Challenges in Head Pose Estimation of Drivers in Naturalistic Recordings Using Existing Tools

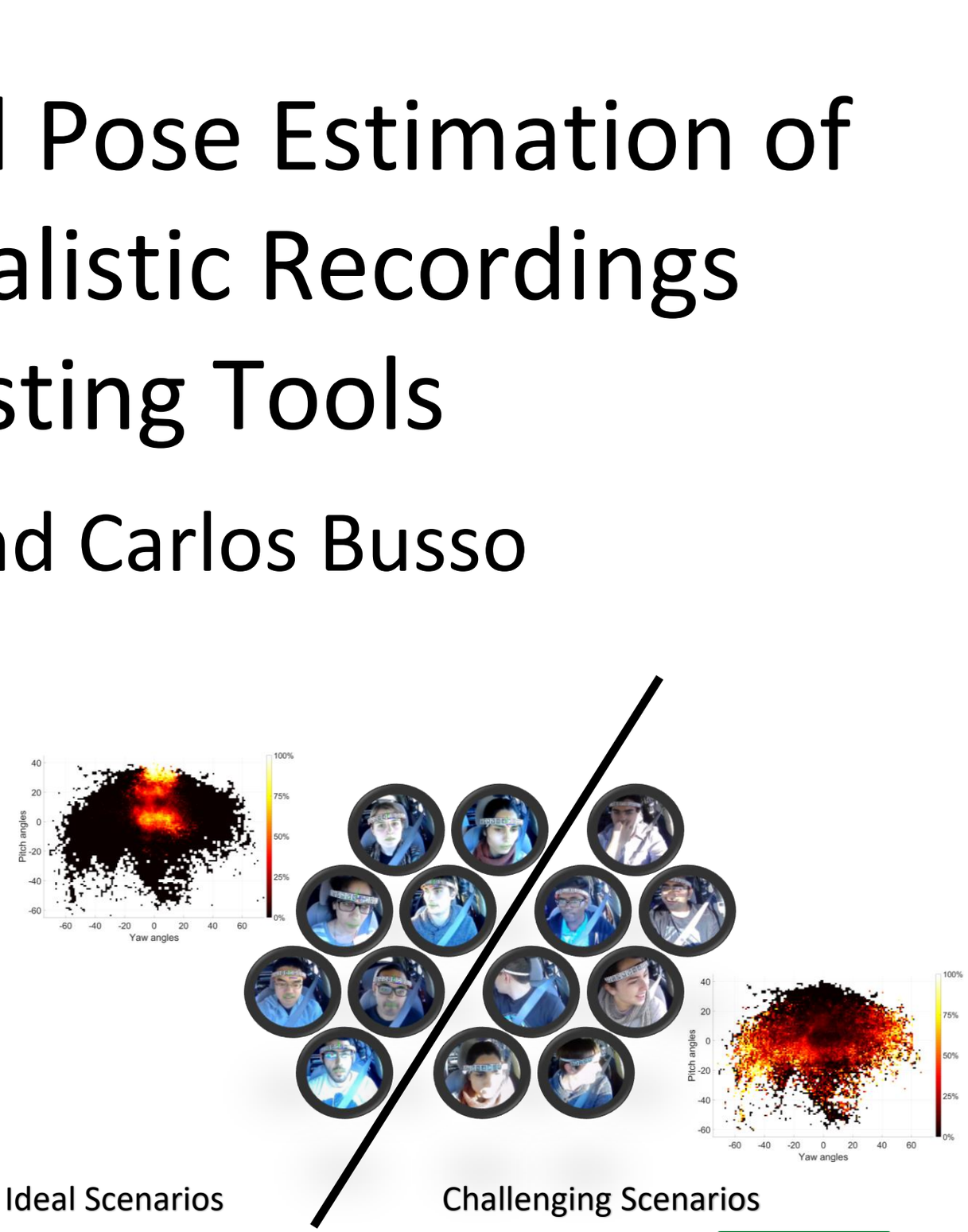
Sumit Jha and Carlos Busso

Multimodal Signal Processing (MSP) Laboratory

Department of Electrical Engineering,

The University of Texas at Dallas,

Richardson TX-75080, USA



Head Pose Estimation

- The position and orientation of the head find useful application in multiple interactive environment
  - Human computer interaction
  - Non-verbal communication
  - Visual attention
- In a vehicle setting
  - Visual attention of driver

Motivations

- Head Pose estimation in a controlled environment with limited head motion is a solved (almost) problem
- Additional challenges in driving environment
  - Wide variation in lighting
  - High head rotations
  - Occlusion

- Questions
  - What are the factors that affect the performance of Head Pose Estimation (HPE) algorithms?
  - What are the conditions where,
    - most algorithms work?
    - most algorithms fail?

Objective

- Use a reference head poses in a naturalistic driving dataset to study factors affecting head pose estimation
  - Glasses
  - Illumination
  - Head rotation

- Isolate frames which are easiest to process and the ones that are the most challenging
  - Ideal Scenario – Frames that always give good estimate
  - Challenging

Outline

- **Tools and Dataset**
- Factors affecting Head Pose Estimation
- Ideal Scenarios and Challenging Scenarios
- Conclusions and Future Work

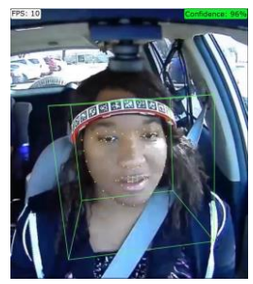
Tools analyzed

- We analyse 3 state-of-the-art head pose estimation tools
  - IntraFace [Xiong et al., 2013]

