4th International Workshop on Perceptual Quality of Systems (PQS 2013)
2-4 September 2013, Vienna, Austria

Does Low Quality Audiovisual Content Increase Fatigue of Viewers?

Sebastian Arndt, Robert Schleicher, Jan-Niklas Antons

Quality & Usability Lab, Berlin Institute of Technology, Germany

sebastian.arndt@telekom.de

Abstract

While most quality studies measure purely the subjectively perceived quality of videos, physiological measurements can give more detailed information about the perception of multimedia consumers. Thus, they provide better information about long-term effects as they are difficult to estimate by purely subjective studies. The current study uses electroencephalography (EEG) to measure effects of fatigue during watching low and high quality videos. Therefore, a usual video was chosen, after partly reducing the bit rate of the video subjects watched it and gave subjective ratings accompanied by measuring EEG. With the obtained data we showed that lower video quality produces a higher percentage of alpha waves in the EEG, thus participants got more fatigued. Since these alpha waves are an indicator for the level of fatigue.

Index Terms: Video Quality, EEG, MOS, Fatigue, Alpha Frequency Power

1. Introduction

As multimedia streaming platforms, such as video on demand services or live streaming offers from providers, are becoming more popular it is important for the providers to use the limited infrastructure most efficiently without neglecting the wish of their customers for good quality. To determine what good quality is subjective user ratings are usually obtained. These video quality tests are usually being done while test participants evaluate the quality on an absolute scale, using no explicit reference for their judgment (ACR) or using an explicit reference stimulus (DCR). These methods are described and specified in detail in ITU-T Recommendation P.910 [1], they are being used for evaluation of short sequences with a length of 10s. For longer sequences the single stimulus continuous quality evaluation (SSCQE) is being used as mentioned in ITU-R BT.500 [2]. Here subjects give continuously a conscious judgment on a slider. All these subjective evaluation methods have in common that a conscious judgment is being asked for. Which might appear distracting especially in the case for longer video sequences since test participants might have to look on the slider to keep track of its position. Furthermore, no assumption about the internal processing and how this judgment is formed in particular can be done while purely analyzing these subjective data. In simplified auditory and (audio)visual quality tests electroencephalography (EEG) has proven to be a valid complement method to assess the quality perception for short sequences, as parameters of event-related potentials (ERP) are correlated with stimulus quality [3]. These ERPs are elicited while subjects are exposed to external stimuli and vary in their amplitude in general due to differences in e.g. intensity. In some studies ERPs were even more sensitive to distortions than subjective tests, similar brain pattern during the perception of strong degraded stimuli and below threshold stimuli which have not triggered a behavioral response were shown on the basis of ERPs. [4].

The quality of stimuli with longer duration, i.e. in the range of minutes, is not only relevant for choosing the one services over an other but also have an effect on the consumers cognitive state, such as arousal and fatigue, as these are rather long-term effects. Fatigue can be measured by using the alpha range EEG signals. Furthermore, the theta band is indicative for the state of fatigue and drowsiness as shown for example in the context of car drivers [5]. These cognitive processes are significantly affected also during perception of long low quality stimuli, as shown previously in an auditory study [6]. Here subjects were presented an audiobook while the brain activity was recorded via EEG. In that study EEG has proven to be a valid instrument to assess these states and showed that listeners got more fatigued while listening to a low quality stimulus in comparison to the high quality version. This was shown by analyzing changes of the alpha frequency band power using a fixed band-pass filter (8 - 13 Hz), as the proportion of alpha waves is an indicator of fatigue and impaired information processing [7]. Long term changes might also be detectable by other physiological measures such as skin conductance or heart rate variability, but these measures tend to be more imprecise in general.

In this paper we conducted a similar experiment to the auditory study, using audiovisual material and manipulating the video track only. For EEG analysis we used an improved individual alpha frequency band extraction. Which is using an individual alpha band range for each subject, as the alpha band is variable over subjects [8]. Additionally, the theta band was analyzed as it gives additional information about the cognitive state.

