Wireless Pers Commun (2015) 84:1133–1149
DOI 10.1007/s11277-015-2680-z

Implementation of Multimodal Interactive Continuous Scoring for 3D Quality of Experience

Jiwoo Kang[1] · Taewan Kim[1] · Sanghoon Lee[1]

Published online: 19 May 2015

Abstract When watching videos in 3D space, viewers perceive dynamic levels of quality of experience accompanied by visual immersion. To measure these dynamics, a reliable methodology is needed to gauge subjective viewer experience. This paper proposes a new methodology called multimodal interactive continuous scoring of quality (MICSQ). MICSQ is comprised of device interaction processes between the 3D display and tablet assessment tool and human interaction processes between the subject and tablet. When MICSQ device interaction takes place over wireless network protocols, such as TCP/IP or Bluetooth, it efficiently handles the diverse viewing environment. Therefore, there is a high degree of freedom to perform subjective assessment in certain viewing environments in terms of multimodal cues (aural and tactile senses), diverse illumination conditions including darkness, handheld portability over wireless networks, and real-time recording. Moreover, it is also possible for multiple subjects to simultaneously perform assessments in a large space, such as a movie theater. For the simulations, the server application in the 3D display was developed in Java, and the tablet device client application was developed with a mobile software development kit and functions optimally in commercial tablets. The experimental results demonstrate that MICSQ shows a higher reliability than the conventional single stimulus continuous quality evaluation method through the proposed implementation on a commercial tablet PC.

Keywords 3D QoE · Mobile assessment device · Wireless quality assessment · Visual discomfort · Interactive continuous quality evaluation

✉ Sanghoon Lee
slee@yonsei.ac.kr

Jiwoo Kang
jwkang@yonsei.ac.kr

Taewan Kim
enoughrice21@yonsei.ac.kr

[1] Multi-dimensional Insight Laboratories, Department of Electrical and Electronic Engineering, Yonsei University, 703, 3rd Engineering Building, Seoul 120-749, Korea

1 Introduction

Recent 3D technological developments have led to a significant increase in the demand for 3D content [1, 2]. Various concerns about high-quality, visually comfortable 3D content have therefore been raised in terms of high-end service, including in related areas that lie beyond the scope of content, such as the design of 3D displays and 3D movie theaters [2–5].

However, unlike 2D content, 3D content is affected by numerous factors, such as photosensitive viewer seizures, visual-induced viewer motion sickness, visual fatigue, and image distortions [1–4]. These factors involve complex visual features; therefore, a solid subjective methodology for measuring viewer quality of experience (QoE) is essential.

3D QoE is much more complicated to understand than 2D QoE because the former is affected by complex characteristics of the human visual system and individual differences therein. Therefore, it is difficult to design a reliable objective metric for 3D QoE prediction, a topic which remains relatively unexplored. Indeed, brain activity-measuring devices, including the electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI), and optometric clinically based measurements, such as AC/A ratio and CA/C ratio, have been used instead of objective models [1, 6]. In particular, approaches to 3D subjective assessments have been inherited from traditional approaches to 2D subjective assessments. However, it is not certain whether their results are sufficiently reliable to be used as a reference because the 3D viewing environmentwhich involves an intensive immersion of users wearing 3D glasses in dark lightingis quite different from the 2D environment. It is therefore essential to develop a more reliable subjective assessment methodology for 3D viewing addressing the characteristics of human perception, display mechanism, and viewing environment. Recently, international standardization activity on 3D display, human factors, and image safety has been reported [7]. Despite this effort, the specific subjective 3D QoE assessment environment, including an assessment tool, remains insufficient. To date, subjective 3D assessment methods have continued to be adapted from subjective 2D methods. The 2D subjective assessment is regarded as a one-way framework in which the subject provides a rating of a test video. However, there are limitations to applying this framework to a 3D subjective assessment because inevitable problems would occur. In the immersion problem, for example, subjects are typically absorbed in entertaining 3D video content and forget the assessment process [8]. Therefore, subsequent problems arise, such as viewer unawareness of what they are rating. In addition, a loss of viewer concentration may occur during long assessment periods, which causes inaccurate synchronization between the test video and rating result. Moreover, the slow reaction speed of the assessment interface also causes a degradation of assessment reliability. In general, subjective 3D QoE assessments are conducted in a dark room; therefore, the subject may experience difficulties in the recording process. The subjective result is subsequently an inaccurate reference as a baseline for comparison with objective 3D QoE models.

It is imperative to guide subjects to prevent unintended problems in subjective 3D QoE assessments. One solution, for example, is managing interaction between the subject and assessment environment to minimize distractions during the assessment task. Specifically, assessment tools are needed to intelligently detect the concentration loss of a subject and to provide warnings using beeps and vibration. This type of interaction management should unquestionably increase the reliability of the assessment task. In addition, to guarantee subject convenience, it would be optimal to conduct the assessment in a wireless environment via a portable device, such as a tablet PC. This setup would enable the subject to conduct the assessment with a high degree of freedom. To simultaneously conduct the assessment with multiple subjects, wireless portability should be implemented in the subjective 3D QoE assessment methodology. This capability would enable implementation of a more reliable real-time cooperative structure between client and server. Each client should send the subjective ratings to the server, which in turn must store the results and play the subsequent test videos while maintaining the time synchronization between them.

