Test

Gesture-Sound Interaction and Embodied Music Cognition

from embodied performance to data to interaction design using machine learning

Federico Ghelli Visi

Universität der Künste Berlin

Sound Studies and Sonic Arts | Summer 2022

Gesture and Sound

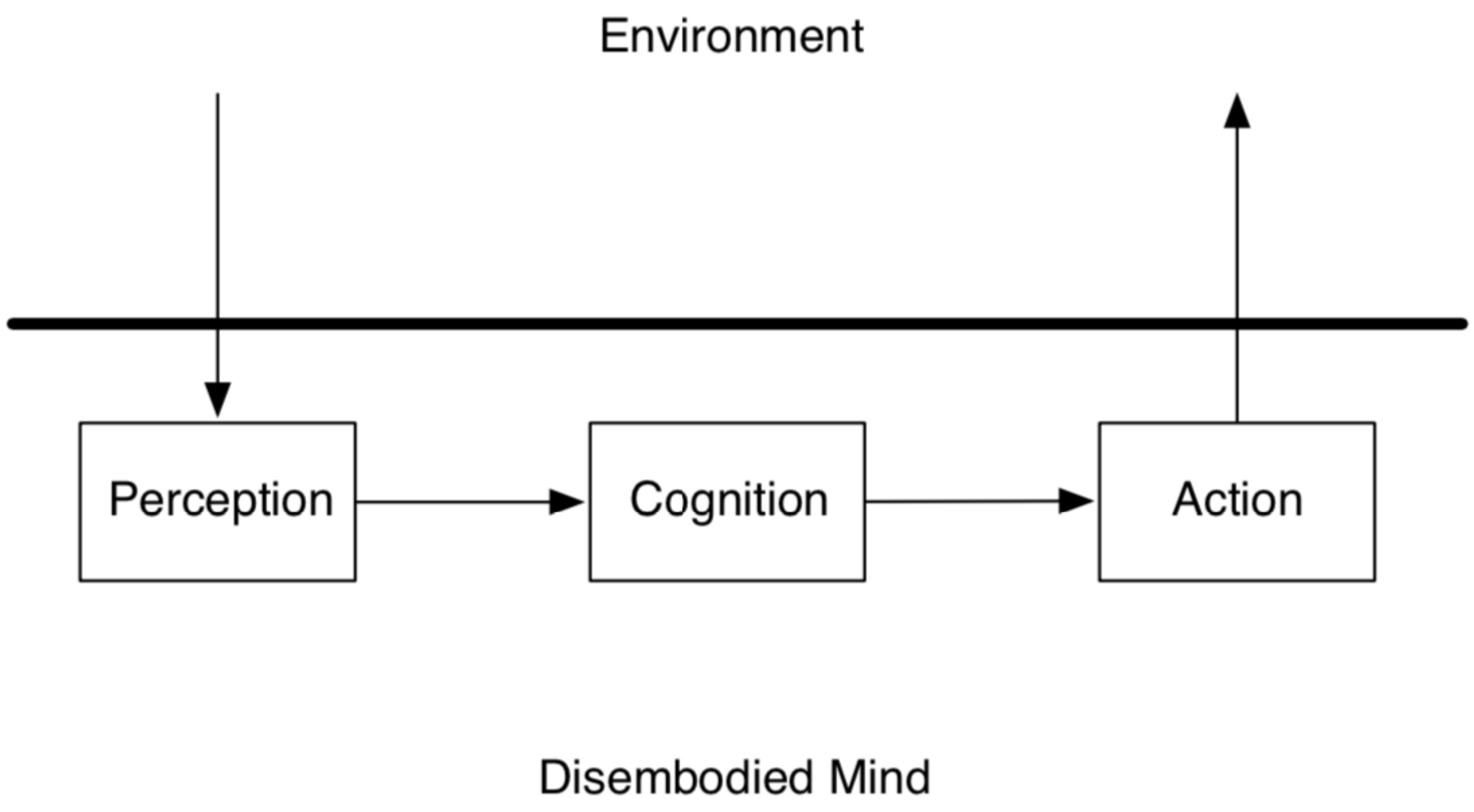

steim

Embodied Music Cognition

- Gesture and corporeal imitation are fundamental constituents of musical expression
- A repertoire of gestures linked with the resulting experiences and sensations: a *gesture/action-oriented ontology*
- Experiencing sound through embodied action.

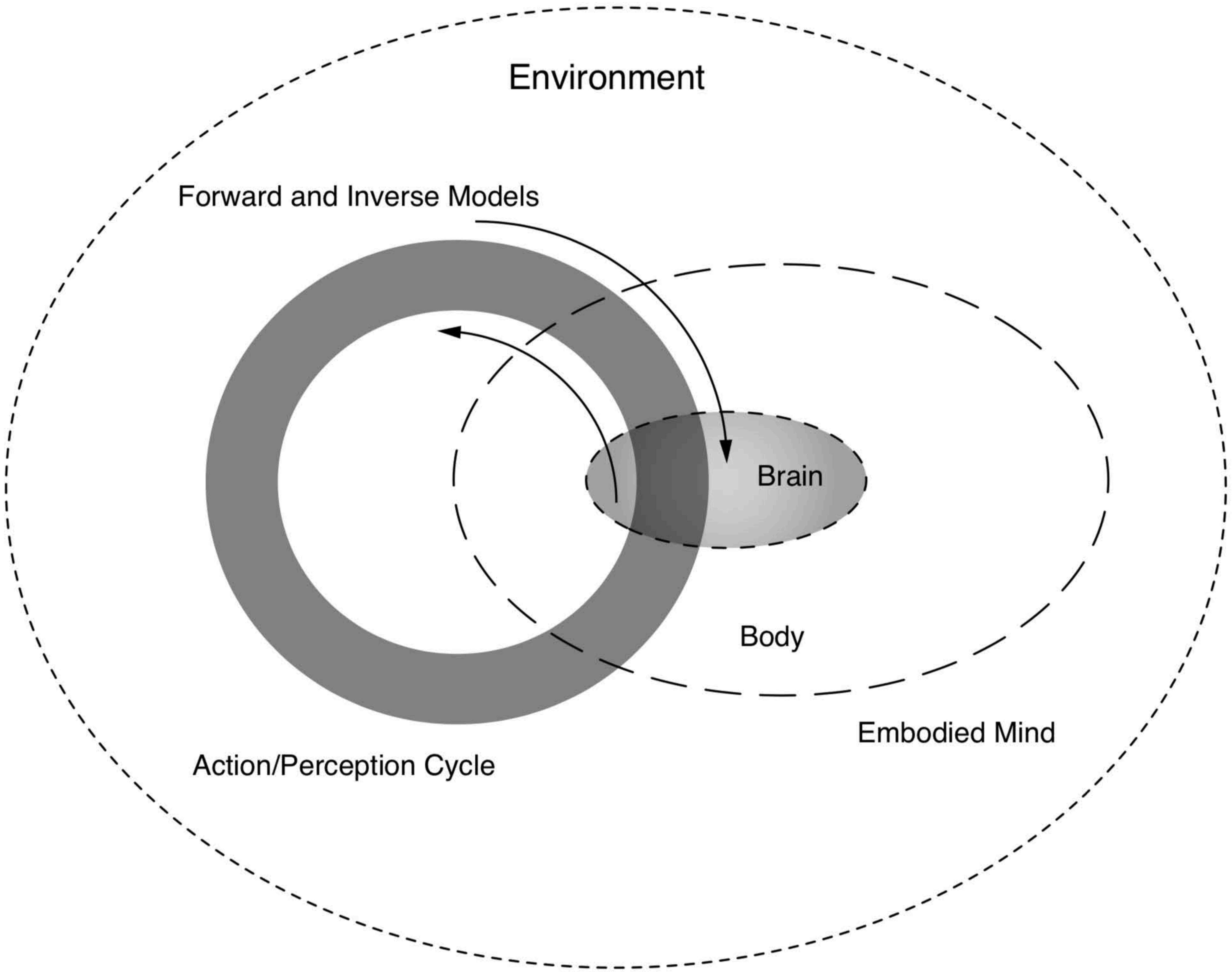
- Godøy, R. I., & Leman, M. (Eds.). (2010). Musical Gestures: Sound, Movement, and Meaning. Routledge.
- Leman, M. (2008). Embodied Music Cognition and Mediation Technology. Cambridge, MA, USA: The MIT Press.
- Schiavio, a., & Menin, D. (2013). Embodied Music Cognition and Mediation Technology: A critical review. Psychology of Music, 41(6), 804–814. https://doi.org/10.1177/0305735613497169

Dis-embodied model



A graphical representation of a traditional cognitivist disembodied model of cognition. Mind and environment are clearly separated and they interact through a unidirectional process. Cognition is segregated between perception and action.

Embodied model



A graphical representation of embodied cognition.
Visi, F. (2017). Methods and Technologies for the Analysis and Interactive Use of Body Movements in Instrumental Music Performance. PhD Thesis. University of Plymouth. Retrieved from https://pearl.plymouth.ac.uk/handle/10026.1/8805

4E approach

conceives of the musical mind
as Embodied, Embedded, Extended,
and Enactive.

van der Schyff, D., Schiavio, A., Walton, A., Velardo, V., & Chemero, A. (2018). Musical creativity and the embodied mind. Music & Science, 1, 205920431879231. https://doi.org/10.1177/2059204318792319

Embodied

Considering the mind as embodied means rethinking the boundaries between the neural and extra-neural (e.g., metabolic, thermodynamic, and muscular, among others) factors that drive cognitive processes. From this perspective, ***the brain becomes a part of a larger network that involves the nervous system and the sensorimotor capacities of the entire organism*** *(e.g., Gallagher, 2005, 2011). In a sense, therefore, separating brain and body, perception and action, experience and behavior, may in fact be a largely artificial move that offers only limited accounts of what mental life really entails (Hurley, 1998, 2001; Thompson, 2007).*

van der Schyff, D., Schiavio, A., Walton, A., Velardo, V., & Chemero, A. (2018). Musical creativity and the embodied mind. Music & Science, 1, 205920431879231. https://doi.org/10.1177/2059204318792319

Embedded

As we have begun to consider, the body does not simply provide biological support for an otherwise detached brain that "commands" behavior. Rather, it participates in driving cognitive processes. But this does not happen in a vacuum. Indeed, ***embodied minds are parts of broader physical and socio-cultural systems that shape and are shaped by the agents who inhabit them.*** *As such, a growing number of scholars also consider cognition as embedded within such systems. This approach draws inspiration from the influential work of psychologist James Gibson (1966,1979), who explored* ***perception as an "ecological" process*** *(see Mace, 1977).*

van der Schyff, D., Schiavio, A., Walton, A., Velardo, V., & Chemero, A. (2018). Musical creativity and the embodied mind. Music & Science, 1, 205920431879231. https://doi.org/10.1177/2059204318792319

Enactive

The enactive dimension describes how organisms and their environments mutually determine each other (Varela, Thompson, & Rosch, 1991). Most centrally, this perspective highlights the active role living creatures play in developing patterns of (sensorimotor, neural, metabolic, interactive) activity that allow them to maintain a viable existence. ***Such sets of meaningful activities constitute what enactivists refer to as "sense-making," which is ultimately equated with "cognition"*** *(see Thompson & Stapleton, 2009). The enactive approach, therefore,* <u>***replaces the more traditional input–output model of mind with a more relational story—where an agent's ongoing history of interactivity (structural coupling) with the environment becomes central to his or her mental life.***</u>

van der Schyff, D., Schiavio, A., Walton, A., Velardo, V., & Chemero, A. (2018). Musical creativity and the embodied mind. Music & Science, 1, 205920431879231. https://doi.org/10.1177/2059204318792319

Extended

The "extended" dimension of the 4E framework holds that "[biological and non-biological] features of the environment can co-constitute the mind" (Hutto & Myin 2013, p. 139; see also, Clark, 2010). Think, for example, of a percussionist improvising on an arrangement of instruments. In the process of enacting a meaningful relationship with this collection of instruments, ***<u>the instruments must become part of the musician's cognitive ecology.</u>*** *[…] This is to say that, when needed,* ***<u>tools and objects from the environment can become integrative parts ofmental life and the creative processes</u>*** *that go along with it (see Malafouris, 2008, 2013, 2015).*

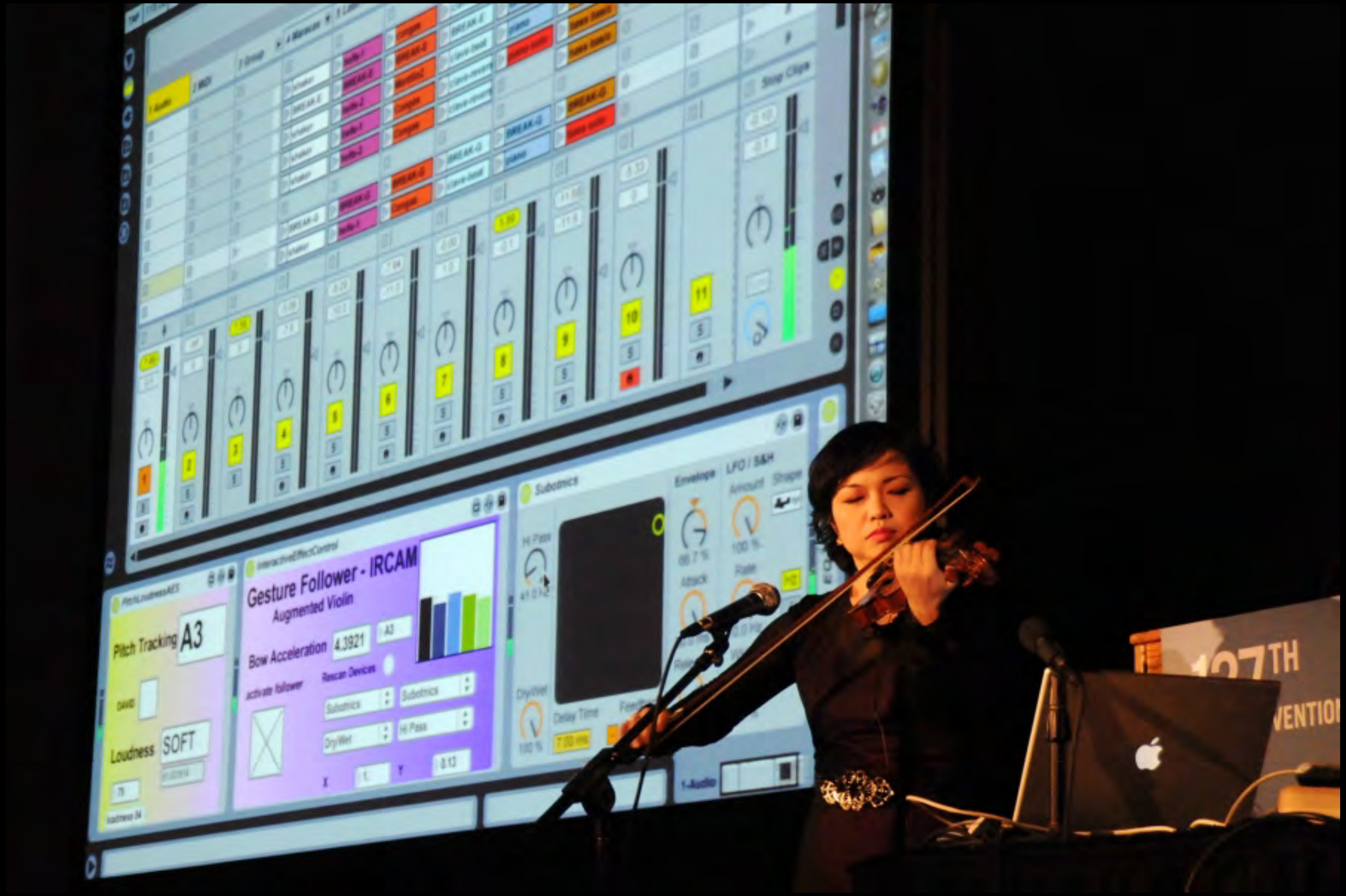
van der Schyff, D., Schiavio, A., Walton, A., Velardo, V., & Chemero, A. (2018). Musical creativity and the embodied mind. Music & Science, 1, 205920431879231. https://doi.org/10.1177/2059204318792319

