Fringe Benefits of the H.264/AVC

Ricardo L. de Queiroz, Rafael S. Ortis, Alexandre Zaghetto, and Tiago A. Fonseca

***Abstract*— H.264/AVC is the newest, state-of-the-art, video compression standard. Recently adopted, it leads to substantial performance improvement compared to other existing standards. In this paper we want to explore less known features of the H.264/AVC that were not the main intent of the standard. We begin by showing a few comparisons as examples of how the H.264/AVC outperforms other coders. Nevertheless, along with better video compression, the H.264/AVC brought us a few unexpected bonuses. H.264/AVC's intra-frame coder is so efficient that we can make it operate as a still image coder that outperforms the existing state-of-the-art still image compression standard (JPEG-2000). For that to work, one has to operate the H.264/AVC in intra frame and compress a single frame. Many results are shown comparing it with JPEG-2000 for still image compression. In effect, the intra-frame prediction mode is so efficient that interframe prediction can be simplified if necessary. We will show that one can turn off motion estimation and yet H.264/AVC can outperform MPEG-2. For that, we simply set all motion vectors to be zero. The interframe prediction is made to "compete" with the intraframe predictions. Results are presented to demonstrate the superiority of H.264/AVC with no motion estimation against MPEG-2 for sequences with moderate motion.**

I. INTRODUCTION

The newest video coding standard [1] was recently adopted in a joint effort of the Motion Picture Experts Groups (MPEG) from the International Standards Organization (ISO) and of the Video Coding Experts Group (VCEG) from the International Telecommunications Union (ITU). The union of the MPEG and VCEG was named the Joint Video Team (JVT). The resulting standard was filed as H.264 at ITU and as MPEG-4 Part 10 at ISO. The coder is codenamed advanced video coder (AVC) and is also referred as H.264, AVC, MPEG-4 Part 10. Here we refer to it as H.264/AVC.

The H.264/AVC coder has been well explained in the literature [2]–[7]. Many papers have illustrated its performance showing many comparative results against coders such as MPEG-2 and H.263. All results point to at least a factor of two improvement over previous standards. The reasons why the H.264/AVC is so good are many small improvements over previous methods. Each improvement brings a small coding gain, adding up to significant gains.

In this paper we want to highlight advantages of the H.264/AVC that are not well known or advertised. Apart from the factor-of-two improvement over other standards, there are a few fringe benefits that come with the AVC package. These fringe benefits are the use of the H.264/AVC as a still image coder and its use without motion estimation for lower complexity applications. We will first show comparative results for the H.264/AVC against other coders and then proceed to discuss these other features.

The authors are with the Department of Electrical Engineering, Universidade de Brasilia, e-mail queiroz@ieee.org and {ortis,zaghetto,tiago}@image.unb.br. This work was supported in part by a grant from HP Brasil.

II. H.264/AVC AGAINST OTHER STANDARDS

Many features of the H.264/AVC coder are innovative in terms of standards. Among the improvements of H.264/AVC we can mention [2],[3]:

- Quarter-pel motion estimation: motion vectors can have 4×4-times more precision. It demands interpolating all frames in order to carry the motion estimation.
- Arbitrary reference frames. Reference image regions used for interframe prediction can be located theoretically anywhere in the frame sequence, not only the previous frame, or next frame in the bidirectional prediction mode.
- In-loop deblocking filter: blocking can be removed or reduced in reference frames before they are used to predict the current frame. This prevents blocking artifacts of past frames from degrading the motion estimation of the current frame.
- Context-adaptive arithmetic coding: powerful arithmetic coding enables superior entropy coding performance.
- Intraframe prediction mode: efficient intraframe prediction modes can be selected based on rate-distortion performance, either to encode I-frames or to "compete" with interframe prediction.
- Variable size macroblock partition: motion estimation does not need to be carried on a full macroblock at a time. For each macroblock the user can break it into 2 or 4 submacroblocks, which can be further partitioned into 2 or 4 sub-blocks each. In this way, the blocks and sub blocks can better fit to object boundaries.
- Variable block size transform: the 16×16-pixel macroblock residual can be transformed by a combination of 8×8 and 4×4 transforms. One choses the transform set that yields the best rate-distortion performance for each macroblock. The transforms used in H.264/AVC are lower complexity DCT approximations [2].

