Direct Control Interfaces for Robotics

Michael Gleicher
Department of Computer Sciences
University of Wisconsin Madison
Dumb Robots for Smart People
Direct Control Interfaces for Robotics

Michael Gleicher
Department of Computer Sciences
University of Wisconsin Madison
Michael Gleicher
Visual Computing Group

Human **Graphics** Interaction
authoring pictures, videos, animations

Human **Robot** Interaction
robots!

Human **Data** Interaction
visualization, visual analytics, interactive learning
Acknowledgements

Students
Danny Rakita
Pragathi Praveena
Guru Subramani
Mike Hagenow
Bolun Zhang

Co-PIs
Bilge Mutlu
Michael Zinn

External collaborator (NRL)
Laura Hiatt
Dumb Robots for Smart People
Direct Control Interfaces for Robotics

Michael Gleicher
Department of Computer Sciences
University of Wisconsin Madison
Dumb Robots for Smart People

Can we enable (smart) people to work with dumb robots?

Robot intelligence may be **overrated** less important in places where we can effectively use human intelligence and skill.

If we can successfully exploit the person to help the robot.
Folding a Shirt

The robot doesn’t do a great job... but this was 2016

And it’s hard!
You need to see the shirt and understand its geometry
You need to understand the tasks and goals
You need to know how cloth reacts to being pushed and pulled
You need to predict how the shirt will move
You need to choose where to grab and which way to pull things
You need to plan/strategize on how to get the shirt folded
What happened?

To fold a shirt...
You need to see the shirt and understand its geometry
You need to understand the tasks and goals
You need to know how cloth reacts to being pushed and pulled
You need to predict how the shirt will move
You need to choose where to grab and which way to pull things
You need to plan/strategize on how to get the shirt folded
What happened?

To fold a shirt...
You need to see the shirt and understand its geometry
You need to understand the tasks and goals
You need to know the tasks and goals
You need to know the objects
You need to know the positions
This robot didn’t do any of that
The human did the “hard parts”
Should you care?

1. Smart robots will do this eventually. To get there…
   - we need to understand control better
   - we need to learn from people

2. Some tasks will remain automation resistant
   - require too much human judgment
   - too specialized
   - people want to be in control (it feels good to be useful!)
How dumb are the robots?

All the robots in this talk are:

**Deaf** – they don’t sense anything (rough touch in 1 demo)

**Dumb** – (both senses of the word)
  they don’t speak (provide feedback other than action)
  they just follow directions

**Blind** – no vision (except crude vision in 1 case)

Future work will fix this – but we need to start somewhere
Outline

Direct Control (Mimicry Interfaces)
  - Why it works (perceptually / UX)
  - How it works (technical issues)

Improving Direct Control
  - Extending to bi-manual
  - Helping with visibility

Learning from Direct Control
  - How to get information from people (input devices)
  - How to use this information (constraint inference)
Mimicry: Direct Control Tele-Operation

Making the robot do what you do.


People can do cool things ...

Not just (experienced) graduate students

Experimental participants ("off-the-street") with no experience
Become competent right away
Why?

Conjecture:
If the interface is “natural enough” then people can apply their existing skills, abilities, and knowledge.

but it has to be “natural enough”
otherwise, the interface gets in the way
What is natural enough?

Closed loop – immediate feedback

Kinesthetically direct mapping

It needs to “feel” like its doing what you are doing
Why do I believe this?

Experiments
- other tele-operation interfaces (touchpad, stylus, ...)
- other robot modalities (kinesthetic teaching)

- breaking naturalness
  - latency
  - slowness
- bad control mappings
Break things differently

Latency
Slowness

Lead to different strategies
Intuition: It’s a perceptual hack!

It feels as if the robot is following your hand.

If it feels like using your hand, you’ll just use your hand and you know how to do that.

If it feels like the robot is just doing what your hand does, you will adapt what your hand does.