    Supervised Gradient Descent (SGD) to track non-linear features associated with each landmarks

  - OpenFace [Baltrusaitis et al., 2016]

    Conditional Local Neural Fields(CLNF) which learns the landmark shape and appearance variations



  - Zface [Jeni et al., 2015]

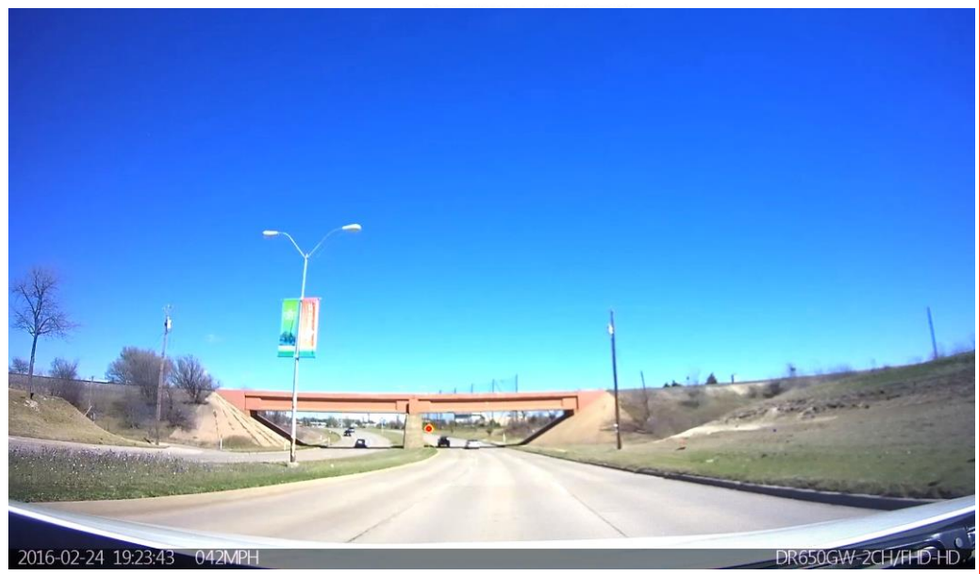
    Iteratively build a 3D mesh from the 2D landmarks to register a dense model

T. Baltrusaitis, P. Robinson, and L. P. Morency, "OpenFace: An open source facial behavior analysis toolkit," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1–10, Lake Placid, NY, USA, March 2016. IEEE.

L. A. Jeni, J. F. Cohn, and T. Kanade, "Dense 3d face alignment from 2d videos in real-time," in *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, volume 1, pages 1–8, Ljubljana, Slovenia, May 2015. IEEE.

Database

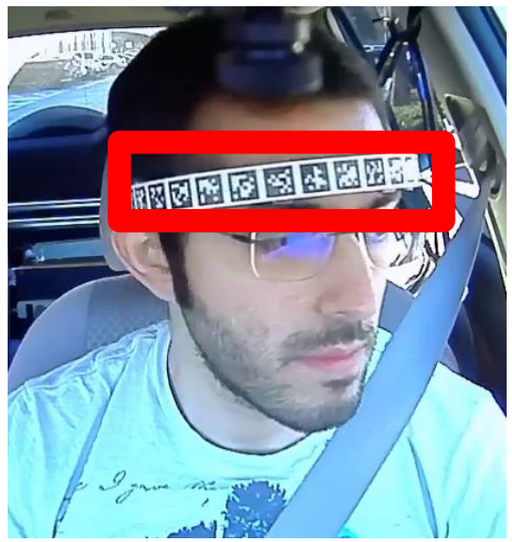
- Collected naturalistic driving data in the UTDrive platform
- Dash Cameras record the road and driver's face
  - Blackvue dr650gw 2 channel
    - Rear camera records the face
    - Front camera records the road
- Data Collected with 16 subjects (10 males and 6 females)
  ~ 6 hours of naturalistic driving



AprilTags for Head Pose Estimation

- AprilTags [Olson, 2011]: 2D barcodes that can be robustly detected in an image

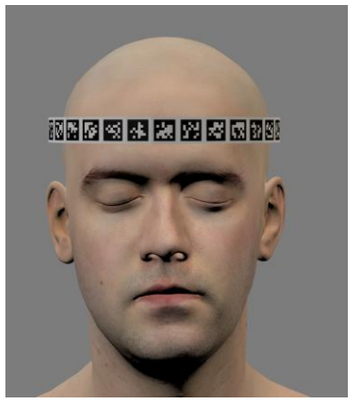
- Headband designed with 17 AprilTags
- Used to establish reference head position and orientation

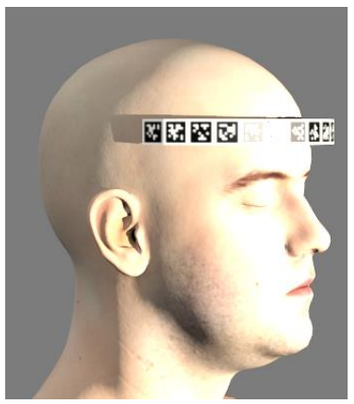


Olson, Edwin. "AprilTag: A robust and flexible visual fiducial system." Robotics and Automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011.

