

초소형 버추얼 프로덕션 환경에서 디지털 휴먼을 이용한 촬영 사례

임재호^{O*}

장민정

전상욱

이수빈

박민수

김유진

(주)엔진비주얼웨이브

{im.jaeho, jang8296, alexchun, dubi0102, minsoo, yujinkim}@its-new.co.kr

Case Study : Cinematography using Digital Human in Tiny Virtual Production

Jaeho Im^{O*}

Minjung Jang

Sang Wook Chun

Subin Lee

Minsoo Park

Yujin Kim

eNgin visual wave Ltd.,

요 약

본 논문에서는 디지털 휴먼을 활용한 버추얼 프로덕션에서의 촬영 사례를 소개한다. 본 촬영 사례는 LED를 활용한 버추얼 프로덕션의 시스템 구성과 디지털 휴먼을 활용한 효율적인 촬영 파이프라인을 다루고 있다. LED를 이용한 버추얼 프로덕션은 주로 LED에 배경을 투사하여 촬영하는 것과 달리 본 사례는 디지털 휴먼을 가상 배우로 사용하여 실제 배우와 소통하는 장면을 촬영하였다. 또한, 실제 배우와 디지털 휴먼의 대화 장면을 실시간 엔진을 이용하여 촬영하기 위해 음성과 텍스트 기반 한국어 립싱크 모션 자동 생성 기술을 적용하여 디지털 휴먼의 발화 애니메이션을 사전에 제작해 사용하였다. 우리는 실시간 엔진을 활용하여 LED 기반의 버추얼 프로덕션 환경에서 실제 배우와 디지털 휴먼을 이용한 짧은 드라마 콘텐츠를 제작함으로써 본 촬영 사례에 대하여 검증하였다.

Abstract

In this paper, we introduce a case study of cinematography using digital human in virtual production. This case study deals with the system overview of virtual production using LEDs and an efficient filming pipeline using digital human. Unlike virtual production using LEDs, which mainly project the background on LEDs, in this case, we use digital human as a virtual actor to film scenes communicating with a real actor. In addition, to film the dialogue scene between the real actor and the digital human using a real-time engine, we automatically generated speech animation of the digital human in advance by applying our Korean lip-sync technology based on audio and text. We verified this filming case by using a real-time engine to produce short drama content using real actor and digital human in an LED-based virtual production environment.

키워드: 버추얼 프로덕션, 디지털휴먼, 영화 프로덕션, VR, AR

Keywords: virtual production, digital human, film production, VR, AR

1. 서론

*corresponding author: Jaeho Im/ eNgin visual wave Ltd.(im.jaeho@its-new.co.kr)

하드웨어 성능과 실시간 엔진 기술의 발달은 비실시간 영역에 있던 컴퓨터 그래픽스를 실시간의 영역으로 확장시켰다. 특히 영화는 다양한 특수효과와 실사 렌더링으로 인한 비실시간 기술의 집약체였으나, 최근 실시간 엔진을 활용한 영화 산업 전반적인 파이프라인이 변화하고 있다. 버추얼 프로덕션은 가상을 의미하는 버추얼과 영화 후반 작업을 담당하는 프로덕션의 합성어로, 기존 파이프라인에서는 컴퓨터 그래픽스를 이용해 특수효과를 제작하고 적용하였던 후반 프로덕션의 영역을 실제 촬영이 이루어지는 온셋(On-Set) 환경으로 앞당겼다. 이는 배우와 촬영 스태프 그리고 감독에게 가상 공간에 대한 명확한 시각적인 결과를 제공함으로써 전통적인 영화 제작 과정에서 발생하는 불확실성을 최소화하는 효율적인 영화 제작 환경을 만들어 냈다. 특히 다수의 LED 패널을 디스플레이 환경으로 사용하여 구축하는 LED 기반 버추얼 스튜디오는 그린 크로마키 촬영을 대체하였다. 기존의 프로덕션 방식에서는 배우와 감독이 가상 공간을 상상하며 촬영하던 환경을 실제 보이는 환경으로 바꾸면서 촬영 현장의 불확실성을 낮추고 배우에게 역할에 대한 높은 몰입감을 제공할 수 있게 되었다 (Figure 1)

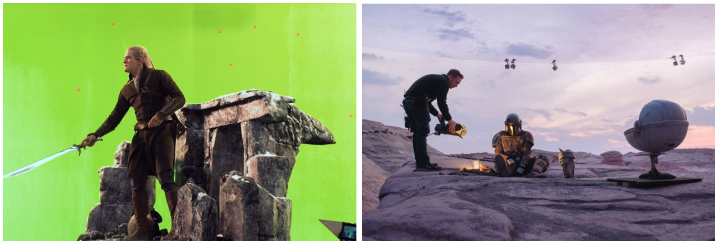


Figure 1 (Left) Traditional chroma key 'The Lord of the Rings' (Right) Virtual production 'The Mandalorian'

촬영 현장의 변화는 영화 제작 파이프라인도 변화시켰다. Figure 2는 전통적인 프로덕션 작업 방식과 버추얼 프로덕션 작업 방식의 차이를 보여준다. 전통적인 프로덕션에서는 (1) 스토리를 기반으로 사전 시각화 전용 디지털 에셋을 제작하고, (2) 제작한 세트에서 실제 배우들이 촬영하는 온셋을 거친 후, (3) 고품질 디지털 에셋 및 CG/VFX를 제작하여 촬영 영상과 합성해 콘텐츠를 완성한다. 반면에 버추얼 프로덕션에서는 고품질 디지털 에셋을 프로젝트 시작 단계에서 제작하여 온셋에서 촬영하고, 시뮬레이션과 같이 실시간성 보장이 어려운 효과들만 최종 합성하여 결과물을 완성하게 된다. 버추얼 프로덕션은 현장에서 촬영한 장면들을 최종 장면으로 활용하기 때문에 전통적인 프로덕션 방식에 비해 후반 작업의 비중이나 중요도가 매우 낮아지는 반면, 촬영 앞 단계 작업의 중요성이 증대되는 차이가 있다.

국내외로 버추얼 프로덕션을 활용한 영화나 드라마가 활발히 제작되고 있으며 관련 연구가 지속적으로 진행되고 있다. 그러나 버추얼 프로덕션의 핵심인 LED는 주로 배경을 송출하는 용도로 그 활용이 매우 제한적이다. 국외 대형 스

튜디오의 경우 실제 인물을 중심으로 근경은 실사 세트로 구성되 중경과 원경은 가상 장면을 제작하여 활용하는 반면, 국내 사례에서는 주로 원경에 배치된 배경을 송출함으로써 현지 출장 촬영(로케이션 촬영)을 대체하는 정도로 사용된다. 따라서 우리는 LED의 활용 사례를 단순히 배경에 국한하지 않고 사람과 상호작용 가능한 디지털 휴먼으로 영역을 확장하는 시도를 하였다.

