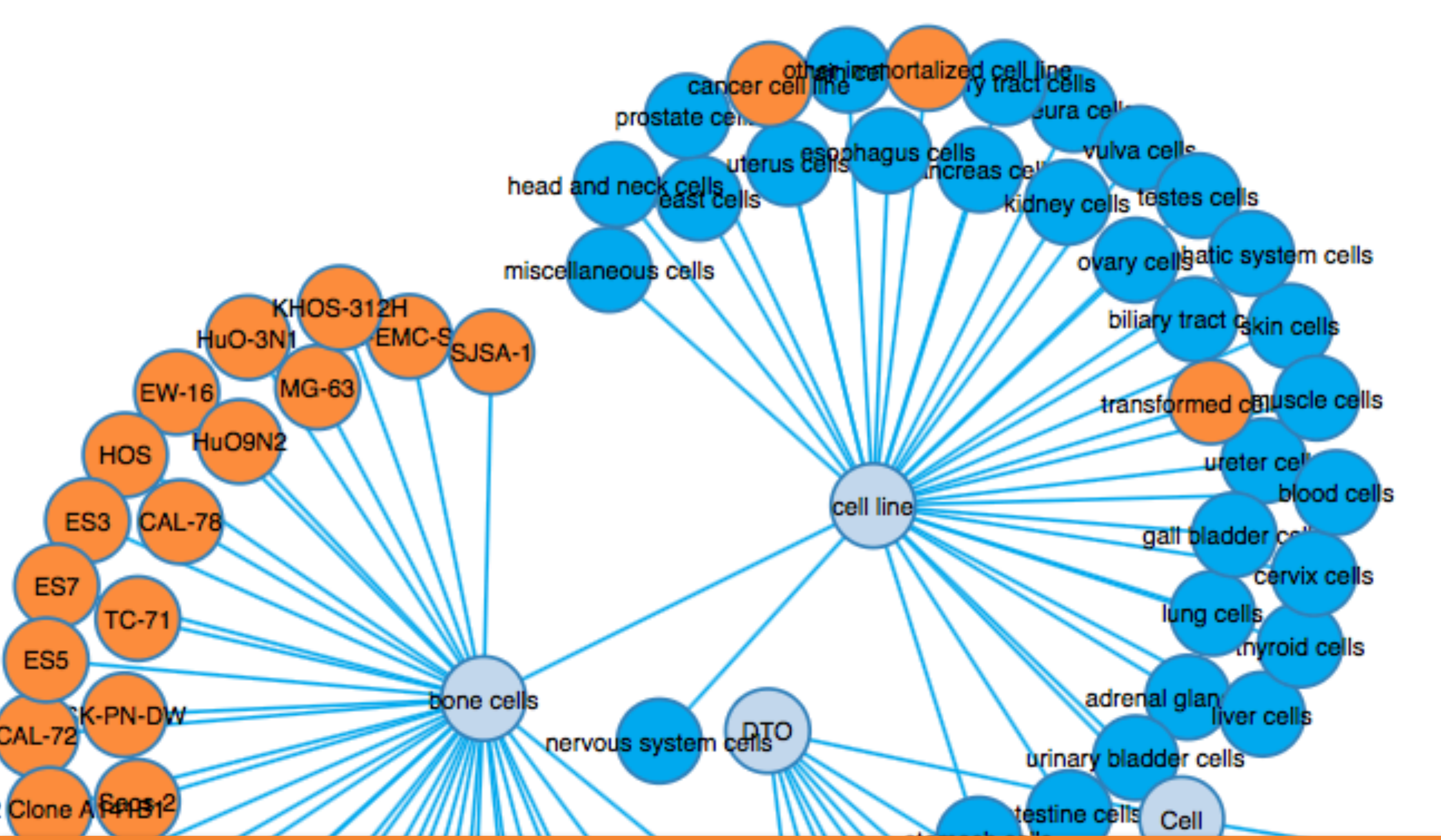
Artificial Intelligence and Robotics for Intelligent Agents

RoboCanes-Lab

Department of Computer Science
College of Arts and Sciences
University of Miami

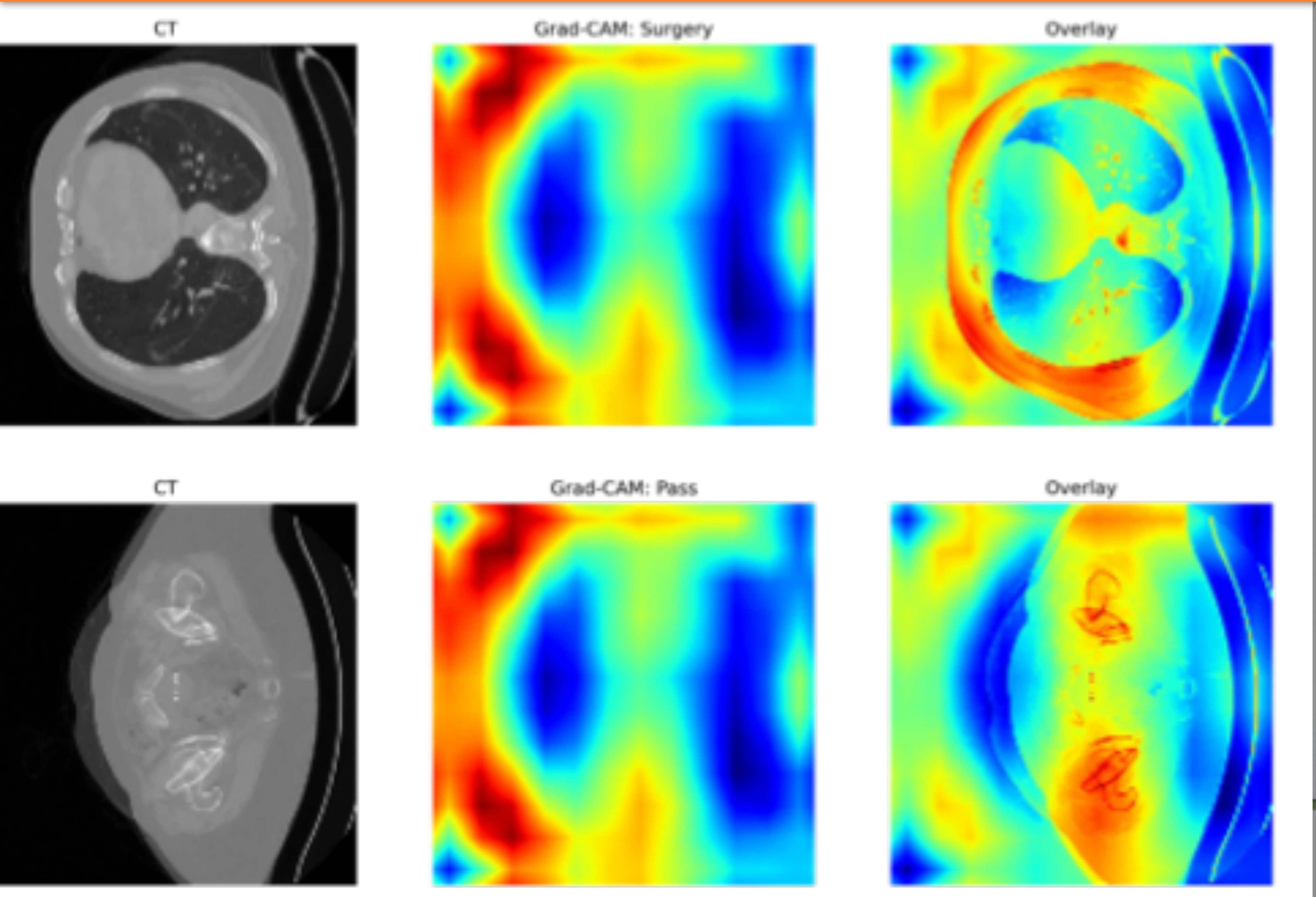
October 2022

UNIVERSITY OF MIAMI
ROBOCANES



Ontologies, Knowledge Engineering, Semantic Web, Spatio-temporal Reasoning

Machine Learning, Medical Imaging, Digital Health, Therapy, FRT



Drive interdisciplinary AI & Robotics research to enable the next generation of robots that work and interact safely alongside humans.

RESEARCH CHALLENGES

- Autonomous robots shall act appropriately in dynamic, real time, and adversarial environments
- One of the biggest challenges in AI and robotics
- Numerous examples: rescue scenarios, home assistance, device assistance (e.g. cars, planes)
- Playing soccer with biped robots as a testbed for development in perception, multi-agent cooperation, complex motions, …
- Service robots for domestic environments
- RoboCup is a landmark project as well as a standard problem

RoboCanes Soccer, 5 NAO robots

RoboCanes@Home, HSR

WIDE RANGE OF RESEARCH CHALLENGES

real time sensor fusion

grasping and manipulation

context recognition

learning

real time planning

multi-agent systems

decision making

human-robot interaction

opponent/user modeling

motor control

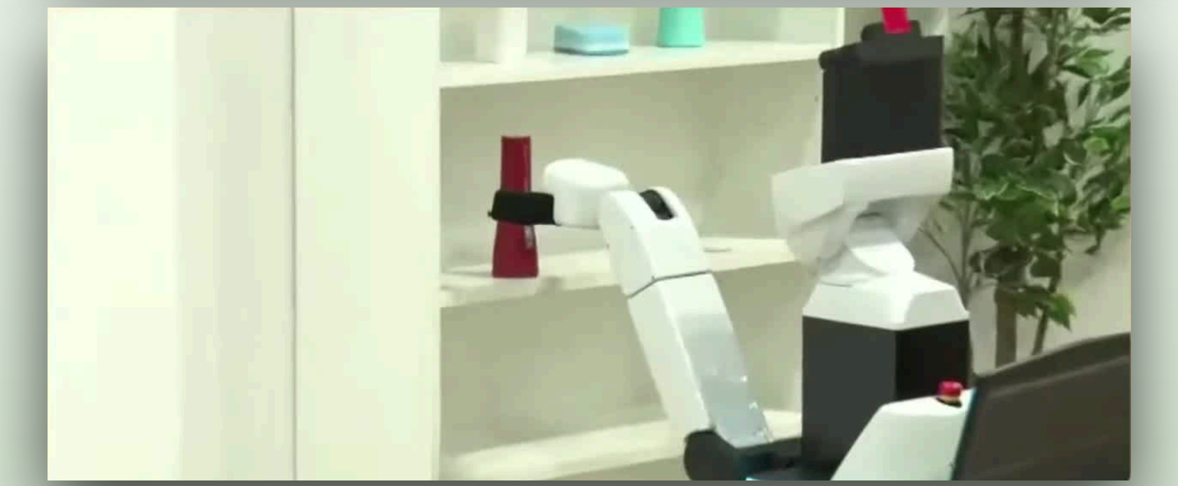
reactive/proactive behavior

and many more...

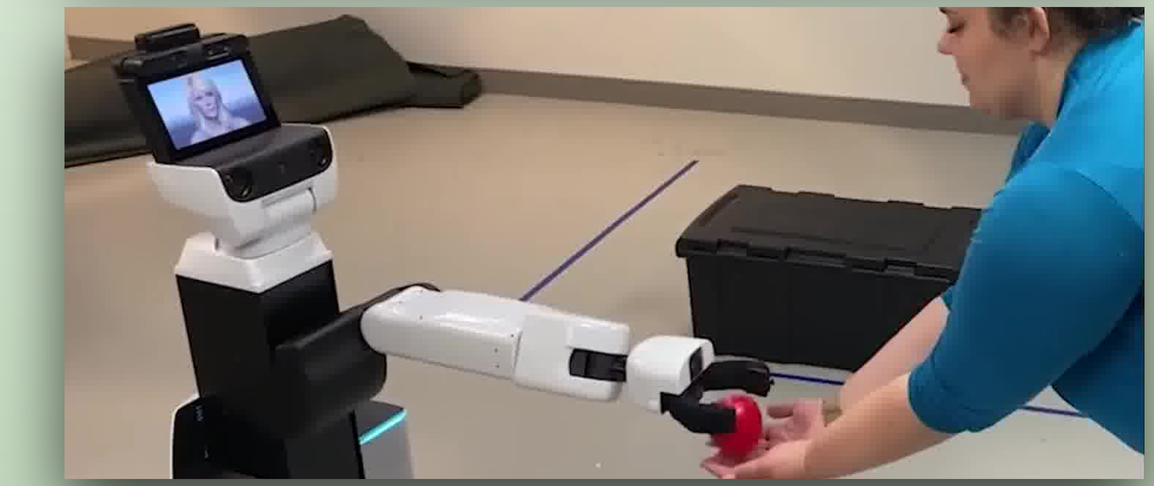
OUTLINE

1. Multi-robot cooperation and communication

2. Manipulation systems



3. Human-Robot Interaction

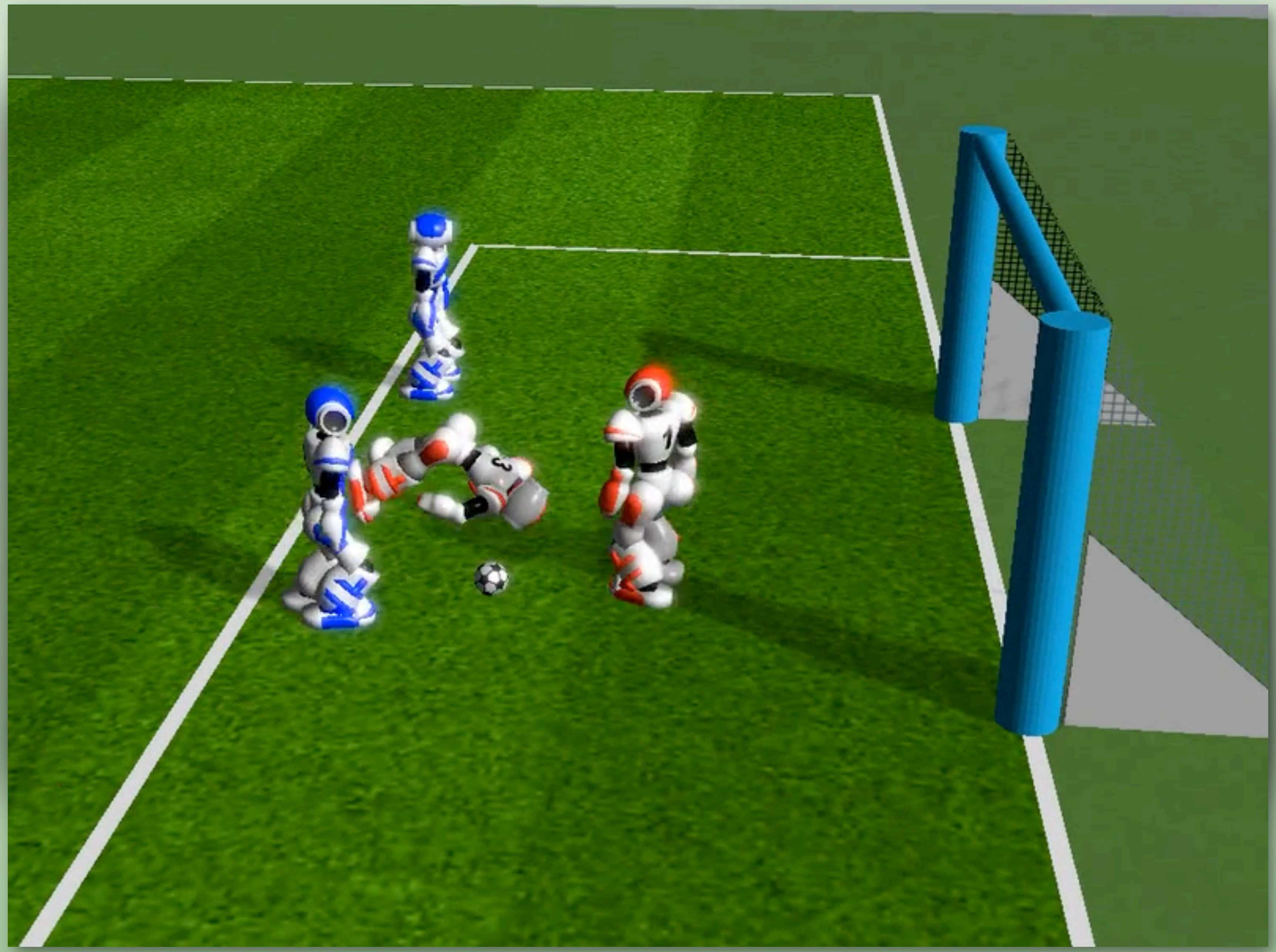


4. Interfaces: VR meets AI & Robotics

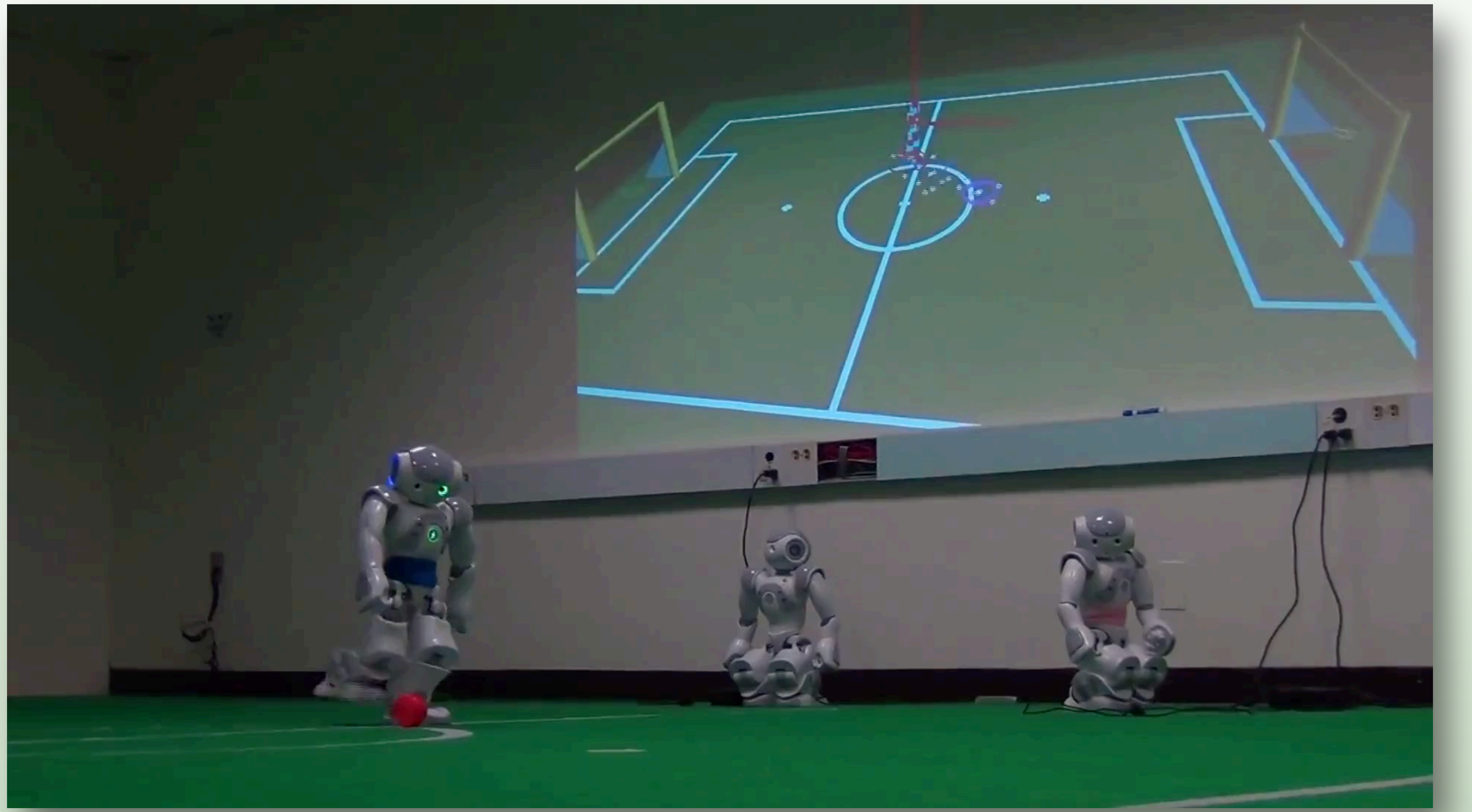
5. Facial recognition on MD criminal justice data

1. MULTI-ROBOT COOPERATION AND COMMUNICATION

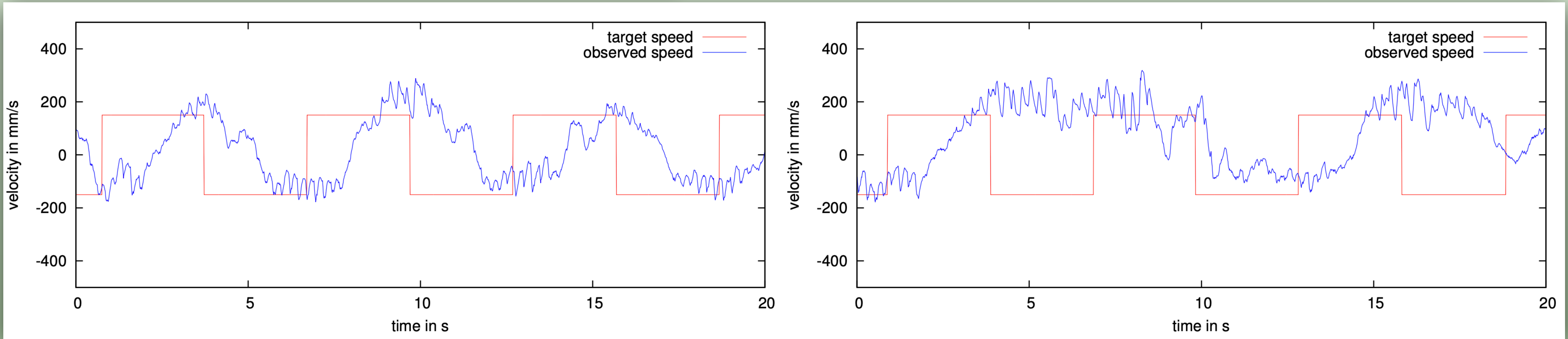
- Task and motion planning
  - Creation of high-level commands and collision-free trajectories to achieve goal
- State estimation and perception
  - Infer relevant quantities from sensor data (field objects, opponents, team mates, contacts/collisions, …)
- Communication
  - Communication with team mates, determine global geometry (e.g. ball position), ad-hoc sub-team building (e.g. offside trap, double-pass)
- Object manipulation
  - Determine good kick positions given relevant constraints (global geometry, local geometry, placement of ball)
- Trajectory generation and control
  - Real-time, reactive generation of control commands to move bipedal robot safely toward goal



