Adaptive transmission for lossless image reconstruction

Elisabeth Lahalle, Gilles Fleury, Rawad Zgheib

ADAPTIVE TRANSMISSION FOR LOSSLESS IMAGE RECONSTRUCTION

Elisabeth Lahalle, Gilles Fleury, Rawad Zgheib

Department of Signal Processing and Electronic Systems, Supélec, Gif-sur-Yvette, France
E-mail : firstname.lastname@supelec.fr
tel: +33 (0)1 69 85 14 27, fax: +33 (0)1 69 85 14 29

ABSTRACT

This paper deals with the problem of adaptive digital transmission systems for lossless reconstruction. A new system, based on the principle of non-uniform transmission, is proposed. It uses a recently proposed algorithm for adaptive stable identification and robust reconstruction of AR processes subject to missing data. This algorithm offers at the same time an unbiased estimation of the model's parameters and an optimal reconstruction in the least mean square sense. It is an extension of the RLSL algorithm to the case of missing observations combined with a Kalman filter for the prediction. This algorithm has been extended to 2D signals. The proposed method has been applied for lossless image compression. It has shown an improvement in bit rate transmission compared to the JPEG2000 as well as the JPEG-LS standards.

***Index Terms*—** adaptive, lossless, compression

1. INTRODUCTION

Lossless compression methods are important in many medical applications where large data set need to be transmitted without any loss of information. Actually, some lesions risk becoming undetectable due to the effects of lossy compression. General lossless compression coders are considered to be composed of two main blocks: a data decorrelation block and an entropy coder for the decorrelated data. Two main tendencies may be noticed for the methods used for the decorrelation step: methods based on wavelet transforms and methods based on predictive coding. They have led to the main compression standards : the JPEG2000 for the former group of methods [1], the JPEG-LS for the latter [2]. Intensive attention is paid to transform based compression methods with many algorithms which perform well regarding the bit rate such as SPIHT [3], QT [4], etc.

All these coders use a uniform transmission of the binary elements to transmit. In a previous paper [5], the design of digital systems based upon non-uniform transmission of signal samples was introduced. The idea behind is to avoid sending a sample if it can be efficiently predicted, e.g. with a prediction error smaller than the quantization one, thus reducing the average transmission bit rate and increasing the signal to noise ratio (SNR). A speech coder based on the Adaptive Pulse Code Modulation (ADPCM) principle and non-uniform transmission of signals have already been proposed in [6]. It uses the Least Mean Square (LMS)-like algorithm [7] for the prediction of the samples that were not sent. However, this algorithm converges toward biased estimations of the model's parameters and does not use an optimal predictor in the least mean square sense. Recently, we proposed a Recursive Least Square Lattice (RLSL) algorithm for adaptive stable identification of non stationary Autoregressive (AR) processes subject to missing data, using a Kalman filter as a predictor [8]. This algorithm is fast, guarantees the stability of the model identified and offers at the same time an optimal reconstruction error in the least mean square sense and an unbiased estimation of the model's parameters in addition to the fast adaptivity to the variations of the parameters in the case of non stationary processes. Non stationnary AR processes can model a large number of signals in practical situations, such as images in the bi-dimensional case [9]. A new lossless image coder based on a non-uniform transmission principle is proposed: it is based on an adaptation of the algorithm proposed in [8] for optimal prediction and identification of 2D AR processes subject to missing observations.

In the following, begin by presenting the non-uniform transmission idea for lossless compression. In a second part, the adaptive algorithm for reconstruction of AR processes with missing observations [8] is described and extended to 2D AR processes. Its integration into a non-uniform transmission system is studied in the third section. Finally, an example illustrates the performances of the proposed system. It is compared to a uniform digital transmission system : the JPEG2000.

