Enhancements to the Opinion Model for Video-Telephony Applications

Jose Joskowicz
Facultad de Ingeniería
Universidad de la República
Montevideo, Uruguay
Phone: +598 2 7110974
josej@fing.edu.uy

J. Carlos López Ardao
ETSE Telecomunicación
Campus Universitario, 36310
Vigo, Spain
Phone: +34 986 8212176
jardao@det.uvigo.es

ABSTRACT
In this paper, we show how the proposed model in ITU-T Recommendation G.1070 "Opinion model for video-telephony applications" cannot model properly the perceptual video quality, especially in the low bit rate range, due to the great variation of MOS values depending on video content. In this work, we present different enhancements to the model, allowing a much better approximation to the perceptual MOS values, knowing only the subjective movement content in the video application, classified in "Low", "Medium" or "High". Studies were made for more than 1500 processed video clips, coded in MPEG-2 and H.264/AVC, in bit rate ranges from 50 kb/s to 12 Mb/s, in SD, VGA, CIF and QCIF display formats. Video clips subjective quality was estimated using one of the quality metrics standardized in ITU-T Recommendation J.144 and ITU-R Recommendation BT.1683.

Categories and Subject Descriptors
H.4.3 [**Information Systems Applications**]: Communications Applications - *Computer conferencing, teleconferencing, and videoconferencing*; C.4 [**Performance of Systems**]: Design Studies, Modeling Techniques

General Terms
Algorithms, Performance, Design, Standardization

Keywords
Video perceptual quality, Video codecs, Video signal processing, VoIP Network design

1. INTRODUCTION
This paper presents enhancements to the ITU-T Recommendation G.1070 "Opinion model for video-telephony applications" [1], related to the video quality estimation. The original model and the proposed enhancements were compared for videos coded in MPEG-2 [2] and H.264 [3] at different bit rates, and in different display formats, including SD (Standard Definition, 720 × 576 pixels), VGA (Video Graphics Array, 640 × 480 pixels), CIF (Common Intermediate Format, 352 × 288 pixels) and QCIF (Quarter Common Intermediate Format, 176 × 144 pixels). Sixteen video sources were used, coded in 96 different formats, varying the codec (MPEG-2 and H.264), the bit rate (from 50 kb/s to 12 Mb/s) and the display format. In total more than 1500 processed video sequences were analyzed to derive the proposed enhancements.

MPEG-2 is widely used in commercial applications for digital TV distribution. It is also used to encode movies and other programs that are distributed on DVD. For these reasons, most of the digital video receivers support it. H.264/AVC is the natural successor to MPEG-2. There is now a very high interest in this new codec, providing better quality at lower bit rates [4].

Based on VQEG (Video Quality Expert Group) work, ITU (International Telecommunication Union) has standardized the ITU-T Recommendation J.144 [5] and ITU-R Recommendation BT.1683 [6] for estimation of the perceived video quality in digital TV applications when the original signal reference is available (Full Reference models). Also, the standardization for the estimation of the perceived video quality in multimedia applications is in process, based on the VQEG Multimedia Reports [7]. Instead of doing subjective tests, we used one of the models standardized by ITU, and developed by the NTIA (National Telecommunications and Information Administration) [8]. With certain error margins, these models predict very well the subjective quality.

We propose four enhancements to the G.1070 model. First, one of the model parameters is suppressed, without loosing performance. Second, we show how the two remaining parameters are highly correlated to video movement content. We propose to define 3 sets of these two parameters, one for low movement content applications, one for medium movement content and one for high movement content applications. Third, one new parameter is added, to take into account the display format. Finally, a generalization to the model is proposed to easily extend it to different codecs, based on the parameters value for MPEG-2. The MSE (mean square error) model with the proposed enhancements are calculated and compared to the original G.1070 model.

The paper is organized as follows: Section 2 briefly describes the video quality estimation model proposed in G.1070. Section 3 describes how perceived quality varies with respect to the bit rate. In Section 4 the model enhancements are presented, and contrasted to the original model. Section 5 summarizes the main contributions.

2. VIDEO QUALITY ESTIMATION IN ITU-T RECOMMENDATION G.1070

The ITU-T Recommendation G.1070 [1] describes a computational model for point-to-point interactive videophone applications over IP networks that is useful as a QoE (Quality of Experience) and QoS (Quality of Service) planning tool for assessing the combined effects of variations in several video and speech parameters that affect the perceived quality. The model takes into account the speech and the video perceived quality [9], and combines both in an integration function for overall multimedia quality [10]. Speech quality estimation is mainly based on the ITU-T Recommendation G.107 [11], known as the E-Model. Video quality estimation V_q is calculated as shown in Equation (1).

$$V_q = 1 + I_c e^{-\frac{P_{plv}}{D_{Pplv}}} \quad (1)$$

where I_c represents the basic video quality, determined by the codec distortion and is a function of the bit rate and frame rate, P_{plv} is the packet loss rate and D_{Pplv} expresses the degree of video quality robustness due to packet loss, and can also depend on bit rate and frame rate.

