 VQEG Meeting Minutes - Stockholm July 2014

# Monday AM, July 7

## Group Updates

### more

with some text

# TUesday AM, July 8

Short discussion about the possible next meeting. INTEL has offered to host in Carlifornia. Suggestion to put it together with the SPIE Electronic Imaging, but this year this was a problem with people not being able to stay away to long. Together with VPQM this could be almost 3 weeks. VPQM is not decided yet.

## QART

### Overview of QART

Mikolaj gives an overview of the QART project, see presentation. QART targeting Task Recognition Video (TRV). This is partially standardized in P.912, but this is not sufficient based on the results in the project.

Can you force people to describe their experiment? Could that be put into the recommendation? Maybe not limit it to licence plate recognition. Description of the particular task at hand is more important. Ask just one thing at the time and simple questions. Also take into account the training of the subjects. Also study how professionals are using the videos in practise e.g. doctors. In sign language studies are limited to a small subset and then extrapolate to a larger set. This may not be appropriate. The standards says that entire ranger. Does that means 6 sigma or 9 sigma or. This should possibly be clarified.

Non clear paragraph should be clarified. Is the camera parameters specified. That may be necessary to clarified. Action point to call for input and get feedback from expert eg Qualcomm

Impact of position of choice button. This is not specified in the standard. Upper right corner was used more often when people of uncertain. But it may also have been the selection of the alternative in the list. If the alternatives are not equally likely the experiment is biased.

Comparison scale with “about the same” is abused even if people can identify. Should warn against “unsure”. Leaving out “unsure” may give you more accurate data, but you need equal probability of alternative. A solution could be to introduce a confidence scale, but psychologist are using forced-choice so that would most likely be the best route. Pre-test could confirm the distribution.

Possibility to use Fuzzy logic. Specify the task first then the analysis can be specified.

Give input to QART if you are not happy about P.912.

People are not motivated. What is the solution. Idea gamification. Payment is one solution and motivating the subjects another.

Suggestion to update P.912 to refer to P.913 instead of P.910.

Statistical analysis should be specified in the Recommendation. Maybe also put in post-screening rules.

A more detailed document has been submitted to VQEG. The plan is to submit to ITU meeting in September 2014. Try to reach out to those who use the P.912 in advance to get them involved in the process, which may help not giving problem in the future.

## AVHD

How should AVHD proceed in the audio-visual part? Collaborative is suggested as the best. Reference best practise is also suggested.

**Decision:** will go for collaborative approach

How should HD proceed? Most are working on adaptive streaming among proponent, so what is the real interest in a pure HD test. Is there any interested in a quick HD test. The idea is to go for UHD

**Decision:** Tabled validating HD video only models

HD method for new codecs (VP, HEVC..) may be validated within the UHD project

Going through the test plan of audio-visual testing.

Editing is done directly in the test plan. No decisions is taken at this point, rather the full document will be approved when it has reached a stable state.