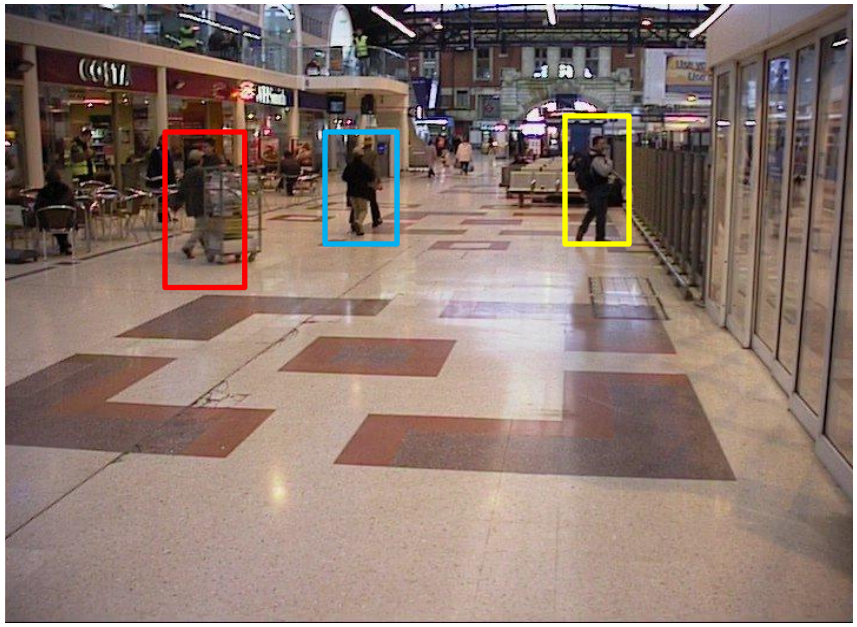
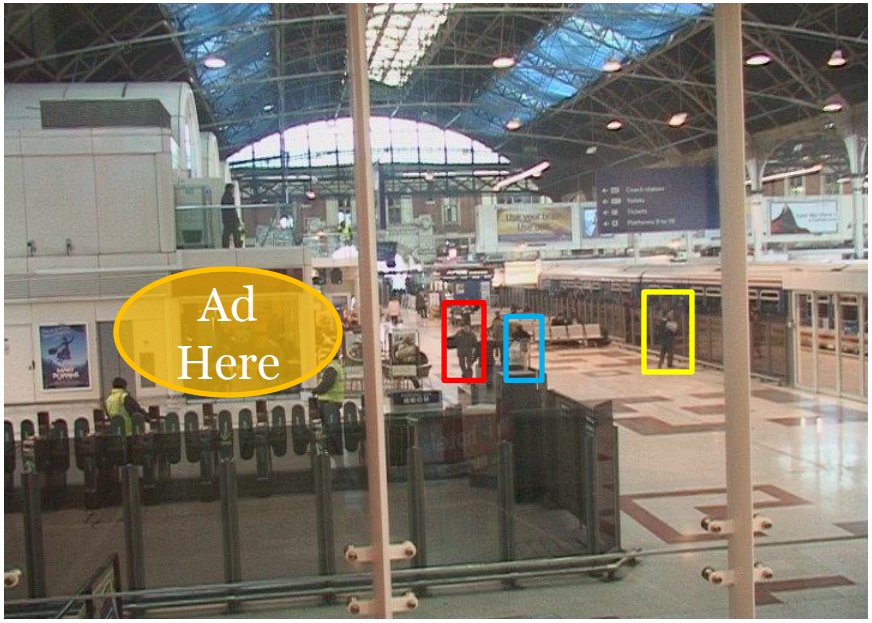
# Gaze Estimation in Camera Networks

Peter Malmgren









Source: [Pets 2007 Surveillance Dataset]

# Related Work: Parts based models

Their data:

Our data:



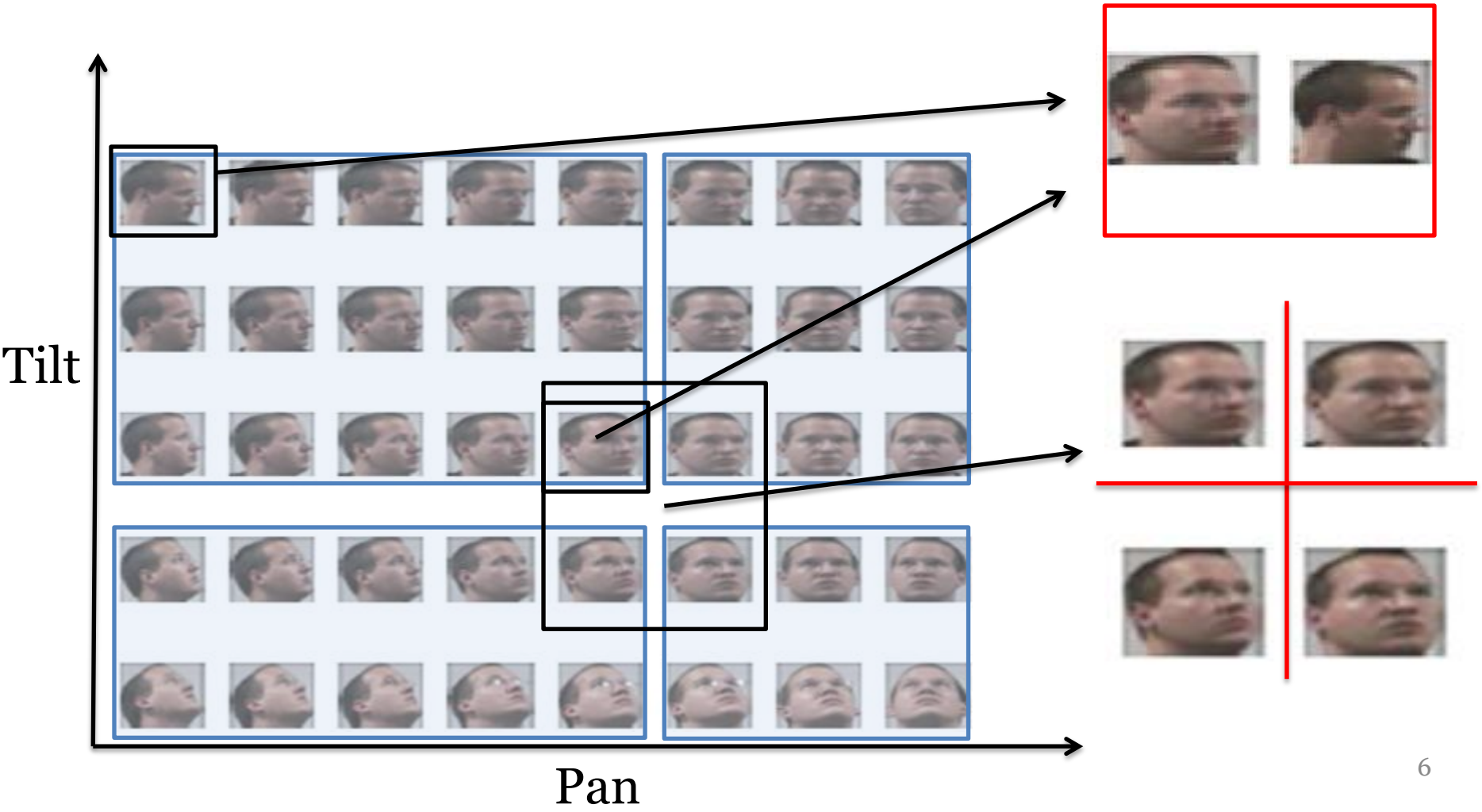
Source: Face Detection, Pose Estimation and Landmark Localization in the Wild

