

## 버추얼 휴먼 가수의 기술적 발전 연구

이 용 수<sup>1</sup> · 이 철 희<sup>2\*</sup><sup>1</sup>경희대학교 포스트모던음악학과 석사과정<sup>2\*</sup>경희대학교 포스트모던음악학과 조교수

# Study on the Technological Development of Virtual Human Singers

Yong-Su Lee<sup>1</sup> · Chul-Hee Lee<sup>2\*</sup><sup>1</sup>Master's Course, Department of Postmodern Music, KyungHee University, Yongin 17104, Korea<sup>2\*</sup>Assistant Professor, Department of Postmodern Music, KyungHee University, Yongin 17104, Korea

### [요 약]

본 논문은 버추얼 휴먼 가수의 기술적 발전에 대한 연구이다. 최근 4차 산업혁명의 영향으로 가상과 현실의 융합이 가속화되면서 버추얼 휴먼 산업 시장이 점차 확대되고 있다. 그중 버추얼 휴먼 가수는 엔터테인먼트 업계에서 특히 주목받고 있다. 따라서 본 연구는 과거 버추얼 휴먼 가수 제작 기술의 한계점과 기술의 발전 현황을 파악하고 향후 전망에 대해 제시하고자 시작되었다. 연구 결과 버추얼 휴먼 가수 제작에 사용되는 기술은 새로운 기술의 개발, 기존 기술의 한계점 보완의 과정을 거쳐 모션캡처, 음성합성, AI기반 딥러닝, 딥리얼 기술 적용 등 지속적인 발전을 해왔으며 현재도 연구가 지속적으로 진행되고 있음을 알 수 있었다. 본 연구의 내용을 통해 앞으로 새롭게 등장할 버추얼 휴먼 가수의 활동 영역, 새롭게 등장하는 기술, 향후 전망에 대해 예상해 볼 수 있게 될 것이다.

### [Abstract]

This paper presents a study on the technological development of virtual human singers. The virtual human industry market is gradually expanding with the accelerated convergence of the virtual space and reality owing to the influence of the 4th industrial revolution. Specifically, virtual human singers are drawing attention in the entertainment industry. Therefore, this study first identifies the limitations of existing virtual human singer production technology and the current technological developments and presents future prospects. Studies have shown that technologies used to produce virtual human singers, such as motion capture, voice synthesis, AI-based deep learning, and deep real technology, have continued to advance through the development of new technologies and supplementation of existing technologies. The findings of this study can provide insights into predicting the activities of new virtual human singers, emerging technologies, and future prospects.

**색인어** : 인공지능, 딥러닝, 가상 인간, 버추얼 휴먼 가수**Keyword** : Artificial Intelligence, Deep Learning, Virtual Human, Virtual Human Singer<http://dx.doi.org/10.9728/dcs.2023.24.6.1261>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 29 May 2023; Revised 15 June 2023

Accepted 19 June 2023

**\*Corresponding Author; Chul-Hee Lee**

Tel: +82-31-201-2863

E-mail: [ch@khu.ac.kr](mailto:ch@khu.ac.kr)

## I. 서론

4차 산업혁명의 영향으로 가상과 현실의 융합이 가속화되고 있다. 사물인터넷, 인공지능, 딥 러닝, 모션캡처, 5G 등 디지털 기술의 발전과 코로나 팬데믹으로 인한 비대면 상태가 지속되며 급부상한 메타버스는 사장되었던 버추얼 휴먼 산업의 재조명을 불러왔다[1]. 버추얼 휴먼은 사람처럼 이름, 성별, 나이 등 구체적인 특징을 갖고 있고 일반적인 인플루언서 처럼 SNS를 통해 일상을 공유하고 댓글과 메시지를 주고받는 등 실제 인간과 같은 소통을 할 수 있다. 또한 버추얼 휴먼은 장소와 시간의 제약, 회사와의 계약 문제, 각종 루머 등의 제한점에서 실제 인간과 달리 비교적 제약 없이 활동할 수 있다는 장점이 있어 엔터테인먼트 업계에서 특히 주목받고 있다[2]. 그 대표적 사례는 버추얼 휴먼 가수이다. 버추얼 휴먼 가수는 1990년대 초창기 모델이라고 볼 수 있는 ‘다테 쿄코’, ‘아담’의 개발을 시작으로 지속적으로 제작되고 있다. 컴퓨터 그래픽스 기술과 AI 기술에 관한 연구가 진행되고 발전하면서 버추얼 휴먼 가수의 완성도 역시 향상되고 있다[3].

본 연구는 버추얼 휴먼 가수의 기술적 발전에 관한 연구이다. 따라서 연구의 범위와 방법은 다음과 같다. 먼저 선행 연구 검토를 통해 버추얼 휴먼의 개념을 정의하고 버추얼 휴먼 가수의 역사를 살펴본다. 다음으로 버추얼 휴먼 가수의 사례를 살펴본다. 본 연구에서 제시하는 버추얼 휴먼 가수의 사례별 주요 제작 기술은 크게 버추얼 휴먼 가수의 시초, 모션 캡처 기술의 적용, 음성 합성 기술의 적용, AI 기술 기반 제작 총 네 가지로 분류했다. 제시한 버추얼 휴먼 가수의 사례를 바탕으로 기술적인 발전과 한계점 및 향후 전망에 대해 논의한다.

## II. 이론적 배경

### 2-1 버추얼 휴먼의 개념과 정의

버추얼 휴먼이란 ‘Virtual(가상)’과 ‘Human(사람)’이 결합된 합성어로 컴퓨터 그래픽을 통해 제작된 스크린 상의 2D, 3D의 가상 인물을 뜻한다[4]. 컴퓨터 그래픽을 활용해 만들어진 버추얼 휴먼은 실제 인간과 유사한 모습으로 이름, 성별, 나이, 출신 등 구체적인 배경을 갖고 있다. 버추얼 휴먼은 활동 영역에 따라 라이브 스트리밍에서 활동하는 ‘버추얼 스트리머’, 유튜브에서 활동하는 ‘버추얼 유튜버’, SNS와 광고에 활동하는 ‘버추얼 인플루언서’, 가수와 같은 음원 활동을 하는 ‘버추얼 가수’ 등 다양한 유형으로 분류된다[5].

### 2-2 버추얼 휴먼 가수의 역사

버추얼 휴먼 가수의 탄생에는 1990년대 IT산업의 발전, PC의 대량 보급, 음악시장의 변화와 같은 시대적 흐름이라는

배경이 있다. 특히 국내의 경우 개인용 컴퓨터 연도별 보급이 1998년에서 2000년까지 매년 50% 이상 급격하게 증가했다.

표 1. 국내 개인용 컴퓨터 연도별 보급현황[6]

Table 1. Distribution status of personal computers by year[6]

Year	1997	1998	1999	2000	2001
Supply number (Thousand unit)	1,921	1,851	2,881	4,602	3,834
Rate of increase (%)	2.9	-3.6	55.6	59.7	-16.7

한편 국내에서는 90년대 초 서태지와 아이들의 등장 이후 새로운 장르의 음악과 퍼포먼스, 의상, 스타일 등의 영향으로 10대 수요층이 증가하게 되었다. 10대 수요층을 공략하기 위해 SM엔터테인먼트, 대성기획(현 DSP미디어)을 필두로 이들의 취향을 반영하여 기획 단계에서부터 회사가 적극적으로 개입하여 이른바 기획형 가수를 제작하게 된다[7]. 또한 디지털 미디어의 발전으로 대중들이 라디오보다 TV화면에서 가수를 접하는 일이 잦아지면서 가창력뿐만 아니라 퍼포먼스, 비주얼 등 외적인 모습은 대중가수에게 중요한 요소가 되었다. 이렇게 대중음악의 흐름이 변화하는 시점에 IT업계에서 컴퓨터 그래픽 기술을 기반으로 하는 버추얼 휴먼 가수가 주목받기 시작했다. 제작 과정에서 가수의 외모, 모션 캡처를 이용한 실수 없는 퍼포먼스, 화려한 뮤직비디오 연출 등을 제작자가 모두 기획할 수 있다는 점에서 버추얼 휴먼 가수는 매력적인 기획형 가수이자 IT 상품이였다. 당시 버추얼 휴먼의 제작에 있어 필요한 소프트웨어 가격이 하락했고 이는 컴퓨터 그래픽 제작 기술이 발전하던 시기와 맞아떨어졌다. 1995년도 PC의 운영체제가 DOS(disk operating system)에서 Windows로 변화했고 PC의 보급이 확산하면서 워크스테이션에서만 운영되던 3D 제작툴이 PC에도 설치할 수 있었다. 특히 3D studio max, Softimage, Maya 등 3D 제작툴의 가격이 하락하면서 국내 IT기업이 사용하기 용이했다. 버추얼 휴먼을 제작하는 기술도 빠른 속도로 발전했는데 Nurbs와 Polygon 모델링이 가능해졌고 UV좌표를 활용한 Mapping 기술도 개발되었다. 또한 PC의 빠른 발전은 렌더링을 빠른 속도로 가능하게 했고 PC의 확산은 인터넷을 통한 개인 간의 기술 공유와 맞물려 컴퓨터 그래픽 기술의 발전을 촉진했다[3].

