

Cheong Un Lee

SEEING SOUND

How to generate visual artworks by analysing a music track and representing it in terms of emotion analysis and musical features?

Thesis Supervisor

Matti Niinimäki

Thesis Advisor

Nuno Correia, Matti Niinimäki

NEW MEDIA DESIGN AND PRODUCTION

MEDIA LAB HELSINKI

SCHOOL OF ARTS, DESIGN AND ARCHITECTURE

AALTO UNIVERSITY

2020

Acknowledgements

I wish to express my sincere appreciation to my supervisor, Matti Niinimäki, and advisor, Nuno Correia. Without their persistent help, the goal of this project would not have been realized. I want to say "thank you" to Pipsa Asiala, who encouraged me to initiate this thesis project in her production clinic course. Also, I sincerely appreciate the following people for their support: Philip Dean, Markku Reunanen, Rasmus Vuori, Teemu Leinonen, Koray Tahiroğlu, Rupesh Vyas, Zach Dodson and David Hall.

I want to say thank you to my friends at Aalto University. When I was an exchange student, you guys were supporting and encouraging me. Without your help, I could not imagine finishing my study and thesis. Special thanks to Thu Nguyen and Hanna Årström. I must thank my lovely wife Sae Young Lee. Without her support (mental, motivational, and financial), I may never have completed this thesis.

Finally, I would like to thank the Lord, my Father, for guiding me and being with me throughout this journey.

Author Cheong Un Lee

Title of thesis SEEING SOUND

“How to generate visual artworks by analysing a music track and representing it in terms of emotion analysis and musical features?”

Department Media

Degree programme New Media Design and Production

Year 2020

Number of pages 67

Language English

Abstract

Music and visual artwork are a valuable part of our daily life. Since both media induce human emotion, this thesis demonstrates how to convert music into visual artwork such as generative art. Especially, the project shows the method of connecting music emotion to the theme of colour. This thesis describes the human emotional model based on arousal and valence. Also, this thesis explains how colour affects our emotion. In order to connect music emotion into the colour theme, this thesis shows the method to retrieve music information which includes arousal and valence of the music. In order to generate visual artwork from the music, this thesis demonstrates the implementation of working software that integrates music emotion and musical characteristics such as frequency analysis. Besides, this thesis presents how to apply generative artwork into our daily life products. This thesis discusses learning outcomes from the project based on practice-based research methodology. Also, this thesis introduces a further plan related to AI.

Keywords generative art, music visualisation, emotion in music, processing, openframeworks, FFT analysis

Chapter 1. Introduction	5
1-1. Research Question	6
1-2. Motivation	7
1-3. Contents of the Thesis	10
Chapter 2. Literature Review	11
2-1. Emotional Model	12
2-2. Elements of Music	13
2-3. MIR(music information retrieval)	14
2-4. Colour and Emotions	16
Chapter 3. Related Work	19
3-1. Generative Art and Relevant Practitioners.	19
3-2. Music Visualization	27
Chapter 4. System Design and Implementation	31
4-1. Methodology	31
4-2. The initial idea and prototype stage	33
4-2-1. Real-time FFT analysis data visualisation	33
4-2-2. Representation of a song by a line.	36
4-2-3. Representation of a song by a curved line drawing	37
4-2-4. Representation of a song by a wavy line drawing	38
4-2-4. Applying a colour for the drawings	40
4-3. First Iteration	42
4-3-1. Visualisation Design	43
4-3-2. Software Implementation	44
4-3-3. Exhibition	47
4-4. Second Iteration	48
4-4-1. Visualisation Design	48
4-4-2. Software Implementation	51
4-5. Example applications	54
4-5-1. Generative artwork in daily life	54
4-5-2. Visualisation of Spotify music chart	55
Chapter 5. Discussion	56
Chapter 6. Conclusion and Future Plans	58
Appendix	60
References	65

Chapter 1. Introduction

Music and visual factors are the fundamental art pieces in our daily life. Through improvements with Internet technology and data compression algorithms, we can listen to high-quality music with portable devices like MP3 players, and smart mobile devices. Indeed, music exists everywhere in our daily life. For instance, many people listen to music while they are running, exercising and even studying. The main reason why music is valuable for our daily life is that music can induce emotions, and it affects the listener [1]. Because, in music, the variation of rhythm, melody, pitch, harmony, and other elements of musical structure is continuously varied. For instance, particular rhythm and beat are known elements that affect body rhythms and movements directly [1]. Visual factors also stimulate our emotions effectively. Because, like music, the fundamental elements of images such as colour, shape, and light induce our emotion [2]. For instance, colour always exists in our daily lives, whether we perceive it or not every single second. It is widely studied that colour is the most important factor to stimulate our emotions and feelings [3]. For instance, the red colour induces a sense of excitement, and the blue colour is related to comfort and security [3].

Music and image have a common characteristic that induces our emotions. When music is being described in a graphical representation, it is called music visualisation[3]. Music visualisation is related to the broader field of information visualisation, which is a way of representation of data to support people in thinking and communication. Typically, maps or diagrams are the most popular examples of data visualisation [3]. Information visualization has also been applied to art. A common approach is to use generative techniques. The term generative art describes any art practice which is created by an autonomous system such as a computer program written by the artists [4].

The study of emotional representation in music has been already studied since ancient Greece [5]. Moreover, in modern research, there have been various studies

analysing emotions in music based on numerical representation. A known emotional model has been adopted widely. This model is Russell's 'A circumplex model of affect', which explains emotions in a two-dimensional coordinate plane. In this model, for classifying emotions in numerical representation, there are two-axis on the plane, and each axis represents valence and arousal [6]. To obtain a numerical representation of emotions in musical contents, music Information Retrieval (MIR) techniques can be used. By using MIR, we could retrieve emotional values (such as happy, angry, sad, and relaxed) from music [7].

This thesis describes a method of music visualisation based on generative art. The visualisation consists of two main categories which are the theme of colour and the shape of the drawing. In this thesis project, I would like to demonstrate how to generate an image that is created by musical factors such as frequency analysis, and the colour of the image is mapping to the emotions in the music. The literature review chapter that follows shows what is the current technology to analyse emotion in music, and presents generative art.

1-1. Research Question

In modern life, it is easy to find information visualisation such as weather forecasts and stock market statistics. Instead of showing just numbers, visualising data is an efficient way to describe the characteristics of data [3]. It is also used in science to understand a phenomenon. However, in the past decade, another type of visualisation method has been widely developed. It consists of artworks based on data visualisation; for instance, these artworks based on data have been introduced at institutions like Whitney Museum of American Art and the San Francisco Museum of Modern Art [8]. In this respect, visualising music is a practical example of artworks based on data visualisation that is very important for this thesis project. However, much research and many projects based on music analysis are focused on how to

create meaningful data from music. For instance, Bergstrom et al. made software to visualise the structure of music; it reads MIDI (Musical Instrument Digital Interface) and visualises using triangular images shown in figure 1 [9]. However, they did not use colour to express the characteristics of the structure of music. This brings me to the main research question:

How to generate visual artworks by analysing a music track and representing it in terms of emotion analysis and musical features?

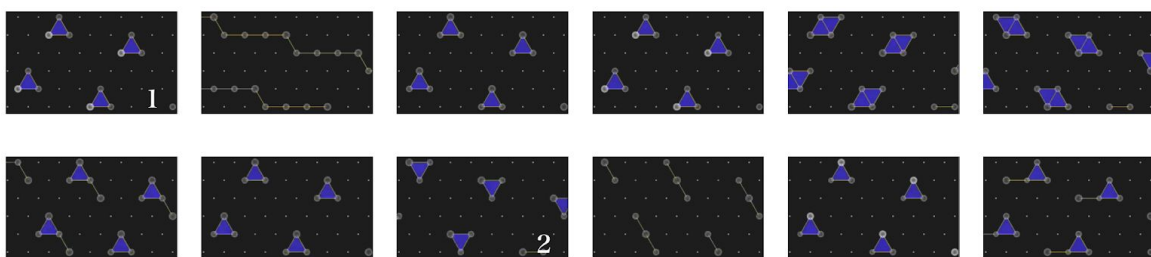


Figure 1: The first measures of J.S. Bach's chorale, Warum sollt ich mich denn grämen [4]

1-2. Motivation

When I was a child, I was always surrounded by music because my parents like to listen to music. I remember that I was very interested in the VU (volume unit meter) on the audio amplifier that included a needle that reacts to a piece of music [4]. It seemed like it would respond to the current volume of the music. That movement

gave me a remarkable impression of dynamic visuals related to the music while I was listening to it. I began to realize people can see the music.

I worked for more than ten years for the semiconductor and LCD industry as a software developer. A few years ago, by chance, I came across a new media art installation, and I was shocked at how the installation gave me a strong impression. Since that time, I joined a design school in Korea and studied subjects related to design such as typography, form, colour, and design thinking, etc. In my first exhibition at the school, I made a 3D sculpture that represents the voice of 'hello'. With my exhibition, the audience could see and touch the 3D sculpture of the voice. I think the reason why I made it was related to my childhood experience with music and volume unit meter.

In 2017, I decided to go to Aalto University as an exchange student, specifically wanting to study abroad because I knew there would be a great opportunity to extend my artistic experience with diverse people from all over the world. I was supposed to stay in Helsinki for one semester. However, I tried to study one more semester because the curriculum of Media Lab was very impressive. There was no limitation for the project topic and idea, and I was able to do whatever I wanted, which is very different from my previous school. Hence, I decided to apply for a master's degree programme in the Media department. Eventually, I was accepted into Aalto's New Media Design and Production programme.

At Aalto University, I joined various courses related to generative art, sound synthesis, and interactive art. Naturally, I began to integrate sound with generative art. For example, I made software that has many particles moving according to the frequency of the music in real-time that gave a sense of visual dynamics to the audience. Many sound visualizations are based on real-time. It means that the visuals react to the sound while it is playing. One day, a new idea for sound visualization just crossed my mind. The idea was about drawing an image by the music itself because once the music play is done, then visualization is invisible any more. Hence, in 2019, I made software to read the frequency data of music, and it

draws a generative line, with curves based on the frequency. Of course, each music has a unique frequency itself. Therefore the final results differed although I used the same algorithm in my software. For instance, rhythmical music creates a dynamic pattern in an image. On the other hand, soft ballad music generated smooth lines, curves on an image.

Furthermore, I made 3-dimensional sculptures of the songs with size 4cm * 10cm. I also made a small audio device, when the audience put the 3D sculpture on top of the audio then the audience could listen to the music which was represented by the 3D sculpture. Eventually, I did my first solo exhibition at Aalto University for this project. It was a great moment when I could show my artistic results to many people at Aalto University.

During my exhibition, the audience asked me about the selection of the colour of my results, and I answered, 'The colour is selected by myself'. With these questions, one thing I realized was that the audience was interested in the colour; hence, I had a plan to improve my project regarding the colour selection. While a visual representation of music could convey the dynamics of the music, the colour is also touching our emotions and feelings. Hence, I searched for a way to analyze the emotions from a piece of music. I have found that there are many ways to do that, for instance, I could use an open-source library (Essentia) to read musical emotions. Also, Spotify provides a web API (Application Programming Interface) to retrieve the musical features. In my thesis project, to create a music feature retrieval library is not the main topic, so, I used the Spotify web API to retrieve the emotion in music. I am expecting that the colours chosen by emotions in a piece of music could bring more emotional impression to the audiences.

There is a more artistic music visualisation that expresses musical characteristics during the music playing, and it is broadly used for live music performance or computer software like iTunes. However, this way of visualisation depends on the music data from a specific time while the music is playing. Hence every single frame

of images is volatile. Also, there is no clear information about how the colours are chosen.

I have seen various music visualisation projects and researches, what I found is that most of the music visualisation is focused on either expressing music data in a non-artistic or entertaining way. Especially emotional analysis is missing, which is related to colour representation. Because emotion is the common element in both music and visual, the emotion could be expressed in colour. So, colour is an essential element when music is translated into visuals. In this thesis project, the emotions in music determine the colour scheme, and the shapes are decided based on musical features within music data. For instance, the musical features could be rhythm, melody, tempo and frequency analysis, while the emotional information in music determines the colour of the picture. In addition, musical information such as frequency data sets the shape of the visualisation. Hence, I am expecting that this project could bring another way of music visualisation based on both emotions and musical features.

1-3. Contents of the Thesis

The thesis consists as follows:

Chapter 1. Introduction, includes motivation, research questions and the contents of this thesis.

Chapter 2. Literature Review, begins with an overview of the emotional model that is used for selecting the colour of this project. It then continues to explain the musical features that are used to draw a visualization of music. This chapter contains how to retrieve musical elements with MIR (Music Information Retrieval). It continues to describe the examples of existing libraries of MIR and what sort of musical features

could be extracted from the libraries. Also, this chapter describes the relationship between colour and emotions.

Chapter 3. Related Work, this chapter describes generative art and also shows examples of generative art projects by known artists and their works. Finally, this chapter explains the term of music visualization, history of it and examples of visualization projects.

Chapter 4. System Implementation, demonstrates the detailed information of the thesis project, for instance, software development tool explanation, system structure, class structure, and the final results.

Chapter 5. Discussion, reflects on the overall thesis project, in particular the emotional relationship between the colour and music, based on the result of the survey from Chapter 4. Also, it explains the learning outcomes of this thesis project.

Chapter 6. Future plans and Conclusion, reflects on how to improve the project for the future.

Chapter 2. Literature Review

This chapter describes the literature review of the thesis project. First, it shows Russell and Thayer's emotion model based on valence arousal. Secondly, it explains the elements of music in detail and how they contribute to our emotions while listening to music, and how to retrieve emotions in music based on the music features. The last part of this chapter describes how colour is linked to the emotional response of people.

2-1. Emotional Model

In 1980, Russell proposed the circumplex model of emotion, as shown in figure 2, which is a circular structure including two axes. In this 2-dimensional model, the horizontal axis represents pleasant (valence) feeling, and the vertical axis is for intense (arousal) emotions [5]. For instance, pleasant emotion is increased when the value of the horizontal axis is located on the right side of this model. The anger emotion is located on the left side of the horizontal axis, and it is more intense on the vertical axis. In this proposal, Russell insisted that emotion could be translated for the result of our complex perception. However, he used an adjective to describe emotions in his model, and there are some overlapped similar words with each other and ambiguous expressions in the adjective.

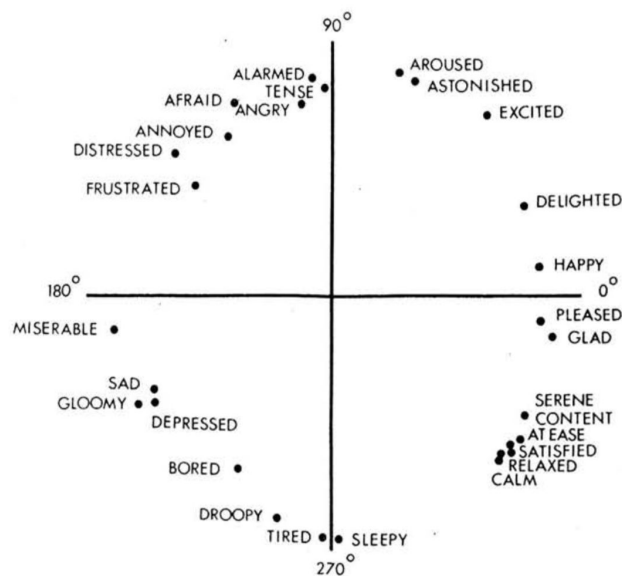


Figure 2: Russell's circumplex model [6]

By contrast, in 1989, Thayer proposed a model of mood [6] shown in figure 3. In this model, the horizontal axis represents positive and negative emotions, whereas the

vertical axis represents the energetic and silent emotion [7]. There have been various research projects based on these models to retrieve music information [8].

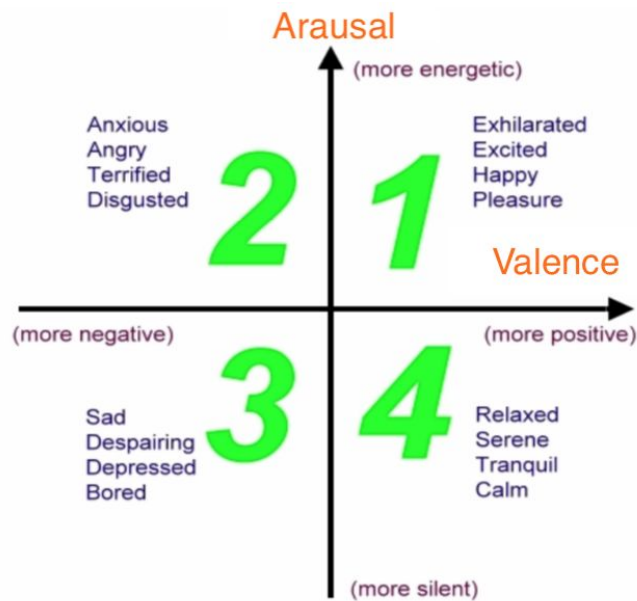


Figure 3: Thayer's model of mood [9]

2-2. Elements of Music

Emotion in music could be described as a combination of valence and arousal, and various emotion perceptions are frequently correlated to different patterns of acoustic cues [10]. For instance, high arousal is related to the emotions of anger, agitation, restlessness, violence, and low arousal is connected to the emotions of relaxing, gentle, soothing, tranquil [11]. The arousal is related to these elements of music which are tempo (fast/slow), loudness (high/low), pitch (high/low) and timbre (bright/soft). The valence is related to mode (major/minor) and harmony (consonant/dissonant). For instance, to propose positive valence, loud chords and high-pitched chords are more relevant than soft chords and low-pitched chords [2]. Hence, the combination of positive valence and low arousal may activate a feeling of

safety [10]. There are core music elements related to emotions such as rhythm, dynamics, melody and timbre.