The maximum basic video quality I_c for each bit rate b (the video quality without packet loss) is expressed in Equation (2). Combining Equation (1) and (2), the video quality V_q without packet loss is expressed in Equation (3).

$$I_c = v_3 \left(1 - \frac{1}{1 + \left(\frac{b}{v_4} \right)^{v_5}} \right) \quad (2)$$

$$V_q = 1 + v_3 \left(1 - \frac{1}{1 + \left(\frac{b}{v_4} \right)^{v_5}} \right) \quad (3)$$

According to the ITU-T Recommendation G.1070, coefficients v_3, v_4 and v_5 are dependent on codec type, video display format, key frame interval, and video display size, and must be calculated using subjective video quality tests. Provisional values are provided only for MPEG-4 in QVGA (Quarter VGA, 320 × 240 pixels) and QQVGA (Quarter QVGA, 160 × 120 pixels) video formats.

In [12], the same model presented in equation (3) is proposed for IPTV services, and coefficient values are provided for the H.264 codec.

3. PERCEIVED QUALITY AS A FUNCTION OF THE BIT RATE

The most reliable form for measuring the perceived video quality of a video clip is through subjective tests, where typically a reference and a degraded video sequence are presented to different viewers, and opinions are averaged. The MOS (Main Opinion Score) or the DMOS (Difference Mean Opinion Scores) are the metrics typically used in these tests. Different kinds of subjective tests can be performed, based on ITU-R Recommendation BT.500-11 [13] and ITU-T Recommendation P.910 [14].

Many efforts have been made, and are currently ongoing, in order to develop objective video quality models. Based on the work performed by the VQEG, ITU has standardized in the ITU-T Recommendation J.144 [5] and ITU-R Recommendation BT.1683 [6] different full reference objective video quality models, which has been proved to be statistically equivalent between them. Among the standardized algorithms are the proposed by the NTIA [8] from U.S.A., the Yonsei University from Korea [15], the Telecommunications Research and Development Center (CPqD) from Brazil [16] and the British Telecom (BFTR) from England [17]. All these algorithms are statistically equivalent between them, but none is statistically equivalent to the "perfect model" (the one who is statistically equivalent to the subjective test results).

For each video clips pair (original and degraded), the algorithms provides a Video Quality Metric (VQM), with values between 0 and 1 (0 when there are no perceived differences and 1 for maximum degradation). Multiplying this value by 100 a metric is obtained which corresponds to the DSCQS (Double Stimulus Continuous Quality Scale) [13] and can be directly associated with the DMOS.

The error obtained using the standardized models with respect to subjective tests can be estimated in +/- 0.1 in the 0-1 scale. This means that the order of magnitude of the standardized algorithm error is 0.1 in a DMOS scale from 0 to 1.

Instead of doing subjective test, we used in this work the model proposed by NTIA standardized in ITU-T J.144, available in [18]. The DMOS values returned from the NTIA model can be related to the MOS using Equation (4). The interpretation of the MOS values is presented in Table 1. MOS varies between 1 (Bad quality) and 5 (Excellent Quality). MOS errors, using this model, can be estimated in +/- 0.4 in the 1-5 scale (4 times the DMOS error).

$$MOS = 5 - 4DMOS \quad (4)$$

The video clips detailed in Table 2, available in the VQEG web page [19], were used. The original and the coded video clips were converted to non-compressed AVI format in order to be compared with the NTIA model.

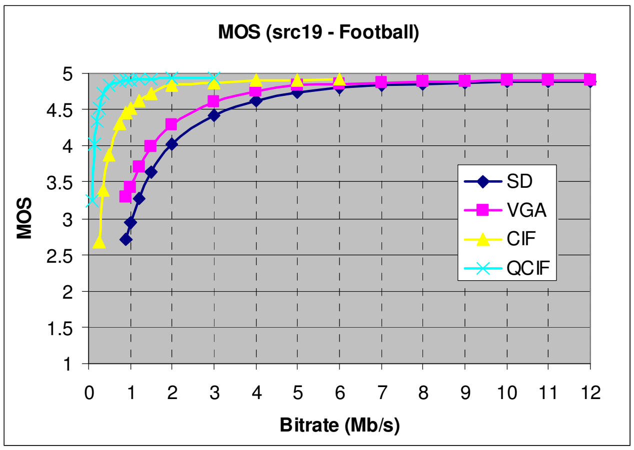
Figure 1 shows how the perceived quality varies (measured as MOS and DMOS) as a function of the bit rate, keeping constant all other coding parameters, for the clip "Football" (src 19), coded in MPEG-2 in different display formats. The figure shows the typical behavior for any video clip:

a) The perceived quality is higher (the DMOS is lower, MOS is higher) for higher bit rates.

b) For the same quality, higher bit rates are needed for larger displays.