Movement into Data


ASUS

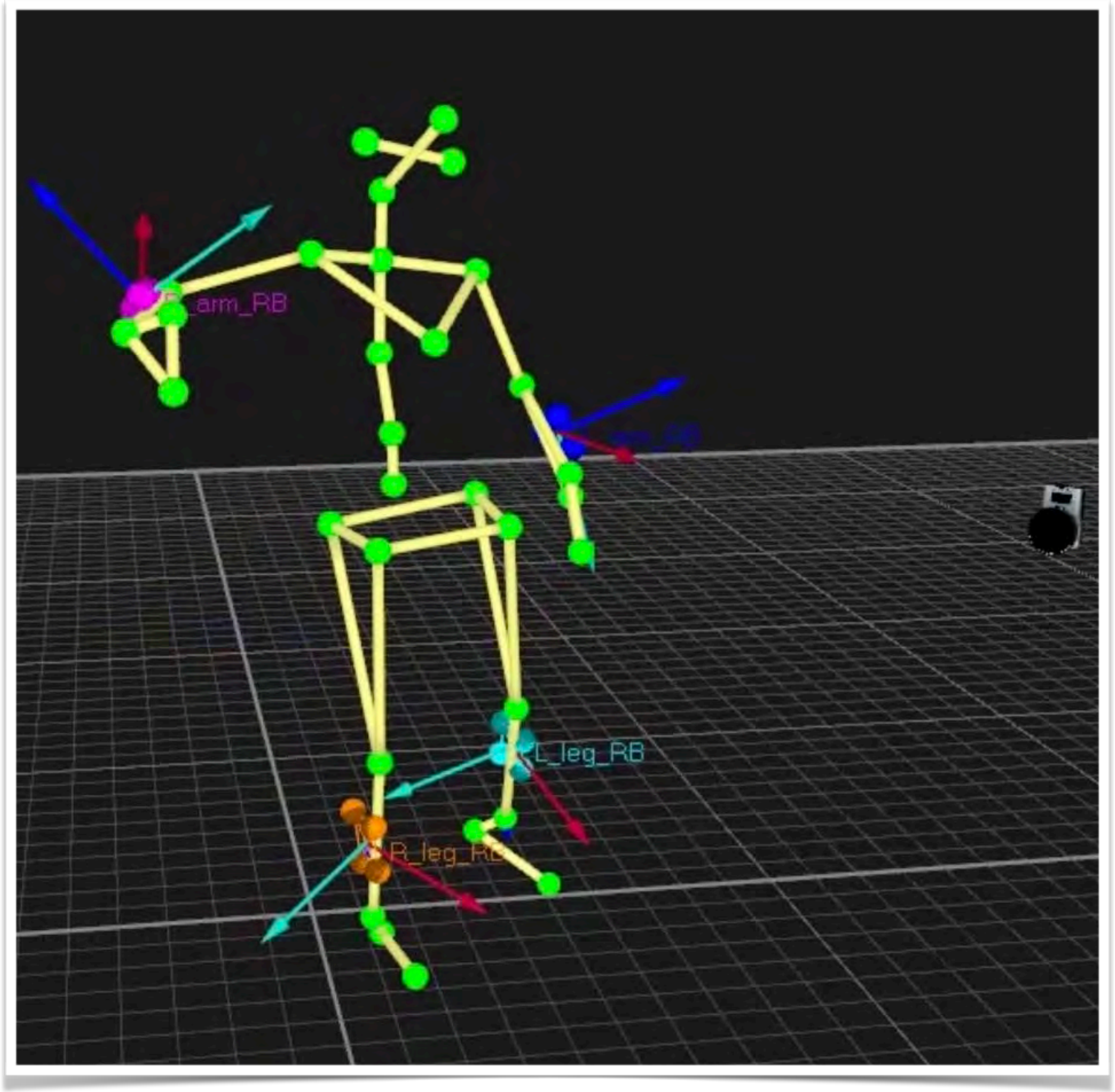
MUGIC sensors

Mari Kimura


Gesture Follower - IRCAM
Augmented Violin
Bow Acceleration
4.3921
Pitch Tracking
A3
Loudness
SOFT

Motion Descriptors

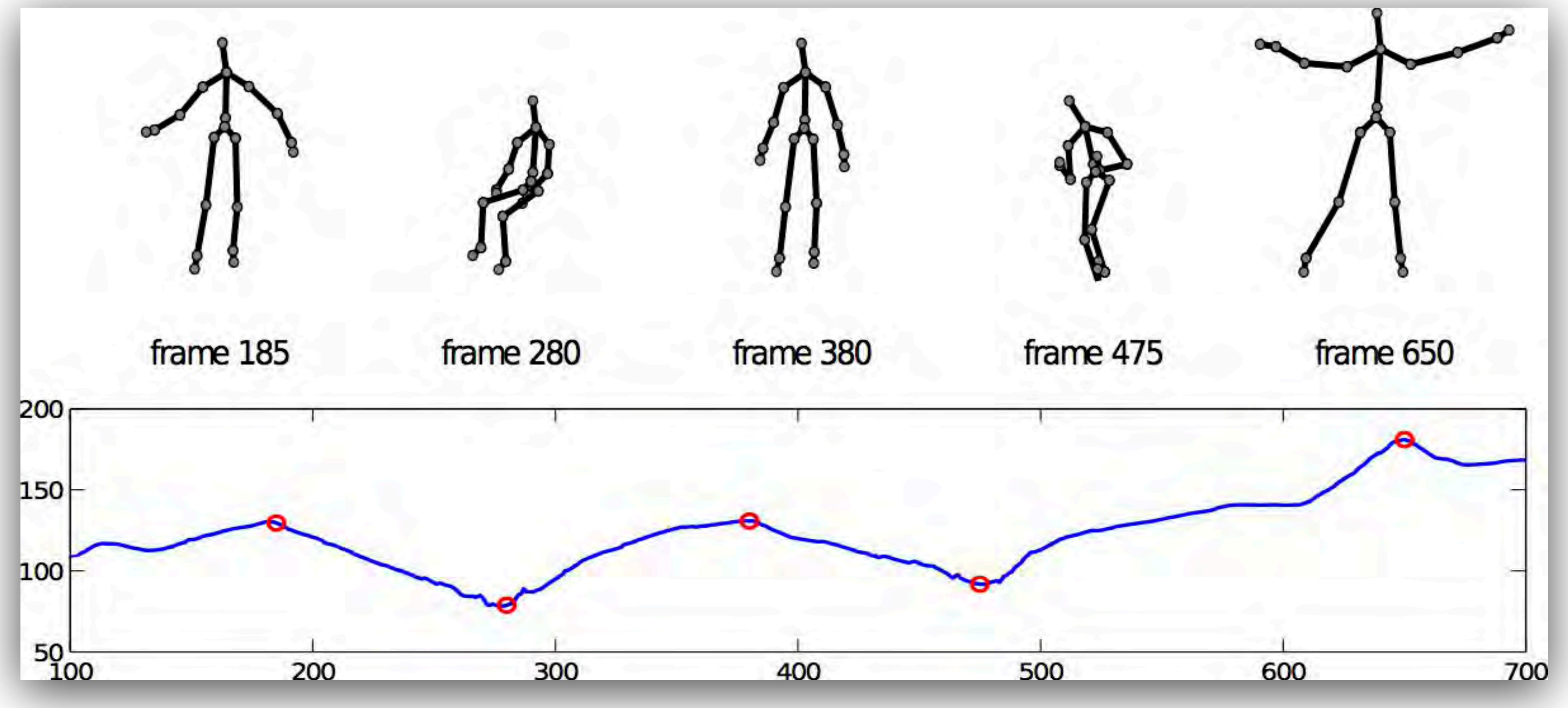
- Motion descriptors extract meaningful quantities from raw data.
- Often used in expressive movement analysis, interaction design, as features for machine learning.



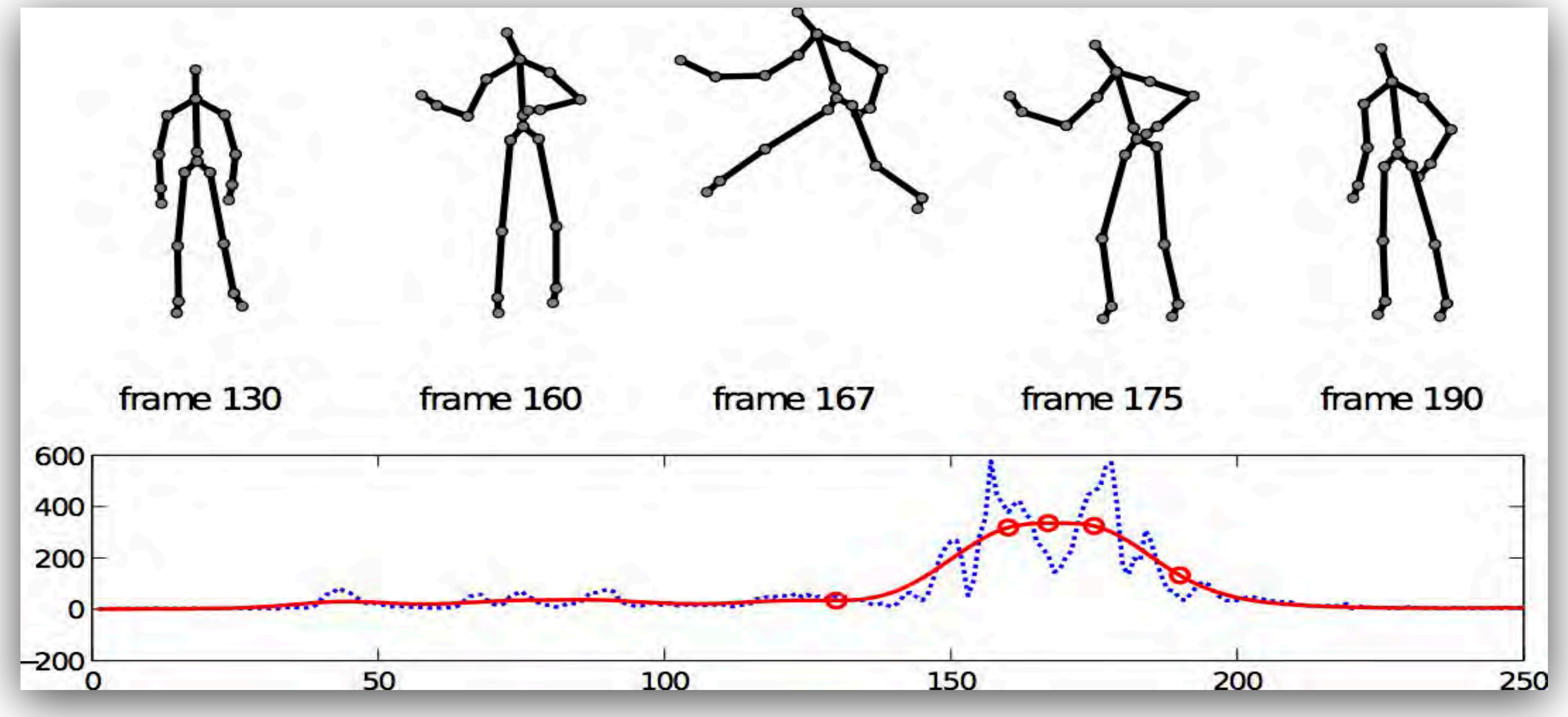
- Piana et al.. 2016. *Adaptive Body Gesture Representation for Automatic Emotion Recognition.*
- Glowinski et al. 2011. *Toward a Minimal Representation of Affective Gestures.*
- Larboulette & Gibet. 2015. *A review of computable expressive descriptors of human motion.*
- and more!

Full-body motion analysis

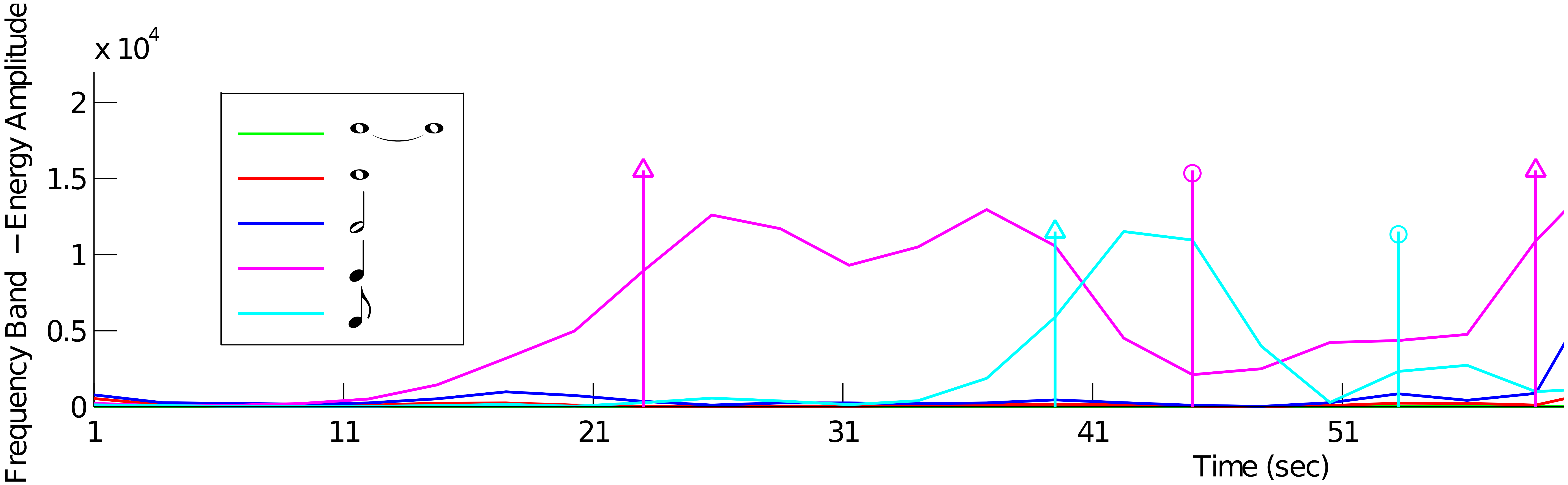
Contraction Index



Quantity of Motion

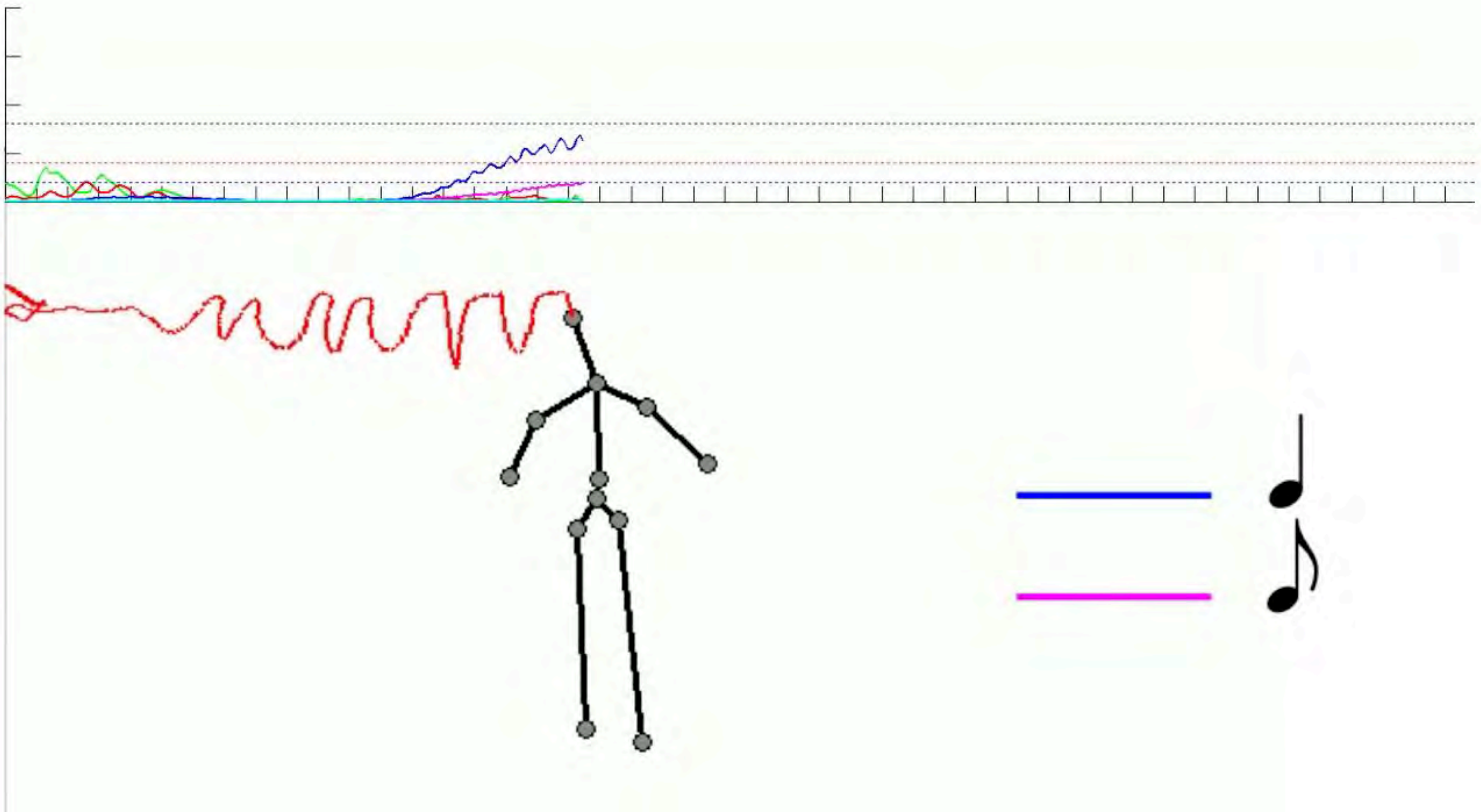


Periodic Quantity of Motion (PQoM)