CONTROL (1): DYNAMIC ADAPTIVE WALKING ENGINE



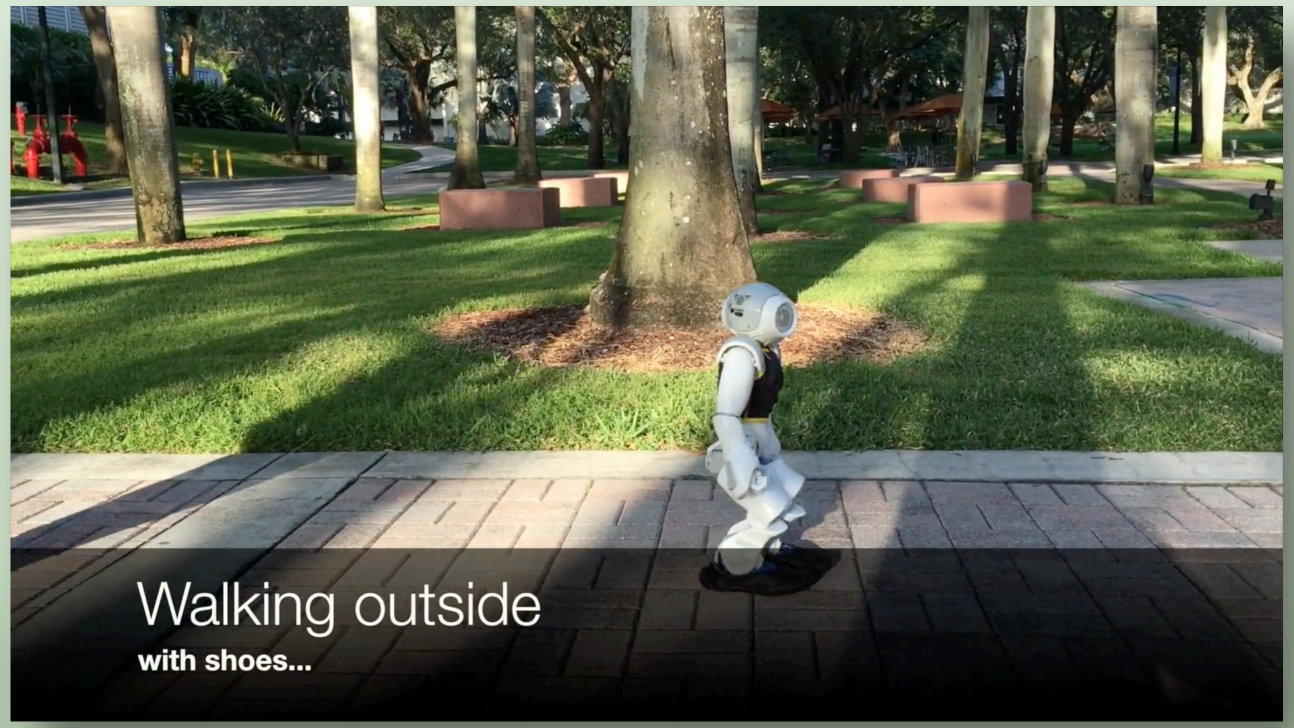
- First bipedal walk, open-loop, parameter prior optimized, no changes during runtime
- Becomes unstable due to changes (e.g. motor temperature)

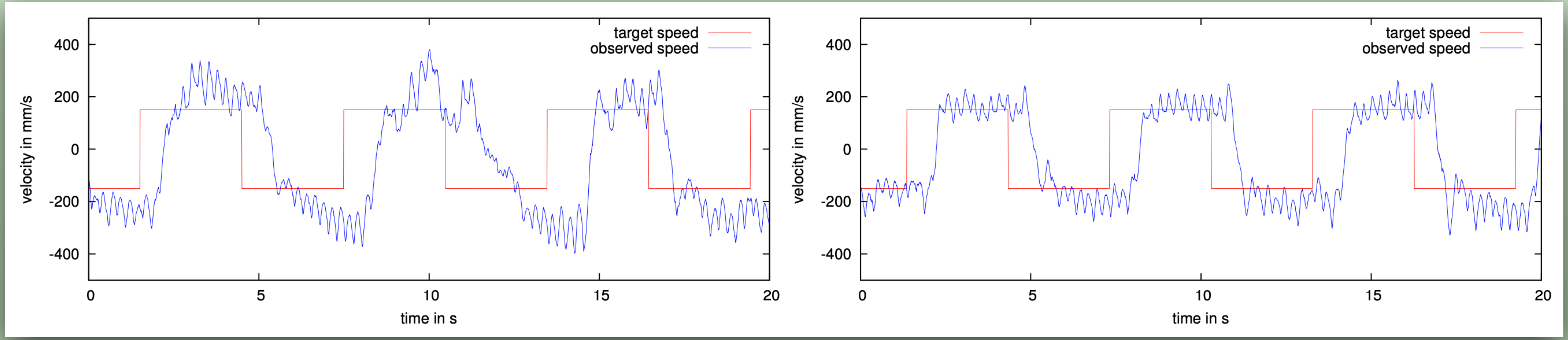


Forward/backward walk. Slow acceleration for stability (left), higher acceleration limit less stable (right)

CONTROL (1): DYNAMIC ADAPTIVE WALKING ENGINE

- LIPM-based closed-loop walk
- Adaptation by optimizing parameter of model in real-time onboard
- Result: new, stable, fast, and energy-efficient walk



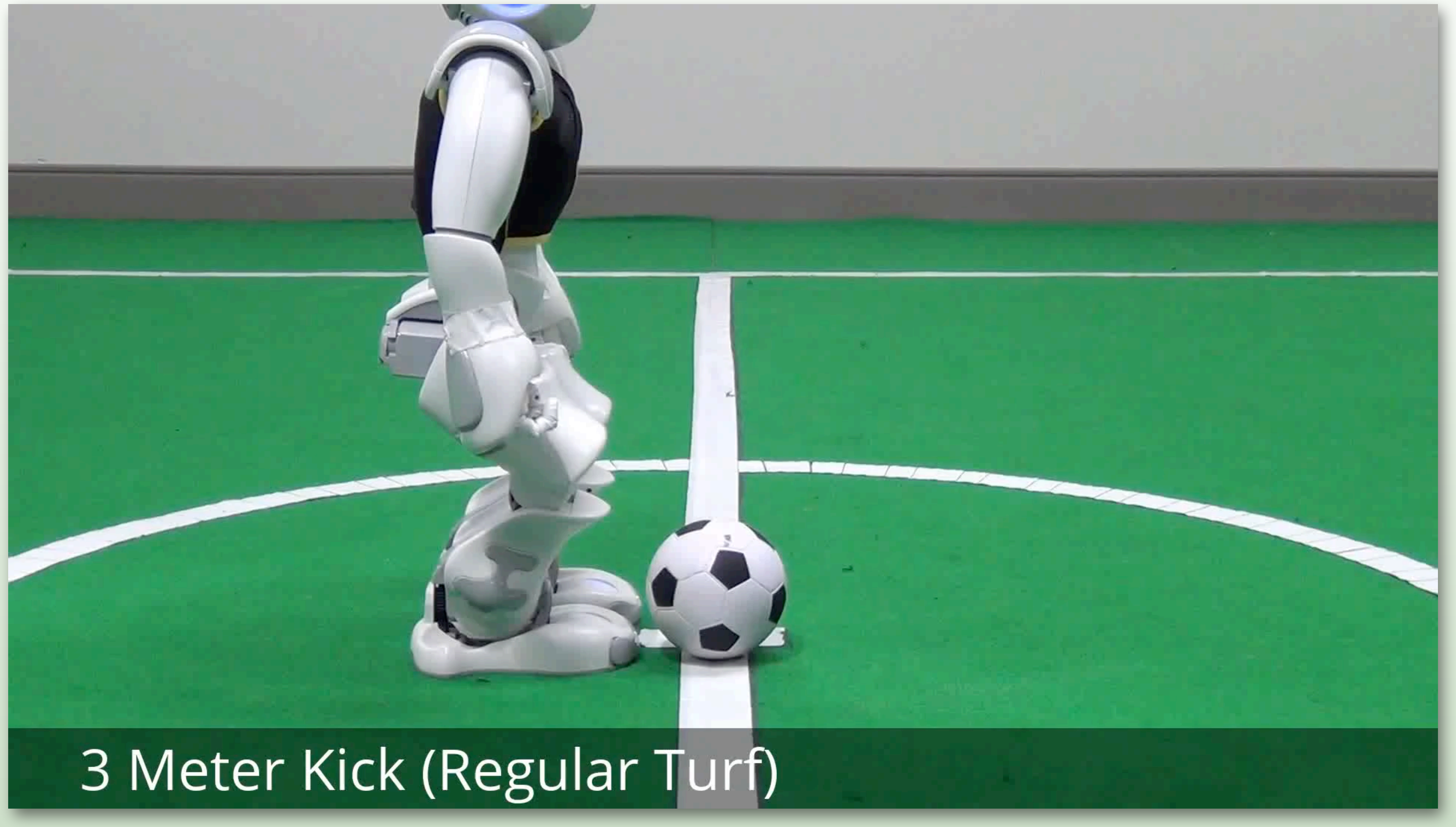


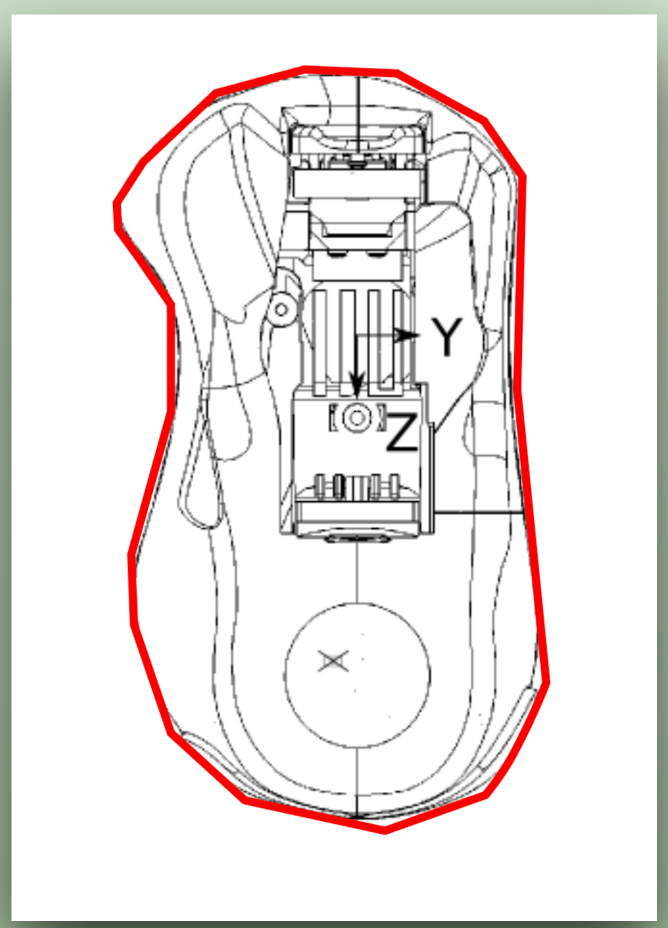
Forward/backward walk with 15cm/s with new walk. No optimization (left), with optimization (right)

Seekircher & Visser (*KI: German AI Journal, 2016*)

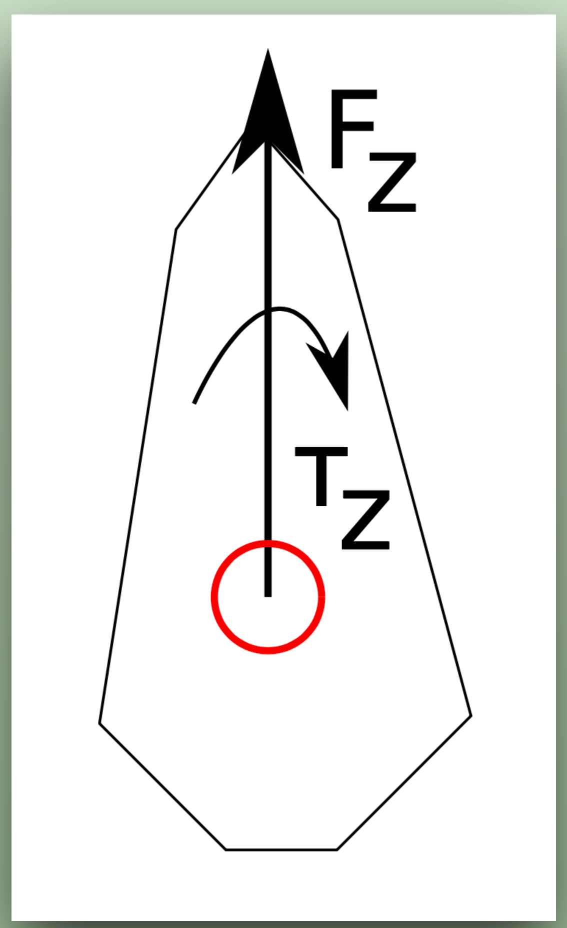
CONTROL (2): OMINDIRECTIONAL KICK

- Kick trajectory generation, arbitrary direction, no prior input or knowledge of the kick parameters
- Key idea: Zero Moment Point (ZMP) based preview controller that minimizes the ZMP error
- Covariance Matrix Adaptation Evolution Strategy (CMA-ES) for model optimization.

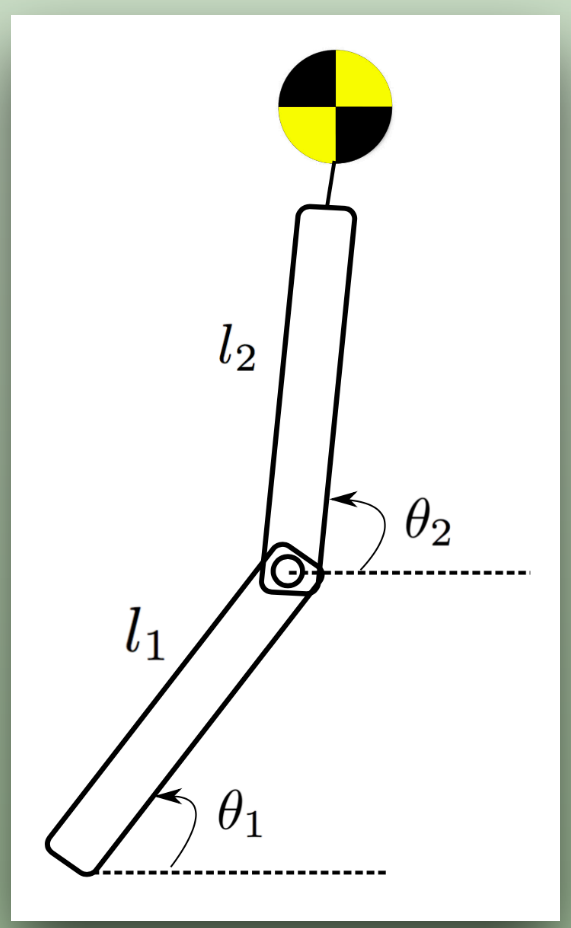




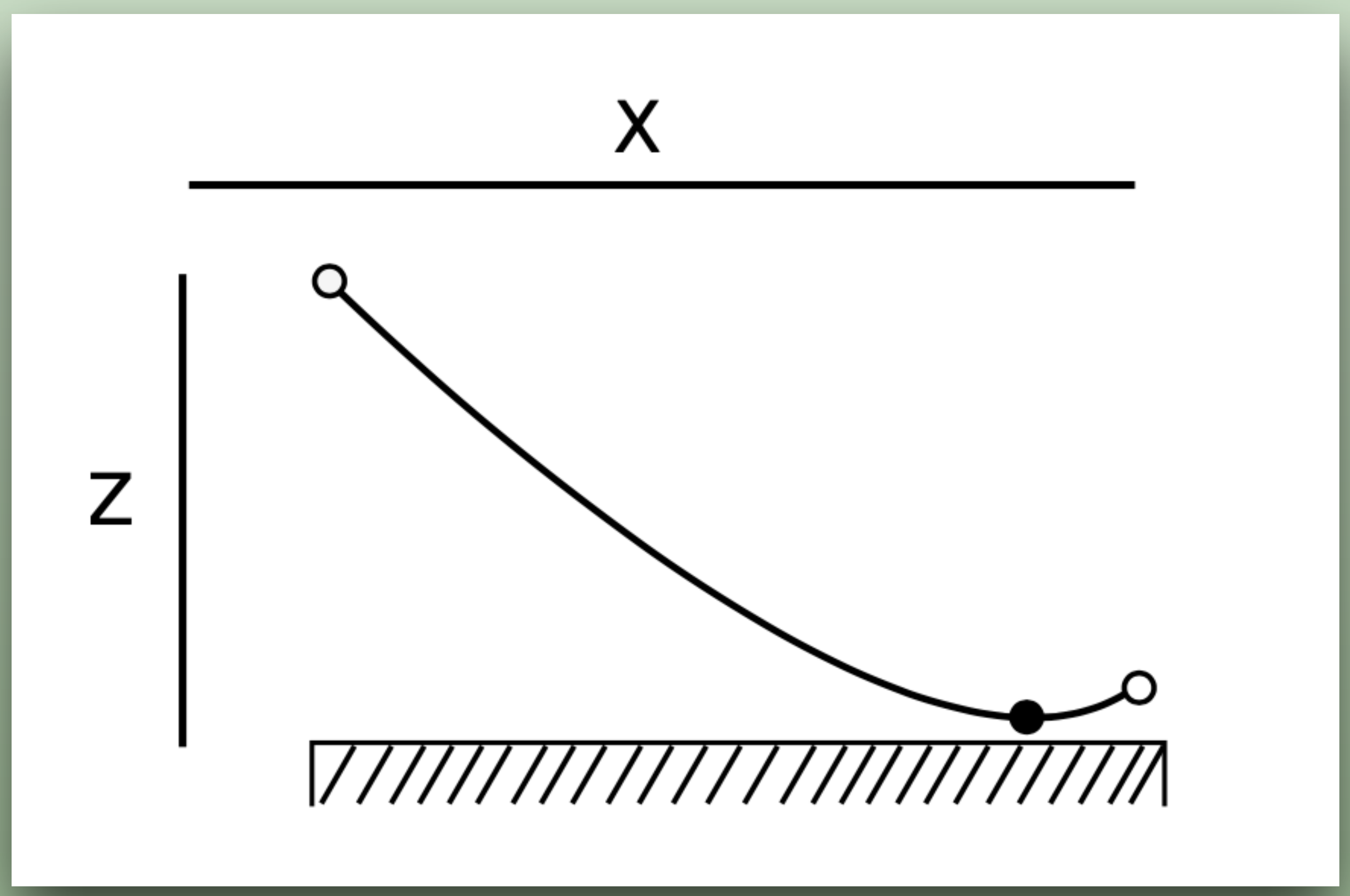
Support polygon (convex hull)



