석사학위논문

확장현실 협업 환경을 위한 장면 합성 프레임워크 설계

- 현실 장면 기반의 가상 장면 생성 -

2024년

한 성 대 학 교 대 학 원

컴 퓨 터 공 학 과

컴 퓨 터 공 학 전 공

나 기 리

석 사 학 위 논 문
지도교수 김진모

확장현실 협업 환경을 위한 장면 합성 프레임워크 설계

- 현실 장면 기반의 가상 장면 생성 -

Framework of Scene Synthesis for Collaborative XR: From MR-based Real Scenes to VR-based Virtual Scenes

2024년 6월 일

한 성 대 학 교 대 학 원

컴 퓨 터 공 학 과

컴 퓨 터 공 학 전 공

나 기 리

석 사 학 위 논 문
지도교수 김진모

확장현실 협업 환경을 위한 장면 합성 프레임워크 설계

- 현실 장면 기반의 가상 장면 생성 -

Framework of Scene Synthesis for Collaborative XR: From MR-based Real Scenes to VR-based Virtual Scenes

위 논문을 공학 석사학위 논문으로 제출함

2024년 6월 일

한 성 대 학 교 대 학 원

컴 퓨 터 공 학 과

컴 퓨 터 공 학 전 공

나 기 리

나기리의 공학 석사학위 논문을 인준함

2024년 6월 일

심사위원장 계 희 원 (인)

심 사 위 원 허 준 영 (인)

심 사 위 원 김 진 모 (인)

국 문 초 록

확장현실 협업 환경을 위한 장면 합성 프레임워크 설계

한 성 대 학 교 대 학 원
컴 퓨 터 공 학 과
컴 퓨 터 공 학 전 공
나 기 리

현실을 기반으로 하는 혼합현실 사용자와 컴퓨터 그래픽을 통해 생성된 환경을 기반으로 하는 가상현실 사용자가 정교하고 유기적인 협업 및 상호작용을 수행하기 위해서는 현실과 가상이 동일한 좌표를 기준으로 대응되는 공간과 장면 생성이 필요하다. 본 연구는 현실 장면으로부터 가상 장면을 정확하고 효과적으로 생성하는 장면 합성 프레임워크를 제안한다. 제안하는 프레임워크는 현실 공간을 구성하는 3차원 구조와 객체의 구성 정보를 파악하기 위한 기하학적 분석, 객체의 종류와 특징을 분류하여 가상 장면 생성에 필요한 정보를 계산하기 위한 의미론적 해석으로 구성된다. 현실 공간의 3차원 기하학적 정보를 분석하기 위해 마이크로소프트 홀로렌즈 2 기기와 Scene Understanding SDK((Software Development Kit)를 사용하여 현실 장면을 스캔하고, 공간을 구성하는 객체의 메시 정보를 추출한다. 또한, Azure 커스

텀 비전을 활용하여 객체의 종류를 인식하고 분류하는 의미론적 해석을 수행한다. 현실 공간을 구성하는 객체의 종류를 태그로 분류하고, 분류된 모든 객체의 기하학적 정보를 연결하는 라벨링 작업을 수행한다. 현실 공간을 구성하는 객체 데이터를 활용하여 3차원 가상 객체를 자동으로 배치하고, 회전과 크기에 대한 미세 조정을 거쳐 현실과 대응되는 가상 장면을 생성한다. 이를 기반으로 제안하는 프레임워크의 효율성과 성능을 검증하기 위해 작업별로 시간을 측정하였다. 결과적으로 각 단계별 10분 내외의 시간으로 현실을 기반으로 하는 가상 장면을 효과적으로 생성할 수 있음을 확인하였다. 또한, 본 연구는 현실을 기반으로 정교하게 설계된 가상 장면으로부터 혼합현실 사용자와 가상현실 사용자가 같은 공간에 함께 참여하여 협업하고 상호작용할 수 있는 체험 환경을 제시하는 것을 목표로 한다. 이를 위해 크리에이티브 스튜디오 공간에서 교육 활동을 목적으로 하는 콘텐츠를 제작하고, 제안하는 프레임워크를 통해 제작된 가상 환경이 가상현실 사용자의 경험에 미치는 영향을 분석하기 위한 설문 실험을 진행하였다. 실험을 통해 현실을 기반으로 만들어진 가상 장면이 몰입, 집중 그리고 적응 등 사용자의 경험에 긍정적인 영향을 미친다는 것을 확인하였다.

【주요어】 장면 합성, 홀로렌즈, 확장현실, 디지털 트윈, 가상 환경

목 차

표 목 차

그 림 목 차

Ⅰ. 서 론

증강현실(Augmented Reality)은 실제 환경 위에 가상 사물이나 정보를 합성하여, 실제 사물처럼 보이도록 하는 기술이다. 혼합현실(Mixed Reality)은 여기서 더 나아가 가상과 현실 정보를 결합하여 융합시키는 공간을 만드는 기술로, 현실과 상호작용할 수 있다는 증강현실의 장점에 몰입감을 전해줄 수 있는 가상현실의 장점을 함께 살려 진화된 가상 세계를 만드는 기술이다. 현실을 기반으로 가상 환경, 정보, 그래픽 등을 합성하여 몰입감 높은 새로운 체험 환경을 제공하는 증강현실, 혼합현실 기술은 교육, 의료, 공학 등 다양한 영역에 적용되고 있다(Zhu et al., 2014). 특히 혼합현실을 활용한 애플리케이션은 가상 세계와 현실 세계를 자연스럽게 혼합하고, 두 세계 사이의 유기적인 상호작용을 통해 향상된 몰입과 경험을 제공하는 것을 목표로 한다. 이를 위해서는 실제 세계의 물리적 공간을 디지털화하여 가상 공간과 현실 장면을 정확하게 연결할 수 있는 기술이 필요하다. 기존의 방법론이나 저작 도구로는 추적 기반, 마커 기반, 또는 지오메트리 기반 등의 방법이 있으며, 이는 현실 세계의 주변 환경을 파악하고 가상 공간과의 관계를 적절히 조정한다. 최근에는 의미론적 참조를 이용해서 증강현실 경험을 생성할 수 있는 ScalAR과 같은 연구가 진행되기도 하였다(Qian et al., 2022).

현실 세계의 디지털화는 가상현실, 혼합현실을 아우르는 확장현실 기술의 발전과 더불어 빠르게 진행되고 있다. 풍경, 도시, 건축 등 실외 환경은 물론 박물관, 문화 공간과 같은 실내 환경까지 디지털화되고 있다. 이는 접근성, 경제성 등의 제약으로 인해 현실 세계에 직접 방문하는 것이 어려운 상황이 많기 때문이다. 이러한 이유로 디지털 트윈을 기반으로 실제 공간과 가상 공간을 함께 탐색할 수 있는 기술도 연구되고 있다. 이때, 가상현실 사용자는 실제 공간 정보를 기반으로 하는 가상 공간을 자유롭게 탐색하고 체험할 수 있지만, 스마트폰을 활용한 증강현실 사용자는 비교적 표현에 제약이 많다. 따라서, 실제와 가상을 결합한 공간에서 협업을 고려한 체험 환경을 구축하기 위해서는 혼합현실 장비를 사용하는 것이 적합할 수 있다.

디지털화된 가상 환경은 가상현실, 증강현실 그리고 혼합현실까지 아우르는 확장현실 기술을 기반으로 사용자가 가상 환경과 유연하고 직관적이며, 현실과 유사하게 상호작용할 수 있는 체험 환경으로 연구되어 발전되고 있다. 또한, 다수의 몰입형 사용자가 협업하는 협업 가상 환경에 관한 연구(Carlsson et al., 2002)와 함께 비몰입형 사용자와 가상현실 사용자, 증강현실 사용자와 가상현실 사용자 등이 같이 원격으로 상호작용하는 비대칭 가상 환경에 관한 연구도 현재까지 폭넓게 진행되고 있다(An et al., 2023). SF(Science Fiction) 문학에서 유래한 확장 가상 세계인 메타버스(Metaverse)를 확장현실로(eXtended Reality) 구현하는 과정에서 더욱 향상된 몰입과 함께 다양한 플랫폼의 사용자가 참여하는 협업 확장현실에 관한 연구들도 다양한 산업 응용을 고려하여 진행되고 있다. 특히, 현실과 가상을 연결하는 디지털 트윈 기반에서 현실 세계 사용자와 가상 세계 사용자가 함께 협업하기 위해서는 동일한 공간과 장면을 생성하는 것이 필요하다. 가상 장면을 생성하기 위해서는 3Ds Max, Blender와 같은 그래픽 소프트웨어를 통해 디자이너가 직접 제작하는 방법이 일반적이지만, 현실에 기반한 정교하고 정확한 작업을 위해서는 많은 시간과 노력이 요구된다. 이러한 이유로 장면 합성을 위한 다양한 연구들이 진행되었다. 이에 관련하여 확률론적 최적해 기법 등의 방법(Yu et al., 2011), 생성형 모델을 비롯한 딥러닝을 활용한 방법(Zhang et al., 2020) 등이 있다. 하지만 대부분의 장면 합성에 관한 연구들은 가상 장면 생성에 집중되어 있고 현실을 충분히 고려하지 않는다. 이와 달리, 현실 세계의 물리적 공간을 실시간으로 감지하고 보행 가능한 영역에 맞춰 가상 공간을 인스턴스화 하여 현실에 기반한 장면을 생성하는 연구 등도 진행되었다(Cheng et al., 2019). 하지만 이는 가상현실 사용자만을 고려한 것이며, 혼합현실 사용자와 가상현실 사용자가 함께 참여하는 협업을 지원하는 장면 생성은 아니다.

본 연구는 혼합현실 기반의 현실 장면으로부터 가상현실 사용자가 함께 참여하는 가상 장면을 정확하고 효과적으로 생성하는 프레임워크를 제안한다. 현실 세계에서 참여하는 혼합현실 사용자의 체험 공간은 사용자를 기준으로 하는 상대적 좌표를 기준으로 이루어진다. 이에 반해 가상현실 사용자는 가상

공간의 정의된 원점을 기준으로 공간을 판단하기 때문에 접속 시점의 사용자 위치에 영향을 받지 않는다. 따라서 두 사용자의 원활한 상호작용을 지원하는 확장현실 협업 환경을 위해서는 같은 고정 좌표를 기준으로 장면을 해석하는 과정이 요구된다. 이를 위해, 마이크로소프트(Microsoft)의 Scene Understanding SDK를 활용하여 현실 세계의 상대적 좌표와 3차원 기하학적 정보를 분석하고, 이를 바탕으로 가상 공간에서의 기하학적 정보를 계산하여 두 공간의 좌표를 일치시킨다. 또한, 현실과 동일한 가상 장면을 만들기 위해 현실 세계를 이루는 객체들의 세부적인 정보를 바탕으로 가상 장면을 구성해야 한다. 이를 위해, Azure 커스텀 비전(Custom Vision)을 활용하여 현실 공간을 이루는 객체들을 인식 및 추론하는 과정을 수행한다. 그리고 사전에 분류된 가상 객체들을 현실 객체에 맞춰 자동으로 배치하여 현실과 유사한 연출의 가상 장면을 생성하는 프레임워크로 정리한다. 협업 확장현실을 위해 정리된 프레임워크를 활용해서 현실과 좌표계가 일치하는 가상 장면을 구성하고, 하나의 공간으로 대응시킨 체험 환경을 구축한다. 혼합현실 사용자를 중심으로, 가상현실 사용자가 함께 참여하고 의사소통하는 과정에서 향상된 몰입과 효과적인 정보 전달을 목표로 한다. 이는 접근성, 경제성 등의 이유로 쉽게 경험하지 못하는 환경을 전제로 하여 현실 기반의 가상 세계를 구성하고 가상현실 HMD(Head Mounted Display)를 통해 사용자들이 참여할 수 있도록 한다. 그리고 현실 세계에 존재하는 혼합현실 사용자는 실제와 가상이 동기화된 객체를 조작 및 제어함으로써 가상현실 사용자에게 사실적인 정보 전달을 가능하게 한다. 본 연구의 기여는 다음과 같다. 본 연구는 현실과 가상 공간이 동일 좌표 기준으로 해석 가능하도록, 현실 세계의 공간과 객체 정보를 분석하여 현실과 대응되는 가상 장면을 빠르고, 효과적으로 생성하는 프레임워크를 설계한다. 그리고 정량적 실험을 통해 홀로렌즈(HoloLens) 2를 사용해서 현실 세계를 분석하는 과정에서부터 현실과 대응되는 가상 장면이 생성되기까지의 작업 시간을 분석하여 효율성을 확인한다. 그리고 현실과 가상 공간이 대응되는 협업 확장현실 체험 환경을 구축한다. 홀로렌즈 2를 통해 현실 세계에서 참여하는 혼합현실 사용자와 현실 기반으로 생성된 가상 공간에 메타 퀘스트 2 HMD를 통해 참여하는 가상현실 사용자가 함께 공유할

