The Impact of Video Encoding Parameters and Game Type on QoE for Cloud Gaming: a Case Study using the Steam Platform

Abstract—In this paper we analyze the impact of different game types and video adaptation strategies on Quality of Experience (QoE) in cloud gaming. We focus on image quality and frame rate as parameters to be manipulated in the video encoder in different ways in order to meet bandwidth constraints of the network link. Our hypothesis is that under certain bandwidth conditions, maximization of the QoE for players will be achieved with different combinations of video encoding parameters (i.e., frame rate and image quality) in dependence of the game type. We test our hypothesis through a series of subjective tests on the streaming solution of the Steam In-home streaming platform . For subjective tests we focus on players with previous gaming experience. Obtained results may be used to drive effective QoE optimization strategies at the cloud gaming server.

I. INTRODUCTION

Game streaming or cloud gaming is a service that allows on-demand streaming of game content from the server to the client. There are three main cloud gaming concepts: 1) streaming 3D graphics, 2) streaming video, and 3) hybrid approaches. In this paper we will focus on the streaming video approach which is implemented in multiple commercial applications while the other two approaches are still in the experimental phase [1], [2], [3].

Cloud gaming has several benefits: it reduces client hardware requirements by a large margin, enables playing on various hardware and software platforms, and alleviates the need for customized game versions on different client platforms. On the other hand, online game play is a highly interactive service with strict real-time requirements. In the streaming video approach, the server executes the game logic, renders the virtual 3D scene, encodes the video, and sends the video stream to the client in real time. The client is responsible for decoding the video and capturing the player’s commands. Hence, video streaming in cloud gaming, as in other interactive online applications, requires continuous game content to be presented to end users without any interruptions in content flow. Meeting latency requirements becomes very challenging (e.g., under 150ms of RTT is needed for good quality for First Person Shooter games [4]), with the need to calculate game state, render the virtual scene, and encode/decode the video stream. There is not enough time for the video encoder to optimize the bandwidth size of the sent video stream. Moreover, conventional methods for diminishing the effects of poor network conditions on streaming media (such as buffering data for display) are not applicable in the context of cloud gaming. Consequently, the impact of heterogeneous and variable network conditions on the end user QoE of cloud gaming has lately been addressed by various researchers [5], [6], [7], [8].

As compared to “traditional” online gaming, game streaming results in a significant increase in the network requirements necessary to secure a high level of Quality of Experience (QoE) [9], with traffic requirements increasing up to two or three orders of magnitude. For example, World of Warcraft, a popular Massively Multiplayer Online Role-Playing Game (MMORPG), has a bandwidth requirement of approximately 55 kbps for the most complex in-game situations (with many players and Non-Player Characters (NPCs) participating) [10], while games from one of the cloud gaming platforms - OnLive, can require up to 5.6 Mbps [11].

With available network resources varying over time, subject to issues such as varying access network conditions or a varying number of players accessing a bottleneck link, there is a need for efficient and dynamic service adaptation strategies on the game server to meet different bandwidth availabilities. The problem we are addressing in this paper is determining how to best adapt the video encoding parameters of the game video stream in light of decreased bandwidth availability, while maximizing the end user QoE. Furthermore, we aim to investigate whether games of different types need unique combinations of video streaming parameters for QoE maximization. Adaptation of streaming parameters based on the bandwidth constraints which can arise in the network can be grouped into two major categories: reduction of the frame rate and reduction of the image quality. Recent studies conducted by Hong et al. [12] have made significant contributions in identifying optimal adaptation strategies for ongoing cloud gaming sessions. Further, the case of in-home streaming and Valve’s commercial Steam platform (intended for streaming games from a powerful PC in a LAN to other weaker devices such as laptops or tablets) confirms that the solution to this problem has still not been found. In particular, in the first iterations of Steam’s In-home streaming, with reduction of bandwidth the frame rate of the stream was reduced, while in more recent iterations only the image quality is reduced while the frame rate is kept at a fixed 60 frames per second (fps).

In this paper we report on an empirical user study examining the impact of frame rate and image quality settings under bandwidth constraints on the end user QoE in the case of game streaming. We further aim to answer whether or not this impact is linked with different game genres. In particular, we perform tests in a controlled lab environment using Steam’s streaming platform [13]. Further, while a number of previous studies have been conducted using the open source GamingAnywhere platform [14], [15], [16], we have opted to use Steam in order to conduct test using both commercial software, and also to enable comparison of different game streaming platforms.

