An H.264/AVC to SVC TemporalTranscoder in Baseline Profile
Digest of Technical Papers

Rosario Garrido-Cantos[1], Jan De Cock[2], José Luis Martínez[1], Sebastiaan Van Leuven[2], Pedro Cuenca[1],
Antonio Garrido[1], and Rik Van de Walle[2], *Member, IEEE*

[1] Albacete Research Institute of Informatics, University of Castilla-La Mancha, Albacete, Spain
[2] Department of Electronics and Information Systems - Multimedia Lab, Ghent University - IBBT, Ghent,
Belgium

Abstract—**Scalable Video Coding provides temporal, spatial and quality scalability using layers within the encoded bitstream. This feature allows the encoded bitstream to be adapted to different devices and heterogeneous networks. This paper proposes a technique to convert an H.264/AVC bitstream in Baseline profile to a scalable stream which provides temporal scalability. Applying the presented approach, a reduction of 65% of coding complexity is achieved while maintaining the coding efficiency.**

I. INTRODUCTION

The demand of multimedia contents is growing every day. It is necessary to compress digital video information to reduce the amount of data and therefore minimizing the resources required for transmission and storage while preserving the video quality. These efficient multimedia communications are highly desirable for practical consumer applications. Nowadays, one of the most popular and efficient compression schemes is H.264/AVC [1].

The encoded video are transmitted over heterogeneous networks and received by devices with different decoding capacities (display, processing power, memory capabilities, etc). It is desirable that the contents can be adapted to these different networks and a wide range of terminals. This is possible using scalable video coding schemes.

Recently, the scalable extension of H.264/AVC was standardized. This extension is denoted as Scalable Video Coding (SVC) [1][2] and can provide temporal, spatial, and quality scalability (or a combination of these). SVC bitstreams are organized in layers (one base layer and one or more enhancement layers). The base layer represents the lowest frame rate, spatial resolution and quality while the enhancement layers allow increasing the frame rate, resolutions and quality. The bitstream is adaptable to the channel bandwidth or the terminal capabilities by truncating the undesired enhancement layers.

Today, most of the video contents are still created in a single-layer format (H.264/AVC video streams) so they cannot benefit from this scalability. This fact leads to requirements for developing alternative techniques to enable video adaptation between non-scalable and scalable

bitstreams. Several techniques have been proposed for introducing scalability in compressed bitstreams. The majority of the proposals are related to quality-SNR scalability [3][4], although there are some proposals for spatial [5] and temporal scalability [6].

In this paper, an efficient technique is proposed for transcoding H.264/AVC bitstreams in Baseline profile (P-picture based) without temporal scalability to SVC bitstreams providing temporal scalability with two temporal layers. This scalability is achieved using hierarchical P-pictures [7] in the SVC encoder. The efficiency during transcoding is obtained by reusing as much information as possible from the original bitstream, such as motion information. The presented technique is capable of reducing coding complexity by 65%, hereby outperforming earlier solutions such as in [6]. In contrast to [6], we show that our proposal can be successfully applied to a wide range of test sequences with varying motion characteristics and resolutions.

II. TRANSCODING FROM H.264/AVC TO SVC

The idea behind the proposed transcoder consists of reusing Motion Vectors (MVs) collected in the H.264/AVC decoding algorithm (as part of the transcoder) to accelerate the SVC Motion Estimation (ME) process which is one of the most time consuming parts of the SVC encoder.

The MVs generated by H.264/AVC represent, approximately, the amount of movement of the MB or sub-MB partition and can be reused to accelerate the SVC ME process by reducing the search area tracing a circumference with radio the length of the incoming MV. As the MB partitioning can be different for the same picture in H.264/AVC and SVC, a one-to-one mapping between previously calculated H.264/AVC MVs and the SVC MVs is often not possible. To overcome that situation, the average of the incoming MVs of a determined MB of H.264/AVC is used.

This proposal is only applied to the enhancement temporal layers while the base layer will be decoded and re-encoded.

III. IMPLEMENTATION RESULTS

In this section, results from the implementation of the proposal described in previous section are shown. Test sequences with varying characteristics were used, namely Foreman, Harbour, Football and Mobile in CIF resolution (30 Hz) and QCIF resolution (15 Hz).
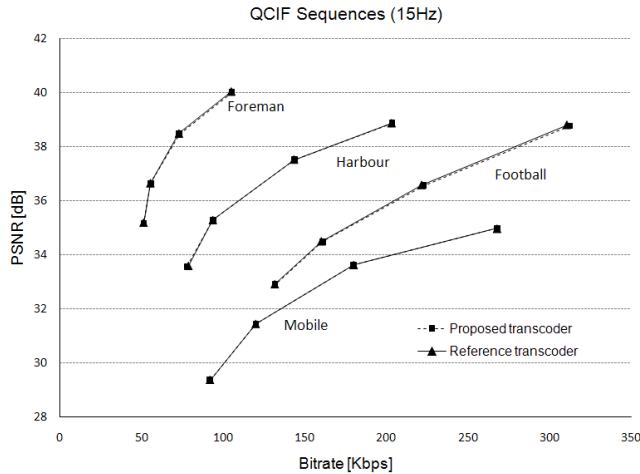
Fig. 1. Rate-Distortion performance of test sequences in QCIF resolution

Fig. 2. Rate-Distortion performance of test sequences in CIF resolution

TABLE I.
RD PERFORMANCE AND TIME SAVINGS FOR QCIF RESOLUTION

RD performance and time savings of AVC/SVC transcoder QCIF (15 Hz) - GOP = 2				
Sequence	ΔPSNR (dB)	ΔBitrate (%)	TS (%)	
			Full Seq.	Partial
Foreman	-0.009	0.37	42.92	63.71
Harbour	0.007	-0.04	60.68	89.97
Football	-0.021	0.79	17.76	26.59
Mobile	0.003	-0.09	58.83	87.13
Average	-0.005	0.26	45.05	66.85

