Assessment of visually induced motion sickness in immersive videos

Huiyu Duan *, Guangtao Zhai, Xiongkuo Min, Yucheng Zhu, Wei Sun, and Xiaokang Yang

Institute of Image Communication and Network Engineering,
Shanghai Jiao Tong University,
Shanghai, China
{hyuduan, zhaiguangtao, minxiongkuo}@gmail.com, {zyc420, sunguwei, xkyang}@sjtu.edu.cn

Abstract. Immersive videos are important components of virtual reality. However, when watching immersive videos, simulator sickness (SS) sometimes occurs. In addition, when there are visual oscillations (VOs) in immersive videos, visually induced motion sickness (VIMS) will make the videos' quality of experience (QoE) worse. Most of existing VIMS studies under controlled experimental conditions use virtual patterns, e.g., black and white stripes. However, QoE of immersive real scenes are far more complicate than visual patterns. Therefore, in this paper, we mainly concern QoE degradations caused by VIMS in immersive videos with real scenes. To get controlled VOs, we add shake to panoramic videos captured by still camera and conduct subjective experiments among 15 individuals. We find that the velocity (frequency) of VOs, the time we immersed and the scenes' contents will all influence the level of VIMS. And the QoE of immersive videos with real life content will be influenced correspondingly.

Keywords: immersive videos; simulator sickness; visual oscillations; visually induced motion sickness; QoE

1 Introduction

Virtual reality (VR) technology is a new display technology, which will bring real experience to individuals. VR can provide users a new interactive experience. The contents of virtual reality can approximately be divided into three kinds: composed completely by graphics, entirely come from reproduction of real world image and blend graphics from virtual world and images from real world. Immersive videos are video recordings where the scene in every direction is recorded at the same time and they are important components of virtual reality. In some cases, when using head-mounted displays (HMDs) to experience

* This work was supported by the National Science Foundation of China (61422112, 61371146, 61521062, 61527804), National High-tech R&D Program of China (2015AA015905), and Science and Technology Commission of Shanghai Municipality (15DZ0500200).

immersive videos, users may undergo simulator sickness (SS). Usually SS will be slight when immersive videos are shot by still camera. However, when immersive videos are captured by moving camera, visual oscillations (VOs) could lead to visually induced motion sickness (VIMS). Quality of experience (QoE) of immersive videos captured by still camera mainly depends on the quality of videos. Meanwhile QoE of immersive videos captured by moving camera mainly depends on the quality of videos and the level of VIMS. We will mainly discuss VIMS in this paper.

Image and video quality assessment has always been a hot topic in recent years [6, 12, 13, 7, 22]. With the fast development of video technologies, studies of video quality assessment also make much progress [4, 19, 21, 20]. Furthermore, when evaluating a video, Besides video quality, QoE of individuals should be considered as well. Nowadays, video technologies have developed into virtual reality era and virtual reality environment is regarded as a new generation of media of video display. Therefore, it is important to study the QoE of immersive videos.

SS is a byproduct of modern simulator technology and sometimes occurs in high-fidelity simulators. SS tends to happen when users immerse into virtual reality environment which is a new type of stereoscopic high-fidelity simulator. The level of SS is measured by simulator sickness questionnaire (SSQ) [9] in our experiments. QoE of immersive videos with VOs is mainly relevant to the quality of immersive videos and the level of VIMS of immersive videos. Therefore, It is important to introduce the concept of VIMS. Motion sickness is general in passengers and it is definitely an unpleasant experience. Unfortunately, this unpleasant experience does not exclusively affect passengers. A person exposed to visual environments with visual oscillations may also sense motion sickness [8, 15]. This kind of case is often called VIMS. Symptoms of VIMS mainly include dizziness, nausea, fatigue, eye strain etc. Some studies have shown that these symptoms would impair user's visual experience and performance, such as [1, 14].

There are many studies research frequency responses of VIMS. In VIMS studies, frequency of VOs could be controlled in two methods:

- Changing the amplitude of VOs while keeping the rms velocity constant.
- Changing the rms velocity of VOs while keeping the amplitude constant.

There were many past studies such as [2, 3] researching the frequency responses of VIMS. Chen et al. [2] summarised the parameters and findings of VOs in these studies. As shown in [2], the frequency responses of VIMS are influenced by motion axis [11], scene content [17] and other latency [10]. Schrater et al. [16] have shown that human visual systems are more sensitive to velocity than acceleration. [2] concluded that the level of VIMS mainly relevant to rms velocity rather than amplitude. Based on the conclusion above, we design an experiment in which we change the rms velocity while keeping the amplitude constant.

