

디지털 휴먼의 제작 기술과 현황

□ 서영호 / 광운대학교

요약

디지털 휴먼(Digital Human)이라고도 알려진 가상 휴먼은 인간처럼 보이고 행동하는 컴퓨터 생성 3차원(3D) 인공지능(AI) 캐릭터를 말한다. 가상 세계에서 실제 사람을 대변하는 인공지능 인간도 가상 인간으로 볼 수 있다. 가상 인간은 1990년대 처음 등장했지만 기술력 부족으로 곧 자취를 감췄다. 1995년 세계 최초의 3D 모델링 가상 아이돌 다테 쿄코가 일본에서 데뷔하여 상당한 센세이션을 일으켰다. 그러나 그녀는 인간이 아닌 애니메이션 캐릭터에 가까웠기 때문에 인기를 유지하지 못하고 순식간에 사라졌다. 한편, 최근에는 인공지능(AI)이나 컴퓨터 그래픽(CG) 등 첨단 기술로 더욱 실감나는 가상 인간이 등장하고 있다. 가상 인간을 제작하는 방법은 다양하지만, 가장 일반적인 방법은 실제 사람을 촬영한 후 CG를 통해 얼굴을 3D 모델링하거나 AI를 사용하여 새 얼굴을 생성하여 얼굴을 변경하는 것이다. 기본 외모를 준비한 후 알고리즘을 이용해 데이터를 분석하는 비지도 머신러닝 방식인 딥러닝을 이용해 표정 등 다른 디테일을 더한다. 그러나 가상 인간의 전신을 3D 모델링하는 것은 막대한 비용과 작업 부하로 인해 아직 일반적이지 않다[1].

1. 가상 인간의 개요

최근 가상 인간이 확산하는 이유를 분석한 결과 2019년 코로나바이러스 감염증(COVID-19)의 발병으로 사람들은 집에 머무르는 시대에 접어들었고 온라인 생활이 보편화되었다. 나아가 인공지능이나 가상 현실에 관한 디지털 기술의 발전은 메타버스의 대두로 이어졌다. 메타버스 산업의 성장과 함께 사람들은 가상 인간에 더 많은 관심을 갖기 시작했고, 일부에서는 가상 인간을 다음 트렌드

로 간주한다. 또한 디지털 기기에 익숙하고 대부분의 시간을 소셜 미디어로 보내는 현대 가상 인간 확산에 기여한 최근 트렌드를 도출했다. 또한 광고에 가상 인간이 자주 등장하는 것도 한몫했다. 기업은 완전한 통제권을 가질 수 있기 때문에 가상 인간을 상업적 모델로 사용하는 것을 선호할 수 있다[1].

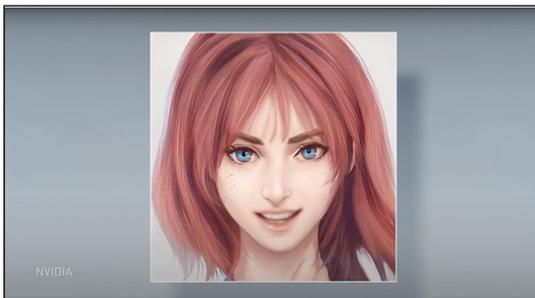
Nvidia에서는 디지털 휴먼 즉 가상 인간은 이미 미디어와 엔터테인먼트 분야에 있지만 디지털 휴먼에 대한 수요는 증가하고 있으며 산업 전반에 걸쳐 있다고 이야기하였



<그림 1> 다테 코코

다. 이에 대한 이점과 잠재적 사용 사례를 들 수 있다. 우선, AI 디지털 어시스턴트의 사용은 의료 및 소매와 같은 산업에서 큰 잠재력을 가지고 있다. 의료 전문가의 경우 디지털 어시스턴트가 교육 및 절차를 개선하는 데 도움이 될 수 있다. 의사는 실제와 같은 시뮬레이션으로 수술할 수 있으며 실제 수술을 수행하기 전에 최상의 결과를 얻기 위해 시뮬레이션을 수백 번 수행할 수 있다. 소매업에서 AI 디지털 지원은 보다 개인화된 경험을 제공하여 고객 서비스를 향상시킬 수 있다. 이것이 작동하려면 AI 디지털 어시스턴트가 구두 의사 소통에 대한 큰 이해력이 필요하다. 이것은 사람들이 필요한 작업을 수행할 수 있도록 디지털 어시스턴트와 더 잘 상호 작용하고 대화하도록 돕는 데 핵심이다.

건축 및 제조와 같은 산업에서 일하는 다른 회사의 경우 디지털 트윈은 팀이 공장과 도시에서 건물에 이르기까지 대규모 환경에서 작업자와 사람을 시뮬레이션하는 데 도움이 된다. 디지털 휴먼의 도움으로 기업은 정확한 시



<그림 2> 여성 AI 디지털 비서

물레이션을 통해 위험을 평가하고 환경을 예측하여 물리적 건물이 최적으로 설계되도록 할 수 있다. 디지털 휴먼의 지능 요구 사항은 AI 디지털 어시스턴트와 다르다. 디지털 휴먼을 가상 환경에 배치할 때 공장 노동자든 고층 빌딩을 걷는 관광객이든 사람처럼 탐색하고 행동하는 방법을 알아야 한다.

디지털 휴먼은 합성 데이터 생성을 개선하는 데 도움이 될 것이다. AI의 경우 데이터 및 훈련 신경망이 핵심이다. Synthesis AI, Microsoft, DataGen, Epic Games 및 Reallusion과 같은 회사는 이미 AI 모델을 교육하기 위해 3D 디지털 휴먼 데이터를 캡처하고 합성하는 작업을 하고 있다. 그러나 우리는 특히 미래를 위해 지속적으로 더 많은 데이터가 필요하다. 인공지능이 성장하기 위해서는 합성 데이터 생성이 핵심이고, 이를 확장하기 위해서는 디지털 휴먼의 합성 데이터 생성이 중요하다[2].