IT기술의 발달, 음악시장의 변화와 함께 1990년대 새로운 소비층인 X세대가 등장했다. X세대는 베이비붐 세대 이후 태어난 세대를 지칭하는 말로 우리나라의 X세대는 산업화의 영향으로 물질적, 경제적으로 비교적 안정적인 환경에서 성장했다는 세대적 특성을 갖는다. 그 영향으로 X세대는 기존의 가치, 관습에서 자유롭고 싶어 하고 자신이 좋아하는 분야에 집중하는 특징을 보인다[8]. 이런 특징은 X세대가 상품을 필요에 의해서만 구매하는 기능적 소비가 아닌 상품의 이미지, 즐거움, 소비행위를 통한 자신의 존재감 표현 등 기호적 소비패턴을 보이게 했다. X세대의 이런 소비패턴은 가수의 음반뿐만 아니라 엔터테인먼트적인 상품성을 소비하는 행위와 맞물

렸다. 결론적으로 1990년대 후반 새로운 소비패턴을 가진 X 세대의 등장, 디지털 미디어의 발달, 가정의 대량 PC 보급, IT 기술의 발달, 음악시장의 변화와 같은 시대적 흐름 속에서 버추얼 휴먼 가수가 등장했다.

### III. 버추얼 휴먼 가수의 기술

#### 3-1 버추얼 휴먼 가수의 시초

##### 1) 다테 료코

가수 다테 료코는 1996년 일본의 연예 기획사 호리프로의 주도로 탄생한 세계 최초의 버추얼 휴먼 가수이다. 다테 료코는 1990년대 중반 일본의 게임 제작사인 코나미에서 제작한 연예 시뮬레이션 게임 「두근두근 메모리얼」의 흥행의 영향을 받아 기획되었다[9]. 다테 료코는 ‘DK-96’이라는 이름으로 디지털 음원 <Love communication>을 발매하며 데뷔했다. 다테 료코의 음반 판매량은 약 3만 장을 기록했고 보컬을 비롯한 인터뷰, 라디오 등에 사용된 음성은 실제 사람의 목소리를 사용했다. 다테 료코의 외형은 후디니(Houdini)라는 컴퓨터 그래픽 기술을 사용해 제작되었다[10].

후디니는 Side Effects社에서 개발된 노드를 기반으로 한 워크플로우 형식의 3D 제작 프로그램이다. 노드는 특정 기능을 갖는 하나의 블록을 뜻하며 이 노드 블록들이 쌓여 각각의 위치에서 정해진 일을 처리하고 결과를 만들어 낸다. 이런 노드 기반 워크플로우 형식은 작업 과정을 직관적으로 볼 수 있어 문제가 생겼을 때 쉽게 파악할 수 있다는 점과 한번 작업했던 프로젝트와 비슷한 작업을 진행할 경우 빠른 속도로 작업할 수 있다는 장점을 갖는다[11].

외형이 모델링 되면 움직임은 만드는 작업이 이루어지는데 당시 다테 료코 제작에 사용된 모션 캡처 기술은 많은 비용과 인적 자원을 요구했다. 또한 당시 기술력의 이유로 다테 료코는 자연스러운 움직임, 목소리, 표정, 감정 표현을 구현하는데 있어서 한계가 있었다. 결론적으로 많은 제작 기간과 비용, 당시 기술력의 문제는 다테 료코의 활동에 영향을 끼쳤고 대중 매체에 노출되는 횟수가 점차 줄어들었다. 이러한 이유로 다테 료코에 대한 대중의 관심이 사라지고 소모가 줄면서 이는 흥행 실패의 원인이 되었다. 하지만 다테 료코는 인간 가수처럼 음반을 발매하고 미디어에서 활동한 최초의 버추얼 휴먼 가수라는 점에 의의가 있으며 이후 수많은 버추얼 휴먼 형태의 가수가 등장하는 계기가 되었다.

##### 2) 아담

가수 아담은 아담 소프트에서 제작하여 1998년 1집 앨범 <세상엔 없는 사람>으로 데뷔한 국내 최초의 버추얼 휴먼 가수이다. 아담은 실제 인간처럼 성격, 성향, 인간성 등이 부여되어 있다. 아담의 외형은 3D 그래픽 기술을 통해 제작되었고 목소리는 가수 박성철이 참여한 것으로 알려졌다[10].

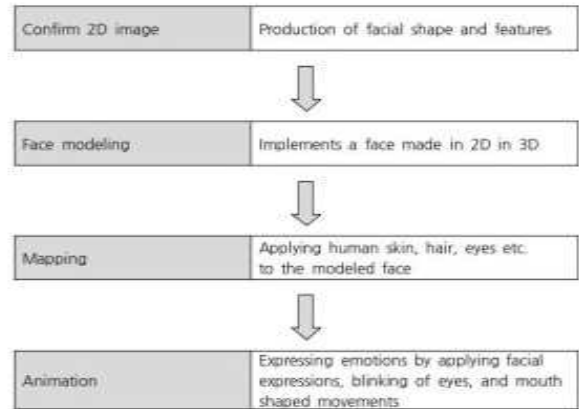


그림 1. 아담의 제작 과정  
Fig. 1. Adam's production process

아담의 제작에 사용된 소프트웨어는 Softimage, 3D studio max, Alias로 제작과정은 위의 그림 1과 같다. 해당 제작 과정은 약 3개월에 걸쳐 이루어졌으며 사람의 형상을 하고 있기 때문에 비교적 부드러운 곡면 표현이 가능한 Softimage의 Nurbs 방식이 사용되었다. 아담의 얼굴을 제작하기 위해 먼저 전체적인 형태를 만들고 세부적인 이목구비와 표정을 제작했다. 얼굴의 모델링이 끝나고 아담의 기획 의도, 분위기를 고려해 몸과 의상을 제작했다. 이 과정을 거쳐 아담의 얼굴과 외형 모델링이 끝나고 다음으로 입 모양을 제작했다. 버추얼 휴먼 가수의 특성상 노래를 립싱크로 불러야 하기 때문에 먼저 크게 벌린 입, 작게 벌린 입, 수직 방향으로 벌어진 입, 수평 방향으로 벌어진 입, 왼쪽이나 오른쪽으로 기울어진 입, 위 또는 아래 방향으로 치우쳐진 입 등 기본적인 입 모양을 제작하고 이 입 모양들을 조합해 말하는 모습, 표정을 만들어냈다. 이후 이미 촬영된 배경에 완성된 아담의 모습을 합성하는 작업을 통해 움직임을 만들어 냈다. 움직임이 거의 없는 부분은 수작업으로, 움직임이 많고 동작이 큰 부분은 모션 캡처를 활용해 작업했다. 모션 캡처는 사람이 16개의 센서를 몸에 부착하고 움직이거나 동작을 취하면 그 데이터를 받아 작업하는 방식으로 진행되었다. 그러나 당시의 기술로는 데이터의 떨림이 있어 사람이 일일이 데이터를 수정해야 하는 번거로움이 있었다[12]. 이러한 기술적 한계는 버추얼 휴먼 가수의 제작에 많은 인력과 시간, 막대한 금액이 필요하게 했다. 따라서 제작사 아담소프트는 아담의 광고, 인터뷰 등 활동을 위한 제작비, 노력 등에 비해 큰 성과를 얻기 힘들었고 이는 자연스럽게 아담이 대중 매체에 점차 사라지는 원인이 되었다. 아담의 흥행 실패에 대해 메타버스 엔터테인먼트를 운영하는 네오엔터테인먼트 권택준 대표는 “지금도 full 3D 캐릭터 제작은 고가여서 대다수의 버추얼 휴먼이 바디와 패션은 실존 모델을 촬영한 뒤 얼굴만 가상 얼굴을 합성한다. 예전 기술력으로 자연스러운 퀄리티의 3D 그래픽을 만드는 건 정말 상상 초월이었을 것이며, 여러모로 세상을 많이 앞서간 데 따른 결과이다.”라고 설명했다[13].

