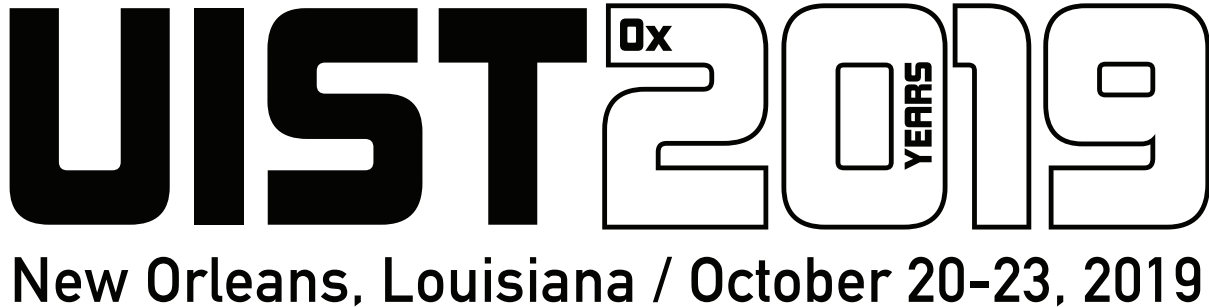
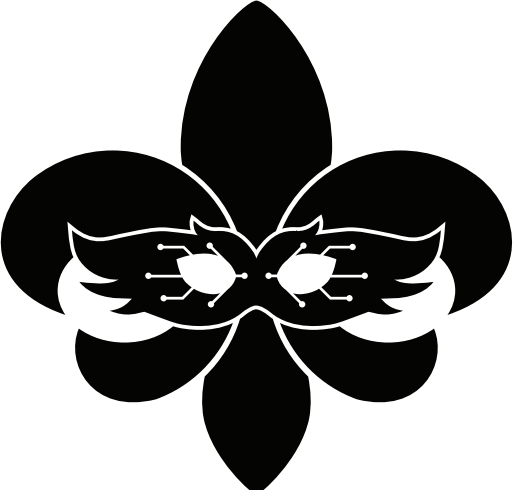


StateLens: A Reverse Engineering Solution for Making Existing Dynamic Touchscreens Accessible

Anhong Guo, Junhan Kong, Michael Rivera, Frank F. Xu, Jeffrey P. Bigham

*Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University*



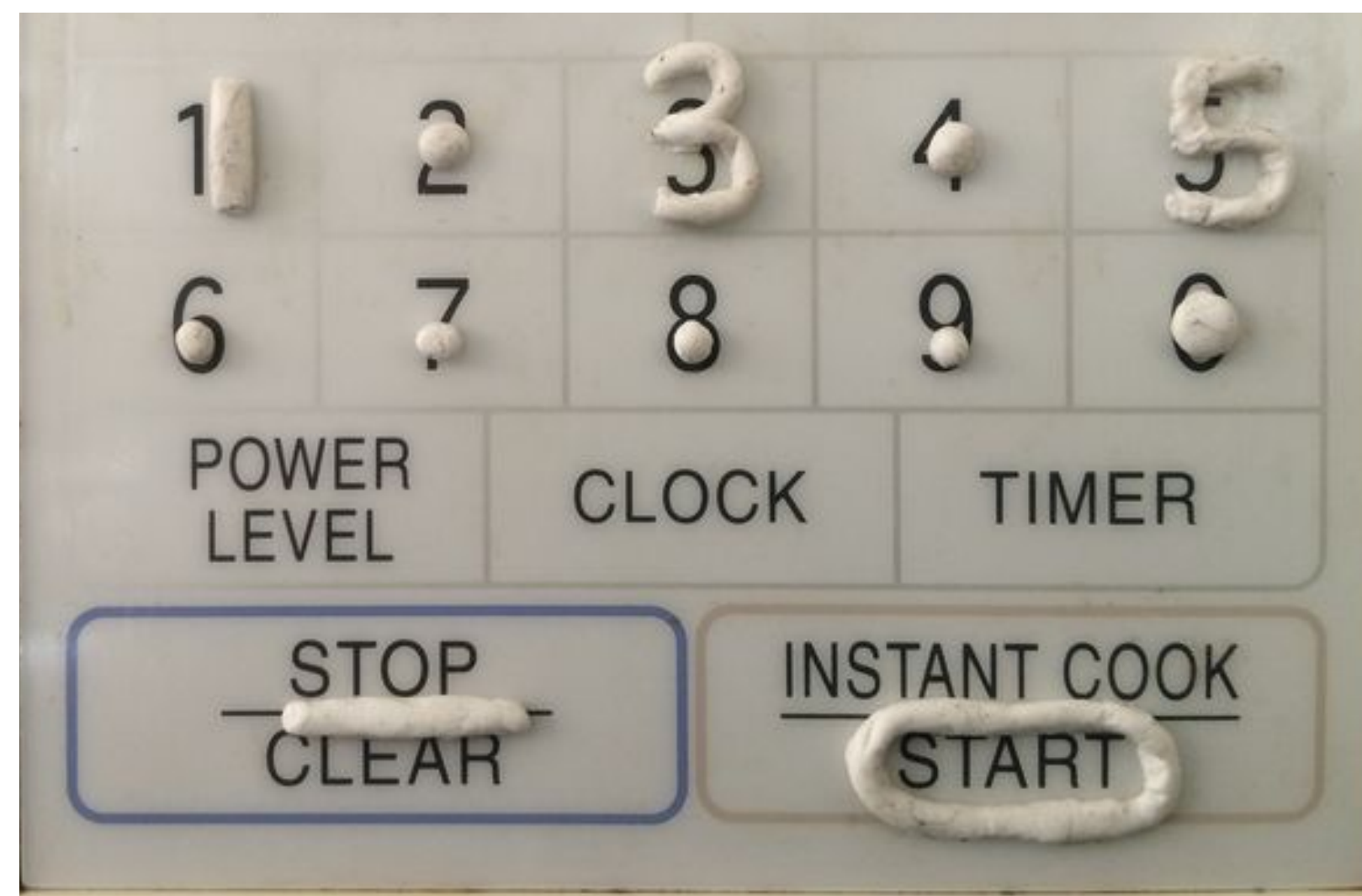


Static Interfaces: VizLens

Guo et al. "Vizlens: A robust and interactive screen reader for interfaces in the real world." *UIST 2016*.



Tactile Markings

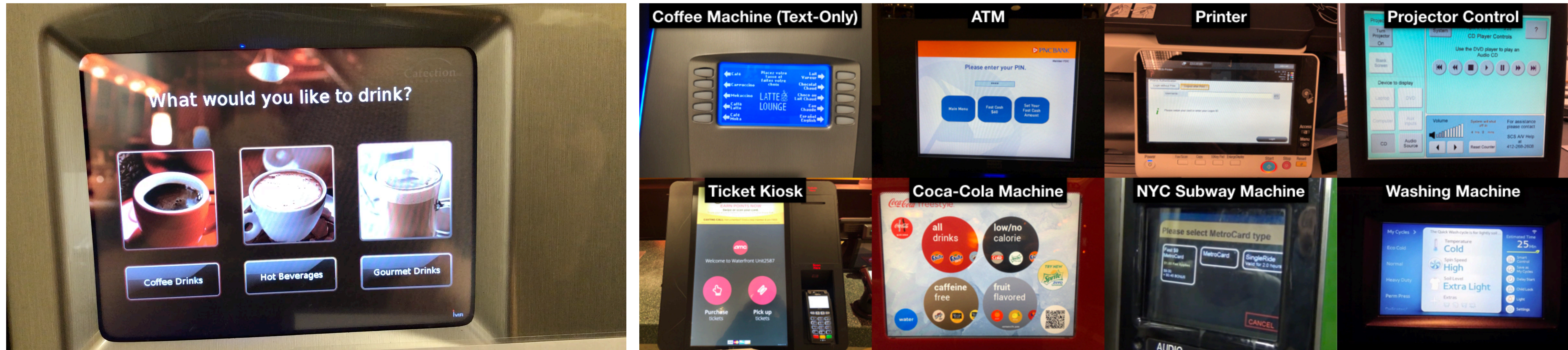


Static Interfaces: Facade

Guo et al. "Facade: Auto-generating tactile interfaces to appliances." *CHI 2017*.



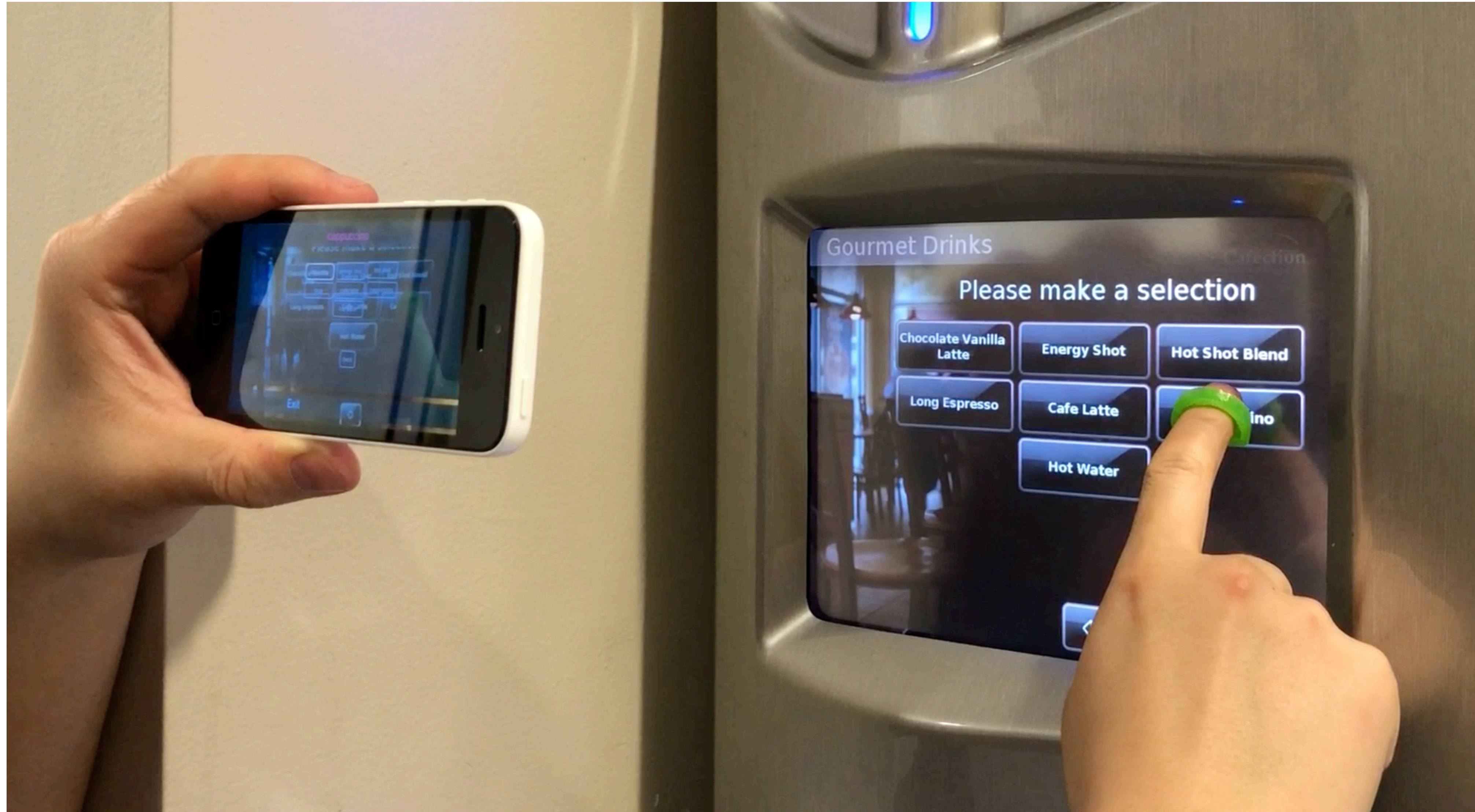
Dynamic Interfaces



Challenges:

- Visual user interfaces change \longrightarrow **Static tactile overlays won't work**
- Interactions often occur over multiple screens \longrightarrow **Hard to search and navigate**
- Easy to accidentally trigger actions while exploring \longrightarrow **Cannot touch and explore**

StateLens



Blind user completing task with app guidance

3D-printed Finger Cap for risk-free exploration

Blind user with wearable camera

Inaccessible Coffee Machine Mockup

StateLens iOS app: at large, press it... move down slowly...



