

Model validation of channel zapping quality

Robert Kooij^{1,2}, Floris Nicolai², Kamal Ahmed¹ and Kjell Brunnström³

¹TNO Information and Communication Technology, Delft, the Netherlands

²Delft University of Technology Delft, Faculty of EEMCS, the Netherlands

³Netlab: IPTV, Video and Display Quality, Acreo AB, Kista, Sweden

ABSTRACT

In an earlier paper we showed, that perceived quality of channel zapping is related to the perceived quality of download time of web browsing, as suggested by ITU-T Rec.G.1030. We showed this by performing subjective tests resulting in an excellent fit with a 0.99 correlation. This was what we call a lean forward experiment and gave the rule of thumb result that the zapping time must be less than 0.43 sec to be good (> 3.5 on the MOS scale). To validate the model we have done new subjective experiments. These experiments included lean backwards zapping i.e. sitting in a sofa with a remote control. The subjects are more forgiving in this case and the requirement could be relaxed to 0.67 sec. We also conducted subjective experiments where the zapping times are varying. We found that the MOS rating decreases if zapping delay times are varying. In our experiments we assumed uniformly distributed delays, where the variance cannot be larger than the mean delay. We found that in order to obtain a MOS rating of at least 3.5, that the maximum allowed variance, and thus also the maximum allowed mean zapping delay, is 0.46 sec.

Keywords: video quality, video metrics, standards, multimedia, subjective, objective

1. INTRODUCTION

Traditional telecom operators, cable TV operators and other network providers are eager to improve their offered services in order to offer added value to their customers. Since differentiating in the area of offered bandwidth becomes harder and harder, new approaches are made by these companies in order to guarantee revenues in the future. The most promising way to realize this is by offering additional services such as digital television, probably together with internet access and telephony, into bundled packages. The combination of these three services into one service bundle is called “triple-play”. In order to interest customers for digital television, the overall quality of experience (QoE) of this service should be at least equal to, or better than, the quality they are used to from analogue TV services. For instance from analogue cable TV, people are used to a reliable television service, offering good audio/video quality and almost instantaneous channel switching. Degradation in the overall QoE due to the performance of these important quality measures can only be partly compensated for by means of the extra services offered or by lowering the price of the service.

Offering short channel zapping times (a.k.a. switching time) is one of the most challenging aspects of digital television. A lot of research has been performed in order to reduce zapping delays. However, up till today, zapping times in IPTV cannot compete with zapping times that people are used to from the services over traditional technologies.

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. In Kozamernik and Vermaele (2005)¹ zapping time is called satisfactory if it is below 1 second. This is consistent with the common guidelines for response time of computer systems². Minimum quality requirements for a lot of aspects related to IPTV have been specified by both the ITU³ and the DSL Forum⁴. However in the ITU document there are no recommendations at all related to zapping delay times, while in the DSL forum document it is recommended to limit zapping delay to an arbitrary maximum of 2 seconds. Additionally it is noticed in the document that providers should strive to zapping times in the order of 1 second. The reason for these vague quality requirements is that up till today there is not much knowledge about the human perception of zapping delay times.

To be able to specify minimum levels for the overall QoE requirements for IPTV, knowledge about the relation between human perception and zapping delay times is needed. Moreover this knowledge enables equipment manufacturers and service providers to make well-considered trade-offs, which have to be made due to current technological and/or financial limitations.

This paper report forms a part of series of studies that the authors have performed and will perform in order to get a more complete understanding of the relationship between zapping time and QoE. A more detailed description of the experiments described in this paper can be found in Nicolai (2008)⁵,

2. ZAPPING MODEL

In an earlier paper we showed⁶, that perceived quality of channel zapping is related to the perceived quality of download time of web browsing, as suggested by ITU-T Rec.G.1030. We showed this by performing subjective tests resulting in the model shown below, see Figure 1 and the equation below.

$$MOS = \max\{\min\{-1.02 \cdot \ln(\text{Zapping Time}) + 2.65, 5\}, 1\}$$

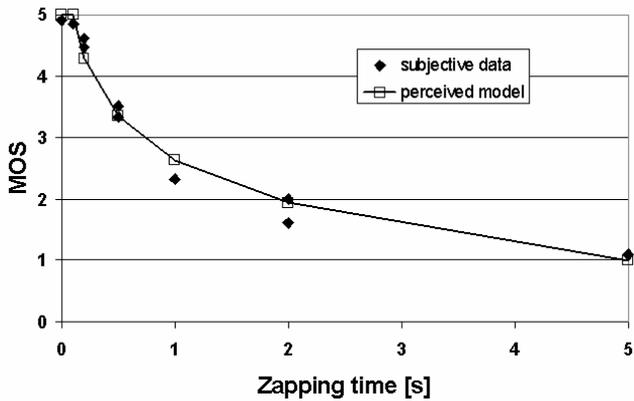


Figure 1: Model and the MOS of the subjective data

The fit was excellent with a 0.99 correlation between subjective data and the model.