당시 아담의 1집 앨범은 20만 장의 판매량을 기록하고 여러 방송사의 섭외 요청, 광고 문의 등 많은 관심을 받는 듯 보였으나 당시 기술력의 한계, 제작에 필요한 인력, 시간, 금액 등을 이유로 오랜 기간 활동을 이어 나가지 못했다. 그러나 1990년대 최초의 버추얼 휴먼 가수 다테 쿄코의 사례를 시작으로 버추얼 휴먼 가수 산업이 큰 관심을 받고 있던 시점에서 국내 최초로 컴퓨터 그래픽 기술을 활용해 시도한 버추얼 휴먼 가수라는 점에 의의가 있다.

### 3-2 모션 캡처

#### 1) 사이다

가수 사이다는 1998년 예스네트에서 제작한 국내 3호 버추얼 휴먼 가수이다. 사이다는 출시 당시 16세 여중생이라는 설정을 두고 1집 앨범 <진실이 싫어>로 정식 데뷔했다. 사이다는 10대 여성 소비자를 타겟으로 키 168cm, 체중 45kg, 걸스카우트 입단과 같은 설정으로 또래 여성들이 동질감을 느끼게 하고 그들을 대표해 당시 청소년기 여성들이 느끼는 심리를 표현하고자 했다. 그 예로 <진실이 싫어>의 가사를 보면 여성이 외모로 인해 차별받는 당시 사회적 분위기를 비판하는 청소년의 반항 심리를 표현한 것으로 보인다[14].

사이다는 버추얼 휴먼 가수와 인간의 움직임이 결합된 점이 핵심 요소이며 이에 활용된 것이 모션 캡처 기술이다. 당시 댄스음악의 경향은 따라 하기 쉬운 동작과 가사의 특성을 담은 안무가 중요했다. 댄스음악은 음악적인 요소를 안무로 표현해 관객에게 듣는 즐거움 외의 시각적인 볼거리를 제공한다. 댄스음악에서 안무의 중요도는 사이다를 제작하는 데 있어 모션 캡처 기술이 필수적인 요소가 된 원인이 되었다. 사이다 제작에 활용된 모션 캡처 기술은 순간적인 움직임을 3차원적으로 분석하고 이를 자유롭게 편집하여 새로운 움직임으로 재창조할 수 있게 했다[3].

버추얼 휴먼 가수를 받아들이는 수용자들이 이들에게 불쾌함, 이질감 등을 느끼는 원인 중 하나가 인간과 흡사한 외형을 하는 버추얼 휴먼 가수의 실제 인간과 다른 어색한 움직임에서 오는 불쾌한 골짜기 때문일 것이다. 그만큼 버추얼 휴먼 가수의 인간과 같은 자연스러운 움직임은 대중들이 실제 인간과 같은 가수로서 받아들일 수 있는 중요한 요소이다. 따라서 모션 캡처 기술을 적극 활용한 사이다 제작 사례는 이후 버추얼 휴먼 가수의 움직임을 보다 인간의 움직임에 가깝게 하기 위한 연구의 깊이를 더한 계기가 되었다.

#### 2) 인즈

가수 인즈는 2000년 11월 14일 엔터테인먼트 전문 포털 업체인 인즈닷컴에서 제작한 댄스 그룹으로 여성 보컬 가이아, 남성 래퍼 미르, 로봇의 외형을 한 댄스 담당 킨으로 구성되었다. 인즈의 외형은 3D 제작 툴인 Maya를 활용해 모델링되었다. 약 1년의 제작 기간이 소요되었고 타이틀곡 <Get Away>를 포함해 총 12곡이 수록된 앨범과 뮤직비디오로 데

뷔했다. 인즈는 미술 해부학의 개념을 적용한 디자인, 3D 모델링, 모션 캡처를 활용한 섬세한 움직임 구현하여 초창기 버추얼 휴먼 가수인 아담에 비해 더욱 현실감 있는 색감과 질감을 구현했다[15]. 인즈의 <Get Away> 뮤직비디오를 살펴보면 이전의 버추얼 휴먼 가수보다 움직임이 훨씬 더 자연스러운 것을 확인할 수 있다[16]. 그 이유는 모델링 과정에서 인간의 근육과 골격의 실제적 움직임 등 인체 해부학의 개념을 적용하여 모델링했기 때문이다. 따라서 인즈는 이전의 버추얼 휴먼 가수보다 비교적 인간의 움직임과 가까운 자연스러운 움직임을 구현할 수 있었다. 이처럼 인즈 제작에 사용된 모션 캡처 기술은 진일보했으나 노래 가사와 입 모양의 싱크가 맞지 않거나 표정이 없는 얼굴이 자연스러운 움직임과 대조되어 오히려 이질감을 느끼게 한다는 점은 여전히 아쉬운 수준이었다.

#### 3) 리밋

가수 리밋은 2001년 리얼스페이스에서 제작한 버추얼 휴먼 가수로 1집 음반 <또 하나의 나>를 발매하며 데뷔했다. 리밋은 표정과 입 모양을 표현하기 위해 리얼스페이스가 자체적으로 개발한 TTFE(언어 입력형; Text To Facial Expression) 실시간 표정 애니메이션 시스템을 사용한 것이 특징이다. 이 시스템은 먼저 인간의 텍스트를 통해 행복, 슬픔, 분노와 같은 기본적인 인간의 감정과 이를 표현하는 표정 이미지를 제공하여 학습하게 한 후 사용자가 텍스트를 입력하면 입력된 정보를 기반으로 인간의 표정이 바뀌면서 영향을 받는 미세한 변화, 근육의 수축과 이완, 각 발음과 단어에 나타나는 연속성을 파악해 실시간으로 표정을 도출해 낼 수 있다. 이로 인해 기존 버추얼 휴먼 가수 제작 시 긴 시간이 소요 됐던 표정 및 언어 구사 표현을 자동화하여 제작 기간을 단축했다[17],[18]. 리밋은 이 시스템을 활용해 자동화된 표정 연기와 노래 가사와 싱크가 맞는 입 모양이 특징인데 이는 1집 음반 타이틀 곡 <독불>의 뮤직비디오를 통해 확인할 수 있다[19]. 앞서 언급한 인즈의 사례와 비교해 보면 인즈의 경우 <Get Away> 뮤직비디오에서 립싱크 장면이 거의 등장하지 않고 노래 가사와 입 모양의 싱크가 맞지 않는다. 표정은 뮤직비디오가 끝날 때까지 무표정에 가깝다. 반면 리밋의 경우 <독불> 뮤직비디오에서 립싱크 장면이 인즈에 비해 많이 등장하는데 노래 가사와 입 모양의 싱크가 자연스럽다. 또한 뮤직비디오 중 클로즈업 구도가 자주 사용된다. 이를 통해 표정이 없었던 인즈에 비해 리밋은 비교적 많은 표정을 구현해 낸 것을 확인할 수 있다. 이런 입 모양과 표정들은 인즈를 비롯한 기존 버추얼 휴먼 가수들 대비 이질감을 감소시켰다. 따라서 버추얼 휴먼 가수 리밋의 제작에는 TTFE 시스템을 활용해 인간의 손을 거치지 않고 표정을 자동으로 만들어 낼 수 있는 시스템이 개발되어 제작 기간을 줄이고 노래와 입 모양의 싱크를 맞게 구현해 냈다는 점에서 의미가 크다.