<그림 3> 공장 AI 에이전트

II. 가상 인간 제작 기술

디지털 휴먼 제작에는 볼류메트릭(Volumetric), CG (Computer Graphics), 모션 캡처(Motion Capture), AI(Artificial Intelligence) 등 다양한 기술이 사용된다. 볼류메트릭 방식은 실제 사람을 디지털 카메라로 360도 촬영하여 빠르게 3D 디지털 휴먼 이미지를 생성할 수 있다. 마이크로소프트(Microsoft)사의 볼류메트릭 스튜디오는 100대의 카메라를 이용한 초당 60프레임 촬영을 통해 고

품질의 3D 모델링이 가능하다.

CG를 활용한 디지털 휴먼 제작 방식은 3D 이미지를 만드는 렌더링(Rendering)에 상당한 시간과 비용이 소요되는 작업이나, 작업 중 결과물의 실시간 확인과 수정이 가능한 실시간 렌더링(Realtime Rendering)이 도입되면서 제작 효율성이 크게 개선되고 있다. 오프라인 렌더링은 1프레임(frame)의 게임 이미지를 생성하는 데 일반적으로 몇 분~몇 시간이 소요되나, 실시간 렌더링은 1초당 수십 프레임 이상의 이미지 생성이 가능하다. CG와 딥러닝(Deep Learning) 기술이 합쳐진 ‘뉴럴 렌더링(Neural Rendering)’도 디지털 휴먼 제작에 사용되고 있다. 뉴럴 렌더링은 2D 이미지를 학습하여 고품질의 3D 이미지 결과를 낼 수 있어 복잡한 모델링의 부담을 덜 수 있다[3]. 엔비디아(NVIDIA)에서 공개한 뉴럴 렌더링 모델인 ‘인스턴트 NeRF(Instant NeRF)’는 다양한 각도에서 촬영된 소수의 2D 인물 이미지만으로도 수 밀리초(1/1,000초) 안에 3D 인물 이미지 구현이 가능하다.

디지털 휴먼이 사람처럼 움직이는 모습을 구현하기 위해서는 ‘모션 캡처(Motion Capture)’ 기술이 사용된다. 이 기술은 센서/적외선 등을 사용하여 인체나 물체의 움직임을 추적해 디지털 형태로 기록한다. 마커(marker)를 통해 움직임을 추적하는 마커 방식, 센서를 이용한 마커리스(markerless) 방식, 카메라를 통해 획득한 이미지에 기반하여 AI로 움직임을 추정하는 이미지 기반(또는 비전 기반) 방식 등이 사용된다. 인간과 같은 생생한 감정 표현과 움직임 구현을 위해서 페이스 캡처(facial capture), 리깅(rigging), 애니메이션 작업 등이 더해지기도 한다. 자

이언트스텝(Giantstep)이 개발한 디지털 휴먼 ‘빈센트(Vincent)’는 리얼타임 렌더링, 모션 캡처, 머신러닝을 활용해 실제 사람의 표정 변화를 실시간으로 캡처하여 3D 인물 표정 변화로 구현한 바 있다.

광고나 영화 등에 쓰이는 세밀한 디지털 휴먼 구현은 아직 전문가의 영역이지만, 클라우드(Cloud) 기반의 디지털 휴먼 제작 플랫폼들이 등장하면서 일반인들도 디지털 휴먼을 제작하거나 활용할 수 있는 길이 열리고 있다. 언리얼 엔진(Unreal Engine)의 메타휴먼 크리에이터(Meta Human Creator)[14]가 제공하는 다양한 형태의 디지털 휴먼 프리셋(Preset)으로 수 분 내에 디지털 휴먼 제작이 가능하다. 국내에서는 클레온, 딥브레인AI, 마인즈랩 등이 디지털 휴먼 제작 플랫폼을 서비스하고 있다.

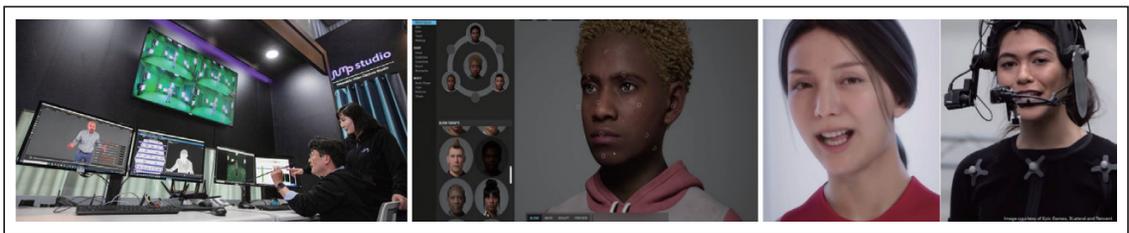
III. 가상 인간 제작 기술 및 절차 소개

1. 가상 인간 정의 및 디자인

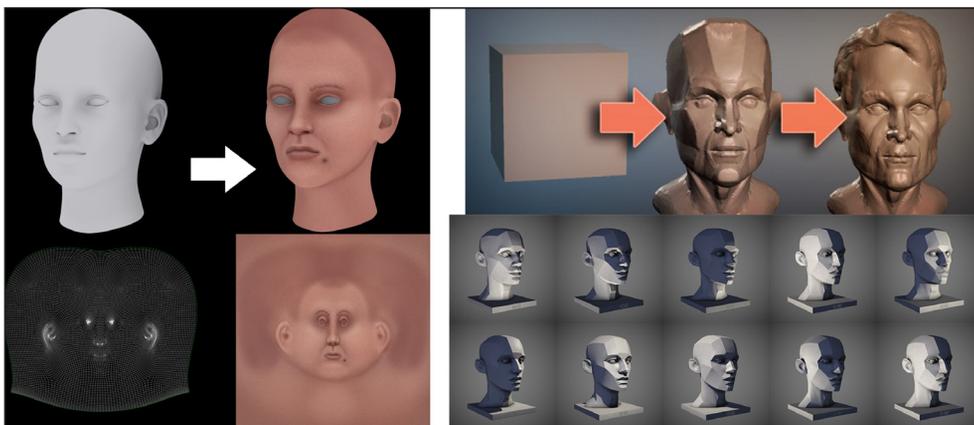
제작하고자 하는 가상 인간에 대한 전체적인 디자인 즉, 외형과 특징, 비율 등을 결정한다.

