

# Méthodes adaptatives

## appliquées au codage prédictif

### intra et inter-images

Adaptive methods applied to intra and interframe predictive coding



#### Jean-Christophe PESQUET

Laboratoire des Signaux et Systèmes (CNRS, UM 0014), École Supérieure d'Électricité, Plateau du Moulon, 91192 GIF-SUR-YVETTE CEDEX.

Ingénieur SUPELEC (promotion 87). Il a obtenu, en septembre 1987, une bourse BDI, pour effectuer une thèse de Doctorat au Laboratoire des Signaux et Systèmes. Ses principaux centres d'intérêt sont les communications numériques, le filtrage adaptatif et le traitement d'images.



#### Georges TZIRITAS

Laboratoire des Signaux et Systèmes (CNRS, UM 0014), École Supérieure d'Électricité, Plateau du Moulon, 91192 GIF-SUR-YVETTE CEDEX.

Ingénieur Institut Polytechnique d'Athènes 1977. DEA Traitement du Signal ENSIEG 1978. Thèse Docteur-Ingénieur INPG 1981 "Communications dans un canal aléatoire dispersif". Thèse Docteur ès Sciences 1985 "Contribution à l'étude des tests d'hypothèses concernant l'opérateur covariance d'un processus gaussien". Chercheur CNRS depuis 1982; activités de recherche en analyse du mouvement dans une séquence d'images et en codage adaptatif d'images.

### RÉSUMÉ

Cet article traite de l'introduction d'algorithmes adaptatifs, dans le codage MIC-Différentiel des images.

Nous généralisons des méthodes unidimensionnelles de filtrage adaptatif transverse et proposons un procédé pour adapter un quantificateur à la dynamique d'un signal bidimensionnel.

Des simulations démontrent ensuite l'intérêt de ces techniques dans le cas de codages intra et inter-images. Ainsi, par rapport à un codeur fixe, on obtient un gain de 40%, sur la distorsion quadratique.

Enfin, nous présentons des modifications des algorithmes qui permettent de les rendre très robustes aux erreurs de transmission.

### MOTS CLÉS

Codage intra et inter-images, prédiction adaptative, quantification adaptative.

### SUMMARY

*This paper deals with the use of adaptive algorithms in the DPCM coding of images.*

*We generalize one-dimensional methods of transversal adaptive filtering and propose a means of adapting a quantizer to the dynamic range of a two-dimensional signal.*

*Simulations prove the interest of these techniques for the intra and interframe cases. So, in comparison with a fixed coder, a 40% improvements is obtained for the quadratic distortion.*

*Finally, we present some modifications to the algorithms which make them robust to transmission errors.*

### KEY WORDS

*Intra and interframe coding, adaptive prediction, adaptive quantization.*

## 1. Introduction

Dans un grand nombre d'applications, l'utilisation d'images numériques nécessite la transmission (et le stockage) de grandes quantités de données. Le but du codage est de résoudre ce problème en opérant une compression de l'information qui exploite la redondance présente dans ces images.

Parmi les diverses méthodes existantes, une solution possible est le codage prédictif où l'on transmet la différence  $e$  entre le signal à coder  $x$  et une estimation  $\hat{x}$  de ce signal. Un moyen de réalisation classique est le codage MIC-D dont le principe est rappelé sur la figure 1.

Les autres signaux du codeur, représentés sur ce schéma, sont l'erreur de prédiction quantifiée  $\bar{e}$  et l'image locale du signal décodé  $\tilde{x}$ . Pour être transmis, les niveaux pris par  $\bar{e}$  sont transformés en codes par l'opérateur  $C$ . Après altération par d'éventuelles erreurs de transmission et après une opération  $D$ , inverse de  $C$ , les signaux obtenus au décodeur sont  $\bar{e}'$ ,  $\hat{x}'$  et  $\tilde{x}'$ . Pour de plus amples informations, on trouvera, dans [1], une présentation détaillée des propriétés de cette structure.

Deux remarques méritent cependant d'être faites. Tout d'abord, la prédiction étant effectuée à partir des échantillons « précédents », on voit que l'on introduit une relation d'ordre entre les pixels, ce qui correspond assez bien aux procédés technologiques de balayage des images.

En second lieu, si l'on raisonne en terme de théorie de l'information, la détermination des caractéristiques du codeur se ramène à un problème de minimisation de l'entropie du signal transmis, sous la contrainte d'une distorsion inférieure à un certain seuil de tolérance.

Dans le cas d'un prédicteur fixe, on utilise généralement une modélisation autorégressive (AR) du signal bidimensionnel [2], conduisant à une prédiction transverse. Cette approche se traduit par une faible complexité dans la mise en œuvre mais n'est pas très satisfaisante pour les raisons suivantes. D'une part, elle nécessite souvent l'introduction de connaissances *a priori* sur les propriétés statistiques du signal d'image. D'autre part, elle ne tient pas compte de la variété des classes d'images existantes et des non-stationnarités au sein d'une même image.

Pour éviter ces inconvénients, différents schémas adaptatifs ont été proposés. Il est notamment envisageable d'opérer, en chaque point, un choix entre plusieurs prédicteurs fixes, suivant le contenu de l'image [3]. Une autre possibilité consiste à découper les données en plusieurs blocs pour lesquels on résout les équations de Yule-Walker, afin de déterminer les coefficients de prédiction optimaux. On est alors amené à transmettre, comme informations supplémentaires, les valeurs de ces coefficients [4].

L'approche adaptative que nous avons adoptée diffère des deux précédentes. En effet, les coefficients de prédiction sont choisis dans un ensemble continu de valeurs. De plus, l'adaptation est réalisée en chaque point, sans qu'il soit nécessaire de transmettre les valeurs de ces coefficients. Ceci simplifie la procédure de transmission.

## 2. Méthodes adaptatives

### 2. 1. ADAPTATION DU PRÉDICTEUR

Étudions, tout d'abord, en l'absence de quantificateur, l'adaptation du prédicteur 2D transverse que nous avons utilisé.