### 3-3 음성 합성

#### 1) 사이아트

사이아트는 마니주 엔터테인먼트에서 제작한 버추얼 휴먼 가수이다. 외형은 컴퓨터 그래픽으로 만들었으나 목소리는 실제 사람의 육성을 녹음하여 사용했던 기존의 버추얼 휴먼 가수들과 달리 사이아트는 외형뿐만 아니라 목소리도 컴퓨터로 제작되었다. 목소리 제작에는 음성 합성 프로그램인 보컬로이드가 사용되었다. 보컬로이드는 사용자가 가사와 멜로디를 입력하면 프로그램 내에 저장된 목소리가 노래를 만들어 내는 데 이를 두고 사이아트를 제작한 마니주 엔터테인먼트에서는 다음과 같이 언급했다. “사이아트가 작동하는 컴퓨터 시스템에 우리가 가사를 입력할 경우 사람의 음성과 흡사하게 그대로 노래가 재현된다. 멜로디를 바꿔 입력해도 이에 적응해 곧바로 노래가 울려 퍼진다.” 보컬로이드를 활용한 사이아트는 3D 영상을 통해 노래하는 모습과 스피커를 통해 노래를 동시 송출하여 노래를 실시간으로 부르는 것처럼 보이게 한다[20].

보컬로이드는 크립톤 퓨처 미디어가 개발한 음성 데이터베이스 소프트웨어를 아마하에서 자사의 음악 소프트웨어에 사용하면서 출시되었다. 기존의 음성 소프트웨어는 인간의 특정 소리나 발음의 성문을 분석하여 모델링하는 피지컬 모델링 방식이었기 때문에 인위적인 느낌이 강했다. 반면 보컬로이드는 사용자가 멜로디에 가사를 입력하면 소프트웨어에 저장된 음성 데이터를 합성하여 사용자가 원하는 소리를 출력하는 라이브러리 방식이다[21]. 다만 사이아트는 국내에서 개발한 버추얼 휴먼 가수임에도 노래 가사가 영어로 이루어져 있고 발음의 정확도가 다소 낮아 전달력이 떨어지는 모습을 보인다. 그런데도 사이아트는 이전 버추얼 휴먼 가수가 외형은 컴퓨터 그래픽으로 제작되고 목소리는 실제 사람이 참여했던 것과 달리 국내에서 외형과 목소리 모두 온전히 컴퓨터로 제작된 첫 번째 사례라는 점이 의미가 있다.

#### 2) 하츠네 미쿠

하츠네 미쿠는 2007년 일본의 음악 소프트웨어 제작 업체 크립톤 퓨처 미디어에서 개발하여 판매한 보컬로이드이자 버추얼 휴먼 가수이다. 크립톤 퓨처 미디어는 음성 합성 프로그램을 기초로 성우 후지타 사키의 음성 데이터 기반의 보컬로이드2를 출시했고 이 소프트웨어에 일러스트레이터 KEI가 미소녀 이미지의 캐릭터를 입히면서 하츠네 미쿠가 탄생하게 되었다. 사용자가 소프트웨어에 가사와 멜로디를 입력하면 하츠네 미쿠가 노래를 부르는 형식이다[22].

하츠네 미쿠는 이전의 1세대 보컬로이드가 흥행에 실패한 데 반해 2014년 타임지가 선정한 가장 영향력 있는 가상 캐릭터 8위에 선정되기도 하며 오랜 기간 소비되고 있다. 1세대와 2세대 보컬로이드는 기술적인 면에서 차이점이 없는데 하츠네 미쿠가 특이 큰 관심을 얻게 된 이유는 2세대 보컬로이드부터 캐릭터성이 부여되었기 때문이다. 당시 크립톤사는 하츠네 미쿠의 저작권을 제한하지 않고 사용자가 2차 창작물 제

작, 복제, 상연, 상영, 캐릭터 명칭 사용 등 2차 창작에 대한 가이드라인을 명시하며 자율권을 보장했다. 그로 인해 사용자들은 하츠네 미쿠 이전의 캐릭터들이 갖고 있던 2차 창작에 대한 제약에서 벗어나 자유롭게 2차 창작물을 생산했고 하츠네 미쿠에서 아래 그림2와 같이 많은 캐릭터가 파생되게 했다. 사용자들은 하츠네 미쿠를 포함해 파생된 캐릭터들을 일본의 동영상 플랫폼 Nico Nico, Piapro를 통해 공유하기 시작했고 이는 사용자들이 직접 생산자가 되어 하츠네 미쿠라는 캐릭터의 특성을 함께 만들어 간다는 특징을 갖게 했다[23].

NAME	Derived Characters	Characteristics
Hachune Miku		The character featured in the video known as "Ivan Polka," popularly recognized as the "Parody Song," has garnered a high level of awareness and is utilized as a quasi-official character associated with Hatsune Miku.
Akita Nehru		This character is derived from the personification of negative comments about Hatsune Miku, represented by an image of a high school girl. It exhibits a tendency to undermine the popularity of Hatsune Miku.
Yowane Haku		This character is one of the derivatives from Hatsune Miku, and its name carries the meaning of "Spitting Weak Sounds." It was created to embody the feelings of fans who purchased Vocaloid due to Hatsune Miku but couldn't achieve musical completeness.
Yuki Miku		This character was unveiled by Crypton Future Media, the creator of Hatsune Miku, at the 61st Sapporo Snow Festival. It is now widely known as one of Hokkaido's iconic characters.
Hatsune Miku		This character is a male version derived from the gender-swapped concept of Hatsune Miku. It is portrayed as a sibling to Hatsune Miku and serves the role of a relatively scarce male vocaloid.

그림 2. 하츠네 미쿠의 대표적인 파생 캐릭터  
Fig. 2. The representative derivative characters of Hatsune Miku

하츠네 미쿠 이전의 버추얼 휴먼 가수들은 인간의 외형을 완벽히 구현해 내려고 시도해 왔지만, 당시 기술력으로는 그에 부합한 결과물을 만들어내지 못했고 이는 대중들이 버추얼 휴먼 가수에게 불쾌한 골짜기를 느끼게 했다. 그러나 하츠네 미쿠는 외형을 애니메이션 캐릭터화해 인간을 닮은 가수가 아닌 하나의 캐릭터로 보게 했다. 또한 2차 창작물의 생산과 확산으로 많은 캐릭터가 파생되었고 마니아층을 형성하는데 성공했다. 하츠네 미쿠는 콘서트, 광고, 홍보 모델 등 현재까지 다양하게 활동하며 관심 받고 있다. 이런 하츠네 미쿠의



홍행은 이후 제작된 버추얼 휴먼 가수 시유의 탄생에 영향을 준다.

### 3) 시유

시유는 2011년 10월 SBS 아트팩과 아마하의 합작으로 제작된 한국어를 구사하는 최초의 보컬로이드이자 버추얼 휴먼 가수이다. 시유는 음성 합성 엔진 보컬로이드3를 사용해 만들어졌고 가수 글램의 멤버 다희의 음성 데이터를 기반으로 제작되었다. 보컬로이드 기술을 기반으로 하는 시유는 사용자가 소프트웨어에 가사와 멜로디를 입력하면 내재되어 있는 음성 데이터를 합성하여 사람의 목소리로 노래가 출력되고 여기에 안무 영상을 결합하여 내보낸다. 시유는 2012년 7월 22일 'SBS 생방송 인기가요'에서 걸그룹 글램과 함께 무대에 올랐는데 3D 홀로그램 형태로 등장했다[24].

