A General Parametric Model for Perceptual Video Quality Estimation

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

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

***Abstract*— In this paper a general parametric model is proposed which provides estimation for the perceived quality of video, coded with different codecs, at any bit rate and display format. The proposed model takes into account video content, using an objective estimation of the spatial-temporal activity, based on the average SAD (Sum of Absolute Differences) for the clip. Studies were made for more than 2000 processed video clips, coded in MPEG-2 and H.264/AVC, in bit rate ranges from 25 kb/s to 12 Mb/s, in SD, VGA, CIF and QCIF display formats. The results shows that the proposed model fits very well to the perceived video quality, in any combination of codec, bit rate and display format.**

***Index Terms*—Video perceptual quality, Video codecs, Video signal processing, VoIP Network design**

I. Introduction

Video and multimedia applications are growing fast. In the massive market, different providers are offering video and multimedia applications to the end users, including cable television, Internet service providers, and traditional and emerging telephony carriers, among others. In the corporate market telephony applications are well established, and different video applications are emerging (video-phones, video conferencing, etc.). In these challenging scenarios, it is critical to guarantee an appropriate QoE (Quality of Experience) for the end user, according to the application to be developed. 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. With digital video coding and distribution, new artifacts are presented, affecting the video perceived quality, and the final QoE.

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.

Picture metrics, or media-layer models, are based on the analysis of the video content. These 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.

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. In this work, we make a comparison between different parametric models for predicting the perceived video quality estimation due to coding degradations, and based on the results, a general parametric model is proposed, applicable to a wide range of applications, display formats, codecs and video content.

The paper is organized as follows: Section 2 describes a summary of different Video Quality Metrics. Section 3 describes how the perceived video quality varies as a function of the bit rate. Section 4 presents different existing parametric video quality estimation models. In Section 5 the proposed perceptual video quality estimation model is presented, and the parameters are calculated for MPEG-2 and H.264/AVC codecs, in SD, VGA, CIF and QCIF display formats. Section 6 summarizes the results and main contributions.

II. Video Quality Metrics

The most reliable form for measuring the perceived video quality of a video clip is the subjective tests, where typically video sequences are presented to different viewers, and opinions are averaged. The MOS (Mean 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 Recommendations ITU-R BT.500-11 [1] and ITU-T P.910 [2].

Subjective tests are difficult to implement, and takes considerable time. For these reasons, different objective video quality metrics have been used for video quality evaluation and estimation. Historically, the PSNR (Peak Signal-to-Noise Ratio) picture metric has been used for the evaluation of video quality models. It is now accepted that such quality measures does not match the "perceived" quality by human viewers [3]. Based on VQEG (Video Quality Expert Group) work, ITU has standardized the Recommendations ITU-T J.144 [4] and ITU-R BT.1683 [5] for estimation of the perceived video quality in digital TV applications when the original signal reference is available (FR models). VQEG is working in models evaluation for the estimation of the perceived video quality in multimedia [6] and HDTV (High Definition TV) [7] applications, and is also working in bit-stream and hybrid models evaluation [8].

The models proposed in the Recommendation ITU-T J.144 perform quality comparisons between the degraded and the original video signal, and can be classified as FR models. For each video clip pair (original and degraded), the algorithms provide a VQM (Video Quality Metric), 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) [1] and can be directly related to the DMOS. The statistical error between the average subjective DMOS and the predicted DMOS using Recommendation ITU-T J.144 models can be estimated in +/- 0.1 in the 0-1 scale [9][10].

Instead of doing subjective test, we used in this work the model proposed by NTIA standardized in Recommendation ITU-T J.144, available at [11]. The DMOS values returned form the NTIA model can be related to the MOS using Equation (1). The interpretation of the MOS values is presented in Table 1. MOS errors, using this model, can be estimated in +/-0.4 in the 1-5 scale (4 times the DMOS error).