수 있는 체험 환경으로 제작한다. 현실을 기반으로 하는 혼합현실 사용자는 현실과 가상이 동기화된 객체를 조작, 제어하는 과정을 통해 정의된 정보를 가상현실 사용자에게 전달하고, 현실을 기반으로 하여 체험하는 가상공간에 참여하는 가상현실 사용자를 대상으로 제작한 체험 환경이 사용자의 경험에 유의미한 차이를 미치는지 확인하기 위해 정성적 실험을 진행하였다. 그리고 본 연구는 정량적 실험과 정성적 실험을 통해 제시하고 있는 프레임워크와 제작한 확장현실 체험 환경의 장단점 및 특징 등을 명확하게 분석하고자 한다.

Ⅱ. 관련 연구

2.1 비대칭 협업 가상 환경

스마트폰의 발전에 따른 모바일 가상현실과 모바일 증강현실 기기의 발전, 다양한 가상현실 HMD의 보급으로 인해 여러 사용자가 동일한 가상 공간에서 각기 다른 특성과 참여 방식 등을 고려하여 다양한 형태로 협업 및 상호작용하는 가상 환경에 관한 연구가 진행되고 있다(Grandi et al., 2018). 모바일 증강현실에서 3D 객체를 공동으로 조작하기 위한 핸드헬드 기반의 인터페이스를 제안하는 연구나 가상 환경에서 원격 사용성 테스트를 위한 새로운 방법을 제안하고, 그 효율성과 몰입도 높은 경험, 생산성을 분석하는 연구가 수행되었다(Madathil et al., 2017). 여러 사용자가 동일한 기기, 상호작용 방법을 사용하는 대칭형 가상 환경과 달리 서로 다른 플랫폼의 사용자가 협업 환경에서 다양한 방식으로 상호작용하는 비대칭 가상 환경에 관한 연구도 주목받고 있다. 가상 객체에 대한 공동 조작 작업에서 모바일 증강현실과 가상현실 사용자가 협업하는 환경에서의 성능을 분석하는 연구와 같이 비대칭 협업 환경에 대한 접근이 진행되고 있다(Grandi et al., 2019). 여기에는 사용자의 기기와 플랫폼, 상호작용 방식의 차이뿐만 아니라, 사용자의 특성에 맞는 서로 다른 능동적 역할을 부여하여 차별화된 경험과 향상된 존재감을 제공하는 새로운 접근법의 연구들도 포함된다. 더불어, 비대칭 협업 환경(증강현실과 가상현실)에서 여러 시각화(FoV(Field-of-View) frustum, Eye-gaze ray, Head-gaze ray)가 공동 작업자의 주의와 행동에 중요한 정보를 제공하는지를 분석하는 사용자 연구도 진행되었다(Piumsomboon et al., 2019). 모바일 증강현실 사용자와 가상현실 사용자가 공동 작업을 수행할 때 각자의 입력 방식에 적합한 상호작용 및 인터페이스를 설계하는 연구가 있었고(Cho et al., 2021), 최근에는 비몰입형, 몰입형 사용자뿐만 아니라 모션 캡처(Motion Capture)를 활용한 가상 아바타까지 함께 참여하는 확장된 비대칭 가상 환경

구축을 위한 개발 환경 및 템플릿을 제안하는 연구도 진행되었다(Cho et al., 2023). 이러한 연구들을 토대로 비대칭 가상 환경을 다양한 산업 분야에 응용할 수 있는 방향으로 제시하기도 한다(Feld et al., 2021). 또한 협업 확장현실 환경에서 디지털 트윈을 통한 원격 유지 관리 지원 등 산업용 메타버스 애플리케이션의 특성과 활용 방향도 제시되었다. 본 연구는 혼합현실 사용자와 가상현실 사용자가 함께 참여하는 비대칭 가상 환경에서 현실과 가상이 정확하게 대응되는 디지털 트윈 기반의 몰입형 가상 환경을 구축하는 것을 목표로 한다.

2.2 장면 합성

디지털 트윈 관점에서 협업 확장현실 환경을 구성하기 위해서는 현실에 기반하여 같은 공간과 경험을 공유할 수 있는 정확하고 효과적인 가상 장면을 생성하는 방법이 필요하다. 그러나 다양한 상황, 배경, 목적을 가진 가상 장면을 매번 Blender나 3Ds Max와 같은 저작도구를 이용하여 제작하는 것은 경제적이지 않을 뿐 아니라 시간적으로도 많은 비용이 발생한다. 이러한 이유로 MCMC(Markov chain Monte Carlo) 알고리즘과 같은 확률적 표본 생성 방법 및 전역 최적화 문제를 계산하기 위한 다양한 알고리즘 등을 활용하여 실내외 장면 합성 자동화에 관한 연구가 이루어지고 있다. 가구 배치의 자동 최적화를 제안한 연구에서는 Metropolis-Hastings 상태 검색 단계를 활용하여 모의 담금질(simulated annealing)을 통해 비용 함수를 최적화하여 자연스러운 장면 합성을 진행하였다. 또한, 몬테카를로 샘플러(Monte Carlo sampler)를 사용하여 밀도 함수를 빠르게 샘플링하여 레이아웃을 생성하거나(Merrell et al., 2011), 인코딩하는 새로운 MCMC 알고리즘을 제안하여 오픈 월드 레이아웃 합성 문제에 대한 가능성을 탐색하기도 하였다(Yeh et al., 2012). 이밖에도, 확률론적 최적화 방식의 비효율적 구조를 해결함을 목적으로 물리학 기반의 연속 레이아웃 합성 기술을 제안하는 연구와(Weiss et al., 2018), 딥러닝을 응용하여 객체의 크기나 모양 같은 속성을 기반으로 객체를 규칙적으로 재배열하기 위한 기술을 제안하는 연구가 진행되었다(Zhang et

al., 2020). 이와 더불어, 인간의 움직임과 일치하는 가구 레이아웃을 생성하는 실내 장면의 생성 모델인 MIME(Mining Interaction and Movement to infer 3D Environments)을 제안하는 연구가 수행되었다(Yi et al., 2023). 최근에는 발전된 방식으로, 충분한 여유 공간을 확보하여 실내 장면 레이아웃을 최적화하기 위해 심층 강화 학습을 적용한 연구도 진행되었다(Sun et al., 2024). 하지만, 장면 합성과 연관된 대부분의 방법들은 생성되는 가상 장면의 결과(객체 배치, 레이아웃 등)에만 주목하고 있다. 제안하는 프레임워크는 현실과 동일한 객체 배치, 레이아웃 합성을 목적으로 현실 정보를 분석하여 가상 장면을 자동으로 생성하는 구조를 포함하고자 한다.

2.3 증강현실 경험을 위한 현실 정보 분석

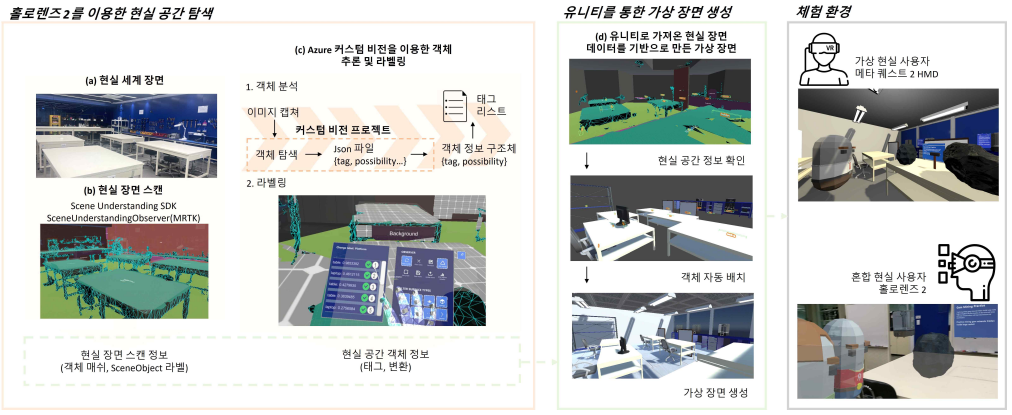
증강현실 경험은 실제 세계에 가상 정보를 합성하여 디지털 증강을 직관적으로 제공한다. 현실에 기반하여 가상 객체와 정보를 합성하기 위해서 추적기반, 마커기반, 기하학 기반, 의미기반 등의 방법이 활용된다. 추적기반 증강현실 경험에서는 모바일 증강현실에서의 3D 스케치를 통해 현장에서 3D 개념 설계를 쉽게 할 수 있는 Mobi3DSketch와 같은 시스템이 제안되었다(Kwan et al., 2019). 마커 기반 증강현실 경험은 중간 유형의 매체(마커, 이미지 타깃 등)를 통해 가상 공간과 실세 세계를 연결하는데, 이를 위해 대표적인 개발 도구로는 AR Core와 Vuforia가 있다. 이와 관련하여, 마커의 감지 속도를 높이기 위한 시스템도 연구되었다(Francisco et al., 2018). 기하학 기반 증강현실 경험은 현실 공간의 물리적 객체의 기하학적 특징을 추출하여 이를 기반으로 가상 환경과 결합하는 방법이다. 의미 기반의 증강현실 경험은 물리적 객체의 의미와 공간적 속성을 바탕으로 증강현실을 렌더링하는 방법이다. Lang(Lang et al., 2019) 등은 혼합현실 기기를 통해 가상 에이전트와 상호작용할 때 실제 장면의 의미를 고려하여 사용자와 상호작용할 수 있는 가상 에이전트의 위치를 지정하는 새로운 방법을 제안했다. 또한, Qian(Qian et al., 2022) 등은 가상현실에서 의미적으로 적응 가능한 증강현실 경험을 위하여 ScalAR이라는 통합 워크플로우를 제안했다. 이러한 연구들은 현실 정보

를 분석하여 증강현실 경험을 향상시키는데 기여하고 있다.

