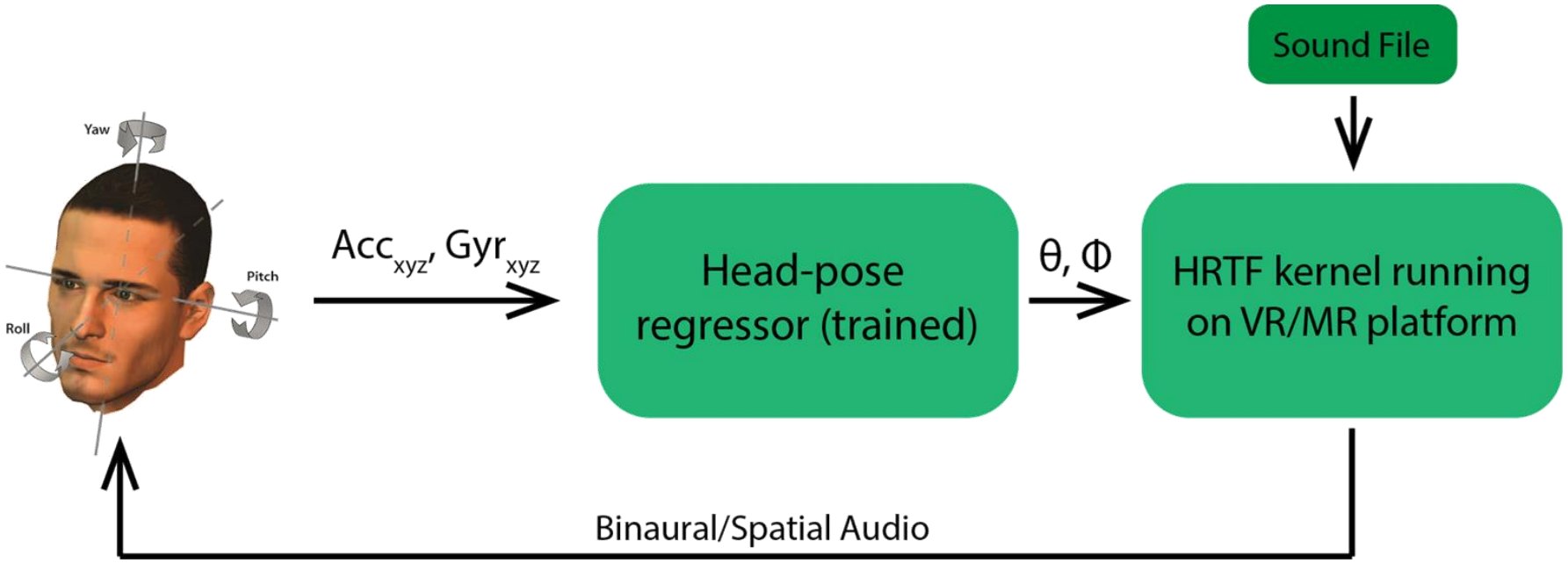
ECE M202A (11/25/2019) Mid Term Presentation