In the earlier study we assumed “fixed delay times” in channel zapping i.e. the switching time did not vary from time to time. Five different video clips were shown as animated gifs within a web page in different layers. Each video clip had a resolution of 720x575 and lasted 10 seconds. Switching between layers in HTML is extremely fast, pretty close to 0 ms. Different switching times were implemented in javascript, so they were easily adjustable.

Test persons were asked to take place behind a laptop. On this laptop they could switch between five channels by clicking with the mouse on the buttons 1 to 5 on the computer screen, see Figure 2. We refer to this as “lean forward zapping”. However, this is not the way people usually watch TV. Would the quality be perceived differently if a user sits in a chair using a remote control? We will call this “lean backward zapping”. Additionally a question was whether there is a difference in perception if people use “arrow up/arrow down” zapping instead of “channel by number” zapping. Furthermore, it is often the case that the zapping time is not constant. Therefore, it is also useful to obtain an understanding of how people perceive variation in zapping times.

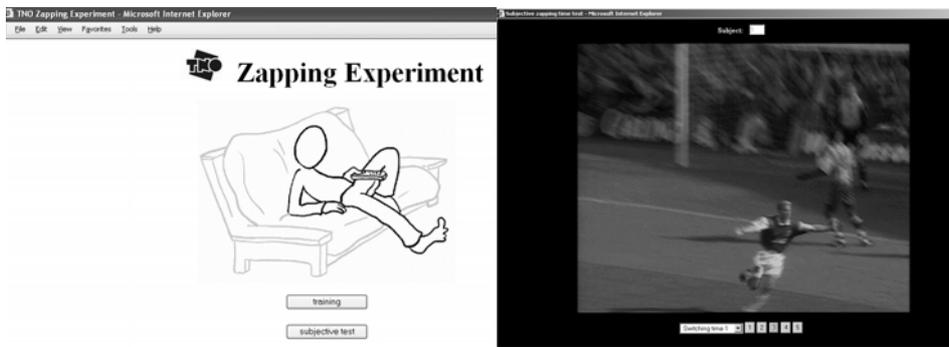


Figure 2: Opening page (left) and web page that contains 5 channels (right).

3. METHOD

Three experiments with different settings for switching times were conducted, using 25 test persons. These tests were carried out in two sessions. During the first session, the tests for the first two experiments were carried out. There was a break of approximately ten minutes between the subsequent tests. After this session the subjects were asked to fill in a short questionnaire. The last experiment was conducted during a second session, about two weeks later.

3.1. Experimental setup

The test setup for the “lean backward” experiments was selected to be comparable to the test setup in the “lean forward” tests. Although the laptop used was newer than that used in the earlier experiments, the screen size and screen resolution of the laptops were equal, namely 1024x768 pixels. Also the same tool for playing video and simulating channel switching was used as in the earlier tests, described briefly above or in Kooij et al (2006)⁶. However, it was extended with remote control functionalities, which was implemented using Bluetooth’s Human Interface Devices (HID) protocol on a mobile phone (Figure 3).

The experiments were conducted for most of the subjects in a lab with “living room” conditions at the TNO office in Delft. For some subjects the tests were done at other locations with a surrounding similar to the living room at the TNO office. These tests were also performed in “comfortable seats”, the distance of the display was approximately 5H (80cm) and lighting conditions were adjusted to be similar to lighting conditions in the TNO office. Furthermore the test equipment was identical in all experiments; this means that the same laptop and remote control were used in all tests. Finally, the brightness level of the display was identical in all experiments and test subjects could adjust the vertical angle of the display to obtain the best possible image quality.



Figure 3: Mobile phone used as remote control

3.1.1. Zapping with number buttons and fixed delay time

The subjects were seated in a comfortable sofa/chair (“lean backward”). A remote control (mobile phone) was used to switch between channels. The rest of the experiment was analogous to the earlier experiments.

The training part consisted of three different switching times. These were 0, 2 and 5 sec corresponding with the shortest, medium and longest switching time in the experiment. For each of the three zapping times selected, the subjects were asked to zap between channels as often as needed to get familiar with typical zapping times as well as with the ITU-T 5-point Absolute Category Rating (ACR) scale⁷ i.e. Excellent, Good, Fair, Poor and Bad as shown in Table 1.

In the actual experiment the subjects assessed ten different scenarios for the zapping time. The ten different scenarios represented a randomly ordered sequence of the following zapping times: 0, 0.1, 0.2 (2x), 0.5 (2x), 1, 2 (2x) and 5 s. These times are equal to those used in the previous “lean forward” performed experiment. Some zapping times were repeated to enable testing the consistency. For each option the user could zap as often as needed by pressing the numbers 1 to 5 on the remote control that corresponded to the channel that the subject wanted to switch to. Judging all the zapping times in one test took about ten minutes for each test subject.

