Signal Processing: *Image Communication* 4 (1992) 457–476
Elsevier

Predictive coding of images using an adaptive intra-frame predictor and motion compensation

G. Tziritas and J.C. Pesquet
Laboratoire des Signaux et Systèmes (C.N.R.S.), École Supérieure d'Electricité, Plateau de Moulon, 91192 Gif-sur-Yvette Cédex, France

Received 7 May 1990
Revised 3 July 1991

Abstract. In this article we present an image predictive coding method using both intra- and inter-frame predictors. The intra-frame predictor is an adaptive FIR filter using the well-known LMS algorithm to track continuously spatial local characteristics of the intensity. The inter-frame predictor is motion-adaptive using a pel-recursive method estimating the displacement vector. Weight coefficients are continuously adapted in order to favor the prediction mode which performs better between intra-frame and motion compensation mode. It is a backwards adaptation which does not necessitate the transmission of an overhead information. Neither the weight coefficients nor displacement vectors are transmitted. Apart from the quantized prediction error, it may be necessary to transmit the detection of a discontinuity of the displacement vector. For the examined image sequence a significant improvement is obtained in comparison with only adaptive intra-frame or only motion compensation mode. We give and discuss the extension of a known adaptive quantizer for 2D signals. A crucial problem in predictive coding, particularly with adaptive techniques, is the sensitivity to transmission errors. A method ensuring the self-adjustment of the decoder in the presence of transmission errors, which do not affect the pixel synchronization, is proposed for the intra-frame mode. Neither overhead information nor error-correcting codes are needed.

Keywords. Predictive image coding; adaptive prediction; motion compensation; adaptive quantization; channel errors.

1. Introduction

Among different image coding methodologies, predictive coding can be simply implemented and produces good results at higher rates. In this paper, we are interested in adaptive methods of predictive coding. The block diagram of the predictive coder/decoder discussed in this paper is shown in Fig. 1. It contains an adaptive predictor and an adaptive quantizer, the adaptation being realized using the quantized prediction error $\bar{e}$, which is the transmitted information only. In Fig. 1, x is the original signal, $\bar{x}$ is the signal reconstructed by the decoder, $\hat{x}$ is the predicted signal and e is the prediction error. As a result of eventual transmission errors the received prediction error may be different, $\bar{e}'$, and similarly will be concerning the predicted, $\hat{x}'$, and the reconstructed, $\bar{x}'$, signals.

Knowing that there exist simultaneously spatial and temporal redundancies, we consider a hybrid structure of intra- and inter-frame prediction (Fig. 2). The spatial part of the predictor is a linear filter. Given that the statistical characteristics of the luminance vary spatially and that this signal is characterized by many nonstationarities, the predictive filter must be space-varying. An approach which permits tracking of the variations is an adaptive filter. A method of adapting the 2D filter consists of switching between different filters, switching based on a classification rule. Sometimes this technique requires the transmission of

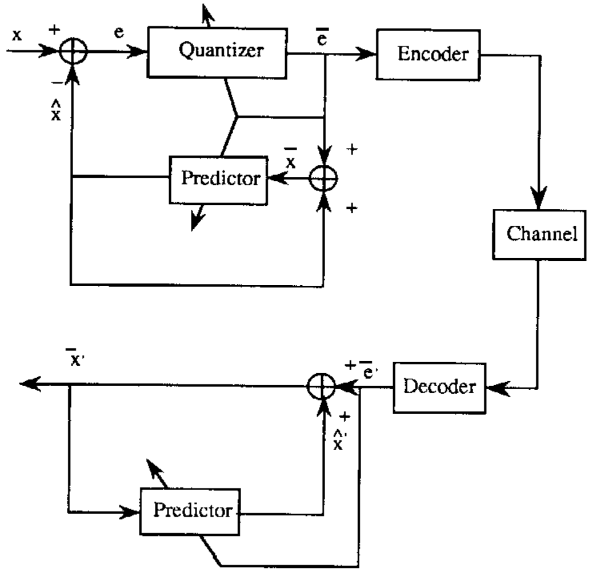

Fig. 1. The structure of the adaptive predictive coder/decoder.

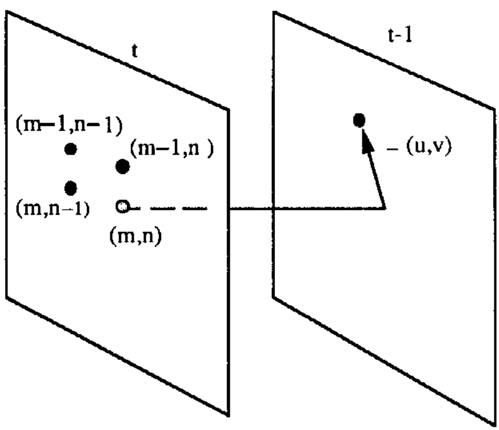

Fig. 2. An adaptive intra-inter-frame predictor.

overhead information, i.e. the class of each pixel. Another adaptive method is to identify the optimum linear filter for a block of the frame using a least squares criterion [11]. The coefficients of the filter must also be transmitted in this case. To avoid the transmission of additional information we here use a continuously adapted filter, the adaptation being based on previously reconstructed pels. The adaptive algorithm used is that of Least Mean Squares (LMS) first introduced by Widrow. The temporal part of the predictor is adapted using an estimator of the displacement vector. A pel-recursive estimator is used, similar to the one of Walker and Rao [17] and of Cafforio and Rocca [3]. We discuss the displacement estimator in more detail in Section 2.

Thus we have an adaptive filter for spatial prediction and a motion adaptive inter-frame prediction. It is obvious that in some regions one type of the described predictions performs better than another, in the sense that the prediction error is minimized. We must then adapt between these two types of prediction. We here propose to continuously adapt a weight coefficient for the inter-frame predictor and the coefficients of a traversal intra-frame filter. The predictor can then be written as

$$\hat{I}(m, n; t) = \sum_{(k,l)\in S} a(k, l)\bar{I}(m-k, n-l; t) + b\bar{I}(m-\hat{v}, n-\hat{u}; t-1), \tag{1}$$

where I is the luminance or the intensity, $\bar{I}$ is the reconstructed intensity, m (respectively n) is the row (respectively column) index, t is the temporal coordinate, S is a quarter-plane or a nonsymmetric half-plane domain, $\hat{u}$ (respectively $\hat{v}$) is the horizontal (respectively vertical) component of the estimated displacement vector. The coefficients of the spatial filter $\{a(k, l)\}$ and the weight coefficient b are continuously adapted using the LMS algorithm.

The prediction error which is quantized and transmitted is not stationary and its probability distribution is unknown. A reduction of the distortion is obtained, if the quantizer is not fixed, but adapted to the statistics of the prediction error. An extension of known techniques in 1D predictive coding is given in this paper for a scalar three-level quantizer.

It is known that predictive coding is sensitive to transmission errors, even with constant length codes. The sensitivity to transmission errors is greater if the predictor is adaptive and this adaptation is based on the prediction error. A topic also studied in this paper is the adjustment of the decoder in the presence of transmission errors. We only consider errors which do not affect the pixel synchronization and we demonstrate that the self-adjustment of the decoder can be obtained using some regularization constraints.