10.21437/PQS.2013-14

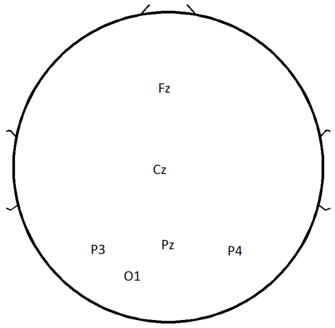

Figure 1: *Electrode setup used during the measurement. Topview; top depicts nose, sides ears.*

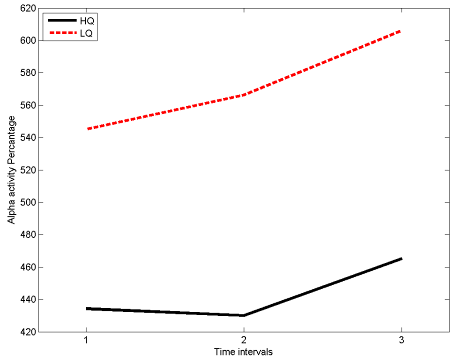

Figure 2: *Grand average of percentual alpha frequency band power changes for both qualities (HQ and LQ) and the three parts within each quality level. Data is extracted from electrode Cz.*

2. Methodology

2.1. Stimulus

The stimulus was a 39 min audiovisual clip as part from the video 'Ocean' with a full HD resolution of 1920 by 1080 pixel. It was showing sea life scenes from below and above the sea level. Thus the video contained brighter and darker scenes. The video was divided into two parts, each with ca. 19 min duration. One part retained the original video bit rate, as taken from the Bluray with 40,000 kbps, named HQ, for high quality, from now on. The other part had a reduced bit rate of 2,000 kbps, named LQ, for low quality. This bit rate reduction was implemented by using the x264 encoder choosing the option of constant bit rate for each part. It was randomized whether the first or the second part of the video clip was distorted. The audio track was not manipulated in either part (HQ and LQ). The audio was uncompressed with a sampling rate of 48 kHz and 32 bit resolution resulting in to a bit rate of 1,536 kbps.

2.2. Experiment

Participants were seated 1.80 m away from the screen in a comfortable arm chair their task was to focus on the video, about every 6 min a pop-up appeared asking for a quality rating. The used scale ranged from 0 (bad) to 10 (excellent) in discrete steps following the ITU-T Rec. P.910 [1]. The test was performed in an ITU-T Rec. P.910 [1] conform room with daylight lighting and gray curtains. The video was presented on a 42" LCD screen.

2.3. EEG

Participants were attached to a six channel EEG system from Brainproducts with electrodes at the locations Fz, Cz, Pz, P3, P4, and O1 according to the 10-20 international standard system (see also Figure 1). As reference electrode the right mastoid was used.

Subject	Lower α	Higher α
1	11	13
2	10	12
3	11	13
4	9	11
5	9	11
6	9	1
7	8.5	10.5
8	11	13
9	11.5	13.5
10	10	12
11	10	12
12	10	12
mean	9.35	11.35

Table 1: Overview of extracted individual alpha levels for each subject.

3. Results

Twelve German participants (3 female, 9 male, average age 26) took part in this study. All of them had corrected-to-normal vision, which was was confirmed with a vision test using the Snellen chart for visual acuity and Ishihara charts for a test of color blindness at the beginning. None reported any hearing impairment or health issues.

3.1. EEG

For analysis EEG data was band-pass filtered (1 - 40 Hz). The data recorded during the presentation of the two parts (HQ and LQ) was divided in a baseline interval, containing the first 2 min, and then into three equally spaced parts (each 5.5 min) which roughly correspond to the rating sections.

