

WAVELET-BASED ENCODING FOR HD APPLICATIONS

Livio Lima, Francesca Manerba, Nicola Adami, Alberto Signoroni, Riccardo Leonardi

DEA - University of Brescia, via Branze 38, I-25123 Brescia, Italy
email: {firstname.surname}@ing.unibs.it

ABSTRACT

In the past decades, most of the research on image and video compression has focused on addressing highly bandwidth-constrained environments. However, for some applications of high resolution and high quality image and video compression, as the case of High Definition Television or Digital Cinema, the primary constraints are related to quality and flexibility. This paper presents a comparison, for video sequence coding, between scalable wavelet-based video codecs and the state of the art in single point encoding and it investigates the performances in terms of compression efficiency obtained using temporal decomposition with respect to pure intra coding.

1. INTRODUCTION

Concerning HD formats, two different applications can be proposed: broadcast and not broadcast. Broadcast applications, as the available bandwidth is limited, are not suitable for a high quality scenario. However there are many non broadcast applications that use HD format for which the available bandwidth is more than the broadcast dedicated one. Examples of applications that will contribute to improve the HD diffusion can be: movie production, Digital Cinema - movie distribution and exhibition, HD-DVD, internet streaming and distribution, HD video games, medical, military and surveillance applications. Let us consider for example the new optical devices such as for example, HD-DVD or BlueRay: nowadays, it is possible to find BlueRay disks of 200 GB that can be used with both BlueRay devices and the new PlayStation3 and Xbox360.

In the domain of non broadcast applications we compare scalable wavelet-based video codecs and the state of the art in single point encoding demonstrating that wavelet-based codecs can give comparable results. The aim of this work is also to show that applying a motion compensated temporal filtering it is possible to obtain better performances with respect to pure intra coding even with constant quality constraints. In Section 2 a brief overview of STP-tool wavelet-based video codec is presented; in Section 3 a comparison with other codecs in image coding applications is presented. Then in Section 4 a possible scenario where scalable coding could be used is described; in Section 5 the results of video coding applications

are presented and finally in Section 6 some conclusions are drawn.

2. THE STP-TOOL CODEC

The codec used during the tests, and developed in the Telecommunication Laboratories of the University of Brescia, is an implementation of the STP-tool scheme, which could be also found in MPEG VidWav reference system [1]. The STP-tool is a 2D+t+2D scalable architecture with an original inter-scale prediction mechanism.

The STP-tool software implementation tested in this work employs a temporal module with a Hierarchical B Picture temporal decomposition and unconstrained MCTF. The entropy encoder used to compress the texture is an implementation of the wavelet-based EMDC algorithm (2D) described in [2] which produces a progressive bit-stream with embedded rate distortion optimization. The coding efficiency of this codec is comparable with the one offered by JPEG2000 [3] so the results obtained in this work can be expected also in the case the EMDC module is replaced with JPEG2000 entropy encoder. A more exhaustive comparison between STP-tool and JPEG2000 can be found in [4], so we will not perform a direct comparison between the two tools but only between a wavelet-based and non wavelet-based codecs.

3. IMAGE CODING

Wavelet technology has been successfully applied in video coding only in the last few years and still now, only in some cases it can be considered comparable to last generation standards such as H.264/AVC [5]. This is due not only to less work spent on wavelet video coding with respect to block-transform-based technology but also on the resolution and quality of video material. Let us consider the performances of JPEG2000, STP-tool and H.264/AVC, with FRExt extension, for still image coding at different resolutions. In table 3 the mean encoding gain of H.264/AVC compared with wavelet encoders is presented for different image classes. A positive gain means that H.264 outperforms wavelet codecs. Based on this test and on previous works [6][7], we can conclude that: in video coding at lower resolutions (QCIF and CIF) H.264/AVC outperforms JPEG2000, at medium resolutions

still image coding (from 200k to 1M pixel) and 4CIF video the performances are comparable, in still image coding and video at very high resolutions STP-tool and JPEG2000 outperform H.264/AVC. Based on this consideration and on the fact that the block based motion estimation and the encoding of motion compensation residue (that has different features from natural images) help video encoders based on block transforms (like DCT), it is difficult to obtain at standard resolutions a wavelet based video codec with comparable performances with respect to last generation standards, such as H.264/AVC; but for the same reasons we can reasonably expect to obtain comparable performances for HD applications.

Image class	gain [dB]
QCIF and CIF frame	+1/1,5
4CIF res. frame	+0/0,5
image at medium res. (<1Mpixel)	0
image at high res. (>1Mpixel)	-0,5/1
HD res. frame	-0,5

Table 1. Coding gain of H.264/AVC compared with wavelet-base encoders.

