Multimedia Music with Dance

Dance the Music

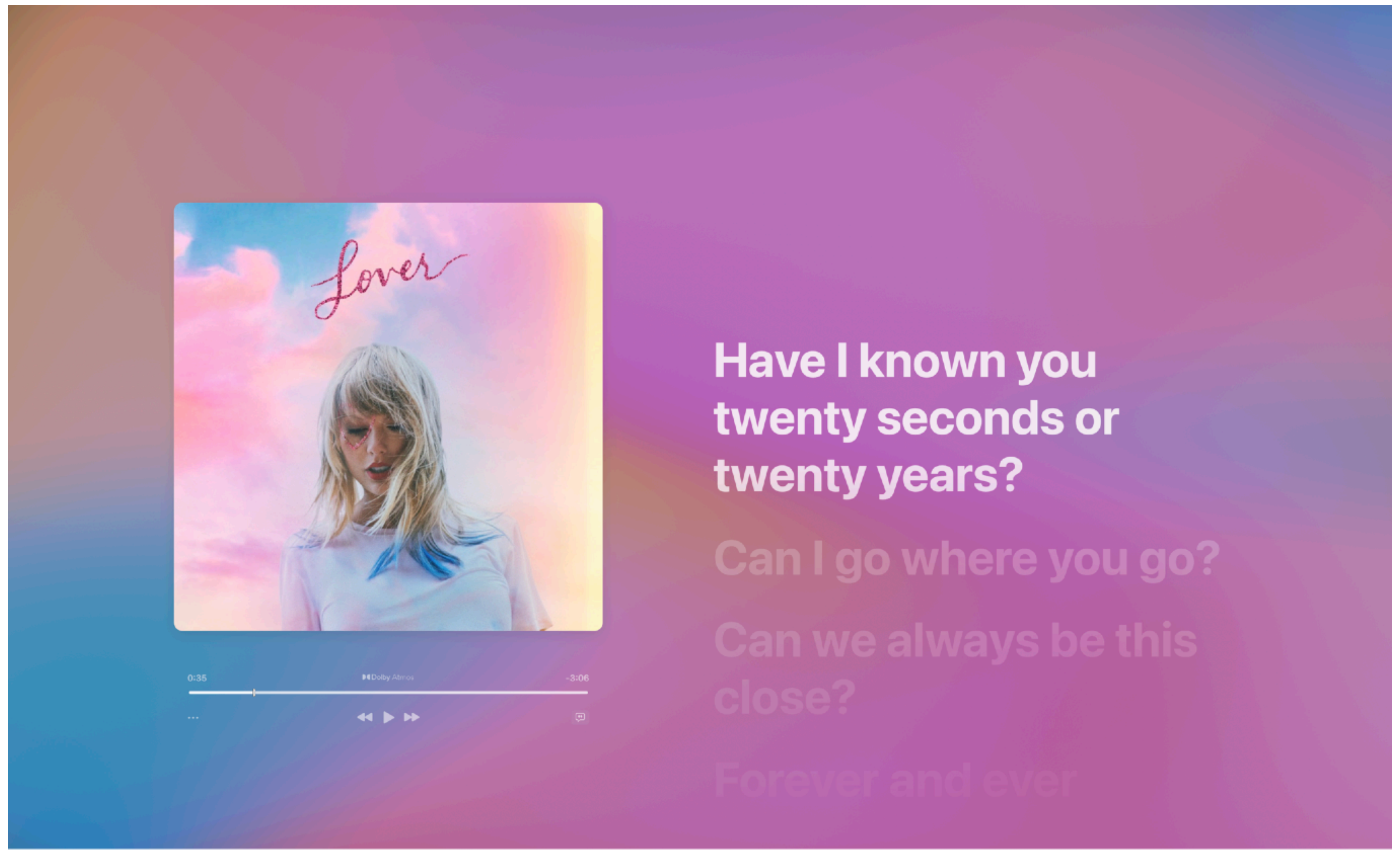
Gwangyu Lee

Gypsy Jazz

Tango


CROSLEY


A
N.R.
IN
OUT
TDK
High Bias 70µs EQ
SA
90


Lover
Have I known you
twenty seconds or
twenty years?
Can I go where you go?
Can we always be this
close?
Forever and ever

What if the dancers could "Dance the music" rather than "Dance to the music"?

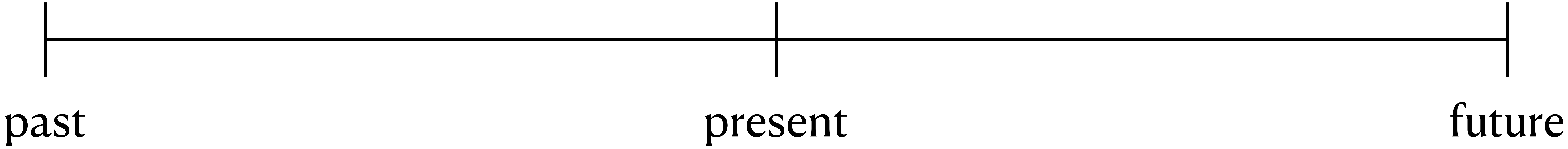
Interactive Dance

Tchaikovsky

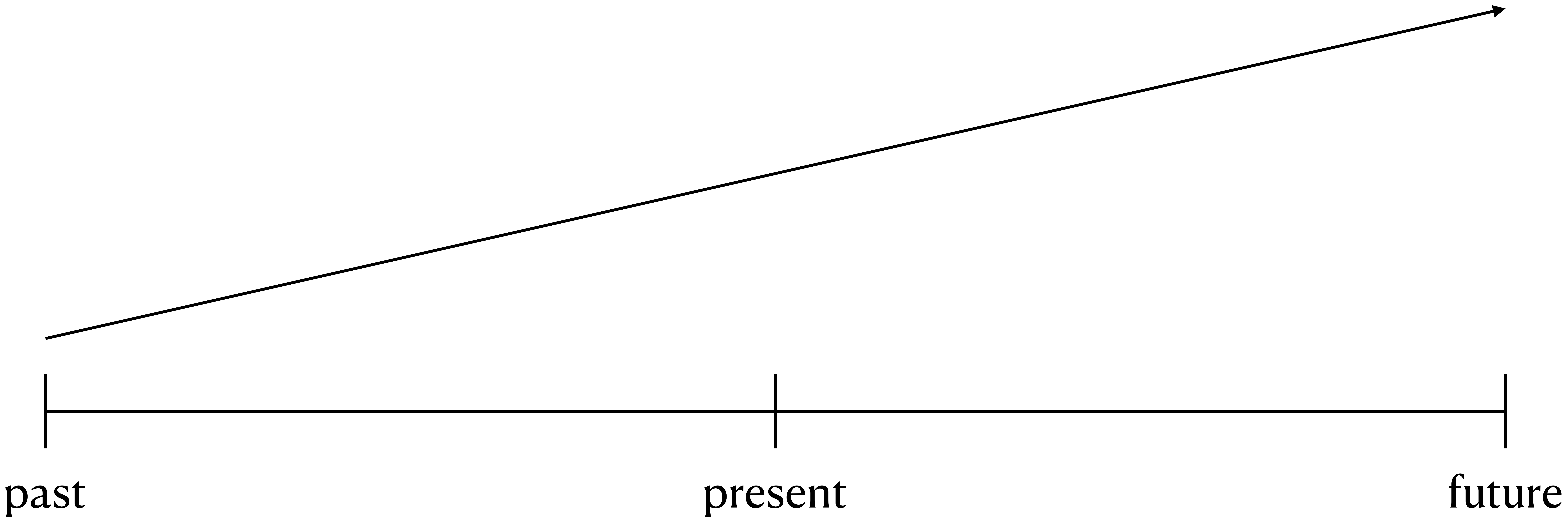
“Audio content and visual content in multimedia cannot be broken down into two independent elements”

Why Interactive Dance?

Linear-time

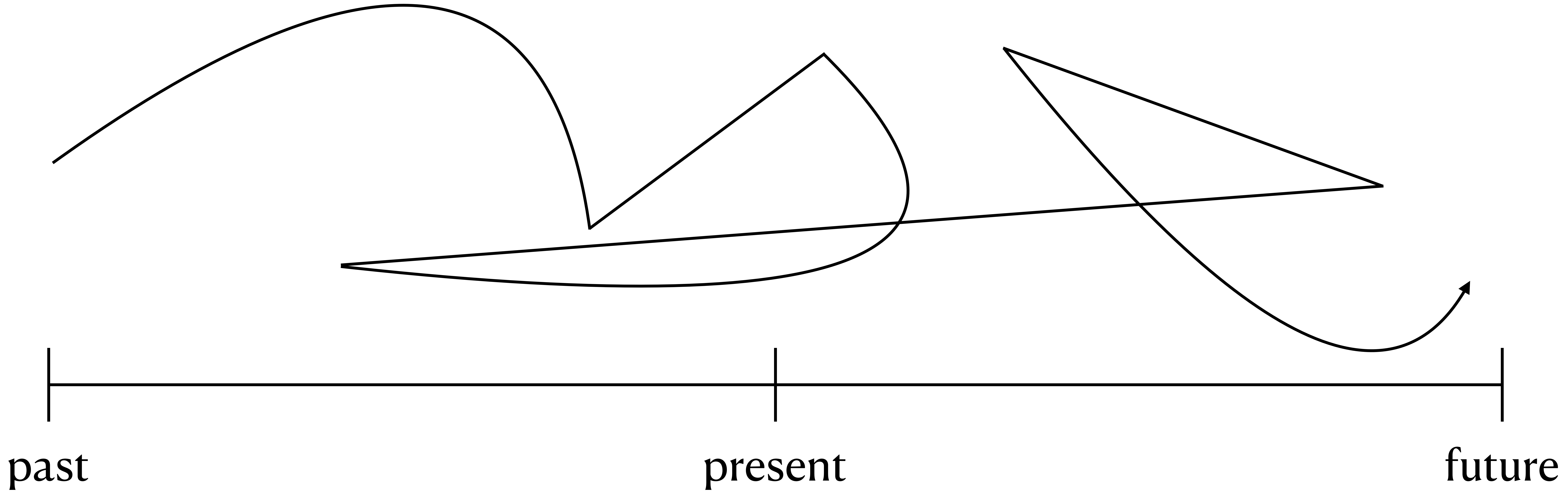


Linear-time



Real-time

Real-time



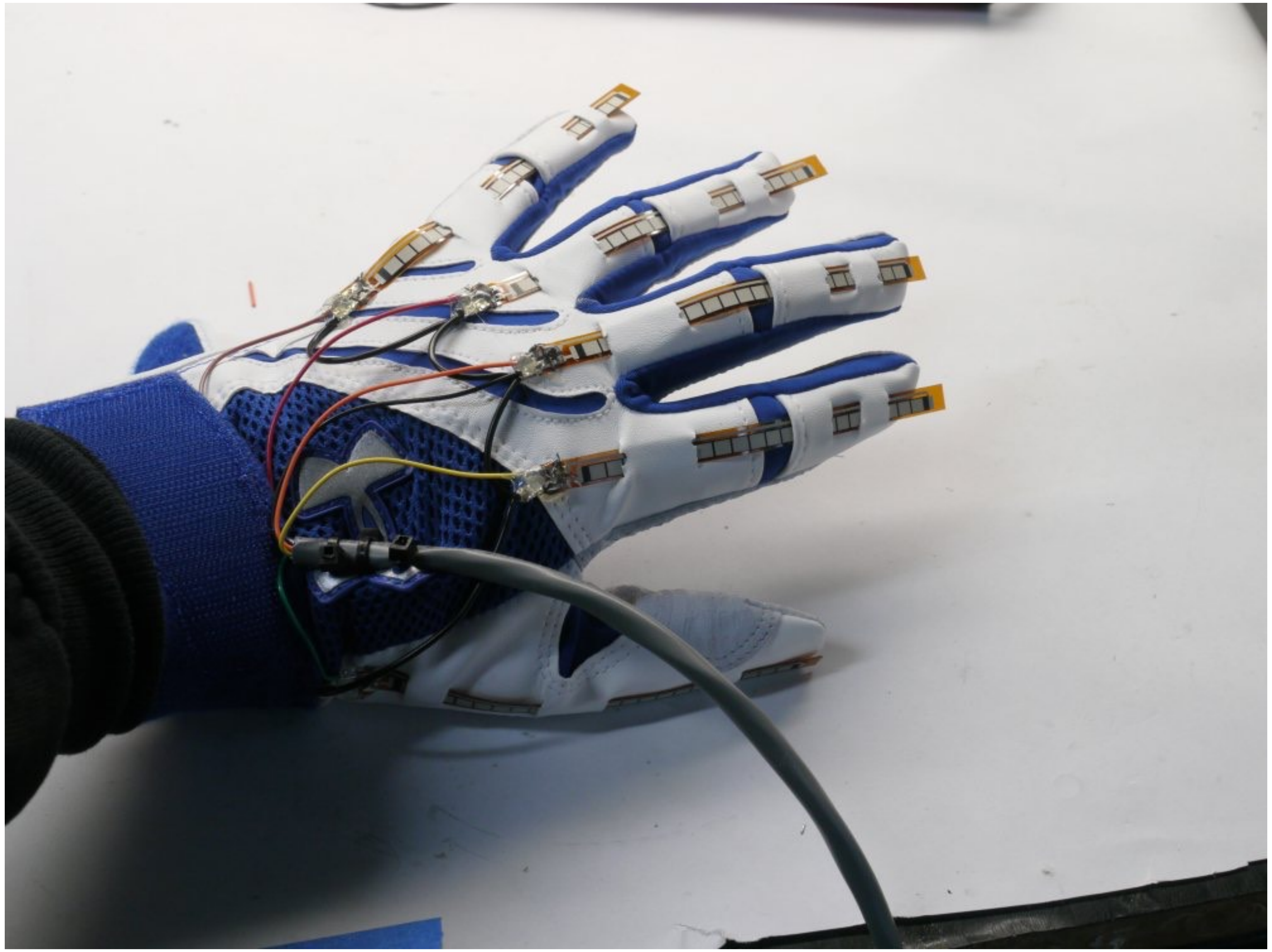
Motion-Tracking Technologies

Motion-Tracking Technologies

- Inside-In
- Inside-Out
- Outside-In
- Camera-based motion-tracking systems(Computer Vision)