Si l'on désigne par  $m$  l'indice de ligne (entre 0 et  $M-1$ ) et par  $n$  l'indice de colonne (entre 0 et  $N-1$ ), le signal prédit s'écrit :

$$(1a) \quad \hat{x}(m, n) = \sum_{(i, j) \in S} h_{i, j}(m, n) x(m-i, n-j)$$

où  $S$ , le support du filtre utilisé, est du type demi-plan asymétrique. Cette équation peut aussi s'exprimer sous la forme d'un produit scalaire :

$$(1b) \quad \hat{x}(m, n) = \mathbf{H}(m, n)^T \mathbf{X}(m, n)$$

où  $\mathbf{H}(m, n)$  est le vecteur des  $h_{i, j}(m, n)$  et  $\mathbf{X}(m, n)$  celui des  $x(m-i, n-j)$ . Par exemple, si  $S$  est un domaine quart de plan et le prédicteur d'ordre 3, on peut prendre :

$$\mathbf{H}(m, n) = [h_{0,1}(m, n), h_{1,0}(m, n), h_{1,1}(m, n)]^T$$

$$\mathbf{X}(m, n) = [x(m, n-1), x(m-1, n), x(m-1, n-1)]^T$$

L'adaptation de  $\mathbf{H}(m, n)$  est réalisée à partir de l'erreur de prédiction  $e(m, n)$ , en cherchant à minimiser

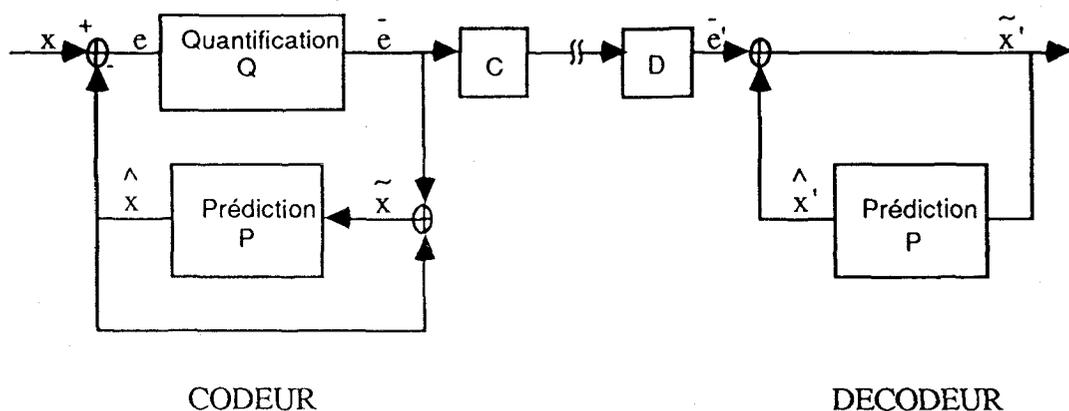


Fig. 1.

sa valeur quadratique moyenne,  $E\{e^2(m, n)\}$ , par un algorithme de gradient stochastique [5]. Des méthodes plus sophistiquées de moindres carrés ont également été essayées mais n'apportent véritablement un gain en performance que lorsque l'ordre du prédicteur est élevé [6]. Nous transposons ici les équations obtenues dans le cas unidimensionnel :

$$(2) \left\{ \begin{array}{l} \text{1D} \\ \hat{x}(n) = \mathbf{H}(n-1)^T \mathbf{X}(n) \\ \mathbf{H}(n) = \mathbf{H}(n-1) + \mu e(n) \mathbf{X}(n) \\ \text{extension 2D} \\ \hat{x}(m, n) = \hat{\mathbf{H}}(m, n)^T \mathbf{X}(m, n) \\ \mathbf{H}(m, n) = \hat{\mathbf{H}}(m, n) + \mu e(m, n) \mathbf{X}(m, n) \end{array} \right.$$

$\hat{\mathbf{H}}(m, n)$  est une estimation des coefficients de prédiction qui est obtenue à partir des vecteurs  $\mathbf{H}(m, n)$  déjà calculés. On peut comme dans [7], choisir :

$$(3) \hat{\mathbf{H}}(m, n) = \begin{cases} \mathbf{H}_0 & \text{pour } m=n=0 \\ \mathbf{H}(m, n-1) & \text{pour } n \neq 0 \\ \mathbf{H}(m-1, N-1) & \text{pour } m \neq 0 \text{ et } n=0 \end{cases}$$

Ceci signifie que l'on utilise le point précédent dans l'ordre de balayage pour estimer les coefficients de prédiction. Il semblerait plus logique de prendre  $\hat{\mathbf{H}}(m, 0) = \mathbf{H}(m-1, 0)$  mais il nous est apparu que ce choix semble poser des problèmes de vitesse de convergence.

La principale difficulté de l'algorithme ainsi décrit est la détermination du pas d'adaptation  $\mu$ . Ce dernier doit être compris entre 0 et  $2/LP_x$  ( $L$  : ordre du filtre,  $P_x = E\{x^2\}$ ), pour assurer la convergence [5]. Le choix précis de  $\mu$  fait alors l'objet d'un compromis : il doit être choisi assez grand pour permettre un oubli rapide des conditions d'initialisation et assurer une bonne capacité de poursuite et il doit être assez petit pour obtenir un bruit de fluctuation faible.

En présence d'un quantificateur,  $\mathbf{X}(m, n)$  et  $e(m, n)$  ne sont plus disponibles au décodeur (fig. 1). Or, pour éviter une divergence entre les signaux codés et ceux décodés, la prédiction est réalisée de la même façon au codeur et au décodeur. Les équations (2) sont alors modifiées pour n'utiliser que les échantillons codés/décodés et les valeurs de  $\bar{e}$  transmises :

$$(4a) \left\{ \begin{array}{l} \hat{x}(m, n) = \hat{\mathbf{H}}(m, n)^T \tilde{\mathbf{X}}(m, n) \\ \mathbf{H}(m, n) = \hat{\mathbf{H}}(m, n) + \mu \bar{e}(m, n) \tilde{\mathbf{X}}(m, n) \end{array} \right.$$

$\hat{\mathbf{H}}(m, n)$  étant toujours obtenu par l'expression (3).

Sous cette forme, il apparaît un couplage prédiction-quantification qui est étudié dans [1].

