

Classification neuronale des fonds marins par modélisation AutoRégressive bidimensionnelle

H. Thomas⁽¹⁾, C. Collet⁽¹⁾, G. Burel⁽²⁾ et K. Yao⁽¹⁾

⁽¹⁾Bâtiment des Laboratoires, GTS, Ecole-Navale, Lanvéoc-Poulmic,
29 240 Brest Naval, France
Email : < name >@ecole-navale.fr

⁽²⁾LEST-UBO, 6 av. V. Le Gorgeu, BP 809,
29285 Brest Cedex, France
Email : burel@lest-gw.univ-brest.fr

RÉSUMÉ

Cette étude présente une méthode simple et efficace de caractérisation de textures couplée à un modèle neuronal dans le but d'obtenir une classification des fonds marins en imagerie Sonar. Un modèle autorégressif (AR) linéaire est appliqué à l'image Sonar pour en extraire des attributs caractéristiques significatifs. Les performances de la méthode autorégressive sont évaluées à l'aide d'un classifieur neuronal, de type Perceptron MultiCouche (PMC), qui distingue quatre types de fonds : cailloux, dunes, rides de sable et sable. Les bons taux de classification rencontrés pour chaque type de fond sur une base distincte de la base d'apprentissage sont présentés. Une originalité de l'approche réside dans le couplage performant obtenu entre modélisation AR et module neuronal pour l'application concernée.

ABSTRACT

This paper presents a simple and efficient method of texture characterization with neural classifier to recognize sidescan sonar images. A linear AutoRegressive (AR) model is applied to the sonar images in order to extract significant features. The AR modeling performances are evaluated through a neural classifier, a MultiLayer Perceptron (MLP), which is able to recognize four seafloors: stones, dunes, ripples and sand. Good results in classification have been observed on a set of real images : some of them are presented in this article.

An originality of this approach is due to the association between AR model and neural classifier.

1 Introduction

Le sonar est un outil privilégié pour obtenir une représentation des fonds marins. La cartographie des fonds marins et l'aide à la détection et/ou reconnaissance d'objets reposant sur les fonds sont deux applications de l'étude menée sur la classification des fonds en imagerie Sonar. L'image Sonar fait apparaître des zones dites *d'ombre* qui traduisent l'absence de signal réverbéré et des zones dites *de fond* qui caractérisent la réverbération du signal sur le fond marin. Dans cette étude, nous développons un système de classification supervisée d'images Sonar permettant la distinction entre quatre types de fonds marins : cailloux, dunes, rides de sable et sable (Fig1).

La procédure de classification se décompose en deux étapes principales : l'extraction d'attributs pertinents et la classification de motifs. La constitution d'une base d'apprentissage robuste est également un point important dans le cas d'un algorithme de classification neuronale supervisée. Etant donnée l'importance de l'étape de caractérisation de textures, nos efforts se sont portés sur les méthodes d'analyse de textures. Parmi les méthodes d'analyse de texture les plus communément employées émergent les méthodes statistiques. Haralick[6] présente en 1973 la méthode des matrices de cooccurrence et suggère quatorze paramètres pouvant être extraits de ces matrices caractérisant les dépendances spatiales de niveaux de gris. Galloway[4] explicite la méthode des longueurs de plage en 1975. De bonnes synthèses sur les mé-

thodes statistiques se trouvent dans [5] et [12]. Depuis, d'autres auteurs ont utilisé une combinaison de plusieurs méthodes pour classifier des images Sonar[10]. L'inconvénient de ces méthodes, outre le choix empirique d'une sélection de paramètres les plus représentatifs, est leur incapacité à générer des images à partir des paramètres extraits. Il s'ensuit ainsi une très vraisemblable perte d'informations lors du passage de l'image au vecteur-paramètres.

Dans cette étude nous proposons une méthode de caractérisation d'images Sonar haute résolution simple, efficace, qui n'a encore jamais été à notre connaissance appliquée aux images Sonar et qui permet de synthétiser des images[3]. Il s'agit de la modélisation autorégressive bidimensionnelle dont la formulation est donnée dans [7].

Nous associons à cette méthode d'extraction d'attributs pertinents un classifieur neuronal de type PMC, préféré à un classifieur de type bayésien parce qu'il ne nécessite aucune hypothèse sur les lois de probabilités et parce qu'il offre une grande robustesse face au bruit.

Le paragraphe 2 explicite la méthode d'extraction d'attributs caractéristiques appliquée aux images Sonar.

