

Performance evaluation of Motion-JPEG2000 in comparison with H.264/AVC operated in pure intra coding mode

Detlev Marpe^a, Valeri George^b, Hans L. Cycon^b, and Kai U. Barthel^b

^aFraunhofer-Institute for Telecommunications, Heinrich-Hertz-Institute (HHI)

^bFachhochschule für Technik und Wirtschaft (FHTW), University of Applied Sciences

ABSTRACT

Recently, two new international image and video coding standards have been released: the wavelet-based JPEG2000 standard designed basically for compressing still images, and H.264/AVC, the newest generic standard for video coding. As part of the JPEG2000 suite, Motion-JPEG2000 extends JPEG2000 to a range of applications originally associated with a pure video coding standard like H.264/AVC. However, currently little is known about the relative performance of Motion-JPEG2000 and H.264/AVC in terms of coding efficiency on their overlapping domain of target applications requiring the random access of individual pictures. In this paper, we report on a comparative study of the rate-distortion performance of Motion-JPEG2000 and H.264/AVC using a representative set of video material. Our experimental coding results indicate that H.264/AVC performs surprisingly well on individually coded pictures in comparison to the highly sophisticated still image compression technology of JPEG2000. In addition to the rate-distortion analysis, we also provide a brief comparison of the evaluated coding algorithms in terms of complexity and functionality.

Keywords: Video coding, wavelet-based coding, Motion-JPEG2000, H.264, AVC

1. INTRODUCTION

Certain applications like image sequence processing in medical imaging or professional motion picture production and archiving require a video coding tool, which allows random access to each individual picture. Motion-JPEG2000¹ as an extension of the JPEG2000 still image coding standard² provides this feature along with the ability to support both lossy and lossless representations within the same bitstream. Moreover, additional features of JPEG2000 like spatial and SNR scalability or region-of-interest coding may also be considered to be important for some applications in digital video coding. However, the success of a coding standard relies primarily on its improved coding efficiency when compared to other existing or emerging coding standards. Although several studies^{3,4} have been focused on rate-distortion (R-D) performance comparisons of JPEG2000 with its predecessors in the still image coding domain, little is known about the performance of Motion-JPEG2000 compared to its competitors in the area of video coding. On the other hand, the currently released video coding standard H.264/AVC^{5,6} has been shown to provide a major breakthrough with regards to compression efficiency.⁷ But here again, most of the currently published studies were intended to evaluate the overall R-D performance of H.264/AVC in comparison to prior video coding standards such as H.263 or MPEG-4 Part 2 (Visual).

In this paper, we investigate the performance of Motion-JPEG2000 and H.264/AVC on their joint region of applicability. By using H.264/AVC in pure intra coding mode, a fair comparison with Motion-JPEG2000 is achieved in terms of coding efficiency for individually coded pictures. For the purpose of gathering additional information on the potential of wavelet-based image coding, the proprietary state-of-the-art PACC codec⁸⁻¹¹ has been tested in competition with the two standards. Our simulation results were obtained by using video material covering a large variety of resolutions and other characteristics like the underlying progressive or interlaced timing of the original capturing process.

^a D.M.: e-mail: Detlev.Marpe@fraunhofer.hhi.de; url: <http://bs.hhi.de/users/marpe>; phone: +49 30 31002-619; fax: +49 30 3927200; postal: Image Processing Department, Fraunhofer-Institute HHI, Einsteinufer 37, 10587 Berlin, Germany. ^b V.G., H.L.C., and K.U.B.: email: [vgeorge,hcycon,barthel]@fhtw-berlin.de; url: <http://www.fhtw-berlin.de>; phone: +49 30 54699-363, +49 30 5019-2416; fax: +49 30 54699-329; postal: University of Applied Sciences (FHTW Berlin), Treskowallee 8, 10313 Berlin, Germany.

As a general trend of the outcome of our experiments, we observed a superior R-D performance of H.264/AVC on lower resolution video like CIF as well as on ITU-R 601 conforming interlaced video. For progressively scanned video at medium to high resolutions, both H.264/AVC and Motion-JPEG2000 perform virtually at the same R-D level with a slight advantage of PACC in terms of PSNR, whereas for very high video resolutions of 1920×1080 pels per picture both wavelet-based coders significantly outperform H.264/AVC.

The organization of the paper is as follows. In Section 2, we give a brief description of the evaluated coding algorithms. Section 3 contains the experimental results of our R-D performance study along with a discussion of the main results. Further aspects related to the support of additional functionalities and their use in potential applications as well as a short discussion of the complexity involved in the implementation of the evaluated coding algorithms can be found in Section 4.

