



# LAR Video : Lossless Video Coding with Semantic Scalability

Erwan Flécher, Samir Amir, Marie Babel, Olivier Déforges

## ► To cite this version:

Erwan Flécher, Samir Amir, Marie Babel, Olivier Déforges. LAR Video : Lossless Video Coding with Semantic Scalability. Sep 2006, pp.227-230. hal-00132965

**HAL Id: hal-00132965**

**<https://hal.science/hal-00132965>**

Submitted on 23 Feb 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# LAR VIDEO: Lossless Video Coding with Semantic Scalability

Erwan Flécher<sup>1</sup>, Samir Amir<sup>1</sup>, Marie Babel<sup>2</sup>, Olivier Déforges<sup>2</sup>

IETR UMR CNRS 6164, Image and Remote Sensing Group, INSA of Rennes

20 avenue des Buttes de Coësmes, 35043 Rennes, FRANCE

e-mail: <sup>1</sup>{eflecher, samir}@ens.insa-rennes.fr <sup>2</sup>{mbabel, odeforge}@insa-rennes.fr

**Abstract**—Baptized “LAR Video”, method proposed in this paper describes a new lossless video coding algorithm with advanced semantic scalability. Motion estimation and compensation steps are first achieved to produce the well known displaced frame difference (DFD). The basic idea is to apply on this residual error a pyramidal decomposition based on an efficient scalable image compression technique, named LAR-APP. Resulting from a progressive data broadcasting, image-sequence can be scalably rebuilt in the decoder at different spatial resolution level. The given experimental results show that the proposed solution, in addition to the scalability, achieves good compression performances.

## I. INTRODUCTION

Video coding is used by various applications requiring efficient and fast coding tools that allow high compression and low distortion rate. These applications have to work with heterogeneous and time variable data network properties. To deal with this problem, a scalable video stream is commonly used [1]. Scalability defines compression algorithms that combine several processing levels [2] in order to hierarchically describe the source. Low bit-rate video coding can be achieved with non-reversible tools. However, in some applications, resulting damages are unbearable especially in telemedicine and movies making context. Methods that combine both efficient lossless compression and scalability are still uncommon. In order to meet the requirement, a new scalable video coding algorithm, based on an efficient multiresolution approach for still image compression (LAR-APP) has been developed.

The LAR (Locally Adaptive Resolution) based on a variable block-size decomposition, leads to an efficient lossy image compression technique [3]. Wrought on this coding assumption, three LAR-like pyramidal decomposition techniques namely LAR-APP [4], Interleaved S+P [5] and RWHT+P [6] followed it. Section II describes the LAR approach and pertinent features of the three succeeding approaches. Section III presents the foundation and the results of a lossless scalable video coder which takes advantage of the residual error coding by the LAR-APP. Finally, we conclude in section IV.

## II. LOSSLESS LAR METHOD FOR STILL IMAGE

### A. Basic principle

Image can be represented as a superposition of a basic information and a local texture. Thus with the same scheme, a low bit rate image (basic information) can be transmitted

with or without its additional error image. The LAR method is based on a two-layer coding that intrinsically gives at least two progressivity levels (Fig. 1).

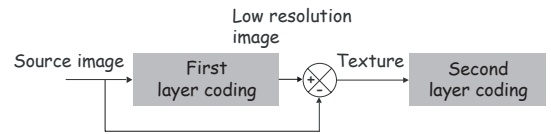


Fig. 1. Basic description of the LAR coder: two-layer coding

First stage of the image coding by LAR method is to generate a low resolution image (in terms of visual quality). This results in the construction of a locally-resolution image, determined by means of a quadtree data structure. Through this decomposition, the pixel size gives implicitly the nature of blocks: indeed, small ones are located on contours whereas large ones are situated on smooth areas. Now, let  $QP[N_{max}..N_{min}]$  represent the quadtree partition where  $N_{max}$  and  $N_{min}$  are respectively the upper and lower limit of square block size. Without the second layer, this coder clearly aims high compression ratio. The low bit-rate rebuilt image is visually acceptable thanks to quadtree partition that accommodates the variable block-size as a function of the original image context. For higher bit-rate, first stage is followed by a refinement layer that, if no quantification is applied, allows a lossless information compression. The semantic scalability property is enabled by a joint use of both the two-layer coding and the quadtree partitioning. This principle is heart of and was enriched by the following three scalable approaches.

