Title	Emophoto : augment emotion information in photo to improve the sharing and reminiscing experience
Sub Title	
Author	沈, 玮玮(Shen, Weiwei) 加藤, 朗(Katō, Akira)
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2019
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2019年度メディアデザイン学 第716号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002019-0716

Master's Thesis
Academic Year 2019

EmoPhoto: Augment Emotion Information in Photo to Improve the Sharing and Reminiscing Experience

Keio University
Graduate School of Media Design

Weiwei SHEN

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

Weiwei SHEN

Master's Thesis Advisory Committee:

Professor Akira Kato	(Main Research Supervisor)
Senior Assistant Professor Marcos Sadao Maekawa	(Sub Research Supervisor)

Master's Thesis Review Committee:

Professor Akira Kato	(Chair)
Senior Assistant Professor Marcos Sadao Maekawa	(Co-Reviewer)
Professor Nanako Ishido	(Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2019

EmoPhoto: Augment Emotion Information in Photo to Improve the Sharing and Reminiscing Experience

Category: Design

Summary

Photos are the most common carrier of human memory and emotions. Currently photo sharing through social networking service is more and more popular among people with the growth of smartphone service. However, there are also problems that people have to browse so many image information every day, and many people are amateur so their photos are not well designed to express their feeling.

In this paper, EmoPhoto system is designed to embed emotion information in metadata of photos and translate it into visual and audio effect to arouse empathy of viewers. The system was developed based on WeChat Mini-Program and the proof of concept was conducted to verify the efficiency of the corresponding effect.

Due to the feedback from participants, visual effect of negative emotions performed very good to express the corresponding emotion while the effect of positive emotions failed. Audio information was greatly appreciated in the test, surprisingly. In the whole, according to the result of the test, the effect of EmoPhoto to promote further communication among people could be expected.

Keywords:

Photo Sharing, Emotion Embedding, Social Networking Service, Empathy, Communication

Keio University Graduate School of Media Design

Weiwei SHEN

Contents

List of Figures

List of Tables

Acknowledgements

I feel so indebted to many people during the time writing this paper and my life in KMD. Firstly, I would like to give my special thanks to Professor Akira Kato, my main research supervisor, for his constant help with my questions and being kind to guide my research from the very beginning. I am also grateful to my sub supervisor Marcos Sadao Maekawa, my sub supervisor, for his sincere advice and guidance on this paper.

My thanks also go to all the participants who took part in the validation test. They not only helped testing the system patiently, but also provided precious feedback and suggestions for me. I would like to thank all the members of Network Media for being kind and lovely to me. Besides, thank all the friends who have been working together with me these days for relieving the hard time and the companionship.

Lastly, I want to thank my family for supporting me in the past years.

Chapter 1
Introduction

1.1. Background

Photography is an important carrier for people to record life and share memory with friends. In old days, people take pictures, have photos developed and collect them to make albums. Sometimes they show photos to family and friends to share memory with them; sometimes they enjoy photos that are taken several years ago and reminisce the old time.

Nowadays it is easier and easier to take photos with smartphones. As long as your smartphone is functioned with camera, you can take pictures freely at any time. There is no need to develop photos because people can keep and share photos digitally through social networking service (SNS) now. People post photos on public social media or send them through private messenger. There are various kinds of platforms that are popular among young generation these years, such as Instagram (summarized in 2.1.1), Facebook (summarized in 2.1.2), LINE (summarized in 2.1.3), etc..

However, there are also issues that many photos have so limited information that the actual feeling cannot be conveyed to the viewer. There is too much information to browse and absorb on SNS that sometimes people do not have enough time to understand the message behind the photo. Also, when people look back at their old photos, they cannot always remember the situation themselves. Photos can record the moment, but the surrounding information also matters to the photo taker. When sending photos to friends, they may not be able to truly understand the feeling of the photo taker due to many factors, which will lead to a bad communication among people in SNS.

1.2. Research Question

Photo sharing through social networking service is getting more and more common in daily life. People have to obtain tons of information a day from online applications and many of them interact on social media only by the "like" button. Gradually, people become insensitive to messages and photos from friends and tend to take in only the information superficially.

For people who are not good at taking photos, their photos might be hard for viewers to catch the most attractive information. It might not be able to convey the feeling of the photo taker. Except from this, even with the convenience of smartphone camera, in some places photography is not allowed. In that case people have to take photos out of the place or other objects instead. How can people express their feeling with a photo that is far from their experience?

The research would like to call attention to a further communication among people. When viewing the photo of others', people can get more emotion information other than content of photo or text explanation. When looking back old photos of their own, people can also remember their feeling at that time.

1.3. Research Purpose

It is assumed that emotion information can make people be more empathetic on other's experience. The purpose of the research is to design how to augment emotion information into photos and reproduce it, and to verify its efficiency. It is expected that the design can promote people to communicate deeper with others by photo sharing and to build a more reliable relationship.

EmoPhoto, a Mini-Program in WeChat, whose name means emotion information embedded photo, serves as a platform to store and translate the emotion information and present the information by other effects. Comparing with public social networking applications, people's connection is close and intimate in private chatting platforms. Hence, people are more likely to have stronger wish to share their feelings behind by sending photos in private chat.

The structure of the research is shown as the Figure 1.1. The system can be divided into two parts, augmentation of emotion information and presentation of

emotion information. The biggest challenge of the paper is how to present the corresponding emotion when people are viewing the shared photos. Visual, audio and some other approaches are possibly available for the research.

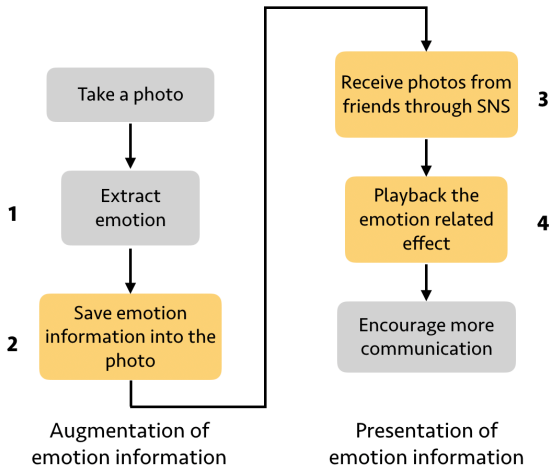


Figure 1.1 System structure of the research

After users take photos and extract emotion information, the system will save the emotion information into the photo and translate the information into other forms. When their friends receive the photo sent through the private messenger, they can enter EmoPhoto system and interact with the photo to playback the effect. If the system is efficient enough, deeper and more motivated communication can be expected, which can improve the relationship among people.

1.4. Scope of Research

As shown in Figure 1.1, it is supposed to be 4 parts in this research. For Part 1 *Extract Emotion*, the ideal solution is to recognize human emotion automatically and embed it into photos. There are related researches are like emotion recognition through physiological signals, electroencephalography (EEG), facial expressions, etc.. However, considering the accuracy of the recognition algorithms, it is hard to gain a precise result and will cause errors to following steps. Hence, in this paper, manual input will be employed to embed emotion information.

Consequently, the main scope of this paper is focusing on Part 2, 3 and 4 in Figure 1.1. The issues that will be discussed in the following chapters are saving of emotion information, photo sharing through SNS, and interaction to playback the embedded effects.

1.5. Contribution

It is said that good photography can convey emotion to viewers. But for many photos taken by common people, the images sometimes cannot express meaning as they expect them to be. There are many researches on emotion and photos, mostly focusing at image emotion recognition, in a scientific way, and their test photos are usually professional photos looked beautiful.

In this paper, emotion is not recognized by algorithm but is presented through different senses. The paper discusses the possibility that emotion of images can be stored, translated, and can arouse similar feelings in viewers. With the development of technology and online social interaction, the importance of deep communication is somehow neglected. EmoPhoto is designed to induce viewers' emotions and let them be more empathetic on other's experience, which eventually lead to a further communication and build an intimate and reliable relationship.

1.6. Structure of the Thesis

- **Chapter 1 Introduction**

  Introduce the background of the research, the question, the purpose, and the contribution of the design.

- **Chapter 2 Literature Review**

  Present the literature review and related work about photo sharing, emotion on photos, colors to enhance emotions, audio to enhance emotions and some emotion models.

- **Chapter 3 Concept & Methods**

Explain the concept of the design, including the component to realize the design and designing process.

- **Chapter 4 Proving Concept & Result**

  Conduct a proof of concept to verify the efficiency of the design and analyze the data and feedback according to the users' assessment.

- **Chapter 5 Discussion**

  Find out the value of the design and propose several ideas for optimization in future work.

Chapter 2
Literature Review

To construct a better system, researches from different fields as theory evidence for the design will be introduced in this chapter.

Current platforms like Instagram, Facebook, LINE are popular among young generation. Every day there are numerous photos uploaded to these platforms. Owing to the convenience of these services, people are motivated to upload photos online. It is also a way for people to build community and feel connected.

Colors are commonly used to express emotions in many different cultures. There are also many researches studying the color-emotion relationship. Audio has been applied in films, video games to create certain atmosphere and induce emotions in people, which can be another medium in this paper to help arouse empathy in users.

Emotion is the basis of the research. There are conceptual models and quantitative models in emotion research. Different models have different advantages in studies. Therefore, it is necessary to build a emotion model that meets the requirement of EmoPhoto in the following design.

2.1. Current Platforms

As the social networking service (SNS) grows, there are more and more different user-targeted services appeared on the smartphone platform. Most of them have the function to upload photos and share with friends. Here are some most popular and typical instances among them.

2.1.1 Instagram

Instagram (see Figure 2.1(a)) is one of the most popular photo and video-sharing social networking services. Users can take photos, apply filter to the original photo and upload it to the networking service. People can add text explanation or emoji stickers to enrich the information of the photo. Viewers can browse photos of users who they are following in their timeline. They can interact with the posts by "like" button and comment the photo. By September 2018, the number of pictures and videos uploaded to Instagram has reached over 100 million[1].

Similar service: Flickr, Pinterest

2.1.2 Facebook

Facebook (see Figure 2.1(b)) is one of the biggest technology companies that provide online social media and social networking service in the world now. Users get connected with their classmates, working partners and other acquaintances. They post text, photos and multimedia of their own and share it with other users as "friends" publicly. People respond on these posts by comment or "like" button to show their opinions.

As the most popular social networking service in the world, the number of photos uploaded to Facebook is also enormous, which reached 4,000 per second by 2018[2].

Similar service: Twitter

2.1.3 LINE

LINE (see Figure 2.1(c)) is used more likely as a messenger application than social networking service. Users can chat freely on the application and send text, photos or videos to each other. Since it is a more private service, some people tend to user LINE to build deeper connections with friends. People interact with each other directly through the private chat.

LINE is growing with more extended features, like digital wallet, as well.

1 https://www.omnicoreagency.com/instagram-statistics/

2 https://www.omnicoreagency.com/facebook-statistics/

Similar service: WhatsApp, WeChat

Figure 2.1 Current public and private photo sharing platforms

2.2. Photo Sharing on SNS

As introduced in 2.1, there are tremendous number of users active on different social network service uploading photos. Traditionally people share memory with the printed photos and decorate their albums. The target audience of photo sharing is usually friends who are able to meet in person. But as the convenient online service on mobile devices grows, the "ease of sharing photos" and "high quality of photos displayed" have enormously increased people's motivation towards photo sharing [1]. Besides, a research of Chei Sian Lee et al. [2] which studied the motivation of photo sharing on Instagram found that users' photo sharing was mainly motivated by the need for self-representation (a form of creating and experimenting with identities) and status-seeking (how sharing photos helps one to attain status among peers).

Other than SNS like Instagram, a permanent photo sharing community, on instant messengers like Snapchat, young people share photos in order to talk about images or communicate visually. It is also a common way to build intimacy

on SNS [3]. Sending photos is similar to sending text as a communication tool, but with richer information.

Due to technological affordances and social needs, the behavior of photo sharing has become an important approach for people to build community and feel connected with others.

2.3. Photo Emotion Recognition & Classification

Photos, or images are important visual contents for people to communicate in daily life. A photo contains a lot of information that is visible and perceivable. Usually since the photo gives a specific image for viewers, it is more easily to induce certain emotions in people.

2.3.1 Relationship Between Images and Emotions

There have been many works studying the relationship between photos and emotions. The emotions that different people perceive from a photo can be influenced by multiple factors, visual content, social context, temporal evolution, location influence, etc. [4].

Thus there are also various researches working on the analysis of image emotion information. Mao et al. [5] analyzed the affective characteristics of different images. Color vision histograms of images and their basic statistical properties are analyzed, and fractal dimensions (FDs) of these images are estimated by wavelet transforming method. Then semantic difference (SD) method is adopted to evaluate images of different affective features. Through the research, it can be proved that the FDs of images were related with the image's affective properties.

However, the feelings that the research studied about are relatively simple, from dull, harmonious, to upset. The mussier the FD of a image is, the more upset it will make people feel, and vice versa.

2.3.2 Image Emotion Classification

Although the feeling upon a photo is very subjective, it will be helpful if images can be retrieved based on emotional content. Thus there are some algorithms

developed to classify image emotion. Machajdik et al. [6] developed methods to extract and combine low-level features representing the color, texture, composition, and content of an image, and use these for image emotion classification. The experiment requires a large amount of photos for testing and training, whose data sets are from: (1) the International Affective Picture System (IAPS), (2) a set of 807 artistic photographs downloaded from an art sharing site, (3) a set of 228 peer rated abstract paintings. The test results show that the classifier has the best performance when it basically detects the portraits, that is, it can infer the emotion from facial expressions. But for artistic photos where faces do not determined the emotion of a whole photo, it has to rely on color features. This approach has advantage on photos with semantic-based features, such as the recognition of facial expressions, or certain common symbols (e.g. hearts).

A further study was conducted by Rao et al. [7] in 2016. The team proposed a new deep network that learns multi-level deep representations for image emotion classification (MldrNet). As introduced before, image emotion can generally be recognized through image semantics, image aesthetics, and low-level visual features. The proposed MldrNet combines deep representations of different levels to effectively classify the emotion types of different kinds of images, such as abstract paintings and web images. Comparing with other former approaches, the research improved a lot, especially when the emotions of the images are mainly conveyed by mid-level and low-level visual features, such as color, texture and image aesthetics.

The biggest problem, however, is that these approaches for image emotion classification all aim at the viewer of the photos. Besides, the test photos are almost artistic photographs that are professional and beautiful. Usually the color and texture are well designed. The image semantics is well considered.

But the target of this paper is the emotion of photo takers. For common people, the photos do not always contain with elements that can accurately express their feeling. Even for photos with human faces, the facial expressions are not equal to the emotion of the photo taker. Consequently, in order to convey a certain emotion to viewers, it is necessary to augment the emotion information into photos to make it easier for viewers to obtain the emotion information.

2.4. Emotion Detection Algorithms

As mentioned in 1.4, there are multiple ways to detect human emotions by scientific approaches. When people are in different moods, people will behave differently in physiological level. It can be easily inferred from daily experience that when people are upset, they tend to feel weaker than normal and when people feel nervous, their heart beats will be faster than usual. During the research, two kinds of approaches were studied to explore its feasibility in this design.

