

MODELISATION D'IMAGES PAR DES MODELES AUTOREGRESSIFS 2D  
NON-STATIONNAIRES EXCITES PAR DES TRAINS D'IMPULSIONS

M. BARLAUD, L. BLANC-FERAUD, P. MATHIEU, F. DUBUS

LABORATOIRE DE SIGNAUX ET SYSTEMES UA 814 DU CNRS UNIVERSITE DE NICE  
41 BD NAPOLEON III - 06041 NICE CEDEX

RESUME

En vue d'un codage efficace d'images noir et blanc, nous présentons un nouvel algorithme de compression numérique d'image digitalisée. Celle-ci est modélisée par la sortie d'un filtre prédictif 2D évolutif excité par un train d'impulsions.

En première partie est présenté le filtre autorégressif 2D. La non-stationnarité de l'image est prise en compte par la variation des paramètres du modèle en fonction de la position du pixel estimé.

L'entrée multipulse du modèle est ensuite estimée par deux méthodes différentes. La première est basée sur la connaissance des erreurs de prédiction du modèle qui représente les contours des formes de l'image. La deuxième est une simplification de la méthode d'ATAL.

Ces différentes techniques ont été testées sur une image (256x256) pixels et les résultats expérimentaux sont présentés.

INTRODUCTION

La compression numérique d'images est très utile pour la transmission d'images vidéo, et permet une réduction de la place mémoire pour le stockage. Ces avantages peuvent être très intéressants, notamment si l'on a affaire à une suite d'images variant dans le temps.

Plusieurs méthodes de compression sont déjà bien connues et très utilisées pour la transmission de signaux 1D de parole. On distingue principalement trois techniques différentes : le codage en sous bande, le codage par transformée et le codage prédictif. Une méthode de codage en sous bande a été étendue au cas 2D [5] et s'est avérée efficace. Le taux de compression obtenu est de 1/8, et l'image est restituée avec un rapport signal sur bruit (SNR) de 32,5 DB. Des méthodes de représentation de l'image par des modèles A.R. prédictifs ont aussi été proposées [1],[2]. L'image est considérée dans ce cas comme la sortie d'un filtre prédictif 2D dont les entrées sont obtenues par codage de l'erreur de prédiction, utilisant des techniques multipulses d'ATAL (MPE) [6] ou de DEPRETERRE (RPE) [8].

Ces modèles supposent l'image stationnaire sur des fenêtres régulièrement définies a priori. Nous proposons une modélisation n'utilisant pas cette hypothèse, par un modèle autorégressif 2D évolutif.

Différentes méthodes d'estimation de la position et de l'amplitude des pulses composant l'entrée du modèle sont étudiées. Une analyse est faite sur la dégradation de l'image reconstituée en fonction du nombre de pulses retenus et de la technique utilisée, et des résultats sont présentés.

SUMMARY

This paper describes a new efficient algorithm for the encoding of images. The images are modeled as the output of a non-stationary 2D linear prediction filter excited by a 2D multipulse input.

In the first part the non-stationary 2D autoregressive filter is presented. The space-varying parameters of the model are approximated by linear decomposition on vertical and horizontal functions basis.

The second part presents two new multipulse techniques for input coding. The input is an approximation of prediction residual that represents image edges. The first method is directly based on the prediction knowledge. The second is a simplification of ATAL's method.

These methods have been tested on a gray-level image of (256 x 256) pixels and experimental results are presented.

MODELISATION A.R.

La représentation d'une image noir et blanc digitalisée, par un modèle autorégressif à deux dimensions consiste à supposer que l'intensité  $s(n,m)$  d'un pixel (élément de l'image) de ligne  $n$  et de colonne  $m$ , est fonction linéaire des intensités d'un certain nombre de pixels du passé. La notion de passé pour un signal échantillonné 2D découle naturellement de la manière usuelle de lire les  $N \times N$  pixels composant une image: de haut en bas et de gauche à droite. Etant donné un pixel  $(n,m)$  "présent", le passé et le futur sont donnés par les demi-plans asymétriques [1] représentés figure 1. La relation linéaire est :

