# RATE-DISTORTION OPTIMISED QUANTISATION FOR HEVC USING SPATIAL JUST NOTICEABLE DISTORTION

*André S. Dias[1], Mischa Siekmann[2], Sebastian Bosse[2], Heiko Schwarz[2], Detlev Marpe[2], Marta Mrak[1]*

[1]British Broadcasting Corporation – Research and Development, UK
[2] Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute, Germany

## ABSTRACT

Due to the higher requirements associated with Ultra High Definition (UHD) resolutions in terms of memory and transmission bandwidth, the feasibility of UHD video communication applications is strongly dependent on the performance of video compression solutions. Even though the High Efficiency Video Coding (HEVC) standard allows significantly superior rate-distortion performances compared to previous video coding standards, further performance improvements are possible when exploiting the perceptual properties of the Human Visual System (HVS). This paper proposes a novel perceptual-based solution fully compliant with the HEVC standard, where a low complexity Just Noticeable Distortion model is used to drive the encoder's rate-distortion optimised quantisation process. This technique allows a simple and effective way to influence the decisions made at the encoder, based on the limitations of the HVS. The experiments conducted for UHD resolutions show average bitrate savings of 21% with no visual quality degradations when compared to the HEVC reference software.

***Index Terms*****—** Just Noticeable Distortion, Rate-Distortion Optimised Quantisation, Perceptual Video Compression, HEVC, UHD

## 1. INTRODUCTION

With the increasing popularity of Ultra High Definition (UHD) video and its emerging adoption in widely used video services, new challenges for storing and transmitting video arise. The requirements in terms of transmission bandwidth and storage capacity of UHD video content are significantly higher. Thus, the successful distribution of UHD video content is highly dependent on the performance of the video compression solutions supporting them.

The state-of-the-art High Efficiency Video Coding (HEVC) standard [1], also known as ITU-T recommendation H.265, developed by the Joint Collaborative Team on Video Coding (JCT-VC), is able to achieve remarkable video compression performances with respect to its predecessor, Advanced Video Coding (AVC) – H.264. However, in order to better accommodate the needs of more demanding video formats, higher compression efficiency can be achieved by exploiting the properties and limitations of the Human Visual System (HVS).

In the past decades, typical video coding solutions mainly focused on optimising compression efficiency according to the differences between the original and reconstructed pictures. The most popular and advanced video compression solutions typically run a Rate-Distortion Optimisation (RDO) algorithm at the encoder to select the best coding modes and other essential coding elements to build the encoded bitstream. Typically, the decisions during the RDO process are made by evaluating both the expected bitrate and the expected quality of the output video signal after reconstruction, measured according to the differences between the original and reconstructed frames.

Since the main objective of video communication systems is to present perceptually satisfying video information to the final user, it makes sense to optimise the compression efficiency of video compression solutions according to the perceptual properties of the HVS.

Many studies and experiments have been conducted in the past years aiming at better understanding the way humans perceive visual information. The concept of Just Noticeable Distortion (JND) is based on the assumption that the HVS shows different sensitivities to different types of visual information. Image characteristics such as spatial frequency, texture patterns and luminance variations play an important role in the way images are perceived by the human brain. JND models aim to quantify these differences and provide thresholds for image elements under which changes are not perceived by human viewers.

JND models are therefore a valuable asset when trying to adapt video coding solutions according to the perceptual properties of the HVS. In this paper, a novel technique to integrate the properties of the HVS into HEVC-based video compression solutions is proposed. This technique is based on a simple, yet effective JND model, which is used to improve the way choices are made by the Rate-Distortion Optimised Quantisation (RDOQ) tool used in the reference HEVC encoder.

Since the RDOQ process operates at the encoder without influencing the syntax of the bitstream, its operation does not have to be standardised. Therefore, the proposed technique can be easily integrated in any HEVC-based video compression solution without compromising the compliance with the standard. Furthermore, the adopted JND model was selected and adapted targeting

978-0-9928626-3-3/15/$31.00 ©2015 IEEE 

low computational complexity, allowing its smooth integration into an HEVC encoder without significant complexity increase.

The remainder of this paper is organised as follows. Section 2 gives an overview of the most relevant background work on JND models applied to video compression solutions. Section 3 describes the adopted JND model, including the constraints that led to its selection. Section 4 describes the proposed JND-driven RDOQ solution used to drive the encoder's decisions according to the limitations of human perception. Section 5 presents the performance results achieved by the proposed technique and finally Section 6 concludes this paper with some final remarks.