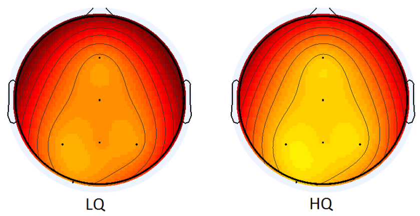

Figure 3: *Grand average of percentual alpha frequency band power changes for both qualities (left low quality and right high quality). Data is averaged over all measured time points.*

The EEG data was analyzed using a power spectral density estimate using Welch's method [9]. The obtained spectral data were divided into the corresponding frequency bands as mentioned in [10] and summed up to one value per band. In case of the alpha band the corresponding range was computed on a single individual subjects basis. Since the alpha frequency range can be affected by inter individual variation [7]. The individual alpha peak for each participant was extracted in the range from 7 to 13 Hz using the data recorded during the baseline interval. Around the frequency with maximum power value on the aforementioned frequency band the individual alpha band (IAB) was set with 1 Hz above and 1 Hz below the extracted value (see Table 1 for detailed inspection). Resulting in an average alpha of 9.35 to 11.35 Hz for all subjects. The power in the IAB for the HQ and LQ condition was computed and set in relation with the baseline, as percentage (see Figure 2).

A spatial distribution of alpha activity can be seen in Figure 3. For analysis purposes we concentrated on the changes in the electrodes in the back, since these show bigger changes in IAB between the two tested conditions as can be seen. It was observed that lower quality caused a higher IAB activity than the qualitative better part. A repeated measure ANOVA was calculated and yielded significance for the factor *time* ($F(11,1) = 4.66$, $p \leq 0.05$) at electrode P4 and the factor *quality* ($F(11,1) = 5.02$, $p \leq 0.05$) for electrode O1.

In case of theta band a general trend could be observed. The lower quality part of the video leads to a higher theta power compared to the undisturbed high quality part. Although this effect is not statistically significant at least a major trend can be observed at electrode O1 for the factor *quality* ($F(11,1) = 4.12$, $p = 0.07$) (see Figure 4).

3.2. Subjective Ratings

The obtained quality ratings show that the better the quality was the higher the MOS ratings were, significant ($F(12,1) = 71.86$, $p \leq 0.01$) (see Figure 5) as expected and should serve as a sanity check measure in this study.

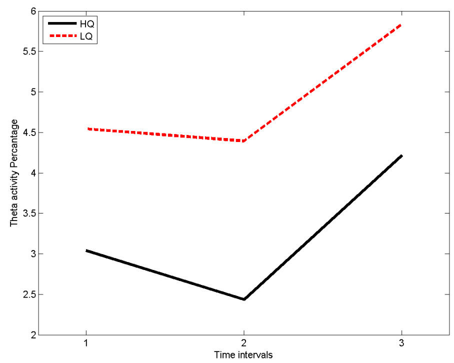

Figure 4: *Grand average of percentual theta frequency band power changes for both qualities (HQ and LQ) and the three parts within each quality level. Data is extracted from electrode O1.*

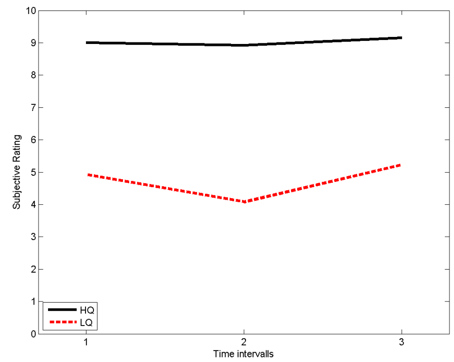

Figure 5: *Subjective rating during the test. For each time point, different lines depict the two different qualities.*

There is no statistical significant difference between the ratings from subjects who were exposed first the HQ and then the LQ sequence compared to the ones who first saw the LQ part and then the HQ ($F(1,5) = 0.8$, $p = 0.41$). Also there was no statistical interaction effect for *quality*sequence* ($F(1,5) = 3.16$, $p = 0.14$). However, by trend the HQ sequence was evaluated better when subjects first saw the degraded sequence (LQ) and subsequent the original than vice versa. Furthermore, subjects tend to rate the LQ sequence worse if they saw the HQ sequence first. Latter difference is larger than the former mentioned (see also Figure 6).