Assessment of QoE is a necessary component of VR technology. Immersion and interactive are two important influence factors of QoE. There are two main factors which will influence the sense of immersion: the quality of VR environment such as resolutions, frame rate. and SS. SS sometimes could also cause VIMS. In this paper, we talk about the effect of SS and VIMS to immersion, which will finally influence the QoE of immersive videos. We conducted a rigorous experiment to obtain subjective scores of VIMS and SS. In this paper, to simulate the VO in virtual environment, we add translation shake in two directions: left-and-right axis and up-and-down axis. We change the velocity of VOs while keeping the amplitude constant. Similarly, time and scenes are also the factors we considered, which will be described detailedly in Section II. By using our results, researchers could study the objective assessment of QoE of immersive videos with VOs easily.

This topic is of great significance and our experimental design could guide subsequent researchers. Firstly, the subjective assessment results of immersive videos with VOs could facilitate the research of objective assessment. In the second place, through the rating of VIMS, we propose to limit the VOs of VR contents under a certain range. We can grade the immersive videos and give a warning when an immersive video may cause uncomfortable VIMS. Meanwhile, the time we immersed in virtual environment should be limited too. The length of this time is relevant to the contents. Last but not least, Our results could direct game development to some extent. Though we use real scenes to research, VIMS of virtual scenes may be similar to that of real scenes.

The rest of this papers is organised as follows. Section 2 describes detailed experimental procedures. Section 3 shows and analyses the experimental results, and make a discussion. Section 4 concludes the paper.

2 Experimental Procedures

In our experiment, there are 10 raw videos. Then we add shake with different levels to them. We conducted an experiment covered 15 subjects, which were healthy college students. There are many items should be noticed when conducting this experiment. Detailed items will be discussed later.

2.1 Shooting and Processing

10 raw video streams are captured by Insta 360 4K Spherical VR Video Camera with resolution of 4096×2048 and file format MPEG-4, which are fit for playing in virtual reality environment. All 10 scenes are shot in one university campus with still camera. To avoid the influence of sound, we eliminate the audio track of the videos. And then 15 seconds' video streams are cut from 10 raw videos. The length of time of each video stream comes from the preliminary experiment, which will be explained detailedly in Section 2.

A total of 15 subjects are exposed to sinusoidal motion VOs along left-and-right axis and up-and-down axis. The methods we used to add shake along left-and-right axis or up-and-down axis are similar. We use up-and-down axis for illustration. Fig. 1 shows the model we used. Usually, we paste a frame of immersive video on a sphere to play. When we watch immersive videos, the image on the sphere could reflect the scenes all around. From Fig. 1, we notice that when the amplitude of VOs is small, we could stimulate VOs through shifting the virtual camera in the sphere. The new frames are stitched on the sphere around the new position of virtual camera. The new latitude of each pixel is computed as follows:

$$\theta = \arctan \frac{r_{xoz}}{y - \Delta y} \quad (y > \Delta y), \tag{1}$$

$$\theta = \pi - \arctan \frac{r_{xoz}}{\Delta y - y} \quad (y < \Delta y), \tag{2}$$

$$\theta = \pi/2 \quad (y = \Delta y). \tag{3}$$

In the expressions above, r_{xoz} is computed as:

$$r_{xoz} = x^2 + z^2, \tag{4}$$

where x, y, z is the coordinate of the pixels on the sphere; Δy is the shift distance of virtual camera. When we get the latitude (θ) of the position of new pixels, the new coordinate of x, y, z could be computed correspondingly.

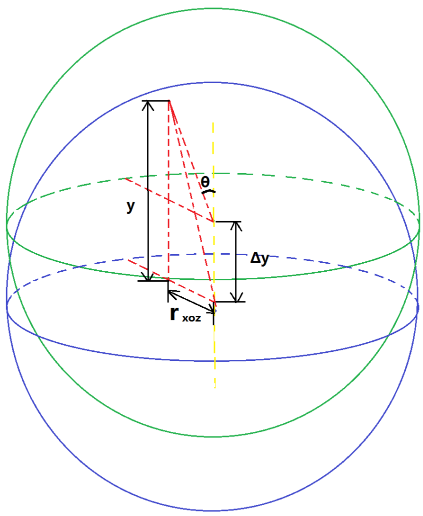

Fig. 1. The model of adding shake we used (y is the distance between the original pixel and the plane xoz, Δy is the shift distance of virtual camera, r_{xoz} is the distance between the pixel and the y axis, θ is the angle of y axis, center of new sphere and the pixel.)

