

A comparative study of JPEG 2000, AVC/H.264, and HD Photo

Francesca De Simone^a, Mourad Ouaret^a, Frederic Dufaux^a, Andrew G. Tescher^b, Touradj Ebrahimi^a

^aEcole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

^bAGT Associates, CA, USA

ABSTRACT

In this paper, we report a study evaluating rate-distortion performance between JPEG 2000, AVC/H.264 High 4:4:4 Intra and HD Photo. A set of ten high definition color images with different spatial resolutions has been used. Both the PSNR and the perceptual MSSIM index were considered as distortion metrics. Results show that, for the material used to carry out the experiments, the overall performance, in terms of compression efficiency, are quite comparable for the three coding approaches, within an average range of $\pm 10\%$ in bitrate variations, and outperforming the conventional JPEG.

Keywords: Image Compression, JPEG 2000, AVC, H.264, HD Photo

1. INTRODUCTION

Despite an increase in the available storage capacity and communication channels bandwidth, there is still a need for compression of digital pictures in many professional and consumer applications. One of the reasons behind this need is that at the same time, progress in acquisition technologies now allows for higher resolution and better quality pictures, producing very large volumes of data if no compression is performed. Today, several alternatives for compression of digital pictures exist to choose from. Beside the well-known and largely deployed JPEG compression, other approaches include JPEG 2000, AVC/H.264, and the recently announced HD Photo.

HD Photo is a recent compression algorithm developed by Microsoft for digital imaging applications and under consideration for standardization by JPEG committee in JPEG XR [1] [2]. HD Photo is characterized by a block-based image compression scheme. The codec design aims at optimizing image quality and compression efficiency while at the same time requiring low-complexity in encoders and decoders implementations. As a result, even if it uses many of the same fundamental building blocks as in other traditional image and video compression schemes, e.g. color conversion, transform, quantization, coefficient scanning and entropy coding, different design objectives and taking advantage of the latest progress in the field of compression, have led to different solutions when compared to other state of the art. In particular, HD Photo uses a reversible Lapped Bi-orthogonal Transform (LBT) and an alternative coefficient coding approach.

The state of the art in technologies for compression of still pictures and video are found in the JPEG 2000 [3] and the Advanced Video Coding (AVC/H.264, MPEG-4 Part 10) [4]. Some researchers have suggested the use of AVC/H.264 Intra coding for efficient compression of still pictures.

JPEG 2000 is a wavelet-based compression standard for still images, but also used for compression of image sequences such as those in digital cinema. It has been designed by the Joint Photographic Experts Group (JPEG) committee who was also behind the original JPEG image compression. Beside outperforming the original DCT-based JPEG standard in terms of compression efficiency in many situations, JPEG 2000 offers a large number of features useful in multimedia applications.

AVC/H.264 is a video compression standard noted for providing excellent video quality at low bit rates. It has been developed by the ITU-T Video Coding Experts Group (VCEG) together with the ISO/IEC Moving Picture Experts Group (MPEG) as the product of a collective partnership effort known as the Joint Video Team (JVT). It is based on a block-based integer DCT transform. It performs spatial prediction for Intra frame coding and temporal motion estimation for Inter frame coding to improve the compression efficiency. In Intra frame coding, each frame is encoded on its own

without using any information from its neighbouring frames. Thus, in Intra frame coding, the AVC/H.264 could be used for compression of still images.

In [5], [6], [7] and [8] AVC/H.264 Intra and JPEG 2000 approaches have been compared for compression of video sequences. In [5], a performance evaluation of AVC/H.264 Main Profile (MP) Intra and JPEG 2000 was conducted. It is reported that AVC/H.264 MP Intra performs better than JPEG 2000 in terms of rate-distortion characteristics for low and intermediate resolution frames, with a gain in PSNR of around 0.5 ~ 2.0 dB. The same work indicates that JPEG 2000 performs better for higher resolution sequences, with a gain of around 0.5 ~ 1.0 dB in PSNR. AVC/H.264 High Profile (HP) Intra and JPEG 2000 were compared in [6] considering high resolution video sequences. The experimental results show that AVC/H.264 HP Intra offers rate-distortion gains of around 0.2 ~ 1.0 dB in PSNR over JPEG 2000 for the test data used in that study. In [7], JPEG 2000 has been compared to both AVC/H.264 Intra MP and HP profiles considering a set of video sequences with various spatial resolutions. It is shown that JPEG 2000 is comparable with AVC/H.264 HP Intra with around 0.1 dB differences in PSNR, in favor of AVC/H.264 HP Intra, for high spatial resolution video sequences. Moreover, JPEG 2000 outperforms AVC/H.264 MP Intra with gains of around 0.1~1.0 dB in PSNR. For intermediate and low spatial resolution video sequences, both profiles of AVC/H.264 Intra outperform JPEG 2000. Finally, [8] has compared the performance between JPEG 2000 and AVC/H.264 HP Intra considering a set of high definition video sequences and better parameter optimization for encoding when compared to [7]. Results show quite comparable performance between the two coding approaches, while in some cases AVC/H.264 HP Intra outperforms JPEG 2000.