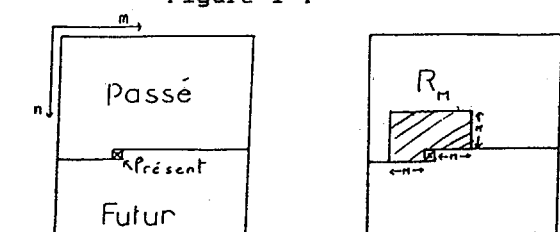
$$s(n,m) = \sum_{(n-1,m-k) \in R} a(n,m) * s(n-1,m-k) + r(n,m) \quad (1)$$

où  $R$  est une région du passé définie par :

$$R = \{ (n-1,m-k) / (1=0, 1 \leq k \leq M) \cup (1 \leq l \leq M, -M \leq k \leq M) \}$$

$M$  est l'ordre du modèle. Des études expérimentales [2] ont montré qu'un modèle d'ordre 2 ou 1 suffisait pour la modélisation.

Figure 1 :



Causalité demi-plan asymétrique

Région de support du modèle (R)



La non stationnarité du signal est prise en compte en faisant varier les paramètres

$a_i$  en fonction des coordonnées  $(n, m)$  du pixel que l'on estime. En supposant que cette variation est assez lente, on peut écrire les 4 paramètres du modèle à l'ordre 1, comme combinaisons linéaires de fonctions de base verticales et horizontales  $f$  et  $g$  :

$$a_i(n, m) = \underline{a}^T \begin{bmatrix} 1 \\ f(n) \\ g(m) \end{bmatrix}$$

$$\text{où } \underline{a}^T = [ a_0, a_{V1}, \dots, a_{Vp}, a_{H1}, \dots, a_{Hp} ]$$

$$f(n) = [ f_1(n), \dots, f_p(n) ]$$

$$g(m) = [ g_1(m), \dots, g_p(m) ]$$

Nous avons retenu pour fonction de base les polynômes de Legendre décalés et modifiés [3], qui donnent expérimentalement de très bons résultats en dimension 1 sur des signaux de parole [4]. Les 4 paramètres variables du modèle d'ordre 1 sont ainsi remplacés par un vecteur de  $4 \times (1 + 2p)$  paramètres constants

$$\underline{a}^T = [ a_0, \dots, a_{4T} ] \quad \text{qu'il faut estimer.}$$

$$\text{En posant } S_{n,m}^T = [ s_1, \dots, s_{4T} ] \quad \text{avec}$$

$$s_{kl}^T = s_{kl}(n, m) [ 1 \quad f(n) \quad g(m) ]^T, \quad \text{et } s_{kl}(n, m)$$

le facteur correspondant au paramètre  $a_{kl}$  (ou  $a_{kl}$ ) dans la relation (1), celle-ci s'écrit vectoriellement :

$$(2) \quad s_{n,m}^T = S_{n,m}^T \underline{a} + r(n, m)$$

On estime le vecteur des paramètres  $\underline{a}$  en minimisant le critère quadratique

$$\sum_{n,m=1}^N r(n, m)^2 \quad \text{où } r(n, m) \text{ est l'erreur de}$$

prédiction définie par la relation (2).

Le nombre de fonctions de base dans chaque direction peut être choisi égal à 2 ou 3, une augmentation de  $p$  au delà de 3 n'apportant pas d'amélioration significative sur le signal reconstitué.

En prenant  $p = 3$ , nous obtenons un modèle unique pour toute l'image, composé de 28 paramètres. Les méthodes prédictives proposées jusqu'à présent dans la littérature, requièrent l'estimation de modèles A.R. stationnaires sur des fenêtres généralement de dimension  $(32 \times 32)$  ce qui donne pour une image  $(256 \times 256)$ , un total de 256 paramètres (à l'ordre 1). Le modèle évolutif décrit réduit d'un facteur 10 le nombre de paramètres à estimer, et évite de plus les effets de rupture dus au découpage arbitraire de l'image.