2.4.1 Vital Signs Based Emotion Detection

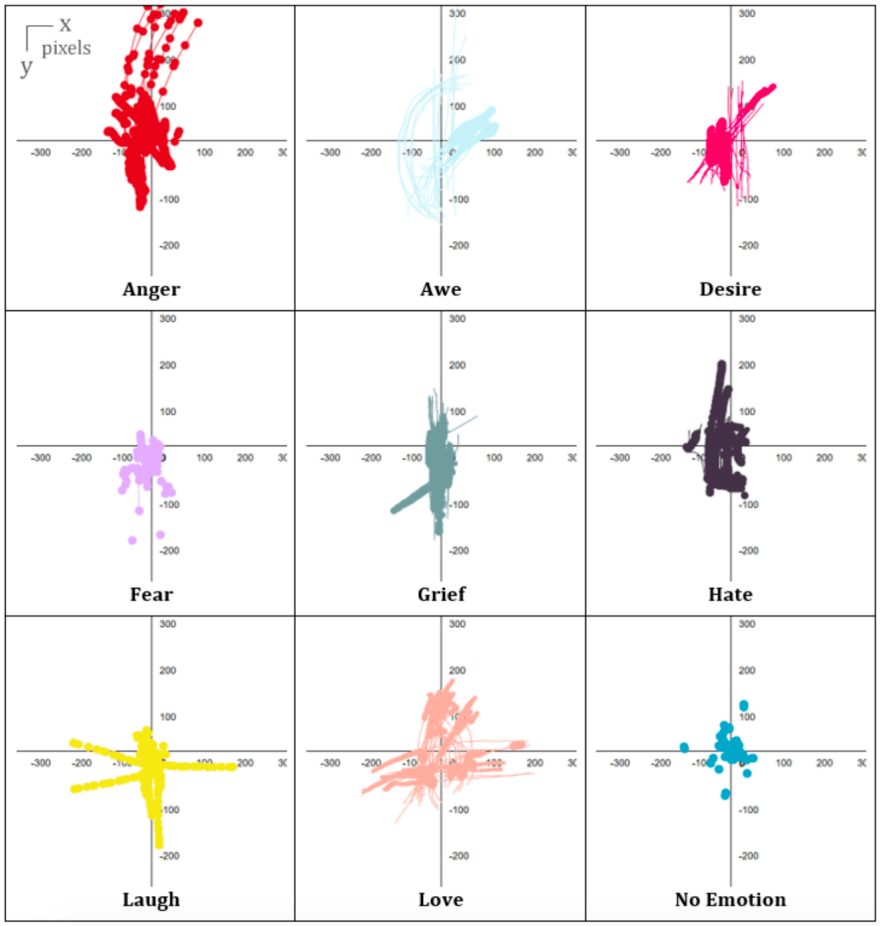
The work of Ménard et al. [8] focused on emotion recognition using biofeedback sensors, which measure heart rate and skin conductance. Participants were shown a series of 12 video-based stimuli that have been validated by a subjective rating protocol. The skin conductivity and the heart-rate were measured by the TEA sensor and the Nonin sensor. The collected signals were processed with a Fourier Transform to get the Fourier coefficients, and then classify these coefficients with the Support Vector Machine. The computed classification rate of skin conductivity was 85.57% and the computed classification rate of heart rate was 89.58%.

Despite of the accuracy, the participants have to wear electronic equipment with them. It is inconvenient in actual application to ask users to put on devices. Besides, when taking photos in daily experience, it is hard to keep the user in a calm status. They might be walking, or climbing stairs. It will greatly influence the accuracy of the emotion detection.

2.4.2 Touchscreen Emotion Recognition

A recent research of Alicia Heraz et al. [9] chose force-sensitive screens to recognize emotions. The participants, who are good imagers and emotionally aware, were trained to express each emotion on a force-sensitive smartphone. Researchers collected hundreds of thousands of their touches, and applied feature selection and machine learning techniques to detect emotions from the coordinates of participants' finger touches, amount of force, and skin area, all as functions of time. For example, Figure 2.2 shows a single frame taken from participants' finger expres-

sions of each emotion at a random instant t on the (x,y) axis. The experiment achieved a high 91% classification accuracy.



(Source: Alicia Heraz and Manfred Clynes [9])

Figure 2.2 Patterns of emotional expressions on the (x,y) axis

However, the technology is still too new to be put into practice. There are too many requirements for participants and their smartphone devices. The participants need to be trained to express each emotion as precisely as possible and the smartphone must be force-sensitive enough to detect the finger motions. For general users, it may also decrease the accuracy in application.

Due the limitation of the emotion recognition algorithm, this paper will use manual input instead of emotion recognition in order to present the emotion in-

formation in a more accurate way.

2.5. Colors to Enhance Emotion

There has been hundreds of researches on the relationship between colors and emotions. Nevertheless, the cultural difference is an inevitable influence factor that determines how people perceive from a certain color.

A cross culture color study of Hupka et al. [10] was conducted in Gersnany, Mexico, Poland, Russia, and the United States. Despite the cultural difference, black and red are associated with anger and black is associated with fear universally, which shows part of the common feature of human being's sensitivity. There is also experiment dealing with colors and emotions in videogames [11]. From their experiment, the emotional responses of player were measured and analyzed. It indicates that red elicits a highly-aroused, negative emotion while yellow elicits a positive emotion in players.

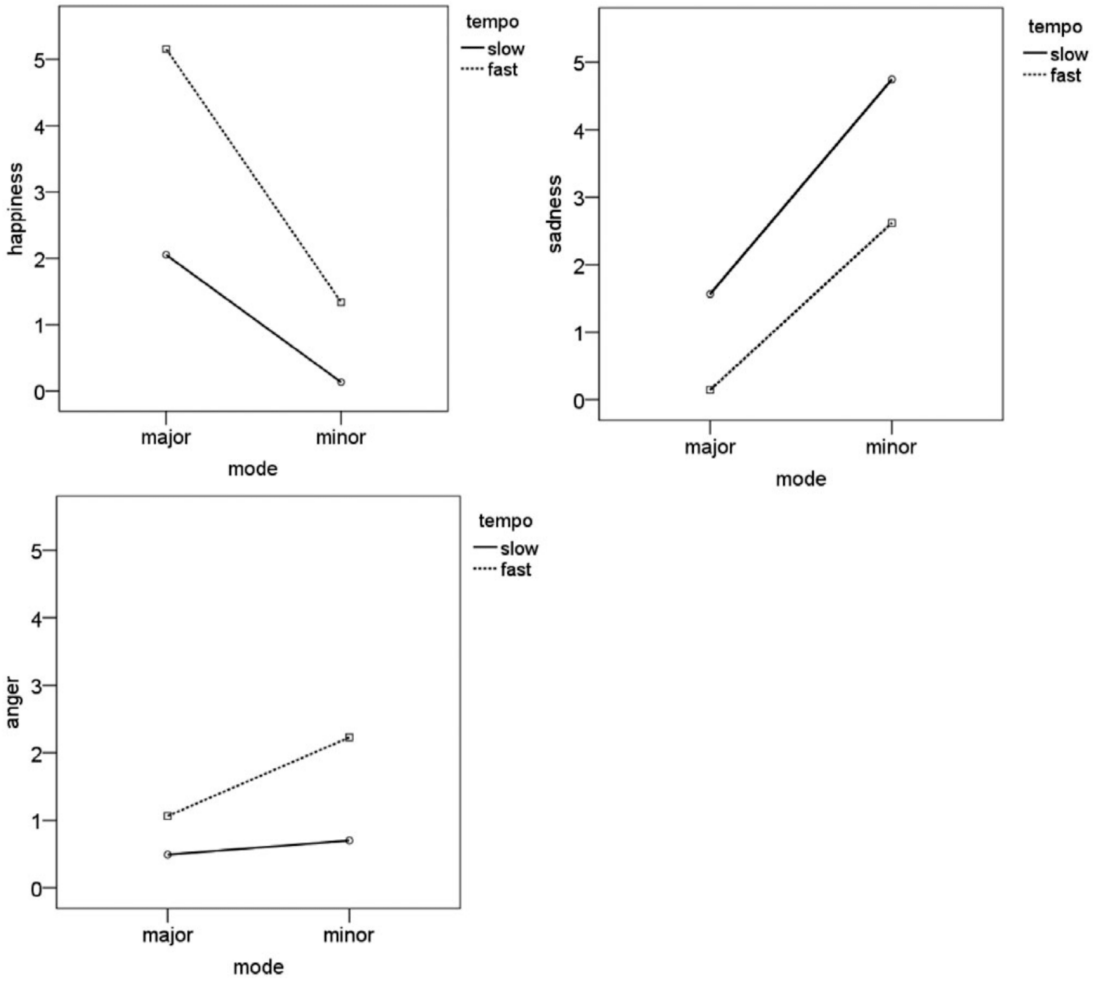
Furthermore, a series of psychophysical experiments were conducted to derive color emotion models for single colors and two-color combinations by Ou et al. [12–14]. Activity, weight, and heat, as the color-emotion factors were applied in research of Martin Solli et al. [15], which proved that people in general perceive color emotions in similar ways for multicolored images. Learning from the studies, we can see that color combinations can provide a more accurate and rich method to deliver emotions than single color.

As colors are essential factors to change people's perception on images, some image processing researches are aimed at transfer color to evoke different emotions as well. Work of He [16] and Liu [17] are both focusing on proposing a color transfer framework to convey rich emotions. Their efforts are important inspiration of my design.

2.6. Audio to Enhance Emotion

Sound effect serves as an audio content in many different media to emphasize their artistic contents. Its wide use is also due to its effect on inducing emotions in many ways. There are a lot of researches on emotion induction through music [18, 19]

that music can induce basic emotions in people, such as happiness, sadness, and fear. Concerning cultural difference around the world, several cross-cultural experiments are conducted as well. Hauke Egermann et al. [20] compared Canadian and Congolese pygmies listeners and found that emotion arousal was quite common between the two cultures. Thomas Fritz et al. [21] had native African and Western participants to study and verified that happy, sad, and scared emotions in Western music can be recognized universally.



(Source: Juslin and Lindström [22])

Figure 2.3 Significant two-way interactions between tempo and mode in listeners' mean ratings of (a) happiness, (b) sadness and (c) anger

Pitch, tempo, and timbre are generally considered to be factors that influence people's perception on different music [23–26]. According to research of Patrik N. Juslin and Erik Lindström [22], different features appeared to be important to different emotions (e.g. mode for "sadness", timbre for "anger", articulation for "fear"). For instance, Figure 2.3 shows how the two-way interactions between tempo and mode induce different emotions.

2.7. Emotion Models

Research on emotions never ends. There are a lot of theories trying to classify emotions of human being. Some of them are conceptual models while some are quantitative models.

In 1989 Johnson-Laird and Oatley [27] have come up with an analysis of emotion words and stated that there are five basic emotions: happiness, sadness, anger, fear, and disgust. Later in an argument of Paul Ekman [28], he mentioned that surprise expressions were recognised across literate cultures. Surprise was distinguished from anger, disgust and happiness, but was difficult to be distinguished from fear.

Thayer's emotion model [29] applied two dimensions of energy and stress. In the model, the four basic emotions Anxious, Excited, Content, Depressed are measured quantitatively (see Figure 2.4).

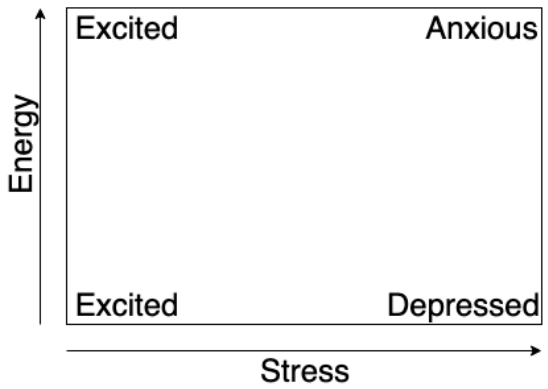


Figure 2.4 Thayer's emotion model

A three-dimensional PAD (Pleasure-Arousal-Dominance) [30] was also devel-

oped to represent emotion state by Mehrabian. The model has been used to study nonverbal communication in psychology.

Chapter 3
Concept & Methods

Based on the literature review of Chapter 2, a series of design component are determined to build up the prototype.

A fieldwork was carried out before the design. The interviews are about how people take photos and share with friends, and how they feel when looking back at old photos.

The design components are built based on the literature review in the previous chapter, including the basic emotions, color combinations to represent different emotions, Exif data to store emotion information, and WeChat Mini-Program as a platform for running the prototype.

The design of the system is divided into two parts: augmentation of emotion information and presentation of emotion information, in which the details of design are introduced. To verify the efficiency of the design, a pre-test was also conducted to test the color combinations. On the basis of the its test result and feedback, the design was optimized.

In the implementation, two journey maps are used to describe the user journey of both sides and it helps to understand users movement better. The prototype was built in WeChat Mini-Program and can functionally embed emotion information into photos and present with visual and audio effects.

3.1. Fieldwork Research

The research started from real life experience. Before the design, it is necessary to have a deeper investigation. Fieldwork among different people was carried out to know more exactly how people share photos and their feeling with friends. Three interviews are as follows.

The main interview questions are:

- When do you usually take photos? How do you feel when taking photos?
- Do you usually share photos with friends? Will you explain your feeling to them?
- How do you feel when look back at your old photos?
- How do you feel when people show you their old photos?

Case 1 Aged 24, Interviewee A is a female graduate school student. She said she took photos quite frequently in life, especially in travelling. Interesting and beautiful things motivate her to press the shutter so she generally feels happy when taking photos. However, except for some special events, she doesn't post photos to public SNS very much. She tends to send photos through private messenger to friends or family and introduce the time and place of the photo.

Talking about old photos, "I want to destroy photos that I didn't look good." She laughed, "But photos of my childhood will be good memory for me. " Besides, it is always interesting and fun to see other's old photos.

Case 2 Interviewee B is a 52-year-old housewife. She said she loved to take photos during travelling for memories and she usually felt happy and excited for it. She will post photos both publicly and privately but she is not good at express her feeling. Usually she just introduced the time and place. She also explained how she felt when looking back at old photos, "The old photos are precious and good memory for me because it is not easy and not cheap to take pictures when I was young. "

Case 3 Interviewee C is a 45-year-old junior school teacher. He usually takes photos when sightseeing and some special events. Also as a teacher he often takes pictures for students when they are taking part in some activities. He found that it is important to keep beautiful and ephemeral moment in life. He likes to share photos with friends and express his emotion in a few lines. As for old photos, he felt so nostalgic and moved by the growth of his students. Sometimes he will send old photos to his students or friends and enjoy talking about old times with them. "Memories are beautiful, aren't they?"

3.2. Components

The research is to augment emotion information into photos so that people are able to perceive more emotion information other than primitive pictures and be more empathetic on other's photos.

The following chapter designed how to augment emotion information into photos and how to translate emotion information to other form.

3.2.1 Basic Emotions

As stated in Chapter 2, there are many applicable emotion models in past researches. As the goal of EmoPhoto is to provide a convenient service to embed emotion information into photos, conceptual models are more likely to be understood by users. Consequently, in this paper the basic emotion argument of Paul Ekman [28] was employed to build the foundation of the system, which includes *Happiness, Surprise, Anger, Disgust, Fear*, and *Sadness*.

In order to grasp the emotion model intuitively, a hexagon model (see Figure 3.1) was applied to reflect the conceptual model into graph. Value of each emotion ranges from 0 to 1.

