스마트폰 증강현실 환경에서 피라미드 비콘을 활용한 딥러닝 영상기반 실내측위 시스템

안현우, 문남미
호서대학교 컴퓨터공학과
e-mail:nammee.moon@gmail.com

Deep Learning Image-based Indoor Positioning System using Pyramid Beacon in Smartphone Augmented Reality Environment

An Hyeon Woo*, Moon Nammee**
Hoseo University

요 약

디지털화된 현실 환경을 증강현실속에서 투영시키기 위해선 증강현실 디바이스의 측위가 필수적이다. 하지만 대부분의 측위 방식이 측위 대상 디바이스에 대해 별도의 하드웨어나 센서를 요구하는데 이를 스마트폰 환경에서 충족시키기란 매우 힘든 일이다. 이에 본 논문은 스마트폰 환경에서 별도의 하드웨어를 요구하지 않는 딥러닝 영상기반 실내 측위 시스템을 제안한다. 제안하는 시스템은 측위를 위하여 설계된 피라미드형의 비콘을 활용하며 실시간에 가까운 피드백을 구현하기 위해 딥러닝 기법을 활용한 탐지를 진행한다. 본 논문에서는 상기한 두 개의 요소를 포함한 제안 시스템의 구성요소들을 설명하고 학습 방법과 비콘의 자세 측정 방법, 최종 측위 프로세스 등 전반적인 측위 프로세스를 설명한다.

1. 서론

최근 통신속도 발달과 스마트폰 고성능화에 따라 스마트폰 기반 증강현실과 그에 대한 요소기술들이 주목받고 있다. 몰입도 높은 증강현실 구현을 위해선 요소기술들의 구현이 견고하게 뒷받침되어야 한다. 일례로 3D 모델을 안정적으로 투영시키기 위한 영상합성 기술의 안정성이 떨어진다면 흔들림, 이질감, 부정확한 렌더링 등의 문제들에 의해 몰입도가 떨어지는 결과를 초래할 수 있다. 따라서 몰입도 높은 증강현실 구현을 위해선 정밀하고 안정적인 요소기술들의 구현이 필요하다.

증강현실 환경에서 디지털화 된 현실 세계를 투영하는 시스템을 설계할 경우 가장 중요한 요소기술은 측위 기술이다. 만약 부정확한 측위 방식을 기반으로 시스템을 구현한다면 예외상황에 대한 처리를 포함하지 않은 이상 콘텐츠의 표현이 측위 오차에 따라 떨어진 위치에 진행되거나 측위 주기에 따라 노이즈의 영향을 크게 받게된다. 스마트폰 환경의 경우 이러한 문제가 크게 가시화되는데 주된 이유로는 대부분의 스마트폰 디바이스의 경우 정밀한 측위를 보조할 하드웨어의 설치가 힘들거나 요구되는 센서가 존재하지 않기 때문이다.

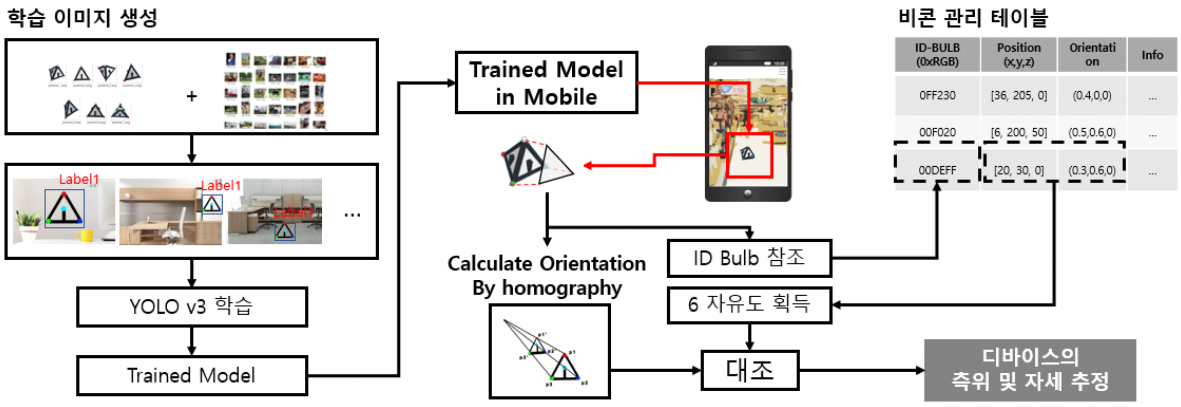
본 논문에서는 이러한 제약사항을 극복하고 스마트폰 환경에서 정밀한 측위가 진행 가능한 시스템을 제안한다. 제안하는 측위 시스템은 탐지에 적합하도록 설계된 피라미드 비콘과 비콘의 절대좌표와 비콘 ID 정보를 관리하는 관리서버, 딥러닝 모델의 일종인 YOLO v3를 활용하여 비콘을 인식하는 인식부, 비콘과의 상대적인 거리와 호모그래피를 활용하여 비콘의 자세를 측정하는 검출부, 마지막으로 획득한 정보를 관리서버로부터 받은 정보와 대조하여 위치를 측정하는 측위부로 이루어져 있다.

