

# GAN 기반 이미지 합성을 통한 3차원 증강 자세 추정

박찬, 문남미  
호서대학교 컴퓨터공학과  
chan.park941003@gmail.com, nammee.moon@gmail.com

## 3D Augmented pose estimation through GAN based image synthesis

Chan Park, Nammee Moon  
Dept. of Computer Science and Engineering, Hoseo University

### 요 약

2차원 이미지를 통한 자세 추정의 경우 관절이 겹치거나 가려져 있는 등의 인식 저해 요소로 인하여 자세 추정 정확도가 감소하는 한계가 있다. 본 논문에서는 GAN을 통해 2차원 이미지를 3차원으로 증강한 뒤 자세를 추정하는 기법을 제안한다. 제안하는 방법은 2차원 이미지의 평면좌표 값에서 GAN을 통해 노이즈 벡터  $z$ 축 값과 피사체에 투영되는 빛의 방향 값을 반영한 3차원 이미지를 만든다. 이러한 이미지 합성 과정을 거친 후 DeepLabCut을 사용해 관절 좌표를 추출하고 자세 추정 및 분류를 진행한다. 이를 통해 2차원에서의 자세 추정 정확도 향상을 기대할 수 있으며, 향후 이를 기반한 이상행동 탐지 분야에서 적용할 수 있다.

### 1. 서론

자세 추정은 이미지 또는 영상에서 관절 위치 등을 파악하여 사람이나 동물이 취하는 자세를 추정하는 것이다. 이는 모션 분석, 상호작용 등 여러 분야에서 사용하고 있다[1]. 2차원 이미지를 통한 자세 추정의 경우 2차원 이미지의 한계인 깊이(Depth)의 모호함, 폐색(Occlusion) 등의 요인으로 인해 추정 정확도가 감소하는 한계가 있다[2]. 이를 보완하기 위해 3차원 모델을 통한 자세 추정을 제안하고 있다[1]. 그러나 3차원 모델을 구현하기 위해서는 양안 입체 영상 카메라 등의 장비를 통해 움직임을 녹화해야 하는 한계가 있다.

최근에는 이미지 합성(Image synthesis)을 통해 2차원 이미지를 3차원 모델로 증강하고, 이를 기반으로 정면에서 찍은 사진을 측면 사진으로 구현하는 연구가 제안되고 있다[3]. 본 논문에서는 2차원 이미지 자세 추정에서 발생하는 정확도 하락의 한계점을 보완하기 위해 GAN(Generative Adversarial Networks)을 사용하여 이미지 합성을 진행하여 3차원으로 증강한다. 이후 증강한 3차원 모델을 기반으로 자세 추정을 진행하고자 한다.

### 2. 관련연구

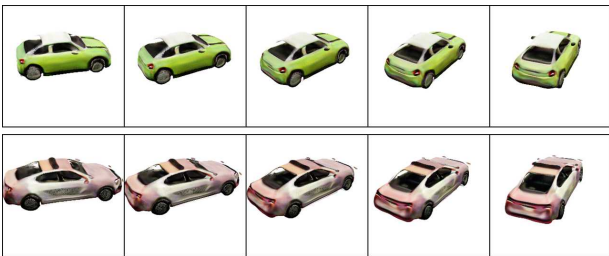
#### 2-1. 자세 추정

자세 추정은 이미지 또는 영상 데이터를 사용하여 인공지능을 통해 각 신체 부위를 탐색한 뒤, 이를 골격(Skeleton) 등으로 나타내는 방법이다[4]. 2차원 자세 추정 연구로는 적외선, RGB-D 두 개의 카메라와 기계학습을 사용하여 쥐의 골격을 나타내고 이를 기반으로 쥐의 자세 추정을 진행한 연구가 있다[5]. 다른 연구는 곤충 또는 동물의 이미지에 Stacked DenseNet 딥러닝 모델을 사용해 골격을 나타내는 연구를 진행하였다[6]. 그러나 2차원 단순 자세 추정은 번짐, 초점 흐림 등의 한계로 인해 추정에 어려움이 있다[7].

이로 인해 양안 입체 영상 카메라인 키넥트(Kinect)를 사용하여 3차원 모델을 구현하고, 자세 추정을 진행하는 연구가 제안되었다[1]. 하지만 3차원 자세 추정을 위해 별도의 특수 카메라를 사용해야 하는 한계가 있다. 본 논문에서는 이러한 한계점을 보완하기 위해, 3차원 동물 자세 추정 연구에 사용된 DeepLabCut을 사용하여[8], 증강한 3차원 모델의 자세 추정을 진행하고자 한다.