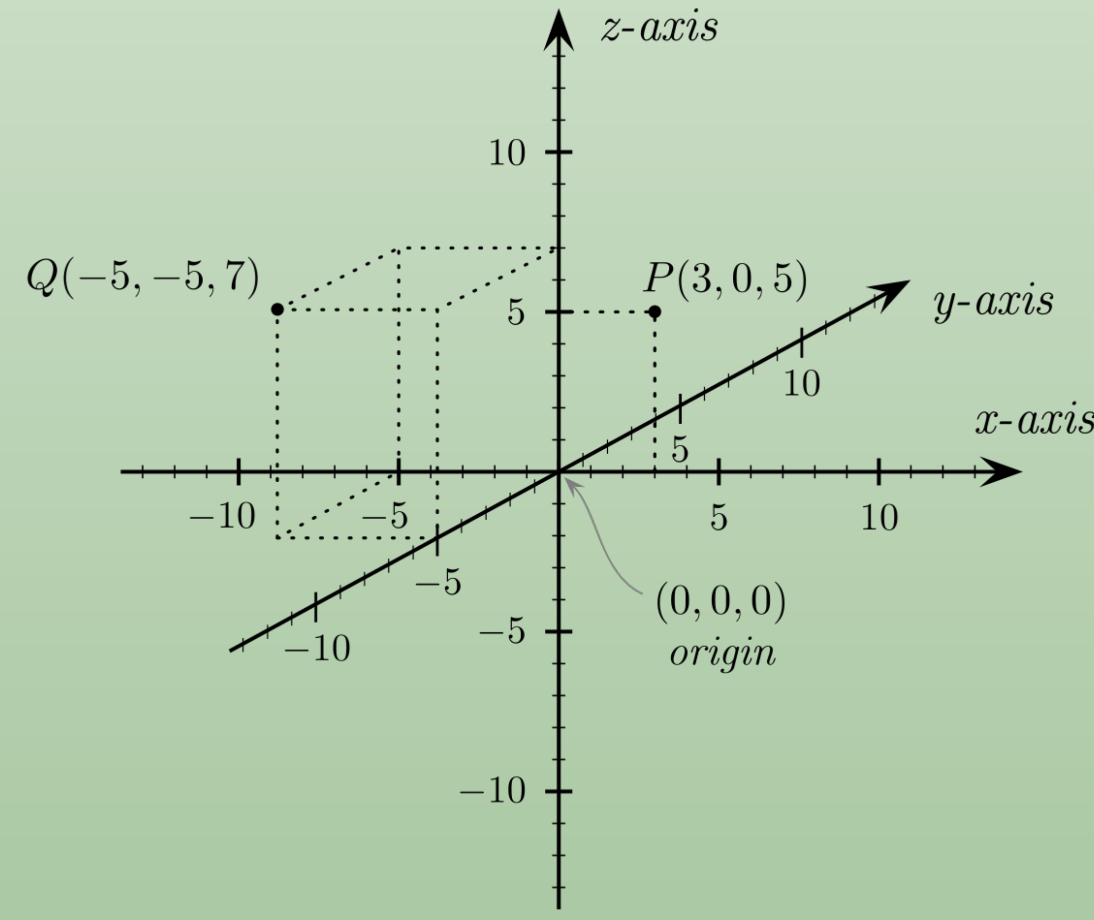
ZMP location



Approximation of inverse kinematic model



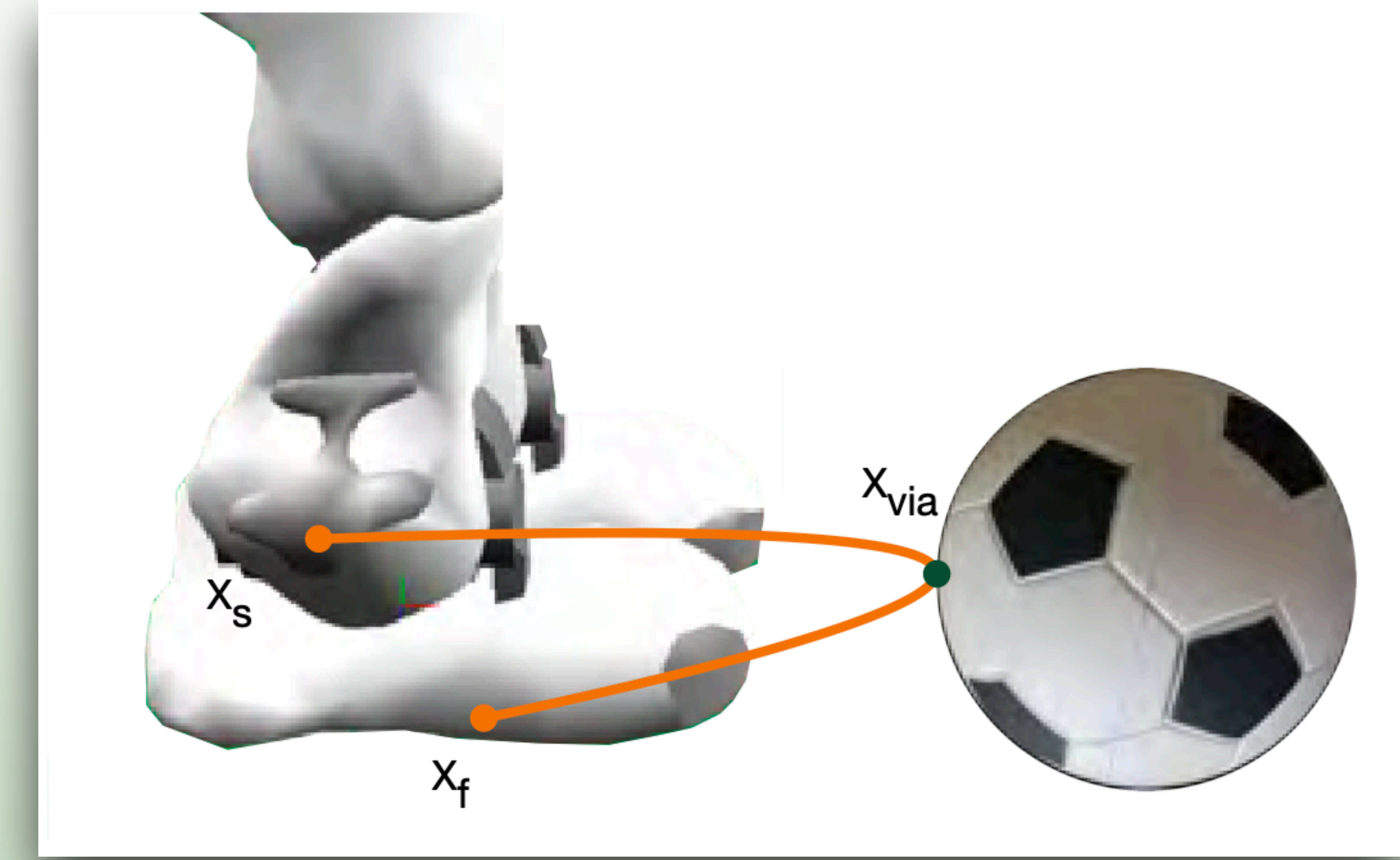
Two polynomials for front kick



Peña, Masterjohn & Visser (*RoboCup Symposium 2017*)

CONTROL (3): WALK-KICK ENGINE

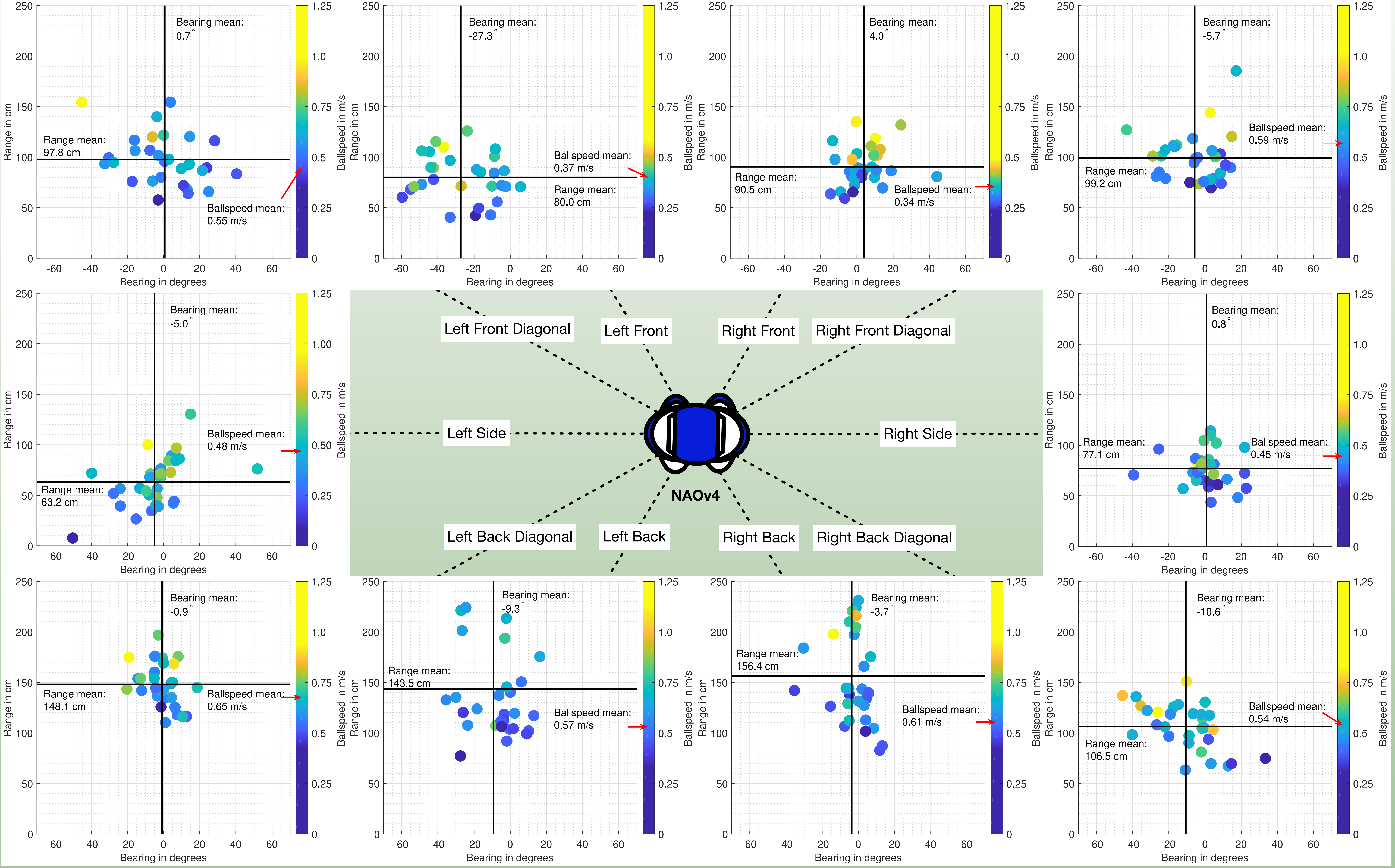
- Walk-kick frame-work generating a kick trajectory
  - any direction, no prior input, while walking
  - guaranteeing reaching a reference trajectory
- Uses kick interpolators from **dynamic kick engine** and walk trajectories generated from **adaptive walking engine** to generate motions
- Reliable in terms of the kick directions and stability of the robot overall (< 1% falling rate)
- Experiments verify that the walk-kick trajectories were consistent with an average absolute bearing of < 6° within any given direction.



Walk-kick motion where the via point is the point of contact with the ball.

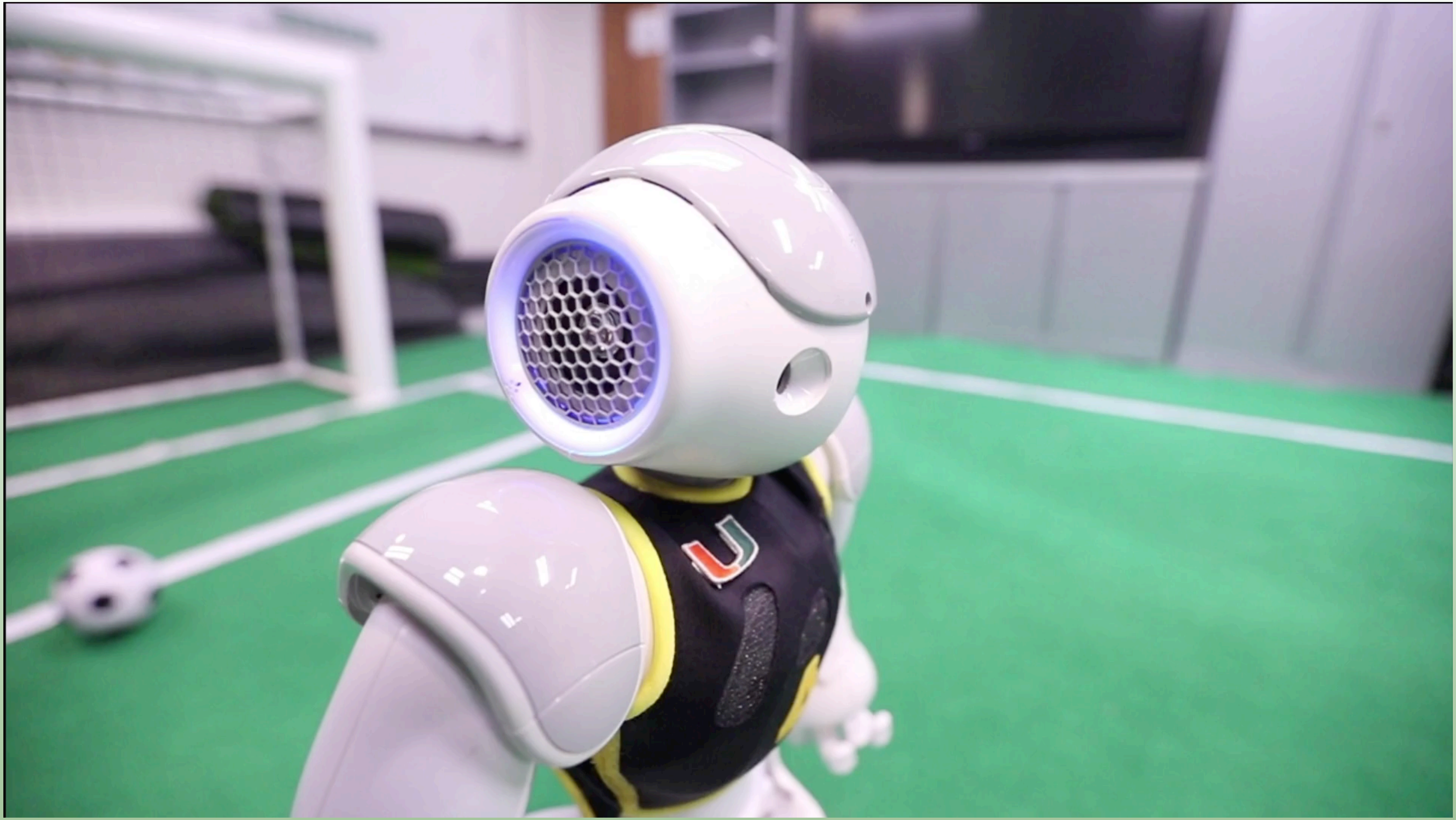
Peña & Visser (*RoboCup Symposium 2018)*

CONTROL (3): WALK-KICK ENGINE



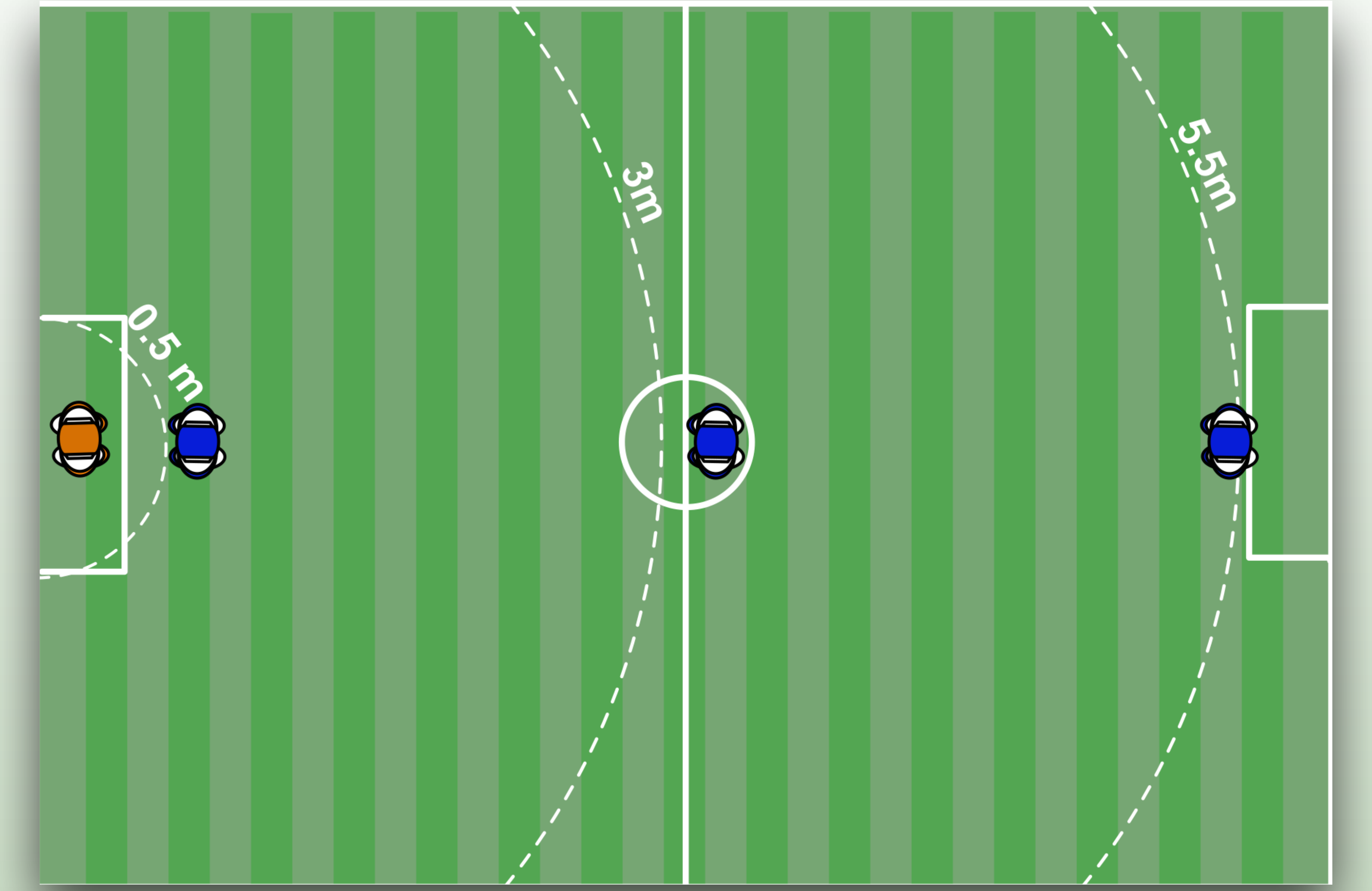
Peña & Visser (*RoboCup Symposium 2018*)