Table 1. MOS to perceived quality relation

MOS Value	Quality
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad



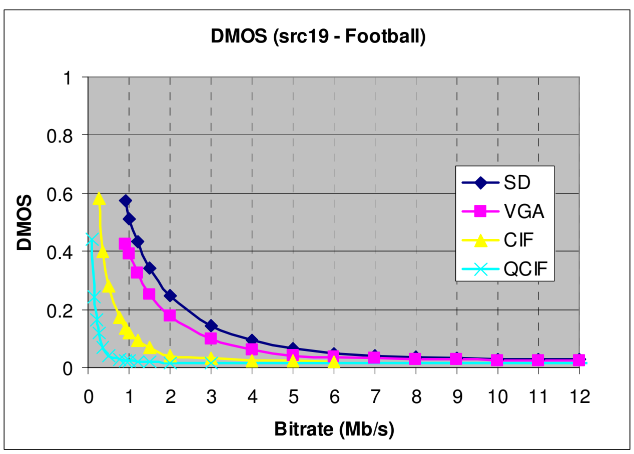


Figure 1. Perceived quality, as MOS and DMOS, using one of the ITU-T J.144 models for the clip "Football" coded in MPEG-2 as a function of the bit rate, for display formats SD, VGA, CIF and QCIF.

Table 2. Source video clips used

Source	Name	Source	Name
src 2	Barcelona	src 14	New York 2
src 3	Harp	src 16	Betes pas betes
src 4	Moving graphic	src 17	Le point
src 5	Canoa Valsesia	src 18	Autums leaves
src 7	Fries	src 19	Football
src 9	Rugby	src 20	Sailboat
src 10	Mobile & Calendar	src 21	Susie
src 13	Baloon-pops	src 22	Tempete

Figure 2 shows the relation between MOS and bit rate, for all the clips detailed in Table 2, coded in MPEG-2 (using the coding parameters detailed in Table 3), in SD display format. MOS values were derived from DMOS, using Equation (4). DMOS values were calculated using the NTIA Model.

Table 3. MPEG-2 and H.264 coding parameters

MPEG-2	H.264
Profile/Level: MP@ML Max GOP size: 15 GOP Structure: Automatic Picture Structure: Alw Frame Intra DC Precision: 9 Bit rate type: CBR Interlacing: Non-Interlaced Frame Rate: 25 fps	Profile/Level: High/3.2 Max GOP size: 33 2 B Pictures between I&P Entropy Coding: CABAC Subpixel mode: ¼ Pixel Bit rate type: CBR Interlacing: Non-Interlaced Frame Rate: 25 fps

As can be seen, all the clips have better perceived quality for higher bit rates, as can be expected. In MPEG-2, in SD, for bit rates higher than 6 Mb/s all the clips have an almost "perfect" perceived quality (DMOS less than 0.1, MOS higher than 4.5). At 3 Mb/s all the clips are in the range between "Good" and "Excellent". However for less than 3 Mb/s the perceived quality strongly depends upon the clip content. For example at 2 Mb/s, MOS varies between 3.6 and 4.8, and at 0.9 Mb/s MOS varies between 1.9 (between "Bad" and "Poor") and 4.2 (between "Good" and "Excellent"). Is common to use MPEG-2 at 3.8 Mb/s in SD IPTV commercial applications, where the perceptual quality is near "Excellent" for all video clips. However, at low bit rates there are high differences in the perceived quality for identical coding conditions, depending on video content. Similar considerations can be made for other display formats (i.e. VGA, CIF and QCIF).

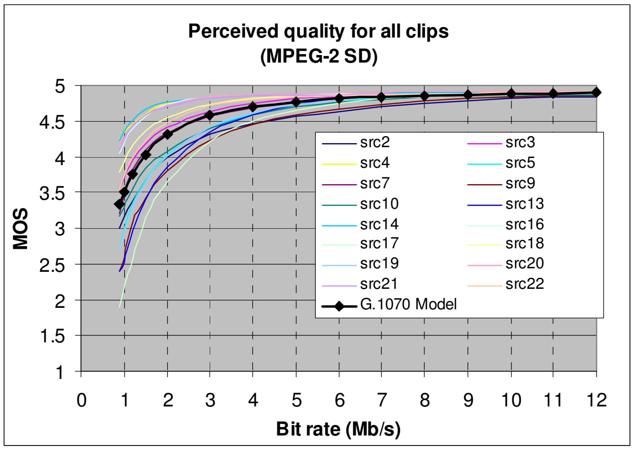


Figure 2. Perceived quality, as MOS, using one of the ITU-T J.144 models for the all the clips, coded in MPEG-2 as a function of the bit rate, for SD display format.

