Quality impact of Audio on Audio-visual Quality in the context of IPTV.

Marie-Neige Garcia[1], Alexander Raake[1]
[1] Quality and Usability Lab, Deutsche Telekom Laboratories, Berlin University of Technology

Abstract

This paper investigates the quality impact of audio in an audiovisual context. It tries to answer a number of questions based on the analysis of audio-only, video-only and audio-visual subjective test results: What is the quality impact of audio on the overall audio-visual quality in the context of IP-based TeleVision (IPTV)? Does this impact depend on the audio-visual content type, e.g. soccer with speech on noise? Does it depend on the degradation type: compression artifacts or transmission errors? Does it depend on the video resolution, i.e. Standard Definition (SD, 720 x 576) or High Definition (HD, 1920x1080)? The paper discusses these questions based on the results of perception tests. This paper also shows how the results of this analysis are used for modeling the audio-visual quality.

Introduction

In order to achieve a high degree of user satisfaction for IPTV services, we developed an audio-visual quality model. The model estimates the audio-visual quality of the service as perceived by the user and has been developed based on the results of intensive perception tests. One interesting aspect related to this topic is the quality impact of audio on the overall Audio-Visual (AV) quality, and does this impact depend on the content type. In other words, should the model be the same for all AV contents (e.g. TV news with speech vs. video clip with music) or should the prediction of the model depend on the audio-visual content type, reflecting the relative importance of the media, audio vs. video, depending on the content? Indeed, we can expect that degradations occurring on audio might have a higher impact on the overall quality in TV news than in sport-programs. In addition, the influence of the degradation itself might play a role in the quality impact of audio on the overall AV quality: for a given audio quality, the interaction with a given video quality might yield different AV quality, depending on the audio degradation and/or video degradation type. The influence of the content on the AV quality has been explicitly taken into account in both Ries [1] and Hands [2] papers, in which they propose either different regression coefficients or model types of the AV quality models, depending on the video content type. However, those models do not consider the influence of the degradation type. Moreover, they do not cover typical IPTV video format and degradation types, such as the effect of transmission errors. In this paper, we will first describe in the "Experimental Design" Section the perception tests we conducted for developing the AV quality model. Then, in the "Data Analysis" Section, we will analyze the quality impact of audio on the overall quality, and the influence of the content type, degradation type and video format - on this impact. We will also study how these impact and influences translate in terms of modeling.

Experimental Design

Audio, video and AV subjective tests have been conducted using audio-only, video-only and AV sequences, respectively. The source material comprises five AV contents of 16 s duration each. Video-only and AV test were conducted for two video resolution separately: High Definition (HD, 1920x1080 pixels) and Standard Definition (SD, 720x576 pixels). They are representative of different TV-programs: A: Movie trailer/speech on music, B: Interview/speech, C: Soccer/speech on noise, D: Movie/classical music and E: Music video/pop music with singer. All tests follow standard procedures (ITU-T P.800, ITU-R BT-500, ITU-T P.910). An “absolute category rating”was used for collecting subjective quality judgements. The subjects rated the quality using the continuous 11-point quality scale recommended in ITU-T P.910. 24 subjects participated in each test, and each subject was allowed to participate in only one test. The stimuli are obtained by off-line processing of the five uncompressed AV contents at various codecs, bit-rates and packet-loss-rates. More information on the perception tests can be found in [3].

Data Analysis

Results were collected from the three subjective tests presented in the previous Section. For each subjective test, the averaged Mean Opinion Score (MOS) was computed over all subjects and transformed to the 100-point-model-scale using the conversion defined in ITU-T G.107. We will first focus on the quality impact of audio on AV quality using ratings averaged over all contents and degradations, thus not considering the influence of the content type and degradation type. Then we will analyze the influence of the content type, of the degradation type and how those influences are reflected in the modeling.

In order to have a first impression of the quality impact of audio on the overall AV quality, we compute the correlation of the overall AV quality with the audio quality, the video quality, the interaction between audio quality and video quality, using the ratings averaged over all contents (see Table 1, column "all"). It can be observed that for both SD and HD, the interaction between audio and video qualities is predominant for the AV quality (SD: correlation = 0.97; HD: correlation = 0.96). The video quality (SD: correlation = 0.78, HD: correlation = 0.81) seems to have more impact on the overall AV quality than the audio quality (SD: correlation = 0.54; HD: correlation = 0.49), especially for HD. This latter statement is expected and shows that the impact of video quality on AV quality increases with the video format.