2. OVERVIEW OF EVALUATED CODING ALGORITHMS

In this section, we first provide a brief description of the three coding schemes, which were evaluated in our performance study on intra coding. As already mentioned above, the focus of our attention is on a comparison between the standardized coding algorithms of Motion-JPEG2000 and H.264/AVC, where the latter has been restricted to pure intra coding. As a consequence, we completely ignore the aspects of temporal prediction in H.264/AVC for the purpose of this study.

2.1. Motion-JPEG2000

Motion-JPEG2000¹ as Part 3 of the JPEG2000 image coding standard is based on the core coding technology of the "baseline" JPEG2000 Part 1.² There are excellent review papers^{12,13} as well as a comprehensive textbook on JPEG2000 by Taubman and Marcellin,¹⁴ to which the interested reader is referred for further information. In this paper, we restrict our review of JPEG2000 to a brief summary of its main distinctive features.

The basic coding algorithm of JPEG2000 is a three-step transform coder consisting of a discrete wavelet transform (DWT), a uniform quantization with a central deadzone, and an embedded bitplane coder for the quantized transform coefficients. The latter operates independently on rectangular blocks of quantized transform coefficients, so-called codeblocks, by performing a context-adaptive binary arithmetic coding in three fractional bit-plane passes. The generation of individual embedded bitstreams for each codeblock is commonly referred to as tier-1 (T1) coding in the JPEG2000 literature.^{13,14} The organization of JPEG2000 codestreams, also called the tier-2 (T2) coding process, is such that the bitstreams associated to a number of sub-bitplanes of individual codeblocks can be aggregated into packets and layers depending on the desired support of different possible modes of progression. The ordering of sub-bitplane bitstreams of a collection of codeblocks can be performed in a Lagrangian rate-distortion optimization process subject to a given constraint on the desired bitrate.^{13,14}

In a kind of pre-processing step, JPEG2000 supports the partitioning of the input image into rectangular, non-overlapping spatial regions on a regular grid, so-called tiles. Since each image can consist of multiple (spectral) components, such as e.g. RGB, tiling is applied to each component separately but in a spatially consistent way. The resulting tile components are treated independently in the subsequent coding process, although a joint R-D optimization across different components and/or tiles may be performed in the final T2 coding process.^{13,14}

Motion-JPEG2000 is capable of processing individual fields of a video sequence consisting of pictures with two interleaved fields, which were captured at different time instants. However, the underlying JPEG2000 coding tools are effectively agnostic with respect to the specific nature of the processed picture, i.e., coding of a single field utilizes the same coding primitives as coding of a whole frame. This is in contrast to the H.264/AVC video coding standard, which offers a more fine-tuned set of coding tools for the processing of interlaced video as will be discussed for the restricted case of pure intra coding in the next section.

2.2. H.264/AVC

An intra coded picture in a video coder is distinguished by using no other information for coding than that contained in the picture itself. Typically, intra coding is based on a transform coder that is enhanced by applying some kind of inter-block or inter-sample prediction within one picture or slice. In contrast to previous video coding standards such as H.263 or MPEG-4 Part 2, where the prediction for intra pictures is conducted in

the transform domain, prediction in H.264/AVC uses spatially neighboring samples of previously coded blocks. There are two types of prediction modes for the luminance samples: the so-called Intra_4x4 mode based on predicting each 4×4 block separately, and the Intra_16x16 mode for predicting a whole 16×16 macroblock. In Intra_4x4 mode, the encoding process can choose between nine prediction modes, one of which represents a plane DC prediction, and the remaining eight modes operate as spatially directional predictors corresponding to eight different angles.⁶ Prediction according to Intra_16x16 mode, which is well suited for smooth image areas, utilizes four prediction modes, as well as the separate intra prediction mode for the chrominance samples of a macroblock. Note that intra prediction across slice boundaries is not allowed in order to keep all slices independent of each other.

The transform part of H.264/AVC utilizes – similar to previous video coding standards – a block transform of the prediction residual. However, instead of using the popular 8×8 discrete cosine transform (DCT), H.264/AVC employs a multiplication-free, separable integer transform based on a 4×4 transform block size, which is suitable for implementation in 16-bit arithmetic. This low-complexity design leads to a subjectively pleasing reduction in ringing artifacts, and, at the same time, inverse transform mismatch problems are avoided.¹⁵ For the four DC coefficients of each chrominance component, an additional 2×2 transform is applied. If a macroblock is coded in Intra_16x16 mode, a similar 4×4 transform is performed for the 4×4 DC coefficients of the sixteen 4×4 transform blocks of the luminance signal, which correspond to a whole macroblock. This cascading of block transforms is equivalent to an extension of the length of the transform basis functions, thus leading to a better decorrelation of the signal in smooth image areas. For the quantization of transform coefficients, H.264/AVC uses scalar quantization, which can be controlled by selecting for each macroblock the quantization parameter (QP) out of 52 values. The QP values are arranged such that there is an increase of the quantization step size by approximately 12% when incrementing the QP value by one. Prior to entropy coding, the quantized transform coefficients of a block are generally scanned in a zig-zag fashion.