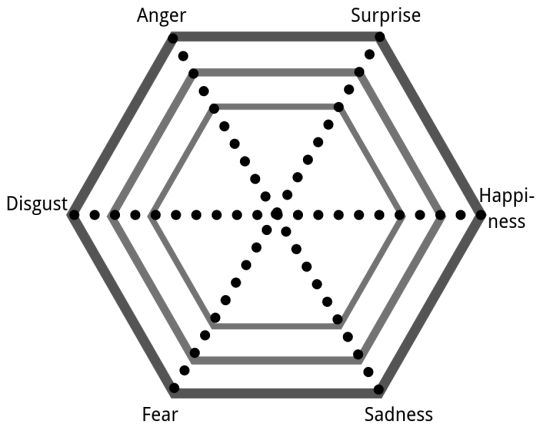


Figure 3.1 Basic emotions that employed in the paper

3.2.2 Color Combinations

Color is an important tool to express emotions in many cultures. Also, as mentioned in Chapter 2, comparing with representing emotions by single color, it is more accurate to express emotions by multiple color combinations.

According to the research of Nippon Color & Design Research Institute Inc. [31], color combinations can be used to express people's different emotions and characters. The projective method is applied in the study to learn the specific color combinations that are widely used to express human emotions. Considering the basic emotions mentioned in the previous section, the following color combinations (see Figure 3.2) are designed as a color mapping between emotion and colors.

Surprise, as one of the basic emotions, is rare and hard to be expressed by color. Consequently, another sensory, vibration, is employed to represent the emotion in EmoPhoto. It is expected that a sudden vibration during viewing photos will directly induce surprise in users.

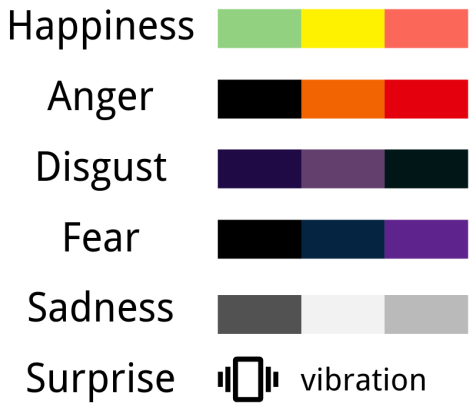


Figure 3.2 Color combinations of each basic emotion

3.2.3 Exif Data

Exchangeable image file format (Exif) is a standard that specifies the formats for multiple media files. The Exif image image file specification [32] stipulates the method of recording image data in files, and specifies the following items:

- Structure of image data files,
- Tags used by this standard,
- Definition and management of format versions.

In the paper, Tag "UserComment" and "RelatedSoundFile" are employed to store emotion information into photo.

3.2.3.1 UserComment

A tag for Exif users to write keywords or comments on the image besides those in ImageDescription, and without the character code limitations of the ImageDescription tag [32]. The information of the tag is shown in Table 3.1.

Table 3.1 Tag Information of UserComment

Tag ID	37510
Type	UNDEFINED
Count	Any
Default	None

3.2.3.2 RelatedSoundFile

This tag is used to record the name of an audio file related to the image data [32]. In the paper the related sound effect audio file's name is written in the Exif data. The information of tag RelatedSoundFile is shown in Table 3.2.

Table 3.2 Tag Information of RelatedSoundFile

Tag ID	40964
Type	ASCII
Count	-
Default	-

3.2.4 WeChat Mini-Program

WeChat Mini-Program is a feature launched by WeChat in 2017, which is an application within an application (WeChat). Each user of WeChat is able to get access to the Mini-Program. The pros and cons of Mini-Program are:

Pros:

- Convenient for existing users. For the Mini-Program is embedded in an existing chatting service, there is no need to register or sign in.
- Sent through chat. The Mini-Program is allowed to share pages or sent through chatting, which makes it easy to share information with friends.

Cons:

- Specific users. Only users with WeChat account can access to it.
- Limited features. WeChat requires that the coding package of a Mini-Program has to be less than 2MB. Hence some functions are difficult to be realized merely by Mini-Program.

A platform embedded inside a chatting application is more likely to get utilized by users rather than a brand new independent application. Due to its accessibility, Mini-Program is considered to be the most appropriate platform for the service to be developed.

3.3. System Design

There are two main parts of the design: augmentation of emotion information and presentation of emotion information.

The first part, augmentation of emotion information, discussed about the method for users to embed emotion information, tried to apply an original expression of emotion and save the emotion information into Exif data of photos. The section of presentation of emotion information explored approaches to retouch the photos according to the embedded emotion information, and selected audio files connected with the emotion.

3.3.1 Augmentation of Emotion Information

3.3.1.1 Emotion Embedment

One of the issues is that how to embed emotion information conveniently and naturally in photos. Since the service acts similarly to some photo retouch application, the photo editor functions are taken into consideration as the prototype for the emotion embedment.

Figure 3.3 shows the photo editor of iPhone. Just like people adjust saturation from minimum to maximum in the editor, it is also possible that people can upload photos to the system and adjust emotion information from 0 to 1, value of the six basic emotions, by emotion editor in EmoPhoto.

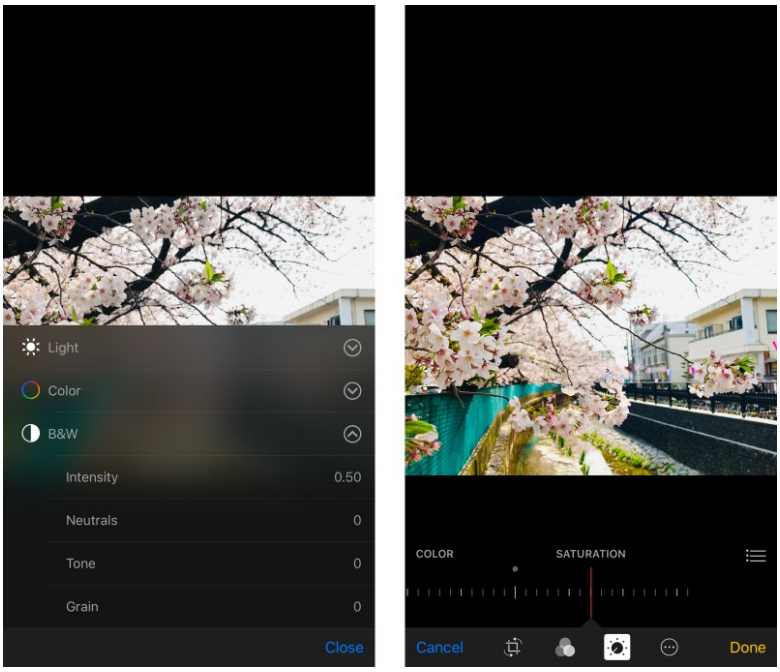


Figure 3.3 Photo editor of iPhone

3.3.1.2 Information Saving

1. Array expression of emotion

   After the system ensured the emotion of the photo taker, our task is to save the emotion information into the photo. As stated above, Exif data is the

metadata of photos and there are several tags available to be written. In the paper, **UserComment** and **RelatedSoundFile** are selected as tags to store emotion related information.

As introduced in Section 3.1.1, a hexagon emotion model is applied as Fig.3.1. According to the emotion model, human's basic emotion will be stored in the following array expression. The parameter in the array are as Equation 3.1:

$$Emotion = \{[Happiness, Surprise, Anger, Disgust, Fear, Sadness]\} \tag{3.1}$$

in which Happiness, Surprise, Anger, Disgust, Fear, Sadness $\in$ [0, 1]. In the following paper, for the convenience of expression, H is used to stand for *Happiness*, Sp for *Surprise*, A for *Anger*, D for *Disgust*, F for *Fear*, and Sd for *Sadness*. The default status in Eq. 3.2 means that the person feels nothing at the moment while all the six emotions are 0.

$$Emotion\{[H, Sp, A, D, F, Sd]\} = \{[0, 0, 0, 0, 0, 0]\} \tag{3.2}$$

When the person had a great meal and he felt super happy and a bit surprise, his emotion can be expressed as Eq. 3.3:

$$Emotion\{[H, Sp, A, D, F, Sd]\} = \{[0.8, 0.3, 0, 0, 0, 0]\} \tag{3.3}$$

in which Happiness = 0.8, Surprise = 0.3.

The array expression is saved into the Exif tag **UserComment**.

2. Emotion change during photo taking

   In my assumption, photos can be roughly divided into the following two types: usual situations and unusual situations (see Fig.3.4). In the usual situation, one of the biggest motivations for people to take photos is the record of life. For examples, a photo of today's beautiful sunny sky, or a meal cooked by yourself. In daily situations, it is hard to guarantee that people will definitely share and look back at the photos. Many photos will be deleted later. So the possibility to record, share, and look back a certain photo shows the significance of the photo.

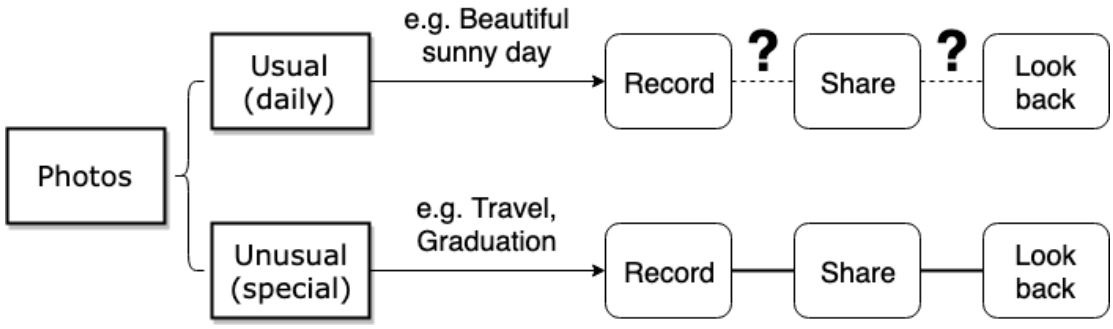


Figure 3.4 Two main situations of photos

While in an unusual situation, people will be more likely to share photos with friends, also look back at the photos to reminisce the old time. Since the time when photography is a rare and expensive event for people, it has served as a medium to memorize important moment of life. Even until now, people still cherish the photos taken during special moment, like travelling or graduation. With the convenient SNS tools, they can easily share these moments with friends, as well.

It can be inferred that the value of photography is more continuous rather than instant. The corresponding emotion upon a photo can possibly change over time. Imagine when you take out an album of your junior high school that you graduated over 20 years ago, it must be a mixture of feelings, in which some could be happiness and some could be nostalgia. It is assumed that the change of emotion is also important information for people. It can also be stored by array expression of emotion, such as Eq. 3.4

$$Emotion\{[H, Sp, A, D, F, Sd]\} = \{[0.7, 0, 0, 0, 0, 0.2], [0.1, 0, 0, 0, 0, 0.8]\} \tag{3.4}$$

The expression shows that the person felt quite happy and a little sad at first, and later when he looked back at the photo the strongest emotion he felt is sadness. More emotion change examples are shown in Table 3.3.

Human emotions are extremely complicated and it is rare to have purely one emotion in a situation. As analyzed above, emotions are expressed in a mathe-

Table 3.3 Emotion change in different situations

Situation during taking photo	Situation when looking back at the photo	Emotion{[H, Sp, A, D, F, Sd]}
Before the meal	After the meal and its taste was within expectation	{[0.8, 0.2, 0, 0, 0, 0]}
During graduation	Several years after the graduation	{[0.7, 0, 0, 0, 0, 0.2], [0.1, 0, 0, 0, 0, 0.8]}
Travelling with girlfriend	After breaking up with the girlfriend	{[0.6, 0.3, 0, 0, 0, 0], [0.1, 0, 0, 0, 0, 0.7]}

matical way. Among all the six parameters (basic emotions), the emotion whose corresponding value is the largest of all is considered to be the dominant emotion of the individual. For example, the emotion expression in the situation "During graduation" from Table 3.3 means that the dominant emotion was happiness (0.7) at first and changed to sadness (0.8) several years later.

Since it is difficult to present multiple emotions in the same image, in this paper only the dominant emotion will be selected to be transferred into other forms.

3.3.2 Presentation of Emotion Information

3.3.2.1 Photo Retouch

As introduced in 3.2.2, we applied color combinations to reflect the related emotions (see Figure 3.2). When the photo taker had multiple feelings the moment taking pictures, the dominant emotion is selected as the key emotion to be translated. For example, in Eq. 3.3, happiness outweighs other basic emotions so that the photo will be retouched by color combinations of happiness.

To retouch the tone of the original color properly, a program was designed to replace the color of each pixel through Python. In the research, HSV color space was employed for better modeling. The HSV color space (Figure 3.5) is an alternative representation of the RGB color model and attempts to characterize colors according to their hue, saturation, and value (brightness) [33].

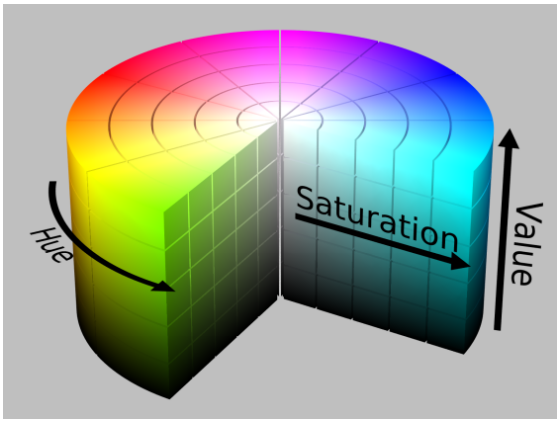


Figure 3.5 HSV color model

In Table 3.4 the six basic emotions and their corresponding HSV data of color combinations are listed.

Table 3.4 Emotions and their corresponding HSV data of color combinations

Emotion	HSV1	HSV2	HSV3
Happiness	(107,39,82)	(57,100,100)	(5,65,98)
Surprise	-	-	-
Anger	(0,0,0)	(25,100,95)	(357,100,90)
Disgust	(262,86,27)	(286,43,43)	(183,96,9)
Fear	(0,0,0)	(209,94,25)	(274,75,55)
Fear	(0,0,32)	(0,0,95)	(0,0,73)

To reserve the content information in the photo maximally, the design chose to alter the hue attribute of each pixel and keep the saturation and value attribute same as the original photo. Otherwise the photo will be consist of only three colors and be unrecognizable for users. In the program, the retouch is processed as follows:

- Get RGB attributes of each pixel and convert them to HSV
- Choose the closest Hue from the color combination of corresponding emotion
- Convert the color back to RGB

The Python code is shown in Appendix A.

3.3.2.2 Audio Effect

Audio is another method for people to express emotions. Based on the literature review in Chapter 2, different audio files are selected to represent each emotions (see Table 3.5).

Table 3.5 Audio description of each emotion

Emotion	Audio Title	Features
Happiness	Dream harp [34]	Fast tempo, major mode, high pitch,,bright timbre, smooth and fluent rhythm of harp
Anger	Cellar Door Slam Shut Sound [35]	Sound effect of a cellar door slamming shut
Disgust	Squish [36]	Sound effect of great squishy, or squelchy things
Fear	Creepy Background [37]	Fast tempo, dissonance, fast vibrato rate
Sadness	Sadness cosmo [38]	Slow tempo, minor mode, low pitch, dull timbre

There are music excerpts and sound effects among the audio files. Music excerpts are selected according to the musical features analyzed in [22]. The audio *Dream harp* is a 2-second harp excerpt whose features are like: fast tempo, simple and consonant harmony, bright timbre, etc., which correlated with happiness much. Audio *Creepy Background* for fear is fast tempo, dissonant, fast vibrato rated, which close to features of music related to fear. *Sadness cosmo*, an excerpt of piano and cello, also matches features of sad music: slow tempo, low pitch, dull timbre, etc..