The organization of the paper is as follows. In Section 2 the algorithm of motion estimation is briefly presented and discussed. In Section 3 the algorithm of adaptive updating of the coefficients is given. In Section 4 the adaptive quantizer used in this paper is presented and discussed. Section 5 presents constraints and modifications used to obtain the self-adjustment of the decoder in the case of transmission errors. Section 6 gives some results of the algorithms proposed in this paper to illustrate and evaluate their performances.

2. Motion compensation

The displacement vector is estimated for each pixel. The algorithm used here to estimate the displacement vector is similar to that of Walker and Rao [17] and Cafforio and Rocca [3]. It is a pel-recursive intensity-based algorithm which is presented below. The estimator is composed of three parts. The first part is an a priori estimator using the estimated values of the displacement vector in the causal neighborhood of the current pixel. The second part is a detector of discontinuities on the velocity field. The third part is an a posteriori estimator based on the minimization of the Displaced Frame Difference (DFD).

2.1. A priori estimation of the displacement vector

The a priori estimator used here is based on the hypothesis that the displacement vector is slowly varying. Therefore it is assumed that the motion is locally translatory. The simplest a priori estimation (u^0, v^0) at pixel (m, n) under this assumption, knowing the already estimated values $(\hat{u}, \hat{v})$ of the displacement vector, is obtained from the previous pixel in the raster scan order:

$$\begin{bmatrix} u^0(m, n) \\ v^0(m, n) \end{bmatrix} = \begin{bmatrix} \hat{u}(m, n-1) \\ \hat{v}(m, n-1) \end{bmatrix}, \quad \text{for } n > 0, \tag{2}$$

and for the first column:

$$\begin{bmatrix} u^0(m,0) \\ v^0(m,0) \end{bmatrix} = \begin{bmatrix} \hat{u}(m-1,0) \\ \hat{v}(m-1,0) \end{bmatrix}, \quad \text{for } m>0.$$

2.2. Detection of a discontinuity

As the estimation method is based on the intensity using local measurements, the algorithm may estimate false displacements. Discontinuities on the real velocity vector field also exist, because in natural scenes independent 3D rigid movements, or edges between different surfaces which are subject to the same 3D rigid motion may exist. An efficient estimate of the velocity vector necessitates the joint detection of all these discontinuities. This detection must be based on the correctness of the a priori estimation of the displacement vector and consequently on the prediction error of the intensity. The prediction error based on the displacement estimation is the displaced frame difference given in

$$e^0(m,n) = I(m,n;t) - I(m-v^0, n-u^0; t-1). \tag{3}$$

Two approaches, which are described below, can be considered.

(1) The test of discontinuity is realized at the current pixel using the current a priori DFD. If a discontinuity is detected, a reset code is transmitted to set the a priori estimation to zero. To determine the test we assume that $I(m,n;t)$, knowing the displacement vector, is a Laplacian (or Gaussian) distributed variable with mean $I(m-v, n-u; t-1)$ and standard deviation σ ($\equiv$hypothesis 0: no discontinuity). Under hypothesis of discontinuity $I(m,n;t)$ is assumed Laplacian-distributed (or respectively Gaussian) with mean $I(m,n;t-1)$ and the same standard deviation. The test is determined by maximizing the probability a posteriori. We suppose known probabilities of discontinuity (p_1) and no discontinuity ($p_0 > p_1$). Some calculations give the following test: in the Laplacian case a discontinuity is detected if

$$|I(m,n;t) - I(m-v^0, n-u^0; t-1)| - |I(m,n;t) - I(m,n;t-1)| > \frac{\sigma}{\sqrt{2}} \ln \frac{p_0}{p_1}, \tag{4}$$

and in the Gaussian case if

$$(I(m,n;t) - I(m-v^0, n-u^0; t-1))^2 - (I(m,n;t) - I(m,n;t-1))^2 > 2\sigma^2 \ln \frac{p_0}{p_1}. \tag{5}$$

(2) The test of discontinuity is realized in a causal neighborhood S_d of the current pixel. If a discontinuity is detected, the a priori estimation is set to zero, but there is no need to transmit a reset code. Using similar assumptions as in case (1), we obtain the following test: a discontinuity is detected in the Laplacian case if

$$\sum_{(k,l)\in S_d} |I(m-k, n-l; t) - I(m-k-v^0, n-l-u^0; t-1)| - |I(m-k, n-l; t) - I(m-k, n-l; t-1)|$$

$$> \frac{\sigma}{\sqrt{2}} \operatorname{card}[S_d] \ln \frac{p_0}{p_1}, \tag{6}$$

card$[S_d]$ being the cardinal of set of pixels S_d, and in the Gaussian case if

$$\sum_{(k,l)\in S_d} (I(m-k;n-l;t)-I(m-k-v^0,n-l-u^0;t-1))^2$$
$$-(I(m-k,n-l;t)-I(m-k,n-l;t-1))^2>2\sigma^2\,\mathrm{card}[S_d]\ln\frac{p_0}{p_1}. \tag{7}$$

2.3. A posteriori estimation of the displacement vector

Independently of the detection of a discontinuity the displacement vector must be updated at each point. The criterion which is optimized is the square of the a posteriori displaced frame difference:

$$e(m,n)=I(m,n;t)-I(m-v,n-u;t-1) \tag{8}$$

under some regularization constraints. Finally the following quadratic form must be minimized:

$$J(u,v)=e^2(m,n)+\lambda[(u-u^0)^2+(v-v^0)^2], \tag{9}$$

where λ is a regularization constant measuring the confidence on the a priori estimation in connection with the estimation error. In reality this criterion is not directly quadratic on the unknown parameters u and v. A first order development is admitted in order to obtain the linearization of $e(m,n)$,

$$e(m,n)=e^0(m,n)+g_x(u-u^0)+g_y(v-v^0), \tag{10}$$

where $g_x=I_x(m-v^0,n-u^0;t-1)$ and $g_y=I_y(m-v^0,n-u^0;t-1)$ are the horizontal and the vertical gradients of the intensity. Using this approximation we obtain the following solution:

$$\begin{bmatrix}\hat{u}(m,n)\\ \hat{v}(m,n)\end{bmatrix}=\begin{bmatrix}u^0(m,n)\\ v^0(m,n)\end{bmatrix}-\frac{e^0(m,n)}{\lambda+g_x^2+g_y^2}\begin{bmatrix}g_x\\ g_y\end{bmatrix}. \tag{11}$$

In practice the reconstructed image intensity is used to determine g_x and g_y, in order to have exactly the same estimator at the decoder. For the same reason the DFD is calculated using the reconstructed intensity.

2.4. Convergence properties of the algorithm

Another way to interpret (11) is that of recursive stochastic adaptive algorithms, which have interesting performance in tracking the eventually changing displacement vector. The algorithm based on (2) and (11), without the discontinuity detection, is known as the Normalized Least Mean Square algorithm (NLMS). In [2] it is proven that the NLMS algorithm according to (2) and (11) is almost surely and mean square exponentially convergent, if $\{g_x, g_y\}$ constitutes

(a) a stationary ergodic process having positive-definite covariance, or
(b) a nonstationary φ-mixing process with mixing constants $\{\varphi_{m,n}\}$ satisfying

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sqrt{\varphi_{m,n}}<\infty$$

and

$$\lim_{M\to\infty, N\to\infty} \frac{1}{MN} \sum_{m=0}^{M} \sum_{n=0}^{N} \mathrm{E}\left\{\begin{bmatrix} I_x^2(m,n) & I_x(m,n)I_y(m,n) \\ I_x(m,n)I_y(m,n) & I_y^2(m,n) \end{bmatrix}\right\}$$

being positive-definite (E being the notation for the average). Let us remind that a φ-mixing process is one for which the distance future is weakly dependent upon the present and that the mixing constants measure this dependence.