## 2. BACKGROUND WORK

The first advances made in exploiting the HVS properties using JND models were made for still images by Ahumada and Peterson in [2], where data from previous psychophysical experiments were used to define a model for visibility thresholds when using Discrete Cosine Transform (DCT) decomposition of images. Later, Watson [3] proposed the so called DCTune model, where the model described in [2] was improved by considering image dependent parameters, notably considering luminance and contrast masking effects. These models aimed to specify perceptually optimised quantisation matrices for JPEG image compression.

In 2005, Yang et. al. [4] proposed a method for pre-processing prediction residuals based on a pixel domain JND model introduced in [5]. The pixel domain JND model was used to reduce the prediction residual prior to the transform operation. This method was developed for the MPEG-2 TM5 encoder.

Later in 2009, Mak et. al. [6] proposed a similar suppression approach to the one in [4], but based on a transform domain JND model. The technique consisted of discarding the residual coefficients whose absolute values were lower than the JND thresholds. This technique was integrated on an H.264/AVC encoder.

Later, Chen et. al. [7] proposed a method for macroblock (MB) quantisation adjustment in H.264/AVC based on the pixel domain JND model in [5]. This JND model was combined with a foveation model to take into account both threshold visibility and visual eccentricity. The method was used at the MB level to select the optimal Quantisation Parameter (QP) and the Lagrangean multiplier in the RDO process according to the model.

In 2011, Naccari et. al. [8] proposed an H.264/AVC-based perceptual video codec using the JND model defined in [9] to adaptively select, at the encoder, the quantisation step of each transform coefficient. At the decoder, a method was proposed to predict the right quantisation step to use for inverse quantisation, to avoid additional signalling bitrate. Due to the required adaptation in the decoder operation, this technique is not compliant with the H.264/AVC standard. This technique was further extended to an HEVC video codec in [10].

In 2013, Naccari et. al. [11] proposed a new perceptual video coding tool used to adjust the quantisation step of each transform coefficient based on the HVS luminance masking effects. The technique was designed for an efficient transmission of the additional luminance masking parameters and low-complexity implementation.

More recently, in 2015, Kim et. al. [12] proposed a solution fully compliant with the HEVC standard where the model in [9] was adjusted to cope with the different transform sizes used in HEVC. The modified JND model is then used to lower and suppress the values of the transform coefficients before quantisation. An average bitrate reduction of around 16% with negligible subjective quality loss was reported.

This paper presents an alternative approach to integrate a simple spatial JND model in the encoding process of an HEVC encoder capable of significantly reducing the associated bitrates and preserving the output subjective quality. The complexity introduced by the proposed technique is very low, making it particularly suitable for UHD video formats.

## 3. ADOPTED SPATIAL JND MODEL

The proposed solution in this paper adopts a JND-model to modify the choices made at the encoder according to the limits of visual perception. The adoption of a low-complexity model was therefore essential to enable the proposed solution to be used in practical video compression applications. A brief description of the adopted JND model is given in this section.

For complexity reduction purposes, the model in [9] was selected and adapted to the different transform sizes allowed in HEVC using the method for the adaptation of the spatial summation effect in [13]. Nevertheless, since the proposed integration technique of the JND model into an HEVC encoder is model-independent, the selected model can be replaced by a more accurate and sophisticated model depending on the complexity restrictions of the target application.

For a given transform block $n$, the JND threshold, $T_{JND}(n,i,j)$, associated to the transform coefficient with indexes $(i,j)$ is defined as

$$T_{JND}(n,i,j) = T_B(n,i,j) \cdot F_{LM}(n) \cdot F_{CM}(n,i,j). \quad (1)$$

As seen in (1), the JND threshold $T_{JND}(n,i,j)$ is given by the product of a base threshold $T_B(n,i,j)$, a luminance masking factor $F_{LM}(n)$ and a contrast masking factor $F_{CM}(n,i,j)$. The following subsections briefly describe each of these components of the adopted JND model.