In [10], [11] and [11] AVC/H.264 HP Intra and JPEG 2000 were compared for still image compression. The experimental results in [9] show that AVC/H.264 HP Intra outperforms JPEG 2000 in terms of PSNR considering monochromatic images. In particular, a similar comparison in [10] shows that, for monochromatic images, the performance of the two approaches are identical in term of PSNR, and that JPEG 2000 has a gain of around 1 dB in PSNR over AVC/H.264 HP Intra when its 8x8 transform is disabled, indicating the impact of this mode in such conditions. In [11], simulations were performed considering different objective metrics as well as subjective quality evaluations. As a result, it has been observed that the subjective image quality of AVC/H.264 Intra is superior to JPEG 2000. Regarding the objective quality, an advantage has been reported in favor of AVC/H.264 HP Intra in terms of PSNR in the range of 0.15 dB to 0.76 dB for monochromatic images and in the range of 0.02 dB to 0.30 dB for color images. The comparison results for monochromatic images in terms of Structural SIMilarity (SSIM) index [12] are also in favor of AVC/H.264 HP Intra. For color images, the results on the Y components are reported to be better for JPEG 2000 in terms of SSIM index measurement metric. Both in [10] and in [11] the evaluations have been performed on a small set of images, which could limit their conclusions, and consistency, in general situations.

Regarding HD Photo, at the best of our knowledge, comparison with state of the art technologies for compression of still images are provided in [13] and [14] only. Comparison of HD Photo performance with JPEG and JPEG 2000 in terms of PSNR and SSIM index are provided in [13]. According to these results, JPEG 2000 outperforms HD Photo both in PSNR and SSIM. In [14] HD Photo is compared to AVC/H.264 HP Intra in terms of PSNR considering still images. This work reports that AVC/H.264 HP Intra provides a PSNR performance comparable to HD Photo.

In this paper, the performance of HD Photo, JPEG 2000 and AVC/H.264 High 4:4:4 Intra Profile are evaluated for a set of ten high definition 4:4:4 color images. The paper is structured as follows. First, the compression algorithms are briefly described in Section 2. Then, the set of high definition images and the encoder parameters used are specified in Section 3. The results of the simulations in terms of PSNR on luminance component and in terms of Mean Structural SIMilarity (MSSIM) index [12] considering all components are presented in Section 4. Finally, we draw some concluding remarks in Section 5.

2. COMPRESSION ALGORITHMS

2.1 JPEG 2000

The JPEG 2000 standard makes use of the Discrete Wavelet Transform (DWT). It also supports a number of features such as multi-resolution representation, Region Of Interest (ROI) coding, error resilience and a flexible file format. The fundamental building blocks of JPEG2000 are: pre-processing, DWT, uniform quantizer with dead-zone, adaptive binary arithmetic coder, and bit stream organization [3].

In the pre-processing stage, an inter-component transformation is used to decorrelate the color data. There are two possible transforms. Both transforms operate on the first three components of an image tile with the implicit assumption that these components correspond to Red, Green, and Blue (RGB). One transform is the Irreversible Color Transform (ICT), which is identical to the traditional RGB to $YCbCr$ color transformation and can only be used for lossy coding. The other transform is the Reversible Color Transform (RCT), which is a reversible integer-to-integer transform that approximates the ICT for color decorrelation and can be used for both lossless and lossy coding. After color conversion, the DWT is applied to the processed samples. The DWT produces a multi-resolution image representation. Furthermore, it achieves good compression due to its energy compaction and the ability to de-correlate the image across a large scale. The resulting wavelet coefficients are quantized using a uniform quantizer with a central dead-zone. It is shown that this quantizer is optimal for a continuous signal with a Laplacian distribution such as DCT or wavelet coefficients. The coefficients are gathered in subbands. Each subband is partitioned into small rectangular blocks called codeblocks and each independently coded by an Adaptive Binary Arithmetic encoder. Finally, the output of the arithmetic encoder is organized as a compressed bit-stream which offers a significant degree of flexibility. This enables features such as random access, Region of Interest coding, and scalability. This flexibility is achieved partly through the various structures of components, tiles, subbands, resolution levels, and codeblocks. For more details about the JPEG 2000 standard, the reader can refer to [3].

2.2 AVC/H.264 Intra

AVC/H.264 Intra is based on the block-based integer DCT transform and takes advantage of the spatial correlation to improve the coding efficiency. The Intra coding of a macroblock consists in four main steps: spatial prediction, DCT transform, scalar quantization, and entropy coding [4].

If a block is encoded in intra predictive mode, a prediction block is formed based on previously encoded and decoded blocks. This prediction block is subtracted from the current one prior to encoding. In the Main Profile, there are a total of 9 optional prediction modes for each 4x4 luminance subblock, 4 optional modes for a 16x16 luminance block and 4 modes for each 8x8 chrominance block. The encoder selects the prediction mode for each block that minimizes the prediction residual. The macroblocks are then transformed using a 4x4 separable integer transform with properties similar to that of DCT.

In the High Profile or AVC/H.264 Fidelity Range Extensions (FR Ext), the encoder chooses adaptively between the 4x4 pixels and the 8x8 pixels block size for the transform of the luminance samples. Using the 8x8 transform enables the preservation of fine details and textures which generally require larger basis functions.

A non-uniform scalar quantization is then used to quantize the transformed coefficients where each macroblock has a Quantization Parameter (QP). Finally, the quantized coefficients are entropy coded. AVC Intra supports two modes for entropy coding: Context-Adaptive Variable Length Coding (CAVLC) and Context-Adaptive Binary Arithmetic Coding (CABAC). With CAVLC, multiple VLC tables are available and the encoder switches among them based on previously encoded syntax elements. CABAC is based on an arithmetic coding approach. Using arithmetic coding, each symbol of the alphabet can be assigned a non-integer number of bits, therefore outperforming VLC tables. Furthermore, a context model is built based on the statistics of previously transmitted syntax elements in order to better estimate conditional probabilities. This allows for the adaptation to non-stationary statistics. For these reasons, CABAC achieves substantially better coding performance when compared to CAVLC. However, it requires significantly higher computational complexity.