To minimize distractions between viewing and assessing tasks and to maximize the degree of freedom for subjective 3D QoE assessments, this paper propose a novel method called multimodal interactive continuous scoring of quality (MICSQ). This method considers the characteristics of human perception, display mechanism, and viewing environment of the 3D display. In particular, the main difference between the proposed method and conventional ones is that the proposed method is a two-way interactive framework involving a device interaction process between the 3D display and an assessment tool and a human interaction process between the subject and device. With this method, it is possible for multiple users to simultaneously perform a subjective assessment. For the simulations, the server application in the 3D display is comprised of Java, and the tablet device client application was built with a mobile software development kit to function optimally in commercial tablets. When one or more clients connect to the server via a wireless network, such as TCP/IP or Bluetooth, the server application sends all information for the assessment to the client; it then receives each client result and integrally stores and manages them. The client helps prevent the subject from experiencing immersion and concentration loss problems by providing multimodal cues (aural, visual and haptic). The synchronization process guarantees that the delay between the server and client never exceeds a threshold of maximally allowed delay.

2 Motivation and MICSQ Overview

2.1 Subjective 3D Video Assessment

The subjective assessment of 3D video can be classified as explorative study, psychophysical scaling, or questionnaire formats. Traditionally, explorative studies on 3D displays are conducted by gathering focus group opinions after subjects view a test sequence [1]. Psychophysical scaling is also widely used for methods that are performance-

Table 1 Previous subjective 3D video assessment researches

Method	Reference paper	Topic
DSCQS	[9]	Visual quality for asymmetric stereo video coding
	[10]	Visual quality for stereoscopic perceptual video coding
	[11]	Visual quality and stereo sense assessment of stereo images
SSCQE	[7]	Assessment methods of 3D TV (continuous-/retrospective-/single-assessment)
	[12]	Effects of parallax distribution and depth motion on visual comfort
	[13]	Relation among visual quality, naturalness, and depth perception
	[14]	Assessment of presence, depth and naturalness for a stereoscopic TV program

oriented (used to execute a certain task) and appreciation-oriented (used to measure degree of appreciation). In general, for the appreciation-oriented method, double stimulus continuous quality scale (DSCQS) and single stimulus continuous quality evaluation (SSCQE) are mainly used in subjective 3D video assessment studies because of the time-variant characteristics of 3D content, as shown in Table 1. However, although these methods have been applied in many 3D research areas, their approach is almost the same as those of methods described in 2D recommendation documents [15]. In addition, the questionnaire method has been widely applied as a general means to assess the degree of video quality and visual discomfort [1].

2.2 Overview of MICSQ

As shown in Fig. 1, if there is no particular mention of the test procedure, the test sequence and assessment interface may appear on the same display for SSCQE, even if the display is stereoscopic. Unintended problems may occur in the occlusion area between the fully displayed test sequence and rating interface. Therefore, for reliable subjective 3D video assessment, it is ideal to not display the assessment tool on the 3D display. Nevertheless, the accuracy of the subjective methodology has not been empirically measured.

The goal of MICSQ is to minimize unintended assessment errors between viewing and assessment. As shown in Fig. 2a, SSCQE is regarded as a one-way framework because the subject assigns a subjective score while watching the test video. Unlike such a framework, as shown in Fig. 2b, MICSQ is regarded as an interactive framework that consists of two interactions: (1) device interaction between the tablet(s) used as an assessment tool and the 3D display over a wireless network protocol and (2) human interaction between the tablet(s) and the subject(s) using vibration and sound from the tablet to the subject and scoring the subjective results on the tablet. In addition, using a wireless protocol, it is possible for multiple subjects to simultaneously assess 3D QoE in a large space, such as a movie theater.

(a) (b)

Fig. 1 Common interfaces of subjective 2D assessment and an example interface of subjective 3D assessment

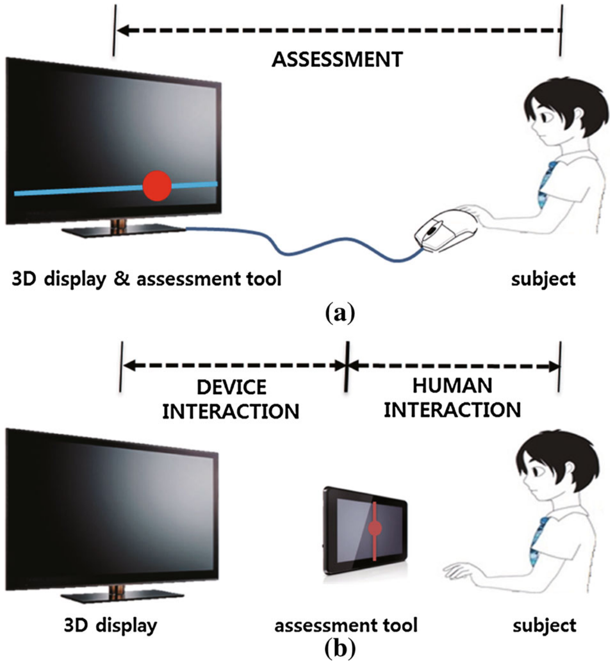

Fig. 2 Conceptual diagrams of conventional and proposed MICSQ protocols for subjective 3D video QoE assessment

3 MICSQ Methodology for Subjective 3D Video Assessment

MICSQ is designed for subjective 3D QoE assessment based on the interaction process among the 3D display, assessment tool(s), and subjects. It is activated by the human-interaction and device-interaction processes as follows.