Au décodeur, on a :

$$(4b) \left\{ \begin{array}{l} \hat{x}'(m, n) = \hat{\mathbf{H}}'(m, n)^T \tilde{\mathbf{X}}'(m, n) \\ \mathbf{H}'(m, n) = \hat{\mathbf{H}}'(m, n) + \mu \bar{e}'(m, n) \tilde{\mathbf{X}}'(m, n) \end{array} \right.$$

Avec ces expressions, il est facile de montrer que, s'il n'y a pas d'erreur de transmission et que si les conditions initiales sont les mêmes, on obtient l'alignement

du codeur et du décodeur :  $\mathbf{H}'(m, n) = \mathbf{H}(m, n)$  et  $\hat{x}'(m, n) = \hat{x}(m, n)$ .

## 2.2. ADAPTATION DU QUANTIFICATEUR

On distingue principalement deux types d'altérations créées par le codage [8] :

- le bruit granulaire visible dans les zones uniformes et lié à la résolution limitée du quantificateur;
- le phénomène de saturation apparent dans les zones où les niveaux extrêmes de quantification sont atteints.

Les différents types d'erreurs ne sont donc pas uniformément répartis dans l'image et leurs influences sont perçues différemment par l'œil. Aussi une approche moyenne, consistant à déterminer un quantificateur optimal minimisant un critère du type :  $E\{|e-\bar{e}|^k\}$ , se révèle être assez limitée.

Une façon de lever la difficulté est de chercher à modifier le critère d'optimisation afin de tenir compte des propriétés de la vision [9]. On peut également concevoir le quantificateur à partir de tests subjectifs [8].

Une démarche parallèle [10] consiste à rendre le quantificateur adaptatif. Dans cette méthode, les niveaux de quantification sont modifiés suivant la dynamique locale de l'erreur. Dans les zones uniformes, la quantification est choisie plus fine que dans les régions de contours où l'on enregistre de forts gradients de luminance et des erreurs de prédiction plus élevées. La stratégie que nous avons adoptée entre dans cette dernière catégorie.

Pour la conception du quantificateur, nous avons supposé que l'erreur de prédiction suit une loi de Laplace ( $p_E(e) = (1/\sqrt{2}\sigma_e) \exp(-(\sqrt{2}/\sigma_e)|e|)$ ). Une estimation  $s_e$  de l'écart-type permet de normer  $e$  et de se ramener ainsi à un signal  $e_1$ . Puis, nous avons calculé les niveaux de décision et de quantification d'un quantificateur de Max [11],  $Q_1$ , opérant sur  $e_1$ . Bien évidemment, en sortie du dispositif, il faut effectuer une opération inverse qui est une multiplication par  $s_e$ . Au décodeur, qui reçoit un code correspondant à  $\bar{e}_1 = Q_1[e_1]$ , il faut réaliser le même type de dénormalisation.

Pour estimer l'écart-type  $\sigma_e$ , deux approches doivent être distinguées. On peut faire l'estimation à partir de  $e(m, n)$ . On parle alors d'adaptation « avant » et,  $e(m, n)$  n'étant pas connu du décodeur, il faut transmettre  $\sigma_e$  comme information supplémentaire. L'autre possibilité consiste à réaliser une adaptation « arrière » en estimant  $\sigma_e$  à partir de  $\bar{e}(m, n)$ . On obtient alors une estimation de moins bonne qualité mais le décodeur peut retrouver, de lui-même, la valeur de l'erreur de prédiction quantifiée. Ce type de fonctionnement est illustré par la figure 2.

On peut prendre comme estimée de la variance :

$$(5) \sigma_e^2(m, n) = (1-\lambda_1)(1-\lambda_2) \times \sum_{i=0}^m \sum_{j=0}^n \lambda_1^{n-j} \lambda_2^{m-i} \bar{e}(i, j)^2$$

$\lambda_1$  (resp.  $\lambda_2$ ) est un facteur d'oubli horizontal (resp. vertical) compris entre 0 et 1.

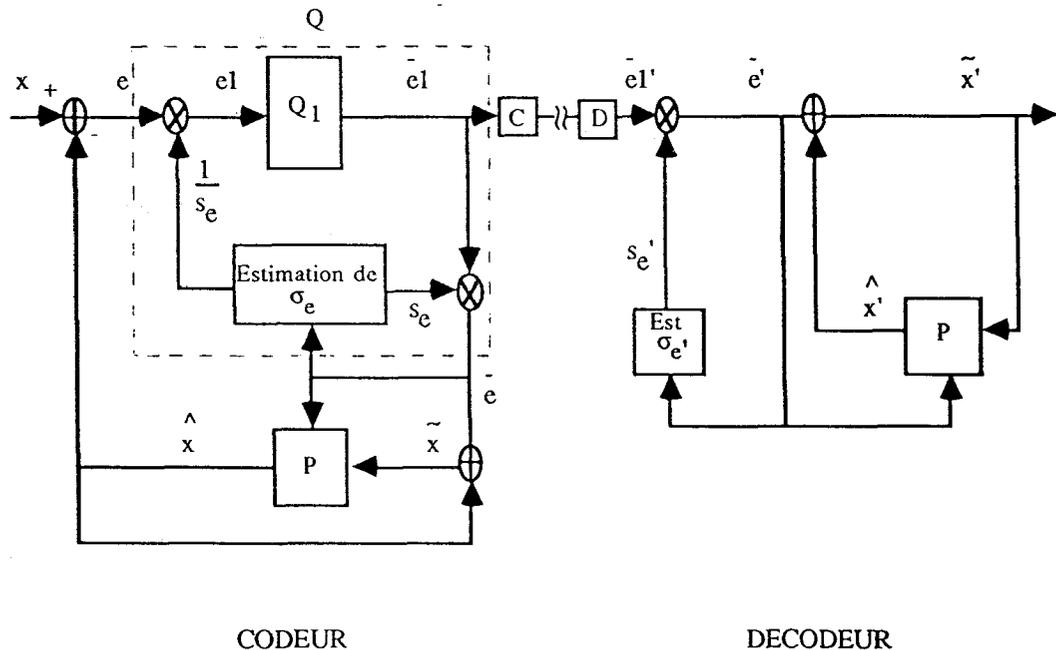


Fig. 2.

On ne fait donc qu'une estimation quart de plan, mais celle-ci a le mérite de conduire à une formulation réursive simple de  $\sigma_e^2(m, n)$ .

