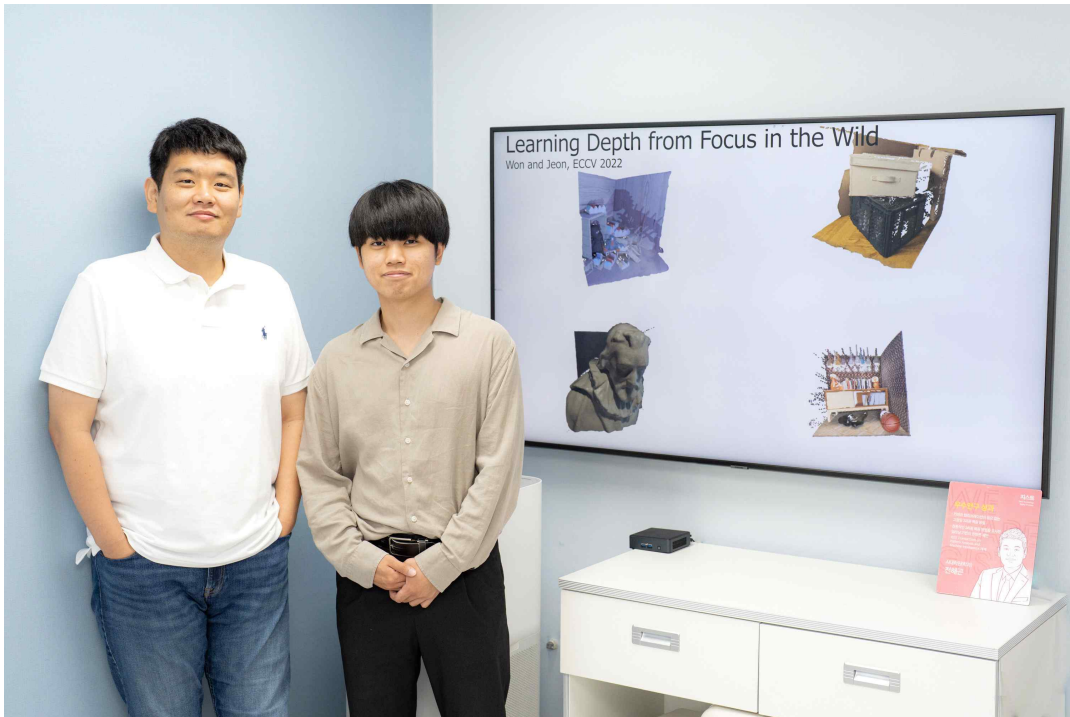


영상 속 3차원 '깊이 정보', AI 알고리즘으로 더 정확히 알아낸다 !

- AI 대학원 전해곤 교수팀, 유럽 컴퓨터 비전 학술대회(ECCV)에서 오는 10월 발표
- 초점거리 기반 세계 최고 성능의 깊이 정보 추정...VR·AR·메타버스 등 응용 분야 활용



▲ 왼쪽부터 전해곤 교수, 원창연 석박통합과정생

지스트(광주과학기술원·김기선 총장) 연구진이 영상 속 '깊이 정보'를 인공지능(AI) 알고리즘을 통해 정확하게 추정할 수 있는 기술을 개발했다.

이번 성과를 적용하면 알고리즘을 통해 영상 속에서 보다 정확한 깊이 정보를 얻어 낼 수 있어 3D 가상현실(VR)·증강현실(AR) 영상 품질 개선, 메타버스, 영상 기반 인지 탐지 등과 같은 컴퓨터 비전 응용 분야의 발전에 기여할 것으로 기대된다.

영상 속의 깊이 정보는 AR·VR의 3차원 공간, 자연스러운 이미지 합성 등 응용하는 여러 분야에서 핵심 요소로 이용되는데, 만약 깊이 정보가 없으면 3차원 공간에 대한 정보가 없어 응용 단계에서 부자연스러운 영상이 구현될 수밖에 없다.

최근 DSLR 카메라나 스마트폰 카메라의 자동 초점 기능을 통해 얻어진 초점 영상들을 활용해 영상의 깊이 정보를 얻는 연구가 구글 등 글로벌 IT 기업이나 여러 대학에서 진행되고 있는데,

기존 방법론들은 좁은 피사계 심도*를 갖는 단일 영상 기반의 깊이 정보를 추정하는 것이 대부분이며, 자동 초점 영상이 정렬*이 되어 있다고 가정하거나 영상 데이터의 특징에 따라 사용자가 직접 추가 설정을 통해 자동 초점 영상을 정렬해야 하기 때문에 깊이 정보의 정확도가 떨어지는 한계가 있다.