Inside-In

Inside-In



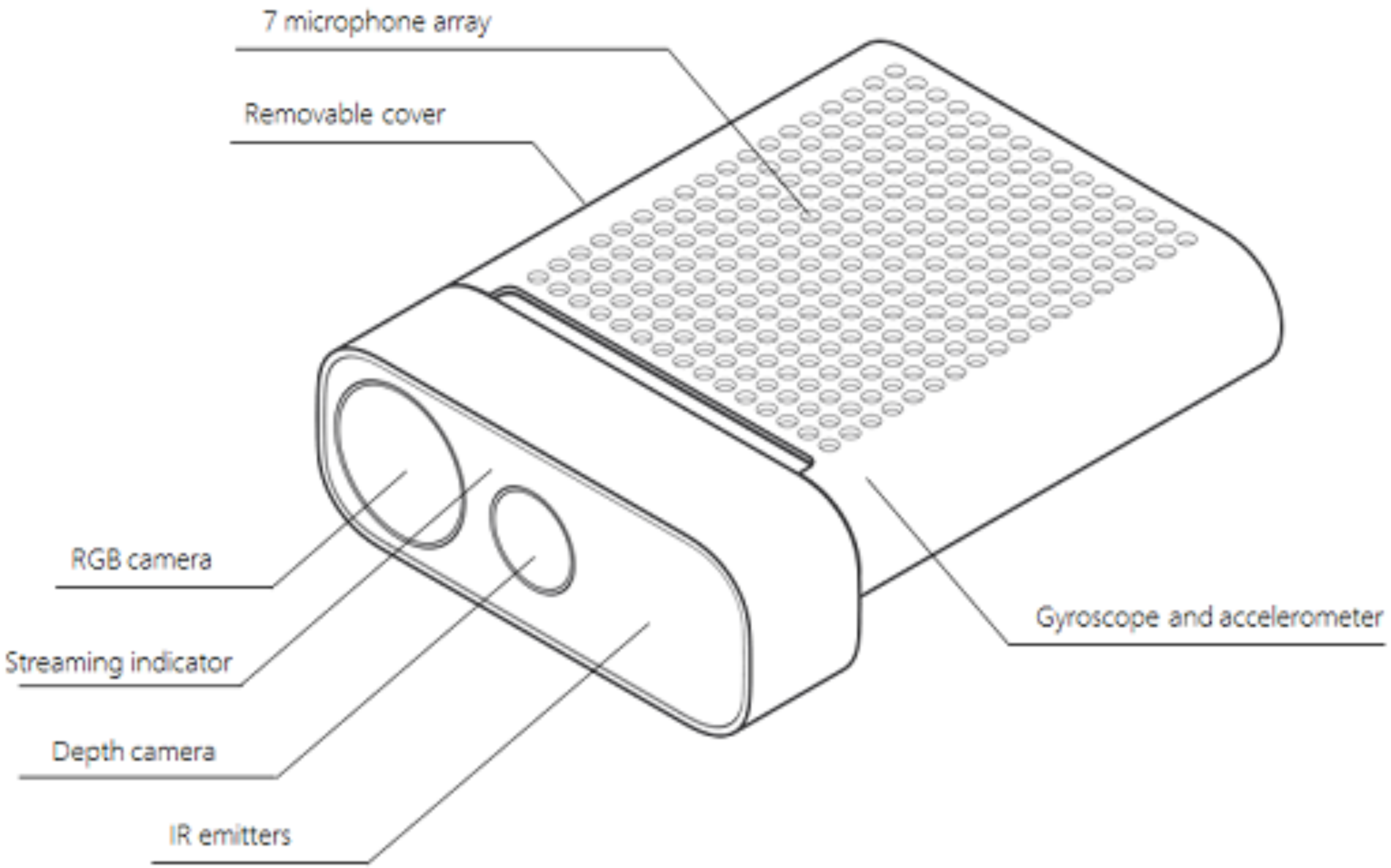
Inside-Out


https://youtu.be/S-T8kcSRLLo?feature=shared&t=419

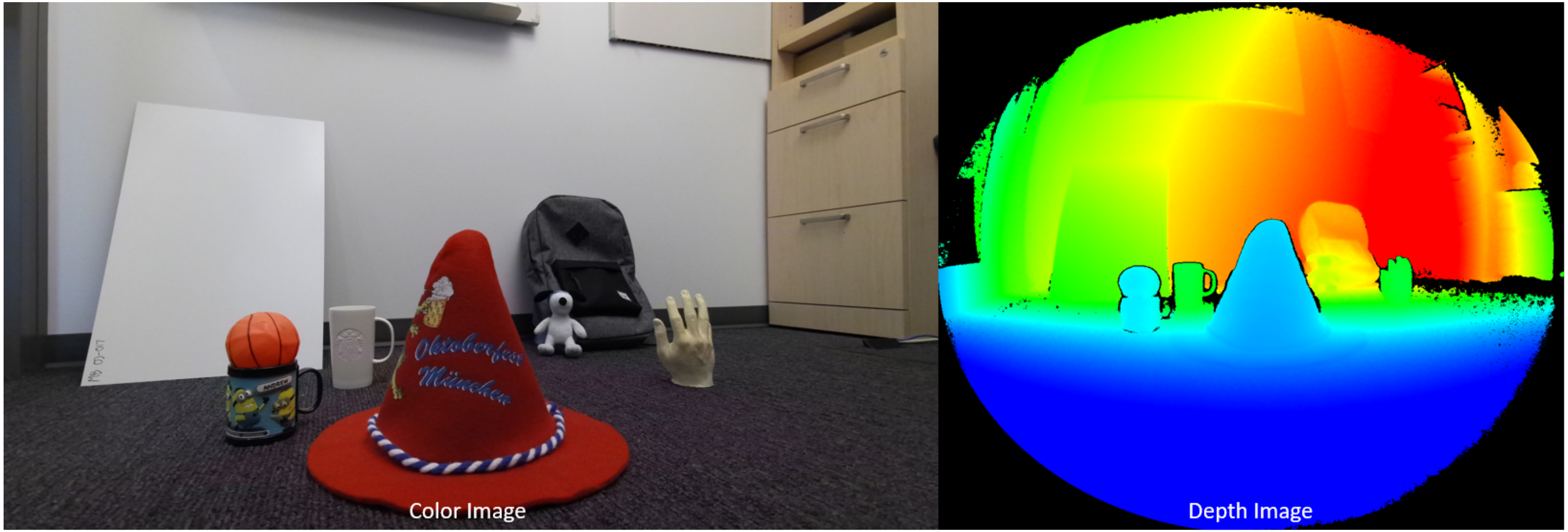
Outside-In

Outside-In

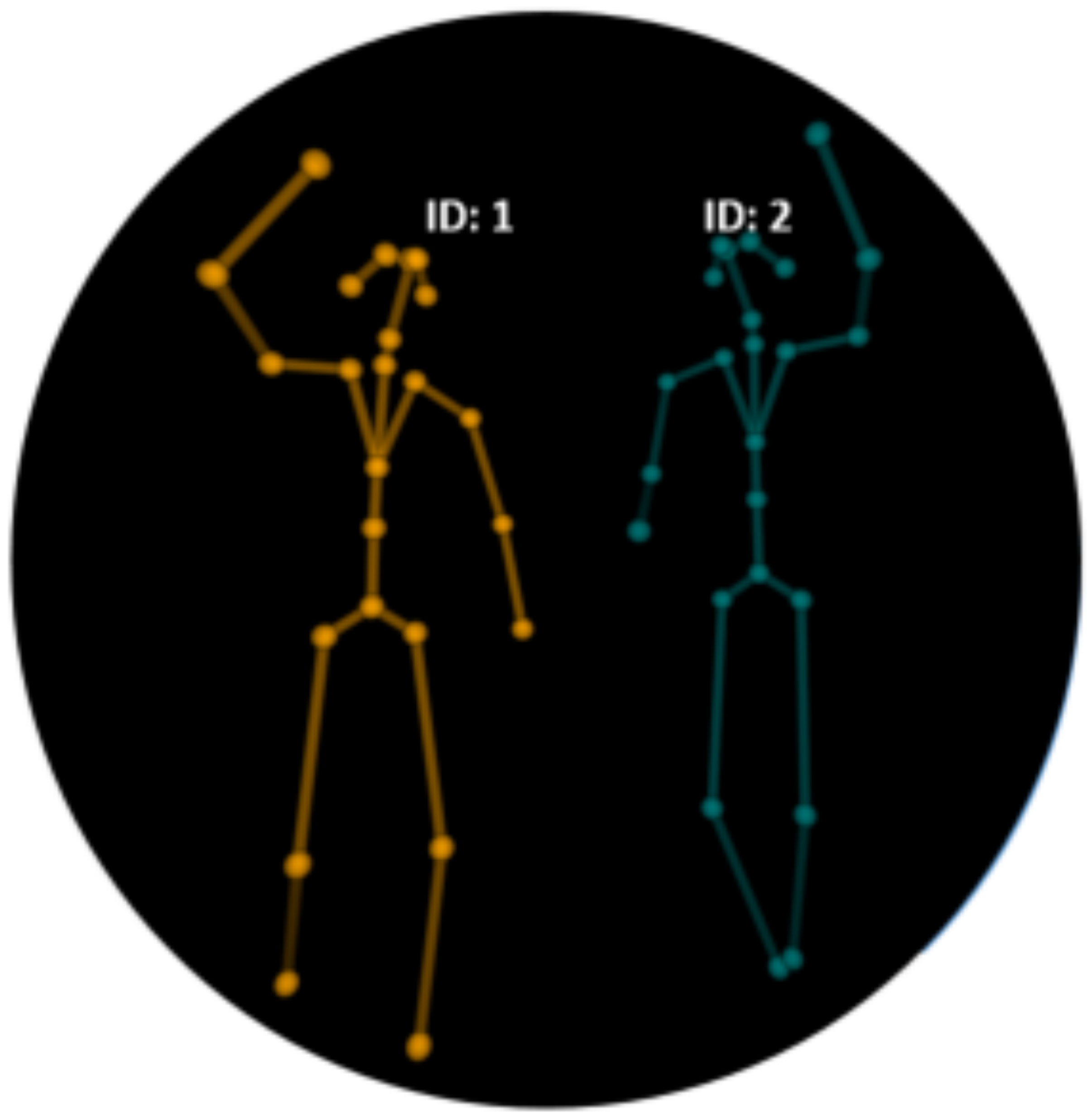
Azure Kinect DK

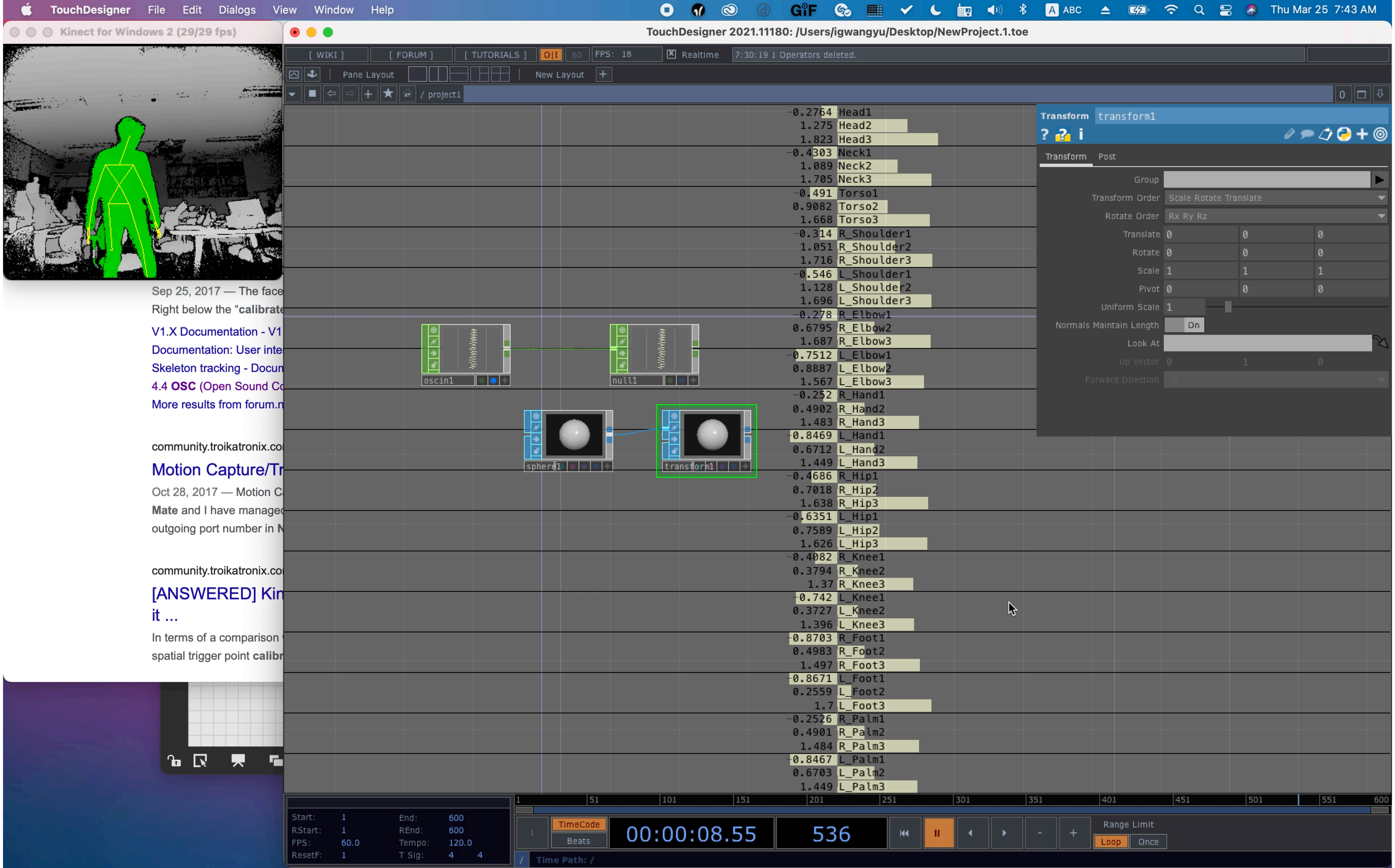


Azure Kinect DK



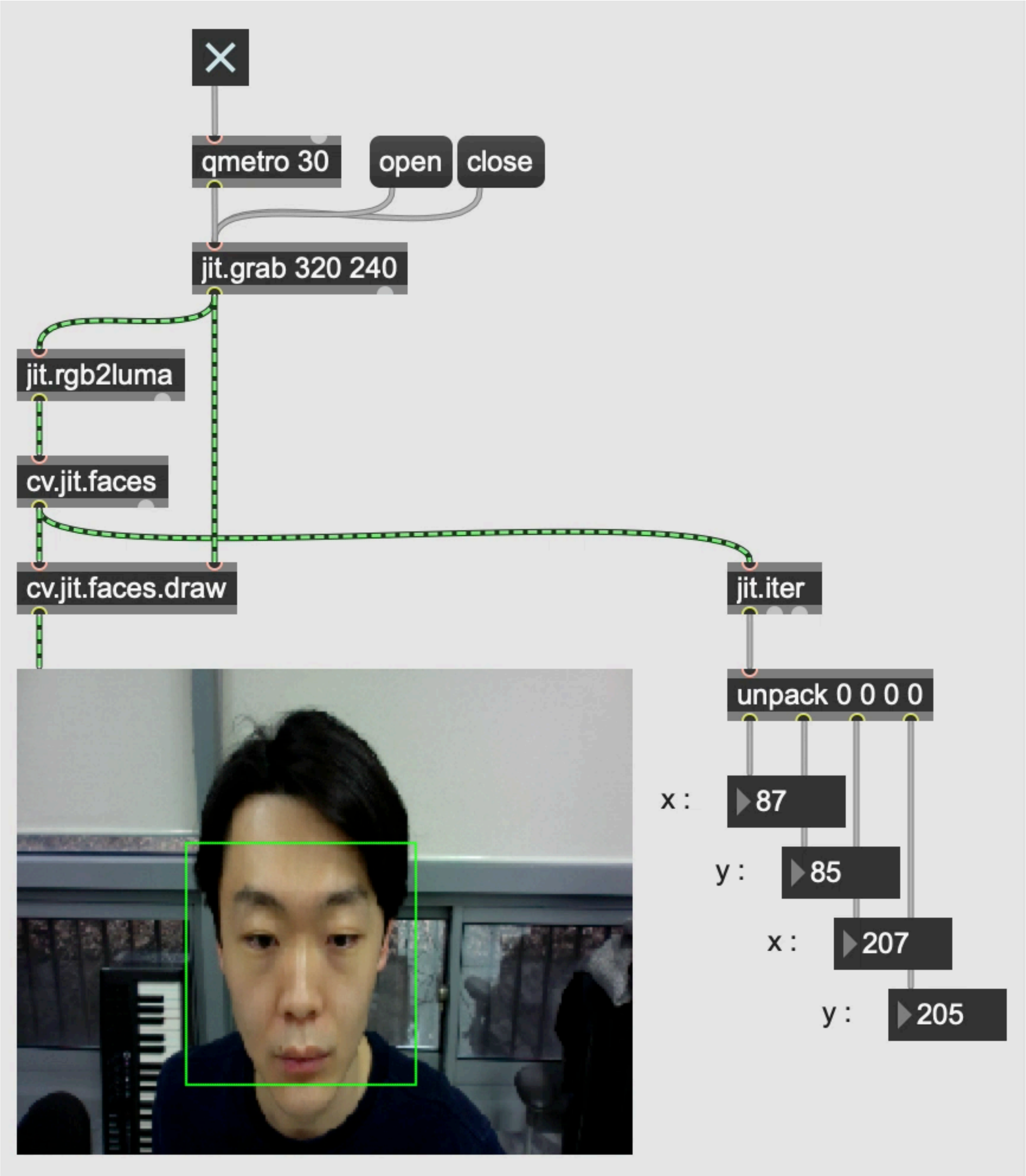
Azure Kinect DK



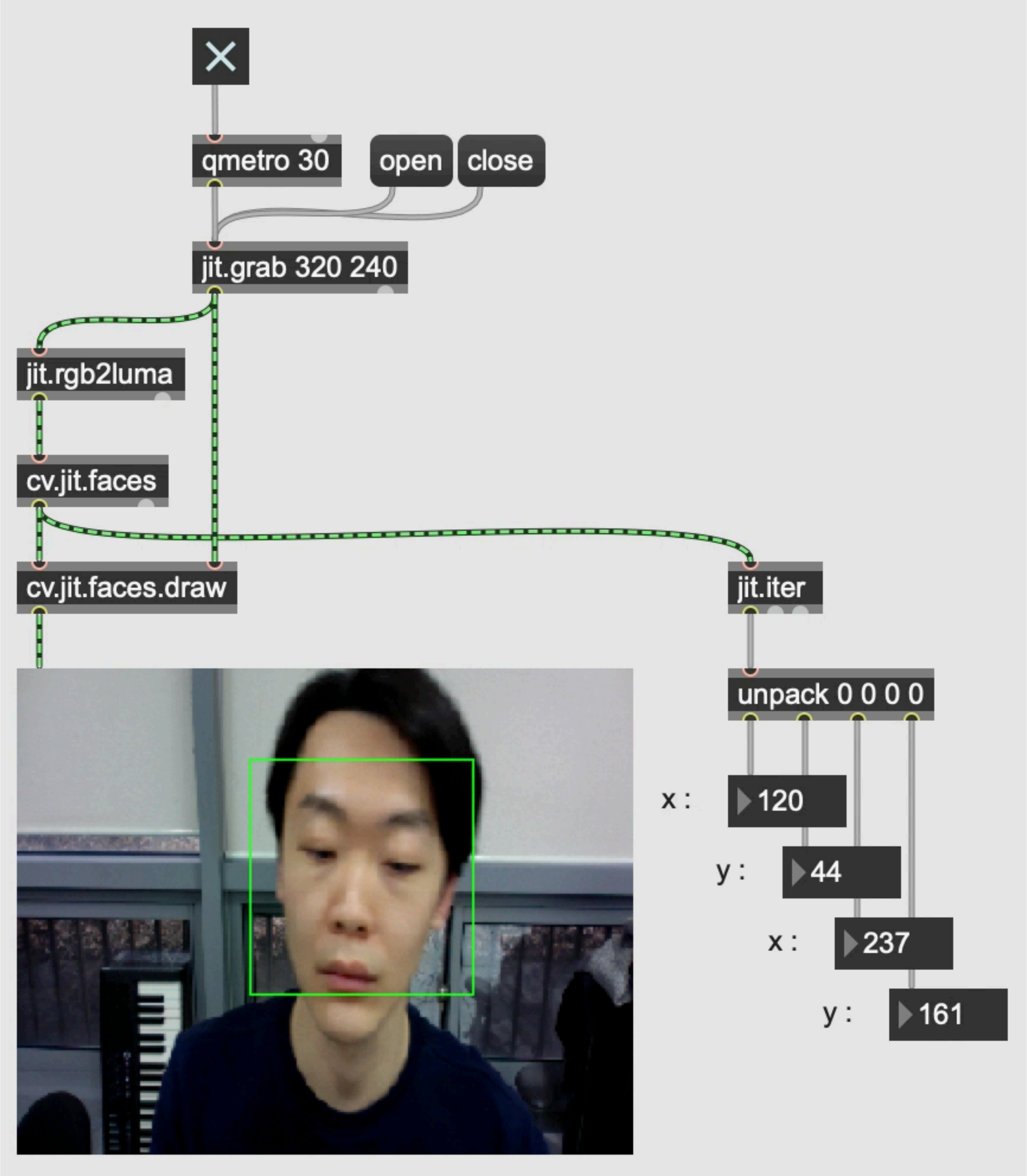


TouchDesigner File Edit Dialogs View Window Help
Thu Mar 25 7:43 AM
Kinect for Windows 2 (29/29 fps)
TouchDesigner 2021.11180: /Users/igwangyu/Desktop/NewProject.1.toe
[WIKI] [FORUM] [TUTORIALS] FPS: 18 Realtime 7:30:19 1 Operators deleted.
Pane Layout New Layout
/ project1
-0.2764 Head1
1.275 Head2
1.823 Head3
-0.4303 Neck1
1.089 Neck2
1.705 Neck3
-0.491 Torso1
0.9082 Torso2
1.668 Torso3
-0.314 R_Shoulder1
1.051 R_Shoulder2
1.716 R_Shoulder3
-0.546 L_Shoulder1
1.128 L_Shoulder2
1.696 L_Shoulder3
-0.278 R_Elbow1
0.6795 R_Elbow2
1.687 R_Elbow3
-0.7512 L_Elbow1
0.8887 L_Elbow2
1.567 L_Elbow3
-0.252 R_Hand1
0.4902 R_Hand2
1.483 R_Hand3
-0.8469 L_Hand1
0.6712 L_Hand2
1.449 L_Hand3
-0.4686 R_Hip1
0.7018 R_Hip2
1.638 R_Hip3
-0.6351 L_Hip1
0.7589 L_Hip2
1.626 L_Hip3
-0.4082 R_Knee1
0.3794 R_Knee2
1.37 R_Knee3
-0.742 L_Knee1
0.3727 L_Knee2
1.396 L_Knee3
-0.8703 R_Foot1
0.4983 R_Foot2
1.497 R_Foot3
-0.8671 L_Foot1
0.2559 L_Foot2
1.7 L_Foot3
-0.2526 R_Palm1
0.4901 R_Palm2
1.484 R_Palm3
-0.8467 L_Palm1
0.6703 L_Palm2
1.449 L_Palm3
oscin1
null1
sphere1
transform1
Transform transform1
Transform Post
Group
Transform Order Scale Rotate Translate
Rotate Order Rx Ry Rz
Translate 0 0 0
Rotate 0 0 0
Scale 1 1 1
Pivot 0 0 0
Uniform Scale 1
Normals Maintain Length On
Look At
Up Vector 0 1 0
Forward Direction
Sep 25, 2017 — The face
Right below the "calibrate
V1.X Documentation - V1
Documentation: User inte
Skeleton tracking - Docun
4.4 OSC (Open Sound C
More results from forum.n
community.troikatronix.co
Motion Capture/Tr
Oct 28, 2017 — Motion C
Mate and I have manage
outgoing port number in N
community.troikatronix.co
[ANSWERED] Kin
it ...
In terms of a comparison
spatial trigger point calibr
Start: 1 End: 600
RStart: 1 REnd: 600
FPS: 60.0 Tempo: 120.0
ResetF: 1 T Sig: 4 4
TimeCode Beats
00:00:08.55
536
Range Limit
Loop Once
Time Path: /

