 한 정점 데이터를 메모리상에 불러오고 실행된 정점 셰이더 프로그램에 의해 가공된 데이터는 프레임 버퍼에 올라가 화면상에 나타나게 된다. 가공된 버퍼 데이터는 연결된 view 에 렌더링 되어 출력 된다. 서버에서 수행된 복원 결과는 매개변수 데이터를 전송받아 3 차원 모델을 변형시키거나 정점으로 이루어진 오브젝트 데이터를 전송받아 모델을 재출력 한다.

2.2. 3 차원 인체 복원

3 차원 인체 복원은 모바일로부터 전송받은 2 차원 입력 영상에 대하여 딥러닝을 적용하여 자세와 체형 정보를 추정한 결과를 이용하여 수행된다. 자세 복원은 관절 위치를 추정하고 3 차원 관절로부터 2 차원으로의 투영을 통해 위치의 거리를 최소화하여 최적의 자세 매개변수를 최적화하는 방법을 통해 이루어진다. 측정된 매개변수 값은 Rodrigues 공식에 의해 각 관절의 회전 반경을 결정하게 된다. 체형 복원은 2 차원 입력 영상에 대해 최신 개체 추정 기법인 Mask RCNN 기법과 GrabCut 기법을 활용하여 가공된 인체의 윤곽 정보와 3 차원 모델의 윤곽 정보를 정합하여 최적의 매개변수 값을 찾는다[3]. 두개의 마스크를 비교하기 위해서는 EMD(Earth Mover's Distance)공식을 사용하여 두 윤곽 데이터 사이의 최소 비용을 계산한다. 복원 기법에 의해 추정된 결과 값은 3 차원 인체 변형 모델의 자세와 체형 매개변수에 접근하여 변형된 오브젝트를 얻게 된다. 복원된 결과는 매개변수 값에 대한 데이터와 오브젝트 데이터로 이루어지며 이를 모바일로 재전송 하여 출력한다.

3. 실험 결과

본 논문은 Intel i7-7700 3.6GHz CPU 와 NVIDIA GeForce GTX 1080 Ti 를 장착한 컴퓨터와 쉘컴 스냅드래곤 835 가 장착된 갤럭시 노트 8 에서 수행하였다. 모바일로부터 서버로의 데이터 전송은 약 5 초 이내로 소요되었으며 서버에서의 3 차원 인체 복원은 약 2분 20초의 시간이 소요되었다. 서버에서 인체 데이터를 전송받아 모바일상에서 변형 시 평균 200 초(약 3 분 20 초)가 소요되었으며 오브젝트 데이터를 전송받아 렌더링 할 시 평균 150 초(2 분 30 초) 가 소요되었다.

그림 2 는 모바일상에서 구현된 3 차원 인체 모델을 나타낸다.

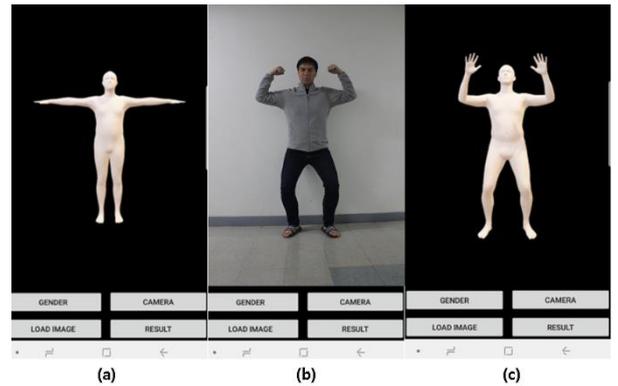


그림 2. 모바일상에 구현된 (a)기본 인체 모델, (b)입력 영상, (c) 3 차원 복원 결과

(a)는 매개변수 기반의 인체 변형 모델의 기본 자세를 나타내고 있으며 (b)의 입력영상에 대해 서버에서 수행된 인체 복원 결과에 대하여 (c)에서와 같이 매우 유사한 자세 및 체형의 결과를 확인할 수 있다.

4. 결론

본 논문에서는 기존의 3 차원 인체 복원 연구를 모바일상에 이식하여 실시간으로 취득한 입력 영상에 대하여 복원을 수행하는 기법을 제안한다. 데이터 통신을 통해 입력 영상을 서버로 전송하고 복원 과정을 수행하여 디바이스의 한계를 극복한다. 이러한 접근 방식은 모바일상에서의 3 차원 인체 모델 및 복원 결과에 대해 다양한 분야에 활용 가능성을 보인다.

감사의 글

본 논문은 CoAsia Holdings 의 지원을 받아 수행된 연구입니다.

참고문헌

- [1] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "SMPL: a skinned multi-person linear model," *ACM Trans. on Graphics*, vol. 34, no. 6, pp. 248, Nov, 2015.
- [2] F. Bogo, A. Kanazawa, C. Lassner, and P. Gehler, "Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image," *Proc. ECCV*, pp. 561-578, October 2016.
- [3] J. S. Han, M. R. Cho, and I. K. Park, "Moving human shape and pose reconstruction from video," *Proc. of Korean Society of Broadcast and Media Engineers Fall Conference*, pp. 66-68, November 2018.