

Title	VIBO : a service for augmenting self-expression through interactive music experiment
Sub Title	
Author	Ren, Luchi(Ōkawa, Keiko) 大川, 恵子
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2021
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2021年度メディアデザイン学 第888号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002021-0888

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

Master's Thesis
Academic Year 2021

VIBO: A Service for Augmenting Self-expression
through Interactive Music Experiment



Keio University
Graduate School of Media Design

Luchi Ren

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

Luchi Ren

Master's Thesis Advisory Committee:

Professor Keiko Okawa (Main Research Supervisor)

Senior Assistant Professor

Junichi Yamaoka (Sub Research Supervisor)

Master's Thesis Review Committee:

Professor Keiko Okawa (Chair)

Senior Assistant Professor Junichi Yamaoka (Co-Reviewer)

Professor Matthew Waldman (Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2021

VIBO: A Service for Augmenting Self-expression through Interactive Music Experiment

Category: Design

Summary

Music is an important part of life for many people, especially in this new normal era when everyone is living a more isolated life. As music lies deeply in human being's nature, our desire towards music is not only about consuming but also about expressing. While so many people are trying to express their music expression desire, these desire is difficult to be fulfilled in a home setting, which enhances the feeling of loneliness. Therefore, this research proposes a service called VIBO aiming at augmenting the music self-expression via reactive musical functions.

The experience that VIBO provides is basically to let users get more automatic response from both the music and the environment to enhance their musical expression through capturing their emotional change which is expressed unconsciously by their facial expression, body movement and voice. This research aims at people living alone or in a quarantined status.

The result shows that VIBO can provide a positive effect on the user's emotion status and its potential to be developed in many scenarios.

Keywords:

emotional expression, music, COVID-19, service design, kinect

Keio University Graduate School of Media Design

Luchi Ren

Contents

Acknowledgements	vii
1 Introduction	1
1.1. Background	1
1.1.1 Music as a Well-being Solution	1
1.1.2 Human Nature of Music Expression	2
1.2. Research Proposal	2
1.3. Contribution	3
1.4. Research Questions	3
1.5. Thesis Structure	3
2 Literature Review and Related Works	5
2.1. Literature Review	5
2.1.1 Music and Emotion Expression	5
2.1.2 Unconscious Emotion Expression	6
2.1.3 Music Listening Habits	8
2.2. Related Works	11
2.2.1 Tangible Installation - Aura	11
2.2.2 Mobile Application - Cove	12
Notes	13
3 Preliminary Study	14
3.1. Interview	14
3.1.1 Interview 1	14
3.1.2 Interview 2	15
3.1.3 Insights	18
3.2. Workshop	18

3.2.1	Basic Flow	18
3.2.2	Participant	19
3.2.3	Emotion Indicator	20
3.2.4	Music Program	20
3.2.5	Feedback	22
3.2.6	Discussion & Insights	22
Notes	24
4	Design and Implementation	25
4.1.	Design Concept	25
4.2.	Prototyping	26
4.2.1	User Flow	26
4.2.2	Functions	27
4.2.3	Emotion Classification	28
4.2.4	User Interface	29
Notes	30
5	Evaluation	31
5.1.	User Test	31
5.1.1	Participants	31
5.2.	Music list	32
5.2.1	Set-up	32
5.2.2	Other Preparations	33
5.3.	Flow	34
5.3.1	Function by Function Introduction	36
5.4.	Data Analysis	37
5.5.	Semi-structured Interview	39
5.6.	Conclusion	44
6	Discussion	47
6.1.	Limitation	47
6.2.	Future Work	48
	References	51

Appendices	56
A. Pre-study Interview 1	56
B. Pre-study Interview 2	57

List of Figures

2.1	Picture of Aura	11
2.2	Picture of Cove	13
3.1	Mizu's quarantine hotel room	16
3.2	Miyu's Quarantine Activities	17
3.3	Proposed Workshop Flow	19
3.4	Practiced Workshop Flow	20
3.5	Relationship between Listening to Music and Mood Change . . .	21
3.6	Workshop Setting	21
3.7	Screenshot of Participant's Reaction	22
4.1	Proposed User Flow	27
4.2	Proposed User Interface	30
5.1	Ustertest Setting	34
5.2	Applied functions	35
5.3	User-test Structure	36
5.4	Comparison chart of "Surprise" & "Happy" values	38
5.5	Function by Function test: Value of "Happy" & "Surprise" . . .	40
5.6	Sample of interface change: At the beginning	44
5.7	Sample of interface change: Approaching the end	45
6.1	Switch for turn it off	48
6.2	Integrated Product	49
6.3	Usage simulation	49

List of Tables

3.1	Evaluation Form	22
5.1	User 1: Mizu	41
5.2	User 2: Viola	41
5.3	User 3: Chou	41

Acknowledgements

I would like to thank my supervisor, Keiko Okawa, for her guidance through each stage of the process. Thanks Keiko Sensei for every opportunities she gave me to participate in many research activities, from which I gained not only learned a lot from excellent seniors but also had many happy time with Sensei and my Global Education family. I really cherish every memory we had together.

I would also express my appreciation to Professor Yamaoka Junichi and Senior Assistant Professor Marcos Sadao Maekawa for every advice on how to develop my research, especially when I met the bottleneck. I gained a lot of new visions and really enjoyed every discussion with them.

I would also like to say a big "Thank you" to my Global Education peers. Thanks my lovely girls for 2019 Autumn batch. Thank you for listening to all my worries and giving me so many useful advice. Thanks my wonderful seniors, Leandro, Silmi, Isa, Hera, and Moe who had been encouraging me all the way along.

Above ground, I would like to express my gratitude to my parents and Rico who always give me power to be positive, trust to move on, the strongest backing and love, without which I will never overcome the difficulties in these two years. Thank you.

Chapter 1

Introduction

1.1. Background

1.1.1 Music as a Well-being Solution

In the year 2020, COVID-19 pandemic broke changed people's daily life rapidly and globally (Sundarassen et al. 2020). Going against the grain of globalization, people are living a more isolated life with new social norms and regulations for curtailing the spread of COVID-19 virus, such as social distance and lockdown. Negative emotions greatly increase while people's satisfaction towards their life decreased to a great extent (Mazza et al. 2020, ?, ?, ?).

Music has long been proved as an effective method for emotion regulation and increasing happiness (DeNora 2000, Van Goethem and Sloboda 2011). The positive role of music was further demonstrated during the COVID-19 period. Ziv and Hollander-Shabtai reported that there was an increase in the consumption of music as well as an increase in the intensity of emotional reaction to music during the lockdown period (Ziv and Hollander-Shabtai 2021). Moreover, researches also reports that music-listening is an effective strategy for coping with negative affect, especially compared with other media such as video games and TV shows (Sachs et al. 2021).

As someone who is also undergoing a new change in social norms of behaviour, I have also spent the past year listening to music to alleviate the boredom of life. For example, I usually listening to the music while cleaning the floor with a Hoover. When the music comes a climax and I feel the strong resonance with the song, I sometimes cannot help but imagine the Hoover in my hand as a microphone, humming with it, or swaying my body to the rhythm of the music.

With more and more people trying to adapt to the new social norms brought by

COVID-19 situation, there are more new services coming out to facilitate this shift process. A service which could help people have a more interesting music-listening experience is also considerable to be meaningful in this situation.

1.1.2 Human Nature of Music Expression

Human being's communicative musicality is innate (Hargreaves et al. 2002, Littleton et al. 2004). In Malloch and Trevarthen's research, infants used their innate communicative musicality to interact with their caregivers which researchers suppose to be the building blocks of more temporal arts in the future (Malloch and Trevarthen 2018). However, despite expressive musicality lies in human's nature, the nature is difficult to be realized in a social and practical context. Musicians can utilize instrument or digital composing software to express themselves through music. However, as for amateurs who have no music education background, this kind of desire always stops at a limited level, such as imaging or imitating. Therefore, unlike the creative musical expression of musicians whose expression can be appreciated by the audience and responded to accordingly, the musical expression of ordinary people in their daily lives often tends to be ignored.

1.2. Research Proposal

In light of the background discussed above, a service called VIBO is proposed for people living alone to enjoy a more interactive music-listening experience in this paper. VIBO provides users an experience where they can interact with music through their own unconscious emotional expressions. Through this service, the users can receive visual response based on the emotional signals they expressed, which made a more augmented musical experience. The service VIBO consists of two parts, a tangible mobile phone peripheral of small size which enables interactive functions, and a smartphone application which provides an interface to help controlling all the interactive functions on the product. With the facilitation of VIBO, every individual music listener is supposed to interact with the music they are listening to and have a happy experience in this New Normal Era.

1.3. Contribution

The main contributions of this work can be summarized as follows:

- Creation of a fun tool to enjoy music in a more augmented way.
- Provides a solution that also allows music lovers to fully interact with music and express themselves.
- Provides a solution for people who live alone for long periods of time to be more aware of their mood changes.

1.4. Research Questions

The motivation of this study is to enhance the well-being through daily music experience of people who lives alone, and ultimately to achieve the goal where all people can enjoy music in a more augmented way. Therefore, three research questions are posed in this research:

- Research Question 1: Does real-time musical and environmental changes makes the music listening experience more immersive and enjoyable?
- Research Question 2: Does the visualization of the interaction let the users have a better understanding of their mood condition?
- Research Question 3: In what a music experience augmenting service can be implemented smoothly into target user's everyday life? (User experience)

1.5. Thesis Structure

This thesis is divided into 5 chapters as follows:

- Chapter 1:
This chapter serves as the introduction of the whole research which includes background, research proposal, contribution and research questions.

- Chapter 2:

This chapter will build up the whole theoretical framework in this research. In the literature review part, music theories, psychological theories, habits of music listeners will be discussed. In the related work part, both tangible installations and application software will be reviewed.

- Chapter 3:

This chapter introduces the preliminary study of an interview and an online workshop to try out some ideas. The results and findings will be applied to the prototyping and evaluation section.

- Chapter 4:

This chapter introduces the design concept and the design process of VIBO. The implementation part involves hardware, software and graphic design.

- Chapter 5:

This chapter talks about the user test and its evaluation. The two sets of evaluation was conducted. The first one is the analysis of data from Kinect sensor. The second one is a semi-structured interview based on a pre-prepared evaluation form.

- Chapter 6:

This chapter introduces about three limitations of this research and the thoughts about future work.

Chapter 2

Literature Review and Related Works

2.1. Literature Review

2.1.1 Music and Emotion Expression

This research would like to design a service that facilitates average music listeners to become the subject of their musical emotional expression in their music listening process. In order to frame a relatively comprehensive vision, this section will start with the emotional expressions of music creators and subsequently shift the focus to music consumers, thus clarifying the direction of focus in the design of this research.