2. 관련연구

이 단락에서는 간략하게 비콘을 인식하는 딥러닝 기술의 개요와 다양한 측위 시스템들을 소개한다.

2-1. YOLO v3

영상에서 객체를 검출하는 방법은 크게 영상의 자질을 이용한 알고리즘 기반 방식과 학습된 인공신경망을 이용한 방법으로 나뉘어진다. 영상의 자질을 이용한 알고리즘 기반 방식은 오탐의 여지가 있는 다양한 상황에 대해 일일이 반영할 수 없는 한계가 존재하며 목표 객체의 특징에 따라 검출 방식의 전반적인 수정이 필요할 수 있다는 단점이 존재한다. 반면 인공신경망 기반 방식의 경우 학습 데이터의 양과 질에 따라 오탐의 여지를 줄일 수 있으며



(그림 1) 측위 시스템의 오버뷰

악조건에 대해 효율적인 검출이 가능하다.

인공신경망 기반 다중 객체 검출 방식에는 Faster R-CNN과 YOLO, R-FCN(Region-based Fully Convolutional Network), SSD(Single Shot Detector) 등이 존재하며 각 검출 방식은 대체적으로 정확도와 fps(frames per second) 상에 반비례하는 특성 갖고 있다. 본 연구에서는 스마트폰 환경에서 피라미드 비콘을 실시간에 가깝게 검출하여야 하므로 다른 방식들에 비해 검출 속도가 빠르고 모델의 용량이 상대적으로 적은 YOLO v3을 사용하기로 결정하였다[1].

2-2. 센서 기반 측위

센서 기반 측위 기술은 대표적으로 Fingerprinting 방식과 Dead Reckoning방식이 존재한다. Fingerprinting 방식으로 잘 알려진 Wifi Fingerprinting은 디바이스 주변으로 포진된 3개 이상의 AP에게서 받는 신호 세기량을 토대로 삼변측량을 진행하거나 미리 구성된 라디오맵과 대조하여 유사한 패턴의 신호 세기량을 통해 위치를 측정하는 방식이다. 이 방식은 Wifi 기지국의 개수가 많으면 많을 수록 정밀도가 높아지지만 AP와 측위 디바이스 사이의 방해물, 온도, 습도 등의 환경 변화에 민감한 단점을 갖고 있다[2].

Dead Reckoning 방식은 측위 디바이스에 포함된 가속도, 자이로, 지자기, 기압센서 등 디바이스의 움직임 및 환경 변화를 포착하는 여러 센서 데이터를 활용하여 일정 시간동안의 움직임을 파악하는 방식이다. 하지만 측위의 정밀도가 센서의 성능에 종속적이며 측위의 시간이 길어질수록 오차가 누적된다는 단점을 갖고 있다.