This convergence property means that, if the real 2D motion is translatory, and the linear model according to (10) is valid, then the algorithm according to (2) and (11) converges to the real displacement vector. It also means that, if the real displacement vector is varying, the algorithm according to (2) and (11) converges to a vector which ensures the minimization of the DFD.

3. Adaptative intra–inter-frame prediction

Two methods can be used in order to adapt the predictor to the spatiotemporal characteristics of the signal. The first uses a forward segmentation where the best predictor is selected from a set of predictors. For example, an inter-frame predictor may be better for parts of the frame without motion or with slow motion, and an intra-frame predictor may be better for parts with fast motion. The second method uses a backward scheme of adaptation of the predictor without segmentation. The predictor is a weighted sum of predictors and the predictor coefficients are adapted continuously, pel by pel, to the local characteristics of the intensity. In this paper we adopt the second method. We use four directive predictors: a horizontal, a vertical, a diagonal and a temporal one. The last one is based on displacement estimation. In (1), $a(0,1)$ is the horizontal coefficient, $a(1,0)$ the vertical, $a(1,1)$ the diagonal and b the temporal. A similar scheme is used by Pirsch. The work of Pirsch [16] is limited to the case of only two predictors: an intra-frame using three fixed coefficients and an inter-frame without motion compensation.

To adapt the predictor coefficients we use the Least Mean Square (LMS) algorithm for its simplicity and good performances in nonstationary situations. The LMS adaptive filter has been used in one-dimensional predictive coding, and in a wide range of one-dimensional signal processing applications. A thorough study of adaptive FIR filters is given by Macchi and Bellanger [10]. An extension of the LMS algorithm for two-dimensional signals is discussed by Hadhoud and Thomas [8]. Alexander and Rajala [1] used the LMS algorithm in image coding and obtained a reduction of the distortion of about 6 dB compared with a fixed coefficient DPCM. We also have given results of applying the LMS algorithm in intra- and inter-frame image coding, with or without adaptive quantization [14].

The stochastic gradient algorithm minimizes the mean square prediction error

$$\mathrm{E}\{(I(m,n;t) - \hat{I}(m,n;t))^2\}. \tag{12}$$

The coefficients (here three intra-frame and one inter-frame coefficients) are updated by the following expression:

$$\begin{bmatrix} a(0,1) \\ a(1,0) \\ a(1,1) \\ b \end{bmatrix}_{(m,n)} = \begin{bmatrix} a(0,1) \\ a(1,0) \\ a(1,1) \\ b \end{bmatrix}_{(m,n-1)} + \mu \tilde{e}(m,n) \begin{bmatrix} \tilde{I}(m,n-1;t) \\ \tilde{I}(m-1,n;t) \\ \tilde{I}(m-1,n-1;t) \\ \tilde{I}(m-\hat{v}, n-\hat{u}; t-1) \end{bmatrix}, \tag{13}$$

where $\bar{e}$ is the quantized prediction error and μ is known as the adaptation step-size. The updated coefficients at pixel (m, n) are used to predict the intensity at the next pixel $(m, n+1)$. It has been proved that without quantization the inequality

$$0<\mu<\frac{2}{\nu P_I} \tag{14}$$

(ν: number of coefficients, P_I: power of the intensity signal) is necessary to ensure the convergence [10]. In fact, the adaptation step size must be big enough to quickly forget the initial conditions and to have good tracking properties and it must be small enough to obtain a low steady state mean-square error. The adaptation step size we use is about $1/(\nu P_I)$. This value of the step-size maximizes the convergence speed of the LMS algorithm [10].

We must note that the gradient term on the right-hand part of (13) uses the quantized prediction error and previously transmitted and reconstructed pels. Thus, in absence of transmission errors exactly the same operation can be realized by the decoder.

4. Adaptive quantization

In order to achieve an interesting compression rate a scalar three-level quantizer is used to quantize the prediction error. In image predictive coding two types of distortion may be visible:
- granular or roundoff errors in homogeneous zones,
- slope overload errors in edge elements.

4.1. A three-level adaptive quantizer

Also knowing that the prediction error is not a spatially stationary process, an adaptive quantizer may track the statistics of the prediction error for minimizing the quantization error (Fig. 3). The parameter $s(i, k)$ completely characterizes the quantizer, when the prediction error is supposed zero-mean. It is adapted by a backwards technique, which means that only the previously quantized values of the prediction error

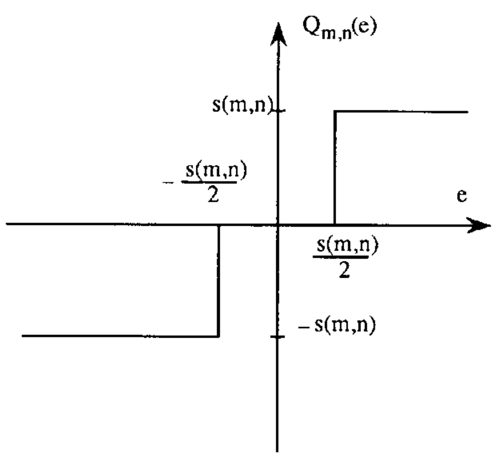

Fig. 3. A three-level adaptive quantizer.

are used. For one-dimensional DPCM schemes this technique was studied by several authors [4–7, 9]. Here the parameter $s(m, n)$ is updated by the following algorithm:

$$s(m,n)=\frac{z(m,n-1)z(m-1,n)}{z(m-1,n-1)}, \quad m\geqslant 0 \text{ and } n\geqslant 0, \tag{15}$$

$$\bar{e}(m,n)=Q_{m,n}(e(m,n)), \tag{16}$$

$$z(m,n)=F\left(\frac{\bar{e}(m,n)}{s(m,n)}\right)s(m,n), \tag{17}$$

with $z(-1,n)=z(m,-1)=z(-1,-1)$ for all $m\geqslant 0$ and $n\geqslant 0$, and $F(\cdot)$ a positive valued function which satisfies

$$F(1)=F(-1)>1, \qquad F(0)<1, \tag{18}$$

$F(1)$ is called the step-size adaptation parameter. A possible choice for these parameters is $F(0)F(1)=1$. The parameter $z(m, n)$ may be interpreted as the state variable of the quantizer.

