Monitoring of Audiovisual Quality by Key Indicators (MOAVI)

Emmanuel Wyckens
Orange Labs
Rennes, France
Email: emmanuel.wyckens@orange.com

Mikołaj Leszczuk
AGH University of Science and Technology
Kraków, Poland
Email: leszczuk@agh.edu.pl

Silvio Borer
SwissQual
Zuchwil, Switzerland
Email: Silvio.Borer@swissqual.com

***Abstract*—Automating quality checking is currently based on finding major video and audio artefacts. The processing is performed on the video and audio signal and/or the transmitted bitstream. The Monitoring Of Audiovisual quality by key Indicators (MOAVI) subgroup of the Video Quality Experts Group (VQEG) is an open collaborative project for developing no-reference models for monitoring audiovisual service quality. MOAVI is a complementary, industry-driven alternative to Quality of Experience (QoE) used to automatically measure audiovisual quality by using simple indicators of perceived degradation. The goal is to develop a set of key indicators (including blocking effects, blurring effects, freeze/jerkiness effects, ghosting effects, slice video stripe errors, aspect ratio problems, field order problems, photosensitive epilepsy flashing effects, silence, and clipping) describing service quality in general (the list is not closed, but the major artefacts are presented), and to select subsets for each potential application. Therefore, the MOAVI project concentrates on models based on key indicators contrary to models predicting overall quality. The main motivation of this approach is to overcome the limitations of current models predicting overall quality in real world applications. As the MOAVI project is at the beginning, it first focuses mainly on video quality. Measurements can be applicable according to the availability of access points along the video chain (video head-end, server of content delivery, terminal, encrypted or not).**

***Index Terms*—Monitoring, audio, video, QoE, KPI**

I. Introduction

Current Quality of Experience (QoE) models of the No-Reference (NR), like the reported in related research work [1], address measuring quality of networked multimedia, using objective parametric models. Unfortunately these models still do not perform well in predicting overall audiovisual QoE. Therefore a complementary, industry-driven alternative used to measure the quality automatically by using simple perceived indicators can to be proposed. Consequently, the Monitoring Of Audio-Visual quality by key Indicators (MOAVI) [15] subgroup of the Video Quality Experts Group (VQEG) [2], an open collaborative project for developing NR models for monitoring audiovisual service quality, develops such a set of key indicators.

This paper is organized as follows. Section II describes related limitations. In section III, the MOAVI Project is discussed. Section IV presents MOAVI's key indicators. Section V concludes the paper and details the future work.

II. Background and State-of-the-Art

This section presents limitations of state-of-the-art Full-Reference (FR), Reduced-Reference (RR) and NR metrics for standardised models, as well as automating quality checking.

A. Limitations of FR, RR and NR for standardised models

Most of the models in the recommendations were validated using one of the following hypotheses:

- Frame freezes up to 2 seconds,
- No degradation at the beginning or at the end of the video sequence are allowed,
- No skipped frames,
- The video reference should be clean (no spatial or temporal distortions),
- Minimum delay is supported between video reference and video (sometimes with constant delay),
- The up or downscaling operations are not always taken into account,
- Most models are based on measuring conventional blurriness, blockiness and jerkiness artefacts for producing predictive Mean Opinion Scores (MOS).
  The most of algorithms producing the MOSp scores is a mix between blur, block, jerkiness metrics. The weighting between each indicator could be a simple mathematical function. If one of these indicators is not correct, the global predictive score is completely wrong.
  The another indicators mentioned in MOAVI are not taken into account (ghosting, slice error) for MOSp.

Whereas the history regarding ITU-T Recommendations is shown in Table I, metrics based on video signal only are shown in Table II.

TABLE I
The history regarding ITU-T Recommendations.

Type of Model	*Format*	*Recommendation*	*Year*
FR	SD	J.144 [3]	2004
FR	QCIF/CIF/VGA	J.247 [4]	2008
RR	QCIF/CIF/VGA	J.246 [5]	2008
FR	SD	J.144 [3]	2004
RR	SD	J.249 [6]	2010
FR	HD	J.341 [7]	2011
RR	HD	J.342 [8]	2011
Bitstream	VGA–HD	In progress	Exp. 2013
Hybrid	VGA–HD	In progress	Exp. 2013

TABLE II
SYNTHESIS OF FR, RR AND NR MOS MODELS (BASED ON: [9]).

