Interleaved S+P Pyramidal Decomposition with Refined Prediction Model

Marie Babel, Olivier Déforges, Joseph Ronsin

INTERLEAVED S+P PYRAMIDAL DECOMPOSITION WITH REFINED PREDICTION MODEL

Marie Babel, Olivier Déforges, Joseph Ronsin

IETR UMR CNRS 6164, Image and Remote Sensing Group / INSA of Rennes
20, av. des Buttes de Coësmes, CS 14315, 35043 RENNES, FRANCE
Contact : {mbabel,odeforge,ronsin}@insa-rennes.fr

ABSTRACT

Scalability and others functionalities such as the Region of Interest encoding become essential properties of an efficient image coding scheme. Within the framework of lossless compression techniques, S+P and CALIC represent the state-of-the-art. The proposed Interleaved S+P algorithm outperforms these method while providing the desired properties. Based on the LAR (Locally Adaptive Resolution) method, an original pyramidal decomposition combined with a DPCM scheme is elaborated. This solution uses the S-transform in such a manner that a refined prediction context is available for each estimation steps. The image coding is done in two main steps, so that the first one supplies a LAR low-resolution image of good visual quality, and the second one allows a lossless reconstruction. The method exploits an implicit context modelling, intrinsic property of our content-based quad-tree like representation.

1. INTRODUCTION

Due to the extensive use of low-bandwidth networks such as the Internet, image compression methods tend to provide a scalable solution for data coding. Specific applications like telemedicine indeed need progressive transmission of images, in both terms of resolution and quality, by means of an unique and simple compression scheme.

For these reasons, although full resolution coding such as the CALIC method [1] is usually used in the field of lossless compression of greyscale images, pyramid structures are still attractive processes. If DPCM schemes and reversible transforms are generally not implemented at the same time, S+P [2] algorithm proposes a multiresolution decomposition of images, combining S-Transform and prediction.

The LAR (*Locally Adaptive Resolution*), based on a variable block-size decomposition, leads to an efficient lossy image compression technique [3]. In [4] we have introduced a minimal-redundancy pyramidal decomposition (also called LAR-APP) where the successive prediction steps take advantage both of a 360˚ type neighborhood and an implicit context modeling.

This work proposes an original image coding solution using a pyramidal approach based on the LAR-APP method combined with a particular implementation of the S-Transform, so that an Interleaved S+P algorithm is generated.

This paper is organized as follows. The following section introduces the basic LAR coder. In section 3 we describe the general principles of the new Interleaved S+P method, before presenting in the next section the predictors used in the DPCM scheme. The section 5 shows the results in association with the implicit context modelling.

2. LOSSY LAR METHOD FOR GREYSCALE IMAGES

The LAR compression method - firstly designed for lossy greyscale images coding - is a two-layer codec : a spatial codec and a complementary spectral one. The spatial coder provides a low bit rate compressed image whereas the spectral one encodes the texture. This basic scheme has been improved through the self-extracting regions of interest representation. Extensions also have permitted to address middle and high quality encoding and region-level chromatic images encoding [5]. Furthermore the quality of the low resolution LAR image has been evaluated and recognized to be better than Jpeg-2000 [3].

2.1. Spatial coder

The basic principle of the spatial coder is that the local resolution, i.e. the pixel size, can depend on the local activity. This leads to construct a variable image resolution, determined by means of a quad-tree data structure. The final size of pixels (from 16×16 to 2×2) is computed thanks to a local gradient estimation. Through this decomposition, the pixel size gives implicitly the nature of the blocks: indeed, small ones are located on contours whereas large ones are situated on smooth areas (Fig. 1.b). Perceptible blocks artifacts in homogenous areas are easily removed by a efficient post-processing [6] (Fig. 1.d).

2.2. Spectral coder

In order to obtain higher image quality, the texture (whole error image) can be encoded by the spectral coder that uses a DCT adaptive block-size approach [7]. The use of adapted square size allows a content-based scalable encoding scheme: for example, edge enhancement can be made by only transmitting the AC coefficients of small blocks.

The Interleaved S+P algorithm, presented in the following, is aimed to provide a lossless representation of images. This new method keeps the same main structure as the LAR one, namely a two-layer codec, so that the first part of the compressed bitstream allows the LAR block image reconstruction. A major improvement lies in the fact that the pyramidal decomposition makes possible a scalable transmission of compressed data.