There are many literature focusing on the musical expression of emotion from the composers and performers' perspective. For example, in many cases in the field of music appreciation, sharp contrasts between "long" and "short" notes are supposed to be happy while high pitch with fast tempo is analyzed as fear (Juslin and Laukka 2004). Boccherini's Adagio of the String is interpreted to provide audience a feeling of being side-by-side with the orator and "exchanging disjoint sweet nothings in a twilit garden" (Le Guin 2005). Besides classical music which are designed to be listened attentively and comprehended in a certain pattern (Olteteanu 2010), impromptu jazz performers are also bringing their own interpretations and emotions to bring the written music back to live (Scruton 1999).

However, when regarding music as a type of media which conveys message from one side to the other, how the audience consume music also matters. The basic model of communication consists of the process of encoding and decoding and involves three parties which are message sender, message itself and message receiver (Hall 2003). Music is not an exception in this system. The value that

composers and performers created needs to be realized by their listeners (Graham 2005). Roe suppose that rather than passively listening to music, music listeners are more likely to actively use music to serve some emotional needs (Roe 1985).

Many mainstream music products on the market are providing the audience a passive music experience, such as Net-ease Cloud Music or Spotify. However, based on the researches discussed above, there are also cases where listeners consume music pieces actively. In other words, instead of emotion being directly affected by the emotional signals in a piece of music that the composers would like to convey, there are also grounds on which music arouses the emotion to some extent while audience themselves are also able to comprehend the music in their own way (Robinson 1994), confer their emotional condition to the music and finally express their emotion in a musical pattern, such as moving the body in a certain tempo or the facial expressions change in response to the music rhythm (Seitz 2005). Since a part of the pleasure of enjoying music pieces comes from our own emotional experience (Robinson 2007), this research infers that if the active emotional interaction of music listeners is enhanced during the process of listening to music, their music experience will also be improved.

This research would like to utilize this pattern of active emotional expression in music listening activity to design a service which turns music consumers from passive audience to active audience.

2.1.2 Unconscious Emotion Expression

Based on the theoretical discussion in the last section, listener of music can gain a better recognition of their current mood in the process of fitting their mood with that of the music. However, manual operations during the listening process, such as switching the playlist, is preferred (Kamalzadeh et al. 2012). In order to design a service which do not bother the existing behavioral pattern of the target users, unconscious emotional expression will be focused in this section. By the end of this section, the most suitable pattern which can be utilized in the design of this research proposal will be discussed.

Although in the previous section we mentioned that people are active in their use of music to express emotions, not all active expressions of emotion are conscious. In this research, Öhman's discussion of conscious and unconscious (Öhman 1999)

is used to distinguish user's behavior. To define whether a behavior is conscious, one critical criterion is whether it is perceptible or accessible to introspection (Öhman 1999). Therefore, when an external happening is processed consciously by someone, these events are likely to be noticed or able to be recalled afterwards, while an unconscious processing can also be something that people response to the outward but is difficult to be kept or memorized (Owren et al. 2005).

In people's daily life there are many moments when people express their emotion consciously or unconsciously. For example, people use language to express happiness when feeling fulfilled or being cared. Besides verbal emotion expression, people also utilize non-verbal channels to express their feeling, such as facial expression, body movements, postures and touch (Collier and Collier 2014). Although there are cultural differences, we can sometimes see people use the V-shape gesture to express their happiness of success in some cultures.

Both V-shape gesture and verbal emotion expressions are something that people express intentionally. There are also times when people are expressing their emotional status but they do not realize. For example, when people feel surprised, tiny facial changes, such as raising eyebrows, eye-widening and jaw dropping are supposed to happen regardless of people's cultural background (Hager et al. 2002). Another example is when people is suffering from strong grief, his or her may get into the condition where the muscles flaccid or heads hangs on the contracted chest (Darwin 1965). These reactions happen in an unconscious pattern because someone does not choose to do it.

In this research, the unconscious emotion expression is mainly supposed to be captured from facial expressions while some body movement expressions are also captured to serve as the facilitation. There are two reasons for this decision. The first reason is that facial expression is innate. Facial expression has been argued to be a part of a person's constitutional makeup which can be controlled only to a limited extent (Collier and Collier 2014). Freud also mentioned this aspect of facial expression, saying, "If his lips are silent, he chatters with his fingertips, betrayed oozes out of every pore" (Elson 2001). Even though some facial expressions will eventually be masked intentionally by oneself, it will still show up at the very beginning as "micro-momentary facial expressions" (Ekman and Friesen 1969).

The second reason is facial expression can express a wide range of emotions. It is supposed to be able to express over nine kinds of emotions, including happiness, sadness, fear, anger, interest, surprise, disgust, contempt and shame (Izard 1977). Also, with the aid of body movement, more accurate and dedicated emotional status can be analyzed (Collier, 2014). Even though not all body movement is expressive, it sometimes can help hint at a more dedicate or complex emotional status than verbal description (Collier and Collier 2014).

Therefore, facial expression is supposed to be utilized to be the primary emotion signal source while body movements are used as a complement for facial expression. The implementation of this idea will be introduced in detail in the Chapter 3.

2.1.3 Music Listening Habits

Before discussing about which approach to realize the interaction between users and music in detail, in order to construct a clear vision about the target users of this research, this section would like to briefly review the general music listening habits in modern society. especially the behavioral and mental changes when listening to the music before and after the COVID-19 pandemic.

Music has an influence in the development of identity and individuals think of it as an essential piece of what their identity is and plays an important role in creating happiness and relaxation (DeNora 2000, Van Goethem and Sloboda 2011). With the development of internet and technologies, music is becoming ubiquitous in people's daily life. In our daily life, many people use music frequently as a background music for different activities, like studying, working, doing sports, or housework (Sloboda and O'neill 2001). Despite the fact that music may not be the focus of attention in these activities, individuals have their way of picking up and matching music to their different cognitive and emotional needs (Ziv and Hollander-Shabtai 2021). The process of matching their mood status to various music is also supposed to be the process of have a better recognition of the listener's current affect (Skånland 2013). Private and portable music services, such as MP3s or mobile phone, provide users more convenience in choosing the music that fit their mood at the moment (Skånland 2013). This enhances one's insight of his or her emotional intelligence which plays an important role in

mental health (Csikszentmihalyi and Csikszentmihalyi 2006). This research also seeks to capitalize on the strengths of portable music products to further enhance the emotional cognitive experience of listening to music, thereby enhancing the user's sense of well-being. This research also seeks to capitalize on the strengths of portable music products to further enhance the emotional cognitive experience of listening to music, thereby enhancing the user's sense of well-being.

Besides mental habits discussed above, physical habit is also an important perspective to be considered when doing design. The following part will review it from three aspects, which are music listening frequency, music listening devices and music listening scenario.

- **Frequency**

Statistics show a high frequency of music consumption among the global population. Take the three countries with the world's top three populations as an example. In 2019, a year before global COVID-19 pandemic, according to a research done by Statista Research Department¹, over half of the population in the US listens to music every day. While in the same year, it is reported that people spend about 17.7 hours per week on music listening in China, 19.1 hours per week in India². During the time when COVID-19 pandemic break out and caused lockdown and quarantine all over the world in 2020, researches reported an increase in music consumption in many regions (Ziv and Hollander-Shabtai 2021, Cabedo-Mas et al. 2020, Ribeiro et al. 2021).

- **Device**

In the past ten years, personal digital music library service as well as subscription based on-demand music streaming service is becoming more and more popular (Kamalzadeh et al. 2012). Spotify alone has 356 million users³. However, before 2020, according to the static by ifpi⁴ in 2019, radio which own 29% of music listener was still the most popular device for music listening while smartphones and computers only had 27% and 19% of the population. This ranking changed with the interruption of COVID-19, the percentage for using digital devices to listening to music was reported to have crossed the 50% threshold and surpassed that of radio for the first

time in the US⁵. Moreover, on-demand listening on apps such as Spotify or Apple Music was increased by 20% in the first two month of 2020⁶, which indicates that more people prefer to use digital devices to listen to music in COVID-19 situation.

- **Scenario**

Music is present in many scenarios of people's daily life, for example music classes, night club or BGM in the department store. Among all these scenarios, working, socializing, driving and taking public transportation were the four main activities taken when listening to music (Lamont and Webb 2010) while this order changed in to commuting, working, doing house work and exercising in a research in 2012 (Kamalzadeh et al. 2012). However, with the impact of COVID-19, "Working From Home" is becoming more and more popular. In 2020, there are 16% of the companies were practising fully remote work patterns while if "hybrid working companies" are taken into consideration 56% of the companies around the world allows remote work⁷. In the United States alone, this phenomenon has directly resulted in a reduction of 60 million hours of commuting time per day⁸. In the meanwhile, these commuting hours are going to other parts of life, for example more working hours within the same work-day spans⁹. Therefore, this research infers that of the four activities mentioned above that would primarily be accompanied by music, the proportion of time spent at work and commuting both has changed significantly during the COVID-19 pandemic situation.

Although many studies have explored the psychological and behavioral characteristics of music listeners, in order to gain a more detailed and subjective understanding, this study conducted a pre-study to deepen the understanding of users before entering the formal design process. This research will combine the research findings discussed above and the findings from the pre-study of this research to try to design a service that matches the user's music listening. The user experience design will be introduced in detail in Chapter 3.



Figure 2.1 Picture of Aura

2.2. Related Works

In this section, two related works will be introduced and analyzed. The two related works are of three types: tangible installation, software and a combination of hardware & software. After the discussion of these related works a more clear direction of how to design the service in the proposed research will be constructed at the end of this section.

2.2.1 Tangible Installation - Aura

Aura¹⁰ is an audiovisual installation created by Studio Nick Verstand presented on Dutch Design Week in 2017. In the design of this installation, biosensors are utilized to gather emotion data from audience which ultimate turn into pulsing light compositions. Walking into the dark space where Aura is installed, there was music playing in order to induce emotions. Then the wearable sensors on visitors' bodies capture basic data information such as heart-rate variability and turn it into lights of different shapes and colors to reinterpret the emotion.

This design made it possible to see each other's emotion as well as give an immersive experience of seeing one's emotion. Aura inspired the present research from two aspects. The first one is the materialisation of internal emotion into

external physical presentation. The second one is the way of using color and shape to present the internal feeling. As we discussed above, listeners will often actively listen to the music and adjust their emotions to match the music. This device offers the possibility for the listener to learn more about how he or she adjusts emotionally and resonates with the music from more perspectives. Thus completing a self-communication process of expressing emotions (encoding) and understanding the expression of emotions (decoding).

2.2.2 Mobile Application - Cove

Cove¹¹ is a mobile application designed to improve users' mental health by making music published by Humane Engineering Ltd. It allows people without any professional music education to create music based on their moods. The basic user flow of Cove is that 1) users define their current mood using pre-prepared adjectives 2) an musical canvas is then created on which users place blocks randomly to compose. 3) then the user can save the music piece with text descriptions and pictures.

