# Just-Noticeable Difference-Based Perceptual Optimization for JPEG Compression

Xinfeng Zhang, *Member, IEEE*, Shiqi Wang, *Member, IEEE*, Ke Gu, *Member, IEEE*, Weisi Lin, *Fellow Member, IEEE*, Siwei Ma, *Member, IEEE*, and Wen Gao

***Abstract*—The Quantization table in JPEG, which specifies the quantization scale for each discrete cosine transform (DCT) coefficient, plays an important role in image codec optimization. However, the generic quantization table design that is based on the characteristics of human visual system (HVS) cannot adapt to the variations of image content. In this letter, we propose a just-noticeable difference (JND) based quantization table derivation method for JPEG by optimizing the rate-distortion costs for all the frequency bands. To achieve better perceptual quality, the DCT domain JND-based distortion metric is utilized to model the stair distortion perceived by HVS. The rate-distortion cost for each band is derived by estimating the rate with the first-order entropy of quantized coefficients. Subsequently, the optimal quantization table is obtained by minimizing the total rate-distortion costs of all the bands. Extensive experimental results show that the quantization table generated by the proposed method achieves significant bit-rate savings compared with JPEG recommended quantization table and specifically developed quantization tables in terms of both objective and subjective evaluations.**

***Index Terms*—Image compression, just-noticeable difference (JND), quantization table, rate-distortion.**

## I. Introduction

BLOCK-BASED transform is widely utilized in lossy image/video compression standards to reduce the correlations among pixels in each block. The transformed coefficients in each block are further quantized according to the given quantization steps to remove the perceptual redundancy. Since human visual system (HVS) shows different sensitivities in terms of frequencies, quantization steps are usually configured with different values for different frequency bands to achieve better compression performance without incurring perceptible distortions. The quantization table recommended by JPEG standard [1] tends to preserve low-frequency information and discard high-frequency details, because HVS is less sensitive to the information loss in high-frequency bands.

However, the JPEG standard quantization table only considers the HVS features while the relationship between the distortions and rates has been largely ignored. In [2], Eernawan and Nugraini derived new quantization tables according to a primitive psychovisual threshold to achieve the optimal balance between the quality of reconstructed images and compression rates. Though significant performance improvement has been achieved, the quantization tables are also fixed for different kinds of images. Ratnakar and Livny [3] derived the rate-distortion optimization-based quantization tables for DCT coefficients according to the statistical characteristics of individual images, which achieves significant coding gain. Yet, they employed the mean square error (MSE) to quantify the quantization distortions, which has been repeatedly proven to be poorly correlated with HVS.

Based on the recent researches [4], human perceived distortions are discretely characterized by some jumps, which can be well explained by just-noticeable difference (JND). Therefore, potential coding gains can be further expected by allocating bits according to the JND-based distortions. Bai *et al.* [5] first introduced the spatial-temporal JND model into multiple description coding by only encoding the visual information that cannot be predicted well within the JND tolerance, and this method achieved better bit rate and perceptual visual-distortion performance. In this letter, we propose a JND-based quantization table generation method by optimizing the rate-distortion costs for all the DCT bands, where the JND-based distortion metric is utilized instead of MSE. By utilizing the first-order entropy to estimate the coding rates, we formulate the quantization table generation as a joint unconstraint rate-distortion optimization problem for all the DCT bands using Lagrange multiplier method. The Lagrange multiplier is estimated for every band based on the statistics of the input images, and a greedy method is proposed to search the optimal quantization steps for every DCT band according to the target distortions.

The remainder of this letter is organized as follows. Section II introduces the DCT domain JND model. Section III provides the detailed description of the proposed JND-based quantization table derivation method for JPEG compression. Experimental

Manuscript received August 20, 2016; revised December 6, 2016; accepted December 15, 2016. Date of publication December 19, 2016; date of current version January 5, 2017. This work was carried out at the Rapid-Rich Object Search (ROSE) Laboratory, Nanyang Technological University, Singapore. The ROSE Laboratory was supported by the National Research Foundation, Singapore, under its Interactive Digital Media Strategic Research Programme. This work was also cooperated with Peking University, and in part supported by the National Basic Research Program of China (973 Program) under Grant 2015CB351800. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yao Zhao.

