Video Streaming into Virtual Worlds:

the effects of virtual screen distance and angle on perceived quality

Pengpeng Ni[1,2], Fredrik Gaarder[2], Carsten Griwodz[1,2], Pål Halvorsen[1,2]
[1]Simula Research Laboratory, Norway
[2]Department of Informatics, University of Oslo, Norway
{pengpeng, fredrga, griff, paalh}@ifi.uio.no

ABSTRACT

There is an increasing trend to include streamed video data in 3D environments. Such environments allow potentially several concurrently visible videos on a single display device, and consequently, network and processing bottlenecks. As a first step towards an avoidance of such problems, we have performed subjective assessments using a 3D application prototype to determine how positioning of video in the 3D environment influences the user perception of reduced-quality videos. Using video clips from several genres, we have compared the influence of various ways to reduce video quality and users' perception of degraded quality. We evaluated the influence of distance and angle of the placement.

Categories and Subject Descriptors

H.5.1 [**Multimedia Information Systems**]: Video

General Terms

Human Factors, Measurement, Experimentation

Keywords

User perception, Video quality, 3D virtual world

1. INTRODUCTION

Video streaming and 3D virtual worlds are becoming more mainstream as the availability of Internet and bandwidth is increasing. We see also the increasing use of systems that stream video content into online virtual worlds. As shown in figure 1, there are now often screens in the virtual world showing dynamic content, e.g., dynamic advertisements in games like Anarchy Online and virtual movie theaters systems like in PlayStation Home. This is a big, fast growing market [10], but at the same time, it raises the challenge of delivering more dynamic content to the end-users.

In our work, we are looking at how to adapt the video stream presented in a 3D world in order to reduce the required bandwidth. However, compared to traditional streaming to an exclusively used display, there are several aspects that influence the users' perceived quality in the 3D environment. For example, only parts of the total display surface are used for the streamed content. In figure 2, we show how the avatars' distance and angle to the screen influence the display. We conjecture that transmitting a full quality video is not necessary and investigate how display characteristics can be exploited to reduce the rate of the stream without the user noticing the degraded quality.

Figure 1: Virtual world video screen

Previous research has been performed to understand the relation between perceived video quality and human-related factors. The effects of picture quality and frame rate downscaling have for example been evaluated in [7]. In [6], studies were conducted to determine viewing preferences on mobile devices. Both investigations focused on video consumption in real life, which is different from viewing 3D scenes.

We start with perceived quality to investigate the effects of distance and angle of the display, and we have implemented a fully functional 3D application displaying video. We streamed video clips of two different genres in a virtual world, where it was displayed at 4 different distances and 4 different angles from the viewpoint. In these settings, we presented 4 versions of the clips at different qualities and 4 at different resolutions to assessors. A total of 2047 individual tests were performed, in which assessors indicated whether they noticed a difference between a full-quality clip and a reduced-quality clip in an identical viewpoint setting. The results were statistically analyzed to see if some conclusions could be drawn regarding the downscaling of video streamed into virtual worlds.

2. EXPERIMENT DESIGN

2.1 Implementation

Our prototype is implemented in the framework of the Object-oriented Graphics Rendering Engine (OGRE) [2] us-

MM'09, October 19–24, 2009, Beijing, China.

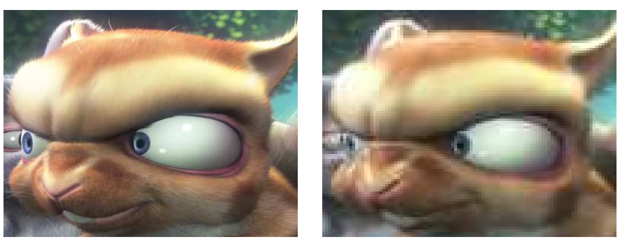
(a) Distance 1400 wu (b) Distance 2100 wu (c) Distance 2800 wu

(d) Angle 25° (e) Angle 50° (f) Angle 75°

Figure 2: Sample screenshots from the prototype

(a) 100% quality (b) 75% quality (c) 50% quality (d) 25% quality

Figure 3: Frame samples at different video quality levels (Quantization level increase)



(a) 100% resolution (b) 25% resolution

Figure 4: Different quality levels (resolution reduction)

ing the graphics engine to create the 3D environment. Videos are decoded for presentation in the environment using FFmpeg [1]. Decoded videos can be played back on one or more surfaces within the 3D environment (called virtual screens or video surfaces). The user's view of the virtual screen can be changed dynamically as shown in figure 2 as the avatar moves through the online world.

