Analyzing the Relationship Between Head Pose and Gaze to Model Driver Visual Attention

Sumit Jha and Carlos Busso

Multimodal Signal Processing (MSP) Laboratory

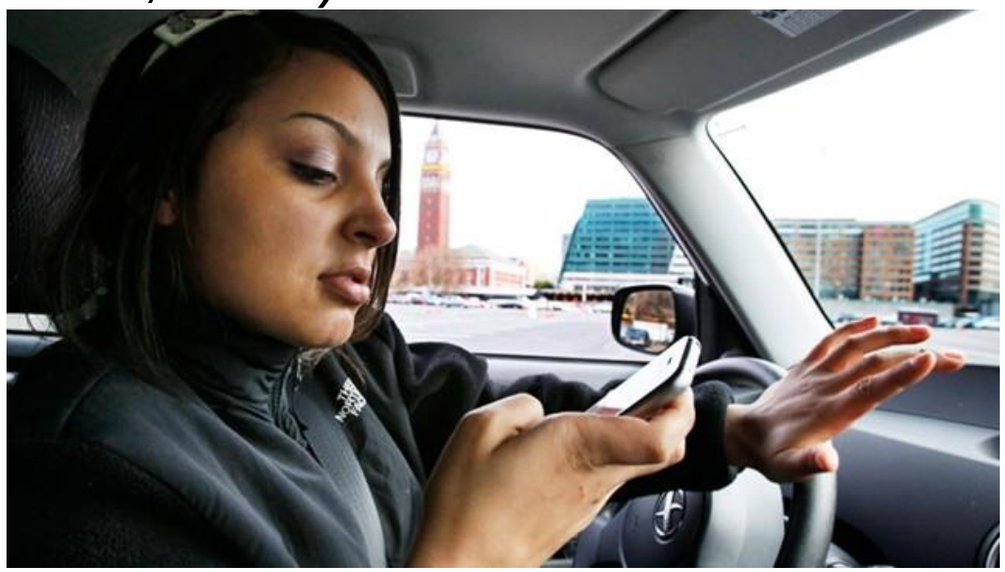
Department of Electrical Engineering,

The University of Texas at Dallas,

Richardson TX-75080, USA

Drivers' Visual Attention

- Primary driving related task
  - Mirror checking actions (Li and Busso, 2016)
  - Lane change
  - Turns and cross sections



- Secondary tasks
  - Mobile Phones and In-vehicle entertainment unit
  - Co-passengers in the car
  - Billboards and other distractions from the environment

Motivations

- Gaze detection is a challenging problem in car environment
- It is often approximated by head pose [Lee et al., 2011]
- Coarse direction of driver's gaze is enough for most in-vehicle applications [Tawari & Trivedi, 2014; Doshi & Trivedi, 2009]

Left mirror

Right mirror

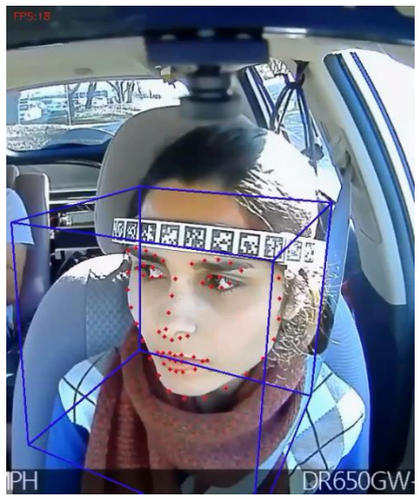
Rear mirror

- Goal of this study is to analyze the relationship between gaze and head pose

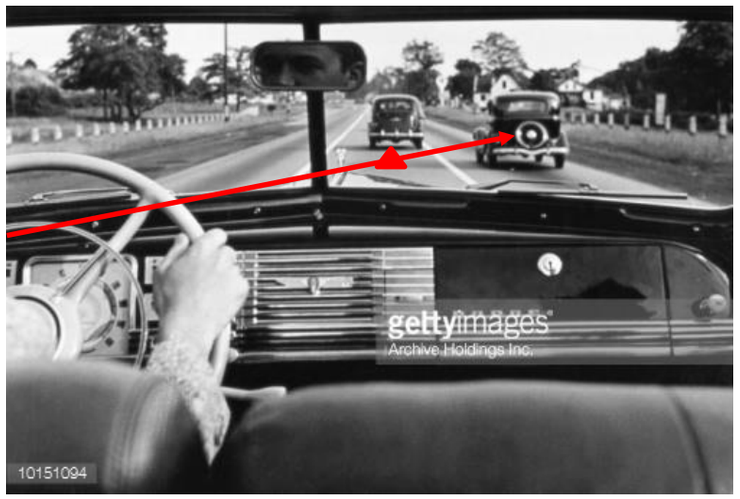
Objective

- Questions
  - How well can we estimate the head pose in a real world driving environment?
  - How well does the head pose of the driver predict his/her gaze (visual attention)?
  - How much does the head pose varies when the driver is looking at a certain direction?