현실의 한정된 공간에 대해서 인지하고 일부 객체에 대한 의미만을 해석해서는 혼합현실 사용자와 가상현실 사용자가 동일한 공간에서 협업 및 의사소통 하는데 제약이 따를 수 있다. 즉, 현실에 기반하여 같은 공간에 존재하고 있다는 경험을 제공하기 위해서는 공간 전체에 대한 해석과 협업, 상호작용에 필요한 객체의 의미를 함께 분석하는 과정이 필요하다. Sra(Sra et al., 2017)는 실내 장면을 3D로 캡처하고 걷기 가능한 영역을 매핑하여 가상 환경을 생성함으로써 실제 걷기를 가능하게 하는 Oasis를 제안하였다. Tian(Trian et al., 2023)은 가상 복제품이 작업 공간의 3D 재구성을 통해 혼합현실의 원격 협업을 향상 시킬 수 있는 방법을 탐색하였다. 이밖에도 디지털 트윈을 사용하여 가상현실에서 증강현실 데이터를 시각화하는 Corsican Twin을 제안하기도 하였다. Ning and Pei(Ning et al., 2024)는 물리적 공간에 따라 가상현실 장면을 동적으로 재배치하고, 물리적 제약과 상호작용 작업을 원활하게 하는 새로운 접근 방식으로 연구하였다. 이러한 연구들은 양방향 혼합현실 시스템의 설계 공간을 탐구하는 Loki 시스템을 제안하거나(Kumaravel et al., 2019), 증강현실과 가상현실을 결합하여 현장 사용자와 원격 사용자가 비대칭 가상 환경을 통해 물리적 공간과 디지털 트윈을 동시에 탐색할 수 있는 방향을 제시하기도 하였다(Schott et al., 2023). 본 연구는 혼합현실 사용자와 가상현실 사용자가 동일한 공간에서 함께 원격으로 참여하여 협업 및 상호작용하기 위해서 현실과 가상이 대응되는 장면 합성과 함께 실험적 협업 확장현실 환경으로 교육용 비대칭 가상 환경을 제안하고자 한다.

Ⅲ. 장면 합성 프레임워크 설계

장면 합성 프레임워크는 현실을 기반으로 가상 장면을 효과적으로 생성하기 위한 일련의 과정이다. 이는, 유기적으로 연결되어 있는 확장현실 협업 환경을 제작하는 것을 목적으로, 현실과 가상의 공간이 동일 좌표를 기준으로 형성되도록 현실 공간의 기하학적 정보를 분석하고, 현실 공간에 있는 객체들을 의미론적으로 해석하여 정보를 수집하는 과정을 포함한다. 현실 공간에서 가상 장면을 생성하기 위해 요구되는 현실 공간의 장면 정보와 객체 정보를 획득하기 위해 홀로렌즈(HoloLens) 2 혼합 현실 장치를 사용했고, 해당 데이터는 유니티 3D(Unity 3D) 엔진으로 통합하여 가상 장면을 생성하는데 사용되었다. 또한, 현실 장면의 공간 추론을 통해 기하학적 형태로 반환된 데이터를 얻기 위해서 마이크로소프트의 공식 문서를 참고하여 Scene Understanding SDK를 활용했다. 기하학적 분석을 완료한 후, 현실 공간을 구성하는 각 객체에 대한 세부 정보를 탐색하기 위해 Azure의 AI 서비스 중 하나인 커스텀 비전(Custom Vision)을 사용하여 객체에 대한 의미론적 해석을 통해 필요한 객체 정보를 계산한다. 이렇게 추출된 현실 공간의 기하학적 정보와 의미론적 객체 정보는 각각 정해진 형식의 파일로 추출되고 이를 다시 유니티 3D 엔진으로 통합하여 객체 배치 과정을 통해 최종적으로 가상 장면을 생성한다. 그림 1은 제안하는 장면 합성 프레임워크의 전체적인 구조를 나타낸 것이다.



[그림 1] 장면 합성 전체 프레임워크 프로세스

(a) 현실 세계 장면

(b) Scene Understanding을 이용한 현실 장면 스캔

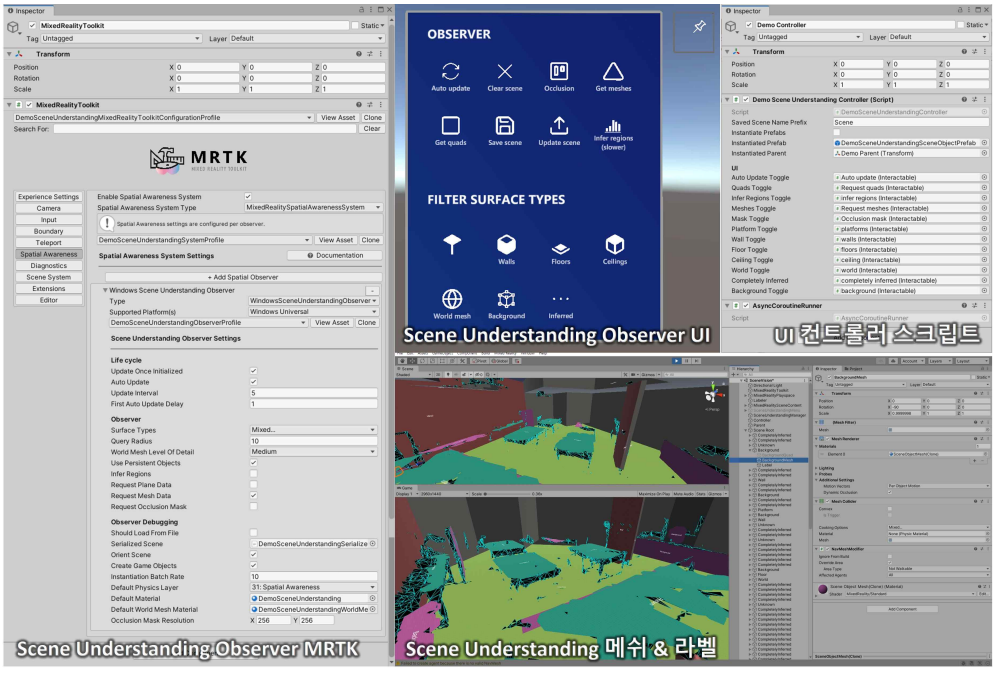
(c) Azure 커스텀 비전을 이용한 객체 추론 및 라벨링

(d) 현실 장면 데이터를 기반의 가상 장면 제작 과정

3.1 현실 공간의 기하학적 분석

현실을 기반으로 한 가상 장면을 자동으로 생성하기 위해서는 먼저 현실 공간에 대한 3차원 기하학적 정보가 필요하다. 이를 위해 본 연구는 Scene Understanding SDK를 활용하여 현실 세계의 기하학적인 공간 정보를 추론한다. Scene Understanding(장면 이해)은 혼합현실 장비를 착용한 사용자가 위치하고 있는 공간을 정밀하게 검사(scan)하고, 탐색된 공간에 대한 벽, 천장, 바닥, 사물 등의 기하학적 형태를 메쉬(mesh) 정보로 반환한다. 이 때, Scene Understanding SDK와 함께 혼합현실 개발 도구인 MRTK 2(Mixed Reality Toolkit)로부터 지원되는 Scene Understanding Observer를 통합하여 사용하였다. Scene Understanding Observer는 3차원 UI(User Interface)를 통해 Scene Understanding 기능을 보다 쉽게 사용할 수 있도록 보조해준다. 장면 스캔 시작 및 초기화, 반환되는 객체 유형 선택, 탐색된 데이터 정보 저장 등이 이에 속한다.

본 연구에서는 홀로렌즈 2를 기반으로 Scene Understanding Observer를 사용하여 현실 장면을 탐색하는 작업 환경을 구동하였다. 홀로렌즈 2에 통합된 작업 환경을 실행하면 Scene Understanding 스캔 기능이 동작한다. 사용자가 위치 하고 있는 좌표를 기준으로 그 주위에 있는 현실 공간에 대한 기하학적 정보를 탐색하고, 그 결과는 현실 공간을 맵핑한 월드 메시(World mesh)를 포함하여 벽(Wall), 바닥(Floor), 천장(Celling), 사물(Platform), 배경(Background) 등의 객체가 SceneObject로 정의된다. 이는 각각 다른 색으로 표현되고, 사각형(Quad)으로 구성된 메시(Mesh)와 SceneObject의 이름으로 된 라벨(Label)과 함께 대응되어 시각화된다. 분석이 끝난 현실 공간의 데이터는 가상 장면 생성에 필요한 기하학적 정보로 사용한다. 그림 2는 마이크로소프트 홀로렌즈 2 장비를 토대로 유니티 3D 엔진 환경에서 제안하는 프레임워크 개발을 위해 사용한 MRTK 2 통합 개발 환경과 관련 속성 및 작업의 일부를 보여주고 있다.



[그림 2] 유니티 3D 엔진에서의 MRTK 2 및 Scene Understanding SDK를 활용한 통합 개발환경 구축과 관련 속성, 작업 과정의 예

3.2 객체의 의미론적 해석

현실 공간의 기하학적 분석을 마친 후에는 현실 공간을 이루는 각 객체가 무엇인지 판단하는 작업이 필요하다. Scene Understanding만으로는 현실 공간을 구성하는 객체의 유형(예: 책상, 의자, 벽장 등)을 구체적으로 구분하기 어렵기 때문에 객체의 의미론적 해석 과정은 현실과 통일된 가상 장면을 생성하기 위해서 추가적으로 필요한 작업이다. 이를 위해 본 연구는 Azure의 커스텀 비전 AI 서비스를 이용해서 현실 장면을 구성하는 객체를 인식하고 분류한다. 커스텀 비전은 간단한 인터페이스를 사용하여 이미지 분류 모델을 쉽게 만들 수 있도록 지원하고, 학습된 모델을 다른 장치에서 유연하게 활용할 수 있도록 배포해주는 기능 등을 포함한 서비스이다. 본 연구는 제안하는 프레임워크를 통해 현실 기반의 가상 장면을 생성하기 위한 실험 공간으로, 다양한 공구와 장비를 활용하여 창의적 작업이 가능한 첨단 실습 공간인 크리에이티브 스튜디오를 선택한다. 해당 공간은 전문가의 도움이 필요하거나 고가의 장비들로 구성되어 일반적으로 접근하기 어려운 환경이기 때문에 본 연구의 목적과 부합하는 특징을 갖는다. 현실 공간과 이를 기반으로 생성된 가상 장면을 토대로 혼합현실 사용자와 가상현실 사용자가 함께 참여하여 상호작용하는 협업 확장현실 환경을 구축하는 것이 또 하나의 핵심 목표이다.

크리에이티브 스튜디오를 구성하는 주요 객체(3D 프린터 (3Dprinter), 보관장 (cabinet), 의자 (chair), 실험 책상 (desk), 망치 (hammer), 노트북 (laptop), 선반 (shelf), 테이블 (table), 공구함 (toolbox), 칠판 (whiteboard))를 총 10가지로 나누어 구분하였고 각 객체를 인식하기 위한 커스텀 비전 학습 데이터를 구축하였다. 이를 위해 인식 대상 객체를 태그로 정의하고, 각 태그마다 최소 15개 이상의 이미지를 학습 데이터로 수집하여 객체 인식을 위한 학습 과정을 설계한다. 이때, 학습 이미지는 시각적 다양성을 고려하여 카메라 각도와 조명 등을 다양하게 구성한다.

