

파운드 푸티지 장르가 결합된 StyleGAN 이미지 기반 영상 제작

박지혜¹ · 송해민¹ · 박진완^{2*}

¹중앙대학교 첨단영상대학원 영상학과 예술공학 석사과정

^{2*}중앙대학교 첨단영상대학원 영상학과 예술공학 교수

The Production of StyleGAN Image-based Video Combined the Found Footage Genre

Ji-Hye Park¹ · Hae-Min Song¹ · Jin-Wan Park^{2*}

¹Master's Course, The Graduate School of Advanced Imaging Science, Multimedia and Film, Chung-Ang University, Seoul 06974, Korea

^{2*}Professor, The Graduate School of Advanced Imaging Science, Multimedia and Film, Chung-Ang University, Seoul 06974, Korea

[요약]

본 연구는 파운드 푸티지 장르가 결합된 StyleGAN 이미지 기반 영상 제작 절차를 기술한 논문이다. 코로나19 이후 마스크 착용이 일상화되면서, 사람들은 '코로나19 이전과 이후'라는 변화로부터 야기되는 이질감을 인식하기 시작하였다. 이 같은 변화를 표현하기 위해 인공지능 기술을 기반으로 상반된 두 상황의 인물들의 대화에서 의도적인 '소통의 격차'를 보였고, 이질감을 극대화시킬 수 있는 실험영화 장르인 파운드 푸티지를 결합하여 뉴노멀 시대 우리의 모습을 재조명하는 작품을 제작하였다. 또한 화상회의 플랫폼 '줌'에서 코로나19 이전과 이후로 구분된 두 인물의 대화를 보여 관람자에게 우리의 '변화'를 객관적으로 진단해볼 수 있는 계기를 마련하였다. 결과적으로 인공지능을 활용한 실험적인 작품을 보인 점에서 예술공학적 시사점을 가진다.

[Abstract]

This study describes the process of the video production based on StyleGAN image combined with the found footage. We informed by the situation in which wearing a mask has become a routine and it symbolizes the change before and after COVID-19, cause of difference. In order to express this change, 'Communication Gap' was expressed through dialogues between characters in two opposite situations based on Artificial Intelligence, and maximize the sense of difference, we produced a work that re-illuminates the appearance by combining the genre. In addition, we can objectively examine our 'change' by showing the communication barriers between the two beings divided on the 'Zoom meeting', and it provided the opportunity. Finally the study has produced implication of Art and Technology, in that shows the work using Artificial Intelligence.

색인어 : 인공지능, 딥러닝, 영화, 파운드 푸티지, 융합

Keyword : Artificial Intelligence, Deep Learning, Movie, Found Footage, Convergence

<http://dx.doi.org/10.9728/dcs.2021.22.12.1941>



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 13 October 2021; **Revised** 10 November 2021

Accepted 10 November 2021

***Corresponding Author, Jin-Wan Park**

Tel: +82-2-850-5710

E-mail: jinpark@cau.ac.kr

I. 서론

최근 코로나19(COVID-19; 신종 코로나 바이러스 감염증) 이후 뉴노멀(New Normal) 시대가 부상하게 되며, 일상의 대부분이 새롭게 변화되기 시작하였다. 대표적으로 우리는 바이러스에 대한 우려로 마스크 착용이 보편화되어 익숙해진 광경을 보이고 있으며, 이를 통해 ‘마스크’라는 존재는 코로나19 이전과 대비되는 변화를 가지적으로 직감하게 되는 매개체로서 현재의 모습을 상징하는 존재가 되었다. 이 같은 현상에 착안하여 바라보았을 때, 코로나19 이전의 미디어에서 다수 인물들이 마스크를 착용하지 않은 채로 소통하는 장면은 코로나19 이후의 우리에게 이질감을 선사하게 된다. 반대로, 코로나19 이전의 시점에서 코로나19 이후에 해당하는 현재 마스크를 착용한 인물들이 생경하게 느껴질 것으로 판단된다.

이러한 사회적 배경을 포착한 본 연구는 인공지능 기술인 GAN 기반 이미지로 정의되는 ‘인공지능이 바라본’ 우리의 모습을 이야기하고자 한다. 이 과정에서 제작한 작품인 <Face New Normal>을 통해 코로나19 이후 우리의 모습을 재조명하는 계기를 마련함으로써 직면하게 된 시대의 현실과 변화를 가지적으로 표현해보고자 한다. 이때 표현되는 두 장면의 교차적인 대비는 시대의 변화를 직감할 수 있는 장치로서의 역할을 수행하게 된다. 시간이 다른 두 시퀀스(sequence)의 배치를 통해 불확실성을 극대화하기 위해 실험영화 장르인 파운드 푸티지 기법을 결합하는 시도를 보였다.

본 연구에서는 작품 <Face New Normal>을 제작 과정에서 온라인에서 수집한 코로나19를 인지하지 못하는 코로나19 이전 인물들의 일상적인 회의 모습을 파운드 푸티지 기법이 결합되어 콜라주하는 연출을 보였다. 이후 인공지능 기술인 딥러닝 기반의 StyleGAN을 통해 생성한 마스크를 착용한 코로나19 이후 인물들의 이미지 데이터를 가상 얼굴 이미지로 구현하여 이를 비대면 화상 회의 플랫폼인 줌 화면에 적용하였다. 결과적으로 이들이 의미 없는 대화를 나누는 장면을 의도적으로 연출함으로써 소통의 장애를 보였으며, 감정적 요소로서 답답함과 같은 예술적 충격을 선사하는 것을 고안하였다.

II. 딥러닝 기반의 StyleGAN 기술

2-1 인공지능과 딥러닝, StyleGAN

인공지능(AI; Artificial Intelligence)은 인간이 할 수 있는 사고와 학습을 기계를 통해 실현한 기술이다. 또한, 컴퓨터 정보기술의 한 분야로서 기계가 인간이 보이는 생각과 판단을 유사하게 할 수 있도록 고안된 것이다[1]. 인공지능 기술은 1940년대를 시작으로 1950년 영국의 컴퓨터 과학자 앨런 튜링(Alan Turing)이 튜링 테스트(Turing test)의 튜링기계(Turing’s Machine)를 통해 인공지능 기술을 탑재한 컴퓨터를 판별하기 위한 시도로부터 기반을 마련하게 되었다[2].