- Quantity of motion in relation to the rhythmic subdivisions in the musical excerpts
- PQoM is obtained by decomposing the motion capture signal into frequency components
- Useful for exploring resonance and entrainment

Periodic Quantity of Motion (PQoM)

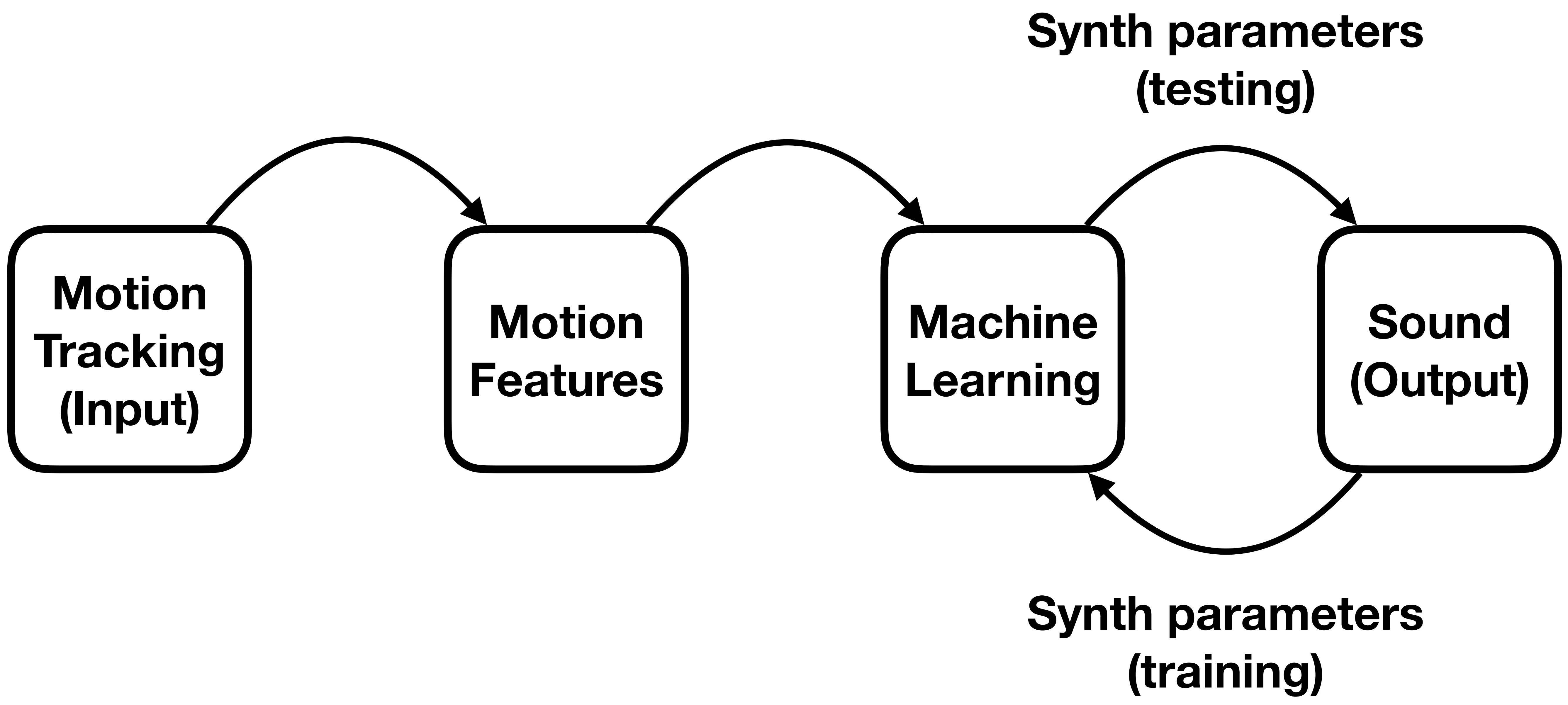



sonif_2.qtm - Qualisys Track Manager - [Project: TEST_1]
File Edit View Play Capture AIM Tools Window Help
Labeled trajectories tracked
2D
ID
R_leg_RB
L_leg_RB
Messages
2018-05-30 10:54:38 File closed
2018-05-30 10:54:38 File loaded from D:\QTM projects\TEST_1\Data\dance_3-2_proto_V2\sonif_3.qtm
2018-05-30 11:00:37 File closed
2018-05-30 11:00:37 File loaded from D:\QTM projects\TEST_1\Data\dance_3-2_proto_V2\sonif_2.qtm
2018-05-30 11:06:07 File saved as D:\QTM projects\TEST_1\Data\dance_3-2_proto_V2\sonif_2.qtm
NIME_video_modosc_getData (unlocked)
File Edit View Object Arrange Options Debug Extras Window Help
150%
udpreceive 1313 CNMAT
Getting descriptor data out
o.route /modosc/points/L_arm_RB/vel_mag /modosc/groups/limbs_gr/ci
1.380
Left arm velocity
0.00207
Contraction index
for limbs group

Interactive Machine Learning

Interactive Machine Learning Model

What: control gestures



What: sounds

<u>Show</u> the computer how gesture and sound are related by using machine learning

Fiebrink, R. A., & Caramiaux, B. (2018). The Machine Learning Algorithm as Creative Musical Tool. (R. T. Dean & A. McLean, Eds.), The Oxford Handbook of Algorithmic Music (Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190226992.013.23

GIMLeT - Gestural Interaction Machine Learning Toolkit

Demo

GIMLeT

Gestural Interaction Machine Learning Toolkit

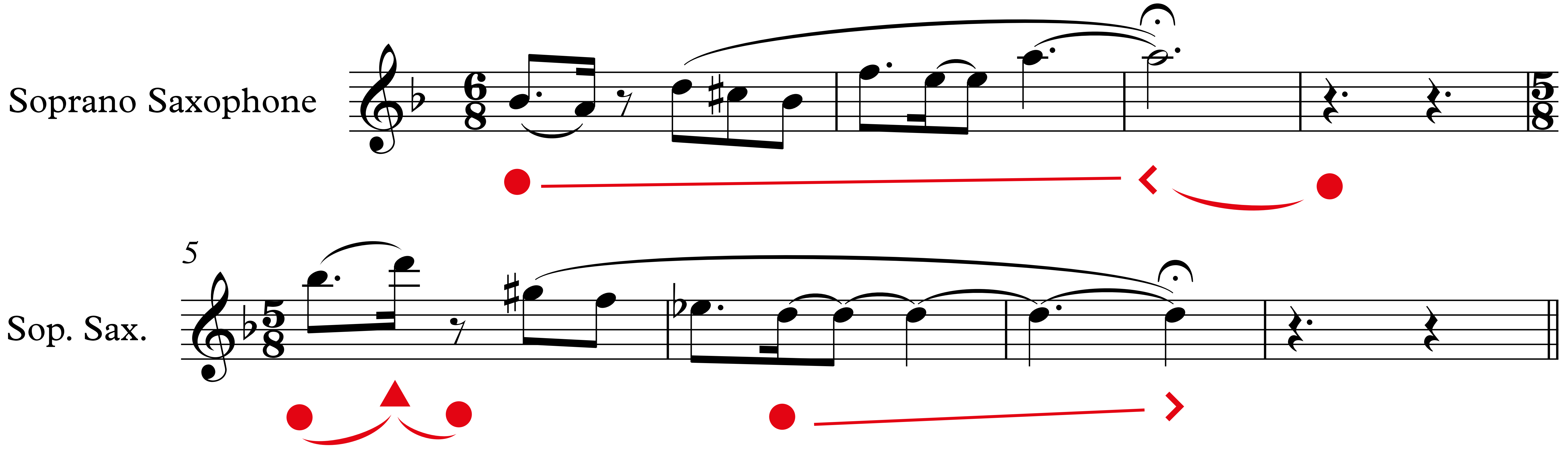
- https://github.com/federicoVisi/GIMLeT

- Video Tutorials
  - Installation and linear regression with artifical neural networks: https://youtu.be/Dace1sHy1IM
  - Gesture following with PoseNet and GVF: https://youtu.be/GoNqiCvVgoY

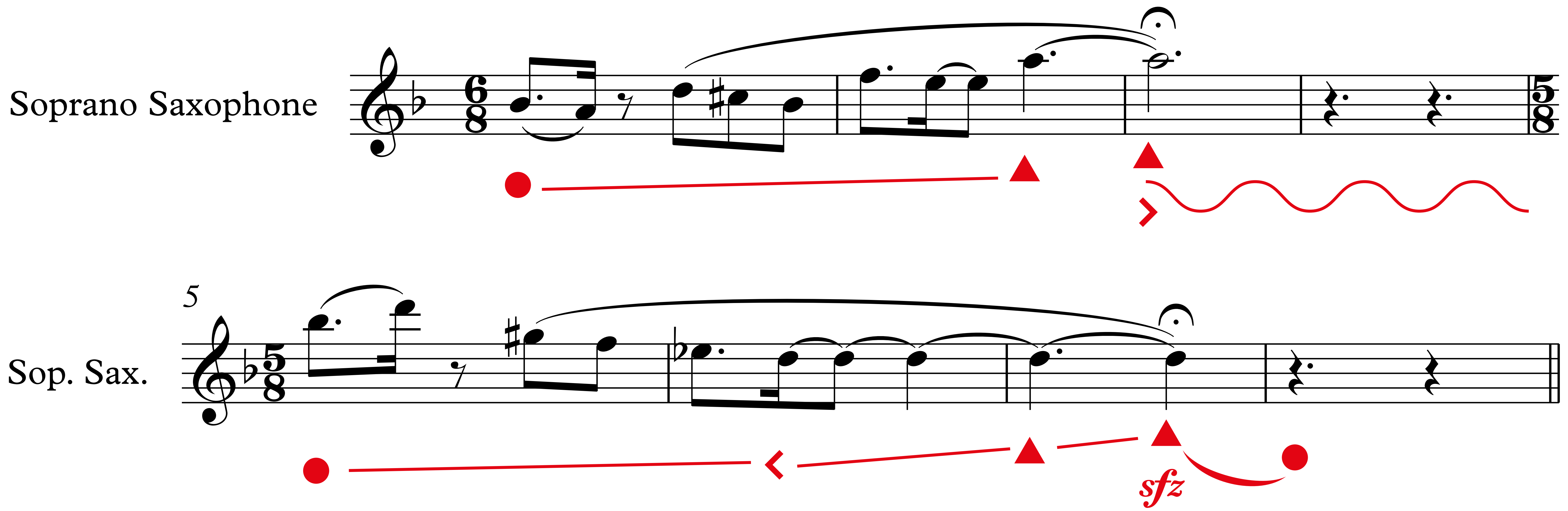
Artistic Practice:
on Movement and Notation

11 degrees of dependence

Section 3a

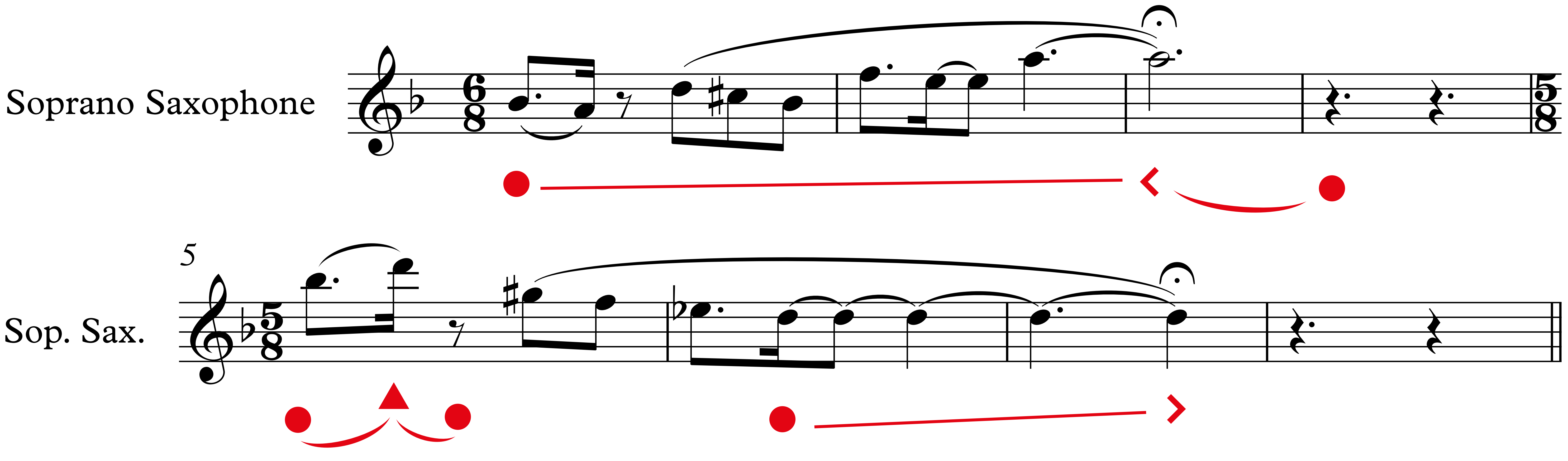


Section 3a (2nd time)

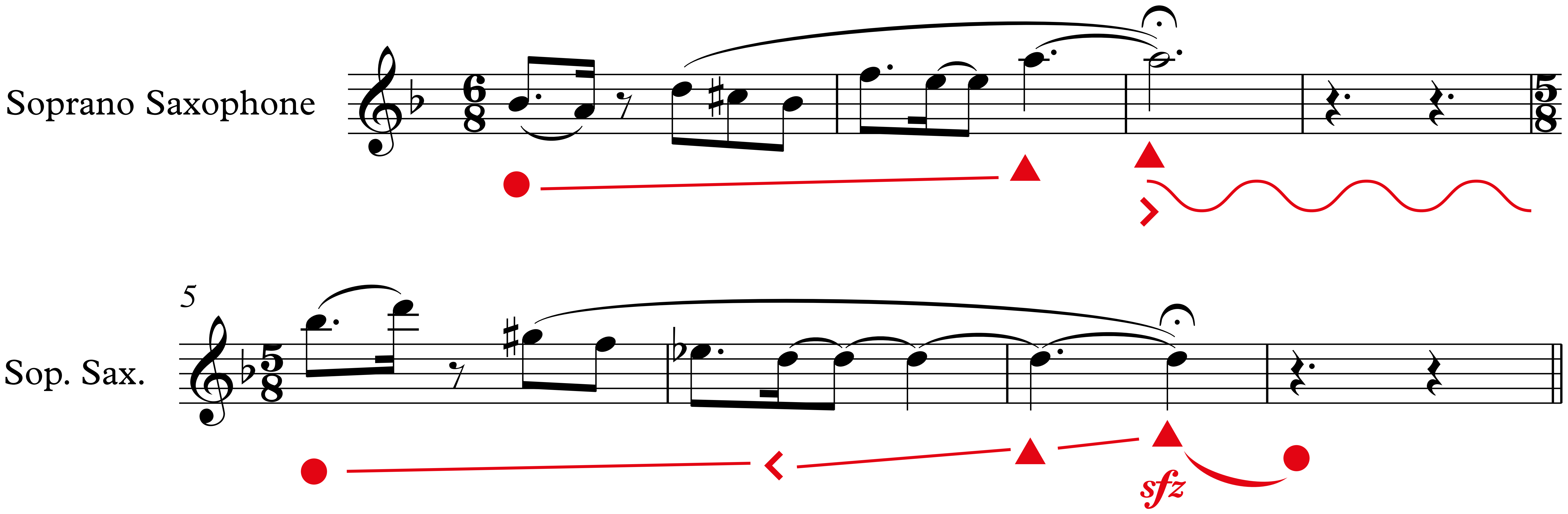


11 degrees of dependence

Section 3a



Section 3a (2nd time)



Connecting sound, bodies, and actions

Noise Bodies (1965) (archival photograph from Schneemann's Noise Bodies performance work, 3rd Annual Festival of the Avant Garde, Judson Hall, New York, NY, August 28, 1965). Performers: James Tenney and Carolee Schneemann.