Head Pose Estimation



Gaze Detection



Outline

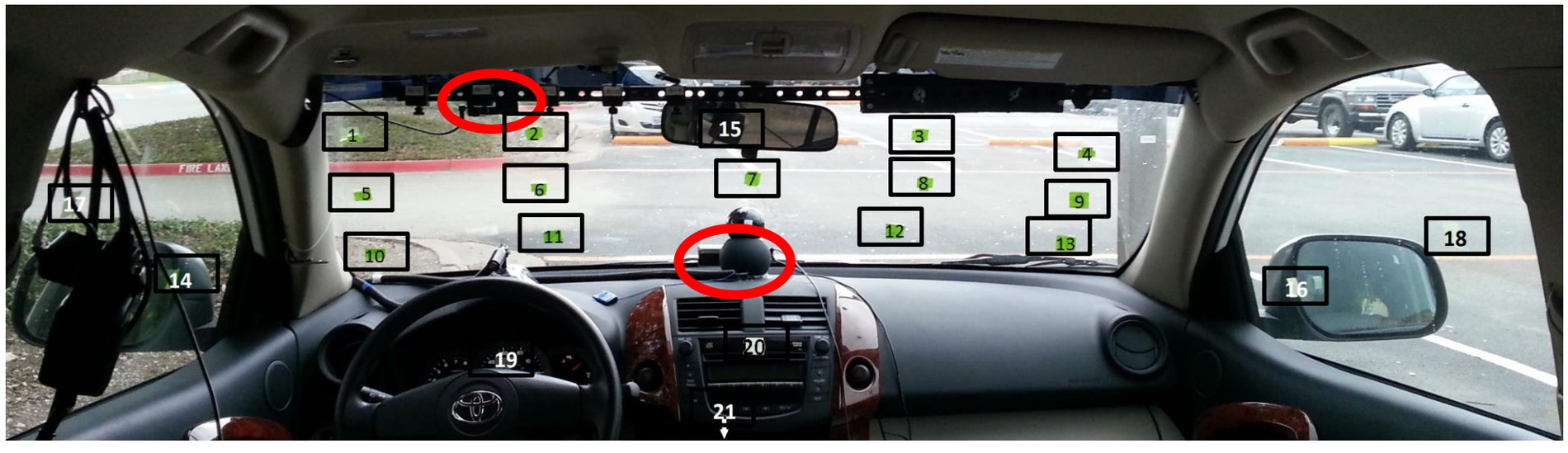
- **Data collection**
- Performance of head pose estimation
- Gaze estimation using linear regression
- Study of eye movement bias
- Conclusion

Data Collection

- To relate the facial image to ground truth gaze locations
- UTDrive platform

- Dash Cameras used instead of the on-board equipment
  - (Blackvue dr650gw 2 channel)
    - 2 channel camera
    - with WiFi, GPS and accelerometer

Experimental Setup



- Rear camera → Face
- Front camera → Road

- Markers placed at the wind
  16), side windows (17-18),
  radio (20), and gear (21)
- Data collected with 16 subjects (10 males, 6 females) in three phases.