Entropy coding in H.264/AVC relies on two alternative methods. The basic entropy coding mode employs a zero-order Exp-Golomb code, which in the case of coding quantized transform coefficients is extended by the so-called Context-Adaptive Variable-Length Coding (CAVLC). The CAVLC method switches between various VLC tables depending on already coded syntax elements, and therefore achieves a better match of the VLC and the actually given conditional probabilities.⁶ For achieving a significantly improved coding efficiency than that provided by CAVLC possibly at the expense of higher complexity, Context-based Adaptive Binary Arithmetic Coding (CABAC) as the second entropy coding mode of H.264/AVC is the alternative choice.¹⁶

For interlaced material, the H.264/AVC design permits the choice of coding each picture by (1) combining the two fields together (frame coding mode), (2) coding the two fields separately (field coding mode), or (3) combining the two fields to one single frame and adaptively choosing for each pair of vertically adjacent macroblocks the frame or field coding mode. The latter option is referred to as macroblock-adaptive frame/field (MBAFF) coding, whereas the choice between options (1) and (2) is called the picture-adaptive frame/field (PAFF) coding. Note that the underlying coding processes of H.264/AVC in frame and field mode are specified in a similar way but with some important individual properties reflecting the anticipated specific geometric and statistical nature of frames and fields, respectively.⁶

2.3. PACC

As an alternative wavelet-based still image coder, we examined the so-called PACC approach. Although already published in 1997,⁸ this coding algorithm can still be viewed as one of the representatives of state-of-the-art technology in wavelet-based image coding. It was later refined in subsequent publications^{9,10} and has also been successfully applied to both motion-compensated video coding^{9,11} and combined lossy/lossless still image coding.¹⁰

PACC is a three-stage entropy coding process consisting of the steps (1) partitioning, (2) aggregation and (3) conditional coding. The initial partitioning process splits the quantized wavelet representation into three sub-sources related to significance, magnitude and sign of each quantized wavelet coefficient. In the (optional) step (2), an aggregation of insignificant coefficients across different scales (resolution levels) is performed by using the well-known instrument of zerotrees. In the final conditioning stage (3) of the PACC entropy coding method, an appropriate context model for the actual coding process in the subsequent adaptive arithmetic coder

is assigned to each element of the three sub-sources. Empirically optimized prototype templates have been designed for each of the three sub-sources.⁹ Note that in contrast to JPEG2000, PACC is not operating on bitplanes of quantizer indices, and therefore, the resulting bitstreams of PACC are not ordered in an embedded way that enables SNR progression. PACC, however, does support progressive transmission by spatial resolution and spectral component.

Table 1. Test sequences used in the rate-distortion performance analysis.

Name	Resolution	No. of used frames	Characteristics of content
Paris	CIF	100	Moderate spatial and color detail
Bus	CIF	260	Moderate spatial detail
Mobile&Calender	$720 \times 576i$	100	High spatial and color detail
Canoe	$720 \times 576i$	100	Moderate to high spatial and color detail
Crew	$1280 \times 720p$	300	Moderate spatial detail
Harbour	$1280 \times 720p$	300	High spatial detail
Vintage Car	$1920 \times 1080p$	250	Moderate to high spatial detail; film grain noise
Book	$1920 \times 1080p$	220	Moderate spatial detail; film grain noise

3. RATE-DISTORTION PERFORMANCE ANALYSIS

3.1. Test cases and test material

In our coding experiments, we examined four different test cases, each relating to a particular video resolution. The first experiment addresses low-resolution CIF video (352×288 pels) typically used in videoconferencing and video streaming applications. The second experiment evaluates R-D performance for interlaced-scan standard definition television (ITU-R 601) video sequences at resolutions of 720×576 pels (25 Hz) while the third experiment is devoted to the evaluation of 60 Hz progressive-scan high-definition (HD) sequences at 1280×720 pels (720p) resolution. In the fourth experiment, video sources at even higher resolutions of 1920×1080 pels (1080p) progressively scanned at 25 Hz are tested. Table 1 summarizes the input sequences for all investigated test cases. Note that all tested sequences are given in YUV 4:2:0 color format, where the two chrominance components (U,V) are down-sampled by factor of two in each spatial dimension.

