Karol Augustyniak
HSLU – MUSIK (MA COMPOSITION)
Algorithmic Composition

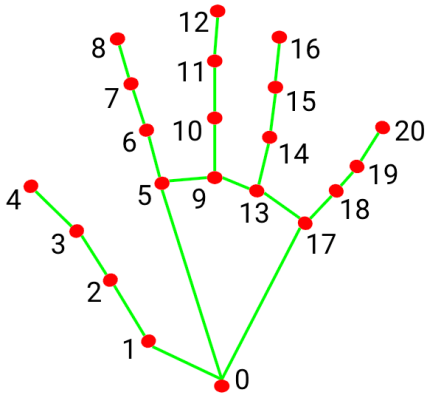
MOTION TO MUSIC

1. Introduction

Motion to Music is an interdisciplinary project combining music and programming. It integrates different technological environments (including Python, JavaScript, MAX and OBS with Virtual Camera). The project was created to serve as an enhancement of an existing piece of music titled <<Anima>>. The piece was composed for flute and audio playback. Motion to Music adds live electronics and live video up to the total by means of the grain delay VST plugin. The reason for naming the project <<Motion to Music>> is that live audio processing is dependent on the flutist's hands movement. By using A.I. technology a performer's hands are tracked analysing the coordinates of the position of the hands. All the materials along with the code and patch can be found here: https://github.com/poko09/motion_to_music

2. Technical Background

The core of the project was developed using the Python programming language. The hand tracking system was implemented using the MediaPipe library[1]. MediaPipe is an open-source, cross-platform framework developed by Google for building multimodal machine learning pipelines. It offers several pre-built solutions for common machine learning tasks, such as Face Detection, Pose Detection, and Hand Tracking.



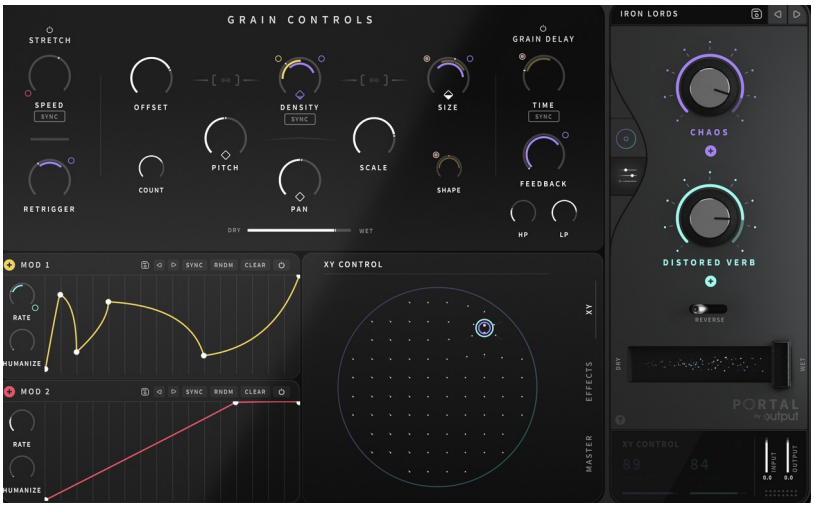
0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

As an example of the operation of this library, the picture above demonstrates how a pre-trained AI model detects 21 landmarks responsible for identifying muscles in the hand. Each landmark provides

[1] https://pypi.org/project/mediapipe/ (entrance date: 20.06.2024)

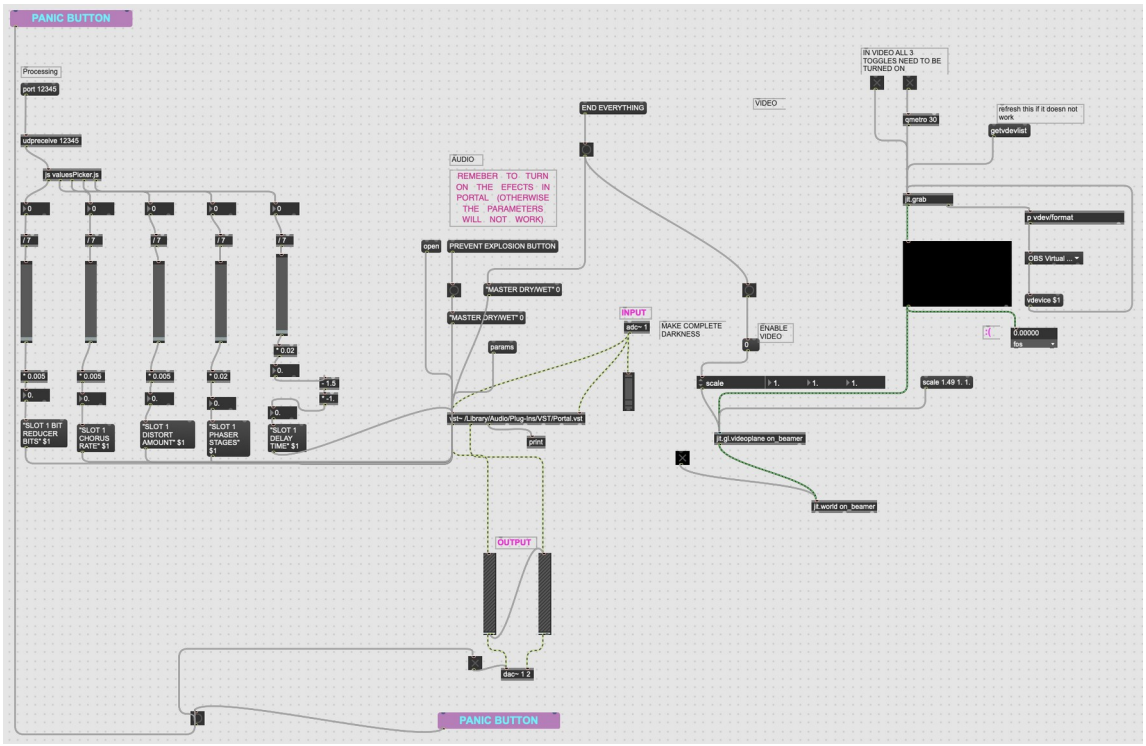
the (x, y, z) coordinates of specific points on the hand. These coordinates are relative to the input image or video frame, with (x, y) representing the pixel positions and z representing the depth (distance from the camera). In the Motion to Music project, only the horizontal and vertical values were used. These figures were important in processing the audio in the piece.