X. Zhang, S. Wang, and W. Lin are with the Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore 639798 (e-mail: xfzhang@ntu.edu.sg; wangshiqi@ntu.edu.sg; wslin@ntu.edu.sg).

K. Gu is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: guke@ntu.edu.sg).

S. Ma and W. Gao are with the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing 100871. China (e-mail: swma@pku.edu.cn; wgao@pku.edu.cn).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2016.2641456

1070-9908 © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

results and analyses are reported in Section IV, and Section V concludes the letter.

## II. DCT Domain JND Model

JND refers to the visibility threshold below which the changes cannot be detected by the majority of subjects [6]. In recent years, numerous works about JND modeling were proposed in spatial and frequency domains. Especially, the frequency domain methods mainly rely on the DCT transform [7], [8], as DCT has been widely utilized in most of the existing image/video compression and processing fields [9]–[11].

The general form of the DCT domain JND is calculated on the blocks with specific size. In this letter, the DCT domain JND model on $8 \times 8$ blocks is utilized considering that JPEG compression takes $8 \times 8$ nonoverlapped blocks as the basic coding units. In [12], Wei and Ngan proposed the JND model for $8 \times 8$ blocks, which is expressed as

$$T'(m,n,u,v) = T_{\text{basic}}(u,v) \times \alpha_{\text{lum}}(m,n) \times \alpha_{\text{cm}}(m,n,u,v) \tag{1}$$

$$T_{\text{basic}}(u,v) = \frac{s}{\phi_u \phi_v} \cdot \frac{\exp(cw_{uv})/(a + bw_{uv})}{r + (1-r)\cdot\cos^2\varphi_{uv}} \tag{2}$$

where $\alpha_{\text{lum}}$ and $\alpha_{\text{cm}}$ are the luminance adaptation and contrast masking, respectively, and $(m,n)$ is the block coordinate and $(u,v)$ is the DCT band index. The parameter $s = 0.25$ is the spatial summation effect factor, $1/(r + (1-r)\cdot\cos^2\varphi_{uv}$ accounts for the *oblique* effect, where $r$ is set to 0.6 empirically. $\phi_u$ and $\phi_v$ are the DCT normalization factors and $\varphi_{uv}$ indicates the direction angle of the DCT band $(u,v)$:

$$\phi_m = \begin{cases} \sqrt{1/N}, & m = 0 \\ \sqrt{2/N}, & m > 0 \end{cases} \tag{3}$$

$$\varphi_{uv} = \arcsin\left(\frac{2w_{u0}w_{0v}}{w_{uv}^2}\right). \tag{4}$$

In (1), $w_{uv}$ is the spatial frequency of the band $(u,v)$, which is calculated as follows:

$$w_{uv} = \frac{1}{2N}\sqrt{(u/\theta_x)^2 + (v/\theta_y)^2} \tag{5}$$

where $\theta_x$ and $\theta_y$ are the horizontal and vertical visual angles of a pixel, which are calculated according to the viewing distance and display width of a pixel on monitor [12].

The parameters $\alpha_{\text{lum}}$ and $\alpha_{\text{cm}}$ are the image content dependent variables. It is reported that HVS is more sensitive to medium brightness areas than that dark or bright areas. Therefore, Wei and Ngan formulated an empirical function in (6) to describe the luminance adaptation effect:

$$\alpha_{\text{lum}} = \begin{cases} (60 - B_{k,\text{avg}})/150 + 1, & B_{k,\text{avg}} \le 60 \\ 1, & 60 < B_{k,\text{avg}} < 170 \\ (B_{k,\text{avg}} - 170)/425 + 1, & B_{k,\text{avg}} \ge 170 \end{cases} \tag{6}$$

where $B_{k,\text{avg}}$ is the average intensity of the $k$th block. The contrast masking $\alpha_{\text{cm}}$ describes the perception of one visual component at the presence of another one [13]. In general, it is calculated by dividing the DCT blocks into three classes with descending order of the HVS sensitivity, and then the masking factor can be determined based on inter- and intraband masking [7], [14].