2. NON-UNIFORM TRANSMISSION SYSTEM FOR LOSSLESS RECONSTRUCTION

The proposed system uses predictive coding and non-uniform transmission to reduce the bit rate transmission. An AR signal modeling is considered for the prediction. Let x_n be the amplitude of the signal at time n. The prediction of a sample will be noted $\hat{x}_{n,P}$ and the prediction error $e_{n,P} = x_n - \hat{x}_{n,P}$. In the receiver, a sample x_n is predicted using the estimated model parameters at time $n-1$, $\hat{\mathbf{a}}_{n-1}$, and the available samples. The key ideas of the proposed system are the following. If $e_{n,P} \approx 0$, x_n is replaced by $\hat{x}_{n,P}$ in the receiver without any loss, requiring only one bit flag to be transmitted for the first and the last sample where $e_{n,P} \approx 0$. If an efficient prediction method for non-uniformly sampled data is used, the above situation occurs many times during the transmission. This is the case for example outside the region of interest of the image where the sample value is constant or null. The whole number of transmitted samples is thus considerably reduced. As some of the samples are not transmitted, the receiver has to deal with the problem of online identification and reconstruction of signals subject to missing samples. The probability law of the prediction error of the image to transmit is then used to adapt the number of bit coding the prediction error in the case where it is non zero.

3. PREDICTION/RECONSTRUCTION FOR NON-UNIFORMLY SAMPLED DATA

3.1. Kalman RLSL algorithm

Let $\{x_n\}$ be an AR process of order L with parameters $\{a_k\}$, and $\{\epsilon_n\}$ the corresponding innovation process of variance σ_ϵ^2. The loss process is modeled by an i.i.d binary random variable $\{c_n\}$, where $c_n = 1$ if x_n is available, otherwise $c_n = 0$. Let $\{z_n\}$ be the reconstruction of the process $\{x_n\}$. If x_n is available $z_n = x_n$, otherwise, $z_n = \hat{x}_n$, the prediction of x_n. In order to identify, in real time, the AR process subject to missing data, the algorithm proposed in [8] can be summarised as follows. The reflection coefficients of the lattice structure are determined by minimizing the weighted sum of the quadratic forward, $f_t^{(l)}$, and backward, $b_t^{(l)}$, prediction errors :

$$E_n^{(l)} = \sum_{i=1}^{n} w_{n-i} \left(f_n^{(l)2} + b_n^{(l)2} \right). \tag{1}$$

A Kalman filter provide an optimal prediction of the signal using the AR estimated parameters. These parameters are deduced from the estimated reflection coefficients using the Durbin Levinson recursions. At time $n+1$, the first line of the matrix A of the state space representation of an AR process is built with $\hat{\mathbf{a}}_n^{(L)\top}$, the vector of the parameters estimated at time n. The matrix is then named A_{n+1}.

$$\begin{aligned} A_{n+1} &= \left[\begin{array}{ccc|c} \hat{a}_{1,n}^{(L)} & \dots & \dots & \hat{a}_{L,n}^{(L)} \\ \hline 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{array} \right], \\ P_{n+1|n} &= A_{n+1} P_{n|n} A_{n+1}^\top + R_\epsilon, \\ \hat{\mathbf{x}}_{n+1|n} &= A_{n+1} \hat{\mathbf{x}}_{n|n} \\ \hat{y}_{n+1|n} &= c_{n+1} \hat{x}_{n+1|n} \end{aligned} \tag{2}$$

If x_{n+1} is available, i.e. $c_{n+1} = 1$,

$$K_{n+1} = P_{n+1|n} \mathbf{c}_{n+1} (\mathbf{c}_{n+1}^\top P_{n+1|n} \mathbf{c}_{n+1})^{-1}, \tag{3a}$$

$$P_{n+1|n+1} = (I_d - K_{n+1} \mathbf{c}_{n+1}^\top) P_{n+1|n}, \tag{3b}$$

$$\hat{\mathbf{x}}_{n+1|n+1} = \hat{\mathbf{x}}_{n+1|n} + K_{n+1} (y_{n+1} - \hat{y}_{n+1|n}) \tag{3c}$$

The predictions of the previous missing data up to time $n - L + 1$ are updated thanks to the filtering of the state in equation (3c). It is convenient now to calculate all the variables of the lattice filter since the last available observation at time $n - h$, where $h \geq 0$ depends on the observation pattern. At each time t, for $n - h + 1 \leq t \leq n + 1$, the recursive equations of the RLSL algorithm given by (5) are applied to estimate the different reflection coefficients $\hat{k}_t^{(l)}$ and prediction errors $\hat{f}_t^{(l)}, \hat{b}_t^{(l)}$ for $1 \leq l \leq L$. The values of the forward and backward prediction errors are initialized using the updated estimates of the missing samples (those contained within the filtered state $\hat{\mathbf{x}}_{n+1|n+1}$), i.e. $\hat{f}_t^{(0)} = \hat{b}_t^{(0)} = \hat{x}_{t|n+1}$.