"It was a noisy collage. We improvised together regarding what made sound and what gestures would produce varieties of sound. The way my kinetic theatre pieces developed was that parameters were set in terms of certain kinds of duration, position and action and then from studying those we would improvise."

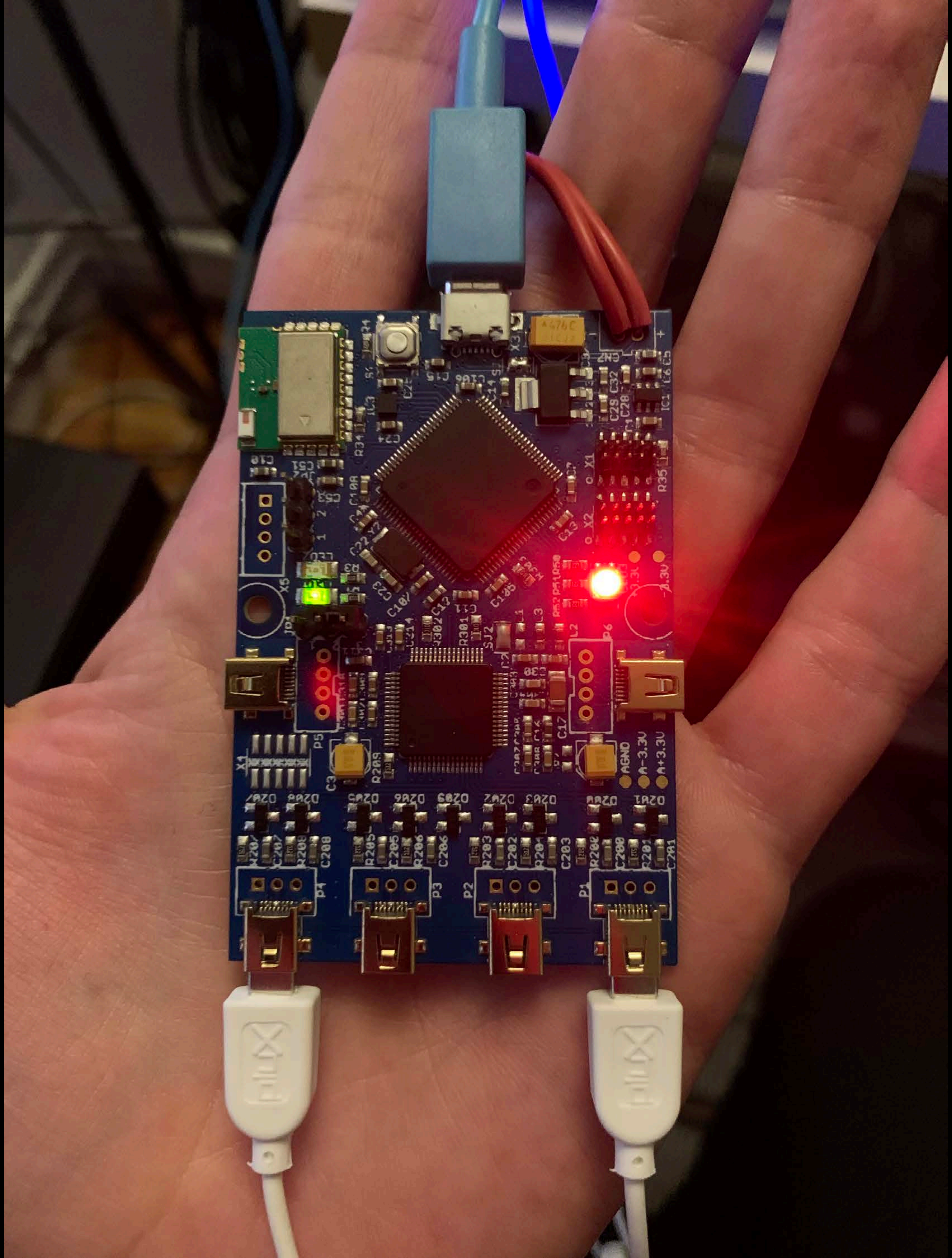
– Carolee Schneemann

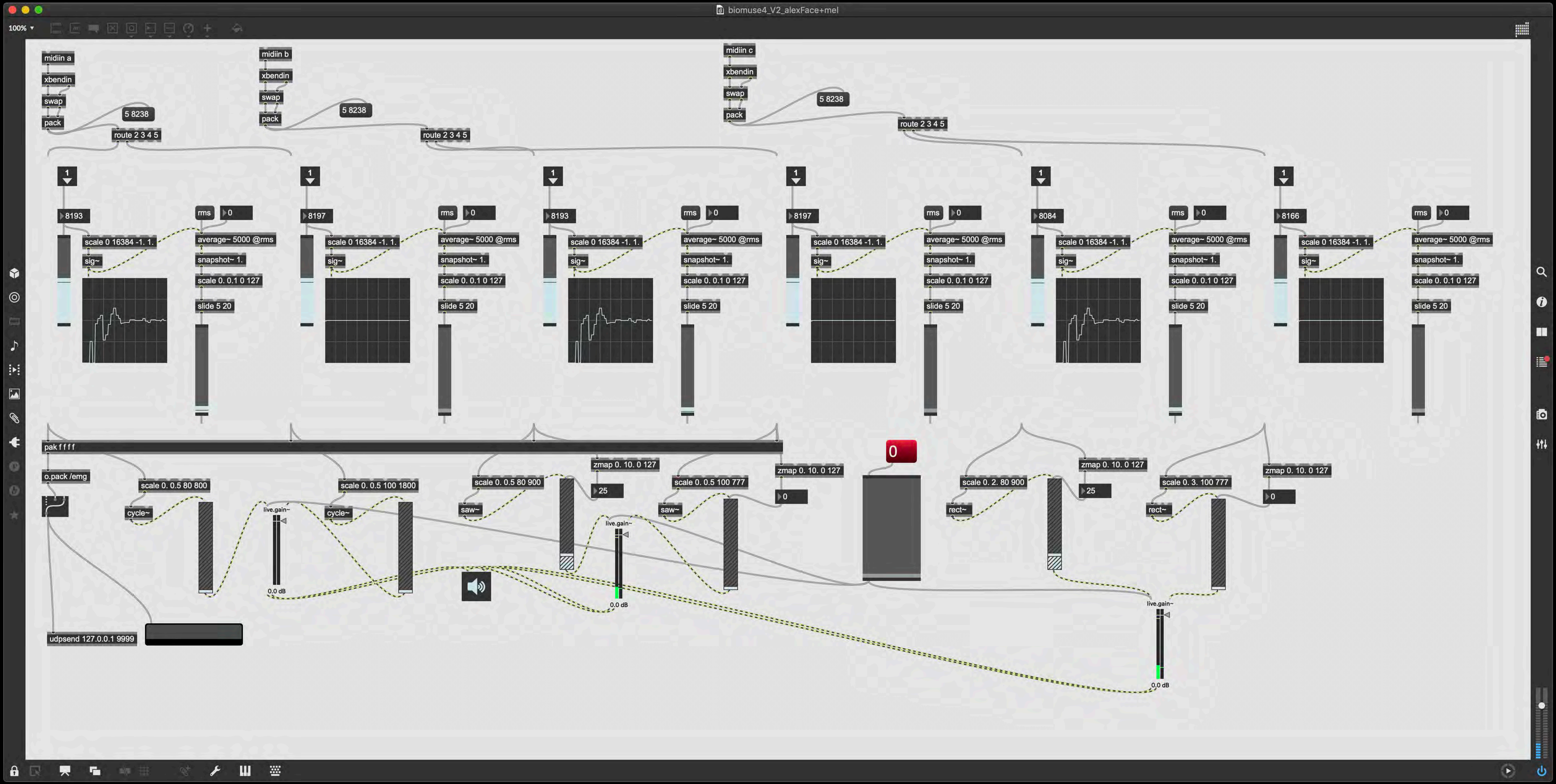
Notes on Fuseology:
Carolee Schneemann Remembers James Tenney
https://bordercrossingsmag.com/article/notes-on-fuseology

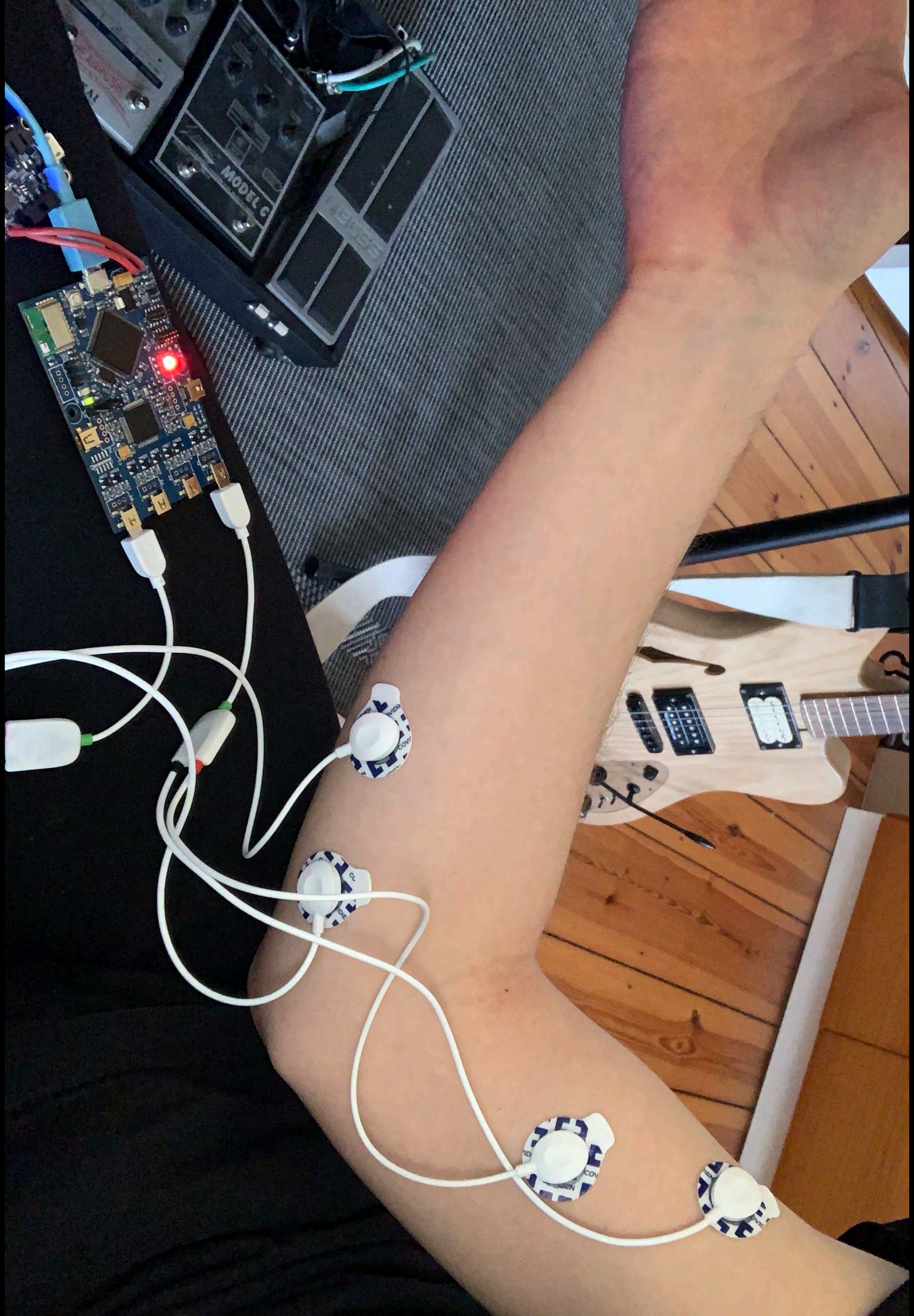
Noise Bodies (1965) (archival photograph from Schneemann's Noise Bodies performance work, 3rd Annual Festival of the Avant Garde, Judson Hall, New York, NY, August 28, 1965). Performers: James Tenney and Carolee Schneemann.

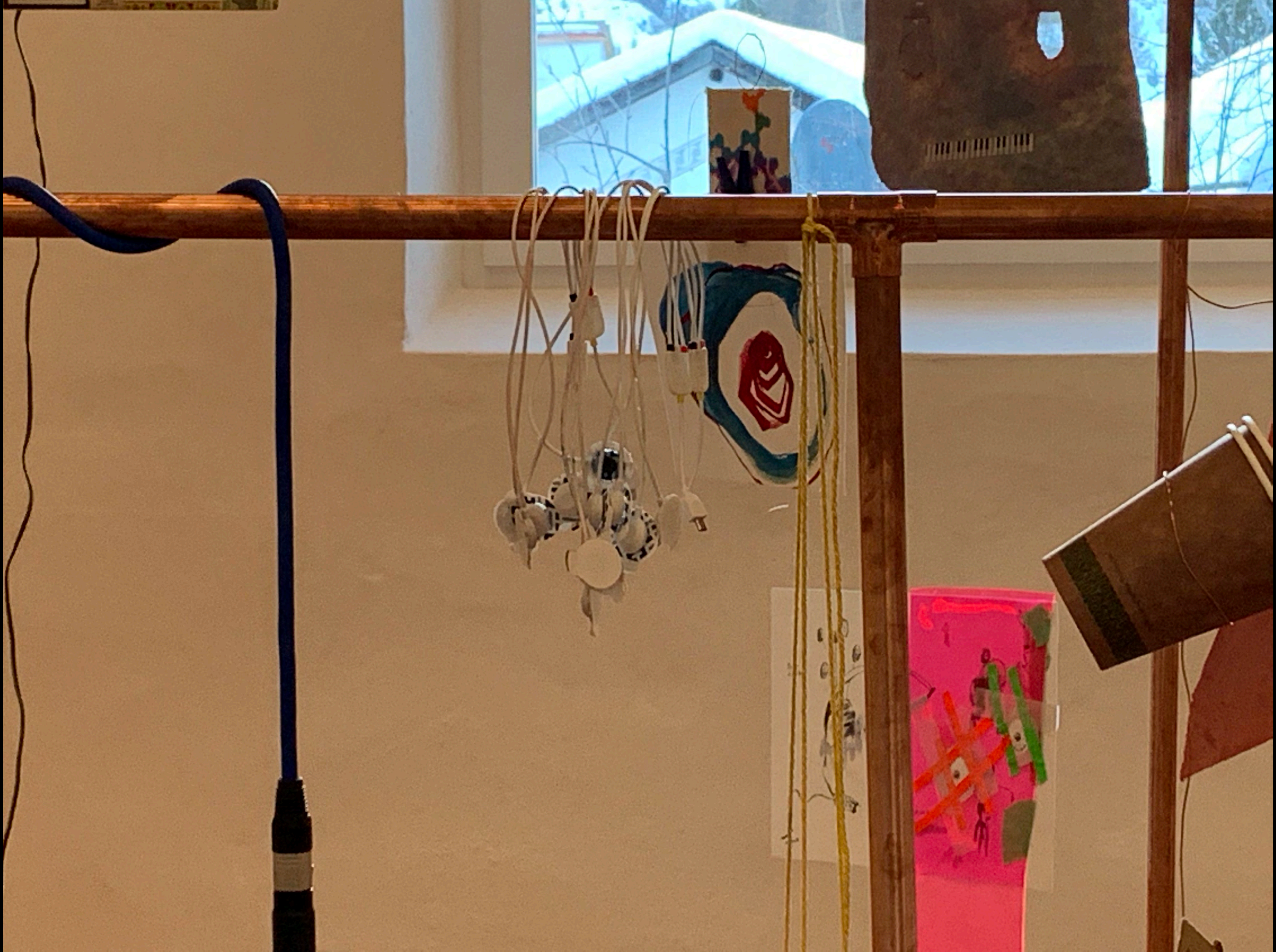
Atau Tanaka

https://youtu.be/G6H1J2k--5I?t=103




biomuse4_V2_alexFace+mel
100%
midiin a
xbendin
swap
pack
5 8238
route 2 3 4 5
midiin b
midiin c
8193
8197
8084
8166
rms
scale 0 16384 -1. 1.
sig~
average~ 5000 @rms
snapshot~ 1.
scale 0. 0.1 0 127
slide 5 20
pak f f f f
o.pack /emg
scale 0. 0.5 80 800
cycle~
live.gain~
0.0 dB
scale 0. 0.5 100 1800
scale 0. 0.5 80 900
saw~
zmap 0. 10. 0 127
25
scale 0. 0.5 100 777
scale 0. 2. 80 900
rect~
scale 0. 3. 100 777
udpsend 127.0.0.1 9999



Exploring Gesture-Sound Mappings Using Reinforcement Learning

Assisted Interactive Machine Learning

What is it?