4. MODEL ENHANCEMENTS

ITU-T Recommendation G.1070 does not take into account the video content, because it is a parametric packet layer model. The best that the model can estimate is the *average* video quality for different contents, at each bit rate. But, looking at Figure 2, it can be seen that video quality strongly depends on video content, especially for low bit rates. Best values for v_3, v_4 and v_5 were calculated for all the video clips and are presented in Table 4. The G.1070 Model curve is also shown in Figure 2. It can be seen how, at low bit rates, the MOS differences between the model and the actual values can be very high for some clips (more than 1.4 in the 1-5 MOS scale).

Table 4. Best values for v_3, v_4 and v_5 for all clips in MPEG-2

v_3	v_4	v_5	MSE
3.94	0.708	1.6	0.35

4.1 Number of Parameters

Equation (3) has 3 parameters (v_3, v_4 and v_5). The maximum MOS quality, according the Equation (3) is $V_q = 1 + v_3$. Taking into account that the maximum value for MOS, by definition, is 5, v_3 can be defined as $v_3 = 4$. The calculated values for v_3 in different scenarios are very close to 4. In Table 4, the best value was 3.94. In the ITU-T Recommendation G.1070, provisional values for v_3 are 3.8 and 3.5, for different display formats. In [12], the best estimation for v_3 is 3.8.

We propose to set $v_3 = 4$, having this value a clear interpretation: it represents the value that sets the maximum MOS value equal to 5. Making this setting, only two parameters are leaving in Equation (3).

4.2 Movement Content

Assuming $v_3 = 4$, the best values form v_4 and v_5 can be obtained for each clip. Table 5 shows the values of v_4 and v_5 that best fits Equation (3) to each curve in Figure 2, as well as the MSE (Mean Square Error), sorted by v_4.

Table 5. v_4 , v_5 values that best fits to the actual NTIA curves

Src	Name	Mov	v_4	v_5	MSE
14	New York 2	Low	0.252	1.200	0.0441
4	Moving graphic	Low	0.252	1.200	0.0288
21	Susie	Low	0.290	1.200	0.0476
20	Sailboat	Low	0.290	1.240	0.0334
16	Betes pas betes	Low	0.328	1.280	0.0537
18	Autums leaves	Low	0.442	1.280	0.0395
3	Harp	Medium	0.594	1.400	0.0336
22	Tempete	Medium	0.594	1.480	0.0365
7	Fries	Medium	0.708	1.440	0.0603
10	Mobile & Calendar	Medium	0.784	1.320	0.0172
2	Barcelona	High	0.860	1.240	0.0228
5	Canoa Valsesia	High	1.012	1.600	0.0465
19	Football	High	1.050	1.680	0.0302
9	Rugby	High	1.240	1.600	0.0647
13	Baloon-pops	High	1.240	1.840	0.0356
17	Le point	High	1.506	2.040	0.0686

In Figure 2, very similar behaviors can be seen for many clips. For example, clips for src 4, 14, 16, 18, 20 and 21 have practically identical behaviors, and all of them have low movement content scenes. Table 5 shows a subjective estimation for the clip movement content, classified into the 3 groups: "Low", "Medium" and "High". It can be seen a correlation between v_4 and the movement content. Instead of trying to model all clips using only one curve, we propose to use 3 different sets of values for v_4 and v_5 depending on the movement content, classified into "Low", "Medium" and "High".

Table 6 shows the values for v_4 and v_5 that best fits to all the clips in each group. The maximum MSE is 0.17 (for the group "High Movement"). These MSE values are between 2 and 4 times better than the MSE obtained trying to model all the clips with only one curve (Table 4).

The main objective of the ITU-T Recommendation G.1070 is for QoE/QoS planning, and it is a packet layer model. Movement content cannot be derived from the packet layer, where only aspects related to the network (such as packet loss) can be calculated. Using only one set of parameters values will predict, in the low bit rate range, MOS values that will not be accurate, due to the high MOS variation with respect to video content. But, using the 3 different set of values proposed, the Model can predict different and more accurate MOS values for different applications, knowing or estimating the average movement content for the application.