4. Conclusion

From the data we obtained during the EEG measurement we conclude that due to the higher IAB portion in the EEG signal during the LQ condition participants had a higher cognitive engagement. This elevated engagement was caused by the fact that participants had to compen-

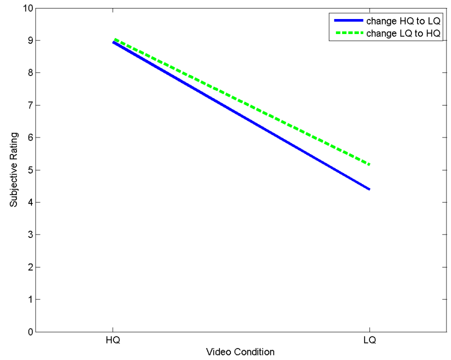

Figure 6: *Subjective rating during the test. Ratings during one quality block were averaged. The different lines depict whether subjects first saw HQ part or first saw the LQ part.*

sate for the lower quality on a neuronal level and thus got more fatigued in comparison to the HQ part. This can be seen throughout all measured time points since the IAB power is always higher (see Figure 2). Second, we conclude that there was a time on task effect. Thus, no matter in which condition, the IAB power increased over the three measured time points significantly. Which means that participants got more fatigued the longer they watch the video. This effect was stronger for the LQ part than for HQ part. However, this effect is a common phenomenon and is known as time-on-task effect from psychology. Similar trend can be observed for the theta band.

To summarize, in this experiment we could show that a quality reduction on long-term audiovisual stimuli has an effect not only on the subjective judgment of observers but also on internal cognitive processes. In this study we assessed the state of fatigue for this purpose. This is in line with previous research using purely auditory stimulation [6] and could on the long run lead to an autonomic estimation of the cognitive state without asking subjects for ratings. This could be of special interest in the case of e.g. video streaming and video conferencing. In the latter case a more vigilant state is desirable for customers, resulting in a more positive perception of calls and for service providers this could lead to longer usage periods of their services.

To better correlate subjective with physiological signals it might be advantageous to assess the state of fatigue from the subject with help of an additional questionnaire scale.

5. Acknowledgments

We are thankful to Ahmad Abbas for his support during data acquisition, and Sebastian Möller and Savvas Argyropoulos for their technical advices concerning video manipulation. This work was co-funded by the German Research Foundation (DFG - 1013 Prospective Design of Human-Technology Interaction) and Bernstein Focus: Neurotechnology Berlin (BFNT-B) which was supported by the Federal Ministry of Education and Research (BMBF) FKZ 01GQ0850.

6. References

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

[2] ITU-R Recommendation BT.500-12, Methodology for the subjective assessment of the quality of television pictures, 2009,International Telecommunication Union, Geneva

[3] Arndt, S. and Antons, J.N. and Schleicher, R. and Möller, S. and Curio, G., Perception of low-quality videos analyzed by means of electroencephalography, IEEE Quality of Multimedia Experience (QoMEX), 2012

[4] Antons, J.N. and Schleicher, R. and Arndt, S. and Möller, S. and Curio, G., Analyzing Speech Quality Perception Using Electroencephalography, Selected Topics in Signal Processing, IEEE Journal of, 2012, 6, 721 - 731

[5] Lal, S. and Craig, A., A critical review of the psychophysiology of driver fatigue, Biological psychology,173–194, 2001, Elsevier

[6] Antons, J.N. and Schleicher, R. and Arndt, S. and Möller, S. and Curio, G., Too tired for calling? A physiological measure of fatigue caused by bandwidth limitations, IEEE Quality of Multimedia Experience (QoMEX), 2012

[7] Klimesch, W., EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis, Brain research reviews, 29, 169–195 ,1999, Elsevier

[8] Klimesch, W and Doppelmayr, M and Russegger, H and Pachinger, Th and Schwaiger, J, Induced alpha band power changes in the human EEG and attention, Neuroscience letters, 1998,Elsevier

[9] Welch, P., The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, Audio and Electroacoustics, IEEE Transactions on, 70–73, 1967, IEEE

[10] M. Coles and M. Rugg, Electrophysiology of Mind: Event- Related Brain Potentials and Cognition, chapter Electroencephalography and High-Density Electrophysiological Source Location, Oxford University Press, 1995