2. 모델링(Modeling)

3D 소프트웨어를 이용하여 작업을 하는 단계이다. 대표적인 소프트웨어로는 Blender, Maya, 3DS Max, ZBrush 등이 있다. 이 단계에서 디자인한 가상 인간을 생



<그림 4> 3차원 가상 인간 생성 (좌) SKT, (중) 언리얼 메타휴먼 크리에이터, (우) 3D Lateral



<그림 5> 가상 인간 제작 세부 기술 (좌) 텍스처, (우상) 모델링 및 스컬핑, (우하) 셰이딩

성하게 된다. 기초가 되는 베이스 메쉬를 생성하며 인간의 기본적인 형태를 만들게 된다. 스컬핑(Sculping) 작업도 진행하며 3D 모델에 세부적인 모양을 만든다. 피부 주름, 눈, 코, 입, 머리카락 등을 세부적으로 교정하고 디테일을 추가하는 것이 포함된다.

3. 텍스처 및 셰이딩(Texturing & Shading)

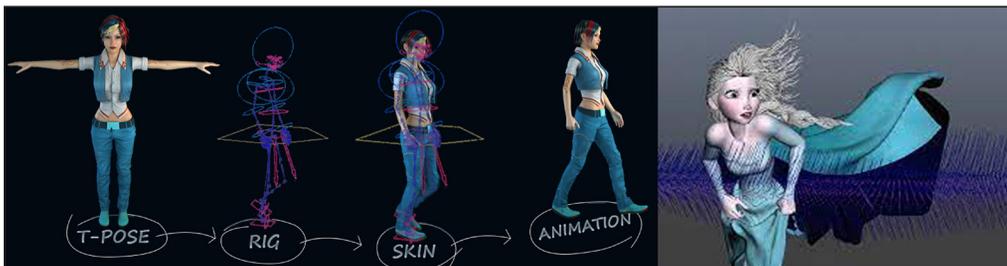
이 단계에서는 3D 모델에 색상과 텍스처를 추가한다. Substance Painter, Mari 등의 툴이 자주 사용된다. 텍스처는 3D 모델 UV에 맞추어 매핑되는 2D 이미지이다. 텍스처는 모델의 색상, 반사도, 질감 등을 정의한다. Shading은 이 텍스처와 함께 어떻게 빛을 받고 반사하는지 결정한다. 이를 통해 피부, 옷, 머리카락 등의 재질을 정의한다.

4. 리깅(Rigging)

Rigging은 모델에 뼈대를 만들고, 그 뼈대를 통해 모델이 움직일 수 있게 하는 과정이다. 뼈대를 통해 모델이 움직일 때 Skin Weighting 기술이 이용되며 뼈대 움직임에 맞추어 모델이 움직이는 정도를 결정한다. 이 과정에서는 디테일한 부분, 예를 들어 얼굴 표정이나 손가락 움직임 등을 정의한다.

5. 애니메이션(Animation)

Animation 단계에서는 이 뼈대를 이용하여 가상 인간이 움직이는 모션을 생성한다. 애니메이션에는 모델의 뼈대 움직임 뿐만 아니라 옷, 헤어 등에 대한 시뮬레이션도 진행한다. 머리카락이나 털은 매우 복잡한 구조를 가지고



<그림 6> 가상 인간 제작 세부 기술 (좌) 리깅, 스킨웨이팅 및 애니메이션 (우) 헤어 및 의상 시뮬레이션

있으므로, 이를 시뮬레이션하는 과정은 매우 세부적인 작업을 필요로 한다. Hair & fur simulation은 각각의 머리 카락이나 털을 정확하게 모델링하고, 그들이 움직이는 방법을 시뮬레이션한다. Cloth Simulation을 통해 옷은 자연스럽게 트레이핑되고 움직인다.

6. 렌더링(Rendering)

가상 인간을 최종적으로 시각화하는 과정이다. 이 과정에서는 3D 모델, 텍스처, 빛, 그림자 등을 모두 합쳐서 최종 이미지를 만든다. Displacement & bump mapping 기술은 3D 모델의 표면에 더 많은 디테일을 추가하는 데 사용된다. Displacement mapping은 실제로 3D 모델의 표면을 변형시키고, Bump mapping은 빛과 그림자를 사용하여 표면에 디테일을 추가하는 일종의 환상을 만든다. Subsurface Scattering 기술은 피부나 왁스와 같이 빛이 약간 투과하는 물질을 더 실재감 있게 렌더링하는 기술이다.

IV. 리타게팅, 모션 캡처 및 애니메이션

1. 가상 인간과 애니메이션

가상 인간 애니메이션은 캐릭터의 얼굴 표정, 몸의 움직임,

근육의 움직임 등을 섬세하게 표현하여 인간의 특징을 현실적이고 자연스럽게 재현해 낼 수 있도록 하는 데 중요한 역할을 한다. 가상 인간의 움직임은 모션 캡처 기술이나 3D 애니메이션 소프트웨어 등 다양한 기술과 기법을 활용하여 구현된다. 가상 인간 애니메이션은 인간과 거의 동일한 수준의 동작과 표정을 갖추는 것을 목표로 하며 이를 위해 다양한 연구 및 개발이 지속적으로 이루어지고 있다.