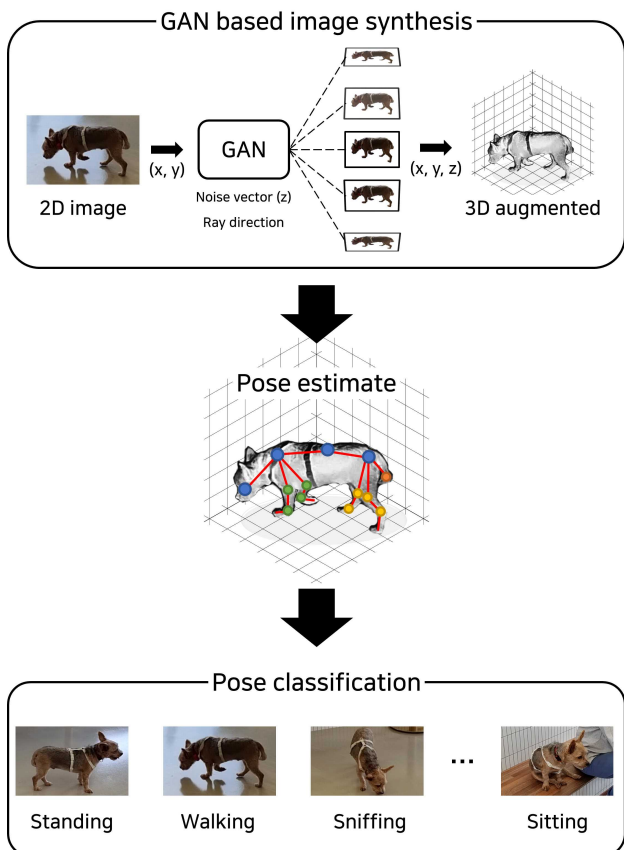
2-2. 이미지 합성을 통한 차원 증강

이미지 합성을 통한 차원 증강은 단일 또는 다수의 이미지 데이터를 이미지 합성을 통해 3차원으로 구현하는 방법이다. 이러한 이미지 합성 연구로는 다양한 각도로 촬영된 2차원 이미지를 학습한 GAN을 사용해 z축에 임의의 노이즈 벡터를 추가하고, 피사체에 투영되는 빛의 방향 값 등을 적분 계산하여 이를 3차원 모델로 구현하는 렌더링(Rendering)을 진행한다[9]. 본 논문에서는 이를 기반한 이미지 합성을 통한 차원 증강 과정을 거쳐 자세 추정을 진행하고자 한다.



(그림 1) GAN 기반 이미지 합성을 거친 3차원 이미지

3. GAN 기반 이미지 합성을 통한 자세 추정



(그림 2) GAN 기반 이미지 합성 및 자세 추정 구성도

본 논문에서는 (그림 1)과 같이 연구를 진행한다. 이미지 합성 및 반려동물의 자세 추정을 위해, 2차원 이미지 데이터는 Stanford Dogs Dataset과 Oxford-IIIT Pet Dataset을 사용한다. 각 데이터셋의 구성은 아래 <표 1>과 같다.

<표 1> 연구에 사용할 데이터셋 구성

	Stanford Dogs Dataset	Oxford-IIIT Pet Dataset
개체 수	120종	37종 (개 + 고양이)
사진 수	20,580장	7,400장

2차원 평면좌표로 이루어진 이미지 데이터를 학습시킨 GAN을 사용하여 z축에 임의의 노이즈 벡터를 추가하고, 피사체에 투영되는 빛의 방향 값을 적분 계산하여 3차원 모델 렌더링을 진행한다. 생성된 3차원 데이터를 기반으로 DeepLapCut을 통해 각 관절 부위의 좌표와 해당 좌표값을 이은 직선을 도출하여 골격을 만들고 이를 통해 자세 추정을 진행한다. 이후 자세별 분류를 진행하고 추정 정확도를 산출한다.

4. 결론

GAN을 통해 2차원 이미지에 z축 노이즈 벡터와 피사체에 투영된 빛의 방향 값을 추가함으로써 3차원으로 증강한 뒤 자세를 추정하는 방법을 제안하였다. 이를 통해 2차원 이미지를 통한 단순 자세 추정 방법보다 정확도 향상을 기대할 수 있으며, 향후 이를 기반한 이상행동 탐지 분야에 적용할 수 있을 것이다.

ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부와 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2019-0-01834)

참고문헌

[1] Wang, J, Tan, S, Zhen, X, Xu, S, Zheng, F, He, Z, Shao, L, "Deep 3D human pose estimation: A review", Computer Vision and Image Understanding, 210(103225), p. 1-21, 2021.

- [2] Ji, X, Fang, Q, Dong, J, Shuai, Q, Jiang, W, Zhou, X, "A survey on monocular 3D human pose estimation", *Virtual Reality & Intelligent Hardware*, 2(6), p. 471-500, 2020.
- [3] Nguyen-Phuoc, T, Li, C, Theis, L, Richardt, C, Yang, Y. L, "Hologan: Unsupervised learning of 3d representations from natural images", In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, p. 7588-7597, 2019.
- [4] Zheng, C, Wu, W, Chen, C, Yang, T, Zhu, S, Shen, J, Kenhtarnavaz, N, Shah, M, "Deep learning-based human pose estimation: A survey", *arXiv preprint arXiv:2012.13392*, p. 1-35, 2020.
- [5] De Chaumont, F, Ey, E, Torquet, N, Lagache, T, Dallongeville, S, Imbert, A, Legou, T, Sourd, L, Anne-Marie, Faure, P, Bourgeron, T, Olivo-Marín, J. C, "Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning", *Nature biomedical engineering*, 3(11), p. 930-942, 2019.
- [6] Graving, J. M, Chae, D, Naik, H, Li, L, Koger, B, Costelloe, B. R, Couzin, I. D, "DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning", *Elife*, 8(47994), p. 1-42, 2019.
- [7] Liu, Z, Chen, H, Feng, R, Wu, S, Ji, S, Yang, B, Wang, X, "Deep dual consecutive network for human pose estimation", In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, p. 525-534, 2021.
- [8] Nath, T, Mathis, A, Chen, A. C, Patel, A, Bethge, M, Mathis, M. W, "Using DeepLabCut for 3D markerless pose estimation across species and behaviors", *Nature protocols*, 14(7), p. 2152-2176, 2019.
- [9] Chan, E. R, Monteiro, M, Kellnhofer, P, Wu, J, Wetzstein, G, "pi-gan: Periodic implicit generative adversarial networks for 3d-aware image synthesis", In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, p. 5799-5809, 2021.