

Towards a Practical PTZ Face Detection and Tracking System



Yinghao Cai, Thang Dinh and Gérard Medioni
 Integrated Media Systems Center
 University of Southern California

iCampus/MediaQ

iWatch ✓

CT

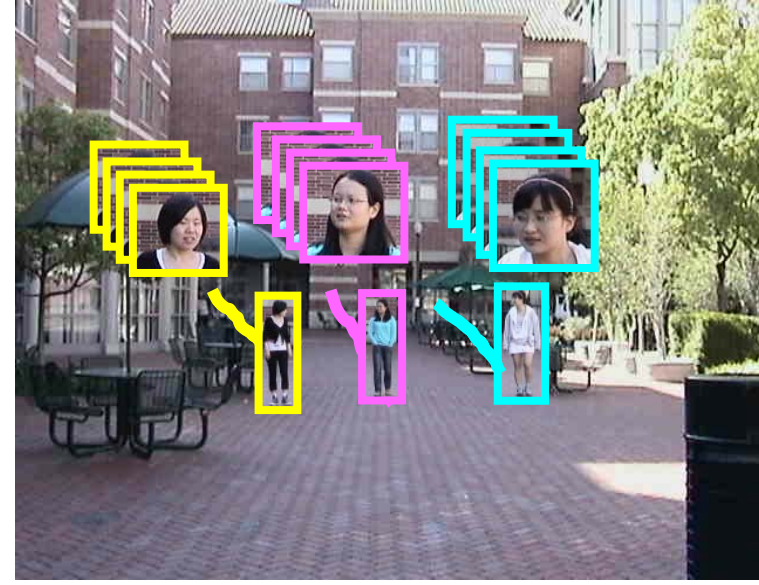
Introduction

❖ Goal