In the High Profile, three sets of context models are added in CABAC for the 8x8 transform coding. Meanwhile, CAVLC is used by regrouping the 8x8 transform coefficients into groups of 4x4. Furthermore, AVC/H.264 HP offers several interesting features such as scaling matrices for perceptually tuned and frequency-dependent quantization specified at the encoder, and the lossless coding capability. To encode high quality video content, the Joint Video Team (JVT) added five new profiles, High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra, CAVLC 4:4:4 Intra and High 4:4:4 Predictive profiles. They include several improvements such as better compression efficiency, better lossless coding, and support for higher bit depths up to 14 bits per sample. In this work, the High 4:4:4 Intra profile is used. With this profile, more effort is put into encoding of all color components; while in other profiles used to encode 4:2:0 or 4:2:2 video content, the emphasis is put on the luminance component. For high quality video, it is reported that the bit rate reduction with the new 4:4:4 profiles is around 20% with respect to a straight forward extension of H.264/AVC to the 4:4:4 format without including the enhanced chrominance compression capability. For more details on AVC/H.264 please refer to [4], [9], [15].

2.3 HD Photo

The HD Photo aims at achieving state-of-the-art image compression, while simultaneously keeping the encoder and the decoder complexities lower. HD Photo is based on a block-transform. Its coding structure, which shares some similarities with traditional image coding techniques, is composed of the following steps: color conversion, reversible Lapped Bi-orthogonal Transform (LBT), flexible quantization, inter-block prediction, adaptive coefficient scanning, and entropy coding of transform coefficients [1] [2]. The distinguishing features are the LBT and advanced coding of coefficients.

HD photo supports a wide range of pixel formats, including High Dynamic Range (HDR) and wide gamut formats, as well as a number of color formats including RGB and CMYK. Internally, HD photo is defining six YUV formats, the encoder selecting the most suitable during encoding. Color and bit depth conversion are performed to convert from external to internal formats whenever required.

To convert spatial domain image data to frequency domain, HD Photo uses an integer hierarchical two stage LBT which is based on a flexible concatenation of two operators: the Photo Core Transform (PCT) and the Photo Overlap Transform (POT). PCT is similar to the widely used DCT and exploits spatial correlation within the block. However, it fails to exploit redundancy across block boundaries and may introduce blocking artifacts at low bit rates. To alleviate these drawbacks, POT is designed to exploit the correlation across block boundaries. Hence, POT improves compression efficiency, while at the same time reducing blocking artifacts.

The transform is performed in a two-stage hierarchical structure. For the sake of simplicity, we consider the case of the luminance channel. At the first stage, a 4x4 POT is optionally applied, followed by a compulsory 4x4 PCT. The resulting 16 DC coefficients of all 4x4 blocks within a 16x16 macroblock are grouped into a single 4x4 DC block. The remaining 240 AC coefficients are referred to as the High Pass (HP) coefficients. At the second stage, the DC blocks are further processed. Another optional 4x4 POT is first performed on the DC blocks, followed by the application of a compulsory 4x4 PCT. This yields 16 new coefficients: one second stage DC coefficient and 15 second stage AC coefficients, referred to as the DC and Low Pass (LP) coefficients respectively. The DC, LP and HP bands are then quantized and coded independently. All transforms are implemented by lifting steps. The chrominance channels are processed in a similar way. Whenever POT and PCT are concatenated, the transform becomes equivalent to LBT. For more details on LBT, the reader can refer to [16].

In order to enable the optimization of the Quantization Parameters (QP) based on the sensitivity of the human visual system and the coefficient statistics, HD Photo uses a flexible coefficient quantization approach. Namely, QP can be varied based on the spatial location within the image (e.g. at frame, tile or macroblock level), the frequency band and the color channel.

To further improve compression efficiency, inter-block coefficient prediction is then used to remove inter-block redundancy in the quantized transform coefficients. More specifically, three levels of inter-block prediction are supported: prediction of DC sub band coefficients, prediction of LP sub band coefficients, and prediction of HP sub band coefficients. This prediction step can be compared to the Intra prediction used in AVC/H.264.

Adaptive coefficient scanning is then used to convert the 2-D transform coefficients within a block into a 1D vector to be encoded. Scan patterns are adapted dynamically based on the local statistics of coded coefficients. Coefficients with higher probability of non-zero values are scanned earlier.

Finally, the transform coefficients are entropy coded. For this purpose, advanced entropy coding using adaptive VLC table switching is used. A small set of fixed VLC tables are defined for each syntax element. The best table is then selected based on local coefficient statistics. The choice of table is computed based on previously coded values of the corresponding syntax element. Furthermore, HP coefficients are layered coded. More precisely, the coefficients are partitioned into two components by adaptive coefficient normalization. After the partition, the significant information is entropy coded and the remainder is signaled using fixed length codes.

For more details on HD Photo, the reader is referred to [1] [2].

3. TEST MATERIAL AND EVALUATION METHODOLOGY

3.1 Pictures

Table 1 represents the set of test pictures used in our tests. The images have been taken from the Microsoft test set in [13]. The border-pixels of each original image have been cut in order to obtain images with width and height multiple of 16 pixels so as to be able to compress them with the available software implementations without any adaptations. It is important to note that none of the compression schemes under study requires such a limitation in terms of the sizes of images to be compressed.

<i>Image name</i>	<i>Resolution (4:4:4)</i>	<i>Description</i>
P01crop	4064x2304	flowers: high textured content
P02crop	2256x1504	boats: both details and uniform texture
P03crop	4064x2304	kid with balloon: human details and uniform texture
P04crop	2256x1504	mountain house: details and uniform texture
P05crop	4064x2304	car1: details
P06crop	4064x2304	car2: details
P07crop	2592x1936	woman with complex background
P08crop	2592x1936	woman with textured background
P09crop	4064x2304	dog with uniform background
P10crop	2128x2304	skyscraper

Table. 1. High Resolutions images used for the evaluation.

