Combining the Effects of Frame Rate, Bit Rate, Display Size and Video Content in a Parametric Video Quality Model

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 analyze how the frame rate affects the video perceived quality for different video contents. Combining these results with our previous work, a new parametric model is proposed for video quality estimation, including the effects of frame rates, bit rate, display size and video content. The performance of the new proposed model was evaluated for video clips coded in H.264/AVC at different frame rates (from 5 to 25 fps), different bit rates (from 25 kb/s to 6 Mb/s), and in different display formats, including VGA, CIF and QCIF. In total more than 670 processed video sequences were analyzed to derive the proposed model.

The new proposed model is compared with other four recently proposed parametric models that take into account the frame rate in the video quality estimation, concluding that the proposed model has better performance than the other four.

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
Video and multimedia applications are growing fast in the market. These applications include videophones, videoconferencing systems, video on demand and IPTV, among others. In these new services, is critical to guarantee an appropriate QoE (Quality of Experience) for the end user, according to the offered application. QoE can be defined as the overall performance of a system, from the user perspective. Many factors can affect the QoE, depending on the application and users expectations. Video quality is one of the most important aspects to consider in the user QoE for multimedia applications. With digital video coding and distribution, new artifacts are presented, affecting the video perceived quality, and the final QoE.[1].

Different evaluations and standardized efforts have been made, and are currently ongoing, in order to derive objective models and algorithms to predict the perceived video quality in different scenarios [2] [3].

Picture metrics, or media-layer models, are based on the analysis of the video content. This metrics can be classified into FR (Full Reference), RR (Reduced Reference) and NR (No Reference) models. In the first one, FR models, the original and the degraded video sequences are directly compared. In the RR models, some reduced information about the original video is needed, and is used along with the degraded video in order to estimate the perceived video quality. NR models are based only in the degraded video in order to make an estimation of the perceived video quality.

Data metrics, or packet-layer models, are based on network information (i.e. IP packets). This metrics can be classified into packet-header models, bit-stream-layer model and hybrid models. The packet-header models use only general information about the network (i.e. packet loss rates), and does not take into account packet contents. Bit-stream-layer models can access IP packets payload, and extract some media related information. Hybrid models use a combination of the other methods.

Parametric models predicts the perceived video quality metrics based on some reduced set of parameters, related to the encoding process, video content and/or network information. These models typically present a mathematical formula, representing the estimation of the perceived video quality as a function of different parameters. Parametric models can be applied to packet-layer models, media-layer models or a combination of both.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
LANC 2011, October 12-13, 2011 Quito, Ecuador
Copyright © 2011 ACM ISBN 978-1-4503-0886-1/11/10... $10.00

One of the fundamentals factors affecting the perceived video quality is the degradation introduced by the encoding process. Different parametric models have been proposed, in order to predict the perceived video quality based on some encoding parameters. However, most of them are applied to some specific applications, display formats or codecs, and are not valid (or were not tested) in other environments.

Towards the development of a general parametric model, in our previous works [4][5][6][7], the combined effects of bit rate, display size, codec and video content were evaluated, and a parametric model was proposed for video quality estimation. In this work, we analyze the effects of the frame rate in the perceived video quality for clips coded in H.264/AVC [8] in VGA (Video Graphics Array, 640 × 480 pixels), CIF (Common Intermediate Format, 352 × 288 pixels) and QCIF (Quarter Common Intermediate Format, 176 × 144 pixels) display sizes, and our previous proposed model is extended, in order to take into account these results.

The paper is organized as follows: Section 2 describes the video quality metric used in this work. Section 3 describes how perceived quality varies with respect to the bit rate, and the previous proposed model is presented. In Section 4 the effects of frame rate are thoroughly studied, and a new parametric model is presented. In section 5 the results are evaluated and the performance of the proposed model is compared with other published parametric models. Section 6 summarizes the main contributions.

2. VIDEO QUALITY METRIC

In this work we used the "Low Bandwidth Reduced Reference model" proposed by NTIA (National Telecommunications and Information Administration), standardized in Recommendation ITU-T J.249 [9] as the VQM (Video Quality Metric). The performance of this model for SD (Standard Definition) display size and 25 / 30 fps was well established in the VQEG (Video Quality Experts Group) RR/NR TV evaluations [10]. The result of these evaluations shows Pearson Correlation values from 0.82 to 0.88 between the subjective scores and the scores predicted from the NTIA model. The Pearson correlation metric evaluates the precision of the prediction. It varies from 0 to 1, where 1 indicates a direct relationship and 0 indicates no relationship at all. In this case, 0.82-0.88 indicates a high correlation between the values. The NTIA model was originally designed and trained also for smaller display sizes and lower bitrates, and the overall performance of the model was presented in [11].

For each video clips pair (original and degraded), the NTIA model provides a 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) [12] and can be directly associated with the DMOS (Difference Mean Opinion Scores).