Table 6. v_4 , v_5 values for each group for MPEG-2, SD

Movement	v_3	v_4	v_5	**MSE**
Low Movement	4	0.328	1.28	0.080
Medium Movement	4	0.670	1.40	0.11
High Movement	4	1.164	1.64	0.17

4.3 Display Format

Figure 3 shows the relation between MOS and bit rate, for all the clips detailed in Table 2, coded in MPEG-2 in different display formats (SD, VGA, CIF and QCIF). As can be seen, same clips have similar curve shapes in all the display formats, if the bit rate scale is compressed or expanded. A bit rate scale factor *a* can be added to the model, as detailed in Equation (5).

$$V_q = 1 + v_3 \left(1 - \frac{1}{1 + \left(\frac{ab}{v_4} \right)^{v_5}} \right) \quad (5)$$

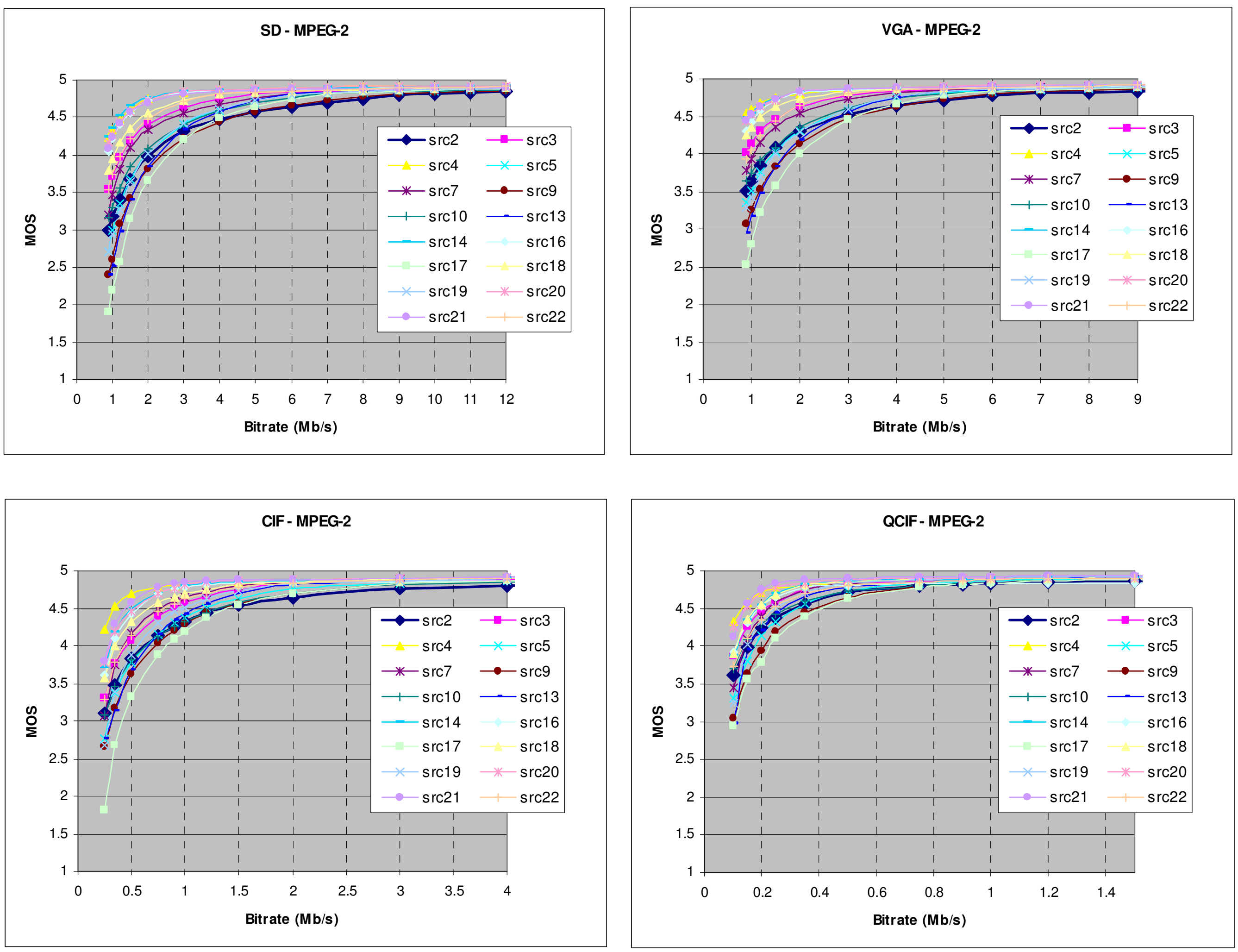


Figure 3. Perceived quality, as MOS, using one of the ITU-T J.144 models for the all the clips, coded in MPEG-2 as a function of the bit rate, for SD, VGA, CIF and QCIF display format

The best values for a can be obtained for each display format. For example, for CIF, this value can be calculated as follows: For each MOS, a value of a can be calculated as the ratio between the bit rates of CIF and SD for this MOS. For example, if MOS=3 for 1 Mb/s in SD and for 0.28 Mb/s in CIF, then, in this case, a=1/0.28=3.57 for CIF. The same procedure can be done for all the clips, and for many MOS values. Then all the a values can be averaged, in order to obtain only one value for the display format.Similar calculations can be performed for VGA, QCIF and any other display format.