### B. Pyramidal approaches

LAR-APP [4], Interleaved S+P [5] and RWHT+P [6] are three methods for still images compression developed in our lab. They are unified algorithms of compression that combine prediction in a enriched context and scalability in terms of resolution and quality level. These methods take advantage of the two-layer LAR codec and add the multiresolution concept. The pyramidal decomposition is ordered by two successive descent processes. The first decomposition pass in the pyramid refines only small blocks located on contours and models homogeneous areas with larger blocks. The conditional splitting is illustrated by figure 2 in which block-size devolves from  $8 \times 8$  to  $2 \times 2$ .

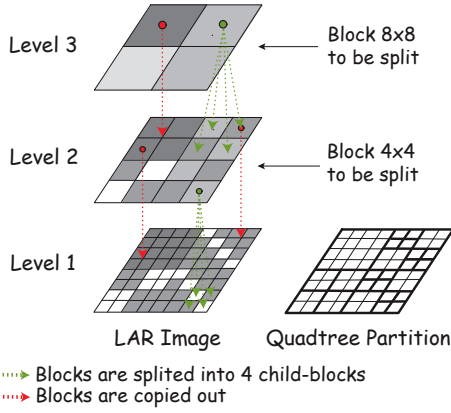


Fig. 2. Pyramidal decomposition of the LAR image conditionally to the quadtree partition  $QP[8..2]$

In this first pyramidal decomposition, semantic scalability is realised using grid information. The content-based local information increases the quality of the low bit-rate images which are rebuilt on each level of the pyramid. If the first decomposition intends to reconstruct the low resolution image (LAR-image), the second one processes the local texture information. All blocks at the current level which have not be encoded during the first pass are decomposed by the refinement layer.

Context modeling has become a key factor in lossless compression efficiency [7]. Global entropy can be reduced when different classes of symbols following the same law can be isolated. Our methods provide a straightforward separation of the laws at two different levels: block-sizes and decomposition level in each pass. Consequently, due to intrinsic property of our coder, an implicit context modelling is realised, leading to dramatically reduce the global entropy.

### C. Decomposition methods

Previously we briefly explained the common working principles of the three pyramidal methods. The purpose of this section is to discuss their own way to reversibly split each block into four child-blocks. The scalable DPCM prediction of the LAR-APP method relies on successive descending processes that benefit from context-based information in the causal image (intra-level prediction) and the subsampled image (inter-level prediction). Contrary to LAR-APP, Interleaved S+P and RWHT+P methods work in a transform space which is more suitable to obtain high compression ratio. The main foundation of Interleaved S+P relies on a special implantation of S-Transform. Originality of this algorithm stands in its feature of efficiently predicting the transformation coefficients from two interleaved pyramids. The third method named RWHT+P, uses a reversible form of the formal Walsh Hadamard Transform applied on  $2 \times 2$  blocks.

Interleaved S+P and RWHT+P are very efficient pyramidal solutions for lossless compression and outperform LAR-APP and states-of-the-art references CALIC [8] and S+P [7]. In spite of that, because of its implementation simplicity, our

video coder uses LAR-APP method for now. Results with better compression ratios utilizing Interleaved S+P and RWHT+P methods shall be reported in near future.

## III. LOSSLESS LAR VIDEO CODING

In order to remove temporal redundancy between successive frames, both motion estimation and motion compensation are implemented in video coding standards. The common idea is to predict the current image by a motion compensated reference known to the decoder. A regular partition (MPEG-4) or a variable block-size decomposition (H.264 [9]) is first carried out. Then for each block, motion estimator seeks the best match in a reference image according to a similarity measure. The distance from the best-match block in the reference is represented by a motion vector that is coded and sent to the decoder. The difference between current and compensated image is usually named displaced frame difference (DFD) and has to be efficiently transmitted. In order to remove spatial redundancy, a transform is applied on this residual error. Commonly, tools used by video coding standard supply a robust compression but do not give an acceptable solution in the lossless coding context.

Joint use of the video coding scheme and the predictive pyramidal approach (section II) aims to meet two fundamental requirements. On one hand, a new algorithm with resolution and quality scalability is proposed. On the other hand, user has at his disposal a simple and unified solution for both lossless and lossy compression. Processing residual error with the LAR-APP method is the solution: a unique algorithm is used to perform in reversible and non-reversible manner (Fig. 3).