In order to make an independent validation of the NTIA "Low Bandwidth RR Model", we compared the model results with the subjective scores obtained for 6 different video clips, coded at different bit rates, in CIF display format, with frame rates from 3.75 to 30 fps. The video clips and the subjective scores used were obtained from the Quality Assessment Database of the Video Lab from Polytechnic Institute of NYU [13]. The results are plotted in Figure 1. The Pearson Correlation value between the subjective scores and the NTIA "Low Bandwidth RR Model" is 0.91 and the RMSE (Root Mean Square Error) is 0.14. These results are better than the obtained in the VQEG model evaluation for SD display size and 25/30 fps. Using these results, the error of the NTIA model with respect to subjective scores can be estimated in +/- 15%.

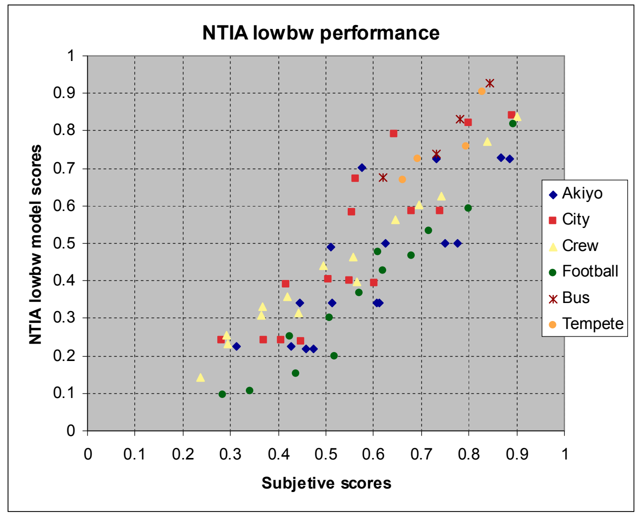


Figure 1. NTIA lowbw model evaluation

According to the previous results, we decided to use in this work the model proposed by NTIA, standardized in ITU-T J.249, available in [14] as the benchmark VQM. The DMOS values returned from the NTIA model can be related to the typical 5 points MOS (Mean Opinion Score) using Equation (1). The interpretation of the MOS values is presented in Table 1. MOS varies between 1 (Bad quality) and 5 (Excellent Quality).

$$MOS = 5 - 4DMOS \quad (1)$$

Table 1. MOS to perceived quality relation

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

3. PERCEIVED QUALITY AS A FUNCTION OF THE BIT RATE AND VIDEO CONTENT

The video clips detailed in Table 2, available in the VQEG web page [15], were used for the following sections.

Figure 2 shows how the perceived video quality varies as a function of the bit rate, keeping constant all other encoding parameters, for the clip "Rugby" (src 9), coded in H.264/AVC, at

25 fps, in different display sizes. The figure shows the typical behavior for any video clip:

a) The perceived quality is higher for higher bit rates.

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

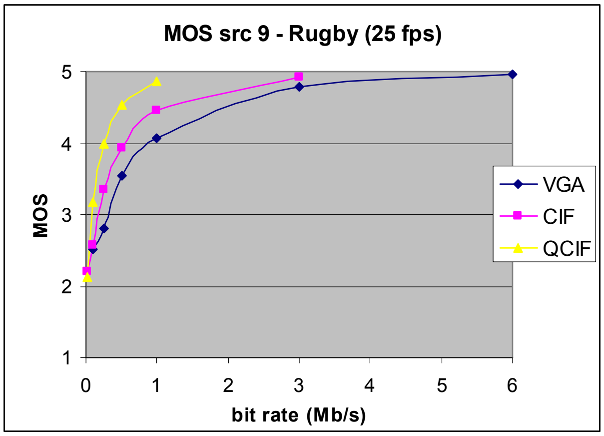


Figure 2. MOS vs bit rate for scr 9 "Rugby" at VGA, CIF and QCIF, at 25 fps

Figure 3 shows the relation between MOS and bit rate, for all the clips detailed in Table 2, coded in H.264/AVC (using the coding parameters detailed in Table 3), at 25 fps, in VGA display format. MOS values were derived from DMOS, using Equation (1). DMOS values were calculated using the NTIA Model.