Using this definition, the best values for the coefficient a are presented in Table 7 for SD, VGA, CIF and QCIF. Other display formats can be included, as needed. The best values for v_3, v_4 and v_5 were re-calculated, for all the clips in the four display formats, and are presented in Table 8. These values are very similar to the values presented in Table 6 (only for SD display format), as expected.

Table 7. Best values for *a*

Display Format	*a*
SD	1
VGA	1.4
CIF	3.2
QCIF	10.8

Including the parameter a, the same values of v_3, v_4 and v_5 can be used for different display formats, making the model more general. The parameter a is independent from the function used to model the perceived quality, and can be calculated as expressed in this section.

Table 8. v_4 and v_5 values that best fits to each group for MPEG-2 for all display formats

Movement	v_3	v_4	v_5	MSE
Low Movement	4	0.366	1.32	0.095
Medium Movement	4	0.67	1.36	0.097
High Movement	4	1.088	1.56	0.15

4.4 Codecs

According to G.1070 Model, the v_3, v_4 and v_5 parameters must be calculated for each particular codec, using subjective tests and least square errors approximations. There are many different codecs in the market, and each codec can have its own profile and configurations. Currently, G.1070 only provides provisional values for MPEG-4 in QVGA and QQVGA display formats.

We propose a different approach to include different codecs in the model. MPEG-2 is widely used, and in many cases establishes a lower limit for the video quality, at a given display format and bit rate. Many other codecs have better performance than MPEG-2 (i.e. MPEG-4, H.263, H.264, etc.) [20] [21] [22].

For the same clip, at the same bit rate and display format, the relation between the perceptual quality for a given codec and MPEG-2 is the enhancement factor from the codec with respect to MPEG-2. We will call this enhancement factor k, using the definition presented in Equation (6). Then, the model can be modified, using the best v_3, v_4 and v_5 values for MPEG-2, and including the factor k for other codecs, as shown in Equation (7).

$$k = \frac{MOS_{CODEC} - 1}{MOS_{MPEG-2} - 1} \qquad (6)$$

$$V_q = 1 + kI_c \qquad (7)$$

The value of k has been computed for H.264/AVC, using all the video clips listed in Table 2, at different bit rates and in different formats (using the coding parameters detailed in Table 3). Figure 4 shows the relation between k and the "scaled" bit rate (i.e. $a.b$).

This relation can be modeled with equation (8), proposing an exponential model.

$$k = 1 + k_1 . e^{-k_2 ab} \qquad (8)$$

Where a depends on the display format as detailed in Table 7, and k_1 and k_2 must be calculated in order best fit equation (8) to the actual values.

Using the source clips detailed in Table 2, coded in the different display formats and bit rates, the best values for k_1 and k_2 were calculated, and are the following:

$k_1 = 1.36$, $k_2 = 1.93$

For higher bit rates, k tends to 1, meaning that the H.264 codec is in average equivalent to MPEG-2 for high bit rates. On the other hand, for low bit rates, H.264 is in average much better than MPEG-2, regarding the perceptual quality obtained for the same bit rate.

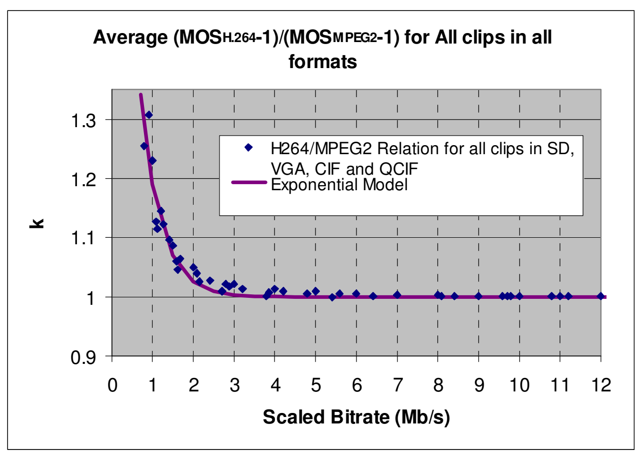


Figure 4. Perceived quality relation between MPEG-2 and H.264 as a function of the scaled bit rate (i.e., *a.bitrate*), averaged for all the video clips in all display formats

All the proposed enhancements can be including in the model, as shown in Equation (9).