Computer Vision

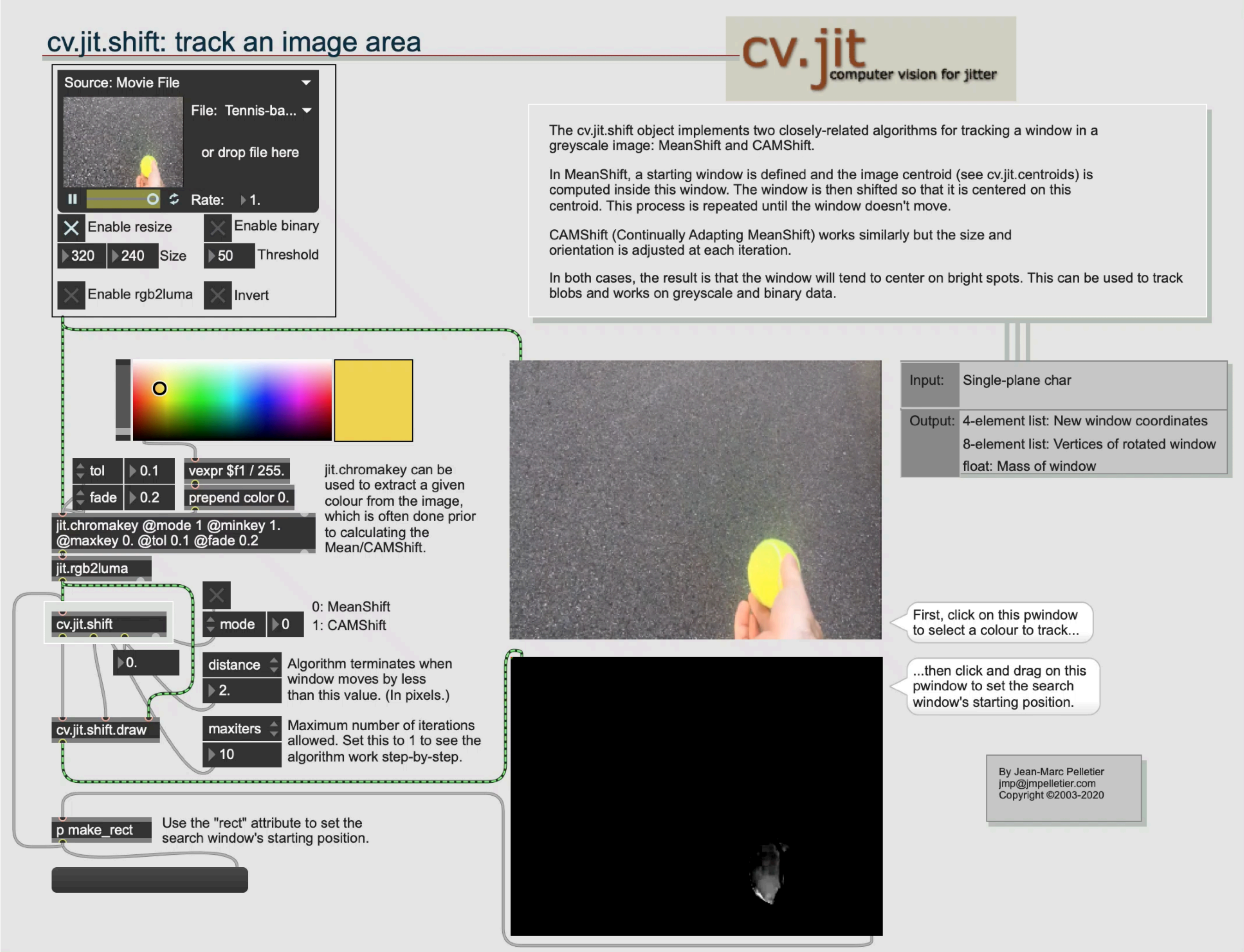


Computer Vision



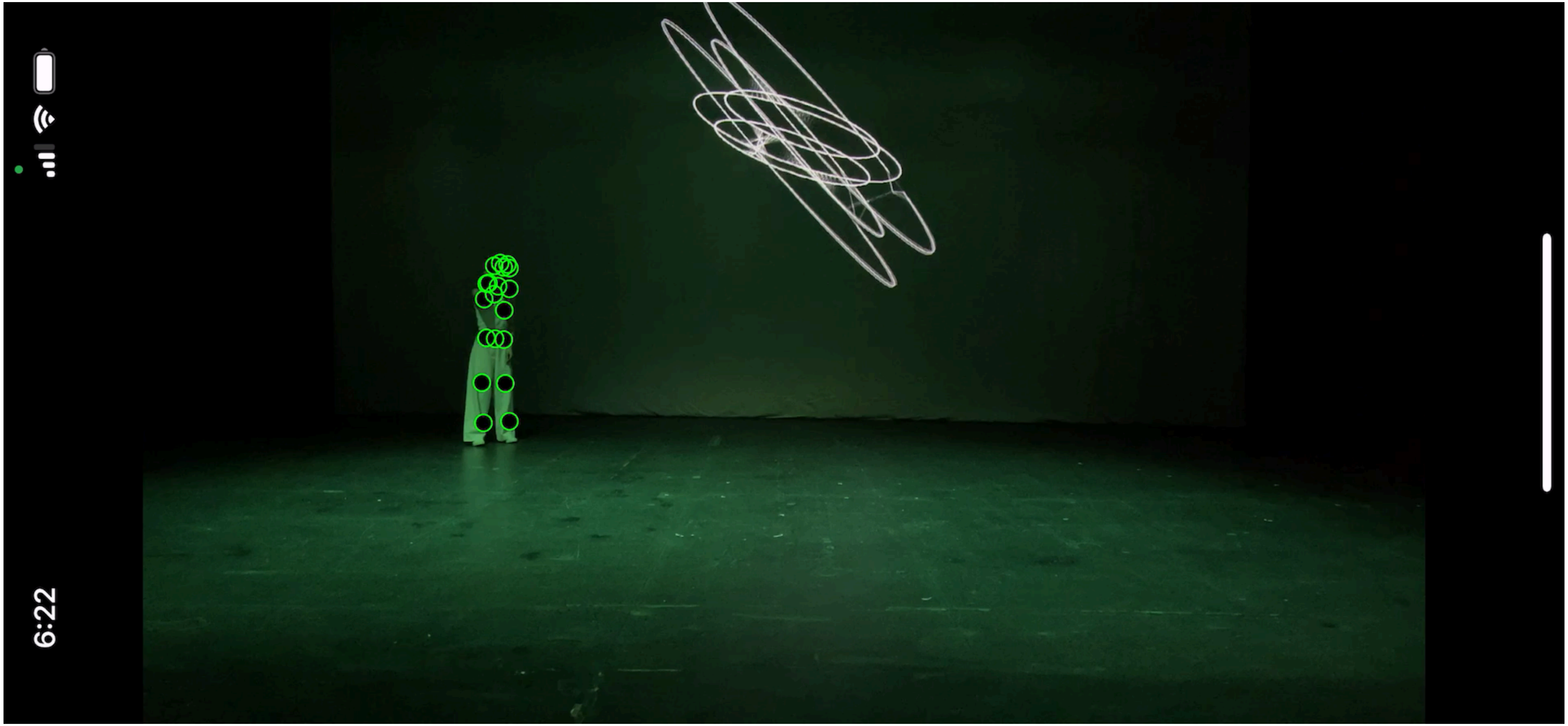
Link

Computer Vision



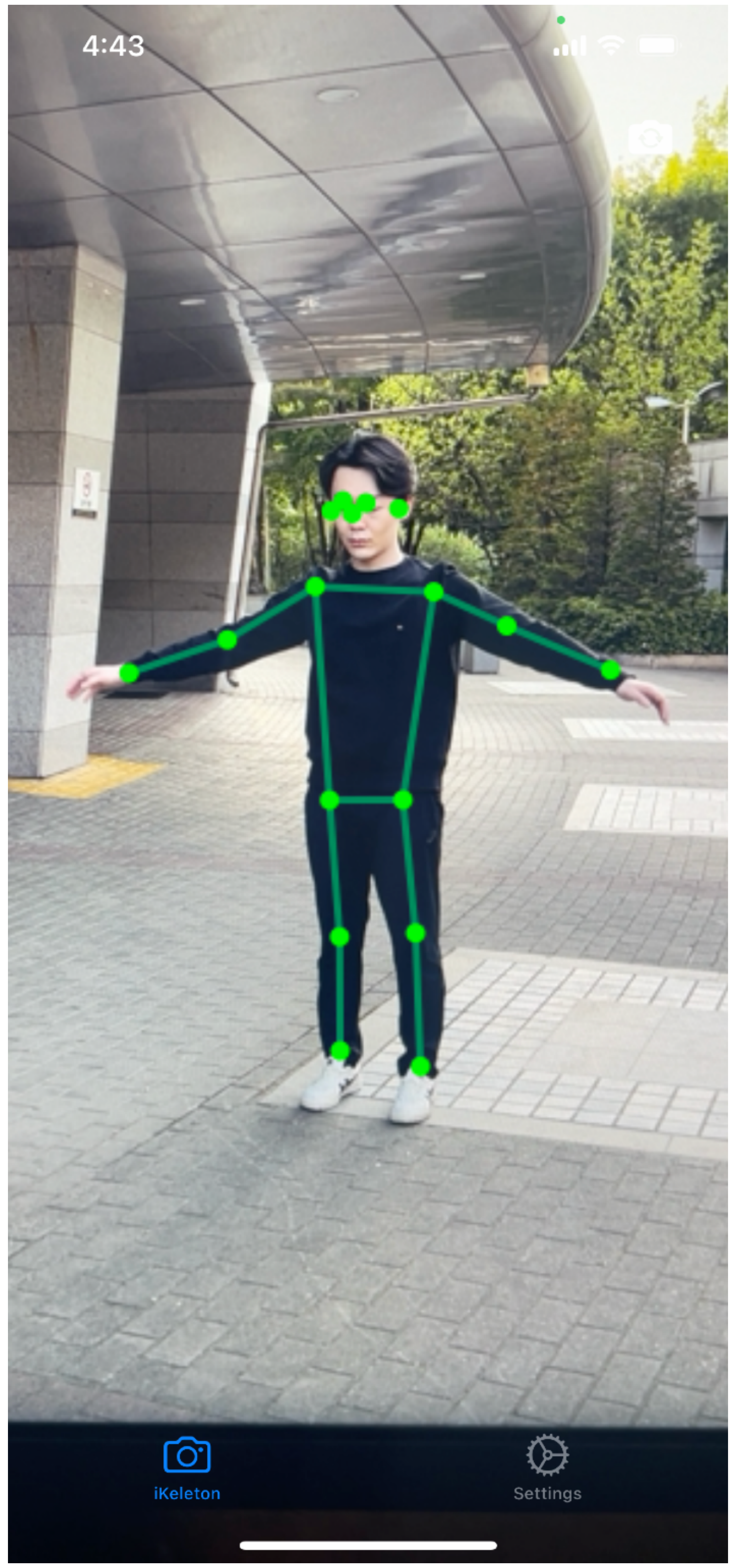
Link

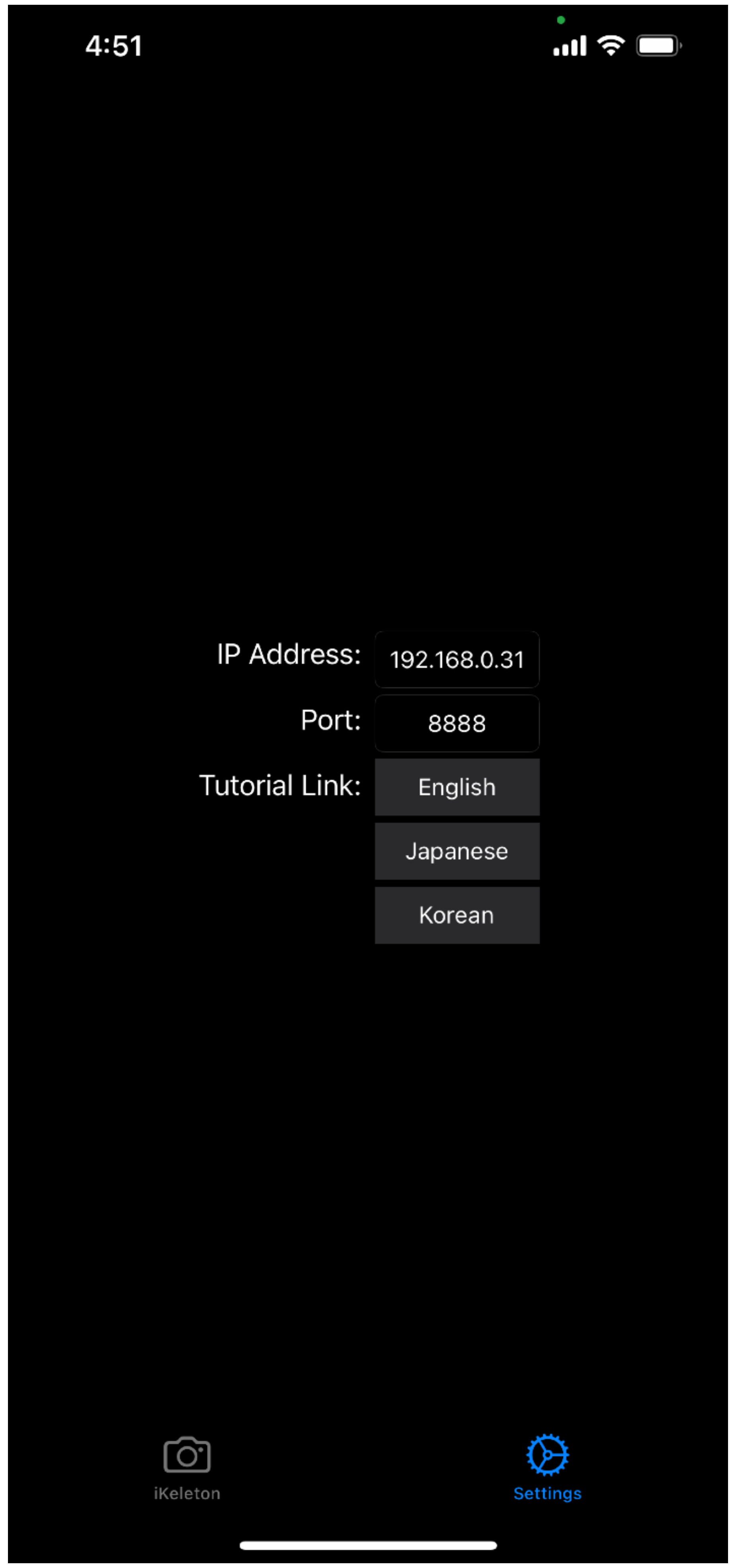
Computer Vision & Machine Learning



Link

Computer Vision & Machine Learning





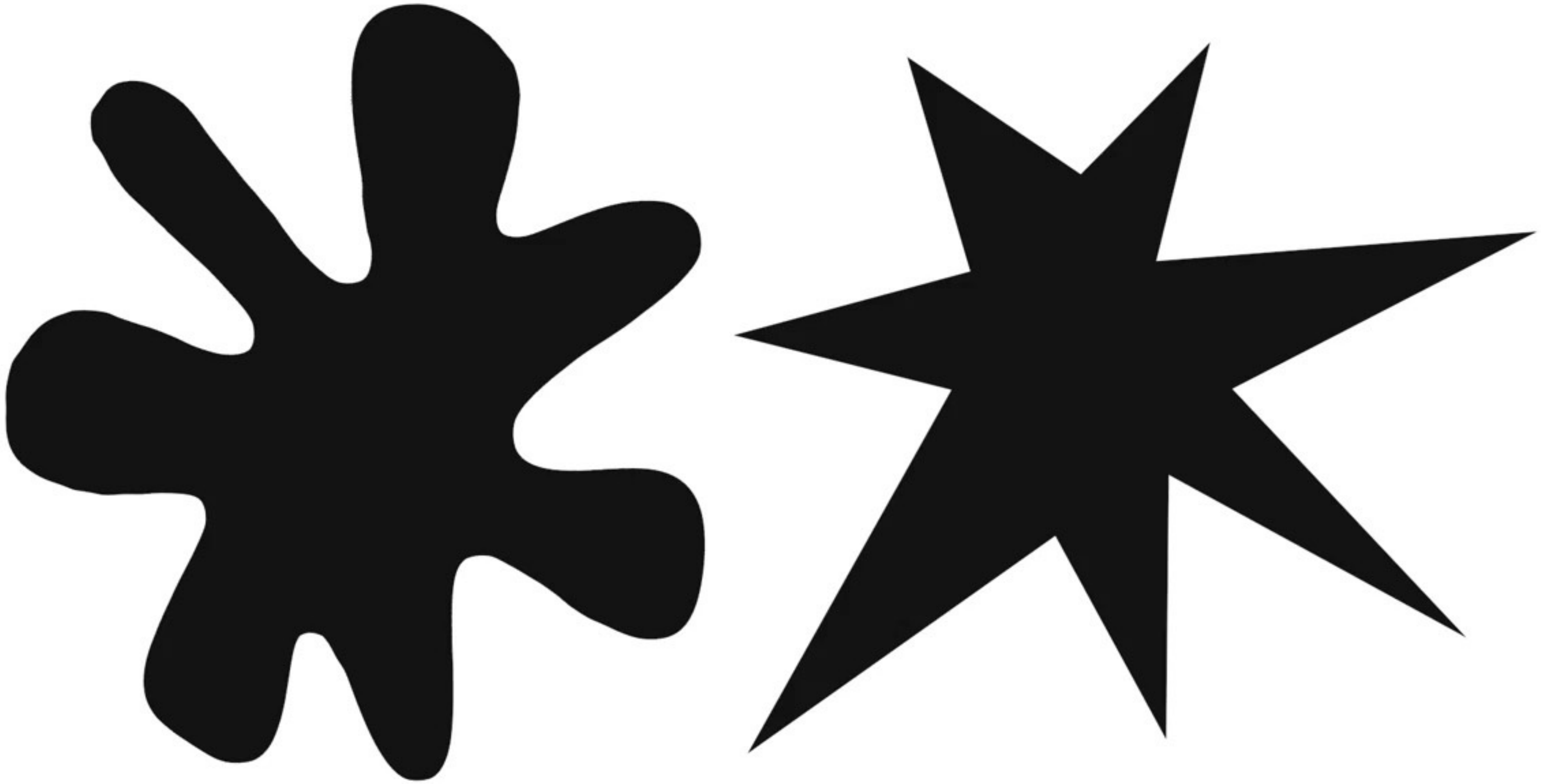
Link

Mapping Motion To Music

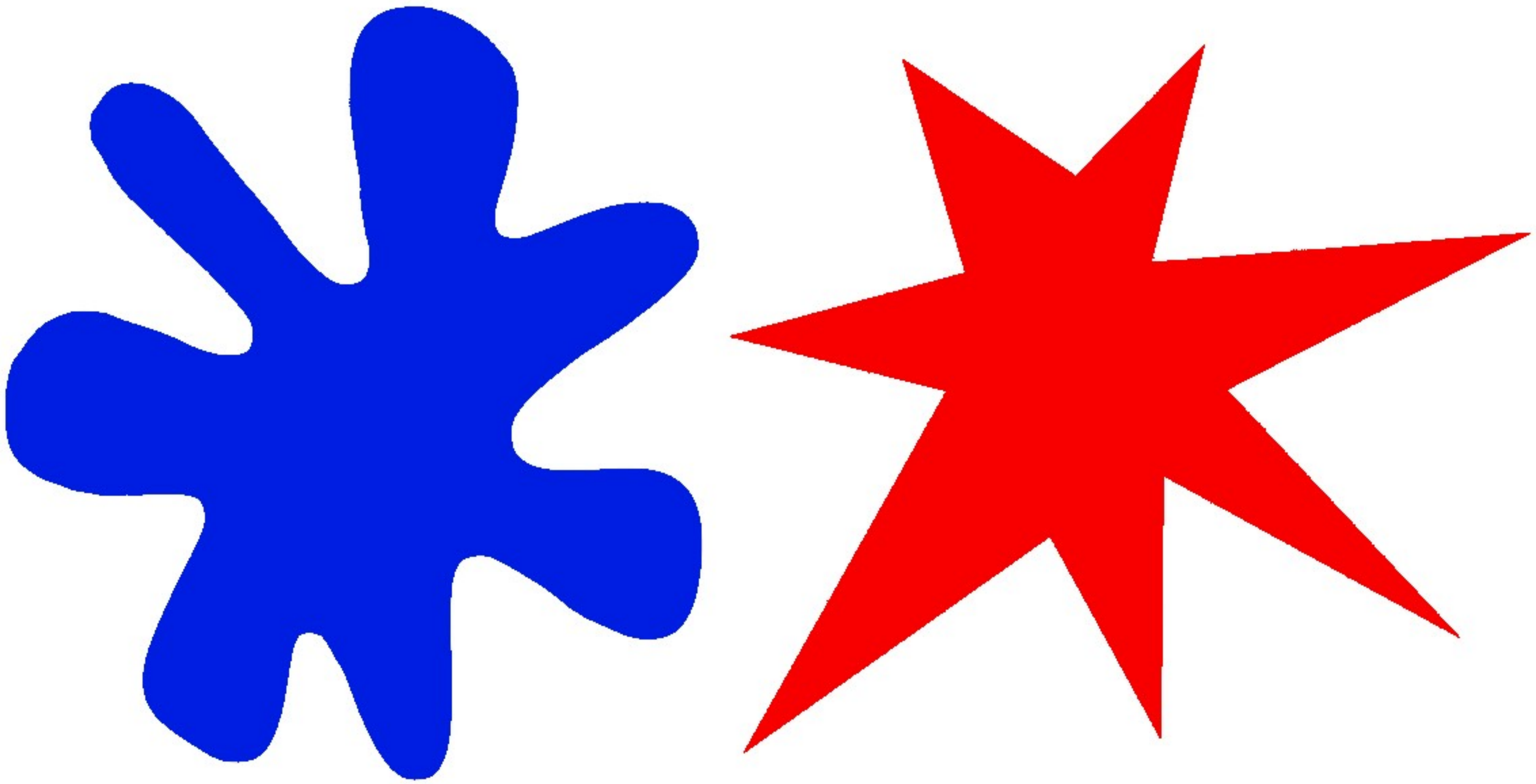