Le paragraphe 3 décrit la structure et le fonctionnement du classifieur neuronal ainsi que la composition des bases d'apprentissage et de test. Les résultats expérimentaux et leurs analyses sont donnés pour divers vecteurs d'attributs autorégressifs au paragraphe 4. Ce dernier compare aussi l'approche autorégressive avec une autre méthode statistique basée sur les matrices de cooccurrence.

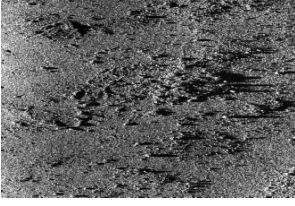


Fig1.a – cailloux



Fig1.b – dunes



Fig1.c – rides de sable

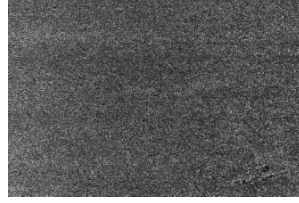


Fig1.d – sable

FIG. 1 — Images représentatives des fonds marins. Un pixel représente une zone de l'ordre de 100cm^2

2 Extraction d'attributs

2.1 Description du modèle AR 2D

Un modèle AR 2D décrit linéairement les dépendances spatiales d'un pixel de l'image avec ses voisins.

Plusieurs types de voisinages sont possibles : causal, semi-causal, non causal... La formation ligne par ligne des images Sonar, l'existence de corrélations entre lignes et entre pixels d'une même ligne ainsi que les dépendances spatiales correspondant à la scène physique suggèrent l'emploi d'un voisinage V non-causal :

$V = \{\text{points de coordonnées relatives } (m, n) \text{ telles que } (m, n) \in [-M \dots + M]^2 \text{ et } (m, n) \neq (0, 0)\}$

Soit $y(k, l)$ un pixel de l'image de taille $K \times L$ ($0 \leq k < K$, $0 \leq l < L$), $y(k, l)$ s'écrit[2][7] :

$$y(k, l) = - \sum_{\substack{m=-M \\ m \neq 0}}^M \sum_{n=-M}^M a(m, n) \cdot y(k-m, l-n) + e(k, l) \quad (1)$$

avec :

$y(k-m, l-n)$: pixel voisin de $y(k, l)$ de coordonnées relatives m et n
 $a(m, n)$: coefficient AR associé au voisin $y(k-m, l-n)$
 $e(k, l)$: erreur de prédiction sur le point $y(k, l)$.

Remarque : la non-considération de l'image sous forme d'un tore induit une contrainte sur l'ensemble des pixels $y(k, l)$ utilisés dans le modèle : $\{y(k, l)\}$ est tel que $0 \leq k-m < K$ et $0 \leq l-n < L$, $\forall (m, n) \in V$.

2.2 Estimation des paramètres

La répartition des luminances des pixels d'une image Sonar n'étant pas gaussienne mais hybride (gaussienne pour l'ombre et de Rayleigh pour les zones réverbérées[9]), une méthode des moindres carrés (méthode de la Pseudo-Inverse) est préférée à la solution du Maximum de Vraisemblance décrite dans [8] qui se place dans un cadre gaussien.

La relation 1 peut s'écrire :

$$y = -Ha + e$$

où y : vecteur des points $y(k, l)$ de l'image

H : matrice des points $y(k-m, l-n)$ de l'image, voisins de $y(k, l)$

a : vecteur des coefficients AR $a(m, n)$

e : vecteur des erreurs de prédiction.

Nous recherchons le vecteur des paramètres a qui minimise $\|y + Ha\|^2$ i.e. la variance de l'erreur de prédiction.

Pour cela, la méthode de la Pseudo-Inverse est utilisée et la solution du système est telle que :

$$H^T y = -H^T H a$$

La matrice $(H^T H)$ est une matrice carrée de sorte que le système est un système de type "n équations n inconnues" dont la solution est obtenue par une méthode de Gauss ou de Gauss-Jordan.

2.3 Sélection des paramètres

Soient \mathcal{A} l'ensemble des paramètres $a(m, n)$:

$\mathcal{A} = \{a(m, n) \text{ tels que } (m, n) \in [-M \dots + M]^2 - \{(0, 0)\}\}$
 et ξ la variance de l'erreur de prédiction donnée par :

$$\xi = \frac{1}{(K-2M) \times (L-2N)} \sum_{k,l} e^2(k, l)$$

Une série d'analyses conduites sur des vecteurs-paramètres extraits d'images Sonar de taille 64×64 pixels a guidé le choix du vecteur d'attributs optimal.