Many papers have dealt with benchmarking the H.264/AVC against other video compression standards. Here, we briefly show a few results comparing the H.264/AVC with other coders such as MPEG (MPEG-1) [8], MPEG-2 [9], MPEG-4 (MPEG-4 Part 2) [10], DivX [11], and H.263 [12]. Peak signal-to-noise ratio (PSNR) plots are shown for popular sequences in Fig. 1. In this figure, the PSNR was computed over the whole sequence and the quantization parameter was varied in order to vary the bit-rate. We only used CIF resolution sequences at 30 frames per second (fps). These test sequences come in color so that bitrate includes all color channels and PSNR was computed in the YCbCr color space [2]. Further plots are shown in Fig. 2. Our simulation results are very similar

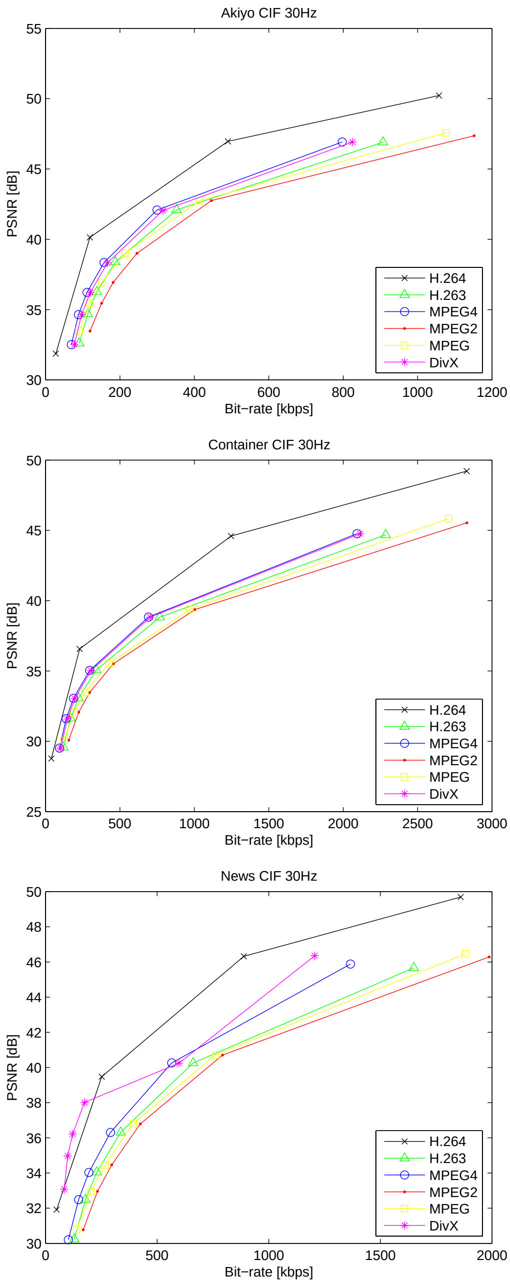

Fig. 1. PSNR (dB) plots for a few popular color video sequences. The sequence names are indicated. Resolutions are CIF (spatial) and 30 fps (temporal). PSNR was computed in YCbCr space.