As can be seen from Table 1, the test set contains a total of 10 high definition color images in 4:4:4, totaling four different spatial resolutions, and content similar to those typically produced by digital imaging applications.

3.2 Encoders parameters

For the AVC/H.264, the publicly available JVT-V206 software implementing the High 4:4:4 Intra profile was used [17]. Each image of the set was first converted into a YUV 4:4:4 one-frame video sequence and coded as an I-frame by varying the Quantization Parameter (QPISlice) between 0 and 51. The encoder_intra.cfg configuration file provided with the JVT-V206 software was used, setting the following encoder parameters values:

- ProfileIDC = 244 (High 4:4:4 Intra Profile).
- LevelIDC = 51 (Level 5.1).
- YUVFormat = 3 (YUV 4:4:4 format).
- SeparateColorFlag = 0 (common mode).
- BitDepthLuma = 8.
- BitDepthChroma = 8.
- RDPictureDecision=1 (enable rate-distorsion optimization).
- RDPictureIntra=1.
- RDPSliceWeightOnly=0.

- RDOptimization=1 (rd-optimized mode decision).
- UseExplicitLambdaParams = 1 (enable the use of explicit lambda parameters).
- LambdaWeightIslice = 0.5.
- EnableIPCM=1 (IPCM mode enabled).

The CABAC entropy coder and the de-blocking filter are enabled by default in the encoder_intra.cfg configuration file.

The software KAKADU version 5.2.6 was used for the JPEG 2000 compression of each original RGB image with 24 bits per pixels, with the following settings [18]:

- Codeblock size of 64x64 (default).
- One tile per frame (default).
- 5 decomposition levels (default).
- Visual Frequency Weighting: switched-off.

The base step parameter QStep was varied between 0.01 and 8.0 bits per pixel.

The HD Photo results were obtained by the HD Photo Device Porting Kit version 1.0 available at [19]. The default parameters settings of the HD Photo encoder/decoder modules were used to encode each original RGB image with 24 bits per pixels [20]. The quality parameter (q) was varied between 2 and 255 (lossy coding) and we considered:

- No chroma sub-sampling (Y_CC_g 4:4:4): d=3 (default).
- One level overlapping: l=1 (default).

3.3 Evaluation

The performance of HD Photo, JPEG 2000 and AVC/H.264 High 4:4:4 Intra Profile was measured by coding each test image and by computing the PSNR and the SSIM [12] index between the original input image and the coded/decoded output image. To facilitate a useful comparison with the YUV 4:4:4 output of the JVT-V206 encoder, input and output RGB images were transformed to YUV 4:4:4. The PSNR between the original and the compressed image has been thus evaluated on the Y, U and V components. The results in terms of PSNR considering the Y component of each image have been plotted along with the bit per pixel values, evaluated on the compressed bit-stream.

The structural similarity indexing visual quality measure represents an alternative image quality assessment approach. It is based on the assumption that the human visual system is highly adapted to extract structural information from a scene, and measures the structural distortion, considering it as a good approximation of the perceived image distortion. The SSIM index metric is computationally simple and has been found to have a good correlation with the subjective perceived image quality [12] [21]. The SSIM indexing algorithm is applied using a sliding window approach which results in a SSIM index quality map of the image. The average of the quality map is called Mean SSIM index. In this context, the SSIM index is computed as MSSIM index considering a weighted summation of the MSSIM indexes evaluated on the Y, C_b and C_r components of the image according to the following formula:

$$MSSIM = W_y MSSIM_y + W_{C_b} MSSIM_{C_b} + W_{C_r} MSSIM_{C_r}$$

where the weights are selected to be $W_y=0.8$, $W_{C_b}=W_{C_r}=0.1$ [21]. The YUV and RGB 4:4:4 images, outputs of the encoders, were thus transformed to YC_bC_r 4:4:4 in order to compute the MSSIM index as defined above. The results in terms of MSSIM have been plotted along with the bit per pixel values in addition to PSNR, in order to provide an alternative to provide quantitative estimates for the quality of color renditions and a more subjectively relevant metric.

4. PERFORMANCE EVALUATION

In this section, the rate-distortion plots in terms of PSNR and MSSIM index for HD Photo, AVC/H.264 High 4:4:4 Intra Profile and JPEG 2000, for our set of test images are analyzed. The considered range of PSNR values is between 30 and 50 dB. Depending on the considered image, this PSNR range corresponds to a bit per pixel range between 0.01 to 5 bpp

for uncompressed original images with 24 bpp. The performance in terms of MSSIM index are thus plotted considering the same ranges of bit per pixels values. The results for the JPEG coding have been also introduced in the graphs, as reference plot.