3.1.2. Zapping with arrow up/down buttons and fixed delay time

In the next experiment the subjects were asked to zap using “arrow up/arrow down” buttons (the upper left- and right keys on the mobile phone), instead of the number buttons. The same setup as in the “lean backward” tests above was used with the same rating scale.

For zapping between non-adjacent channels we used the implementation which is also the most common one in STBs today. The TV set or STB starts fetching a “new channel” as soon as it gets a “channel switching” command from the remote control. The “new channel” will be displayed as soon as it is available unless another “push of a button” has been detected. In that case the STB immediately starts fetching the “new channel” which corresponds to the “last pushed button”. For instance, if a user wants to switch from channel 1 to 5 he can just push the “channel up” button 4 times after another. And, if it takes a fixed time of one second to fetch a new channel, then the total zapping time would be equal to the time to push the button four times plus the one second fetching delay (ch1 -> push button 4 times -> 1 sec -> ch5), as illustrated in Figure 4.

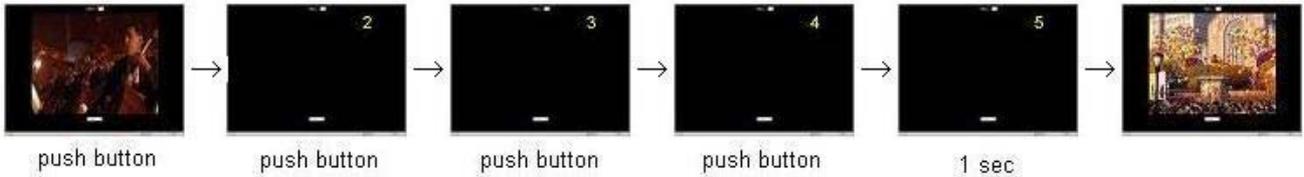


Figure 4: Implementation of arrow up/arrow down zapping

The training part of this experiment consisted of the same three zap times as those used above.

In the actual experiment test subjects were asked to follow the same procedures as in the previous experiments. The same zapping times was also used. However, test subjects were asked to zap both channel-by-channel as well as skipping several channels, i.e. to push the button several times after another in order to zap directly from e.g. channel 1 to channel 5.

3.1.3. Zapping with “variable delays”

We wanted also to investigate how people perceive channel zapping if the delay varies from time to time.

The total zapping delay consists of the sum of various contributions from different sources. Each individual source contributes some delay according to some distribution function. For example, assume that I-frames arrive at an STB at fixed intervals. The delay that is introduced due to waiting time for an I-frame depends on the exact moment that a switch between channels is initiated. If this happens right before the arrival of an I-frame the waiting time will be very short because the new I-frame will arrive almost immediately at the STB. However, if a channel switch is initiated right after the arrival of an I-frame, then the STB has to wait almost a whole interval for a new I-frame to arrive. Since the inter I-frame delay are usually dominant, we expect the inter I-frame delay to be approximately uniformly distributed.

Since we want to obtain a relation between user perception and the variance in the delays, it is necessary to select multiple values for the variances. Only in this way it might be possible to observe a relation between the two. In order to keep a limited amount of different options (under 20) in the experiment, it was decided to use only three different mean delay times (denoted by μ), that is $\mu = \{0.5, 1, 2\}$ sec. For each mean time a “sufficient” number of settings for the variance (var). These are: $\mu = 0.5$ sec, var = $\{0, 0.2, 0.5\}$ sec; $\mu = 1$ sec, var = $\{0, 0.2, 0.5, 1\}$ sec; and $\mu = 2$ sec, var = $\{0, 0.2, 0.5, 1, 2\}$ sec.

Similar to the previous experiments this experiment was preceded by a short training session. During the training session, test subjects were able to get familiar with the typical settings for the switching times that could be expected during the experiment. Additionally it gave users the possibility to get familiar with ITU-T 9 point ACR scale (ITU-T Rec. P.910), shown in Table 1.

Table 1: ITU-T 5 and 9-point scale

5 point scale	9 point scale	Perceived quality
5	9	Excellent
	8	
4	7	Good
	6	
3	5	Fair
	4	
2	3	Poor
	2	
1	1	Bad

Since in total there were twelve different settings for the zapping behaviour in the actual experiment, it was decided to use a higher resolution scale in comparison with the previous experiments. To convert the 9-point scale values to the 5-point scale values the following formula is used:

$$\frac{9\text{-point scale value} + 1}{2} = 5\text{-point scale value} \quad (1)$$

For example; a rating of 8 on the 9-point scale leads to a $(8+1)/2 = 4.5$ on the 5-point scale.