홀로렌즈 2에서 크리에이티브 스튜디오를 구성하는 객체로 학습된 커스텀 비전 프로젝트를 연동하기 위한 작업이 필요하다. 앞선 과정을 통해 학습되어

만들어진 커스텀 비전의 이미지 분류 모델을 배포(Publish)하면 URL(Uniform Resource Locator)(Prediction-Endpoint)과 Prediction-Key가 생성되는데 이를 통해서 외부에서 커스텀 비전 프로젝트로의 접근이 가능하다. 따라서, 생성된 정보(URL, Prediction-Key)를 통해 홀로렌즈 2에서 커스텀 비전 프로젝트를 사용할 수 있도록 연동한 후 객체 분석 결과를 받기까지의 과정은 다음과 같은 흐름으로 진행된다.

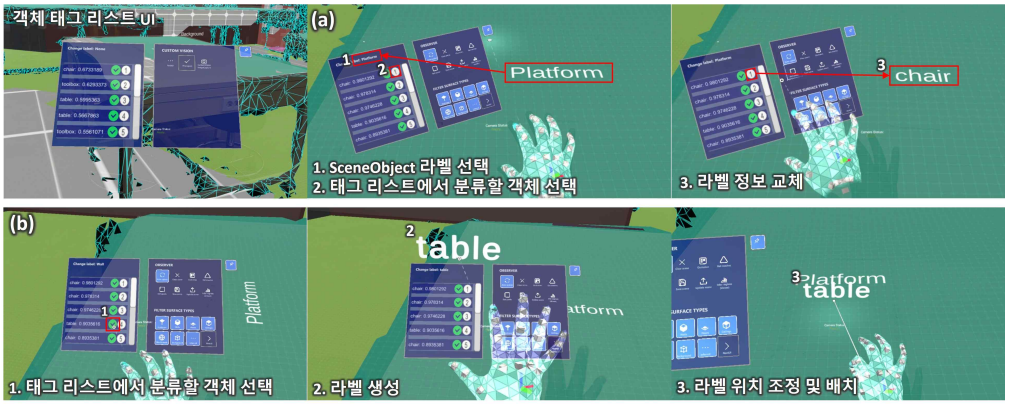
(1) 홀로렌즈 2에서 이미지 캡처 UI 버튼을 선택, 홀로렌즈 카메라로 캡처된 이미지가 커스텀 비전 프로젝트 전송

(2) 전송된 이미지를 커스텀 비전 프로젝트에서 분석, 학습된 객체들의 태그를 기준으로 분류하여 인식된 태그를 비롯한 여러 정보와 함께 반환

(3) 반환된 파일에서 필요한 정보만 가져오기 위해 역직렬화 과정을 거친 후, 태그와 확률 값(probability)만 구조체 형태로 저장하여 객체 분류하는데 사용

커스텀 비전 프로젝트를 통해서 객체의 인식을 모두 마치면 홀로렌즈 2에서 사용자가 직접 인식된 객체를 추론, 분류하는 라벨링 과정을 수행한다. 앞선 기능을 통해 인식된 객체 태그는 확률 값이 높은 순으로 정렬되어 사용자에게 UI의 형태로 제공한다. 가령, 홀로렌즈 2를 착용한 혼합현실 사용자의 시점에서 책상과 의자가 보이는 상황에 연동 기능을 실행하면 인식된 태그(e.g., desk, chair)를 태그 리스트 UI에서 확인할 수 있다. 이를 기반으로 사용자는 라벨링 작업을 수행한다. 먼저, Scene Understanding을 통해 기하학적 메시와 함께 SceneObject의 이름으로 된 라벨 (e.g., 사물 (Platform))이 분류가 된 경우, 정의된 SceneObject 라벨 정보를 가져와서 태그 리스트에 있는 분석된 객체 정보와 교체하는 방식으로 라벨링을 수행한다(그림 3(a)). 이와 달리 SceneObject라벨은 분류되지 않고 기하학적 메시 정보만 있는 객체의 경우에는 직접 라벨을 설정해 주어야 한다. 라벨 데이터가 없는 객체(사물 등)이기 때문에 인식된 태그로 리스트에서 선택하여 라벨을 먼저 생성하고, 분류하고

자 하는 객체의 적절한 위치에 직접 배치한다(그림 3(b)). 기본적으로 기하학적 메시와 함께 기본 태그 정보를 갖춘 모델에 인식된 태그로 교체하는 방법(그림 3(a))이 라벨 선택, 생성, 위치 조정 및 배치 과정을 거지는 두 번째 방법(그림 3(b))과 비교하여 효율적이다. 하지만, Scene Understanding을 통해 분류되는 SceneObject 라벨은 다소 작은 부피를 차지하는 객체(예: 노트북, 공구함)에는 라벨이 생성되지 않기 때문에 두 번째 방법을 추가해주었다.

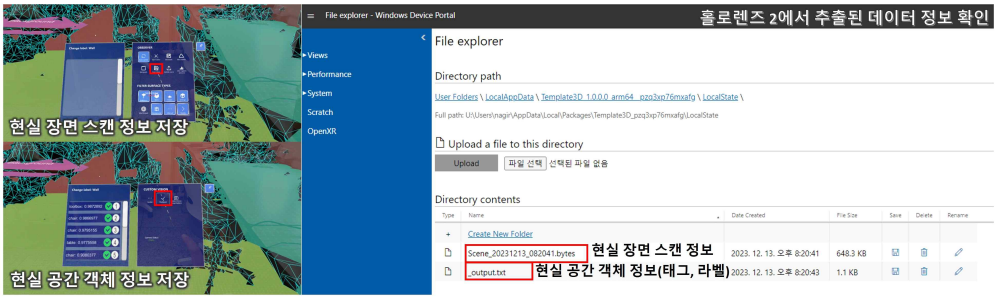


[그림 3] 라벨링 과정

(a) Scene Understanding으로 만들어진 SceneObject라벨 정보와 객체 태그 리스트의 라벨 정보 교체

(b) 객체 태그 리스트에서 인식된 태그 선택 후 라벨 생성, 배치

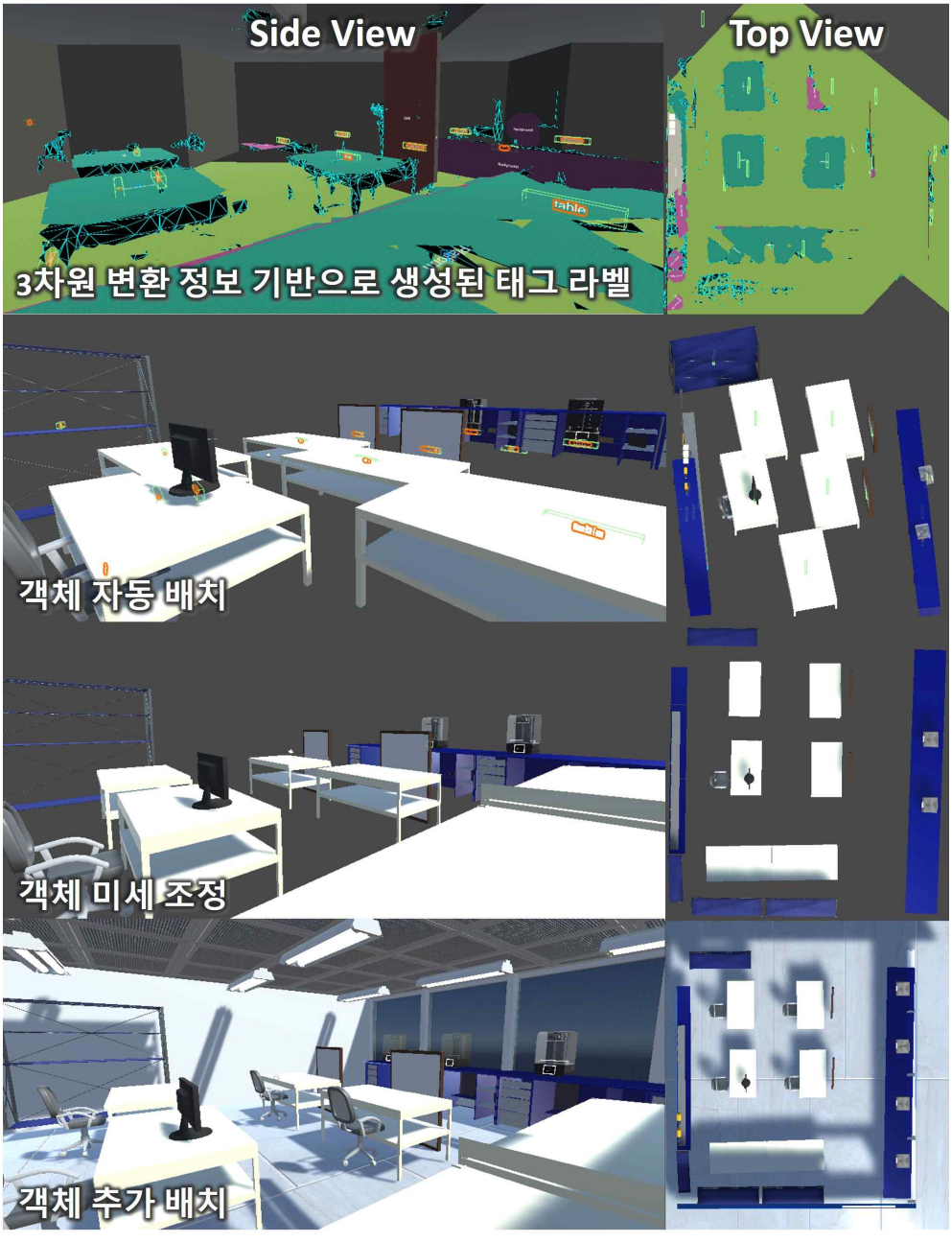
현실 세계의 객체를 인식하여 반환된 객체 정보를 라벨로 만든 후, 이를 객체의 위치에 대응시키는 라벨링 작업을 통해 객체의 종류 및 3차원 변환 값을 추출한다. 이러한 정보는 가상 장면을 만들 때 사용할 수 있는 파일로 변환된다. 그림 4는 객체의 의미론적 해석 과정을 통해 얻은 객체 분류 정보와 이전에 탐색한 장면 정보를 추출하는 과정을 보여주고 있다. 추출된 정보는 홀로렌즈의 Window Device Portal(https://(홀로렌즈 IP(Internet Protocol))) 주소를 통해 확인 가능하다.



[그림 4] 현실 장면 스캔 정보 및 현실 객체 정보 추출

3.3 객체 자동 배치

현실 공간의 기하학적 분석을 통해 얻은 현실 장면 스캔 정보(메쉬와 SceneObject라벨)와 객체의 의미론적 해석을 통해 추출한 객체 정보(태그와 3차원 변환 값(위치))를 바탕으로 가상 장면을 구성한다. 단, 현실 장면과 유사한 가상 장면을 생성하기 위하여 현실 객체와 대응되는 가상 객체에 대해 디자인하고, 정의된 태그에 맞게 객체를 분류하는 사전 작업이 필요하다. 그리고 추출된 태그와 3차원 변환 정보를 토대로 가상 장면을 구성하는 객체를 자동으로 배치하는 과정을 구현한다. 이전 과정을 통해 얻은 현실 공간에 대한 정보들과 이를 이용한 가상 객체 자동 배치 과정을 통해 디자인 툴을 사용해서 가상 장면을 직접 만드는데 소비하는 시간을 줄이는 방안을 제시하고자 하였다. 이 과정에서 객체의 크기, 방향 정보 등에 대한 미세한 조정은 필요할 수 있다. 하지만 이는 전체적인 현실 공간의 스캔 정보를 기반으로 작업하는 것이기 때문에 정확하고 빠른 장면 생성을 도울 수 있다. 그림 5는 추출된 객체 정보를 기반으로 자동으로 배치된 장면 정보와 미세한 조정을 거친 최종 결과물을 보여준다.



