ON THE BENEFIT OF PARAMETER-DRIVEN APPROACHES FOR THE MODELING AND THE PREDICTION OF SATISFIED USER RATIO FOR COMPRESSED VIDEO

Jingwen Zhu, Patrick Le Callet, Anne-Flore Perrin, Sriram Sethuraman, Kumar Rahul

▶ **To cite this version:**

Jingwen Zhu, Patrick Le Callet, Anne-Flore Perrin, Sriram Sethuraman, Kumar Rahul. ON THE BENEFIT OF PARAMETER-DRIVEN APPROACHES FOR THE MODELING AND THE PREDICTION OF SATISFIED USER RATIO FOR COMPRESSED VIDEO. IEEE International Conference on Image Processing, Sep 2022, Bordeaux, France. hal-03788606

HAL Id: hal-03788606
https://hal.science/hal-03788606

Submitted on 26 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON THE BENEFIT OF PARAMETER-DRIVEN APPROACHES FOR THE MODELING AND THE PREDICTION OF SATISFIED USER RATIO FOR COMPRESSED VIDEO

Jingwen Zhu[1], Patrick Le Callet[1], Anne-Flore Perrin[2], Sriram Sethuraman[3], Kumar Rahul[3]

[1]Nantes Université, Ecole Centrale Nantes, CNRS, LS2N, UMR 6004, Nantes, France
[2]CAPACITÉS SAS, Nantes, France [3]Amazon Prime Video, Bangalore, India

ABSTRACT

The human eye cannot perceive small pixel changes in images or videos until a certain threshold of distortion. In the context of video compression, Just Noticeable Difference (JND) is the smallest distortion level from which the human eye can perceive the difference between reference video and the distorted/compressed one. Satisfied-User-Ratio (SUR) curve is the complementary cumulative distribution function of the individual JNDs of a viewer group. However, most of the previous works predict each point in SUR curve by using features both from source video and from compressed videos with assumption that the group-based JND annotations follow Gaussian distribution, which is neither practical nor accurate. In this work, we firstly compared various common functions for SUR curve modeling. Afterwards, we proposed a novel parameter-driven method to predict the video-wise SUR from video features. Besides, we compared the prediction results of source-only features based (SRC-based) models and source plus compressed videos features (SRC+PVS-based) models.

Index Terms— Video Quality Assessment, Just Noticeable Difference, Satisfied User Ratio

1. INTRODUCTION

With the rapid growth of multimedia demand, Picture Wise Just Noticeable Difference (PW-JND) [1, 2, 3, 4] and Video Wise Just Noticeable Difference (VW-JND) [5, 6, 7, 8] have been investigated from human perceptive aspects in order to provide high quality of experience for end-users with limited storage and internet bandwidth. Furthermore, JND estimation is helpful for visual processing tasks such as visual signal enhancement and perceptual quality evaluation.

JND depends on 3 factors : (1) display setting, *e.g.,* viewing distance [9], monitor profiling, *etc.*; (2) subjects; (3) image/video contents [10]. In this work, we concentrated on how different video contents impact the location of JND in the context of video compression. Factor (1) and (2) are fixed and are not investigated.

For a given visual content, JND of different subjects will be different [10]. Wang *et al.*[11] proposed to conduct the subjective test of JND with respect to a viewer group other than the very few experts (golden eyes) for the worst-case analysis, because the group-based quality of experience (QoE) is closer to the realistic applications. Satisfied-User-Ratio (SUR) curve can be derived from this group-based JND value. SUR curve is defined as the Q-function supposing that the group-based JND follows Gaussian distribution [5]. Intuitively, the value of SUR curve at a certain distortion level d, is the percentage of the group users who cannot perceive any difference between the reference stimuli and the distorted stimulus whose distortion level is smaller than d, *i.e.,* these users are satisfied. At a given threshold p for SUR, the corresponding distortion level is defined as p%SUR instead of the misleading notation p%JND in previous works [5, 6, 7, 12].