2.2 Preliminary Experiments

Fifteen healthy individuals aged between 20 and 27 years old participated in our experiment and all participants consented to take part in our formal experiments and preliminary experiments. Before the first preliminary experiment, we had investigated the physical conditions of all subjects. All subjects responded that they have normal or corrected-to-normal vision, intact vestibular function. And two people reflected that they have slight travel sickness. All participants were told that once they want to stop they could take off the HMD immediately. In the first preliminary experiment, we discuss the length settings of the time of each video stream. From the feedback, we noticed that the emergency of VIMS usually happen in 15 seconds. Although as time goes, symptoms of VIMS and SS would increase, the variation would be small after 30 seconds. Additional research on this point is needed. In the second preliminary experiment, we concern the velocity settings of shake. Velocity settings of shake are important in our experiments. We use percentage average velocity to describe the velocity settings. It could be computed by:

$$v_{pav} = \frac{\Delta y}{r \cdot t_{\Delta y}} \quad . \tag{5}$$

Several levels of shake are manipulated with one scene. Based on this preliminary experiment, we selected five different velocity levels from them according to rough boundary. The procedure of this preliminary experiment is similar to the formal experiment, and detailed items could be found in Section 3.

2.3 Formal Experiments

In order to investigate the influence of velocity, time and scenes on VIMS when watching immersive videos, we conducted an experiment to obtain faithful data. There are many items should be noticed when conducting the experiments, such as rest time, drink time and so on. There are 110 videos altogether in our experiments. 10 raw videos with no VOs captured from 10 different scenes. 50 videos with VOs along up-and-down axis and another 50 videos with VOs along left-and-right axis. Five different levels of percentage average velocity are listed in the Table 1. The device we used as the virtual reality head-mounted display is HTC Vive. The order of 10 scenes subjects exposed to was randomised. The model we used to rate dizziness and nausea is similar to the 7-point Likert scale used in [5]. However, in our preliminary experiment, Some subjects reported that they only have vertigo symptoms, some other subjects reflected that they only have nausea symptoms, while others said that they have both vertigo symptoms and nausea symptoms. Since our final goal is to evaluate the QoE of immersive videos with VOs, we redefined the 7-point scale: 0 - no symptom, 1 - any slight symptom (dizziness), 2 - mild symptoms but no vertigo and nausea, 3 - mild vertigo or mild nausea, 4 - mild to moderate vertigo or mild to moderate nausea, 5 - moderate vertigo or moderate nausea but can continue and 6 - moderate vertigo or moderate nausea and want to stop. The experiment was terminated when the level of 6 was reached whenever during the trial.

Before each experiment, subjects were requested to eat normal food and keep good spirits. And they were notified the process of experiment and the rating announcements before the first experiment. Detailed rating scale were explained in this procedure, such as the contents of dizziness and the difference between dizziness and vertigo. All subjects were informed that they could quit the experiment at any moment and stop whenever they want. The contents of SSQ were also explained. First of all, subjects were immersed into 10 raw immersive videos with no VOs. At the end of each video, they should give the rating of VIMS. When this group of immersive videos were all played, they were asked to fill in the SSQ. After a rest time, the next group of 10 videos with VOs along up-and-down axis or left-and-right axis were played. Similarly, at the end of each video stream, we recorded the rating score given by each subject. After rating of each group of immersive videos, subjects were asked to fill in the SSQ. And then we provided a break time. The length of break time should be appropriate. In our experiment, at least 30 minutes' break time were needed while according to the feedback of subjects, it could be extended. During the rest time, subjects could drink some water, but other drinks are not permitted. The order of video groups with different velocity levels was random and the videos' order in each group is random too. To ensure the length of time, subjects were demanded to experience the whole video. The trial was conducted with subjects standing. Because when we experience VR environment, we often stand rather than other position so as to look around convenient.

Table 1. velocity settings of the experiment.

level	1	2	3	4	5
percentage average velocity (%/s)	77.333	40.000	25.777	20.000	13.333
frequency (Hz)	1.933	1.000	0.644	0.500	0.333

3 Data Analysis

Before analyzing the data, there is an important circumstance to say. There are 15 individuals involved in our experiment, However, 3 subjects did not complete the trial because the physical cause. Detailed illustration will be discussed in Subsection 3.2.