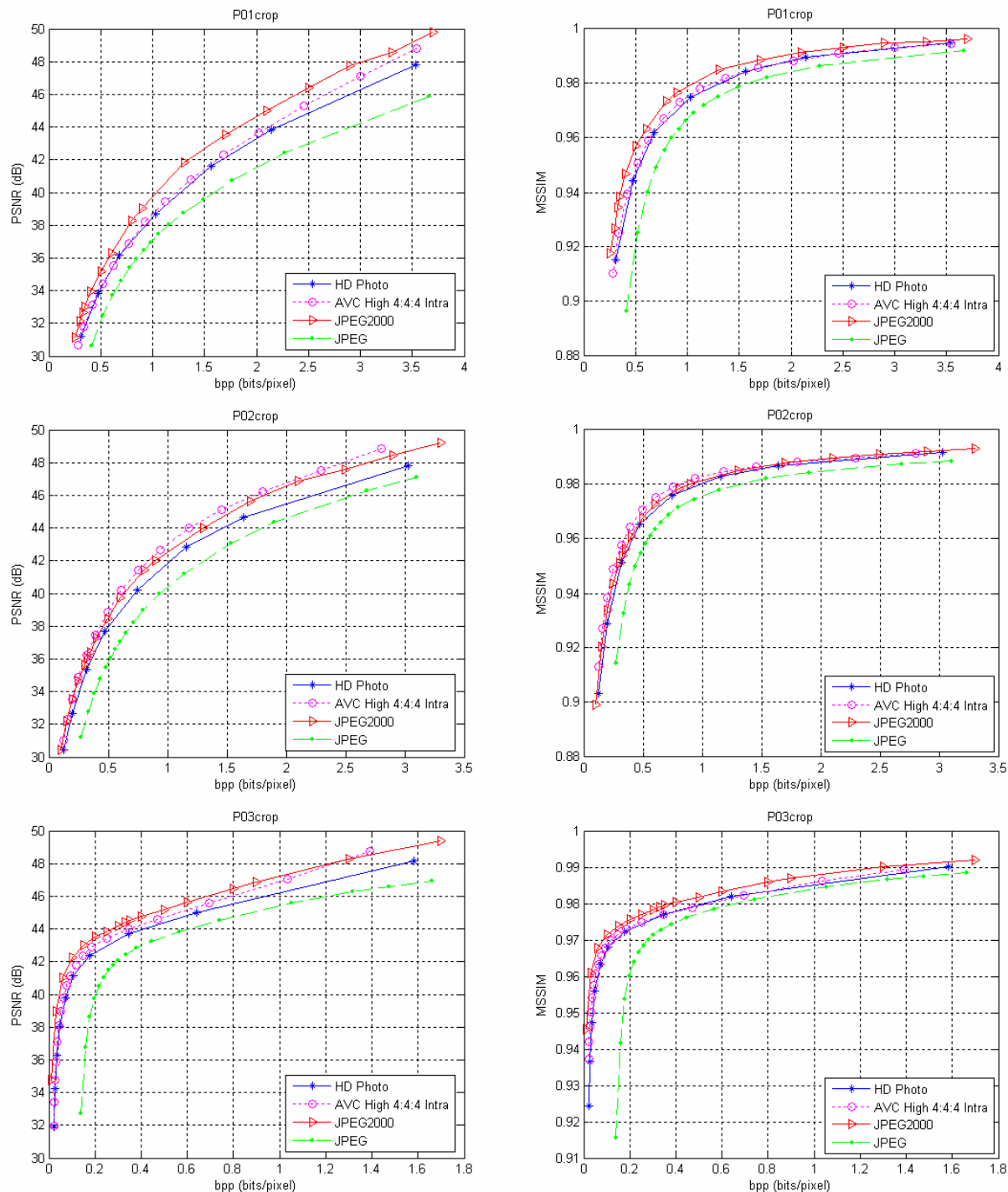


Fig. 2. Rate-Distortion performance using PSNR and MSSIM metrics, for images P01crop, P02crop and P03crop.

