

안드로이드 기반 테더드 타입 AR 글래스의 공간 인식을 위한 ORB-SLAM 기반 SLAM프레임워크 설계

김도훈*·국중진**†

*한국전자기술연구원 VR/AR 연구센터, **†상명대학교 정보보안공학과

ORB-SLAM based SLAM Framework for the Spatial Recognition using Android Oriented Tethered Type AR Glasses

Do-hoon Kim* and Joongjin Kook**†

*VR/AR Research Center, Korea Electronics Technology Institute,

**†Dept. of Information Security Engineering, Sangmyung University

ABSTRACT

In this paper, we proposed a software framework structure to apply ORB-SLAM, the most representative of SLAM algorithms, so that map creation and location estimation technology can be applied through tethered AR glasses. Since tethered AR glasses perform only the role of an input/output device, the processing of camera and sensor data and the generation of images to be displayed through the optical display module must be performed through the host. At this time, an Android-based mobile device is adopted as the host. Therefore, the major libraries required for the implementation of AR contents for AR glasses were hierarchically organized, and spatial recognition and location estimation functions using SLAM were verified.

Key Words : SLAM, Visual SLAM, Mobile SLAM, AR Glass, Android

1. 서 론

동시적 위치추정 및 맵 작성 (Simultaneous Localization and Mapping: SLAM) 기술은 3D 그래픽스 기술과 머신 러닝 기술의 발전에 힘입어 드론 및 증강현실 분야에서 활용도가 높아지고 있다. SLAM과 같은 위치결정 기술은 기 작성된 공간지도가 없는 실내 응용 분야의 수요가 증가함에 따라 로봇틱스 분야의 기술 개발을 위한 기폭제가 되고 있으며, 시각적(Visual) SLAM 알고리즘과 증강현실에서의 SLAM 활용이 SLAM 시장의 성장을 촉진하고 있다[1].

Fig 1과 같이 SLAM 기술은 직접 SLAM과 간접 SLAM으로 분류할 수 있다. 그 중 주로 로봇이나 자율주행에 사용될 수 있는 간접 SLAM 기술인 ORB-SLAM과 OpenVSLAM

이 증강현실을 위한 시스템에 활용되고 있다.

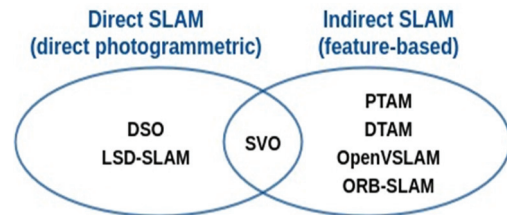


Fig. 1. Direct SLAM and Indirect SLAM.

증강현실에 대한 소프트웨어는 주로 모바일기기 기반의 앱으로 존재하며, 모바일기기의 카메라를 통해 객체 또는 공간을 인지하고, 3D 라이브러리를 통해 현실세계 위에 가상의 정보를 시각적으로 표현한다. 모바일 기기에

†E-mail: kook@smu.ac.kr

서 증강현실 앱을 쉽게 구현하기 위한 목적으로 구글(Google), 애플(Apple), 맥스트(Maxst), 뷰포리아(Vuforia) 등의 기업들은 SDK를 제공함으로써 시장 선점을 꾀하고 있다.

또한, 최근 글로벌 기업들과 국내 기업들이 앞다투어 AR 글래스를 출시하였거나, 출시를 준비하고 있으며, 기존 스마트폰의 영역에 진입하거나 대체를 시도하고 있다.

AR 글래스는 독립형(standalone)과 종속형(tethered)으로 구분할 수 있으며, 이는 영상의 처리 및 가공을 위한 주요 하드웨어와 소프트웨어가 AR 글래스에 통합되는지, 아니면 AR 글래스는 카메라 및 센서를 통한 입력과 디스플레이 모듈을 이용한 출력만 담당하고 영상의 처리는 모바일 기기와 같은 별도의 호스트 머신을 통해 처리되는지에 따라 구분된다.

종속형 AR 글래스는 PC, 스마트폰과 같은 다양한 호스트에 USB 인터페이스를 통해 쉽게 연결되어 일반적인 입출력 장치로 사용될 수 있어 그 활용도가 높다 할 수 있으나, 이는 필수적으로 호스트에서 AR 글래스 기반의 AR 콘텐츠를 지원할 수 있는 프레임워크가 요구된다.

