

# 사회적 거리 두기를 위한 스테레오 영상과 스켈레톤

## 정보기반 객체 간 거리 추정 방법\*

장규진<sup>01</sup> 배현재<sup>2</sup> 김영민<sup>3</sup> 김영남<sup>1</sup> 김진평<sup>1</sup>

<sup>1</sup>차세대융합기술연구원 컴퓨터 비전 및 인공지능 연구실

<sup>2</sup>성균관대학교 소프트웨어학과

<sup>3</sup>인천대학교 경제학과

gjjang@snu.ac.kr, jason0425@snu.ac.kr, winston1214@naver.com, hwarangmoon@snu.ac.kr, jpkim@snu.ac.kr

## Stereo Image and Skeleton Information based Inter-Distance Estimation for Social Distancing

GyuJin Jang<sup>01</sup> HyunJai Bae<sup>2</sup> YoungMin Kim<sup>3</sup> YoungNam Kim<sup>1</sup> JinPyung Kim<sup>1</sup>

<sup>1</sup>Computer Vision and Artificial Intelligence laboratory, Advanced Institute of Convergence Technology

<sup>2</sup>Department of Computer Science and Engineering, Sungkyunkwan University

<sup>3</sup>Department of Economics, Incheon National University

### 요 약

최근 COVID-19와 같은 감염전파율이 높은 전염병 확산 방지를 위해 공공장소에서 최소 1m 이상 거리를 유지하는 사회적 거리 두기(Social Distancing)를 통하여 개인 또는 집단 간 접촉을 최소화하는 대책이 시행되고 있다. 그러나 위험도가 높은 다중이용시설이나 불특정 다수가 밀집하는 대중교통 등 공공장소의 관리가 현실적으로 어려우며, 모노 카메라로 구성된 기존 CCTV영상을 이용한 거리두기 모니터링 사례에서는 거리추정에 오차가 발생하여 문제점을 보인다. 이를 보완하기 위해서 본 논문에서는 다중객체를 대상으로 하는 스테레오 비전과 스켈레톤 정보를 이용한 거리추정 방식을 제안한다. 입력된 영상에서 YOLO를 통해 탐지된 객체 경계상자(Bounding Box)에서 관절 키포인트(Keypoint)를 추출하고 2차원 좌표의 스켈레톤 정보를 생성한다. 스테레오 영상에 대한 카메라 캘리브레이션과 시차(Disparity)를 이용하여 깊이 정보(Depth Information)를 획득하고 이를 통해서 3차원 공간좌표로 변환한다. 3차원 키포인트의 좌표 값에서 객체별 중심점을 계산하고 유클리디안 거리(Euclidean Distance)를 적용하여 사람들 간의 거리를 추정한다. 1m를 기준으로 약 10cm 이내의 거리 추정오차를 보였으며, 위의 방식을 기반으로 다중객체 사이의 거리를 추정하여 사회적 거리 두기 위반을 탐지할 수 있는 시스템을 제시한다.

### 1. 서 론

COVID-19가 2019년 12월에 중국에서 처음으로 발병한 이후 1년이 지난 지금까지 전 세계적으로 유행하며 팬데믹(pandemic) 상황을 유지하고 있다. 공공장소 및 다중이용시설에서 사람들 사이의 감염확산을 막기 위하여 세계보건기구(WHO; W.H. Organization)는 사람 사이의 접촉 가능성을 감소시키고자 물리적 거리두기 일환으로 대략 2m(6 feet) 이상을 권고한다. 우리나라의 경우에는 거리두기 1단계부터 실내/외에서 최소 1m 거리를 유지해야 한다. 공공장소와 같은 밀집 지역에서 물리적 거리두기를 감시하기 위한 인력배치 및 CCTV, 드론 등 다양한 방식을 활용하나 실제로 운영인력이 부족한 상황이다.

최근, 스마트시티(Smart City) 분야에서는 문제해결을 위해 Computer Vision, AI기술을 이용한 사회적 거리두기 모니터링 등 공공안전을 위한 기술연구가 활발히 진행되고 있다[1][2].

객체 사이의 거리를 추정하기 위해서는 우선 영상 내에서의 객체 탐지를 통해 사람이 존재하는 정확한 위치를 찾아야 한다. 현재 많이 사용되는 객체를 탐지방법은 딥러닝 기반의 YOLO[3], Mask R-CNN[4] 등이 있다.

