Doctoral Dissertation

Shibaura Institute of Technology

Healthcare System focusing on Emotional Aspect

-Using Augmented Reality, Facial Expression, and ECG Signal-

September 2015

Somchanok Tivatansakul

HEALTHCARE SYSTEM FOCUSING ON EMOTIONAL ASPECT

-USING AUGMENTED REALITY, FACIAL EXPRESSION, AND ECG SIGNAL-

BY

SOMCHANOK TIVATANSAKUL

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

IN THE
FUCTIONAL CONTROL SYSTEMS
GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
SHIBAURA INSTITUTE OF TECHNOLOGY

SEPTEMBER 2015

Declaration of Authorship

I, Somchanok TIVATANSAKUL, declare that this thesis titled, “Healthcare system focusing on emotional aspect using augmented reality, facial expression and ECG signal,” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at Shibaura Institute of Technology.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at Shibaura Institute of Technology or any other institution, this has been clearly stated.
- Where I have consulted the published work of other, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with other, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ____________________________________

(Somchanok TIVATANSAKUL)

Date: _______________________________________

Abstract

In this doctoral dissertation, I proposed a new healthcare system focusing on emotional aspects to cope with daily negative emotional health.

To make the healthcare system more attractive and effective, I proposed a breathing control application for a relaxation service using a deep breathing technique and augmented reality to increase user relaxation and decrease user stress. My evaluation results suggested that the breathing control application for the relaxation service decreased stress and the augmented reality helped users decrease stress more quickly than only the deep breathing technique.

To make the healthcare system more effective and intelligent, I integrated emotion recognition by facial expression. To increase the accuracy of existing facial emotion recognitions, I proposed the Complementary Directional Ternary Pattern (CDTP) algorithm to recognize emotions. CDTP was designed to decrease the size of DTP feature vectors and to reduce the feature redundancy in pattern representation while maintaining the integrity of the extracted feature. My evaluation results indicated that CDTP with SVM recognized basic emotions with high accuracy and performance.

However, emotion recognition from facial expression gives confusion issue on recognition. To solve this and increase the accuracy and the efficiency of the emotion recognition by facial expression, I integrated emotion recognition from ECG signal. To recognize emotions from ECG signal, I adapted LBP, LTP and CLTP, all of which are favorable local pattern description methods for emotion recognition by facial expression. My evaluation results indicated that LBP and LTP effectively extracted ECG features with high accuracy using a k-NN classifier.

To apply my proposed system in a real environment, I built a prototype of the emotional healthcare system. The experimental results showed that the breathing control application of the relaxation services also effectively decreased negative emotions. The emotion recognition by facial expression effectively recognized negative emotions but recognizing some emotions caused some confusion. The results also indicated that the integration of emotion recognition by facial expression and ECG signal addressed the confusion issue and increased the accuracy of the facial emotion recognition. However, simply recognizing negative emotions might not be adequate to provide the relaxation service. Therefore, stress detection from ECG signal was integrated in this prototype to recognize stress with emotion recognition. Thus, emotion recognition and stress detection had a potential for the emotional healthcare system to provide assistance when users experienced stress or negative emotions. After improving the prototype's efficiency, it became more suitable for practical usage.

Based on this dissertation's results, I achieved all research goals to construct an attractive, effective and intelligent emotional healthcare system using augmented reality, emotion recognition by facial expression and ECG signal, and stress detection from ECG signal. I obtained a new design of an emotional healthcare system, a new relaxation service with augmented reality, a new feature extraction approach (CDTP) for recognizing emotions from facial expression, new feature extraction approaches (LBP and LTP) for emotion recognition from ECG signal, a new real-time emotional healthcare system using real-time emotion recognition by facial expression and ECG signal, and real-time stress detection from ECG signal. The findings of this dissertation are also applicable to other research and systems, such as a new feature extraction algorithm (CDTP) for recognizing emotions from facial expression and new feature extraction approaches (LBP and LTP) for emotion recognition from ECG signal.

"IF YOU CAN DREAM IT, YOU CAN DO IT"

WALT DISNEY

"ความพยายามอยู่ที่ไหน ความสำเร็จอยู่ที่นั่น"

(WHERE THERE IS A WILL, THERE IS A WAY)

THAI QUOTE

Acknowledgements

Firstly, I would like to sincerely thank my supervisor, Prof. Michiko Ohkura, for her contribution of time, valuable advice, guidance, understanding, and patience during my research study. The joy and enthusiasm she has in the research have inspired me not only to my academic life, but to my personal life as well. She has always been there and willing to listen and help everyone in the laboratory. Her valuable advice has kept me on track and helped me to the completion of this research, for which I truly appreciate.

I would also like to express my sincere gratitude to Prof. Tiranee Achalakul of Department of Computer Engineering of King Mongkut's University of Technology Thonburi (KMUTT) who has provided an excellent opportunity for me to pursue my doctoral degree and conduct this collaborative research between Shibaura Institute of Technology and KMUTT. I would not have been here without her support and encouragement. Additionally, I would like to thank Prof. Santitham Prom-on and Prof. Suthathip Maneewongvata from KMUTT for their kindly support and advice.

I would like to express my appreciation to all committee members (Prof. Ken Tomiyama, Prof. Eiji Kamioka, Prof.Masaomi Kimura, Prof. Hiroaki Fukuda and Prof. Michiko Ohkura) for serving as my committee members even at hardship. I also want to thank you for letting my defense be an enjoyable moment with invaluable suggestion and comments.

Furthermore, I would like to thank Hybrid Twining Program and Japanese Government Scholarship (Monbukagakusho). Their financial support throughout my academic career has also made this research possible. I would especially like to thank my participants who volunteer for my experiment. I also would like to thank the development team; Nattawat Chanthaphan (for emoticon virtual model) Supadchaya Puangpontip, Gantaphon Chalumporn and Yada Kankanokkuland (for LDP), I would like to thank Prof. Makoto Mizukawa and Dr. Kanlaya Rattanyu for International affective picture system and I also thank Sarunya Kanjanawattana and Tipporn Loahakangvalvit for their guidance about Data mining. Furthermore, I am really appreciated all Thai friends at SIT and my lab members at the Information System Engineering Laboratory (SIT) and CASTLAB (KUMTT) for their kind help during the past three years, their friendship and good humor which have made my life in Japan and both laboratories much more enjoyable.

Lastly and most importantly, I would like to express my deepest thanks to my family. Words cannot express how grateful I am to my parents, my two little sisters, my little brother, my boyfriend and his family for all their love and care. Their support and encouragement have helped me along the way. They are such a big part of my life and I am beyond grateful to have them.

Contents

List of Figures

LIST OF FIGURES

List of Tables

Abbreviations

AAM	=	Active appearance model
ANG	=	Anger
ANOVA	=	Analysis of variance
ANS	=	Autonomic nervous system
AR	=	Augmented reality
AuBT	=	Augsburg biosignal toolbox
ASM	=	Active shape model
BPM	=	Beats per minute
BEMD	=	Empirical mode decomposition
CBT	=	Cognitive behavioral therapy
CDTP	=	Complementary directional ternary patterns
CLTB	=	Complementary local ternary patterns
CI	=	Confident interval
CK+	=	Extended Cohn-Kanade dataset
DIS	=	Disgust
DTP	=	Directional ternary patterns
DWT	=	Discrete wavelet transform
ECG	=	Electrocardiography
EDR	=	ECG-derived respiration
EEG	=	Electroencephalography
EMG	=	Electromyogram
F	=	F-test
FDR	=	Fisher discriminant ratio

FEA	=	Fear
GMM	=	Gaussian mixture model
GSR	=	Galvanic skin response
HAP	=	Happiness
HCI	=	Human computer interaction
HF	=	High frequency
HMD	=	Head-mounted display
HOC	=	Higher order crossing
HOC-EC	=	HOC-emotion classifier
HR	=	Heart rate
HRI	=	Human robot interaction
HRV	=	Heart rate variability
IIS	=	Internet information services
IMF	=	Intrinsic mode functions
JAFFE	=	Japanese female facial expression dataset
KDEF	=	Karolinska directed emotional faces dataset
k-NN	=	k-nearest neighbor
LBP	=	Local binary patterns
LDA	=	Linear discriminant analysis
LDP	=	Local directional patterns
LF	=	Low frequency
LTP	=	Local ternary patterns
Mdiff	=	Mean difference
M	=	Mean
MFCC	=	Mel frequency cepstrum coefficient
MSFs	=	Modular spectral features
NEU	=	Neutral
P	=	P-value
PCA	=	Principle component analysis
PLP	=	Perceptual linear predictive
QDA	=	Quadratic discriminant analysis
RBF	=	Radial basis function
RMSSD	=	Root mean square of standard deviation of NN interval

RSP	=	Respiration change
SAD	=	Sad
SC	=	Skin conductivity
SDNN	=	Standard deviation of normal-to-normal interval
SE	=	Standard Error
SMS	=	Short Message Service
SUR	=	Surprise
SVM	=	Support vector machine
TP	=	Total power
VAM	=	Vera am Mittag database
VCR	=	Video cassette recording
VR	=	Virtual reality
VDT	=	Visual display terminal
WHO	=	World health organization

Chapter 1

Introduction

This chapter introduces the motivation underlying my doctoral dissertation, “Healthcare system focusing on emotional aspect using augmented reality, facial expression and ECG signal.” This chapter also presents the existing research issues and problems that my dissertation addresses as well as its solutions. Furthermore, it describes my dissertation’s goals and how to achieve them. Finally, I summarize this chapter and explain an organization of my dissertation.

1.1 Motivation

In recent years, the design and implementation of smart home, intelligent space and healthcare systems have become very popular in the field of human computer interaction

(HCI) and human robot interaction (HRI). The basic idea of such systems is to automatically monitor both the environments and the humans in them to provide assistance and services (Figure 1.1). Several systems provide more support of the physical aspects of people and downplay the emotional aspects. Less research has focused on user emotions than on physical aspects. However, emotional health is as important as physical health [1-2].

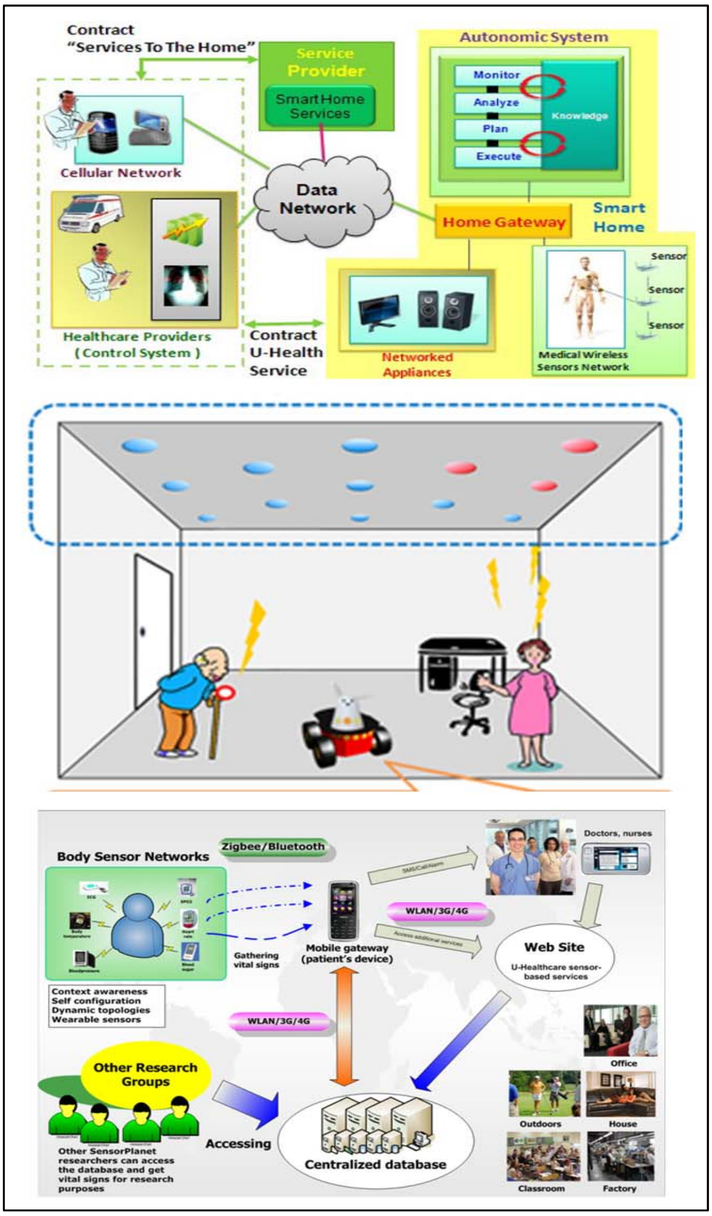


Figure 1.1 Examples of smart homes [3], intelligent spaces [4] and healthcare systems [5] focusing on physical aspects of humans

Human generally express emotions when they experience good or bad situations. Positive and negative emotions can affect human behaviors, thoughts, and feelings. Positive emotions include happiness, pleasure, and satisfaction. Negative emotions include sadness, anger, anxiety, and fear [6]. People with more positive emotions can control themselves and address their negative emotions. However, people with more negative emotions experience difficulty controlling and decreasing such emotions (Figure 1.2). Therefore, negative emotions can lead to social and mental health problems.

According to the World Health Organization (WHO) [7], worldwide suicide rates have increased by 60% in the last 45 years [8]. Unfortunately, Japan ranks among the top-ten countries. One reason is such negative emotions as anxiety, sadness, boredom, and stress that are often caused by compulsory studying and excessive work.

Figure 1.2 People with negative emotions and stress [9-10]

To address such problems, many researches have emphasized user emotions to support patients in psychological, phobia, anxiety, and stress disorders [11-14]. These proposed systems diagnose and examine users to provide treatment plans and services for improving their emotional and mental states.

However, the emotional health of ordinary people should be concerned as well. Some people cannot cope with or control daily negative emotions. Some might experience difficulty controlling or relieving stress. In the worst cases, severe psychological symptoms might develop and then they might become psychological patients. Therefore, they probably need therapy or counseling to help them cope with daily stress or negative emotions (Figure 1.3).

To cope with negative emotions in daily life and improve the emotional health of ordinary people, I propose a new healthcare system focusing on emotional aspects using augmented reality (AR). I apply AR to display virtual objects in real environments to make my system more attractive because it enables the interaction between users and virtual objects in real environments and outperforms virtual reality. AR examples are shown in Figure 1.4.

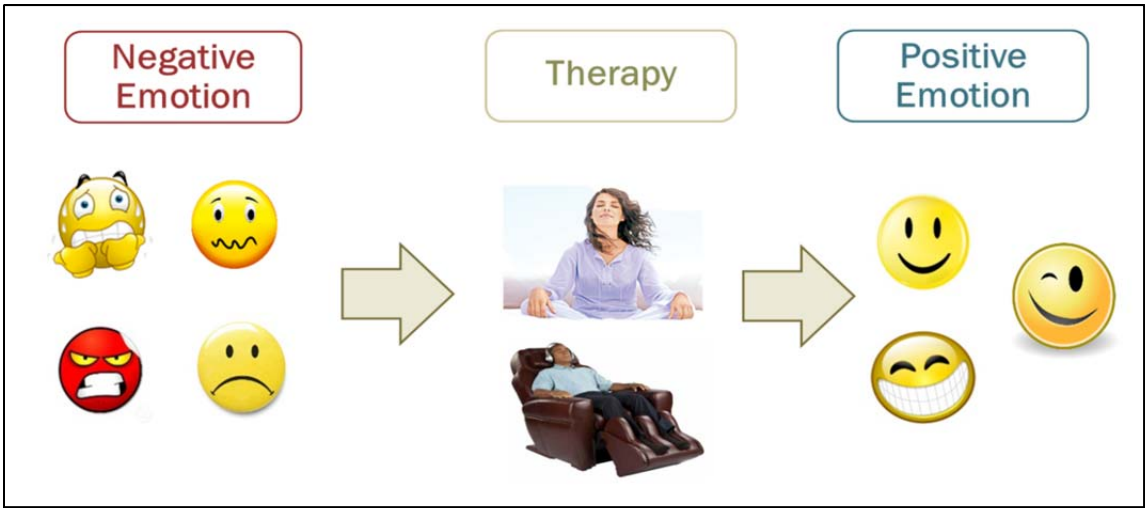


Figure 1.3 Therapy for decreasing negative emotions and increasing positive emotions [15-17]

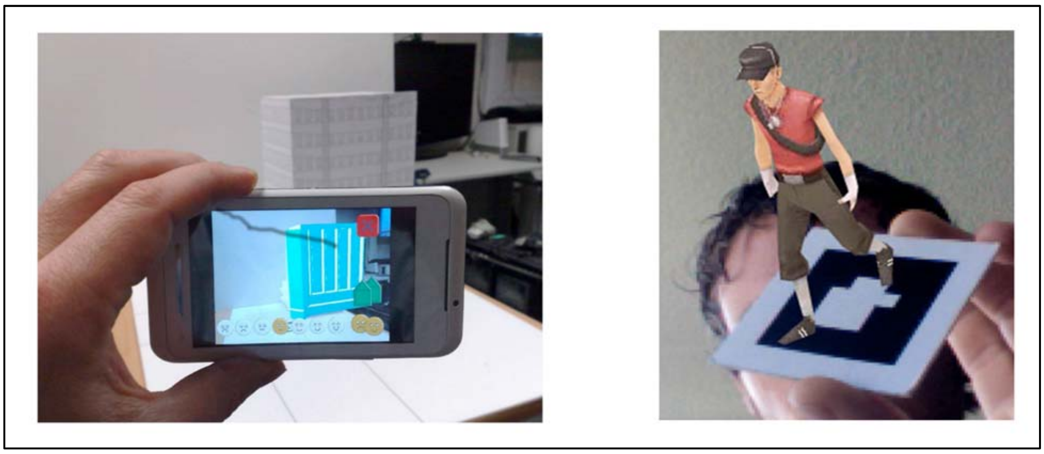


Figure 1.4 Examples of AR applications [18-19]

1.2 Problem Statement

In Human Computer Interaction (HCI) fields, even though several healthcare systems have been proposed for various kinds of patients, few systems are designed for ordinary people. These systems still undergo some issues that must be improved such as coping with the emotional health problems of ordinary people, or recognizing current emotional states. This section describes the problems of healthcare systems, relaxation services, and emotion recognition for proposing the healthcare system focusing on emotional aspects using augmented reality, facial expression and ECG signal to solve the problems.

1.2.1 Healthcare systems

Nowadays, the existing healthcare systems in the HCI concentrate more on providing physical healthcare at the expenses of emotional healthcare. Few emotional healthcare systems provide treatment to support only patients. However, the emotional/mental state of ordinary people is also critical and it should be concerned as well. To maintain good emotional/mental health, systems are necessary that help people relieve negative emotions and stress and encourage relaxation in daily life [20].

In healthcare systems, the key factor for successfully addressing mental/emotional health problems is to increase user interest in using such systems. If healthcare systems are attractive, effective and intelligent, users will want to use them. They will successfully help users cope with their problems. Moreover, the treatment or therapy contents provided to users as well as the content sources are also important to make systems more attractive and effective. Generally, healthcare systems provide text, images, and sounds; however, since using these sources might not be attractive enough, as the results, treatment or therapy might not be effective. To increase the attractiveness of the content and the effectiveness of the treatment, one alternative for healthcare systems is to provide content by displaying virtual objects in the real world.

Furthermore, to make treatment more effective and healthcare systems more intelligent, the user emotional/mental states should be aware. Thus, one necessary feature in healthcare systems is real-time emotion recognition that can recognize user emotions in short-time durations to provide assistance when users feel bad. Real-time emotion recognition is critical because it can activate the system to provide appropriate treatment/therapy based on user emotional states [21]. However, real-time emotion recognition in healthcare systems that provide feedback is not as common as in other fields such as robotics, entertainment and the automotive industry. Real-time emotion recognition must be integrated into healthcare systems. Various sources also exist for recognizing such emotions as facial expressions, speech, body gestures or biological signals. Recognition sources must not be neglected; if they are more appropriate, recognition rate will be more accurate.

1.2.2 Relaxation services

Several relaxation applications in HCI fields employ many techniques to encourage user relaxation and reduce stress or negative emotions. However, these existing relaxation applications only provide text, sound, or image content for relaxing users. As mentioned in Section 1.2.1, not only the content itself is important but a source to display is also needed to make the application more effective and more attractive. Technology that displays virtual objects in the real world must be integrated with treatment content to improve the attractiveness of relaxation applications.

1.2.3 Emotion recognition

The popularity of research on emotion recognition by facial expressions continues to increase. Several researchers have proposed various feature extraction methods to improve the accuracy of emotion recognition by facial expression in both offline and real-time processes. Although no techniques can achieve 100% accuracy, the Gabor wavelet filter [22], Active Appearance Model (AAM) [23-24], Local Binary Pattern (LBP) [25-26], Directional Ternary Pattern (DTP) [27] are popular for recognizing emotions because they achieve higher accuracy or better performance (execution time). For real-time processes, simply providing high accuracy is not enough; real-time emotion recognition also needs good performance. However, one issue is that such feature extraction approaches as AAM and DTP [23, 24, and 27] with higher accuracy take more time to recognize emotions;

those with good performance cannot recognize emotions with adequate accuracy, such as LBP. To address this issue, a new feature extraction method with high accuracy and performance must be proposed to improve the existing methods.

Emotion recognition using ECG signal is another favorable technique to recognize human emotions. In general, signal processing or medical diagnoses with statistical methods are applied to extract ECG features [28]. However, the accuracies of these techniques depend on the ability to correctly detect the components of such ECG signal as RR and QRS intervals, etc. In a real-time process, if the ECG signal is weak or contains noise, these techniques might have difficulty correctly detecting the ECG components and fail to get high accuracy, since ECG signal contains such noise as power line interference, baseline wander, muscle noise, and motion artifacts, all of which are caused by ECG sensors or user movements during recording the signal [29]. Therefore, applying other techniques to extract the ECG features from the entire ECG signal without the detection of RR, or QRS intervals or other ECG components is one option for real-time emotion recognition from ECG signal.

1.3 Research Questions

Regarding my problem statements, several issues of the healthcare system focusing on emotional aspects for ordinary people must be solved to design and build the system that can provide effective emotional services based on the current emotional states of users. Such system must be intelligent by automatically recognizing user emotions. Based on these issues, I investigated the following questions.

1. How should the healthcare system focusing on emotional aspect be designed and which devices, applications and services are necessary for maintaining user emotional health?
2. How can this system relax users, which techniques should this system apply, and how should the application be designed to be more attractive and effective to support users?
3. How can this system detect user emotions with high accuracy and performance?

To answer these questions, I research, design, implement, and evaluate services, applications, and emotion recognition. After that, I implement the prototype of this system by combining each part. Finally, I experimentally evaluate the prototype in a real-time process to confirm its efficiency and effectiveness for recognizing user emotions and stress to provide appropriate services and maintain the emotional health of users.

1.4 Research goals and contribution

Regarding my problem statements and research questions, the new key point that existing healthcare systems should take into consideration is the daily emotional/mental health of ordinary people. Healthcare systems should improve the attractiveness of their therapy content and improve their intelligence to maintain good emotional/mental health for users. These ideas influence me to conduct this dissertation which researches on the design and implementation of the healthcare system focusing on emotional aspect for ordinary people. The research goals and contribution are described below.

1. This dissertation proposes a new healthcare system focusing on emotional aspect.
2. This dissertation proposes a relaxation service using augmented reality that must be designed and implemented to effectively relax users and improve the attractiveness of the treatment contents. The service's effectiveness for decreasing stress and negative emotions must be verified.
3. Emotion recognition by facial expression and ECG signal must be applied for the real-time recognition of user emotional states from both user appearances (outside) and biological signals (inside) with high accuracy and performance before/after providing appropriate services.
4. Finally, the prototype of healthcare system focusing on emotional aspect must be implemented and evaluated to confirm its effectiveness.

1.5 Research Requirements

In this section, I describe the requirements to achieve my first three research goals.

1. I propose a new healthcare system focusing on emotional aspect because little research has emphasized on emotional health of ordinary people, as I explained in the motivation part (Section 1.1). This system is designed to support people at daily workplace by providing services that allow them to interact with virtual objects in real environments to feel more positive emotions and to decrease negative emotions or stress that might reduce workplace productivity. Negative emotions can also affect working ability and attitudes.
2. I propose and design a relaxation service using augmented reality. The relaxation service should be designed first in this emotional healthcare system because relaxation is critical for building daily positive attitudes and good health to address negative emotions and stress that cause several diseases. I also design this relaxation service with augmented reality to improve its attractiveness to encourage users to use the service. Although other technologies can also improve the user interface's attractiveness by images, 2D animation, or virtual reality. However, I choose augmented reality because it enables better interaction between users and virtual objects in real environments than the others, since it can display virtual objects in the real world. The service's design should also combine treatment techniques with AR to decrease negative emotions and stress.
3. To make the emotional healthcare system more intelligent, the emotional/mental states of users must be aware. Therefore, I integrate real-time emotion recognition so that my system can realize user emotional states to provide appropriate treatment/therapy in short-time duration with high accuracy and good performance. First, I apply emotion recognition by facial expression to my system because a facial expression is a natural user interface and is the most common and easiest way for humans to express emotions. The recognition of the emotional state of users using facial expressions in workplaces is easier than other sources, such as speech. However, emotion recognition by facial expression has limitations. For example, users need to express their emotions on their face; if they do not, the system is not

intelligent enough to provide appropriate services. Therefore, I apply emotion recognition and stress detection from ECG signal because they can analyze and capture internal user emotions and stress. ECG signal can also be utilized to analyze other healthcare information. Since, ECG sensors might be designed as daily accessories in the near future, the issue of attaching them to user bodies might be solved.

1.6 Organization of dissertation

This dissertation consists of nine chapters including this one. Figure 1.5 shows the relationship between my proposed healthcare system and the organization of Chapters 3 to 7. The following are the organization of the remaining Chapters 2 to 9.

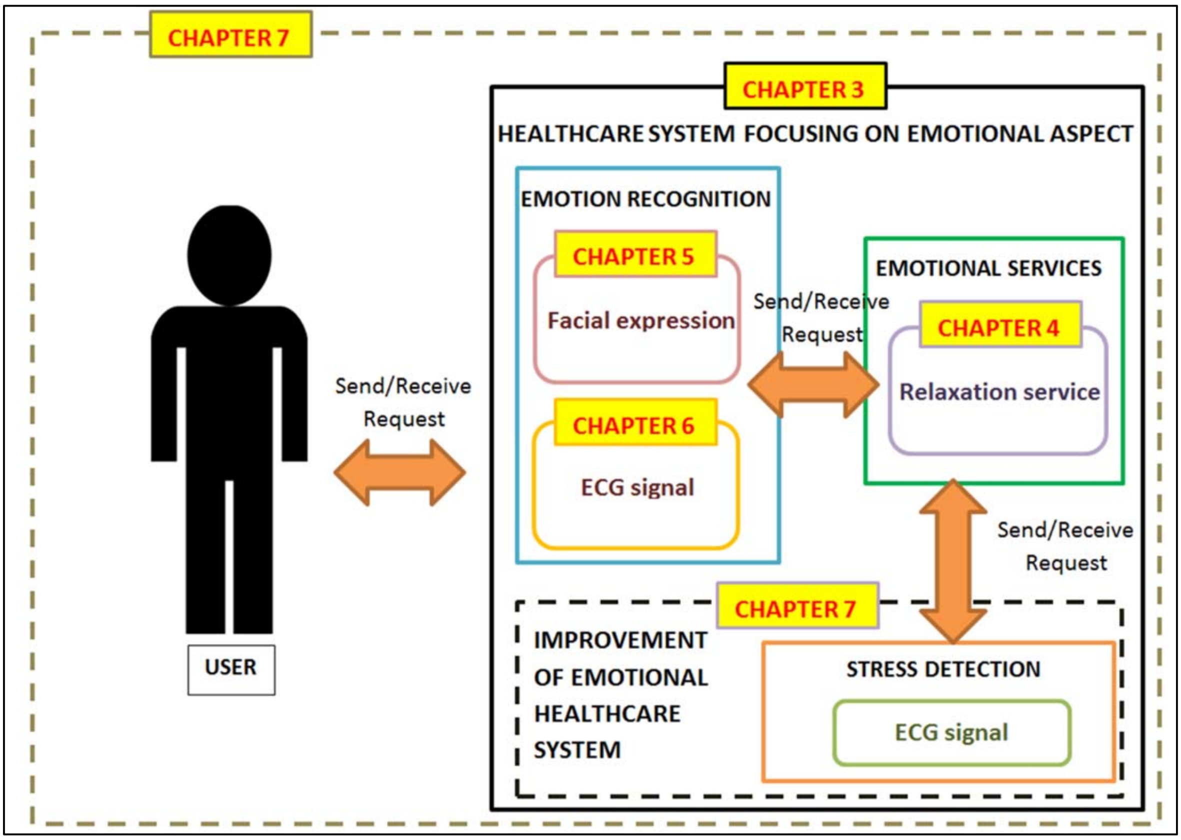


Figure 1.5 Organization of dissertation

Chapter 2 presents a literature review of the current researches in healthcare systems, relaxation services, and emotion recognition in the HCI field that contributed to this dissertation. This survey is critical to achieve my dissertation goals.

Chapter 3 presents the overall system and framework designs of my proposed healthcare system focusing on emotional aspects and describes its necessary support services, devices and modules.

To maintain user emotional states, Chapter 4 presents the design, implementation and evaluation of a relaxation service, which is designed and implemented to reduce user stress and such negative emotions as sadness, anger and fear. After that, I conducted two experiments to evaluate how effectively the relaxation service decreased stress.

Chapter 5 presents emotion recognition by facial expression because, before the healthcare system can maintain user emotional states, it needs emotion recognition that recognizes user emotions. I designed a new facial feature extraction approach that improves the existing approaches and implemented it to fit a real-time process. I also conducted three experiments to confirm the accuracy and the effectiveness of my new feature extraction approach.

Based on some limitations of emotion recognition by facial expression, emotion recognition using ECG is applied to improve accuracy. Chapter 6 presents emotion recognition using ECG. The design and implementation of selected feature extraction approaches for emotion recognition using ECG signal are explained. Then, I conducted the performance evaluation to confirm the efficiency of the emotion recognition using ECG signal.

To allow users to experience and evaluate my proposed healthcare system, Chapter 7 presents a prototype that integrated the relaxation service (Chapter 4) with emotion recognition (Chapters 5 and 6). I conducted two experiments to evaluate the relaxation service and real-time emotion recognition by facial expression and ECG signal. Based on the results, I improved the prototype with stress detection from ECG signal and evaluated the improved prototype again.

Chapter 8 discusses this dissertation especially how it was researched, built, evaluated, and solved. This chapter also discusses its new findings.

Chapter 9 concludes this dissertation and summarizes its processes, including the motivation to propose its topic that focused on the emotional healthcare system, its goals, design, implementation, and evaluation of the necessary applications in it as well as the proposed system’s prototype. Finally, I explain the dissertation results, its benefits, and future work.

Chapter 2

Literature Review

2.1 Healthcare systems focusing on emotional aspects

Healthcare systems in HCI are applications or systems which integrate several technologies to provide assistants, services and treatment to maintain users in both physical and mental health. Several researches about healthcare systems focus on physical aspect. Below is a review of the researches that focuses on emotional aspect.

Yuanchao et al. proposed MoodMiner, which is a mood assessment system [30] that uses phone-based sensing techniques to sense mobile phone sensors and communication activities the daily moods of users. The following are the mobile phone sensors:

- An accelerometer measures acceleration on three axes.
- A microphone senses the sounds from the environment and the user’s voice

- GPS detects the user’s location.
- An ambient light sensor detects the environment’s brightness and identifies the phone’s position such as in bag or in a pocket.

They also developed an application based on an Android platform to collect sensor data and communication logs to send such information to the server side which collects all the user data and creates a personalized model for preliminary analysis. This application shows user behavior related to the mobile phones, such as the frequency of sending messages, distance traveled, and so on. This application also shows the daily assessments of the mood of patients.

Nakajima et al. developed a relax/refresh system that applies virtual reality technology [31]. Their system consists of a massage lounger, a head-mounted display (HMD), a standard video cassette recording (VCR), and an interface circuit that controls the massage chair and generates stereographic images. This system provides visual, aural and physical simulations that promote three stages: encouraging sleeping, and refreshing.

- In the encouraging sleep stage, this system provides relaxing images (2D), music with body vibrations and soft massages to users.
- When the user falls asleep (sleeping stage), this system dims the user’s view, lower the music’s volume, and stops the massages.
- When the user wakes up (refresh stage), the system provides refreshing images (3D) with loud music and strong massages.

This system relaxes and refreshes users and reduces their stress. They also performed the experiment to measure electrocardiogram data and calculate the activity level of the sympathetic nervous system and the parasympathetic nervous system to indicate relaxed and refreshing state. In the experiment, the participant performed some kind of specified visual display terminal (VDT) work for 40 minutes three times. After they finished performing one VDT work, they took a rest for 10 minutes by using VR relax/refresh system, sitting, or continue to perform VDT work. The results indicated that the VR relax and refresh system is effective for relaxation while using it and for refreshment after using it.

Zhang Q. and research group [32] proposed e-phychotherapy. They developed a platform which consists of Internet Psychological Counseling (IPC) and Computer-aided

Counseling System (CACS). IPC has online and offline counseling methods. The online counseling method includes four basic operations: diagnosis, psychological, training, appointment and psychotherapy. These operations can operate between the group of users and the therapists, or between the individual and the therapist. The offline counseling method can support the users via forum and SMS that the users can communicate and consult with others who have same problems by using this method. Another module is CACS which is the computer assisted system for e-therapy. This system manages the information and provides multi-media materials and tool for counseling process. Moreover, it provides the self-help treatment module that the users are guided by the programs with or without therapists in order to improve their mental health and solve their own problems. For the self-help treatment, they applied the computer-aided cognitive behavior therapy (CCBT). The cognitive behavior therapy uses Zen and Buddhist philosophy for the treatment process.

From my literature review, a few researches have focused on user emotions compared with researches that focus on physical aspects. Many researches which emphasize user emotions support psychological, phobic, anxiety, and stress disorders of patients. These researches provide services to treat users and improve their emotional health. However, the emotional health of people should not be ignored. Some people become psychological patients because they cannot cope with negative emotions in their daily life. Therefore, designing a system to cope with the negative emotions of users in daily life is very promising. Thus, I propose a new healthcare system that detects user emotions and offers support when they are suffering from negative emotions to improve their positive emotional states.

2.2 Relaxation service

Relaxation is the most important for building positive attitude and good health in daily life. Several researchers have designed and implemented relaxation services and applications to release stress. Below are some examples of relaxation services.

Lee et al. designed stress evaluation and personal relaxation system [33]. Stress evaluation consists of two parts: hardware and software. Hardware part is integrated PPG sensor and date acquisition. Software part is developed using LabVIEW to detect R-Peak, calculate R-R interval in order to analyze personal stress using LF/HF ratio of heart rate variability (HRV). Personal preference relaxation system is implemented to relax users from visual, auditory and olfactory stimulus such as natural scenes video, natural sounds, lemon and lavender of essential oil.

Pioggia et al. proposed Interreality, which is a technology-based approach to assess and treat stress [11]. This approach, which is based on cognitive behavioral therapy (CBT), provides a 3D virtual world to assist patients who are suffering from stress and stress-related disorders. It consists of the following three virtual islands:

- The learning island teaches patients how to relax and improve their stress management skills.
- The community island provides real-life examples to help patients understand themselves and, reduce avoidance behaviors and unrealistic thinking. Patients are encouraged to discuss and share their experiences.
- The experience island practices controlled exposure, emotional management, general decision making, and problem solving skills.

Interreality also applies bio and activity sensors to provide patients behaviors and health state to therapists so that they can monitor their patients. With this system, therapists can treat and assess their patients.

Many relaxation services apply visual, audio or virtual reality stimulus to release users’ stress. To improve the effectiveness and attractiveness of existing relaxation services, the source to display the content is very important. Therefore, applying a new technology to improve relaxation service is very promising. Thus, I apply augmented reality (AR) which is technology that combines virtual world and real world to produce a new environment. AR enhances the interface by displaying two- or three-dimension computer graphics over real objects [34]. In general, AR is used with head-mounted display, computer and smart phone. AR enables the interaction between users in real world and virtual objects in virtual world [34-36]. Nowadays, various AR applications have been developed in the field of Game, Education, Entertainment, Navigation, Industrial Design, and Medication [34-35]. However, there is no relaxation service that applies AR with treatment content to relax users.

2.3 Emotion recognition in HCI

Emotion recognition or emotion detection is essential and useful in human computer and human robot interaction because emotions indicate human feelings and needs [6]. Emotion recognition in HCI enables the compute and its application to interact with human in natural way. Furthermore, the computer and its application can analyze and interpret user emotion to provide appropriate responses.

The following are the three main categories of techniques that have been proposed to recognize and classify user emotions:

- Speech emotion recognition: identifies emotional states using voice analysis [36-39]. These speech features such as pitch, formants, and short-term energy [28] are useful for emotion recognition. Thus, several feature extraction methods have been proposed to extract these features. These features, however, are very distinctive among different people and finding a standard set of numbers to correctly identifying emotions is very difficult [28]. Speech is moreover language specific. While capturing emotional state through the controlled experiment may be possible, implementing it in a real-time system for healthcare is still not practical. Moreover, the main limitation of this approach is that speech is necessary to recognize user emotion. Therefore, if users do not speak, his/her emotion cannot be determined.
- Emotion recognition from facial expressions: recognizes human emotion from facial muscle movement, the movement of eyes, mouths or eyebrows movement and facial texture. Several studies have applied computer vision systems to automatically analyze and recognize changes in facial motion from visual information [40-42]. This approach has some limitations that users are required to animate their emotions with their face, and a camera must capture the users' frontal view in order to detect and recognize user emotions. However, facial expression is the common and easiest way to express emotions for human. In addition, it is most visible modality to perceive emotions among humans around the world for natural communication in various social situations [43]. Furthermore, facial expression is more useful for detecting emotions from speech disorders as well as diagnosing early state of psychological disorders by recognizing an inability to express certain

emotional facial expressions [44].

- Emotion recognition using biological signals: analyzes and recognizes human emotional state from biological signals such as electroencephalography (EEG), electrocardiography (ECG), temperature, and galvanic skin response (GSR). This approach can recognize users' emotion when the users are silent or do not show their emotion on their face or outside appearance. The use of biological signals is more suitable for laboratory settings. Even though the techniques can provide generally quite accurate results, the limitation of the adoption comes from the fact that the sensors can be expensive and they must be attached to the body [45-50].

2.4 Emotion recognition by facial expression

This section reviews two main classes [51] of facial feature extraction, namely, Geometric-based and Appearance-based methods. In both of the approaches, computer vision and image processing are applied to extract facial features. Then, the expressions are recognized through classification techniques.