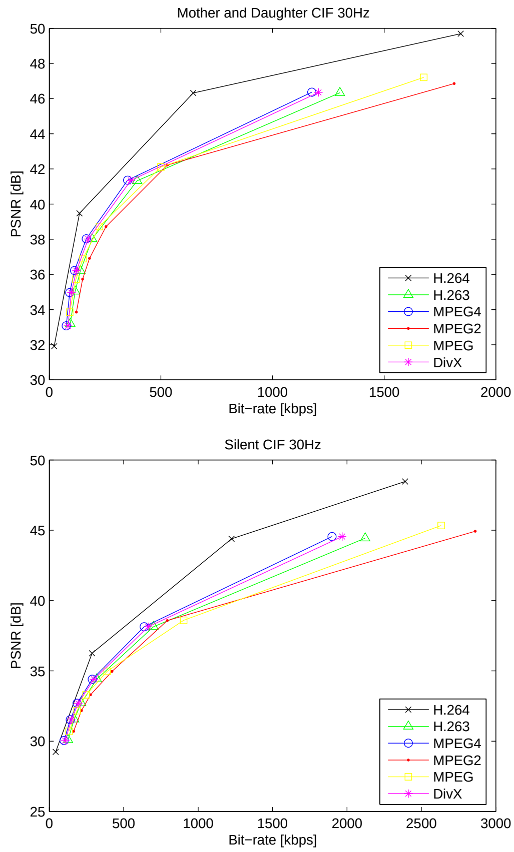

Fig. 2. PSNR (dB) plots for two popular color video sequences. The sequence names are indicated. Resolutions are CIF and 30 fps. PSNR was computed in YCbCr space.

to those other presented in the literature. Note the large gains brought by the H.264/AVC. Also note that there are large gains even over MPEG-4 which is a relatively recent standard [13].

III. H.264-Intra as a still image coder

H.264 is a video compression standard and it was not conceived to be applied as a still image compression tool. Nevertheless, the many coding advances brought into H.264, not only set a new benchmark for video compression, but they also make it a formidable compressor for still images [14]. One of the components of these advances is the intra-frame macroblock prediction method, which, combined with the context-adaptive binary arithmetic coding (CABAC), turns the H.264 into a powerful still image compression engine. If we set our H.264 coder implementation to work on a sole "intra" frame it will behave as a still image compressor. We refer to this coder as H.264-Intra. The big surprise is that it also outperforms previous state-of-art coders such as JPEG-2000 [16]. This is a surprise to many because it was not meant to be an image coder at all. However, results are consistent and unisone.

Sample PSNR results are shown in Fig. 3, were we show PSNR curves comparing H.264-Intra and JPEG-2000 for a few images (popular "Lena" and "Goldhill" gray images along with image "compound1" from the JPEG-2000 test set) and a plot of PSNR differences between methods. Results clearly show that H.264-Intra consistently outperforms JPEG-2000. We carried tests for many other images and the results are consistent.

Gains of the H.264-Intra over JPEG-2000 are typically in the order of 0.25dB to 0.5dB in PSNR for pictorial images. By themselves these results are excellent improvements. However, the H.264-Intra seems to have an extra capacity of adapting itself to heterodox content. For compound images (mixed pictures and text) the PSNR gains are more substantial, even surpassing the mark of 3 dB improvement! Without a doubt the H.264-Intra sets a new level of performance that will likely influence future still image coding standards. Furthermore,