What makes it feel right?
The perceptual hack

Exact position matching doesn’t matter
motions are relative
and details only matter when you slow down

Responsiveness matters
Continuity matters (if I move continuously, ...)

So how do we do this?
Mimicry-based Teleoperation
Approach Overview

Hand 6-DOF Space → Spatial Mapping → End Effector 6-DOF Space → Inverse Kinematics → Robot Joint Angle Space
Not Traditional Inverse Kinematics

Position and orientation goals are not the only goals we don’t need precise matches!

Other things are important too:
smoothness / continuity
responsiveness
avoid self-collisions and singularities

Make tradeoffs (multi-objective optimization)
\[ f(\theta) = w_p f_p(\theta) + uw_o f_o(\theta) + w_j f_j(\theta) + w_e f_e(\theta) \]

Position Tracking

\[ w_p \text{ low} \]
\[ w_p \text{ medium} \]
\[ w_p \text{ high} \]
\[ f(\theta) = w_p f_p(\theta) + uw_o f_o(\theta) + w_j f_j(\theta) + w_e f_e(\theta) \]

Orientation Tracking

\( w_o \) low  \quad w_o \text{ medium}  \quad w_o \text{ high}
\[ f(\theta) = w_p f_p(\theta) + u w_o f_o(\theta) + [w_j f_j(\theta)] + w_e f_e(\theta) \]

Minimized Joint Velocities

\( w_j \) low

\( w_j \) medium

\( w_j \) high
\[ f(\theta) = w_pf_p(\theta) + uw_ofo(\theta) + w_jf_j(\theta) + wef_e(\theta) \]

Minimized End Effector Cartesian Velocity

\( w_e \) low \hspace{1cm} \( w_e \) medium \hspace{1cm} \( w_e \) high
Singularity Avoidance

\[ c(\theta) := \sqrt{|\det(J(\theta)J(\theta)^T)|} - s_{\text{min}} \geq 0 \]
Self-Collision Avoidance

TRAC-IK

Relaxed-Mimicry (Ours)

\[ l_i \leq \theta_i \leq u_i \]
Keep away from bad things

Keep manipulability above minimum (away from singularities)
Keep distance to self-collision above minimum

Efficient formulations as constraints

Use a neural net to approximate the self-collision distance function.
Importance-Based Inverse Kinematics
Importance-Based Inverse Kinematics

\[ f(\theta) = w_p f_p(\theta) + uw_0 f_0(\theta) + w_j f_j(\theta) + w_e f_e(\theta) \]

Hand Velocity High
\[ f(\theta) = w_pf_p(\theta) + uw_o f_o(\theta) + w_j f_j(\theta) + w_e f_e(\theta) \]

*Importance-Based Inverse Kinematics*  
Hand Velocity *Low*
How do we balance competing objectives?

Use constraints for high-priority

Hard to balance weights for different objectives
  different scales
  different falloffs

Use shaping functions!
Relaxed IK

IK (position/orientation matching) is just one goal

Self-collision and singularity avoidance are priorities

Everything else is a tradeoff
  be flexible (allow for different objectives)
  be dynamic (tune weights responsively)

And do it fast...
Relaxed IK

Very highly engineered, open source library uses automatic differentiation, standard non-linear solvers

v.1 Python
v.2 Julia
v.3 Rust (coming soon)

Very flexible
provide robot kinematics and geometry...
Robustness

Fails less than other solvers
Fails more gracefully
Preserves continuity of path
Square Tracing

RelaxedIK

Trac-IK

Real-time Motion Planning
Part 1 Takeaways
Mimicry seems magical – uses the human’s skills
Relaxed IK gives a robust, flexible implementation
just use it (it’s open source)

Can we use the flexibility of the optimization framework to do
even better tele-operation?

Outline

Direct Control (Mimicry Interfaces)
  - Why it works (perceptually / UX)
  - How it works (technical issues)

Improving Direct Control
  - Extending to bi-manual
  - Helping with visibility

Learning from Direct Control
  - How to get information from people (input devices)
  - How to use this information (constraint inference)
Assisted / Extended Mimicry Tele-Operation

Have the robot help you out as you control things


How about two hands?

