PREPOSE: Security and Privacy for Gesture-Based Programming

Lucas Silva Figueiredo  
Federal University of Pernambuco

Benjamin Livshits, David Molnar, and Margus Veanes  
Microsoft Research
GESTURE SENSORS

Microsoft Kinect V1
24M - 2013

Microsoft Kinect V2
20M - 2016
VR + GESTURE SENSING
AR + GESTURE SENSING
GESTURE SENSORS — ALL SORTS AND SIZES

- Microsoft Kinect V1
- Microsoft Kinect V2
- Asus Xtion Pro Live
- Creative Senz3D
- Intel SR300
- Intel R200
- Project Tango
- Structure.IO
- Leap Motion
- DUO M
- DUO MLX
- DUO MC
PROBLEM: PRIVACY
APPLICATIONS CAN ACCESS SENSITIVE DATA
PROBLEM: SAFETY

Just because you can ask the user to make a gesture doesn't mean you should.

Prepose can check if an app asks the user to move in a dangerous way.
PROBLEM: CONFLICTS

Do any apps try to override the system attention gesture?

Could any **two gestures** be triggered at the same time?
// Punch Gesture
if ( vHandPos.z - vShoulderPos.z > fThreshold1 &&
    fVelocityOfHand > fThreshold2 ||
    fVelocityOfElbow > fThreshold3 &&
    DotProduct( vUpperArm, vLowerArm ) > fThreshold4 )
{
    bDetect = TRUE;
}

PROBLEM: BUILDING NEW GESTURE RECOGNIZERS

Machine Learning Approaches

- Modeling Variations
- ML Knowledge
- Data Set Size
- Recording and Tagging
- Editing the Gesture
- Recognition Time
- Detailed Feedback
- Conflicts
PREVIOUS APPROACHES

Limited access (Jana et al. 2013 x2, Vilk et al. 2015)

- Color frames
  - Visuals of the environment and users
- Depth frames
  - Shapes which can be used to automatically detect objects and furniture
- Silhouettes
  - Presence, clothing, weight
- Skeletons
  - Processed information about the user

Add noise (Jana et al. 2013)
PREPOSE CONCEPT

Prepose interpreter and runtime

MSR Z3 constraint solver

Trust boundary

Prepose Code

Gesture Events

App 0
App 1
App 2
App 3
App 4

Skeleton
EXAMPLE: THERAPY

- Feedback
- Gesture description
- Repetitions

Raise your left leg to the side and return.
You need to keep your left leg straight.

2/5 reps
Hip Abduction
OUR APPROACH

raise your left leg to the side

action

body part

direction
UNDERSTANDING A PHRASE IN PREPOSE

rotate your left elbow 20 degrees to your left,
Our prototype hard codes verb implementations, but they could be sub-gestures created manually.

"20 degrees" is an argument to the function "rotate".

"left" is an argument to the function "rotate."

actions are functions of the recognized skeleton.

rotate your left elbow 20 degrees to your left,
OUR WORK: PREPOSE

GESTURE fourth-position-en-avant:

POSE cross-legs-one-behind-the-other:
  put your left ankle behind your right ankle,
  put your left ankle to the right of your right ankle.

POSE high-arc-arms-to-right:
  point your arms down,
  rotate your right arm 70 degrees up,
  rotate your left elbow 20 degrees to your left,
  rotate your left wrist 25 degrees to your right.

EXECUTION:
  cross-legs-one-behind-the-other,
  high-arc-arms-to-right.
CHECKING SAFETY WITH Z3

Safety Check
- Spine Constraint
- Elbows Constraint
- Shoulders Constraint
- ...

assume that it is safe

GESTURE gesture: EXECUTION:
- step1,
- step2.

Body
- head x,y,z
- neck x,y,z
- spine x,y,z
- ...

ask if it is still safe
if yes
Do any apps try to override the system attention gesture?

Could any two gestures be triggered at the same time?
## EVALUATION

### Expressiveness

<table>
<thead>
<tr>
<th>Application</th>
<th>Gestures</th>
<th>Poses</th>
<th>LOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Therapy</td>
<td>12</td>
<td>28</td>
<td>225</td>
</tr>
<tr>
<td>Ballet</td>
<td>11</td>
<td>16</td>
<td>156</td>
</tr>
<tr>
<td>Taichi</td>
<td>5</td>
<td>32</td>
<td>314</td>
</tr>
</tbody>
</table>

### Speed

<table>
<thead>
<tr>
<th>Performance</th>
<th>Online</th>
<th>Offline</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 FPS</td>
<td>Matching time: 3 ms</td>
<td>Safety: 725 ms</td>
</tr>
<tr>
<td></td>
<td>Pose transition: 91 ms</td>
<td>Conflicts: 52 sec</td>
</tr>
</tbody>
</table>
CONCLUSION

1) Prepose: a language for expressing gestures
2) A runtime for creating always-on training applications using Prepose
3) Support for user privacy built into system design
4) An approach to checking gesture safety and conflicts between gestures
5) Ability to check gestures and their composition using Z3 at app submission
FUTURE WORK

New actions/primitives to improve expressiveness
Make use of time based constraints (slowly, fastly)
Reusable gestures as assets on a store
Example applications
Add sensors: Leap Motion, Intel Real Sense, etc
PrePose Language

// real case scenario
APP therapy_gestures:

// precision and body planes as coordinate system
GESTURE shoulder_abduction:
  POSE arm_down:
    point your right arm down,
    rotate your right arm 20 degrees to your right.