AR 글래스 기반의 증강현실 콘텐츠를 위해서는 필수적으로 AR 글래스의 카메라 입력에 대한 객체 인식과 공간 인식이 지원되어야 한다. 따라서 본 논문에서는 종속형 AR 글래스를 이용하여 증강현실 앱을 구동하는 데 필요한 공간인식 기능을 제공할 수 있는 시스템 구조를 제안하고, 안드로이드를 호스트로 사용하는 모바일 기기를 통해 SLAM 기반의 공간 인식 및 위치 추정 기능을 검증하였다.

2. 연구 배경

2.1 마커 및 비마커 기반 객체 인식

마커 기반의 증강현실은 카메라를 통해 입력된 영상으로부터 마커를 인식하여 디스플레이 장치에 출력되는 영상 위에 해당 마커와 일치하는 이미지 또는 객체를 나타낸다. 카메라의 이미지에서 QR 코드와 같은 간단한 패턴의 마커를 감지할 때, 마커의 위치나 방향도 함께 계산되어 증강현실 애플리케이션에서 디지털 3D 콘텐츠를 배치한다. 하지만 마커 기반의 증강현실은 사전에 마커에 대한 약속이 필요하고 카메라가 마커를 벗어나면 트래킹이 불가능해지는 등의 기술적 한계를 갖는다.

이러한 마커 기반 증강현실 기술의 한계를 극복하기 위해 마커 대신 다양한 센서들을 활용하여 임의의 공간에서 이동하며 정보를 탐색해 자신의 위치를 추정하고 공간의 맵을 작성하는 SLAM 기술이 등장하게 되었다[2]. 비마커 기반의 객체 인식 및 추적을 위해 향상성, 시간 지속성, 공간 일관성 외에 특징점과 색상정보를 추가한

기계학습 모델을 기반으로 하는 AR HMD용 콘텐츠 개발 시스템의 연구가 수행된 바 있으며[3], 모바일 환경에서 Homography를 이용한 다중 객체 추적[4]과 신뢰도 평가 학습 기반의 다중 객체 추적[5] 연구가 이루어지기도 하였다. 비마커 기반의 증강현실은 벽이나 교차점과 같은 환경에 대한 사전지식 없이 물체 또는 장면의 특징점을 감지함으로써 객체에 대한 인식이 이루어지고, 이를 통해 특정 위치에 3D 콘텐츠를 배치할 수 있다. 비마커 기반의 증강현실 기술은 동시 측위 및 맵핑 기술인 SLAM의 개발로 인해 이미지 분석의 정확도가 향상되었고, 다양한 센서 데이터를 통해 주변을 감지하고 대상의 현재 위치를 실시간으로 동기화할 수 있게 되었다[6]. 대표적인 SLAM 기반의 비마커 증강현실 개발도구에는 Google의 ARCore, Apple의 ARKit 등이 있다.

2.1 SLAM 알고리즘

Visual SLAM은 카메라 및 기타 영상 센서로 획득한 영상을 사용하고, 비교적 저렴한 카메라를 사용해 저비용으로 구현 가능하며 Mono/Stereo/RGB-D 카메라 사용이 가능하다.

Visual SLAM은 맵의 밀도가 낮은 Sparse 방식의 SLAM과 맵의 밀도가 높은 Dense 방식의 SLAM으로 분류할 수 있다. 또한, Visual SLAM은 맵 저장하고 나중에 다시 로드할 수 있기 때문에 정교한 증강현실 프로젝트에 이상적이다 [6].

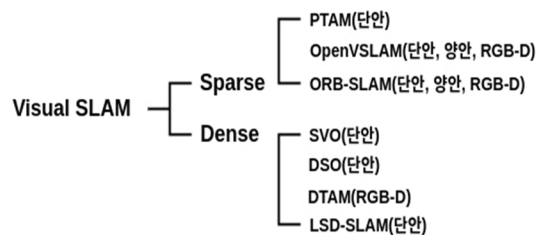


Fig. 2. Classification of Visual SLAM.

