Analysis of the JND-Suppression Effect in Quantization Perspective for HEVC-based Perceptual Video Coding

Jaeil Kim and Munchurl Kim

Department of Information and Communications, KAIST, South Korea {jaeil1203, mkimee}@kaist.ac.kr

***** Corresponding Author: Jaeil Kim, Munchurl Kim

Received April 5, 2014; Revised June 13, 2014; Accepted November 19, 2014; Published February 28, 2015

** Regular Paper*

***Abstract*:** Transform-domain JND (Just Noticeable Difference)-based for PVC (Perceptual Video Coding) is often performed in quantization processes to effectively remove perceptual redundancy. This study examined the JND-suppression effects on quantized coefficients of transform in HEVC (High Efficiency Video Coding). To reveal the JND-suppression effect in quantization, the properties of the floor functions were used for modeling the quantized coefficients, and a JND-adjustment process in an HEVC-compliant PVC scheme was used to tune the JND values by analyzing the JND suppression effect. In the experimental results, the bitrate reduction decreases slightly, but the PSNR and perceptual quality are improved significantly when the proposed JND adjustment process is applied.

***Keywords*:** Perceptual video coding, Quantization, HEVC

1. Introduction

Recently, the standardization of next-generation video coding, called High Efficiency Video Coding (HEVC) [1], has been finalized with the goal of improving its coding efficiency more than 2 times compared to that of H.264/AVC [2]. To improve the coding performance, the following advanced techniques have newly been adopted for HEVC: a flexible structure of the Coding Unit (CU) with skip modes; squared and asymmetric prediction units (PUs) with advanced motion vector prediction and merge modes; a residual quad-tree transform structure with 4×4, 8 ×8, 16× 16, and 32×32; 33 directional intra prediction with planar and DC modes; Transform Skipping Mode (TSM) for 4×4 inter and intra prediction mode; and additional in-loop filtering of sample adaptive offset [1]. Fig. 1 shows the CU, PU and TU partition structures of HEVC. On the other hand, such efforts toward improving the coding efficiency have led to dramatic increases in the encoder complexity, and further improvement of the coding efficiency becomes increasingly difficult in a rate-distortion (MSE: Mean Squared Error) sense, even with more elaborate coding tools at the price of complexity increases.

Another axis toward coding efficiency improvement is perceptual video coding (PVC), where perceptual distortion can be incorporated instead of MSE. Therefore, considerable effort has been made in this direction in the sense of PVC [3-7]. In PVC, how to effectively incorporate visual perception models into the coding process is one of the most important issues in reducing the perceptual redundancy. Among the techniques for reducing the perceptual redundancy, JND is an efficient model to assess the perceptual redundancy of human visual systems (HVS) [3-7]. In recent years, several methods have focused on PVC based on HVS [3-7]. In [3], a JND-adaptive residue pre-processor in the image-domain was proposed for a motion-compensated residue signal with a nonlinear additive model to improve the compression performance. *Chen et al.* [5] introduced a foveation JND model to improve the existing spatial and temporal JND models. The proposed JND model is used for macroblock quantization adjustment in H.264/AVC.

Naccari *et al.* [6] reported a decoder-side JND model estimation to avoid additionally required signal bits for the adjustment of the quantization matrices. Based on this, they proposed a H.264/AVC-based perceptual video coding architecture. In [7], a perceptual visual model was applied for the earliest reference software version of HEVC, called Test Model under Consideration (TMuC), to improve the coding performance. On the other hand, all previous methods are only applied for fixed and small-

