

Fusion multi-attributs d'images sonar

Isabelle QUIDU¹, Jean-Philippe MALKASSE¹, Pierre VILBÉ², Gilles BUREL²

¹Thomson Marconi Sonar

Route de Sainte Anne du Portzic, 29601 BREST cédex

isabelle.quidu@tms.thomson-csf.com, jean-philippe.malkasse@tms.thomson-csf.com

²L.E.S.T. - FRE CNRS 2269, Université de Bretagne Occidentale

6, avenue Le Gorgeu, BP 809, 29285 BREST cédex

Pierre.Vilbe@univ-brest.fr, Gilles.Burel@univ-brest.fr

Résumé – L'utilisation de sonar haute résolution pour la guerre des mines permet d'envisager la classification d'un objet immergé à partir de l'ombre acoustique portée sur le fond. Le levé des ambiguïtés inhérentes à la classification monovue peut être réalisé en utilisant plusieurs images présentant les ombres acquises selon des points de vue distincts. Nous proposons ainsi d'effectuer une classification supervisée multi-vues par fusion des attributs caractéristiques extraits d'images sonar constituant trois points de vue imprécis (angles de vue inconnus) d'une même cible. La méthode repose sur l'analyse de la répartition d'attributs caractéristiques extraits d'images sonar haute résolution dans le cadre de la théorie des ensembles flous et la définition de modèles de fusion adaptés. L'évaluation de ces modèles a été réalisée sur une base d'images simulées avec cinq types de mines.

Abstract – Using high resolution sonar for mine hunting enables to classify objects laying on the seabed from their cast shadow. A solution to alleviate ambiguities inherent in a single view process consists in using several images emanating from different points of view. We propose to perform a supervised multiview classification by fusing features computed over three sonar images showing imprecise points of view of the same object. Our data fusion scheme focused on the analysis of features is based on the theory of fuzzy sets and aims at providing suitable fusion models. These models have been tested over a simulated database including five types of mines.

1. Introduction

Dans le contexte de la guerre des mines, les mines détectées peuvent être classifiées à partir de leur ombre portée. L'image obtenue au moyen de sonars actifs à haute résolution, c'est-à-dire dont les dimensions de la cellule de résolution en distance et en gisement sont inférieures aux dimensions des objets recherchés, fait apparaître l'ombre acoustique que l'objet porte sur le fond. Le processus de classification consiste en trois étapes : la segmentation, l'extraction des attributs et enfin la classification. La classification sur ombre est un problème classique de reconnaissance de formes. La notion de forme englobe la région intérieure de l'ombre segmentée et son contour. La qualité actuelle des images sonar haute résolution permet de caractériser finement les ombres segmentées en calculant L attributs ou critères discriminants (composantes du vecteur caractéristique) en vue de l'étape finale de classification parmi K classes possibles.

Classifier en mono-vue suppose que l'on se situe dans un cas favorable (fond horizontal, faible relief, ombre révélatrice de l'objet ciblé...). En réalité, un certain nombre d'ambiguïtés peuvent subsister et rendent parfois la classification mono-vue non pertinente. On propose alors de traiter ces **incertitudes** en **fusionnant** les informations ou attributs extraits de N images I_n présentant les ombres acquises selon des points de vue distincts c'est-à-dire effectuer une **caractérisation multi-vues**. Ces données sont en outre **imprécises** puisque les angles de prise de vue sont quelconques. La **théorie des ensembles flous** permet de représenter les classes (caractérisées par les attributs) par des ensembles flous et d'associer à tout élément inconnu un degré d'appartenance à

chacune d'entre elles. Quel que soit l'attribut considéré, cette donnée simple est aisément interprétable. L'article vise à confronter différents modèles flous destinés à la fusion multi-attributs pour la classification robuste de mines. Le paragraphe 2 situe le problème de la fusion d'informations dans le cadre de la classification sur ombre. Les différents modèles testés sont décrits dans le paragraphe 3. Le paragraphe 4 donne l'ensemble des résultats expérimentaux.