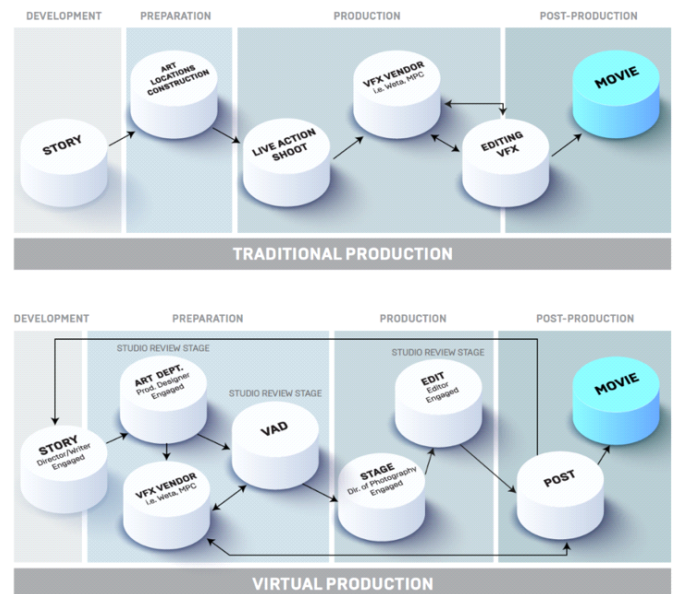


Figure 2 Traditional vs. Virtual production in film (Source: The Virtual Production Field Guide)

본 논문에서는 최근 크게 주목받고 있는 디지털 휴먼과 버추얼 프로덕션, 두 기술을 결합한 촬영 사례를 소개한다. LED 기반 버추얼 프로덕션 촬영 환경에서의 조명, 밝기와 같은 기계적인 한계점을 파악하고 디지털 휴먼을 투사하여 고품질의 결과를 얻기 위한 방법론을 소개한다. 본 촬영 사례는 LED 기반 버추얼 프로덕션에서 디지털 휴먼과 실제 배우가 함께 등장하는 짧은 영상 콘텐츠를 직접 촬영함으로써 고품질의 결과를 얻기 위한 우리의 방법론을 증명하였으며, 이에 디지털 휴먼을 버추얼 프로덕션 환경에서 사용하기 위한 효율적인 파이프라인을 소개한다.

2. 관련 사례

2.1 버추얼 프로덕션

그린 크로마키 환경에서 촬영하고 후반 합성을 하는 그린 스튜디오 촬영 방식은 전통적인 파이프라인에서 CG/VFX를 합성하기 위해 사용하는 촬영 기법 중 하나이다. 버추얼 프로덕션 기술이 소개된 지금도 널리 사용되고 있으며, 실시간 합성을 위해 많은 준비가 필요하거나 가상의 생물이 등장하고, 등장인물을 디지털 에셋으로 대체해야 하는 경우에 사용된다. 그린 스튜디오 촬영은 전통적인 후반 합성 방식

과 버추얼 프로덕션을 결합한 실시간 합성 방식으로 분류되며, 후자는 그린 스크린 라이브(Green Screen Live)라는 이름으로 불린다. 두 기술의 차이는 실사와 가상 공간의 합성이 촬영 현장(온셋)에서 이루어지는지, 프로덕션으로 데이터가 이전된 후에 진행되는지에 따라 분류된다.

국내에서는 버추얼 프로덕션의 시작인 On-stage 사전 시각화 기술이 개발되었으나 그 활용도가 높지 않다. 2014년 영화 제작 파이프라인인 Production 및 Post-production 공정과 연동되는 실시간 저작 소프트웨어인 사전 시각화 도구 'Farsee'가 개발되었지만, 하드웨어 및 구동에 많은 제약이 있어 실제 상용화에 성공하지 못하였고 영화 현장에도 활용되지 못하였다. 2011년 부산영상위원회가 설립한 아시아 최초 버추얼 스튜디오인 '부산 3D 프로덕션센터-디지털베이'의 온셋 사전 시각화(On-set Pre-visualization) 기술은 소프트웨어와 하드웨어 대부분이 해외 수입에 의존하였고, 스튜디오 현장 배경 합성 솔루션 수준으로 실시간으로 배경의 환경을 편집하지 못한다는 한계점이 있다.

LED를 이용한 버추얼 프로덕션 기술은 SIGGRAPH 2019년 Real-Time Live 세션을 시작으로 빠르게 발전하고 있다. 에픽게임즈는 2019년 자사의 언리얼 엔진을 활용한 LED 기반 버추얼 프로덕션 기술을 소개하였는데, 실시간으로 구동되는 촬영 현장 기술에 국내외 대규모 프로덕션들이 앞다투어 기술 개발에 뛰어들었다. ILM은 스타워즈 IP 기반의 '만달로리안'을 LED 기반 버추얼 프로덕션을 이용하여 촬영하였다. 실제 공개되는 콘텐츠에 적용한 최초 사례로, LED 기반 버추얼 프로덕션의 이름은 ILM에서 사용하는 볼륨(Volume)으로 대중에게 알려졌다.

국내에서는 XON의 XR 스튜디오가 LED 기반 버추얼 스튜디오를 최초로 시작하였다. XON의 스튜디오는 카메라 위치 추적과 LED 패널의 캘리브레이션을 위하여 디스가이즈(Disguise)를 이용하였다. 디스가이즈 자체 기능으로 LED가 존재하지 않는 영역까지 3D 그래픽을 연장하는 XR 기술을 접목하여 뮤직비디오 등의 콘텐츠 촬영이 가능하였으나, 렌더링, 카메라 트래킹과 같은 촬영 필수 기술들을 디스가이즈를 통해 송출해야 했기 때문에 빠른 카메라 움직임에는 지연(Latency)이 발생하는 한계점이 있었다. 또 다른 국내 버추얼 스튜디오로는 VA 코퍼레이션의 약 3,000평 규모의 하남 버추얼 스튜디오가 있으며, 텍스터 또한 파주에 D1 스튜디오를 개관하였다. 비브 스튜디오는 곤지암에 VIT 스튜디오를 개관하여 BTS 등 국내 아이돌 뮤직비디오 촬영을 통해 버추얼 프로덕션 활용도를 높이고 있다.

연구 기술 개발 측면에서 LED 기반 버추얼 프로덕션 기술이 공개된 이후, 환경 측면에서의 연구[1]와 촬영 효율성을 위한 연구[2]가 진행되었다. 환경 분야에서는 기존의 그린 스튜디오 및 실사 세트를 제작하는 경우 재활용이 불가능한 쓰레기 및 전력 사용량이 높아짐에 따라 환경 파괴가 가속되는 반면, 버추얼 프로덕션을 사용하는 경우 자원의 재활

용 및 에너지 절약 측면에서 큰 효율을 볼 수 있기 때문에 환경 보호를 위하여 버추얼 프로덕션을 장려해야 한다고 주장하였다. 위 연구는 영화의 촬영 및 기술을 다양한 측면에서 접근한 좋은 예이긴 하지만 촬영에 대한 솔루션을 다루지는 않았다.

넷플릭스는 SIGGRAPH 2022에서 버추얼 프로덕션 환경에서 카메라에 촬영되는 영상의 결과와 사람의 눈에 보이는 결과가 크게 다른 것에 착안하여 둘 사이의 간극을 최소화하기 위한 색상/채도 조절 기술을 공개하였다. 해당 기술을 오픈소스로 공개하였으나, 버추얼 스튜디오를 보유한 프로덕션마다 LED에 대한 상태 및 제어 방식이 다르기 때문에 일반화하여 적용하기는 어려운 한계점이 있다.

영상 합성 프로그램인 NUKE를 개발하는 파운더리 솔루션은 SIGGRAPH 2022에서 NRT(Near Real Time) 연구를 공개하였다[3]. 위 논문에서 공개한 기술은 딥러닝 기반 연구로, 저품질의 가상환경에 대해 실사 인물 및 물체가 없는 상태로 가상 공간을 송출하는 배경을 촬영한 뒤, 실제 인물을 촬영하고 고품질 렌더링 샷과 실제 인물 촬영 데이터를 NUKE를 이용하여 최종 합성하는 방식으로 소형 스튜디오에서도 사용 가능한 기술이다. 하지만 위 기술의 경우 NUKE라는 소프트웨어를 필수적으로 사용해야 하는 단점이 있으며, 대부분을 후반 합성에 의존해야 하기 때문에 온셋 환경에서 최종 결과를 만들고자 하는 버추얼 프로덕션의 근본적인 사용 의미에 부합하지 못하는 한계가 있다.

