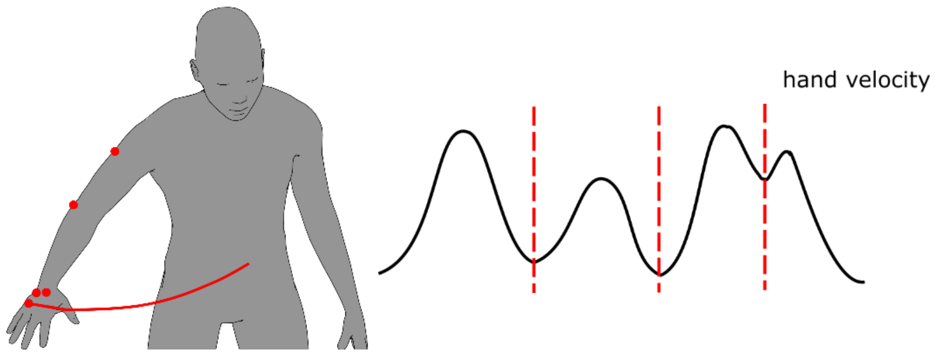
Detection and Recognition of Human Manipulation Building Blocks

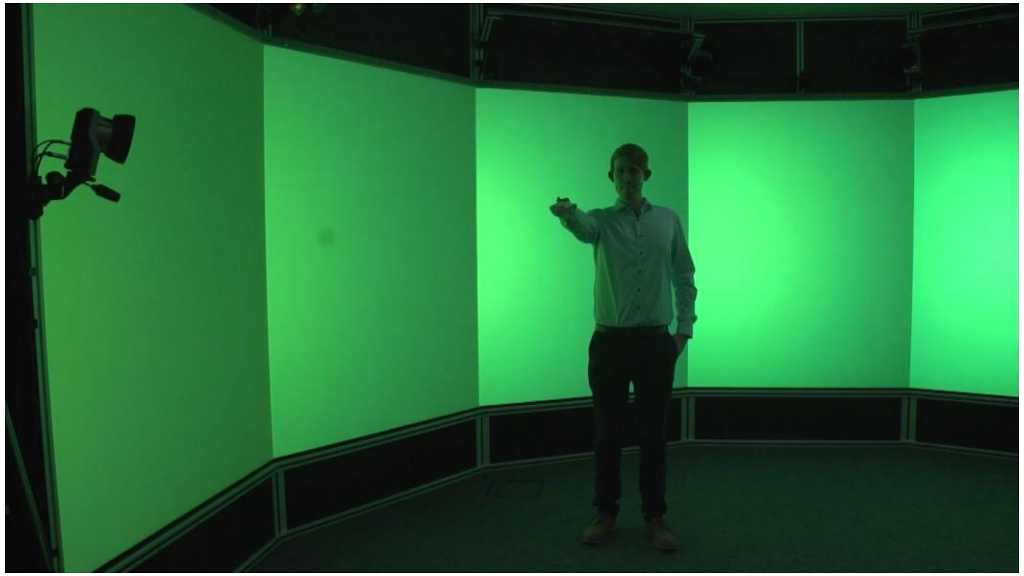
LISA GUTZEIT

24.01.2023



Motivation

Learning from Humans in Robotics



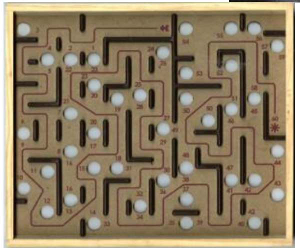


Image: health.uottawa.ca

- robotic learning through interaction with humans
- use principles of human leaning to learn new robotic behavior
  - divide complex learning problems into simple subtasks [Adi-Japha, 2008; Sakai, 2003]
  - combine building blocks to different complex movements [Graybiel, 1998]
  - re-use learned movements in other (related) tasks

How to find these subtasks/building blocks?

Overview

Goal: *“Development of methods to detect building blocks in human manipulation movements that can be used to generate robotic behavior“*

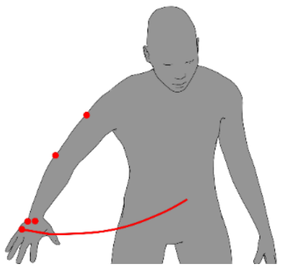
1. detection of building blocks in human manipulation movement
   - unsupervised segmentation of human movement trajectories
2. few-shot recognition of building blocks
   - classification of human movement segments with a small number of training data
3. generate robotic movements based on human example
   - imitation of human movement for a robotic system

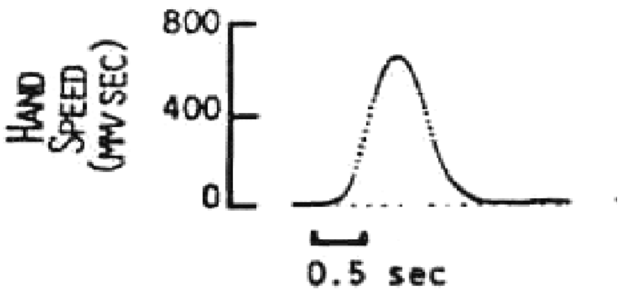
Background

Generation of Goal-directed Movements in Humans

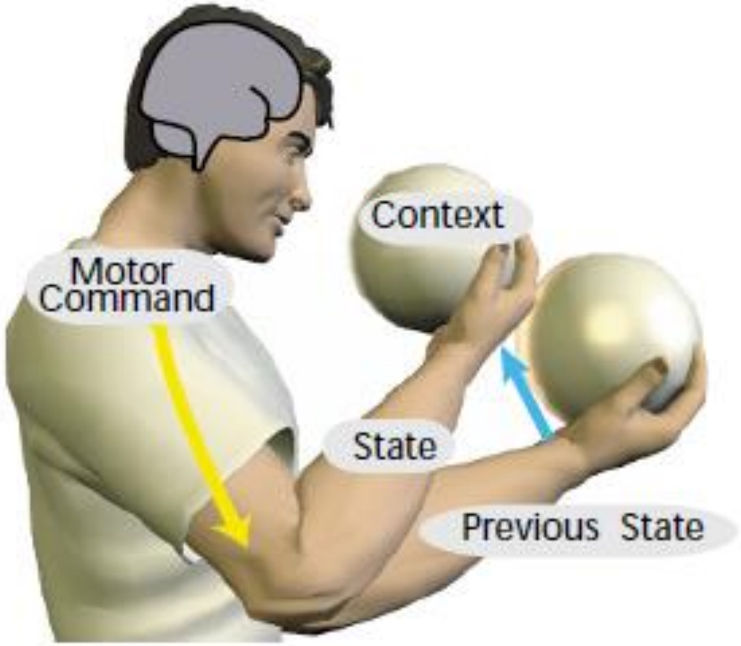
Goal-directed Manipulation Movement:

