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 high bandwidth- constrained environments. However, for high resolution and high quality image and video compression, as in the case of High Definition Television (HDTV) or Digital Cinema (DC), the primary constraints are related to quality and flexibility. This paper presents a comparison between scalable wavelet-based video codecs and the state of the art in single point encoding and it investigates the obtainable compression efficiency when using temporal correlation with respect to pure intra coding.

1. INTRODUCTION

Concerning HD formats, two different scenarios can be considered: broadcast and non broadcast. In broadcast conditions, the available bandwidth is limited, so that they are not suitable for high quality signal reproduction. However there are many non broadcast scenarios that use the HD format for which the available bandwidth is more than the broadcast one. Examples of applications that are contributing to improve the HD development are: movie production, Digital Cinema - movie distribution and exhibit, 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 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 scenarios we compare scalable wavelet-based video codecs and the state of the art in single point encoding demonstrating that wavelet-based codecs can achieve comparable performances. The aim of this work is also to show that by applying motion compensated temporal filtering it is possible to obtain better performance with respect to pure intra coding even under 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 video codecs is reported. Then in Section 4 a possible scenario where scalable coding could be used is described; in Section 5 the experimental results are discussed. Finally in Section 6 some conclusions are drawn.

2. THE STP-TOOL CODEC

The used codec for the performed tests has been designed independently and is an implementation of the STP-tool scheme as described in MPEG VidWav reference software [1]. The STP-tool is a 2D+t+2D scalable architecture with an original inter-scale prediction mechanism.

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

3. IMAGE CODING

Wavelet technology has been successfully applied for video coding only in the last few years and till now, it has shown comparable to last generation standards such as H.264/AVC [5] only in some cases. This is due in part to less effort spent on wavelet video coding with respect to block-transform-based technology. It is also due on the resolution and quality of the coded video material. We shall compare the performance of JPEG2000, STP-tool and H.264/AVC in pure intra mode with FRExt extension, for still image coding at different resolutions. In Table 3 the mean encoding gain of H.264/AVC compared to wavelet encoders is presented for different image classes. A positive gain means that H.264 outperforms the 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 at 4CIF resolutions the performance are comparable. In still image coding and video at very high resolutions STP-tool and JPEG2000 outperform H.264/AVC. Based on these considerations and on the fact that the encoding of motion compensated residual generated by block-based motion estimation (that has different characteristics with respect to natural images) is facilitated by a block based approach(such as the DCT) to handle residual information, it is difficult to achieve at standard resolutions comparable performance with a wavelet based video codec when compared to the last generation video coding standards, such as H.264/AVC.

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 in pure intra mode compared to wavelet-based encoders for image coding.

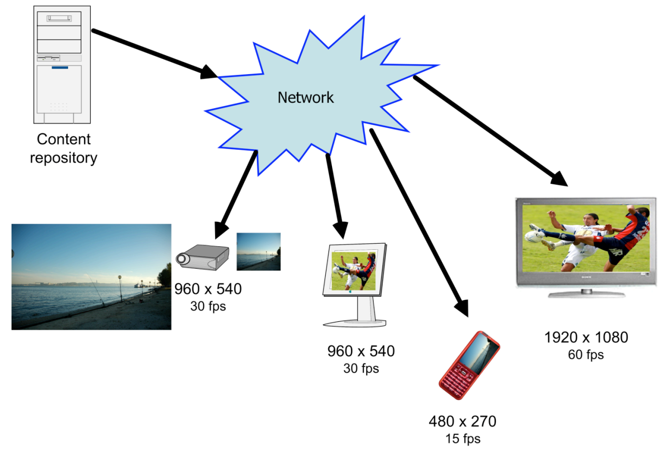

Fig. 1. Scalable video coding in multimedia entertainment.

4. POSSIBLE APPLICATION SCENARIO

In this section a possible application scenario involving HD applications for home multimedia entertainment is presented. Due to current hardware limitations it targets 2k as maximum spatial resolution (see Figure 1), but it can be easily scaled up to higher resolutions. For home applications, image sequences are required to be highly compressed and easily delivered to different devices either through wireless or wired communication network. While new storage and networking technologies are going to be capable of handling bigger amounts of data, content demand is also increasing therefore coding efficiency remains crucial for the efficiency of the multimedia systems. For example, according to the latest DCI specifications [8], encoding with JPEG2000 three hours of 2k resolution content at 24 fps and with a rate of 250 Mbit/sec requires about 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 are distributed to different devices through a network. A new paradigm of content distribution can be used thanks to the features provided by scalable bit-streams. Instead of having simulcast transmission of each single targeted stream, the scalable coded bit-stream is transmitted only once and each device selects or ask (exploiting coding policies) for the only packets containing information required for its own targeted decoding point. For example, an HDTV will decode full bit-stream to display the highest operating point (frame-rate, resolution, quality), while a projector could consider a lower point in terms of resolution, frame-rate and quality (960×540, 30 fps, PSNRY $>$ 38dB). Moreover mobile device shall take only packets containing low resolution and frame-rate video (480×270, 15 fps).