It is interesting that on which part the user put the music block, the music comes out make sense, as if telling people no matter how they feel, their feeling is justified. Moreover, as an application designed especially for mental health usage, this application has been approved by several medical institutions such as NHS and Good Health.

The point that Cove inspired me most is that it allows vague definition for complex emotions. Wessman and Ricks' study shows that the richness and diversity of feelings varies people to people, and named this "affective complexity" (Lindquist and Barrett 2008). Therefore, it is difficult for everyone to define their emotion at a precise level. Also, forcing people to define emotion may also cause mental burden on users. As an application for maintaining mental health, this application provide a good solution to turn language definition into shapes, colors and sounds, which touches the recognition of emotion from more sensitive perspectives.

However, as an application for mental health, functions of Cove are designed based on psychological considerations. The proposed design would like to focus more on entertaining usage while mental well-being as an additional possible value. Therefore, a design that fits users' music consuming behavioral habits and gives

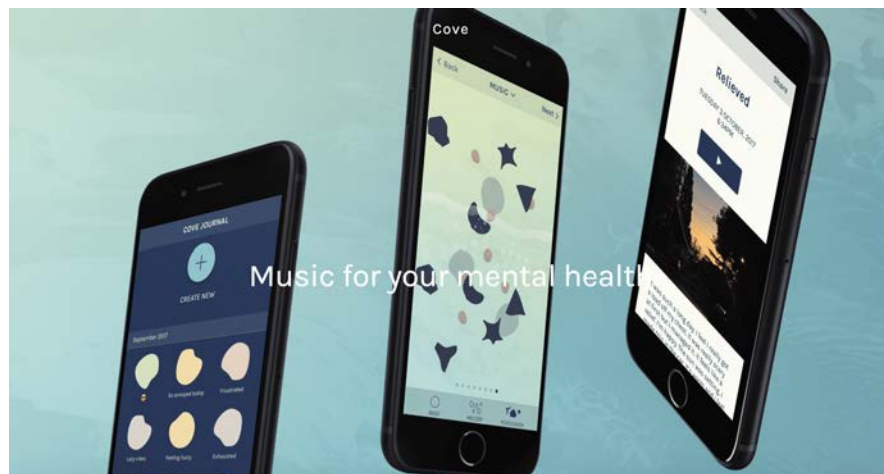


Figure 2.2 Picture of Cove

users fun is also required in the present research.

Notes

- 1 <https://www.statista.com/statistics/749666/music-listening-habits-age-usa/>
- 2 <https://www.ifpi.org/wp-content/uploads/2020/07/Music-Listening-2019-1.pdf>
- 3 <https://newsroom.spotify.com/company-info/>
- 4 <https://www.ifpi.org/wp-content/uploads/2020/07/Music-Listening-2019-1.pdf>
- 5 edisonresearch.com/digital-device-listening-crosses-the-50-threshold-during-covid-19-disruptions/
- 6 <https://www.intheknow.com/post/music-listening-pandemic-covid19-nielson-billboard-spotify/>
- 7 <https://resources.owllabs.com/state-of-remote-work>
- 8 <https://voxeu.org/article/how-americans-use-time-saved-working-home>
- 9 <https://hbr.org/2020/12/where-did-the-commute-time-go>
- 10 <https://www.dezeen.com/2017/11/25/aura-installation-translates-emotions-into-beams-of-light-studio-nick-verstand-dutch-design-week/>
- 11 <http://www.cove-app.com/#about>

Chapter 3

Preliminary Study

The pre-study includes two parts. The first part is an in-depth qualitative study of people’s music listening behavioral habits in COVID-19 period. Two people were interviewed remotely. The questions are prepared based on the researches discussed in Chapter 2. The second part is a small workshop to test some possible functions and the general design concept in advance. Valuable feedback and insights were gained through this mini workshop.

3.1. Interview

3.1.1 Interview 1

The first interviewee, Mizu, is a 20-year-old college student from China (Appendix B). The interview was conducted on the 5th day of her collective quarantine life in China in February 2021. The quarantine regulation in China is very strict especially for people going back to China from foreign countries. All arrivals in China must be in mandatory quarantine at a hotel room for 14 days and are not allowed to set foot outside their rooms during the quarantine period otherwise the quarantine period will be recalculated. During quarantine, the quarantined person will not come into direct contact with anyone, and all food and water will be distributed without contact by being placed at the door of the quarantined room at the specified time.

For the case of the first interviewee, she flew from Japan to China to spend her spring vacation with her family. During the interview, her mental state, daily routine, habits about music listening, her expectation for music services in her quarantine days were talked about.

The most problem she's facing is boredom and lack of passion brought by the boredom. Music was an important part of her quarantine life she mentioned. Since she's alone all the time, sounds from the speaker made her feel not so lonely. Music is her companion when doing many activities, such as waking up, doing homework, washing clothes etc. "In particular", she mentioned, "I will definitely listen to the song while I am in the shower. I love the reverb of listening to singing voices in the bathroom. "

As for the music services she's using, Netease Cloud Music is the only one, since she has created many personal playlists on it and been using it for many years. Sometimes functions such as Heartbeat Playlist (a function that the system automatically generates a recommended music list based on the user's historical preferences) and Karaoke are used but not often, because these functions sometimes require many extra operation that interrupts the general music listening experience.

She's expecting a service that could "let music play with her" from present music services. "Sometimes when I listen to a song, I can remember the scene when I was listening to it under stage during the concert. Then I recall more about the concert memories and even want to re-watch the video of that concert. But when the music ends and I switch to a completely unrelated song in the single playlist, I get very frustrated or manually stop the new song.", she said. "But listening to that song all the time is not what I want either. The touch and surprise from the stage cannot be sensed from recorded pieces."

3.1.2 Interview 2

The second interviewee was a 26-year-old Japanese girl (Appendix B). By the time she was doing this interview, she has experience Japan's state of emergency twice. In response to the Japanese government's declaration of a state of emergency, measures were taken to prevent the spread of the new strain of coronavirus, including requests for companies to work from home and for restaurants to shorten their hours of operation. In the midst of this situation, the government requested that people entering and returning to Japan recuperate at home or in hotels for 14 days. Within the period of home quarantine, people are allowed to live with their families but not to leave their homes.



Figure 3.1 Mizu's quarantine hotel room

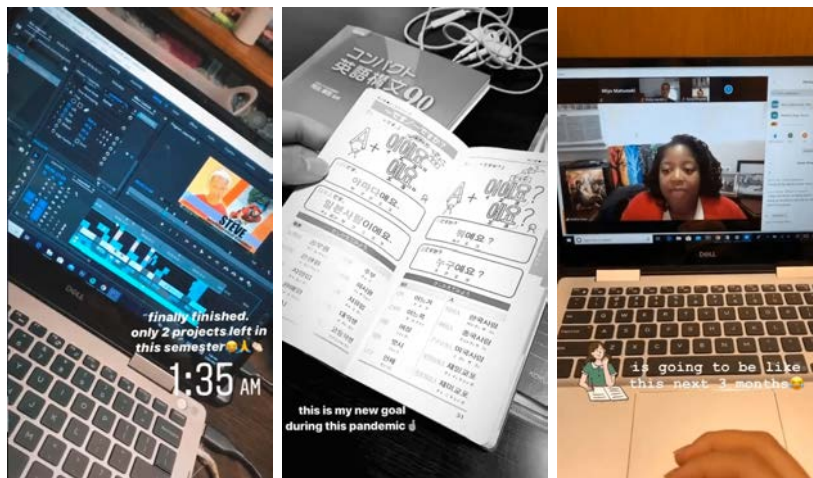
Miyu is an international student studying film in the United States. She returned to Japan after the first state of emergency was declared. She experienced 14 days of strict home quarantine and nearly 2 month of stay-at-home life. She described her life of that at time as:

"I was with my family, so I didn't feel particularly lonely, but it was still boring. In this boring time without any surprise, I tried to do many things I don't usually do, like organizing the attic and renewing the garden."

The biggest problem Miyu faced is her negative thought of self-denying. Three reasons led to this problem. The first one is that she's not doing enough valuable input and output at home compared with at school. The second reason is that the pandemic interrupted her future career plan. The third reason is that she had too much free time to overthink about everything.

To overcome this situation, she tried mainly two strategies. The first thing is to fill her schedule with planned study time. The second thing is to distract herself with activities that allow you to relax. In both strategies, music played an important role.

Firstly, she has the habit of listening to music while she studies. She listens to classical music when she is doing something that requires a lot of brain power, such as learning English. She listens to R&B when she is doing mechanical and repetitive work, and her attention is sometimes directed to the task at hand and



(a) Studying skills (b) Studying Korean (c) Online Class

Figure 3.2 Miyu's Quarantine Activities

sometimes to the music. This will enable her to study for a long time without feeling tired.

The second point is that singing while taking the shower helps her empty her brain and have a good sleep. She especially likes to use Amazon Music's lyric scrolling and music video functions while taking a bath. With the aid of the echo in the bathroom, she feels like she's enjoying karaoke which is something she cannot enjoy in that stay-at-home condition.

Another function she loves most in her quarantine days is the Shuffle function of Amazon Music. She get to know many new songs of her taste through this function. She said, "This feature sometimes pushes me songs that I heard inadvertently when I was studying in the US. The atmosphere conveyed by the music will remind me of the happy times before the new crown broke out, so even if it's a song I didn't know before, I'll still enjoy listening to it."

Once a song made her think of a nightclub she had been to, and it made her want to go again. So she would collect that music and listen to it when she was bored.

3.1.3 Insights

Different from the statistics in Chapter 2, these two interviews provide a more subjective and in-depth emotional view of the target users. Based on two interviews, a more vivid image of the target user was sketched out, which can be concluded into the following 3 points:

- Young people whose richness of life has been seriously affected by the new social norms of behavior brought by COVID-19, such as social distance or state of emergency.
- People who use music listening as a positive method for emotion regulation and seek for pleasure or relaxation from music listening behavior.
- People who want to fly out of the current real life and go deeper in the imagination of the world in music.

To further analyse the needs of target users, the following key words are raised to describe: 1) Enjoyable 2) Immersive 3) Surprising 4) Convenient to use 5) Interactive but not interrupting 6) Cares about users' emotion (mood). Therefore, a music service that can automatically interact with user's emotional change is proposed in this research.

Based on the analysis above, a rough structure of design was created and a small workshop was conducted to test it's functionality.

3.2. Workshop

In this workshop, some of the ideas generated based on the preliminary researches were briefly tested, including the internal logic, functionality and user flow of the product. In order to prevent the spread of the COVID-19 virus while fully simulating an environment of enjoying music alone, this workshop was conducted remotely with the aid of Zoom. This workshop was conducted in February 2021.