During the training session test subjects were offered four different settings for the “zapping behaviour”. The four different settings were: $\mu = 0$ sec, var = 0 sec; $\mu = 0.5$ sec, var = 0.5 sec; $\mu = 1$ sec, var = 1 sec; and $\mu = 4$ sec, var = 0 sec. Again test subjects were asked to zap as often as they felt necessary in order to get familiar with the different settings and with the ITU-T 9-point ACR scale.

During the actual experiment, test subjects were asked to asses 15 different settings for the “zapping behaviour”. The fifteen different options in the experiment represented a randomly ordered sequence of the settings combinations of mean delay and variance listed above. In this randomly ordered sequence, the combinations ($\mu = 1$ sec, var = 0 sec), ($\mu = 1$ sec, var = 1 sec) and ($\mu = 2$ sec, var = 2 sec) were presented twice

Zapping was possible by using the “arrow buttons” on the mobile phone.

4. RESULTS

4.1. “Lean back” zapping experiment

In the table to the left in Figure 5 the results of the “lean backward tests” are shown in numbers. The MOS values are an average over 23 scores. Two of the scores from in total 25 “test subjects” were excluded because one had a low correlation with the overall MOS and the other had not followed the procedure.

The results are visualized in Figure 5. For the zapping times which were used twice in the experiment, we took the average of both MOS values. Notice the apparent logarithmic relation between the MOS values and the zapping delay times. This is in accordance with the results from the initial “lean forward” experiment.

zapping time [s]	MOS (LB)	Standard deviation
0	5.0	0.21
0.1	4.7	0.47
0.2	4.6	0.50
0.2	4.6	0.73
0.5	4.0	0.64
0.5	3.8	0.65
1	2.9	0.76
2	2	0.74
2	1.9	0.63
5	1	0.21

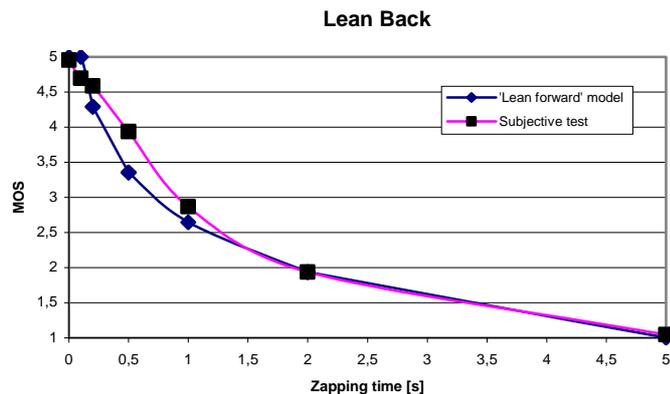


Figure 5: Results of the “Lean backward” zapping using “number buttons”

4.2. “Arrow up/arrow down” zapping experiment

In the table to the left in Figure 6 the results of the “arrow up/arrow tests” are shown. The MOS values are an average over 23 scores. Two of the scores from in total 25 “test subjects” were excluded for the same reasons as mentioned above. A plot of the results is shown in Figure 6. The figures show very similar results as those obtained from the “lean backward” “number zapping” experiment.

zapping time [s]	MOS	Standard deviation
0	4.9	0.46
0.1	4.7	0.47
0.2	4.5	0.58
0.2	4.4	0.51
0.5	3.9	0.63
0.5	3.8	0.54
1	2.8	0.69
2	2.1	0.63
2	1.8	0.65
5	1	0

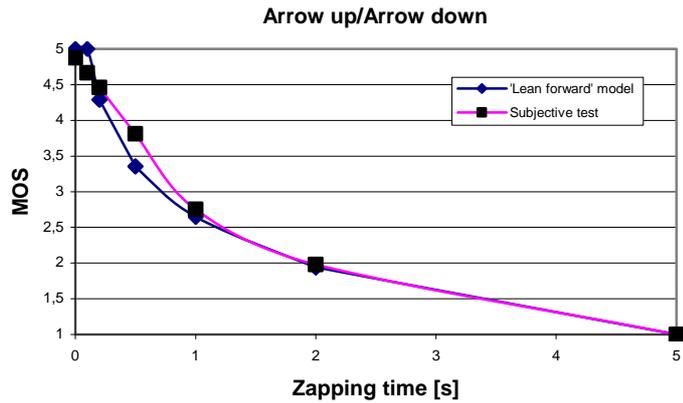


Figure 6: Results of the “arrow up/arrow down” zapping experiment

4.3. “Variable zapping delay” experiment

The MOS scores that were obtained from the “variable delay experiment” are shown in the table to the left in Figure 7. Also the standard deviation in the opinion scores for each of the zapping time settings is included in the table. The MOS values are the averages of the ratings from 25 test subjects.