Rhythm is the element of music related to specific patterns in time. It is expressed in terms of tempo, meter, or phrasing. While we are listening to music, there is a constant pulse or tempo that is what we mentally recognise into meter [12]. Tempo in music is linked to arousal and valence. For instance, high arousal is easily perceived by a piece of music with a fast tempo. However, the firm tempo is related to negative valence, while the smooth flowing tempo is commonly associated with positive valence [10]. Dynamics, often called energy, is strongly related to the perception of arousal [10]. It could be included in very soft musical passages, as well as loud passages, and is expressed as large or wide dynamics range. Dynamics can vary rapidly or gradually, such as crescendo or decrescendo [12]. Melody is often called musical line, it is a composite of pitch and rhythm and it is treated to be the main theme of music composition. Melody could be described in a variation of its contour (rising or falling) and the size of the intervals of it [12]. Timbre is often called sound quality or tone colour; timbre is an essential element to recognise the differences between one instrument and another [12]. For instance, if you sing the note "C" right after playing a "C" on the piano, the pitch of these two "C" produces the same pitch. However, you could easily recognize there is a quality difference between your voice and the piano.

2-3. MIR(music information retrieval)

The music industry has grown rapidly in the past several years through improved techniques in both software and hardware. We have access to a large number of music collections which we never had before. Consequently, various research on music information retrieval has been done [6]. This section introduces practical examples of MIR. Dmitry et al. developed an open-source C++ library called Essentia

to analyse audio-based music information retrieval under the Affero GPL license. This library contains a large set of reusable algorithms, which provide various statistical characterisations of data and high-level music descriptors in real-time of music source. For instance, Essentia has a rhythm descriptor which detects tempo changes, the statistics of the BPM histogram of a song and tonal descriptor gives pitch salience function [23]. Eventually, these elements are strongly related to detecting arousal value based on Thayer's model.

The tonal descriptor provides the harmonic pitch-class profile (HPCP) and the sequence of codes present in a song, and these are related to valence. The library supports multi-platform operating systems such as Linux, Mac OS X, and Windows. Also, there is a wrapper library of Essentia for openFrameworks as well [24]. Chen et al. proposed a music emotion detecting system that is based on a deep Gaussian process. The system provides two major parts: feature extraction and classification. In the feature extraction part, there are five types of features selected for representing the music signal, which are rhythm, dynamics, timbre, pitch and tonality. In the classification part, a deep GP (Gaussian process) is utilized to recognize emotion [12]. In order to extract the acoustical feature from a piece of music, this system used a MIR software called MIR toolbox [13].

Malik et al. studied the emotion recognition of musical tracks in 2-dimensional valence-arousal (V-A) emotional space. They proposed a method integrated with convolutional (CNN) and recurrent neural networks (RNN) for continuous prediction of emotion in V-A model. In this study, they utilized one CNN layer followed by two branches of RNNs trained separately for arousal and valence [27]. There are more music feature extraction tools developed and presented to the community. Moffat et al. evaluated ten existing audio feature extraction libraries such as Aubio, Essentia, Audio, MIR Toolbox, and Meyda, etc. They evaluated these tools with the Cranfield Model and presented example use cases for users to decide which tools are most suitable. Hence, their evaluation allows software engineers or researchers to easily select appropriate music feature extraction tools [28].

Most of the MIR systems which have been described in this chapter analyse music files with formats such as mp3 or wav directly. As an example of MIR implementation, Spotify provides the Spotify Web API for collecting music features. This API is based on a simple REST (Representational State Transfer) principle. The Spotify Web API endpoints return JSON (JavaScript Object Notation) metadata about music artists, albums, and tracks, directly from the Spotify Data Catalogue. One of the important functions in this API is the ability to get music information on a track. There are several meaningful objects provided such as key, mode (major or minor), energy (0.0 ~ 1.0), valence (0.0 ~ 1.0) and more. As described earlier in the emotional model section, Spotify Web API allows the user to analyse emotion in music based on the result of the Spotify Web API. Eventually, energy is equivalent to arousal [29].

This thesis project does not aim to re-develop a library such as Essentia because there are already well-made libraries which are widely used for music information retrieval. Therefore, the thesis project will use the Spotify Web API to retrieve music information.

2-4. Colour and Emotions

In artistic projects, colour is commonly used to affect how the audience feels when they look at the artistic works. It means that the colour is not only used to reveal the fact of an object, for instance like the tree is green, but also the colour is related to the emotional response of the audience [14]. The specific relationship between colour and emotions or objects have developed in human history. To understand these relations is an effective way to approach people's hearts because the emotions are much stronger than logical thoughts based on data and figures [15].

The meanings of colour are based on psychological effects, biological conditioning, and cultural developments. For instance, when we see something related to red colour, it could be associated with a warm feeling, or green is connected to nature and energy. Those colours are deeply engraved in our brain because it is visible in our daily life. Also, colour is a significant fact for survival for human history, for instance, we have learned that we should pay attention to bright colour because dangerous animals or plants often have a bright colour [15]. Also, colour preference might be related to cultural aspects. In 1968, Choungourian found that the most preferred colours in the USA were red and blue. However, those colours were not preferred in other cultures [16]. The meaning of red colour is good luck in China, but it is the colour of mourning in South Africa. The colour of mourning is black in Western countries. However, it is white in some East Asian countries [15].

Colour is fundamentally generated by light that is a sort of energy. The scientists have revealed that human psychological experiences could be changed when we perceive different colours. Therefore, colour can induce various forms of human emotion effects by wavelength [17]. For instance, a long-wavelength colour produces stimulus effects; like red is a vivid colour. On the other hand, a colour with short-wavelength has a comforting effect; for instance, a blue colour lowers the blood pressure and pulse. Orange colour makes the feeling of friendliness, pink express softness, and grey indicates professionalism [18]. Claudia Cortes researched a large amount of list of attributes of positive and negative characteristics, shown in figure 5. Shirley Willett insisted on the colour codification of emotions. This model is used as a guide for the critical role of interpreting the emotions of colours shown in Figure 4 [19]. The circle outside contains positive traits, the second circle represents the six emotions, the inner circle has the negative characteristics, and the centre circle represents depression that is blended to all negative traits [17].

Color	Positive trait	Negative trait	Emotion
Red	active, emotional	offensive, embarrassed	anger, love
Orange	ambition	tiring	joy, determination
Yellow	lively, energetic	cautious	fear, Happiness/joy
Green	calm, neutral	greedy, sick	faith, greed
Blue	faithful, traditional	depressed	confident, sadness
Purple	leadership, passive	arrogant, sorrow	introspective, melancholic

Figure 5: Claudia Cortes color extraction [19]



figure 4: Colour codification of emotions [19]

In 2005, Ståhl applied Itten's colour theory to a circular colour model [20]. Her theory was an intuitive and detailed way to link the relationship between various emotions to corresponding colours based on Russell's Circumplex Model of Affect theory[21] shown in Figure 6.

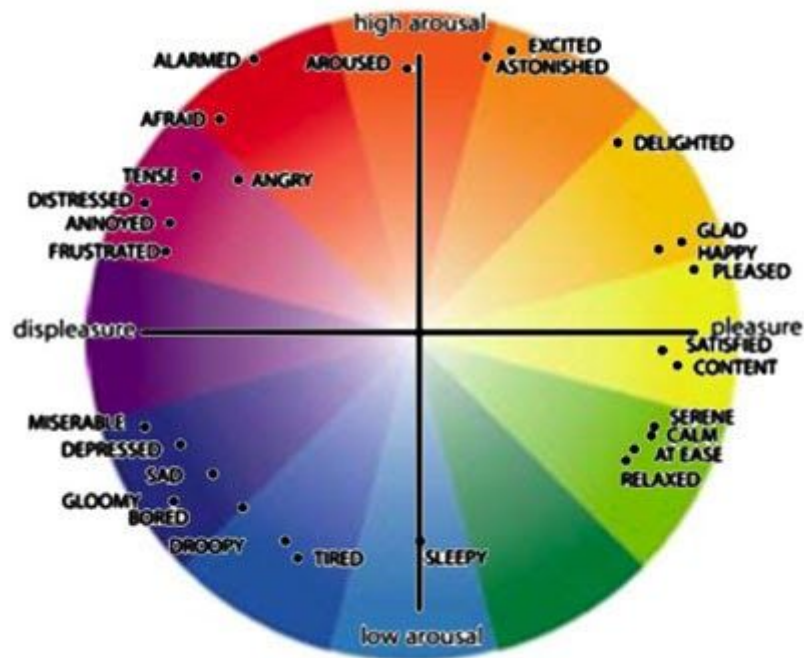


Figure 6: Ståhl's colour wheel fits Russell's Circumplex Model of Affect [21]

Chapter 3. Related Work

3-1. Generative Art and Relevant Practitioners.

The term generative art indicates any art practice which is created by an autonomous system, such as a computer program written by artists [22]. The early phase of generative is often called computer art, and it began around the 1960s. In this early stage of generative art, computers were only allowed to be used by engineers and mathematicians because only large industrial or university laboratories had mainframe computers. In 1965, a graduate student in mathematics at the University of Stuttgart, Frieder Nake, created computer-generated art [23]. This artwork was created using a computer algorithm. The starting point of his algorithm was inspired by Klee's exploration of proportion and the relationship between the vertical and horizontal lines, shown in figure 7. The final production was drawn by a pen plotter. A plotter is a mechanical device controlled by a computer, and it uses a pen or brush, moving in the x and y-axis simultaneously to draw images on paper [24]. Vera Molnár is one of the pioneers of computer and algorithm arts who was interested in the relationship between order and disorder. In 1974, she produced generative art (Dés)Ordres shown in figure 8 and used simple geometric shapes created by the computer algorithm. In order to disrupt the consistency of the squares randomly, she changed the parameters in her algorithm [24]. The final outcome was drawn by the plotter.

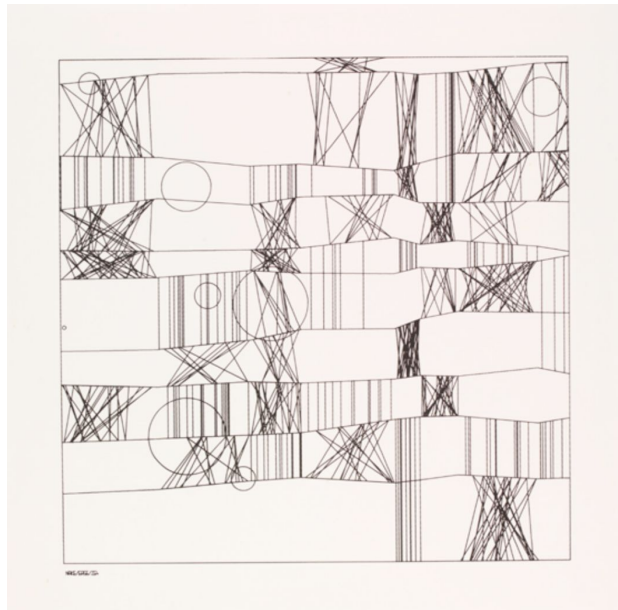


Figure 7: Frieder Nake, 'Hommage à Paul Klee, 13/9/65 Nr.2', 1965.
Screenprint after a plotter drawing. Museum no. E.951-2008.
Given by the American Friends of the V&A through the generosity of Patric Prince [18].

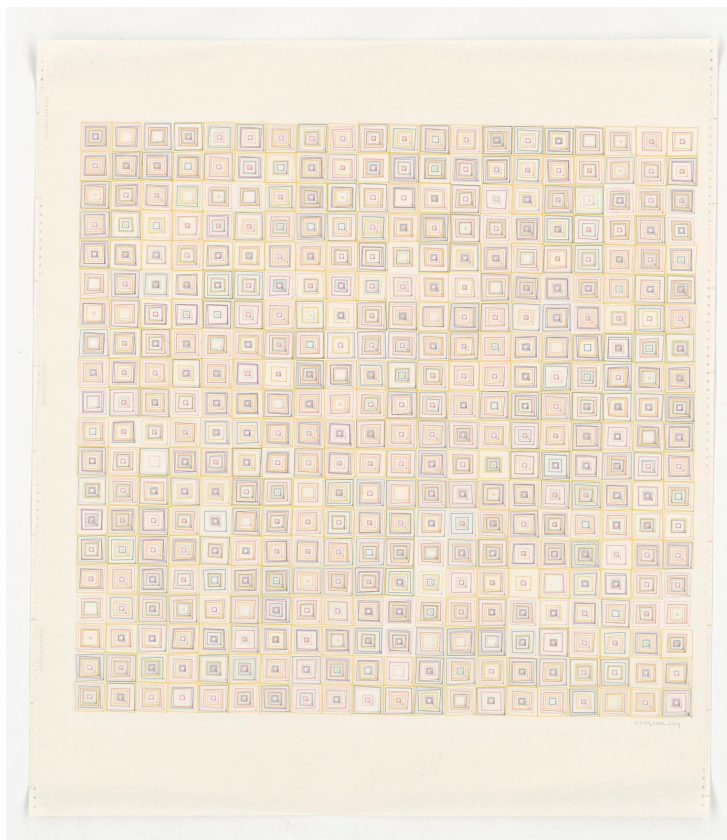


Figure 8. (Dés)Ordres, Molnar Vera, Paris, 1974

Nowadays, the technology of both computer hardware and software has been tremendously improved, compared to the early stage generative art. More people can make a generative art piece with software development tools such as Processing and openFrameworks. Both Processing and openFrameworks were created for media artists and designers to implement their creative projects in visual, audio and interactive media. Processing originally included a minimal set of functions based on the Java language. However, Processing has a library which can be extended by anyone who wants to contribute it to the Processing community [25].

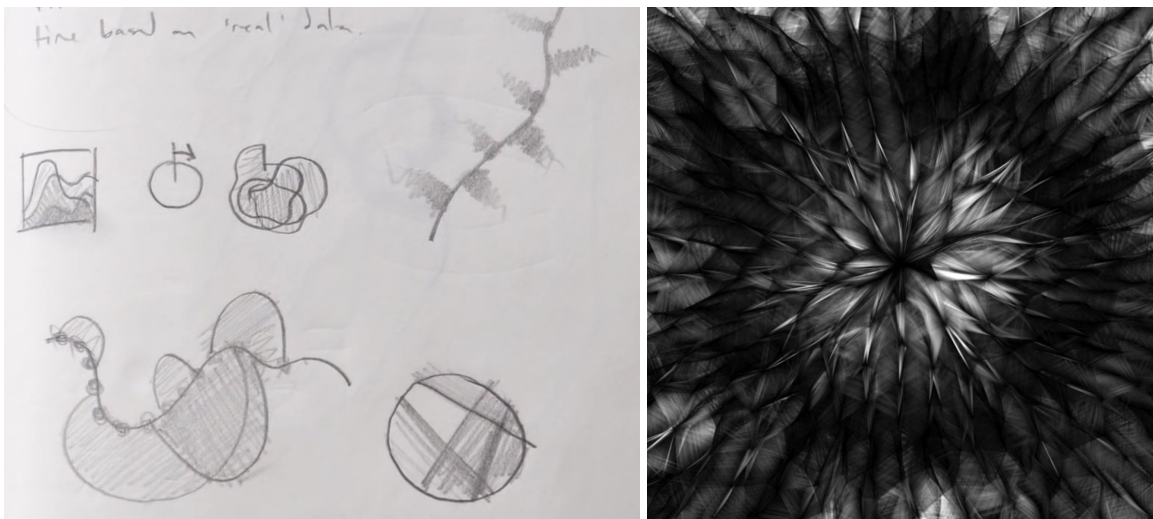
In 2003, Tarbell created images based on known algorithms such as Henon Phase diagrams and with his invented algorithm shown in figure 9. In his artwork, the paintings are reminiscent of urban patterns [26].



Figure 9: Substrate software by Jared Tarbell (2003) [22].

Casey Reas is one of the inventors of Processing. Processing is an open-source programming language that can be used for creating visual artworks through code [27]. Reas' works, consisting of software, prints, and installations, have been highlighted in various solo and group exhibitions at museums and galleries in many locations in the United States, Europe, and Asia [28]. Reas is an artist, educator, and

open-source developer. For him, the software is the most natural medium to work with. The way he works with code is the following: He sketches his idea on paper (example shown in figure 10), then he translates it into computer code. Unlike writing poems, for example, his concept is created by a logical procedure using the computer algorithm. Once the initial image is made, he does many iterations with different algorithms and variables, until it becomes a better image. In his opinion, there is never really a goal, however, there are many different results that could exist.



Reas' idea sketches [25].

