Applied Ergonomics 62 (2017) 158–167

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

# A wellness platform for stereoscopic 3D video systems using EEG-based visual discomfort evaluation technology

Min-Koo Kang [a], Hohyun Cho [b], Han-Mu Park [b], Sung Chan Jun [b], Kuk-Jin Yoon [b, *]

[a] Imaging Media Research Center, Korea Institute of Science and Technology, Seoul, 136-791, South Korea
[b] School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Seoul, 136-791, South Korea

ARTICLE INFO

*Article history:*
Received 20 January 2016
Received in revised form
7 December 2016
Accepted 28 February 2017
Available online 6 April 2017

*Keywords:*
Visual discomfort
Stereoscopic 3D
Wellness platform

ABSTRACT

Recent advances in three-dimensional (3D) video technology have extended the range of our experience while providing various 3D applications to our everyday life. Nevertheless, the so-called visual discomfort (VD) problem inevitably degrades the quality of experience in stereoscopic 3D (S3D) displays. Meanwhile, electroencephalography (EEG) has been regarded as one of the most promising brain imaging modalities in the field of cognitive neuroscience. In an effort to facilitate comfort with S3D displays, we propose a new wellness platform using EEG. We first reveal features in EEG signals that are applicable to practical S3D video systems as an index for VD perception. We then develop a framework that can automatically determine severe perception of VD based on the EEG features during S3D video viewing by capitalizing on machine-learning-based braincomputer interface technology. The proposed platform can cooperate with advanced S3D video systems whose stereo baseline is adjustable. Thus, the optimal S3D content can be reconstructed according to a viewer's sensation of VD. Applications of the proposed platform to various S3D industries are suggested, and further technical challenges are discussed for follow-up research.



## 1. Introduction

Image-safety issues such as photosensitive epilepsy have been studied for decades (ISO IWA3, 2005; Ujike, 2009). In particular, visual discomfort (VD) has been cited as a bottleneck for stereoscopic 3D (S3D) video technology (Lambooij et al., 2009). Although many technology-induced VD factors can be eliminated with advancements in hardware technology, factors intrinsic to VD, such as vergence-accommodation conflict (VAC), persist as long as stereoscopy is used as 3D technology (Kim et al., 2014; Sweeney et al., 2014; Vienne et al., 2014). Therefore, S3D content needs to be reconstructed in consideration of the viewer's VD perception for their safety (Hoffman et al., 2008; Onural, 2007).

To address this issue, the Moving Picture Experts Group (MPEG) has developed an advanced stereoscopic video system (ASVS) (ISO/IEC JTC1/SC29/WG11, 2009), as shown in Fig. 1. With this system, the stereo baseline can be adjusted by selecting from a range of virtual views. As a result, S3D content can be reconstructed adaptively based on the adjusted stereo baseline. A video-plus-depth format is used to synthesize multiple virtual views at intermediate viewpoints between the original left and right views (Arican et al., 2009; Ndjiki-Nya et al., 2011). The original left view and one of the virtual views are then projected to the viewer as a final stereoscopic video pair. However, the challenge is to select the best virtual view among all candidates, and one way of doing so is to refer to the VD experienced by viewers.

In light of the rising demand for VD evaluation technology, a variety of both subjective and objective methods have been researched. Subjective VD evaluations involve assessing VD directly by asking subjects about VD symptoms (ITU, 2002; Kuze and Ukai, 2008). This approach has an advantage of evaluating the VD actually experienced by viewers so that it is often used to collect ground-truth data. However, it requires high costs for experiments, and the manual process restricts its applications. For more practical usage, objective VD evaluations present the alternative to the subjective approach. In 3D video-signal processing, a quantitative degree of VD can be computed based on pre-defined models of VD factors (e.g., disparity differences, object motion, etc.) (Choi et al., 2010; Ha and Kim, 2011; Jung et al., 2012). Unlike subjective

* Corresponding author.
*E-mail addresses:* kangmkgb@gmail.com (M.-K. Kang), AugustCho@gist.ac.kr (H. Cho), hanmu@gist.ac.kr (H.-M. Park), scjun@gist.ac.kr (S.C. Jun), kjyoon@gist.ac.kr (K.-J. Yoon).

http://dx.doi.org/10.1016/j.apergo.2017.02.022
0003-6870/

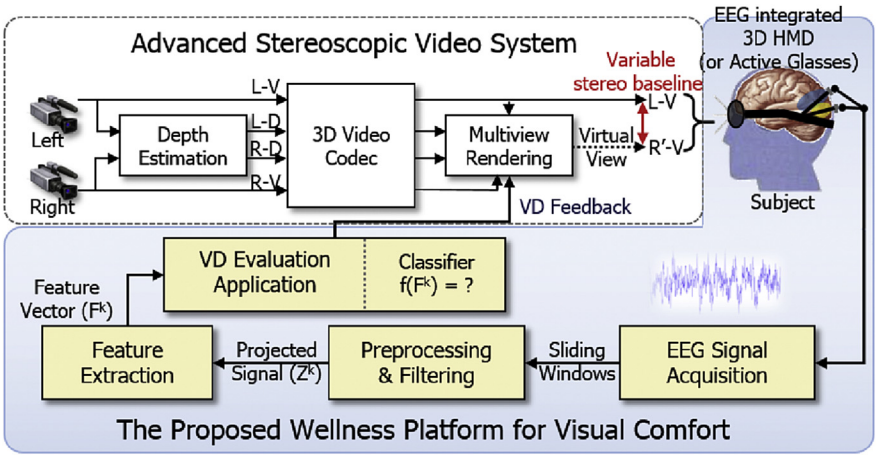


**Fig. 1.** Advanced stereoscopic video system (ASVS) with the proposed wellness platform to facilitate comfort with S3D video. The ASVS mitigates VD by changing the stereo baseline. The proposed wellness platform provides VD feedback via an analysis of real-time EEG responses as a criterion for the best stereo baseline.

methods, this approach is inexpensive, quick, and immune to the influence of external stimuli. However, it is difficult to measure personal factors (e.g., stereo acuity, user experience, etc.) and environmental factors (e.g., viewer position, lighting conditions, etc.) despite their significant effect on VD perception. Another objective approach exploits research in cognitive neuroscience to interpret abnormal bio-signals from viewers as an index for VD perception (Schmidt et al., 1979; Stern et al., 1979; Ukai and Kato, 2002; Hagura and Nkajima, 2006; Hsieh et al., 2010). Because abnormal bio-signals are directly elicited by viewers, this approach is able to evaluate the VD actually perceived by viewers, regardless of the hidden factors that affect VD perception.

Especially, thanks to the recent advances of the brain imaging technology, many investigations using features of brain signals have noticeably succeeded in assessing 3D visual discomfort or 3D visual fatigue. Among them, functional magnetic resonance imaging (fMRI) has begun to even localize the human brain regions associated with the sensation of visual discomfort (Backus et al., 2001; Georgieva et al., 2009; Durand et al., 2009; Tsao et al., 2003), and in-depth investigations have been accomplished owing to its high performance (Jung et al., 2015). However, when developing an application for S3D video systems, we should consider many practical issues such as price, safety, interruption of natural viewing. For this reason, electroencephalography (EEG) is regarded one of the most promising candidates for developing an application because of its advantages for practical use. First of all, EEG is cost effective. Although it relies on measurement performed on the scalp, EEG has elaborate temporal resolution so that real-time monitoring is possible during S3D video viewing. Also, EEG is one of the most non-invasive brain-imaging modalities so that it is proper to applications for daily life. In addition, EEG does not require a special space that blocks interferences, and its volume is getting smaller. Nowadays, even portable EEG systems (EMOTIV EPOC and InteraXon Muse headsets) are used for ergonomic research (Laghari et al., 2013).

In this paper, we suggest a concept of wellness-platform that could lead to comfort and safe S3D viewing in coordination with the ASVS. To this end, we exploited the EEG modality and its advantages, while overcoming the technical challenges of existing EEG-based methods (see details in Section 2.2). Fig. 1 depicts the overall concept of the proposed wellness platform. It first captures a viewer's brain waves via EEG, and extracts reliable VD feedback from the EEG data. Then, the VD feedback is referred to reconstruct the S3D content adaptively according to the viewer's VD perception. Notice that the main focus of this paper is to put the well-known EEG-based VD evaluation technology to practical use, instead of finding advances with respect to neuroscience.