2. Fusion de données en imagerie sonar

La fusion de données est l'utilisation conjointe des données pour l'aide à la prise de décision. Elle vise à traiter des informations imprécises (défaut quantitatif de connaissance) et incertaines (conformité à la réalité) [1][2]. En ce qui concerne le problème de classification multi-vues sur ombre, le principe de fusion apporte de la robustesse et renforce la crédibilité des résultats en gérant l'imprécision liée à la prise de vue de même que les incertitudes liées aux irrégularités du fond.

On dispose de N images représentant l'ombre portée du même objet avec des angles de vue distincts. On note f , l'ombre portée (forme) sur laquelle porte la décision. La décision consiste à lui affecter la classe C_k parmi les $K=5$ classes d'objets recherchés, soient 'cylindre', 'sphère', 'mine Manta', 'mine Sigel' et 'mine Rockan'. La décision sur f est prise à partir d'informations $a_i(f)_{i \in \{1, \dots, L\}}$ que constituent nos L attributs. La fusion s'appuie sur ces informations par l'intermédiaire de **mesures** $M_k^n(\mathcal{O})_{(k,n) \in \{1, \dots, K\} \times \{1, \dots, N\}}$, reliant les attributs sur f extraits de l'image I_n à une décision potentielle C_k sur f . Le **modèle de fusion** nécessite de disposer simultanément de toutes les informations recueillies à partir

de N images. La mesure fusionnée relative à la décision C_k est obtenue grâce à un **opérateur de fusion F** permettant d'affecter l'objet ciblé de forme f à la classe C_k si :

$$M_k(f) = \max_{l \in \{1, \dots, K\}} \{M_l(f)\} \text{ où } M_l(f) = F(M_l^n(f), 1 \leq n \leq N)$$

3. Fusion floue [2][3]

Les informations de haut niveau que nous manipulons sont les **attributs caractéristiques** calculés sur chacune des ombres segmentées des images. Dans le contexte de la théorie des ensembles flous, l'ensemble des exemples, ou configurations possibles d'observation pour l'objet de la classe C_k , déterminent un sous-ensemble flou auquel on peut associer une fonction d'appartenance μ_k [4]. Le degré d'appartenance mesure la ressemblance d'un élément avec l'ensemble flou en induisant l'incertitude concernant le concept « élément de classe C_k ». L'intérêt des ensembles flous est de pouvoir assigner plusieurs classes à un individu donné avec des degrés d'appartenance différents et de repousser ainsi l'étape ultime de décision.

La **modélisation** nous permet de passer des informations extraites des images à une représentation mathématique liée à la théorie des ensembles flous. Le modèle permet d'affecter à chaque image courante I_n de la base de généralisation (élément inconnu à classifier) une mesure $M_k^n(f)$ à la classe C_k via une distance liant ses attributs a_i^n aux attributs a_i^k des exemples de la classe C_k de la base d'apprentissage ou d'entraînement. Il est alors étroitement lié à la caractérisation de l'ensemble flou suivant que sa configuration permet ou non de le réduire à un *élément idéal*. La pertinence de la distance associée au modèle dépend également de la distribution des nuages de points des ensembles flous pour l'application courante.

Compte tenu des données à fusionner et de la procédure d'acquisition de ces données, notre choix se porte naturellement sur des opérateurs F de compromis correspondant à un comportement prudent [6]. En l'occurrence, l'opérateur de moyenne semble bien adapté à notre problème.

3.1 Par rapprochement à un élément idéal

Si l'on considère que la base est suffisamment importante, on peut en déduire, pour chaque classe, un élément idéal ou prototype. Le degré d'appartenance ou mesure $M_k^n(f)$ à la classe C_k est fonction de sa distance à l'élément idéal caractérisé par une ou plusieurs grandeurs statistiques calculées sur la base d'apprentissage.

