Title	AffectiveWear : Toward Measuring Facial Expression in Daily Life
Sub Title	
Author	正井, 克俊(Masai, Katsutoshi) 稲見, 昌彦(Inami, Masahiko)
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2015
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2015年度メディアデザイン学 第417号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002015-0417

Master's Thesis

Academic Year 2015

AffectiveWear: Toward Measuring Facial Expression in Daily Life

Graduate School of Media Design,
Keio University

Katsutoshi Masai

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

Katsutoshi Masai

Thesis Committee:

Professor Masahiko Inami	(Supervisor)
Associate Professor Kai Kunze	(Co-supervisor)
Professor Hiroyuki Kishi	(Co-supervisor)

Abstract of Master's Thesis of Academic Year 2015

AffectiveWear: Toward Measuring Facial Expression in Daily Life

Category: Science / Engineering

Summary

Facial expressions can tell us a lot, from subtle non verbal communication to hints about a person 's emotional states. This thesis investigates how to detect facial expressions in realistic scenarios. To this end I developed a research prototype of eye glasses. The eyewear prototype can recognize wearer 's 7 facial expressions (Neutral, Smile, Laugh, Disgust, Angry, Surprise and Sad). The device is fully integrated, and its appearance is close to normal eyewear. It can recognize and record facial expressions in real life over the day. The photo reflective sensors integrated into our device are small, affordable and accurate enough to capture a subtle facial movement. The photo reflective sensors measure the distances between multiple spots of the eyewear frame and the skin surface of a person 's face. After normalization, the inputs are applied to Support Vector Machine algorithm in order to recognize facial expression. The latest prototype includes 17 sensors to evaluate optimal placement. With 7 Japanese and 1 French, the accuracy is 98% if the data is recorded in one session. For one user, with 5 recording on different days, the accuracy is 89%.

Keywords:

Facial Expression, Wearable Computing, Eyewear Computing, Sensing Technique, Communication, Emotion

Graduate School of Media Design, Keio University

Katsutoshi Masai

Acknowledgements

First and foremost I offer my sincerest gratitude to my supervisor, Prof. Masahiko Inami, who has always given me an interesting feedback with his unique insights and knowledge whilst allowing me the room to work in my own way.
Next, I would like to appreciate Prof. Kai Kunze, who supported me throughout my thesis with his patience and kindness. The time we discuss was alway meaningful and insightful.
I would like to appreciate Prof. Hiroyuki Kishi for his guidance about my thesis.
I would like to thank faculty members and staff of Reality Media. Without their help, I couldn't finish my research project.
I would like to thank members of Reality Media. We spent long time with doing research and having fun.
I would like to thank collaborative research partners. They gave me good insight to deepen my research.
Finally, I would like to thank my family. They supported and encouraged me with best wishes. Without their help, I couldn't have had a precious research experience and couldn't study abroad.

Table of Contents

List of Figures

List of Tables

1. Introduction

Do you know how often and how long you smiled yesterday? Have you ever been aware of your facial expressions while you are talking to others?

Our facial expression changes spontaneously over a course of day, and we are not usually aware of our own facial expression in daily life. Facial expressions change unconsciously depending on our mental and physical condition. Also Facial expressions can change when we interact with the environment, they give us the insight into how people think and feel. I think that the facial expressions are on the boundary between our inner self, body and the environment. Therefore, we can better ourselves understand our mind by tracking our facial expressions in daily life as Andy Clark asserts that our mind comes from the interaction among our brain, body and the environment [2].

Our facial expression reflects on our emotional state or our mental health. For instance, if people have a sad face, they may be tired physically or mentally. Also, our emotional state reflects on our facial expression. According to Niedenthal, individuals who were led to smile evaluated the cartoons as funnier than did participants whose smiles were blocked [17]. From the experiments, it can be said that people who smile many times over a course of a day enjoy a happy life. Therefore, by tracking our facial expression for a long time, it can be a valuable data to understand our mental condition and our facial expression habits.

When we communicate with people, we may change what and how we say depending on their facial expression because facial expression can tell us a clue of non-verbal communication. For example, when we talk to somebody and find

them have a angry face, we may try to change the subject to release their tension. Another example is that if people show us a smile when we talk to them, we think they accept or agree with us. Then, we can be relaxed and talk more naturally. Besides that, laughter is said to be contagious to others like yawning [19]. Thus, facial expression plays an important role during social interaction, conveying rich nonverbal information to others. By being aware of our own facial expression, we can improve communication skills which are important to have better social interaction with people.

I would like to achieve the system that can recognize facial expression in daily life so as to improve the user's long term behaviour and their productivity.

Wearable device can track our daily activities such as how many steps we take, how long we work out, and how well we sleep. Recently, wearable devices are getting popular and accessible to many people because IC chips are getting cheaper, smaller and faster. U.K.-based Juniper Research forcasts that the size of wearable market will jump from US$1.4 billion in 2013 to US$19 billion in 2018 [26].

With the wearable device designed properly, we can gain the information only when we need and use them naturally as if it is a part of ourselves as Andy Clark puts it " a new transparent technologies increasingly blur the already fuzzy boundary between the user and her tools for thought " [3]. There are many commercial products that are designed naturally to give information to help our daily decisions. Jawbone UP2 (Figure 1.1) can record users' daily steps and sleep quality, then based on the information, the device can suggest when they should go to bed [11]. Apple Watch (Figure 1.2) can record how often users have stood up to take a break from sitting and can notify when they should stand up [1]. They are small and light so that it is comfortable for users to bring for a day. As most wearable devices allow users to track our daily activities so as to manage and improve them, this approach is appropriate for facial expression recognition in daily life.

There are various types of wearable devices: wristband type, watch type, head-

Figure 1.1: *Jawbone UP2*

Figure 1.2: *Apple Watch*

band type, eyewear type and so on. I would like to focus on eyewear type because of 2 reasons. First, in order to capture facial expression, a device has to be worn on user's head. Second, eyewear is an accessories accepted in public and it can be put on over a course of a day.

Previously, as one of the most famous wearable device Google Glass(Figure 1.3), eyewear devices didn't look like natural eyewear. Google Glass has a mounted display and camera on a glasses. It was novel because it allows users to get and share information with hands free. However, the project of Google Glass in the status quo is failure because the embedded camera can cause privacy issues and the design is not socially acceptable to wear in daily life. On the other hand, JINS MEME is a good example. It is invented by Japanese eyewear company JIN [12]. The device aims at understanding inner self. It has an electrooculography (EOG) sensor that can detect 8 direction of eye movement and a blink and accelerometer which can detect head motion. With these information, JINS MEME can notify how tired or how sleepy users are and how concentrate users are. The design of

eyewear is natural and it can be socially acceptable. In this thesis, I explore smart eyewear that can recognize facial expression in daily life will be explored.

Figure 1.3: *Google Glass*

1.1. Contributions

In this thesis, AffectiveWear is presented(Figure 1.4). It is the eyewear prototype that can recognize the wearer's facial expressions in daily life. Facial expression recognition is done by a number of researchers. Most dominant method for facial expression recognition is camera-based approach in computer vision field. However, with this method, facial expression recognition is feasible only at the certain places and it is difficult to capture users continuously. In this thesis, the different approach is taken to solve these problems by integrating photo reflective sensors into an eyewear. The biggest contribution is making the eyewear prototype that can recognize a wearer's 7 facial expressions. The eyewear can capture facial expression regardless of face direction. Yet, there are also many advantages over camera-based systems.

- Low cost sensing: 17 photo reflective sensors cost around 5 dollars in total.
- Simplicity: The recognition system is simpler than the usual process of camera-based systems and don't need a high performance processor.

- Privacy: Since the eyewear is a personalized item, user can use the information for user's sake

In addition, facial expression recognition in daily life can give the rich information provided by nonverbal communication to computers. As an interface of facial expression, our device can provide nonverbal information not only to users but also to computing systems. If those systems can use the information about user's facial expression, they make a better performance by giving users appropriate feedback.

Figure 1.4: *User wears AffectiveWear*

1.2. Outline

In the following I describe the outline of this thesis. The next chapter gives the reader the background about facial expression recognition and related research, to get a better understanding of the field and give a basis for understanding this thesis work. Followed by a Chapter about Affectivewear where I describe the concept and the iterative process of prototype design. Subsequently, Implementation chapter provides the reader the system overview and hardware and software

implementation of the latest prototype. Evaluation chapter conifirms the accuracy of facial expression recognition. Last Chapter discusses the limitation, the possible application and future direction of the thesis.