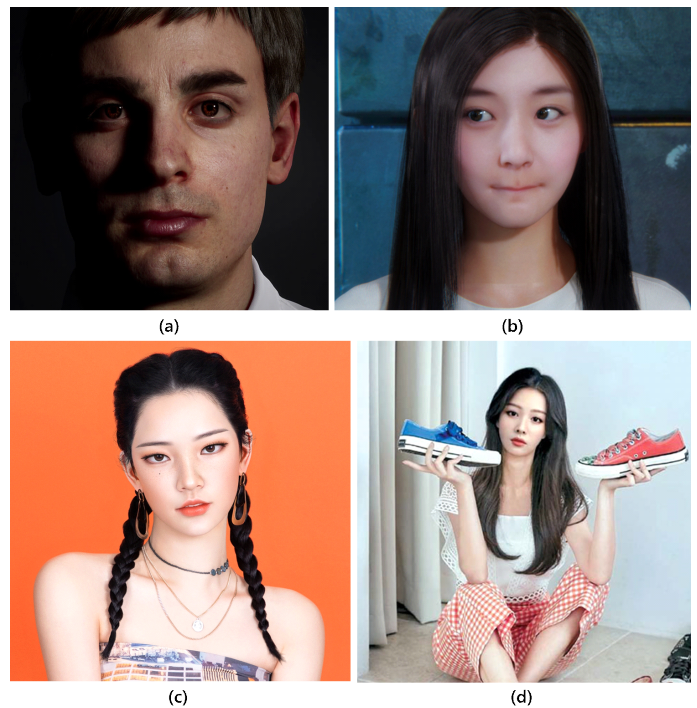


Figure 3 3D digital human examples. (a) Vincent (b) Sua (c) Rogy (d) Lucy

이처럼 실시간 엔진을 활용한 버추얼 프로덕션 기술은 촬영

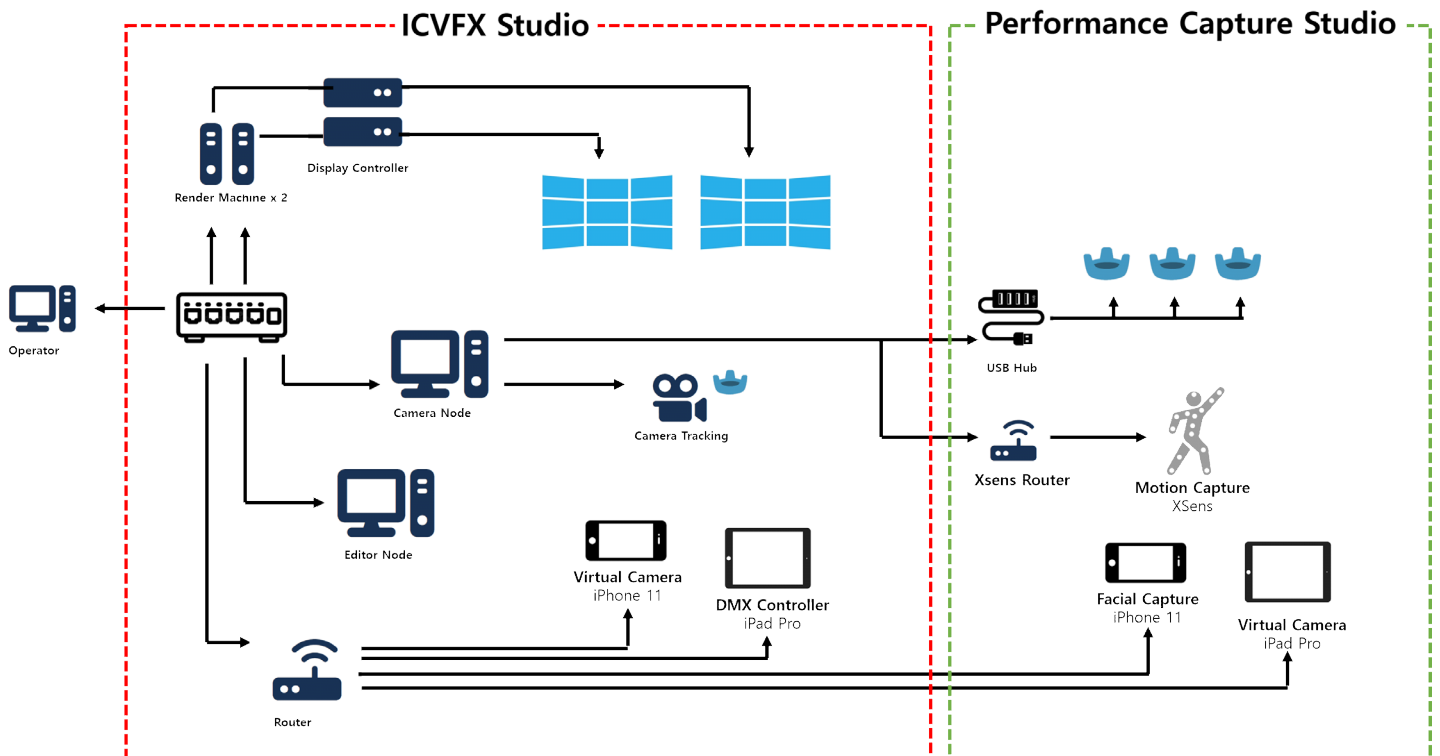


Figure 4 Virtual production system overview

영 환경 개선과 효율적인 영화 제작 파이프라인을 위하여 지속적으로 연구가 진행되고 있으나, 주로 사람과 상호작용하는 근경은 실사 세트에 대체하고 이외의 근경이나 사람 뒤에 존재하는 중경과 원경은 가상 세트로 구축하여 카메라의 포커스를 이용한 촬영 기법으로 활용하고 있다. 이에 따라 실제 사람이 디지털 휴먼과 함께 등장하는 장면이나 디지털 크리처가 등장하는 장면을 사람과 함께 촬영하는 시도는 거의 이루어지지 않고 있다.

2.2 디지털 휴먼

디지털 휴먼 제작은 영화 CG/VFX 산업에서 지속적인 요구가 있었으며, 고품질의 디지털 휴먼을 제작하기 위해 많은 연구가 진행되었다. 비실시간 디지털 더블(실제 인물의 디지털 복제)의 경우 2017년 영화 '로건'을 통해 고품질의 사실적인 디지털 휴먼이 공개되었으며, 이후 유니티(Unity)를 기반으로 개발된 버추얼 휴먼 '수아(온마인드)'나 언리얼 엔진(Unreal Engine) 기반의 '빈센트(자이언트스텝)' 사례를 통해 청중과 상호작용 가능한 실시간 디지털 휴먼이 공개되었다. 최근에는 로지(Sidus Studio X), 루시(롯데홈쇼핑) 등 다양한 버추얼 인플루언서들이 공개되면서 대중들과 소통 가능한 디지털 휴먼이 개발되었다(Figure 3).