It is well known that subjective test is expensive and time-consuming, especially for VW-JND. Therefore, it is crucial to develop VW-JND prediction methods in order to predict the encoding parameter (*e.g,* Quantization Parameter (QP)) corresponding to a given SUR threshold. The issue has been addressed in [5], who proposed a model to predict SUR curve by using support vector regression (SVR) under the assumption that the individual JND points of a group users follow a normal distribution. Their model infers SUR values from VMAF [13] Quality Degradation Features concatenated with Masking Effect Features [14, 15] and is trained on the large scale JND dataset of compressed video [12]. The 75%SUR point can then be derived by the predicted SUR curve. [6] is the extended work of [5], the *2nd* and *3rd* JND points are predicted using 3 different settings in which the reference inputs of the predictor are different. Instead of predicting encoding parameter QP as SUR profile, Zhang *et al.* [7] proposed a novel perceptual model to predict SUR versus bitrate, which is more widely used in practice. Three kinds of features, Masking features, re-compression features and basic attribute features, are extracted from original reference video to build a feature vector, which will be used to conduct a Gaussian Processes Regression (GPR) to predict SUR.

However, previous works assume that the individual VW-JND of a group viewers follows Gaussian distribution, which might not be the optimal modeling method. Besides, when predicting SUR curve, previous works are computationally expensive because they extract features from SRC and from every encoded PVS and predict the individual SUR score of

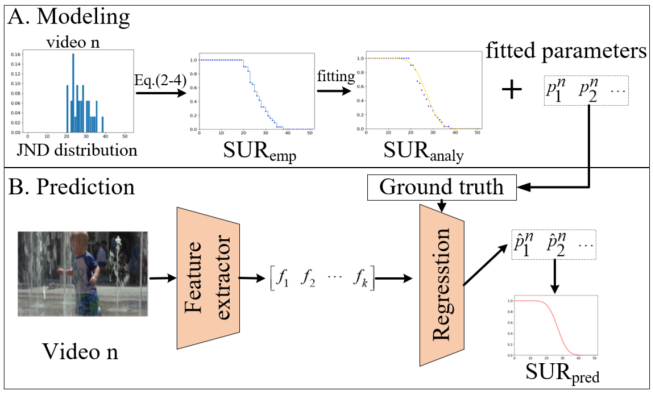

Fig. 1: Illustration of the pipeline of SUR and JND modeling (A) and prediction (B)

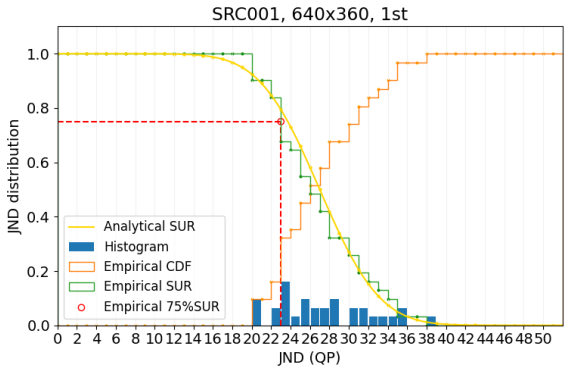

Fig. 2: Distribution of group-based VW-JND (blue bar); empirical and analytical SUR (in green and yellow respectively) and empirical 75%SUR (red circle) of SRC001 (360p) in VideoSet [12]

each PVS to derive the SUR curve. Therefore, we first investigate the modeling of SUR curve and then propose a novel SUR prediction model only based on SRC.

2. MODELING & PREDICTION OF SUR

In this work, we firstly investigated the modeling of the group-based VW-JND rather than using a simple normality test as in the previous works [12, 11], in order to find the mathematical model that best fits SUR. Afterwards, we proposed a SUR prediction method via predicting the model parameters obtained by the modeling fit, which is called parameter-driven model, in preference to the commonly used point-by-point models. The entire pipeline is shown in Fig.1