# Related Work: Detection models


Source: Pointing 2004 ICPR Workshop

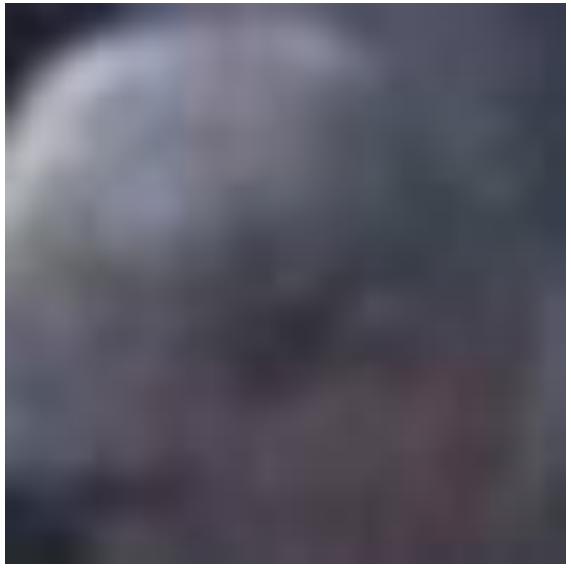
# Problem 1: Arbitrary Boundaries

These are now in the same category!



# Problem 2: Similar appearance

Looking left

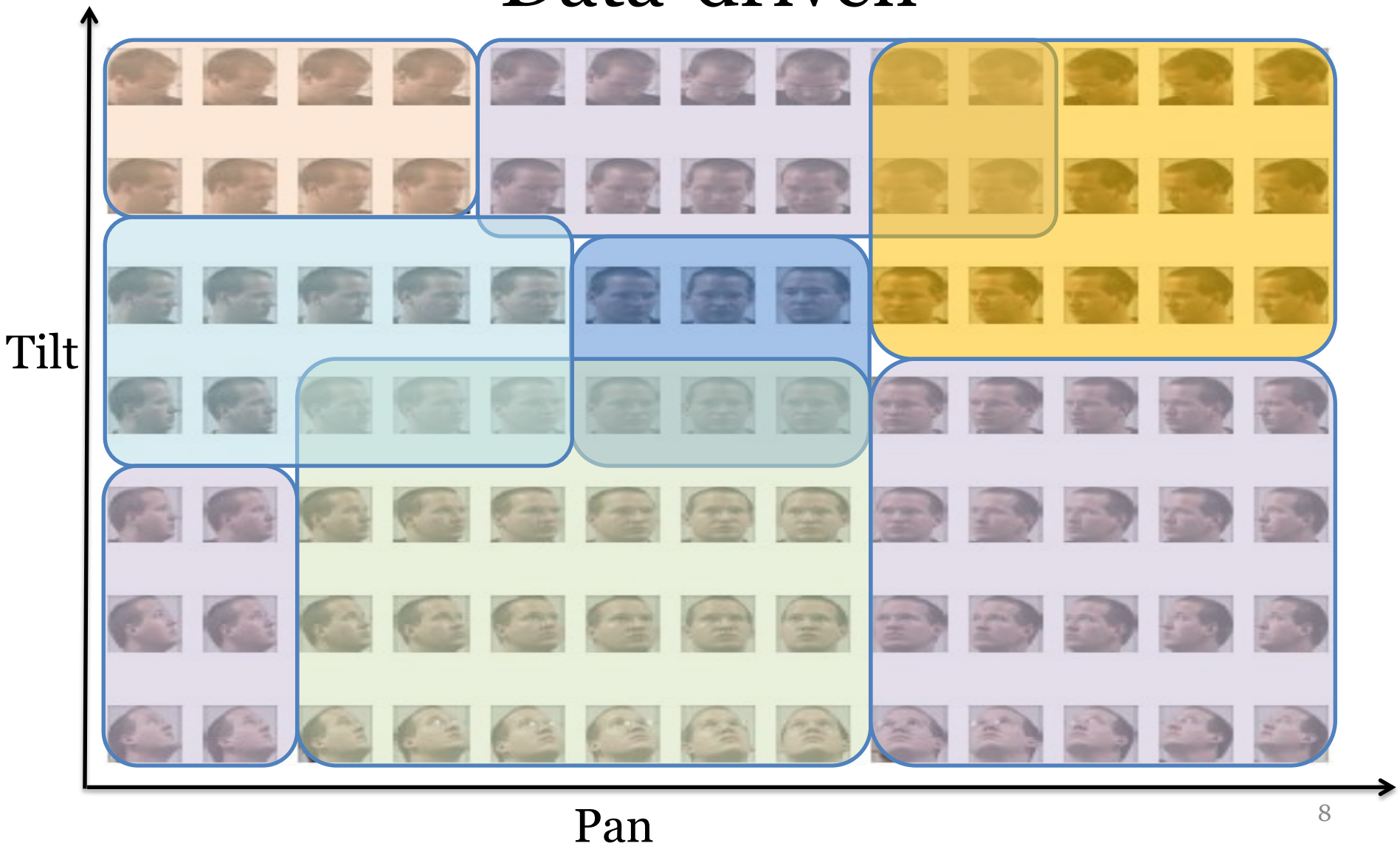