- infinite number of possible trajectories to move
- CNS always chooses a similar one
- deliberative actions





Mussa-Ivaldi & Solla, 2004



Wolpert & Ghahramani, 2000

Theory: CNS plans the movement so that the end effector moves a long an approximately straight path with a smooth, **bell-shaped velocity profile** [Shadmer & Wise, 2005]

Why segmentation? → effective learning; detection of building blocks

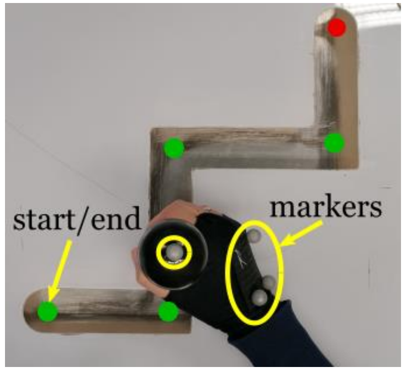
How to find building blocks? → e.g. bell-shaped velocity profile of the hand (manipulation movement)

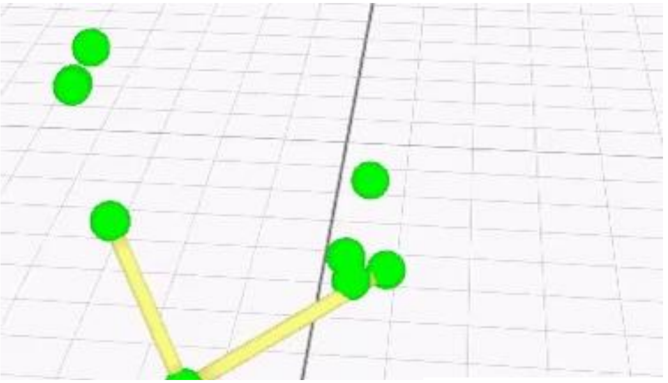
Definitions:

- **Manipulation Movement:** *Hand or arm movement executed to manipulate an object, e.g. "pick" or "place".*
- **Building Block:** *Central movement entity of manipulation which can be combined with other building blocks to solve different tasks, characteristic: bell-shaped hand velocity.*
- **(Movement) Action:** *Concatenations of multiple building blocks, e.g. "pour water into cup".*

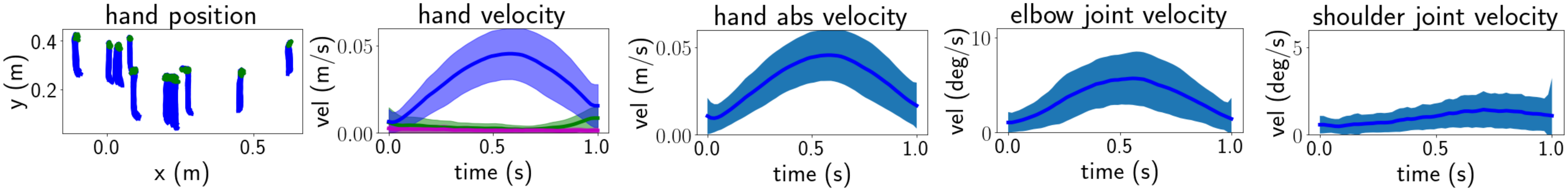
Segmentation into Building Blocks

Regularities in Manipulation Movements – step movements

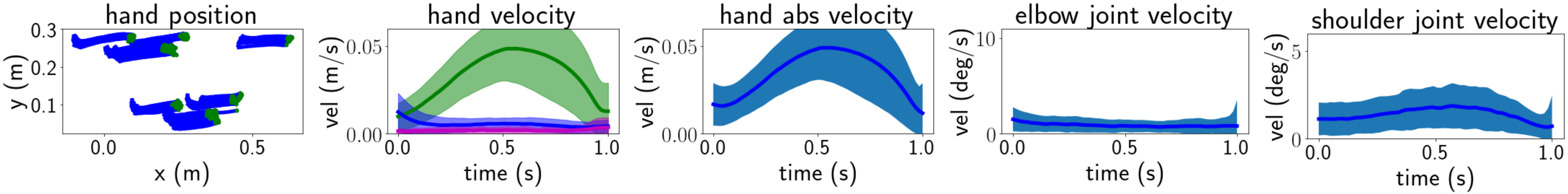




“down” movements:

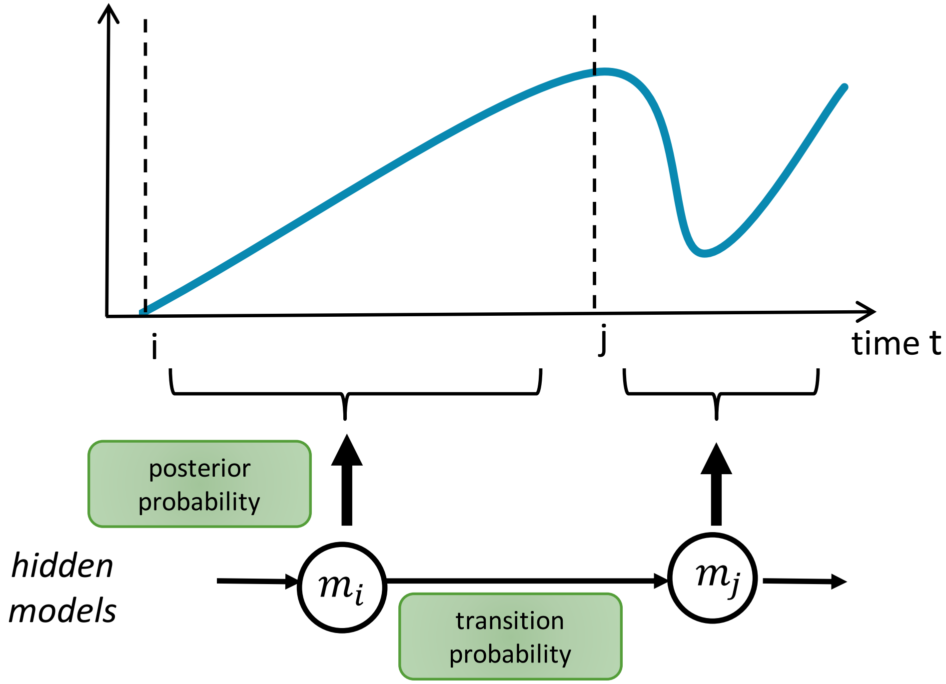


“left” movements:



Segmentation into Building Blocks

- Velocity-based Multiple Change-point Inference (**vMCI**)
  - infers change-points in a time series
  - resulting segments are characterized by a bell-shaped velocity profile
  - based on multiple change-point inference (MCI) algorithm introduced by [Fearnhead & Liu,2007]



data model:

$$y_{i+1:j} = \sum_{k=1}^{q} \beta_k \phi_k + noise,$$

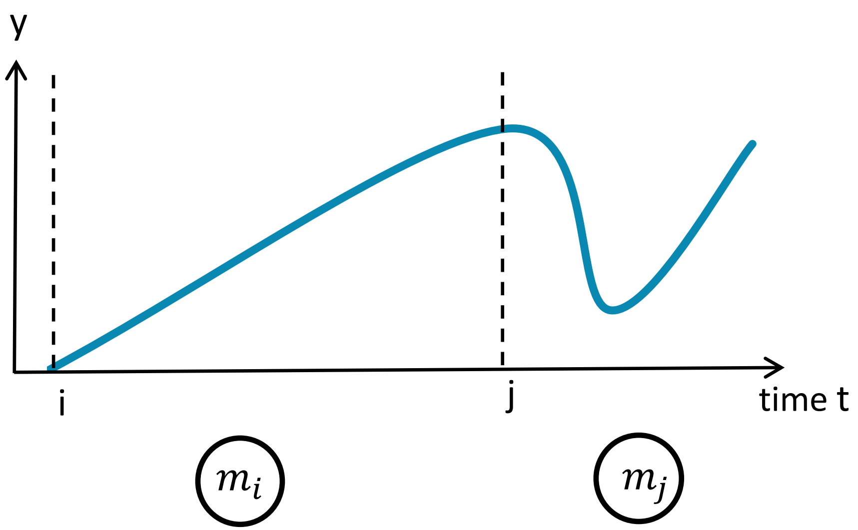
$y_{i+1:j}$	data time i - time j
β_k	weights
ϕ_k	basis function
q	model order

\+ online inference of segments
\+ probabilistic, i.e., small movement variations are modeled
\+ unsupervised, hyper-parameters can be calculated from the data

Segmentation into Building Blocks

Velocity-Based Multiple Change-Point Inference (vMCI) [Senger et al. 2014, Gutzeit & Kirchner 2022]

position model:

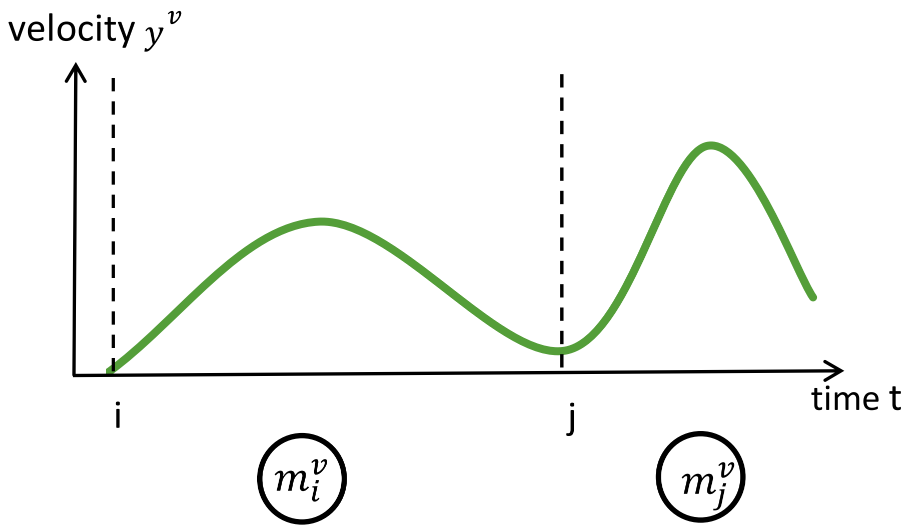


$$y_{i+1:j} = \sum_{k=1}^{q} \beta_k \phi_k + noise,$$

auto-regressive basis function: $\phi_k(x_t) = \beta_k(x_{t-1})$

$y_{i+1:j}$ data time i - time j
β_k weights
q model order

velocity model:



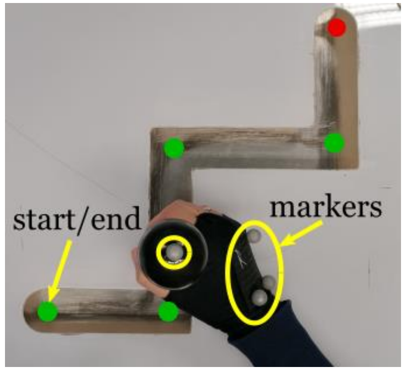
$$y_{i+1:j}^v = \alpha_1 \phi_v + \alpha_2 + noise,$$

radial basis function: $\phi_v(x) = exp\left\{-\frac{(c-x)^2}{r^2}\right\}$

$y_{i+1:j}^v$ velocity data time i - time j
α_1, α_2 weights
c, r center and width of the basis function