StateLens Overview

1. Generate the State Diagram

Reconstructing interface structure using usage videos

2. Access the State Diagram

- Conversational agent
- Interactive audio guidance

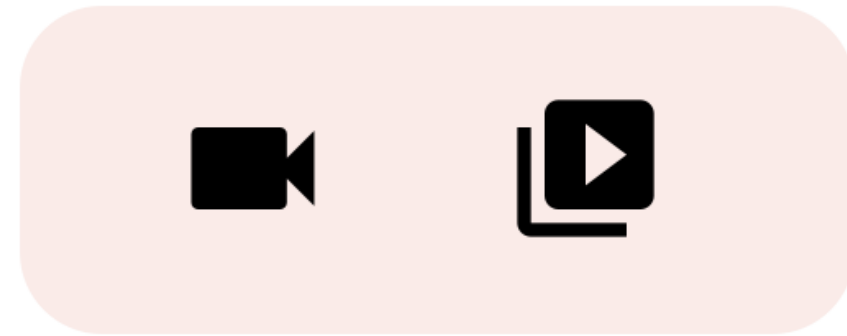
3. Risk-free exploration* accessories

Allowing blind users to freely explore without accidentally trigger touches

*Kane et al. Slide rule. *ASSETS 2008*.

1. Generate the State Diagram

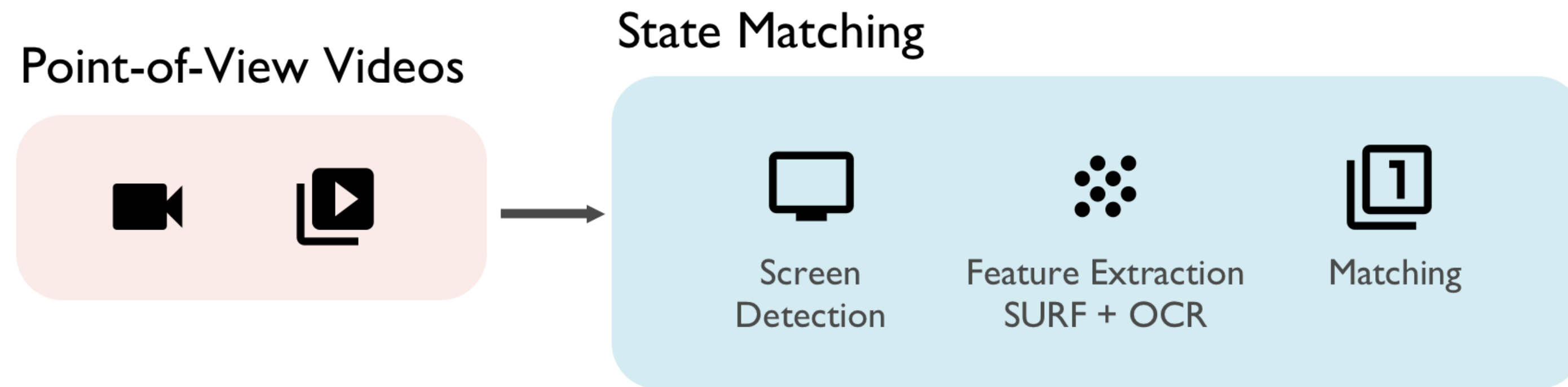
Point-of-View Videos



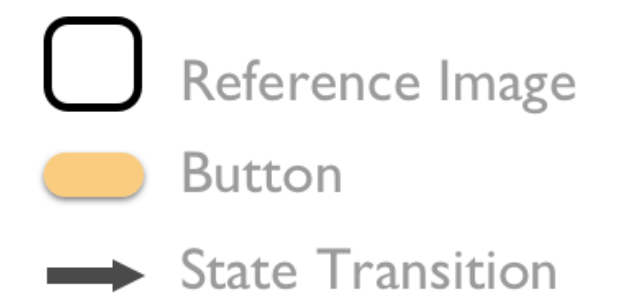
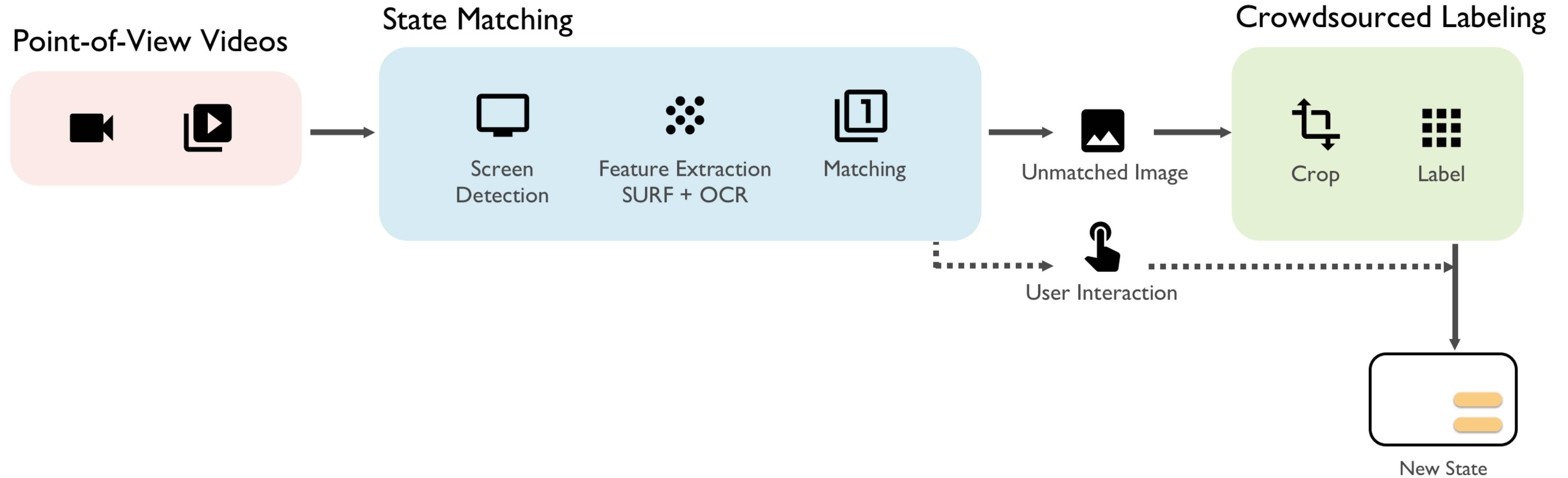
Sources of Usage Videos

- IoT and surveillance cameras
- Sighted volunteers using mobile/wearable cameras
- Manufacturers to share videos
- Online repositories of demo/tutorials