Performance of AprilTag system

- Accuracy of AprilTag based detection
- Rendered the band on a head in virtual environment
- Studied the performance in various quality of rendering and adding external effects like illumination

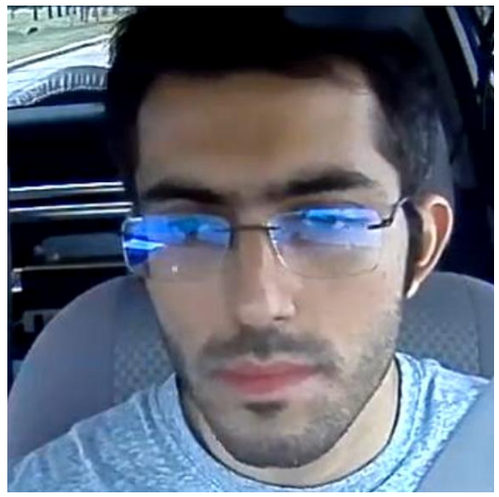




Data condition	Median angle error
High Quality render (3840 x 2160)	0.89°
Medium Quality render (960 x 540p)	1.26°
Data with added illumination	2.69°

Affect of Headband on HPE

- Head band occludes a part of forehead which can confuse HPE
- 7 subject collected without headband for comparison
- Frames missed by each algorithm



	From AprilTag	IntraFace	OpenFace	Zface
With AprilTag	5.3 %	27.3 %	24.1 %	21.9 %
Without AprilTag	-	19.0 %	21.8 %	8.9 %

Outline

- Tools and Dataset
- **Factors affecting Head Pose Estimation**
- Ideal Scenarios and Challenging Scenarios
- Conclusions and Future Work

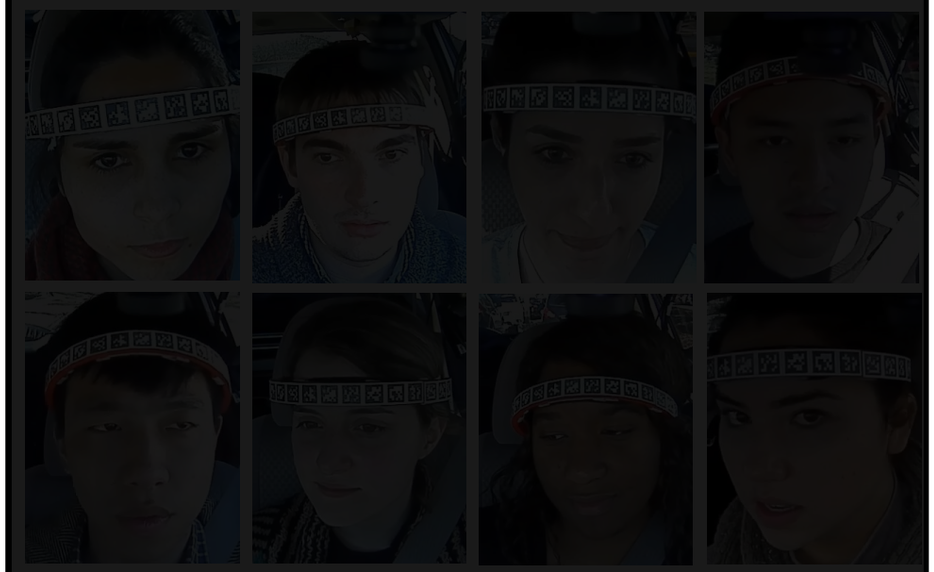
Factors affecting Head Pose Estimation

- We study factors that effects Head Pose Estimation in driving environment
  - Glasses
  - Illumination
  - Head Rotation

Occlusion due to glasses

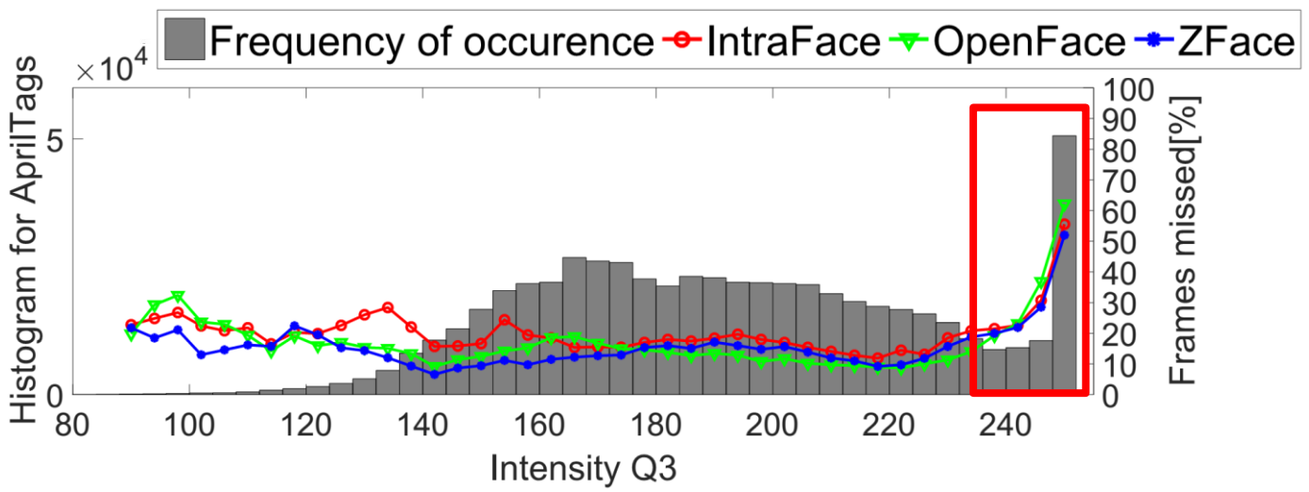
- Glasses occlude the face affecting performance of HPE
- Percentage of the total frames that failed detection by each algorithm

Method	No Glasses	Glasses with thick frame	Glasses with Normal frame
IntraFace	15.40%	67.70%	18.50%
OpenFace	12.10%	67.50%	13.70%
Zface	8.80%	64.10%	13.80%