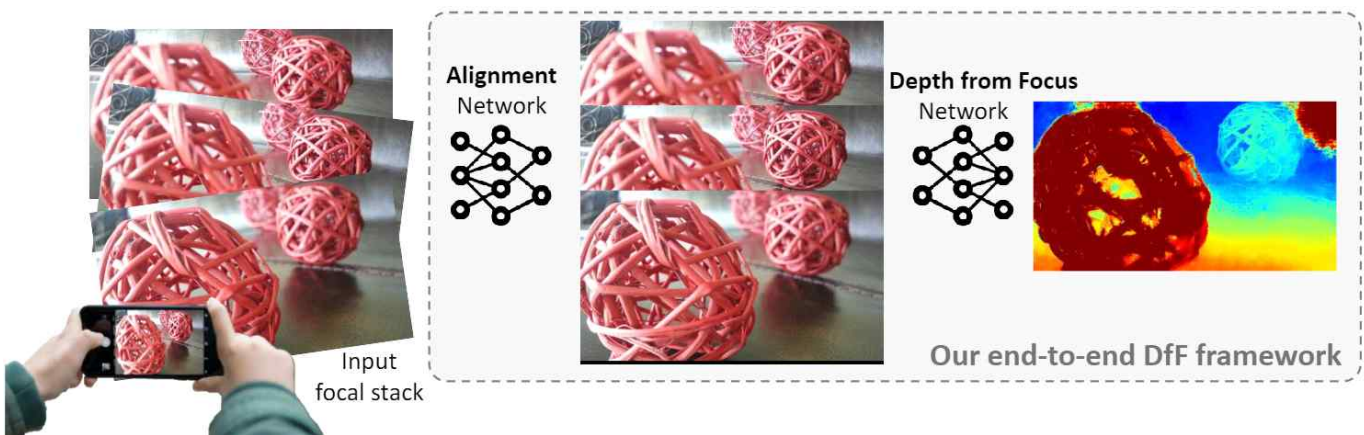
* **피사계 심도(depth of field):** 피사체를 중심으로 전후(前後)로 초점이 맞는 거리의 정도. 피사체의 거리가 멀고 조리개를 좁힐수록 심도는 깊어진다.

* **정렬(alignment):** 이미지 상에서 피사체의 위치를 다른 이미지에서의 위치와 동일하게 맞추는 과정

지스트 AI대학원 전해곤 교수 연구팀은 AI 기법을 통해 자동 초점 기능에서 사용하는 영상 전체를 이용해 깊이 정보를 보다 정확히 추정하는 알고리즘을 고안했다.

연구팀은 카메라에서 출력해주는 메타데이터에서 자동 초점 기능을 통해 얻어진 영상을 인공지능 네트워크로 정렬하고, 이를 이용해 깊이 정보까지 추정하는 엔드-투-엔드 기법*을 세계 최초로 제안했다.