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

TABLE I. MOS TO PERCEIVED QUALITY RELATION

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

III. Perceived Video Quality As a Function of the Bit Rate

Sixteen video clips, available in the VQEG web page [12], were used. These video clips spans over a wide range of contents, including sports, landscapes, “head and shoulders”, etc. 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 the relation between MOS and bit rate, for all the used clips, coded in MPEG-2 (using the coding parameters detailed in Table 2), in SD display format. MOS values were derived from DMOS, using Equation (1). DMOS values were calculated using the NTIA Model.

TABLE II. MPEG-2 AND H.264 CODING PARAMETERS

MPEG-2	H.264
Profile/Level: MP@ML Max GOP size: 15 GOP Structure: Automatic Picture Structure: Always Frame Intra DC Precision: 9 Bit rate type: Constant Bit Rate Interlacing: Non-Interlaced Frame Rate: 25 fps	Profile/Level: High/3.2 Max GOP size: 33 Number of B Pict between I and P: 2 Entropy Coding: CABAC Motion Estimated Subpixel mode: Quarter Pixel Bit rate type: Constant Bit Rate Interlacing: Non-Interlaced Frame Rate: 25 fps

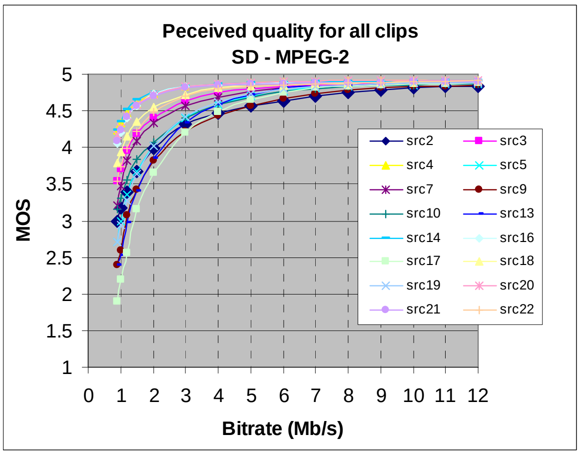

Fig. 1. Perceived Quality as a function of the Bit Rate

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 (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 codecs (i.e. H.264/AVC) and display formats (i.e. VGA, CIF and QCIF).

IV. Existing Parametric Video Quality Models

The curves in Figure 1 represent the perceived quality variation, in function of the bit rate, for clips coded in MPEG-2, due to coding degradations only. Similar curves are obtained with H.264/AVC. These curves can be modeled by different type of relations between the bit rate and the MOS. By definition MOS values can have values between 1 and 5. We will define I_c as the video quality determined by the encoding parameters, and V_q as the estimation for MOS, with the relation presented in Equation (2). I_c varies between 0 and 4, and V_q between 1 and 5.

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

Different functions for I_c were published, each one applicable for some specific conditions (i.e. display formats, codecs, applications, etc.), as described in the following paragraphs.

In [13] an exponential model is presented, for IPTV applications coded in MPEG-2 and H.264, in SD and HD display formats. The proposed model in this paper, called "T-V-Model", is presented in Equation (3), using the parameters names provided in the referred paper.

$$I_c = a_3 - a_1 e^{-a_2 b} \qquad (3)$$

where I_c represents the video quality, determined by the codec distortion, b is the bit rate and a_1, a_2 and a_3 are the three model parameters. The influence of video content in the "T-V-model" is qualitatively described in [14], showing that some spatial and temporal features must be taken into account in order to predict the perceived quality, but a specific quantitative model is not proposed in the mentioned paper.

In [15] the exponential model showed in Equation (4) is presented, for videos coded in MPEG-4 in CIF and QCIF display formats, for multimedia applications that distribute audiovisual content over 3G/4G (3rd/4th generation) networks.

$$I_c = [PQ_H - PQ_L](1 - e^{-\alpha(b - BR_L)}) + PQ_L - 1 \qquad (4)$$

where I_c represents the video quality determined by the codec distortion, b is the bit rate and PQ_H, PQ_L, BR_L and α are the four model parameters. In this work, video quality was evaluated using the MPQoS (Mean Perceived Quality of Service) VQM, a metric based on a Picture Quality Measurement [16]. The model parameters depend on the spatial and temporal activity level of the video clip.

