

다중 타겟, 다중 카메라 추적을 활용한 인공지능 판매점

황영진¹, 박수진¹, 이수언¹, 이승현¹, 황광일

¹인천대학교 임베디드시스템공학과

yeongjin.gongjin@gmail.com, pppssujin@gmail.com, miantndjs@naver.com, fhzktm@naver.com

AI Market using Multi-Target, Multi-Camera Tracking

Yeong-Jin Hwang¹, Su-Jin Park¹, Soo-Eon Lee¹, Seung-Heyon Lee¹, Gwang-Il Hwang

¹Dept. of Embedded-Systems Engineering, Incheon National University

Abstract

무인 매장의 수가 점점 증가하며 보안과 비용 절감을 위한 솔루션들이 요구되고 있고, 시스템은 이를 해결하기 위한 다중 객체 추적 알고리즘을 제안한다. 한 대의 카메라에서 여러 사용자를 인식하고, 사용자의 정보를 서버에 저장해 주변 카메라에 전달함으로써 적은 카메라 수로 많은 사용자를 추적할 수 있게 한다.

1. Introduction

최근 사람이 없는 매장을 흔하게 볼 수 있다. 특히 기존에는 무인 빨래방, 아이스크림 전문 매장 등에 무인 매장이 한정되었다면 최근에는 이동통신사, 편의점, 과일 가게, 건어물 매장, 밀키트 매장, 프린트 카페, 정육점, 펫샵 등 종류도 다양해졌다. 낮에는 유인매장, 밤에는 무인매장으로 시간에 맞춰 유동적으로 조정하는 하이브리드 방식의 매장이 생기고 있다. 이렇게 무인 매장은 점점 우리 삶에 깊이 스며들고 있다.

국내에서 무인 매장이 확대되는 가장 큰 이유는 인건비 절감을 통한 비용 감소이다. 특히 최근에는 코로나로 인해 매출이 줄어 일손을 구하기 어렵게 되었고, 타 매장과의 경쟁에서 가격 경쟁력을 높이기 위해 무인 매장을 찾게 되는 것이다.

무인 매장은 소비자들이 어느 시간에도 이용할 수 있다는 장점이 있다. 코로나로 인해 영업시간을 제한했던 당시에 무인 매장은 24시간 운영이 가능했다. 또한 아직도 코로나 상황이 심각한 만큼 코로나 문제로 사람을 접하는 것이 꺼려지는 고객들, 사람을 대하는 데 부담을 느끼는 소비자들이 편하게 이용할 수 있다는 것도 중요한 요인이다.

대한민국뿐 아니라 전 세계적으로도 지속해서 무인점포 개발에 힘쓰고 있다. 미국에서는 2016년 아마존(Amazon)에서 세계 최초로 아마존 고(Amazon Go)라는 무인 점포를 만들어 선보였다. 중국의 유통 기업인 빙고박스(BingoBox)는 중국 내 가장 많은 무인 편의점을 보유하며 무인 매장의 영역을 넓

혀가고 있다. 이외에도 중국 알리바바의 타오카페, 일본 세븐일레븐 등 다양한 국가에서 무인 점포에 대한 사업을 확장해나가고 있다.

이러한 무인 매장의 가장 큰 단점은 도난에 취약하다는 것이다. CCTV를 사용해 실시간으로 모니터링되고 있지만, 무인 매장의 특성상 즉각적인 대처를 하기는 어렵다. 실제로 국내 자영업자 온라인 커뮤니티에서는 무인 점포 사장들의 절도에 대한 고발이 끊이지 않고 있다. 그러나 범죄 대응에 무게를 둘수록 비용이 늘기 때문에 비용 절감이 최우선인 국내 유통업계 입장에서는 골치 아픈 문제가 아닐 수 없다. 이 글에서는 범죄에 대처하면서 비용을 줄이는 방안에 대해 제안하고자 한다.