// poses as pure transformations from previous state
GESTURE arm_down,
  POSE rot_up:
    rotate your right arm 20 degrees up.

  POSE rot_down:
    rotate your right arm 20 degrees down.

EXECUTION:
  arm_down,
  //using repetitions to get further
  rot_up,
  rot_up,
  rot_up,
  rot_up,
  rot_up,
  rot_down,
  rot_down,
  rot_down,
  rot_down.

shoulder_abduction

arm_down  rot_up  rot_up  rot_up  rot_up  rot_up  rot_down  rot_down  rot_down  rot_down  rot_down  rot_down

FPS = 30.0 Time = 00:10:50.1354902
completed 2 times
DOWNLOAD IT AT

github.com/microsoft/prepose

We take pull requests!
GESTURE RECOGNITION POWERS KINECT APPS
GESTURE SENSORS

Microsoft Kinect V1
24M - 2013

Microsoft Kinect V2
20M - 2016

Depth View:   Skeletal View:

Frames per Second:
30
PROBLEM: PRIVACY IS A MAJOR CONCERN

Privacy concerns threaten to overshadow Microsoft's new console
By Brian Crecente on Jun 05, 2013 at 11:14a @crecenteb

Xbox One will still function, even without Kinect plugged in (update)
By Dave Tech on Aug 12, 2013 at 5:42p @dottach

The Xbox One brings with it a required peripheral packed with microphones and cameras that can monitor a person's every
KINECT ENABLES NATURAL USER INTERFACE (NUI)

Microsoft Kinect V2

20M - 2016
SMART HOUSES/WORKPLACES
PREPOSE CONCEPT

Prepose interpreter and runtime

Skeleton

Trust boundary

Domain Specific Language

Gesture Events

App 0
App 1
App 2
App 3
App 4
EXPRESSIVENESS: BALLET

Performance dance

Body precision

Reusable partial poses

APP ballet:

GESTURE first-position:
POSE stand-straight:
  point your spine, neck and head up.
POSE point-feet-out:
  point your right foot right,
  point your left foot left.
POSE stretch-legs:
  align your left leg,
  align your right leg.
//POSE put-ankles-together:
//connect your right ankle to your left ankle.
POSE low-arc-arms:
  point your arms down,
  rotate your elbows 15 degrees up,
  rotate your left wrist 5 degrees to your right,
  rotate your right wrist 5 degrees to your left.
//POSE first-position-composed:
  //stand-straight,
  //point-feet-out,
  //stretch-legs,
  ///</POSE put-ankles-together,
  //low-arc-arms.
EXPRESSIVENESS: TAICHI

Martial arts

Complex gestures

Relation between joints extensively explored

Time constraints

APP taichi_gestures:

GESTURE starting:

POSE stand_straight:
point your spine, neck and head up.

POSE starting_legs:
point your legs down,
rotate your right leg 10 degrees to your right,
rotate your left leg 10 degrees to your left,
point your feet to your front.

POSE starting_arms:
point your arms down.

POSE inhale_arms:
point your arms to your front.

POSE transition_arms:
put your left wrist in front of your left shoulder,
put your right wrist in front of your right shoulder.
//put your left wrist near to your left shoulder,
//put your right wrist near to your right shoulder,
//move your spine should 10 centimeters to your back.

POSE bend_your_knees_slightly:
don't align your left knee and your left ankle,
don't align your right knee and your left ankle.
CONCLUSION

Novel way to write gesture recognizers
Alternative to existing machine learning approaches
Architecture protects privacy
Expressive enough for a wide class of gestures
Resulting recognizers run quickly
First gesture recognizers with precise static analyses
COMPLEMENTARY TO MACHINE LEARNING

- Declarative
- Close to English
- Easy to understand and refine
- Fast coding and prototyping
- Reusable components
- Open to different interpretations of the underlying actions

Analysis Power
- Safety
- Conflict Detection
- Target Generation
FOCUS: TRAINER APPLICATIONS

Always on

Monitors user’s progress

Only needs to report progress or completion

May run concurrently (e.g. passive health monitoring)
PRINT SCREENS
PHYSICAL THERAPY
COMPLEX GESTURES
PREPOSE: PRIVACY, SECURITY, AND RELIABILITY FOR GESTURE-BASED PROGRAMMING

Lucas Silva Figueiredo
Benjamin Livshits
David Molnar
Margus Veane