3.2.1 Basic Flow

As discussed in the previous sections, this study aims to design a new service that allows emotions to interact with music for people living a quarantined life. Based

Basic Flow for Online Workshop

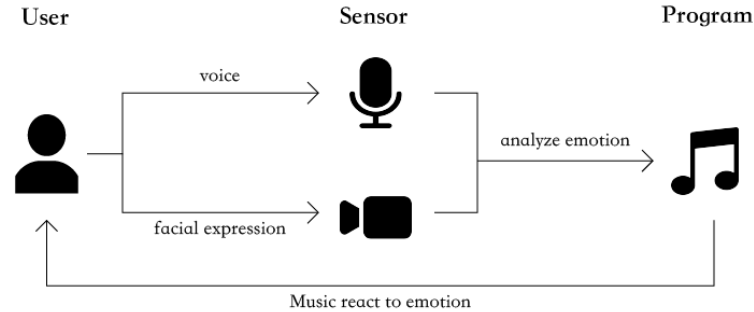


Figure 3.3 Proposed Workshop Flow

on a preliminary research of literature and users, it can be inferred that a design that could provide users with an unexpected interactive experience without any manual action was in line with their current needs. In this workshop, a rough flow of realising the interaction of emotional expression to change the music mix will be tested to the level of acceptance of the emotional response and to get some initial feedback for designing the first prototype.

The basic flow is as follows:

- 1) The user listens to the music with a digital device on his/her own with his/her usual habits.
- 2) Sensors such as camera and speaker capture the unconscious emotion signals expressed by the user and send it the inner program.
- 3) The program analyzes the emotion and changes the original music to react to user's emotion.

3.2.2 Participant

The participant this time is a 24-year-old master student from China called Galyn. She is living alone in a rent apartment with her cat and taking online classes from her Master's school in Australia. She is knows nothing about music composing

Basic Flow for Online Workshop

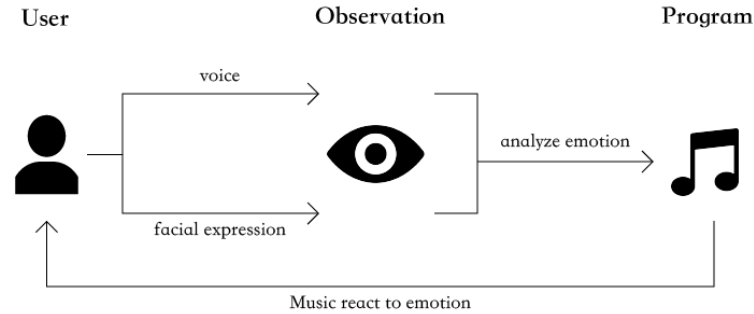


Figure 3.4 Practiced Workshop Flow

but had many experience about music, such as learning the piano, going to music concerts and etc.

3.2.3 Emotion Indicator

Distinct emotions such as happiness, sadness and anger are the typical emotions that are most often expressed, perceived and induced in listeners (Juslin and Laukka 2004). Since the aim of this design is to make users get happier in the music listening experience, this design will focus on reinforcing the positive emotions of the user. The detection of emotion is realized by observation instead of using sensors in this time's experiment.

Considering the possible mis-recognition caused by the researcher's observation of emotions, an emotion detection open API called Empath¹ was used to double check the emotion showed from the participant during the whole process.

3.2.4 Music Program

The emotion indicator in music change referred to the result of Masatomo's research which concludes music features and its effect on listeners' emotion as the following picture (Naito 2006):

Two music packages are prepared for this part. The first one is a piece of original music with many separate remix tracks to change the tempo and complexity of

Diagram of Tempo, Melody and Emotion

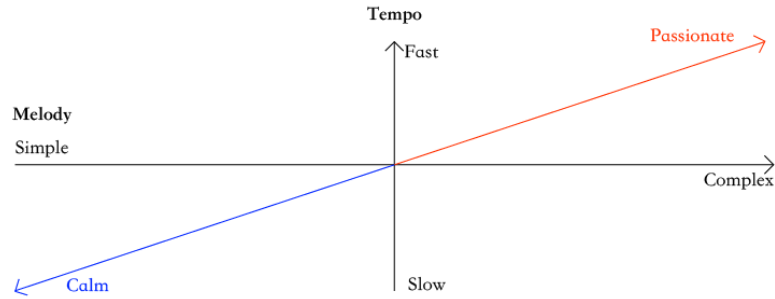


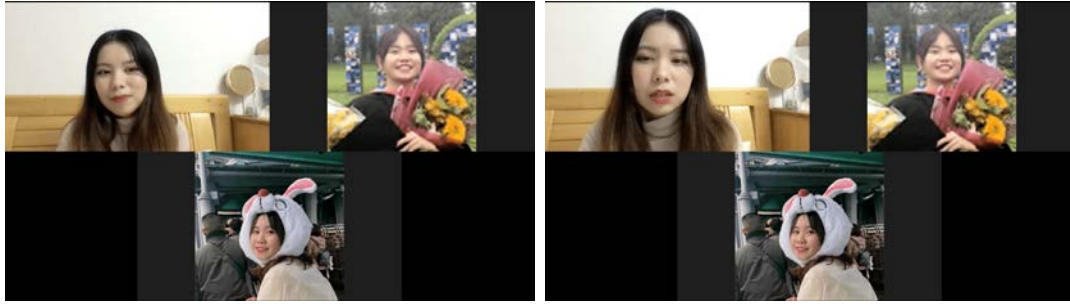
Figure 3.5 Relationship between Listening to Music and Mood Change



Figure 3.6 Workshop Setting

the music. The second one is her favourite piece of commercial music with several sound effects collected from the music concert on which that song was performed.

As the observer observes the participant's mood gradually becoming happier with the music (e.g. a gradual smile on the expression or shaking shoulders unconsciously, etc.), the observer matches this change to gradually increase the richness of the music, such as adding remix tracks or sound effects. Technically, adding remix tracts is realized by Adobe Premiere and adding sound effects is realized by both Adobe Premiere and Edjing. The implementation setting of this time is concluded as Figure 3.6.



(a) Part 1

(b) Part 2

Figure 3.7 Screenshot of Participant's Reaction

3.2.5 Feedback

The workshop ran for two rounds. Round 1 used the music track package of the original music. Round 2 used the music track package of the participant favorite song. The methodology for getting feedback is a semi-structured interview. The participant was firstly asked to fill out an evaluation form (Table 3.1). Interview questions were asked based on the scores she give for each item (Appendix C).

		Table 3.1 Evaluation Form		
	Engagement	Excitement	Responsiveness	Satisfaction
Round 1	8	8	5	4
Round 2	3	4	5	3

3.2.6 Discussion & Insights

Participant's feedback was discussed from the following 3 perspectives in the interview:

- 1. Difficulties in capturing voice data.

Even though before the workshop the participant reported that she has the habit of whispering along the music while listening, it was hardly to capture any voice data during the whole process. This is because the general setting

made her unable to act naturally as her usually does, since what she is facing is a Zoom interface which does not let her feel alone.

- 2. Difference in the feedback of the two rounds.

For the first round, it entertained the participant successfully because the music is in sync with her rising emotions. As for the second round, she thinks the additional sound ruined her favourite music. Two reasons can explain this difference. The first one is that she did not have any expectation on the original music piece while for her favourite song she had strong expectation of what should happen at which point of time. The second reason is that the original song was a whole with all the remix tracks together while the second one is adding additional factors to the whole. In other word, changing the integrity of the music arbitrarily can spoil the listener's experience.

- 3. Failure in fulfilling user's imagination of music experience.

With the bright room any without any visual hint, the participant reported that it was difficult for her to have an augmented experience which makes her feel like under the stage. "That was even not as exciting as watching to music videos in my bedclothes", she said. The lack of ambience environmental setting makes the ambient sound effects seem more rigid. Previous researches and reports also showed that lighting at a live concert helps a lot in conveying emotions and creating the proper atmosphere (Bonde et al. 2018). In the update of functions, proper environmental feedback will be paid more attention to.

Insights gained from this workshop can be concluded as follows:

- More comfortable settings
- No interruption from the observer.
- More environmental interactions for creating the atmosphere.
- An interface specially designed for this experience.

The insights gained from this time's workshop will be practiced in the next prototype.

Notes

1 <https://www.webempath.com/>

Chapter 4

Design and Implementation

In this chapter, the design concept will be introduced first to clarify the value of this design. Then the general design process is implemented following the principles raised up in the design concept. Finally, design outcome will be described to from four aspects, which are outlook, functions, user experience flow and mechanism.

4.1. Design Concept

This service aims at synchronizing emotion change, music and atmosphere which can deliver a more augmented music experience where users can enjoy music in a more responsive way. By using this service, users are supposed to have a more enjoyable music listening experience when staying home alone.

To achieve this goal, based on the discussion above, three principles are supposed to be raised up before doing the design process. Which are described as follows:

- This design would serve an entertaining service which people would like to utilize in their daily life, rather than a serious mental health treatment toolkit. Thus, improving mental well-being status is not the ultimate goal of this research, but providing an enjoyable experience which can be widely accepted.
- This design aims at enhancing the music experience through the interaction between unconscious emotional expression and music. Therefore, how to capture the unconscious emotional expression in a way that does not interrupt the user's unconsciousness is one of the focus in the design process.

- Instead of heavy installations in large scale, this design is supposed to be handy and difficult to be noticed. Therefore, the users will not be bother with being interrupted visually while listening to the music in their private space. Moreover, an program which runs at the backstage will be less easy to be paid attention to which provides a more close-life user experience.

4.2. Prototyping

Although the feedback given by the users in the first try was not satisfying, still a lot of valuable experience were gained, such as further adjusting the user flow design to make the overall process more in line with the user's behavioural habits. At the same time, some of the user needs and inspirations from the preliminary literature review and interviews were not realised in the first attempt, for example facilitating user's wish of starting a Karaoke wherever he or she is.

Therefore, in the making of the prototype, a design that contains a new user experience flow and several updates on functions called VIBO is proposed. The name VIBO comes from the word "vibe". Vibe as a verb has two meanings, enjoying oneself by listening to popular music and giving out an atmosphere (Dictionary 1989). This word is also quite frequently used in music industry in recent years. In this section, the overall user flow will be introduced first and then process of building the prototype will be introduced from several perspectives, including the main proposed functions, the technology applied and the interactive interface.