2. Related Works

To value the contributions of this work, we explore related research in this chapter. First, we take a look into computer vision technologies enabling facial expression recognition. This is the largest area with a complementary sensing modality. Afterwards, we will look at the use of photorefelctive sensors, the sensor used by affective wear, for recognizing skin deformation. Finally, we discuss the relation between emotion and facial expressions.

2.1. Facial Expression Recognition with Camera-based System

There are a number of works related to automatic facial expression recognition in computer vision. According to Tien et al. the general approach to automatic facial expression analysis (AFEA) consists of three steps: face acquisition,facial data extraction and representation, and facial expression recognition. [22]. One of AFEA approach proved the high accuracy 91.5% of Recognizing 6 basic expressions [15] using Cohn-Kanade Database [13]. However, there are three problems with camera-based system for usability in daily life. First, the biggest problem is tracking. Most works use a simple, passive camera installed in the environment. With this camera setting, it is hard to track users for a long time. Also, it is difficult to track their facial expression when users move constantly or there is an occlusion between a camera and users. Second, a camera system mostly uses a clear image of face from frontal view. A facial image acquired from non-frontal

view makes difficult to recognize facial expression. In real environments, a frontal-view face is not always available. Next, a camera-based system does not fit to wearables: camera needs a certain distance from user to acquire facial image. If user wear wearable camera for facial expression recognition, the design of the device can be weird and would not be acceptable to society(Figure 2.1). To sum up, a camera-based system is not suitable to analyze facial expression continuously in daily life.

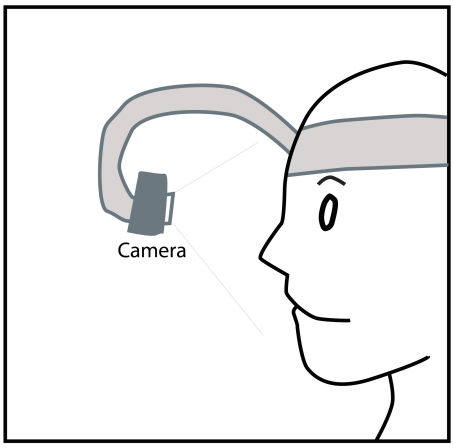

Figure 2.1: *Possible Design of Wearable Camera for Wearer's Facial Expression Recognition*

2.2. Photo Reflective Sensors and Human Skin Movement Sensing

A photo reflective sensor is a proximity sensor. It consists of infrared LED and photo transistor. It can measure a distance between a sensor and an object by quantifying the amount of Infrared (IR) LED that sensors emit and the object reflect. In general, a photo reflective sensor is used to measure close distance. On the other hand, the range of its application is wide because photo reflective sensor can capture the subtle change of distance and its size is small (i.e. SG-105: 2.7 mm x 3.2 mm x 1.4 mm). For example, Ogata et al. leverages skin

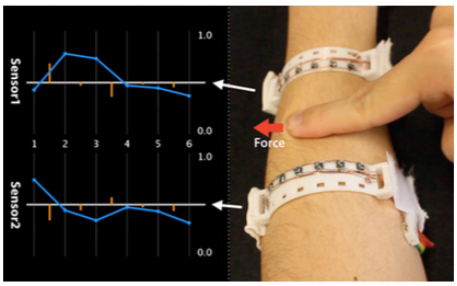

Figure 2.2: *SenSkin*

Figure 2.3: *Me-Me switch*

deformation to use skin as an interface(SenSkin) [18]. This work uses two arrays of 6 photo reflective sensors to detect the gestures (pinch, touch, pull-up, pull-down, pull-right, pull-left) on skin surface(Figure 2.2). Taniguchi et al. made Me-Me Switch that is a earphone based interface and in which a photo reflective sensor is integrated. It can detect the subtle change of an ear caused by eye movements and tongue movements [21]. Concretely, it can recognize 6 eye movements and 2 tongue movement.

Fukumoto et al. applied 2 photo reflective sensors to a pair of glasses [7]. It can detect smile, laughter and other states by measuring the movement of eye tail and cheek. The aim of this work is indexing interesting/enjoyable events on a recorded video. Their mechanism achieved the accuracy 74% / 94% for detecting smile/laughter respectively.

Nakamura et al. proposed a device for controlling the volume of augmented reality information by the glabella effect [16]. One photo reflective sensor is used to detect the movement of eyebrow.(Figure: 2.5) The purpose is intuitive and seamless control of the information using natural movement of eyebrow when users try to focus and stare at one part.

Those works measure the movement of certain area. In my thesis, we use a number of sensors to recognize facial expression which comes from the movements

Figure 2.4: *Smilie/Laughter Recognition System*

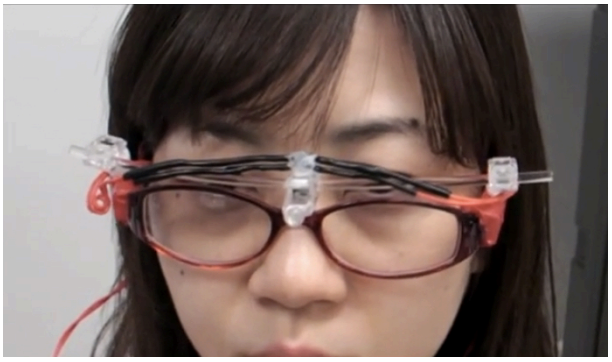

Figure 2.5: *Glabellar Fader*

of various muscles related to eyelid, eyebrow, nose, cheek and mouth. From the previous work, I conclude the size of a photo reflective sensor and its characteristics is suitable for facial expression recognition in daily life.

2.3. Wearable on Face

First approach trying to detect facial expression with a wearable device is Expression Glass [20].(Figure 2.6) It can recognize specific expression (confusion/interest) with better accuracy than random guess by measuring the facial muscle movement using piezoelectric sensors. It is wired, but showed the potential of the eyewear computing. Gruebler et al. designed the wearable device which can read positive facial expressions using facial EMG signals [8]. As seen in Figure 2.7, it has to be attached to the side of face, but it can record the user's affective state for more than 4 hours with high accuracy. It can be used during therapeutic interventions and to support medical professionals.
Fukumoto et al. used only 2 photo reflective sensors considering the usability of daily life. Ishiguro et al. proposed Aided eyes that can enhance human memory [9]. By measuring the eye movement while recording a video, the user can

Figure 2.6: *Expression Glass*

Figure 2.7: *Wearable Device to Capture Positive Expression via EMG signal*

know what they looked at and what they were interested in when they look back over the video. The system includes small phototransistors, infrared LEDs, and a video camera. It can be used in daily life because the entire system is attached to the glasses. It showed the importance of mobility for wearable sensing. Kimura et al. proposed the eyeglass-based hands-free videophone [14]. It can yield one frontal face image with fish-eye cameras and capture facial expressions. They simulated that 4 fish-eye cameras with 250-degree field of view can cover 83 % of the frontal face. The system also can be used for visual life-logging. But, the system is wired, and the size and weight was not comfortable for practical use(Figure 2.9). Considering the wearing comfort of the device, JINS MEME which is a commercial smart eyewear is well designed [12]. The design, size and weight of the smart eyewear is almost same as a normal eyewear(Figure 2.10). Based on these works, AffectiveWear system needs to be designed.

2.4. Facial Expressions and Emotion

Darwin firstly suggested the facial feedback hypothesis [4]. This theory explains that facial expressions and emotions are interrelated each other. It is well explained by James Williams's quote“ We don't laugh because we're happy, we're happy because we laugh. [10] ” The theory is applied to many works. Tsujita et

Figure 2.8: *Aided Eyes*

Figure 2.9: *Eyeglass-based Hands-free Videophone*

Figure 2.10: *JINS MEME*

al. proposed the system to make people smile more [24]. They made the fridge that user can't open without smiling. They also evaluated the effect with user study. Yoshida et al. proposed the emotion evoking system [27]. The mirror system changes the user's facial expression artificially so that the artificial facial expression influence the actual one.

As an application of our device, there is a great potential to investigate this theory in order to improve user's mental health.

3. AffectiveWear

AffectiveWear(Figure 3.1) aims at recognizing facial expression in daily life. AffectiveWear can give users the log of their facial expressions in daily life so that they can realize their cognitive or emotional state. It can also show real-time prediction so that user can enjoy subtler way of nonverbal communication in digital world. Hence, AffectiveWear needs to capture basic facial expressions. On top of that, AffectiveWear must be possible to be used in daily life. In order to make it socially acceptable, the design of AffectiveWear needs to look like eyewear.