### 3.1. Base Threshold

The base threshold accounts for the different sensitivity of the HVS to distortions added to different spatial frequencies. For a given transform block size $N$, $T_B(n,i,j)$ is given by

$$T_B(n,i,j) = S(N) \cdot \frac{1}{\phi_i \phi_j} \cdot \frac{H(f_{i,j})^{-1}}{r+(1-r)\cdot \cos^2 \varphi_{ij}}. \quad (2)$$

where $H(f_{i,j})$ is the Contrast Sensitivity Function (CSF), $S(N)$ is the spatial summation effect, $\phi_j$ and $\phi_i$ are the DCT normalisation factors and the term $r + (1 - r) \cdot \cos^2 \varphi_{ij}$ accounts for the different sensitivity of the HVS regarding directionality. All the parameters in (2) were computed as in [9], with the exception of the CSF and $S(N)$. The adopted CSF is given by

$$H(f_{i,j}) = (1 - a + \frac{f_{i,j}}{f_0})e^{-\left(\frac{f_{i,j}}{f_0}\right)^p}, \tag{3}$$

where $f_{i,j}$ represents the spatial frequency, computed as in [9], and $f_0 = 1.7377$, $a = 1.0465$ and $p = 0.6937$ are the best fitting parameters to a CSF of this type, according to the experiments conducted in [14] for a dataset of 43 image patterns. The parameters used in [9] were not considered in this case since they were empirically estimated based on a fixed transform size experiment ($8 \times 8$). The $S(N)$ factor compensates for spatial summation, which accounts for the effect of having simultaneous distortions over a range of spatial frequencies in a given frame area. Similarly to [13], the spatial summation effect was modelled as

$$S(N) = N^{-\frac{2}{\lambda}}, \tag{4}$$

in order to adapt the base threshold to the transform size used. In (4), the parameter $\lambda$ was set to 1.873 according to the experiments conducted in [13].

### 3.2. Luminance Adaptation Factor

The luminance adaptation factor accounts for the fact that visibility thresholds depend on the average brightness level of a given block. The HVS is less sensitive to changes in brighter and darker backgrounds and therefore the visibility threshold in these conditions can be increased.

As in [9], for a given transform block $n$, the luminance adaptation factor is given by

$$F_{lum}(n) = \begin{cases} \frac{(60 - \bar{I})}{150} + 1, & \bar{I} \le 60 \\ 1, & 60 < \bar{I} < 170, \\ \frac{(\bar{I} - 170)}{425} + 1, & \bar{I} \ge 170 \end{cases} \tag{5}$$

where $\bar{I}$ denotes the average luminance intensity value of the pixels inside block $n$.

### 3.3. Contrast Masking Factor

The contrast masking factor accounts for the reduction of visual sensitivity in one visual component in the presence of another. Typically, distortions are more difficult to notice when introduced in areas where texture energy is high. Given this, a contrast masking factor is used to elevate the threshold of each coefficient in a given block depending on the texture characteristics of the visual content in this area.

For the purpose of computing $F_{CM}(n, i, j)$, the Canny edge detector [15] is first applied to the whole frame and for a given DCT transform block size $N$, each block is classified as a Plane, Edge or Texture block according to

$$Block\ type = \begin{cases} Plane, & \rho_{edge} \le \alpha \\ Edge, & \alpha < \rho_{edge} \le \beta, \\ Texture, & \rho_{edge} > \beta \end{cases} \tag{6}$$

where $\alpha$ and $\beta$ are empirically set to 0.1 and 0.2, respectively, and $\rho_{edge}$ is the density of edge pixels inside the block identified by the Canny edge operator. For a given coefficient with indexes $i$ and $j$ inside block $n$, the final elevation factor is given by

$$F_{CM}(n, i, j) = \begin{cases} 1, & for\ Plane\ or\ Edge \\ 2.25, & for\ (i^2 + j^2) \le 2N\ in\ Texture. \\ 1.25, & for\ (i^2 + j^2) > 2N\ in\ Texture \end{cases} \tag{7}$$

Contrarily to the contrast masking factor in [9], the term introduced following the Foley-Boynton [16] method was not considered in the proposed approach since it required the computation of the transform coefficients of the original frame, increasing this way the complexity of the overall solution.

## 4. JND-DRIVEN RATE-DISTORTION OPTIMISED QUANTISATION

The method to integrate the selected JND model into the reference HEVC encoder consists of modifying the RDOQ process according to the thresholds defined by the JND profile described in the previous section. In this section, a brief description of the RDOQ process is first given, followed by the description of the proposed modifications to turn it into a perceptually adjusted tool.