## III. JND-Based Quantization Table

### A. JND-Based Distortion Model and Rate Model

For an image $\mathcal{I}$, the transformed coefficient of band $(u,v)$ in $k$th block is denoted as $F_k(u,v)$. The quantization distortions in each frequency band based on the traditional distortion metric, MSE, can be calculated as follows:

$$D(u,v) = \frac{1}{K}\sum_{k=1}^{K}(F_k(u,v) - \hat{F}_k(u,v))^2 \tag{7}$$

$$\hat{F}_k(u,v) = \text{round}\left(\frac{F_k(u,v)}{Q(u,v)}\right) * Q(u,v) \tag{8}$$

where $K$ is the amount of the nonoverlapped blocks in image $\mathcal{I}$, and $Q$ is the $8 \times 8$ quantization table.

The MSE metric is a continuous function of quantization steps, which is poorly correlated with human perceptual quality. It has been discovered that human perceptual quality is perceived in a discontinuous manner, which is piecewiseconstant for a specific range of MSE [4]. The constant value can be interpreted as the visibility threshold or JND value. Therefore, in this letter, we directly redefine the distortion metric by directly soft-thresholding MSE values according to the JND as a threshold for each coefficient, which is formulated in (9). Specifically, $d_k(q,u,v)$ represents the quantized coefficient distortion of band $(u,v)$ in the $k$th block when quantization step is $q$, and $D_q(u,v)$ is the JND-based distortion for band $(u,v)$. Since JND reflects the perceptual tolerance to distortions for different image content, larger distortions (corresponding to MSE) can be assigned to these areas with higher JND values to saving more coding bits without obviously degrading image visual quality:

$$d_k(q,u,v) = \begin{cases} 0, if|F_k(u,v) - \hat{F}_k(u,v)| \le T_k(u,v) \\ (|F_k(u,v) - \hat{F}_k(u,v)| - T_k(u,v))^2 \\ if|F_k(u,v) - \hat{F}_k(u,v)| > T_k(u,v) \end{cases} \tag{9}$$

$$D_q(u,v) = \frac{1}{K}\sum_{k=1}^{K} d_k(q,u,v). \tag{10}$$

In JPEG compression, the quantized coefficients are further encoded by huffman entropy coding to remove their statistical redundancy. In order to avoid time consuming coding procedure, we utilize the first-order entropy of quantized coefficients in each band to estimate the consumed bits,

$$R(u,v) = -K\sum_{k=1}^{K}(p(C_k(u,v))\log_2(p(C_k(u,v))) \tag{11}$$

where $p(x)$ is the probability of symbol $x$, which is equal to the statistical frequency. $C_k(u,v)$ is the quantization index for coefficient of band $(u,v)$ in $k$th block.

### B. Quantization Table Optimization

Rate-distortion optimization plays an important role in optimizing image and video compression algorithms [15]–[19]. Based on the JND-based distortion model and rate model, the quantization table derivation problem can be formulated as the following constraint rate-distortion optimization problem at a given distortion level:

$$\min \sum_{u,v=0}^{7} R(u,v), s.t. \sum_{u,v=0}^{7} D(u,v) \le D_t \tag{12}$$

where $D_t$ is the target distortion for the compressed image. Based on the Lagrangian optimization techniques, the constrained optimization problem in (12) can be reformulated as a unconstrained minimization problem,

$$\min \quad J(Q) = \sum_{u,v=0}^{7} D(u,v) + \lambda_{u,v} R(u,v). \tag{13}$$

where $\lambda_{u,v}$ is the Lagrange multiplier, which is derived based on the statistics on many images offline and the same for all the bands in traditional methods [3], [15]. Actually, the Lagrange multiplier is also a content-related variable. Therefore, in the proposed method, we estimate the Lagrange multiplier for each band according to the distortion and rate function for the given image.

