ETRI

2012
Electronics and
Telecommunications
Trends

실사 인체 아바타 생성 및 활용 동향

Realistic Human Avatar Creation and Its Applications

김호원 (H.W. Kim) CG기반기술연구팀 선임연구원
김도형 (D.H. Kim) CG기반기술연구팀 선임연구원
이승욱 (S.W. Lee) CG기반기술연구팀 선임연구원
최진성 (J.S. Choi) CG기반기술연구팀 팀장
구본기 (B.-K. Koo) 영상콘텐츠연구부 부장

사용자 중심
차세대콘텐츠기술 특집

Ⅰ. 서론
Ⅱ. 실사 인체 아바타 생성 기술
Ⅲ. 실사 아바타 활용 동향
Ⅳ. 결론

아바타(avatar)는 온라인 게임이나 채팅, 가상 공간 등에서 사용자의 역할을 대신하는 애니메이션 캐릭터를 뜻하는 말로 1990년대 인터넷 보급과 더불어 활발히 도입되기 시작하였으며 2000년대에 실제 사람의 사실성보다는 3D 그래픽 캐릭터의 관점에서 세컨드라이프(Second Life)나 리니지 등의 서비스에서 활용되었다. 또한 컴퓨터 그래픽스 기술의 발달로 인해 보다 사실적인 인체 표현이 가능한 아바타 기술이 영화나 CF를 중심으로 활발히 활용이 되었다. 최근 컴퓨터 그래픽 기술과 컴퓨터 비전 기술의 결합으로 인한 시너지 효과 및 센서 기술과 컴퓨팅 하드웨어의 발전으로 인해 실제 사용자의 외형이나 움직임을 사실적으로 표현할 수 있는 실사 인체 아바타에 대한 연구와 산업으로의 응용이 활발히 진행되고 있다. 본고에서는 이러한 실사 인체 아바타의 생성 기술과 아바타를 응용한 활용 기술에 대해 살펴보고자 한다.

© 2012 한국전자통신연구원

Ⅰ. 서론

1990년대 인터넷의 보급 및 활성화에 따라 2000년대 세컨드라이프(Second Life)와 같은 가상 세계 서비스나 리니지와 같은 온라인 게임 혹은 온라인 채팅 서비스 등을 통해 온라인상에서 사용자의 역할을 대신하는 애니메이션 아바타에 대한 수요가 급증하였다. 이 시기에는 서비스 제공자가 이미 만들어 놓은 기성품 성격의 아바타를 사용자가 선택하거나, 사용자가 얼굴이나 헤어 등 미리 만들어진 파트들을 사용자 선택에 의해 조합하는 수준에서부터 사용자의 얼굴 사진을 아바타에 인위적으로 반영하는 정도로 (그림 1)과 같이 사실성과는 다소 거리가 있는 형태의 아바타 서비스가 진행되었다.

(그림 1) 인체 아바타(좌:Second Life, 우: 리니지)

반면, 제한된 컴퓨팅 리소스와 런타임 환경인 온라인상에서 서비스되는 비사실적 아바타들과는 달리 3D 애니메이션 등의 영화산업 분야에서는 보다 전문적이고 영화 장면에 특화되어 있긴 하지만 (그림 2)와 같이 실제 인체를 사실적으로 모방한 외형과 정교한 인체 형상변형에 대한 표현이 가능한 실사형 인체 아바타 혹은 디지털 액터가 "베오울프", "중천", "아바타" 등 많은 영화 제작에 활용되었다.

이러한 실사형 아바타는 컴퓨터 그래픽스 기술 및 컴퓨팅 파워의 발달과 레이저나 패턴광을 사용한 인체 형상 스캔 기술 및 광학식 마커 등을 사용하는 모션 캡처 기술의 발달, 그리고 오토데스크사의 "마야(Maya)"나 "맥스(Max)"와 같은 3D 저작툴의 발달을 기반으로 전문 디자이너들의 수작업 공정에 의해 주로 제작되었고 현재까지도 대부분 이러한 제작 파이프라인이 산업 현장에 활용되고 있다.

(그림 2) 실사 인체 아바타(베오울프)

하지만 대부분의 전통적인 제작 파이프라인들은 디자이너의 숙련도와 오랜 제작기간이 필요함과 동시에 사실적인 아바타 표현을 위해서는 많은 컴퓨팅 파워를 필요로 하기 때문에 영화나 CF 등과 같이 특수 목적에 국한되어 있어, 일반인을 대상으로 한 실사형 인체 아바타 서비스는 아직까지 현실적인 어려움이 있다.

2007년 아이폰을 필두로 한 스마트폰의 보급과 초고속 인터넷 확산, IPTV의 도입 등의 융합미디어 환경이 도래함에 따라 사용자들은 인터넷이나 PC를 통한 e-커머스(commerce)나 TV를 통한 t-커머스 등의 다양한 전자상거래를 통해 의류나 장신구들을 온라인상에서 구매하게 되었고, 실사 수준의 게임들을 일반 가정의 PC나 TV, 혹은 스마트 단말들을 통해 즐길 수 있게 되었다.