2.1. Modeling of SUR

When modeling the SUR of VW-JND, previous works [12, 11] conducted Jarque-Beta test [16] to verify the normality of the VW-JND position of every subject. However, when revisiting the original distribution of VW-JND annotations, we found that Gaussian distribution is not necessarily the best modeling of the VW-JND distribution (Fig.2). Therefore, we first clarify the definition of empirical SUR curve and p%SUR. Afterwards, we set the Complementary Cumulative Distribution Function (CCDF) of different candidate distributions (*e.g.,* Gaussian, Sigmoid, Weilbull, *etc.*) as model functions to find the best fit function of the empirical SUR curve. The model functions were named as "analytical SUR". For a given content m, assuming that there are N reliable subjects' VW-JND annotation, VW-JND of a subject "n" is denoted by j_n^m. VW-JND of N subjects can be denoted by J^m as

$$J^m = [j_1^m, j_2^m, ..., j_N^m]\,. \tag{1}$$

Considering J^m as a discrete random variable, the Probability Mass Function (PMF) of J^m is defined by

$$p(x) = P(\text{JND} = \text{x}) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\,(j_i = x), \tag{2}$$

where $\mathbf{1}(c)$ is an indicator function that equals to 1 if the specified binary clause c is true. Thus, the empirical Cumulative Distribution Function (CDF) can be calculated from the PMF:

$$\text{CDF}_{\text{emp}}(x) = P(\text{JND} \leq \text{x}) = \sum_{\omega<\text{x}} \text{p}(\omega). \tag{3}$$

Because J is discrete, the CDF increases only by jump discontinuities, as shown in Fig.2 (orange curve). The empirical SUR (green curve in Fig.2) is the CCDF of J:

$$\text{SUR}_{\text{emp}}(x) = 1 - \text{CDF}_{\text{emp}}(\text{x}). \tag{4}$$

The empirical p%SUR is defined as:

$$p\%\text{SUR}_{\text{emp}} = \min\left\{x \,|\text{SUR}_{\text{emp}}\,(x) \leq p\%\right\}, \tag{5}$$

where the range of x is the range of distortion levels. For instance, the range of x is [0, 51] with step equals to 1 in VideoSet. The analytical SUR curve and p%SUR are calculated by Eq.(6) and (7), $f(x)$ is the Probability Density Function (PDF). Contrary to the empirical SUR, the analytical SUR is a continuous function.

$$\text{SUR}_{\text{analy}}\,(x) = 1 - \text{CDF}_{\text{analy}}(x) = 1 - \int_{-\infty}^{\infty} f(x)dx. \tag{6}$$

$$p\%\text{SUR}_{\text{analy}} = \underset{x\in\{0,1,...51\}}{\arg\min}\ (\text{SUR}_{\text{analy}}(x) - p\%) \tag{7}$$

After computing the empirical SUR from the VW-JND annotations, we fitted the discrete points in empirical SUR with 8 model functions (Table1) for each video. It can be easily proved that the CDF (Eq.(3)) of a distribution is monotonic non-decreasing, thus SUR (Eq.(4)) is monotonic non-increasing. Therefore, monotonic constraint was applied during least-squares optimization for polynomial model function. In addition to MAE and RMSE, we use $\Delta p\%\text{SUR}_{|\text{E}-\text{A}|}$ (Eq.(8)) to evaluate the candidate model functions, where $p\%$ is set to 75%. The experiment results will be detailed in Section 3.1.

$$\Delta p\%\text{SUR}_{|\text{E}-\text{A}|} = |p\%\text{SUR}_{\text{emp}} - p\%\text{SUR}_{\text{analy}}| \tag{8}$$

Table 1: Summary of candidate model functions. (NB para is the number of parameters in model function)

Name	Model function	NB para
Polynomial-3	$f(x) = \sum_{k=0}^{n} a_k x^k$	4
Polynomial-4		5
Gaussian	$1 - \frac{1}{2}\left(1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right)$	2
2-para-logistic	$1 - \frac{1}{1+e^{-(x-\mu)/s}}$	2
4-para-logistic	$f(x) = b + \frac{L}{1+e^{-k(x-x_0)}}$	4
Weibull	$e^{-\left(\frac{x}{\lambda}\right)^k}$	2
Gumbel	$1 - e^{-e^{-(x-\mu)/\beta}}$	2
Rayleigh	$e^{\frac{-x^2}{(2\sigma^2)}}$	1

2.2. Prediction of SUR

We revisited the SUR prediction model proposed by Wang *et al.* [5] namely the baseline. Furthermore, we analysed the two main drawbacks of the baseline model and proposed solutions to solve these issues.