Figure 3.1: *AffectiveWear*

3.1. Photo Reflective Sensors

For AffectiveWear system, photo reflective sensors are used to recognize facial expression. Photo reflective sensors have 3 good points. First, photo reflective sensors are small enough to be integrated into an eyewear frame. It enables AffectiveWear look like everyday glasses. Second, photo reflective sensors are not invasive and not required to touch on Face. Third, photo reflective sensors can recognize basic facial expressions based on Facial Action Coding System and principle that are described in this section.

3.1.1 Facial Action Coding System

Facial Action Coding system(FACS) is a system to measure human facial expression with action units [6]. FACS defines a facial expression that human can make. For example, characteristics of angry facial expression is 1)eyebrows down and together, 2) eyes grale, and 3) narrowing of lips. Action units represent muscular activity that produces facial appearance changes. Action units include the movement of head, eyelid, eyebrow, nose, cheek and mouth. Among those of the action units, the movements of eyelid, eyebrow, nose and cheek cause the 3 dimensional skin deformation around eyes(Figure 3.2). The movement of a mouth also somewhat changes the skin shape around eyes because the mouth movement influences a form of cheek. Therefore, sensors are put uniformly to a front frame of our eyewear device. In short, the sensors integrated in the eyewear can detect most of the movements that are crucial for facial changes related to facial expressions that should be recognized by AffectiveWear recognize [23].

3.1.2 Principle

AffectiveWear leverage skin deformation caused by the movement of facial muscles. When users move their facial muscles, 3 dimensional movement of skin oc-

TABLE 1
Upper Face Action Units and Some Combinations

NEUTRAL	AU 1	AU 2	AU 4	AU 5
Eyes, brow, and cheek are relaxed.	Inner portion of the brows is raised.	Outer portion of the brows is raised.	Brows lowered and drawn together	Upper eyelids are raised.
AU 6	AU 7	AU 1+2	AU 1+4	AU 4+5
Cheeks are raised.	Lower eyelids are raised.	Inner and outer portions of the brows are raised.	Medial portion of the brows is raised and pulled together.	Brows lowered and drawn together and upper eyelids are raised.
AU 1+2+4	AU 1+2+5	AU 1+6	AU 6+7	AU 1+2+5+6+7
Brows are pulled together and upward.	Brows and upper eyelids are raised.	Inner portion of brows and cheeks are raised.	Lower eyelids cheeks are raised.	Brows, eyelids, and cheeks are raised.

TABLE 2
Lower Face Action Units and Some Combinations

NEUTRAL	AU 9	AU 10	AU 12	AU 20
Lips relaxed and closed.	The infraorbital triangle and center of the upper lip are pulled upwards. Nasal root wrinkling is present.	The infraorbital triangle is pushed upwards. Upper lip is raised. Causes angular bend in shape of upper lip. Nasal root wrinkle is absent.	Lip corners are pulled obliquely.	The lips and the lower portion of the nasolabial furrow are pulled pulled back laterally. The mouth is elongated.
AU15	AU 17	AU 25	AU 26	AU 27
The corners of the lips are pulled down.	The chin boss is pushed upwards.	Lips are relaxed and parted.	Lips are relaxed and parted; mandible is lowered.	Mouth stretched open and the mandible pulled downwards.
AU 23+24	AU 9+17	AU9+25	AU9+17+23+24	AU10+17
Lips tightened, narrowed, and pressed together.				
AU 10+25	AU 10+15+17	AU 12+25	AU12+26	AU 15+17
AU 17+23+24	AU 20+25			

Figure 3.2: *Example of Action Units*

curs. Various facial muscles move when users change their facial expressions. Therefore, the distance between a spot of the eyewear and skin surface on face changes with different facial expression. This distance is measured with photo reflective sensors. As seen in Chapter 2.2, the approach to sense skin movement with photo reflective sensors is taken in various works. data is collected at various points of an eyewear front frame. The data collected from many sensors is applied to machine learning algorithm to recognize facial expressions.

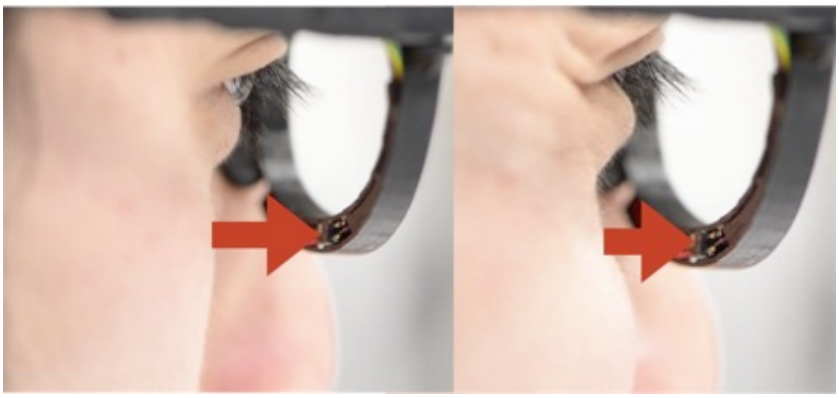

Figure 3.3: *Principle*

3.2. Iterative Design

In order to design AffectiveWear which aims at being used in daily life, 3 conditions should be fulfilled: high accuracy of facial expression recognition, robustness in various situations, and usable design in daily life.

The high accuracy is fundamental for using this information for our applications. I think at least 90% accuray is recquired. Evaluation of the accuracy was performed using iPython notebook.

Facial expression recognition should be done at various places we live daily. The most dominant place for us to live is indoor condition such as office, school and living room. These places have different lighting condition. lighting condition also

changes depending on night or day. Active lighting was applied to consider these for practicality.
The appearance of the device must fit in with daily usage. First, the frame of AffectiveWear should be as thin as possible. Next, sensors should be integrated naturally into the frame. In addition, the device should be worn comfortably. Usual smart glass is not well designed in this aspect. However JINS MEME is designed well enough for daily use. The design of AffectiveWear should look natural as JINS MEME. The process of making was applied so that the prototype can achieve natural appearance as an eyewear.

3.2.1 The 1st Prototype

The first prototype (Figure 3.4) focused on proof of principle described in chapter 3, facial expression recognition with photo reflective sensors. I attached 6 photo reflective sensors (SG-105 [Kodenshi Corp]) to a commercial eyewear frame. The photo reflective sensors are connected to Arduino Uno (A basic model of Arduino that is a rapid prototyping tool). 330 Ohm resistors were chosen for the LED of photo reflective sensors. 33 K Ohm resistors were chosen for the transistor to transmit the sensor output to an A/D converter. The resistance value was chosen so that it can measure the distance between eyewear and face appropriately. Since a photo reflective sensor is a proximity sensor, if resistor value of LED was large, long distance (10 mm or longer) is difficult to measure. The value was adjusted by changing the value of a variable resistor while the author worn the eyewear and checked the raw data. The applied voltage was 5V. It proved the possibility of facial expression recognition, but the design aspect was ignored.

3.2.2 The 2nd Prototype

From 1st prototype to 2nd prototype, the design of eyewear has changed by iteration. The Figure 3.5 shows the examples of during iterative design. By

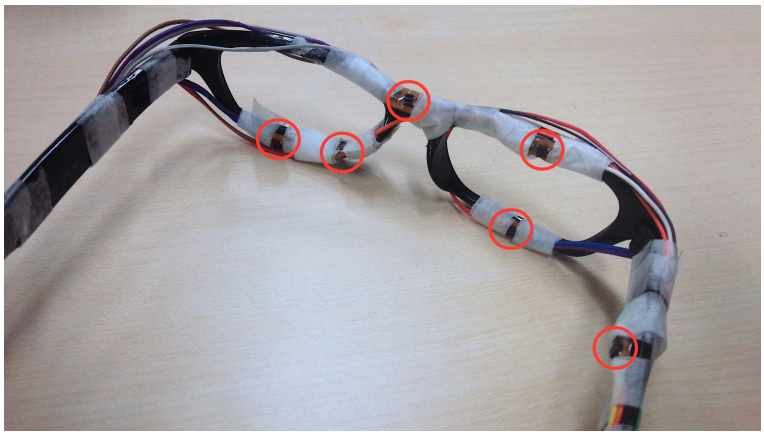

Figure 3.4: *The 1st Prototype*

iterating a process of making, the gap between physical thing and thing in the head was decreased.

Figure 3.5: *Examples during Iterative Design*

The 2nd prototype (Figure 3.6) with 8 photo reflective sensors was developed for the proof of concept, facial expression recognition and usable design in daily life. The prototype consists of Arduino Fio, Xbee (wireless module), Li-po Battery and 8 photo reflective sensors. 8 sensors were used because Arduino Fio has 8 analog inputs. The applied voltage is 3.3V because of the characteristic of Arduino Fio. The resistor value was changed accordingly considering a) the distance between sensors and face b) the battery consumption: A 180 - 240 Ohm resistor for the LED of photo reflective sensor and a 62 K Ohm resistor for the transistor to transmit the sensor output to an A/D converter.