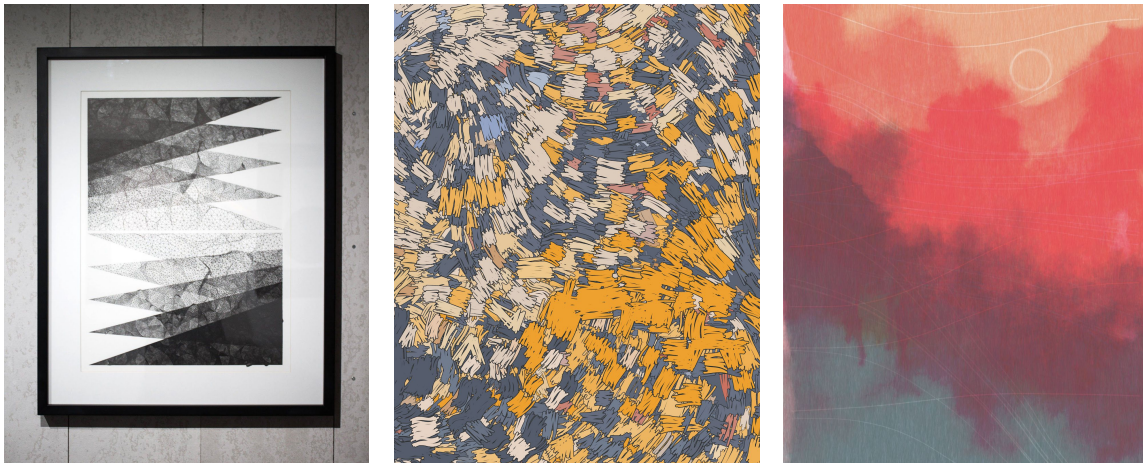
The Process Compendium [29].

Figure 10: Reas' idea sketches and the result

The main topic of his work is emergence. He uses a simple algorithm and things come out which are entirely unexpected output. One of his main concerns is that Processing is used across different fields, allowing people to collaborate on the same software development environment. Also, he has been working with an artist to create project mapping on building walls, which means that he has been moving away from the computer screen to public space [25].

Tyler Hobbs is a generative artist and software developer from Austin, Texas. He created software to paint abstract images. His artworks shown in figure 11 are inspired by painting, vegetation, and naturally appearing patterns. In his artist statement, he reveals that the interplay of randomness and structure is the most

core value of his artwork. He believes generative artwork is especially adapted for an examination in detail.



Turn It Around (Progress 2D) 2018

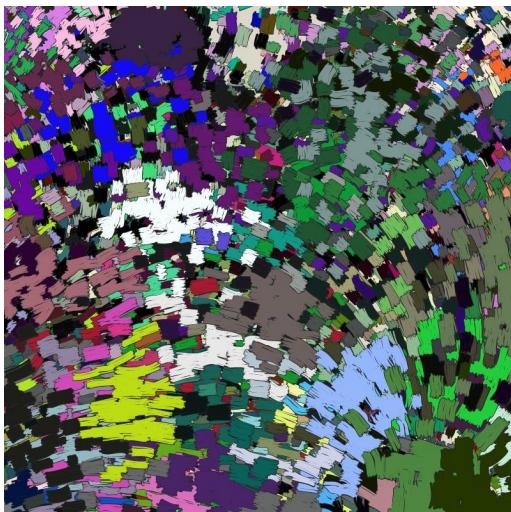
Non-Repeating 2017

Linear II (Welcome Back) 2017

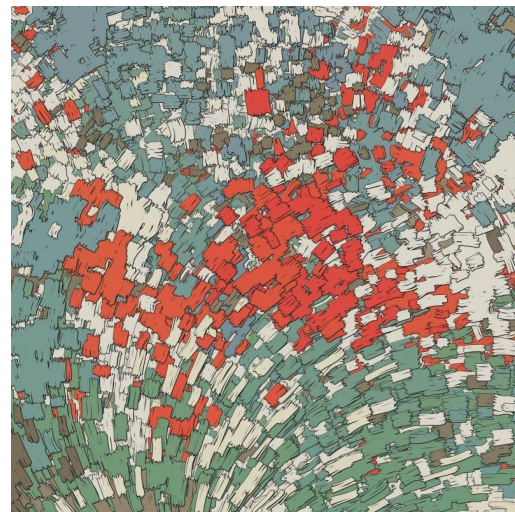
Figure 11: Various of Tyler's artwork

Furthermore, he has written several articles about colour and randomness, which allow improving the result of generative artworks. In generative artwork, he strongly recommends using HSB (Hue, Saturation, Brightness) colours instead of RGB. On the one hand, HSB is closer to human perception, and it is easier to use for creating gradient colour expression. On the other hand, RGB is mainly designed for the machine. Another core advice of using colour is to avoid the random colour of RGB because many practical cases are demonstrated that using random RGB colour looks worse than a colour set which is carefully selected. The example is shown in figure 12. Once programming is done, the artist becomes a curator because there could be hundreds of images which should be manually selected. Hence, randomness allows for a new approach to creating art. Mutability gives freedom of experimentation to artists. It means that there is a low cost to re-generate their artworks. They can try a different style, colour sets, and processes to create digital artworks that are different than the conventional way. In both artwork and software engineering, reuse is widespread. Artists can reuse colour sets, algorithms, composite design and much more. This reusability brings an extra toolset to artists. They can rearrange existing algorithms to another module to generate different

results without solving a problem again and again. The scale is one of the most significant advantages of algorithmic artwork, and artists can increase or decrease the size with small changes in the code. These algorithms also affect the randomness to ensure diversity of other elements of the artwork, for instance, colour sets and resolution [30].



With random colour choices



With hand-picked colours

Figure 12: Random colour example

Zach Lieberman is an artist, educator and software developer based in New York City. He is a well-known generative artist who writes code for creating poetic artworks. Among other things, he is one of the inventors of openFrameworks, an open-source platform based on C++ development, which is a popular tool for generative art, like Processing [30]. He started a process of creating daily sketches on the 1st of January 2016, shown in figure 13. These are small project suites for daily coding that he posts on Instagram. Doing this daily coding, he does not focus on a final goal, but he tries many iterations. It means that he tries to do the same thing again and again until he finds the solution to improve the results. Because he realised that creativity is not just waiting for a light bulb to turn on in his head, it seems to him like meditation. He spent most of his first year of sketching working with blobs, which are essentially particles that repel away from each other, connected by a loop of springs to avoid being mixed together [32]. He created

software to experiment with this blobs behaviour shown in figure 14. Also, his artworks are selling on the internet shown in figure 15, and this is a valuable example of how generative artworks extended to the commercial market.

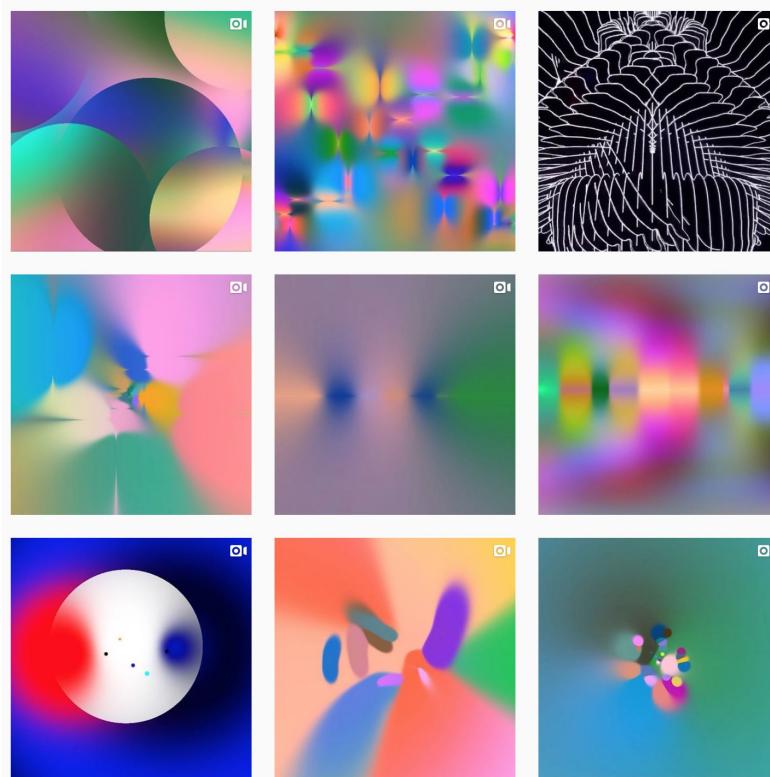


Figure 13: Daily sketches on his Instagram

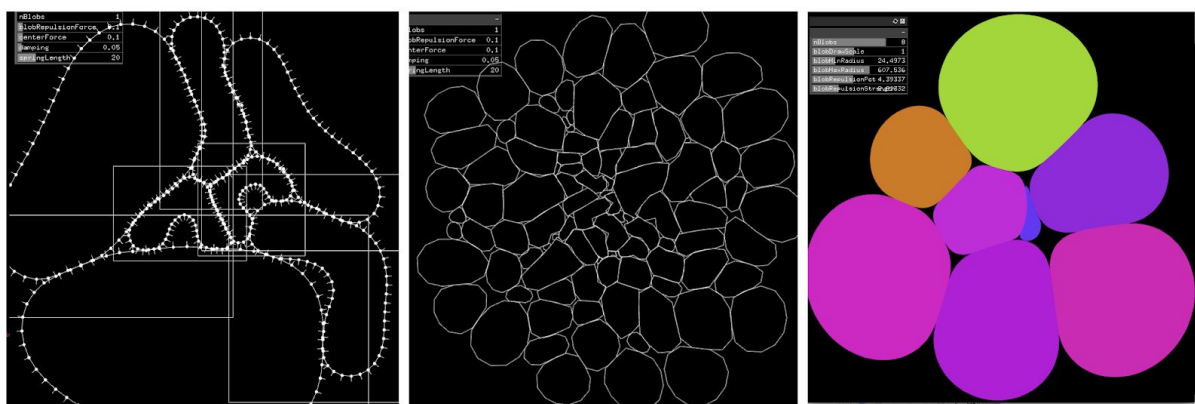


Figure 14. Blobs experimenting tool



extruded blob #1
\$30.00



extruded blob #2
\$30.00



blob family
\$30.00

Figure 15: People can purchase his artworks

Lia is an Austrian artist and educator. Since 1995, she has produced generative artworks, becoming one of the pioneers of software and net art. Her artworks stretch across video, performance, software, installations, sculpture, projections and digital applications. Lia's work is characterized by minimalist quality and affinity with conceptual art, which combines the aesthetics of traditional painting with digital images and algorithms shown in figure 16 [31].

No. 042 (re-built No. 012 with archival ink)
Indian ink on Hahnemühle paper, natural white, 170g/m² | 30cm x 30cm, printed area: ~20cm x ~20cm



No. 041 ●
Indian ink on Hahnemühle paper, natural white, 170g/m² | 30cm x 30cm, printed area: ~20cm x ~20cm



No. 040
Indian ink on Hahnemühle paper, natural white, 170g/m² | 30cm x 30cm, printed area: ~20cm x ~20cm



No. 039 ●
Indian ink on Hahnemühle paper, natural white, 170g/m² | 30cm x 30cm, printed area: ~20cm x ~20cm



No. 038 ●
Pearlescent pigments (Königsgold) on Hahnemühle paper, natural white, 170g/m² | 30cm x 30cm, printed area: ~18cm x ~18cm



No. 037 ●
Pearlescent pigments (Königsgold) on Hahnemühle paper, natural white, 170g/m² | 30cm x 30cm, printed area: ~18cm x ~18cm



Figure 16: Unique plotter drawings, based on code

3-2. Music Visualization

Algorithmic music visualization relates to generative art. Like most generative art, it is done by autonomous systems that use a particular algorithm to control parameters to draw an image. Musical features such as rhythm, frequency analysis, volume, and mode (major, minor) are often used as parameters for music visualization. Music visualization is applied in many different art fields, such as performance (music, opera, dance) and entertainment (musical dramas, movies) [32].

Music visualization has a long history. Ancient Greek philosophers such as Aristotle and Pythagoras hypothesized that there must be an association between sound scale and colour. Around 1730, The Jesuit, Father Louis Bertrand Castel invented an Ocular Harpsichord which had an extra frame above a harpsichord. The frame included many small coloured-glass windows with curtains attached by pulleys to the keys. Hence, each time the player pushes a key, the audience could see the related colour [33], as shown in Figure 17 [34]. In the last few decades, there have been tremendous improvements in technology. Eventually, music visualization is also improved by using techniques such as electronic visualization. The early stage of electronic music visualization emerged for entertainment purposes. In 1976, Atari released a system called 'Atari Video Music (Model C240)' which is the earliest commercial digitized music visualizer, shown in Figure 14. This console device was supposed to connect to a record player and a television set. There were twelve buttons and five nodes, which allowed users to select the custom created visual effects. The effects displayed on the television set were the hallucinatory forerunner of software-based music visualizations. The colour was selected from each source channel. While the music was playing, many visual arrangements in the form of diamond-shaped objects were visualized in real-time [35].



Figure 17: A caricature of Louis-Bertrand Castel's "by Charles Germain de Saint Aubin[36]

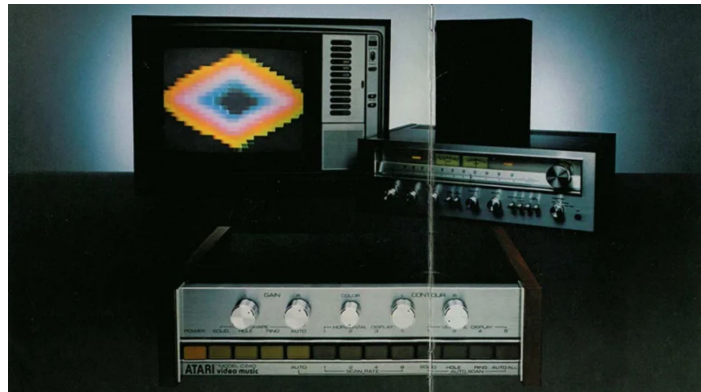


Figure 18: Atari Video Music [34]

Nowadays, there is sophisticated music visualization software that can be seen in computer music players, for instance, in Windows Media Player and iTunes. These players decode digital audio files (Wav or Mp3) and visualize the characteristics of music such as beat and/or frequency data by specific algorithms, as shown in figure 19.

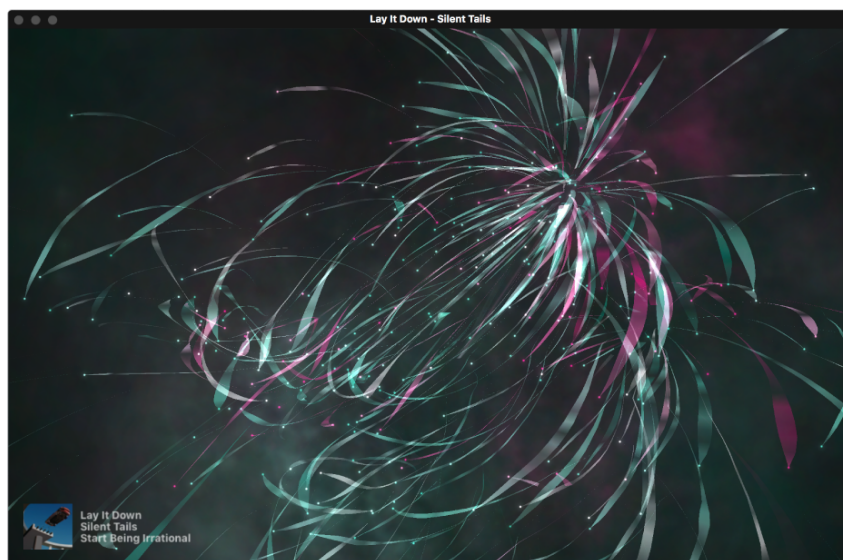


Figure 19: iTunes visualizer [37]

Music visualization can often be seen in contexts outside of computer monitors, such as public art installations. In 2011, Mexican-Canadian electronic artist Rafael Lozano-Hemmer created 'VOICE ARRAY'. This installation project includes many

LEDs that spread in a few meters wide, as shown in Figure 20. When the audience speaks into an intercom, their voice is converted into a signal that flashes the LEDs. Then this particular pattern is saved to the array, and it keeps playing back the recordings in the array. Every new recording is stacked on the new array index [38]. Peter Crawley uses stitched Illustrations to express waveforms of popular songs, and he stitches wave data of a song on a piece of material, such as cardboard. In Figure 21, the song 'Whole Lotta Love' is represented visually, and each horizontal line represents one minute of the song [40]. Canadian artist Nicolas Bernier created audio-visual installations called Frequencies (light quanta). The concept of this project is to create experimental audio-visual installations as shown in Figure 18 based on the idea of relationships with quantum physics and sound, as shown in figure 22. He used 100 transparent acrylic panels which have engraved the smallest energy of frequency by the laser cutter. These acrylic panels show a superimposing in the three-dimensional structure for the pattern of energy. While the sound is playing, each acrylic panel is lit by led lights [42].