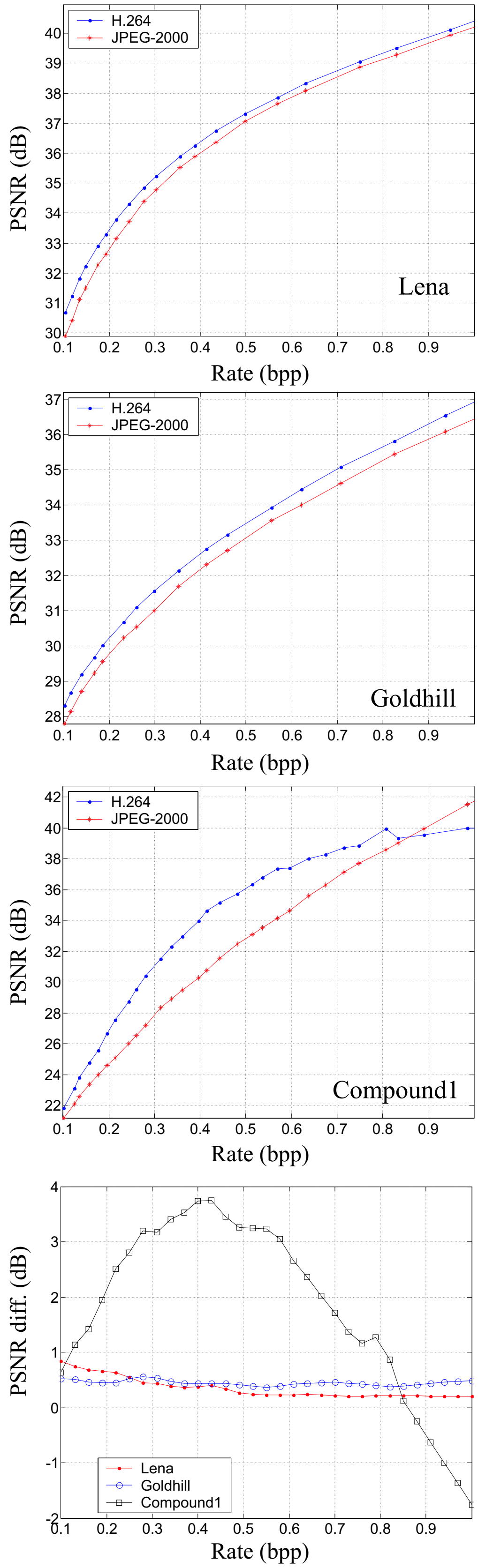

Fig. 3. PSNR (dB) plots comparing H.264-Intra against JPEG-2000 for a few monochrome images. At the bottom, the differential PSNR (relative to JPEG-2000) is shown for all the images to facilitate the comparison between coders.

since it is not a dedicated image coder, many improvement on the H.264-Intra can still be made to make it more efficient either in terms of distortion or complexity. Improvements and add-ons to the H.264-Intra are a new path for research in image compression.

IV. H.264/AVC WITHOUT MOTION ESTIMATION

The intraframe prediction modes of the H.264/AVC are so efficient that not only it enables an efficient still image coder, but it also allows one to completely turn off the motion estimation (ME) in H.264/AVC and yet sustain manageable losses. H.264/AVC is a hybrid codec where temporally predictive coding is associated with spatial transform coding, just like all previous video coding standards before it. In order to have a hybrid coder without ME we assume all motion vectors to be zero, but we are free to use any frame available as a reference, i.e. in H.264/AVC we are not constrained to use only the immediate neighbour frame. Furthermore, in the case of B-frames, one can use direct average of other frames, including the previous and next frames in the sequence.

In effect for coders such as MPEG-1 we would obtain a simple temporal DPCM, where the prediction error is spatially transformed. In H.264/AVC it is not quite so because intra-frame prediction is also efficient. For each macroblock, the coder tests all intra-frame prediction modes, then it tests the interframe prediction using the previous macroblock as a predictor. One can also set some B-frames and for those also test the average of previous and next frames. In the end, the encoder selects the prediction that yields the least error (if we assume the quantizer parameter Q_p is fixed). This is the major difference: interframe prediction is just one mode of prediction and all modes are tested. In H.264/AVC for popular test sequences, one can easily find macroblocks that are better estimated by its spatial neighbors rather than by ME. As a result for many sequences, H.264/AVC without any ME still outperforms video coders with full motion estimation such as MPEG-2.

In order to explain why the H.264/AVC without ME (H.264/WME) works so well, we may divide the frame regions (or macroblocks) into three types depending on the motion: still (S) regions without motion; cut (C) regions where motion is very abrupt or does not exist such as sharp scene cuts, occlusions etc., i.e. macroblocks where ME will likely fail; motion (M) regions where the moderate motion is manageable by a good ME procedure. In S regions, there is no motion and our zero motion vector interframe prediction approach works perfectly. In C regions motion estimation would fail anyway and intraframe prediction would be selected regardless of using or not ME. Only at M regions the H.264/AVC would benefit from ME. Where there is motion, the interframe prediction will likely outperform intraframe one. However, in these regions, H.264/WME will be forced to chose among intraframe prediction modes.

