

INTRA PREDICTION VERSUS WAVELETS AND LAPPED TRANSFORMS IN AN H.264/AVC CODER

Rafael G. de Oliveira and Ricardo L. de Queiroz

Departamento de Engenharia Elétrica
Universidade de Brasília
Brazil

rafael@image.unb.br

queiroz@ieee.org

ABSTRACT

H.264/AVC is the latest video coding standard and, among other things, it uses a DCT-like transform and intra prediction modes. We are studying the possibility of replacing the modified DCT stages by lapped and wavelet transforms. Since those transforms have overlap, intra-frame prediction is not feasible, because of its block-recursive nature. Hence, intra-frame prediction is turned off. In essence, this paper contains a comparison among lapped (wavelet) transforms and linear prediction schemes, within the AVC framework. Results indicate that lapped transforms can outperform the intra prediction scheme, specially for high definition images.

Index Terms— H.264/AVC, wavelets, intra prediction, lapped transforms, image coding

1. INTRODUCTION

In 2004, the Joint Video Team adopted a new video coding standard, the H.264/AVC[1]-[3]. It is now considered the state of the art in video compression, which was accomplished by the adoption of a number of innovative features like: quarter pixel motion estimation precision, arbitrary reference frame, variable size macroblock partition, in-loop deblocking filter, intra-frame prediction, context-adaptive arithmetic coding (CABAC) and variable block size transforms (4×4 and 8×8).

Even though H.264/AVC was meant to be a video coding standard, when used in intra frame coding, it works as a formidable still image coder. Surprisingly, it outperforms the JPEG 2000 [4]-[6], considered the state of the art in image compression standards [7]. Most of the performance improvement may be attributed to the last three of the above cited features, combined with rate-distortion optimization (RDO).

Since early development on lapped transforms [8], they are being compared to the DCT and wavelets. For example [9], there are many comparisons among transforms such as

the discrete cosine transform (DCT) [10], the 9/7-tap wavelet transform (WT) used in JPEG-2000 [7],[11], and lapped transforms (LT) such as the lapped orthogonal transform (LOT) [12], the generalized LOT (GenLOT) [13], and the generalized lapped biorthogonal transforms (GLBT) [14]. The results point to a small but important improvement achieved by using lapped transforms for all codecs tested. The comparisons were carried for JPEG [15] and SPIHT [16] coders using proper adaptations [17],[18]. Recently, WT vs. LT comparisons have also been carried in the context of JPEG-2000[19].

Despite all these comparisons, there is not much information in the literature available about the comparison between intra prediction and lapped transforms or wavelets. This is the focus of this work. Only gray scale images and intra-coded frames are considered here.

2. INTRA PREDICTION, LAPPED TRANSFORMS AND WAVELETS AS COMPETITORS

In block transforms like the DCT, only pixels inside the block are considered for the transform. That may be interesting for many reasons. However, it does not take advantage of correlation between neighbor blocks and may cause blocking effects, a discontinuity on the edge of the blocks due to quantizations errors. In H.264/AVC, this last effect is avoided by the use of a deblocking filter.

In order to exploit the redundancy among neighbor blocks, H.264/AVC has intra-prediction modes that use previously encoded pixels to predict the current block. Only the residue, that is, the difference between the predicted and the actual pixels, is encoded. The process is recursive. This hinders parallel processing. However, it makes possible to vary the size of the block. H.264/AVC has 9 modes of predictions for 4×4 -pixel blocks, 9 for 8×8 blocks and 4 for 16×16 blocks.

On the other hand, lapped transforms and wavelets already have supports larger than the block being encoded. Their filters are non-recursive and have finite impulse response, which allows for parallel processing. In Fig. 1 it is

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, Brasil, under grant 474912/2006-0

illustrated the support available to encode one block using wavelets, lapped transforms and intra prediction. The transforms are computed just once, while the intra-prediction tests are repeated exhaustively for each mode for every macroblock or subpartition of it.

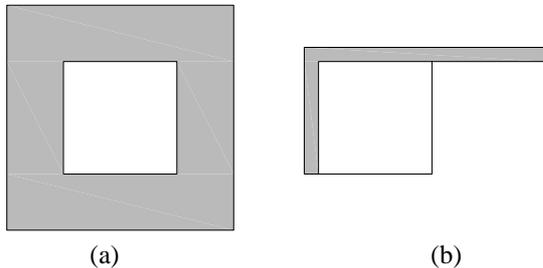


Fig. 1. Support available for encoding one block. (a) wavelets and lapped transforms (b) intra-prediction.

3. LAPPED TRANSFORMS AND WAVELETS IN H.264/AVC

In order to use transforms other than the modified DCT, a few modifications had to be carried.