* **엔드-투-엔드 기법:** 알고리즘의 입력과 출력 사이에 있는 모든 과정을 하나의 모델로 학습 및 추론하는 딥러닝 기법

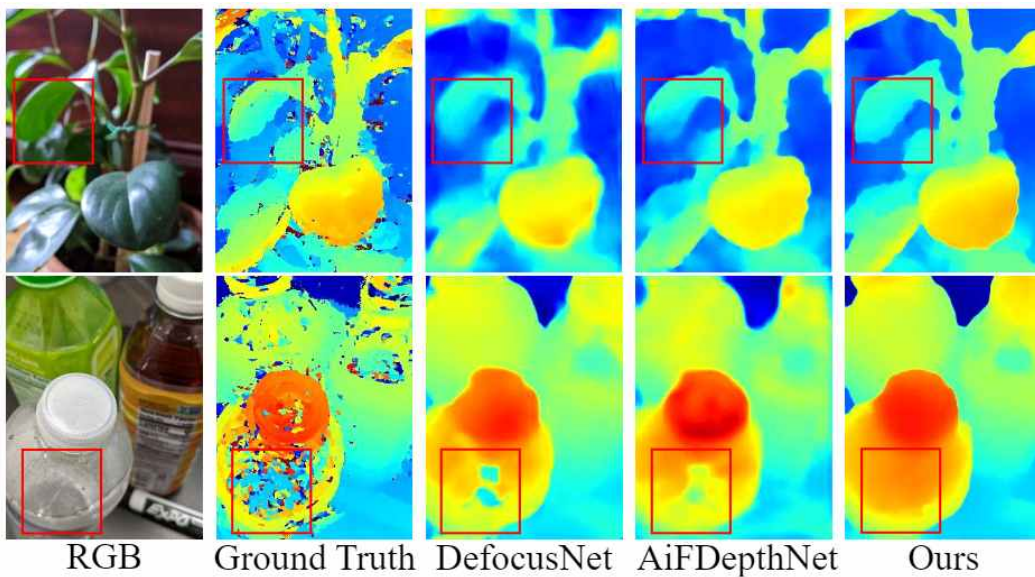


▲ **본 연구에서 제안한 초점거리 기반 깊이 정보 추정 네트워크 구조.** 깊이 정보를 추정하는 네트워크(Depth from Focus network)만 다루었던 기존 연구들과 다르게, 여러 초점에서 촬영된 영상들의 피사체 위치를 동일하게 맞춰주는 정렬 네트워크(Alignment network)와 정확한 깊이 정보를 추정하는 네트워크(Depth from Focus network)가 하나로 합쳐진 엔드-투-엔드 구조를 제안한 것이 본 연구의 특징이다.

촬영한 초점 영상들은 다음과 같은 방법으로 정렬된다. ① 이미지 상이 보이는 정도, 즉 화각(field of view)을 메타데이터로 각 초점 별로 계산한 후 ② 초점 영상들을 특정 화각에 맞게 잘라내 촬영한 모든 영상들의 화각을 동일하게 맞춰준다. ③ 연구팀이 설계한 인공지능 네트워크로 초점 영상들을 추가로 정렬시켜서 잘 정렬된 초점 영상들을 최종적으로 얻을 수 있다.

특히, 이번 성과를 통해 측정된 깊이 정보는 공인 벤치마크*에서 오차와 정확도를 측정하는 총 20개의 척도 중 17개에서 1위를 기록하고 있으며, 다양한 카메라 및 환경 조건에서도 일관적으로 우수한 성능을 보였다.

* **공인 벤치마크:** 본 연구는 독일 뮌헨공과대학교(Technical University of Munich) 연구진의 논문(ACCV2018 학술대회)이 제공한 데이터셋(DDFF 12 Scene)에서 성능을 측정하기 위한 벤치마크(<https://competitions.codalab.org/competitions/17807>)를 활용했다. 이 벤치마크는 실측 깊이 정보와 알고리즘으로 추정된 깊이 정보 간의 오차 및 정확도를 수치화한 19개의 척도와 알고리즘의 깊이 정보 추론 속도를 측정하는 1개의 척도를 통해 알고리즘의 성능을 객관적으로 평가하고 있다.



RGB Ground Truth DefocusNet AiFDepthNet Ours

▲ 맨 왼쪽(RGB)의 경우 일반적인 카메라로 찍을 때 나오는 영상 화면, 나머지 오른쪽 4개의 이미지는 깊이 정보를 컬러 영상으로 나타냄

전해곤 교수는 "이번 연구는 기존에 수동적으로 해결했던 초점거리 영상 정렬과 깊이 정보 추정 알고리즘의 한계를 극복하는 인공지능 모델을 제안한 것"이라며 "인공지능 네트워크를 통해 얻은 깊이 정보를 활용하면 메타버스, 증강현실, 영상 기반 인지 탐지와 같은 응용 분야 발전에 기여할 것으로 기대한다"고 말했다.

시대학원 전해곤 교수(교신저자)와 석박통합과정 원창연(제1저자) 학생이 주도한 이번 연구는 과학기술정보통신부(연구개발특구)와 광주광역시의 지원으로 지스트가 주관하는 '인공지능 기반 메타버스 구현을 위한 융·복합 문화 가상 스튜디오' 과제와 과학기술정보통신부의 'AI혁신 허브' 과제의 지원을 받아 수행되었으며, 인공지능·컴퓨터비전 분야 세계 최고 학회인 European Conference on Computer Vision(ECCV)에서 오는 10월 발표될 예정이다.

논문 및 저자 정보

1. 발표 예정 학회 / 논문명

- European Conference on Computer Vision (ECCV)
 - 인공지능/컴퓨터비전 분야 최우수 학술대회
 - 한국정보과학회 기준 최우수 학술대회(S급)
- 논문명 : Learning Depth from Focus in the Wild