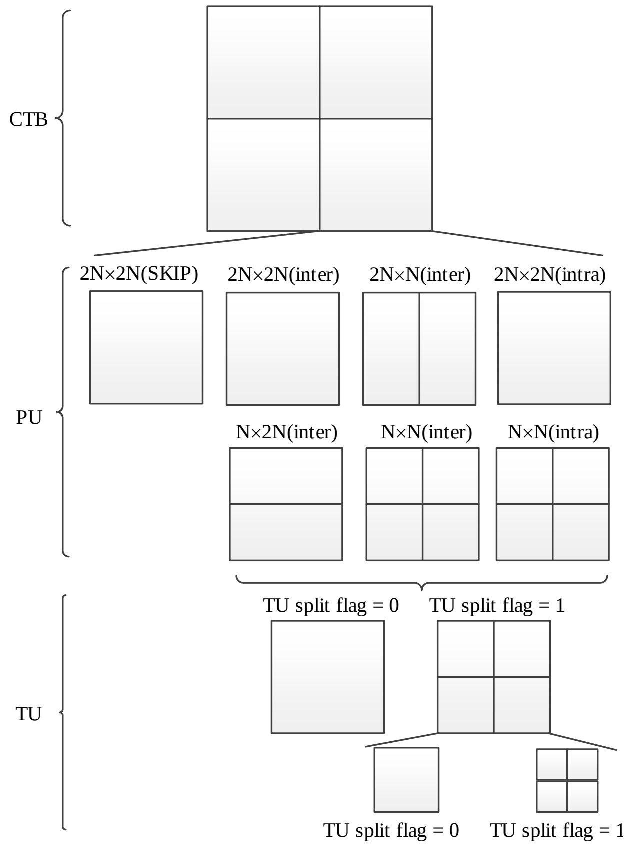


Fig. 1. CU, PU and TU Partition Structures.

sized transforms (4x4 and 8x8). Luo *et al.* [8] proposed a PVC scheme determining the optimal quantization levels limited in JND error. Nevertheless, to find the optimal quantization levels, this PVC scheme required a recursive structure that involved significant computational complexity. A previous study [9] proposed a JND-based HEVC-compliant PVC scheme with simple computation complexity. On the other hand, the effects of JND suppression and quantization errors were not investigated for human perception. Therefore, this study analyzed the JND suppression with a quantization error limited in JND and modified the proposed JND value.

The remainder of this paper is organized as follows: Section 2 briefly describes the previous JND-based HEVC-compliant PVC scheme [9] and proposes a new PVC scheme with a JND-suppression adjustment taking the effects of JND suppression and quantization in Section 3 into account. Section 4 presents the performance evaluations and discussion of the effect of the proposed scheme. Finally, this paper is concluded in Section 4.

2. The proposed HEVC-Compliant PVC scheme

To consider the quantization and JND suppression effects, a simple model of quantization process was defined as follows:

$$\left|l_{n,i,j}\right| = \left\lfloor \left|w_{n,i,j}\right| / q + f \right\rfloor \tag{1}$$

where $w_{n,i,j}$ and $l_{n,i,j}$ are the transformed and quantized transform coefficients at (i, j) in block n, respectively, and q and f are the quantization step size and rounding offset value for quantization process, respectively, and $\lfloor\ \rfloor$ is the floor function (called integer operation). HEVC and H.264/AVC use a shift operation for quantization process that are different from Eq. (1) to avoid division operation which causes mismatch between quantization processes for the encoder and decoder, and has heavy computation complexity. In this paper, the simple quantization model in Eq. (1) was used to analyze the effects from the quantization process and JND suppression for simple derivation.

2.1 HEVC-Compliant Transform-domain JND suppression

The PVC schemes can be categorized into two: standard-incompliant PVC schemes and standard-compliant PVC schemes. The standard-incompliant PVC schemes have been applied for quantization and inverse quantization in encoder processing and inverse quantization in decoder processing. Therefore, the related works for standard-incompliant PVC schemes require the modification of encoders and decoders to reduce the perceptual redundancy in quantization. On the other hand, the related works for standard-compliant PVC schemes have only been applied for quantization in encoder processing. The standard-incompliant PVC schemes have more freedom to reduce the bitrate, but cannot be applicable in real applications of the video coding standard. The standard-compliant PVC schemes can be applicable in actual video coding encoders in a compliant manner with video coding standards, such as H.264/AVC and HEVC.