4.2.1 User Flow

The user flow is updated to work with more functions. Based on the feedback from pre-study, users would like a more environmental feedback to fulfill the imagination in music listening experience. Therefore, functions such as ambient brightness adjustment and ambient light effects are added to the general user flow. However, the general logic of this flow is not changed too much. The closed loop in which people's emotion interact with the program and then influenced by the program is called the affective loop (Sokolova and Fernández-Caballero 2015). Within this loop, the user first show their unconscious emotion expressions to the sensor, after the detection of emotion signals, audio and environmental responses will start to

User Flow of the Prototype

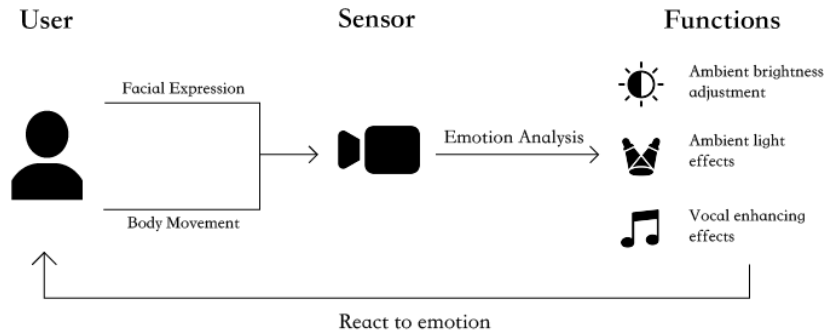


Figure 4.1 Proposed User Flow

react to the user. Therefore, the user will get to know about their emotion status from more perspectives as well as being influenced by the environment. The flow can be concluded as Figure 3.8.

Kinect Sensor is used as the camera sensor in this flow. There are two reasons for Kinect Sensor here. The first one is that unlike fixed camera on the computer, Kinect sensor can be place in many positions to fit user's listening habit. For example, if the user likes drawing while listening to music, a computer camera is hard to capture the face since it's towards the downside. But a Kinect sensor can achieve this while connected to the computer program. The second reason is that VIBO also involves body-movement detection and the depth camera of Kinect is able to facilitate this detection well.

4.2.2 Functions

Three main functions are proposed in the design of VIBO, which are ambient brightness adjustment, ambient light effects and vocal enhancing effect.

- **Ambient light effects**

Lighting effects will be projected to the environment to give positive signal of positive emotion, so that to achieve a more engaged and augmented music listening experience. Colors in principal hues are adapted when prepare

for the lighting since they convey positive emotions most (Kaya and Epps 2004). When situated in a space, people tend to attune their emotion to the lighting in the atmosphere (Edensor 2015).

- **Ambient brightness adjustment**

This function is proposed to further simulate the environment of a live concert. The environmental lighting will go darker with the rise of the user's "High" emotion (defined later). Theoretical support can be discussed from two perspectives. Firstly, lighting works darkness when inducing people's emotion in a space. Therefore, this function can work with the first one to provide more immersion (Edensor 2015). Secondly, in the music listening experience, when other redundant visuals are darkened, the focus of sensation will shift and will focus more on the music (Edensor 2015).

- **Vocal enhancing effect** This function is proposed based on users' wish on singing along while performing as the subject while listening to the music. Therefore, preferred characteristics such as echoes and vocal reduction were applied here. The sensor when detect the mouth shape changes of the user's to detect when to let this function work.

4.2.3 Emotion Classification

In the previous workshop, observation and an open API were used to tell the participant's emotion statement. However, observations are too subjective while the API checking is not something real-time. In order to achieve the real-time objective emotion identification, this prototype applied a real-time vision program which use convolutional neural network (CNN) to realise the emotion classification (Arriaga et al. 2017).

In this program, emotion is classified into 7 categories, which are angry, disgust, fear, happy, sad, surprise and neutral. Everytime the program detects the user's face, it will print out the values of the 7 types emotions contained in this facial expression. Examples are like:

- angry: 0.14160264 disgust: 0.03391444 fear: 0.18253383 happy: 0.064284734
sad: 0.1999726 surprise: 0.091957904 neutral: 0.31961435

This program fits the requirement of this prototype because it includes distinct emotions that music researches always discuss and it allows real-time feedback. Working in pair with the camera on the computer or cellphone, this program can realise the detection of unconscious emotional expressions in a way that does not bother the user.

Another part to be involved in the emotion classification part is the body movement. In the pre-studies, participants mentioned that they sometimes moves their body to the rhythm while listening to music. Therefore, real-time emotion detection in body movement is also used as a method to do mutual corroboration with facial expression. The emotion detection in body movement is achieved by machine learning. Data in UCF 101 data set were trained to realise the function¹.

4.2.4 User Interface

The aim of designing this interface is to let users have a way to know about what their current emotion is. There are two principles for designing this interface. The first one is make it looks more like a musical service interface instead of emphasizing too much on the mood, so that the interface will not bother users' music listening habits. The second one is to make it easy to use and easy to understand. In this way, users can spend less time on getting used to the function and pay more attention to the music they are listening.

As the outcome these thought, a slider with one side defined to be "Chill" and the other side defined to be "High" while in the middle of the slider is the state of "Neutral". The user interface of Spotify was referred to in order to let the design looks more close to user's common impression of a music service. There's no need for users to operate anything on this interface. It works only as a reflection of the user's current mood state. With the mood changing from Chill to High the tone of the background color will also change from cold color to warm color.

Despite this research would like propose a service which could let music listeners feel happier in a music listening experience, it is not to say that unhappy mood condition is something bad. Therefore, in order to avoid verbal discrimination of negative emotions, in the interface design process, negative mood are concluded as "Chill" and happy emotion is concluded as "High" and presented as "Yeah " to make it looks more relaxing.



Figure 4.2 Proposed User Interface

Notes

- 1 crcv.ucf.edu/data/UCF101.php

Chapter 5

Evaluation

5.1. User Test

In order to see whether with the facilitation of augmented music tools, user's music experience can be enhanced, a series of user test was done with the facilitation of a mock-up set-up.

5.1.1 Participants

The user test was done with 3 participants who define themselves as music lovers. They share three features in common. The first one is that they all have the habits of enjoying music while doing other tasks in their daily life. The second one is that they all have the desire to react to the music while listening, for example shaking legs or singing along with the song. The third one is that they basically listening to the music in order to get rid of bored or enjoy a leisure time. The detailed profile of the three participants are as follows:

- Mizu: Female/20 years old/Art University Student
Her habit of listening to music is to listen while tidying up stuffs.
- Chou: Male/25 years old/KMD Alumni
His habit of listening to music is to listen while doing graphic design tasks.
- Chen: Female/23 years old/KMD Student
Her habit of listening to music to listen while doing make-ups or tidying up stuffs.

5.2. Music list

Three music lists were created on Spotify according to the music taste of the three participants. The researcher asked the three participants to provide 5 titles of the songs that they usually listen to during the stay-at-home period. The music list created is as following:

- Music list of Mizu: 1)Luo hua liu shui (Eason Chan)
2)Zui jia sun you (Eason Chan)
3)hope is a dangerous thing for a woman like me to have (Lana Del Rey)
4)Because of you (Mayday)
5)Sora ni uta e ba(amazarashi)
- Music list of Chou: 1)Pretty boy (Joji)
2)Test drive (Joji)
3)ONLY (Niki)
4)Real Love (Beatles)
5)Midsummer Madness (88RISING)
- Music list of Chen: 1)Lover Boy 88 (Phum Viphurit&Higher Brothers)
2)SLOW DANCING IN THE DARK (Joji)
3)A RA SHI(Arashi)
4)Aozora(aiko)
5)Sakura no toki(aiko)

The songs are gathered on the testing day and was gathered by asking the question of: Please list the five songs you would like to listen to with your mood now.

5.2.1 Set-up

The user test was launched in one of the rooms in my apartments. Three participants were invited to come and conduct the user test one by one. The whole set-up includes four devices and a software.

Hardware includes:

- Laptop:
Running the emotion classification program, playing the music and recording the user's voice.
- Kinect Sensor:
Located on the desktop to capture the participants' facial expression and body movement. Work together with program on the laptop.
- Spotlight projector:
For projecting ambient lighting effects.
- Video projector:
For projecting music videos sometimes.

Software includes:

- PyCharm:
For running the emotion classification program.
- Spotify:
For playing music and creating music lists.
- Echo:
For making the echo and vocal elimination effect.

5.2.2 Other Preparations

Instead of the set-up, music playlist and the environmental setting were also prepared in order to bring the users a more close-to-life music experience.

For the playlist, each participant was asked to provide 5 6 songs before the test to create a personal playlist in advance. The songs are required to be the ones that like to listen to during the pandemic period. And the songs are preferred to be the ones that they are listening to while doing other stuffs such as washing clothes etc recently. A personal playlist was created on Spotify for each participant so that they can enjoy the whole experience seamlessly.

For environment settings, the environment setting varies from each participant. For Mizu, a box of messed up ear-rings were provided for her to re-arrange during

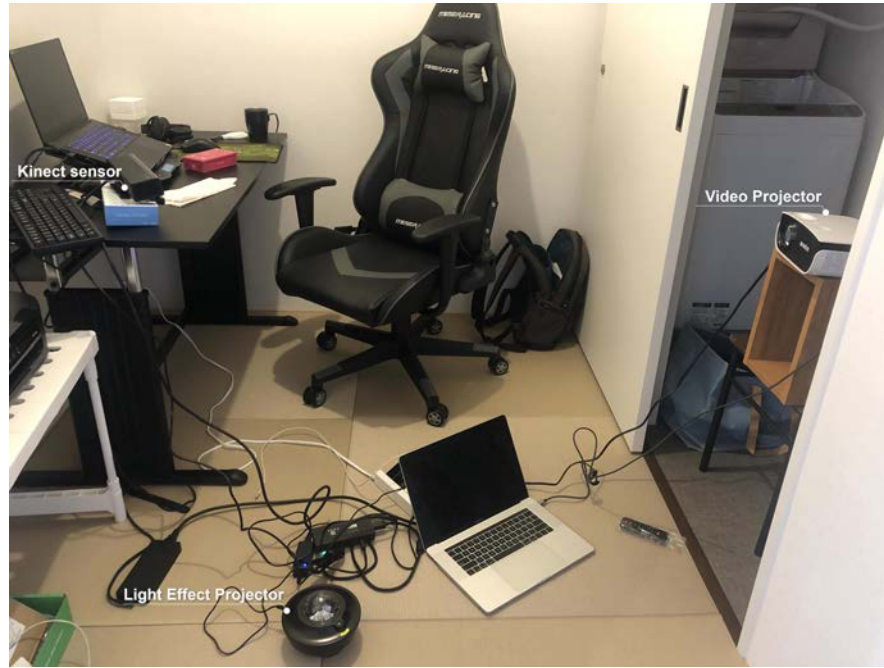


Figure 5.1 User test Setting

the test. For Chou, he was asked to design a random poster while listening playlist. For Chen, a set of make-up toolbox was prepared and let she enjoy it as if she is going to attend an online meeting.