As shown in Fig.(3), the input of the baseline model is the uncompressed source video (SRC). SRC is firstly compressed with different encoding parameters (*e.g.*,QP 1 to 51) to get a series of PVS (Processed Video Sequence). Afterwards, SRC and all PVS are segmented to small video patches both spatially and temporally to extract features from the eye fixation level. Two types of features are extracted from the segmented patches: Masking effect and Quality degradation (M and Q in Fig.(3)). Masking effect is a measure of the spatial and temporal randomness [14, 15]. A high randomness masks distortions for human eye, *i.e.*, it indicates the human eye hardly perceive the difference. Quality degradation is calculated based on the difference of quality scores (*e.g.*, VMAF) between SRC and PVS. The masking effect and quality degradation feature vectors of one video are the histogram and cumulative curve of its video patches, respectively. When extracting the Quality degradation features, only video patches with significant quality degradation were selected to compute the final feature vector. The two feature vectors for SRC and each PVS are concatenated and are used to predict SUR scores by regression.

The baseline model is computational expensive as individual SUR score of each PVS of a SRC must be computed to derive the SUR curve prediction. Accordingly, features from a SRC and its PVSs (*i.e.,* 52 sequences) have to be computed. This is the first drawback of the baseline model. The second drawback is that the $\mathrm{SUR}_{\mathrm{pred}}$ curve of baseline model is not monotonic non-increasing because every individual point of SUR curve is predicted separately. The basic meaning of SUR is: for a given distortion level x, its SUR value is what percentage of subjects are satisfied, *i.e.,* what percentage of subjects do not perceive the difference between the reference video and all the distorted video whose distortion level is less than x. Therefore, it is not reasonable that when the distortion level increases, the SUR value increases.

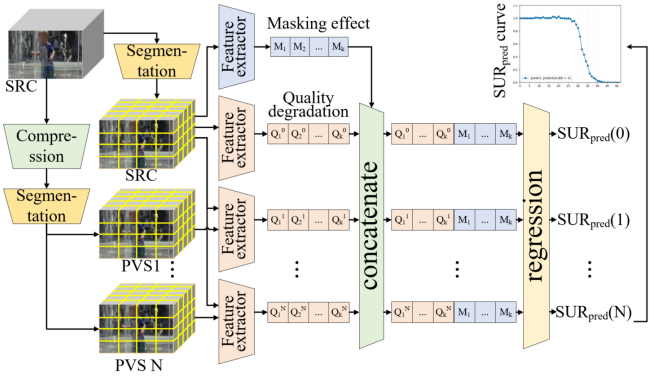

Fig. 3: Illustration of baseline SUR prediction model

To address these issues, we proposed a straightforward solution with the help of the modeling parameters of SUR in Section 2.1. The pipeline of the proposed method is shown in Fig.1. Instead of predicting SUR scores of every PVS and geting the SUR curve accordingly, the parameters which describe the SUR curve (e.g., σ and μ for Gaussian) are predicted. Only SRC is used for prediction, hence these models are called **SRC-based** model. Masking effect features are extracted from SRC and Support Vector Regression (SVR) [17] is used for regression. Although SRC-based models are preferred in real-life applications, it is still interesting to understand how important is the quality degradation information from PVSs to the prediction of SUR and JND. Therefore, we also investigated **SRC+PVS-based** model, where masking effect and quality degradation features of every PVS are concatenated into one vector for regression to predict the modeling parameters.

