Performance Evaluation of a Perceptual Ringing Distortion Metric for Digital Video

Zhizhong Zhe[1], Hong Ren Wu[1,3], Zhenghua Yu[2], Tim Ferguson[1] and Damian Tan[1]

1. School of Computer Science and Software Engineering, Monash University, Victoria 3800, Australia.
    {zhezhiz,hrw,timf,dmt}@csse.monash.edu.au
2. Motorola Australian Research Centre, Sydney, Australia. zhyu@ieee.org
3. School of Computer Engineering, Nanyang Technological University, Singapore.

ABSTRACT

This paper evaluates a perceptual impairment measure for ringing artifacts, which are common in hybrid MC/DPCM/DCT coded video, as a predictor of the mean opinion score (MOS) obtained in the standard subjective assessment. The perceptual ringing artifacts measure is based on a vision model and a ringing distortion region segmentation algorithm, which is converted into a new Perceptual Ringing Distortion Metric (PRDM) on a scale of 0 to 5. This scale corresponds to a modified Double-Stimulus Impairment Scale variant II (DSIS-II) method. The Pearson correlation, the Spearman rank order correlation and the average absolute error are used to evaluate the performance of the PRDM compared with the subjective test data. The results show a strong correlation between the PRDM and the MOS with respect to ringing artifacts.

Keywords: Ringing Artifact, Digital Video Quality, Perceptual Distortion Metric for Digital Video.

I. INTRODUCTION

The quality assessment of the hybrid MC/DPCM/DCT (motion compensated/differential pulse code modulation /block based discrete cosine transform) coded video has become an increasingly important issue in digital video coding and communications [1-5]. The absence of appropriate and commonly acceptable human visual system (HVS) based objective digital video quality measures has led to an inadequate assessment of the visual performance of existing video coding and communications systems.

Digital video coding distortions introduced by using standard hybrid coding algorithms have been well understood and classified [6,7]. In the design of objective quality and impairment metrics, it is highly desirable to identify the nature of occurring distortions and to quantify the quality degradation caused by different types of distortions. This allows a more detailed analysis and tuning of the performance of video quality. The investigations into quantitative metrics for assessment of the different artifacts are essential in video coding performance evaluation, coder design and overall picture quality assessments. Blocking impairment metrics for both still images and digital videos have been investigated and reported [8-11]. Particularly, the PBDM (Perceptual Blocking Distortion Metric) based on a human vision model for digital video is shown to have high correlation with subjective data in terms of the Pearson correlation, the Spearman rank order correlation and the average absolute error [11]. A vision model based ringing artifact measure was proposed in [12] for digital video.

In this contribution, the ringing artifact measure is transformed into a new video Perceptual Ringing Distortion Metric (PRDM) on a scale of 0 to 5 corresponding to a modified Double-Stimulus Impairment Scale variant II (DSIS-II) method. Section II introduces the perceptual ringing distortion metric. Its performance evaluations are presented in Section III, followed by conclusions.

II PERCEPTUAL RINGING DISTORTION METRIC

As shown in Figure 1, the PRDM consists of four stages: spatio-temporal decomposition, ringing region segmentation, contrast gain control, and detection and pooling.

A. Ringing Region Segmentation

To measure the perceived distortion due to ringing, a segmentation algorithm is devised to identify regions with ringing artifacts. Ringing artifacts are fundamentally related to the Gibb's phenomenon, when quantization of individual coefficients results in high-frequency irregularities of the reconstructed block. The mosquito effect could also be treated as the temporal version of ringing artifact, where high frequency fluctuations in areas around high contrast edges will appear due to coarsely quantized higher frequency AC coefficients. The ringing
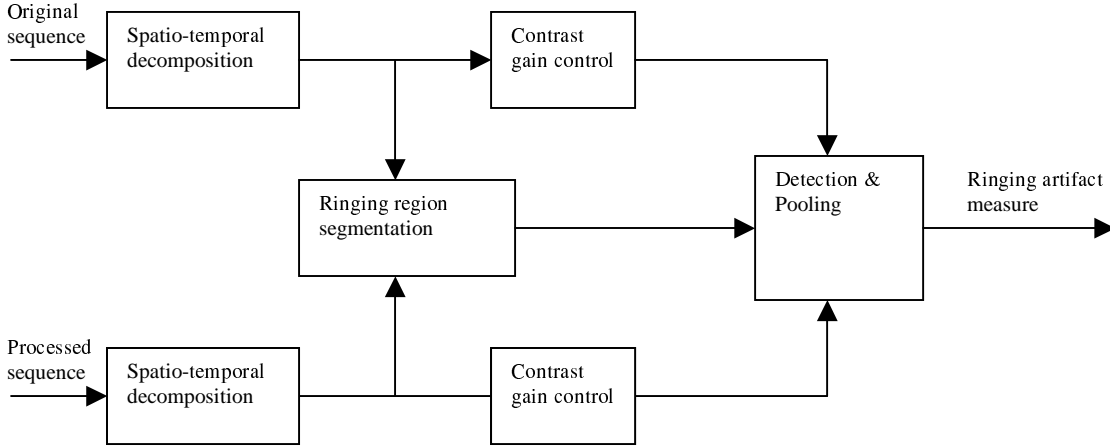
Figure 1. Block diagram of the Perceptual Ringing Distortion Metric.

