SPIRAL SCAN IN VIDEO COMPRESSION

Łukasz Błaszak, Marek Domański

Poznań University of Technology, Institute of Electronics and Telecommunications, ul. Piotrowo 3A, 60-965 Poznań, Poland,
phone: +(48) 61 66 52 762 and +(48) 61 66 52 870, fax: +(48) 61 66 52 572,
email: lukasz.blaszak @multimedia.edu.pl, domanski@et.put.poznan.pl, web: www.multimedia.edu.pl.

ABSTRACT

The paper describes a proposal to replace the raster (linear) scan of macroblocks by a new spiral scan. Implications of such a scan for coding efficiency and implementation complexity are considered. The considerations are verified by some experiments. The conclusion is that both coding efficiency and implementation complexity are nearly the same for both raster and spiral scan of macroblocks. The spiral scan is considered as a tool for low-complexity SNR (quality) scalability as well as for coding of regions of interest. In particular, good subjective quality may be obtained when the spiral scan is used for fine-grain quality scalability.

1. INTRODUCTION

Recently, video compression came to a turning point when there appeared a family of advanced techniques with improved coding efficiency. The new video coding systems like AVC/H.264 [1,2] and VC1/WM 9 [3] continue the development of hybrid codecs with motion-compensated prediction loop. Significantly increased compression performance is obtained as a cumulative effect of many improvements. Development of all those modifications was aimed at increased compression performance rather than other functionalities provided by the previous version of MPEG-4 video encoding [4,5].

Now, the video compression community is approaching the next development step, i.e. embedding advanced functionalities, like scalability and coding of regions of interest into advanced video coding systems. Because of its prospective wide applications in heterogeneous communication systems, scalable video coding is already a subject of standardization activities in ISO/IEC (MPEG group [6,7]) and ITU (recently jointly with ISO/MPEG within Joint Video Team). Moreover, encoding of regions of interest was defined as a requirement for coming video coding standard [6]. This requirement was practically still not fulfilled by recent standardization proposals.

This paper deals with both fine-granularity SNR (quality) scalability as well as coding of regions of interest in advanced video coding techniques. Our approach exploits an observation that both classic and advanced video coding systems proceed images on mackroblock basis. In this paper, the spiral scan of macroblocks (already announced in [8,9]) is considered as a tool for scalability and coding of regions of interest.

2. REGION OF INTEREST CODING USING SPIRAL SCAN

In still image compression, encoding of a region of interest (RoI) with higher fidelity is a useful mean of bit allocation that allows for better quality of those image portions that are more important for a viewer. Region of interest coding is already adopted in the approved international image compression standard JPEG2000 [10,11]. On the other hand, in video compression technology, there is a lack of such a functionality even in the advanced video coding techniques [1,2]. Here, we propose a tool for such a functionality.

In video coding, simplicity of the definition of a region of interest would be important as it is related to a low number of bits needed to encode the location and the borders of RoI. A straightforward way to define a region of interest (RoI) is to define its central point (centre of interest – CoI), and to define the region relative to this point (Fig. 1).

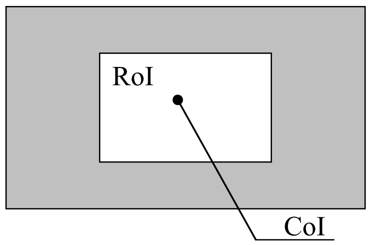

Fig. 1. Region of interest (RoI) with its centre of interest (CoI).

In video production, cameramen usually track the most interesting objects in such a way that they remain in central portions of the consecutive pictures. Therefore, it is useful to have the default position of CoI just in the geometrical centre of pictures. If needed, the current position of CoI may be defined relative to this central point.

For coding purposes, a picture as well as a region of interest may be viewed as an ordered collection of macroblocks. Here, in this paper, it is proposed that this order would be defined by a spiral scan (Fig. 2). The spiral scan is very similar to the water-ring scan [12,13] proposed as a tool for fine granularity scalability.

