

Research Article

Real-Time Mobile Emotional Content Player Using Smartphone Camera-Based PPG Measurement

Haena Lee¹ and Dong Keun Kim² 

¹Identity and Access Management Development Team, NCSOFT Corporation, Gyeonggi-do, Republic of Korea

²Department of Intelligent Engineering Informatics for Human, Sangmyung University, Seoul, Republic of Korea

Correspondence should be addressed to Dong Keun Kim; dkim@smu.ac.kr

Received 12 June 2018; Revised 26 February 2019; Accepted 9 April 2019; Published 5 May 2019

Academic Editor: Salvatore Carta

Copyright © 2019 Haena Lee and Dong Keun Kim. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A real-time mobile content player was developed that can recognize and reflect emotions in real time using a smartphone. To determine effective awareness, a photoplethysmogram (PPG), which is a biological signal, was measured to recognize emotional changes in users presented with content intended to induce an emotional response. To avoid the need for a separate sensor to measure the PPG, PPG signals were extracted from the red (*R*) values of images acquired by the rear camera of a smartphone. To reflect an emotion, the saturation (*S*) and brightness (*V*) levels, which are related to the ambience of a content, are changed to reflect the emotional changes of the user within the content itself in real time. Arousal- and relaxation-inducing scenarios were conducted to validate the effectiveness. The sample *t*-test results show that the average peak-to-peak interval (PPI), which is the time interval between the peaks of PPG signals, was significantly low when viewing the content under the arousal-inducing scenario as compared to when watching regular content, and it was determined that the emotion of the user was led to a state of arousal. Ten university students (five males and five females) participated in the experiment. The users had no cardiac disease and were asked not to drink or smoke before the experiment. The average PPI was significantly higher when the content was viewed in the relaxation-inducing scenario compared to regular content, and it was determined that the emotion of the user was induced to a state of relaxation. The designed emotional content player was confirmed to be an interactive system, in which the video content and user concurrently affect each other through the system.

1. Introduction

Owing to the advancements in information technology, smartphones have become popular devices for accessing services in various fields. In particular, the demand for services using multimedia content is increasing, and research works show that such multimedia content induces an emotional response from the users. For example, one study reported that fear-related content induces the emotional responses of fear, arousal, and displeasure, and relaxation-related content induces the emotional response of relaxation [1–9]. Currently, however, users want to use multimedia content provided according to their circumstances or emotional state rather than uniform content that does not consider their emotional responses [10, 11]. In other words, recent studies have focused on methods to induce the

emotional responses of users and reflect those responses within the content, just as the content can induce emotional responses in the users. Emotional content has been defined as content that not only affects the emotion of the user but also changes the audiovisual elements of the image, such as tone, brightness, contrast, and the strength and speed of sound, within the content according to the user's emotion [12, 13]. In this regard, recent studies have been recently conducted on emotional content that can interact with users by recognizing their emotional states and changing the elements of the content accordingly.

Conversely, in conventional studies, the emotional states of users have been recognized through biological signal analyses. Such studies have a limitation in that an attachment type sensor is required to measure the biological signals for effective awareness. Furthermore, the contents

are usually provided by changing the color element to reflect an emotion; however, users may be biased toward a particular color element. Hence, based on the connotations associated with specific colors, there is a risk that the original intended meaning of a particular content can be changed considerably. In addition, there have been no studies on changing the saturation (S) and brightness (V) of content in a smartphone in real time. Therefore, there is need for a system that can recognize and reflect user emotions in real time on a smartphone and does not require without a special sensor. Accordingly, a real-time mobile emotional content player was developed in this study. A user's emotion can be recognized by measuring and analyzing the photoplethysmogram (PPG) signals through a camera in a smartphone configuration. A PPG is an optically obtained physiological signal that can be used to detect blood volume changes [14]. PPG signals can also be used to monitor cardiac cycles and circulatory conditions such as heart rates (HRs). Furthermore, the emotion of a user can be reflected by changing the S and V of the content in real time according to the acquired user emotions. The emotional content player, which recognizes and reflects emotions in real time, can provide interactive content whereby the users and content can interact with each other.

In this study, a real-time mobile emotional content player was developed to provide interactive content that can affect and be affected by users concurrently when content is viewed on a smartphone configuration. To clarify the purpose of this study, we suggested the following hypotheses:

- (i) The mobile user can measure HRs using a rear camera
- (ii) The measured HRs can be used to reflect the emotional changes of users
- (iii) The emotional changes affect the changes in V and S values of video content in the smartphone in real time

An emotion induced by a particular content affects the emotional changes in the users. To determine a user's affective awareness, PPG signals are measured by obtaining the R value of video frames from a smartphone camera in real time. The peak-to-peak intervals (PPIs) of PPG signals are calculated using a sliding window technique and are used to calculate HR. In addition, based on the average PPI of the user, the user's emotional state is determined to be either a state of arousal or relaxation through a comparison with the current PPI. The arousal and relaxation emotions of a user acquired in this manner are reflected in the content in real time. To acquire the emotions for each frame of video, the demuxing and decoding processes are applied using FFMPEGLibrary. In this study, because FFMPEGLibrary was coded in the C programming, the native development kit (NDK) method was used on the mobile operating system design. The NDK is a tool that allows developers to program Android mobile devices in C or C++ languages. It is useful for supporting the CPU intensive operations, such as mobile video game, and signal processing and controlling the hardware in smartphone devices. Thus, the emotions of the user can be reflected by changing the S and V values for every frame of video content

without the need for a special sensor. Furthermore, it was determined that a user and content can interact with each other through this system. Arousal- and relaxation-inducing scenarios are proposed and verified experimentally.

2. Related Work

2.1. Emotional Responses and Physiological Signals. A subjective method using a survey and an objective method using physiological responses are applied to measure and determine the emotions of a user. In a physiological response application, the response is involuntary and cannot be controlled by an individual; hence, quantification is convenient, and objective measurements are possible [14–16]. Therefore, to recognize and determine the user emotions, many studies are being conducted using biological signals, which are physiological responses. By extracting the signal features from biological signals, a psychological emotional state can be determined [14–18].

Studies have been conducted using the central nervous system responses or automatic nervous system responses from biological signals [18–23]. Electroencephalograph (EEG) signals are usually used in studies on central nervous system responses [19–23]. A study was conducted that implemented an emotion inference system by measuring, analyzing, and evaluating an EEG while a user was experiencing an audio and video stimulus [19–23]. Furthermore, a study that measured the biological signals of an EEG and an electrocardiograph (ECG) displayed the emotional states of tension, irritation, anger, joy, and surprise on the power spectrum graph of the signals [24, 25]. However, to measure EEG signals, the electrical activities of the brain can be measured only by accurately attaching a sensor to a fixed position on the head; consequently, user discomfort occurs owing to movement limitations [13,15–22].