$$V_q = 1 + 4k\left(1 - \frac{1}{1 + \left(\frac{ab}{v_4}\right)^{v_5}}\right) \qquad (9)$$

Where k depends on the codec, with the following values:

$k = 1$ for MPEG-2

$k = 1 + k_1.e^{-k_2.a.b}$ for H.264

Figure 5 shows the perceived quality for clips coded in MPEG-2 and in H.264, in SD and CIF display format, with high and low movement content, and the curve derived from equation (9) using the values for a, v_3, v_4 and v_5 detailed in Table 7 and Table 8 respectively. As can be seen, with the proposed enhancements, the model fits very well with the actual MOS values, with a MSE lower than 0.22, computed for more than 1500 processed video clips coded in MPEG-2 and H.264, in SD, VGA, CIF and QCIF, and at different bit rates.

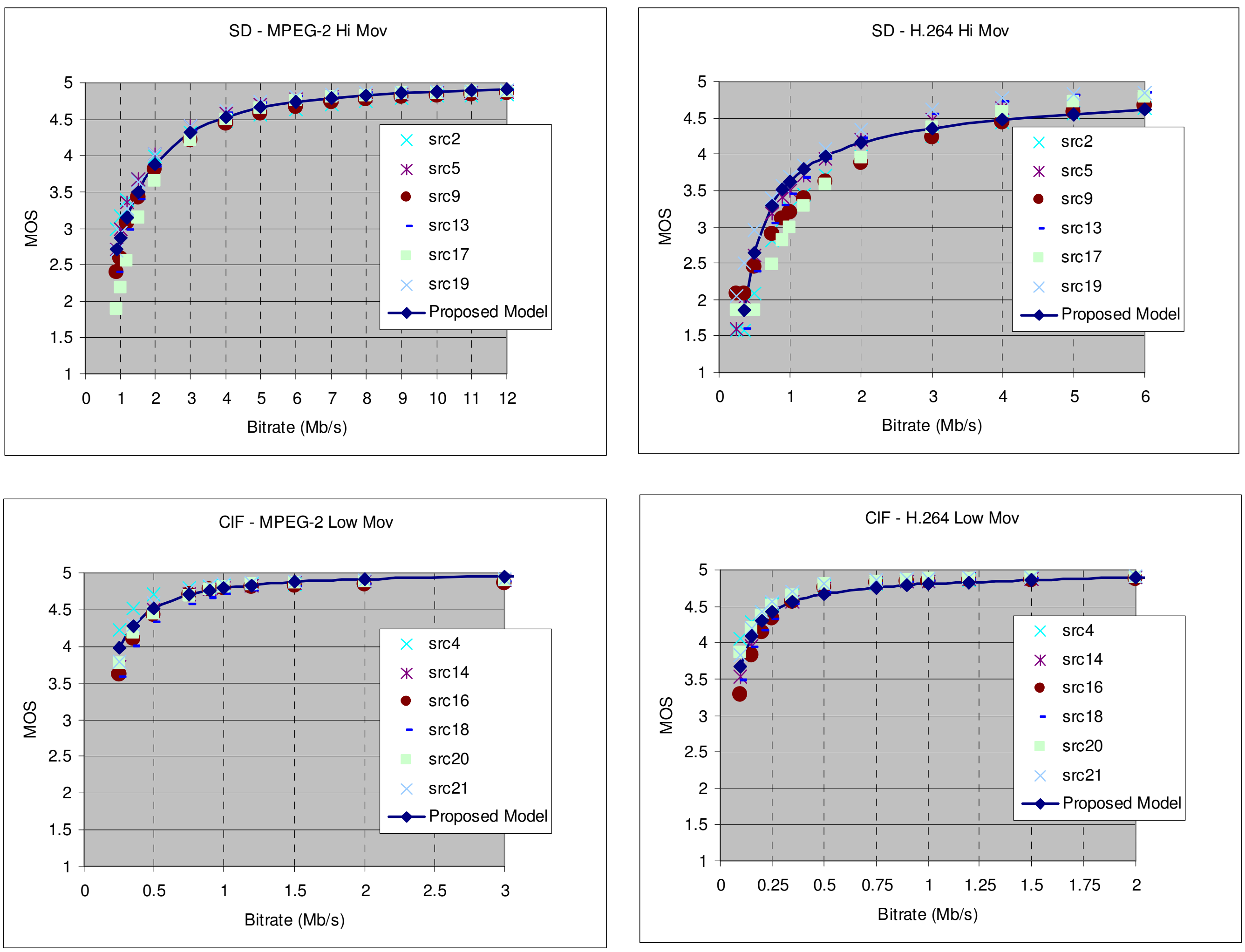


Figure 5. Examples of perceived quality computed with ITU model and estimation using the proposed model, for different codecs and display formats.