		Type of ITU-T Model		
		FR	*RR*	*NR*
Resolution	*HDTV*	J.341 [7]	n/a	n/a
	SDTV	J.144 [3]	n/a	n/a
	VGA	J.247 [4]	J.246 [5]	n/a
	CIF	J.247 [4]	J.246 [5]	n/a
	QCIF	J.247 [4]	J.246 [5]	n/a

The related research work [1] addresses measuring multimedia quality in mobile networks with an objective parametric model. Closely related is a current standardization activity at ITU-T SG12 on models for multimedia and IPTV based on bit-stream information. SG12 is now working on models for IPTV. Q.14/12 is responsible for these projects, which are provisionally called P.NAMS (non-intrusive parametric model for assessment of performance of multimedia streaming) and P.NBAMS (non-intrusive bit-stream model for assessment of performance of multimedia streaming). P.NAMS utilizes only packet-header information (e.g., from IP through MPEG2-TS), while P.NBAMS is allowed to use the payload information (i.e., coded bit-stream) [10]. However, this work has been focused on the overall quality (in MOS units), while MOAVI is focusing on Key Performance Indicators (KPI).

The MOAVI project could explore human behaviour on longer period and propose adapted model with enhanced SSCQE methods.

Most of the recommended models are based on a global quality evaluation of the video sequences as in P.NAMS and P.NBAMS projects. The predictive score is correlated to subjective score obtain with global evaluation methodologies (SAMVIQ, DSCQS, ACR, etc.). Generally, the duration of video sequences is limited to 10 s or 15 s in order to avoid a forgiveness effect (the observer is not enabling to score properly the video after 30 s and can give more weight to artefacts occurring at the end of the sequence). When one model is deploying for monitoring of video services, the global scores are provided for fixed temporal windows and without any acknowledgement of the previous scores.

B. Automated Quality Checking

Automating quality checking is currently based on finding major video and audio artefacts. The processing is performed on the video signal and/or the bit-stream. Quality checking can be conducted before, during, and/or after the encoding process. However, no MOS is provided, detailing (among others):

- Blocking effects,
- Freeze, jerkiness effects,
- Ghosting effects,
- Slice error or video stripes,
- Aspect ratio conformity,
- Field order conformity,
- Photosensitive flashing epilepsy (ITU-R BT.1702 [11]).

III. INTRODUCTION TO MOAVI

The MOAVI subgroup of the VQEG is an open collaborative project for developing NR models for monitoring audiovisual service quality. The goal is to develop a set of key indicators (including blocking effects, blurring effects, freeze/jerkiness effects, ghosting effects, slice video stripe errors, aspect ratio problems, field order problems, photosensitive epilepsy flashing effects, silence, and clipping) describing service quality in general, and to select subsets for each potential application. Therefore, the MOAVI project concentrates on models based on key indicators contrary to models predicting overall quality.

MOAVI is a complementary, industry-driven alternative to QoE used to measure the audiovisual quality automatically by using simple perceived indicators. The Table III tries to summarize some difference between the complementary projects. The perceived indicators should have a robust prediction performance with a minimum operational restriction. Targeted services include video on demand (VoD) and live broadcast services (satellite, IPTV, digital terrestrial television). The video indicators can be based on analysing the video signal only, or on using parametric (bit-stream) or hybrid measurements (bit-stream + video signal).

TABLE III
SUMMARY OF SOME DIFFERENCE BETWEEN THE COMPLEMENTARY PROJECTS.

Item	*MOAVI*	*P.NAMS*
Approach	No signal reference, parametric, hybrid	Parametric only
Result	Artefact detection NOT MOSp quality	MOSp
Use cases	Not necessary to access to bit-stream, video Head-end monitoring, VoD quality checking on the source and/or coded sequences of content providers, end-user monitoring	Need to access to bit stream and buffering access and size buffer, limited to QCIF-QVGA resolution for mobile application, end-user monitoring
Performance	Should be better than MOSp without any video reference	Under investigation in P.NAMS project
Signal access	Can support Conditional Access, can support and FEC & ARQ	Need to know FEC & ARQ models

Measurements can be applicable according to the availability of access points along the video chain (video head-end, server of content delivery, terminal, encrypted or not).

The MOAVI activities are split into 4 steps:

1) Maintaining a list of potential real-world applications for visual quality monitoring.
   As a result, some additional artefact definitions could be submitted to ITU-T G.100 [12].
2) Identifying the main video indicators taken into account in customer acceptability.
   Some artefacts definitions are existing on ITU-T P.930 [13], but the definitions should be updated for ITU-T H.264 [14] (AVC) and future ITU-T H.265 (HEVC) encoding technologies. The contributors are invited to

suggest the most representative perceived indicators. During this step, participants can also propose some appropriate subjective tests for each indicator relating to user acceptability of ITU-T G.100 [12].

3) Designing the indicators according to 3 categories.
According to the result obtained during the previous step, participants help to design each indicator for one or more of the categories:
   a. Based on the visual signal,
   b. Based on the parametric signal,
   c. Based on the hybrid signal.

   If possible, the models will be designed by using video sequences collected under operational conditions.
4) Performance evaluation of the indicators.
The performance of each indicator will aim to maximize the true prediction (true positives/negatives) and minimize the false prediction (false positives/negatives). Statistical instruments may include precision, recall, specificity, sensitivity, accuracy, F1-score, etc. Furthermore, if DCR (Degradation Category Rating)-like scores are collected, regular MOS-like statistical analysis will be applicable as well.