Among the related works, Luo's PVC scheme [8] proposed a standard-compliant JND-suppression scheme by modifying the quantization process only in the encoder sides. In this scheme, the quantized coefficients, $l'_{n,i,j}$ were suppressed in the direction by JND as

$$\left|l'_{n,i,j}\right| = \left|l_{n,i,j}\right| - h_{n,i,j} \tag{2}$$

where the tuning value, $h_{n,i,j}$, is derived from the following equation:

$$\begin{aligned} & h_{n,i,j} = \max h \\ & s.t. 0 \le h \le \left|l_{n,i,j}\right|, h \in Z \\ & e_{n,i,j}(h) / T_{n,i,j} \le JND_{n,i,j} \end{aligned} \tag{3}$$

In Eq. (3), $T_{n,i,j}(h)$ is the probability of detecting distortion based on a visual saliency map and $e_{n,i,j}(h)$ is the error function between transformed coefficient and JND-suppressed reconstruction coefficient in terms of the tuning value, h, for the quantized coefficient. This means that Luo's PVC scheme reduces the quantized coefficient until the error by quantization and JND-suppression are the lowest, and can control error less than the JND value. On the other hand, Luo's PVC scheme needs to find a quantized JND-suppression level recursively, requiring high computational complexity.

To solve this problem, a previous work [9] proposed a JND-suppression method as follows:

$$\left|l'_{n,i,j}\right| = \begin{cases} \left\lfloor \dfrac{\left(\left|w_{n,i,j}\right| - JND^{scale}_{n,i,j}\right)}{q} + f \right\rfloor, & \left|w_{n,i,j}\right| > JND^{scale}_{n,i,j} \\ 0, & otherwise \end{cases} \tag{4}$$

where $JND^{scale}_{n,i,j}$ is a scaled JND value to consider a normalization factor for the HEVC transform in [9]. In Eq. (4), subtraction operation is also needed for the PVC-based quantization process. Nevertheless, perceptual distortion can be detected when the error by quantization and JND-suppression is larger than a JND value. Therefore, these effects are considered in the following chapter.

3.2 Analysis of JND suppression in terms of the quantization effects

To consider the error effect by quantization and JND suppression and improve the previous PVC scheme, this section presents the proposed HEVC-compliant JND-suppression PVC scheme by considering the JND suppression effect with quantization error, as shown in Fig. 2.

The proposed JND adjustment process for a transform-domain JND model tunes the quantized level according to its JND value, quantization step size and rounding offset. Firstly, a lower bound is defined as follows:

$$\left|\hat{w}'_{n,i,j}\right| \le \left|\hat{w}_{n,i,j}\right| - JND_{n,i,j} \tag{5}$$

where $\hat{w}_{n,i,j}$ and $\hat{w}'_{n,i,j}$ are the reconstruction coefficient and JND-suppressed reconstruction coefficient with JND adjustment. In the right-hand side, the reconstruction coefficient is used rather than the transformed coefficient because the purpose of the proposed PVC scheme is to reduce the difference between the JND-suppressed reconstruction coefficient and reconstruction coefficient to a level virtually undetectable by the human visual system.

From Eq. (1), Eq. (5) can be redefined as follows:

$$q\left\lfloor \left(\left|w_{n,i,j}\right| - \left(JND_{n,i,j} - \alpha\right)\right)/q + f \right\rfloor \le \left|\hat{w}_{n,i,j}\right| - JND_{n,i,j} \tag{6}$$

where α is the tuning value to adjust the JND value, $JND_{n,i,j}$. To consider the rounding offset, f, for the transformed coefficient, $w_{n,i,j}$, and JND value, $JND_{n,i,j}$, in Eq. (6), Eq. (7) can be defined as

$$q\left\lfloor \left(\left|w_{n,i,j}\right| - \left(JND_{n,i,j} - \alpha'\right)\right)/q + 2f \right\rfloor \le \left|\hat{w}_{n,i,j}\right| - JND_{n,i,j} \tag{7}$$

where α' can be defined in

$$\alpha' = \alpha - f \cdot q \tag{8}$$

To find the optimal α, the floor function in [10] can be used as follows:

$$\lfloor x+y+z \rfloor \le \lfloor x \rfloor + \lfloor y \rfloor + \lfloor z \rfloor + 2 \tag{9}$$

for real numbers, x, y, z. From Eq. (9), Eq. (7) can be rewritten as follows:

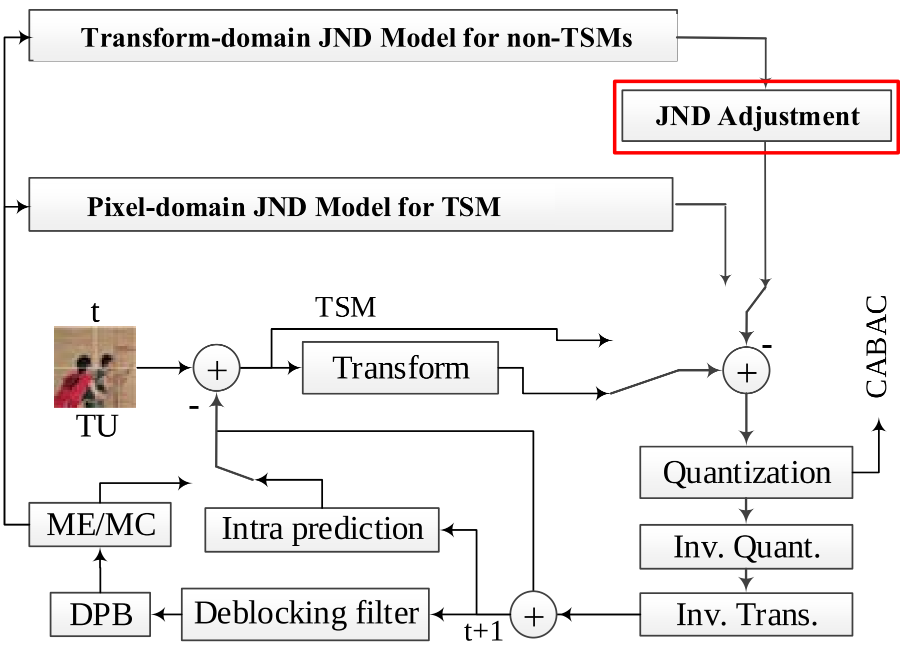


Fig. 2. Proposed PVC scheme considering JND suppression effect.

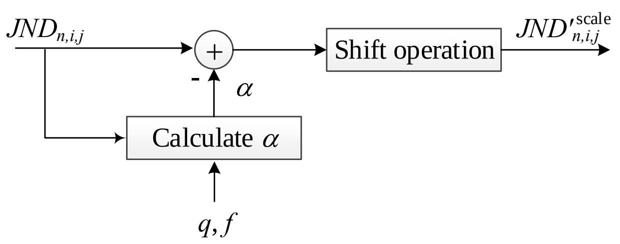


Fig. 3. Proposed JND adjustment process.

$$q\left\lfloor \left(\left|w_{n,i,j}\right| - \left(JND_{n,i,j} - \alpha'\right)\right)/q + 2f \right\rfloor$$
$$\le q\left\lfloor \left|w_{n,i,j}\right|/q + f \right\rfloor - q\left\lfloor \left|JND_{n,i,j}\right|/q + f \right\rfloor + q\left\lfloor \alpha'/q \right\rfloor + 2q$$
$$\le \left|\hat{w}_{n,i,j}\right| - JND_{n,i,j} \quad (10)$$

and

$$\left|\hat{w}_{n,i,j}\right| - J\hat{N}D_{n,i,j} + q\left\lfloor \alpha'/q \right\rfloor + 2q \le \left|\hat{w}_{n,i,j}\right| - JND_{n,i,j} \quad (11)$$
$$-J\hat{N}D_{n,i,j} + +q\left\lfloor \alpha'/q \right\rfloor + 2q \le -JND_{n,i,j} \quad (12)$$

