

디지털 휴먼 기술 동향과 문화체육관광분야 활용 전망

황재인
(한국과학기술연구원
AI·로봇연구소
책임연구원)

CONTENTS

- I. 디지털 휴먼 연구 정의 및 특징
- II. 디지털 휴먼 연구 및 기술 동향
- III. 디지털 휴먼의 활용 분야
- IV. 결론

문화정보 이슈리포트
2021-7호(제27호)

디지털 휴먼 기술 동향과 문화체육관광분야 활용 전망

황재인(한국과학기술연구원, 책임연구원)

요약

최근에 컴퓨터 그래픽스, 컴퓨터 비전, 인공 지능 기술의 발전과 더불어 디지털 휴먼 기술에 많은 관심이 집중되고 있다. 이러한 디지털 휴먼 기술은 모델링 및 상호작용 기술, 감정 생성 기술들이 다양하게 적용이 된다. 디지털 휴먼을 모델링 하는 방법은 볼류메트릭 캡처, 단일 이미지 기반, 영상 합성 기반 방식 등이 있다. 그리고 디지털 휴먼 상호작용은 대화 생성, 제스처 생성, 행동 생성, 환경과의 반응 등을 설계하는 것들이 있다. 또한 디지털 휴먼의 감정 생성 기술은 상대에 맞는 표정이나 대화 등의 창작에 필요하다.

디지털 휴먼 기술은 다양한 분야에서 사용되고 있다. 특히나 문화 콘텐츠 및 실감 커뮤니케이션 방면에서 적극적으로 활용되고 있는데 버추얼 인플루언서나 원격 회의 등의 분야에서 사용되고 있다.

또한 메타버스에서 디지털 휴먼의 중요도는 더욱 높아질 것으로 예상된다. 메타버스에서 가상 인간 혹은 실제 사람들 대변하는 디지털 더블의 기반 기술로 디지털 휴먼 기술이 적극적으로 활용될 것으로 예상된다. 그리고 인공지능 기술의 발전과 동시에 디지털 휴먼은 외형 뿐만 아니라 지능적인 측면에서 인간과 유사하게 될 것으로 예상되며 이는 사람과 같은 실재감을 주는 디지털 휴먼이 될 것으로 기대된다.

미래에는 디지털 휴먼이 동반자가 되는 버추얼 컴패니언으로 활용이 될 것으로 예상되며 인간의 디지털적인 연속체인 메타 휴먼의 기반 기술로 사용될 것으로 생각된다.

※ '문화정보 이슈리포트'의 내용은 작성자의 의견으로 한국문화정보원의 공식적인 입장과 다를 수 있습니다.

I. 디지털 휴먼 연구 정의 및 특징

1. 디지털 휴먼 정의

- 디지털 휴먼(Digital Humans)은 사실과 같은 수준의 인간 3차원 모델에 인간 수준의 지능까지 포함한 형태를 궁극적인 목표로 하고 있음, 이 카테고리에서 버추얼 휴먼(Virtua Humans)과 디지털 더블(Digital Doubles)로 나눌 수가 있음
 - 버추얼 휴먼은 일반적으로 가상으로 창조해낸 디지털 휴먼으로 실제 존재하는 사람이 아닌 창조된 자기 자신의 정체성을 보유하고 있음
 - 디지털 더블은 실제 사람을 사실적으로 복제해서 사용하는 경우로 외형이나 행동 패턴 등을 복제해서 사용함, 이는 실제 존재하는 사람을 최대한 비슷하게 모사하는 것을 목표로 하고 있음
 - 디지털 휴먼은 외형적인 형상, 행동, 언어, 지능까지도 함께 구현되어야 함, 궁극적으로는 사람과 구별이 되지 않을 정도의 외모의 자연스러움과 상호작용 능력을 보유하는 것을 목표로 하고 있음