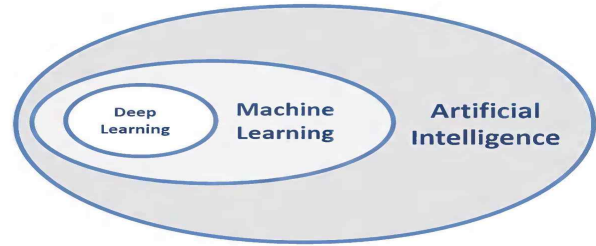


그림 1. 인공지능과 머신러닝, 딥러닝의 관계 시각화 [3]
Fig. 1. The Position of Machine Learning and Deep Learning in AI

이러한 인공지능을 구성하는 요소로는 [그림 1]에서 시각화된 바와 같이 머신러닝(Machine Learning)과 딥러닝(Deep Learning)이 있다. 머신러닝은 데이터 기반 예측을 위한 도구로서 가장 효과적인 인공지능 기술이며, 이에 속한 여러 가지 알고리즘 중 사람이 생각하는 인지 처리 과정을 유사하게 처리하는 방식을 차용하여 최대의 효과를 내는 것이 딥러닝이다[1].

언급한 딥러닝의 특징은 우리 생활과 산업에서 활발히 사용되어 효과를 발휘하는 중이다. 이에 해당하는 대표적인 예로 생성적 적대 신경망(GAN; Generative Adversarial Network)이 있다. GAN은 이안 굿펠로(Ian Goodfellow)에 의해 제안되었으며[4], 딥러닝을 기반으로 한 가상 이미지를 생성할 수 있는 알고리즘이다[1]. 이와 같은 이미지 처리기술로서 사람과 동물 같은 객체의 얼굴 이미지를 현실과 유사하게 생성할 수 있는 특징을 가진다. 이 기술은 크게 이미지 생성(Generative)과 감별(Discriminative) 모델로 구분되며, 용어에서 시사하는 바와 같이 ‘생성’이라는 공통된 목적을 달성하기 위해 경쟁 및 상호보완적인 과정을 거쳐게 된다. 예를 들어 감별 모델이 구별하기 어려운 진짜와 가짜 이미지를 생성 모델이 찾는 방식으로 협업하여 이미지를 생성하게 되는 것이다.

그러나 GAN의 한계로 모델 학습의 불안정성 문제가 대두되면서[4], [6], 이러한 한계를 극복하기 위한 모델이 제안되기 시작하였다. 특히 2018년 엔비디아(NVIDIA) 연구원에 의해 StyleGAN(Style-Generative Adversarial Network)이 새롭게 제안된 것이 대표적이다[5], [7].



그림 2. StyleGAN으로 생성한 가상 이미지 [8]
Fig. 2. The Virtual Images generated by StyleGAN

위의 [그림 2]는 StyleGAN으로 존재하지 않는 가상 이미지를 생성한 결과이다. 사람의 성별과 연령을 종합하여 같은 행으로부터 서로의 머리와 피부색을 추출하여 도출한 것이다 [6]. 해당 모델은 StyleGAN1과 이를 보완한 StyleGAN2로 나누어지는데, 후자의 StyleGAN2는 물방울 형태의 무늬가 발견되는 한계가 대두되었던 StyleGAN1을 개선한 형태를 보인다[6], [7], [9].

이러한 특성에 주목하여 현재에는 새로운 이미지 생성물의 상업적인 활용 가능성에 관심이 이어지고 있는 상황이다. 특히 예술작품의 경우 인공지능을 활용한 이미지 기술이 적용된 영상 및 영화 제작 작업 시도가 크게 증가하고 있는 추세이다.

2-2 인공지능 기반 예술작품 제작 사례 분석

본 연구에서는 인공지능 기술을 이용한 예술작품 제작 사례를 총 2가지로 분석하였다.

1) <컴퓨터는 인간을 어떻게 상상하는가?>

첫 번째로, 후안 마티노 무라(João Martinho Moura)의 ‘<컴퓨터는 인간을 어떻게 상상하는가?(How Computers Imagine Humans)>’에서는 컴퓨터 비주얼 노이즈와 인공지능 기반 얼굴 탐지 시스템을 통해 컴퓨터가 보이는 “상상된” 얼굴 이미지를 연출하는 새로운 상호작용을 보였다. 구현된 [그림 3]은 인간이 인간을 탐지하기 위해 기계에 입력한 데이터, 즉 ‘지식’을 시각화한 것을 보인다. 두 대의 컴퓨터를 통해 보이는 대비와 이에 따른 두 장면의 연출은 관람자로 하여금 ‘기계는 사고할 수 있는가’에 대한 인식을 제고함과 더불어 우리 사회에 미치는 영향을 탐색해 볼 수 있는 계기를 마련하였다[그림 3].

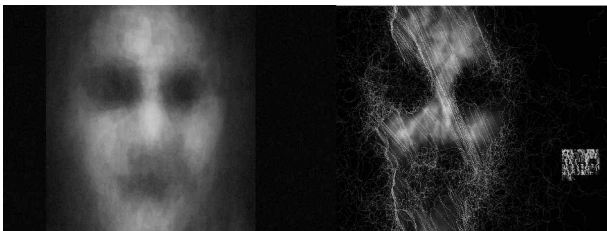


그림 3. <컴퓨터는 인간을 어떻게 상상하는가?> [10]
Fig. 3. <How Computers Imagine Humans?>

2) <The Giver of Names>

두 번째로, <The Giver of Names>라는 [그림 4]의 작품에서는 ‘생각하는 기계’라는 상황을 설정하여 컴퓨터 시스템이 물건에 이름을 부여하는 방식으로 진행된다. 예를 들어, 왼쪽 [그림 4]에 있는 빈 받침대에 관람자가 물건을 올려놓으면 비디오카메라가 관찰하여 외곽선, 색과 질감을 분석해 이미지를 처리한다. 그 결과, 생각하는 기계와 대리자를 통해 소통하는 기계를 보여주게 된다.

