

모바일기반 증강현실 사용자 인터랙션 프레임워크

*정성욱, 김주영, 박상현, 이병규

통신미디어연구소 한국전자통신연구원

{brcastle, kimjy1113, shpark12, bglee86}@etri.re.kr

A Mobile-based User Interaction Framework in Augmented Reality Environment

*Sung-Uk Jung, Juyoung Kim, Sangheon Park, Byunggyu Lee

Communications & Media Research Laboratory, ETRI

요약

메타버스 서비스가 점차적으로 부각됨에 따라 이와 관련한 가상현실, 증강현실 관련 기술은 메타버스 환경을 구축하기 위한 요소 기술로서 필요성이 대두되었다. 기존 기술의 경우 다른 종류의 증강현실 기기의 입력 포맷의 차이로 인해 가상증강현실 환경에서 동시에 사용자가 인터랙션을 체험하는데 제약이 되고 있다. 이를 극복하기 위해, 고가의 설치형 하드웨어 기반 추적 시스템을 구축하여 활용하고 있으나, 공간 및 범용성을 고려한다면, 일반 사용자가 이용하고 있는 범용성 모바일 기기를 활용한 인터랙션 기술 개발 필요하게 되었다. 본 논문에서는 위와 같은 제약점을 해결하기 위해 모바일 기기를 활용하여 다수 사용자 및 원격 사용자가 증강현실을 체험할 수 있는 프레임워크 구조 및 원격사용자가 참여할 수 있는 방안에 대해서 기술하였다.

I. 서론

메타버스 서비스가 점차적으로 부각됨에 따라 이와 관련한 VR(Virtual Reality), AR(Augmented Reality) 관련 기술은 메타버스 환경을 구축하기 위한 요소 기술로서 필요성이 대두되었다. 특히 메타버스 환경에서 VR기반의 사용자 참여 및 단말기를 통한 사용자 참여는 제한적으로 이루어지고 있으나 AR기반으로 현실세계를 반영한 사용자 참여는 활발히 연구되고 있는 추세이다 [1-2].

기존 기술의 경우 다른 종류의 증강현실 기기(스마트폰, AR글래스, VR기기)의 입력 포맷의 차이(RGB, RGBD, IMU, Depth)로 인해 가상증강현실 환경에서 동시에 사용자가 인터랙션을 체험하는데 제약이 되고 있다. 이를 극복하기 위해, 고가의 설치형 하드웨어 기반 추적 시스템을 구축하여 활용하고 있으나 [3], 공간 및 범용성을 고려한다면, 일반 사용자가 이용하고 있는 범용성 모바일 기기를 활용한 인터랙션 기술 개발 필요하게 되었다. 이와 관련해서, 모바일환경에서 정밀 인터랙션을 구현하기 위해서는 사용자의 스켈레톤 인식을 통해 자유도가 높은 인터랙션이 가능해야 하며, 다수 사용자의 인터랙션을 동시에 처리하는 기술이 필요하게 된다.

Google의 ARCore의 경우 영상 및 추가 센서를 이용한 템스 맵을 생성하여 객체 합성을 가능하게 하는 ARCore Depth API 공개하였지만, 템스 맵 생성을 위한 디바이스의 프로세스를 다량으로 사용하며 추정치에 의한 심도맵 기반으로 위치 측위 시스템과의 연동 및 스케일 조정 부분은 제한된 기능만을 제공한다 [2]. Apple의 ARKit의 경우에도 카메라와 LiDAR 센서를 이용하여 공간 구축과 사용자의 몸, 손 자세 추정 API 제공 평면기반의 공간 인식을 이용해서 가상객체를 증강하는 방식으로 복합공간을 인식하여 가상객체와의 인터랙션은 제한된 기능만 제공하고 있다 [1]. Niantic의 경우 RGB 영상을 이용하여 숨겨진 지면 형상과 범위를 예측하는 Footprints 기술을 개발하여 인터랙션을 위한

공간정보와 바닥 표면 검출을 통한 증강현실 서비스가 가능하지만, 위치 추정기술을 제공하지 않고 있다 [4]. 사용자의 동작을 인식하기 위한 영상에서의 스켈레톤 추출 기술은 딥러닝 네트워크를 활용하여 사용자의 스켈레톤을 추출하여 동작을 인식하고 있으나, 현재는 PC기반의 고정밀 위치 추적기술에 대한 연구가 주로 이루어지고 있고, 모바일 디바이스에서의 적용은 리소스의 부족으로 인해 제한되고 있다.