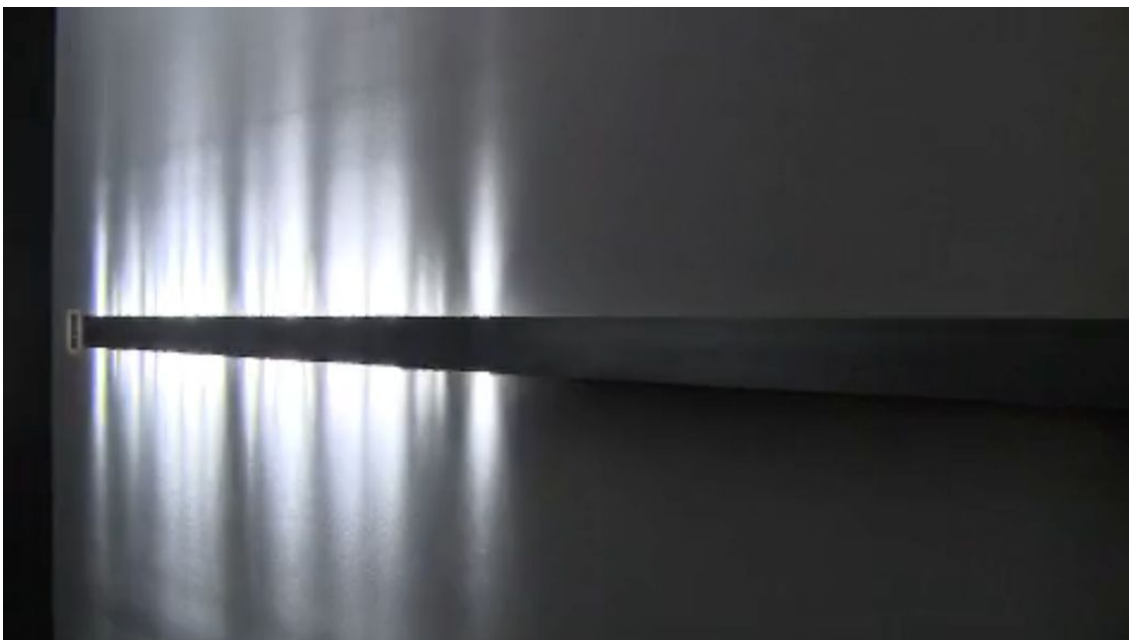


Figure 20: Museum of Contemporary Art, Sydney, Australia, 2011 [39]



figure 21: Visualization of 'Whole Lotta Love'

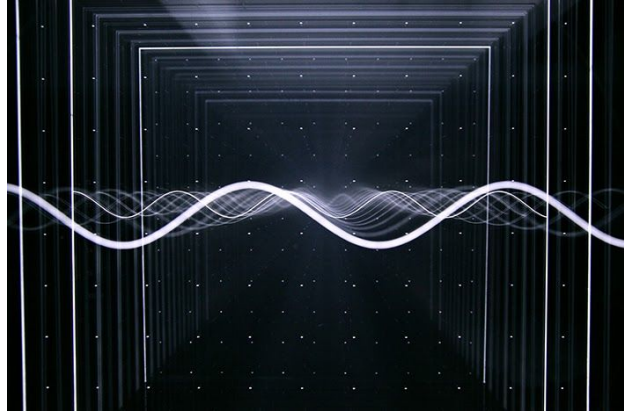


Figure 22: *Frequencies (light quanta)*, 2014 – 2015

Chapter 4. System Design and Implementation

4-1. Methodology

This thesis follows the practice-based research methodology. In order to extend knowledge of a particular discipline and area of creative practice, this methodology reveals how creative work influences it [40]. Candy defines practice-based research as an “investigation undertaken in order to gain new knowledge partly by means of practice and the outcomes of that practice”, and this contribution to knowledge can be demonstrated by creative outcomes such as an artefact [41]. Also, like any other research, practice-based research should include “textual analysis or explanation to support its position and to demonstrate critical reflection” [42]. The result of this thesis project is an artefact, and this thesis includes a supporting document which is demonstrating critical reflection.

This practice-based research project has progressed based on an iterative development process. Each iteration has a few cycles of implementation and improvement. Chapter 4-2 shows the initial idea of how to visualise a song into an image and read real-time data from the song, for instance, to read frequency data from a song by using FFT (fast Fourier transform) analysis. It also reveals the weakness of reading real-time data while the song is playing. In this stage, there were a few iterations processed to make a better visualisation outcome. In order to move to the next step (chapter 4-3), the results were evaluated.

In practice-based research, re-evaluation and refinement of the research question is a common occurrence. Skains mentioned that “there may be significant effects on the composition process of continuing contextual research in theory and creative

works". During the process of this thesis, Skains' quote holds true for this project as well. He also argues that

Practice-based research is given to exploration and significant moments of discovery, which are largely unpredictable at the start of the project. Thus serendipity can lead to new perspectives on the research, reshaping the project goals throughout. [47]

During the implementation of this thesis project, there was no major issue to reconsider the aim of the project. However, there has been a new perspective obtained to refine the necessary functionalities, and this affected the whole iteration process of this thesis project.

This thesis shows the final stages of the music visualisation project, and it serves as the final stage of practice-based research, aiming for a "substantial contextualization of creative work" mentioned by Candy [41]. In addition, Skains argues that in addition to the unexpected serendipitous connections gained during the creative process, argument formation and exegesis are the final steps in the practice-based research methodology [43]. This thesis project follows that stage and provides a practical method and insights to create a music visualiser, and the result of this thesis project was used for an art exhibition (chapter 4-3-3). For future research, Skains argues that for a deeper understanding of creative artefacts and the domain, post-textual analysis and critical research is required [43]. The consideration for future improvement is provided in Chapter 5 and 6 [43].

About the timeline of this thesis project, the first development period was from December 2018 to February 2019. In this period, there were iterations to experiment with visualisation styles, and the initial software architecture was implemented. The first exhibition of this thesis project was on 23 Feb - 03 Mar 2019 at Väre gallery V1, Aalto University. After the exhibition, the project has improved the mappings to the colour theme from emotions in a song. The latest esthetic visualisation outcome of

this thesis project is inspired by the early 1970s generative art (Dés)Ordres by Vera Molnar [44].

4-2. The initial idea and prototype stage

In November 2018, I joined one of the courses at Aalto University called Production Clinic by Pipsa Asiala, who is Producer and Tutor at Aalto University. The main trigger to start this thesis project was preparing for the resulting exhibition, encouraged by her.

4-2-1. Real-time FFT analysis data visualisation

For the first trial, I tried to read FFT analysis data of a song and visualise it with a basic bar chart style in real-time. FFT analysis data is an important method for visualising a song into an image, as shown in figure 23. To acquire FFT analysis data,



Figure 23: real-time Visualisation: FFT analysis(128 bands) data into an image

I used the 'analyze()' function of the FFT class provided by the sound library of Processing. This function calculates the frequency spectrum in every single frame

while the song is playing. The maximum number of frequency bands for the FFT class is 512 as default. However, I used 128 bands for this prototyping stage because 128 bands are enough to visualise the song. I could not find valuable data from more than 128 bands, because each band would otherwise translate into a meaningless small number. Figure 24 shows how to load and play a song and read FFT analysis data. However, there was a problem with Processing source code related to data smoothing. I used FFT analysis directly to draw the bars in every single rendering frame. It created an uncomfortable visual result for the eyes because the speed of variation of each bar is too fast without filling the visual gap between previous and current data. This issue could affect the whole thesis project. Hence, I used the smoothing algorithm “ $\text{sum}[i] += (\text{fft.spectrum}[i] - \text{sum}[i]) * \text{smoothingFactor}$ ” to create a smoother visualisation of the frequency spectrum [48]. This algorithm calculates the difference between current and previous data, and it applies the difference to the next frame data. Eventually, I had a better smoothing visualisation result shown in Figure 25. Of course, a smaller smoothing factor creates a smoother image. However, there is a critical side effect of delayed visualisation for the real-time music visualisation because, for the audience, it is important how much the visualisation is synchronised with music in real-time. If music and visualisation are not synchronised, then the audience could lose their immersion. During this thesis project does not aim to develop real-time music

```
import processing.sound.*;

SoundFile sample;
FFT fft;

int bands = 128;
int barWidth = 10;
float scale = 10;

public void setup() {

    size(1200, 800);

    sample = new SoundFile(this, "The Four Horsemen.wav");
    sample.loop();

    fft = new FFT(this, bands);
    fft.input(sample);
```

```

}

public void draw() {

    background(30, 30, 30);
    fill(255);
    noStroke();

    fft.analyze();

    for (int i = 0; i < bands; i++) {

        rect(i*barWidth+100, height, barWidth, -fft.spectrum[i]*100*scale);
    }
}
}

```

Figure 24: Processing source code for real-time music visualisation

visualiser; hence, the usage of the smoothing factor is not limited. However, it could be easily lost the shape of the frequency spectrum could be easily lost if the smoothing factor is too small, therefore in this thesis project, I used the smoothing factor between 0.05 and 0.2, which is valuable for creating practical images.

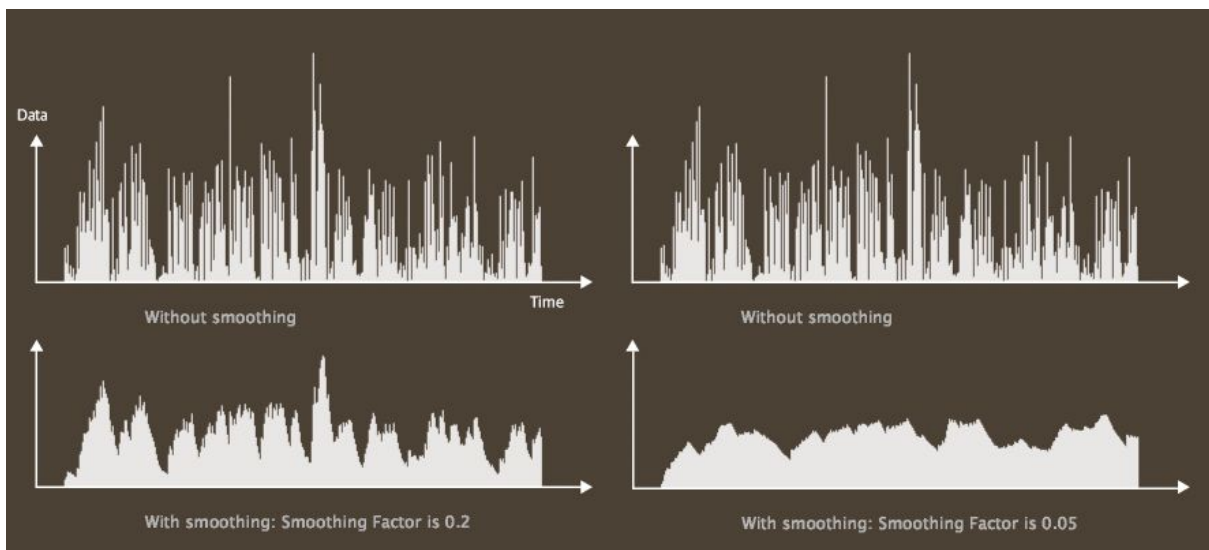


Figure 25. Comparison of different smoothing factors

4-2-2. Representation of a song by a line.

Unlike real-time music visualisation, this thesis project aims to generate non-volatile visual artworks. Hence, the next step for the prototyping was to experiment with various music visualisations that represent characteristics of a song. The first trial was to generate an image, including the overall frequency form of a song shown in Figure 26. In order to do that I used the lines to represent it, where one single line is the average of bands in a single frame. The reason why I used the line at the beginning is that the line is the most fundamental of all the elements of artistic creation such as drawing, painting, and sketching ideas [45]. The x-axis represents the time of a song and the colour is monochrome. As a result, it is easy to recognise the differences between the two songs, the first song(when she loved me) is pop-ballad and the second one(DNA) is hip-hop/rap. The most contrast is about the pattern of frequency spectrum which represents the rhythm/dynamics of the songs. While the first song has rhythm and idle periods repeatedly, the second song represents more rhythmic parts. From this result, I obtained an idea of how to use FFT analysis data and the average of FFT data bands.

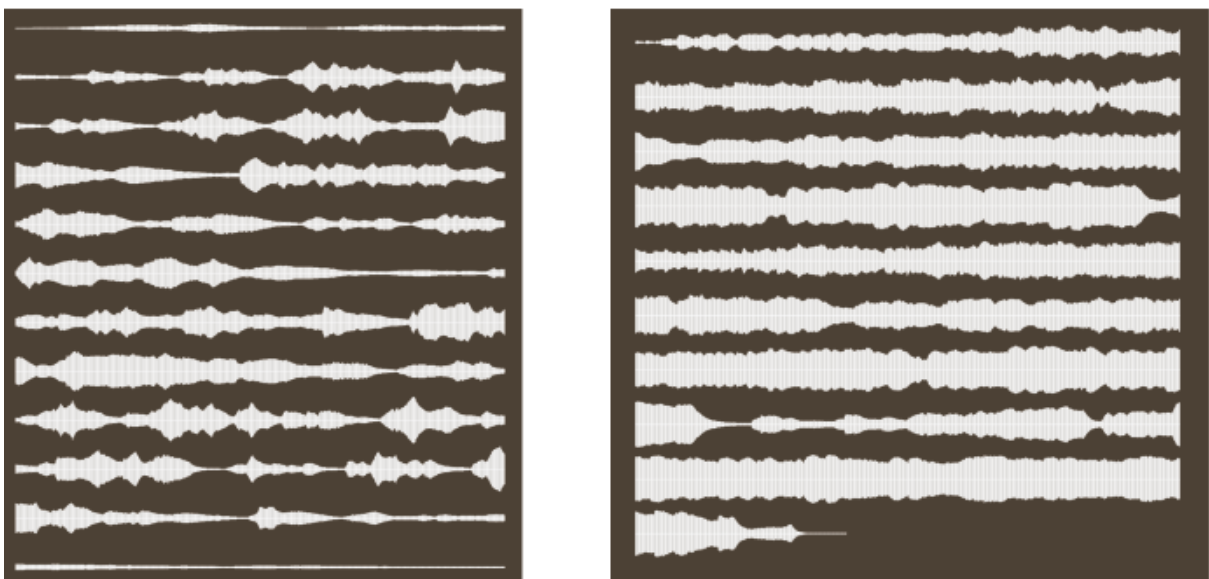


Figure 26: Comparison of the pattern of two songs.
The left song is "when she loved me", and the right is "DNA"

4-2-3. Representation of a song by a curved line drawing

Like a simple line, the curved line is also a fundamental element of drawing, and this line can be bent and vary direction continuously, it can be effortlessly wavy or spiral. These curved lines induce a feeling of comfort and ease [46]. Processing provides a function called 'curveVertex(x, y, z)' shown in Figure 27 to generate a curved line [47]. This function is the implementation of Catmull-Rom splines [48]. In order to implement the curved line drawing, I made a circle shape which has 12 vertices. Each vertex is mapped to a band of FFT analysis data where the vertex would move toward the normal vector depending on the value of the data. It means that lower data moves the vertex close to the centre of the circle; higher data moves to the opposite position, as shown in figure 28. However, there is a difference between line and curved line drawing to represent a song which is about time.

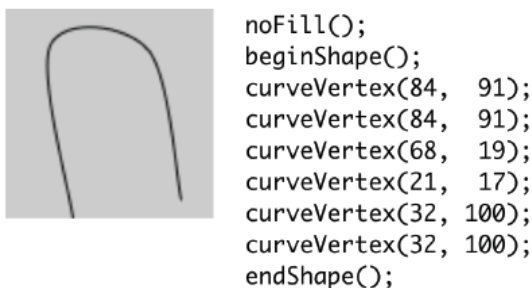


Figure 27: Processing example of curveVertex function

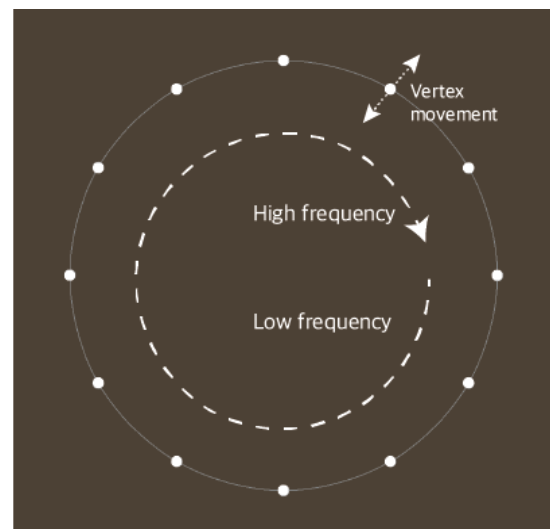


Figure 28:
The idea of representing a song in curved lines.

While in the line drawing, every single vertical line is mapped to the time sequence, the time in the curved drawing is superimposed based on the centre of the circle. Hence, I had to manage the opacity of each drawing; otherwise, the final image would be just white. It is easy to change opacity in Processing. The result is shown in

figure 29, while the pop-ballad song is represented in less movement in curved line drawing, hip-hop/rap song has more movement and energy.

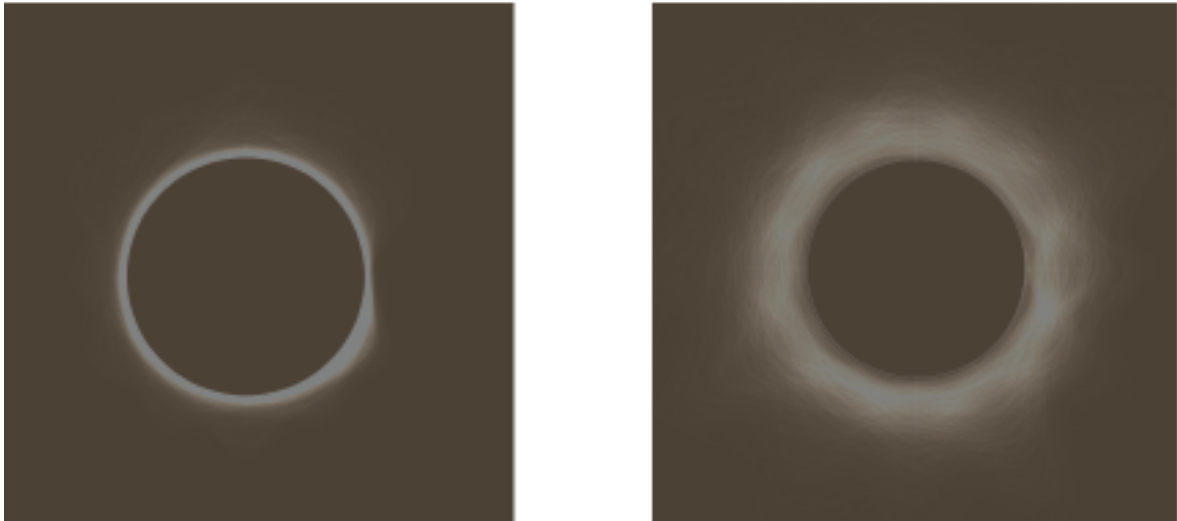


Figure 29: Comparison of the curved line of two songs.
The left song is "when she loved me", and the right is "DNA"

4-2-4. Representation of a song by a wavy line drawing

I also tried to represent a song with a wavy curved line, and this is simply an extended version of the previous trial. Unlike the circle shape, each wavy line represents a time towards the lower part of the screen, where the vertices are affected by the data of the FFT analysis band. In this experiment, it was required to manage the opacity of each drawing. Similarly to other results, it shows the differences between the two songs shown in figure 30, and the source code is shown in figure 31.