2-3. 영상 기반 측위

영상을 기반으로 디바이스의 위치를 측위하는 기술에는 여러 방식이 존재한다. 일례로 LED의 깜빡임을 사람은 인지하지 못할 정도로 다르게 조정하여 모바일 디바이스의 전면 카메라에 촬영된 천장 LED를 토대로 측위하는 기술이 있다[3]. 이는 오차가 cm단위로 몰입도 있는 AR 구현에 적용할 수 있다. 하지만 디바이스의 카메라가 LED 를 향해 끊김없이 촬영하고 있어야 하며 보다 정확한 위치 측정을 위해선 영상 내에 두 개 이상의 LED Bulb를 포착하고 있어야 한다.

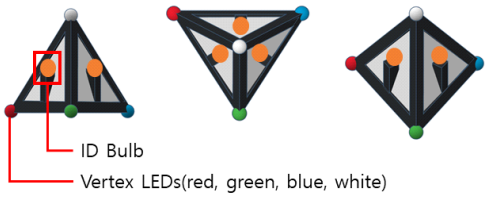
또 다른 방식으로는 마커 기반 측위가 존재한다. 마커란 QR 코드와 같은 기하학적 패턴으로 이루어진 표시물로써 각 마커의 절대 좌표와 자세 추정을 통해 위치를 측정하는 기술이다[4].

위 두 기술과 같은 영상 기반 측위 기술의 경우 센서 융합 기반 측위에 비해 상대적으로 높은 정밀도를 갖고 있으며 유지보수 또한 용이한 편이나 다른 측위 보조기술이 존재하지 않는다면 측위를 위해 지속적으로 목표물을 촬영하고 있어야 하고 측위 성능이 카메라 성능에 영향을 받는 등 여러 제한 사항이 존재한다.

3. 본론

제안 시스템은 그림 1과 같이 이루어져 있다. 사전에 YOLO v3를 학습하고 스마트폰 환경에 학습된 모델을 포함시켜 카메라에 포착된 비콘의 Bounding box를 검출하는 용도로 사용하며 검출된 영역을 활용해 각 비콘의 면이 취하고 있는 회전 상태를 도출, 서버의 비콘 관리 테이블의 정보와 대조를 통해 측위를 진행하게 된다.

피라미드 비콘의 경우 그림 2와 같이 탐지와 자세 추정에 특화되도록 설계되어 진행된다. 피라미드 비콘의 구조는 영상에서 기하학적 특징을 보다 쉽게 탐지하기 위해 세개의 꼭지점에 각각 다른 색상의 LED가 존재하며 각 면의 가운데에는 해당 비콘을 판별할 ID Bulb가 포함되어 있다. ID Bulb는 색상 스펙트럼에서 모바일 디바이스가 인식 가능한 단위로 분할하여 할당된다. 각 LED의 색은 2차원 공간에서의 3차원 물체 회전 변환을 설명하기 위한 Homography의 특징점으로 활용된다[5]. 또한 ID Bulb를 이용하여 비콘 서버에서 해당 비콘의 절대좌표와 각 비콘



(그림 2) 피라미드 비콘 구성도

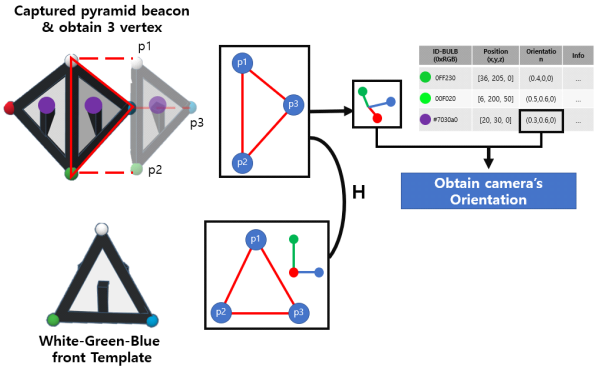
이 갖는 정보를 매핑하고 측위부에서 대조할 비콘을 찾는데 활용된다. 이를 통해 시스템은 여러개의 비콘을 동시 관리 할 수 있으며 나아가 증강현실 디바이스와 동적인 상호작용 또한 가능하다.