The parameter $z(m, n)$ is bounded in order to avoid that $s(i, k)$ converges to zero or diverges:

$$0<\delta\leqslant z(m,n)\leqslant\Delta. \tag{19}$$

The adaptive quantizer with this maximum and minimum constraint is known in the 1D case as the saturating adaptive quantizer and studied in [5]. Because $z(m, n)$ is bounded, $s(m, n)$ is also bounded:

$$\frac{\delta^2}{\Delta}\leqslant s(m,n)\leqslant\frac{\Delta^2}{\delta}. \tag{20}$$

4.2. Convergence properties of the adaptive quantizer

We now consider the case of a fixed 2D predictive filter. We shall study the convergence properties of the adaptive quantizer assuming that the predictive filter is matched. For this reason according to the results of [5], we must prove that the space-averaged mean-absolute value of the reconstructed signal and of the prediction error are bounded. We give these bounds in the following. If the predictive filter is stable, then the reconstructed 2D signal is deterministically bounded. For this, consider a recursive predictive filter which satisfies

$$\bar{x}(m,n)=\sum_{(k,l)\in S}\sum a(k,l)\bar{x}(m-k,n-l)+\bar{e}(m,n). \tag{21}$$

Knowing that $\bar{e}(m, n)$ is absolutely bounded by Δ^2/δ, the result is obvious. In Appendix A we prove that, if the filter is quarter-plane (with $a(k, l)=0$, for $k>K$ and $l>L$), a bound for $\bar{x}(m, n)$ is given by

$$|\bar{x}(m,n)|\leqslant\frac{\Delta^2}{\delta}\inf_{r_1,r_2}\frac{1}{(1-r_1)(1-r_2)}\sup_{\theta_1,\theta_2}\left|1-\sum_{\substack{k=0\\|k|+|l|\neq 0}}^{K}\sum_{l=0}^{L}a(k,l)r_1^{-k}r_2^{-l}\,\mathrm{e}^{\,\mathrm{j}(k\theta_1+l\theta_2)}\right|^{-1}, \tag{22}$$

where $(r_1\,\mathrm{e}^{\mathrm{j}\theta_1}, r_2\,\mathrm{e}^{\mathrm{j}\theta_2})$ belongs to the convergence region in the z-transform domain with $0<r_1<1$, $0<r_2<1$, $0\leqslant\theta_1<2\pi$ and $0\leqslant\theta_2<2\pi$. This expression relates the bound for $\bar{x}(m, n)$ to the coefficients of the filter $\{a(k, l)\}$.

Let us now consider the prediction error under the hypothesis of a matched predictive filter, that is, we suppose

$$x(m,n)=\sum_{\substack{k=0\\|k|+|l|\neq 0}}^{K}\sum_{l=0}^{L}a(k,l)x(m-k,n-l)+w(m,n), \tag{23}$$

where $w(m,n)$ is a 2D white noise. We shall demonstrate that the space-averaged mean-absolute prediction error is bounded. It is relatively easy to show that

$$e(m,n)=\sum_{\substack{k=0\\|k|+|l|\neq 0}}^{K}\sum_{l=0}^{L}a(k,l)e(m-k,n-l)+\omega(m,n),\quad m\geqslant 0 \text{ and } n\geqslant 0, \tag{24}$$

with

$$\omega(m,n)=w(m,n)-\sum_{\substack{k=0\\|k|+|l|\neq 0}}^{K}\sum_{l=0}^{L}a(k,l)\bar{e}(m-k,n-l).$$

The solution of this 2D linear difference equation on $e(m,n)$ depends on the input $\omega(m,n)$ and on the initial conditions. Let us consider the following initial conditions for the predictor: $\hat{x}(m,n)=0$, for $m<0$ or $n<0$. Then, $e(m,n)=x(m,n)$, for $m<0$ or $n<0$. Let us write

$$e(m,n)=e^{(s)}(m,n)+e^{(0)}(m,n) \tag{25}$$

where the part $e^{(s)}(m,n)$ depends on the input $\omega(m,n)$ with zero initial conditions, and $e^{(0)}(m,n)$ is the solution of the homogeneous 2D linear difference equation resulting from (24), with $\omega(m,n)=0$, depending on the initial conditions only. Concerning the first term $e^{(s)}(m,n)$, we can write

$$e^{(s)}(m,n)=\sum_{\substack{k=0\\|k|+|l|\neq 0}}^{K}\sum_{l=0}^{L}a(k,l)e^{(s)}(m-k,n-l)+w(m,n)-\sum_{\substack{k=0\\|k|+|l|\neq 0}}^{K}\sum_{l=0}^{L}a(k,l)\bar{e}(m-k,n-l). \tag{26}$$

Let us consider the impulse response $h(k,l)$ of the above recursive filter. We can write

$$e^{(s)}(m,n)=\sum_{k=0}^{m}\sum_{l=0}^{n}h(k,l)\left(w(m-k,n-l)-\sum_{\substack{k_1=0\\|k_1|+|l_1|\neq 0}}^{K}\sum_{l_1=0}^{L}a(k_1,l_1)\bar{e}(m-k-k_1,n-l-l_1)\right). \tag{27}$$

Knowing that the quantized prediction error is bounded by Δ^2/δ, we obtain the following bound:

$$\mathrm{E}\{|e^{(s)}(m,n)|\}\leqslant\left(\mathrm{E}\{|w(m,n)|\}+\frac{\Delta^2}{\delta}\sum_{\substack{k=0\\|k|+|l|\neq 0}}^{K}\sum_{l=0}^{L}|a(k,l)|\right)\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}|h(k,l)|. \tag{28}$$

The second term $e^{(0)}(m,n)$ due to the initial conditions is studied in Appendix B, where it is proved that its space-averaged mean absolute value converges to zero, i.e.

$$\lim_{M\to\infty,N\to\infty}\frac{1}{MN}\mathrm{E}\left\{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|e^{(0)}(m,n)|\right\}=0. \tag{29}$$

Using (25), (28) and (29), we finally find that

$$\lim_{M\to\infty, N\to\infty} \sup \frac{1}{MN} \mathrm{E}\left\{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} |e(m,n)|\right\}$$

$$\leqslant \left(c_w + (|a(0,1)| + |a(1,0)| + |a(1,1)|)\frac{\Delta^2}{\delta}\right)\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} |h(k,l)|, \tag{30}$$

where $c_w = \mathrm{E}\{|w(m,n)|\}$.

The above results, according to the study of [5], permit to establish the existence of a unique stationary joint distribution of the input and decoded 2D processes.

5. Decoding in the presence of transmission errors

The main problem with predictive coding, and particularly with adaptive methods, was found to be its instability in the presence of transmission errors. It is important to ensure the robustness of the decoder without rate augmentation. The criterion of adaptation of the predictor must then be modified. We study this problem in the case of intraframe coding without motion compensation ($b=0$).

The algorithm presented in Section 3 may be divergent in the presence of transmission errors. A modification of the initial algorithm, constraining the output of the predictor, leads to the stabilization of the filter in the decoder. The criterion minimized for the constrained LMS algorithm is

$$J_s = \mathrm{E}\{e^2(m,n)\} + \eta\mathrm{E}\{(\hat{I}(m,n) - \hat{I}_f(m,n))^2\}, \tag{31}$$

where $\hat{I}_f(m,n)$ is the output of a fixed coefficient filter, whose inverse is a stable filter, and η is a positive constant which may be interpreted as a regularization factor.