3.2. Test conditions and software implementations of evaluated codecs

For our coding simulations, we used the following software implementations and specific choices of encoding parameters for the three coding algorithms:

- *Motion-JPEG2000*: Verification Model (VM) of Motion-JPEG2000, software version 2.1 built on top of JPEG2000 VM software version 8.6 has been used. The default encoding options for maximizing the R-D performance of a JPEG2000 Part 1 compliant bitstream output were used:
 - One tile per picture (no tiling)
 - Five levels of wavelet decomposition
 - 9/7-tap default biorthogonal wavelet filter kernel
 - Explicit quantization by step size control
 - Codeblock size of 64×64 samples
 - Sequential coding mode and full effort coding mode (disabling of parallel and lazy coding mode features)
 - One single quality layer

In addition, for a fair comparison of the R-D performance of Motion-JPEG2000 the fairly high amount of header data (typically greater than 170 bytes per picture) has been neglected for the bit-rate calculation.

- *H.264/AVC*: For encoding the progressive-scan source material, the Joint Model (JM) of the Joint Video Team (JVT), software version 7.1 has been used, while for encoding of the interlaced material a software implementation of a Main Profile compliant H.264/AVC encoder developed by ourselves (Fraunhofer-Inst. HHI) has been employed. Both encoders were following the same Lagrangian R-D optimizing strategy as proposed in.⁷ The CABAC entropy coding mode¹⁶ was chosen for all H.264/AVC coding experiments. Only one slice per picture was permitted and for the coding of interlaced source material the macroblock-adaptive frame/field (MBAFF) coding mode has been chosen.
- *PACC*: For the PACC approach, the software version, which was already used for the simulations reported in¹⁰ has been used. This implementation includes a four-level standard discrete wavelet decomposition using the 9/7-tap filter. Due to the lack of interlace support, PACC was only applied to progressive-scan source material.

As an objective performance measure, the average PSNR of the luminance component over all frames in a given sequence was chosen for representation of the visual distortion. The average PSNR of the chrominance components (U,V), however, has been adjusted for different encoder outputs to each other such that the R-D curves representing the average (U,V) distortion versus the average (total) bit-rate were nearly identical for all three encoders within a small tolerance. This objective was achieved by first producing the R-D curve related to the Motion-JPEG2000 encoder and then iteratively varying the quantization parameter offset for the chrominance components both for the H.264/AVC encoder and the PACC encoder until the chrominance R-D curves for all three encoders were approximately matching.

By forcing the R-D curves of the averaged (U,V) component to be virtually identical, it is sufficient to present for each test case and test sequence one R-D curve for each encoder, which shows the average luminance (Y) PSNR over the average bit-rate. The R-D curves for each encoder were generated in several coding passes by encoding in each pass the whole test sequence with a specific fixed quantization step size for the luminance and, in the case of H.264/AVC and PACC, by adjusting the chrominance quantization step size according to the above given procedure.

3.3. Test results and discussion

Figure 1 shows the R-D plots for the two selected test sequences of the first and second experiment. As an indication of the typically observed large performance gap between a pure intra coder and a full operating video coder using all kind of temporal prediction modes (in P or B pictures), the R-D diagram for the ‘Paris’ sequence also contains the R-D curve obtained for the H.264/AVC by using only one single I-picture at the beginning of the test sequence followed by two B-pictures, which were inserted between each P-picture. A bit-rate increase by a factor of 7 – 10 is observed for the pure intra coder in this test case of a typical videoconferencing scene, which is a high price to be paid in terms of coding efficiency for the functionality of random access of each individual frame.

However, under the premise that the support of such a feature is worth the substantial loss in R-D performance, our investigation clearly indicates that H.264/AVC outperforms Motion-JPEG2000 for the case of pure intra coding in the first two test cases as well. For the test sequences in CIF resolution, a gain in Y-PSNR of 0.5 – 1.5 dB in favor of H.264/AVC is observed, while the PACC coder shows a R-D performance somewhere located in the middle between both standardized codecs for this kind of test material.

For the test case of ITU-R 601 compliant TV interlace material, we generated two R-D curves for the Motion-JPEG2000 encoder: one curve by encoding each picture as a whole frame (denoted as frame coding) and a second curve by encoding the two fields of each picture separately (field coding). As can be seen from the R-D graphs in the bottom row of Fig. 1, the relative R-D performance of both coding modes depends largely on the nature of the given test material. For the ‘Mobile&Calendar’ sequence, a single picture typically shows a high amount of spatial consistency across adjacent rows, which were captured at different time instances, since in this sequence the camera is zooming slowly and the rather slow motion of the objects are covering only a small area of each picture. This, however, is in contrast to the characteristics of the ‘Canoe’ sequence, which includes a fast camera pan as well as a fast moving object in the scene, leading to a reduced degree of statistical dependency between