2.2 Video Material

We selected two video sequences from popular genres. The sequences are the same used in earlier studies on video quality assessment [3, 8], allowing for a comparison with earlier results. From each sequence, we extracted an 8 second clip (200 frames) without scene cuts. The first video (figure 3) is a sports clip, featuring a waterskiing man. It has an original spatial resolution at 640 by 352 pixels, and a maximum bitrate at about 1500 Kbps after re-encoding by FFmpeg. The second clip (figure 4) featured parts of a scene from the open source, short animated film "Big Buck Bunny" by the Blender Foundation [9]. The chosen footage shows three furry squirrels plotting against the bunny, with some camera panning and detailed background foliage. The clip's original resolution is 1280 by 720 pixels and maximum average bit-rate at about 4576 Kbps.

To generate video clips with smaller size (bitrate), quality impairments were injected into copies of the tested video sequences to varying degrees. In our experiment, the two video clips were downscaled, reducing required bitrates alternately by using a larger quantization parameter and by reducing spatial resolution. Three levels of each downscaling option were tested, at 75%, 50% and 25% of the original spatial resolution and maximum bitrates. In figure 3 and 4, different quality levels are shown for the two clips (zoomed to see the differences on small thumbnails in this paper). It might be hard to see on these small examples, but in general, the quality reduction resulted in blockiness and blurriness.

2.3 Procedure

We conducted a field study at different locations (campus, living rooms, classrooms). All assessors watched the clips on the same laptop. The visual setup was a screen resolution of 1024 by 768 pixels on nineteen inch LCD monitors with aspect ratio 4:3, which natively matched the aspect ratio of the raw rendered output. To test how viewing distance and angle affect users' perception of video quality, 4 distances and 4 angles were selected ranging from 700 to 2800 world units (wu) and 0° to 75° (see figure 2), respectively, where

700 wu and 0° provided the assessors with a frontal view of the virtual screen.

The subjective test design was based on the double-stimulus impairment scale method (DSIS) [4, 5] which implies that test clips are presented in pairs: the first clip presented in each pair is always the unimpaired reference, while the second is the same clip with reduced quality. Video clip pairs were chosen from randomly selected viewing distances and angles, and the reference clip was compared with versions containing random kinds and levels of impairments. Combining 12 different viewing condition with 6 levels of impairment, there were 72 different test scenarios for each test sequence. A binary rating scale was used for testing under which viewing conditions the assessors could detect quality loss. Binary tests are preferred by many assessors in our previous field studies [3, 8] since it enhances the certainty of the voting given its discriminatory power.

To assess the videos, we used people, mostly university students, that volunteered when asked to participate. Our assessment was performed at the main campus of Oslo University, where there is no bias with respect to the field of interest of the students. The participants spent an average of 15 minutes on the test. Brief instructions were given to assessors prior to their test session, which informed them that they should compare two video clips and that one of the two might be of inferior quality compared to the other one. Assessors expressed their opinion by clicking one of the two buttons of a mouse connected to the laptop. When the voting was over, the assessors moved on to the next randomly selected test case by clicking a control button.

3. RESULTS

We have gathered in total 2047 answers. Combing the ratings from the two genres (both video clips shows the same trends), there are at least 20 ratings for each test case, which meets the minimum requirement of sampling size [4, 5]. For statistical analysis, we ran binomial tests to see if a significant majority of answers indicated detectable quality degradations. Results are reported as detectability of quality loss with 0.05 confidence intervals.