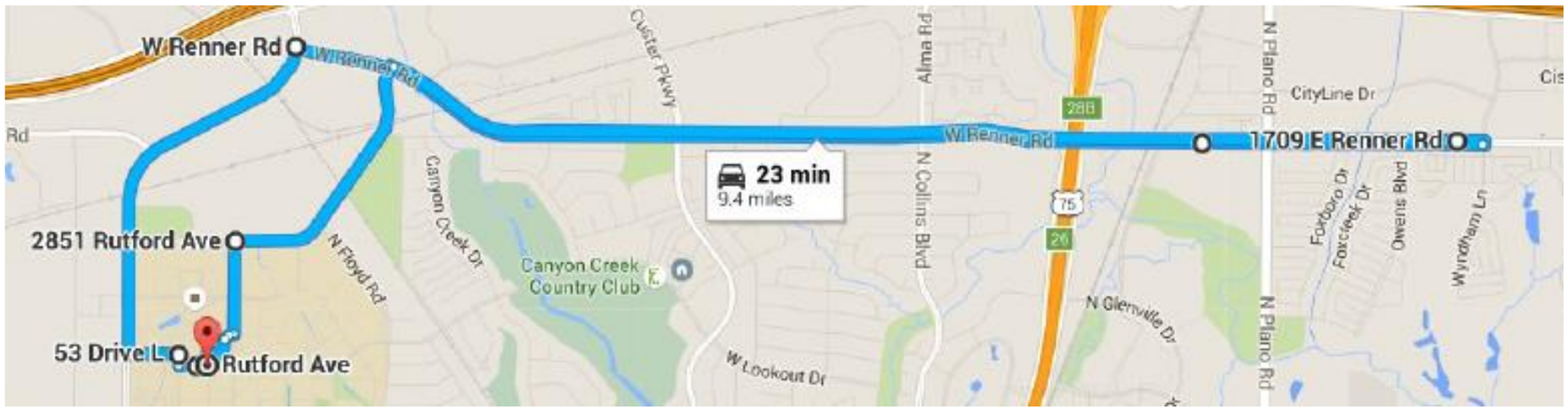
Phase 1
(Natural Gaze – Parked Vehicle)

- Collected in a parked car
- Subject asked to look at each point multiple times

- Natural variability in head pose without the constraint of driving task
- The driver familiarizes to the core task

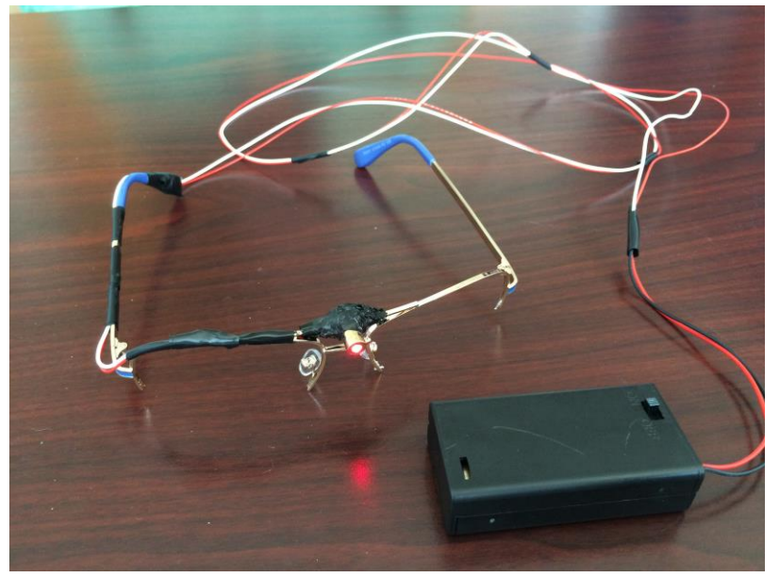
Phase 2
(Natural Gaze - Driving)

- Collected when the subject is driving the car
- Subject asked to look at points
- Data collected in a straight road with minimum maneuvering task



Phase 3
(Controlled Gaze – Parked Vehicle)

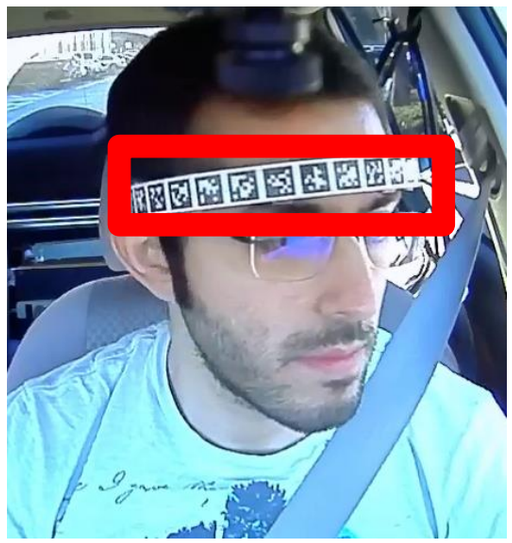
- Direct head pose toward markers
  - Head pose ≈ gaze
  - No bias due to eye movement