시유는 하즈네 미쿠 홍행의 영향을 받아 인간의 외형이 아닌 캐릭터의 모습으로 제작되었다. 보컬로이드를 활용해 제작된 시유 이전의 버추얼 휴먼 가수 사이아트는 모든 노래를 영어로만 부를 수 있었지만, 시유는 한국어로 된 노래를 부를 수 있다. 그 이유로 기술적인 큰 차이점이 없는 보컬로이드 1세대, 2세대와 달리 보컬로이드가 3세대로 발전하면서 일본어, 영어, 한국어, 스페인어, 중국어 등 다양한 언어 라이브러리를 추가하여 해당 언어로 음성 합성이 가능하게 했다. 다만 어렵게도 실제 사람의 음성 데이터를 기반으로 제작된 시유의 노래를 들어보면 목소리에서 인위적인 느낌이 강하게 느껴진다. 그러나 시유는 보컬로이드를 사용해 처음으로 한국어가 가능한 버추얼 휴먼 가수를 만들었다는 점, 기존 보컬로이드 엔진의 음정이 바뀌는 구간이나 빠른 템포에서 나타나는 음색 변질 문제가 개선되어 사이아트에 비해 비교적 발음이 정확해 가사 전달력을 높였다는 점에서 의미가 있다.

## 3-4 AI 기반

### 1) 로지

로지는 콘텐츠 크리에이티브 기업 싸이더스 스튜디오에서 제작한 버추얼 인플루언서이다. 그는 SNS를 통한 인플루언서 활동과 광고 모델로 활약하다 2022년 2월 22일 싱글 앨범 <Who Am I>를 발매하며 가수로 데뷔했고 발매 9일 만에 뮤직비디오 조회 수 90만 회를 달성하며 큰 관심을 받았다[25].

로지는 딥러닝 기술을 기반으로 제작되었다. 인공지능이 사람의 얼굴을 합성하고 이를 분석해 존재하지 않는 가상의 얼굴을 만들어낸다. 또한 인공지능이 딥러닝을 통해 음성 데이터들을 합성하고 입력된 텍스트를 음성으로 추출하며 여기에 얼굴, 입의 위치, 입 모양, 표정 등을 입힌다. 딥러닝을 통한 이미지 제작 과정에서 더욱 현실감 있는 이미지를 위해 GAN(생성적 대립 신경망; Generative Adversarial Network)이 사용되는데 두 개의 인공지능이 역할을 나누어 경쟁하며 학습하는 것을 뜻한다. 하나는 계속 가상의 이미지를 생성하고 다른 하나는 생성된 이미지를 현실감이 있는지

판단한다. 이 두 인공지능이 이렇게 서로 정보를 주고받으면서 경쟁하여 더욱 현실감 있는 가상의 이미지를 만들어 낸다[26]. 이렇게 이미지가 만들어지면 실제 인간을 모델로 촬영된 영상에 얼굴만을 합성하여 모델링 하는 디지털 3D 더블 기술을 통해 로지가 탄생하게 된다. 디지털 3D 더블 기술을 활용해 버추얼 휴먼 가수를 제작하게 되면 얼굴을 제외한 외형은 실제 사람이기 때문에 현실감 있게 느껴지게 한다. 그 예로 신한 라이프 광고 영상에 등장하는 로지를 보면 춤을 추거나 움직이면 그에 따라 갈비뼈와 복부 근육이 변화함을 알 수 있다[27]. 모션 캡처 기술을 사용해 움직임이 다소 어색했던 이전 버추얼 휴먼 가수들의 한계를 해결했다고 볼 수 있다.

### 2) 한유아

한유아는 엔터테인먼트 기업 스마일게이트와 리얼타임 콘텐츠 솔루션 기업 자이언트스텝이 협업하여 제작했다. 한유아는 SNS를 통해 버추얼 인플루언서로 활동하다 2022년 4월 12일 <I Like That>을 발표하며 버추얼 휴먼 가수로서 데뷔했다. 제작사 스마일게이트는 한유아의 목소리를 구현하기 위해 불특정다수의 음성 데이터를 취합한 후 인공지능으로 합성했다고 밝혔다[28]. 또 다른 제작사 자이언트스텝은 한유아의 3D 모델링 작업에 있어서 실제 사람의 몸에 얼굴을 합성하는 방식, 사진이나 영상에 얼굴을 합성하는 딥페이크 방식 두 가지의 기술을 활용했다고 알렸다. 한유아의 <I Like That> 뮤직비디오는 발표 10일 만에 조회 수 700만 회를 달성했다[29].

한유아는 데뷔곡 <I Like That>을 시작으로 2022년 11월 14일 <나의 외로움이 날 부를 때>, 2023년 3월 6일 <보랏빛 향기>를 발매하며 꾸준히 가수 활동을 이어오고 있다. <I Like That> 뮤직비디오를 보면 한유아가 가상의 세계에서 열차를 타고 현실 세계로 들어오는 연출이 있는데 이는 가상과 현실의 벽을 허물고 버추얼 휴먼 가수가 실제 사람들과 춤을 추고 노래하는 것을 표현했다고 볼 수 있다[30]. 이 뮤직비디오는 실제 인간의 몸에 한유아의 얼굴을 합성하는 방식으로 제작되어 실제 인간과 움직임이 큰 차이가 없는 것을 확인할 수 있다. 이를 통해 버추얼 휴먼 가수의 움직임 표현에 있어 기존의 사례들보다 기술적인 발전이 있었음을 알 수 있다. 또한 <보랏빛 향기>의 뮤직비디오는 한유아가 등장하지 않고 한유아가 직접 그린 이미지를 배경으로 하여 제작되었는데 이는 버추얼 휴먼의 활동 영역이 점차 확대되고 있음을 의미한다[31].

### 3) 이터니티

버추얼 휴먼 아이돌 이터니티는 AI 그래픽 전문기업 펄스나인에서 제작한 걸그룹이다. 이터니티는 지구와 평행한 시간을 갖는 행성 아이아(AIA)에서 11명의 여전사가 사랑데이터를 찾아 나선다는 세계관과 설정을 갖고 총 11명의 버추얼 휴먼으로 제작되었다. 이터니티의 제작사 펄스나인 박지은 대표는 이터니티 제작에 사용된 대표적인 기술이 딥러닝 AI, 딥러

얼 Live라고 소개했다. 먼저 딥리얼 AI 기술은 기존 버추얼 휴먼 가수 제작 시 얼굴 모델링 과정에서 인간이 하나씩 2D 픽셀로 그려서 만드는 작업을 자동화한 기술로 다음 그림3과 같이 사용자가 원하는 데이터를 AI에 제공하면 3초 만에 가상의 인물을 만들어 내는 자동화된 기술이라고 설명했다.



그림 3. 딥리얼 AI를 활용해 가상 인물을 만드는 과정[32]  
 Fig. 3. The process of creating a virtual person using Deep Real AI[32]

제작사는 이 딥리얼 AI 기술을 활용해 20년가량의 K-pop 아이돌 얼굴 데이터를 제공한 후 탄생한 가상의 얼굴들을 일명 ‘AI심쿵챌린지’라는 이상형 월드컵을 진행해 최종 11명을 선정했다고 밝혔다. 그중 이터니티의 멤버 제인은 아리랑 라디오 ‘Super K-pop’에 출연해 생방송 보이는 라디오를 진행한다거나 웹드라마 ‘안녕하세요’에 출연하고 YTN ‘뉴스라이더’에 출연해 첫 생방송 촬영을 진행하기도 했다. 이 과정에서 사용된 기술이 딥리얼 Live기술이다. 딥리얼 Live기술은 만들어진 가상의 얼굴을 실제 사람의 얼굴에 초당 30프레임으로 실시간 페이스 스왑을 하는 것을 말한다. 이터니티는 2021년 3월 1집 앨범 <I’m Real>의 발매를 시작으로 2022년 10월 4집 앨범 <DTDTGMGN>을 발매하며 꾸준한 활동을 이어오고 있다[33]. 이터니티는 AI 기술 기반으로 제작된 타 버추얼 휴먼 가수에 비해 목소리를 사람이 직접 참여한 것으로 알려져 있다. 그러나 딥리얼 AI 기술과 딥리얼 Live를 활용한 실시간 페이스 스왑으로 버추얼 휴먼 가수 최초로 생방송에 참여하고 EBS 드라마 ‘공상가들’에 주연으로 참여하는 등 활동 범위를 넓히고 있다.