Il est, en effet, immédiat que :

$$(6a) \quad \sigma_e^2(m, n) = \lambda_1 \sigma_e^2(m, n-1) + \lambda_2 \sigma_e^2(m-1, n) - \lambda_1 \lambda_2 \sigma_e^2(m-1, n-1) + (1-\lambda_1)(1-\lambda_2) \bar{e}(m, n)^2$$

pour  $m \neq 0$  et  $n \neq 0$

Dans les cas particuliers de la première ligne et de la première colonne, on peut prendre :

$$(7) \quad s_e^2(m, n) = \begin{cases} \frac{\lambda_1 \sigma_e^2(m, n-1) + \lambda_2 \sigma_e^2(m-1, n) - \lambda_1 \lambda_2 \sigma_e^2(m-1, n-1)}{\lambda_1 + \lambda_2 - \lambda_1 \lambda_2} & \text{si } m \neq 0 \text{ et } n \neq 0 \\ \sigma_e^2(0, n-1) & \text{si } m=0 \text{ et } n \neq 0 \\ \sigma_e^2(m-1, 0) & \text{si } m \neq 0 \text{ et } n=0 \end{cases}$$

On calcule ensuite :

$$(8) \quad \bar{e}(m, n) = K s_e(m, n) Q_1 \left[ \frac{e(m, n)}{K s_e(m, n)} \right]$$

Le facteur K permet ici d'introduire un degré de liberté par rapport au modèle statistique retenu pour  $e_1(m, n)$ .

Enfin, on peut mettre à jour  $\sigma_e^2(m, n)$  à partir de  $s_e(m, n)$  et  $\bar{e}(m, n)$  :

$$(9) \quad \sigma_e^2(m, n) = \begin{cases} (\lambda_1 + \lambda_2 - \lambda_1 \lambda_2) s_e^2(m, n) + (1-\lambda_1)(1-\lambda_2) \bar{e}(m, n)^2 & \text{si } m \neq 0 \text{ et } n \neq 0 \\ \lambda_1 s_e^2(0, n-1) + (1-\lambda_1) \bar{e}(0, n)^2 & \text{si } m=0 \text{ et } n \neq 0 \\ \lambda_2 s_e^2(m-1, 0) + (1-\lambda_2) \bar{e}(m, 0)^2 & \text{si } m \neq 0 \text{ et } n=0 \end{cases}$$

$$(6b) \quad \sigma_e^2(m, n) = \begin{cases} \lambda_1 \sigma_e^2(0, n-1) + (1-\lambda_1) \bar{e}(0, n)^2 & \text{pour } m=0 \text{ et } n \neq 0 \\ \lambda_2 \sigma_e^2(m-1, 0) + (1-\lambda_2) \bar{e}(m, 0)^2 & \text{pour } m \neq 0 \text{ et } n=0 \end{cases}$$

En fait, puisque l'on a besoin d'une estimation de l'écart-type avant de disposer de  $\bar{e}(m, n)$ , il faut procéder en trois étapes.

Tout d'abord, on fait une estimation *a priori* de  $\sigma_e^2(m, n)$  :

L'implantation de cette méthode nécessite la mémorisation de  $N+1$  valeurs  $\sigma_e^2(m, n)$ .

Notons qu'un algorithme plus simple peut également être utilisé pour adapter le quantificateur [12]. Mais ce dernier s'est révélé moins performant lors de nos simulations.

### 3. Codage intra-image

Nous avons appliqué les méthodes précédentes au cas du codage intra-image.

Nous nous sommes alors limités à un prédicteur d'ordre 3. En effet, il avait été remarqué, dans plusieurs études, que le gain en performance devenait faible au-delà de cet ordre. Nous avons, par ailleurs, pu vérifier ce fait, lors des simulations. D'un point de vue théorique, on sait, de plus, que ce type de

codage « voiture -00 »

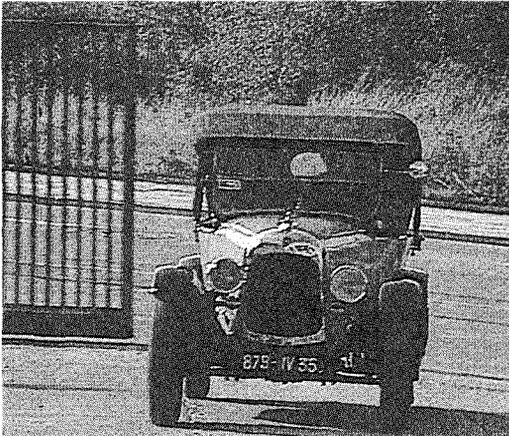


image originale

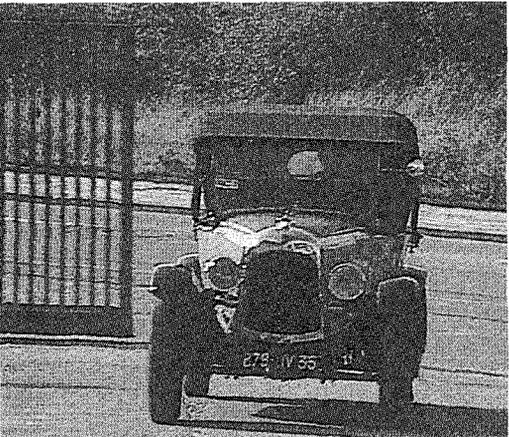
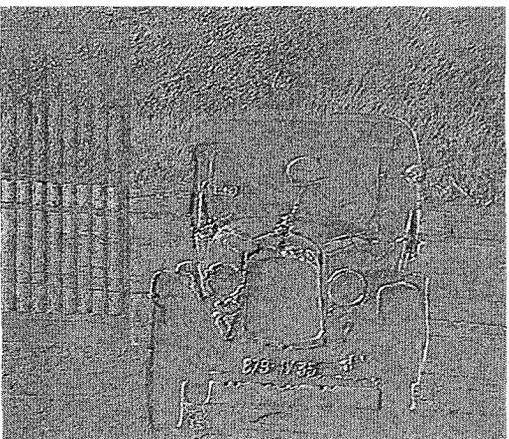


image codée



erreur de prédiction  
(histogramme dilaté d'un facteur 2,2)

codage « statue »