- Difficult to achieve naturally
  - Used a glass frame with laser mounted at the center
- Subjects point at the target marks with the beam

AprilTags for Head Pose Estimation

- Head pose estimation challenging in driving environment
- AprilTags (Olson, 2011): 2D barcodes that can be robustly detected in an image
- Headband designed with 17 AprilTags
- Useful for robust detection of head pose across conditions

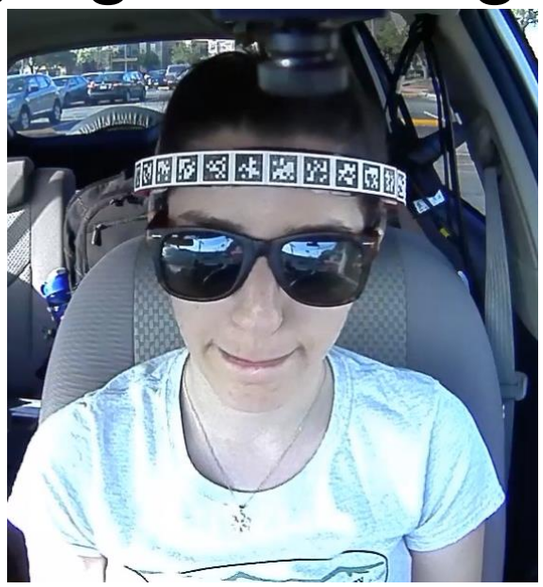


Outline

- Data collection
- **Performance of head pose estimation**
  - **Question 1: How well can we estimate the head pose in a real world driving environment?**
- Gaze estimation using linear regression
- Study of eye movement bias
- Conclusion

Performance of Head pose Estimation Algorithm

- Head Pose estimation challenging in driving environment
  - Wide variation in lighting
  - High head rotations
  - Occlusion



- We Study a state-of-the-art head pose estimation algorithm (HPA) (Baltrusaitis et al. 2013)
  - Representative performance with respect to other good head pose estimation algorithms

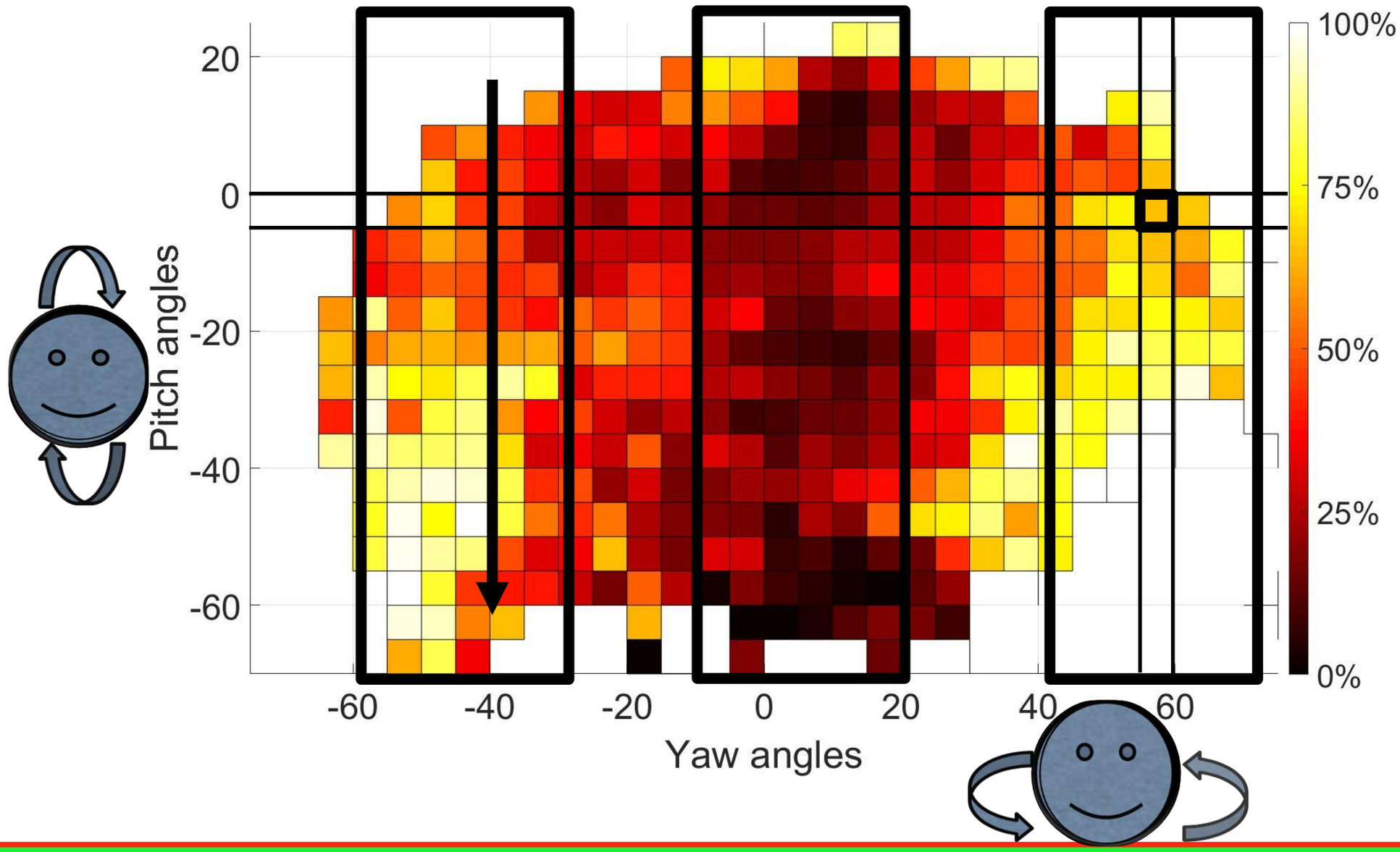
Performance of Head Pose Estimation Algorithm (HPA)