Similarly, studies utilizing automatic nervous system responses from biological signals have also been conducted. ECG, PPG, and skin temperature (SKT) signals are mainly used in automatic nervous system responses. A study was conducted on acquiring emotion information in real time by measuring the biological signals of a PPG, SKT, and galvanic skin response (GSR), and a real-time emotion information sharing system was developed [10–13, 20–32]. However, to measure ECG signals, a sensor has to be attached close to the heart, and accordingly, there are disadvantages in that the user may feel repulsion and experience a limitation in their activities [32]. PPG and SKT signals can be measured relatively easily by attaching a sensor to the fingertip; however, like other methods, there is a disadvantage: the use of such a sensor is always required to recognize an emotion. Hence, to solve this problem, this study seeks a means to conveniently acquire biological signals without repulsing the user through the use of a smartphone camera instead of a sensor. PPG signals are measured from biological signals, which can be acquired through the R values of a camera image.

2.2. Emotional Model and HSV Model. An emotion is a complex emotional state inside a human being [31], and

various emotions are expressed mainly in two-dimensional (2D) space. In particular, Russell proposed a 2D emotion model, shown in Figure 1, and stated that an emotion distribution can be expressed through two axes, i.e., the X -axis for pleasure/displeasure and the Y -axis for arousal/relaxation. In this study, because PPG signals can be measured using a smartphone camera, the emotions of the users are evaluated using PPG signals. In a previous study, it was verified that PPG signals can be used as an objective metric for determining states of arousal or relaxation [14–16]. Therefore, in this study, arousal and relaxation, which are on the Y -axis of the 2D model, are also evaluated using PPG signals.

Conversely, the color models studied with respect to emotion are mainly divided into a red, green, and blue (RGB) model and a hue, saturation, and value (HSV) model. The RGB model is a method for expressing colors by applying the additive color principle of the three primary colors of light. Red (R), green (G), and blue (B) each has a value of 0–255, and all colors can be expressed through RGB ratios [32–34]. Studies have been conducted on the relationship between an emotional state and RGB.

In addition, Yang et al. converted colors of a certain region to intensify the emotions of a content [33]. As a result, it was concluded that the emotions are intensified and the sense of immersion is increased.

The HSV model representation is similar to the way that humans perceive colors. In the HSV model, hue (H) has a range of 0–360° and is a ring-shaped relative alignment angle, having an R of 0° (R having the longest wavelength). The saturation (S) has a range of 0–100% and shows the thickness of a color. The value (V) has a range of 0–100% and shows the level of brightness. Other studies have examined the relationship between an emotional state and HSV. Jang et al. extracted feature points using the intrinsic information possessed by an image and designed a model that classifies the emotions based on these points [35]. Lee et al. confirmed whether the emotion of the user is influenced through color changes in the lighting and derived a procedure for using the lighting color as a design element [36]. In addition, through an experiment using three types of medium (paper, textile, and a computer monitor), Seok found that there is a positive correlation (between saturation and pleasure and between saturation and excitement, through an experiment) on the characteristics of emotional responses according to the changes in color saturation [37].

Previous studies were also conducted on the relationship with emotion by focusing mainly on color elements such as R, G, B, and H. However, if a color element is substantially changed, there is a limitation: a user may feel a sense of repulsion. Furthermore, because the color of an image has an associated meaning shown through the color itself, there is a risk that the originally intended meaning of a particular content can be changed significantly when its color is changed. When combined, the S and V are expressed as a tone, which is mainly related to the overall ambience of a particular content. Therefore, changing the S and V levels may not necessarily mean changing the intended meaning of the content; rather, there is a possibility that the ambience of

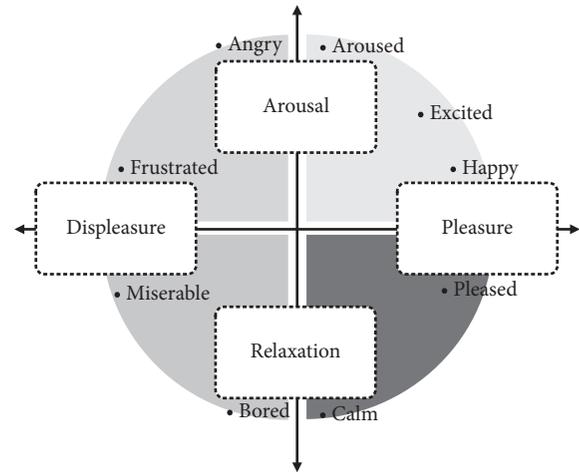


FIGURE 1: Diagram of the two-dimensional emotion model.

the content can be brought out naturally. Hence, this study will examine the relationship among S , V , and arousal/relaxation, in which repulsion and its influences are relatively small and previous research is insufficient.

2.3. Emotional Contents. Emotional content refers to a content that can accelerate, maintain, or reduce the existing emotional state of a user by recognizing the state, i.e., pleasure, displeasure, arousal, or relaxation, and based on that recognized state, change the audiovisual elements of the content accordingly [38]. In other words, a content that understands the user's emotions and responds appropriately through an emotional communication between the content and user can be called an emotional content [12].

Many studies have been conducted with respect to emotional content. For example, when viewing an emotional content, the emotional changes of the user are measured through SKT, GSR, PPG, and EEG sensors, and based on the results, a larger emotional response is induced by providing a visual feedback to the content [13, 30, 31]. As a result, a larger emotional response is induced in a three-dimensional (3D) image compared to a 2D image and in a bright image compared to a dark image. In addition, Lee and Kim classified the emotions of a user through the biological PPG, SKT, and GSR signals and, based on these classification, verified the effect of a real-time interactive emotional content player by changing and providing the image contents according to the predefined color model rules [26]. Yang et al. transformed a partial image to intensify the emotion of a content and confirmed through subjective evaluations that, compared to a regular content, emotional contents intensify users' emotions and increase their sense of immersion [33].

In conventional studies, systems have been developed that require a separate part for recognizing the emotion of a user through a sensor and another for viewing emotional content. Thus, the disadvantages are as follows: two systems are required, and communication has to be established. Therefore, in this study, a smartphone was used to recognize the emotional state of the user as either arousal or relaxation. Based on this application, a real-time mobile emotional

content player was developed to play an emotional content that changes in terms of S and V , the visual elements of the content.

3. Implementation of Mobile Emotional Content Player

3.1. System Design and PPG Signal Processing. When a light is projected onto an area of the human body with thin skin, such as the tip of a finger or an ear lobe, some of the light is absorbed in the blood, bone, and tissue, and the remaining light penetrates the skin. Therefore, a change in the intensity of the penetrated light can be said to be a reflection of the change in blood flow [14–17]. Here, PPG signals can be measured because the amount of light absorbed by the skin, according to heart contraction and relaxation, differs owing to changes in blood flow.

A PPG signal measurement using the absorption of light follows the Beer–Lambert law. This law shows a correlation between the absorbed and penetrated light when the projected light passes through a homogeneous medium [14, 41]. That is, when a light produced from a light emitting diode (LED) is projected onto the skin of a peripheral body part, some of the light passes through the blood and is transferred to a sensor, and some of the remaining light is absorbed by the blood. Hence, this law states that the sum of the intensity of light transferred to a sensor and the intensity of light absorbed by the blood is the same as the intensity of light when originally projected [14, 41]. This law can be expressed through the following equation:

$$I_t = I_0 \times e^{-\varepsilon(\lambda)cd}, \quad (1)$$

where I_t is the intensity of penetrated or reflected light and I_0 is the intensity of light when absorbed or incidented. In addition, $\varepsilon(\lambda)$ is the extinction coefficient corresponding to the optical absorption rate of a tissue, c is the concentration of light absorbed by the tissue, and d is the length of the optical path [14, 41]. When a certain amount of light is projected onto the skin of a peripheral body part in a fixed posture, $\varepsilon(\lambda)$ and c are generally constant and d changes according to the contraction and relaxation of the heart. The factor that changes d is the change in blood flow at the peripheral body part; the PPG can be detected using this factor [14, 41].