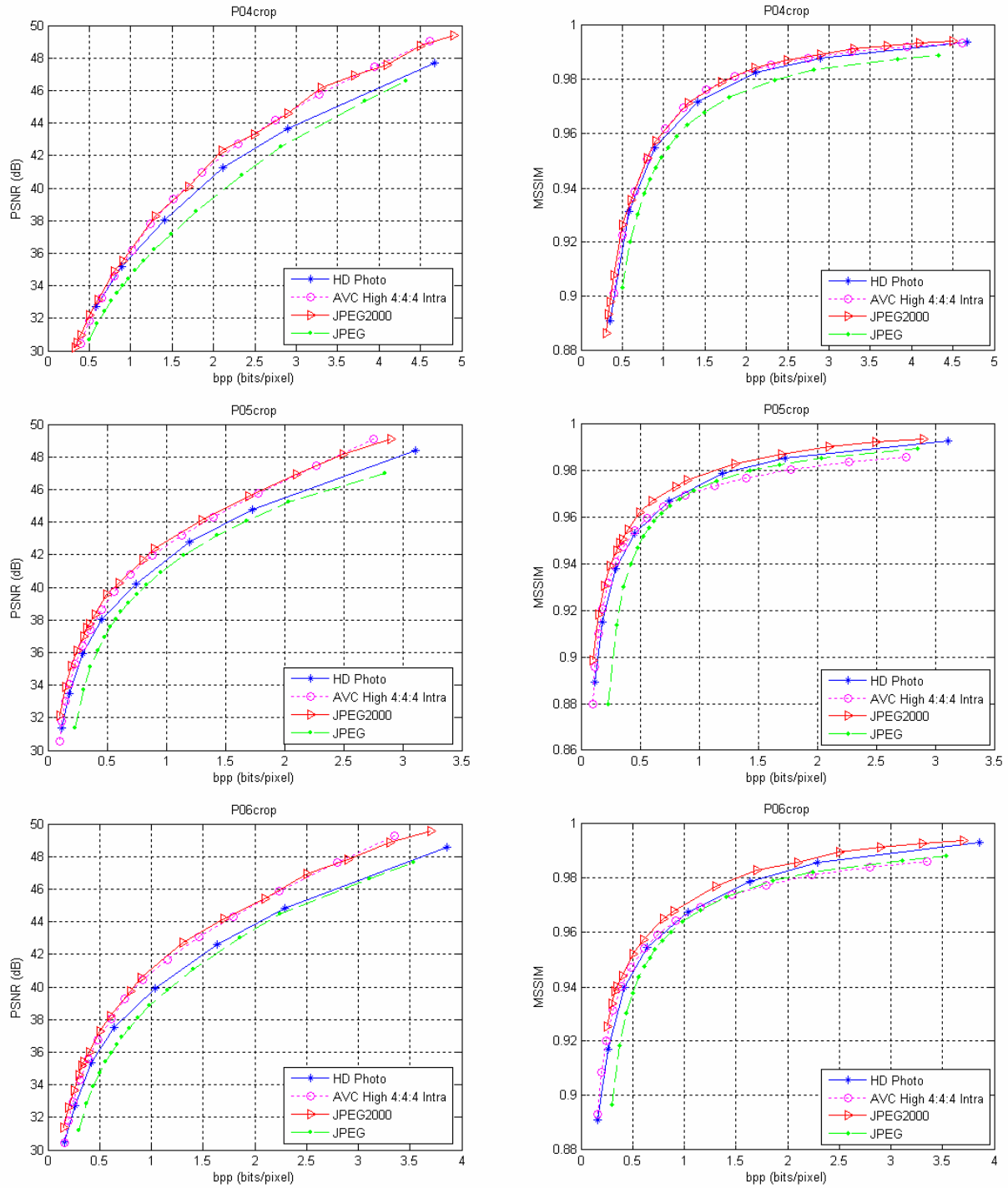


Fig. 2. Rate-Distortion performance using both metrics, PSNR and MSSIM, for images P04crop, P05crop and P06crop.

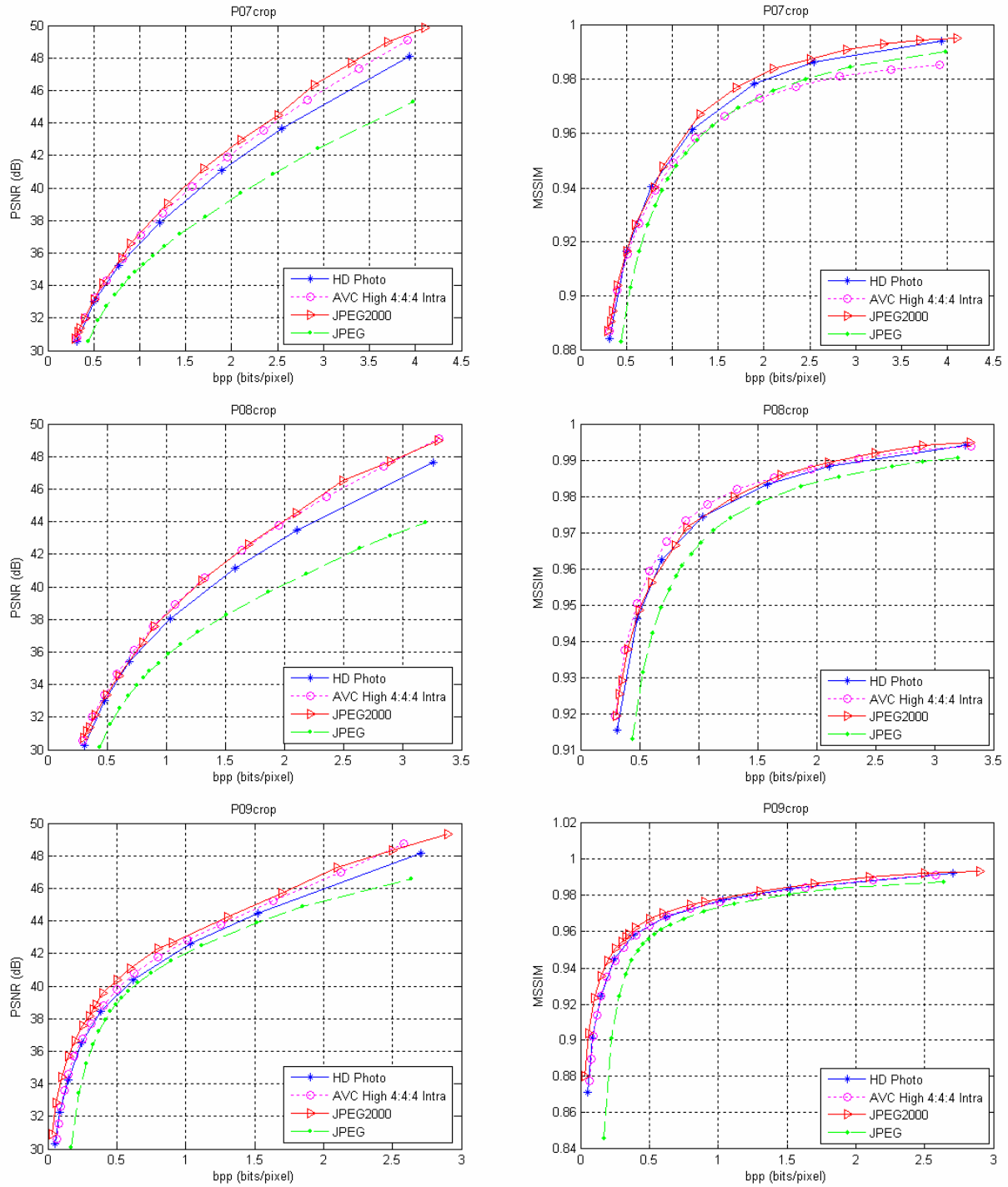


Fig. 3. Rate-Distortion performance using PSNR and MSSIM metrics, for images P07crop, P08crop and P09crop.