2.4.1 Geometric-based approaches

Geometric-based approaches rely on the geometry of the human face. The techniques extract shapes, points and locations of facial components, including eyes, eyebrows, mouth and nose. Once the features are extracted, the distances between them are computed to form feature vectors [52]. In addition, several geometric-based approaches apply model-based techniques to detect and locate facial points [23, 24, and 53]. The model-based techniques [54] consider shape, appearance or pose to construct 2D or 3D models in order to extract necessary features. These techniques can produce high accuracy results because they extract both facial points and appearance features. However, geometric-based approaches and model-based techniques require high resolution sources and are generally time consuming. The approaches are thus not suitable for a real-time system [52]. Examples of geometric-based approaches with model-based techniques can be listed as follows:

- Active Shape Model (ASM)

  ASM [53] uses the landmark points that are determined from the training data. The points are then transformed using the principle component analysis (PCA) technique to locate the main variations of the training data. These variations are later used to construct a shape model. ASM, in addition, constructs a gray-level appearance model obtained from pixels around the landmark points and perpendicular to the contour. Finally, both models are used to formulate the facial features.
- Active Appearance Model (AAM)

  AAM [23-24] analyzes shapes and textures of facial images. It combines shape and texture variations together to create a model by fitting shape and appearance components using PCA. It requires face images with facial points that indicate the necessary regions, including facial region, eyes, eyebrows, mouth and nose. AAM requires the texture at each point to build a facial model, which is used to extract the necessary features.

2.4.2 Appearance-based approaches

Appearance-based approaches represent facial textures by extracting the change in face appearance and skin texture. Examples of appearance-based approaches include:

- Gabor wavelet filters

  Gabor wavelet filter [22] considers optimal localization features in both spatial and frequency domains. Its important parameters are orientation and frequency of patterns. Gabor filter applies Gaussian functions to modulate the amplitudes of sinusoids. Feature vectors can then be extracted using Gabor filter with different frequencies and orientations. The filter has been widely applied in many pattern recognition applications.
- Curvelet transform

  Curvelet transform [42] has been adopted mostly in texture classification and image enhancement. Curvelet transform is different from wavelet transform because it considers not only points, but also curves and lines in an image in order to extract the feature vectors. The feature vectors are constructed from curvelet coefficients by performing curvelet filter with different angles and scales.

- Local Binary Patterns (LBP)

  Local binary pattern (LBP) [25-26] is a widely used texture description method in pattern classification. LBP operator encodes the information of curves, edges, spots, and other local features as binary numbers by processing 3×3 neighborhood pixel mask with the center value. In LBP, the image is divided into 7×6 rectangular regions, each of which corresponds to a 256-level histogram. All resulting histograms are then concatenated to create a global profile of the entire face as a feature vector.

- Local Directional Patterns (LDP) [55]

  Instead of encoding the intensity of each image pixel, local directional pattern (LDP) [56] operator applies Kirsch eight-directional edge detector to compute the edges and then encodes the texture information by considering the edge responses. For images with noise or non-monotonic illumination changes, LDP is more robust than LBP because LDP uses edge responses that are more stable than intensity values when generating the binary patterns.

- Directional Ternary Patterns (DTP)

  Directional ternary patterns (DTP) [27] operator applies an edge detector to compute the edge responses in eight directions similar to LDP. DTP instead applies Robinson eight-directional edge detection rather than the Kirsch detector. Another difference is that DTP operator uses a three-value code (-1, 0, 1) instead of a two-value code (0 or 1). As a result, the smooth and high edge responses can be differentiated and a threshold value can be used to form positive and negative binary patterns.

Due to the complexity of the geometry-based approaches, the appearance-based approaches are more suitable in real-time environment. Among them, DTP has highest accuracy because the negative different values between edge responses are considered and the noise sensitivity is lowered than those of LBP and LDP [27]. In addition, the encoding of the edge responses is more stable than that of the intensity values in obtaining the correct binary patterns from noisy images [25, 56, and 27]. DTP is also more accurate than the Gabor wavelet and Curvelet filters in different image resolutions [57]. I then chose to work on an improved version of DTP to fit the needs of my real-time facial emotion recognition system.

2.5 Emotion recognition using ECG signal

An ECG measures the heart's electrical activity over a period of time. Several techniques using ECG have been proposed to recognize human emotions.

- Agrafioti et al. [58] analyzed ECG patterns for emotion recognition and proposed ECG signal decomposition techniques using a bivariate extension of Empirical Mode Decomposition (BEMD), which extracts the ECG features from the instantaneous frequency and local oscillation of Intrinsic Mode Functions (IMF) for recognizing the emotions of individual subject (subject-dependent). They conducted two experiments to evaluate how well their system detected positive and negative emotions as well as mental stress. They also analyzed and classified the valence, arousal, passive and active emotions that are useful for various kinds of emotions detection with ECG.
- Rattanyu and Mizukawa recognized and classified six emotions: anger, fear, disgust, sadness, neutral and joy [48]. They applied a diagnosis method that uses both inter-beat (HR and RR-interval) and within-beat (PR-interval, QRS-interval, ST-interval, QT interval, PR-segment, and ST-segment) for feature extraction and calculated five types of statistical data: maximum, minimum, medium, mean, and standard deviation. They applied the K Nearest Neighbor (k-NN) and Linear Discriminant Analysis (LDA) methods for emotion classification. The experiment result showed that the eleven features approach (inter-beat and within-beat) had higher classification accuracy than the traditional three features approach (only inter-beat).

Many researches have proposed feature extraction approaches using signal processing or such medical diagnosis techniques as heart rate variability or PQRST interval with statistical methods to recognize emotions from ECG [59-62]. Signal processing techniques have been adopted to extract facial features for emotion recognition by facial expressions. Therefore, some facial feature extraction techniques might be able to extract ECG features. Additionally, I found that one popular facial feature extraction technique: local pattern description has not been applied to extract ECG features yet. Since, this method can also extract facial features with high accuracy. I adopted local binary and ternary patterns (LBP and LTP) [25, 63] to extract the ECG features to recognize emotions. These methods will be described in Chapter 6.

2.6 Emotion recognition using speech, EEG and GSR

In this section, I survey feature extraction techniques of speech, EEG and GSR to recognize users' emotional state

2.6.1 Speech

Several researches proposed various techniques to extract speech feature to recognize emotions. For example:

- Utane and Nalbalwar recognized five basic emotions: angry, happy, sad, surprise and neutral by extracting both prosodic and spectral features such as pitch, energy, Mel frequency cepstrum coefficient (MFCC) from 20ms-30ms frame-lengths [64]. Speech features were then classified using Gaussian mixture model (GMM) and Support vector machine (SVM). The accuracies were around 70% for both GMM and SVM. GMM was better than SVM to classify angry and sad speech. However, SVM was more accurate to classify neutral speech.
- Wu et al. [65] proposed modular spectral features (MSFs) to extract emotions from speech. In their experiment, MSFs, two short-term spectral features (MFCC and PLP) and 75 prosodic features such as pitch, intensity, delta-pitch, delta-intensity, zero-crossing rate were extracted from Berlin emotional speech database and Vera am Mittag (VAM) database. Fisher discriminant ratio (FDR) was applied to eliminate irrelevant features. Only necessary features were classified using SVM with radial basis function (RBF) kernel. In the first step, they compared their MSF with MFCC and perceptual linear predictive (PLP) coefficients. The results

indicated that MSF with Linear discriminant analysis (LDA) produced 85.6% of accuracy when recognizing seven emotions: anger, boredom, disgust, fear, joy, neutral and sadness. Then, MSF, MFCC and PLP features were combined with prosodic features to recognize emotions. The evaluation results showed that MSF produce higher accuracy than MFCC and PLP. In conclusion, the highest accuracy was 91.6% which produced by extracting MSF and prosodic features with LDA and classifying using SVM.

2.6.2 Electroencephalography (EEG)

EEG is brain wave activity recording from scalp. This technique measures and performs emotional feature extraction based on EEG. Several feature extraction and emotional classification method have been proposed. For example:

- Murugappan and his research group used Discrete Wavelet Transform (DWT) for EEG feature extraction [46]. For classifying emotion, they performed two simple classification methods: k-NN and LDA methods. They classified five emotions: disgust, happy, surprise, fear and neutral. The experiment result showed that the proposed feature extraction method based on DWT achieved the maximum average classification rate of 83.26% using k-NN and 75.21% using LDA compared to conventional features (Power, Std Dev, and Varinace).
- Petrantonakis and Hadjileontiadis employed higher order crossing (HOC) for feature extraction and emotion classification method, called HOC-emotion classifier (HOC-EC) [66]. They classified six basic emotions: happiness, surprise, anger, fear, disgust and sadness. The experiment result showed that HOC-EC achieved a 62.3% using quadratic discriminant analysis (QDA) and 83.33% using support vector machines (SVMs) compared to other feature extraction methods (Statistical-Based Features and Wavelet-Based Features).

2.6.3 Galvanic Skin Response (GSR)

GSR measures electrical conductivity of the skin. The changes in the electrical properties of the skin is due to the activity of sweat glands that caused by emotional stimulus. The sweat glands are controlled by the sympathetic nervous system. Thus GSR is an appropriate technique to indicate physiological arousal. Many researches also applied GSR to measure user arousal level to recognize user emotion. For example:

- Kim et al. proposed emotional recognition system using short-term monitoring of physiological signals [47]. Their system integrates various sensors to measure electrocardiogram, skin temperature variation and electro dermal activity for classifying emotions. The Electro dermal activity (EDA) is used to measure galvanic skin response or skin conductivity in order to indicate arousal level of user.
- Nakasone et al. have researched in emotion recognition [67]. They integrated GSR approach to recognize arousal level of human. They classified four emotional states in high arousal level: fear, frustrated, excited, and joyful. The emotional states in low arousal level are sad and relaxed.

2.7 Comparison between three approaches in emotion recognition

As mentioned in section 2.3 that there are three approaches to recognize emotions in HCI: by facial expression, by speech and by biological signals. Although, these three approaches have some limitations, they are state-of-the-art and popular techniques in which several researches have been proposed to recognize emotions.

Emotion recognition by facial expressions and speech are suitable to be used in everyday life applications because facial expression and speech are natural user interfaces in which users can interact to applications in such a natural way. Moreover, any sensor is not required to attach on users' body that is different from the emotion recognition by biological signal. For cross-cultural applications, facial expression is more appropriate than speech because languages are different among cultures that affects to voice quality of utterance speech. Hence, emotion recognition by speech is difficult to define languages-independent features that can produce high performance and accuracy when considering the same emotional features with different languages. Whereas emotion recognition by facial expressions can provide stable performance and high accuracy when apply similar facial features using facial expressions from different nationality subjects [25]. Therefore, the adoption of facial emotion recognition is much simpler and more practical to integrate with living and working spaces. In addition, latest development in imaging technology has made the availability and the cost of imaging devices to be affordable. Devices like webcam and CCTV can nowadays be found in normal settings. This makes it possible to create such an integrated intelligent working and living space. Moreover, facial expressions are the most visible modality [68] and are, to a certain extent, universal despite our idiosyncratic and cultural differences [43-44]. Unlike speech, facial data can be obtained during all kinds of activities.

However, in some systems such as healthcare system, using only emotion recognition by facial expression might not be enough. If it is necessary to recognize users' emotions for treatment but they don't express emotions on their facial expression, integrating emotion recognition by facial expression with biological signal is more appropriate than using only facial expression. Moreover, biological signal can be utilized to measure other healthcare information.

2.8 Healthcare system with emotion recognition

Emotion recognition in healthcare system is usually designed to analyze emotional state of patients or elderly for reporting to caregivers, families or doctors in order to provide any assistance. Below are the examples of healthcare system with emotion recognition.

Lisetti and research group developed multimodal intelligent affective interfaces for tele-home healthcare applications [69]. They integrated emotion recognition for interpreting patients' emotional state which is important to the caregiver in order to provide better assistance. They built the model of patients' emotions (happy, sad, frustrated, angry and afraid) from speech, text, facial expressions, body temperature, galvanic skin response and heart rate which is useful for clinicians to understand patients' emotions.

Jiang et al. created a wearable home healthcare system with emotion recognition for the elderly [70]. Their emotion recognition based on physiological signals such as heartrate, body temperature and skin conductance. When the physiological signals were abnormal, the system sent alarm message via SMS to family or doctor of the elderly.

From the above literature survey, I realize that

- Many researches on e-healthcare have focused on physical aspect of ordinary people. Thus, emotional healthcare system should be proposed.
- The researches on relaxation service have provided therapy or treatment with visual, audio or virtual reality. No relaxation service applied the augmented reality with therapy or treatment before. Thus, relaxation service with augmented reality should be proposed.
- Emotion recognition in healthcare system is usually applied to assess patients' emotional state for clinicians, therapists or doctors. However, in the other fields such as game, education or automobile, the emotion recognition is usually applied to assess the emotional state of ordinary people for automatic feedback control system. Thus, emotion recognition in the healthcare system should be applied for automatic providing appropriate feedback or service to users.

Therefore, I propose a new emotional healthcare system which integrates the emotion recognition to recognize users' emotional state to provide appropriate emotional services with augmented reality as a feedback in order to cope with users' negative feelings. This is a novelty of this dissertation.

Chapter 3

Design of healthcare system focusing on emotional aspects

This chapter presents the design of the proposed emotional healthcare system. This system aims to support ordinary people at workplace in daily life in order to cope with negative emotions and improve their emotional health. The overall system and framework designs are described as below.

3.1 Overall system design

I design my scheme as a web-based system to support users in workplaces by easily accessing the system using the web browsers on personal computers, tablets, or smartphones (Figure 3.1). While users are working on personal computers, tablets, or smartphones, their emotions are recognized and classified using facial expressions, speech and biological signals which are detected by webcam, microphone and biological sensors respectively. When they have negative emotions, the system suggests that they should take a break by providing one of three services (relaxation, amusement and excitement services) based on their type of negative emotions to decrease it and improve positive emotions. For example, if user is sad the system will provide amusement service. After a service is provided, they can select which application they would like to use. To activate the application such as breathing control application in relaxation service, AR marker is employed in front of the camera to display a virtual object whose purpose is to encourage such positive emotions as relaxing, enjoying, and excitement. Additionally, for some applications, they need to interact with virtual objects using Kinect, which detects their gesture. I believe that interactions with virtual objects may decrease negative emotions effectively.

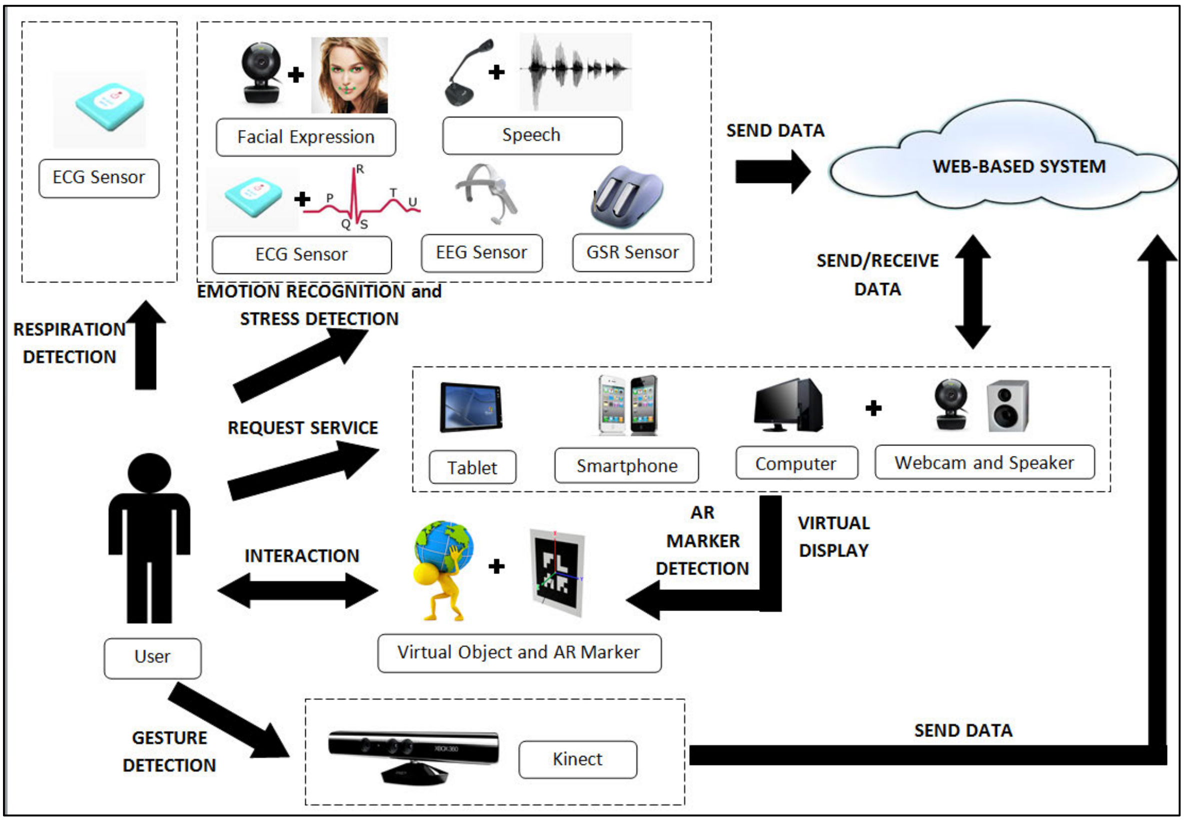


Figure 3.1 Overall system design [71-72]

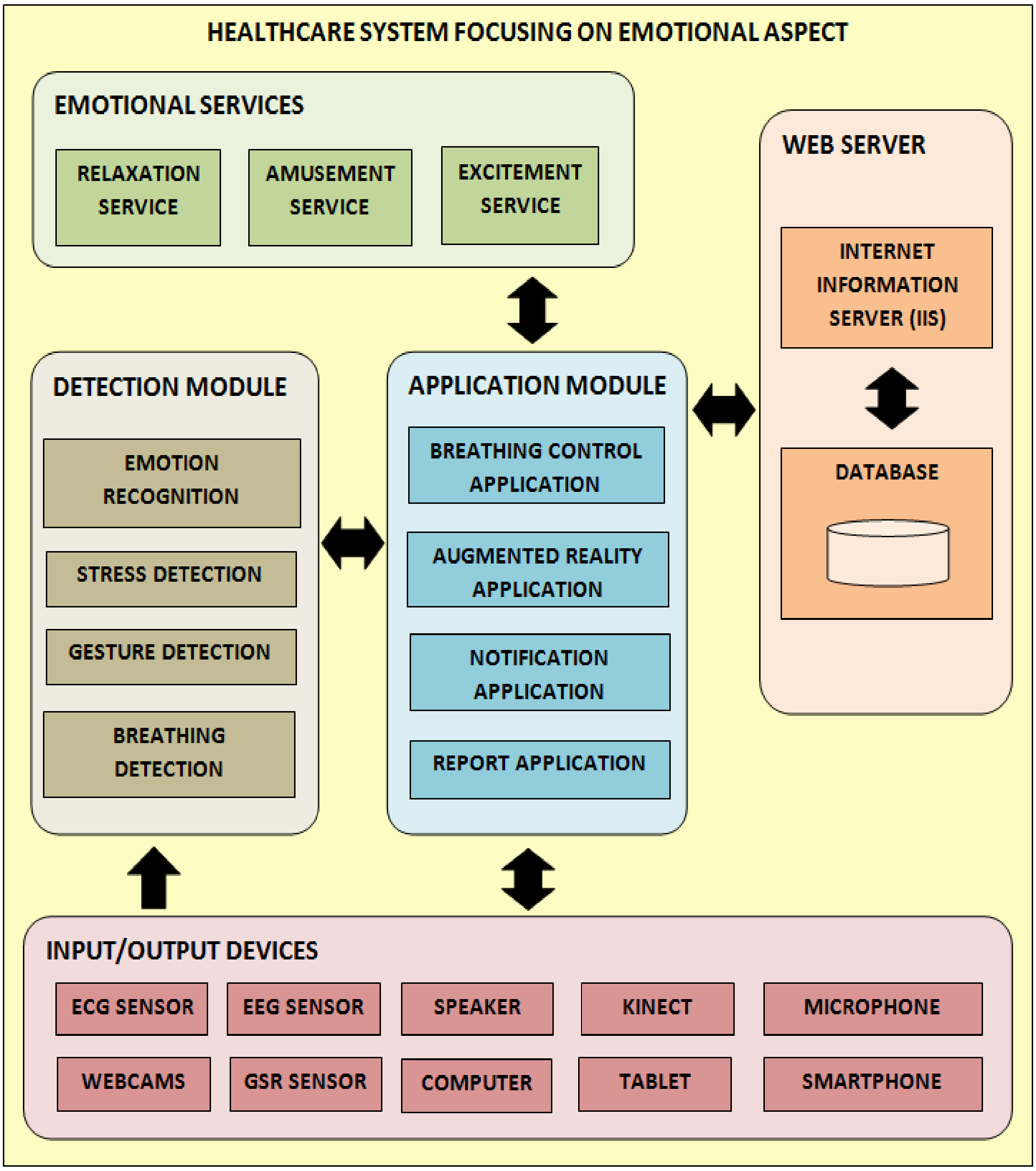


Figure 3.2 Framework design

3.2 Framework design

The framework design of this system is shown in Figure 3.2. This healthcare system consists of six parts: I/O devices, emotional services, a detection module, an application module, a database and a web server as described below.

3.2.1 I/O devices

There are eleven I/O devices: a personal computer, webcams, a speaker, a microphone, a Kinect, a tablet, a smartphone, an EEG sensor, an ECG sensor and a GSR sensor.

- The personal computer, tablet and smartphone are used to access the system via web browsers.
- The webcams are used for detecting AR marker to build a virtual object and for recognizing user's emotion from facial expressions.
- The microphone is used for detecting emotions from speech.
- The EEG sensor which measures voltage fluctuations from electric ions within the brain's neurons can be used for recognizing emotions (Rattanyu & Mizukawa , 2011).
- The ECG sensor measures the heart's electrical activity over period of time [48]. I apply RF-ECG sensor (Figure 3.1) which is a small and light-weight sensor (40*35*7.2 mm) to record the ECG signal with 204 Hz of sampling rate. ECG signals can be interpreted as the heart rate in beats per minute (BMP) and can be analyzed using heart rate variability (HRV) technique to detect stress as well as emotions. Moreover, the ECG signal can be converted to be respiration signal. Therefore, I applies ECG sensor to detect users' heart rate, respiration signal and emotions.
- The GSR sensor measures galvanic skin response which can be applied to recognize stress and active-passive emotions [48].
- The Kinect analyzes and detects such users' gesture as hands, fingers and legs.
- The speaker is used for playing music and notification.

3.2.2 Emotional services

In this system, emotional services are the services that create and provide any information to induce users' expected emotions in positive ways. This system is designed to provide three essential services: relaxation service, amusement service, and excitement service in order to encourage positive emotions and decrease negative emotions while they are working to improve their well-being in our society.

- *Relaxation service*

Relaxation is the most important for building positive attitude and good health in daily life because it can address any negative emotions and stress that are the causes of any diseases, since stress and negative emotions stimulate high heart rate, blood pressures, breathing rate, etc. When people know how to relax and when they experiences with relaxation, they can better sleep and get more energy to improve their positive attitude, and ability to deal with any problems as well as negative emotions and stress [73]. Finally, they can work effectively.

This service aims to make users feel more relaxed by providing applications with stress management techniques (Figure 3.3) such as

- o Virtualizing in relaxed scenes e.g. beach, sky, meadow, etc.
- o Massaging, meditation, and deep breathing with relaxed music

Figure 3.3 Examples of relaxation service [16-17, 74-75]

However, providing only relaxation service might not be enough to support all people because some of them might need to experience other emotions to make them feel better such as fun or exciting. Therefore, this system provides two more services to increase and encouraged users' amusement and excitement.

- *Amusement service*

This service aims to make users feel more fun by providing applications with augmented reality (Figure 3.4) such as

- o Playing with virtual pets such as cat, dog, rabbit, etc.
- o Finding the hidden things

Figure 3.4 Examples of amusement service [76-78]

- *Excitement service*

This service aims to make users feel more exciting (Figure 3.5) such as

  - Touching bubble or opening the door, then excited things happen with sound.

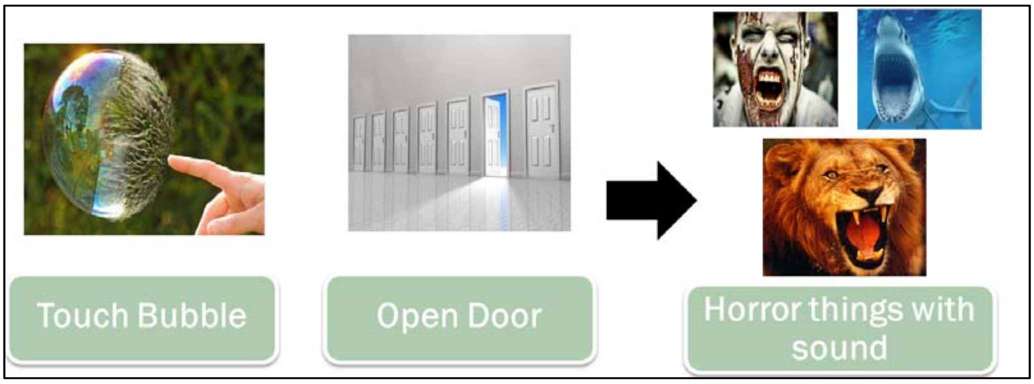


Figure 3.5 Examples of excitement service [79-83]

3.2.3 Detection module

Detection module contains several detection applications to analyze, activate and support emotional services. There are four detection applications:

- *Emotion recognition* measures, analyzes and recognizes user emotions from facial expressions, speech and biological signal. Six basic emotions (joy, surprise, anger, sadness, disgust and fear) together with neutral state can be recognized.
- *Stress detection* analyzes and recognizes user stress from biological signal such as ECG signal.
- *Gesture detection* detects the activities of users' hands, legs, or fingers such as point,

touch, hit, or kick using Kinect to allow an interaction between users and virtual objects in real environments.

- *Breathing detection* converts ECG signal to respiration signal in order to detect and measure user respiratory rate from the ECG sensor.

3.2.4 Application module

The application module consists of various kinds of applications that is designed to support and stimulate expected emotions to users. Currently, there are four applications in the application module:

- *The augmented reality application* detects and analyzes AR markers to display related 3D virtual models.
- *The breathing control application* applies deep breathing technique of stress management to decrease user negative emotions and stress by allowing users to perform deep breathing with various kinds of virtual music boxes.
- *The notification application* suggests users to take a break by providing appropriate services related to their current emotional state.
- While users access the healthcare system, *the report application* displays such biological information as heart rates, and current emotional states using emoticons. Furthermore, this health information are recorded and reported to users' colleagues, their leader, their family members, and their doctors.

3.2.5 Database

I create a database to collect information for each service such as the users' personal information, their emotion status logs based on facial expression, biological signals and so on. This collected information will improve the system reliability. Example of tables in database as is shown below.

- *User Information Table* stores users' personal information such as system users, their doctor and their family members, their colleges, their leader and so on.
- *Heart Rate Table* stores information about users' heart rate every one minute.

- *Time Breath Table* stores respiration rate for breathing detection.
- *Emotion Table* stores users' emotional state from each facial expressions, biological signal or speech.
- *Notification Table* stores notification list about notification time, suggestion service in order to collect notification log.
- *Relaxation Service Table* stores information log when users use relaxation service, such as the emotion before and after using service, the duration when using service and so on.
- *Amusement Service Table* stores information log when users use amusement service, such as the emotion before and after using service, the duration when using service and so on.
- *Excitement Service Table* stores information log when users use excitement service, such as the emotion before and after using service, the duration when using service and so on.
- *Report Table* stores statistical information for daily such as users' daily emotional state, how often they use each service, etc.

3.2.6 Web server

This system is designed to use an internet information server (IIS) as a web application server. IIS is a web server application and set of feature extension modules created by Microsoft for using with Microsoft Windows [84]. It supports HTTP, HTTPS, FTP, FTPS, SMTP and NNTP.

3.3 The advantages of this emotional healthcare system

Several emotional healthcare systems using virtual reality have been proposed to diagnose, examine and treat users (patients) with some mental health problem such as psychological, phobia, anxiety and stress disorders [11-14]. These systems also integrated biological sensors to monitor patients' behaviors and their health states. They build virtual environment to help addressing patients' mental help problem. However, I found that only few systems have been designed for supporting ordinary people, and no system using augmented reality to address and improve mental/emotional health. Therefore, my system integrates augmented reality to induce positive emotions and decrease negative emotions of ordinary people at workplace in daily life. The advantage of augmented reality is it combines virtual world with real world that users can experience virtual objects in real world. This might help increasing their positive emotion better than using only virtual reality because augmented reality is better to enable the interaction between users and virtual objects in real environment as I described in Chapter 1.

Moreover, I found that real-time emotion recognition in the healthcare systems for providing feedback is not common comparing with other fields such as robot, entertainment and automobile. Therefore, this system also integrates the emotion recognition to analyze and recognize users' emotional state in short-time duration for providing assistance when users feel bad.

The advantages of this system are as follows:

- Real-time detection of users' current emotions and stress from facial expressions, speech together with biological signals.
- Utilizing augmented reality in order to display virtual objects to improve their emotional state.
- Applying such mental health technique as deep breathing techniques of stress management to implement new breathing detection algorithm to detect and control users' respiration.
- Users can access the web-based system anywhere and anytime using personal computer, smartphone and tablet.

3.4 Example when user severely suffers negative emotions and stress from works

Alice is an accountant; she needs to deal with several documents and receipts about budgets for working, meeting and so on. Every day, she checks the documents and uses Microsoft excel to write the report about transaction from budget, fee, charge, etc. Sometime, she needs to discuss with her colleagues about some unclear documents that make she gains a lot of stress. She might undergo troubles from her works, colleagues, leaders and customers that make she has negative emotions. During she works in front of her personal computer, the emotional healthcare system recognizes that she has negative emotions from her facial expressions and ECG signal. Then, one of three services are served based on her emotional state e.g. relaxation service. The relaxation service encourages her relaxation by allowing hers to perform deep breathing for five minutes. After that, the system will recognize her emotional state again. If she feels better, the service will be stops. However, she can continue using the service as she likes.

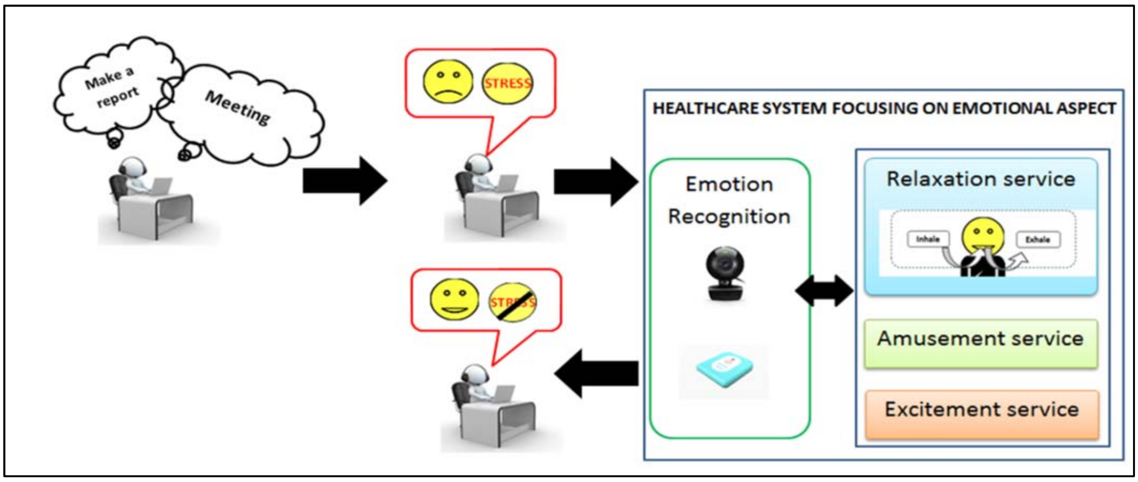


Figure 3.6 Example when using healthcare system focusing on emotional aspects [85]

3.5 Discussion

The emotional healthcare system aims to decrease negative emotions or stress because not all people can cope with these emotions and this can lead to social and mental health problems, since one preliminary reason of these problems is negative emotions and stress in daily life as I mentioned in Chapter 1. This system is designed to integrate many sensors, applications and services to support users. Although, integration of biological sensors required users to attach on their body such as ECG sensor, but these sensors might be designed as wearable accessories in the near future. Therefore adapting them to use in daily life is possible. With biological sensors, augmented reality, emotion recognition, and stress detection, this system is designed to be more effective, attractive and intelligent to improve users' emotional state when they feel bad.

3.6 Summary

This chapter explains system design in overall and explains each device and application in framework designs. This system aims to support normal people when they experience with negative emotion or stress in daily life. It is designed as a web-based system that is easy to access anywhere and anytime. To improve system's efficiency and intelligence, eleven I/O devices and various kinds of applications are integrated to the system. The I/O devices feed or display useful information from/to users. The applications apply several techniques to analyze and recognize their emotional status. Furthermore, three services (relaxation, amusement, and excitement services) are designed to support users in daily life. To make these services more effective and more attractive, augmented reality is applied to let users interact with virtual objects in real environments to get different positive emotions and decrease their negative emotions.

Chapter 4

Relaxation Service

The relaxation service aims to make users feel more relaxed by applying augmented reality (AR) with stress management techniques. To support relaxation service, I design one application: breathing control application. This chapter presents the design, implementation and evaluation of breathing control application in relaxation service. From the evaluation results, the breathing control application is improved to fit the users' requirements.

4.1 Basic design

In psychology field, the deep breathing technique of stress management is one technique that effectively increase relaxation, and reduce stress as well as negative emotions by let users control their respiration [86]. In general, people with stress always quickly and shallowly breathe [87]. When they control their respiration to get rid of the stress or negative emotions, their shallow breath will be converged to regular breath and they will feel more relaxed [86-88]. Therefore, the breathing control application applies deep breathing technique of stress management to increase relaxation. This application supports users when they experience stress or any negative emotions from society, work and so on. The application also includes a virtual music box to assist such deep breathing. Virtual objects and music can help users quickly and easily feel more relaxed. The workflow of this application (Figure 4.1) is as below.

1. Emotional healthcare system suggests users by providing the relaxation service with the breathing control application.
2. Users can select the breathing control application in the relaxation service by web browsers on a tablet, a smartphone or a personal computer.
3. A request is sent to the web-based healthcare system.
4. Users can start using this application by showing an AR marker by camera.
5. The application detects the AR marker.
6. The application displays a virtual music box, which slowly turns and plays music.
7. User inhales and exhales in harmony with the turning of the music box to control his breathing.
8. While user is controlling his breathing, the ECG sensor on his chest records ECG signals.
9. ECG data is sent to system and convert it to respiration signal in order to detect user’s respiration.
10. If their breathing isn’t in harmony with the turning of virtual music box and their breath rate isn’t suitable for deep breathing, the application suggests that user breathes more slowly or quickly. The application continues to support users until they feel more relaxed.

4.2 Design detail and implementation

4.2.1 Breathing control application

The breathing control application is shown in Figure 4.2. There are four main components as below.

- *Animation guide (a):* guides users how long they should inhale and exhale.
- *Application suggestion (b):* gives advice so that users breathe more deeply or shortly.
- *Music (c):* provides three types of classical music that users can choose using arrows.
- *3D virtual music box (d):* slowly turns around and play music.

Regarding the workflow of breathing control application, it needs to send request to augmented reality application in order to detect AR marker and display a virtual music box. Moreover, it also sent request to breathing detection to convert ECG signal to respiratory signal in order to detect inhalation and exhalation.