• Taille du voisinage :

L'examen de plusieurs vecteurs de paramètres obtenus pour plusieurs tailles de voisinage $(M, N) \in \{(1, 1), (2, 2), (3, 3), (4, 4)\}$ a montré que le voisinage devait être au moins d'ordre 2 $((M, N) = (2, 2))$ pour que la distinction entre les classes soit possible. Le nombre de paramètres AR à prendre en compte dans le cas d'un voisinage d'ordre 4 $((M, N) = (4, 4))$ est considérable. Par souci de garder un nombre de paramètres raisonnable tout en conservant assez d'information, nous avons retenu les deux voisinages $(2, 2)$ et $(3, 3)$ i.e de taille 5×5 pixels et 7×7 pixels.

• Symétrie :

Nous avons constaté une symétrie quasi-totale des paramètres autorégressifs par rapport au paramètre de coordonnées

$(m, n) = (0, 0)$ pour toutes les images Sonar et pour tous les voisinages V considérés : $a(m, n) \approx a(-m, -n)$.

La prise en compte de cette caractéristique permet de réduire de moitié le nombre de paramètres AR et \mathcal{A} devient \mathcal{A}' tel que :

$$\mathcal{A}' = \mathcal{A} - \{a(-m, -n)\}$$

• Intérêt de la variance de l'erreur de prédiction :

Fig2 représente la variance de l'erreur de prédiction pour 280 images réparties en quatre classes équiprobables.

L'examen de la figure montre que ce paramètre est très discriminant pour les trois classes "cailloux", "dunes" et "sable" dans la mesure où les courbes caractérisant chaque classe ne se chevauchent pas. La distinction de la courbe "rides de sable" est moins évidente.

Devant le caractère discriminant manifeste de ce paramètre, celui-ci est ajouté à l'ensemble \mathcal{A}' devenant \mathcal{A}'' tel que :

$$\mathcal{A}'' = \mathcal{A}' + \{\xi\}$$

Les quatre ensembles de paramètres \mathcal{A} , \mathcal{A}' , \mathcal{A}'' et $\mathcal{A}''' = \{\xi\}$ constituent les vecteurs caractéristiques de l'étude. Leurs performances sont évaluées au 4^e paragraphe.

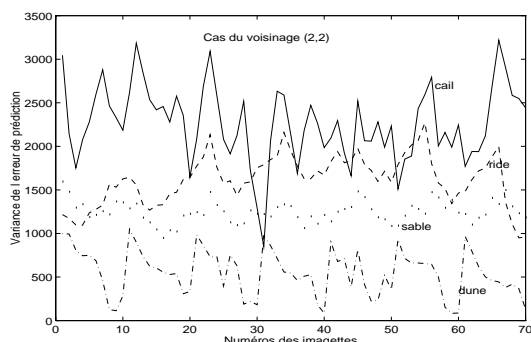


FIG. 2 — Valeurs de la variance de prédiction pour 280 images réparties en 4 classes, dans le cas d'un voisinage (2,2), i.e de taille 5×5 pixels

3 Classification neuronale supervisée

L'étape de classification a pour but d'affecter une classe à une image caractérisée par un ensemble de paramètres.

Notre choix s'est porté sur un classifieur neuronal supervisé de type Perceptron MultiCouche entraîné par un algorithme de rétropropagation du gradient[1], pour ses performances en classification d'images fortement bruitées[10].

- *Structure du PMC :*

Le PMC est un réseau de neurones à couches dont les nombres de neurones constituant respectivement la couche d'entrée et la couche de sortie sont égaux au nombre d'attributs caractéristiques et au nombre de classes à distinguer. Le nombre de couches cachées et de neurones par couche cachée est variable. Les meilleurs taux de reconnaissance pour cette application sont obtenus avec une seule couche cachée contenant 10 neurones cachés.

- *Fonctionnement :*

Un PMC fonctionne en deux temps. La phase d'apprentissage est fondamentale. Elle consiste à faire apprendre au réseau les différentes configurations des vecteurs de paramètres caractéristiques des images de fonds que l'on souhaite classifier. Pour cela, on utilise la règle usuelle de rétropropagation du gradient de l'erreur (erreur encourue entre la réponse calculée par le réseau et la réponse désirée qui est connue). La phase de test ou de généralisation consiste à présenter un vecteur de paramètres inconnu (*i.e.* que le réseau n'a pas appris) en entrée du réseau et à observer la réponse qu'il fournit (*i.e.* le type de fond reconnu).

- *Bases d'apprentissage et de test :*

La base d'apprentissage doit être représentative des différents fonds que l'on souhaite reconnaître et significative (types de fonds + différents aspects pour un fond donné). Ainsi sont intégrés dans cette base d'apprentissage des échantillons distincts entre eux et représentatifs des quatre fonds à distinguer. Les échantillons de la base de test sont distincts des échantillons de la base d'apprentissage.