Looking right



Source: IIT (Italian Institute of Technology) Head Orientation Dataset

# Our approach: Data-driven

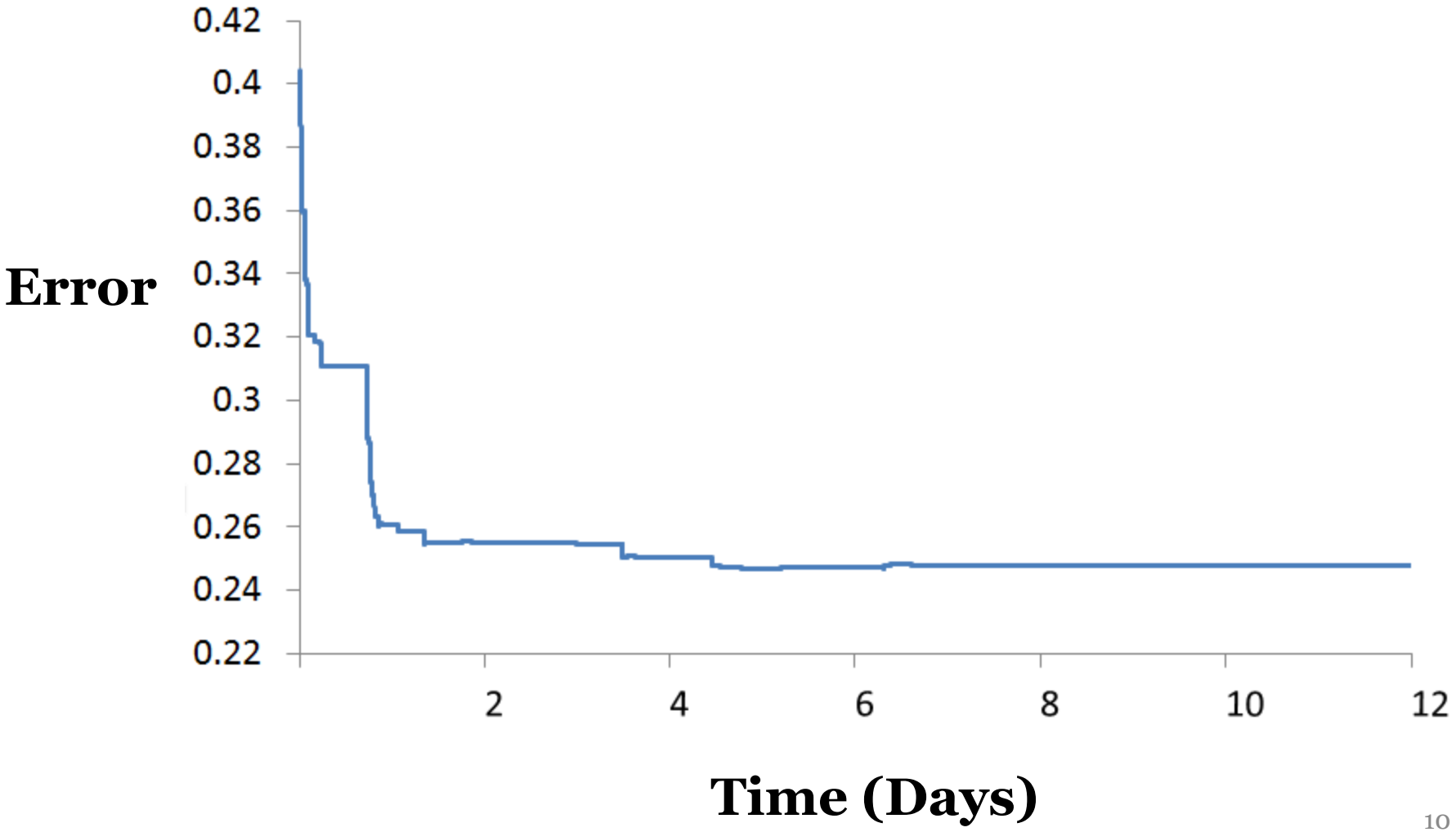


# How do we learn the groups?

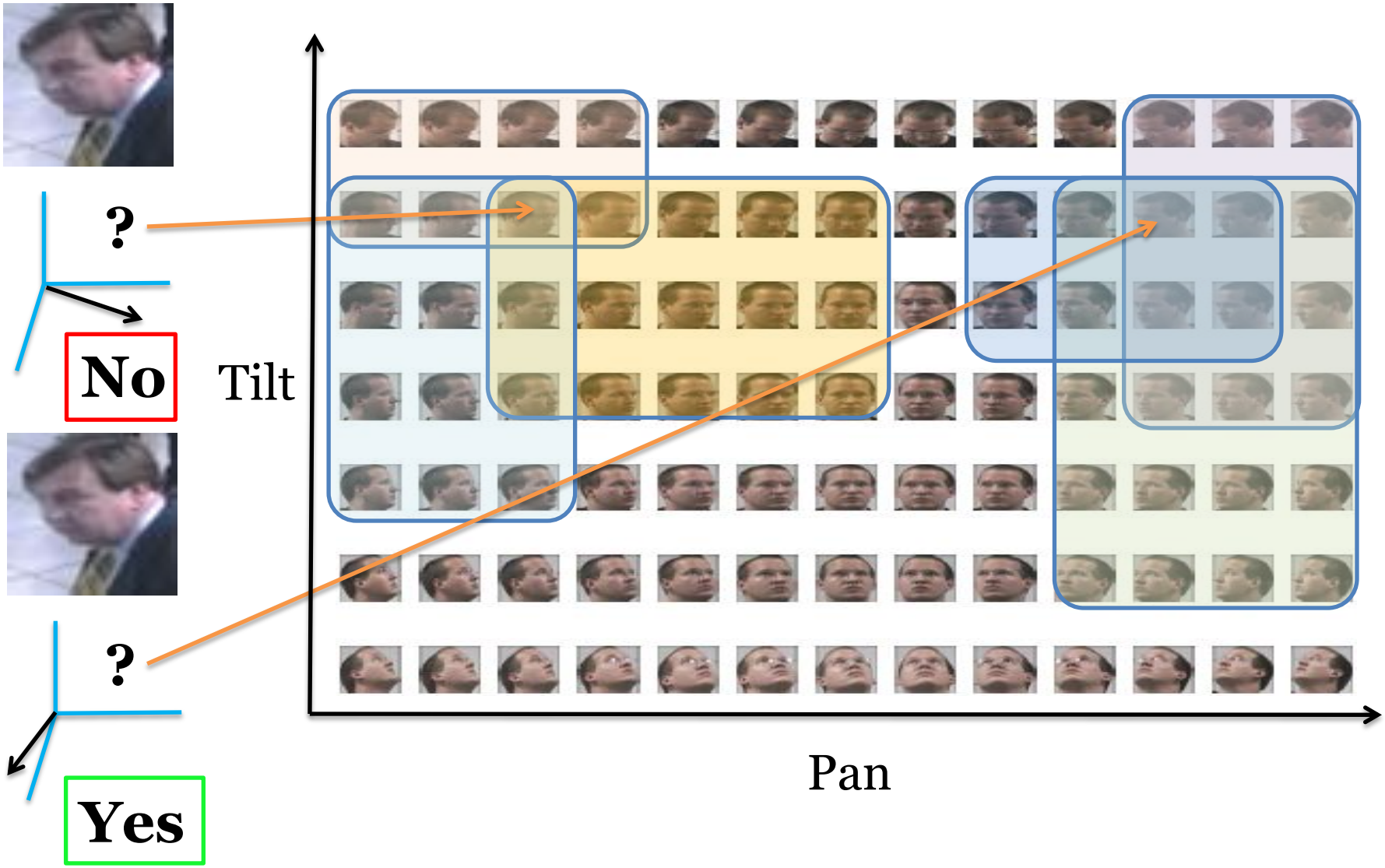
40,000 faces (IDIAP Head Pose Dataset)
Head-mounted tracker

