2015 IEEE 16th International Conference on Information Reuse and Integration

PERCEPTUALLY LOSSLESS IMAGE CODING BASED ON FOVEATED JND

Yiming Li, Hongyi Liu, Zhenzhong Chen
School of Remote Sensing and Information Engineering, Wuhan University, China
Email: zzchen@whu.edu.cn

***Abstract*—Removing perceptual redundancy plays an important role in image compression. In this paper we developed a foveated just-noticeable-difference (FJND) model to quantify the perceptual redundancy in the image and integrate it in the H.265/HEVC intra encoding framework to provide a perceptually lossless image coding method. The experiments demonstrate that the proposed method effectively increases the compression performance.**

***Keywords*-Perceptual coding; H.265/HEVC; FJND; lossless.**

I. Introduction

How to remove redundancy in the image for effective image compression has played an important issue in image applications. JPEG or JPEG2000 have been developed for image compression. Moreover, there are many fields need lossless images to guarantee the accuracy of their analysis, like remote sensing image, medical image, etc. Therefore, efficiency lossless compression algorithm is desirable. Some lossless image compression algorithms have been developed, such as JPEG_LS. However, in most applications people just need perceptually lossless image by which we can process and analyse the image for our applications. So removing perceptual redundancy in the image to provide perceptually lossless image coding is desired and can further improve efficiency of the compression.

In this paper, we propose a perceptually lossless image compression method based on H.265/HEVC intra coding technique and Foveated Just-Noticeable-Difference (FJND) model. H.265/HEVC, which means high efficiency video coding, was jointly developed by ISO/IEC Moving Picture Experts Group (MPEG) and ITU-T Video Coding Experts Group (VCEG) as ISO/IEC 23008-2 MEPG-H Part 2 and ITU-T H.265[1]. It can not only compress videos but also images efficiently. As during the standarization, H.265/HEVC has adopted some effective coding tools for efficient image compression. So we utilize its intra coding functionality as still image coding framework. With this efficient coding framework, we aim to integrate a perceptual model which can effectively quantify the perceptual redundancy in the image therefore we can use it in perceptually lossless image compression. During these past few years, several perceptual models have been proposed with the better understanding of HVS such GBVS model for visual attention, JND model for visual sensitivity, etc. Foveated JND model can be developed by designing a JND model with the considerations on humans visual foveation characteristics. Due to the non-uniform distribution of photoreceptor cells on human retina, the human visual system has different visual acuity to perceive image details [2]. The experiments of [4] proved that FJND has better approximation to human visual perception. Therefore, we propose to apply the foveated JND model to describe the smallest detectable difference of the image. Then, we propose a perceptually lossless image coder based on H.265/HEVC and FJND to compress the image. If we guarantee the maximum distortion in the compressed image is smaller than the detectable difference by the FJND model provided, we can get a perceptually lossless image.

The rest of the paper is organized as follows. Section II provides a brief introduction of H.265/HEVC and how to quantify perceptual redundancy by foveated JND model. Section III describes the details of the proposed perceptually lossless image coding. Section IV shows the experiments and Section V concludes the paper.

II. FRAMEWORK

A. *Introduction of H.265/HEVC*

H.265/HEVC is the latest video compression standard, compared with H.264/AVC, which can improve the compression performance up to 50% with the same visual quality [6]. The improvement benefits from various aspects such as quadtree structure of the coding unit, sample adaptive offset (SAO), advanced motion vector pre-diction (AMVP), etc [1]. Coding tree unit (CTU) is a new structure, which replace 16×16 pixel macroblocks in H.264/AVC. In H.265/HEVC, the picture is divided into CTUs which size are 64×64, 32×32, 16×16 and 8×8. The division of CTU according to different video texture adaptively that increases the coding efficiency. In addition, H.265/HEVC defines PU (Prediction Unit) for prediction coding and TU (Transform Unit) for transform. In inter prediction, H.265/HEVC allows advanced motion vector prediction(AMVP) to improve coding efficiency while a merge mode for motion vector coding is used, as well [7]. In addition, sample adaptive offset (SAO) is added in H.265/HEVC to reconstruct the signal [1]. In intra prediction, H.265/HEVC uses 33 direction modes compared to 8 directional modes by H.264/AVC [1]. This feature can guarantee more precise prediction direction.