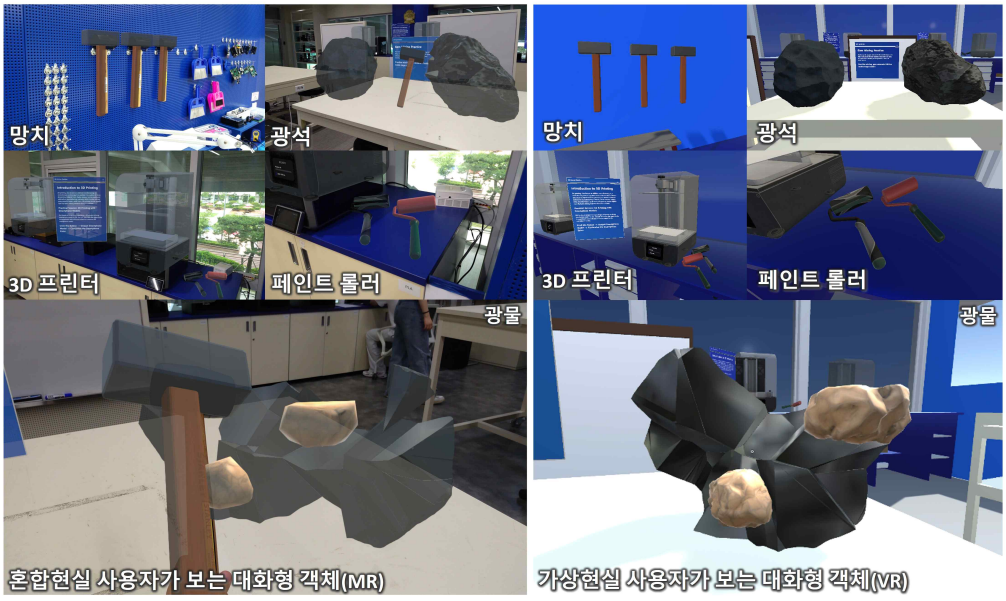
[그림 5] 가상 장면 생성 과정

Ⅳ. 확장현실 협업 환경 구축

본 연구는 혼합현실 장비를 활용하여 계산된 현실 세계로부터 가상현실 사용자가 참여 가능한 가상 장면을 생성하는 프레임워크를 제안한다. 또한, 현실과 대응되는 가상 공간으로부터 혼합현실 사용자와 가상현실 사용자가 함께 참여하는 협업 확장현실 환경을 제시하는 것이 본 연구의 궁극적인 목표이다. 이를 위해 본 연구는 크리에이티브 스튜디오를 현실과 가상의 사용자가 협업 및 상호작용할 수 있는 실험 공간으로 설정한다. 혼합현실 사용자는 홀로렌즈 2를 사용하여 현실 세계에 참여하고, 가상현실 사용자는 메타 퀘스트 2 (Meta Quest 2) HMD를 착용하여 생성된 가상 장면에 참여한다. 또한, 다중 사용자 간의 유기적인 협업, 상호작용, 그리고 정확한 정보 동기화를 위해 유니티 3D 엔진을 기반으로 PUN(Photon Unity Networking) 패키지를 통합하여 구현한다.

협업 확장현실 환경의 목적은 크리에이티브 스튜디오와 같이 첨단 실습 환경에서 교육을 진행하기 어려운 학생들이 현실 기반의 몰입도 높은 교육 활동을 수행할 수 있는 새로운 방향을 제시하는 것이다. 이를 위해 대학교의 창의적 실험 공간인 크리에이티브 스튜디오를 협업 공간으로 활용하도록 설계하였다. 협업 활동은 3D 프린터 교육과 광물 채석 교육으로 구성한다. 우선, 3D 프린터 교육은 3D 프린터를 이용해서 스마트폰 모델을 출력하고, 페인트 롤러를 사용하여 모델의 색을 칠하는 실습 교육 과정으로 이루어진다. 광물 채석 교육은 주어진 광물을 크리에이티브 스튜디오에 배치된 망치로 깨고, 부시는 과정을 혼합현실 사용자가 직접 시범을 통해 보이면서 광석을 캐는 과정을 현실감 있게 교육하는 것이다. 협업 활동에서 혼합현실 사용자는 정의된 현실 세계 도구를 직접 사용하여 교육 활동을 수행하고, 이와 대응되는 가상 객체를 연결하여 혼합현실 사용자는 현실에서, 가상현실 사용자는 가상에서 동일한 교육 활동을 진행한다. 따라서, 현실과 대응되는 가상 객체를 대화형 객체로 정의하여 현실 객체의 위치, 방향, 크기에 맞춰 가상

객체를 사전에 배치한다. 크리에이티브 스튜디오에 정의된 대화형 객체는 3D 프린터, 페인트 롤러, 광물, 광석, 그리고 망치이다. 그림 6은 이러한 대화형 객체를 혼합현실 사용자와 가상현실 사용자의 시점으로 나누어 보여준다. 혼합현실 사용자가 정의된 대화형 객체를 활용하여 현실 세계에서 행동을 수행하면, 가상현실 사용자에게는 대응되는 가상 객체의 행동이 정확하게 동기화되어 보이게 된다. 또한, 가상현실 사용자는 UI를 통해 교육에 대한 기본적인 개념 및 설명을 확인할 수 있다.



[그림 6] 혼합현실 사용자와 가상현실 사용자의 시점에서 보는 대화형 객체 (망치, 광석, 광물, 3D 프린터, 페인트 롤러)

제안하는 협업 확장현실 환경에서의 교육 콘텐츠는 다음과 같은 구조를 갖는다. 먼저, 혼합현실 사용자와 가상현실 사용자는 교육 환경에서 서로 다른 역할을 맡게 된다. 홀로렌즈 2를 착용한 혼합현실 사용자는 교육자의 역할을 수행하며 손을 활용하여 직접 대화형 도구들을 조작하면서 2가지 교육 활동을 직접 실습한다. 반면, 가상현실 사용자는 피교육자로서 혼합현실 사용자가 수행하는 교육을 받는다. 그림 7은 본 연구에서 제작된 협업

확장현실 환경에 참여하는 사용자를 보여준다. 혼합현실 사용자는 대화형 객체를 조작하여 3D 프린터로 모델을 출력하고, 출력된 모델의 색을 변경하는 등의 과정을 직접 수행한다. 또한, 망치를 사용하여 광물을 깨고, 광석을 캐는 작업을 직접 시연하여 현실에 기반한 활동을 함께 공유한다. 가상현실 사용자는 컨트롤러를 활용하여 움직임을 조작할 수 있지만, 가상 객체를 직접 조작하는 등의 직접적인 참여는 제한하여 혼합현실 사용자의 교육 활동에만 집중하도록 하였다.

[그림 7] 협업 확장현실 체험 환경에 참여한 사용자

Ⅴ. 실험 및 분석

제안하는 장면 합성 프레임워크와 이를 기반으로 혼합현실 및 가상현실 사용자 참여를 고려한 협업 확장현실 통합 개발 환경 구축을 위해, 본 연구는 유니티(Unity) 2020.3.12f1을 기반으로 혼합현실 개발 도구(Mixed Reality Toolkit)와 오큘러스 통합 SDK(Oculus Integration SDK)를 사용하여 구현하였다. 혼합현실 사용자는 홀로렌즈 2를, 가상현실 사용자는 메타 퀘스트 2 HMD를 사용한다. 통합 개발 환경과 실험을 위한 PC 환경은 Intel Core i7-10875H, 16GB RAM, Geforce RTX 2060 GPU를 탑재하고 있다.

본 연구는 혼합현실 제작 환경의 관점에서 제안하는 프레임워크의 성능과 효율성을 확인하기 위한 실험으로, 현실 장면으로부터 가상 장면 생성까지의 일련의 작업 시간을 기록하여 정량적 실험 정보를 확인하고, 이에 대한 장단점, 특징, 문제점 등을 종합적으로 분석한다. 또한, 제안하는 프레임워크를 활용하여 혼합현실 및 가상현실 사용자가 함께 참여하는 크리에이티브 스튜디오 실험 콘텐츠를 직접 제작하고, 현실을 기반으로 생성되는 가상 환경이 가상현실 사용자에게 미치는 경험에서의 영향을 분석하기 위한 설문 실험을 진행한다.

5.1 혼합현실 생성자 관점의 정량적 실험

제안하는 프레임워크는 현실 공간의 기하학적 분석과 객체의 의미론적 해석의 두 프로세스로 구성된다. 이를 위해 마이크로소프트에서 제공하는 Scene Understanding SDK를 사용하여 현실 공간의 기하학적 분석을 수행한다. 그리고 Azure 커스텀 비전 AI 서비스를 활용하여 객체의 의미론적 해석을 구현한다. 본 연구에서 실험을 위해 제작한 크리에이티브 스튜디오는 총 10개의 객체 태그(3D 프린터, 보관장, 의자, 책상, 망치, 노트북, 선반, 테이블, 공구함, 칠판)로 정의되는데, 이는 상호작용 객체 또는 현실 공간에서 큰 부피를 차지하는 객체들을 우선순위로 설정하였다. 각 태그당 41~53개의 이미지를 수집하여 총 472개의 이미지로 학습을 진행하였다. 실험을 위해 사용된 크리에이티브 스튜디오 공간의 면적

은 8m x 10m이고, 10개의 태그로 분류된 객체는 현실 공간 정보를 수집하기 시작한 순간을 기점으로 총 55개(3D 프린터 7개, 보관장 2개, 의자 14개, 책상 3개, 망치 4개, 노트북 1개, 선반 3개, 테이블 6개, 공구함 13개, 칠판 2개)이다. 라벨링 작업에서 책상(desk)이 3개 연속으로 한 줄로 배치된 경우, 이를 한 묶음(set)으로 인식하여 라벨을 한 번에 배치하였다.

첫 번째 정량적 실험은 제안하는 프레임워크를 통해 현실 공간을 스캔하는 것부터 시작하여 대응되는 가상 장면을 생성하기까지의 단계별 작업 과정 시간을 기록하는 것이다(표 1). 작업 환경은 현실 공간에서의 작업과 추출 및 계산된 정보를 토대로 대응되는 가상 장면을 제작하는 가상 환경에서의 작업으로 구분된다. 우선, 현실 공간의 기하학적 분석과 객체의 의미론적 해석을 토대로 라벨링 작업을 통해 가상 장면 생성에 필요한 현실 장면 데이터를 추출하고 계산하기까지의 시간이다. 기록된 시간은 평균 8분 내외로 기록되었다. 소요 시간의 대부분은 라벨링 작업(7분 이상)에서 발생하였다. 이는 인식된 정보를 토대로 사용자가 수작업을 통해 객체의 라벨을 설정하기 때문에, 현실 공간의 크기와 작업해야 할 객체의 수에 비례하여 시간이 소요된다. 특히, 라벨링 작업에서 두 번째 방법을 사용할 경우가 많아질수록 소요 시간이 증가한다. 라벨을 교체하기만 하면 되는 첫 번째 방법에 비해 두 번째 방법에서는 수행해야 할 수작업의 양이 많기 때문이다. 따라서, 두 번째 방법의 시간을 단축할 수 있는 개선된 방법(예: 분류 대상 객체의 중심 좌표에 라벨이 자동 배치되는 방법)을 구현하면 시간을 단축할 수 있을 것으로 판단된다. 다음으로는 현실 세계에서 추출된 객체 정보(태그, 변환)를 활용하여 각 객체의 태그 라벨과 가상 객체를 가상 세계에 자동으로 배치하는 시간은 평균 1분 정도 소요되었다. 이 과정도 객체의 수에 비례하는 시간이지만, 하드웨어 사양과 연관성이 높기 때문에 시간 개선 방법을 고려하기는 어려운 부분이 있다. 이후 배치된 가상 객체의 회전과 크기 값을 미세 조정 해준다. 현재의 3차원 변환 정보는 위치(position)만 추출되기 때문에 회전과 크기는 Scene Understanding을 통해 계산된 객체의 메시 정보 위에서 수작업으로 미세 조정하게 된다. 이러한 이유로, 제안하는 프레임워크에서 가장 많은 시간인 9분 정도가 소요되었다. 객체 미세 조정에 대한 작업 시간은 그래픽 디자인이나 유니티 3D 인터페이스에 대한 숙련도에 따라 다르기 때문에 상대적인 결과라고 볼 수 있다.

