Challenges in Smartphone-based Video QoE Assessment "In-the-wild"

Canberk Temiz

880809-4679

Submitted in fulfillment of the requirements for the degree of Bachelor of Science

Towards Department of Communication Systems

September, 2014

Department of Communication Systems

Blekinge Tekniska Högskola

Supervisor: Selim Ickin

Examiner:

Abstract

Quality of Experience (QoE) of applications is highly important since a poor application performance and a low QoE may cause a user to stop using it (e.g., shift to a “better” one), or change the network operator due to misinterpretation of the underlying reason. Assessing video QoE in-the-wild is highly challenging, and the challenges have to be identified earlier to overcome and prevent any circumstances that new comers, e.g., new researchers in the area, may experience. In this thesis, a set of challenges in assessing video QoE in the wild, and user behavior during video streaming, are presented. Ultimate goal of this thesis is to provide newcomers and researchers a practical view on our conducted video QoE studies with real users in the wild. Moreover, the adapted solutions for the challenges are presented. We used a video streaming tool, VLQoE that is able to display and in parallel quantify temporal impairments. A user study consists of a user survey, face-to-face interview, and a data collection via VLQoE. In this thesis, only the results of the first two are presented and we discuss on the challenges.

Table of Contents

1 Introduction

Video streaming is a popular application on the mobile terminals. Everyday people stream video through Internet to get information or just to have fun [6]. It's quite entertaining to stream a video while travelling or waiting for somebody [7]. This leads to a general question; what is the perceived quality of the user on the streamed video? Has the quality of the video satisfied the user? One keyword of this thesis is QoE, and that is stands for a measurement used to determine how well that network or service is satisfying the end user's requirements. QoE (Quality of Experience) is by definition "is the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user's personality and current state" according to [1]. It is important to understand end-user QoE according to [8], user reactions and satisfactions related to the particular product such as a smartphone, service, or an application. Essential point is to provide a good enough quality to the end-user.

QoE of applications is highly important as the low QoE may cause a user to discontinue using the application (e.g., shift to a "better" one), or change the network operator due to misinterpretation of the underlying reason. Thus, QoE of applications has become so important to consider both for the application developers and also at the network operator. According to Laghari et al. "Today, humans are quality meters, therefore it is crucial to consider their needs, perceptions, and expectations with respect to a particular product, service, or application. Because human pleasure carries a great value" [2]. In order to understand the QoE on smartphone-based real-time video applications, it is important to investigate it on the smartphone user interface, which is the layer that a user actually sees and perceives the app quality [3]. Based on the QoE management strategies[24], it is suggested for an application to adapt to use minimal network bandwidth and in parallel maintain QoE by knowing the parameters that users are most sensitive to. The ITU-T (International Telecommunication Union Standardization Sector) has defined QoE as the "overall acceptability of an application or service, as perceived subjectively by the end user" [4]. The recent research suggests that in order to truly meet end-user requirements and expectations, QoS mechanisms may need to be complemented with more user-centric approach [5]. In this thesis work, user study is conducted to quantify the influence temporal impairments of a video stream on the end-user perceived quality. A video is shown to the users that streamed through 3G and the users are asked to give their opinions. In this thesis, the user behavior during a video stream, how the perceived quality influenced user in realistic user environment are presented.

QoE on mobile terminals, especially on smartphones have always been a challenge as there is a lack of ground truth method to follow. We tried to stick to the well-known QoE assessment methodologies in the state of the art, however due to the higher degree of freedom in the smartphone context, we have experienced various challenges. The user's feedback on the experiment methodology is often neglected; however this is extremely important in order to improve the future user study methodologies on the smartphone. In other words, the users in the user studies are expected to be fully focused on the given task. Obtained feedback signifies whether the study is boring, or enjoying, and also it gives insights on how the experiment method can be improved. The result of this study basically aims to help newcomers in the field planning and designing these kinds of in the wild smartphone based studies. Therefore, while providing a practical view on how to conduct user studies, it is also aimed to provide an overview on a set of challenges

that has been overcome in the user studies. Results and solutions of the challenges would help newcomers, and this helps them to overcome situations when they face similar cases.

Our user study consists three main phases. First phase comprises a user survey, which has a specific questionnaire form. This form was used to collect demographics such as age, occupation, as well as background knowledge. This helps us to understand experience of users and the interactions in-between smartphone. Second phase is the video streaming phase, in which after we train the user by explaining the user about the test procedure, we enable them to stream video via 3G. Users were asked to rate the video and five-scale MOS (Mean Opinion Score) [10, 11] values to quantify QoE where '1' refers to the "bad" and '5' to "excellent" in video streaming quality. Third, a short face-to-face interview is conducted with the users. This leads to learn more about qualitative feedback received from the users during the whole experiment and novel user feelings about the study. The users were also asked to give an overall score in a 5-level ACR scale. We Therefore this study allows us to closely experience video QoE on the terminals. To stream the sample video, we used VLC player [12] with a new functionality called VLQoE [3]. The tool enables us to record user interface, application-level, network-level metrics.