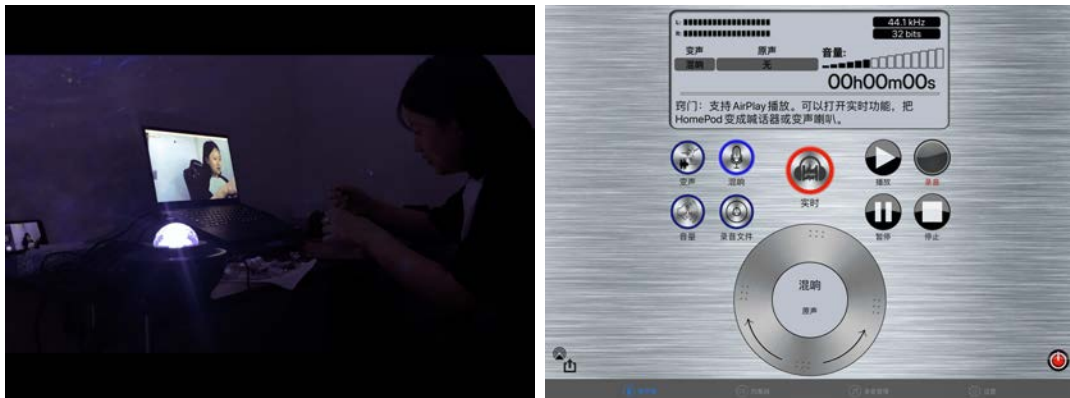
5.3. Flow

Three participants are invited to the set-up room on three days to take the test. The basic structure of user test is described in the following figure Figure 5.2. During the test B, the functions are tested in the sequence of 1) Ambient brightness adjustment, 2) Ambient light effects 3) Vocal enhancing effects. Therefore, in the following parts, Function 1, 2, 3 is used for this sequence as well. The three functions were tested one by one. When one function is reflecting to the user's emotion change, two other functions are not supposed to be working at the same time, but staying at the certain state. For example, when the ambient light effects are added to the environment the ambient brightness is stable and no vocal enhancing effects will be triggered.



(a) Ambient light effects 1

(b) Ambient light effects 2



(c) Ambient brightness adjustment

(d) Vocal enhancing effect

Figure 5.2 Applied functions

User-test Structure

	Test A (Control group)	Test B (Experimental group)
Features	In the morning Listening without any facilitation	In the afternoon Listening with all functions gradually added on
Commonalities	Same playlist, same play sequence, same room doing same things while listening	

Figure 5.3 User-test Structure

5.3.1 Function by Function Introduction

Function 1: Ambient brightness adjustment

The first function was to darken the lighting of the room when the user’s emotion is detected to be changing with the music. This was realised by using the remote controller of the lamp of the room. Based on the data gathered from test A in the morning, as the users get more into the condition of enjoying the music, their positive emotions tends to get grow slightly. This test would like to testify whether with the facilitation of Function 1, the users can focus more on the music listening experience and gain more positive emotion.

Function 2: Ambient light effects

The second is ambient light effects. In this function, lighting effects will be projected to the environment to give positive signals of positive emotion in order to achieve a more engaged listening experience. The signals include lights in principle hues and twinkling with the rhythm. According to the literature discussed above, principle hues lighting provide positive emotion signals which can respond to the users’ emotion change (Kaya and Epps 2004). Also, the twinkling effects are added to the beats of the music at the climax part. This is to visualize the increase of complexity of music by adding the complexity of the environmental elements, so

that the users can use more sensations to experience the music and trigger their unconscious emotions. Previous research from KMD also proves that ambient light change according to the user's emotion can provide a better recognition of the user's mood condition and facilitate in a positive way (Yamamoto 2020).

Function 3: Vocal enhancing effects

The third function is vocal enhancing effect. The function is proposed based on users' wish of singing along while performing. When the user starts to sing, the original vocal in the song will be eliminated and echo effects will be triggered. A Chinese software called Mei Yin Bian Sheng Qi was used here to realise the function. The software was set to imitate the sound effect of a Karaoke box so when users sing something aloud, an echo will be triggered. Adobe Audition was used for eliminating the original vocal in the music track, .

In the following parts, the basic data analysis results will be described first. Then the result from the semi-structured interview will be introduced. Finally, a further comprehensive analysis will be conducted based on the results from both results.

5.4. Data Analysis

In each session, the changes in sensor values are recorded in detail. The program of expression recognition is set to print emotion parameters every 0.2 seconds whenever a face is recognized. Data analysis is conducted in two ways. The first one is to compare the average values of "High" emotions throughout the experiment in both the control group and the experimental group. The second one is to see the changing trend of participants' emotion when facilitated by the three functions of VIBO. The values of "Surprise" and "Happy" are the focus of this data analysis.

Analysis 1

Analysis 1 aims at comparing the differences in test A and test B. In this analysis, all data in Figure 5.4 presents the comparison between test A and test B. The

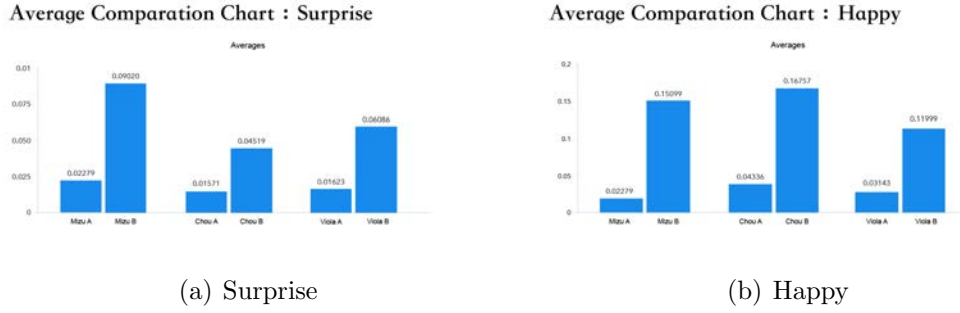


Figure 5.4 Comparison chart of "Surprise" & "Happy" values

average value of both "Surprise" and "Happy" increased greatly in test B compared with test A, which proves that the three functions of VIBO can facilitate the emotion increase of "Happy" and "Surprise" compared with just listening to recorded music. Moreover, the average value of "Happy" increased greater than that of "Surprise". Whether this means that VIBO have better effects on "Happy" emotions will be discussed further in Analysis 2 part.

Analysis 2

This part aiming at doing an in-depth analysis of the effect of each functions in VIBO by using data. The data is recorded function by function, but the general time length for each function is controlled to be basically in similar length. The curves in Figure 5.5 is generated based on the trend of the value change in the scatter chart.

In the six graphs about the emotion changes of "Surprise" and "Happy", we can notice a very large individual difference between each participant as well as each function. The analysis will be conducted function by function while considering the individual differences.

In the two graphs about Function 1 (function of Ambient brightness adjustment), all three participants generally showed a slow upward trend in the parameters of both "Surprise" and "Happy". Even though in the graph of Mizu, the value of "Surprise" showed a trend of decrease at the end of the test session, the overall upward trend in the first 72 seconds is already proof enough that this feature has brought some surprises to users during use.

In the two graphs about Function 2 (function of Ambient light effect), all three participants generally showed a slow upward trend in the parameters of both "Surprise" and "Happy" as well. Similarly as Function 1, the same participant showed a trend of decrease in "Surprise" at the end of the test session, but the general trend is relatively smooth. Moreover, a quick increase in the value of "Happiness" of Chou was observed. To rule out that this rapid increase was due to other factors from the experiment settings, the researcher reviewed the video record of the experiment and confirmed that the change was due to VIBO's function.

In the two graphs about Function 3 (function of Vocal enhancing effect), the trend was difficult to conclude. Function 3 is that every time when the participant wanted to whisper the sing on his or her own, the original vocal in the song will be diminished a little and the whisper will be enhanced with the facilitation of the software. By reviewing the experiment video while referring to the data, it is found that every time when the participant activates the function, his or her value of "Happy" and "Surprise" will increase suddenly and then gradually get back to calm again. Moreover, two participants' data showed a generally growing trend of "Happy". This proves that Function 3 worked for these two participants. However, for one participant, both the value "Surprise" and "Happy" decreased a little bit during the process, which means that Function 3 showed a negative effect on his positive emotion. The differences in effect on different participant for Function 3 will be further discussed later.

5.5. Semi-structured Interview

A semi-structured interview was launch at end of each user test day. The interview basically has two parts. The first one is to let the users fill out an evaluation form to rate each function. Then, questions are asked based on the scores given by the participants. The result given by the three participants are listed as the following three tables (See Table 5.1, Table 5.2, Table 5.3). There are two principles for this form. The first one is that higher score means higher evaluation. The second one is that 5 is the standard score for comparing with the normal music listening experience. If the user has better experience than normal times, scores over 5 are



Figure 5.5 Function by Function test: Value of "Happy" & "Surprise"

supposed to be rated and vise versa.

Table 5.1 User 1: Mizu

	Engagement	Excitement	Responsiveness	Satisfaction
Function 1	8	8	6	7
Function 2	7	7	7	8
Function 3	8	8	8	8

Table 5.2 User 2: Viola

	Engagement	Excitement	Responsiveness	Satisfaction
Function 1	8	7	5	7
Function 2	7	7	6	8
Function 3	7	8	7	8

Table 5.3 User 3: Chou

	Engagement	Excitement	Responsiveness	Satisfaction
Function 1	6	8	5	6
Function 2	6	7	6	7
Function 3	2	3	5	2

Although the participants filled out the questionnaire without knowing the results of the Kinect data, the results of the questionnaire generally matches the results of Kinect data. All three participants rated Function 1 and Function 2 with a score over 5. However, for Function 3, the participant whose data didn't show an increase in "Happy" and "Surprise" rated it with quite low scores. Based on these results, several questions are asked.

Q1: Why you rated these three functions with these scores in the column of "Engagement" and why is it more/less engaging compared with the normal music listening experience?

- Mizu: I was surprised at first and gradually focused more on the music and my emotion data on the interface slider. The pattern of "Me respond to

music to change to environment” really made me felt more concentrated on the music listening experience, so I rated ”Engagement” with a quite high score.

- Viola: I didn’t expect this would happen but this gave me a similar feeling to that of the moment when a concert is going to start and the star is going to show up and feel that the star is me. This interface made me felt I’m the center of this experience.
- Chou: Really good to have the videos and light effects around me. I feel like accompanied by all these elements around.

Q2: Why you rated these three functions with these scores in the column of ”Excitement” and why is it more/less exciting compared with the normal music listening experience?

- Mizu: My favourite singer’s video showed up around me when both my concentration on music and the music itself is coming to a climax. That was exciting.
- Viola: Even though the sound effect is not really good but it was exciting to have voice echo when I want to whisper along the song. Also with the ambient light setting, I felt I was in the spotlight.
- Chou: The surrounding light effect was interesting. They surprised me a lot at the beginning, and made me felt I was in a Karaoke box, especially within this tiny room. Since I haven’t been to a Karaoke for long, the mood was exciting.

Q3: Why you rated these three functions with these scores in the column of ”Responsiveness” and why is it more/less responsive compared with the normal music listening experience?

- Mizu: I think I was basically calm, so it’s good to see every effect was calm and slightly changing as well. I sometimes need to shift my focus on to the tasks I was doing, the calm pattern was good facilitation to calm me down when I’m in trouble with the task.