Hence,

- For $t = n - h + 1$ to $n + 1$
  - Initialize for $l = 0$

$$\hat{f}_t^{(0)} = \hat{b}_t^{(0)} = \hat{x}_{t|n+1}, \hat{k}_t^{(0)} = 1, \tag{4}$$

  - For $l = 1$ to $\min(L, n)$

$$C_t^{(l)} = \lambda C_{t-1}^{(l)} + 2 \hat{f}^{(l-1)_t} \hat{b}_{t-1}^{(l-1)}, \tag{5a}$$

$$D_t^{(l)} = \lambda D_{t-1}^{(l)} + \hat{f}_t^{(l-1)2} + \hat{b}_{t-1}^{(l-1)2}, \tag{5b}$$

$$\hat{k}_t^{(l)} = -\frac{C_t^{(l)}}{D_t^{(l)}}, \tag{5c}$$

$$\hat{f}_t^{(l)} = \hat{f}_t^{(l-1)} - \hat{k}_t^{(l)} \hat{b}_{t-1}^{(l-1)}, \tag{5d}$$

$$\hat{b}_t^{(l)} = \hat{b}_{t-1}^{(l-1)} - \hat{k}_t^{(l)} \hat{f}_t^{(l-1)}, \tag{5e}$$

  - end
- end.

The AR parameters at time $n+1$, $(\hat{a}_{i,n+1}^{(L)})_{1 \leq i \leq L}$, are deduced from the reflection coefficients $(\hat{k}_{n+1}^{(l)})_{1 \leq i \leq L}$ using the Durbin Levinson recursions. owever if x_{n+1} is absent, $c_{n+1} = 0$, the predicted state, $\hat{\mathbf{x}}_{n+1|n}$, is not filtered by the Kalman filter, and the parameters are not updated since the reflection coefficients $(\hat{k}_{n+1}^{(l)})_{1 \leq l \leq L}$ are not yet calculated,

$$K_{n+1} = 0, \tag{6a}$$

$$P_{n+1|n+1} = P_{n+1|n}, \tag{6b}$$

$$\hat{\mathbf{x}}_{n+1|n+1} = \hat{\mathbf{x}}_{n+1|n}, \tag{6c}$$

$$\hat{\mathbf{a}}_{n+1}^{(L)} = \hat{\mathbf{a}}_n^{(L)}. \tag{6d}$$

The cost function minimized by this algorithm is the weighted mean of all quadratic prediction errors. When a sample is missing, the prediction error can not be calculated, it is replaced by its estimation. Indeed, recall that in order to update the reflection coefficients at a time n, the lattice filter variables must have been calculated at all previous times. Therefore, using this algorithm, the lattice filter variables are estimated at all times even when a sample is missing. Consequently, this algorithm presents an excellent convergence behavior and have fast parameter tracking capability even for a large probability of missing a sample. The computational complexity of this algorithm is found to be $O((1-q)NL^2)$, where q is the bernoulli's probability of losing a sample, N is the size of the signal and L the order of the AR model.

3.2. Adaptation to 2D signals

A first solution to use the previous algorithm for 2D signals is to use the classical video scanning of the image in order to get a 1D signal. However, only a 1D decorrelation is achieved using this method.
In order to get a 2D decorrelation of the image, a 2D AR predictor $\hat{x}_{i,j}$ of the sample $x_{i,j}$ (7) must be used in addition to the video scanning of the image.