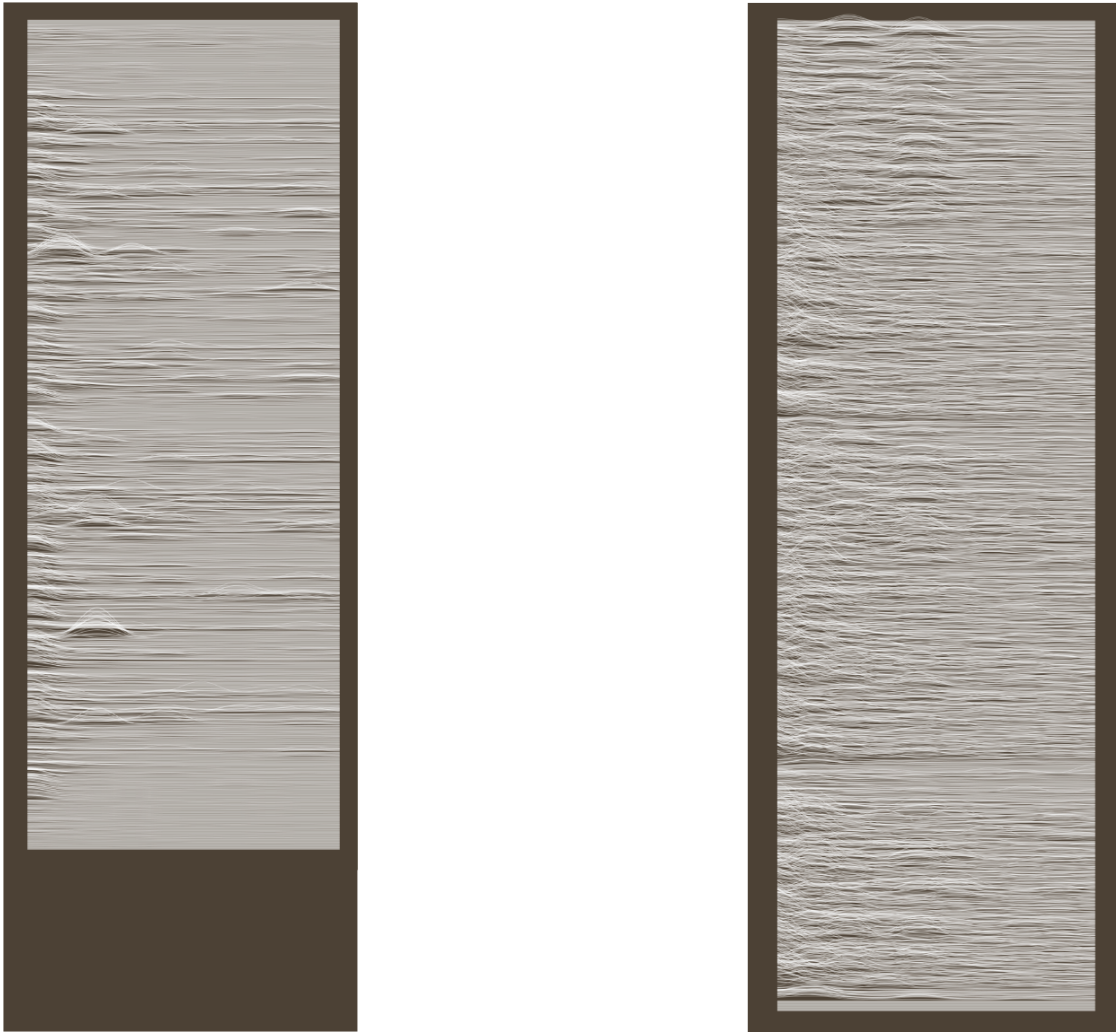


Figure 30: Comparison of the wavy line of two songs.

```

void drawWavy() {

  float data = 0;
  float weight = 1;
  float sizeOfEllipse;
  int start, end;
  float x1, y1, x2, y2, r;
  float angle = radians(360 / basePoint.length);

  start = 0;

  end = sectionInterval-1;

  noFill();

  for (int j = 0; j < basePoint.length; j++) {

    data = avgSectionLine(j*sectionInterval, end) * (weight);
  }
}

```

```

weight += 0.05;

data = sum[j*2];

basePoint[j].x = 0-data * 50;
basePoint[j].y = (basePoint.length - j) * 50;

start += sectionInterval;
end += sectionInterval;
}

stroke(100,0,100,80);

translate(curX, 100);

beginShape();

for (PVector vector : basePoint) {

    curveVertex(vector.x, vector.y);
}

endShape();
}

```

Figure 31: Processing code of wavy lines.

4-2-4. Applying a colour for the drawings

For all the results of previous trials, the colour was monochrome. Although in this phase I did not have an idea of how to map the emotion in music to the colour theme, I wanted to apply some colour to make better results out of the drawings. In order to do that, I could easily map some colour to FFT analysis data, for instance: higher-value data could result in more reddish colour, or lower-value data could be related to yellowish colour. It means that somehow I should affect colour gradually, which is a very similar concept with the gradient. There are several colour models, and the RGB(red, green, blue) is one of the most well-known ones. RGB is an additive colour model, and this colour model is broadly used for displaying images in electronic systems, such as television and computer monitors [49]. In Processing, the

default colour mode is RGB. Whenever users want to change the colour of stroke or fill they can use one of these functions "stroke(v1, v2, v3, alpha)" or "fill(v1, v2, v3, alpha)". The parameter v1 = Red, v2 = Green, v3 = Blue and alpha is the opacity of the colour. Although the RGB colour model is widely used, one of the drawbacks of this colour model is that it is not easy to select colour intuitively. Whereas HSB (hue, saturation, brightness) colour model is based on human perception, and it gives the advantage over RGB, that is selecting colour intuitively [50]. The HSB colour model has a cylindrical geometry as shown in figure 32, with hue, starting with 0° representing primary red, passing through the green primary at 120° and the blue primary at 240°. Saturation varies based on how much white colour is added, and the brightness varies by adding black colour [51]. Hence, the HSB colour model provides an easy way to choose the colour for neighbours gradually, also to obtain a complementary colour by just adding 180° [52].

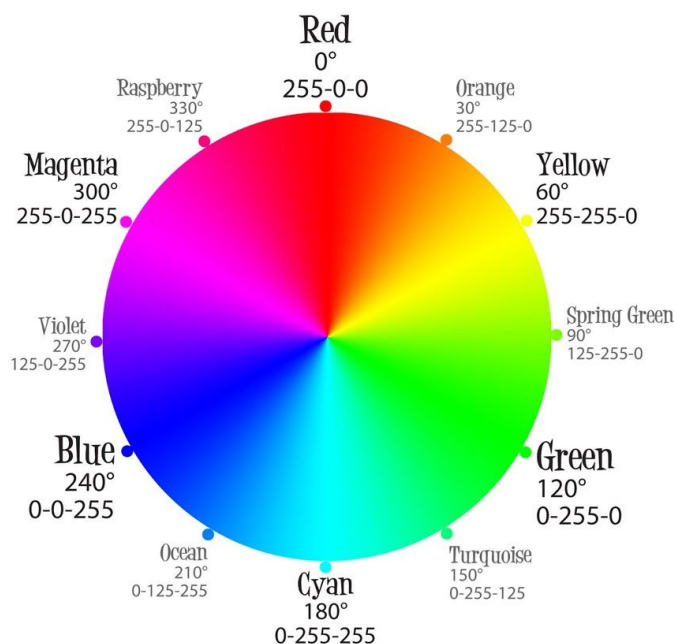


Figure 32: The HSB colour model [49]

In order to use the HSB colour model, Processing provides a function called "colorMode(mode, max1, max2, max3)". The mode can be either RGB or HSB, in HSB colour model, the variable max1 represents the hue, max2 the saturation, and max3

the brightness. I added a simple algorithm to the line drawing project shown in figure 33, and the hue is mapped to the FFT analysis data, with the angle starting at 20°. So, the lower-value data represents the more reddish colour, and higher-value data has become a more greenish colour. Although the colour was not mapped to the emotion of the songs yet, the shape of each song represents the rhythmic movement of the songs, and also the colour is supporting it.

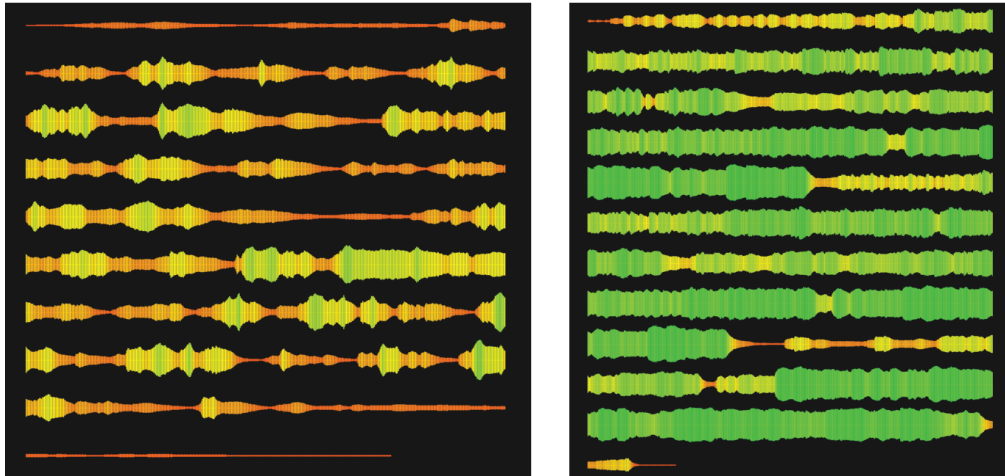


Figure 33: Comparison of the HSB colour.

The left song is "when she loved me", and the right is "DNA"

4-3. First Iteration

This chapter describes the design and implementation of the first iteration of the generative sound visualisation project for the exhibition at Aalto University. Based on the previous prototyping stage, I have learnt how to use FFT analysis data in practice and developed various visualisation styles such as line, circle and a curved line. Also, I have realised that the HSB colour model is an efficient way to choose a colour naturally compared to the RGB colour model.

4-3-1. Visualisation Design

For the visualisation design for the exhibition, I have decided to use the wavy style in chapter 4-2-4 to represent a song in generative artwork. The wavy line style is a practical way to generate FFT analysis data into an image naturally. The colour was the HSB colour model, and each line was mapped to the FFT analysis data. In order to generate a wavy image, I had to manage the width of each line related to the FFT analysis band data. For instance, lower-value data generate a narrow width of the line and higher-value data create a thicker line shown in figure 34. Also, for blending lines (with colour), I used the superimposed technique learnt from the prototyping stage. This means that when the current y position reaches the height of the display window then y position is set to the top position. Hence, the benefit of this idea is that I could generate a more dynamic, and natural-looking image regardless of the length of the song compared to the result of the previous prototyping stage.

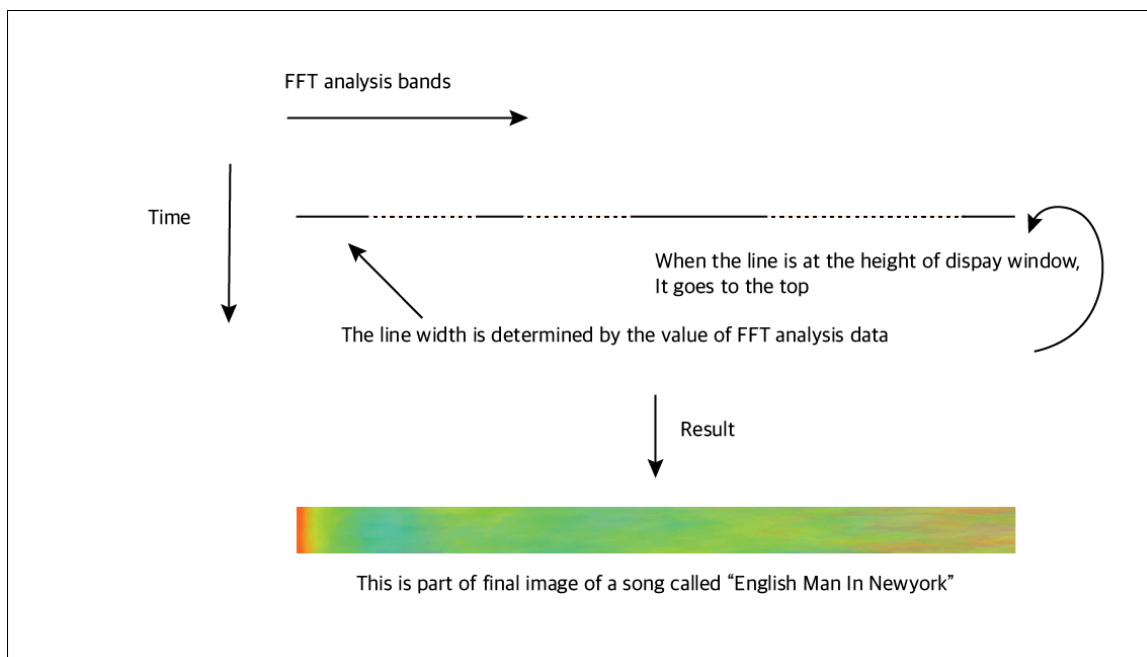


Figure 34: Visualisation Design

4-3-2. Software Implementation

In the software implementation stage, I decided to use openFrameworks instead of Processing, because in general, the real-time rendering speed of openFrameworks is faster. The overall system block diagram is shown in figure 35.

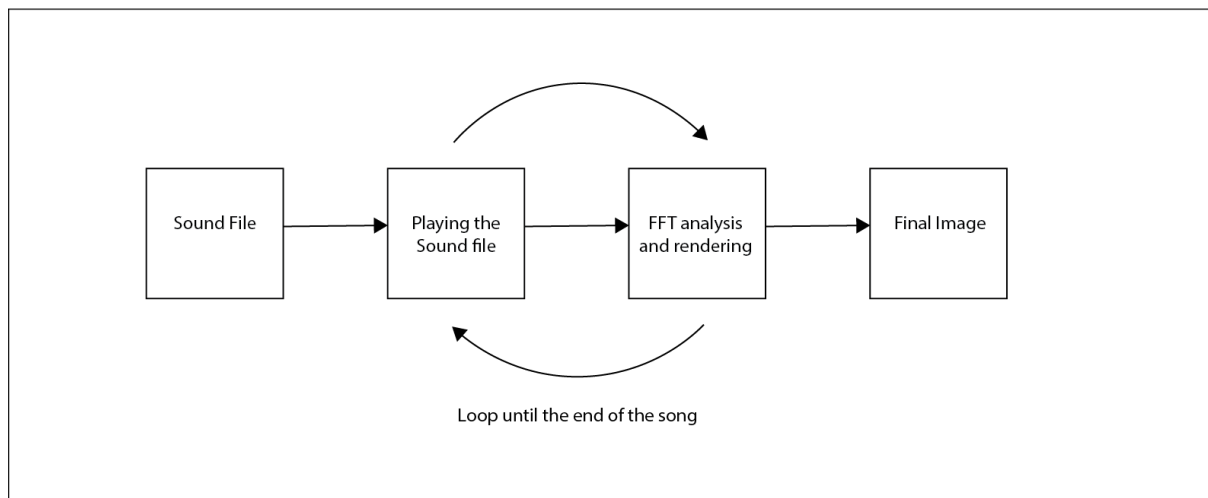


Figure 35: Overall system block diagram

Next, I designed software architecture based on UML (Unified Modeling Language) expression shown in figure 36. UML is a standard way to visualise the design of the software system, and it is intended for expressing the field of software engineering in common language [55]. There are two main classes designed. The first one is called 'JLSound' and the second one is 'JLPainter'. The purpose of JLSound class is to load a sound file (this could be mp3, wav etc.) and generate FFT data while the song is playing by using a function of 'ofSoundPlayer' class. Like Processing, to retrieve FFT spectrum data is easy in openFrameworks by using the 'ofSoundGetSpectrum(int nBands)' function. The parameter nBands is the number of spectrum bands to return, and the maximum is 512. The second class is JLPainter, and it contains skeleton functions to generate an image from FFT spectrum data. JLPainter class is the root class that provides a common data set and functionalities. However, this class does not contain the actual function body. Instead, it provides virtual functions (shown in figure 37). The main reason why I made the root class was that I wanted to create various visualisation classes by

inheriting from the JLPainter class shown in figure 36. Inheritance is one of the powerful concepts of OOP (Object-oriented programming) [56]. Hence, the inherited class (called child class) should implement the virtual functions from their parent class shown in figure 37. The benefit of this technique is about standardisation, and it means that many classes could be inherited from the parent class. All the inherited classes have the same function name, but the implementation of each function of classes could differ. Hence, the caller class or function does not need to consider their logic of inherited class inside. This concept is polymorphism in OOP concept [53]. Based on the result of the first iteration, I made four different song visualisations shown in figure 38.

