Anonymisation in egocentric video

A brief summary

Samuel Albanie 04/2022

Outline

- Motivation

- Gait and its significance for egocentric video anonymity

- Attack: Gait recognition

- Attack: Combining egocentric footage with third person footage

- Attack: Hand gestures

- Defence: An anonymisation strategy for egocentric videos

- Summary

Motivation
Why care about anonymisation in egocentric video?

Egocentric vision is on the rise

Egocentric (first person) video footage is increasingly popular, thanks to devices like the GoPro camera series.

Use-cases: law enforcement, extreme sports, geriatric care, life-logging etc.

Videographers may think that sharing egocentric footage does not reveal their identity.

Combat operations

(alleged) robbery

Does egocentric video reveal videographer identity?

Egocentric videos may leak identity

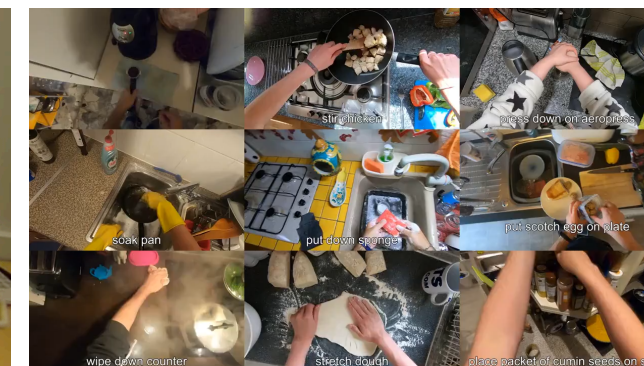**Recent works:** suggest that wearers of egocentric video cameras may have a false sense of security.

Techniques: recognition of gait, hands, correlating visual content against third person cameras

Availability of research data

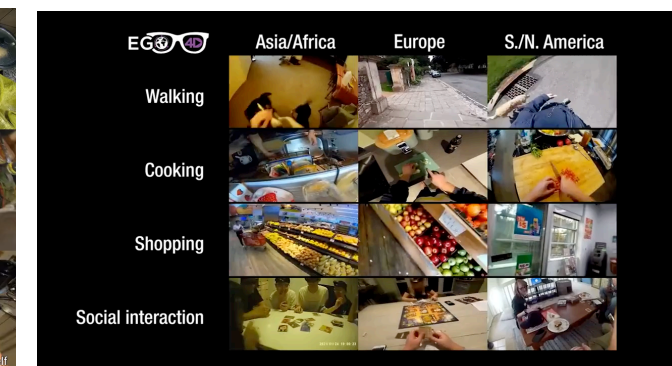**Egocentric** research datasets

GTEA (2011)
35 minutes

EPIC-KITCHENS-100 (2021)
100 hours

Ego4D (2021)
3670 hours

References/Image/Videos credits

https://commons.wikimedia.org/wiki/File:Montage_d%27une_Gopro_HD_sur_un_masque_de_chasse_sous-marine.jpg

Y. Hoshen and S. Peleg, "An egocentric look at video photographer identity." CVPR (2016)

D. Thapar, A. Nigam and C. Arora. "Anonymizing Egocentric Videos." CVPR (2021)

A. Fathi et al. "Learning to recognize objects in egocentric activities." CVPR (2011)

D. Damen et al. "Rescaling Egocentric Vision: Collection, Pipeline and Challenges for EPIC-KITCHENS-100" IJCV (2021)

K. Grauman et al. "Ego4d: Around the world in 3,000 hours of egocentric video." *arxiv* (2021)

Outline

- Motivation
- Gait and its significance for egocentric video anonymity
- Attack: Gait recognition
- Attack: Combining egocentric footage with third person footage
- Attack: Hand gestures
- Defence: An anonymisation strategy for egocentric videos
- Summary

Human gait
Introduction

What is gait and how does it affect anonymisation?

Gait definition: "manner of walking or running; bearing" (Collins Dictionary)

Significance: gait may provide cues about identity

CERES: "Great Juno, comes, I know her by her gait", The Tempest [Act 4]

Characteristics for identification

Gait may enable remote identification over a large distance

Gait may be difficult for an individual to purposefully disguise

Limits of gait identity cues

Gait may be:

• obscured by certain clothing (e.g. skirts)

• affected by footwear

• affected by walking surface

Gait is also influenced by physical condition:

• pregnancy

• fatigue

• injury

• drunkenness

Human perception of gait
Early work in psychology

Perception of human motion (1973)

Johannson (1973) studied perception of complex biological motion

Inspired by Wertheimer's (1923) illustration of the *law of common fate*

with dot patterns, 10 bright spots were used to characterise motion

Trajectories were generated by an actor wearing lightbulbs (or

reflectors) attached to their primary joints.
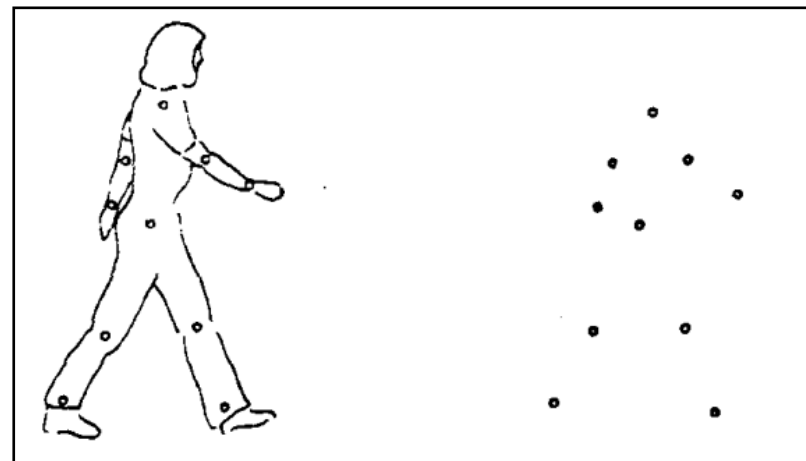
Walking
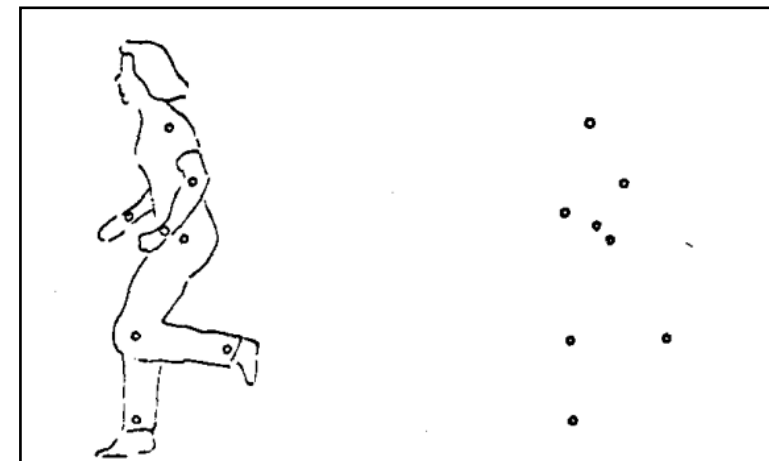
Running

Conclusion: humans can recognise activities such as walking, cycling,

dancing robustly from short clips of bright spots.

Gait as a cue for identity (1999)

Stevenage et al. (1999) conducted controlled experiments to see if humans

could identify individuals from their gait

Motivation: *identify bank robber whose face was hidden during robbery*

Experiments:

(1) Subjects learned to recognise six walkers (matched for build and

height and fully dressed in black) in various lighting conditions

Identification was successful even when point-lights were used

(2) Further experiments mimicked an "identity parade" based on gait, and

showed that participants could recognise identity far above chance.

Conclusion: humans can recognise individuals from their gait without

body shape information, given a small number of training examples.