We need to coordinate the hands
It takes more than 2 one-handed systems

We must help with coordination
How to do this...

Determine what kinds of coordination are necessary
  watch videos – get ideas
  track people – get patterns
  common “assistance modes”
Determine how to assist
  rigid coupling – one hand drives
  self-handover – relative motions
Figure out how to detect when to engage
  learn from examples

Laura will say more tomorrow
A different problem

How to see what you’re doing?

Hard enough when you’re there

Real applications are remote
Array of Static Cameras

Cameras

Camera Views
Participant Video

- No occlusions by manipulation robot
- View blocked by grasped objects
Our Solution

**Dynamic** – the camera moves to best show action
   there is no one right view – it changes

**Autonomous** – not under user control
   too busy doing manipulation

Use a second robot to hold the camera
Approach 1 (Fall 2017):
Camera as a second optimization

Point the camera at the hand
use leading (prediction) and framing
avoid occlusions (of the other arm)
keep motion smooth
move closer as hand slows

This is naturally Relaxed-IK
The camera robot doesn’t use video – we have robot geometry
Motion Retargeting Optimization

User Motion Input

Live Video Stream

Camera Robot Configuration (per update)

Manipulation Robot Configuration (per update)

Camera Robot Motion Optimization
Camera Distance

✓ Camera should move in for detail when user is exhibiting slow control
Camera Distance

✔ Camera should move out for context when user moves robot quickly
Control Frame

Absolute Frame

Relative Frame
Organize Pills
Even better camera control

Let the camera robot see (look for AR tag on gripper)

avoid occlusions
should always be able to see hand and “goal”

Optimize both robots together

keep manipulation visible

User control

searching and exploring, nudging, control over distance
Takeaways

Improve tele-operation by letting the robot help control the second hand (even if it has a camera)
Make use of the flexibility of optimization
Keep the naturalness of the primary hand


Outline

Direct Control (Mimicry Interfaces)
- Why it works (perceptually / UX)
- How it works (technical issues)

Improving Direct Control
- Extending to bi-manual
- Helping with visibility

Learning from Direct Control
- How to get information from people (input devices)
- How to use this information (constraint inference)
Getting and Using Natural Demonstrations

If naturalness is so magical...
How can we get good, natural demonstrations?
What information can we get from them?


Guru Subramani, Michael Hagenow, Bolun Zhang, Michael Zinn and Michael Gleicher. Robust Replay of Human Demonstrations using Identified Constraints. *Submitted for publication.*
How to capture human demonstrations?

Watch people do it (naturally)
  use their hands

Have people use “robot demonstration methods”
  teach pendants
  kinesthetic teaching
Kinesthetic Teaching
Why not the hands?

Hard to instrument (track)
Hard to instrument (forces)

Hand is different than gripper
Hard to map
Implausible strategies

Hand is overkill
Sensing and actuation

Photo by [Meghan Schiereck](https://unsplash.com) on [Unsplash](https://unsplash.com)
Photo by [Mat Reding](https://unsplash.com) on [Unsplash](https://unsplash.com)
Photo from Robotiq Website
Desirable properties in a demonstration method

Efficient – Prevent wasteful use of resource
Subjective performance – Demonstrator perception
Facile – Demonstrator ease of use
Amenable to analysis – Post-processing
Desired demonstrations – Experimenter objectives
Affords quality demonstrations – Data quality
Easy to learn – Process to proficiency
Preference – Demonstrator liking
Feedback – Access to demonstration performance
Plausibility – Equivalent strategies
Feasibility – Correspondence
Instrumentable – Measurement capabilities
Showing tasks to robots (common approaches)

**Hand Demonstrations**
- Easy for demonstrator
- High quality demonstrations
- Hard to instrument
- Hard to map to robot

**Kinesthetic Teaching**
- Hard for demonstrator
- Poor quality demonstrations
- Easy to instrument (with caveat)
- Feasible mappings (with caveats)
Is there a middle ground?