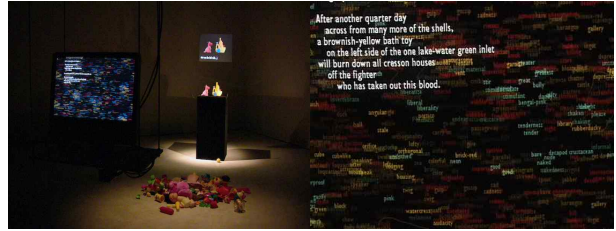


그림 4. <The Giver of Names> 받침대(왼쪽) 그 결과(오른쪽) [11]
Fig. 4. <The Giver of Names> Pedestal(left) and its result(right)

이 과정을 통해 관람자들은 기계의 시선으로 사물을 바라보고 이에 대한 판단을 거치게 되면서, 이들이 기계에 투영하였던 기존의 시각을 반추할 수 있는 계기를 마련하게 된다. 이에 따라 해당 작품에서 표현된 ‘기계는 생각하는 방식’은 인공지능 예술 작품에 대한 새로운 가능성을 시사하며, 그 이후의 가능성 또한 조망해볼 수 있는 계기를 마련하였다[11], [12].

본 연구에서는 분석한 예술작품에서 공통적으로 부여되는 감정을 인공지능이라는 존재에 따른 인간의 두려움으로 조명하였다. 이와 같은 이질적인 감정은 인공지능 혁신에 대한 불안에 직결된 감정이라고 할 수 있다. 선행 연구에 따르면 인공지능이 인간사회에 참여하게 되었을 때 인간에게 혼란스러움을 제공할 수 있다는 것이 이를 증명한다[13]. 이는 ‘강한 인공지능’을 지칭하면서 인간과 유사하게 자의식을 갖고 감정을 느끼며 추론을 할 수 있는 능력을 통해 인간에게 두려움이라는 감정을 불러일으키게 되는 요인이 된다[14]. 따라서 본 연구에서는 이러한 특성에 착안하여 관련된 장르적 특성을 결합한 작품을 제작하고자 하였다.

III. 파운드 푸티지

3-1 실험영화 장르로서 파운드 푸티지

1) 파운드 푸티지의 정의 및 특징

파운드 푸티지(Found Footage)는 페이크 다큐멘터리(Fake Documentary)의 한 종류이다. 이는 실험영화 제작자들이 보편적인 촬영 도구인 ‘카메라(Camera)’를 생략하는 방식으로 제작하는 영상에 대한 호기심에 촉발되어, 점차 하나의 장르로서 자리 잡게 되었다. 그 결과 예술가로서의 창조적 자유와 권리를 소유할 수 있는 영역으로 확장되기 시작하였다.

용어에서 시사하는 바와 같이 우연히 ‘발견된’ 저작권이 만료되었거나 출처가 모호한 영상, 사진, 이미지, 음성 등을 재구성해 콜라주(Collage)하는 방식의 집약체로서, ‘아카이브 시네마(Archive Cinema)’ 또는 ‘기록 영화’의 특성을 내재하고 있다.

해당 기법을 활용하는 실험영화 제작자들은 관람자의 몰입을 증진시키기 위해 이를 적용한 영화를 개개인의 입장에서 자연스럽게 받아들일 수 있도록 그것을 제작자의 심리 상태와 동일시하는 방법으로 고안하였다. 이에 따라 파운드 푸티

지가 적용된 실험영화에서는 관람자에게 이질적인 느낌을 선사해 서서히 ‘불편한 영화’로 인식 가능함과 동시에 모호한 세계를 체험할 수 있도록 하는 효과를 발휘하게 된다[15].

이 같은 기법은 ‘디지털 기기 속 사람들’의 생활 양상을 보이는 영화에 주로 활용되고 있다. 이는 미디어에서 해당 장르를 ‘스크린 라이프(Screen Life)’로 지칭하는 것으로부터 확인할 수 있으며[16], 이러한 특성을 반영한 영화의 경우 등장 인물들의 관계가 스크린 속 ‘소셜 네트워크 서비스(SNS; Social Network Service)’ 속 웹 카메라를 비롯한 영상 플랫폼 화면으로 연출된 배경으로 그려진다[17].

특히 코로나19 여파로 영화 산업이 침체를 겪게 되면서 적은 예산으로도 작품성을 극대화 하고자 하는 효율적인 시도로 해당 기법을 적용하는 사례가 두드러지기 시작하였다. 이에 대한 성공 사례는 스크린 라이프 특성이 직접적으로 반영된 영화 <서치>(Searching)가 대표적이다. 영화의 흥행 성공 이후 ‘서치’의 기획 및 제작에 참여한 티무르 베크맘베토브(Timur Bekmambetov)가 국내 영화 <롱디>(Long Distance) 제작에 참여하게 된다는 점과 더불어 무당 가문에서의 세 달간의 기록을 파운드 푸티지 장르로 보인 영화 <광중>(The Medium)이 공포 영화의 새로운 지평을 열었다는 긍정적인 평가를 받으며, 그 효과를 증명하는 바 있다[18].

이 같은 특성에 착안한 본 연구에서도 현재 코로나19 이후의 시점에서 코로나19 이전의 미디어에서 다수의 인물들이 마스크를 착용하지 않은 채로 대화를 나누는 장면을 파운드 푸티지 장르를 결합하여 보임으로써 이를 보는 사람으로 하여금 ‘이질감’을 선사하는 것을 극대화하기 위한 전략으로서 선정하였다.

3-2 파운드 푸티지 장르의 작품 제작 사례 분석

본 연구에서는 파운드 푸티지 장르의 작품 제작 사례를 총 2가지로 분석하였다.

1) 영화 <서치> (Searching)

2018년 영화감독 아니쉬 샤건티(Aneesh Chaganty)가 연출한 영화 <서치>는 실종된 딸을 찾는 아버지의 이야기를 다룬다. 이 과정에서 모든 장면은 일인칭 시점으로 연출되었으며, 한정되었으나 익숙한 공간인 컴퓨터 운영체제(OS; Operating System), 모바일(Mobile), SNS와 캠코더(Camcorder) 등으로도 신선한 연출을 보일 수 있음을 시사하였다[그림 5].