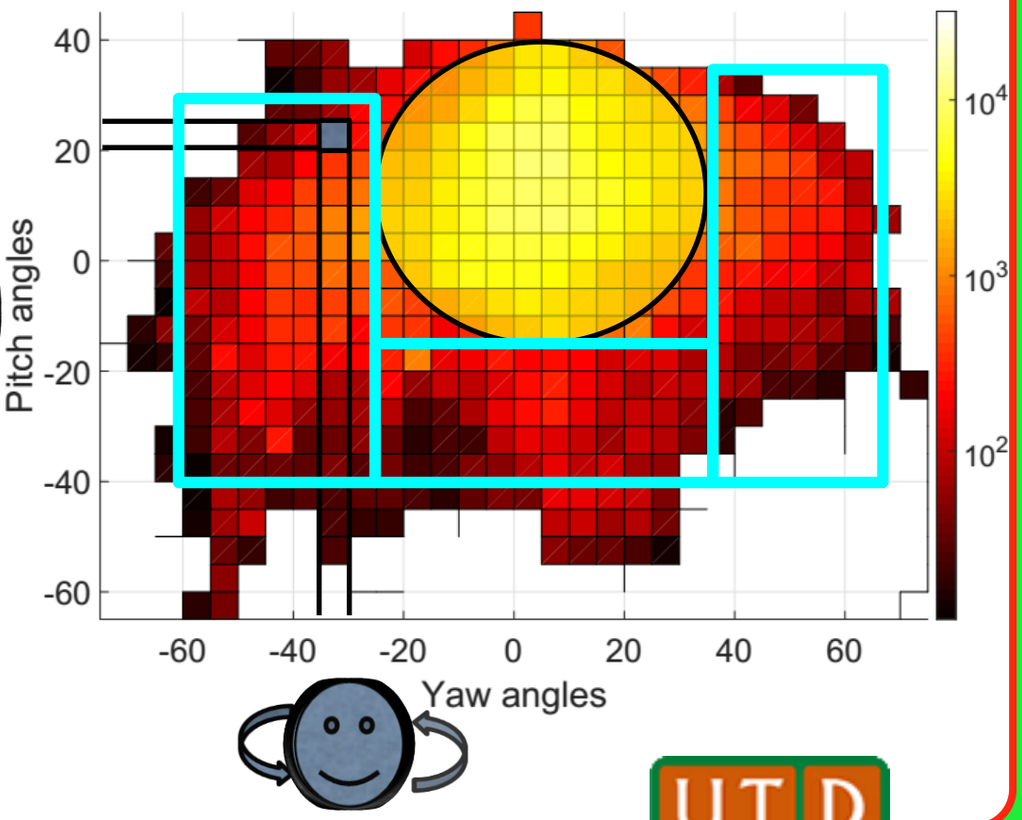
Effect of Illumination

- Both high and low illumination affects the quality of image
  - We study the effect of saturation of the face image
  - Partial or total saturation of face image
- Performance depends on the third quartile of the face image
  - Third Quartile of the intensity(Q3) is high when part of the face is saturated



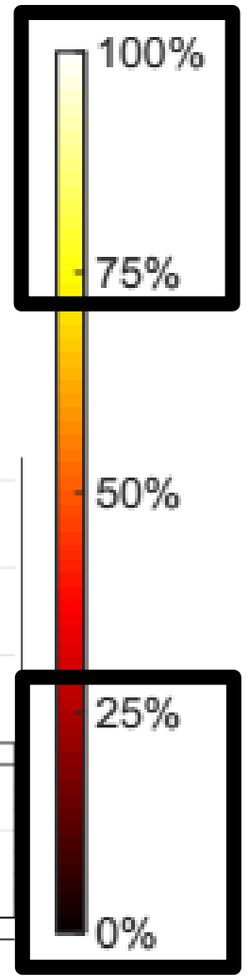
Effect of head rotation

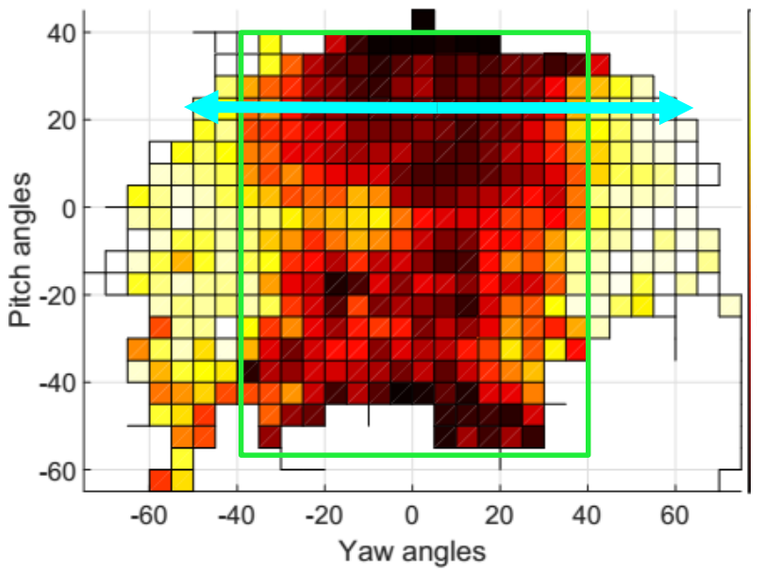
- Face detection and head pose estimation affected by high head rotation
  - Most tools only work for frontal and semi-frontal faces
- Naturalistic driving scenario
  - Distribution of head poses
    - Bright – High frequency
    - Dark – Low frequency
  - Most of the time head pose is frontal
  - The robustness is more crucial when head pose is non frontal