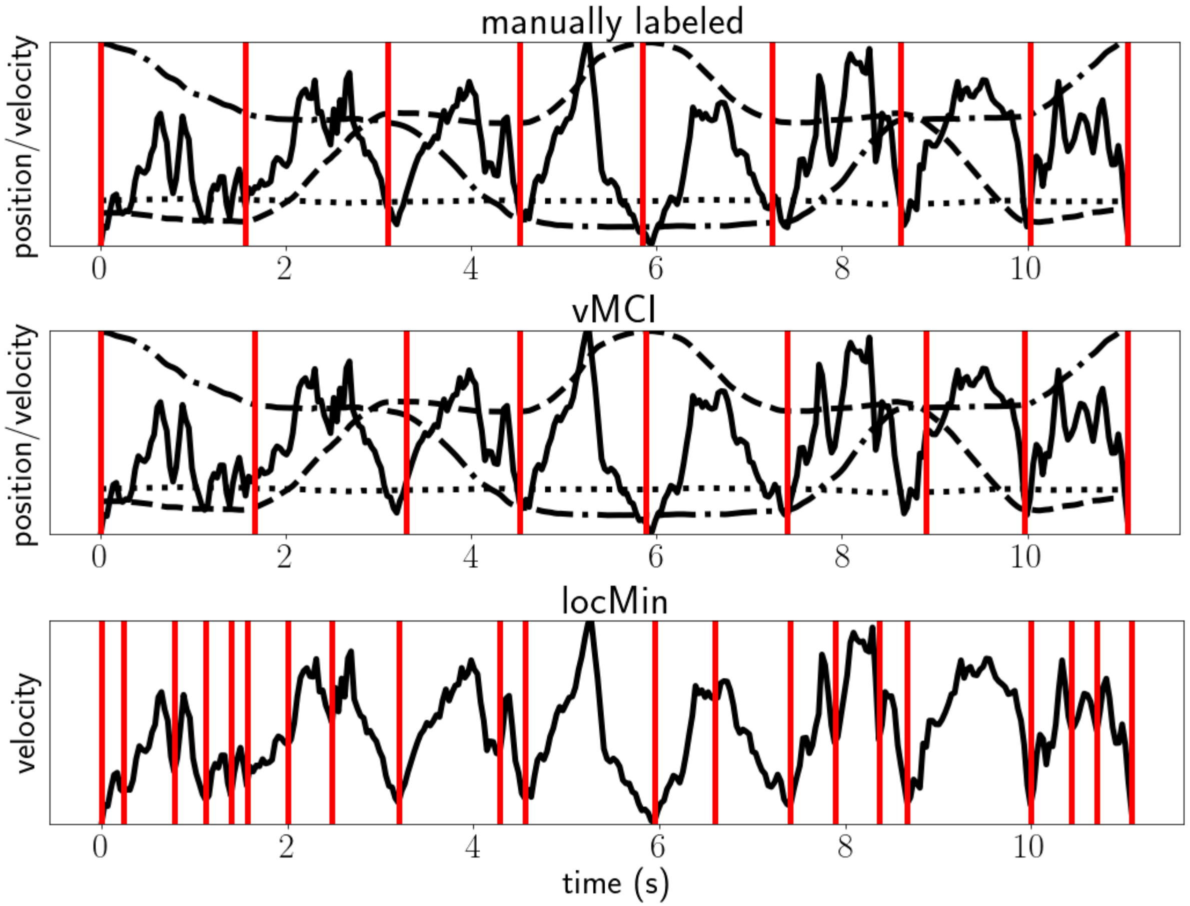
Segmentation into Building Blocks

vMCI - Evaluation on step movements



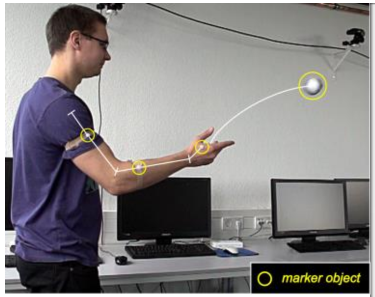
average results on 171 demonstrations

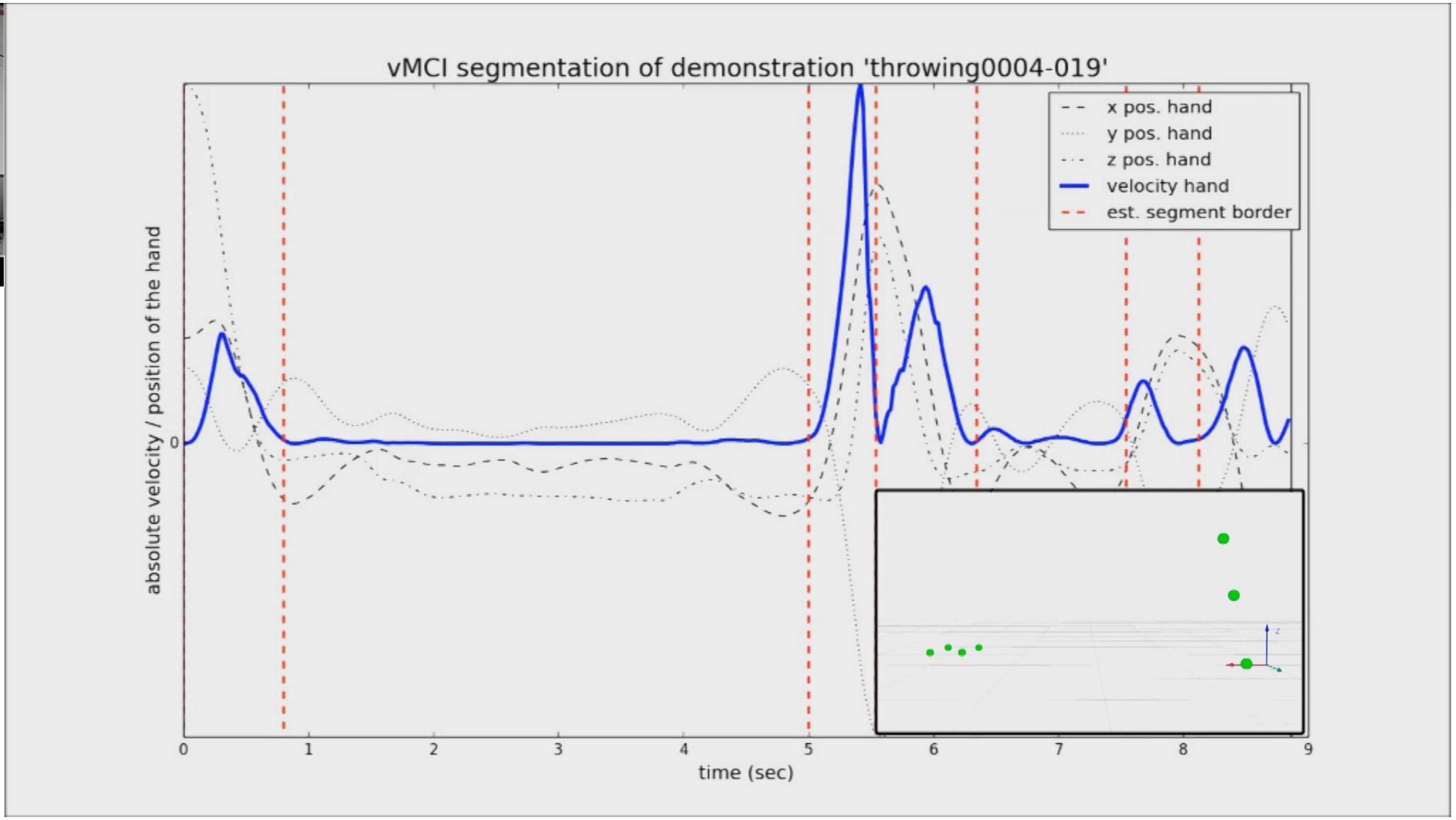
	F1-score	TP (opt.: 7)	FP (opt.: 0)
vMCI	**0.85**	6.0	**1.1**
MCI	0.63	4.0	1.2
locMin*	0.83	**7.0**	3.7
BPARHMM [Fox et al. 2010]	0.81	8.9	3.1
ProbS [Lioutikov et al. 2017]	0.18	0.18	1.9