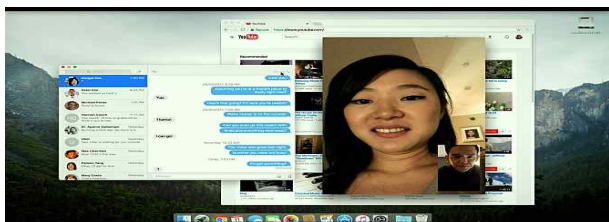


그림 5. 영화 <서치> (2018) [19]

Fig. 5. <Searching> (2018)



그림 6. 영화 <언프렌디드: 친구삭제> (2014) [20]

Fig. 6. <Unfriended> (2014)

2) 영화 <언프렌디드: 친구삭제> (Unfriended)

2014년 영화 <언프렌디드: 친구삭제>는 온라인에서 영상 통화를 기반으로 이루어지는 페이크 다큐멘터리 형식의 공포 영화이다. SNS가 발전은 SNS가 사람의 목숨을 위협하는 도구가 될 수 있다는 것을 시사하면서 이것이 간접적으로도 관람자에게 공포감을 부여할 수 있는 장치가 될 수 있음을 보였다. 특히 ‘의미 없는’ 마우스 커서 움직임, 운영체제 등 도구 및 환경의 상태 그 자체가 등장인물들의 심리 상태를 은유적으로 표현되는 도구가 된다는 것으로부터 창의적이고도 섬세한 연출을 가능하게 했다는 긍정적인 평을 받게 되었다[그림 6].

IV. 딥러닝 기반의 StyleGAN 기술

4-1 작품 <Face New Normal> 기획

1) 작품 <Face New Normal> 기획

본 작품은 현재 우리 사회가 직면한 코로나19 이후 새롭게 부상한 뉴노멀 시대에 착안하여, 작품 <Face New Normal>을 제작하였다. 이에 따라 코로나19 이후 변화한 우리의 모습을 인공지능의 관점으로 진단함으로써 뉴노멀 이후 우리 사회의 변화 양상을 객관적으로 확인해보는 것에 의의가 있다.

또한, 두 장면의 대비를 통해 보이게 되는 변화양상을 대비되는 인공지능의 모습은 기존 영화에서 그려지는 ‘인간의 경계를 넘나드는’ 존재로서의 역할을 수행하게 된다는 점에 주목하여[21]. 앞서 언급한 기획의도를 표현하기 위해 코로나19 이전과 이후를 포괄한 총 두 종류의 외형 데이터를 시각화하여 대비시키는 연출을 기획하였다. 구체적으로 앞서 언급한 파운드 푸티지 장르를 결합해 코로나19 이전과 이후 인공지능이 학습한 현재 우리 모습을 화상회의 플랫폼인 ‘줌(Zoom)’ 화면에 투영해 연출해보고자 한다.

2) 디자인 기획

작품의 기획의도를 표현하기 위해 비대면 시대의 대표적인 화상회의 플랫폼인 줌의 사용자 환경(UI; User Interface)을 적용한 화상회의 상황 연출을 기획하였다. ‘줌’은 2020년 코로나19로 인해 대면 회의가 어려워지면서 뉴노멀 시대를 대표하는 화상회의 플랫폼의 역할을 수행하고 있다[그림 7].

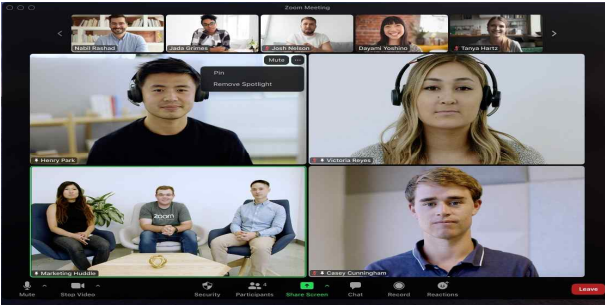


그림 7. 줌 [22]
Fig. 7. Zoom

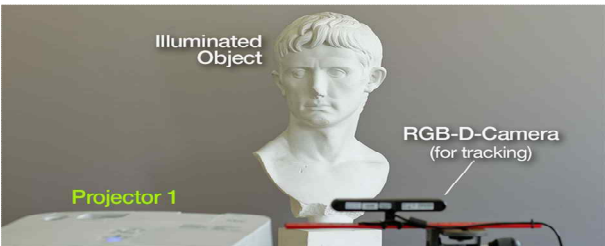


그림 8. 프로젝션 맵핑 시스템 구성
Fig. 8. The Configuration of Projection Mapping System

3) 기술 기획

본 작품은 두 가지 그래픽 작업과 프로젝션 맵핑, 촬영 작업을 기획하였다. 먼저 그래픽 작업에서 첫 번째 과정에서는 우리 일상생활의 회의 모습을 편집하는 것이다. 두 번째 과정에서는 현재 우리 모습을 사진으로 촬영 후 도출된 데이터를 수집하고 이것을 인공지능 기술인 StyleGAN을 이용해 학습시키는 방법이 병행된다. 이를 통해 인공지능이 본 우리 모습을 시각화하는 작업이 진행된다. 프로젝션 맵핑은 [그림 8]과 같이 하드웨어를 구성하여 StyleGAN을 이용해 시각화한 영상을 사람 형태의 흉상 모델에 전사하는 작업이다.

4) 연출 및 시놉시스 기획

본 작품의 콘셉트는 줌을 배경으로 한 화상회의이다. 회의 내용을 용이하게 전달하기 위한 도구로서 대사가 필요하지만, 본 작품에서는 대사를 제외한 상황과 인물의 표정을 통해 작품의 기획의도를 설정하였다. 본 작품의 기획 의도에 따른 예상 결과물은 아래 [그림 9]와 같다.

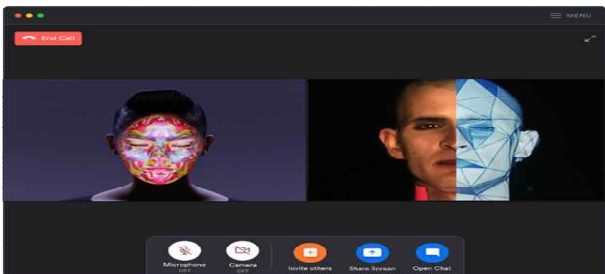


그림 9. 연출 예시
Fig. 9. The Example of the Scene

4-2 작품 <Face New Normal> 제작

본 작품의 제작 과정은 [그림 10]과 같이 연출안 제작, 디자인, StyleGAN을 이용한 영상 제작, 프로젝션 맵핑, 촬영 및 편집 총 5단계로 진행된다.

