Multimedia Quality of Experience for Target Recognition Applications

Mikołaj Leszczuk and Lucjan Janowski

Introduction

A decade ago, the telecommunications industry believed that high-performance Quality of Service (QoS) techniques would resolve any recurrent problems of low-quality multimedia services. However, within a few years, it became clear that optimization of QoS parameters such as throughput, packet loss, delay, or jitter is not the best way of improving the quality experienced by users. The problem of low bandwidth can be compensated for by more efficient codecs. The impact of packet loss is strongly dependent on their distribution, and the use of redundancy coding and transmission. For many applications, buffering multimedia data streams can alleviate major delays and jitter.

Since discovering that QoS is not an adequate metric of network quality, most proposals have suggested that quality should be measured at the user level. This process was named Quality of Experience (QoE) [1] [2]. Such measurement calls for special structures (frameworks) for integrated assessment of the quality of video sequences [3]. These structures are increasingly being filled with solutions that attempt to model overall quality, operating at the intersection of QoS and QoE areas [4] or only in the area of QoE. However, it has become obvious that such a general approach simply does not work for many visual applications such as target recognition (utility) applications (video surveillance, telemedicine/remote diagnostics, fire safety, backup cameras, games, etc.) [5] [6].

In fact, QoE – the way quality of multimedia services is perceived – depends on a number of objective and subjective contextual parameters [7]. Only a full understandingof these parameters makes it possible to obtain results consistent with the expectations of service users, and, consequently, to optimize quality, but that is usually only possible when strong limitations are placed on the QoE modelling application. [8]. Unfortunately, the large number of contextual parameters means this research question is still open.

Target Recognition Video

In many visual applications, the quality of the motion picture is not as important as the ability of the user to perform specific tasks for which the visual system was created, given the processed video sequences. Such sequences are called Target Recognition Video (TRV). Regardless of the different ways in which the concept of TRV quality is understood, its verification is necessary to perform dedicated quality testing. The basic premise of these tests is to find TRV quality limits for which the task can be performed with the desired probability or accuracy.

Such tests are usually subjective tests (psychophysical experiments) with a group of subjects. Unfortunately, due to the complexity of the issue and our relatively low degree of understanding of human cognitive mechanisms, computer modelling of TRV quality has not yet achieved satisfactory results beyond very limited areas of application.

Given the use of TRV, qualitative tests do not focus on the subject's satisfaction with the quality of the video sequence, but instead they measure how the subject uses TRV to accomplish certain tasks. Purposes of this may include:

- Video surveillance – recognition of vehicle license plate numbers
- Telemedicine/remote diagnostics – correct diagnosis

- Fire safety – fire detection
- Backup cameras – parking the car
- Games – spotting and correctly reacting to a virtual enemy

The human factor is a significant influence; therefore it is necessary to ask questions on the procedures that must be followed to make a subjective assessment of TRV quality. In particular, questions arise on:

- Method of selecting the TRV source from which the test TRV (with degraded quality) arises
- Subjective testing methods and the general manner of conducting the psychophysical experiment
- Method of selecting a group of subjects in the psychophysical experiment, especially identification of any prior knowledge of the task
- Training subjects before the start of the experiment
- Conditions in which the test will be carried out
- Methods of statistical analysis and presentation of results

Methods for Subjective Evaluation of TRV

The questions formulated in the previous section are addressed by Recommendation ITU-T P.912 [9] "Subjective Video Quality Assessment Methods for Recognition Tasks", published in 2008. In addition, Recommendation P.912 organizes terminology related to subjective TRV testing, introducing appropriate definitions for the methods of testing (psychophysical experiments).

Unfortunately, Recommendation P.912 is only the first step in the standardization of methods of subjective TRV testing. In the opinion of the authors, based on research results (their own and independent) and observations conducted during numerous experiments with TRV, many claims of Recommendation P.912 are formulated at too high a level of

generality. What's more, selected statements are not supported by research results and are significantly disputable. In this situation, the authors propose amendments to Recommendation P.912. We would like to invite all researchers working on TRV-related topics to join us in the process of improving P.912.

References

[1] E. Cerquiera, S. Zeadally, M. Leszczuk, M. Curado, A. Mauthe, "Recent advances in multimedia networking", *Multimedia Tools and Applications*, vol. 54, pp. 635-647, 2011

[2] M. Grega, L. Janowski, M. Leszczuk, P. Romaniak, Z. Papir, "Estimation of the perceived quality of service (QoE) for multimedia communication", *Review of Telecommunications, Telecommunication News*, vol. 81, p 142-153, 2008

[3] M. Mu, P. Romaniak, A. Mauthe, M. Leszczuk, L. Janowski, E. Cerquiera, "Framework for the integrated video quality assessment", *Multimedia Tools and Applications*, vol. 61, pp. 787-817, 2012

[4] M. Leszczuk, L. Janowski, P. Romaniak, Z. Papir, "Assessing quality of experience for high-definition video streaming packet loss under diverse patterns", *Signal Processing-Image Communication*, vol. 28, pp. 903-916, 2013

[5] M. Leszczuk, "Image quality for utility applications: definitions, testing, standardization and current trends", *Review of Telecommunications, Communications News*, vol 83, pp. 242-246, 2010

[6] S. Moeller, A. Raake (ed.), *Quality of Experience: Advanced Concepts, Applications and Methods*, Springer, Berlin 2013

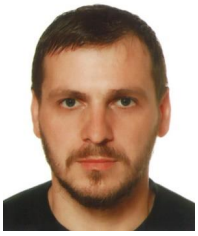
[7] P. Le Callet, S. Moeller, A. Perkis (ed.), *QUALINET White Paper on QoE Definitions*, European Network on QoE in Multimedia Systems and Services (COST Action IC 1003), Dagstuhl 2012

[8] M. Leszczuk, "Optimizing task-based video quality: a journey from psychophysical experiments Subjective

quality is objective optimization", *Multimedia Tools and Applications*, vol. 68, pp. 41-58, 2014

[9] ITU-T *P.912: Subjective video quality assessment methods for recognition tasks*, Geneva 2008

Mikołaj Leszczuk, PhD, is an assistant professor at the Department of Telecommunications, AGH University of Science and Technology (AGH-UST), Krakow, Poland. He has participated actively in several national and European projects. He is a member of the VQEG Board and a co-chair of VQEG QART (Quality Assessment for Recognition Tasks) and MOAVI (Monitoring of Audio Visual Quality by Key Indicators) Group. *The Polish National Centre for Research and Development funds his work, Contract No. C2013/1-5/MITSU/2/2014 under the international EUREKA programme: MITSU – Next Generation Multimedia Efficient, Scalable and Robust Delivery.*



Lucjan Janowski *is an assistant professor at the Department of Telecommunications, AGH University of Science and Technology, in Krakow, Poland. He is a Co-Chair of the VQEG JEG-Hybrid project (http://www.its.bldrdoc.gov/vqeg/projects/jeg/jeg.aspx).*