For a spiral scan, its aspect may be defined by the relation between lengths of horizontal and vertical runs of macroblocks. For a given spiral scan aspect, a region of interest is uniquely defined by the starting point of the spiral scan (i.e.

centre of interest) and the number of macroblocks within the region of interest.

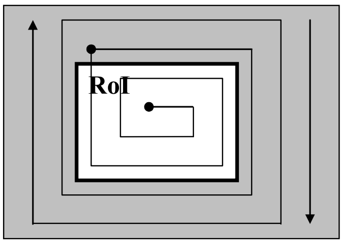

Fig. 2. Spiral scan for RoI coding.

The above mentioned way of defining a RoI needs the number of macroblocks to be transmitted only. This number of macroblocks may be represented using no more than 11 bits for standard-definition video.

The macroblocks inside RoI are encoded with higher fidelity while those outside RoI are coded with lower fidelity, e.g. with higher values of the quantizer scale factor.

3. SPIRAL SCAN IN SCALABLE VIDEO CODING

The spiral scan may be also used as a tool to obtain fine-granularity quality (SNR) scalability. Let assume that a coder produces a layered video representation. The layers represent various levels of quality or various spatial and/or temporal resolutions. The low-resolution (or low-quality) base layer bitstream in embedded in the total video representation, i.e. may be extracted from the overall bitstream.

The individual bitrates of the base layer and the enhancement layer may be set using bitrate control mechanism in the corresponding layers. If video is broadcasted to many receivers, it may happen that the communication links to individual users differ significantly. In such a case, fine-granularity quality scalable video coding is very profitable as it allows to extract bitstreams with bitrates exactly matched to throughputs of individual communication links. Such a bitstream may be extracted as the base layer and a part of the enhancement layer. In the enhancement layer, the part representing outer portions of pictures is disregarded (Fig. 3). In this way, the visually less important macroblocks are decoded from the low-resolution base layer only (Fig. 4).

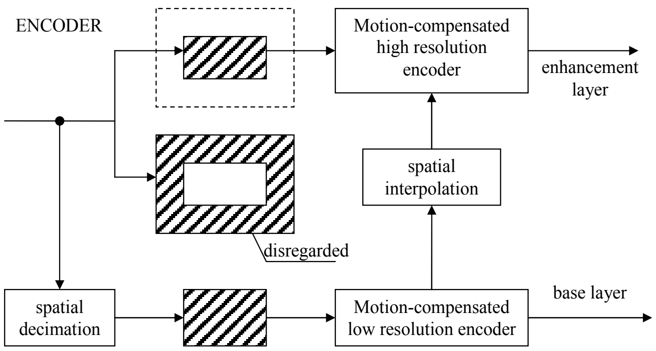

Fig. 3. Quality (SNR) scalability with region of interest encoding.

Therefore, the region of interest is reconstructed with higher quality while the other parts images are decoded with lower quality.

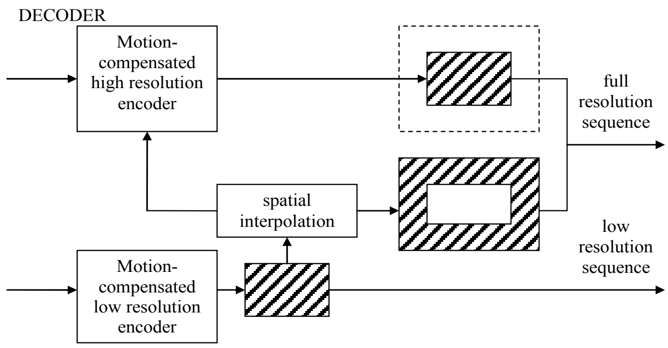

Fig. 4. Decoding of a quality scalable bitstream with region of interest decoding.

The spiral scan provides not only region of interest encoding but also fine-granularity quality scalability. For an image, encoded are macroblocks starting from the centre of interest. The stream of encoded macroblocks may be cut in an arbitrary point thus giving very small scalability granules of order of one compressed macroblock, i.e. about 50 – 200 bits. The cutting point defines the size of the high-quality area about the centre of interest in the decoded pictures.