80,000 non-faces

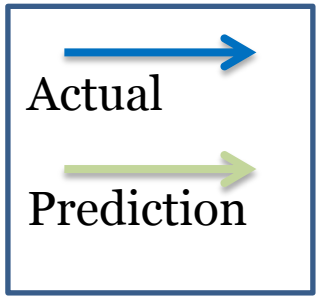
# Training time



# Determining head pose



# Results

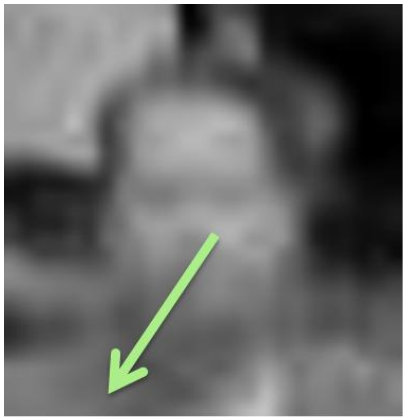
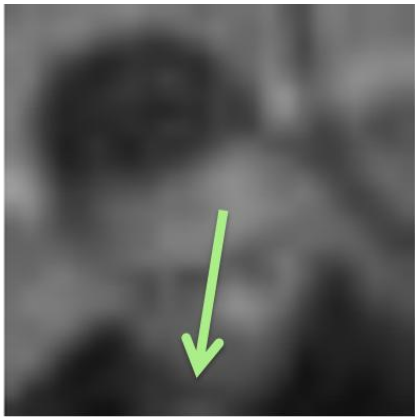


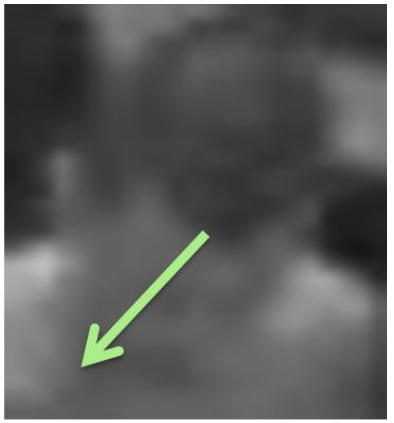

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Pan | -46.2° | -15° | | -6.8° | -20° | -9.6° | -10° |
| Tilt | 9.3° | -10° | | -10.9° | -10° | -26.8° | -5° |

## Tested on 2000 images

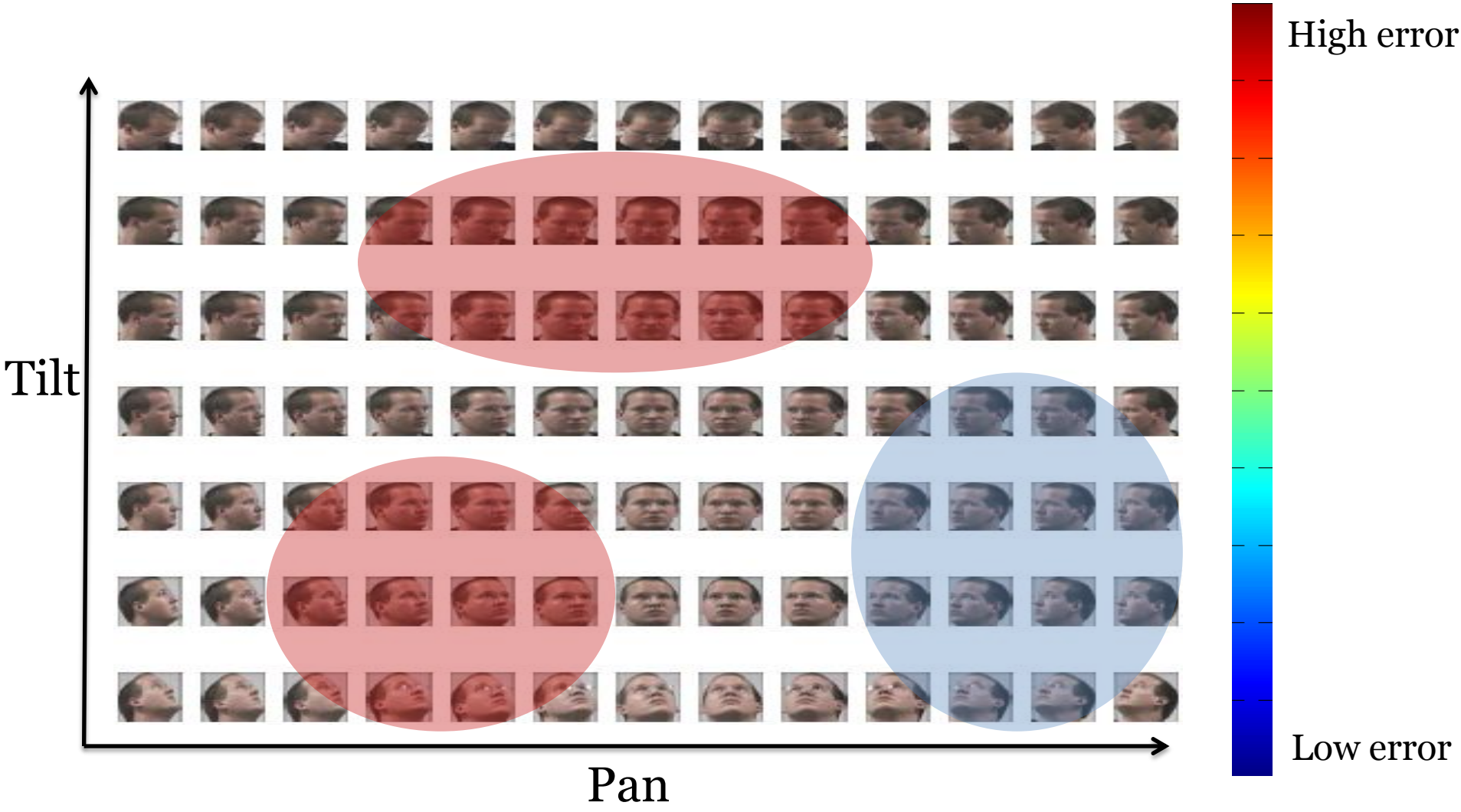
| Median Tilt Error | Median Pan Error |
|---|---|
| 29.8° | 10.3° |