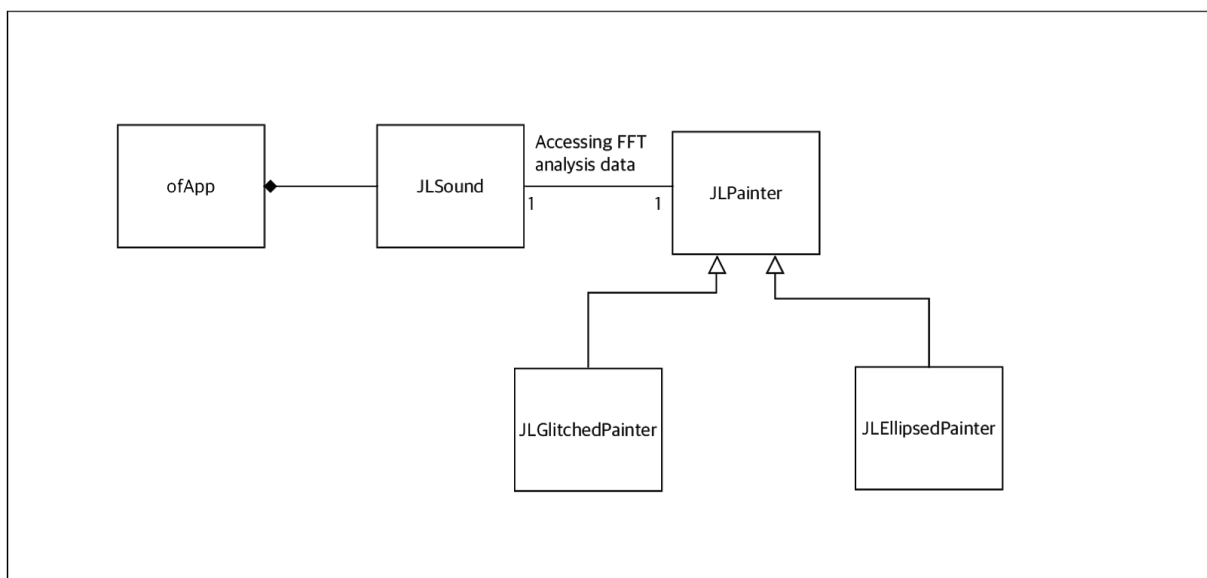


Figure 36: Class diagram

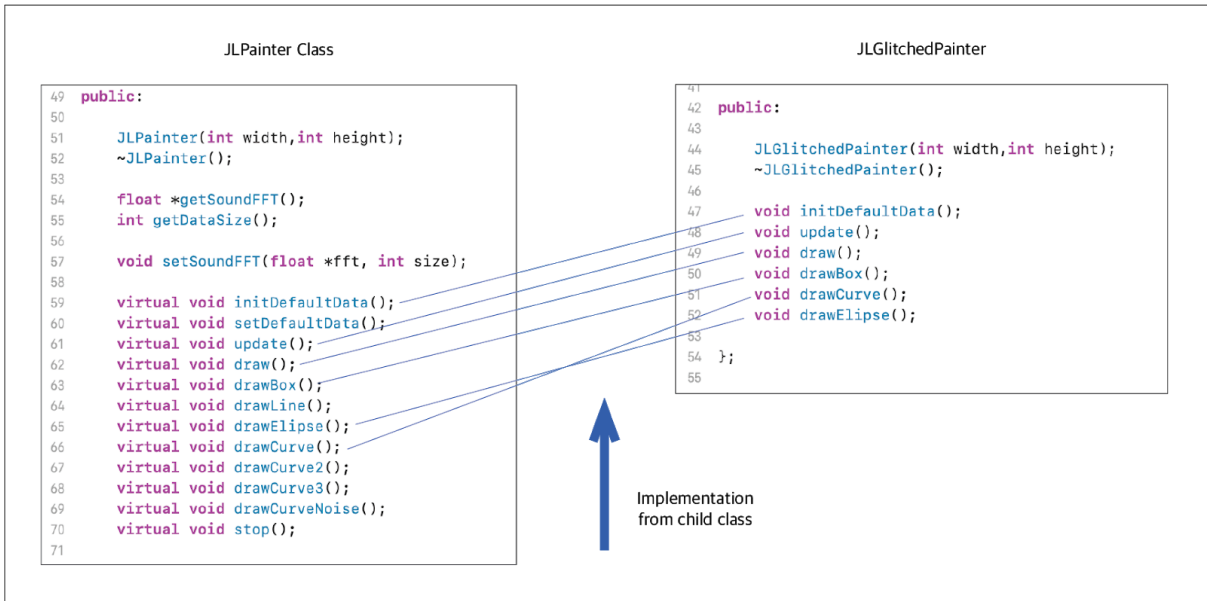


Figure 37: Example of polymorphism in OOP concept

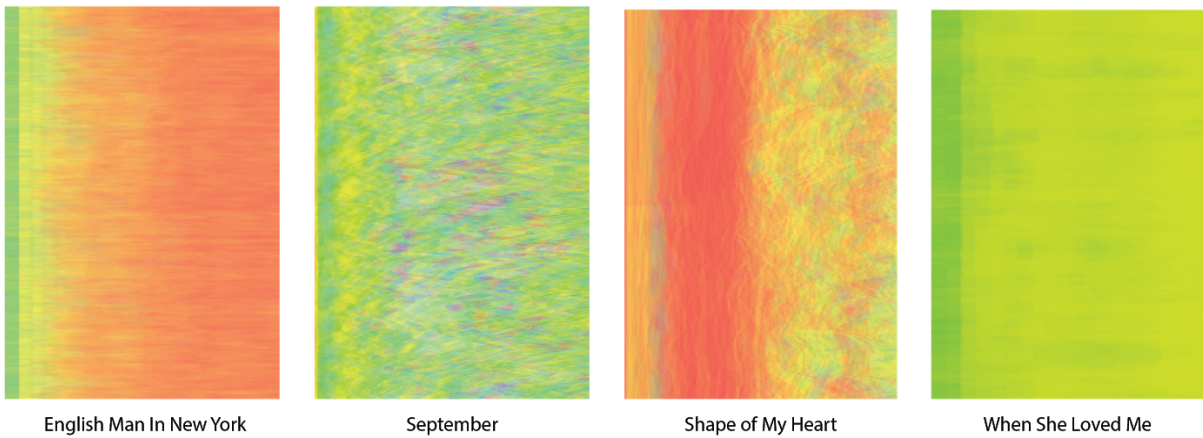


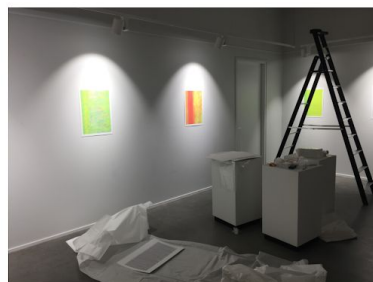
Figure 38: Sound visualisation for the exhibition

4-3-3. Exhibition

This exhibition aimed to give the audience a new experience to perceive a song in visual interpretation, as shown in figure 39. Through my exhibition, I focused on making this invisible data more visual, observing how the audiences could potentially interpret sound differently if it was transformed into another medium, such as a generative image. The exhibition period was 23 Feb - 03 Mar 2019 at Väre gallery V1 of Aalto University. The final outcome for this exhibition is the printed image on the paper, so I visited the print lab at Aalto University. There is a printing master who provided a professional print quality for the exhibition. I used the A2 size paper. I did not make an extra frame on it because It was thick enough to hang on the wall. It was my first solo exhibition, and even though I had no idea what is the optimal height for the hanging of my images on the wall, then I found the rule of thumb for hanging artwork is 58 inches on centre.



The exhibition poster



Preparation



58 Inches from the floor

Figure 39: Photos of the exhibition

4-4. Second Iteration

This chapter demonstrates the continuing development of the music visualisation based on the first iteration. Specifically, it has improved how to connect the emotion in the song into a colour theme. In order to do that, I used the emotion colour wheel created by Ståhl [19]. Also, I tried to improve the quality of the aesthetic visualisation of this thesis project.

4-4-1. Visualisation Design

The visualisation design for the second iteration aimed to improve the colour scheme and aesthetic. The emotional colour wheel by Ståhl is based on Russell's Circumplex Model of Affect theory [20]. Several kinds of research are described in chapter 2-3 about MIR (music information retrieval) conducted to find the emotion in a song, and I choose the API provided by Spotify. The main reason I selected the API of Spotify is that Spotify is one of the biggest music content providers, and it has been used widely, and they provide enough information for this thesis project. In figure 40, it is shown that Spotify provides several meaningful information, such as energy and valence. The meaning of energy is the same as arousal and valence in Russell's model. The data range of both arousal and valence from Spotify API is from 0 to 1, which is not matched to Russell's model. Hence, I mapped those data to the range between -1 to 1 that is the proper data range for both Russell's model and Ståhl's emotional colour wheel [20].

I am inspired by Molnar Vera's generative artwork which is described in chapter 3-1. In her project, she separated the area by rectangular shapes, and each shape has various rectangular objects inside. Her artwork brought me a new idea of visualisation for the second iteration. First, I separated the area of the drawing by rows and columns. Each block has multiple polygon shapes which represent the FFT

analysis spectrum, and the size is mapped to the average of total bands. Hence, one block could have more than 100 frames of FFT analysis spectrum data depending on the number of rows and columns shown in figure 41.

Get Audio Features for a Track

Get Audio Features for a Track **DOCs**

Description Get Audio Features for a Track **DOCs**

Endpoint <https://api.spotify.com/v1/audio-features/{id}>

HTTP Method GET

OAuth Required

id *

50o8QmrUrzQRbdyfx5Lez

OAuth Token

BQBgg8cDfj_9XuKfuwRqQQfp8VHBcfJMPme3H...

GET TOKEN

TRY IT **FILL SAMPLE DATA**

```

{
  "danceability": 0.287,
  "energy": 0.128,
  "key": 5,
  "loudness": -14.672,
  "mode": 1,
  "speechiness": 0.0439,
  "acousticness": 0.983,
  "instrumentalness": 0,
  "liveness": 0.107,
  "valence": 0.23,
  "tempo": 77.345,
  "type": "audio_features",
  "id": "50o8QmrUrzQRbdyfx5Lez",
  "uri": "spotify:track:50o8QmrUrzQRbdyfx5Lez",
  "track_href": "https://api.spotify.com/v1/tracks/50o8QmrUrzQRbdyfx5Lez",
  "analysis_url": "https://api.spotify.com/v1/audio-analysis/50o8QmrUrzQRbdyfx5Lez",
  "duration_ms": 185240,
  "time_signature": 5
}

```

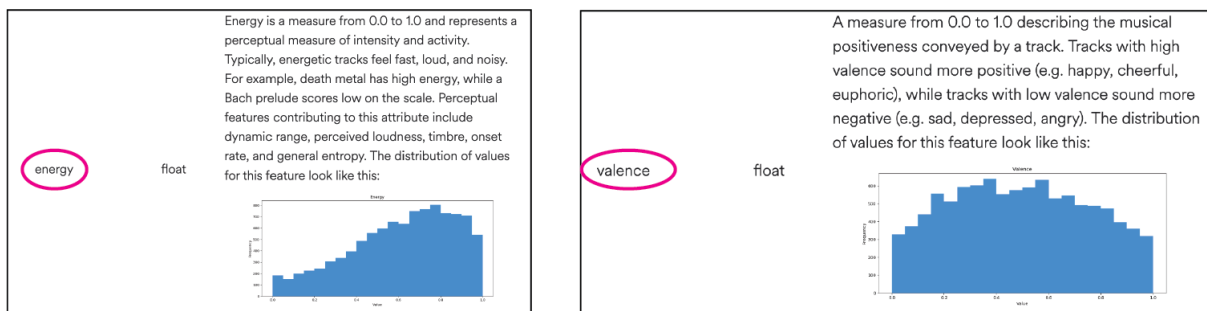


Figure 40: The music feature retrieval results of “When she loved me” by using Spotify Web API

In order to find the corresponding colour based on Russell’s emotional model, I mapped the arousal and valence to the emotional colour wheel. Hence, I could retrieve the primary emotional colour of a song. Then I choose four more similar colours for supporting the primary colour. Each supporting colour is selected by adding a specific angle to the primary colour such as +16°, +32°, -16° and +53°, and it could be easily done in the HSB colour model. I developed a software which helps to find the primary and supporting colour easily, as shown in figure 42; otherwise, I had to calculate every HSB colour set manually, which would be wasting time unnecessarily.

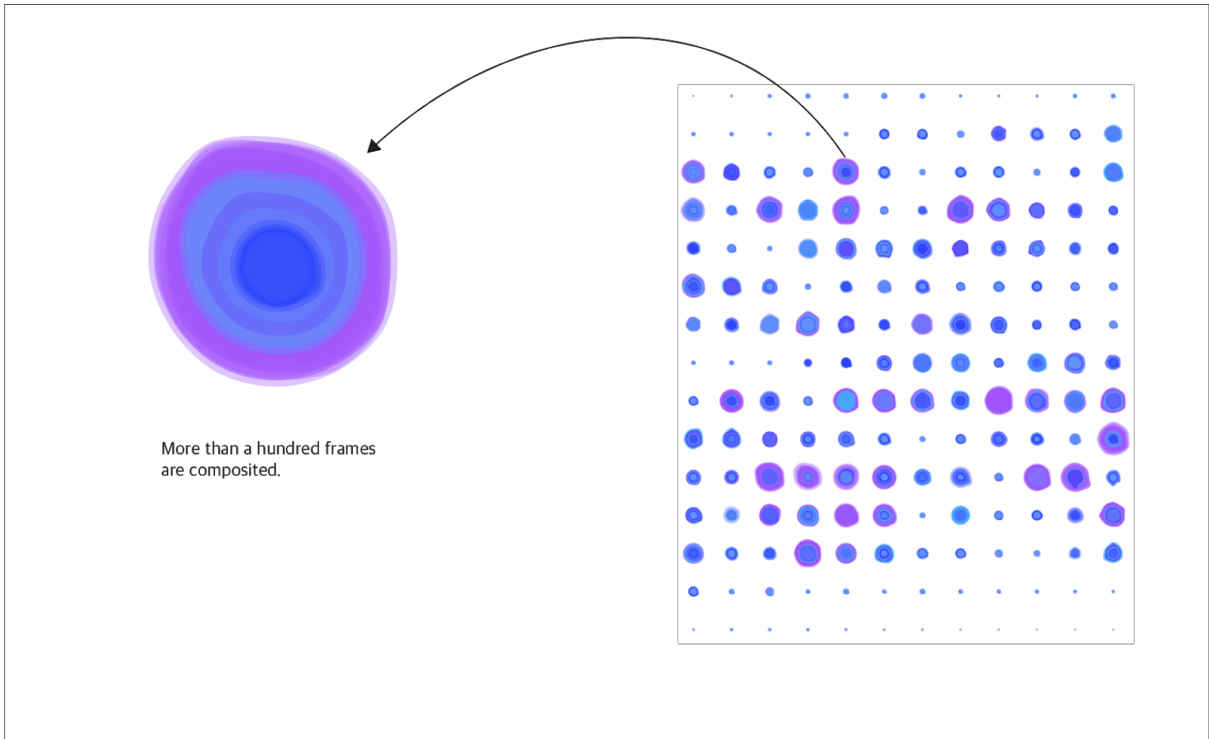


Figure 41: Visualisation idea of the second iteration

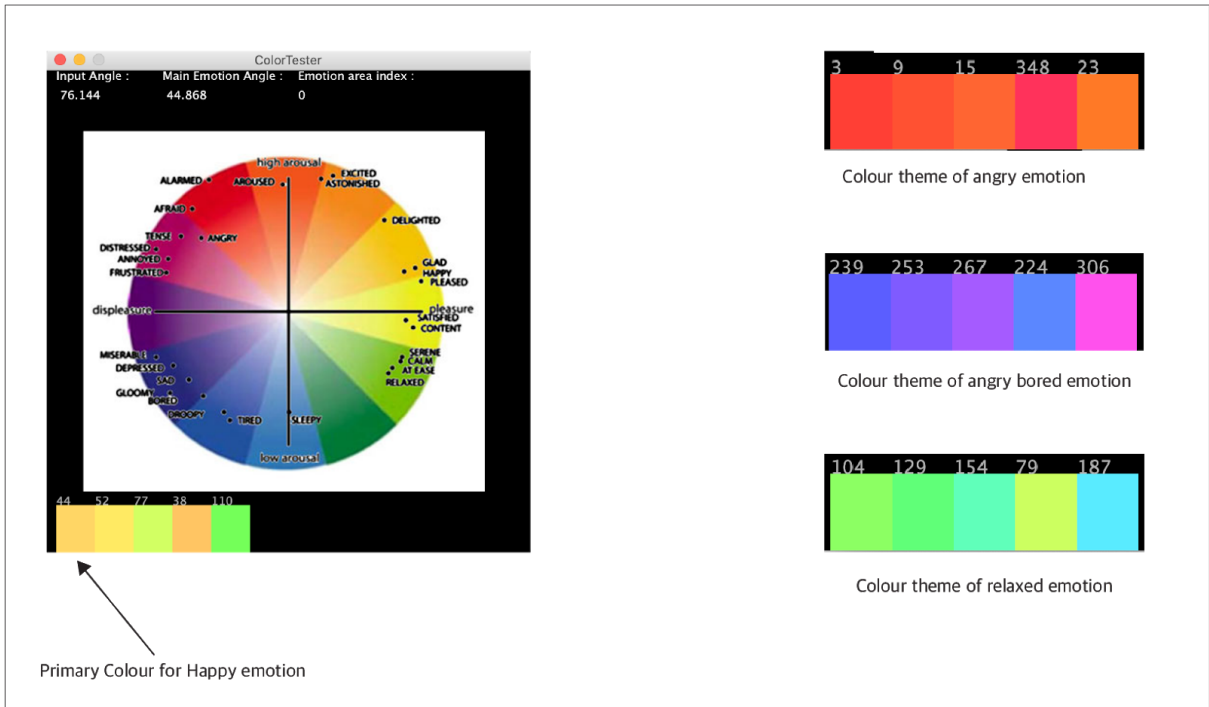


Figure 42: Various emotional colour themes

4-4-2. Software Implementation

In the second iteration, I changed the data gathering method, away from real-time FFT analysis. It runs a pre-processing stage before rendering the image. The FFT analysis for the song runs only once and saves every single FFT analysis spectrum data as a text file. Then I could read the text file to visualise a song, which saves much testing and developing time. For instance, in the first iteration, I had to wait until the song finished playing. However, in the second iteration, it just took less than 30 seconds for each song, depending on the complexity of the rendering algorithms. The system block diagram is shown in figure 43. Also, there was another benefit of pre-processing: the size of the final image is no longer limited. In the first iteration, the size is limited by the performance of the computer, because the rendering speed depends on it; otherwise, I would lose many frames. However, with the pre-processing idea, I could generate a large size (ex. 22900×28600) image, and it would be of adequate size for high-resolution printing.