## ESTIMATION DE L'ENTREE MULTIPULSE

### 1. Présentation, méthode optimale :

Le modèle estimé, il faut ensuite en déterminer l'excitation. L'entrée du modèle est une approximation de l'erreur de prédiction  $r(n, m)$ , définie en (2), obtenue grâce à un nombre limité d'impulsions. Le choix de l'entrée est un élément déterminant dans la qualité de la synthèse et le taux de réduction des données, entre lesquels il faut faire un compromis. On appellera taux de réduction le rapport entre le nombre de pixels ayant une valeur non nulle à l'entrée et le nombre total de pixels d'une image.

Le champ d'excitation est estimé successivement sur des fenêtres  $(16 \times 16)$  (ou  $(8 \times 8)$ ) partitionnant l'image, en minimisant sur chacune d'entre elles l'énergie de l'erreur entre l'image réelle et l'image synthétisée, engendrée par les impulsions.

Avec des notations à un seul indice pour simplifier l'écriture, on note  $e_t$  l'erreur au point  $t$  entre le signal réel et le signal reconstitué, après placement de  $m$  impulsions de position  $t$  et d'amplitude  $w(i)$  ( $i=1, \dots, m$ ). On a :

$$e_t = s_t - \sum_{i=1}^m h(t, t_i) * w(i)$$

où  $h(t, t_i)$  est la réponse impulsionnelle du filtre au point  $t$ , déclenché au point  $t_i$ , et  $s_t$  signal réel.

$E(m)$  l'énergie correspondante est :

$$E(m) = \sum_t (e_t)^2$$

En supposant que le nombre  $m$  de pulses à déterminer sur chaque fenêtre est fixé, la méthode optimale pour déterminer les positions et les amplitudes, consiste à calculer les amplitudes optimales, c'est-à-dire celles qui minimisent  $E(m)$ , pour tous les emplacements possibles des pulses sur le carré d'étude. Les positions étant fixées, la minimisation de  $E(m)$  par rapport aux  $w(i)$  s'obtient simplement, en annulant les dérivées partielles de  $E(m)$  et conduit à un système de  $m$  équations linéaires à résoudre [6], [7].

Cette minimisation ne pose donc pas de difficultés. Cependant cette procédure n'est pas applicable, car le nombre de configurations possibles pour positionner les pulses est très important. Ce choix trop large de positions à tester rend l'algorithme très complexe et très lourd en calculs. Son utilisation n'est de ce fait pas envisageable.

Les procédures utilisées en image comme en parole sont donc sous-optimales, et consistent à estimer séparément les positions, puis les amplitudes des pulses. La recherche de l'emplacement des entrées apparaît comme un problème majeur, un bon positionnement demandant énormément de calculs et étant primordial pour la qualité du signal reconstitué.

On rappelle dans la suite les deux solutions sous optimales les plus utilisées.

**2. Méthodes sous-optimales (Rappels) :**

La première est la méthode MPE (Multi-Pulse Excitation) proposée par ATAL [6], qui consiste à déterminer les impulsions séquentiellement de la façon suivante :

Si  $m$  impulsions sont déjà placées, l'impulsion  $(m+1)$  positionnée en  $t$  et d'amplitude  $w(m+1)$  à déterminer, donne une nouvelle erreur :

$$e_t^{m+1} = e_t^m - h(t, t_{m+1}) * w(m+1)$$

L'énergie  $E(m+1)$  à minimiser s'écrit en fonction de  $E(m)$  :

$$E(m+1) - E(m) = 2 w(m+1) \sum_t e_t h(t, t_{m+1}) + w(m+1)^2 \sum_t h(t, t_{m+1})^2$$

(3)

$w(m+1)$  est obtenu en annulant la dérivée partielle de  $E(m+1)$  par rapport à  $w(m+1)$  :

$$w(m+1) = \frac{\sum_t e_t h(t, t_{m+1})}{\sum_t h(t, t_{m+1})^2}$$

(4)

et en remplaçant (4) dans (3) on obtient :

$$E(m+1) = E(m) - \frac{(\sum_t e_t h(t, t_{m+1}))^2}{\sum_t h(t, t_{m+1})^2}$$

(5)

Le rapport formé de l'intercorrélation du carré de l'erreur avec la réponse du filtre sur la corrélation de la réponse impulsionnelle (5) doit être maximum en  $t$

afin que  $E(m+1)$  soit minimum. Toutes les possibilités pour  $t$  seront testées, l'emplacement retenu étant celui donnant un rapport maximal. Une fois  $t$  calculé, on peut soit déduire  $w(m+1)$  de (4), soit recalculer globalement les amplitudes des  $(m+1)$  pulses. La deuxième solution donne de meilleurs résultats, mais demande évidemment plus de calculs.