〈그림1〉 디지털 휴먼, 버추얼 휴먼, 디지털 더블의 분류

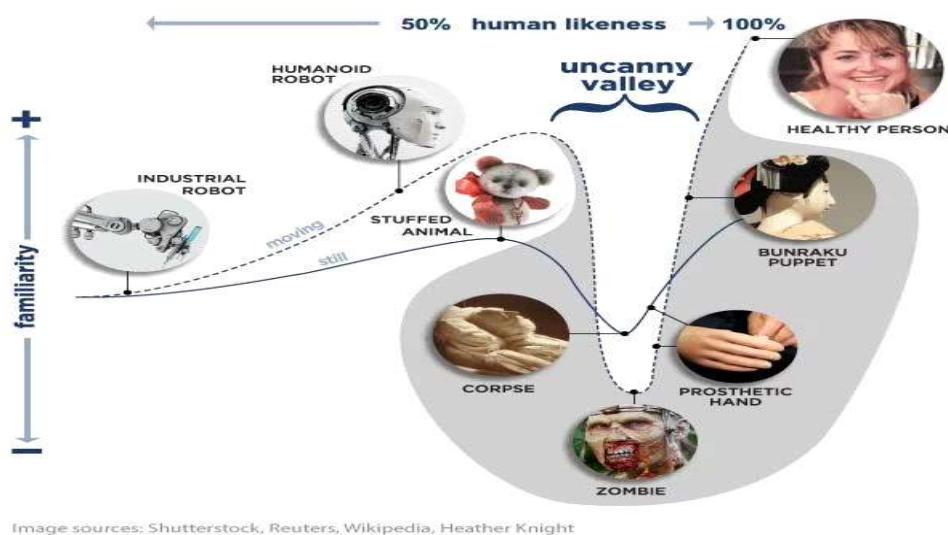
자료: Alvaro Lamarche-Toloz, Feb 12, 2020

<https://virtuals.co/digital-humans-virtual-humans-differences-overview/>

2. 디지털 휴먼 특징

○ 실재감과 불쾌한 골짜기

- 일반적으로 디지털 휴먼은 사람과 유사한 형태로 구현하고자 하는 경우가 많음. 이는 디지털 휴먼의 사실성(realism)을 높임으로써 실제로 함께 존재하는 실존감 (co-presence)과 상호작용하는 사회적 존재감(social presence)을 높여서 사용자에게 만족감과 효용성을 극대화하기 위함임
- 그러나 실제 사람과 유사하게 구현될 경우에 불쾌한 골짜기(Uncanny Valley) 이슈가 제기 될 수 있음, 불쾌한 골짜기란 1970년에 모리 마사히로가 제시한 이론으로 사람의 모습과 흡사해질수록 호감도가 증가하다가 어느 정도에 도달하게 되면 강한 거부감을 느끼고 이를 넘어서서 인간과 거의 구별이 불가능한 정도가 되면 호감도가 다시 증가해서 인간처럼 감정을 가질 수 있다는 것임
- 불쾌한 골짜기 논란을 피하기 위해서는 크게 두 가지 접근법이 있을 수 있음
 첫째는 사람이 보기에 실제 사람이 아니라는 사실을 명확히 알 수 있도록 외형을 표현하는 방법이 있음, 예를 들어 토이스토리의 경우 나오는 캐릭터들이 마치 사람처럼 얘기하고 행동하지만 장난감의 외형을 가지고 있기 때문에 사람들이 불쾌함을 느끼지 않음, 이 방법은 광고, 영화 등에서 많이 활용되었음
 두 번째 방법은 실제 인간과 구별이 되지 않을 정도로 정교하게 외형 및 상호작용을 설계하는 방법임, 현재 많은 연구들은 이 수준을 목표로 하고 있음



〈그림2〉 불쾌한 골짜기 이론

자료: <https://towardsdatascience.com/the-uncanny-valley-in-game-design-6a6c38a36486>

II. 디지털 휴먼 연구 및 기술 동향

1. 디지털 휴먼 모델 생성 기술

○ 디지털 휴먼의 모델을 생성하는 기술은 크게 세 가지로 나눌 수 있음. 사람의 형상을 3차원 픽셀의 집합으로 만드는 볼류메트릭 캡처 기술, 단일 이미지로 애니메이션이 가능하도록 메시(Mesh)와 텍스처(Texture)로 만드는 휴먼 모델링 기술, 마지막으로 딥러닝으로 임의 시점 및 다양한 표정과 자세의 사람 영상을 만드는 뉴럴 휴먼 렌더링 기술이 있음

