Enabling Fine-Grained Permissions for Augmented Reality Applications With Recognizers

Suman Jana\(^1\)*, David Molnar\(^2\), Alexander Moshchuk\(^2\), Alan Dunn\(^1\)*, Benjamin Livshits\(^2\), Helen J. Wang\(^2\), and Eyal Ofek\(^2\)

\(^1\)University of Texas at Austin  
\(^2\)Microsoft Research  
* Work done during internship at Microsoft Research

USENIX Security 2013
Augmented Reality (AR): the new frontier!

SoundWalk app on Layar AR browser
http://www.layar.com/layers/clicmobiletestlayar/
Augmented Reality (AR) : the new frontier!

EA Sports Active 2
Augmented Reality (AR) : the new frontier!

Google Glass navigation
AR application pipeline

1. recognize objects
2. transform
3. render

You have arrived.
AR application pipeline

Recognize objects
Transform
Render

Current AR applications do it all by themselves!

Topic of today’s talk
All these stages need platform support (HotOS’13)
Privacy concern: unrestricted access
Privacy concern: unrestricted access

all apps see this

tons of sensitive information!
Functionality concern: one app at a time

AR tour guide

AR navigator

can’t run concurrently
Scalability concern: must scale to multiple concurrent applications
What can we do with today’s abstraction?

- A lot of repetitive work across apps

2x overhead for 2 Kinect Apps

- Hard to maintain usable frame rate

Diagram:

1. **App 1**
   - Recognize objects
   - Transform
   - Render

2. **App 2**
   - Recognize objects
   - Transform
   - Render
RECOGNIZERS: A NEW ABSTRACTION
Introducing recognizers

- A new platform abstraction to recognize real-world objects – generates events that apps can subscribe to
- Fine-grained control over detected objects
- Can enforce least privilege – each app must obtain permissions to access each recognizer
Scalability with multiple recognizers

Explicit data flow

Sample directed acyclic graph of recognizers

- RGB
- Face
- 3D model
- Depth
- Skeleton
- Hand
- Foot
Benefits of explicit dataflow

Face
RGB
Depth
Skeleton
Hand
Foot

3D model
offload to cloud

better scheduling
lazy invocation
## How do recognizers help?

<table>
<thead>
<tr>
<th>Concerns</th>
<th>Role of recognizers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy concern</td>
<td>Recognizer based permissions allow enforcing least privilege</td>
</tr>
</tbody>
</table>
### How do recognizers help?

<table>
<thead>
<tr>
<th>Concerns</th>
<th>Role of recognizers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy concern</td>
<td>Recognizer based permissions allow enforcing least privilege</td>
</tr>
<tr>
<td>Scalability concern</td>
<td>Recognizers allow computation sharing across apps</td>
</tr>
<tr>
<td></td>
<td>Offload individual recognizers</td>
</tr>
</tbody>
</table>
Explaining recognizers: privacy goggles

- Visual way to explain information given by each recognizer to an application

Privacy goggles for skeleton recognizer
Recognizer-based AR architecture

- App subscribes to skeleton recognizer
- AR platform delivers recognizer events to app
- AR platform
- Privacy goggles
- Recognizers
IMPLEMENTATION
## Example applications

<table>
<thead>
<tr>
<th>Application</th>
<th>What it does</th>
<th>Recognizers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hand cursor</td>
<td>Control cursor with hand movements</td>
<td>Skeleton</td>
</tr>
<tr>
<td>Facial movement detector</td>
<td>Visualize tracked faces</td>
<td>Face detector</td>
</tr>
<tr>
<td>Room Scanner</td>
<td>Find flat surfaces</td>
<td>3D Model, Depth</td>
</tr>
</tbody>
</table>
EVALUATION
Evaluation criteria

• Privacy: How many applications need access to raw video and sensor data?

• Scalability
  • Performance of concurrent applications
  • Performance of outsourced AR computation

• Usability
  • Is the information released less sensitive than raw data?
  • Do users understand privacy goggles?
Few apps need raw data

Only 4 recognizers cover ~90% of shipping Xbox apps

<table>
<thead>
<tr>
<th>Recognizer</th>
<th>% Apps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skeleton</td>
<td>94.3%</td>
</tr>
<tr>
<td>Person Texture (PT)</td>
<td>25.3%</td>
</tr>
<tr>
<td>Voice Commands (VC)</td>
<td>3.44%</td>
</tr>
<tr>
<td>Hand Position (HP)</td>
<td>5.74%</td>
</tr>
<tr>
<td>Video Clip</td>
<td>3.4%</td>
</tr>
<tr>
<td>Picture Snap</td>
<td>1.1%</td>
</tr>
<tr>
<td>Voice Intensity</td>
<td>1.1%</td>
</tr>
<tr>
<td>Voice Modulation</td>
<td>1.1%</td>
</tr>
<tr>
<td>Speaker Recognition</td>
<td>1.1%</td>
</tr>
<tr>
<td>Sound Recognition</td>
<td>1.1%</td>
</tr>
<tr>
<td>Basketball Tracking</td>
<td>1.1%</td>
</tr>
<tr>
<td>Skeleton+PT+VC</td>
<td>82.75%</td>
</tr>
<tr>
<td>Skeleton+PT+VC+HP</td>
<td>89.65%</td>
</tr>
</tbody>
</table>

recognizers used by 87 shipping Xbox Applications
Released information less sensitive

10 surveys with 50 respondents each about recognizer output

Consider the two pictures below. Which picture contains “more sensitive” information?

86% of the users said the left one is more sensitive
Privacy goggles communicate

- 152 respondents
- 80% identified that the app could see body position
- 47% identified that the app could see hand positions
Sharing recognizers works!

> 25 fps for up to 6 apps
Recognizer offloading

Kinect fusion recognizer needs an extremely powerful GPU to run

doesn’t work

4.46 fps
Recognizer offloading

Kinect fusion recognizer needs an extremely powerful GPU to run

4.17 fps

offloaded recognizer

4.46 fps
Minimizing recognizer false positives

- Recognizers are not perfect (yet) - false positives can lead to information leakage

AR platforms can apply simple heuristics like combining multiple recognizers to decrease false positives

OpenCV face recognizer

OpenCV face recognizer & Kinect depth filter
Summary

- New AR paradigm needs new platform abstractions
- Recognizers
  - Help in enforcing least privilege
  - Allow sharing of computation across apps
  - Allow efficient offloading of heavyweight computation
  - False positives in recognizers can be dealt using heuristics at the platform level
- Privacy goggles - visual permission management
Future work

- recognize objects
- transform
- render

tons of exciting new stuff! come talk to us if interested
Thanks!