Evaluation of Quality of Experience for 3D Videos based on Quality of Service

Ivett Kulik
Department of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary
kulik@tmit.bme.hu

Abstract: Provisioning 3D video stream-based services online in an acceptable quality, even in wireless access environment is a big challenge for Future Internet service providers. Characterizing the necessary Quality of Service requirements is hard, since only a few empirical results are known about the user perceived 3D quality. In this paper a statistical analysis of subjective perception of 3D stereoscopic video Quality of Experience (QoE) are investigated with respect to network level QoS. The network is configured to demonstrate a real environment, thus GPON-based aggregation is used. Our results show characteristics of QoE-QoS relationship in case of 3D video playback. We also tackle the challenge by carrying out GPON-based transport network with IEEE802.11n standard based WiFi access measurements focusing QoE of 3D content. And according to our results we propose cubic fitting function for modeling QoE-QoS relationship in case of throughput degradation.

Keywords: 3D stereoscopic video, Quality of Experience-QoE, Quality of Service-QoS, GPON-based network, WiFi network, Mean Opinion Score-MOS, subjective evaluation

1 Introduction

The Internet has approached an historic turning-point when mobile platforms and applications poised to replace the fixed-host/server model that has dominated since its inception. The existing Internet architecture has been designed for efficient communication but not for real-time data distribution. Exponential growing of smart mobile devices with Internet access and the user needs to be “always connected” definitely indicate that the Internet has become the core mobile communication environment for business, entertainment, educational, social and human interactions.
In the past decades new network architectures and protocols have been proposed sketching the idea of Future Internet. Paul et al. [1] presented a comprehensive survey on the networking research on network architecture for future networks and next generation Internet. The articular network neutrality aspect where users are able to access any web content and to use any applications according to their choice without restrictions or limitations becomes the biggest challenge for Internet Service Providers (FISP).

FISP has to prepare for capability to support multiple types of terminals, hosts and nodes, protocols and applications. The major design goals of FISP networks are: 

- **mobility** as the norm with dynamic host and network mobility at scale;
- **robustness** with respect to intrinsic properties of wireless medium;
- **trustworthiness** in the form of enhanced security and privacy for both mobile networks and wired infrastructure;
- **usability** features such as support for context-aware pervasive mobile services, evolvable network services, manageability and economic viability.

The Future 3D Media Internet generates a significant part of research work recently, which should be designed to overcome current limitations of network architecture, involving content and service mobility, new forms of 3D content provisioning etc. [15][3]. A seamless delivery of 3D video streams means that the provider needs to be able to observe and react quickly on Quality of Service (QoS) problems in transport network and the importance of Quality of Experience (QoE) appears as well. QoE is customer-centric metrics while QoS is network-centric. Human perception of video streams is best characterized in term of QoE which looks at the streaming content from the standpoint of end users. Today, in the fast resolution increasing era, mobile-phone owners commonly watch movie trailers or whole films on their small favorite devices, while customer satisfaction will remain dominant criteria for future applications. Consequently, appropriate QoS support of service providers and satisfactory level of 3D video QoE at the client side provided through the wireless access for mobile handhelds remain big challenge for Future Internet researchers, as well. Investigation of QoE characteristics based on QoS degradation for 3D multimedia contents delivery is in focus recently. The assessment of QoE in multimedia services can be performed either by subjective or objective methodologies [2].

More research subjects had brought into focus the QoE and QoS [3][12] or evaluation of stereoscopic images [4][11][6][10] but more investigations are needed for appropriate QoE provisioning in wireless network based networks. The Gigabit Passive Optical Network (GPON) GPON transport based test-bed with wireless client access is an appropriate representation of an environment for measurements and recent research works in evaluation of QoE for 3D multimedia delivery by means of QoS in Future Internet wireless access scenarios.