Fig. 1. (a) Original image. (b) Variable block sizes representation (subsampled grid : 0.032 bpp). (c) reconstructed low-resolution image (0.4bpp) resulting from first pyramid descent process. (d) Post-processed image.

3. INTERLEAVED S+P PYRAMID CONSTRUCTION

In this section, we first introduce the general principles of our algorithm before setting out the construction of the pyramid. In the following, z_0 and z_1 denote the S-transformed coefficients in such a way that if (u_0, u_1) is a couple of value, we have:

$$
\begin{aligned}
z_0 &= \lfloor (u_0 + u_1)/2 \rfloor, \\
z_1 &= u_1 - u_0.
\end{aligned}
\tag{1}
$$

3.1. General principles

The LAR pyramidal decomposition is based on an adaption of the Wu predictor described in [8]. For full resolution image, errors are coded by means of three interlaced sampling of the original image. By this way, we tend to obtain a spatial configuration of 360° type context surrounding a given pixel so that the resulting prediction error is drastically reduced.
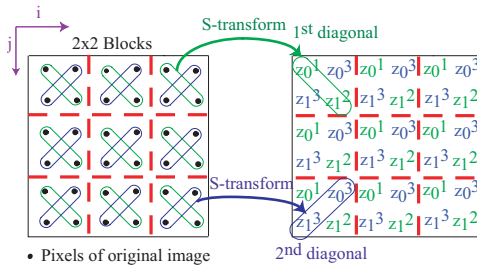
Fig. 2. Original application of the S-Transform

The general principle of the Interleaved S+P algorithm is as follows. The first step consists in applying the 1D S-Transform on the 2 vectors formed by 2 diagonally adjacent pixels in a 2×2 block, as depicted in figure 2.

We note z_i^k, where $i \in \{0, 1\}$ and $k \in \{1, 2, 3\}$, the z_i coefficient coded through the kth pass.
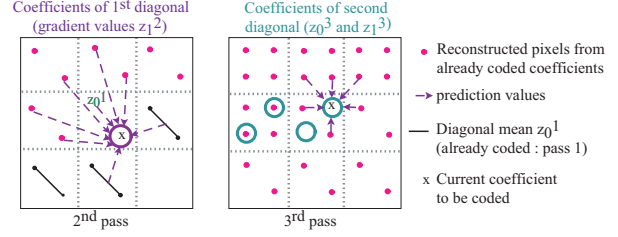
Fig. 3. Second and third pass of the transformed coefficients prediction process.

The first pass encodes through a classical DPCM system an uniform subsampled image formed by the average of two diagonally adjacent pixels within each 2x2 block (z_0^1 transformed coefficients).

Then the second pass predicts the z_1^2 transformed coefficients in order to reconstruct the value of the two pixels of the first diagonal. At this stage, the 360° type prediction context consists in the already known values of the current pass and the diagonal means coded by the first pass.

Finally, the third pass encodes the remaining half of the original image composed of the set of the z_0^3 and z_1^3 S-coefficients. Once again, thanks to the reconstructed pixels resulting from the two previous passes, a completely spatially enclosing and adjacent context is available to predict the transformed pixel (Fig. 3).

3.2. Pyramid construction - Interleaving

Let I the original image be of size $N_x \times N_y$. The multiresolution representation of an image is described by the set $\{Y_l\}_{l=0}^{l_{max}}$, where l_{max} is the top of the pyramid and $l = 0$ the full resolution image.

As an extension of the Wu method, four blocks $\frac{N}{2} \times \frac{N}{2}$ are gathered into one block $N \times N$ valued by the average of the two blocks of the first diagonal (first S-pyramid on fig. 4), that let us write the following equation:

$$
\begin{cases}
l = 0, & Y_0(i,j) = I(i,j); \\
l > 0, & Y_l(i,j) = \left\lfloor \frac{Y_{l-1}(2i,2j) + Y_{l-1}(2i+1,2j+1)}{2} \right\rfloor,
\end{cases}
\tag{2}
$$

with $0 \leq i \leq N_x^l$, $0 \leq j \leq N_y^l$, where $N_x^l = N_x/l$ and $N_y^l = N_y/l$.
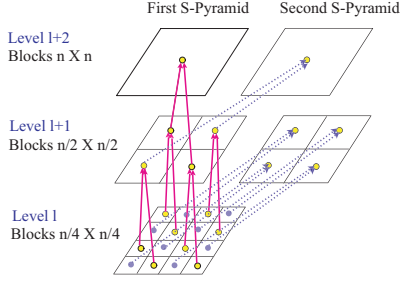
The transformation of the second diagonal of a given 2×2 block can also be seen as a second S-pyramid, where the pixel values depend on the ones existing at the lower level of the first S-pyramid. Interleaving is in this way realized.