Since the problem in (13) is a joint optimization problem for 64 frequency bands, it is difficult to directly search across the solution space. Considering that the coefficients in each band are quantized independently, we first split the overall minimization problem into a 64 suboptimization problems according to the frequency bands. For each frequency band, the suboptimization problem can be solved by setting the partial derivative of $J(Q(u,v))$ with respect to $Q(u,v)$ to 0,

$$\frac{\partial J(u,v)}{\partial Q(u,v)} = \frac{\partial D(u,v)}{\partial Q(u,v)} + \lambda(u,v)\frac{\partial R(u,v)}{\partial Q(u,v)} = 0. \tag{14}$$

Then, the $\lambda(u,v)$ for band $(u,v)$ is calculated as

$$\lambda(u,v) = -\frac{\partial D(u,v)}{\partial R(u,v)}. \tag{15}$$

Based on (15), the variable $\lambda_{u,v}$ reflects the distortion decrease rate with respect to the increase of coding rate. To show its variations, we quantize the DCT coefficients of *Lena* with quantization steps from 1 to 255, and calculate the entropy of the quantized coefficients for every band respectively. Fig. 1 shows the rate-distortion curve for different bands, where the JND-based distortions are calculated based on (9) and (10), and the coding bits for each band are estimated according to the entropy of quantized coefficients. We can see that for different bands the rate-distortion curves have different decrease rate at the same quantization step. Therefore, the optimal quantization step achieves the largest distortion decrease rate while minimizing the rate at the given distortion level. In this letter, $\lambda_{u,v}$ is approached by using the finite-difference methods as follows:

$$\lambda_q(u,v) = -\frac{D_{q+1}(u,v) - D_q(u,v)}{R_{q+1}(u,v) - R_q(u,v)} \tag{16}$$

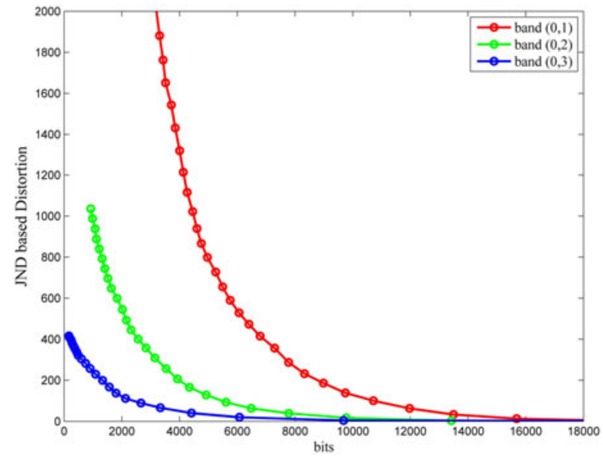


Fig. 1. The rate-distortion curves for different bands.

where $q$ is the quantization step for band $(u,v)$. Since the JND-based distortions reflect the perceivable changes between different quantization steps, it makes the $\lambda_q(u,v)$ more consistent with distortion change rate perceived by HVS.

Considering the practical application of image compression, the quantization steps are finite, and constrained into the range of $[1, 255]$ in this letter. For a given distortion $D_t$, we first calculate the quantization distortions and rates for each band with quantization steps from 1 to 255. Then, the Lagrange multiplier $\lambda_{u,v}$ is further calculated according to (16). An array variable $\Lambda[u][v]$ is utilized to store the $\lambda_{Q(u,v)+1}(u,v)$. For each time, we search the band with minimum $\Lambda[u][v]$, and increase the corresponding quantization step by 1, i.e., $Q(u,v) = Q(u,v) + 1$, and update $\Lambda[u][v] = \lambda_{Q(u,v)+1}(u,v)$ with new quantization value. This procedure is performed iteratively until the distortion achieves the target value. In order to reduce the search range, an initial quantization table is first generated according to the JPEG recommended quantization table. The detailed algorithm is described in Algorithm 1.