For emotion of anger and disgust, sound effect rather than music is chosen to present the corresponding emotion, which is expected to give a certain image of situation for listeners and induce emotions in them. The sound effect of anger is a door slamming shut while the other one is sound of squelchy things.

3.3.2.3 Vibration

As mentioned in 3.2.2, due to the difficulty of representing emotion of surprise visually, the sense of touch, vibration, is introduced in the design.

Considering the possibility that one vibration is too instant for users to notice, **a continuous double vibration** is employed in the system.

3.4. Pre-test to Verify the Color Combinations

A pre-test is carried out to test the effect of the color replacement and what people perceive from the photo. Based on the result of the pre-test, the color combinations and the retouch approach can be improved better.

3.4.1 Pre-test Design

The most difficult but important issue in the research is how to reflect emotion correctly in the photos. As stated in 3.2, the original photos were converted into the corresponding color tones by the emotion definition (Table 3.4). However, the current color combination is summarized based theoretical study and its actual effect need to be verified. Three different photos (see Figure 3.6) are selected and converted the color of each photo into 5 patterns. *Surprise*, as we have introduced before, is hard to be expressed visually by colors. Hereby in the pre-test the paper only studied the other 5 color combinations. There were 11 people taking part in the small investigation and their opinions were collected. The expectation of the pre-test is that participants can perceive the certain emotion from each retouch of photo to some extent. Figure 3.7 shows one of the instances of the test photos that converted color tones according to their corresponding emotions. As listed below each photo, the figure consists of the original photo, happiness version, anger version, disgust version, fear version, and sadness version.

The investigation showed the participants each photo with comparison to its original one, like Figure 3.8, and asked what emotion they perceived from the retouched photo by comparing the two photos.

There are totally 15 retouched photos based on 3 original photos. In each comparison, participants are asked to rank the degree of each five emotion they

Figure 3.6 Photos used in the pre-test

(Original Photo credits: Pixabay)

Figure 3.7 Instance of the five different patterns to retouch the photo

Figure 3.8 Retouched photos (below) are shown by comparing with original one (above)

	0	1	2	3	4	5
Happiness	○	○	○	○	○	○
Anger	○	○	○	○	○	○
Disgust	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Sadness	○	○	○	○	○	○

Figure 3.9 Rank the degree of each emotion which ranges from 0 to 5

felt (see Figure 3.9).

3.4.2 Result & Evaluation

The age range of the 11 participants is between 20 and 35 years old, and the average age is 26. Ten of them are female. The average grading result of all participants is displayed in Appendix B, from which we can see the result is far below the expectation.

Suppose that the emotion whose average grade is the highest is the strongest emotion that most participants perceived, from the 15 sets of comparison we can find that 6 sets matched with the intended emotion. Among the remaining 9 sets, there are 2 sets of comparison in which the average grade of intended emotion is the second strongest. Roughly half of the photos can be perceived as expected.

Among the result of five different emotions, photos of *Sadness* performed the best and all three photos are perceived correctly as *Sadness*. On the other hand, photos of *Happiness* performed worst in the investigation.

At least one good news is that it is quite efficient to express sadness by black and white photos. Black and white colors make photos look faded and old which lead people feeling nostalgic. However, green from the color combination of *Happiness* is easily perceived as disgust which greatly influenced the combination's effect. Also according to participants' feedback, the context of photos themselves influenced how people felt a lot. For example, when the photo shows a delicious meal, most people feel happy and content regardless of the slight change of color.

Therefore, the prototype design is modified in the following research and is expected to perform better than the previous design.

3.5. Implementation

3.5.1 Journey Map

In order to build up a prototype that meets the requirement of users, the journey maps of sender and receiver are considered. This approach is to organize users' routine better and understand their actions when using the service. The user journey of both sides are as follows.

3.5.1.1 User journey of the photo taker

1. Photograph

   Take photos by smartphone as usual in daily life to record moment.

2. Embed

   Upload photos to EmoPhoto and input the emotion data through interface of the program. The photo retouch will be processed automatically through the program.

3. Share

   Share the photo through the platform and select friends to share with. The link will be sent through chatting.

4. Communicate

   After friends entering the photo display page and viewing the embedded effect, they can communicate further upon the photo about their feeling and experience.

The journey map is shown in Figure 3.10.

3.5.1.2 User journey of the viewer

1. Enter

   Receive the link from the photo taker through chatting and enter the page.

2. Interact

   Tap the photo to interact with it. Watch the visual effect, listen to the audio effect, and feel the vibration.

3. Communicate

   Tell their feeling or understanding towards the photo to the photo taker. Talk about the images or experience.

The journey map of the viewer is shown in Figure 3.11.

Figure 3.10 Journey map of the photo taker

Figure 3.11 Journey map of the viewer

3.5.2 Prototype

The system consists of two brunches: the retouch of photos and the platform of reproducing and sharing. As introduced in previous section, the retouch of photos will be implemented by Python and the platform will be built on WeChat mini-program. In the following section, the detail of each brunch will be explained.

3.5.2.1 Photo Retouch

The approach to retouch photos is optimized based on the pre-test result (Chap. 3.4) and a different color combination (Figure 3.12) is applied as "happiness" in this plan.

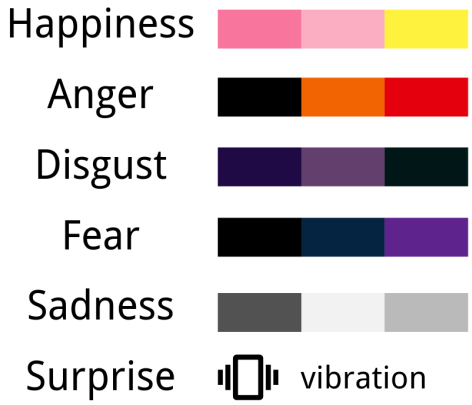


Figure 3.12 Optimized color combination applied in the prototype

Table 3.6 shows the renewed HSV data of the color combinations

In order to decrease the negative effect of green color, pink color is chosen in place of the former combination. Also the method to retouch color of each pixel is changed because in the earlier design when applying the hue attributes of the target color directly to the source photo, the retouch results did not vary much from each other. In the renewed retouch system, **overlay** blend mode was selected to optimize the output. Here is the definition for overlay blend mode from Adobe Photoshop [39]: Multiplies or screens the colors, depending on the base color. Patterns or colors overlay the existing pixels while preserving the highlights and

Table 3.6 Renewed emotions and their corresponding HSV data of color combinations

Emotion	HSV1	HSV2	HSV3
Happiness	(342,53,97)	(345,31,98)	(56,75,100)
Surprise	-	-	-
Anger	(0,0,0)	(25,100,95)	(357,100,90)
Disgust	(262,86,27)	(286,43,43)	(183,96,9)
Fear	(0,0,0)	(209,94,25)	(274,75,55)
Fear	(0,0,32)	(0,0,95)	(0,0,73)

shadows of the base color. The base color is not replaced, but mixed with the blend color to reflect the lightness or darkness of the original color.

However, when applying *Sadness* combinations to overlay the original photo, the output photo is merely darker in brightness, whose visual impact is less than the black and white effect. Consequently, in the retouch program, *Sadness* is separated from the other emotions and the photos will be applied the former retouch filter.

Figure 3.13 displays the optimized retouch result of the photo. Both the content information of original photo and the emotion related color information are emphasized in the result. The code is attached in the Appendix C.

3.5.2.2 Mini-Program

The mini-program EmoPhoto is developed by the official developing tool of WeChat. Here is the page structure including photo taking and playback function from the mini-program in Figure 3.14. There are two buttons on the top page of EmoPhoto. Tap the button “Camera” and the mini-program will jump to the camera page for users to take photos. Tap the button “Album” and the page will jump to the album page where photos will be displayed.

Figure 3.15 shows the flow of taking photo and selecting emotion information. Fig.3.15(a) is built-in camera function in mini-program. In Fig.3.15(b) users can confirm the photo just taken and choose “User Photo” button. In the edit page of Fig.3.15(c) users are able to set the score of each emotion. In the example picture

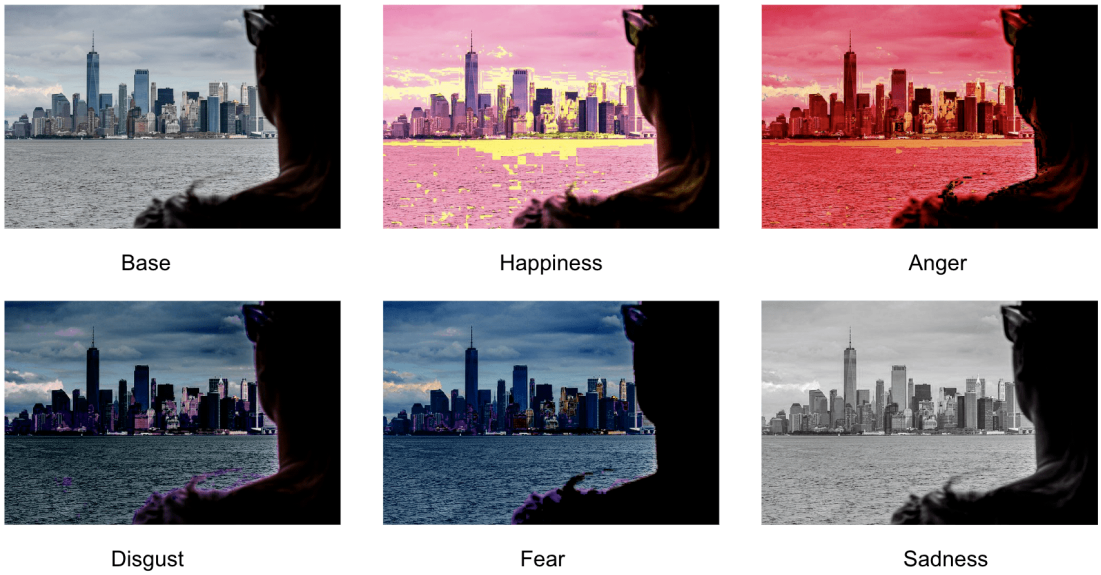


Figure 3.13 Optimized photo retouch of the five different emotion patterns

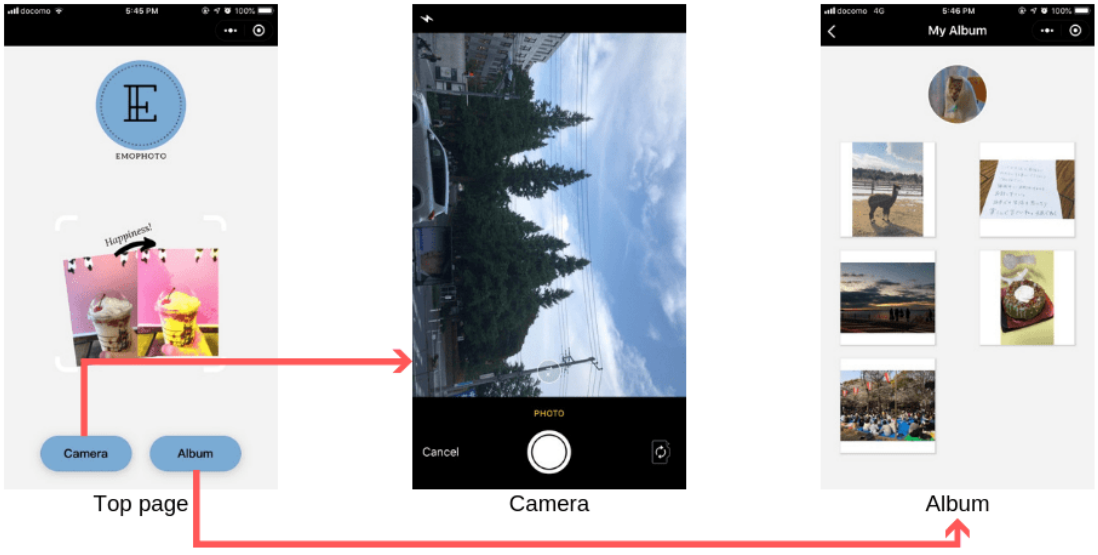


Figure 3.14 Page structure of the mini-program

Sadness is set to be 0.8. Lastly, when user click "Save" button on the top-right corner, the photo will be uploaded to the server and saved (see Fig.3.15(d)).

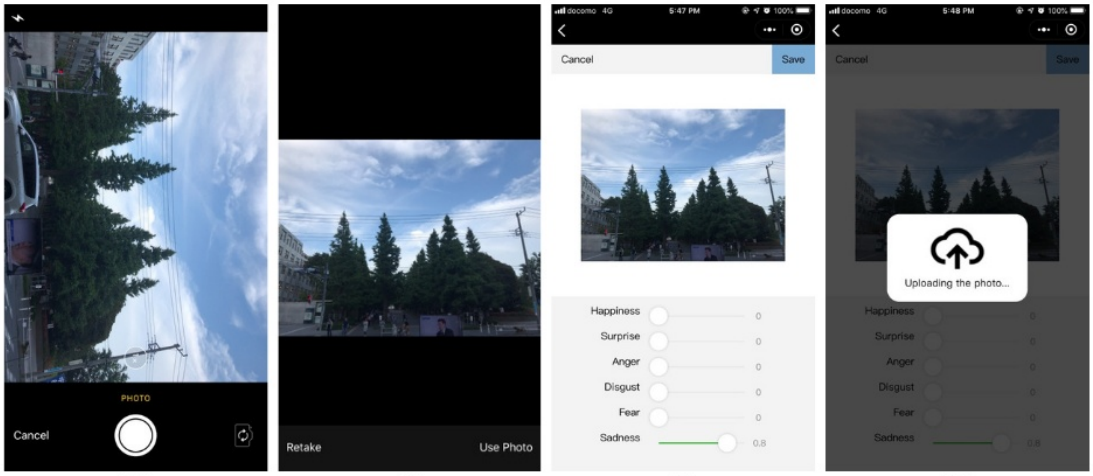


Figure 3.15 Flow when users take photos and choose emotion information

Figure 3.16 is the flow of review and interaction with the photo. After uploading the photo successfully (Fig.3.15(d)), EmoPhoto will automatically jump to the page of user's photo gallery (Fig.3.16(a)). The photo just taken is displayed at the tail of the photo list. Click the photo and enter the photo display page (Fig.3.16(b) and 3.16(c)). Since it takes a few seconds to read the stored Exif data embedded in the photo, there will be a loading window floating in front of the photo until the data gets loaded completely.

Fig.3.16(c) shows the original photo. Tap the photo, and EmoPhoto program will display the retouched photo and colored frame according to the emotion information embedded. The emotion related sound effect will be played at the same time. Tap again then the photo will return into the original version.

Figure 3.17 shows the flow when users want to share photos with their friends. Tap the button "Share with friends!", and WeChat application will show the personal list of friends, from which users can select their friend to send the photo. From the received link in Fig.3.17(c), other friends can enter the mini-program, view the photo, and interact with the photo as well.