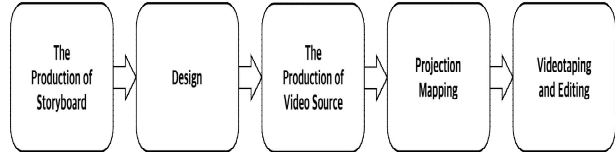


그림 10. 제작 과정
Fig. 10. The Process of Production

1) 연출안 제작

본 작품의 핵심이 되는 기획의도는 코로나19 이전과 이후의 대비를 강조하여 표현하는 것에 있다. 이러한 목적에 따라 해당 시놉시스에서는 ‘코로나19’ 자체 사건에 대한 반응이 중점이 되며, 작품 내에서 코로나19를 마주하는 인물의 반응을 객관적으로 확인해 볼 수 있다는 것이 장점으로 작용할 것으로 판단되었다. 이를 위한 전략으로 영화나 영상물에 나오는 회의 영상을 수집할 계획을 수립하였다. 이를 위해 온라인에서 코로나19 이전 사람들이 회의하는 영상을 찾아 편집한 후, 코로나 19 이후의 사람들의 데이터로 구성된 인공지능 대화 영상을 연출하는 전략을 수립하였다.

2) 디자인

기존의 사용자 환경인 줌을 적용한 디자인을 기획한 후, 구체적으로 줌 회의 참여자가 좌우로 배치되는 연출을 진행하였다. 이때 왼쪽 영상에 배치된 사람의 경우, 코로나19 이전 우리 모습을 표현했으며, 오른쪽 영상에 배치된 사람의 경우 인공지능이 학습한 데이터 결과로서 현재의 우리 모습을 표현하는 것을 계획하였다. 또한 작품의 몰입을 극대화하기 위해 이외의 컴포넌트와 요소들을 생략하는 과정을 거쳤다. 이에 대한 예시는 ‘2) 디자인 기획’의 [그림 7]과 ‘4) 연출 및 시놉시스 기획’의 [그림 9]와 같다.

3) StyleGAN을 이용한 소스 영상 제작

본 작품은 자연스러운 연출을 위해 다양한 인공지능 기술 모델의 분석을 진행하였다. 이 과정에서 논의된 기술로 이미지 생성에 주로 사용되는 GAN이 채택되었다[4]. 채택 과정에서는 이미지 인식에 주로 사용되는 합성곱 신경망(CNN; Convolutional Neural Network)[23]과 문자열 등의 순차적인 정보 인식에 주로 사용되는 순환 신경망(RNN; Recurrent Neural Network)[24]이 고려되었다. 최종적으로 본 연구의 기획의도에 부합한 ‘이미지 생성’의 GAN을 채택하게 되었다.

표 1. 논의된 딥러닝 모델

Table 1. Deep Learning Model Discussed

	GAN	CNN	RNN
Definition	Generative Adversarial Network	Convolutional Neural Network	Recurrent Neural Network
Feature	Image generation	Image Recognition	Information Recognition
Example	Generate the Facial Image	Alphago	Machine Translation

또한, GAN 기술 중 StyleGAN을 이용해 진행하였다. StyleGAN은 엔비디아에서 발표한 네트워크로, 이미지 스타일을 변환하거나 가상 이미지 생성에 효과적으로 활용 가능한 특성을 가진다[7], [9]. 이외에 논의된 다른 기술로는 CycleGAN(Cycle-Generative Adversarial Network)이 있는데, 해당 기술은 두 개의 데이터 셋(Data set)이 혼합되어 이미지 형태를 유지한 결과를 도출할 수 있는 특징을 가졌다[25]. 반면 StyleGAN은 기존 GAN 모델보다 뛰어난 품질의 얼굴 이미지를 생성할 수 있는 성능을 기반으로 하며[5], 이를 통해 새로운 가상 얼굴을 학습된 결과물로 보일 수 있다는 가능성이 내재되어 있다. 이에 착안하여 시각적으로 연출의 폭을 넓힐 수 있다고 판단하여 최종적으로 채택하게 되었다. 결과적으로 제작되는 영상 속에서 코로나19 이전과 이후 상황을 보이는 과정에서 명확한 대비를 그릴 수 있을 것이라고 판단되었다.

표 2. 논의된 GAN 모델

Table 2. GAN Model Discussed

	StyleGAN	CycleGAN
Feature	Generate the virtual image	Generate the mixed image
Learning Result	· Transfer the image style · Generate the new virtual Image	Generate the image by Mixing two data sets

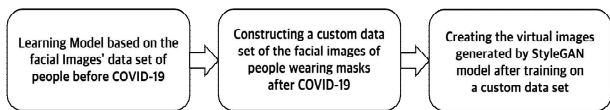


그림 11. StyleGAN 학습 과정 추진 체계

Fig. 11. The Learning Process Structure of StyleGAN

최종적으로 [표 2]에 대한 논의를 기반으로 한 개발 과정에서는 온라인에서 수집한 코로나19 이후 마스크를 착용한 사람들의 이미지 1,000장으로 학습한 가상 얼굴 이미지 영상을 구현하였다. 완성된 결과물은 약 600개이며, 얼굴이 지속적으로 변형되는 영상 형식으로 구현된다. 이를 캡처한 이미지는 [그림 12]와 같으며, 해당 이미지가 구성되어 동적으로 변형되는 모습을 보이는 영상은 구글 코랩(Google Colab) 서비스를 통해 가상 이미지를 생성할 수 있는 인공 신경망(Artificial Neural Network)인 StyleGAN2를 기반으로 생성하였다. 그 결과 코로나19 이후 사람들의 이미지를 기반으로 학습해 구현한 가상 얼굴 이미지를 구현하였다.