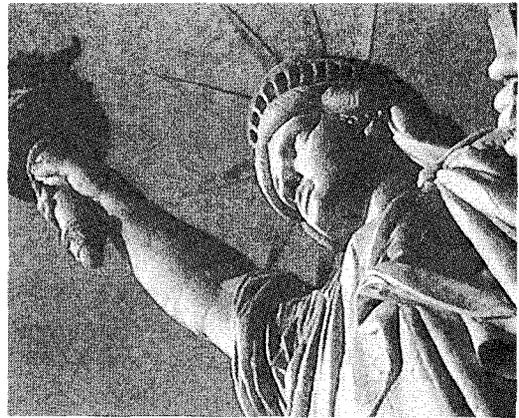


image originale



image codée



erreur de prédiction  
(histogramme dilaté d'un facteur 5)

prédicteur conduit à une erreur spatialement décorrélée, dans le cas d'un modèle à covariance exponentielle séparable

$$(\text{Cov}[x(m, n), x(p, q)] = \sigma_x^2 \rho_1^{|m-p|} \rho_2^{|n-q|}).$$

On a alors l'expression suivante :

$$(10) \quad \hat{x}(m, n) = \hat{h}_{0,1}(m, n) \tilde{x}(m, n-1) + \hat{h}_{1,0}(m, n) \tilde{x}(m-1, n) + \hat{h}_{1,1}(m, n) \tilde{x}(m-1, n-1)$$

Les simulations que nous avons effectuées à partir de ce prédicteur, ont porté sur deux images monochromes de types différents : l'une représente la statue de la liberté, l'autre est extraite de la séquence « voiture » (normalisation européenne COST 211 bis). Le format de ces images est de  $480 \times 512$  pixels pour la première et de  $512 \times 512$  pixels pour la seconde, les niveaux de gris allant de 0 à 255.

Le critère d'évaluation de la quantité de ces images est donné par la distorsion quadratique moyenne :

$$(11) \quad D = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [x(i, j) - \tilde{x}(i, j)]^2$$

ou, de façon équivalente, par  $\sqrt{D}$ .

Nous précisons aussi l'entropie  $H(\bar{e})$  de  $\bar{e}(m, n)$  et  $P_e$  la puissance moyenne de  $e(m, n)$ .

L'entropie  $H(\bar{e})$  est une borne inférieure de la longueur moyenne  $\bar{L}$  du code utilisé pour transmettre chaque valeur de l'erreur quantifiée. On peut également indiquer le taux de compression défini ici par  $8/\bar{L}$ . Quant à  $P_e$ , elle renseigne plus particulièrement sur la qualité de la prédiction.

Notons que, dans le cas des images, on traite un signal qui n'est pas centré. Il semblerait donc logique d'ajouter à l'équation (10) un terme supplémentaire pour tenir compte de la moyenne. D'un point de vue pratique, si (10) est utilisée directement, on trouve, avec des valeurs usuelles des coefficients de prédiction, que  $E\{e(m, n)\} \approx 1$ . Ceci peut entraîner une dissymétrie dans la répartition des erreurs qui peut être gênante dans le cas d'un quantificateur à caractéristique impaire. Pour remédier à cet inconvénient, la méthode la plus simple consiste à ramener la plage de variation du signal à un intervalle centré  $[-127, 128]$ , à l'entrée de la boucle de codage. On vérifie alors expérimentalement que l'erreur de prédiction est centrée et que des méthodes plus élaborées d'estimation de  $E\{x(m, n)\}$  ne conduisent pas à des résultats très supérieurs.

Pour ce qui concerne le quantificateur, nous nous sommes restreints à trois niveaux de quantification. Ceci permet d'éviter une partie des inconvénients liés aux quantificateurs à deux niveaux, bien connus pour leurs applications dans la modulation  $\Delta$ . Ce choix se révèle surtout intéressant, si l'on adopte un codage entropique pour la transmission. Dans nos simulations, nous n'avons cependant pas étudié cet aspect qui nécessiterait la gestion d'un buffer de transmission pour uniformiser le débit.

On trouvera dans les tableaux I et II les principaux résultats obtenus. L'avantage des techniques adaptati-

ves apparaît d'autant plus clairement que les prédicteurs fixes utilisés pour les deux types d'images n'étaient pas les mêmes. La « statue » était mieux modélisée par un prédicteur tel que  $h_{0,1}=0,96$ ;  $h_{1,0}=0,95$ ;  $h_{1,1}=-0,91$  alors que pour l'image « voiture », dont les textures sont très fines et le bruit de source important, on obtenait de meilleurs résultats avec  $h_{0,1}=0,628$ ;  $h_{1,0}=0,443$ ;  $h_{1,1}=-0,081$ .

TABLEAU I

(\* $h_{0,1}=0,96$ ;  $h_{1,0}=0,95$   $h_{1,1}=-0,91$ )  
Image statue : moyenne : 125,5 (signal avant codage),  
puissance : 17960.

	P* fixe Q fixe	P adaptatif Q fixe	P* fixe Q adaptatif	P adaptatif Q adaptatif
$\sqrt{D}$	6,7	6,0	6,4	4,8
$P_e$	136	80	136	71,0
$H(\bar{e})$ (bits/points)	1,2	0,8	1,5	1,4

TABLEAU II

(\*\* $h_{0,1}=0,628$ ;  $h_{1,0}=0,443$ ;  $h_{1,1}=-0,081$ )  
Image voiture -  $\infty$  : moyenne : 111,6, puissance : 14280.

	P** fixe Q fixe	P adaptatif Q fixe	P** fixe Q adaptatif	P adaptatif Q adaptatif
$\sqrt{D}$	11,5	10,4	10,0	9,0
$P_e$	355	302	338	287
$H(\bar{e})$ (bits/points)	1,0	1,2	1,5	1,5

Parallèlement, dans les deux cas, la valeur optimale du pas d'adaptation du prédicteur était  $\mu \approx 10^{-5}$  et les caractéristiques du quantificateur adaptatif étaient :  $\lambda_1 \approx \lambda_2 \approx 0,3$  et  $K \approx 1,1$ . Ces paramètres ont fait l'objet d'une optimisation numérique, pour chaque cas.

Il faut souligner que l'image obtenue par les méthodes adaptatives est de bonne qualité et que l'adaptation du quantificateur joue un rôle primordial à cet effet.

