

See-Direct 서비스를 위한 이동성 객체 인식 기법

이형석, 이재호*, 김도형**, 류철***
한국전자통신연구원

hyslee@etri.re.kr, *bigleap@etri.re.kr, **dhkim@etri.re.kr, ***ryuch@etri.re.kr

Technique to recognize movable object for See-Direct service

Hyung-Seok Lee, Jae-Ho Lee*, Do-Hyung Kim**, Cheol Ryu***
Electronics and Telecommunications Research Institute

요약

스마트 안경과 같은 기기에서 사용자의 시선에 있는 대상 객체를 인식하고 연관 서비스를 제공하는 것이 중요할 것이다. 본 논문은 이러한 See-Direct 서비스를 위하여 이동 가능한 객체를 인식하는 방법을 제안하고 성능을 실험하였다. 본 연구에서 카메라를 통한 시선 상의 이미지로부터 객체의 유형을 판별하는 딥러닝 신경망과 객체의 위치를 표현하고 보여주기 위한 기술로 구글 ARCore의 앵커를 활용하여 객체가 이동되었을 때 이를 추적하고 인식할 수 있다.

I. 서론

스마트 안경과 같은 기기[1]에서 사용자의 시선에 있는 대상 객체를 인식하고 통신하는 기술[2]이 중요하다. 시선의 객체들을 제어하거나 관련 정보를 제공하는 See-Direct 서비스[2]를 위한 대상 객체의 인식은 먼저 같은 공간에 있는 같은 유형의 서로 다른 대상을 구별할 수 있는 능력이 필요하다. 카메라 시선의 이미지로부터 객체의 분별하는 방법만으로는 이러한 문제를 해결하는 데는 한계가 있다. 만일 객체의 위치가 고정되어 있다면 주변의 이미지 차이를 활용하여 구별할 수 있을 것이지만, 이를 위해서는 같은 유형이라고 하더라도 주변 배경이 다른 대상들의 이미지를 확보하여 신경망을 학습이 필요하다. 만일 객체들이 이동 가능하다면 위치가 바뀔 때마다 이미지 데이터 획득과 신경망 학습을 다시 해야 하는 문제가 있어 만일 이동이 빈번한 상황에서는 실제 적용이 어렵다.

본 논문에서는 이동 가능한 객체를 인식하기 위하여 이미지 기반의 딥러닝 신경망 기술과 더불어 공간상의 위치를 판별할 수 있는 기술을 활용하는 한 방법을 제시한다. 이미지 기반의 객체 유형 인식을 위해서는 inception v3 딥러닝 신경망을 사용하였다[3]. 공간상의 위치를 인식하고 추적하기 위해서 동시적 위치 추적 및 지도 작성(SLAM, Simultaneous localization and mapping) 기술을 제공하는 구글 ARCore 앵커 기술을 활용하였다. 이를 활용하여 객체에 대한 공간상의 위치 정보를 추적하여 이동이 있는 경우라 하더라도 인식할 수 있다.

II. 본론

본 논문연구의 시스템 구성은 이미지 기반의 딥러닝 신경망을 활용한 객체 유형 인식 및 객체 관련 정보를 관리하는 서버와 객체 인식을 수행하는 단말로 구성된다. 연구를 위하여 8개의 유형에 대하여 각각 100매 이상의 이미지 데이터를 촬영 수집하여 서버의 신경망을 학습하였다. 8개의 유형은 램프 2종, 스피커 3종, 프로젝터, 프린터, 전자레인지 각각 1종으로 구성된다.

이동 가능한 객체의 위치 정보는 수시로 변동 가능하므로 서버에서 객체의 다른 속성들과 따로 분리하여 관리한다. 변하지 않는 객체의 속성으로는 ID, 이름, 유형, 통신 주소 등이 포함된다. 통신 주소는 블루투스

MAC 주소이다. 하나의 공간은 방(room)으로 표현되는데 서버에 저장되는 객체의 위치 정보는 소속된 방과 앵커 ID로 표현된다.



그림 1. 대상 객체 등록 화면

단말에서 대상 객체들 정보를 서버에 등록하고 이들의 초기 위치를 설정하는 단말용 앱을 또한 개발하였다. 그림 1은 이 앱을 통한 객체를 등록하는 것을 보여준다. 그림 2는 객체 초기 위치를 ARCore 앵커를 기술로 기반으로 등록하는 것을 보여준다. 등록된 객체의 정보는 서버에 저장되어 이를 인식 앱에서 이를 활용할 수 있게 된다.



그림 2. 대상 객체 위치 정보 등록 화면

인식 앱의 동작은 사용자가 방을 선택하면서 시작된다. 그러면, 서버로부터 해당 방에 존재하는 객체들의 위치 정보인 앵커 ID를 얻어 온다. 이를 가지고 ARCore에서 계속해서 주변을 스캔하여 얻은 시각적 데이터로부터 지정된 방에서의 앵커 객체들을 찾아낸다.