같은 시기에 사람의 눈의 기능에 해당하는 영상 센서들을 기반으로 물체의 형상을 복원하고 움직임을

추출하는 등의 컴퓨터 비전 분야의 영상 기반 복원 관련 기술들이 기존의 컴퓨터 그래픽스 기술과 유기적으로 결합함으로써, 인체 아바타 생성을 위해 기존의 액티브 방식의 형상 스캔이나 광학식 마커를 이용한 모션 캡처와 같은 고가 장비의 도움 없이도 다소 많은 시간과 컴퓨팅 파워를 요구하지만 사실적인 인체의 형상과 모션을 캡처할 수 있는 기술들이 활발히 연구되었다.

이러한 급변하는 기술 변화의 가운데 최근 2010년에 Xbox 360 게임기의 사용자 인터페이스로 등장한 키넥트(Kinect)라는 영상 센서는 기존 영상 센서들이 제공하는 컬러 영상과 더불어 인간의 시각기능 중 중요한 부분에 해당하는 영상 내 물체의 깊이 정보, 영상 내 인체의 움직임 정보를 별도의 컴퓨팅 파워의 도움 없이 실시간으로 제공해 주어 종래의 기술 접근들에 있어 많은 변화들을 촉발시켰다.

본고에서는 이러한 센서류의 발전과 인체 형상이나 모션의 처리 기술의 발전에 따라 실사형 인체 아바타의 생성이 영화와 같은 특정한 분야뿐만 아니라 일상생활에 적용될 수 있도록 하는 다양한 최근 기술동향과 이렇게 생성된 실사형 인체 아바타가 활용될 수 있는 응용 분야를 살펴보고자 한다.

II. 실사 인체 아바타 생성 기술

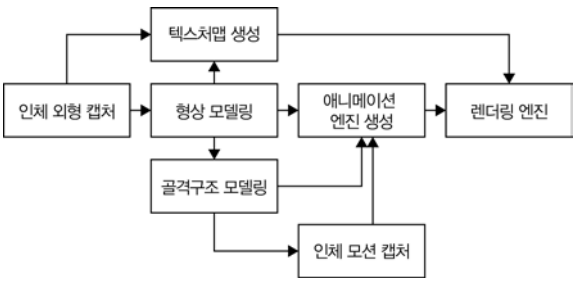
아바타(avatar)는 컴퓨터 사용자 스스로를 묘사하는 것으로, 컴퓨터 게임에서는 2차원 혹은 3차원 캐릭터의 형태로, 인터넷 포럼이나 커뮤니티에서는 2차원 아이콘으로, 초기 머드 게임에서는 문자열 구조로, 사용자가 스스로의 모습을 부여한 물체로 정의할 수 있다. 현재 대부분의 아바타는 사용자 자신을 사람이나 동물로 표현하고 2차원이나 3차원의 애니메이션이 가능한 캐릭터 형태로 활용되고 있다.

본 장에서는 사실적 인체 표현을 다루는 실사 인체 아바타의 생성에 관련된 최근의 기술동향을 살핀다.

1. 실사 인체 아바타 생성 개요

실사 인체 아바타는 실제 사람의 모습과 움직임을 3차원 애니메이션 기술을 활용해 컴퓨터상에서 마치 사람의 움직임을 그대로 촬영해 옮겨 놓은 듯이 사실적으로 표현할 수 있고 사용자가 원하는 대로 다양한 동작과 표정을 제어할 수 있는 3차원 애니메이션 캐릭터로 정의할 수 있다.

이를 위해서는 (그림 3)과 같이 실제 사람의 외형을 디지털화하는 3차원 형상 모델링과 사람의 외형을 촬영한 사진을 이용한 텍스처맵(texture map) 생성, 실제 사람의 골격(skeleton)을 근사화한 골격구조(skeletal structure) 모델링, 근사화된 골격구조의 움직임에 따라 모델링된 3차원 형상을 실제 사람의 움직임처럼 변형할 수 있는 애니메이션 엔진, 변형된 형상을 주변 조명환경에 맞게 사실적으로 렌더링하는 렌더링 엔진이 필요하다. 이상의 순차적 과정들이 실제 대부분의 실사 인체 아바타를 생성할 때 주로 사용되는 제작 파이프라인이며, 3차원 형상 스캔과 모션 캡처, 텍스처용 사진촬영 데이터를 기반으로 대부분 모델러나 애니메이터와 같은 전문가들의 수작업 공정에 의해 마야나 맥스와 같은 전문 3차원 저작툴을 이용해 제작된다.



(그림 3) 실사 인체 아바타 생성 파이프라인

2. 실사 인체 아바타 생성 기술동향

본 절에서는 실사 아바타 생성 파이프라인을 구성하는 각 요소 기술별로 기존의 접근 방식과 최근의 기술 트렌드들을 살핀다.

