MelodyPainter: Draw the Melody in Virtual Reality

Kaiming Cheng
University of Virginia
Charlottesville, VA, USA
kc4jd@virginia.edu

ABSTRACT
The introduction of computer-based musical systems has largely changed our perception of creating music. However, most music creation systems are firmly rooted in existing conceptual models like a keyboard or a digital audio workstation (DAW). Virtual reality applications provide an alternative to existing musical expression methods by making use of the expressive capabilities of motion controls as well as the fluidity of virtual environments.

In this paper, I present MelodyPainter, a virtual reality-based composition software that transforms the user’s motion into a creative melody. It explores new possibilities for musicians to express their musical ideas more intuitively and for people with nonmusical training to improvise their musical ideas freely. This paper also discusses control and programming methodologies as well as the specific mapping algorithm used to link three-dimensional coordinates with musical characteristics.

CCS CONCEPTS
• Human-centered computing → Virtual reality; Haptic devices; Gestural input;

KEYWORDS
Digital audio, Audiovisual interaction, Unity, Design methods

1 INTRODUCTION
Our relationships and perceptions with sound and space are incorporated into a multiplex system, bounded not only by the natural law of physics, but also our personal cognitive understanding of our sound world. In the history of instrument design and performance, there exists a rich embodiment of physical interaction and gesture correlation, from pressing and plucking the tuned strings to blowing air into the flute. Each gesture serves a specific purpose to output a certain frequency or tune quality.

With the introduction of virtual reality, humans are now able to reform their visual and auditory systems into a created space. The rich data produced in virtual reality platforms offer a great opportunity to create an innovative sonic representation to the outside world. We do not have to be limited by the former boundaries, but that also means there is no standard on how interaction and generated sound should relate to each other. In other words, we can map any data generating process to any aspect of sound generation, which leads to the essential question: How should we best map the sound generating processes and the visual elements with the motion and gesture input from the users.

There have been numerous applications and algorithms transforming motion data into music notes,[6–8] but not many on the design principal of mapping motion and sound in virtual space. This paper describes the MelodyPainter, a vr instrument that which combines the gaming engine Unity with audio engine Max/MSP and Ableton for the creation of expressive melodies as well as artworks. I offer a new way of coupling motion and action in virtual spaces to music generation processes by creating and controlling sounds through the mapping of parameters of motion, thus the sonic experience can be closely integrated with the visual elements, creating a successful cognitive immersion experience. I also outline various related approaches to design virtual reality music software, articulate the workflow, the implementation of MelodyPainter and explain the evaluation criteria with results.

2 BACKGROUND
The use of virtual reality engines for music and sound generation has become increasingly common as generations of musicians who have grown up with readily accessible game systems, internet access, and personal computers seek to bring together visually immersive graphical game-worlds, wide-area networks, interactive control methodologies and musical performance systems. Because this project focuses on music composition sandboxes, it is necessary to understand existing virtual reality experiences involving artistic creation and music composition.

2.1 Tilt Brush
Figure 1 shows the interface of Tilt Brush, a 3D painting application for room-scale vr software developed and published by Google.[10] Users are presented with a palette of various brushes and colors with which they can paint in a virtual environment. Movement of the handheld controller in 3D space creates brush strokes that follow in the virtual environment. Users can draw with impossible materials such as fire, snow and audio waveform. While this application is not directly related to music composition, it provides an example of mapping users gesture with different types of visual elements. Based on that, MelodyPainter links users gesture with audio output, which brings a different level of immersion.

Figure 1: Tilt Brush
2.2 UDK-OSC
UDK-OSC is a great approach to link an audio engine to a graphics engine via a network communication protocol such as Open Sound Control.[5] This project is designed by Rob Hamilton and used in many works in the NIME community [1, 6] and largely inspires MelodyPainter. Figure 2 shows the overall work-flow as it was designed to bring together real-time procedural sound synthesis, spatialization and processing techniques from the realm of computer music with the visually immersive networked multi-player environments of commercial grade gaming engines. It sets the foundational mapping schemata between action in virtual space and the generation of musical sound. Gestures, motions and actions generated by actors in game-space are analyzed and transformed in real-time into control messages for complex audio and musical software systems.

![Figure 2: UDKOSC processes OSC input to control avatar and camera motion while generating OSC output representing avatar and skeletal mesh location, rotation and action/state data.](image)

2.3 Chunity
Chunity is a programming environment for the design of interactive audiovisual games combining the capabilities of Unity and ChucK.[1] It embodies an audio-driven approach that integrates audio programming and graphics programming, taking advantage of strongly-timed audio programming features of the ChucK programming language and the real-time graphics engine in Unity. There are a lot of similarities between Chunity and MelodyPainter, but MelodyPainter focuses more on the artistic expression in painting based on the user’s motion.