3.1 Human Interaction of MICSQ

The human interaction process provides a main difference from conventional ones by using a tablet. The concentration loss may interrupt an accurate task of continuous assessment (describe as concentration loss problem), and the immersion to 3D video can also cause the obtaining of unreliable assessment results (describe as immersion problem). To resolve these problems, the assessment tool of MICSQ makes sight, hearing and touch cues to subjects consistently during the assessment as shown in Fig. 3. The tablet may lead the subjects to preventing loss of concentration by using the periodic flickering, beep sound and vibration. In order to reduce the immersion problem, the tablet adjusts the cycle timing of the flickering, beep sound and vibration randomly and announces the rating score in every second. Through announcing the score, subjects can perceive their own rating score unwittingly, even if the subjects are fully absorbed in watching test sequence. In particular, although the standard has not been officially discussed yet, the haptic interface which is a tactile feedback technology between the subject and the tablet is very appropriate as a human interaction process. A variety of multimodal cues have been proposed and discussed in previous researches. The physical and cognitive interactions between the

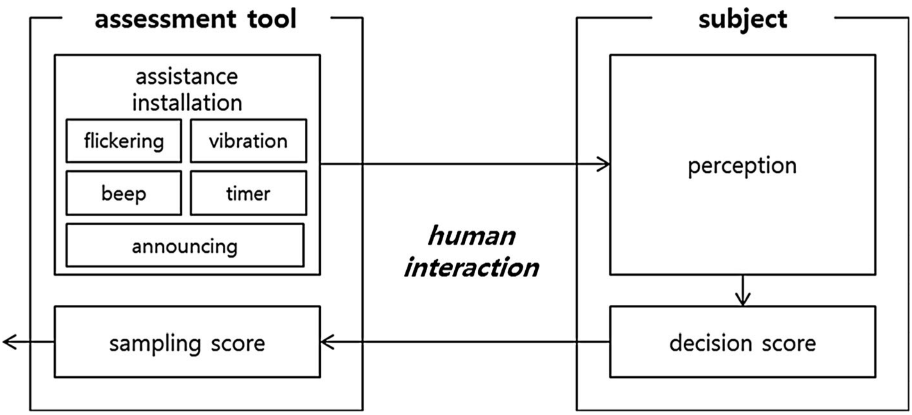

Fig. 3 Human interaction process of MICSQ

responses to haptic, sight and hearing cues tend to mutually enhance the efficacy of each modality [16]. In particular, the haptic feedback technology has the potential to minimally intrude upon the subjects awareness while supplying multimodal cues for guidance, control, and distraction reduction. Arranged items of the assistant device against the concentration loss and immersion problem are summarized in Table 2.

3.2 Device Interaction of MICSQ

The device interaction process is conducted using the interaction process between the server and tablet in real time. PC is used as a server and the assessment tool is used as a client. Almost all tablet PCs based on general purpose operating system (GPOS) can be used as the assessment tool.

The two important roles of the server are playing a test sequence and storing the rating score obtained by subjects. To maintain long-term synchronization, the tablet sends the server the time duration from the beginning to the ending of the assessment, as shown in Fig. 4. However, when the subject mistakenly stops the assessment task, the device interaction process causes the assessment task to cease and restart from the beginning of the test sequence. In addition, based on the viewing distance, wireless network protocols such as Wireless LAN (IEEE 802.11) and Bluetooth (IEEE 802.15.1) can be used interchangeably for communications between server and client for MICSQ.

The server requires applications to play test sequences and manage the assessment task through communications with the client. Moreover, the client requires an application that interacts with both subject and server. These applications of server and client cooperatively

Table 2 Multimodal assessment tool cues

	Concentration loss problem	Immersion problem
Sight stimulus	Periodic flickering	Shortened flickering
Aural stimulus	Periodic beep sound	Randomized beeping
		Announcing scores
Haptic stimulus	Periodic vibration	Randomized vibration

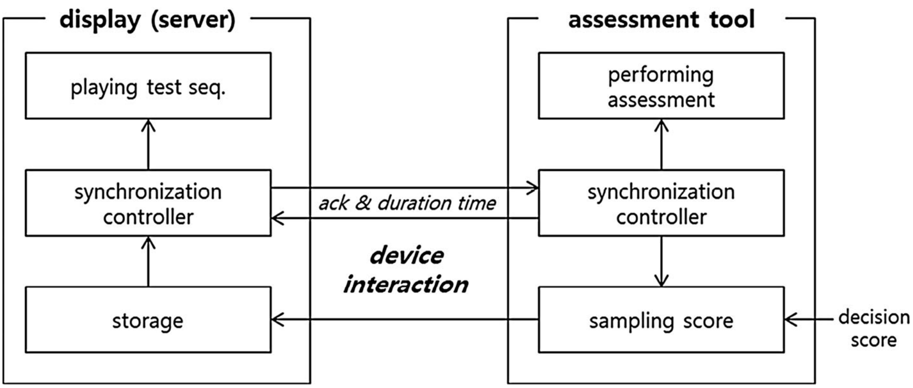

Fig. 4 Device interaction process of MICSQ

execute the subjective assessment based on predefined protocols. Subjective scores are recorded continuously in the server at every pre-defined sampling interval according to the ITU recommendation (a typical sampling rate is 2 Hz; 2 samples per 1 s) [15].