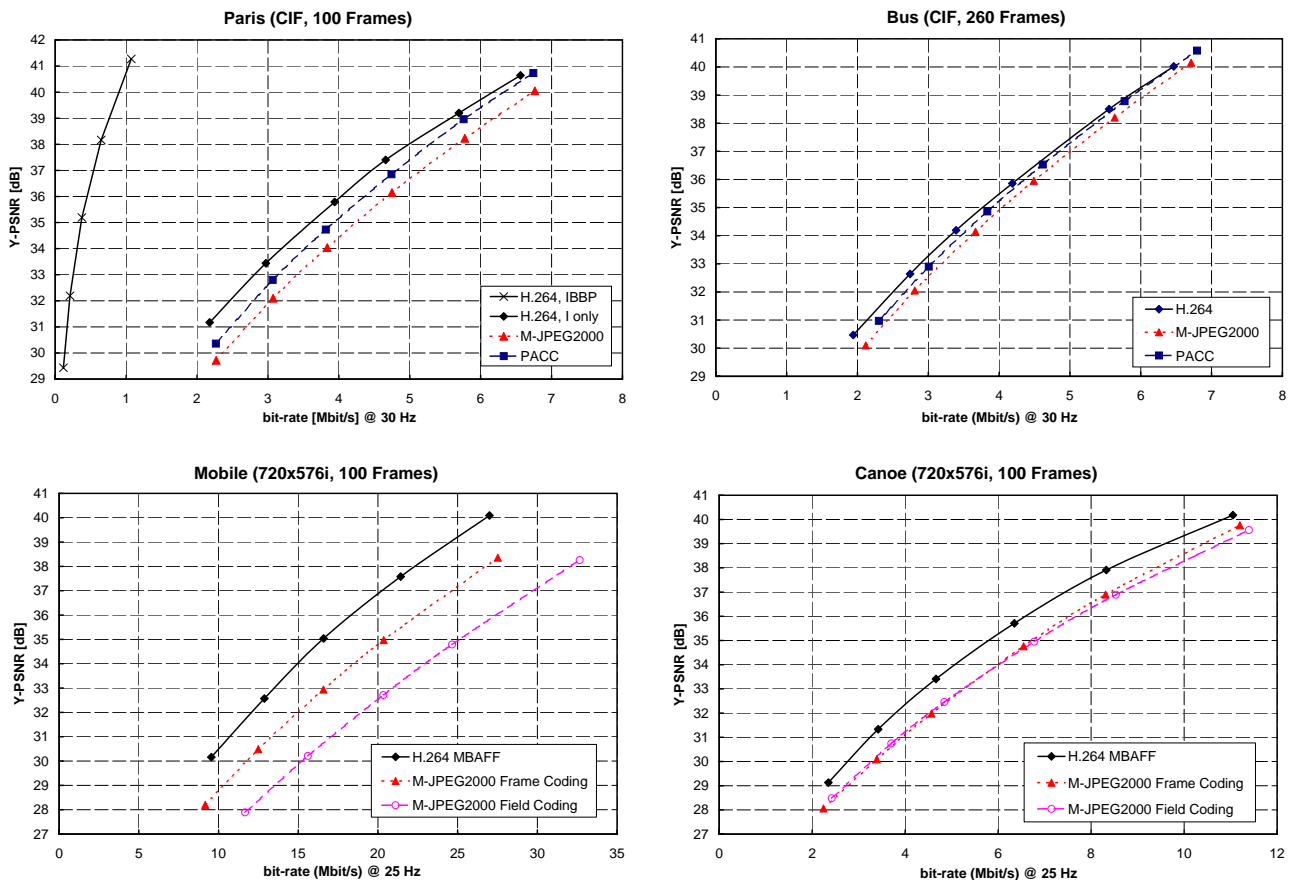


Figure 1. Rate-distortion curves for CIF (top row) and ITU-R 601 (bottom row) test material.

two adjacent rows that belong to different fields. In the latter case, field-based coding is favorable, whereas in the former case, frame-based coding usually leads to better results. This behavior is largely reflected in the R-D curves of the bottom row in Fig. 1, although it should be mentioned that frame-based coding for the case of the ‘Canoe’ sequence results in subjectively worse reconstructions than field-based coding, even if the related R-D curves are very close. Picture-adaptive frame/field (PAFF) coding may improve the R-D behavior of Motion-JPEG2000, but this kind of encoding option was not supported by the Motion-JPEG2000 VM 2.1.

For the case of H.264/AVC, MBAFF usually provides the most efficient way of encoding interlaced source material. This method of locally adapting the frame/field coding decision is a unique feature of H.264/AVC as already explained in Section 2.2, and it was reported to reduce the bit-rate in the range of 15% compared to PAFF, especially for source material with mixed regions.⁶ When comparing H.264/AVC MBAFF and the best performing coding mode of Motion-JPEG2000 in the outcome of our experiments related to our second test case, a significant gain in Y-PSNR of 0.5 – 2 dB for the H.264/AVC codec can be observed, as shown in Fig. 1.