artifact is most evident along high contrast edges, if either side of the edge is in the area of generally smooth texture [6]. Thus the ringing region map can be generated through detecting the boundaries of the regions with smooth texture and complex texture. In this work, we adopted the ringing region segmentation algorithm designed by Yu et al [12]. At first, the localized variance (of 5x5-pixel block) is calculated for all pixels in the image. Then the variance is low-pass filtered to remove possible noises. If the variance at a pixel location is less than a low threshold, the pixel is classified as in the smooth texture region. If the variance is higher than a high threshold, it is grouped as in the complex region. The complex region includes high contrast edges and complex textures. Other unclassified pixels are further classified into either the smooth texture region or the complex region by a region growing algorithm. At last, the ringing region is detected as the boundaries between the smooth texture region and the complex region.

B. Ringing Impairments Measure

As shown in Figure 1, the original and the processed video sequences are input into the metric and decomposed respectively by the spatio-temporal filterbanks. Since ringing distortions are most evident in the temporal sustained channel [11], only the temporal low-pass channel is processed for further spatial decomposition and distortion summation. This reduces the computational complexity. After the decomposition, ringing dominant regions in the video sequences are segmented and an associated ringing region map is generated. The Teo-Heeger contrast gain control model is implemented for pattern masking in the same stage [13]. At the detection stage, the normalized responses to the two input sequences are subtracted. The pooling stage combines the differences over all the spatial frequency and orientation channels between the original and processed sequences using Minkowski summation to obtain an overall distortion measure and the Just Noticeable Difference (JND) map. In the JND map, each pixel value represents the strength of noticeable distortion at its corresponding spatial location. The JNDs in the ringing regions, as defined by the ringing region map, are summed up to obtain the ringing artifact metric. Currently, the calculation of the PRDM is applied to the luminance component only.

III. EXPERIMENTS AND PERFORMANCE EVALUATION

Both subjective and objective tests have been conducted to evaluate the performance of the PRDM in terms of correlations and prediction errors between subjective and objective data.

A. Subjective Test

Two test sequences, Mobile & Calendar (MC) and Table Tennis (TT), are selected from all VQEG and ANSI video sequences [3] based on their suitability in rendering different degrees of the ringing impairment after the MPEG-2 encoding. The two original sequences are MPEG-2 encoded [14] at ten different bit rates as shown in Table I, to cover the full range of ringing impairments. The two different scene sequences are presented using an interleaved order. The sequences with the same scene and different bit rates are presented in a pseudo-random order.

TABLE I
BIT RATES OF TEST SEQUENCES (Mbps)

1.5	1.75	2.0	2.25	2.5	2.75	3.0	4.0	5.0	7.5

The Double-Stimulus Impairment Scale variant II (DSIS-II) method as defined in ITU-R BT.500 [15] is employed in the subjective test, with a rating system modified from the five-grade scale to a continuous range between zero and five. This continuous scale corresponds to the normal DSIS-II five-grade impairment scale and is used to avoid quantization errors. It is shown in Figure 2. The panel consists of five expert viewers and two non-expert viewers. The assessors are required to vote on the degree of ringing distortions only, and are trained before the subjective test for differentiating ringing artifacts from other distortions. This is achieved using four processed sequences – Sailboat (1.0 Mbps and 3.0 Mbps) and Tempete (1.0 Mbps and 2.0 Mbps).
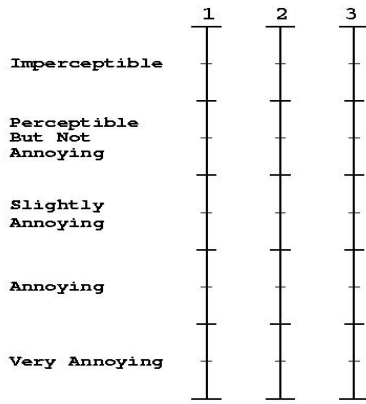