이처럼 코로나 팬데믹을 지나며 다양한 버추얼 인플루언서가 등장하였으나 완전한 3D 기반의 디지털 휴먼은 극소수이며, 대부분 얼굴 교체 방식의 디지털 휴먼으로 제작되었다. 얼굴 교체 알고리즘으로 개발된 디지털 휴먼은 3D 디지

털 휴먼보다 빠르게 제작이 가능하며 실시간으로 대중과 소통이 가능한 장점이 있는 반면, 필수적으로 해당 디지털 휴먼을 연기하는 연기가 필요하고, 극심한 조명의 변화나 격렬한 액션에는 사용이 어려운 한계점도 존재한다.

디지털 휴먼은 주로 영화나 드라마에서 디지털 더블로 등장하며, 실제 사람이 하기 어려운 스텐트 액션이나 신체 일부가 파손되거나 소실되는 역할에 주로 사용된다. 디지털 휴먼은 성능과 품질 사이에 교체비용(Trade-off)이 존재하는데, 고품질로 렌더링할수록 실시간성을 상실하기 때문에 고품질을 요구하는 영화나 드라마에서 디지털 휴먼을 적극적으로 사용하기는 어렵다. 국내 버추얼 인플루언서를 운영하는 기획사에서는 자사에서 제작한 버추얼 인플루언서의 인지도를 높이기 위하여 드라마에 출연시키는 시도가 증가하고 있으나, 중요도가 낮은 장면에 비중이 낮은 역할만 가능하다는 한계가 있다.

3. 시스템 개요

이번 장에서는 본 촬영 사례에 사용된 두 가지 주요 기술들에 대한 개요를 소개한다. 먼저 3.1에서는 본 촬영 사례에 사용된 버추얼 스튜디오의 공간 구성을 설명한다. 기존의 버추얼 스튜디오들이 대규모 공간을 활용하였던 것에 반해, 우리는 중소 프로덕션에서 운용 가능한 초소형 버추얼 스튜디오를 구성하는 방법을 제안한다. 이어서 3.2에서는 본 촬영 사례에 사용된 디지털 휴먼 배우 제작 공정을 다룬다. 우리는 최소한의 인원과 스캔 시스템을 활용해 고품질의 디

지털 휴면을 효율적으로 제작하였다.

3.1 초소형 버추얼 스튜디오

Figure 4는 우리의 버추얼 스튜디오 시스템 환경을 보여준다. ICVFX는 In-Camera VFX의 약자로, LED 기반 버추얼 프로덕션을 표현하는 단어이다. 우리의 시스템은 제작된 가상공간을 송출하기 위한 ICVFX 스튜디오와 디지털 휴면의 얼굴 표정 및 몸동작을 실시간으로 전달하기 위한 퍼포먼스 캡처 스튜디오로 구성된다. 본 촬영에서는 디지털 휴면의 몸 움직임이 크지 않기 때문에 퍼포먼스 캡처 스튜디오는 사용하지 않았다. 우리의 버추얼 스튜디오는 에픽게임즈의 버추얼 스튜디오 가이드에 따라 구성하였으며, ICVFX 스튜디오의 각 요소는 다음과 같다.

Operator

버추얼 스튜디오를 구성하기 위해서는 많은 수의 PC가 요구된다. 그중 오퍼레이터 PC에서는 네트워크로 연결된 각각의 PC를 관리하고 실행하거나 프로그램을 종료시키는 일을 수행한다. 오퍼레이터 PC에서 전체 시스템은 에픽게임즈에서 버추얼 프로덕션을 위해 개발한 상용 솔루션인 스위치보드(Switchboard)를 사용한다.

Rendering Machine

렌더링 머신은 대형 LED에 화면을 송출하기 위해 사용한다. 디스га이즈와 같은 미들웨어(Middleware)를 사용하지 않는다면 일반적으로 다수의 LED로 화면을 송출하기 위해 디스플레이 제어부가 존재하며, 렌더링 머신에서는 디스플레이 제어부와 HDMI 케이블을 통해 영상 데이터를 전달한다.

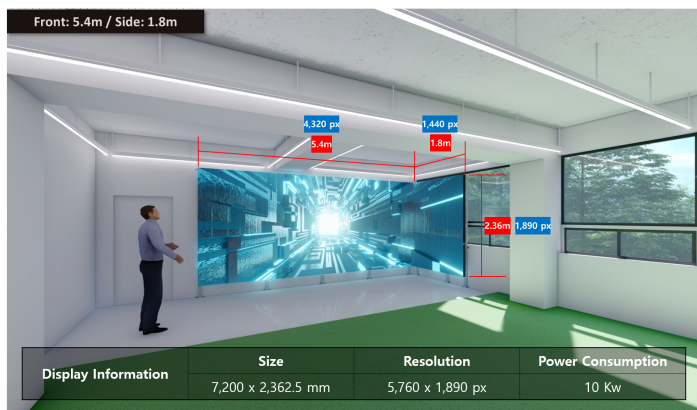


Figure 5 LED display in our virtual production system

Figure 5는 우리의 버추얼 스튜디오에 설치된 LED를 보여준다. 영화에 사용되는 가상환경은 실사에 가까운 초고품질 렌더링을 요구한다. 비록 우리의 버추얼 스튜디오가 크지 않더라도 한 대의 렌더링 머신으로는 고품질 장면을 실시간으로 렌더링하는 것을 보장하기 어렵다. Figure \ref{fig:vp_overview}에서 볼 수 있듯이 우리는 LED 화면을 분할하여 각각의 LED에 하나의 렌더링 머신을 할당하는 분할 렌더링

방식을 사용하며, 이를 통해 고품질 가상환경을 LED 화면에 실시간으로 송출하는 것이 가능하다.

Camera Tracking

영상 촬영을 위해서는 실제 카메라와 가상공간의 카메라가 서로 동일한 동작을 할 필요가 있다. AR 및 XR 공연이 크게 인기를 얻으면서 RedSPY, Mo-Sys, NCam과 같은 물리적인 좌표를 가상공간에 전달하여 동기화하는 카메라 장비들이 공개되었으며, 좌표를 획득하는 방식에 따라 이미지 처리 기반 방식과 마커 기반 방식으로 구분된다. 이미지 처리에 기반한 좌표 추적 방식을 사용하는 대표적인 장비는 NCam이다. 하지만 이미지 처리 기반 방식은 영상 내 특징점이 적은 LED 기반 버추얼 스튜디오에는 적합하지 않다. 마커 기반 방식의 대표적인 장비는 RedSPY와 Mo-Sys가 있다. RedSPY는 AR 공연에 특화되어 다양한 미들웨어들을 지원하지만, 미들웨어를 배제한 우리의 시스템에는 적합하지 않다. Mo-Sys는 높은 화면 동기화와 줌 및 포커스를 실시간 가상 동기화 제어가 가능한 장점이 있으나, 대규모 환경에 적합하기 때문에 우리의 초소형 버추얼 스튜디오에서 사용하기 어렵다.

우리는 본 촬영 사례에서 초소형 버추얼 스튜디오 환경에서 사용하기 용이한 VR 장비 기반의 카메라 추적 시스템을 사용한다. 바이브 트래커(VIVE Tracker)는 단일 구성으로 5m x 5m 크기의 공간 내 장비 위치 및 회전 값 추적이 가능하며, 바이브 스테이션(VIVE Station)을 추가 장착하면 더 넓은 공간도 추적이 가능하다는 장점이 있다. 특히 층고가 낮은 스튜디오 환경에서는 마커 기반 트래킹 방식보다 효율적이며 별도의 미들웨어를 거치지 않아도 된다는 장점이 있다.