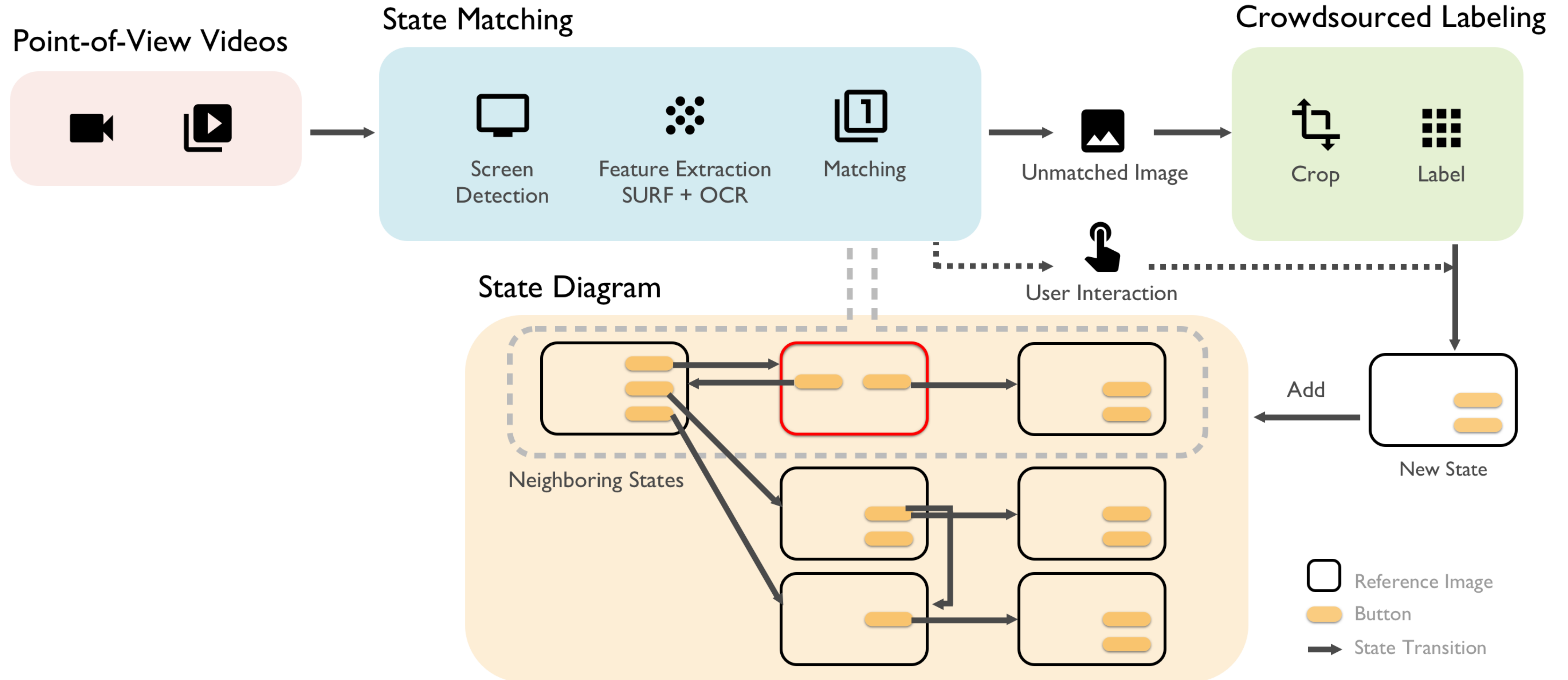
1. Generate the State Diagram



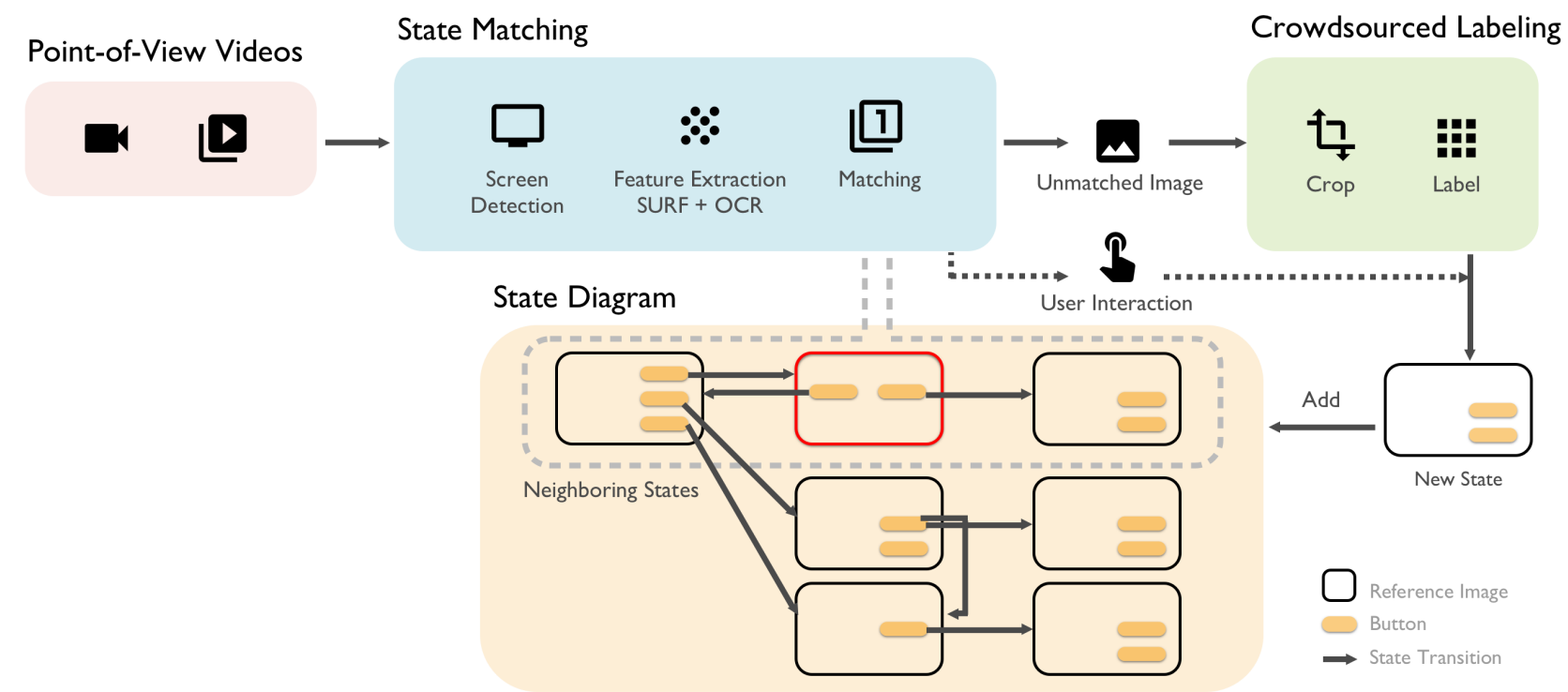
1. Generate the State Diagram



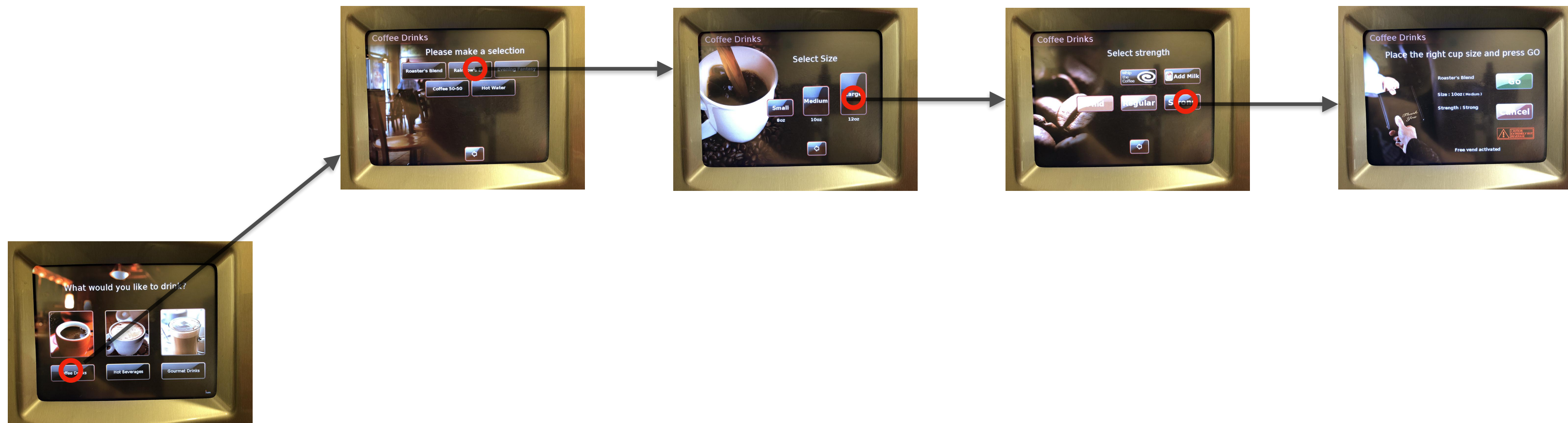
1. Generate the State Diagram