Remarquons, enfin, que par rapport à une configuration quantificateur et prédicteur fixes, les méthodes adaptatives conduisent environ à un accroissement d'un facteur 5 du nombre de multiplications à effectuer.

#### 4. Codage inter-images

Des simulations ont également été effectuées sur une séquence d'images. On doit alors modifier la structure du prédicteur pour utiliser à la fois les informations spatiales et temporelles. Pour limiter le nombre de mémoires nécessaires, nous n'avons utilisé que l'image précédant celle à coder.

Par ailleurs, nous avons constaté expérimentalement qu'il était possible de se limiter à neuf coefficients de prédiction au niveau de l'image précédente (fig. 3).

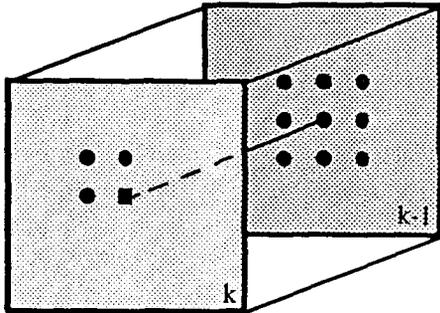


Fig. 3.

Si l'on désigne par  $k$  le numéro de l'image, l'équation de prédiction prend la forme suivante :

$$(12) \quad \hat{x}(m, n, k) = \sum_{\substack{i=0 \\ (i,j) \neq (0,0)}}^1 \sum_{j=0}^1 \hat{h}_{i,j,0}(m, n, k) \tilde{x}(m-i, n-j, k) + \sum_{i=-1}^1 \sum_{j=-1}^1 \hat{h}_{i,j,1}(m, n, k) \tilde{x}(m-i, n-j, k-1)$$

Le choix d'un tel support, pour le filtre, a une interprétation physique. Il correspond à une possibilité de prédiction d'un déplacement limité à une translation verticale et/ou horizontale d'au plus 1 pixel. On a alors une information sur la direction locale du mouvement. On attend de l'adaptation qu'elle privilégie la prédiction temporelle dans les zones à faible mouvement et qu'elle conduise à une estimation spatiale dans les zones subissant des déplacements importants. Cette démarche est à rapprocher des méthodes où est faite une sélection adaptative de prédicteurs intra ou inter-images [13].

Pour donner une idée de l'amélioration apportée par ce type de prédiction, nous avons comparé deux méthodes de codage inter-images. Les deux expériences ont été menées, sur les deux premières images de la séquence « voiture », avec un quantificateur à trois niveaux. Si l'on transmet la différence entre les deux images successives, on a  $\sqrt{D}=16$  alors qu'en utilisant le procédé que nous proposons, on obtient  $\sqrt{D}=7,4$ .

Une simulation complète de la transmission des 20 premières images de cette même séquence a également été menée, en utilisant une prédiction et une quantification adaptatives (fig. 4). La première image fait l'objet d'un codage intra-image alors que, dans les suivantes, la prédiction temporelle utilise l'image précédemment transmise donc dégradée. On trouve en moyenne, dans ces conditions,  $\sqrt{D}=8,4$  et  $H(\bar{e})=1,5$ .

distorsion

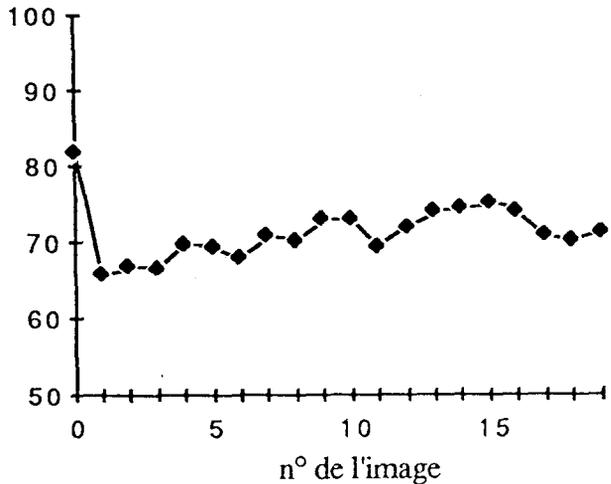


Fig. 4. - Codage de la séquence « voiture ».

### 5. Décodage en présence d'erreurs de transmission

Les techniques de codage prédictif sont sensibles aux erreurs de transmission. En effet, un examen de la figure 1, montre que le décodeur a une structure bouclée. Ceci induit un effet de propagation des perturbations sur  $\tilde{x}'(n)$  créées par les erreurs de transmission. Dans le cas d'un schéma MIC-D fixe,  $\tilde{x}'(n)$  est le résultat d'un filtrage linéaire récursif de  $\bar{e}'(n)$  et ce type de problème est résolu en assurant la stabilité asymptotique de l'inverse du prédicteur.

Lorsque l'on introduit des algorithmes adaptatifs, l'effet des erreurs de transmission peut être encore plus critique. Les perturbations sur les signaux entraînent alors une différence entre les comportements des opérateurs adaptatifs du codeur et du décodeur. Ce phénomène peut ainsi accentuer les erreurs sur les signaux décodés.

#### 5.1. ALIGNEMENT DES PRÉDICTEURS

L'algorithme du gradient, que nous avons présenté dans le paragraphe 2.1, conduit à un comportement divergent du décodeur, si le canal est bruité. Une seule erreur de transmission suffit alors pour rendre une image non reconnaissable.