4. INTERLEAVED S+P PYRAMID DECOMPOSITION - REFINED PREDICTION MODEL

The Interleaved S+P pyramidal decomposition process results from the extension of the previously described prediction method to the whole pyramid. In this section we describe the implemented prediction scheme.

4.1. Pyramid decomposition principles - Notations

The reconstruction of a 2×2 block located on a given level l is realized in two successive passes. First, the S-transform is applied

Fig. 4. Construction of the pyramid

Fig. 5. Decomposition: extraction of LAR block image data

on pixels $Y_l(2i, 2j)$ and $Y_l(2i + 1, 2j + 1)$ so that

$$z_0^{l,1}(2i, 2j) = \left\lfloor \frac{Y_l(2i, 2j) + Y_l(2i + 1, 2j + 1)}{2} \right\rfloor,$$

$$z_1^{l,2}(2i + 1, 2j + 1) = Y_l(2i, 2j) - Y_l(2i + 1, 2j + 1). \quad (3)$$

We immediately can notice that the $z_0^{l,1}(2i, 2j)$ coefficient is equal to the value of the reconstructed pixel obtained in the upper level of the pyramid:

$$z_0^{l,1}(2i, 2j) = Y_{l+1}(i, j). \quad (4)$$

In a similar way, pixels from the second diagonal of a block are processed as follows:

$$z_0^{l,3}(2i + 1, 2j) = \left\lfloor \frac{Y_l(2i, 2j + 1) + Y_l(2i + 1, 2j)}{2} \right\rfloor,$$

$$z_1^{l,3}(2i + 1, 2j + 1) = Y_l(2i + 1, 2j) - Y_l(2i, 2j + 1). \quad (5)$$

As shown in paragraph 2.2, the entire pyramid is decomposed in two successive descent processes. If the first one is intended to reconstruct the LAR block image, the second one encodes the image texture: the grid information is thus exploited (see figure 5). For this purpose, an image, denoted Siz, containing the size information is elaborated from our quad-tree partitioning. In figure 1.c, one can see the block image as a result of first partial pyramid decomposition.
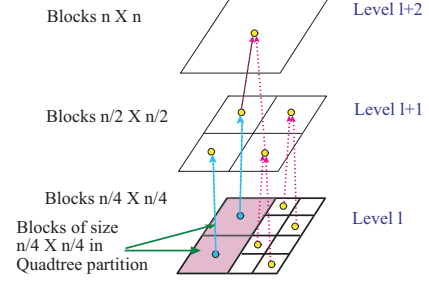
4.2. Top of the pyramid - Classical DPCM

On the highest level of the pyramid, we applied the first pass so that the $Y_{l_{max}}(i, j) = z_0^{l_{max}-1,1}(2i, 2j)$ coefficient values are predicted. For this purpose, we use the classical MED predictor implemented in the compression method LOCO-I [9].

4.3. LAR block image processing

Reconstructing the LAR block image means that for a given level of the pyramid, a block is processed only if the corresponding block size is lower or equal to the level size (*i.e.* $Siz(x \times l, y \times l) \leq 2^l$). Conversely, if $Siz(x \times l, y \times l) > 2^l$), the block values are copied out so that $Y_l(2i, 2j) = Y_l(2i + 1, 2j) = Y_l(2i, 2j + 1) = Y_l(2i + 1, 2j + 1) = Y_{l+1}(i, j)$.

Note that $\breve{z}$ stands for the estimated value of transformed coefficient z.