The remainder of this paper is organized as follows. In Section 2, we review recent EEG-based VD evaluation methods and the technical challenges in their practical use. Our proposed viewing experimental framework is presented in Section 3, and discriminative EEG components are investigated in Section 4. We then present a sophisticated framework that can automatically determine VD perception from the discriminative EEG components during S3D viewing by capitalizing on machine-learning-based brain-computer interface technology in Section 5. Eventually, we discuss some guidance for future studies in Section 6, and present our conclusions in Section 7.

## 2. Background

### 2.1. EEG-based VD evaluation technology

In addition to the advantages of the EEG modality with respect to its practical use, EEG is well known for a significant and reliable bio-signal reflecting mental fatigue (Murata et al., 2005; Liu et al., 2010; Kar et al., 2010), and two types of EEG components are mainly investigated: temporal components and oscillatory components. Table 1 summarizes typical characteristics of specific EEG components widely used in cognitive neuroscience. In general, cognitive loading such as processing syntactic errors induces mental fatigue (Vos et al., 2001), and people have difficulty for processing visual stimuli or making decisions when they suffer from mental fatigue (Murata et al., 2005; Liu et al., 2010). As a result, if significant mental fatigue is perceived, amplitude of the corresponding temporal EEG components is decreased while their latency is delayed. Thus, in cognitive neuroscience, mental fatigue can be measured by comparing the behavior of temporal components before and after an event happening. Feature oscillatory components reflect various states of a subject. For example, the proportion of low frequency bands such as delta ($\delta$), theta ($\theta$), and alpha ($\alpha$) increases while the proportion of higher frequency bands such as beta ($\beta$) decreases as alertness declines (Trejo et al., 2005). In this point of view, EEG spectral changes are often investigated for mental fatigue evaluation. Most typically the power of alpha band in a fatigue condition is compared with that of alpha band in the relaxation condition, and fatigue occurrence can be judged when the power of alpha band is significantly attenuated (Reisman, 1997).

**Table 1**
Characteristics of major temporal/oscillatory electroencephalography (EEG) components.

| Category | Name | Detection range | Reflection |
|---|---|---|---|
| Temporal Component | P100 | A positive peak around 100 ms | Visual stimulus processing |
| | P300 | A positive peak around 300 ms | Decision making (stimulus evaluation or categorization) |
| | P600 | A positive peak around 600 ms | Syntactic phenomena processing |
| | Delta | 0.1–3 Hz | Deep sleep, unconsciousness of mind |
| | Theta | 4 - 7 Hz | Light sleep, drowsiness |
| Oscillatory Component | Alpha | 8 - 14 Hz | Wakeful relaxation with closed eyes |
| | Beta | 15 - 30 Hz | Active, busy, or anxious thinking and active concentration |

Recently, ratios of feature frequency bands (e.g, $(\alpha+\theta)/\beta, \alpha/\beta$, etc.) and their power changes are also examined in a more sophisticated way for better accuracy (Jap et al., 2009; Li et al., 2012).

On the basis of the above fundamentals, active attempt to evaluate perceptual VD with EEG have been made in S3D video processing field, and they have indeed succeeded in showing discriminative brain responses to the VD perception capitalizing on various EEG components. For instance, Emoto et al. (2005) provided evidence that the VAC problem in S3D displays results in VD. They tested the visually evoked cortical potential (VECP)[1] before and after an hour-long period of S3D video viewing. They found that the P100 latency[2] is also a reliable index for VD in response to various vergence loads. Li et al. (2008) investigated the effects of binocular parallax (2D or 3D) and the presentation time (3 min or 30 min) on VD perception by measuring the event-related potential (ERP).[3] Their results confirmed that the behavior of the P700 component is much more discriminative than the conventional P300 component.[4] In addition, the distinctiveness of the P700 component becomes greater in the 3D condition than the 2D condition and in the 30-min condition than the 3-min condition, respectively. Similarly, Mun et al. (2012) found that significantly reduced P600 amplitudes and delayed P600 latencies[5] after watching an hour of S3D video via a ERP test. In addition to the methods based on the above temporal components, it was confirmed that VD can also be detected by comparing the power of feature frequency bands before and after watching 2D and S3D content (Li et al., 2008; Frey et al., 2014; Cho et al., 2012; Chen et al., 2013). For example, Li et al. (2008) investigated the oscillatory EEG components during 30-min 2D/S3D viewing tests, and analyzed the high-frequency spectrum, specifically beta bands (12–30 hz). Usually, the beta band reflects the state of brain stress induced by anxiety or active concentration. Similarly, they found that beta frequency bands increased after S3D video viewing on most EEG channels. Chen et al. (2013) found that the alpha and beta bands significantly decreased, and that the delta band significantly increased as subjects viewed S3D content, whereas the theta rhythm remained unchanged.

### 2.2. Challenges to an ASVS as a wellness platform

In addition to the potential of EEG-based VD evaluation technology, the following technical challenges can enhance its functionality as an objective VD evaluation tool. Current methods mostly rely on comparing the subject's state before and after watching a long S3D test sequence. Thus, real-time VD evaluations as viewers watch S3D video are needed. Furthermore, discriminative EEG components are mostly derived from EEG responses to two different conditions: visually comfortable (2D) stimuli, and visually uncomfortable (3D) stimuli. Consequently, the ability to assess VD quantitatively is needed. Developing a systematic platform that can automatically reproduce discriminative EEG components and interpret them in real time should be also taken into account.

Therefore, our study involves applying EEG-based VD evaluation technology to a wellness platform for ASVSs. There are three specific goals to our study: **(G1)** to elicit feature EEG components after only a few seconds of viewing, without using external stimuli; **(G2)** to show that specific EEG components reflect quantitative changes in VD, rather than mere differences in EEG responses caused by any conditional changes; and **(G3)** to develop a platform that automatically extracts EEG components to generate viewer feedback that reflects VD perception.

## 3. Viewing experiment

A new viewing experiment is proposed to achieve the above goals. For G1, a single-trial test was conducted repeatedly. Specifically, stimulus viewing and EEG recording were simultaneously conducted for a period of 6 s during one stimulus projection. For G2, test stimuli with differences only in terms of binocular disparity were investigated in a strictly controlled experiment. For G3, a subjective VD evaluation using a simple questionnaire followed each stimulus projection. This helped not only to analyze EEG responses, but also to construct the ground-truth data used for the machine learning process of our wellness platform.

### 3.1. Apparatus

Experiments were executed inside a darkroom equipped with a passive 47″ 3DTV (47LW6500, LG Electronics Inc.), and all recording and command transmissions were executed by observers from outside the darkroom while monitoring the subject through an IR camera, as shown in Fig. 2. The darkroom helped to eliminate unexpected changes to the viewing environment, and the IR camera enabled the observers to react promptly to unforeseen behavior such as epileptic seizures. The viewing distance was fixed at 2.5 m. For EEG signal acquisition, a multi-channel system (WEEG-32, Laxtha Inc., Daejeon, Korea) was used. EEG electrodes were attached to subjects' scalps according to the international 10–20 system, where 18 channel locations (F3–6, Fz, C3–6, Cz, P3–6, Pz, O1–2, and Oz) were used at a sampling rate of 512 Hz.

[1] VECP measures the time required for a visual stimulus to travel from the eye to the brain, and it is a useful index for detecting optic nerve problems.

[2] P100 is a positive peak observed at 100 ms after the onset of the stimulus. Its delay reflects fatigue of the interrelated extraocular and intraocular muscles and the central nerve of the brain.

[3] ERP measures the brain's response to perceptual experience, and it is used to observe cognitive processing. Subjects are asked to react to target stimuli hidden as infrequent occurrences among a series of frequent stimuli. In Li et al. (2008), a close-distance random dot stereogram (RDS) surface and a long-distance RDS surface were used as the infrequent stimulus (20%) and the frequent stimulus (80%), respectively.