Kick Drum vs Fade Out

Link

Kiki & Bouba

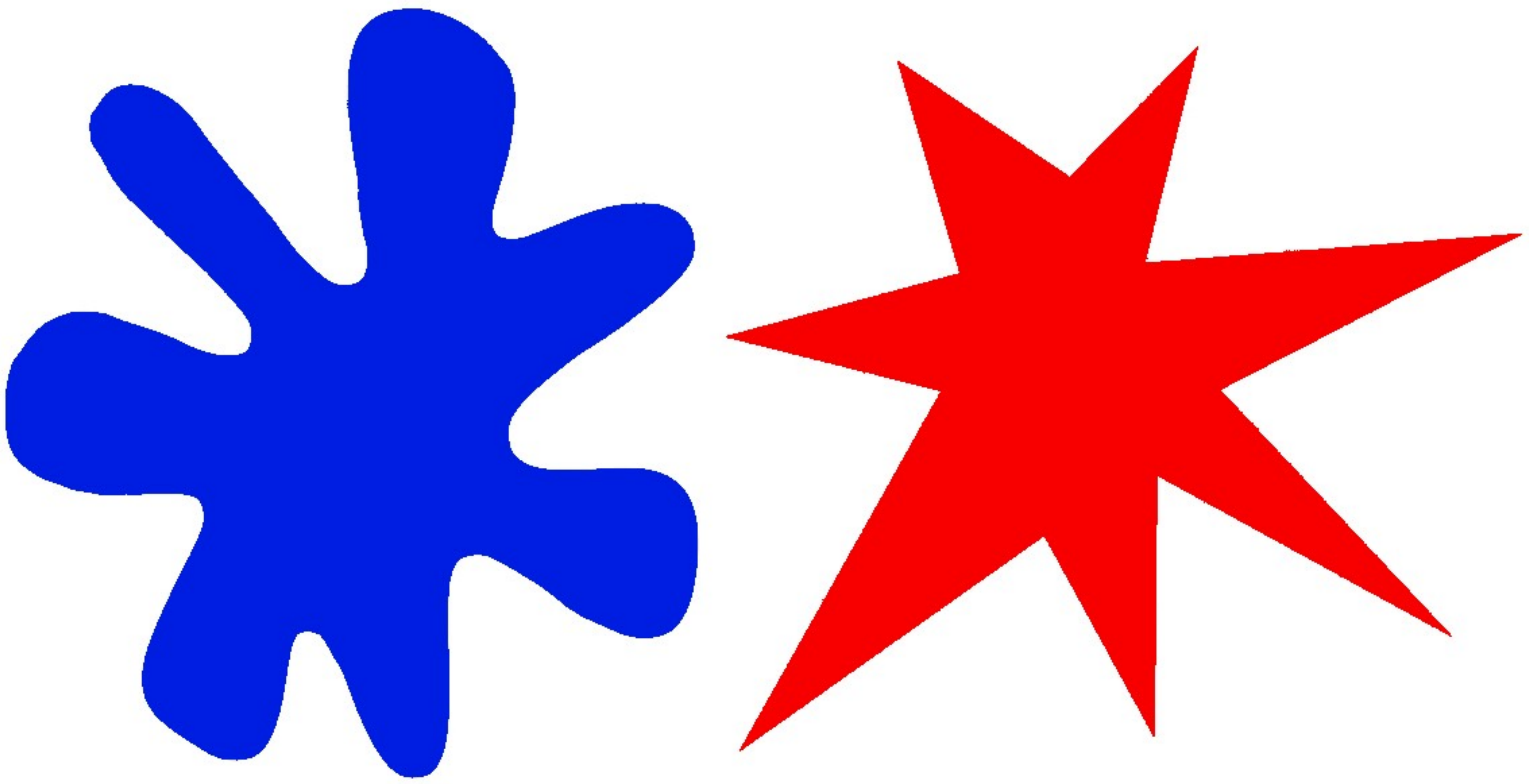


Link



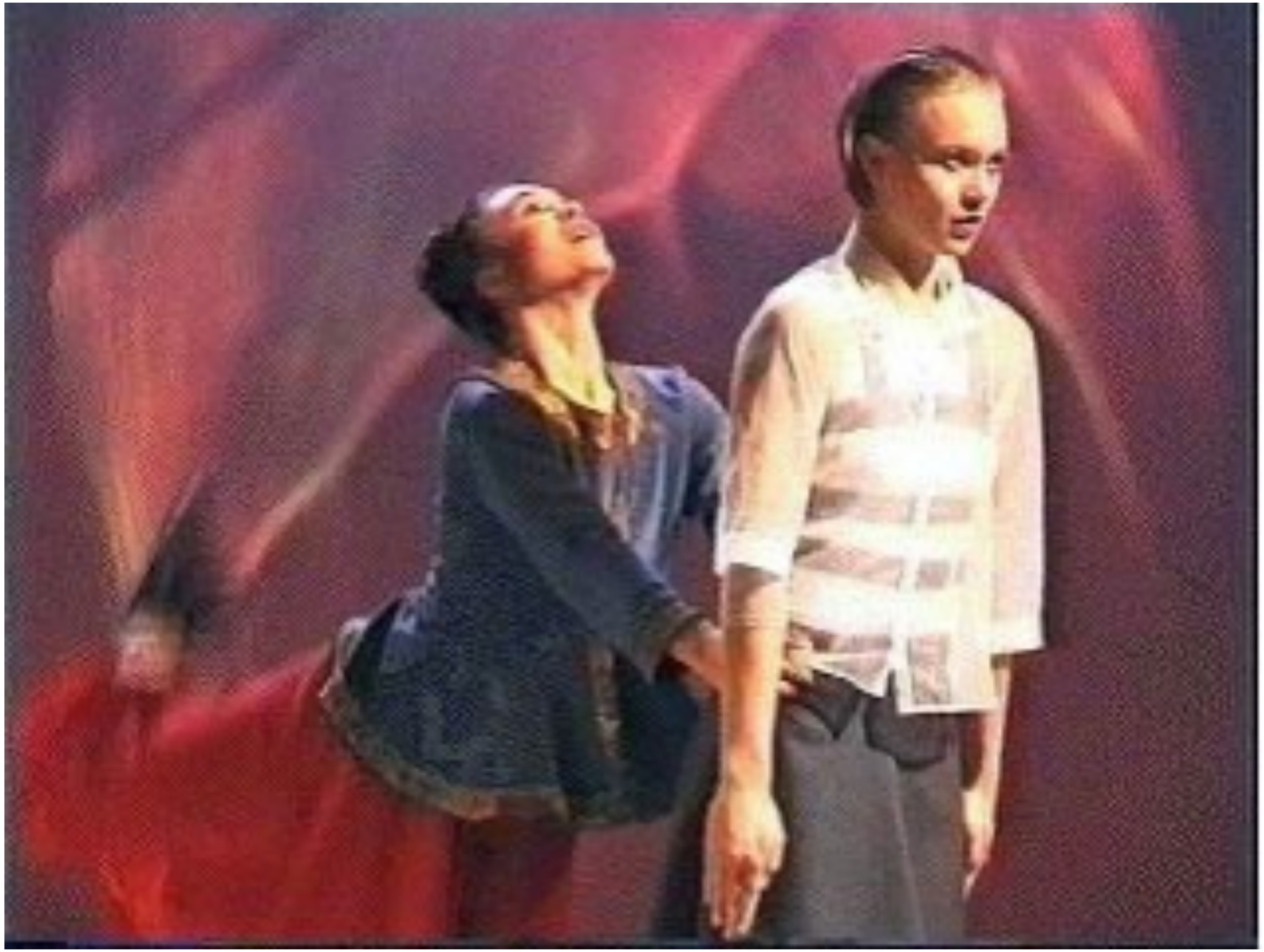
Link

Bouba & Kiki



Link

Sisters



http://waynesiegel.dk/wp-content/uploads/2013/08/Sisters.pdf

Pandora Project

http://waynesiegel.dk/?page_id=2703

Pandora Project

- Workshop
- Experiments
- Sketches

Liquid

Who

The Box

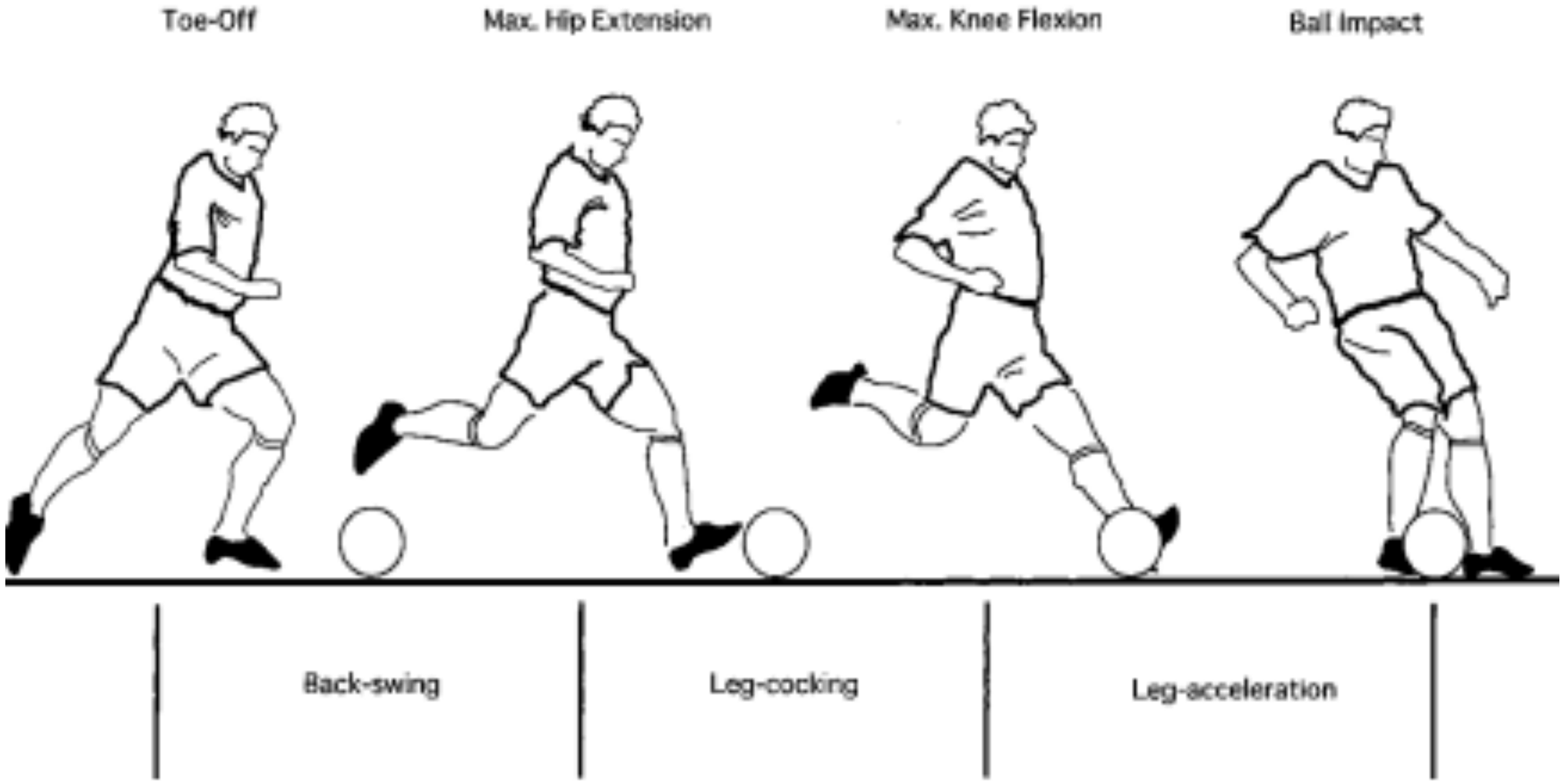
Fire

Circles