○ 볼류메트릭 캡처 기반 디지털 휴먼 모델링

- 볼류메트릭 캡처 기술은 많은 수의 카메라를 사용하여 공간상의 3차원 픽셀 단위로 형태와 색상을 촬영하는 방식임. 대표적인 예로 서던캘리포니아 대학교(University of Southern California)와 구글에서 공동 개발하여 발표한 Relightables 시스템이 있음. 이 시스템은 기존에 USC에서 보유하고 있던 Light Stage의 최신 버전으로 331개의 LED와 4112x3008 해상도의 58개의 컬러 카메라, 32개의 적외선 카메라를 이용해서 60fps의 속도로 고해상도 노멀맵(Normal Map), 텍스처(Texture), 깊이맵(Depth Map) 등을 생성해 냄

이 시스템으로 생성한 모델은 매우 정교한 디테일까지 보유하면서 주변 광원에 맞게 새로 렌더링 할 수 있는 장점이 있음



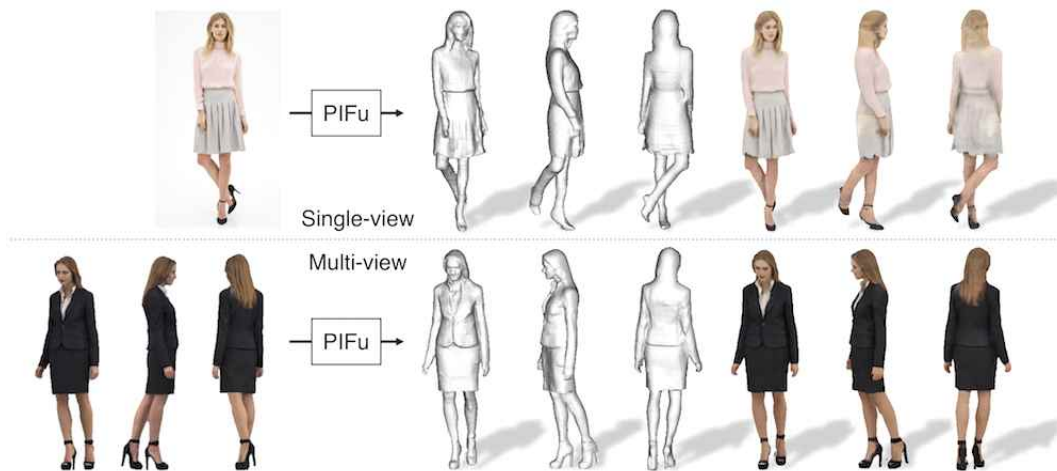
〈그림3〉 The Relightables 시스템

자료: Kaiwen Guo et al.,

“The Relightables: Volumetric Performance Capture of Humans with Realistic Relighting.” SIGGRAPH Asia 2019

○ 단일 이미지 기반 디지털 휴먼 모델링

- 여러 뷰에서 촬영을 한 여러 장의 이미지로 다중시점기하학(Multi-view Geometry)을 이용하여 디지털 휴먼을 모델링 하든지 혹은 3차원 카메라를 사용하여 직접적으로 모델링 하는 방식이 주로 사용되는 방법이지만 최근에는 한 장의 이미지만으로 3차원을 모두 복원하는 방법이 소개되고 있음. USC와 Facebook에서 2020년 CVPR에 발표한 PiFu 방식은 한 장의 정면 전신 사진에서 전신 모델 및 애니메이션 가능한 모델을 생성함, 정면 사진만 있지만 딥러닝 기법을 활용하여 뒷면의 모양까지 생성이 가능함

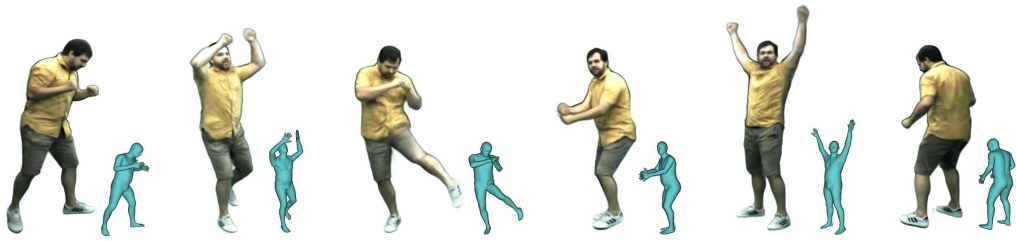


〈그림4〉 단일 이미지 기반 모델링 PiFu

자료: Saito et al, "PiFu: Pixel-Aligned Implicit Function for High-Resolution Clothed Human Digitization.", ICCV 2019