5. RESULT EVALUATION

In this section the results obtained by the coding of two HD sequences with spatial format 1920×1080(which is close to the 2k format) and 50 fps are presented. The sequences, CrowdRun and OldTownCross, have been downloaded from the European Broadcasting Union website [9] where they are classified as sequences with high and medium coding complexity respectively. The executed tests are presented in terms of rate-distortion curves only considering the luminance component. Additionally, in order to best fit high quality requirements, the encoded bit-streams have been generated to minimize the PSNR fluctuation between adjacent frames in the reconstructed sequences. It is worth noting that due to the temporal decomposition, the above constrain usually determine a slight decreasing bias of the coding performance in terms of average PSNR. In order to show the flexibility of our wavelet encoder for the HD scenario (previously described), we have performed three different tests using a combined scalability configuration [10]. A 960×540 lower resolution base layer sequence has been generated with a wavelet filter. A simple bit-stream organization policy has been adopted where enhancement layer are built starting from maximum quality points of the base layer. Even if this can penalize the decoding step of the low (and less interesting) quality points, it keeps the scalable architecture simple. In a first test, in order to point out the gain that can be reached exploiting the video temporal correlation, the base-layer results obtained using intra coding mode are compared with the results obtained using a hierarchical B-picture decomposition structure with 4 frame GOPs, between the proposed architecture and the JM11.0 software [11], which is H.264/AVC reference software. In a second test we concentrate on intra coding performance and compare STP-tool in combined scalability and JM11.0 at 1920×1080 resolution. Here the STP-tool base layer has been tested both in intra mode and with temporal

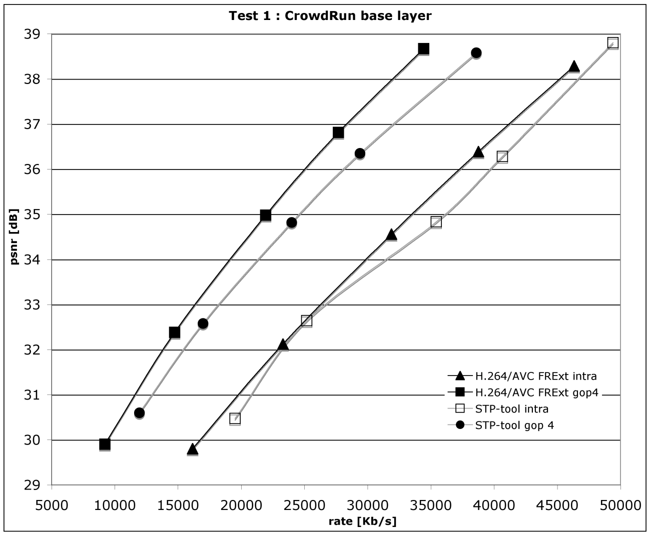

Fig. 2. Comparison between STP-tool and H.264/AVC for the sequence "CrowdRun" at 960×540.

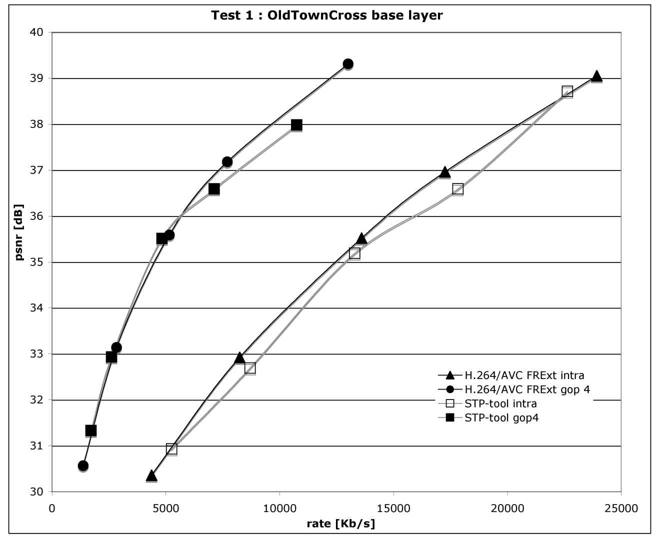

Fig. 3. Comparison between STP-tool and H.264/AVC for the sequence "OldTownCross" at 960×540.

decomposition as in test 1. Since H.264/AVC is not a scalable coder the sequence at 1920×1080 resolution required a new whole coding because cannot be exploited the base layer of test 1. In the last test the STP-tool has been used in combined scalability with a 25 fps base layer and a 50 fps related to full spatial resolution, in order to show the performance obtained with a fully-scalable bit-stream in a realistic scenario.

5.1. Test 1

In Figure 2 the curves related to the "CrowdRun" sequence at 960×540 resolution are shown. In the case of the adopted STP-tool solution the sequence has been coded only once and a single "extractor" has been used to extract 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's important to note that both the codecs have similar gain in exploiting temporal redundancy by using motion compensation.