2. 모션 캡처 기술

가상 인간 애니메이션의 가장 대표적인 기술인 모션 캡처는 몸에 센서를 부착하거나 카메라 또는 다른 장치를 사용하여 사람의 동작을 실시간으로 tracking하고, 데이터로 기록해 가상 캐릭터에 반영 및 애니메이션하는 데 사용되는 기술이다. 기존에는 animator가 프레임마다 캐릭터의 움직임을 조정하는 등 수작업을 필요로 했지만, 모션 캡처를 사용하면 사람의 실제 동작을 기록하여 자연스럽게 보면서 섬세한 움직임 데이터를 자동으로 얻을 수 있다. 모션 캡처는 크게 광학식 모션 캡처와 비광학식 모션 캡처 방식으로 나눌 수 있다. 광학식 모션 캡처는 카메라와 마커를 사용하여 동작을 캡처하는 방식이다. 캡처 대상인 사람이나 물체에 마커를 부착하고 최소 두 개 이상의 카메라가 이 마커를 인식하도록 설정되며 각 카메라는 동일한 지점에서 캡처 대상을 찍도록 조정된다. 이후 캡처된 이미지에서 마커의 위치를 인



<그림 7> 물탱크에서 수중 모션 캡처를 진행하는 '아바타2' 촬영 장면



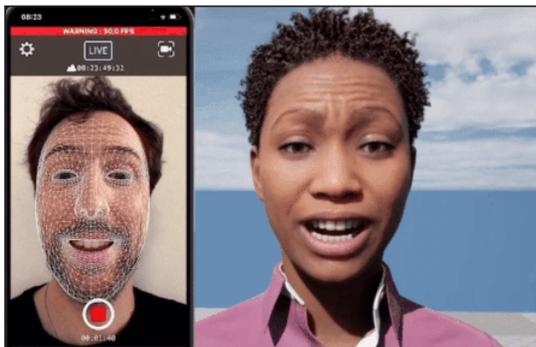
<그림 8> 반사를 막기 위해 물탱크의 수면 위를 흰색 공들로 덮고 촬영한 장면

식하고 삼각측량법(triangulation)을 사용하여 대상의 삼차원 좌표를 계산한다[3]. 비광학식 모션 캡처는 가속도 센서, 자이로센서 등의 센서를 사용하여 신체의 관절 및 주요 부위에 부착된 전용 슈트를 통하여 캡처 대상의 움직임, 회전, 방향 등 동작을 캡처하는 방식이다. 광학식과 달리 마커를 사용하지 않으며, 사용자의 움직임을 센서로 감지하여 추적한다. 광학식 모션 캡처에는 VICON, Optitrack 등이 있으며, 광학식 장비는 사람이나 물체 무엇이든 캡처가 가능하며 다수의 인원으로 대규모 작업을 할 수 있으며 3D 공간에서 좋은 퀄리티를 가진다는 장점이 있는 반면 초기 구축 비용 및 운영비 등 높은 비용을 요구하며 카메라에 의존도가 높기 때문에 충분한 조명이 없거나 반사되는 물체가 많은 실외 환경에서는 정확한 마커 인식이 어렵기 때문에 촬영 장소에 제한을 받는다는 단점이 있다[4]. 비광학식 모션 캡처 장비에는 대표적으로 Xsense, Rokoko, Perception Neuron이 있으며, 이러한 비광학식 장비는 광학식에 비해서 상대적으로 비용이 낮으며 광학식과 달리 카메라에 의존하지 않기 때문에 움직일 수 있는 공간만 있으면 야외와 같은 촬영 장소에 제약이 없다는 장점이 있다. 하지만 광학식과 비교했을 때, 광학식 모션 캡처는 카메라와 마커를 사용하여 고정밀한 위치 추적과 정확한 삼차원 좌표 계산이 가능한 반면 비광학식 모션 캡처는 센서를 통해 움직임을 감지하고 처리하기 때문에 정밀도가 떨어질 수 있다는 단점이 있다. 또한 움직이는 거리가 길어질수록 오차가 쌓이며 공간 구성이 잘 안 되는 문제가 있다[5]. 모션 캡처 기술은 영화 산업에서 많이

사용되고 있는 기술 중 하나인데, 실제 배우의 동작을 외계인, 동물 등 특수한 캐릭터의 움직임으로 재현하는 데 유용하다. 광학식 모션 캡처의 가장 대표적인 예로는 영화 ‘아바타(Avatar)’가 있다. 2022년 12월 14일자로 개봉한 영화 ‘아바타2’는 수중 모션 캡처에 도전했다는 점에서 기존 모션 캡처와는 차이점을 가진다. 그동안 물을 배경으로 한 영화는 대부분 건식 촬영 기법(dry for wet)으로 제작되었다. 하지만 영화 ‘아바타2’는 수중에서 배우들이 고정된 카메라에서 적외선을 반사하는 마커가 있는 슈트를 입었으며 촬영할 때 배우의 모든 복잡한 동작을 포착하기 위해 무대 주변에 120대의 카메라를 설치하였다[6]7.

3. 가상 인간 얼굴 애니메이션

얼굴에 움직임을 구현하는 방식에는 얼굴 캡처(Facial Capture)와 얼굴 리깅(Facial Rigging)이 있다. 얼굴 캡처는 얼굴 표정과 움직임을 실시간으로 추적하여 가상 캐릭터에 적용하는 기술이다. 가성비 좋은 아이폰과 같은 기기를 사용하여 Live Link(Unreal Engine) 또는 Face Capture(Unity) 등의 Apple 앱을 통해 얼굴 표정을 실시간으로 추적하고 Unreal Engine이나 Unity와 같은 특정 엔진을 통해 데이터를 가상 캐릭터에 쉽게 적용할 수 있다. 얼굴 리깅은 3D 캐릭터 모델의 얼굴을 구성하는 다양한 부위(눈, 입, 미간 등)에 대해 움직일 수 있도록 애니메이션 뼈대를 만들어 표정 및 얼굴 움직임의 다양한 변화를



<그림 9> Unreal Engine의 Live Link를 이용한 얼굴 애니메이션



<그림 10> Unity의 Face capture를 이용한 얼굴 애니메이션



<그림 11> Capture의 예시 화면

제어할 수 있도록 해 주는 것이다.