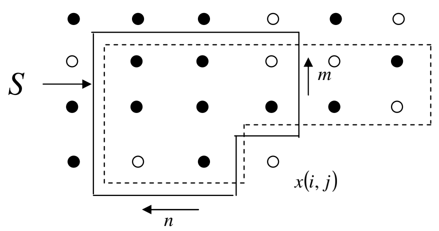

Fig. 1. AR 2D: prediction support

$$\hat{x}_{i,j} = \sum_{n,m \in S} \hat{a}_{n,m} x_{i-n,j-m} \tag{7}$$

In order to integrate this 2D AD predictor into the previous algorithm, the first line of the A matrix is built with the $\hat{a}_{n,m}$ parameters, and the regressor vector $[x_{n-1} \dots x_{n-L}]^\top$ is replaced by $[x_{i-1,j} \dots x_{i-n,j-m} \dots x_{i-p,j-q}]^\top$. The renumbering task excepted, to built the A matrix, the computational time of these 2D algorithm is similar to the 1D one.

4. PROPOSED ADAPTATIVE TRANSMISSION ALGORITHM

In this section, we propose to use the algorithms discussed in section 3 as efficient predictors in the non uniform transmission system proposed in section 2 in order to minimize the number of bit to transmit. At each time n, knowing all transmitted samples and using the same identification and reconstruction method as the one used in the receiver, the transmitter evaluates the signal reconstruction performance in the receiver. This can be done by comparing the receiver prediction error, $|e_{n,P}|$, with different thresholds, $\mathcal{S}_1 \approx 0, \mathcal{S}_2, ..., \mathcal{S}_i$. Thus, if the receiver is able to reconstruct the sample without error (error greatly smaller than the quantification error $(1e^{-5})$), only a one bit flag is transmitted to indicate the first and the last missing sample. The number of thresholds $\mathcal{S}_i$ and their values are chosen according to the probability law of the prediction error to transmit only the B_i bits required to code the prediction error for each threshold. The proposed coding decoding algorithm can be summarized, at a time n, as:

- In the transmitter:
  . $e_{n,P} = x_n - \hat{x}_{n,P}$
  . if ($|e_{n,P}| = 1e^{-5}$ and $|e_{n-1,P}| > 1e^{-5}$ or $|e_{n,P}| > 1e^{-5}$ and $|e_{n-1,P}| = 1e^{-5}$), one bit flag is transmitted,
  . else if $|e_{n,P}| < \mathcal{S}_2$,
  . if $|e_{n,P}| < \mathcal{S}_3$, B_3 bits are transmitted,
  . else B_2 bits are transmitted,
  . else B_1 bits are transmitted.

- In the receiver, the method described in 3 is used for adaptive identification and reconstruction of a signal subject to missing data: if a new sample is received, the AR parameters are updated. Otherwise, the missing sample is predicted in terms of the past available samples and the current estimation of the parameters.

5. SIMULATIONS

The performances of both proposed methods are compared to the JPEG2000. The first method uses a 1D AR model of order 3 of the signal. In the second method, the image is modeled by a 2D AR process of order $(2,2)$. The performances of the different methods are evaluated in term of bit rate (in bpp) on CT images. The PSNR is computed for the proposed methods to show the lossless reconstruction of the image. The PSNR which have been reached for all the simulations corresponds to the infinity value. Table 1 shows the results for CT images of (512x512x12) bits presented in figures 2, 3 and 4 (Images courtesy of Dr Kopans, MGH Boston, USA. Tomosynthesis investigational device from GE Healthcare (Chalfont St Giles, UK)). In these images the prediction error is in most of the case small (lower than 32), but for the pixels of the edge of the ROI the prediction error requires 12 bits to be coded. Consequently, the following values are chosen for the number of bit to code the prediction error : $B_1 = 13$, $B_2 = 8$, $B_3 = 6$.

6. CONCLUSION

A new digital transmission system for lossless image reconstruction has been proposed. It is based on a non-uniform transmission principle and on extensions to 2D of the algorithm proposed in [8] for real time identification and reconstruction of AR processes subject to missing data. The pro-

