CSE 591: Human-aware Robotics

Instructor: Dr. Yu ("Tony") Zhang

Location & Times: CAVC 359, Tue/Thu, 9:00--10:15 AM
Office Hours: BYENG 558, Tue/Thu, 10:30--11:30AM

September 15, 2016

Slides adapted from Aude G. Billard and Abhishek Pal

This set of slides borrows from various online sources; it is used for educational purposes only.
Human-robot Communication

- NLP
- Gesture recognition
- Learning from demonstration
Outline

➤ Gesture recognition

  • Gesture type
  • Technologies behind

➤ Learning from demonstration

  • Problems in LFD
  • Symbolic level
  • Trajectory level
  • Refinement
  • Other issues
What are Gestures
Outline

- Gesture recognition
  - Gesture type
  - Technologies behind

- Learning from demonstration
  - Problems in LFD
  - Symbolic level
  - Trajectory level
  - Refinement
  - Other issues
Gesticulation:
Spontaneous movements of the hands and arms that accompany speech.
Types Of Gestures

- Language-like gestures:
  Gesticulation that is integrated into a spoken utterance, replacing a particular spoken word or phrase.
Types Of Gestures

- Pantomimes

Pantomime

- is an art of dramatic representation by means of facial expressions and body movements rather than words. Pantomime, or mime, has always played a part in theater.

- Pantomime, or dumb show, was essential to commedia dell'arte, an improvised comedy that arose in 16th-century Italy and spread throughout Europe.
Sign languages:
Linguistic systems, such as American Sign Language, which are well defined.
Facial Gesture Recognition

https://www.youtube.com/watch?v=n8wJ8tjmnmU
Eye Gesture Recognition

1. The eye tracker incorporates near-infrared microprojectors, optical sensors and image processing.

2. Microprojectors create reflection patterns on the eyes.

3. Image sensors register the image of the user, the user’s eyes, and the projection patterns, in real time.

4. Image processing is used to find features of the user, the eyes and projection patterns.

5. Mathematical models are used to exactly calculate the eyes’ position and the gaze point.
Outline

- Gesture recognition
  - Gesture type
  - Technologies behind

- Learning from demonstration
  - Problems in LFD
  - Symbolic level
  - Trajectory level
  - Refinement
  - Other issues
Gesture Sensing Technologies:

**Contact type:**
- Touch based gestures

**Non-Contact:**
- Device Gesture Technologies
- Vision-based Technologies
- Electrical Field Sensing
Device Gesture Technologies

Device Sensing Technology

Device-based techniques use a glove, stylus, or other position tracker, whose movements send signals that the system uses to identify the gesture.

The glove is equipped with a variety of sensors to provide information about hand position, orientation, and flex of fingers.
Vision Based Technologies

(a) Original  (b) Single camera  (c) Stereo camera

(d) Depth information  (e) Thermal camera
Electrical Field Sensing
Algorithms

Spatial Gesture Models

3D Model-Based
- Skeletal
- Volumetric
  - NURBS
  - Primitives
  - Super-quadrics

Appearance-based
- Deformable 2D templates
- Image sequences

Ira A. Fulton
Schools of Engineering
Arizona State University
1. Latency
Image processing can be significantly slow creating unacceptable latency for video games and other similar applications.

2. Lack of Gesture Language
Different users make gestures differently, causing difficulty in identifying motions.

3. Robustness
Many gesture recognition systems do not read motions accurately or optimally due to factors like insufficient background light, high background noise etc.

4. Performance
Image processing involved in gesture recognition is quite resource intensive and the applications may found difficult to run on resource constrained devices.
Outline

- Gesture recognition
  - Gesture type
  - Technologies behind

- Learning from demonstration
  - Problems in LFD
  - Symbolic level
  - Trajectory level
  - Refinement
  - Other issues
INCREMENTAL LEARNING OF GESTURES
BY IMITATION IN A HUMANOID ROBOT

SYLVAIN CALINON AND AUDE BILLARD

LASA LABORATORY - EPFL
CH-1015 LAUSANNE, SWITZERLAND
HTTP://LASA.EPFL.CH

ACM/IEEE International Conference on Human-Robot Interaction.
March 9-11, 2007. Washington DC, USA.
What is Imitation
What is Imitation

• “True” imitation: Ability to learn new actions not part of the usual repertoire
• The appanage of humans only, and possibly great apes

• Whiten & Ham, *Advances in the Study of Behaviour*, 1992
• Savage & Rumbaugh, *Child Devel*, 1993
Characteristics of Imitation

- Innate Facial Imitation (newborns → 3 months)
  Tongue and lips protrusion, mouth-opening, head movements, cheek and brow motion, eye blinking
- Delayed imitation up to 24 hours
  → Imitation is mediated by a stored representation

Meltzoff & Moore, *Developmental Psychology*, 1989
• Infants aged 14 months.
• Children imitate new action to achieve the same goal only if they consider it to be the most rational alternative.

Characteristics of Imitation

- 18-months infants
- Learn from unsuccessful examples

Meltzoff, *Dev. Psychol. 31, 1995.*
Advantages: When is Imitation useful?

- It is a powerful means of transferring skills
- It speeds up the learning process by showing possible solutions or conversely by showing bad solutions
- Natural interface
Disadvantages: When is Imitation not useful?