3.1 VIMS Ratings and Elements Analysis

When discussing the ratings of VIMS, we should find out the main factors which will influence the level of VIMS first. We discovered that when we experience immersive videos with VOs, the velocity(frequency) of VOs, the time we immersed and the scenes' contents will all influence the level of VIMS. The rating of VIMS could be impacted relevant. Therefore, in this subsection, we will evaluate the level of VIMS and analyse the effect of three aspects.

First of all, the Mean opinion Scores(MOS) of each videos was computed as follows:

$$MOS_j = \frac{\sum_{i=1}^{N} m'_{ij}}{N} \quad . \tag{6}$$

In the expression above, N is the number of subjects and m'_{ij} is the score assigned by subject i to video j with various conditions.

Before computing the MOS, there is an important step :Removing outliers. Sometimes, a few subjects will give a score which is far away the mean value. And these outliers should be removed. However, because of the specificity of this

Table 2. Overall MOS of VIMS of 10 scenes with different VOs along up-and-down axis or along left-and-right axis.

orientation		Up-and-down axis					Left-and-right axis				
frequency (Hz)		0.333	0.500	0.644	1.000	2	0.333	0.500	0.644	1.000	2
MOS of ratings	scene1	0.63	1.72	2.27	3.363	4.45	0.63	1.72	2.27	3.363	4.45
	scene2	0.81	1.63	2.63	3.45	4.36	1	1.27	2.09	3.09	4.27
	scene3	1.09	2	2.72	3.27	4.54	1.36	1.45	2.45	3	4.72
	scene4	1.36	2.09	2.63	3.54	4.90	1	1.72	2.63	3.72	4.72
	scene5	1.18	1.72	2.36	3.18	4.72	0.81	1.36	2.27	3.27	4.54
	scene6	0.81	1.81	2.54	3.27	4.81	0.63	1.54	2.18	3.36	4.72
	scene7	1.18	1.90	2.54	3.36	5.09	0.72	1.27	2.45	3.18	5
	scene8	1.63	2.18	2.72	4.09	5.27	1.09	1.81	2.63	3.81	5.09
	scene9	1.54	1.90	2.45	3.90	5.45	0.81	1.54	2.27	3.27	5.09
	scene10	1.27	2	2.45	4	5.09	1	1.36	2.27	3.63	5

experiment, some subjects could give special scores indeed. So we mainly remove outliers produced by unconscious fault. To avoid removing useful data, we would observe overall data of this subject and make sure whether it is natural before removing it. We used 3σ principle to remove outliers.

In this experiment, amplitude was kept a constant of 0.1. Therefore, the variation of frequency is corresponding to the change of velocity. Table 2 shows all MOS of the VIMS. From the table, we noticed that as the increase of frequency (velocity), the level of VIMS would also increase. And it is obvious that the scene could influence the rating of VIMS to some extend. The axis of VOs will also influence the experimental results reflected from the Table 2. Detailed explanation will be given later.

Fig. 2 shows the MOS of one scene ratings. The trend could be observed obviously. Through analyzing the overall situations in the experiment, we found that exposed to immersive videos with VOs along up-and-down axis could cause the larger VIMS than VOs along left-and-right axis, especially in the situation that the velocity of VOs is small. The reason we believe is that when subjects experience the immersive videos with VOs along left-and-right axis, they could waggle with the shake which could alleviate the VIMS.

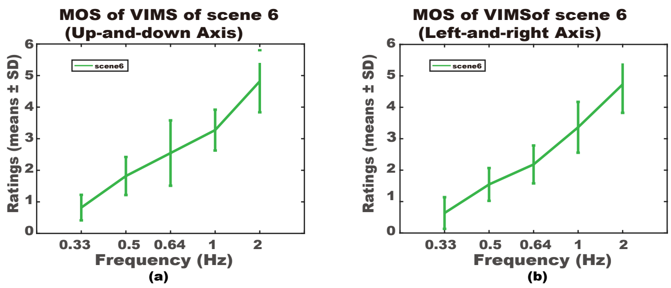

Fig. 2. Rating of scene 6 with different velocity (frequency).