In order to asses this technique of quality scalability, the authors have performed several experiments with test video sequences using default setting of the centre of interest in the geometrical centre of pictures. The results of subjective quality assessment are astonishingly good. For some sequences, like City, even cutting out of 40-50% of a bitstream may be very difficult for a viewer to perceive. For other test sequences, like Crew, some objects attract attention in the marginal portions of images, and a perceived subjective quality is much worse. Nevertheless, for most video sequences, the scheme works very well in terms of subjective quality. For most sequences, most of viewers are unable to perceive significant loss of quality even by cutting out 30-40 % of a bitstream.

Then, quality (SNR) scalability using spiral scan was implemented within scalable video model (SVM) software [14] that serves as reference for algorithm development in MPEG.. The experimental results prove that the subjective quality was quite high for the technique proposed. For the spiral scan, good subjective quality was verified by the independent tests [15].

4. EMBEDDING OF SPIRAL SCAN INTO STANDARD VIDEO BITSREAM HIERARCHY

Video data exhibit hierarchical structure with some levels: sequence, group of pictures, frame, group of slices, slice, macroblock, and block. This structure is applicable to coding with spiral scan. The level of slices may be the most questionable. For the spiral scan, the consecutive slices may be defined as pieces of a series of macroblocks that are located along the spiral scan (Fig. 5).

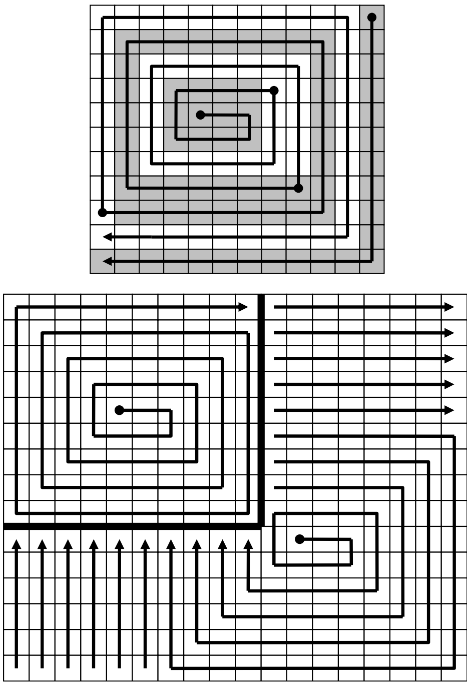

Fig. 5. Possible structures of slices in an image with the spiral scan.

Therefore, the spiral scan allows for standard bitstream syntax with modified semantics. The optional syntax modification is related to additional information on the spiral aspect that should be sent for the case that the spiral aspect is different from the one that may be calculated from frame aspect. Other optional information is related to the shift of the centre of interest (CoI) with respect to the geometrical centre of pictures. Both pieces of information may be sent in the sequence header or in the group of pictures header.

5. COMPARISON OF CODING EFFICIENCY FOR SPIRAL AND RASTER SCAN

Extensive experimental comparisons have been made in order to compare compression efficiency for the raster (classic) and the spiral scan of macroblocks (Fig. 6).

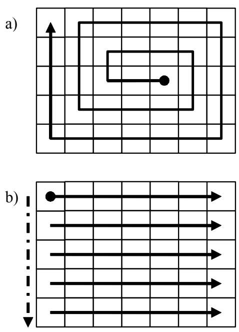

Fig .6. a) Spiral scan of macroblocks, b) Raster scan of macroblocks.

The comparative tests made for single-layer (nonscalable) codecs prove that this performance is very similar for both scans. The rate-distortion curves are undistinguishable for both scans for many test sequences (Figs. 7, 8). Therefore, a clear conclusion is that the spiral scan does not decrease compression performance of an AVC/H.264-compliant codecs.

When the spiral scan order was used by switching on the flexible macroblock order (FMO) mode, i.e. without proposed context modification, then there was a significant decrease of coding efficiency. The bitstream overhead, depending on the sequence, was up to 30 % and the average overhead for this mode was 13%.Thus, the context modification for spatial prediction and motion vector prediction highly influence on the compression efficiency.