- Analysis performed on all the frames when the subject was driving
- Frames detected by the HPA compared to the AprilTag

HPA / AprilTag	Face detected	Face not detected	
Tag detected	73.2%	21.51%	**94.71%**
Tag not detected	2.25%	3.03%	**5.28%**
	75.45%	**24.54%**	

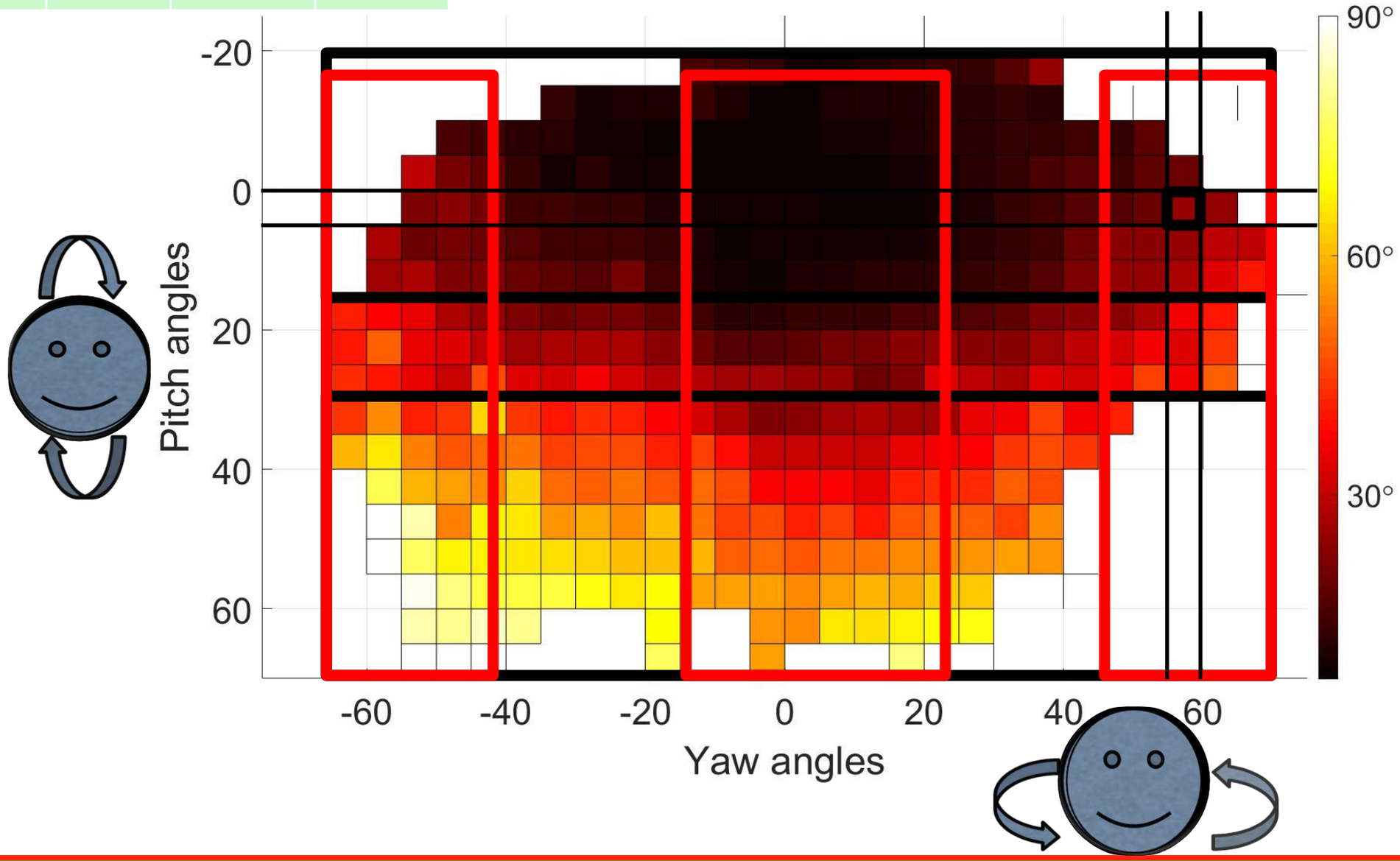
Percentage of Frames Missed by the HPA at Different Angles