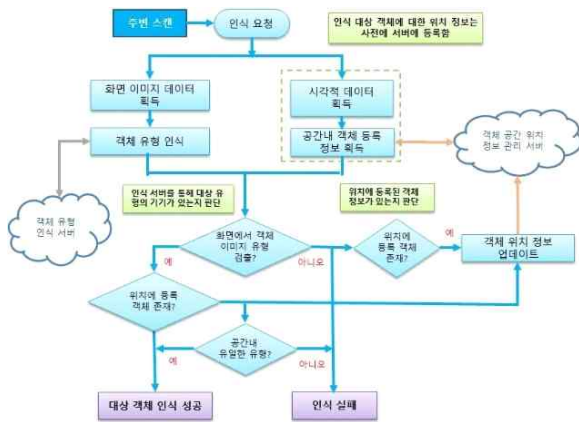


그림 3. 이동형 대상 객체 인식 절차도

See-Direct 서비스를 위한 이동형 대상 객체의 인식 절차는 그림 3에서 보여준다. 사용자가 화면에 보이는 사각의 관심 영역 안에 원하는 객체를 위치시키고 인식을 요청하면 해당 이미지를 획득하여 서버로 유형 인식을 요청한다. 동시에 이미지 내에 기준에 얻은 앵커가 있는지 검사한다. 서버로부터의 객체 유형 인식 결과와 기준에 등록된 객체 위치(앵커)의 존재 유무를 종합하여 인식 결과를 처리한다. 이미지로부터 인식된 객체의 위치에 이미 얻은 앵커가 존재한다면 최근 위치가 등록된 후 이동이 없는 상태를 의미한다. 그림 4는 인식 앱에서 방을 선택하는 화면과, 실린더형 램프 객체의 위치가 서버에 저장된 정보와 같은 경우에 이를 인식하고 램프를 켜는 제어를 하는 화면을 보여준다.



그림 4. 대상 인식 및 제어 동작 화면



그림 5. 이동된 대상 인식 동작 화면

앵커는 존재하지만 해당 위치에 객체가 없는 경우 다른 곳으로 이동해 갔음을 의미한다. 또, 화면상의 이미지에 객체가 검출되었는데 해당 앵커가 존재하지 않는다면 다른 곳에서 이동해 온 것을 의미한다. 같은 유형의 이동해 간 객체가 있다면 이에 대한 객체의 위치를 새로운 앵커로 등록하고 이를 서버에 저장한다. 그림 5는 객체가 이동되었을 때 새로운 위치를 등록하는 동작 화면을 보여준다.

만일 같은 유형의 이동해 간 객체가 여러 개 있다면 이들 중에서 사용자가 선택할 수 있게 한다. 이렇게 함으로써 이동 가능한 대상 객체들을 인식할 수 있다.

본 연구를 위한 실험으로 4개 유형의 총 6개의 대상 객체를 사용하였다. 이들은 사각램프A, 사각램프B, 실린더 램프, JBL스피커A, JBL스피커B, LG스피커다. 각 대상마다 10번의 인식 실험을 하였으며, 이 중 5번은 동록한 위치에 고정난채 하였고, 나머지 5번은 이동 후에 하였다.

평가 방법은 대상 객체에 대한 인식 매칭값이 주어진 임계값 이상일 경우만 인식 성공으로 처리하였다. 카메라 시선에 객체에 대응되는 앵커를 인지하면 위치 인식 성공으로 처리하였다.

	기기 매칭값 0.7 이상 / 위치인식		
	기기	위치	인식성공
인식 성공 수 (기기 임계값 만족) = 정인식 수 + 오인식 수	56	60	56
정인식 수 (기기 인식에 성공한 경우)	56	60	56
오인식 수 (기기를 오인식한 경우)	0	0	0
인식 실패 수 (기기 인식 임계값을 만족시키지 못하는 경우) = 전체 실험수(=60) - 인식 성공 수	4	0	4
인식률 (%) = 정인식 수 / 전체 실험수(=60) X 100	93.33	100.00	93.33
오인식률 (%) = 오인식 수 / 전체 실험수(=60) X 100	0.00	0.00	0.00
미인식률 (%) = 인식 실패 수 / 전체 실험수(=60) X 100	6.67	0.00	6.67

그림 6. 대상 인식 실험 결과

그림 6에서 임계값을 0.7로 실험한 결과를 보여준다. 총 60번의 실험에서 기기 유형 인식은 56번 성공하여 93.33%의 인식률을 보여주었다. 4번의 실패는 색상이 어두운 JBL 스피커에서 나왔는데 다른 객체로 오인식하지는 않고 미인식의 결과가 나왔다. 이는 객체의 이미지 색상이 어두워 약간 어두운 환경에서 배경과의 차이를 잘 분별하지 못한 것에서 기인한 것으로 추정된다. 객체의 위치 인식은 60번 모두 100% 성공하였다.

III. 결론

본 논문연구에서는 유형 분별 딥러닝 기법과 SLAM 기술을 활용하여 이동 가능한 대상 객체를 인식하는 한 방법을 제시하였다. 실험 결과에서 제시한 방법이 적절한 환경에서 유효하게 동작하는 것을 알 수 있었다. 그러나, 본 연구에서 사용한 단순 유형 분별 기반으로 사용자가 인식을 위한 동작을 취할 때만 이미지를 통한 유형 검출을 수행하므로 객체의 이동을 지속해서 추적하는 데 다소 단점이 있었다. 이를 개선하기 위한 차기 연구로 YOLO와 같이 화면상의 객체들을 다중 검출할 수 있는 기술을 활용한다면, 주기적으로 화면을 스캔하면서 객체들의 이동 상황을 지속해서 업데이트하는 방법을 생각해 볼 수 있다.

ACKNOWLEDGMENT

본 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (2018-0-00226, 포스트 스마트폰 시대를 대비한 Trusted Reality 핵심기술 개발)

참고문헌

- [1] K. Song, G. Kim, T. Kim, C. Ryu, S. Park, J.H. Lee, J.K. Lee, and S. Hwang, "Trusted Reality Technology, from a Post-Smartphone Perspective," *Electronics and Telecommunications Trend*, 2018.
- [2] 김선애, 김영훈, 김수창, 이문식, 방승찬. (2013). 보는 대로 바로 연결하는 시선(視選)통신 기술. *한국통신학회지(정보와통신)*, 31(1), 89-93
- [3] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, Zbigniew Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)