- Intra prediction, which is recursive and incompatible with lapped transforms and wavelets, was disabled.
- 8x8 transform mode was forced to all macroblocks.
- The quantization tables were changed (quantization steps were made equal for all coefficients). This was necessary because the transform and the quantization are performed together in H.264/AVC.
- RDO was disabled in order to make a fair comparison between intra-prediction and the other transforms. It is being studied the use of RDO to LT and WT.

For M -band transforms, like LT or DCT, a simple substitution of the coefficients is sufficient. In order to improve the performance and to exploit the redundancy among neighbor blocks, an additional 2×2 Hadamard decomposition was performed on the DC coefficients of each macroblock. For HD images, we also tested a 4×4 Hadamard decomposition to DC coefficients of groups of 2×2 macroblocks.

In dyadic transforms, like wavelets, as the coefficients are grouped by subbands, some reordering of the coefficients is necessary. Coefficients are grouped by location as shown in Fig. 2, where four level decompositions are performed, and the resulting blocks are encoded. In a similar way, for HD images, five levels of wavelet decomposition were tested.

The tests were performed at JM 10.2 using the High profile and CABAC as the entropy coder.

4. RESULTS

The rate distortion curves (PSNR in dB vs. bit rate in bits/pel or bpp) for the popular 512×512 -pixel image Barbara are

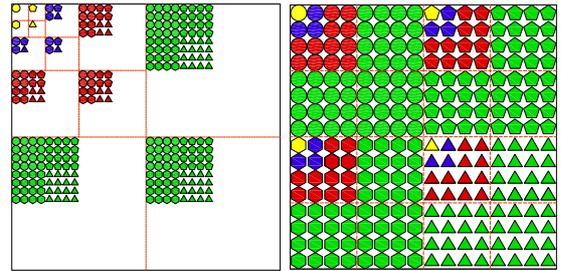


Fig. 2. Coefficients reordering for dyadic transforms.

shown in Fig. 3. Since it is difficult to compare the different curves in Fig. 3, we re-plotted the data as differential PSNR relative to the standard H.264/AVC without RDO. Results are shown in Fig. 4.

The results for three other popular images, Lena, Baboon and Goldhill, of the same dimensions (512×512 pixels) are shown in Figs. 5, 6, and 7, respectively. We also tested high definition (1920×1080 pixels) video sequences such as Riverbed, Rush_hour, Station2 and Sunflower. Only the first frame was encoded and the results are presented in Figs. 8 through 11.

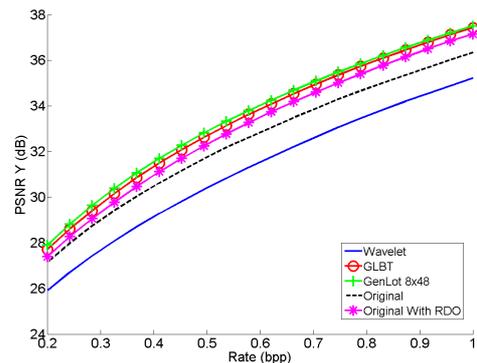


Fig. 3. Plots PSNR curves for image Barbara. Rate is given in bits per pixel (bpp).

5. CONCLUSION

The results corroborate the previous knowledge that lapped transforms outperform wavelets in image coding for many scenarios. Among all tested images only in the “Sunflower” case wavelets outperformed lapped transforms. The results are consistent with results presented in the literature using other coders. It is possible that the AVC context models do not work properly for dyadic transforms. As expected, in the HD images, the performance increases with the number of decomposition levels.

For high definition frames, the results surprisingly show us that lapped transforms can also outperform the intra-

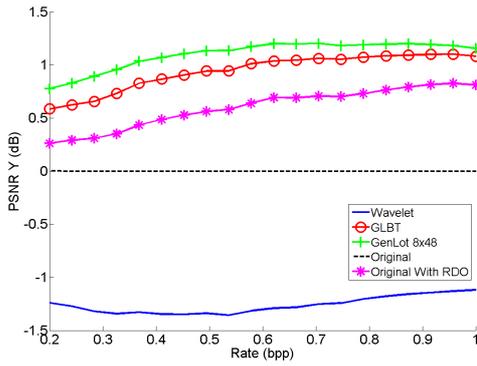


Fig. 4. Differential PSNR curves for image Barbara relative to the performance of the original H.264/AVC.