In addition to the significance test, we examined the normalized share of positive ratings in each sample from all test cases. The shares in samples that show significant detectability and non-detectability are $\geq 73\%$ and $\leq 32\%$; while the shares in the samples that show insignificant effect fall into the interval between $[30\%, 68\%]$. Due to different sampling size (number of ratings), there is a small overlap between the percentage ranges of the three sampling groups. The average sampling size in table 1 and 2 is 28.4. Since the test scenarios were randomly chosen by the prototype, the number of ratings is not identical for each test case. Hence, some samples in the tables have smaller size than others, which may affect the result of the statistical analysis. By re-checking the sampling size, we found that some samples showing insignificant effect contain noticeably fewer than average ratings. We note, however, that the ratings are distributed around the significance bound of $[30\%, 68\%]$, but significance can most likely be achieved by adding a small number of additional samples.

3.1 The Effects of Distance

Table 1 shows the test results where viewing distance is the factor being tested. The viewing angle was kept constant (at 0° and 25°) while the distance was stepwise increased. We grouped the results according to the applied downscaling techniques and displayed those results separately in table 1a and 1b. Correspondingly, the second column in the two tables indicates the amount of impairment injected into the target clips relative to one aspect of video quality measure, either quantization levels or spatial resolution.

a) Changing quantization levels

Angle	Impairment Degree	Distance (wu) 700	1400	2100	2800
0°	25%	○	-	(-)	-
	50%	○	-	-	-
	75%	+	○	○	-
25°	25%	○	-	-	-
	50%	○	○	-	-
	75%	+	○	○	-

b) Downscaled spatial resolution

Angle	Impairment Degree	Distance (wu) 700	1400	2100	2800
0°	25%	○	-	-	-
	50%	(+)	-	-	(-)
	75%	+	+	○	○
25°	25%	-	-	-	-
	50%	+	-	-	-
	75%	+	+	○	-

Table 1: Detectability of quality loss versus distance (+ detectable, - non-detectable, ○ not significant, (*) not significant but having small sample size).

Despite some insignificant results and inconsistencies, it is quite apparent that the distance does impact people's perception of quality loss. In table 1b, some ratings indicate significant detectability when the distance is closer to the virtual screen. Meanwhile, the significance is decreased as the distance increases. Similar trends can also be found in table 1a. At the distance of 2800 wu, the results show that the majority of people do not detect quality loss even at the highest amount of quality reduction, i.e., we can reduce the rate by 75% of the original, saving a significant amount of resources.

It is also clear that the degree of impairment and the distance mutually affect user-perceived video quality. Looking at the ratings for an angle of 25° and spatial downscaling, people are unlikely to detect the quality loss when the downscaling by 50% or less at a distance 1400 wu or more. On the other hand, more than 25% downscaling at a distance of 700 wu was easily detected by most of our assessors.

3.2 The Effects of Angle

Table 2 shares the setup of table 1, but instead of virtual viewing distance, the viewing angle became the testing factor. Not all of the results in table 2 are significant, and we infer that the insignificance is related to the amount of quality loss and assessor's individual preference. Note also that the screen is rotated around its center of the video instead of its close edge.

The results show that there is a weak tendency that the significance is decreased as angle grows, i.e., the quality can be reduced with a growing angle. However, a general conclusion is not possible based on table 2 in terms of the effect that the viewing angle might have on quality perception. One potential explanation for this is that video appears very different when shown on a virtual screen with an angle. Most people are not used to watch video in this fashion, which may have triggered feelings of big differences, even though the differences were very subtle. Another explanation may be related to the QGRE 3D engine used by our prototype which might cause the closest side of the virtual screen being trimmed as the angle increases due to its proximity. However, more work is required in this area.

a) Changing quantization levels

Distance	Impairment Degree	Angle 0°	25°	50°	75°
700 wu	25%	∘	∘	∘	+
	50%	∘	∘	∘	∘
	75%	+	+	+	+
1400 wu	25%	-	-	∘	-
	50%	-	∘	-	∘
	75%	∘	∘	∘	-

b) Downscaled spatial resolution

Distance	Impairment Degree	Angle 0°	25°	50°	75°
700 wu	25%	∘	-	-	-
	50%	(+)	+	+	∘
	75%	+	+	+	+
1400 wu	25%	-	-	∘	-
	50%	-	-	∘	-
	75%	+	+	+	∘

Table 2: Detectability of quality loss versus angle (+ detectable, - non-detectable, ∘ not significant, (*) not significant but having small sample size).