○ 뉴럴 휴먼 렌더링 기술

- 최근 들어 딥러닝 기술의 발달과 더불어 디지털 휴먼도 최신 딥러닝 기법을 적용하는 경우가 자주 발생하고 있음. 그 중 대표적인 것이 NeRF(Neural Radiance Fields)를 적용하는 방법임
- 독일 MPI와 페이스북은 SMPL 신체 모델과 NeRF 영상 합성 기술을 융합하여 임의 시점 및 동작 렌더링이 가능한 뉴럴 액터 기술을 개발하였음. 이 방법을 이용하면 2차원 이미지들을 사용하여 학습을 한 후에 해당 상대에 대한 고화질의 합성 결과를 임의의 시점이나 임의의 자세로 생성이 가능함. 이 기술을 이용하면 한 명의 학습결과를 이용하여 다른 사람들의 이미지도 임의 시점 및 자세 렌더링의 적용이 가능함
- 복잡한 모델링이 필요하거나 값비싼 장비로 대상을 촬영할 필요가 없기 때문에 앞으로 다양하게 적용될 것으로 보임. 특히나 실감 만남 등에 사용할 때 보다 손쉽게 적용이 가능할 것으로 보임



〈그림5〉 NeRF를 사용한 디지털 휴먼 렌더링

자료: Habermann et al, “Neural Actor: Neural Free-view Synthesis of Human Actors with Pose Control”, SIGGRAPH Asia 2021

2. 디지털 휴먼 상호작용 기술

○ 대화 상호작용 기술

- 디지털 휴먼 중에서 실제로 형체를 가지면서 대화도 가능한 종류를 ECA(Embodied Conversational Agent)라고 부름. 이런 ECA는 사용자들이 함께 무언가를 하고 있다는 감각인 사회적 실재감(Social Presence), 사교적인 감각(Social Richness)을 더 많이 느낄 수 있게 함¹⁾
- ECA의 대화는 자연어 처리(Natural Language Processing) 기술들과 맞닿아 있음, 구글의 DialogFlow나 마이크로소프트의 Power Virtual Agents 등과 같은 서비스들이 있으며, GPT-3(Generative Pre-trained Transformer)-3*나 BERT**(Bidirectional Encoder Representations from Transformers) 모델을 사용하여 자연스럽게 대화가 가능한 기술들이 점차 발생하고 있음. ECA의 대화는 이러한 챗봇(chatbot)의 기술을 뛰어 넘어 상황과 관련된 (context-related) 인간 유사 대화를 해야 하는 필요가 있음. 최근에는 이런 지능형 가상 에이전트(intelligent virtual agent) 기술이 점차 현실화되고 있음

* GPT-3 : 딥러닝을 이용해 인간다운 텍스트를 만들어내는 자기회귀 언어 모델

** BERT : 자연어처리를 위한 트랜스포머 기반 양방향 인코더 기술

1) K. Kim, N. Norouzi, T. Losekamp, G. Bruder, M. Anderson, and G. Welch, “Effects of patient care assistant embodiment and computer mediation on user experience,” in International Conference on Artificial Intelligence and Virtual Reality, IEEE, 2019.

- 디지털 휴먼의 의사소통은 언어적인 행동(verbal behavior)과 비언어적인 행동(nonverbal behavior)으로 나눌 수 있음, 언어적인 행동은 직접 대화나 수화 등 의미를 전달하기 위한 행동으로 볼 수 있고 비언어적인 행동은 대화 이외에 의사소통을 위해 사용되는 동작임. 비언어적인 행동의 예로는 시선 응시, 눈 깜빡임, 얼굴 표정, 손 움직임, 고개 움직임, 몸 움직임 등으로 자신의 의도를 전달하게 됨
- 이러한 비언어적인 행동을 생성하기 위해서는 실제 대화에서 어떤 비언어적인 행동들이 발생하였는지를 관찰해서 반영하는 방법과 기계학습 등을 통해서 규칙을 생성하는 방법이 있음. 최근에는 영상에서 대화하는 사람의 자세나 표정과 그 대화 내용을 같이 학습을 하여 입력으로 대화가 들어오면 적절한 비언어적 행동을 생성하는 방식들이 자주 쓰이고 있음²⁾