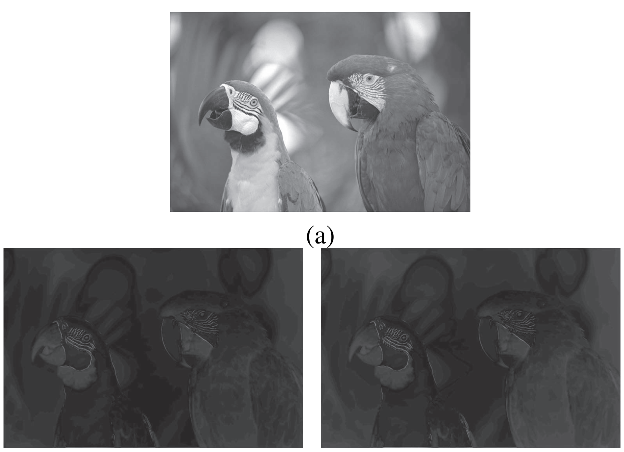
This work was supported in part by National Natural Science Foundation of China (No. 61471273) and Natural Science Foundation of Hubei Province of China (No. 2015CFA053).

978-1-4673-6656-4/15 $31.00 © 2015 IEEE
DOI 10.1109/IRI.2015.21

B. Foveated JND

Just-Noticeable-Difference (JND) measures the smallest detectable difference between two signals therefore can be utilized to quantify the perceivable distortion in the noise contaminated image. Since the visual signal is ultimately perceived by HVS [8], [9], JND thresholds could hence be used to determine optimum quantization step sizes for different parts of the image for perceptually lossless image coding thus achieve the highest perceptual quality at a given bit-rate,

The JND estimation can be accomplished through modeling the relationship between visual sensitivity and spatial-temporal masking effects [10], [11]. Various computational models have been proposed in both spatial and frequency domains to exploit the perceptual redundancy of HVS. SJND (spatial domain JND) assumes that each pixel of the image is projected on the fovea area of HVS, where it can be perceived under the highest visual acuity. However, the visual acuity will decrease when visual stimulus is projected out of the fovea region. Thus the SJND could only provide a local visibility threshold. In [3], [4], a foveated JND (FJND) has been proposed to measure the global visibility threshold of the whole image by incorporating the viewing eccentricity to SJND model. FJND enhances the SJND model by accounting for the relationship between decreased visibility corresponding to the increased eccentricity. FJND can better exploit the perceptual redundancy in the area of less visual significance to observers. A comparison between SJND and FJND is shown in Fig.1.



(a)

(b) (c)

Figure 1. Fig.1 (a) Original image. (b) Map of SJND (x5 for display). (c) Map of FJND (x5 for display).

In this paper we utilize the FJND model proposed in [4] to quantify distortion visibility of images. To obtain the FJND threshold, the SJND of a given image is firstly calculated according to luminance contrast and spatial masking.

$$SJND(x,y) = \max\{f1(bg(x,y), mg(x,y)), f2(bg(x,y))\} \tag{1}$$

where $f1(bg(x,y), mg(x,y))$. and $f2(bg(x,y))$ are functions to estimate the spatial masking and luminance contrast. Detailed illustration can be found in [4].

In order to describe the foveation behavior of Human Visual System, a foveation model F is developed based on the experiments which obtain the visibility threshold due to the foveation property of HVS. The F is both background luminance and eccentricity dependent, and it tends to increase with the increased viewing eccentricity. The FJND threshold of pixel $P(x,y)$ is defined as Equations (2) to (4) shown:

$$FJND = SJND(x,y) \cdot F(x,y,e) \tag{2}$$

$$e = \tan^{-1}(\frac{d}{v}) \tag{3}$$

$$d = \sqrt{(x - x_f)^2 + (y - y_f)^2} \tag{4}$$

where e denotes the viewing eccentricity from the fixation point $P(x_f, y_f)$ to pixel $P(x,y)$, and v is the physical viewing distance between observers and monitor. Further explanations of F is available in [4]. The foveation model F assumes that the fixation point is known. To obtain fixation point, we utilize GBVS model [12] to locate conspicuous visual area of the given image and then calculate the fixation point through the GBVS saliency model. Since there may exist multiple fixation points in practice, the foveation model F can be estimated by only considering the closet fixation point that results in the smallest eccentricity e and minimum foveation weight for F.

III. PERCEPTUALLY LOSSLESS IMAGE CODING HEVC

In H.265/HEVC intra coding, the image is divided into many blocks which range from 64×64 to 8×8. The encoder adaptively choose different depth to split the units, which ranges from 0 to 3. In addition, if the depth is 3, the CU block can continually be divided into 4 blocks and these are the smallest size of PU, which size are all 4×4 pixels. When the image is divided into 64×64 size of unit which depth is 0 initially. Then the encoder choose the most probable mode (MPM). In each MPM (3 or 8), it pre-split the unit into four 32×32 size block and choose the best prediction angle by calculating rate-distortion (R-D) cost, and then, the encoder choose whether split as this predict angle or not by rate distortion optimization (RDO). It repeats these steps (maximum up to depth 3) until reach the best R-D cost. The steps are shown in Fig.2.