4. DISCUSSION

The two tested downscaling techniques are responsible for different types of distortion. A reduced spatial resolution will cause blurrier images while a larger quantization parameter will bring blocky artifacts. By comparing the result grouped by downscaling technique, we get the impression that it is easier for people to detect blurriness than blockiness in virtual worlds. Combing with the bit saving enabled by the respect downscaling method, it may give us an optimal way to prioritize the downscaling options, which takes user's experience into concern.

Distance is the dominant factor influencing user's perception according to this study. We observed that the reduction of virtual viewing distance is close to matching the linearity of increasing bitrate requirement for maintaining user's satisfaction. However, perceived video quality is not linear for people and the perception is usually expressed by binary or ordinal data [3,8]. It is more feasible to find the distance from where given amount of distortion is not detectable. Our experiment indicates that a *detectability threshold* exists which can be used for dynamic bitrate adaptation. For example, we identified the detectability threshold for 75% of impairment at about 2800 wu, while for 50% of impairment the threshold may exist not too far away than 1400 wu.

For the effect of viewing angle, a similar *detectability threshold* may also exist as there are some indications that quality (and resource requirement) can be reduced as the angle to the virtual screen increase. But, we needs further investigation and justification in terms of additional experiments.

Further, we have tested different types of content with varying detail and motion. The effects varies slightly for different content characteristics to the degree, but share similar tendency for the detectability of quality loss.

5. CONCLUSION

To investigate the effects of distance and angle of the display, we have implemented a fully functional 3D application displaying video. In our experiments, we have discovered that distance and angle between the avatar and the virtual video screen do influence the ability to perceive video quality. However, it is difficult to give an exact general approximation of the degree of the possible quality reduction and how to downscale. Our results indicate that decreasing picture quality performs better than reducing the spatial resolution with respect to bitrate adaptation.

To make general downscaling guidelines, more questions like "what is the required 100% quality?" and "what combinations of downscaling techniques (like picture quality, spatial resolution, frame rate, etc.) are appropriate?" must be asked. Furthermore, tests varying more parameters must be also performed. Currently, we are therefore designing further experiments where we will try more content types in combination with various approaches for quality downscaling using MPEG4 scalable video codec.

6. REFERENCES

[1] FFmpeg. *http://www.ffmpeg.org/* (2009).

[2] OGRE. *http://www.ogre3d.org/* (2009).

[3] Eichhorn, A., and Ni, P. Pick your Layers wisely - A Quality Assessment of H.264 Scalable Video Coding for Mobile Devices. In *IEEE International Conference on Communications* (2009).

[4] International Telecommunications Union. *ITU-T P.910. Subjective video quality assessment methods for multimedia applications*, 1999.

[5] International Telecommunications Union - Radiocommunication sector. *ITU-R BT.500-11. Methodology for the subjective assessment of the quality of television picture*, 2002.

[6] Knoche, H. O., and Sasse, M. A. The sweet spot: how people trade off size and definition on mobile devices. In *MM '08: Proceeding of the 16th ACM international conference on Multimedia* (New York, NY, USA, 2008), ACM, pp. 21–30.

[7] McCarthy, J. D., Sasse, M. A., and Miras, D. Sharp or Smooth?: Comparing the Effects of Quantization vs. Frame Rate for Streamed Video. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2004), pp. 535–542.

[8] Ni, P., Eichhorn, A., Griwodz, C., and Halvorsen, P. Fine-Grained Scalable Streaming from Coarse-Grained Videos. In *Proc. of NOSSDAV* (2009).

[9] The Blender Foundation. Big Buck Bunny. *http://www.bigbuckbunny.org* (2008).

[10] Yankee Group. Advertising and Games: 2007 In-Game Advertising Forecast. *http://www.yankeegroup.com/ResearchDocument.do?id=16395*.