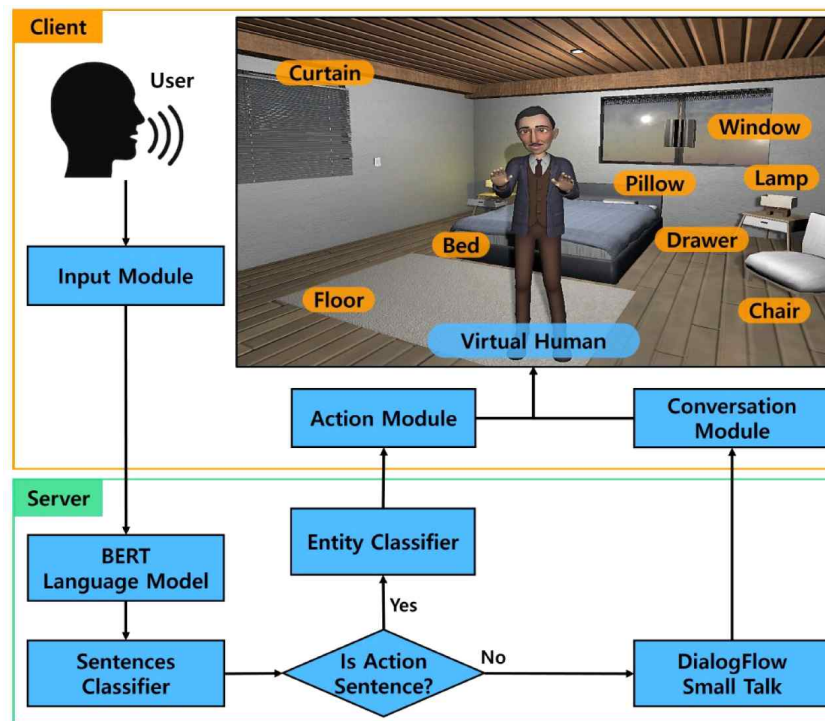
○ 행동 상호작용 기술

- 디지털 휴먼의 행동을 제어하는 대표적인 방법은 모션 캡처(Motion Capture)가 있음, 다양한 동작 애니메이션 클립을 모션 캡처를 이용하여 저장한 후에 필요할 때 이를 재생하는 방법이 가장 일반적임
- 이 방법은 실시간 동작이 필요 없다면 게임과 같이 동일한 동작을 반복하는 것으로 충분할 경우에 사용이 가능함. 요즘은 모션 캡처 장비와 기술이 대중화되어 쉽고 저렴하게 모션 캡처와 얼굴 캡처가 가능해짐, HTC Vive의 트랙커를 이용해서 몸 동작을 캡처한다든지 페이스리그(FaceRig)와 같은 제품을 이용해서 간단하게 얼굴 표정을 읽을 수가 있음. 이렇게 실시간으로 몸동작 및 얼굴 모션을 캡처하는 경우 실시간 디지털 휴먼 인터랙션이 가능해서 버추얼 유튜버(Virtual Youtuber)와 같은 응용 분야에 적용이 가능함
- 디지털 휴먼과 상호작용하는 방법 중에 대화 이외에 특정한 행동을 하도록 유도할 수 있음. 이를 위해서는 행동 저작이 필요한데 행동 저작을 위한 유명한 언어로 BML(Behavior Markup Language)이 있음. BML은 디지털 휴먼의 행동을 정의하는 마크업 언어로 연속적으로 행동을 정의할 수가 있음³⁾

2) G. Ali & M. Lee & J.I. Hwang, "Automatic text-to-gesture rule generation for embodied conversational agents," Computer Animation and Virtual Worlds, 2020

3) Aggarwal, Priti & Traum, David. "The BML Sequencer: A Tool for Authoring Multi-character Animations". Intelligent Virtual Agents 2011

- 최근에는 대화 모델을 이용하여 사용자가 음성으로 말하면 디지털 휴먼이 이를 행동으로 실행하는 방법들이 개발되고 있음, 음성 대화를 분석해서 Entity들을 추출하여 어떤 대상으로(Subject), 어떤 행동을(Action), 어디에서(Position), 어디로(Target) 하는지를 분석하여 수행에 옮김



〈그림6〉 디지털 휴먼 행동 인터랙션,

자료: Kim et al, "Auto-generating Virtual Human Behavior by Understanding User Contexts", IEEE VR 2021