In this time, I created the front frame of the eyewear and the outer box for

Arduino Fio and Li-po battery with 3D CAD modeling software(Rhinoceros 5) and 3D printer (Uprint). The model can be seen in Figure (3.7The 3D printer was used in order to realize the natural appearance as an eyewear. 3D model can make the design of the front frame fit in with the size of photo reflective sensors so that the sensors can be integrated into the frame naturally. The temple is reused from commercial eyewear frame. The elastic nose pad of anti-pollen glasses was used from J!NS because this is made of silicon and adjustable to various shape of nose. I applied eyewear band to make the position of the eyewear stable. 8 photo reflective sensors are integrated into the eyewear. It was made wireless for its mobility. The eyewear communicates with the laptop via zigbee.

Figure 3.6: *The 2nd Prototype*

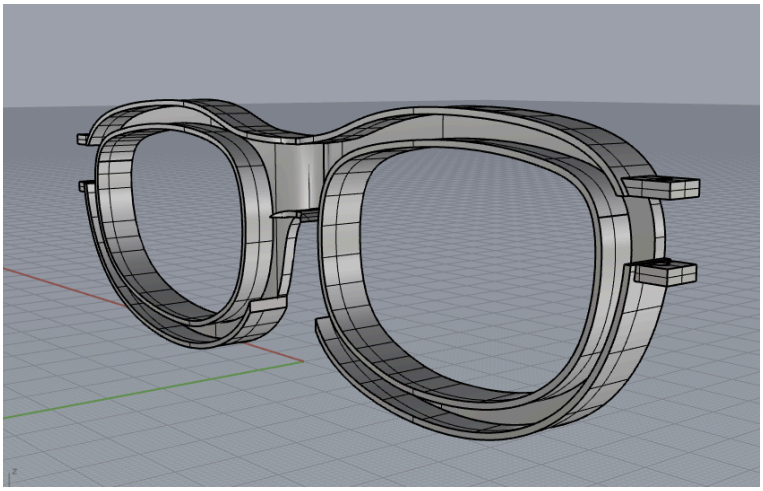

Figure 3.7: *Computer generated 3D model for AffectiveWear*

4. Implementation

4.1. System Overview

The Figure 4.1 shows the overview of the system. AffectiveWear system works with 4 steps. First, sensors that are integrated into an eyewear frame capture facial gestures. Second, those data are processed in Arduino. Then, processed data are sent to Processing. Finally, based on those data, facial expression is predicted and visualized in Processing,. The detail is described from two perspectives, hardware and software. It can detect the wearer's 7 facial expressions (neutral, disgust,

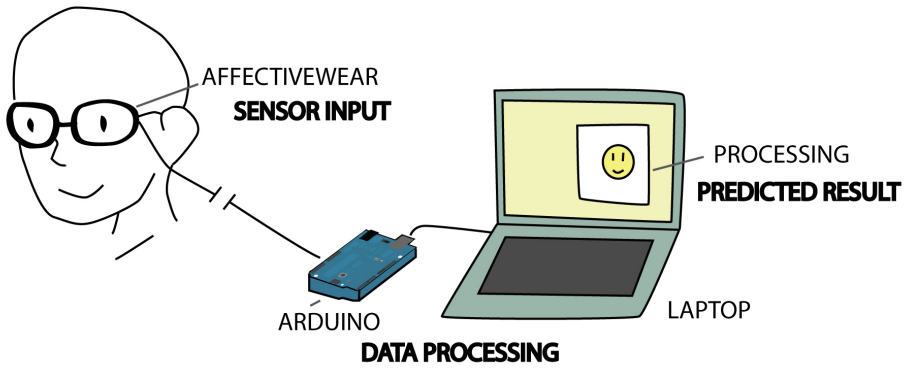

Figure 4.1: *System Overview*

angry, smile, laugh, sad, and surprise, Figure 4.2). 7 facial expressions are picked out from six universal emotions [5](happy, anger, disgust, sad, surprise and fear).

A happy facial expression is divided into two stages: smile and laugh.

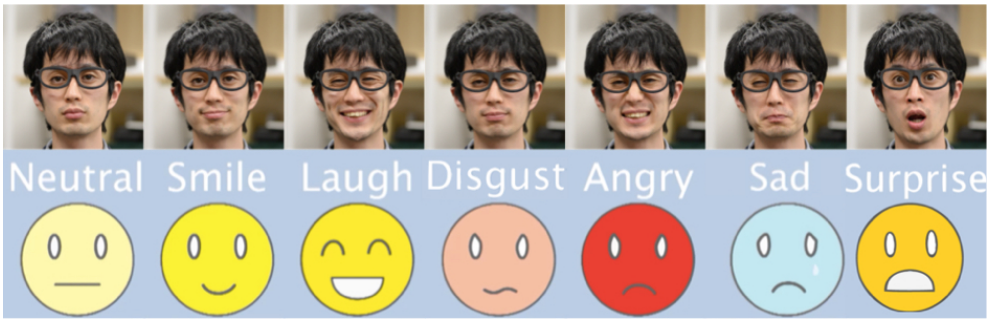

Figure 4.2: *Recognizable Facial Expressions*

4.2. Hardware Implementation

The latest prototype is improved in terms of accuracy from the 2nd prototype. The main difference from the 2nd prototype is that 17 photo reflective sensors was used that is the maximum number for the placement instead of 8 and applied the circuit to reduce the influence of ambient light. In order to reduce the influence of ambient light, The transistor TIP31 is used to turn on and off LED of photo transistors. With this method, it is possible to acquire both sensor data that is not influenced by ambient light and data that is influenced by ambient light. For this prototype, Arduino Uno with 3.3V is used. The same resistor value as 2nd prototype was picked out. Since the analog input of Arduino Uno is only 6, A multiplexer was applied to increase the number of analog inputs. A multiplexer has a logic circuit that uses 4 digital inputs, and by switching them on and off, it allows 16 channels to share 1 analog input.

For this prototype, a CAD board was made with Eagle software because cables take space and make the appearance worse. Figure 4.3 shows the model created in Eagle. The board was cut with PCB printer. The PCB board got thinner than the previous prototype and the distance between face and the sensor as well (Figure 4.4). The front frame is redesigned to make the position of eyewear stable

on face. 3D model of a front frame and a nose pad were made (Figure 4.5). In this model, a folding mechanism of temple is made. The different material was used for a front frame and a nose pad since the latest 3D printer Objet can adjust the compounding ratio of materials (rubber and ABS). Concretely, 80% ABS and 20% rubber material for front frame and 60% ABS and 40% rubber material for nose pad are used. The aim of mixing rubber is to make it more elastic so that users can wear comfortably. The 3D model can be seen in Figure 4.6.