References/Image credits: G. Johansson. "Visual perception of biological motion and a model for its analysis." *Perception & psychophysics* (1973) (this is the image source for walking/running figures)
Max Wertheimer "Untersuchungen zur Lehre von der Gestalt II." *Psychologische Forschung 4* (1923)
Stevenage, Sarah V., Mark S. Nixon, and Kate Vince. "Visual analysis of gait as a cue to identity." Applied Cognitive Psychology (1999)

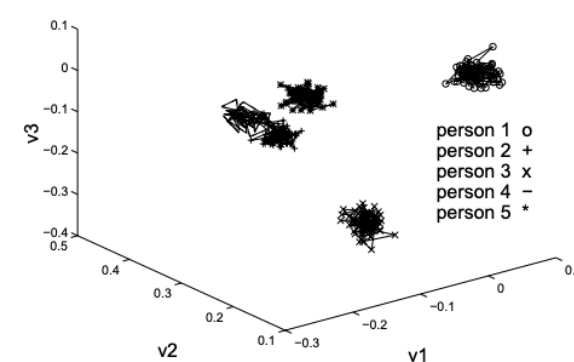# Automatic recognition of gait
Image-based and device-based methods

Low-level gait cues (1999)

Nixon et al. (1999) explore 2 methods in a small study

1. Statistical measurement (PCA + CST)

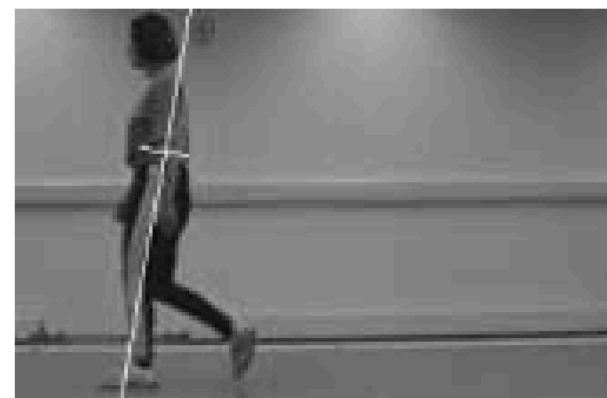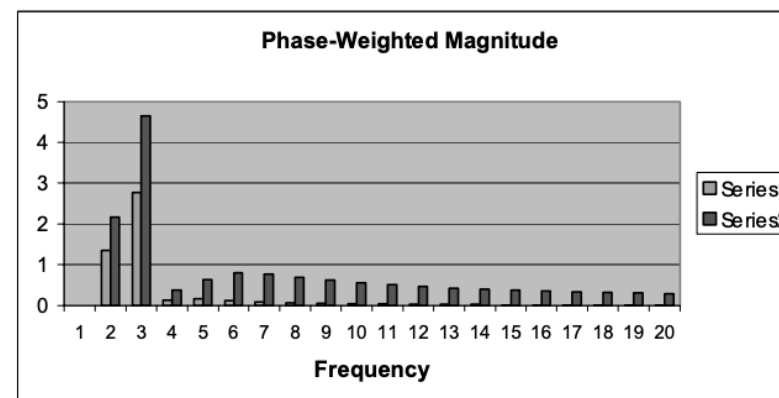
Eigengait

Canonical Space Transform

2. Feature-based measurement (Edges + HT + FT)

Edge detection

Spectral Signature

Results: promising on 10-way identity classification

Gait recognition from silhouettes (2003)

Wang et al. (2003) showed that silhouettes can be employed to identify individuals from their gait.

1. Frame-level segmentation

2. Silhouette tracking across frames

3. PCA is used on boundary features to compress sequences
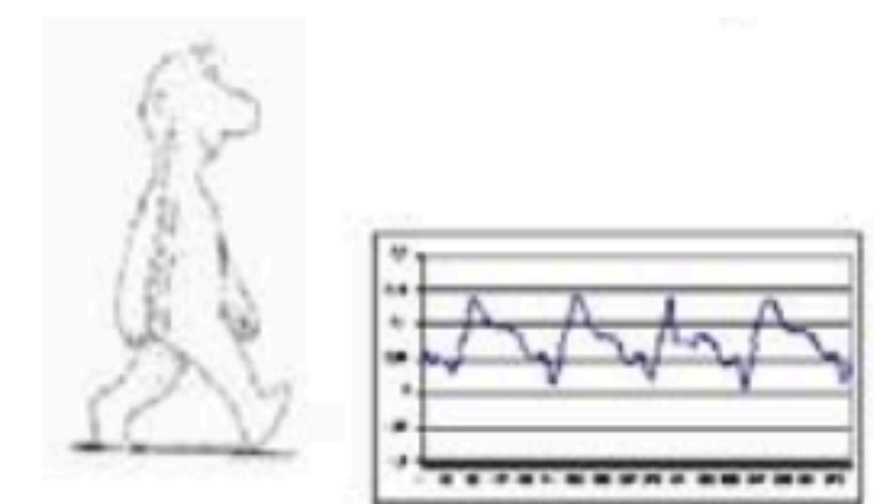
4. Classifier based on nearest neighbours

NLPR dataset

20 individuals

240 sequences

3D Accelerometers (2005)

Motivation: mobile phones are ubiquitous - gait may be useful for authentication.

Method: Identify gait using 3D accelerometers via template matching.

Experiments: Promising results across 36 individuals walking at different speeds.

References/Image credits: M. S. Nixon, J. N. Carter, D. Cunado, P. S. Huang, and S. V. Stevenage. Automatic gait recognition. Biometrics (1999)
L. Wang, T. Tan, H. Ning, & W. Hu. Silhouette analysis-based gait recognition for human identification. TPAMI (2003)
J. Mantyjarvi et al. "Identifying users of portable devices from gait pattern with accelerometers." ICASSP (2005)

Outline

- Motivation

- Gait and its significance for egocentric video anonymity

- Attack: Gait recognition

- Attack: Combining egocentric footage with third person footage

- Attack: Hand gestures

- Defence: An anonymisation strategy for egocentric videos

- Summary

An egocentric look at video photographer identity

Hoshen and Peleg, CVPR 2016

Overview

Key idea: egocentric video effectively serves as a head mounted visual gyroscope that can capture body motion. Consequently, egocentric video can reveal videographer identity in much the same way that gait recognition does.

Experiments: Recognition and verification on
- FPSI dataset of six individuals (Fathi, 2012)
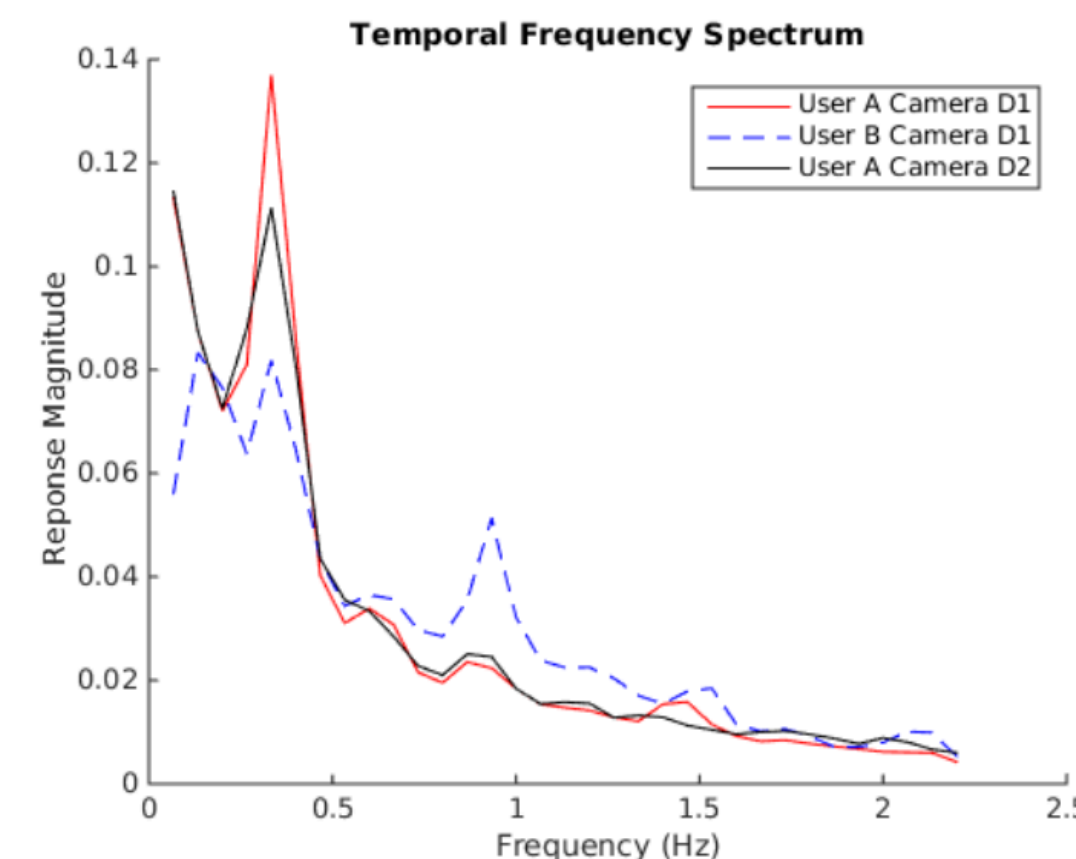- A new EVPR dataset (36 individuals) collected with head mounted GoPros
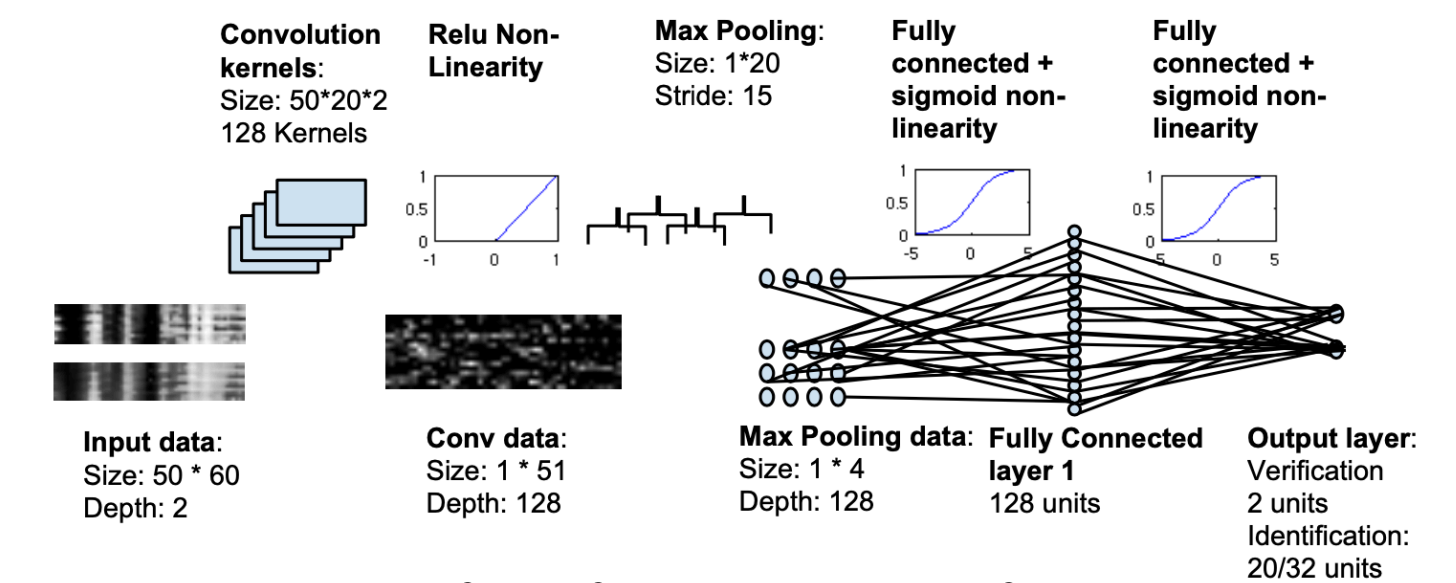