It can be seen that the "T-V Model" represented in Equation (3) and the exponential model represented in Equation (4) are the same, with the parameters relation detailed in Equations (5). These two equivalent models will be called the "Exponential Model" for reference in the present paper.

$$\begin{aligned} a_1 &= (PQ_H - PQ_L)e^{\alpha BR_L} \\ a_2 &= \alpha \\ a_3 &= PQ_H - 1 \end{aligned} \qquad (5)$$

In [17], the relation between the bit rate and the DMOS is modeled with Equation (6), for MPEG-2 and H.264, in different display formats.

$$DMOS = \frac{m}{k.(ab)^n} \qquad (6)$$

where b is the bit rate, a is a constant that depend on the display format (SD, VGA, CIF or QCIF), k is related to the codec (i.e. k=1 for MPEG-2, and k is a function of ab for H.264) and m and n are the other model parameters. In this paper, it is shown that the model parameters m and n are related to the subjective movement content of the clip. Video clips are classified in three classes, according to the subjective movement content: High, Medium and Low movement content. Using the relation between MOS and DMOS from Equation (1), I_c can be defined as described in Equation (7). We will call this the "*m-n* Model" for reference in this paper.

$$I_c = 4\left(1 - \frac{m}{k(ab)^n}\right) \qquad (7)$$

Another model can be found in Recommendation ITU-T G.1070 [18], which describes a computational model for point-to-point interactive videophone applications over IP networks. This Recommendation is based on the work performed by NTT (Nippon Telegraph and Telephone) Service Integration Laboratories [19][20]. The model has 3 parameters, which depend on codec type, video display format, key frame interval, and video display size. Provisional values are provided only for MPEG-4 in small display formats (QVGA and QQVGA). In [21], the same model presented in Recommendation ITU-T G.1070 is proposed for video quality estimation in IPTV services, in HD (High Definition, 1440 x 1080 pixels) display format. In this case, parameters values are provided for the H.264 codec.

Enhancements to the Recommendation G.1070 were proposed in [22], with the model presented in Equation (8)

$$I_c = 4k\left(1 - \frac{1}{1 + \left(\frac{ab}{v_4}\right)^{v_5}}\right) \qquad (8)$$

where I_c represents the video quality determined by the codec distortion, b is the bit rate, a is a constant that depend on the display format and v_4 and v_5 are the other model parameters. The coefficient k is related to the codec, with k=1 for MPEG-2 and k is a function of ab for H.264. We will call this the "Enhanced G.1070 Model" for reference in the present paper. In this model ([22]), it is shown that the relation between I_c and the bit rate not only depends on the codec used and the display format, but strongly depends on video content, specially for low bit rates. In this paper, using similar ideas presented in the "*m-n* Model" in [15], video clips are classified in three classes, according to the subjective movement content, and the model parameters v_4 and v_5 are calculated for each class (High, Medium and Low movement content). The parameters a and k

do not depend on movement content according to the mentioned paper.

V. Proposed Perceptual Video Quality Model

As has been shown in section 4, different models for the relation between the perceived video quality and the bit rate have been proposed, each one applicable to some specific codec, application or display format. Nevertheless, same or equivalent models have been presented, in different works, for different multimedia applications. This is the case of the "Exponential Model", applied for MPEG-2 and H.264 with high definition and large screens (IPTV applications in SD and HD display formats) in [13] and for MPEG-4 with low definition and small screens (3G/4G applications in CIF and QCIF display formats) in [15]. This is also de case for the "G.1070 Model", applied to video telephony applications with small display formats (MPEG-4, QVGA and QQVGA) in [18] and for IPTV (H.264, HD) in [21]. On the other hand, the "*m-n* Model" and the "Enhanced G.1070 Model" are applicable to different kind of applications, from small to large display formats, and for different codecs. In these two models, and also in the "Exponential Model", it is shown that some of the model parameters depend on video content, but none of them presents a direct way to objectively derive the model parameters from the video content. In the rest of the current section, a comparison between the different models is presented, and for the selected model, a direct relation between the model parameters and video content is proposed.

