Proceedings of the Fifth IASTED International Conference
COMMUNICATION SYSTEMS AND NETWORKS
August 28–30, 2006, Palma de Mallorca, Spain
ISBN Hardcopy: 0-88986-606-6 / CD: 0-88986-608-2

PERCEIVED QUALITY OF CHANNEL ZAPPING

Robert Kooij[1,2], Kamal Ahmed[1], Kjell Brunnström[3]
[1] TNO Information and Communication Technology, Delft, the Netherlands, {robert.kooij,kamal.ahmed}@tno.nl
[2] Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science
[3] Acreo, Kista, Sweden, kjell.brunnstrom@acreo.se

ABSTRACT
The end user experience of service quality is critical to the success of a service provider's IPTV deployment program. A key element involved in validating IPTV quality of experience (QoE) is how quickly and reliably users can change TV channels, often referred to as channel zapping. Currently there is no knowledge about the explicit relation between zapping time and the user perceived quality as expressed as a Mean Opinion Score (MOS). We have proposed a model where the MOS depends on the zapping time on a logarithmic scale. In order to validate our model we have conducted a number of subjective tests in order to get insight in this relation. The tests were performed by 21 test subjects at TNO and at Acreo. It turns out that the correlation between the subjective data and the perceptive model is very high (0.99). Therefore we conclude that our perceptive model is very useful for assessing the perceived quality of zapping. The result is not limited to IPTV but applies to any type of TV service with channel switching times that are not instantaneous, regardless whether it is terrestrial or satellite based.

KEY WORDS
Zapping, perceived QoS, subjective tests, MOS

1. Introduction

The end user experience of service quality is critical to the success of a service provider's IPTV deployment program. A key element involved in validating IPTV quality of experience (QoE) is how quickly and reliably users can change TV channels, often referred to as channel zapping. Although it is well-known that zapping time is an important factor influencing QoE there is not much information about the requirements for zapping time (a.k.a. as switching time). In [1] zapping time is called satisfactory if it is below 1 second while according to [2] usual figures for a simple change of channels can be 1-2 seconds. A recent literature study revealed that currently there is no knowledge about the explicit relation between zapping time and the user perceived quality as expressed as a Mean Opinion Score (MOS).

The main contribution of this paper is a model that maps zapping time to perceived quality in terms of MOS.

Note that because IPTV is an eminent part of triple play offerings the performance of channel zapping is of great importance to all providers that offer triple play.

In addition it can be remarked that zapping time also plays an important role for other communication devices that can operate with a remote control such as DVD players, satellite, or terrestrial receivers, etc.

2. Suggestion for the mapping between zapping time and MOS

In the absence of a mapping between zapping time and MOS one could use the results presented in ITU-T Rec. G.1030 [3] where a mapping is suggested from response and download times to perceived web browsing quality.

Analogous to [3] we suggest a general mapping between zapping time to MOS by defining a minimum (*Min*) and a maximum (*Max*) zapping time and to use a logarithmic interpolation between these extreme zapping times. If we write MOS = *a – b ln(Zapping Time)*, fill in MOS = 5 (excellent quality) for *ZappingTime = Min* and MOS = 1 (bad quality) for *ZappingTime = Max* we obtain for zapping times between *Min* and *Max*:

$$MOS = 4\left(\frac{\ln(ZappingTime) - \ln(Min)}{\ln(Min) - \ln(Max)}\right) + 5 \qquad \textbf{(1)}$$

We choose the values *Min* and *Max* based upon observations about quality perception related to response time reported in [4]:

- 0.1 second is about the limit for having the feel that the system is reacting instantaneously,
- 1 second is about the limit for the user's flow of thought to stay uninterrupted, even though the user

does lose the feeling that the service is operating directly, an important limit for interactive services

- 10 seconds is the limit for keeping the user's attention focused

Based upon these perceptual regions we choose *Min = 0.1s* and *Max = 5 s*. Plugging these values into the Eq. (1) leads to the following mapping between zapping time and MOS, see also Figure 1.

$$MOS=\max\{\min -1.02*\ln(Zapping\ Time)+2.65,\ 5\},1\}. \quad \mathbf{(2)}$$

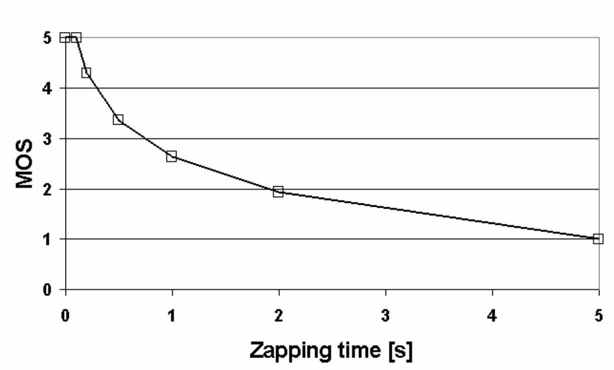

Figure 1 MOS as function of zapping time

We thus propose the following method for determining the perceived quality of channel zapping for a given TV delivery system:

- Measure the Zapping Time for the system
- Determine the MOS by applying the model given by Eq. (2)

In the following section we will describe the outcome of a subjective experiment in a lab environment that has been used to validate the perceptive model suggested in Eq. (2). In the subjective experiments test subjects have rated the quality of zapping time.