Editor

버추얼 프로덕션의 장점은 촬영 환경에서 실시간으로 장면을 편집하고 디지털 에셋들의 위치를 이동할 수 있다는 점이다. 우리는 에픽게임즈에서 버추얼 프로덕션을 위해 공개한 멀티 유저 에디팅(Multi User Editing) 기술을 이용하여 실시간으로 다수의 작업자가 촬영 환경에서도 가상공간을 편집하는 것이 가능하다. 멀티 유저 에디팅 기능을 통해 동일한 네트워크 대역에 위치한 워크스테이션들이 하나의 장면에 참가하여 실시간으로 조명, 디지털 에셋 등을 제어할 수 있다.

3.2 디지털 휴면



Figure 6 Our scan to metahuman pipeline. Our pipeline extended from Song's research[4]

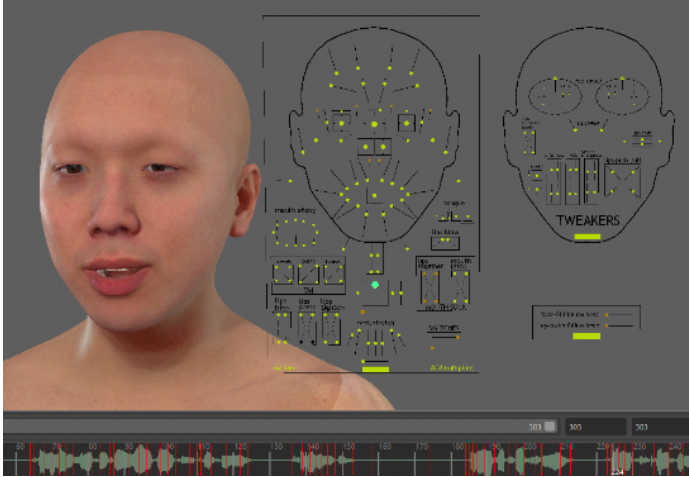


Figure 7 Speech animation generated by our lip-sync technology

[4]의 연구에서는 페이스 스캔 장비를 활용하여 디지털 더블을 제작하였으며, 이때 에픽게임즈의 메타휴먼 기술에 기반하여 페이스 리깅된 디지털 휴먼을 활용하였다. 이 연구에서는 페이스 스캔 장비를 통해 실존 인물의 얼굴을 디지털로 변환하고, 베이스 메쉬를 스캔 메쉬에 정합하는 Mesh Registration 과정을 소개하였다. 또한, 페이스 애니메이션을 위한 Rig Transfer 기술을 함께 다루었다. 우리의 사례에서는 디지털 휴먼의 얼굴뿐만 아니라 전체 몸의 형상이 필요하기 때문에 기존 연구의 방법을 그대로 활용하되 적용 범위를 몸으로 확장하여 사용하였다 (Figure 6).

한편 가상 배우인 디지털 휴먼이 대사를 말하려면 자연스러운 발화 애니메이션이 필요하다. 디지털 휴먼의 발화 애니메이션을 자동 제작하기 위해 우리는 [5]의 연구 기법을 활용하였다. 이 연구에서는 특정 언어의 발화에 필요한 대표적인 음소와 각각의 음소에 대응되는 입술 또는 혀 움직임의 시각적 표상인 독화소(Viseme)를 정의하였다. 우리의 사례에서는 메타휴먼의 페이스 리깅에 맞춰 한국어 독화소 구현에 필요한 리깅 컨트롤러와 파라미터 값을 사전에 정의한 후, 보다 개선된 동시조음 모델을 적용하여 독화소를 최

적화함으로써 음운환경에 따라 같은 음소도 다양한 입 모양으로 실현되도록 하였다 (Figure 7). 이 과정에서 음성과 텍스트를 입력받아 강제음성정렬 수행을 통해 음소 정렬 정보를 획득하였는데, 여기서 텍스트는 짧은 드라마 콘텐츠상에서 가상 배우의 대사를 의미하며, 음성의 경우 배우가 극중 상황에 어울리는 톤으로 연기하듯이 녹음하여 확보하였다.

4. 콘텐츠 촬영 사례

이전 장에서 언급하였듯이 우리의 버추얼 스튜디오는 크기가 작기 때문에 근경과 중경이 많이 배치되는 장면을 이용한 촬영은 불가능하다. 따라서 근경과 중경이 최대한 적게 등장하는 장면 연출이 필요하다. 일반적으로 우주 배경은 버추얼 프로덕션에서 많이 사용되는 소재 중 하나이다. 우주 배경은 (1) 배경이 단조롭고 (2) CG에 대한 거부감이 적으며 (3) 근경과 중경이 거의 등장하지 않아 세트 제작에 많은 비용이 소요되지 않는다. 이와 같은 특성은 소규모 버추얼 스튜디오에서도 쉽게 콘텐츠를 촬영할 수 있는 장점이 있기 때문에 우리는 달을 배경으로 하는 시나리오를 기획하여 본 촬영 사례에 활용하였다.

콘텐츠 촬영을 위한 전체 절차는 Figure 8과 같다. 시나리오 작성이 완료되면 시나리오에 따라 가상환경과 실제 물리 환경, 하드웨어를 구성한다. 하드웨어는 3.1장에서 언급한 방법에 따라 2개의 렌더링 머신 기반의 버추얼 스튜디오를 구성하였다. 본 촬영 사례는 기존 하드웨어의 구성 환경과 차이가 없기 때문에 동일한 환경을 사용하였으나, 시나리오 및 촬영 사례에 따라 3.1장에서 구성한 하드웨어에 모션 캡처 장비를 추가하거나 디스га이즈와 같은 미들웨어를 추가하여 XR을 활용할 수도 있다.

촬영 시에는 카메라 구도에 따라 LED 외부가 드러나야 하는 경우가 발생한다. 특히 사람의 다리를 촬영하거나 전신 구도를 촬영하는 경우 필연적으로 바닥이 노출될 수밖에 없다. 디스га이즈를 활용한 XR 기술은 LED 화면 외부를 가상 환경으로 채워줄 수 있으나 우리의 하드웨어 구성에는 적합하지 않으며, 디스га이즈를 운영하기 위한 전문인력이 추가로 필요하다는 한계점이 있다. 따라서 대부분의 스튜디오에

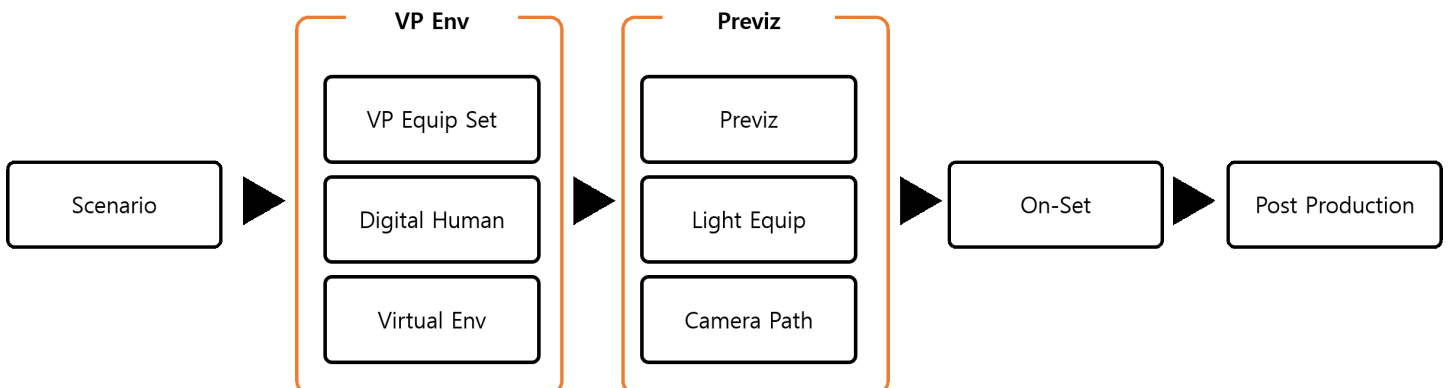


Figure 8 Content production pipeline in virtual production system

서는 LED와 연결된 바닥면은 실제 세트를 제작하여 사용한다. 우리도 본 촬영 사례를 위하여 LED를 벗어난 영역에 실제 물리 세트를 제작하였다. 우리는 달 표면을 묘사하기 위하여 페스티로폼을 활용하였으며, 가상공간 구성은 시나리오를 기반으로 제작된 스토리보드에 맞춰 제작하였다. 우주 기지 내부와 외부, 2가지 배경을 활용하되 외부 환경은 기지를 배경으로 하는 환경과 지구에서 달을 바라보는 환경, 총 2개의 환경을 제작하였고 3개의 카메라 동선을 사용하였다.