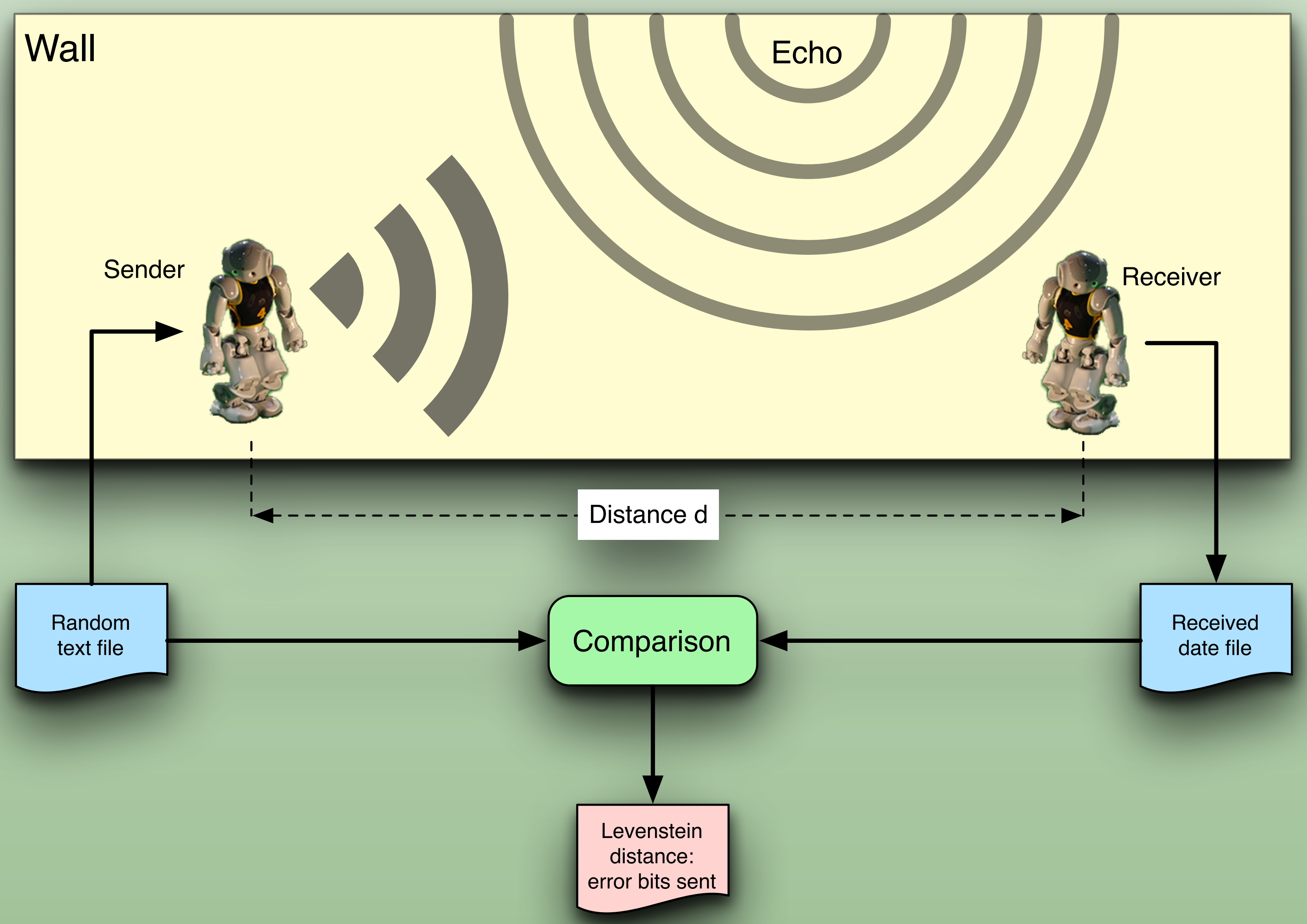
CONTROL (3): WALK-KICK ENGINE



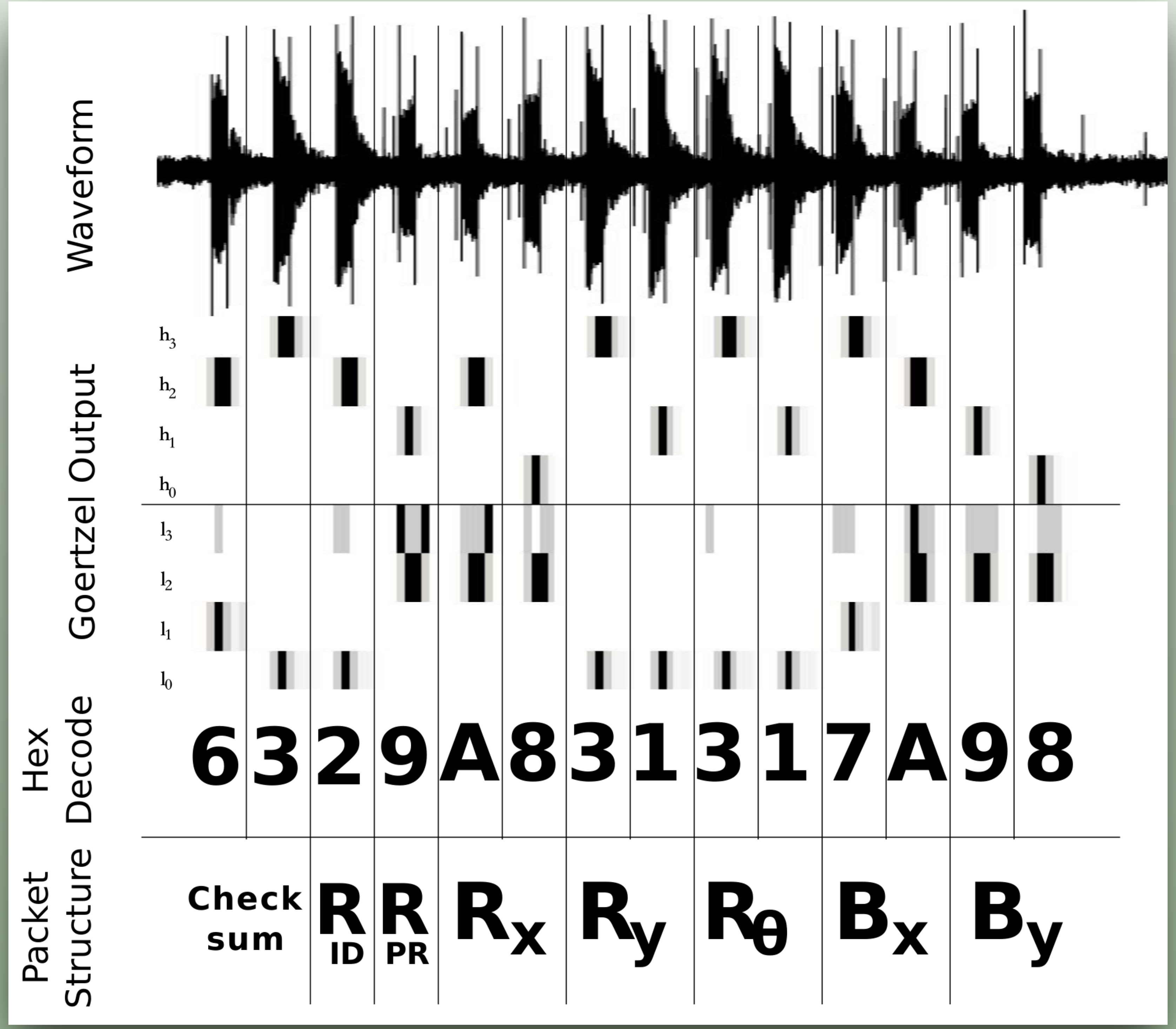
Peña & Visser (*RoboCup Symposium 2018*)

COMMUNICATION (1): DTMF BROADCASTING





- Multi-agent broadcasting based on fixed length DTMF messages



Poore et al. 2015 (*FLAIRS*)

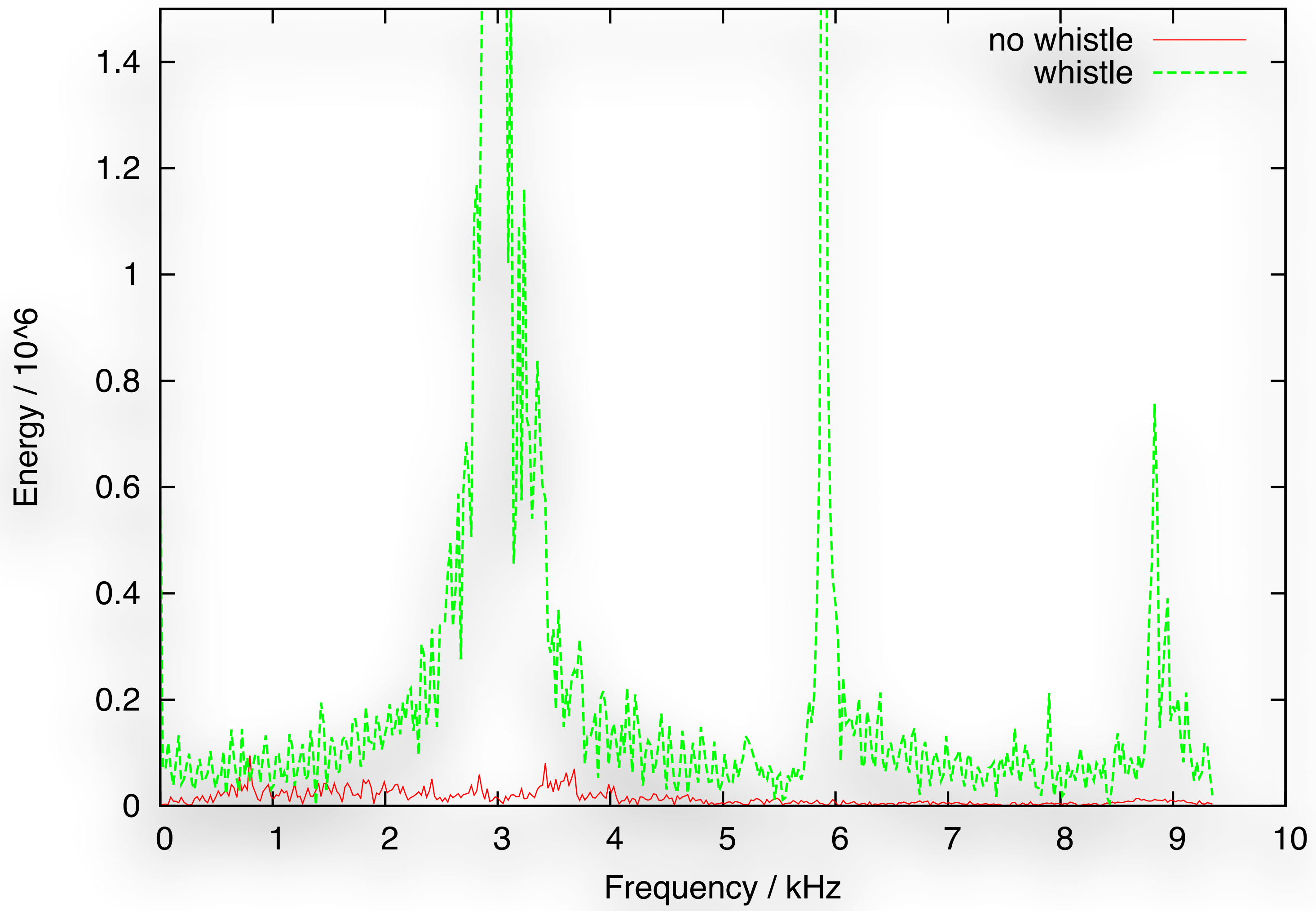
COMMUNICATION (1): DTMF BROADCASTING

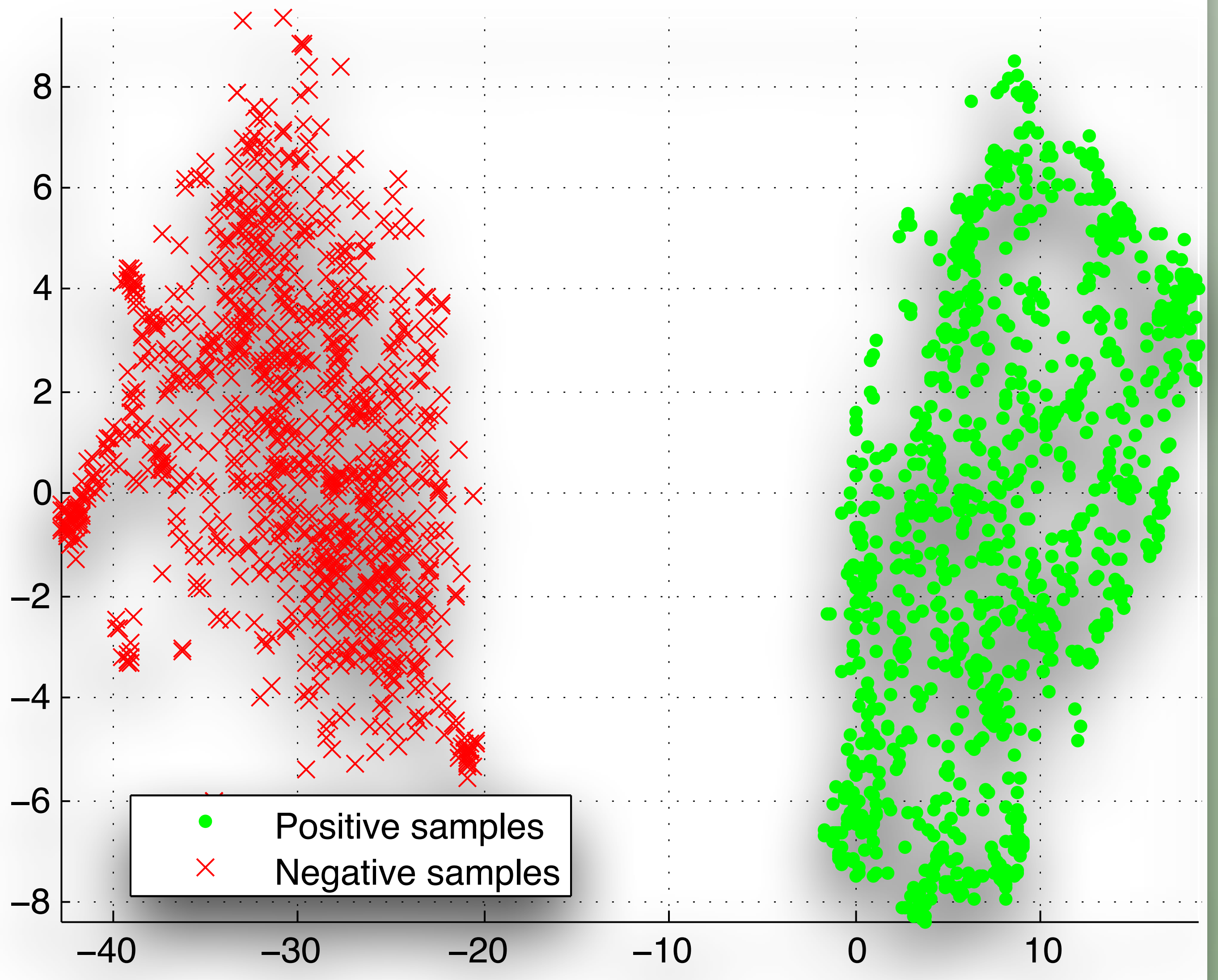
- Multi-agent broadcasting based on fixed length DTMF messages

Poore et al. 2015 (*FLAIRS*)

COMMUNICATION (2): WHISTLE DETECTION

- Based on Logistic Regression with L2-norm Regularization

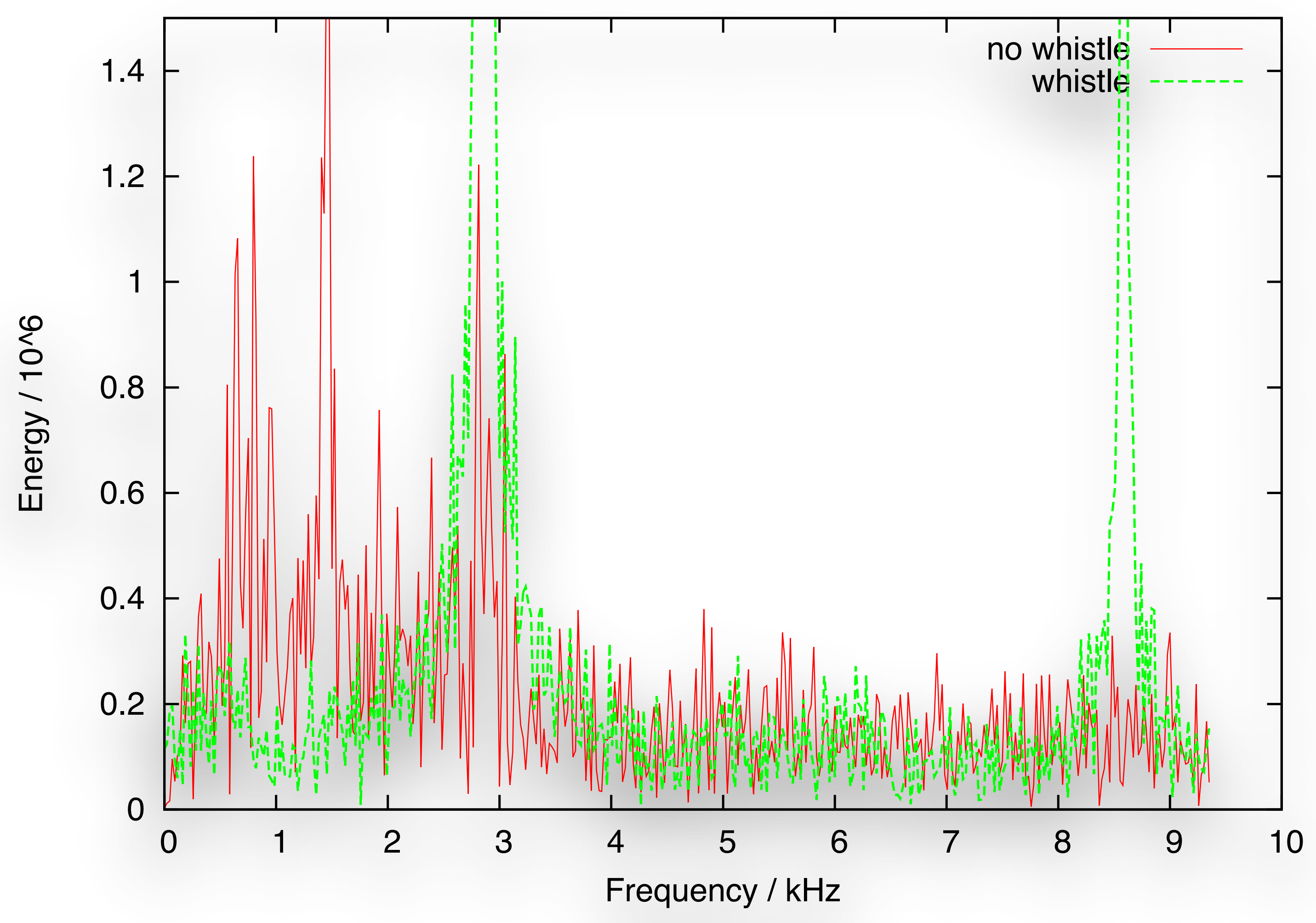


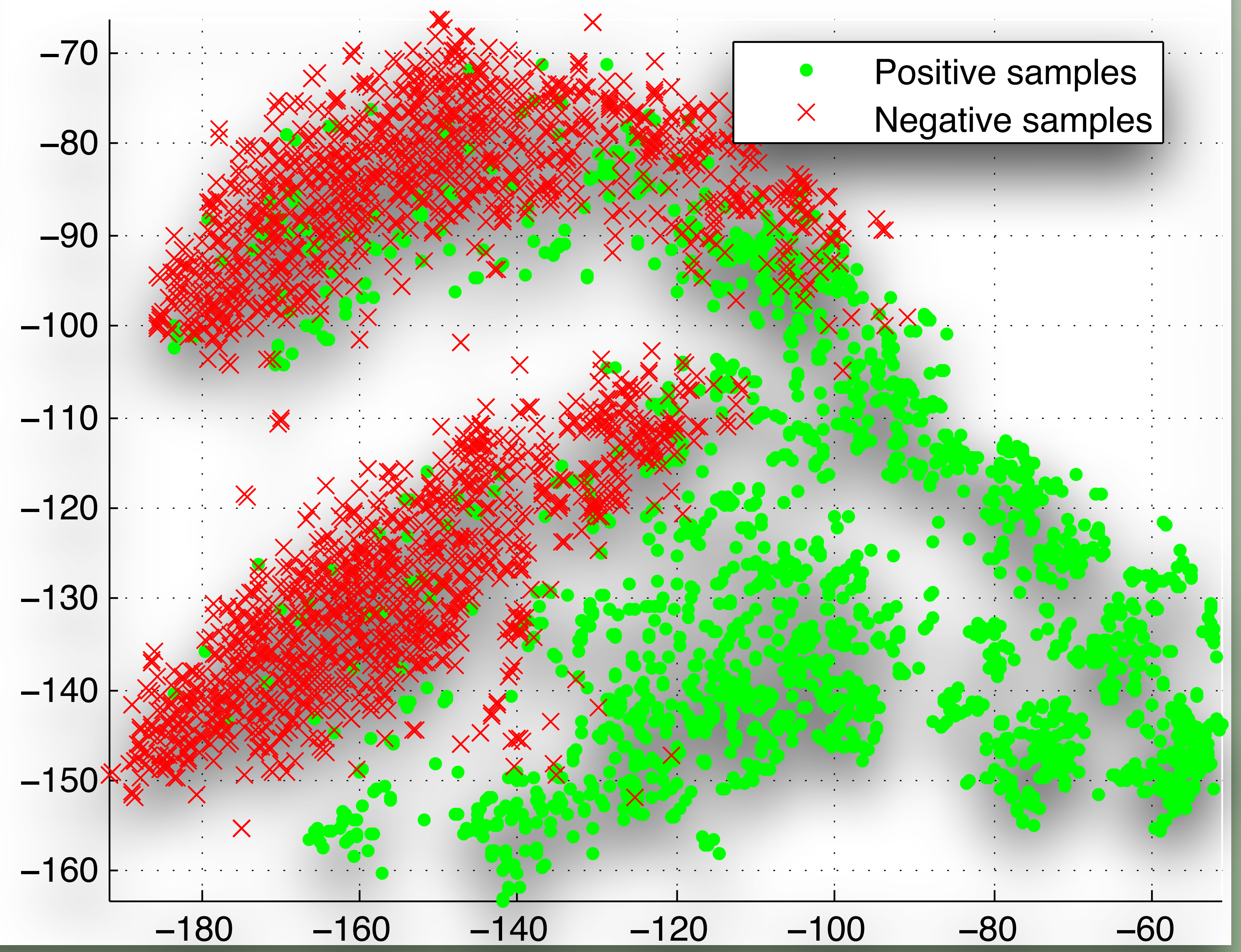


Poore et al. 2014 (*RoboCup Symposium*)

COMMUNICATION (2): WHISTLE DETECTION

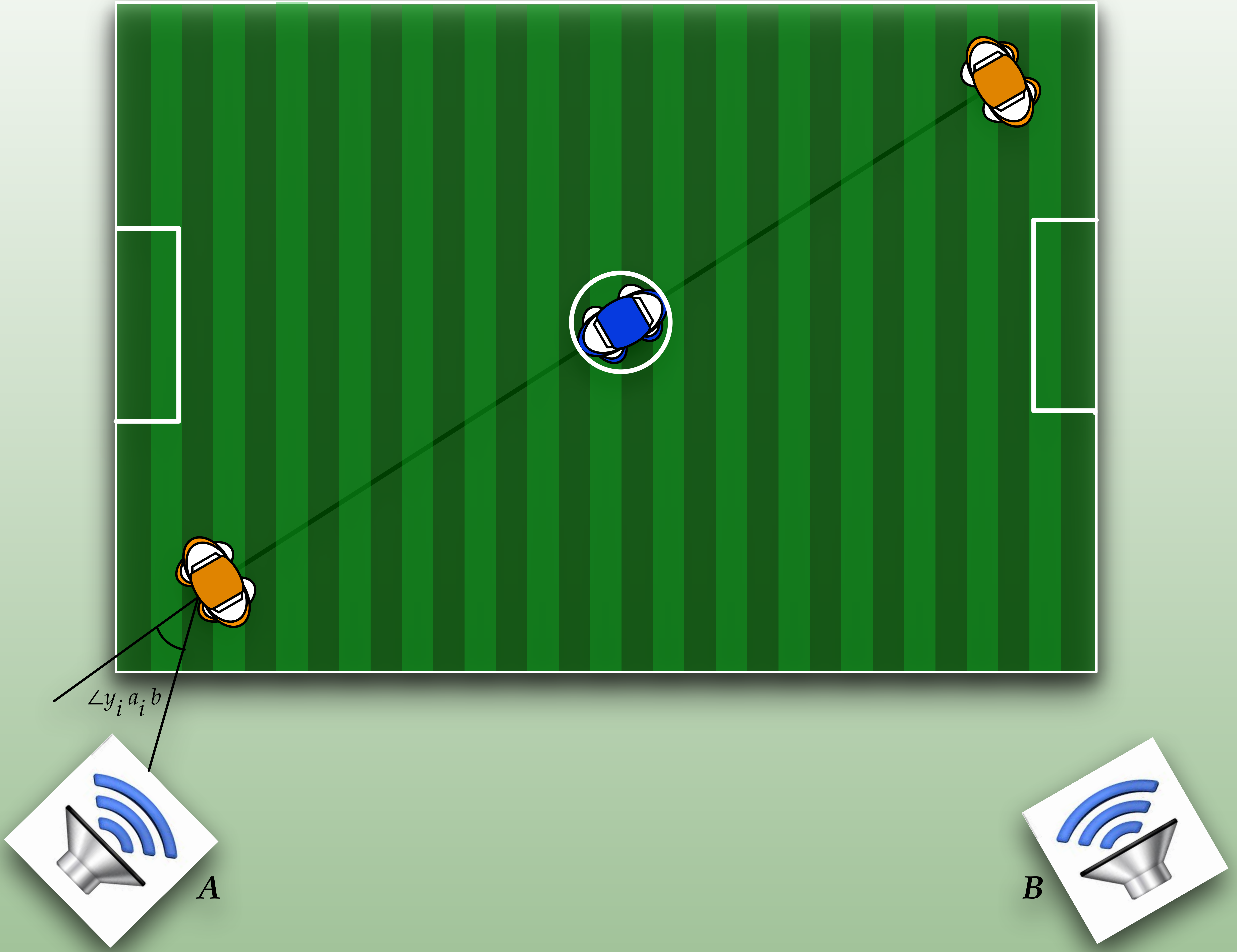
- Based on Logistic Regression with L2-norm Regularization