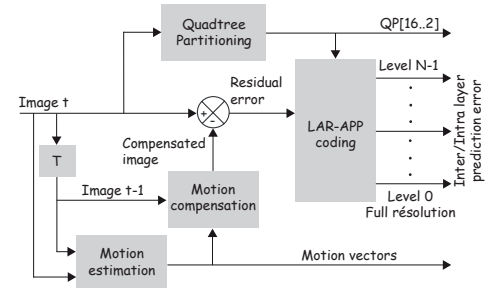


Fig. 3. Basic scheme of the lossless LAR video coder. N decomposition levels for the residual error compression by LAR-APP

### A. Motion estimation

Block matching algorithm (BMA) is a simple and efficient motion estimation method. Full search of the best matching provides optimal solution (in conformity with a similarity criterion) but proves to be very power consuming. Belonging to the zonal search family, EPZS (Enhanced Predictive Zonal Search) [10] overcomes this drawback with a priori knowledge of the block shift. Based on the blocks motion assessment in a causal neighbourhood and a motion reference memory, the displacement of the current block is predicted. Though DFD coding is already not precisely explained, it is nevertheless

interesting to evaluate the influence of the motion model support on residual error entropic cost. Figure 4 illustrates the effect of the regular grid-size variation (block size:  $16 \times 16$ ,  $8 \times 8$  and  $4 \times 4$ ) on the residual error coding by LAR-APP pyramidal approach. Resulting from lossless coding of Foreman image-sequence, mentioned entropies show that the best trade-off is given by a  $8 \times 8$  block-size support.

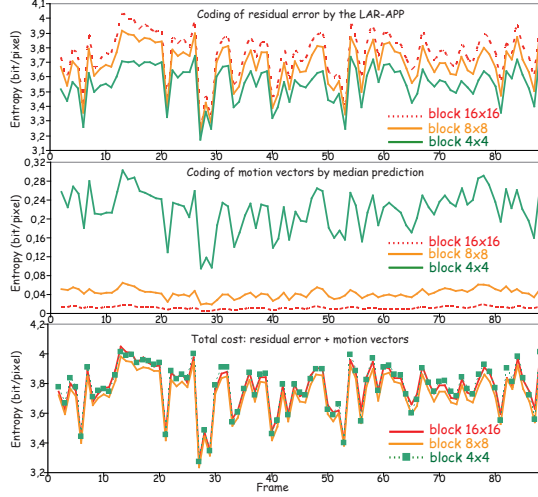


Fig. 4. Effect of the motion model grid-size variation on the residual error coding by the lossless LAR-APP (Foreman sequence)

Naturally, the cost of decreasing DFD entropy with accurate prediction is the increased motion vector entropy. In the low bit-rate context, solutions similar to H.264 [9] approach, prove to be appropriate [11]. Since there exists a trade-off between the costs of motion vectors and residual error, an optimal partition with variable block-size can be found. However in the reversible coding context, advantage provided by these methods, shall be mask in the error entropic dynamics. Consequently, adopted solution in our coder achieves a motion estimation on a grid of  $8 \times 8$  block-size.

### B. Residual error coding

In order to reap profit of the semantic scalability and the context modeling, quadtree partition (with spatial criterion) is here directly applied on images of the video sequence. Image partitioning and scalable decomposition of the bitstream can be described with a block diagram (Fig. 5). In accordance with the method described in section II, the lossless compression of the residual error is obtained by carrying out a full pyramidal decomposition. All blocks of the LAR grid are entirely split into the finest resolution. The coding with semantic scalability separates small blocks information and texture-based data. In this way, the first pass rebuilds an image with well-drawn contours. Thus global diagram defines for each pyramid level, a minimal block-size that conditions the decomposition process. Associated to  $QP^{[16..2]}$ , the scheme (Fig. 5) shows this principle when the minimal block-size is equal to  $2^n$  at pyramidal level  $n$ . Due to the similarity of evaluation process at the decoder, only inter/intra layer prediction errors are transmitted in a scalable bitstream.

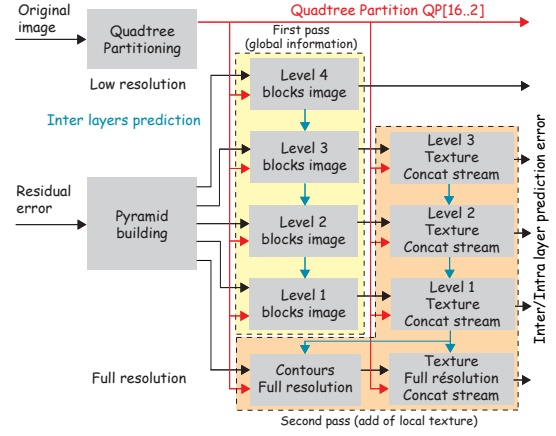


Fig. 5. Scalable coding of the residual error by the predictive pyramidal LAR-APP approach