특히 [그림 12]에서 표현된 ‘불확실하고 기묘한 이미지’는 사실적인 모습의 결과를 보이는 것보다 상대적으로 연출의 폭을 넓힐 수 있는 요소라고 판단되었다. 이 같은 과정을 거친 후, 해당 영상을 통해 얻은 이미지는 프로젝션 맵핑(Projection Mapping)을 이용해 오브젝트에 투사하는 계획을 수립하였다.



그림 12. StyleGAN 학습 과정 추진 체계

Fig. 12. The Learning Process Structure of StyleGAN

4) 프로젝션 맵핑 시스템 구성 및 실행

프로젝션 맵핑은 ‘2) 디자인 기획’에서 언급한 인공지능이 학습한 코로나19 이후에 해당하는 현재 우리의 모습을 맵핑해 촬영하는 것을 목표로 진행했다. 이를 진행하기 위해 프로젝션 맵핑 프로그램인 매드매퍼(MadMapper)와 이를 구동하기 위한 노트북, 빔 프로젝터를 병행 사용하였다. 이를 위해, 앞서 ‘3) StyleGAN을 이용한 영상 개발’에서 보인 결과물인 [그림 12]를 영상소스로 활용하였다.

프로젝션 맵핑 실행 결과는 아래 [그림 13]과 같다. [그림 13]은 StyleGAN으로 구현한 가상 얼굴 이미지를 조각상에 그대로 투사하여 촬영한 결과이다. 그림에서 표현된 것과 같이 코로나19 이후 ‘마스크를 착용한’ 우리 모습의 조각상을 보이게 된다.



그림 13. 정면 사진(왼쪽)과 측면 사진(오른쪽)

Fig. 13. The Frontal(Left) and Side Images(Right)

5) 프로젝션 맵핑 시스템 구성 및 실행

영상 편집 및 제작 과정에서는 ‘4) 프로젝션 맵핑 시스템 구성 및 실행’에서 도출된 결과물을 촬영, 편집하는 과정이 이루어졌다. 해당 영상 표현을 위해 온라인에서 코로나19 이전 사람들의 회의 영상을 수집하였다. 수집의 범위는 드라마와

영화 속 회의 클립 영상으로 설정하여, 이를 콜라주 방식으로 표현하였다.

또한, 촬영한 결과물을 화상회의 플랫폼 중의 사용자 환경을 적용하여 코로나19 이후 우리 모습이 내재된 데이터를 학습한 인공지능의 모습을 연출하였다.

이에 따라 컴퓨터에서 Zoom을 실행 후 본격적으로 화상회의에 참여하게 되면서 대화를 이어가려는 본 작품은 이를 실험 영화 장르로서 ‘파운드 푸티지’ 기법을 적용하여 코로나19 이전에 제작된 영상물에서 인물들이 대화하는 장면과 코로나19 이후의 모습을 상징하는 인공지능 비디오표를 배치한 후, 오디오를 빨리 감기하는 방식으로 이들이 나누는 대화와 소통이 ‘의미 없게’ 보일 수 있도록 연출하였다. 이러한 연출을 진행한 이유는 ‘의사소통’이 원활하게 진행되지 않는 모습을 보이면서 소통이 되지 않는 상황을 표현하기 위한 것이다. 이를 ‘소음’으로 인식할 수 있는데, 소음은 코로나19 이전과 이후의 ‘격차(Gap)’를 보여주는 하나의 증거로 작용하게 된다. 그 결과 우리는 많이 변화했음을 보이게 되는 것이다. 동시에 최종적으로 이를 감상하는 관람자가 제 3자로서 본 작품의 장르가 파운드 푸티지 특성의 실험영화로 연상될 수 있도록 제작하였다.

또한, 이는 앞서 설정한 소통의 격차로 초래한 결과로서 자연스럽게 [그림 15]에서의 ‘모두에 대해 회의 종료’, ‘나가기’ 버튼을 누르는 결과를 보이게 된다. 이는 예술적 충격으로서의 ‘답답함’을 표출한 행동으로 볼 수 있다.

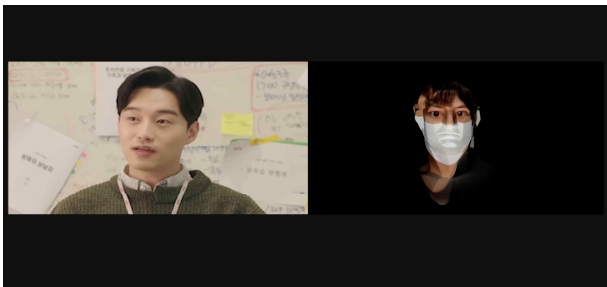


그림 14. 코로나19 이전과 이후 사람들의 데이터 기반 인공지능 대화 모습 영상 연출
 Fig. 14. The Video Production of Conversation Data-based AI with People Before and After COVID-19

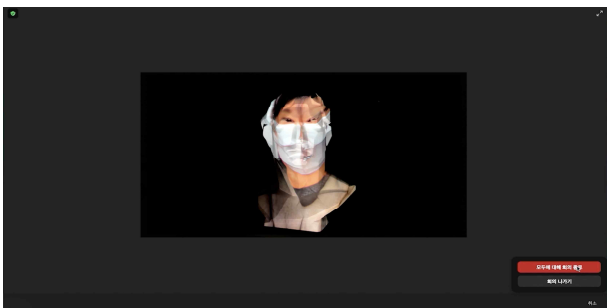


그림 15. 대화를 종료하기 위해 ‘모두에 대해 회의 종료’, ‘나가기’ 버튼을 누르는 최종 장면
 Fig. 15. The Final Scene of Pressing ‘End Meeting For All’ and ‘Exit’ Buttons to Cut the Conversation

4-3 작품 <Face New Normal> 결과 및 고찰

작품 <Face New Normal>의 결과로 알 수 있는 점은 사람과 인공지능 사이에서의 ‘원활하지 않은 소통’이다. 이를 위해 본 작품은 인공지능 기술 중 딥러닝 기반의 가상 이미지로 연출한 실험영화를 보인 것이 관람자 경험을 극대화할 수 있는 시도임을 보인다.