### 4.1. Rate-Distortion Optimised Quantisation

The RDOQ process [17] consists of optimising the choice of the level obtained after quantising a given transform coefficient, considering both the introduced distortion and the associated bitrate. When the RDOQ tool is not used, the nearest integer rounding rule is used by the reference HEVC encoder to round a given quantised coefficient to the nearest integer level, $L$. Even though this rounding process minimises the distortion introduced by quantisation, choosing a different quantised level may be beneficial when considering also the associated bitrate. Therefore, when RDOQ is enabled in the version of the reference HEVC software (HM 16.2) used in this paper, the levels $L$, $L - 1$ and 0 are considered and the mode that shows the lowest Rate-Distortion (RD) cost is selected. Figure 1 shows an example of the reconstructed values corresponding to the levels tested by the RDOQ process.

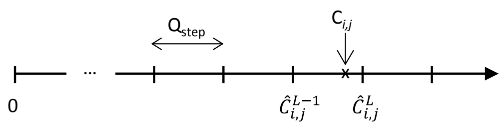


**Figure 1.** Candidates tested when using the RDOQ process to quantise a given transform coefficient $\boldsymbol{C_{i,j}}$.

The cost of each level tested by the RDOQ process, $J$, is computed according to

$$J = D_x + \lambda \cdot R_x, \tag{8}$$

where $D_x$ is the distortion introduced by the selection of a given candidate level $x$ (i.e, $L$, $L-1$ or 0), $\lambda$ is the Lagrangean multiplier and $R_x$ is the bitrate associated with each level being tested. In (8), the distortion, $D_x$, is the square of the error introduced by the quantisation process, $E_x$, given by

$$E_x = \left| C_{i,j} - \hat{C}_{i,j}^{x} \right|. \tag{9}$$

It is important to recall that the HEVC standard only specifies the syntax of the encoded bitstream and the decoding process. Thus, adjusting the quantised levels to minimise the RD cost is a decision made at the encoder and therefore any rule for selecting the quantised levels can be applied for this purpose without sacrificing compliance with the standard.

### 4.2. JND-Driven RDOQ

The JND profile defines a threshold for each transform coefficient that represents the maximum amount of distortion that can be added to a given coefficient without being perceived by the HVS. It is therefore possible to modify the value of $D_x$ according to this threshold and take into consideration the limitations of the HVS when computing the cost of each optimised level being tested.

Assuming that $T_{JND}(n, i, j)$ denotes the visibility threshold of the $(i, j)$th coefficient of a given transform block $n$, the proposed modified distortion, $D'_x$, to be used in cost computation of each candidate level, is computed based on a different error, $E_x{}'$, given by

$$E_x{}' = \begin{cases} 0, & if\ E_x \le T_{JND}(n,i,j) \\ E_x - T_{JND}(n,i,j), & if\ E_x > T_{JND}(n,i,j) \end{cases}. \tag{10}$$

In practice, replacing $D_x$ for $D'_x$ in the cost computation means that any distortion lower than that allowed by the JND threshold for the coefficient being quantised should be considered null in the RDOQ cost computation, since this distortion is not perceptually noticeable. In case this distortion is higher than the threshold, only the difference between these two values should be considered in the cost computation.

## 5. PERFORMANCE EVALUATION

Experiments were performed to assess the bitrate reduction capabilities of the proposed solution. The experiments were performed for the first 100 frames of 3 UHD test sequences and 3 HD test sequences. The experiments were conducted under Random Access conditions with the HEVC reference software HM 16.2 for four different QPs. The results of the proposed technique implemented on top of the reference software were compared with the reference software. In both cases, RDOQ was enabled. The results obtained are shown in Table 1. For all the results shown in Table 1, the decoded sequences were evaluated and no visual quality degradation was observed comparing with the decoded output of the HEVC reference software, despite the small PSNR losses.

The proposed JND-driven RDOQ technique is able to significantly reduce the bitrate for lower QPs in all sequences, especially for the three UHD sequences tested, where this reduction can go up to 62%. Higher reductions are expected in lower QPs since lower quantisation steps increase the number of cases where the quantisation error is lower than the JND threshold.

As expected, a small loss in terms of PSNR is introduced when using the proposed JND-driven RDOQ solution. Nonetheless, all decoded sequences were observed and no visual quality degradations were identified. Since the main target of the JND-driven RDOQ technique is to perceptually optimise the performance of the RDOQ decisions in an HEVC encoder, the PSNR loss is not as relevant as the subjective output video quality of the decoded sequences.

For higher quality RD points, the extra complexity introduced by the proposed technique is compensated by a reduction in the number of non-zero coefficients to encode, leading to even lower overall encoding times in the case of UHD sequences. For the remaining QPs, the overall additional complexity introduced for all sequences by the proposed technique is in general low (average encoding time penalty of 8%).

