SUBJECTIVE QUALITY ASSESSMENT OF MOBILE 3D VIDEOS

Zhiwei Fan, Tingting Jiang

National Engineering Laboratory For Video Technology, School of EECS, Peking University

ABSTRACT

With the development of mobile network and 3D display technology, watching 3D videos on portable devices through mobile network is to be a trend. In this paper, we study the influence of Quantization Parameters (QP), resolution and packet loss on the QoE of mobile 3D videos. Side-by-side formatted 3D video sequences of different QP and resolution, transmitted through networks of different packet loss ratios (PLR) are produced, and two subjective assessment experiments have been made on two portable devices. Each video sequence is tested according to four measures, including clearness, depth perception, comfort level and overall quality. The results show that QP and resolution have a greater impact on mobile 3D videos' clearness than depth perception and that the influence of packet loss is not as stable as the other factors.

***Index Terms*—** mobile, 3D, QoE, transmission, subjective

1. INTRODUCTION

With the development of mobile network, watching videos on mobile or portable devices through mobile network is becoming more and more popular. In the meanwhile, with the development of 3D display technology, 3D videos are widely appreciated, and people want to watch 3D videos more than in the cinema. As some commercially available portable devices currently support auto-stereoscopic display function, watching mobile 3D videos on portable devices is to be the trend. [1] tried to examine users' needs and requirements for mobile 3D television and videos in a psychological way, and users' desires for having 3D content everywhere could be proved. In that case, it is important to find a way to offer mobile 3D videos with high QoE, and getting knowledge of the QoE of mobile 3D videos is a first step.

There have been a number of databases related to 3D videos or mobile videos, while a database specifically designed for both 3D and mobile videos is quite rare. [2–4] each described databases of 3D videos that encompassed a range of environments, lighting conditions, textures, motion, coding conditions, capture parameters, etc. However, none of them was designed for mobile displays and the impairments due to transmission were not introduced. [5] presented a database specifically designed for 2D mobile TV quality assessment, and [6] presented a subjective assessment database of videos with lost slices due to network impairments. But they were both designed for 2D videos. Therefore a new database of both mobile and 3D videos is really in need.

There are a series of artifacts that influence the QoE of mobile 3D videos, caused in different stages of 3D video production and delivery, including content creation, conversion to the desired format, coding/encoding, transmission, and visualization on 3D display [7]. And the artifacts that are most pronounced on portable displays have been analyzed in [8].

In the process of content creation, unnatural disparity between the views in the stereo-pair is the most common and annoying artifact for 3D videos, which leads to bad performance in depth perception. Improper camera geometry and position often cause this problem. [2] offered a database, where sequences of each scene were captured with different camera distances, and a subjective test had been made on it. Unfortunately, the test was made on a computer monitor, and no such experiments on portable displays have been found. Besides, changing the resolution of sequences or the size of displays can introduce unnatural disparity as well [8]. It means that a 3D video with excellent QoE played on large displays like TV screens, may perform badly, especially in depth perception, on small displays like phones. In that case, the impact of displays size and video resolution on QoE of mobile 3D videos, especially on depth perception, is an important research issue, as many videos played on mobile devices are originally created for larger displays, for example 3D movies.

In the process of format conversion and coding/encoding, more artifacts, compression artifacts for instance, are introduced. [9] compared the performance of two video formats (frame-sequential and side-by-side) and two codecs (JMVC and JM) for 3D videos. [10, 11] evaluated four different coding methods and two codec profiles for mobile 3D-TV, finding that Multi View Coding and Video + Depth Coding performed better than the others, and that the high profile could provide the same quality as low profile at a lower bit rate. Compression artifacts are the main artifacts in the process of coding, on which various researches have been done. [12,13] explored the influence of varying depth and compression artifacts on the QoE of mobile 3D video content, showing that compres-

This work is partially supported by NSFC (61390514, 61103087, 91120004, 61210005)

sion artifacts had a larger impact on the QoE than varying depth range. However, the compressed artifacts in the above experiments are usually distinguished by QP, while another important factor impacting the data rate of sequences, which is resolution, has hardly been considered.

For videos displayed on mobile devices, artifacts produced during transmission make a great impact. There are already a number of related works on 2D videos [14, 15], while related works on 3D videos are few, as there are rare real systems specially supporting 3D video delivery. [16] examined the impact of packet loss on 3D videos, however, the impact of packet loss was simplified as frame loss, which did not accord to the fact. [17, 18] made subjective assessment experiments on 3D sequences with different degradations caused by transmission errors such as packet loss and video freeze, while the study targeted a new approach of subjective quality assessment, rather than the effect of transmission errors. In addition, the 3D sequences of the experiments above were all displayed on large 3D displays, even requiring glasses, which meant the displays were neither portable nor auto-stereoscopic, not appropriate for mobile use.