[4] P300 is a positive peak observed at 300 ms after the onset of the stimulus. Its delay reflects cognitive impairment in decision making when evaluating or categorizing stimuli.

[5] P600 is a positive peak observed at 600 ms after the onset of the stimulus. Its delay reflects processing syntactic errors.

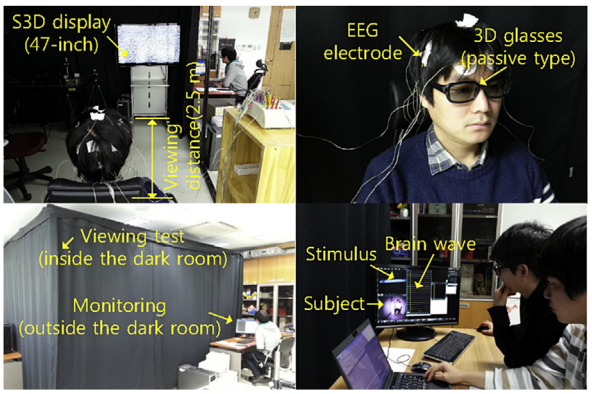


**Fig. 2.** System setup. A subject watches the test stimuli inside a darkroom equipped with a passive 47″ 3DTV and an IR camera. Observers control the proposed viewing test and watch the subjects from outside the darkroom through the IR camera.

### 3.2. Subjects

Nine healthy subjects all male and right-handed with an average age of 26.67 ± 3.08 years (indexed as S1, S2, …, S9) took part in the proposed viewing test. We provided a detailed summary of the experimental procedure to each subject in advance. The experiment was approved by the Institutional Review Board at the Gwangju Institute of Science and Technology.

### 3.3. Test stimuli and procedure

In order to evaluate VD induced by the VAC problem, we generated five images of a random-dot stereogram (RDS) with the same shape at varying depths: a circle with radius of 3.89°, as shown in Fig. 3. The RDS merely provides the impression of depth via disparity information, thus removing other factors to S3D content that can affect VD perception. The test disparity levels ranged from 2D to an outer binocular fusion limit: the minimum disparity (RDS0), the maximum disparity (RDS4), and the intermediate disparities (RDS1–RDS3) around the thresholds of Percival's zone of comfort (Lambooij et al., 2009; Wopking, 1995) and the zone of clear single binocular vision (Ukai and Kato, 2002; Sheard, 1934), as shown in Fig. 4.

The test procedure is depicted in Fig. 5. Each trial consisted of a stimulus period (SP) and a rest period (RP). During the SP, the subject observed an RDS stimulus that was randomly selected from among the five RDSs shown in Fig. 4, and the subject's EEG response was simultaneously recorded. During the RP, the subject rated the stimulus in terms of the discomfort experienced, as shown in Fig. 5 (b). The questionnaire was a simplified version of Hoffman et al. (2008); Kim et al. (2014). Moreover, an inter-trial period (ITP) was inserted between the RP and the SP. The duration of the ITP varied from 0 s to 1 s to prevent subjects from guessing the next stimulus.

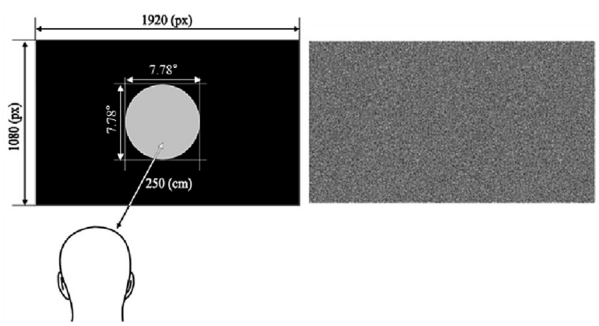


**Fig. 3.** Test RDS stimulus: (left) a conceptual image constructed in a viewer's brain, (right) a real RDS image in a S3D display.

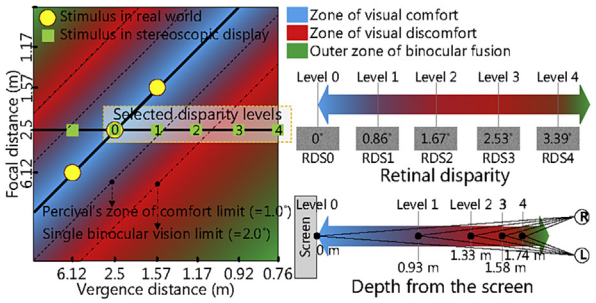


**Fig. 4.** Selection of test stimuli and human binocular vision characteristics according to changes in binocular disparity.

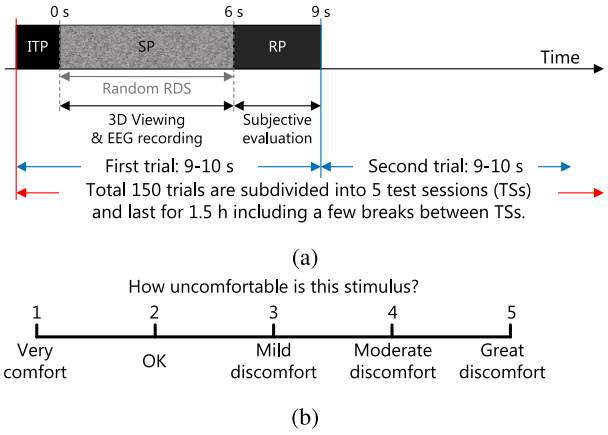


**Fig. 5.** (a) Proposed test procedure and the stimulus presentation timing. (b) A questionnaire with a visual discomfort chart for subjective evaluations of VD.

Notice that VD is a cumulative quantity depending on the presentation time (Li et al., 2008), and that a sequence effect occurs whereby a projected stimulus is affected by the previously viewed stimulus. Therefore, we divided the 150 trials (30 trials per disparity degree) into five test sessions (TSs) to give the subjects a short break (2 min) between TSs. Further, we randomly changed the presentation sequence of the test stimuli with the same frequency for each stimulus during a TS, and repeated the TSs with the same test stimuli in a different sequence to negate the sequence effect. Each TS lasted approximately 5 min, and the total experiment lasted approximately 1.5 h including periods for the introduction of the experiment, EEG setup, and all breaks.

## 4. EEG response analysis

The results of the subjective evaluation were analyzed first as the reference. Spectral, spatial, and temporal EEG-features were then analyzed with respect to their feasibility in S3D video systems.

### 4.1. Subjective VD evaluation

In Fig. 6, the top-left shows the overall average VD score at each RDS level ranging from 1.167 (comfortable) to 4.774 (uncomfortable). RDS1, within the comfortable zone, was assigned a value of 2.563. RDS2, within the binocular fusion area, was assigned a value of 3.486. RDS3, beyond the fusion limit, was assigned a value of 4.358. These results show that our experiment is consistent with