일반 영상 카메라를 이용한 모션 캡처 방식은 고가의 장비와 센서를 필요로 하지 않는다는 측면에서 비용 절감을 할 수 있으며, 그러한 특수한 센서를 부착한 수트나 장비를 입을 필요가 없어 사용자들에게 편리함을 제공할 수 있다는 점에서 유용하다. 하지만 모션의 정확한 추적이 어려울 수 있으며, 빠른 속도나 복잡한 움직임의 경우 제한이 있을 수 있다는 단점이 있다. 일반 영상 카메라를 이용한 모션 캡처 소프트웨어로는 Capture가 있다. Capture는 카메라를 방에 배치하고 소프트웨어를 설치하면 카메라에서 데이터를 캡처하고 움직임을 쉽게 추적할 수 있다. 또한 동영상상을 기반으로 3D 모션 데이터를 추출하는 Move.AI 등도 있다[8][9].

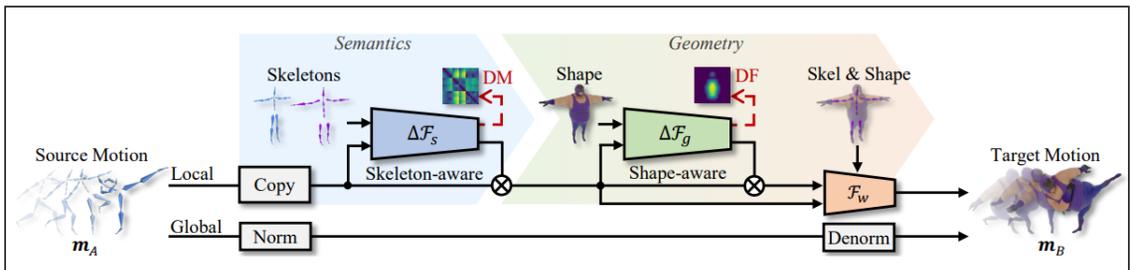
리타게팅은 한 캐릭터의 동작을 다른 캐릭터에 적용하여 동일한 움직임을 구현하는 기술이다. 각 캐릭터는 캐릭터마다 고유한 뼈대 구조를 가지기 때문에 bone의 길이와 개수가 다를 수 있다. 그렇기 때문에 동일한 애니메이션

션을 만들더라도 캐릭터 구조에 맞게 각각 새롭게 제작해 주어야 해서 많은 시간과 노력을 필요로 한다. 하지만 리타게팅 기술을 사용하면, 기존의 동작 데이터를 재사용할 수 있어 새로운 동작을 생성하는 데 드는 시간과 비용을 절약해 작업을 효율적으로 처리할 수 있다. 리타게팅 방식의 캐릭터 애니메이션에는 Unreal, Unity 엔진을 이용한 모션 리타게팅이 있다. Unreal에서는 IK Solver를 이용하여 source와 target 간에 bone chain을 생성하여 생성된 chain을 매핑하는 방식으로 source의 동작을 target에 리타게팅할 수 있으며, Unity에서는 Animator 컴포넌트와 Mecanim 시스템을 사용하여 모션 리타게팅을 구현할 수 있다. 또한 딥러닝 기반의 모션 리타게팅 방식도 있는데, 이 방식은 인공지능을 사용하여 source 모션과 target 캐릭터 사이의 관계를 학습하고 source의 모션을 target 캐릭터의 모션으로 변환한다. Skinned Motion Retargeting with Residual Perception of Motion Semantics & Geometry 논문에서는 새로운 Residual RETargeting 네트워크(R2ET) 구조를 제안하여 source 동작을 target 골격과 모양에 점진적으로 맞추도록 조정한다[11].

V. 가상 인간의 후보정 및 렌더링

1. 후보정(포스트 프로덕션)의 소개

포스트 프로덕션 또는 포스트 프로세싱은 녹음 및 녹화,



<그림 12> 딥러닝 기반의 모션 리타게팅 논문에서 제안하는 네트워크 구조

영화, 비디오, 디지털 아트의 제작 과정 중 하나를 가리키는 말로, 실제 촬영이 모두 끝난 뒤에 이루어지는 생산 작업을 통틀어 말하는 용어이다. 특히 가상 인간 제작에서의 **포스트 프로덕션**이란 초기 제작 단계에서 생성한 가상 인간을 보다 현실적으로 만들기 위해 음향 효과, 모션 편집, 텍스처 및 조명 조정을 하는 작업이다. 일반적으로 가상 인간의 포스트 프로덕션은 캡처 또는 생성, 모션 편집, 얼굴 애니메이션 편집, 텍스처 및 조명 조정, 음향 효과 추가, 시각 효과 추가의 단계로 이루어진다. 먼저 캡처/생성 단계에서는 실제 인간의 동작, 얼굴 움직임, 음성을 캡처하거나 생성한다. 이후 모션 편집 단계에서 모션 데이터를 편집하여 원하는 동작을 만들어 낸다. 동작의 흐름, 속도, 자세 등을 조정하거나 다양한 동작을 결합하는 등의 작업을 거쳐 자연스러운 움직임을 만들어 낸다. 얼굴 애니메이션 편집 단계에서는 얼굴 표정, 입 모양, 눈 깜빡임을 조정하여 가상 인간의 얼굴 표현을 개선한다. 이 단계에서는 표정 블렌딩 기술을 사용하여 다양한 표정을 만들 수 있다. 텍스처 및 조명 조정 단계에서는 가상 인간의 피부, 옷감, 머리카락 등의 텍스처를 조정하고 조명 설정을 통해 실제 환경에 더 잘 어울리도록 만든다. 이 단계에서는 광원의 위치, 강도, 색상 등을 조정하여 가상 인간이 실제 환경에서 자연스럽게 보이도록 한다. 이후 음향 효과 추가 단계에서 가상 인간의 음성을 녹음하거나 음성 합성 기술을 사용하여 목소리를 만들고, 이에 적합한 음향 효과를 추가한다. 이 단계에서는 음성 녹음, 음성 합성, 음향 편집 등을 수행하여 가상 인간의 목소리를 현실적으로 만