Features

Frame t   Frames s+1 to s+60
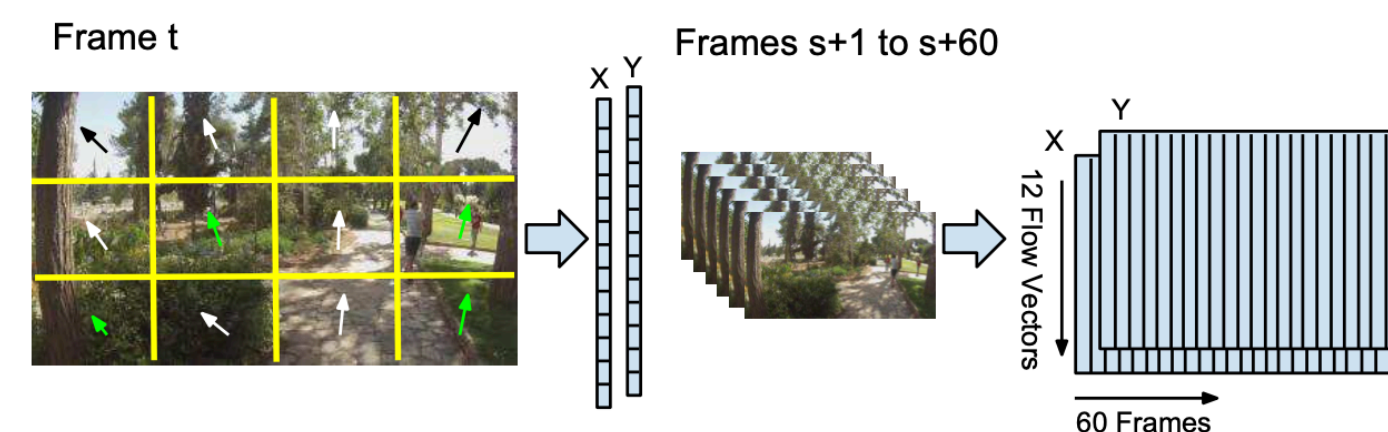
Features: sparse optical flow vectors

Enable camera-independent identity recognition

Method

Use a CNN on flow feature vectors for videographer verification and identification.

Results: 91% 20-way recognition accuracy on a test set recorded 7 days after the training samples.

Failure cases: motion blur from sharp head turns; Large moving objects causing unusual flow patterns.

References/Image credits: Y. Hoshen and S. Peleg, "An egocentric look at video photographer identity." CVPR (2016)
A. Fathi, J. Hodgins and J. Rehg. "Social interactions: A first-person perspective." CVPR (2012)

Outline

- Motivation

- Gait and its significance for egocentric video anonymity

- Attack: Gait recognition

- Attack: Combining egocentric footage with third person footage

- Attack: Hand gestures

- Defence: An anonymisation strategy for egocentric videos

- Summary

Ego2Top: Matching Viewers in Egocentric and Top-view Videos

Ardeshir and Borji, ECCV 2016

Overview

**Key idea:** surveillance cameras and drone footage are often captured from a top-down view.

These videos can be correlated with egocentric recordings made by multiple individuals to solve "who is who?"

Egocentric Videos | Top-view Video

Experiments: show promising results on 50 test scenarios (both indoor and outdoors) with different numbers of humans.

Performance improves as the number of egocentric camera recorders increase.