The scores are calculated using the horizontal x-coordinate values on the tablet screen. The screen is uniformly divided into horizontal sections in five intervals with equally spaced marks [bad]–[poor]–[sufficient]–[good]–[excellent]. These intervals represent the assessment criteria. As shown in Fig. 5d, the vertical lines, interval lines between each criterion, current voting score, and elapsed time since the sequence began are simultaneously represented on the screen. The range of subjective scores is set from 0 to 10, with an

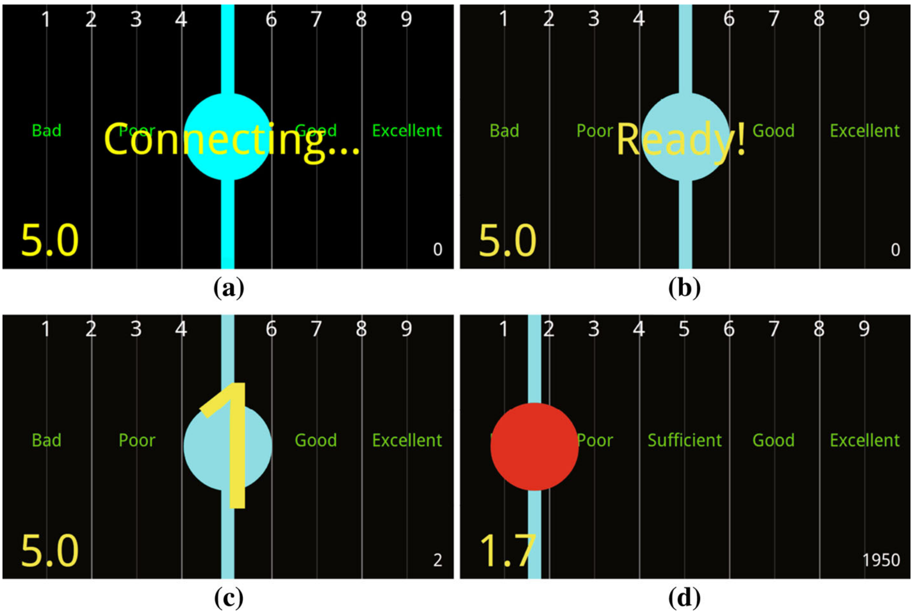

Fig. 5 Clients UIs at each process of device interaction process; **a** connecting process, **b** ready process, **c** counting process, **d** assessment process

initial score of 5; it continuously varies when the subjects finger moves to the left or right on the screen. At the same time, the client calculates the score at every sampling time and robustly sends this information to the server.

The functions of the server application include display of the stereoscopic test sequences and management of the assessment processes through communications with the client. It does not matter whether both functions are implemented in a single application or if each function is implemented in a different application and the applications interact with each other by predefined key events or messages. Before the assessment, the server should know the number of sequences, a playlist of sequence videos, playing times of sequence videos, voting intervals of each sequence, and the maximum allowed delay between the server and client. In MICSQ, there is no information relating to the assessment in the client; rather, the server sends the data required for the assessment to the client at each assessment. This is useful when the assessments are conducted using several assessment tools. In such a case, each assessment tool requires only one assessment application and obtains all required data needed for the assessment from the server, to which it sends scores.

3.3 Device Interaction Signal Flow

The overall MICSQ process is divided into four sub-processes: initial, ready, counting and synchronization, and assessment. These sub-processes are outlined in Fig. 6. The control signals shown in the figure are named for convenient understanding.

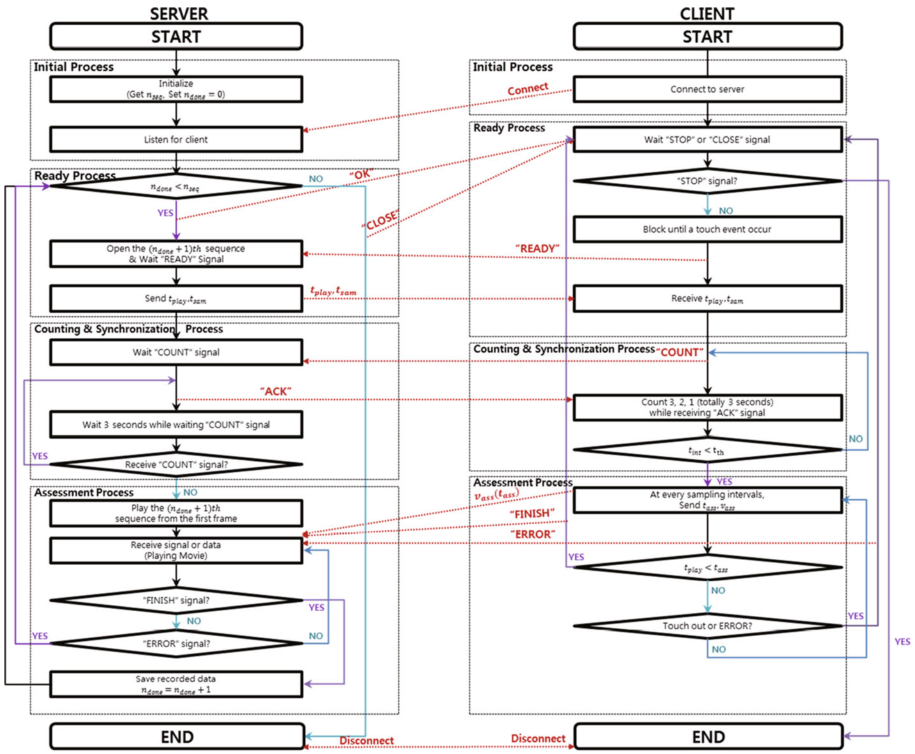

Fig. 6 Signal flow of MICSQ protocol

In the initial process, the server application must be started before the client is ready. The server reads the number of sequences to be assessed (n_{seq}), initializes the number of sequences already assessed (n_{done}) to zero, and then waits until the client connects. Table 3 presents pseudo code of the initial process.