Double-blind submission
We have conducted tests involving two distinct games, a popular First Person Shooter game known as *Serious Sam 3*, and an adventure platform game known as *Bastion*. We vary image quality and frame rate settings under fixed bandwidth constraints to evaluate whether there is a difference between these games and whether different adaptation strategies should be applied for each game when aiming to maximize QoE under bandwidth restrictions. Obtained results have shown that the impact of frame rate reduction on QoE is dependent of the game genre. Furthermore, we aim to model QoE as a weighted linear combination of QoE features, namely perceived graphics quality and perceived game fluidity (referring to the perception of the smoothness in the rendering of the virtual scene) for the two different game types that have been addressed.

The paper is organized as follows. In Section II we give a brief overview of related studies. Section III outlines the test methodology used for conducting empirical studies. Obtained results and proposed QoE models are presented in Section IV, while Section V presents concluding remarks and directions for future studies.

II. RELATED WORK

Previous research has shown that the network traffic generated by different cloud gaming platforms differs significantly (between 2 and 14 Mbit/s) [15]. Further, several works [15], [17], [18] have shown that different games offered by a cloud gaming provider have different network traffic characteristics (e.g. bandwidth requirements for OnLive have been proven to vary more than double [11]). The results reported in [19] analyze the traffic characteristics of 18 games delivered via a cloud gaming platform and differing in type, perspective, and input characteristics. Also, it has been established that the game type has an influence on the reported QoE under the same network conditions [20], and that cloud gaming systems adapt their service according to the available bandwidth [18]. However, an open research question which remains is how to adjust the game streaming parameters under certain bandwidth constraints to maximize users’ QoE, and to what extent this adaptation strategy should vary between games of different types/genres.

With regards to service adaptation, various video adaptation techniques have only as of very recently been studied in order to improve perceptual quality. For example, the concept of a Game Attention Model has been introduced in [21] for the purpose of reducing bitrate, whereby different macro blocks in a game frame are encoded with different qualities based on their estimated importance from a player’s perspective. In [22], the authors identify different objective (rendering, video encoding, mobile network) and subjective factors (graphics quality, video quality, perceived response time) influencing the QoE of cloud mobile rendering applications. The impacts of encoding and network parameters on QoE were addressed in their previous work [23].

In their recent work, Hong et al. [12] address the issue of optimal adaptation strategies for ongoing cloud gaming sessions. Following empirical evidence collected via a large-scale crowdsourced user study and involving subjective scores for three different games, the results provide guidelines on adapting video bitrate and frame rate in light of bandwidth reductions. Gamer experience is studied in terms of the implications of different games and video configurations (bitrate and frame rate) on overall MOS scores, graphics quality scores, and interactivity scores. A key finding is that for high bit rates, higher frame rates lead to better overall scores, while for lower bitrates, higher frame rates lead to overall lower scores (attributed to degraded graphics quality in the case of there being more video frames to encode). In our study, we address the same research question, however with the following differences with respect to [12]: (1) we use the commercial Steam platform for test purposes rather than the GA platform, (2) we conduct tests in a controlled lab environment involving only highly skilled players (n=15), (3) we collect ratings of overall QoE, perceived graphics quality, and perceived fluidity, and (4) we conduct tests involving an FPS game and an action role playing game.

With respect to QoE modeling, while previous work has modeled cloud gaming QoE in terms of network performance parameters [5], [16], [24], we are also interested in modeling QoE as a function of perceived QoE features. In a detailed gaming taxonomy proposed by Möller et al. [25], the following quality features are identified: interaction quality (also linked to playability), playing quality (addressing game learnability and intuitivity), aesthetic aspects, and overall player experience. Quality features have also been addressed by Hong et al. [12], whereby the authors present overall QoE scores as weighted linear combinations of perceived graphics quality and perceived interactivity scores, showing very little variation in weight coefficients across three different game types. Our tests further look into this issue, with results showing that there are indeed differences when it comes to different game genres.