2. Proposal

한편 아마존고와 같은 기존 무인 매장에서 사용하는 사용자 추적 방식은 고객 한 명 당 센서가 부착된 카메라 한 대가 사용되기 때문에, 매장 내 많은 인원을 수용할 경우 이에 대한 비용이 만만치 않다는 문제가 있다. 이를 해결하기 위해 우리는 멀티 카메라 멀티 트래킹(Multi-Camera Multi-Object tracking) 방식을 도입한 무인 판매 시스템을 제안한다. 기존 연구[1], [2], [3]에서는 효율적이고 정확한 객체 추적을 위한 다양한 연구를 보여주지만 이를 다중 객체 추적에는 적용할 수 없다. [4], [5] 연구는 실시간 다중 객체 추적에 대한 효율적인 알고리즘을 제안하지만, 단일 카메라 또는 비디오를 사용하는 경우로 제한된다. 본문에서는

다중 카메라를 사용한 다중 객체 추적을 활용하여 이를 무인 매장 모니터링에 적용하고자 한다. 기존 무인매장에서는 하나의 카메라로 하나의 대상을 인식하여 트래킹하는 대신 사용자 수만큼의 카메라가 필요했다. 그러나 멀티 카메라 멀티 트래킹 방식을 적용할 경우, 한 대의 카메라에서 여러 사용자를 인식하고, 사용자가 이동할 경우 이 정보를 주변 카메라에 전달하여 트래킹 상태를 유지할 수 있으므로 적은 카메라 수로도 많은 사용자를 추적할 수 있게 된다. 이를 무인 매장 시스템에 적용하게 되면 카메라를 필요 이상으로 설치하게 됨으로써 발생하는 초기 비용을 줄일 수 있게 된다. Implementation에서는 이에 대한 구현을 설명한다.

3. Implement

3.1 Duke MTMC-reID Dataset

Duke MTMC-reID (Duke Multi-Tracking Multi-Camera ReIdentification) Dataset[6]은 이미지 기반 ReIdentification을 위한 Duke MTMC Dataset의 하위집합으로 1080p, 60FPS의 8개 카메라에서 14시간 이상의 동기화된 감시 비디오로 만들어졌다. 데이터셋은 Bounding-Box에 의해 비디오에서 이미지를 추출했으며 702개의 ID의 16,522개의 훈련 이미지, 다른 702개의 ID의 2,228개의 쿼리 이미지 및 17,661개의 갤러리 이미지로 구성 되어있다.

3.2 Tracking

Single-Camera Tracking : Deep Sort[7]를 사용하여 3frame 이전의 영상 이미지에서 현재 frame의 개체의 위치를 예측하고 측정한다. Deep Sort의 주요 기능인 Kalman filter Bounding Box 요소로 데이터를 처리한다. 3frame 마다의 영상 이미지 정보를 가져와서 각 사람마다 바운딩 박스를 그리고, 특정 ID를 부여한다.

Kalman filter state – Bounding Box : 영상에서 가져온 이미지에서의 개체들은 $[x, y, a, h, vx, vy, va, v, h]$ 의 요소를 가진다. (x, y) 는 Bounding Box의 중심 위치, a 는 가로-세로 비율, h 는 높이 정보를 나타낸다. vx, vy, va, vh 는 각 요소들의 속도를 나타낸다.

Multi-Camera Tracking : 우리는 여러 제약조건을 제시하고 있다. 1. frame 한 카메라에서의 다른 좌표에 있는 사람이 동일한 사람일 수 없다. 2. 한 카메라에서 비교적 중앙에 위치한 사람은 갑자기 없어지거나 threshold 이상의 좌표로 갑자기 이동할 수 없다. 3. 서로 다른 카메라가 촬영중인 장소에서 겹치는 공간을 제외하면 서로 다른 카메라가 촬영 중인 시간 또는 frame 이 t_1 또는 $frame_1$ 이라고 가정할 경

우 t_1 또는 $frame_1$ 에 있어서 같은 사람이 다중 카메라뷰에 비춰질 수 없다.

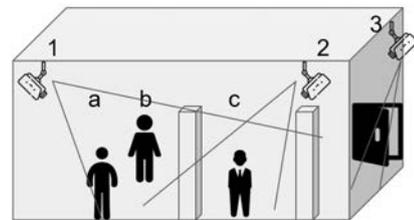
우리가 제시한 제약조건 내에서 오류가 발생할 경우를 검출해내며 이 오류 검출 과정은 2 회이상 반복하고 있다. Single-Camera Tracking에서 ID를 통해 사람을 구분하고 ID에 따라 제약조건 내에서 오류를 검토하고 있다.