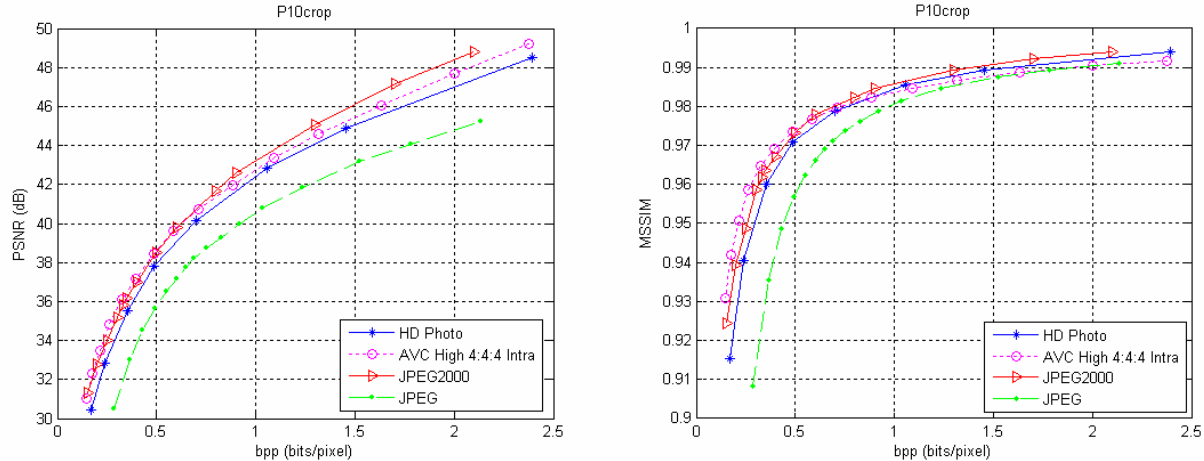


Fig. 4. Rate-Distortion performance using PSNR and MSSIM metrics, for image P10crop.

In terms of PSNR on the Y component, JPEG 2000 outperforms HD Photo for all the images. The average gain, considering the entire bit per pixel range and all the image test set, is in the range 0.5~ 1.7 dB.

AVC/H.264 High 4:4:4 Intra Profile performs well compared to HD Photo. In general, it can be observed a PSNR difference on most images in favor of AVC/H.264 High 4:4:4 Intra Profile in the range 0.1~ 1.2 dB. Considering the images P01crop, P04crop and P09crop, for smaller bitrates HD Photo has comparable performance and at some points slightly outperforms AVC/H.264 High 4:4:4 Intra Profile with an average maximum gain of 0.2 dB.

Comparing the JPEG 2000 and the AVC/H.264 High 4:4:4 Intra Profile performance, JPEG 2000 in most cases outperforms AVC/H.264 High 4:4:4 Intra Profile, reaching an average maximum gain of 1.4dB on the complete set of test images. In particular, on image P01crop, JPEG 2000 always outperforms AVC/H.264 High 4:4:4 Intra Profile with a gain in the range 0.4~ 1.3 dB. On the other hand, considering all the other test images, there are some points where AVC/H.264 High 4:4:4 Intra Profile has a slight gain with respect to JPEG 2000, with an average maximum value of 0.2 dB considering all the images. In particular, on image P02crop and P08crop, AVC/H.264 High 4:4:4 Intra Profile and JPEG 2000 have similar performance and considering the image P10crop, for bitrates lower than 0.6 bpp AVC/H.264 High 4:4:4 Intra Profile outperforms JPEG 2000, and vice-versa for higher bitrates.

These results are correlated to the results in [13], [14], with small value differences due to the different color spaces, image pre-processing and encoding parameters used.

Since PSNR is not the best quality metric in terms of correlation with the subjective visual quality perception, and the information on chrominance components have not been taken into account, the comparison between the algorithms has been also performed in terms of MSSIM, as explained in the previous subsection. The results show that JPEG 2000 performance is still the best, but the HD Photo results are also more competitive when compared to the case with PSNR. In particular, on the images P05crop, P06crop and P07crop, HD Photo outperforms AVC/H.264 High 4:4:4 Intra Profile. AVC/H.264 High 4:4:4 Intra Profile shows the best performance on P02crop and P08crop. Finally, considering the images P10crop, for bitrates lower than 0.6 bpp, AVC/H.264 High 4:4:4 Intra Profile outperforms JPEG 2000, while for higher bitrates, JPEG 2000 outperforms AVC/H.264 High 4:4:4 Intra Profile, similarly to the results obtained with PSNR. Finally, HD Photo outperforms AVC/H.264 High 4:4:4 Intra Profile for bitrates above 0.7 bpp.

5. CONCLUSIONS

In this paper, a simple evaluation methodology has been used to compare the compression performance between JPEG 2000, HD Photo and AVC/H.264 High 4:4:4 Intra Profile for compression of still images. To do so, rate-distortion characteristics of each codec were produced for a number of high resolution, 4:4:4 color images. The distortion metrics were chosen to be the widely used PSNR measure and the Mean SSIM (MSSIM) index metric. The PSNR was used to compare the quality of luminance components of the tested images. MSSIM brings additional information about the

quality of the chrominance quality, but also is an alternative metric to compare the visual distortions produced by each approach under study.

It is important to note that comparison of compression efficiency is a non-trivial and delicate task. Therefore, drawing conclusions should be made with a lot of care, and by providing many details about the environment, test conditions, and material used to generate the results. First, none of the approaches under study imposes a rigorous constraint on the way the encoding is performed. Experience shows that depending on the choice of the encoding strategy, different encoders implementing a same coding approach (and hence producing compressed bitstreams, decodable by the same decoder), can produce significantly different results. So, when comparing coding approaches, it is more adequate to talk about performance comparisons between specific encoder implementations rather than general coding approaches. The nature of the test material and the number of test images can also have a significant impact on the conclusions. In summary, the conclusions presented below should be considered as a snap shot in time of the performance comparison of three specific codecs, and for the specific test material selected for the study.