4) 메이브

메이브는 카카오 엔터테인먼트와 메타버스 엔터테인먼트의 합작으로 제작된 버추얼 휴먼 아이돌이다. 시우, 제나, 타이라, 마티 총 네 명의 버추얼 휴먼으로 이루어져 있으며 2023년 1월 25일 MBC ‘쇼!음악중심’에서 데뷔곡 <PANDORA>를 공개하며 정식 데뷔했다. 메이브가 공개한 데뷔곡은 멜론 최신곡 차트 5위에 진입하고 뮤직비디오는 공개 3주 만에 1,400만 뷰를 달성하며 큰 관심을 얻었다. 메이브는 현실적이고 정교한 비주얼을 위해 화면에 보일 3차원의 공간상의 빛, 위치, 색상 등의 데이터를 계산하여 화면에 실시간으로 데이터를 시각화해 주는 풀 3D 리얼타임 렌더링 기술이 사용되었다. 또한 디자인 제작에 참여한 제작사 메타버스 엔터테인먼트는 메이브의 이미지 제작에 텍스트를 이미지로 변환해 주는 스테이블 디퓨전 방식과 GAN 방식을 사용했다

고 설명했다[34]. 메이브의 무대는 기성 아이돌과 같이 정확한 군무 중에도 네 명의 멤버가 세부적인 동작이 조금씩 상이한 것을 확인할 수 있다[35]. 이는 방송, 뮤직비디오 등에서 자연스러운 움직임을 표현하기 위해 각각 멤버의 움직임을 실제 댄서들의 움직임을 모션 캡처하여 표현했기 때문이다. 이러한 메이브의 각 멤버의 디테일한 움직임 표현과 정확한 군무, 풀 3D 리얼타임 렌더링 기술을 활용하여 높은 완성도의 풀 3D 버추얼 아이돌을 구현했다는 점에서 버추얼 휴먼 가수 제작 기술이 발전했다고 볼 수 있다.

IV. 결론

본 논문은 버추얼 휴먼 가수의 기술적인 발전에 관한 연구이다. 1990년대 IT산업의 발전과 PC의 보급으로 등장했던 버추얼 휴먼 가수는 당시 기술력의 한계로 큰 성과를 얻지 못했다. 그러나 현재까지 버추얼 휴먼 가수의 제작은 지속되고 있고 4차 산업혁명의 영향으로 기술의 발전이 가속화되고 있다. 또한 메타버스의 등장과 AI 기술의 급성장으로 버추얼 휴먼 가수의 대중화 가능성이 재조명 되는 시점이다. 따라서 본 연구는 과거부터 현재에 이르기까지 버추얼 휴먼 가수의 기술적인 발전에 대해 알아보았다.

본 연구에서 제시한 버추얼 휴먼 가수의 유형은 총 네 가지로 제작 기술 유형에 따라 버추얼 휴먼 가수의 시초, 모션 캡처 기술의 적용, 음성 합성 기술의 적용, AI기술 기반 제작으로 분류했다. 각 유형별로 선정한 버추얼 휴먼 가수의 사례는 다음과 같다. 첫째, 버추얼 휴먼 가수의 시초 유형에는 세계 최초의 버추얼 휴먼 가수인 다테 료코와 국내 최초의 버추얼 휴먼 가수 아담을 선정했다. 둘째, 모션 캡처 기술이 적용된 유형에는 사이다, 인체 해부학의 개념이 적용된 인즈, TTFE 기술이 추가된 리밋을 선정했다. 셋째, 음성 합성 기술이 적용된 유형에는 국내에서 처음으로 외형과 음성 모두 컴퓨터로 제작된 사이아트, 현재까지 활발히 활동중인 하츠네 미쿠, 국내 처음으로 한국어 구현이 가능한 시유를 선정했다. 넷째, AI기술을 기반으로 제작된 유형에는 딥러닝 기술을 기반으로 한 로지와 딥러닝, 딥페이크 기술로 제작된 한유아, 딥리얼 AI 기술과 딥리얼 Live를 활용한 페이스 스왑이 사용된 이터니티 마지막으로 모델링 제작 과정에서 스테이블 디퓨전과 GAN 방식이 활용된 메이브를 선정했다.

**표 2.** 버추얼 휴먼 가수 제작에 사용된 기술  
**Table 2.** Technology used in the production of virtual human singers

Virtual Human Singer	Face modeling	Voice	Movements
Date Kyoko	3D Modeling	Human	Motion Capture
Adam	3D Modeling	Human	Motion Capture
Cyda	3D Modeling	Human	Motion Capture
Intz	3D Modeling	Human	Motion Capture
Limit	3D Modeling, TTFE	Human	Motion Capture
Sciart	3D Modeling	Vocaloid	Motion Capture
Hatsune Miku	3D Modeling	Vocaloid	Motion Capture
SeeU	3D Modeling	Vocaloid	Motion Capture
Rozy	Deep learning	Deep learning	Digital Double
Han Yua	Deep learning	Deep learning	Digital Double, Deep Fake
Eternity	Deep Real	Human	Deep Real Live Face Swap
MAVE:	Full 3D, Stable Diffusion, GAN	Human	Motion Capture

위의 표 2는 본 논문에서 제시한 버추얼 휴먼 가수의 제작 과정에서 제작 기본 요소인 얼굴 모델링, 목소리, 움직임 제작에 사용된 기술을 정리한 표이다. 초기 모델인 다테 쿄코와 아담의 사례에서 리밋의 사례까지 이들의 얼굴 모델링, 목소리, 움직임 제작에 있어 사용된 기술은 큰 차이점이 없다. 그러나 인츠의 사례에서 외형 모델링에 인체 해부학의 개념이 적용되면서 이전 사례들보다 비교적 사람의 움직임에 가깝게 구현되었다. 하지만 노래 가사와 입 모양의 싱크가 맞지 않거나 무표정에 가까운 얼굴 등 아쉬운 점이 있었다. 이후 리밋의 사례에 TTFE 기술이 적용되면서 버추얼 휴먼 가수의 입 모양, 표정 제작에 소요되는 시간이 감소하고 노래 가사와 입 모양의 싱크를 맞게 하여 이전 인츠의 사례의 한계점인 얼굴 표정, 입 모양, 노래와의 싱크 등을 보완했음을 알 수 있다.

초기 버추얼 휴먼 가수들의 외형은 컴퓨터 그래픽스 기술로 제작되었지만 노래, 인터뷰, 광고 등에서 쓰이는 목소리는 실제 사람이 직접 참여하는 방식으로 구현해 왔다. 그러나 사이아트, 하츠네 미쿠, 시유의 사례부터는 음성 합성 기술인 보컬로이드를 사용하여 목소리까지 컴퓨터로 구현해 낼 수 있게 되었다. 다만 시유의 사례까지 버추얼 휴먼 가수의 얼굴 모델링에 있어서 사람이 직접 디자인에 개입해야 하는 번거로움이 있었다.

2020년대에 이르러서 로지와 한유아의 사례를 보면 딥러

닝 기술을 통해 인공지능이 버추얼 휴먼 가수의 얼굴 모델링과 목소리 구현까지 스스로 학습을 통해 만들어 냈다. 또한 이전의 버추얼 휴먼 가수의 움직임이 모션 캡처 기술을 중심으로 만들어졌다면 로지와 한유아는 실제 사람의 몸에 얼굴만을 합성하는 디지털 더블 기술과 딥페이크 기술이 사용된 것을 확인할 수 있다. 한편 이터너티는 딥리얼, 딥리얼 Live 기술을 활용해 생방송과 드라마에 출연해 실제 사람들과 호흡을 맞췄고 메이브는 풀 3D 리얼타임 렌더링 기술을 통해 음악 방송 무대에 올랐다. 이 둘의 목소리는 같은 AI 기술 기반으로 제작된 로지와 한유아의 사례와 달리 실제 사람이 참여한 것을 알 수 있다. 이를 통해 기존 버추얼 휴먼 가수의 얼굴 모델링, 움직임, 목소리 구현, 표정의 어색함 등에서 나타났던 기술적 한계점이 시대적 흐름에 따라 모션 캡처, 음성 합성, AI 기반 딥러닝, 딥리얼, 딥리얼 Live 등의 기술 적용으로 보완되고 변화한 것을 확인할 수 있었다.