This thesis is structured as follows. In Section 2, we review the Background and Related Work on video QoE, and basically discuss on the existing video streaming protocols and the MOS. We present the user study methodology in Section 3. The results are given in Section 4. This involves the challenges experienced during the user study. The experiences faced can help to guide future smartphone based user studies in the wild. The thesis is concluded in Section 5.

The research questions in this thesis are as follows:

- **What are the subjective perceived qualities of experience on real time video streaming on the smartphone?**
- **What are the most important challenges in video QoE assessment in the wild?**

2 Background and Related Work

The previous works on QoE mainly focus on the audio-visual quality. A. Catellier *et al. [14]* shows how subject's perception of multimedia quality differs when content is viewed on different mobile devices. In our work, we used the same smartphone to avoid these kinds of problems. Ickin *et al.* [13] investigates a similar way to understand perceived video quality. Instead of one video, three different movies of different qualities from three different sources are examined and presented to users. A. Catellier *et al. [14]* also presented a three level rating scale such as; video quality levels ("excellent", "fair", and "bad"). Moreover, it is remarked that when a video screen of the same size is used, statistically similar MOS is obtained. It is also indicated that the device screen plays a huge role on the perceived QoE. Furthermore, to compare multimedia content or communication systems regarding quality from an end-user perspective, subjective quality assessment methodologies have been developed [15,16]. Ultimate aim of these methodologies is to obtain user opinion ratings that are comparable across different studies and a common recommendation is usage of the MOS. On the other hand, Tobias *et al.* [17] expresses that averaging opinion scores, e.g., MOS, might smooth out the extreme e.g., very high or very low, user ratings, which are important to consider as well. Thus, they emphasize the importance of considering the Standard deviation of Opinion Scores (SOS). Martin *et al.* shows different aspects of QoE estimations, they defined Pseudo-Subjective Quality

Assessment (PSQA) a generic methodology for estimating QoE for media applications [18]. Furthermore, Tao *et al.* [19] addresses the problem of perceptual quality measurement on individual coded video frames distortions caused by both lossy source coding and lossy transmission. The results of the distortions are examined and as a result it creates a block-wise-full-reference quality metric for a single distorted video frame.

The most important difference is that this thesis comprises a user study conducted in order to present the challenges faced during the video quality assessment with respect the temporal impairments on smartphone. This thesis presents a novelty of all the experienced challenges during the user studies obtained.

2.1 Video Streaming Protocols

Nowadays, one of the most network and energy demanding applications is network-based video streaming [21]. For example, it might require high amount of bandwidth, processing power, good enough screen brightness. Related to this manner, a research involving a subjective study with the focus on inter-picture time is conducted. In this study, we stream a sample video via 3G on Android smartphone using the VLQoE tool [3]. Video clips on the mobile terminals can be streamed via two main application layer protocols: Real Time Streaming Protocol (RTSP) [22] or Hypertext Transfer Protocol (HTTP) [23]. For instance RTSP is better for long movies, because there is no file stored. Thus, RTSP is implemented for live multimedia streaming. On the other hand HTTP is used for multimedia streaming and it commonly works based on progressive download. File is downloaded to user's device, but can't start playing before it is partially downloaded. The characteristics of each protocol with respect to the network layer, such as the burst size and duration, are supposed as different parameters, and this difference is expected to influence the application performance, and eventually the end-user perceived QoE [3]. Thus, we included both type of streaming in our study.

2.2 MOS (Mean Opinion Score)

MOS (Mean Opinion Score) is used in subjective assessment of the user-perceived QoE throughout the thesis. MOS is a test method that has been used for decades in telephony networks to obtain the human user's view of the quality from 1 (bad) to 5 (excellent). It is the arithmetic mean of all the individual scores, and MOS has a range from 1 (bad) to 5 (excellent). The users were asked to rate video quality according to MOS ratings.

3 Methodology

In this section, the description of the VLQoE tool, the user survey, the video streaming experiment, the face-to-face interviews, and the experiment testbed used in the video streaming experiment are given.

3.1 VLQoE Tool Description

Functionalities are implemented within the VLQoE application. The most important part of this study belongs to VLQoE user interface. As shown in Figure 1 below, there are 5 buttons, which are related with the MOS rating scale (1,2,3,4,5). In addition, there is a "freeze" button as well beside the rating buttons. There is an interface provided for user to choose the streaming protocols (RTSP/HTTP).