The best parameters values for the "Exponential Model", for the "Enhanced G.1070 Model" and for the "*m-n* Model", as well as the RMSE (Root Mean Square Error) were calculated for each curve in Figure 1. The maximum RMSE values were found for the video clip "Rugby" (src9), which has very high movement content, and are presented in Table 3, with the corresponding parameter values for each model. In Figure 2 the perceived video quality derived from one of the ITU-T standardized models (NTIA), and the estimated using "Exponential Model", "Enhanced G.1070 Model" and "*m-n* Model", with the values presented in Table 3, is showed for this clip.

The three models fit well to the actual values, but the "Enhanced G.1070 Model" has some advantages regarding the other models. First, the "Enhanced G.1070 model" has lower RMSE values than the other two models for all the clips. Second, for lower bit rates, when $b \rightarrow 0$, in the "Exponential Model" $V_q \rightarrow (1+a_3-a_1)$, in the "*m-n* Model" $V_q \rightarrow -\infty$, while in the "Enhanced G.1070 Model" $V_q \rightarrow 1$ (for any parameters value). The minimum MOS value is, by definition, 1, as derived from the "Enhanced G.1070 Model" for any parameters values. For these reasons, we have selected a model based on the "Enhanced G.1070 Model" in this work, in order to derive the model parameters from video content. The proposed model is presented in Equation (9), leaving the parameters name provided in the "Enhanced G.1070 Model".

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

TABLE III. Models Comparision

Model	Parameters	RMSE
Enhanced G.1070	v_4=1.24 v_5=1.6 a=1 k=1	0.65
Exponential Model	a_1=4.50 a_2=0.77 a_3=3.75	0.79
m-n Model	m=0.56 n=0.99 a=1 k=1	0.77

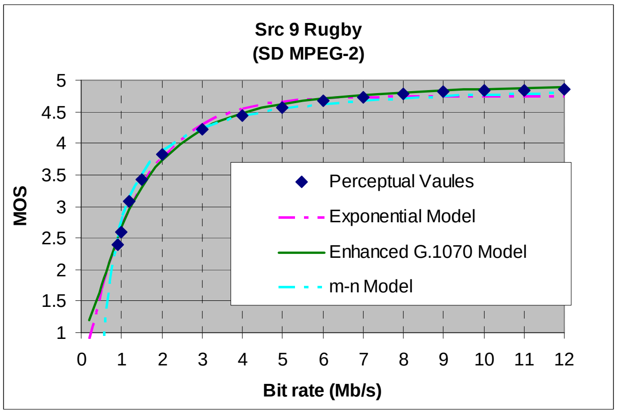

Fig. 2. Perceived quality for the clip Rugby (src 9), coded in SD in MPEG-2, and values derived from different models

As has been shown in [14] and [15] ("Exponential Model"), [17] ("*m-n* Model") and [22] ("Enhanced G.1070 Model"), there is a strong relation between MOS and video content, for a given codec, display format and bit rate. This can be confirmed looking at Figure 1. In MPEG-2, SD display format, for less than 3 Mb/s there are high variations in MOS values for the same bit rate. 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"). Similar behaviors can be found for different codecs and for different display formats. For this reason video content must be taken into account in order to estimate the perceived MOS, for a given codec, display format, bit rate and frame rate.

In [14], for the "Exponential Model", only a qualitative analysis is presented, regarding the relation between video content and perceived quality. In [15], ("Exponential Model"), it is shown that the model parameters can be derived from only one parameter, related to the spatial-temporal activity of the clip. But this paper did not use a standardized video quality metric, and did not present how to directly derive this parameter from the video clip. An indirect method is presented, based on the quality evaluation of the clip coded with a high bit rate.

In [17] ("*m-n* Model") and [22] ("Enhanced G.1070 Model") video clips are classified according to the subjective movement content into three classes (High, Medium and Low movement content), but it is not described how to derive the video clip movement content based on objective parameters.

Using the "Enhanced G.1070 Model", the values of a, v_4 and v_5 that best fits Equation (9) to the perceived quality for all the clips coded in MPEG-2 and H.264/AVC, in SD, VGA, CIF and QCIF display formats were calculated. Different estimations for the video spatial-temporal activity were evaluated, founding a strong correlation between the v_4 and v_5 parameters with the average SAD (Sum of Absolute Differences). 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.