4. POSSIBLE APPLICATION SCENARIO

In this section a possible application scenario involving HD applications for home multimedia entertainment is presented: it targets 2k as maximum spatial resolution due to current hardware limitations but it can be easily up scaled to higher resolution content (see Figure 1). For home applications, image sequences are required to be highly compressed and easily delivered to different devices either through wireless or wired communication network. Even if new storage and networking technologies are going to be capable of handling big amounts of data, it will still be crucial to efficiently compress multimedia content in order to improve the storage and bandwidth requirements of the system. For example, according to the latest DCI specifications [8], encoding three hours of 2k resolution content at 24 fps and rate of 250 Mbit/sec with JPEG2000 requires around 370 GB of storage space. By improving compression efficiency lower bit-rates can be considered in order to obtain the required quality. In the given scenario compressed data available in the content repository will be distributed to different devices through a network. A new paradigm of content distribution will be used thanks to the features provided by scalable bit-streams. Instead of having a separate streaming for each single device, the bit-stream related to a specific content is transmitted only once and each device takes only packets containing information required for decoding. For example, an HDTV can decode full bitstream to display the highest operating point (frame-rate, resolution, quality), to display video on a projector we can consider scalable bit-stream containing video of high resolution, frame-

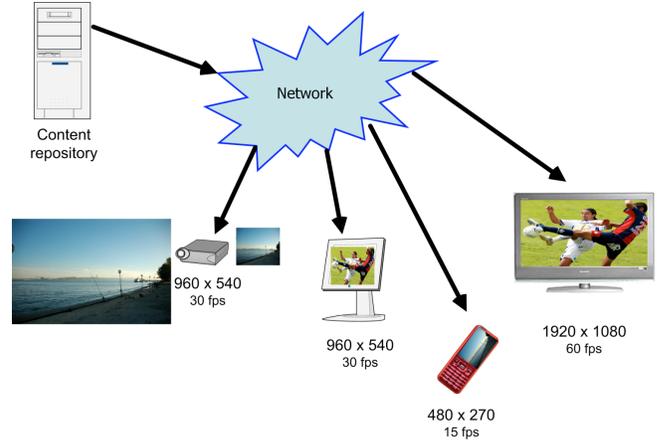


Fig. 1. Scalable video coding in multimedia entertainment.

rate and quality (960×540 , 30 fps, PSNRY > 38dB) while mobile device will take only packets containing low resolution and frame-rate video (480×270 , 15 fps).

5. RESULT EVALUATION

In this section the results obtained coding two HD sequences in the format 1920×1080 , that is the closest to the 2k format, at 50 fps are presented. The sequences, CrowdRun and Old-TownCross, have been obtained from the European Broadcasting Union website [9] where are classified as sequences with respectively high and medium coding complexity. The executed tests are presented with their rate-distortion curves only considering the luminance component. Additionally, in order to best fit high quality requirements, the encoded bit-streams has been generated to minimize the PSNR fluctuation between adjacent frames in the reconstructed sequences, even if when temporal filtering is applied, this constraint penalizes the coding efficiency, at least in terms of average PSNR.

In order to show the flexibility of our wavelet encoder for HD scenario previously described, we have performed three different tests in combined scalability configuration [10], so a sequence with resolution 960×540 , which constitutes the base layer, has been generated with a wavelet filter. The used combined scalability scheme constraints to extract the base layer maximum quality point from each level of the enhancement layer. It is clear that this particular configuration penalizes the decoding step of the lower quality points of the enhancement layer, especially in the case of a small bit-rate difference with the higher quality point of the base layer.

In the first test, in order to point out the gain that can be reached exploiting the video temporal correlation, the base-layer results obtained in intra mode are compared with the results obtained using a hierarchical decomposition structure with 4 frame GOPs and they are compared with the JM11.0 software [11], which is H.264/AVC reference software.

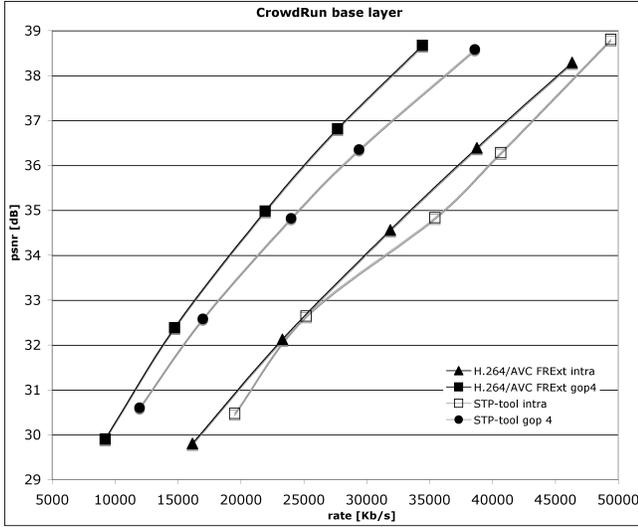


Fig. 2. Comparison between STP-tool and H.264/AVC FRExt for the sequence “CrowdRun” at 960×540 resolution.