# Results from camera network

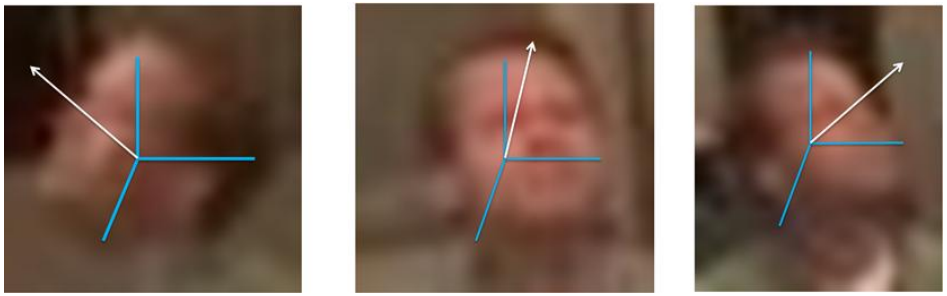



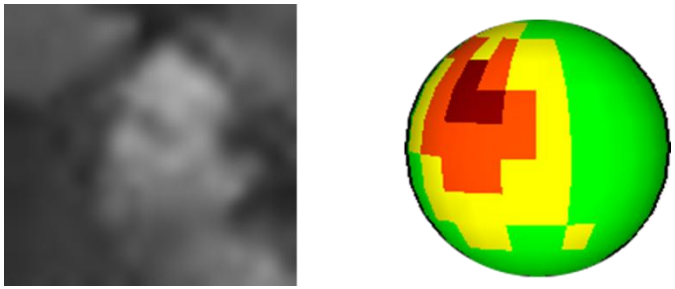


# Error map

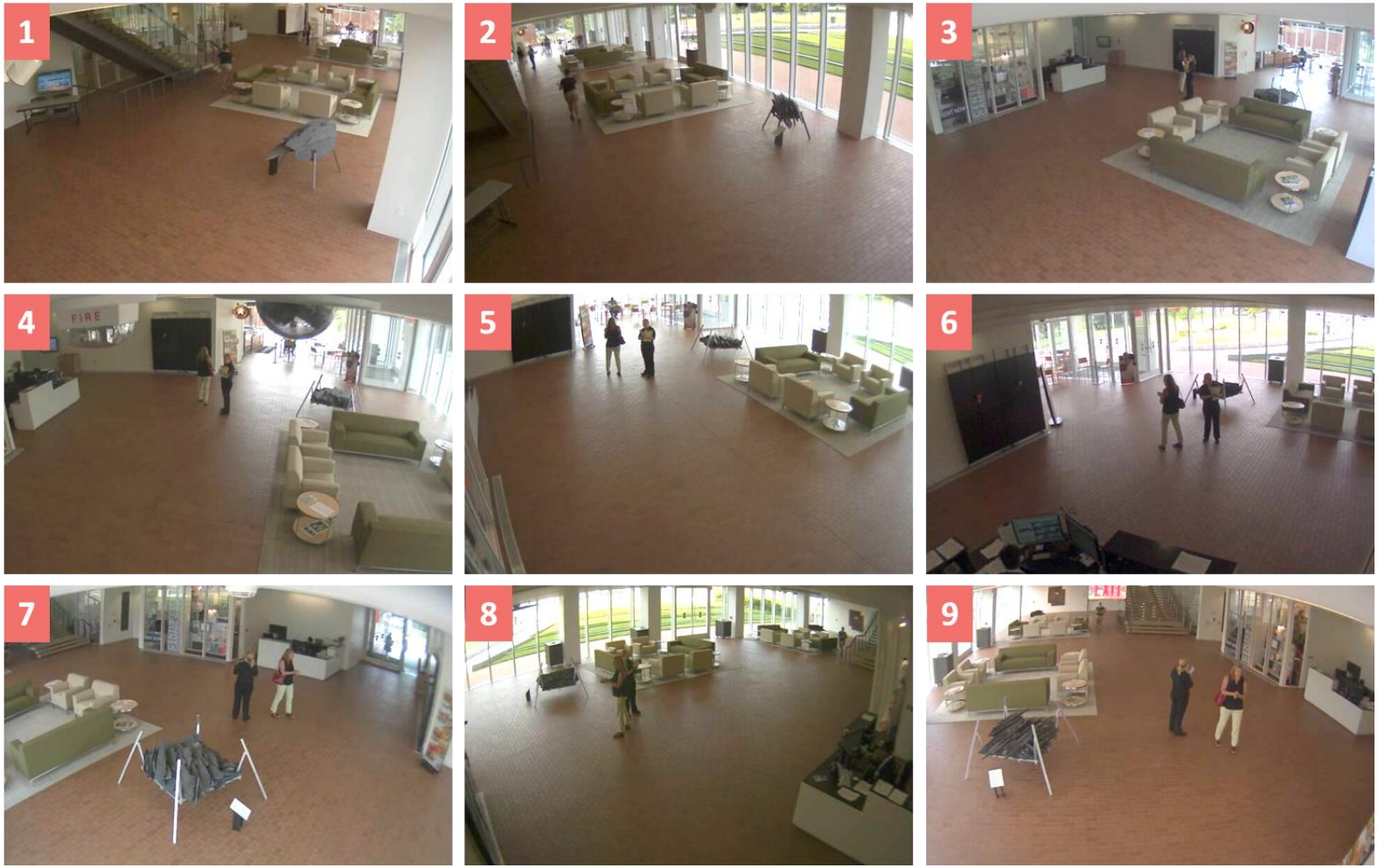


# Combine camera predictions



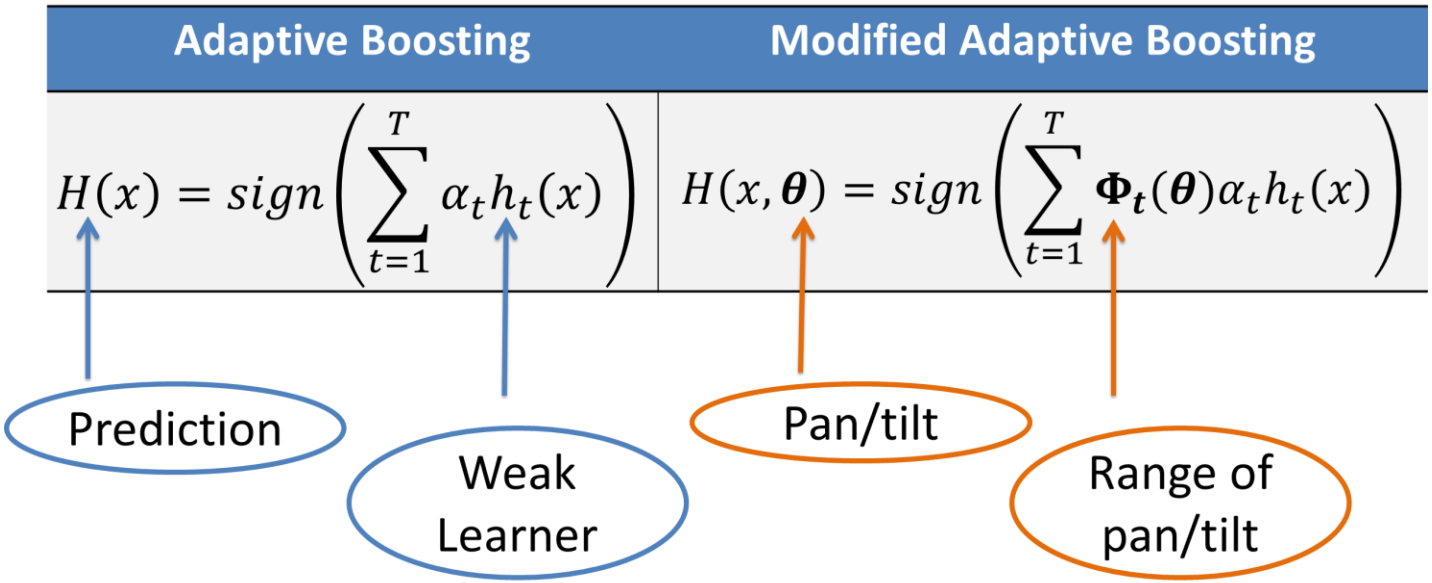


# Future work: Integrate into existing camera networks



# Questions?

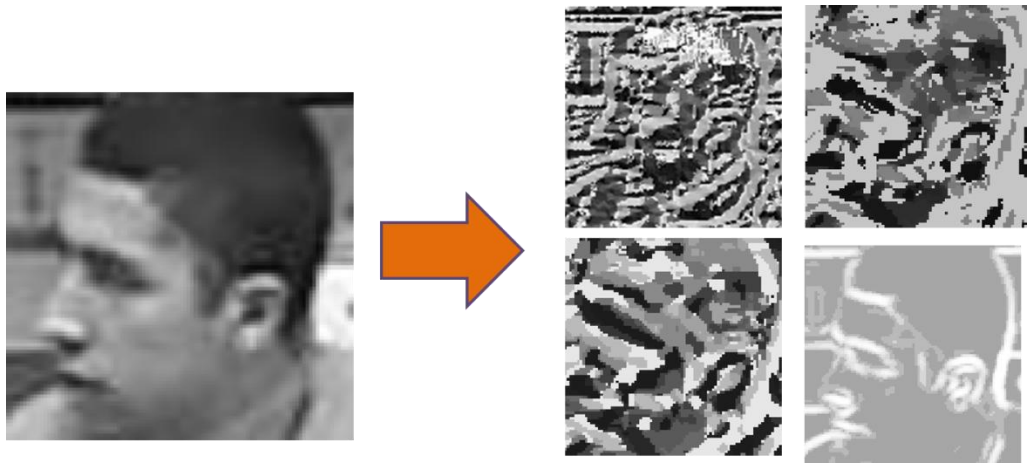