Table 2. Video clips used

Src	Name	Avg SAD/ pixel	Mov
2	Barcelona	4.243	High
3	Harp	3.457	Med
4	Moving graphic	0.684	Low
5	Canoa Valsesia	5.148	High
7	Fries	3.632	Med
9	Rugby	6.164	High
10	Mobile & Calendar	3.600	Med
13	Baloon-pops	5.656	High
14	New York 2	1.386	Low
16	Betes pas betes	1.804	Low
17	Le point	8.256	High
21	Susie	1.251	Low
22	Tempete	3.315	Med

Table 3. H.264/AVC encoding parameters

H.264/AVC
Profile/Level: High/3.2
GOP size: 33
2 B Pictures between I&P
Entropy Coding: CABAC
Subpixel mode: ¼ Pixel
Bit rate type: CBR
Interlacing: Non-Interlaced

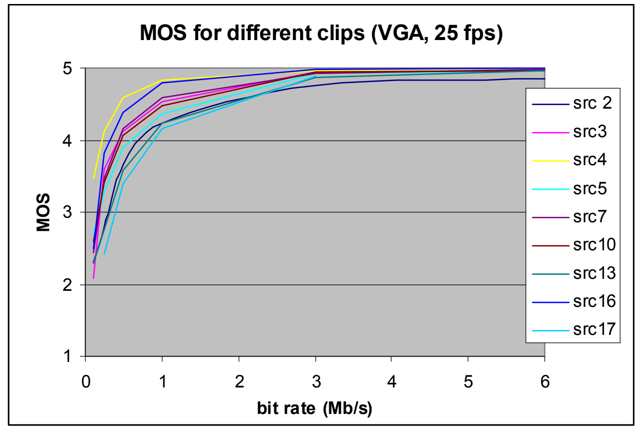


Figure 3. MOS vs bit rate for different clips at 25 fps.

In our previous works [6][7] we proposed a parametric model for video quality estimation due to encoding degradation according to Equation (2)

$$V_q = 1 + I_c \qquad (2)$$

were $\boldsymbol{V_q}$ is the video quality estimation (the predicted MOS or MOS_p) and $\boldsymbol{I_c}$ is the video quality due to the encoding process, with values between 0 and 4. $\boldsymbol{I_c}$ was defined as shown in Equation (3)

$$I_c = v_3 \left(1 - \frac{1}{1 + \left(\frac{ab}{v_4} \right)^{v_5}} \right) \qquad (3)$$

were $\boldsymbol{I_c}$ is the video quality due to the encoding process, $\boldsymbol{a}$ depends on display size according to Table 4, $\boldsymbol{b}$ is the bit rate (in Mb/s) and $\boldsymbol{v_3}$, $\boldsymbol{v_4}$ and $\boldsymbol{v_5}$ are model coefficients.

Table 4. *a* values for different display sizes

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

This model was derived using video clips coded in bit rate ranges from 50 kb/s to 12 Mb/s, in SD, VGA, CIF and QCIF display formats at 25 fps.

The coefficient v_3 is the maximum video quality achievable by the encoder, for high bit rates (i.e., when $b \rightarrow \infty$). In [6] we proposed to set $v_3 = 4$ (looking into Figure 2 and Figure 3, for high bit rates, all clips tend to have MOS = 5 ($v_3 = 4$) for all video contents and for all display sizes). In the same work, we showed that the coefficients v_4 and v_5 depends on the spatial and temporal activity of the video clip, and can be derived from the average SAD (Sum of Average Differences) per pixel of the clip, according to Equation (4). SAD is a simple video metric used for block comparison and for moving vectors calculations. Each frame is divided into small blocks (i.e. 8x8 pixels) and for every block in one frame the most similar (minimum SAD) block in next frame is find. This minimum sum of absolutes differences is assigned as the SAD for each block in each frame (up to the n-1 frame). Then all the SAD values are averaged for each frame and for all the frames in the clip, and divided by the block area, for normalization. This value (average SAD/pixel) provides an overall estimation about the spatial-temporal activity of the entire video clip.

$$\begin{aligned} v_3 &= 4 \\ v_4 &= c_1 s^{c_2} + c_3 \qquad (4) \\ v_5 &= c_4 s^{c_5} + c_6 \end{aligned}$$

Combining (3) and (4), the previous proposed model depends only on three parameters: the bit rate b, the display size a and the average SAD/pixel s. The coefficients $c_1..c_6$ can be different for each codec (i.e. MPEG-2, H.264). The effects of frame rate variations were not included in the model (the frame rate was fixed to 25 fps). These effects are presented in next section.

4. PERCEIVED QUALITY AS A FUNCTION OF THE FRAME RATE

Figure 4 shows how the perceived quality varies as a function of the bit rate, keeping constant all other coding parameters, for the clip "Rugby" (src 9), coded in VGA H.264/AVC for different frame rates. Similarly, Figure 5 shows the same information for the clip "New York" (src 14).

These figures show that the maximum video quality achievable, for high bit rates (i.e., when $b \rightarrow \infty$) depends on the frame rate and on video content. For 25 fps, the maximum MOS is always 5, but for lower frame rates, the maximum MOS decreases. On the other hand, for low bit rates, video quality can be improved using lower frame rates for a given bit rate. For a given bit rate, at lower frame rates, each frame can be encoded with higher quality. This is specially perceived for low bit rates, and as a result, the overall perceived video quality can be better for lower frame rates. In Figure 5 this effect can be seen for bit rates lower than 250 kb/s.