마지막으로 제작된 가상 장면의 완성도를 높이기 위해 객체나 조명을 추가하는 보정 작업으로 약 5분 정도의 시간이 소요되었다. 이 작업은 프레임워크의 핵심 작업 시간과는 연관성이 낮지만, 협업 확장현실 환경에서 가상현실 사용자의 시각적 몰입을 향상하기 위한 중요한 요소가 될 수 있기 때문에 추가적으로 기록하였다.

작업 도구	단계별 작업 과정	시간
홀로렌즈 2 (현실 환경)	현실 공간 탐색 및 객체 분류 (현실 공간의 기하학적 분석 및 객체의 의미론적 해석 작업)	8분
유니티 3D (가상 환경)	객체 분류 정보(태그 라벨, 3차원 변환)를 활용한 객체 자동 배치	1분
	객체 미세 조정 (회전, 크기 정보 조정)	9분
	완성도를 위한 보정 작업 (조명, 추가 객체 배치 등)	5분

[표 1] 장면 생성 단계별 시간 기록

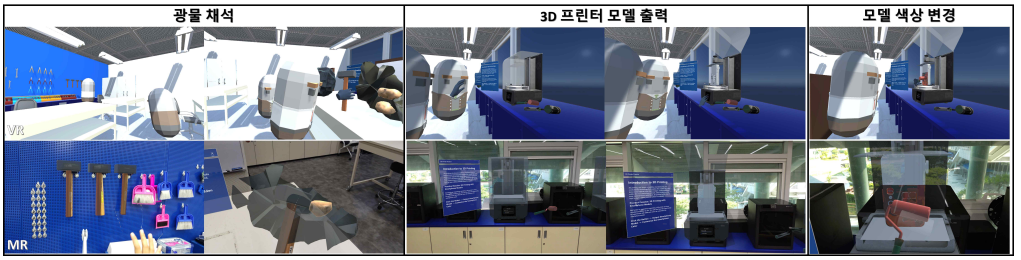
제안하는 프레임워크를 통해 현실 공간을 분석하고, 이를 기반으로 3차원 가상 장면을 생성하는 일련의 작업 시간을 토대로 효율성을 확인하기 위한 기존 연구와의 비교 분석을 진행하였다. 본 연구의 프레임워크와 목적이 정확하게 부합하지는 않지만, 현실 공간을 기반으로 가상 환경을 구축하는 기존의 연구들이 진행되어왔다. Sra et al.[Oasis]는 현실 실내 장면을 3D로 캡처하고 걷기 가능한 영역을 매핑하여 생성된 가상 환경에서 실제 걷기를 가능하게 하는 Oasis를 제안하였다. Wang et al.[PointShopAP]은 현실 공간 스캔 결과를 토대로 사용자가 공간적 맥락에서 디자인 아이디어를 빠르게 프로토타입화 할 수 있는 PointShopAR을 제안하였다. 두 연구 모두 현실 공간을 토대로 가상 환경을 구성하는 연구로 포인트 클라우드를 기본 표현으로 사용하는 방식이다. 우선, 현실 세계를 구성하는 객체의 수의 경우, 본 연구는 55개인 반면 Wang et al.[]의 PointShopAR은 평균 10개(최대 26개)로 2~5배 차이가 발생한다. 공간 규모는 PointShopAR과 비슷한 영역을 갖는 것을 확인할 수 있었다. 효율성에 측면에서

주목할 만한 점은 현실 공간을 스캔하고 가상 장면을 생성하기까지의 장면 계산 시간이다. PointShopAR은 장면 스캔에서 평균 48초(최대 1분 34초), 편집을 통해 장면을 완성하기까지 평균 21분 26초(최대 38분 48초)가 나타난 것으로 기록되었다. 중요한 점은 PointShopAR은 객체에 대한 의미론적 해석 없이 3차원 기하학적 정보만을 계산하고, 장면을 생성하는 시간이다. 이에 반하여 제안하는 프레임워크는 장면을 구성하는 객체의 의미론적 해석 과정까지 포함하여 평균 8분, 이를 기반으로 상호작용을 고려한 3차원 가상 장면 완성까지 평균 23분이 소요되었다. Sra et al.[]의 Oasis는 구체적인 객체 수나 정확한 공간 규모를 명시하지 않았고, 장면 계산까지 2-4분 기록된 것으로 나타난다. Oasis의 경우는 가상현실 사용자가 상호작용에 필요한 객체를 탐지하고 분류하는 전처리 작업이 있다. 이 작업에 필요한 소요 시간이 인스턴스 수에 따라 20분~1시간, 또는 2분이 소요되는 것을 알 수 있다. 하지만, 큰 차이는 걷기 가능한 영역만을 계산하기 때문에 현실 공간의 전체적인 기하학적 구조를 종합적으로 분석하지는 않아 정확한 비교 분석에는 어려움이 있음을 알 수 있었다. 또한, 의미론적 해석에 사용된 객체 역시 의자 1종류 뿐이기 때문에 상호작용 객체가 늘어난다면 작업에 소요되는 시간은 증가할 것으로 예측된다. 본 연구 역시, 작업자의 경험이나 숙련도가 요구되는 미세조정, 보정 작업 등은 소요 시간이 증가 될 요소가 충분히 존재한다. 그럼에도, 기존의 연구들과 비교해서 많은 수의 객체를 분석하는 것은 물론 의미론적 해석을 통해 가상현실 사용자와의 다양한 협업 및 상호작용을 고려하고 있다는 장점이 있다는 것을 확인할 수 있었다. 그리고 작업 시간 역시 충분히 효율적인 범위내에 있음을 판단할 수 있었다. 하지만, 보다 명확한 활용 가능성을 검증하기 위하여 PointShopAR과 같이 제안하는 프레임워크를 통해 디자이너 전문가의 작업 시간을 기록하여 비교 분석하는 실험을 진행할 계획이다.

5.2 가상현실 사용자 관점의 정성적 실험

본 연구는 현실에 기반하여 정확히 대응되는 가상 장면을 생성하고, 이를 기반으로 혼합현실과 가상현실 사용자가 같은 공간 정보에서 정교하고 유기적인 협업을 수행할 수 있음을 확인하기 위해 크리에이티브 스튜디오를 협업 확장현실 환

경으로 제작하였다. 제작된 콘텐츠는 3D 프린터 교육과 광물 채석 교육의 두 가지 주제로 구성되며, 혼합현실 사용자는 현실 기반의 교육자이고 가상현실 사용자는 피교육자로 콘텐츠에 참여한다. 본 연구는 현실을 기반으로 구성된 가상 환경에서의 경험이 가상현실 사용자에게 미치는 영향을 분석하기 위하여 설문 실험을 진행하였다. 그림 8은 사용자의 시점에 따라 구분된 크리에이티브 스튜디오 콘텐츠의 교육 과정을 보여준다. 혼합현실 사용자는 현실 공간에서 실제 장비와 도구를 활용하여 교육을 수행하며, 가상현실 사용자는 현실을 기반으로 생성된 가상 장면에서 현실과 대응되어 동작하는 객체와 사용자의 행동을 통해 콘텐츠의 교육 과정을 체험하게 된다. 이 실험을 통해 가상현실 사용자가 현실에 기반한 가상 환경에서 얼마나 몰입감을 느끼고, 교육 효과를 높일 수 있는지에 대해 종합적으로 분석하고자 하였다. 본 연구에서 크리에이티브 스튜디오 제작에 사용된 모델은 에셋 스토어(Asset Store)의 그래픽 리소스를 활용하였다.



[그림 8] 실험 진행 과정

5.2.1 실험 설계

본 연구에서 가상현실 경험에 관한 설문 실험은 22세부터 29세까지의 총 9명(여자 3명, 남자 6명)을 대상으로 진행되었다. 참가자 중 가상현실 콘텐츠 체험 경험이 있는 사용자는 7명이었다. 설문 실험은 객관식과 주관식으로 구분하여 진행되었다. 먼저, 가상현실에서의 몰입, 집중, 그리고 가상 환경으로의 적응을 분석하기 위해 PQ(Presence Questionnaire)(Witmer et al., 2005)의 항목 중 현실감(realism), 탐색 가능성(possibility to examine), 그리고 자기 평가 성과(self-evaluation of performance) 항목 가운데 7개를 선택하여 본 연구의 목적에 맞게 객관식 설문을 구성하였다. 설문은 7점 척도로 진행되었으며, 질문에 대한

값을 기록하였다. 표 2는 가상현실 사용자 경험 분석을 위해 정리된 문항을 나타낸 것이다.

항목		질문
현실감(realism)	1번	환경의 시각적 요소가 당신을 얼마나 몰입하게 했습니까?
	2번	공간을 통한 객체들의 움직임에 대한 느낌이 얼마나 흥미로웠습니까?
	3번	가상 환경에서의 경험은 실제 세계의 경험과 얼마나 일치합니까?
	4번	가상 환경 내에서 움직이는 감각이 얼마나 흥미로웠습니까?
	5번	가상 환경 경험에 당신은 얼마나 몰입하였습니까?
탐색 가능성 (possibility to examine)	6번	여러 시점에서 얼마나 효과적으로 객체를 검사할 수 있었습니까?
자기 평가 성과 (self-evaluation of performance)	7번	가상 환경 경험에 얼마나 빨리 적응하였습니까?

[표 2] 가상현실 사용자 경험 분석을 위한 설문 문항

실험에 참여하는 참가자들은 가상 환경의 목적, 환경(혼합현실 사용자와의 상호작용), 교육 내용(3D 프린터 교육과 광물 채석 교육), 조작(컨트롤러를 통한 이동) 등에 대한 기본적인 안내를 받는다. 체험 환경에 참여하는 혼합현실 사용자는 교육자로서 교육 콘텐츠를 순서대로 진행하고, 가상현실 사용자는 제공된 콘텐츠를 체험한 후 주어진 문항에 대해 1차 설문에 응답한다. 여기서 중요한 점은 1차 설문이 종료된 후에 체험한 가상 환경이 현실을 기반으로 제작된 가상 장면이라는 사실과 함께, 교육을 주도한 교육자가 정해진 패턴으로 수행된 에이전트가 아닌, 혼합현실 사용자가 현실에서 직접 움직이며 가상 객체와 동기화된 현실 객체를 실시간으로 조작하면서 교육을 수행하였다는 사실을 영상과 함께 고지한다는 것이다. 그 후, 경험과 인식의 변화가 발생했는지를 확인하기 위해 동일한 문항

으로 2차 설문을 진행한다. 이때, 영상은 홀로렌즈 2를 통해 녹화된 영상과 3인칭 시점에서 혼합현실 사용자의 행동을 촬영한 영상으로 구성된다.