The next crucial element in the project was the usage of Open Sound Control (OSC), which is a communication protocol designed for networking sound synthesizers, computers, and other multimedia devices. The key features of OSC are high performance, leveraging the UDP protocol, and compatibility. It was necessary to have reliable and stable data transmission. Data collected through a Python application was then sent via UDP to MAX. In a MAX patch the numbers were unpacked with the help of a script written in a JS file - valuesPicker.js. Subsequently the data was standardised and forwarded to the chosen parameters of a Portal VST plugin[2]. Parameters selected for audio processing are bit reducer, chorus, distortion, phaser, and delay time. The Portal VST plugin, developed by Output, is an advanced granular synthesis effect plugin designed for transforming audio in creative and unique ways. Unlike traditional granular synthesis, which can often produce abstract and experimental sounds, Portal resynthesizes sounds and maintains their original musicality, making it suitable for a wide range of applications from subtle enhancements to drastic transformations.



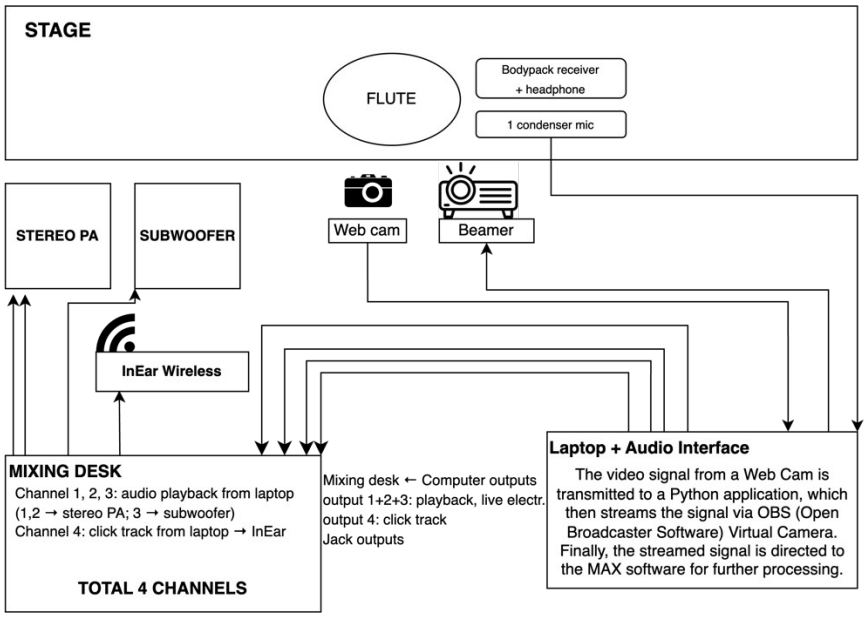
The image above illustrates the functionality and effects of Portal. The decision to use this VST plugin stems from the inherent difficulty in controlling parameters altered by hand movements, which often result in chaotic outcomes. Portal's grain delay effect is particularly effective in harnessing this chaos, transforming it into musically coherent and intriguing results.

[2] https://output.com/products/portal?utm_source=google&utm_medium=search&utm_campaign=Perpetual-Acquire-SER-Branded-Buys-Prospect-INTL-Brand&utm_term=portal%20plugin&utm_content=126088242387&gad_source=1&gclid=Cj0KCQjw1qO0BhDwARIsANfnkv8bYTno27eC8SqUE7SYAKMYoFS6UmcNSgsMS4uxoS6FL4EiJeS0e_waArgtEALw_wcB (entrance date: 01.07.2024)



The screenshot above illustrates the functionality of the piece. The left and middle sections handle audio processing, while the right section manages video transmission and reception. As <<ANIMA>> is a multimedia piece, the visual components are equally important. The output of a Python application, which records the flutist's hands and marks red dotted points, is transmitted via OBS Virtual Camera and then forwarded to MAX to be projected through a beamer.

The final technical setup for the piece is depicted in the chart below. It comprehensively maps both the visual and audio connections, showcasing how different components interact.



3. Conclusion

Finally, the project was performed in Luzern during the New Music Days concert. The audio processing was applied in the middle part of the piece, where the sound of the live electronics was dependent on the movement of the flutist's hands. Both the sound and the visual component, projected via beamer, were synchronized and presented to the audience in the middle part of the piece. While the motion tracking system was activated, the flutist improvised within given limitations and a range of notes, which were elaborated during rehearsals.

The most relevant problem for the piece was the data flow and synchronization between all the software mentioned in the previous section. Due to the different program specifications, a significant delay in projecting the visuals was noticed. The number of frames per second fluctuated between 90 and 120. Additionally, the project was performed on a 2021 M1 16 GB RAM MacBook Pro, which was not an appropriate computer for the complexity of the project. Thus, the live-video part was not perfect and should be improved.

The technical aspect of the piece was challenging for my level of experience and skills, and the outcome was somewhat frustrating. However, I enjoyed the process of building my first multimedia project with advanced technology. Indeed, it did not meet my expectations in the end, but I learned a lot and gained the motivation to improve this idea.