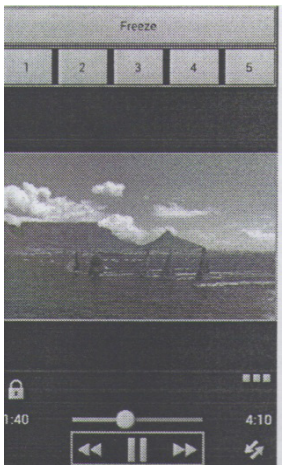

Figure 1: Snapshot from the VLQoE player. [3]

All experiments are conducted on an Android based smartphone terminal.

3.2 User Survey (Questionnaire form)

The user survey is employed during the study. In order to improve the representativeness of our study, we tried to find as many participants as possible. Before the video streaming session, we asked them to fill out questionnaire form. A questionnaire form is filled by the user (see Appendix for the complete form) and then the experiment procedure is explained to the users. And participants were asked to fulfill the forms. We tried to be clear about the survey as much as possible before we apply it to the users. For each user, we collected the demographics and background information based on their daily life interaction with video streaming on the smartphones. It reveals general background about each user. There were in total 30 participants (5 female) in the user study. Most of them are students at BTH. Amongst all the subjects, 28 of them are within 21-30 years age range. And the rest, S4 and S7 were within 31-40 years range.

The background information of the subjects, *S*, is presented in Table 1. The user's phone type are given in rows 1 and 2. Figure 2 presents the distribution on which interface the users stream video on their own smartphones in daily life. 13 subjects are streaming via only WiFi, 5 subjects using streaming via 3G and 6 subjects stated that they use both streaming sources while watching video on the smartphone in daily life. Among all of the participants 4 subjects claim that they use local streaming. Moreover, 23 subjects watch a video on the smartphone at least five minutes per day. Users experience freezes mostly while streaming via 3G on the smartphones as given in Figure 3.

3.3 Video Streaming Experiment

During the experiment what we want from the user is to rate for the picture quality by the means of video jerkiness and freezes, *i.e.*, temporal impairment. They were asked to rate if receiving pictures/frames are in well sequence. So, while the users watch the video stream, they were asked to score it approximately every 10 – 15 seconds. Five level MOS rating scale was used in order to rate the temporal quality. Furthermore, the user was asked to press "freeze" button to indicate an evidence of a visual freeze on the display any time during the video play out.

Phone type: Iphone	S1, S2, S4, S5, S6, S8, S13, S14, S17, S20, S21, S26, S28
Phone type: Android	S3, S7, S9, S10, S11, S12, S15, S16, S18, S19, S22, S24, S25, S27