In the ready process, the server determines if the number of sequences already assessed (n_{done}) is equal to or greater than the number of sequences to be assessed (n_{seq}) at the initial time. If all the test sequences are already rated, the server sends the CLOSE signal so that the client finishes the assessment. If there are sequence(s) remaining to be assessed, the server opens a sequence and sends the OK signal, indicating its readiness for the assessment. When the server receives the READY signal from the client, it then sends the play time of the current sequence (t_{play}) and the scoring interval (t_{sam}) to the client.

In the ready process, the client waits for either the CLOSE or OK signal. If the former arrives, the client finishes the assessment; if the latter arrives, the client waits until the subject touches the screen. When the subject touches the screen, the client sends the READY signal to the server to notify the client that the subject is ready for the assessment. Table 4 presents pseudo code of the ready process.

In the counting and synchronization process, as soon as the client receives the play time of the sequence and the scoring interval from the server, it sends the COUNT signal to the server and undergoes a three second count conveying the upcoming start of the sequence assessment to the subject, as shown in Fig. 5c. It is necessary not only for the subject to prepare the assessment of the sequence, but for the synchronization to guarantee the maximum delay between the server and client.

Three seconds after the client sends the COUNT signal, the assessment begins; 3 s after the server receives the COUNT signal, it begins playing the sequence. Therefore, the time difference (t_{syn}) between when the assessment begins and when the sequence begins playing is regarded as the time interval between when the client sends the COUNT signal and when the server receives that signal, as shown in Fig. 7.

The server receives the COUNT signal from the client; it then responds to the client with the ACK signal. When the client receives that signal, it calculates the interval time (t_{syn}) between the transmission of the COUNT signal and the reception of the ACK signal. It is always true that tint is greater than t_{syn}, as described in Fig. 7. In other words, it is guaranteed that the time difference between when the server starts playing the sequence and when the client begins the assessment is always less than t_{int}. Thus, if tint is less than the pre-defined maximum delay (t_{th}) between the server and the client, both the client and the server return to the beginning of the counting and synchronization process and repeat these steps until the maximum guaranteed delay is achieved. Table 5 shows pseudo code of the counting and synchronization process.

During the assessment process, the subject continuously scores the degree of perceived QoE. At every scoring interval (t_{sam}), the client sends the score (v_{ass}) to the server, and the server records the received scores until it receives either the FINISH or ERROR signal.

Table 3 Pseudo code of client and server at initial process

Server	Client
$n_{seq} = 0$, $n_{done} = 0$	*servSock* = connect(*server IP*)
clntSock = listen(*port*)	receive(*servSock*, *buf*)

Table 4 Pseudo code of client and server at ready process

Server	Client
// Label of READY location	// Label of READY location
READY:	READY:
if ($n_{done} \geq n_{seq}$) {	// *buf* can has either CLOSE or OK
send(*clntSock*, CLOSE);	if (*buf* == CLOSE) {
finishServer();	disconnect(*serverIP*);
}	finishClient();
	}
send(*clntSock*, OK);	
openSeqence(n_{done}+1);	// block until a touch event occurs
	while (*touch event is no occur*);
// receive READY signal	
receive(*clntSock*, *buf*);	send(*servSock*, READY);
send(*clntSock*, t_{play});	receive(*servSock*, *buf*);
send(*clntSock*, t_{sam});	t_{play} = atoi(*buf*);
	receive(*servSock*, *buf*);
	t_{sam} = atoi(*buf*);

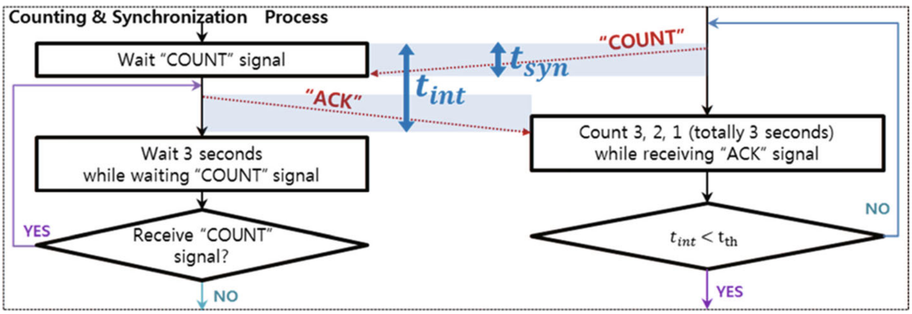

Fig. 7 The flow chart of synchronization process

When the elapsed time (t_{play}) after the assessment sequence begins is greater than the sequence play time (t_{ass}), the assessment sequence is successfully finished. Therefore, the server sends the FINISH signal to notify the client that the assessment sequence is successfully completed; and the server returns to the connecting process to assess the next sequence. However, if a touch release mistakenly occurs during the assessment, the client sends the ERROR signal to notify the client that the assessment sequence should be repeated from the beginning; it returns to the connecting process to again assess the current

Table 5 Pseudo code of client and server at counting and synchronization process

Server	Client

// receive COUNT signal
receive(*clntSock*, *buf*);
 	
do {
    send(*clntSock*, ACK);
 	
    // receive a signal for 3000ms
    r = receive(*clntSock*, *buf*, 3000);
 	
} while (r != TIMEOUT);
// TIMEOUT means there is
// no signal received from the client
	t_{start} = getCurrentTime();
	
	send(*servSock*, COUNT);
	
	for ($i = 3$; $i > 0$; $i--$) {
	    showCount(i);
	
	    // receive the ACK signal
	    receive(*servSock*, *buf*, 0);
	
	    t_{finish} = getCurrentTime();
	    $t_{int} = t_{finish} - t_{start}$;
	
	    if not ($t_{int} < t_{th}$) {
	        send(*servSock*, COUNT);
	        $i = 3$;
	    }
	}

sequence. When the server receives either the FINISH or ERROR signal, it stops recording. In the case of the FINISH signal, the recorded data is stored by the server, which returns to the ready process by increasing the number of sequences (n_{done}). Subsequently, the server proceeds to assess the next sequence or completes the assessment task. However, if the server receives the ERROR signal, it returns to the ready process for again assessing the current sequence. Table 6 shows pseudo code of the assessment process.