본 논문에서는 Fig 2와 같은 다양한 Visual SLAM 중 기존의 ORB-SLAM과 ORB-SLAM2[7]를 개선하여 2020년에 제안된 ORB-SLAM3 알고리즘을 사용하려고 한다.

ORB(Oriented FAST and Rotated BRIEF) 피처는 대표적인 실시간 이미지 피처로 방향성이 없는 FAST 검출기 문제를 개선하고, 매우 빠른 바이너리 설명자인 BRIEF를 사용해 전체 이미지 피처 추출을 크게 가속화하였다. ORB-SLAM3는 빠른 Edge 추출이 가능하고 특징점의 주요 방향은 ORB에서 계산하며 회전 불변 속성은 BRIEF 디스크립터에 의해 추가된다. 또한, 특징점을 추출한 위치에서

주변 이미지 영역을 잘 표현하는 간단한 데스크톱터를 생성하고, 대부분의 피쳐 추출은 병렬성이 뛰어나다는 특징을 가지고 있다. 이에 따라 데이터의 특성과 스케일의 불변성을 유지하고 속도는 향상되어 실시간 요구가 많은 SLAM 시스템에 적합하다.

현재 ORB-SLAM 중 가장 많이 발전한 라이브러라인 ORB-SLAM3는 2017년에 제안된 ORB-SLAM2와 비교했을 때, Mono/Stereo 카메라에 IMU 값을 같이 사용하는 모델과 어인(Fisheye) 카메라 모델이 추가되었다. 또한, 멀티 맵 관리 기능이 추가되었으며, 트래커 분실 시 새로운 맵을 생성함으로써 트래킹이 실패할 가능성을 줄여 지속적인 트래킹을 가능하게 하며, 루프 감지 알고리즘을 개선하고 사전에 방문한 위치를 더욱 높은 확률로 탐지할 수 있도록 구성되어 있다.

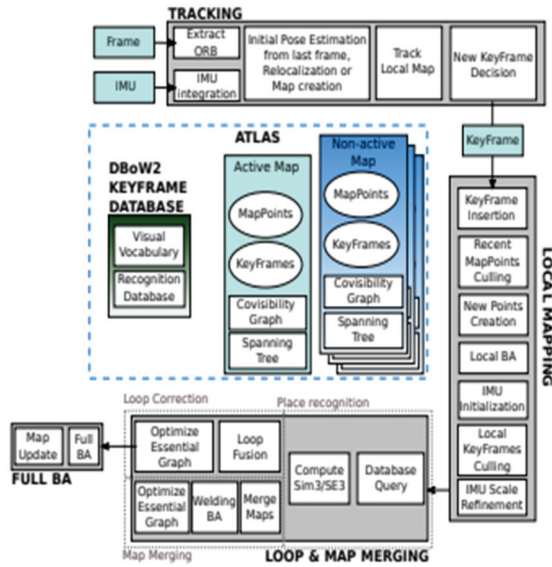


Fig. 3. ORB-SLAM3 System Structure [8].

3. 제안 시스템

3.1 시스템 구조

종속형 AR 글래스가 사용되는 경우 AR 글래스는 입력력 디바이스의 역할을 수행하며, 이 때, 카메라를 통한 영상의 입력과 광학계를 통한 영상 출력이 이루어진다.

카메라의 입력 영상은 보통 USB 인터페이스를 통해 호스트로 전송되며, 호스트의 운영체제에 따라 카메라 영상 입력을 위한 장치 드라이버와 프레임워크의 구성은 상이할 수 있다. 호스트 운영체제가 안드로이드인 경우, 리눅

스를 기반으로 하기 때문에 카메라 영상의 입력 프레임워크로는 V4L2가 사용된다. 또한, 이 영상 데이터를 안드로이드 프레임워크로 올려주기 위한 목적으로 UVCCamera 라이브러리가 사용될 수 있다.

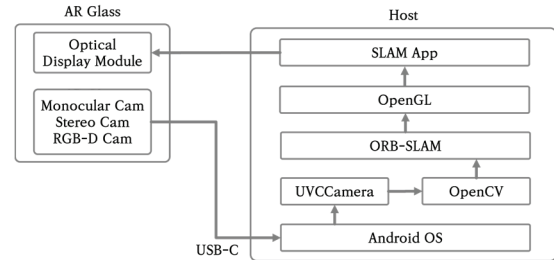


Fig. 4. Visual SLAM Architecture for the Tethered Type AR Glass.