가. 인체 형상 캡처 기술

인체의 형상을 디지털화하는 기술은 크게 레이저나 패턴광을 이용한 액티브(active) 방식과 영상 카메라만을 이용하는 패시브(passive) 방식으로 나누어 볼 수 있다.

1) 액티브 방식 형상 캡처

액티브 방식은 레이저나 패턴광을 물체에 투사하고 물체의 표면에 투사된 패턴을 광원과 일정거리 떨어진 영상 카메라로 촬영하여 삼각측량 방식으로 물체의 3차원 위치를 결정함으로써 인체의 형상을 캡처하는 기술이다.

전신을 캡처할 수 있는 대표적인 제품으로는 (그림 4)와 같이 왼쪽부터 차례로 레이저 광을 사용하는 Cyberware사의 제품과 패턴광을 사용하는 Space-vision사의 제품 등이 있다. 그 외에도 최근 공항 검색대에 활용되어 이슈가 되고 있는 옷을 투시할 수 있는 밀리미터파를 사용하는 L-3 Communication사의 제품들이 있다.

이러한 제품들을 통해 전신의 형상 정보를 수많은 3차원 포인트들로 캡처할 수 있으며, 이들 포인트 정보들을 애니메이션에 용이하도록 구성된 삼각형 혹은 사각형 메시(mesh)구조로 근사화하는 작업을 통해 인체의 전신 형상을 모델링한다.

액티브 방식은 밀리미터 이하의 형상 복원 정밀도를 제공하는 많은 상용 제품들이 출시되고 있으며, 영화나 계측 산업 등의 광범위한 분야에서 활용되고 있다. 다만 인체의 전신 스캔에 적용될 경우, 최대 수십 초에서 수 초의 스캔 시간이 소요되며 스캐닝 시간 동안 인체의 미세한 움직임들이 스캔에 반영되는 문제와 고가의 시스템 비용이 소요되는 단점이 있다.

이러한 상황에서 2010년에 등장한 (그림 5)의 키넥트 센서는 액티브 방식의 방법에 있어 큰 변화점을 제시하고 있다. 기본적인 동작 원리는 패턴광 방식과 유사하지만 기존 센서들이 가시광 대역의 패턴을 사용하고 대부분의 스캐닝 시간이 초 단위 이상으로 소요되는 반면, 키넥트 센서는 적외선 대역의 특수 패턴 방식을 사용하여 비록 해상도는 640×480 정도로 낮지만 초당 30 프레임의 깊이 정보를 컬러 정보와 함께 제공함으로써 인간의 시각기능과 유사한 영상정보의 실시간 획득이 저렴한 센서로 가능해진 것이다. 기존의 컴퓨터 비전 분야의 큰 기술적 걸림돌 하나에 대해 컴퓨팅 파워를 거의 필요로 하지 않으며 실시간성을 보장하는 저렴한 상용 솔루션이 등장한 것이다.

이로 인해 기존의 관련 기술들이 키넥트라는 센서를 접목하기 위해 다양한 시도들을 진행 중이며, 최근 2012년 2월에는 마이크로소프트사에서 공식적으로 PC 환경에 대한 SDK(Software Development Kit) 지원을 시작하여 보다 광범위한 분야에 활용이 기대된다.

(a) Cyberware (b) Spacevision (c) L-3 Communication

(그림 4) 액티브 방식 전신 캡처 스캐너들

(그림 5) 키넥트 센서[1]

그 외에 밀리미터파를 이용하는 스캐닝 방식이 활발히 연구되고 있다. 이 방식은 옷을 입은 상황에서 인체의 내부 형상에 대한 스캐닝이 가능하기 때문에 최근 공항 검색대에 보안 목적으로 많이 활용되고 있으며, 가상 의복 체험과 같은 의복 피팅 서비스에서 사용자의 의복 탈의의 번거로움을 해소하면서 실측 기반의 서비스를 제공하기 위한 목적으로 활용이 시도되고 있다.

2) 패시브 방식 형상 캡처

패시브 방식은 복원 대상체를 둘러싸며 배치된 멀티카메라에서 얻어진 다시점 영상을 이용하여 컴퓨터 비전의 스테레오(stereo)나 체적(volumetric) 기반의 복원과 같은 영상 기반 복원 기법을 이용하여 대상체의 형상 정보를 복원하는 방식[2],[3]으로 고가의 액티브 방식 스캐너를 대체할 목적으로 주로 연구되고 있다.

패시브 방식은 액티브 방식과 달리 산업용 카메라나 DSLR(Digital Single Lens Reflex) 카메라와 같은 영상 센서들을 사용하기 때문에 인체의 전신에 대해 순간 촬영이 가능하여 (그림 6)과 같이 다양한 자세를 취하는 인체나 연속적인 움직임을 가지는 인체를 캡

(그림 6) MIT, 스탠포드 대학의 패시브 방식의 형상 캡처 기술[2],[3]

처할 수 있으며 액티브 방식에 비해 정밀도는 떨어지지만 보다 사실적인 외형을 복원할 수 있다.