To acquire high resolution face images as well as trajectories from multiple people using a single PTZ camera.

❖ Challenges

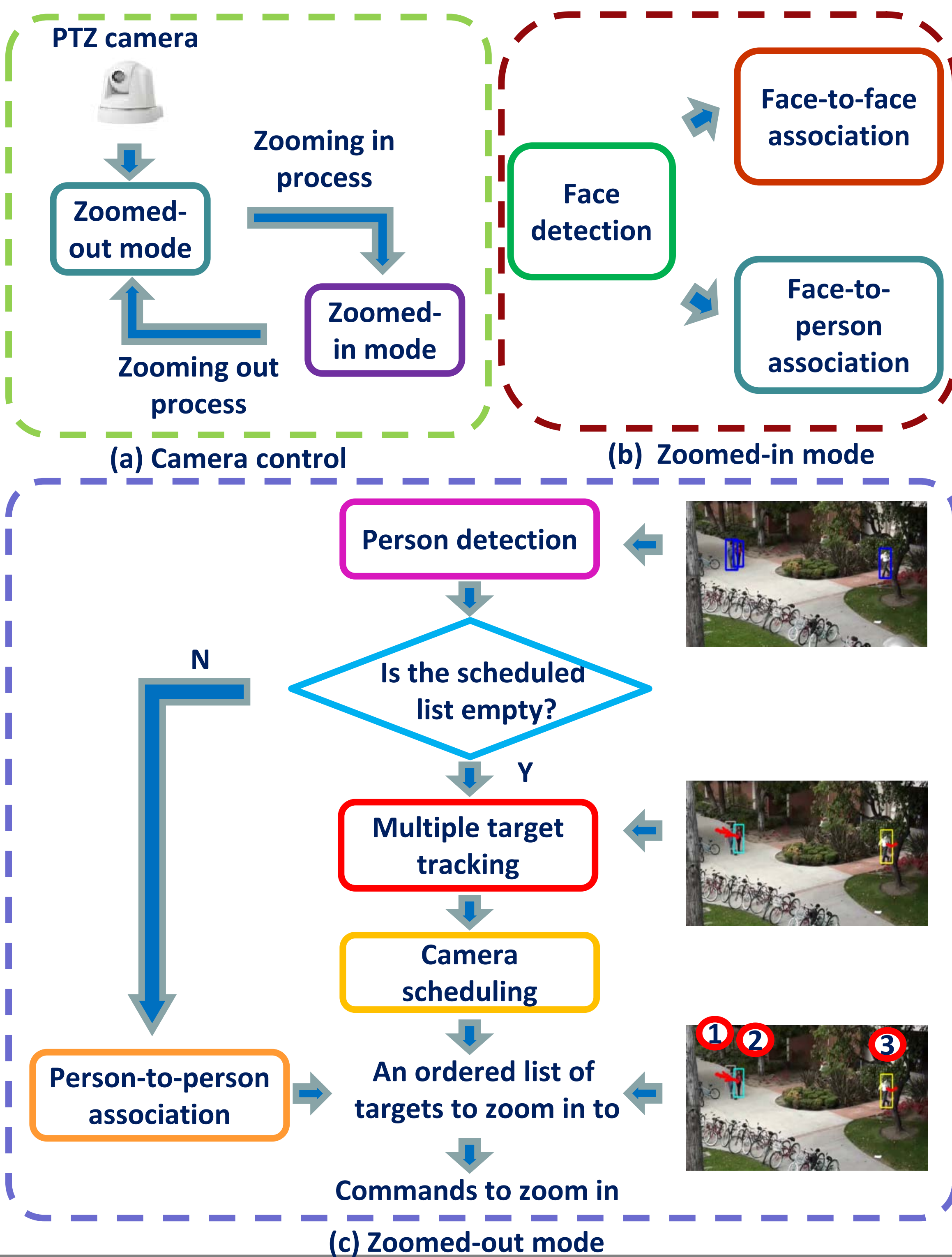
- ✓ Detection and tracking in real time.
- ✓ Efficient camera scheduling.
- ✓ Multiple people with similar appearances.
- ✓ Slow mechanical response of the PTZ camera.