In the process of visualization on 3D display, limitations of the display technology used often cause artifacts. [19, 20] explored different portable devices for 2D videos, and [21] explored auto-stereoscopic displays for mobile devices. Besides, it is possible that a stereo-video stream needs to be rescaled on the receiving device, while rescaling can introduce impairments to disparity as content capture [7].

From above, there are some limitations in the existing researches. For example, the impact of resolution changing on both depth perception and clearness for mobile 3D videos is still unclear, and the existing subjective experiments are seldom conducted on commercially available portable auto-stereoscopic devices. In this paper, we present a subjective research on the influence of QP, resolution and packet loss on the QoE of side-by-side formatted 3D videos. Sequences of different scenes, QP and resolutions, transmitted through networks of different PLR, have been tested on two commercially available two-view auto-stereoscopic devices. The tests are divided into two parts. The first subjective assessment experiment explores the impact of QP and resolution on QoE of mobile 3D sequences, and the second is mainly made to test the influence of packet loss. The experimental results show that without packet loss, there is a positive monotonous relation between videos' resolutions and QoE, and a negative monotonous relation between videos' QP and QoE. However, with packet loss introduced, it is possible that a 3D video of large QP and low resolution performs better than a 3D video of small QP and high resolution, under the same transmission condition. Besides, the artifacts of resolution and QP have a larger impact on mobile 3D videos' clearness than those on depth perception.

This paper is organized as follows. In Section 2, the experimental setup and procedure of the two subjective tests are described. The obtained results of the two experiments and their analysis are presented in Section 3, and the final conclusion is made in Section 4.

(a) Phone (b) MP5 player (c) a Simulated Living Room

Fig. 1. Display devices and environment.

2. SUBJECTIVE ASSESSMENT EXPERIMENT

2.1. Subjective Test 1

The first experiment was made to test the influence of QP and resolution on QoE of mobile 3D videos.

2.1.1. Participants

24 participants (age: 18-25) were randomly selected from the students in the school. 14 of them were male, and 10 were female. They were all screened for normal or corrected to normal visual acuity, color vision and stereo vision. And they were all non-expert viewers with a marginal experience of 3D video viewing.

2.1.2. Test Apparatus and Environment

The experiment was conducted in a simulated living room as shown in Fig. 1, which was closer to a real environment to use a mobile device than a standard room described in ITU-T Rec. P.911 [22] and ITU-R Rec. BT.500 [23]. The simulated living room contained an office chair and a table, together with some other furniture and decorations like a bookshelf and a couple of sofas. Though generally quite quiet, background noise was occasionally audible (e.g., a car driving past).Besides, lighting was simulated as indoor daytime case. During the tests, the participants were asked to sit on the chair in front the table, and they were encouraged to adjust their poses, the angle of the display, and the distances between the display and themselves as they wanted [20]. And two two-view auto-stereoscopic displays were tested in the first experiment. Their parameters are shown as Table 1.

2.1.3. Test Material

The source videos were from [24]. And the database contained 6 scenes as shown in Fig. 2, including various indoor

Table 1. Parameters of the devices used in the experiments.

Device	Marque	Resolution	Size	Displays
Phone	LG P920	800x480	4.3 in	parallax barrier
MP5 player	aigo PMP887	1280x768	8.0 in	parallax barrier

Fig. 2. Six scenes for test 1.

and outdoor scenes with a large variety of colors, textures, moving objects and depth structures. The contents of the 6 scenes were analyzed by evaluating the Spatial Information (SI) and Temporal Information (TI) indexes on the luminance component of each sequence as shown in Fig. 3. Both left and right views of the 3D videos were offered independently. We transformed them into side-by-side format for our experiment, and then they were compressed with different QP and resolutions by software 'x264'. There were 4 different QPs (25, 30, 35, 40), and 6 different resolutions (400x228, 600x342, 800x452, 640x360, 960x540, 1280x720). And the first three resolutions were used for the phone, and the last three resolutions were used for the MP5 player. Therefore, 144 (6 scenes x 4 QP x 6 resolutions) sequences were produced, and each sequence lasted for 10 seconds. All the 24 sequences of the scene 'bike' were selected for training.