키넥트의 등장으로 액티브 방식 대비 장점들이 다소 약해지기는 하였지만, 실측 인체 아바타 생성 관점에서는 DSLR 카메라와 같은 초고해상도의 영상정보의 접근 용이성과 순간 촬영의 장점, 저렴한 비용 때문에 아직 많은 연구들이 진행되고 있으며, 키넥트와 고해상도 영상 센서를 함께 사용하는 이기종 센서 간 융합을 통해 계산시간과 복원 정밀도를 동시에 해결하기 위한 접근들도 많이 시도되고 있다.

MIT와 스탠포드 대학[2],[3]의 형상 캡처 방식은 액티브 스캐너나 다중시점 스테레오 복원을 통해 생성된 액터의 템플릿 메시 모델을 사용하며, 다시점 영상의 전경 영역 정보와 전경 영역의 표면 컬러 정보를 이용하여 템플릿 모델을 피팅(fitting)하는 방식으로 다양한 자세의 인체 외형을 복원한다. 따라서 복원과정에 많은 컴퓨팅 파워와 시간이 소요되며, 정밀도가 액티브 방식들에 비해 떨어진다는 단점이 있다.

현재 ETRI에서 진행되고 있는 형상 캡처 방식은 액터를 촬영한 다시점 영상에서 영상 내의 액터 영역에 해당하는 전경 영역을 자동으로 검출하고, 추출된 다시점 영상의 전경 영역을 3차원 복셀(voxel) 공간에 역투영(back-projection)하여 액터의 3차원 외형을 체적 기반으로 복원한다. 이렇게 하여 얻어진 비주얼 헐(visual hull)이라는 복원 외형을 기반으로 다시점 영상의 컬러 정보를 추가적으로 이용하여 포토헐(photo hull)이라는 보다 실제 액터의 형상에 근접한 외형 정보를 복원한다[4].

이렇게 복원된 복셀 기반의 외형 정보에서 실제적인 인체 아바타의 외형 모델링을 위해 (그림 7)과 같은 외형 전이(transfer)라는 자동 변환과정을 수행한다. 외형 전이는 실제 콘텐츠 제작에 사용되는 간결한 메시구조를 가지는 표준 템플릿(template) 인체모델

(그림 7) 실사 인체 외형 모델링(ETRI)

을 이용하여 템플릿 모델의 메시 토폴로지(topology)를 유지하면서 메시를 구성하는 각 정점(vertex)들의 3차원 위치를 자동으로 변형하여 복원된 체적 기반의 복셀 형상을 템플릿 모델의 메시구조로 표현한다.

외형 전이 과정 이후, 촬영된 다시점 영상 정보를 이용하여 템플릿 모델의 텍스처맵 상의 각 픽셀에 해당하는 표면 컬러 정보를 전이된 메시 형상과 다시점 카메라 간의 가시성(visibility) 테스트를 통해 가시성을 가지는 다시점 영상에서 자동으로 추출하여 변경함으로써 실사 인체 아바타 외형 모델링을 수행하고 있다.

나. 인체 모션 캡처 기술

인체의 다양한 자세나 연속적인 모션을 캡처하는 기술은 인위적인 마커(marker)의 사용유무에 따라 마커 방식과 마커리스(markerless) 방식으로 나누어 볼 수 있다.

1) 마커 방식 모션 캡처

마커 방식 모션 캡처는 마커의 속성에 따라 자기(magnetic) 방식이나 광학(vision), 자이로(giro) 방식 등이 있으며, 이중 광학 방식이 가장 많이 사용되고 있다.

광학식 마커 방식은 카메라에서 보내는 적외선 광에 반사되는 마커들을 여러 대의 카메라가 삼각측량 방식으로 마커의 3차원 위치를 정하는 고전적인 패시브 방식에서 최근 마커에 직접 LED 광을 적용하여 마커의 LED가 발산하는 빛을 여러 대의 카메라가 촬영하는 액티브 방식으로 발전하고 있다.

기본적인 동작 원리는 마커 방식에서 주로 사용되는 적외선 광에 반사된 혹은 발광된 빛에 의해 영상에 투영된 마커를 영상처리를 통해 마커의 영상 내 위치를 얻어내고 카메라 보정정보를 이용한 영상 특징점들의 삼각측량으로 마커의 3차원 위치를 얻어낸다. 따라서 기존의 패시브 방식은 조명에 대한 제어가 가능한 비교적 어두운 실내에서만 적용이 가능하였고, 밝은 조명을 필요로 하는 영상 카메라와 함께 장면에 대한 촬영이 필요한 용도에는 활용이 어려웠다.