Mean time [s]	Variance	MOS	Standard deviation
0.5	0	7,84	0,99
0.5	0.2	7,64	0,86
0.5	0.5	7,04	1,10
1	0	5,88	1,54
1	0	6,12	1,42
1	0.2	5,8	1,61
1	0.5	5,8	1,19
1	1	5	1,38
1	1	5,08	1,53
2	0	3,32	1,57
2	0.2	3,32	1,44
2	0.5	3,04	1,17
2	1	3,2	1,55
2	2	2,68	1,35
2	2	2,88	1,81

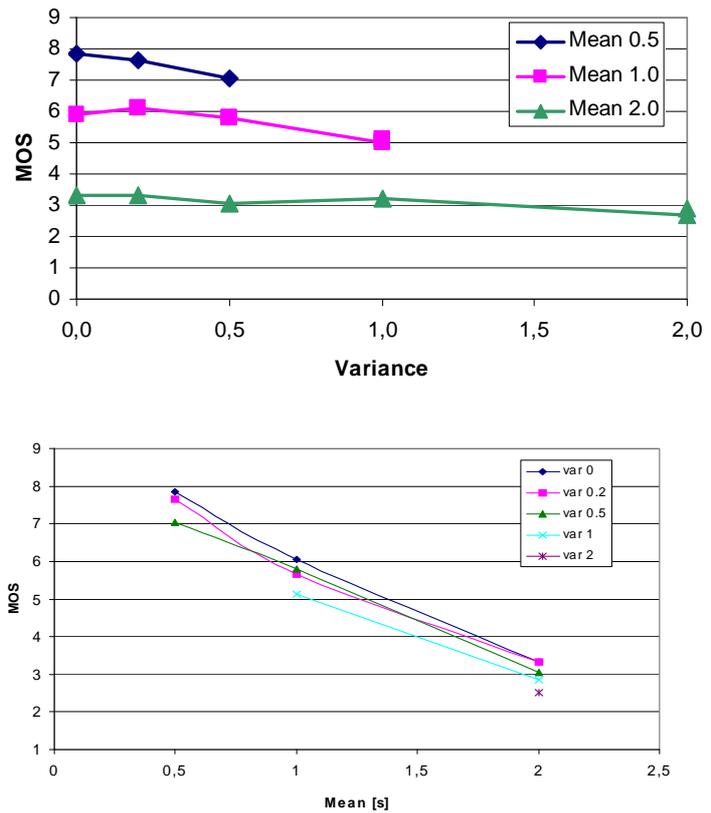


Figure 7: Results of the “variable delays” experiment

The results are visualized as two plots in Figure 7. In the top most plot the results for each of the mean times against the amount of variation in the switching times is shown. The three different graphs do not show a clear relation between

MOS scores and the amount of variation in the switching time. It can be observed from the figure that the MOS scores show a decrease of about 1 point, between zero variation and the maximum variation, for each of the mean switching times. The actual decrease of the MOS values at maximum variation are 0.80 for mean 0.5 sec, 0.92 for mean 1 sec and 0.80 for mean 2 sec. The average decrease of the MOS value at maximum variance is 0.84 points.

The bottom plot in Figure 7 shows the MOS for different amounts of variance in the switching times against the mean delay in the switching times. Note the apparent linear relation between the MOS values and the mean switching times. This contradicts with the results that we obtained from the former experiments, from which an apparent logarithmic relation could be observed. It could further be noted in the “variable delays” experiment that regardless of the amount of variance in the switching times, the MOS values for each of the mean delay times are higher than in the former experiments. This also holds after conversion of the 9-point scale MOS values to 5-point scale values.

5. DISCUSSION

In Figure 8 the results of the three “fixed delay” experiments are plotted together. Two important observations can be made from this figure. First of all the correspondence between the “lean back” (LB) number zapping experiment and the “arrow up/arrow down” (UD) is very apparent in this graph. The correlation between the obtained MOS values from these two experiments is higher than 0.99. It can be concluded from this observation that apparently people do not perceive zapping delays different whether they are using number buttons or arrow buttons for zapping.

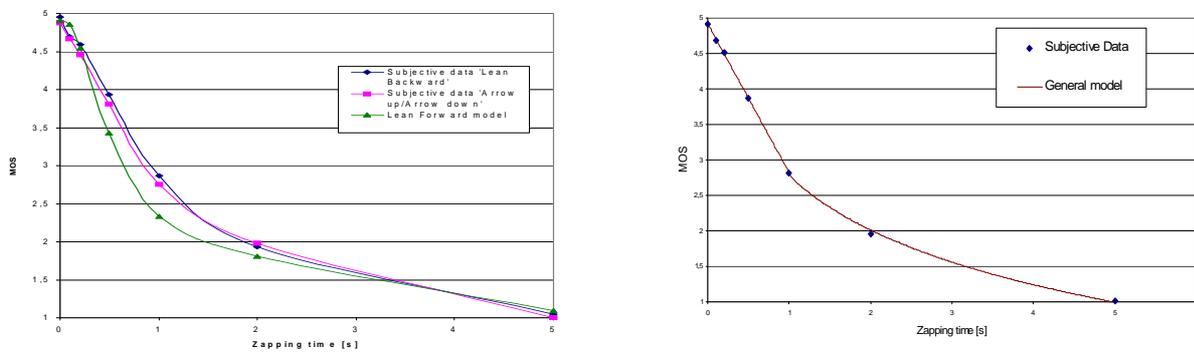


Figure 8: To the left, results from all three “fixed delays” experiments and to the right the general zapping model (fixed delays only)