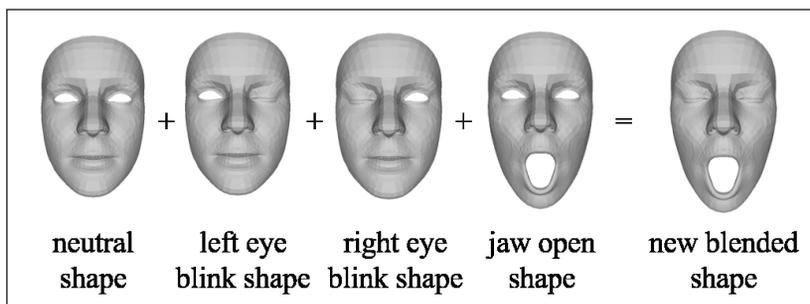
든다. 마지막으로 시각적인 효과를 추가하여 가상 인간을 더 생동감 있게 만든다. 입자 효과, 렌더링, 컬러 그레이딩 등을 통해 화면에 흥미로운 시각적인 효과를 부여하고, 최종 영상을 완성한다[15].

2. 모션 편집

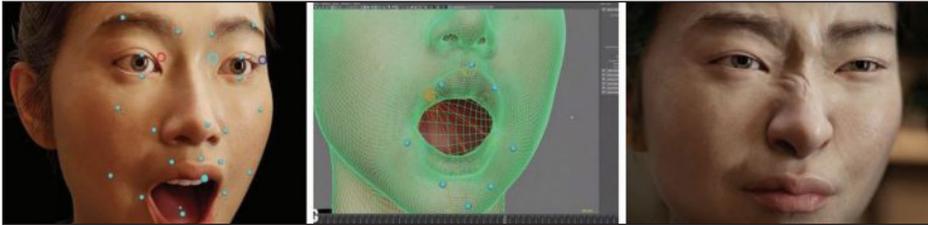
모션 편집은 가상 인간의 동작을 조정하고 편집하는 과정으로, 동작 조정, 동작 결합, 동작 수정 과정으로 세분화된다. 동작 조정은 캡처한 동작이나 생성된 모션 데이터를 시각적 또는 공간적으로 조정하여 원하는 동작을 만드는 작업으로, 예를 들어 동작의 흐름을 부드럽게 만들거나 더 빠르게 하거나, 동작의 크기를 조절하는 등의 작업이 포함된다. 동작 결합은 여러 동작을 결합하여 복잡한 동작을 만들어 내는 작업으로, 예를 들어 걷기와 뛰기를 조합하여 달리기 동작을 만들거나, 다양한 손 동작을 결합하여 제스처를 생성한다. 동작 수정은 캡처한 동작이나 생성된 모션 데이터를 조작하여 보다 자연스러운 움직임을 만드는 작업이다. 너무 강한 움직임을 완화하거나, 관절 각도를 조정하여 자연스러운 자세를 만들 수 있다.

3. 얼굴 애니메이션 편집

얼굴 애니메이션 편집은 가상 인간의 얼굴에 표정과 움직임을 만드는 과정으로, 가상 인간의 감정을 표현하고 시청자의 몰입감을 높일 수 있다. 얼굴 애니메이션 편집은



<그림 13> blendshape 방식의 얼굴 애니메이션



<그림 14> 크래프트톤이 공개한 얼굴 리깅 방식의 얼굴 애니메이션(좌: [17], 중: [18], 우: [19])



<그림 15> 크래프트톤이 공개한 얼굴 리깅 작업이 적용된 가상 인간

표정 제어, 입 모양 조정, 눈 깜빡임으로 세분화될 수 있다. 표정 제어는 가상 인간의 표정을 조정하는 기술로 주로 표정 블렌딩(blending) 기술과 리깅 기술이 사용된다. 표정 블렌딩은 다양한 표정의 형태(shape)를 만들어 두고, 이들을 섞어(blend) 조합된 표정을 만들어 낸다. 예를 들어 왼쪽 눈 감기, 입술 오므리기 표정을 혼합하여 씩크하며 뽀뽀하는 표정을 생성할 수 있다. 입 모양 조정 기술은 가상 인간이 말하는 동안 입술의 움직임을 자연스럽게 만드는 작업이다. 발화 속도, 강세, 발음 등을 고려하여 입 모양을 조절한다. 눈 깜빡임 작업은 가상 인간의 눈 깜빡임을 제어하여 더 현실적인 표현을 구현한다[16].

페이셜 리깅 기술은 얼굴을 여러 부분으로 나누고, 각 부분에 적절한 가중치를 매핑하여 자연스러운 얼굴 근육

과 관절의 움직임을 만드는 방식으로, 다양한 표정의 구현이 가능하다. 뼈대(bones), 제어자(controllers), 혹은 머슬(muscles)과 같은 제어 시스템을 통해 얼굴의 다양한 부위의 제어가 가능하다. 이 제어 시스템을 조작하여 표정과 표현을 구성할 수 있다.