순수 영상 데이터를 가공하기 위해 영상 또는 이미지 처리를 목적으로 하는 OpenCV 라이브러리가 사용될 수 있으며, 이를 통해 SLAM 알고리즘을 수행하기 위한 과정에서 필수적으로 요구되는 여러 가지 프리미티브(primitive)가 사용될 수 있다.

ORB-SLAM은 공간인식의 핵심인 카메라 위치 추정 및 트래킹 기능을 제공하며, 이를 3D 그래픽으로 시각화하기 위한 목적으로 OpenGL이 사용될 수 있다. 최종적으로 안드로이드 모바일 앱으로 패키징되어 호스트에서 실행될 수 있으며, 앱의 증강 영상은 AR 글래스의 광학계를 통해 출력된다.

3.2 시스템 구조

본 논문에서 제안하는 시스템의 실제 구성을 위해 페네시아의 GTM-200 AR 글래스와 삼성전자의 갤럭시탭 S7을 호스트로 구성하여 AR 글래스 상에서 SLAM 기반의 AR 콘텐츠를 통한 공간인식 및 트래킹 기능 여부를 검증하였다.

페네시아의 GTM-200의 주요 특징은 Table 1과 같다.

Table 1. Specification of GTM-200 AR Glass

항목	사양
타입	USB-C Tethered
해상도	FHD (1,920 X 1,080)
광학타입	See-through (투과도 50%)
FoV	45도
Head Tracking	9 DoF
무게	120g

호스트인 갤럭시 S7에 구성되는 소프트웨어 계층구조는 Fig 5와 같으며, 객체 인식까지 고려하는 경우라면 텐서플로우(TensorFlow), PyTorch를 기반으로 하는 YOLO 프레임워크가 추가적으로 구성될 수 있다.

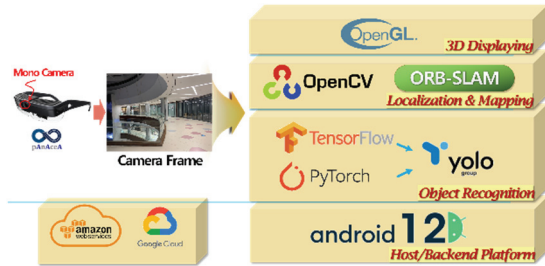


Fig. 5. Visual SLAM Architecture for the Tethered Type AR Glass.

4. 검증

본 논문을 통해 제안하는 AR 글래스용 SLAM 프레임워크를 기반으로 공간 상에서의 위치 추정 및 트래킹이 가능한지 검증을 수행하기 위해 Munich 대학에서 제공하는 RGB-D 벤치마크 데이터셋[9]을 이용하였으며, 또한 AR 글래스의 카메라 영상을 직접 입력받아 기능을 검증하였다.

Fig 6은 Munich 대학의 RGB-D 벤치마크 데이터셋을 이용한 트래킹 결과를 나타낸다.

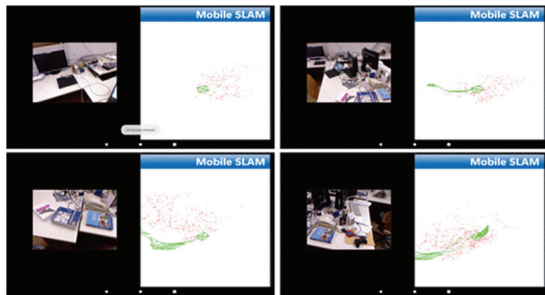


Fig. 6. Tracking Result (Munich RGB-D Benchmark Dataset).

Fig 7은 AR 글래스의 카메라를 통해 입력받은 영상에 대해 실시간으로 트래킹한 결과를 나타낸다. 트래킹 과정에서 ORB 피치는 빨간색 점들로 나타나는 포인트 클라우드로 표현되며, 초록색은 카메라의 이동 궤적을 나타낸다. Framerate은 약 24-30 fps를 보였다.