3. EXPERIMENTS AND RESULTS

3.1. Modeling

In our experiment, VideoSet [12] was used to evaluate the modeling and prediction of SUR and JND. VideoSet include 220 5-second SRCs in 4 resolutions. Each SRC is encoded with H.264 codec with QP from 1 to 51. More than 30 subjects participated into the viewer group for each SRC and each individual VW-JND annotation was publicly available as well. As mentioned in Section 2.1, we used the individual VW-JND annotations of each SRC to generate the $\mathrm{SUR}_{\mathrm{emp}}$ and $75\%\mathrm{SUR}_{\mathrm{emp}}$ (Eq.(2)-(5)) and every discrete points of $\mathrm{SUR}_{\mathrm{emp}}$ were used to fit the model functions listed in Table 1. Scikit-learn linear regression was used for polynomial fittings and monotonic constraints were applied by using Polyfit[1]. Non-linear least squares from SciPy were used for other model functions.

For every SRC in 4 resolutions of VideoSet ($220 \times 4 = 880$ SRC in total), we calculated the MAE and RMSE be-

[1]https://github.com/dschmitz89/Polyfit

Table 2: Mean of MAE, RMSE and $\Delta 75\%\text{SUR}_{|\text{E}-\text{A}|}$ for different model functions with VideoSet [12]

Name	MAE	RMSE	$\Delta 75\%\text{SUR}_{\|\text{E}-\text{A}\|}$
Polynomial-3	0.1204	0.1466	5.0614
Polynomial-4	0.1085	0.1338	4.7420
Gaussian	0.0147	0.0253	0.6625
2-para-logistic	0.0156	0.0250	0.5875
4-para-logistic	0.0164	**0.0236**	**0.5761**
Weibull	**0.0138**	0.0240	0.6761
Gumbel	0.0220	0.0343	0.5977
Rayleigh	0.1451	0.1703	8.9114

tween SUR_{emp} and $\text{SUR}_{\text{analy}}$ and also the difference between empirical and analytical 75%SUR : $\Delta 75\%\text{SUR}_{|\text{E}-\text{A}|}$ (Eq.(8)) with different fitting functions. The results are shown in Table 2. It can be observed that the CCDF of Gaussian distribution is not the best modeling for SUR. 4-para-logistic model function outperforms the other candidate model functions both in RMSE and $\Delta 75\%\text{SUR}_{|\text{E}-\text{A}|}$.

3.2. Prediction

Each prediction model was evaluated on videos using 5-fold cross validation. Radial basis function kernel was used for SVR. We firstly compare the baseline model (in-house implementation) and 3 SRC-based parameter-driven models with different model functions. Fig.4 shows the SUR prediction results of 2 SRCs. The dashed lines in orange, green and blue are the analytical SUR curves obtained from fitting to the red points in empirical SUR with Gaussian, 2-p-logistic and 4-p-logistic respectively. Plain lines with dots are the predicted SUR of the 3 SRC-based models obtained by predicting the parameters of the corresponding dashed lines. The purple line is the SUR prediction of baseline model. Difference between **P**redicted and **A**nalytical SUR (denoted $\Delta\text{SUR}_{|\text{P}-\text{A}|}$) is the MAE between them. $\Delta\text{SUR}_{|\text{P}-\text{A}|}$ indicates the error between ground truth (analytical SUR curve) and prediction SUR curve, but the modeling error (between empirical and analytical SUR curve) are not considered. Therefore, difference between **P**redicted and **E**mpirical SUR (denoted $\Delta\text{SUR}_{|\text{P}-\text{E}|}$) is evaluated as well. $\Delta 75\%\text{SUR}$ is evaluated in the same way. The results are shown in Table 3. The 3 SRC-based parameter-driven models outperform the baseline model both in ΔSUR and $\Delta 75\%\text{SUR}$. The prediction errors between Gaussian and logistic parameter-driven models are quite close. However, the 4-para-logistic which has the smallest modeling error performs worse than Gaussian and 2-para-logistic.

