EFFECT OF CONTENT CHARACTERISTICS ON QUALITY OF EXPERIENCE OF ADAPTIVE STREAMING

Samira Tavakoli[1], Muhammad Shahid[2], Kjell Brunnström[3], Narciso García[1]

[1]Universidad Politécnica de Madrid, [2]Blekinge Institue of Technology, [3]Acreo Swedish ICT

ABSTRACT

The growing popularity of adaptive streaming-based video delivery nowadays has raised the interest about the user's perception when experiencing quality adaptation. The impact of the video content characteristics on user's perceptual quality has already become evident. The aim of this study is to investigate the influence of this factor on the user's quality of experience in adaptive streaming scenarios. Our results show that the perceptual quality of adaptation applied on videos with high spatial and low temporal amount of activity is significantly lower compared to the other content types.

Index Terms— Adaptive Streaming, Quality of Experience, Content Characteristics

1. INTRODUCTION

Nowadays adaptive streaming technology is employed by numerous content providers to adapt the quality of delivered video to the network bandwidth fluctuations and users' device capabilities. Adaptation can be done by switching between different quality representations of the video content (e.g. in encoding parameters, frame-rate, or frame resolution) partitioned into small chunks, and through various strategies. Using adaptive streaming offers significant advantages such as less video playback interruption which results in higher user quality of experience (QoE). Besides study on adaptation technical development, investigating on the factors which affects on adaptive streaming QoE is crucial to optimize the provided quality. One of such studies is presented in [1] where a subjective assessment of QoE of different scenarios for video quality adaptation in response to bandwidth variations was performed. Mainly, the different level of perceived quality by switching to alternative encoding bitrates in different frequency and amplitude was studied.

Because of using long sequences in subjective evaluation of adaptive streaming, it would be difficult to perform full factorial studies. Therefore in order to properly interpret the experimental results it is necessary to take the objective characterization of the content viewed by the people into account.

There are many studies such as [2] that show the influence of the content characteristics, namely spatial information (SI) and temporal information (TI) as defined in ITU-T Rec. P910 [3], on perceptual video quality. SI describes the level of details in a single image, while TI indicates the degree of similarity between two subsequent images of a given sequence. It was shown that videos with different amount of spatial details and motion dynamics receive different values of mean opinion score (MOS) under similar compression rates. Moreover, various artifacts can be masked by SI and TI present in the relevant areas/intervals of a video.

Motivated by above findings, in this paper the perceptual quality of adaptation strategies with respect to the spatial and temporal characterization of the processed video sequences (PVS) used in our previous experiments is studied. After brief overview of the experiments, our main findings are presented in the rest of the paper.

2. EXPERIMENT OVERVIEW

Adaptation scenarios

The goal of our previous study [1], was to investigate the impact of different adaptation scenarios on QoE of the users. To this aim different strategies to increase and decrease the video quality were proposed as the Hypothetical Reference Circuits (HRC). They consisted of rapid and gradual change of the encoding bitrate using four adaptation streams (quality levels from 5 Mbps to 600 kbps) and two different chunk sizes (2 s and 10 s). Furthermore, the QoE of each of the adaptive streams was also examined to explore the perceptual difference between distinct quality levels. Seven source videos (SRC) in various genres and with different spatial and temporal information were used to investigate the impact of video content on QoE of the twelve HRCs.

Subjective experiment

The MOS data used in the current study was collected from three subjective experiments. One experiment was done in the Acreo Swedish ICT lab using single stimulus method with the setup presented in [1] and 5 levels absolute category rating (ACR) scale for evaluation. The other two experiments were conducted in the lab of Universidad Politecnica de Madrid (UPM), with the same equipment and setup as the first lab, but different assessment methodology. Similar methodology as [4] was employed in UPM while using the same scale and questions as the one used in Acreo (a more elaborate description of the experiments in UPM is in preparation). After
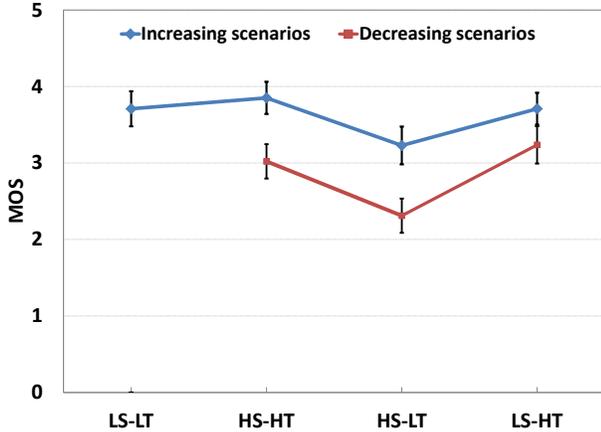
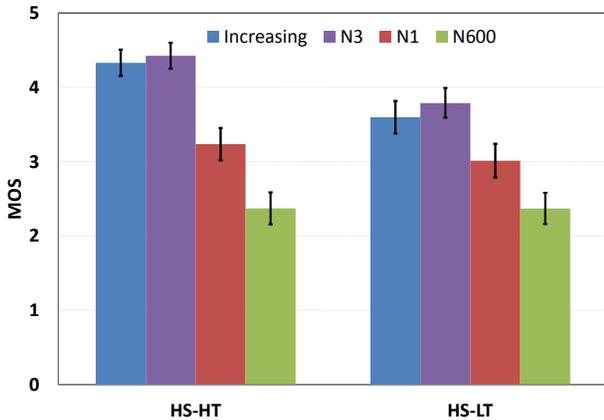**Fig. 1**. MOS for decreasing scenarios in different content