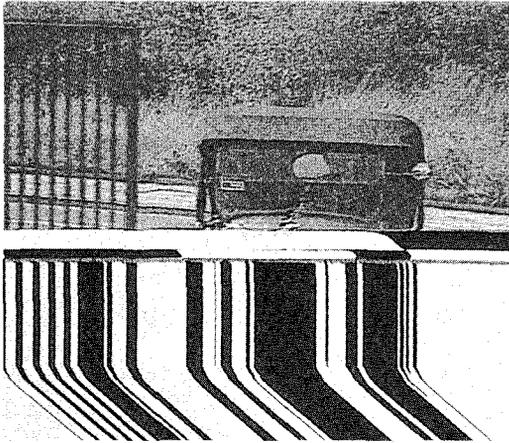
On peut penser limiter la divergence des coefficients du prédicteur en introduisant une contrainte de régularisation, au niveau du critère  $J_R(m, n)$  à minimiser. On a alors :

$$(13) \quad J_R(m, n) = E \{ e^2(m, n) \} + \frac{\delta}{\mu} \mathbf{H}_R(m, n)^T \mathbf{H}_R(m, n)$$

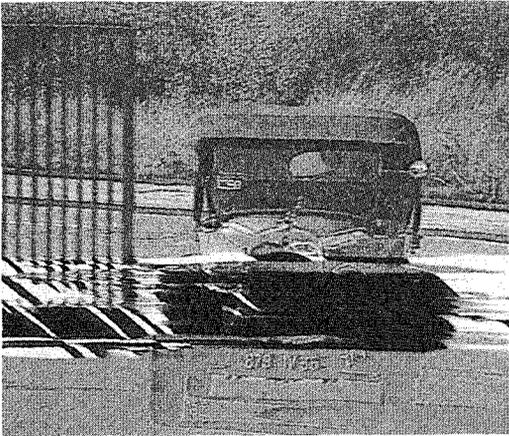
où  $\delta$  est une constante positive réelle.

$\mathbf{H}_R(m, n)$  est le vecteur des coefficients du prédicteur

décodage de « voiture - 00 »



Algorithme du gradient, 1 erreur  
(quantificateur fixe, erreur au centre de l'image).



Algorithme régularisant, 1 erreur  
( $\delta = 10^{-2}$ , quantificateur fixe, erreur au centre de l'image)



Algorithme stabilisant, 260 erreurs  
(quantificateur adaptatif modifié)

utilisé dans cette approche régularisante et, *a priori*, est variable si les signaux sont non stationnaires. A partir de (13), on aboutit à un algorithme de gradient avec facteur de fuite et les équations (4) sont modifiées de la façon suivante :

$$(14) \left\{ \begin{aligned} \hat{x}(m, n) &= \hat{\mathbf{H}}_{\mathbf{R}}(m, n)^T \tilde{\mathbf{X}}(m, n) \\ \mathbf{H}_{\mathbf{R}}(m, n) &= (1 - \delta) \hat{\mathbf{H}}_{\mathbf{R}}(m, n) \\ &\quad + \mu \bar{e}(m, n) \tilde{\mathbf{X}}(m, n) \end{aligned} \right.$$

Les résultats obtenus à partir de ces expressions sont meilleurs visuellement que les précédents mais l'alignement du codeur et du décodeur ne parvient pas à être rétabli. La persistance d'un biais sur les coefficients de prédiction et les signaux du codeur et du décodeur a pu être démontrée pour des signaux sinusoïdaux à une dimension [1].

Cependant il est possible d'aboutir à un algorithme de stabilisation plus performant, en cherchant également à modifier le critère minimisé [14], [1] :

$$(15) J_S(m, n) = E \{ e^2(m, n) \} + \alpha E \{ [\hat{x}(m, n) - \hat{x}_f(m, n)]^2 \}$$

où  $\alpha$  est une constante réelle positive et,

$$(16) \hat{x}_f(m, n) = \mathbf{H}_f^T \tilde{\mathbf{X}}(m, n)$$

est le résultats d'une prédiction fixe de  $x(m, n)$ .

Lorsque le prédicteur a tendance à diverger, le terme ajouté au critère devient grand, ce qui conduit à une diminution de  $|\mathbf{H}_S(m, n)|$ , la norme du vecteur des coefficients de prédiction.

Dans un contexte stationnaire, la minimisation de  $J_S$  donne, comme solution optimale :

$$(17) \hat{\mathbf{H}}_S = \frac{1}{1 + \alpha} \hat{\mathbf{H}} + \frac{\alpha}{1 + \alpha} \mathbf{H}_f$$

où  $\hat{\mathbf{H}}$  est la solution optimale approchée par l'algorithme du gradient stochastique.

On voit donc que  $\hat{\mathbf{H}}_S$  est le barycentre de  $(\hat{\mathbf{H}}, 1)$  et  $(\mathbf{H}_f, \alpha)$ .

Ceci signifie que  $\mathbf{H}_f$  est une estimation *a priori* des paramètres de prédiction et que  $\alpha$  peut être interprété comme un degré de confiance sur cette valeur. Un terme jouant un rôle similaire à  $\mathbf{H}_f$  avait déjà été introduit dans [15], dans le cas d'un choix adaptatif entre plusieurs prédicteurs fixes.

Il apparaît aussi que l'on introduit un biais

$$\hat{\mathbf{H}} - \hat{\mathbf{H}}_S = \frac{\alpha}{1 + \alpha} (\hat{\mathbf{H}} - \mathbf{H}_f),$$

par rapport au problème de minimisation de  $E \{ e^2(m, n) \}$ .

Tout calcul fait, l'algorithme stabilisant prend la forme :

$$(18) \left\{ \begin{aligned} \hat{x}(m, n) &= \hat{\mathbf{H}}_S(m, n)^T \tilde{\mathbf{X}}(m, n) \\ \mathbf{H}_S(m, n) &= \hat{\mathbf{H}}_S(m, n) + \mu [\bar{e}(m, n) \\ &\quad - \alpha (\hat{\mathbf{H}}_S(m, n) - \mathbf{H}_f)^T \tilde{\mathbf{X}}(m, n)] \tilde{\mathbf{X}}(m, n) \end{aligned} \right.$$

Pour éprouver la robustesse de cet algorithme, nous avons simulé le fonctionnement du décodeur, en présence de taux d'erreur importants. Nous avons réparti les erreurs, dans l'image, de façon aléatoire, selon une loi uniforme. Dans une approche simplificatrice, nous avons supposé que le code correspondant à l'erreur

quantifiée appartenait à un alphabet ternaire. Pour chaque code erroné, nous avons imposé une transition équiprobable de la valeur réelle vers l'un ou l'autre des deux autres codes. Les résultats ainsi obtenus sont donnés dans le tableau III. Ils montrent qu'avec un quantificateur fixe, et en choisissant  $\alpha=0,7$ ,  $h_{f0,1}=0,7$ ,  $h_{f1,0}=0,7$  et  $h_{f1,1}=-0,49$ , la qualité des images n'est pas affectée jusqu'à un taux d'erreur de  $10^{-3}$ .