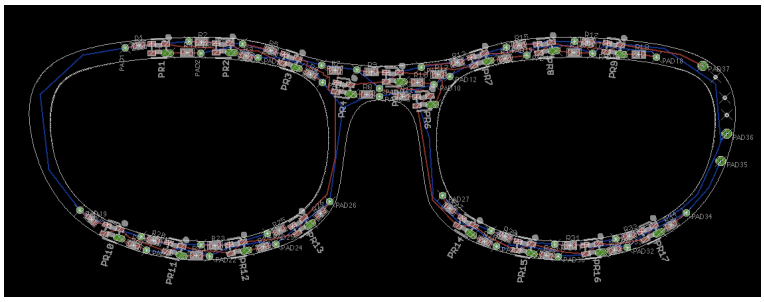

Figure 4.3: *The Latest Prototype AffectiveWear CAD: Eagle*

Figure 4.4: *Abobe: With only Cables, Below: With PCB Printer*

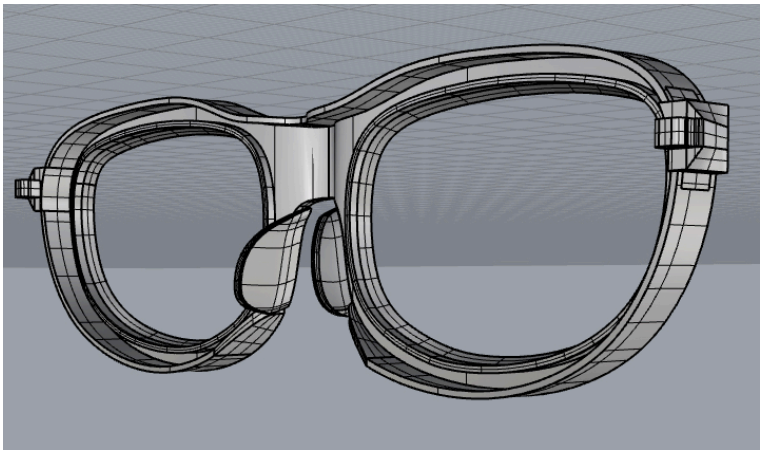

Figure 4.5: *The Latest Prototype AffectiveWear CAD: Rhinoceros*

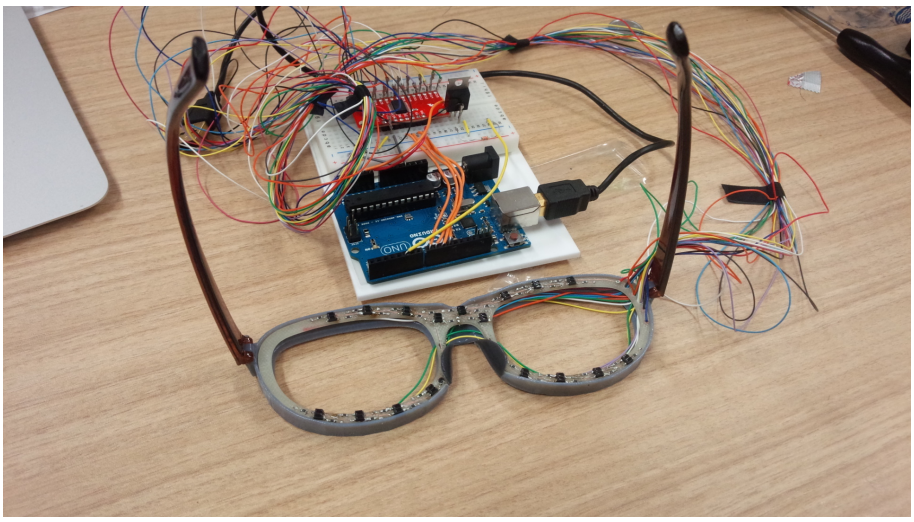

Figure 4.6: *The Latest Prototype AffectiveWear*

4.3. Software Implementation

4.3.1 Arduino

In Arduino environment, input data is processed after acquired from sensors, then sent to Processing. Since digital switch is applied to turn on and off LED of photo reflective sensors in order to get the sensor value with little influence of ambient light, data in both times is collected when switch is on and off. Each sensor value is put in different array A and B after averaged with 5 data in a row. The subtraction of B from A is smoothed out again by averaging 5 continuous data. The aim of averaging method is to reduce the effect of noise from ambient light. Since it takes small fraction of time to switching on and off, the delay method was applied before collecting sensor values: 3 micro second after switching on and 2 micro second after switching off. The frequency of data acquisition is around 40 Hz. All of data is sent as a 10 bit data (0 - 1024).

4.3.2 Processing

In processing environment which is appropriate for interactive programming and visualization, the data is processed with normalization and machine learning, and the result is visualized.

Data Processing

The data sent from Arduino is normalized by following steps.

- 1) The standard of each data(BaseLineValue) is set as 0.5. This is mostly done when users wear AffectiveWear device and make neutral expression.
- 2) The range(Range) of normalization is fixed by calibration. while user dynamically moves facial muscles, the maximum value(Max) and minimum(Min)

value of each data is set. With these values, the normalized value(NSV) is calculated as follows.

$$Range = (Max - Min) * 0.8 + 40 \tag{4.1}$$

(each constant is fixed experimentaly)

$$\begin{cases} NSV = 0.5 + (SensorInput - BaseLineValue)/Range \\ if\ NSV > 1,\ then\ NSV = 1 \\ if\ NSV < 0,\ then\ NSV = 0 \end{cases} \tag{4.2}$$

The aim of using the constant and variable for determining Range is to make most of the advantage of calibration while reducing the unwanted effect caused by calibration since there is a tradeoff with calibration.

- Advantage: Calibration can determine the range which is different according to the position each sensor located. This can improve the accuracy because calibration can set the appropriate weight for each sensor value which should be different.
- Disadvantage: The position of the device is not always stable, which change the max and min value decided by calibration because 1) every time user wear the eyewear, the position of the eyewear can be changed slightly. 2)If user moves too much dynamically, that moves the eyewear. Also, sensor value doesn't have linear relationship with actual distance. These factors may give inappropriate weight for each sensor.

- 3) Normalized data is saved as csv file with the label (1-7, 1: Neutral, 2: Smile, 3: Laugh, 4: Disgust, 5: Angry, 6: Surprise, 7: Sad) where users choose as a desired output for the normalized input. For each label, input from one frame is used.

After the normalization, Support Vector Machine (SVM) model was implemented for the categorization with psvm library. SVM is a machine learning technique that is used in broad field. SVM model was selected because SVM has a good generalization capability by maximizing the margin of each class. SVM can predict facial expression for the input with the trained dataset. In addition to the trained data which includes the output label and normalized input, the calculated value from the adjacent sensor is also used for SVM. The calculation formula is shown below

$$(S1 - S2)/2 + 0.5 \tag{4.3}$$

Using this formula is to reduce the influence of the device position by considering the centroid of all sensor values. The parameter of SVM is picked out as follows.

- C =500
- Gammma = 0.0001
- Kernel = LINEAR

From the perspective of users, user has to do the procedure as below.
1) User wear the device. 2) Set the standard of sensor data set when user make natural expression by pressing (b).
3) User normalizes the range of the movement of facial muscles by moving them dynamically while calibration state is ON. Calibration state is changed by pressing (c)
4) The dataset for each facial expression is created while user makes the facial

expression and press the number corresponding to each facial expression (Neutral:1,Smile:2,Laugh:3,Disgust:4,Angry:5,Surprise:6,Sad:7) while learning state is ON. Learning state is changed by pressing (l)
5) After running SVM by pressing (r) and change prediction state into ON by pressing (p), user checks if facial expression is categorized correctly. If not, more training is employed.

Data Visualization

For real-time prediction and data visualization, user interface was made, which is shown in Figure 4.7.

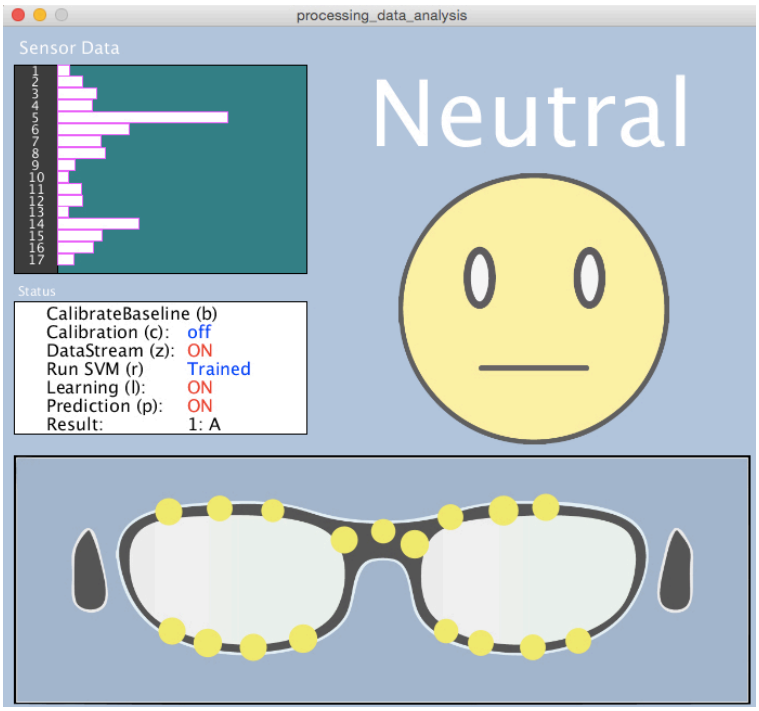

Figure 4.7: *User Interface for AffectiveWear System*

The bar graph in top left part of the user interface shows the raw data from each sensor. It 's for checking if all sensor work or not. The picture in the top right

part shows the result of prediction. Based on training data, it can show real-time prediction. In the bottom part, the size of yellow circle on eyewear picture changes depending on the normalized data from each sensor. The position of yellow circle corresponds to the position of each sensor. It is to understand visually which part of sensors response mostly when user change their facial expressions.

The main machine learning algorithm is based on the work from Masa Ogata.

5. Evaluation

5.1. Initial User Test

For a initial test, the 3rd prototype which has 17 sensors on it was used. The test was performed with 2 volunteers.