5.2.2 현존감 설문 분석

표 3은 실험에 참여한 가상현실 사용자들이 표 2의 항목에 대해 응답한 결과를 보여준다. 전반적으로 참가자들은 현실 공간을 토대로 가상 장면이 생성되었고, 교육을 이끈 주체가 현실 세계의 혼합현실 사용자임을 인지하였을 때 현실과 유사한 몰입 (1차: 5.711, 2차: 6.333), 콘텐츠에 대한 집중 (1차: 5.667, 2차: 6.000), 그리고 환경에 대한 적응 (1차: 6.333, 2차: 6.444)에서 향상된 결과를 보였다. 이를 통해 현실을 기반으로 생성된 장면에서 현실 세계의 사용자가 직접 행동하는 것을 인지하면 사용자들에게 더 큰 흥미를 유발하고, 실제와 유사한 경험을 높일 수 있다는 것을 확인할 수 있었다. 또한, 가상 환경에 대한 집중도를 높이고, 환경에 좀 더 빠르게 적응하도록 하는 측면에서도 현실에 대한 인지가 영향을 미칠 수 있음을 확인할 수 있는 좋은 근거가 될 수 있을 것으로 판단된다. 또한, 일원 분산 분석(one-way ANOVA)을 통해 통계적 유의성을 확인하였다. 모든 항목에 대해서 유의미한 차이가 발생하지 않았지만, 현실과 유사한 몰입에 있어서는 참가자의 수를 늘려 실험을 진행한다면 유의미한 차이가 나타날 수 있을 것이라는 기대를 해볼 수 있었다.

항목		현실 인지 전	현실 인지 후
현실감(realism)	1번	5.889	6.444
	2번	5.778	6.000
	3번	5.222	6.111
	4번	5.778	6.667
	5번	5.889	6.444
	총합	5.711	6.333
탐색 가능성 (possibility to examine)	6번	5.667	6.000
자기 평가 성과 (self-evaluation of performance)	7번	6.333	6.444
Pairwise Comparison			
현실감(realism)		F(1, 16) = 3.1360, p = 0.0956	
탐색 가능성(possibility to examine)		F(1, 16) = 2.0000, p = 0.1765	
자기 평가 성과 (self-evaluation of performance)		F(1, 16) =0.1081, p = 0.7466	

[표 3] 현실 인지를 기반으로 한 가상현실 사용자의 경험 비교 분석 결과

5.2.3 주관적 설문 분석

통계적 비교 분석 과정에서 유의미한 차이가 발생하지 못한 이유로, 본 연구는 현실을 기반이라는 것을 알았음에도 인식에 변화가 없는 참가자들이 영향을 미쳤을 것이라고 판단하였다. 따라서, 본 연구는 현실에 대한 인지가 가상현실 경험에서의 인식 변화 여부에 대한 주관적 경험을 응답 받아 분석해보고자 하였다. 다음은 주관식으로 진행된 설문 문항이다.

(1) 가상 환경을 경험한 후에 해당 환경이 현실 기반이라는 것을 알고 인식에 대한 변화가 있었을 때: 가상 환경을 경험하기 전과 현실 기반으로 만들어진 가상 환경이라는 사실을 알았을 때의 기대와 경험 후의 인식이 어떻게 다른지, 어떤 측면에서 그렇게 느꼈는지 서술하시오.

(2) 가상 환경을 경험한 후에 해당 환경이 현실 기반이라는 것을 알고 인식에 대한 변화가 없었을 때: 그 이유와 어떤 측면에서 그렇게 느꼈는지, 만약 다른 환경(예, 공장, 실험실 등)이어도 동일하게 느꼈을 것 같은지 서술하시오.

우선, 인식 변화가 있었는지에 대해서 응답을 받았다. 그 결과 전체 참가자 9명 중 7명은 인식 변화가 있었다는 응답을 하였다. 전반적으로 현실을 기반으로 생성된 가상 장면이라는 사실이 가상 환경에 대한 흥미를 높였고, 몰입에 긍정적인 영향을 미칠 수 있다고 답했다. 또한, 교육 과정에 대한 적응, 이해도를 높이는데도 긍정적인 영향을 미쳤다는 의견이 있었다. 이외에도, 현실 세계 혼합현실 사용자가 교육자로서 아바타의 행동을 수행하였다는 사실이 캐릭터가 어색하다고 느꼈던 첫 인상을 자연스럽다는 인식으로 변화시켰다는 의외의 의견도 있었다. 기대 효과의 측면에서 공간에 대한 이용도를 높이거나 실험, 교육에 있어 제안했던 2가지 활동 외에도 다양하게 활용할 수 있을 것 같다는 생각을 제시하기도 하였다. 중요한 점은 유의미한 인식의 변화가 없었다는 2명의 의견이었다. 이와 관련하여, 혼합현실 사용자의 행동이 현실을 기반으로 하고 있다고 하여도 피교육자의 입장에서 교육 행동에 대한 결과가 예측이 가능하여 큰 변화를 주지 못한다는 점을 이유로 들었다. 또한, 궁극적으로 가상현실 사용자는 현실 장면을 보는 것이 아닌 가상 장면에 대한 시각적 정보만 제공되는 상황이고, 내용 또한 다르지 않기 때문에 인식의 변화가 없었다는 점이었다. 이를 통해 가상현실 경험을 하기 전에 현실을 기반으로 체험한다는 사실을 먼저 인지 시키는 것이 오히려 긍정적인 영향을 미치는데 도움을 줄 수 있을 것이라는 점도 고려해볼 수 있었다.

Ⅵ. 한계 및 토론

본 연구는 현실 세계를 기반으로 정교하게 설계된 가상 장면을 생성하기 위한 프레임워크를 제안하였다. 또한, 이를 기반으로 혼합현실 사용자와 가상현실 사용자가 함께 참여하는 협업 확장현실 환경으로 교육 콘텐츠를 직접 제작하고, 현실을 기반으로 한 가상 환경에서의 가상현실 사용자 경험을 분석하고자 하였다. 하지만 제안하는 프레임워크의 효율성과 사용자 평가에서의 한계가 존재하였다. 다음은 본 연구의 한계와 함께 논의한 내용이다.

6.1 미세 조정 작업에서의 효율성

정량적 실험을 통해 확인된 문제 중 가장 주목할 만한 부분은 가상 장면의 미세 조정 작업이다. 이는 실험 결과에서도 언급된 바와 같이 개발자와 제작자의 숙련도에 따라 차이가 크게 발생할 수 있는 작업이라는 점이다. 따라서, 미세 조정 작업에서의 효율성을 높이기 위해서는 현실 세계로부터 추출된 객체 정보 가운데 변환 값을 보다 정교하게 설계하는 것이 필요하다. 현재는 3차원 기하학적 정보 가운데 위치(position)에 대한 정보만으로 장면을 자동 배치하고 있기 때문에, 이와 함께 회전과 크기 정보도 함께 추출하는 구조로 보완한다면 미세 조정 작업 시간은 크게 단축될 수 있을 것으로 기대된다. 향후에는 3차원 변환 정보를 추출하는 것에 대한 개선 방향을 보완할 계획이다.

6.2 수동적인 가상현실 사용자의 참여

가상현실 사용자 경험 분석을 위해 제작된 콘텐츠는 현실을 기반으로 제작된 가상 환경이 가상현실 사용자의 몰입, 집중, 그리고 적응에 미치는 영향을 확인하는 것이 주요 목적이었다. 그러나, 제작된 교육 목적의 콘텐츠에서 가상현실 사용자는 피교육자로서 수동적인 참여로 제한하고 있기 때문에 역할로 인한 부정적인 경험 효과를 배제할 수 없다. 가상현실 사용자의 활동에 현실에서 동작하고 있는

객체와 동기화된 가상 객체를 직접 제어하는 등 보다 능동적으로 체험환경에 참여할 수 있는 역할을 추가한다면 현실에서 활동하고 있는 혼합현실 사용자와 연결된 느낌이 직접적으로 느껴질 수 있기 때문에 현실을 인지하였을 때 경험에서의 유의미한 차이를 유도할 수 있었을 것이다. 또한, 협업 확장현실 환경에서 가상현실 사용자의 중요성도 확인해 볼 수 있었을 것이다. 따라서, 현실을 기반으로 하는 협업 확장현실 체험 환경에서 혼합현실, 가상현실 사용자의 역할이 보다 다양하게 제공될 수 있는 콘텐츠를 제작하고, 이를 기반으로 사용자 경험을 분석함으로써 다양한 응용 방향으로의 유의미한 실험 결과를 도출할 수 있도록 보완하고자 한다.

6.3 탐색할 현실 장면 규모

본 연구에서의 현실 공간 규모는 8m x 10m로 사무실 공간과 비슷한 정도의 크기이다. 해당 공간의 현실 정보를 수집하는 데 드는 시간이 본 논문에서는 8분 정도 소요되었지만, 그 규모에 따라 해당 시간이 변동될 수 있고, 규모가 큰 공간일 경우 해당 공간을 이루는 현실 물체의 수만큼 객체 인식을 위해 미리 학습시키는 시간 역시 늘어나기 때문에 전체적인 작업 시간은 현실 공간에 상당히 큰 영향을 받는다. 따라서, 제안하는 프레임워크는 사무실이나 회의실과 같은 한 공간을 대상으로 이용하기에 적합하며, 건물 전체나 큰 규모의 공간에 활용하기에는 시간 소모가 늘어날 수 있기 때문에 제한적일 수 있다.

Ⅶ. 결론

본 연구는 혼합현실 사용자 기반의 현실 장면으로부터 가상현실 사용자가 경험하는 가상 장면으로 연결되는 협업 확장현실 체험 환경 구축을 위한 새로운 프레임워크를 제안하였다. 이는 현실과 가상이 정확하고, 정교하게 연결되어 참여할 수 있는 체험 환경을 위해 현실 공간을 분석하고 이를 기반으로 가상 장면을 자동으로 생성하는 것을 목적으로 하였다. 제안하는 프레임워크는 확장현실 협업 환경 장면 합성을 위한 현실과 가상이 대응되는 동일 좌표 설정을 위해 현실 공간에 대한 기하학적 정보를 분석하는 것과 현실 공간을 구성하는 객체들의 유형 및 특징 등의 정보를 분류하는 과정을 구분하여 정의하였다. 이 과정에서 유니티 3D 엔진 내의 Scene Understanding SDK와 Azure 커스텀 비전을 활용하여 현실 세계 정보를 추출하였고, 현실 세계 정보를 바탕으로 가상 장면 자동 생성을 위하여 사전에 정의된 3차원 가상 객체를 자동으로 배치하는 과정을 설계하였다. 궁극적으로, 제안하는 프레임워크를 활용하여 혼합현실 사용자와 가상현실 사용자가 같은 장면으로 구성된 공간에서 협업하고 상호작용할 수 있는 체험 환경 제작을 본 연구의 핵심 목표로 하였다. 따라서, 우선 제안하는 프레임워크를 활용하여 현실 세계로부터 대응되는 가상 장면을 효율적으로 생성할 수 있는지에 대한 여부를 확인하기 위해 작업 시간 측정에 관한 정량적 실험을 수행하였다. 프레임워크의 주요 작업을 중심으로 시간을 측정하였고, 제작자(또는 개발자)의 그래픽 디자인 또는 편집 작업의 숙련도에 의존하는 미세 조정 작업을 제외하고 효율적인 구조로 설계되었음을 확인할 수 있었다. 또한, 제안하는 프레임워크를 활용하여 혼합현실 사용자과 가상현실 사용자가 함께 참여하는 체험 환경을 제작하였다. 제작된 가상 장면은 크리에이티브 스튜디오 공간으로, 창의적인 교육 활동을 수행할 수 있고 쉽게 접하기 어려운 공간을 선택하여 가상현실 기술이 요구되는 체험 환경을 구축하고자 하였다. 그리고, 현실을 기반으로 제작된 가상 장면에서의 체험 활동이 가상현실 사용자 경험에 미치는 영향을 분석하기 위하여 설문 실험을 진행하였다. 설문 실험은 몰입, 집중, 그리고 적응을 확인할 수 있는 문항으로 구성하여 진행하였고, 실험 결과 현실을 기반으로 생성된 가상 장면에서의