반면 액티브 마커 방식의 경우에는 각 마커가 고유한 데이터 값을 가지고 직접 발광하기 때문에 밝은 실내나 야외 환경에서도 마커의 인식률이 높으며 원거리에서도 안정적인 모션 캡처가 가능한 장점이 있다. 실제 이러한 액티브 방식을 이용하여 영화 아바타에서는 실내의 작은 공간에서만 활용 가능하던 광학식 모션 캡처 기술을 밝은 대규모 실내 스튜디오에 적용하였다. 최근 모션테크놀로지사의 상용 모션 캡처 시스템인 Xtive 시스템[5]은 쉽게 야외 환경에서 정밀한 모션 캡처를 지원하는 기능을 제공하고 있다.

그 외에 또 다른 시도로는 영화 캐리비안의 해적 촬영에 사용된 특수 패턴 방식을 들어 볼 수 있다. (그림 8)과 같이 AR(Augmented Reality) 기술에서 주로 활용되는 간단한 영상처리를 통해 쉽게 검출가능하고 각 특징에 대해 고유한 인덱스를 인식할 수 있는 특수 패턴을 액터의 몸에 부착한 후, 다수의 가시광 카메라로 이를 촬영하고 영상처리를 통해 패턴의 위치

(그림 8) 영화 캐리비안의 해적의 모션 캡처 시스템

와 정보를 인식하여 패턴의 3차원 위치를 추정하는 방식으로 실제 자연 조명의 야외 세트장 촬영에 활용하였으며, 촬영된 비디오 프레임에서 액터의 자세를 분석해 CG 캐릭터를 합성하여 장면을 제작하였다.

또한 최근에는 인체의 바디 모션 캡처뿐만 아니라 보다 사실적인 얼굴의 표정 모션 캡처에도 많은 연구가 진행되고 있다. 앞에서 설명한 패시브나 액티브 방식의 마커들은 실제 콩알 정도의 크기로 얼굴에 직접 장착할 경우 자연스러운 표정연기가 어려운 단점이 있어, 얼굴 표정 캡처를 위해서는 주로 패시브 마커를 소형화하여 활용하였다. 하지만 이 방법 역시 얼굴에 직접 마커들을 부착해야 하고 마커들 간의 간섭 때문에 자연스럽고 섬세한 표정을 캡처함에 있어 한계가 있다.

영화 아바타의 경우에는 얼굴의 디테일한 표정을 잡기 위해 모션 수트(suit)를 입은 배우의 얼굴에 (그림 9)와 같이 마커 대신 특수 페인팅을 피부 표면에

(그림 9) 영화 아바타의 이모션 캡처 시스템

직접 적용하고, 얼굴 전면에 부착된 카메라를 이용해 얼굴 특징들의 움직임을 감지하여 배우의 섬세한 감정까지도 동시에 캡처하는 이모션(emotion) 캡처 기술을 적용하였다.

2) 마커리스 방식 모션 캡처

주로 마커 방식이 적외선 조명과 적외선 필터를 적용한 카메라들을 이용하는 반면 마커리스 방식은 특별히 제작된 마커들을 사용하지 않기 때문에 주로 가시광 환경에서 일반적인 비디오카메라를 이용한다.

기본적인 동작 원리는 모션을 캡처하고자 하는 대상의 표면에 있는 자연 특징의 추적 및 영상처리를 통해 삼각측량 하거나, 대상체의 움직임에 따라 동기화되어 획득된 다시점 비디오 영상에서 프레임별로 복원된 인체 형상 정보에 대해 프레임 간 체적의 변화분을 검출하여 모션을 캡처하는 방식이다.

자연 특징을 직접적으로 사용하는 방식은 현재 학계를 중심으로 활발히 연구되고 있는 부분으로 아직 영화와 같은 콘텐츠 제작에 적용되기에는 보다 많은 기술개발이 필요해 보인다.

현재 ETRI에서 진행되고 있는 마커리스 모션 캡처 방식은 (그림 10)과 같이 연속된 다시점 비디오 프레임에서 프레임별로 다시점 전경 영상 정보를 이용해 비주얼헐 외형을 복원하고, 인접한 프레임 간에 발생하는 체적의 변화분을 표현할 수 있도록 인체모델의 해당 관절의 모션 값을 추정함으로써 액터의 모션 캡

(그림 10) 마커리스 모션 캡처(ETRI)

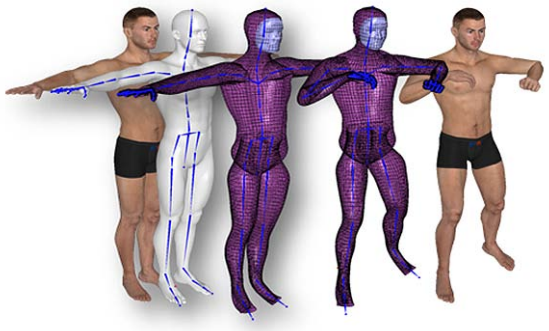
처를 수행하고 있다. 가로, 세로 각 10m의 공간에 배치된 20대의 HD(High Definition) 비디오 카메라를 이용하여 20 관절을 가지는 인체모델에 대해 초당 15 프레임 정도의 준실시간 모션 캡처가 가능하며, 계속적으로 다개체에 대한 적용과 속도향상에 대한 연구를 진행하고 있다.