○ 환경 상호작용 기술

- 혼합 현실(Mixed Reality)과 같은 분야에 디지털 휴먼을 사용할 때 사람이 아닌 환경과 디지털 휴먼의 상호작용 또한 고려되어야 할 사항임. 이런 경우에는 주위 환경에 대해서 인식(Recognition)하고 분할(Segmentation)하는 것이 중요한 기술적 요소임, 인간이 주위 환경을 인지하고 회피하거나 상호작용을 하는 것과 마찬가지로 디지털 휴먼도 이러한 기능을 보유하여야 함
- 주위 사물을 인식하고 분할하기 위해서 카메라와 같은 센서를 이용할 수 있는데 깊이 카메라와 같은 센서가 있다면 보다 손쉽게 인식 및 분할이 가능함. 최근에는 Mask RNN과 같은 딥러닝 기법의 발달로 보다 손쉽게 이러한 인식 및 분할이 가능해지고 있으며 또한 거의 실시간에 가까운 인식 및 분할이 가능함

- 아래 그림과 같이 환경을 인식하고 분할할 시에 실제 물체와 디지털 휴면이 상호작용하게 할 수가 있음. 즉 디지털 휴면이 실제 물체가 그 위치에 있는 것처럼 그 사이를 피해 간다든지 실제 물체를 밀어서 치우려고 노력한다든지 하는 행동이 가능함. 또한 가상 물체와 실제 물체가 충돌했는지를 검출하는 충돌 검사(collision detection)이나 실제 물체가 가상 물체를 가리게 하는 가림 현상(occlusion) 등도 구현이 가능함. 충돌 검사나 가림 현상이 구현되지 않으면 디지털 휴면이 실제 사물을 뚫고 지나간다든지 뒤로 지나가지만 실제 물체에 가려지지 않는다든지 하는 문제가 발생하게 됨



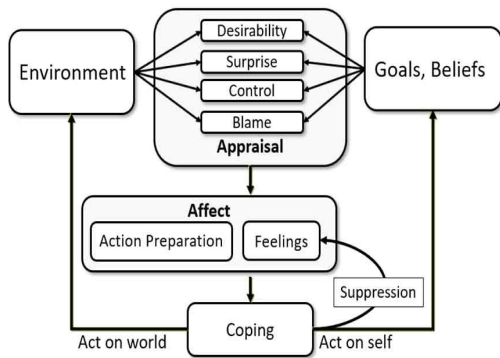
〈그림7〉 환경 인식 및 분할과 디지털 휴먼의 인터랙션,

자료: Kim et al, "Silhouettes from Real Objects Enable Realistic Interactions with a Virtual Human in Mobile Augmented Reality", Appl. Sci. 2021, 11, 2763.

3. 디지털 휴먼 감정 생성 기술

○ 감정 계산 모델링 (Computational Modeling of Emotion)

- 디지털 휴먼의 감정을 모델링 하는 것은 디지털 휴먼의 표정이나 행동 등을 정의하기 위한 기반 작업으로 매우 중요함. 디지털 휴먼의 감정 모델링은 결국 인간의 감정을 모사하는 것으로 귀결이 됨. 디지털 휴먼에 적용하기 위해서 감정을 계산하여야 하고 이는 감정 계산 모델링(Computational Modeling of Emotion)이라는 연구 분야로 발전하게 됨
- 감정을 계산하는 이론에는 크게 세 가지 정도의 방법이 있는데 이산감정이론(Discrete Theories of Emotion), 둘째는 차원감정이론(Dimensional Theories of Emotion) 마지막으로 평가감정이론(Appraisal theories of emotion)이 있음
- (이산감정이론) 생물학적으로 정해진 몇 가지의 이산적인 핵심 감정이 있다는 이론으로, 예를 들어 분노나 슬픔 같은 핵심 감정들을 이동하면서 느낀다는 것임
- (차원감정이론) 감정이 연속적인 것으로 몇 가지의 요소를 결합해서 특정한 감정이 표현되는데 마치 다차원의 한 점처럼 여러 요소의 값이 합쳐져서 정해지기 때문에 연속적인 형태로 감정이 변화한다는 이론임. 예를 들어 PAD 감정모델이 대표적인데 기쁨(Pleasure), 각성(Arousal), 지배(Dominance)의 세 가지 요소가 적정 비율로 결합이 되면 특정 감정이 발생한다는 이론임
- (평가감정이론) 특정 감정은 정신적인 과정(Mental Process)이 일어나서 그 결과로 발생된다고 제시하고 이 과정에서 내면 혹은 환경의 평가(Appraisal)가 중요한 요소로 작동한다고 보고 있음
- 이 세 가지 방법 모두 디지털 휴먼의 감정 모델링에 사용이 될 수 있으나 다양한 연구 그룹에서는 평가감정이론을 많이 채택하여 디지털 휴먼의 감정의 변화를 계산하고 있음. 예를 들어 USC ICT에서는 EMA(EMotion and Adaptation)이라는 평가 모델을 사용하고 있는데 이를 이용해서 디지털 휴먼을 제작하여 활용하고 있음