Further to the above considerations, it is important to note that other parameters such as the complexity and resource requirements on both encoder and decoder, as well as additional features offered by a given coding scheme can have significant impact in the performance of a given approach, and should also be taken into account.

The results of our study show that overall, all three codecs under study provide comparable compression efficiency for the conditions and the test material used to carry out the experiments. In cases, JPEG 2000 and AVC/H.264 High 4:4:4 Intra Profile can slightly outperform HD Photo with gains around 0.5~1.7 dB and 0.1~1.2 dB in PSNR respectively, when only considering the luminance component of the image in YUV 4:4:4.

Future work consists in comparing the visual quality of the results produced by different codecs by mean of subjective tests. In addition, carrying out tests with more efficient encoders could improve the results produced by some of the approaches under the study. Furthermore, a more extensive set of images, accurately selected in order to consider contents with different critical features, can also identify strengths and weaknesses of each approach. Finally, a comparison of complexity, memory requirements and power consumption between the codecs under study should be performed in order to produce a better understanding of their relative performance.

ACKNOWLEDGEMENTS

This work has been performed in the context of the EU funded FET project DISCOVER and EU funded Network of Excellence VISNET-II. The authors would like to acknowledge the help from Gary Sullivan, Alexis Tourapis, Haoping Yu and Karsten Suehring for producing the results obtained for AVC/H.264.

REFERENCES

1. S. Srinivasan, C. Tu, Zhi Zhou, D. Ray, S. Regunathan and G. J. Sullivan (Microsoft Corporation), "An Introduction to the HD Photo Technical Design", JPEG document WG1 N4183.
2. S. Srinivasan, C. Tu, S. L. Regunathan, R. A. Rossi, Jr., G. J. Sullivan, "HD Photo: a new image coding technology for digital photography," *Applications of Digital Image Processing XXX, Proceedings of SPIE*, vol. 6696, San Diego, CA USA (August 2007).
3. A. Skodras, C. Christopoulos, T. Ebrahimi, "The JPEG 2000 still image compression standard", *Signal Processing Magazine, IEEE* vol. 18, Issue 5, 36-58 (September. 2001).
4. T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 7, 560-576 (July 2003).
5. D. Marpe, V. George, and H. L. Cycon, "Performance evaluation of Motion-JPEG2000 in comparison with H.264/AVC operated in intra coding mode", *Wavelet Applications in Industrial Processing, Proceedings of SPIE*, Rhode Island, USA (October 2003).

6. P. Topiwala, "Comparative study of JPEG2000 and H.264/AVC FRExt I- frame coding on high-definition video sequences", *Optical Information Systems III, Proceedings of the SPIE*, vol. 5909, 284-292 (2005).
7. M. Ouaret, F. Dufaux, T. Ebrahimi, "On comparing JPEG2000 and Intraframe AVC", *Proceedings of SPIE, Applications of Digital Image Processing XXIX*, San Diego, USA (August 2006).
8. M. Ouaret, F. Dufaux, T. Ebrahimi, "On comparing image and video compression algorithms", *International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM 2007)*, Scottsdale, Arizona, U.S.A, (January 2007).
9. Joint Video Team of ITU-T and ISO/IEC, "Draft Text of H.264/AVC Fidelity Range Extensions Amendment," Doc. JVT-L047, (September 2004).
10. D. Marpe, S. Gordon, and T. Wiegand, "H.264/MPEG4-AVC Fidelity Range Extensions: Tools, Profiles, Performance, and Application Areas", *IEEE International Conference on Image Processing*, Genova, Italy, (September 2005).
11. S. Cho, Z. Bojkovic, D. Milovanovic, J. Lee, and J. Hwang, "Image quality evaluation: JPEG2000 versus Intra-only H.264/AVC High Profile", *FACTA UNIVERSITATIS Elec. Energ.*, vol. 20, no. 1, 71-83 (April 2007).
12. Z. Wang, A. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", *IEEE Trans. on Image Processing*, vol. 13, no. 4, 600-612 (April 2004).
13. Thomas Richter, Simulations available at the JPEG committee document register.
14. J. Ridge and K. Willner (Nokia Corporation), "ISO/IEC JTC 1/SC 29/ WG 1 4285: H.264/AVC intra-only", ISO/IEC JTC 1/SC 29/ WG 1, (June 2007).
15. G. J. Sullivan, H. Yu, S. Sekiguchi, H. Sun, T. Wedi, S. Wittmann, Y. Lee, A. Segall, and T. Suzuki, "New standardized extension of MPEG-4AVC/H.264 for professional-quality video applications," *ICIP 2007*, San Antonio, USA, (September 2007).
16. H. S. Malvar, "Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts," *IEEE Trans. Signal Processing*, 1043-1053, (April 1998).
17. Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, "Joint 4:4:4 Video Model (JFVM) 6 Reference Software," Doc. JVT-V206, (January 2007).
18. <http://www.kakadusoftware.com>.
19. <http://www.microsoft.com/windows/windowsmedia/forpros/wmphoto/default.aspx>.
20. Microsoft Corporation, "HD Photo: Device Porting Kit Specification" version 1.0 (June 2006).
21. Z. Wang, L. Lu, and A. Bovik, "Video quality assessment based on structural distortion measurement", *Signal Processing: Image Communication*, vol. 19, 121-132 (2004).