Table 1: Summary of common features of users

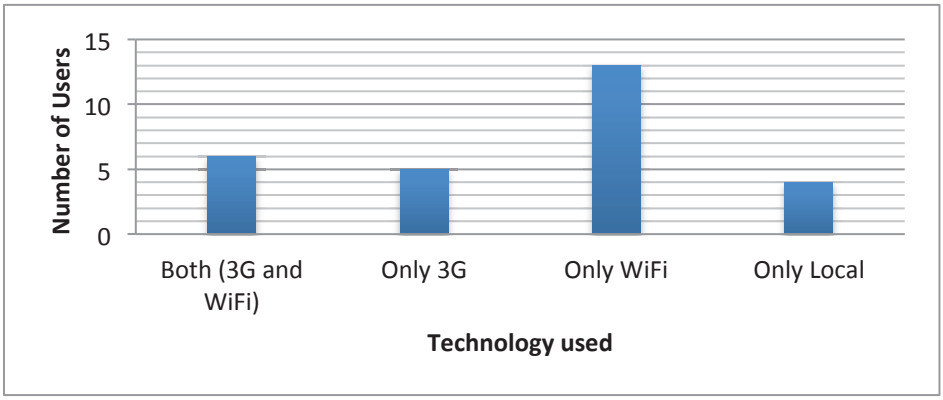

Figure 2: Often used interface for video streaming

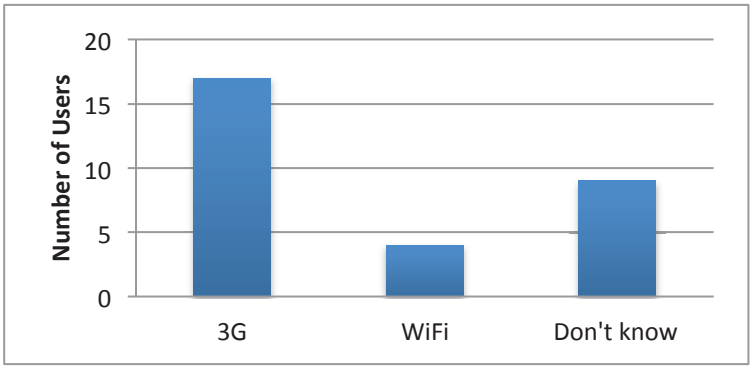

Figure 3: Often experienced freezes per wireless access technology

The VLQoE user interface helps to record user rating during the experiment. Each user first watched RTSP-based video and then the HTTP based video. We tried to conduct experiments in such conditions that the users might experience in the daily life. Finally we conducted 60 studies as each user had to watch 2 videos. In the beginning of each user session, there was a “training session”. The aim is to teach the procedure to the users and make them get used to the smartphone, the user interface, e.g., video application and buttons and also experience a baseline video quality. To make it equal, the users streamed same video first with RTSP and then with HTTP and the videos are displayed with 196 x 117 pixels resolution. The reason for using two different streaming protocols is to improve the representation of the video streaming dataset. We do not aim to compare the two protocols, but this is scheduled for future work. The video was approximately 250 seconds long. Additionally the focus is kept on visual temporal impairments not the audio. So we switched off audio in order to enable the users to focus only on the visual impairments. Thus, this let users to

focus only on visual impairments. And the screen brightness is set to constant value for all user experiments. Five subjects watched the video at home, and 25 subjects at the university. All of the experiments are conducted indoor. In addition, all subjects watched the video while sitting.

3.4 User Interview

We conducted a face-to-face interview with the users after they watched the two videos, i.e., one with the RTSP and the other with HTTP. We collected the overall MOS score based on the two videos, and also asked them regarding the freezes in the two videos. In addition, we asked them to provide feedback on the study itself including the content of the video, background experience in video on the smartphone, willingness to attend the study, user interface of VLQoE, and repetition of the same video.

3.5 Experiment Testbed

VLQoE tool had been used to record user rating during the experiment. We used the same video for the RTSP and HTTP part. The tool that we used in the experiments was installed to Samsung Galaxy S (version Android 4.2.2). The video was located at a dedicated Ubuntu streaming server, which is located in BTH university campus. And then it was streamed to the smartphone via RTSP and HTTP protocols as given in Figure 4.

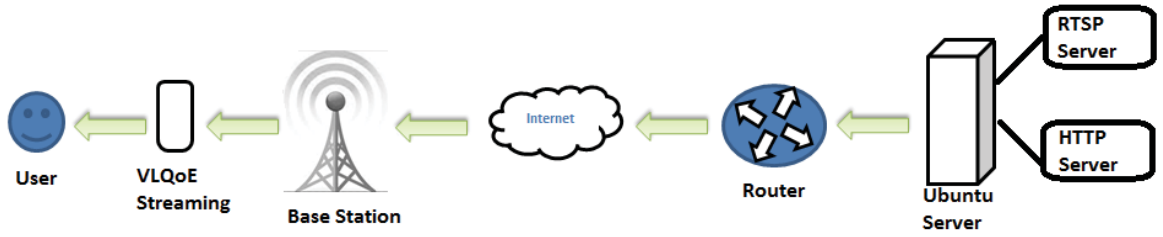

Figure 4: Experiment Testbed [3]

4 Results

In Section 4.1, we present the user reflection on the study obtained via face-to-face interview. Then, the overall challenges experienced during the study are summarized in Section 4.2.

4.1 User's Reflection on the Study

Subject #1, (S1) was an undergraduate student from Karlskrona/Sweden. She was so **kind** and didn't reject us. She was a bit **hesitant** while she was rating the video. She was having a **hard time to press the rating buttons**. Also, she mentioned, "second video was better", which is HTTP-based.

S2 was a male participant from Karlskrona. We met him while he was **having his coffee**. He was an undergraduate student of Management. In addition he was so **excited to involve**. Although he commented that the **video length** was **a bit long**, he was determined to finish it.

S3 was a female participant from Telecommunication department of Karlskrona/Sweden. It was easy for her to adjust herself into the experiment. At the end of the session, she stated that "it didn't freeze a lot, and generally it was fine."

S4 was a male, Swedish M.Sc Student from Karlskrona/Sweden. Although he was **not willing to involve,** because of his lectures, than he is convinced when he heard the survey is just for **10 minutes**. Telecommunication and technology was **not his field of interest**. Additionally, he would like to make a comment about the repetition of the same video twice and asked "Am I going to **watch the same thing one more time**? ".

S5 was an undergraduate student from BTH/Sweden. He would like to participate in this study before I asked him, which indicates that he was enthusiastic to participate in the study. Furthermore, he was **not interacting** well with the user interface. In parallel, his rating speed was very low.

S6 was an undergraduate student from BTH/Sweden. She was so busy at the first time before we start to explain the survey. She is a nurse from health department. At the beginning of the RTSP video, she was **hesitant** to rate it, than she **got used to it**, and **begin to rate more often**. So at the first moments of the video, she missed to rate some freezes and bad pictures. Moreover, she was **happy to help** and stated "it was good feeling to being a part of such studies."

