

•Hongkui Wang<sup>1</sup>, Li Yu<sup>1</sup>, Xiatao Tang<sup>2</sup>, Haibing Yin<sup>3</sup>, and Junhui Liang<sup>1</sup>

<sup>1</sup>School of Electronic Information and Communications Huazhong University of Science and Technology, <sup>2</sup>Zhejiang Special Equipment Research Institute, <sup>3</sup>School of Communication Engineering, Hangzhou Dianzi University

## ABSTRACT

The just noticeable distortion (JND) has been widely applied in perceptual image/video compression. However, the existing JND models are not accurate enough, which results in the degradation of perceptual quality. In this paper, we propose a quantization distortion and JND (QD&JND) compensation based perceptual video coding (PVC) scheme to compress videos. Experimental results show that the proposed scheme achieves bit rate saving significantly with better perceptual quality.

## INTRODUCTION

With the explosive growth of video data, video compression has been confronted with enormous challenges. In order to achieve highly compact representation for videos, perceptual video coding (PVC) has made great process and numerous perception-based compression technologies [1-2] have been proposed with the in-depth study of the human visual system (HVS).

For a visual stimulus, previous perceptual studies have indicated that human eyes can not perceive the distortion below a certain visibility threshold. This threshold is so-called as the just noticeable distortion (JND) [3]. Because the JND threshold directly reflects the characteristics of the HVS, numerous JND-based compression tools have been applied into perception-based image/video applications. Nevertheless, existing JND-based PVC schemes have following shortcomings. 1. Most JND models were constructed based on the fixed block size, such as  $8 \times 8$ . Therefore, these models are not sufficient for the prevailing video coding standards with variable block-sized processing. 2. The HVS is an extremely complex system which includes more than two dozen interconnected visual cortex areas. The visual signal processing is difficult to imitate and the existing JND models are not accurate enough. To improve this situation, we propose a JND compensation based PVC scheme for improving the quality of the reconstructed video. In the encoder, the JND threshold of each block is estimated at first. Then, the residue is suppressed with a quantization distortion and JND (QD&JND) based filter. In the decoder, a QD&JND-based compensation scheme is developed to improve the quality of reconstructed videos.

## THE PROPOSED PVC SCHEME

Usually, the HVS is sensitive to plain and contour areas, and insensitive to texture areas. In this paper, we calculate the mean absolute difference (MAD) of each block and find that the intensity of MAD increases in turn in plain, texture and contour areas. Thus, we propose a MAD-based block classification method to divide image blocks into plain, contour and texture types. The intensity of the

MAD for each block is calculated by

$$\tau = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |I(i, j) - \bar{I}| / (KN^2)$$

$K$  is the dynamic range of pixel intensity (255 in 8-bit image format).  $I(i, j)$  is the pixel intensity.  $\bar{I}$  is the average intensity of the block. The block type is determined with two parameters ( $\tau_{pt} = 0.01$  and  $\tau_{pc} = 0.05$ ) as follows:

$$\text{block type} = \begin{cases} \text{Plain}, & \tau \leq \tau_{pt} \\ \text{Texture}, & \tau_{pt} < \tau \leq \tau_{pc} \\ \text{Contour}, & \tau > \tau_{pc} \end{cases}$$

In pixel-based JND estimation, the model is constructed mainly by considering the luminance adaptation (LA) and the contrast masking (CM) effects. The JND threshold for the LA effect ( $J_{LA}$ ) is calculated based on the background luminance of each pixel. In this paper, the background luminance of each pixel is represented by the average intensity of the current block. So, the LA-JND threshold of  $(i, j)$ -th pixel is obtained as follows:

$$J_{LA}(i, j) = \begin{cases} 17 \times (1 - \sqrt{\bar{I}/127}) + 3, & \text{if } \bar{I} \leq 127, \\ 3 \times (\bar{I} - 127) / 128 + 3, & \text{Otherwise.} \end{cases}$$

The JND threshold for the CM effect ( $J_{CM}$ ) is usually estimated based on the contrast intensity of each pixel. In this paper, we apply the MAD intensity to measure the contrast intensity of each block. Because of the selective response of neurons, the JND threshold of the contour block is comparatively smaller. Therefore, an inhibitory factor ( $\eta$ ) is set to suppress the contrast intensity for contour blocks. The JND threshold of CM effect is estimated by

$$J_{CM}(i, j) = \beta \cdot \eta(n) \cdot \tau(n)$$

$n$  is the block index.  $\eta(n)$  is the inhibitory factor which is set 1/3 for contour blocks and 1 for plain and texture blocks.  $\tau(n)$  is the MAD intensity of the  $n$ -th block.  $\beta$  is scaling factor and is set 30 based on our experiments. Finally, the JND threshold of the current coding block is obtained with NAMM model by

$$J(n) = J_{LA}(n) + J_{CM}(n) - C^{gr} \cdot \min(J_{LA}(n) + J_{CM}(n))$$

$C^{gr}$  is the gain reduction parameter and is set to 0.3. In the encoder, the residue is suppressed with a QD&JND-based filter. The filter is designed by

$$R'(i, j) = \text{sign}(R(i, j)) \cdot \max(0, R(i, j) - \max(0, J_E(n) - D(Qstep)))$$