IV. MOAVI's Key Indicators

As an alternative to MOS, the following list of key indicators has been tentatively planned to be considered under MOAVI:

- Compression artefacts (blocking and blurring effects),
- Compression artefacts (flickering and ringing effects),
- Blurring effects,
- Freeze (jerkiness) effects,
- Ghosting effects,
- Slice error or video stripes,
- Aspect ratio conformity,
- Field order conformity,
- Photosensitive epilepsy or flashing effects (ITU-R BT.1702 [11]),
- Tearing effects.

In the next step, literature review will be done in order to find out the state-of-the-art research results on defining quantitative threshold of perceptibility of individual distortions. The list is not closed, but the major artefacts are presented.

V. Conclusions and Next Steps

This project is still in its infancy. There are no details on the proposed architecture available yet.

The MOAVI Project URL is [15]. Questions should be sent to the MOAVI Co-Chairs. Currently, 9 people have been involved into this activity.

In the next step (until the next VQEG meeting, winter 2012/2013), literature review will be done in order to find out the state-of-the-art research results on defining quantitative threshold of perceptibility of individual distortions. Methods for measuring distortions will be analysed as well. Then, psychophysical experiments will be conducted for distortions, for which, quantitative thresholds are missing. As the result, the thresholds are going to be contributed to the research community (by means of a published scientific paper).

VI. Acknowledgements

The Mikołaj Leszczuk's research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement №218086 (INDECT) and from AGH University of Science and Technology under Department of Telecommunications statutory work.

References

[1] J. Gustafsson, G. Heikkila, and M. Pettersson, "Measuring multimedia quality in mobile networks with an objective parametric model," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, oct. 2008, pp. 405 –408.

[2] *The Video Quality Experts Group*, VQEG, July 2012, http://www.vqeg.org/.

[3] *ITU-T J.144, Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference*, International Telecommunication Union Recommendation, 2004. [Online]. Available: http://www.itu.int/rec/T-REC-J.144-200403-I

[4] *ITU-T J.247, Objective perceptual multimedia video quality measurement in the presence of a full reference*, International Telecommunication Union Recommendation, 2008. [Online]. Available: http://www.itu.int/rec/T-REC-J.247-200808-I

[5] *ITU-T J.246, Perceptual isual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference*, International Telecommunication Union Recommendation, 2008. [Online]. Available: http://www.itu.int/rec/T-REC-J.246-200808-I

[6] *ITU-T J.249, Perceptual video quality measurement techniques for digital cable television in the presence of a reduced reference*, International Telecommunication Union Recommendation, 2010. [Online]. Available: http://www.itu.int/rec/T-REC-J.249-201001-I

[7] *ITU-T J.341, Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference*, International Telecommunication Union Recommendation, 2011. [Online]. Available: http://www.itu.int/rec/T-REC-J.341-201101-I

[8] *ITU-T J.342, Objective multimedia video quality measurement of HDTV for digital cable television in the presence of a reduced reference signal*, International Telecommunication Union Recommendation, 2011. [Online]. Available: http://www.itu.int/rec/T-REC-J.342-201104-I

[9] E. Wyckens, "Proposal studies on new video metrics," in *VQEG Hillsboro Meeting*, A. Webster, Ed., Orange Labs. Hillsboro: Video Quality Experts Group (VQEG), December 2011, p. 17.

[10] A. Takahashi, K. Yamagishi, and G. Kawaguti, "Global standardization activities recent activities of qos / qoe standardization in itu-t sg12," *Ntt Technical Review*, vol. 6, no. 9, pp. 1—5, 2008.

[11] *ITU-R BT.1702, Guidance for the reduction of photosensitive epileptic seizures caused by television*, International Telecommunication Union Recommendation, 2005. [Online]. Available: http://www.itu.int/rec/R-REC-BT.1702

[12] *ITU-T G.100/P.10, Vocabulary for performance and quality of service*, International Telecommunication Union Recommendation, 2006. [Online]. Available: http://www.itu.int/rec/T-REC-P.10-200607-I

[13] *ITU-T P.830, Subjective performance assessment of telephone-band and wideband digital codecs*, International Telecommunication Union Recommendation, 1996. [Online]. Available: http://www.itu.int/rec/T-REC-P.830-199602-I/en

[14] *ITU-T H.264, Advanced video coding for generic audiovisual services*, International Telecommunication Union Recommendation, 2012. [Online]. Available: http://www.itu.int/rec/T-REC-H.264-201201-P/en

[15] E. Wyckens, S. Borer, and M. Leszczuk, *MOAVI (Monitoring of Audio Visual Quality by Key Indicators)*, VQEG, July 2012, http://www.its.bldrdoc.gov/vqeg/projects/moavi/moavi.aspx.