Segmentation into Building Blocks

vMCI - online application



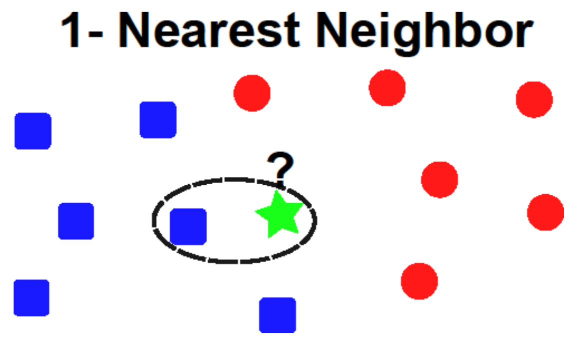


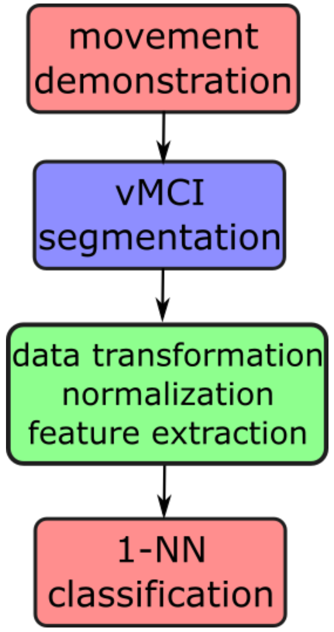
Few-shot Movement Recognition

1-NN Approach

- requirements:
  - minimal manual efforts (small number of parameters, small training data set sizes)
  - high accuracy in detection of distinct movement
  - generalization to different subjects

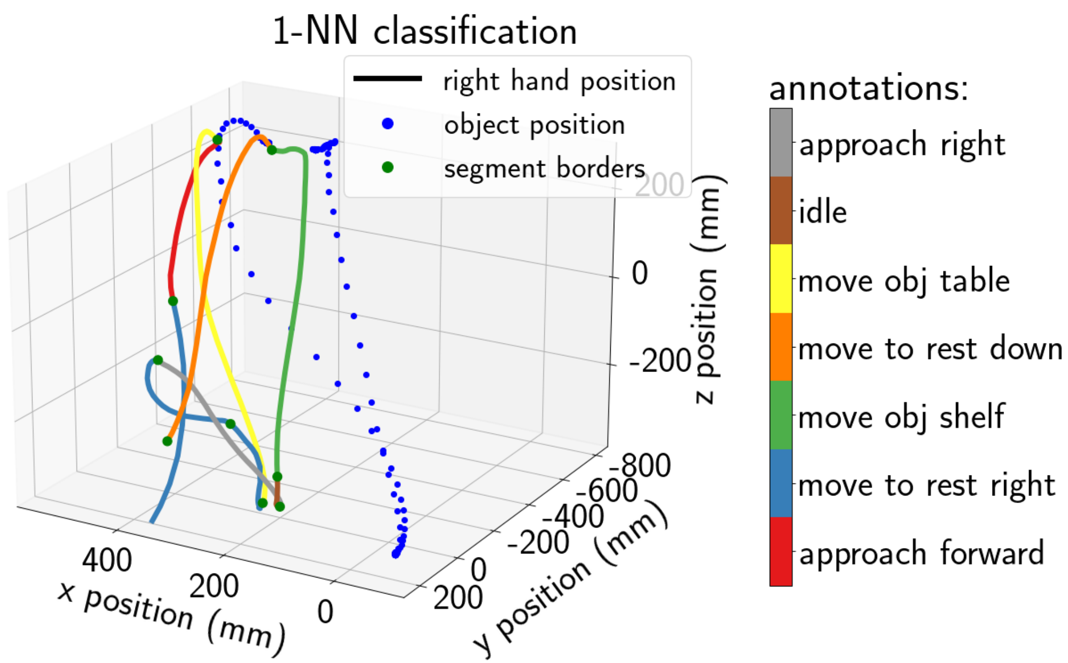
- 1-Nearest Neighbor (1-NN) classification:
  - features: hand and arm markers
  - transformation of data into coordinate system located at the back of the demonstrator

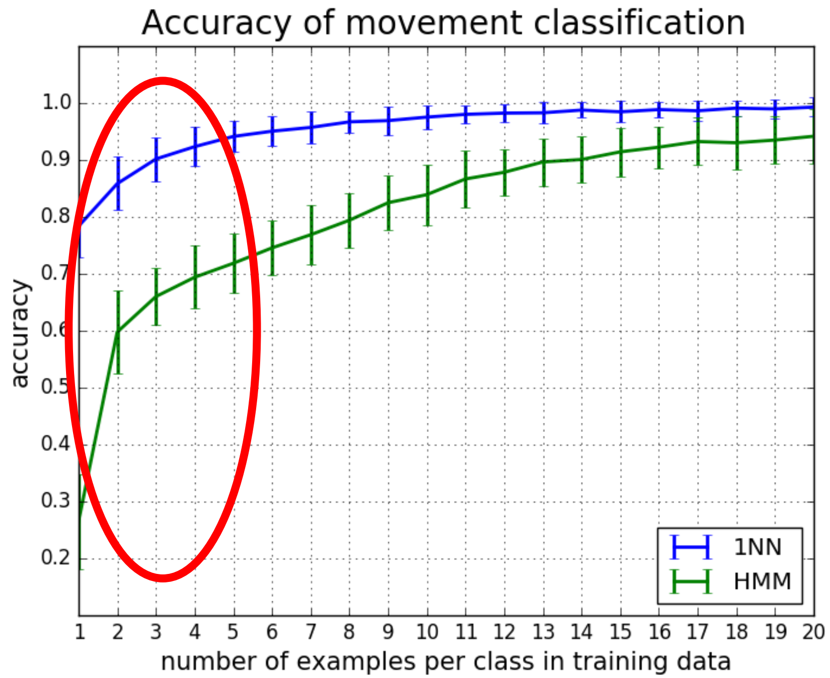




Few-shot Movement Recognition

Experiments: Comparison 1-NN vs. Hidden Markov Models (HMMs) on pick-and-place movements