3.1.1 Mesure vectorielle

On désire rapprocher l'image courante I_n de vecteur caractéristique $V_n = (a_i^n)_{i \in \{1, \dots, L\}}$ de chacune des classes C_k pour $k \in \{1, \dots, K\}$ caractérisées par leur élément idéal $\bar{V}_k = (\bar{a}_i^k)_{i \in \{1, \dots, L\}}$ et la matrice de covariance Σ_k calculés à partir de la base d'apprentissage de N_k éléments. Le degré d'appartenance est alors inversement relié à la distance entre

le vecteur caractéristique de l'image et celui de l'élément idéal [7].

Description du modèle

Soit $d(V_n, \bar{V}_k)$, la distance entre l'image courante I_n et l'élément idéal de la classe C_k , le degré d'appartenance de l'image I_n à la classe C_k est défini par :

$$M_k^n = \exp(-d(V_n, \bar{V}_k)) \Rightarrow \text{modèle } m_i$$

où $d = d(V_n, \bar{V}_k) = (V_n - \bar{V}_k) \Sigma_k^{-1} (V_n - \bar{V}_k)^T$ est la distance de Mahalanobis et $\Sigma_k = \left(\sigma_{ij}^k \right)_{i,j=1 \dots L}$ la matrice de covariance.

3.1.2 Mesure multi-attributs

Pour i donné, les attributs a_i calculés pour les objets d'une même classe C_k constituent un sous-ensemble flou dont l'élément idéal est caractérisé par un attribut représentatif pour cette classe C_k . Le degré d'appartenance de l'image courante à une classe est alors la combinaison de L degrés d'appartenance calculés *relativement à chaque attribut a_i* .

Définition de la mesure

$$M_k^n = \frac{1}{L} \sum_{i=1}^L M_k^n \Big|_{a_i}$$

Description des modèles

▪ Premier modèle basé sur la S-fonction de Zadeh [4]

Dans le cadre de la théorie des ensembles flous, de nombreuses fonctions d'appartenance ont été introduites pour représenter explicitement les informations imprécises. On rappelle que l'ensemble des attributs caractéristiques de la base d'images forment un espace caractéristique flou. Le problème revient à associer à chaque sous-espace relatif à un attribut donné d'une classe donnée a_i^k une fonction d'appartenance $\mu_i^k(n)$ et de l'identifier simplement à la mesure recherchée $M_k^n \Big|_{a_i}$. La S-fonction de Zadeh permet de réaliser un seuillage progressif suivant la valeur de l'attribut en posant, pour $x > 0$:

$$S(x, \alpha, \beta, \gamma) = \begin{cases} 0.0 & \text{si } x < \alpha \\ 1.0 & \text{si } x \geq \gamma \\ 2(x - \alpha)^2 / (\gamma - \alpha)^2 & \text{si } \alpha \leq x < \beta \\ 1 - 2(x - \gamma)^2 / (\gamma - \alpha)^2 & \text{si } \beta \leq x < \gamma \end{cases}$$

$\Delta\beta = \gamma - \alpha$ est la *zone floue* et β le *point de croisement*.

Afin de déterminer le degré de compatibilité ou d'appartenance de l'attribut courant a_i^n à l'attribut moyen \bar{a}_i^k associé à l'élément idéal, on définit une application composée :

$$\mu_i^k(a_i^n) = 1 - S\left(\frac{|a_i^n - \bar{a}_i^k|}{\sigma_i^k}, \alpha, \beta, \gamma\right) \in [0, 1]$$

et
$$M_k^n \Big|_{a_i} = \mu_i^k(a_i^n) \Rightarrow \text{modèle } m_2$$

$$m(d(k)) = [1/d(k)] / \left[\sum_{l=1}^K 1/d(l) \right]$$

Cette mesure a l'intérêt de relier la notion de frontière floue d'un ensemble donné à ses grandeurs statistiques (moyenne et écart-type) en autorisant une marge $\Delta\beta$ plus ou moins grande autour de la valeur moyenne (ici, $(\alpha, \gamma) = (1, 1.5)$).