Method

Egocentric Videos | Graph Formation | Graph Matching | Assignment

Top-view Video
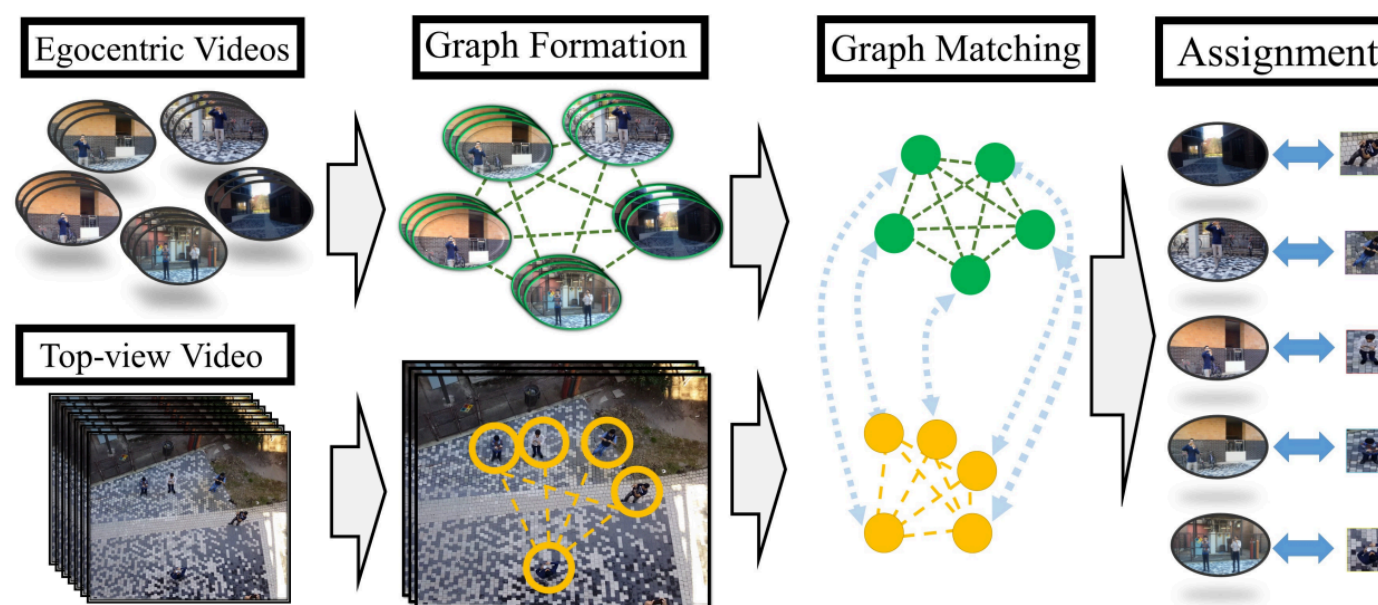
Building the top-down graph:
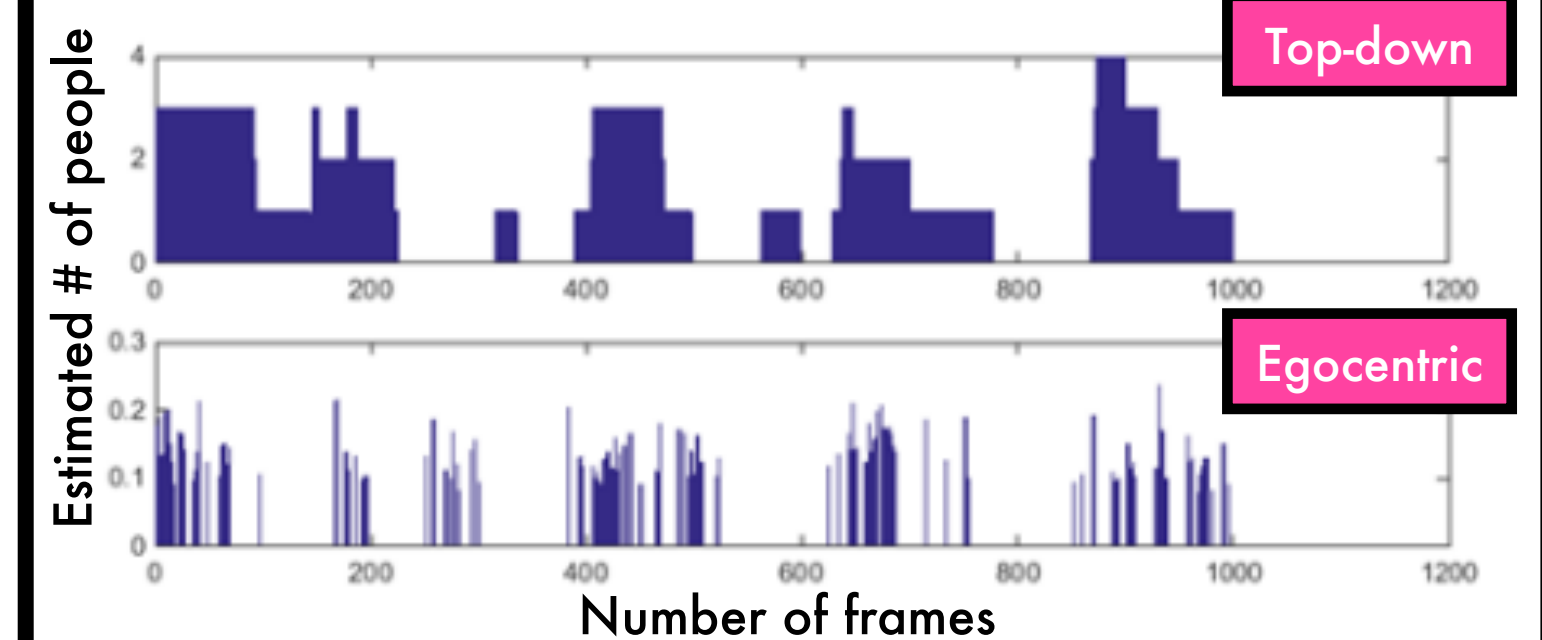
1. Use multi-object tracking in the top-down view to compute trajectories

2. Node representations include: overlaps between field-of-view and number of (visible) people

Further Details

Building the egocentric graph:

1. Estimate number of people with object detector
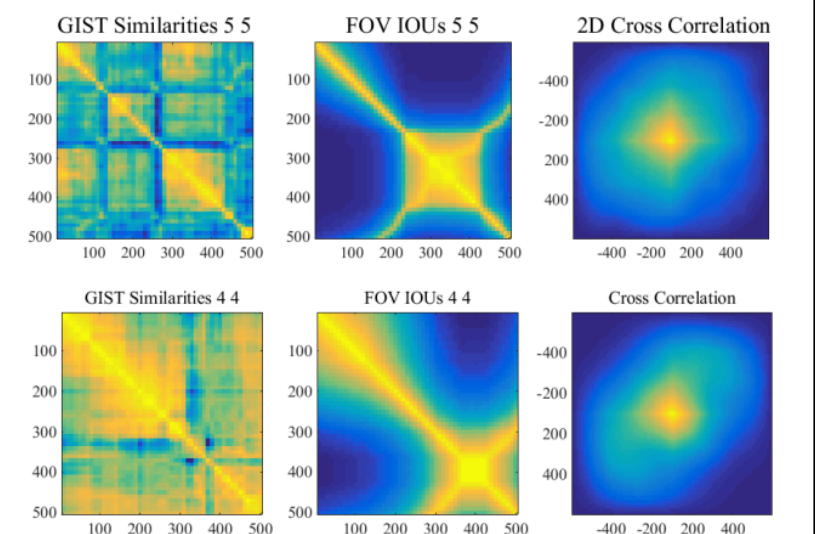

Top-down

Egocentric

Estimated # of people / Number of frames

2. Use GIST features for similarity between frames.

Given GIST similarities and FOV overlaps 2D cross correlation is used for best alignment

GIST Similarities 5 5 | FOV IOUs 5 5 | 2D Cross Correlation

GIST Similarities 4 4 | FOV IOUs 4 4 | Cross Correlation

Spectral graph matching is used to match graphs

References/Image credits: S. Ardeshir and A. Borji. "Ego2top: Matching viewers in egocentric and top-view videos." ECCV (2016)
(Multi-object tracking used in this work) C. Dicle, I. C. Octavia and M. Sznaier. "The way they move: Tracking multiple targets with similar appearance." ICCV (2013)
(GIST features used in this work) A. Torralba. "Contextual priming for object detection." IJCV (2003)
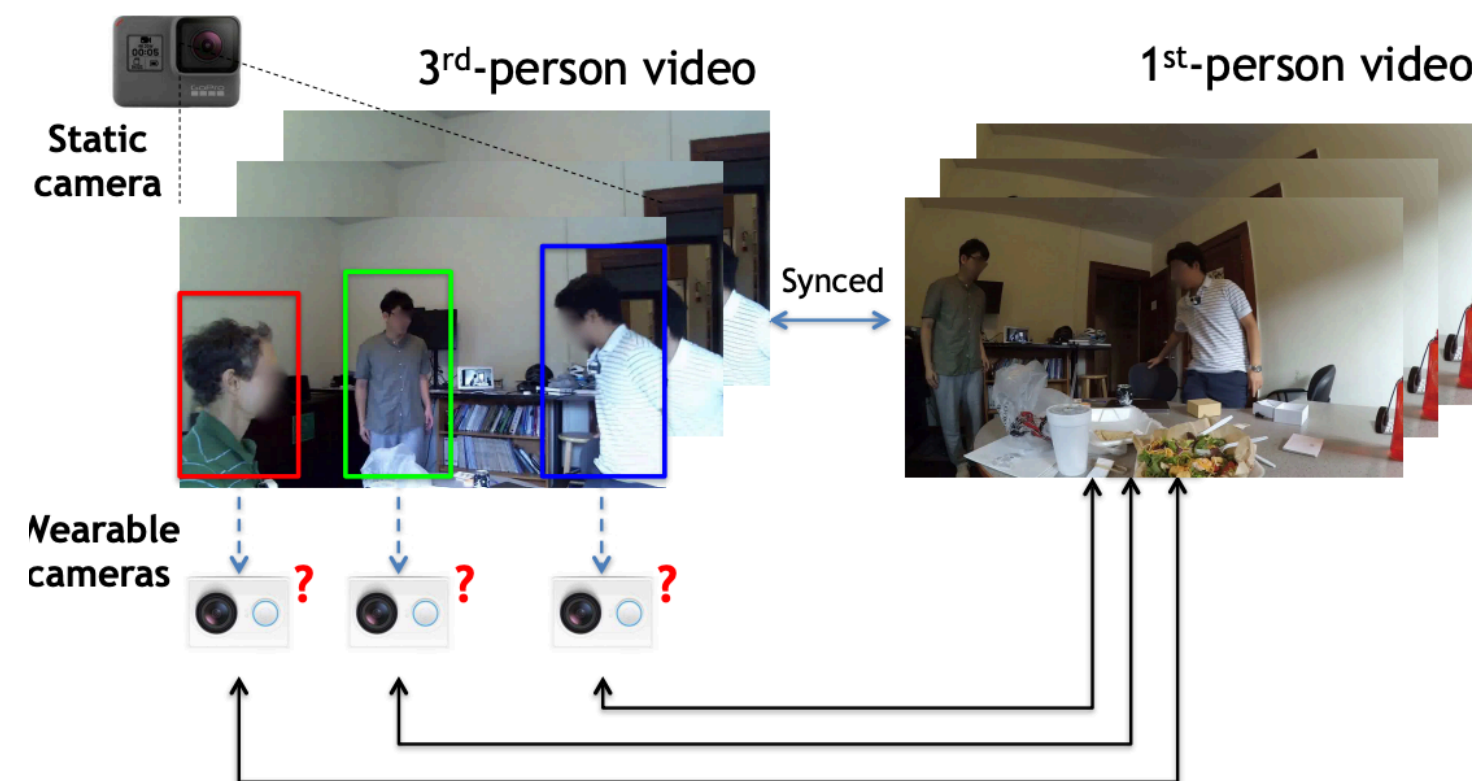
Identifying First-person Camera Wearers in Third-person Videos