For stationary signals, the minimization of the above criterion gives the following solution:

$$\boldsymbol{H}_s = \frac{1}{1+\eta}\boldsymbol{H} + \frac{\eta}{1+\eta}\boldsymbol{H}_f, \tag{32}$$

where $\boldsymbol{H}$ (vector containing the filter coefficients) is the optimal solution resulting from the LMS algorithm. Thus $\boldsymbol{H}_s$ is the barycenter of $(\boldsymbol{H}, 1)$ and $(\boldsymbol{H}_f, \eta)$. Therefore $\boldsymbol{H}_f$ may be interpreted as an a priori conservative estimation of the predictor coefficients with confidence degree η. A similar approach is introduced in [12] for another type of adaptive predictor. The updated equation of the constraind stochastic gradient algorithm is

$$\begin{bmatrix} a(0,1) \\ a(1,0) \\ a(1,1) \end{bmatrix}_{(m,n)} = \begin{bmatrix} a(0,1) \\ a(1,0) \\ a(1,1) \end{bmatrix}_{(m,n-1)} + \mu(\bar{e}(m,n) - \eta(\hat{I}(m,n) - \hat{I}_f(m,n)))\begin{bmatrix} \bar{I}(m,n-1;t) \\ \bar{I}(m-1,n;t) \\ \bar{I}(m-1,n-1;t) \end{bmatrix}. \tag{33}$$

A study of this algorithm is given in [13], where its performance is evaluated, and in [15], where its robustness is proved for some specific signals.

If an adaptive quantizer is used, it is also vulnerable to transmission errors. It must be modified, to avoid or to dissipate the effects of transmission errors. In [7] a technique is introduced for this purpose in the case of a 1D adaptive quantizer. We use the same technique here by modifying (15) to diminish the effect of a transmission error:

$$s(m, n) = \frac{z(m, n-1)^{\gamma} z(m-1, n)^{\gamma}}{z(m-1, n-1)^{\gamma^2}}, \tag{34}$$

where $0 < \gamma < 1$. In fact, the above equation gives a linear recursive relation on $\ln \sigma(m, n)$:

$$\ln z(m, n) = \gamma \ln z(m-1, n) + \gamma \ln z(m, n-1) - \gamma^2 \ln z(m-1, n-1) + \ln F\left(\frac{\bar{e}(m, n)}{s(m, n)}\right). \tag{35}$$

The 2D linear recursive filter on $\ln z(m, n)$ is separable and obviously stable for $0 < \gamma < 1$. Thus effects of errors on $F(\bar{e}(m, n)/s(m, n))$ will be dissipated.

The theoretical results of [13] and [15], and simulation results of the following section demonstrate that the decoder is readjusted if the transmission errors do not affect the pixel synchronization. This supposes that fixed length codes are used and only bit inversions are considered. The case of variable length codes is more difficult; it is not considered here.

6. Simulations and results

The coder described in this paper has been simulated for the 'CAR' sequence of images provided by CCETT (COST 211 bis European normalization). Figure 4 gives the first image of the sequence. The car is in movement and the camera pans the scene. A strong additional noise disturbs the intensity and different types of spatial details are present in the scene. To appreciate the importance of the motion, the difference between the first two frames is given in Fig. 5. The empirical standard deviation of the difference is 30.1.

The quality criterion is given by the mean square distortion

$$D = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{k=0}^{N-1} (I(m, n) - \bar{I}(m, n))^2, \tag{36}$$

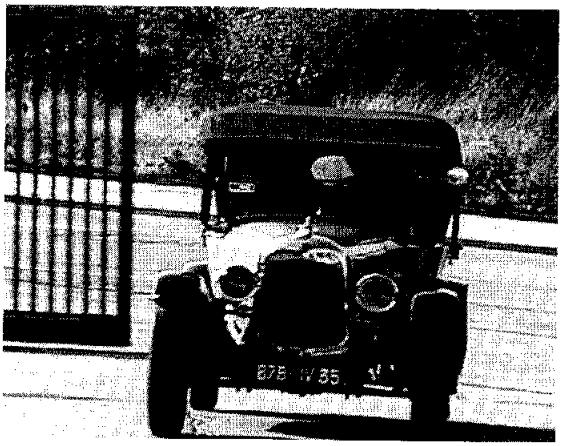

Fig. 4. The first frame of the sequence 'CAR' (CCETT).

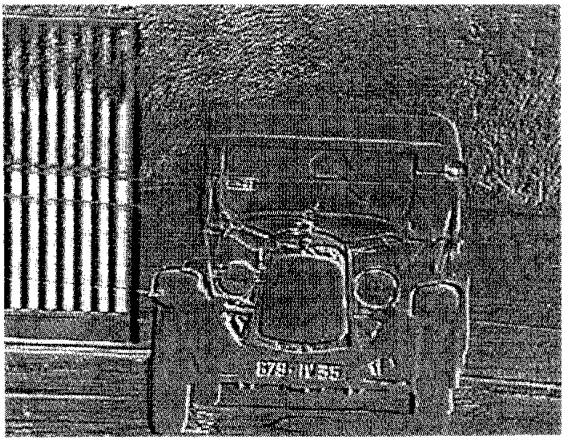

Fig. 5. The difference between the first two frames.

or equivalently by $\sqrt{D}$. The power of the prediction error is given in order to appreciate the performance of the predictor. Finally, the entropy of the quantized prediction error is given, this quantity being closely linked to the compression rate.

Table 1 gives numerical results for 60 fields of the sequence using the adaptive quantizer of Section 4. We distinguish two cases of predictors using a motion estimator: in case A the coder transmits the addresses of points of discontinuity and in case B it does not transmit them. The detection of discontinuities is described in Section 2, and the rate of discontinuities detected for the frame of Table 1 was about 1%. P_e is the power of the prediction error using the adaptive quantizer and H is the entropy of the 3-level quantized prediction error (H also includes the discontinuity side information in the case of coders A). SNR is the signal-to-noise ratio given in dB and defined by

$$\mathrm{SNR} = 10 \log_{10} \frac{255^2}{D}. \tag{37}$$

In our simulations each field is composed of 264 lines and 674 points per line. For the obtained value of the entropy and for fields of 288 by 720 points, the bit-rate is about 15 Mbits/sec for the luminance component. The pictures given in the paper are frames composed of 512 by 512 points. Figure 6 gives the distortion for the 60 fields of the sequence using the hybrid predictor of case A, except the first two fields which are coded using an adaptive intra-field predictor.

Table 1

Numerical results for the 'CAR' sequence

	Hybrid prediction *A*	Hybrid prediction *B*	Motion compensation *A*	Motion compensation *B*	Adaptive intraframe
$\sqrt{P_e}$	13.3	13.9	16.0	16.2	16.4
$\sqrt{D}$	7.5	7.8	8.6	8.7	9.0
SNR	30.7	30.3	29.5	29.3	29.1
H	1.33	1.33	1.36	1.36	1.41

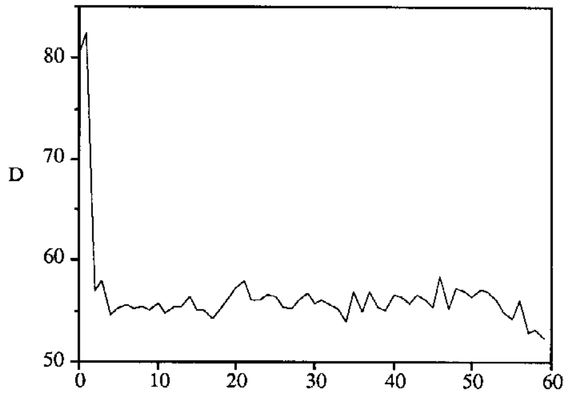

Fig. 6. Distortion of the 'CAR' sequence using the hybrid coder A.