❖ Highlights

Explore the limits to what can be expected from a **single** PTZ camera.

Framework



Main Approach

Zoomed-out mode

❖ Person detection

- ✓ Joint ranking of granules features (JRoG) [1].
- ✓ A simulated annealing (SA) step and Real Adaboost for feature selection.

❖ Multiple target tracking

- Multiple instances of "context tracker" [2].
- ✓ Exploring "Supporters" and "Distracters" for tracking.

❖ Camera scheduling

- A weighted Round Robin approach.
- ✓ People moving towards the camera.
- ✓ People about to exit the field of view.

❖ Person-to-person association

- To re-acquire a person in the list after zooming out.
- ✓ Color and location affinity + Hungarian algorithm.



Fig. Multiple target tracking results

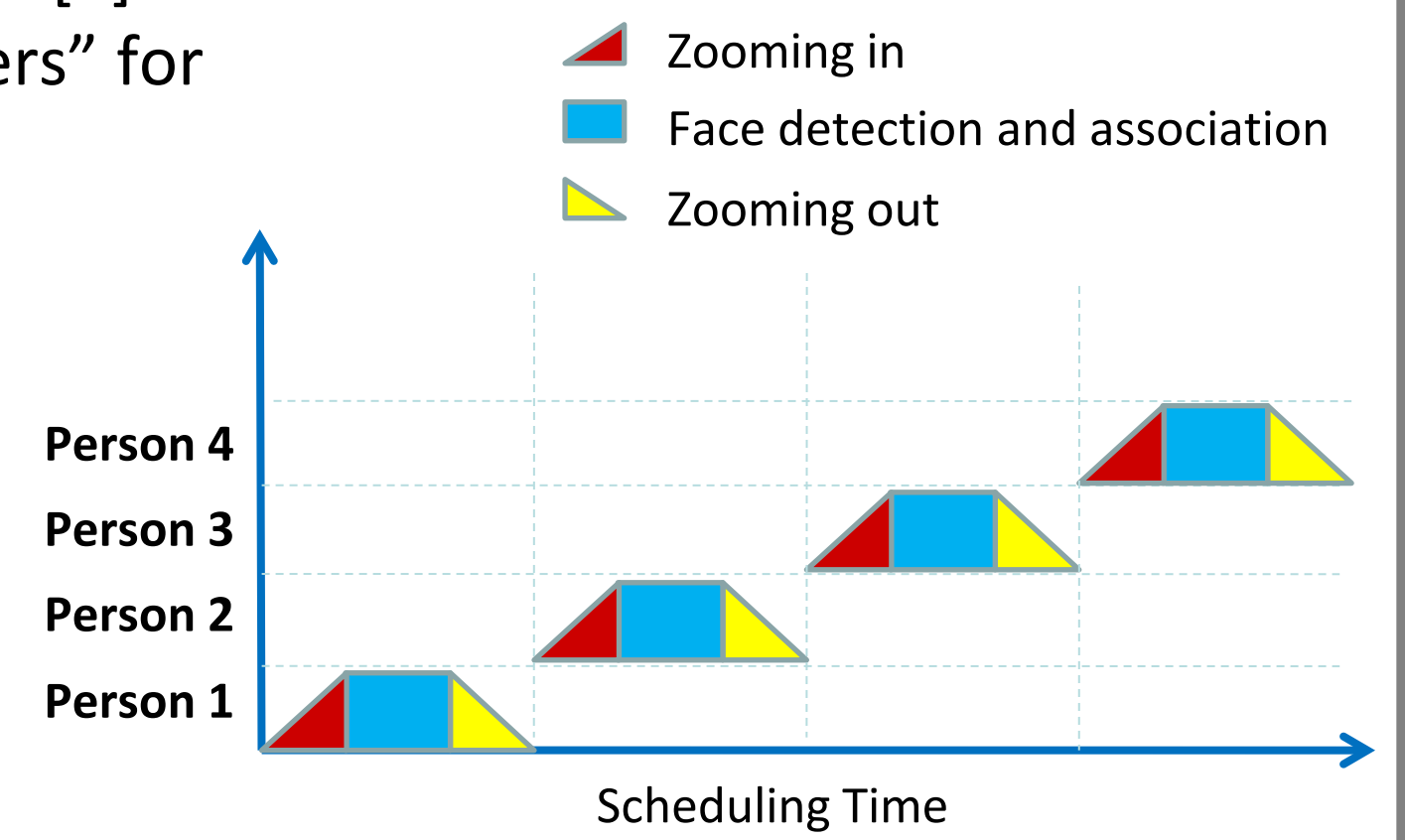


Fig. Camera scheduling

Zoomed-in mode

❖ Face detection

- ✓ Illumination invariant Local Structure Feature [3].
- ✓ 20 fps on 640*480 images.

❖ Face-to-face association

- ✓ Color histogram + LBP histogram.
- ✓ Hungarian algorithm.

❖ Face-to-person association

- ✓ Color histogram + Hungarian algorithm.