A new term can be added to Equation (2) in order to include the frame rate effect, as expressed in Equation (5)

$$V_q = 1 + I_c I_f \qquad (5)$$

were I_f represents the frame rate effects on the video quality. For 25 fps $I_f = 1$, but for lower frame rates, I_f depends on the frame rate and video content.

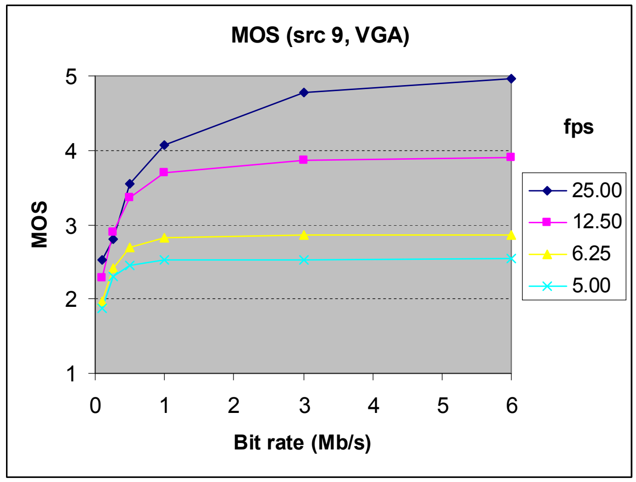


Figure 4. MOS vs bit rate for scr 9 "Rugby" at different frame rates

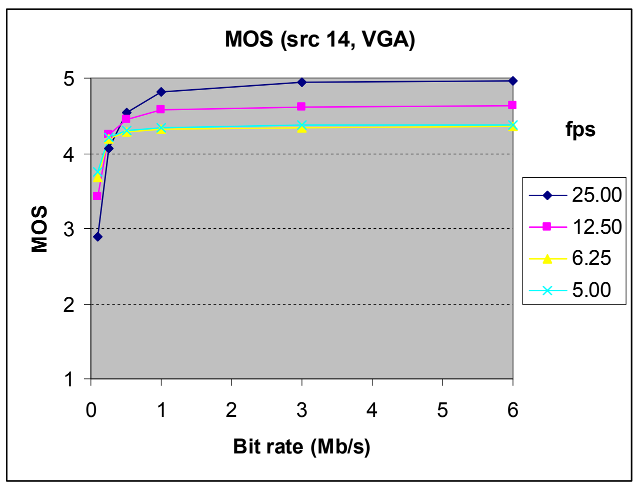
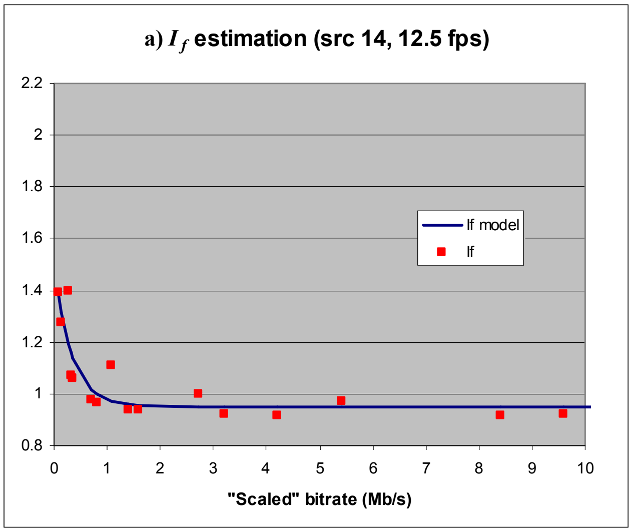


Figure 5. MOS vs bit rate for scr 14 "New York" at different frame rates

Best values for I_f were calculated for different clips (src 2, 3, 4, 5, 7, 9, 10, 14, 16, 17) and different display sizes (QCIF, CIF, VGA) coded in H.264/AVC with bit rate from 25 kb/s to 6 Mb/s and frame rate from 5 to 25 fps. Figure 6 a, shows a plot of I_f as a function of the "scaled bitrate" (**a**.bitrate) for 12.5 fps for the clip src 14 ("New York"), and Figure 6 b, shows similar information for 6.25 fps. For all other clips and frame rates, similar curves are obtained. These curves can generally be modeled using Equation (6)

$$I_f = d_1 + d_2 e^{-d_3 ab} \qquad (6)$$

Were a depends on display size, according to Table 4, b is the bit rate, and d_1, d_2 and d_3 are three coefficients that can still depend on frame rate and video content.