2.1.4. *Test Procedure*

All the participants were asked to finish all the tests on a device before starting the tests on the other device, and the order of the two devices was random. The tests on one device were performed as following.

The participants first watched all of the 12 different sequences of scene 'bike' as training. The training session was used for the participants to get familiar with the devices and get knowledge of QoE range of mobile 3D videos displayed on this device. Then the participants were asked to test the rest 5 scenes including 60 sequences, one scene after another. When videos were displayed on the devices, they were rescaled to the resolution of the device as similar as possible, keeping the horizontal-to-vertical proportion, which meant there might exist subsampling on the receiving device. As the QoE of 3D video is a multidimensional concept [25], participants were asked to give scores for 4 measures, including clearness, depth perception, comfort level and overall quality. All the four subjective measures are single items, using 5-point ACR scale. Participants were encouraged to have a short rest after the tests of a scene were over. And on average, it took 50 to 60 minutes for one participant to finish the whole test 1. Besides, all the tests of different scenes, QP and resolutions for each participant were performed in a random order.

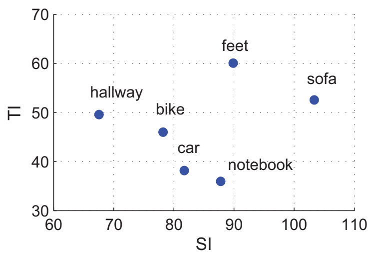


Fig. 3. Spatial Information(SI) and Temporal Information (TI) indexes of the selected video sequences for test 1.

2.2. Subjective Test 2

The second experiment is designed to test the influence of packet loss.

2.2.1. *Participants*

There are 20 participants in total in which 16 participants (age: 18-25) were randomly selected from the students in the school, and 4 participants (age: 32-55) were teachers. 10 of them were male, and 10 were female. They were all screened for normal or corrected to normal visual acuity, color vision and stereo vision. And they were all non-expert viewers with a marginal experience of 3D video viewing.

2.2.2. *Test Apparatus and Environment*

The settings of test apparatus and environment were just the same as that in the first experiment.

2.2.3. *Test Material*

Three scenes were selected from the movie 'Avatar' as test materials, whose screenshots were shown in Fig. 4. And the three scenes contained various motion. Each scene lasted for 10 seconds and was compressed with different QP and resolutions by x264. There were 6 parameter combinations, and the first three were for the phone, while the last three were for the MP5 player as shown in Table 2.

It was worth noting that we encoded the video using multiple slices per frame, and that each slice was packed in one or several packets, where each packet contained roughly the

Fig. 4. Left views of the three scenes for Test 2.

Table 2. Resolution and QP of test videos for Test 2

Parameter	Phone			MP5 player		
Width	600	800	800	960	1280	1280
Height	342	452	452	540	720	720
QP	25	25	30	25	25	30

same number of bytes (approximately 200 bytes per packet) [26]. Then packet loss was introduced to the videos. Different from the previous related works which simulated the transmission by softwares, we managed to record the received video on the client end after transmission through a real lossy link. Our experiment setup contained two computers connected with single ethernet link. One of the computer acted as a server running Darwin Streaming Server, the other as a client playing and recording received video. We used traffic control command from the server, which ran Linux, to set the PLR (0.0%, 0.2%, 0.5%, 0.8%), in which way, the ethernet was able to simulate mobile networks. Therefore, a total of 72 (3 scenes x 6 QP and resolution combinations x 4 PLRs) sequences were recorded. And they were tested for all the observers in test 2.

Another scene lasting for 30 seconds selected from the movie was used for training. In total, eight sequences of different qualities were produced. And the concrete parameters were as Table 3.

2.2.4. Test Procedure

The test procedure of subjective test 2 was similar to test 1. Two devices were tested one by one, and for the test of each device, three scenes were tested one by one after training. And 12 different sequences were tested for one scene on one device. Participants were asked to give scores ranging from 1 to 5 to each sequence on four measures, including clearness, depth perception, comfort level and overall quality. And on average, it took 30 to 40 minutes for one participant to finish the whole test 2.

Table 3. Parameters of the training videos for two devices.

Device	Phone			
Resolution	600x342	600x342	800x452	800x452
QP	25	30	25	30
PLR	0.2%	0.8%	0.0%	0.5%
Device	MP5 players			
Resolution	960x540	960x540	1280x720	1280x720
QP	25	30	25	30
PLR	0.2%	0.8%	0.0%	0.5%

Table 4. SROCC of overall quality against the other measures.