	Face detected	Face not detected	
Tag detected	73.2%	21.51%	**94.71%**
Tag not detected	2.25%	3.03%	**5.28%**
	75.45%	**24.54%**	



Mean Absolute Angle Difference between AprilTags and HPA

	Face detected	Face not detected	
Tag detected	73.2%	21.51%	**94.71%**
Tag not detected	2.25%	3.03%	**5.28%**
	75.45%	**24.54%**	



Outline

- Data collection
- Performance of head pose estimation
- **Gaze estimation using linear regression**
  - **Question 2: How well does the head pose of the driver predict his/her gaze (visual attention)?**
- Study of eye movement bias
- Conclusion

Linear Regression Model for Gaze Estimation

- Investigate the linear relationship between head pose and gaze location
- Model Trained
  - $x_0 = a_0 + \underbrace{a_1 x + a_2 y + a_3 z}_{\text{Position}} + \underbrace{a_4 \alpha + a_5 \beta + a_6 \gamma}_{\text{Orientation}}$
- Driver independent partition
  - 10 training, 6 testing

Linear Regression (Contd.)

- R-squared value

	Phase 1 (Natural-Parked)		Phase 2 (Natural-Driving)		Phase 3 Controlled*	
	Train	Test	Train	Test	Train	Test
x_0	0.78	0.77	0.69	0.73	0.91	0.87
y_0	0.36	0.12	0.36	0.16	0.66	0.31
z_0	0.25	0.10	0.24	0.12	0.31	0.25

* Head Pose ≈ Gaze

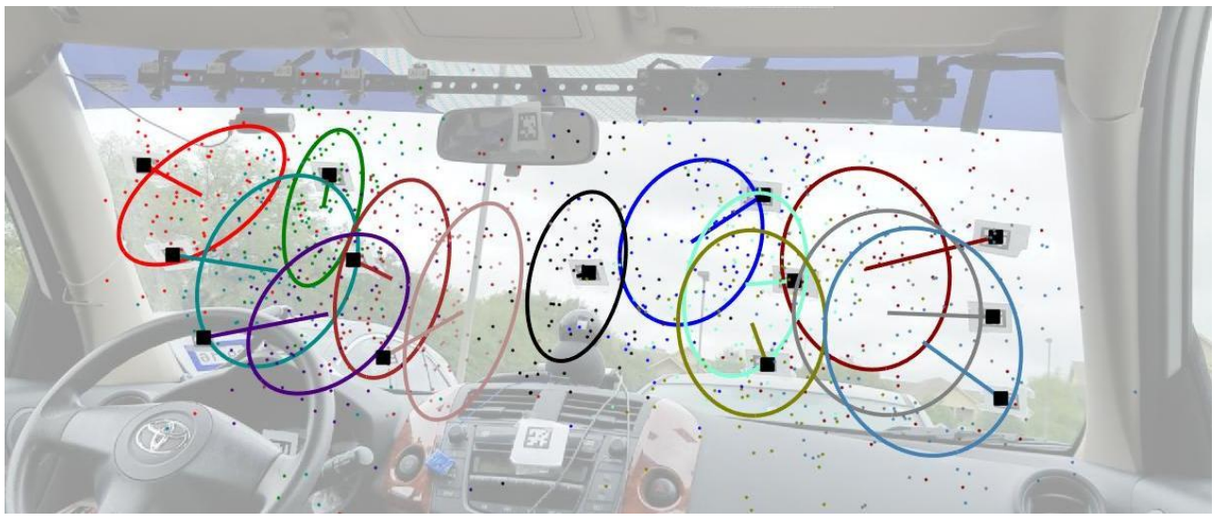
- High correlation in Horizontal direction ➔ But deterministic prediction of gaze not possible
- Low R^2 values of y ➔Low predictability in pitch direction
- High values in Phase III ➔ No eye movement therefore more predictability`