TABLEAU III

*Décodage de voiture -00  
avec quantificateur fixe*

		Taux d'erreur			
		0	$10^{-4}$	$10^{-3}$	$10^{-2}$
D	Prédicteur fixe $H(\bar{e})=1$	133	134	144	234
	Algorithme stabilisant $H(\bar{e})=1,2$	103	104	111	177

### 5. 2. MODIFICATION DU QUANTIFICATEUR

En présence d'erreurs de transmission, l'estimateur de variance utilisé pour la quantification adaptative ne pose pas de problème de stabilité dans la pratique. Il est cependant nécessaire de le modifier légèrement pour parvenir à ce résultat. En effet, pour une erreur donnée sur le code transmis, (8) montre que l'erreur maximale sur  $\bar{e}(m, n)$  est fonction de  $s_e(m, n)$ . Lorsque cette valeur est grande, l'erreur peut être importante et ne devient négligeable qu'après un nombre élevé d'échantillons. Par rapport au cas où le quantificateur est fixe, on peut donc considérer, que le nombre d'erreurs est multiplié par un facteur pouvant être élevé.

Un moyen de limiter ce phénomène est d'imposer une borne supérieure  $S_{\max}$  à  $s_e(m, n)$ . Ainsi l'erreur maximale sur  $e(m, n)$  est bornée. Pour éviter que  $s_e(m, n)$  ne puisse devenir nul dans (8), nous lui avons également imposé une borne inférieure  $S_{\min}$ .

Le tableau IV montre que l'on obtient ainsi un algorithme très robuste si l'on prend  $S_{\min}=10$  et  $S_{\max}=40$ .

TABLEAU IV

*Décodage de voiture  $-\infty$  avec quantificateur adaptatif  
(prédicteur stabilisé)*

		Taux d'erreur			
		0	$10^{-4}$	$10^{-3}$	$10^{-2}$
D	Quantificateur r modifié $H(\bar{e})=1,4$	77	78	86	164

## 6. Conclusions-perspectives

Nous avons montré l'intérêt de l'introduction de méthodes adaptatives dans le codage prédictif intra et inter-images.

Nous obtenons, pour la séquence « voiture » des résultats visuels de qualité tout à fait acceptable, pour un taux de compression de l'ordre de 5. Par ailleurs, la complexité de la mise en œuvre reste faible.

Nous avons également proposé des variantes des algorithmes adaptatifs qui sont peu sensibles aux erreurs de transmission. Ainsi nos simulations indiquent que les performances du codage se maintiennent jusqu'à un taux d'erreur de  $10^{-3}$ .

Dans la suite de cette étude, nous pensons qu'il serait intéressant d'utiliser un filtrage adaptatif, non plus transverse, mais récursif. De plus, des travaux théoriques sont en cours sur la stabilité des méthodes de décodage, en présence d'erreurs de transmission.

## Remerciements

Nous tenons tout particulièrement à remercier Odile Macchi, Madeleine Bonnet et Christine Uhl pour la pertinence des remarques dont elles nous ont fait part, au cours de cette étude.

*Manuscrit reçu le 26 février 1988.*

## BIBLIOGRAPHIE

- [1] M. BONNET, Codage numérique des signaux par quantification et prédiction adaptatives couplées, *Thèse d'Etat*, Université Paris-Sud, janvier 1988.
- [2] A. K. JAIN, Advances in Mathematical Models for Image Processing, *Proceedings of the IEEE*, 69, n° 5, May 1981, p. 502-528.
- [3] K. KRETZ, Codage MIC-Différentiel à prédiction adaptative en télévision : qualité visuelle, réduction du débit et susceptibilité aux erreurs de transmission, *Annales des Télécommunications*, n° 7-8, juillet-août 1982, p. 299-234.
- [4] P. A. MARAGOS, P. W. SCHAFER et R. M. MERSEREAU, Two Dimensional Linear Prediction and its Application to Adaptive Predictive Coding of Images, *IEEE Transactions on ASSP*, 32, n° 6, December 1984, p. 1213-1229.
- [5] O. MACCHI et M. BELLANGER, Le point sur le filtrage adaptatif transverse, *Onzième colloque GRETSI*, juin 1987, p. G1-G14.
- [6] J.-C. PESQUET et G. TZIRITAS, Adaptive Methods Applied to Intra and Interframe Coding, *Visual Communications and Image Processing III, SPIE*, 1988.
- [7] S. T. ALEXANDER et S. A. RAJALA, Image Compression Results Using the LMS Adaptive Algorithm, *Correspondence, IEEE Transactions on ASSP*, 33, n° 3, June 1985, p. 712-714.
- [8] P. PIRSCH, Design of DPCM Quantizers for Video Signals Using Subjective Tests, *IEEE Transactions on Communications*, 69, n° 7, July 1981, p. 990-1000.
- [9] D. K. SHARMA, Design of Absolutely Optimal Quantizers for a Wide Class of Distorsion Measures, *Transactions on Information Theory*, 24, n° 6, November 1978, p. 693-702.
- [10] N. S. JAYANT et P. NOLL, *Digital Coding of Waveforms*, Prentice Hall, 1984.
- [11] J. MAX, Quantizing for Minimum Distorsion, *IEEE Transactions on Information Theory*, 6, March 1960, p. 7-12.

- [12] L. H. ZETTERBERG, S. ERICSSON et H. BRUSEWITZ, Interframe DPCM with Adaptive Quantization and Entropy Coding, *IEEE Transactions on Communications*, 30, n° 8, August 1982, p. 1888-1899.
- [13] H. G. MUSMANN, P. PIRSCH et H. J. GRALLERT, Advances in Picture Coding, *Proceedings of the IEEE*, 73, n° 4, April 1985, p. 523-548.
- [14] M. JAIDANE-SAIDANE, Stabilité des prédicteurs récursifs adaptatifs: application à la numérisation des signaux téléphoniques, *Thèse d'État*, Université Paris-Sud, juillet 1987.
- [15] N. F. MAXEMSCHUK et J. A. STULLER, Reduction of Transmission Error Propagation in Adaptively Predicted DPCM Encoded Pictures, *The Bell System Technical Journal*, 58, n° 6, July-August 1979, p. 1413-1423.