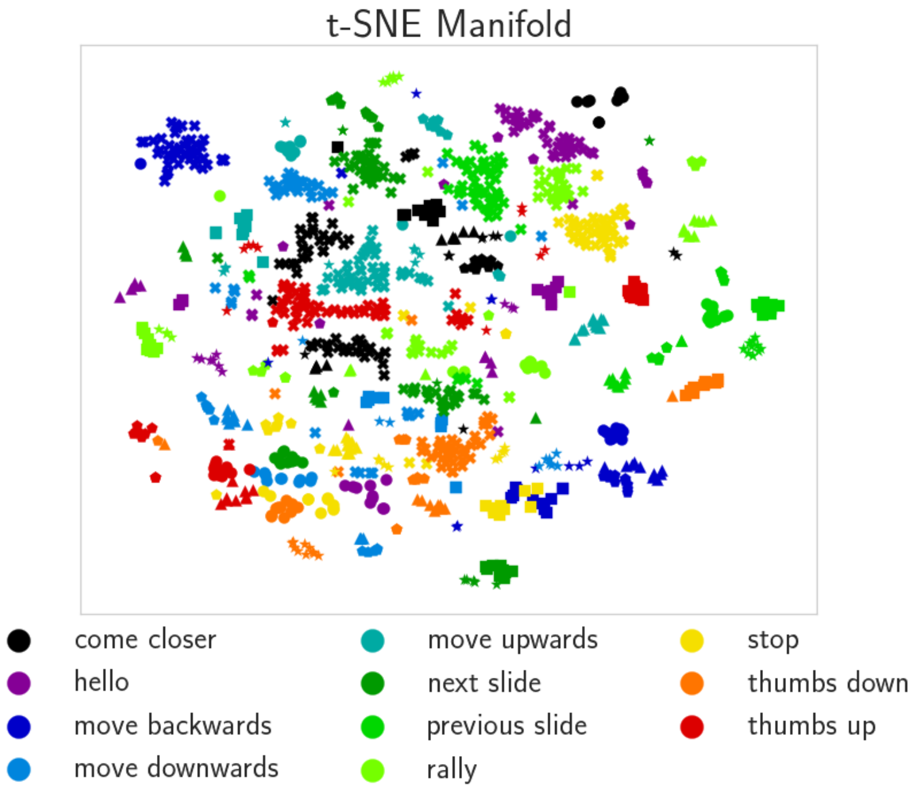


Few-shot Movement Recognition

Experiments: Comparison of Long-Short Term Memory Networks (LSTMs), HMMs and k-NN

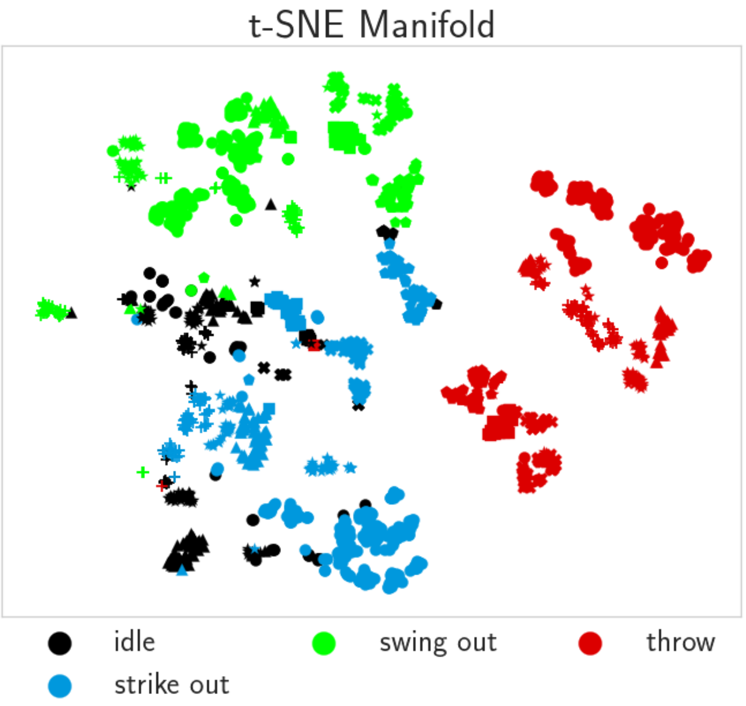
gesture data

- 11 gestures, 6 subjects
- approx. 95 demonstrations of each gesture



stick-throwing data

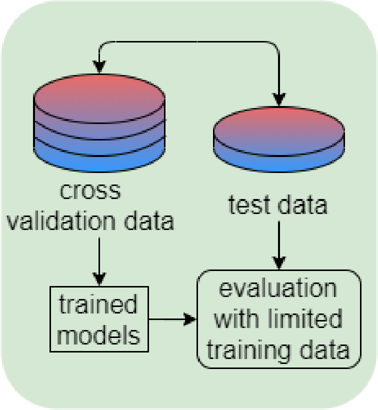
- 4 building blocks, 7 subjects
- in total 697 throwing demonstrations



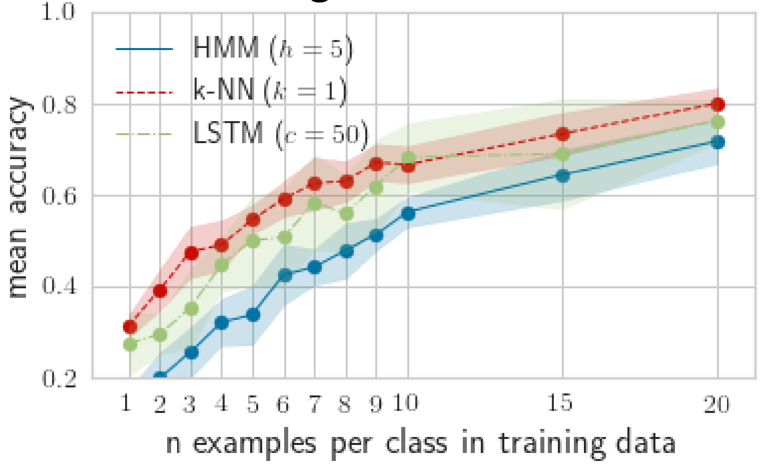
Few-shot Movement Recognition

Experiments: Comparison of LSTMs, HMMs and k-NN on data of different complexity - results