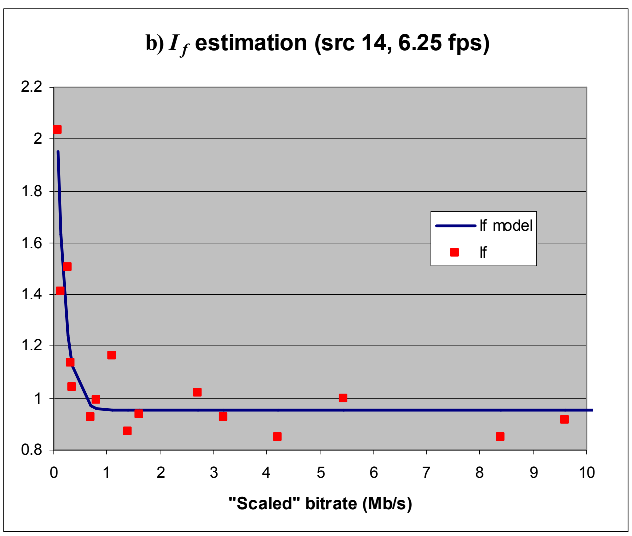


Figure 6. I_f as a function of the "scaled bitrate" (*a*.bitrate) for the clip src 14 at 12.5 fps (a) and 6.25 fps (b)

The coefficient d_1 is the maximum I_f value, for high bit rates (i.e., when $b \to \infty$). We have previously showed that this limit depends on video content and frame rate. Figure 7 shows the best d_1 values for different video clips as a function of the frame rate.

The relation between d_1 and frame rate and video content can be modeled with Equation (7)

$$d_1 = 1 + k_1 s (f_{\max} - f) \qquad (7)$$

Were f is the frame rate (in fps), f_{max} is 25 fps, s is the average SAD per pixel and k_1 is a constant. For any video content, $d_1 = 1$ for $f = f_{max}$. For a given frame rate (lower than f_{max},), d_1 depends on video content according to (7).

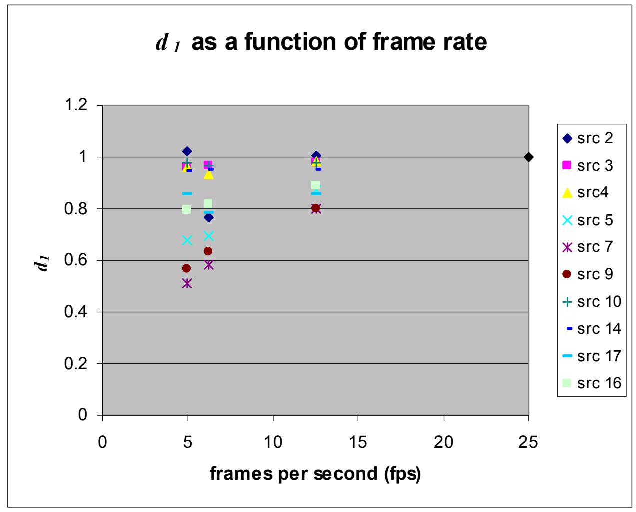


Figure 7. d_1 vs frame rate for all the clips

Figure 8 shows the variation of d_2 as a function of frame rate, averaged for all the clips. In a similar way, Figure 9 shows the variation of d_3 as a function of frame rate, averaged for all the clips. These relations can be modeled with Equation (8)

$$\begin{aligned} d_2 &= k_2 (f_{\max} - f) \\ d_3 &= k_3 (f_{\max} - f) \end{aligned} \qquad (8)$$

Were f is the frame rate (in fps), f_{max} is 25 fps, and k_2, k_3 are constants.