S7 is a worker in Sweden. He is addicted to the technology. He **prefers his smartphone to his personal computer**. He didn't have any problem adapting to the user interface. And he commented, "It is a **nice video** that you choose already, its **worth to watch**." However he was experiencing some freezes on the video, which he thinks is **disturbing**. In addition he stated that, "there are a lot of things happening in the video, this makes it easy to freeze." This indicates that he has some technical knowledge or experience on the video streaming applications.

S8 is an undergraduate student in Sweden. Before the session, he said "I'm **not much a technical person**, but I would like to be in a study." He was **not hesitating to rate** the video when it comes to bad pictures. However it was challenging to conduct a study to this subject. He was **only focusing on the bad quality** pictures. This indicates that not very bad quality can be tolerable, and also supports the well known fact that "No news is good news" While answering the questions on the survey, he doesn't know which one has he experiences more freezes during the video when he watches via Internet.

S9 is an undergraduate male from Karlskrona/Sweden. He was the most concentrated one amongst the whole participants. It is observed that the freezes made him loose his focus. He stated "So, now you need my eyes for observing this, I'm **happy to involve** in this project." In addition, he enjoyed what he watched and he even made couple of whistles.

S10 is a M.Sc. student of Blekinge Tekniska Högskola/Sweden. He was very **focused** on what he was doing and he remained **silent** till the end of the video. He commented on some pictures that how much he likes the quality.

S11 is an Undergraduate student from department of Computer Science. He was willing to involve in the study. The most often watches the videos through the local storage. He stated that **HTTP streamed video had better pictures quality** than RTSP video.

S12 is a M.Sc. Student of the BTH/Sweden from the department of Telecommunication. On the session the subject got **bored** for seeing the **same video twice.** Moreover he was **happy to help** a person who studies the same department with himself.

S13 is a M. Sc. Student from Karlskrona/Sweden. He was enthusiastic about participating in the study.

S14 is a male participant from Karlskrona/Sweden. He is an undergraduate student of Engineering Department. It was a short session, he didn't comment about the survey and the experiment. He was **unfocused and slacker on the second video**. And he said that HTTP video had more freezes.

S15 is a M. Sc. Student from Karlskrona/Sweden. However, he had some hard times to interact with our experiment smartphone interface. He thinks RTSP video stream was "slower", he claimed that he had more freezes there, and moreover he believes freezes are acceptable even though it affects the quality.

S16 is a M. Sc. Student from Karlskrona/Sweden. After the experiment, he would like to comment on freezes. And he said "**freezes are more noticeable than blurring images**". But generally he **liked** the quality of both.

S17 is an undergraduate student from Department of Telecommunication from Karlskrona/Sweden. He was **willing to participate** in this study. He stated HTTP video was **easy to watch**, because he calims **RTSP video was lagging more than HTTP.**

S18 is a M. Sc. Student from Karlskrona/Sweden. Despite, he was not shy to rate the quality and indicate the freezes. In addition the smartphone was charging during the experiment. He stated that HTTP is streaming was very bad. He had hard times to choose a rating for the bad pictures. He was **glad to involve** in the study.

S19 is a M. Sc. Student from Karlskrona/Sweden. He lost his focus on the HTTP video, because of the **repetition of same video**. Furthermore, his HTTP video due to **initial delay** started late. He stated **HTTP video had more fluent pictures**.

S20 is a M. Sc. Student from Computer Department of Karlskrona/Sweden.

S21 is a M. Sc. Student from Karlskrona/Sweden. Before starting to this session, I had **hard time to convince** him. First, he didn't want to spend time on it. It's told that it's going to be only ten minutes, then he was convinced. Before the HTTP session started, he thought study had finished. All in all, he **wasn't happy with the video durations** of the videos.

S22 is a M. Sc. Student of Karlskrona/Sweden. He was willing to involve in this study. There were no freezes on the RTSP session, and he was a bit shy to rate in the RTSP session. He added some comments during the session such as; it didn't freeze much, and moreover, in the beginning of HTTP video, it was **laggy** but afterwards **it settled down**. And the user used interaction buttons more on the HTTP video.

S23 is a M. Sc. Student from Karlskrona/Sweden. As told so, he was concentrated during the whole two sessions. HTTP video was freezes more than RTSP version. He mentioned, only the beginning and the ending moments of the video was **spotty** on the RTSP.

S24 is a M. Sc. Student of Karlskrona/Sweden. The user thinks **RTSP video had more freezes than HTTP**. And moreover she likes the quality of the HTTP video.

S25 is a M. Sc. Student of Karlskrona/Sweden. In addition she is now a smartphone user. User is told to be focus on the. However, in summary, user was satisfied from the video quality.