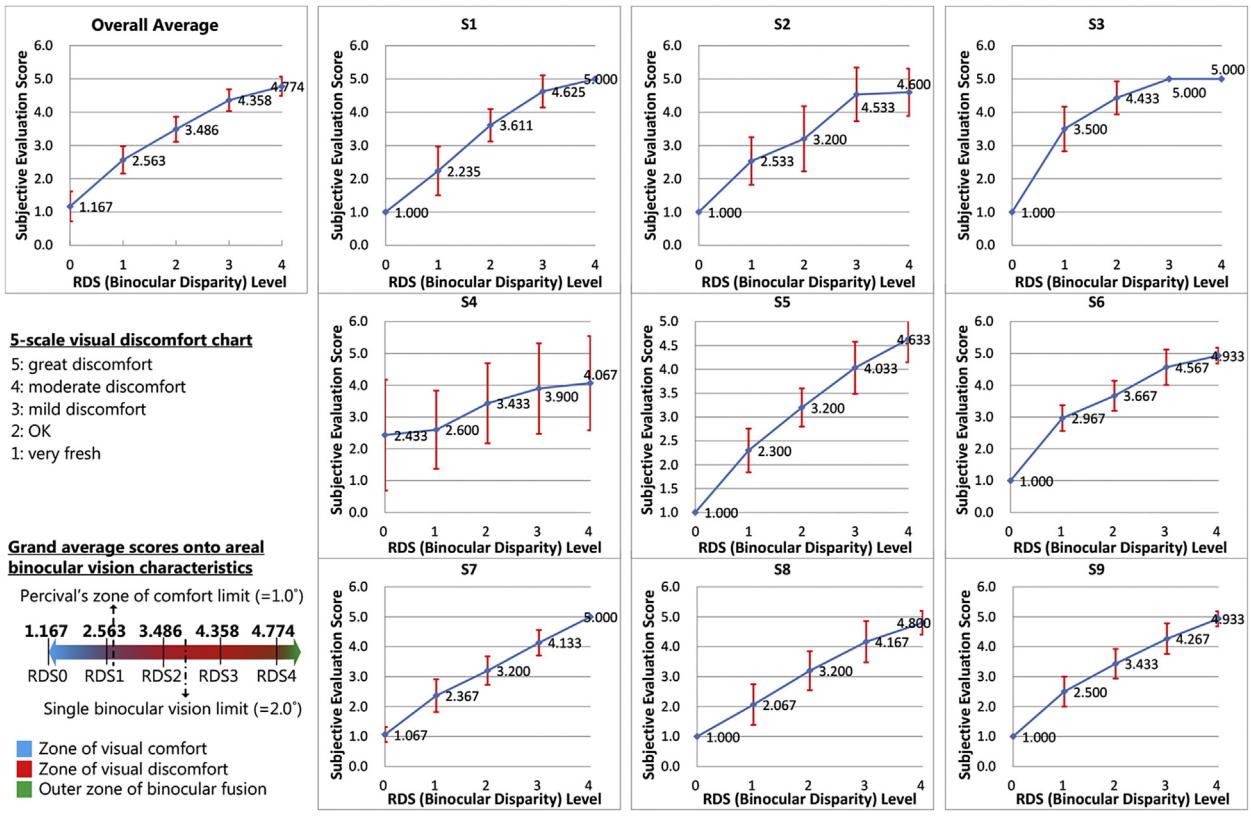


**Fig. 6.** Results of the subjective VD evaluation. Each subject (S1-S9) evaluated 150 trials of the five RDSs (30 trials for each RDS level). Note that the overall average VD increased logarithmically with increases to the stimulus level (in terms of retinal disparity), and the subject-specific results vary slightly with respect to the dynamic range and the sensitivity against the test stimuli.

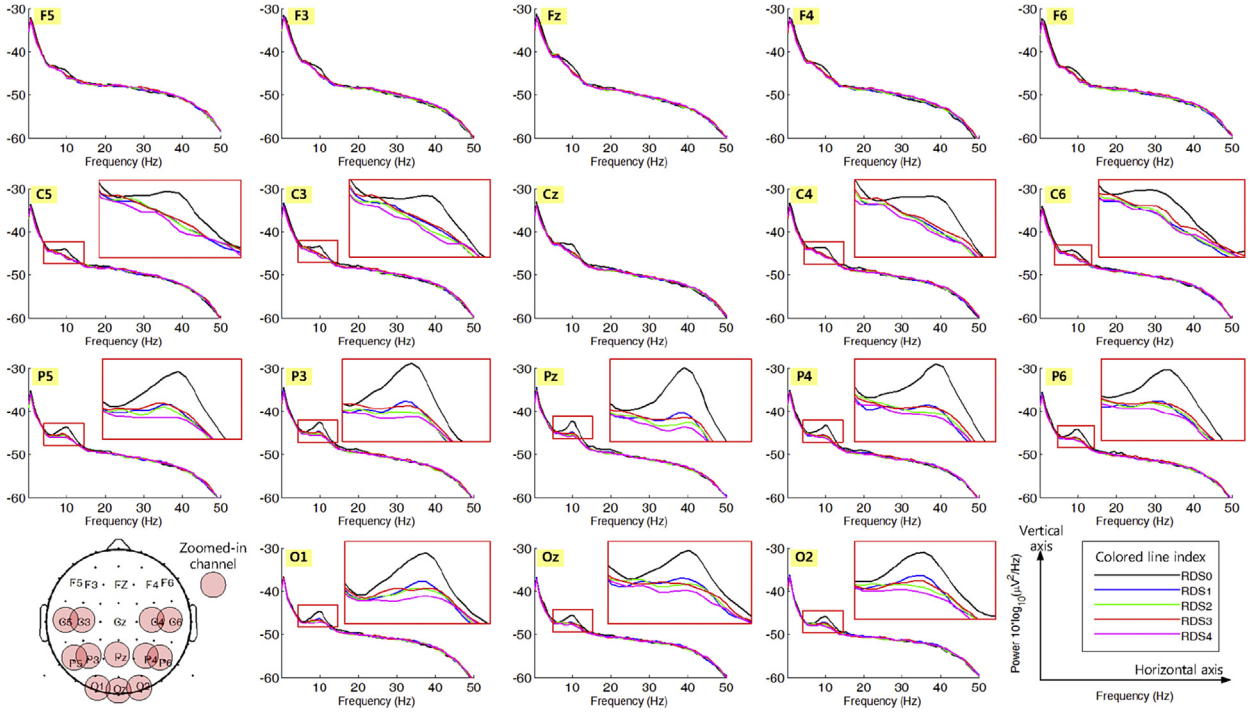


**Fig. 7.** Overall average log spectra of EEG data for all subjects (S1–S9). The spectra are distinguished by color according to the RDS levels. There is significant spectral attenuation of 3D stimuli (RDS1–4) against the 2D stimulus (RDS0) in the alpha and beta frequency bands (8–30 Hz), and around the visual association cortex (Pz), the primary visual cortex (Oz), and the sensorimotor cortex of the eyes (C3 and C4).

the literature (Ukai and Kato, 2002; Lambooij et al., 2009; Sheard, 1934; Wopking, 1995), and the excessive disparity in S3D content clearly results in significant VD with S3D displays. In addition, the subject-specific graphs prove that VD can vary depending on individual factors such as stereo acuity, despite the same S3D content and viewing environment.

### 4.2. Oscillatory EEG response

Fig. 7 shows the overall average log spectra of the EEG data across all subjects. It shows that significant spectral attenuation occurred against the 3D stimuli (RDS1–4), compared to the 2D stimulus (RDS0) in the alpha (8–14 Hz) and beta (15–30 Hz) EEG bands. The attenuation was mainly elicited around the visual association cortex (Pz, $p = 0.015$ in alpha and $p = 0.129$ in beta), the primary visual cortex (Oz, $p = 0.024$ in alpha and $p = 0.307$ in beta), and the sensorimotor cortex of the eyes (C3 and C4, $p = 0.033$ in alpha and $p = 0.467$ in beta). A Student's t-test ($P < 0.05$) was performed to clarify the distinctiveness of the alpha and beta EEG bands. The P-values confirm that the alpha band is significantly distinct compared to the other bands. In addition, these results show that the degree of spectral attenuation increases roughly in proportion to the RDS level.

### 4.3. Spatial EEG response

Fig. 8 shows the overall average topographies of $r^2$ values across all subjects. The $r^2$ is a statistical measurement of the extent to which the means of the two distributions differ in relation to their variances (Blankertz et al., 2008; Schalk et al., 2004). We used this measurement to show discrimination in the spatial domain (the EEG electrode channel), where the band-passed spectrum (8–30 Hz) differences between two stimulus distributions (RDS0 vs. one of RDS1-4) were used to condense features from the above analysis in the frequency domain. These topographies confirm that the Pz channel (the visual association cortex) is the most discriminative among all EEG channels. The visual association cortex is linked with the unnatural processing of 3D information, as discussed in Hoffman et al. (2008); Lambooij et al. (2009); Wopking (1995).

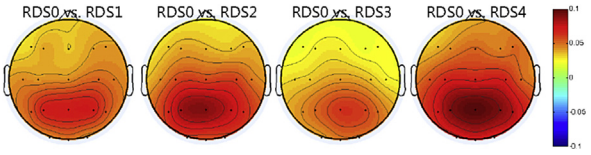


**Fig. 8.** Overall average topographies of $r^2$ values, showing how strongly the means of two stimulus distributions (RDS0 vs. one of RDS1-4) differ in relation to their variances.