In general, a PPG signal graph looks like that shown in Figure 2. Different information can be obtained from the PPG signals. When the heart contracts, the amount of blood increases; hence, more light is absorbed in the blood, and the intensity of the penetrated light is decreased.

Therefore, the shape of the PPG signal is decreasing in a cycle graph. When the heart relaxes, the amount of blood is relatively decreased, and less light is absorbed in the blood. Therefore, the intensity of the penetrated light shows an increasing shape within the cycle. However, the amplitude is the measured height from a peak, which is the highest point of the current cycle in a PPG signal, to a trough, which is the lowest point of the previous cycle; in addition, the PPI is the measured distance between the peak of the current cycle and peak of the previous cycle. The amplitude and PPI values

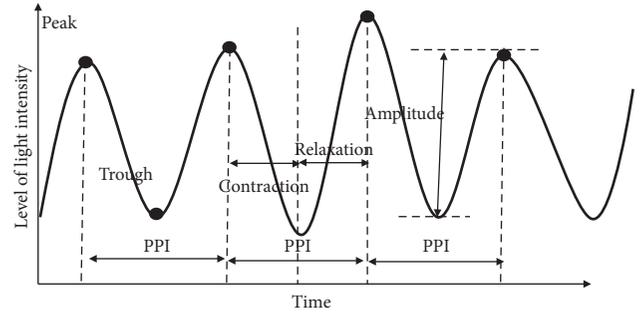


FIGURE 2: Example of the photoplethysmogram (PPG) graph.

change continuously according to the state of the user. In particular, when the PPI decreases, it indicates that the sympathetic nerve has been activated and the heart beat is fast, which indicates that the user is close to a state of arousal. Conversely, when the PPI increases, it indicates that the parasympathetic nerve has been activated and the heart beat is slow, which indicates that the user is close to a state of relaxation [14, 39–44]. That is, the HR can be calculated through changes in the PPI, and the psychological state of the user can be determined as either arousal or relaxation.

PPG signal measurement sensors are mainly divided into penetration- and reflection-type sensors. A penetration-type sensor consists of an LED that projects light and a photodiode that measures the intensity of light that penetrates the skin; the light projection and receiving parts face each other. A reflection-type sensor consists of an LED that projects light and a photodiode that measures the intensity of light reflected after hitting the skin; the light projection and receiving parts are placed side by side [41].

The flash of a smartphone can be a substitute for the LED of a sensor, and a smartphone’s rear camera can be a substitute for the photodiode of a sensor. Because the flash and camera of a smartphone are located side by side, they can be viewed as a reflection-type sensor; therefore, PPG signals can be measured from a smartphone based on the principle of a reflection-type sensor. In other words, when the light from a flash is projected onto the skin, some of the light is absorbed according to the change in blood flow, and some of the remaining light is reflected and used by the smartphone camera. PPG signals can be acquired from images of fingers measured in this manner. Accordingly, because the image quality of a camera can be an important element, a preliminary experiment was conducted considering the three properties of resolution, bitrate, and flash with respect to image quality.

In a previous study, the results of a correlation analysis of PPG signals acquired through the sensor and PPG signals acquired using the R, G, B, H, S, and I components of smartphone camera images showed that the signals acquired through the R component had the highest correlation [45]. Therefore, in this study, the PPG signals were measured from camera images using the R component values.

3.2. Video Image Processing. A video sequence is literally a collection of moving images and can be described as a series

of images composed of one stationary image after another. One stationary image is called a frame, and the number of frames played per second is called the frame rate. At 12 frames per second and faster, the human visual system cannot recognize individual stationary images and perceives them as a moving image. As the frame rate decreases, the video size decreases. However, there is a tendency for images to be shown discontinuously. Thus, as the frame rate increases, the pictures appear smoother; however, the video size tends to increase. Therefore, compression is necessary to create a small-size high-quality video sequence. As shown in Figure 3, converting an original video sequence, comprising video pixel data, into a compressed video sequence, i.e., a video bitstream, using a certain codec is called encoding and decompression conducted as a reverse process is called decoding. A codec used in these processes refers to the software or algorithm that can compress the media information. Typical video codecs include H.264 and DivX, and a typical audio codec is mp3. Putting an encoded video sequence in a certain format and converting it into a video file is called muxing, and the reverse process of extracting video and audio data from such a format is called demuxing. The format used in this process is usually called a container, which is also known as the file type or extension. A container acts as a box producing both a video codec and an audio codec in a single file; typical container formats include AVI and mp4.

3.3. System Implementation. To calculate the HR from PPG signals and classify it as either a state of arousal or relaxation, a signal processing procedure is conducted out using the sequence in Figure 4.

In the first sequence, a peak must be detected from the PPG signals. However, in the PPG signals extracted from the R values, the baseline is continuously changed, and moreover, noise exists in the signals. This noise can result from the user's breathing or other tiny movements that occur when PPG signals are being measured. Because noises can induce errors in all processes, from peak detection to HR calculation and the determination of arousal/relaxation, a removal process must be applied before calculating the PPI.

A PPG signal is mainly activated in a low-frequency band, and noise is mainly activated in a high-frequency band [15]. Therefore, noise mixed in a PPG signal can be removed using a low-pass filter, resulting in a clean PPG signal [15, 42]. Typically, when a low-pass filter is used, a moving average filter is also used. A moving average filter is a tool for calculating an average value by using partial filter-sized data from recent data measurements, instead of using all measured data [42]. An equation related to a moving average filter is shown in the following equation:

$$R_n = \frac{1}{n} \sum_{m=0}^{n-1} R_m. \quad (2)$$

In this study, a moving average filter with a filter size of seven was used to correct the baseline and remove any noise, and the R values calculated frame by frame from the camera

images were used as the input data. When PPG signals acquired from the R values are inputted from right to left, the filtering is performed by calculating the average of seven R values for the equivalent filter size.

Figure 5(a) shows the PPG signals before applying the moving average filter, and Figure 5(b) shows the PPG signals after the filter is applied. The X -axis in each graph is time, and the Y -axis shows the R values ranging from 0 to 255. To observe the filtering changes intuitively, the intervals of the Y -axis were consistently maintained. As a result, the baseline was maintained more consistently in Figure 5(b) than in Figure 5(a), and most of the noise was removed. Although it may appear that the signals were not produced well in Figure 5(b), when the intervals were divided into smaller sizes, it was confirmed that the signals were properly produced.

Accurate peaks can be detected through the filtered PPG signals. In this study, the sliding window method was used to detect accurate peaks in real time. In general, the sliding window method is used to control the flow of packets communicated between two hosts, and the number of packets matching the window size are sent concurrently instead of one by one. Thereafter, this method sends the next packets by moving the window by as much as the resolution. Using this principle, the sliding window method was used in this study with a window size of 30 and resolution of one. After reading the number of filtered PPG signals equal to the window size, a peak is detected in one window. The window moves horizontally by one resolution of spacing, and peaks are continuously detected.