## V. 논의

본 논문에서는 버추얼 휴먼 형태의 가수가 처음 등장한 이후 현재까지 제작된 버추얼 휴먼 가수 중 본 연구에서 제시한 제작 기술에 따른 네 가지 유형에 부합하는 사례의 분석을 통해 기술적 발전에 대해 연구하였다. 그 결과 버추얼 휴먼 가수 제작에 있어서 기술적인 발전과 시도는 끊임없이 지속되어 왔으며 현재도 계속해서 발전하고 있음을 확인할 수 있었다. 과거 한계점이 명확했던 버추얼 휴먼 가수의 제작 기술이 현재 인공지능 딥러닝을 적용한 기술들을 활용하여 대중들이 버추얼 휴먼 가수가 실제 사람이라고 착각할 만큼의 수준으로 발전했다. 또한 음성 데이터를 기반으로 이미 사망한 사람의 목소리를 비슷하게 구현하거나 빅히트 뮤직과 하이브에서 론칭한 아티스트 ‘미드나트(MIDNATT)’의 사례처럼 남성 목소리를 기반으로 여성 목소리를 새롭게 만들어 내거나 녹음 후 다양한 언어의 발음을 정확하게 교정할 수도 있다[36].

이처럼 버추얼 휴먼 산업 시장이 점차 확대되고 있는 시점에서 딥러닝을 적용한 딥보이스와 딥페이크, 모창 AI 기술, 보이스 디자인, 다국어 발음 교정과 같은 새로운 기술들의 등장은 앞으로의 버추얼 휴먼 가수가 딥러닝을 활용한 학습을 통해 숏소리, 바이브레이션을 포함해 인간만이 가능했던 감정 표현을 가능하게 하고 언어의 제약에서 자유롭게 여러 국가로 활동 범위를 넓히리라 전망한다.

실제로 2021년 SBS에서 방송한 ‘SBS 신년특집 세기의 대결 AI vs 인간’에서 가수의 음성 데이터를 입력하면 그 가수의 목소리로 노래를 부르는 이른바 ‘모창 AI’가 소개된 바 있다. 이 모창 AI는 가수의 호흡과 바이브레이션, 숏소리까지 학습하여 똑같이 구현이 가능하다. 음의 높낮이를 조절하는 목, 발음을 가능하게 하는 입 이 두 가지를 사용해 노래하는 인간의 방식을 모창 AI의 학습 알고리즘에 적용해 음정과 발음을 분



리하여 학습한다. 모창 AI가 학습하는데 걸리는 시간은 약 10분 정도이며 학습 방법은 앞서 로지의 사례에서 설명한 GAN이 사용된다. 먼저 AI가 학습할 수 있는 가창 데이터를 제공한다. 가창 데이터가 AI에 전달되어 학습이 완료된 후 다음으로 노래 선곡만 해주면 보컬AI와 이를 가르치는 AI로 나뉘어 학습하는데 학습이 반복될수록 완성도가 더 높아진다. 이날 방송에서는 가수 아이유(IU)가 부르는 박효신의 <야생화>와故 김광석이가 부르는 김범수의 <보고싶다>를 공개했다[37].

또한 최근 Youtube를 포함한 인터넷 포털에서 ‘AI 커버송’이라는 제목을 한 영상이 대중들에게 큰 관심을 받고 있다. 원곡 가수가 아닌 다른 가수의 목소리, 이미 사망한 가수의 목소리를 통해 노래를 커버하는 영상이다. 영상은 걸그룹 블랙핑크 멤버 지수의 솔로 곡 <꽃>을 같은 그룹 멤버 제니의 목소리로, 걸그룹 뉴진스의 노래 <Ditto>를 Ariana Grande의 목소리로 재연한다. 심지어 걸그룹 FIFTY FIFTY의 <Cupid>는 이미 사망한 Michael Jackson의 목소리로 구현하는 등 이외에도 수많은 버전의 커버곡들이 존재한다. 이는 최근 AI기술이 발전하면서 ‘AI 딥보이스’를 사용해 만들어진 영상들이다[38]. 딥보이스는 앞서 한유아의 사례에서 설명한 딥페이크 기술과 비슷한 개념이다. 딥페이크는 영상 합성이라면 딥보이스는 음성을 합성하는 기술이다. 딥러닝 기술을 기반으로 음성 데이터를 수집하고 분석해 더욱 정교한 목소리를 구현해 낼 수 있다. 딥보이스를 활용한 AI 커버송은 사용자가 원하는 노래를 이미 사망한 가수의 목소리나 듣고 싶은 가수의 목소리로 들어볼 수 있다는 장점이 있다. 그러나 AI 커버송 제작 과정에서 타인의 목소리를 무단으로 사용해 목소리를 불법 복제하거나 커버에 사용되는 원곡의 저작권 문제 등 양면성이 존재한다. 최근 일반인도 쉽게 접할 수 있는 딥페이크, 딥보이스를 활용할 수 있는 소프트웨어가 늘어나고 있다는 점 역시 저작권 문제 발생을 증가시킬 가능성이 농후하다. 이를 통해 불법 복제된 음성 및 영상과 AI 커버송을 만드는 방법에 관한 내용들이 Youtube와 TikTok 등 각종 플랫폼을 통해 공유되고 높은 조회수를 기록하고 있기 때문이다. 이에 따라 특정인의 외모와 음성의 불법 도용으로 인한 초상권 침해, 저작권법 위반 등의 문제 증가가 우려된다.

버추얼 휴먼 가수 제작 기술이 발전되고 제작 접근성이 향상됨으로 인해 향후 다양한 버추얼 휴먼 가수의 등장이 예상된다. 이는 대중에게 다양한 콘텐츠 제공할 수 있으며 기존에 없던 새로운 시장을 개척할 기회가 될 수도 있다. 하지만 버추얼 휴먼 가수 제작에 사용되는 여러 기술을 전문 제작사뿐만 아니라 일반인도 손쉽게 이용 가능해진 만큼 특정인의 초상권, 저작권 침해 등 기술의 오용에 대한 대책 역시 마련되어야 할 것이다.

## 참고문헌

- [1] H. R. Yun, “A Study on the Virtual Idol Activation Strategy Using Game Characters: Focused on K/DA,” Master’s Thesis, The Graduate School of Culture and Contents, Hanyang University, Ansan, August 2022.
- [2] Editorial Department (Editor), “The Evolution of Virtual Influencer,” *Marketing*, Vol. 55, No. 9, pp. 16-27, September 2021.
- [3] D. W. Kim, “The Evolution of Cyber Singers from a Coevolutionary Perspective of Humans and Machines,” *Cartoon&Animation Studies*, No. 39, pp. 261-295, 2015. <https://doi.org/10.7230/KOSCAS.2015.39.261>
- [4] D. C. Park and D. E. Lee, “A Study on the Storytelling Strategy for Virtual Human’s Uncanny Valley Offset,” *The Journal of Korea Game Society*, Vol. 22, No. 6, pp. 43-56, December 2022. <http://dx.doi.org/10.7583/JKGS.2022.22.6.43>
- [5] J. W. Kim and H. W. Cha, “Effect of Virtual Human Live Streaming Motivation and Virtuality Recognition, and Flow on Continuous Use Intention and Intention of Interaction,” *The Korean Journal of Advertising*, Vol. 33, No. 4, pp. 89-123, May 2022. <https://doi.org/10.14377/KJA.2022.5.31.89>
- [6] Korea Electronics Association, “Computer Penetration in Korea in 2001,” *Journal of Korean Electronics Association*, Vol. 22, No. 6, pp. 1-10, June 2002.
- [7] J. A. Kim, “A Study on the Limits of Teen Idols-dominated Korean Popular Music Market: Focused on the Influences on the Birth of Teen Idols,” Master’s Thesis, The Graduate School of Journalism, Korea University, Seoul, February 2011.
- [8] Naver Knowledge Encyclopedia of Counseling. Generation X [Internet]. Available: <https://terms.naver.com/entry.naver?docId=5677628&cid=62841&categoryId=62841>
- [9] Media Today. [Seong Sang-min’s Cultural Reversal] From ‘Adam’ to ‘Another World Idol’, the Virtual Worldview Evolves. [Internet] Available: <http://www.mediatoday.co.kr/news/articleView.html?idxno=303549>
- [10] S. R. Park, “A Study on the Virtual Idol in Metabus Environment: Focusing on the ‘Isegye Idol’,” Master’s Thesis, The Graduate School of Culture and Technology, Sangmyung University, Seoul, August 2022.
- [11] N. S. Park, (Flower of VFX, for FX) *Houdini*, Gangwon: Hillbook, pp. 14-26, 2017.
- [12] D. S. Kim, “A Review of the Current Status and the