4. 텍스처 후보정

텍스처 후보정은 텍스처의 색상, 명암, 촉감, 패턴 등을 보정하며, 특히 가상 인간 피부의 주름, 주근깨, 점 등의 세부 사항 추가에 이용된다. 일반적으로 텍스처 보정 작업은 수동으로 디자이너에 의해 영상 편집 소프트웨어로 진행되며, 주로 Adobe Photoshop, Substance Painter, Mari,



<그림 16> Unreal Engine 메타휴먼 크리에이터의 주근깨 추가

PixPlant, GIMP 소프트웨어들이 주로 사용된다. 주근깨를 추가하는 작업을 손쉽게 구현하도록 Unreal에서는 메타휴먼 크리에이터의 기능으로 지원하고 있다.

컴퓨터 비전과 인공지능의 발전으로 자동화된 텍스처 후보정 기술도 연구되고 있으며 대표적으로 NeuTex 논문이 있다. 해당 논문에서는 3D 모델의 표면에 2D 이미지를 매핑하는 texture mapping 네트워크를 학습하고, 기하학 성분(x,y,z)과 텍스처를 분리함으로써, 사용자가 2D 텍스처 맵을 수동으로 편집 가능하게 하는 편의성을 제공한다[20].

5. 조명 조정

조명 조정은 가상 환경에서 가상 인간이 더욱 현실적으로 보이도록 조명 설정을 조절하는 과정이다. 이 과정은 광원의 위치, 강도, 색상 등을 조절하는 작업으로 세분화할 수 있다. 광원의 위치 설정 과정은 광원의 위치를 조절하여 가상 인간에게 올바른 하이라이트와 그림자를 적용하는 과정으로, 가상 인간이 실제 환경에 존재하는 것처럼 보이게 한다. 광원 강도 조절은 강도를 높여 밝은 조명 효과나 강도를 낮춰 낮은 조명 효과를 얻을 수 있다. 광원의 강도를 조절하여 가상 인간의 피부, 옷감 등에 적절한 반사를 생성하고, 장면의 분위기와 조명 조건에 맞게 조명을 조정한다. 조명 색상 조정을 통해 가상 인간과 장면에 원하는 분위기와 톤을 부여한다. 예를 들어 따뜻한 빛을 사용하여 온화하고 따뜻한 분위기를 연출하거나, 차가운 빛



<그림 17> SMCU(SM 엔터테인먼트의 Culture Universe)

을 사용하여 시원하고 냉정한 분위기를 연출할 수 있다.

VI. 가상 인간 제작 관련 기업 현황

1. 스튜디오 클론

포스트 프로덕션을 수행하는 기업에는 ‘스튜디오 클론’이 있으며, 최근 SM에서 인수하였다. 해당 기업은 영상 편집, Mixing, DI(Digital Intermediate), OAP(On Air Production) 등의 윈스톱 솔루션을 확립한 대표 콘텐츠사로 평가받고 있다. 한편 SM은 지난해 7월 설립한 스튜디오 광야를 기반으로 VR 콘서트 플랫폼 기업 어메이즈VR과의 합작법인 ‘스튜디오 A’ 구축과 함께 시각특수효과(VFX) 제작, 가상 인간, 포스트 프로덕션, 가상 현실 제작 등 차세대 콘텐츠 IP 역량 강화에도 박차를 가하고 있다[21].

2. 오모션

가상 인간의 모션 캡처 개발 기업에는 ‘오모션’이 있으며, 볼류메트릭 캡처 방식으로 가상 인간을 생성 후, 비전 기반의 모션 캡처 기술을 통해 일반인도 손쉽게 가상 인간을 생성 가능한 솔루션을 제공한다[22].



<그림 18> 오모션의 가상 인간



<그림 19> 딥브레인AI의 가상 인간[24]

3. 딥브레인AI

한국의 스타트업인 딥브레인AI는 인공지능(AI) 기술을 통해 AI 휴먼을 개발하고 공급하고 있다. AI 휴먼은 기존의 가상 인간과는 달리 실제 사람과 매우 유사하게 만들어진다. 딥브레인AI는 입 모양을 실시간으로 구현하는 어려움을 극복하기 위해 ‘실시간 합성’ 기술을 개발하여 가상 인간의 품질을 높였다. 이 기술은 사람의 음성에 따라 입 모양을 자연스럽게 변화시키는 것으로, 높은 화질과 해상도로 입 모양을 구현하는 것을 가능하게 한다. 이러한 실시간 합성 기술은 금융 및 컨택센터와 같은 분야에서 활용될 수 있으며, 딥브레인AI는 이를 위한 서비스를 준비 중이다. 이 회사는 실시간 합성 기술을 갖춘 가상 인간 제작 기업으로서 특별한 위치에 있다고 평가받고 있다. 현재 한국의 제1금융권, 유통사, 방송사 등에 이 기술을 공급했으며, 미국과 중국을 포함한 다른 국가들에게도 제공하고 있다[23].

VII. 결론

본고에서는 기술의 발전과 함께 다양한 분야와의 협업 및 융합을 통해 많은 사회적 관심을 끌고 있는 가상 인간에 대하여 필요성 및 제작 기술을 중심으로 탐색하고 분석하였다. 가상 인간의 제작은 대화, 학습, 경험 등 현실

과 상호 작용하여 새로운 서비스와 인터페이스를 창출하는 데 중요한 역할을 하며 실제 사람과 유사한 모습과 움직임을 갖는 것을 목표로 한다. 이를 위해서는 정교한 3D 모델링 기술과 얼굴 표정 및 몸짓의 자연스러운 애니메이션 기술이 필수적인 요구 사항이다. 가상 인간은 다양한 산업 분야에서 많은 필요성과 활용 가능성을 가지고 있는 만큼 향상된 기술을 이용하여 사회, 의료, 교육 등 여러 분야에 적용되어 긍정적인 변화 및 발전을 이룰 것으로 기대된다. 이에 본고에서는 가상 인간 제작 분야의 현재 기술을 파악하고 가상 인간이 실제 어떻게 활용되고 있는지 적용 사례 등을 조사하여 현재 동향 및 그 기여도를 살펴보고자 하였다.