This contribution is publishing a few results of subjective tests carried out by participants focusing to describe the relationship between QoE and QoS for 3D
contents delivery in a real network environment. Obviously network level QoS parameters like throughput, delay, jitter and packet loss affect user level QoE parameters. At first we carried out experiments based on subjective testing of 3D video files where 50 participants observed QoE changing due to degradation of QoS parameters. Results of this experiment are published in [15]. We went on our experiments and this contribution shows results of QoE – QoS relationship investigation when one video file was observed in 3D and 2D type of visualization, as well. 40 users watched videos with QoS degradations while jitter increased and throughput decreased. Second part of this paper describes a few results of experiment where 36 participants assessed quality of the 3D video content when the network was a representative combination of GPON-based transport network and IEEE802.11n standard based WiFi access.

The paper is structured as follows. In Section 2 we explain the network environment. Section 3 describes method of measurements Section 4 discusses results in case of GPON environment. Section 5 shows a few results with WiFi network on the client side. Finally this paper is concluded in Section 6.

2 The network environment

Based on the 3D multimedia transport requirements the appropriate test network was planned and realized. Basically, the multimedia server is connected with a broadband and reliable connection and 3D video contents were transferred through the network in unicast mode using TCP transport. Types of encoding and compression affect the demand of bandwidth in case of multimedia content transport. The used average bandwidth could be between 10Mbit/s and 20Mbit/s by stream or more but in case of higher motion level scenes even 40Mb/s throughput are needed. Videos were displayed by the Nvidia Vision Player v1.6.

The GPON based transport network was efficient with 2.5Gbit/s download speed and 1.5Gbit/s upload speed [7] by providing broadband and responsible access to video server with 3D multimedia streams. The whole GPON-based network architecture with wireless sub-networks on client side is shown in Figure 1.

The GPON-based transmission network consist four components: Optical Line Terminal (OLT) on the provider side, Optical Network Terminal (ONT) on the customer side, optical cables for connecting and passive splitters that can split optical signals in split ratios 1:2. The OLT and ONT devices are managed by the Siemens EM-PX manager client. Hardware configuration of the server and clients are shown in the Table 1.
Figure 1 The GPON-based network with WiFi sub-networks for 3D video streams investigation

<table>
<thead>
<tr>
<th>CLIENT</th>
<th>Components</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Intel Core 2 Quad, Q8300, 2.5GHz</td>
<td>Needs: At least Intel Core 2 Duo, or AMD X2 Athlon</td>
</tr>
<tr>
<td>Video-card</td>
<td>NVIDIA GeForce GT 240</td>
<td>Needs: 8 series, 9 series or 200 series NVIDIA video-card</td>
</tr>
<tr>
<td>Memory</td>
<td>4GB RAM</td>
<td></td>
</tr>
<tr>
<td>Spectacles</td>
<td>Nvidia 3D Vision</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SERVER</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motherboard</td>
<td>Asus P5B Deluxe</td>
</tr>
<tr>
<td>Processor</td>
<td>Intel Core 2 Duo, 2.13GHz</td>
</tr>
<tr>
<td>Memory</td>
<td>1 GB RAM</td>
</tr>
</tbody>
</table>

Table 1 Hardware configuration of the client and server

Video server was responsible for 3D and 2D video files storing and sharing what was guaranteed by the VLC program. The WANulator software simulated different Internet conditions such as delay, jitter or packet loss providing the
proper QoS degradation level in the transport network and bandwidth limitation was set Netlimiter.

![Network architecture diagram](image)

**Figure 2** Network architectures of the experiment without and with the WiFi sub-networks

Figure 2 shows network architectures of the experiment for both scenarios: at first when 2D and 3D videos were delivered and watched on the PCs connected directly to the GPON and secondly when 3D video was transferred through the GPON to clients with WiFi 802.11n access to the transport network.

### 3 Method of Measurements

The common practice to estimate user perception from network-level performance criteria is to conduct large experiments in a controlled environment. The QoE can be affected by many factors. Network features which refer to QoS metrics packet loss, delay, jitter, reordering, bandwidth limitation and also multimedia features which include higher levels' specific parameters like: coding, quantization, bit-rate, frame-rate and motion level could have an effect on the QoE [12].