SROCC	Clearness	Depth Perception	Comfort
test1_Phone	0.987	0.968	0.991
test1_MP5	0.988	0.984	0.988
test2_Phone	0.975	0.947	0.948
test2_MP5	0.985	0.983	0.982

3. RESULTS AND ANALYSIS

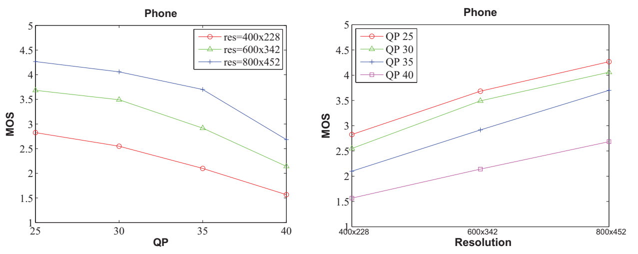
Outlier detection has been conducted on the results of both two experiments [23], and no participant has been discarded. There were four measures tested in the two subjective experiments, including clearness, depth perception, comfort level and overall quality. Table 4 shows the SROCC of the MOS of overall quality against the other three measures, and we can find that they are highly correlated.

For the results of the first experiment, the influence of QP and resolution is analysed, while the results of the second experiment are mainly analysed for the influence of packet loss.

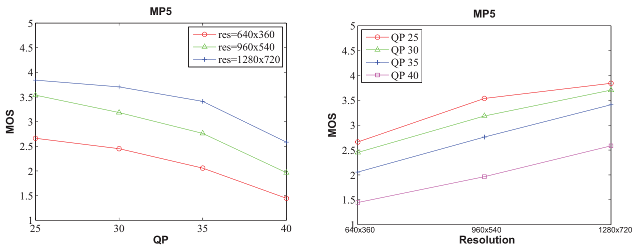
3.1. Subjective test 1

Fig. 5 shows how the average MOS of the five different scenes' overall quality decreases with the QP increasing and the resolution decreasing. And it can be found that the relation between the MOS and the two variables is quite monotonous, without any exceptions.

Unlike clearness, depth perception is a new feature of 3D videos compared with 2D videos. We try to compare the impact of QP and resolution on these two measures. Fig. 6 (a) is a scatter plot of the MOS on clearness (MOSc) against the MOS on depth perception (MOSd) for all the sequences tested in the first test. We can find that the two kinds of MOS are highly linear correlated. The fitting straight-line has a slope of 0.79 smaller than 1. It means that the impairments on the mobile 3D videos' depth perception are more tolerated than those on clearness. According to the fitting straight-line, MOSd is 1.66, when MOSc is 1, and MOSd is 4.88 when MOSc is 5. It means that when the impairments are severe, videos perform better on depth perception than on clearness; when the impairments are gentle, videos perform similarly on the two measures . It shows that the artifacts due to QP and resolution have a larger impact on clearness than on depth

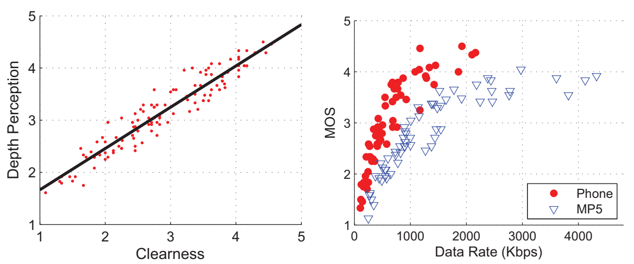


(a) Impact of QP and Resolution on Phone Video



(b) Impact of QP and Resolution on MP5 Video

Fig. 5. The impact of QP and resolution on mobile 3D videos for both two devices.



(a) (b)

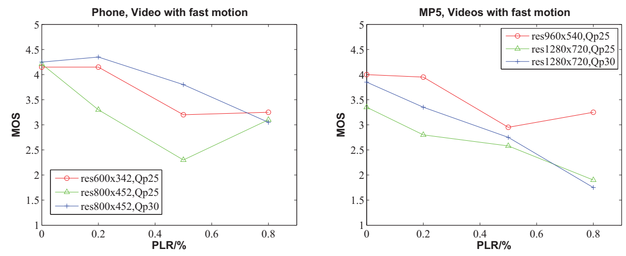
Fig. 6. (a)The scatter plot of the MOS on clearness against the MOS on depth perception. (b)The MOS of overall quality on two devices under different data rates.

perception, although the MOS on the two measures are highly linear correlated. We suspect that it is partly because that participants are more sensitive on the perception of clearness than depth, as the participants are more experienced in watching 2D videos, where only clearness is perceived.