In a sentence:

Assisted Interactive Machine Learning (AIML) is an interaction design method based on **Deep Reinforcement Learning** developed for the purpose of **exploring the many possible mappings** between two heterogeneous spaces (e.g. gesture and sound synthesis).

Why?

- Machine learning has made designing live interfaces for artistic expression more accessible, allowing rapid creative experimentation and enabling new creative relationships with machines.

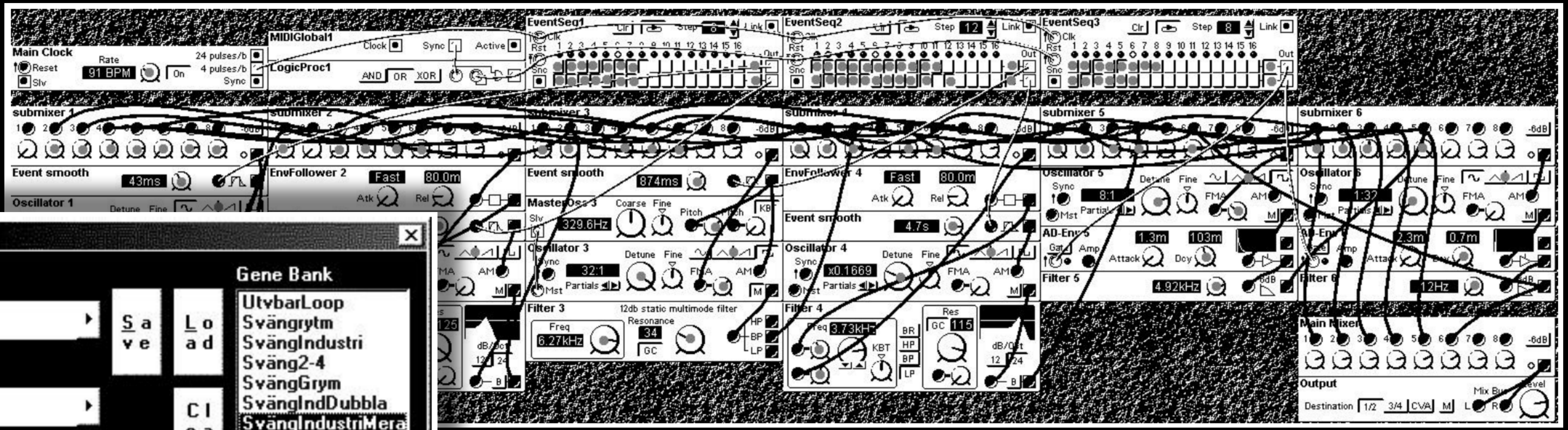
- Digital interfaces can be programmed in so many different ways (there are **many possible mappings**).

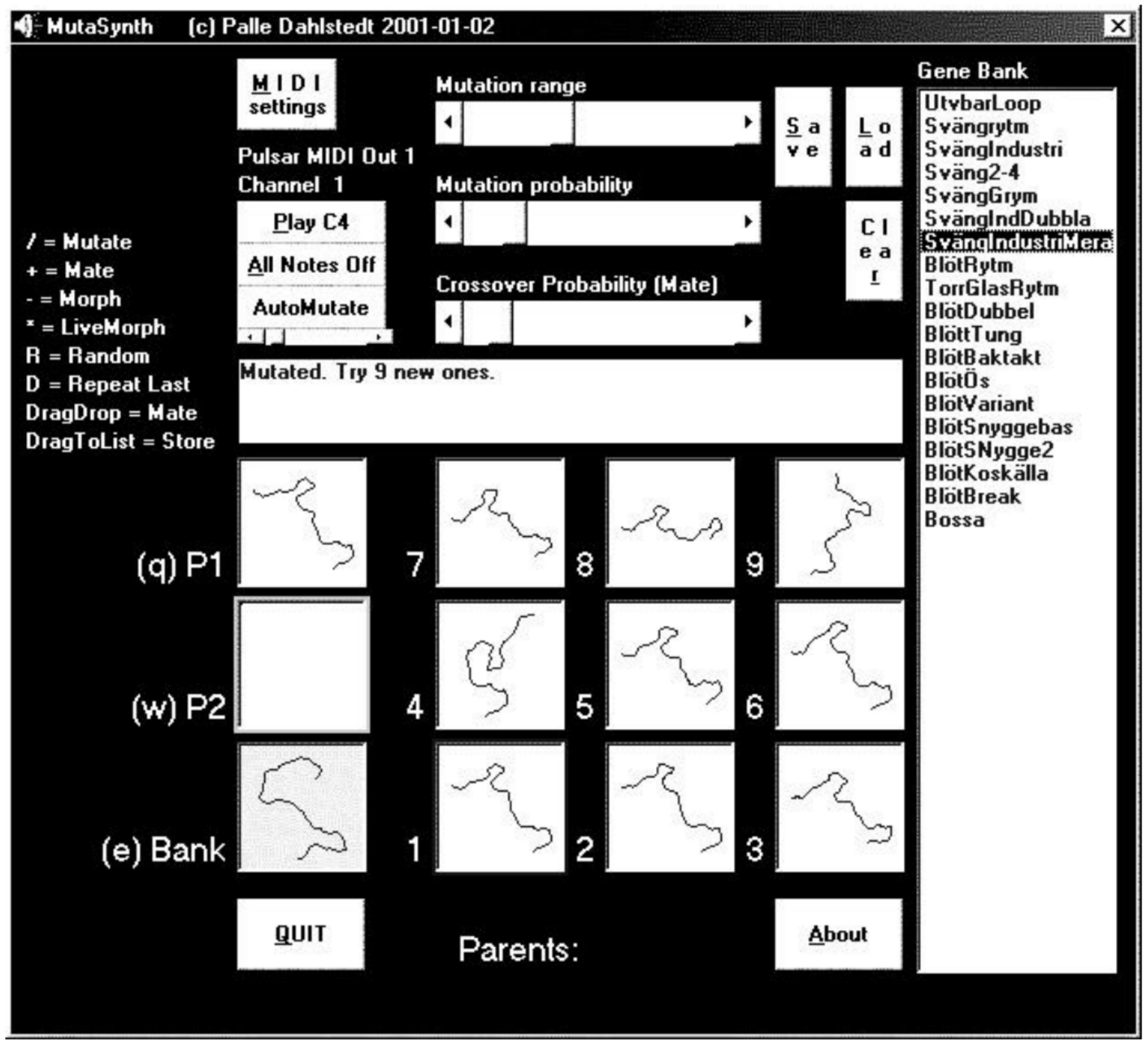
- **So, what if I could collaborate with an AI to explore mappings I wouldn't normally think of?**

Why? (cont.)

- Supporting exploration of complex spaces (such as sound synthesis)
- In artistic practice: able to surprise and help breaking creative deadlocks
- Learning: understanding how things work through co-exploring
- Demystifying AI/ML by using it <u>actively</u> to build something

Exploring Complexity through Emergent Behaviours: Artificial Agents in Computer Music

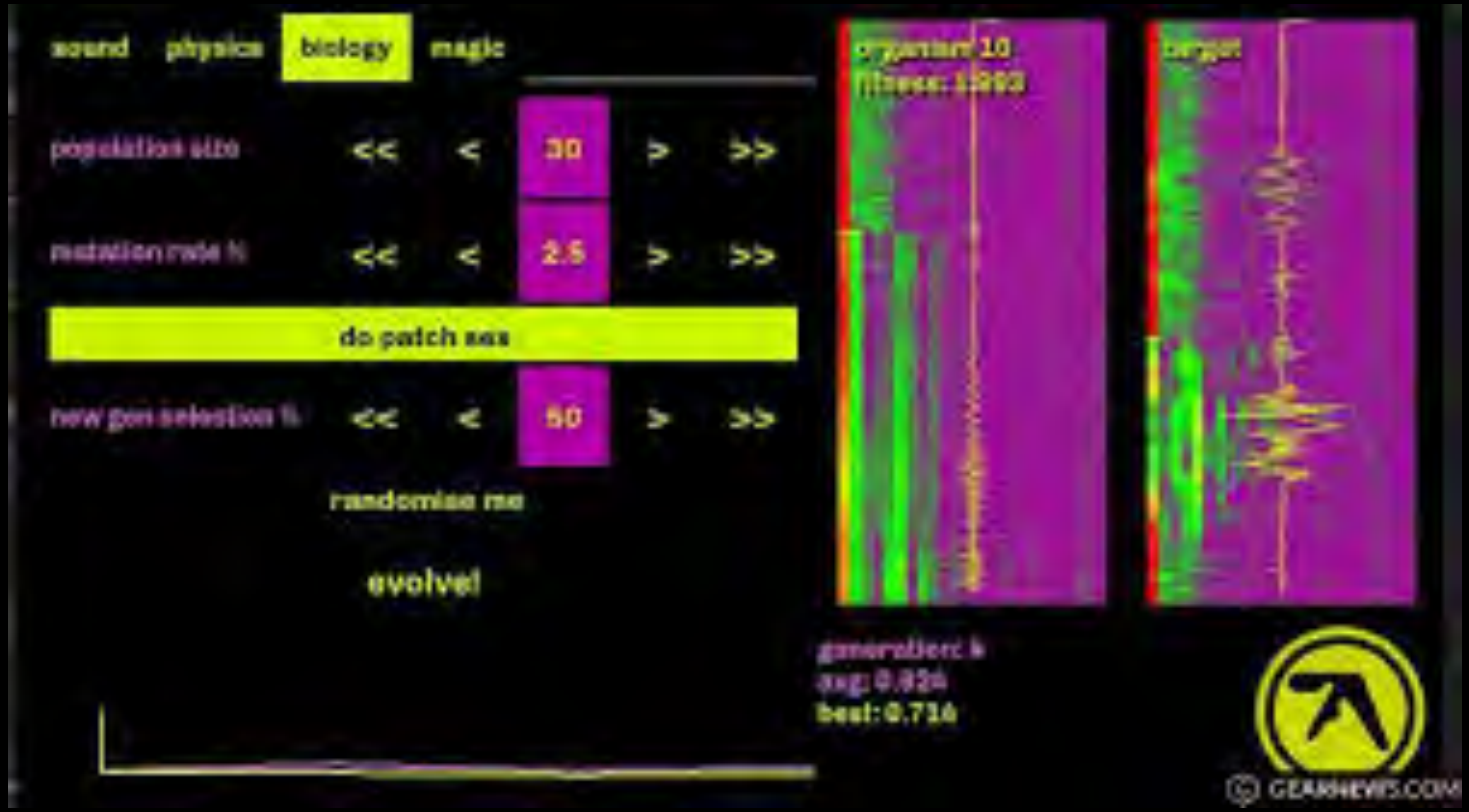




“Interactive evolution as a compositional tool makes it possible to create <u>surprisingly complex sounds</u> and structures in a very quick and simple way, while keeping a feeling of control.”

Dahlstedt, P. (2001). Creating and exploring huge parameter spaces: Interactive evolution as a tool for sound generation. In Proceedings of the 2001 International Computer Music Conference.

Genetic algorithms for exploring complex sound parameter spaces



https://fo.am/activities/midimutant/

Review and more recent approaches (also deep networks!):
Yee-King, M. J., Fedden, L., & D'Inverno, M. (2018). Automatic Programming of VST Sound Synthesizers Using Deep Networks and Other Techniques. IEEE Transactions on Emerging Topics in Computational Intelligence, 2(2), 150–159.

Explore Together: Reinforcement Learning

Reinforcement Learning

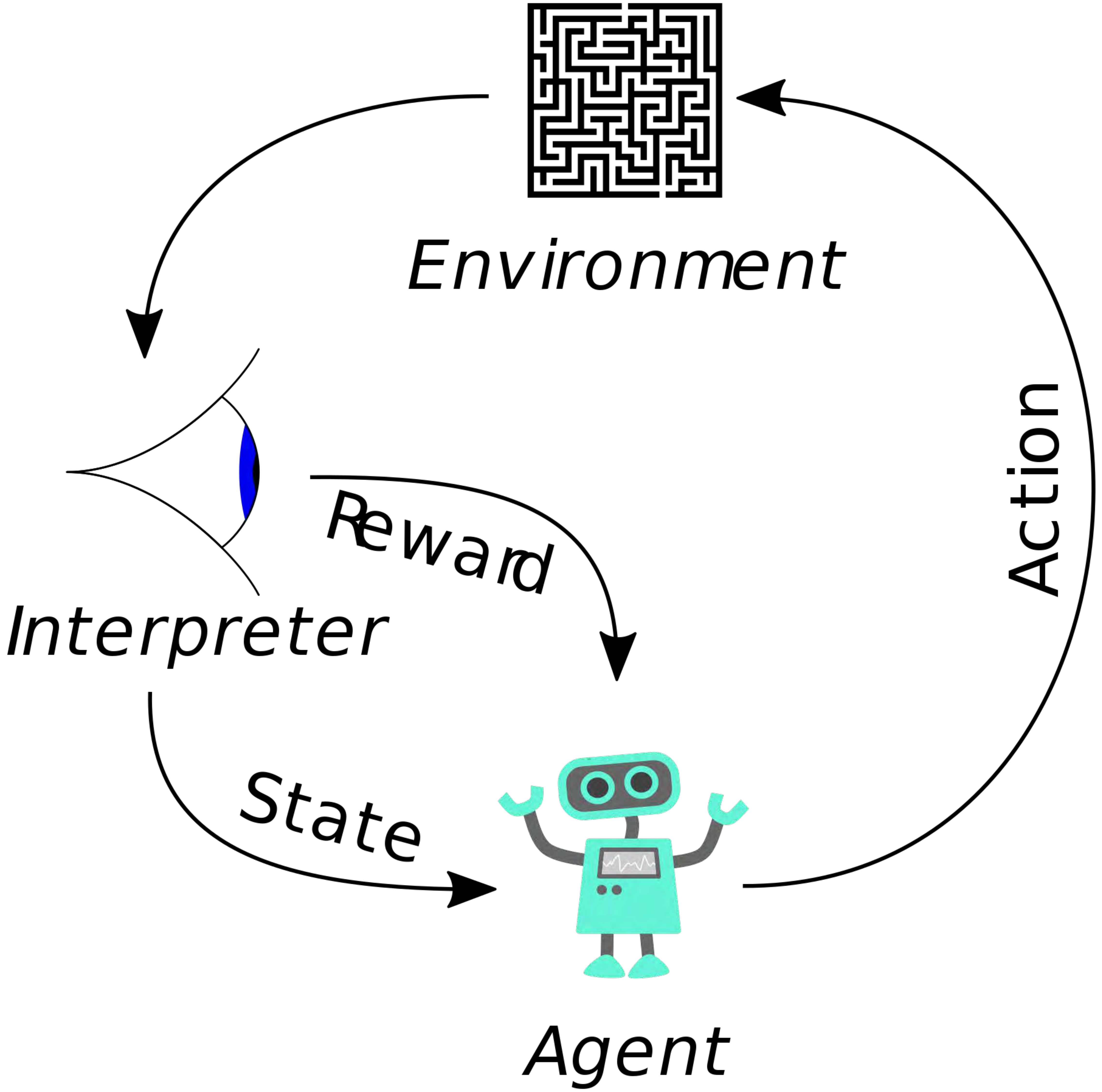
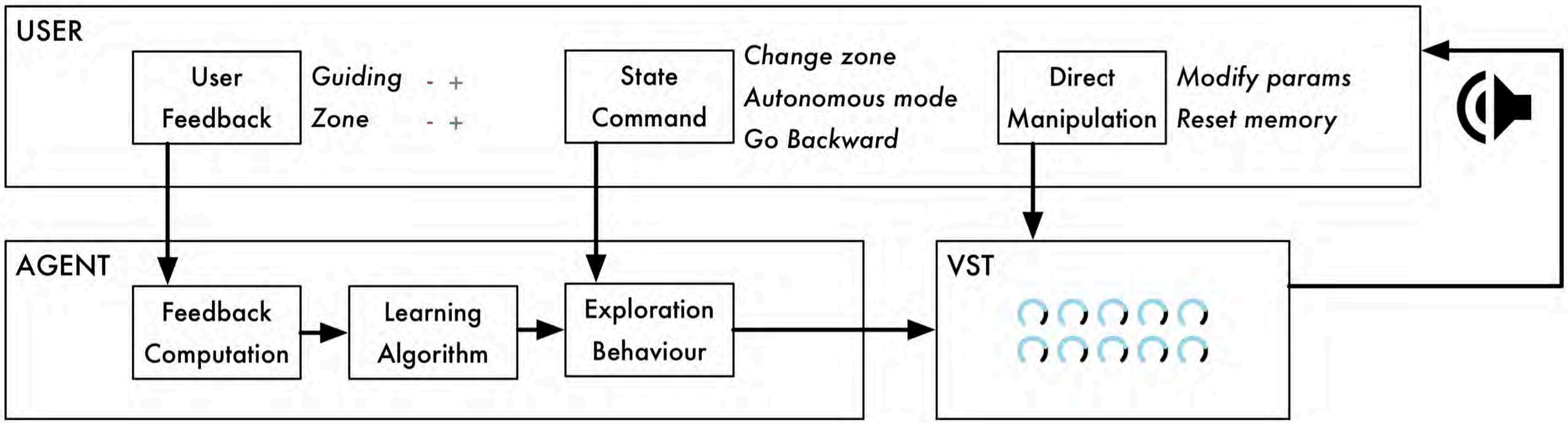