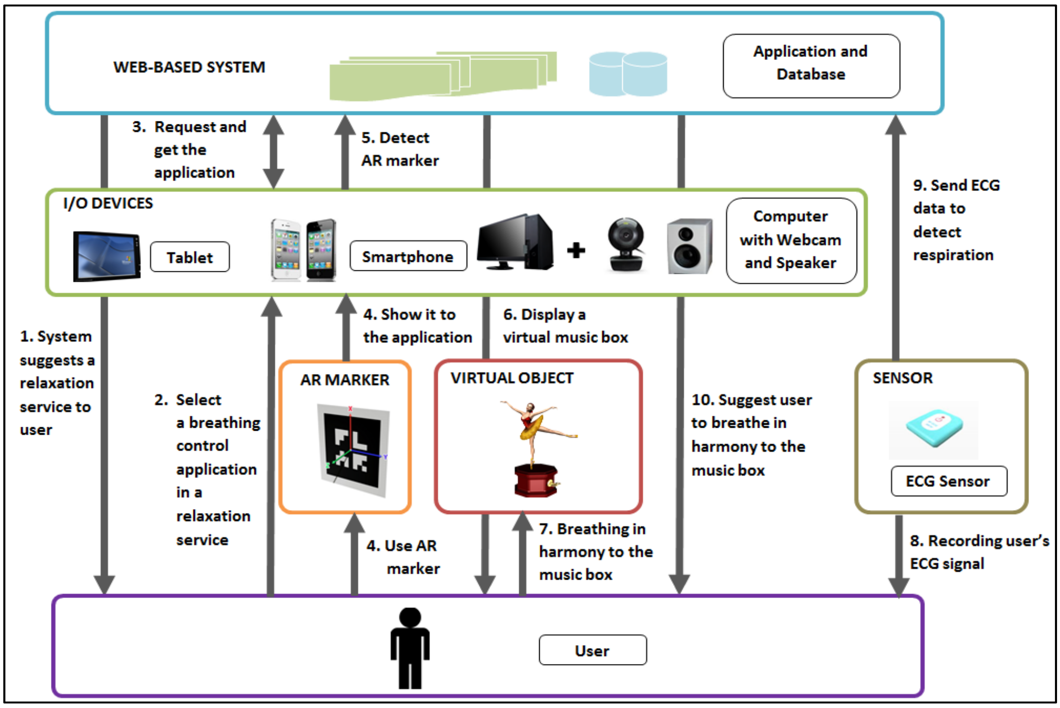


Figure 4.1 Workflow of breathing control application [89-90]

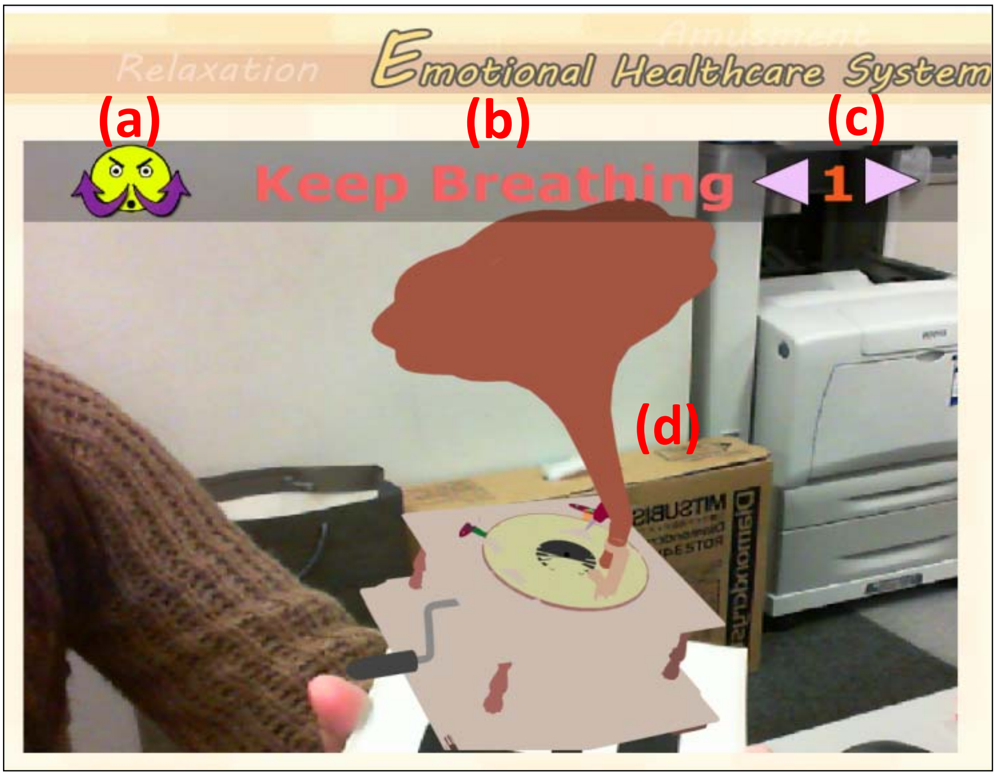


Figure 4.2 User interface of breathing control application

4.2.2 Augmented reality application

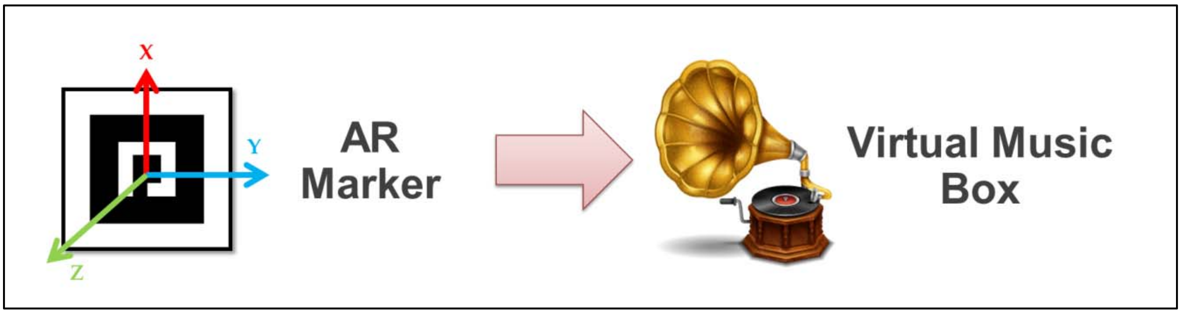


Figure 4.3 Process of augmented reality application [90-91]

This application is implemented based on FLARManager which is a lightweight framework that creates and builds augmented reality applications for Flash more easily [19]. According to #4 and #5 of workflow, with this application, the breathing control application can detects AR marker to render and display a virtual music box [92] with sound for users as shown in Figure 4.3.

4.2.3 Breathing detection

This breathing detection utilizes ECG sensor to analyze and convert ECG signal for detecting respiration signal using ECG signal processing that is called ECG-derived respiration (EDR) as shown in Figure 4.4.

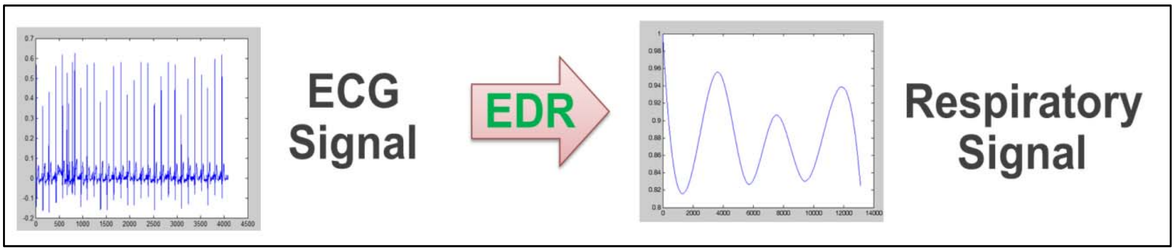


Figure 4.4 Process of breathing detection

I employ the RR amplitude technique to derive respiratory signal from ECG signal using the estimation of the R-wave amplitude modulation [93]. This method is reliable and can produce most accurate respiratory signal when comparing with other methods [94-96]. The estimation of the R-wave amplitude modulation method [93] performs low pass and notch filer to remove noise. QRS Detection is performed to detect QRS Interval. Baseline wander noise is then removed. After that, cleaned ECG signal is produced to detect peak amplitude and location of each RR-Interval to estimate respiratory signal. Then the interpolation is performed to smooth signal and down-sampled sampling rate to get respiratory signal as show in Figure 4.5.

Since, the accuracy of the based method [93] can be improved if the QRS interval and R-peak amplitude can be detected more accurately. Therefore, I improve the accuracy of the based method by accurately detecting QRS interval and R-peak using low pass filter with 11 Hz cut-off frequency, high pass filter with 5 Hz cut-off frequency, differentiation, moving average and appropriate thresholds. Although, I just apply some further techniques and tune some threshold values to accurately detect QRS interval and R-peak, but it is not easy to find the appropriate one. To compare between the improved method and the based method, I use ECG signal and real respiratory signal from Fantasia Database of PhysioNet [97] as shown in Figure 4.5a and 4.5b respectively. As a result, an estimated respiratory signal from improved method (Figure 4.5d) is quite similar to real respiratory signal

(Figure 4.5b) and have similar number of breathes/minute. Thus, the improved method is better than based method (Figure 4.5c) to estimate respiratory signal from ECG signal.

After the respiratory signal is produced, this application performs to detect inhalation and exhalation signal and calculate time of inhalation and exhalation as shown in Figure 4.6.

According to #8 of workflow, the breathing control application will use this application to detects users' respiratory signal.

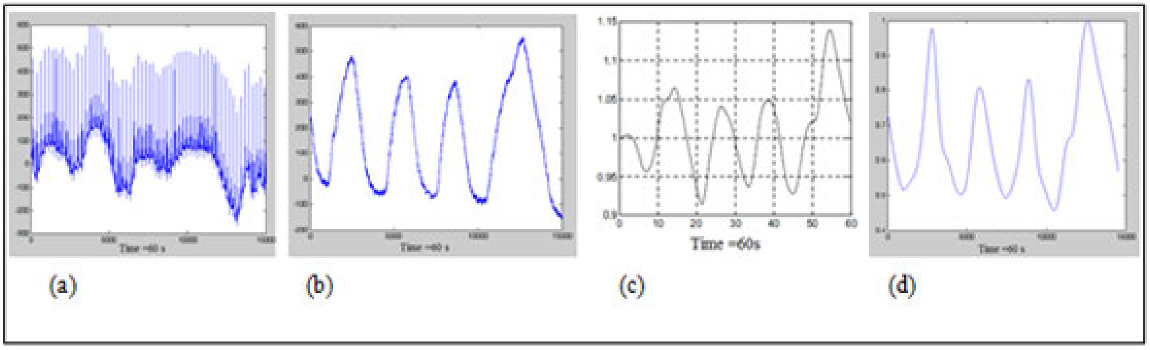


Figure 4.5 ECG and respiration signal: (a) ECG signal [97], (b) Actual respiratory signal [97], (c) Estimated respiratory signal [93], (d) Improvement of estimated respiratory signal

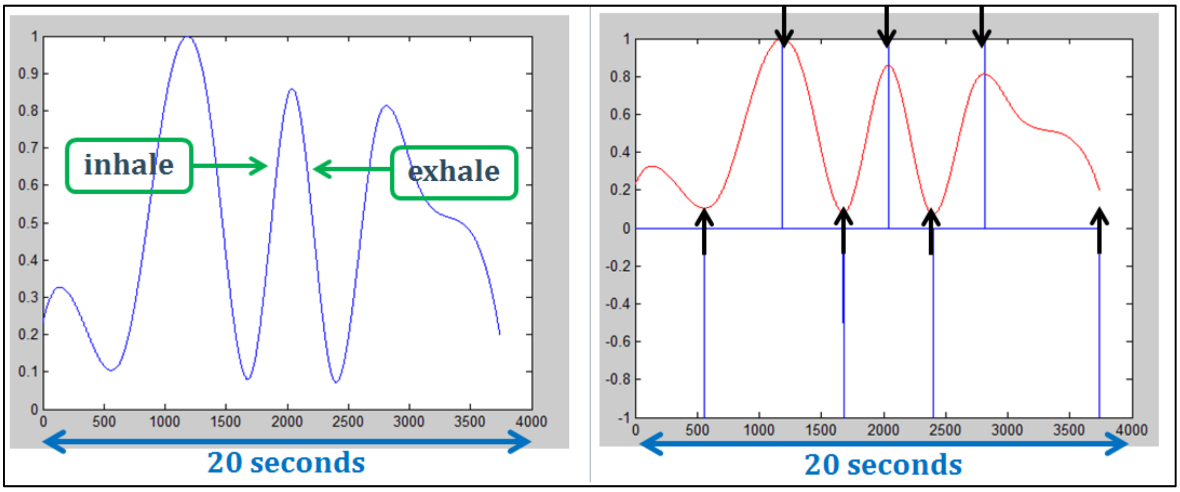


Figure 4.6 Detection of inhalation and exhalation signals

4.3 Experiment 1: Evaluation of service's effectiveness

4.3.1 Objective

This experiment aims to evaluate the effectiveness of breathing control application to decrease stress and how participants feel during using this application.

4.3.2 Participants

After I design and implement the breathing control application, I perform an experiment with seven graduate students from graduated school of Engineering, Shibaura Institute of Technology. The participants are three Thai females, one Malaysian female, two Thai males and one Malaysian male. No participant is familiar with augmented reality application. All participants already knew how to deep breathing. However, no one uses deep breathing technique to increase relaxation when they experience negative emotions or stress in daily life.

4.3.3 Tools and Materials

This experiment checks whether the breathing control application effectively decreases stress and evaluates the participant feelings when they use it. The tools for evaluating the effectiveness of the application and participants feeling are described below.

A. Stress Measurement

The effectiveness of the breathing control application is evaluated by measuring the stress level of each participant. The stress level is measured by performing salivary amylase test.

Salivary amylase is one of the most important enzymes in saliva. The release of this amylase is governed by the activation of the autonomic nervous system (ANS). When human experiences stress, autonomic activation is high; the salivary amylase may increase [98]. Therefore, several researches have performed the experiments and have confirmed that the salivary amylase is significant to indicate the stress level [98-100]. Thus, the salivary amylase test is very useful to measure the stress.

In this experiment, all participants perform salivary amylase test using Cocoro Meter as shown in Figure 4.7. Cocoro Meter is an equipment to measure the salivary amylase [KU/L] from saliva and classify it into stress levels [101].

Figure 4.7 Cocoro meter [101].

B. Evaluations of participants feeling

I evaluate participants' feeling when they use the breathing control application to decrease their stress. I design the questionnaires with a 5-point Likert scale where five is the highest score (strongly agree) and one is the lowest score (strongly disagree). The questions in questionnaire are shown in Table 4.1.

Table 4.1 Questions for evaluation of user feeling

#	Questions
1	While using this application, did you feel relaxed?
2	While using this application, did you feel comfortable?
3	While using this application, did you feel bored?
4	While using this application, did you feel sleepy?
5	While using this application, did you like it?

4.3.4 Experimental procedure

Regarding the objective of this experiment, I design the experiment procedure to increase participants' stress first. After that, let them use the breathing control application to decrease stress in order to measure the effectiveness of the application. After finishing the experiment, participants answer questionnaire to evaluate their feeling.

The experiment procedure is separated into two periods.

- *Stress Increase Period:* each participant need to calculate Math exercises about addition, subtraction, multiply and division. I provide 70 Math problem solving questions; all participants need to answer the questions as much as they can within 20 minutes. The experimenter read each question twice in English. The participants mentally calculate and answer in English. Mental calculation might be able to increase their stress.
- *Stress Decrease Period:* the participants use the breathing control application for ten minutes. They need to put ECG sensor in order to perform deep breathing and controlled their breathing. They listen to the classical music and watch the slow turning of virtual music box. The deep breathing techniques of stress management, the classical music and turning of 3D object might be able to decrease their stress.

Finally, the processes to perform experiment are as follows:

1. The participants perform salivary amylase test to measure their current stress.
2. The stress increase period is served to the participants for 20 minutes in order to increase their stress.
3. After that, the participants perform salivary amylase test again. The stress decrease period is served to the participants for ten minutes in order to decrease their stress and increase their relaxation.
4. Every taking stress decrease period for five minutes, the participants perform salivary amylase to check their stress.
5. Finally, the participants answer the questionnaire to evaluate their feeling and write some comments.

4.3.5 Results

This experiment aims to evaluate two aspects: effectiveness aspect by measuring stress and user feeling aspect by questionnaires. The results of the stress measurement and questionnaires are shown below.

A. Stress measurement results.

I measure salivary amylase value [KU/L] of all participants four times during experiment. The collected salivary amylase data are normalized by subtracting the value of first measurement of each participant. The normalized data are shown in Figure 4.8. From the measured salivary amylases of all participants, an increase in the measured data indicates that the participants become more stressed. If the measured data decreased, the participants become more relaxed.

Table 4.2 shows the number of participants whose stress increases, decreases or doesn't change, after taking each three periods: Stress Increase Period, Stress Decrease Period for five minutes and Stress Decrease Period for ten minutes. The results indicate that after Stress Increase Period, four participants gain more stress, but two participants' stress is decreased. However, after Stress Decrease Period for five minutes, five participants' stress is decreased and finally, all participants' stress decreases from the beginning after they experience the breathing control application for ten minutes (Stress Decrease Period for ten minutes)

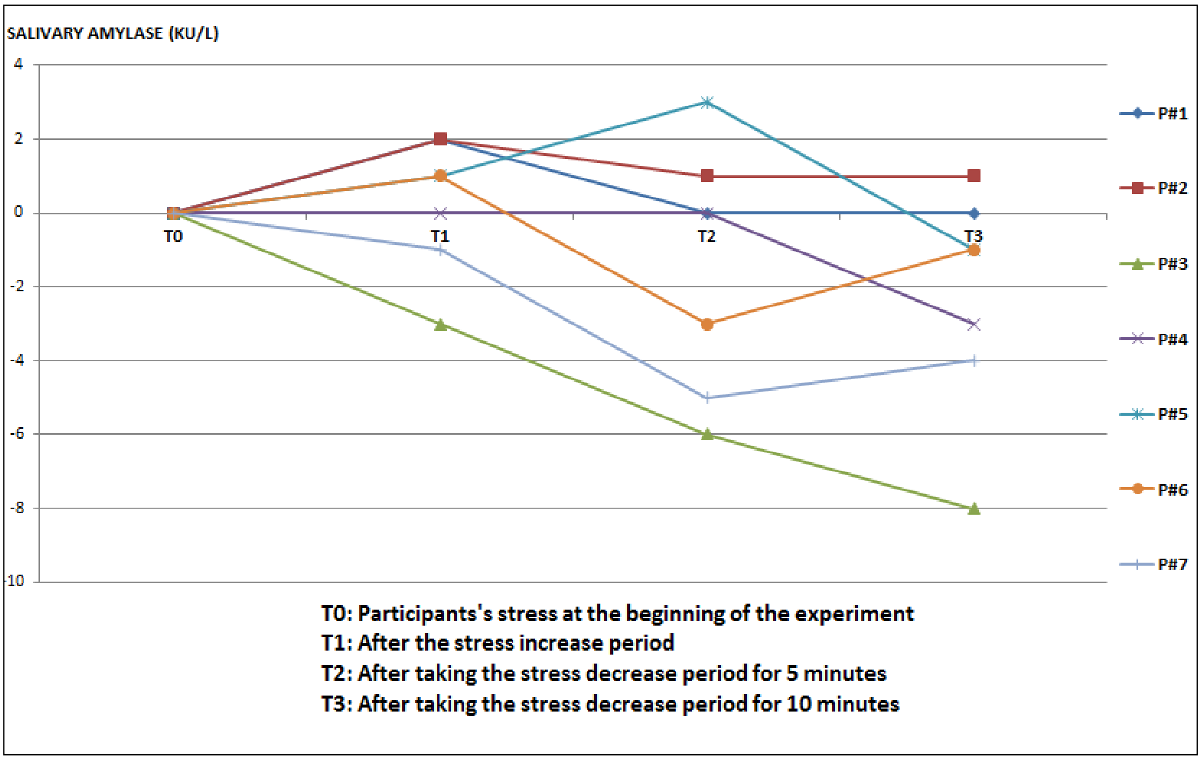


Figure 4.8 Normalized salivary amylase data.

Table 4.2 Analysis result of stress measurement

Stress / Period	Increase	Decrease	No change
After Stress Increase Period (T1-T0)	4	2	1
Taking Stress Decrease Period for 5 minutes (T2-T1)	1	5	1
Taking Stress Decrease Period for 10 minutes (T3-T1)	0	7	0

B. Questionnaire results

I analyze the questionnaire results by using statistic method to determine how users feel when they experience the breathing control application to decrease their stress and increase their relaxation. T-test is applies to determine the results are significant different from population mean or not. Figure 4.9 shows the averages of scores with standard deviations for five questions of the questionnaire. Table 4.3 shows statistical results of questionnaire using t-test. Only the score of question #4 is significantly high ($p<0.05$). It indicated that participants significantly feel much sleepy after experiencing the relaxation service with breathing control application. Although, the results cannot judge that they feel relaxed and comfortable or not, but I obtain the tendencies of the participants when using this application as follows:

- They feel relaxed and comfortable.
- They also like the application when they experience it.
- They feel a little bored.

Moreover, the participants give comments about the application as follows:

- It needs more different kinds of music.
- The 3D virtual music box does not look very realistic.
- There are too many components to focus on with different interface position. They prefer to concentrate on only one position.

Based on the experimental results and participants' comments, the breathing control application needs to re-design for increasing ease of use and efficiency of the application to make users relaxed and comfortable.

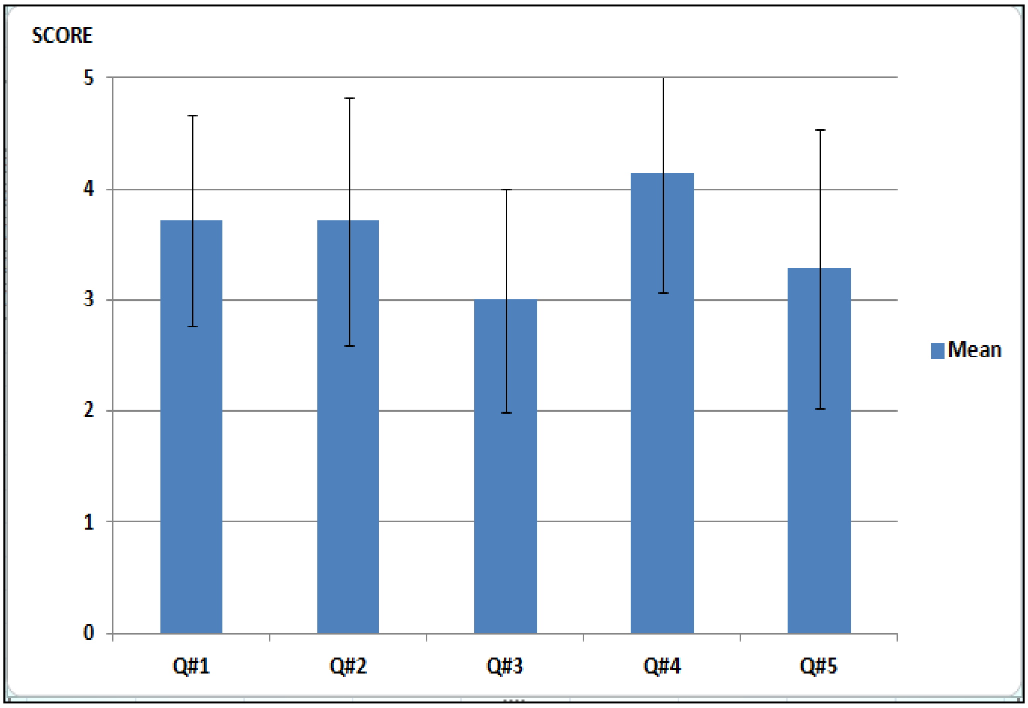


Figure 4.9 Mean scores with standard deviations of five questions.

Table 4.3 Analysis of questionnaire results

Question	Mean	S.D.	t	P-Value
#1	3.71	.95	1.987	.094
#2	3.71	1.11	1.698	.140
#3	3.00	1.00	.000	1.000
#4	4.14	1.06	2.828	.030*
#5	3.28	1.25	.603	.569

*: P<0.05

4.3.6 Discussion

Regard to the salivary amylase test results, some participants don't have stress after they mentally calculate the Math solving problem for 20 minutes. The reason might be the problem is too easy for them. Moreover, some participants don't feel relaxed after Stress Decrease Period for five minutes because the duration to perform deep breathing might be short. Even though, stress level of all participants is decreased after Stress Decrease Period for ten minutes. Thus, the breathing control application can decrease stress and increase relaxation because finally all participants become more relaxed.

The questionnaire results indicate that participants feel much sleepy but they do not indicate that the participants have other emotions or not when the participants use the application because only the result of question #4 is significant. However, the tendency of the participants feeling is obtained that they feel relaxed, comfortable and a little bored respectively. Moreover, the comments from all participants are very useful for the improvement of this application.

In this experiment, I don't evaluate how much the participants feel relaxed and how much their stress decrease. I evaluate only the application can relax participants and decrease their stress. From the results, I can conclude that the application is effective enough to reduce stress and increase relaxation.

However, the effectiveness of Augmented Reality should be clarified, since only deep breathing technique is effective to decrease stress. Therefore, I conduct the second experiment to clarify and confirm the effectiveness of Augmented Reality to help decreasing stress.

4.4 Experiment 2: Confirmation of the augmented reality's effectiveness

4.4.1 Objective

I conduct this experiment to clarify the effectiveness to reduce stress and increase relaxation of AR in breathing control application.

4.4.2 Participants

Six participants are performed this experiment. The participants in this experiment are different from the previous experiment. They are two females and four males. Three participants are undergraduate students from KMUTT, Thailand and the others are graduate students from Shibaura Institute of Technology, Japan. No participant is familiar with augmented reality application. All participants already knew how to deep breathing. However, no one uses deep breathing technique to increase relaxation when they experience negative emotions or stress in daily life.

4.4.3 Tools and Materials

Experimental tools for this experiment are similar to the previous experiment. The Cocoro Meter is used to measure the stress levels [101]. The questionnaire (Table 4.1) is also used to evaluate the participants feeling.

4.4.4 Experimental procedure

I evaluate the effectiveness of AR by comparing the participants' stress level and feeling while they experience the breathing control application with and without AR.

I prepare two sub-experiments as follows:

- Sub-experiment A: use breathing control application with augmented reality (Figure 4.2). Same as previous experiment in section 4.3.
- Sub-experiment B: use breathing control application without augmented reality

(Figure 4.10). There is only deep breathing technique of stress management. (Only (a) and (b) components in Figure 4.2).

The sub-experiment procedure is similar to previous experiment on section 4.3.4

To perform the experiment, I equally divide the participants into two groups. Each group performs in different condition of the experiment to eliminate order effects.

- First group: three participants perform sub-experiment A and then B.
- Second group: three participants perform sub-experiment B and then A.

Figure 4.10 Breathing control application without AR

4.4.5 Results

A. Stress measurement results

Table 4.4 Analysis result of stress measurement between the breathing control application with and without augmented reality.

	With augmented reality (Sub-exp. A)			Without augmented reality (Sub-exp. B)		
Stress / Period	Increase	Decre-ase	No cha-nge	Increase	Decre-ase	No cha-nge
After Stress Increase Period	6	0	0	5	1	0
Taking Stress Decrease Period for 5 min.	1	3	2	2	1	3
Taking Stress Decrease Period for 10 min.	1	5	0	2	3	1

The results in Table 4.4 show the number of participants whose stress increases, decreases or doesn't change, after taking each three periods (Stress Increase Period, Stress Decrease Period for five minutes and Stress Decrease Period for ten minutes) for both sub-experiment A and B. The results show that the numbers of participants whose stress decreased during sub-experiment A are higher than the numbers of participant during sub-experiment B in both five and ten minutes. This indicates that the AR is more effective to decrease stress than only deep breathing.

B. Questionnaire results

Paired t-test is applied to analyze the questionnaire results. Table 4.5 shows statistical results from paired t-test. The results of questions #3, #4, and #5 are significantly different ($P<0.05$). So, the participants feeling is summarized as follows:

- They feet less bored and less sleepy when they use the breathing control application with AR.
- They like the breathing control application with AR more than the application without AR.
- However, the results cannot judge that they really feel relaxed and comfortable. Therefore, the breathing control application needs to re-design for improvement.

From these experimental results, I can confirm that the AR is effective enough to decrease stress and increase relaxation.

Table 4.5 Analysis Result of Questionnaire between the breathing control application with and without augmented reality.

Question	With AR		Without AR		Paired Samples Test			
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**t**	**P-Value**
#1	3.50	1.05	2.33	1.03	1.167	1.83	1.557	.090
#2	3.33	1.03	3.00	.632	.333	.816	1.000	.181
#3	1.67	.816	4.33	.516	-2.667	1.21	-5.394	.001*
#4	2.67	1.21	4.00	1.09	-1.333	1.03	-3.162	.012*
#5	3.50	1.05	2.33	.816	1.167	1.33	2.150	.042*

*: $P<0.05$

4.4.6 Discussion

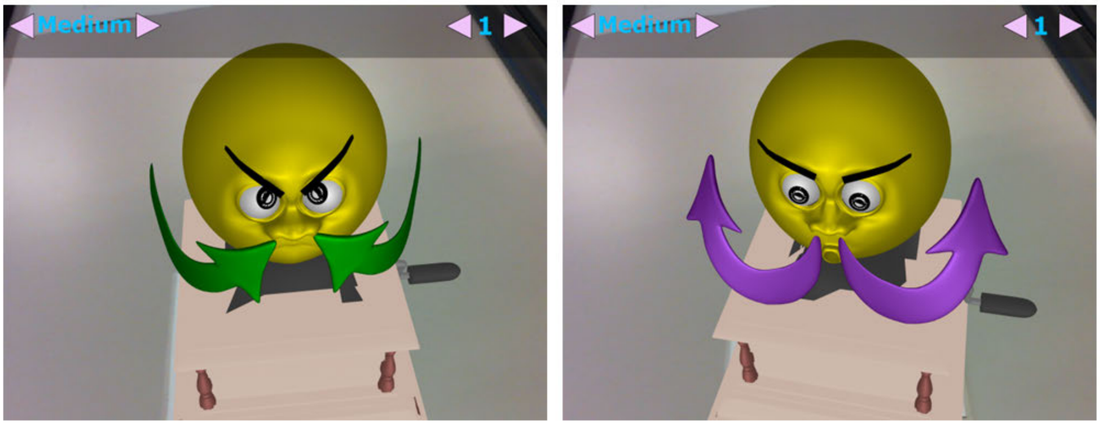
Regard to the salivary amylase test results from both sub-experiment A and B, many participants gain more stress after stress increase period. This means that mental calculation of Math problem solving can increase people stress. After Stress Decrease Period for five and ten minutes, all participants don't feel relax. However, the number of participants who feel relaxed is increasing from the beginning of the experiment until after using the application with and without AR for five and ten minutes. This indicates that performing deep breathing for long time can effectively increase relaxation. Furthermore, the results show that many participants feel more relaxed after using the application with AR than without AR. Therefore, it can be said that AR helps decreasing stress more quickly than only deep breathing because AR (virtual music boxes and music) can stimulate users to concentrate on deep breathing by following the virtual music boxes.

Although, the questionnaire results of #1 and #2 which specific participants' relaxation and comfort, are not significantly different between application with and without AR, but the tendency can be obtained that the participants feel more relaxed and comfortable when using the application with AR. Moreover, the questionnaire results of #3 and #4 indicate that participants significantly feel much bored and sleepy when using the application without AR because they perform only deep breathing without any virtual music box and music. Thus, AR can make them feel more relaxed, comfortable, and fun when perform deep breathing and they also like AR. Therefore, AR is effective to reduce stress and increase relaxation in breathing control application.

4.5 Re-design of breathing control application

According to previous experimental results and participants' comments, the breathing control application is improved based on the following issues.

- Several participants focus only on the animation guide and (Figure 4.2 (a)). They have less concentration on the 3D virtual music box. So I change the animation guide process by re-creating it as a virtual object and combine it with 3D virtual music box. It is called a training model (Figure 4.11), which is only displayed for two a half minutes. After that, the different kind of virtual music box is displayed.



(a) Inhalation through nose (b) Exhalation through mouth

Figure 4.11 Training model

- Participants also focus more on the application suggestions that informs them every 20 seconds to breathe more deeply or shortly. I change position of the application suggestions to display on the virtual music box to allow users to increase their concentration on the virtual music box.
- I change the deep breathing manner. Participants need to slowly inhale through their nose for three seconds to five seconds and slowly exhale through their mouth for three seconds to five seconds [102].
- Since the participants feel tired when they breathe by following the animation guide, I change it to display only two and a half minutes to guide how to deep breathing and they need to adjust the suitable length of their deep breathing.
- Since the participants feel bored because the application provides only classical music, so I provide several kinds of music.

The improvements of breathing control application are shown in Figure 4.12. It now has the following five main components.

- *The breathing speed (a):* provides five speeds for deep breathing: slowest (breath rate is 5 seconds), slow (breath rate is 4.5 seconds), medium (breath rate is 4 seconds.), fast (breath rate is 3.5 seconds), fastest (breath rate is 3 seconds).
- *The virtual model (b):* provides three different kinds of music boxes: Emoticon,

Gramophone [103] and Castle [104].

- *The music (c):* provides classical, jazz and pop music which users can choose by using arrows.
- *The 3D virtual music box (d):* slowly rotates and plays music.
- *The application suggestion (e):* gives advice about breathing more deeply or quickly.

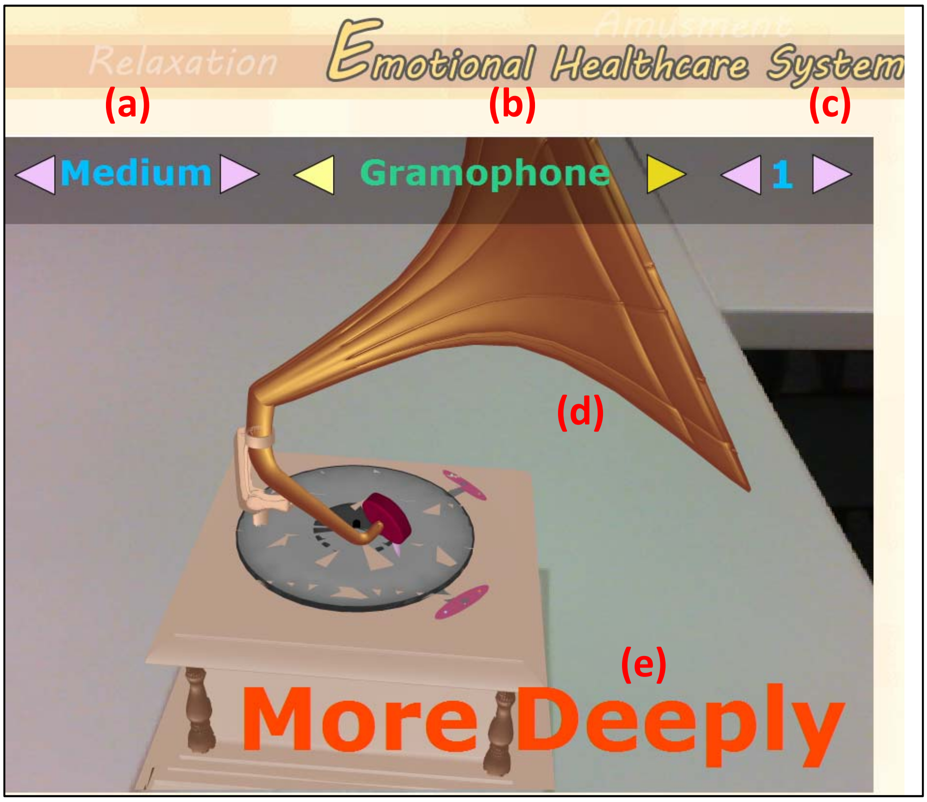


Figure 4.12 New design of breathing control application

4.6 Summary

For the emotional healthcare system, the first service that I design, implement and evaluate is the relaxation service. This is the most important service to support users because relaxation can build positive attitude and good health in daily life. Moreover, it can reduce negative emotions and stress.

I newly design and develop the breathing control application using AR in relaxation service; this application applies deep breathing technique of stress management to increase relaxation and decrease stress or negative emotions. The application includes a virtual music box using AR technology to assist such deep breathing. The application integrates the ECG sensor to convert ECG signal to respiratory signal in order to detect the respiration rate and suggest how to deep breathing in an appropriate way. The experimental results indicate that the breathing control application with AR can help decreasing stress and increasing relaxation, better than using only deep breathing. Moreover, the augmented reality, one of my system features, can help improving user relaxation. Finally I re-design the breathing control application based on the comments from all participants to improve ease of use, user relaxation and comfort, and avoid boring.

Chapter 5

Emotion recognition by facial expression

Emotion detection or recognition is essential and useful in human computer and human robot interaction applications because emotions indicate feelings and needs [6]. To recognize user emotions and provide appropriate services, the emotional healthcare system applies the emotion recognition by facial expression because facial expression is one of the most powerful, natural and easy methods for human to express and communicate emotions and intentions [6].

As mentioned in the literature surveys (Chapter2), there are several approaches to detect emotions from facial expression. However, only few approaches that fit such requirements of the emotional healthcare system as high accuracy and high speed for real-time processing. Thus, I improve one of appearance-based feature extraction approach. The basic idea is to improve the feature representation so that the feature vector size is reduced while maintain or even improve the detail of the pattern. This chapter presents the design, implementation and evaluation of the emotion recognition by facial expression using my improved feature extraction approach.

5.1 Support Vector Machine classification as background knowledge

SVM was originally proposed by Vladimir N. Vapnik in 1995 [105-106]. SVM is a well-known supervised learning model for both linear and nonlinear data which uses a hyperplane to separate the classes and maximize the margin between the classes. The hyperplane is built from margins and the nearest data points called support vectors which are the critical elements of the training set. The mapping process is governed by a kernel function, which can be linear or non-linear.

Suppose I have training dataset {xi, yi}, i=1, 2, …, n, $y_i \in \{-1,1\}$ and $x_i \in R^d$. [107-108]. The data point xi belongs to class yi where the classes are 1 or -1. The hyperplane is constructed to separate between these two classes in which the data point x that lie on the hyperplane satisfy $w \cdot x + b = 0$, where w is normal vector to the hyperplane and $|b|/||w||$ is the distance from the hyperplane to the origin.

In the case of linearly separable dataset, two hyperplane are built to separate the classes and maximize the margin between the classes that can be formulated as below.

$$xi \cdot w + b \geq +1 \text{ for } yi = +1$$

$$xi \cdot w + b \leq -1 \text{ for } yi = -1$$

In the case of non-linear classification, Bernhard E. and his research group suggested to apply the non-linear kernel function to maximize the margin between hyperplanes. The non-linear kernels for pattern recognition problem are as follows:

- Polynomial function of degree p in the data: $K(x, y) = (x \cdot y + 1)p$
- Gaussian radial basis function: $K(x, y) = e^{-||x-y||^2/2\sigma^2}$
- Hyperbolic tangent function: $K(x, y) = \tanh(\kappa x \cdot y - \delta)$

Moreover, SVM can process as multi-class classifier by combining multiple binary SVM classifiers. There are two common approaches for multi-class classification: one against all approach and one against one approach. One against all approach uses the binary SVM classifiers to separate data points of each one class from the other classes. One against one approach uses binary SVM classifiers to separate data point of each pair of classes. Several researches in computer visions and pattern recognition adopted both linear and non-linear SVM classification to train and classify the dataset because SVM can produce higher performance and accuracy.

5.2 Design and implementation

5.2.1 Workflow of emotion recognition by facial expression

Emotion recognition by facial expressions recognizes and interprets human emotions from facial textures, and the movement of facial muscles, eyes, mouth or eyebrows. The workflow of the emotion recognition by facial expression is shown in Figure 5.1.

The training process takes facial expression image dataset, collected from professional actors and, in each image, face regions are identified and used to extract facial features. EmguCv face detection algorithm using Haar Cascades [109] is used to identify face regions. EmguCv face detection detects eyes and face region that appear in an image. This method can detect multiple facial regions, but to apply this method in the emotion recognition, only the most prominent facial features are used for the real-time processing purpose. As shown in Figure 5.2, only facial features that are related to emotions are included. Particularly, ears are automatically excluded by this emotion recognition since they bear no relation to facial expression.

Support vector machine (SVM) is selected as classification methods then iteratively trained using all data points in feature vector that belong to one of seven classes in order to construct the hyperplane. The hyperplane is built to separate between the classes in which each data point that lie on the hyperplane should satisfy linear method. Then, in real-time processing, SVM uses the hyperplane to categorize input data into appropriate class. After training, in the real-time recognition process, the emotion recognition captures and detects users' face from each frame of the video input. The facial feature vectors are then extracted and used with SVM to identify the most likely emotion class from the seven classes (neutral, happiness, sadness, anger, disgust, fear, and surprise).