https://www.gwangyulee.com/2022/01/circlesfor-dance-piano-real-time-sound.html

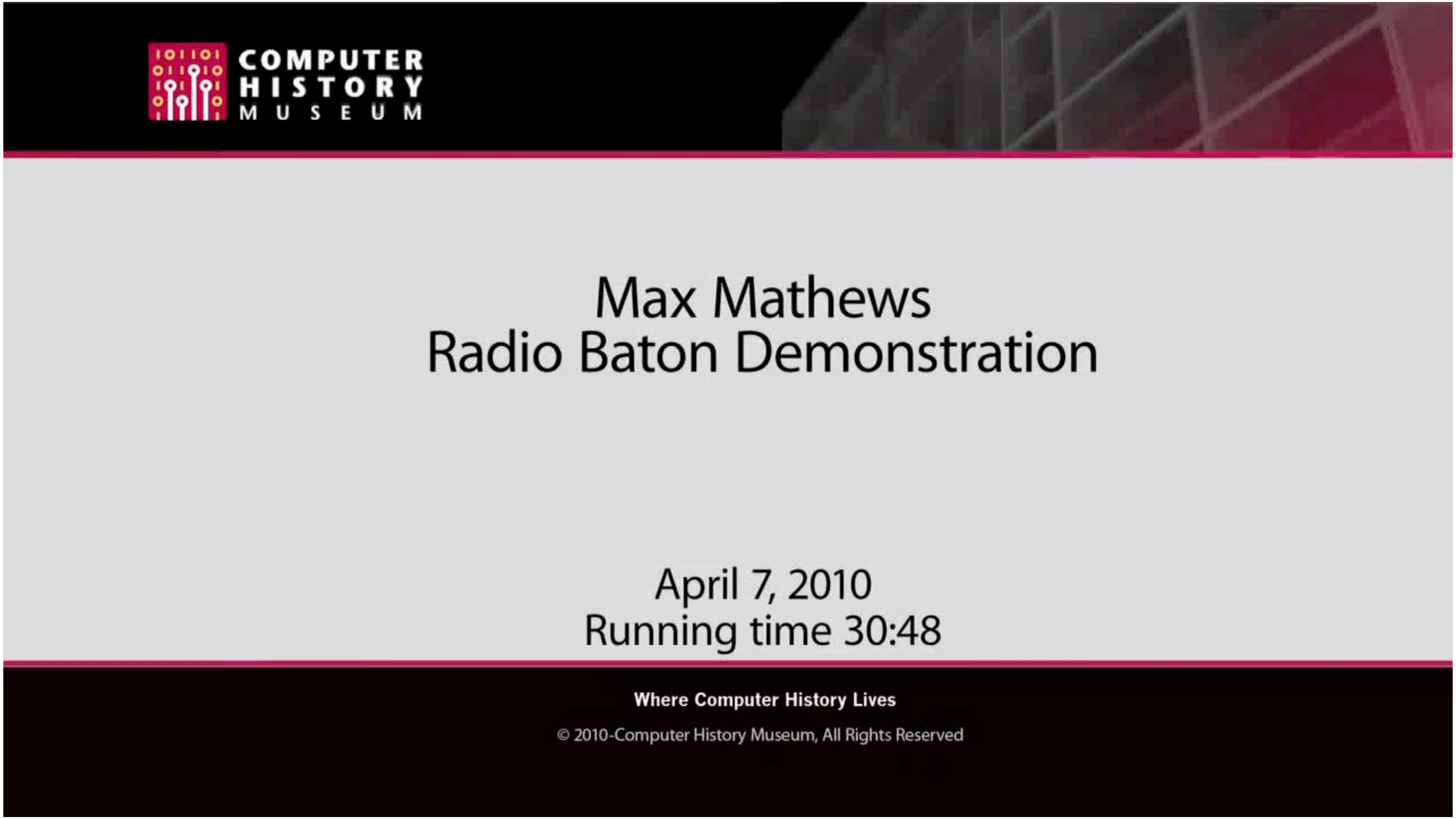
Gesture Interaction

Gesture



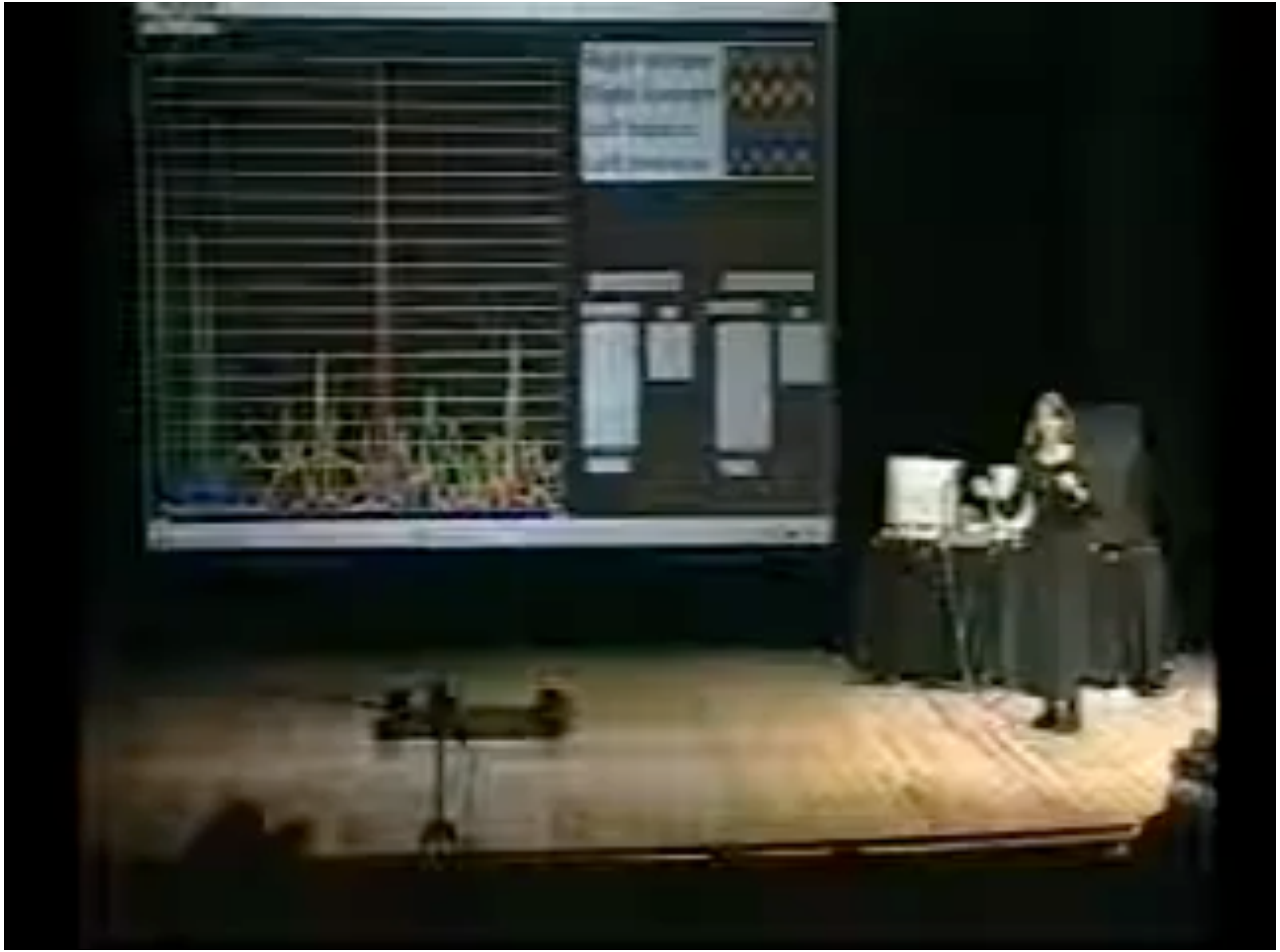
Link

Radio Baton



https://youtu.be/3ZOzUVD4oLg?feature=shared&t=344

Conductor Jacket



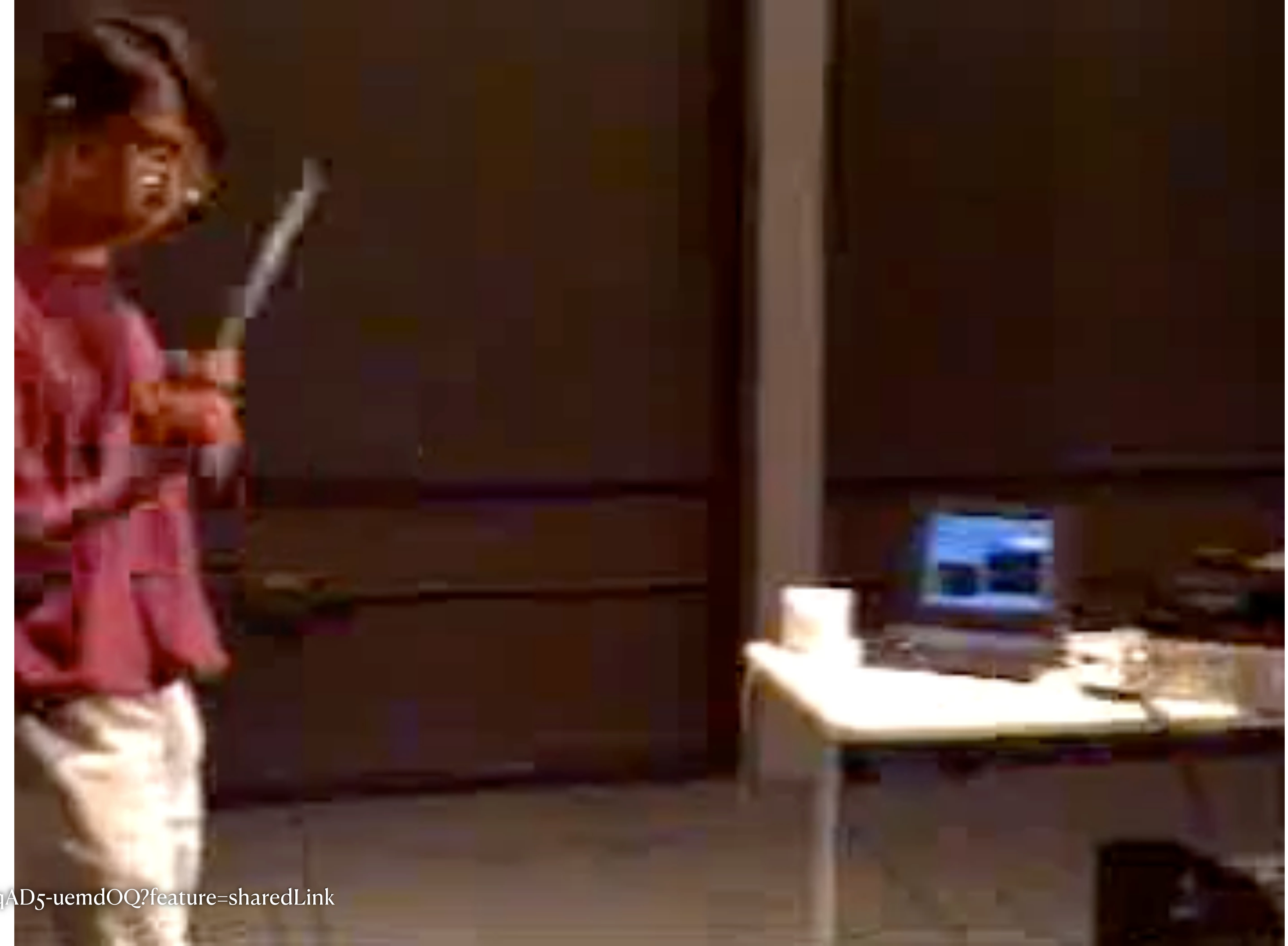
https://youtu.be/Lnpj7MSRFYM?feature=shared

UBS Virtual Maestro (aka "Conductor Hero")

Rationale for Selecting Wii Remote