Multimedia sequences (undistorted and distorted contents as well) can be scored by the Mean Opinion Score (MOS) in case of subjective evaluation, which is the core of our experiments. The Mean Opinion Score (MOS) [17] quality scale method used to be applied for voice and video traffic scale (shown in Table 2). Reference
sequence quality can be also grade by MOS for more detailed results but usually only outcome has to be done.

<table>
<thead>
<tr>
<th>Score</th>
<th>Sequence quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Excellent</td>
</tr>
<tr>
<td>4</td>
<td>Good</td>
</tr>
<tr>
<td>3</td>
<td>Regular</td>
</tr>
<tr>
<td>2</td>
<td>Bad</td>
</tr>
<tr>
<td>1</td>
<td>Awful</td>
</tr>
</tbody>
</table>

Based on the first hand experience of our testing [15] we prepared an investigation about QoE-QoS relation not only for 3D video streams but for 2D contents as well. Our goal was to use statistical analysis to obtain more information on the relationship between the degradation of QoS parameters and QoE evaluations.

Table 2 MOS Quality Scale

In both cases (in GPON environment and in WiFi network topology, as well) participants watched a short part of 3D stereoscopic Avatar film with features shown in Table 3 and had to evaluate following questions about quality during video watching focusing on the empirical quality of the video.
1) Rate continuity of the video content.
2) Rate the quality of picture. Did you notice disintegration of picture?
3) How did you assess the 3D experience on the whole?
4) How did you feel conformity between the picture and voice?
5) What was the quality like on the whole?

Table 3 Features of the investigated 3D video

Order of these points was also essential. First 4 points were about QoE from various point of view. And the last one was about QoE on the whole which is usually much more complicated than only recapitulation of first 4 points. We also asked users to weight their answers for the correct statistical analysis. These weights helped to calculate the weighted average for representation of QoE - QoS relationship based on subjective tests.
4 Test results in GPON environment

We gathered some basic demographic information. 40 users (37 men, 3 women, 16 spectacled from them and with average age 22) attended this experiment. They watched the 3D stereoscopic Avatar film trailer mentioned above and the same part of film in the 2D implementation. A short part was enough because the goal was the QoE estimation and no the content assessment [14].

Two types of degradation was made on the 3D and 2D video file as well. And test users scored videos in case of follow scenarios by the MOS:

1) Reference undistorted video files
2) Videos disturbed only by jitter increase
3) Videos disturbed by bandwidth limitation and jitter increase

The value of bandwidth limitation was calculated based on the maximum bandwidth demand which was around 40Mb/s for the 3D content in case of the highest motion level scenes. The mean value of used bandwidth was around 32Mb, so we set the bandwidth threshold to 32 Mb/s which caused throughput limitation. This value was set by the Netlimiter software on the each client.

Value settings of these these scenarios are shown in the Table 3 and Table 4.

<table>
<thead>
<tr>
<th>QoS setting</th>
<th>Type of video</th>
<th>Values refer to every measuring</th>
<th>1. test</th>
<th>2. test</th>
<th>3. test</th>
<th>4. test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitter</td>
<td>2D</td>
<td>9400 packets + 470 burst for jitter; Bandwidth limit. none</td>
<td>Jitter: 100 ms</td>
<td>Jitter: 120 ms</td>
<td>Jitter: 140 ms</td>
<td>Jitter: 160 ms</td>
</tr>
<tr>
<td>Jitter</td>
<td>3D</td>
<td>9400 packets + 470 burst for jitter; Bandwidth limit. none</td>
<td>Jitter: 90 ms</td>
<td>Jitter: 100 ms</td>
<td>Jitter: 120 ms</td>
<td>Jitter: 160 ms</td>
</tr>
</tbody>
</table>