Figures 7, 8 and 9 give the prediction error before quantization for the hybrid adaptive, the motion compensated and the adaptive intraframe predictor, respectively. The prediction error is dilated by a factor 2 in these three figures.

If the rate of transmission errors is about 10^{-3}, the square root of the distortion is 9.1 at the coder, using the stabilization method of Section 5, and 9.3 at the decoder for frame 'CAR_00'. This distortion is practically visually imperceptible. In Fig. 10 reconstructed at the decoder frame 'CAR 00' is given, when the pure LMS algorithm is used for adapting an intra-frame predictor and when only one transmission error occurs. In Fig. 11 is given the reconstructed at the decoder frame 'CAR_00', when the modified LMS algorithm is used for adapting an intra-frame predictor and when the rate of transmission errors is about 10^{-3}. We give in Fig. 12 frame 'CAR 00' as it is reconstructed at the decoder, when a block of 256 transmission errors occurred at the 255th line of the frame. The square root of the distortion is 9.5. All these results illustrate the small sensitivity of the proposed adaptive method to transmission errors.

7. Conclusion

An adaptive intra-inter-frame predictive coding method is presented. The simulation of the proposed method for a very critical image sequence illustrated a certain improvement in comparison with

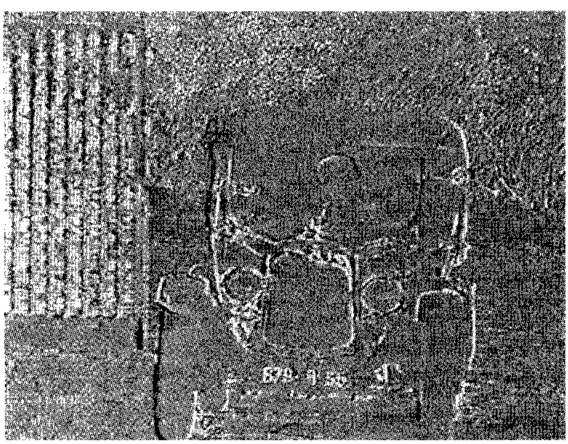

Fig. 7. The prediction error before quantization for the hybrid adaptive predictor.

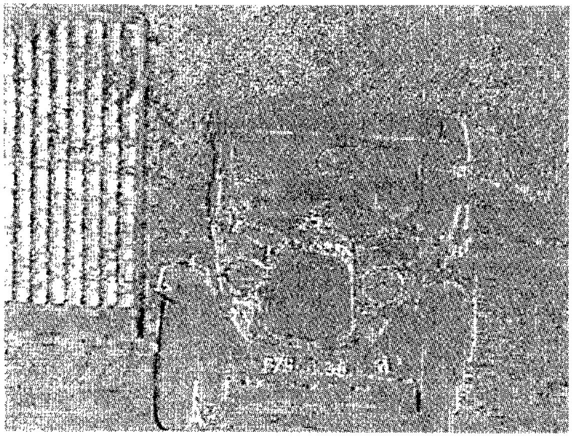

Fig. 8. The prediction error before quantization for the motion-compensated predictor.

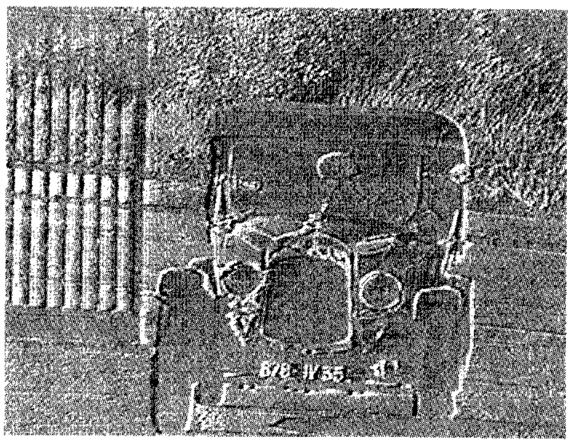

Fig. 9. The prediction error before quantization for the adaptive intra-frame predictor.

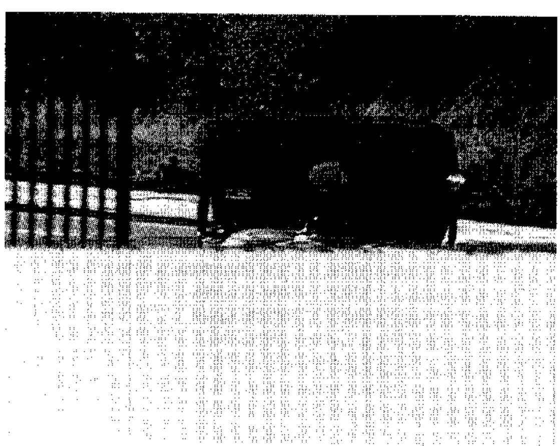

Fig. 10. The decoded frame 'CAR_00' with the LMS intra-frame predictor after one transmission error.

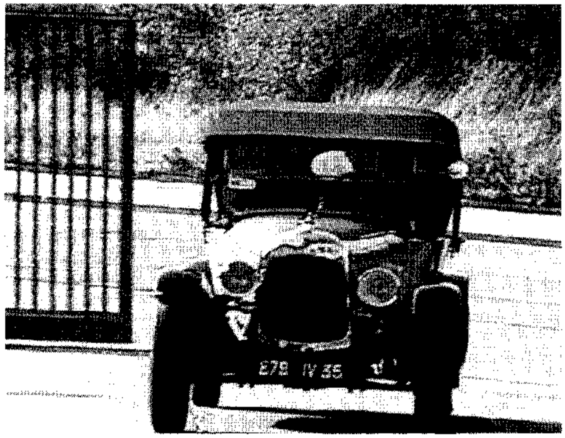

Fig. 11. The decoded frame 'CAR_00' with the modified LMS intra-frame predictor with a rate of transmission errors of 10^{-3}.

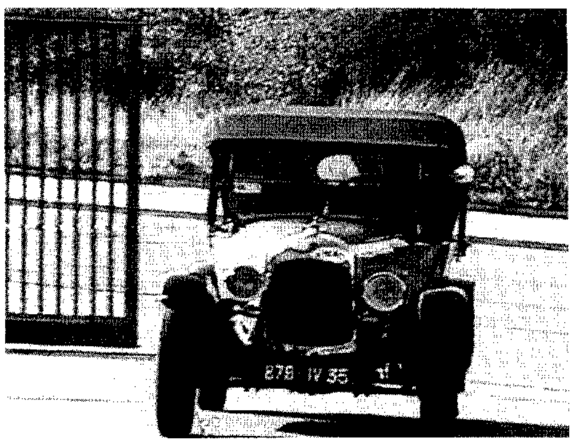

Fig. 12. The decoded frame 'CAR_00' with the modified LMS intra-frame predictor after a block of 256 transmission errors.

non-adaptive methods or only motion compensation techniques. Only a small complexity increase from motion compensation pel-recursive methods is needed. The crucial problem of transmission errors is considered in the case of adaptive intraframe prediction. Only some algorithmic complexity is added in order to make the decoder in practice non-sensitive to channel errors. Knowing that pel predictive methods produce good results at higher rates, which are improved using adaptive techniques, we think that replacing the pel prediction by block prediction, and scalar quantization by vector quantization, good results could be obtained at lower bit-rates.