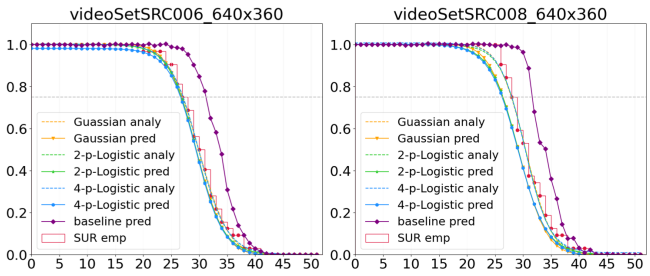

Fig. 4: Examples of SUR prediction results comparison between SRC-based models and baseline model

Table 3: Averaged prediction error comparison between baseline model and 3 SRC-based parameter-driven models.

RES	Model name	ΔSUR		$\Delta 75\%\text{SUR}$	
		$\|\text{P}-\text{A}\|$	$\|\text{P}-\text{E}\|$	$\|\text{P}-\text{A}\|$	$\|\text{P}-\text{E}\|$
360p	baseline	0.0769	0.0799	4.3682	4.3773
	2-p-Gaussian	**0.0459**	**0.0480**	2.4773	2.5864
	2-p-Logistic	0.0462	0.0489	**2.4455**	**2.5682**
	4-p-Logistic	0.0496	0.0515	2.4591	2.5909
540p	baseline	0.0786	0.0812	4.3182	4.2909
	2-p-Gaussian	**0.0397**	**0.0428**	2.1182	2.1045
	2-p-Logistic	0.0398	0.0437	**1.9727**	**2.0955**
	4-p-Logistic	0.0435	0.0458	2.0045	2.1000
720p	baseline	0.0783	0.0820	4.2864	4.2909
	2-p-Gaussian	**0.0433**	**0.0447**	**2.1636**	**2.2045**
	2-p-Logistic	0.0435	0.0459	**2.1636**	2.2364
	4-p-Logistic	0.0467	0.0476	**2.1636**	2.2318
1080p	baseline	0.0801	0.0834	4.6000	4.5591
	2-p-Gaussian	0.0412	**0.0431**	2.3455	2.2136
	2-p-Logistic	**0.0409**	0.0440	**2.1182**	**2.1773**
	4-p-Logistic	0.0439	0.0455	2.1455	2.1727

We also compared SRC-based model and SRC+PVS-based model, the results are shown in Table 4. It can be observed that adding quality degradation information from PVSs will improve the prediction of both SUR and 75%SUR, but with a cost of encoding SRC to 51 PVSs.

Table 4: Averaged prediction error comparison between SRC-based and SRC+PVS based model on 1080p with Guassian modeling.

Model	ΔSUR		$\Delta 75\%\text{SUR}$	
	$\|\text{P}-\text{A}\|$	$\|\text{P}-\text{E}\|$	$\|\text{P}-\text{A}\|$	$\|\text{P}-\text{E}\|$
SRC-based	0.0412	0.0431	2.3455	2.2136
SRC+PVS-based	**0.0377**	**0.0412**	**2.0727**	**2.1409**

4. CONCLUSION

In this paper, we proposed a novel pipeline for SUR modeling and prediction to predict the optimal encoding parameter only from SRC. Experiment results show that the proposed parameter-driven model (2-p-Logistic for instance) improves the mean SUR prediction error to 0.046, reducing it by 43.64% compared with the baseline and reduces the mean 75%SUR prediction error from 4.38 QP (baseline) to 2.27 QP. Furthermore, compared with SRC-based model, the SRC+PVS-based model slightly improves the mean prediction error of SUR curve and 75%SUR by 0.0019 and 0.0727 QP respectively, which means the quality degradation features from PVSs are not crucial to SUR prediction.

5. REFERENCES

[1] Huanhua Liu, Yun Zhang, Huan Zhang, Chunling Fan, Sam Kwong, C-C Jay Kuo, and Xiaoping Fan, "Deep learning-based picture-wise just noticeable distortion prediction model for image compression," *IEEE Transactions on Image Processing*, vol. 29, pp. 641–656, 2019.

[2] Tao Tian, Hanli Wang, Lingxuan Zuo, C-C Jay Kuo, and Sam Kwong, "Just noticeable difference level prediction for perceptual image compression," *IEEE Transactions on Broadcasting*, vol. 66, no. 3, pp. 690–700, 2020.