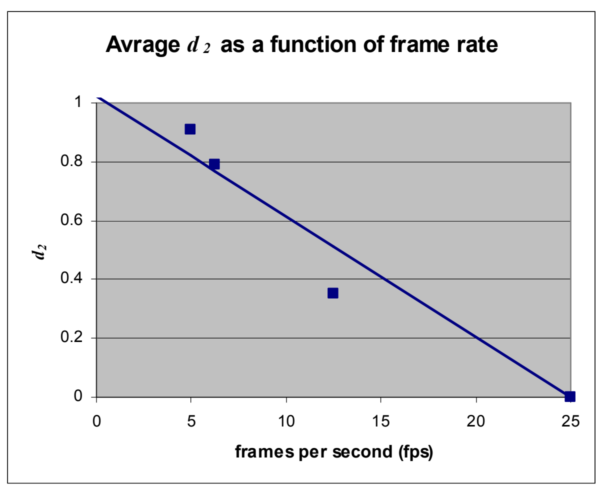


Figure 8. Average d_2 vs frame rate for all the clips

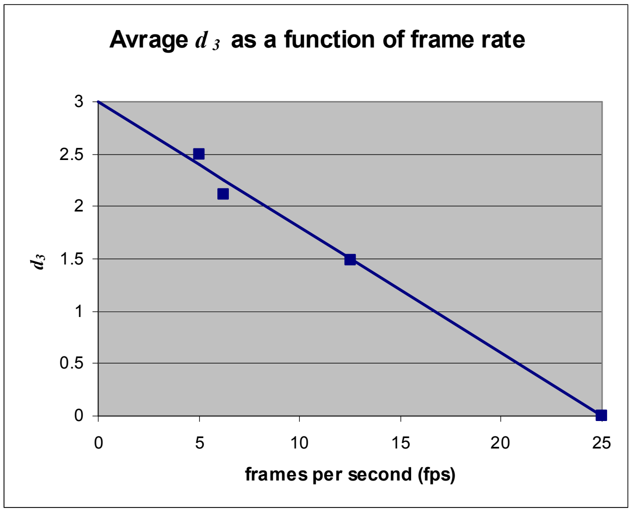


Figure 9. Average d_3 vs frame rate for all the clips

Combining Equations (6), (7) and (8), I_f can be expressed as Equation (9):

$$I_f = 1 + (f_{\max} - f)(k_1 s + k_2 e^{-k_3 (f_{\max} - f) ab}) \qquad (9)$$

were f is the frame rate (in fps), f_{max} is 25 fps, s is the average SAD per pixel, a depends on display size according to Table 4, b is the bit rate, and k_1, k_2, k_3 are constants coefficients.

5. RESULTS AND COMPARISION TO OTHER MODLES

The final model consists in Equation (5), with I_c according to Equation (3) (with v_3, v_4 and v_5 according to equation (4)) and I_f according to Equation (9). The best model coefficients were calculated for clips src 2, 3, 4, 5, 7, 9, 10, 14, 16, 17 (included in Table 2), coded in H.264/AVC in VGA, CIF and QCIF display sizes at different bit rates and with frame rates from 5 to 25 fps. First, the best values for c_1.. c_6 were calculated, for 25 fps. Then, the best k_1.. k_3 values were calculated using different frame rates. Table 5 resumes the best values for all the coefficients. Figure 10 shows the dispersion comparing the MOS obtained using the proposed parametric model with the MOS according to the standardized NTIA model, for the clips src 2, 3, 4, 5, 7, 9, 10, 14, 16, 17. The dotted lines corresponds to a +/- 15%, the estimated error of the NTIA model. The Pearson correlation between the values is 0.90, and the RMSE is 0.36. Only 8% of the points are outside the +/- 15% range.

Table 5. Best values for the model coefficients

Coef	c_1	c_2	c_3	c_4	c_5	c_6
Value	0.030	1.24	0.15	0	0	1.00

Coef	k_1	k_2	k_3
Value	-0.0015	0.041	0.12

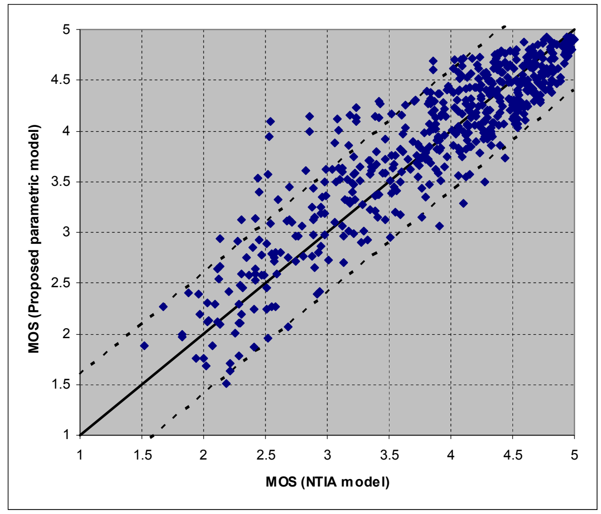


Figure 10. Dispersion between proposed parametric model and standard NTIA model

The model was applied to other 3 clips (src 13, 21 and 22), not used in the "training" process (i.e., not used to obtain the model coefficients). These clips includes scenes with different spatial temporal contents, one with low spatial temporal activity, one with medium spatial temporal activity and one with high spatial temporal activity. The results are showed in Figure 11, obtaining a Person Correlation of 0.95 and a RMSE of 0.044, with only 1.6% of the points outside the +/- 15% range. As can be seen, very good results are obtained with the proposed parametric model.