이러한 객체 탐지방법 결과는 경계 상자(Bounding Box)로 이미지 내에서 탐지된 객체 주위를 직사각형 형태로 표시하는 것을 의미하며 이를 통해 객체의 위치정보를 획득하게 된다. 경계 상자 형식의 사람 감지 경우 다양한 신체 포즈 변화(예 : 앉기, 구부림 등)가 고려되지 않아 정확하지 않은 객체 중심위치(좌표)를 얻게되며 이를 거리 추정에 활용할 경우 물리적(사회적) 거리 분석에 오차가 발생할 수 있다[5]. 다음은 탐지된 객체의 중심좌표를 이용한 객체 간의 거리 추정과정을 진행한다.

CCTV에서 일반적으로 사용하는 단안 카메라에서는 원근변환(Perspective Transformation)방식을 이용하여 객체 사이의 거리를 추정한다[6]. 그러나 카메라가 설치된 위치에 따라 성능에 영향을 받으며 2차원 이미지로 투사되었기 때문에 깊이 정보(Depth Information)가 손실되어있는 상태에서 정확한 거리 추정에는 한계가 있다[7].

본 논문에서는 향상된 다중객체 사이 거리 추정을 위해 다음과 같은 방식을 적용한다. 먼저 실시간 처리가 가능한 AlphaPose[8]를 선정하여 관절에 대한 키포인트(Keypoint)를 추출하고 2차원 좌표인 스켈레톤 정보를 생성한다. 그리고 스테레오 비전(Stereo Vision)을 이용하여 깊이 정보를 구하고 앞선 키포인트의 좌표를 3차원 공간좌표로 변환한다. 다음으로 3차원으로 변환된 스켈레톤 정보를 통해 정확한 객체별 중심위치를 구한다. 객체별 중심위치에 대하여 유클리디안 거리(Euclidean Distance)를 계산하고 사람들 간의 거리를 추정한다. 이런 방식을 이용하여 사회적 거리(Social Distancing)가 1m 이

\* 이 논문은 행정안전부 극한재난대응기술개발사업의 지원을 받아 수행된 연구임(2020-MOIS31-014)

하일 때 경고할 수 있는 자동화 된 시스템을 제안한다.

## 2. 사회적 거리두기 위반 탐지 시스템

본 논문에서 제안하는 사회적 거리두기 위반 탐지 시스템의 동작 순서는 그림 1과 같다.

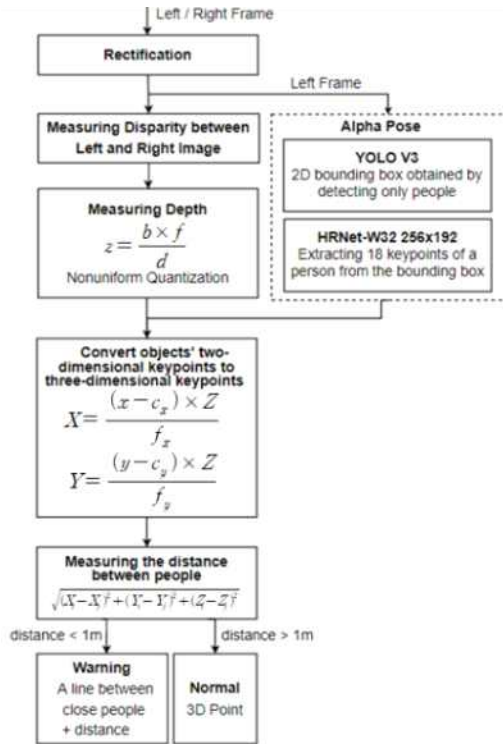


그림 1 사회적 거리 위반 탐지 시스템 동작과정

스테레오 비전에서 입력받는 좌/우 영상에서 대응점 위치의 차이를 이용하여 시차(Disparity)를 계산한다. 좌측영상에 Alpha-Pose를 적용하고 대상 객체에서는 최대 18개의 2차원 관절 키포인트를 추출한다.

$$depth(z) = \frac{focal\ length(f) \times baseline(b)}{disparity} \quad (1)$$

다음으로 수식(1)에서 추출된 키포인트에 해당하는 깊이(Depth)를 계산하기 위해 시차를 입력값으로 대입한다. 시차와 깊이는 반비례 관계임으로 Baseline이 멀수록 정밀한 깊이 정보를 얻을 수 있으나, 본 논문에서는 사용하는 ZED2 Camera는 Baseline이 고정된 형태이기 때문에 시차 편차를 보정하기 위한 양자화(Quantization)를 수행하여 깊이 오차를 감소시켰다. 이후, 카메라 캘리브레이션을 통하여 추출된 키포인트를 3차원 공간좌표로 변환하였다. 객체별 추출된 최대 18개 이내의 키포인트로 이루어진 스켈레톤 정보에서 객체를 대표하는 중심위치를 구하기 위해 수식(2)를 적용하며 N은 관절포인트 개수이다.

$$X = \frac{1}{N} \sum_{n=1}^N X_n, \quad Y = \frac{1}{N} \sum_{n=1}^N Y_n, \quad Z = \frac{1}{N} \sum_{n=1}^N z_n \quad (2)$$

$$Distance_{ij} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2 + (Z_i - Z_j)^2} \quad (3)$$

객체별 중심위치 정보에 수식(3)과 같이 유클리디안 거리(Euclidean Distance)를 적용하여 객체 간 사이의 거리를 추정하게 된다.