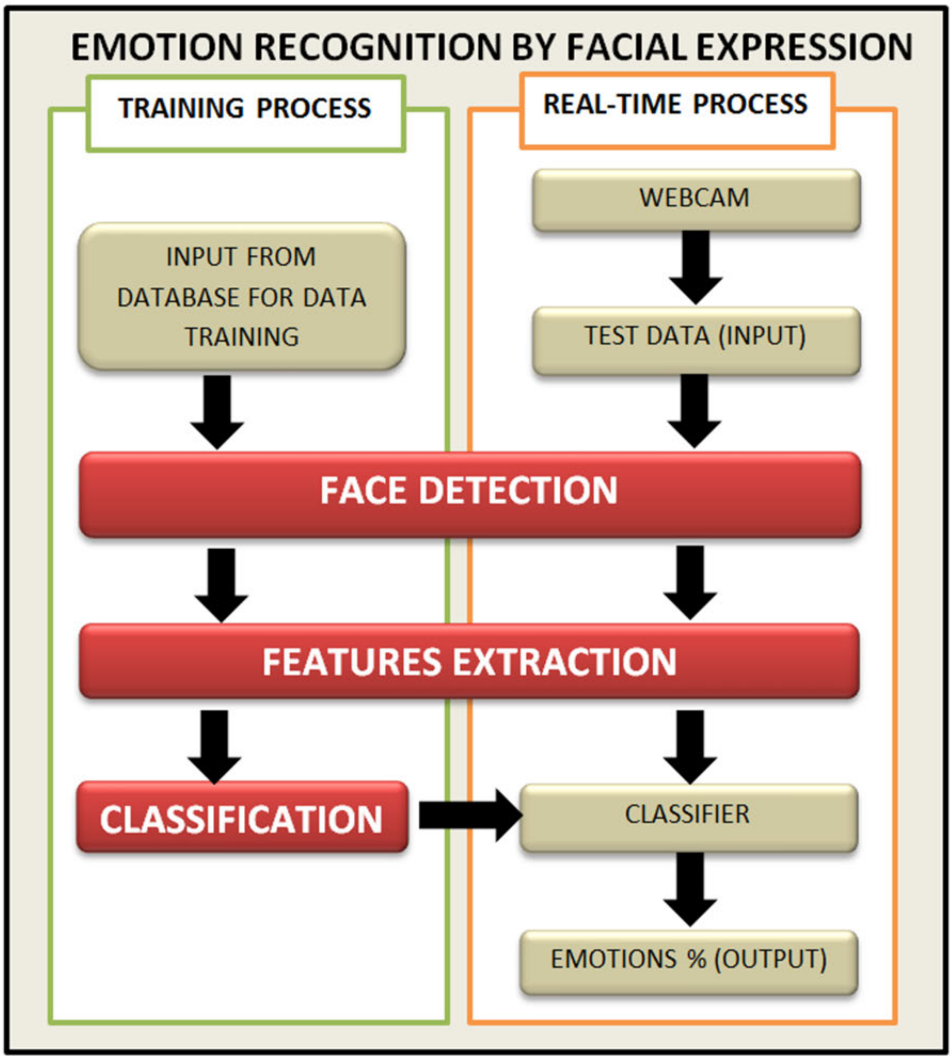


Figure 5.1 Workflow of emotion recognition by facial expression.

Figure 5.2 Emotion-related facial region identification [110,111,24]

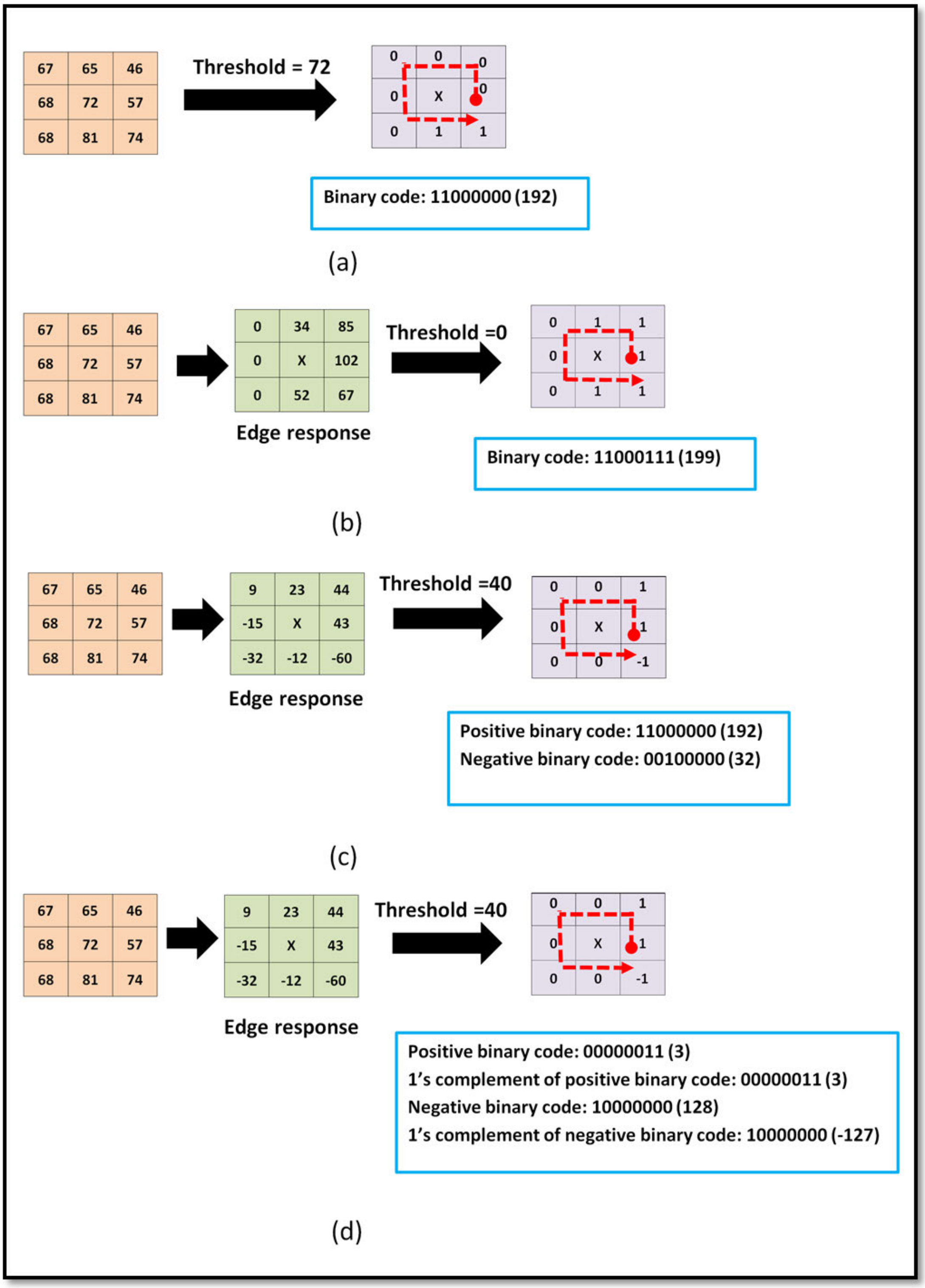


Figure 5.3 Feature extraction methods, including (a) LBP operator with threshold of 72 [25], (b) LDP operator with threshold of 0 (mk when k = 3) [56], (c) DTP operator with threshold of 40 [27], (d) My approach operator with threshold of 40.

5.2.2 Previous feature extraction approaches

Regarding the literature review of the emotion recognition by facial expression (Chapter 2), I found that the appearance-based approaches are more suitable in a real-time environment than the geometry-based approaches because of the complexity of the geometry-based approaches. Among several appearance-based approaches, DTP has highest accuracy because the negative different values between edge responses are considered and the noise sensitivity is lowered than those of LBP and LDP [27]. In addition, the encoding of the edge responses is more stable than that of the intensity values in obtaining the correct binary patterns from noisy images [25, 56 and 27]. DTP is also more accurate than the Gabor wavelet and Curvelet filters in different image resolutions. I then choose to improve DTP to fit the needs of the real-time facial emotion recognition. The LBP, LDP and DTP operators are shown in Figure 5.3 (a - c)

5.2.3 My new feature extraction approach: complementary directional ternary pattern

Feature extraction is an important process for extracting facial features that can represent changes in facial expressions caused by emotions. My feature extraction improves the Directional Ternary Pattern (DTP), which is an appearance-based approach. To improve its performance, the size of DTP feature vector should be decreased. However, reducing its size would affect its effectiveness. To achieve such a reduction and maintain the accuracy, it is crucial to first take a look at how the feature vector is constructed.

A feature vector of appearance-based approach is generally formed from 256-level histograms, created by encoding eight neighboring pixels as an eight-bit binary pattern and transforming it to an unsigned decimal number for mapping to each bin of histograms (Figure 5.3). Such a transformation of topological arrangement that many elements are zeroes into binary codes could make the representation quite sparse. So there is a room to improve the accuracy and performance of the emotion recognition by considering the binary coding scheme that is more compact. Since the purpose here is to encode the feature vector into binary codes, and the most important thing about this feature vector is its regional contrast, not on whether the edge occurs on which bit locations. Here I devise a strategy to code the bit arrangement in such a way that the occurrences of edges that are complementary to one another are grouped in the same bin. Considering the following binary coding techniques:

- Signed Magnitude: this is the simplest way to represent a sign of numbers by applying a left-most bit as a sign bit. If left-most bit is 0, it indicates a positive number. If left-most bit is 1, it indicates a negative number. For example, 00000001 is +1 but 10000001 is -1.
- One's Complement: the left-most bit is a sign bit to indicate positive (0) and negative (1) numbers. Positive numbers are similar to ordinary binary numbers. However, negative numbers are represented in different way by convert 0 to 1 or vice versa. For example, 00000001 is +1 but 11111110 is -1.
- Two's Complement: positive numbers are similar to ordinary binary numbers. However, negative numbers are obtained by calculating one’s complements first and then add one to it. For example, 00000001 is +1 but 11111111 is -1.

These three techniques modify binary codes when generate negative numbers. Two’s complement is actually the derivative of one’s complement. Also, only one’s complement technique does not change patterns or types of binary numbers that represent patterns of edge, spot, etc. in an image. Therefore, it can be used to represent feature that is complementary to one another. This is the basic idea of my proposed feature vector, Complementary DTP (CDTP). CDTP reduces the sparseness in feature representation by using one’s complement coding and absolute operator to obtain a bin that represents a specific complementary pattern. This halves the feature size from 256 to 128 while maintains the complementary patterns in the same bin (Figure 5.4) in order to reduce the redundancy which might be able to increase the accuracy [112].

Similar to DTP, my approach encodes the information of corners, edges, spots, flats, and other local features as binary numbers [25] (Figure 5.4) to generate a binary pattern by comparing the edge responses of neighboring pixels around the center pixel to the specific threshold and then splitting the binary into a positive binary pattern and negative binary pattern to compute the complementary code of positive and negative binary patterns. Absolute value of the decimal code is then used to identify the position of histogram bins as shown in Figure 5.4. These feature vectors of positive and negative binary patterns are concatenated together to construct the full feature vector of that region. Finally, full feature vectors of all regions are concatenated together to build one global feature vector for emotion classification.

Below are steps to extract facial feature using my approach from one facial image (Figure5.5):

1. After face detection, CDTP divides the facial region image into 7*6 rectangular regions.
2. For each region and each pixel, it applies edge detector to get edge response in eight directions.
3. To encode each pixel, it then compares eight neighborhood pixels with threshold to generate positive and negative binary codes. The threshold is set based on the standard deviation of edge responses from smooth face regions
4. Then, it calculates the decimal representative values of both positive and negative binary codes using one's complement technique.
5. After finding the representative values of one's complement of positive and negative binary codes from all pixels in each region, it constructs 128-level histograms of one's complement of positive and negative binary patterns as feature vector with 256 lengths
6. Finally, it concatenates the feature vectors from all regions together to construct the global feature vector for emotion classification.

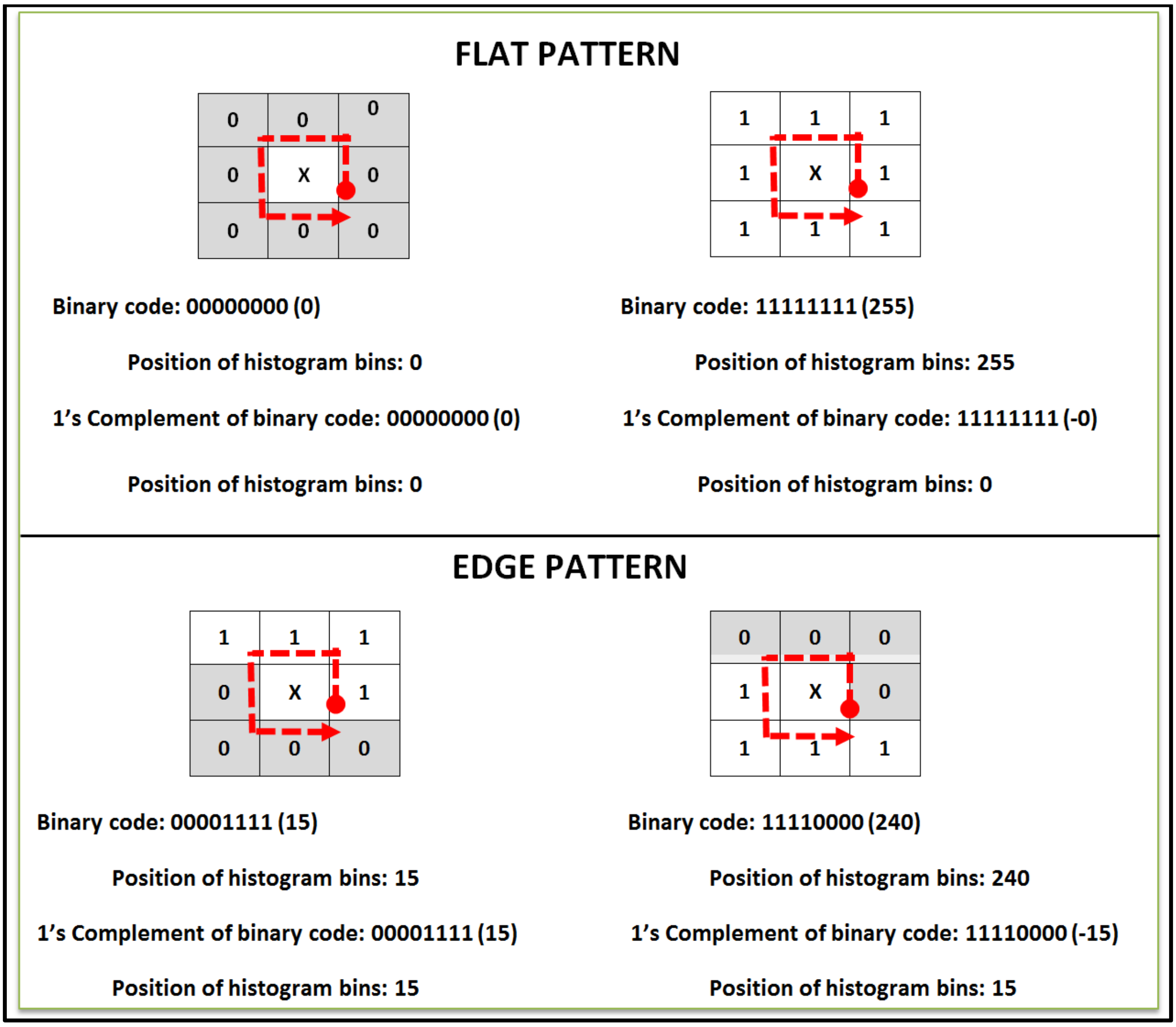


Figure 5.4 Type of Binary Pattern [25].

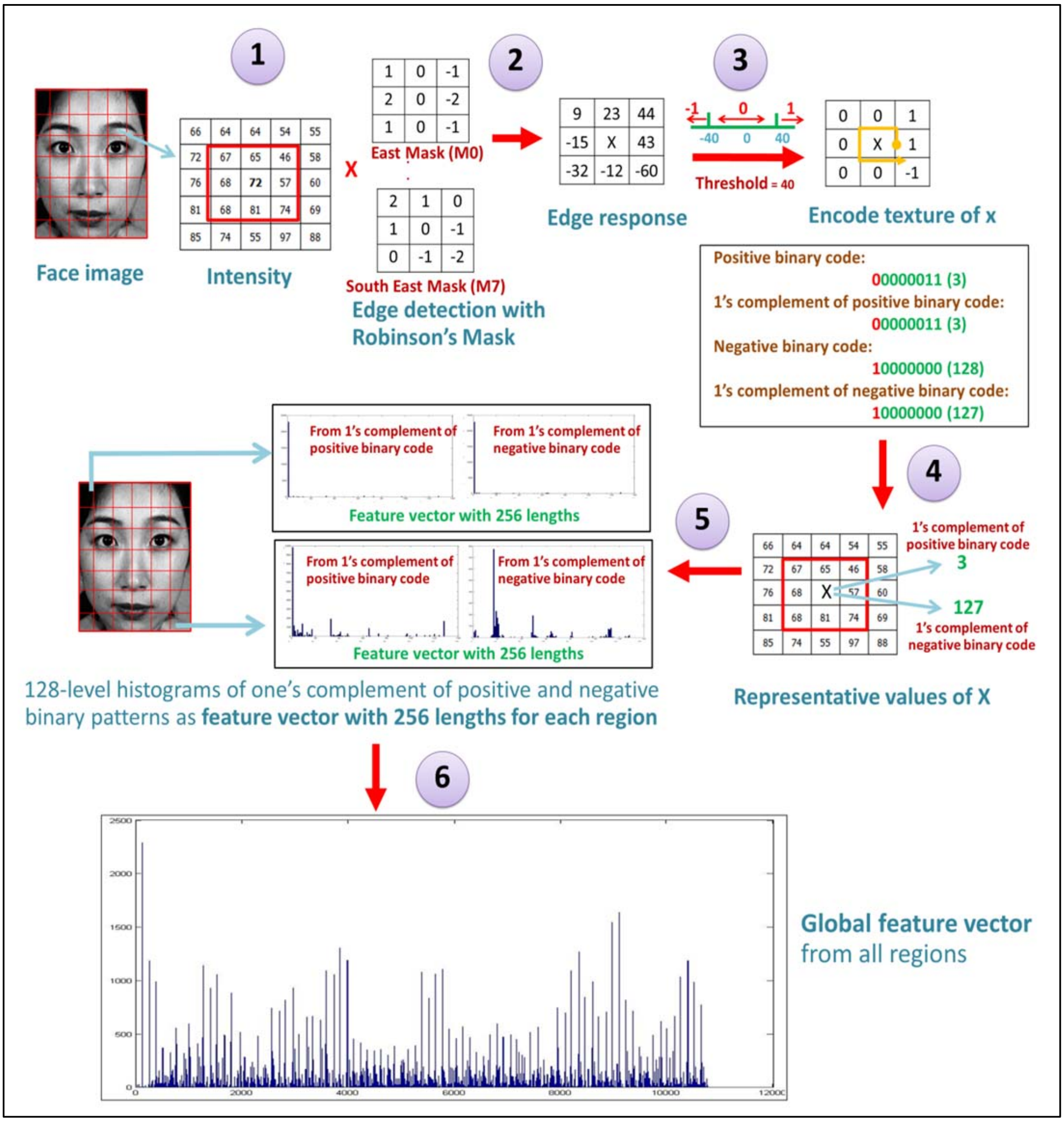


Figure 5.5 CDTP process [110]

5.3 Evaluation

In this section, the proposed method is evaluated and compared against other methods in terms of both accuracy and performance aspects. Through a series of experiments, the evaluation aims to address the following research questions: (1) Is the proposed method significantly improve the accuracy of emotion classification? (2) Does the improvement affect the performance of the emotion recognition? (3) Is the proposed method still robust when using it either with only partial face data or on images with different resolutions?

5.3.1 Datasets

Universal emotions in facial expressions can be classified into seven basic emotions: neutral, happiness, sadness, anger, disgust, fear and surprise, as shown in examples in Figure 5.6. These emotions are what people of most nationalities similarity expressing on their face [113]. Four datasets that are designed with these basic emotions are used to assess the accuracy and performance of the proposed method.

1. Japanese Female Facial Expression (JAFFE) dataset [110], which consists of 213 images of ten female subjects in six basic emotions and a neutral expression. Each image on JAFFE dataset was rated by 60 Japanese subjects to assign the emotion label. Thus they are categorized based on their facial expression into 31 happiness, 29 disgust, 32 fear, 30 anger, 31 sadness, 30 surprise and 30 neutral.
2. Karolinska Directed Emotional Faces (KDEF) dataset [111], which consists of 4900 pictures from 70 amateur actors (35 females and 35 males whose ages are between 20 and 30 years old). Each subjects display facial expression in seven emotions (happiness, disgust, fear, anger, sadness, surprise and neutral) together with five different angles (full left, half left, straight, half right and full right). The emotion labels of KDEF are compatible and similar to those of CK+ and JAFFE. In this thesis, I selected facial expressions from straight face with seven emotions. Thus for KDEF, there are 490 facial expressions in total, including 70 happiness, 70 disgust, 70 fear, 70 anger, 70 sadness, 70 surprise and 70 neutral.
3. An extended Cohn-Kanade Dataset (CK+) [24] is an extension of the Cohn-Kanade Dataset that consists of 593 sequences of 123 subjects that are FACS (Facial Action Coding System) coded at the peak frame. All sequences are from the neutral face to the peak expression. However, only 327 of the 593 sequences have emotion sequences because the peak expression of 327 sequences fits the prototypic emotion definition that was validated with reference to the FACS Investigators Guide and confirmed by visual inspection by emotion researchers. Each peak

expression of each sequence is assigned to have only one emotion. Therefore, I selected 309 emotion sequences for which each peak expression is labeled as one of the six basic emotions. In each emotion sequence, I selected three peak expressions. I also selected one neutral expression for each subject from the emotion sequences. Therefore, I selected 1033 images in total (207 happy, 177 disgust, 75 fear, 135 angry, 84 sad, 249 surprise and 106 neutral).

4. Since the numbers of selected images from JAFFE and KDEF datasets for each emotion are similar, I combine both datasets as JAFFE-KDEF to construct combined dataset between Asian and European subjects.

Figure 5.6 Seven basic emotions in facial expression from (a) Japanese Female Facial Expression [110], (b) Karolinska Directed Emotional Faces [111], (c) extended Cohn-Kanade dataset [24].

5.3.2 Evaluation procedure

A series of computational experiments is conducted to address each of the research questions posted earlier. Classification accuracy and performance are the measurements that will be taken in each experiment to assess the effectiveness and efficiency of the proposed method in comparison to other previously proposed methods. The accuracy is measured by statistically comparing the predicted emotions with those originally annotated. The performance comparison is done by comparing the running time of the proposed method with others.

SVM is selected as the main classifier in this study because of its high accuracy as shown in Table 5.1. Low accuracy in k-NN suggests that the feature may be too noisy to be cover by similarity-based classification.

Table 5.1 Average recognition accuracies from LBP, LDP, DTP, and CDTP using all datasets comparing between SVM and k-NN when k = 1

Classifier	Accuracy
SVM	75.00%
K-NN	53.56%

Experiments here are offline experiments (experiment 1-4) which aim to evaluate the effectiveness and efficiency of the proposed method in comparison to other methods. In each offline experiment, ten-fold cross-validation is used to test the intra-dataset generalizability. RapidMiner Studio 5 [114] is used to design the analysis processes and perform SVM pattern classification.

Sets of facial features based on different approaches (my approach, LBP [25], LDP [56] and DTP [27]) are extracted from the training and testing datasets. Additionally, to determine the most suitable edge detection algorithms for CDTP, I test two candidate methods which are implemented in DTP and LDP respectively, including Robinson eight-directional edge detection used in DTP (henceforth, CDTP-A) and Kirsch eight-directional edge detection used in LDP (henceforth, CDTP-B). Analysis of variance (ANOVA) is used to statistically analyze related factors in each experiment. Further post-hoc analysis in each experiment is done using Scheffé's method.

5.3.3 Experiment 1: Testing accuracy of emotion recognition

The aim of this experiment was to determine the best classifier configuration that would result in the highest classification accuracy and to assess the accuracy of my approaches (CDTP-A and CDTP-B) in comparison to other feature extraction methods, including LBP, LDP, and DTP. Two different comparison strategies used in SVM, including one-against-one and one-against-all, were evaluated in four datasets used in this study. I found that that there is a significant interaction between the comparison strategies and the datasets [$F(3,32) = 56.994$, $p = .00$] as shown in Figure 5.7.

For JAFFE and CK+ datasets, the one-against-all strategy (JAFFE: M = 78.03, SE = 4.26; CK+: M = 93.65, SE = 1.50) has a higher accuracy than the one-against-one (JAFFE: M = 65.44, SE = 3.43; CK+: M = 89.71, SE = 1.79). On the other hand, the one-against-one for KDFE (M = 67.55, SE = 2.40) and JAFFE-KDEF (M = 60.71, SE = 3.04) produces higher accuracy than the one-against-all (KDEF: M = 62.25, SE = 1.83; JAFFE-KDEF: M = 58.00, SE = 1.48). Therefore, to use the most optimal classifiers, one-against-all strategy was applied to JAFFE and CK+ datasets and one-against-one was used for KDEF and JAFFE-KDEF.

Table 5.2 shows the averaged classification accuracies and standard errors comparing between different feature extraction methods applied to JAFFE, KDEF or CK+ datasets. For JAFFE, my approach (CDTP-A) was the most accurate method comparing to others. Figure 5.8(a) shows the confusion matrix of the classification accuracy of CDTP-A. Consistently high recognition rates can be observed in all emotions. However, the accuracy of sad expression was less than the others because it was confused with neutral expression.

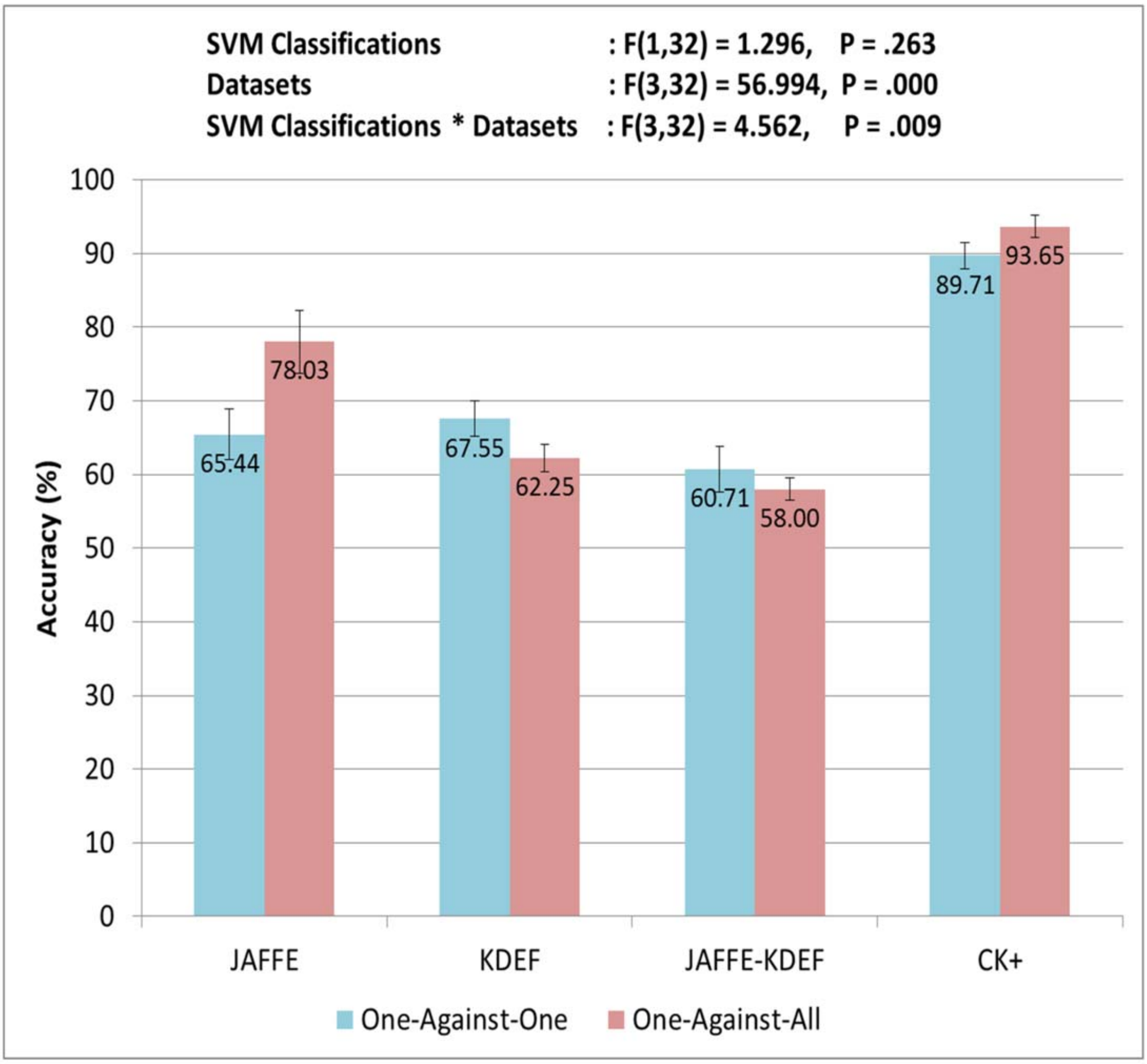


Figure 5.7 Analysis results between SVM and dataset using two-way ANOVA.

	ANG (%)	DIS (%)	FEA (%)	HAP (%)	SAD (%)	SUR (%)	NEU (%)
ANG	**86.67**	6.90	0	0	3.23	0	0
DIS	6.67	**93.10**	6.25	0	3.23	0	0
FEA	0	0	**81.25**	3.23	0	0	0
HAP	0	0	0	**93.54**	6.45	3.33	0
SAD	3.33	0	3.125	0	**74.19**	0	3.33
SUR	0	0	3.125	0	0	**90.00**	0
NEU	3.33	0	6.25	3.23	12.90	6.67	**96.67**

(a) Confusion matrices from JAFFE
using one-against-all of linear SVM

	ANG (%)	DIS (%)	FEA (%)	HAP (%)	SAD (%)	SUR (%)	NEU (%)
ANG	**78.57**	10	4.29	0	5.72	0	2.86
DIS	5.71	**75.71**	0	4.28	0	0	0
FEA	4.29	2.86	**74.29**	10	20	28.57	4.28
HAP	0	0	1.42	**82.86**	0	0	0
SAD	7.14	11.43	10	1.43	**57.14**	1.43	11.43
SUR	0	0	5.71	0	0	**67.14**	1.43
NEU	4.29	0	4.29	1.43	17.14	2.86	**80**

(b) Confusion matrices from KDEF
using one-against-one of linear SVM

	ANG (%)	DIS (%)	FEA (%)	HAP (%)	SAD (%)	SUR (%)	NEU (%)
ANG	**62.00**	6.06	1.96	0	0	0	1.00
DIS	11.00	**72.73**	0	0.99	1.98	0	0
FEA	3.00	6.06	**68.63**	6.93	24.75	21.00	5.00
HAP	0	0	0.98	**84.16**	0.99	1.00	0
SAD	8.00	8.08	10.78	1.98	**45.54**	0	5.00
SUR	0	0	4.90	0.99	0	**65.00**	0
NEU	16.00	7.07	12.75	4.95	26.73	13.00	**89.00**

(c) Confusion matrices from JAFFE-KDEF
using one-against-one of linear SVM

	ANG (%)	DIS (%)	FEA (%)	HAP (%)	SAD (%)	SUR (%)	NEU (%)
ANG	**98.52**	0.00	1.33	0.48	4.76	0.00	16.04
DIS	0.00	**100.00**	0.00	0.00	0.00	0.00	0.94
FEA	0.00	0.00	**98.67**	0.00	0.00	0.00	0.00
HAP	0.00	0.00	0.00	**99.52**	0.00	0.00	0.00
SAD	0.00	0.00	0.00	0.00	**85.71**	0.00	4.72
SUR	0.74	0.00	0.00	0.00	0.00	**98.39**	1.89
NEU	0.74	0.00	0.00	0.00	9.52	1.61	**76.42**

(d) Confusion matrices from CK+
using one-against-all of linear SVM

Figure 5.8 Confusion matrices from my facial emotion recognition. (ANG = Anger, DIS = Disgust, FEA = Fear, HAP = Happiness, SAD = sad, SUR = Surprise, NEU = Neutral): (a) with JAFFE using one-against-all of linear SVM, (b) with KDEF using one-against-one of linear SVM, (c) with JAFFE-KDEF using one-against-one of linear SVM , (d) with CK+ using one-against-all of linear SVM.

Table 5.2 Comparison of classification accuracies between my approaches and other methods on different datasets

Feature Extraction Methods	**JAFFE (one against all)**	**KDEF (one against one)**	**JAFFE-KDEF (one against one)**	**CK+ (one against all)**
LBP	62.90 ± 0.39	66.33 ± 0.23	59.46 ± 0.24	87.70 ± 0.11
LDP	76.56 ± 0.49	72.65 ± 0.23	64.58 ± 0.25	94.87 ± 0.04
DTP	84.09 ± 0.49	62.24 ± 0.25	51.49 ± 0.21	95.45 ± 0.05
CDTP-A	87.77 ± 0.49	62.86 ± 0.32	58.46 ± 0.23	95.65 ± 0.05
CDTP-B	78.81 ± 0.59	73.67 ± 0.14	69.56 ± 0.18	94.58 ± 0.07

On the other hand, CDTP-B is more accurate than other approaches when classifying the emotions from KDEF and JAFFE-KDEF datasets. From the recognition rates of seven-class emotional expression classification as shown in Figure 5.8(b) and 5.8(c), the accuracy of sad expression is less than the others because it is confused with neutral and fear classes for both datasets.

Additionally, my approach (CDTP-A) was the most accurate method comparing to others for CK+. The confusion matrix is shown in Figure 5.8(d). The recognition rates of all emotions are very high. It can recognize disgust expression up to 100%. However, the accuracy of neutral expression is lower than others because it confused with anger expression.

When analyzing all possible combinations of datasets, I found that different feature extraction methods do not significantly affect the accuracy [$F(4, 15) = .258$, $p = 0.900$]. My approach is also not either significantly better or worse than others, but nevertheless the mean accuracy of my CDTP-B ($M = 79.16$, $SE = 5.48$) is the best among approaches. Thus, in general, CDTP-B may be more preferable than other approaches since it has highest accuracy.

5.3.4 Experiment 2: Testing performance of the emotion recognition

This experiment aims to assess the performance of my approach comparing to others. This is done by comparing the execution time of feature extraction and classification methods, separately and in combination. A testing image with resolution of 246×308 is used in the process by repeating feature extraction and classification 100 times to gain measurable time scale. The results are statistically analyzed using one-way ANOVA. I found that time requires to extract features depends significantly on the methods [$F(3, 396) = 426.46$, $p < 0.001$] as shown in Figure 5.9.

Post-hoc comparisons using the Scheffé's method indicate that the DTP-based methods, including DTP and CDTP requires more computational times than LBP [DTP: Mdiff = -239.86, 95% CI[-265.89, -213.84], $p < 0.001$]; Mdiff = -235.20, 95% CI[-261.22, -209.18], $p < 0.001$] and slightly more than LDP [DTP: Mdiff = -90.76, 95% CI[-116.78, -64.74], $p < 0.01$; (Mdiff = -86.10, 95% CI[-112.12, -60.08], $p < 0.001$)]. This is because DTP-based methods spend much time on calculating positive and negative binary pattern to construct feature vector. The complexity of my approach, however, does not significantly differ from that of DTP (Mdiff = 4.66, 95% CI[-21.36, 30.68], $p = 0.947$). This means that the additional complexity of the improvement does not significantly increase the feature extraction time from the general DTP method.

Surprisingly, the performance of classification is significantly depending on the feature extraction methods [$F(3, 396) = 368.28$, $p < 0.001$]. The post-hoc results indicate that my approach makes the classification significantly faster than DTP (Mdiff = 5.91, 95% CI [5.28, 6.54], $p < 0.001$. Moreover, it also indicates that DTP is significantly slower than LBP (Mdiff = -3.94, 95% CI [-4.57, -3.31], $p < 0.001$), LDP (Mdiff = -2.13, 95% CI [-2.76, -1.50], $p < 0.001$), because of its feature vector size.

When considering the total computation time, feature extraction methods do indeed affect the total execution times [$F(3, 396) = 424.62$, $p < 0.001$]. My approach requires more time to execute than both LBP (Mdiff = -233.23, 95% CI[-259.43, -207.03], $p < 0.001$) and LDP (Mdiff = -82.32, 95% CI[-108.52, -56.12], $p < 0.001$). Comparing to DTP, the computation time of the approach is not significantly different (Mdiff = 10.57, 95% CI [-15.63, 36.77], $p = 0.60$), but it is slightly faster. The reason that DTP-based methods are slower than LBP and LDP is because they spend much more time than others in constructing feature vectors by considering positive and negative binary patterns. But it makes the classifiers more accurate.

