Correlation between XR content watching time and concentration/stress according to biometric data acquisition of XR multimodal devices users

Hyoung Lee
d Jong-Bae Lee, HyungKi Son, Beom-Ryeol Lee, Wook-Ho Son

*hyoung0708@etri.re.kr
1CG/Vision Research Section, Content Research Division, Hyper-Reality Metaverse Research Laboratory, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Republic of Korea
2VR/AR Content Research Section, Content Research Division, Hyper-Reality Metaverse Research Laboratory, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Republic of Korea

Keywords: XR multi-modal devices, Human factor, Statistical analysis of data.

ABSTRACT
In this study, we develop a novel multimodal device capable of collecting biometric data from viewers of XR content, including EEG signals, pupil size changes, eye blink signals, head movement data, and hand movement data. Using the various biometric data collected, correlations between concentration and stress factors and feature information are verified using statistical analysis methods. Finally, the correlation of concentration/stress according to the time of viewing the XR content is visualized to verify the meaningful relationship.

1 Introduction
The recent development of eXtended Reality (XR) based metaverse technology has enabled many users to live a better life in the metaverse world. Applications of metaverse technology in education, healthcare, military, business and leisure continue to expand. With this interest, various augmented reality devices are being developed, and many studies are being conducted to consider the human factors of users. In particular, there has been a lot of research into the relationship between viewing time of Virtual reality (VR)-based content and user concentration and stress.

1.1 Related work
There is an ongoing body of research that correlates the viewing of VR and Augmented Reality (AR)-based content with a variety of biometric data acquired from XR multimodal device, including head, hand and eye movements, pupil micro-size changes, eye blinks and EEG information.

When experiencing virtual content, brain waves in certain parts of the brain change significantly over time [1], which can have a variety of positive effects, such as improving memory and concentration and increasing motivation [2]. While increased concentration can be stressful for some people [3], the educational benefits are more positive, and researchers are exploring ways to improve concentration by changing audio based on gaze [4, 5] or only displaying important information when the user is focused [6]. In the AR environment, concentration has a positive effect on users, but due to the nature of displaying mixed virtual and real information in front of the eyes, concentration changes regardless of gaze [7], and changes in brainwaves can be associated with various biometric data. This paper aims to verify the correlation between users’ concentration and stress by relating content to vital signs in an XR-based environment.

1.2 Proposal concept
Our study, a new type of XR multimodal device has been developed. The device can acquire various XR-based biometric data from the wearer while watching XR content. The user wearing the developed XR multimodal device performs experimental sessions and surveys according to a set protocol while watching XR content. The raw data obtained from the XR multimodal device is subjected to a refinement process to extract feature information. Finally, each feature information is subjected to statistical analysis through one-way analysis of variance (ANOVA) and post-hoc Tukey HSD test to visualize the results of correlation between XR content viewing time and concentration/stress.

2 Method
2.1 Multi-modal XR device
For the acquisition of biometric data, this study integrates the HoloLens 2, the Pupil Labs eye tracker and the NeuroHarmony S20 as shown in Figure 1. The feature of multimodal data acquired by the unified XR multimodal device is that each biometric data is made into a common feature using a learning (deep learning) model instead of simple merging, which facilitates feature extraction and effectiveness analysis [6]. First, HoloLens 2 (MS) is a see-through holographic lens-based display.
It features head, hand and eye tracking via sensors and 6 DoF position tracking.

The Pupil Labs eye tracker has the following characteristics: eye tracking accuracy 0.60 degrees, wide angle lens 139 degrees X 83 degrees (1080p), sampling frequency 200Hz @ 192 x 192 px. The Pupil Labs eye tracker can be attached to the Holo Lens 2 as an add-on.

Finally, the NeuroHarmony S20 is a 2-channel portable EEG device with Bluetooth wireless connectivity, 50Hz bandwidth, 250Hz sampling rate and 24-bit digital resolution. In this study, the NeuroHarmony S20 was disassembled for integration into the XR multimodal device, and aluminum shielding was applied to minimize external noise [8] before integration into the Holo Lens 2 device.

2.2 Reference Contents

It is necessary to create XR content using XR multimodal device to test the relationship between the viewing time of XR content and the wearer’s concentration and stress. In this study, we used Unity3D (ver. 2019.03.01f) to create XR content and considered the following factors: a) the wearer can compare the EEG before and after interaction, b) the difficulty level of the content can be set to adjust the effects of concentration and stress, c) there is a concentration element in the creation of XR content, and d) the interaction element between the user of the XR multimodal device and the content is reflected.

The XR content created has a total of 3 levels of difficulty and 5 minutes (300 seconds) of interaction for each experimental session. Wearing the XR multimodal device, the subject sees the initial screen shown in Figure 2. Ten buttons are placed on a virtual board, and in the middle the subject sees a display board with cognitive features to help them press the correct buttons, while the left and right sides show the time spent and the correct score, respectively. Each level of content is performed in a randomized order with full explanation and practice by the experimenter. The user wearing the XR multimodal device is prompted to press the correct button by visual stimulation (interfacing) of colors or letters corresponding to the instructions presented on the display board, and the content provides auditory feedback through beeps for correct answers and beeps for incorrect answers.

The brainwave characteristics of the produced XR content are different before and after the start of the experimental session, and the factors of concentration and stress can be collected as data by setting the difficulty and viewing time of the XR content. In addition, the XR content provided at the same time as the subject’s XR content enables interaction between the content and the user through notification sounds and display panel content.