3.3 Data Split

Duke MTMC-reID Dataset에서 데이터 셔플링 과정을 거치고 60% train, 10% Validation, 30% test dataset으로 분할하여 모델 학습을 하였다.

3.4 System Implementation

아래 그림과 같은 system이 있다고 가정할 때, 우리는 위 정보들을 통해 학습된 모델을 사용하여 실시간 비디오 데이터를 3frame 마다 받아와서 각 카메라들에 보이는 유저에 고유 ID를 부여한다. 각 유저들을 Tracking 하며 구매하고자 하는 상품을 선택했을 때 서버에 [ID] list에서 product들을 append한다. 최종 3번 Camera에서 특정 좌표를 통해 사라지는 유저 ID를 파악하고 최종적으로 list에 추가된 상품들을 서버에 저장되어 있는 [$Product_1, Price_1, Product_2, Price_2, \dots$] list에서 각 상품의 가격정보를 받아와서 최종합계를 한다.



4. Conclusion

최저 임금의 상승으로 인건비 절감을 위해 각종 산업에서 무인 매장이 증가하며 이에 따라 생활 범죄와 비용 절감을 위한 기술들이 요구되고 있다.

기존 무인 매장에 적용된 사용자 추적 방식의 경우에는 고객 한 명 당 센서가 부착된 카메라 한 대가 사용되기 때문에, 매장 내 많은 인원을 수용할 경우 이에 대한 비용 부담의 문제가 있다. 이를 해결하기 위해 우리는 멀티 카메라 멀티 트래킹(Multi-Camera Multi-Object tracking) 방식을 도입한 무인 판매 시스템을 제안한다.

시스템은 Duke MTMC-reID Dataset을 기반으로 모델 학습을 진행하며, Deep Sort를 사용해 3frame 이전의 영상 이미지에서 현재 frame의 개체의 위치

실시간으로 예측하고 측정한다. 한 대의 카메라에서 여러 사용자를 인식하고, 사용자의 정보를 서버에 저장해 주변 카메라에 전달함으로써 적은 카메라 수로도 많은 사용자를 추적할 수 있게 된다.

Reference

- [1] Xi Chen¹, Zuoxin Li, Ye Yuan, Gang Yu, Jianxin Shen, Donglian Qi. State-Aware Tracker for Real-Time Video Object Segmentation. CVPR 2020.
- [2] Xiao Wang, Xiujun Shu, Zhipeng Zhang, Bo Jiang, YaoWei Wang, Yonghong Tian, Feng Wu. Towards More Flexible and Accurate Object Tracking with Natural Language: Algorithms and Benchmark. CVPR 2021.
- [3] Xiao Wang, Jianing Li, Lin Zhu, Zhipeng Zhang, Zhe Chen, Xin Li, YaoWei Wang, Yonghong Tian, Feng Wu. VisEvent: Reliable Object Tracking via Collaboration of Frame and Event Flows.
- [4] Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, Xinggang Wang. ByteTrack: Multi-Object Tracking by Associating Every Detection Box. ECCV 2022.
- [5] Nicolai Wojke, Alex Bewley, Dietrich Paulus. Simple Online and Realtime Tracking with a Deep Association Metric. 2017 IEEE International Conference on Image Processing (ICIP). Beijing, China. 2017.
- [6] Ergys Ristani, Francesco Solera, Roger S. Zou, Rita Cucchiara, Carlo Tomasi. Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking. ECCV 2016 Workshop on Benchmarking Multi-Target Tracking In European. 2016.
- [7] Nicolai Wojke, Alex Bewley, Dietrich Paulus. Simple Online and Realtime Tracking with a Deep Association Metric. 2017 IEEE International Conference on Image Processing (ICIP). Beijing, China. 2017.

"본 논문은 과학기술정보통신부
정보통신창의인재양성사업의 지원을 통해 수행된
ICT멘토링 프로젝트 결과물입니다."