The relations between the v_4 and v_5 parameters with the average SAD per pixel are depicted in Figure 3 and in Figure 4. In Figure 3, the subjective movement content is graphically showed, confirming that low values for SAD/pixel are related to low spatial-temporal activity and high values are related to high spatial-temporal activity or movement content. An estimation of v_4 and v_5 for MPEG-2 and H.264, as a function of the average SAD/pixel can be performed as

$$v_4 = c_1 s^{c_2} + c_3$$
$$v_5 = c_4 s^{c_5} + c_6 \qquad (10)$$

where s is the video average SAD/pixel. The best values for $c_1..c_6$ and for a are presented in Table 4 and Table 5.

With these relations, the video quality estimation presented in Equation (9) only depends on the encoded bit rate and the spatial-temporal activity of the video clip, measured as the average SAD/pixel, calculated for each frame in 8x8 blocks and averaging for all the frames in the clips (for clip duration of the order of 10 seconds, as the ones used in this work).

TABLE IV. c_i VALUES FOR EACH CODEC

Codec	c_1	c_2	c_3	c_4	c_5	c_6
MPEG-2	0.208	0.95	0.036	0.036	1.52	1.17
H.264/AVC	0.150	0.95	0	0.030	0.68	1.20

TABLE V. a VALUES FOR EACH DISPLAY FORMAT

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

The dispersion between the MOS values derived using Equations (9) and (10) and the perceived MOS values (using the NTIA VQM standardized in Recommendation ITU-T J.144), for the sixteen video clips used, coded in MPEG-2 and H.264, in SD, VGA, CIF and QCIF display format, with bit rates from 25 kb/s to 12 Mb/s are plotted in Figure 5. In this figure, each point represents a video clip coded in a specific combination of codec, bit rate and display format. It is worth noting that subjective rating scales have ranges of 1 unit, in the 1-5 MOS scale. On the other hand, the NTIA algorithm standardized by the ITU has errors in the order of +/- 0.4 regarding to MOS measures of subjective quality. In Figure 5, the dotted lines represent the estimated +/- 0.4 error margin of the NTIA model. Only 27 from the 2064 points are outside the dotted lines, meaning that the predicted MOS values (using the proposed model) have the same degree of precision than the video quality metric used.