In Figure 3 the same test have been performed on the "OldTownCross" sequence at 960×540 resolution. In this case STP-tool performance are nearer to that of H.264/AVC FRExt and due to the lower "coding complexity" of the sequence a higher coding gain using temporal decomposition has been achieved with respect to the "CrowdRun" sequence. It is important to remember that FRExt provides single point encoding so we are comparing a tool that enables SNR scalability and the use of the coded sequence as a base layer with a non scalable codec.

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 the above intra-coded and inter-coded lower resolution base layers. The results are shown in Figure 4. With intra-coded base layer the loss in coding efficiency of wavelet encoders with respect to H.264/AVC can be associated to a price to pay to the high flexibility introduced by the full scalability, while it can be interesting to note that with inter-coded base layer, a gain in coding efficiency with respect to the intra-coded case is present due to the coding gain in the base layer. This solution allow to maintain high coding performance and makes it possible to preserve typical intra-coding benefits such as low-delay requirements.

5.3. Test 3

In the last test STP-tool in combined scalability has been used with base layer and full resolution sequences at different frame rates. This can better fit the application scenario presented in section 4 where heterogeneous devices are present. In this configuration temporal decomposition has been used both at the base layer and at full resolution. Figure 5 shows the results of the "CrowdRun" sequence at 1920×1080 and 50 fps obtained upon a 25 fps base layer. Also in this case we can show that the results obtained with our scalable wavelet-based video codec are comparable with those produced using the H.264/AVC standard.

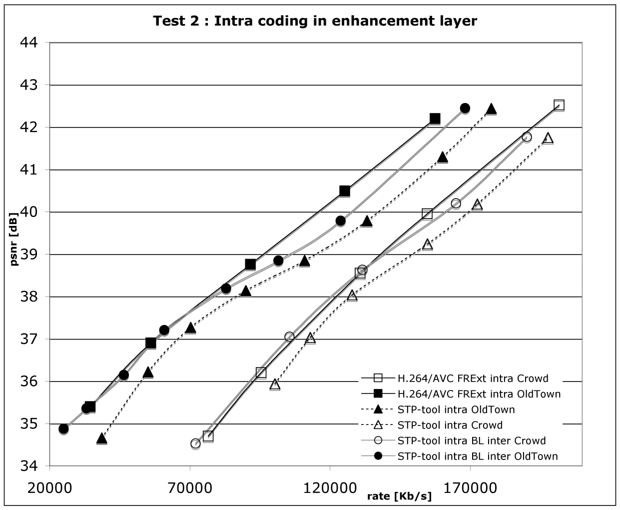

Fig. 4. Intra coding comparisons between STP-tool and H.264/AVC at 1920×1080 with or w/o inter-coded base layer.

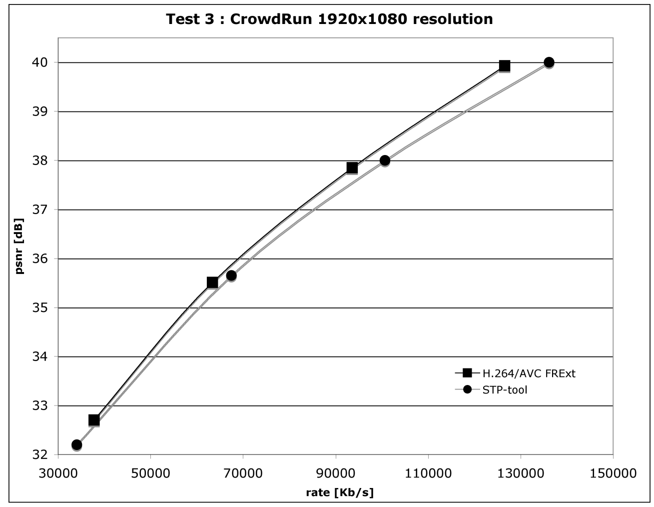

Fig. 5. Comparison between STP-tool and H.264/AVC with inter-coding at 1920×1024.

6. CONCLUSIONS

The aim of this work was to test scalable wavelet-based video coding on HD video material. The obtained results have been produced under test conditions that fit HD coding scenario needs. It has been shown how the exploitation of temporal redundancy can significantly decrease the coding bit-rate even under the unfavorable constraint of a near constant quality. Moreover, it has been shown that for HD applications that use 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-coding case, and this encourages further work in wavelet video coding. Results have been obtained with an entropy coding technology that provide performance comparable to JPEG200. Therefore similar conclusions can surely be expected for a similar scalable video coding architecture which employ JPEG2000 compliant technology for texture compression. It would be of particular interest 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, "Efficeient wavelet-based video compression," Tech. Rep., ISO/IEC JTC1/SC29/WG1, N3954, Assisi, Italy, July 2006.

[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, Geneve, Swizerland, 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.dcimovies.com/-DCI_Digital_Cinema_System_Spec_v1.pdf.

[9] "European boadcasting 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/.