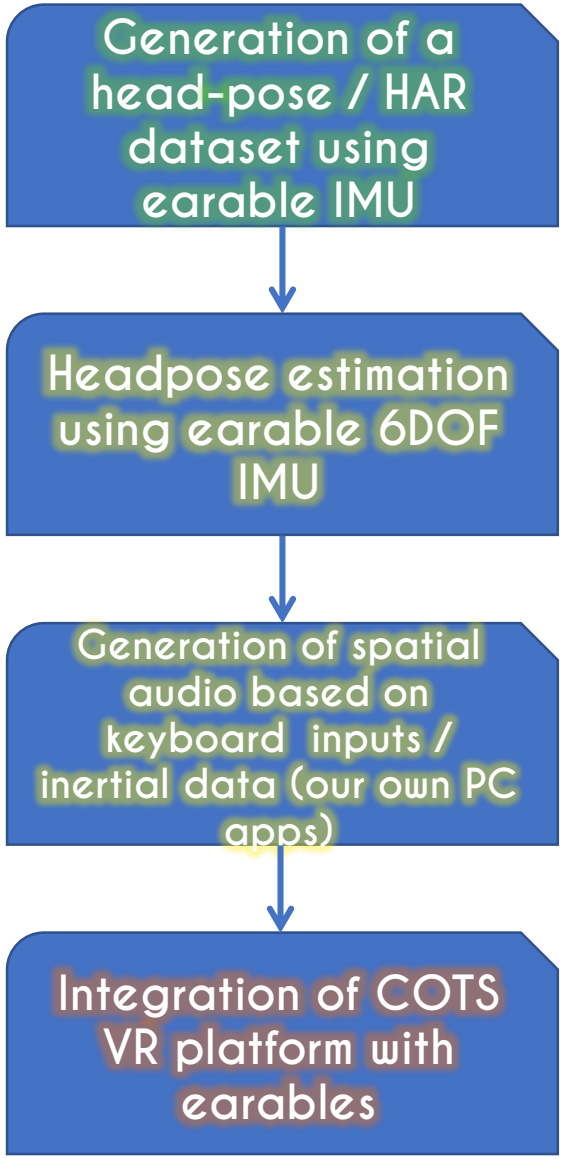
MREarable: Integrating Spatial Audio in a Mixed Reality Environment through Earable Sensor Modalities