From the decoder point-of-view, reconstruction of the source image is a two stage-work (Fig. 6). For each level of the pyramid, residual error is firstly decoded and added to the same-size compensated image. For this, pyramids of the compensated image and the residual error are formulated on the same base. Naturally, a coder with full scalability option requires a progressive stage of motion prediction/compensation. Though this step is not yet realised, some applications can now be implemented. In the reversible context, pyramid has to be totally decoded in order to rebuild a lossless full resolution image. Therefore, browsing through image-sequences at different resolution levels is one of our coder's functionalities.

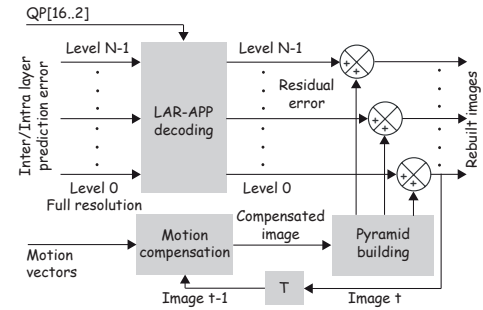


Fig. 6. Basic scheme of the lossless LAR video decoder. N decomposition levels for the residual error compression by LAR-APP

Figure 7 presents successive images resulting from the decoding steps of Football sequence. These describe the progressively rebuilt source-image in accordance with the process described in the preceding paragraph. Part two of this figure shows the rebuilt source-image with two intermediate quality levels. It proceeds from the sum of the full resolution compensated image and a DFD extracted from two intermediate levels. In fig 7.b, resulting from level 1 of the first descending process, subsampled error is interpolated. In fig 7.c, the full resolution error is obtained by decomposition of the contours (blocks  $2 \times 2$ ) only. Results show that the small blocks decomposition is sufficient to improve significantly the visual quality.

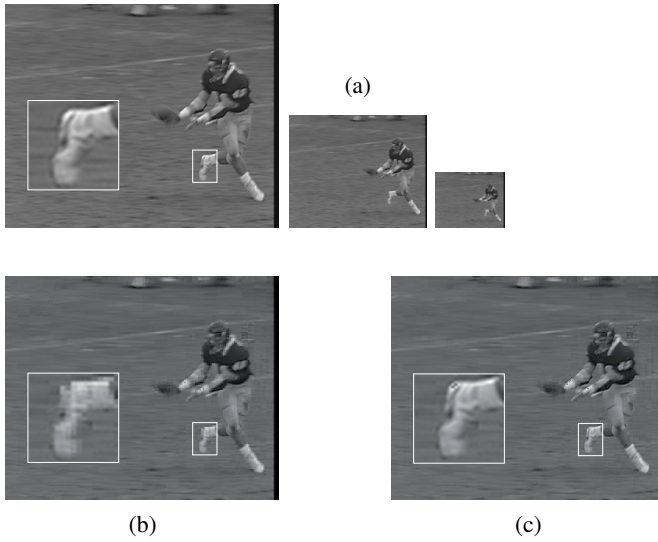


Fig. 7. (a) Progressive reconstruction of the source image from the pyramid of the residual error: lossless, full resolution: 3.80 bpp, level 1: 1.01 bpp, level 2: 0.37 bpp. Intermediate reconstruction from the residual error of the first pass: (b) error of the level 1 + interpolation (c) error of the level 1 + contours decomposition

We just presented our video coder and some intermediate images, it is now interesting to evaluate the coder performances. Table I summarises results of coding various test-sequences with the LAR-APP approach. The APP-Inter (in the video context) and the APP-Intra (similar to still image compression) work with two different types of data and are compared with reference to an entropic criterion. Experimental results show that the DFD's coding by LAR-APP outperforms the APP-Intra mode. The given entropies indicate the possibility of getting better results with Interleaved S+P or RWHT+P coder in near future.

Sequence	Entropy (bpp) - CIF format (352 × 288)			
	APP Intra	Residual error	APP Inter	Motion vectors
Foreman	4.30	<b>3.66</b>	3.69	0.042
Mother	3.60	2.91	<b>2.85</b>	0.035
Mobile	5.86	5.17	<b>5.04</b>	0.030
Football	4.09	4.33	<b>3.81</b>	0.081
Tempete	5.35	4.68	<b>4.51</b>	0.033
CoastGuard	5.36	4.66	<b>4.40</b>	0.025

TABLE I

INTRA/INTER MODE COMPARISON OF THE LOSSLESS LAR-APP

Figure 8 better describes the results presented in table I for Football and MotherandDaughter sequences. Both the zero-th order entropy of the DFD (named residual error in table) and that resulting from the lossless APP are showing in a temporal evolution context. Offering scalability function by level, LAR-APP also proves to be a efficient compression tool.