▪ **Second modèle basé sur les histogrammes**

A partir de l'histogramme ou répartition des valeurs prises par un attribut a_i donné, on forme la fonction de répartition F_H associée à l'histogramme à partir des valeurs cumulées :

$$F_H^k(x) = N_b^k(x) / N_{b_{tot}}^k$$

soit le nombre d'exemples $N_b^k(x)$ de la classe C_k dont la valeur de l'attribut est inférieure ou égale à x rapporté au nombre total d'exemples $N_{b_{tot}}^k$ de cette même classe.

Soit $med(a_i^k)$, la valeur médiane caractérisant l'élément idéal. L'objectif étant de comptabiliser le nombre d'exemples N_k' de la base d'apprentissage pour la classe C_k dont la valeur de l'attribut a_i^k est plus proche de la valeur optimale $med(a_i^k)$ que la valeur courante a_i^n :

$$N_k' = |1 - 2F_H^k(a_i^n)| \times N_{b_{tot}}^k$$

Cependant, le recouvrement entre les classes peut être important et doit être pris en compte. Pour ce faire, il nous faut connaître, pour chaque classe $C_k \neq C_{k'}$, le nombre d'exemples $N_{k'}'$ dont la valeur de l'attribut est comprise entre les bornes a_i^n et $med(a_i^k)$ soit :

$$N_{k'}' = |F_H^{k'}(a_i^n) - F_H^{k'}(med(a_i^k))| \times N_{b_{tot}}^{k'}$$

Finalement, le modèle est défini par la mesure :

$$M_k^n \Big|_{a_i} = \max \left\{ 0, 1 - \left(\frac{N_k'}{N_{b_{tot}}^k} + \frac{\sum_{k', k' \neq k} N_{k'}'}{\sum_{k', k' \neq k} N_{b_{tot}}^{k'}} \right) \right\} \Rightarrow \text{modèle } m_3$$

3.2 Par rapprochement à l'ensemble flou

Dans certains cas, il est plus prudent d'élaborer la mesure d'appartenance à partir de l'ensemble ou d'une partie des N_k exemples de chacune des classes de la base d'apprentissage.

Définition de la mesure

$$M_k^n = m[d(k)] ; \text{ où } d(k) \text{ est la distance à l'ensemble flou}$$

Description des modèles

Afin de mesurer le degré de compatibilité de l'élément inconnu à la classe C_k , on rapporte la distance $d(k)$ à l'ensemble des distances $d(l)$, $l=1, \dots, K$, soit [5] :

▪ **Distance $d_1 \Rightarrow$ modèle m_4**

Chacun des ensembles flous de la base d'apprentissage associés aux K classes peut être représenté par un nombre limité de points particuliers représentatifs. Ces points sont les centroïdes de $N_C=10$ sous-ensembles ou *clusters* obtenus par la méthode des fuzzy C-means généralisée par Bezdek [8]. La distance à l'ensemble flou représentatif de la classe C_k est alors :

$$d_1(k) = \min_{j \in [1, N_C]} \|V_n - v_j^k\| \begin{matrix} \text{Distance euclidienne minimale} \\ \text{entre l'élément inconnu et chacun} \\ \text{des } N_C \text{ clusters (leurs centroïdes } v_j \\ \text{de la classe } C_k \end{matrix}$$

▪ **Distance $d_2 \Rightarrow$ modèle m_5**

A partir de la fonction de répartition associée à l'histogramme (cf. §3.1.2), on peut définir une nouvelle distance basée sur la configuration des exemples de l'ensemble flou au voisinage de l'attribut courant a_i^n . Ainsi, pour une classe C_k donnée et attribut par attribut, on s'intéresse à la densité ρ_i^k , ou inversement la dispersion, des individus au voisinage de a_i^n :

$$d_2(k) = \sum_{i=1}^L \frac{1}{\rho_i^k(a_i^n)}$$

4. Résultats expérimentaux

Nous avons évalué ces modèles sur une base d'images simulées comptabilisant plus de 5000 prises de vues représentant cinq classes ou types de mines choisis. Ces vues sont prises au cours d'une trajectoire circulaire du sonar autour de l'objet (même altitude, même distance). Nous proposons de classifier les objets à partir de $N=3$ vues seulement dont les angles de prise de vue sont inconnus. En outre, trois tailles différentes d'objets ont été considérées pour les mines cylindriques et sphériques.