To detect a peak, primary differentiation was conducted within one window; a point where the slope sign changes is regarded as a peak. However, because the window moves in one resolution spacing, the same peaks can be detected in the window. Therefore, only when a peak is different from a previous peak, it is detected as a corrective peak, and the time of this point is saved. Figure 6 is a flow chart of the peak detection process.

After detecting the peaks, the PPI, which is the time interval between peaks, can be calculated. The difference of the n th peak at the $(n + 1)$ th peak is determined to be the PPI at the time point of n [38].

A graph showing the continuous changes in PPI calculated in this manner is called the pulse rate variability (PRV) [37]. HR refers to the number of contractions and relaxations repeated by the heart in one minute (60 s) [44].

Conversely, the emotion of a user can be determined as either arousal or relaxation through changes in the PPI. If a user is in a state of arousal, the PPI of the PPG signals is decreased because the sympathetic nerve is activated and the HR is increased. Conversely, if a user is in a relaxed state, the PPI of the PPG signals is increased because the parasympathetic nerve is activated and the HR is decreased [14, 15]. Therefore, after setting the average PPI of a user as a reference, if the current PPI becomes lower than the average PPI, the emotional state of the user may be determined as arousal. If the current PPI becomes higher than the average PPI, the emotional state of the user may be determined as relaxation.

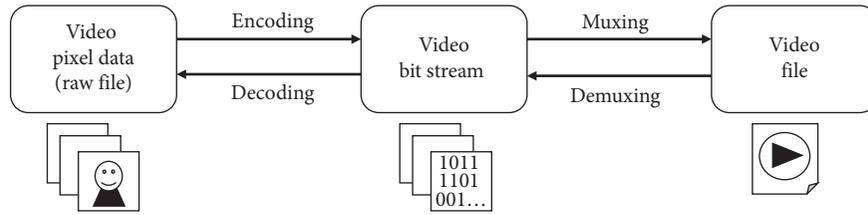


FIGURE 3: Example of video processing steps: encoding, decoding, muxing, and demuxing.

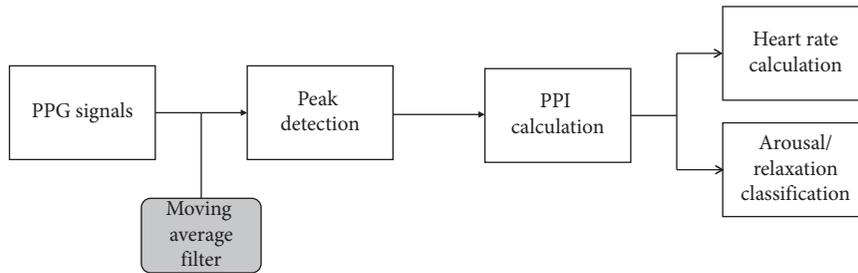


FIGURE 4: Procedure for processing PPG signals to calculate HR and classify arousal and relaxation.

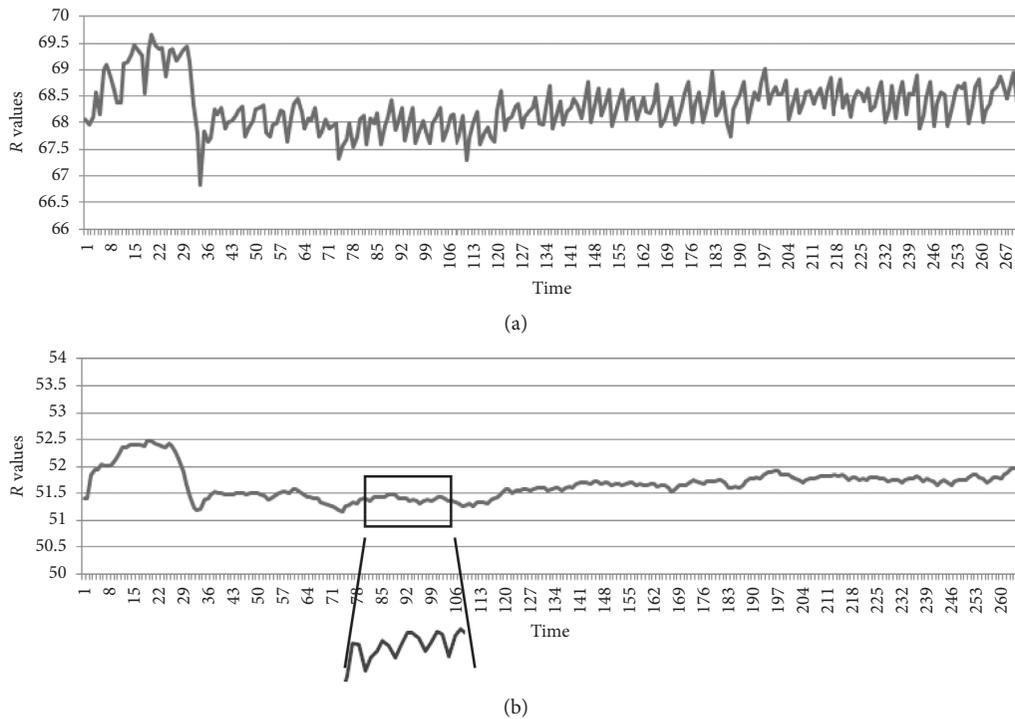


FIGURE 5: Moving average filter for PPG signals: (a) before applying the moving average filter; (b) after the filter is applied.

To change the *S* and *V* levels in real time, the original pixel data have to be accessed such that the video file can be read one frame at a time. Afterward, the *S* and *V* levels are changed and shown on the screen through a color space conversion of the pixel data. The demuxing and decoding processes have to be applied to access the pixel data in a video. The demuxing process extracts a video data stream from a video file, and the decoding process extracts the original pixel data from the video data stream using a certain

video codec. A specific FFMPEGLibrary is needed for these processes: (a) through the libavformat library, a video file goes through the demuxing process for a video bitstream; and (b) through the libavcodec library, a video bitstream goes through the decoding process for the pixel data.

However, the pixel data are YUV pixel 0data, which have to be converted into HSV pixel data to adjust the *S* and *V* levels of the content. Hence, after the conversion into RGB pixel data, they are first converted into HSV pixel data,

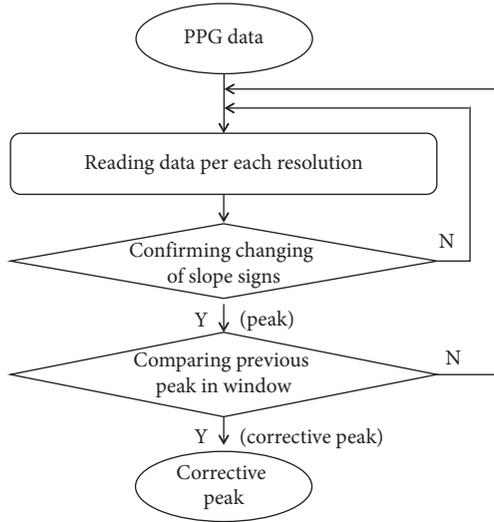


FIGURE 6: Flow chart of the PPG signal peak detection process.

thereby changing the S and V values. They are then converted into RGB pixel data again so that the content can be displayed on the screen. This process is shown in Figure 7.