Another important observation that can be made from the figure is the difference between the graphs for the “remote control” experiments and the initial “lean forward” (LF) experiment. If we assume that both tests were carried out under similar circumstances, or at least that differences in the circumstances did not have much influence on the test subjects, we can draw an important conclusion from these results: people are more tolerant to zapping times when in a “lean backward” position compared to the case where people are in a “lean forward” position.

A possible explanation for this could be that people are more focused if they are in a “lean forward” position. In the “lean forward” tests people were positioned as if they were working behind a computer at a desk. In the lean backward test, subjects were using a “wireless” remote control for zapping. Because of this, people were able to position themselves in a much more comfortable and relaxed way.

A second explanation for the different results could be that different intermediate times have been used in training sessions of the “lean forward test” and the “remote control tests”. In the initial “lean forward” test an intermediate time of 0.5 seconds was used during the training period. This is in contrast with the intermediate time of 2 seconds that was used in the training session of both “lean backward” experiments.

Since in practice, zapping is primarily done with the use of a remote control, we can use the results of both “lean backward” experiments to generate a “general zapping model for fixed delays”. For this purpose we took the average over the combined data of both lean back experiments, which is shown to the right in Figure 8 as “subjective data”. With the use of regression analysis we retrieved a model that shows a correlation higher than 0.99 with the subjective data. The Root Mean Square Error (RMSE) between this model and the “subjective data” is 0.038. The model looks as follows:

$$\text{MOS} = \begin{cases} -2.1 * x + 4.92 & , \text{ if } 0 \leq x < 1.04 \\ -1.11 * \ln(x) + 2.78, & \text{ if } 1.04 \leq x < 4.97 \\ 1 & , \text{ if } 4.97 \leq x \end{cases} \quad (2)$$

where x denotes the zapping time in seconds, see Figure 8 (right).

To obtain the MOS rating for a certain zapping delay from this model, one should fill in the value of the “mean” zapping delay time.

From 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 ITU-T Rec.P.800, we need to ascertain that Zapping Time < 0.67 sec!

This is a noteworthy result, since if we compare this to the maximum tolerable delay of 0.43 s that was obtained from the “lean forward” zapping experiment, it can be concluded that the constraints on zapping delays are less stringent in the case of “lean backward” zapping.

In Figure 9 a graph is shown of the previous obtained data from the “fixed delays” experiments together with the results from the variable delay experiment. For this purpose the results of the “variable delays” experiment were converted to a 5-point rating scale with the use of Eq. (1). From this figure it is clear that people were not able to grade zapping times very consistently when the settings of the experiment were largely changed. In all experiments there were settings for the zapping times with 0.5 s, 1 s and 2 s mean delay with zero variance. However, it can be noted from the figure that test persons gave higher ratings in the “variable delays” experiment in comparison with the ratings in the first experiments. The reasons for these differences are not very clear. A possible explanation could be the different rating scale that was used in the “variable delay” experiment. Especially in The Netherlands people are used to a 10-point rating scale used in the Dutch education system. It could be that people unconsciously tend to use this 10-point rating scale. Since this leads to higher ratings this could partially explain the higher MOS values in the “variable delays” experiments. Also people might be more cautious to give the really low ratings in the 9-point scale. This would also lead to a bias in the MOS values of the “variable delays” experiment.

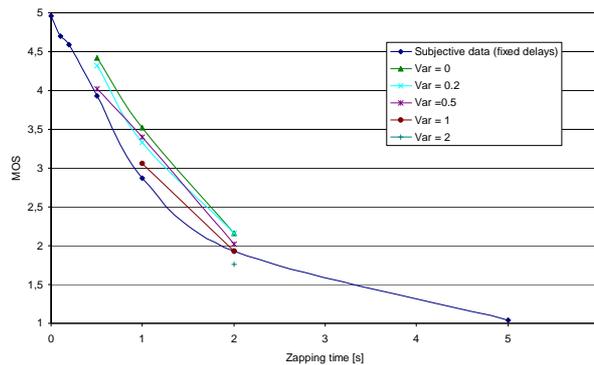


Figure 9: Combined graph of “general model for fixed delays” and “variable delay” results

However, even if the previous mentioned reasons are the cause for the higher ratings in the “variable delay” experiment, there is still no explanation for the fact that the first experiments led to a more or less “logarithmic relation”, while the “variable delay” shows a linear relation between MOS and zapping delays.