In the second test we compare STP-tool in combined scalability and JM11.0 at 1920×1080 resolution considering pure intra coding. For STP-tool we use the base layer both in intra mode and with temporal decomposition as in test 1. Since H.264/AVC does not allow the scalable coding the sequence at 1920×1080 resolution has been coded separately from the base layer of test 1.

In the last test we use the STP-tool in combined scalability with the base layer at 25 fps and the full resolution at 50 fps, in order to show the gain obtained using temporal decomposition both in the base layer and at full resolution.

5.1. Test 1

In Figure 2 the curves obtained for the “CrowdRun” sequence at 960×540 resolution are shown. In the case of our wavelet-based STP-tool the sequence has been coded only once and an Extractor Tool has been used to extract, starting from the same coded stream, different streams with variable PSNR from 30dB to 40dB while in the case of H.264/AVC each working point has to be coded separately. It is possible to notice that both the codecs can better compress the stream when exploiting temporal redundancy using motion compensation.

In Figure 3 the same test has been performed on the “OldTownCross” sequence at 960×540 resolution. In this case it is possible to notice that STP-tool performs near H.264/AVC FRExt and the same conclusion as in the latter case about temporal redundancy can be drawn, but with a better coding gain using temporal decomposition compared with “CrowdRun” sequence. It is important to remember that FRExt provides single point encoding so we are comparing a tool that enables SNR scalability for the base layer with a non scalable codec.

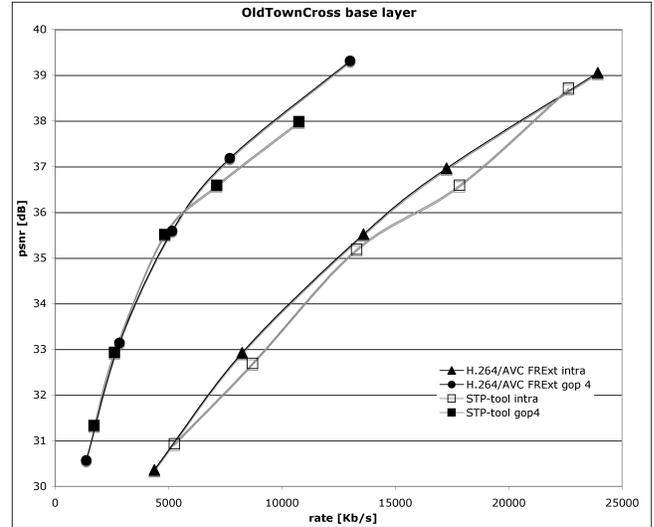


Fig. 3. Comparison between STP-tool and H.264/AVC FRExt for the sequence “OldTownCross” at 960×540 resolution.

5.2. Test 2

The two sequences used for the previous test have also been encoded at 1920×1080 resolution. In the case of STP-tool, intra coding has been performed using both intra-coded and inter-coded base layer, and the results are shown in Figure 4. In intra-coded base layer the loss in coding efficiency of wavelet encoders with respect to H.264/AVC is due to scalable scheme, that decreases the performances, while it can be interesting to note that, in inter-coded base layer case, a gain in coding efficiency with respect to intra-coded case is present due to the coding gain in the base layer obtained with temporal decomposition. So in this case we have reached comparable performances with H.264/AVC FRExt maintaining the advantages of intra-coding, as for example low delay encoding/decoding and random access.

5.3. Test 3

In the last test STP-tool in combined scalability has been used with the base layer and the full resolution sequence at different frame rates. This can better simulate the application scenario presented in section 4 where different capability devices are present. In this particular configuration we show the gain in coding efficiency using the temporal decomposition both in the base layer and at full resolution. In Figure 5 we show the results of “CrowdRun” sequence at 1920×1080 at 50 fps obtained starting from a 25 fps base layer and even in this case we can show that the result obtained with a scalable wavelet-based video codec is comparable with the H.264/AVC standard.

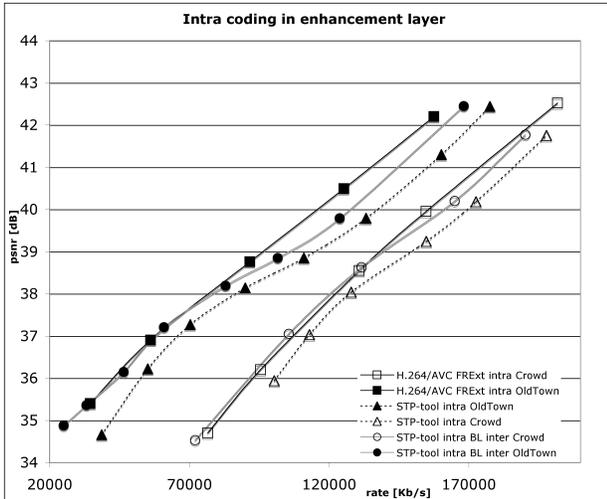


Fig. 4. Comparison between STP-tool and H.264/AVC FRExt at 1920×1080 resolution.