Except the velocity (frequency) and the orientation of VOs, the contents of scenes could also influence the rating of VIMS. There are 10 scenes in our experiment. And through the comparison of the MOSs of 10 scenes, we believe that the brightness of the environment, the movement speed of the surroundings and the distance between the scene and the camera could all influence the level of VIMS. Fig. 3 shows the MOSs of scene 2 and scene 4. From the figure, we noticed that nearly all of the average ratings of scene 2 is less than that of scene 4 whatever the velocity (frequency) or orientation of VO are. The main difference between scene 2 and scene 4 is that the shooting place in scene 4 is football playing while that of scene 2 is pedestrian. The people in two videos have different movement speed. So we could conclude that the scene with faster moving objects will cause more serious VIMS. By similar methods, we derived that as the increase of brightness of immersive videos with VOs, the VIMS will be more serious. Though if there were no VOs, the QoE would increase. The distance between the scene and the camera could also influence the level of VIMS. The level of VIMS will increase when shortening the distance. Additional research is needed in this direction.

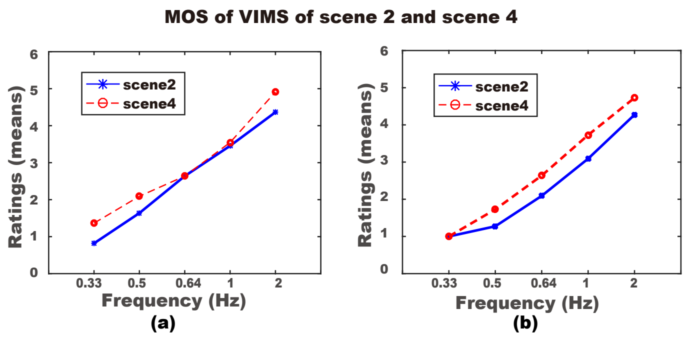

Fig. 3. The comparison between scene 2 and scene 4 with different velocity (frequency)((a)up-and-down axis,(b)left-and-right axis.).

Because the order of the videos in each group is random, in this experiment, we could obtain the data of the relation between VIMS rating and immersion time. As time pass by, the level of VIMS would increase in a short time and then fluctuate around a certain value. The main cause of fluctuation is the contents of the scene. That is an interesting findings, we believe that an abrupt shake which could up to a certain level will cause an abrupt VIMS.

3.2 Exceptional Case

Although we considered many Precautions, some exceptional cases still happen. First of all, there are 12 people completing the trial including one individual with slight travel sickness. 3 subjects did not complete the experiment altogether because of the discomfort. Another subject with slight travel sickness only experienced one group of immersive videos with the minimum VOs and then he said it is unbearable. The situation of other two subjects may be better but they also can not stand finally. Indeed, Motion sickness is common in passengers and has occupied a great proportion [18]. Therefore, we suggest that we should grade the contents in VR environment. The scores given by 3 subjects who did not complete the trial also make sense. Further research is needed.

From the feedback of subjects, some individuals report that they think the VOs along left-and-right axis will cause more VIMS than VOs along up-and-down axis. Some subjects feel that when the velocity (frequency) is too large, they feel less dizziness than lower velocity because they think the VO is untruthful.

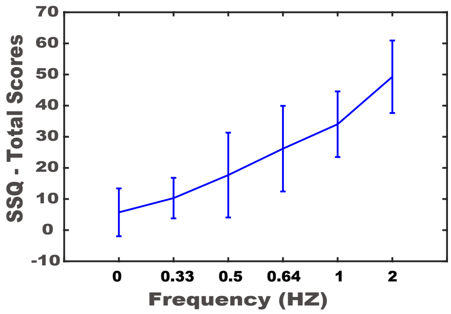

Fig. 4. Frequency responses of SSQ results.

3.3 SSQ scores

Before and after each trial participants were required to fill in an SSQ. As shown in Fig. 4, we noticed that the increases in velocity (frequency) resulted in a significance increase in the SSQ total scores.

4 Conclusion

In this paper, we conducted an experiment to evaluate the QoE of immersive videos with VOs. In order to get videos with controlled VOs, we add shake to panoramic videos captured by still camera through spherical transform. VOs along up-and-down axis or left-and-right axis in our experiment were added. Changing velocity while keeping the amplitude constant will cause different levels of dizziness. We find that the velocity(frequency) of VOs, the time we immersed and the scenes' contents will all influence the level of VIMS when we experience immersive videos. In detail, increases in velocity (frequency) could result in a significance increase in the level of VIMS and SSQ total scores. Similarly, increasing brightness and movement speed or shortening the distance between scene and camera will also increase VIMS. With time pass by, VIMS would increase in a short time and then fluctuate around a certain value.