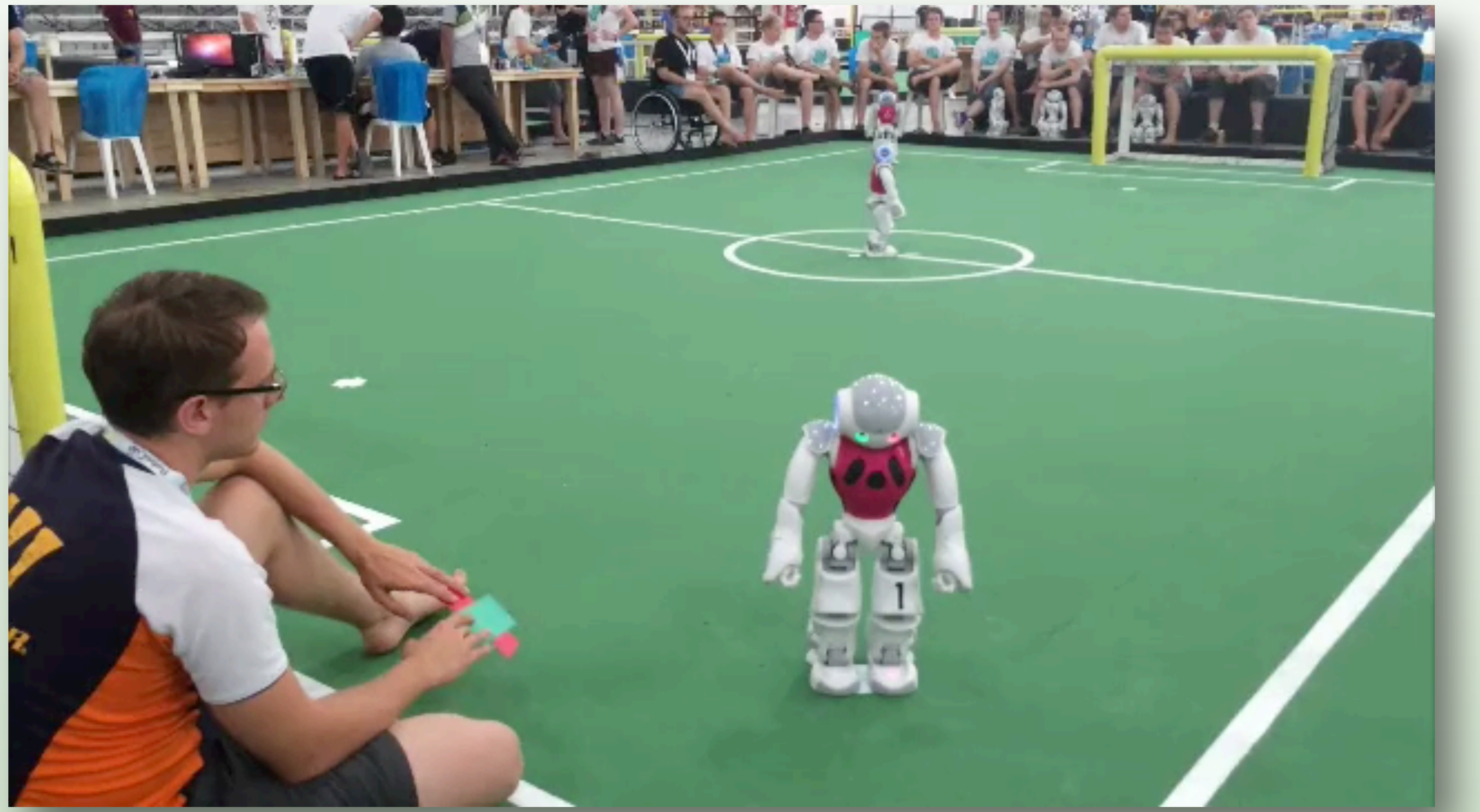
Poore et al. 2014 (*RoboCup Symposium)*

COMMUNICATION (2): WHISTLE DETECTION



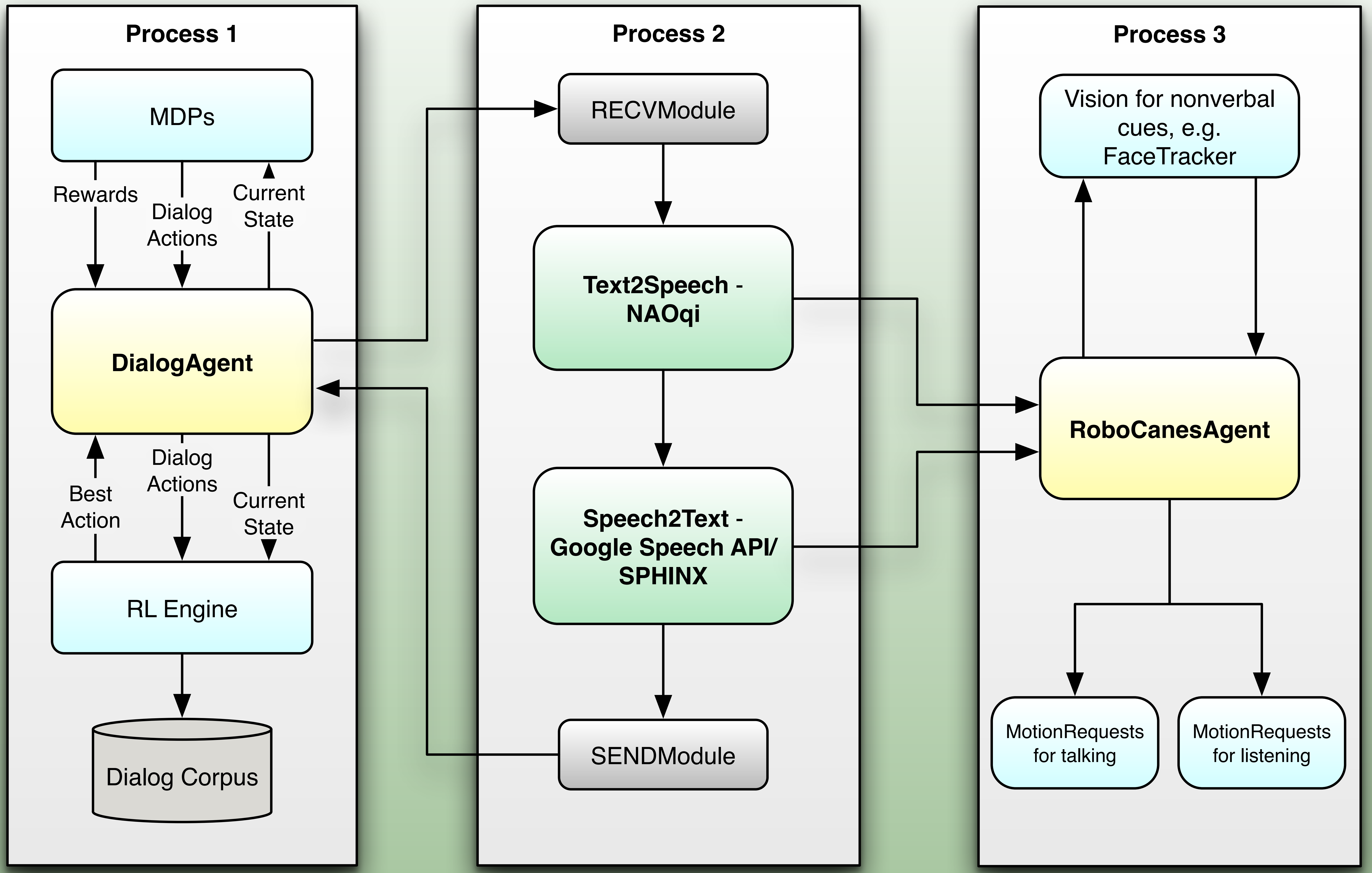
Poore et al. 2014 (*RoboCup Symposium*)

COMMUNICATION (2): WHISTLE DETECTION



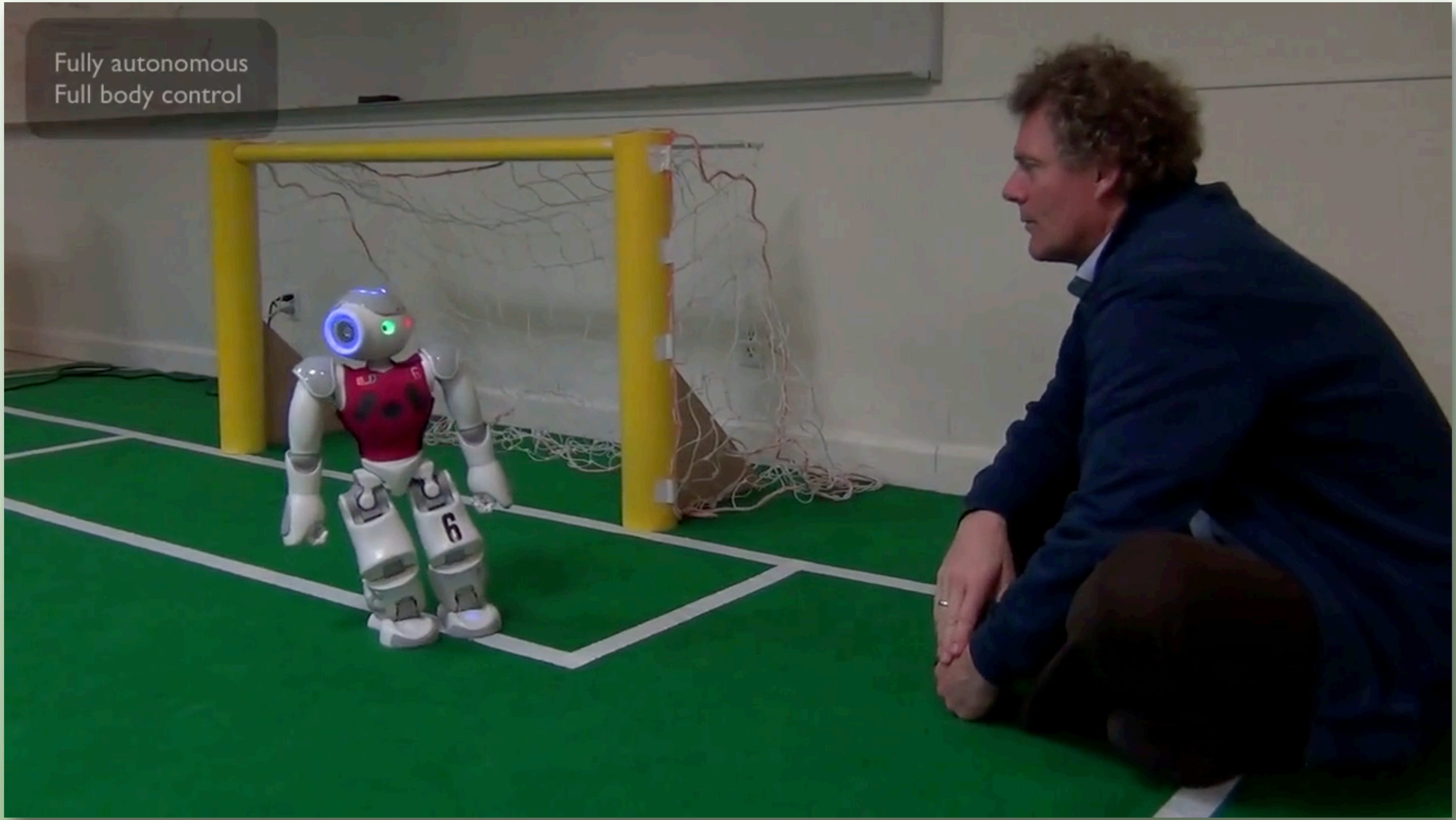
Poore et al. 2014 (*RoboCup Symposium*)

COMMUNICATION/CONTROL (3): SPOKEN DIALOG/HRI



Abeyruwan et al. 2014 (*AAAI Symposium*)

COMMUNICATION/CONTROL (3): SPOKEN DIALOG/HRI

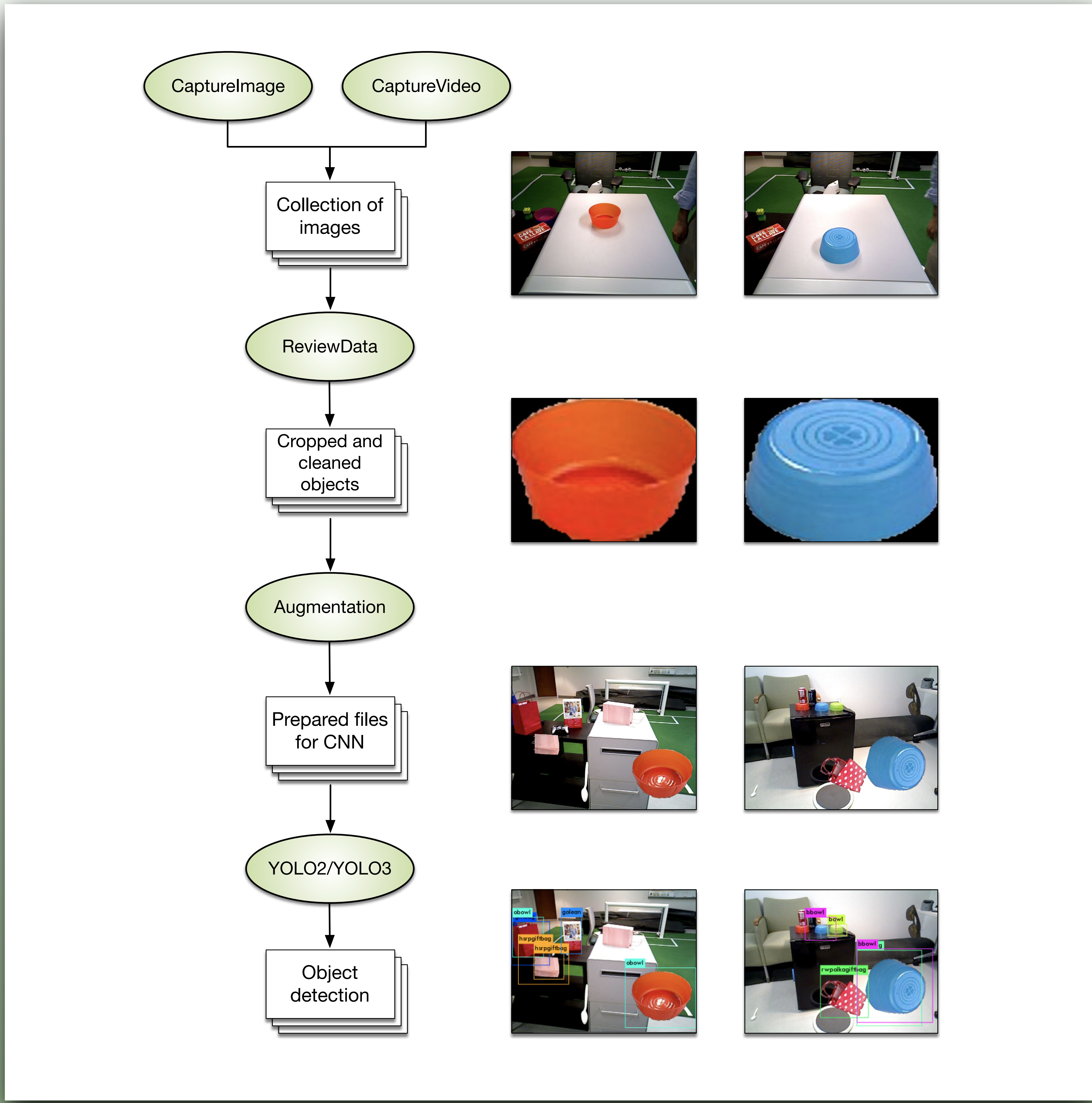


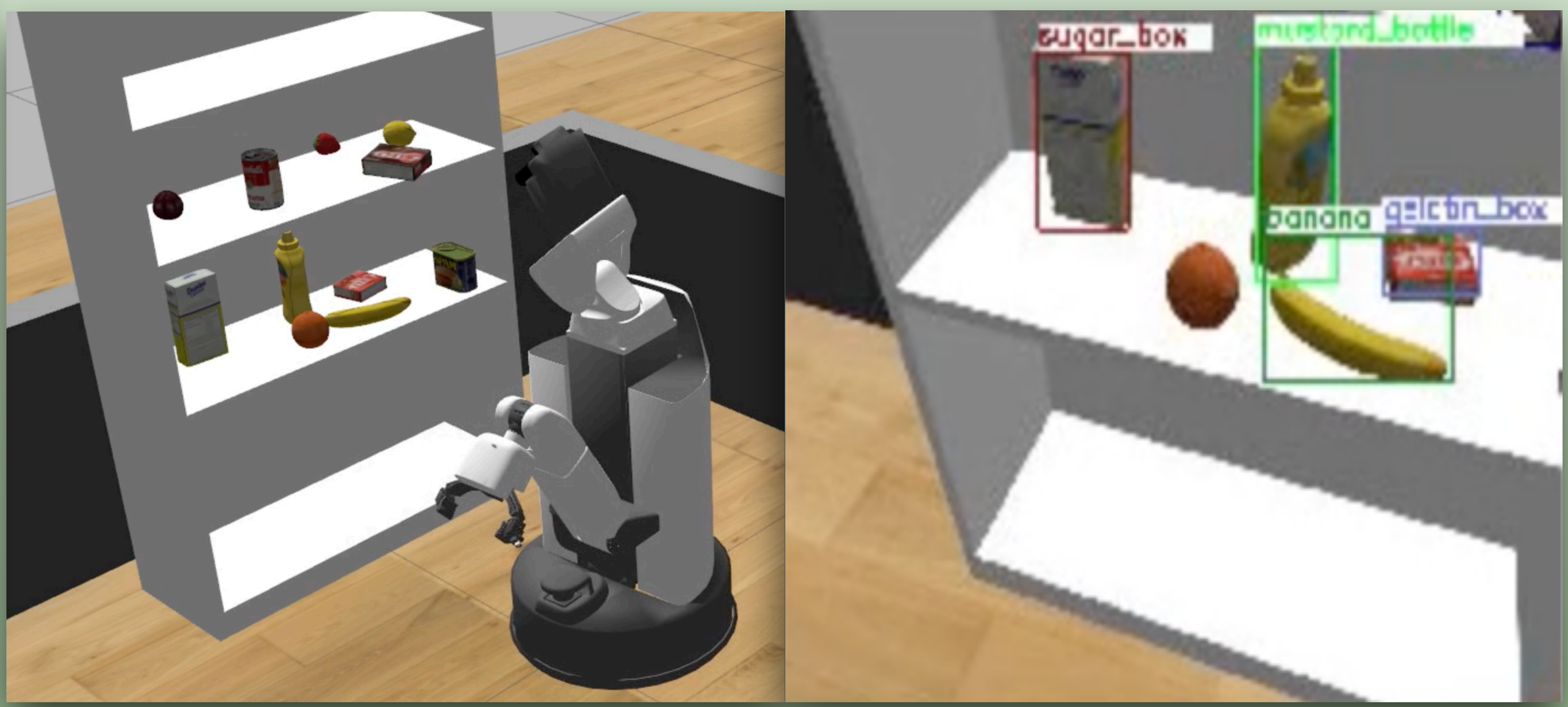
Abeyruwan et al. 2014 (*AAAI Symposium)*