3 Experiment & Results

The overall experimental protocol for this study is shown in Figure 3. The trial guide will fully explain the purpose and method of this study to the wearer of the XR Multimodal device throughout the trial preparation and pre-training process, and will ensure that the subject is fully familiar with the wearing of the XR Multimodal device and basic pre-training. The protocol sequence is repeated three times: experimental session, questionnaire and rest. A room free of external noise such as electromagnetic waves was selected for the integrated EEG device, and the experiment was conducted in the same room within a fixed time period. The average brightness is an office environment (250 lux) and the average temperature is 26 degrees. The questionnaire used in this study is the SIM-TLX (NASA), Focus of Attention questionnaire, which can subjectively assess the level of concentration and stress [9].
Using an integrated XR multimodal device, two clinical subjects were tested for 3 minutes each in the closed-eye, open-eye and interaction conditions. To ensure the reliability of the data obtained, the same procedure was repeated at the same time and place for five days, and the data were averaged. The analysis used the time-frequency domain analysis method, and after converting the EEG signal into a spectrogram, the mean and variance of the signal intensity by frequency band were determined. The results of the time-frequency analysis method showed that the intensity and variance of the signal increased in the order of closed-eyes state, open-eyes state and interaction state with unified XR device in the spectrogram of the gamma band (25Hz~45Hz) of the measured EEG, as shown in Figure 4.

Fig. 4 EEG measurements with closed-eye, open-eye, and XR interaction content

In addition, the EEG data could be measured and extracted without distortion even with the EEG meter disassembled and integrated. The X, Y and Z axes of each graph represent the measurement time, band strength and frequency band of the EEG, respectively.

3.2 Correlation study experiments

In this study, we conducted a study of XR content with 10 subjects wearing an XR multimodal device. The 10 participants are aged between 29 and 39 years (3 women) and a questionnaire were used to collect user information such as physical condition, age group, experience with XR and choice of main handle.

Participants sit in a chair in front of a desk and wear an integrated XR multimodal device. Prior to the start of the experiment, an explanation of the experiment and a practice session allow participants time to acclimatize to the augmented reality-based content environment and to familiarize themselves with the interaction of the XR content by pressing buttons provided in the virtual space. At the start of the experiment, the content is presented for 300 seconds at a randomized level of difficulty according to the protocol sequence. For each experimental session, users complete a questionnaire for 180 seconds after the experiment and take a 60-second break before proceeding to the next experimental session.

In this study, each biometric data was saved as a .csv file to create a dataset, and each feature data was processed for statistical analysis. The data acquired in the XR-based environment were subjected to a one-way ANOVA and Tukey HSD test, and the feature information that showed statistical significance with concentration were the β and γ bands of the EEG, eye blinks, small changes in pupil size, and head and hand movements. In addition, in experiment with XR content, the features associated with stress are the δ and γ bands of the EEG, head and hand movements, and pupil dilation. The feature information correlated with concentration or stress was determined to be correlated for features that met the p-value of 0.005 (99.5% confidence) or less using the statistical analysis method of one-way ANOVA. Moreover, the results of the SIM-TLX and the Focus of Attention Questionnaire showed that the level of concentration and stress increased as the difficulty of XR content and viewing time increased. Finally, the correlation between XR content viewing time and concentration/stress along with significant feature information can be visualized as shown in Figure 7.
Figure 6 shows 10 items (EEG_delta, EEG_theta, EEG_beta, EEG_gamma, (2) pupil micro-size change, eye blink, and head-hand-eye movement) that are significantly correlated with XR content viewing time (elapsed_time) and concentration/stress in an 11x11 Metrix, and the correlation between each item is expressed by the intensity of the color. In this study, the correlation between head and eye movements, β and γ bands of EEG, pupil micro size changes, θ bands of EEG, eye blinks and hand movements were identified as the feature information that correlated with the viewing time of XR content. This means that head and eye movements are reduced as XR content is viewed longer. This can be understood as an increase in concentration and stress as the time spent watching XR content increases. In addition, the decrease in the values of the β and γ bands of the brainwave according to the time spent watching XR content is an objective proof that human conscious activities, especially β waves, which appear when we are anxious, nervous and concentrating, and γ waves, which are more emotionally edgy (stressful) or related to advanced cognitive information processing, affect concentration and stress. Similarly, the rate of change of pupil size and blink frequency decreases with the time spent viewing XR content, which can be interpreted as the user’s concentration in interacting with or viewing the content increasing over time.

4 Conclusion

In this paper we have verified and analyzed the correlation between the time spent watching XR content and the user’s concentration/stress in an XR environment. For this purpose, we developed an integrated XR multimodal device that can acquire EEG, pupil, head, hand and eye movement information, and confirmed the basic characteristics of each biometric data through experiments. We designed and produced XR content with cognitive load, interaction elements, and difficulty adjustment, and conducted a pilot test with 10 people according to the experimental protocol. The collected biometric data was analyzed using one-way ANOVA analysis and Tukey HSD test.

According to the results of the experiment using XR-based content, the feature information related to concentration is the β and γ bands of EEG, eye blinks, small changes in pupil size, and head and hand movements. In addition, the stress related features are the δ and θ bands of the EEG, head and hand movements, and small changes in pupil size. Using a visualized correlation matrix correlating the feature information related to concentration and stress with the time spent viewing XR content, the correlation was confirmed in the following order: head and eye movements, β and γ bands of the EEG, micro changes in pupil size, θ bands of the EEG, blinking and hand movements.

In future research, it is believed that it will be possible to study the configuration of the interaction-oriented XR device environment by using more biometric data to show the correlation between stress and concentration according to XR content viewing time.

Acknowledgement

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2022-0-00137, XR User Interaction Evaluation and Employing Its Technology)

References