The results of our experiments for the two selected test sequences of the third and fourth test case are shown in Figure 2. From the two R-D diagrams in the top row of Fig. 2, we draw the conclusion that H.264/AVC and Motion-JPEG2000 perform virtually at the same level for the 720p HD sequences with a slight advantage of Motion-JPEG2000 over H.264/AVC at lower bit-rates and vice versa at higher bit-rates. Overall, the PACC codec performs best with an improvement of 0.5 – 1 dB Y-PSNR relative to its worst performing competitor.

A clear advantage of Motion-JPEG2000 over H.264/AVC can be observed from the R-D curves plotted in the diagrams of the bottom row in Fig. 2 for the test case of 1080p. For this very high-resolution content,

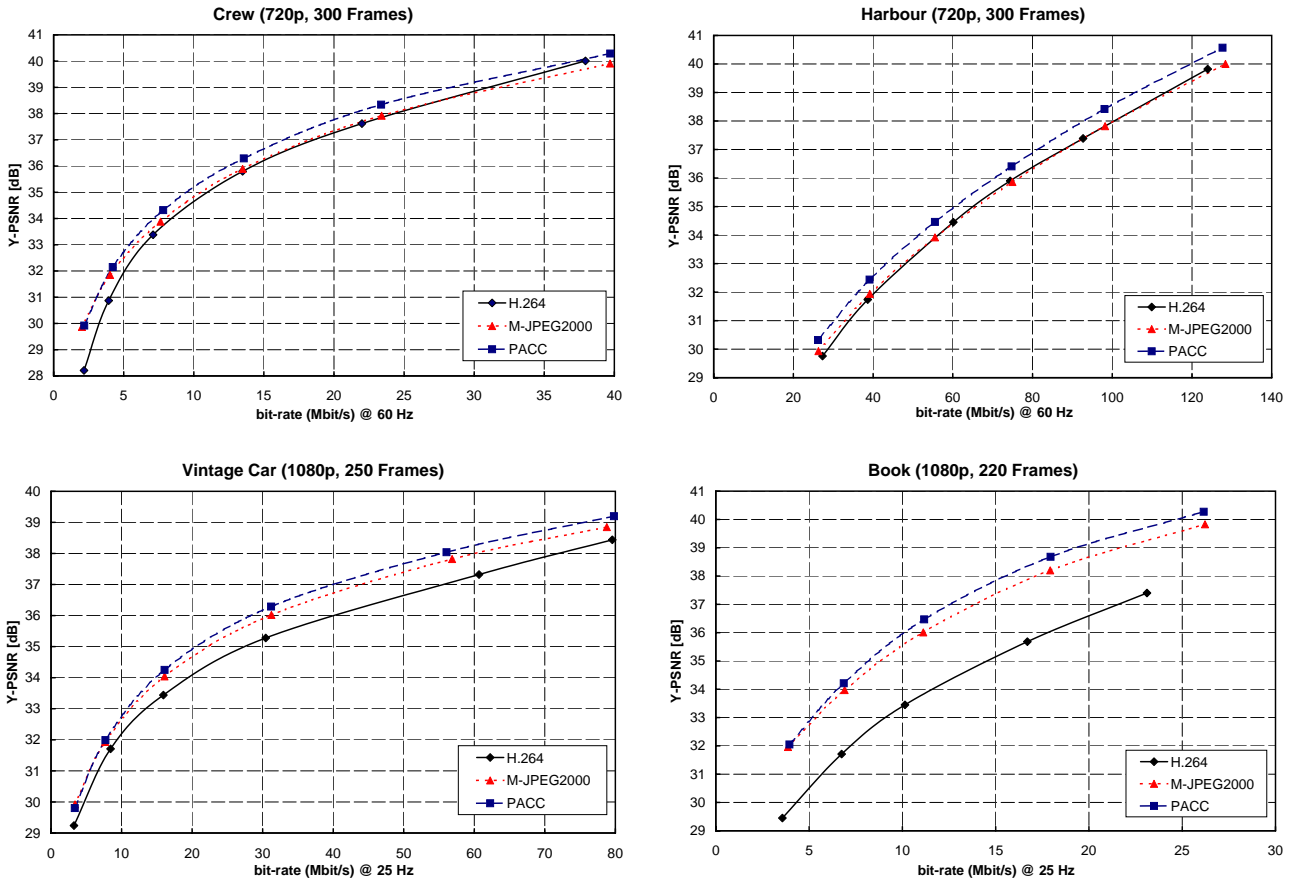


Figure 2. Rate-distortion curves for 720p (top row) and 1080p (bottom row) test material.

Motion-JPEG2000 achieved PSNR gains in the range of 0.5 – 2 dB relative to H.264/AVC. Additional gains of up to 0.5 dB PSNR were obtained by the PACC codec, in particular at higher bit-rates.