경험이 몰입을 높이고, 흥미로운 경험을 제공하는 등 긍정적인 영향을 미칠 수 있다는 것을 확인할 수 있었다. 특히, 현실 세계와 가상 세계의 정보가 동기화된 환경에서 혼합현실 사용자가 현실 객체를 조작함으로써 가상현실 사용자에게 사실적인 정보 전달이 가능하다는 점에서 협업 확장현실 환경의 발전 가능성과 활용성 등을 검증하였다. 본 연구를 통해 현실에서 접하기 어려운 환경을 가상으로 재현하여 교육, 제조업, 의료 등 다양한 분야에서 활용될 수 있음을 시사한다. 더 나아가, 이러한 기술이 더욱 보완되고 발전된다면 현실과 가상이 융합된 새로운 형태의 디지털 경험을 제공할 수 있을 것으로 기대된다. 뿐만 아니라, 다양한 산업 분야에서 혁신적인 변화를 가져올 수 있으며 사용자들에게 보다 높은 몰입감과 함께 현실과 유사한 새로운 경험을 제공할 수 있을 것이다.

참 고 문 헌

1. 국외문헌

Carlsson, C., Hagsand, O. (2002). DIVE A multi-user virtual reality system, Published in: Proceedings of IEEE Virtual Reality Annual International Symposium, 394-400.

Cheng, L., Ofek, E., Holz, C., Wilson, A. D. (2019). V R o a m e r : Generating On-The-Fly VR Experiences While Walking inside Large, Unknown Real-World Building Environments, Published in: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 359-366.

Cho, Y., Kang, J., Jeon, J., Park, J., Kim, M., & Kim, J. (2021). X-person asymmetric interaction in virtual and augmented realities. Computer Animation and Virtual Worlds, 32(5), e1985.

Cho, Y., Park, M., & Kim, J. (2023). XAVE: Cross-platform based Asymmetric Virtual Environment for Immersive Content. IEEE Access. 11, 71890 - 71904.

Duval, T., Fleury, C. (2009). An asymmetric 2D Pointer / 3D Ray for 3D Interaction within Collaborative Virtual Environments, Web3D 2009, Jun 2009, Darmstadt, Germany. 33-41.

Feld, N., Weyers, B. (2021). Mixed Reality in Asymmetric Collaborative Environments: A Research Prototype for Virtual City Tours, Published in: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 250-256.

Francisco J. R. R., Rafael, M. S., Rafael, M. C., (2018). Speeded up detection of squared fiducial markers, Image and Vision

Computing 76, 38-47.

Grandi, G. J., Debarba, G. H., Hemdt, I., Nedel, L., Maciel, A. (2018). Design and Assessment of a Collaborative 3D Interaction Technique for Handheld Augmented Reality, Published in: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 49-56.

Grandi, G. J., Debarba, G. H., Maciel, A. (2019). Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities, Published in: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 127-135.

Ibayashi, H., Sugiura, Y., Sakamoto, D., Miyata, N., Tada, M., Okuma, T., Kurata, T., Mochimaru, M., Igarashi, T. (2015). Dollhouse VR: a multi-view, multi-user collaborative design workspace with VR technology, SIGGRAPH Asia 2015 Emerging Technologies, November 2015, Article 8, 1-2.

Kumaravel, B. T., Anderson, F., Fitzmaurice, G., Hartmann, B., Grossman, T. (2019). Loki: Facilitating Remote Instruction of Physical Tasks Using Bi-Directional Mixed-Reality Telepresence, UIST '19: Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, October 2019, 161-174.

Kwan, K. C., Fu, H. (2019). Mobi3DSketch: 3D Sketching in Mobile AR, CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, May 2019, Paper 176, 1-11.

Lang, Y., Liang, W., Yu, L. F. (2019). Virtual Agent Positioning Driven by Scene Semantics in Mixed Reality, Published in: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 767-775.

Madathil, K. C., Greenstein, J. S. (2017). An investigation of the efficacy

of collaborative virtual reality systems for moderated remote usability testing, Volume 65, November 2017, 501-514.

Merrell, P., Schkufza, E., Li, Z., Agrawala, M., Koltun, V. (2011). Interactive furniture layout using interior design guidelines, ACM Transactions on Graphics, Volume 30, Issue 4, Article 87, 1-10.

Ning, B., Pei, M. (2024). Task and Environment-Aware Virtual Scene Rearrangement for Enhanced Safety in Virtual Reality, Published in: IEEE Transactions on Visualization and Computer Graphics, Volume 30, Issue 5, May 2024, 2517 - 2526.

Qian, X., He, F., Hu, X., Wang, T., Ipsita, A., Ramani, K. (2022). ScalAR: Authoring Semantically Adaptive Augmented Reality Experiences in Virtual Reality, Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, Article 65, 1-18.

Oppermann, L., Buchholz, F., Uzun, Y. (2023). Industrial Metaverse: Supporting remote maintenance with avatars and digital twins in collaborative XR environments, CHI EA '23: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, April 2023, Article 178, 1-5.

Piumsomboon, T., Dey, A., Ens, B., Lee, G., Billinghurst, M. (2019). The Effects of Sharing Awareness Cues in Collaborative Mixed Reality, Front Robot AI 6:5.

Schott, E., Makled, E. B., Zoeppig, T. J., Muehlhaus, S., Weidner, F., Broll, W., Froehlich, B. (2023). UniteXR: Joint Exploration of a Real-World Museum and its Digital Twin, VRST '23: Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology, October 2023, Article 25, 1-10.

Sra, M., Sergio, G. J., Maes, P. (2017). Oasis: Procedurally Generated Social Virtual Spaces from 3D Scanned Real Spaces, Published in:

IEEE Transactions on Visualization and Computer Graphics, Volume 24, Issue 12, 01 December 2018, 3174 - 3187.

Sun, J. M., Yang, J., Mo, K., Lai, Y. K., Guibas, L., Gao, L. (2024). Haisor: Human-aware Indoor Scene Optimization via Deep Reinforcement Learning, ACM Transactions on Graphics, Volume 43, Issue 2, Article 15, 1-17.

Tian, H., Lee, G. A., Bai, H., Billinghurst, M. (2023). Using Virtual Replicas to Improve Mixed Reality Remote Collaboration, Published in: IEEE Transactions on Visualization and Computer Graphics, Volume 29, Issue 5, May 2023, 2785 - 2795.

Tycho T. D. B., Angelica, M., Tinga and Max, M. (2021). Learning in immersed collaborative virtual environments: design and implementation, 5364-5382.

Wang. Z., Nguyen. C., Asente. P., Dorsey. J. (2023). PointShopAR: Supporting Environmental Design Prototyping Using Point Cloud in Augmented Reality, Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, April 2023, Article 34, 1-15

Weiss, T., Litteneker, A., Duncan, N., Nakada, M., Jiang, C., Yu, L. F., Terzopoulos, D. (2018). Fast and Scalable Position-Based Layout Synthesis, Published in: IEEE Transactions on Visualization and Computer Graphics, Volume 25, Issue 12, 01 December 2019, 3231 - 3243

Yeh, Y. T., Yang, L., Watson, M., Goodman, N. D., Hanrahan, P. (2012). Synthesizing open worlds with constraints using locally annealed reversible jump MCMC, ACM Transactions on Graphics, Volume 31, Issue 4, Article 56, 1-11.

Yi, H., Huang, C. H. P., Tripathi, S., Hering, L., Thies, J., Black, M. J. (2023). MIME: Human-Aware 3D Scene Generation, Proceedings

of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, 12965-12976.

Yu, L. F., Yeung, S. K., Tang, C. K., Terzopoulos, D., Chan, T. F., Osher, S. J. (2011). Make It Home: Automatic Optimization of Furniture, ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2011, v.30, (4), July 2011, article no.86, v. 30, 3767-3777.

Zhang, Z., Yang, Z., Ma, C., Luo, L., Huth, A., Vouga, E., Huang, Q. (2020). Deep Generative Modeling for Scene Synthesis via Hybrid Representations, ACM Transactions on Graphics, Volume 39, Issue 2, Article 17, 1-21.

Zhu, E., Hadadgar, A., Masiello, I., Zary, N. (2014). Augmented reality in healthcare education: an integrative review, PeerJ 2:e469. Arrangement, ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2011, v.30, (4), July 2011, article 86, 3767-3996.

2. 기타 자료

Microsoft. (2022). Mixed Reality, Design, Scene Understanding.

Microsoft. (2022). Mixed Reality and Azure Service. Azure Custom Vision.

Microsoft. (2022). Mixed Reality, MRTK2, Spatial awareness, Scene Understanding observer.

Meta. (2022). Meta quest. Meta Technologies, LLC.

UnityTechnologies. (2019). Unity engine. Unity Technologies.

ABSTRACT

Framework of Scene Synthesis for Collaborative XR: From MR-based Real Scenes to VR-based Virtual Scenes

Na, Gi-Ri

Major in Computer Engineering

Dept. of Computer Engineering

The Graduate School

Hansung University

To facilitate sophisticated and organic collaboration and interaction between mixed reality (MR) users based on real-world environments and virtual reality (VR) users operating within computer-generated environments, the seamless alignment and generation of corresponding spaces and scenes between reality and virtuality are essential. This study proposes a scene synthesis framework aimed at accurately and effectively generating virtual scenes from real-world scenes. The proposed framework comprises geometric analysis to understand the three-dimensional structure of the real space and semantic interpretation to classify objects by type and features, facilitating the calculation of

information necessary for virtual scene generation. To analyze the three-dimensional geometric information of real spaces, Microsoft HoloLens 2 device and Scene Understanding SDK are utilized to scan real scenes and extract mesh information of objects (such as walls, ceilings, and objects). Subsequently, semantic interpretation is performed using Azure Custom Vision to recognize and classify object types. Objects' types (features) in the real space are classified into tags, and labeling tasks are conducted to connect the geometric information (position) of all classified objects. Leveraging the object data (tags, transformations) from the real space, three-dimensional virtual objects are automatically placed, followed by fine adjustments for rotation and size, resulting in the creation of virtual scenes corresponding to reality. Based on this, the efficiency and performance of the proposed framework are evaluated by measuring the time required for each task, revealing the ability to effectively generate virtual scenes based on reality in approximately 10 minutes. Furthermore, this research aims to create educational content in creative studio spaces, fostering collaboration and interaction between MR and VR users in the same space. Additionally, a survey experiment is conducted to analyze the impact of the virtual environments created through the proposed framework on VR user experience. The results confirm the positive effects of real-based virtual scenes on user immersion, concentration, and adaptation, highlighting their potential for enhancing user experiences.

【Key words】 Scene Synthesis, HoloLens, eXtended Reality, Digital Twin, Virtual Environment