Cette méthode, appliquée à l'image représentée par un modèle stationnaire par bloc, donne, avec un taux de réduction de 1/6.4, une image reconstituée ayant un SNR d'environ 29,5 DB [2].

Dans le cadre de notre étude, il n'est pas possible d'utiliser une telle méthode, car la non stationnarité du modèle rend nécessaire le calcul en tout point de la réponse impulsionnelle du filtre déclenché en un point quelconque de la fenêtre.

Nous proposons en troisième partie une simplification de la méthode, permettant son application avec le modèle évolutif retenu.

Une autre technique utilisée est la méthode RPE (Regular Pulse Excitation) de DEPRETERRE. Les pulses sont positionnés a priori sur l'image, au sommets de figures géométriques (carrés, losanges) régulièrement réparties. Le nombre de possibilités de positionnement des pulses est ainsi extrêmement réduit, et il est possible de calculer les amplitudes optimales pour chaque configuration, afin de déterminer celle donnant l'erreur minimale [2]. Horne obtient avec cette méthode sur une image noir et blanc de (256 x 256) pixels, un SNR de 27,2 DB pour une réduction de 1/4.

**3. Méthodes proposées**

Nous avons remarqué précédemment que les différentes techniques connues sont difficilement applicables dans le cas d'un modèle évolutif 2D, le nombre de calculs devenant beaucoup trop important. Nous proposons deux méthodes sous optimales différentes, qui nous permettent d'estimer l'excitation de notre modèle.

**- Méthode basée sur les résiduels :**

La première est une méthode rapide, qui demande très peu de calculs, dans laquelle nous déterminons, en une seule fois, toutes les positions des pulses, et ensuite leur amplitude. Le placement des impulsions est effectué par seuillage sur les résiduels  $r(n, m)$ , ou erreur de prédiction à un pas (2). En effet, ces résiduels représentent l'information manquante au modèle, information qu'il nous faut restituer. Le modèle correspond à la texture de l'image, les résiduels représentent les contours des formes (figure 2). Il semble ainsi naturel de placer les entrées à l'endroit de ces contours, c'est-à-dire au maximum de  $r(n, m)$ .

Sur chaque carré (16x16) d'une image noir et blanc (256x256), 42 pulses ont ainsi été positionnées par seuillage du signal  $r(n, m)$ . Leur amplitude est ensuite estimée par minimisation du critère quadratique de l'erreur. L'image reconstituée est de bonne qualité visuelle (figure 3), le SNR étant de 27,96 db pour un taux de compression de 1/6. L'avantage principal de cette méthode est sa simplicité et sa rapidité.

**- Méthode d'ATAL simplifiée :**

La technique de positionnement proposée par ATAL requiert le calcul, en tout point de la fenêtre, de l'intercorrélation de la réponse impulsionnelle avec le signal d'erreur sur la corrélation de la réponse impulsionnelle (5), l'emplacement choisi étant celui donnant un rapport maximum. Comme nous l'avons fait remarquer, en non stationnaire, il n'est pas possible d'utiliser cette méthode trop coûteuse en temps de calculs. Aussi nous proposons pour positionner l'impulsion  $m+1$ , de ne calculer ce rapport que pour des entrées situées dans un voisinage du maximum de l'erreur  $e_t$ , et non sur la fenêtre entière.

Ainsi le domaine de recherche du maximum est restreint d'un facteur 7, la dimension des fenêtres de travail étant pour cet algorithme de (8x8), et le nombre de points testés de 9.

Des essais ont montré la validité de cette méthode sous-optimale, grâce à laquelle nous avons reconstitué l'image de la figure 4, dont le SNR est de 28,43 DB. Le taux de compression atteint est de 1/10.