The overall advantage of both wavelet-based coders in terms of R-D performance for this kind of source material may be attributed to two phenomena. First, the inter-pixel correlation typically increases with increasing resolution such that the better decorrelating properties of the wavelet transform in largely smooth image areas may be increasingly beneficial. Secondly, the spatial prediction step in the H.264/AVC encoder may suffer from the film grain noise, which is typically observed for our tested 1080p source material, and which in the case of the wavelet-based coders is mostly well conserved in the lower significant bitplanes of the wavelet coefficients.

4. COMPLEXITY, FUNCTIONALITIES, AND APPLICATIONS

We conclude our performance study by a brief consideration of certain aspects regarding the complexity of the evaluated algorithms. In addition, we address some of the specific key functionalities that are provided by each of the three coding algorithms beyond compression, and finally, discuss how potential applications may benefit from those functionalities.

A profound and comprehensive complexity study of the evaluated coding algorithms is certainly far beyond the scope of this paper. Providing only a rough estimation of complexity in terms of memory bandwidth/access and/or computational cost is already a difficult task, since it depends heavily on the specific implementation platform. From a decoder centric point of view, the major bottleneck of all three algorithms seems to consist in the entropy decoding unit, which is mainly due to the inherently sequentially organized arithmetic decoding process.

In a complexity study related to different aspects concerning hardware-based implementations of JPEG2000, to which different companies contributed, it was consistently reported that the T1 entropy decoder accounts for more than 60% of the hardware complexity of a JPEG2000 compliant decoder implementation.¹⁷ For a non-bitplane oriented entropy coding approach like CABAC or PACC, this problem is partly alleviated, although not substantially. Another determining factor in complexity is the amount of memory transfer, which, at least from a conceptual point of view may have a significantly lower impact in the case of the block-based hybrid coder of H.264/AVC.

Besides these rather general considerations, it is worth mentioning that standardized codecs typically provide a rich set of instruments for putting certain restrictions on the encoding parameters such that some kind of complexity scalability can be achieved. Examples of these parameters are given by the choice of the wavelet kernel, the tile size or the codeblock size in the JPEG2000 environment, or alternatively, the slice size or the number of prediction modes in the H.264/AVC setting. However, it should be noted that by imposing such restrictions on the encoding process the R-D performance is generally negatively affected compared to the non-constrained case that we considered in our R-D performance analysis.

Apart from the pure compression task, image and video coding standards must also support a number of functionalities required by potential applications of the coding standards. Here we must distinguish between basic functionalities and optional value-added features. The most important basic technical feature that must be provided by any multimedia standard is a flexible interface to the broad variety of storage media and transmission layers. For that purpose, Motion-JPEG2000 defines its own MJ2 file format,¹ which is inherited from the JPEG2000 JP2 file format² and which is strongly related to the MPEG-4 MP4 file format. Compared to that, H.264/AVC provides a more flexible design of a so-called network abstraction layer (NAL),⁵ which customizes the video coding layer, i.e., the coded representation of the video data according to the different needs of the specific transport or storage media, like e.g., any kind of IP-based service, MP4 file format, or MPEG-2 Systems. Another important basic feature that must be provided by any coding standard is robustness to data errors/losses. H.264/AVC includes a rich set of tools particularly designed for that purpose.⁶ Motion-JPEG2000 provides also a couple of coding modes for improving error resilience, which are inherited from JPEG2000.¹⁴

Concerning the value-added features, ‘scalability’ is the most prominent feature of Motion-JPEG2000. Usually, we refer to this feature as the ability of extracting different resolutions, fidelities, components or spatial locations from a single compressed bitstream. This is a unique functionality of JPEG2000/Motion-JPEG2000, which has no correspondence on the H.264/AVC side, apart from the flexible macroblock ordering (FMO) feature, which allows some kind of region-of-interest support. In general, however, scalability features also have an impact on the R-D performance. For instance, rate or SNR-scalability may cause a drop in PSNR of up to 0.5 dB compared to the single layer case as tested in our experiments.¹² Another distinctive feature provided by Motion-JPEG2000 is the capability of generating a single bitstream, from which both a lossless and a number of lossy representations can be extracted. But here again, there is a price to be paid for providing this feature, which typically consists in a significant loss in R-D performance of 0.5 – 1 dB PSNR for the lossy reconstruction(s) due to the less efficient 5/3-tap reversible wavelet filter.¹²

Taken altogether, the range of target applications, where the two standardized codecs of Motion-JPEG2000 and H.264/AVC may compete is not very large. One field of applications may be given by future devices, where the codec is already built in for other purposes such as e.g., digital still cameras with a burst capturing mode suitable for JPEG2000/Motion-JPEG2000, or digital camcorders with a still image capturing mode more suitable for H.264/AVC. PC-based video capturing may represent another potential joint application domain, although there might be a tough competition with proprietary video codecs already built-in in most PCs.