S26 is a M. Sc. Student from Karlskrona/Sweden. He is very kind to me and **respectful** to study. However, this didn't affect the participant. On the video session, he faced more freezes during HTTP video, and he said "there is no lag on RTSP video, I like the smoothness". After the video sessions he mentioned that he was glad to participat.

S27 is a M. Sc. Student of Karlskrona/Sweden. Despite that he was about to have his lunch, he wanted to participate. When he would like to stream a video on the smartphone, he decision depends on **length of the video**, if it is a short video, he prefers 3G, and for the long ones he is connecting to the WiFi. In addition he very serious to what he was doing, very nice usage of freeze button. He didn't miss any freezes at all. His opinion was fair for both of the videos. He remarked that both videos are almost in same quality.

S28 is a current undergraduate student of Karlskrona/Sweden. In fact this participant would not like to see any freezes or lags because he was having fun while watching the sample video. He was great **fan of sailing**, he remarked that "what a good choice of video" during the experiment. He preferred HTTP video, and bad pictures, freezes on RTSP were the reason behind it. In summary he was **happy** to watch a video, which comes from his field of interest, and he was glad to help to another student of BTH.

S29 is a M. Sc. Student of BTH/Sweden. He accepted our survey offer. Despite of the freezes on the first video, he **didn't use the freeze button**.

S30 is a M. Sc. Student from Karlskrona/Sweden. First of all he didn't use the freeze button at all, despite having freezes. He had some comment during the video session such as; "after I rate it, it doesn't come any sound from smartphone", and he also remarked that freezes that happen on the smartphone depends on the signals which comes from a 3G operator, and how we are going to rate it? In addition he thinks there was no difference between RTSP and HTTP videos. On the whole, he prefers RTSP version of video, he said that it had better quality. He had more experienced freezes during the HTTP video.

The distribution of the overall mean opinion scores (MOS) of the videos after the video streaming is finished is given in Figure 5. There were no MOS values with 1's and 2's; and the dominant MOS was 4. Twelve of the users especially stated during the interview that they were happy to participate in this study. It is clear that majority of the users can distinguish the quality per wireless access technology.

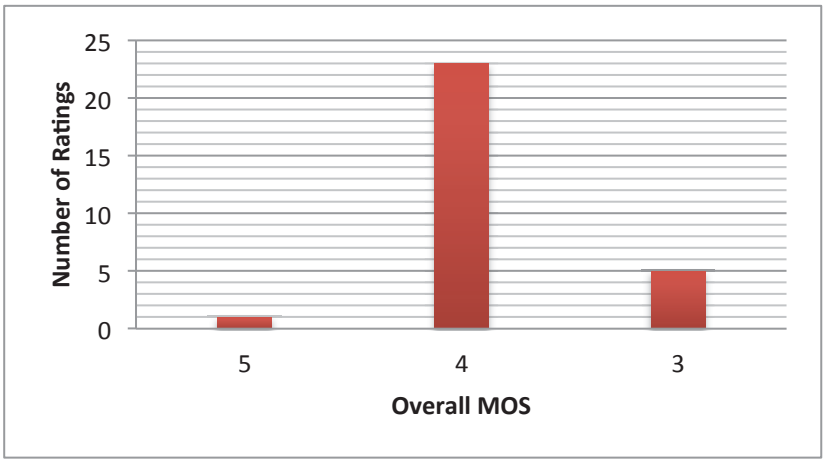

Figure 5: Overall MOS values received during the interview

4.2 Qualitative Feedback on Video Impairments

In addition, we investigate temporal frame quality however some users gave us feedback about the spatial frame quality. And the most used words among spatial impairment and temporal impairment are shown in Table 2.

Spatial Impairment	Temporal Impairment
"Blurring"	"Freezes"," jerkiness"," slower", "settled-down"

Table 2: Spatial and Temporal Impairments

This section summarizes the user feedback regarding with temporal and spatial impairments. In the tables below, users' feedbacks about the video section are presented and matched with the users' comments. When a user comments about "freeze", "slower" or "settled-down", the feedback is related to the temporal impairment. Feedback such as "blurring" relate to the spatial impairments during the video streaming. The subjects whose feedbacks are related to the temporal and spatial impairments are given in Table 3 and Table 4, respectively. 11 users explicitly commented about the temporal impairments and 3 users complained about the spatial impairments.

	Freezes	Slower	Settled-down
Temporal	S2, S3, S9, S14, S17, S23, S29	S15, S18	S4, S22

Table 3: Temporal Impairment

	Blurring
Spatial	S7, S16, S30

Table4: Spatial Impairment

4.3 Challenges during "In-the-wild" Studies.

As shown in the Table 5, the challenges that we have identified are classified into seven categories.