First S-Pyramid. As the $z_0^{l,1}$ value is already known from the upper level of first S-pyramid, we just have to predict the transformed coefficient $z_1^{l,2}$, gradient value of the first 2×2 block diagonal. In the LAR block image processing context (high activity areas), it turns out that linear prediction is more efficient to estimate $z_1^{l,2}$:

$$\breve{z}_1^{l,2}(2i + 1, 2j + 1) = 2.1 \left[0.9 Y_{l+1}(i, j) + \frac{1}{6} \Big(Y_l(2i + 1, 2j - 1) \right.$$
$$+ Y_l(2i - 1, 2j - 1) + Y_l(2i - 1, 2j + 1) \Big)$$
$$- 0.05 \Big(Y_l(2i, 2j - 2) + Y_l(2i - 2, 2j) \Big)$$
$$\left. - 0.15 \Big(Y_{l+1}(i, j + 1) + Y_{l+1}(i + 1, j) \Big) - Y_{l+1}(i, j)) \right]. \quad (6)$$

Second S-Pyramid. The third pass of adapted Wu predictor encodes the coefficients from second diagonal. Estimation of $z_0^{l,3}$ uses inter- and intra-level prediction from reconstructed values, so that

$$\breve{z}_0^{l,3}(2i + 1, 2j) = \alpha_0 \frac{1}{4} \Big(Y_l(2i - 1, 2j + 1) + Y_l(2i, 2j + 2)$$
$$+ Y_l(2i + 2, 2j) + Y_l(2i + 1, 2j - 1) \Big) \quad (7)$$
$$+ \beta_0 \widehat{z}_0^{l,1}(2i, 2j),$$

where $(\alpha_0, \beta_0) = (0.25, 0.75)$, and $\widehat{z}_0^{l,1}$ is the reconstructed value of coefficient $z_0^{l,1}$. The estimated value $\breve{z}_1^{l,3}(2i, 2j)$ is computed as follows:

$$\breve{z}_1^{l,3}(2i, 2j + 1) = \alpha_1 \Big(Y_l(2i - 1, 2j + 1) + Y_l(2i, 2j + 2)$$
$$- Y_l(2i + 1, 2j - 1) - Y_l(2i + 2, 2j) \Big)$$
$$- \beta_1 \Big(Y_l(2i - 1, 2j) + Y_l(2i - 1, 2j + 2)$$
$$- Y_l(2i, 2j - 1) - \breve{Y}_l(2i, 2j + 1) \Big). \quad (8)$$

with $(\alpha_1, \beta_1) = (3/8, 1/8)$. $\breve{Y}_l(2i, 2j + 1)$ corresponds to the Wu predictor [8] for third pass applied to the pixel $Y_l(2i, 2j + 1)$.

4.4. Texture processing

The second descent process of the pyramid allows to reconstruct the texture. On each level, pixel of size $Siz(x \times l, y \times l) > 2^l$ are

here evaluated. Once again, we can distinguish prediction of first S-pyramid from the second one.

First S-Pyramid. Texture blocks are characterized by their low local activity. Thus gradient values $z_1^{l,2}(2i, 2j)$ are difficult to estimate through linear prediction. This is the reason why we use median values extracted from close neighborhood.

Let $m_e(u_1, u_2, \ldots, u_n)$ be the median value of a set of n values $(u_1, u_2, \ldots, u_n)$. The estimation of coefficient $z_1^{l,2}(2i + 1, 2j + 1)$ located in a texture area is

$$\check{z}_1^{l,2}(2i+1, 2j+1) = \frac{1}{4}\Big(m_e\big(Y_l(2i-2, 2j), Y_l(2i, 2j-2),$$
$$Y_l(2i-1, 2j-1)\big) \quad (9)$$
$$+ m_e\big(Y_{l+1}(i+1, j), Y_{l+1}(i, j+1), Y_{l+1}(i+1, j+1)\big)\Big).$$

Second S-Pyramid. Context is there sufficient to evaluate precisely coefficient values, thanks to linear predictors. $z_0^{l,3}$ are then estimated through the equation 7, where $(\alpha_0, \beta_0) = (0.37, 0.63)$. Predicted coefficients $\check{z}_1^{l,3}(2i, 2j)$ are obtained by the application of the relation 8, where $(\alpha_1, \beta_1) = (1/4, 0)$.