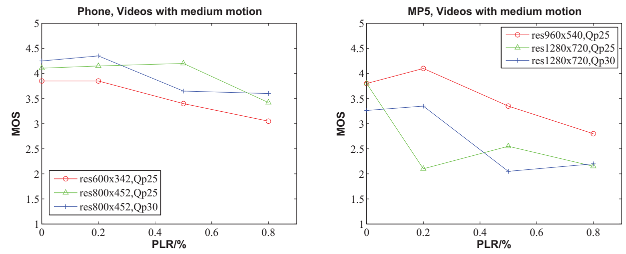
Fig. 6 (b) compares the overall qualities on two devices, and it shows that videos on the phone own higher overall qualities than videos on the MP5 player, when the videos have the same data rate. As the two devices use the same display technology of parallax barrier, the screen's size and production technology should be the main cause for the different performances on two different devices.

3.2. Subjective test 2

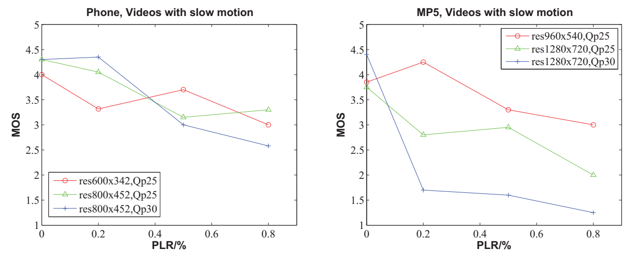
Fig. 7 shows how videos' overall quality changes with the PLR increasing. We can find that the overall quality is on



(a) (b)



(c) (d)



(e) (f)

Fig. 7. The impact of packet loss ratio on overall quality of different videos

a declining trend when PLR increases, while exceptions are not rare. The reason may be that the packet loss is random although the PLR is set, and that a burst of packet loss may lead to a great impact on the QoE.

An interesting but confusing discovery is that the sequences with highest resolution and smallest QP do not perform the best under the same PLR. For example, in Fig. 7 (a), the sequence of highest resolution (800x450) and smallest QP (25) performs the worst, when the PLR is larger than 0.0%. A possible reason may be that the sequence of higher resolution and smaller QP owns a higher data rate, in which case, one slice of the sequence often contains more bytes, and is split into more packets for transmission when the packet size is fixed, which makes the slice easier to be lost.

4. CONCLUSION

The purpose of these two tests is to investigate the effect on QoE of mobile 3D videos when videos are compressed with different QP, subsampled with different resolutions, and transmitted through networks of different PLR. We regard 3D videos' QoE as a multidimensional concept composed of clearness, depth perception, comfort and overall quality. We find that the final MOS on the four measures are highly correlated, and artifacts of QP and resolution changing have a larger impact on clearness than on depth perception. As for packet loss, the impact is not stable, as packet loss is random even with a fixed PLR. But a network of higher PLR still tends to produce more impairments to the QoE of mobile 3D videos. Moreover, there is a discovery that it is possible that the videos with highest resolution and smallest QP do not perform the best under the same PLR.

5. REFERENCES

[1] Dominik Strohmeier, Mandy Weitzel, and Satu Jumisko-Pyykkö, "Use scenarios: mobile 3D television and video," in *IS&T/SPIE Electronic Imaging*, 2009, pp. 72560L–72560L.

[2] Lutz Goldmann, Francesca De Simone, and Touradj Ebrahimi, "A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video," in *IS&T/SPIE Electronic Imaging*, 2010, pp. 75260S–75260S.

[3] Xing Liyuan, You Junyong, T. Ebrahimi, and A. Perkis, "Stereoscopic quality datasets under various test conditions," in *QoMEX*, 2013, pp. 136–141.

[4] M. Urvoy, M. Barkowsky, R. Cousseau, Y. Koudota, V. Ricorde, P. Le Callet, J. Gutierrez, and N. Garcia, "NAMA3DS1-COSPAD1: Subjective video quality assessment database on coding conditions introducing freely available high quality 3D stereoscopic sequences," in *QoMEX*, 2012, pp. 109–114.

[5] Werner Robitza, Yohann Pitrey, Matej Nezveda, Shelley Buchinger, and Helmut Hlavacs, "Made for mobile: a video database designed for mobile television," in *Sixth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, 2012.