5. CONCLUSION

The video quality estimation proposed in ITU-T Recommendation G.1070 "Opinion model for video-telephony applications" cannot model properly the perceptual video quality, especially in the low bit rate range, due to the great variation of MOS values depending on video content. In this work, enhancements have been proposed to the model, allowing a much better approximation to the perceptual MOS values, knowing only the subjective movement content in the video application, classified in "Low", "Medium" or "High".

The original 3-parameters model was reformulated, reducing it to 2-parameters, with a strong relation to movement content.

Two new parameters were added to the model, with a clear interpretation: One is related to the display format, and the other is related to the codec performance compared to MPEG-2. With these additions, the model can be easily extended to different display formats and codecs. The two new parameters can be calculated independently, and with no relation to the other model parameters.

The proposed model enhancement have been evaluated, for MPEG-2 and H.264, in SD, VGA, CIF and QCIF display formats and in the range from 50 kb/s to 12 Mb/s. Sixteen video sources were used, coded in 96 different formats, varying the codec, the bit rate and the display format. In total more than 1500 processed video sequences were analyzed. The parameters values have been calculated. The result shows that the new model fits well with respect to the perceptual video quality estimations.

6. REFERENCES

[1] ITU-T Recommendation G.1070 Opinion model for video-telephony applications, April 2007

[2] ISO/IEC 13818-2:2000. Information technology – generic coding of moving pictures and associated audio information: Video.

[3] ITU-T H.264 Advanced Video Coding for Generic Audiovisual Services, March 2005

[4] Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra: Overview of the H.264 / AVC Video Coding Standard, IEEE Transactions on Circuits and Systems For Video Technology, Vol 13, July 2003

[5] ITU-T Recommendation J.144 Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference, February 2004

[6] ITU-R Recommendation BT.1683 Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference, January 2004

[7] Final Report of VQEG's Multimedia Phase I Validation Test, 19 September 2008

[8] Margaret H Pinson and Stephen Wolf: A New Standardized Method for Objectively Measuring Video Quality, IEEE Transactions on Broadcasting, Volume 50, Issue 3, September 2004, pp. 312-322

[9] Kazuhisa YAMAGISHI and Takanori HAYASHI: Opinion Model for Estimating Video Quality of Videophone Services, IEEE Global Telecommunications Conference, Nov. 27 2006

[10] Takanori Hayashi, Kazuhisa Yamagishi, Toshiko Tominaga, and Akira Takahashi: Multimedia Quality Integration Function for Videophone Services, IEEE Global Telecommunications Conference, 26-30 Nov. 2007

[11] ITU-T Recommendation G.107 The E-model, a computational model for use in transmission planning,March 2005

[12] K Yamagishi and T Hayashi: Parametric Packet-Layer Model for Monitoring Video Quality of IPTV Services, IEEE International Conference on Communications 2008 (ICC 08), 19 May 2008

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

[14] ITU-T Recommendation P.910. Subjective video quality assessment methods for multimedia applications, September 1999

[15] Sungdeuk Cho, Jihwan Choe, Taeuk Jeong, Wonseok Ahn, and Eunjae Lee: Objective video quality assessment, Optical Engineering Vol. 45 (1), January 2006

[16] Alengar Lotufo, R Da Silva, W D F Falcao, A X Pessoa: Morphological image segmentation applied to video quality assessment, IEEE Proceedings in Computer Graphics, Image Processing and Vision, SIGGRAPI Proceedings, pp 468-475, October 1998

[17] Alexandre J Bourret, David S Hands, Damien Bayart, Andrew G Davies: Method and System for Video Quality Assessment, US Patent No. 2006/0152585 A1, July 13, 2006

[18] Video Quality Metric (VQM) Software [Online]. Available at: www.its.bldrdoc.gov/n3/video/vqmsoftware.htm

[19] VQEG Phase I Test Sequences. [Online]. Available at: ftp://vqeg.its.bldrdoc.gov/SDTV/VQEG_PhaseI/TestSequences/Reference/

[20] Anthony Joch, Faouzi Kossentini, Heiko Schwarz, Thomas Wiegand, Gary J. Sullivan: Performance Comparison of Video Coding Standards Using Lagrangian Coder Control, IEEE IFIP 2002

[21] Wiegand, T. Schwarz, H. Joch, A. Kossentini, F. Sullivan, G.J: Rate-constrained coder control and comparison of video coding standards, IEEE Transactions on Circuits and Systems for Video Technology, July 2003

[22] Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra: Overview of the H.264 / AVC Video Coding Standard, IEEE Transactions on Circuits and Systems For Video Technology, Vol 13, July 2003