III. METHODOLOGY

The study consisted of an hour and a half long gaming sessions that were conducted in a laboratory environment as shown in Figure 1. As previously stated, we opted to use Valve’s Steam In-Home streaming platform in order to test high quality commercial grade video streaming quality. As we were focused on assessing the impacts of different bitrates and fps settings on QoE, we excluded further loss and delay degradations and conducted tests in a controlled network environment. Steam desktop clients were installed on all three PCs (Windows 7 desktops, each with Intel 3.3 GHz i3 processor, 4GB RAM and GIGABYTE Radeon R7 250), which included implementation of the In-Home Streaming service. PC 1 acted as a server of the Steam In-Home Streaming service, while the other two PCs (PC 2 and PC 3) were used as clients of the same Steam server.
service. Two games from different game genres were played in this study: *Serious Sam 3 (SS3)* as an example of a fast paced first-person shooter game, and *Bastion* as a representative of the action role-playing genre, thus including into the study games that differ in camera perspective, graphics style and quality, gameplay pace and the intensity of user interaction.

Given that previous studies have shown that player experience is an important factor influencing QoE ratings [6], we conducted tests using a homogeneous group of skilled and experienced players, as such players have been shown to be the most demanding in terms of QoE requirements. The participants were 15 male adults, aged between 23 and 32 (average age 26, median age 27). Prior to testing, participants were asked to report their previous gaming experiences, with emphasis on considered games in the study. 87% reported having prior experiences with playing SS3 and 40% of the participants reported having previously played Bastion.

In this experiment, we manipulated two video encoding parameters that have been shown to have a high impact on cloud gaming experience [12]: fps and bitrate. By setting the bitrate value and changing the value of FPS we indirectly control the image quality due to H264 coder’s rate control mechanism. It is a typical trade-off mechanism, setting a lower fps value leads to a less smooth game experience and can introduce jenkiness of the graphics during gameplay, but enables the use of more bits per pixel in a single frame thus increasing the image quality. Our goal is to see at which point the balance of these two parameters yields the highest QoE value for a given bandwidth threshold. Manipulation of these video encoding parameters was done using Steam’s developer console (manipulation of frame rate) and In-Home Streaming client GUI settings (manipulation of bitrate).

Four levels of video fps were used during the experiments: 15 fps, 20 fps, 25 fps and 30 fps. These video frame rate values were chosen based on previous study [26] which showed that serious degradation of the gaming experience and user’s in-game performance occurred when video frame rate was below 15 fps. Also, during testing sessions prior to this study, we noticed that the Steam platform does not support streaming of SS3 with frame rates lower than 15 fps, so we used this as a minimum value. We further selected three levels of video bitrate for testing purposes: 3 Mbps (minimum bitrate enabled by Steam), 5 Mbps, and 10 Mbps. With four fps levels, three bitrate levels, and two different games, we collected ratings for a total of 24 different test conditions. All conditions were tested by each participant, with the sequence of test scenarios being randomly selected for every player to avoid possible bias of manipulated video encoding parameters. The participants were instructed that they were playing games using the Steam streaming service. At the beginning of each game session, participants were given a small amount of time (tutorial phase) to familiarize themselves with a chosen map and gameplay mechanics of each game. The first 12 test scenarios consisted of playing one round of Serious Sam 3 survival mode on a single map. While the participants were playing on one of the client PCs (PC 2 or PC 3), a test administrator changed video encoding parameters on the second client PC for the next test scenario. Each of the twelve SS3 test scenarios was 2-3 minutes long, depending on how long a given participant lasted on the map without dying. After finishing each test scenario, players were instructed to fill out a questionnaire about their perceived graphics quality, perceived fluidity, and overall quality of experience (all reported on a 5-point MOS scale). In addition, participants were asked to report their willingness to continue playing under the current test conditions (yes/no). The test administrator further noted the amount of time participants spent during this test scenario. After filling out the questionnaire, players switched to playing on the other client PC under different test conditions (only one video stream was active at once). The second half of the test scenarios consisted of playing Bastion’s score mode on a single map. One run of Bastion on a single map contained three checkpoints where participants paused their gaming progress (approximately after 2 minutes of playing time), filled out a questionnaire and continued playing on the other PC. Once again, the test administrator recorded the number of player deaths and obtained game score during each test scenario. The entire gaming session lasted approximately 1.5 hours.