2. MANIPULATION SYSTEM

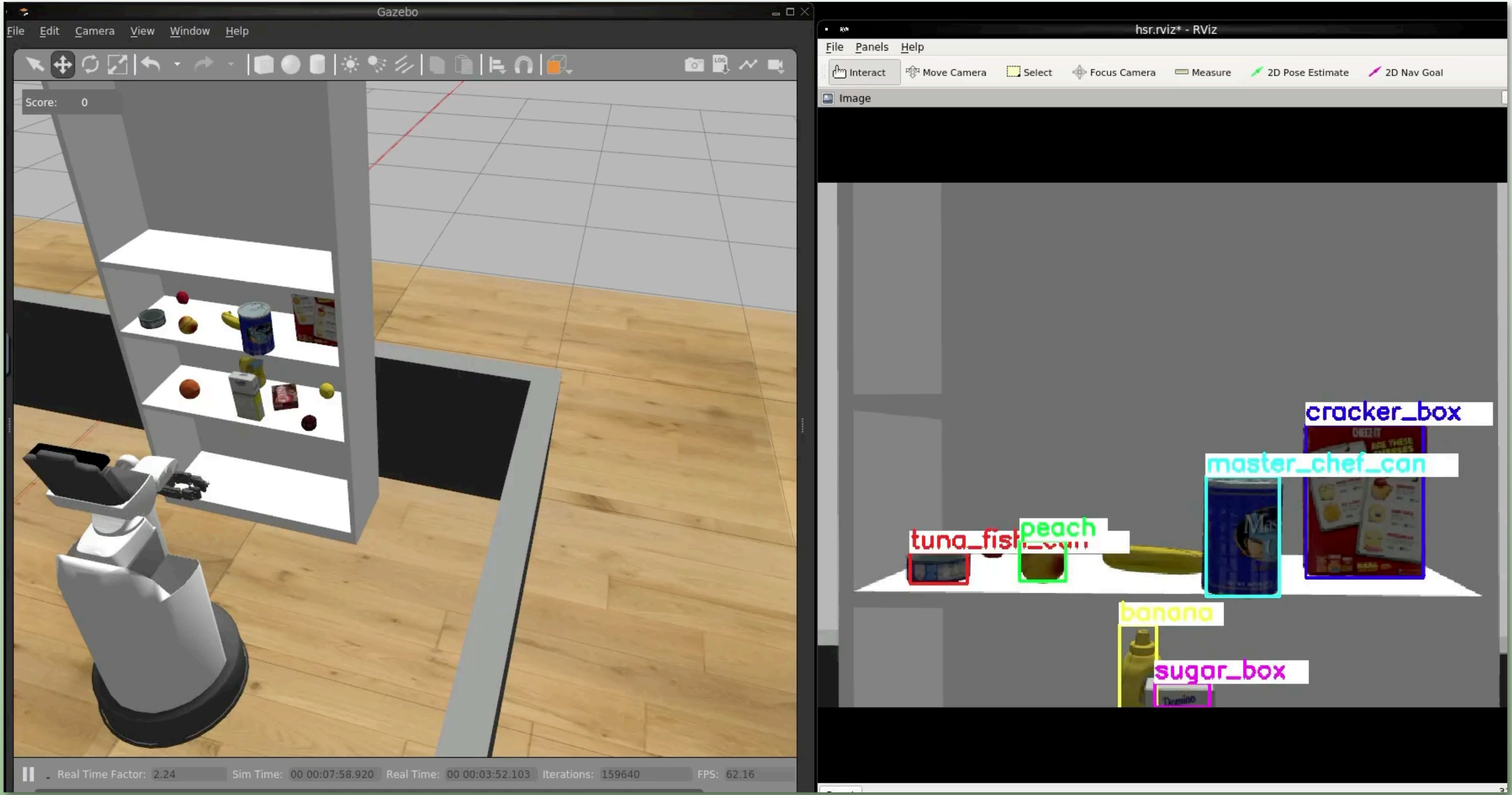
- Task and motion planning
  - Creation of high-level commands and collision-free trajectories to achieve goal
- State estimation and perception
  - Infer relevant quantities from sensor data (objects, drawers, manipulators, contacts/collisions, …)
- Object grasping and placement (pick-and-place)
  - Determine good grasps for objects given relevant constraints (gripper opening, local geometry, placement)
- Trajectory generation and control
  - Real-time, reactive generation of control commands to move robot (or parts) safely toward goal

STATE ESTIMATION AND PERCEPTION

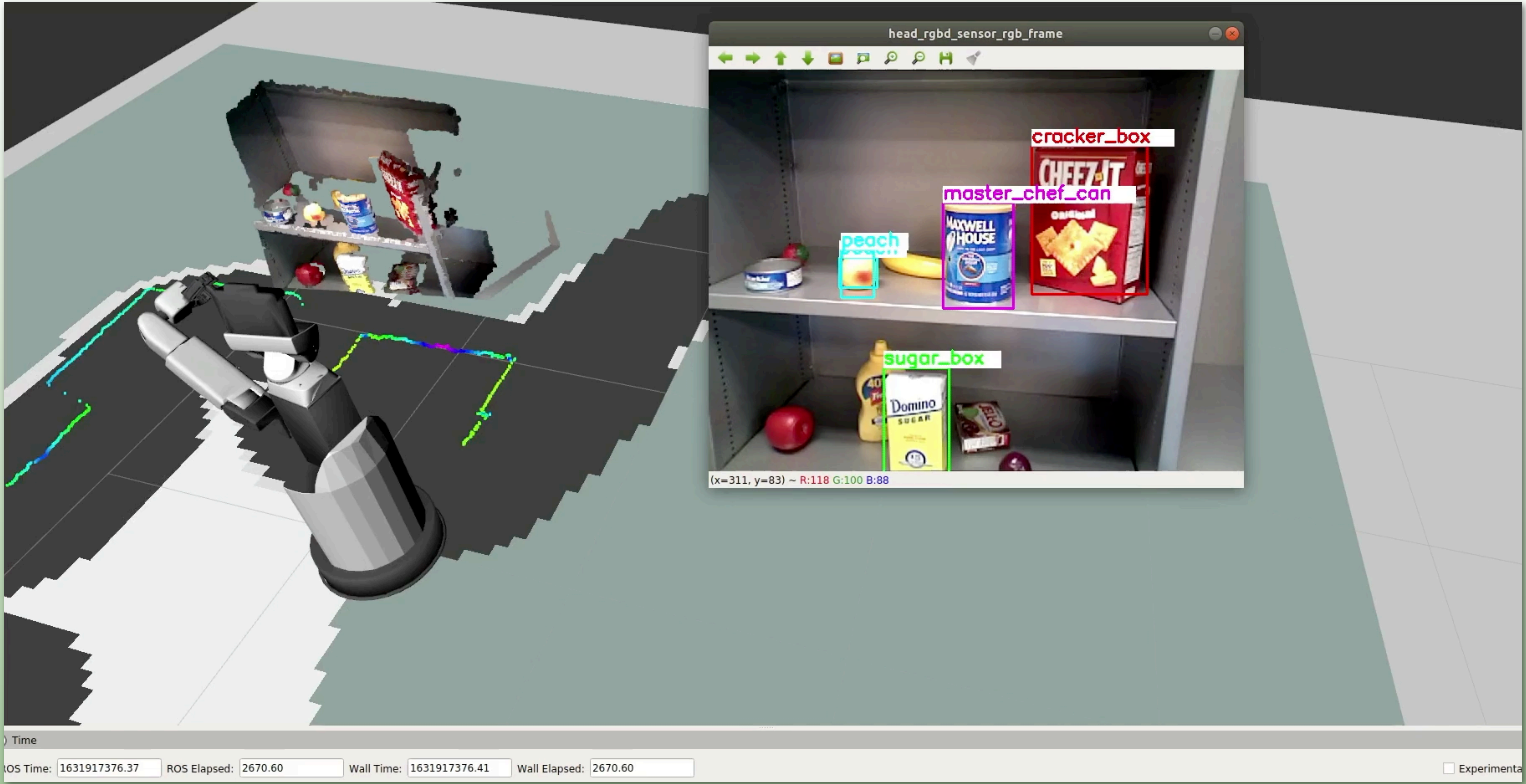




STATE ESTIMATION AND PERCEPTION

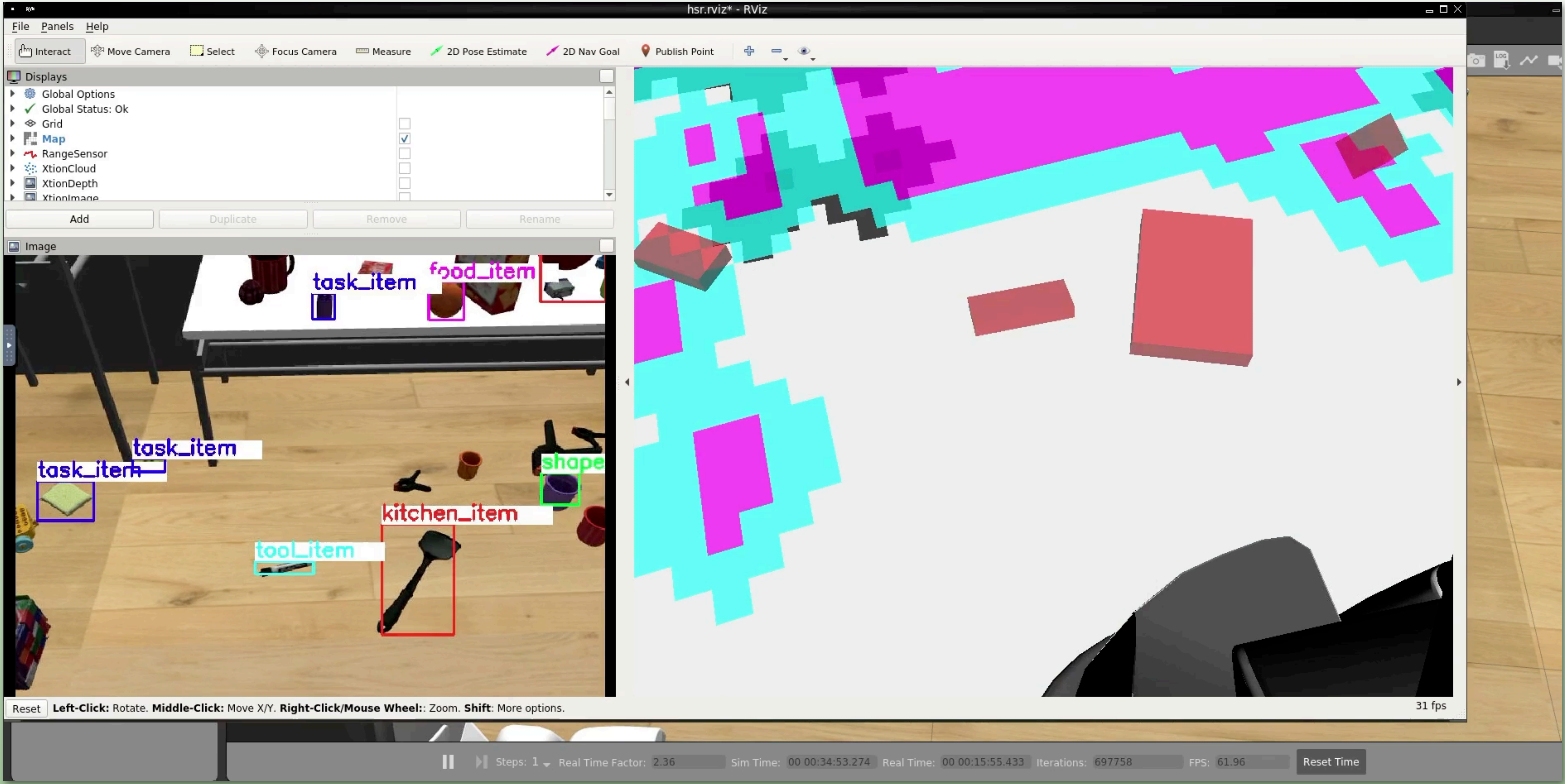


STATE ESTIMATION AND PERCEPTION



STATE ESTIMATION AND PERCEPTION

STATE ESTIMATION AND PERCEPTION

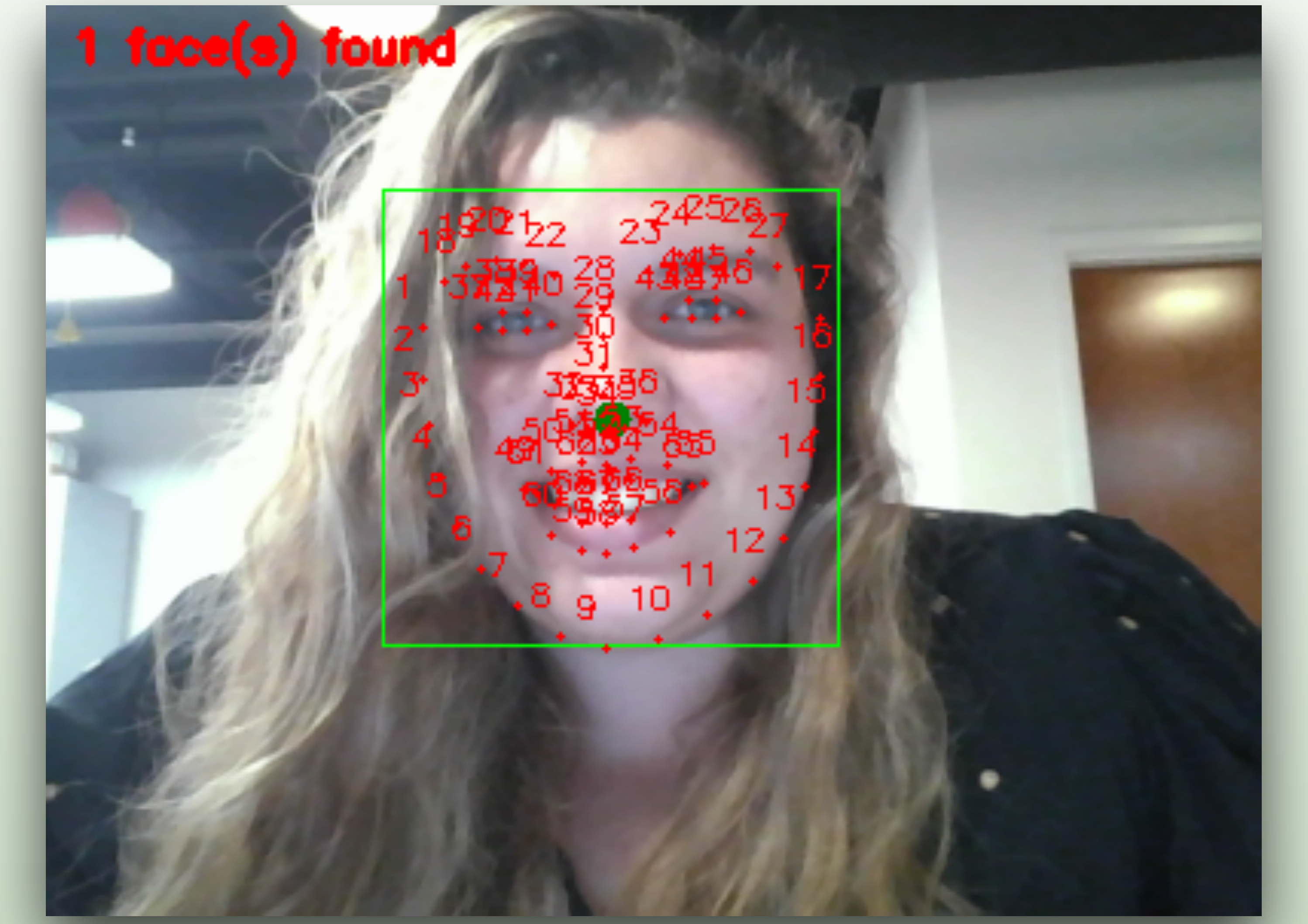


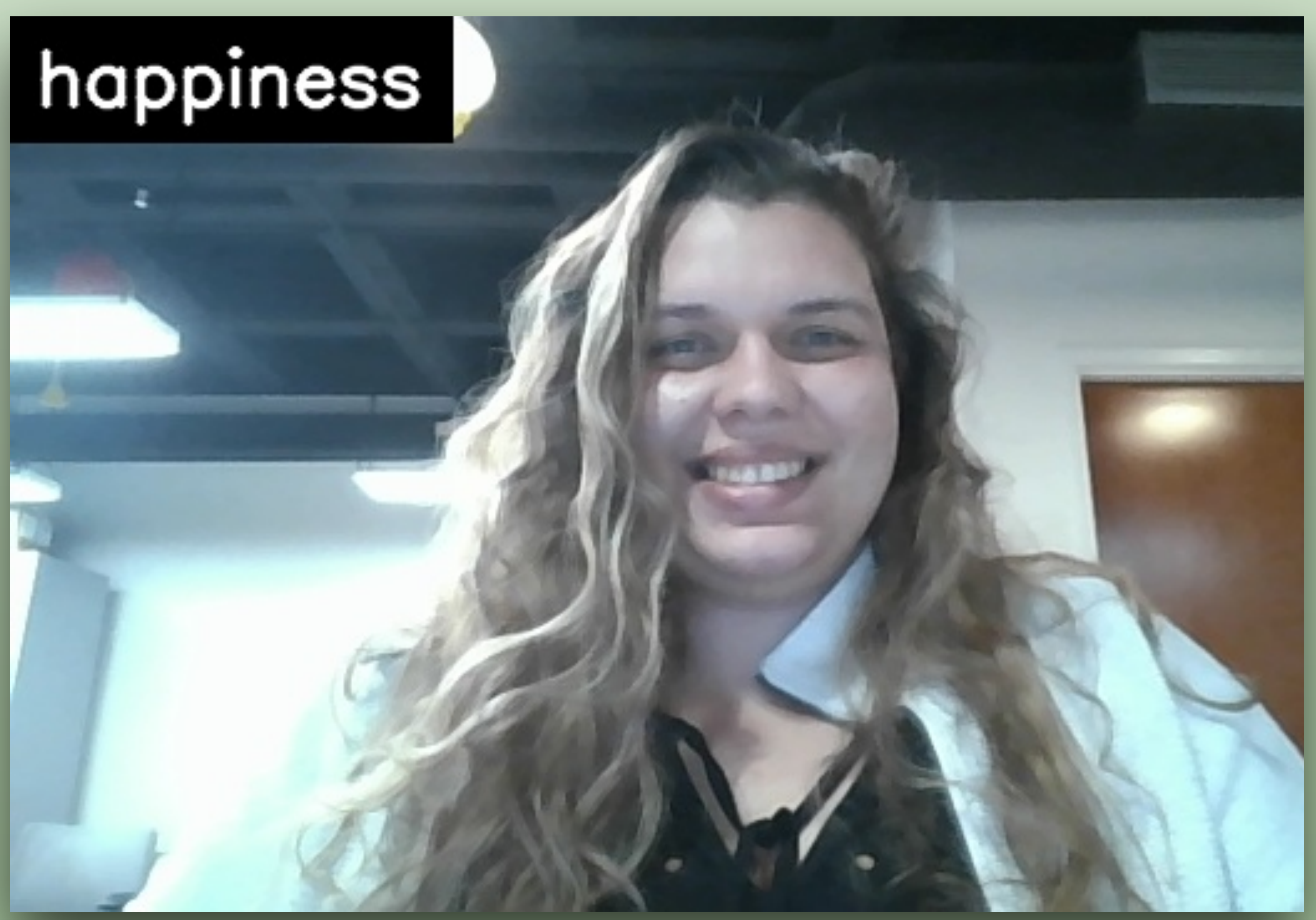
3. HUMAN ROBOT INTERACTION

- Task and motion planning
  - Creation of high-level commands and collision-free trajectories to achieve goal
- State estimation and perception
  - Infer relevant quantities from sensor data (human faces, human poses, emotions, NLP, race, ethnicity, …)
- Communication
  - Building rapport, interact socially, show affect, infer cultural differences, speech recognition, text2speech, people recognition
- Trajectory generation and control
  - Real-time, reactive generation of control commands to move robot (or parts) safely toward goal (people tracking, cleanup, go-get)