## IV. Experimental Results

In this section, we verify the proposed JND-based quantization table optimization method with JPEG standard quantization table, denoted as *JPEG*, and the psychovisual threshold-based quantization table in [2], denoted as *PSY-Table*. We randomly select ten images from the Kodak standard test images, and compress their luminance with different quantization tables. Table I shows the rate reduction when images are compressed at the same quality measured by PSNR. The proposed method achieves about 16.7%$\sim$22.0% bit rate savings for test images, and 18.3% bit rate savings on average compared with JPEG standard. Fig. 2 illustrates the RD curves to show the compression performance in a large bit-rate range. We can see that the proposed method achieves significant coding gains on a large bit rate range. Fig. 3 shows the subjective quality comparison of the close-up of *Bike* compressed with different quantization tables at 0.68 bpp. We can see that the image generated with our quantization table is more visually pleasant, especially less ringing artifacts are perceived at the back of the rider.

**Algorithm 1:** JND based Quantization Table Derivation**:**

**Input**: Original image, $\mathcal{I}$, and the target distortion for the compressed image, $D_t$

**Initialization**: Calculate the initial quantization table according to JPEG standard, $Q_0$;

**JND calculation**: Calculate the DCT domain JND value for each coefficient according to Eqn. (1) and (2);

**JND Based distortion calculation**: Calculate the JND based distortion according to Eqn.(9) and (10), $\{D_q(u,v)|1 \le q \le 255, 0 \le u, v \le 8\}$;

**Rate estimation**: Estimate rates with Eqn.(11), $\{R_q(u,v)|1 \le q \le 255, 0 \le u, v \le 8\}$;

**Lagrange multiplier estimation**: Estimate Lagrange multiplier with Eqn.(16), $\{\lambda_q(u,v)|1 \le q \le 255, 0 \le u, v \le 8\}$;

Update the current distortion variable $D$, and $\Lambda[0\cdots 8][0\cdots 8]$ with $Q_0$ ;

**while** $D \le D_t$ **do**

  Search $\Lambda$ to find the band $(u_0, v_0)$ with minimum value;

  Update $Q$: $Q(u_0,v_0) = Q(u_0,v_0) + 1$;

  Update $D$:

  $D = D + (D_{Q(u_0,v_0)}(u_0,v_0) - D_{Q(u_0,v_0)-1}(u_0,v_0))$;

  Update $\Lambda$: $\Lambda[u_0][v_0] = \lambda_{Q(u_0,v_0)+1}(u_0,v_0)$

**end**

**Output**: Quantization Table, $Q$

TABLE I
JPEG COMPRESSION PERFORMANCE COMPARISON WITH DIFFERENT QUANTIZATION TABLES

| Images | Hat | Motor | Boat | Alfred | Wharf |
|---|---|---|---|---|---|
| PSY-Table | −5.6% | −4.7% | −5.1% | −4.5% | −5.0% |
| Proposed | −22.0% | −17.1% | −17.3% | −17.1% | −17.0% |
| Images | Mountain | Statue | Country | Lighthouse | Ridinghome |
| PSY-Table | −4.9% | −5.1% | −4.6% | −5.3% | −5.1% |
| Proposed | −20.2% | −17.0% | −18.7% | −16.7% | −19.8% |

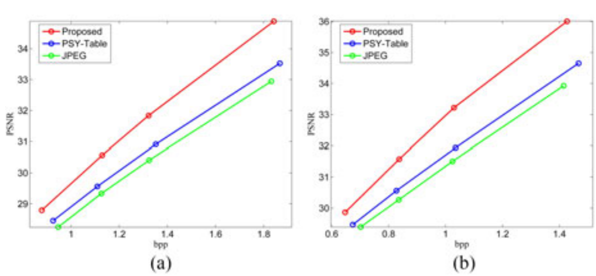


Fig. 2. The RD curve for images *Alfred* and *Ridinghome* compressed by JPEG with different quantization tables. (a) *Alfred*. (b) *Ridinghome*.