4.1 디지털 휴먼 배우 제작

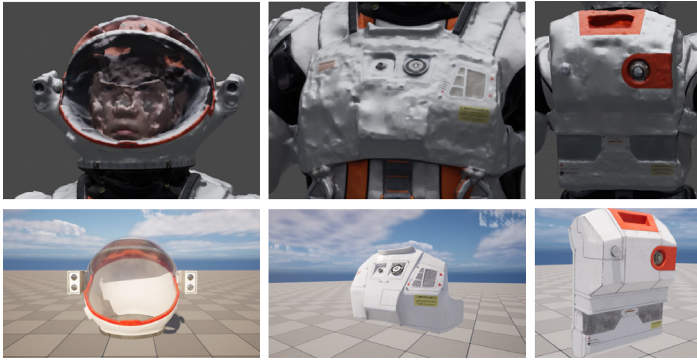


Figure 9 Results of refining digital human assets. (Top) Scan results (Bottom) Assets created by restoring lost detail

우리는 사실적인 디지털 휴먼을 제작하기 위해 우주복을 대여하여 포토그래메트리(Photogrammetry) 기반 풀바디 스캔(Full-Body Scan) 시스템을 이용하였다. 다만 포토그래메트리 기반 스캔 솔루션은 조명 반사가 심한 영역에 대해서는 정확한 데이터를 생성하지 못한다는 한계점이 있다. 편광필터를 적용하여 조명에 의한 하이라이트 부분은 일부 개선 가능하지만, 투명한 재질의 물체나 반사가 심한 부분, 플라스틱과 같이 단단한 물체 재질을 메쉬로 변환하는 데는 데이터 소실이 발생한다. 우리의 사례에서는 고품질의 디지털 휴먼 제작 시 (1) 디지털 휴먼의 얼굴은 [4]의 기법을 이용하여 제작하고, (2) 투명 재질이 포함된 헬멧은 모델링을 통해 에셋을 제작하였으며, (3) 하이라이트에 의해 소실된 단단한 재질의 에셋을 추가 제작함으로써 데이터 소실 및 저품질 메쉬를 개선하였다 (Figure 9).

모니터를 통해 보이는 색상 결과는 LED 설정에 따라 다른 색상으로 보이는 경우가 있다. 우리의 사례는 우주 배경이 강하고 방향성 있는 광원에 의해 강한 대비를 보여주는 특징이 있다. Figure 10은 후반 작업을 최소화하기 위해 디지털 휴먼의 머티리얼을 수정하여 실시간으로 변화하는 모습을 촬영함으로써 LED에서 보이는 디지털 휴먼을 사실적으로 렌더링할 수 있는 색상을 효율적으로 찾는 과정을 보여준다. 본 논문에서 사용한 디지털 휴먼의 최종 결과물은 Figure 11과 같다.

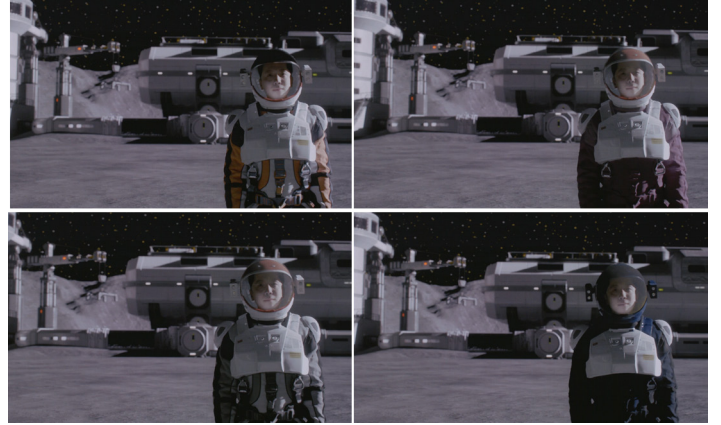


Figure 10 The process of efficiently finding colors that can realistically render a digital human on LEDs



Figure 11 The final result of digital human generated with our full-body scanning system, used in virtual production content

4.2 프리비즈 제작

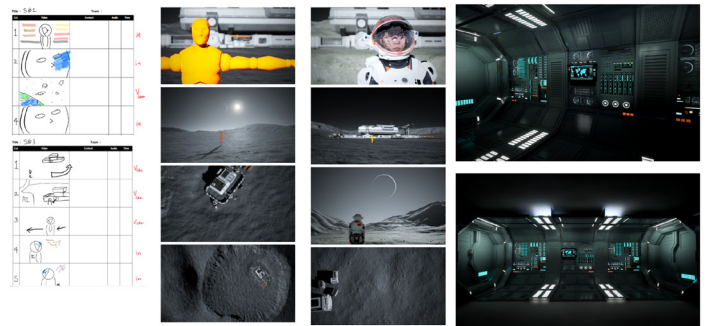


Figure 12 Our storyboard-based virtual environment

가상환경과 디지털 휴먼 제작이 완성되면 디지털 에셋을 활용한 프리비즈를 제작한다 (Figure 12). 프리비즈는 본 촬영에 들어가기 전에 카메라의 구도 및 배우의 동선을 확인하기 위해 제작하는 비교적 낮은 퀄리티의 3D 애니메이션

을 의미한다. 전체적으로 낮은 퀄리티의 에셋을 사용하기 때문에 전통적인 프로덕션 제작 방식에서는 프리비즈에 사용된 디지털 에셋을 그대로 활용하는 것이 불가능하며, 최종 결과물 산출을 위해 고품질 디지털 에셋을 추가로 제작해야 한다.

반면에 버추얼 프로덕션의 프리비즈는 전통적인 프로덕션 제작 방식의 프리비즈와 그 목적은 동일하지만, 기존에 존재하지 않던 LED에 가상환경 투사 구도 점검이 추가된다. 또한, 실시간 엔진을 통해 촬영되기 때문에 기존의 프리비즈와 같이 낮은 품질의 에셋을 사용하지 않고 최종 촬영에 사용될 수 있는 품질의 디지털 에셋을 배치 및 렌더링하여 제작한다. 따라서 한 번의 에셋 제작으로 프리비즈부터 최종 결과물 렌더링까지 활용할 수 있다.