Team Members: **Swapnil Sayan Saha, Vivek Jain and Siyou Pei**

Presenter: **Swapnil Sayan Saha**

Overall Project Goals and Specific Aims

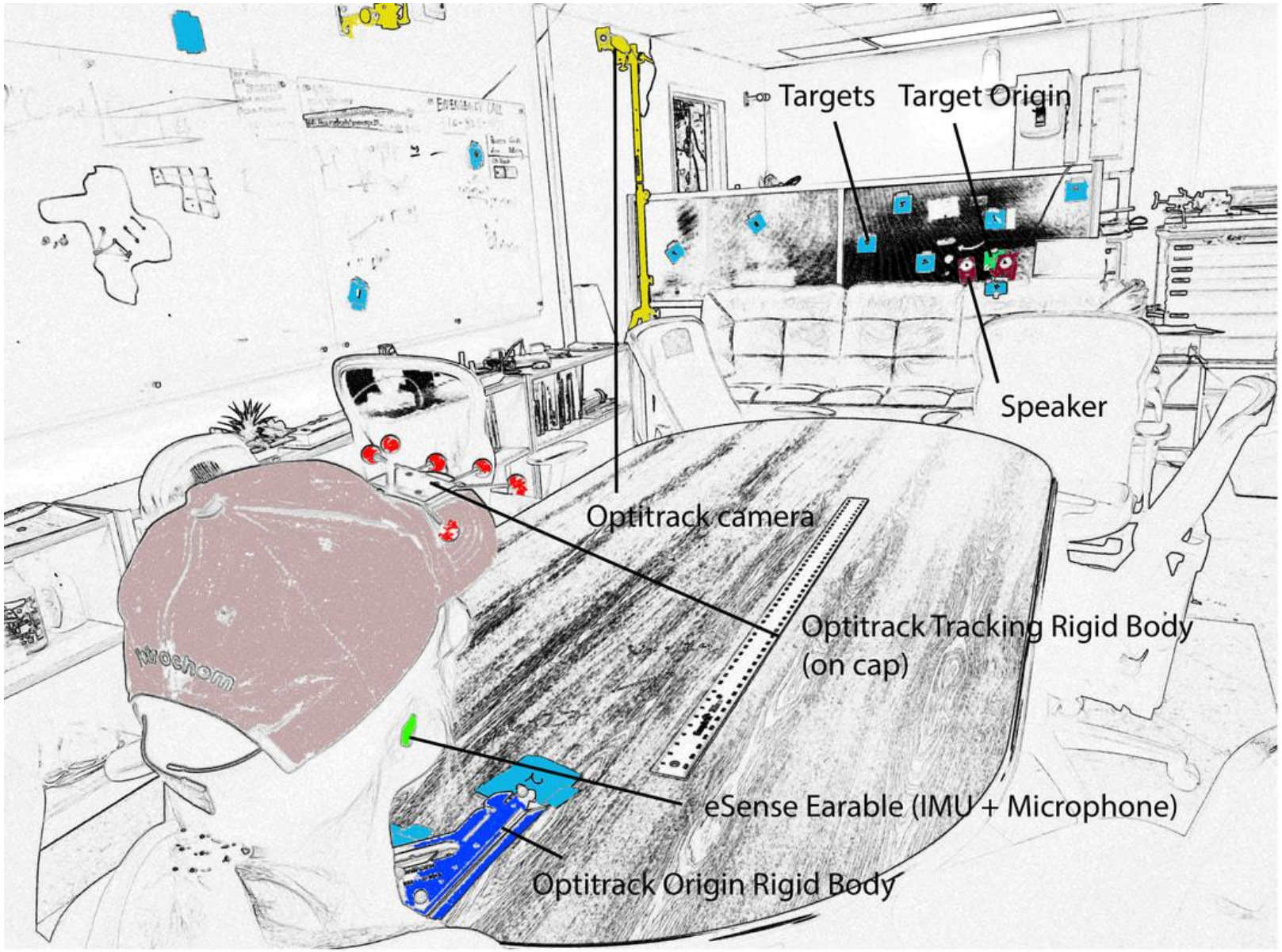




Deliverables:

- Use trained head-pose estimator to generate spatial audio from a stationary sound source in real time (ongoing)
- Integration of VR platform with eSense earables for perception-processing-feedback loop.

Technical Approach



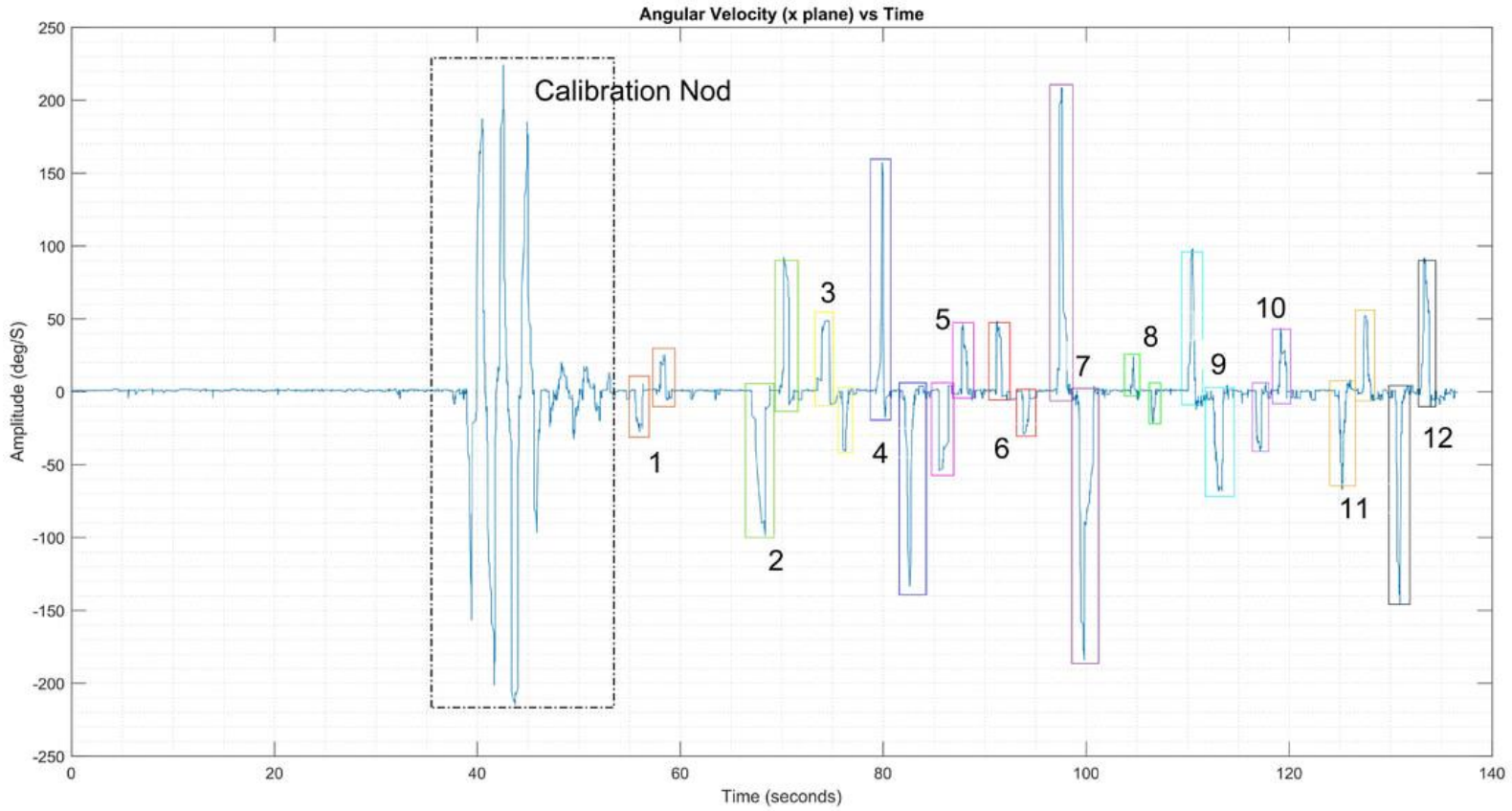
Two types of head-motion: 1. O-T-O 2. O-T1-T2-O