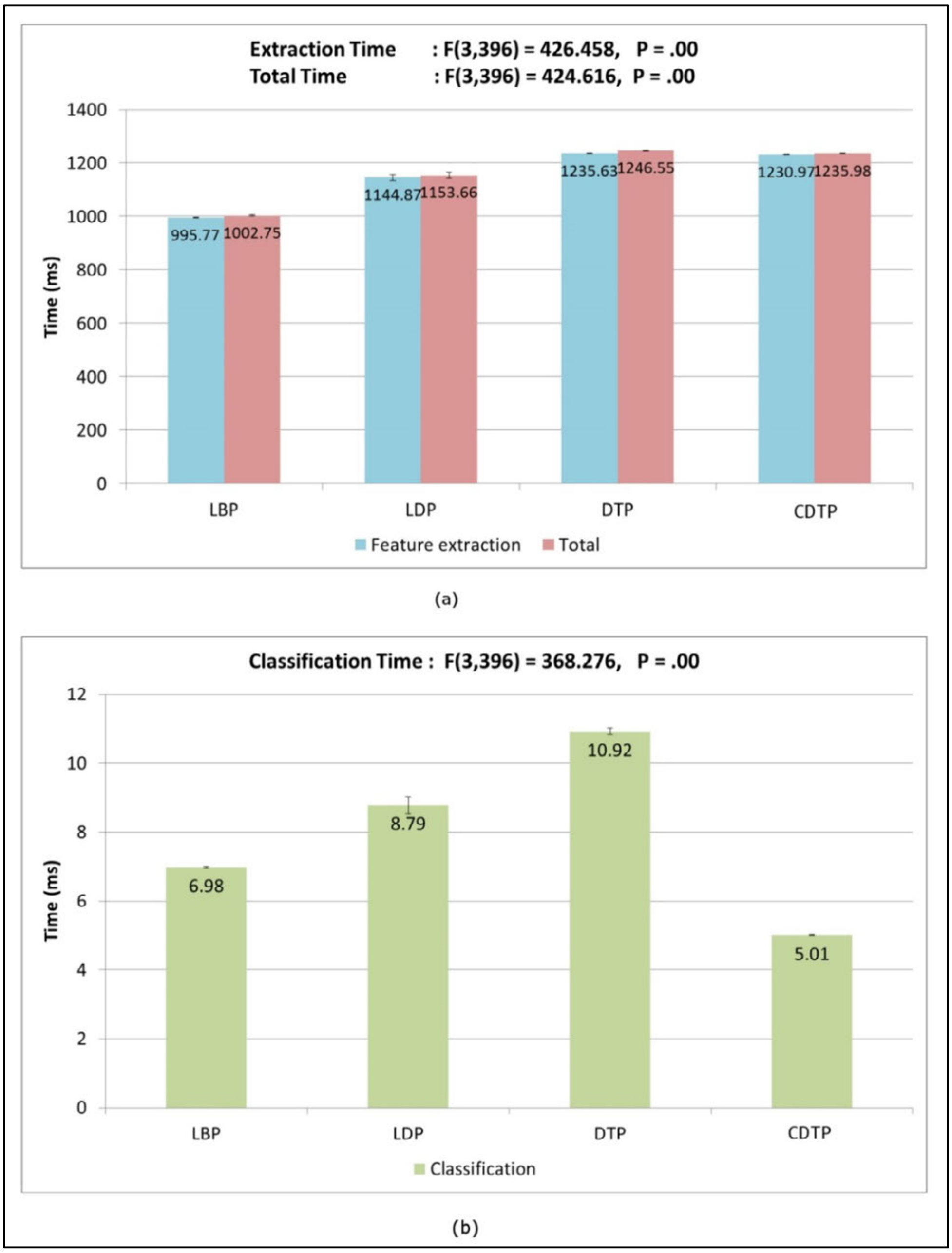


Figure 5.9 Statistical results of time. Mean and standard errors of execution time for feature extraction, classification and total times of LBP, LDP, DTP and CDTP: (a) Extraction and Total time, (b) Classification Time

5.3.5 Experiment 3: Testing partial face images and resolutions

This experiment aims to evaluate the robustness of my approach to recognize emotions using only partial face data, including upper and lower regions as shown in Figure 5.10(a) and 5.10 (b), respectively. For the upper region, the necessary facial components are eyes and eyebrows. The recognition results are shown in Table 5.3, that my approach is better than the others in JAFFE, KDEF and JAFFE-KDEF. For CK+, my approach is less accurate than DTP. While for the lower region which the necessary facial components are mouth and cheek, my approach is a bit better than the other for CK+ but it is less accurate than DTP for JAFFE dataset and LDP for KDEF and JAFFE-KDEF dataset as shown in Table 5.4. However, DTP approach that is the most accurate for JAFFE is worse than the others in KDEF, while LDP approach that is the most accurate for KDEF and JAFFE-KDEF is worse than my approach in JAFFE. These results suggest that the accuracy might vary depending on the datasets. Although my approach does not produce best result for all datasets but the accuracies of my approach are consistently ranked among the best methods for all datasets. This suggests that my approach is more stable than LDP and DTP.

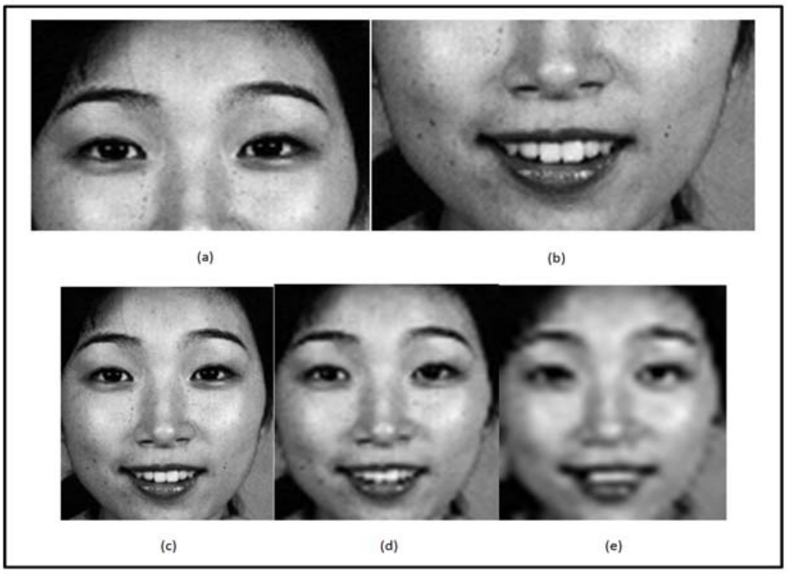


Figure 5.10 Images for robustness evaluation [110] (a) Upper region of face, (b) Lower region of face, (c) Normal resolution, (d) 60×70 Resolution, (e) 30×35 Resolution

Table 5.3 Results of emotion classification of upper region of face using linear SVM

Feature Extraction Methods	JAFFE Dataset	KDEF Dataset	JAFFE-KDEF Dataset	CK+ Dataset
LBP	52.06± 0.54	55.51± 0.32	45.53± 0.13	74.63 ± 0.08
LDP	64.24± 0.55	59.18± 0.34	50.50± 0.20	82.29 ± 0.06
DTP	73.87± 0.57	48.78± 0.33	44.53± 0.26	85.67 ± 0.08
CDTP-A	78.90± 0.38	49.16± 0.29	50.36± 0.21	84.61 ± 0.08
CDTP-B	69.50± 0.68	61.02± 0.39	56.19± 0.14	83.44 ± 0.07

Table 5.4 Results of emotion classification of lower region of face using linear SVM

Feature Extraction Methods	JAFFE Dataset	KDEF Dataset	JAFFE-KDEF Dataset	CK+ Dataset
LBP	41.80± 0.67	61.63± 0.36	53.61± 0.24	85.77 ± 0.08
LDP	58.20± 0.59	64.08± 0.27	58.60± 0.18	94.49 ± 0.08
DTP	74.57± 0.59	51.63± 0.36	46.51± 0.29	94.39 ± 0.08
CDTP-A	70.35± 0.65	52.65± 0.30	49.50± 0.30	94.58 ± 0.07
CDTP-B	60.52± 0.57	63.47± 0.27	58.30± 0.17	93.81 ± 0.08

Additionally, I also analyze the interaction between feature extraction approaches and regions of face on the accuracy using two-way ANOVA. In terms of interaction between factors, each approach performs similarly accuracy when classifies emotions from each region of face as shown in Figure 5.11. $F(8, 45) = 0.014$, $p = 1.00$]. The results show that the choice of regions do not affect the accuracy significantly [$F(2, 45) = 2.522$, $p =$

0.092]. Post-hoc comparison results using Scheffé's method indicated that using full face produced higher accuracy than from partial faces in both upper (Mdiff = 11.49, 95% CI[-1.97, 24.94], p = 0.109) and lower (Mdiff = 8.56, 95% CI[-4.90, 22.02], p = 0.283) regions of face. In summary, the result of this experiment indicates that my approach is more stable than the others to recognize emotions from both full and partial faces because CDTP-B (M = 71.91, SE = 3.83) is better than the others when considering the total average accuracies from all regions of face.

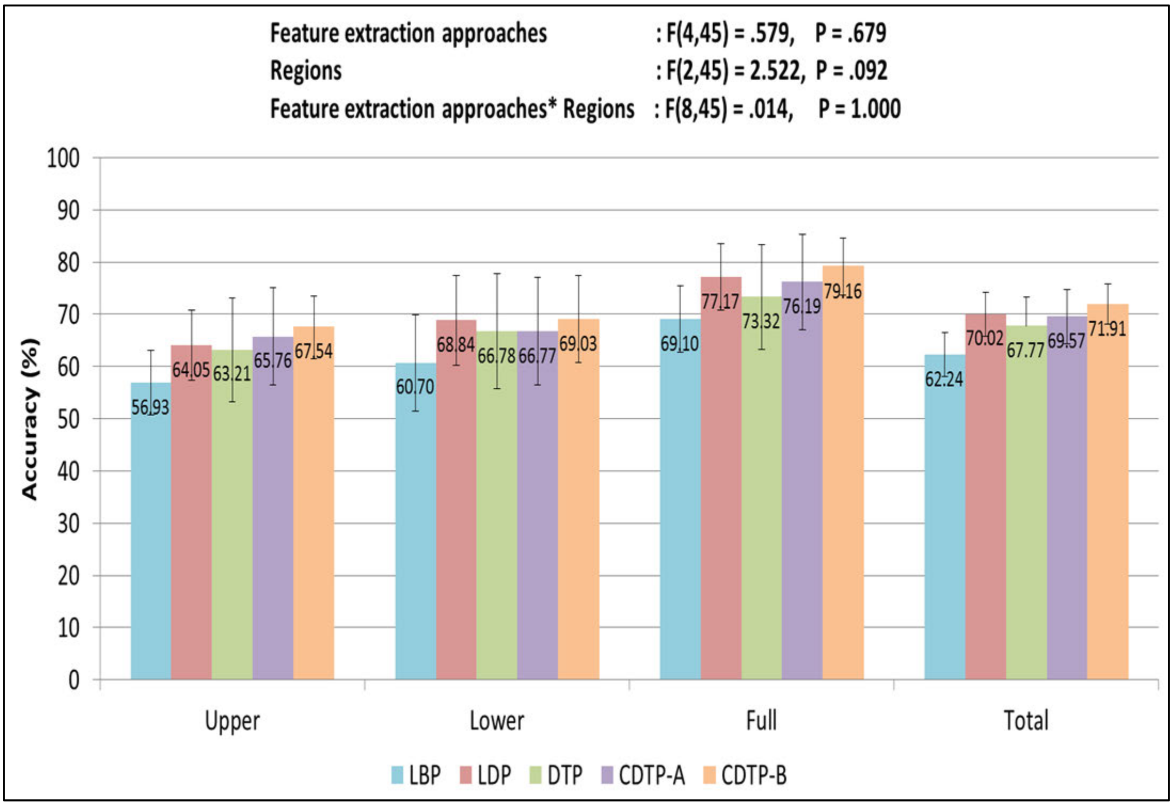


Figure 5.11 Analysis results between feature extraction approaches and region using two-way ANOVA.

I further test the robustness of different feature extraction methods when different image resolutions are used as shown in Figure 5.10(c)-5.10(e). The JAFFE, KDEF and JAFFE-KDEF and CK+ datasets are used in this test. The results as shown in Table 5.5 indicates that CDTP-B is the best approach for classifying emotions from normal resolution for JAFFE, KDEF and JAFFE-KDEF, but it is not the best for 60×70 resolutions. For 30×35 resolutions, no approach gets highest accuracy for these three datasets but CDTP-B consistently rank among the best methods. This is because it is more accurate than DTP, which is the best approach for JAFFE, when uses with KDEF and JAFFE-KDEF, and it also is more accurate than LDP, which is the best approach for KDEF, when uses with JAFFE and JAFFE-KDEF. This suggests that my approach is more stable than LDP and DTP for these three datasets. For CK+ dataset, my approach (CDTP-A) is more accurate than the others when classifying emotions from all image resolutions. In summary, CDTP-B is the best method for recognizing emotions from all image resolutions of all datasets as shown in Figure 5.12.

Table 5.5 Results of 7-class emotion expression classification of combined dataset in different image resolutions using linear SVM

Feature Extraction Methods	**JAFFE Dataset**			**KDEF Dataset**			**JAFFE-KDEF Dataset**			**CK+ Dataset**		
	Normal	**60×70**	**30×35**	**Normal**	**60×70**	**30×35**	**Normal**	**60×70**	**30×35**	**Normal**	**60×70**	**30×35**
LBP	62.9±0.4	61.1±0.8	59.6±0.5	66.3±0.2	56.0±0.6	53.5±0.3	59.5±0.2	60.30±0.2	50.9±0.2	87.70±0.11	89.25±0.10	86.64±0.08
LDP	76.6±0.5	72.8±0.7	72.2±0.9	72.6±0.2	67.6±0.3	62.4±0.3	64.6±0.2	65.13±0.2	60.0±0.2	94.87±0.04	96.23±0.06	91.10±0.07
DTP	84.1±0.5	86.0±0.6	78.5±0.7	62.2±0.2	62.9±0.2	60.8±0.2	51.5±0.2	63.15±0.2	62.6±0.2	95.45±0.05	95.36±0.06	91.32±0.05
CDTP-A	87.8±0.5	84.5±0.6	78.4±0.7	62.9±0.3	62.6±0.3	55.7±0.3	58.5±0.2	64.44±0.2	61.0±0.2	95.65±0.05	96.32±0.04	92.64±0.05
CDTP-B	78.8±0.6	85.9±0.5	77.5±0.8	73.7±0.1	66.9±0.2	61.8±0.3	69.6±0.2	63.57±0.2	63.4±0.2	94.58±0.07	95.65±0.08	92.55±0.07

Additionally, when I consider only the combined dataset (JAFFE-KDEF) which is more suitable for real-time emotion recognition that can detect emotion from various people with different nationalities, CDTP-B is less accurate than LDP but it is better than DTP when testing with 60×70 resolution dataset. For others image resolutions, CDTP-B gets the highest results. Thus, CDTP-B is general more robust than other approaches to classify emotions from combined dataset with normal resolution and lowest resolution sources (30×35). Therefore, CDTP-B is the best approach and more suitable than the others to detect emotions from combined dataset with different image resolution sources for real-time process. Additionally, in terms of resolution, there are no significant effects of feature extraction approaches [$F(4, 45) = 0.902$, $p = 0.471$], resolutions [$F(2, 45) = 0.508$, $p = 0.605$] and their interaction [$F(8, 45) = 0.035$, $p = 1.000$] on the accuracy as shown in Figure 5.12.

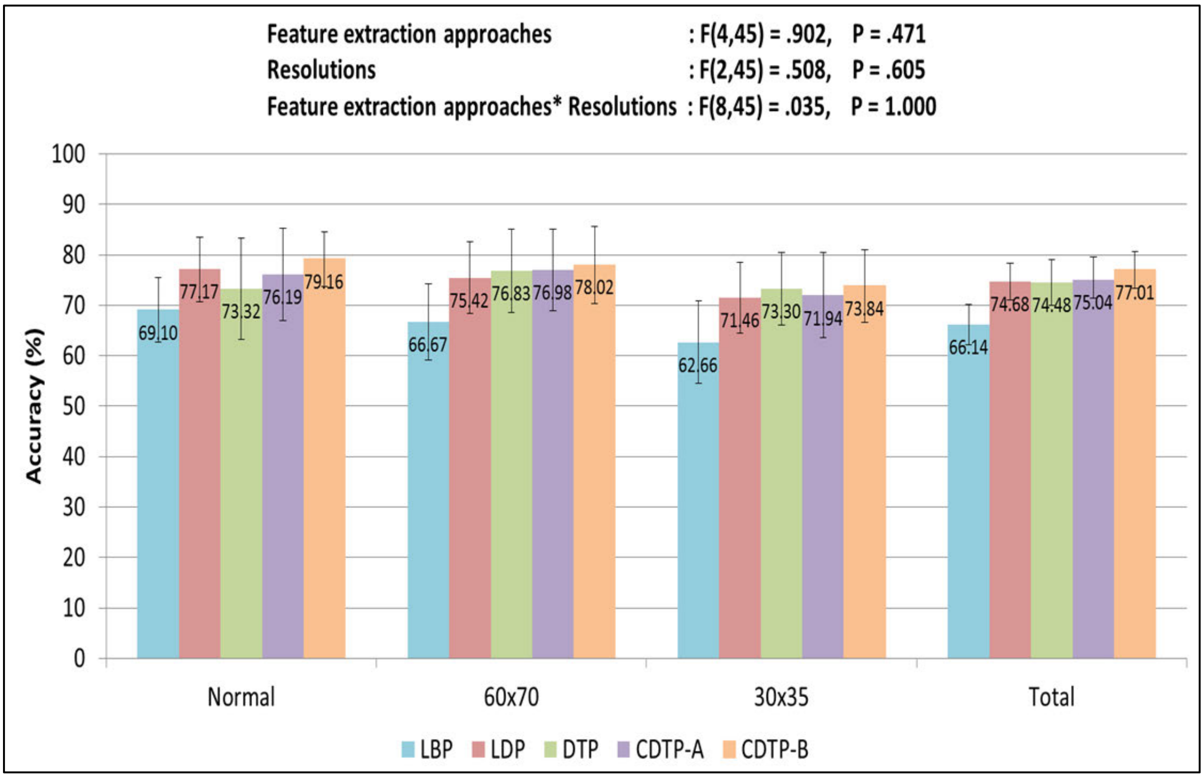


Figure 5.12 Analysis results between feature extraction approaches and resolution using two-way ANOVA.

5.4 Discussion

For emotion recognition by facial expression, feature extraction and classification methods are very important to produce higher accuracy and reliable performance. I design, implement and evaluate the feature extraction approach to improve the accuracy and performance of the emotion recognition by facial expression for my emotional healthcare system.

To improve the performance and accuracy of DTP, my approach (CDTP) decreases the size of DTP feature vector and reduces the feature redundancy in pattern representation in such the way that CDTP applies one's complement technique to calculate the representative value (unsigned decimal value) from positive and negative binary numbers to form the feature vector that represented changes in facial expressions caused by emotions. Moreover, it can assign a similar representative value for similar type of binary patterns as shown in Figure 5.4 and then collects it in the same bin of histogram (same feature) to construct a feature vector.

To classify emotions with the high accuracy and performance, I apply SVM for the emotion recognition by facial expression because SVM is a powerful classifier in fields of computer vision and image processing [106-107]. Since a facial image is a high-dimensional data and SVM is designed to handle high-dimensional data [115]. Furthermore, image features are sparse that contain a lot of zeroes. Linear SVM will ignore zero features when classifying emotions but for k-NN, their accuracies are lower than SVM because the feature vectors might be too noisy to be covered by similarity-based classification. Moreover, LBP, LDP and DTP adapted SVM as classifier [25, 56, and 27]. So I choose SVM for evaluating my approach and the results of ten-fold cross validation confirm that SVM can classify emotions with high accuracy. The effectiveness and efficiency are evaluated in three experiments using JAFFE, KDEF, JAFF-KDEF and CK+ datasets.

The first experiment, CDTP-A produces higher accuracy than other approaches when test with JAFFE and CK+ datasets, and CDTP-B is more accurate than others when evaluates with KDEF and JAFFE-KDEF. CDTP-A is nevertheless less accurate than LDP when evaluates with KDEF and JAFFE-KDEF dataset, and CDTP-B is less accurate than DTP when evaluates with JAFFE. This is because they apply different edge detection. When I compare the approaches that apply similar edge detection, I found that CDTP-A is more accurate than DTP and, CDTP-B is also better than LDP for all datasets. Moreover, I found that Robinson edge detection is more robust for JAFFE and CK+ datasets. Kirsch edge detection is more robust for KDEF and JAFFE-KDEF datasets. Thus, the appropriate edge detection may improve the accuracy because the edge detection is the first important step to emphasize the texture information such as curves, edges as well as spots. This experiment confirms that my approach is better than the others because of its highest average accuracy. From Figure 5.8, the confusion matrices of my approach shows that recognition of sad expressions is worse than the others because of more confusion with neutral, fear and happy expressions. However, it recognized neutral and happy expressions with the highest accuracies more than 93%, 80% and 84% for JAFFE, KDEF and JAFFE-KDEF respectively. For CK+ dataset, the recognition of neutral and sad expressions was worse than the others because the recognition of neutral expression was more confused with angry, sad, and surprised expressions, and the recognition of sad expression was more confused with neutral and angry expressions. However, it can recognize disgusted expression up to 100% and other emotions more than 98%

The second experiment aims to evaluate the performance of emotion recognition. The results indicate that DTP-based methods (DTP and CDTP) requires more computational times than LBP and LDP for considering both positive and negative binary patterns to address noise sensitivity issue which increases the accuracy. However, CDTP requires less computation times than DTP because it forms a feature vector and classifies emotion faster than DTP since the size of CDTP feature vector is decreased.

Finally, the third experiment is set to confirm the robustness of my approach by classifying emotions from partial face images or low resolutions. The results indicate that most approaches' performances are reduced when recognized emotions from partial face, low resolution images and combined dataset (JAFFE-KDFE). Nevertheless, CDTP-B is more stable and robust than the others because it always produces high accuracies for all cases.

In general, the emotion recognition using high-resolution images should be more accurate than using low-resolution images. However, the analysis results from ANOVA indicated that facial emotion recognition from images with different resolutions did not significantly affect the accuracy because even though the input images had different resolutions but extracted feature vectors were normalized to be suitable for emotion classification. Furthermore, the mean accuracies of DTP and CDTP-A showed that the emotion recognition using low-resolution images (60*70) was a little more accurate than using the normal one. Therefore, the emotion recognition from higher-resolution images might not be better than the lower one. This might be because of the selected edge detection technique, since the edge detection is the first important step to emphasize the texture information for emotion recognition. Moreover, I found that DTP and CDTP-A whose accuracies of low-resolution images were a bit better than the normal one, applied similar Robinson edge detection, but CDTP-B whose accuracy of normal-resolution images was better than the lower one, applied Kirsch edge detection. Thus, Robinson edge detection might be better than Kirsch edge detection to capture texture information from lower resolution images.

These evaluation results in total suggest that CDTP-B is more applicable than the others for real-time facial emotion recognition because it is more accurate to recognize emotions from combined datasets with full face and normal resolution images. Moreover, it is more robust than the others to recognize emotions from partial regions of face and low resolution images that might be occurred when recognizing emotions in real environment.

In summary, the evaluation results show that the classification performance and the accuracy of emotion recognition using CDTP are increased. This indicates that decreasing size of DTP feature vector effectively increases the classification performance. Furthermore, reducing redundancy in pattern representation by combining similar patterns into similar features of histogram can increase the accuracy. Therefore, I can expect CDTP to effectively recognize emotions.

5.5 Summary

I design Complementary Directional Ternary Pattern (CDTP), a new appearance-based feature extraction approach to fit the needs of my real-time emotion recognition by facial expressions. My approach improves the DTP by applying the one's complement technique to calculate the representative value from positive and negative binary numbers to generate their 128-level histograms in order to construct feature vectors with 256 lengths for emotions classification. This technique reduces the size of DTP feature vector by half and the size of CDTP feature vector was similar to feature vectors of LBP and LDP. Moreover, CDTP also reduces the redundancy in pattern representation by assigning similar representative value to similar patterns using the one's complement technique when it forms a feature vector. I then evaluate the performance of my approach with CK+, JAFFE, KDEF and JAFFE-KDEF datasets using linear SVM. The experimental results which evaluate the accuracy and the performance confirm that CDTP with Kirsch eight-directional edge detection and classify by one-against-one of linear SVM is more applicable than the others for using in real-time processing. It should be noted that, to accurately recognizing emotions, users need to clearly and exaggeratedly express the emotions on their faces. Otherwise, recognition of some similarity in facial expressions caused by different emotions, become confused and complicated to correctly classifying emotions. Therefore, I address these issues of the emotion recognition to improve its accuracy by applying ECG signal to recognize emotion, with the emotion recognition by facial expressions.

Chapter 6

Emotion recognition using ECG signal

Regard to confusion issues of emotion recognition by facial expression when recognizing emotions from similar facial expressions caused by different emotions, I design to integrate emotion recognition using ECG signal to address those issues and to increase the accuracy of the emotion recognition for the emotional healthcare system. To recognize emotion from ECG signal, I apply local pattern description methods to extract ECG emotional patterns features, since no research in emotion recognition using ECG signal applied this technique before. Additionally, this technique is popular to extract facial emotional feature because it produces high accuracy and high performance as I described in Chapters 2 and 5. This chapter presents the design, implementation and evaluation of the emotion recognition from ECG signal using pattern description methods.

6.1 K-nearest neighbor classification as background knowledge

K-nearest neighbor (k-NN) is a simple classification method [116] that learns and remembers all training data to build a feature space. Then, it classifies each testing data into one class by majority voting from its k-nearest trained neighbors. The k-NN still has some issues about choosing k-value. If the k-value is too small, the classification results might be sensitive to noise. On the other hand, if k-value is too large, too many trained neighbors with different classes might be included. In order to reduce the sensitivity of the choice of k-value, the k-NN applies weighted voting that the training data is weighted and vote by its distance [117]. In general, to define the appropriate k-value, the k-NN need to be run many time with different k-value and select the best k-value that produces the highest accuracy and performance [118].

6.2 Design and implementation

6.2.1 Workflow of emotion recognition

Real-time emotion recognition is divided into two parts: by facial expressions (Chapter 5) and by ECG signal. The workflow of real-time emotion recognition proceeds as follows (Figure 6.1).

1. Emotion recognition by facial expressions (Chapter 5)
    - Finds a user’s face from video frames (input).
    - Extracts facial features and normalizes them to form feature vectors.
    - Classifies user emotions into one of seven classes (neutral, joy, sadness, anger, disgust, fear and surprise) using a classifier generated from a training process.
    - Calculates percentage of each emotion for further analysis.
2. Emotion recognition by ECG signal
    - Records user’s ECG signals (input).
    - Removes noise.
    - Extracts ECG features and normalizes them to form feature vectors.

- Classifies user emotions into one of three classes (joy, anger and sadness) using a classifier generated from a training process.
- Calculates percentage of each emotion for further analysis.

After both parts produce results, the system combines them into a final result using a fusion method by equally averaging results from facial expression and ECG signal.

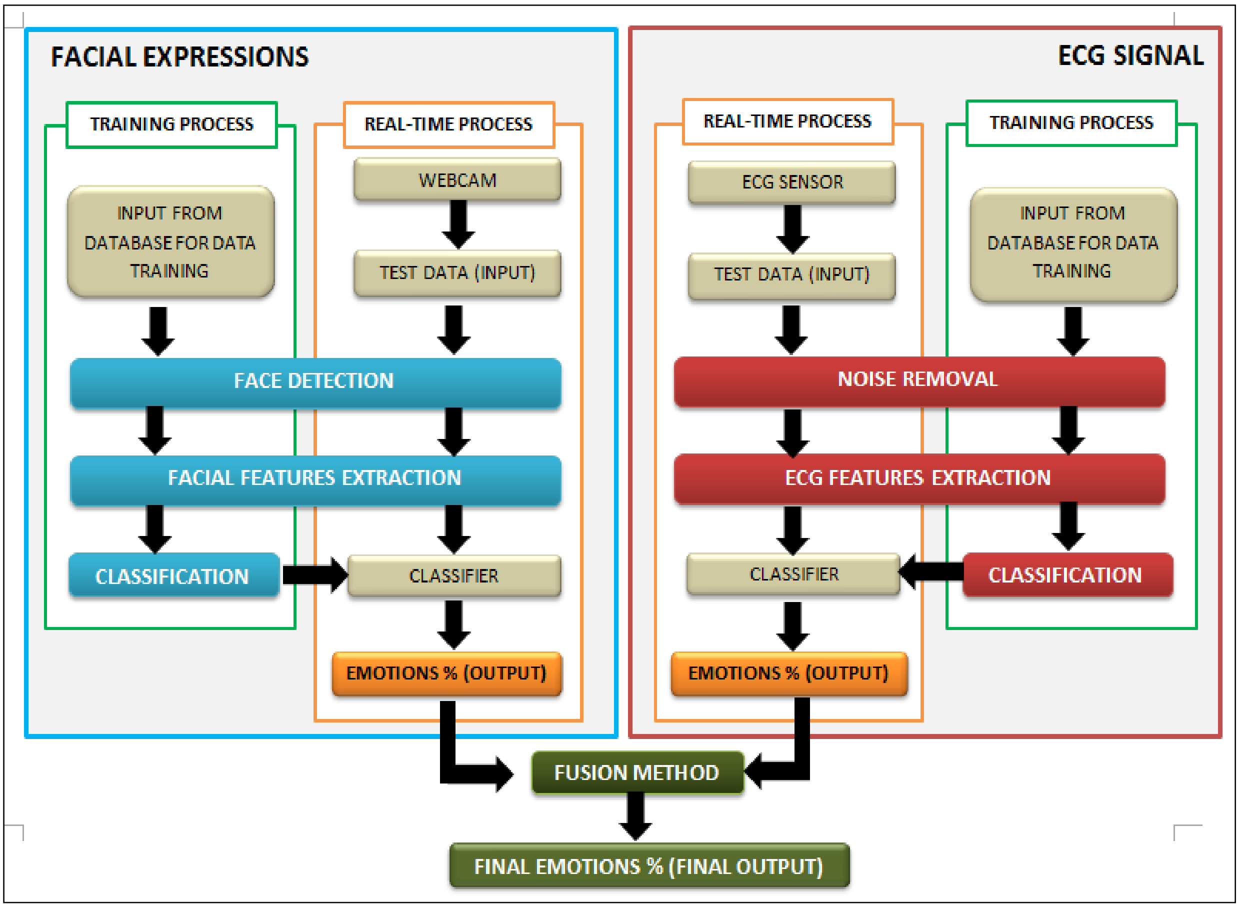


Figure 6.1 Emotion recognition from facial expression and ECG signal

6.2.2 The emotion recognition using ECG signal

To improve the accuracy and reliability of the emotion recognition by facial expressions in healthcare system, I applied the emotion recognition using ECG signals because ECG signal is directly affected by mental states, including emotions [48]. This section describes the preprocessing, feature extraction and classification of the emotion recognition using ECG signal.

6.2.2.1 Preprocessing of ECG signal

Since an ECG signal contains various noises, I applied a second-order IIR notch filter to remove power-line and narrow band noises due to 60 Hz caused by ECG sensor or user motion artifact based on specification of ECG sensor. Then, I applied Butterworth low-pass filter to remove higher frequency noise with 60 Hz cut-off frequency.

6.2.2.2 Feature extraction methods

The emotion recognition using ECG signal adapts local pattern description methods to extract the ECG features. LBP, LTP, LDP and DTP are successful methods to extract facial features with favorable performance. However, I select only LBP and LTP methods because LDP and DTP improved from LBP and LTP respectively by applying eight-directional edge detection and this edge detection is not necessary for ECG signal, since signal is already an edge. Therefore, these two methods are selected to extract features for the emotion recognition using ECG signal. Furthermore, I apply the one's complementary of my approach (CDTP) on Chapter 5 to LTP as CLTP in order to extract emotional feature for ECG signal. For facial expression, LBP, LTP and CDTP extract such facial features as spot, flat, edge, etc. For ECG signal, the ECG feature of directions and patterns are extracted. The process of LBP, LTP and CLTP (Figure 6.2) are as follows:

A. Local binary pattern

LBP [25] is a widely used texture description method in pattern classification for computer vision and image processing. I adapt an LBP operator to extract the ECG features as shown in Figure 6.2(a). First, it divides signal into frames. For each frame, it encodes the information of the ECG signal as binary numbers by comparing the center value with its four previous and subsequent values, and assigns them to 0 or 1. If those eight values exceed the threshold (center value), their binary codes are 1; if they are less than or equal to the threshold, their binary codes are 0. After that, it constructs a 256-level histogram for each frame as a feature vector.

B. Local ternary pattern

LTP [63] is an extension of LBP. The LTP operator (Figure 6.2(b)) also compares the center value with its four previous and subsequent values. However, unlike LBP that it assigns a three-value code (-1, 0, 1) instead of a two-value code (0 or 1). To assign the three-value code, this method considers ECG signals into two cases: positive and negative. If the ECG value is greater than or equal zero, it is a positive case. If the value is less than zero, it is a negative case. The following are the conditions for assigning ternary codes:

- For positive cases: if these eight values exceed a positive threshold (positive center value), their binary codes are 1; if they are less than the positive threshold, their binary codes are 0.
- For negative cases: if these eight values exceed a negative threshold (negative center value), their binary codes are 0; if they are less than the negative threshold, their binary codes are -1.

After that, these ternary values can be used to form positive and negative binary patterns. Then, a 512-level histogram is constructed as a feature vector.

After that, these ternary values can be used to form positive and negative binary patterns. Then, it constructs a 512-level histogram as a feature vector.

C. Complementary local ternary pattern

Complementary local ternary pattern (CLTB) (Figure 6.2(c)) improves from LTB by applying one's complement technique to reduce the size of LTP feature vectors by half, which is the same concept as CDTP (Chapter 5). Moreover, when encoding the ECG signal to build binary patterns using one's complement, the complementary binary patterns are collected in to the same bin during forming the 256-level of histogram as feature vector (Figure 6.2).

6.2.2.3 Classification method

To classify emotions from ECG signals, I applied the K-nearest neighbor (k-NN) [116]. K-NN is a simple classification method that learns, remembers and weights the ECG training data to build a feature space. When it classifies the ECG testing data, it considers the distance of the trained data and chooses the k-nearest one to be an appropriate emotional class from three classes (joy, anger, and sadness) by majority voting.

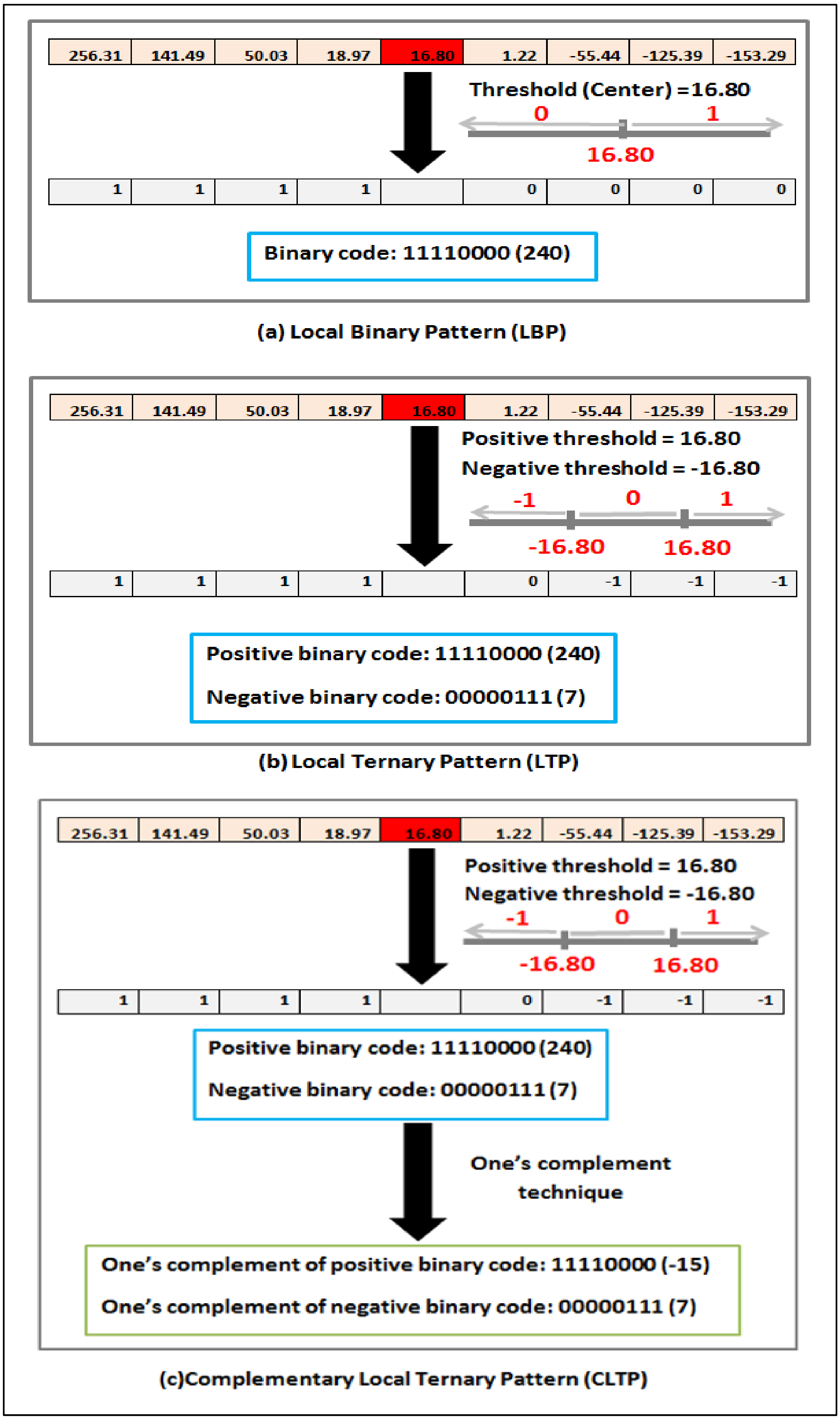


Figure 6.2 Feature extraction methods

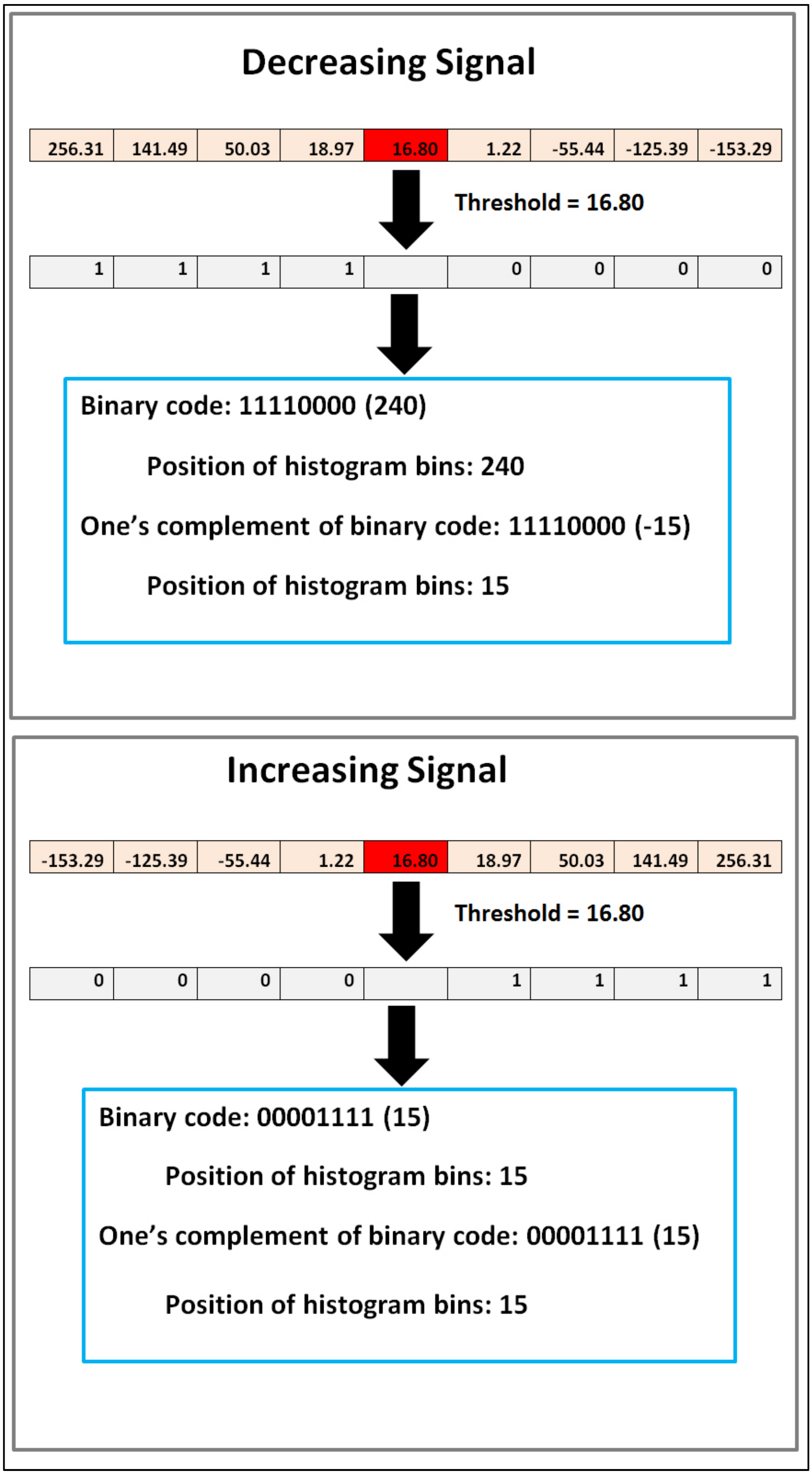


Figure 6.3 Decreasing and increasing signal extract features by CLTB

6.3 Performance evaluation

6.3.1 Objectives

I conduct the performance evaluation of the emotion recognition using ECG signal with LBP, LTP and CLTP feature extraction methods. The following are the objectives:

- To evaluate the effectiveness of LBP, LTP and CLTP for emotion recognition using ECG signals.
- To find the appropriate frame-length, frame-shift and k-value of k-NN that produces the highest accuracy for LBP, LTP and CLTP.