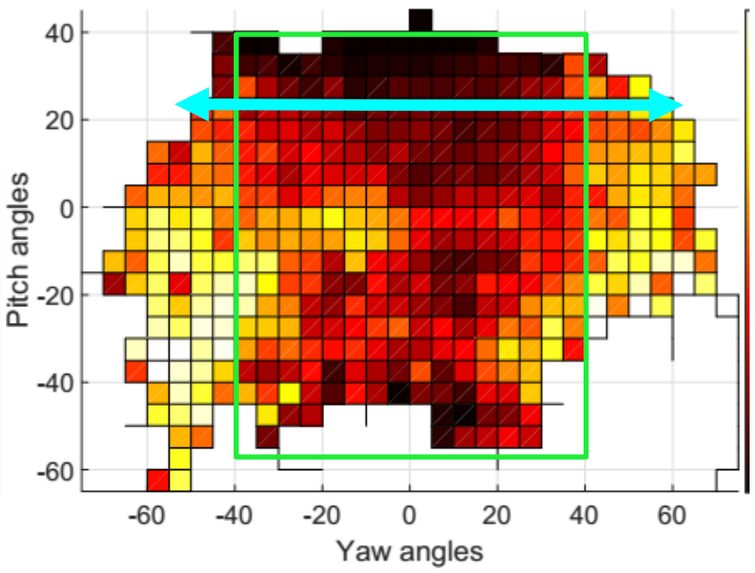
Percentage of Frames Missed by the HPEs at Different Angles

- Analysis of percentage of face missed by HPE at different angles
  - Bright pixels – most frames not detected by HPE
  - Dark pixels – few frames not detected by HPE