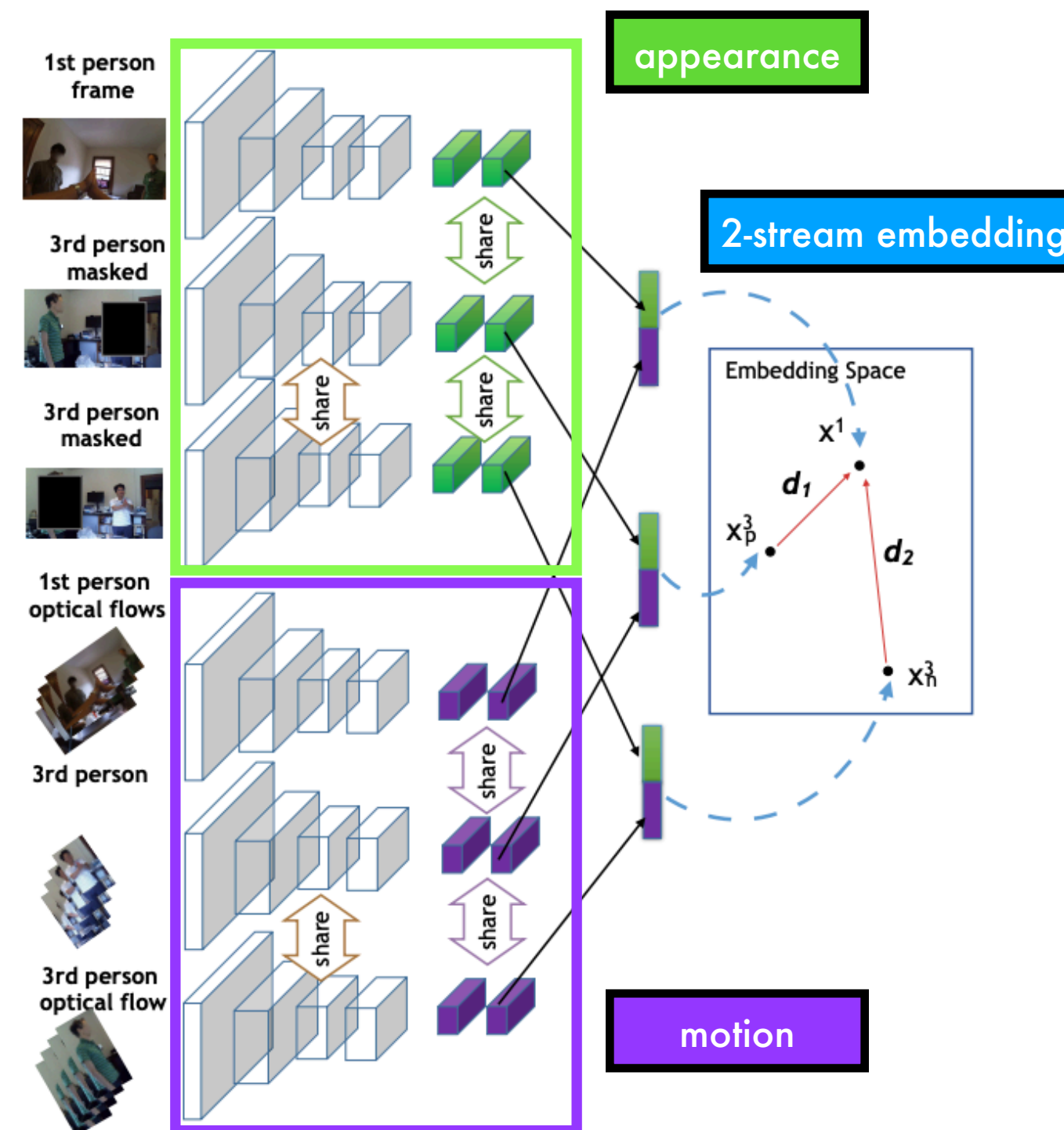## Fan, Lee, Xu, Singh, Lee, Crandall and Ryoo, CVPR 2017

Overview

Goal: find correspondences between first-person camera wearers and third-person static camera videos

Static camera
3rd-person video
Synced
1st-person video
Wearable cameras ? ? ?

Key idea: video from first-person and third-person views will share visual evidence (similar appearance/motion) We can therefore learn an embedding to capture similarities across views.

Embedding framework

Method: Train a dual encoder with a triplet loss

appearance
2-stream embedding
motion

1st person frame
3rd person masked
3rd person masked
1st person optical flows
3rd person
3rd person optical flow
share

Embedding Space
x^1
d_1
x_p^3
d_2
x_n^3

Experiments
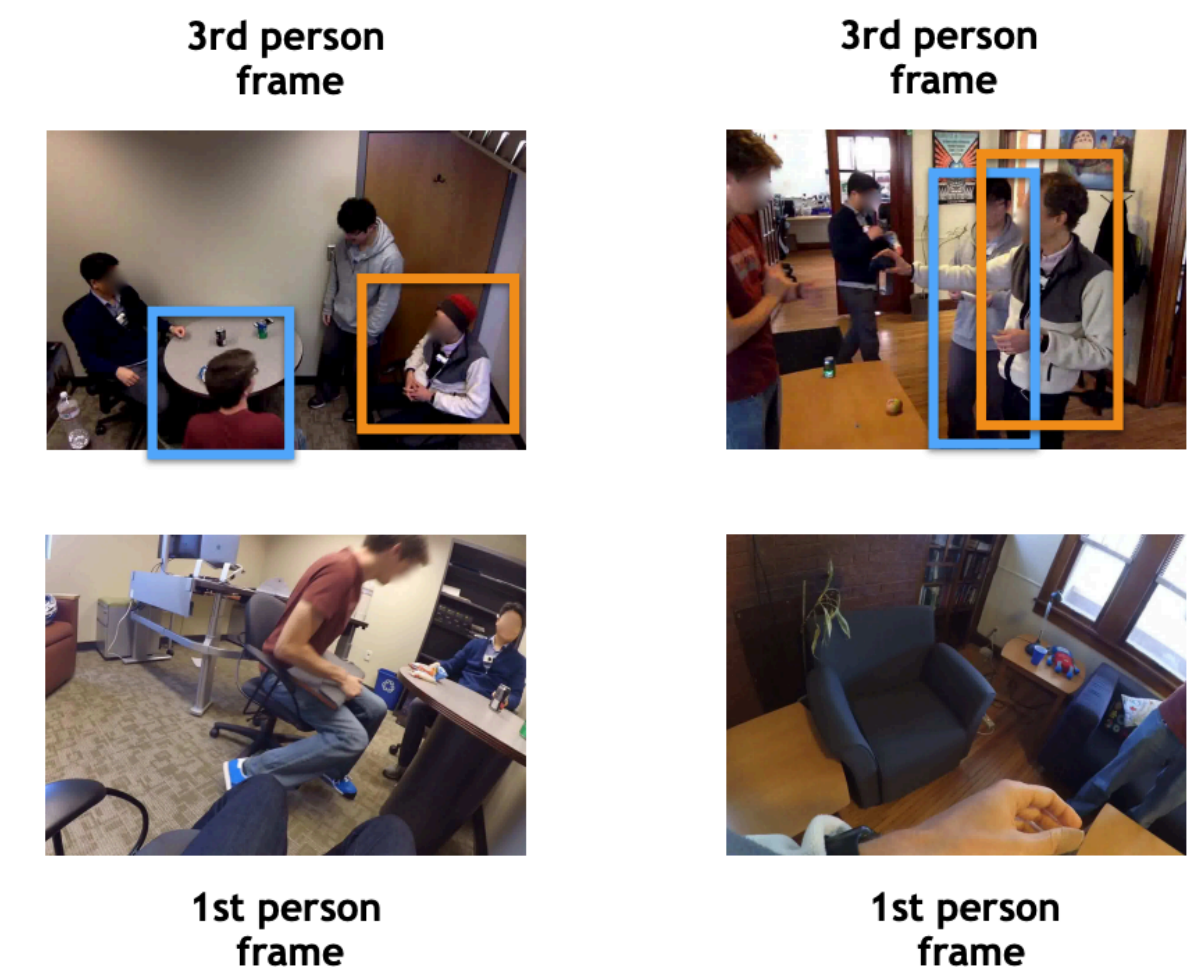
Data: activities in six indoor environments