1. research participants wear AffectiveWear.
2. research participants set the baseline for neutral facial expression.
3. research participants are asked to make intentional facial expressions about 7 states related to basic emotion. They make facial expressions based on the word an experimenter gives to them. The order of recording is numerical order from 1 to 7.
4. while making each facial expression for 5 -10 seconds, the experimenter labels the number which corresponds to each facial expression. This procedure is for creating the training data.
5. Run the SVM model.

Data wasn't collected from this experiment, yet there is a hardware issue that causes the position of eyewear unstable. Users sometimes can't imagine what facial expression looks like only with a word. Thus, the pictures were prepared so that research participants can tell what each facial expression looks like. In addition, AffectiveWear didn't work on the person with a chiseled face.

5.2. Main Experiment

Based on the initial user test, the procedure of the experiment and the target of the experiment were changed as below. The latest prototype was used. All recording was performed at indoor laboratory environment. 7 volunteers participated this experiment. Their background is: 6 male,2 female, age: 22-26, average age is 24.7, 7 Japanese, 1 French. The experiment was taken place in laboratory setting (Figure 5.1,Figure 5.2)

1. research participants wear AffectiveWear.
2. research participants set the baseline for neutral facial expression.
3. research participants move their face dynamically for the calibration.
4. If the position of eyewear changes when research participants moves their facial expression, the experimenter asks them to hold the eyewear to make the position of AffectiveWear stable during the experiment.
5. research participants are asked to make intentional facial expressions about 7 states related to basic emotion by mimicking the picture shown below (Figure 5.3) [25] .The order of recording is numerical order from 1 to 7. Before recording each state, there has an interval where user adjusts the position of the eyewear and makes neutral expression and set the baseline again. This procedure is for reducing noise between each states.
6. while making each facial expression for 5 -10 seconds, the experimenter presses the number which corresponds to each facial expression. This procedure is for creating training data.
7. Run the SVM model.

8. The author asks research participants to make each facial expression at random to check the predicted results corresponds to their facial expression. If not, the process 5-7 was repeated again.

Figure 5.1: *Experiment Environment 1*

Figure 5.2: *Experiment Environment 2*

5.3. Result

The summary of result can be seen in Table 5.1.

The collected data was analyzed in iPython Notebook and used scikit-learn library.

5.3.1 User Dependency

In order to see if the accuracy of facial expression recognition would change depending on research participants, the analysis for each participant was performed. First, 50% of all data that was randomly selected was trained to SVM. The data that was not used as a training data from each participant is used as test dataset of each participants. SVM was same machine learning algorism as the one used in Processing(SVM, kernel = linear, C = 500, gamma = 0.0001). Next, the predicted results of SVM with test dataset was compared to actual facial expression label

User	A	B	C	D
Gender	Male	Male	Female	Male
Age	26	26	23	25
Nationality	Japanese	Japanese	French	Japanese
DataSet	488	519	528	542
Accuracy with 8 sensors	0.80	0.67	0.92	0.98
Accuracy with 17 sensors	0.97	0.90	1.0	1.0
User	E	F	G	H
Gender	Male	Male	Male	Female
Age	24	26	22	23
Nationality	Japanese	Japanese	Japanese	Japanese
DataSet	531	803	456	417
Accuracy with 8 sensors	0.83	0.65	0.83	0.84
Accuracy with 17 sensors	0.97	0.97	1.0	1.0

Table 5.1: Experiment summary

of test dataset. The formula to decide the accuracy is

$$(PredictedResult/ActualLabel) * 100(\%) \tag{5.1}$$

The result is that if 17 sensors are used, the accuracy got high(90% was minimum). The accuracy was also calucularated with 8 sensors. 8 sensors are selected by the feature selection algorithm SelectKBest. The algorithm doesn't make sure the best sensor number, but it saves the time of calculation, and gives a meaning index of sensor placement. The result is 67% accuracy with User C (minmum) and 98% accuracy with User D.

5.3.2 Sensor Placement

With the dataset collected from all participants, the appropriate position to place sensors and sufficient amount of sensors for detecting glasses wearer's 7 facial expressions was analyzed. The dataset was divided into training data set(70%) and test data set(30%) at random. SelectKBest algorithm was used to pick out the features. The relationship between the placement of each sensor and the sensor number can be seen in Figure 5.4. The Figure **??**. shows the relation between the number of sensors and the accuracy of categorization.

The number of sensors improves the accuracy dramatically until 8 sensors are used. With 8 sensors, the accuracy of categorizing 7 facial expressions becomes 90%. If more than 8 sensors are used, the result improve slowly according to the number of sensors. With all 17 sensors, the accuracy is 98%.From this experiment, the trade-off was confirmed between the accuracy of Facial Expression recognition and the number of sensors. If 8 sensors are used, the optimal placements that the algorithm showed are 4,5,8,9,11,12,14,17 as depicted in Figure5.6. As predicted, the algorithm showed well-balanced placement of sensors. The order of sensor number picked out is as follows: 11, 12, 9, 14, 8, 17, 5, 4, 16, 2, 3, 1, 15, 10, 13, 6, 7.