In Eqs. (11) and (12), $J\hat{N}D_{n,i,j}$ is the reconstructed JND value. Finally,

$$k \le J\hat{N}D_{n,i,j}/q - JND_{n,i,j}/q - 2q \quad (13)$$

where k is defined in

$$k = \left\lfloor \alpha'/q \right\rfloor \quad (14)$$

From Eqs. (8) and (14), the range of α can be derived from the definition of floor function in [10] as follows:

$$k \le \alpha'/q < k + 1, \quad (15)$$

and

$$qk + qf \le \alpha < kq + q + fq\ . \quad (16)$$

From Eq. (13), the optimal k can be found first as

$$k = \left\lfloor J\hat{N}D_{n,i,j}/q - JND_{n,i,j}/q - 2q \right\rfloor \quad (17)$$

From Eq. (16), α can be found as the smallest value as

$$\alpha = \left\lceil qk + qf \right\rceil \quad (18)$$

where $\lceil\ \rceil$ is the ceiling function. As shown Fig. 3, the proposed JND adjustment process uses the JND value, $JND_{n,i,j}$, quantization step size, q, and rounding offset, f, to derive α in Eq. (15) for tuning the value for JND.

3. Experimental Results

To evaluate the performance improvement achieved by the proposed JND-suppression adjustment scheme, the proposed model in Eqs. (14) and (15) were implemented into the HM 11.0 [11] and the JND-based HEVC-compliant PVC scheme with JND-suppression adjustment process was compared with the original HM 11.0 and the JND-based HEVC-compliant PVC without a JND-suppression adjustment in the TU blocks. For the experiments, four FHD (Full High Definition – 1920×1080) test sequences of the *ParkScene* (*PkS*), *Cactus* (*Ca*), *BQTerrace* (*BQT*) and *BasketballDrive* (*BDv*) and four FWVGA (Full Wide Video Graphics Array – 832×480) test sequences of *RaceHorses* (*RH*), *BQMall* (*BQM*), *PartyScene* (*PtS*) and *BasketballDrill* (*BDl*) were used, which are all in 4:2:0 color format. As the test conditions, the default Low-Delay (LD) main configuration [11] and the first 100 frames were used for all experiments.

The original HM11.0, PVC (PVC without JND-suppression adjustment process) and PVC_{AP} (PVC with JND-suppression adjustment process) were compared with respect to PSNR, bitrate reduction and encoding time. The encoding times were measured on a PC with an Inter 2.5GHz HexaCore processor with 32GB RAM. The bitrate reduction (*ΔBitrate*), the encoding time (*ΔTime*) and the PSNR difference (*ΔPSNR*) were calculated as

$$\Delta Bitrate = \frac{Bitrate_{pro} - Bitrate_{ref}}{Bitrate_{ref}} \times 100 \quad (16)$$

$$\Delta Time = \frac{Time_{Pro} - Time_{ref}}{Time_{ref}} \times 100 \quad (17)$$

$$\Delta PSNR = PSNR_{pro} - PSNR_{ref} \quad (18)$$

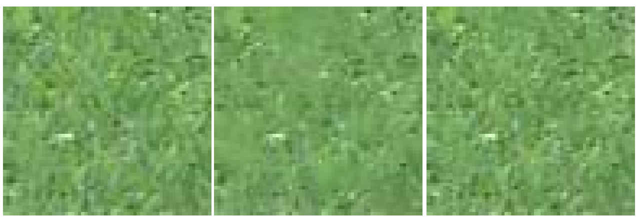
Table 1 presents the performance comparison for PSNR and bitrate reduction between PVC and PVC_{AP} under the LD configuration. As shown in Table 1, the average bitrate reduction was 23.66% for PVC and 18.53% for PVC_{AP}, compared to the original HM 11.0. This suggests that the PVC outperforms PVC_{AP} with 5.13% more bitrate reduction but PVC_{AP} outperforms PVC with a 0.39dB PSNR increase. As shown in Table 1, PVC_{AP} increases the encoder complexity by 29.9% in execution compared to the HM 11.0 whereas PVC increases 22.7 %, which has similar computation complexity.