- Prospect of the Cyber Character,” Master’s Thesis. The Graduate School of Chosun University, Gwangju, February 2000.
- [13] THE DAILYPOST. Sad Retirement of Cyber Singer ‘Adam’ “Zero Possibility of Comeback?” [Internet]. Available:  
<https://www.thedailypost.kr/news/articleView.html?idxno=88399>
- [14] Bugs. I Don't Like the Truth of Cider's Song (The Lyrics) [Internet]. Available: <https://music.bugs.co.kr/track/83724>
- [15] ZDNET Korea. INTZ.COM, Introduces Cyber Dance Group 「INTZ」 [Internet]. Available:  
<https://zdnet.co.kr/view/?no=0000010030341>
- [16] Youtube. INTZ "Get Away" music video [Internet]. Available:  
<https://www.youtube.com/watch?v=I56XY8Fz4NU>
- [17] OhmyNews. Cyber Character Generation Change Declaration [Internet]. Available:  
[https://m.ohmynews.com/NWS\\_Web/Mobile/at\\_pg.aspx?CNTN\\_CD=A0000047467#cb](https://m.ohmynews.com/NWS_Web/Mobile/at_pg.aspx?CNTN_CD=A0000047467#cb)
- [18] Youtube. From Text to Facial Expression [Internet]. Available:  
<https://www.youtube.com/watch?v=YJqqJ5pMzw>
- [19] Youtube. The Music Video for "Dok Bull" by LIMIT [Internet]. Available:  
<https://www.youtube.com/watch?v=IhOuH0NWkHo>
- [20] Sports Kyunghyang. ‘100% Cyber Singer’ Sciart, Unveiled on the 13th [Internet]. Available:  
[https://sports.khan.co.kr/entertainment/sk\\_index.html?cat=view&art\\_id=200709042144383&sec\\_id=540301&pt=nv](https://sports.khan.co.kr/entertainment/sk_index.html?cat=view&art_id=200709042144383&sec_id=540301&pt=nv)
- [21] E. H. Cho, “A Consumption of the Idol as a Sign in the Digital Era- A Study on a Virtual Idol, ‘Hatsune Miku’-,” *Global Cultural Contents*, Vol. 0, No. 29, pp. 203-218, August 2017.
- [22] X. W. Dou and J. H. Yang, “Innovative Business Model from the Perspective of Digital Cultural Industry-Centered on the Virtual Idol ‘Hatsune Miku’,” *The Journal of the Convergence on Culture Technology*, Vol. 7, No. 4, pp. 823-831, November 2021.  
<http://dx.doi.org/10.17703/JCCT.2021.7.4.823>
- [23] D. H. Kim, “A Study for Characteristic Analysis of Vocaloid Virtual Character -Focused on Idol Character Hatsune Miku-,” *The Korea Society of Illustration Research*, Vol. 19, No. 56, pp. 27-36, September 2018.  
<http://dx.doi.org/10.37379/jksir.2018.56.3>
- [24] Sports Kyunghyang. Do you Know Vocaloid? Korea’s First Vocaloid ‘SecU’ Broadcast Debut! [Internet]. Available:  
[http://sports.khan.co.kr/entertainment/sk\\_index.html?cat=view&art\\_id=201207221742573&sec\\_id=540301&pt=nv](http://sports.khan.co.kr/entertainment/sk_index.html?cat=view&art_id=201207221742573&sec_id=540301&pt=nv)
- [25] Dong-A Ilbo. 24 Years Since ‘Adam’...Cyber Singer ‘Rozy-Han Yua’ Heats Up the Music Industry [Internet]. Available:  
<https://www.donga.com/news/article/all/20220304/112147383/1>
- [26] YTN. [Head-to-head match] How Was the Hot Virtual Human ‘Rozy’ Created These Days? [Internet]. Available:  
[https://www.ytn.co.kr/\\_ln/0101\\_202107091949461873](https://www.ytn.co.kr/_ln/0101_202107091949461873)
- [27] Youtube. Adding Astonishment to Life at Shinhan Life [Internet]. Available:  
[https://www.youtube.com/watch?v=y8v\\_UXdBQtw](https://www.youtube.com/watch?v=y8v_UXdBQtw)
- [28] TECH M. Han Yoo-ah’s ‘I Like That’ Is Finally Out...Will It Surpass Cyber Singer ‘Adam’? [Internet]. Available:  
<https://www.techm.kr/news/articleView.html?idxno=96301>
- [29] The JoongAng. The Reason Why ‘Game King’ Kwon Hyuk-bin Fell for Virtual Human ‘Han Yoo-ah’ [Internet]. Available: <https://www.joongang.co.kr/article/25085785>
- [30] Youtube. The Music Video for Han Yoo Ah's Song "I Like That" [Internet]. Available:  
[https://www.youtube.com/watch?v=oyFySZK\\_z\\_U](https://www.youtube.com/watch?v=oyFySZK_z_U)
- [31] Youtube. Han Yoo Ah's Lyric Video for "Scent of Violet" [Internet]. Available:  
<https://www.youtube.com/watch?v=s3AljY-IWv8>
- [32] Youtube. [TechM Conference] "Virtual Girl Group Project Eternity Made by Your Pick," CEO Park Ji-eun of Pulse9 [Internet]. Available:  
<https://www.youtube.com/watch?v=xOFxomaaaFo>
- [33] Daily Consumer Economic News. [ICT Changers] Park Ji-eun, CEO of Pulse 9 “Eternity is 11 Female Warriors Looking for Love Data” [Internet]. Available:  
<http://www.dailycnc.com/news/articleView.html?idxno=212858>
- [34] The JoongAng. After 1 AI Girl Group, 10 Substitutes... Gold Spoon vs. Dirt Spoon Too [Internet]. Available:  
<https://www.joongang.co.kr/article/25141430#home>
- [35] Youtube. MAVE: (Mave) - PANDORA | Show! Music Core | MBC230218 Broadcasting [Internet]. Available:  
<https://www.youtube.com/watch?v=IDFQdA2uroM>
- [36] etoday. Hybe Newcomer Midnatt Revealed...‘Voice Design’ Realizes ‘Freedom’ of Male and Female Voices [Internet]. Available:  
<https://www.etoday.co.kr/news/view/2249172>
- [37] Youtube. IU's Rendition of Park Hyo Shin's 'Wild Flower'! The Secret behind the AI Technology in the Mock Performance? | AI vs Humans (SBS Broadcast) [Internet]. Available:

<https://www.youtube.com/watch?v=7JR2ehHia04>

[38] NEWSIS. Bruno Mars on 'Hype Boy' and Cupid'?  
[Internet]. Available:  
[https://newsis.com/view/?id=NISX20230512\\_0002301195  
&cID=50101&pID=50100](https://newsis.com/view/?id=NISX20230512_0002301195&cID=50101&pID=50100)



**이용수(Yong-Su Lee)**

2021년~현 재 : 경희대학교 대학원 (음악학석사과정)

※ 관심분야 : 인공지능, 딥러닝, 버추얼 휴먼, 사이버 가수 등



**이철희(Chul-Hee Lee)**

2009년 : 경희대학교 Post Modern음악전공 (음악학사)  
2013년 : 경희대학교 아트·퓨전디자인 대학원 (음악학석사)  
2020년 : 경희대학교 대학원 (응용예술학박사)

2007년~2021년: 마인드오픈너 음악감독

2021년~현 재: 경희대학교 포스트모던음악학과 조교수

※ 관심분야 : 컴퓨터음악(MIDI), 작곡(Composition), 사운드디자인(Sound design) 등