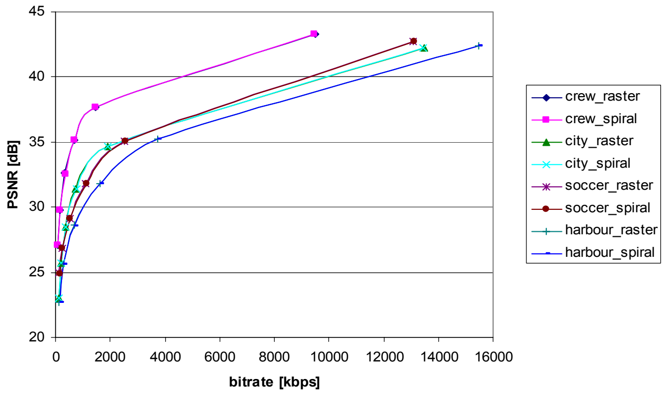

Fig. 7. Comparison of the ratser and spiral scans for IPPPP 30 Hz sequences with 4CIF resolution (704×576).

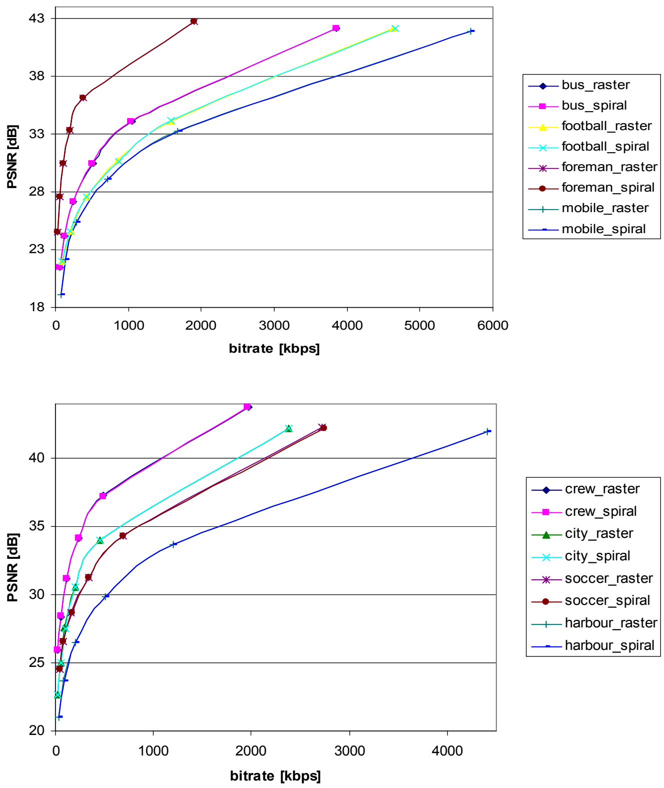

Fig. 8. Comparison of the raster and spiral scans for IPPPP 30 Hz sequences with CIF resolution (352×288).

6. IMPLEMENTATION COMPLEXITY

Careful study shows that implementation complexity is very similar for both scans. The study was done for several coding scenarios. Exemplary results for AVC-compliant intra-frame coding are shown in Table 1. Similarly, the number of operations needed for inter-frame prediction with spiral scan is less then 1% higher than that for the raster scan (Table 2). The only additional operations needed for the spiral scan in inter-frame mode are related to macroblock reading order modification, block order modification and neighborhood retrieving.

Table 1. The comparison of estimated number of operations needed for go through the prediction 4×4 path for raster scan and spiral scan for intra prediction.

oparation	number of operations	
	raster scan	spiral scan
+, -, <<, >>, ++, --, !, &, |	8200	8250
*, /, %	2250	2300
comparison	1700	1750
memory access	11300	11350

Table 2. The comparison of estimated number of operations needed for the 16×16 interframe prediction for the raster scan and the spiral scan (block match search with range of 64 pixels).

oparation	number of operations	
	raster scan	spiral scan
+, -, <<, >>, ++, --, !, &, |	266000	266300
*, /, %	2000	2150
comparison	267500	268000
memory access	266800	267200

In particular, for the spiral scan, the program flow is somewhat more complicated because contexts must be adapted to the current direction of processing. Different contexts for inter- and intra-frame prediction and context-adaptive entropy coding are used for individual directions of macroblock processing. The code length for full implementation is about 5% greater for the spiral scan solution in compare to the standard raster scan solution.

