

# Flexible Robot Error Detection Using Natural Human Responses for Effective HRI

Maia Stiber

Department of Computer Science, Johns Hopkins University

## Robot Errors Can Cause Harm

Unmanaged robot errors damage user trust and task performance, leading to user reluctance in collaborating with robots—underscoring the importance of effective error management strategies.

## Need for Flexible Robot Error Detection

Traditional error detection methods in HRI often use task or error-specific information and lack versatility, highlighting the need for a more flexible approach that can robustly detect robot errors across various tasks, scenarios, and error types.

## Potential of Social Signals

Errors elicit social signals from humans. Social signals serve as observable manifestations of a human's mental model of the interaction that they are involved in, conveying information about errors, tasks, and the robot.

**Aim: Enabling human-aware, error-aware robots through the use of social signals exhibited in response to robot errors to promote effective physical HRI.**

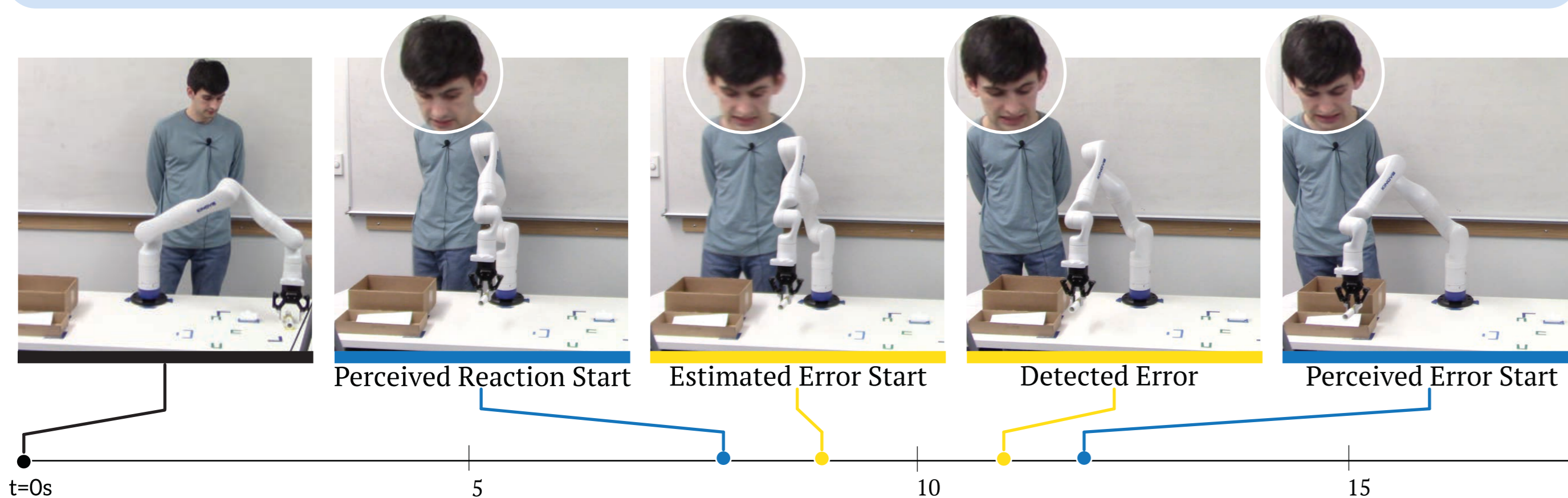
### 1 Understanding Social Signals in Response to Unexpected Errors

### 2 Operationalizing Social Signals for Flexible Automatic Error Detection

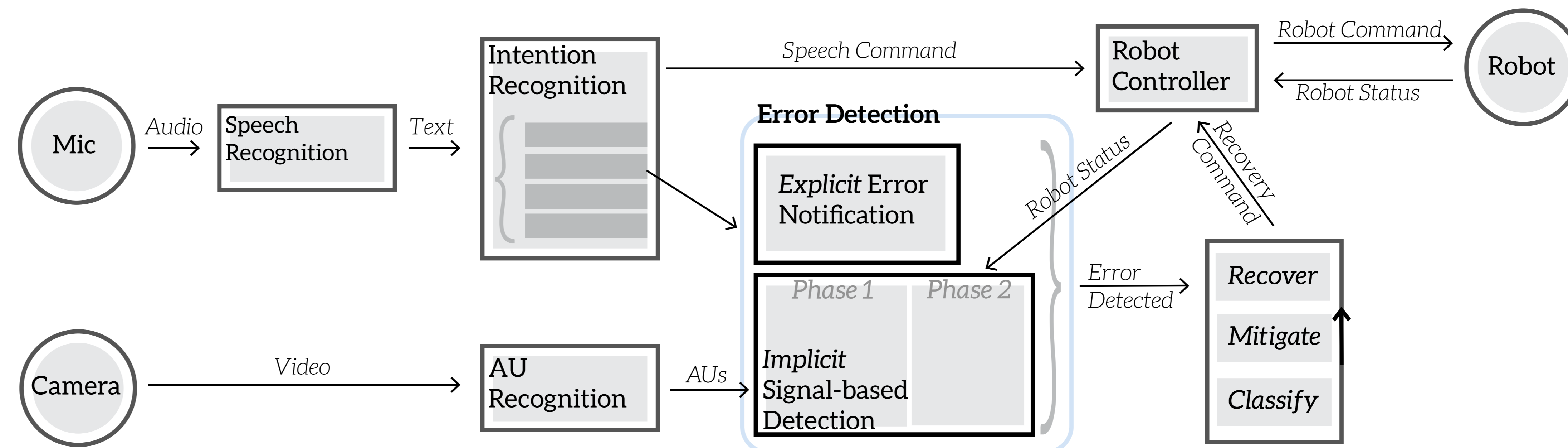
#### Social Response to Unexpected Errors