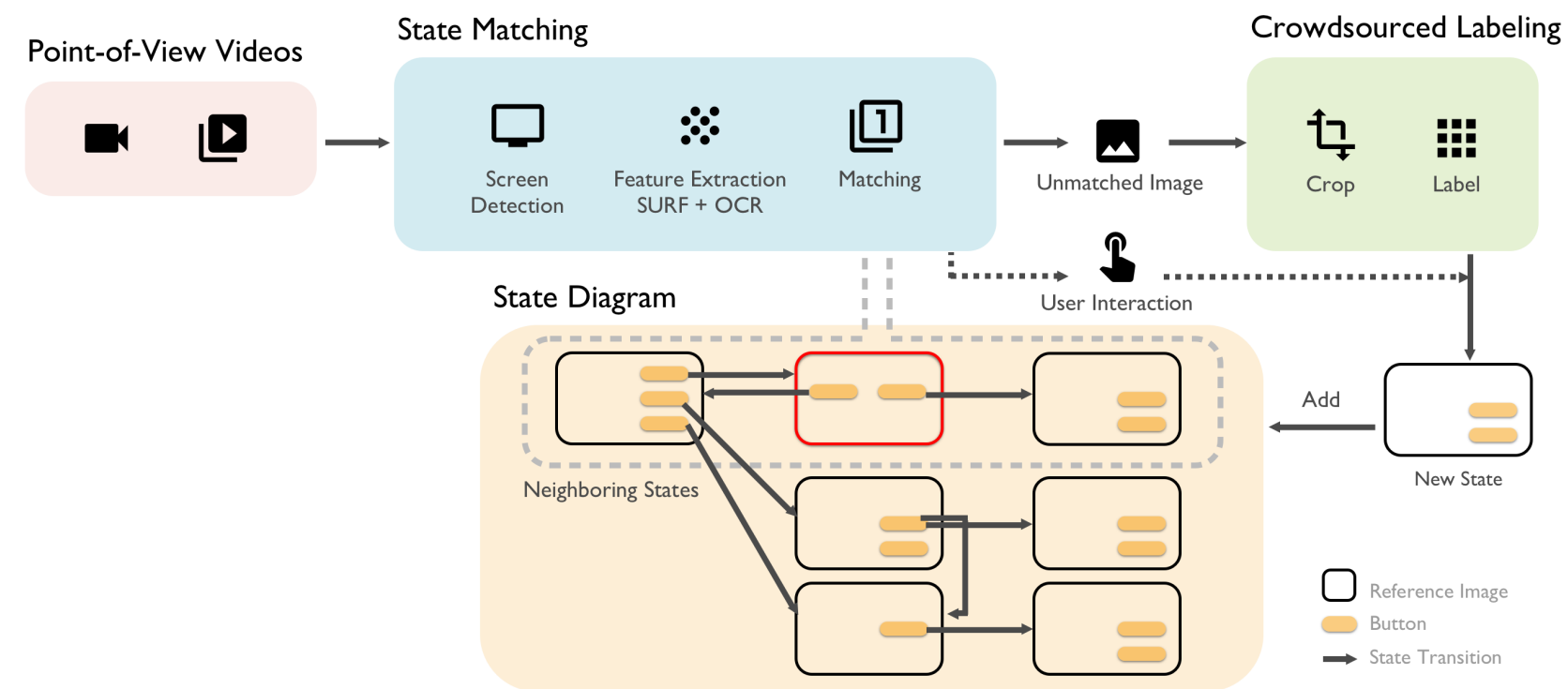
1. Generate the State Diagram



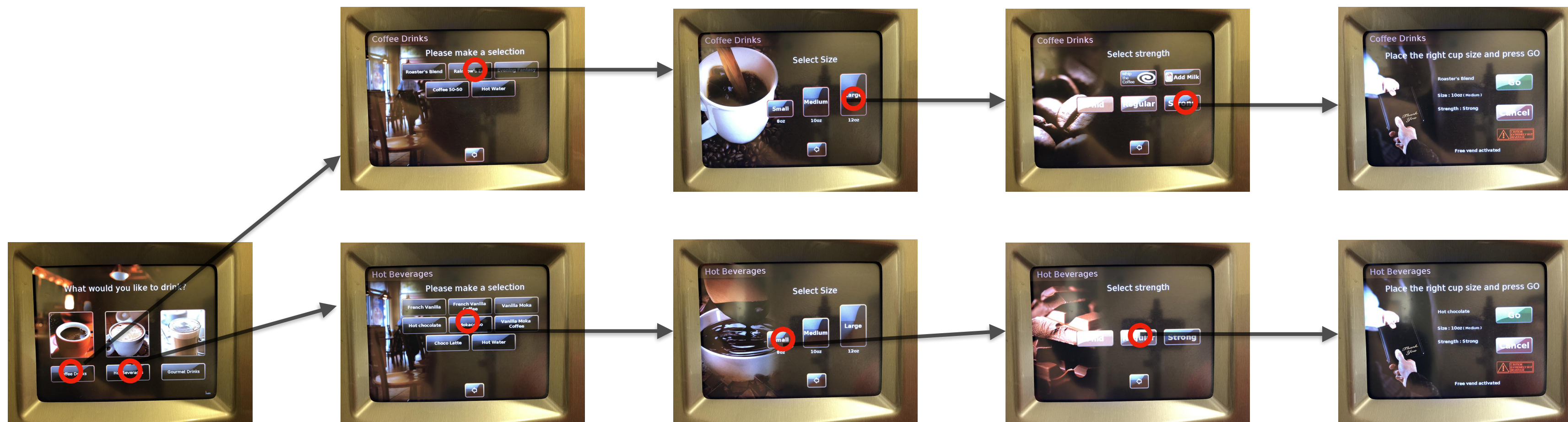
Video 1



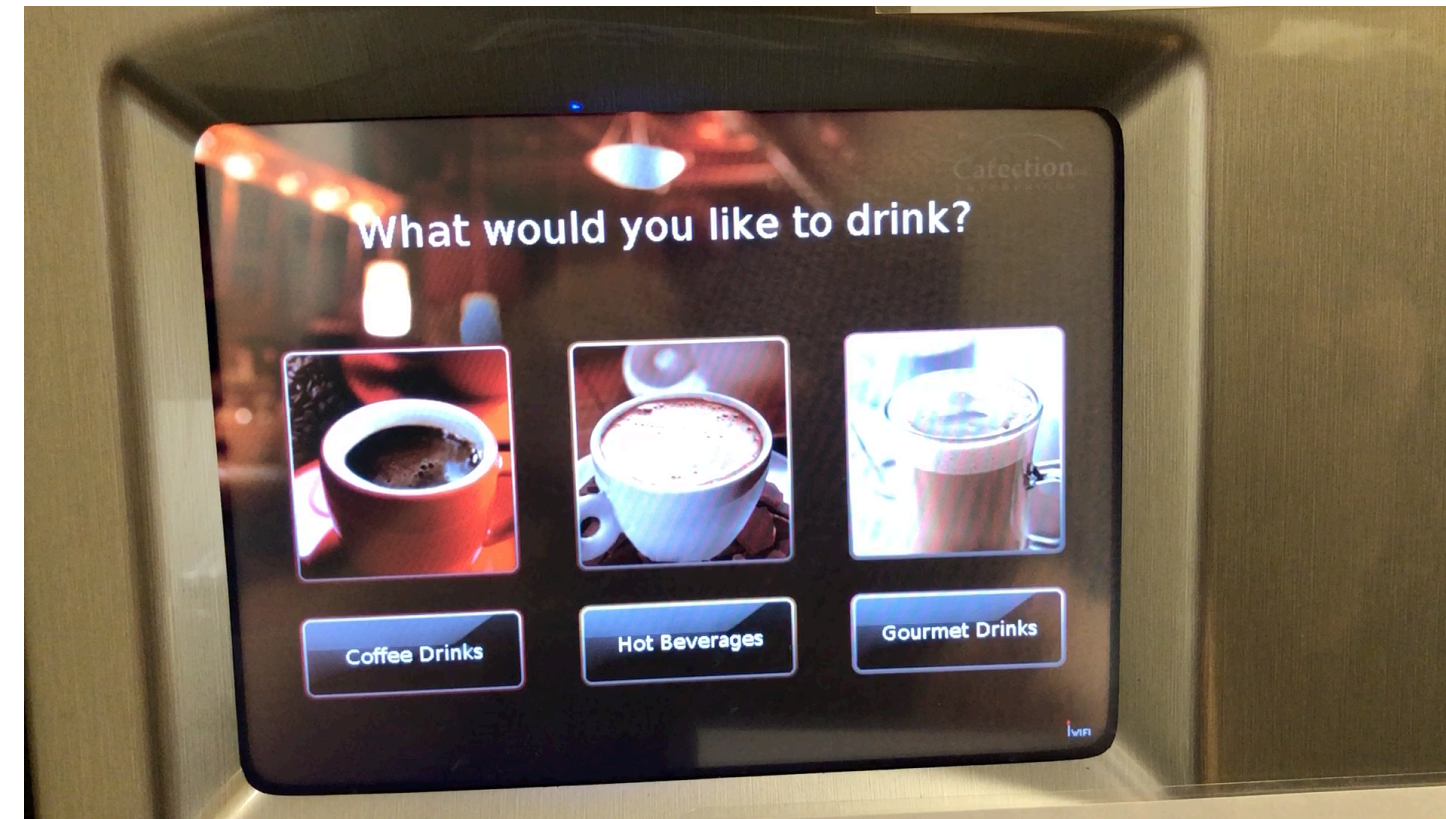
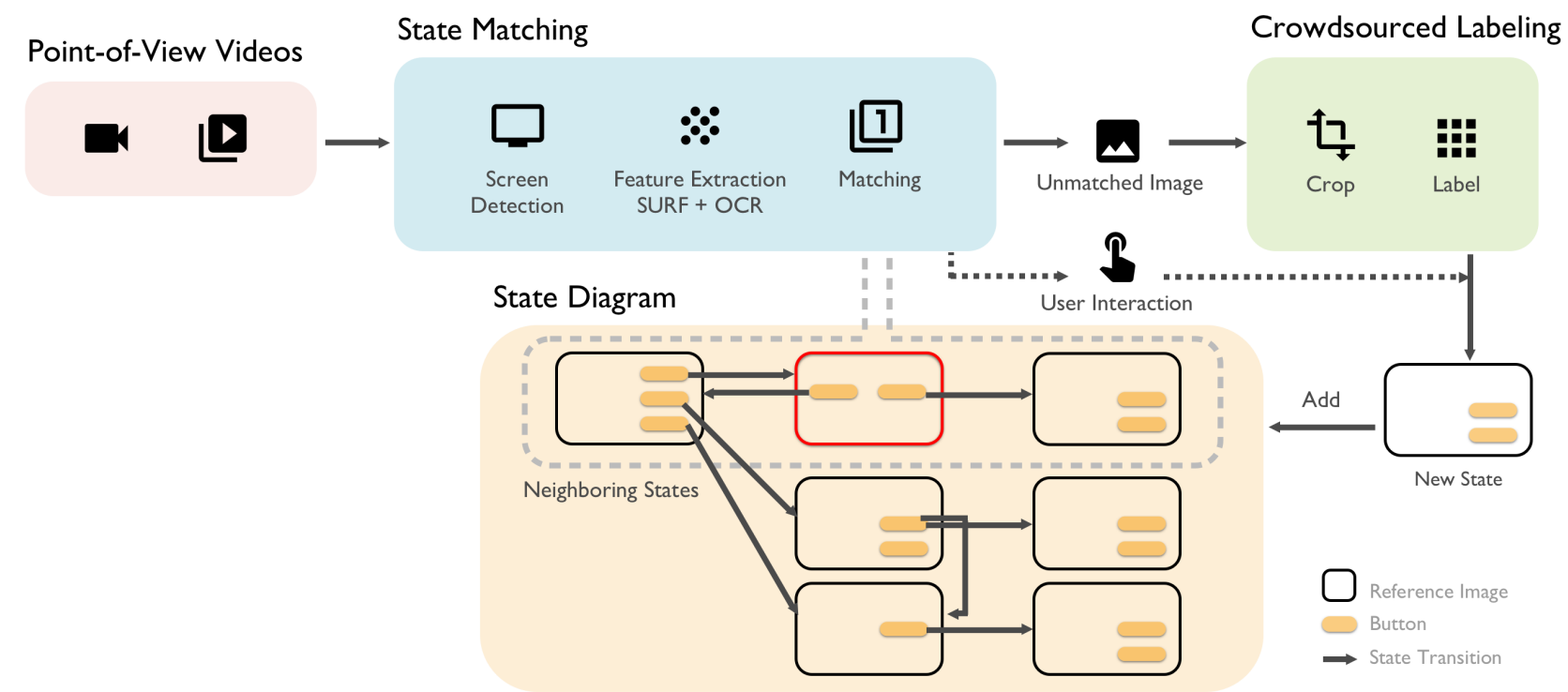
1. Generate the State Diagram