4 Etude expérimentale

Une série de tests a été menée pour :1) vérifier la symétrie des paramètres ; 2) étudier l'influence de la taille du voisinage ;

3) vérifier l'intérêt de l'attribut variance de l'erreur de prédiction ; 4) étudier la pertinence des paramètres retenus et 5) comparer la méthode autorégressive avec une autre méthode statistique basée sur les matrices de cooccurrence.

Les bases d'apprentissage et de test se composent respectivement de 301 et 288 échantillons qui sont tous distincts. Chaque type de fond y est représenté de manière quasi équiprobable et chaque échantillon-image a une taille de 64×64 pixels. Pour l'application, nous avons utilisé le module neuronal du logiciel de traitement d'images *Allegory* développé par THOMSON. Pour chaque série de test, nous faisons varier deux paramètres de configuration du réseau de neurones : le nombre de neurones dans la couche cachée (unique) et le nombre de présentations de l'ensemble de la base d'apprentissage devant le réseau. Les résultats sont donnés pour le nombre de neurones et pour le nombre de présentations optimaux (ceux-ci correspondent à un taux de généralisation global maximal pour le test donné).

- *Symétrie des paramètres :*

Le tableau1 donne les taux de reconnaissance par classe (en %) obtenus pour les deux tailles de voisinage (2,2) et (3,3) et pour les ensembles \mathcal{A} et \mathcal{A}' .

La réduction de \mathcal{A} en \mathcal{A}' ne modifie quasiment pas les taux de reconnaissance. Toute l'information prise en compte par le réseau est contenue dans \mathcal{A}' .

classes	cailloux	dunes	rides	sable	global
voisinage (2,2)	79	94	98	96	92
1/2 voisinage (2,2)	79	92	96	98	91
1/2 voisinage (3,3)	90	98	98	100	96
voisinage (3,3)	92	98	98	100	97

TAB. 1 — Taux de reconnaissance (en %) pour 4 tailles de voisinage. Mise en évidence de la symétrie des paramètres.

- *Taille du voisinage :*

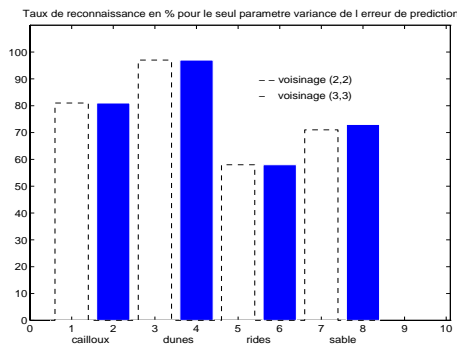
TAB.1 met également en évidence la supériorité d'un voisinage (3,3) par rapport à un voisinage (2,2). Les résultats pour un voisinage (3,3) et surtout pour un voisinage (2,2) montrent une petite faiblesse du système à reconnaître les fonds de type cailloux. Celui-ci est souvent confondu avec la classe sable.

- *Variance de l'erreur de prédiction :*

Ce seul paramètre permet de distinguer nettement les classes cailloux et dunes et dans une moindre mesure la classe sable (TAB.2). Par contre cet attribut ne convient pas pour la reconnaissance de la classe rides de sable qui est confondue avec la classe cailloux et surtout la classe sable. Cette constatation est à rapprocher des remarques faites en 2.3. Les taux de reconnaissance par classe diffèrent très peu d'une configuration de voisinage à une autre.

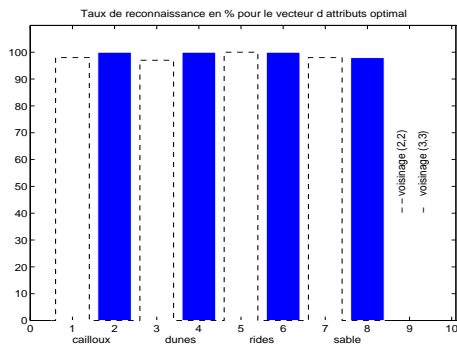
- *Performance du vecteur \mathcal{A}'' :*

L'ensemble \mathcal{A}'' est le vecteur de paramètres qui donne les meilleurs résultats. TAB.3 donne les taux de reconnaissance obtenus sur la base de test pour chaque classe et pour les deux voisinages. Le bon pouvoir de discrimination du paramètre variance de l'erreur de prédiction pour la classe cailloux a permis de rehausser le taux de reconnaissance des images de ce type.



TAB. 2 — Taux de reconnaissance (en %) pour les voisinages (2,2) et (3,3) en considérant le seul paramètre variance de l'erreur de prédiction

Les taux de reconnaissance diffèrent peu d'une configuration de voisinage à une autre.