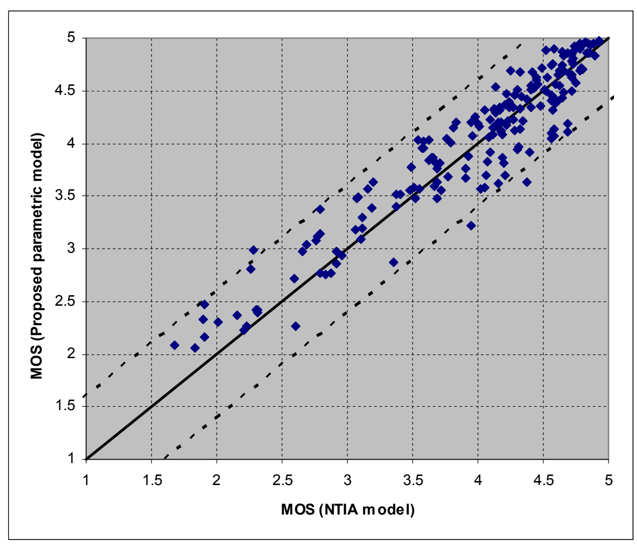


Figure 11. Dispersion between proposed parametric model and standard NTIA model for 3 clips not used in the training process

Other parametric models have been proposed in recent years. In [16] M. Ries et al. proposed a model for video quality estimation according to Equation (10)

$$I_c = A + Bb + \frac{C}{b} + Df + \frac{E}{f} \qquad (10)$$

were b is the bit rate, f is the frame rate, and A, B, C, D, E are the model coefficients. The authors proposed to classify the video clips according to the video content, and for each class, a different set of coefficients are used.

We have classified the clips used in this work (Table 2) in three different classes, according to the spatial temporal activity: High, Medium and Low, and for each class, the best A, B, C, D, E were calculated for VGA display size. The overall Pearson correlation obtained for the M. Ries et al. model was 0.77, with an RMSE of 0.51. With this model, 21% of the points are outside the +/- 15% range.

In [17] A. Khan et al. proposed a model for video quality estimation according to Equation (11)

$$I_c = a_1 + a_2 f + a_3 \ln(b) \qquad (11)$$

were b is the bit rate, f is the frame rate, and a_1, a_2, a_3, are the model coefficients. The authors proposed to classify the video clips in three categories: Slight Movement, Gentle Walking and Rapid Movement.

We have classified the clips used in this work (Table 2) in the same three categories, according to the spatial temporal activity, and for each class, the best a_1, a_2, a_3, were calculated for VGA display size. The overall Pearson correlation obtained for the A Khan et. al model was 0.76, with an RMSE of 0.54. With this model, 25% of the points are outside the +/- 15% range.

In [18] Yen-Fu Ou et al. proposed a model for video quality estimation according to Equation (12), based in the analysis of subjective tests in CIF and QCIF display formats, for frame rates between 6 and 30 fps.

$$I_c = I_{c_{\max}} \frac{1 - e^{-c\frac{f}{f_{\max}}}}{1 - e^{-c}} \qquad (12)$$

were f is the frame rate, f_{max} is the maximum frame rate (i.e. 25 or 30 fps), I_{cmax} is the video quality obtained for the video clip coded at f_{max} and c is the model coefficient. The authors show that the coefficient c is in some way related to the video content, but no explicit mathematical relation is provided. This model assumes that the video quality is always degraded for lower frame rates, for any bit rate. We have shown in previous sections, that video quality can be improved for lower frame rates and low bit rates, due to the fact that each frame can be encoded with higher quality.

Again we used the same classification of the clips used in this work (Table 2) in three categories, according to the spatial temporal activity, and for each class, the best value for c was calculated for VGA display size. The overall Pearson correlation obtained for the Yen-Fu Ou et al. model was 0.70, with an RMSE of 0.70. With this model, 31% of the points are outside the +/- 15% range.

Takanori Hayashi et al. proposed a video quality estimation model [19], standardized in ITU-T Recommendation G.1070 [20]. The proposed video quality model can be expressed as Equation (13)

$$I_c = I_{c_{\max}} e^{-\frac{\left(\ln(f) - \ln(v_1 + v_2 b)\right)^2}{2(v_6 + v_7 b)^2}} \qquad (13)$$

Were b is the bit rate, f is the frame rate, v_1, v_2, v_6, v_7, are model coefficients and I_{cmax} is similar to Equation (3), with a =1. In this model, video content is not taken into account, so all the coefficients are calculated averaging all the video clips. The best v_1, v_2, v_6 and v_7 were calculated for VGA display size (for best fitting with all the clips detailed in Table 2). The overall Pearson correlation obtained for the ITU-T G.1070 model was 0.80, with an RMSE of 0.80. With this model, 32% of the points are outside the +/- 15% range.

In Table 6 a summary of the models performance is presented. The proposed model has higher Pearson correlation, lower RMSE and lower percentage of points outside the +/-15% than the other evaluated models.

Table 6. Models performance comparison

Model	Pearson Correlation	RMSE	Points outside +/- 15%
Proposed model	0.90	0.36	8%
M. Ries	0.77	0.51	21%
A. Khan	0.76	0.54	25%
Yen-Fu Ou	0.70	0.70	31%
ITU-T G.1070	0.80	0.80	32%

6. CONCLUSION

In this work we studied the influence of frame rate, bit rate, display size and video content in the perceived video quality. A new parametric model for perceptual video quality estimation was proposed, which provides a very good estimation to the MOS values, for different bit rates, frame rates, display sizes and video content. The spatial-temporal activity, related to the video content, is derived from the average SAD per pixel of the clip. SAD is a simple video metric, commonly used for block comparison and for moving vectors calculations. The proposed model depends on four parameters: bit rate b (in Mb/s), frame rate f (in fps), display size a (according to tabulated values for each display size) and video content s (measured as the average SAD/pixel of the original video clip). The parameters are combined with a set of nine fixed coefficients that must be adjusted for each codec.