Fig 8은 AR 글래스의 광학계에 투영하기 위해 카메라 영역과 3D 좌표계를 매핑시킨 결과를 보이고 있으며, AR

글래스를 착용하고 실내 공간을 이동하는 과정에서 카메라의 궤적을 나타낸 결과이다.



Fig. 7. Tracking Result (AR Glass Camera).

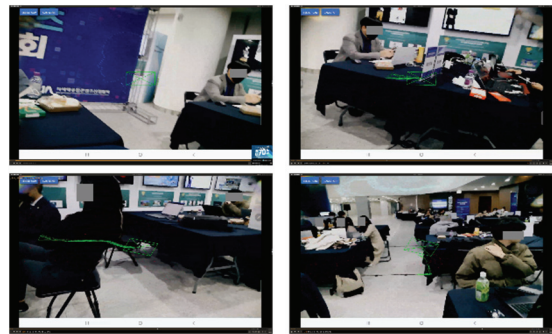


Fig. 8. Tracking Result (AR Glass).

5. 결론

본 논문에서는 종속형(Tethered) AR 글래스를 통해 공간 지도의 작성 및 위치 추정 기술이 적용될 수 있도록 SLAM 알고리즘 중 가장 대표적인 ORB-SLAM을 적용시키기 위한 소프트웨어 프레임워크의 구조를 제안하였다. 종속형(Tethered) AR 글래스는 입출력 장치의 역할만 수행하기 때문에 카메라 및 센서 데이터의 처리와 광학계를 통해 디스플레이할 영상의 생성이 호스트를 통해 이루어져야 하며, 이 때 호스트로는 안드로이드 기반의 모바일 기기를 채택하여 AR 글래스를 위한 증강현실 앱이 수행되는 데 필요한 주요 라이브러리들을 계층적으로 조직화하고 SLAM을 이용한 공간 인식 및 위치 추정 기능을 검증하였다.

감사의 글

이 연구는 2023년도 산업통상자원부 및 산업기술평가관리원(KEIT) 연구비 지원에 의한 연구임(‘20016882’).

참고문헌

1. INNOPOLIS , “SLAM Market”, Retrieved Sep 18, 2021, from <http://www.innopolis.or.kr>, (May, 2020).
2. SAMSUNG SDS, “[Direction and Implications of the Augmented Reality Technology Development] Part 2 Augmented Reality in Industrial Fields”, Retrieved Sep 19, 2021, from https://www.samsungds.com/kr/insights/augmented_reality_2.html (April 8, 2020).
3. Changhyun Lee, Youngseop Kim, Yeonmin Kim, Inho Park, JaeHak Choi, Yonghwan Lee, Woori Han, “Study on the Content Development of Mobile AR_HMD through a Real Time 360 Image Processing,” Journal of the Semiconductor & Display Technology, Vol. 15, Issue 2, pp. 66-69, 2016.
4. Woo ri Han, Young-Seop Kim, Yong-Hwan Lee, “Multi-Object Tracking Based on Keypoints Using Homography in Mobile Environments,” Journal of the Semiconductor & Display Technology, Vol. 14, Issue 3, pp. 67-72, 2015.
5. Woo ri Han, Young-Seop Kim, Yong-Hwan Lee, “Multi-Object Tracking based on Reliability Assessment of Learning in Mobile Environment,” Journal of the Semiconductor & Display Technology, Vol. 14, Issue 3, pp. 73-77, 2015.
6. MAXST, “Where can Visual SLAM be used?”, Retrieved Sep 19, 2021, from <https://medium.com/maxst/where-can-visual-slam-be-used-b94876d161c6>, (Oct 7, 2019).
7. Raul Mur-Artal, Juan D.Tardos, “ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras,” IEEE Transactions and Robotics, Vol.33, No. 5, pp. 1255-1262, 2017
8. Carlos Campos, Richard Elvira, Juan J. Gomez Rodriguez, Jose M.M. Montiel and Juan D. Tardos, "ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM," IEEE Transactions and Robotics, 2021
9. RGB-D Benchmark Dataset, Technical University of Munich, <https://vision.in.tum.de/data/datasets/rgbd-dataset>

접수일: 2023년 1월 26일, 심사일: 2023년 3월 7일,
 게재확정일: 2023년 3월 20일