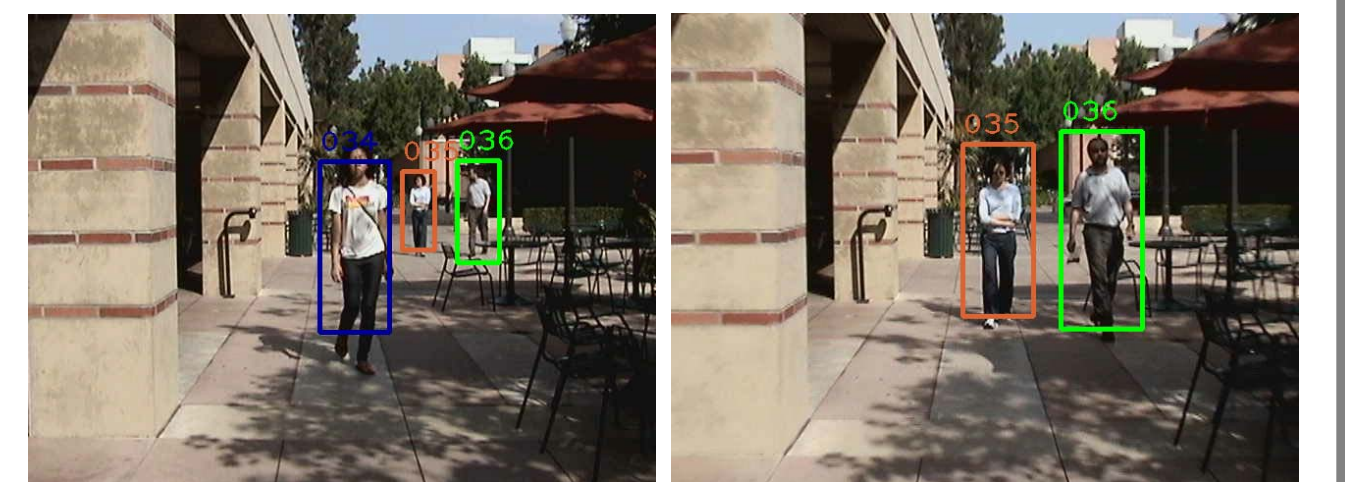


Fig. person-to-person association results

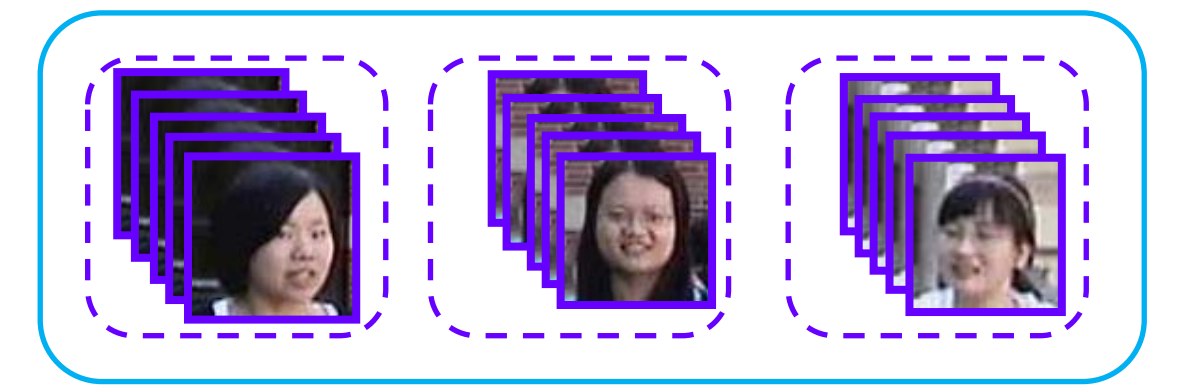


Fig. face-to-face association results

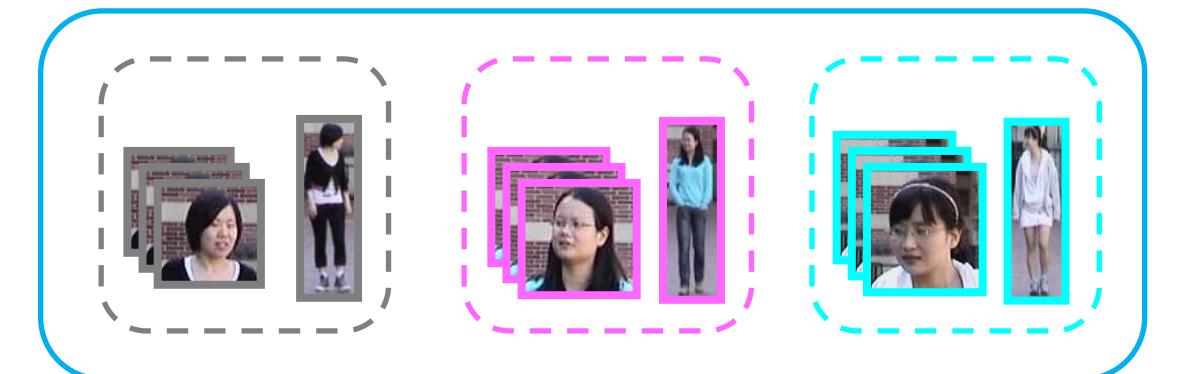


Fig. face-to-person association results

Zooming in/out process

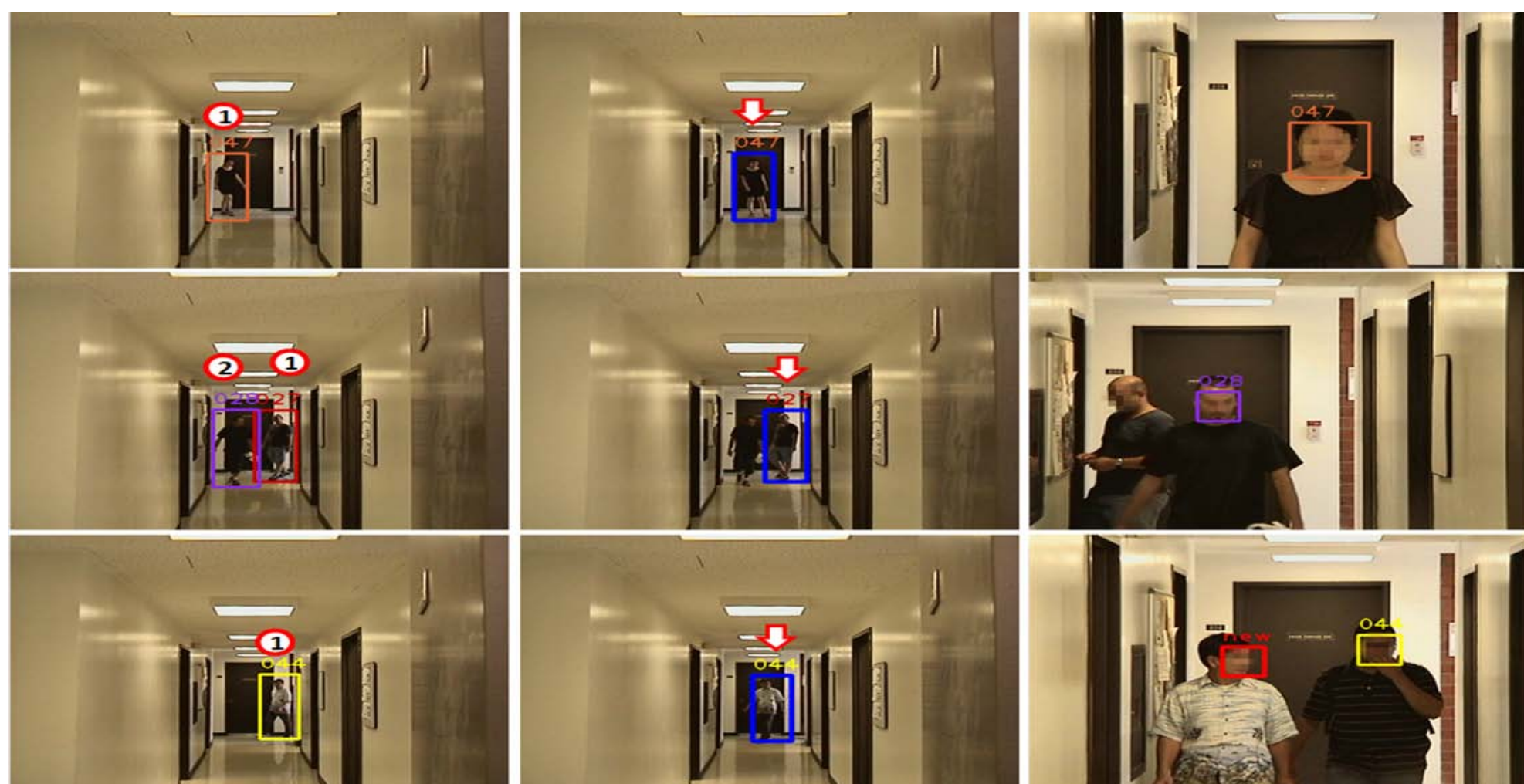
- ✓ Slow mechanical response of the PTZ camera.
- ✓ Compensate for the motion during the delay.

References:

1. C. Huang and R. Nevatia. High performance object detection by collaborative learning of joint ranking of granules features. In CVPR, 2010.
2. T. B. Dinh, N. Vo, and G. G. Medioni. Context tracker: Exploring supporters and distracters in unconstrained environments. In CVPR, 2011.
3. B. Froba and A. Ernst. Face detection with the modified census transform. In FG, 2004.

Experimental Results

Five hours of continuous operation, 19 faces and trajectories from 21 people.



(a) Person detection and tracking (b) Select a person to Zoom in (c) Face detection and association

One hour of continuous operation, 19 faces and trajectories from 25 people.



(a) Person detection and tracking (b) Select a person to Zoom in (c) Face detection and association