Chaque image est normalisée en rasant, segmentée puis caractérisée par quatre attributs hybrides (cf. FIG. 1) : critères géométriques associés à des invariants de forme auparavant utilisés en classification mono-vue [9].

Après un mélange aléatoire de la base, elle est divisée en base d'apprentissage et base de généralisation. Après entraînement, on calcule, pour chacun des modèles, les mesures moyennes $\bar{M}_k = \sum_{objets} M_k$ obtenues par sous-

ensemble d'objets de la base de généralisation. Le TAB. 1 permet de comparer les performances obtenues par modèle.

L'ensemble des modèles présentent de bonnes performances en particulier les modèles m_4 et m_5 . Ces deux derniers modèles font appel à une analyse des données qui s'effectue *hors ligne* c'est-à-dire une fois pour toute en début de traitement. L'information relative à l'ensemble des éléments de chacune des classes de la base d'apprentissage

est ainsi avantageusement résumée. L'analyse des histogrammes est toutefois séduisante par sa simplicité comparée à l'analyse par *clustering*. Quant au modèle m_3 également performant et ce, quelle que soit la classe (contrairement à m_1 et m_2), il exige de prendre en compte le recouvrement entre les histogrammes. Cette précaution garantit de bonnes performances mais au prix d'un temps de calcul plus important.

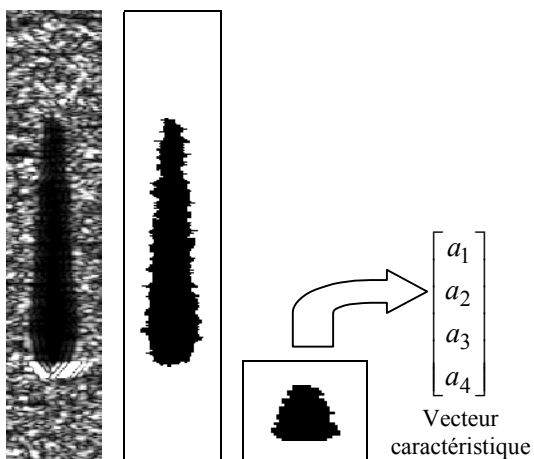


FIG. 1 : Segmentation, normalisation, extraction d'attributs

TAB. 1 : performances en pour-cent

Modèles classes	m_1	m_2	m_3	m_4	m_5
C_1 (cylindre)	99.4	99.9	95.7	100	99.4
C_2 (sphère)	100	99.9	100	100	100
C_3 (Manta)	100	97.7	94.3	100	100
C_4 (Sigeel)	64.7	71.7	89.3	100	100
C_5 (Rockan)	99.3	98.7	95.7	98.7	93
Bilan	92.7	93.6	95	99.7	98.5

La FIG. 2 illustre les mesures \bar{M}_k obtenues pour chaque sous-ensemble d'objets (cylindres, sphères, mines Manta, Sigeel et Rockan) testés de la base de généralisation conditionnellement à la classe C_k en fonction des $K=5$ classes pour les cinq modèles testés. Chaque diagramme comporte donc, pour chaque sous-ensemble, un lot de mesures représentées par des barres colorées dont la hauteur reflète le degré de compatibilité $\in [0,1]$ avec une classe donnée C_k . Outre les performances obtenues en terme de bonne classification, un autre critère apparaît à la vue de ces diagrammes : un modèle est d'autant plus fiable que ses mesures permettent d'éviter toute classification litigieuse. En d'autres termes, on privilégiera un modèle qui associe à un objet donné une mesure maximale (plus forte compatibilité à une classe donnée) bien supérieure aux mesures annexes (compatibilités avec les autres classes). En ce sens, le modèle m_4 est remarquable avec, quelque soit l'objet à classer, une mesure maximale de valeur supérieure à 0.5 et des mesures annexes bien plus faibles.