Table 3 Parameters values for jitter degradation

Results of reference tests (undistorted video file watching) showed that people who have watched 3D movie or video before this experiment (36 person) perceived the 3D contents less qualitative than the rest of them (4 person). The average value of 3D experience (3. point in the questionnaire) was 3.83 (almost 4 - good quality) which was very good score on the whole.
Table 4 Parameters values for throughput limitation + jitter

After evaluation of averages we counted the weighted average based on weights answers gathered from users and we could assign one QoE value to every certain value of QoS parameters. If an answer was given a larger weight by the user that meant that this feature (one of 1-5 points above) was more important for the user. Summary of these information are shown in the Table 5.

<table>
<thead>
<tr>
<th>QoS setting</th>
<th>Type of video</th>
<th>Values refer to every measuring</th>
<th>1. test</th>
<th>2. test</th>
<th>3. test</th>
<th>4. test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth limit. + Jitter</td>
<td>2D</td>
<td>9400 packets + 470 burst for jitter; Bandwidth 32 Mb/s</td>
<td>Jitter: 100 ms</td>
<td>Jitter: 120 ms</td>
<td>Jitter: 140 ms</td>
<td>Jitter: 160 ms</td>
</tr>
<tr>
<td></td>
<td>3D</td>
<td>9400 packets + 470 burst for jitter; Bandwidth 32 Mb/s</td>
<td>Jitter: 90 ms</td>
<td>Jitter: 100 ms</td>
<td>Jitter: 120 ms</td>
<td>Jitter: 160 ms</td>
</tr>
</tbody>
</table>

Table 5 Summary of weighted values

We can clearly recognize QoE deterioration based on increase of QoS. Observers watched content on two PCs simultaneously and separately connected to GPON by two WiFi access points. NVPv1.6 player was set up with 440ms de-jittering buffer and it was not changed during the whole experiment.
Figure 3 shows QoE degradation based on jitter increasing by using interpolation lines in case of 2D and 3D video.

Applying the method of least squares we got the next solutions:

- 2D: \(1.41046 \times 10^{-6} x^3 - 0.000558526x^2 + 0.0398079x + 4.88527\) \( (6)\)
- 3D: \(4.19435 \times 10^{-6} x^3 - 0.00116773x^2 + 0.0644745x + 4.22993\) \( (7)\)

QoE-QoS relationship shows cubic correlation and the sensibility is more conspicuous in case of 3D video.

Figure 3 shows confidence interval (CI) of QoE values, where normal distribution is applied and 90% confidence interval and the critical value was calculated for this 90% CI. Lines of averages are plotted with bold lines and margins of CI are plotted with dashed lines. In case of 3D video the CI is more descending.
Jitter value 90ms was the threshold for the 3D video and jitter value 100ms was the threshold for the 2D video when the vision quality was still good without jerkiness and freezing during the watching. The quality rapidly broke down from this point and participants were not satisfied with the quality due to jerkiness and later even freezing picture. This method of evaluation was used in case of jitter increase and throughput limitation in the same time when the threshold values kept on 90ms and 100ms jitter value but fell down rapidly from this point.

5 Test results in WiFi environment

In this experiment 36 participants attended (34 men and 2 women) who study at the Budapest University of Technology and Economics. 18 of them were spectacled and their mean age was 22.14, the youngest student was 20 years while the oldest was 27. 32 participants have watched 3D movies before the tests.

Observers watched content on two PCs simultaneously and separately connected to GPON by two WiFi access points. When people watched 3D stereoscopic content on two PCs simultaneously playing was not fully fluent especially in higher motion level scenes, even in case without any QoS parameter degradation.
in the transport. Simultaneously two wireless configuration is investigated and loaded condition of them significantly effected our measurements which could appear in real network as well. Using WiFi channel-13 caused a medium load, while channel-3 showed an extremely crowded wireless condition.