#### IV. CONCLUSION

Scalable video coder presented in this paper is an efficient solution for both lossless and lossy compression. Enriched

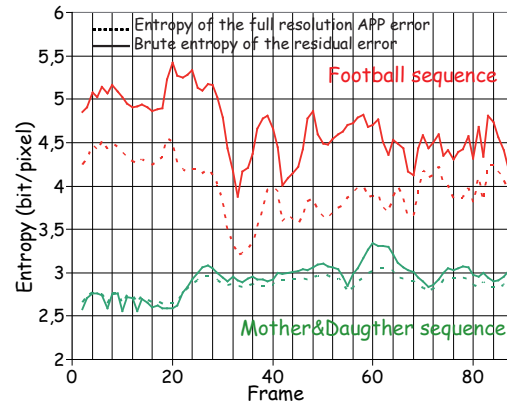


Fig. 8. Comparison between entropy of the lossless LAR-APP and the zero-th order entropy of the DFD on Football and MotherandDaughter sequences

by the semantic scalability, DFD's coding by the LAR-APP presents good compression properties in a reversile context. Through several examples, we showed that its performances are encouraging for our futur works. In order to increase the compression ratio, a short aim is to substitute the LAR-APP by Interleaved S+P or RWHT+P method. Coding of the residual error with a pyramidal approach is not a sufficient condition to give our coder a global scalability. Thus, the second aim is to realise a method with full scalability which includes an hierarchical motion vectors coding.

#### REFERENCES

- [1] G. Marquant, "Représentation par maillage adaptatif déformable pour la manipulation et communication d'objets vidéo," Ph.D. dissertation, Université de Rennes 1, Rennes, Dec. 2000.
- [2] N. Cammas, "Codage vidéo scalable par maillages et ondelettes 2D," Ph.D. dissertation, Université de Rennes 1, Rennes, Nov. 2004.
- [3] O. Déforges and J. Ronsin, "Region of Interest Coding for Low Bit-Rate Image Transmission," in *Proc. International Conference on Multimedia and Expo ICME'2000*, vol. 1, 30 July-2 Aug 2000, pp. 107 – 110.
- [4] M. Babel, O. Déforges, and J. Ronsin, "Lossless and Lossy Minimal Redundancy Pyramidal Decomposition for Scalable Image Compression Technique," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'03*, vol. III. Hong Kong: Conference cancelled - Invited paper in ICME 2003, April 6-10 2003, pp. 249–252.
- [5] —, "Interleaved S+P Pyramidal Decomposition with Refined Prediction Model," in *IEEE International Conference on Image Processing, ICIP'05*, vol. 2, Genova, Italy, September 2005, pp. 750–753.
- [6] O. Déforges, M. Babel, and J. Motsch, "The RWHT+P for an improved lossless multiresolution coding," in *European Signal Processing Conference, EUSIPCO'06*, To be published, September 2006.
- [7] A. Said and W. A. Pearlman, "An image multiresolution representation for lossless and lossy image compression," *IEEE Trans. on Image Processing*, vol. 5, pp. 1303–1310, September 1996.
- [8] X. Wu, N. Memon, and K. Sayood, "A Context-based, Adaptive, Lossless/Nearly-Lossless Coding Scheme for Continuous-Tone Images (CALIC)," 1995.
- [9] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video Coding with H.264/AVC: Tools, Performance, and Complexity," *IEEE Circuits and Systems Magazine*, vol. 4, no. 1, pp. 7–28, First Quarter 2004.
- [10] A. M. Tourapis, "Enhanced predictive zonal search for single and multiple frame motion estimation," in *Proc. SPIE Vol. 4671, p. 1069-1079, Visual Communications and Image Processing 2002*, C.-C. Jay Kuo; Ed., C.-C. J. Kuo, Ed., Jan. 2002, pp. 1069–1079.
- [11] J. Zhang, M. Ahmad, and M. Swamy, "Quadtree Structured Region-Wise Motion Compensation for Video Compression," *IEEE Transactions On Circuits And Systems For Video Technology*, vol. 9, no. 5, pp. 808–822, August 1999.