Video 2



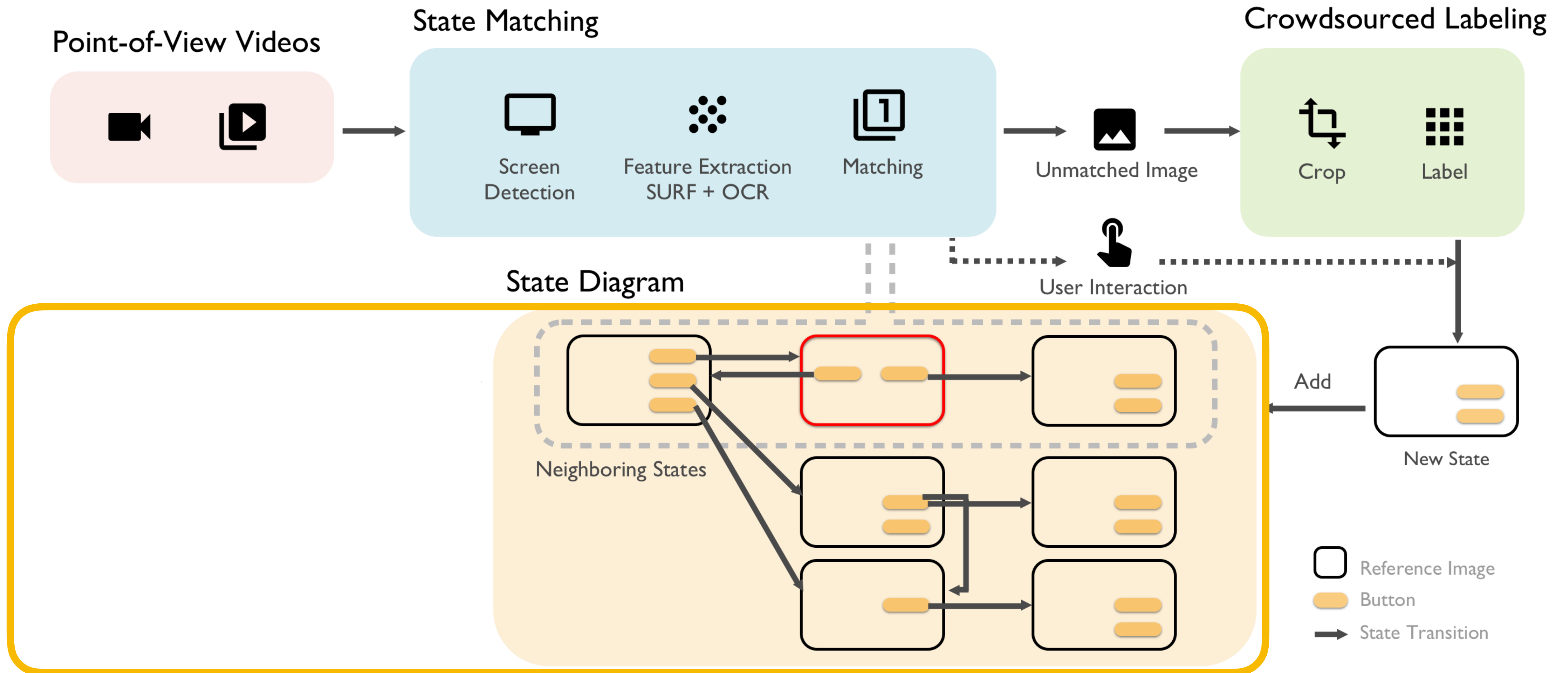
1. Generate the State Diagram



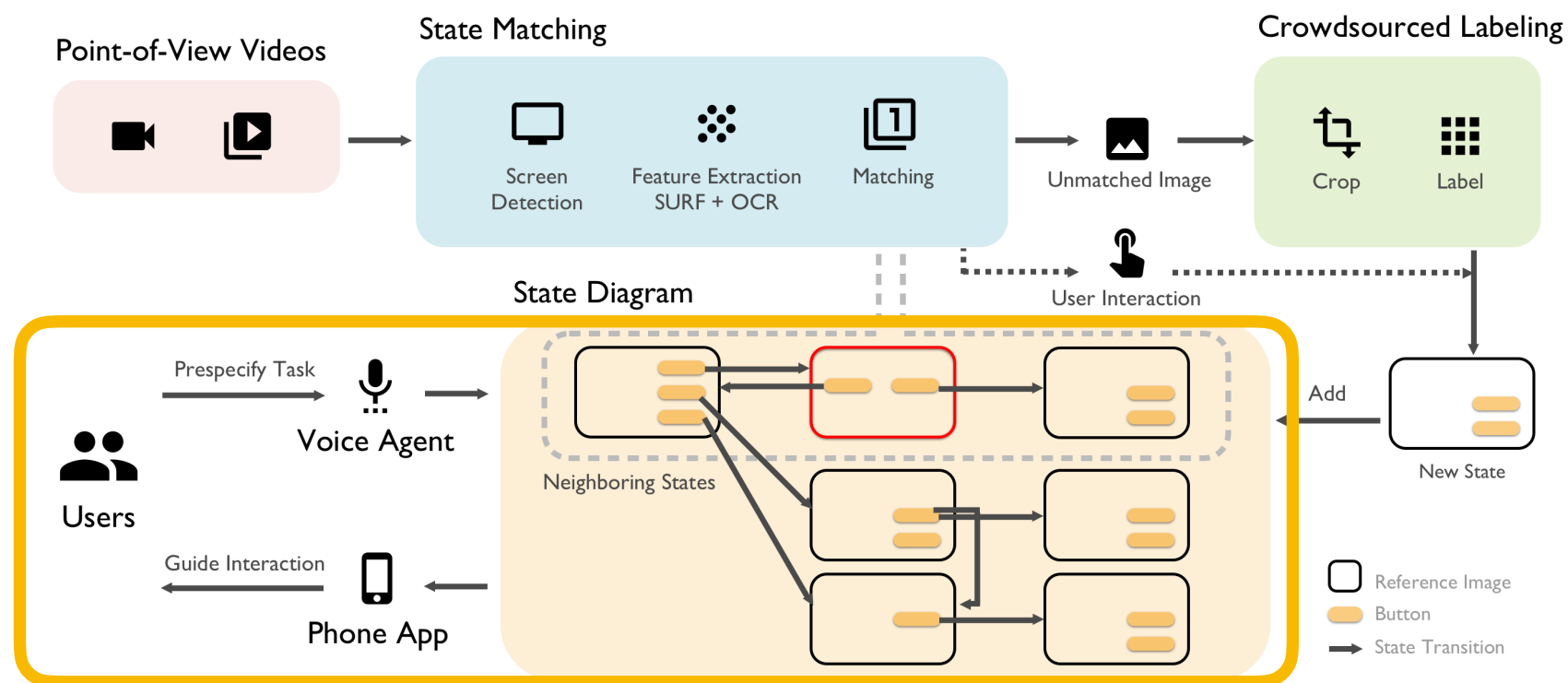
Video 3



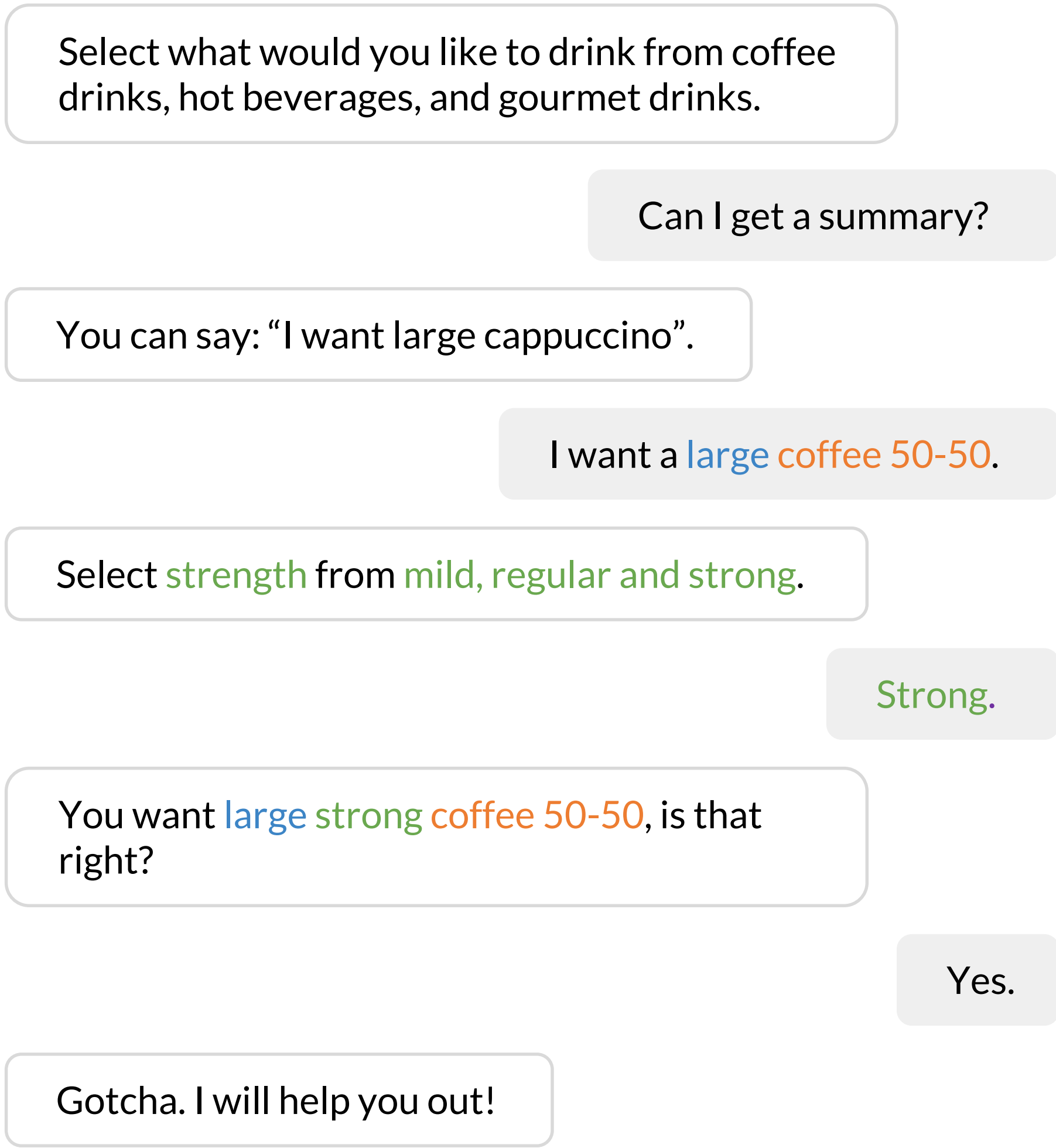
2. Access the State Diagram



2. Access the State Diagram



Conversation Agent Example - Coffee Machine



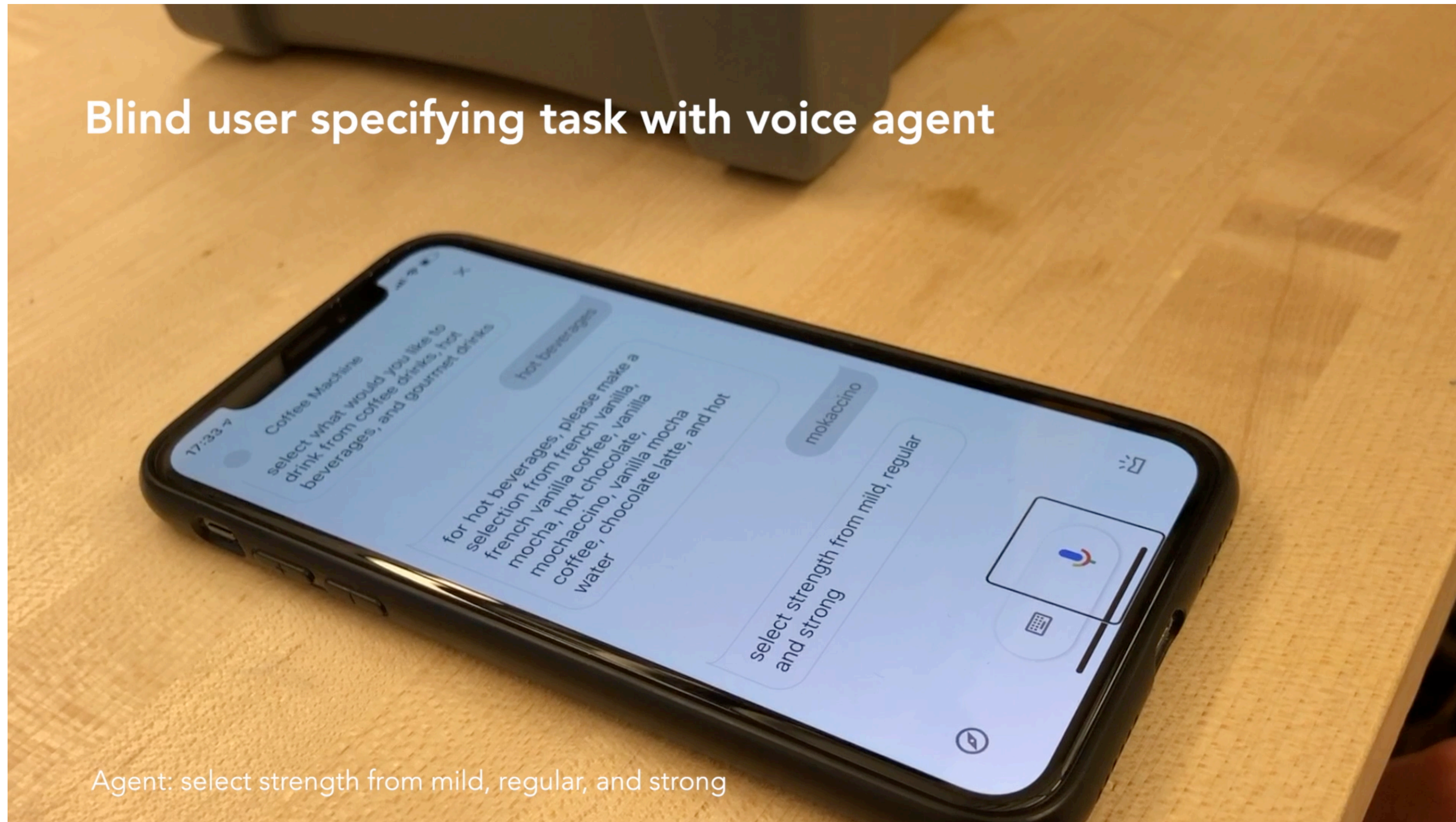
- 1 Welcome message from the initial state
- 2 Summary by aggregation
- 3 Parse required parameters:
size = large
coffee_type = coffee 50-50
- 4 Prompt missing parameter:
strength = ?
- 5 Ask for confirmation
- 6 Proceed to guidance