Challenge	Subjects
Length of the video	S2, S21, S27
Repetition of video	S4, S12, S19
Lost data	S19
Unwilling users	S4, S14, S5, S12, S21
User interface	S1, S16, S29
Rating only bad quality	S8

Table 5: Summary of challenges

Unwilling users: It is indeed a challenge to find users for the study. The users wouldn't like to spend their time to this work, and that narrows the users participation. There were five unwilling users, but they were convinced after the length of the study and the importance of his/her contribution were told.

Study duration: Length of the video is one factor that extends the duration of the study. Before each session, 16 users asked how long does it going to take to finish the survey? If they are not comfortable with the duration, they say they don't want to spend time for the survey. For that reason some users decline the invitation. The ones that were pleased with the content of the video did not complain about the length. In total 3 users complained about the study duration.

Repetition of video: The most challenging situation was showing the same video twice. Users become bored in the second video. Therefore the focus, which they keep for the first part, is reducing on the second video. Amongst all users 3 of them were annoyed by the repetition of the same video.

User interface: It takes some time for users to understand what exactly they will need to do during video session, even though it had explained clearly before the study. User interface of the video player needs to be well explained before hand. After the first video session, most of the users learn exactly what needs to be done. Therefore, participants become more aware on the second video. It's hard to obtain this conscious on the first video. More detailed and clear information was given to participants before the sessions in order to limit this problem.

Rating only bad quality: Instead of MOS rating scale, some users used their self-scales. For instance, when they face with bad picture, they directly go for the lowest score "1". Next, when they perceive even worse picture or freezes right after the previous one, there is no score lower than 1. Therefore users might miss some bad qualities and it is getting away from being a fair score table for that particular user. It is vital to understand how the number scheme works in MOS. This is avoided during the training session before the study.

In contrast, there were positive feedback; some users were willing to help science and participate in the study such as S6, S9, S10, S26. In addition, subject S7 and S28 explicitly stated that they enjoyed the video content.

5 Conclusion

This thesis presented the feedback obtained regarding the user studies itself, while assessing the video QoE in user's natural environments. The results obtained focuses on the QoE of users on the video streaming quality with respect to the temporal impairments. We present the user study methodology that consists of user survey, experiment phase (*i.e.*, video streaming and collecting user QoE), and the user interview. In addition, this thesis highlights the obtained challenges. First we gave an overview of user behavior during the video phase and next, we presented challenges that faced during the study. The key challenges in smartphone based video QoE studies are related to the length and repetition of the video. Choosing a shorter video with an interesting content, might also indirectly decrease the number of unwilling users. The ultimate aim is to provide practical view on such studies, and help new researchers in the area to reflect on their own studies.

In order to find more users, some social media platforms might be used. Therefore, user's participation would highly increase and more data can be obtained after all. The number of participants, the variety of context and different kinds of users, which has different aspect of the perceived quality, will increase the diversity in the user profile. This way, the quality of the data can be more representative. Such studies would also enable to assess the quality of the conducted studies and to objectively evaluate the results collected on the future studies. However, in those remote studies, the reliability of the results might be an issue as it is not possible to interview the subjects.

We also recommend that it might be more motivating for users to participate in the user study to enable subjects to choose their own video instead of the presented "sample video". At least video could be more entertaining to the user if it interests the users. That method would decrease the number of bored users, and might improve the quality of the results. On the other hand, we had to have full control on the video quality. Other alternative could be to provide users a set of prerecorded videos with various contents and genre. However, then the user studies would not be much realistic.

References

[1] Qualinet White Paper on Definitions of Quality of Experience, in Proceedings of the Output version of the Dagstuhl seminar 12181, Version 1.1, Dagstuhl, June 2012. Available at: http://www.academia.edu/3001209/Qualinet_White_Paper_on_Definitions_of_Quality_of_Experience

[2] K.Laghari, N-Crespi, K. Connelly (2012) "Toward total quality of experience: a QoE model in communication ecosystem,"IEEE Communication Magazine, vol. 50, no. 4, pp.58-65. April 2012

[3] SelimIckin, Markus Fiedler, KatarzynaWac, Patrik Arlos, Canberk Temiz, Khadija Mkocha. VLQoE: Video Quality of Experience Instrumentation on the Smartphone, Multimedia Tools and Applications. Springer US. April, 2014. DOI=10.1007/s11042-014-1919-0.

[4] International Telecommunication Union ,ITU-T Recommendation P.10/G.100 Series P: Telephone Transmission Quality, Telephone Installations, Local Line Networks. July, 2008

[5] Sabina Barakovic, Lea Skorin-Kapov. Survey and Challenges of QoE Management Issues in Wireless Networks. Journal of Computing Networks and Communications Volume 2013. December, 2012. DOI=10.1155/2013/165146

[6] Dr. Yipeng Zhou. Analysis of Movie Replication and Benefits of Coding in P2P VoD.The Chinese University of Hong Kong. August, 2012.

