Robot skill synthesis through human sensorimotor learning

Jan Babič
Jožef Stefan Institute, Slovenia
Grasping vs. full body motion
Outline

Human sensorimotor learning

• Ball swapping
• Grasping
• Reactive postural control
• Concluding remarks

Robot skill synthesis
Robots in everyday life...

- Robots in daily life require **new methods** for synthesis of skillful behaviour
- Classical approach requires experts, and lot of expert work hours. How could non-experts teach robots is an active research topic in robotics:
  - Teaching by demonstration
  - Robotic imitation
  - ...

- To make the task as **natural and easy** for the human teacher
- The human provides an initial demonstration but is **NOT** part of the motor control loop
The paradigm

- Use human sensorimotor learning ability to obtain robot behaviors
- Include the human in the control loop
- May ask human to do extensive training
- Utilize the human brain as the adaptive controller
Sensorimotor learning

- Sensorimotor learning is fundamental for adaptive and intelligent behavior
- Driving a car
- Using a pair of chopsticks
- Using a computer mouse
- ...

![Diagram of sensorimotor learning process]

- Feedback to human sensory system (f)
- Feedback Interface
- Robot state (s)
- Human Motion (m)
- Feedforward Interface
- Motor command (u)
For autonomous operation, the key issue is transferring the control policy learnt by human to the robot.
Why should this paradigm work?

• The ability of the brain to learn novel control tasks by forming *internal models*. The robot can be considered as a tool (e.g. as driving a car, playing an instrument, using chopsticks)

• The flexibility of the *body schema*; extensive human training modifies the body schema so that the robot can be naturally controlled
Ball swapping is a suitable task for testing the proposal since it is complex and not straightforward to manually program on a robotic hand.

work of Erhan Oztop
Ball swapping interface

Feedback to human: DIRECTVISION

Human ~ Adaptive Controller

Feedback to human sensory system (f)

Feedback Interface

Robot Learning: Learn $\pi: s \rightarrow u$

Robot state (s)

Human Motion (m)

Feedforward Interface

Motor command (u)

Human ~ Adaptive Controller

Feedback to human sensory system (f)

Feedback Interface

Robot Learning: Learn $\pi: s \rightarrow u$

Robot state (s)

Human Motion (m)

Feedforward Interface

Motor command (u)
Human control of the robot

- Human hand movement
- Data capture
- VizualEyez
- Finger tip positions
- Gifu Hand reference frame
- Calibration
- Inverse kinematics
- Filtered data
- PD control
- System information
- Human control of the robot
- Gifu Hand actuation
- Desired joint angles
- Raw joint angles
- Inverse kinematics
- Central controller
- Input driven
- 30 Hz
- 10 Hz
- Gifu Hand controller
- Hand status
- Finger tip positions
- Gifu Hand joint angles
Human learning...
Finally human learns to swap balls
Offline analysis & improving performance

A. Original finger joint trajectories

B. Smoothing & Linear interpolation

C. Kicks superimposed on to (B)

D. Speed-up, then apply (B) and (C)
Ball swapping at x2 speed up

Open loop control

\[ u = \pi \text{(time)} \]
Ball swapping with visual feedback

In collaboration with Jan Steffen from Bielefeld University

Color tracking

Color tracker developed in house by Ales Ude

Learning Technique: Unsupervised Kernel Regression (UKR)

policy \( u = f(s, v) \)

u: desired joint angles
s: finger joint angles
v: position of the balls

\( u = f(s, v) \)
Ball swapping: feedback vs. open loop

Closed loop (feedback) control
\[ u = \pi(\text{angles, ball positions}) \]

Open loop control
\[ u = \pi(\text{time}) \]
Extending the paradigm to visual grasping

in collaboration with Brian Moore
Visual grasping

in collaboration with Emre Ugur
Analysis and preliminary findings

Limited success in simulation to robot transfer

The skill obtained in the simulator was satisfactory. Transfer to the real robot had limited success.

General observation: For efficient grasp skill generation there should be low level tactile controller at the fingers (work in planning 😊).
Reactive postural control

- teach the humanoid robot to counteract external postural perturbations
- choices of feedback interface:
  - abstract visual feedback
  - motion of the support polygon
  - force impulses at human’s COM
The key factor for muscle activation during postural control is COM information [Lockhart et al, 2007]
An interface between robot COM and human COM
"COM force" interface
Reactive postural control

- Goals:
  - teach the robot to counteract external postural perturbations
  - on-line learning
  - gradually transfer control responsibility from human to autonomous robot controller
Principles

Feedback Interface \(\rightarrow\) Autonomous Controller \(\rightarrow\) Feedforward interface

sensory stimulation \(\rightarrow\) human joint positions

training data \(\rightarrow\) training data

sensory information \(\rightarrow\) robot joint positions

Influence Weighting
To teach the robot the demonstrated task we used Locally Weighted Projection Regression [Vijayakumar et al. 2005]

LWPR offers incremental learning and is among the fastest regression techniques [Nguyen et al. 2011]

As opposed to global regression (GPR) which uses entire training set, LWPR partitions the training set into more sections

Each section is described by a local model:

\[ \bar{y}_k = \bar{x}_k^T \theta_k \]

Influence of each model is determined by the weights characterized by Gaussian kernel:

\[ w_k = \exp\left( -\frac{1}{2} (x - c_k)^T D_k (x - c_k) \right) \]

The output prediction for an input \( x \) is a sum of contributions from all models weighted by \( w_k \):

\[ \hat{y} = \frac{\sum_{k=1}^{N} w_k \bar{y}_k(x)}{\sum_{k=1}^{N} w_k} \]
Responsibility transfer

- The influence weighting algorithm calculates the mean square error (MSE) between the human reaction and predicted reaction over a period $T$ during the demonstration.
- The maximum MSE is set as a reference for the weighting criterion:
  \[ C = \frac{MSE_{\text{total}}}{MSE_{\text{max}}} \]
- The criterion is used to weight the human influence and the influence of the autonomous controller.
- The output that is controlling the robot is calculated by:
  \[ y = Cy_{\text{human}} + (1-C)y_{\text{predicted}} \]
- If the MSE fails to improve over $N$ periods the algorithm disconnects the human from the control loop.
- At that point the robot is considered trained.
Responsibility transfer
Stability algorithm + manipulation task

- Combine learned skill with an additional arbitrary task
- Stability algorithm can influence the manipulation task but only if necessary.
- Manipulation task must not influence the postural stability of the robot.
- Null space exploration

in collaboration with Leon Žajpah
Concluding remarks...

Our work so far indicates that

- Obtaining robot skill via human sensorimotor learning is a viable approach

- Since the paradigm reverse engineers the control policy obtained by the brain, the behaviors obtained should be natural and human-like

- Help build smart prosthetics that can be controlled intuitively via high-level signals or brain machine interface (BMI)

- Shed light on mechanisms of internal models, agency and body image
  - Help ameliorate impairments related to these brain mechanisms
  - Offer new design principles for robot self-exploration and learning
THANK YOU FOR YOUR ATTENTION!

Collaborators:

Erhan Ozttop, Ozyegin University, Turkey
Luka Peternel, Jožef Stefan Institute, Slovenia
Joshua Hale, Cyberdyne, Japan
Mitsuo Kawato, ATR, Japan