In order to further verify the performance of the scheme, a subjective quality evaluation experiment is conducted to compare the quality of these images compressed by JPEG with different quantization tables based on a two-alternative-forced-choice (2AFC) method. This method is widely used in psychophysical studies [20], where a subject is shown a pair of images and asked to choose the one with better quality based on his/her experience in each trial. In the subjective experiment, we compare the bit rates of compressed images with similar visual quality. Fifteen subjects were invited to compare the quality of images, and each pair is repeated four times with random order. Therefore, 40 2AFC results are obtained for each subject. The statistical results are shown in Tabel II, where the score corresponds to the percentage regarding the favor of the images compressed by JPEG and the proposed method. According to the results in Table II, the proposed method achieves about 18.5% bit rate savings compared with JPEG recommended quantization table, when the images compressed based on the proposed quantization table have the same visual quality or even better quality compared with JPEG standard images.

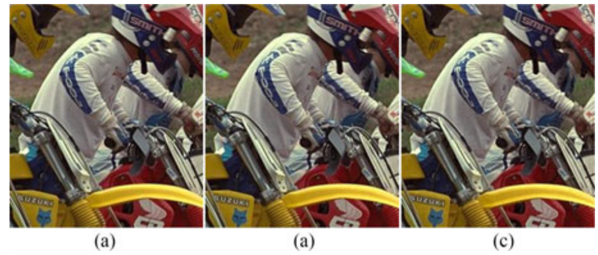


Fig. 3. The subjective quality comparison for *Bike* compressed by JPEG with different quantization tables (cropped for better visualization). (a) JPEG; (b) PSY-Table; (c) The proposed method.

TABLE II
BIT-RATE REDUCTION FOR IMAGES COMPRESSED BY JPEG DIFFERENT QUANTIZATION TABLES

| Images | JPEG | | Proposed | | bit-rate saving |
|---|---|---|---|---|---|
| | Score | bpp | Score | bpp | |
| Hat | 60.0% | 0.608 | 40.0% | 0.495 | 18.6% |
| Motor | 45.0% | 2.167 | 55.0% | 1.722 | 20.6% |
| Boat | 48.3% | 1.266 | 51.7% | 1.035 | 18.3% |
| Alfred | 21.7% | 2.271 | 78.3% | 1.842 | 18.9% |
| Wharf | 40.0% | 1.551 | 60.0% | 1.257 | 18.9% |
| Mountain | 41.7% | 1.853 | 58.3% | 1.479 | 20.2% |
| Statue | 45.0% | 0.612 | 55.0% | 0.525 | 14.2% |
| Country | 43.3% | 1.860 | 56.7% | 1.421 | 23.6% |
| Lighthouse | 76.7% | 0.694 | 23.3% | 0.592 | 14.8% |
| Ridinghome | 48.3% | 0.900 | 51.7% | 0.744 | 17.4% |
| Overall | 47.0% | 1.378 | 53.0% | 1.111 | 18.5% |

## V. CONCLUSION

In this letter, we have proposed a JND-based JPEG quantization table optimization method. The DCT domain JND value is utilized to threshold the traditional mean square errors to model the perceived distortions by HVS. The optimal quantization table is derived by minimizing the rate-distortion costs for all the bands by using Lagrange optimization techniques, and the Lagrange multipliers are estimated from every frequency band individually according to the change rate of distortions with respect to coding rates. Both objective and subjective experiments verify that significant coding gains can be achieved based on the quantization table generated by the proposed method.

## REFERENCES

[1] G. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consum. Electron.*, vol. 38, no. 1, pp. xviii–xxxiv, Feb. 1992.