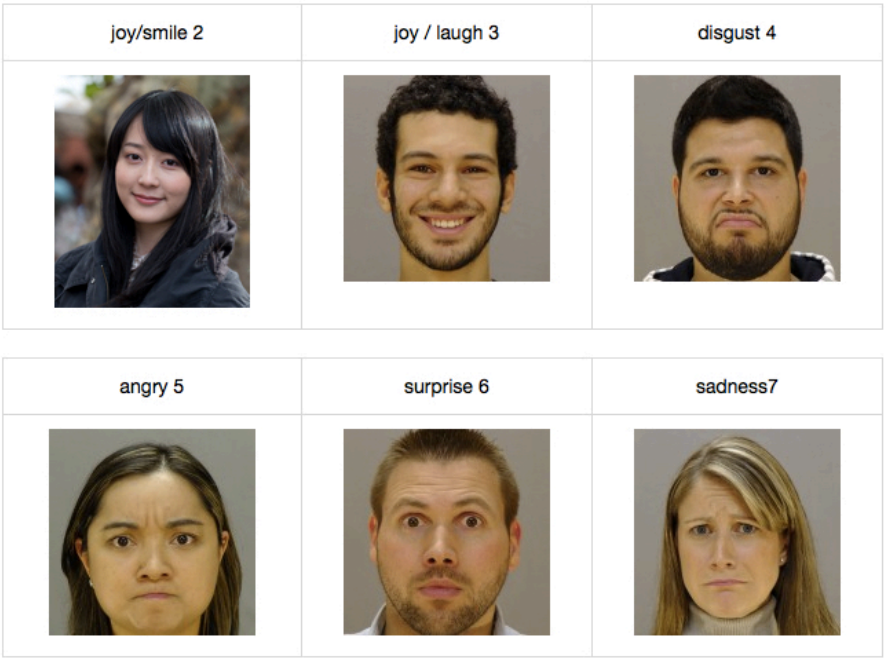

Figure 5.3: *Facial Expression Table*

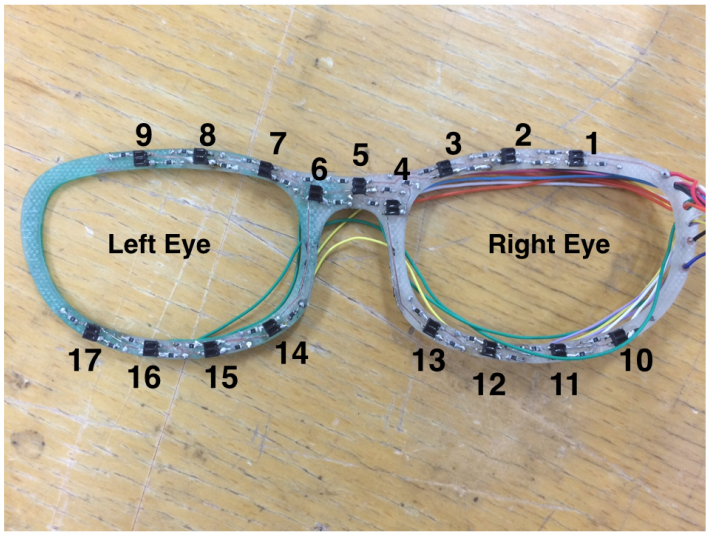

Figure 5.4: *The Placement and Sensor Number*

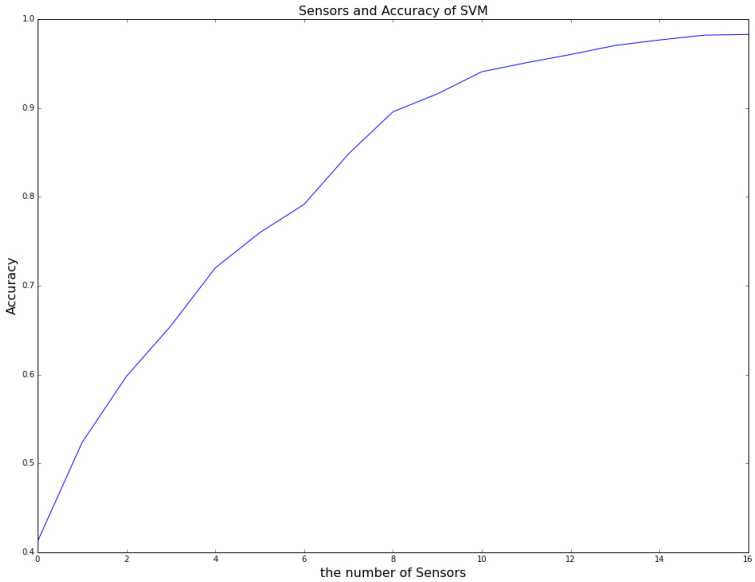

Figure 5.5: *The Relationship between the Accuracy and Number of Sensors*

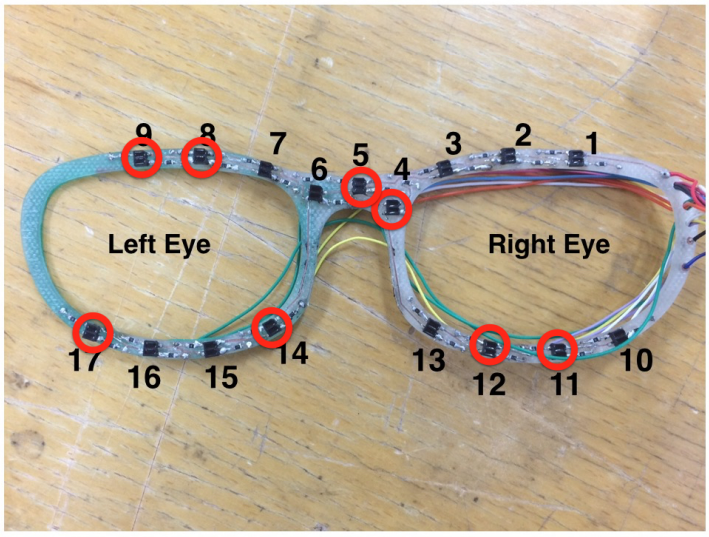

Figure 5.6: *The Best 8 Sensors Placement*

5.3.3 Difference among all states

Each facial expression state has each sensor value characteristics. data of all 8 participants was used. The Figure 5.7 shows the mean and standard deviation of normalizatoin values of each facial expression.

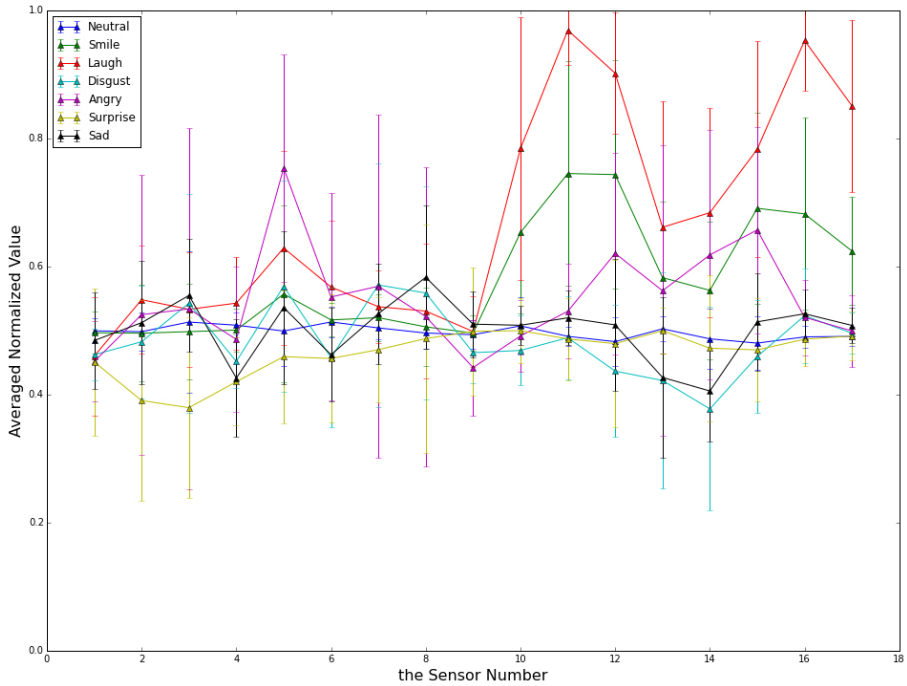

Figure 5.7: *Normalized Sensor Value depending on Each State*

5.3.4 Different time

In order to confirm the consistency of our system, multiple experiment was performed for one user. facial expression data for training sets was collected 5 times on different days. In total, 2429 data was collected. 817 data for test set was also collected in twice on another day. The result was 89% accuracy with 17 sensors. The table shows The confusion matrix of this result.
The result shows that 21% of disgust label is confused as smile or laughter. 33% of laugh label is also mislabeled as smile.

		Predicted Result						
		Neutral	Smile	Laugh	Disgust	Angry	Surprise	Sad
	Neutral	110	0	0	0	0	4	0
	Smile	0	124	0	0	0	0	0
Actual	Laugh	0	26	52	0	0	0	0
Value	Disgust	17	35	4	216	0	0	0
	Angry	0	0	0	0	172	0	0
	Surprise	1	0	0	0	0	41	0
	Sad	0	0	0	2	0	0	65

Table 5.2: Confusion Matrix

6. Conclusion and Future Work

6.1. Discussion

6.1.1 Intentional Facial Expression

Facial expression used for categorization is intentional and not in real life setting. Future work should consider the design of data-collection method in real setting and record for long hours. Yet, AffectiveWear is the significant step because the data collected in our experiment can be also used for spontaneous facial expression because there is a similarity between spontaneous facial expression and intentional one.

6.1.2 Personalization

Since Eyewear is a personalized item, personalization of the design of the eyewear is required. AffectiveWear couldn't fit into diverse people because each user has their own facial features such as the shape of nose and the size of the eye. It makes the position of our eyewear unstable. This requires continuous calibration to captrue facial expression and makes difficult for recording for long hours. Our device works with the Asian face like the author has, and it doesn't work with Western face. In future work, this fact should be considered to create different design of the eyewear to cover diverse people.
In addition, user needs to calibrate and and train to collect data before AffectiveWear can categorize user 's facial expressions with the stable hardware. With

the current work, it takes 2-3 minutes to calibrate. Individual training is required because user has their own way of facial expression and the database was not big enough for all divers users. Future work should create external dataset to reduce the time of calibration.

6.1.3 Intensity of Facial Expression

Since a support vector machine model which is a categorization algorithm, each input is predicted as one of the defined facial expressions. Therefore, the fuzzy facial expression which may be mixed up with a couple of facial expressions is recognized as one clear facial expression. For example, a bitter smile maybe sometimes categorized as disgust, and at other time categorized as smile. This categorization is not accurately reflecting on user's facial expression. Besides, the input from the facial gesture such as talking can influence the prediction of facial expressions, and can be categorized as one of nonneutral facial expressions, which is not desirable. Future work should tackle to categorize a smile into a natural smile, a biter smile and a false smile.

6.1.4 Contexts of Facial Expression

The meaning of facial expression can be different depending on the context. If user thinks hard while reading the book, his eyebrow moves and his facial expression looks like angry though he is just serious, concentrated. Or, people sometimes makes a false smile which does not imply the same meaning as a smile when people are happy. To understand the context of facial expression, the analysis of chronological data is needed. In addition, it is meanigful to integrate other modalities or sensors such as temperature sensor or humidity sensor since our device is wearable. Future works should consider what sensors should be integrated so that the device captures the meaning behind the facial expressions deeply.

6.1.5 Hand Gesture to Face

Facial expression can be recognized only when user doesn't touch their face. If user touchs their face with hands, then this action moves the facial muscles. These movement influence the sensor value. Hence, touching your face keeps from recognizing facial expression correctly. As a future work, it might be possible to use our device as a ḣand gesture to faceẗecognition.

6.2. Usage and Application

In this section, one implemented application (HMD) and possible application idea are discussed.