스컬핑 및 덤핑 등의 기술을 이용하여 가상 인간의 형태와 세부 사항을 더욱 정교하게 구현할 수 있고 리타게팅 및 모션 캡처 등의 애니메이션 기술의 도입으로 현실적이고 자연스러운 움직임이 구현되었으며, 후보정 기술을 통해 시각적 품질이 향상되었다. 이에 따라 앞으로의 가상 인간은 현실과 거의 구별할 수 없을 정도의 더욱 정교하고 자연스러운 형태와 움직임, 사실적인 시각적 품질로 사람들과 상호 작용할 때보다 현실적인 경험을 제공할 수 있을 것이다. 따라서 이를 위해 지속적인 연구가 필요하며 새로운 방법론과 기술의 개발을 통해 효율적이고 혁신적인 가상 인간의 생성과 응용이 추진될 것으로 예상된다.

참 고 문 헌

- [1] 강민주, Virtual Humans Are Coming!, <http://skt.skku.edu/news/articleView.html?idxno=1316>
- [2] Simon Yuen, An Era of Digital Humans: Pushing the Envelope of Photorealistic Digital Character Creation, <https://developer.nvidia.com/blog/an-era-of-digital-humans-pushing-the-envelope-of-photorealistic-digital-character-creation/>
- [3] CVPR 2020: Tutorial on Neural Rendering, https://juststhies.github.io/posts/neuralrenderingtutorial_cvpr/
- [4] UEAZ, Motion Capture (모션 캡처) 02. 모캡장비를알아보자(관성식 #1), <https://www.youtube.com/watch?v=D9hhTgn7Z2g>
- [5] UEAZ, Motion Capture (모션 캡처) 03. 모캡장비를알아보자(관성식 #1), <https://www.youtube.com/watch?v=zbzw-y2Y27Q&t=2s>
- [6] 주간동아, 실제보다 더 실제 같은 ‘아바타2’의 수중 장면 촬영 비밀은?, <https://weekly.donga.com/List/3/08/11/3845140/1>
- [7] goodword, 아바타2 제작 기술, CG, 물의 길 놀라운 비주얼 이면에 숨겨진 기술 탐색, <https://pixelpicture.tistory.com/7541>
- [8] UEAZ, Motion Capture (모션 캡처) 04. 모캡장비를알아보자(딤러닝), https://www.youtube.com/watch?v=w45_l9nZJu8&t=243s
- [9] Theodore McKenzie, Move AI: A New Software for Markerless Motion Capture, <https://80.lv/articles/move-ai-a-new-software-for-markerless-motion-capture/>
- [10] 아이뉴스24, “가상인간이 모델로”...슈퍼랩스, 뷰티 브랜드 어뮤즈 앰버서더 제작, <https://news.nate.com/view/20230502n17719?mid=n0105>
- [11] Skinned Motion Retargeting with Residual Perception of Motion Semantics & Geometry(CVPR 2023), <https://arxiv.org/abs/2303.08658>
- [12] 가상인간 제작 기술 참고문헌
- [13] Handbook of Virtual Humans, Nadia Magnenat-Thalmann, Daniel Thalmann 편집
- [14] 메타 휴먼 크리에이터, <https://metahuman.unrealengine.com/>
- [15] NeuTex 논문 “NeuTex: Neural Texture Mapping for Volumetric Neural Rendering”: https://openaccess.thecvf.com/content/CVPR2021/papers/Xiang_NeuTex_Neural_Texture_Mapping_for_Volumetric_Neural_Rendering_CVPR_2021_paper.pdf
- [16] blendshape 방식의 얼굴 애니메이션 출처: me
- [17] 크래프톤의 얼굴 리깅 그림1 포함된 기사 url: <https://cm.asia.co.kr/article/2022033109014408969>
- [18] 크래프톤의 얼굴 리깅 그림2 포함된 기사 url: <https://www.newspim.com/news/view/20220415000531>
- [19] 크래프톤의 얼굴 리깅 그림3 포함된 기사 url: <http://mbiz.heraldcorp.com/view.php?ud=20220208000920&a=99>
- [20] Unreal의 주근깨 변경 url: <https://docs.metahuman.unrealengine.com/ko/metahuman-creator-face-controls/>
- [21] sm엔터의 포스트 프로덕션 기업 인수 기사 url: <https://www.etnews.com/20230426000113>
- [22] 가상인간 모션 캡처 개발 기업 오모션 뉴스 영상 url: https://www.youtube.com/watch?v=9EV4C_gALeA
- [23] 딥브레인 AI 기사 url: <https://www.newstheai.com/news/articleView.html?idxno=3974>
- [24] 딥브레인 AI의 가상 인간 그림 url: <https://www.aitimes.com/news/articleView.html?idxno=150502>

저 자 소 개



서영호

- 1999년 : 광운대학교 전자재료공학과 학사
- 2001년 : 광운대학교 전자재료공학과 석사
- 2004년 : 광운대학교 전자재료공학과 박사
- 2004년 ~ 2005년 : 한국전기연구원 전기정보망그룹 연구원
- 2005년 ~ 2007년 : 한성대학교 정보통신공학과 교수
- 2008년 ~ 현재 : 광운대학교 전자재료공학과 교수
- 2010년 ~ 현재 : 광운대학교 실감미디어연구센터 센터장
- 2021년 ~ 현재 : 오모션 주식회사 CTO
- 주관심분야 : 영상처리, 3D 그래픽스, 디지털홀로그램, 광신호처리