본 논문에서는 위와 같은 제약점을 해결하기 위해 범용 모바일 기기를 활용하여 다수 사용자 및 원격 사용자가 증강현실을 체험할 수 있는 환경 및 시스템 구조에 대해서 기술하고자 한다.

II. 본론

본 논문에서 구현한 모바일기기 기반 AR 인터랙션 시스템 구성도는 크게 모바일 정밀분석 모듈, 모바일 카메라 위치 추적 모듈, 모바일 경량화 모듈, 원격 사용자 지원 모듈로 나뉜다. 각 모듈의 입력 정보는 스마트폰, AR글래스, VR기기 등의 센서정보 (RGB 영상, IMU 데이터, GPS정보 등)를 포함한다

1. 모바일기기 기반 AR인터랙션 프레임워크



<그림 1> 모바일기기 기반 AR 인터랙션 프레임워크

<그림 1>는 모바일기기 기반 AR 인터랙션 시스템 구성도를 나타낸다. 증강현실 인터랙션 시스템 구성도는 크게 모바일 정밀분석 모듈, 모바일 카메라 위치 추적 모듈, 모바일 경량화 모듈, 원격 사용자 지원 모듈로 나뉜다.

모바일 정밀 동작 모듈에서는 입력된 센서 정보로부터 사용자의 스켈레톤 추출, 손동작 인식, 그리고, 스켈레톤과 사용자의 손을 병합하는 과정을 수행한다. 이 모듈에서는 Outside-in, Inside-out 뷰에서의 스켈레톤을 추출하게 된다. 예를 들면, 사용자가 AR글래스를 착용하고 있는 경우 사용자가 바라보는 뷰에서 사용자 손의 스켈레톤을 추출하고, 다른 사용자의 바디 스켈레톤을 추출한다. 모바일 카메라 위치 추적 모듈에서는 입력된 센서정보를 기반으로 체험공간 맵을 구축하고, 공간에서의 실객체의 분할 및 인식을 수행하며, 모바일기기 카메라의 공간상 정밀 위치를 추적하는 기능을 수행한다. 사용자 스마트폰을 이용해서 공간 스캔을 통한 공간맵을 구축하고, 실시간으로 스마트폰의 3차원 위치를 추적하게 된다. 모바일 경량화 모듈에서는 상기 두 모듈이 모바일기기에서 동작할 수 있도록 모바일 덤러닝 경량화 기능 지원과, 착용형 디스플레이 (AR글래스, VR기기 등)에서 동작할 수 있는 지원을 담당하게 된다. 원격 사용자 지원 모듈은 원격의 사용자가 로컬 장소에 접속하여 증강현실을 체험할 수 있는 기능을 제공한다. 이 모듈에서는 로컬공간에서의 공간맵과 영상을 이용해서 원격 사용자에게 로컬뷰 볼 수 있도록 하는 원격공간을 가상화할 수 있는 기능과 원격사용자의 동작을 인식하고, 가상 아바타를 통해 로컬공간에 증강시키는 기능을 수행한다