Appendix A. Bounding the output of a 2D quarter-plane linear filter

Firstly we prove the following proposition.

PROPOSITION. Let h(m, n) be the impulsional response of a first quarter-plane 2D linear filter and H(z, w) its transformation in the (z, w) domain. If the filter is stable, then

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}|h(m,n)| \leq \inf_{r_1,r_2}\frac{1}{(1-r_1)(1-r_2)}\sup_{\theta_1,\theta_2}|H(r_1\,\mathrm{e}^{\mathrm{j}\theta_1}, r_2\,\mathrm{e}^{\mathrm{j}\theta_2})|, \tag{A.1}$$

where $(r_1\,\mathrm{e}^{\mathrm{j}\theta_1}, r_2\,\mathrm{e}^{\mathrm{j}\theta_2})$ belongs to the region of convergence of $H(z, w)$, with $0<r_1<1$ and $0<r_2<1$.

PROOF. The impulsional response is related to its transformation in the (z, w) domain by

$$h(m,n) = -\frac{1}{(2\pi)^2}\oint_{C_1}\oint_{C_2} H(z,w)z^{m-1}w^{n-1}\,\mathrm{d}z\,\mathrm{d}w, \tag{A.2}$$

where C_1 and C_2 are closed curves belonging to the convergence region. Let us put $z=r_1\,\mathrm{e}^{\mathrm{j}\theta_1}$ and $w=r_2\,\mathrm{e}^{\mathrm{j}\theta_2}$ with $0<r_1<1$, $0<r_2<1$, $0\leq\theta_1<2\pi$ and $0\leq\theta_2<2\pi$. Then we can write

$$h(m,n) = \frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} H(r_1\,\mathrm{e}^{\mathrm{j}\theta_1}, r_2\,\mathrm{e}^{\mathrm{j}\theta_2})r_1^m r_2^n\,\mathrm{e}^{\mathrm{j}(m\theta_1+n\theta_2)}\,\mathrm{d}\theta_1\,\mathrm{d}\theta_2. \tag{A.3}$$

From this expression we obtain the following inequality:

$$|h(m,n)| \leq \frac{r_1^m r_2^n}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi} |H(r_1\,\mathrm{e}^{\mathrm{j}\theta_1}, r_2\,\mathrm{e}^{\mathrm{j}\theta_2})|\,\mathrm{d}\theta_1\,\mathrm{d}\theta_2, \tag{A.4}$$

and, also,

$$|h(m,n)| \leq r_1^m r_2^n \sup_{\theta_1,\theta_2}|H(r_1\,\mathrm{e}^{\mathrm{j}\theta_1}, r_2\,\mathrm{e}^{\mathrm{j}\theta_2})|. \tag{A.5}$$

The summation over the quarter-plane gives

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}|h(m,n)| \leq \frac{1}{(2\pi)^2(1-r_1)(1-r_2)}\int_0^{2\pi}\int_0^{2\pi} |H(r_1\,\mathrm{e}^{\mathrm{j}\theta_1}, r_2\,\mathrm{e}^{\mathrm{j}\theta_2})|\,\mathrm{d}\theta_1\,\mathrm{d}\theta_2. \tag{A.6}$$

Then we have

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}|h(m,n)| \leq \frac{1}{(1-r_1)(1-r_2)}\sup_{\theta_1,\theta_2}|H(r_1\,\mathrm{e}^{\mathrm{j}\theta_1}, r_2\,\mathrm{e}^{\mathrm{j}\theta_2})|. \tag{A.7}$$

This is true for any (r_1, r_2) satisfying the above conditions, and thus the proposition is demonstrated. □

If the filter is recursive quarter-plane, for which

$$H(z,w)=\left(1-\sum_{\substack{k=0\\|k|+|l|\neq 0}}^{K}\sum_{l=0}^{L}a(k,l)z^{-k}w^{-l}\right)^{-1}, \tag{A.8}$$

then we can write

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}|h(m,n)|\leqslant \inf_{r_1,r_2}\frac{1}{(1-r_1)(1-r_2)}\sup_{\theta_1,\theta_2}\left|1-\sum_{\substack{k=0\\|k|+|l|\neq 0}}^{K}\sum_{l=0}^{L}a(k,l)r_1^{-k}r_2^{-l}\,\mathrm{e}^{-\mathrm{j}(k\theta_1+l\theta_2)}\right|^{-1}. \tag{A.9}$$

Appendix B. Study of the prediction error due to the initial conditions

An upper bound of $\mathrm{E}\{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|e^{(0)}(m,n)|\}$ is derived which allows to prove that the space-averaged mean-absolute value of $e^{(0)}(m,n)$ vanishes to zero. In order to simplify the demonstration and the notation, we consider the particular case where $L=K=1$, in (24). Then, we have

$$\begin{aligned}&e^{(0)}(m,n)=a(1,0)e^{(0)}(m-1,n)+a(0,1)e^{(0)}(m,n-1)+a(1,1)e^{(0)}(m-1,n-1), && \text{if } m\geqslant 0 \text{ and } n\geqslant 0,\\ &e^{(0)}(m,n)=e(m,n), && \text{if } m<0 \text{ or } n<0.\end{aligned} \tag{B.1}$$

Let us define

$$E^{(0)}(z,w)=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}e^{(0)}(m,n)z^{-m}w^{-n}. \tag{B.2}$$

It is easy to prove that the previous (z,w) transform satisfies

$$\begin{aligned}&[1-a(1,0)z^{-1}-a(0,1)w^{-1}-a(1,1)z^{-1}w^{-1}]E^{(0)}(z,w)\\&\quad=a(1,0)e(-1,0)+a(0,1)e(0,-1)+a(1,1)e(-1,-1)\\&\qquad+\sum_{n=1}^{\infty}[a(1,0)e(-1,n)+a(1,1)e(-1,n-1)]w^{-n}\\&\qquad+\sum_{m=1}^{\infty}[a(0,1)e(m,-1)+a(1,1)e(m-1,-1)]z^{-m}.\end{aligned} \tag{B.3}$$

The above relation is also equivalent to

$$\begin{aligned}e^{(0)}(m,n)=\;&a(1,0)h(m,0)e(-1,n)+a(0,1)h(0,n)e(m,-1)+a(1,1)h(m,n)e(-1,-1)\\&+\sum_{k=1}^{n}[a(1,0)h(m,k)+a(1,1)h(m,k-1)]e(-1,n-k)\\&+\sum_{k=1}^{m}[a(0,1)h(k,n)+a(1,1)h(k-1,n)]e(m-k,-1).\end{aligned} \tag{B.4}$$