3. Experiment

As a test setup, a laptop (Pentium 4, 8 MB, OS Windows XP, screen resolution 1024x768 pixels) was used with a mouse, serving as the TV set, and the remote control device, respectively. In order to assess the quality of zapping times we have built our own tool which is completely based upon HTML. Actually we started out using real video clips (in avi, wmv or mpeg formats) but we found out that it was virtually impossible to realise accurate and consistent zapping times in the sub 200 ms range. Our solution was that we used animated gifs within a web page using layers. First we made 5 animated gifs from video content. Each video clip had a resolution of 720x575 and lasted 10 seconds. Then we constructed a HTML page that contains the five animated gifs in different layers. This means that the gifs are in the HTML page but only one is visible at the time. Switching from one clip to the other just means that another gif becomes visible. Switching between layers in HTML is extremely fast, pretty close to 0 ms. Different switching times are implemented in javascript, so they are easily adjustable. One drawback is that the animated gifs do not contain audio. It is possible to add audio but then we encountered synchronization problems so we left out the audio al together.

The opening page of the experiment is depicted in Figure 2.

Figure 2 Opening page

On the opening page the test subjects can choose between the training part and the actual subjective test. After going through the training part the test subjects were shown five video clips, played in a continuous loop in a local web page (using Internet Explorer 6.0) where only one of the five clips is visible at any moment. Each scene corresponded to a channel, and the user could switch between these five channels, by pressing a button numbered 1 to 5 on the web page. Whilst zapping from one channel to the other, a black screen was displayed, which contained in the right upper corner the number (in an 18 points font) of the target channel.

The experiment did not contain "arrow-up/arrow-down" zapping. However, we expect the results to be comparable as the channel zapping considered in this paper. It is left for further study to validate this claim.

Figure 3 Web page that contains 5 channels

Each scene represented typical TV content: an orchestra scene, two film trailers, a cartoon scene, and a sports scene. Figure 3 shows the content of channel 3 which is obviously a sports scene.
The test subjects consisted of 13 people at TNO ICT in Delft, the Netherlands, and 8 people at Acreo in Stockholm, Sweden. The test subjects varied in age, gender and experience.

3.1 Training part

During the training part the test subjects were offered the three switching times 0, 2 and 5s, corresponding with the shortest, medium and longest switching time in the experiment in order to make them acquainted with the ITU-T Absolute Category Rating scale [5], see Table 1. With a particular switching time selected, the subject was then asked to switch between the channels as often as needed and subsequently to get used to the perceived quality of the zapping time.

Table 1 The Absolute Category Rating scale, used for the assessment of the zapping times

Mean Opinion Score	Explanation
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

3.2 Actual experiment

During the actual experiment, in each zapping time scenario, a random zapping time, unknown to the test subject, was generated from the following values: 0, 0.1, 0.2 (2x), 0.5 (2x), 1.0, 2.0 (2x), 5.0 s. The motivation for using some zapping times twice is to test the consistency of the user responses. The subject was then asked to switch between the channels as often as needed and subsequently to assess the perceived quality of the switching time according to the ACR ITU-T scale given in Table 1. The same procedure was done for the remaining nine zapping time scenarios. For each test subject judging all ten scenarios took about half an hour.
The result of the experiments is shown in Table 2, where the MOS is obtained by averaging the results of all 21 test subjects per zapping time. The complete subjective data set is given in the Appendix.

Table 2 Results of subjective experiments

Zapping time [s]	MOS	Standard deviation MOS
0	4.9	0.30
0.1	4.9	0.36
0.2	4.6	0.50
0.2	4.5	0.51
0.5	3.5	0.60
0.5	3.3	0.48
1	2.3	0.66
2	1.6	0.92
2	2.0	0.71
5	1.1	0.44

Using Table 2 we can validate the perceptive model suggested in Eq. (2). This validation is visualised in Figure 4.