https://youtu.be/eJuIS2-wdfU?feature=shared

Brain Opera: Digital Baton



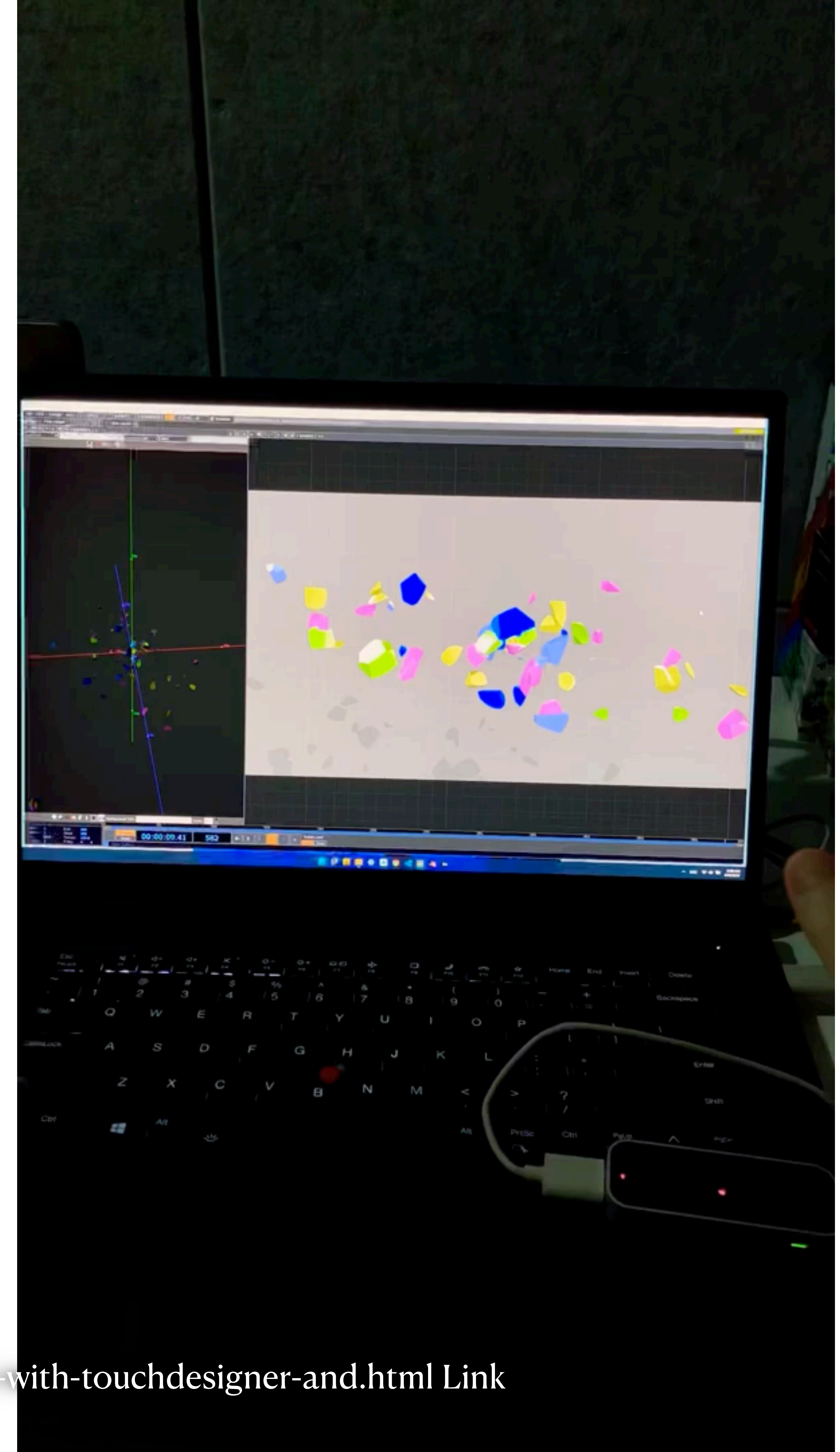
https://youtu.be/qAD5-uemdOQ?feature=sharedLink

Brain Opera: Sensor Chair

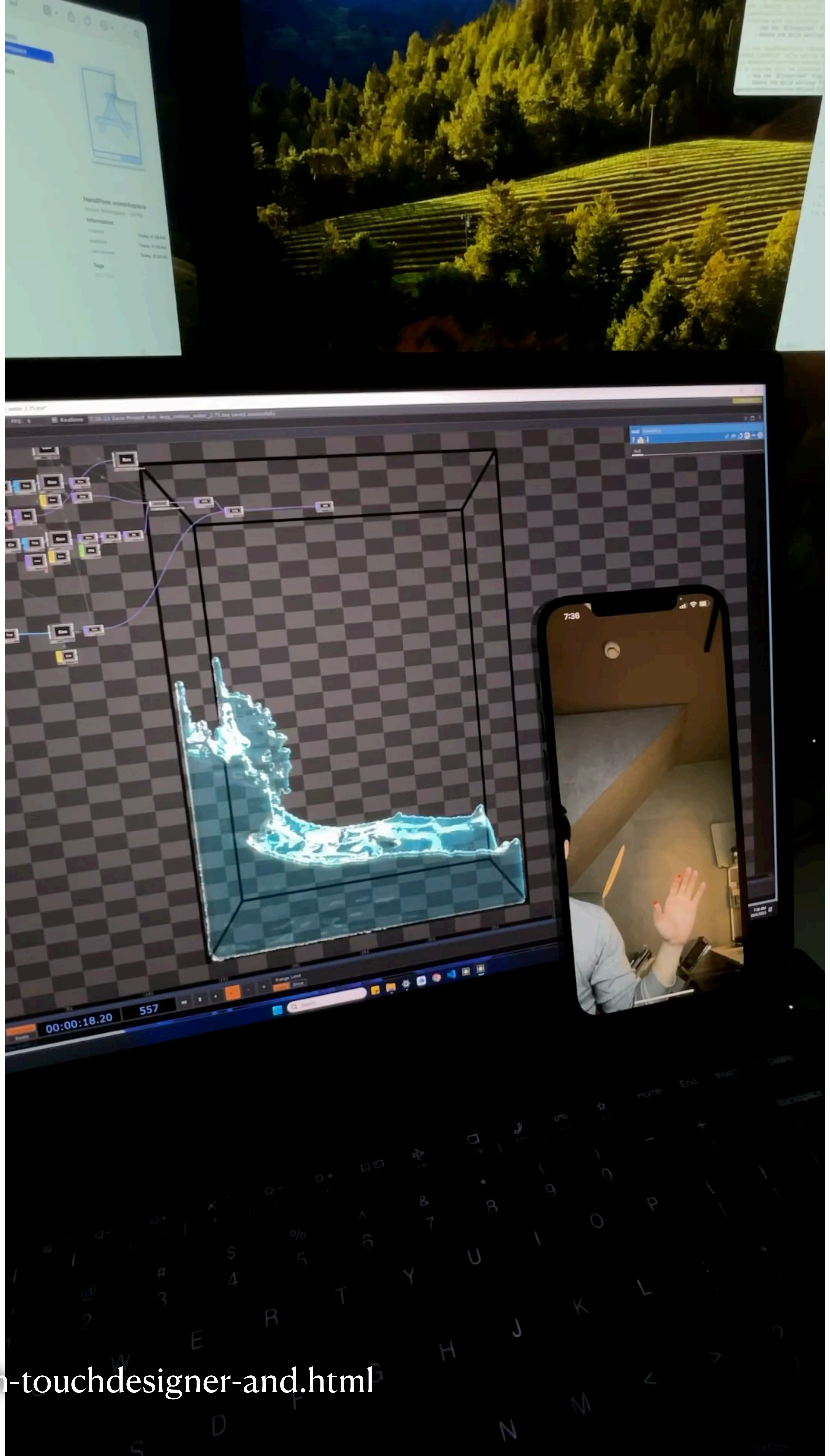
https://youtu.be/qAD5-uemdOQ?feature=sharedLink

Leap Motion

Link



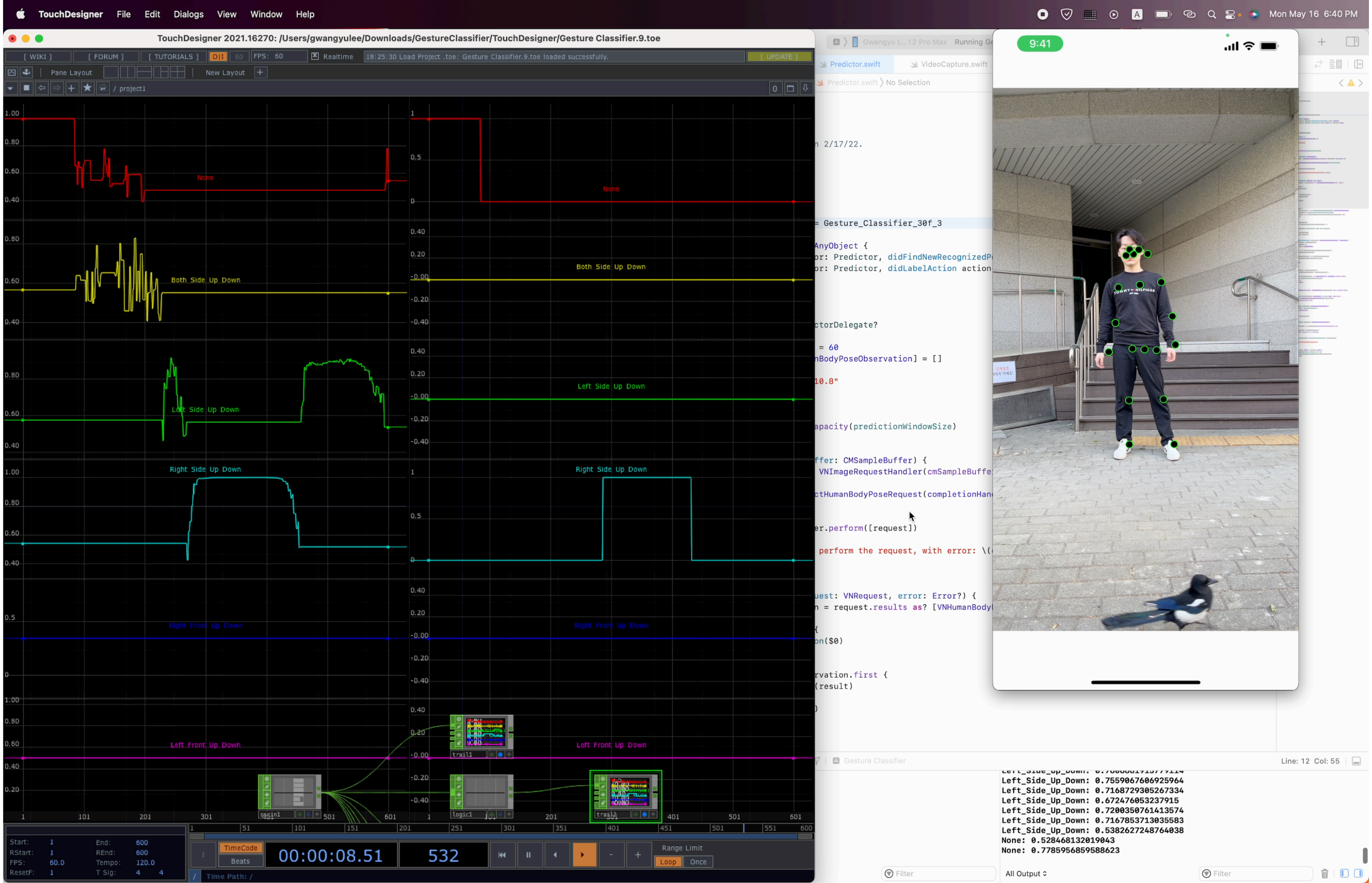
https://www.gwangyulee.com/2022/04/destroyed-sphere-with-touchdesigner-and.html Link