Outline

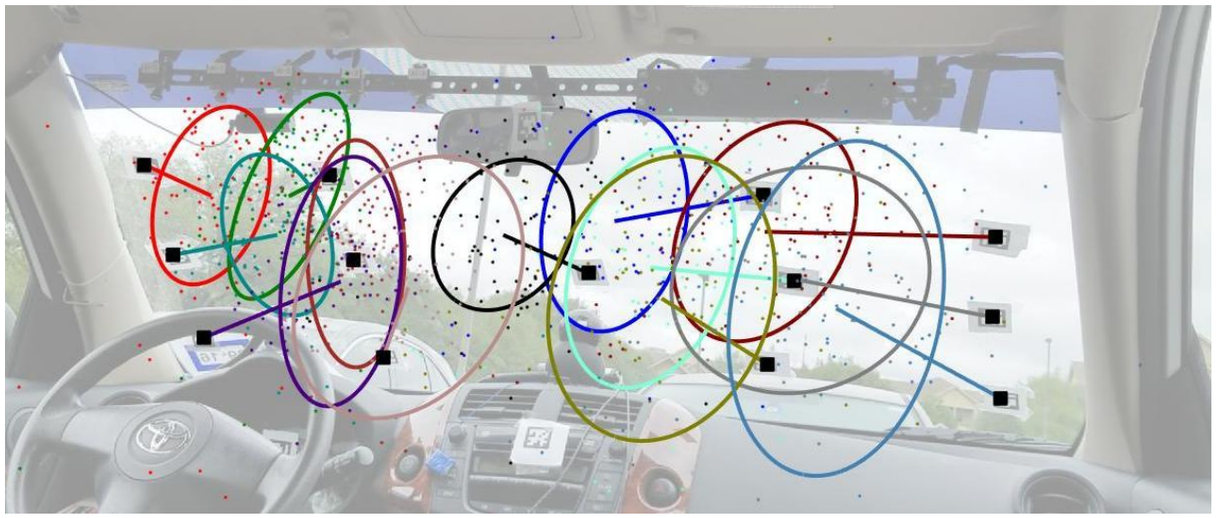
- Data collection
- Performance of head pose estimation
- Gaze estimation using linear regression
- **Study of eye movement bias**
  - **Question 3: How much does the head pose varies when the driver is looking at a certain direction?**
- Conclusion

Study of Eye Movement Bias

- Projected the head direction on the windshield
- Ellipse representing the standard deviation of the head pose
- Distance between the ellipse and the gaze point is the average bias due to the eye movement