4 MICSQ Reliability Simulation Result

4.1 Assessment Environment

4.1.1 Viewing Environment

To perform subjective 3D QoE assessment using MICSQ, 46 in. polarized stereoscopic display with a 1920 × 1080 resolution and display height of 0.6 m was used. The viewing distance was set to 1.8 m following ITU-R Rec. 500-13 [15]. A commercial tablet PC was used as an assessment tool and was fixed on the table at a distance of 0.4 m from the display. In addition, the subjective assessment was conducted under constant room and background illumination conditions as shown in Fig. 8.

Table 6 Pseudo code of client and server at assessment process

Server	Client

playSequence(n_{done}+1);
while (true) {
    receive(*clntSock*, *buf*);
 	
if (*buf* == FINISH or ERROR)
    break;
else
    recordData(*data*, *buf*);
}
 	
if (*buf* == FINISH) {
    saveData(*data*);
    $n_{done} = n_{done} + 1$;
}
 	
clearData(*data*);
goto READY:
	while(true) {
	    v_{ass} = getCurrentScore();
	    send(*servSock*, v_{ass});
	
	    if($t_{play} < t_{ass}$) {
	        send(*servSock*, FINISH);
	        goto START:
	    }
	    else if (*touch error occur*) {
	        send(*servSock*, ERROR);
	        goto READY:
	    }

Fig. 8 Apparatus of experimental equipment for MICSQ

4.1.2 Subjects

Thirty subjects (21 males and 9 females) participated in the assessment, in which six of the subjects were involved in 3D research. The ages of the subjects ranged from 24 to 31 years, with an average age of 27. All subjects had normal or good visual acuity >1.25 (the Landolt C-test) and a good stereoscopic acuity <60 arc (the RANDOT stereo test). If the rating score of a subject differed too greatly from those of the others, the subjects result was handled as an outlier following the rejection procedure in [17].

4.1.3 Stimuli

Subjective assessment is conducted for visual comfort on 20 test videos from 3D video databases made available by IEEE-SA [7] and EPFL [18]. All sequences had a high-definition resolution (1920 × 1080) with a play length of 10 s. In addition, the sequences from the databases had a frame rate of 30 and 25 fps, respectively.

4.1.4 Procedure

Each subject was asked to assess the degree of visual comfort experienced when viewing the above test videos using both MICSQ and SSCQE protocols. For the benchmark, SSCQE was conducted using a mouse while the test sequence and rating scores were displayed on the same display.

4.2 MICSQ and SSCQE Comparison

Because the assessment was conducted using multiple subjects, the mean opinion score (MOS) was obtained as

$$d_k = \sum_{j=1}^{N} s_{jk}/N \tag{1}$$

where N is the number of subjects, ($N = 40$) and sjk is the subjective score by subject j for test sequence k. To obtain the statistical reliability of the collected data, confidence intervals (CIs) on the MOS values are utilized. Based on the MOS of all subjects, the CI of was computed according to the Students t-distribution:

$$CI_k = t\left(\frac{1-\alpha}{2}, N\right) \cdot \frac{\sigma_k}{\sqrt{N}} \tag{2}$$

where $t((1-\alpha)/2, N)$ is the t-value corresponding to a two-tailed t-Student distribution with $N - 1$ degrees of freedom; is the standard deviation of a single test condition across subjects and a desired significance level.

Figure 9 depicts the subjective 3D QoE (visual comfort) assessment scores and their CIs for three representative test sequences. The values of the MOS using MICSQ and SSCQE were quite similar. However, as shown in the second column, the average length of the CIs for MICSQ was shorter than that for SSCQE at each sample. Moreover, the standard deviation of the MICSQ was also smaller than that of SSCQE, as described in Table 7, which implies that the subjective assessment of MICSQ has a higher reliability than SSCQE.

In general, due to the 3D immersion of human viewers, the reaction of a subject becomes slower and the concentration paid to the assessment decreases. Therefore, the subjects do not accurately assess the 3D video and response time is delayed. However, as shown in Fig. 9, subjective scores of MICSQ demonstrate a wider dynamic range of human perception than those of the SSCQE. Therefore, it is hypothesized that the availability of