https://www.gwangyulee.com/2023/10/real-time-wave-with-touchdesigner-and.html

Link

Machine Learning



TouchDesigner
File
Edit
Dialogs
View
Window
Help
Mon May 16 6:40 PM
TouchDesigner 2021.16270: /Users/gwangyulee/Downloads/GestureClassifier/TouchDesigner/Gesture Classifier.9.toe
[WIKI]
[FORUM]
[TUTORIALS]
FPS: 60
Realtime
18:25:30 Load Project .toe: Gesture Classifier.9.toe loaded successfully.
[UPDATE]
Pane Layout
New Layout
/ project1
None
Both Side Up Down
Left Side Up Down
Right Side Up Down
Right Front Up Down
Left Front Up Down
trail1
oscin1
logic1
trail2
Start: 1
End: 600
RStart: 1
REnd: 600
FPS: 60.0
Tempo: 120.0
ResetF: 1
T Sig: 4 4
TimeCode
Beats
00:00:08.51
532
Range Limit
Loop
Once
Time Path: /
Predictor.swift
VideoCapture.swift
No Selection
n 2/17/22.
= Gesture_Classifier_30f_3
AnyObject {
or: Predictor, didFindNewRecognizedP
or: Predictor, didLabelAction action
ctorDelegate?
= 60
nBodyPoseObservation] = []
10.8"
apacity(predictionWindowSize)
ffer: CMSampleBuffer) {
VNImageRequestHandler(cmSampleBuffe
ctHumanBodyPoseRequest(completionHan
er.perform([request])
perform the request, with error: \(
uest: VNRequest, error: Error?) {
n = request.results as? [VNHumanBody
on($0)
rvation.first {
(result)
9:41
Gesture Classifier
Line: 12 Col: 55
Left_Side_Up_Down: 0.7559067606925964
Left_Side_Up_Down: 0.7168729305267334
Left_Side_Up_Down: 0.672476053237915
Left_Side_Up_Down: 0.7200350761413574
Left_Side_Up_Down: 0.7167853713035583
Left_Side_Up_Down: 0.5382627248764038
None: 0.528468132019043
None: 0.7785956859588623
Filter
All Output

An interactive system using gesture recognition for multimedia performance

Gwangyu Lee
MARTE Lab.
Dongguk University
Jung-gu Pil-dong, Seoul, Korea
frd.lee@icloud.com

Jun Kim
MARTE Lab.
Dongguk University
Jung-gu Pil-dong, Seoul, Korea
music@dongguk.edu

Abstract

The production of interactive multimedia performances by combining music, visuals, and dance has been attempted using various equipment. This study aims to develop an interactive system using machine learning to classify gestures. To keep the system simple, an iPhone and CoreML were used with an app to classify gestures and send the detected gestures and level of correspondence through a network. The data received were used to control music and visuals on a computer.

Keywords

Gesture recognition; Machine learning; CoreML; Music; Visuals; Dance.

1. Introduction

Music, visuals, and dance are different art forms, although they are inseparable. Most music performances go with visuals and dance; it is difficult to find a music performance that does not include visuals or dance.
However, in most performances, dancers follow the music, taking their cues from predetermined directions while visuals are independent. What if music, visuals, and dance could interact in a single multimedia performance [5]? Such systems have been attempted in the past [1-3, 5-12]. For example, visuals have been changed according to the amplitude of the music or the movement of the dancer, music changed according to the dancer's movement [14, 15], and so on. Consider an interactive multimedia performance of the dancer's movement. If the dancer's hands are raised, the music and visuals change. If the composer wants to use the movement only in a particular part, the operator should turn off the signal of the movement in the rest of the piece. If so, what if the computer could identify the dancer's gesture and use it in the interactive multimedia performance?
This study aims to create an interactive system for multimedia performance using gesture recognition. The programs used are Max/MSP[1] and TouchDesigner[2] on a computer, with machine learning used for gesture recognition.

2. Machine Learning Model Design

This study uses machine learning to identify a specific gesture. Machine learning, a field of artificial intelligence, uses computer algorithms that automatically improve

[1] Max/MSP, also known as Max/MSP/Jitter, is a visual programming language for music and multimedia developed and maintained by the San Francisco-based software company Cycling '74.

[2] TouchDesigner is a node-based visual programming language for real-time interactive multimedia content, developed by the Toronto-based company Derivative.

Link

through experience. In this study, we chose CoreML, which applies a machine-learning algorithm to a training dataset to create a model, allowing the iOS app to run the models locally.

Figure 1. CoreML workflow[3]

2.1 Rationale for Selecting CoreML

A machine-learning program using a computer-vision system that receives the motion signal in real time is required to identify the dancer's motion. It can be implemented using an ordinary webcam or Azure Kinect DK [2, 4, 13]. However, there are some reasons that we chose to use CoreML and the iPhone's camera in this study. First, they keep the system simple. If we used an ordinary camera and machine learning, we would need one more program on the computer. In contrast, a camera and machine learning can be integrated into one iPhone.

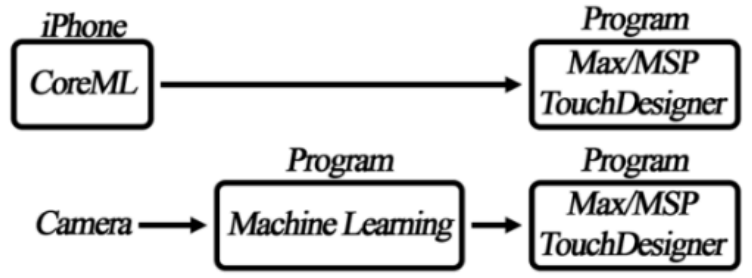


Figure 2. System architecture comparison

Second is the possibilities available with mobile phone cameras. Cell phone cameras are developing tremendously quickly, and a machine-learning model using a LiDAR sensor or more could be applied in future studies.

[3] https://developer.apple.com/documentation/coreml

The last reason is portability. Anyone with an iPhone can implement the system without limits on time or place.

2.2 Rationale for Using Machine Learning

Why use machine learning? Compare previous methods with the data from machine learning.
For example, in the motion of raising a hand, what is the difference between mapping data by obtaining coordinates and recognizing a specific gesture using machine learning? Machine learning can distinguish the direction the hand is moving. Of course, one could write a program to distinguish the direction. However, machine learning is simpler and more effective.
The change in the y-axis is the same when raising the hand to the side and raising it forward. The result is the same even if the speed value of the movement is obtained. However, machine learning allows the capture of various data by separating the components of the hand or other parts of the body's movement.

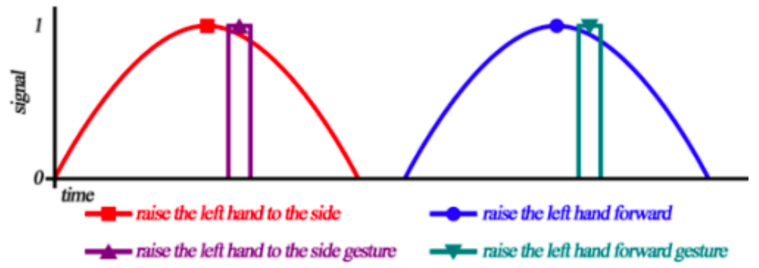


Figure 3. Comparison between y-axis measurement and machine learning

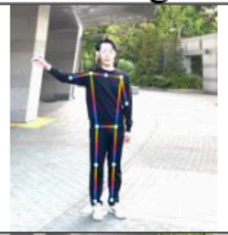
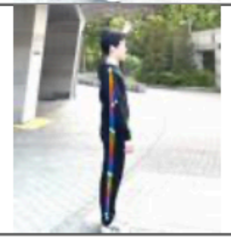
2.3 Dataset for Machine Learning

We had to prepare videos for the machine-learning dataset. This study classified six gestures (raising both hands to the side, raising the left hand to the side, raising the right hand to the side, raising the left hand

Link

forward, raising the right hand forward, and no movement).
Table 1 shows examples of sideways gestures. Videos taken from various angles were edited at two-second intervals. The gesture of raising the hand forward was organized similarly. There were 40 or more videos per gesture.

Table 1. Examples of sideways gestures from the dataset

Dataset		
Raise both	Raise left	Raise right
		
		
		

2.4 Results of Machine Learning

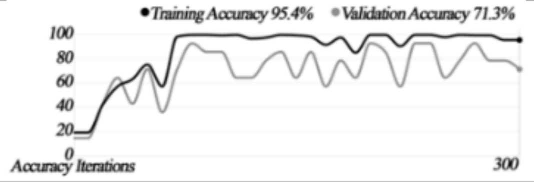
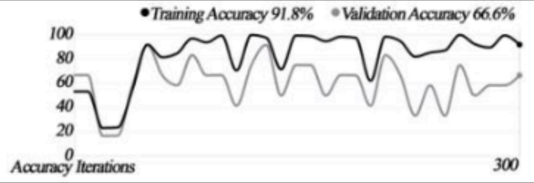
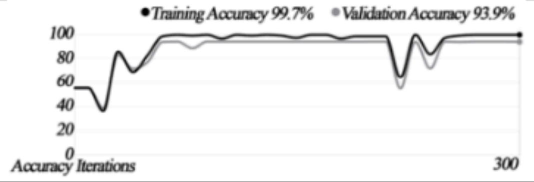
We built three machine-learning models, each of which went through 300 iterations. The results and graphs are as follows.

Table 2. Gesture classification by models

Model 1	raise both hands to the side raise the left hand to the side raise the right hand to the side raise the left hand forward raise the right hand forward no movement
Model 2	raise the left hand to the side raise the right hand to the side raise the left hand forward raise the right hand forward no movement
Model 3	raise both hands to the side raise the left hand to the side raise the right hand to the side no movement

The training accuracy of all the models was greater than 90%, although the graphs show that Models 2 and 3 found it slightly difficult to distinguish between forward and sideways movement.

Table 3. Accuracy graphs by models

Model 1	
Model 2	
Model 3	

3. Implementation

We had to build an iOS app to use the machine-learning model. The system in this study was implemented on an iPhone 12 Pro Max.

3.1 App Operation

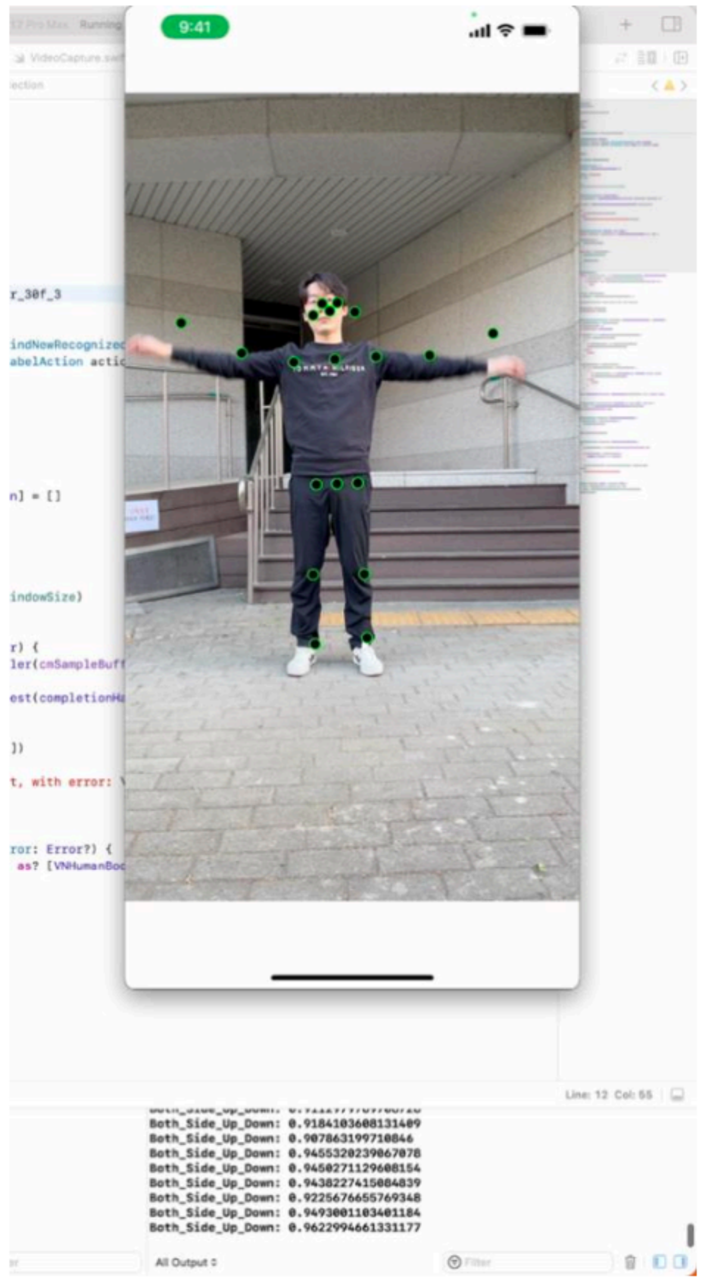


Figure 4. Screen and console of the app

3.2 Data Capture

The steps in this section were as follows. The gesture was classified on the iPhone. Data identifying the gesture detected (labeled) and the correspondence rate (confidence level) via OSC[4]. Therefore, we needed to obtain the current gesture first and then send it to Max/MSP and TouchDesigner via OSC.

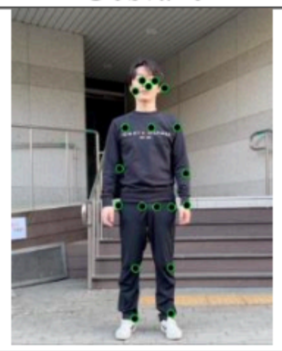
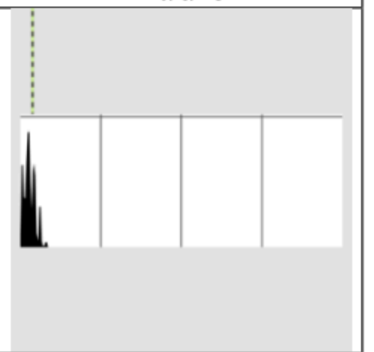
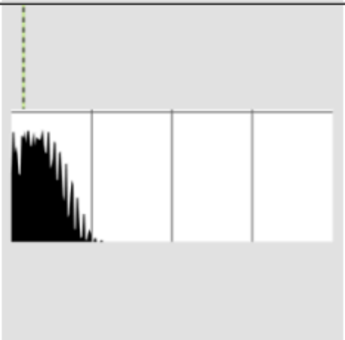
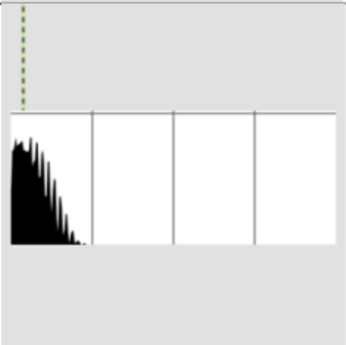
[4] Open Sound Control (OSC) is a protocol for networking sound synthesizers, computers, and other multimedia devices for musi-cal performance or show control purposes.

4. Multimedia Control

4.1 Control of Musical Instruments

We could change the instrument's timbre when we recognized a particular gesture.