[6] N. Staelens, G. Van Wallendael, R. Van de Walle, F. De Turck, and P. Demeester, "High definition H.264/AVC subjective video database for evaluating the influence of slice losses on quality perception," in *QoMEX*, 2013, pp. 130–135.

[7] A. Boev, D. Hollosi, A. Gotchev, and K. Egiazarian, "Classification and simulation of stereoscopic artifacts in mobile 3DTV content," in *IS&T/SPIE Electronic Imaging*, 2009, p-p. 72371F–72371F.

[8] A. Boev, A. Gotchev, and K. Egiazarian, "Stereoscopic artifacts on portable auto-stereoscopic displays: what matters?," *Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, 2009.

[9] K. Yamagishi, L. Karam, J. Okamoto, and T. Hayashi, "Subjective characteristics for stereoscopic high definition video," in *QoMEX*, 2011, pp. 37–42.

[10] Dominik Strohmeier and Gerhard Tech, "Sharp, bright, three-dimensional: open profiling of quality for mobile 3DTV coding methods," in *IS&T/SPIE Electronic Imaging*, 2010, pp. 75420T–75420T.

[11] Dominik Strohmeier and Gerhard Tech, "On comparing different codec profiles of coding methods for mobile 3D television and video," in *Proceedings of the International Conference on 3D Systems and Applications (3DSA10)*, 2010.

[12] Tong Zhu, Lina Karam, and Tuyet-Trang Lam, "Subjective assessment of compressed 3D video," in *Sixth International Workshop on Video Processing and Quality Metrics for Consumer Electronics VPQM*, 2012.

[13] S Jumisko-Pyykko, Tomi Haustola, Atanas Boev, and Atanas Gotchev, "Juxtaposition between compression and depth for stereoscopic image quality on portable auto-stereoscopic display," in *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, 2011, pp. 1–4.

[14] F. De Simone, M. Naccari, M. Tagliasacchi, F. Dufaux, S. Tubaro, and T. Ebrahimi, "Subjective assessment of H.264/AVC video sequences transmitted over a noisy channel," in *QoMEX*, 2009, pp. 204–209.

[15] Fadi Boulos, Benoît Parrein, Patrick Le Callet, David S Hands, et al., "Perceptual effects of packet loss on H. 264/AVC encoded videos," in *Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM)*, 2009.

[16] L. Pinto, J. Carreira, S. Faria, N. Rodrigues, and P. Assuncao, "Subjective quality factors in packet 3D video," in *QoMEX*, 2011, pp. 149–154.

[17] Jesús Gutiérrez, Pablo Pérez, Fernando Jaureguizar, Julián Cabrera, and N Garcia, "Subjective evaluation of transmission errors in IPTV and 3DTV," in *Visual Communications and Image Processing*, 2011, pp. 1–4.

[18] J. Gutierrez, Pe, x, P. rez, F. Jaureguizar, J. Cabrera, and N. Garcia, "Validation of a novel approach to subjective quality evaluation of conventional and 3D broadcasted video services," in *QoMEX*, 2012, pp. 230–235.

[19] A. Khan, Sun Lingfen, J. Fajardo, I. Taboada, F. Liberal, and E. Ifeachor, "Impact of end devices on subjective video quality assessment for QCIF video sequences," in *QoMEX*, 2011, pp. 177–182.

[20] A. Catellier, M. Pinson, W. Ingram, and A. Webster, "Impact of mobile devices and usage location on perceived multimedia quality," in *QoMEX*, 2012, pp. 39–44.

[21] Atanas Boev and Atanas Gotchev, "Comparative study of autostereoscopic displays for mobile devices," in *IS&T/SPIE Electronic Imaging*, 2011, pp. 78810B–78810B.

[22] ITU-T Recommendation P.911, "Subjective audiovisual quality assessment methods for multimedia applications," in *International Telecommunications Union, Tech. Rep*, Geneva, 1999.

[23] ITU-R Recommendation BT.500-13, "Methodology for the subjective assessment of the quality of television pictures," in *International Telecommunications Union, Tech. Rep*, Geneva, 2012.

[24] "3D video quality assessment," `http://mmspg.epfl.ch/3dvqa`.

[25] V. Kulyk, S. Tavakoli, M. Folkesson, K. Brunnstrom, K. Wang, and N. Garcia, "3D video quality assessment with multi-scale subjective method," in *QoMEX*, 2013, pp. 106–111.

[26] Thomas Stockhammer and Miska M Hannuksela, "H. 264/avc video for wireless transmission," *Wireless Communications, IEEE*, vol. 12, no. 4, pp. 6–13, 2005.