The proposed perceptually lossless image coding is developed based this framework. We assume human cannot perceive any noise below the foveated JND threshold [4]. So, if we want to reduce more bits and guarantee the distortion of each position in the image is smaller than the detectable difference provided by the foveated JND model, we compare the sum of squares due to error (SSE) of this block than the

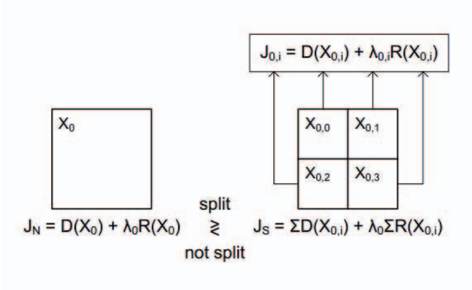


Fig.2 The Split Flow by RDO [13].

foveated JND value in the same position. As Equation (5) shown, if we obtain the higher QP value we can reduce more bits but still get the perceptually lossless image.

$$\max\{QP\}s.t.\ SSE(i,j) \leq JND(i,j) \tag{5}$$

$$SSE = \sum_{i=1}^{n} w_i(y_i - y_i^{'})^2 \tag{6}$$

We use a binary-search algorithm to find the optimal QP value and SSE is no larger than the sum of detectable difference which foveated JND model provided in the same size, and then, store the QP value so that we can get both minimum bit rate and conform the foveated JND model. The algorithm is shown in Algorithm (1). And then we do this operation in each depth which range from 0 to 3. So that we can utilize the HEVC encoder RDO rule that shown as Equation (7) to determine the best CU split that we can get the best balance bewtween the distortion and bitrate.

$$\min J_N = D(X_0) + \lambda_0 R(X_0) \tag{7}$$

Besides, it is worth noticeable that MPM and lambda are based on the current QP value. So if we change the QP we should also recalculate the MPM and λ. To solve this, we preset some reference QP values (1, 6, 11, 17) to calculate their corresponding MPMs, then add them into the MPM list if they are different. For recalculating λ, we calculate and update the value after every time QP changed. By this scheme, we obtain relative MPM and λ estimate veracity and relative low complexity compared with the iterative loop algorithm.

IV. Experimental Results

To compare the effectiveness of the proposed perceptually lossless coding algorithm, the original H.265/HEVC and proposed method are implemented. Subjective tests have also been conducted to verify the performance of our proposed algorithm. The original HM method is used by setting its QP value to 0 to represent lossless comparison. We used different standard JPEG test images. In these test images, they include scene, person, object and others. The resolution of these pictures are 576×768 or 768×576.

Algorithm 1 Calculate $QPvalue$

Require:
original image:I;
Foveated JND map of the image:S;
parameter to control threshold:
QP_0,$begQP$,$midQP$,$lastQP$;

set parameter
  $QP_0 \leftarrow originalHEVCQPvalue$
  $begQP \leftarrow QP_0$
  set $lastQP$
for each $CU(i,j)$ in I **do**
  $FJND(i,j)$ = get sum value of (i,j) block in S
  while $begQP \leq lastQP$ **do**
    compute:
    set $midQP$=$begQP + lastQP$
    distortion $D(I, J.midQP)$ of this CU(i,j) in this QP by H.265/HEVC framework
    if $D(i,j,midQP) == FJND(i,j)$ **then**
      break;
    end if
    if $D(i,j,midQP) < FJND(i,j)$ **then**
      set $begQP = midQP + 1$
    else
      set $midQP = midQP - 1$
      set $lastQP = midQP$
    end if
  end while
end for