From the results in Table 1, it is clear that the proposed JND-driven RDOQ solution shows higher bitrate reduction capabilities when the target qualities are high. The solution is able to reduce the bitrates by reducing the amount of perceptually irrelevant visual information in the decoded sequences, providing the same output perceptual quality for significantly lower bitrate.

**Table 1.** JND-driven RDOQ performance.

| Sequence | QP | HM 16.2-RDOQ Bitrate [kb/s] | HM 16.2-RDOQ Y PSNR [dB] | JND-RDOQ Bitrate [kb/s] | JND-RDOQ Y PSNR [dB] | Bitrate saving | PSNR diff. [dB] | Enc. time diff. |
|---|---|---|---|---|---|---|---|---|
| Show Drummer 3840x2160 @ 60 Hz | 22 | 62823 | 38.27 | 26614 | 37.86 | -58% | -0.41 | -1% |
| | 27 | 8224 | 37.52 | 7446 | 37.47 | -9% | -0.05 | 9% |
| | 32 | 3827 | 36.92 | 3767 | 36.89 | -2% | -0.03 | 10% |
| | 37 | 2098 | 36.00 | 2082 | 35.98 | -1% | -0.02 | 10% |
| Homeless Sleeping 3840x2160 @ 60 Hz | 22 | 85844 | 37.38 | 32302 | 36.82 | -62% | -0.56 | -4% |
| | 27 | 8276 | 36.52 | 6277 | 36.48 | -24% | -0.04 | 8% |
| | 32 | 2810 | 36.06 | 2743 | 36.04 | -2% | -0.02 | 9% |
| | 37 | 1393 | 35.41 | 1380 | 35.40 | -1% | -0.01 | 10% |
| Young Dancers 1 3840x2160 @ 60 Hz | 22 | 61201 | 40.38 | 30206 | 39.22 | -51% | -1.16 | 0% |
| | 27 | 10726 | 38.74 | 7086 | 38.55 | -34% | -0.20 | 7% |
| | 32 | 3021 | 38.05 | 2821 | 38.02 | -7% | -0.04 | 9% |
| | 37 | 1623 | 37.29 | 1615 | 37.26 | 0% | -0.03 | 9% |
| BasketballDrive 1920x1080 @ 50 Hz | 22 | 17254 | 39.30 | 13502 | 38.93 | -22% | -0.36 | 4% |
| | 27 | 6071 | 37.70 | 5740 | 37.57 | -5% | -0.13 | 11% |
| | 32 | 2884 | 35.92 | 2829 | 35.84 | -2% | -0.07 | 12% |
| | 37 | 1537 | 33.97 | 1522 | 33.94 | -1% | -0.03 | 11% |
| BQTerrace 1920x1080 @ 60 Hz | 22 | 39832 | 37.99 | 26556 | 36.94 | -33% | -1.05 | 2% |
| | 27 | 10001 | 35.54 | 8489 | 35.32 | -15% | -0.22 | 9% |
| | 32 | 3654 | 33.79 | 3491 | 33.69 | -4% | -0.11 | 11% |
| | 37 | 1672 | 31.76 | 1650 | 31.71 | -1% | -0.05 | 11% |
| Cactus 1920x1080 @ 50 Hz | 22 | 20816 | 38.43 | 15924 | 37.96 | -24% | -0.47 | 6% |
| | 27 | 6791 | 36.75 | 6363 | 36.57 | -6% | -0.19 | 11% |
| | 32 | 3230 | 34.84 | 3159 | 34.74 | -2% | -0.09 | 13% |
| | 37 | 1675 | 32.65 | 1654 | 32.60 | -1% | -0.05 | 13% |

To further evaluate the performance of the proposed solution for higher qualities, an alternative perceptual quality metric was also used to evaluate the quality of the decoded sequences, in an attempt to have a more perceptually oriented evaluation. The selected metric to additionally evaluate the quality of the decoded sequences was the Video Quality Metric (VQM) [18], which is a standardised metric that according to [18] shows a better correlation with Mean Opinion Score (MOS) tests than PSNR. In contrast to PSNR, the lower the VQM value, the higher the quality of the sequence being evaluated. The results obtained are shown in Table 2.