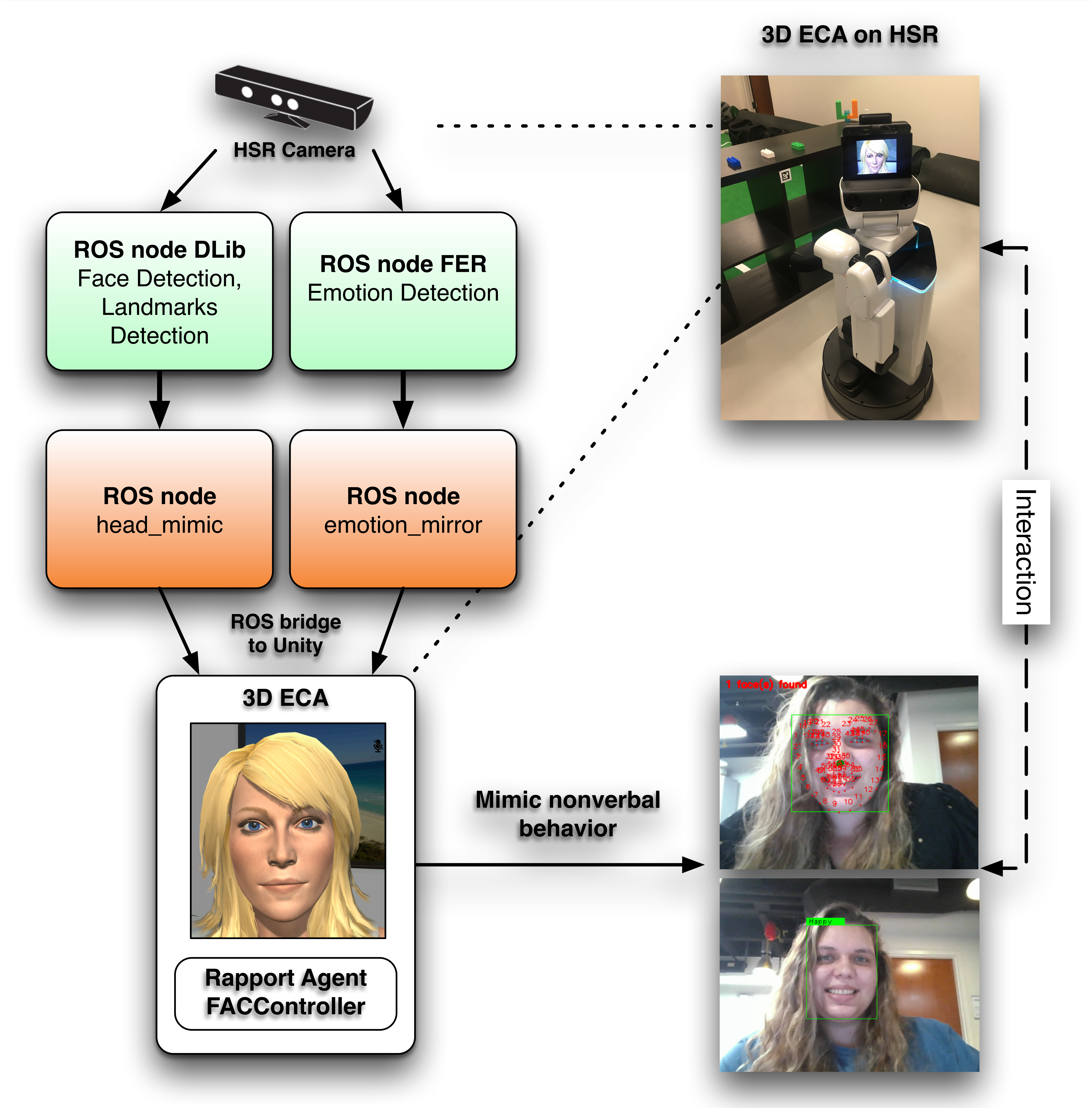
COMMUNICATION: BUILDING RAPPORT

- Rapport as result of a combination of socio-cultural-emotional complex processes, e.g. unconscious:
  - mutual attentiveness (mutual gaze, mutual interest, focus during interaction)
  - positivity (e.g., head nods, smiles, friendliness, and warmth)
  - unconscious coordination (e.g., postural mirroring, synchronized movements, balance, and harmony)
- Focus here on coordination/mirroring of
  - head movements and
  - facial emotions





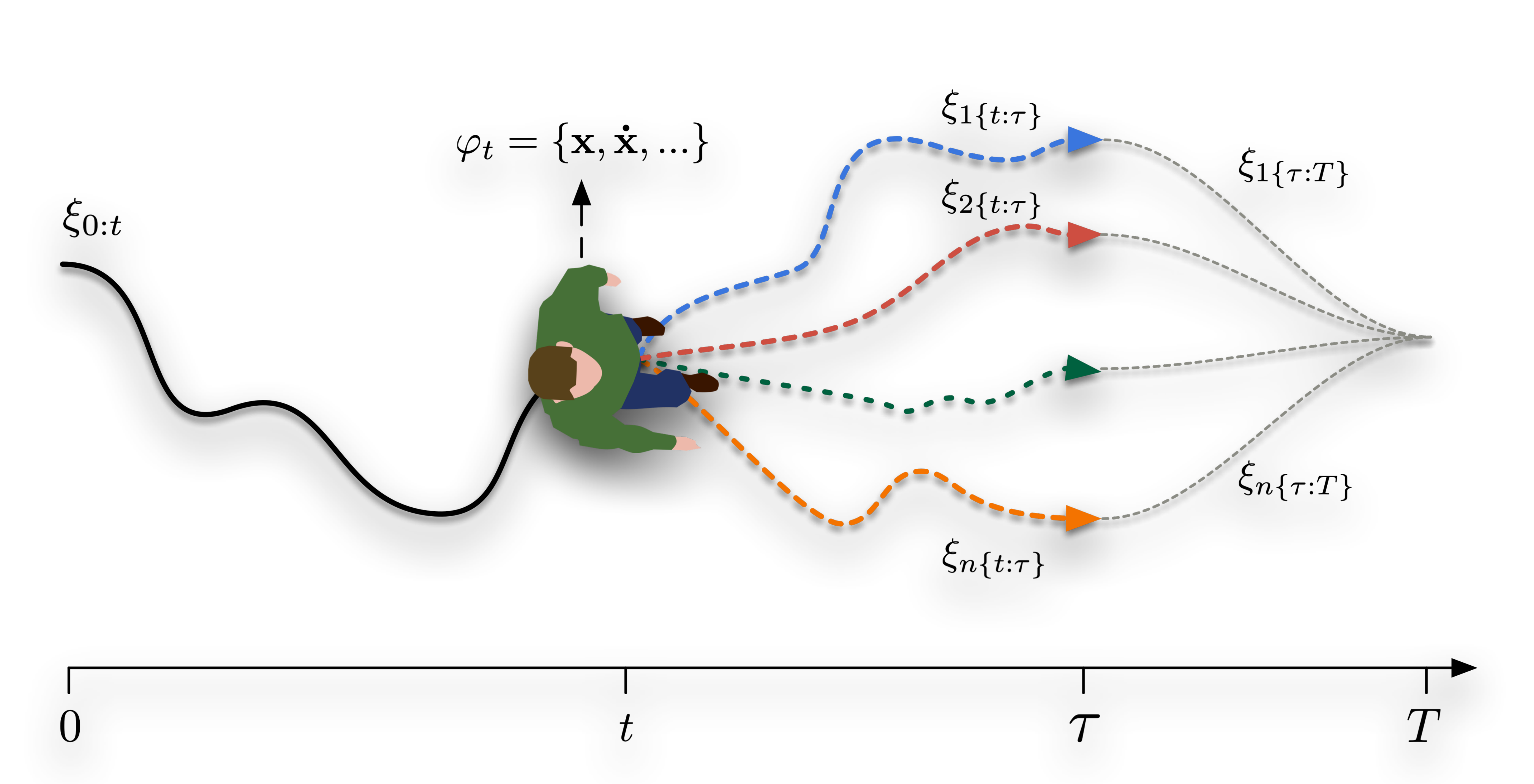
COMMUNICATION: BUILDING RAPPORT



COMMUNICATION: BUILDING RAPPORT

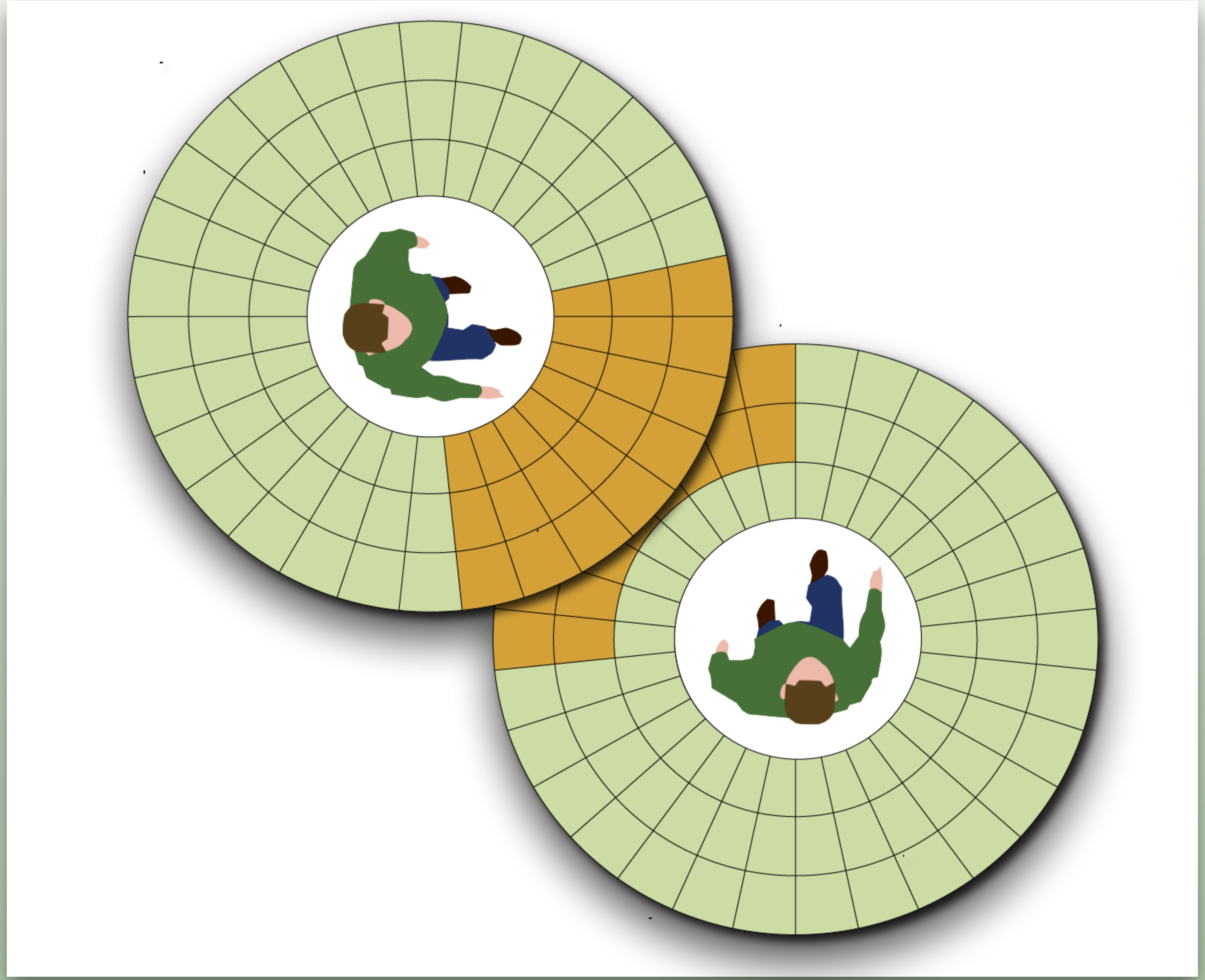
STATE ESTIMATION AND PERCEPTION: HUMAN POSES

- Tracking and predicting humans in 3D space
- Novel probabilistic framework in which multiple models can be fused into a circular probability-map to forecast human poses
- ITP: Inverse Trajectory Planning



Peña & Visser (*KI: German AI Journal, 2020)*

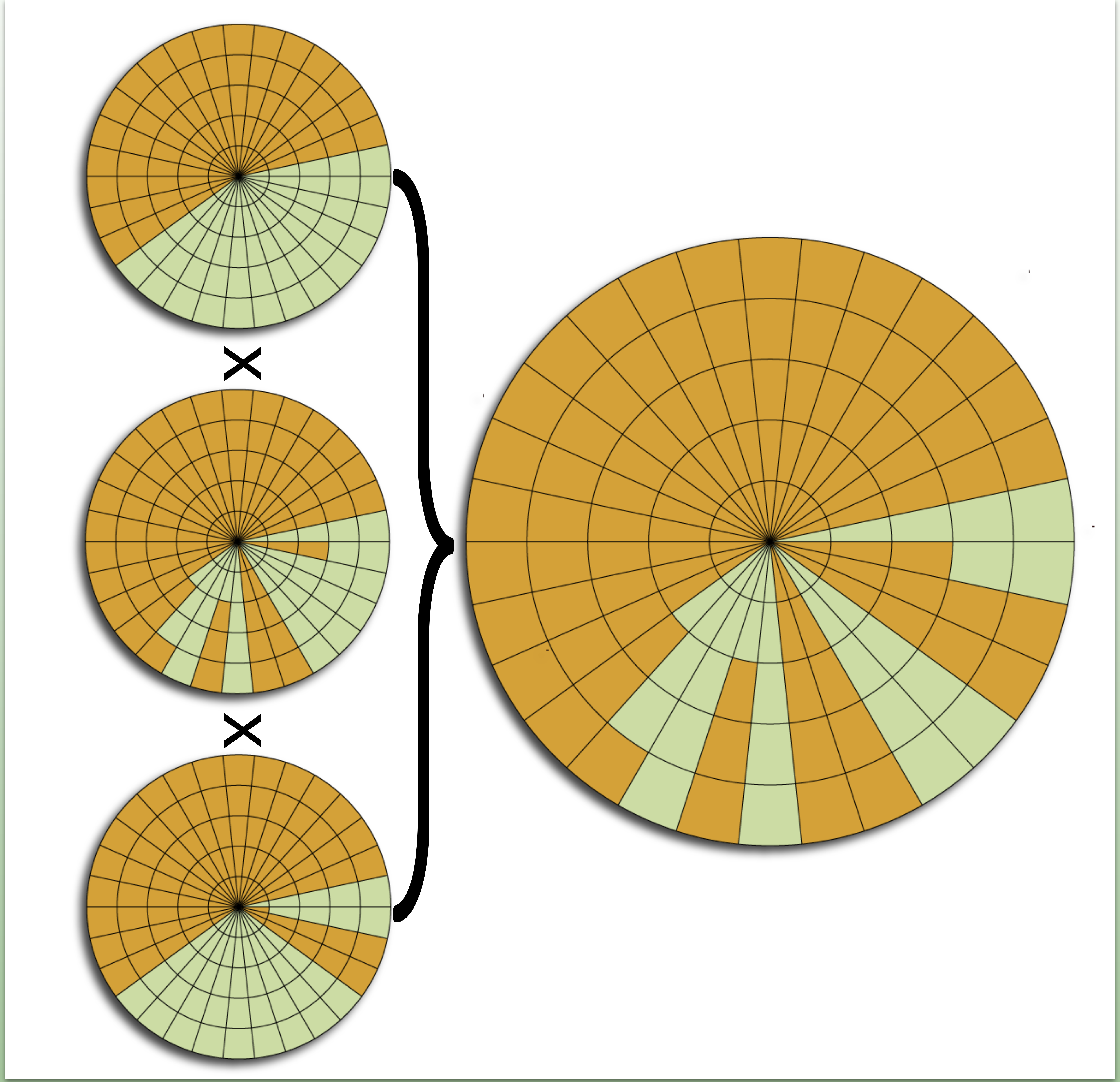
- Circular probability-maps for the social force model
- ITP allows interaction between two people or two probability-maps



Peña & Visser (*KI: German AI Journal, 2020*)

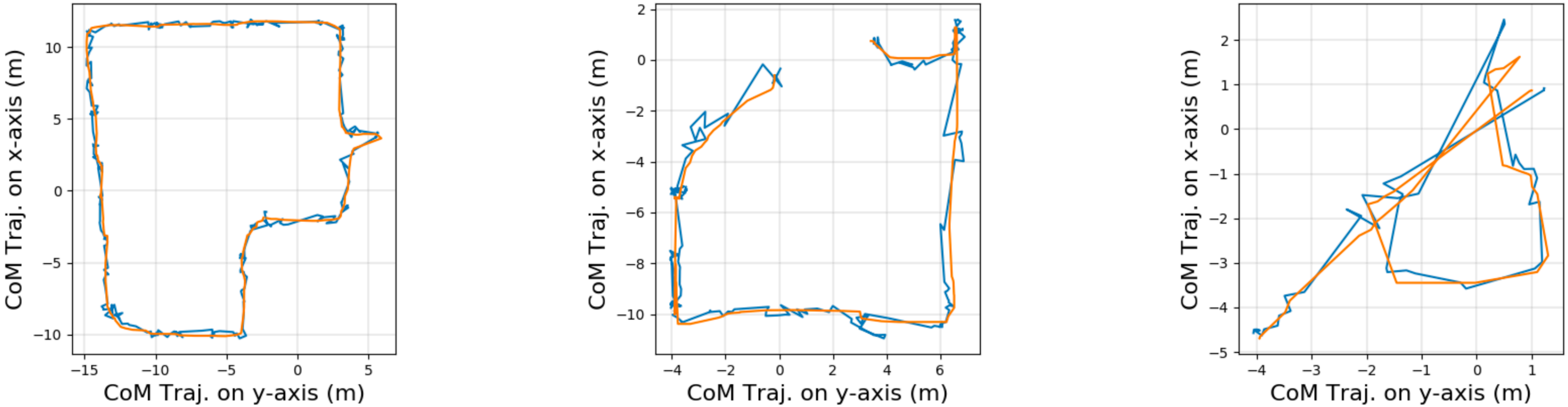
STATE ESTIMATION AND PERCEPTION: HUMAN POSES

- Lidar-based model for 2D obstacles
- OctoMap model for 3D obstacles
- Social force model for other humans
- Result: Heading model as prediction of future states



Peña & Visser (*KI: German AI Journal, 2020*)

STATE ESTIMATION AND PERCEPTION: HUMAN POSES



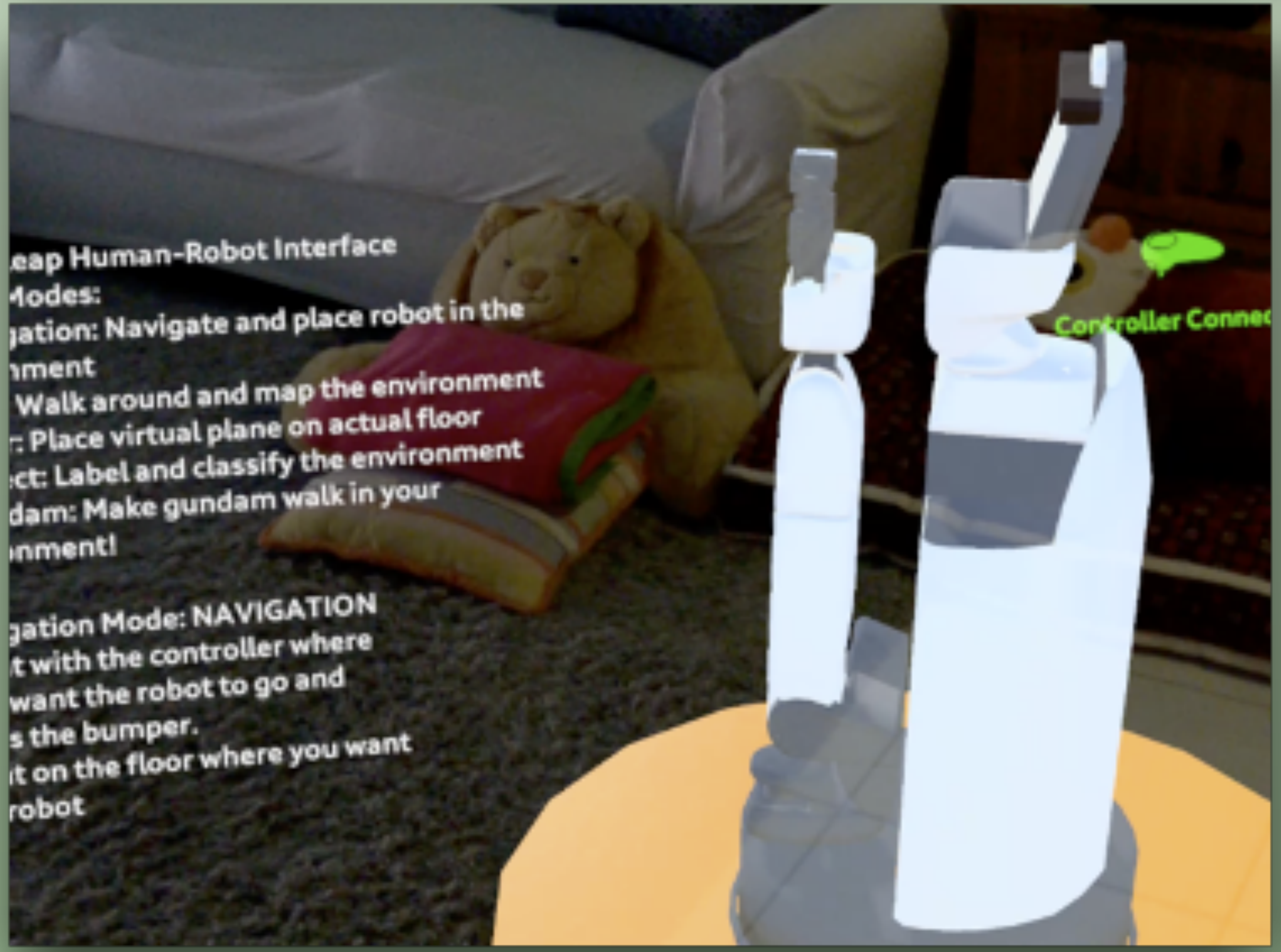
Peña & Visser (*KI: German AI Journal, 2020)*

STATE ESTIMATION AND PERCEPTION: HUMAN POSES

Peña & Visser (*KI: German AI Journal, 2020*)