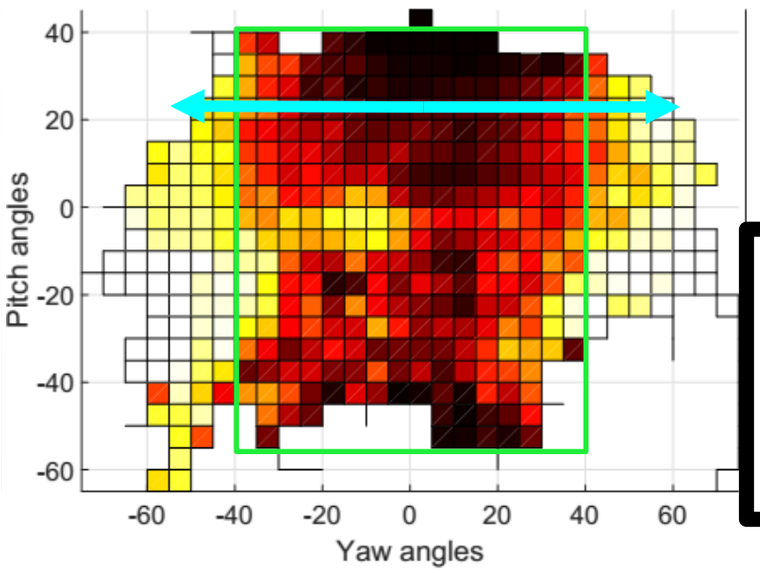




IntraFace



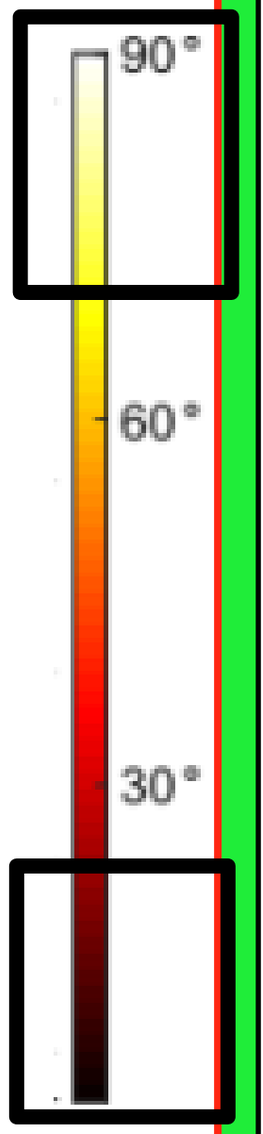
OpenFace

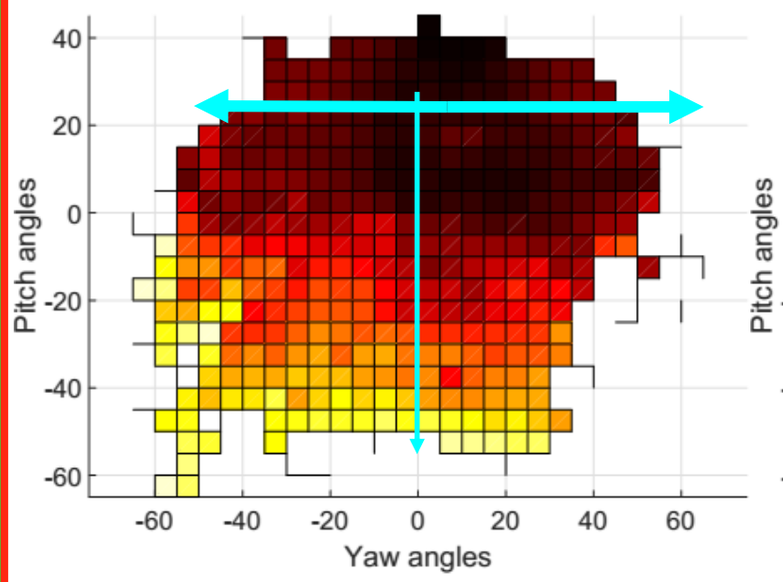


ZFace

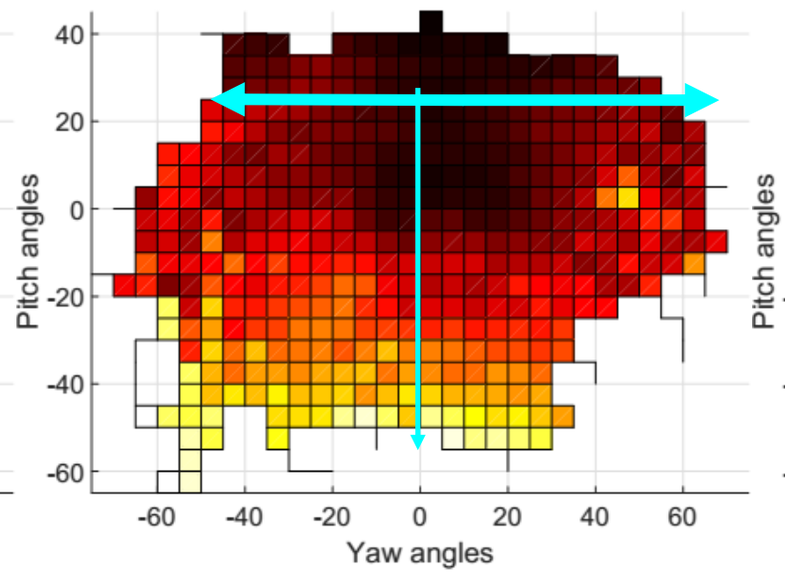
Difference in Angle between estimates from AprilTags and HPEs

- Difference in estimation for the frames detected by each Algorithm
  - Bright Pixels – Large difference in angular estimation
  - Dark Pixels – Small difference in angular estimation

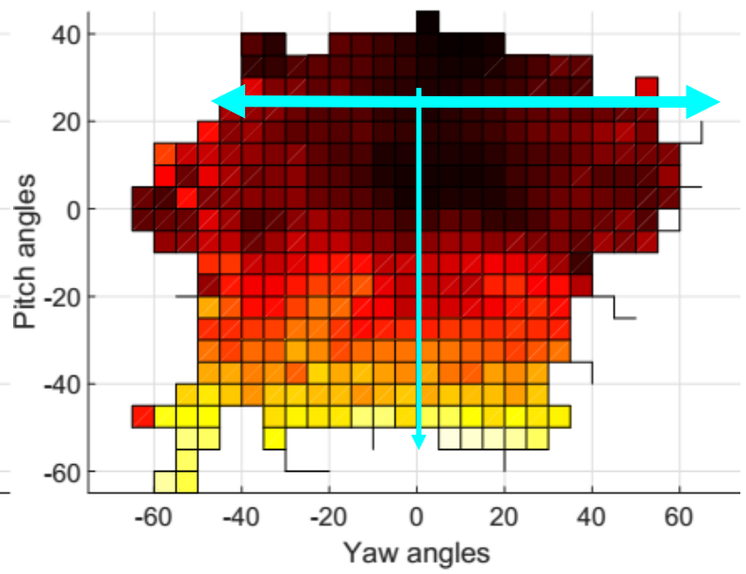




IntraFace



OpenFace



ZFace

Outline

- Tools and Dataset
- Factors affecting Head Pose Estimation
- **Ideal Scenarios and Challenging Scenarios**
- Conclusions and Future Work

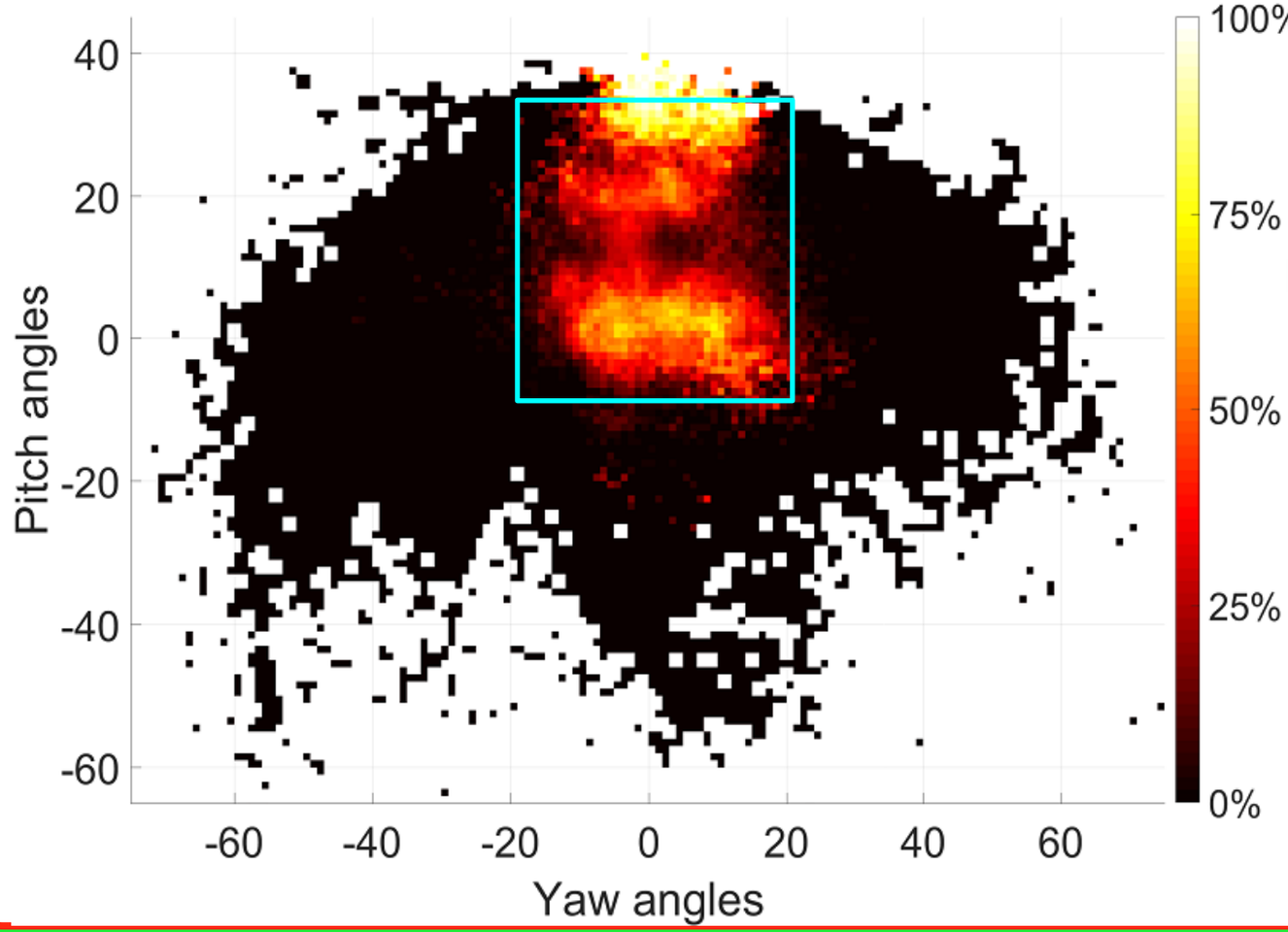
Ideal Scenarios(IS) and Challenging Scenarios(CS)