[3] Xuelin Shen, Zhangkai Ni, Wenhan Yang, Xinfeng Zhang, Shiqi Wang, and Sam Kwong, "Just noticeable distortion profile inference: A patch-level structural visibility learning approach," *IEEE Transactions on Image Processing*, vol. 30, pp. 26–38, 2020.

[4] Hanhe Lin, Mohsen Jenadeleh, Guangan Chen, Ulf-Dietrich Reips, Raouf Hamzaoui, and Dietmar Saupe, "Subjective assessment of global picture-wise just noticeable difference," in *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2020, pp. 1–6.

[5] Haiqiang Wang, Ioannis Katsavounidis, Qin Huang, Xin Zhou, and C-C Jay Kuo, "Prediction of satisfied user ratio for compressed video," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 6747–6751.

[6] Haiqiang Wang, Xinfeng Zhang, Chao Yang, and C-C Jay Kuo, "Analysis and prediction of jnd-based video quality model," in *2018 Picture Coding Symposium (PCS)*. IEEE, 2018, pp. 278–282.

[7] Xinfeng Zhang, Chao Yang, Haiqiang Wang, Wei Xu, and C-C Jay Kuo, "Satisfied-user-ratio modeling for compressed video," *IEEE Transactions on Image Processing*, vol. 29, pp. 3777–3789, 2020.

[8] Sehwan Ki, Sung-Ho Bae, Munchurl Kim, and Hyunsuk Ko, "Learning-based just-noticeable-quantization-distortion modeling for perceptual video coding," *IEEE Transactions on Image Processing*, vol. 27, no. 7, pp. 3178–3193, 2018.

[9] E Nakasu, K Aoki, R Yajima, Y Kanatsugu, and K Kubota, "A statistical analysis of mpeg-2 picture quality for television broadcasting," *SMPTE journal*, vol. 105, no. 11, pp. 702–711, 1996.

[10] Joe Yuchieh Lin, Lina Jin, Sudeng Hu, Ioannis Katsavounidis, Zhi Li, Anne Aaron, and C-C Jay Kuo, "Experimental design and analysis of jnd test on coded image/video," in *Applications of digital image processing XXXVIII*. International Society for Optics and Photonics, 2015, vol. 9599, p. 95990Z.

[11] Haiqiang Wang, Weihao Gan, Sudeng Hu, Joe Yuchieh Lin, Lina Jin, Longguang Song, Ping Wang, Ioannis Katsavounidis, Anne Aaron, and C-C Jay Kuo, "Mcl-jcv: a jnd-based h. 264/avc video quality assessment dataset," in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2016, pp. 1509–1513.

[12] Haiqiang Wang, Ioannis Katsavounidis, Jiantong Zhou, Jeonghoon Park, Shawmin Lei, Xin Zhou, Man-On Pun, Xin Jin, Ronggang Wang, Xu Wang, et al., "Videoset: A large-scale compressed video quality dataset based on jnd measurement," *Journal of Visual Communication and Image Representation*, vol. 46, pp. 292–302, 2017.

[13] Zhi Li, Anne Aaron, Ioannis Katsavounidis, Anush Moorthy, and Megha Manohara, "Toward a practical perceptual video quality metric," *The Netflix Tech Blog*, vol. 6, no. 2, 2016.

[14] Sudeng Hu, Lina Jin, Hanli Wang, Yun Zhang, Sam Kwong, and C-C Jay Kuo, "Compressed image quality metric based on perceptually weighted distortion," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5594–5608, 2015.

[15] Sudeng Hu, Lina Jin, Hanli Wang, Yun Zhang, Sam Kwong, and C-C Jay Kuo, "Objective video quality assessment based on perceptually weighted mean squared error," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 9, pp. 1844–1855, 2016.

[16] Carlos M Jarque and Anil K Bera, "A test for normality of observations and regression residuals," *International Statistical Review/Revue Internationale de Statistique*, pp. 163–172, 1987.

[17] Alex J Smola and Bernhard Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.