### 4.4. Temporal EEG response

Fig. 9 shows the overall average correlations between envelopes of the band-passed spectrum and stimulus levels (RDS1–4). The correlations were computed using the algorithm proposed by Blankertz et al. (2008), where the greatest value in channel-by-time matrices reveals the channel and the time that are the most correlated with each RDS label. Thus, these results further clarify the temporal behavior of the target EEG components (8–30 Hz at Pz channel). Moreover, they confirm that the target EEG components maintain sufficient correlations during the SS (6 s), except during the first second after the onset of the stimulus. This indicates that the target EEG components are a sustainable response to VD. Thus, real-time VD evaluations can be made by capturing EEG responses at any point during continuous S3D video viewing.

## 5. Detecting VD

In this section, a framework for determining whether a subject is currently feeling severe VD is described, with the ultimate aim of achieving **G3**. The proposed framework was developed on the basis of the conventional BCI technology (Schalk et al., 2004). It was optimized based on the experimental results and EEG analysis described in the above sections. The entire process is depicted in Fig. 10.

### 5.1. Overall framework

To construct training data, we divided the EEG response acquired in Section 4, into a set of visual comfort (VC) responses and a set of VD responses based on the subjective VD evaluation. We assumed that the VD set has values greater than 2 (OK) according to the definition of the five-scale VD chart. Then, we applied ten-fold cross validation to avoid overfitting and bias. We used 70% of each set as training data and 30% as testing data without the subjective-evaluation-based labels. A support vector machine (SVM) (Cortes and Vapnik, 2008) was trained to distinguish the two EEG response sets (VC/VD) with discriminative EEG feature vectors from the training data. During the testing phase, the EEG feature vectors were examined by the trained classifier to determine whether stimuli would induce VD. Because severe spatial and spectral data variations are common in raw EEG signal processing, subject-specific optimization should be conducted before extracting the EEG feature vectors. To this end, the raw EEG signals were roughly band-pass filtered and down-sampled during both training and

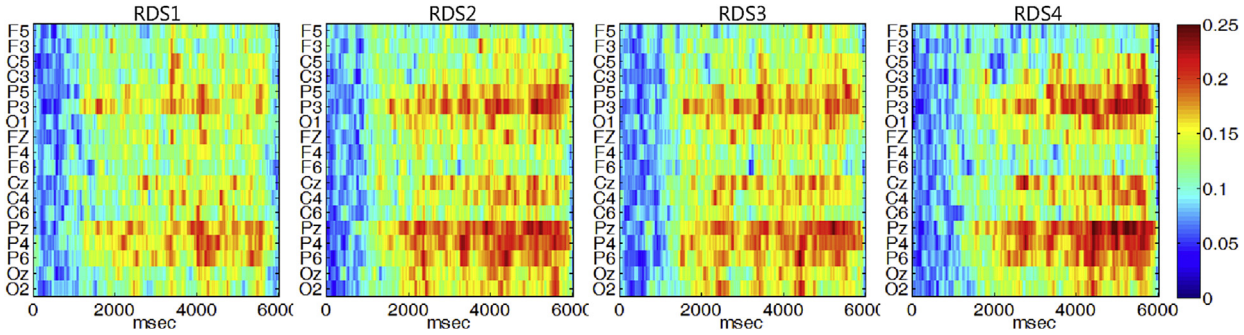


**Fig. 9.** Overall average correlations between envelopes of the band-passed (8–30 Hz) signals and the stimulus levels (RDS1–4). These matrices depict the spatio-temporal behavior of brain oscillations. Note that sufficient correlations are maintained during the SS at the Pz channel. This implies that the target EEG components can be acquired continually as the subject views S3D video, except in the interval (0–1 s).

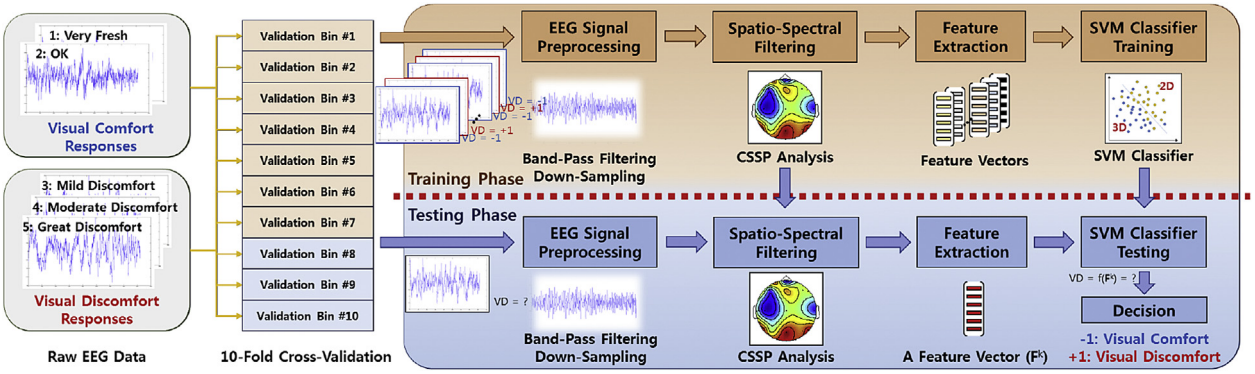


**Fig. 10.** Proposed method for detecting single-trial-based VD. For the training data, raw EEG responses are divided into a VC group and a VD group based on the subjective evaluation. After ten-fold cross validation, a classifier that distinguishes the two groups is trained using a support vector machine (SVM) (Cortes and Vapnik, 2008). Unknown EEG responses are then examined using the classifier. The common spatio-spectral pattern (CSSP) was analyzed (Lemm et al., 2008) to optimize unstable EEG signals in each subject during both training and testing phases.

testing. Then, the vectors were finely optimized with common spatio-spectral pattern (CSSP) analysis (Lemm et al., 2008) during the training phase. This CSSP filter was then reused during the testing phase.

### 5.2. Variable EEG signal optimization

In Section 4, we reported that the discriminative EEG components corresponding to VD were observed in $\alpha$- and $\beta$-bands (8–30 Hz) in the visual cortex (Pz channel). In practice, however, it is difficult to specify the most representative frequency in candidate bands because the optimal bands vary slightly among subjects. In addition, the most discriminative spatial position also varied because the attached electrodes were not attached in exactly the same place for all subjects.

Thus, a variety of BCI frameworks generally use subject-adaptive filters and statistical optimization techniques. One of the most popular spatial filters is the common spatial pattern (CSP) algorithm (Ramoser et al., 2000). It extracts the most discriminative EEG channel by maximizing a variance ratio between two different classes. Suppose $\mathbf{X}^k = (X^k_{c,t}), c = 1, \ldots, C$ and $t = t_0, \ldots, T$ are band-passed EEG potentials for the $k$th trial, where $c$ and $t$ are indices for the EEG electrodes and instants of time, respectively. The goal of the CSP algorithm is to obtain a spatial filter matrix $W$ that projects the original signal $\mathbf{X}^k$ into each of the discriminative spatial bases. It is computed by maximizing the objective function in 1, where $Y^k$ is the class label (comfortable or uncomfortable) for the $k$th trial.

$$W = arg\max_{W} \frac{\left\| WX^k_{\{k:Y^k=+1\}} \right\|^2}{\left\| WX^k_{\{k:Y^k=-1\}} \right\|^2} \tag{1}$$

Although the CSP algorithm can handle the spatial variation in EEG signals, the spectral variation remains troublesome. As an alternative, the CSSP algorithm (Lemm et al., 2008) can be considered. It simultaneously handles both spatial and spectral variations by concatenating the time-delayed version of the original signals. Therefore, we adopted the CSSP algorithm in the proposed framework because the target EEG components are mostly observed in the spatio-spectral domain, and because the target EEG components suffer from the same spatial and spectral variation problems. However, the initial frequency bands for the CSSP algorithm are subject to error, and finding the optimal spectral basis is computationally expensive. To this end, we added two processes to the CSSP algorithm. First, band-pass filtering was used to extract the candidate frequency bands, 8–30 Hz, as input frequency bands. Second, down-sampling was performed to reduce the complexity required in finding the optimal spectral basis.