Results: embeddings improve over simple baselines

Failure cases:

Similar egocentric motion (nodding/sitting down)

Person (spatial appearance) occlusion

3rd person frame
3rd person frame
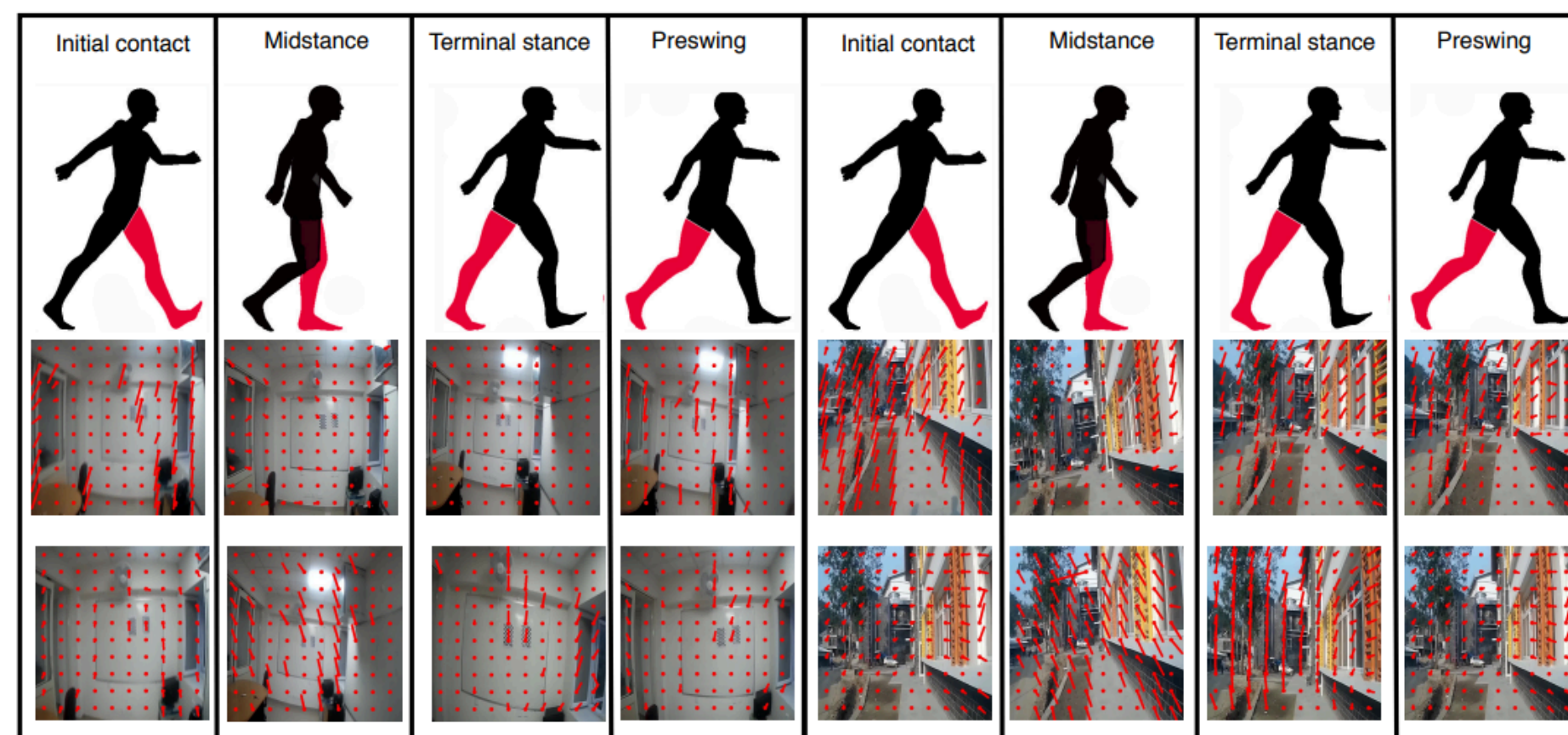1st person frame
1st person frame

Is Sharing of Egocentric Video Giving Away Your Biometric Signature?
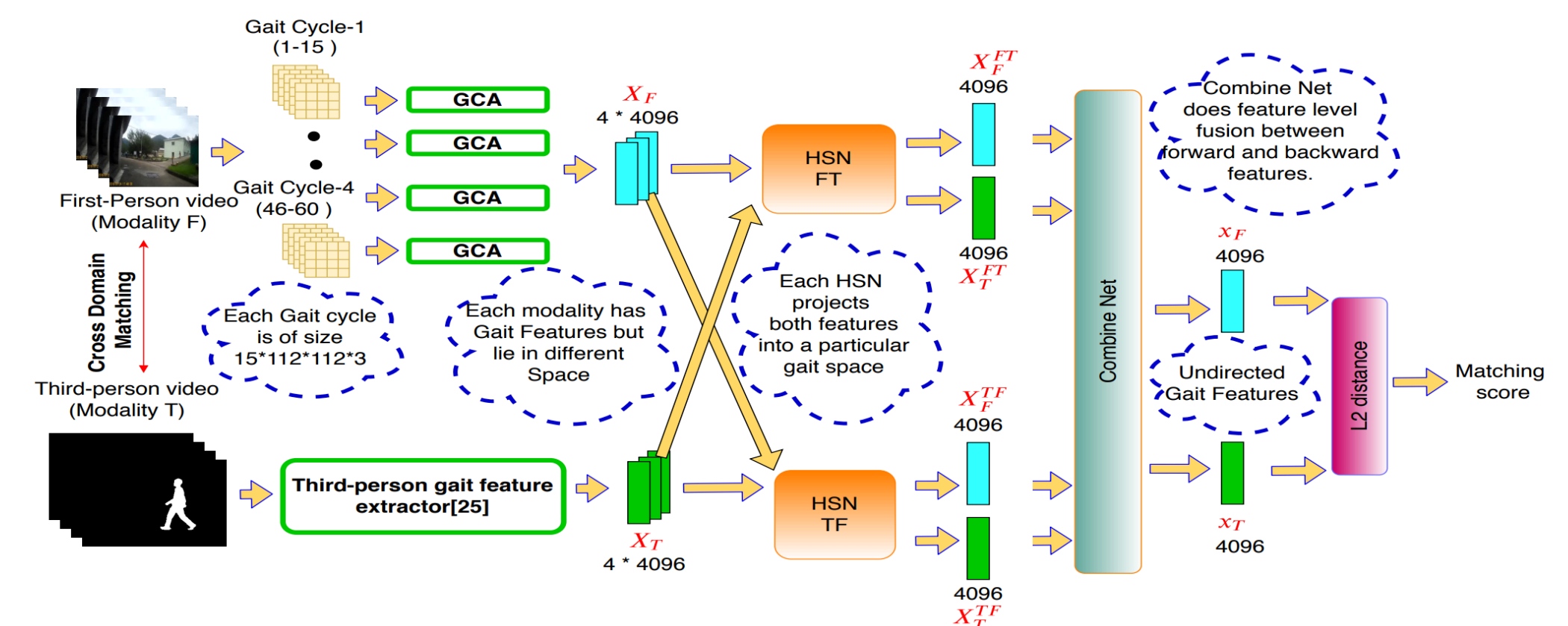
Thapar, Arora and Nigam (ECCV 2020)

Overview

Problem: Can the identity of an egocentric videographer be matched against third person footage with no overlap in content?