최근에 주목할만한 부분은 앞에서도 설명한 키넥트에 의한 모션 캡처다. 마커를 사용하지 않는다는 점에서 (그림 5)의 키넥트를 이용한 모션 캡처도 마커리스의 범주에 해당할 것이다. 기술적인 관점뿐만 아니라 상업적으로 키넥트 센서가 주목을 받는 이유는 일반 조명환경에서 동작하며, 수십만 원대의 단일 센서에서 복수의 인체에 대해 인체별로 실시간의 모션 캡처를 컴퓨터의 리소스를 거의 점유하지 않고 지원하기 때문이다. 비록 센서의 목적이 게임 인터페이스용이라 기존의 고가의 상용 모션 캡처 장비에 비하여 정밀도는 떨어지지만, 가정의 일반적 조명환경에서 인체의 실시간 모션 정보를 활용할 수 있다는 장점 때문에 센서 출시 이후 사용자 친화적 인터페이스 분야에서 많은 연구들이 진행되고 있다.

다. 인체 애니메이션 엔진 기술

액터에 대한 외형 캡처를 통한 모델링 작업이 완료되면, 액터의 체형에 맞게 인체 골격을 근사화한 골격구조를 모델링하게 된다. 이때 골격구조의 근사화 과정은 사용되는 애니메이션 방법과 모델러나 애니메이터의 경험적 요소에 의해 다양한 변화가 가능하다.

골격구조 모델링이 완료되면 근사화된 골격구조에 맞게 모션을 캡처하게 되며, 캡처된 모션을 이용하여 골격구조의 움직임에 따라 모델링된 메시구조의 각 정점들의 위치를 변형시키기 위한 애니메이션 엔진을 생성하게 된다. 대부분은 인체 아바타의 애니메이션은 골격구조의 관절 움직임에 영향을 받는 정점들에 대해 마야와 같은 저작툴을 이용해 가중치(weight)를



(그림 11) 파라메트릭 애니메이션 엔진(ETRI)

부여하고, 관절의 회전 값에 가중치를 적용하여 정점의 위치들을 이동시키는 방식으로 인체의 형상변형을 표현하는 스키닝(skinning) 기법을 이용하고 있다. 얼굴의 경우에는 미리 정의된 다양한 표정의 블렌드(blend) 형상을 생성하고 무표정과 블렌드 형상 간의 보간(interpolation)을 통해 여러 표정을 생성하는 기법이 많이 이용되고 있다.

이러한 기법들은 수작업으로 디자이너가 원하는 방향에 맞게 형상변형을 생성할 수 있지만, 액터의 고유한 외형에 맞게 맞춤형으로 이러한 가중치나 블렌드 형상을 자동으로 생성하는 것은 기술적으로 해결해야 할 많은 이슈들이 있다.

현재 ETRI에서는 다양한 외형적 특성을 지니는 액터의 자연스러운 형상변형을 위해 (그림 11)과 같이 아바타의 골격구조와 결합된 넙스(nurbs) 곡선과 곡면의 제어에 의해 아바타의 각 정점들을 직접 제어하는 방식의 파라메트릭(parametric) 애니메이션 엔진을 개발하였으며, 이를 통해 팔/다리/몸통의 신체부위와 손가락, 얼굴 표정의 애니메이션 제어를 수행하고 있다.

Ⅲ. 실사 인체 아바타 활용 동향

본 장에서는 생성된 실사 인체 아바타의 활용이 가

능한 몇 가지 응용 서비스와 서비스를 위해 필요한 파일포맷 및 표준화 현황을 살핀다.

1. 실사 인체 아바타 응용 서비스

본 절에서는 가상의 캐릭터가 아닌 실제 사용자의 실사형 아바타를 이용하여 기존의 서비스들에 새로운 가치를 창조할 수 있는 계측, 방송 콘텐츠, 게임 콘텐츠 등의 응용 서비스들을 다룬다.

가. 실측 기반 의복 피팅 서비스

가상 의복 피팅 혹은 매직미러 서비스는 의류 시뮬레이션 기술 및 AR 기술을 이용하여 사용자에게 가상의 환경에서 실제 의복을 착용한 효과를 제공하여 자신에 맞는 의복의 사이즈나 디자인을 간접 체험하게 하는 응용 서비스이다.