### 5.3. Feature vector extraction and training

From the CSSP algorithm, we acquire a projection matrix $W$, with which the optimal spatio-spatial filter $\widehat{W}$ is constructed by selecting the $m$ most discriminative bases of $W$ for the class $Y^k = +1$ and the class $Y^k = -1$. Then, $\widehat{W}$ projects the original EEG signal $\mathbf{X}^k$ into the $2m$ spatio-spectral bases, and the projected sample is represented by a $2m$-dimensional projected vector, $\mathbf{Z}^k = (Z^k_{i,t}), i = 1, \ldots, 2m$, as follows:

$$\mathbf{Z}^k = \widehat{W}\mathbf{X}^k \tag{2}$$

Although the projected vector $\mathbf{Z}^k$ retains the optimal channel and frequency band information, it cannot be used directly as a feature vector for the following reasons. First, the recorded EEG potentials varied slightly during a 6 s SS, as shown in Fig. 9. Therefore, converting the time-varying potentials into a single representative value remains a problem. Second, absolute values for these potentials differ for each person and each trial. Therefore, they must be normalized in advance. For these reasons, the conventional feature-extraction method in BCI—viz., log-transformed signal variance—was adopted to convert $\mathbf{Z}^k$ into a tonic feature vector $\mathbf{F}^k$ as in 3, where variances are computed for each filtered channels, $i = 1, \ldots, 2m$.

$$f^k_i = \log\left(var\left(Z^k_i\right)\right) \tag{3}$$

### 5.4. Testing detected VD

For the simulation, we initially selected a wider frequency band (8–40 Hz) than the observed feature frequency bands (8–30 Hz) during band-pass filtering in order to consider possible spectral variations to the raw EEG signals. To down-sample the data, we sub-sampled the original EEG signals every four time-stamps (approximately 8 ms), The EEG signals between 0 and 1 s were skipped in order to strengthen the temporal characteristics. For spatio-spectral filtering, a time-delay parameter $\tau$ was selected empirically to tune the initial frequency band. During feature extraction, we set $m = 3$ to use six-dimensional feature vectors for training and testing the classifier.

Previous works merely distinguish 2D EEG responses from S3D responses (Chen et al., 2013; Li et al., 2008; Cho et al., 2012). By contrast, the proposed method classifies two EEG responses according regarding VD, making it more valuable but also a challenging to extend in future research (Frey et al., 2014; Zou et al., 2015). From this perspective, we detected a set of responses to significant VD—i.e., responses with a subjective evaluation score between 3 and 5—under two different input conditions: with and without the EEG response to the 2D stimulus.

The results are shown in Fig. 11. These results confirm that the proposed framework is considerably accurate—approximately 80% on average—under both conditions. This implies that viewer feedback regarding VD can be provided with high degree of accuracy in ASVSs. Moreover, these results show that VC/VD feedback is elicited by the difference in the magnitude of VD among S3D stimuli. The results for Subjects S3 and S6 are absent under the condition without 2D stimulus. This is because there was an insufficient amount of training data for the VC class, as shown in Fig. 6, where most scores were biased into the VD class.

## 6. Discussion

### 6.1. Subject-specific analysis

As shown in Fig. 6, VD significantly varied between subjects, despite the same test stimuli and viewing environment. This shows the influence of personal factors such as binocular fusion limits and stereo acuity in VD perception. It also suggests that a viewer-interactive approach is needed. In this section, we provide a subject-specific analysis of our results to show how our approach reflects such personal factors.

Unfortunately, we did not previously analyze the subjects' binocular vision characteristics. However, the slopes and the dynamic ranges of each graph in Fig. 6 help to infer the viewers' binocular fusion limits and stereo acuity. For instance, the data for Subject S4 has the gentlest slope with the smallest dynamic range, and S4's evaluation score (4.067) for RDS4 (3.39°) was much lower than the overall average (4.358) for RDS3 (2.53°), which belongs in the outer fusion area. This implies that S4 has the widest fusion limit among all subjects, and likely perceives all RDS images with clear binocular vision. By contrast, the graph for Subject S3 has the steepest slope with the maximum dynamic range (from 1 to 5). Indeed, S3 provided an evaluation score of 5 for RDS3. The evaluation of RDS2 (4.433) was also higher than the overall average for RDS3. These results imply that S3 has the shortest fusion limit, and that RDS2 was likely perceived as a double or blurred image that exceeded S3's binocular fusion limit. The graph for S1 has an intermediate slope with an evaluation score of 4.625 for RDS3, which is higher than the overall average for RDS3. This implies that S1 has an intermediate fusion limit, and that RDS3 might exceed S1's binocular fusion limit and appear as a double or blurred image.

From the above inferences, Fig. 12 compares the log spectra of the three representative subjects (S4, S1, and S3). The results confirm that the degree of spectral attenuation consistently increased according to the relevant RDS level increase when these levels are supposed to be within the subject's binocular fusion limit. When RDS levels exceeded a viewer's binocular fusion limit, the degrees of the corresponding spectra oscillated (rather than attenuated) according to increase in the RDS level. For instance, S4 was judged to have clear single binocular vision across all RDS levels. Indeed, the degrees of spectral attenuation perfectly increased according to the RDS level increase. Likewise, S1 was judged to have clear single binocular vision for RDS0, RDS1 and RDS2. Accordingly, attenuation among the log spectra (black, blue, and green solid lines) corresponding to these RDS images was strongly associated with increased RDS levels. Finally, only RDS1 was judged to fall within S3's binocular fusion limit. Thus, strong spectral attenuation was exclusively observed between RDS0 and RDS1 (the black and blue solid lines), and the other spectra oscillated at around the spectrum for RDS1's EEG response.

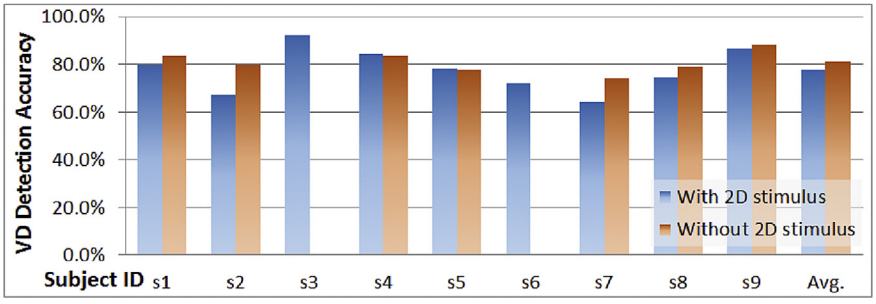


**Fig. 11.** Accuracy of detected VD. We assumed that an EEG response with a subjective-evaluation score between 3 and 5 represents significant VD, and two conditions (with and without an EEG response to 2D stimulus) were compared.

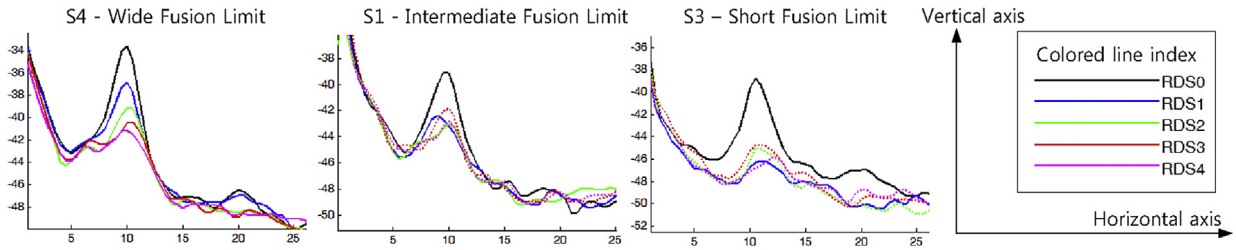


**Fig. 12.** Subject-specific log spectra at the Pz channel. These graphs reveal the differences in EEG responses according to the binocular vision characteristics of three representative subjects (S4, S1, and S3). The colors distinguish EEG responses according to the testing RDS levels, and the line types show whether a testing RDS level is within a subject's binocular fusion limit (solid lines) or beyond it (dotted lines).