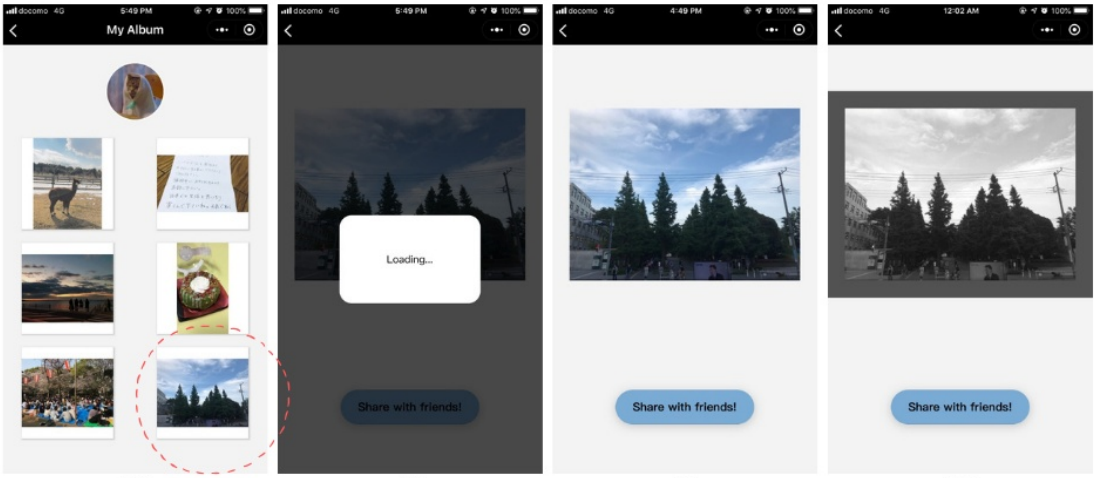


Figure 3.16 Flow when users review and interact with the photo

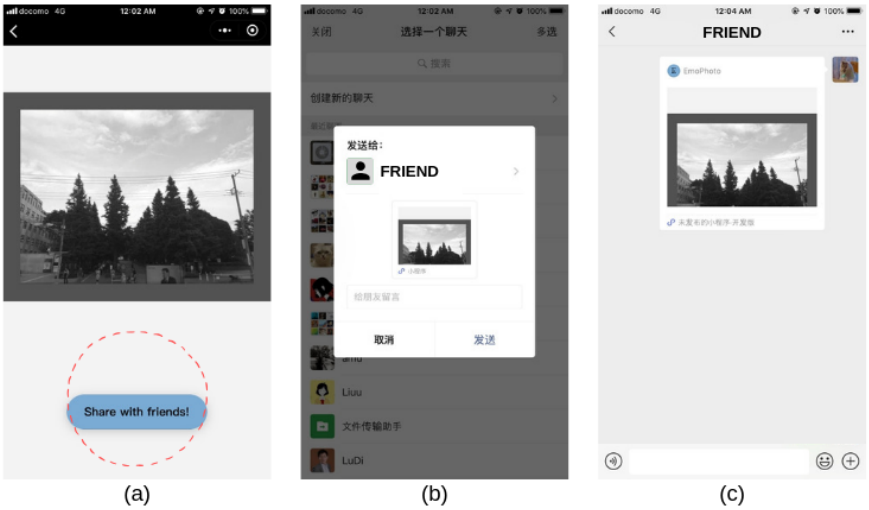


Figure 3.17 Flow when users share the photo with friends

Chapter 4
Proving Concept & Result

As the design component and designing process are introduced in the previous chapter, to verity the feasibility of the system, a proof of concept was conducted in this chapter.

Participants were asked to interact with photos through the Mini-Program EmoPhoto and assess their perception for three times, including photos with no effect, with only visual effect, and with the full effect (visual, audio, and vibration effect). The average of their emotions are calculated to evaluate the efficiency of the effects.

The test also received much feedback on the presentation and their opinions on photo sharing. On the whole, the test obtained many valuable suggestions that can be helpful for further improvement of the system.

4.1. Evaluation Test

4.1.1 Experiment Methods

1. Open the mini-program EmoPhoto on smartphone.

2. Tap and enter display page of each photo and watch the original photo. Assess the 1st impression of their own upon the photo in the assessment sheet based on the hexagon emotion model.

3. Re-enter the display page of each photo. Tap and watch the color change of the photo. Assess the 2nd impression of their own in the assessment sheet.

4. Put on earphones. Re-enter the display page of each photo. Tap the photo, watch the color change, listen to the corresponding sound effect, and feel

the vibration. Assess the user's perception in the assessment sheet based on the hexagon emotion model.

5. Evaluate the user experience and feedback.

The test photos are displayed as Figure 4.1. Participants need to evaluate photos in order from A to F according to the emotion model. The before and after change of each photo is displayed in Figure 4.2. The assessment sheet is attached in Appendix D.

4.1.2 Experiment Result

10 participants took part in the experiment to experience the mini-program EmoPhoto and evaluated each photo with corresponding visual and audio effect. Among all the participants, there are 5 Chinese, 2 Japanese, and 3 Indonesian. 6 of them are female and the other 4 are male. The age of the participants ranged from 23 to 57. Figure 4.3 shows how the participant was playing with EmoPhoto.

According to the result of emotion evaluation from the participants, the average emotion of 10 participants is calculated (Table 4.1). **Embedded Emotion** implies the actual perception of the photographer. **1st Impression**, **2nd Impression** and **3rd Impression**, which indicate the emotion participants perceived during the test, respectively stand for the perception when participants have no interaction with the photo at all (Step 2), the perception when participants obtain only visual change of photos (Step 3), the perception when participants obtain all the effect including visual, audio and vibrations (Step 4).

Graphs can help us better understand and compare the perception result. In this experiment, radar chart is employed to analyze the data in Table 4.1. The chart is shown in Figure 4.4.

4.1.3 Considerations

The radar charts intuitively display how users' perception changed each time when users interacted with the photo.

The embedded and dominant emotion in Photo A is happiness (0.8) and surprise (0.4). As the content of photo itself tends to influence viewers greatly, the first

Figure 4.1 Test photos

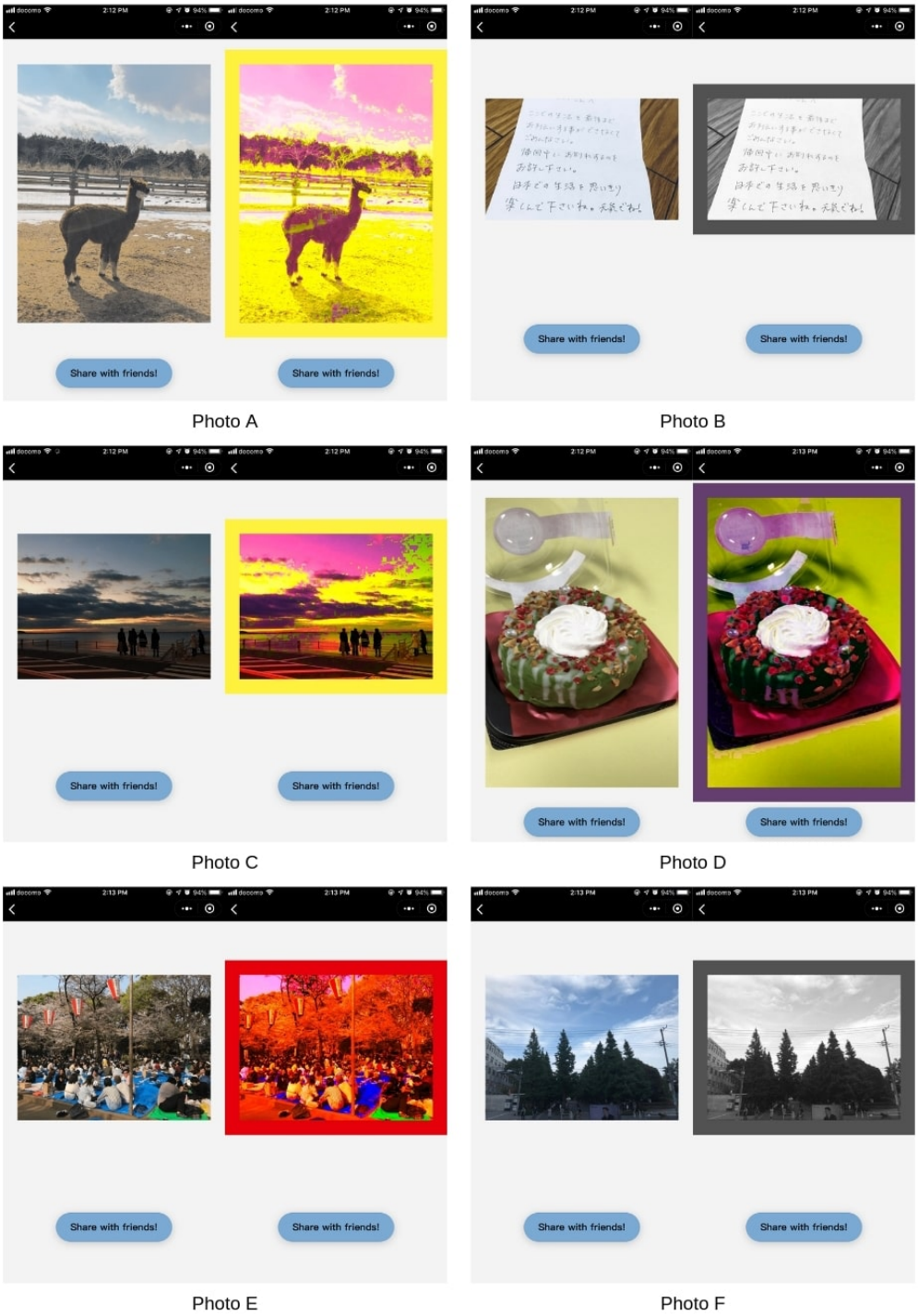


Figure 4.2 Before & after change of each photo

Table 4.1 Average perceived emotion in the experiment

Photo	Embedded emotion {[H, Sp, A, D, F, Sd]}	User 1st impression (No Effect)	User 2nd impression (Visual Effect)	User 3rd impression (Visual + Audio Effect)
A	{[0.4, 0.8, 0, 0, 0, 0]}	{[0.57, 0.02, 0, 0.02, 0, 0.02]}	{[0.19, 0.32, 0.03, 0.12, 0.15, 0]}	{[0.43, 0.25, 0.03, 0.03, 0.06, 0]}
B	{[0, 0.5, 0, 0, 0, 0.6]}	{[0.23, 0, 0, 0.03, 0.1, 0.3]}	{[0.01, 0.03, 0, 0, 0.13, 0.6]}	{[0, 0.04, 0, 0, 0.06, 0.78]}
C	{[0.8, 0, 0, 0, 0, 0]}	{[0.49, 0.03, 0, 0, 0, 0.36]}	{[0.21, 0.13, 0.04, 0.07, 0.23, 0.03]}	{[0.26, 0.21, 0, 0.12, 0.07, 0.06]}
D	{[0, 0.2, 0, 0.7, 0, 0]}	{[0.42, 0.15, 0, 0.13, 0, 0.05]}	{[0.08, 0.11, 0.11, 0.38, 0.19, 0.03]}	{[0, 0.12, 0.07, 0.61, 0.18, 0.04]}
E	{[0, 0, 0.8, 0, 0, 0]}	{[0.58, 0.07, 0.06, 0.08, 0.14, 0.01]}	{[0.19, 0.19, 0.09, 0.12, 0.29, 0.02]}	{[0.06, 0.43, 0.3, 0.07, 0.37, 0]}
F	{[0, 0, 0, 0, 0, 0.8]}	{[0.15, 0, 0, 0, 0.01, 0.11]}	{[0.03, 0, 0, 0.03, 0.02, 0.58]}	{[0, 0, 0, 0, 0, 0.83]}

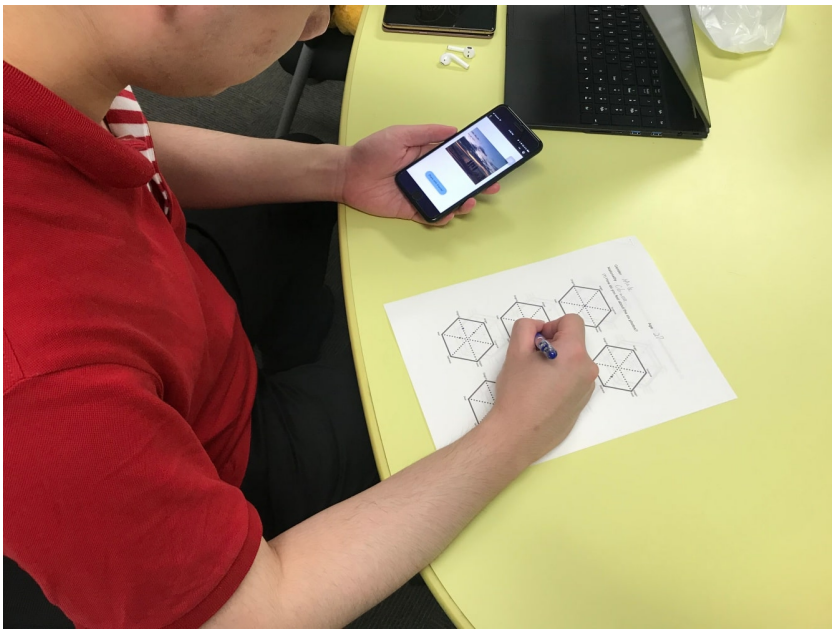


Figure 4.3 User test

impression towards the photo is basically happiness. With the visual effect of the photo in Step 3, some participants felt less happy and more surprised. There are also some people thinking that the colors make it look disgusted. But adding the audio effect of happiness, the perception of happiness was significantly increased while the negative perception (fear and disgust) were decreased.

In test of Photo B, the dominant emotion is surprise (0.5) and sadness (0.6). One interesting finding is that since the photo has some farewell message in Japanese on it, people who can read Japanese generally felt happy or sad while people cannot read Japanese felt nothing or fear. It also indicates that the content of the photo determines people's perception in the first place. When tapping the photo and making it into black and white, similar as the pre-test, the participants' perception were sharply changed into sadness, especially for people who cannot read Japanese. Lastly in Step 4, the perception of sadness is even strengthen with the audio information.