Ensure:
$midQP$

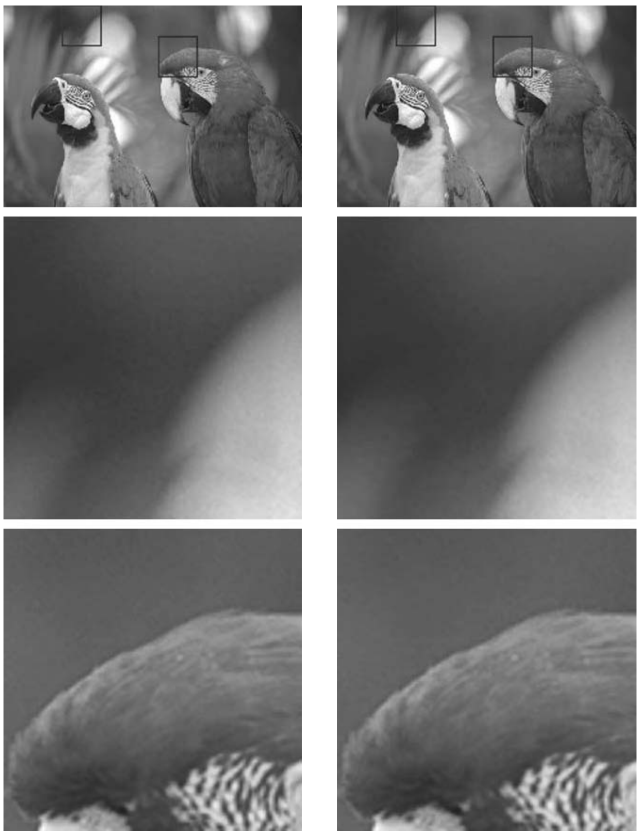
Fig.4 provides the comparisons of the results. We can see that our proposed method consumes less bits and the perceptual area has no visual quality loss. The subjective test protocol we used is the double-stimulus continuous quality scale (DSCQS) protocol which is recommended in Rec. ITU-R BT.500 [14]. The vote is to compare each two stimulus pictures A and B. Stimulus A is randomly the original HEVC result or the picture compressed by our method and stimulus B is the other result. The observers are asked to vote by the DSCQS protocol. The score range from -3 to 3 that -3 means B is much worse than A, and 0 means B is the same quality as A. Table 1 shows the results where PSPNR measures the perceived image quality as suggested in [4].

V. Conclusion

In this paper, a perceptually lossless method is proposed to further improve H.265/HEVC efficiency of image compression. With the model of Foveated JND, we change the original RDO to distortion-based-RDO model. As shown in the test, the proposed method have 56.64% bit rate loss without perceptual quality loss.

Table I
SIMULATION RESULTS

	bpp		PSNR(Y) (dB)		PSPNR(Y)(dB)		MOS	Bit Saving
	HM	Proposed	HM	Proposed	HM	Proposed		
kodim04	5.17	2.34	78.34	48.35	NaN	73.66	0	54.74%
kodim08	6.53	3.59	78.08	48.12	NaN	76.71	0	45.02%
kodim13	7.44	4.38	79.38	48.39	NaN	79.52	0	41.13%
kodim17	5.07	2.21	78.65	48.30	NaN	75.18	0	56.41%
kodim19	5.55	2.61	78.13	48.08	NaN	71.56	0	52.97%
kodim23	4.56	1.67	77.47	48.39	NaN	74.95	0	63.38%
kodim24	5.8	2.98	80.03	48.61	NaN	74.63	0	48.62%
Average								51.75%



(a) (b)

Fig.4 The comparison. (a) HM result (b) our result.

REFERENCES

[1] G. J. Sullivan, J. Ohm, W. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[2] Z. Chen and H. Liu, "JND modeling: Approaches and applications," in *Digital Signal Processing (DSP), 2014 19th International Conference on*. IEEE, 2014, pp. 827–830.

[3] Z. Chen and C. Guillemot, "Perceptually-friendly H. 264/AVC video coding based on foveated just-noticeable-distortion model," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 20, no. 6, pp. 806–819, 2010.

[4] J. R. Ohm, G. J. Sullivan, and H. Schwarz, "Comparison of the coding efficiency of video coding standardsincluding high efficiency video coding (HEVC)," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1669–1684, 2012.

[5] J. L. Lin, Y. W. Chen, and Y. P. Tsai, "Motion vector coding techniques for HEVC," in *IEEE 13th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2011, pp. 1–6.

[6] X. K. Yang, W. S. Lin, Z. K. Lu, E. P. Ong, and S. S. Yao, "Just noticeable distortion model and its applications in video coding," *Signal Processing: Image Communication*, vol. 20, no. 7, pp. 662–680, 2005.

[7] W. S. Geisler and J. S. Perry, "Real-time foveated multiresolution system for low-bandwidth video communication," in *Photonics West'98 Electronic Imaging*. International Society for Optics and Photonics, 1998, pp. 294–305.

[8] C.-H. Chou and Y.-C. Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 467–476, 1995.

[9] X. Yang, W. S. Lin, Z. Lu, E.-P. Ong, and S. Yao, "Motion-compensated residue preprocessing in video coding based on just-noticeable-distortion profile," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 15, no. 6, pp. 742–752, 2005.

[10] Z. Chen and C. Guillemot, "Perceptually-friendly H.264/AVC video coding," in *IEEE International Conference on Image Processing (ICIP)*. IEEE, 2009, pp. 3417–3420.

[11] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Advances in neural information processing systems*, 2006, pp. 545–552.

[12] C. Yeo, H. L. Tan, and Y. H. Tan, "SSIM-based adaptive quantization in HEVC," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2013, pp. 1690–1694.

[13] ITU-R, "BT. 500-11: Methodology for the subjective assessment of the quality of television pictures," *International Telecommunication Union*, 2002.