Image: Megajuice on Wikipedia

Reinforcement Learning

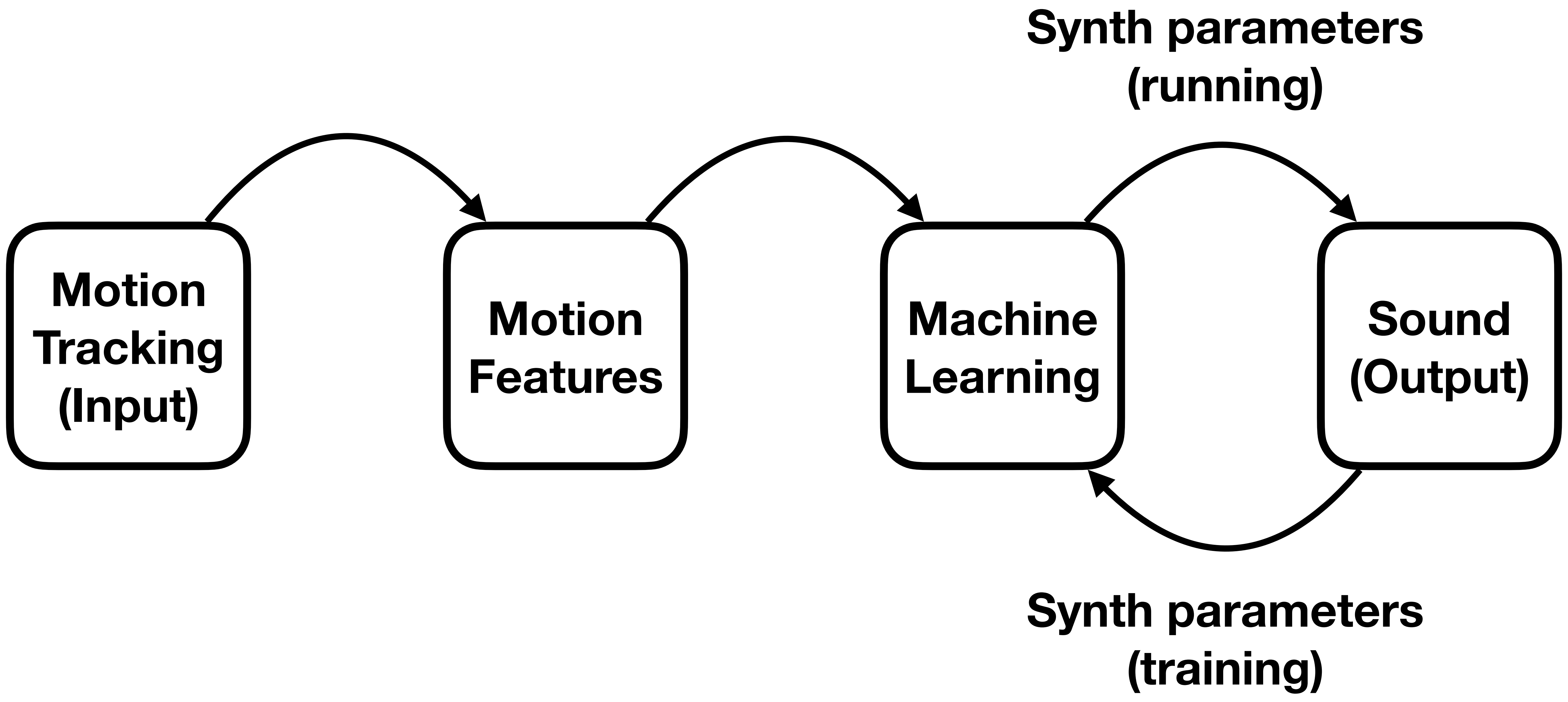
a sound design application



Scurto, H., Van Kerrebroeck, B., Caramiaux, B., & Bevilacqua, F. (2019). Designing Deep Reinforcement Learning for Human Parameter Exploration. ArXiv Preprint. Retrieved from https://arxiv.org/pdf/1907.00824.pdf