4.3 송출 색감 보정

버추얼 프로덕션에서 가장 중요한 작업 중 하나가 실제 세트와 가상공간의 연속성을 만드는 것이다. 이러한 연속성은 가상공간의 색감과 실제 세트의 색감 조절을 통해 만들 수 있다. 하지만 제작이 완료된 세트의 색감을 제어하는 것은 불가능하기 때문에 (1) LED의 밝기, 대비, 색온도를 제어하거나 (2) 가상공간의 머티리얼을 제어하거나 (3) 조명을 통해 색감을 통일시키거나 (4) 카메라 속성을 이용하는 방법을 활용한다. 색감을 통일하기 위해 동일한 구도로 인물과 실사 물체를 배치하고, 가상환경도 구도의 변경 없이 색감 통일 작업을 진행한다.

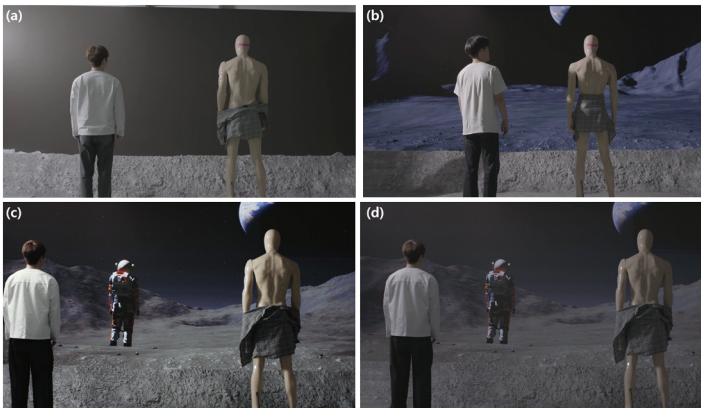


Figure 13 Results of color adjustment steps between a real object and virtual environment. (a) No virtual environment on LEDs (b) Original virtual environment displayed on LEDs (c) Adjust material color in virtual environment (d) Adjust the brightness and contrast of LEDs

Figure 13은 각 색상 조정 단계 결과를 보여준다. 모니터와 워크스테이션에서 보이는 가상환경이 사실적일지라도 LED에 송출하게 되면 색감이나 색온도에 따라 의도하지 않은 색상이 나타날 수 있다 (Figure 13 (b)). 제작된 가상환경의

머티리얼 정보를 수정하는 방법이 가장 직관적이지만 (Figure 13 (c)), 모든 촬영 장면에 따른 가상 물체의 재질을 변경해야 하는 번거로움이 있다. Figure 13 (d)는 LED의 온도를 6,500K, Brightness 80, Contrast 55로 조정한 결과이다. 전체적인 색감이 흐려지지만 가장 빠르고 효율적으로 전체 색감을 맞출 수 있는 장점이 있다.

4.4 카메라 초점 동기화



Figure 14 Moire effect caused by camera focus. (Top) Focus on the real person standing in front of LEDs (Bottom) Focus on the digital human on LEDs

모아레(Moire) 현상은 LED 기반 버추얼 프로덕션에서 쉽게 관찰되지만 해결하기 어려운 문제 중 하나이다. 최근 카메라 렌즈의 성능이 좋아짐에 따라 LED 표면에 초점을 맞출 경우 작은 LED 소자들이 더욱 두드러지게 나타나면서 모아레 현상이 심화되는 문제가 있다. Figure 14는 카메라의 포커스에 따른 모아레 현상의 차이를 보여준다. 상단의 이미지는 LED 앞에 서 있는 실제 사람에게 포커스를 맞춘 결과이며, 하단의 이미지는 LED에 렌더링된 디지털 휴면에 포커스를 맞춘 결과이다. Figure 14에서 볼 수 있듯이 실제 사람에게 포커스를 맞춘 경우 후면 LED에는 모아레가 거의 생기지 않는 반면, LED에 렌더링되는 디지털 휴면에 포커스를 맞춘 경우 디지털 휴면 뒤쪽 배경에 강한 모아레가 생기는 것을 확인할 수 있다.

모아레를 해결하기 위해서는 카메라 렌즈와 LED 사이에 거리를 충분히 확보하는 것이 중요하다. 하지만 본 사례에서는 스튜디오의 전체 크기가 7m x 7m로 작기 때문에 렌즈와 LED 사이에 거리를 확보하는 데 한계가 있다. 또한, 거리를 확보하더라도 초점 거리 조절에 따라 렌즈의 포커스를 LED에 맞추게 되면 모아레가 발생하기 때문에 LED 내 렌더링되는 디지털 휴면을 촬영하는 데 심도 조절이 불가능하다. 우리는 LED 내 디지털 휴면의 심도를 표현하고 촬영

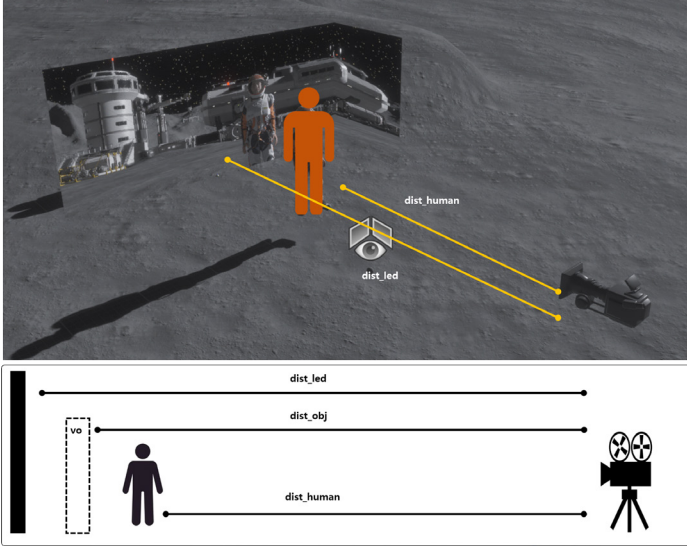


Figure 15 Placing a real person and a digital human in the same position

거리 제약을 완화하기 위하여 실제 초점 거리 조절 기능과 가상 카메라의 초점 거리 기능을 연동하여 사용하였다. 가상 카메라의 심도계와 실제 카메라의 심도계를 연동하기 위해서는 가상 카메라와 실제 카메라의 위치를 동일하게 배치해야 한다 (Figure 15 (Top)).

Algorithm 1 Synchronized camera position

Require: Real world camera distance D

if $D \leq dist_obj$ **then**

Control real world camera focus

else

Control virtual world camera focus

end if

Algorithm 1은 가상 카메라와 실제 카메라의 초점 거리 연동을 보여준다. 두 카메라의 위치가 동일하다는 가정하에 실제 사람이 배치된 위치와 LED 사이에 보이지 않는 가상의 물체(vo)를 가정한다 (Figure \ref{fig:camera_sync} (Bottom)). 실제 사람에서 LED에 렌더링된 디지털 휴먼으로 초점을 이동시킬 때, vo를 기준으로 초점 거리(dist)가 실제 사

람과의 거리(dist_human)보다 작으면 실제 카메라의 초점을 조절하고, 크거나 같다면 가상 카메라의 초점 거리를 조절한다.

5. 결과

본 논문에서 사용된 주요 장비의 사양은 Table 1에 정리하였다. 촬영 데이터 보관을 위해 모든 워크스테이션은 Samsung M.2 2TB의 HDD를 사용하였으며, 네트워크를 통해 렌더링 및 카메라 데이터 송수신을 수행해야 하기 때문에 10 G 네트워크 카드를 사용하였다. 본 촬영을 위해 에픽게임즈 언리얼 엔진 5.1 버전을 사용하였으며, 모든 워크스테이션은 동일한 프로젝트 파일을 보유해야 하기 때문에 SVN을 이용하여 프로젝트 버전을 관리하였다.