일반적으로 인공지능과 이미지 연구에서는 이미지를 구현하는 과정에 있어 정확도에 대한 전략과 시도가 지속적으로 이루어지고 있다. 이는 본 연구에서도 직면한 한계로 남으며, 구체적으로 사람 얼굴 위치와 오차가 발생한 점과 더불어 이미지의 높은 정확도를 달성하지 못한 점이 한계로 남는다.

그럼에도 불구하고, 이 같은 한계는 장점으로 승화되어 인공지능 예술의 관점에서 흥미로운 요소로 작용할 수 있다고 판단되었다. 이는 인공지능 해석의 ‘오류’를 내포함과 동시에 이질감을 표현하는 방법에 있어 적합한 연출로 고려되었기 때문이다. 결과적으로 고정된 얼굴만으로 이루어지는 연출을 보완하기 위해 답답함과 착잡함과 같은 감정적인 요소를 추가하여 예술적 충격을 표현할 수 있게 되었다.

V. 결 론

본 연구는 코로나19 이후 뉴노멀을 반영한 작품한 예술작품을 구현하였다.

이를 구현하기 위한 시도는 다음과 같이 이루어졌다.

첫째, 인공지능 기술로서 가장 큰 효과를 내고 있는 딥러닝 기반의 StyleGAN의 특징을 분석하여, 해당 기술을 통해 생성한 이미지 기반의 예술작품 제작 사례를 분석하여 작품들이 관람자에게 선사하는 이질감이라는 감정을 선사할 수 있다는 효과를 확인하였다.

둘째, 딥러닝 기반 이미지로 구성된 영상을 제작 과정에서 사실적이면서도 이질적인 느낌을 부여하기 위한 시도로 실험 영화 장르인 ‘파운드 푸티지’의 특성을 분석하여 이를 적용하기 위한 발판을 마련하였다.

셋째, 본 연구의 결과로서 코로나19 이전을 상징하는 ‘사람’과 코로나 19 이후 우리의 모습 그 자체를 상징하는 ‘인공지능 가상 이미지’ 사이에서 ‘원활하지 않은 소통’을 보인 영상의 제작 절차를 기술하였다.

한편, ‘인공지능 가상 이미지’는 인공지능 기술이 적용된 ‘데이터 시각화(Data Visualization)’의 결과로서, 최근 다양하게 이루어지고 있는 인공지능과 예술의 융합 가능성에 시도한 점에 의의가 있다. 특히 본 연구에서 인공지능 기술을 활용하여 현재 코로나19로 인해 변화하는 사회상을 표현하는 데이터 시각화 작품의 과정을 절차적으로 기술한 것은 추후 관련 작품 제작자들의 이론적 기반이 될 것이라고 기대한다. 특히 본 연구에서는 인공지능 기술을 응용한 실험영화 장르

로서 파운드 푸티지를 결합한 작품을 보인 것에 있어 독창적인 가치를 지닌다.

그럼에도 불구하고, 본 연구의 한계점은 다음과 같다. 기술적인 한계로서 StyleGAN으로 구현된 가상 이미지가 보이는 오차이다. 구현 과정에서 이미지의 화질과 연관되는 정확도가 낮아 원본과의 구별이 용이하지 않다는 점이다. 하지만 이미지의 정확도를 높이는 방법에 대한 시도는 인공지능 이미지 알고리즘 개발 연구에서 공통적으로 여겨지는 가장 큰 목표이기도 하다[4], [7], [9], [23-25]. 이에 따라 이를 위한 지속적으로 시도 및 논의와 새로운 방법들이 제안되고 있는 상황과 결부된다는 것을 피력하고자 한다. 따라서 본 연구를 기반으로 한 후속 연구에서는 이미지 데이터 학습 횟수를 늘리는 방법과 더불어, 본 연구에서도 논의한 다른 기술을 다양하게 활용해본다면 효과적인 방법을 도출해낼 수 있을 것임을 기대한다.

동시에 본 연구에서는 인공지능 기술의 한계를 ‘인터페이스(Interface)’로서 역할을 수행하게 되는 예술 작품을 매개로 하여 관람자에게 감정적인 요소를 부여할 수 있다[26]는 상호보완적인 결과를 보인 점에서 예술공학적 시사점을 가진다.

최종적으로 코로나19 뉴노멀 시대 ‘관찰자로서 인공지능’의 시점을 통해 진단하여, 이를 비평적인 시선으로 바라본 코로나19 사회상이 반영된 본 작품을 국내외 전시를 통해 상영할 계획을 수립하였다. 관람자에게 예술적 충격을 보이면서 포커스 그룹 인터뷰(FGI; Focus Group Interview)를 통해 관람자의 수용태도와 의견을 듣는 연구가 진행된다면 앞으로의 인공지능 예술 작품 활성화에 기여할 수 있을 것으로 기대된다. 더 나아가 앞으로 이러한 전시가 활성화되어 관람자에게 색다른 형태의 예술작품을 제공하는 계기를 마련함으로써 관련 작품이 지속적으로 연구 및 개발되어지기를 제안한다.

감사의 글

본 연구는 2020년도 중앙대학교 4단계 BK21 인공지능-콘텐츠 미래산업 교육연구단의 지원에 의하여 이루어진 연구로서, 관계부처에 감사드립니다.