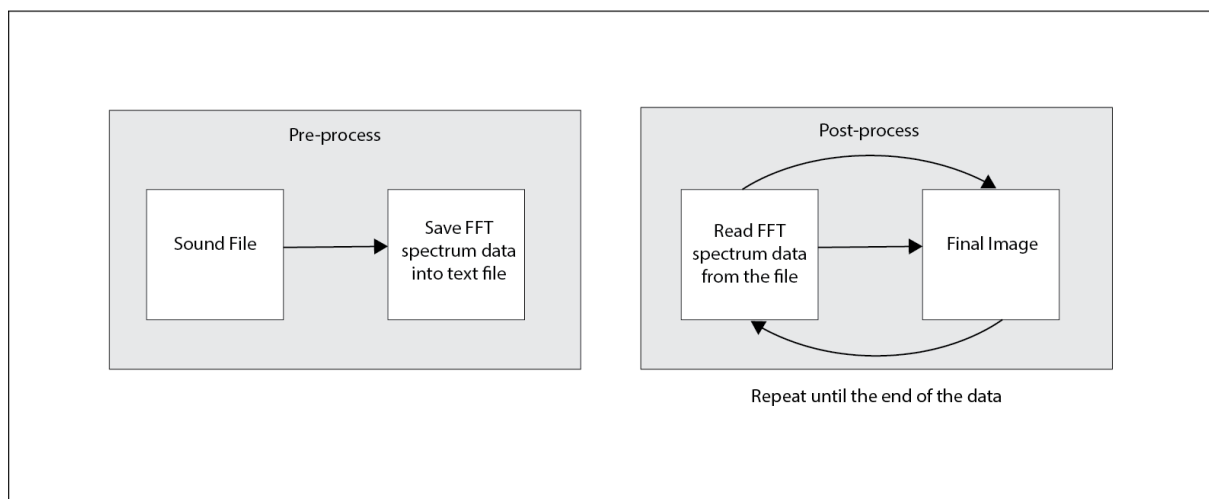


Figure 43: System block diagram for the second iteration

The design and implementation approach of class is a bit different than the first iteration because of the pre-processing for the FFT analysis shown in figure 44. The main difference compared to the first iteration is the role of the sound analyser class called JLSoundAnalyser. This class saves the FFT analysis spectrum data as a text file and provides functionality to read the files from beginning to end. Also, this class

provides several useful functions such as avgSectionLine, avgOfLine, and getData. Those functions are required to implement the visualisation design described in chapter 4-4-1. The biggest challenge for the second iteration is to connect emotion in the song into the colour theme. In order to do that, I made a class called calcColorTheme (float valence, float arousal) shown in figure 45. This function generates colour palettes as an array, based on parameters valence and arousal that comes from Spotify Web API. In table 1 shows the final result applied to the colour theme calculated by calcColorTheme function for each song.

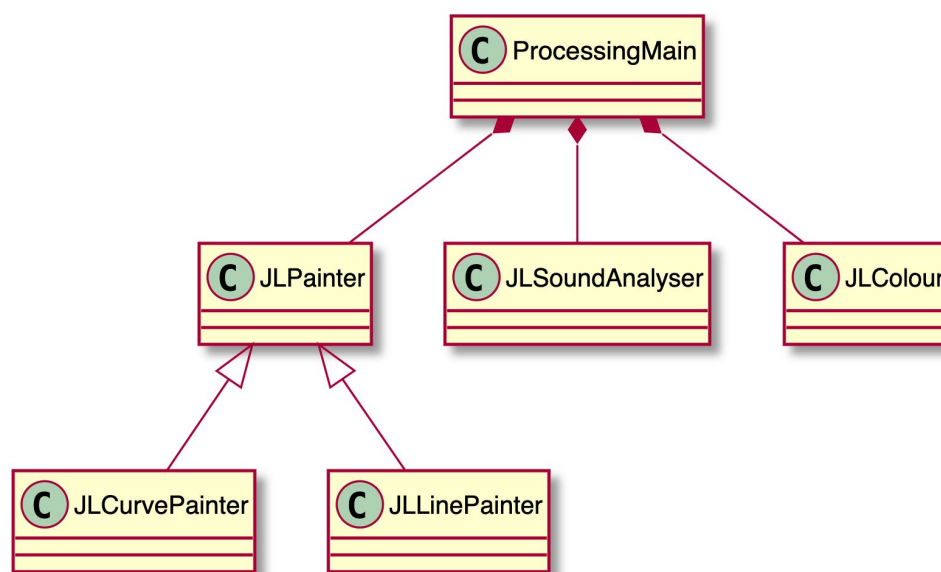


Figure 44: Class diagram of the second iteration

```

float calcColorTheme(float valence, float arousal) {

    float retAngle = 0;
    float distance;
    float angle;
    float tempArousal = arousal;

    valence = map(valence, 0, 1, -1, 1);
    arousal = map(arousal, 0, 1, 1, -1);

    angle = degrees(atan2(arousal, valence));

    angle += 90;

    arousal = map(tempArousal, 0, 1, -1, 1);

    println("Current Valence is "+valence+" Arousal is "+arousal);
  }

```

```

println("Current angle is "+angle);

distance = map(dist(0, 0, valence, arousal), 0, 1.41, 50, 100);

retAngle = getConvertedAngle(angle);
palettes[0].set((int)retAngle, int(distance * 0.76), 100); // main

println("Main emotion angle is "+retAngle);
text("Main Emotion Angle :", 120, 10);
text(retAngle, 120, 30);

retAngle = getConvertedAngle(angle+16);
palettes[1].set((int)retAngle, int(distance * 1), 100);

retAngle = getConvertedAngle(angle+32);
palettes[2].set((int)retAngle, int(distance * 0.95), 100);

retAngle = getConvertedAngle(angle-16);
palettes[3].set((int)retAngle, int(distance * 0.90), 100);

retAngle = getConvertedAngle(angle+53);
palettes[4].set((int)retAngle, int(distance * 0.85), 100);

text(distance, 400, 10);

for (int i=0; i<palettes.length; i++) {
    palettes[i].showMe();
}

return retAngle;
}

```

Figure 45: Source code for generating colour theme





Song title	Arousal	Valence	Emotion	Colour theme
When She loved me	0.128	0.230	Groopy, tired	
DNA	0.774	0.696	Excited	
You Are The Sunshine Of My Life	0.45	0.587	Calm, Relaxed	
The Four Horsemen	0.953	0.335	Alarmed, Afraid	

Table 1: Various colour theme based on the emotion of music

4-5. Example applications

In this chapter, I would like to introduce how this thesis project results could be applied for our daily life not only for the exhibition.

4-5-1. Generative artwork in daily life

As any other artworks, the result of generative art could be exhibited in a public space such as a museum, or gallery. Figure 46 is an example of how the output of this project can be exhibited in a gallery. Also, figure 47 demonstrates that the outcome of this thesis project can be used as a pattern for our daily life product such as a cup and it can be an accessory in a living room. All the photoshop mockups are created by rawpixel.com - www.freepik.com



Figure 46: Example of an exhibition at the gallery



Figure 47: Example of daily life products

4-5-2. Visualisation of Spotify music chart

Since this thesis project generates an image from a song, the colours and shapes differ among songs. The colour theme is especially related to the emotion in the song. Hence, it could be applied for music chart visualisation instead of numbers and a traditional bar chart. Figure 48 shows the top nine played songs on 12th March 2020 from the Spotify chart.



Figure 48: Visualisation of Spotify music chart
The photoshop mockup is created by rawpixel.com - www.freepik.com

Chapter 5. Discussion

The randomness is the core value of generative art because it leads to unpredictable aesthetic results that are generated by a computer algorithm [53]. Hence, the artists are using randomness as a method for generating creative results, such as artworks shown in figure 8,11 by Vera Molnár and Tyler Hobbs. However, for this thesis project, the randomness is not suitable to generate an image from the music, because the aim of this thesis project is that the shape of an image should represent the frequency of the music, this is why I used the FFT analysis data instead of the randomness for visualizing sound.

Processing and openFrameworks are practical tools to implement a software system related to visualisation. Both tools provide access to many premade libraries. The main difference of openFrameworks compared to Processing is that openFrameworks is based on C++ language. The learning curve of C++ is a bit harder than Java language; for instance, in C++ developer should take responsibility for memory management. On the other hand, Java has an automatic memory management feature which is a benefit [58]. However, from the performance point of view, openFrameworks runs a bit faster than Processing. Because Java is an interpreted language, it means that Java source code is translated into machine language in real-time. I used Processing for this thesis project when I needed to make rapid prototyping software. In contrast, I used openFrameworks for drawing an image faster, for instance, the result of the first iteration in chapter 4.3. Hence, to use Processing or openFrameworks absolutely depends on the requirements of the task, the level of coding skill, and the individual's preference.

In the first iteration, there was a drawback for generating an image during a song playback. While I was implementing and testing software, I realised that I have to wait for the results until the song playing is done. It took much time unnecessarily because Processing and openFrameworks provide an FFT function for real-time

analysis of sound. In order to solve this problem, I tried to reduce the unnecessary time for implementing and testing the system by a pre-processing concept in the second iteration. With this approach, instead of generating an image while the song is playing, the sound analysis class saves the FFT spectrum data into a text file. Once the file is created, I do not have to play the song again, I load the text file and use it for generating an image. Hence, I could rapidly get the results by changing various parameters that were mapped in order to experiment with many different styles of visualisation.

The colour theme of this thesis project represents the emotion in the music, and it is based on research such as Russell's 'A circumplex model of affect', and emotional colour wheel by Ståhl [4][19]. This method led me to use MIR (music information retrieval) solutions such as Spotify Web API for analysing music tracks. The drawback I found from my thesis project relates to the lack of freedom of colour selection by the artist. As Tylor mentioned, the colour theme should be selected carefully [30]. The software algorithm entirely decides the colour based on emotion in the music, then where is the creativity and esthetic preference of the artists for selecting the colour theme?

The visual theme of the second iteration was inspired by Vera Molnár's generative artwork, especially (Dés)Ordres. She divided the screen as a grid which has regular width and height. Within each grid point, there is a square that is disrupting the consistency. Eventually, The contrast between order and disorder brought the movement of vibrating the squares [24]. Hence, I thought that I could apply her concept to the visual theme of this thesis project. Unlike she used squares for her artwork, I decided to use a rounded shape, because in this thesis project the number of FFT spectrum bands is more than four, and the rounded shape of many FFT spectrum bands makes a better result.

The drawback I found in the work I have developed is the lack of variety of visualisation. Since the visualisation style of the second iteration is a circular shape and the size of it is related to the average bands of FFT spectrum data, it creates a similar shape and colour among the songs which have the same emotion and music elements. This side effect should be improved in future research.

Chapter 6. Conclusion and Future Plans

Since both music and visual artwork induce emotion, I believe it is worthy of trying to convert audible media into the visual medium. The main goal of this thesis project is to demonstrate how to generate visual artworks based on analysing music tracks and connect the emotion of the music into the colour theme of the visual artwork. In order to implement this thesis project, I had to make sure that I could obtain the numerical data which represents emotional value in music. Also, I had to find out how to map the emotional numerical data into a colour theme in this thesis project.

By using both Processing and openFrameworks, I could quickly generate several images from the songs by the algorithm that I made. This is the main concept of generative art, that is the autonomous system creates artworks [4]. Of course, the visualisation style is a very personal outcome, and it could be perceived differently among audiences.

This thesis project could show another way to visualise the character and emotion of music with generative art. It describes the emotion of music in numerical representation and maps it to colour. Hence, the approach of this thesis project could be applied to various music visualisations, including real-time or non-real-time approaches. Also, to reduce the whole project development time, this thesis project

followed a pre-processing approach, which saves FFT data into a text file. Finally, the project demonstrated how to apply generative artworks into a daily product, not only for an exhibition.

For my future research, I would like to try various visualisation styles and also do a survey on how the colour theme is related to emotion. Also, for visualisation style, I would like to use AI techniques to generate art pieces for music representation. It means that once the AI system (music visualiser) would make a generative artwork, then the other AI system could analyse the result and give feedback to the music visualiser until the shape and colour are a good fit to the emotion in the music.

Appendix

This thesis project can be found online at

<https://github.com/delphi1004/SoundPainter>

The Web API provided by Spotify

<https://developer.spotify.com/documentation/web-api/reference-beta/>

The exhibition archive of the first iteration

<https://johnlee012.com/exhibition/>

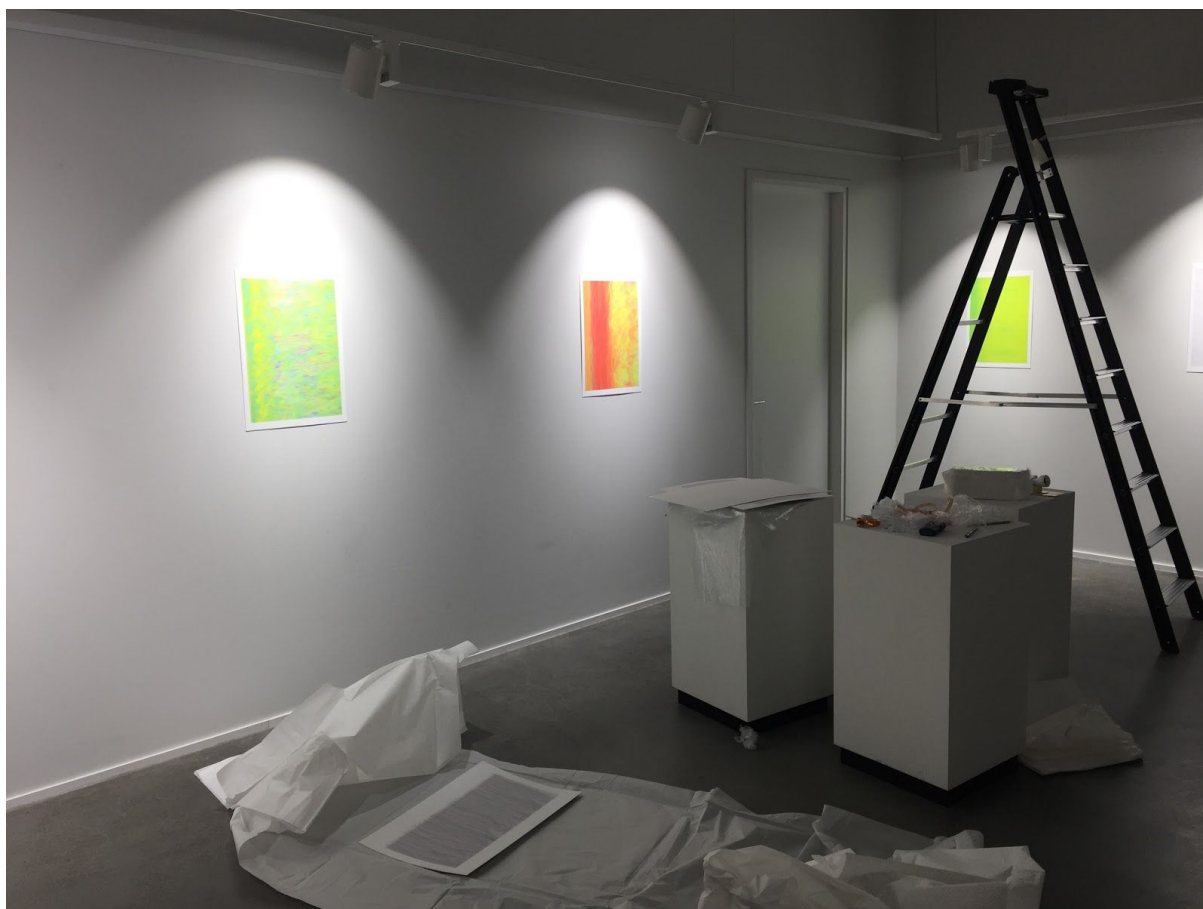


Figure 49: Preparation of exhibition at Aalto University



Figure 50: Sound visualisation of "English Man In New York"