국내에서는 건국대 의상학과를 중심으로 한 아이패션(i-fashion)[6] 사업이 있다. 먼저 액티브 형상 스캔을 이용해 사용자의 외형정보를 계측하고, 이를 기반으로 사용자 아바타를 생성한다. 생성된 아바타에 사용자가 선택한 의상을 착장시키고 의류 시뮬레이션을 통해 사용자의 체형에 대한 의상의 끼임 정도나 어울림 정도를 서비스하고 있으며, 실제 몇 곳의 의류 매장에 적용되어 있다. 이 서비스에서는 사용자가 자신의 아바타를 직접 제어하여 옷의 피팅 여부를 확인할 수 있는 기능은 제공되지 않고 있다.

Fitnect사는 (그림 12)과 같이 키넥트 센서를 이용

(그림 12) Fitnect사의 매직미러 서비스[7]

(그림 13) 물리 기반의 사실적 의류 시뮬레이션[8]

해 사용자가 직접 의류와 장신구들을 선택적으로 적용할 수 있는 인터페이스를 제공하고, 자신의 몸 동작에 따른 의상의 변형을 실시간으로 매직미러에 AR하여 제공하는 서비스를 개발하였다.

국내 기업인 클로버추얼패션사는 (그림 13)과 같이 물리 기반의 사실적인 의류 시뮬레이션 기술을 적용하여 여러겹의 복잡한 의상에 대해서도 사실적인 의복 피팅이 가능한 서비스를 제공하고 있으며, 현재 ETRI와 실사 아바타를 이용하여 실측 기반의 의류 피팅이 가능한 매직미러 서비스를 개발하고 있다.

나. 방송, 영화 콘텐츠 서비스

방송이나 영화 콘텐츠의 경우에는 이미 실사 인체 아바타의 활용이 활발히 이루어지고 있는 분야이다. 본고에서는 콘텐츠 제작을 위해 소요되는 비용과 시간의 관점에서 4D View Solution사[9]의 실사 아바타 서비스를 언급하고자 한다. 해당 아바타 생성 시스템은 주연급 캐릭터보다는 조연급 혹은 군중 캐릭터 정도로 사용이 가능하다. (그림 14)과 같이 TV 드라마

(그림 14) 실사 아바타를 이용한 군중 장면 연출[9]

"The Destiny of Rome"에 적용되어 대규모 군중 장면이나 전투 장면을 위해 크로마키 환경에서 동기화된 다시점 비디오 카메라로 액터의 움직임을 프레임별로 영상 기반 복원한 후 이를 다중 적용하여 군중 장면을 연출하였으며, 이를 위해 100만 유로 정도의 비용으로 50분 분량의 저예산 CG 드라마를 제작, 방송하였다.

다. 실사 게임 서비스

게임은 CG 응용에서 빠질 수 없는 분야로 스포츠 게임, 슈팅게임 등에서 게임의 몰입감과 사실감을 극대화하기 위해 (그림 15)의 예와 같이 실사 캐릭터가 많이 활용되고 있다.

2012년 출시 예정인 넷마블과 애니파크에서 개발한 "마구더리얼[10]" 게임은 한국 프로야구의 대부분의 선수들을 실사 기반으로 모델링하여 실사 아바타를 적용한 야구게임으로, 실제 게임에서 어떤 선수를 모델링했는지를 알 수 있을 정도로 그래픽 퀄리티가 뛰어나다는 평가를 받고있다.

2010년 Visual Concept에서 개발하고 2K Sprots에서 배포한 "NBA2K10"도 대표적인 실사 농구 게임이다. 디자이너에 의한 실사 같은 모델링과 다양한 모션 캡처를 통해 자연스럽고, 사실성과 현장감이 뛰어난 그래픽을 제공하고 있다. 또한 선수 각 개인의 문신을 아바타에 똑같이 새겨넣어 재미를 더해주었고, 얼굴표정을 캡처하여 실사감 및 재미를 더했다.

(a) 넷마블의 마구더리얼 (b) 2K Sport사의 NBA2K10

(그림 15) 실사 아바타가 적용된 게임서비스[7]

2. 파일포맷 및 표준화

실사 기반으로 사용자를 모델링하여 나온 결과물은 인체의 외형, 텍스처, 애니메이션 등의 다양한 정보를 포함하고 있다. 이러한 다양한 파일들을 전송 및 보관하기 위해서는 파일포맷의 표준이 필요하며, 본 절에서는 이를 위한 몇 가지 적용사례를 살핀다.

가. XML 기반의 Full 3D 복원 콘텐츠 메타데이터

2009년 ETRI는 "사용자 인터랙션을 가지는 Full 3D 콘텐츠를 저장하기 위한 메타데이터 및 파일포맷 구조"[11]라는 표준을 정보통신단체표준으로 채택하였다. 이는 복원된 full 3D 콘텐츠를 구성하는 다양한 엘리먼트를 정의할 수 있는 하나의 표준 XML(eXtensible MarkUp Language)을 만들어 이 형태로 파일을 저장, 전송할 수 있도록 도와주는 메타데이터 표준이다. 디자이너에 의해 만들어진 것이든, 스캐너를 이용하든, 모션 캡처 장비를 이용하든 어떠한 방식으로 만들어진 콘텐츠를 링크 형태로 설명하는 구조이다.