### 6.2. Active brain areas

According to literature surveys in neuroscience (Roe et al., 2007; Tyler and Kontsevich, 1995; Parker, 2007), visual information is first processed at the occipital lobe, and then follows the dorsal stream or the ventral stream depending on its purpose as shown in Fig. 13. The dorsal stream terminates in the parietal lobe, and processes information about where the object is located, while the ventral stream terminates in the temporal lobe and processes information about what the object is. On the basis of the above fundamentals, the distinct EEG component observed in our experiment (response around the PZ channel) looks reasonable, and must be captured as the most strengthened feature by mechanisms of stereoscopic processing lined with the dorsal stream. This is because the only variable in our stimuli is binocular disparity that only constructs where information, and the PZ channel is located on the way of the dorsal stream. Therefore, the degree of VD perception could be also explained by the workload of neurons in the dorsal stream to clarify the location information.

The above inference is also reliable when compared to many fMRI-based studies (Backus et al., 2001; Georgieva et al., 2009; Durand et al., 2009; Tsao et al., 2003). In their experiments, S3D stimuli containing only disparity variant were viewed, and distinct responses were observed at the intraparietal sulcus (IPS) regions. As a result, it was revealed that the processing of 3D information by disparity is certainly involved in the IPS regions. In addition, Jung et al. (2015) recently reported an in-depth fMRI analysis that reveals the brain activity linked to the perception of the VD when natural S3D videos are viewed. They found that IPS regions are certainly involved in the processing of excessive binocular disparities, and additionally captured distinct response at eye-movement-related regions (the frontal eye fields (FEF), premotor cortex and IPS). Their observation substantiates that the VAC problem is the basic cause of the VD sensation (Hoffman et al., 2008; Kim et al., 2014; Vienne et al., 2014) in the manner of neuropsychology, and indirectly proves that the abnormal eye-movement was elicited in the VAC condition as reported in Ukai and Kato (2002), where a commercial video refractor was used for capturing the abnormality.

### 6.3. Remaining challenges

The proposed wellness platform is currently limited to distinguishing between two levels of VD sensation. To expand its range, we here present two more major challenges: evaluating quantitative degrees of VD, and studying other relevant factors that contribute to VD.

To quantify VD precisely, the binocular vision characteristics of each viewer should be investigated in advance. Based on this investigation, the levels of S3D stimuli for each viewer need to be calibrated. Then, the customized test stimuli will elicit more meaningful EEG responses as training data to identify the representative features for VD detection. In addition, the training phase should also be extended to a multi-class-based framework in order to train classifiers to quantify VD.

We assumed that the VAC problem is the most critical factor that causes VD. However, actual S3D content is much more complex, with various contributing factors such as a change in the rate of disparity, left and right asymmetry, disparity inversion, etc. These factors are summarized in Table 2. Because these factors contribute to VD, future research is needed to the find discriminative EEG components that correspond to each factor. These should then be combined as a single feature vector. In addition, the following also should be taken into consideration:

- Personal factors: stereopsis disability (strabismus, anisometropia), ametropia, accommodation/convergence amplitude, etc.
- Viewing environment: viewing distance, posture/position, lighting, etc.
- Display factors: display type (active/passive), size, crosstalk, shear distortion, 3D glasses, etc.

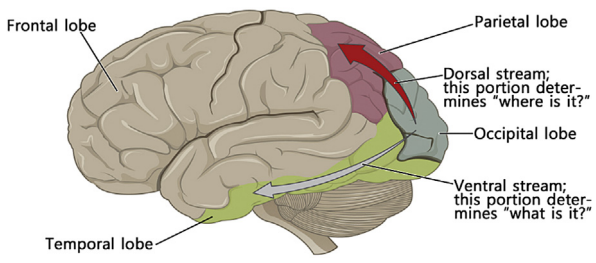


**Fig. 13.** Two-streams hypothesis of the neural processing of human vision (OpenStax College, 2013). First, visual information is processed at the occipital lobe, and then flows to the dorsal or ventral stream depending on its purpose.

### 6.4. Applications in various S3D industries

The proposed wellness platform can be used as a safety measure for both public and personal displays. In public displays, it can provide alarms to indicate the viewers' states. That way, viewers can take the appropriate action before suffering significant VD. In personal displays, a system such as an ASVS can more directly reconstruct the input S3D content according to the viewer's VD feedback. In addition, the proposal can inform safety guidelines for producing S3D content. For instance, a depth chart is essential for safe yet dramatic 3D scenes within a limited depth budget when S3D content is produced. However, the manual creation of such content by a stereographer is time-consuming and dependent on experience. Thus, our platform can help to generate an ideal depth chart more efficiently, by reducing the manual workload. Moreover, hardware manufacturers can utilize the proposed wellness platform to inspect the quality of experience of S3D displays and glasses objectively.

## 7. Conclusion

In this paper, we developed the first wellness platform that exploits EEG to facilitate comfortable S3D video viewing in ASVSs. EEG-based VD evaluations were made by recording raw EEG data repeatedly under natural viewing conditions without any external

**Table 2**
Well-known stereoscopic 3D content factors that cause VD.

| Factors | Details |
|---|---|
| Disparity | Excessive use of absolute disparity<br>Excessive use of relative disparity |
| Disparity's temporal/spatial change | Object motion (local motion)<br>Camera motion (global motion) |
| Stereoscopic image asymmetry | Geometry asymmetry<br>Luminance/chrominance asymmetry |
| Unnatural scene representation | Object structure asymmetry<br>Window edge violation<br>Cardboard effect<br>Depth plane curvature<br>Puppet theater effect |
| Binocular/monocular depth cue conflict | Disparity inversion |

test paradigm such as an ERP and VECP. Nevertheless, we could elicit discriminative EEG components that reflect the perception of VD owing to various EEG signal analysis techniques for extracting spatio-temporal frequency characteristics. On the basis our feature EEG components, we then developed a classifier that could automatically detects severe perception of VD by using a machine-learning-based BCI technology. Although the proposed approach still faces many technical challenges, we believe that it sets the initial grounds for utilizing EEG-based VD evaluation techniques toward various ergonomic S3D applications.

## Acknowledgements

This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2016.