Here,  $R$  is residue,  $J_E(n)$  is the JND threshold of  $n$ -th block estimated in the encoder. In the decoder, the distortion of the reconstructed residual block ( $r$ ) is compensated and the final reconstructed block ( $Rec$ ) is obtained by

$$Rec(i, j) = P(i, j) + r(i, j) + \text{sign}(r(i, j)) \cdot \max(0, J_D(n) - D(Qstep))$$

$P$  is the predictive block.  $r$  is the reconstructed block.  $J_D(n)$  is estimated JND threshold of the  $n$ -th block in the decoder. When the residues are less than the JND threshold, these residues are filtered to zero. The distortions of these residues can not be compensated because the signs and the real filtered residual intensities of these residues are unknowable.

## RESULT

Usually, the performance of JND models can be measured by the ability of hiding distortion. We add almost the same noise to the test sequences by  $I'(i, j) = I(i, j) + \gamma \cdot S_r(i, j) \cdot J(i, j)$ .  $I$  is the original sequence and  $I'$  is the distorted sequence with the noise injection.  $\gamma$  is a parameter applied to adjust the energy of JND noise.  $S_r(i, j)$  is bipolar random noise of  $\pm 1$ . A good performance of a JND model under test means that the distorted video by the JND model shows a lower PSNR value with the same or higher MS-SSIM value against other models under comparison. The test result in the Table 1 shows that the proposed model is more accurate than Wu's model [4]. The proposed PVC scheme is incorporated into the latest version of the HEVC reference codec (HM16.20) with two cases. Case 1: the video is compressed by the proposed filter and without distortion compensation in the decoder. Case 2: the video is compressed by the proposed filter and the distortion is compensated in the decoder. As shown in Table 2, simulation results show that the proposed scheme achieves 2%-11% bit rate saving with negligible perceptual quality reduction and performs better in the smaller Qps. The complexity increases 4.63% and 10.7% for two cases, respectively.

Table 1: Model performance comparisons in terms of PSNR (dB) and MS-SSIM

Class	Wu's Model [21]		Proposed Model		Class	Wu's Model [21]		Proposed Model	
	PSNR	MS-SSIM	PSNR	MS-SSIM		PSNR	MS-SSIM	PSNR	MS-SSIM
A	36.289	0.997	34.996	0.998	D	35.220	0.996	37.384	0.998
B	36.108	0.996	35.364	0.998	E	37.904	0.998	35.347	0.998
C	34.605	0.993	37.870	0.998	F	36.560	0.996	35.026	0.994
Average						36.114	0.996	35.998	0.997

Table 2: The test result of the proposed PVC scheme with different Qps

Qp	Case 1 vs HM16.20				Case 2 vs HM16.20			
	$\Delta R(\%)$	$\Delta P(\text{dB})$	$\Delta M$	$\Delta T(\%)$	$\Delta R(\%)$	$\Delta P(\text{dB})$	$\Delta M$	$\Delta T(\%)$
22	-7.59	-2.447	-0.003	2.84	-5.48	-1.604	-0.002	5.22
27	-4.88	-1.389	-0.003	2.75	-4.49	-1.011	-0.002	9.20
32	-3.20	-0.588	-0.002	4.68	-3.27	-0.582	-0.002	12.75
37	-3.20	-0.405	-0.003	8.26	-3.92	-0.390	-0.003	15.66

## CONCLUSION

In this paper, the QD&JND compensation based PVC scheme was proposed to suppress the residue in pixel domain. We developed a block-level JND estimation model to estimate the JND threshold of the original and reconstructed blocks. In the encoder, the residue is suppressed with the proposed QD&JND-based filter. In the decoder, the distortion of the reconstructed block is compensated. Simulation results show that the proposed QD&JND-based residue filter reduces the bit rate effectively and the proposed QD&JND compensation scheme improves the quality of compressed videos significantly.

## REFERENCES

- [1] Z. Chen and C. Guillemot, "Perceptually-friendly H.264/AVC video coding based on foveated just-noticeable-distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 806-819, 2010.
- [2] X. Zhang, S. Wang, K. Gu, W. Lin, S. Ma, and W. Gao, "Just-noticeable difference-based perceptual optimization for JPEG compression," *IEEE Signal Process. Lett.*, vol. 24, no. 1, pp. 96-100, 2017.
- [3] S. L. Macknik and M. S. Livingstone, "Neuronal correlates of visibility and invisibility in the primate visual system," *Nature Neuroscience*, vol. 1, no. 2, pp. 144-149, 1998.
- [4] J. Wu, L. Li, W. Dong, G. Shi, W. Lin, and C.-C. Jay Kuo, "Enhanced just noticeable difference model for images with pattern complexity," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2682-2693, 2017.