6.3.2 ECG signal dataset

The four emotion corpus of the Augsburg Biosignal Toolbox (AuBT) was applied for this evaluation [119]. This corpus contains four types of physiological data: electromyogram (EMG), electrocardiogram (ECG), skin conductivity (SC) and respiration change (RSP). The data were collected from one subject over 25 days while he listened to four kinds of music to induce emotions of joy, anger, sadness and pleasure. The data length was cropped into two minutes per session per emotion.

To evaluate emotion recognition using ECG signals, I selected only the ECG data (256 Hz) of three emotions (joy, anger and sadness) because they are also recognized by the emotion recognition using facial expressions and are useful for the emotional healthcare system. I used 75 pieces of data: 25 of joy, 25 of anger, and 25 of sadness).

6.3.3 Method

In this evaluation, LBP, LTP and CLTP extracted features from 3s, 5s, 10s and 15s frame-lengths with the 1.5s, 2.5s, 5s, and 7.5s frame-shifts respectively because these frame sizes are suitable for extracting emotions from ECG signal, since emotional details can be observed on ECG signal from 3s to 15s [120].

. K-NN is selected as the main classifier in this study because it is better than SVM as shown in Table 6.1, since the ECG features are not high-dimensional data comparing with image features and k-NN is better for low-dimensional data which is different from SVM that is better for high-dimensional data.

Table 6.1 Average recognition accuracies from LBP and LTP using AuBT dataset comparing between SVM and k-NN when k = 5

Classifier	Accuracy
SVM	66.27%
K-NN	80.10%

Ten-fold cross-validation is performed with a k-NN classifier where k is 5, 10 and 15 using RapidMiner Studio 5 to calculate the recognition accuracy [114]. The evaluated results are presented in next section.

6.3.4 Results

The performance evaluation of LBP, LTP and CLTP is shown on Table 6.2. The results are the accuracies when recognizing three emotions from ECG signals with different frame-lengths, frame-shifts, and classification using different k-values of k-NN. The results indicate that LBP LTP and CLTP produce the highest accuracies around 84.17%, 87.92% and 66.42% respectively, when they extract feature vectors from the input data with a 15s frame-length and a 7.5s frame-shift, and were classified using a k-NN classifier with k=5. However, CLTP which improved from LTP to reduce the feature vector size by applying one's complement technique produces less accuracy than LBP and LTP. Its accuracy is not high when compared with the others. Therefore, I will focus only the results of LBP and LTP. The confusion matrixes of LBP and LTP are shown in Tables 6.3 and 6.4 which explain the accuracies of emotion recognition between the actual and predicted classes and indicate the following results:

- LBP and LTP recognize sadness more accurately than joy and anger.
- LBP and LTP identically recognize joy.

- LTP recognizes anger and sadness better than LBP.
- For both LBP and LTP, recognizing anger is more confused with joy or vice versa.
- Recognizing anger and joy are less confused with sadness.

Furthermore, when I observe the histogram patterns of ECG emotional feature as shown in Figure 6.4, I found that the histogram features of joy, anger and sadness have different patterns. For example:

- The A pattern of joy gets a higher value than the A patterns of anger and sadness.
- The B and C patterns of each emotion have its own pattern.
- The D patterns of anger and sadness have a light difference; the D pattern of anger is lower than the D pattern of sadness. However, both are very different from the D pattern of joy.

In summary, the results from Tables 6.2-6.4 and Figure 6.4 confirm that local pattern description methods which successfully extracted facial emotional features can be adapted to extract ECG emotional features. LBP and LTP with a k-NN classifier can also recognize emotions from ECG signals with high accuracy.

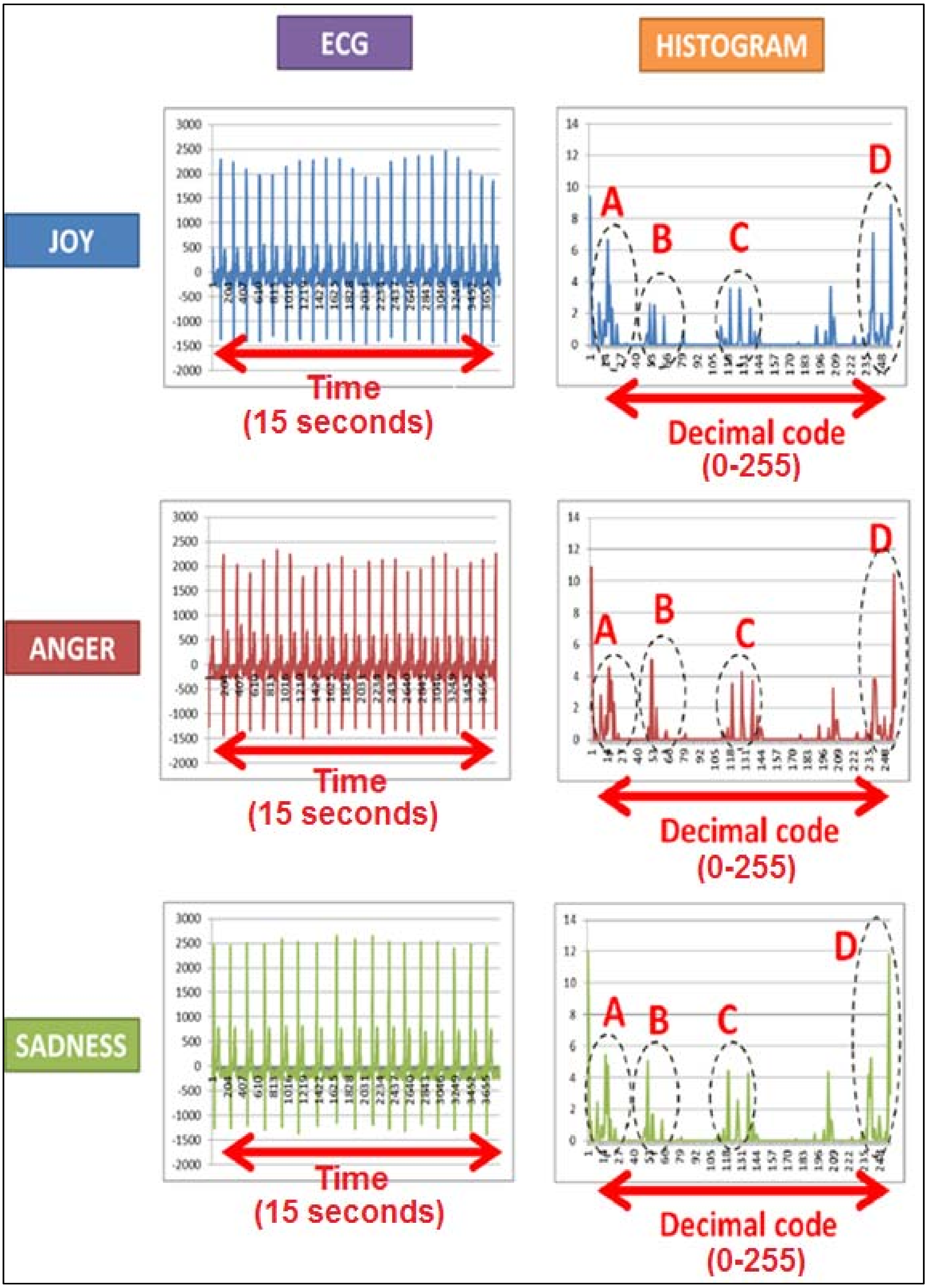


Figure 6.4 ECG signal and histogram features related to joy, anger and sadness.

Table 6.2 LBP, LTP and CLTP accuracies

Feature extraction	Frame-length (s)	Frame-shift (s)	k-value of k-NN	Accuracy
LBP	3	1.5	5	69.22% +/-0.025%
			10	69.33% +/-0.026%
			15	68.82% +/-0.028%
	5	2.5	5	79.11% +/- 0.050%
			10	77.81% +/- 0.044%
			15	75.53% +/- 0.053%
	10	5	5	82.61% +/-0.036%
			10	81.39% +/-0.042%
			15	77.89% +/-0.043%
	15	**7.5**	**5**	**84.17% +/- 0.095%**
			10	82.25% +/- 0.105%
			15	79.67% +/- 0.126%
LTP	3	1.5	5	73.17% +/- 0.021%
			10	72.52% +/- 0.016%
			15	71.43% +/- 0.020%
	5	2.5	5	80.06% +/- 0.038%
			10	79.33% +/- 0.042%
			15	77.67% +/- 0.054%
	10	5	5	84.56% +/- 0.050%
			10	83.94% +/- 0.048%
			15	80.56% +/- 0.062%

	15	**7.5**	**5**	**87.92% +/- 0.107%**
			10	86.42% +/- 0.121%
			15	83.75% +/- 0.100%
CLTP	3	1.5	5	52.10% +/- 0.022%
			10	51.22% +/- 0.025%
			15	51.83% +/- 0.024%
	5	2.5	5	54.83% +/- 0047%
			10	55.67% +/- 0.036%
			15	55.67% +/- 0.037%
	10	5	5	60.72% +/- 0.072%
			10	58.94% +/- 0.072%
			15	58.67% +/- 0.073%
	15	**7.5**	**5**	**66.42% +/- 0.107%**
			10	65.33% +/- 0.121%
			15	64.33% +/- 0.100%

Table 6.3 Confusion matrix of LBP (15s frame-length, 7.5s frame-shift, k=5)

TRUE / **PRED.**	**JOY (%)**	**ANGER (%)**	**SADNESS (%)**
JOY	**85.75**	14.75	4.00
ANGER	8.75	**77.50**	6.75
SADNESS	5.50	7.75	**89.25**

Table 6.4 Confusion matrix of LTP (15s frame-length, 7.5s frame-shift, k=5)

TRUE / PRED.	JOY (%)	ANGER (%)	SADNESS (%)
JOY	**85.75**	11.50	1.00
ANGER	11.25	**82.75**	3.75
SADNESS	3.00	5.75	**95.25**

6.3.5 Discussion

From the evaluation results (Tables 6.2-6.4), I obtained the following:

- The accuracies of LBP, LTP and CLTP are increasing when extracting emotional feature from 3s to 15s of frame-length with 1.5s to 7.5s of frame-shift. This indicates that a 15s of frame-length contains the emotional details more than the other frame-lengths.
- LBP, LTP and CLTP with K-NN classifier can classify emotions from ECG signal with high accuracy and k-NN is better than SVM because the ECG features are not high-dimensional data comparing to image features and it is better for low-dimensional data [117]. Additionally, the evaluation of emotion recognition by facial expression used around 10000-20000 features, but the evaluation of emotion recognition from ECG signal used only 256-512 of features. The emotion recognition by facial expression extracted more features than the emotion recognition from ECG signal. Therefore, the best classification methods can be different because number of feature and number of training example affect to accuracy and performance of classification methods. Thus, k-NN is more suitable for classifying ECG emotional feature and it has the highest accuracy when using k=5.
- CLTP is less accurate than LBP and LTP. Applying one’s complement technique to LTP might not be good for extracting feature vectors from ECG signal. Since, the ECG signal has the direction property that increasing and decreasing signal are different. However, when applying one’s complement technique to form ECG feature

vectors, it encodes and collects ECG signal with different direction into the same bin of histogram (Figure 6.3). This means that CLTP does not consider the direction of ECG signal which is necessary to distinguish between each emotion on extracted feature vector.

- LBP and LTP can recognize emotion from ECG with high accuracy. However, they have some confused recognition among each three emotions (joy, sadness and anger). Additionally, the confusion matrixes shows that joy and anger are more confused with each other, but they are less confused with sadness. This is because joy and anger are active emotions but sadness is passive emotion based on the Circumplex model of affect [121].
- LBP and LTP which are favorable local pattern description methods for facial emotion recognition can be adapted to extract ECG emotional features. They can recognize emotions with high accuracy. The reason of high accuracy is shown in Figure 6.4 that recognizing the specific patterns of different emotions on ECG signal is very difficult because it looks similar. However, after extracting pattern description features as histograms, different emotions produce different patterns on its histogram (A-D on Figure 6.4).
- LTP is more accurate than LBP because LTP considered both negative and positive ECG signal to generate both binary patterns and form its feature vectors which contain more emotional detail than feature vectors of LBP. Therefore, LTP is more suitable for real-time processing.

This performance evaluation applies the ECG dataset that was collected from only one subject. Further evaluation is needed to confirm the effectiveness of LTP for a person-independent in real-time processing.

6.4 Summary

The emotional healthcare system integrates emotion recognition from facial expressions and ECG signal to identify user emotions to provide appropriate services. This chapter explains the emotion recognition using ECG signal. To recognize emotions from ECG signal, I first apply the LBP and LTP which are the originally favorable local pattern description methods for emotion recognition by facial expressions. Additionally, I also apply CLTP which improved from LTP using one's complement technique. LBP, LTP and CLTP are evaluated their performance to recognize emotions from ECG signal. The results indicate that only LBP and LTP effectively extract ECG emotional features and produce high accuracy. Moreover, I can expect the emotion recognition from ECG signal to recognize emotions together with facial expression to produce higher accuracy and performance.

Chapter 7

Evaluation of a real-time prototype of healthcare system focusing on emotional aspect

Regarding Chapter 3 to Chapter 6, I separately design, implement and evaluate relaxation service and emotion recognition by facial expression and ECG signal. However, to apply the proposed system in real environment, I need to build a prototype of healthcare system focusing on emotional aspect in order to evaluate the effectiveness to recognize users' negative emotions and stress, and provide relaxation service. This chapter describes the prototype's workflow and experimental evaluation of user feelings when they experienced the prototype of healthcare system focusing on emotional aspect.

7.1 Heart rate variability as background knowledge

Heart rate variability (HRV) is a variation in time interval between each heart beat which calculated from inter-beat interval time series [122-123]. HRV can be analyzed in time domain or frequency domain. The time domain analysis indicates activity of Autonomic Nervous System (ANS), on the other hand frequency domain analysis indicates balance of ANS. When ANS is balance, HRV will be higher but when it is not balance, the HRV will be lower. HRV is favorable technique to measure and evaluate stress because stress can directly affect to ANS [124]. The following are the examples of HRV parameters which calculate from normal-to-normal (NN) interval to recognize stress.

A. Time domain parameters [124]

In general, time domain parameter measures normal-to-normal interval using statistical methods such as mean, standard deviation, and so on as below.

- SDNN is a standard deviation of the NN interval. The SDNN value depends on the length of measurement in which it is higher when the duration of measurement is long. SDNN is generally measure in millisecond. The higher SDNN indicates good ability to cope with stress. In contrast, lower SDNN indicates the risk of having chronic stress.
- RMSSD is a root mean square of standard deviation of NN interval. RMSSD generally indicates high frequency variation in heart rate.

B. Frequency domain parameters [123]

- Low frequency (LF) is the frequency band of the power spectrum between 0.04 Hz – 0.15 Hz. It generally indicates the activity of sympathetic.
- High frequency (HF) is the frequency band of the power spectrum between 0.15 Hz – 0.4 Hz. It generally indicates the activity of parasympathetic.
- LF/HF ratio indicates the balance of sympathetic and parasympathetic systems.

- Total power (TP) is a short-term total power of power spectrum between 0 Hz – 0.4 Hz. TP measurement reflexes to ANS. When TP is decreased, it indicates the chronic stress. The analysis meaning of TP is similar to SDNN in time domain.

7.2 The workflow of prototype of healthcare system focusing on emotional aspect

The workflow is shown in Figure 7.1.

- First, emotion recognition by facial expression and ECG signal recognizes and analyzes the user's current emotional state (1).
- If a user is experiencing negative emotions, the emotional healthcare system provides the re-design version of relaxation service with a breathing control application to perform deep breathing to decrease them and make him/her feel better (2). (The re-design was described in Chapter 4.5)
- After using the breathing control application for five minutes, the emotion recognition by facial expression and ECG signal re-analyses the user's emotions (3).
- If the user still has negative emotions, the relaxation service is provided again (2, 3).
- If the user feels better, he/she can choose to stop the service (4).

Next sections, I describe the experiments that evaluate the prototype of healthcare system focusing on emotional aspect.

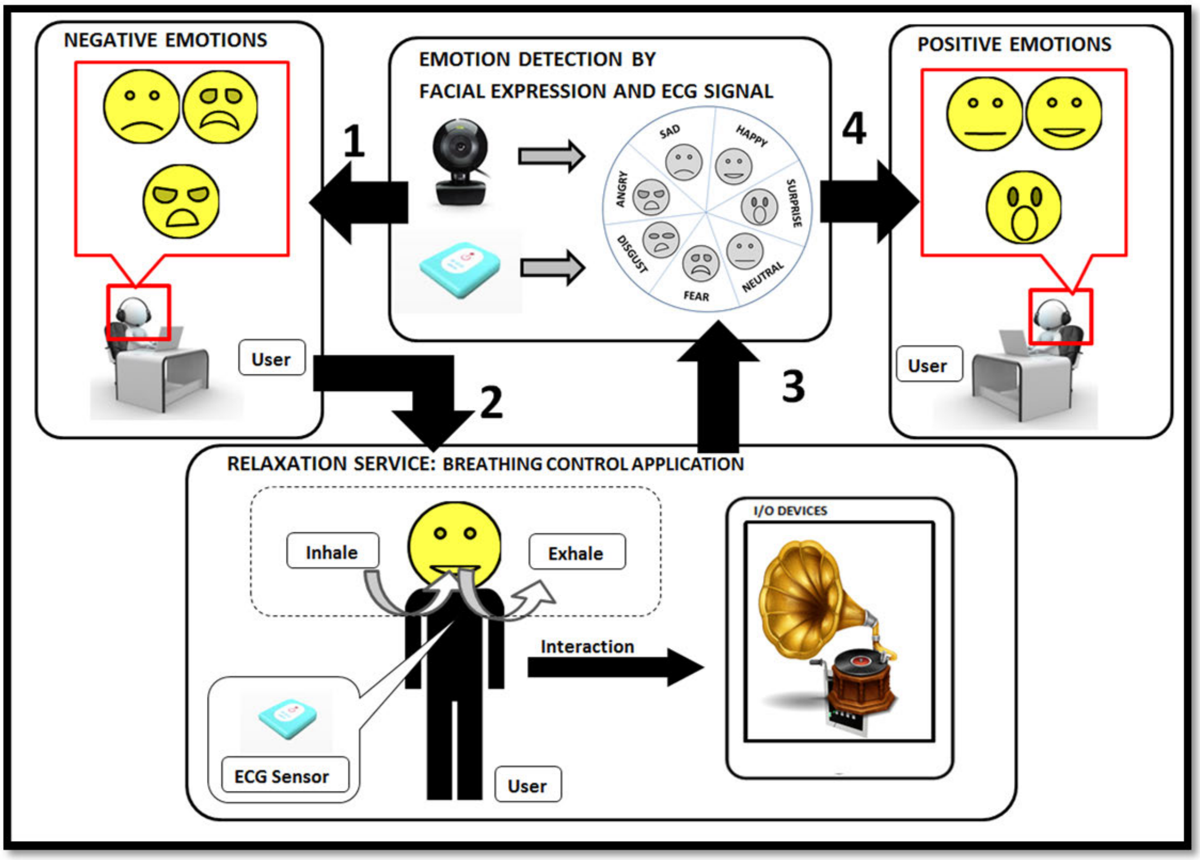


Figure 7.1 Workflow of the healthcare system focusing on emotional aspect [90-91]

7.3 Experiment 1: The relaxation service with emotion recognition by facial expression

7.3.1 Objective

Regarding the emotion recognition by facial expressions in Chapter 5, I have already experimentally evaluated it using only ten-fold cross validation. However, it has not been evaluated in real-time environment yet. Thus, this experiment aims to evaluate the real-time facial emotion recognition. Moreover, this experiment is set to evaluate the effectiveness of relaxation service to decrease negative emotions, since it has been confirmed to effectively decreased stress in Chapter 4; however, it has not been confirmed to reduce negative emotions yet. Therefore, I evaluate both real-time emotion recognition by facial expression and the relaxation service.

7.3.2 Experiment setup

Regarding the experimental results of emotion recognition by facial expressions in Chapter 5, the results suggest that CDTP feature extraction with Kirsch eight-directional edge detection is more applicable than the others for real-time processing. Thus, the real-time emotion recognition by facial expressions applies CDTP-B to extract facial feature and classifies emotions using one-against-one of linear SVM. In this experiment, the SVM is trained using JAFFE database.

JAFFE [110] consists of 213 images of ten females in six basic emotions with neutral faces. Another 60 Japanese subjects rated each image to assign emotion labels. I selected all 213 images (31 happiness, 29 disgust, 32 fear, 30 anger, 31 sadness, 30 surprise and 30 neutral) as training data.

Figure 7.2 shows the re-design version of relaxation service which is used in this experiment.

Figure 7.2 Re-design version of relaxation service

Figure 7.3 shows the experiment setup when participants perform the experiment.

Figure 7.3 Experiment setup

7.3.3 Participants

According to the emotion recognition by facial expression which applies the JAFFE dataset as training data, eight Japanese female subjects, aged 20-21, with no self-reported mental disorders, participated in this experiment. They are the fourth-year students at the Department of Information Science and Engineering, Shibaura Institute of Technology.

7.3.4 Experimental procedure

I experimentally evaluate the accuracy of emotion recognition by facial expressions and the effectiveness of breathing control application using the facial emotion recognition. To perform the experiment, I equally divide the participants into two groups (Groups A and B).

In the first session, Group A is presented with the positive stimulus first while Group B is presented with the negative stimulus. Then, if the participants have negative emotions, the breathing control application in relaxation service is presented to them. Afterward, there is a break before the subsequent session to reduce the effect of prior stimuli toward the next session. In the second sessions, Group A is presented with the negative stimulus, and Group B is presented with the positive stimulus. Then, if the participants have negative emotions, the relaxation service supports them again. After each session, the participants answered a forced-choice questionnaire (Figure 7.4) about their feeling for the video stimulus and relaxation service. These choices are the basic emotions, including happiness, sadness, disgust, surprise, anger, fear, and neutral. During the experiment, the system actively monitored the participant facial expression to predict the emotional response to the video stimuli and relaxation service. Multiple emotions are actively detected by the system.

The following is the steps to perform the experiment:

#1. Participants watched a video for one minute. Participants in Group A watched a happy video about cute cats, but Group B watched a bad video about pus treatment.

#2. While they watched the video, the emotion recognition recognizes their emotions by facial expression.

#3. If the emotion recognition determined that participants had neutral or positive emotions (happy, surprise), the participants skip to procedure #6. However, if the emotion recognition determined that the participants have negative emotions, the participants continue to procedure #4.

#4. The relaxation service provides a five-minute breathing control application with AR.

#5. After five minutes, the participants stopped controlling their deep breathing but continued to watch the AR object and listened to music for ten seconds. During ten seconds, the facial emotion recognition analyzed their emotion from their face.

#6. They answered the questionnaires about their real emotions (Figure 7.4).

#7. They rested for ten minutes.

#8. Then, they watched a video for one minute. Participants in Group A watched a bad video about pus treatment, but Group B watched a happy video about cute cats.

#9. Finally, the participants repeated procedures #2 to #6.

1. In an experiment, after watching the video, how did you feel?

   ☐ Happiness ☐ Sadness ☐ Disgust ☐ Surprise

   ☐ Anger ☐ Fear ☐ Neutral

2. In an experiment, after using the breathing control application, how did you feel?

   ☐ Happiness ☐ Sadness ☐ Disgust ☐ Surprise

   ☐ Anger ☐ Fear ☐ Neutral

Figure 7.4 Questionnaire about feelings (English version)

7.3.5 Results and Discussion

I experimentally evaluate two aspects: the accuracy of emotion recognition by facial expressions and the effectiveness of breathing control application. I also discuss the effectiveness of emotional healthcare system when integrating emotion recognition by facial expressions to provide relaxation service.

A. Accuracy of emotion recognition by facial expression

To evaluate the accuracy of emotion recognition by facial expressions, I compare the detected emotional results from the facial emotion recognition with the participants' real emotional results from questionnaires (Figure 7.4). Table 7.1 shows the average results of detected emotions from the facial emotion recognition and the result of participants' emotion from questionnaires. Even though several participants claim to feel neutral when using the breathing control application, the facial emotion recognition recognize feeling of disgust. Thus, the emotion recognition by facial expression has some confusion between disgusted and neutral expressions.

From the emotion recognition by facial expression with such confusion, the assessment of the system accuracy is done in emotion groups rather than individual emotion in order to increase the system robustness. Therefore, I separate the detected emotions into three groups.

- Neutral emotions: neutral and disgust
- Positive emotions: happiness and surprise
- Negative emotions: fear, anger and sadness

However, only the emotion having the highest probability score is used in identifying the recognized emotion group as show in Table 7.2. The results are then compared with the emotion groups obtained from the questionnaires. Instead of treating each groups as independent, error calculation is done in consideration of the closeness of the emotion groups. Therefore, the error can be calculated from rules in Figure. 7.5. Below is the explanation of the error rule:

- If the questionnaires are either negative or positive but the recognized results are the opposite, the error score of that case is 1.
- If the questionnaires are either negative or positive but the recognized results are neutral or disgust, and vice versa, the error score of that case is 0.5
- If the questionnaires and the recognized results are matched, then the error score is 0.

Table 7.1 Results of participants' emotions from emotion recognition and questionnaires

Partici-pant	Video	Emotional Results			
		Watch a video for one minute		After using a breathing control application for five minutes	
		Emotion recognition	Questionnaire	Emotion recognition	Questionnaire
P1	Happy	Dis 41.46% Ang 58.54 %	Happiness	Dis 87.50 % Fear 12.50%	Neutral
	Bad	Dis 95.12% Ang 4.88%	Fear	Dis 100.00%	Neutral
P2	Happy	Fear 78.05% Dis 21.95%	Neutral	Dis 100.00 %	Neutral
	Bad	Fear 100.00%	Fear	Fear 62.50% Dis 37.50%	Neutral
P3	Happy	Dis 100.00%	Happiness	Dis 100.00 %	Neutral
	Bad	Dis 46.34 % Fear 31.71 % Ang 19.51 %	Fear	Dis 75.00% Fear 25.00%	Neutral
P4	Happy	Fear 73.17% Dis 4.88 %	Happiness	Dis 25.00%	Neutral
	Bad	Fear 78.05% Dis 9.76%	Disgust Surprise Fear	Dis 25.00% Fear 62.50%	Sadness
P5	Happy	Dis 2.44% Ang 95.12%	Happiness	Dis 12.50% Ang 75.00%	Neutral
	Bad	Ang 97.56%	Neutral	Dis 25.00% Ang 75.00%	Fear
P6	Happy	Dis 100.00%	Happiness	Dis 100.00%	Disgust Neutral
	Bad	Dis 100.00%	Disgust	Dis 100.00%	Neutral
P7	Happy	Dis 100.00%	Happiness	Dis 50.00%	Neutral
	Bad	Dis 78.05% Ang 2.44%	Fear	Dis 100.00%	Neutral
P8	Happy	Ang 100.00%	Happiness	Ang 75.00% Dis 25.00%	Happiness
	Bad	Ang 73.17% Dis 19.51% Fear 7.32%	Surprise	Dis 50.00% Ang 50.00%	Happiness

*Dis = Disgust, Ang = Anger

**The summary of average results from emotion recognition by facial expression might not equal 100% because of face detection failure.

Table 7.2 Emotion groups based on emotion recognition by facial expression

Partici-pant	Video	Emotion Groups			
		While watching a video for one minute		After using a breathing control application for five minutes	
		System	Questionnaire	System	Questionnaire
P1	Happy	Negative	Positive	Neutral	Neutral
	Bad	Neutral	Negative	Neutral	Neutral
P2	Happy	Negative	Neutral	Neutral	Neutral
	Bad	Negative	Negative	Negative	Neutral
P3	Happy	Neutral	Positive	Neutral	Neutral
	Bad	Neutral	Negative	Neutral	Neutral
P4	Happy	Negative	Positive	Neutral	Neutral
	Bad	Negative	Negative	Negative	Negative
P5	Happy	Negative	Positive	Negative	Neutral
	Bad	Negative	Neutral	Negative	Negative
P6	Happy	Neutral	Positive	Neutral	Neutral
	Bad	Neutral	Neutral	Neutral	Neutral
P7	Happy	Neutral	Positive	Neutral	Neutral
	Bad	Neutral	Negative	Neutral	Neutral
P8	Happy	Negative	Positive	Negative	Positive
	Bad	Negative	Positive	Negative	Positive

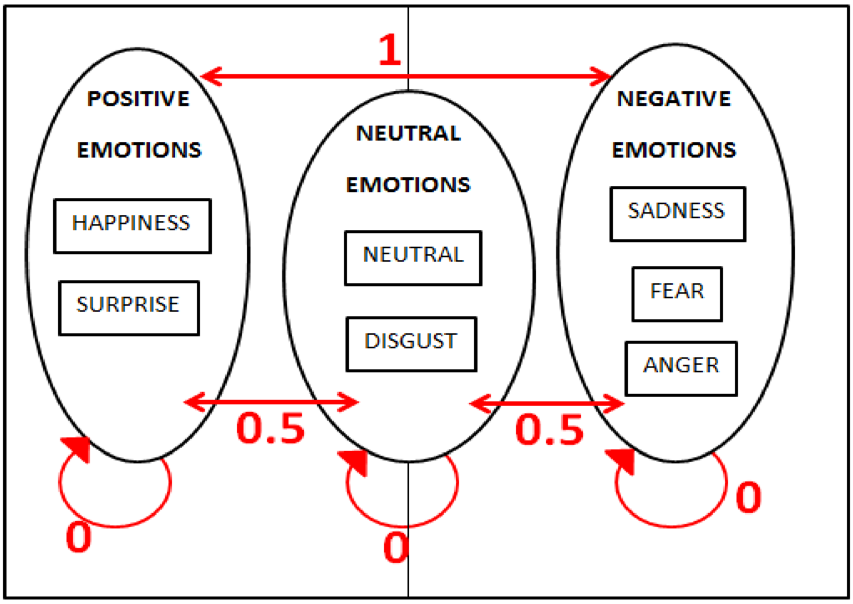


Figure 7.5 Rules for error calculation

From Table 7.2 and Figure 7.5, the error score is 12, but if errors occur in all cases, the score is 32. Thus, the error percentage is ((12*100)/32) = 37.5, and the accuracy is 100% – 37.5% (error) = 62.5%.

When comparing the accuracies from real-time processing and ten-fold cross validation (Table 7.3) to classify positive, negative and neutral emotions, the benchmark and comparison results indicate that neutral expression is more confused with negative expressions. Moreover, the emotion recognition's accuracy is reduced by around 84.05 % - 62.5% = 21.55%. This is because of some issues from real situation as follow:

- Emotion recognition by facial expressions applies a pattern recognition technique to recognize the emotions. Recognition of some similarity in facial expression is caused by different emotions, can become confused and complicate to correctly classifying emotions.
- I apply JAFFE dataset as training data. Since collected data is from only ten Japanese females, perhaps the variation in facial expressions is not in adequate to recognize emotions from several users in real-time processing.

- The emotion recognition has some limitations. It can only accurately recognize emotions if users clearly and exaggeratedly express them on their faces. However, since Japanese females generally hide their emotions and do not express more emotion on their face.

To address this limitation of the emotion recognition by facial expression in order to improve its accuracy I apply multimodal techniques (ECG signals) to detect emotion with the emotion recognition by facial expressions. I will explain in section 7.4.

Table 7.3 Confusion matrices of emotion recognition by facial expression using my approach (CDTP-B) and one against one of linear SVM when classifying JAFFE dataset

TRUE / PRED.	Positive (%)	Neutral (%)	Negative (%)	Accuracy
Positive	**93.44**	36.67	2.46	84.05% +/- 0.15%
Neutral	1.64	**10.00**	0	
Negative	4.92	53.33	**97.54**	

B. Effectiveness of breathing control application

From the results of facial emotion recognition in Table 7.2, several participants feel negative emotions (disgust, anger and fear) while watching the video. After watching it, I provide them with a breathing control application for five minutes. After that, their negative emotions (fear and anger) decrease. However, they still feel disgust, even though this feeling should be neutral because the emotion recognition by facial expressions has some confusion between disgusted and neutral expressions. In addition, the questionnaire results (Table 7.2) confirms that they feel neutral or happy after using the breathing control application. Therefore, the breathing control application effectively decreases some negative emotions and increases some neutral or positive emotions.

C. The emotional healthcare system with the real-time emotion recognition by facial expression

The purpose of the emotional healthcare system is to cope with users' negative emotion in daily life. To clarify the efficiency of the implemented emotion recognition by facial expression to detect negative emotions, I calculate how accurately it recognizes negative emotions for my facial emotion recognition as follows.

From Table 7.2 and Figure 7.5, only the cases of watching the bad video are considered to measure the effectiveness of the emotional healthcare system. The results show that the error score is 4.5 but if errors occur in all the cases, the errors pose to 16. Thus, the error percentage is ((4.5*100)/16) = 28.125%, and the accuracy is 100% – 28.125% (error) = 71.875%. Even though I failed to achieve 100% accuracy, but it adequately and reliably recognized negative emotions for the emotional healthcare system.

7.4 Experiment 2: Real-time emotion recognition by facial expression and ECG signal for the relaxation service

7.4.1 Objective

Regarding the results of experiment 1, the real-time emotion recognition by facial expression has some difficulty and confusion issues when users slightly express emotions on their face. To improve its accuracy, I apply emotion recognition by ECG in Chapter 6 to recognize emotions with emotion recognition by facial expression. I conduct this experiment to evaluate the real-time emotion recognition by facial expression and ECG signal.

7.4.2 Experiment setup

The real-time emotion recognition by facial expressions is set up similar to Experiment 1. The real-time emotion recognition using ECG signal applies LTP to extract ECG feature from 10s of frame-length and 5s of frame-shift. Actually, the result of benchmark test (Chapter 6) indicated that the emotion recognition by ECG signal produced the highest accuracy when extracting feature from 15s of frame-length with 7.5s of frame-shift. However, to apply it in real-time processing, high accuracy and high performance should be considered. I think extracting feature from 10s of frame-length does not consume time as much as extracting feature from 15s of frame-length. Moreover, the accuracies of 10s and 15s of frame-length are not significantly different. Therefore, it is better to extract ECG feature from 10s of frame-length and 5s of frame-shift in real-time processing because it can produce both high accuracy and performance. The emotion recognition by ECG signal classifies emotions using k-NN with k=5. In this experiment, the k-NN is trained using AuBT database (Chapter 6.3.3).

7.4.3 Participants

According to the emotion recognition by facial expression which applies the JAFFE dataset as training data, eight Japanese female subjects, with no self-reported mental disorders, participated in this experiment. They are the fourth-year students and staff at the Department of Information Science and Engineering, Shibaura Institute of Technology.

7.4.4 Experimental procedure

I experimentally evaluate the accuracy of real-time emotion recognition by facial expressions and ECG signal. To perform the experiment, I equally divide the participants into two groups (Groups A and B). The experimental procedure is similar to the previous one.

In the first session, Group A is presented with the positive stimulus first while Group B is presented with the negative stimulus. Afterward, there is a break before the subsequent session to reduce the effect of prior stimuli toward the next session. In the second sessions, Group A is presented with the negative stimulus, and Group B is presented with the positive stimulus. After each session, the participants answered a forced-choice questionnaire (1st question of Figure 7.4) about their feeling for the video stimulus. These choices are the basic emotions, including happiness, sadness, disgust, surprise, anger, fear, and neutral. Finally, they answered the second questionnaire about acceptable error of emotion recognition (Figure 7.6). During the experiment, the system actively monitored the participants' facial expression and ECG signal to predict the emotional response to the video stimulus. Multiple emotions are actively detected by the system.

The following is the steps to perform the experiment:

#1. Participants watched a video for three minutes. Participants in Group A watched a happy video about baby, but Group B watched a bad video about daughter who has dumb father.

#2. While they watched the video, the emotion recognition recognizes their emotions by facial expression and ECG signal.

#3. They answered the questionnaires about their real emotions (1^{st} question of Figure 7.4).

#4. They rested for ten minutes.

#5. Then, they watched a video for three minute. Participants in Group A watched a bad video, but Group B watched a happy video.

#6. After that, the participants repeated procedures #2 to #3.

#7. Finally, the participants answered the questionnaires in Japanese about acceptable error of emotion recognition (Figure 7.6).

Description of our emotional healthcare system in the experiment:

While the participants watched the pictures, the emotion detection by facial expression and ECG recognized their face and ECG signal to analyze their emotions. If the participants have some negative emotions such as fear, the system then provided breathing control application. However, the emotion detection by facial expression and ECG might produce two errors:

1. The participants have negative emotions but the system detects as positive emotions. Then, the system will not do anything to make the participants feel better.
2. The participants have positive emotions but the system detects as negative emotions. Then, the system will provide the application to let the participants deeply breathe to relax them.

Therefore, we set the following questions to ask you about the errors.

	Questions	Ranking 5	4	3	2	1
Which error is better for our emotional healthcare system?						
1	If you have positive emotions but the application detects your emotion is negative. **This one is better.**					
	Reason:					
2	If you have negative emotions but the application detects your emotion is positive. **This one is better.**					
	Reason:					

Figure 7.6 Questionnaire about system errors (English version)

7.4.5 Results and Discussion

This experiment, I evaluate two aspects: the accuracy of the emotion recognition by facial expressions and ECG signal, and the system error that is acceptable for emotional healthcare system.