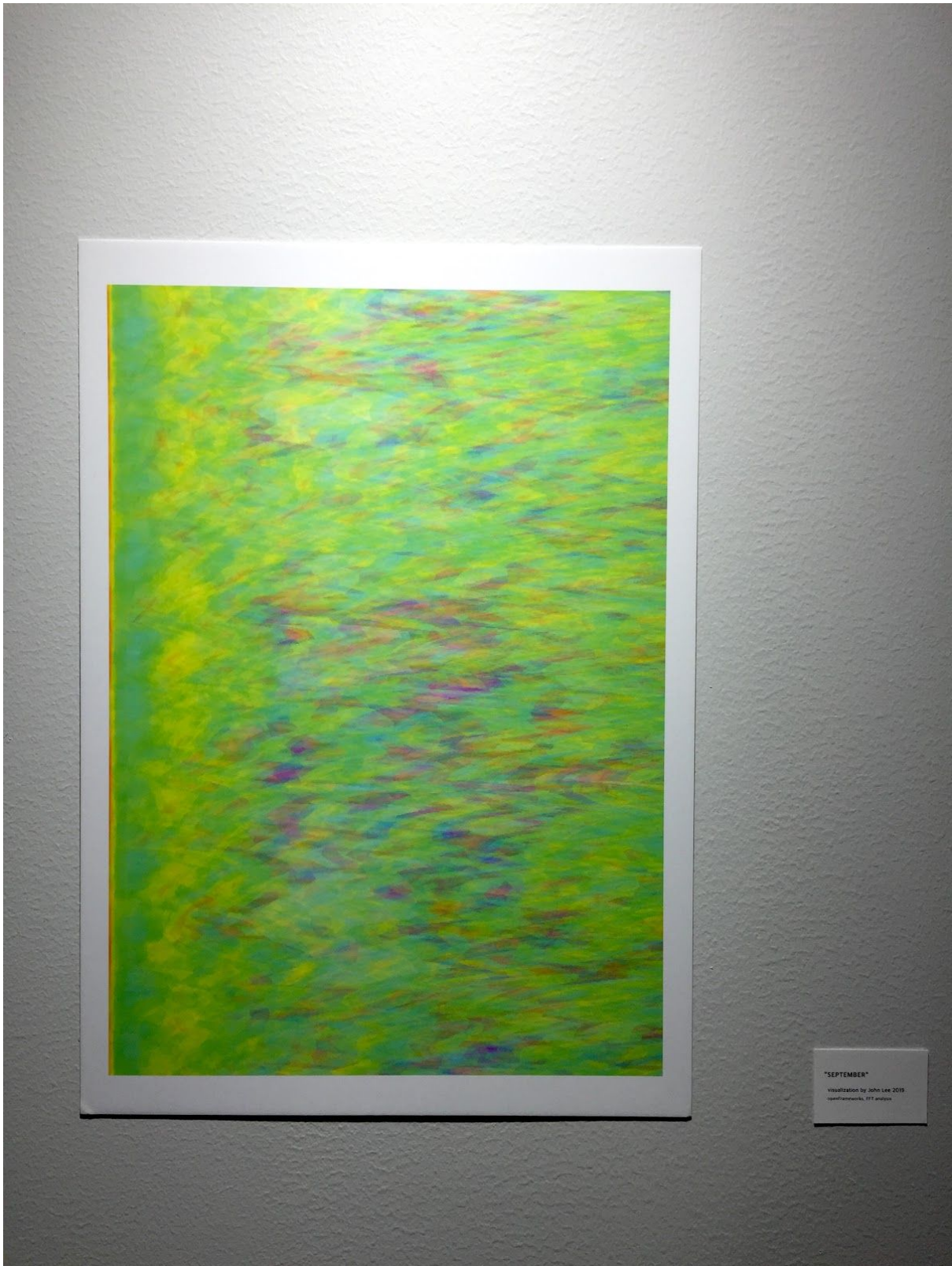


Figure 51: Sound visualisation of "September"

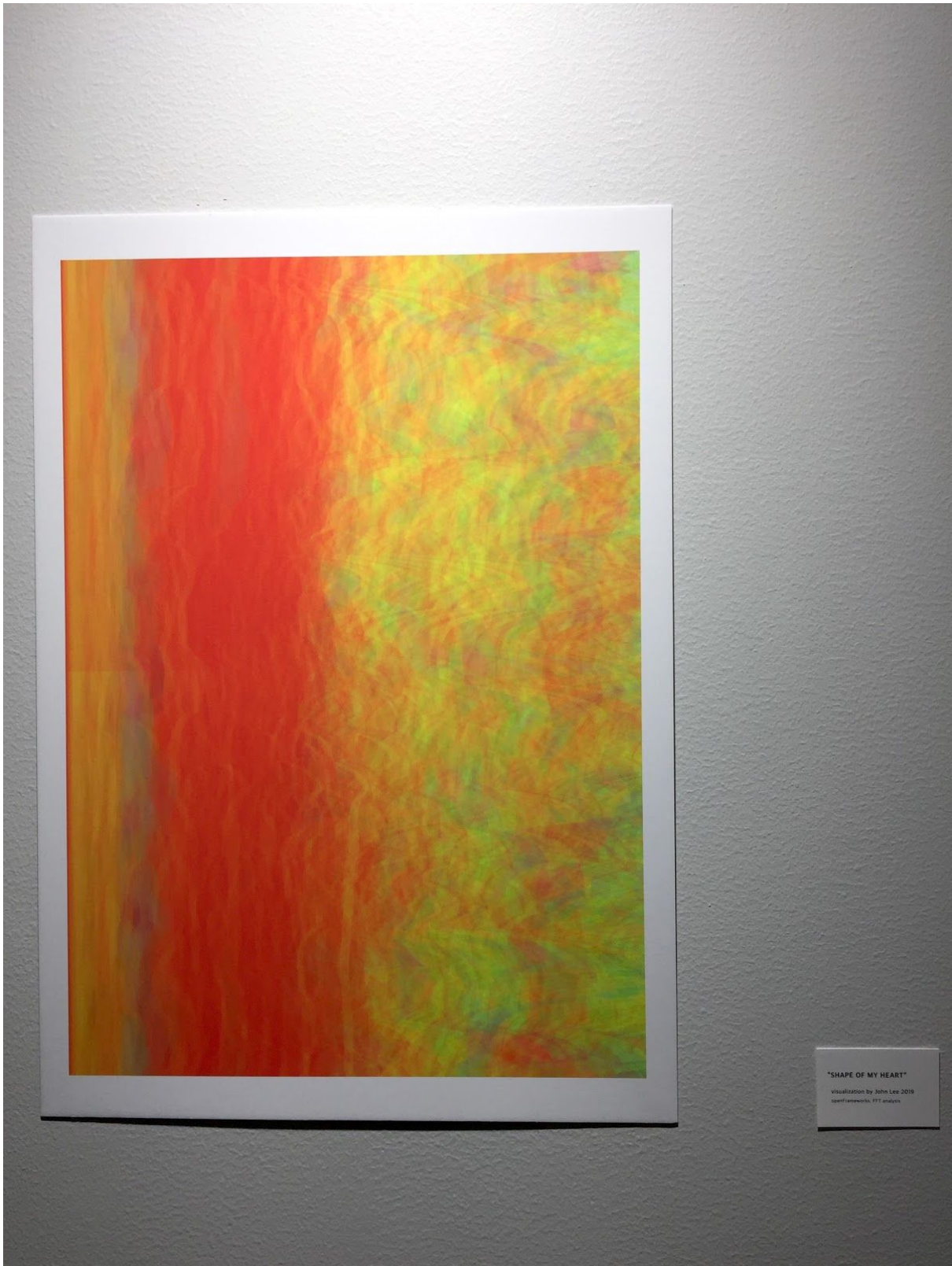


Figure 52: Sound visualisation of "Shape Of My Heart"



Figure 53: Sound visualisation of "When She Loved Me"

References

- [1] K. R. Scherer, "Which Emotions Can be Induced by Music? What Are the Underlying Mechanisms? And How Can We Measure Them?," *J. New Music Res.*, vol. 33, no. 3, pp. 239–251, Sep. 2004, doi: 10.1080/0929821042000317822.
- [2] G. Csurka, S. Skaff, L. Marchesotti, and C. Saunders, "Learning moods and emotions from color combinations," in *Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing - ICVGIP '10*, Chennai, India, 2010, pp. 298–305, doi: 10.1145/1924559.1924599.
- [3] C. Pramerdorfer, "An Introduction to Processing and Music Visualization," p. 17.
- [4] "VU meter - Wikipedia." https://en.wikipedia.org/wiki/VU_meter (accessed Mar. 27, 2020).
- [5] J. Russell, "A Circumplex Model of Affect," *J. Pers. Soc. Psychol.*, vol. 39, pp. 1161–1178, Dec. 1980, doi: 10.1037/h0077714.
- [6] R. E. Thayer, *The biopsychology of mood and arousal*. New York, NY, US: Oxford University Press, 1989.
- [7] N. Kaya, "Relationship between color and emotion: a study of college students," p. 13.
- [8] "(PDF) Feel the Moosic: Emotion-based Music Selection and Recommendation," *ResearchGate*.
https://www.researchgate.net/publication/334084112_Feel_the_Moosic_Emotion-based_Music_Selection_and_Recommendation (accessed Jan. 30, 2020).
- [9] yi-hsuan Yang, C. Chu Liu, and H. Chen, "Music emotion classification: A fuzzy approach," 2006, pp. 81–84, doi: 10.1145/1180639.1180665.
- [10] Y.-H. Yang and H. H. Chen, "Machine Recognition of Music Emotion: A Review," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 3, pp. 1–30, May 2012, doi: 10.1145/2168752.2168754.
- [11] A. Gabrielsson, "Emotion perceived and emotion felt: Same or different?," *Music. Sci.*, vol. 5, no. 1_suppl, pp. 123–147, Sep. 2001, doi: 10.1177/10298649020050S105.
- [12] "Elements of Music." <http://accounts.smccd.edu/mecklerd/mus250/elements.htm> (accessed Feb. 01, 2020).
- [13] Y. Ephraim and H. L. V. Trees, "A signal subspace approach for speech enhancement," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 4, pp. 251–266, Jul. 1995, doi: 10.1109/89.397090.
- [14] "Color and Emotion: The Psychological Appeal of Colors," *Breathing Color Blog*, 21-Sep-2015. <https://blog.breathingcolor.com/color-and-emotion/> (accessed Jan. 20, 2020).
- [15] "Color meanings and the art of using color symbolism," *99designs*, 21-Aug-2019. <https://99designs.com/blog/tips/color-meanings/> (accessed Jan. 20, 2020).
- [16] A. Choungourian, "Color Preferences and Cultural Variation," *Percept. Mot. Skills*, vol. 26, no. 3_suppl, pp. 1203–1206, Jun. 1968, doi: 10.2466/pms.1968.26.3c.1203.
- [17] W.-L. Chang and H.-L. Lin, "The impact of color traits on corporate branding," 2010.
- [18] M. Gobe, M. Gobé, and S. Zyman, *Emotional Branding: The New Paradigm for Connecting Brands to People*, 1st edition. New York: Allworth Press, 2001.
- [19] N. A. Nijdam, "Mapping emotion to color," 2005.
- [20] L. Bialoskorski, J. Westerink, and E. L. van den Broek, "Mood Swings: Design and evaluation of affective interactive art," *New Rev. Hypermedia Multimed.*, vol. 15, pp. 173–191, Aug. 2009, doi: 10.1080/13614560903131898.
- [21] "design," *Lemuria*. <https://cm4group.wordpress.com/tag/design/> (accessed Mar. 11,

- 2020).
- [22] P. Galanter, "What is Generative Art? Complexity Theory as a Context for Art Theory," p. 21.
- [23] G. W. Smith, "An Interview with Frieder Nake," *Arts*, vol. 8, no. 2, p. 69, Jun. 2019, doi: 10.3390/arts8020069.
- [24] "Homage à Paul Klee, 13/9/65 Nr.2 | Nake, Frieder | V&A Search the Collections," *V and A Collections*, 02-Oct-2019. <http://collections.vam.ac.uk/item/O211685> (accessed Oct. 02, 2019).
- [25] "How To Draw With Code | Casey Reas - YouTube." https://www.youtube.com/watch?v=_8DMEHxOLQE (accessed Jan. 04, 2020).
- [26] Æ. B. Fry, C. Reas, and B. Fry, DOI 10.1007/s00146-006-0050-9 ORIGINAL ARTICLE. .
- [27] C. Menezes, "Casey Reas – interview: 'There is an increased understanding that software is central to our lives,'" *Studio International - Visual Arts, Design and Architecture*. <https://www.studiointernational.com/index.php/casey-reas-interview-computer-art-coding> (accessed Jan. 04, 2020).
- [28] "Casey Reas," *Arts at MIT*. <https://arts.mit.edu/cast/symposia/being-material/participants/casey-reas/> (accessed Jan. 04, 2020).
- [29] "REAS.com is a database for Casey REAS." https://reas.com/compendium_b_p/ (accessed Jan. 04, 2020).
- [30] W. Joel, "Zach Lieberman's interactive art creates moments of wonder and surprise," *The Verge*, 10-May-2019. <https://www.theverge.com/2019/5/10/18564182/zach-lieberman-interactive-art-animation-movement-openframeworks-art-club> (accessed Jan. 06, 2020).
- [31] "About – LIA." <https://www.liaworks.com/about/> (accessed Jan. 06, 2020).
- [32] R. Hiraga, F. Watanabe, and I. Fujishiro, "Music learning through visualization," in *Second International Conference on Web Delivering of Music, 2002. WEDELMUSIC 2002. Proceedings.*, Darmstadt, Germany, 2002, pp. 101–108, doi: 10.1109/WDM.2002.1176199.
- [33] "The Dream of Color Music, And Machines That Made it Possible." <https://www.awn.com/mag/issue2.1/articles/moritz2.1.html> (accessed Jan. 11, 2020).
- [34] M. Franssen, "THE OCULAR HARPSICHORD OF LOUIS-BERTRAND CASTEL," p. 63.
- [35] D. W. Fourney and D. I. Fels, "Creating access to music through visualization," in *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*, Toronto, ON, Canada, 2009, pp. 939–944, doi: 10.1109/TIC-STH.2009.5444364.
- [36] "Charles Germain de Saint Aubin," *Wikipedia*. 20-Feb-2019, Accessed: 16-Mar-2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Charles_Germain_de_Saint_Aubin&oldid=884209772.
- [37] "How to trip out with the iTunes visualizer | Macworld." <https://www.macworld.com/article/3074173/how-to-trip-out-with-the-itunes-visualizer.html> (accessed Jan. 11, 2020).
- [38] "Everyday Listening - Sound Art, Sound Installations, Sonic Inspiration." <http://www.everydaylistening.com/articles/tag/visualization> (accessed Jan. 11, 2020).
- [39] "Rafael Lozano-Hemmer - Project 'Voice Array.'" https://www.lozano-hemmer.com/artworks/voice_array.php (accessed Jan. 11, 2020).
- [40] P. Dallow, "Representing creativeness: practice-based approaches to research in creative arts," *Art Des. Commun. High. Educ.*, vol. 2, no. 1, pp. 49–66, Mar. 2003, doi: 10.1386/adch.2.1.49/0.

- [41] L. Candy, "Practice Based Research: A Guide," *CCS Rep.*, vol. V1.0, 2006, Accessed: 21-Jul-2013. [Online]. Available: <http://www.creativityandcognition.com/wp-content/uploads/2011/04/PBR-Guide-1.1-2006.pdf>.
- [42] R. Andrews, E. Borg, S. B. Davis, M. Domingo, and J. England, *The SAGE Handbook of Digital Dissertations and Theses: SAGE Publications*. SAGE, 2012.
- [43] R. L. Skains, "Creative Practice as Research: Discourse on Methodology," *Media Pract. Educ.*, vol. 19, no. 1, pp. 82–97, Jan. 2018, doi: 10.1080/14682753.2017.1362175.
- [44] "(Des)Ordres | Molnar, Vera | V&A Search the Collections," *V and A Collections*, 04-Mar-2020. <http://collections.vam.ac.uk/item/O1193781> (accessed Mar. 04, 2020).
- [45] "Element of Design Line." <http://www.wcs.k12.mi.us/cousino/wcsart/art%20foundatons%20site/line.html> (accessed Mar. 08, 2020).
- [46] "5 Types of Line in Art, Their Meaning And When To Use Them," *YourArtPath*, 23-Oct-2019. <https://yourartpath.com/types-of-line-in-art-meaning> (accessed Mar. 08, 2020).
- [47] "curveVertex() \ Language (API) \ Processing 3+." https://processing.org/reference/curveVertex_.html (accessed Mar. 08, 2020).
- [48] "Introduction to Catmull-Rom Splines." <http://www.mvps.org/directx/articles/catmull/> (accessed Mar. 08, 2020).
- [49] "RGB color model," *Wikipedia*. 04-Mar-2020, Accessed: 09-Mar-2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=RGB_color_model&oldid=943849582.
- [50] F. García-Lamont, J. Cervantes, A. Lopez-Chau, and L. Rodríguez, "Segmentation of Images by Color Features: A Survey," *Neurocomputing*, vol. 292, Mar. 2018, doi: 10.1016/j.neucom.2018.01.091.
- [51] Andrey, "What is a color model?," *Script Tutorials*. <https://www.script-tutorials.com/what-is-a-color-model/> (accessed Mar. 09, 2020).
- [52] "HSL and HSV," *Wikipedia*. 22-Feb-2020, Accessed: 09-Mar-2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=HSL_and_HSV&oldid=942114723.
- [53] L. Barrière, "Exploring Randomness in Digital Art," p. 6.