D'un point de vue opérationnel, ce processus a l'intérêt de pouvoir s'inscrire dans une chaîne de traitement semi-automatique où l'opérateur est alors amené à n'intervenir qu'en cas de litige entre deux classes.

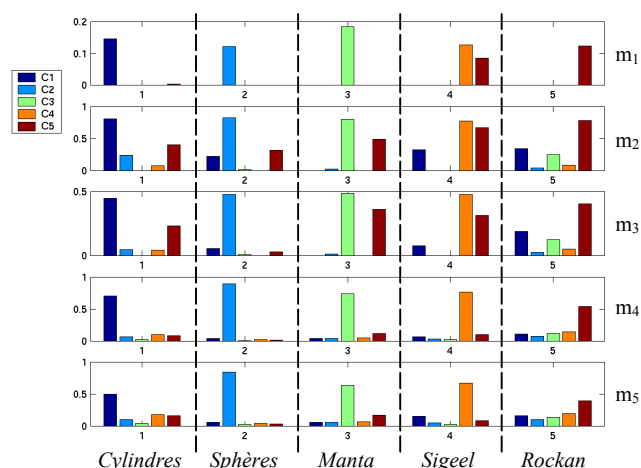


FIG. 2 : mesures \bar{M}_k

5. Conclusion

Dans cet article, nous avons tiré profit de la souplesse des fonctions d'appartenance pour la définition de modèles flous adaptés à la fusion multi-attributs. L'intérêt de la méthode est double. En effet, outre la prise de décision (la classe correspondant à la plus forte mesure est affectée à l'objet ciblé), on y peut associer un degré de confiance : l'ambiguïté entre deux classes est en fait inversement proportionnelle à la différence entre les deux mesures correspondantes.

Références

- [1] I. Bloch, « Fusion de données, ensembles flous et morphologie mathématique en traitement d'images », *Rapport du département Images*, Télécom Paris, 1995.
- [2] I. Bloch, H. Maître, « Fusion de données en traitement d'images : modèles d'information et décisions », *Traitement du signal* 1994, Vol. 11, No. 6, pp. 435-446.
- [3] L.A. Zadeh, « Fuzzy Sets », *Information and Control*, Vol. 8, No. 3, pp. 338-353, june 1965.
- [4] L.A. Zadeh, "Calculus of fuzzy restrictions", in *Fuzzy Sets and Their Applications to Cognitive and Decision Processes*, New York: Acad. Press, Inc, 1975, pp. 1-39.
- [5] S. Daniel, « Fusion multisource appliquée à la reconnaissance d'objets dans le milieu sous-marin », *Thèse de l'université de Rennes I*, novembre 1998.
- [6] I. Bloch, « Information Combination Operators for Data Fusion : A Comparative Review with Classification », *IEEE Trans. on Systems, Man, and Cybernetics – Part A : Systems and Humans*, Vol. 26, No. 1, pp. 52-67, January 1996.
- [7] S.K. Pal, D.K. Dutta Majumder, « Fuzzy mathematical approach to pattern recognition », *John Wiley/Halsted Press*, New York, 1986.
- [8] J. C. Bezdek, « Fuzzy mathematics in pattern classification », *Ph.D. thesis, Cornell University, Ithaca, NY*, 1973.
- [9] I. Quidu, J. Ph. Malkasse, G. Burel and P. Vilbé, « Mine classification using a hybrid set of descriptors », *Proceedings of OCEANS'2000 MTS/IEEE*, Providence, Rhode Island, Sept 2000.