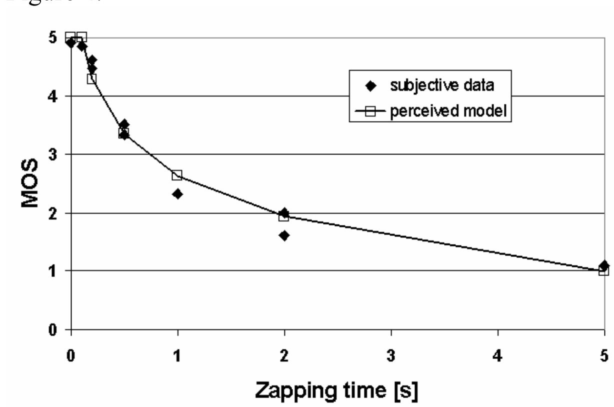

Figure 4 MOS versus Zapping Time

4. Conclusion and further research

It turns out that the correlation between the subjective data and the perceptive model is 0.99 which is very high. In addition, the RMSE (Root Mean Square Error) equals 0.203 while the MCI (Mean Confidence Interval) satisfies 0.234. Thus the evaluation error of our perceptive model for zapping time is less than the statistical reliability of the subjective score, i.e. RMSE < MCI. Therefore we conclude that the perceptive model given by Eq. (2) is very useful for assessing the perceived quality of zapping.

From the model in Eq. (2) we can deduce that in order to guarantee a MOS of at least 3.5, which is considered the lower bound for acceptable quality of service, see [5], we need to ascertain that Zapping Time < 0.43 s.

It should be noted that the logarithmic regression line, shown in Figure 4, is derived under the assumption that *Min = 0.1* and *Max = 5*. If these parameters are different, e.g. due to another contextual situation, then the regression in Eq. (1) still may lead to a good fit. For instance, if we choose *Min = 0.01* and *Max = 3* then the correlation with the subjective data is still high (0.90).

Although the intended target application for our research is IPTV, there was nothing in the approach that was specific for IPTV and therefore our result is not limited to this form of TV delivery. It will also apply to any type of TV e.g. terrestrial and satellite, having longer than instantaneous zapping time.

Operators that offer any form of TV delivery could use the perceptive model (2) to study the impact of certain network settings. If for instance a certain buffer setting leads to a lower zapping time then Eq. (2) can be used to quantify the resulting increase in user experienced quality.

Interesting issues for future research with respect to zapping times include:

- the impact of variation of zapping times
- validation of model in Eq. (1) in case of relaxed ranges for zapping times e.g. 0.5 s – 5 s
- use of video clips with both audio and video
- assessment perceived quality for “arrow-up/arrow-down” zapping
- measurement of zapping times in real-life implementations of IPTV

Acknowledgement

This work is partially supported by the IST-507295 Multi-Service Access Everywhere (MUSE) project, see www.ist-muse.org. The overall objective of MUSE is the research and development of a future low cost, full service access and edge network, which enables the ubiquitous delivery of broadband services. The work reported in the paper is conducted within the MUSE lab trials and demonstration Taskforce (TF4) where extensive effort is put in defining a complete test suite for full-service end-to-end testing.

References

[1] F. Kozamernik & L. Vermaele, “Will Broadband TV shape the future of broadcasting?”, *EBU Technical Review*, April 2005.

[2] Private Communication: Engineer of Imagenio (IPTV of Telefonica) and Dick van Smirren (TV expert at TNO ICT), August 2005.

[3] ITU-T Rec. G.1030, Estimating end-to-end performance in IP networks for data applications, International Telecommunication Union, Telecommunication standardization sector, November 2005

[4] J. Nielsen, “Response Times: The Three Important Limits” (1994). Available: http://www.useit.com/papers/responsetime.html

[5] ITU-T Rec. P.800, "Methods for Subjective Determination of Transmission Quality", International Telecommunication Union, Telecommunication standardization sector, 1996

Appendix

This Appendix contains all Opinion Scores obtained in the subjective experiment, see Table 3.

Table 3 All Opinion Scores obtained in the subjective experiment

	Zapping Time [s]									
Subject	0	0.1	0.2	0.2	0.5	0.5	1	2	2	5
1	5	5	5	5	3	4	2	1	1	1
2	4	5	4	4	3	3	2	1	2	1
3	5	5	5	5	4	3	3	2	2	1
4	5	4	5	4	3	3	2	1	2	1
5	5	5	5	5	4	3	2	2	3	1
6	5	5	5	4	4	3	2	1	2	1
7	5	5	5	4	4	3	2	1	2	1
8	5	5	5	5	4	4	2	2	2	1
9	5	5	4	5	3	3	3	2	2	1
10	5	5	4	4	4	4	3	2	2	1
11	5	5	5	5	4	3	3	2	2	1
12	5	5	4	4	3	3	2	2	2	1
13	5	5	5	5	4	4	2	1	2	1
14	5	4	4	5	3	3	1	1	1	1
15	5	5	5	5	5	4	4	5	4	3
16	5	5	5	4	3	4	2	1	1	1
17	5	5	5	4	4	3	3	2	3	1
18	5	5	5	5	3	4	3	1	2	1
19	5	5	4	4	3	3	2	1	1	1
20	4	4	4	4	3	3	2	2	2	1
21	5	5	4	4	3	3	2	1	2	1