Conversational Agent

- Natural language queries
- Summary based on frequent usage

2. Access the State Diagram

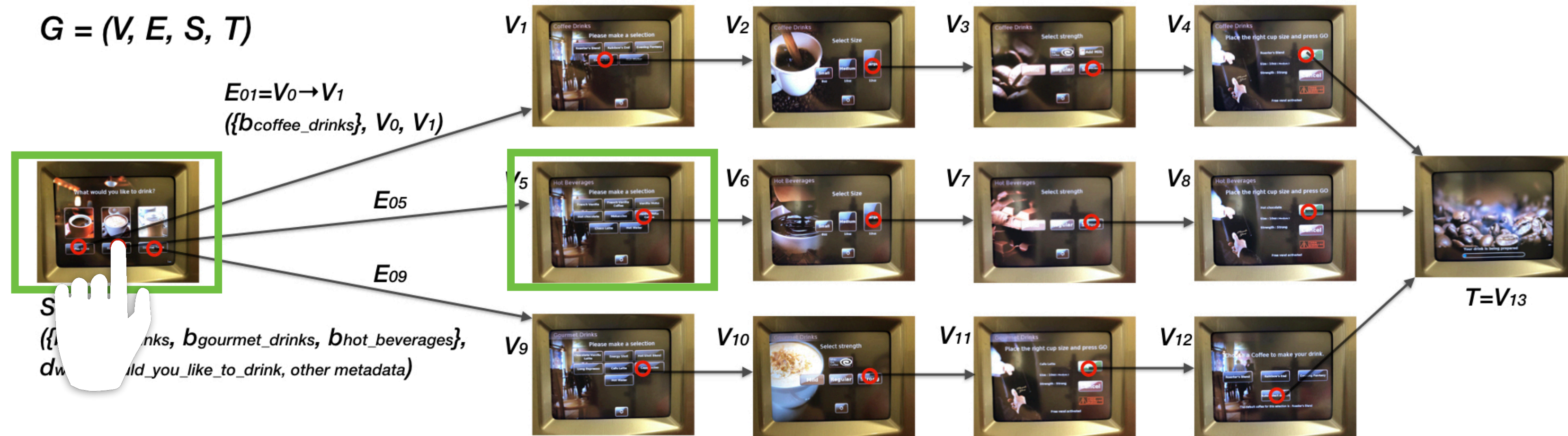
Blind user specifying task with voice agent



Agent: select strength from mild, regular, and strong

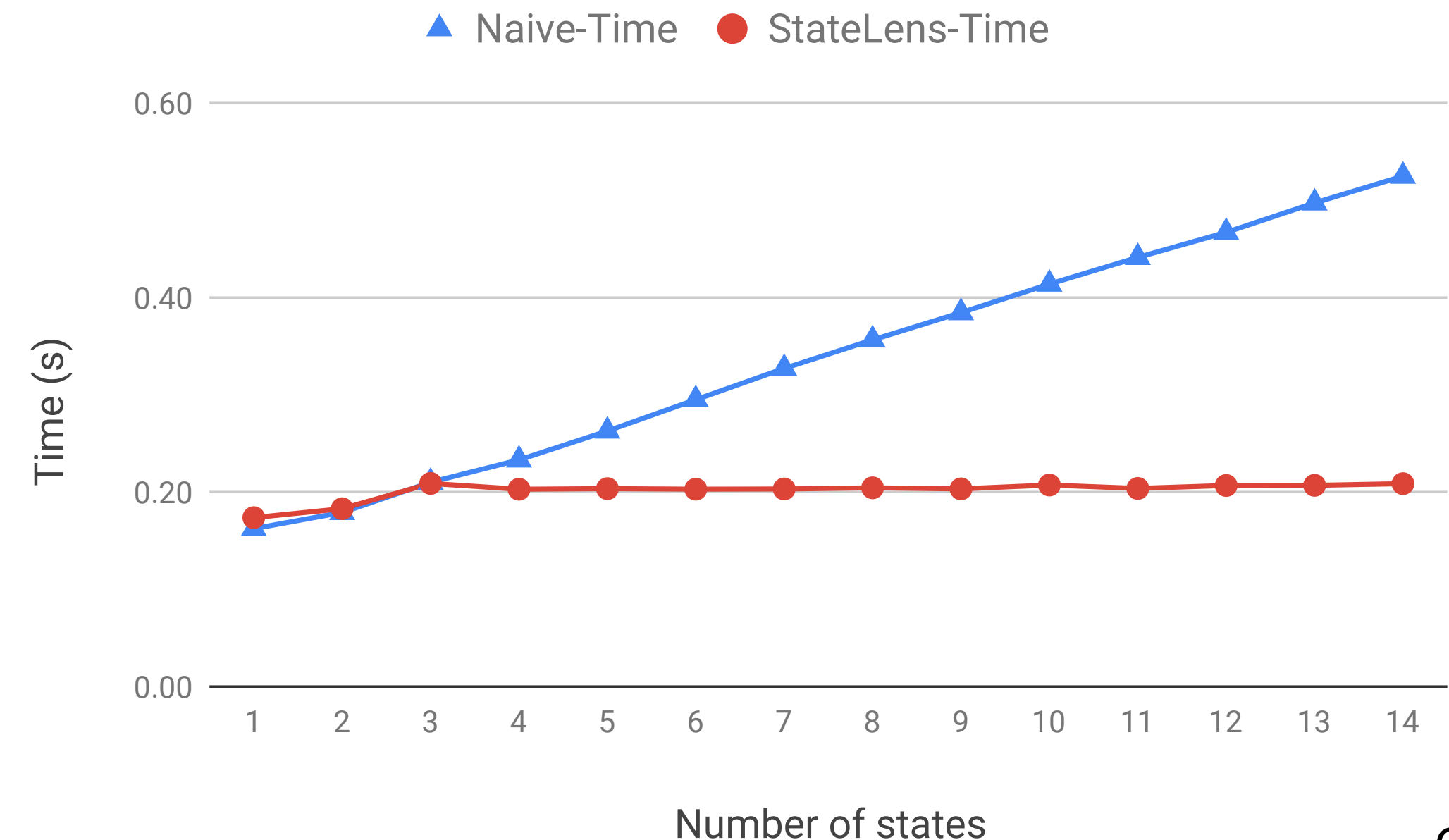
- **Drink type**
Hot beverages
- **Beverage type**
Mochaccino
- **Strength**
Strong
- **Size**
Large
- **Confirmation**

2. Access the State Diagram

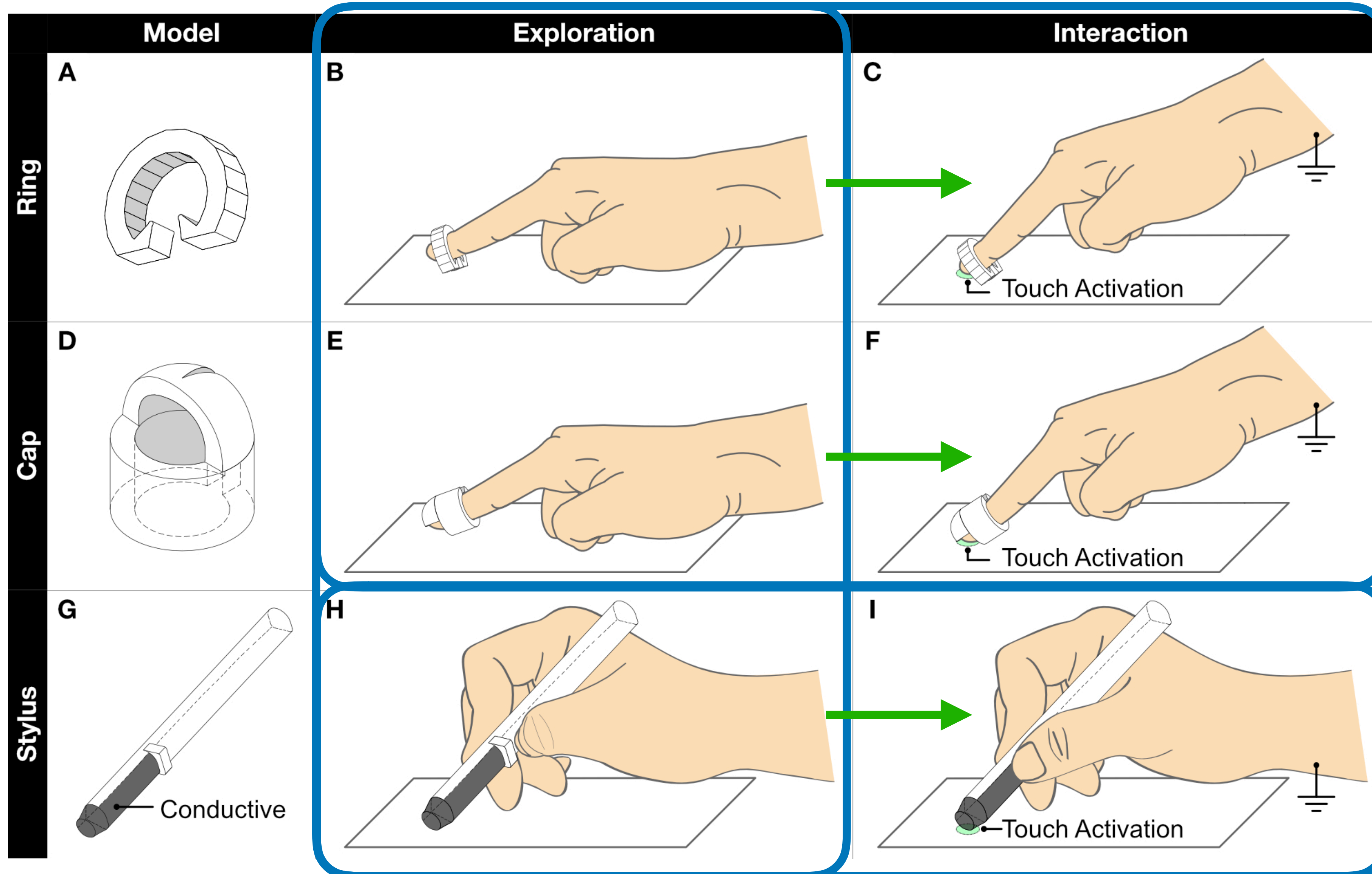


Identify States Efficiently and Robustly

- User's finger location for *expected state*
- *Neighboring states*
- Gradually *expand the search space* to others



3. Accessories for Risk-Free Exploration



Blind user completing task with app guidance

3D-printed Finger Cap for risk-free exploration

Blind user with wearable camera

Inaccessible Coffee Machine Mockup

StateLens iOS app: at large, press it... move down slowly...