IV. RESULTS

The average subjective ratings of overall QoE for both tested games and across all test conditions are shown in Figure 2. Overall QoE for both games during test scenarios with 15 fps is notably lower then during test scenarios with higher values of frame rate, thus confirming findings of previous studies claiming that this frame rate value should be considered a minimum threshold value. Furthermore, no significant differences were observed in ratings for both games between test conditions with 30 fps and 25 fps, thus indicating that even highly skilled players do not notice small frame rate drops while playing games with high overall fps values such as 25-30 fps. It should also be noted that SS3 has on average higher scores of QoE for all fps levels in comparison with Bastion, except for test scenarios with 15 fps. We attribute this to fluidity being a more influential QoE dimension for first-person shooter games as compared to slower paced games. On the other hand, we found that reducing video bitrate under fixed frame rate had more significant degradation on Bastion then on SS3 (keeping in mind that the lowest bitrate tested was 3 Mbps). This can likely be attributed to graphics degradations being more perceptible in Bastion then in SS3.

Figure 3 provides a heat map overview of all collected mean subjective ratings of QoE features for both games during different test conditions. In terms of graphics quality, ratings naturally decreased with decreasing video bitrate values for both games. However, the impact of graphics quality was weaker for SS3 than Bastion, which can be linked to players’ perceptions of graphics impairments being hindered by the fast-paced nature of the game. We further note that our test results show that subjective graphics ratings did not increase with lower frame rates, which is contrary to our initial assumptions and also to the results reported in [12]. This contrary finding could be attributed to different games used in both our study and [12], and/or to the smaller sample size in our study. The data further shows that frame rate has more impact on perceived fluidity while playing SS3 then while playing Bastion, which is expected considering the difference in the dynamics of gameplay. Additionally, we report statistical relationships between overall QoE (and it’s features) and video encoding parameters in Table I. We compute Pearson’s product moment correlation coefficient $r$ to measure linear correlations. The data shows significant positive correlations between overall QoE and framerate for
both games, and between frame rate and perceived fluidity. Video bitrate had significant correlations with perceived graphics quality, while no correlations were found between bitrate and overall QoE for SS3. Furthermore, we report that we did not find any correlation or any statistically significant relationship between objective metrics (survival time for SS3, number of deaths and score for Bastion) and manipulated video parameters/subjective ratings.

TABLE I. CORRELATIONS BETWEEN RATINGS AND VIDEO PARAMETERS (** p-VALUE < 0.01, * p-VALUE < 0.05)

<table>
<thead>
<tr>
<th></th>
<th>Serious Sam 3</th>
<th>Bastion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall QoE</td>
<td>r = 0.652 ***</td>
<td>r = 0.07</td>
</tr>
<tr>
<td>Graphics quality</td>
<td>r = 0.289 ***</td>
<td>r = 0.137 ***</td>
</tr>
<tr>
<td>Fluidity</td>
<td>r = 0.421 ***</td>
<td>r = 0.07</td>
</tr>
</tbody>
</table>

The collected data was used to derive multiple linear regression models (by using ANOVA, Analysis of Variance), modeling overall QoE both in terms of independent manipulated video encoding parameters, and in terms of dependent QoE features for both games, as shown in Figure 4. The goal was to analyze how each of these predictors, and to what magnitude, contributes to the overall QoE. It should be noted that we consider our data as interval data and not ordinal (i.e., we consider that the intervals between points on the rating scales are equal). Also, while results of some test scenarios show skewness and kurtosis, and some of our test cases have not passed normality tests, ANOVA is considered quite robust on non-normality violations [27], and has been used for analysis of this type of data frequently in related work. Nevertheless this information should be taken into consideration when using the obtained model. ANOVA results (of video encoding parameters) for SS3 shows that only frame rate has a significant impact on QoE (p-value < 0.001), while video bitrate is statistically insignificant (p-value > 0.05). We attribute this to the high frequency of game screen changes characteristic for first-person shooter games distracting players from observing graphics degradations caused by lowering video bitrate, and due to players being sensitive to the smoothness of the delivered game video content. This analysis shows that one potential way to increase/preserve perceptive QoE of first-person shooter games is to keep video frame rate at a reasonably high level, even at lower bitrates. The accuracy of the aforementioned prediction model for QoE for SS3 is shown in Figure 5. We further model QoE for SS3 as a weighted linear combination of graphics quality and fluidity (Figure 4). Both predictors have a significant impact on QoE (p-value < 0.001), with fluidity having a significantly stronger impact, as expected from prior findings. The $R^2$ value is 0.72.