A. Accuracy of emotion recognition by facial expression and ECG signal

To evaluate the accuracy of emotion recognition by facial expressions and ECG signal, I compare the detected emotional results from the facial expression and ECG signal with the participants' real emotional results from questionnaires. Table 7.4 shows the average results of the detected emotions from the emotion recognition by facial expression and ECG signal, and the participants' emotions from questionnaires.

According to the previous experimental results, the emotion recognition by facial expression has some confusion between disgusted and neutral expressions. The assessment of the system accuracy is done in emotion groups rather than individual emotion (similar to the previous experiment). The detected emotions are divided into three groups.

- Neutral emotions: neutral and disgust
- Positive emotions : happiness and surprise
- Negative emotions: fear, anger and sadness

From the previous experiment, only emotion having the highest probability score is used in identifying the recognized emotion group. However, to detect symptom of negative feeling for providing the relaxation service, I set standard probability thresholds of each emotion for facial expression and ECG signal in this experiment as shown in Table 7.5. If one of recognized negative emotions has higher score than its threshold, the emotion group is negative group. The reason why I apply this threshold is to increase the effectiveness for detecting negative emotions. To combine the recognition results of emotion recognition by facial expression and ECG signal, I apply OR rule of probability in which negative emotions are "1", neutral and positive emotions are "0" but if there are only neutral and positive emotions, the positive emotions are "1" and neutral emotion is "0" as shown in Table 7.6

The combined results between facial expression and ECG signal are shown in Table 7.7. Then, I compared the results with the emotion groups obtained from the questionnaires using error calculation (Figure 7.5). Table 7.8 shows error score that obtains using error calculation. Then, I calculate the error percentage by dividing error score with maximum error score based on questionnaire results. The maximum error score of all cases, positive case and negative case will be 16, 7 and 8 respectively. After that, the error percentage and accuracies of emotion recognition by facial expression and ECG signal are also calculated and shown in Table 7.8.

Table 7.4 Results of participants' emotions from emotion recognition and questionnaires

Participant	Video	After watching videos			Appropriate Error	
		System		Questionnaire	#1 Pos->Neg	#2 Neg -> Pos
		Face	ECG			
P1	Funny	Dis 97.56%	Hap 97.18% Sad 2.81%	Happy	2	4
	Sad	Dis 85.37% Ang 9.76%	Hap 65.74% Sad 34.26%	Sad, Fear, Anger		
P2	Sad	Dis 12.50%	Hap 82.84% Sad 17.16%	Sad	4	2
	Funny	Dis 21.95%	Hap 92.80% Sad 7.20%	Happy		
P4	Sad	Dis 95.12% Fea 2.24%	Hap 56.04% Sad 43.96%	Sad	4	3
	Funny	Dis 82.93% Fea 14.63%	Hap 3.47% Sad 96.53%	Happiness		
P5	Funny	Fea 95.00% Dis 5.00%	Sad 65.22% Hap 34.78%	Happy	4	2
	Sad	Fea 92.68%	Hap 66.67% Sad 29.17% Ang 4.16%	Sad		
P6	Sad	Dis 40.00%	Sad 70.00% Hap 30.00%	Sad	4	2
	Funny	Dis 67.56%	Sad 64.71% Hap 35.29%	Happy		
P7	Funny	Dis 72.68%	Sad 82.61% Hap 17.39%	Neutral	5	2
	Sad	Dis 31.56%	Sad 52.63% Hap 42.11% Ang 5.26%	Sad		
P8	Sad	Dis 100.00%	Sad 55.00% Hap 45.00%	Sad	3	4
	Funny	Dis 92.86%	Hap 73.68% Ang 21.06% Sad 5.26%	Happy		
P9	Funny	Dis 75.00% Fea 22.50%	Hap 76.92% Sad 23.08%	Happy	2	4
	Sad	Dis 78.04%	Hap 95.00% Sad 5.00%	Sad		

* Dis = Disgust, Ang = Anger, Fea = Fear, Hap= Happy

**The summary of average results from emotion recognition by facial expression might not equal 100% because of face detection failure.

Table 7.5 Standard probability thresholds for facial expression and ECG signal

Emotion recognition	Detected emotions	Standard probability threshold
Facial expression	7 emotions	100/7 = 14.29
ECG signal	3 emotions	100/3 = 33.33

Table 7.6 OR rule of probability for determining emotion groups

Result from emotion recognition by facial expression		Result from Emotion recognition by ECG signal		Assigned emotion groups
Negative	**1**	Positive	0	Negative
		Neutral	0	Negative
Positive	**0**	Negative	1	Negative
	1	Neutral	0	Positive
Neutral	**0**	Negative	1	Negative
		Positive	1	Positive

Table 7.7 Emotion groups based on emotion recognition

Participant	Video	After watching videos				Appropriate Error	
		System			Questionnaire	#1 Pos->Neg	#2 Neg -> Pos
		Face	ECG	Face + ECG			
P1	Funny	Neu	Pos	Pos	Pos	2	4
	Sad	Neu	**Neg	**Neg	Neg		
P2	Sad	Neu	Pos	Pos	Neg	4	2
	Funny	Neu	Pos	Pos	Pos		
P4	Sad	Neu	**Neg	**Neg	Neg	4	3
	Funny	Neg	Neg	Neg	Pos		
P5	Funny	Neg	Neg	Neg	Pos	4	2
	Sad	Neg	Pos	Neg	Neg		
P6	Sad	Neu	Neg	Neg	Neg	4	2
	Funny	Neu	Neg	Neg	Pos		
P7	Funny	Neu	Neg	Neg	Neu	5	2
	Sad	Neu	Neg	Neg	Neg		
P8	Sad	Neu	Neg	Neg	Neg	3	4
	Funny	Neu	Pos	Pos	Pos		
P9	Funny	Neg	Pos	Neg	Pos	2	4
	Sad	Neu	Pos	Pos	Neg		

*Pos= Positive, Neu= Neutral, Neg=Negative

** Results from using thresholds to detect the symptom of negative feeling

Table 7.8 Error and accuracy of emotion recognition by facial expression and ECG signal

Emotion groups		All	Positive	Negative
Error score	Facial expression	8.5	5	3.5
	ECG	6.5	3	3
	Facial expression and ECG	6.5	4	2
Error percentage	Facial expression	53.13	71.43	43.75
	ECG	40.63	42.86	37.50
	Facial expression and ECG	40.63	57.14	25.00
Accuracy (100%- Error Percentage)	Facial expression	46.88	28.57	56.25
	ECG	59.38	57.14	62.50
	Facial expression and ECG	59.38	42.86	75.00

From the recognized results, the emotion recognition by facial expression still confuses when recognizing emotions that effect to its accuracy. It can produce only 46.88% of accuracy. On the other hand, the accuracy of emotion recognition by ECG signal is 59.38%. Therefore, the emotion recognition by facial expression is less accurate than the emotion recognition by ECG signal because participants slightly express emotions on their face that caused some confusion, but ECG signal can capture their internal emotions.

When combining the results from facial expression and ECG signal, the combined results indicate that the emotion recognition from facial expression and ECG signal can increase the accuracy of emotion recognition by facial expression. Even though, the accuracy of emotion recognition by facial expression and ECG signal is only 59.38% but it can recognize negative emotions and symptom of negative feeling around 75%. Moreover, real-time emotion recognition by facial expression can recognize emotions around 1 second/frame. However, emotion recognition from ECG signal can recognize emotions around 10 seconds, since emotional details can be observed on ECG signal from 3s -15s. Therefore, real-time emotion recognition is feasible to recognize emotions for 10 seconds. This real-time emotion recognition by facial expression and ECG signal is beneficial for the relaxation service and the emotional healthcare system to analyze negative emotions to provide assistance.

B. System errors in the emotional healthcare system

Another purpose of this experiment is to get participants' opinion on errors that might be occurs in the emotional healthcare system. I evaluate using the questionnaire in Figure7.6. There are two types of errors that the emotion recognition by facial expression and ECG signal can produce. These errors are reflex to the efficiency of emotional healthcare system.

Error #1: The participants have positive emotions but the system detects it as negative/neutral emotions. Then, the system will provide the application to let them deeply breathe to relax them.

Error #2: The participants have negative emotions but the system detects it as positive/neutral emotions. Then, the system will not do anything to make them feel better.

The questionnaire results which analyzed using independent t-test (Table 7.9) show that there is no significant difference between error #1 and #2. However, mean of error #1 is higher than error #2 which indicates many participants can accept error #1 more than error #2 because if users have negative emotions but the emotion recognition recognizes as positive emotions (error #2), then the emotional healthcare system will not provide any assistance to make them feel better. This means that the system is inefficient. Therefore, the error #1 is acceptable.

The emotion recognition by facial expression and ECG signal has not reach 100% of accuracy yet. It still has errors. From Table 7.8, it can detect negative case better than positive case. Therefore, the results confirmed that this real-time emotion recognition is efficient enough to recognize negative emotion and symptom of negative feeling for the emotional healthcare system to provide assistance.

Table 7.9 T-test comparison about system errors between error#1 and error#2

Acceptable error	Mean	S.D.	t	P-Value
Error #1 : Positive emotions are detected as negative emotions	3.50	1.069	1.213	1.000
Error #2 : Negative emotions are detected as positive emotions	2.88	.991	1.213	

7.5 Improvement of the prototype with emotion recognition and stress detection

7.5.1 Issues of the prototype

From the prototype and its evaluation, I found that the prototype of emotional healthcare system still has two issues as follows:

- After providing the relaxation service, the emotion recognition by ECG signal cannot recognize neutral emotions to confirm the effectiveness of relaxation service because the training dataset that I applied for real-time emotion recognition using ECG signal, contain only three emotions (happiness, sadness and anger). To recognize neutral emotions is very necessary, since many participants feel neutral after experiencing the relaxation service based on the results of experiment 1. To solve this issue, I need to apply another dataset that contains ECG data of neutral emotion in order to confirm the effectiveness of relaxation service.
- Since, the purpose of relaxation service is to decrease negative emotions and stress; however, this emotional healthcare system can recognize only negative emotions but it cannot recognize stress in order to provide relaxation service. To solve this issue, the stress detection should be applied in the prototype.

Therefore, to increase the effectiveness of the prototype and address the above issues, integration of stress detection to the prototype is more preferable because currently, the prototype of emotional healthcare system can detect neutral emotion from users' facial expression but it cannot recognize stress. Thus, integration of stress detection will increase the effectiveness and efficiency of the prototype to provide the relaxation service, since stress is one of mental problems that the healthcare system should be taken into account as well. Furthermore, it might be able to solve the confusion issue of emotion recognition by facial expression. Next section, I explain an improvement of the prototype of emotional healthcare system and I conduct the third experiment to evaluate the prototype to recognize emotions from facial expression and stress from ECG signal in order to provide the relaxation service.

7.5.2 Workflow of improvement of the prototype

Workflow of the improved prototype is shown in Figure 7.7.

- First, emotion recognition by facial expression and stress detection by ECG signal recognize and analyze the user's current emotional state and stress level respectively (1).
- If a user is experiencing negative emotions or stress, the emotional healthcare system provides the re-design version of relaxation service with a breathing control application to perform deep breathing to decrease negative emotions or stress, and make him/her feel better (2).
- After using the breathing control application for a five minute, the emotion recognition by facial expression and stress detection by ECG signal re-analyze the user's emotions and stress (3).
- If the user still has negative emotions or stress, the relaxation service is provided again (2, 3).
- If the user feels better, he/she can choose to stop the service (4).

From the workflow, I improve the effectiveness and efficiency of the prototype by integrating the real-time stress detection from ECG signal to recognize users' stress together with facial emotion recognition to provide the relaxation service. The stress detection applies the standard deviation of the normal-to-normal interval (SDNN) of heart rate variability (HRV) to detect stress because SDNN of HRV is one effective technique for recognizing stress [124]. Additionally, SDNN can indicate the ability of release stress. In general, SDNN is derived from the average of all standard deviations of RR interval in every five minutes. According to the HRV analysis system, when humans have stress, SDNN will be less than 35 ms, but if SDNN is greater than 35 ms, it indicates that humans have a good ability to release stress. However, this SDNN value is calculated from the average of standard deviation of each five-minute analysis within 24 hours. It is suitable for long-time analysis. To detect stress in short period, the appropriate SDNN threshold should be used. Thus, I set preliminary experiment to determine the appropriate SDNN threshold for short-time analysis in the next section.

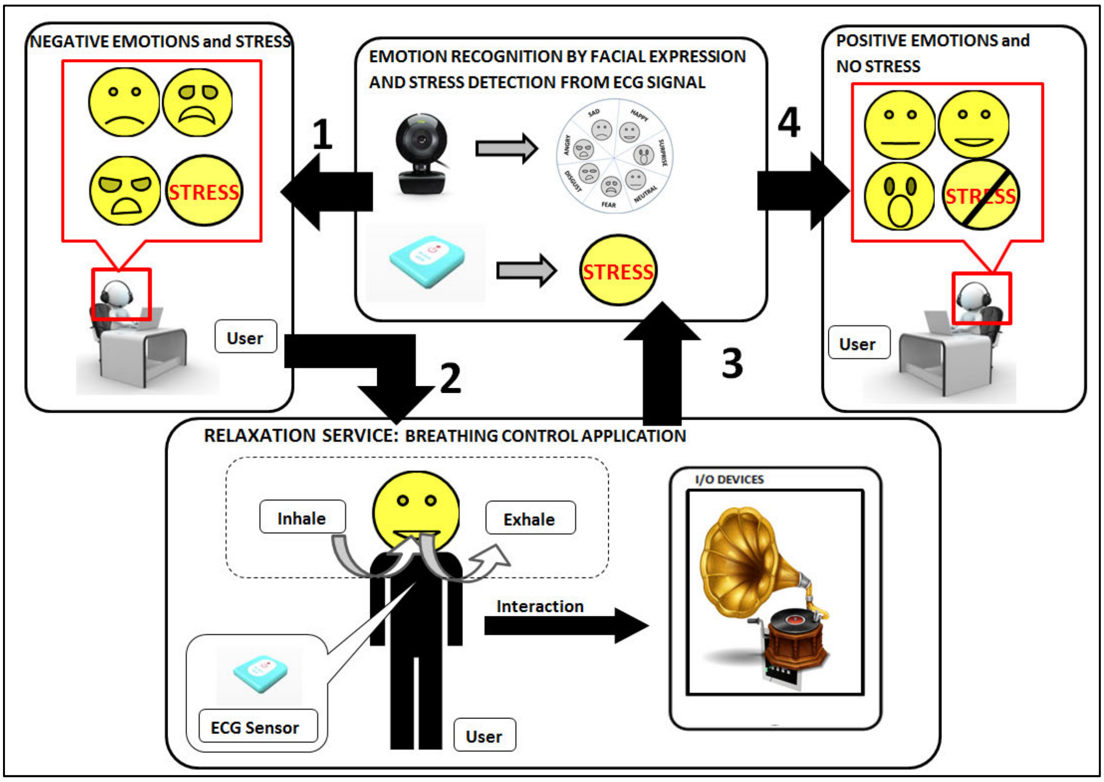


Figure 7.7 Workflow of the improved prototype [90-91]

7.5.3 Preliminary experiment: SDNN threshold

This preliminary experiment is conducted to set the SDNN value that is suitable for short-time analysis. The participants are two Japanese and Thai females. The experiment procedure is as below.

#1. The participants watch 30 negative images from International Affective Picture System (IAPS) [125] for four minutes while their ECG signals are recorded.

#2. They use breathing control application for five minutes.

#3. After that, their ECG signal is recorded for one minute.

#4. Finally, they answer the questionnaire about experiencing stress (Yes or No).

The results of the preliminary experiment are shown in Table 7.10. The SDNN of this experiment is calculated from the average of all standard deviations of RR interval in every ten seconds. The threshold of the SDNN value for detecting stress should be between 29ms-46ms. Therefore, I set SDNN thresholds: 35 ms based on the HRV analysis system for detecting and analyzing stress.

Table 7.10 Results of preliminary experiment about participants' stress

Participants	**While watching images for four minutes**		**After using a breathing control application for five minutes**	
	ECG (SDNN)	**Questionnaire**	**ECG (SDNN)**	**Questionnaire**
P1	28.91	Yes	63.98	No
P2	23.15	Yes	46.59	No

The workflow of stress detection using SDNN (Figure 7.8) is as follows.

- The stress detection performs low pass and notch filter to remove noise from ECG signal.
- QRS Detection is performed to detect the QRS Interval.
- After that, it detects peak amplitude and location of each RR-Interval from the cleaned ECG signal.
- RR Interval is collected within ten seconds

- The stress detection calculates SDNN from the average of all standard deviations of RRI in every ten seconds.
- If SDNN is less than the threshold, it indicates that the user is experiencing some stress. If SDNN is greater than or equal to the threshold, user has no stress.

In the next sections, I describe the experiment that evaluates the improvement of the prototype.

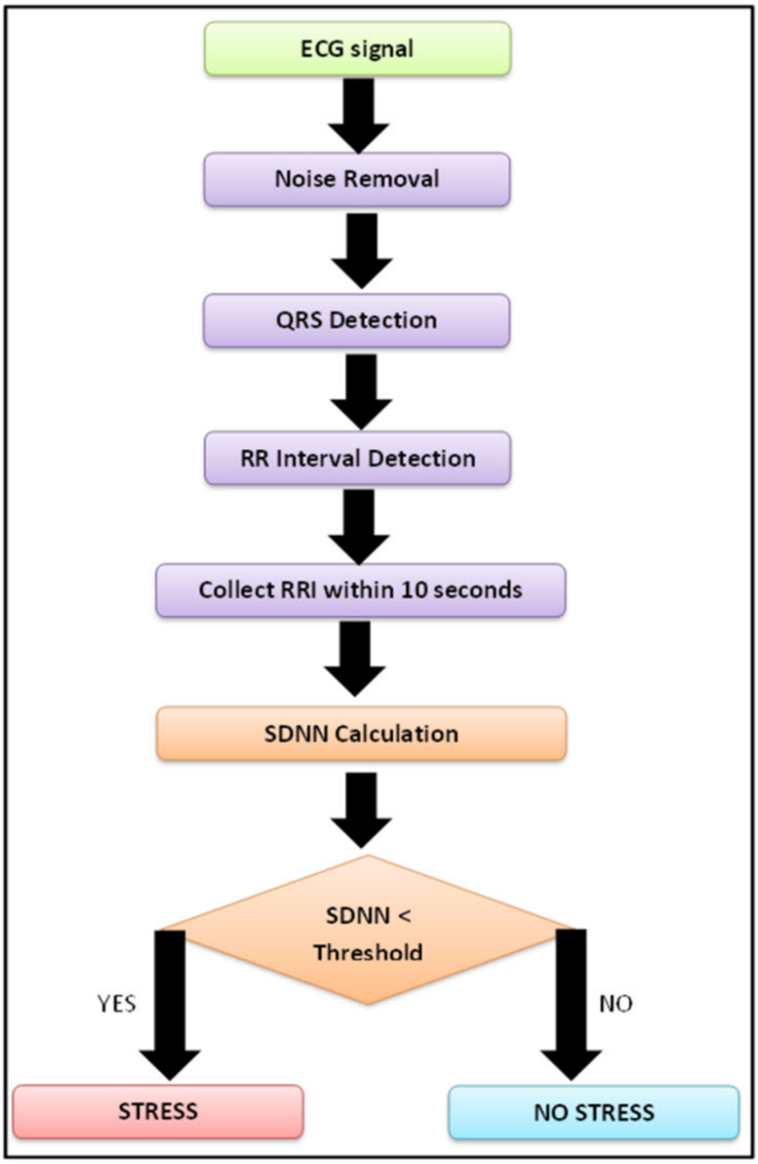


Figure 7.8 Stress detection using ECG signal

7.6 Experiment 3: the efficiency and effectiveness of improvement of the prototype with emotion recognition and stress detection

7.6.1 Objective

This experiment is conducted to measure the accuracy of the real-time stress detection from ECG signal with emotion recognition by facial expressions and to evaluate the prototype of emotional healthcare system. Additionally, this experiment is set to determine the most appropriate SDNN threshold for the prototype.

7.6.2 Experiment setup

The real-time emotion recognition by facial expressions and relaxation service are set up similar to Experiment 1. The real-time stress detection from ECG signal uses SDNN to recognize stress. The SDNN threshold is set to 30ms, 35ms and 40ms in order to find appropriate threshold for this prototype.

7.6.3 Participants

According to the emotion recognition by facial expression which applies the JAFFE dataset as training data, another six Japanese female subjects, ages between 24 to 46 years old, with no self-reported mental disorders, participated in this experiment. They are graduate students and staffs at Shibaura Institute of Technology.

7.6.4 Experimental procedure

I experimentally evaluate the efficiency and effectiveness of the prototype of emotional healthcare system and the accuracy of the real-time stress detection from ECG signal with the emotion recognition by facial expressions.

The participants are presented with the 30 negative images from International Affective Picture System (IAPS) [125] including snakes, spiders, crying babies, etc. Then, if the participants have negative emotions and stress, the breathing control application in relaxation service is presented to them. After that, the participants answer a forced-choice questionnaire (Figure 7.9) about their feeling for the image stimulus and breathing control application in relaxation service. These choices are the basic emotions, including happiness, sadness, disgust, surprise, anger, fear, and neutral. Finally, they answer the second questionnaire about their stress (Figure 7.9). During the experiment, the system actively monitors the participant facial expression and ECG signal to predict the emotional and stress response to the image stimulus and relaxation service. Multiple emotions and stress are actively detected by the system.

The following are the steps to perform the experiment:

#1. Participants watch 30 negative images for four minutes.

#2. While they watch the images, the emotion recognition recognizes their emotions from facial expressions and the stress detection recognizes stress from ECG signal.

#3. If they have negative emotions or stress, the relaxation service with breathing control application is provided to the participants for five minutes.

#4. After they experience the relaxation service, the emotion recognition recognizes their emotions from facial expressions and the stress detection recognizes stress from ECG signal for one minute.

#5. Finally, they answer questionnaire about their experienced emotions and stress (Figure 7.9).

1. In an experiment, after watching the images,
   - How did you feel?

   ☐ Happiness ☐ Sadness ☐ Disgust ☐ Surprise

   ☐ Anger ☐ Fear ☐ Neutral

   - Do you get any stress?

   ☐ Yes ☐ No

2. In an experiment, after using the breathing control application

   (Please skip this question, If you did not use it)

   - How did you feel?

   ☐ Happiness ☐ Sadness ☐ Disgust ☐ Surprise

   ☐ Anger ☐ Fear ☐ Neutral

   - Do you get any stress?

   ☐ Yes ☐ No

Figure 7.9 Questionnaire about feeling and stress (English version)

7.6.5 Results and Discussion

This experiment, I evaluate two aspects: the accuracy of emotion recognition by facial expressions with stress detection from ECG signal, and the efficiency of the improved prototype of emotional healthcare system

A. Accuracy of emotion recognition by facial expression with stress detection from ECG signal

To evaluate the accuracy of emotion recognition by facial expressions with stress detection from ECG signal, I compare the detected emotional results from the facial expression and stress results from ECG signal with questionnaire results. Table 7.11 shows results of the detected emotions from the emotion recognition by facial expression, SDNN values from stress detection using ECG signal, and the questionnaire results. Based on the emotion groups of the emotion recognition by facial expression and SDNN thresholds, Table 7.12 shows emotion groups and stress results.

Table 7.11 Results of participants' emotions and stress from emotion recognition, stress detection and questionnaires

Participants	While watching images for four minutes			After using a breathing control application for five minutes		
	Facial expression	ECG (SDNN)	Questionnaire	Facial expression	ECG (SDNN)	Questionnaire
P1	Dis 63.04%	29.41	Fear / Stress	Dis 100.00%	35.16	Happiness / No stress
P2	Fea 84.32% Dis 13.72%	19.72	Sadness / Stress	Dis 50.00% Fea 20.65%	58.36	Neutral / No stress
P3	Dis 86.28% Fea 11.76%	19.03	Sadness / Stress	Dis 96.60% Fea 1.14%	35.68	Neutral / No stress
P4	Dis 98.00% Fea 2.00%	21.93	Sadness, Disgust, Surprise, Fear, Neutral / Stress	Dis 94.00% Fea 2.00%	47.30	Disgust, Neutral/ No stress
P5	Fea 60.00% Dis 32.00% Ang 8.00%	25.09	Sadness / Stress	Fea 84.78% Dis 8.70% Ang 4.34%	39.25	Neutral / No stress
P6	Dis 100.00%	28.90	Surprise / Stress	Dis 97.22%	33.28	Neutral / No stress

*Dis = Disgust, Ang = Anger, Fea= Fear

**The summary of average results from emotion recognition by facial expression might not equal 100% because of face detection failure.

Table 7.12 Emotion groups and stress results when SDNN thresholds are 30ms, 35ms and 40ms

Period	Participants	Questionnaire	Facial expression	ECG (SDNN threshold = 30)	ECG (SDNN threshold = 35)	ECG (SDNN threshold = 40)
While watching images for four minutes	P1	Negative / Stress	Neutral	Stress	Stress	Stress
	P2	Negative / Stress	Negative	Stress	Stress	Stress
	P3	Negative / Stress	Neutral	Stress	Stress	Stress
	P4	Negative / Stress	Neutral	Stress	Stress	Stress
	P5	Negative / Stress	Negative	Stress	Stress	Stress
	P6	Positive / Stress	Neutral	Stress	Stress	Stress
After using breathing control application for five minutes	P1	Positive / No stress	Neutral	No stress	No stress	Stress
	P2	Neutral / No stress	Neutral	No stress	No stress	No stress
	P3	Neutral / No stress	Neutral	No stress	No stress	Stress
	P4	Neutral/ No stress	Neutral	No stress	No stress	No stress
	P5	Neutral / No stress	Negative	No stress	No stress	Stress
	P6	Neutral / No stress	Neutral	No stress	Stress	Stress
Accuracy (%)			75.00	100.00	91.67	66.67

From the results in Table 7.12 and the error rule (Figure 7.5), the error score of emotion recognition by facial expression is 3, but if errors occur in all cases, it is 12. Thus, the error percentage is ((3*100)/12) = 25, and the accuracy is 100% – 25% (error) = 75%. From the comparison results between stress detection with SDNN threshold = 35 and questionnaire results, 11 out of 12 are correct. Therefore, its accuracy is = 91.67%.

To activate relaxation service when users experience negative emotions or stress, I apply the OR rule of probability (Negative emotions or stress will be 1 and the others will be 0) to the combined result between the emotion recognition by facial expression and stress detection from ECG signal as shown in Table 7.13 and the combination results are shown in Table 7.14

Table 7.13 OR rule of probability for activation of relaxation service.

Emotion Results (questionnaire and emotion recognition by facial expression)		**Stress Results (questionnaire and stress detection using ECG signal)**		**Activation of relaxation service**
Negative	**1**	Stress	1	√
		No stress	0	√
Positive / Neutral	**0**	Stress	1	√
		No stress	0	-

The results in Table 7.14 show that the emotion recognition by facial expression can activate the relaxation service with 66.67% of accuracy. The combined results from the emotion recognition by facial expression and stress detection from ECG signal produce error of only 16.67% which is reduced from 33.33% of error from only facial emotion recognition. Therefore, stress detection from ECG signal can address the confusion issue of the facial emotion recognition. Furthermore, the combination of both is very effective at recognizing negative emotions and stress with 83.33% of accuracy. Additionally, this real-time emotion recognition by facial expression can recognize emotions around 1 second/frame. However, stress detection from ECG signal can recognize emotions around 10 seconds. Thus, real-time emotion recognition and stress detection are feasible to recognize emotions and stress for 10 seconds.

Table 7.14 Combined results to activate relaxation service when SDNN threshold = 35ms

Period	Participants	Facial expression	ECG (SDNN threshold = 35)	Activation of relaxation service		
				Questionnaire	Facial Emotion Recognition	Combined results
While watching images for four minutes	P1	Neutral	Stress	Activated	Not activated	Activated
	P2	Negative	Stress	Activated	Activated	Activated
	P3	Neutral	Stress	Activated	Not activated	Activated
	P4	Neutral	Stress	Activated	Not activated	Activated
	P5	Negative	Stress	Activated	Activated	Activated
	P6	Neutral	Stress	Activated	Not activated	Activated
After using breathing control application for five minutes	P1	Neutral	No stress	Not activated	Not activated	Not activated
	P2	Neutral	No stress	Not activated	Not activated	Not activated
	P3	Neutral	No stress	Not activated	Not activated	Not activated
	P4	Neutral	No stress	Not activated	Not activated	Not activated
	P5	Negative	No stress	Not activated	Activated	Activated
	P6	Neutral	Stress	Not activated	Not activated	Activated
Accuracy (%)					66.67	83.33

B. Efficiency of improved prototype of emotional healthcare system

From the experimental results (Table 7.11 – 7.14), integration of stress detection can improve the efficiency of the prototype for recognizing stress and providing relaxation service. Furthermore, the results confirm that relaxation service is effective at decreasing stress and negative emotions. All participants watched negative images, they experienced some stress with negative emotions but after using the breathing control application, they had neutral or positive emotions without stress.

Table 7.15 Accuracies of stress detection when SDNN thresholds are 30, 35 and 40 ms

SDNN Threshold (ms)	Accuracy (%)
30	100.00
35	91.67
40	66.67

Table 7.16 Combined results to activate relaxation service when SDNN threshold = 30ms

Period	Participants	Facial expression	ECG (SDNN threshold = 30)	Activation of relaxation service	
				Combined results	Questionnaire
While watching images for four minutes	P1	Neutral	Stress	Activated	Activated
	P2	Negative	Stress	Activated	Activated
	P3	Neutral	Stress	Activated	Activated
	P4	Neutral	Stress	Activated	Activated
	P5	Negative	Stress	Activated	Activated
	P6	Neutral	Stress	Activated	Activated
After using breathing control application for five minutes	P1	Neutral	No stress	Not activated	Not activated
	P2	Neutral	No stress	Not activated	Not activated
	P3	Neutral	No stress	Not activated	Not activated
	P4	Neutral	No stress	Not activated	Not activated
	P5	Negative	No stress	Activated	Not activated
	P6	Neutral	No stress	Not activated	Not activated

C. The most appropriate SDNN threshold for the prototype

Since, the SDNN threshold for detecting stress should be between 29ms-46ms. To determine the most appropriate SDNN threshold, I analyzed the experiment results again by setting another two SDNN thresholds: 30ms and 40ms. From Table 7.12, the errors of stress detection are calculated when SDNN thresholds are 30 ms and 40 ms respectively. The errors are 0 (0%), and 4 (33.33%) as shown in Table 7.15. Therefore, the appropriate SDNN threshold for the prototype is 30ms because stress detection achieves 100% of accuracy. Moreover, the accuracy of emotion recognition by facial expression with stress detection from ECG signal can reach 91.67% when SDNN threshold of stress detection is 30 ms as shown in Table 7.16.

7.7 Summary

The prototype of healthcare system focusing on emotional aspect is built to recognize users' negative emotions and stress, and provide relaxation service in real-time process. I implement real-time emotion recognition by facial expression and ECG signal, real-time stress detection from ECG signal in order to provide real-time relaxation service. From the experimental results, the relaxation service is effective to decrease negative emotions. The emotion recognition by facial expression and ECG signal is effective to recognize negative emotions but it cannot recognize neutral emotion without confusion. Therefore, I integrate the stress detection using ECG signal to improve the prototype and the results confirm that combination between the emotion recognition by facial expression and stress detection from ECG signal is very effective to recognize negative emotions and stress. Moreover, stress detection can address the confusion issue of the facial emotion recognition. The integration of stress detection can improve the efficiency of the prototype to recognize stress and provide relaxation service. In summary, this real-time prototype of healthcare system focusing on emotional aspect is effective enough at recognizing negative emotions and stress to provide relaxation service to decrease negative emotions or stress, and increase positive emotions in real-time process.

Chapter 8

Discussion

8.1 Summary of previous chapters

Our modern societies require healthcare systems that focus on emotional aspect to provide assistance and services to encourage positive emotions and decrease negative emotions. To construct such systems, some issues must be addressed. Therefore, I conduct my research questions and goals that were solved and achieved by this dissertation by researching, designing, implementing, and evaluating a healthcare system focusing on emotional aspect, relaxation service, and emotion recognition by facial expression and ECG signal. I summarize my work in each previous chapter as follows:

- I described the motivation and problem statements to set my research questions and goals in Chapter 1.
- In Chapter 2, I reviewed several researches related to my dissertation.
- Next, I designed the overall system and a framework for its implementation in Chapter 3.
- In Chapter 4, I implemented a relaxation service with a breathing control application using augmented reality and breathing detection. I also evaluated the relaxation service's efficiency to decrease stress.
- Next I designed, implemented, and evaluated emotion recognition by facial expression using my proposed feature extraction method (CDTP) in Chapter 5.
- In Chapter 6, I designed, implemented, and evaluated emotion recognition from ECG signal using local pattern description methods.
- Finally, in Chapter 7, I constructed a prototype of healthcare system focusing on emotional aspect by integrating the relaxation service (Chapter 4), the emotion recognition from facial expression and ECG signal (Chapters 5 and 6) and a stress detection from ECG signal (Chapter 7.5). I evaluated the effectiveness of the prototype to recognize users' negative emotions and stress and to provide the relaxation service in a real-time process.

This chapter discusses my research which solved all of my research questions and achieved my research goals.

8.2 Healthcare system focusing on emotional aspect

My first research question is: How should the healthcare system focusing on emotional aspect be designed and which devices, applications and services are necessary for maintaining user emotional health? To answer this question, I set a research goal to propose a healthcare system focusing on emotional aspect and designed it to be more attractive, effective, and intelligent to support users. I created a new design of emotional healthcare system that is more attractive, effective, and intelligent to support ordinary people at their workplaces by assisting them to release their negative emotions or stress that are often caused by compulsive studying or excessive overwork. This system's attractiveness was improved by providing three emotional services (relaxation, amusement, and excitement) with augmented reality or Kinect because the interaction between users and virtual objects in real environments can raise the interest of users for the system. I also improved the system's effectiveness and intelligence by integrating emotion recognition and stress detection from facial expression, speech, and biological signals so that it can provide appropriate services to maintain users' emotional health based on their current emotional states, which are analyzed by emotion recognition and stress detection. This system was designed to integrate biological sensors for many purposes. When these biological sensors are designed as wearable accessories, adapting them for daily life will be more comfortable for users. Additionally, I designed this system to support future implementation and enhancement. For example, breathing detection currently detects user respiration using ECG sensors; if the system integrates a new respiration sensor to measure respiration, just breathing detection needs to be modified to feed the data from respiration sensors instead of ECG sensors. The other services and applications won't need to be modified.

In summary, I confirmed that my new emotional healthcare system's design is more attractive, effective, and intelligent to support users with augmented reality, emotion recognition, and stress detection.

8.3 Relaxation service

The second question is comprised of the following parts: How can this system relax users, which techniques should this system apply, and how should the application be designed to be more attractive and effective to support users? I set a research goal to propose a relaxation service with augmented reality. The service must be designed and implemented to provide effective relaxation and increase the treatment content's attractiveness by combining it with augmented reality. Therefore, I designed and implemented a breathing control application by applying the deep breathing technique of stress management with virtual music boxes using augmented reality because displaying them and allowing users to experience them in real environment would increase the attractiveness of the application. Since, just deep breathing is an effective technique to reduce stress. Thus, in this dissertation, I ensured that the breathing control application with deep breathing and augmented reality more effectively decreased stress than only a deep breathing technique because when users controlled such breathing by following the augmented reality (virtual music boxes and music), the virtual music boxes encouraged them to concentrate more on deep breathing and music can entertain and increase relaxation and comfort.

In summary, I confirmed that my new relaxation service with a breathing control application using AR reduced stress and increased relaxation.

8.4 Emotion recognition

My third research question is: How can this system detect user emotions with high accuracy and performance? To answer this question, I set a research goal in which emotion recognition by facial expression and ECG signal must be applied to real-time recognition of user emotional states from both user appearances (outside) and biological signals (inside) to achieve high accuracy and performance for providing appropriate services.

Feature extraction and classifications are the important steps to recognize emotions. I improved the accuracy and the performance of emotion recognition by facial expression by proposing CDTP to extract facial features and classifying them using SVM. My benchmark evaluation results confirmed that CDTP improved the accuracy and performance over DTP because it effectively decreased the size and the redundancy features of the DTP feature vector. However, emotion recognition by facial expression using CDTP caused some confusion when recognizing similar facial expressions was caused by different emotions.

To address the confusion issue, I integrated emotion recognition using ECG signal because emotions directly affect to ECG signal and ECG signal are controlled by the autonomic nervous system. I applied LBP, LTP, and CLTP to extract the features from the entire signal. I found that ECG signal has emotional local patterns that resemble facial expression in which different emotions produced different patterns on the histogram features. I confirmed that these approaches extracted ECG emotional features with high accuracy. I think such local pattern description approaches might be able to extract emotional features from other biological signals, such as EEG and GSR, because they might have emotional local patterns.

To classify emotions with high accuracy and performance, I selected linear SVM to classify them from facial expression because it was suitable for classifying high-dimensional and sparse features like facial image features. The CDTP features were also linearly separable, since linear SVM can classify emotions with high accuracy. However, I selected k-NN instead of linear SVM for emotion recognition from the ECG signal because it classifies emotions with higher accuracy and faster than SVM. ECG features from LBP and LTP were also low-dimensional and different from the image features. The features were quite unique in which different emotions produced different patterns and similar emotions produced similar patterns. Therefore, ECG features were more suitable for k-NN, which was designed for low-dimensional and similar-based features. The classification methods for emotion recognition from ECG signal were different from emotion recognition by facial expression, since the image and ECG features were different.

In summary, I proposed a new feature extraction method (CDTP) for emotion recognition by facial expression and new feature extraction methods (LBP and LTP) for emotion recognition from ECG signal that recognized emotions with high accuracy and performance.

8.5 The prototype of healthcare system focusing on emotional aspect

Finally, to achieve my last research goal, I implemented and evaluated my prototype of healthcare system focusing on emotional aspect to recognize negative emotions and stress from facial expression and ECG signal to provide the relaxation service. I found the following:

- The relaxation service effectively decreased negative emotions and stress.
- The emotion recognition by facial expression using CDTP caused some confusion.
- Integrating ECG signal to recognize emotions with facial expression increased the accuracy of emotion recognition by facial expression.
- Emotion recognition by facial expression and ECC signal accurately analyzed

negative emotions with adequately high accuracy.

The relaxation service and emotion recognition by facial expression and ECG signal were adequately effective for integrating in a real-time emotional healthcare system.