비콘의 자세 추정을 진행하기 전에, 촬영 이미지에서 비콘의 위치와 영역 검출을 위한 작업이 필요하다. 이는 학습된 YOLO v3에 의해서 수행되며 학습은 비콘 이미지를 그림 3과 같이 회전, 조도변화, 스케일링 되어 합성한 이미지 셋에 의해 진행된다. 합성을 위한 배경 데이터 셋은 VOC 데이터를 참조하였다. 학습 결과 몸체가 가려지거나 과도한 변화가 있는 비콘을 제외하고는 그림 3과 같이 정확하게 검출되었음을 확인하였다.

(그림 3) 테스트 이미지 셋 검출 결과

비콘의 자세 검출은 그림 4와 같이 Homography를 활용하여 진행된다. 먼저 포착된 비콘의 3개의 꼭지점을 추출한다. 서로 다른 색으로 이루어진 추출된 3개의 꼭지점을 특징점 삼아 미리 저장된 3개 꼭지점의 정면 자세 이미지와 Homography의 H행렬을 획득하고 이를 이용해 포착된 비콘의 자세를 추정한다. 마지막으로 획득한 ID Bulb 색상을 활용해 비콘 서버에서 해당하는 비콘의 절대적인 자세 값을 얻어와 포착된 비콘의 자세와 대조하여 카메라의 자세를 검출한다.

상대적 거리의 검출은 스마트폰 카메라의 미리 정의된 화면비를 통해 구한다. 비콘의 각 모서리는 일정한 길이로 설계되어 있다. 모서리의 길이를 스마트폰 화면에서 갖는 길이를 비교하여 거리를 얻을 수 있다. 이를 위해 이전 과정에서 획득한 H 행렬을 활용해 포착된 비콘의 이미지를



(그림 4) 포착된 비콘을 통한 카메라 자세 추정

정면 이미지로 변환하는 과정이 필요하다. 마지막으로 추출한 거리와 카메라 자세 정보를 비콘 관리 서버 내에 존재하는 해당 비콘 좌표와의 차이를 구하면 디바이스의 x,y,z 위치와 자세 정보를 추출할 수 있다.

4. 결론

본 논문에서는 스마트폰 증강현실 환경에서 카메라를 통한 딥러닝 기반 실내 측위 시스템을 제안하였다. 본 시스템은 미리 조형된 피라미드 형태의 비콘과 관리 서버를 활용하여 포착된 비콘과의 상대적인 위치정보를 통해 측위하는 방식으로 비콘을 검출 하기 위한 방식으로 상대적으로 가볍고 빠른 딥러닝 모델인 YOLO v3를 활용하였고, 영상내 3차원 비콘의 변환관계를 설명하기 위하여 homography를 활용하였다. 제안하는 시스템은 피라미드 비콘에 포함된 ID Bulb로 인해 마커 기반과 유사하지만 보다 유연한 관리가 가능한 기술이며 몰입도 높은 증강현실 구현에 도움을 줄 것이다. 본 시스템은 측위를 보조할 다른 기술은 포함하지 않았으며 따라서 카메라 내에 비콘이 포착된 경우에 한해서만 측위가 가능하도록 설계되었다. 때문에 만약 길 찾기와 같은 지속적인 측위가 가능한 기능의 경우에는 Dead Reckoning을 통하여 비콘이 포착되지 않는 공간을 보조 측위하는 등 보다 발전된 기법이 필요하다고 사료된다.

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2017R1A2B4008886).

참고문헌

[1] Redmon, Joseph, Ali Farhadi "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

[2] 이상우, 김선우 "실내 위치측위 기술 동향 및 전망."

한국통신학회지 (정보와통신) 32.2 (2015): 81-88.
[3] Kim, Hyun-Seung, et al "An indoor visible light communication positioning system using a RF carrier allocation technique." Journal of Lightwave Technology 31.1 (2012): 134-144.
[4] Mulloni, Alessandro, et al "Indoor positioning and navigation with camera phones." IEEE Pervasive Computing 8.2 (2009): 22-31.
[5] Sukthankar, Rahul, Robert G. Stockton, and Matthew D. Mullin "Smarter presentations: Exploiting homography in camera-projector systems." Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001. Vol. 1. IEEE, 2001.