Evaluation

- **Technical Evaluation:** 12 touchscreen devices (>70K frames), can accurately reconstruct interface structures from stationary, hand-held, and web videos
- **Feature Effectiveness:** ScreenDetect filtered out irrelevant frames in web videos, OCR retrieved states with only text changes, ScreenDetect+SURF+OCR the best



Evaluation

- **User Study:** 14 blind participants, the complete system successfully enables blind users to access inaccessible dynamic touchscreen devices

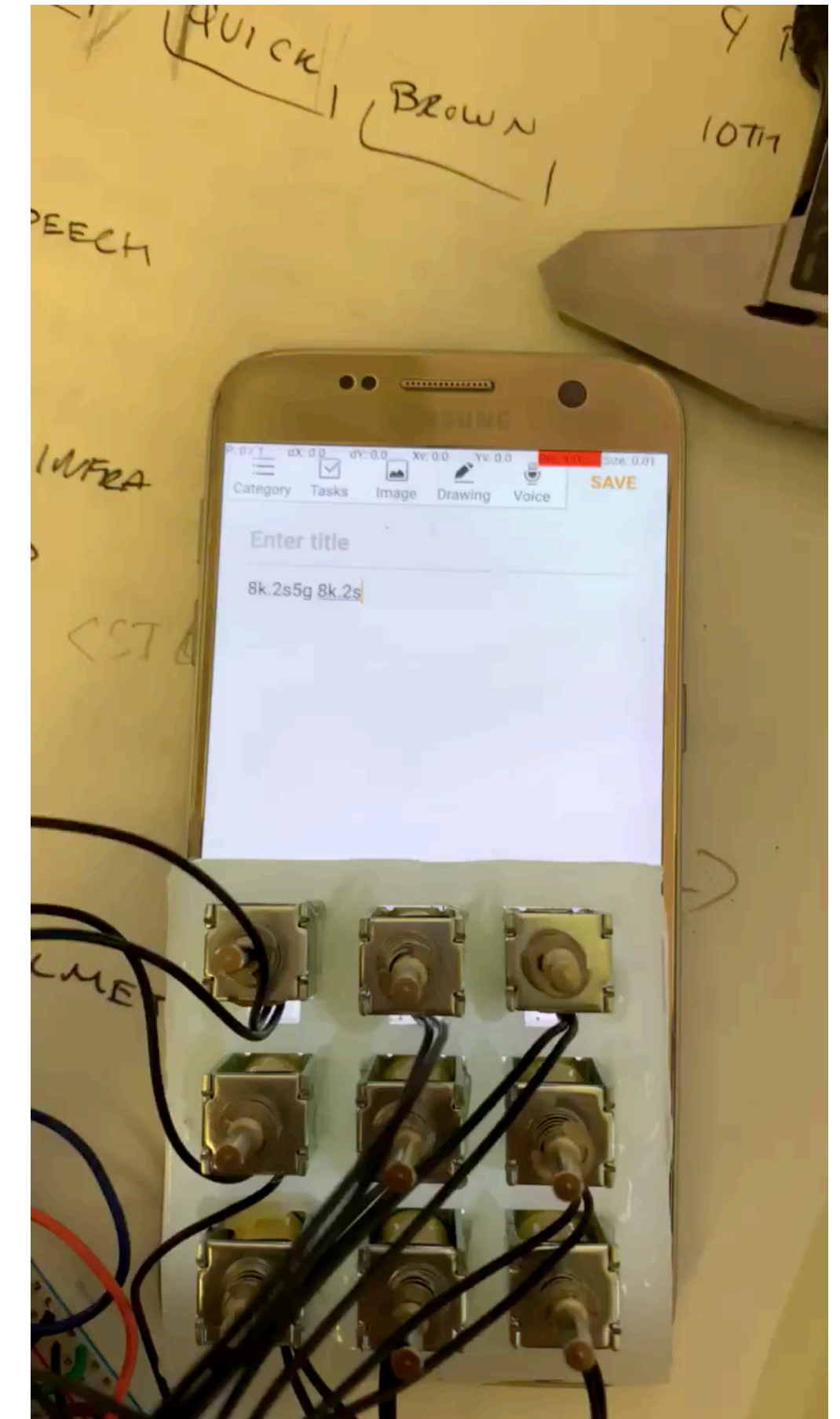
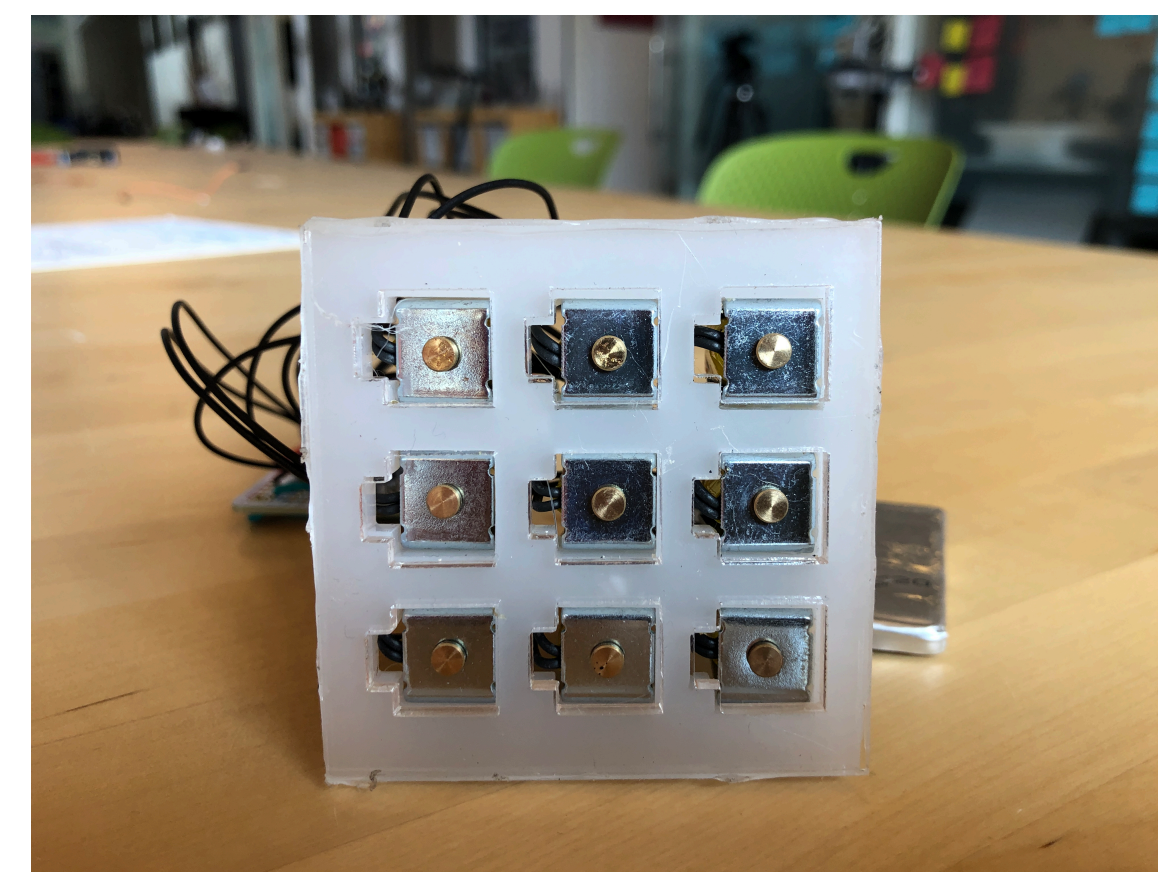
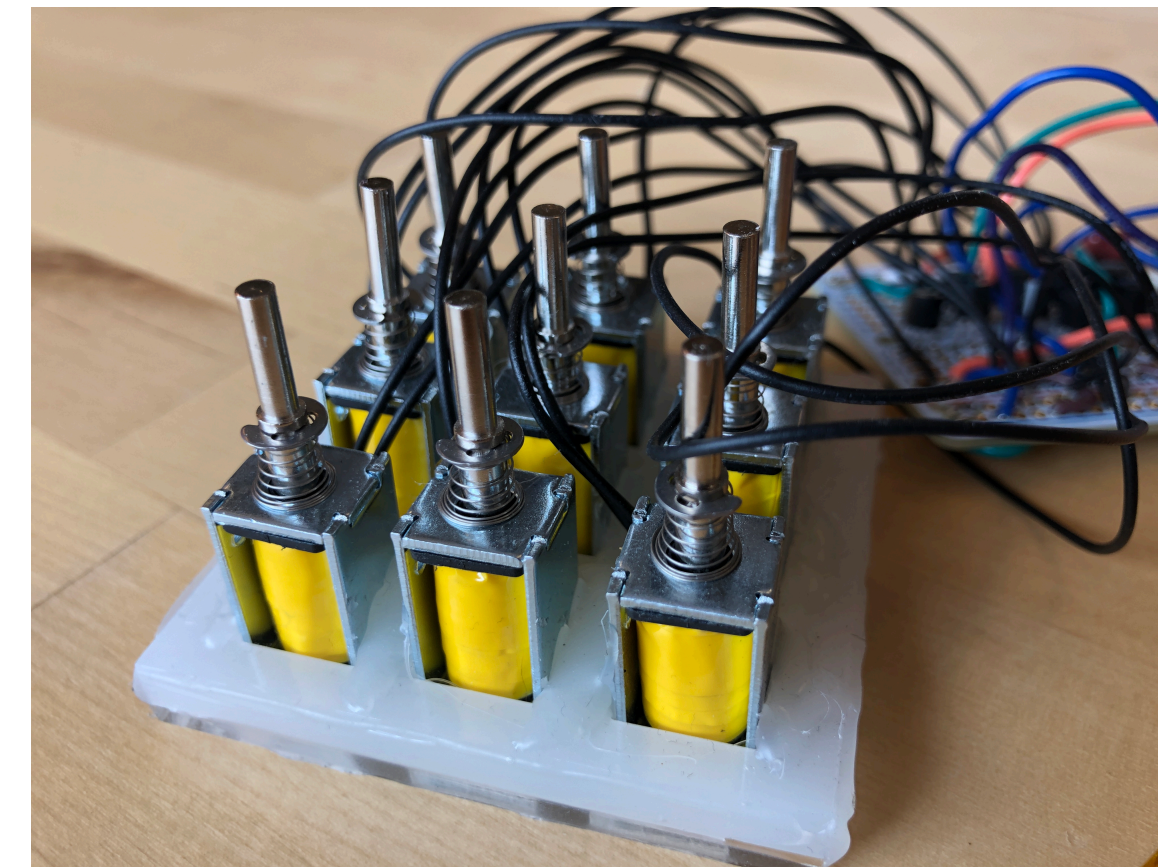
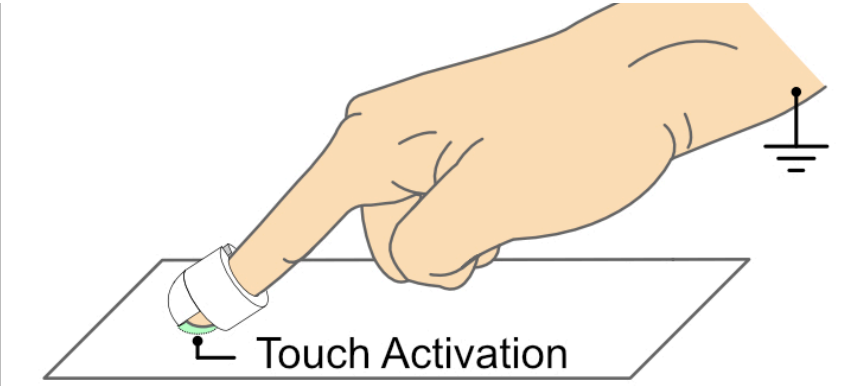
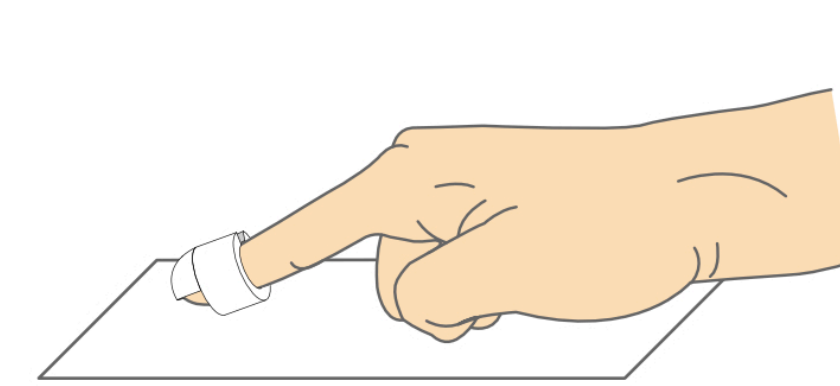
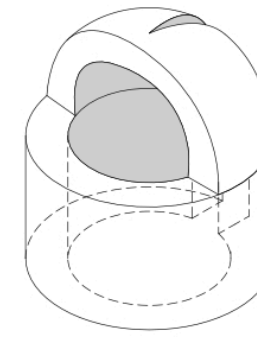
“*[StateLens] gives much more flexibility, so that if the machine itself doesn't have speech, this can cover the instances where you have to interact with a touchscreen. [With StateLens,] there are more tools to access them. This combination opens up more accessibility. (P6)*”