Nevertheless, with the spiral scan, the implementation of quality scalability is much simpler as compared to that proposed in the scalable video model [14] that is currently approved as a reference for standardization activities.

7. CONCLUSIONS

The spiral scan has been proposed as an efficient tool for encoding of regions of interest. Moreover, the spiral scan may be used for low-complexity fine-granularity quality (SNR) scalability.

Replacement of the raster scan by the spiral scan does not deteriorate compression performance and yields negligible complexity increase only. Nevertheless, the complexity of the quality scalability tool may be much lower for the spiral-based solution as compared to the quality scalability tool proposed in the scalable video model.

It is worth to stress very good subjective quality of the frames retrieved from the bitstreams with the bitrate reduced by macroblock partitioning along the spiral scan.

ACKNOWLEDGEMENTS

This work has been supported by the Polish State Committee for Scientific Research.

The authors thank anonymous reviewers for their valuable comments.

REFRENCES

[1] ISO/IEC JTC1/SC29/WG11, Doc. N6540, ISO/IEC 14496-10 *Advanced Video Coding, 3rd Edition*, July 2004.

[2] *Special issue on the H.264/AVC video coding standard*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no 7, July 2003.

[3] J. Liang, M.-C. Lee, J. Ribas-Corbera, *Windows Media Video 9: overview and applications*, Signal Processing: Image Communication, vol. 19, No 9, Oct. 2004, pp. 851-876.

[4] ISO/IEC IS 14496-2, *Coding of Audio-Visual Objects – Part 2: Visual, 2nd Ed.*, 2001.

[5] F. Pereira, T. Ebrahimi (eds.), *The MPEG-4 book*, Prentice Hall PTR, 2002.

[6] ISO/IEC JTC1/SC29/WG11 Doc. N6901, *Working Draft 1.0 of 14496-10:200x/AMD1 Scalable Video Coding*, Hong Kong, Jan. 2005.

[7] ISO/IEC JTC1/SC29/WG11 Doc. N6052, *Applications and Requirements for Scalable Video Coding*, , Brisbane, October 2003.

[8] Ł. Błaszak, M. Domański, R. Lange, *Modified AVC Codec with SNR Scalability Based on Macroblock Hierarchy*, 11th International Workshop on Systems, Signals and Image Processing, IWSSIP'04, Poznań, Sept. 2004, pp. 15-18,.

[9] Ł. Błaszak, M. Domański, R. Lange, A. Łuczak, ISO/IEC JTC1/SC29/WG11 Doc. M11101, *Response to SVC CE2 tasks: testing of SNR scalability technology*, July 2004.

[10] ISO/IEC IS 15444-1. *JPEG2000 image coding system*, 2000.

[11] D. Taubman, M. Marcellin, *JPEG2000. Image compression fundamentals, standards and practice*, Kluwer 2002.

[12] G.H. Park, Y.J. Lee, W.-S. Cheong, K. Kim, J. Kim, Y. K. Lim, *Water ring scan method for H.26L based FGS*, Tech. Rep. B094, ITU-T VCEG | ISO/IEC MPEG (JVT), Feb. 2002.

[13] G.H. Park, K. Kim, *Water ring scan method for FGS video coding schemes*, IEICE Trans. on Communications, vol. E88-B, pp. 835-840, Feb. 2005.

[14] ISO/IEC J TC1/SC29/WG11 Doc. N6717, *Scalable video model version 3.0*, Palma de Mallorca, Oct. 2004.

[15] V. Baronchini, T. Oelbaum, *Subjective test results for the CfP on scalable video coding technology*, Doc. ISO/IEC JTC1/SC29/WG11 M10737, March 2004.