# Modified Adaptive Boosting

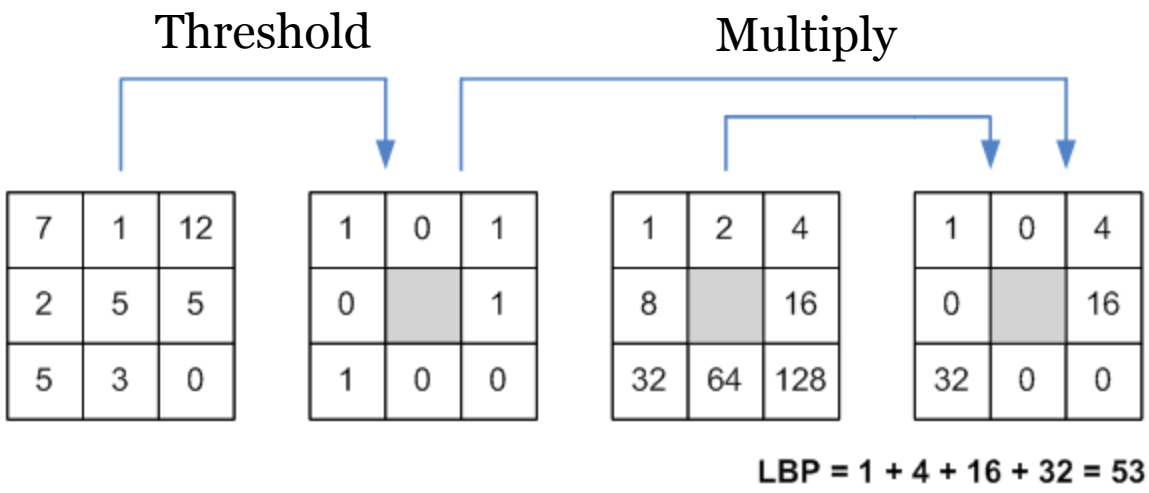



$\boldsymbol{H(x, \theta)}$: Prediction at pan/tilt angle

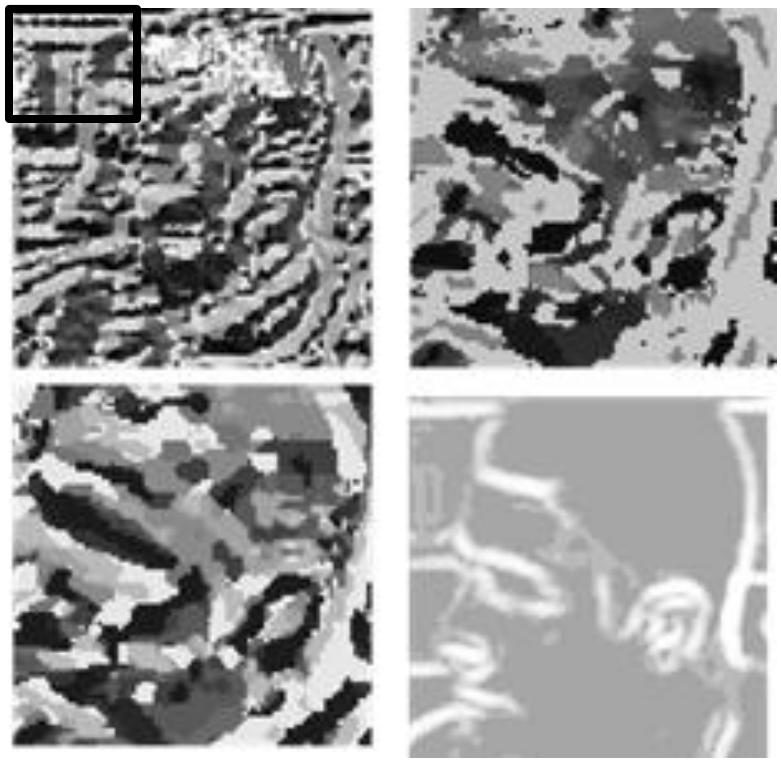
$\boldsymbol{\Phi_t(\theta)}$: { 1 if input is inside pan/tilt range
0 if input is outside pan/tilt range