Figure 2. The continuous scale used in DSIS-II tests.

B. Objective Prediction

The test materials are processed as shown in Figure 1 to obtain the ringing distortion measure, d, which in turn determines the PRDM on a scale between 0 and 5. This corresponds to the subjective test scale, which is the predicted MOS by the objective metric. The PRDM is calculated using the following formula:

$$PRDM = \begin{cases} 5 - d^{0.75} & \text{if} \quad d < 5^{4/3}, \\ 0 & \text{otherwise.} \end{cases}$$

The exponent 0.75 is acquired through a best fit of d to the MOS using the nonlinear least-square regression method. The saturation threshold $5^{4/3}$ is a precaution in case of a very large value of d.

C. Data Analysis

The MOS and the associated confidence intervals are calculated. No subjects have been discarded after the screening process as dictated in the ITU-R BT.500 [15].

Three evaluation metrics are used to compare the PRDM with the MOS for ringing distortion assessment: Pearson correlation for prediction accuracy, Spearman rank-order correlation for prediction monotonicity [16], and the average absolute error between the MOS and the PRDM (E|error|) [11].

The performance of the PSNR is also investigated since the PSNR performs unexpectedly well as a generic video quality metric in VQEG tests [3]. Meanwhile there is no other suitable ringing distortion metric to benchmark against. Presently no standard formula scales PSNR into the range from 0 to 5. However, the PSNR is passed through the logistic fit because it partly serves this purpose by assuming a monotonic nonlinear relationship.

Table II presents the evaluation results of the PRDM and PSNR. The average standard deviations of the subjective data are also shown. Figure 3 illustrates a scatter plot of the PSNR versus the MOS. Figure 4 illustrates a scatter plot of the PRDM versus the MOS. The experimental results show that the PRDM achieves a very good agreement with the MOS and a superior performance than the PSNR in measuring ringing artifacts.

TABLE II
METRIC PERFORMANCE

Metric	Pearson	Spearman	E|error|	Average standard deviation
PRDM(both)	0.9397	0.9805	0.2651	0.4402
PSNR(both)	0.7674	0.7489	0.5358	
PRDM(MC)	0.9642	1	0.2457	0.4219
PSNR(MC)	0.9341	1	0.6659	
PRDM(TT)	0.9945	1	0.2845	0.4649
PSNR(TT)	0.9709	1	0.4056	

IV CONCLUSIONS

Performance evaluation of a new perceptual ringing distortion metric has shown very high correlations and low prediction errors compared with subjective mean opinion scores, providing a better alternative to the traditional objective measures, such as the PSNR.
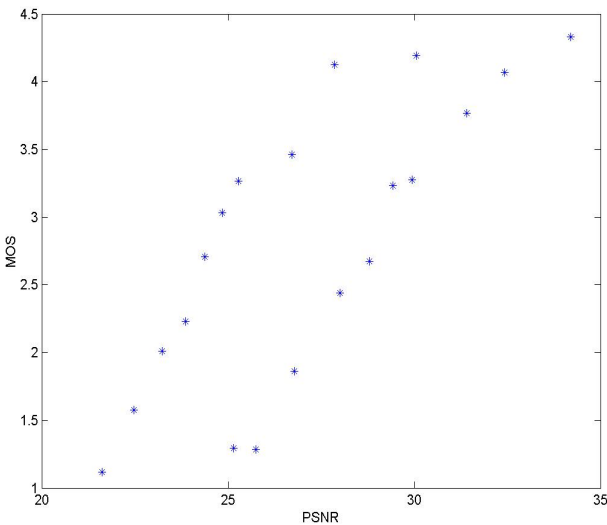
ACKNOWLEDGMENTS