〈그림8〉 (좌) 평가 모델을 사용한 감정 설계, (우) 감정 설계에 기반한 디지털 휴먼 인터랙션

자료: USC ICT, <https://people.ict.usc.edu/~gratch/>

- 이러한 감정 모델은 디지털 휴먼과 사람 간의 신뢰관계를 형성하는데 많은 도움이 됨, 이러한 신뢰관계를 라포(Rapport)라고 부르는데 사용자의 표정 및 동작 등을 컴퓨터 비전 등의 기술로 분석해서 상대의 감정을 파악하고 그에 걸맞은 표정 및 동작을 디지털 휴먼이 생성해서 라포를 획득하게 됨

Ⅲ. 디지털 휴먼의 활용 분야

1. 문화 콘텐츠

○ 버추얼 인플루언서

- 미국의 시장조사 업체인 비즈니스 인사이더 인텔리전스에 따르면 기업이 인플루언서에게 쓰는 마케팅 비용은 2019년 80억 달러(약 9조 1800억 원)에서 2022년 150억 달러(약 17조 2000억 원)로 늘어날 전망이다. 미국 블룸버그통신은 이 중 상당 부분을 가상 인플루언서가 차지할 것이라고 분석하기도 함⁴⁾
- 이러한 버추얼 인플루언서는 실제 사람과 구분이 어려울 정도이며 시공간의 제약도 없고 사생활의 문제 등이 없기 때문에 다양한 방면에 사용이 가능함
- 버추얼 인플루언스는 음악, 영화, 방송 등 다양한 문화 콘텐츠에 사용이 가능할 것으로 보임

	릴 미켈라 19세 여성, LA 거주 인스타 팔로워: 300만 명 제작사: 미국 스타트업 브러드 활동: 샤넬, 프라다, 캘빈 클라인 모델, 싱글 앨범 발매(영국 스포티파이 8위) 수익: 2020년 130억원
	이마 20대 초반 추정, 여성 인스타 팔로워: 34만 명 제작사: 일본 스타트업 AWW 활동: 이케아 광고 모델 수익: 2020년 약 7억원
	루이 리 22세 여성 유튜브 루이커버리(RuiCovery) 구독자: 2만3000명 제작사: 디오비스튜디오 활동: 여행 브이로그, 노래 영상 커버, 쇼핑물 모델
	김래아 23세 여성, 서울 거주 제작사: LG전자 인스타 팔로워: 1만2500명 이상 활동: 싱어송라이터 겸 DJ, 올해 초 CES 2021 프레스 컨퍼런스에 등장

〈그림9〉 버추얼 인플루언서

자료: 시사저널

4) 시사저널, “현실세계로 스며든 가상 인플루언서”, <https://www.sisajournal.com/news/articleView.html?idxno=221677>

○ 버추얼 어시스턴트

- 네온(NEON)은 대화형 에이전트로 개인화된 뉴스를 전해주는 앵커나 쇼핑 호스트 등으로 개발되고 있음, 매장이나 공항 등에서 인간 직원 대신 고객을 응대하는 점원이 될 수 있고 사용자와 대화하고 도움을 주는 존재로 기획됨
- 2016년 창립된 Soul Machine사에서는 학습을 하고 반응을 하는 디지털 휴먼을 개발하고 있음. 주의, 학습, 감지 및 행동을 제어하는 기본인지 프로세스를 시뮬레이션해서 마치 인간이 생각하고 행동하는 것과 같이 구현이 됨, 이를 통해서 인간을 대신하는 가상 에이전트를 제공하고 있음