Taking the absolute value of $e^{(0)}(m,n)$ and using inequality (A.5) of Appendix A, we obtain

$$|e^{(0)}(m,n)| \leqslant c_H \Bigg\{ \alpha^m \beta^n |a(1,1)e(-1,-1)| + \alpha^m \left[|a(1,0)e(-1,n)| + (|a(1,0)|\beta + |a(1,1)|) \sum_{k=1}^{n} \beta^{k-1} |e(-1,n-k)| \right] + \beta^n \left[|a(0,1)e(m,-1)| + (|a(0,1)|\alpha + |a(1,1)|) \sum_{k=1}^{m} \alpha^{k-1} |e(m-k,-1)| \right] \Bigg\}, \tag{B.5}$$

where $c_H = \sup_{\theta_1,\theta_2} |H(r_1 e^{j\theta_1}, r_2 e^{j\theta_2})|$, $\alpha = r_1$ and $\beta = r_2$ $(0<\alpha<1, 0<\beta<1)$. Then, one can write

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |e^{(0)}(m,n)| \leqslant c_H \Bigg\{ \frac{(1-\alpha^M)(1-\beta^N)}{(1-\alpha)(1-\beta)} |a(1,1)e(-1,-1)| + \frac{1-\alpha^M}{1-\alpha} \times \left[|a(1,0)| \sum_{k=0}^{N-1} |e(-1,k)| + \frac{1}{1-\beta} (|a(1,0)|\beta + |a(1,1)|) \sum_{k=0}^{N-2} (1-\beta^{N-k-1}) |e(-1,k)| \right] + \frac{1-\beta^N}{1-\beta} \left[|a(0,1)| \sum_{k=0}^{M-1} |e(k,-1)| + \frac{1}{1-\alpha} (|a(0,1)|\alpha + |a(1,1)|) \times \sum_{k=0}^{M-2} (1-\alpha^{M-k-1}) |e(k,-1)| \right] \Bigg\}. \tag{B.6}$$

Let us consider the following initial conditions for the predictor: $\hat{x}(-1,k) = \hat{x}(k,-1) = 0$, for all $k \geqslant -1$. Then, $e(-1,k) = x(-1,k)$ and $e(k,-1) = x(k,-1)$, for all $k \geqslant -1$. The above equation leads to

$$\mathrm{E}\left\{ \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |e^{(0)}(m,n)| \right\} \leqslant c_x c_H \Bigg\{ \frac{(1-\alpha^M)(1-\beta^N)}{(1-\alpha)(1-\beta)} |a(1,1)| + \frac{1-\alpha^M}{1-\alpha} \left[|a(1,0)|N + \frac{|a(1,0)|\beta + |a(1,1)|}{1-\beta} \left(N-1-\beta \frac{1-\beta^{N-1}}{1-\beta} \right) \right] + \frac{1-\beta^N}{1-\beta} \left[|a(0,1)|M + \frac{|a(0,1)|\alpha + |a(1,1)|}{1-\alpha} \left(M-1-\alpha \frac{1-\alpha^{M-1}}{1-\alpha} \right) \right] \Bigg\}, \tag{B.7}$$

where $c_x = \mathrm{E}\{|x(m,n)|\}$. Finally, we find

$$\mathrm{E}\left\{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|e^{(0)}(m,n)|\right\}$$

$$\leqslant \frac{c_x c_H}{(1-\alpha)(1-\beta)}\left\{(1-\alpha^M)(|a(1,0)|+|a(1,1)|)N + (1-\beta^N)(|a(0,1)|+|a(1,1)|)M - \frac{(1-\alpha^M)(1-\beta^N)}{(1-\alpha)(1-\beta)} \times [\beta(1-\alpha)|a(1,0)| + \alpha(1-\beta)|a(1,0)| + (1-\alpha\beta)|a(1,1)|]\right\}. \quad \text{(B.8)}$$

From this inequality, it is straightforward to show that

$$\lim_{M\to\infty, N\to\infty}\frac{1}{MN}\mathrm{E}\left\{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|e^{(0)}(m,n)|\right\} = 0. \quad \text{(B.9)}$$

Acknowledgment

The authors wish to thank Mrs. O. Macchi for many helpful discussions.

References

[1] S. Alexander and S. Rajala, "Image compression results using the LMS adaptive algorithm", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-33, No. 3, June 1985, pp. 712–714.
[2] R. Bitmead and B. Anderson, "Performance of adaptive estimation algorithms in dependent random environments", *IEEE Trans. Automat. Control*, Vol. AC-25, No. 4, August 1980, pp. 788–794.
[3] C. Cafforio and F. Rocca, "The differential method for motion estimation", in: T.S. Huang, ed., *Image Sequence Processing and Dynamic Scene Analysis*, Springer, New York, 1983, pp. 104–124.
[4] N. Gerr and S. Cambanis, "Analysis of Adaptive Differential PCM of a stationary Gauss-Markov input", *IEEE Trans. Inform. Theory*, Vol. IT-33, No. 2, May 1987, pp. 350–359.
[5] L. Goldstein and B. Liu, "Deterministic and stochastic stability of Adaptive Differential Pulse Code Modulation", *IEEE Trans. Inform. Theory*, Vol. IT-23, No. 4, July 1977, pp. 445–453.
[6] D. Goodman and A. Gersho, "Theory of an adaptive Quantizer", *IEEE Trans. Comm.*, Vol. COM-22, No. 8, August 1974, pp. 1037–1045.
[7] D. Goodman and R. Wilkinson, "A robust adaptive Quantizer", *IEEE Trans. Comm.*, Vol. COM-23, No. 11, November 1975, pp. 1362–1365.
[8] M. Hadhoud and D. Thomas, "The two-dimensional adaptive LMS (TDLMS) algorithm", *IEEE Trans. Circuits and Systems*, Vol. CAS-35, No. 5, May 1988, pp. 485–494.
[9] J. Kieffer, "Stochastic stability for feedback quantization schemes", *IEEE Trans. Inform. Theory*, Vol. IT-28, No. 2, March 1982, pp. 248–254.
[10] O. Macchi and M. Bellanger, "Le filtrage adaptatif transverse", *Traitement du Signal*, Vol. 5, No. 3, 1988, pp. 115–132.
[11] P. Maragos, P. Schaffer and R. Mersereau, "Two dimensional linear prediction and its application to adaptive predictive coding of images", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-32, No. 6, December 1984, pp. 1213–1229.
[12] N. Maxemschuk and J. Stuller, "Reduction of transmission error propagation in adaptively predicted DPCM encoded pictures", *Bell Syst. Tech. J.*, Vol. 58, No. 6, July–August 1979, pp. 1413–1423.
[13] J.-C. Pesquet, O. Macchi and G. Tziritas, "Second order statistical analysis of two constrained LMS algorithms", *EUSIPCO'90*.

[14] J.-C. Pesquet and G. Tziritas, "Méthodes adaptatives appliquées au codage prédictif intra et inter-images", *Traitement du Signal*, Vol. 5, No. 3, 1988, pp. 143–152.
[15] J.-C. Pesquet, G. Tziritas and O. Macchi, "Modified LMS algorithms for robust ADPCM", *Internat. Conf. Acoust. Speech Signal Process. '90*, pp. 1405–1408.
[16] P. Pirsch, "Adaptive intra interframe DPCM coder", *Bell Syst. Tech. J.*, Vol. 61, No. 5, May-June 1982, pp. 747 764.
[17] D.R. Walker and K.R. Rao, "Improved pel-recursive motion compensation", *IEEE Trans. Comm.*, Vol. COM-32, No. 10, October 1984, pp. 1128–1134.