# Local Binary Patterns

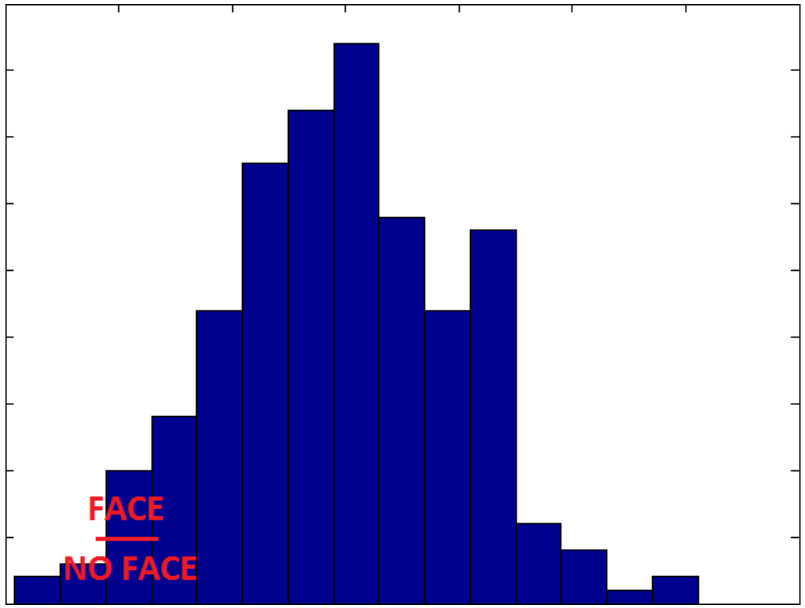




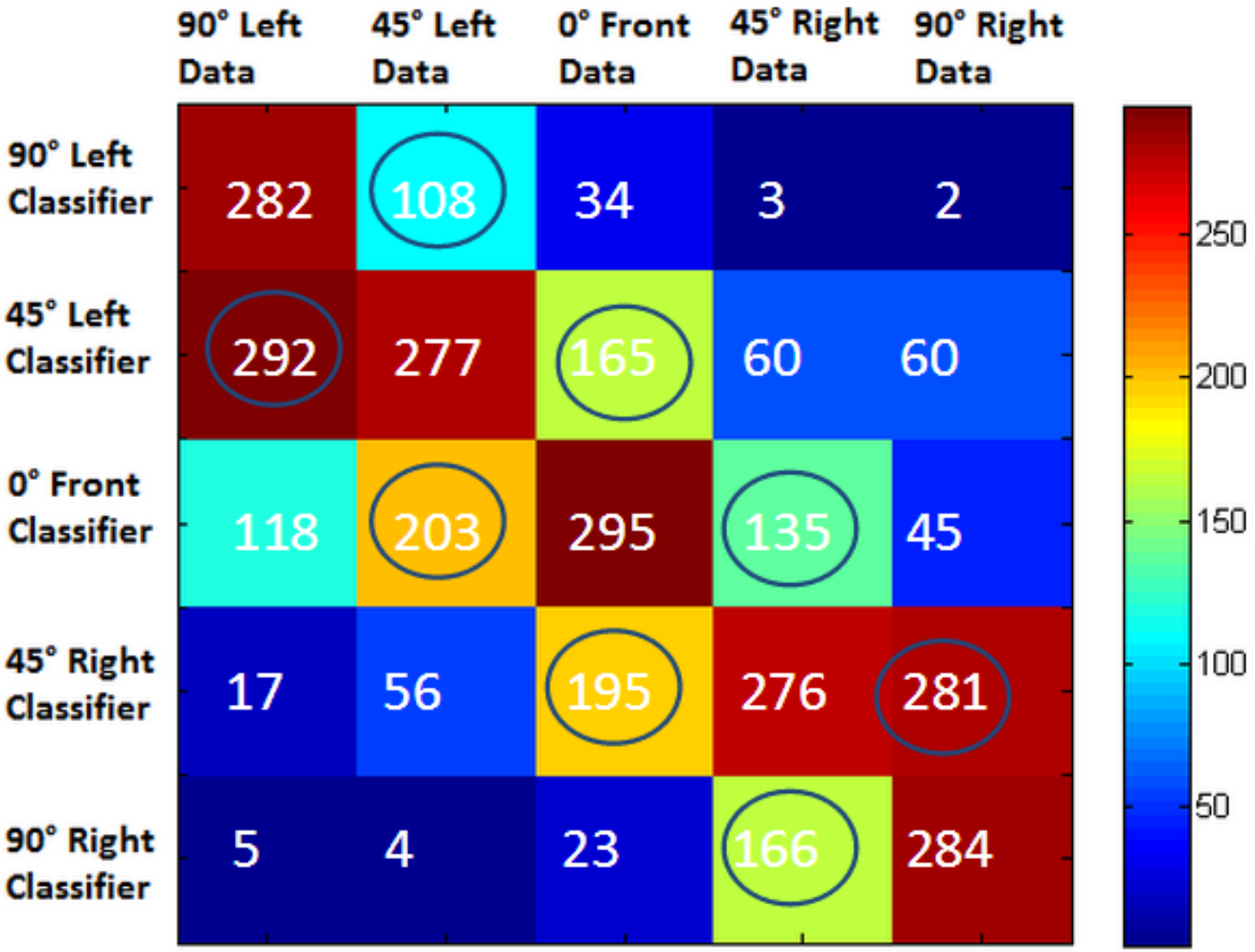
# Local Binary Patterns



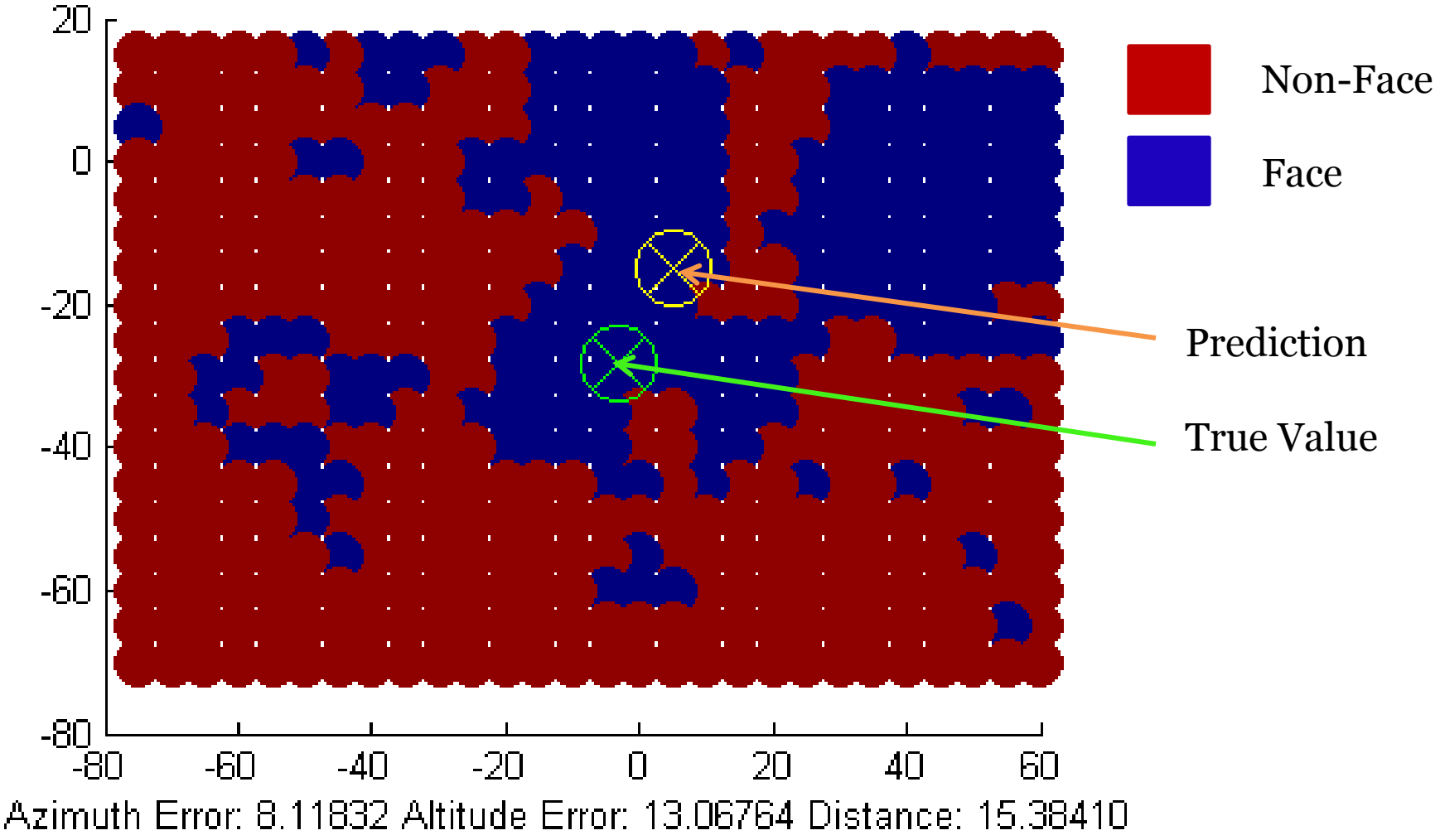
Feature histogram
(for every sub-window)



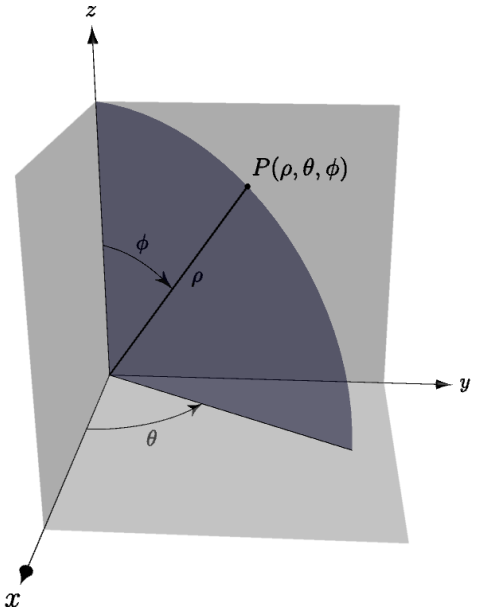
# Unintended Feature Sharing



# Evaluating Entire Gaze Space



# Spherical Coordinates




- Camera ⟶ world
- Head is origin (0,0,0)