References

1. Caroux, L., Le Bigot, L., Vibert, N.: Impact of the motion and visual complexity of the background on players' performance in video game-like displays. Ergonomics 56(12), 1863–1876 (2013)
2. Chen, D.J., Bao, B., Zhao, Y., So, R.H.: Visually induced motion sickness when viewing visual oscillations of different frequencies along the fore-and-aft axis: keeping velocity versus amplitude constant. Ergonomics 59(4), 582–590 (2016)

3. Diels, C., Howarth, P.A.: Frequency characteristics of visually induced motion sickness. Human Factors: The Journal of the Human Factors and Ergonomics Society 55(3), 595–604 (2013)
4. Duan, H., Zhai, G., Yang, X., Li, D., Zhu, W.: Ivqad 2017: An immersive video quality assessment database. In: Systems, Signals and Image Processing (IWSSIP), 2017 International Conference on. pp. 1–5. IEEE (2017)
5. Golding, J.F., Kerguelen, M.: A comparison of the nauseogenic potential of low-frequency vertical versus horizontal linear oscillation. Aviation, Space, and Environmental Medicine (1992)
6. Gu, K., Li, L., Lu, H., Min, X., Lin, W.: A fast reliable image quality predictor by fusing micro-and macro-structures. IEEE Trans. Ind. Electron. 64(5), 3903–3912 (2017)
7. Gu, K., Zhai, G., Yang, X., Zhang, W.: Using free energy principle for blind image quality assessment. IEEE Transactions on Multimedia 17(1), 50–63 (2015)
8. Hettinger, L.J., Berbaum, K.S., Kennedy, R.S., Dunlap, W.P., Nolan, M.D.: Vection and simulator sickness. Military Psychology 2(3), 171 (1990)
9. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. The international journal of aviation psychology 3(3), 203–220 (1993)
10. Kinsella, A.: The effect of 0.2 hz and 1.0 hz frequency and 100 ms and 20-100 ms amplitude of latency on simulatory sickness in a head mounted display (2014)
11. Lo, W., So, R.H.: Cybersickness in the presence of scene rotational movements along different axes. Applied ergonomics 32(1), 1–14 (2001)
12. Min, X., Ma, K., Gu, K., Zhai, G., Wang, Z., Lin, W.: Unified blind quality assessment of compressed natural, graphic and screen content images. IEEE Trans. Image Process. (to appear 2017)
13. Min, X., Zhai, G., Gu, K., Fang, Y., Yang, X., Wu, X., Zhou, J., Liu, X.: Blind quality assessment of compressed images via pseudo structural similarity. In: Proc. IEEE Int. Conf. Multimedia and Expo. pp. 1–6 (Jul 2016)
14. Read, J.C., Bohr, I.: User experience while viewing stereoscopic 3d television. Ergonomics 57(8), 1140–1153 (2014)
15. Reason, J.T.: Motion sickness adaptation: a neural mismatch model. Journal of the Royal Society of Medicine 71(11), 819 (1978)
16. Schrater, P.R., Knill, D.C., Simoncelli, E.P.: Mechanisms of visual motion detection. Nature neuroscience 3(1), 64–68 (2000)
17. So, R.H., Ho, A., Lo, W.: A metric to quantify virtual scene movement for the study of cybersickness: Definition, implementation, and verification. Presence: Teleoperators and Virtual Environments 10(2), 193–215 (2001)
18. Turner, M., Griffin, M.J.: Motion sickness in public road transport: passenger behavior and susceptibility. Ergonomics (1999)
19. Zhai, G.: Recent Advances in Image Quality Assessment, pp. 73–97. Springer International Publishing, Cham (2015), https://doi.org/10.1007/978-3-319-10368-6_3
20. Zhai, G., Cai, J., Lin, W., Yang, X., Zhang, W.: Three dimensional scalable video adaptation via user-end perceptual quality assessment. IEEE Transactions on Broadcasting 54(3), 719–727 (2008)
21. Zhai, G., Cai, J., Lin, W., Yang, X., Zhang, W., Etoh, M.: Cross-dimensional perceptual quality assessment for low bit-rate videos. IEEE Transactions on Multimedia 10(7), 1316–1324 (2008)
22. Zhai, G., Wu, X., Yang, X., Lin, W., Zhang, W.: A psychovisual quality metric in free-energy principle. IEEE Transactions on Image Processing 21(1), 41–52 (2012)