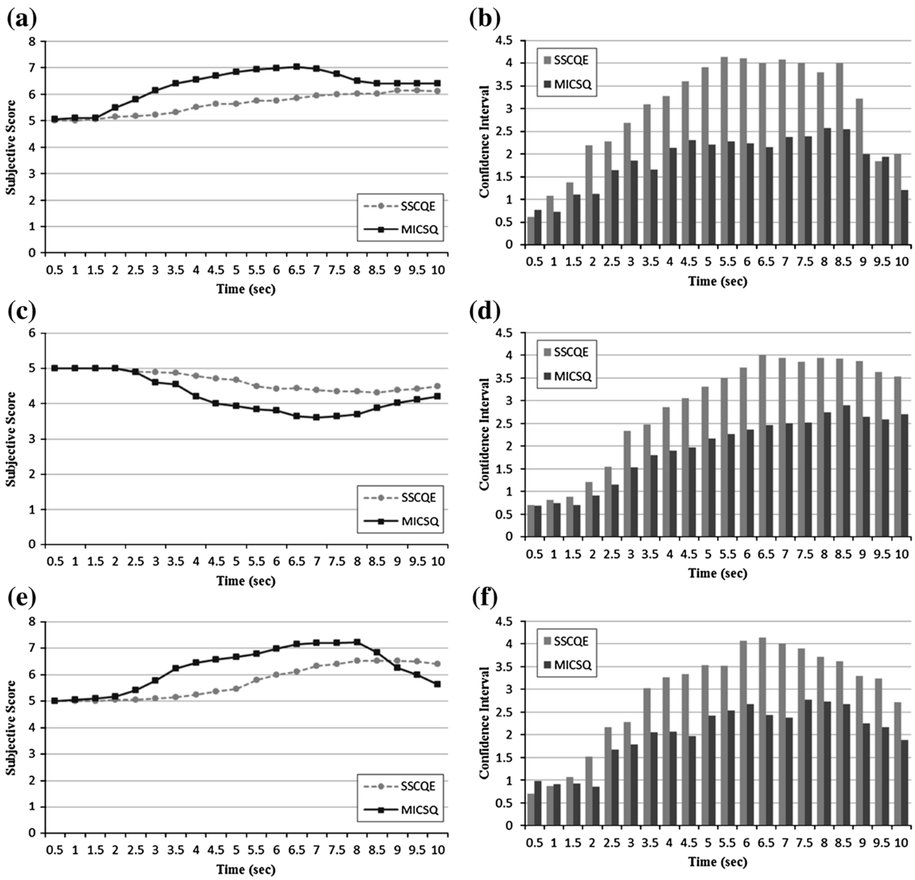

Fig. 9 Subjective scores of **a** 08-02 [EPFL], **c** 11-02 [EPFL], **e** Restaurant1 [IEEE-SA] and their confidential intervals **b** 08-02 [EPFL], **d** 11-02 [EPFL], **f** Restaurant1 [IEEE-SA]

multimodal cues may reduce unfamiliarity relative to the assessment task [19]. In other words, due to sensory reminder prompts, the subject can maintain concentration on the 3D video without distraction from scoring.

5 Conclusions

Unlike 2D scenarios, subjects become deeply involved in 3D visualization. This condition is referred to as an immersion problem in dark space. Therefore, it is difficult to appropriately assess 3D video in an approach similar to that of 2D. This paper has therefore proposed a novel methodology named MICSQ for subjective 3D QoE assessment. Unlike conventional methods, MICSQ utilizes multimodal cues to minimize distractions between viewing and assessment tasks through human/machine interaction. Moreover, to maximize the degree of freedom for assessment, the device interaction between the server and client addresses the diverse 3D viewing environment, even in darkness over wireless network protocols in real time. The implementation of MICSQ based on a mobile platform enables

Table 7 Mean length and standard deviation of CIs

Sequence	Mean		Standard deviation	
	MICSQ	SSCQE	MICSQ	SSCQE
01-02 [EPFL]	1.706	3.010	0.877	1.797
01-05 [EPFL]	1.100	1.594	0.510	0.751
02-02 [EPFL]	1.963	2.857	0.750	1.188
02-05 [EPFL]	2.233	3.364	1.248	1.924
06-02 [EPFL]	1.896	2.540	0.624	1.138
06-05 [EPFL]	1.720	2.243	0.511	0.935
08-02 [EPFL]	1.797	2.870	0.646	1.184
08-05 [EPFL]	1.465	2.382	0.436	0.988
11-02 [EPFL]	1.551	2.454	0.445	0.758
11-05 [EPFL]	1.394	2.182	0.500	0.705
Car1 [IEEE-SA]	1.162	1.745	0.667	1.079
Car2 [IEEE-SA]	0.944	1.324	0.578	0.941
Crosswalk2 [IEEE-SA]	2.014	3.411	1.240	1.443
Libray3 [IEEE-SA]	0.855	1.401	0.437	0.630
Libray4 [IEEE-SA]	2.041	3.114	0.340	0.499
Marathon1 [IEEE-SA]	2.312	3.617	0.877	1.330
Restaurant1 [IEEE-SA]	1.951	2.793	0.676	1.170
Street-lamp1 [IEEE-SA]	1.325	1.998	0.791	1.212
University1 [IEEE-SA]	1.850	2.842	0.582	1.050
University2 [IEEE-SA]	1.534	2.052	0.577	0.934
Average	1.640	2.490	0.666	10.83

multiple users to simultaneously perform a subjective assessment using a commercial tablet PC or smart phone.

With the development of more advanced techniques for 3D signal processing and with the increasing demand for 3D content, this method allow for precise analysis of 3D QoE based on MICSQ, including various aspects of viewer visual comfort.

Acknowledgments This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2015.