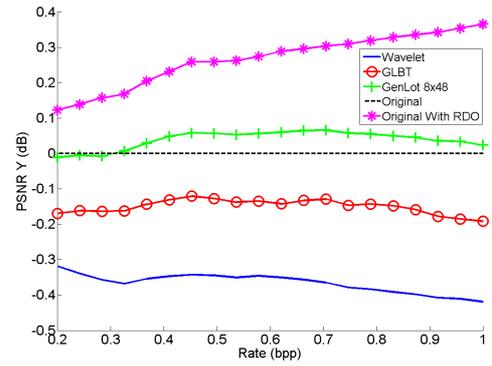


Fig. 6. Differential PSNR curves for image Baboon relative to the performance of the original H.264/AVC.

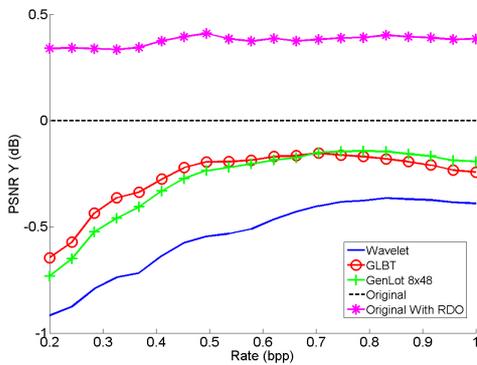


Fig. 5. Differential PSNR curves for image Lena relative to the performance of the original H.264/AVC.

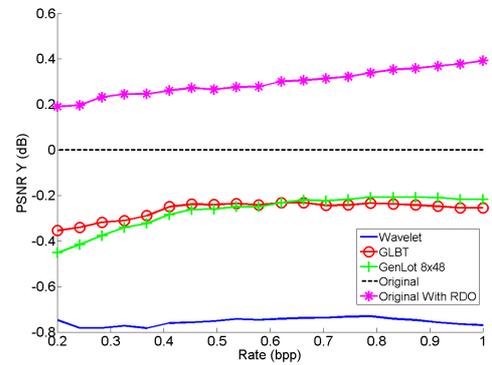


Fig. 7. Differential PSNR curves for image Goldhill relative to the performance of the original H.264/AVC.

prediction scheme used in H.264/AVC. For all HD pictures LT yield better results than the original H.264/AVC (without RDO). For “Riverbed” and “Sunflower” LTs outperforms the standard using RDO and, for the other two images, “Rush_hour” and “Station2”, they present performance similar to H.264/AVC with RDO. The differences observed on the last two can be compensated by the reduction in computational complexity, since all intra modes must be computed before one is chosen, while the LT is only performed once. It is important to remember that RDO was not used for LTs or wavelets so that our results can be further improved.

For the 512×512 pixel images, lapped transforms do not present a clear advantage over intra prediction. The better performance of intra predictions scheme may be caused by its ability to adapt to these images using smaller blocks. The performance of lapped transforms for the image “Barbara” is well known for its typically good performance.

As for testing other transforms for inter-frame coding, this is ongoing research and results will be reported at a future opportunity.

6. REFERENCES

- [1] *Final Draft International Standard of Joint Video Specification. ITU-T Recommendation H.264, ISO/IEC 14496-10 AVC.* March 2003.
- [2] T. Wiegand, G. J. Sullivan, G. Bjntergaard, 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, pp. 560-576, July 2003.
- [3] I. E. G. Richardson, *H.264 and MPEG-4 Video Compression*, Wiley, 2003.
- [4] R. L. de Queiroz, R. S. Ortis, A. Zaghetto, and T. A. Fonseca, “Fringe benefits of the H.264/AVC, Proc. of Intl. Telecom. Symp., Fortaleza, Brazil, pp. 208-212, Sep. 2006.
- [5] 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.
- [6] D. Marpe, V. George, H. L. Cycon, 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.
- [7] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*, Kluwer Academic, 2002.

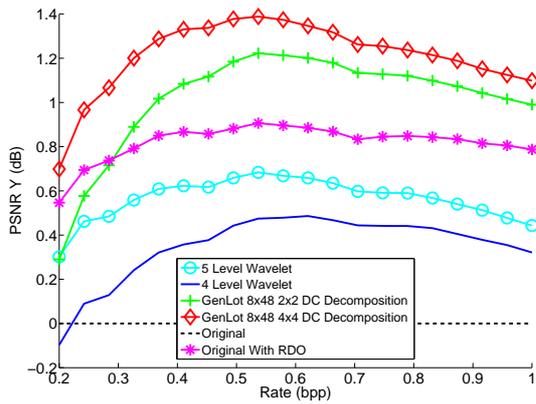


Fig. 8. Differential PSNR curves for first frame of the high definition sequence Riverbed relative to the performance of the original H.264/AVC.