2. 원격 실감 만남

○ VR/AR 원격 회의

- 코로나19 등의 상황으로 비대면 회의와 재택근무가 일상화된 상황에서 현재의 영상 위주의 제한적인 원격 만남에 VR/AR 기술을 적용한 시스템들이 점차 늘어나고 있음, Engage, MeetinVR, AltSpaceVR, VRChat, Rec Room, Hoppin', Cluster, JumpVR, Mozilla Hubs, Spatial 등이 그러한 플랫폼임. 이 중 Spatial은 자신과 닮은 아바타 형태를 사용해서 협업을 하게 되는 플랫폼인데 홀로렌즈, PC, 모바일 폰 등 다양한 플랫폼에서 VR 및 AR을 사용해서 참여가 가능함. 원격 회의 및 협업과 같은 비즈니스 분야에서는 자신의 모습을 그대로 표현하는 디지털 더블 방식이 적합할 것으로 예상됨



〈그림10〉 Spatial을 활용한 원격 회의

자료: <https://spatial.io/>

IV. 결론

1. 디지털 휴먼과 최신 기술의 결합

○ 디지털 휴먼과 메타버스

- 메타버스는 가상과 현실을 결합하는 공간으로 디지털 휴먼이 가상 인간이나 실제 사람을 대변하는 표현 형태로 활용이 됨
- 메타버스에서는 가상 인간과 실제 인간 참여가 혼재되는 형태가 될 것으로 예상되는 바 보다 적극적인 디지털 휴먼 기술이 적용될 것으로 보임. 또한 AR 및 VR 환경에서 디지털 휴먼을 표현하고 상호작용하는 기술은 가상의 인간과 실제 인간과의 구분이 어려울 정도 수준이 될 것으로 예상됨

○ 디지털 휴먼과 인공지능

- 최근 딥러닝 분야의 연구가 발전하면서 디지털 휴먼에도 새로운 전기가 마련되고 있음, 렌더링 분야에서는 NeRF 등의 기술이 적용되면서 근미래에 실제와 유사한 디지털 휴먼의 표현이 가능해질 것으로 보임
- 또한 GPT나 BERT 등의 자연어 처리 관련 기술의 발달로 인간과 유사한 대화를 생성하는 기술이 적용되면서 디지털 휴먼의 대화도 점차 실제 인간과 근접하게 될 것으로 전망됨
- 영상의 동작을 분석해서 디지털 휴먼 대화 동작 및 행동을 생성하는 분야는 비교적 기초적인 단계이나 향후에 보다 자연스러운 동작을 하는 디지털 휴먼을 만드는 데 일조할 것으로 기대가 됨, 디지털 휴먼의 감정을 모델링 하는 것은 다양한 모델이 제시되고 있으나 아직은 연구할 여지가 많은 것으로 고려됨. 그러나 향후 인간의 감정과 유사한 감정을 재현하는 디지털 휴먼이 가능할 것으로 생각됨

2. 디지털 휴먼의 미래 발전 방향

○ 버추얼 컴패니언과 메타휴먼

- 버추얼 컴패니언 (Virtual Companion): 버추얼 컴패니언은 마치 인간과 같이 사람과 교감이 가능한 수준의 디지털 휴먼을 의미함. 즉, 인간과 같은 수준의 감정 표현 및 상대 이해력을 가지고 있어 인간이 버추얼 컴패니언으로부터 정서적인 위안을 얻고 가족과 같은 연대감을 가질 수 있는 정도 수준을 목표로 함. 이는 디지털 휴먼의 감정 모델링 기술 및 상호작용 기술과 깊은 연관성이 있음. USC의 SimSensei 등에서는 디지털 휴먼을 사용해서 상대의 감정을 이해하고 상대의 감정에 맞는 대화를 제시하는 연구들이 진행되었고 인간 수준의 라포가 형성되는 것을 실험으로 보여준 바가 있음
- 메타 휴먼 (Meta Human): 메타 휴먼은 일반적인 디지털 휴먼을 넘어서서 실제 인간의 경험, 지식, 성격까지 복제할 정도의 수준을 의미함. 메타 휴먼 수준에서는 디지털 휴먼이 자아 정체성을 가지고 인간과 구분이 힘든 정도의 인지적, 이성적 사고가 가능해짐, 많은 영화 및 소설 등에서 표현되는 인간의 디지털적인 복제를 통한 연속적인 존재의 구현이 메타 휴먼적인 기술이라고 볼 수 있음. 테슬라에서는 뉴럴링크 등의 연구를 통해서 사람의 뇌와 통신을 해서 이를 전송하는 등의 시도를 하고 있음