## References

Arican, Z., Yea, S., Sullivan, A., Vetro, A., 2009. Intermediate view generation for perceived depth adjustment of stereo video. In: In Proc. SPIE Applications of Digital Image Processing XXXII, 74430U.
Backus, B.T., Fleet, D.J., Parker, A.J., Heeger, D.J., 2001. Human cortical activity correlates with stereoscopic depth perception. J. Neurophysiol. 86 (4), 2054–2068.
Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Muller, K.R., 2008. Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Process. Mag. 25 (1), 41–56.
Chen, C., Li, K., Wu, Q., Wang, H., Qian, Z., Sudlow, G., 2013. EEG-based detection and evaluation of fatigue caused by watching 3DTV. Displays 34 (2), 81–88.
Cho, H., Kang, M.-K., Yoon, K.-J., & Jun, S. C., 2012. Feasibility study for visual discomfort assessment on stereo images using EEG. In Proc. IEEE IC3D, 1–6.
Choi, J., Kim, D., Ham, B., Choi, S., & Sohn, K., 2010. Visual fatigue evaluation and enhancement for 2D-plus-depth video. In Proc. IEEE ICIP, 2981–2984.
Cortes, C., Vapnik, V., 2008. Support-vector networks. Mach. Learn. 20 (3), 273–297.
Durand, J.B., Peeters, R., Norman, J.F., Todd, J.T., Orban, G.A., 2009. Parietal regions processing visual 3D shape extracted from disparity. Neuroimage 46 (4), 1114–1126.
Emoto, M., Niida, T., Okano, F., 2005. Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television. J. Disp. Technol. 1 (2), 328–340.
Frey, J., Pommereau, L., Lotte, F., & Hachet, M., 2014. Assessing the zone of comfort in stereoscopic displays using EEG. In Proc. ACM CHI, 2041–2046.
Georgieva, S., Peeters, R., Kolster, H., Todd, J.T., Orban, G.A., 2009. The processing of three-dimensional shape from disparity in the human brain. J. Neurosci. 29 (3), 727–742.
Ha, K., & Kim, M., 2011. A perceptual quality assessment metric using temporal complexity and disparity information for stereoscopic video. In Proc. IEEE ICIP, 2525–2528.
Hagura, H., & Nkajima, M., 2006. Study of asthenopia caused by the viewing of stereoscopic images: measurement by MEG and other devices. In Proc. SPIE, 60570K–60570K.
Hoffman, D.M., Girshick, A.R., Akeley, K., Banks, M.S., 2008. Vergenece-accommodation conflicts hinder visual performance and cause visual fatigue. J. Vis. 8 (3), 1–30.
Hsieh, P., Colas, J., Kanwisher, N., 2010. Baseline fMRI pattern activity in early visual cortex predicts the initial dominant percept in subsequent binocular rivalry. J. Vis. 10 (7).
ISO IWA3, 2005. Image Safety Reducing the Incidence of Undesirable Biomedical Effects Caused by Visual Image Sequences (Geneva).
ISO/IEC JTC1/SC29/WG11, 2009. Applications & Requirements on 3D Video Coding. N11061, Xian.
ITU, 2002. Methodology for the Subjective Assessment of the Quality of Television Pictures. Recommendation BT.500-11.
Jap, B.T., Lal, S., Fischer, P., Bekiaris, E., 2009. Using EEG spectral components to assess algorithms for detecting fatigue. Expert Syst. Appl. 36 (2), 2352–2359.
Jung, Y.J., Lee, S., Sohn, H., Park, H.W., Ro, Y.M., 2012. Visual comfort assessment metric based on salient object motion information in stereoscopic video. J. Electron. Imaging 21 (1), 011008-1–; 011008–16.
Jung, Y.J., Kim, D., Sohn, H., Lee, S.I., Park, H.W., Ro, Y.M., 2015. Towards a physiology-based measure of visual discomfort: brain activity measurement while viewing stereoscopic images with different screen disparities. J. Disp. Technol. 11 (9), 730–743.
Kar, S., Bhagat, M., Routray, A., 2010. EEG signal analysis for the assessment and quantification of drivers fatigue. Transp. Res. part F traffic Psychol. Behav. 13 (5), 297–306.
Kim, J., Kane, D., Banks, M.S., 2014. The rate of change of vergence accommodation conflict affects visual discomfort. Vis. Res. 105, 159–165.
Kuze, J., Ukai, K., 2008. Subjective evaluation of visual fatigue caused by motion images. Displays 29, 159–166.
Laghari, K., Gupta, R., Arndt, S., Antons, J.N., Schleicher, R., Moller, S., Falk, T.H., 2013. Neurophysiological experimental facility for quality of experience (QoE) assessment. In: In Proc. Int. Conf. QCMan, pp. 1300–1305.
Lambooij, M., IJsselsteijn, W., Fortuin, M., Heynderickx, I., 2009. Visual discomfort and visual fatigue of stereoscopic displays: a review. J. Imaging Sci. Technol. 53 (3), 1–14.
Lemm, S., Blankertz, B., Curio, G., Muller, K.-R., 2008. Spatio-spectral filters for improving the classification of single trial EEG. IEEE Trans. Biomed. Eng. 52 (9), 1541–1548.
Li, H.O., Seo, J., Kham, K., Lee, S., 2008. Measurement of 3D visual fatigue using event-related potential (ERP): 3D oddball paradigm. In: In Proc. 3DTV Conference, pp. 213–216.
Li, W., He, Q.C., Fan, X.M., Fei, Z.M., 2012. Evaluation of driver fatigue on two channels of EEG data. Neurosci. Lett. 506 (2), 235–239.
Liu, J., Zhang, C., Zheng, C., 2010. EEG-based estimation of mental fatigue by using KPCAHMM and complexity parameters. Biomed. Signal Process. Control 5 (2), 124–130.
Mun, S., Park, M.-C., Park, S., Whang, M., 2012. SSVEP and ERP measurement of cognitive fatigue caused by stereoscopic 3D. Neurosci. Lett. 525 (2), 89–94.
Murata, A., Uetake, A., Takasawa, Y., 2005. Evaluation of mental fatigue using feature parameter extracted from event-related potential. Int. J. Ind. Ergon. 35 (8), 761–770.
Ndjiki-Nya, P., Koppel, M., Doshkov, D., Lakshman, H., Merkle, P., Muller, K., Wiegand, T., 2011. Depth image-based rendering with advanced texture synthesis for 3D video. IEEE Trans. Multimedia 13 (3), 453–465.
Onural, L., 2007. Television in 3-D: what are the prospects? Proc. IEEE 95 (6), 1143–1145.
OpenStax College, 2013. Anatomy & Physiology [Online]. Available: http://cnx.org/content/col11496/1.6/.
Parker, A.J., 2007. Binocular depth perception and the cerebral cortex. Nat. Rev. Neurosci. 8 (5), 379–391.
Ramoser, H., Muller-Gerking, J., Pfurtscheller, G., 2000. Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Transactions on Rehabil. Eng. 8 (4), 441–446.
Reisman, S., 1997. Measurement of physiological stress. In: Bioengineering Conference, 1997, Proceedings of the IEEE 1997 23rd Northeast, pp. 21–23.
Roe, A.W., Parker, A.J., Born, R.T., DeAngelis, G.C., 2007. Disparity channels in early vision. J. Neurosci. 24 (44), 11820–11831.
Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R., 2004. BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Trans. Biomed. Eng. 51 (6), 1034–1043.
Schmidt, D., Abel, L.A., Dell'osso, L.F., Daroff, R.B., 1979. Saccadic velocity characteristics: intrinsic variability and fatigue. Aviat. Space Environ. Med. 50 (4), 393–395.
Sheard, C., 1934. The prescription of prisms. Am. J. Optom. 11 (10), 364–378.
Stern, J.A., Boyer, D., Schroeder, D., 1979. Blink rate: a possible measure of fatigue. Hum. Factors 36 (2), 285–297.
Sweeney, L.E., Seidel, D., Day, M., Gray, L.S., 2014. Quantifying interactions between accommodation and vergence in a binocularly normal population. Vis. Res. 105, 121–129.
Trejo, L. J., Kochavi, R., Kubitz, K., Montgomery, L. D., Rosipal, R., & Matthews, B., 2005. EEG-based estimation of cognitive fatigue. In Proc. SPIE, 5797, 105–115.
Tsao, D.Y., et al., 2003. Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. Neuron 39 (3), 555–568.
Tyler, C.W., Kontsevich, L.L., 1995. Mechanisms of stereoscopic processing: stereoattention and surface perception in depth reconstruction. Perception-London 24 (1), 127–154.
Ujike, H., 2009. Estimation of visually induced motion sickness from velocity component of moving image. Proc. Virtual Mix. Real. 5622, 136–142.
Ukai, K., Kato, Y., 2002. The use of video refraction to measure the dynamic properties of the near triad in observers of a 3-D display. Ophthal. Physiol. Opt. 22, 385–388.
Vienne, C., Sorin, L., Blond, L., Huynh-Thu, Q., Mamassian, P., 2014. Effect of the accommodation-vergence conflict on vergence eye movements. Vis. Res. 53, 124–133.
Vos, S.H., Gunter, T.C., Kolk, H.H., Mulder, G., 2001. Working memory constraints on syntactic processing: an electrophysiological investigation. Psychophysiology 38 (1), 41–63.
Wopking, M., 1995. Viewing comfort with stereoscopic pictures: an experimental study on the subjective effect of disparity magnitude and depth of focus. J. Soc. Inf. Disp. 3, 101–103.
Zou, B., Liu, Y., Guo, M., Wang, Y., 2015. EEG-based assessment of stereoscopic 3D visual fatigue caused by vergence-accommodation conflict. J. Disp. Technol. 11 (12), 1076–1083.