3.4. Experiment. The real-time mobile emotional content player measures and analyses PPG signals in a smartphone and classifies the emotional state of the user as either arousal or relaxation. According to the classified emotion, the content is played by changing the S and V values. The arousal- and relaxation-inducing scenarios will be implemented, and the effectiveness of the system will be verified through experiments. In many conventional studies, it is reported that an arousal state corresponds to an increase in S and V , and a relaxation state corresponds to a decrease in S and V [45]. Accordingly, in this study, the rules are set to increase the S and V values of the content to induce an arousal state and decrease the S and V values of the content to induce a relaxed state.

An experiment was conducted to verify the effectiveness of the arousal- and relaxation-inducing scenarios using the real-time mobile emotional content player. The subjects who participated in the verification of the applied system comprised ten university students: five males and five females, with an average age of 25.5. None of the subjects had cardiovascular disease, and they were asked not to drink or smoke the day before the experiment.

As shown in Figure 8, a hand of each subject was measured at the height of their heart for an accurate PPG measurement, and the experiment was conducted by placing their left index finger on the rear camera of the smartphone. The content used in the experiment did not contain any emotional element and is a standard demonstration video with no storyline, as shown in Figure 9.

The arousal-inducing scenario aims to make the user feel a stronger sense of arousal while watching the content by increasing the S and V values when the user is in a relaxed state. The effectiveness of the arousal-inducing scenario will be examined by comparing a regular player that does not change content with the arousal-inducing scenario player.

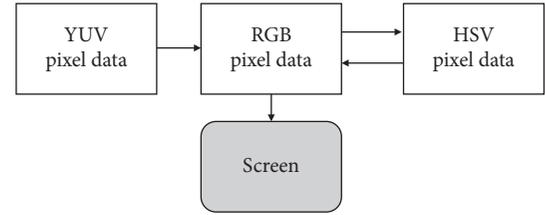


FIGURE 7: Color space changing procedure for saturation and brightness.

The relaxation-inducing scenario aims to make the user feel a stronger sense of relaxation while watching the content by decreasing the S and V values when the user is in an arousal state. The effectiveness of the relaxation-inducing scenario will be examined by comparing a regular player that does not change content with the relaxation-inducing scenario player.

Before starting the experiment, the purpose, procedure, and method of the experiment were explained to the subjects. The left index finger of each subject then touched the rear camera of the smartphone. After measuring the average PPI of each subject for 2 min, the subject watched different versions of the video sequence in the following order: regular playback, arousal-inducing scenario playback, and relaxation-inducing scenario playback for a period of 2 min 30 s each. A 2 min rest was given between the playback periods to minimize the effect of the previous playback.

3.5. Test Scenario. In the arousal-inducing scenario, the content is played, with the intention of further intensifying the arousal emotion. Figure 10 shows the process of the arousal-inducing scenario: (i) while the user is watching a content, changes in their emotional states are affected; (ii) simultaneously, by analyzing the PPG signals measured through the smartphone camera, the emotion of the user is determined as a state of either arousal or relaxation; (iii) only when an emotional state of relaxation is reached are the S and V values of the content increased (in the case of arousal, the content is played without a change in the S and V); (iv) the changed content affects the user again, thereby inducing an emotional change in their arousal.

In the relaxation-inducing scenario, a content with a goal of further intensifying the emotion of relaxation is displayed. Figure 11 shows the process of the relaxation-inducing scenario: (i) while the user is watching the content, changes in their emotional states are affected; (ii) simultaneously, by analyzing the PPG signals measured using a smartphone camera, the emotion of the user is determined as a state of either arousal or relaxation; (iii) only when an emotional state of arousal is reached are the S and V values of the content decreased (in the case of relaxation, the content is played without a change in the S and V); (iv) the change in content affects the user again, thereby inducing an emotional change into a state of relaxation.

4. Results

4.1. Mobile Emotional Content Player. The graphical user interface of the mobile emotional content player developed

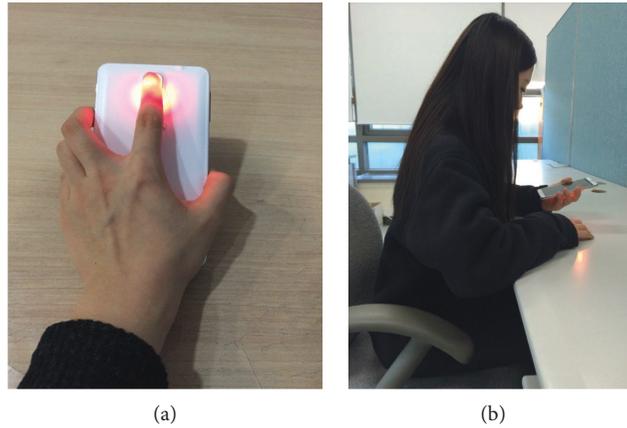


FIGURE 8: Example of experiment setup with a subject.

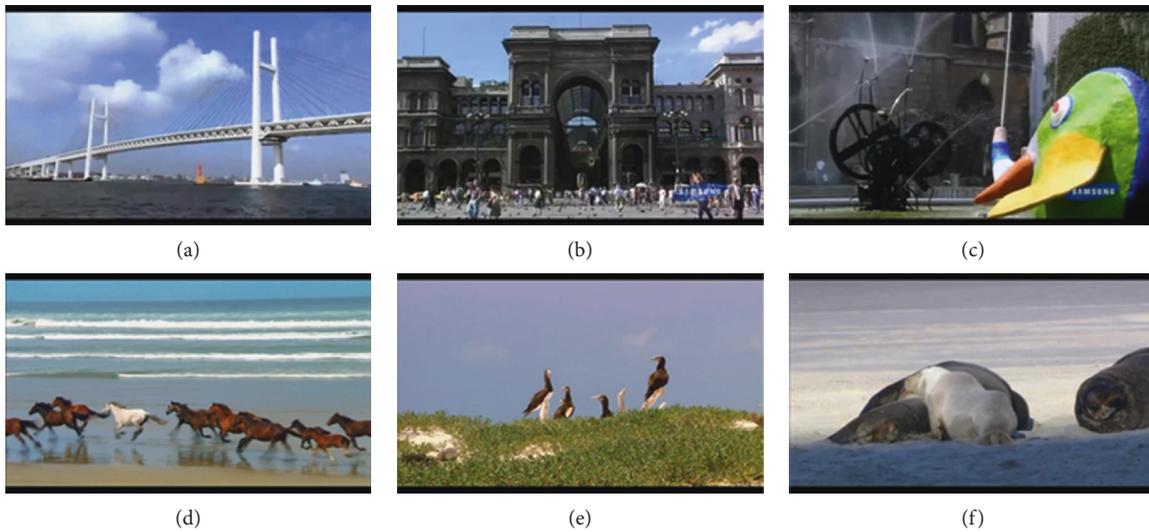


FIGURE 9: Test video clips.