In terms of perceptual verification, however, PVC_{AP} can improve the visual quality significantly, as shown in Figs. 4 and 5. Fig. 4(a) shows the 51st frame of the original *RaceHorses* sequence with an image region (some grass) in the bounded box. Figs. 4(b)-(d) show some image regions of the reconstructed frames by the original HM 11.0, PVC and PVC_{AP}, respectively. From a comparison of the visual quality, the proposed PVC_{AP} produces fewer visual artifacts in the reconstructed region of some grass compared to the PVC. Also, Fig. 5(a) shows the 30st frame of the original *BQMall* sequence with an image region (back of one's head) in the bounded box. Figs. 5(b)-(d) show some image regions of the reconstructed frames by

Table 1. Experimental Results.

Seq.	QP	ΔBitrate(%)		ΔPSNRY(dB)		ΔTime	
		PVC	PVC_{AP}	PVC	PVC_{AP}	PVC	PVC_{AP}
PkS	22	-33.73	-32.75	-2.75	-2.62	23.0	26.9
	27	-23.12	-16.33	-1.73	-1.36	32.6	36.9
	32	-9.00	-7.08	-0.74	-0.58	38.1	40.8
Ca	22	-46.77	-48.37	-2.00	-1.89	13.0	14.9
	27	-19.74	-13.63	-1.32	-0.93	30.3	32.1
	32	-5.76	-5.76	-0.71	-0.47	36.5	37.4
BQT	22	-65.33	-64.13	-2.82	-2.73	7.7	8.5
	27	-40.86	-18.78	-1.12	-0.64	29.2	29.4
	32	-12.09	-7.59	-0.45	-0.32	40.1	37.4
BDv	22	-36.06	-38.09	-1.97	-1.88	23.0	22.6
	27	-15.57	-12.16	-1.46	-1.07	32.2	34.1
	32	-6.57	-3.65	-0.98	-0.56	38.3	38.9
RH	22	-35.84	-34.71	-4.27	-4.01	8.6	19.8
	27	-25.86	-2.67	-2.65	-1.06	16.9	30.8
	32	-5.92	-1.68	-0.91	-0.60	22.7	36.2
BQM	22	-28.34	-27.50	-3.54	-3.32	10.7	22.1
	27	-20.42	-8.99	-2.46	-1.67	18.8	31.8
	32	-7.82	-2.97	-1.30	-0.84	22.8	36.4
PtS	22	-32.45	-30.29	-4.61	-4.31	5.3	17.4
	27	-25.33	-11.08	-2.71	-1.72	16.0	28.2
	32	-12.29	-3.49	-1.13	-0.61	23.6	35.5
BDl	22	-17.40	-23.63	-2.44	-2.31	12.6	26.3
	27	-18.00	-10.96	-1.40	-0.97	19.1	34.9
	32	-18.50	-3.54	-0.55	-0.29	23.9	39.5
Avg	-	-23.66	-18.53	-1.98	-1.59	22.7	29.9

(a)



(b) (c) (d)

Fig. 4. Comparisons of subjective qualities at QP27 (a) original 51st frame of the *RaceHorses* sequence, (b) the reconstructed region of the grass by the original HM 11.0, (b) the reconstructed region of the grass by the PVC, (c) the reconstructed region of the grass by the PVC_{AP}.