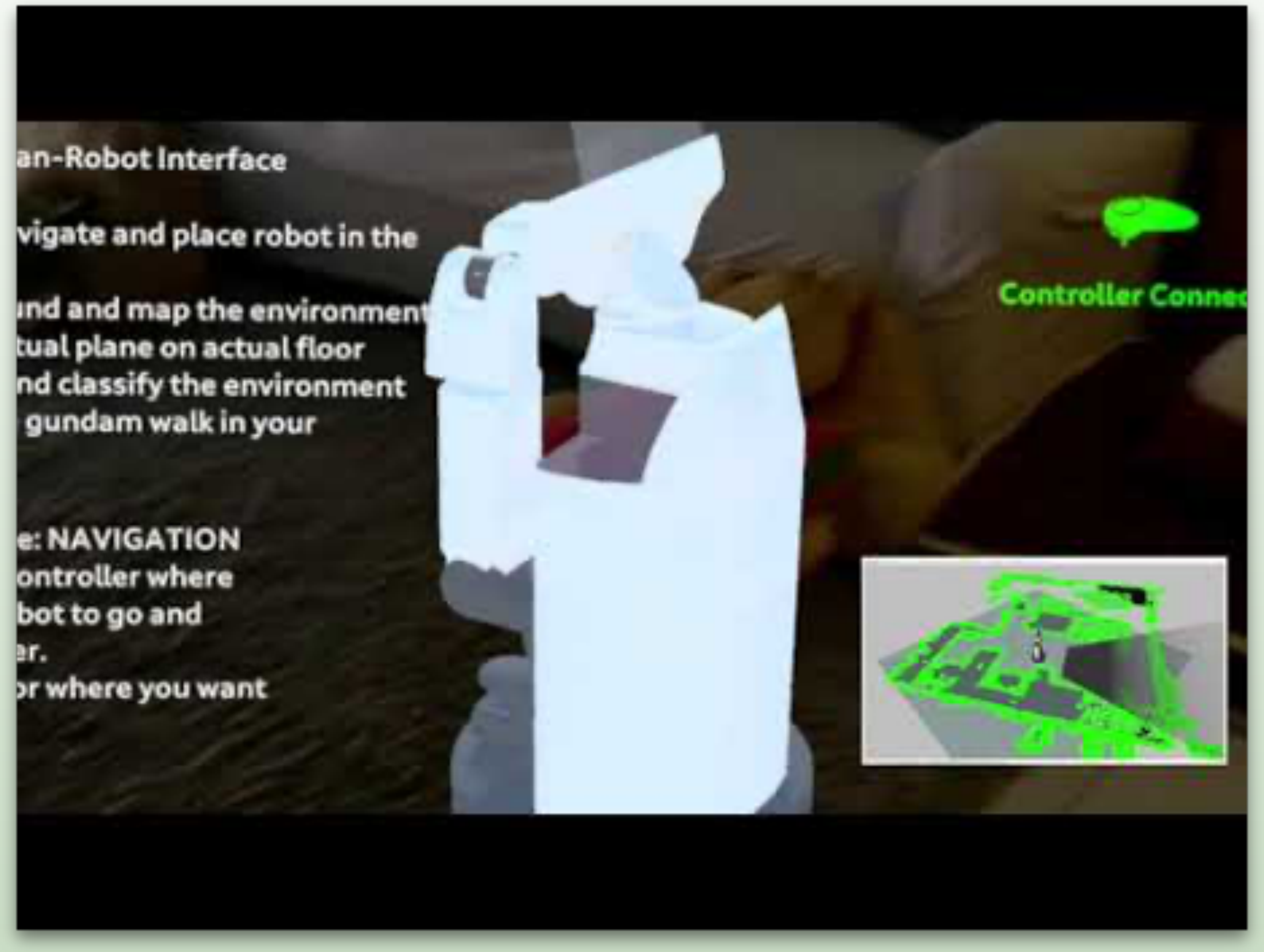
4. INTERFACES: VR MEETS AI AND ROBOTICS

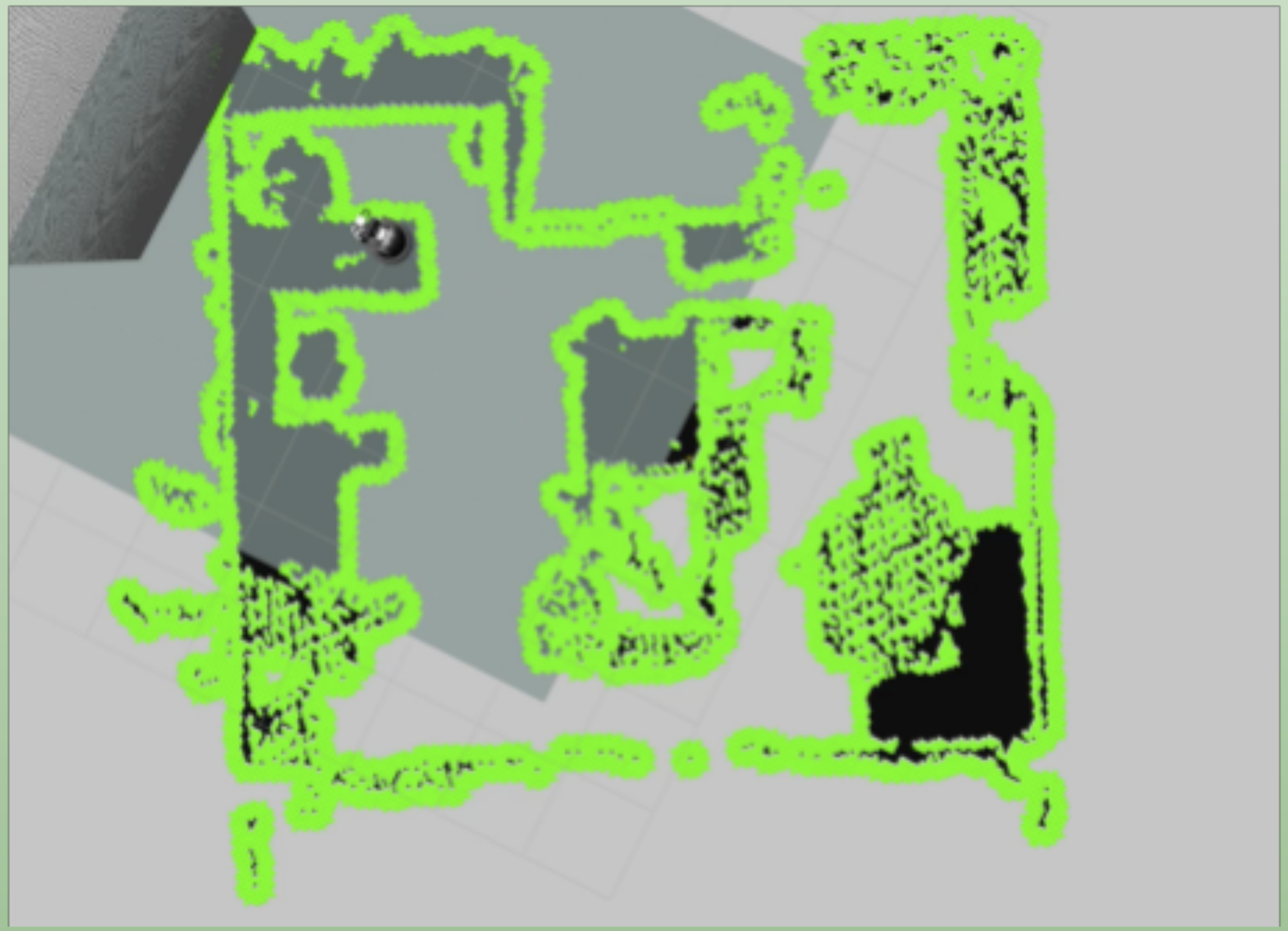
- Implement a human-robot interface that is intuitive and does not require a computer
- Geriatric (not robotics) expert uses magic leap to map the environment and label important features (bed, bathroom, sink, chair, TV, fridge, emergency path, etc.)
- Localization and navigation with no calibration process
- The interface is implemented in Unity Engine
- ROS# is a ROS (Robot Operating System) bridge for Unity
- Unity application communicates with ROS



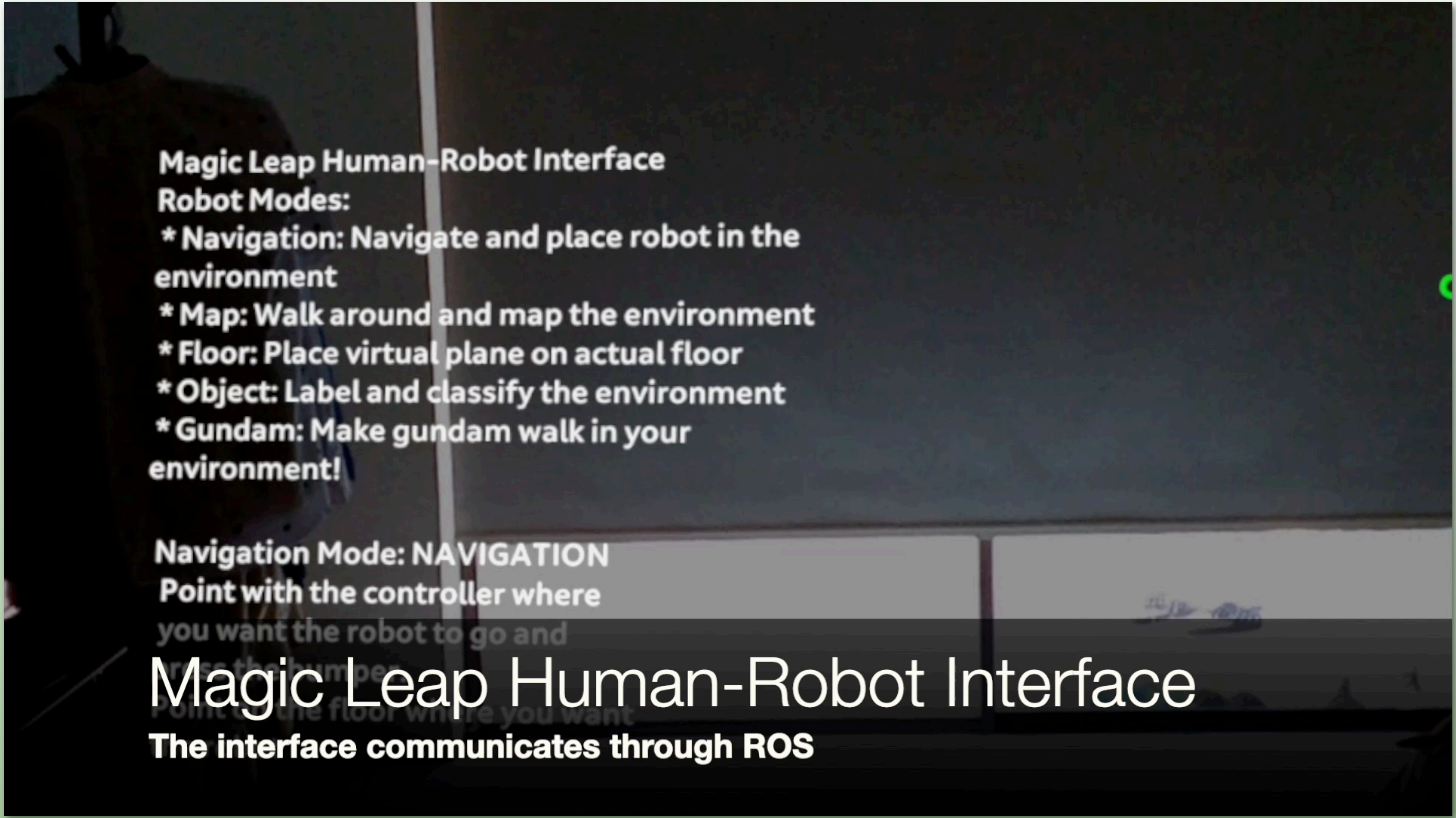
INTERFACES: IMPLEMENTATION (MAPPING AN ENVIRONMENT)

- The meshing feature of Magic Leap is used to create a 2D map
- A virtual camera in the Unity scene is placed on top of the scene generating a birds-eye view of the environment
- The image of the virtual camera is rendered in a texture
- The texture is used by the map server script in Unity to generate a occupancy grid map used by the navigation stack of the robot
- Along with the 2D costmap, the robot sends the transformation from Unity to ROS obtained from Magic Leap's global frame





MAGIC LEAP INTERFACE FOR HSR



5. FACIAL RECOGNITION: DETECTING RACIAL INEQUALITIES IN CRIMINAL JUSTICE

- Recent events have highlighted large-scale systemic (race, gender, skin tone, etc.) disparities in U.S. criminal justice.
- Propose an experimental methodology based on ethical AI principles to generate binary racial categories using mugshots
- Data: ~200K defendants from Miami-Dade County Clerk of Records
- Ground truth: 2 sources
  - Dataset (N=14,177 random images) labeled by official court records single rater
  - Dataset (N=14,018 random images) is formed using consensus-driven racial categorization by multiple raters

Ousmane Bah (left) was arrested but let go after a detective realized he looking nothing like the suspect (right) caught stealing in a from an Apple store. (Photos courtesy of KTVU and Subhan Tariq)

N.Y. Teen Blames Apple's Facial Recognition for Wrongful Arrest, Files $1B Lawsuit

By Tanasia Kenney - April 29, 2019 4084 0

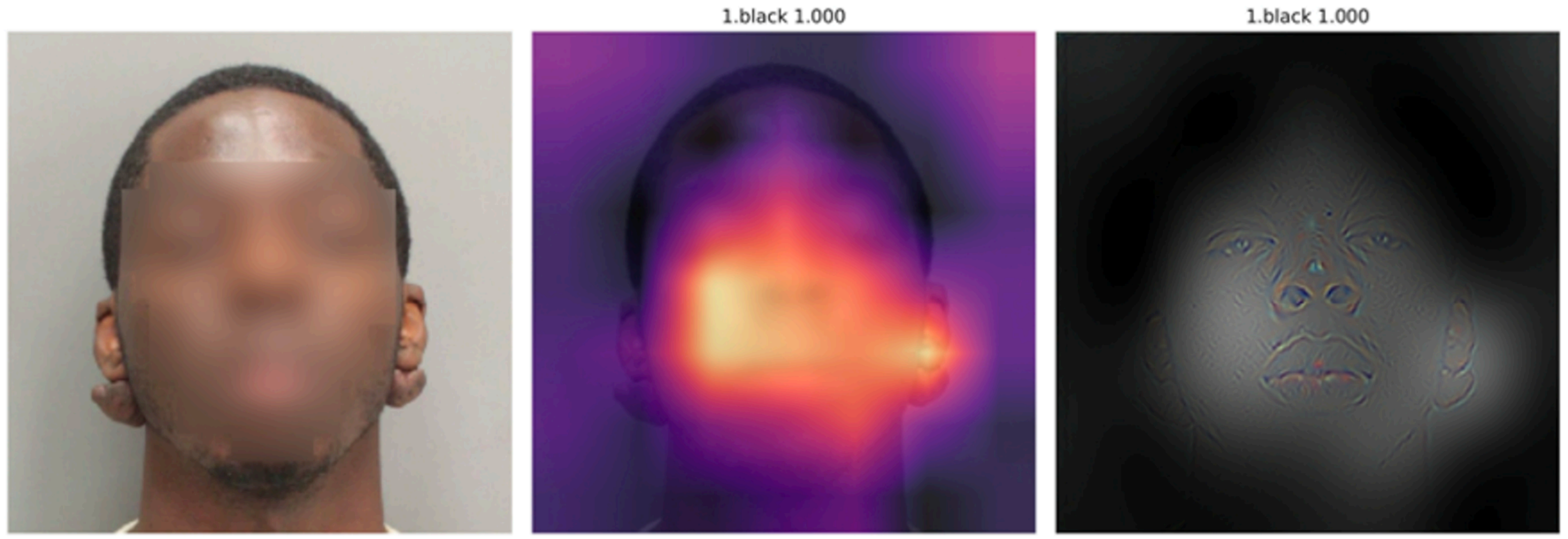
DETECTING RACIAL INEQUALITIES IN CRIMINAL JUSTICE

- First study showed 90%+ accuracy for race (black/white), < 75% accuracy for ethnicity (black/white/black hispanic/white hispanic) with SOA FRT (U-Link project)
- Results show that data preprocessing is crucial

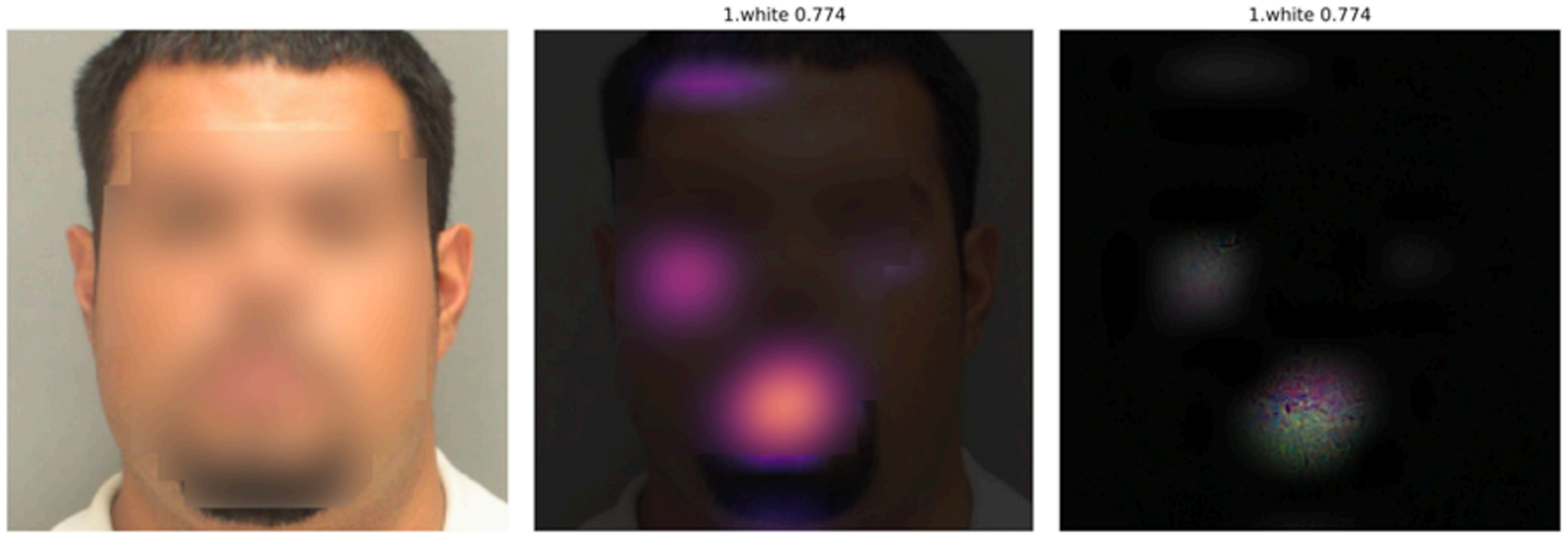
- DL library *fastai* using 7 state-of-the-art computer vision architectures (DenseNet161, ResNet50, InceptionV4, SE-ResNet50, SE-ResNeXt50_32x4d, AlexNet, and VGG19_bn)
- Trained with different preprocessing steps using our 2 sources, test on remaining images
- Highest accuracy is 97.75% (SE-ResNet50 with OpenFace preprocessing)
- Lowest accuracy is 1.28% (InceptionV4 with MTCNN preprocessing)

Dass et al. (*CRV, 2020)*

DETECTING RACIAL INEQUALITIES IN CRIMINAL JUSTICE



(c) “Best” Black mugshot by MTCNN preprocessed court trained SE-ResNet50 model.



(b) “Worst” Black mugshot by OpenFace preprocessed court trained DenseNet161 model.

(e) “Best” Black mugshot by OpenFace preprocessed court trained InceptionV4 model.

(f) “Worst” Black mugshot by OpenFace preprocessed court trained InceptionV4 model.

DETECTING RACIAL INEQUALITIES IN CRIMINAL JUSTICE

- Method
  - Data generation (multiple "ground truths" to tackle **labeling bias)**
  - Data preprocessing (face segmentation, aligning, cropping, pose, illumination, ...)
  - DLM architectures (7 SOA methods with 6 experimental combinations) to tackle **algorithmic bias**
  - Training setup to avoid **representation bias**
  - Testing (model inference and interpretability) to tackle **historical** and **evaluation bias; self-auditing** (using 42 fine-tuned models for 12 test scenarios: 252 in total)

Table 4: A comparison of **validation accuracies** during the training process from seven deep learning-based vision models.

Model	original		OpenFace		MTCNN	
	Courts	Students	Courts	Students	Courts	Students
AlexNet	92.50%	**94.00%**	97.25%	92.25%	92.75%	93.00%
DenseNet161	**97.00%**	93.00%	97.00%	92.25%	96.50%	94.00%
InceptionV4	93.25%	90.25%	90.75%	90.50%	90.75%	91.75%
ResNet50	96.50%	92.50%	97.00%	91.50%	95.25%	93.25%
SE-ResNet50	95.00%	92.75%	**97.75%**	**92.50%**	96.75%	93.25%
SE-ResNeXt50	96.75%	91.75%	97.75%	90.00%	96.75%	94.25%
VGG19	96.00%	89.75%	97.00%	92.25%	**96.75%**	**94.75%**

CONCLUSIONS

- Task and motion planning
- State estimation and perception
- Communication
- Object grasping and placement
- Trajectory generation and control

THANK YOU!


Saminda Abeyruwan


Chloe Arluck


Lloyd Beaufils


Michael Davis


Rahul Dass


Alexander Härtl


Nasir Laskar


Shengxin (Tony) Luo


Joe Masterjohn


Phil McKenna


Katarzyna (Kasia) Pasternak


Pedro Peña


Kyle Poore


Andreas Seekircher


Justin Stoecker


Zishi Wu

With help from the VISAGE group (FIU): Dr. Christine Lisetti, Dr. Mihai Polceanu, Stephanie Lunn; ENIB/FRANCE: Dr. Cédric Buche

- Contact
  - Web: https://www.cs.miami.edu/~visser
  - Email: visser@cs.miami.edu