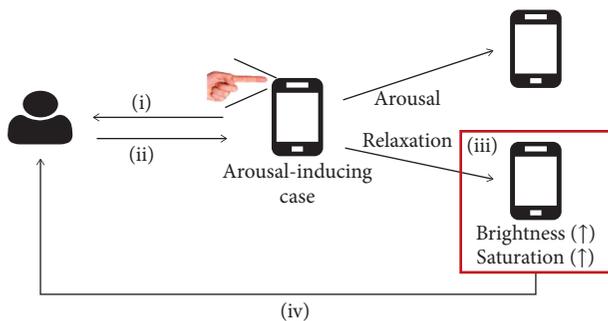


FIGURE 10: Procedure for arousal-inducing scenario.

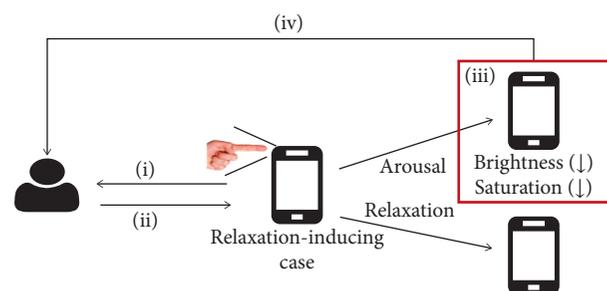


FIGURE 11: Procedure for relaxation-inducing scenario.

in this study is shown in Figure 12. In the upper part is a symbol of a heart, text showing the HR, a camera preview display and a measure button. When the user places their finger on the rear camera of the smartphone and presses the measure button, the video begins playing. While the user is watching the video, their HR is calculated and shown; internally, the emotion of the user is classified as a state of arousal or relaxation, and according to their emotion, the S and V values are changed and the video is displayed.

The real-time mobile emotional content player is mainly divided into two parts. The first part of the player is a PPG measurement and analysis module. PPG signals are extracted by acquiring R values with the highest-quality images from the smartphone camera. After the signals are cleaned through the signal processing application of the moving average filter, the position where the signal slope sign changes is detected as a peak using the sliding window technique; the time at which this change occurs is stored, and the PPI is calculated. The PPI is used in two instances: first, the HR is calculated through the

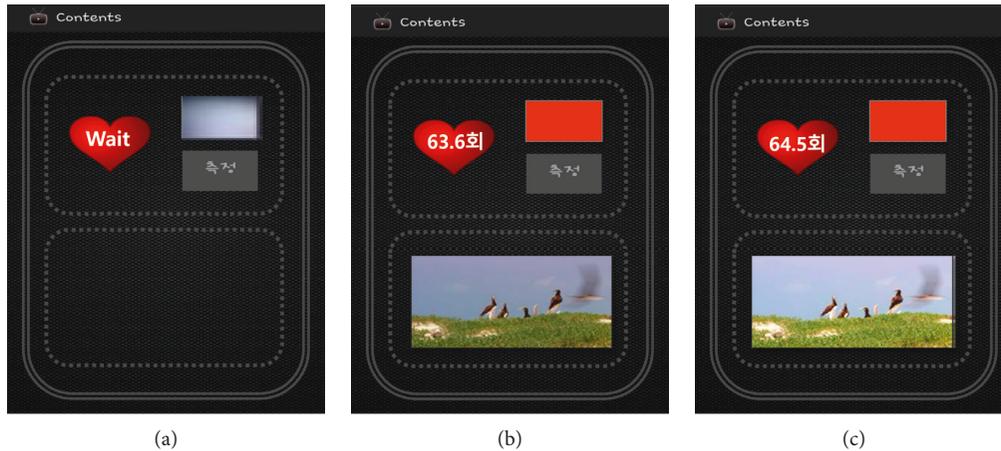


FIGURE 12: Example of the graphical user interface for the designed system.

PPI and displayed, and second, by comparing the current PPI with the average PPI, the emotion of the user is classified as a state of arousal or relaxation.

The second part of the player is a video processing module. Through the demuxing process, a video file is converted into a video bitstream, and through the decoding process, the video bitstream is converted into pixel data. After converting the YUV pixel data into RGB pixel data, the data are converted into HSV pixel data. The S and V values of the video content are changed in accordance with the defined rules based on the state of arousal or relaxation. Converting the data again into RGB pixel data after applying the changes in the S and V values, the video is played on the screen. Figure 13 shows an overall configuration diagram of the real-time mobile emotional content player.

4.2. Experimental Results. The matching sample t -test results of the PPI are as follows. As shown in Figure 14, the average PPI was 0.9158 when the user watched the regular content, whereas it was 0.8758 when they watched the content applying the arousal-inducing scenario, which is a large difference with a level of significance of 99% ($p < 0.000$, $t = 7.388$, $n = 917$). That is, the PPI showed a significantly lower trend when users were watching the content under the arousal-inducing scenario as compared to watching the regular content. It was therefore confirmed that the emotion of the user can be induced to a state of arousal if the content is viewed under the arousal-inducing scenario. Furthermore, as shown in Figure 14, the average PPI was 0.9158 when users were watching the regular content, whereas it was 0.9553 when they were watching the content under the relaxation-inducing scenario, which shows a large difference at a significance level of 99% ($p < 0.000$, $t = -5.856$, $n = 917$). That is, the PPI showed a significantly high trend when users were watching the content under the relaxation-inducing scenario as compared to watching the regular content. Thus, it was confirmed that the user's emotion can be induced to a relaxation state if the content is viewed under the relaxation-inducing scenario.

5. Discussion

In general, with regard to the relationship between the content and the user, when a user is watching a video content, the content affects the emotional change of the user; however, the user does not affect the content, which means the player is unidirectional. In other words, no players have an influence on the content according to the emotional change in the user. In this study, a real-time mobile emotional content player was developed to provide interactive content whereby the content and user affect each other in both directions.

A smartphone camera was used in the real-time mobile emotional content player to measure the PPG. The R values were received from the camera video frame, and the PPG signals were extracted from these R values. From the PPG signals, the PPI was calculated and the HR was displayed. Furthermore, based on the changes in the PPI, the emotion of the user was classified as a state of arousal or relaxation.

The S and V values influence the content according to the emotional change in the user. Because the S and V levels have to be changed by receiving content frames one by one, the demuxing and decoding processes were applied. To change the S and V values in an acquired frame, the color space was converted from YUV to RGB and then from RGB to HSV. In this format, the S and V values were changed, and finally, the color space was converted again from HSV to RGB to display the content on the screen. A surface holder is used to fetch the preview images of a smartphone camera frame by frame and convert the YUV format into the RGB format in real time. In real-time image processing, because the computation amount is large, it can become a burden on the application. Therefore, the load of the application was reduced using a method for displaying the view on a screen through a thread in the background using a surface view.

To change the S and V of a video sequence in real time, which is the goal of this study, the video file has to be received frame by frame through the demuxing and decoding processes, and the S and V have to be changed for every frame. Therefore, a video sequence is set to be played by continuously displaying the frames as bitmaps on the user's view.