Natural – easy, high quality demonstrations – like the hand

Maps nicely to the robot
Inspiration...
Teaching with Tongs!
Lots of graphs like this...
Lots of anecdotes like this...
Kinesthetic teaching gives a feasible robot trajectory…

*Aren’t you concerned about feasibility?*

Yes – it’s a problem
   But K-Teaching really doesn’t solve it (only for replay)

No – we have good mapping methods
   Relaxed IK (pathwise versions) get trajectories
   This is the least of our worries…
What can you get from a demonstration?

Trajectory
Timing

Open/Close
Pressure
Force/Torque
What can you get from a demonstration?

Trajectory
Timing
Open/Close
Pressure
Force/Torque
What can you get from a demonstration?

- Trajectory
- Timing
- Open/Close
- Pressure
- Force/Torque

Two ATI high-precision, 6dof, F/T sensors
Why forces and torques?

They tell us about why a motion happened (not just what it was)
Physical Constraints:
Teaching a robot what NOT to do
Drawing

Pen tip on plane
Pushes down with force

Infer constraint...
Figure out where tip is
Figure out where plane is
How hard to push
Point on plane constraint estimation

Estimated Plane

Estimated Tip Motion

Trajectory

Estimated Tool Tip

Estimated Plane

Ground Truth Plane

0.0005m
Constraints as a vocabulary for motion

Planar constraint  Prismatic constraint  Axial constraint  Point on plane constraint

Planar relaxed constraint  Prismatic relaxed constraint  Axial relaxed constraint  Point on line constraint
In the real world...

axial relaxed constraint

planar relaxed constraint

prismatic relaxed constraint
How do we do this?

Collect forces, torques, and pose
   (wrench = force+torque)
Segment motion into constrained actions

Factor forces (reaction forces)
Fit to geometric models
   find parameters
Select Model
Why force/torque?

Kinematic fitting required samples (135)

Kinematic + wrench fitting required samples (36)

Fit plane

Graph showing error in m against samples.
How to use constraints?

Robust playback (current)
   Generalize from one-shot demonstrations
   Use hybrid (position/force) control against constraints
   Still works if things move

Representation of Action (future)
   Editing
   Combining demonstrations
   Bigger Generalizations
Takeaways

Consider input devices that balance constraints and naturalness
Instrument demonstrations
Infer constraints to describe actions


Guru Subramani, Michael Hagenow, Bolun Zhang, Michael Zinn and Michael Gleicher. Robust Replay of Human Demonstrations using Identified Constraints. *Submitted for publication.*
What’s next?

Use sensing and intelligence to improve tele-operation collisions in complex and dynamic environments better depth perception
Better signaling and feedback motion cues, tactile cues, haptic cues, ...
Using constraints as input and output better inference and application

Applications beyond real-time tele-op
The bigger picture: Why do we need to communicate?

Specify (Learning from Demonstration)  Specify (Direct Control)  Interpret  Monitor
The overall story...

We can enable (smart) people to work with (dumb) robots by...
- Considering how to make “natural” interfaces
- Using relaxed mappings
- Providing basic assistance where people need help
- Providing help with viewpoint choices
- Designing appropriate input devices
- Describing motions with inferred constraints
Acknowledgements

Students
Danny Rakita
Pragathi Praveena
Guru Subramani
Mike Hagenow
Bolun Zhang

Co-PIs
Bilge Mutlu
Michael Zinn

External collaborator (NRL)
Laura Hiatt
Thanks!
To you for listening.
To my students and collaborators.
To the NSF, NASA and UW2020 for funding.

Dumb Robots for Smart People

Direct Control Interfaces for Robotics

Michael Gleicher
University of Wisconsin Madison
gleicher@cs.wisc.edu