2.4 Radio-Baton
The Radio-Baton [2] is a controller for live computer music performances. It is a groundbreaking method of controlling computer-synthesized sound through a predominantly wireless three-dimensional interface as Figure 3 shows. It tracks the motion in three-dimensional space, of the ends of two Batons which are held in the hands of a performer. With the Radio Drum’s surface now wired with receiving antennas and two batons (or drumsticks) outfitted with transmitting antennas, location data for the batons could be registered in X (right-to-left), Y (top-to-bottom) and Z (position above the drum surface) coordinates. The Baton was designed to work with MIDI synthesizers and MIDI-based sequencing and programming software. It emancipated the electronically controlled (and this includes computer-controlled) musical instruments from the keyboard metaphor and set the foundation for new mappings of gesture to sound.

![Figure 3: Radio-Baton Interface](image)

3 MELODYPAINTER
Users hold a wireless VR controller, and they can see the virtual world and the controller from the VR headset. As they hold the button on the controller and wave around, the positions of the controller are transferred to the real-time X, Y, Z coordinates and sent back to the computer, just as Figure 4 shows.

![Figure 4: Architecture of MelodyPainter VR System.](image)

3.1 Melody Algorithm
Every time users press the button on the controller, the program is triggered and creates a new game object. It captures the X, Y, Z coordinates of the VR controller. With a linear transformation, these
coordinates are scaled into proper double values that represent pitch, velocity, and duration. These values are stored into a list object named midiData, sent to OSCHandler and passed into MAX/MSP. The tempo is controlled by the framerate speed in Unity. The default render-rate for Unity is 60 frames per second. The user can control this variable by changing the x coordination of the controller. Every time user releases the controller button, the value of nextActionTime resets to 0. There is a default fading function, which makes the game object become transparent after 5 seconds. This explores the relationship between audio-vision and memory as music is a single direction information flow. Figure 5 shows the interface of MelodyPainter.

3.3 Spectromorphology
The concept of Spectromorphology[9] is a tool for describing and analyzing the listening experience. It refers to the interaction between sound spectra and the ways they change and are shaped through time. The concept of spectral space is often analogical: higher pitch generally equals the spatially higher, and fast tempo might equals to a dense representation form. virtual reality music should not be limited to spatial reality and it can juxtapose an impossible experience in itself. This makes MelodyPainter a unique art. Users can determine the corresponding relationships between the direction of their movements and pitch, velocity and duration. They can also decide if they want to have a contiguous or non-contiguous spatial texture by updating the framerate.

3.4 Conceptual Model
In MelodyPainter, users can draw a line in virtual space. The paintings are transformed into a line of melodic notes which plays a generated rhythm. Thus users with little music experience are able to produce interesting melodies just by making expressive motions. Users get instant feedback visually, as they can see what they create in this virtual space immediately. Google’s Tilt Brush uses the concept of painting to help users understand how to use their application, and this system could be viewed as an extension of this concept. People’s conceptions of real musical instruments must also be taken into account. Since musicians must move quickly in order to play more complicated sequences of notes, users probably expect that moving faster creates faster or more complex music. Users might also have a conceptual model of how pitch maps to space. Some users might expect that pitch maps to the y-axis where a smaller y coordinate is a lower pitch, while others might expect it to map to the x-axis.

4 EVALUATION
Erkut et al. discuss an evaluation method for VRMs based on player engagement. This is defined as “the ability of the player to keep interested in playing the instrument for a significant amount of time.” They also suggest evaluation based on agency, which they define as “the ability to take meaningful actions and see the results of one’s decisions and choices” [4]. I evaluate this application on these criteria as well as on the user’s overall enjoyment of their experience. I use the feedback from the participants to determine the characteristics of effective VR music applications with a room-scale VR setup: a Windows Mixed Reality System.

4.1 User Testing Setup
We used Qualtrics to collect the survey responses. The surveys did not ask for any identifying information, so all responses were anonymized. Participants were asked to sign an Institutional Review Board (IRB) consent form before the experiment. They were given a brief questionnaire before the test to learn about their musical background and virtual reality using experience. Participants were then asked to play music using the VR instrument and answer questions about their experience, the mapping algorithm they prefer and leave their comments.
4.2 Testing Result

The pre-questionnaire contained a series of statements where the participants marked 'N/A - Basic Knowledge - Novice - Intermediate - Expert', '0 - One time or two - Several times - Many times - Regular user', or 'Strongly Disagree - Disagree - Neutral - Agree - Strongly Agree'. The questionnaire includes:

- 'Do you often experience nausea when using VR headsets?'
- 'Do you often experience dizziness when using VR headsets?'
- 'Please indicate your proficiency level with your favorite musical instrument.'
- 'Please indicate your proficiency level with your favorite digital audio workstation. (Software like Ableton Live, Reaper, Logic, FL Studio, etc.)'
- 'How many times have you used a roomscale VR system (Oculus Rift / HTC Vive)?'
- 'How many times have you used a VR headset?'