## 3. 실험 결과

실험 환경은 3.6GHz의 Intel(R) core i7-9700K CPU와 NVIDIA Geforce GTX 1660 Ti Graphic Card, 32GB RAM이 하드웨어 구성 및 PyTorch 1.1.0 버전이 설치된 데스크톱 PC에서 시뮬레이션을 수행하였다.

본 실험에서 객체 사이 거리 추정 성능의 신뢰도를 확인하기 위해 3차원 중심위치정보를 활용한 거리 추정 성능에 대한 검증 수행하였다. 스테레오 카메라로부터 1m부터 4m까지 거리에서 일정 간격(1m)을 두고 위치를 반복적으로 10회 이동하며 거리 추정에 대한 실험을 진행했다. 그림 2의 좌측 이미지는 시차를 0-255 사이로 정규화한 시차맵(Disparity Map)을 나타내고 우측 영상은 표시된 수치는 탐지된 객체에 대한 중심좌표(x, y, z)에 대한 정보를 의미한다. 여기서 z값은 수식(2)에서 깊이 정보를 이용한 거리추정에 대한 결과이다.



그림 2 시차맵(Disparity Map)과 거리 추정 결과(1m 기준)

1m 거리 기준에서 추정거리 오차가 3cm 이하로 높은 신뢰도를 보였다.

두 사람 사이의 사회적 거리 두기(Social Distancing) 위반 탐지 성능을 확인하기 위해서 그림 3과 같이 고정된 사람을 기준으로 1m 반경 영역을 설정하였고 6개 방향에서 두 가지 방식의 실험을 진행하였다.

첫 번째는 고정된 사람을 기준으로 1m 거리를 두고 이동하였을 때 거리 추정결과에 대한 성능을 확인을 위하여 해당 위치에서 10회씩 반복수행하였다.

표1과 같이 6개의 방향 위치에서 객체 간 1m 거리에 대한 추정결과는 평균 오차 0.098m로 우수한 성능을 확인할 수 있다.

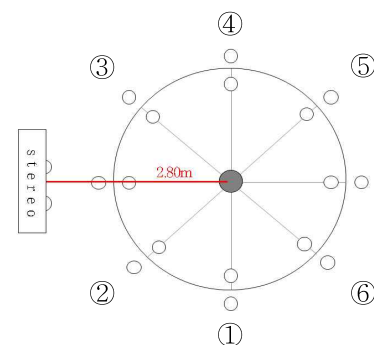


그림 3 객체 사이 거리 추정을 위한 영역 설정

두 번째는 위 결과를 바탕으로 사회적 거리 두기 위반 경우를 탐지하고 경고할 수 있는 시스템을 적용하기 위한 인식성능 실험을 진행하였다. 6개의 지점에서 영역의 안쪽과 바깥으로 이동하며 거리가 1m 미만이거나 1m 이상인 상황을 연출하였으며 이에 대한 인식 가능 여부를 확인하였다.

표 1 6개 지점의 반경 1m 영역에서 대한 거리 추정결과

	①	②	③	④	⑤	⑥
1st	0.99	0.95	0.96	0.86	0.93	0.97
2nd	0.96	0.96	0.97	0.96	1.24	0.98
3rd	0.95	1.05	1.05	0.98	1.02	1.27
4th	1.00	0.97	0.96	1.00	1.11	0.96
5th	0.98	0.95	0.96	0.99	0.95	1.11
6th	1.00	1.13	1.07	1.05	1.27	0.98
7th	1.04	1.01	0.98	1.02	0.98	1.07
8th	1.02	1.12	1.03	1.00	0.99	1.09
9th	0.99	1.01	1.13	1.07	1.20	0.96
10th	1.00	0.96	1.02	0.99	1.00	1.23
Average Error[m]	0.02	0.05	0.05	0.04	0.10	0.10

그림 4는 ①, ②, ③ 지점에 대한 거리 추정 실험화면으로 좌측 이미지는 객체 사이 거리가 1m 이상인 경우이고 우측 이미지는 사회적 거리 유지를 위반한 상황이다. 1m 이하의 간격에 있는 사람들은 빨간선으로 연결하고 추정된 거리를 함께 표시하여 위반상황을 경고한다. 이때 연결점은 스켈레톤 정보를 기반으로 객체의 중심위치를 추출한 것이다. 다중객체를 대상으로 거리 추정 및 위반 여부 인식에 대한 실험도 진행하였으며 그림 5와 같이 위반 탐지가 가능함을 보였다.