Photo C has only emotion of happiness (0.8). The first impression of participants is almost happiness and sadness. However, the perception of emotion changed into a complex of happiness, surprise, fear and a little disgust. Obvi-

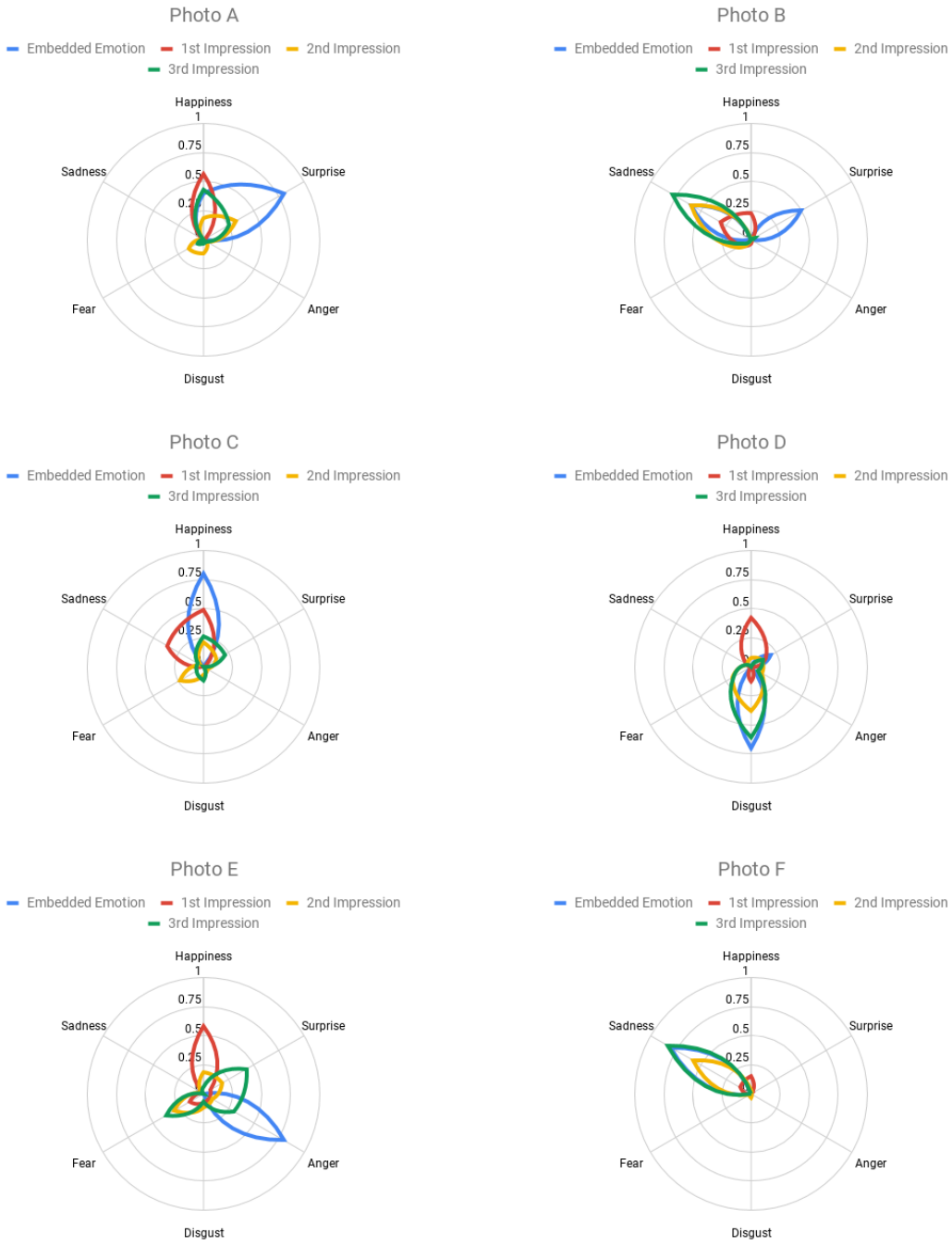


Figure 4.4 Radar chart of the average perception result

ously, the retouch effect of happiness did not look good with the pink and yellow combination. The audio of happy sound effect made fear decreased and surprise increased which modified the perception to some extent.

Photo D, a picture of a green cake, has been embedded emotion of surprise (0.2) and disgust (0.7) in its Exif data. The first impression varied from happiness to disgust, and positive perception was higher than negative ones. After tapping the photo, happiness reduced to less than 0.1, and disgust grew to 3 times higher than first impression. Next, happiness decreased to zero and disgust grew even higher with the assistance of audio effect. We can know from the radar chart of Photo D that the green circle (3rd impression) is the closest to blue one (embedded emotion).

Photo E is a picture of "Hanami" but its embedded emotion is anger (0.8). The first impression is, undoubtedly, almost happiness. Nevertheless, when adding the photo retouch effect of anger, the perception of happiness sharply decreased to one third of its first impression and other negative emotions grew. With the audio playing sound of a door slamming shut which meant to present anger, surprise and anger increased a lot while fear also became higher.

Photo F was taken in front of Hiyoshi station, which is a familiar place for most of the participants. The embedded emotion is sadness (0.8). The perception including happiness and sadness is basically mild. Then the participants tapped the photo and turned it into black and white. The perception of sadness raised to 0.58 and other perception became negligible. Later in Step 4 the audio effect was added and we can easily tell from the chart that sadness grew higher and the shape of the third impression almost overlapped with the one of embedded emotion.

Unfortunately, the vibration of smartphone which is designed to induct surprise in users did not perform as expected. Even though there are vibrations while tapping Photo A, B, and D, some participants took it as an irrelevant reminder and some did not understand its meaning. Comparing with visual and audio changes, the message of vibration is more difficult to take in.

On the whole, test of Photo B and F have received the most accurate and expected result. Fundamentally, the visual change can guide the participants to perceive sadness as the emotion information. Audio effect also played an impor-

tant role to enrich the impression of sadness. The perception did not vary too much from different culture or age. In addition, visual retouch of Photo D and E changed the tendency of participants' perception successfully. When the picture is about something that should be happy, the retouch made people less pleasant obviously and increased more negative perceptions. Although not accurate enough, here audio really helped to strengthen the negative emotion and reduce the positive. However, Photo A and C whose main emotion are happiness both failed to convey the corresponding emotion to users. Especially in Photo C where the photo had higher saturated, it even misled the participants and let them feel more fear towards the photo. But it is also helpful to see that in the two cases, the audio helped to lower the negative emotion and correct the emotion tendency.

4.2. Feedback & Analysis

The experiment was conducted smoothly and received many feedback from participants. Based on the assessment and feedback, the advantages and limitation of the design will be more clear.

4.2.1 Feedback from Participants

After the user test, the participants were asked to evaluate the interaction system and whether they would like to use this kind of service to share photos with friends.

Eight participants found it interesting to have this kind of interaction in photos. For the visual information, different color combinations definitely changed their impression upon the photos, especially the photos whose content and embedded emotions are opposite. Moreover, one participant added that when the color changed her impression but she was not sure if it was correct, the audio helped to complete the emotion of the photo which made her be more convinced of her perception. From the negative feedback, one of the participants thought that it was a bit weird when the emotion of audio did not match with the photo content. For example, the audio of happiness actually sounds a bit surprise to her. Consequently, she felt inconsistent with pictures that look normal. Also,

there is suggestion that the visual change to the photos were too radical and it may be more suitable to modify photos mildly.

When asked about their favourite interaction among visual, audio, and vibration, almost all of them chose audio information. Compared with the simple and plain photo retouch effect, audio effect is more fresh and attractive. Participants felt that audio effect actually gave significant change to the impression of photos. The time delay between the visual change and audio play which is due to the construction of the program has made the contrast even better, as well. From the feedback, there are also voice that she appreciated the black and white pattern most among all the color combinations, which is consistent with the pre-test result. Although both visual and audio information influenced how people perceived photos, audio was more powerful in altering the impression. The perception of color combinations is, inevitably, influenced by individual difference. As mentioned before, many participants did not even realize the vibration. For people who recognized it, they did not make sense of it.

As for sharing photos with friends in this way, eight participants answered "yes". The target to whom they would like to send the photo is mostly close friends. The interaction was fun and I wanted to surprise them, said one participant. Another participant added that she would prefer to send some special photos during travel for it could convey the information more than pictures themselves, while for daily usual photos it would be unnecessary. Besides, there is also opinion that he would like to customize the effect himself and then send to friends. Some people also thought that it is more effective and interesting to send with only the sound effect rather than the visual change.

Regarding any improvements or other interaction, the opinion varies. Four people consider that it is necessary to moderate the visual effect of photos, or only part of the picture rather than the whole picture changes. One thought the photo should change its color with transition to serve as a buffer for users. There is also advice that the photo could have more degrees of modification change from light, intermediate to heavy mode. More importantly, the retouch effect should look more natural. On the other hand, there are four people saying that they are expecting more animation effect on the photo, like adding some GIF pictures or some surprising pop-up image. Additionally, there is some suggestion on audio

effect. Since the audio file is played according to the embedded emotion data in each photo, participants feel less surprised after a few photos of same emotion. It will keep users feeling fresh and interesting in these interaction if there are more choices on sound effect. Besides, one participant believe it is helpful with audio of the actual surrounding content. Meanwhile, there is a problem that audio playing requires the user to put on earphones in advance.

4.2.2 Analysis

4.2.2.1 Findings

From the user test, it is really inspiring to see that the color retouch and audio effect can actually affect how people perceive a photo.

There was a worry that the effect of color modification might varies greatly on people from different countries. For the content in Photo B specially requires culture understanding, the average perception result of people from the three countries upon Photo B was calculated and made the corresponding charts (see Figure 4.5). From the charts we can tell that although the first impression of the three charts varies greatly, the perception of sadness all grew a lot with the visual effect. After adding the audio information, the perception of sadness raised even higher in all three groups.

Though the degree of change between the three groups is different, the tendency is quite similar. We can tell that the perception change affected by visual interaction does not distinguish too much from different cultural background, that is, the influence of culture difference is weaker than individual difference.

During the experiment, audio was surprisingly popular among the participants. In cases that the photo retouch does reflect the corresponding emotion, the audio could strengthen the perception. When the photo retouch does perform well, the audio could correct the perception of people. It is believed that if the audio is more associated with the content of photo, it will help users build a more complete image of the perception.

Figure 4.5 Average perception result on Photo B of people from three countries

The time delay due to the program structure won a unexpected feedback that the delay made the emotion contrast even better. Sometimes the unexpectedness can help enrich the perception and impress users. As for the opinion that audio information could be embedded with the surrounding environment sound, personally it is not always necessary. Except for voice memo specially recorded for the photo, the actual surrounding sound is usually related with the environment rather than the photographer him or herself. For example, some people found it weird that emotion embedded in Photo E was anger, whose content is about "Hanami". Photo E was taken when the photo taker spent over an hour to arrive at a hot spot and could not find any space to rest and admire the flower, feeling exhausted and angry. It is the information that usually cannot be obtained from

the picture itself. In other words, the feeling of awkwardness between the picture and the embedded emotion is exactly the reason why it is important to augment emotion information into photos and also is the meaning of the study.

4.2.2.2 Efficiency of Interactions

Among the six photos, the evaluation result of Photo B and F match with the embedded emotion best. Accordingly, gray colors among all the combinations is the most accurate and efficient color combination to arouse sadness in participants. Between the 2nd and 3rd impression, there is also a great increase of perception, which also proved the efficiency of the related audio information. In the same way from Photo D we can tell that the color combination and audio information of disgust match with the perception and help reproduce the similar emotion for users.

Surprise, as it is difficult to be expressed by colors, was designed to be presented by vibrations in this system and was a total failure. Referring to the evaluation result, the situation that participants felt surprise is generally when the visual and audio expressions have huge contrast with the content of photos, that is, the actual embedded emotion was opposite from the superficial meaning of the picture. Since surprise in actual life is not limited to positive or negative emotions, its related expression should be conducted with more detailed design.

In Photo E where anger was embedded in Exif data, although from the interaction neither visual effect nor audio information managed to deliver the emotion, the perception of happiness decreased to zero and the tendency of the perception was obviously changed into negative way. After the interaction of visual and audio, surprise and fear were significantly high. Possible reason could be that the embedded emotion is opposite to people's first impression which would lead users in feeling surprise and unexpected. Another reason may be simply that the audio itself is too frightening. The intention of the audio that a door slamming shut is to convey the image of an angry person. However, when hearing such a sound effect, it is common to feel scared as a bystander. Based on the reason below, there are still many points that remain to be improved in case of emotion "Disgust". But it could be inspiration to design emotion "Fear" whose efficiency was not tested in this experiment.

4.2.2.3 Limitations

In this test, the nationalities of the participants are China, Indonesia, and Japan, which are all located in Asia. Although there are inevitably differences between these countries, the culture of the three countries are relatively close to each other, which may lead to a close perception for the designed color combinations.

As introduced in 2.4, there are quite similar color expressions for emotions like anger and fear in European and American. Nevertheless, it is hard to assert that every color combination works the same as that for Asian users. Since there are possibilities that the design of color combinations might not be effective for users from European or American countries, it is expected that more and further validation test could be conducted for European and American users.

On the other hand, being popular among the participants, the audio effect has an essential problem that there are too limited audio files to choose from in the present prototype. The audio effect is effective in this test where almost each audio file is used once. If the number of photos is too big that the audio effect for the same emotion has to be used repeatedly, users will soon get tired of the same sound and it may not be able induce the same emotion in users as before. Of course, it is also not interesting enough for users to interact with the photos.

Apart from this, the visual effect remains to be improved further. In the current prototype, when tapping the photo, the color will be changed completely. The change may be too sudden that it might make some users feel uncomfortable. To optimized the user experience, the visual change can be more soft and gradual along with the audio playing. However, there are also worries that it might weaken the emotion information and influence how people perceive from the photos. For future improvement, more sophisticated designs are expected to be employed.

4.2.2.4 Summary

Throughout the result of the proof of concept, efficiency of the negative emotions' interactions is quite successful, especially emotion of sadness and disgust. Both visual and audio information contribute to the perception change in users.

On the contrary, the result for samples of happiness are not favorable. The biggest failure is that when transferring into photos, the color combination of pink and yellow do not look clean. Besides, the photo retouch program itself is

not perfect enough and caused the loss of image quality.

Comparing the effect of visual and audio information, it is obvious that audio effect is quite efficient to alter users' perception. It is also a proof that music or audio is able to induce certain emotion in people. The further improvement of audio effect can be expected.

Lastly, it is a pity that vibration did not work well in the test. It requires a more sophisticated design in it.

Chapter 5
Discussion

Photo sharing on SNS has become popular and common for people. Still, it is necessary to talk about the hiding emotions of the images, as photos are important materials and tools in online communication.

It was expected that with a service that can embed emotion information into photos and reproduce it, it would be easier for people to feel empathetic on other's experience. However, the way to reproduce an emotion is a huge challenge. The result of the experiment showed that the design of EmoPhoto was partially efficient in practice.

5.1. Conclusion

The concept of EmoPhoto was built on the motivation that emotion information can make people perceive similar feeling as the photo taker do and improve communication among people.

Generally, the presentation of negative emotions like sadness and disgust worked well in the experiment, which proved the feasibility of the color combinations. The combinations of black and red, standing for anger, although failed to induce anger in users, was recognized as negative emotions commonly. Failure of the experiment that happiness was not recognized by most participants proved that the photo retouch of happiness emotion still requires a further optimization. It is worthy to study whether to change a better color combination or to improve the algorithm to amend the photo effect.

The comments on audio effect are mostly good. As analyzed in the previous chapter, audio can strengthen the impression and even correct the perception in people. The issue is that there are too limited audio files in EmoPhoto. If a user

uploads photos for several times, the user will soon get used to the same audio effect. For receivers, they will also lose interest on the repeating audio. After people get accustomed to it, it will be difficult to induce the same emotion with the limited audio files.

During the experiment, some participants told me that they felt some mismatch between the photo and its related visual and audio effect. To some extent, the mismatch indicates the necessity of emotion information. Sometimes people have to view hundreds of photos a day on social network without thinking about its inner message. They look at a photo and take it for granted that the message is only about the image. There are countless applications to retouch photos to make them beautiful. But there is also needs to make photos to be understood and make the inner message interesting. For professional photographers, image is the message itself. For daily communication, the image, the story, and the feeling of the photo taker are the message to be conveyed.

5.2. Future Work

Owing to the limitation of the design, the first issue to solve is the photo retouch of "happiness". A beautiful and natural visual presentation is able to please the viewer better and induce more positive emotions in them. Of course, the whole quality of the photo retouching need to be improved. It is also expected to come up with better algorithm to process photos rapidly.

There is also need to make the visual effect change gradually rather than suddenly. For different mood intensity, more distinguishable visual effect can be added to images.

So far the interaction between viewer and the system is quite limited, and it could be another reason that some embedded information cannot be fully perceived by users. In further study, it is essential to improve users' experience and introduce more immersive interaction to the system. More importantly, it is crucial to redesign the presentation for surprise. In some ways it could be the most challenging part of future work.

Being popular in this experiment, the point that audio effect requires users to put on earphones in advance might cause inconvenience in real life utilization.

Hence it will be encouraging if there is a better interaction that do not bother users and perform as good as audio effect.