Key idea: Gait can be estimated both from egocentric footage and from third person footage.

| Initial contact | Midstance | Terminal stance | Preswing | Initial contact | Midstance | Terminal stance | Preswing |

Subject 1

Subject 2

Method and Experiments

Experiments: In addition to studies on FPSI and EVPR, experiments are conducted on a dataset of 12 subjects (with egocentric/3rd person videos)

Results: Evaluations under both closed-set and open-set protocols demonstrate some degree of successful identification on the new dataset.

References/Image credits: D. Thapar, C. Arora, and A. Nigam. "Is sharing of egocentric video giving away your biometric signature?" ECCV (2020)
(FPSI dataset) A. Fathi, J. Hodgins and J. Rehg. "Social interactions: A first-person perspective." CVPR (2012)
(EVPR dataset) Y. Hoshen and S. Peleg, "An egocentric look at video photographer identity." CVPR (2016)

Outline

- Motivation

- Gait and its significance for egocentric video anonymity

- Attack: Gait recognition

- Attack: Combining egocentric footage with third person footage

- Attack: Hand gestures

- Defence: An anonymisation strategy for egocentric videos

- Summary

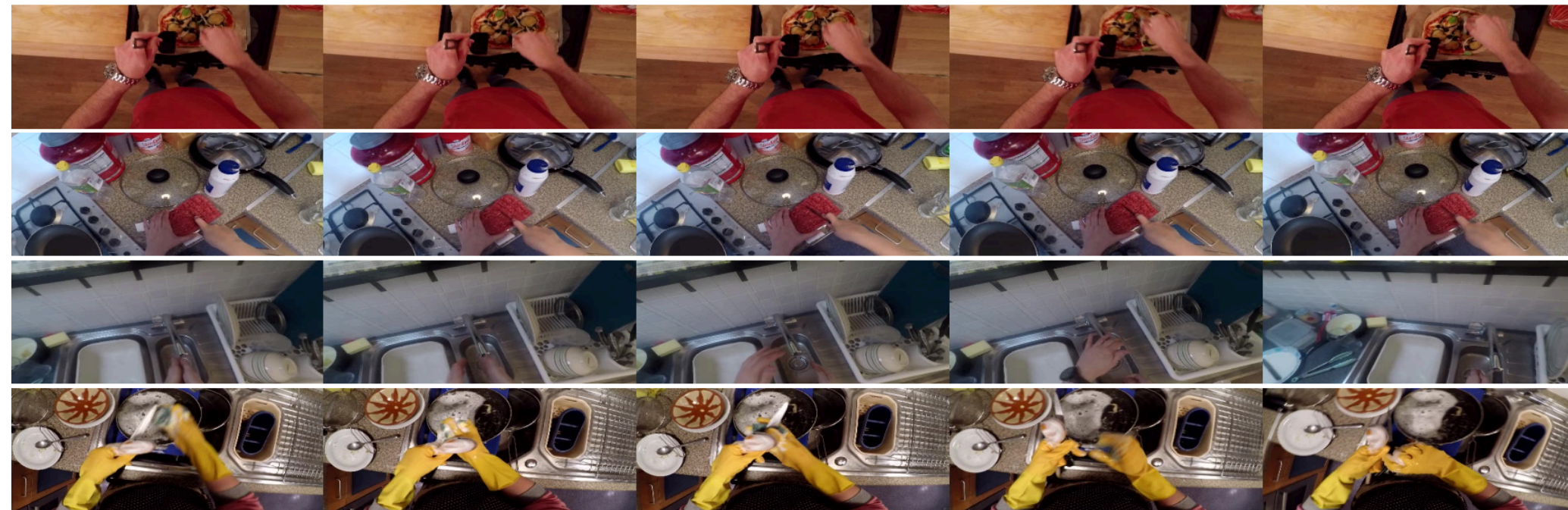# Recognizing Camera Wearer from Hand Gestures in Egocentric Videos
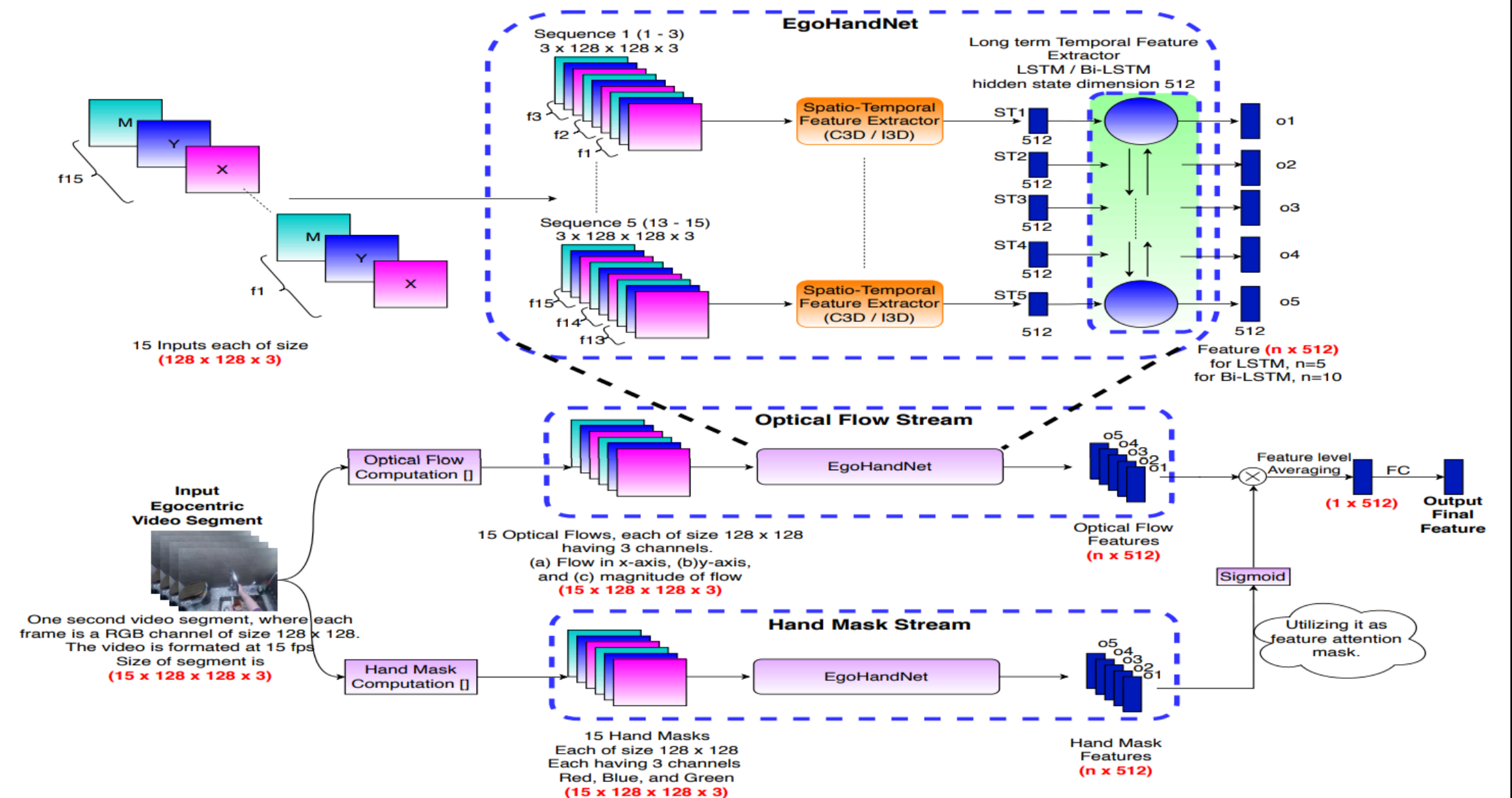
Thapar, Nigam and Arora (ACM MM 2020)