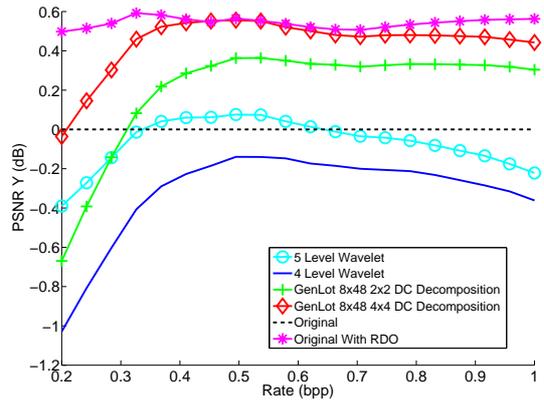


Fig. 10. Differential PSNR curves for first frame of the high definition sequence Station2 relative to the performance of the original H.264/AVC.

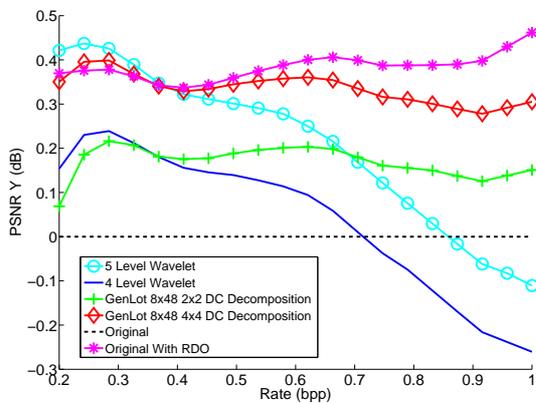


Fig. 9. Differential PSNR curves for first frame of the high definition sequence Rush_hour relative to the performance of the original H.264/AVC.

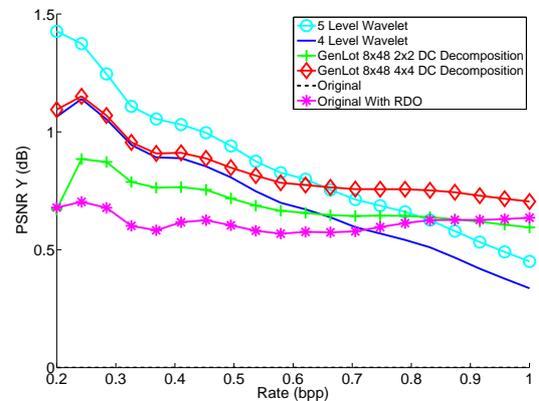


Fig. 11. Differential PSNR curves for first frame of the high definition sequence Sunflower relative to the performance of the original H.264/AVC.

- [8] H. S. Malvar, *Signal Processing with Lapped Transforms*. Norwood, MA: Artech House, 1992.
- [9] R. L. de Queiroz and T. D. Tran, *Lapped Transforms for Image Compression*, Chap. 5 in *The Handbook on Transforms and Data Compression*, edited by K. R. Rao and P. Yip, CRC Press, 2001.
- [10] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, and Applications*, San Diego, CA : Academic Press, 1990.
- [11] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley, MA: Wellesley-Cambridge, 1996.
- [12] H. S. Malvar and D. H. Staelin, "The LOT: transform coding without blocking effects," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-37, pp. 553-559, Apr. 1989.
- [13] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao, "GenLOT: generalized linear-phase lapped orthogonal transforms," *IEEE Trans. on Signal Processing*, Vol. 44, pp. 497-507, Apr. 1996.
- [14] T. D. Tran, R. L. de Queiroz, and T. Q. Nguyen, "Linear phase perfect reconstruction filter bank: lattice structure, design, and application in image coding," *IEEE Trans. on Signal Processing*, Vol. 44, Jan. 2000.
- [15] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Com-*

pression Standard," New York, NY: Van Nostrand Reinhold, 1993.

- [16] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans on Circuits Syst. Video Tech.*, vol. 6, pp. 243-250, June 1996.
- [17] R. de Queiroz, C. Choi, Y. Huh, and K. R. Rao, "Wavelet transforms in a JPEG-like image coder," *IEEE Trans. on Circuits and Systems for Video Tech.*, Vol. 7, pp. 419-424, Apr. 1997.
- [18] T. D. Tran and T. Q. Nguyen, "A progressive transmission image coder using linear phase uniform filter banks as block transforms," *IEEE Trans. on Image Processing*, Vol. 8, pp. 1493-1507, Nov. 1999.
- [19] R. G. de Oliveira, and R. L. de Queiroz, "Lapped transforms in a JPEG 2000 coder," *Simpósio Brasileiro de Telecomunicações, Recife, PE, Brazil, Sep. 2007*