A possible explanation for the observed difference could be the choice of the switching times and the rating scale in the first experiments. The switching times that were used in the first experiment were more or less exponentially increasing, while the rating scale was linear with only 5 different possible values. The first three options, 0 s, 0.1 s and 0.2 s are quite close to each other. The other four options were differing a lot more. It could be that since the difference between the four slower options was more obvious, that people tend to rate these options with the scores 4, 3, 2 and 1, while the first options are mostly graded with a five or occasionally a four. This grading of the switching times would automatically lead to a more or less logarithmic relation between MOS and zapping time. On the other hand, this effect should be compensated since the results are averaged over 25 test persons. Additionally, previous research that was done

on perception of web browsing, see ITU-T Rec.G.1030, also led to a logarithmic relation between MOS and waiting time.

With the results from the various experiments it is possible to compose a “general zapping model” where both mean delays and variance, are included. The “general zapping model” with variance included is based on the “general model for fixed delays”, and thus only applies for the case of “lean backward zapping”. This “general model for fixed delays” defines the relation between MOS and “fixed switching times”.

We use the results from the “variable delays” experiment to extend the model to include variance in switching times. For this purpose we use the fact that the MOS ratings in this experiment show an average decrease of 0.84 points (on the 9-point scale) at maximum variance. In our final “general zapping model” we therefore define the MOS rating at maximum variance to lie 0.42 points (on the 5-point scale) below the zero variance ratings.

For our “general zapping model” we assume a linear relation between the decrease of the MOS rating and the ratio between the actual amount of variance and the maximum amount of variance.

Finally, we should take into account that at very low mean times a decrease in MOS of 0.42 does not seem very plausible. For example: it is not really likely that for a mean of 0.1 s, the MOS will decrease with 0.42 points when the variance is 0.1 s.

For this reason we define the decrease in the MOS rating at maximum variance to increase proportionally with the mean time, for mean times up to 0.42 s. This implies that for mean times lower than 0.42 sec, the decrease of MOS will be less than 0.42 for maximum possible variance.

This definition allows us to complete our “general zapping model”. The model is visualized in Figure 10.

$$\text{Decrease of MOS} = \begin{cases} \text{variance} & , \quad \text{if } \text{mean delay} < 0.42 \\ \frac{\text{variance}}{\text{mean delay}} * 0.42 & , \quad \text{if } \text{mean delay} \geq 0.42 \end{cases} \quad (3)$$

Finally, since the lowest possible MOS score is limited to one, the final MOS rating is obtained by filling in the results of Eq.(2) and Eq.(3) into the following formula:

$$\text{MOS} = \max\{\text{Eq.(2)} - \text{Eq.(3)}, 1\} \quad (4)$$

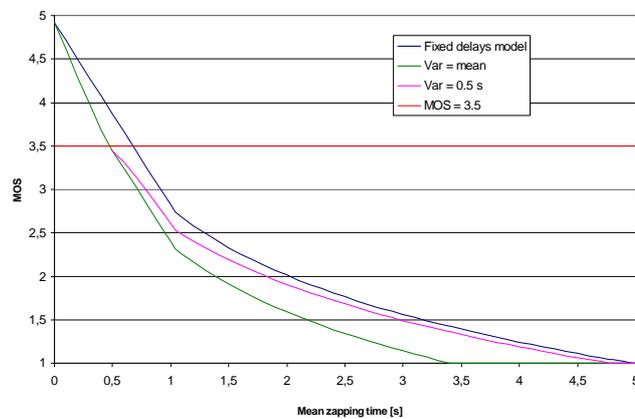


Figure 10: General zapping model (variance included)

For illustration purposes the line for a variance of 0.5 sec is included in the figure, note that this line is not defined for mean zapping times lower than 0.5 sec.

Also a line is drawn where the MOS equals 3.5. In order to guarantee a MOS of at least 3.5, the combination of the average delay and the amount of variance should lie in the area above the “MOS 3.5. line” and is bounded by the two graphs for “zero variance” and “maximum variance”. The “MOS 3.5 line” crosses the “maximum variance” graph at 0.46 sec and the “zero variance” graph at 0.67 s.

6. CONCLUSIONS

It can be concluded from both “lean backward” experiments, namely the “number zapping” experiment and the “arrow zapping” experiment, that there is no difference in the perception of switching times between both zapping methods. The data obtained from the “number zapping” experiment and the data obtained from the “arrow zapping” experiment shows a correlation higher than 0.99. Furthermore it can be concluded, under the assumption that the different training periods did not have an influence on the grading of our test persons, that people are more tolerant to switching delays when they are in a relaxed position. This conclusion is supported by the significant higher rating of zapping times in the “lean backward tests” in comparison with the rating of zapping times in the “lean forward tests”.