Interactive Machine Learning Model

What: control gestures



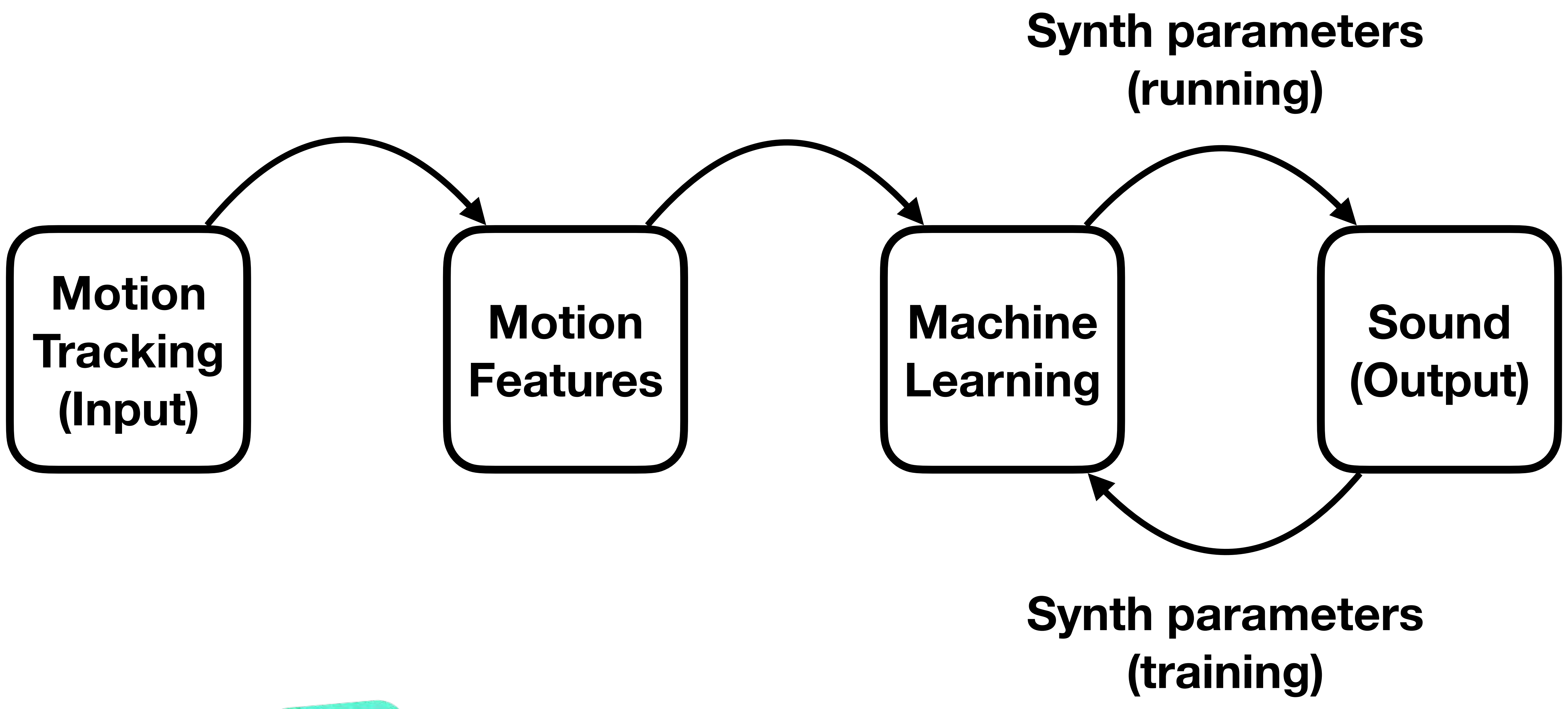
What: sounds

How these are related

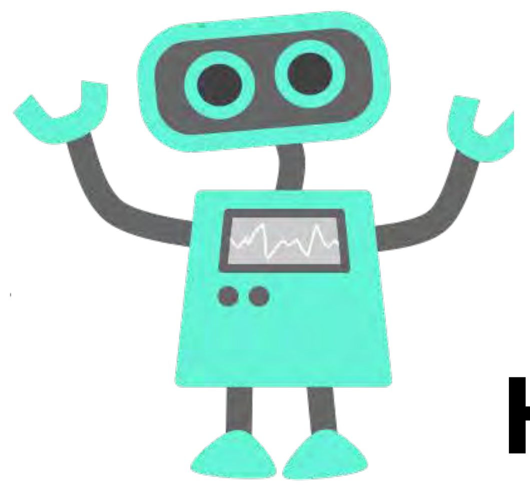
Examples

Assisted Interactive Machine Learning Model

What: control gestures



What: sounds



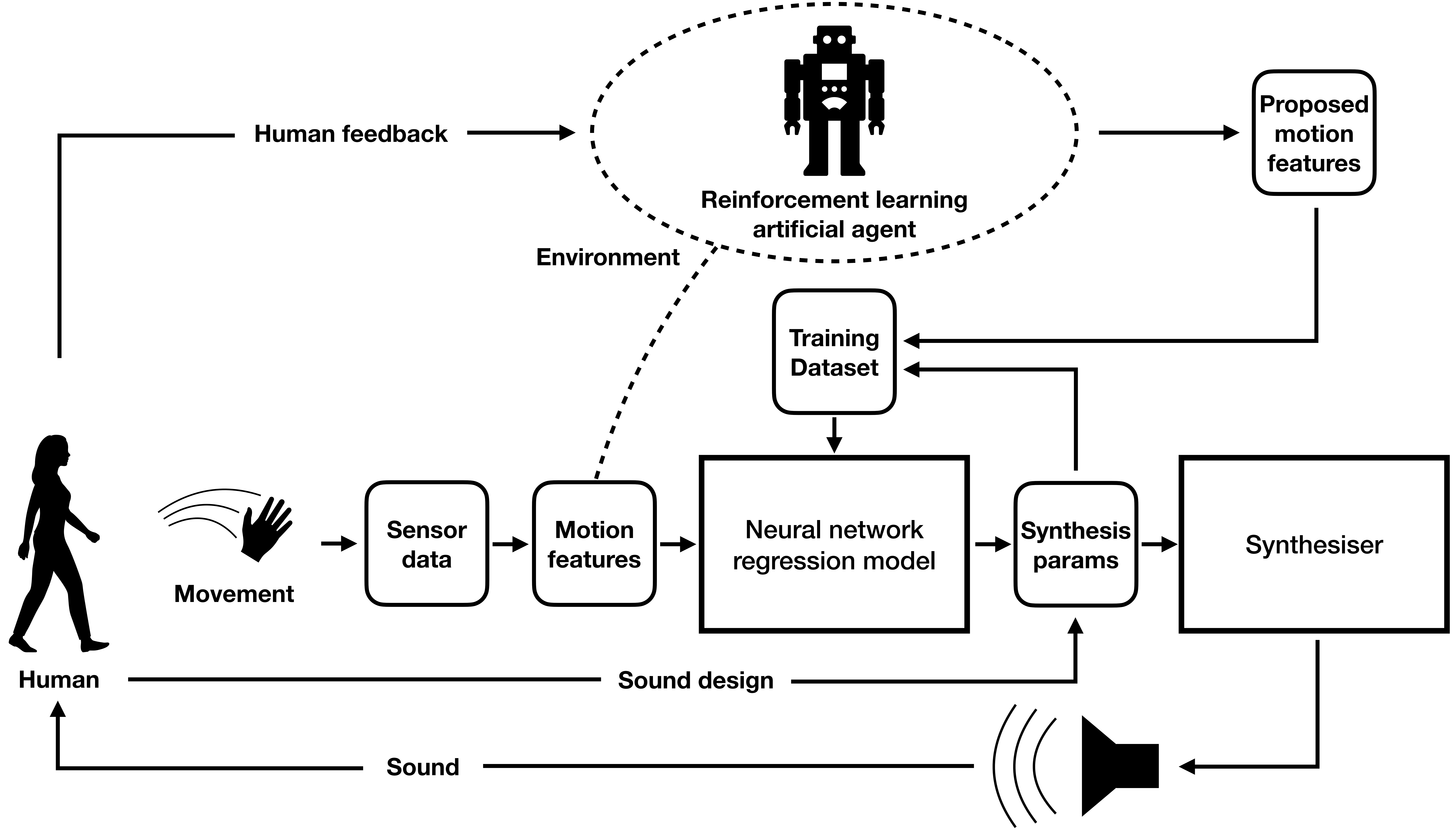
How these are related

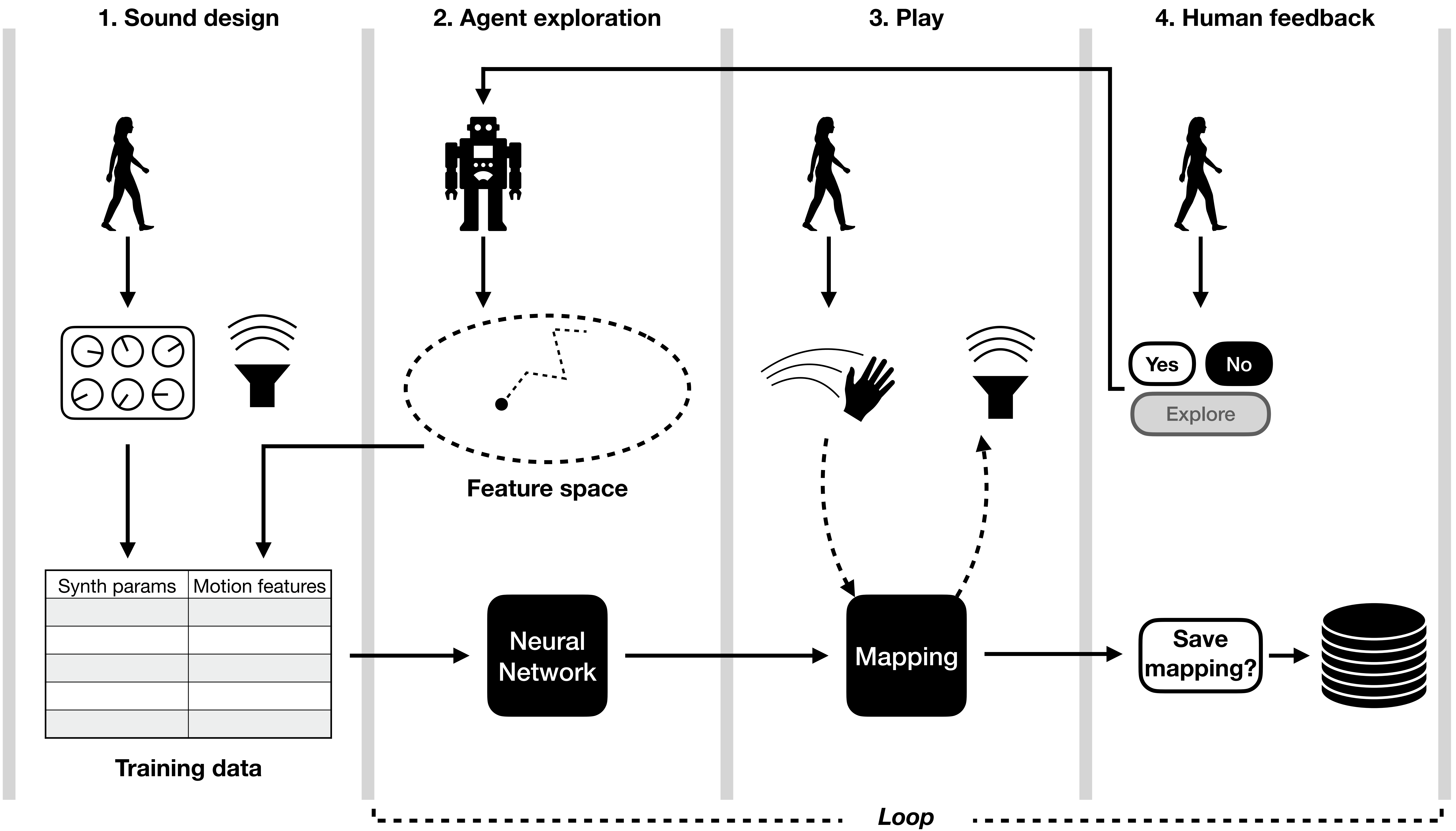
Agent

Interactive Machine Learning

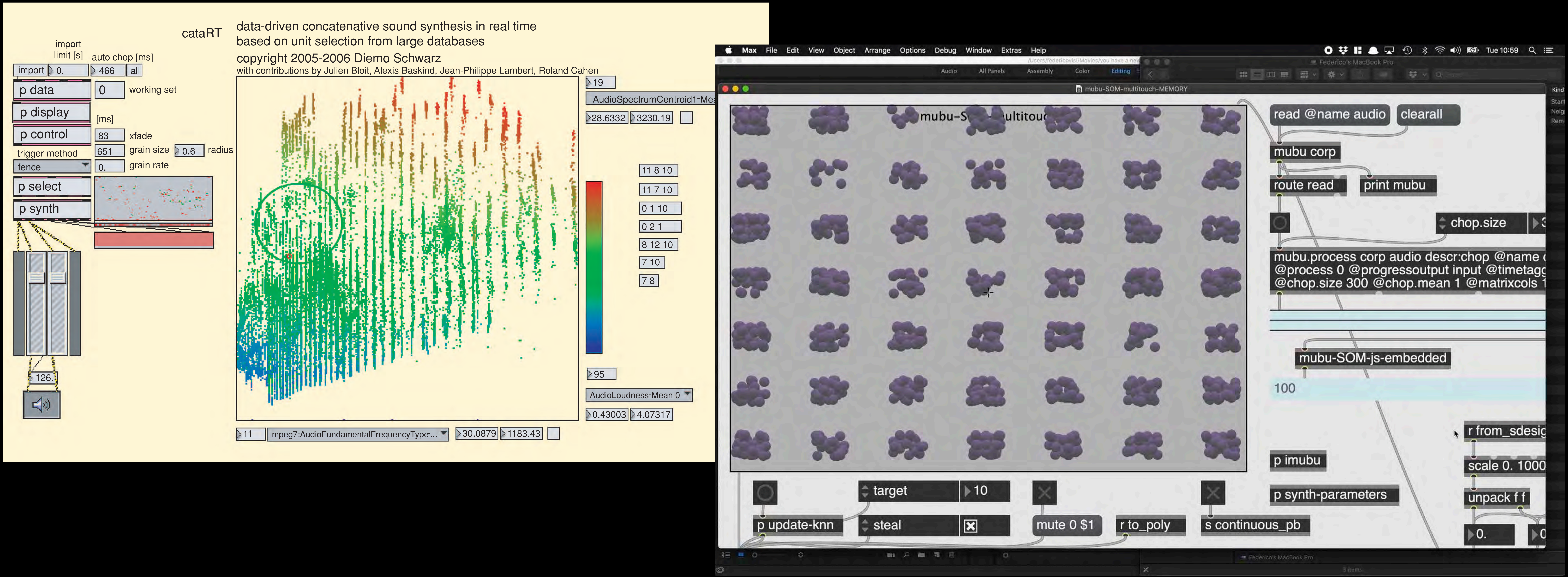
Artificial Agent Exploration

Assisted Interactive Machine Learning


Human feedback
Reinforcement learning
artificial agent
Proposed
motion
features
Environment
Training
Dataset
Movement
Sensor
data
Motion
features
Neural network
regression model
Synthesis
params
Synthesiser
Human
Sound design
Sound


1. Sound design
2. Agent exploration
3. Play
4. Human feedback
Yes
No
Explore
Feature space
Synth params
Motion features
Training data
Neural Network
Mapping
Save mapping?
Loop

Corpus-based concatenative synthesis



Schwarz, D. (2007). Corpus-based concatenative synthesis. IEEE signal processing magazine, 24(2), 92-104.

Margraf, J. (2019). Self-Organizing Maps for Sound Corpus Organization. Masters Thesis. Technische Universität Berlin.

Demo Corpus

Dataification of memories

My phone beeps. A notification on the home screen says "You have a new memory". It happens at times, unsupervised learning algorithms scan your photos and videos, look at their features and metadata, and then you get a nice slideshow of that trip to South America, or those shows you went to while you were in Hamburg or London.

I guess it could be about the way your recent past is lumped into categories, or how different moments are juxtaposed in unexpected ways, or just because you'd rather be the one in charge of deciding when "you have a new memory."

Using AIML to explore sonic memories

https://aqaxa.bandcamp.com/

AQAXA – sonic_memory_1

♡ 1 ···

Auction

This is a sonic memory used by AQAXA in the track 'Knee-jerk Drifter' part of 'Corporeal EP' released on 23 April 2021 on Punch Up Records.