[7] Petteri Repo, Kaarina Hyvonen, Mika Pantzar, Päivi Timonen .HICSS '04 Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 4 - Volume 4January, 2004.

[8] Verdolini, A.; Petrangeli, S. A smartphone agent for qoe evaluation and user classification over mobile networks. *Quality of Multimedia Experience (QoMEX), 2013 Fifth International Workshop on,* On page(s): 230 – 235. July, 2013. DOI=10.1109/QoMEX.2013.6603242

[9] Human-Computer Interaction. Wikipedia. Available at: http://en.wikipedia.org/wiki/Human%E2%80%93computer_interaction

[10] ITU-R Rec. BS.1534-1.Method for the Subjective Assessment of Intermediate Quality Level of Coding Systems. 2005.

[11] International Telecommunication Union, ITU-T Rec. E.800. Definitions of Terms Related to Quality of Service. 2008

[12] VLC Media Player for Android. Available at: http://www.videolan.org/vlc/download-android.html. Accessed March 2013.

[13] Selim Ickin, Lucjan Janowski, Katarzyna Wac, Markus Fiedler. Studying the challenges in assessing the perceived quality of mobile-phone based video. International Workshop on Quality of Multimedia Experience (QoMEX), Australia. July, 2012. DOI=10.1109/QoMEX.2012.6263853

[14] A. Catellier, Pinson, M. ; Ingram, W. ; Webster, A. Impact of mobile devices and usage location on perceived multimedia quality. Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on. July, 2012. DOI=10.1109/QoMEX.2012.6263834

[15] International Telecommunication Union, ITU-T P.800.1Mean Opinion Score (MOS) Terminology. March, 2003

[16] International Telecommunication Union. ITU-T P.910.Subjective video quality assessment methods for multimedia applications. 2008

[17] Tobias Hossfeld, Raimund Schatz, Sebastian Egger. SOS:The MOS is not enough. Quality of Multimedia Experience (QoMEX), 2011 Third International Workshop on. September, 2011. DOI=10.1109/QoMEX.211.6065690

[18] Martin Varela, Jukka-Pekka Laulajainen. QoE-Driven Mobility Management-Integrating the Users' Quality Perception into Network-Level Decision Making. Quality of Multimedia Experience

(QoMEX), 2011 Third International Workshop on. September, 2011. DOI=10.1109/QoMEX.2011.6065703.

[19] Tao Liu, Hua Yang, Stein A., Yao Wang. Perceptual quality measurement of video frames affected by both packet losses and coding artifacts. Quality of Multimedia Experience, 2009.QoMEx 2009.International Workshop on. July, 2009. DOI= 10.1109/QoMEX.2009.5246949.

[20] Mattia Gustarini, Selim Ickin, Katarzyna Wac. Evaluation of Challenges in Human Subject Studies In-the-wild Using Subjects' Personal Smartphones.UbiComp'13, Zurich, Switzerland. 2013.

[21] Joonho Choi, Myungsik Yoo, Mukherjee B. Efficient Video-on-Demand Streaming for Broadband Access Networks. Optical Communications and Networking, IEEE/OSA Journal of (Volume: 2 , Issue: 1). January, 2010. DOI= 10.1364/JOCN.2.000038

[22] RFC2326: Real Time Streaming Protocol (RTSP). Available at: http://www.ietf.org/rfc/rfc2326.txt. Accessed September 2013.

[23] HTTP Live Streaming. draft-pantos-http-live-streaming-13. Version.13. April, 2014. Available at: http://tools.ietf.org/pdf/draft-pantos-http-live-streaming -02.pdf. Accessed September 2013.

[24] Florence Agboma, Antonio Liotta. Quality of experience management in mobile content delivery systems

Appendix

A. User Questionnaire

Your age is between...
21-30
31-40
41-50
51-60
Your occupation...
Academic Staff
Undergraduate Student
M.Sc. Student
Other (Please specify)
How long have been using a smartphone since (start year)?
2009
2010
2011
2012
If you watch videos that are streamed through Internet on the smartphone, I stream through...
WiFi
3G
Local Storage
Please choose a nickname that you will use in the study...
Please type the full brand name of your smartphone.
In order for us to contact you, please write your email address below.
Gender
Female
Male
Do you want to participate in this study.
Yes
No
How long do you watch a video via 3G/WiFi on the smartphone?
I watch, but no more than 2 mins per day
5 mins per day
10 mins per day
30 mins per day
1 hour per day
More than 1 hour per day
Never
I everyday watch a video either via 3G or WiFi on the smartphone
True
False
I often experience more freezes during the video when I watch ait via...
3G
WiFi
I don’t know