(a)

(b) (c) (d)

Fig. 5. Comparisons of subjective qualities at QP27 (a) original 30st frame of the *BQMall* sequence, (b) the reconstructed region of the back of one's head by the original HM 11.0, (b) the reconstructed region of the back of one's head by the PVC, (c) the reconstructed region of the back of one's head by the PVC_{AP}.

the original HM 11.0, PVC and PVC_{AP}, respectively. From a comparison of the visual quality, the proposed PVC_{AP} produces fewer ringing artifacts in the reconstructed region of the back of one's head compared to the PVC.

4. Conclusion

This study examined the effects of JND suppression in terms of the quantization effect in the transform domain, and proposed a JND value adjustment method for a JND-based HEVC-compliant PVC scheme. The bitrate reduction decreased in some QP ranges but the PSNR and visual quality were improved significantly.

Acknowledgement

This study was supported by the IT R&D program of MSIP/IITP (10039199, A Study on Core Technologies of Perceptual Quality based Scalable 3D Video Codecs), Korea.

References

[1] G. J. Sullivan, J.-R. Ohm, W.-J. Han, T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard." *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012. Article (CrossRef Link)

[2] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards – Including High Efficiency Video Coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669-1684, Dec. 2012. Article (CrossRef Link)

[3] Y. Xiaokang, et al., "Motion-compensated residue preprocessing in video coding based on just-noticeable-distortion profile," *IEEE Trans. Circuits Syst. Video Technol.,* vol.15, no. 6, pp. 742-752, Jun. 2005. Article (CrossRef Link)

[4] W. Zhenyu and K. N. 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. Article (CrossRef Link)

[5] 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, Jun. 2010. Article (CrossRef Link)

[6] M. Naccari and F. Pereira, "Advanced H.264/AVC-Based Perceptual Video Coding: Architecture, Tools, and Assessment," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 6, pp. 766-782, Jun. 2011. Article (CrossRef Link)

[7] M. Naccari and F. Pereira, "Integrating a spatial just noticeable distortion model in the under development HEVC codec," *Proc IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 817-820, Prague, Czech Republic, May 2011. Article (CrossRef Link)

[8] Z. Luo, L. Song, S. Zheng, and N. Ling, "H.264/Advanced Video Control Perceptual Optimization Coding Based on JND-Directed Coefficient Suppression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 6, pp. 935-948, Jun. 2013. Article (CrossRef Link)

[9] J. Kim, M. Kim, "An HEVC-Compliant Perceptual Video Coding Scheme based on JND Models for Variable Block-sized Transform Kernels", *IEEE Trans. Circuits Syst. Video Technol.*, to be published, 2015. Article (CrossRef Link)

[10] B. M. Stephen, R. Anthony, Discrete Algorithmic Mathematics, Third Edition, CRC Press, 2005. Article (CrossRef Link)

[11] HM reference software 11.0. Article (CrossRef Link)

Jaeil Kim received his B.S. degree in electrical engineering from the Korea University of Technology and Education, Cheonan, Korea, in 2006. He received a Ph.D. degree in Department of Information and Communications Engineering of Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea. His current research interests include perceptual video coding, hardware friendly video codec design, transform coding for video, image/video processing, and pattern recognition.

Munchurl Kim received his B.E. degree in electronics from Kyungpook National University, Daegu, Korea, in 1989, and has M.E. and Ph.D. degrees in electrical and computer engineering from the University of Florida, Gainesville, in 1992 and 1996, respectively. After his graduation, he joined the Electronics and Telecommunications Research Institute, Daejeon, Korea, as a Senior Research Staff Member, where he led the Realistic Broadcasting Media Research Team. In 2001, he was an Assistant Professor with School of Engineering, Information and Communications University, Daejeon, Korea. Since 2009, he has been an Associate Professor and is now a Full Professor of Department of Electrical Engineering, KAIST, Daejeon, Korea. His current research interests include 2D/3D perceptual video coding, visual perception modeling on 2D/3D video, high dynamic range video processing, machine learning, and pattern recognition.