Analogously, we repeated the same procedure for Bastion, with models also shown in Figure 4. The model shows that both frame rate and bitrate have significant impacts on QoE (p-value < 0.001), with frame rate having a significantly stronger impact. Before conducting experiments, we expected that video bitrate (and with that, graphics quality) would contribute more to overall QoE then video frame rate, in accordance with Bastion being a slower-paced game with very detailed graphics. We also anticipated that our test group would be more affected by degradations of graphics quality induced by lowering the bit rate and by improvements in graphics quality due to lower frame rate. The accuracy of the model is shown in Figure 6. A prediction model for Bastion with gaming QoE features as predictors is shown in Figure 4, with an $R^2$ value of 0.66. As in the case of SS3, both of the QoE features were found to have a significant impact on QoE (p-value < 0.001), with fluidity again having a significantly stronger impact, but not as strong as in the SS3 prediction model.

As previously mentioned, we also asked participants after each test scenario to express their willingness to continue playing the game under the test conditions in a given scenario. Results are shown in Figure 7. For test scenarios that involved
Fig. 4. Summary of linear regression models for QoE with video parameters and QoE features

Fig. 5. Accuracy of predicted QoE ratings vs subjective QoE ratings for Serious Sam 3

Fig. 6. Accuracy of predicted QoE ratings vs subjective QoE ratings for Bastion

playing SS3, we observe that the percentage of players that are willing to continue playing declines with lower frame rates during experiments, with less than 40% of players willing to keep playing at 15 fps (regardless of bitrate down to 3 Mbps). There was no significant difference in the percentage of players that were willing to continue playing between test conditions with 30 and 25 fps, once more confirming the finding that even more experienced players are not aware of game performance degradations at these fps levels. Furthermore, it can be observed that lowering bitrate did not have any effect on the users’ willingness to keep playing, meaning that the streaming platform could without serious repercussions allocate the minimal amount of bandwidth (in this case 3 Mbps) for the addressed game types. In comparison, for test scenarios that involved playing Bastion, we note that by lowering the frame rate, participants’ willingness to play decreased, but that percentage remained on average higher then the percentage of players that did not want to continue playing SS3 under the same test conditions.

V. CONCLUSIONS AND FUTURE WORK

In this paper we presented a user study focused on measuring cloud gaming QoE using the Steam In-home streaming platform for test purposes. We investigated whether the video coder settings (in terms of frame rate and thus graphics quality) which maximize QoE under certain bitrate constrains vary between games. We summarize the following findings:
- Given the highest quality test condition, the average score for both games is 4 out of 5, and over 96% of the players would continue playing in such conditions,
- Lowering frame rate down to 25 fps does not significantly degrade the gaming experience regardless of the game,
- Bitrate reduction has a more significant impact on Bastion then on SS3, while in the case of frame rate the situation is reversed,
- While differences between games exist, fluidity has a more significant impact on QOE for both investigated games then graphics quality under the same bitrate constraints - we did not find cases in which reduction of frame rate resulted in increased QoE,
- Although, First Person Shooter has, at average, higher QoE scores for frame rate of 20 fps and higher, it is much more sensitive to severe degradation of frame rate then an adventure platform game (i.e. 15fps).

In future work we would like to test more players to confirm the findings. We will focus on fine grain tuning in the area between 25 and 20 fps. Further we would like to categorize games based on the temporal and spatial characteristics of the stream video so we could automatically assign them a proper adaptation strategy.

We compare our results regarding continuation of play to the results of the GamingAnywhere platform at the same bitrate reported in [16]. For both Steam and GamingAnywhere at bitrates of 3 Mbps, around 80% of the players were willing to continue playing, although the games under test are different, in [16] the Massively Multiplayer Online Role-Playing Game - World of Warcraft was tested, while in this paper a First Person Shooter and an Adventure Platform game were tested. This confirms that Gaming Anywhere is comparable with a commercial product at speeds of 3 Mbps.

We compare our results regarding continuation of play to the results of the GamingAnywhere platform at the same bitrate reported in [16]. For both Steam and GamingAnywhere at bitrates of 3 Mbps, around 80% of the players were willing to continue playing, although the games under test are different, in [16] the Massively Multiplayer Online Role-Playing Game - World of Warcraft was tested, while in this paper a First Person Shooter and an Adventure Platform game were tested. This confirms that Gaming Anywhere is comparable with a commercial product at speeds of 3 Mbps.
Fig. 7. Percentage of players willing to keep playing under different test conditions

REFERENCES