촬영 현장의 조명 제어는 아리(ARRI)사의 스카이 패널을 3개 대여하여 사용하였으며 촬영에 총 6시간이 소요되었다. 촬영에 투입된 전체 작업 인원은 총 12명으로 일반적인 버추얼 프로덕션에 비해 소규모 인원으로 촬영이 가능하였다. Figure 16은 본 논문에 소개된 초소형 버추얼 스튜디오의 온셋 현장을 보여준다.



Figure 16 Virtual production content filming site in our tiny studio

Equipment	Type	CPU	RAM	GPU
Workstation	Rendering Machine	i7-9700KF	64GB	Quadro RTX A6000
	Edit	AMD Ryzen 9 7950X	128GB	Geforce RTX 4090
	Operator	i7-9700KF	64GB	Quadro RTX A4000
	Camera tracking	i9-10940X	128GB	Geforce RTX 3080
	Position	Pixel pitch	# of panels	
LED	Left Display	1.25mm	42	
	Right Display	1.25mm	42	

Table 1 : Hardware specification for our virtual production system

6. 결론

본 논문은 초소형 버추얼 스튜디오를 구축하고 공간 크기에 맞는 콘텐츠를 기획 및 제작함으로써 버추얼 프로덕션 공정의 효율성을 검증하고, 초소형 스튜디오에서 촬영 가능한 카메라 기술을 제안하였다. 본 촬영 사례를 통해 버추얼 스튜디오에서 촬영 현장 크기에 따라 촬영 가능한 콘텐츠의 종류 및 범위가 달라지는 것을 확인하였다. 비록 스튜디오의 공간상 제약으로 인해 다양한 각도의 촬영 구도를 사용하거나 카메라의 움직임을 다양하게 주는 것이 불가능하였지만, 작은 공간에서 버추얼 프로덕션 콘텐츠를 촬영하기 위한 파이프라인을 구축할 수 있었다. 또한, 스캔 시스템을 활용하여 디지털 휴면을 영상에 등장시키기 위한 효율적인 방법론을 제안하였으며, 기존에는 시도하지 않았던 실제 사람과 디지털 휴면의 동시 촬영을 시도함으로써 디지털 휴면을 사실적으로 표현할 수 있는 LED 색상 제어 방법을 제안하였다.

우리의 사례는 5분 영상을 촬영하기 위해 6시간을 소요하였다는 점에서 촬영 시간을 단축하고 세트 교체 시간 및 비용을 절감한다는 버추얼 프로덕션 본연의 목적에 부합하지는 못하였다. 그러나 초소형 스튜디오를 활용한 콘텐츠 시범 제작 사례를 도출함으로써 버추얼 프로덕션 운영에 대한 경험과 기술을 습득하였다. 대규모 스튜디오는 구축에 많은 자본이 필요하기 때문에 국내 중소 스튜디오에서 버추얼 스튜디오를 운영하기에는 어려움이 있으나, 본 사례를 통해 한정된 장면 및 구도를 활용한 촬영 기법을 소개함으로써 중소 스튜디오에서도 버추얼 스튜디오를 구축하고 운영할 수 있는 기술적 증빙을 수행하였다.

향후 우리의 초소형 버추얼 스튜디오를 활용하여 작은 공간을 효율적으로 활용할 수 있는 다양한 콘텐츠를 기획하고 제작할 것이다. 또한, 카메라 제어, 색상 보정 자동 기술 등을 통해 콘텐츠 제작의 효율을 극대화하는 기술을 개발함으로써 버추얼 프로덕션 기술력을 확보할 계획이다.

감사의 글

이 논문은 2022년도 정부(문화체육관광부)의 재원으로 한국콘텐츠진흥원의 지원을 받아 수행된 연구임 (No. R2022020123, 가상공연 참여를 위한 감정 표현이 가능한 소통형 관객 아바타 생성 기술 개발)

References

- [1] V. Helzle, S. Spielmann, and J. Trottnow, “Green screens, green pixels and green shooting,” in ACM SIGGRAPH 2022 Talks, ser. SIGGRAPH ’22. New York, NY, USA: Association for Computing Machinery, 2022. .
- [2] C. Payne and F. L. Giardiello, ““openvpcal”: An open

source in-camera visual effects calibration framework,” in ACM SIGGRAPH 2022 Talks, ser. SIGGRAPH ’22. New York, NY, USA: Association for Computing Machinery, 2022.

- [3] D. Ring, K. T. Haug, P. Verga, D. Stump, N. Redmond, D. Caffrey, and P. Canning, ““comandante”: Braving the waves with near real-time virtual production workflows,” in ACM SIGGRAPH 2022 Talks, ser. SIGGRAPH ’22. New York, NY, USA: Association for Computing Machinery, 2022.
- [4] J. Song, “Digital double creation with unreal metahuman and facial scan,” 한국컴퓨터그래픽스학회 학술대회, pp. 105-106, 2022.
- [5] J. N. Minjung Jang, Sunjin Jung, “Speech animation synthesis based on a korean co-articulation model,” 한국컴퓨터그래픽스학회 논문지, vol. 26, no. 3, pp. 49-59, 2020

〈 저자 소개 〉



임재호

- 2003 ~ 2010 숭실대학교 글로벌미디어학부 학사
- 2010 ~ 2012 고려대학교 컴퓨터전파통신공학 석사
- 2012 ~ 2017 고려대학교 컴퓨터전파통신공학 박사
- 2018 ~ 2019 ㈜투썸디지털아이디어 수석연구원
- 2019 ~ 현재 ㈜엔진비주얼웨이브 R&D연구소 부소장
- 관심분야 : Virtual production, Digital human platform, Physics based simulation, Natural phenomenon modeling
- <https://orcid.org/0000-0001-8797-2627>



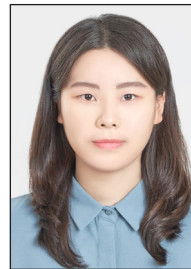
박민수

- 2004 ~ 2010 한경국립대학교 생명공학부 학사
- 2013 ~ 2016 서울과학기술대학교 건설시스템공학부 석사
- 2017 ~ 2019 디지털아이디어
- 2019 ~ 2022 씨제스컬리버스튜디오
- 2023 ~ 현재 엔진비주얼웨이브
- 관심분야 : Computer graphics, Virtual reality
- <https://orcid.org/0009-0000-3417-2597>



장민정

- 2014 ~ 2018 고려대학교 국어국문학과/미디어학부 학사
- 2018 ~ 2020 한국과학기술원 문화기술대학원 석사
- 2021 ~ 현재 ㈜엔진비주얼웨이브 연구원
- 관심분야: Facial animation, Character animation
- <https://orcid.org/0000-0002-7095-8375>



김유진

- 2016 ~ 2019 국가평생교육진흥원 관광학 관광학사
- 2022 ~ 현재 ㈜엔진비주얼웨이브
- 관심분야 : Virtual reality and contents
- <https://orcid.org/0009-0007-6521-1377>



전상욱

- 2008 ~ 2015 숭실대학교 글로벌미디어학부 학사
- 2015 ~ 2018 UdK (Universitat der Kunste)
- 2019 ~ 2020 (주)디스트릭트코리아
- 2020 ~ 현재 (주)엔진비주얼웨이브 Techincial Artist
- 관심분야 : Virtual Production, Unreal Engine
- <https://orcid.org/0009-0008-1665-7839>



이수빈

- 2011 ~ 2017 이화여자대학교 국제사무학/경영학 학사
- 2022 ~ 현재 ㈜엔진비주얼웨이브 Techincial Artist
- 관심분야 : Computer graphics, Virtual production
- <https://orcid.org/0009-0000-6802-6432>