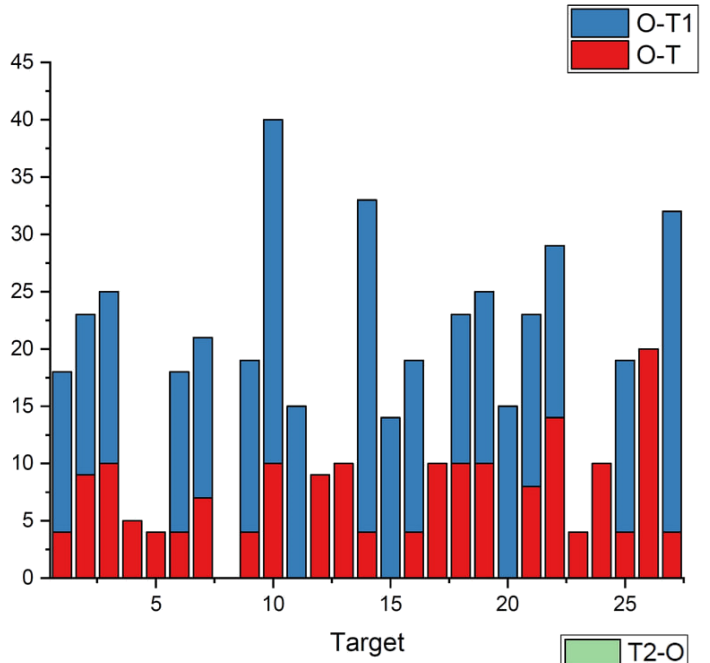
27 targets; 34 (13 + 21) distinct head-poses; 15 subjects; 9 activities