- Viola: It was interesting when I laughed out because of my lipstick, the lighting effect shake as well. With the aid of music around, I feel like in the center of the room.
- Chou: I usually like to work in the darker places to get more immersed in my work and the music. It's good to see the brightness adjustment can fit my emotion change.

Q4: Why you rated these three functions with these scores in the column of "Satisfaction" and why is it more/less satisfying compared with the normal music listening experience?

- Mizu: I enjoyed the music listening experience more than usual I think. I felt the time passed so quick. I was hearing, seeing, doing tasks and experiencing. Not boring at all.
- Viola: Very entertaining as an music experience. Sometimes feel that the darkness fits my emotional needs but doesn't fit the needs of what I'm doing. For example, when I'm drawing my eyeliner, I need the light to be a little bit brighter.
- Chou: I don't really have the habit of whispering while listening to the music. So it was weird when the sound effect is triggered but I was not about to sing along.

Q5: Are there anything disrupting in the general user-test setting and user-test flow?

- Mizu: It was smooth and comfortable. I feel I can listen to the music whole day sitting here. I hope it can have more scenarios to use for example sweeping the floor or washing dishes.
- Viola: Interesting to interact with smiles. The Kinect sensor is a little bit too bad for the whole setting. It would be better it a smaller camera is placed. Also, it seems to be able to facilitate online lives as well, for example, audience emotion can change the room setting or the stage setting.

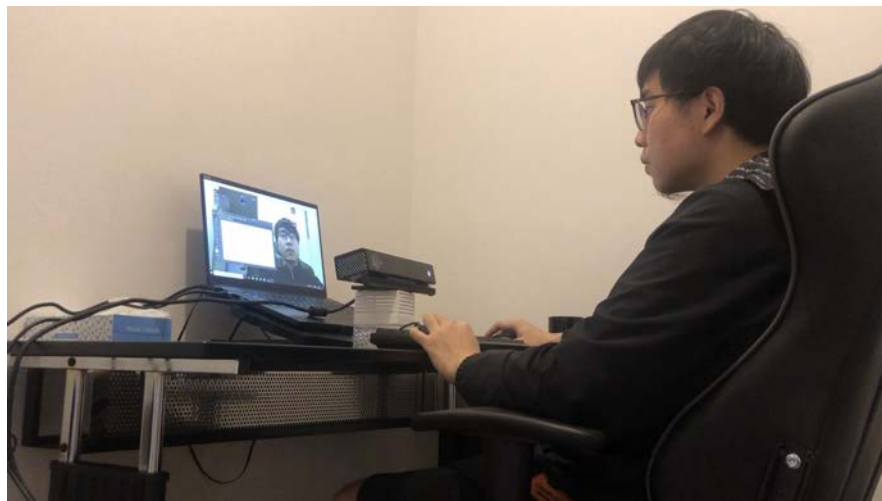


Figure 5.6 Sample of interface change: At the beginning

- Chou: The setting allows me to enjoy music alone and doesn't interrupt my work. But I want more possibilities on operating myself.

5.6. Conclusion

According to the user test presented above, the conclusion of the user test can be summarized from three perspectives: 1)Effectiveness 2) Points to be improved 3)Potential usage.

Effectiveness

Effectiveness of Function 1 and Function 2 was proved through data analysis and the interview feedback. However, the result shows that the effectiveness of Function 3 depends on people and their habits. But since it requires the detection of mouth movement to activate, if we regard Function 3 as a function that people without that habit will never activate, it is a function that won't supposed to be triggered by people without such habit. In other words, this function won't bother the overall experience. Considering this point, VIBO is basically effective for its users.

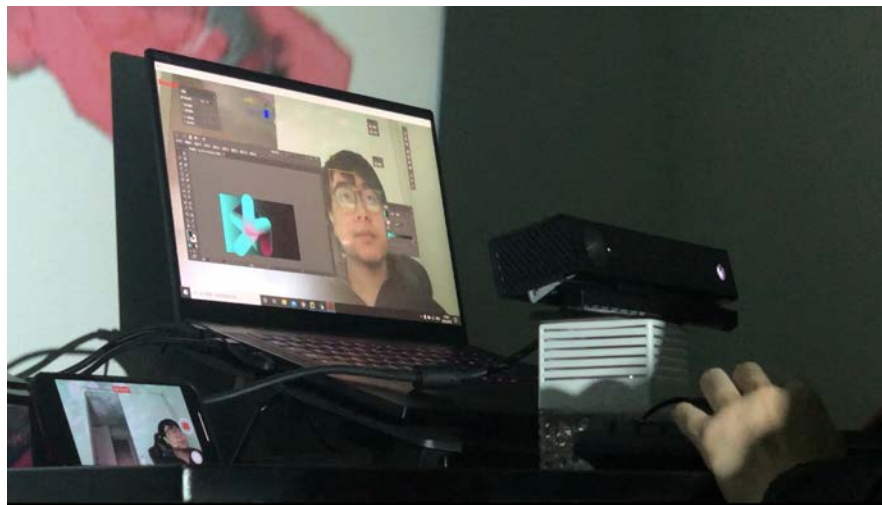


Figure 5.7 Sample of interface change: Approaching the end

Points to be improved

According to the answers from the participants, two points to be improved can be concluded. The first one is lack of interactions that users can manipulate consciously, for example turning the function off in the mid-way or adjusting light brightness according the current needs. The second one is scale of the installation. One of the participant mentioned that the Kinect Sensor bothered her experience to some extent in the interview. Also, when doing the set-up of the installation, the researcher also noticed the weakness of portability of this service. The detailed solutions for these two points will be discussed in more detail future work in Chapter 6.

Potential usage

The surprising finding from the interview is that from the participants' answers two potential usages for VIBO can be indicated. The first one is to develop it into a service or product designed especially for online live concerts. Mizu mentioned in the interview that she had participated in online concerts twice after the break out of COVID-19. She turns off the lights to create the concert mood but still feel lack of connection with the concert. It will be nice if she can use VIBO to create more ambient effects around according to her mood. Also, she mentioned

that when there's a point that requires the audience to sing along with the singer, it will also be nice to use Function 3 at home to let the user feel the echo of her voice which is more engaging than whispering in a low voice. The second one is to apply the functions of VIBO to nowadays' Karaoke boxes to make singing alone a more exciting thing to experience. For example, the ambient lighting effect and the brightness changes as the user's singing is getting into a better condition.

Chapter 6

Discussion

6.1. Limitation

1. **Limited accuracy of emotion recognition.**

In the original paper of the emotion recognition system, two limitations are listed (Arriaga et al. 2017). The first one is that it is biased towards western faces. The second one is that its accuracy goes down when people wear glasses. One of the participants in this research wears glasses, the result may have been affected by this factor.

2. **Insufficient sample size.**

Due to the influence of COVID-19, this research was also supposed to be conducted following the new social manner to the largest extent. Therefore, three participants who has the least risk of being infected were invited to participate (Two had the vaccine injected in China and one lives nearby). However, since three is an odd number, all participants had Test A and Test B in the same sequence. It would be better if even number participants could join, so that the test sequence can be switched for each pair.

3. **Limitations of use scenarios in testing.**

In the proposed user scenario, the service can be enjoyed while doing many activities, such as cooking, cleaning the floor, etc. However, in the user test setting, since Kinect is connected to the laptop, users can only listen to the music while doing things on the desk-top. Luckily, all three participants have the habit of listening to the music on the desk. If a mobile application version of VIBO could be developed, this limitation can be solved to a great extent.

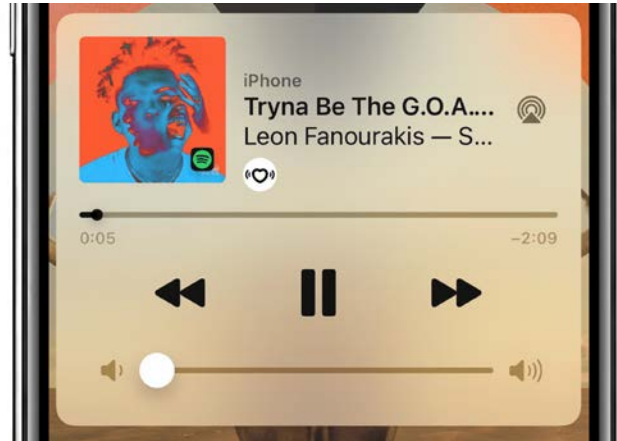


Figure 6.1 Switch for turn it off

6.2. Future Work

Three directions of future work are proposed based on the three points that this version of VIBO needs improve.

1. **Future work 1** The first one is to add an extra button for turning this function on or off on the interface. Figure 6.1 presents a mock-up image of this button on Spotify. Therefore, the user can enjoy or stop the function whenever they want. For example when the user suddenly have to get back to another business, he or she can change the situation very quickly.
2. **Future work 2** The second one is to develop the service into a portable product to let users enjoy it in their comfortable way. Therefore, the whole setting won't be too interrupting and the user scenario can be expanded as well. Since the service works together with software and camera, a peripheral that can be plugged into a cellphone or a computer is proposed. The concept image for this is as Figure 6.2 and Figure 6.3.
3. **Future work 3** The third part of the future work is to plan a more wide-ranged testing with the integration. In the next stage of user test, users of a wide age range and a nationality with music tastes of a wider range will be considered to invited. Therefore, online methods such as sharing platform



Figure 6.2 Integrated Product



Figure 6.3 Usage simulation

and focus groups such as music communities are also possible methods to be considered for the next stage of research.

References

- Arriaga, Octavio, Matias Valdenegro-Toro, and Paul Plöger (2017) “Real-time convolutional neural networks for emotion and gender classification,” *arXiv preprint arXiv:1710.07557*.
- Bonde, Esben Oxholm, Ellen Kathrine Hansen, and Georgios Triantafyllidis (2018) “Auditory and Visual based Intelligent Lighting Design for Music Concerts,” *Eai Endorsed Transactions on Creative Technologies*, Vol. 5, No. 15, p. e2.
- Cabedo-Mas, Alberto, Cristina Arriaga-Sanz, and Lidon Moliner-Miravet (2020) “Uses and perceptions of music in times of COVID-19: a Spanish population survey,” *Frontiers in psychology*, Vol. 11.
- Collier, Gary and Gary James Collier (2014) *Emotional expression*: Psychology Press.
- Csikszentmihalyi, Mihaly and Isabella Selega Csikszentmihalyi (2006) *A life worth living: Contributions to positive psychology*: Oxford University Press.
- Darwin, Charles (1965) *The Expression of the Emotions in Man and Animals (by) Charles Darwin*: University of Chicago Press.
- DeNora, Tia (2000) *Music in everyday life*, Vol. 665: Cambridge University Press Cambridge.
- Dictionary, Oxford English (1989) “Oxford english dictionary,” *Simpson, JA & Weiner, ESC*.
- Edensor, Tim (2015) “Light design and atmosphere,” *Visual Communication*, Vol. 14, No. 3, pp. 331–350.