나. XML 기반의 COLLADA

COLLADA(COLLAborative Design Activity)[12]는 산업현장에서 만들어진 다양한 3D 모델 자원을 공용하기 위해 XML 기반의 텍스트로 만들어진 산업계 표준이다. Sony에서 플레이스테이션 게임을 만들 때 다양한 업계에서 게임 리소스를 만드는 것을 통일하기 위해 만들었으며, 2005년에 Khronos 그룹에 의해 최초로 표준화가 완성된 표준이다. 이 파일은 외형, 텍스처, 스키닝 기반 애니메이션 정보 외에, 키프레임(key frame) 애니메이션, 모핑(morphing) 등 다양한 애니메이션 기법도 포함한다.

COLLADA는 마야, 맥스 등 많은 산업 현장에서 사용되지만 텍스트 기반으로 이루어져 보안 및 파일 크기에 문제가 있는 관계로, 압축표준 전문그룹인

MPEG에서 XML 기반의 파일에 대한 MPEG 압축 표준툴을 지원하는 3DGCM(3D Graphics Compression Model)[13]을 개발하였고, 2011년 2nd edition을 공표하였다.

Ⅳ. 결론

지금까지 본고에서는 실사 인체 아바타를 생성하기 위한 최근의 기술동향과 이의 활용 동향에 대해 살펴보았다.

최근 컴퓨터 비전과 그래픽스 기술의 발전, 다양한 센서들의 발전을 기반으로 보다 사실적인 인체 아바타의 생성이 가능해지고 있다.

일반 사용자들이 실생활에서 본인의 실사 아바타를 생성하고, 이를 이용해 다양한 서비스를 접할 수 있기까지는 기술적 장벽이 존재하는 듯 하지만, 아이폰의 등장으로 시작된 스마트 모바일 시대의 조속한 도래처럼 스마트 단말과 무선 네트워크 속도의 발전은 아바타 생성 관련 기술들의 발전과 함께 머지 않은 미래에 이러한 실사 아바타의 활용이 현재 사진을 찍고 응용하듯이 일상생활으로의 활용을 가속화할 것으로 기대된다.

다만, Ⅲ장 2절에서 살펴 보았듯이 HDTV나 3DTV 등과 같은 표준포맷이 존재하지 않은 상황에서, 생성된 정보 간의 비호환성을 개선하고 상호 호환성을 향상시키기 위한 표준화 노력도 적극적으로 병행되어야 할 것으로 보인다.

용어해설

실사 인체 아바타 실제 사람의 모습과 움직임을 3차원 모델링 및 애니메이션 기술을 이용해 컴퓨터상에서 사실적으로 표현하고 제어할 수 있는 3차원 애니메이션 캐릭터

영상 센서 가시선이나 적외선의 입력에 의해 전달되는 영상정보를 전기신호로 변환해주는 촬상장치로 DSLR 카메라나 비디오 캠코더, 산업용카메라 등을 통칭하여 사용

약어 정리

3DGCM	3D Graphics Compression Model
AR	Augmented Reality
CG	Computer Graphics
COLLADA	COLLAborative Design Activity
DSLR	Digital Single Lens Reflex
ETRI	Electronics and Telecommunications Research Institute
HD	High Definition
SDK	Software Development Kit
XML	eXtensible MarkUp Language

참고문헌

[1] http://www.microsoft.com/en-us/kinectforwindows/

[2] D. Vlasic et al., “Articulated Mesh Animation from Multi-view Silhouettes,” *ACM Trans. Graphics*, vol. 27, no. 3, 2008, pp. 97:1~97:9.

[3] E. de Aguiar et al., “Performance Capture from Sparse Multi-view Video,” *ACM Trans. Graphics*, vol. 27, no. 3, 2008, pp. 98:1~98:10

[4] H. Kim and I. Kweon, “Appearance-Cloning: Photo-Consistent Scene Recovery from Multi-View Images,” *Int. J. Comput. Vision*, vol. 66, no. 2, 2006, pp. 163-192.

[5] http://www.motioncapture.co.kr/

[6] http://www.ifashion.or.kr

[7] http://www.fitnect.com/

[8] http://www.clo.co.kr

[9] http://www.4dviews.com

[10] http://ma9real.netmarble.net/

[11] 한국전자통신연구원, “사용자 인터랙션을 가지는 Full 3D 콘텐츠를 저장하기 위한 메타데이터 및 파일포맷 구조,” TTAK.KO-10.0318, 2009.

[12] COLLADA, Khronos Group. http://collada.org

[13] 3DGCM, “Information Technology — Coding of Audio-visual Objects — Part 25: 3D Graphics Compression Model,” ISO/IEC 14496-25, 2nd ed., 2011. http://www.iso.org