**Table 2.** JND-driven RDOQ performance analysis for lower QPs using VQM.

| | HM 16.2-RDOQ | | | JND-RDOQ | | | Bitrate saving [%] | PSNR diff. [dB] | VQM diff. |
|---|---|---|---|---|---|---|---|---|---|
| | QP | Bitrate [kb/s] | VQM | QP | Bitrate [kb/s] | VQM | | | |
| Homeless Sleeping | 26 | 13743 | 0.0436 | 25 | 10976 | 0.0417 | -20% | -0.18 | -0.0019 |
| Show Drummer | 24 | 28920 | 0.9841 | 22 | 26614 | 0.9797 | -8% | -0.02 | -0.0044 |
| Young Dancers 1 | 22 | 61201 | 1.1772 | 20 | 54485 | 1.1764 | -11% | -0.72 | -0.0007 |

Similarly to the previous results presented in this section, negative values in the bitrate saving column represent bitrate reductions achieved by the JND-driven RDOQ with respect to the HEVC reference software. In the VQM difference column, negative values represent an increase of output video quality according to the VQM metric and negative values in the PSNR difference column represent a quality decrease in terms of PSNR.

From the VQM results in Table 2, it is possible to conclude that for these specific target qualities, the proposed JND-driven RDOQ technique is able to increase the decoded quality of the decoded sequences and, at the same time, reduce the bitrate up to 20% for the tested UHD sequences.

## 6. FINAL REMARKS

This paper presented a novel technique for integrating a JND model into an HEVC encoder, allowing a perceptually-oriented selection of the quantised levels by the RDOQ process of an HEVC encoder. The technique modifies the decisions made at the encoder side, meaning that a fully compliant bitstream is generated with the proposed solution. The results obtained show significant bitrate reductions with respect to the HEVC reference software, for the same perceived output visual quality, especially for UHD video content. The required extra complexity is very low, making this technique suitable for integration into any HEVC encoder.

## ACKNOWLEDGEMENTS

This work was supported by the Marie Skłodowska-Curie Initial Training Network PROVISION (PeRceptually Optimised VIdeo CompresSION) project.

## REFERENCES


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

[2] A. Ahumada and H. Peterson, "Luminance-model-based DCT quantization for color image compression," in *Human Vision, Visual Process., and Digital Display III*, San Jose, CA, USA, Aug. 1992.

[3] A. Watson, "DCTune: A technique for visual optimization of DCT quantization matrices for individual images," *Soc. for Information Display Digest of Technical Papers XXIV,* vol. 24, pp. 946-949, 1993.

[4] X. Yang, W. Lin, Z. Lu, E. Ong and S. Yao, "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.

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

[6] C.-M. Mak and K. N. Ngan, "Enhancing compression rate by just-noticeable distortion model for H.264/AVC," in *IEEE International Symposium on Circuits and Systems*, Taipei, Taiwan, May 2009.

[7] 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, Mar. 2010.

[8] 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, May 2011.

[9] 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, Feb. 2009.

[10] M. Naccari and F. Pereira, "Integrating a spatial just noticeable distortion model in the under development HEVC codec," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Prague, Czech Republic, May 2011.

[11] M. Naccari and M. Mrak, "Intensity dependent spatial quantization with application in HEVC," in *IEEE Int. Conf. on Multimedia and Expo*, San Jose, CA, USA, Jul. 2013.

[12] J. Kim, S. Bae and 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.,* vol. PP, no. 99, Jan. 2015.

[13] S. Bae and K. Munchurl, "A novel generalized DCT-based JND profile based on an elaborate CM-JND model for variable block-sized transforms in monochrome images," *IEEE Trans. Image Process.,* vol. 23, no. 8, pp. 3227 - 3240, Aug. 2014.

[14] A. Watson and A. Ahumada, "A standard model for foveal detection of spatial contrast," *Journal of Vision,* vol. 5, no. 9, pp. 717-740, Oct. 2005.

[15] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 8, no. 6, pp. 679 - 698, Nov. 1986.

[16] J. M. Foley and G. M. Boynton, "New model of human luminance pattern vision mechanisms: analysis of the effects of pattern orientation, spatial phase, and temporal frequency," in *Computational Vision Based on Neurobiology*, Park Grove, CA, USA, 1993.

[17] M. Karczewicz, Y. Ye and I. Chong, "Rate distortion optimized quantization," in *ITU-T SG16/Q6 VCEG, Doc. VCEG-AH21*, Antalya, Turkey, Jan. 2008.

[18] M. Pinson and S. Wolf, "A new standardized method objectively measuring video quality," *IEEE Trans. on Broadcast.,* vol. 50, no. 3, pp. 312-322, 2004.