The NTIA "Low Bandwidth Reduced Reference model", standardized in Recommendation ITU-T J.249 was used as the VQM, against which the parametric models were compared. In this work, the performance of this standardized model was evaluated for low frame rates and small display sizes, showing that this model has a Pearson correlation of 0.91 and the RMSE of 0.14 with respect to subjective test.

The best coefficients for the proposed parametric model were calculated using 10 different video clips including different types of scenes (sports, cartoons, head and shoulders, panorama, and so) coded in H.264/AVC at bit rates from 25 kb/s to 6 Mb/s and frame rates from 5 fps to 25 fps, in VGA, CIF and QCIF display sizes. In total more than 670 processed video sequences were analyzed. The overall model performance was evaluated comparing the quality estimation derived from the proposed parametric model with the standard VQM, obtaining Pearson correlation of 0.90, RMSE of 0.36 and only 8% of the estimations with errors larger than 15% (the estimated error of the VQM with respect to subjective scores).

The proposed parametric model performance was compared to other four recently proposed parametric models that take into account the bit rate and frame rate. The comparison results show that the proposed model outperforms the other four models, with higher Pearson correlation, lower RMSE and lower points outside the +/-15% range.

The performance of the proposed model was also tested against three video clips not used to derive the model coefficients. These clips include scenes with different spatial temporal contents. The results obtained a Person Correlation value of 0.95, an RMSE value of 0.044, and only 1.6% of the points outside the +/- 15% range.

The final result shows that the video quality estimation calculated with the proposed model fits very well with respect to the perceptual video quality estimations derived from the standardized VQM.

7. REFERENCES

[1] H. R.Wu and K. R. Rao: Digital Video Image Quality and Perceptual Coding, Eds. Boca Raton, FL: CRC Press, 2005

[2] Stefan Winkler: Video Quality and Beyond, European Signal Processing Conference (EURASIP), 2007

[3] Stefan Winkler and Praveen Mohandas: The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics, IEEE Transactions on Broadcasting, Vol. 54, No. 3, September 2008

[4] Jose Joskowicz, José-Carlos López-Ardao, Miguel A. González Ortega and Cándido López García: A Mathematical Model for Evaluating the Perceptual Quality of Video, FMN 2009, LNCS 5630, pp. 164–175, 2009

[5] Jose Joskowicz, J. Carlos López Ardao: Enhancements to the Opinion Model for Video-Telephony Applications, LANC'09, Sep 24–25, 2009, Pelotas, Brazil , pp. 87-94

[6] Jose Joskowicz, J. Carlos López Ardao: A General Parametric Model for Perceptual Video Quality Estimation, 2010 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR), June 8-10 2010, Vancouver, Canada

[7] Jose Joskowicz, J. Carlos López Ardao: A Parametric Model for Perceptual Video Quality Estimation, Telecommunication Systems, Springer Special Issue on the Quality of Experience issues in Multimedia Provision, Volume 49: 1 -2

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

[9] ITU-T Recommendation J.249 Perceptual video quality measurement techniques for digital cable television in the presence of a reduced reference, January 2010

[10] VQEG: Validation of Reduced-Reference and No-Reference Objective Models for Standard Definition Television, Phase 1, 2009

[11] Stephen Wolf and Margaret H. Pinson: Low Bandwidth Reduced Reference Video Quality Monitoring System, First Int'l Workshop on Video Proc. and Quality Metrics, Jan 2005

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

[13] Quality Assessment Database of the Video Lab from Polytechnic Institute of NYU [Online]. Available at: http://vision.poly.edu/index.html/index.php?n=HomePage.QualityAssessmentDatabase

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

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

[16] Ries, M.; Crespi, C.; Nemethova, O.; Rupp, M: Content Based Video Quality Estimation for H.264/AVC Video Streaming, Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE, 11-15 March 207, pp. 2668-2673

[17] Asiya Khan, Lingfen Sun and Emmanuel Ifeachor: Content-Based Video Quality Prediction for MPEG4 Video Streaming over Wireless Networks, Journal of Multimedia, Vol 4, No 4 (2009), 228-239, Aug 2009

[18] Yen-Fu Ou, Tao Liu, Zhi Zhao, Zhan Ma, Yao Wang: Modeling the impact of frame rate on perceptual quality of video, 15th IEEE International Conference on Image Processing, 2008, ICIP 2008, pp. 689-692

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

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