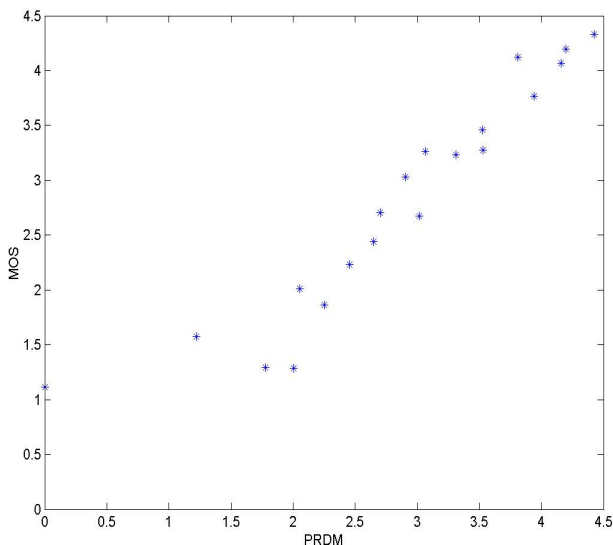Figure 3. Scatter plots of the PSNR versus the MOS.

Figure 4. Scatter plots of the PRDM versus the MOS.

REFERENCES

[1] C. J. van den Branden Lambrecht, "Special Issue on Image and Video Quality Metrics", Signal Processing, Vol. 70, No. 3, November 1998.

[2] H. R. Wu, "Special Session: Testing and Quality Metrics for Digital Video Services", in Proc. of Visual Communications and Image Processing 2000, ed. by K.N. Ngan, T. Sikora and M.-T. Sun, Proc. of SPIE, Vol 4067, pp.741-809, June 2000.

[3] VQEG, ITU-T Study Group 9-Contribution 80, "Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment", COM 980-E, June 2000. (available from ftp.its.bldrdoc.gov.)

[4] J.E. Farrell and C. J. van den Branden Lambrecht, "Special Issue on Translating Human Vision Research into Engineering Technology", in Proc. of IEEE, Vol. 90, No.1, Jan. 2002.

[5] J. Caviedes and S. Mitra, "Special Session: Objective Video Quality Metrics", in Proc. of IEEE ICIP, 2002.

[6] M. Yuen and H. R. Wu, "A survey of hybrid MC/DPCM/DCT video coding distortions," Signal Processing, vol. 70, pp. 247-278, 1998.

[7] ANSI T1.801.02-1995, "Digital transport of video teleconferencing/video telephony signals – Performance terms, definitions, and examples," American National Standard for Telecommunication, 1995.

[8] M. Miyahara and K. Kotani, "Block distortion in orthogonal transform coding-analysis, minimization, and distortion measure", IEEE Transactions on Communications, vol. COM-33, no.1, pp.90-96, Jan. 1985.

[9] S.A. Karunasekera and N.G. Kingsbury, "A distortion measure for blocking artifacts in images based on human visual sensitivity", IEEE Transactions of Image Processing, vol.4, no.6, pp.713-724, June 1995.

[10] H. R. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding", IEEE Signal Processing Letters, vol.4, no.11, pp.317-320, Nov. 1997.

[11] Z. Yu, H. R. Wu, S. Winkler and T. Chen, "Vision-model-based impairment metric to evaluate blocking artifacts in digital video", Proc. of the IEEE, vol.90, no.1, pp.154-169, Jan. 2002.

[12] Z. Yu, H. R. Wu and T. Chen, "A perceptual measure of ringing artifact for hybrid MC/DPCM/DCT coded video", Proc. of the IASTED International Conference, Signal and Image Processing, pp.94-99, IASTED/ACTA Press. 2000.

[13] P. C. Teo and D. J. Heeger, "Perceptual image distortion", Proc. of the SPIE, vol.2179, pp. 127-141, 1994.

[14] MPEG Software Simulation Group, mpeg2encode /mpeg2decode version 1.1, Jun. 1994, available from http://www.mpeg.org/MPEG/MSSG/.

[15] ITU-R Rec. BT 500-10, "Methodology for the subjective assessment of the quality of television pictures", ITU, Geneva, Switzerland, 2000.

[16] VQEG, "Full reference television phase II subjective test plan", Sept. 2002, available from www.vqeg.org.