To increase the system's effectiveness to detect stress, I improved its prototype by integrating stress detection using the SDNN of HRV from ECG signal. As a result, the stress detection recognized stress with higher accuracy. It also increased the accuracy of facial emotion recognition and my prototype's effectiveness and efficiency to provide the relaxation service when users experienced negative emotions or stress.

The real-time prototype was practicable to recognize emotions from facial expression and stress from ECG signal at around 10 seconds. Although, emotion recognition by facial expression recognized emotions at around 1 second, the stress and emotional details were observed from the ECG signal from 3 to 15s. Therefore, I chose to recognize emotions and stress from ECG signal at around 10 seconds for high accuracy and performance. Finally, the results from emotion recognition by facial expression, emotion recognition from ECG signal and stress detection from ECG signal were analyzed using OR rules to activate the relaxation service. Even though the real-time prototype took time to analyze the emotional state of users to provide assistance services, this system was designed for workplaces during working hours. Since this system was not designed for emergency cases, its analysis time is acceptable.

In summary, I confirmed that my prototype of a new emotional healthcare system with real-time relaxation service, emotion recognition, and stress detection effectively and efficiently recognized the negative emotions and stress of users to provide the relaxation service.

8.6 Potential of this research

The novelty of this dissertation is a new emotional healthcare system that integrates emotion recognition and stress detection to recognize user emotional states with high accuracy and performance to provide appropriate emotional services with augmented reality to cope with negative feelings in a real-time process. From this dissertation, I obtained the following new findings:

- New design of an emotional healthcare system.
- New relaxation service with augmented reality.
- New feature extraction approach (CDTP) for recognizing emotions from facial expression
- New feature extraction for emotion recognition from ECG signal by applying feature extraction methods for emotion recognition by facial expression.
- New real-time emotional healthcare system using real-time emotion recognition by facial expression and ECG signal, and real-time stress detection from ECG signal

At this state, since the real-time emotional healthcare system (the current achievement version of my system) can only provide the relaxation service, it does not need to distinguish from among such specific user emotions as sadness, anger, or fear. Its results improved when it separated user emotions into only three groups: positive, neutral and negative. However, I designed and implemented real-time emotion recognition to recognize seven basic emotions (instead of only three emotional groups) to fit the requirements of an ideal system. This will benefit its future enhancement and growth potential, to provide such remaining services as amusement and excitement or new services based on different negative emotions. If emotion recognition can identify such emotions as anxiety, boredom or contempt using speech, EEG, GSR, or gestures with facial expression and ECG signal, the emotion recognition efficiency will improve because different emotions affect the human body in different ways. Then the system's potential will increase because it will be able to recognize and analyze more emotions than just seven classes. For example, if users are sad or bored, the system will provide amusement service.

Below are future possibilities to enhance the system, improve its intelligence, and overcome its limitations:

- Providing more relaxation applications for the relaxation service such as visualizing relaxing scenes.
- Providing more services based on user anticipated emotions, such as amusement and excitement services.
- Applying new training datasets for emotion recognition by ECG signal to recognize more emotions, such as neutral, disgust, surprise, and fear.

- Applying various techniques such as speech, EEG, GSR, and gestures to increase the accuracy and performance of recognizing emotions and stress.
- Applying learning processes from users to improve the accuracy of real-time emotion recognition.
- Improving the system to allow access from smartphones and tablets.

Chapter 9

Conclusion and Future work

In this chapter, I conclude my doctoral dissertation and explain future work to expand my research and recommendations for subsequent steps of this research field.

9.1 Conclusion of research work

I proposed a new healthcare system focusing on emotional aspect for ordinary people such as students or working people to help them cope with their daily negative emotions or stress. Below is a summary of each research goal.

To achieve my first research goal, I designed a more attractive, effective and intelligent emotional healthcare system to support users in daily life. My system's attractiveness was increased by providing three emotional services (relaxation, amusement and excitement) with augmented reality or Kinect because these technologies allowed users to experience and interact with virtual objects in a real environment to get positive emotions and decrease their negative emotions. My system became more effective and intelligent by equipping it with various I/O devices to support three services. I integrated emotion recognition and stress detection from facial expression, speech and biological signals to recognize user emotional states to provide appropriate services and maintain emotional health.

To achieve my second research goal, I proposed a relaxation service to help users relax by decreasing their stress and negative emotions. I designed and implemented a breathing control application for my relaxation service. This application applied a deep breathing technique of stress management with virtual music boxes using augmented reality to assist users to perform deep breathing. My experimental results confirmed that the breathing control application decreased stress. Moreover, the augmented reality helped decrease stress more quickly than just deep breathing.

To achieve my third research goal, I applied emotion recognition by facial expression and ECG signal. For emotion recognition by facial expression, I proposed CDTP by reducing the size, the sparseness, and the redundancy of the DTP feature vectors. Emotion recognition by facial expression using CDTP with Kirsch edge detection and SVM classifier recognized emotions with 79.16% average accuracy. However, some confusion was caused when recognizing similar facial expressions by different emotions. To address this issue, I integrated emotion recognition from ECG signal and applied such local pattern description methods as LBP, LTP and CLTP from facial feature extraction to extract ECG feature. The local pattern description methods recognized emotions with high accuracy because the ECG signal had emotional local patterns in which different emotions produced different patterns on its histogram features. Furthermore, the emotion recognition from ECG signal using LTP with a 15s frame-length, a 7.5s frame-shift, and k-NN when k=5 produced higher accuracy of around 87.92%. Therefore, I expect the emotion recognition from ECG signal to recognize emotions with facial expression to produce higher accuracy and performance in a real-time process.

To achieve my last research goal, I built a prototype of a healthcare system focusing on emotional aspect to recognize negative emotions and stress of users, and provided the relaxation service in a real-time process. I implemented real-time emotion recognition by facial expression and ECG signal, real-time stress detection from ECG signal and real-time relaxation service. I evaluated the effectiveness of my prototype and my results confirmed that its relaxation service effectively decreased negative emotions. Emotion recognition by facial expression and ECG signal effectively recognized negative emotions and symptom of negative feelings with up to 75% accuracy. The combination of emotion recognition by facial expression and stress detection from ECG signal more effectively recognized negative emotions and stress with up to 83.33% accuracy. This ensured that the integration of stress detection improved the prototype’s efficiency to recognize stress and provide the relaxation service. In summary, this real-time prototype of a healthcare system focusing on emotional aspect effectively recognized negative emotions and stress to provide the relaxation service to decrease negative emotions or stress and increase positive emotions in a real-time process.

Finally, I confirmed that I achieved all of my research goals to construct attractive, effective and intelligent healthcare system focusing on emotional aspect using augmented reality, emotion recognition by facial expression and ECG signal, and stress detection from ECG signal. Additionally, the new findings from my dissertation are also applicable to other research and systems that I will describe in future work.

9.2 Future work

My emotional healthcare system has the following limitations. It can only support users with its relaxation service, and the emotion recognition by ECG can only recognize sadness, anger and happiness. It recognizes emotions and stress with some confusion. However, the new findings from my dissertation can be applied in other research fields.

- The concept and new design of an emotional healthcare system are also adaptable for a home healthcare or a smart home for the elderly to maintain emotional states and improve quality of life.
- The new relaxation service with augmented reality might be a useful tool for patients who have symptoms of stress disorder and their caregivers.
- The new algorithm for emotion recognition by facial expression (CDTP) and new algorithms for emotion recognition by ECG signal (LBP and LTP) are applicable to other research fields.

References

[1] S. Stewart-Brown, and R. Layte, “*Emotional health problems are the most important cause of disability in adults of working age: a study in the four counties of the old Oxford region,*” Journal of epidemiology and community health, vol. 51, no. 6, pp. 672-675, 1997.

[2] K. Vaughn, “*Physical Health vs. Mental Health: Is One More Important?*” [Online]. Available at http://www.bjceap.com/Blog/ArtMID/448/ArticleID/93/Physical-Health-vs-Mental-Health-Is-One-More-Important [Accessed 18 June 2015].

[3] J. Kim, H. S. Choi, H. Wang, N. Agoulmine, M. J. Deerv, and J. K. Hong, *“POSTECH's U-Health Smart Home for elderly monitoring and support,”* IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM), pp. 1-6, June, 2010.

[4] K. Morioka, J. H. Lee, and H. Hideki, *"Human centered robotics in intelligent space,"* In Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on, vol. 2, pp. 2010-2015., 2002.

[5] T. H. Laine, “*U-Healthcare System*” [Online]. Available at http://people.ubilife.net/teemu/?page_id=10 [Accessed 7 May 2015].

[6] S. Haq, and Ph. J.B. Jackson, “*Multimodal Emotion Recognition*”, Machine Audition: Principles, Algorithms and Systems, 398-423, 2011

[7] World Health Organization, “*Suicide Prevention*” [Online]. Available at http://www.who.int/mental_health/prevention/en/ [Accessed 7 May 2015].

[8] Befrienders International, “*Suicide Statistics*” [Online]. Available at http://www.befrienders.org/info/index.asp?PageURL=statistics.php [Accessed 7 May 2015].

[9] Jeffgissing, *“Archives For Leadership”* [Online]. Available at http://jeffgissing.com/tag/leadership/ [Accessed 7 May 2015].

[10] Association for Psychological Science, *“Negative emotions in response to daily stress take a toll on long-term mental health,”* ScienceDaily [Online]. Available at www.sciencedaily.com/releases/2013/04/130402150159.htm [Accessed 7 May 2015].

[11] G. Pioggia, N. Carbonaro, G. Anania, A. Tognetti, G. Tartarisco, M. Ferro, D. De Rossi, A. Gaggioli, and G. Riva, *“Interreality: The use of advanced technologies in the assessment and treatment of psychological stress,”* 10th International Conference on Intelligent Systems Design and Applications (ISDA), pp.1047-1051, November, 2010.

[12] W.P. Brinkman, D. Hartanto, N. Kang, D. de Vliegher, I. L. Kampmann, N. Morina, and P. G.M. Emmelkamp, *“A virtual reality dialogue system for the treatment of social phobia”*, In Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts, 1099-1102 ,2012.

[13] H. M. Jani, *“Benefiting from online mental status examination system and mental health diagnostic system,”* 3rd International Conference on Information Sciences and Interaction Sciences (ICIS), 66-70, 2010.

[14] J. Kim, Ch. H. Youn, D. Kim, J. M. Woo, and S. Jung, *“Cost-Minimized E-Health Service for Identification of Mental Stress related Heart Rate Changes,”* 4th IEEE/EMBS International Summer School and Symposium on Medical Devices and Biosensors, 102-106, 2007

[15] *“Emoticon images”* [Online]. Available at http://pulpypics.com/emoticons-for-facebook-and-skype/ [Accessed 18 June 2015].

[16] *“Relaxation image”* [Online]. Available at http://www.lvwomansclub.org/wp-content/uploads/2013/03/a3.jpg [Accessed 18 June 2015].

[17] *“Massage image”* [Online]. Available at http://centerforspinaldisorders.com/wp-content/uploads/2011/05/Picture-61-300x261.png[Accessed 18 June 2015].

[18] M. Allen, R. Holger, and M. Abbott, *"Smart-phone augmented reality for public participation in urban planning,"* Proceedings of the 23rd Australian Computer-Human Interaction Conference. ACM, 2011.

[19] E. Socolofsky, *“FLARManager Augmented Reality in Flash”* [Online]. Available at http://words.transmote.com/wp/flarmanager/ [Accessed 7 May 2015].

[20] M. S. Obaidat, M. Denko, and I. Woungang, eds. *Pervasive computing and networking*. John Wiley & Sons, 2011.

[21] M. Smith, R. Segal, and J. Segal, “*Improving Emotional Health”* [Online]. Available at http://www.helpguide.org/articles/emotional-health/improving-emotional-health.htm [Accessed 7 May 2015].

[22] S. M. Lajevardi, and Z. M. Hussain, *“Automatic facial expression recognition: feature extraction and selection,”* Signal, Image and Video Processing, vol.6, no.1, pp. 159-169, 2012..

[23] T. F. Cootes, G. J. Edwards, and C. J. Taylor, *“Active appearance models,”* IEEE Trans. Pattern Anal. Mach. Intell, vol. 23, no. 6, pp. 681-685, 2001

[24] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, *“The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression,”* in Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 94-101, San Francisco, California, June 13-18, 2010

[25] C. Shan, Sh. Gong, and P.W. McOwan, *“Facial expression recognition based on Local Binary Patterns: A comprehensive study,”* Image Vision Comput, vol. 27, no.6, pp. 803-816, 2009

[26] T. Ahonen, A. Hadid, and M. Pietikainen, *“Face description with local binary patterns: Application to face recognition,”* IEEE Trans. Pattern Anal. Mach. Intell, vol. 28, no.12, pp. 2037-2041.

[27] M.H. Kabir, and F. Ahmed, *“Face recognition with directional ternary pattern (DTP),”* In Proceedings of the 2012 International Conference on Graphic and Image Processing, Singapore, March 14, 2013, 87687E. doi: 10.1117/12.2009559

[28] D. Ververidis, and K. C. Kotropoulos, *“Emotional speech recognition: Resources, features, and methods”*, SPEECH COMMUNICATION vol. 48 No. 9, pp. 1162-1181, 2006.

[29] M. Z. U. Rahman, R. A. Shaik, and D. R. K. Reddy, *“Noise cancellation in ECG signals using computationally simplified adaptive filtering techniques: Application*

to biotelemetry," Signal Processing: An International Journal (SPIJ), vol. 3 , no.5, pp. 1-12, 2009.

[30] M. Yuanchao, X. Bin, B. Yin, S. Guodong, and Z. Run, *"Daily Mood Assessment Based on Mobile Phone Sensing,"* Ninth International Conference on Wearable and Implantable Body Sensor Networks (BSN), pp.142-147, 9-12 May 2012

[31] R. Nakajima, J. Nomura, S. Fukushima, and T. Ojika, *"A relax/refresh system applying virtual reality,"* 3rd IEEE International Workshop on Robot and Human Communication, RO-MAN '94, Nagoya, pp.145-150, 18-20 Jul 1994

[32] Q. Zhang, A. H. Zhu, S. J.Wang, J. L Xin, and W. C. Zhang, *"E-Psychotherapy: Towards an Internet-based mental health service,"* 2011 International Symposium on IT in Medicine and Education (ITME), vol. 2, pp. 191-195, December 2011.

[33] Y. H. Lee, V. Shieh, C. L. Lin, and Y. J. Shiah, *"A Stress Evaluation and Personal Relaxation System Based on Measurement of Photoplethysmography,"* Second International Conference on Robot, Vision and Signal Processing (RVSP), pp. 182-185, December 2013

[34] M. Billinghurst, and H. Kato, *"Collaborative Augmented Reality"* [Online]. Available at http://masters.donntu.org/2012/iii/akchurin/library/article9.htm [Accessed 7 May 2015].

[35] D. W. F. Van Krevelen, and R. Poelman. *"A survey of augmented reality technologies, applications and limitations,"* International Journal of Virtual Reality vo. 9, no. 2 , 2010.

[36] B. Furht, ed. *"Handbook of augmented reality,"* Springer Science & Business Media, 2011.

[37] D. Bitouk, R. Verma, and A. Nenkova, *"Class-level spectral features for emotion recognition,"* Speech Communication, vol. 52, no.7, pp. 613-625, 2010

[38] L. Chen, X. Mao, Y. Xue, and L.L. Cheng, *"Speech emotion recognition: Features and classification models,"* Digital Signal Processing, vol. 22, no.6, pp. 1154-1160 , 2012

[39] X. Cheng, and Q. Duan, *"Speech Emotion Recognition Using Gaussian Mixture Model,"* In International Conference on Computer Application and System Modeling, pp. 1222-1225, 2012

[40] X. Feng, B. Lv, Z. Li, and J. Zhang, *"A novel feature extraction method for facial expression recognition,"* In Proceedings of Joint Conference on Information Sciences, 2006.

[41] A. Hussain, S. M. Khan, M. Nazir, and M. A. Iqbal, *"Survey of various feature extraction and classification techniques for facial expression recognition,"* In Proceedings of the 11th WSEAS International Conference on Electronics, Hardware, Wireless and Optical Communications, and Proceedings of the 11th WSEAS International Conference on Signal Processing, Robotics and Automation, and Proceedings of the 4th WSEAS International Conference on Nanotechnology; World Scientific and Engineering Academy and Society (WSEAS), pp. 138-142, 2012

[42] X. Wu, and J. Zhao, *"Curvelet feature extraction for face recognition and facial expression recognition,"* in Proceedings of the 2010 IEEE Sixth International Conference on Natural Computation (ICNC), vol. 3, pp. 1212-1216, Yantai, China, August 10-12, 2010.

[43] S. Emerich, E. Lupu, and A. Apatean, *"Emotions recognition by speech and facial expressions analysis,"* In 17th European Signal Processing Conference (EUSIPCO 2009), pp. 1617-1621, Glasgow, UK, August 2009,

[44] N. Sebe, I. Cohen, T. Gevers, and T.S. Huang, *"Multimodal approaches for emotion recognition: a survey,"* In Electronic Imaging 2005, International Society for Optics and Photonics, pp. 56-67, January 2005.

[45] X. Ya, L. Guangyuan, H. Min, W. Wanhui, and H. Xiting, *"Analysis of affective ECG signals toward emotion recognition",* Journal of Electronics (China), vol. 27, no. 1, pp. 8-14, 2010

[46] M. Murugappan, N. Ramachandran, and Y. Sazali, *"Classification of human emotion from EEG using discrete wavelet transform,"* J. Biomedical Science and Engineering,vol. 3, no. 4, pp. 390-396, 2010

[47] K. H. Kim, S. W. Bang, and S. R. Kim, *"Emotion recognition system using short-term monitoring of physiological signals,"* Medical and biological engineering and computing , vol. 42, no.3, pp. 419-427, 2004

[48] K. Rattanyu, and M. Mizukawa, *"Emotion Recognition Based on ECG Signals for Service Robots in the Intelligent Space During Daily Life,"* Journal Of Advanced

Computational Intelligence And Intelligent Informatics, vol. 15, no. 5, pp. 582-591, 2011

[49] M. Ohkura, Y. Aoki, and T. Aoto, *"Evaluation of Comfortable Spaces for Women using Virtual Environment -Objective Evaluation by Biological Signals-,"* Kansei Engineering International, vol. 8, no. 1, pp.67-72, 2009

[50] M. Ohkura, M. Hamano, H. Watanabe, and T. Aoto, *"Chap. 18 Measurement of Wakuwaku Feeling of Interactive Systems,"* in Emotional Engineering –Service Development (Ed: Shuichi Fukuda), pp.327-343, Springer, 2010

[51] Y.L. Tian, T. Kanade, and J.F. Cohen, *"Handbook of Face Recognition,"* Springer-Verlag, pp. 487-519, London, 2011

[52] Z. Zhang, M. Lyons, M. Schuster, and S. Akamatsu, *"Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron,"* in Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 454-459, Nara, Japan, April 14-16, 1998,

[53] B. Van Ginneken, A.F. Frangi, J.J. Staal, B.M. ter Haar Romeny, and M.A. Viergever, *"Active shape model segmentation with optimal features,"* IEEE Trans. Med. Imag. vol. 21, no.8, pp. 924-933, 2002, doi: 10.1109/TMI.2002.803121

[54] C. J. Burges, *"A tutorial on support vector machines for pattern recognition,"* Data mining and knowledge discovery, vol. 2, no. 2, pp. 121-167, 1998

[55] B. Fasel, and J. Luettin, *"Automatic facial expression analysis: a survey,"* Pattern Recogn, vol. 36, no.1, pp.259-275, 2003, doi: 10.1016/S0031-3203(02)00052-3

[56] T. Jabid, M. Kabir, and O. Chae, *"Robust facial expression recognition based on local directional pattern,"* ETRI Journal, vol. 32, no. 5, pp. 784-794, 2010, doi: 10.4218/etrij.10.1510.0132

[57] S. Liao, W. Fan, C.S. Chung, and D.Y. Yeung, *"Facial expression recognition using advanced local binary patterns, Tsallis entropies and global appearance features,"* in Proceedings of the 2006 IEEE International Conference on Image Processing (ICIP), pp. 665–668, Atlanta, GA, October 8-11, 2006

[58] F. Agrafioti, H. Dimitrios, and A. K. Adam, *"ECG pattern analysis for emotion detection,"* IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 102-115, 2012.

[59] M. Ohkura, S. Goto, A. Higo, and T. Aoto, *"Relation between Kawaii Feeling and Biological Signals,"* Transactions of Japan Society of Kansei Engineering, vol. 10, no. 2, pp. 109-114, 2011.

[60] Y. Harada, T. Furuya, N. Takahashi, K. Hasegawa, T. Nakazato, and M. Ohkura, *"Content evaluation of exciting feeling by using biosignals,"* Proceeding of AHFE2014, pp. 6931-6936, 2014.

[61] M Ohkura, N. Takahashi, H. Seki, Y. Saito, and K. Hasegawa, *"Study on driver distraction indexes using biological signals applying assessment of information display,"* Proceeding of IDETC-CIE, 2013

[62] M. Ohkura, M. Hamano, H. Watanabe, and T. Aoto, *"Measurement of "wakuwaku" feeling generated by interactive systems using biological signals,"* The Proceedings of the Kansei Engineering and Emotion Research International Conference, pp.2301-2293, Paris, March 2010

[63] T. Xiaoyang, and B. Triggs, *"Enhanced local texture feature sets for face recognition under difficult lighting conditions,"* IEEE Transactions on Image Processing, vol. 19, no. 6, pp. 1635-1650, 2010.

[64] A. S. Utane, and S. L Nalbalwar, *"Emotion Recognition through Speech Using Gaussian Mixture Model and Support Vector Machine,"* emotion, vol. 2, no. 8, 2013

[65] S. Wu, T. H. Falk, and W. Y. Chan, *"Automatic speech emotion recognition using modulation spectral features,"* Speech Communication, vol. 53, no. 5, pp. 768-785, 2011

[66] P. C. Panagiotis ,and L. J. Hadjileontiadis, *"Emotion recognition from EEG using higher order crossings,,* IEEE Transactions on Information Technology in Biomedicine, vol. 14, no.2, pp. 186-197, 2010

[67] A. Nakasone, H. Prendinger, and M. Ishizuka, *"Emotion recognition from electromyography and skin conductance,"* Proc. of the 5th International Workshop on Biosignal Interpretation., 2005.

[68] L. Erlbaum, *“Associates Facets of Emotion: Recent Research,”* Hillsdale, p. 115, New Jersey, 1988.

[69] Lisetti, C., Nasoz, F., LeRouge, C., Ozyer, O., and Alvarez, K., *“Developing multimodal intelligent affective interfaces for tele-home health care,”* International Journal of Human-Computer Studies, vol. 59, no. 1, pp. 245-255, 2003.

[70] Jiang, Z., Lu, L., Huang, X., and Tan, C., *“Design of wearable home health care system with emotion recognition function,”* 2011 International Conference on In Electrical and Control Engineering (ICECE), pp. 2995-2998,September, 2011.

[71] *“Face image”* [Online]. Available at http://www.ifreewallpapershd.com/wp-content/uploads/2015/06/keira-knightley-wallpaper2.jpg [Accessed 18 June 2015].

[72] *“3D Figures”* [Online]. Available at http://img1.tebyan.net/Big/1389/09/46329211372172144326520422248158152113142.jpg [Accessed 18 June 2015].

[73] Heart of Healing, *“Benefits of Relaxation”* [Online]. Available at http://heartofhealing.net/relaxation-wellness/benefits-of-relaxation/ [Accessed 7 May 2015].

[74] *“Beach”* [Online]. Available at http://prudentialpalmsrealty5.agentxsites.com/xSites/Agents/PRUDENTIALPALMSREALTY5/Content/UploadedFiles/beachy%20photo.jpg [Accessed 18 June 2015].

[75] *“Sunset image”* [Online]. Available at http://www.freephotoz.net/images/sunset-at-varkala-beach-kerala-india-desktop-wal-800x600.png [Accessed 18 June 2015].

[76] *“Puppies image”* [Online]. Available at http://wallpapers111.com/wp-content/uploads/2015/01/Maltese-Puppies-Wallpapers11.jpg [Accessed 18 June 2015].

[77] *“Kitten image”* [Online]. Available at https://s-media-cache-ak0.pinimg.com/736x/d6/7c/b1/d67cb10a6173f656a2c317f4824b0c90.jpg [Accessed 18 June 2015].

[78] *“Room image”* [Online]. Available at http://farm9.staticflickr.com/8387/8559614842_8d19ff186e_z.jpg [Accessed 18 June 2015].

[79] *“Bubble image”* [Online]. Available at https://jimtsutsui.files.wordpress.com/2013/02/bubble_burst.jpg [Accessed 18 June 2015].

[80] *“Doors image”* [Online]. Available at http://dailyps.com/wp-content/uploads/2014/08/faith.jpg [Accessed 18 June 2015].

[81] *“Zombie image”* [Online]. Available at http://spi2uk.itvnet.lv/upload/articles/77/77082/images/_origin_helovina-12.jpg [Accessed 18 June 2015].

[82] *“Shark image”* [Online]. Available at http://cmex29.ru/uploads/posts/2014-04/1397068675_akula.jpg [Accessed 18 June 2015].

[83] *“Lion image”* [Online]. Available at https://jerryandgod.files.wordpress.com/2012/11/256-a.jpg [Accessed 18 June 2015].

[84] Microsoft, *“IIS”* [Online]. Available at http://www.iis.net/ [Accessed 7 May 2015].

[85] *“Presentation icon”* [Online]. Available at http://ertrade.ru/images/electroset/presentation_ico.png [Accessed 18 June 2015].

[86] K. Jenkins, *“Take a deep breath”* [Online]. Available at http://www.health.harvard.edu/staying-healthy/take-a-deep-breath [Accessed 7 May 2015].

[87] D. Antai-Otong, *“Creative Stress-Management Techniques for Self-Renewa,l”* Dermatology Nursing, vol. 13, no. 1, pp. 31-37, 2001

[88] *“Deep breathing 'cuts stress as much as massage”* [Online]. Available at http://news.bbc.co.uk/2/hi/health/8563078.stm [Accessed 7 May 2015].

[89] *“Music box image”* [Online]. Available at https://s-media-cache-ak0.pinimg.com/originals/a1/cf/3b/a1cf3b3e596b28125e502127f9ec30de.jpg [Accessed 18 June 2015].

[90] *“AR marker image”* [Online]. Available at http://saqoo.sh/a/wordpress/images/2009/01/marker-axis.jpg [Accessed 18 June 2015].

[91] *“Gramophone image”* [Online]. Available at http://img-fotki.yandex.ru/get/5312/20086847.63/0_ed2c0_97259ff0_M [Accessed 18 June 2015].

[92] *“Virtual Music Box”* [Online], Available at http://www.3dmodelfree.com/models/29549-0.htm [Accessed 7 May 2015].

[93] Sh. P. Arunachalam, and L. F. Brown, *“Real-Time Estimation of the ECG-Derived Respiration (EDR) Signal using a New Algorithm for Baseline Wander Noise Removal,”* Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 5681-5684, 2009

[94] P. M. Vyas, and. M. S. Panse, *“Various Methods to Detect Respiration Rate From ECG Using LabVIEW,”* International Conference on Biomedical Engineering and Assistive Technologies (2010)

[95] J. Boyle, N. Bidargaddi, A. Sarela, and M. Karunanithi, *“Automatic Detection of Respiration Rate From Ambulatory Single-Lead ECG,”* IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 6, pp. 890-896, 2009

[96] D. Widjaja, J. Taelman, S. Vandeput, M. Aka Braeken, R. A Otte, B. RH. V.D. Bergh, and S. V. Huffel, *“ECG-Derived Respiration: Comparison and New Measures for Respiratory Variability,”* Computing in Cardiology, pp. 149-152 2010.

[97] *“Fantasia Database,”* Cambridge, MA; PhysioNet [Online], Available at http://www.physionet.org/cgi-bin/atm/ATM [Accessed 7 May 2015].

[98] UM. Nater, and N. Rohleder, *“Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: current state of research,”* Psychoneuroendocrinology, vol. 34, no. 4, pp. 486-496, 2009

[99] N. Takai, M. Yamaguchi, T. Aragaki, K. Eto, K. Uchihashi, and Y. Nishikawa, *“Effect of psychological stress on the salivary cortisol and amylase levels in healthy young adults,”* Archives of Oral Biology, vol. 49, no. 12, pp. 963-968, 2004

[100] N. Rohleder, U. M. Nater, J. M. Wolf, U. Ehlert, and C. Kirschbaum, *“Psychosocial Stress Induced Activation of Salivary Alpha-Amylase: An Indicator of Sympathetic Activity?”*, Annals of the New York Academy of Sciences; vol. 1032, no. 1, pp. 258-263. 2006.

[101] M. Yamaguchi, T. Kanemori, M. Kanemaru, N. Takai, Y. Mizuno, and H. Yoshida, *"Performance evaluation of salivary amylase activity monitor,"* Biosensors and Bioelectronics, vol. 20, no. 3, pp. 491-497, 2004

[102] K. Jenkins, *"Take a deep breath"* [Online]. Available at http://extension.missouri.edu/silverthreads/10/Deep%20Breathing%20for%20Good%20Health.htm [Accessed 7 May 2015].

[103] *"Virtual Gramophone"* [Online], Available at http://www.3dmodelfree.com/models/29549-0.htm [Accessed 7 May 2015].

[104] *"Virtual Castle"* [Online], Available at http://tf3dm.com/3d-model/castle-43181.html [Accessed 7 May 2015].

[105] T. Prasartvit, B. Kaewkamnerdpong, and T. Achalakul, *"Dimensional reduction based on artificial bee colony for classification problem,"* Bio-Inspired Computing and Applications, eds. by DS Huang, Y Gan, P Premaratne and K Han, Springer Berlin Heidelberg, pp. 168-175, 2012.

[106] C. Cortes, and V. Vapnik, *"Support-vector networks,"* Mach. Learn. 20, pp. 273-297, 1995, doi: 10.1007/BF00994018

[107] C. J. Burges, *"A tutorial on support vector machines for pattern recognition,"* Data mining and knowledge discovery, vol. 2, no.2, pp. 121-167, 1998.

[108] Wikipedia, *"Support vector machine"* [Online]. Available at http://en.wikipedia.org/wiki/Support_vector_machine [Accessed 7 May 2015].

[109] *"EmguCv"* [Online]. Available at http://www.emgu.com [Accessed 7 May 2015].

[110] M. J. Lyons, M. Kamachi, and J. Gyoba, *"Japanese Female Facial Expressions (JAFFE) (2000),"* [Online]. Available at http://www.kasrl.org/jaffe.html [Accessed 7 May 2015].

[111] D. Lundqvist, A. Flykt, and A. Öhman, *"The Karolinska Directed Emotional Faces – KDEF,"* CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

[112] J. Zhan, and H. Loh, *"Using Redundancy Reduction in Summarization to Improve Text Classification by svms,"* Journal of information science and engineering, 25, 591-601, 2009.

[113] P. Ekman, *“Universal facial expressions of emotion,”* California Mental Health Research Digest, vol. 8, no. 4, pp. 151-158, 1970.

[114] RapidMiner, Inc., *“RapidMiner Studio 5”* [Online]. Available at http://rapidminer.com [Accessed 7 May 2015].

[115] W. S. Noble, “*What is a support vector machine?*,” Nature biotechnology, vol. 24, no. 12, pp. 1565-1567, 2006.

[116] B. Sunita, and J. Arora, *“Classification and feature selection techniques in data mining,”* International Journal of Engineering Research and Technology, vol. 1, no.6, ESRSA Publications, 2012.

[117] X. Wu , V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, and D. Steinberg, *“Top 10 algorithms in data mining,”* Knowledge and Information Systems, vol. 14, no.1, pp. 1-37, 2008.

[118] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, *“KNN model-based approach in classification,”* On The Move to Meaningful Internet Systems: CoopIS, DOA, and ODBASE Springer Berlin Heidelberg, pp. 986-996 , 2003.

[119] J. Wagner, J. Kim, and E. André, *“From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification,”* IEEE International Conference on Multimedia & Expo (ICME), 2005.

[120] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, *“Physiological signals based human emotion recognition,”* IEEE 7th International Colloquium on Signal Processing and its Applications (CSPA), pp. 410-415, 2011.

[121] A. Russell James, *“A circumplex model of affect,”* Journal of personality and social psychology, vol. 39, no. 6, pp. 1161, 1980.

[122] SweetWater Health, *“Stress and Heart Rate Variability”* [Online]. Available at http://www.sweetwaterhrv.com/documentation/stressandhrv.pdf [Accessed 7 May 2015]

[123] J. F. Thayer, F. Åhs, M. Fredrikson, J. J. Sollers, and T. D. Wager. *“A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health,”* Neuroscience & Biobehavioral Reviews, vol. 36, no. 2, pp. 747-756, 2012.

[124] MEDICORE, *"Heart Rate Variability Analysis System,"* [Online]. Available at http://medi-core.com/download/HRV_clinical_manual_ver3.0.pdf [Accessed 7 May 2015].

[125] P. J. Lang, M. M. Bradley, and B.N. Cuthbert, *"Affective ratings of pictures and instruction manual,"* International affective picture system (IAPS), Technical Report A-8, University of Florida, Gainesville, FL, 2008.

Publication List

Journal Papers

[1] **S. Tivatansakul**, and M. Ohkura, “Healthcare System Focusing on Emotional Aspect using Augmented Reality- Relaxation Service –”, Transactions of Japan Society of Kansei Engineering, vol. 13, no. 1, pp. 191-201, 2014.

[2] M. Ohkura, **S. Tivatansakul**, and K. Akimoto, “Kawaii Spoons and Heart Rates of Elderly”, The IEICE transactions on information and systems, vol. 97, no. 1, pp. 177-180, 2014 (in Japanese).

[3] **S. Tivatansakul**, and M. Ohkura, “Emotion recognition using ECG signals with local pattern description methods”, International Journal of Affective Engineering (in submission).

[4] **S. Tivatansakul**, S. Maneewongvatana, S. Prom-on, T. Achalakul and M. Ohkura, “Facial expression recognition using complementary directional ternary pattern” (Plan to submit).

International Conference Papers

[5] M. Ohkura, T. Komatsu, **S. Tivatansakul**, S. Charoenpit, and S. Settapat, “Comparison of Evaluation of Kawaii Ribbons between Genders and Generation of Japanese”, in Proceeding of IEEE/SICE International Symposium on System Integration, pp. 824-828, Fukuoka, Japan, Dec. 2012.

[6] **S. Tivatansakul**, and M. Ohkura, “Healthcare System Design Focusing on Emotional Aspects using Augmented Reality – Relaxed Service Design-”, IEEE Symposium on Computational Intelligence in Healthcare and e-Health (CICARE), pp. 88-93, Singapore, Apr. 2013.

[7] **S. Tivatansakul**, and M. Ohkura, “Healthcare System Focusing on Emotional Aspects Using Augmented Reality-Implementation of Breathing Control Application in Relaxation Service”, in Proceedings of the IEEE International Conference on Biometrics and Kansei Engineering (ICBAKE 2013), pp. 218-222, Tokyo Japan, July. 2013.

[8] **S. Tivatansakul**, and M. Ohkura, “Healthcare System Focusing on Emotional Aspect Using Augmented Reality: Control Breathing Application in Relaxation Service”, in Proceeding of The 15th International Conference HCI International 2013, vol. 8005, pp. 225-229, Las Vegas, NV, USA, July. 2013.

[9] **S. Tivatansakul**, T. Achalakul and M. Ohkura, “Healthcare System Focusing On Emotional Aspects Using Augmented Reality -Re-Design of Breathing Control Application in Relaxation Service-”, in Proceeding of The 5th International Congress of International Association of Societies of Design Research (IASDR 2013), pp. 4920-4930, Tokyo, Japan, Aug. 2013.

[10] **S. Tivatansakul**, G. Chalumporn, S. Puangpontip, Y. Kankanokkul, T. Achalakul, and M. Ohkura, “Healthcare System Focusing on Emotional Aspect Using Augmented Reality: Emotion Detection by Facial Expression”, in Proceeding of The 5th International Conference Applied Human Factors and Ergonomics (AHFE 2014), pp. 5674-5683, Krakow, Poland, July 2014.

[11] **S. Tivatansakul**, S. Puangpontip, T. Achalakul, and M. Ohkura, “Emotional healthcare system: Emotion detection by facial expressions using Japanese database”, in Proceeding of 6th Computer Science and Electronic Engineering Conference (CEEC), pp. 41-46, Colchester, UK, Sep. 2014.

[12] **S. Tivatansakul**, and M. Ohkura, “The design, implementation and evaluation of a relaxation service with facial emotion detection”, in Proceeding of 2014 IEEE Symposium Series on Computational Intelligence (SSCI 2014), pp. 40-47, Orlando, Florida, USA, Dec. 2014.

[13] **S. Tivatansakul**, and M. Ohkura, “Emotional healthcare system –The emotion detection using ECG signal with local pattern description methods –”, International Symposium of Affective Science and Engineering (ISASE 2015), Tokyo, JAPAN, Mar. 2015.

[14] T. Laohakangvalvit, M. Ohkura, T. Achalakul and **S. Tivatansakul**, “Electronic Patient Referral System With Data Exchange And Fingerprint Patient Identification”, in the 17th Japan-Korea Joint Symposium on Ergonomics, Tokyo, Japan, June 2015.

[15] T. Laohakangvalvit, M. Ohkura, T. Achalakul and **S. Tivatansakul**, “Electronic Patient Referral System With Data Exchange And Fingerprint Patient Identification”, in the 9th South East Asian Technical University Consortium, Thailand, July 2015.

[16] **S. Tivatansakul**, and M. Ohkura, “Improvement of emotional healthcare system with stress detection from ECG signal”, in the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, Italy, August 2015.

[17] T. Laohakangvalvit, **S. Tivatansakul**, T. Achalakul, and M. Ohkura, “Implementation and Evaluation of Electronic Patient Referral System With Data Exchange And Fingerprint Patient Identification”, in the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, Italy, August 2015.

Workshop Paper

[18] **S. Tivatansakul**, and M. Ohkura , “Healthcare System Design Focusing On Emotional Aspects Using Augmented Reality-Implementation of Control Breathing Application in Relaxed Service-”, in Intensive Workshop of the 7th South East Asian Technical University Consortium, pp. 15-18, Bandung, Indonesia, Mar. 2013.

Domestic Conference Paper

[19] **S. Tivatansakul**, G. Chalumporn, S. Puangpontip, Y. Kankanokkul, T. Achalakul and M. Ohkura, “Healthcare System Focusing on Emotional Aspect using Augmented Reality－Emotion Detection by Facial Expression－”, in The Institute of Electronic Information and Communication Engineers (IEICE), Technical report, vol. 114, no. 67, pp. 151-156, Okinawa, Japan, May. 2014.