Taking into account the outcome of our R-D performance analysis and considering the list of supported features, Motion-JPEG2000 seems to be most suitable for high-quality, high-resolution digital video coding in the field of film postproduction, archiving and distribution, commonly referred to as the ‘Digital Cinema’ chain. In addition, there might be a good chance for Motion-JPEG2000 to be applied in the field of high-resolution medical or satellite imaging.

H.264/AVC, on the other hand, will find its way into applications, where video transmission over potentially error-prone channels with rather limited transmission capacities is the primary task and where the random access of each individual picture is of subordinate importance.

5. CONCLUSIONS

A comparison of Motion-JPEG2000 and H.264/AVC in terms of rate-distortion performance for pure intra coding was presented. Coding simulations were performed on video material covering a large range of resolutions as well as progressive and interlaced scanning type.

As a result, we observed a significantly superior rate-distortion performance of H.264/AVC at lower resolutions and for standard definition television interlaced source material. In the range of progressive-scan, medium to high resolution video, our experiments showed a balanced behavior of both competing codecs with respect to their rate-distortion performance. At very high resolutions and high bit-rates, Motion-JPEG2000 clearly outperforms H.264/AVC on individually coded pictures, at least in terms of averaged PSNR over average bit-rate.

REFERENCES

1. ISO/IEC 15444-3 Motion-JPEG2000 (JPEG2000 Part 3), 2002.
2. ITU-T Recommendation T.800 and ISO/IEC 15444-1 JPEG2000 Image Coding System: Core Coding System (JPEG2000 Part 1), 2000.
3. T. Ebrahimi et al., "JPEG2000 Still Image Coding versus Other Standards", *Proc. SPIE*, Vol. 4115, pp. 446–454, 2000.
4. D. Santa-Cruz, R. Grosbois, and T. Ebrahimi, "JPEG2000 Performance Evaluation and Assessment", *Signal Processing: Image Comm.*, Vol. 17/1, 2002.
5. ITU-T Recommendation H.264 and ISO/IEC 14496-10 MPEG-4 Part 10, Advanced Video Coding (AVC), 2003.
6. T. Wiegand et al., "Overview of H.264/AVC Video Coding Standard", *IEEE Trans. on Circ. and Sys. for Video Technology*, Vol. 13, No. 7, pp. 560–576, July 2003.
7. T. Wiegand et al., "Rate-Constrained Coder Control and Comparison of Video Coding Standards", *IEEE Trans. on Circ. and Sys. for Video Technology*, Vol. 13, No. 7, pp. 688–703, July 2003.
8. D. Marpe and H. L. Cycon, "Efficient Pre-Coding Techniques for Wavelet-Based Image Compression", *Proc. Picture Coding Symposium 1997*, pp. 45–50, 1997.
9. D. Marpe and H. L. Cycon, "Very Low Bit-Rate Video Coding Using Wavelet-Based Techniques", *IEEE Trans. on Circ. and Sys. for Video Technology*, Vol. 9, No. 1 pp. 85–94, 1999.
10. D. Marpe, G. Blättermann, J. Ricke, and P. Maaß, "A Two-Layered Wavelet-Based Algorithm for Efficient Lossless and Lossy Image Compression", *IEEE Trans. on Circ. and Sys. for Video Technology*, Vol. 10, No. 7, pp. 1094–1102, Oct. 2000.
11. D. Marpe and H. L. Cycon, "High-Performance Wavelet-Based Video Coding Using Variable Block-Size Motion Compensation and Adaptive Arithmetic Coding", *Proc. 4th IASTED Int. Conf. on Signal and Image Proc. (SIP)*, pp. 404–409, Aug. 2002.
12. M. W. Marcellin, M. J. Gormish, A. Bilgin, and M. P. Boliek, "An Overview of JPEG2000", *Proc. IEEE Data Comp. Conf.*, pp. 523–541, 2000.
13. M. Rabbani, R. Joshi, "An Overview of the JPEG2000 Still Image Compression Standard", *Signal Processing: Image Comm.*, Vol. 17/1, 2002.
14. D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*, Kluwer Academic, 2002.
15. H. S. Malvar et al., "Low-Complexity Transform and Quantization in H.264/AVC", *IEEE Trans. on Circ. and Sys. for Video Technology*, Vol. 13, No. 7, pp. 598–603, July 2003.
16. D. Marpe, H. Schwarz, and T. Wiegand, "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard", *IEEE Trans. on Circ. and Sys. for Video Technology*, Vol. 13, No. 7, pp. 620–636, July 2003.
17. ISO/IEC JTC1/SC29/WG1 Doc. WG1N2027, "Singapore Core Experiment Report HW Analysis of Coding Options", Feb. 2001.