6. CONCLUSIONS

The aim of this work is to test scalable wavelet-based video coding on new HD video material. The obtained results have been produced under redefined test conditions in order to better fit HD coding needs. Concluding, it has been shown how the exploitation of temporal redundancy can significantly decrease the compressed bit-rate even under unfavorable coding constraints such as the near constant quality. Moreover, it has been shown that for HD applications with scalable wavelet technology, it is possible to obtain comparable performances with respect to single point encoding H.264/AVC standard both in intra and inter case, and this encourages further work in wavelet video coding. Although results have been obtained with experimental software, they can surely be expected for a similar architecture implementation which employ JPEG2000 compliant technology for texture compression. It is our intention to complete tests for combined scalability by also considering the adoption of fast motion estimation techniques, which remains the bottleneck in terms of time needed to compress the original sequence.

7. REFERENCES

- [1] R. Leonardi, T. Oelbaum, and J.-R. Ohm, "Status report on wavelet video coding exploration," Tech. Rep. N8043, ISO/IEC JTC1/SC29/WG11, 76th MPEG Meeting, Montreux, Switzerland, April 2006.
- [2] N. Adami, E. Izquierdo, R. Leonardi, M. Mrak, A. Signoroni, and T. Zgaljic, "Efficient wavelet-based video compression," Tech. Rep., ISO/IEC JTC1/SC29/WG1, N3954, Assisi, Italy, July 2006.

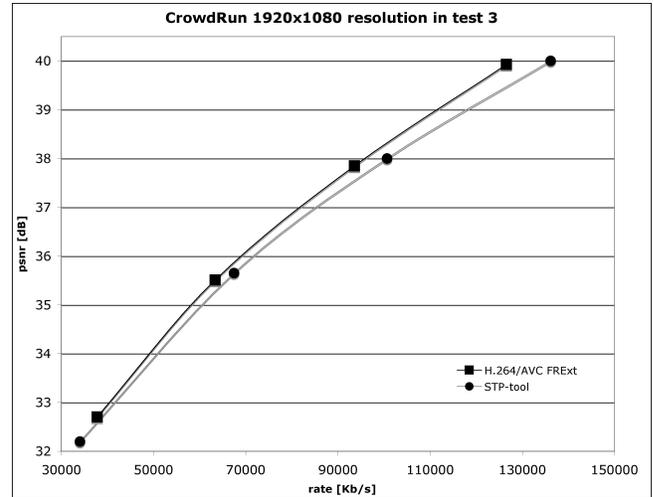


Fig. 5. Comparison between STP-tool and H.264/AVC FRExt for the sequence "CrowdRun" at 1920×1024 resolution.

- [3] ITU-T Recommendation T.800 and ISO/IEC 15444-1, "JPEG2000 image coding system : Core coding system (JPEG2000 part 1)," Tech. Rep., 15444-1, 2003.
- [4] F. Lazzaroni, R. Leonardi, and A. Signoroni, "High-performance embedded morphological wavelet coding," *IEEE Signal Processing Letters*, vol. 10, no. 10, October 2003.
- [5] ISO/IEC 14496-10, "Information Technology - Advanced Video Coding: A codec for video signals which is also called AVC and is technically identical to the ITU-T H.264 standard ISO/IEC," Tech. Rep., 14496-10, Geneva, Switzerland, 2003.
- [6] D. Marpe, S. Gordon, and T. Wiegand, "H.264/MPEG4-AVC Fidelity Range Extension : Tools, profiles, performance, and application areas," in *Proc. of IEEE ICIP05*, Genova, Italy, September 2005.
- [7] M. Ouaret, F. Dufaux, and T. Ebrahimi, "On comparing JPEG2000 and intraframe AVC," in *Proc. of the SPIE Applications of Digital Image Processing*, August 2006.
- [8] Digital Cinema Initiatives, "Digital cinema system specification v1.0," http://www.dcinemovies.com/-DCI_Digital_Cinema_System_Spec_v1.pdf.
- [9] "European broadcasting union," http://www.ebu.ch/en/technical/hdtv/test_sequences.php.
- [10] M. Wien and H. Schwarz, "Testing conditions for svc coding, efficiency and jsvm performance evaluation," in *JVT-Q205 Document*, Poznan, Poland, July 2005.
- [11] "JM software," <http://iphome.hhi.de/suehring/tml/>.