****StateLens is not the ideal solution!***

Future Work

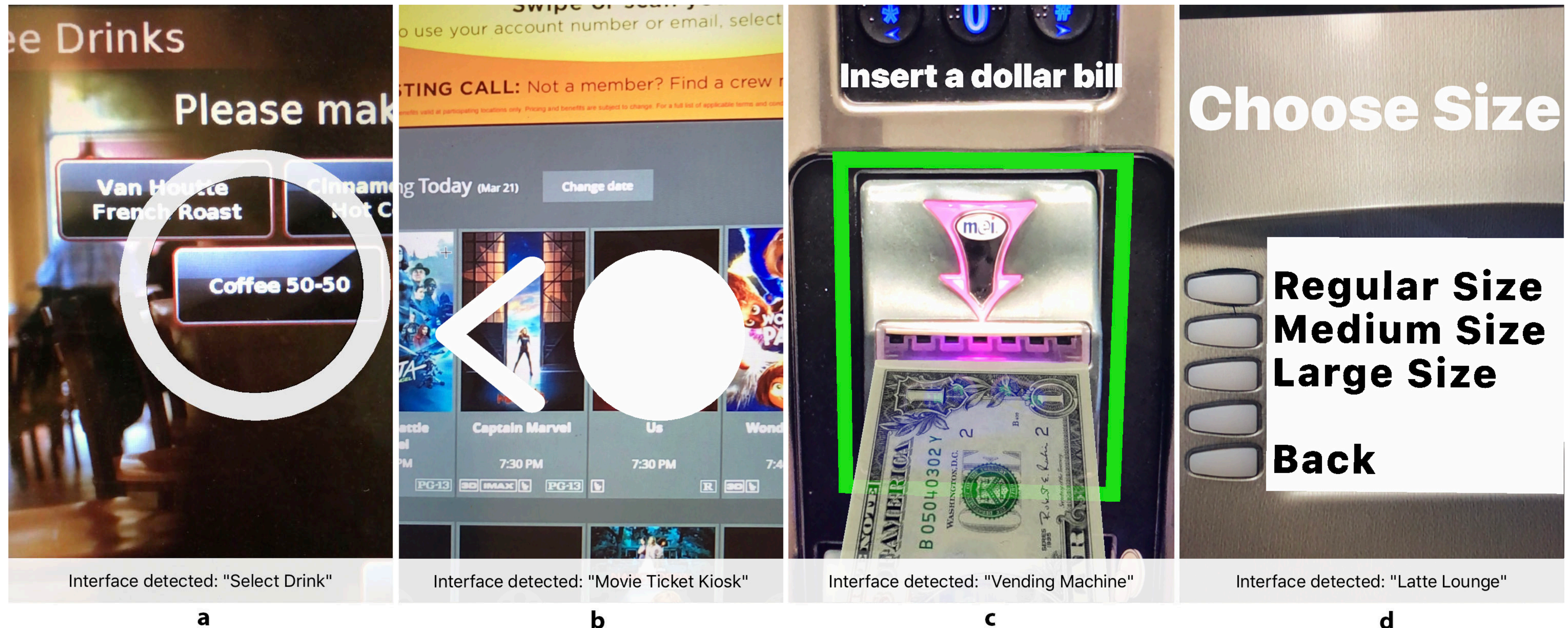
- Automatic Screen Actuation
 - 3D-printed accessories adds “risk-free exploration”
 - “the last (centi-)meter” problem
 - Hardware actuation proxies
 - Brushing interactions for automatic screen actuation

Cap



Future Work

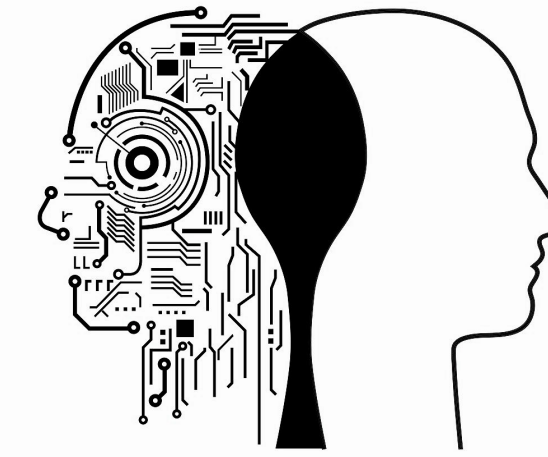
- Cognitive Assistance for Physical Interfaces



StateLens

- A **human-AI system** to make existing dynamic touchscreens accessible
- **Human:** contribute usage videos, interpret user interface
- **Machine:** state diagram, conversation agent to provide interactive guidance
- Broadly augment how people interact with touchscreen interfaces in the real world

Human-AI Systems



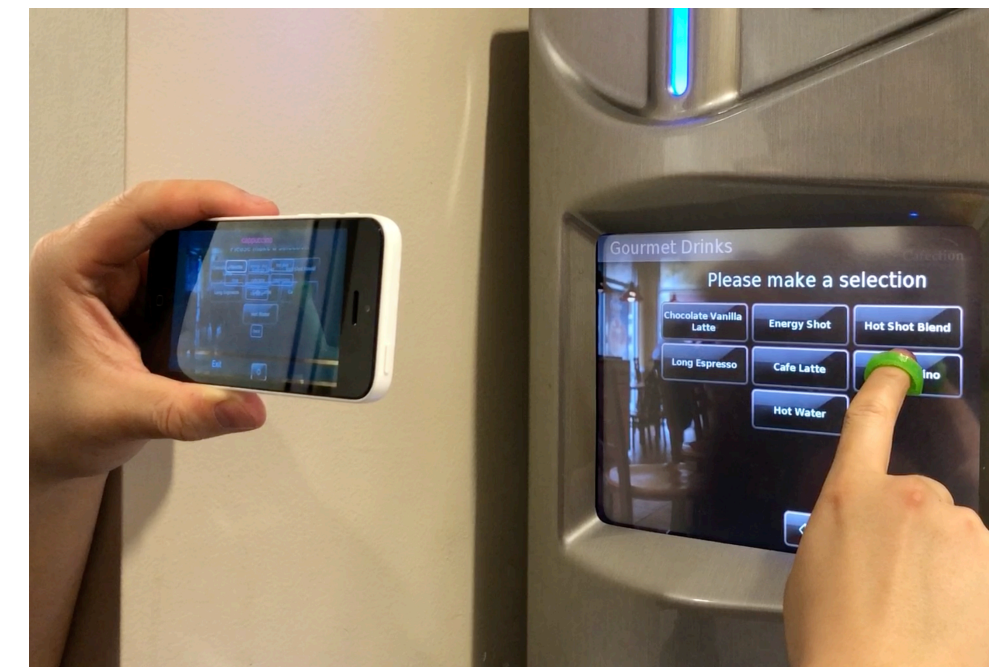
Physical Interfaces Accessibility



VizLens:
interactive screen reader
for physical interfaces
UIST 2016

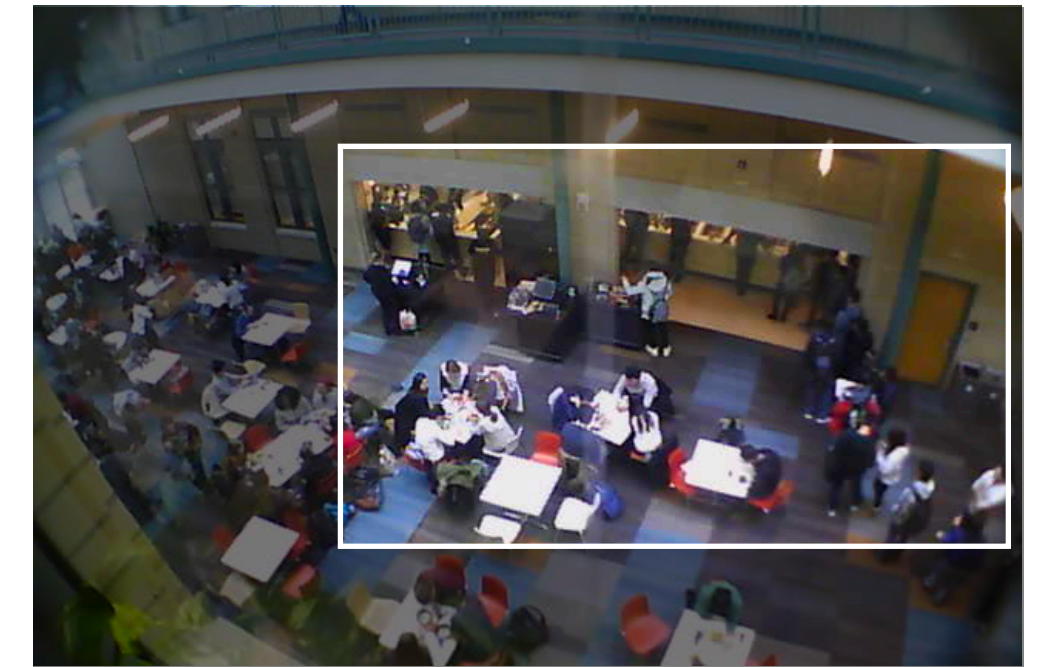


Facade:
auto-generating tactile
interfaces to appliances
CHI 2017



StateLens:
solution for existing
dynamic touchscreens
UIST 2019

Environmental Sensing



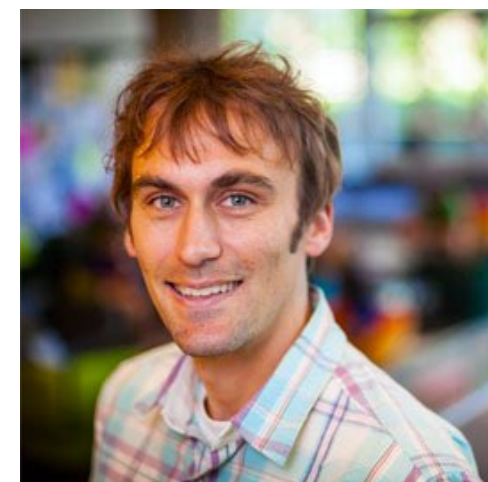
Zensors++:
camera sensing system to
answer real-world question
Ubicomp 2018

StateLens: A Reverse Engineering Solution for Making Existing Dynamic Touchscreens Accessible

Anhong Guo

*PhD candidate, Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University*

anhongg@cs.cmu.edu <https://guoanhong.com>



**Carnegie
Mellon
University**