TAB. 3 — Taux de reconnaissance (en %) pour les voisinages (2,2) et (3,3) en considérant le vecteur de paramètres \mathcal{A}''

• Comparaison avec une autre méthode statistique :

Tab.4 donne les résultats obtenus avec une autre méthode d'extraction d'attributs : les matrices de cooccurrence. Les paramètres utilisés sont décrits dans [11][6]. Les mêmes bases de test et d'apprentissage sont utilisées pour la comparaison. Les taux de reconnaissance sont sensiblement meilleurs avec une modélisation AR de l'image. La méthode des matrices de cooccurrence présente certes l'avantage d'être invariante en rotation mais outre ses moins bonnes performances en classification, elle est beaucoup plus coûteuse en temps de calcul.

classes	cailloux	dunes	rides	sable	global
matrices de cooccurrence	100	95	97	95	97
modèle AR	100	100	100	98	99

TAB. 4 — Comparaison des taux de reconnaissance (en %) obtenus par la méthode des matrices de cooccurrence et par la modélisation AR pour un voisinage (3,3) en considérant le vecteur de paramètres \mathcal{A}'' .

5 Conclusion

Dans cette étude¹, nous avons développé une chaîne de classification d'images Sonar efficace et robuste.

Elle consiste : 1) à extraire des attributs pertinents grâce à l'application d'un modèle AR 2D sur l'image et 2) à utiliser un classifieur neuronal pour reconnaître quatre types de fonds marins : cailloux, dunes, rides de sable et sable.

Cette méthode est efficace parce que les taux de reconnaissance sont particulièrement bons : cailloux : 100%, dunes : 100%, rides de sable : 100% et sable : 98%. Elle est originale puisqu'à notre connaissance la modélisation AR 2D associée à un classifieur neuronal pour reconnaître des fonds marins sur images Sonar n'avait pas été jusqu'alors expérimentée. Toutefois, la modélisation AR 2D d'images Sonar ne fournit pas sous sa forme actuelle des paramètres invariants en rotation *i.e.* une image de rides de sable, par exemple, d'orientation très différente de celle apprise par le réseau, peut ne pas être reconnue comme appartenant à cette classe. Des études sont actuellement en cours pour pallier cet inconvénient.

Références

- [1] G. BUREL. *Réseaux de neurones en traitement d'images : des modèles théoriques aux applications industrielles*. PhD thesis, Université de Bretagne Occidentale, décembre 1991.
- [2] R. CHELLAPPA and R.L. KASHYAP. Statistical inference in gaussian markov random field models. *IEEE Computer Society Conference on Pattern Recognition and Image Processing*, pages 77–80, june 1982.
- [3] R. CHELLAPPA and R.L. KASHYAP. Texture synthesis using 2D noncausal autoregressive models. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-33(1) :194–203, february 1985.
- [4] M.M. GALLOWAY. Texture analysis using gray level run lengths. *Computer Graphics Image processing*, 4(2) :172–199, june 1975.
- [5] R.M. HARALICK. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), may 1979.
- [6] R.M. HARALICK, K. SHANMUGAM, and ITS'HAK DIENSTEIN. Textural features for image classification. *IEEE Trans. Syst., Man, Cybern.*, SMC-3 :610–621, november 1973.
- [7] R.L. KASHYAP. Random field models on torus lattices for finite images. *IEEE Proceedings of the 5th international Conference on Pattern Recognition*, pages 1103–5, december 1980.
- [8] R.L. KASHYAP, R. CHELLAPPA, and A. KHOTANZAD. Texture classification using features derived from random field models. *Pattern Recognition Letters*, 1(1) :43–50, october 1982.
- [9] F. SCHMITT, M. MIGNOTTE, C. COLLET, and P. THOUREL. Estimation of noise parameters on sonar images. *Signal and Image processing, SPIE'96*, august 1996.
- [10] W.K. STEWART, M. JIANG, and M. MARRA. A neural network approach to classification of sidescan sonar imagery from a midocean ridge area. *IEEE Journal of Oceanic Engineering*, 19(2), april 1994.
- [11] H. THOMAS. Détermination d'attributs et classification de fonds par réseaux de neurones en imagerie sonar. Technical Report 5, ECOLE NAVALE, juillet 1995.
- [12] J.S. WESZKA, C.R. DYER, and A. ROSENFELD. A comparative study of texture measures for terrain classification. *IEEE Trans. Syst. Man Cybern.*, SMC-6(4) :269–285, april 1976.

¹Les auteurs remercient le GESMA pour la fourniture des images Sonar ainsi que pour le financement partiel de cette étude.