References

1. Lambooij, M., IJsselsteijn, W., Fortuin, M., & Heynderickx, I. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science Technology*, *53*(3), 30201.
2. Lee, K., Moorthy, A. K., Lee, S., & Bovik, A. C. (2014). 3D visual activity assessment based on natural scene statistics. *IEEE Transactions on Image Processing*, *23*(1), 450–465.
3. Tam, W. J., Speranza, F., Yano, S., Shimono, K., & Ono, H. (2011). Stereoscopic 3D-TV: Visual comfort. *IEEE Transactions on Broadcasting*, *57*(2), 335–346.
4. Daly, S. J., Held, R. T., & Hoffman, D. M. (2011). Perceptual issues in stereoscopic signal processing. *IEEE Transactions on Broadcasting*, *57*(2), 347–361.
5. Zilly, F., Kluger, J., & Kauff, P. (2011). Production rules for stereo acquisition. *Proceedings of the IEEE*, *99*(4), 590–606.
6. Emoto, M., Niida, T., & Okano, F. (2005). Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television. *IEEE Journal of Display Technology*, *1*(2), 328–340.

7. IEEE P3333.1. (2012). *Standard for the quality assessment of three dimensional displays, 3D contents and 3D devices based on human factors*. IEEE Standards Association.
8. Yang, S., Schlieski, T., Selmins, B., Cooper, S., Doherty, R., Corriveau, P., et al. (2012). Stereoscopic viewing and reported perceived immersion and symptoms. *Optometry and Vision Science*, *89*(7), 1068–1080.
9. Yano, S., Ide, S., Mitsuhashi, T., & Thwaites, H. (2002). A study of visual fatigue and visual comfort for 3D HDTV/HDTV images. *Displays*, *23*(4), 191–201.
10. Saygili, G., Gurler, C. G., & Tekalp, A. M. (2011). Evaluation of asymmetric stereo video coding and rate scaling for adaptive 3D video streaming. *IEEE Transactions on Broadcasting*, *57*(2), 593–601.
11. Zhang, L., Peng, Q., Wang, Q. H., & Wu, X. (2011). Stereoscopic perceptual video coding based on just-noticeable-distortion profile. *IEEE Transactions on Broadcasting*, *57*(2), 572–581.
12. Lambooij, M., IJsselsteijn, W., & Heynderickx, I. (2011). Visual discomfort of 3D TV: Assessment methods and modeling. *Displays*, *32*(4), 209–218.
13. Yang, J. C., Hou, C. P., Zhou, Y., Zhang, Z. Y., & Guo, J. C. (2009). Objective quality assessment method of stereo images. In *3DTV Conference* (pp. 1–4), 4–6 May 2009, Potsdam, Germany.
14. Yamagishi, K., Karam, L., Okamoto, J. Okamoto, & Hayashi, T. (2011). Subjective characteristics for stereoscopic high definition video. In *International workshop on quality of multimedia experience (QoMEX)* (pp. 37–42), 7–9 September 2011, Mechelen, Belgium.
15. Recommendation, I. T. U. R. B. T. (2012). *500-13, Methodology for the subjective assessment of the quality of television pictures*. Geneva: International Telecommunication Union.
16. Ricciardi, S., Nappi, M., Paolino, L., Sebillo, M., Vitiello, G., Gigante, G., et al. (2010). Dependability issues in visual–haptic interfaces. *Journal of Visual Language and Computing*, *21*(1), 33–40.
17. Recommendation, I. T. U. R. B. T. (2002). *500-11, Methodology for the subjective assessment of the quality of television pictures*. Geneva: International Telecommunication Union.
18. Goldmann, L., Simone, F. D., & Ebrahimi, T. (2010). A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video. In *SPIE electronic imaging*, 04 February 2010, San Jose, USA.
19. Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: Harper Collins.

Jiwoo Kang received his B.S. degree in Electrical and Electronic Engineering from Yonsei University, Seoul, Korea in 2011. He is currently working toward the joint M.S and Ph.D. from 2011. His research interests include medical image processing, multimedia processing and GPU programming.

Taewan Kim received his B.S. degree and the M.S. degree in Electrical and Electronic Engineering from Yonsei University, Seoul, Korea in 2008 and 2010, respectively. He is currently working toward the Ph.D. from 2010. His research interests include quality assessment of 2D and 3D image and video, 3D video coding, cross-layer optimization and wireless multimedia communications. He has participated in the IEEE standard working group for 3D quality assessment (IEEE P3333.1). He was accepted in the Samsung Humantech Thesis Prize in 2013.

Sanghoon Lee (M'05–SM'12) received the B.S. in E.E. From Yonsei University in 1989 and the M.S. In E.E. from Korea Advanced Institute of Science and Technology (KAIST) in 1991. From 1991 to 1996, he worked for Korea Telecom. He received his Ph.D. in E.E. from the University of Texas at Austin in 2000. From 1999 to 2002, he worked for Lucent Technologies on 3G wireless and multimedia networks. In March 2003, he joined the faculty of the Department of Electrical and Electronics Engineering, Yonsei University, Seoul, Korea, where he is a Full Professor. He was an Associate Editor of the IEEE Trans. Image Processing (2010–2014). He has been an Associate Editor of IEEE Signal Processing Letters (2014-) and Chair of the IEEE P3333.1 Quality Assessment Working Group (2011-). He currently serves on the IEEE IVMSP Technical Committee (2014-), was Technical Program Co-chair of the International Conference on Information Networking (ICOIN) 2014, and of the Global 3D Forum 2012 and 2013, and was General Chair of the 2013 IEEE IVMSP Workshop. He also served as a special issue Guest Editor of IEEE Trans. Image Processing in 2013, and as Editor of the Journal of Communications and Networks (JCN) (2009–2015). He received a 2015 Yonsei Academic Award from Yonsei University, a 2012 Special Service Award from the IEEE Broadcast Technology Society and a 2013 Special Service Award from the IEEE Signal Procoessing Society. His research interests include image/video quality assessment, medical image processing, cloud computing, sensors and sensor networks, wireless multimedia communications and wireless networks.