Figure 5 QoE scores comparison between scenarios with the moderate bandwidth limitation on channel-13: x-axis MOS values in case of bandwidth 40Mb/s and y-axis MOS values in case of bandwidth value 36Mb/s with bandwidth limitation 4Mb/s

Figure 6 QoE scores comparison between scenarios with the high bandwidth limitation on channel-13: x-axis MOS values in case of bandwidth 40Mb/s and y-axis MOS values in case of bandwidth 28Mb/s with bandwidth limitation 12Mb/s
Figure 5 shows linear regression with small deviation between average scores based on observation results in case of small QoS degradation. This means only small differences appeared in scoring between intact and moderately limited playback. In the second case the video was played back with the small QoS degradation, namely bandwidth is limited with 4Mb/s compared to the intact situation. Bandwidth limitation values were calculated on the average demand bandwidth value of 3D stream, which during the 95% of playing time was 32Mb/s except in case of the highest motion level scenes when spine values appeared, exceeded this 32Mb/s value up to 40Mb/s. According to our experiments with respect to the offered load 40Mb/s considered as the highest load in the network, thus considered as intact situation. During the tests bandwidth in the transport is limited. Threshold was set to 36Mb/s when bandwidth limitation was 4Mb/s and so on which caused network QoS degradation during our experiments.

Figure 6 shows comparison of results in intact case and the highest bandwidth limitation setting when bandwidth threshold value was 28Mb/s with bandwidth limitation of 12Mb/s. As we can see the linearity disappeared in this case. When the quality of continuity became unacceptable because of jerkiness and freezing, some participants’ average score still remained above 3 (regular quality). As can be seen on the figure these participants are mostly spectacled and did not assess the bad quality so critically.

Also it is an apparent observation, that only spectacled people scored better the playback with higher bandwidth limitation in both cases depicted in Figure 5 and Figure 6.

In the article [12] the IQX hypothesis is presented which is a natural and generic relationship between QoE and QoS. They demonstrated the feasibility of exponential relationship through a couple of case studies for example measurements results for web browsing in a fast network taken from G.1030. Our experiments show correlation with quadratic and even with cubic model is much better than applying exponential model assuming the limitations of moderate and high crowded channel, channel 3 and channel 13 respectively. The applied models are shown in Figure 7 and Figure 8. This means that QoE-QoS relationship for 3D stereoscopic video playback shows cubic correlation with R square of 0.964 on channel 3.

We can recognize bigger contrast in case of channel 13 shown in Figure 8, where higher QoE were evaluated with better scores at the beginning but from the threshold bandwidth limitation value of 8Mb/s stronger QoE decrease appeared. QoE-QoS relationship also shows cubic correlation with R square 0.993. This difference from the logarithmic approaches found in [18][19] and the correlation model proposed in [12] is caused by 3D video content specifics compared to data centric QoE observations.
In case of services like video transport the continuity is the most significant viewpoint in case of QoE evaluation. We can recognize it from Figure 9 which shows in detail evaluation of QoE degradation caused by bandwidth limitation for each question. From the boxplots the following observations can be made.

1) Rate of continuity was scored the most critically because the highest mean
score was only 2.5 (between regular and bad), which was caused by data sequences stuck in high motion level parts of video. The threshold bandwidth limitation value was 8Mb/s. In case of 10Mb/s and 12Mb/s limitation values the quality of continuity was unacceptable because of jerkiness and freezing occurred during playback.

2) The quality of picture was scored much better than continuity, usually between 4 and 3 (good and regular quality) because blurriness did not appeared during the experiment even in the worst case.

3) Assess of 3D experience on the whole was not so much criticized than the continuity and the best mean score was 3 – good even for 4Mb/s limitation. This point is interesting, because this means people are still accustomed to 2D screening and they are more tolerant in case of 3D quality impairment than in case of video continuity stalling or short jerkiness. And the 3D QoE like depth of picture was not so sensitive to the QoS degradation than the screening continuity.

