UnrealEgo: A New Dataset for Robust Egocentric 3D Human Motion Capture

Hiroyasu Akada[1,2] Jian Wang[1] Soshi Shimada[1] Masaki Takahashi[2] Christian Theobalt[1] Vladislav Golyanik[1]

[1]Max Planck Institute for Informatics, SIC [2] Keio University

ECCV TEL AVIV 2022

See our project page for more details!

Overview

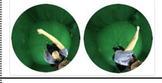
Goal

- Providing a new **large-scale** naturalistic dataset for egocentric 3D human pose estimation, *i.e.*, UnrealEgo

- Evaluating SOTA methods on a **challenging dataset** with a wide variety of motions that can be seen in daily human activities

- Proposing **a new benchmark method** that achieves **state-of-the-art results** on UnrealEgo

Datasets for Egocentric 3D Human Pose Estimation

	Monocular Setting		Stereo Setting		
	Mo2Cap2[1]	xR-EgoPose[2]	EgoCap[3]	EgoGlass[4]	UnrealEgo
Device					
Example Data					
Distance to user's face	~8cm from the head	~2cm from the nose	~25cm from the head	~1cm from the head	~1cm from the head
Number of egocentric views	530k	380k	30k × 2 views	170k × 2 views	450k × 2 views
Number of keypoints	body: 15	body: 25 hand: 40	body: 17	body: 13	body: 32 hand: 40
Image generation	composite	composite	lab environments	lab environments	3D environments
Image quality	low	realistic	real	real	realistic
Motion diversity	middle	middle	low	low	high

UnrealEgo Setup

Glasses-based Setup (Unreal Engine)

- Two fisheye cameras attached on the glasses' frames
- UnrealEgo provides RGB-D images captured at 25 fps with a resolution of 1024×1024 px, 3D keypoints, 2D keypoint heatmaps and height w.r.t. the ground
- The largest dataset for egocentric 3D human pose estimation (903,208 images)



3D Human Characters

- 9 female and 8 male models
- Different types of clothing and skin colour

3D Environment

- A wide variety of indoor and outdoor scenes with different time of day and night

e.g., laboratory, factory, office, cafeteria, park, boulevards, football field, tennis court, football field and Japanese restaurant



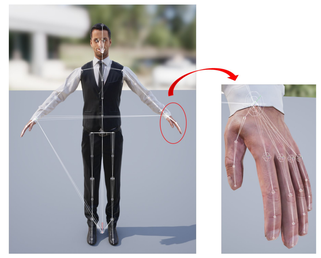

UnrealEgo Motions

Motions

- UnrealEgo contains daily human motions, ranging from simple to highly complex motions, such as walking, crouching, exercising, backflip, to name a few
- 4,841 motions in total: 3,821 for training, 494 for validation, and 526 for testing

Keypoints

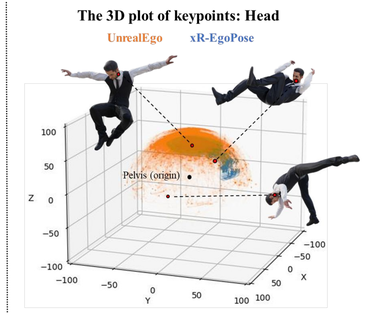
- UnrealEgo uses the default skeleton of Unreal Engine
- UnrealEgo contains 72 keypoints in total: 32 for body and 40 for hands
- We use 16 keypoints for our experiments: neck, upper arms, lower arms, hands, thighs, calves, feet and balls of the feet

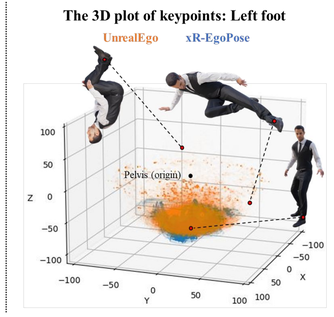


3D skeleton used in UnrealEgo

Distribution of Pelvis-relative 3D Keypoints

- The keypoints of UnrealEgo are more widespread than those of xR-EgoPose [2]

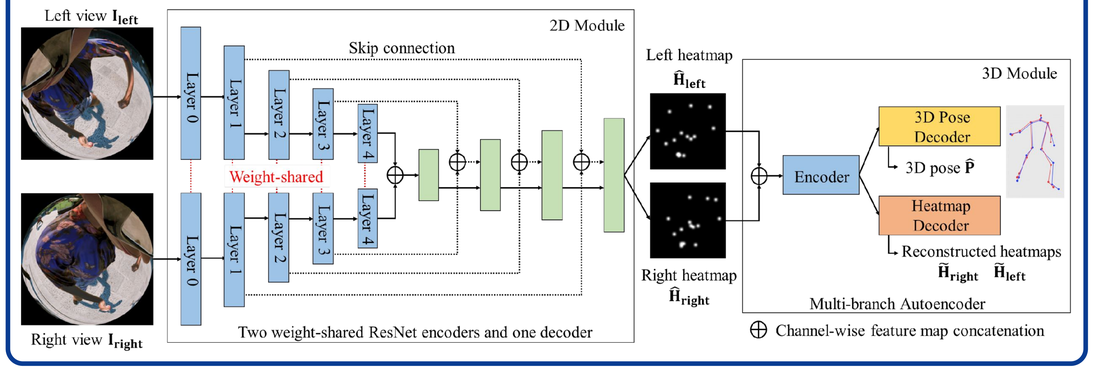




New Method for Egocentric 3D Human Pose Estimation

A Two-step approach:

- 2D module for keypoint heatmap estimation and 3D module for 3D pose estimation



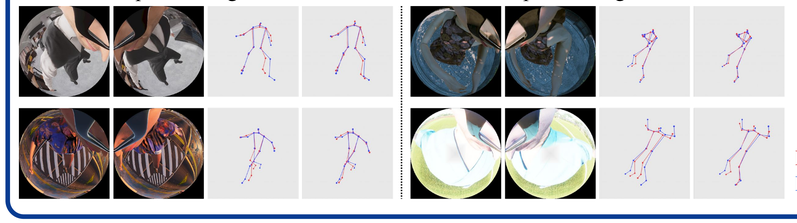
3D Pose Estimation on UnrealEgo

Quantitative Evaluation

Methods	Settings	MPJPE (std)	PA-MPJPE (std)
xR-EgoPose	Monocular	112.86 (1.16) / 123.15 (2.05)	88.71 (0.98) / 96.56 (1.27)
EgoGlass	Stereo	91.44 (0.84) / 107.70 (1.88)	70.21 (0.90) / 84.22 (0.99)
Ours	Stereo	**79.06 (0.25) / 87.31 (0.57)**	**59.95 (0.74) / 64.65 (0.93)**

- Evaluation based on train/test/validation splits of UnrealEgo
- Result with/without ImageNet pre-training
- Our method outperforms the existing method by at least 10%

Qualitative Evaluation



Red: prediction
Blue: ground truth

References

[1] Mo2Cap2; Xu, et al., TVCG 2019
[2] xR-EgoPose; Tome et al., ICCV 2019
[3] EgoCap; Rhodin et al., SIGGRAPH Asia 2016
[4] EgoGlass; Zhao et al., 3DV 2021

Project Page

Check out our project page through the QR code above or URL below for

- Dynamic visualization of UnrealEgo
- Download link to UnrealEgo
- Source code on GitHub

https://4dqv.mpi-inf.mpg.de/UnrealEgo/

Keio University

4D and Quantum Vision Group

max planck institut informatik

SIC Saarland Informatics Campus