6.2.1 Main Usage: Feedback System for User Activity

Capturing the facial expressions in everyday life can help improve our mental health. [24], If our system can let the user gain insights about their facial expressions over days (How often did I smile this week versus last?), tracking these changes enables users to understand more about their frequent unintentional non-verbal clues, helping them to improve their mindsets. Users suffering from depression or other mental disorders might get indications if their state is improving using our system. Also, it is useful to check the emotional state of elderly people or children and keep an eye on them.

6.2.2 Application1: HMD Application

With HMD embedded with our technology, User can reflect their facial expression on their avatar in virtual world(Figure 6.1). Oculus Rift and Unity is used for this application. 7 photo reflectors are used. With Kinect, user also can reflect

their movement. People can enjoy more natural and rich ways of communication in virtual world.

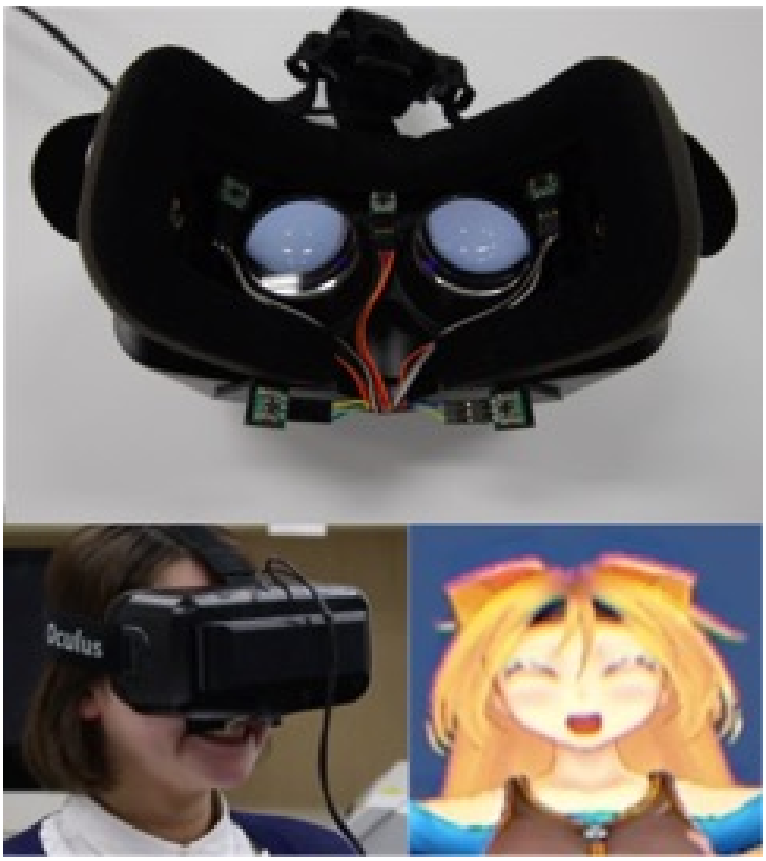

Figure 6.1: *HMD Application*

6.2.3 Application2: Typography and Emoticon

Text message does not always convey enough information. In order to convey the nuance or the context behind the messages, we use emoticons and symbols such as question mark and exclamation mark. Or in Manga, typography of character ' s messages changes according to their emotional state. Typography has the power to change the impression of messages. With our device, user can put emoticon or change their typography of text message according to facial expression while they

are texting. Emoticon is recommended after users end a sentence so that they can make final decision.It can enrich the communication with texts.

6.2.4 Application3: Tagging Facial Expressions

Tagging the facial expressions to images or movies captured using a life-logging system, e.g. the Narrative Clip or a GoPro. By knowing user 's facial expression, user can remember well about the scene user record.

6.3. Conclusion

Toward recognizing facial expression in real setting scenario, an eyewear type prototype is presented in this thesis. It can recognize 7 facial expressions in various places regardless of the wearer's head angle. By using multiple photo reflective sensors, it is proved that facial expression can be captured. Also, the design of device can be natural af if it looks like a normal eyewear. The current prototype allows users to enjoy short-term application such as putting emoticon and HMD application.

The prototype was tested with 8 people, and achieved 98% accuracy to recognize 7 facial expressions with individul training. The design of the current imptemented hardware makes the postion of our prototype sometimes unstable which recquires continuous calibration and makes it difficult to record long hours which is important to investigate our mind from the perspective of facail expression. If this problem is solved, it has a potential to be used in many applications because facial expression contains rich information about social interaction and emotional state.

References

[1] Apple. Apple Watch. `https://www.apple.com/watch/health-and-fitness/`, 2015.

[2] Clark, A. *Being there: Putting brain, body, and world together again.* MIT press, 1997.

[3] Clark, A. *Natural-born cyborgs: Minds, technologies, and the future of human intelligence.* Oxford University Press, 2004.

[4] Darwin, C., Ekman, P., and Prodger, P. *The expression of the emotions in man and animals.* Oxford University Press, USA, 1998.

[5] Ekman, P. Are there basic emotions?

[6] Ekman, P., and Friesen, W. V. Facial action coding system.

[7] Fukumoto, K., Terada, T., and Tsukamoto, M. A smile/laughter recognition mechanism for smile-based life logging. In *Proceedings of the 4th Augmented Human International Conference*, ACM (2013), 213–220.

[8] Gruebler, A., and Suzuki, K. Design of a wearable device for reading positive expressions from facial emg signals.

[9] Ishiguro, Y., Mujibiya, A., Miyaki, T., and Rekimoto, J. Aided eyes: eye activity sensing for daily life. In *Proceedings of the 1st Augmented Human International Conference*, ACM (2010), 25.

[10] James, W. *The Principles of Psychology (Volume 1 of 2).* Digireads. com Publishing, 2004.

[11] JAWBONE. JAWBONE UP2. https://jawbone.com/store/buy/up2, 2015.

[12] JINS. JINSMEME. https://www.jins-jp.com/jinsmeme/en/, 2015.

[13] Kanade, T., Cohn, J. F., and Tian, Y. Comprehensive database for facial expression analysis. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, IEEE (2000), 46–53.

[14] Kimura, S., Fukuomoto, M., and Horikoshi, T. Eyeglass-based hands-free videophone. In *Proceedings of the 2013 International Symposium on Wearable Computers*, ACM (2013), 117–124.

[15] Littlewort, G., Fasel, I., Bartlett, M. S., and Movellan, J. R. Fully automatic coding of basic expressions from video. *University of California, San Diego, San Diego, CA 92093* (2002).

[16] Nakamura, H., and Miyashita, H. Control of augmented reality information volume by glabellar fader. In *Proceedings of the 1st Augmented Human International Conference*, ACM (2010), 20.

[17] Niedenthal, P. M. Embodying emotion. *science 316*, 5827 (2007), 1002–1005.

[18] Ogata, M., Sugiura, Y., Makino, Y., Inami, M., and Imai, M. Senskin: adapting skin as a soft interface. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, ACM (2013), 539–544.

[19] Provine, R. R. Contagious yawning and laughter: Significance for sensory feature detection. *Social learning in animals: The roots of culture* (1996), 179.

[20] Scheirer, J., Fernandez, R., and Picard, R. W. Expression glasses: a wearable device for facial expression recognition. In *CHI'99 Extended Abstracts on Human Factors in Computing Systems*, ACM (1999), 262–263.

[21] Taniguchi, K., Nishikawa, A., Kobayashi, E., Miyazaki, F., and Sakuma, I. Me-me switch: a wearable input device using movement of ears. *Proceedings of the interaction 10* (2010).

[22] Tian, Y., Kanade, T., and Cohn, J. F. Facial expression recognition. In *Handbook of face recognition*. Springer, 2011, 487–519.

[23] Tian, Y.-l., Kanade, T., and Cohn, J. F. Recognizing action units for facial expression analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 23*, 2 (2001), 97–115.

[24] Tsujita, H., and Rekimoto, J. Smile-encouraging digital appliances. *IEEE Pervasive Computing 12*, 4 (2013), 5–7.

[25] VisualEmotion LLC. 7 Universally Recognized Facial Expressions of Emotion. `http://www.facscodinggroup.com/universal-expressions`, 2015.

[26] Wei, J. How wearables intersect with the cloud and the internet of things: Considerations for the developers of wearables. *Consumer Electronics Magazine, IEEE 3*, 3 (2014), 53–56.

[27] Yoshida, S., Tanikawa, T., Sakurai, S., Hirose, M., and Narumi, T. Manipulation of an emotional experience by real-time deformed facial feedback. In *Proceedings of the 4th Augmented Human International Conference*, ACM (2013), 35–42.