- Users consistently exhibit facial action units (AUs) in response to errors
- AUs hold discriminative power to detect errors during physical interactions in a timely manner
- Modeling AUs may be useful in error detection across people, tasks, and error types



#### Active Flexible Robot Error Detection System

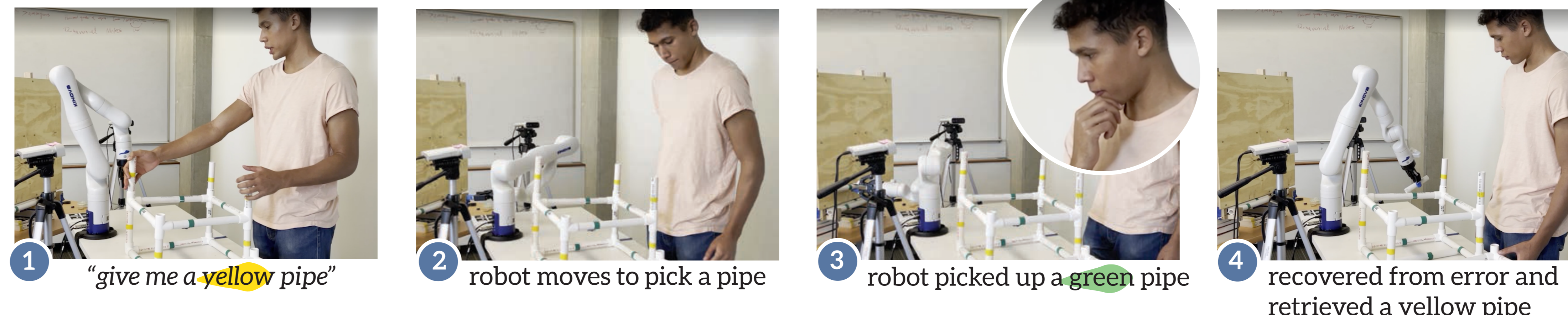


Robotic system employs **active** error detection

- [Phase 1] AUs (implicit signal) and robot status to initially detect an error
- [Phase 2] Verbal confirmation from user responding to a robot query

Social signals have potential to improve on the timeliness of traditional error detection.

Method uses social signals to allow us to capture unexpectedness and variability in complex human-robot interaction.



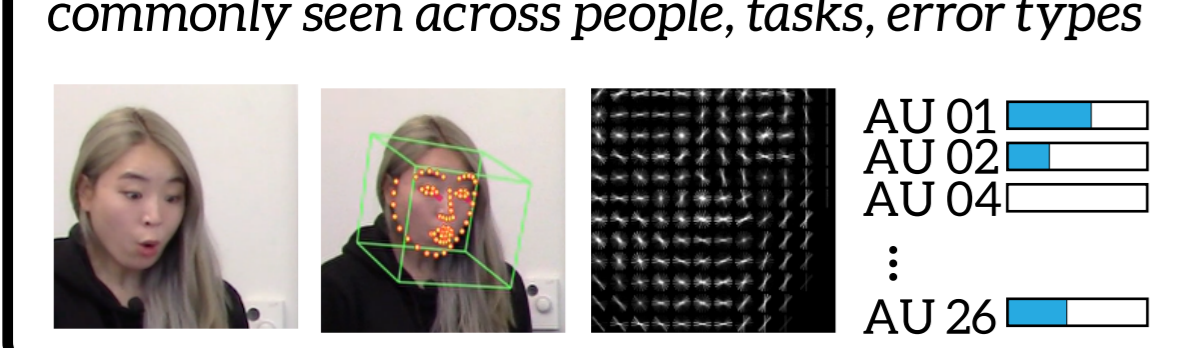
#### Framework for Flexible Error-Aware HRI

Status quo methods for error detection

**User indication**

**Deviated task behavior**  
w.r.t. coded task/domain specific rules

**Natural social reaction**  
commonly seen across people, tasks, error types



**Error Management**

**detection is the critical first step**  
in error management

Detect

Classify Mitigate Recover

#### Contributions

- Illustrate that social signals are exhibited in response to errors across tasks, people, error types, and embodiments.
- Show context influences social signals
- Show social signal modeling can be used across tasks, people, and error types
- A real-time error detection system using flexible error-aware framework
- Support that active automatic error detection is beneficial and users prefer it

#### Future Work

**Social Signals and Error Context**

- Should social signal models be used across embodiments for error detection?
- Does incorporating context improve error detection performance?

**Social Signals in the Wild/Repeated Interactions**

- What social signals are exhibited during natural user interactions in the wild?
- How does social signals exhibited change over time (repeat interactions)?

#### Acknowledgements

This work was supported by the Johns Hopkins Institute for Assured Autonomy and National Science Foundation award #2143704.