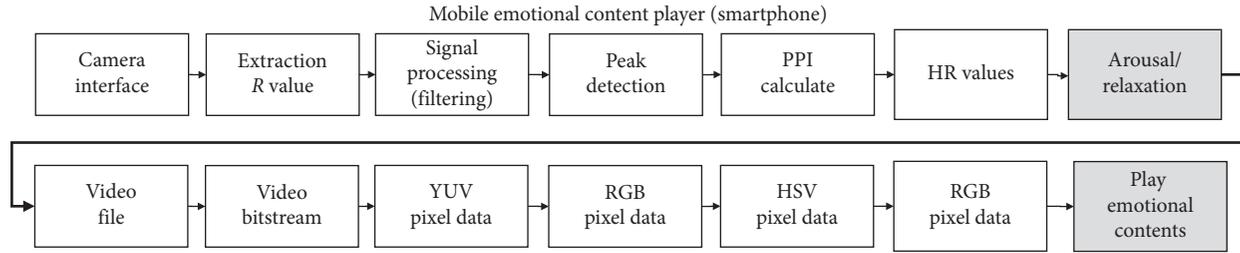


FIGURE 13: Overall process diagram of the mobile emotional content player.

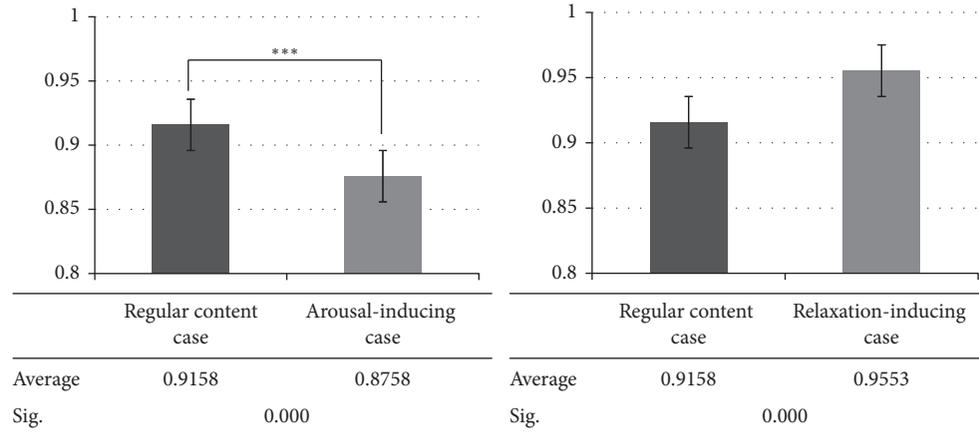


FIGURE 14: Sample t -test results of PPI between regular content case and arousal-inducing case and between regular content case and relaxation-inducing case ($p < 0.01$).

In the experiment, for the average PPIs of the regular content and the arousal- and relaxation-inducing scenarios, a matching sample t -test was applied. The analysis results show that the average PPI was significantly lower when users were watching the content according to the arousal-inducing scenario as compared to watching the regular content. Thus, the arousal-inducing scenario can be said to induce a state of arousal in the user. Furthermore, the average PPI was significantly higher when users were watching the content according to the relaxation-inducing scenario as compared to watching the regular content. It can therefore be said that the relaxation-inducing scenario induces a state of relaxation in the user.

6. Conclusions and Future Work

In conclusion, a real-time mobile emotional content player was developed, which measures a user's emotion through the PPG acquired in a smartphone system without the need for a specific sensor and reflects the emotion by changing the S and V values, thereby forming an ambience of the content naturally according to the measured emotion. It was verified experimentally that the real-time mobile emotional content player is a system in which a content affects the emotion of the user, and the user's emotion concurrently affects the content. The real-time mobile emotional content player allowed a user and a video content to interact with each other.

However, because the contents were viewed through the relatively small screen of a smartphone, there may be a

limited sense of immersion. During the experiment, inducing user's emotions with just two scenarios can be a kind of drawbacks as a limitation. It will be considered as a further study. In addition to video content, the system can also be applied to games, virtual reality, and smart devices. Furthermore, in a future study, it will be possible to utilize the proposed system to recognize other emotions in addition to arousal and relaxation by using various biological signals and applying them to a video content, thereby strengthening or weakening the pertinent emotion according to the reason why the content is provided.

Data Availability

The physiological data generated and analyzed during the current study are not publicly available in order to protect privacy. But, the data are available from the authors upon reasonable request with approval from the IRB of the institute.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1D1A1B03035606).