[2] F. Ernawan and S. H. Nugraini, "The optimal quantization matrices for JPEG image compression from psychovisual threshold," *J. Theoretical Appl. Inf. Technol.*, vol. 70, no. 3, 2014.

[3] V. Ratnakar and M. Livny, "An efficient algorithm for optimizing DCT quantization," *IEEE Trans. Image Process.*, vol. 9, no. 2, pp. 267–270, Feb. 2000.

[4] S. Hu, H. Wang, and C.-C. J. Kuo, "A GMM-based stair quality model for human perceived JPEG images," in *Proc. 2016 IEEE Int. Conf. Acoust., Speech Signal Process.*, 2016, pp. 1070–1074.

[5] H. Bai, W. Lin, M. Zhang, A. Wang, and Y. Zhao, "Multiple description video coding based on human visual system characteristics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 8, pp. 1390–1394, Aug. 2014.

[6] Z. Chen and H. Liu, "JND modeling: Approaches and applications," in *Proc. 2014 19th Int. Conf. Digit. Signal Process.*, Aug. 2014, pp. 827–830.

[7] L. Ma, K. N. Ngan, F. Zhang, and S. Li, "Adaptive block-size transform based just-noticeable difference model for images/videos," *Signal Process., Image Commun.*, vol. 26, no. 3, pp. 162–174, Mar. 2011.

[8] S.-H. Bae and M. Kim, "A novel DCT-Based JND model for luminance adaptation effect in DCT frequency," *IEEE Signal Process. Lett.*, vol. 20, no. 9, pp. 893–896, Sep. 2013.

[9] X. Zhang, R. Xiong, S. Ma, and W. Gao, "Reducing blocking artifacts in compressed images via transform-domain non-local coefficients estimation," in *Proc. 2012 IEEE Int. Conf. Multimedia Expo.*, Jul. 2012, pp. 836–841.

[10] X. Zhang, R. Xiong, X. Fan, S. Ma, and W. Gao, "Compression artifact reduction by overlapped-block transform coefficient estimation with block similarity," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4613–4626, Dec. 2013.

[11] X. Zhang, R. Xiong, W. Lin, S. Ma, J. Liu, and W. Gao, "Video compression artifact reduction via spatio-temporal multi-hypothesis prediction," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6048–6061, Dec. 2015.

[12] Z. Wei and K. Ngan, "Spatio-temporal just noticeable distortion profile for grey scale image/video in DCT domain," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 3, pp. 337–346, Mar. 2009.

[13] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *J. Opt. Soc. Amer.*, vol. 70, no. 12, pp. 1458–1471, 1980.

[14] X. Zhang, W. Lin, and P. Xue, "Just-noticeable difference estimation with pixels in images," *J. Vis. Commun. Image Representation*, vol. 19, no. 1, pp. 30–41, Jan. 2008.

[15] G. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74–90, Nov. 1998.

[16] X. Zhang, R. Xiong, S. Ma, and W. Gao, "New image coding scheme with hierarchical representation and adaptive interpolation," in *Proc. 2011 Vis. Commun. Image Process.*, Nov. 2011, pp. 1–4.

[17] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-motivated rate-distortion optimization for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 4, pp. 516–529, Apr. 2012.

[18] Z. Pan, Y. Zhang, and S. Kwong, "Efficient motion and disparity estimation optimization for low complexity multiview video coding," *IEEE Trans. Broadcast.*, vol. 61, no. 2, pp. 166–176, Jun. 2015.

[19] Z. Pan, J. Lei, Y. Zhang, X. Sun, and S. Kwong, "Fast motion estimation based on content property for low-complexity H. 265/HEVC encoder," *IEEE Trans. Broadcasting*, vol. 62, no. 3, pp. 675–684, Sep. 2016.

[20] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "Perceptual video coding based on SSIM-inspired divisive normalization," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1418–1429, Apr. 2013.