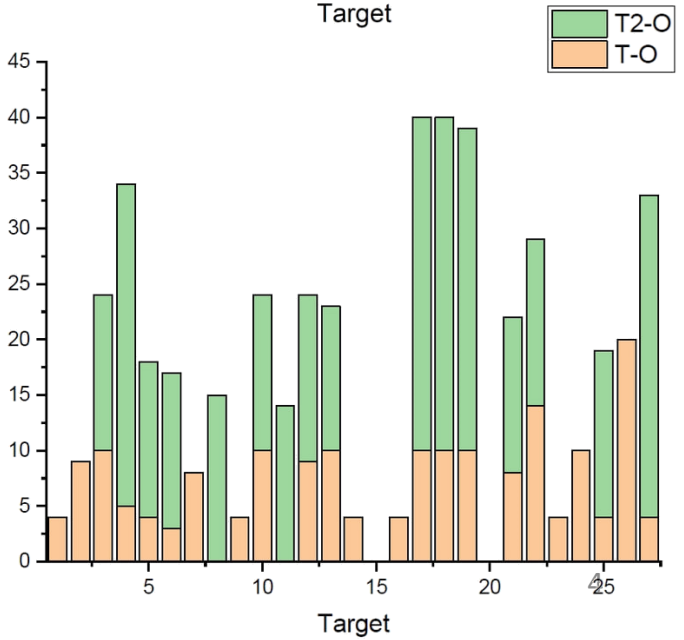
Range: +- 2g and +- 500 deg/S; Adv. & Conn. Intv: 45-55 & 20-30 mS; F_s = 100 Hz, LPF: 5 Hz.

Ground truth / Measurements: Optitrack, earable audio, Leica Disto X3 & video camera

Technical Approach (cont.)







1266 inertial frames: 356 O-T-O, 607 O-T1 and T2-O, 303 T1-T2

Two types of head-motion: 1. O-T-O 2. O-T1-T2-O

Technical Approach (cont.)

Extracted 37 features

$Integration\_Acc_x$	$Integration\_Acc_y$	$Integration\_Acc_z$
$Integration\_Gyr_x$	$Integration\_Gyr_y$	$Integration\_Gyr_z$
$Variation\_Acc_x$	$Variation\_Acc_y$	$Variation\_Acc_z$
$Variation\_Gyr_x$	$Variation\_Gyr_y$	$Variation\_Gyr_z$
$Skew\_Acc_x$	$Skew\_Acc_y$	$Skew\_Acc_z$
$Skew\_Gyr_x$	$Skew\_Gyr_y$	$Skew\_Gyr_z$
$Kurtosis\_Acc_x$	$Kurtosis\_Acc_y$	$Kurtosis\_Acc_z$
$Kurtosis\_Gyr_x$	$Kurtosis\_Gyr_y$	$Kurtosis\_Gyr_z$
$Norm\_Acc_x$	$Norm\_Acc_y$	$Norm\_Acc_z$
$Norm\_Gyr_x$	$Norm\_Gyr_y$	$Norm\_Gyr_z$
$SumOT\_Acc_x$	$SumOT\_Acc_x$	$SumOT\_Acc_x$
$SumOT\_Gyr_x$	$SumOT\_Gyr_x$	$SumOT\_Gyr_x$
TimeWindow		

734 out of 1266 frames grouped into bins of 20 degrees (noisy data)

Elevation: [-140, 180]
Azimuth: [-140, 140]

Validation: 3-fold CV, Holdout ratio: 90:10

Classifier – CNN: 32%, DNN: 40.5% (only on elevation)

Classifier - XGBoost: 74.57% (42.3% without grid search hyperparameter optimization) (only on elevation)

Technical Approach (cont.)

Inputs to HRTF: ϴ, Ø and sound file

Output: Spatial Audio

VR Integration

User can locate sound source with eyes closed!

Demo: Head movement simulation via keyboard

Effects: Binaural, Doppler

unity

SAMSUNG Gear VR Powered by oculus

Current Status and Next Steps

Current Status	Dataset has been built, cleaned, sliced and published.
	Preliminary training on elevation classifier is complete.
	Software developed to mine eSense data on computer (and track head-pose using complementary Kalman Filter).
	Developed a spatial sound test system based on the 3D kart template on Unity platform (inputs: keyboard)

Next Steps	Complete training of elevation and azimuth classifier with all data; training on HAR dataset.
	Improve classification performance of head-pose estimator
	Integration of head-pose estimator with COTS VR platform (Unity).

THANK YOU