Reference/Image credits: D. Thapar, A. Nigam and C. Arora. "Recognizing camera wearer from hand gestures in egocentric videos." ACM MM (2020)
(EPIC kitchens dataset) D. Damen et al. "Scaling egocentric vision: The epic-kitchens dataset." ECCV (2018)
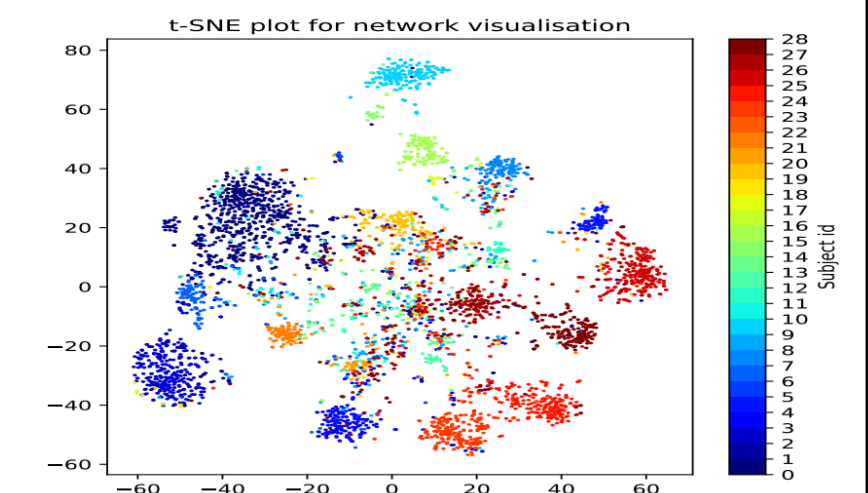
Outline

- Motivation

- Gait and its significance for egocentric video anonymity

- Attack: Gait recognition

- Attack: Combining egocentric footage with third person footage

- Attack: Hand gestures

- Defence: An anonymisation strategy for egocentric videos

- Summary

Anonymising Egocentric Videos
Thapar, Nigam and Arora (CVPR 2021)

Overview

Problem: Can we anonymise egocentric videos against identity attacks?

Key idea: apply perturbations to the video to fool gait/gesture-based techniques for estimating identity.

Goal: perturbations that are (i) effective at protecting identity; (ii) visually imperceptible; (iii) preserve task performance (activity/object recognition)
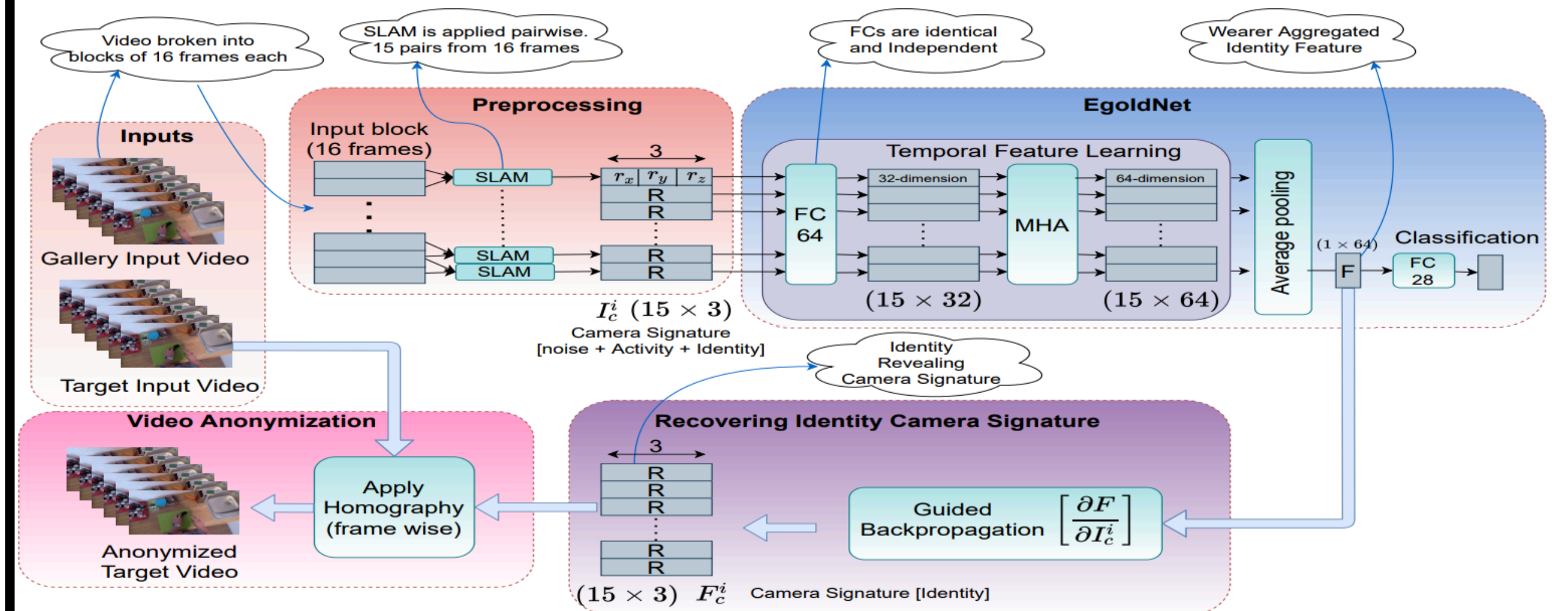

Egocentric Video

Identity removal by adding noise to camera motion
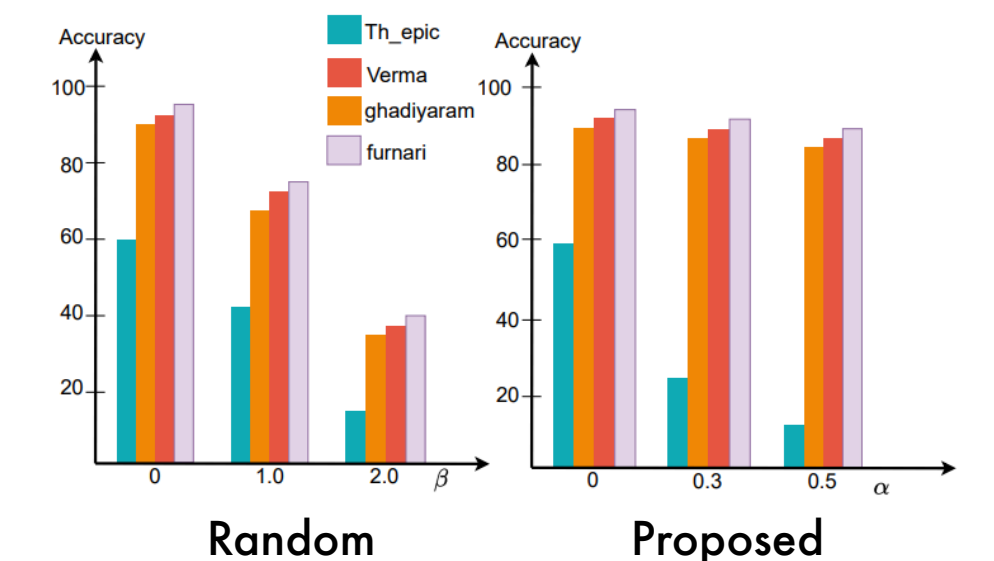
Anonymized egocentric video (Our Proposal)

Method and Experiments

Experiments: EgoIdNet outperforms attacks based on optical flow, particularly in open-set conditions

Perturbations (mostly) preserve task accuracy

Results shown across FPSI, EPIC-Kitchens & IITMD-WFP

Reference/Image credits: D. Thapar, A. Nigam and C. Arora. "Anonymizing Egocentric Videos." CVPR (2021)
(FPSI dataset) A. Fathi, J. Hodgins and J. Rehg. "Social interactions: A first-person perspective." CVPR (2012)
(EVPR dataset) Y. Hoshen and S. Peleg, "An egocentric look at video photographer identity." CVPR (2016)
(IITMD-WFP dataset) D. Thapar, C. Arora, and A. Nigam. "Is sharing of egocentric video giving away your biometric signature?" ECCV (2020)

Outline

- Motivation

- Gait and its significance for egocentric video anonymity

- Attack: Gait recognition

- Attack: Combining egocentric footage with third person footage

- Attack: Hand gestures

- Defence: An anonymisation strategy for egocentric videos

- Summary

Summary

Overview

Egocentric video footage is on the rise

This growth has benefited from better hardware and useful applications

Caution is required: several identity attacks have been demonstrated

Attack cues include:

• Videographer gait

• Hand gestures

• Third person footage

Potential gait anonymisation: selective perturbations to the footage

Limitations and Caveats

The attacks have been studied with relatively small numbers of subjects

Methods mostly make use of deep neural networks that are challenging to interpret: they may be exploiting biases (though Thapar et al. (2020) make an effort to interpret their flow features).

These works are valuable proof of concepts (larger-scale confirmatory studies are a natural next step) indicating the topic deserves further study.

This review: represents an incomplete summary to give a brief introduction to the topic, rather than a comprehensive review.

Reference: D. Thapar, C. Arora, and A. Nigam. "Is sharing of egocentric video giving away your biometric signature?" ECCV (2020)