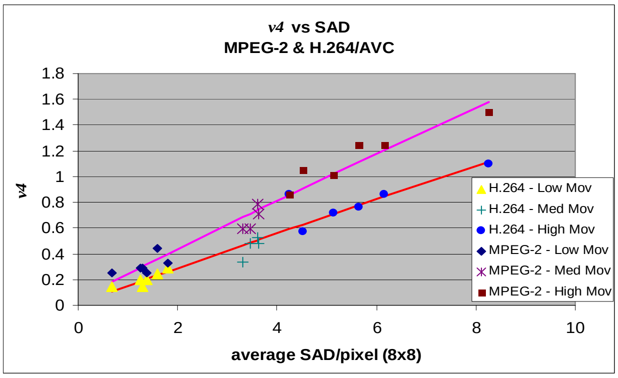

Fig. 3. Relation between v_4 with respect to SAD

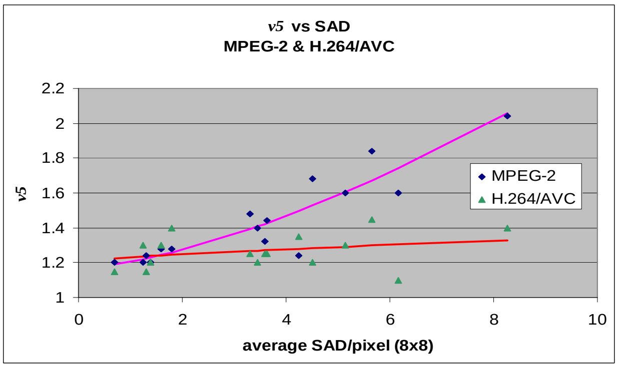

Fig. 4. Relation between v_5 with respect to SAD

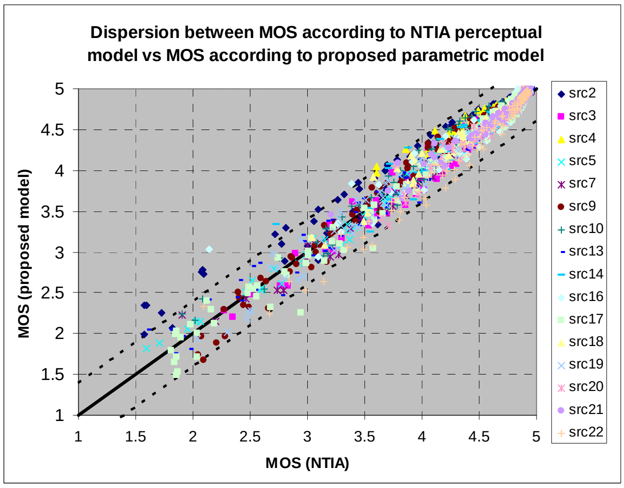

Fig. 5. MOS dispersion in proposed model

VI. CONCLUSION

A parametric model for perceptual video quality estimation was proposed, which provides a very good estimation to the perceptual MOS values, for different codecs, bit rates and display formats, knowing the spatial-temporal activity content in the video application. This spatial-temporal activity 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. It has been shown that low values for SAD/pixel are related to low spatial-temporal activity, while high values are related to high spatial-temporal activity or movement content.

To derive the proposed model, different parametric models already proposed were studied and compared. The proposed model uses only 2 parameters with a clear interpretation. One is the average SAD/pixel (s), and is related to the spatial-temporal activity, with different relations for each codec. The other parameter (a) is related to the display format, and is independent from the codec used. The model can be easily extended to different display formats and codecs.

Instead of doing subjective tests, one of the video quality metrics (VQM) standardized in Recommendations ITU-T J.144 and ITU-R BT.1683 was used. The proposed model has been evaluated, comparing the quality estimation derived from the proposed parametric model with the standard VQM, for MPEG-2 and H.264, in SD, VGA, CIF and QCIF display formats and in the bit rate ranges from 25 kb/s to 12 Mb/s. Sixteen video sources were used, coded in more than 120 different formats, varying the codec, the bit rate and the display format. In total more than 2000 processed video sequences were analyzed. The parameters values for the model have been calculated and are presented in the paper. The 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.

REFERENCES

[1] Recommendation ITU-R BT.500-11, Methodology for the subjective assessment of the quality of television pictures, 06/2002

[2] Recommendation ITU-T P.910. Subjective video quality assessment methods for multimedia applications,09/1999

[3] S Winkler, P Mohandas: The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics, IEEE Trans on Broadcasting, Vol. 54, No. 3, September 2008.

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

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

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

[7] Test Plan for Evaluation of Video Quality Models for Use with High Definition TV Content, Draft Version 3.0, 2009

[8] Hybrid Perceptual/Bitstream Group TEST PLAN, Draft Ver. 1.3, Jan 4, 2009

[9] 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

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

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

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

[13] Raake, A. Garcia, M.N. Moller, S. Berger, J. Kling, F. List, P. Johann, J. Heidemann,: T-V-Model: Parameter-Based Prediction of IPTV Quality. IEEE ICASSP 2008

[14] Garcia, M.N., Raake, A., List, P.: Towards Content-Related Features for Parametric Video Quality Prediction of IPTV Services, Acoustics, Speech and Signal Processing, 2008. ICASSP 2008, pp. 757-760

[15] H. Koumaras, A. Kourtis, D. Martakos and J. Lauterjung: Quantified PQoS assessment based on fast estimation of the spatial and temporal activity level, Multimedia Tools and Applications, Springer Netherlands, Volume 34, Number 3, Sep 2007

[16] J. Lauterjung: Picture Quality Measurement, Proc of the International Broadcasting Convention (IBC), Amsterdam, 1998, pp. 413-417

[17] 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

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

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

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

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

[22] 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