Creator 13% royalties

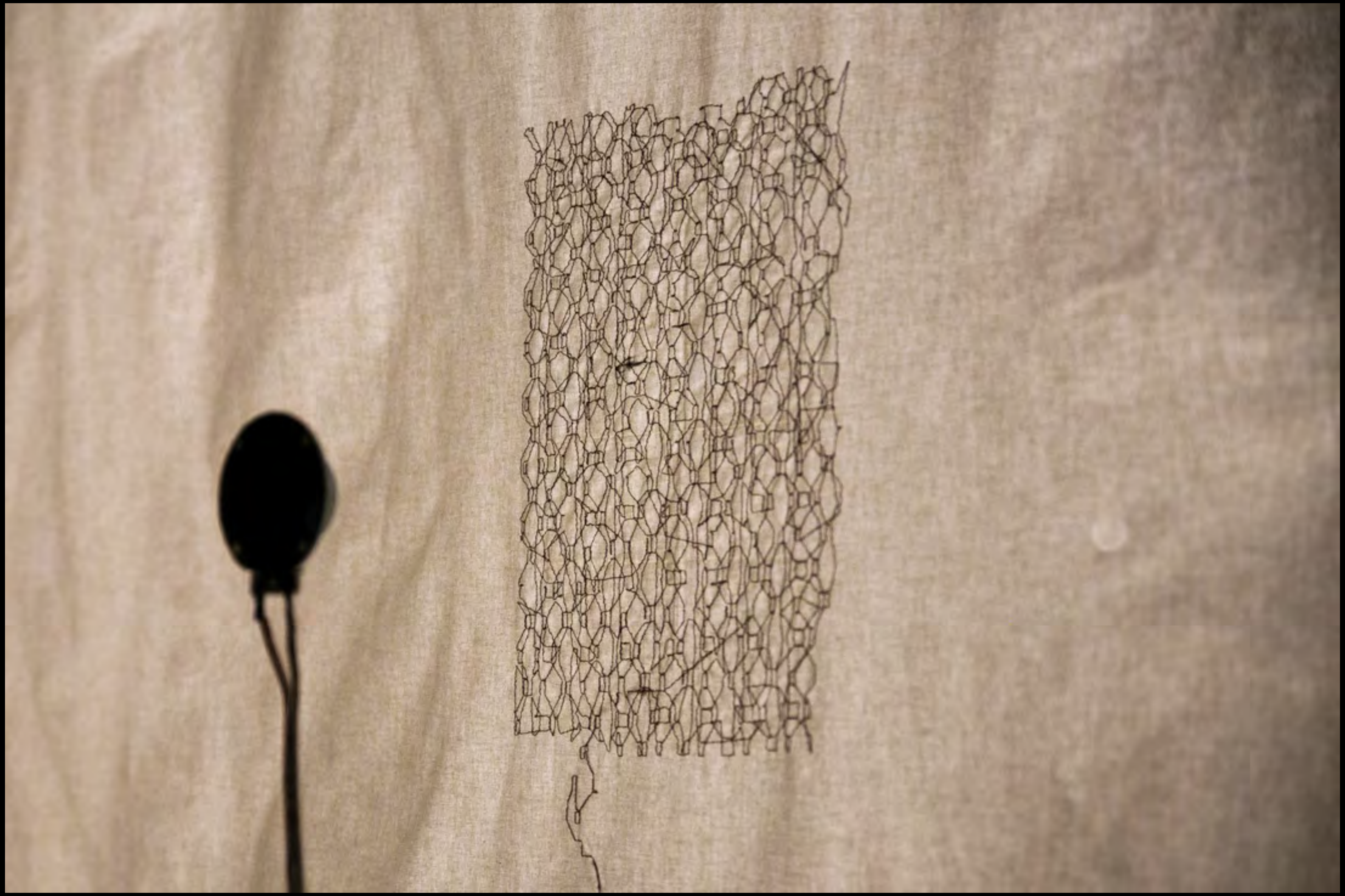
AQAXA

Collection

Rarible

https://aqaxa.bandcamp.com/


designtransfer



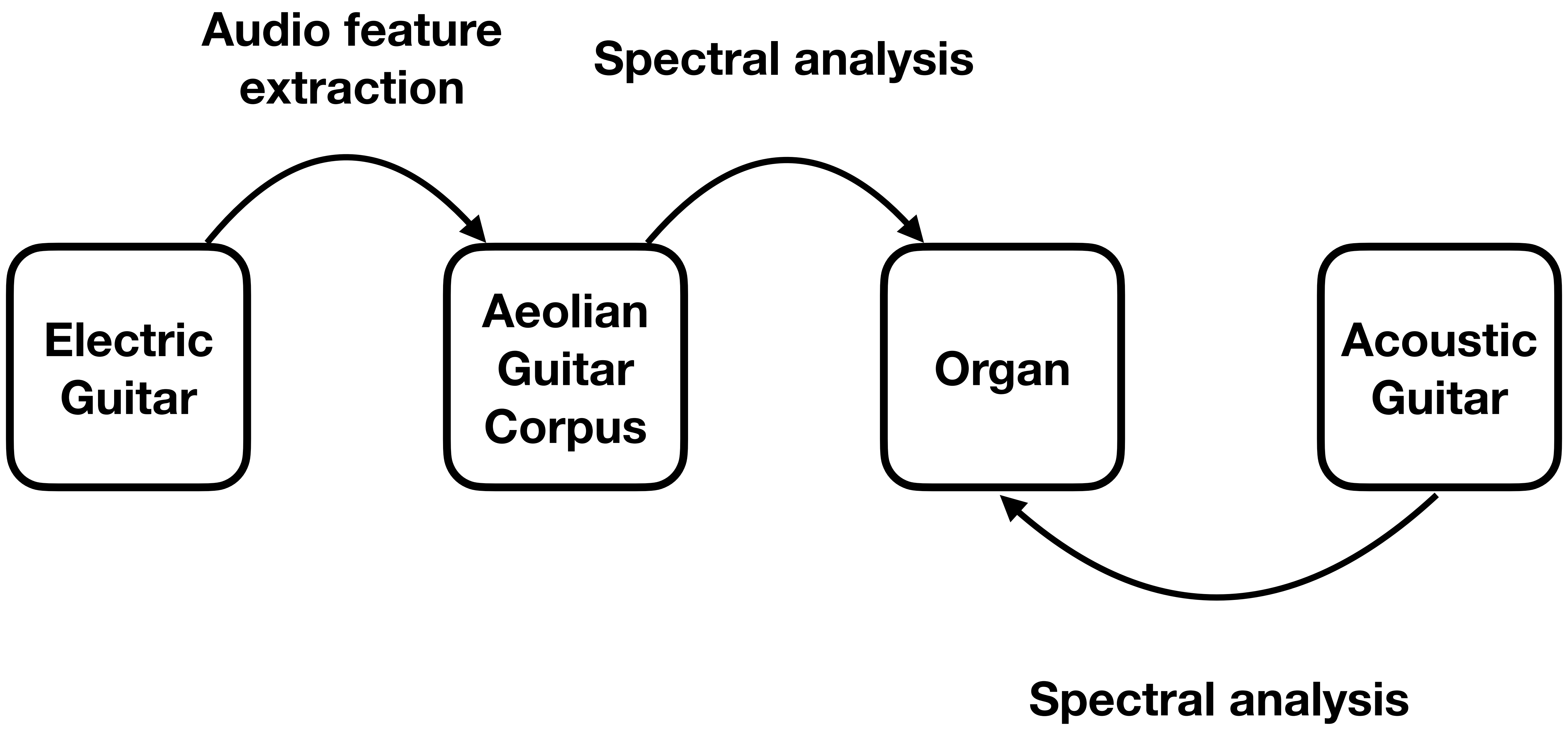
Piteå Archipelago, Sweden

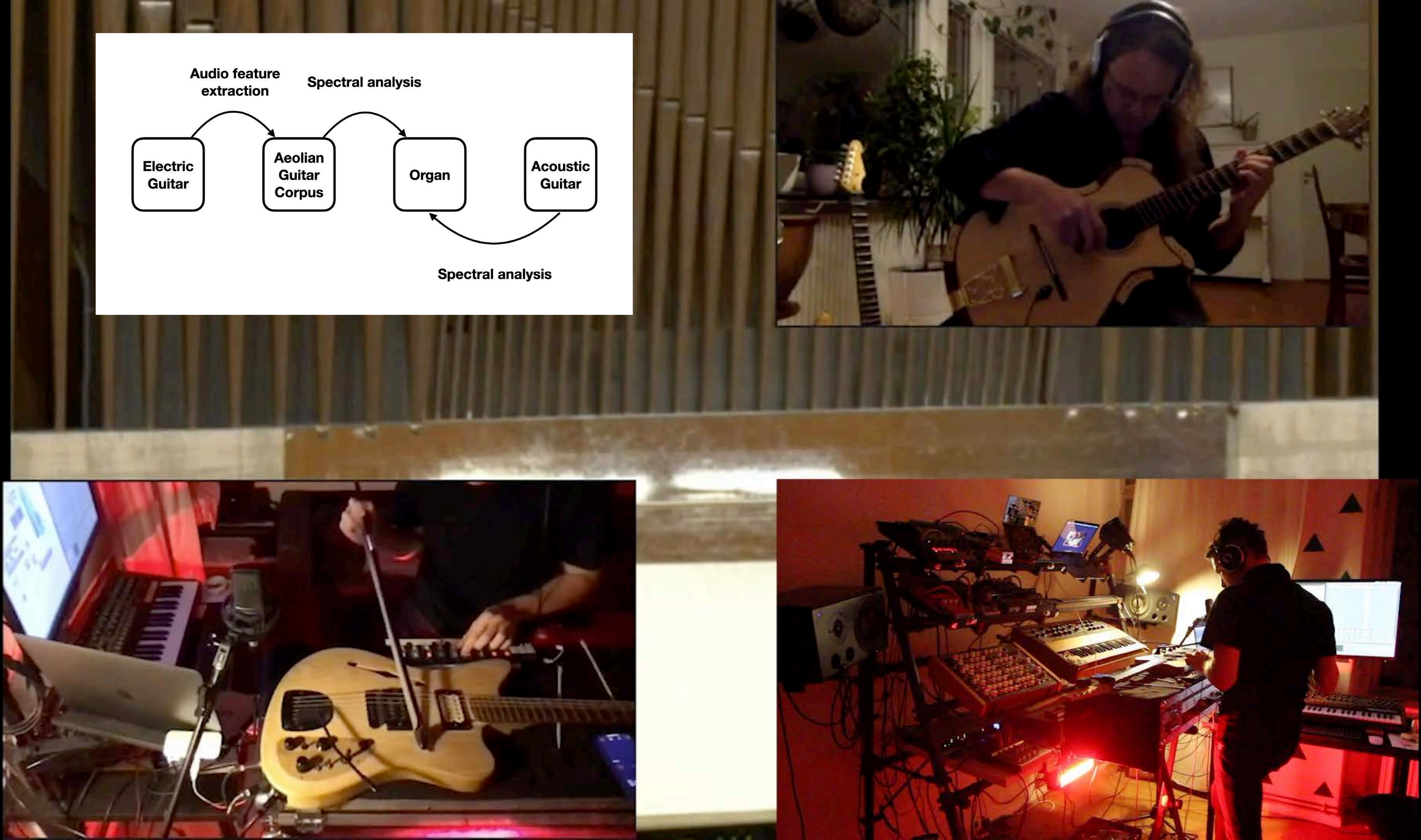
Electric Guitar

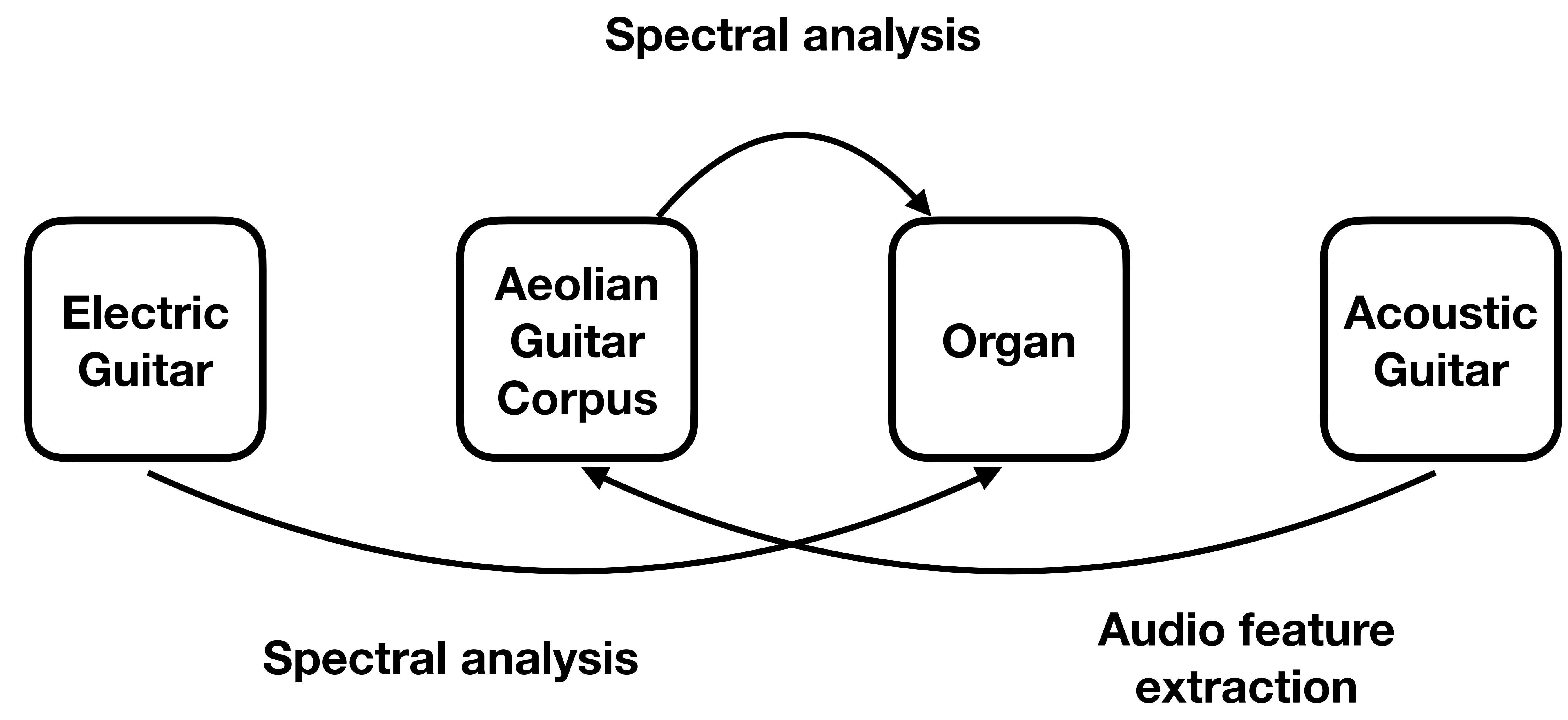
Aeolian Guitar Corpus

Organ

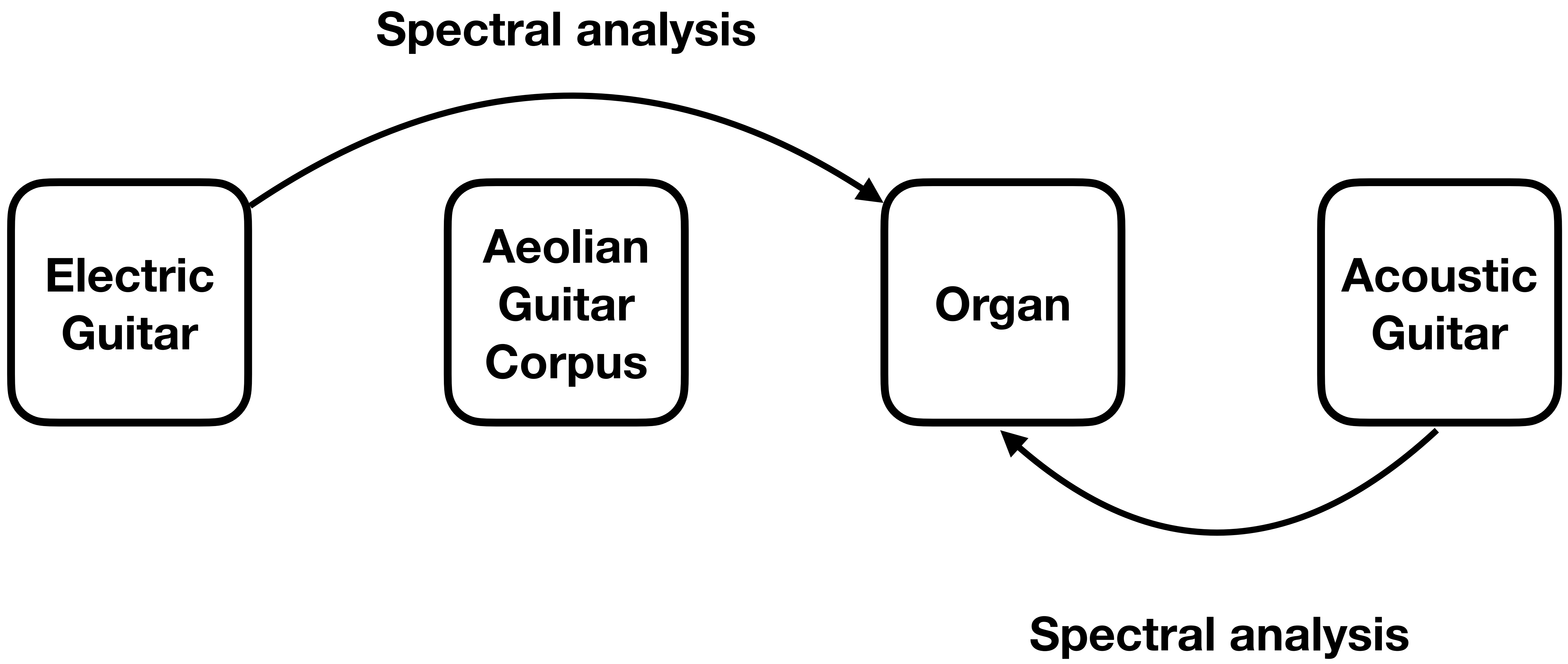
Acoustic Guitar


Audio feature extraction
Spectral analysis
Electric Guitar
Aeolian Guitar Corpus
Organ
Acoustic Guitar
Spectral analysis

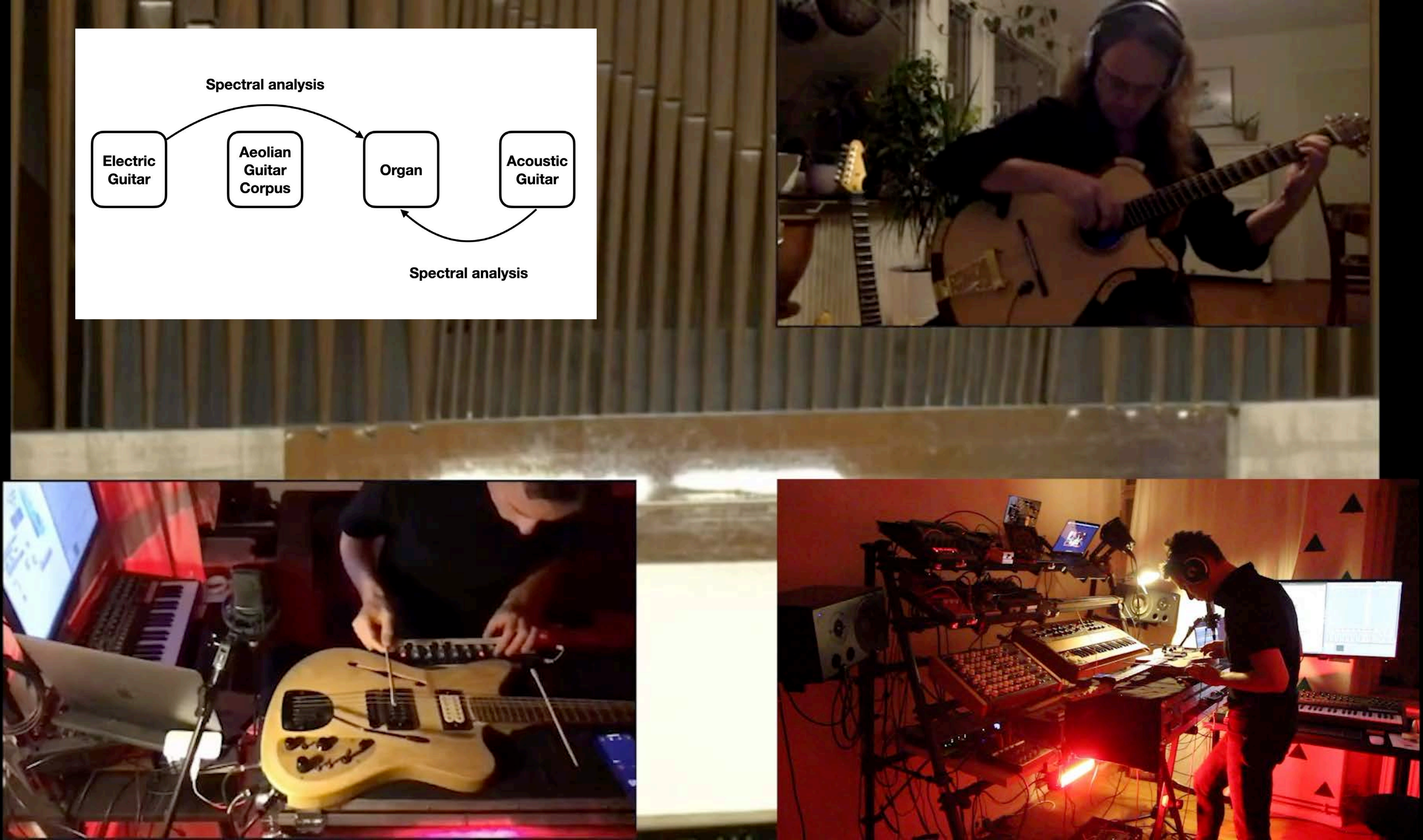

Audio feature extraction
Spectral analysis
Electric Guitar
Aeolian Guitar Corpus
Organ
Acoustic Guitar
Spectral analysis


Spectral analysis
Electric Guitar
Aeolian Guitar Corpus
Organ
Acoustic Guitar
Spectral analysis
Audio feature extraction


Spectral analysis
Electric Guitar
Aeolian Guitar Corpus
Organ
Acoustic Guitar
Spectral analysis
Audio feature extraction



Spectral analysis
Electric Guitar
Aeolian Guitar Corpus
Organ
Acoustic Guitar
Spectral analysis


Spectral analysis
Electric Guitar
Aeolian Guitar Corpus
Organ
Acoustic Guitar
Spectral analysis

TCP/Indeterminate Place quartet
live at NIME 2021
NYU Shanghai + Online

References

Visi, F. G., & Tanaka, A. (2021). **Interactive Machine Learning of Musical Gesture.** In E. R. Miranda (Ed.), *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*. Springer Nature.

Visi, F. G., & Tanaka, A. (2020). **Towards Assisted Interactive Machine Learning: Exploring Gesture-Sound Mappings Using Reinforcement Learning.** In *ICLI 2020 – the Fifth International Conference on Live Interfaces*.

Interactive Machine Learning of Musical Gesture

Federico Ghelli Visi* and Atau Tanaka **

GEMM))) Gesture Embodiment and Machines in Music
School of Music in Piteå, Luleå University of Technology, Sweden

EAVI – Embodied Audiovisual Interaction
Goldsmiths, University of London, UK

mail@federicovisi.com,
a.tanaka@gold.ac.uk

Abstract. This chapter presents an overview of interactive machine learning techniques applied to
tures. We go through the main c
analysing, and applying machi
gestures with the purpose of p
We discuss how different algorit
tasks, including interacting wit
ploring interaction possibilities
conclude the chapter with a des
were employed by the authors fo
thus outlining the implications t
musical practice.

Keywords: machine learning,
analysis, expressive movement,
reinforcement learning, music.

Proceedings of the 5th International Conference on Live Interfaces, NTNU, Trondheim, Norway 2020

Towards Assisted Interactive Machine Learning: Exploring Gesture-Sound Mappings Using Reinforcement Learning

Federico Ghelli Visi[1, 2, *] and Atau Tanaka[2, **]

[1]GEMM))) Gesture Embodiment and Machines in Music
School of Music in Piteå, Luleå University of Technology, Sweden

[2]EAVI – Embodied Audiovisual Interaction
Goldsmiths, University of London, UK

*mail@federicovisi.com
**a.tanaka@gold.ac.uk

Abstract.

We present a sonic interaction design approach that makes use of deep reinforcement learning to explore many mapping possibilities between input sensor data streams and sound synthesis parameters. The user can give feedback to an artificial agent about the mappings proposed by the latter while playing the synthesiser and trying the new mappings on the fly. The design approach we adopted is inspired by the ideas established by the interactive machine learning paradigm, as well as by the use of artificial agents in computer music for exploring complex parameter spaces. We refer to this interaction design approach as Assisted Interactive Machine Learning (AIML). We describe the architecture of an AIML system prototype, a typical workflow for interacting with the agent and obtain gesture-sound mappings. We then present feedback data collected during a demonstration and discuss perspectives for developing the AIML paradigm further, pointing out current limitations. In light of the feedback obtained and the considerations arisen following the use of the system in a multimedia performance piece, we suggest that the implementation and evaluation of new features should take into consideration established creative workflows and take place close to actual artistic practice.

Keywords. Gestural Interaction, Interactive Machine Learning, Reinforcement Learning, Artificial Agents, Sonic Interaction Design

Thank you.

Special thanks to my collaborators in these works:
Stefan Östersjö, Robert Ek, Mattias Petersson, Åsa Unander-Scharin
Atau Tanaka, Michael Zbyszyński, Jonas Margraf , Alex Murray-Leslie, Melissa Logan, Krõõt Juurak