- **Not appropriate**: When a good solution for the teacher is not a possible solution for the learner

- **Disadvantageous**: When it induces you in error - bad teacher (e.g. phobia of spiders)
Outline

- Gesture recognition
  - Gesture type
  - Technologies behind

- Learning from demonstration
  - Problems in LFD
  - Symbolic level
  - Trajectory level
  - Refinement
  - Other issues
The Transfer Problem

Demonstrator

$$\begin{align*}
\theta_1, \theta_2, \theta_3 \\
\theta_4
\end{align*}$$

$$x = (x_1, x_2, x_3)$$

$$\begin{align*}
\theta_5, \theta_6, \theta_7
\end{align*}$$

Imitator

$$\begin{align*}
\theta'_1, \theta'_2, \theta'_3 \\
\theta'_4
\end{align*}$$

$$\begin{align*}
\theta'_5, \theta'_6, \theta'_7
\end{align*}$$

$$\begin{align*}
\vec{x}' = (x'_1, x'_2, x'_3)
\end{align*}$$

Slide from A.G. Billard
What to imitate?

\[ \mathbf{x} = \mathbf{x}' \quad \text{Same Object, same target location} \]

\[ \mathbf{d} = \mathbf{d}' \quad \text{Same direction of motion} \]

\[ \mathbf{v} = \mathbf{v}' \quad \text{Same speed, same force} \]

\[ \mathbf{\theta} = \mathbf{\theta}' \quad \text{Same posture} \]
How to Imitate?
The correspondence problem

Demonstration → Imitation

No solutions (smaller range of motion)

→ Find the closest solution according to a metric
Imitate from Demonstration

Observation of multiple demonstrations

Reproduction of a generalized motion in a different situation

Extraction of a subset of keypoints

Demonstration → Reproduction

Robot programming by demonstration
Aude Billard, Sylvain Calinon, Rüdiger Dillmann, Stefan Schaal
Refinement & Generalization

User feedback

- Demonstration
- Model of the skill
- Ambiguities in the reproduction context?
- Reproduction

Metric for imitation performance

- Demonstrated effect model
- Corresponding effect imitator
- Relative Displacement
- Absolute position
- Extraction of the task constraints
- Application to a new context

Robot programming by demonstration
Aude Billard, Sylvain Calinon, Rüdiger Dillmann, Stefan Schaal
Outline

- Gesture recognition
  - Gesture type
  - Technologies behind

- Learning from demonstration
  - Problems in LFD
    - Symbolic level
    - Trajectory level
    - Refinement
    - Other issues
One-Shot Learning Methods

- Segmentation of demonstration into primitives
- Classification of gestures into predefined states (e.g. grasp, collision)
- Built-in controller for producing sequences of states

- Ritter et al, Rev Neuroscience, 2003
One-Shot Learning Methods

Explicit teaching/learning:
- Reasoning about tasks
- Verbal instructions

State Recognition:
For each sensor a context-dependent model based on background knowledge is provided: ‘opening the refrigerator door’, ‘extracting the bottle’ and ‘closing the door’

Task Reproduction:
Store action sequences in a tree-like structure of macro-operators


Slide from A.G. Billard
Outline

- Gesture recognition
  - Gesture type
  - Technologies behind

- Learning from demonstration
  - Problems in LFD
  - Symbolic level
  - Trajectory level
  - Refinement
  - Other issues
HMM

Robot programming by demonstration
Aude Billard, Sylvain Calinon, Rüdiger Dillmann, Stefan Schaal
Robot programming by demonstration
Aude Billard, Sylvain Calinon, Rüdiger Dillmann, Stefan Schaal
Dynamical Systems

Robot programming by demonstration
Aude Billard, Sylvain Calinon, Rüdiger Dillmann, Stefan Schaal
### Comparison

<table>
<thead>
<tr>
<th>Span of the generalization process</th>
<th>Advantages</th>
<th>Drawbacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbolic level</td>
<td>Allows to learn hierarchy, rules and loops</td>
<td>Requires to pre-define a set of basic controllers for reproduction</td>
</tr>
<tr>
<td>Trajectory level</td>
<td>Generic representation of motion which allows encoding of very different types of signals/gestures</td>
<td>Does not allow to reproduce complicated high-level skills</td>
</tr>
</tbody>
</table>

#### Generalization at a symbolic level:
- Model of the skill
- Pre-determination of the set of controllers required for the skill
- Extraction of pre-defined actions
- Generalization
- Prior knowledge
- Additional information (e.g., social cues)
- Application to a new context

#### Generalization at a trajectory level:
- Model of the skill
- Projection in a latent space of motion
- Generalization
- Prior knowledge
- Additional information (e.g., social cues)
- Application to a new context
Outline

- Gesture recognition
  - Gesture type
  - Technologies behind

- Learning from demonstration
  - Problems in LFD
  - Symbolic level
  - Trajectory level
    - Refinement
  - Other issues
Incremental Refinement

Model of the movement relative to the cylinder

Model of the movement relative to the cube

Reproduction attempts in a new situation

Red cylinder

Yellow cube
Refinement via Iterative Refining

- Require new demonstrations

- If the novel situation **differs Importantly** from the demonstrated ones, then adapting the demonstrated trajectory is no longer sufficient to satisfy the task.

→ Need to relearn the task -- Reinforcement Learning

→ Need to define a new metric – the reward

Solution
Refinement via Reinforcement learning

Robot programming by demonstration
Aude Billard, Sylvain Calinon, Rüdiger Dillmann, Stefan Schaal
Refinement

Robot programming by demonstration
Aude Billard, Sylvain Calinon, Rüdiger Dillmann, Stefan Schaal
Outline

➢ Gesture recognition
  • Gesture type
  • Technologies behind

➢ Learning from demonstration
  • Problems in LFD
  • Symbolic level
  • Trajectory level
  • Refinement
  • Other issues
Human-robot interaction

Robot programming by demonstration
Aude Billard, Sylvain Calinon, Rüdiger Dillmann, Stefan Schaal
Outline

➤ Gesture recognition
  • Gesture type
  • Technologies behind

➤ Learning from demonstration
  • Problems in LFD
  • Symbolic level
  • Trajectory level
  • Refinement
  • Other issues