참고문헌

- [1] T. H. Cho, *Deep-learning for everyone, [KIIP] Korea Institute of Intellectual Property*. Seoul: Gilbut, pp. 1-308, 2017.
- [2] H. S. Choi and Y. M. Shon, “Case Analysis and Characteristics of the Convergence between Artificial Intelligence and Art Creation Activities,” *The Korean Society of Science & Art*, Vol. 28, No. 0, pp. 289-299, March 2017. <https://doi.org/10.17548/ksaf.2017.03.28.289>
- [3] ResearchGate. Natural Language Understanding (NLU) [Internet]. Available: <https://www.researchgate.net/publication/325111111>
- [4] G. Ian, J. P. Abadie, M. Mirza, B. Xu, D. W. Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” *Communications of the ACM*, Vol. 63, No. 11, pp. 139-144, November 2020. <https://doi.org/10.1145/3422622>
- [5] J. Y. Kim, S. A. Hong, Kim and H. M. Kim, “A StyleGAN Image Detection Model Based on Convolutional Neural Network,” *Journal of Korea Multimedia Society*, Vol. 22, No. 12, pp. 1447-1456, December 2019. <https://doi.org/10.9717/kmms.2019.22.12.1447>
- [6] Y. J. Jo, K. M. Bae and J. Y. Park, “Research Trends of Generative Adversarial Networks and Image Generation and Translation,” *[ETRI] Electronics and Telecommunications Trends*, Vol. 35, No. 4, pp. 91-102, August 2020. <https://doi.org/10.22648/ETRI.2020.J.350409>
- [7] T. Karras, S. Laine and T. Aila, “A Style-Based Generator Architecture for Generative Adversarial Networks,” *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA: USA, pp. 4396-4405, June 2019. <https://doi.org/10.1109/CVPR.2019.00453>
- [8] Neurohive. The StyleGAN Code Released: Neural Network for Faces Generation by NVIDIA. [Internet]. Available: <https://neurohive.io/en/state-of-the-art/stylegan-code-released/>
- [9] T. Karras, S. Laine, M. Aittala., J. Hellsten., J. Lehtinen and T. Aila, “Analyzing and Improving the Image Quality of StyleGAN,” *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA: USA, pp. 8107-8116, June 2020. <https://doi.org/10.1109/CVPR42600.2020.00813>
- [10] Art Center nabi. [ISEA2019] Lux Aeterna. [Internet]. Available: http://www.nabi.or.kr/page/board_view.php?brd_idx=1009&brd_id=project
- [11] David Rokeby. Interactive Installations : The Giver of Names (1990-) (in progress) [Internet]. Available: <http://www.davidrokeby.com/gon.html>
- [12] H. Y. Kim, “A Study on Industrial Potential of Artificial Intelligence through the Cases of Film and Artificial Intelligence Art,” *Cartoon and Animation Studies*, Vol. 0, No. 50, pp. 423-452, March 2018. <https://dx.doi.org/10.7230/KOSC.AS.2017.50.423>
- [13] C. S. Rhee and H. J. Rhee, “Expectations and Anxieties Affecting Attitudes toward Artificial Intelligence Revolution,” *JOURNAL OF THE KOREA CONTENTS ASSOCIATION*, Vol. 19, No. 9, pp. 37-46, September 2019. <https://doi.org/10.5392/JKCA.2019.19.09.037>
- [14] D. Y. Kim and S. J. Bae, “A Transdisciplinary and Humanistic Approach on the Impacts by Artificial

Intelligence Technology,” *JOURNAL OF BROADCAST ENGINEERING*, Vol. 24, No. 3, pp. 411-419, May 2019. <https://doi.org/10.5909/JBE.2019.24.3.411>

[15] S. J. Kwon, Anti-aesthetic forms and meanings in found footage experimental films: focused on deconstruction and pastiche in somatic images, Ph.D. dissertation, The Graduate School of Advanced Imaging, Science, Multimedia and Film of Chung-Ang University, Seoul, February 2017.

[16] J. G. Lee, “The directing’s analysis of the film <Searching> - Focusing on fusion of pound footage and traditional genres - ”, *Journal of the Moving Image Technology Association of Korea*, Vol. 0, No. 32, pp. 43-65, May 2020. <https://doi.org/10.34269/mitak.2020.1.32.003>

[17] Korea Business News Agency. ‘Long Distance’ Jang Dong-Yoon and Park Yuna show love in the age of screen life. [Internet]. Available: <http://www.nspna.com/news/?mode=view&newsid=508368>.

[18] CBS Nocutnews. [EN:Terview] ‘The Medium’ entwined fear and controversy, director Banjong directly answered. [Internet]. Available: <https://www.nocutnews.co.kr/news/5590205>.

[19] Naver Movie. <Searching> (2017) [Internet]. Available: <https://movie.naver.com/movie/bi/mi/photoView.naver?code=172425&imageNid=6618594#tab>.

[20] Naver Movie. <Unfriended> (2014) [Internet]. Available: <https://movie.naver.com/movie/bi/mi/photoView.naver?code=127768>.

[21] W. C. Son, “Science and Technology Governance in the Era of Artificial Intelligence,” *CHUL HAK SA SANG : Journal of Philosophical Ideas*, Vol. 0, No. 68, pp. 267-299, May 2018. <https://doi.org/10.15750/chss.68.201805.009>

[22] 24 BOAN24, Zoom, Improves New Feature Updates and Accessibility[Internet]. Available: <https://www.boan24.com/news/articleView.html?idxno=12440>.

[23] L. A. Gatys, A. S. Ecker and M. Bethge, “Image Style Transfer Using Convolutional Neural Networks,” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV: USA, pp. 2414-2423, June 2016. <https://doi.org/10.1109/CVPR.2016.265>

[24] W. Zaremba, I. Sutskever and O. Vinyals, “Recurrent neural network regularization,” arXiv: 1409.2329, 2014.

[25] J. Y. Zhu, T. Park, P. Isola and A. A. Efros, “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks,” *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. pp. 2223-2232, October 2017. <http://arxiv.org/abs/1703.10593>

[26] D. Ma and H. G. Kim, “Media Art of Using AI Technology for Audience Interaction to Generate Aging,” *The Korean Journal of Art and Media*, Vol. 19, No. 1, pp. 117-143, May 2020. <https://doi.org/10.36726/cammp.2020.19.1.117>

박지혜(Ji-Hye Park)



2020년 : 중앙대학교 컴퓨터공학과 (공학사)

2020년~현 재: 중앙대학교 첨단영상대학원 영상학과
예술공학 석사과정
※ 관심분야 : Artificial Intelligence, Human-Computer Interaction, Interface

송해민(Hae-Min Song)



2018년 : 성균관대학교 영상학과 (예술사)

2020년~현 재: 중앙대학교 첨단영상대학원 영상학과
예술공학 석사과정
※ 관심분야 : Artificial Intelligence, Media Art, Virtual Reality

박진완(Jin-Wan Park)



1995년 : 중앙대학교 컴퓨터공학과 (공학사)
1998년 : Pratt CGIM Computer Media (MFA)

2003년~현 재: 중앙대학교 첨단영상대학원 영상학과
예술공학 교수
※ 관심분야 : Art&Technology, Procedural Animation, Data Visualization