So, the performance of the H.264/WME is directly tied to the number of macroblocks that fall within a region that can be characterized as M. If we compare with a coder like MPEG-2 the larger the number of macroblocks labelled

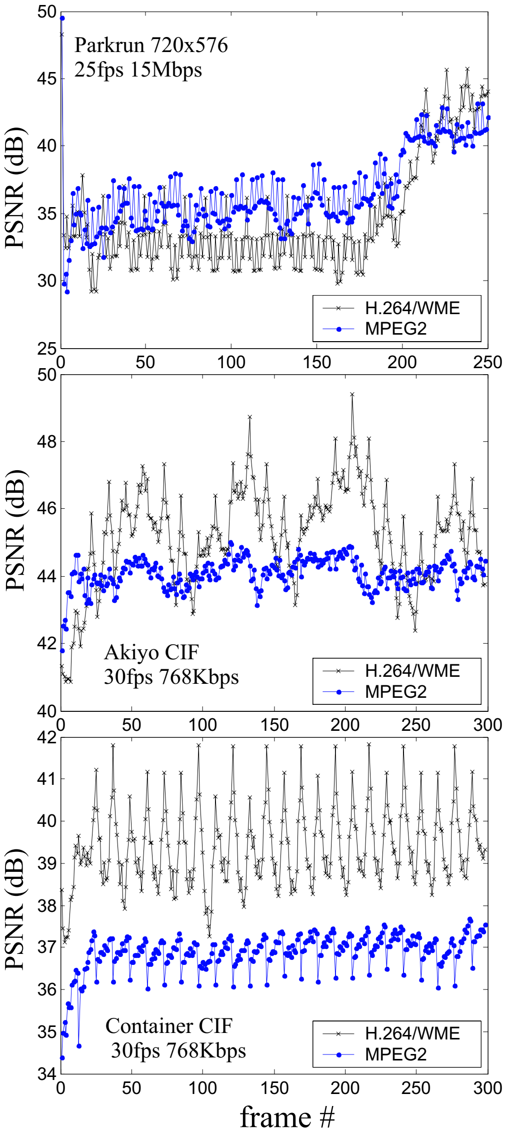

Fig. 4. PSNR (dB) plots from frame to frame comparing H.264/WME and MPEG-2 for a few color sequences. PSNR was computed in the YCbCr space.

as M in the scene, the better MPEG-2 will be, compared to H.264/WME. For contents with low global motion like videoconference, newscast, etc., the H.264/WME coder would outperform MPEG-2.

In Fig. 4 we show a few plots relating the temporal PSNR in dB (taken frame to frame) of sequences compressed with H.264/WME and MPEG-2. The first plot refers to a standard definition sequence "parkrun" which contains strong camera motion and one can say that most of the macroblocks in that scene fall within the M class. This is sort of a worst case scenario. As a result one can see from Fig. 4 (top) that MPEG-2 outperforms H.264/WME by a small margin. For other cases such as "container" and "Akiyo" sequences we do not have constant global movement and the balance of macroblock movement types shifts so that H.264/WME largely outperforms MPEG-2. In Fig. 4 (center and bottom) we can see the large difference between H.264/WME and MPEG-2 in favor of the former. Both of these sequences have CIF/30fps

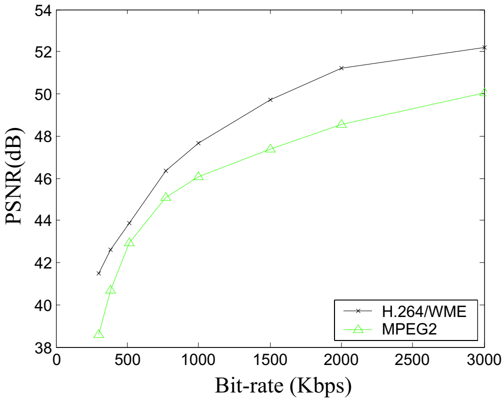

Fig. 5. Average PSNR (dB) plots comparing H.264/WME and MPEG-2 for color sequence "Akiyo" at several bit-rates. PSNR was computed in the YCbCr space.

resolution. The results for these sequences are not accidental nor qualitatively dependent on bit-rate. To show that, Fig. 5 depicts an average PSNR (dB) plot for "Akiyo" at several bit rates, also comparing H.264/WME and MPEG-2. This figure shows the large advantage of H.264/WME over MPEG-2.