Last but not least, it is really happy to see that many participants are willing to share photos with people in this way. When asked about the relationship with the receiver, all the participants chose close and intimate friends. But how can actually the emotion embedded photos help promote communication among friends comparing with normal photos? A further investigation could be conducted to evaluate the quality of the communication when sending the special photo to friends and the feeling of both senders and receivers during the chat. It will also be valuable information to improve the design of EmoPhoto.

References

[1] Anne Oeldorf-Hirsch and S Shyam Sundar. Social and technological motivations for online photo sharing. *Journal of Broadcasting & Electronic Media*, 60(4):624–642, 2016.

[2] Chei Sian Lee, Nur Alifah Binte Abu Bakar, Raudhah Binti Muhammad Dahri, and Sei-Ching Joanna Sin. Instagram this! sharing photos on instagram. In *International Conference on Asian Digital Libraries*, pages 132–141. Springer, 2015.

[3] Jette Kofoed and Malene Charlotte Larsen. A snap of intimacy: Photo-sharing practices among young people on social media. *First Monday*, 21(11), 2016.

[4] Sicheng Zhao, Hongxun Yao, Yue Gao, Guiguang Ding, and Tat-Seng Chua. Predicting personalized image emotion perceptions in social networks. *IEEE transactions on affective computing*, 9(4):526–540, 2016.

[5] Xia Mao, Bin Chen, and Itsuya Muta. Affective property of image and fractal dimension. *Chaos, Solitons & Fractals*, 15(5):905–910, 2003.

[6] Jana Machajdik and Allan Hanbury. Affective image classification using features inspired by psychology and art theory. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 83–92. ACM, 2010.

[7] Tianrong Rao, Xiaoxu Li, and Min Xu. Learning multi-level deep representations for image emotion classification. *Neural Processing Letters*, pages 1–19, 2016.

[8] Mickael Ménard, Paul Richard, Hamza Hamdi, Bruno Daucé, and Takehiko Yamaguchi. Emotion recognition based on heart rate and skin conductance. In *PhyCS*, pages 26–32, 2015.

[9] Alicia Heraz and Manfred Clynes. Recognition of emotions conveyed by touch through force-sensitive screens: Observational study of humans and machine learning techniques. *JMIR mental health*, 5(3):e10104, 2018.

[10] Ralph B Hupka, Zbigniew Zaleski, Jurgen Otto, Lucy Reidl, and Nadia V Tarabrina. The colors of anger, envy, fear, and jealousy: A cross-cultural study. *Journal of cross-cultural psychology*, 28(2):156–171, 1997.

[11] Evi Joosten, Giel Van Lankveld, and Pieter Spronck. Colors and emotions in video games. In *11th International Conference on Intelligent Games and Simulation GAME-ON*, pages 61–65, 2010.

[12] Li-Chen Ou, M Ronnier Luo, Andrée Woodcock, and Angela Wright. A study of colour emotion and colour preference. part i: Colour emotions for single colours. *Color Research & Application*, 29(3):232–240, 2004.

[13] Li-Chen Ou, M Ronnier Luo, Andree Woodcock, and Angela Wright. A study of colour emotion and colour preference. part ii: colour emotions for two-colour combinations. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur*, 29(4):292–298, 2004.

[14] Li-Chen Ou, M Ronnier Luo, Andrée Woodcock, and Angela Wright. A study of colour emotion and colour preference. part iii: Colour preference modeling. *Color Research & Application*, 29(5):381–389, 2004.

[15] Martin Solli and Reiner Lenz. Color emotions for multi-colored images. *Color Research & Application*, 36(3):210–221, 2011.

[16] Li He, Hairong Qi, and Russell Zaretzki. Image color transfer to evoke different emotions based on color combinations. *Signal, Image and Video Processing*, 9(8):1965–1973, 2015.

[17] Shiguang Liu and Min Pei. Texture-aware emotional color transfer between images. *IEEE Access*, 6:31375–31386, 2018.

[18] Daniel Västfjäll. Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae*, 5(1_suppl):173–211, 2001.

[19] Roni Shifriss, Ehud Bodner, and Yuval Palgi. When you ' re down and troubled: Views on the regulatory power of music. *Psychology of Music*, 43(6):793–807, 2015.

[20] Hauke Egermann, Nathalie Fernando, Lorraine Chuen, and Stephen McAdams. Music induces universal emotion-related psychophysiological responses: comparing canadian listeners to congolese pygmies. *Frontiers in psychology*, 5:1341, 2015.

[21] Thomas Fritz, Sebastian Jentschke, Nathalie Gosselin, Daniela Sammler, Isabelle Peretz, Robert Turner, Angela D Friederici, and Stefan Koelsch. Universal recognition of three basic emotions in music. *Current biology*, 19(7):573–576, 2009.

[22] Patrik N Juslin and Erik Lindström. Musical expression of emotions: Modelling listeners' judgements of composed and performed features. *Music Analysis*, 29(1-3):334–364, 2010.

[23] Kate Hevner. The affective value of pitch and tempo in music. *The American Journal of Psychology*, 49(4):621–630, 1937.

[24] Laura-Lee Balkwill and William Forde Thompson. A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music perception: an interdisciplinary journal*, 17(1):43–64, 1999.

[25] E Glenn Schellenberg, Ania M Krysciak, and R Jane Campbell. Perceiving emotion in melody: Interactive effects of pitch and rhythm. *Music Perception: An Interdisciplinary Journal*, 18(2):155–171, 2000.

[26] Julia C Hailstone, Rohani Omar, Susie MD Henley, Chris Frost, Michael G Kenward, and Jason D Warren. It's not what you play, it's how you play it: Timbre affects perception of emotion in music. *The quarterly Journal of Experimental psychology*, 62(11):2141–2155, 2009.

[27] Philip Nicholas Johnson-Laird and Keith Oatley. The language of emotions: An analysis of a semantic field. *Cognition and emotion*, 3(2):81–123, 1989.

[28] Paul Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.

[29] Robert E Thayer. *The biopsychology of mood and arousal*. Oxford University Press, 1990.

[30] Albert Mehrabian. *Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies*. Oelgeschlager, Gunn & Hain Cambridge, MA, 1980.

[31] 日本カラーデザイン研究所. 心を伝える配色イメージ. 講談社, 2008. URL: https://books.google.co.jp/books?id=3DBePgAACAAJ.

[32] Camera & Imaging Products Association et al. Exchangeable image file format for digital still cameras: Exif version 2.3. *CIPA DC-008 Translation-2010*, 2010.

[33] PM Nishad. Various colour spaces and colour space conversion. *Journal of Global Research in Computer Science*, 4(1):44–48, 2013.

[34] Finnolia Productions Inc. Dream harp. URL: https://www.soundeffectsplus.com/product/dream-harp-06/.

[35] Cellar door slam shut. URL: http://soundbible.com/297-Cellar-Door-Slam-Shut.html.

[36] Mike Koenig. Squish. URL: http://soundbible.com/511-Squish-1.html.

[37] Daniel Simion. Creepy background. URL: http://soundbible.com/2165-Creepy-Background.html.

[38] Bruno ph. Sadness cosmo. URL: https://freesound.org/people/Bruno_ph/sounds/109409/.

[39] Blending modes in adobe photoshop. URL: https://helpx.adobe.com/photoshop/using/blending-modes.html.

Appendices

A. Retouching function to replace color of each pixel

Python function to replace each pixel's color hue attribute in the photo with the hue-closest color from the emotion model

```
def retouch(img, num):
    i = 1
    j = 1
    width = img.size[0]
    height = img.size[1]
    for i in range(0, width):
            for j in range(0, height):
                    (r, g, b) = img.getpixel((i, j))
                    (h, s, v) = colorsys.rgb_to_hsv(r/255,
                       g/255, b/255)
                    minDis = 180
                    for m in range(0, 3):
                            hueDis = min(abs(h * 360 - H[m
                               ]), 360 - abs(h * 360 - H[m
                               ]))
                            if hueDis < minDis:
                                    minDis = hueDis
                                    hueNum = m
                    h = H[hueNum] / 360
                    if h == 0:
                            s = 0
                    (rr, gg, bb) = colorsys.hsv_to_rgb(h, s
                       , v)
```

```
img.putpixel((i, j), (int(rr * 255),
    int(gg * 255), int(bb * 255)))
```