Table 4. Control of musical instruments

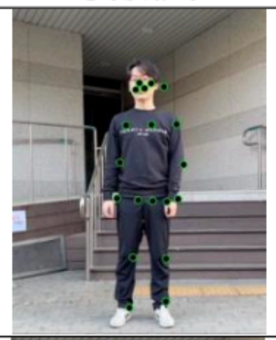
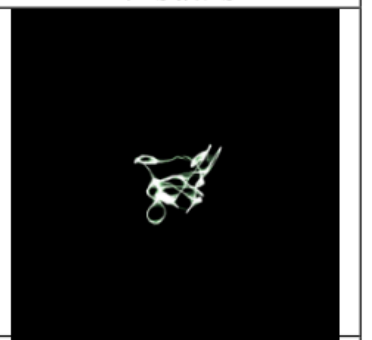
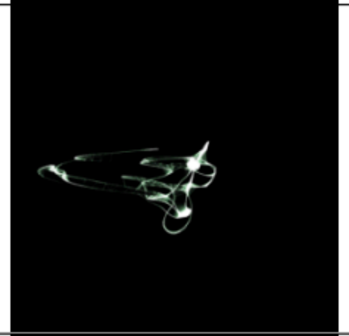
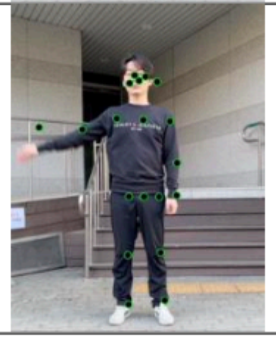
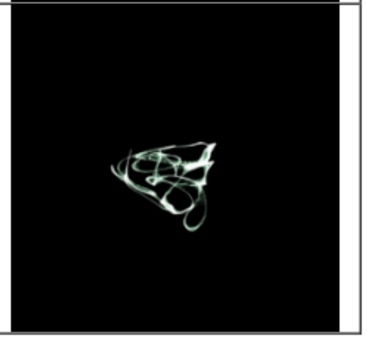
Gesture	Audio
	
	
	

Link

4.2 Control of Visuals

We could change the visuals when we recognized a particular gesture.

Table 5. Control of visuals

Gesture	Visuals
	
	
	

5. Conclusions

Currently, many interactive multimedia performances combine music, visuals, and dance, and artists make various efforts to recognize dancers' movements and gestures. Various equipment is used to create an interactive multimedia performance. We were able to use an iPhone rather than professional equipment. In addition, gestures were classified using machine learning rather than a one-to-one mapping of movements to produce more effective results.

6. References

[1] Behringer, R. (2007, July). Gesture interaction for electronic music performance. In *International Conference on Human-Computer Interaction* (pp. 564-572). Springer, Berlin, Heidelberg.

[2] Bhattacharya, S., Czejdo, B., & Perez, N. (2012, November). Gesture classification with machine learning using kinect sensor data. In *2012 Third International Conference on Emerging Applications of Information Technology* (pp. 348-351). IEEE.

[3] Camurri, A., Coletta, P., Peri, M. F., Ricchetti, M., Ricci, A., Trocca, R., & Volpe, G. (2000, August). A real-time platform for interactive dance and music systems. In *ICMC*.

[4] Chin-Shyurng, F., Lee, S. E., & Wu, M. L. (2019). Real-time musical conducting gesture recognition based on a dynamic time warping classifier using a single-depth camera. *Applied Sciences*, *9*(3), 528.

[5] Dean, T. R. (2012). The Oxford Handbook of Computer Music. Oxford University Press

[6] Deng, L., Leung, H., Gu, N., & Yang, Y. (2011). Real-time mocap dance recognition for an interactive dancing game. *Computer animation and virtual worlds*, *22*(2-3), 229-237.

[7] James, J., Ingalls, T., Qian, G., Olsen, L., Whiteley, D., Wong, S., & Rikakis, T. (2006, October). Movement-based interactive dance performance.
In *Proceedings of the 14th ACM international conference on Multimedia* (pp. 470-480).

[8] Nakra, T. M.,, & Paradiso, J. A. (1997, September). The Digital Baton: a Versatile Performance Instrument. In *ICMC*.

[9] Nakra, T. M., & Picard, R. W. (1998, December). The" conductor's jacket": A device for recording expressive musical gestures. In *ICMC*.

[10] Nakra, T. M., Ivanov, Y., Smaragdis, P., & Ault, C. (2009, June). The UBS Virtual Maestro: an Interactive Conducting System. In *NIME* (pp. 250-255).

[11] Nakra, T. M. (2014). Interactive conducting systems overview and assessment. *The Journal of the Acoustical Society of America*, *135*(4), 2377-2377.

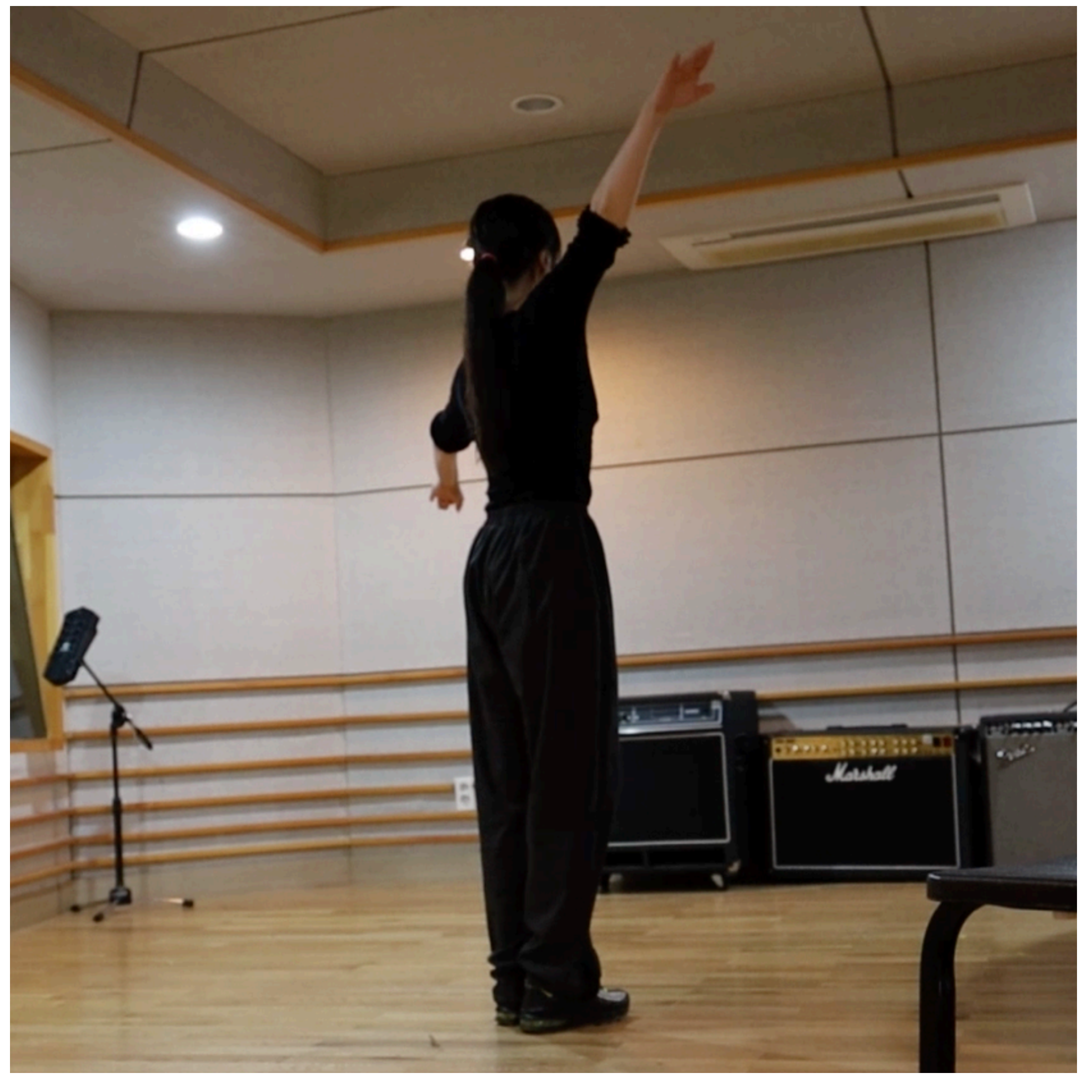
[12] Paradiso, J., & Sparacino, F. (1997). Optical tracking for music and dance performance. *Optical 3-D Measurement Techniques IV*, 11-18.

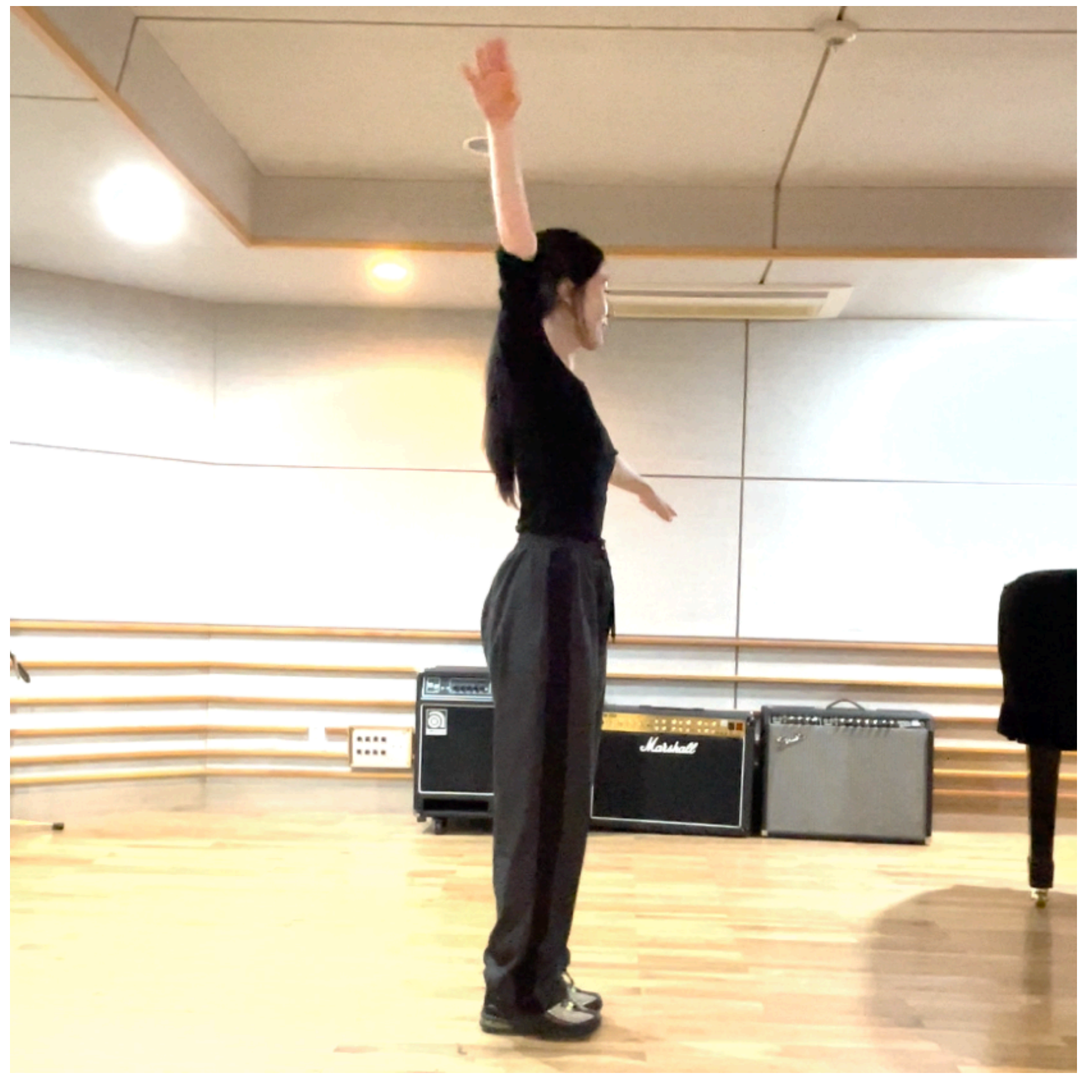
[13] Raptis, M., Kirovski, D., & Hoppe, H. (2011, August). Real-time classification of dance gestures from skeleton animation. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics symposium on computer animation* (pp. 147-156).

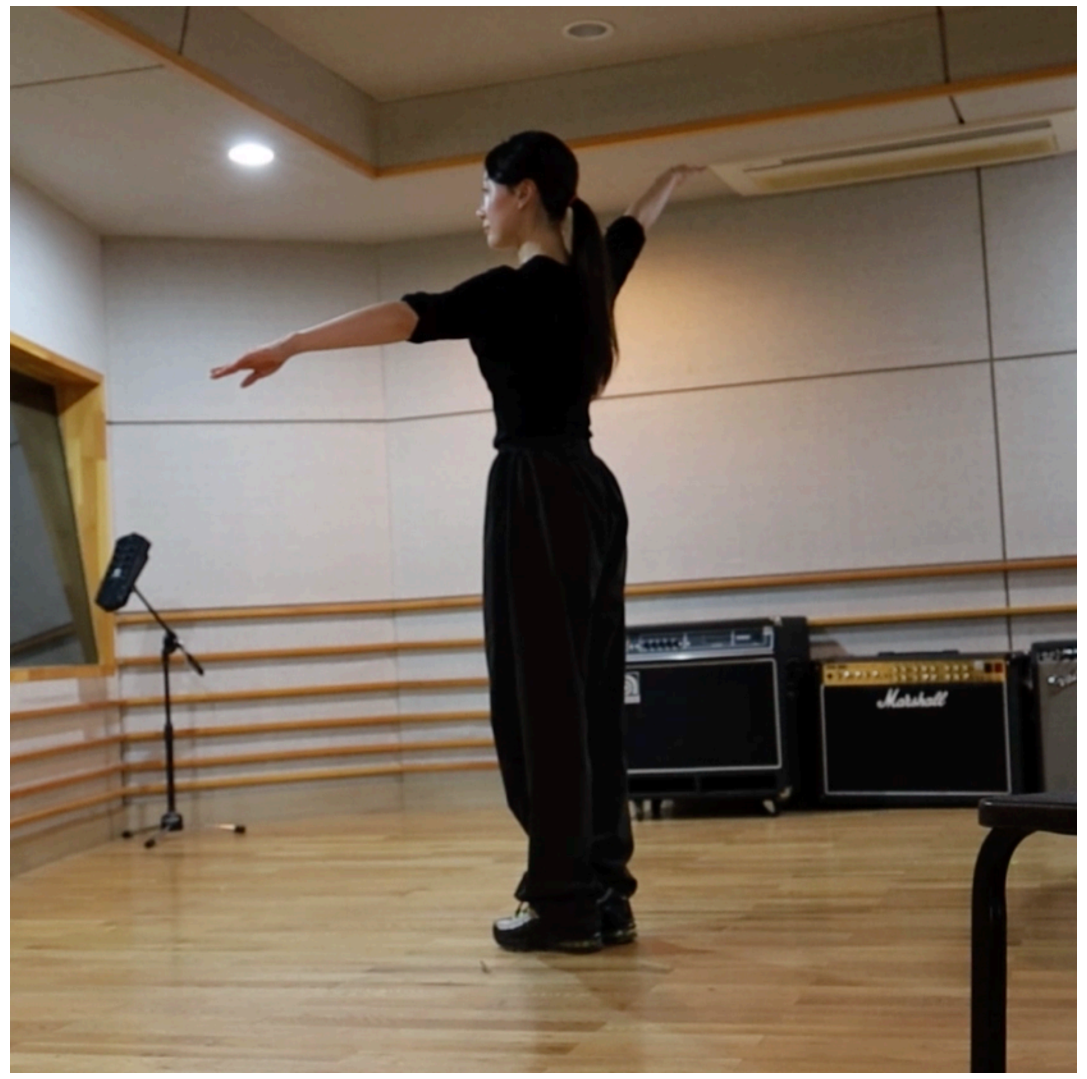
[14] http://waynesiegel.dk/wp-content/uploads/2013/08/Sisters.pdf Accessed 26 June 2021

[15] http://waynesiegel.dk/wp-content/uploads/2013/08/Movement_Study.pdf Accessed 26 June 2021

Link







Link

https://www.gwangyulee.com/2022/11/decimalinteractive-audio-visualization.html