TABLE II.
RD PERFORMANCE AND TIME SAVINGS FOR CIF RESOLUTION

RD performance and time savings of AVC/SVC transcoder CIF (30 Hz) - GOP = 2				
Sequence	ΔPSNR (dB)	ΔBitrate (%)	TS (%)	
			Full Seq.	Partial
Foreman	-0.010	0.52	41.79	63.45
Harbour	0.003	-0.15	60.09	87.26
Football	-0.024	0.95	15.45	25.26
Mobile	0.004	-0.12	56.19	84.74
Average	-0.007	0.30	43.38	65.18

ΔBitrate: Bitrate increase; ΔPSNR: Difference in quality (negative means reduction); TS (Time Saving): complexity reduction

These sequences were encoded using the H.264/AVC Joint Model (JM) reference software [8], version 16.2, with an IPPP pattern with a fixed QP = 28 in a trade-off between quality and bitrate. Then, for the reference results, encoded bitstreams are decoded and re-encoded using the JSVM software, version 9.19.3 [9] with temporal scalability, Baseline profile and different values of QP (28, 32, 36, 40).

For the results of our proposal, encoded bitstreams in H.264/AVC are transcoded using the technique described in Section II.

Fig. 1 and Fig. 2 show *Rate-Distortion* (RD) curves for the SVC bitstreams where it can be seen that our proposal for transcoding is able to approach the RD optimal of reference transcoded (re-encoded) without any significantly loss.

Table I and Table II show ΔPSNR, ΔBitrate and Time Saving (TS) when our technique is applied compared to the more complex reference transcoder. ΔPSNR and ΔBitrate are calculated using [10] and TS is calculated for the full sequence (TS Full Seq.) and for the enhancement temporal layer (where our approach is applied, TS Partial) using the expression $TS (\%) = 100 \cdot (T_{ref} - T_{prop})/T_{ref}$.

IV. CONCLUSIONS

This paper presents an approach for H.264/AVC to SVC transcoding with temporal scalability in Baseline profile. The motion vectors extracted after decoding H.264/AVC bitstream have been used to accelerate the motion estimation process in the e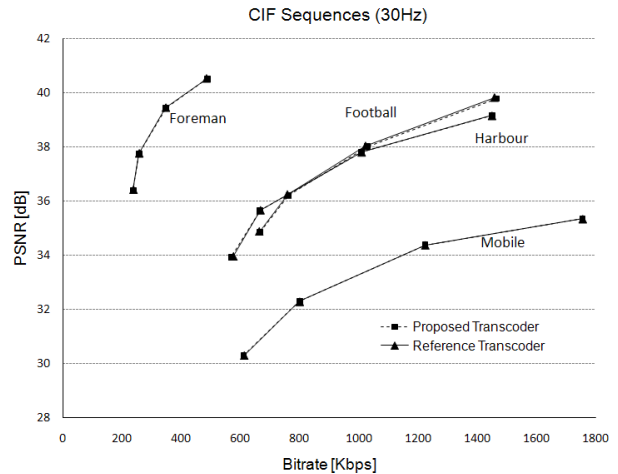nhancement temporal layer of SVC. The experimental results show that our approach is capable of achieving a time saving of 65% with negligible RD loss, hereby outperforming the state-of-the-art solution.

REFERENCES

[1] ITU-T and ISO/IEC JTC 1: Advanced Video Coding for Generic Audiovisual Services. ITU-T Rec. H.264/AVC and ISO/IEC 14496-10 (including SVC extension). March 2009.

[2] H. Schwarz, D. Marpe, T. Wiegand, "Overview of the Scalable Extension of the H.264/MPEG-4 AVC Video Coding Standard". Joint Video Team, Doc. JVT-W132, San Jose, CA, USA, 2007.

[3] H. Shen, S. Xiaoyan, F. Wu, H. Li and S. Li, "Transcoding to FGS Streams from H.264/AVC Hierarchical B-Pictures," *IEEE Int. Conf. Image Processing*, Atlanta, 2006.

[4] J. De Cock, S. Notebaert, P. Lambert and R. Van de Walle, "Architectures of Fast Transcoding of H.264/AVC to Quality-Scalable SVC Streams," *IEEE Transaction on Multimedia* vol. 11 n.7, pp.1209-1224, 2009.

[5] R. Sachdeva, S. Johar and E. Piccinelli, "Adding SVC Spatial Scalability to Existing H.264/AVC Video," *8th IEEE/ACIS International Conference on Computer and Information Science*, Shanghai, 2009.

[6] A. Dziri, A. Diallo, M. Kieffer and P. Duhamel, "P-Picture Based H.264 AVC to H.264 SVC Temporal Transcoding," *International Wireless Communications and Mobile Computing Conference,* 2008.

[7] S. Wenger, "Temporal scalability using P-pictures for low-latency applications," *IEEE Second Workshop on Multimedia Signal Processing,* Redondo Beach, CA, USA, Dec. 1998, pp. 559–564.

[8] Joint Model JM reference software, http://iphome.hhi.de/suehring/tml/download/

[9] Joint Video Team JSVM reference software, http://ip.hhi.de/imagecom_G1/savce/downloads/SVC-Reference-Software.htm

[10] G. Sullivan and G. Bjøntegaard, "Recommended Simulation Common Conditions for H.26L Coding Efficiency Experiments on Low-Resolution Progressive-Scan Source Material". ITU-T VCEG, Doc. VCEG-N81. September 2001.