With the use of regression analysis on the combined data from both “lean backward” experiments we obtained a “zapping model for fixed delays”. To obtain the MOS rating from this model one should fill in the “fixed” zapping delay time. The model looks as follows:

$$\text{MOS} = \begin{cases} -2.1 * x + 4.92, & \text{if } 0 \leq x < 1.04 \\ -1.11 * \ln(x) + 2.78, & \text{if } 1.04 \leq x < 4.97 \\ 1, & \text{if } 4.97 \leq x \end{cases}$$

where x denotes the zapping time in seconds.

From this model we can deduce that in order to have a MOS rating higher than 3.5, which is considered the lower bound for acceptable quality of service, the zapping time needs to be lower than 0.67 sec.

The results of the “variable delay” experiment show, that for equal average delay times, the perception becomes lower when the amount of variance increases. The results show that, for higher mean times, the average decrease in MOS is about 0.42 (on the 5-point ACR scale) if the amount of variance becomes equal to the mean zapping time. For the composition of our “general zapping model” we assumed a negative linear relation between the MOS and the amount of variance as a percentage of the mean delay time. However for mean zapping times lower than 0.42 sec with maximum variance, we assumed the decrease in MOS to be bounded by the mean delay time. This resulted in the “general zapping model”, visualized in Figure 10, where both fixed and varying zapping delay times are included. It takes two steps to obtain the MOS rating for zapping times with a given mean delay and a given variance. First we assume zero variance and we calculate the MOS for the mean delay according to the “model for fixed delays”, shown above. The second step is to subtract the outcome of the following formula from the calculated MOS:

$$\text{Decrease of MOS} = \begin{cases} \text{variance}, & \text{if } \text{mean delay} < 0.42 \\ \frac{\text{variance}}{\text{mean delay}} * 0.42, & \text{if } \text{mean delay} \geq 0.42 \end{cases}$$

Finally, since the lowest possible MOS score is limited to one, the final MOS rating is obtained by filling in the results of Eq.(2) and Eq.(3) into the following formula:

$$\text{MOS} = \max \{ (\text{Eq.(2)} - \text{Eq.(3)}), 1 \}$$

From this model we can deduce that the maximum acceptable average delay should be less than 0.46 sec when the variation in the delay is equal to the mean delay. In case that the variation in the delay reduces to zero than the constraint on the maximum tolerable delay is relaxed to 0.67 sec.

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

- Distribution of zapping times in real life digital TV implementations; Since in real (digital) TV implementations zapping delay times might be different from uniformly distributed.

- Impact of varying zapping times that are non-uniformly distributed; If real (digital) TV implementations show different zapping time distributions the impact of this should be investigated.
- Zapping experiments where both audio and video are present in the content (possibly with a sync delay between audio and video); In reality, audio is included in the video stream. Zapping delays for the audio part of the content might show different zapping delay times than the video part of the content.

7. ACKNOWLEDGEMENT

This work is partially supported by the CELTIC Project RUBENS. The objective of RUBENS is to define and evaluate a new broadband access infrastructure that offers personalized Quality of Experience in a flexible and scalable way for a large variety of applications, delivery models, and devices, see also <http://wiki-rubens.celtic-initiative.org>. The fourth author's work has been financed by VINNOVA (The Swedish Governmental Agency for Innovation Systems), which is hereby also gratefully acknowledged.

8. REFERENCES

1. Kozamernik, F. and Vermaele, L., "Will Broadband TV Shape the Future of Broadcasting?", European Broadcasting Union, (2005)
2. Nielsen, J., Response Times: The Three Important Limits [on-line], <http://www.useit.com/papers/responsetime.html>. Accessed: 30 June 2008
3. ITU-T, "Quality of Experience Requirements for IPTV", FG IPTV-DOC-0118, ITU Focus group on IPTV. Working Document, http://ties.itu.int/ftp/public/itu-t/fgiptv/readonly/Previous_Meetings/20070723_Geneva/Documents/FG%20IPTV-DOC-0118e.doc (visited Jan 2009) , (2007)
4. DSL Forum, "Triple Play Services Quality of Experience (QoE) Requirements and Mechanisms", Technical Report TR-126, <http://www.dslforum.org/techwork/tr/TR-126.pdf> (visited Jan 2009), (2006)
5. F.P. Nicolai, "Perceived Quality of Zapping: Impact of remote control and variance ", TNO Report 34068, (2008)
6. Kooij, R., Ahmed, K., and Brunnström, K., "Perceived quality of channel zapping", *Proc. of 5th IASTED International Conference on Communication Systems and Network, August 28-30, 2006*, (2006)
7. ITU-T, "Subjective Video Quality Assessment Methods for Multimedia Applications", ITU-T Rec. P.910, International Telecommunication Union, Telecommunication standardization sector, (1999)