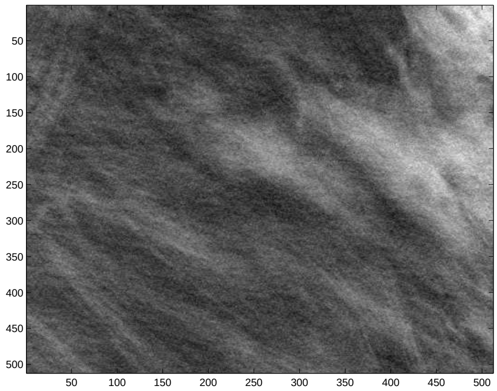

Fig. 2. CT1 image

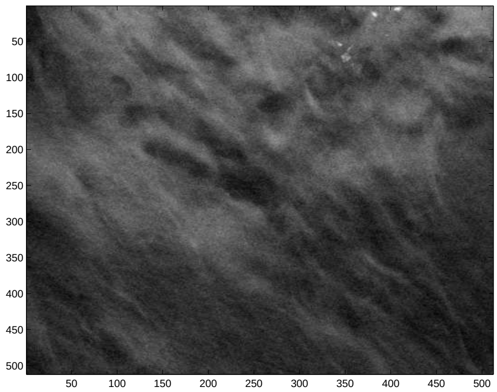

Fig. 3. CT2 image

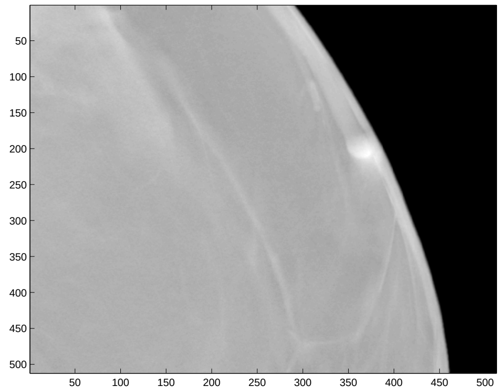

Fig. 4. CT3 image

posed methods, applied on CT images, has shown in their two forms (2D as well as 1D) an improvement in bit rate comparing to the JPEG2000 and JGPEG-LS standards. Comparing to the JPEG2000, significant gains for lossless compression are reached: 3.4% for CT3 image up to 4.6% for CT1 image. Comparing to the JPEG-LS, the most significant gains (2.7% up to 3.6%) are reached for CT2 and CT1 images where the RLE coding of the JPEG-LS is not used.

7. REFERENCES

[1] ISO/IEC 15444-1, "Information technology - jpeg2000 image coding system," JPEG2000 standard, Part 1-Core coding system, 2000.

[2] ISO/IEC 14495-1, "Information technology - lossless and near-lossless compression of continuous-tone still images," JPEG-LS standard, Baseline, 2000.

[3] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitionning in hierarchical trees," *IEEE Trans. on Circuits and systems for Video Technology*, vol. 6, pp. 243–250, June 1996.

[4] A. Munteanu and J. Cornelis, "Wavelet based lossless compression scheme with progressive transmission capability," *International Journal of Imaging Systems and Tecnology*, vol. 10, pp. 76–85, January 1999.

[5] S. Mirsaidi, G. Fleury, and J. Oksman, "Reducing quantization error using prediction/non uniform transmission," in *Proc. International Workshop on Sampling Theory and Applications*. IEEE, 1997, pp. 139–143.

[6] E. Lahalle and J. Oksman, "ADPCM speech coder with adaptive transmission and ARMA modelling of non-uniformly sampled signals," in *5th Nordic Signal Processing Symposium, CD-ROM proceedings, Norway*. IEEE, 2002.

[7] S. Mirsaidi, G. Fleury, and J. Oksman, "LMS like AR modeling in the case of missing observations," *IEEE Transactions on Signal Processing*, vol. 45, pp. 1574–1583, June 1997.

[8] R. Zgheib, G. Fleury, and E. Lahalle, "Lattice algorithm for adaptive stable identification and robust reconstruction of non stationary ar processes with missing observations," *IEEE Transactions on Signal Processing*, vol. 56, pp. 2746–2754, July 2008.

[9] N. S. Jayant and P. Noll, "Digital coding of waveform, principles and applications to speech and video," Prentice Hall, 1984.

Table 1. Comparison of the three methods in bit rate (in bpp) for CT images of (512x512x12) bits:

Method	$CT1$	$CT2$	$CT3$
1	6.53	6.67	5.10
2	6.45	6.61	5.10
JPEG2000	6.76	6.89	5.28
JPEG-LS	6.69	6.79	5.15