4) Conformity between picture and voice was scored with the biggest deviation and was acceptable except in the last two scenarios with 10Mb/s and 12Mb/s bandwidth limitation values when due to heavy continuity degradation voice quality also rapidly fell off.

5) The quality of 3D video watching like on the whole was scored with big deviation even in case of no bandwidth degradation. Some people scored it with 4 (good) but some even with 2 (bad) therefore even the best mean score is only under three (less than regular quality), representing the most subjective part of the experiments.

![Boxplots mean scores with deviation for each bandwidth limitation values separately and clustered by questions: 1-continuity, 2-quality of picture, 3-3D experience, 4-conformity between picture and voice, 5-3D video vision quality on the whole](image)

Figure 9 Boxplots mean scores with deviation for each bandwidth limitation values separately and clustered by questions: 1-continuity, 2-quality of picture, 3-3D experience, 4-conformity between picture and voice, 5-3D video vision quality on the whole
Consequently, participants were the most sensitive to the fluidness and continuity of scenes and the 3D experience was less important when they evaluate the subjective video quality.

6 Conclusion

Within this paper a complex subjective test method of QoE investigation of 3D stereoscopic video files has been introduced. The GPON network with its capacity was suitable for efficient transport of these contents even in unicast mode.

At first the relationship between QoE and QoS was shown based on the gathered results for 3D stereoscopic multimedia content compared with results of the 2D implementation of the same content. Evaluation of data was carried out by the IBM Statistics software. QoS metrics like jitter and throughput limitation disturbance were demonstrated by tests results which showed cubic correlation in both cases. The quality of 3D presentation like depth impression is influenced by multimedia features as well and dynamic, lots of movement sections in video are more sensitive to the QoS degradation.

In Future Internet research one significant concept is to obtain network neutrality by extending of heterogeneity in network architecture and service support. In the second part of this article are presented some results of subjective test results of QoE-QoS relationship character with 36 participants in suitable environment representing common future environment, GPON-based transport network + WiFi sub-network based on IEEE 802.11n in the 2.4GHz band on the client side. Characteristics of QoE degradation were shown and analyzed on gathered MOS scores of participant experiments. The good quality guarantee is more complex in case of WiFi access because QoE is influenced by the nature of wireless technology (like bandwidth limitation of multiple clients or channel interferences) and by QoS level in transport network, as well. Robustness of 3D content, QoS degradation and limitation of WiFi network together cause stronger QoE deterioration on the client side.

The goal was to compare gathered experiments with exponential fitting function based on IQX hypothesis [12] in case of vision quality investigation of 3D stereoscopic video delivery through WiFi network. Applying cubic fitting function to measurements results leads to better correlation with $R^2$ values 0.964 and 0.993 than exponential fitting function with $R^2$ values 0.765 and 0.745 in the investigated bandwidth limitation interval. This different result was caused by 3D video content delivery service investigation and QoE assess by users subjectively.
Our results show that the fluidness and permanent continuity of video-streams is the most important aspect for good QoE. Primary importance of QoE investigation in wireless network environment is came to the front due to the worldwide growing of video-streams presentation on smart small mobile devices and results of this contribution could be helpful for ISPs in case of 3D based multimedia services providing.

In future more measurements and investigation are needed with various QoS disturbances like delay, jitter and packet loss in wireless environment and with explicit channel parameters like the WiFi Access Category, Beacon time, Max. Agg. Frames as long as the resulting A-MPDU fits within the configured TXOP limit, etc. considered. The goal is the mathematical modeling of functional relationship between QoE and QoS metrics which is needed for optimal solution of 3D stereoscopic video contents delivery with appropriate display quality.

Acknowledgement

This work was supported by NKTH-OTKA grant CNK77802.

References


[8] Zilly F., Müller M., Eisert P., Kauff P. „The Stereoscopic Analyzer – An Image-Based Assistance Tool for Stereo Shooting and 3D Production” ICIP 2010 – IEEE International Conference, Hong Kong, September 2010