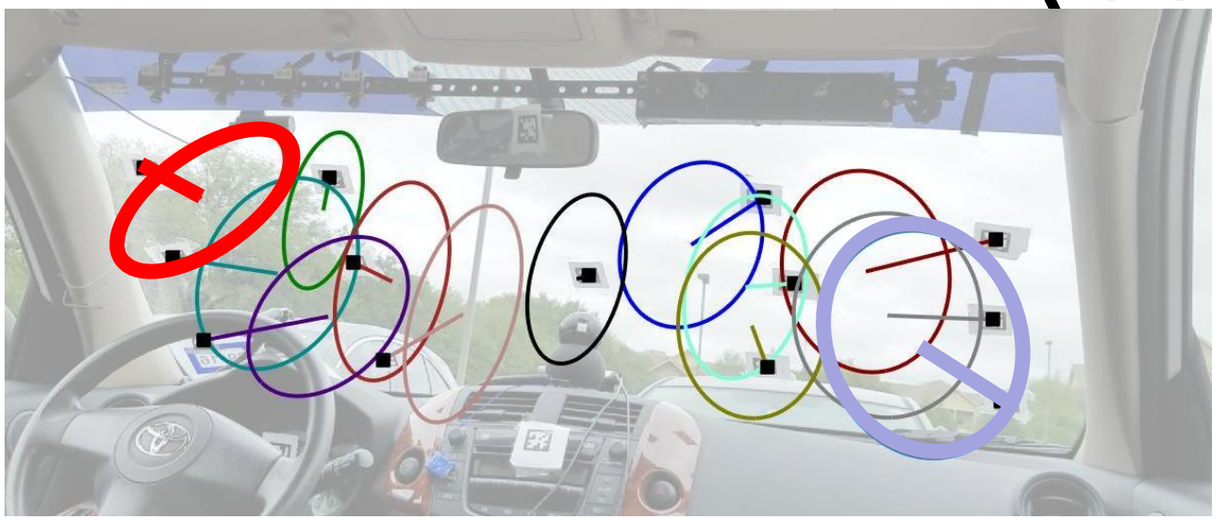


Phase 1 (Parked)

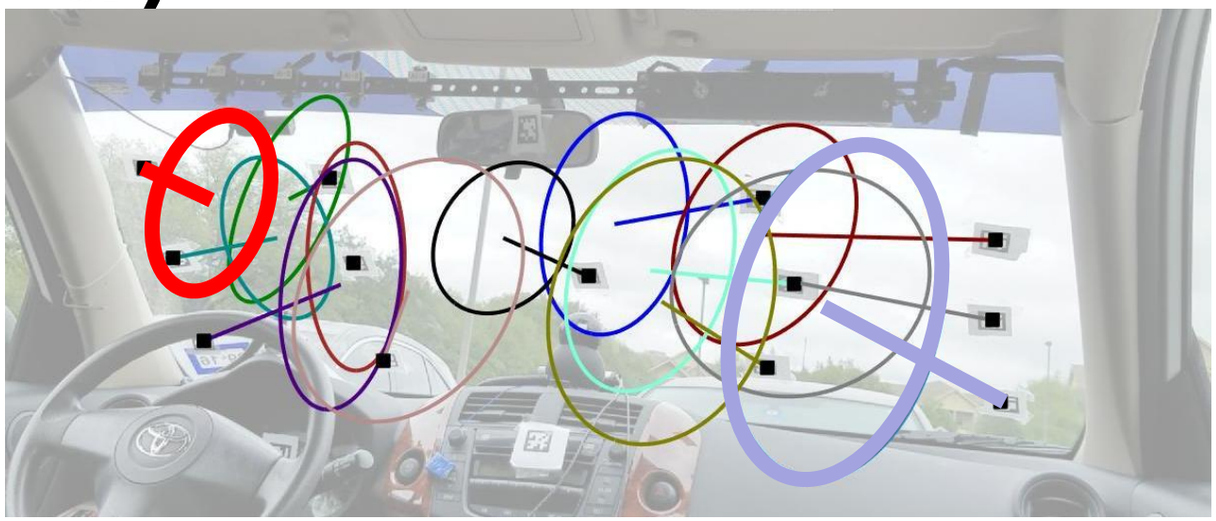


Phase 2 (Driving)

Study of Eye Movement Bias (cont.)



Phase 1 (Parked)



Phase 2 (Driving)

- Observations
  - More variance (hence less predictability) when driving
  - More variance when looking away from the front.
  - The bias increases as the direction moves away from the frontal pose

Conclusions

- How well can we estimate the head pose in a real world driving environment?
  - At high yaw angles detection rate goes down
  - At high pitch angles the difference between the angles goes up
- How well does the head pose of the driver predict his/her gaze (visual attention)?
  - While there is strong correlation (horizontal direction) a deterministic model may not be possible
- How much does the head pose varies when the driver is looking at a certain direction?
  - Variation in head pose and the bias due to eye movement increases when looking further away from the front.

Thank you!

Questions?

msp.utdallas.edu

Nanxiang Li and Carlos Busso, "Detecting drivers' mirror-checking actions and its application to maneuver and secondary task recognition," IEEE Transactions on Intelligent Transportation Systems 17 (4), 980-992.

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

Baltrusaitis, T., P. Robinson, and L.-P. Morency (2013, December). Constrained local neural fields for robust facial landmark detection in the wild. In-Proceedings of the IEEE International Conference on Computer Vision Workshops, Sydney, Australia, pp. 354-361. IEEE.