References

- [1] Y. J. Eum, J. S. Eom, and J. H. Sohn, "Difference of facial skin temperature responses between fear and joy," *Korean Journal of the Science of Emotion & Sensibility*, vol. 15, no. 1, pp. 1–8, 2012.
- [2] M. Kim, K. S. Park, D. Kim, and Y. Cho, "Emotional intelligent contents: expressing user's own emotion within contents," in *Entertainment Computing-ICEC 2011*, pp. 391–394, Springer, Berlin, Germany, 2011.
- [3] C. Huang, W. Gong, W. Fu, and D. Feng, "A research of speech emotion recognition based on deep belief network and SVM," *Mathematical Problems in Engineering*, vol. 2014, Article ID 749604, 7 pages, 2014.
- [4] P. Partila, M. Voznak, and J. Tovarek, "Pattern recognition methods and features selection for speech emotion recognition system," *Scientific World Journal*, vol. 2015, Article ID 573068, 7 pages, 2015.
- [5] H. Muthusamy, K. Polat, and S. Yaacob, "Improved emotion recognition using Gaussian mixture model and extreme learning machine in speech and glottal signals," *Mathematical Problems in Engineering*, vol. 2015, Article ID 394083, 13 pages, 2015.
- [6] S. Gong, Y. Dai, J. Ji, J. Wang, and H. Sun, "Emotion analysis of telephone complaints from customer based on affective computing," *Computational Intelligence and Neuroscience*, vol. 2015, Article ID 506905, 9 pages, 2015.
- [7] K. Yano and T. Suyama, "A novel fixed low-rank constrained EEG spatial filter estimation with application to movie-induced emotion recognition," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 6734720, 12 pages, 2016.
- [8] W. Wei and Q. Jia, "Weighted feature Gaussian kernel SVM for emotion recognition," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 7696035, 7 pages, 2016.
- [9] W. Wei, Q. Jia, Y. Feng, and G. Chen, "Emotion recognition based on weighted fusion strategy of multichannel physiological signals," *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 5296523, 9 pages, 2018.
- [10] D. K. Kim, Y. J. Cho, and K. S. Park, "Mediating individual affective experience through the emotional photo frame," *Journal of Ambient Intelligence and Smart Environments*, vol. 7, no. 1, pp. 99–113, 2015.
- [11] B. Ovčjak, M. Heričko, and G. Polančič, "How do emotions impact mobile services acceptance? A systematic literature review," *Mobile Information System*, vol. 2016, Article ID 8253036, 18 pages, 2016.
- [12] M. Y. Kim, K. S. Park, D. K. Kim, and Y. J. Cho, "A study of the effects on users' emotional changes by 2D and 3D emotional video contents," *Korean Society for Computer Game*, vol. 25, no. 3, pp. 165–174, 2012.
- [13] K. S. Park, Y. J. Cho, and D. K. Kim, "A framework for the creating, expressing and sharing of user's emotion," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 9, no. 1, pp. 425–442, 2012.
- [14] E. J. Kim, M. C. Whang, J. H. Kim et al., "A study on evaluation of human arousal level using PPG analysis," *Journal of the Ergonomics Society of Korea*, vol. 29, no. 6, pp. 925–932, 2010.
- [15] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," *Current Cardiology Reviews*, vol. 8, no. 1, pp. 14–25, 2012.
- [16] D. Snagita, P. Saurabh, and M. Madhuchanda, "Real time heart rate detection from PPG signal in noisy environment," in *Proceedings of 2016 International Conference on Intelligent Control Power and Instrumentation*, Kolkata, India, October 2016.
- [17] M. Kim, T. Kim, S. Jang, D. Ban, B. Min, and S. Kwon, "Noise-robust algorithm for PPG signal measurement," *Journal of Korean Institute of Communications and Information Sciences*, vol. 38, no. 12, pp. 1085–1094, 2013.
- [18] Y. Huang, J. Yang, P. Liao, and J. Pan, "Fusion of facial expressions and EEG for multimodal emotion recognition," *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 2107451, 8 pages, 2017.
- [19] N. Jatupaiboon, S. Pan-ngum, and P. Israsena, "Real-time EEG-based happiness detection system," *Scientific World Journal*, vol. 2013, Article ID 618649, 12 pages, 2013.
- [20] M. Kim, M. Kim, E. Oh, and S. P. Kim, "A review on the computational methods for emotional state estimation from the human EEG," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 573734, 13 pages, 2013.
- [21] H. Cai, J. Han, Y. Chen et al., "A pervasive approach to EEG-based depression detection," *Complexity*, vol. 2018, Article ID 5238028, 13 pages, 2018.
- [22] N. Zhuang, Y. Zeng, L. Tong et al., "Emotion recognition from EEG signals using multimodal information in EMD domain," *BioMed Research International*, vol. 2017, Article ID 8317357, 9 pages, 2017.
- [23] R. Munoz, R. Olivares, C. Taramasco et al., "Using black hole algorithms to improve EEG-based emotion recognition," *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 3050214, 21 pages, 2018.
- [24] D. K. Kim, Y. J. Cho, and K. S. Park, "Comparative analysis of affective and physiological responses to emotional movies," *Human-Centric Computing and Information Science*, vol. 8, no. 1, 2018.
- [25] S. H. Hwang, C. H. Park, and K. B. Shim, "Pattern classification of bio-information to percept human emotion," *Proceedings of Korean Institute of Intelligent System Conference*, vol. 15, no. 3, pp. 306–311, 2005.
- [26] H. Lee and D. K. Kim, "Development of real-time reactive emotion image contents player system to induce the user's emotion," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 18, no. 1, pp. 155–161, 2014.
- [27] J. Kim, D. Kim, S. Ann, S. Park, and M. Whang, "Brain function connectivity analysis for recognizing different relation of social emotion in virtual reality," in *Human-Computer Interaction. Towards Intelligent and Implicit Interaction*, pp. 441–447, Springer, Berlin, Germany, 2013.
- [28] D. K. Kim, S. Ahn, S. Park, and M. Whang, "Interactive emotional lighting system using physiological signals," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 4, pp. 765–771, 2013.
- [29] D. Kim, J. Kim, E. Lee, M. Whang, and Y. Cho, "Interactive emotional content communications system using portable wireless biofeedback device," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 4, pp. 1929–1936, 2011.
- [30] J. H. Kim, S. M. Ahn, J. U. Park, M. C. Whang, and D. K. Kim, "The XML based real-time emotion information sharing system using physiological signals," *Journal of Korean Institute of Information Technology*, vol. 10, no. 3, pp. 205–211, 2012.
- [31] C. J. Kim, M. C. Whang, J. H. Kim et al., "A study on arousal evaluation by cardiovascular response," *Journal of the Ergonomics Society of Korea*, vol. 11, pp. 444–447, 2009.
- [32] D.-G. Oh, B.-S. Min, S.-O. Kwon, and H.-J. Kim, "Effective PPG signal processing method for detecting emotional stimulus," *Journal of Korean Institute of Communications and Information Sciences*, vol. 37, no. 5, pp. 393–402, 2012.

- [33] H. K. Yang, H. S. Yang, and K. H. Min, "Enhancing emotion of visual contents based on emotion model for elevating absorption degree of games," *Korean Society for Computer Game*, vol. 25, no. 2, pp. 171–176, 2012.
- [34] J. W. Song, "Content-based image retrieval using HSV color and uniform local binary patterns," *Journal of Korean Institute of Information Technology*, vol. 12, no. 6, pp. 169–174, 2014.
- [35] J. K. Chang, S. T. Ryoo, and M. J. Jung, "Real-time image emotion retrieval for custom-made contents service," *Korean Society for Computer Game*, vol. 24, no. 2, pp. 169–176, 2011.
- [36] E. S. Lee and H. J. Suk, "The emotional response to lighting hue focusing on relaxation and attention," *Journal of Korean Society of Design Science*, vol. 25, no. 2, pp. 27–39, 2012.
- [37] H. J. Seok, "A study on a correlation between color chroma and appetitive motivation across media," *Journal of Korean Society of Design Science*, vol. 5, pp. 188–189, 2007.
- [38] M.-Y. Kim, D.-K. Kim, and Y.-J. Cho, "A study on the development of the interactive emotional contents player platform," *Journal of the Korean Institute of Information and Communication Engineering*, vol. 14, no. 7, pp. 1572–1580, 2010.
- [39] S. H. Jeong, "Development direction of emotional contents through analysis of successful cases from applying emotional technology," *Korean Journal of the Science of Emotion & Sensibility*, vol. 15, no. 1, pp. 121–132, 2012.
- [40] J. Kim, M. Whang, and J. Woo, "The research on adaptive process for emotion recognition by using time-dependent parameters of autonomic nervous response," in *Human-Computer Interaction. New Trends*, pp. 77–84, Springer, Berlin, Germany, 2009.
- [41] H. Lee, M. Kim, M. Whang, and D. K. Kim, "Evaluation of PPG signals regarding to video attributes of smart-phone camera," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 19, no. 4, pp. 917–924, 2015.
- [42] I.-B. Jung and K.-H. Kim, "A study of PPG wave and pulse measurement on radial artery using digital potentiometer and exponentially weighted moving average filter," *Transactions of The Korean Institute of Electrical Engineers*, vol. 63, no. 7, pp. 962–967, 2014.
- [43] H. M. Lee and D. J. Kim, "Development of a human sensibility evaluation and biofeedback system using PPG," *Transactions of the Korean Institute of Electrical Engineers*, vol. 57, no. 6, pp. 1087–1094, 2008.
- [44] J.-S. Kim and S.-H. Kim, "Design and implementation of a customized contents service system for relieving users' stress," *Journal of the Korea Contents Association*, vol. 11, no. 2, pp. 101–112, 2011.
- [45] J. Scheirer and R. Picard, "Affective objects," Technical Report No. 524, Perceptual Computing Section, MIT Media Laboratory, Cambridge, MA, USA, 2010.



Hindawi

Submit your manuscripts at
www.hindawi.com