- Ekman, Paul and Wallace V Friesen (1969) "Nonverbal leakage and clues to deception," *Psychiatry*, Vol. 32, No. 1, pp. 88–106.
- Elson, Miriam (2001) "Silence, its use and abuse: A view from self psychology," *Clinical Social Work Journal*, Vol. 29, No. 4, pp. 351–360.
- Graham, Gordon (2005) *Philosophy of the arts: An introduction to aesthetics*: Routledge.
- Hager, Joseph C, Paul Ekman, and Wallace V Friesen (2002) "Facial action coding system," *Salt Lake City, UT: A Human Face*, p. 8.
- Hall, Stuart (2003) "Encoding/decoding," *Television: critical concepts in media and cultural studies*, Vol. 4, No. 1, pp. 43–53.
- Hargreaves, David J, Dorothy Miell, RAR MacDonald, RAR MacDonald, D Hargreaves, and D Miell (2002) "Musical identities," *MUSIC PSYCHOLOGY*, p. 759.
- Izard, Carroll E (1977) "Differential emotions theory," in *Human emotions*: Springer, pp. 43–66.
- Juslin, Patrik N and Petri Laukka (2004) "Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening," *Journal of new music research*, Vol. 33, No. 3, pp. 217–238.
- Kamalzadeh, Mohsen, Dominikus Baur, and Torsten Möller (2012) "A survey on music listening and management behaviours."
- Kaya, Naz and Helen H Epps (2004) "Relationship between color and emotion: A study of college students," *College student journal*, Vol. 38, No. 3, pp. 396–405.
- Lamont, Alexandra and Rebecca Webb (2010) "Short-and long-term musical preferences: what makes a favourite piece of music?" *Psychology of Music*, Vol. 38, No. 2, pp. 222–241.
- Le Guin, Elisabeth (2005) *Boccherini's body: An essay in carnal musicology*: Univ of California Press.

- Lindquist, Kristen A and Lisa Feldman Barrett (2008) “Emotional complexity..”
- Littleton, Karen, Dorothy Miell, and Dorothy Faulkner (2004) *Learning to collaborate, collaborating to learn: Understanding and promoting educationally productive collaborative work*: Nova Science Publishers Inc.
- Malloch, Stephen and Colwyn Trevarthen (2018) “The human nature of music,” *Frontiers in psychology*, Vol. 9, p. 1680.
- Mazza, Cristina, Eleonora Ricci, Silvia Biondi, Marco Colasanti, Stefano Ferracuti, Christian Napoli, and Paolo Roma (2020) “A nationwide survey of psychological distress among Italian people during the COVID-19 pandemic: immediate psychological responses and associated factors,” *International journal of environmental research and public health*, Vol. 17, No. 9, p. 3165.
- Naito, M (2006) “The study of relationship between listening to music and mood change-Musical tempo and melody, preference music influence on a mood scale and relaxation state,” *Nihon University GSSC Journal*, Vol. 7, pp. 441–450.
- Öhman, Arne (1999) “Distinguishing unconscious from conscious emotional processes: Methodological considerations and theoretical implications,” *Handbook of cognition and emotion*, pp. 321–352.
- Oltețeanu, Ion (2010) *Understanding Music Theory: Meaning, Self-Consciousness, and Emotional Expressiveness*: Addleton Academic Publishers.
- Owren, Michael J, Drew Rendall, and Jo-Anne Bachorowski (2005) “Conscious and Unconscious Emotion in Nonlinguistic Vocal Communication..”
- Ribeiro, Fabiana Silva, João Paulo Araújo Lessa, Guilherme Delmolin, and Flávia H Santos (2021) “Music listening in times of COVID-19 outbreak: a Brazilian study,” *Frontiers in Psychology*, Vol. 12.
- Robinson, Jenefer (1994) “The expression and arousal of emotion in music,” *The Journal of Aesthetics and Art Criticism*, Vol. 52, No. 1, pp. 13–22.

- Robinson, Jenefer (2007) “Can music function as a metaphor of emotional life?” *Philosophers on music: Experience, meaning, and work*, pp. 149–177.
- Roe, Keith (1985) “Swedish youth and music: Listening patterns and motivations,” *Communication Research*, Vol. 12, No. 3, pp. 353–362.
- Sachs, Matthew, Sarah Hennessy, Jonas Kaplan, and Assal Habibi (2021) “Music and Mood Regulation 1 during the early-stages of the COVID-19 Pandemic.”
- Scruton, Roger (1999) *The aesthetics of music*: Oxford University Press.
- Seitz, Jay A (2005) “Dalcroze, the body, movement and musicality,” *Psychology of music*, Vol. 33, No. 4, pp. 419–435.
- Skånland, Marie Strand (2013) “Everyday music listening and affect regulation: The role of MP3 players,” *International journal of qualitative studies on health and well-being*, Vol. 8, No. 1, p. 20595.
- Sloboda, John A and Susan A O’neill (2001) “Emotions in everyday listening to music,” *Music and emotion: Theory and research*, Vol. 8, pp. 415–429.
- Sokolova, Marina V and Antonio Fernández-Caballero (2015) “A review on the role of color and light in affective computing,” *Applied Sciences*, Vol. 5, No. 3, pp. 275–293.
- Sundarasan, Sheela, Karuthan Chinna, Kamilah Kamaludin, Mohammad Nurrunnabi, Gul Mohammad Baloch, Heba Bakr Khoshaim, Syed Far Abid Hossain, and Areej Sukayt (2020) “Psychological impact of COVID-19 and lockdown among university students in Malaysia: Implications and policy recommendations,” *International journal of environmental research and public health*, Vol. 17, No. 17, p. 6206.
- Van Goethem, Annelies and John Sloboda (2011) “The functions of music for affect regulation,” *Musicae scientiae*, Vol. 15, No. 2, pp. 208–228.
- Yamamoto, Daisuke (2020) “Design for assisting method of emotional homeostasis with ambient medium in everyday space,” URL: <file:///Users/ada/Downloads/K040001001-00002020-0792.pdf>.

- Ziv, Naomi and Revital Hollander-Shabtai (2021) “Music and COVID-19: Changes in uses and emotional reaction to music under stay-at-home restrictions,” *Psychology of Music*, p. 03057356211003326.

Appendices

A. Pre-study Interview 1

Q1. Why are you are living a quarantined life now?

I flew from Japan to China to spend my spring vacation with my family. Since my hometown doesn't have direct flight from Japan, so I'm doing my quarantine in Shanghai.

Q2. How do you feel ?

I felt excited on the first day because the condition was not bad. But I began to get bored with everything here from the next day. This is my 5th day. I feel super bored. Sometimes I look out of the window just hope time could fly faster. My mental state is terrible, I'm thinking a lot about negative things and no passion to do anything. For example, I always don't know what is the next thing to do after finishing the breakfast. Also, even though I know I need to do something, such as washing socks and doing homework, I just don't move on to doing it. Seems that my bank of happiness is not enough. My daily routine is very regular and boring. When to have meals and when to take a temperature test were fixed. Every day was the same thing over and over again. Although I'm not strongly dissatisfied with life, I'm gradually losing enthusiasm and expectation of life, even losing interest in snacks and TV dramas I used to love, and spending most of the time watch Tom and Jerry series again and again to charge myself, or probably just to kill the time.

Q3. In what scenarios do you usually listen to music?

Music was an important part of my daily life, especially in these days. Since I'm alone all the time, sounds from the speaker makes me feel not so lonely. Music is

my companion when doing many activities, such as waking up, doing homework, washing clothes etc. Since there's nobody around, I use the speaker of smartphone to play it aloud. I will definitely listen to the song while I am in the shower. I love the reverb of listening to singing voices in the bathroom. I will definitely listen to the song while I am in the shower. I really like the reverb of listening to singing in the bathroom. At other times too, I often can't help but sing along as I listen to the music.

Q4. What music service do you usually use?

I use Netease Cloud Music most, because it has recorded my tastes and the music library I created over the past ten years. It knows me well, even though I'm not satisfies with some new functions, I'm still using it.

Q5. What's your expectation about future music services?

Sometimes I use functions such as Heartbeat Playlist (a function that the system automatically generates a recommended music list based on the user's historical preferences) and Karaoke on Netease Cloud Music but not often, because these functions sometimes require many extra operation that interrupts the general music listening experience. I'm expecting a service that could "let music play with me" from present music services. Sometimes when I listen to a song, I can remember the scene when I was listening to it under stage during the concert. Then I recall more about the concert memories and even want to re-watch the video of that concert. But when the music ends and I switch to a completely unrelated song in the single playlist, I get very frustrated or manually stop the new song.

B. Pre-study Interview 2

Q1. Why are you living a quarantine life?

I was studying abroad in a film academy in the US before the pandemic. Due to the pandemic, all the classes became online. As a film student, I really want to have a chance to work on spot. Another reason is that because of the pandemic,

I lost my part time job as an Airbnb host assistant. Therefore, I decided to come back to Japan to find more chances. Japan has the in-bound quarantine policy. So I was at home for 14 days. Right after the quarantine period, Japanese government announced the state of Emergency. Therefore, I'm still at home.

Q2. How do you feel about your quarantine life?

I was with my family, so I didn't feel particularly lonely, but it was still boring. I had a lot of time on my hands, so I did all the things I don't usually do, like organizing the attic, renewing the garden or cutting fish. But since I had too much time but few things to do, things get worse afterwards. I cannot help thinking myself negatively. For example, I denied my abilities and my future. When I have nothing to do, I always think I'm useless. I was quite worried about my career plan. The pandemic totally interrupted it. I was neither having enough input nor doing anything practical in Japan. And I just have too much time to overthink about everything. Therefore, I'm not feeling well.

Q3. What methods did you try to overcome this situation.

The first thing I did to overcome this mental situation was to study various things, such as Korean language and video editing. Then, to relax myself, I also took a bath and listened to music when feeling bad.

Q4. Are there any specific behavioral patterns when you listen to the music?

I don't particularly sit still and listen to music. For example, when I need to use my brain, like studying English or doing homework, I listen to classical music, and when I'm not using my brain for work, like video editing, I listen to R&B or rap. Also, when I take a bath, I always play a song on my phone that I can sing, because the echo in the bath is amazing and it makes it easier to sing. It makes the bathroom like a private Karaoke box.

Q5. What music service do you usually use?

I used to use Spotify when I was in the US, but now I use Amazon Music most. I didn't subscribe to either of the service. So when I was using Spotify, random promotion music was really hard to bear. In Amazon Music I use the Shuffle function a lot, but it only recommend new songs of my favorite genre, so it's quite enjoyable.

Q6. In what situation do you enjoy music most?

I don't have a clear criteria for felling love with a new song. But because I came back from studying in the U.S. and immediately went into isolation, I can't help but devote more emotion and attention on listening whenever I hear music that reminds me of the time before. For example, when I happen to hear a song I heard in a nightclub in America, my mind will be filled with memories of that time, and I will wish to be in that scene.