V. Conclusions

Our motivation to write this paper was not to present a new technique, but to highlight a couple of extra (fringe) benefits of the recent video coding standard H.264 that were not explored in the literature. These benefits were not made by design, but they are not accidental either, since they are fruits of a well conceived compression scheme.

The H.264-Intra can be used as a new benchamrk in image compression. It easily outperforms JPEG-2000. Of course, H.264/AVC outperforms MPEG-2 by a large margin, nevertheless the H.264/WME can be used as a lower cost version to rival MPEG-2 in complexity too.

Research directions are to simplify H.264-Intra and H.264/WME to make them competitive with JPEG-2000 and MPEG-2 in terms of complexity, but without sacrificing the performance. Another research direction is to use a H.264-Intra as a benchmark and try to improve it for upcoming image processing standards.

References

[1] Final Draft International Standard of Joint Video Specification. ITU-T Recommendation H.264, ISO/IEC 14496-10 AVC. March 2003.

[2] I. E. G. Richardson, *H.264 and MPEG-4 Video Compression*, Wiley, 2003.

[3] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. on Circuits and Systes for Video Technology*, Vol. 13, No. 7, pp. 560–576, July 2003.

[4] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Trans. on Circuits and Systes for Video Technology*, Vol. 13, No. 7, pp. 657–673, July 2003.

[5] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards" *IEEE Trans. on Circuits and Systes for Video Technology*, Vol. 13, No. 7, pp. 688–703, July 2003.

[6] G. Sullivan et al., "The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions," *Proc. SPIE Conference on Applications of Digital Image Processing XXVII*, Special Session on Advances in the New Emerging Standard: H.264/AVC, August, 2004.

[7] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: tools, performance, and complexity," *IEEE Circuits and Systems Magazine*, pp. 7–28, (First Quarter) March 2004.
[8] J. L. Mitchell, W. B. Pennebaker, C. Fogg, and D. J. LeGall, *MPEG Video Compression Standard*," Chapman and Hall, New York, USA, 1997.
[9] B. G. Haskell, A. Puri, A. N. Netravalli, *Digital Video: An Introduction to MPEG-2*, Chapman and Hall, NewYork, USA, 1997.
[10] *The MPEG 4 book*, F. Pereira and T. Ebrahimi Editors, IMSC Press, 2002.
[11] http://www.divx.com
[12] ITU-T Rec. H.263, "Video coding for low bit rate communication," Version 1: Nov. 1995, Version 2: Jan. 1998, Version 3: Nov. 2000.
[13] ISO/IEC JTC1, "Coding of audio-visual objects Part 2: Visual," ISO/IEC 14496-2 (MPEG-4 visual version 1), Apr. 1999; Amendment 1 (version 2), Feb. 2000; Amendment 4 (streaming profile), Jan. 2001.
[14] D. Marpe, V. George, and T. Wiegand, "Performance comparison of intra-only H.264/AVC HP and JPEG2000 for a set of monochrome ISO/IEC test images," Contribution JVT ISO/IEC MPEG and ITU-T VCEG, JVT M-014, Oct. 2004.
[15] D. Marpe, V. George, H. L. Cyconb, and K. U. Barthel, "Performance evaluation of Motion-JPEG2000 in comparison with H.264/AVC operated in pure intra coding mode," Proc. SPIE , Vol. 5266, pp. 129-137, Feb. 2004.
[16] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice,* Kluwer Academic, 2002.