#### 4. 결론

본 논문에서는 스테레오 영상과 스켈레톤 정보를 활용한 실시간으로 다중객체(사람) 간의 거리를 추정하는 방식을 제안하고 이를 적용한 사회적 거리 두기 위반 자동 탐지 시스템을 설계하였다. 거리 추정 및 물리적 거리위반 탐지 성능은 실험을 통하여 입증하였다. AlphaPose를 이용한 스켈레톤 정보로 정확한 객체 중심위치 얻고 스테레오 비전으로 얻은 깊이 정보 통해 3차원 좌표로 변환하여 사람 간의 거리를 유클리디안 거리로 추정한 결과 오차가 작음을 보였다. 차후 기존 CCTV가 설치되는 동일한 인프라 환경에 적용하기 위해서 다중객체의 겹침이 발생할 때도 거리 추정에 대한 정확도 향상 방안에 관련한 연구도 필요하다.

#### 참고 문헌

- [1] Jaiswal, R., Agarwal, A., Negi, R., "Smart solution for reducing the COVID-19 risk using smart city technology." IET Smart Cities 2.2 (2020): 82-88.
- [2] Saponara S, Elhanashi A, Gagliardi A. "Implementing a real-time, AI-based, people detection and social distancing measuring system for Covid-19." J Real Time Image Process. 21:1-11 (2021)
- [3] Joseth Redmon, et al. "You Only Look Once : Unified, Real-Time Object Detection." Proceedings of the IEEE International Conference on Computer Vision. 2016.
- [4] He Kaiming, et al. "Mask R-CNN." Proceedings of the IEEE International Conference on Computer Vision, 2017.
- [5] Aghaei, M., Bustreo, M., Wang, Y., Bailo, G., Morerio, P.,Bue, A.D "Single Image Human Proxemics Estimation for Visual Social Distancing." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021.
- [6] Yang, D., Yurtsever, E., Renganathan, V., Redmill, K. A., Özgüner, Ü. "A vision-based social distancing and critical density detection system for covid-19." arXiv preprint arXiv:2007.03578 (2020): 24-25.
- [7] Zaarane, A., Slimani, I., Al Okaishi, W., Atouf, I., Hamdoun, A., "Distance measurement system for autonomous vehicles using stereo camera." Array 5 (2020): 100016.
- [8] Fang Hao-Shu, et al. "Rmpe: Regional multi-person pose estimation." Proceedings of the IEEE International Conference on Computer Vision. 2017.

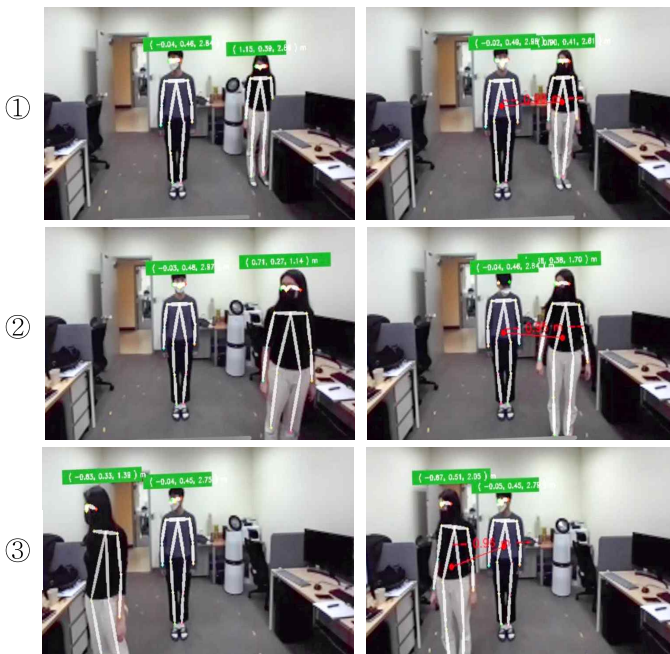


그림 4 사회적 거리 두기 위반 탐지 결과



그림 5 다중객체의 사회적 거리 두기 위반 탐지 결과