5. RESULTS - IMPLICIT CONTEXT MODELLING

In table 1, each of the test images was losslessly compressed using the proposed Interleaved S+P. The resulting first order entropy can be directly compared with the ones obtained when using CALIC and S+P methods. Note that no entropy encoder was implemented there. The Interleaved S+P algorithm, includind quad-tree partitioning, outperforms noticeably classical S+P and is, on the whole, equivalent to CALIC in terms of entropy. In addition, our method provides a scalable image encoding both in terms of resolution and quality. Moreover, functionalities such as free segmentation map and regions of interest processing are still available.

Thanks to our specific non-uniform subsampling, entropy is drastically reduced, as shown in the table 1. Indeed global entropy can be reduced when different classes of symbols following the same law can be isolated. Our method provides a straightforward separation of the laws: for each level of the pyramid, we distinguish texture information from LAR block image, pass 2 from pass 1, and z_0 coefficients from z_1 values. Consequently, due to our content-based grid representation, an implicit context modelling is realized.

| Image | Entropy (bpp) | | | |
	Interleaved S+P - No partition	**Interleaved S+P**	CALIC	S+P
Barb2	5.31	4.96	4.93	5.04
Cafe	5.76	5.43	5.37	5.42
Gold	4.82	4.64	4.65	4.73
Hotel	4.91	4.66	4.57	4.97
Lena	4.46	4.30	4.33	4.33
Peppers	4.74	4.57	4.58	4.67
tools	5.78	5.52	5.53	5.73
us	4.01	3.55	3.60	3.78
Average	4.97	**4.70**	**4.70**	4.83

Table 1. First-order entropy (bit/pixels) of the proposed Interleaved S+P (with quad-tree partitioning or not), CALIC and S+P.

6. CONCLUSION

This paper has presented an original technique for scalable lossy and lossless compression. Based on the LAR method, this solution uses the S-transform in such a manner that a refined prediction context is available for each estimation steps. The knowledge of the quad-tree data structure is sufficient to improve significantly the compression ratio for the lossless coding, by gathering and coding separately the information of the same nature (transform coefficients, low resolution LAR image, contours...).

This Interleaved S+P algorithm performs significantly better than the classical S+P transform, and is globally equivalent to CALIC. However, the proposed technique not only allows progressive reconstruction of images, but also provides a free segmentation map at both the coder and decoder so that regions of interest can be easily supported.

Our current research focuses on the adaptation of the LAR scheme to the video. Other work concerns the elaboration of a joint source/channel image coding approach, as a solution for secured transmission of medical images through low-bandwidth networks like the Internet.

7. REFERENCES

[1] X. Wu and N. Memon and, "A Context-based, Adaptive, Lossless/Nearly-Lossless Coding Scheme for Continuous-Tone Images (CALIC)," *International Standards Organization working document, ISO/IEC SC29/WG 1/N256*, vol. 202, 1995.

[2] A. Said and W. Pearlman, "Reversible Image Compression via Multiresolution Representation and Predictive Coding," in *Visual Communication and Image Processing*. SPIE, 1993, vol. 269, pp. 664–374.

[3] O. Deforges and J. Ronsin, "Region of Interest Coding for Low Bit-Rate Image Transmission," in *Proc. ICME'2000*, July 2000, vol. 1, 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*, Hong Kong, April 6-10 2003, Conference cancelled - Invited paper in ICME 2003, vol. III, pp. 249–252.

[5] O. Deforges and J. Ronsin, "Supervised Segmentation at Low Bit Rates for Region Representation and Color Image Compression," in *Proc. ICME'2002*, Aug. 26-29 2002, vol. 1, pp. 665 − 668.

[6] D. Muresan and T. Parks, "Optimal Recovery Approach to Image Interpolation," in *ICIP'01*, 2001, vol. 3, pp. 848–851.

[7] C. Chen, "Adaptive transform coding via quad-tree based variable block-size DCT," in *ICASSP'89*, 1989, pp. 1854–1856.

[8] X. Wu, "Lossless Compression of Continuous-Tone Images, via Context Selection, Quantization and Modelling," *IEEE Trans. on Image Processing*, vol. 6, no. 5, pp. 656–664, May 1997.

[9] M.J. Weinberger, G. Seroussi, and G. Sapiro, "Loco-i : a low complexity, context-based, lossless image compression algorithm," *Proc. Of the IEEE Data Compression Conference*, pp. 141–150, 1996.