Fig. 2. QoE of increasing scenarios vs. N3, N1 and N600

screening, the observations of 63 subjects (20 from Acreo and 43 from both experiments in UPM) were considered for the current study.

3. SELECTED RESULTS

To explore the effectiveness of the spatial and temporal characteristics of the content on perception of the adaptation strategies, the PVSs used for the evaluation of proposed scenarios were classified by using a combination of spatial and temporal perceptual information of the content [2]. Through this classification, each PVS can belong to one of these categories: low spatial- low temporal (LS-LT), low spatial- high temporal (LS-HT), high spatial- low temporal (HS-LT), or high spatial- high temporal activity (HS-HT).

Statistical tests (Fisher's Least Significant Difference, LSD, and Tukey Honestly Significant Difference, HSD) on the MOS data were performed considering the existed PVS classes used for each of the adaptation scenario type (it was found that in some of the scenarios not all four classes were used for the evaluation). In this paper because of the space limitation, only the aggregated results are presented.

The MOS comparison of PVS classes used for evaluating the decreasing scenarios (aggregated of all the decreasing

scenarios) is shown in Fig. 1. A statistically significantly lower quality (HSD-p & LSD-p < 0.05) have been perceived in decreasing the quality of those PVS with high spatial and low temporal level of activity (HS-LT) compared to the other content classes. Also, the MOS of quality increasing scenarios in this specific content class has been statistically significantly lower than the other content classes.

Regarding the perceptual quality of the adaptive streams, the HS-LT content class when encoded in 3 Mbps (N3) and 5 Mbps (N5) has been perceived significantly worse than the other content encoded with the similar encoding parameter. About evaluating the videos encoded in 1 Mbps (N1) and 600 kbps (N600), the PVSs used in the experiments belonged on either HS-HT or HS-LT classes. Here, no statistically significantly difference has been found on QoE of these two content classes in each of these encoding scenarios.

It was presented in [1] that the perception of increasing scenarios in some of the video sequences was not significantly different than experiencing N1, and in a few cases, than N600. By analyzing the result through content classes, we found that increasing the quality for the contents of HS-HT and HS-LT classes has provided significantly better perceptual quality compared to N1 and N600 (Fig. 2). Furthermore, the QoE of increasing scenarios have not been significantly different than N3.

4. CONCLUSION

Because of using longer sequences in the subjective evaluation of adaptive streaming, it would be difficult to perform full factorial studies. Therefore to explain the experimental results it is necessary to take the objective characterization of the content viewed by the people into account. Our results show that the QoE of adaptation in the content with high spatial and low temporal amount of activity is significantly lower compared to the other content types. Furthermore, since the SI and TI values are in the same range also for the constant bitrate case, we can expect a similar behavior for the content itself (SRC) in general. Our analysis also shows that the applied classification can be useful for the test design similar to the presented study.

5. REFERENCES

[1] S. Tavakoli, K. Brunnström, K. Wang, B. Andren, M. Shahid, and N. Garcia, "Subjective quality assessment of an adaptive video streaming model," in *IS&T/SPIE Image Quality and System Performance XI*, Feb. 2014.

[2] N. Cranley and L. Murphy, "Incorporating User Perception in Adaptive Video Streaming Systems," *Digital multimedia perception and design*, pp. 242–263, 2006.

[3] "Subjective video quality assessment methods for multimedia applications," September 1999, ITU-T Recommendation P910.

[4] S. Tavakoli, J. Gutierrez, and N. Garcia, "Subjective quality study of adaptive streaming of monoscopic and stereoscopic video," *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 4, pp. 684–692, April 2014.