2. 시스템 환경 구성



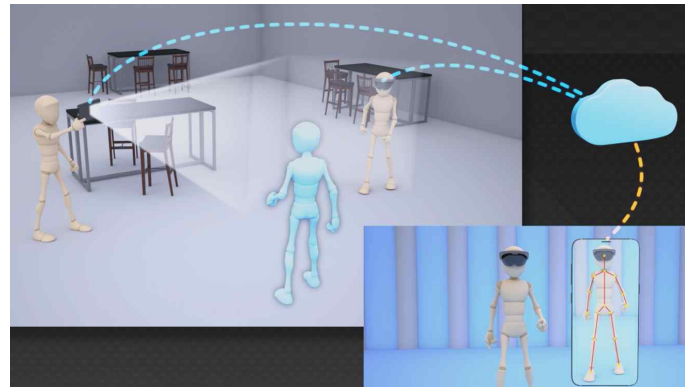
<그림 2> 모바일기기 기반 증강현실 시스템 환경

<그림 2>는 기술한 시스템을 이용하여 5m × 5m의 공간에 증강현실 체험존을 구축하는 경우를 나타낸다. 먼저 환경은 3대 이상의 모바일기기를 공간상에 고정되게 거치한다. 그리고, 사용자는 AR글래스를 착용한다. 서버는 모바일기기 및 AR글래스와 통신하면서 데이터를 주고받는다. 사전 작업으로 체험공간의 공간맵을 구축하고, 서버 및 각 기기에 전송한다.

거치된 모바일기기는 모바일 정밀동작 모듈의 기능을 이용해서 체험자의 3차원 스켈레톤을 실시간으로 추출하여 서버로 보내고, 서버에서는 각 모바일기기에서 추출된 체험자의 3차원 스켈레톤을 합성하여 정밀한 사용자의 3차원 스켈레톤을 재생성하고, 각각의 모바일 디바이스로 재전송한다. 또한 거치된 모바일기기는 모바일 카메라 정밀추적 모듈을 이용해서 각 기기의 위치를 실시간으로 서버에 전송한다. 이는 초기 설치시 카메라 캘리브레이션 과정을 생략할 수 있고, 약간의 이동에도 강인하게 기기의 위치 추적이 가능하게 된다. 사용자가

착용한 AR글래스에서는 모바일 카메라 정밀추적 모듈을 이용해서 실시간으로 바라보는 뷰의 3차원 위치를 추적하고, 서버로 그 정보를 전송하게 된다. 가상 콘텐츠의 경우에는 사전에 구축한 콘텐츠를 게임엔진(Unity3D, Unreal 등)을 활용해서 사용자가 착용한 AR글래스에 사용자의 뷰와 행동에 맞추어 실시간으로 증강시킨다.

3. 가상 원격 사용자 참여



<그림 3> 원격사용자 참여 예시

<그림 3>은 원격 사용자가 참여하는 경우를 보여주는 예시이다. 원격사용자는 사용자 앞에 카메라 뷰에 전신이 나오도록 모바일기기를 거치한다. 거치된 기기에서는 모바일 정밀동작 모듈의 기능을 이용해서 실시간으로 3차원 스켈레톤을 추출한다. 추출된 스켈레톤은 로컬서버로 전송한다. 로컬 서버에서는 가상 아바타에 스켈레톤을 매핑하여 로컬체험공간에 증강시킨다. 서버는 양방향으로 로컬 체험공간에 증강된 콘텐츠를 원격사용자의 AR글래스 및 로컬사용자의 AR글래스에 증강시킨다.

III. 결 론

본 논문에서는 가상증강환경에서 범용 모바일 디바이스를 이용하여 다수 사용자 및 원격 사용자가 참여할 수 있는 프레임워크에 대해서 기술하였다. 모바일 카메라 위치 추적, 사용자 정밀동작 인식, 그리고, 경량화 모듈을 통해서 로컬 사용자만이 아니라 원격 사용자가 참여할 수 있는 구조에 대해서 정의 하였다. 향후에는 실외 환경에서 사용자의 가림현상에 강인한 메타버스 요소기술을 개발하는 것이 필요하다.

ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No.2021-0-00230, '실·가상 환경 해석 기반 적응형 인터랙션 기술 개발')의 일환으로 수행되었음.

참 고 문 헌

- [1] Apple augmented reality developer, ARKit 5, 2021 from: <https://developer.apple.com/augmented-reality/arkit>
- [2] Google augmented reality developer ARCore, 2017 from: <https://developers.google.com/ar?hl=en>
- [3] NaturalPoint, Inc. from: <https://optitrack.com>
- [4] Niantic, Inc. from: <https://nianticlabs.com/>