Slider or glove? Proposing an alternative quality rating methodology

S.Buchinger, W.Robitza, P. Hummelbrunner, M.Nezveda, M.Sack
H.Hlavacs
University of Vienna, Faculty of Computer Science, Austria
Department of Distributed and Multimedia Systems

Introduction

Continuous subjective multimedia quality assessment is generally performed by using a slider [1] based on a continuous scale ranging from 0 to 100. The context of use of such a slider is limited because it needs to be positioned on a horizontal surface to enable a one handed manipulation.

A glove containing motion sensors at different points of the hand represents a suitable alternative for rating multimedia quality continuously. After some initial calibration the maximum value of 100 can be rated when the hand is fully open while 0 is represented by a closed fist. In contrast to the slider, where it happens that test persons check the slider position by a glance on the device for a certain amount of time, the glove itself does not lead to any distraction. Furthermore scores rated with the glove generally result to be as precise as the slider. In some cases when using a highly reduced scale (to five values only), this precision cannot be reached.

Another important aspect consists of the fact that the average time needed to perform a rating is lower when using the glove instead of the slider.

The user satisfaction seems to be highest for the non hidden slider with the possibility to check the current position at the cost of missing a not ignorable part of content. However, the glove has still been judged to be a good measurement method.

Investigation

To compare the rating methodologies under investigation we have set up an experiment: Two different videos containing a black background and numbers in the foreground produced with a random number generator in the range between 0 and 100 with a time interval of 5 seconds have been shown to the user. Video 1 contains only numbers of the set {0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100} and video 2 only numbers of the set {0, 25, 50, 75, 100}.

Each time a new number appears on the screen the user had the task to find the rating value displayed on the screen,

Figure 1: Rating with the slider

Figure 2: rating with a hidden slider

Figure 3: Rating with the glove

with the slider visible on the table as depicted in Figure 1, with a hidden slider as shown in Figure 1 and with the glove as illustrated in Figure 3.

During the ratings performed with a visible slider, the distraction time caused by glancing to the slider has been roughly measured by using a simple stoppwatch. Until now, six valid test persons participated at this user studies and they were additionally asked to rate the measuring methods

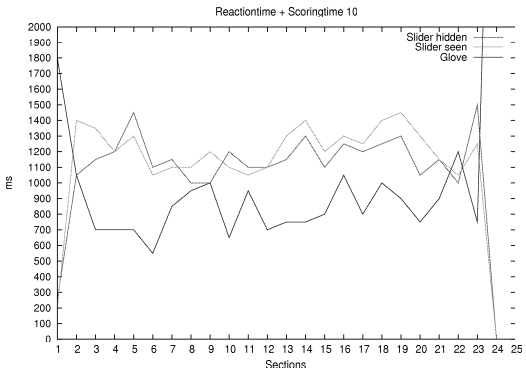

Figure 4: Average rating delay

by using the categorical scale of five grades: excellent, good, fair, poor, bad as well as to add some justification for their choice.

Data analysis

Collected data of all test persons has been batched first in time intervals of 50 ms and then averaged in order to obtain values for one single fictive test persons. Excluding the start and the end of each experiments where ratings tend to be more erratic, the standard deviation is below 10, hence 10% of the available scale for the video containing only numbers of the set {0, 25, 50, 75, 100} and around 15% for the video containing only numbers of the set {0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100}

The obtained data has been analysed according to the following criteria:

Rating delay: time needed to perform the rating. It starts when a new number is displayed and ends when the rating value stabilizes.

Precision: the deviation of the rated score to the displayed value, i.e., the absolute value of the difference between the value displayed on the screen and the average performed rating in this interval.

Distraction period: amount of time spent looking at the slider instead of the video.

User satisfaction: results obtained by analysing the questionnaires.

Results

The rating delay results to be much lower when using the glove. Figure 4 depicts the average rating delay of all test persons observing the video 1. The curve representing the glove is situated widely below the curves depicting the slider ratings. This effect is even increased for video 2.

The precision of the glove is ranged between the hidden and the open slider for video 1 at and average deviation of 9.44, but unfortunately for video 2 ratings do not seem to be precise for this sample of test users. The slider contains a scale reaching from 0 to 100 and the general aim is to rate continuously, therefore we estimate that video 2 does not fully represent the focus of the slider rating.

Figure 5: Average rating delay

The distraction period highly depends on the character of the user. Comments taken from the questionnaires reveal that persons who have a high desire to be precise have a much higher distraction time than more intuitive persons.

Furthermore the glove seems to be suitable for outdoor experiments that may be needed in addition to test performed in the lab to meet QoE requirements such as the awareness of the customers context and expextation. For example mobile multimedia is consumed on the way as well as at home. Figure 5 should give an idea about how outdoor tests could be realised.

Technical details

The glove used for the experiments is a model by the manufacturer 5DT Technologies. There are various models available for right and left hands, yet for the tests, only the left handed version has been used. Models come with different sensor counts. The left hand glove features five sensors, whereas the right hand glove uses a total of 14 sensors (two per finger).

The glove values can be read by USB or serial connection with the appropriate software. For the experiments, a software was written in C+ +, using the supplied glove software development kit (SDK). Sensor values are polled in intervals which can be set by the system's sleep () function.

1. REFERENCES

[1] ITU-R BT.500-11. Methodology for the subjective assessment of the quality of television pictures. ITU-R BT.500-11, 2002.