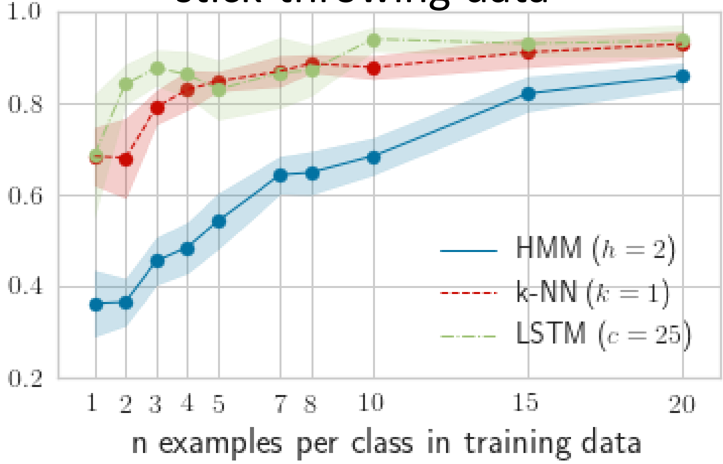
Experiment 1:



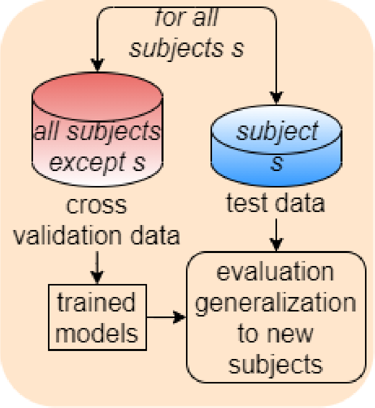
gesture data



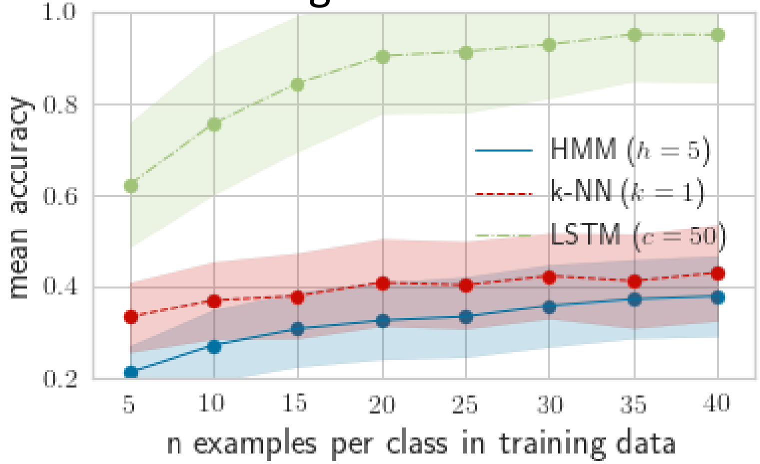
stick-throwing data



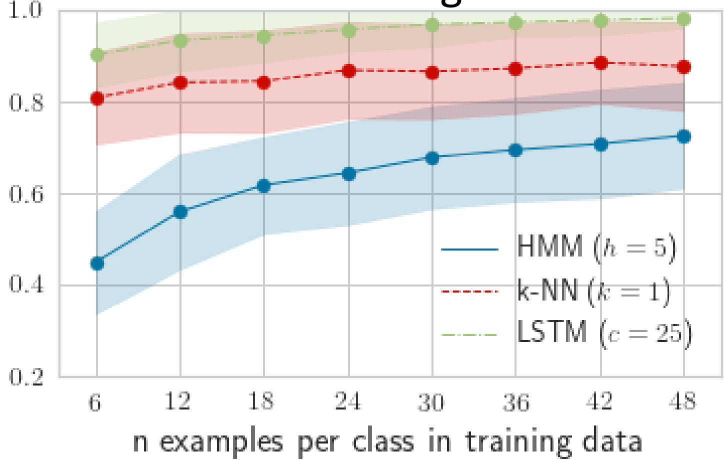
Experiment 2: generalization



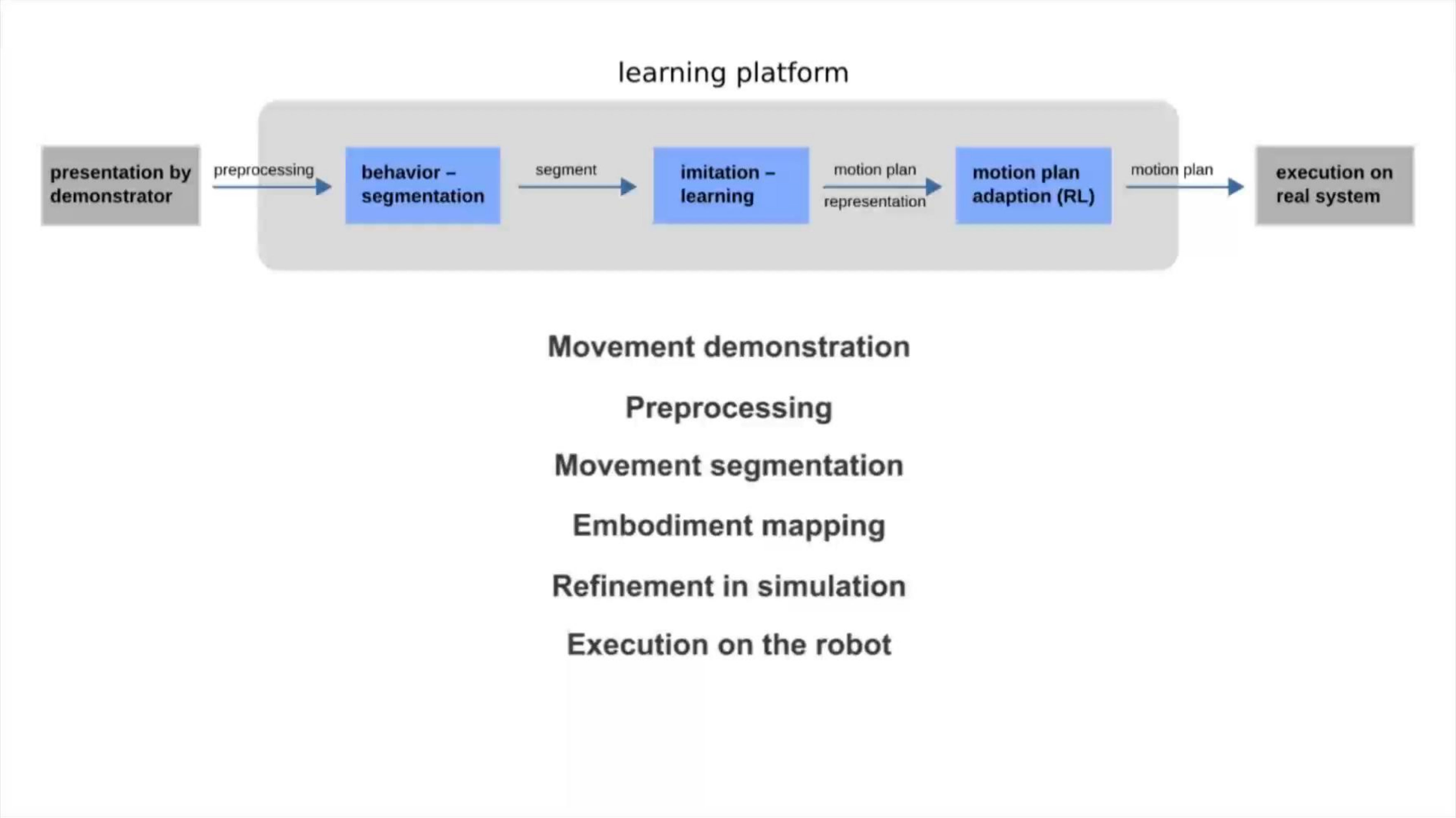
gesture data



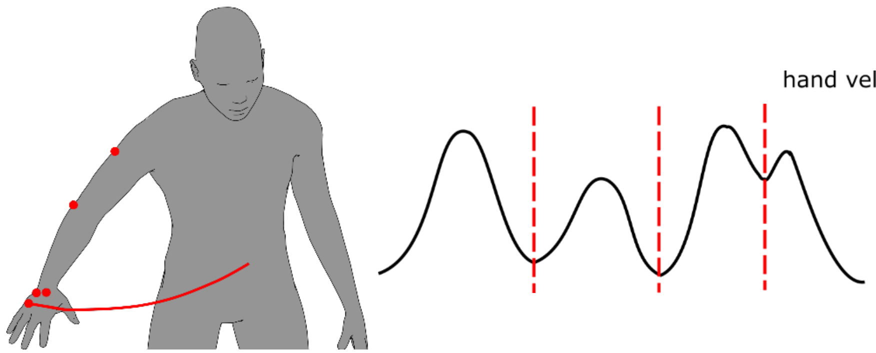
stick-throwing data



Applications
Robotic Learning from Demonstration



Thank you!



List of publications

Segmentation approach:

- Gutzeit & Kirchner: *“Unsupervised Segmentation of Human Manipulation Movements into Building Blocks“,* IEEE Access, 2022
- Senger et al.: *“Velocity-based Multiple Change-point Inference for Unsupervised Segmentation of Human Movement Behavior“,* ICPR 2014

Segment Classification:

- Gutzeit et al.: *“Simple and Robust Automatic Detection and Recognition of Human Movement Patterns in Tasks of Different Complexity“,* Physiological Computing Systems, Springer, 2016
- Gutzeit: “*A Comparison of Few-Shot Classification of Human Movement Trajectories*”, ICPRAM 2021
- Gutzeit & Kirchner: *“Automatic Detection and Recognition of Human Movement Patterns in Manipulation Tasks“*, PhyCS 2016

Robotic Learning from Demonstration:

- Gutzeit et al.: “*Automated Robot Skill Learning from Demonstration for Various Robot Systems*”, KI 2019
- Gutzeit et al.: *“The BesMan Learning Platform for Automated Robot Skill Learning“,* Frontiers in Robotics and AI, 2018

References

[Adi-Japha et al. 2008] “A shift in Task Routines during the Learning of a Motor Skill: Group-Averaged Data May Mask Critical Phases in the Individuals’ Acquisition of Skilled Performance", Journal of Experimental Psychology: Learning, Memory and Cognition, vol.24, pp. 1544-51, 2008.

[Graybiel, 1998] “The basal ganglia and chunking of action repertoires", Neurobiology of Learning and Memory, vol.70, pp. 119–136, 1998.

[Fearnhead and Liu 2007] "On-line Inference for Multiple Change Point Models", Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 69, pp. 589-605, 2007.

[Fox et al. 2010] "Sharing Features among Dynamical Systems with Beta Processes", Neural Information Processing Systems 22, 2010.