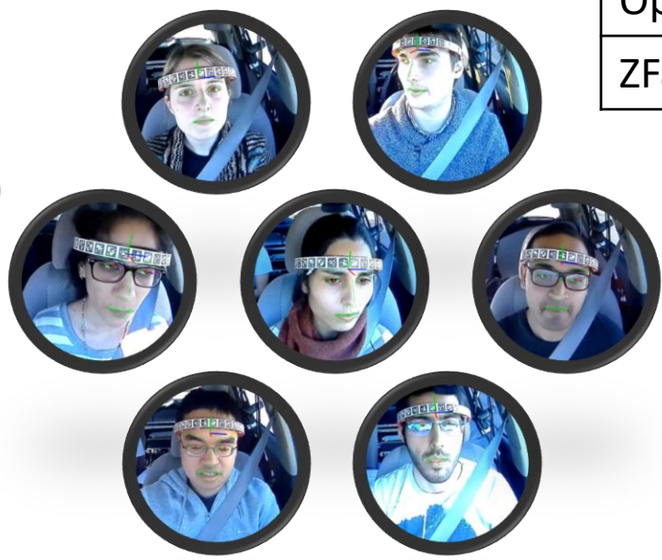
- We extract two types of frames from the database
  - Ideal Scenarios (IS) : Frames successfully detected and estimation error less than 10°
  - Challenging Scenarios (CS) : Frames that failed detection by all the three HPEs
- Helps in design of more robust algorithms that work for challenging cases