B. Grading result of pre-test

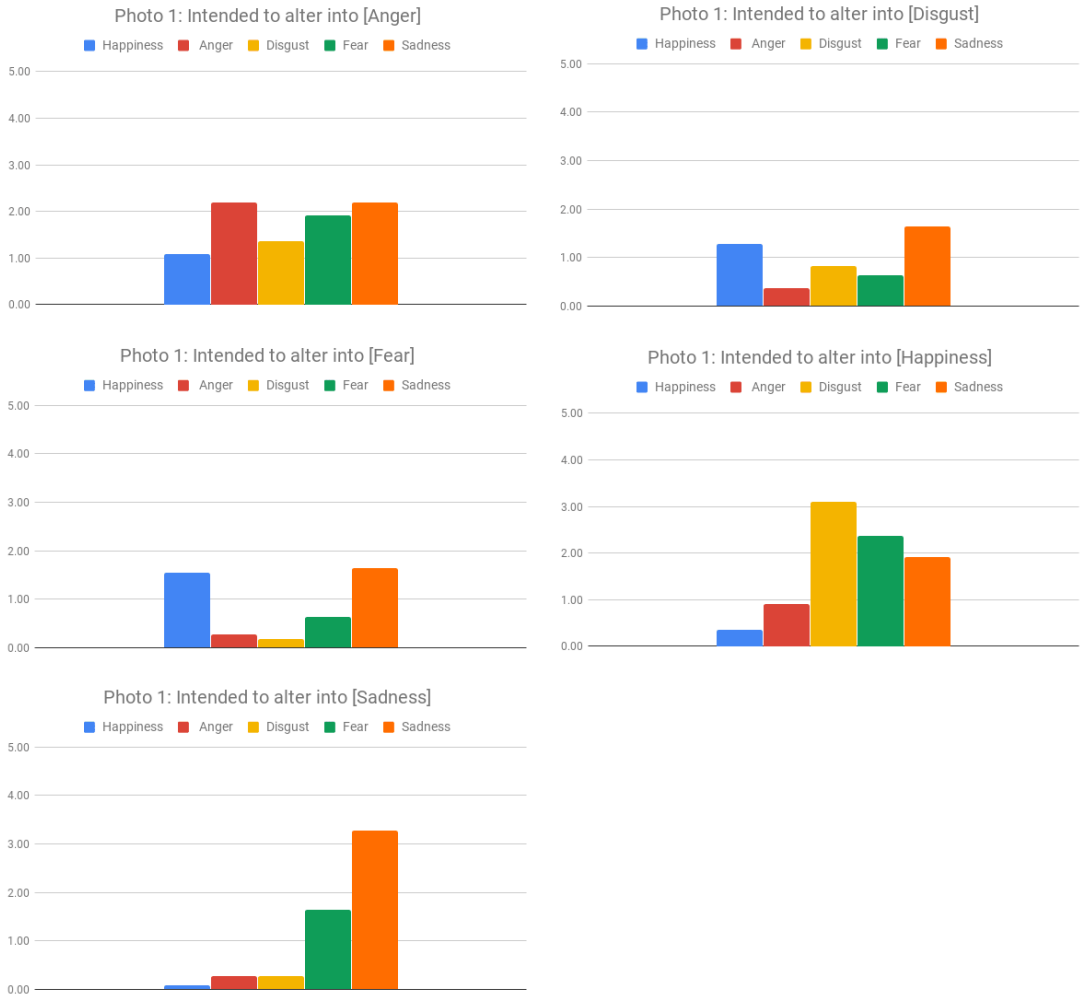


Figure B.1 Average grading result of Photo 1

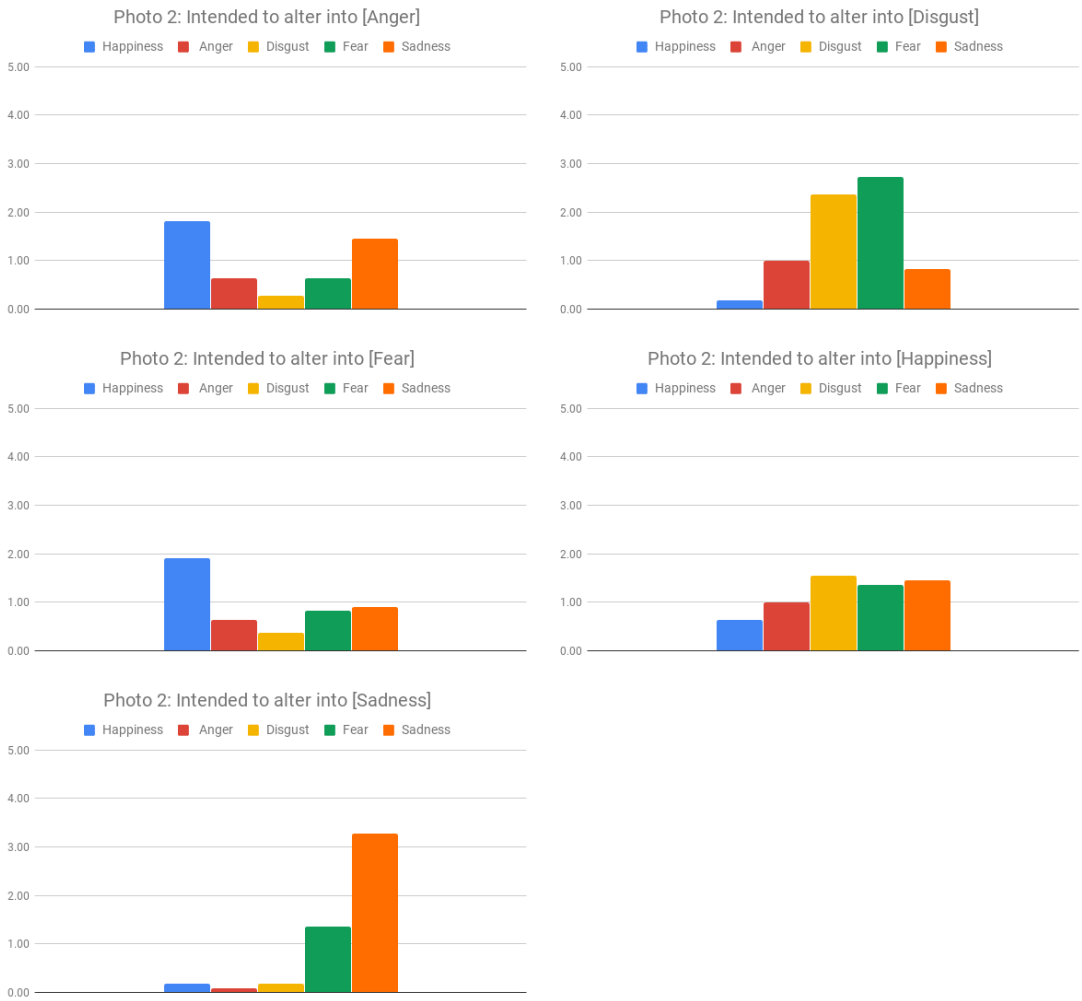


Figure B.2 Average grading result of Photo 2

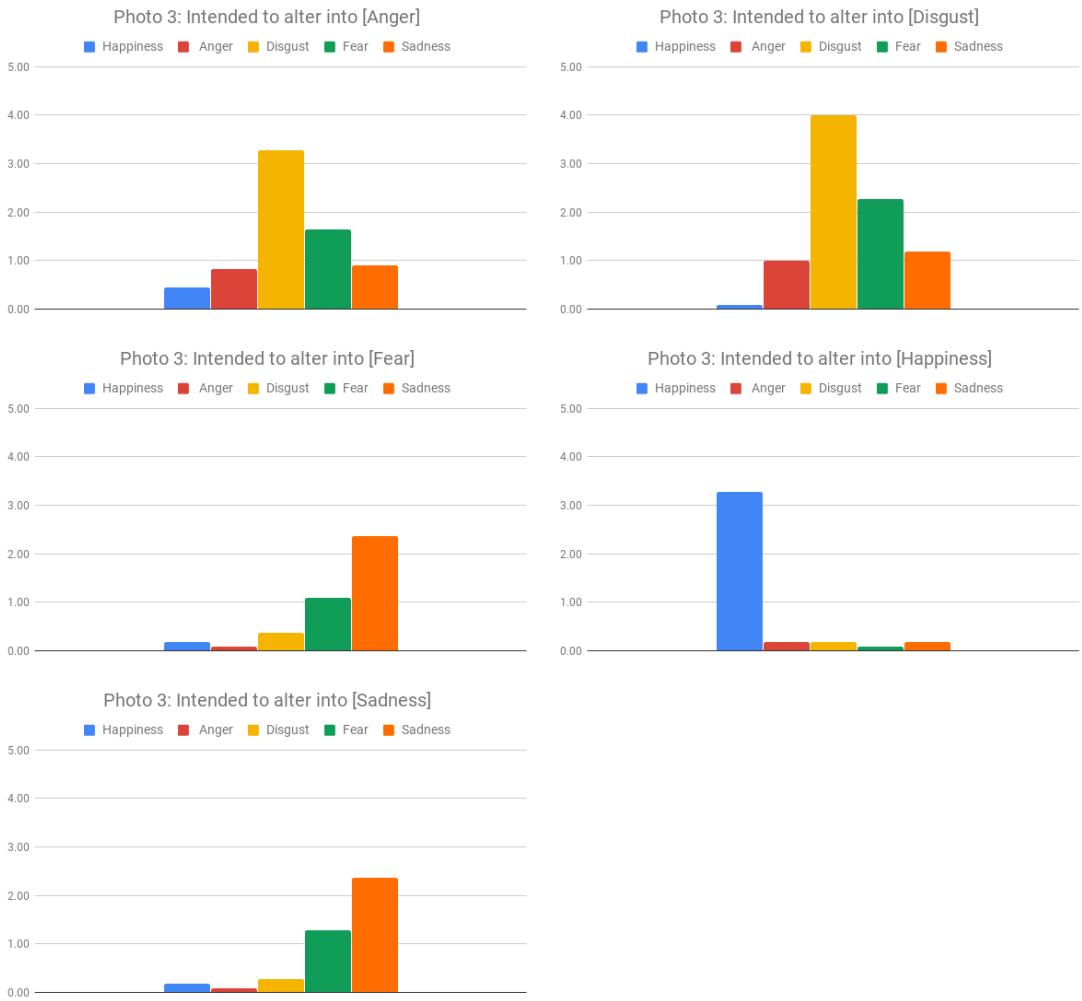


Figure B.3 Average grading result of Photo 3

C. Optimized retouching function

Optimized Python function to modify color of each pixel using overlay blend mode

```
def retouch(img, num):
    i = 1
    j = 1
    width = img.size[0]
    height = img.size[1]
    if int(num) < 6:
            for i in range(0, width):
                    for j in range(0, height):
                            (r, g, b) = img.getpixel((i, j)
                                )
                            (h, s, v) = colorsys.rgb_to_hsv
                                (r/255, g/255, b/255)
                            minDis = 180
                            for m in range(0, 3):
                                    hueDis = min(abs(h *
                                        360 - H[m]), 360 -
                                        abs(h * 360 - H[m]))
                                    if hueDis < minDis:
                                            minDis = hueDis
                                            hueNum = m
                            ovl = colorsys.hsv_to_rgb(H[
                                hueNum] / 360, S[hueNum] /
                                100, V[hueNum] / 100)
                            src = (r / 255, g / 255, b /
                                255)
                            output = overlay(np.array(ovl),
                                 np.array(src))
                            (rr, gg, bb) = (output[0],
                                output[1], output[2])
                            img.putpixel((i, j), (int(rr *
                                255), int(gg * 255), int(bb
                                * 255)))
```

```
        elif int(num) == 6:
                for i in range(0, width):
                        for j in range(0, height):
                                (r, g, b) = img.getpixel((i, j)
                                    )
                                (h, s, v) = colorsys.rgb_to_hsv
                                    (r/255, g/255, b/255)
                                h = 0
                                s = 0
                                (rr, gg, bb) = colorsys.
                                    hsv_to_rgb(h, s, v)
                                img.putpixel((i, j), (int(rr *
                                    255), int(gg * 255), int(bb
                                    * 255)))

def overlay(img_1, img_2):
    mask = img_2 < 0.5
    output = 2 * img_1 * img_2 * mask + (1-mask) * (1- 2 * (1-
        img_1)*(1-img_2))
    return output
```

D. The assessment sheet for evaluation test

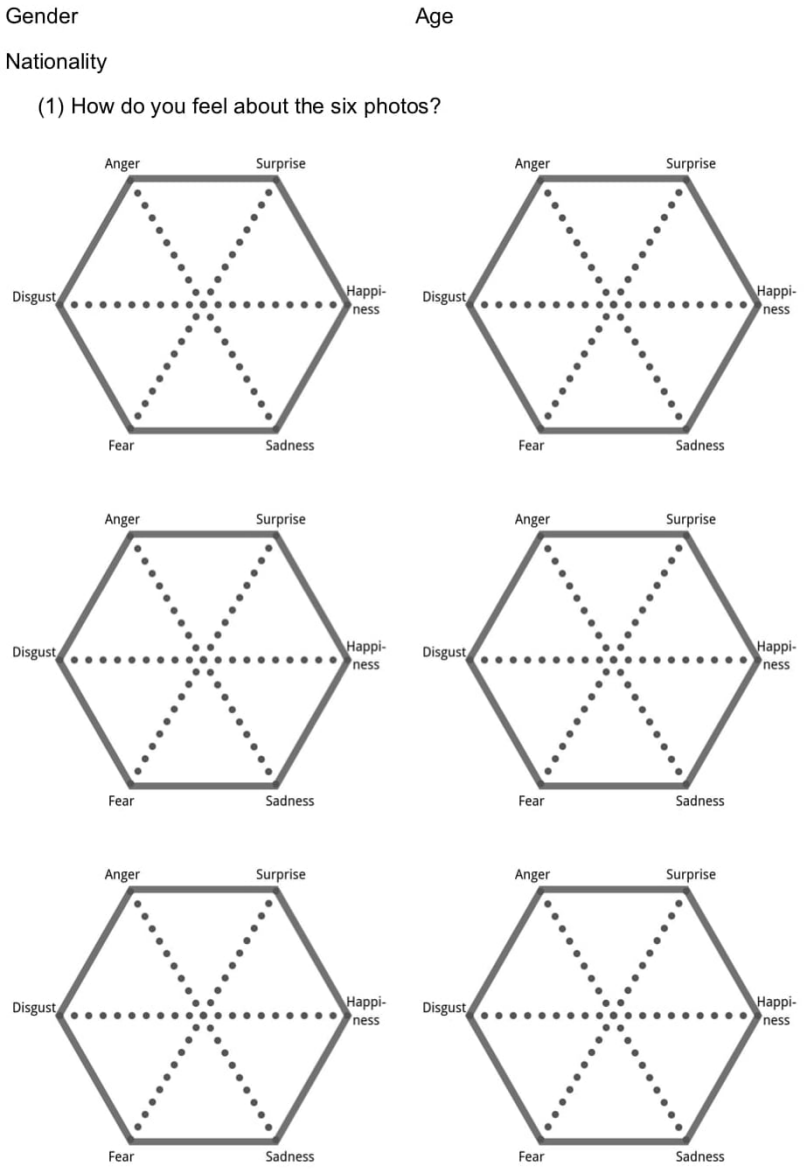


Figure D.1 Assessment sheet for participants to measure their perception during interaction

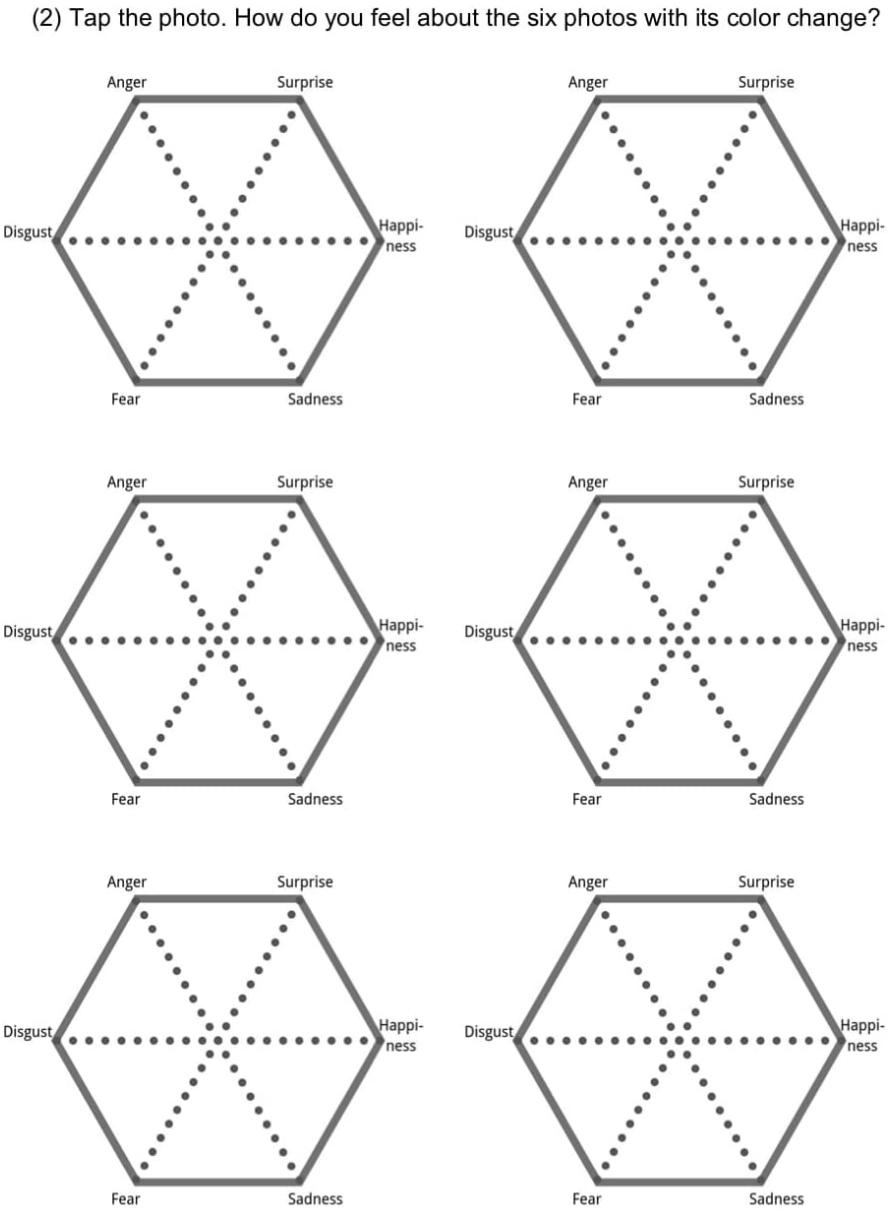


Figure D.2 Assessment sheet for participants to measure their perception during interaction

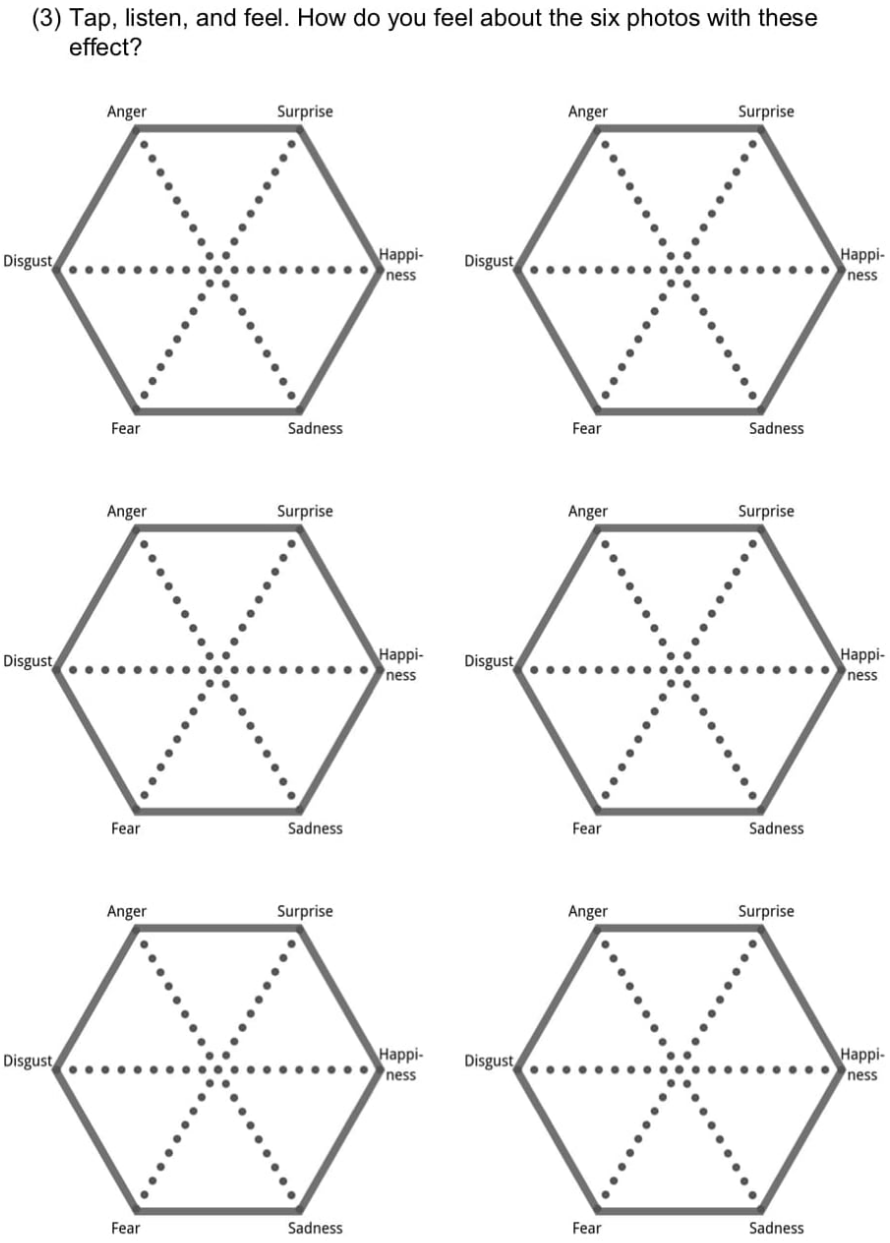


Figure D.3 Assessment sheet for participants to measure their perception during interaction

(4) How do you feel about these kind of interaction? Or did the effects change your impression on the photo?

(5) Among visual, audio, vibration information, which one do you appreciate most?

(6) Would you like to share photos with friends in this way?

(7) What other interaction do you like to add to photo?

Figure D.4 Assessment sheet for participants to measure their perception during interaction