Ce très bon résultat est dû au fait que cet algorithme, contrairement au précédent,



a l'avantage d'offrir une meilleure répartition des entrées sur l'image. En effet, lors de la reconstruction, les impulsions étant calculées séquentiellement, il est possible d'arrêter la procédure de recherche des entrées sur un carré lorsque l'image reconstituée est bonne. Ceci se fait par un simple test de la valeur du SNR sur le carré étudié par rapport à une valeur de référence. Ainsi les régions homogènes de l'image, bien reconstituée par le modèle, ne comportent que très peu d'entrées, et les parties à fortes variations d'intensité, plus difficiles à reconstruire, peuvent bénéficier de plus de pulses.

C'est un réel avantage puisque l'image reconstruite par cette technique (figure 4) est de qualité équivalente à celle de la figure 3 (28,43 DB contre 27,96 DB), mais pour un taux de compression bien meilleur : 1/10, au de 1/6.

### CONCLUSION

La méthode simplifiée d'ATAL permet d'obtenir, avant codage, la meilleure compression d'informations pour des images noir et blanc digitalisées. Cependant, malgré les simplifications apportées, elle reste lourde et le calcul des entrées est très long.

La méthode basée sur les résiduels est par contre beaucoup plus rapide. On peut donc espérer qu'une méthode basée sur les résiduels, et rendue adaptative sur le nombre d'entrées par fenêtre, donne de bons résultats. Celle-ci allierait les qualités des deux algorithmes proposés : rapidité, et réduction maximum du nombre d'entrées nécessaires à la reconstruction de l'image. Cette méthode est actuellement à l'étude.

### REFERENCES

- [1] M. KAUFMAN, J.W WOODS, S. DRAVIDA, M. TEKALP - "Estimation and identification of two dimensional images". IEEE Trans. on automatic control, vol AC 28 n°7 July 1983.
- [2] C. HORNE, K. JAINANDUNSING, E.F DEPRETERRE - "Image coding based on 2D linear prédiction and 2D multipulse excitation". EUSIPCO LA HAYE 1986.
- [3] G. ALENGRIN, M. BARLAUD, J. MENEZ - "Unbiased parameter estimation of non stationary signals in noise" - IEEE ASSI October 1986.
- [4] R. CHARBONNIER, M. BARLAUD, G. ALENGRIN, J. MENEZ, J. ZERUBIA - "Identification methods for non-stationary signals" - EUSIPCO, LA HAYE 1986.
- [5] J.W WOODS, S.D O'NEIL - "Subband coding of images". IEEE Signal Processing vol ASSP 34, n°5 October 1986.
- [6] B.S ATAL and J.R REMDE - "A new modele of LPC excitation for producing natural sounding speech at low bit rates". Proc. IEEE Int. Conf. Acoust. Speech and Signal Processing, pp 614-617 (April 1982)
- [7] M.C OMNES CHEVALIER, Y. GRENIER, G. CHOLLET - "Codage multi-impulsionnel pour la restauration de parole par des modèles évolutifs". GRETSI -Mai 85, pp 887,892.

IMAGE  
ORIGINALE



FIG.2

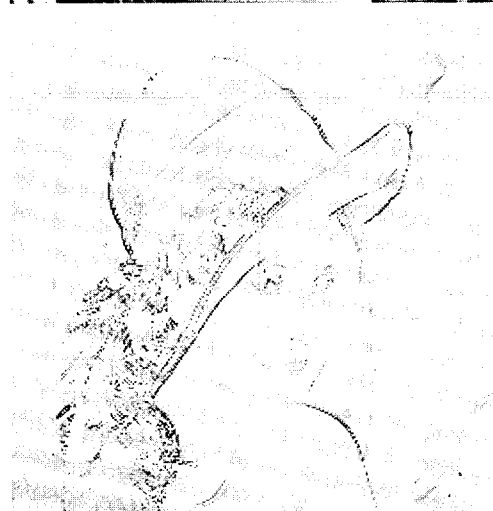


FIG.3



FIG.4