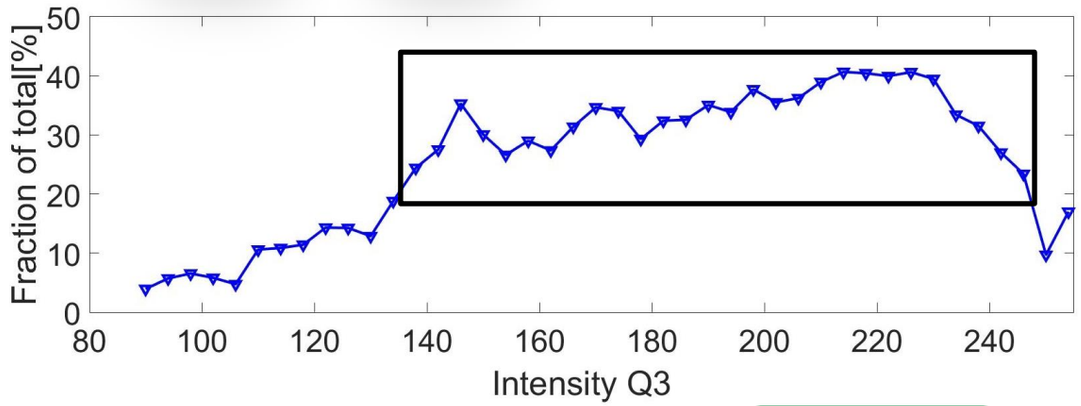
Ideal Scenarios

- Distribution of Ideal frames at different rotation angles and illumination

IntraFace	✓
OpenFace	✓
ZFace	✓



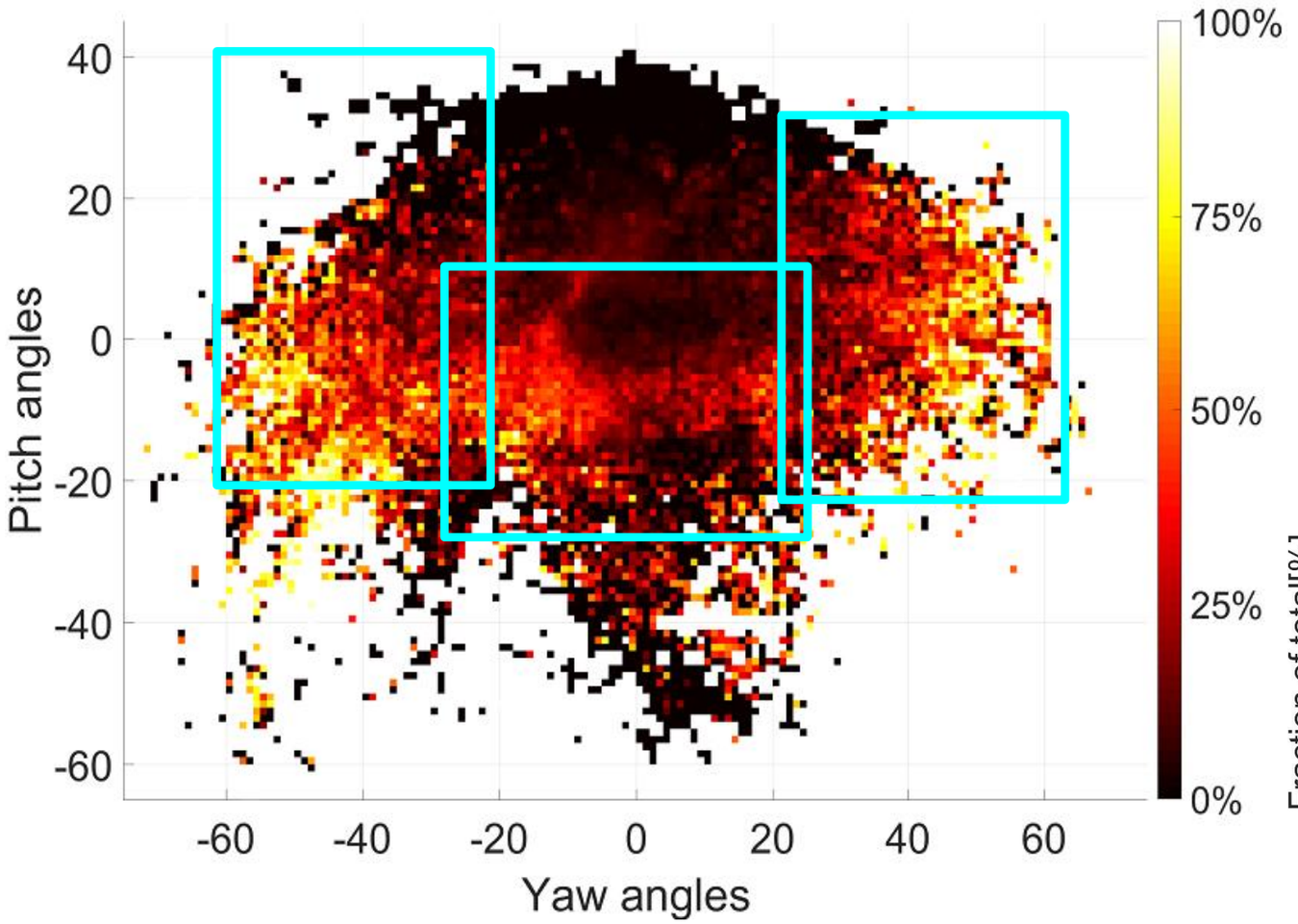


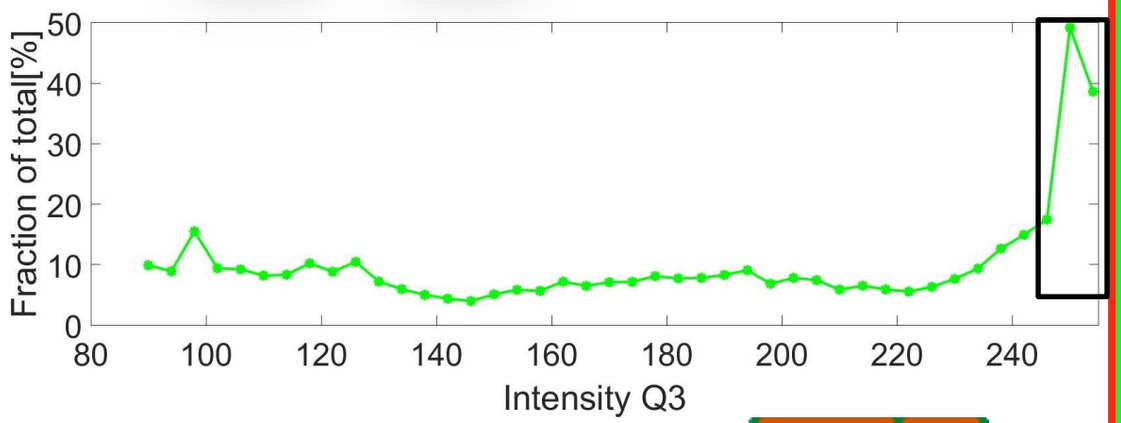


Challenging Scenarios

- Distribution of Challenging frames at different rotation angles and illumination

IntraFace	X
OpenFace	X
ZFace	X





Outline

- Tools and Dataset
- Factors affecting Head Pose Estimation
- Ideal Scenarios and Challenging Scenarios
- **Conclusions and Future Work**

Conclusions and Future Work

- Open access face processing tools have limited reliability in naturalistic driving environment
- Reliable estimation of head pose can be useful in designing smart systems in car
- Future Work
  - A more robust reference system with minimal obtrusion
  - Investigate and evaluate other modalities such as depth sensing cameras
  - Extend the database with more subjects under varying conditions

Thank you!

msp.utdallas.edu