[Lioutikov et al, 2017] “Learning movement primitive libraries though probabilistic segmentation”, International Journal of Robotics Research, vol. 36, pp. 879-894, 2017.

[Mussa-Ivaldi & Solla, 2004] “Neural primitives for motion control”, IEEE Journal of Oceanic Engineering, vol. 29(3), pp. 640–650., 2004.

[Sakai, 2003] “Chunking during human visuomotor sequence learning", Experimental Brain Research, vol. 152, pp. 229-242. 2003.

[Shadmer & Wise, 2005] “The Computational Neurobiology of Reaching and Pointing – A Foundation for Motor Learning", MIT Press, 2005.

[Wolpert and Ghahramani, 2000] “Computational principles of motor neuroscience”, Nature Neuroscience, 3, pp. 1212–1217, 2000

References 2

[Liu et al., 2018] “Skeleton-based action recognition using spatio-temporal LTSM networks with trust gates”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, pp.3007-3021, 2018

[Lo Presti & La Cascia, 2016] “3D skeleton - based human action classification: A survey”, Pattern Recognition, vol. 53, pp.130-147, 2016

[Morasso, 1981] "Spatial Control of Arm Movements", Experimental Brain Research, vol. 42, pp. 223-227, 1981.

[Mussa-Ivaldi et al., 1994] "Linear combinations of primitives in vertebrate motor control", Proc. Natl. Acad. Aci. USA, vol. 91, pp. 7534-7538, 1994.

[Mussa-Ivaldi & Bizzi, 2000] “Motor Learning through the combination of primitives”, Philosophical transactions of the Royal Society of London, Series B, Biological sciences, vel. 355, pp. 1755-69, 2000.

[Poppe et al., 2010] “A survey on vision-based human action recognition”, Image and Vision Computing, vol. 28, pp. 976-990, 2010

[Sakai, 2003] “Chunking during human visuomotor sequence learning", Experimantal Brain Research, vol. 152, pp. 229-242. 2003.

[Shadmer & Wise, 2005] “The Computational Neuobiology of Reaching and Pointing – A Foundation for Motor Learning", MIT Press, 2005.

[Schaal, 2002] "Arm and Hand Movement Control", The Handbook of Brain Theory and Neural Networks, MIT Press, pp. 110-113, 2002.

[Stefanov et al., 2010] “Online intention recognition in computer-assisted teleoperation systems”, Proceedings IEEE International Conference on Robotics and Automation, pp. 5334-5339, 2010

[Wolpert and Ghahramani, 2000]] “Computational principles of motor neuroscience”, Nature Neuroscience, 3, pp. 1212–1217, 2000