	Qav	**all**	**A**	**B**	**C**	**D**	**E**
	Qa	0.49	0.49	0.46	0.45	0.45	*0.57*
HD	Qv	0.81	0.87	0.80	0.86	0.84	*0.69*
	Qa·Qv	0.96	0.94	0.94	0.94	0.95	*0.92*
	Qa	*0.54*	*0.57*	*0.52*	*0.48*	*0.48*	*0.61*
SD	Qv	0.78	0.73	0.77	0.81	0.79	*0.67*
	Qa·Qv	0.97	0.94	0.95	0.92	0.96	*0.96*

Tabelle 1: Correlation of the audio and video quality, and their interaction with the overall quality (SD and HD)

Averaging the results over all contents might hide that for some contents, the above statements are not valid anymore. As a consequence, we compute the same correlations as above, but using ratings per content (see Table 1, columns "A" to "E"). For the contents "A" to "D", the same observation as for "all" contents can be made. For content "E", the quality impact of audio seems to be higher than the other content, and close to the quality impact of video (SD: correlation (Qav,Qa)= 0.61; HD: correlation (Qav,Qa) = 0.57), especially for SD, confirming the impact of the video format.

One more aspect to be considered is the influence of the degradation type on the quality impact of audio on the overall AV quality. In our case, degradation types refer to audio and video compression, and audio and video packet loss. We want to know if, for instance, for given audio and video qualities, we obtain different AV quality values if we have audio compression instead of audio packet loss. For addressing this issue, we will refer to a previous work described in [3] and in which we compare two AV quality models: the Quality-based (Q-based) model, shown in equation (1), which estimates the AV quality Qav based on the audio quality (Qa) and the video quality (Qv), and the Quality-Factor-based (QF-based) model, presented in equation (2), and which is based on audio and video quality factors ($Icod_X$,$Itra_X$, $X \equiv A$: audio, $X \equiv V$: video). $Icod_X$ refers to the quality impact of compression while $Itra_X$ refers to the quality impact of the packet loss.

$$Qav = \alpha + \beta \cdot Qa + \gamma \cdot Qv + \mu \cdot Qa \cdot Qv \quad (1)$$

$$\begin{aligned} Qav = & \; Qavo \\ & -a \cdot Icod_A - b \cdot Icod_V - c \cdot Icod_A \cdot Icod_V \\ & -d \cdot Itra_A - e \cdot Itra_V - f \cdot Itra_A \cdot Itra_V \\ & -g \cdot Icod_A \cdot Itra_V - h \cdot Icod_V \cdot Itra_A \quad (2) \end{aligned}$$

Contrary to the Q-based model, the QF-based model assumes that the quality impact of audio depends on the degradation type. For instance, if in the equation (2), the regression coefficients b and e corresponding to $Icod_A$ and $Itra_A$ are significantly different, which is the case (see [3]), we can conclude that the quality impact of audio on the overall AV quality depends on the degradation type.

For taking into account the influence of the content type in the modeling, we apply both the Q-based and QF-based model on the ratings averaged over all contents – content-blind models – and on the ratings per content – content-based models. We compare the obtained regression coefficients taking into account their confidence intervals: if the confidence intervals of two regression coefficients overlap, the regression coefficients are not considered to be different. If the confidence interval overlaps the '0' value, the regression coefficient is not considered to be different from 0. Results of the regressions are shown in Table 2. As expected from the observations we previously made on the correlations, when we do not consider the influence of the degradation type – Q-based model – the audio quality for SD has an impact on the AV quality only when interacting with the video quality ($\mu \neq 0$ while $\beta = 0$). For HD, the video quality is predominant over the audio quality, yielding a significant regression coefficient γ, except for content E, for which the audio and video quality are more balanced. When taking into account the degradation type – QF-based model – the audio quality alone does have an impact on the overall AV quality, and there is also an influence of the content, yielding different coefficients values for the terms linked to audio and video compression ($Icod_A$ and $Icod_V$).

		HD	**SD**
Q	blind	$\beta = 0$	$\beta = \gamma = 0$
Q	based	A to D: $\beta = 0$	All: $\beta = \gamma = 0$
		E: $\beta = \gamma = 0$	
QF	blind	cf. $\neq 0$	cf. $\neq 0$
	based	A to D: cf $\neq 0$	All: $c = 0$
QF		E: $b = 0$	a: C≠E;b: C$\neq B, D, E$

Tabelle 2: Influence of the degradation type and of the content type on modeling; SD and HD;Q- and QF-based model;blind: Content-blind; based:Content-based; cf: all Regression Coefficients. Contents A,B,C,D,E

Conclusion

Based on the analysis of a series of perception tests, we came to the conclusion that, in the IPTV context, when not considering the influence of the content type and degradation type, the audio has a quality impact on the overall AV quality mainly when interacting with the video quality ($Qa.Qv$). We have shown that the content type, the degradation type and the video format have an influence on the quality impact of audio on the overall AV quality. The selected contents and listening environment were representative of the IPTV domain. We plan to extend this study to additional contents in which the encoding complexity of the video sequences is lower than in the five already studied contents, and to conduct tests with headphones, instead of loudspeakers. At last, we plan to validate the various obtained models on unknown database.

Literatur

[1] Ries, M et al.: Audiovisual Quality Estimation for Mobile Streaming Services, IEEE ISWCS (2005)

[2] Hands, D.S.: A Basic Multimedia Quality Model, IEEE Trans. on Multimedia (2004)

[3] Garcia,M-N. and Raake. A.: Impairment-factor based audio-visual quality model for IPTV, QOMEX (2009)