These questions first help us ensure overall safety since some VR applications have a risk of motion sickness or vertigo. They also help us investigate the correlation between the degree of acceptance to MelodyPainter and their previous VR/music experience.

The post-questionnaire contains statements about users experience where the participants marked 'Strongly Disagree - Disagree - Neutral - Agree - Strongly Agree' on the following questions. I codified these responses to represent numbers 1 through 5.

- 'Do you feel immersed in the virtual environment?'
- 'Do you feel that you are free to produce the sounds that you wanted to produce?'
- 'Are you satisfied with the result of your improvisations?'
- 'Do you enjoy the experience?'

Also, participants are asked to describe which mapping strategy they feel more comfortable with:

- 'I prefer using x-coordinate (move horizontally) to determine pitch'
- 'I prefer using y-coordinate (move vertically) to determine pitch'
- 'I prefer using z-coordinate (move inwardly) to determine pitch'

4.3 User Feedback

Many users found it satisfying to play MelodyPainter:

- 'This project is amazing!! Would even be cooler if one could change the color / texture of brush.'
- 'I liked it a lot, it felt smooth and fluid – the sound with the motions'
- 'It has the potential for a whole stage type set-up.'

Meanwhile, other users express their expectations for future developments and virtual reality in general.

- 'Need to see how it operates with user selection of sounds'
- 'I felt like I need more precise control on the note I am creating.'
- 'Would be better if I can control the articulation.'

Based on the feedback, 10% of the participants felt dizziness before when using VR headset, 80% of the participants have a music training background and all of the participants played with virtual reality before. The mean score for post-questions is 3.93 with a variance of 0.46 given 5 is the maximum score. These questions were not intended as a quantitative assessment, but rather as another useful qualitative way to gauge how people felt in using the tool.

40% of the participants pick x-coordinate as their intuitive prefer method for changing pitch, whereas 60% of the participants pick y-coordinate.

Ultimately, users seemed empowered by this tool. At the same time, it is clear that much can be improved both in terms of existing features and in terms of making MelodyPainter more satisfying to use.

5 CONCLUSION AND FUTURE WORK

For this research, I created not only a new VR music instrument but also to understand its implications. Since all instruments encourage particular ways of playing resulting from the physical attributes, my evaluation attempted to better understand the ways of working as MelodyPainter suggests. However, understanding the ways of working is not straightforward, as it is not always easily reducible to its constituent parts. Such understanding involves the overall aesthetics of using the tool, what it allows one to do, and the manners in which it suggests itself, as well as the domain(s) it engages with.

The interactive audiovisual design is an inherently complex and messy domain. It entails working simultaneously with interaction, sound, and graphics to create a single coherent experience. It asks the designer to reconcile their conceptions and intentions with the idiosyncrasies of the underlying tools, while working with two different programming paradigms.

As a programming paradigm, MelodyPainter encourages ways of working that mesh well with its conception as a state-of-the-art art engine. It’s work-flow, while complex, has successfully make connections between MAX, Unity and Ableton possible. Meanwhile, MelodyPainter provides a specific way of thinking about time and concurrency as musical constructs. It explores the relationship between audio-vision and memory as music is a piece of single direction information, and inspires new audio creativity with the real-time visual cue.

In thinking critically about MelodyPainter as such a hybrid tool, I have observed both limitations and useful affordances not found elsewhere. The inherent tension and sense of complexity in mixing two disparate paradigms (e.g., graphics vs. audio; C# vs. digital audio workstation) are evident in the users’ feedback.

Future work may consider to create a collaborative mode, which adds the potential for an ensemble level performance. People will be able to play this VR software and create art simultaneously across the world. It would be desirable to conduct a user testing on people with non-vr experience, to see how they react to this new interface of musical performance. It would also be of interest to improve the melody algorithm based on more parameters, for example the acceleration speed of the controller. Audio examples related to this paper are available online at https://soundcloud.com/kaisupreme/sets. The example includes several composition ideas using different instrument packages from Ableton.
6 ACKNOWLEDGMENTS

This research is supervised by Professor Mark Sherriff and Professor Leah Reid.

REFERENCES

[10] Ungerleider, N. Google’s tilt brush is the first great vr app. 2016.

A APPENDIX

Figure 7: User’s reaction on MelodyPainter during NIME 2018