

Approche bayésienne et fonctions de croyance continues pour la classification

Bayesian approach and continuous belief functions for classification

Anthony Fiche Arnaud Martin
ENSIETA, E³I² EA3876

2 rue François Verny 29806 Brest Cedex 9
fichean@ensieta.fr, Arnaud.Martin@ensieta.fr

Résumé :

L'intérêt des fonctions de croyance pour la classification a été démontré ces dernières années avec des fonctions de masse discrètes. Les approches bayésiennes sont cependant performantes lorsque les distributions peuvent être estimées correctement. Récemment des travaux ont montré la possibilité d'étendre les fonctions de croyance discrètes dans le cas continu pour des applications en classification.

Nous proposons dans ce papier d'appliquer et comparer les résultats obtenus par une approche bayésienne et les fonctions de croyance continues et discrètes en vue de caractériser les fonds marins à partir d'images sonar. Les distributions des caractéristiques de l'image sont estimées de la même façon à partir d'un mélange de gaussiennes pour l'approche bayésienne et pour l'approche issue des fonctions de croyance continues.

Mots-clés :

Fonctions de croyance continues, Mélange de gaussiennes, classification.

Abstract:

The theory of belief functions in discrete domain has been employed with success for pattern recognition. However Bayesian performs well provided that once the probability density functions are well estimated. Recently the theory of belief functions has been formalized in continuous domain especially for pattern recognition.

We propose in this paper to apply and compare the obtained results with a Bayesian approach and with the continuous and discrete belief functions to characterizing the seabed from sonar images. The distribution of the image texture parameters are estimated in the same way for Bayesian approach and continuous belief functions approach, based on a Gaussian mixture.

Keywords:

Continuous belief functions, Gaussian mixture, pattern recognition.

1 Introduction

La théorie des fonctions de croyance développée par A. Dempster [3], puis par

G. Shafer [10] a trouvé ces dernières années de nombreuses applications en particulier en classification [4, 7].

Les approches bayésiennes restent cependant performantes dès lors que l'estimation des distributions de probabilité est bonne. La théorie des probabilités est de plus largement employée pour des problèmes d'estimation.

La théorie des fonctions de croyance souvent présentée comme une extension de la théorie des probabilités a pourtant peu été appliquée dans le cas de problèmes d'estimation ou de combinaison de phénomènes continus, même si les bases théoriques ont été proposées dans [16]. Récemment plusieurs travaux ont été proposés pour étendre les bases des fonctions de croyance du domaine discret au domaine continu [15, 5]. En particulier des solutions ont été proposées pour résoudre des problèmes de classification à partir des fonctions de croyance continues [9, 1].

Dans ce dernier cas, comme dans le cas bayésien se pose le problème de l'estimation des distributions des caractéristiques descriptifs des données à classer. En classification, il est courant d'estimer une distribution de probabilité quelconque par un mélange de gaussiennes. Les deux approches vont être comparé dans le cadre de la classification à partir d'images SONAR. Cependant, dans le cas probabiliste, il y aura besoin d'évaluer la proportion de chaque type de sédiment par rapport aux autres, appelée

couramment probabilité *a priori*.

Ces deux approches ont été comparées sur des données générées [1], elles seront comparées ici sur des données réelles avec de plus une approche de classification crédibiliste [4] qui a déjà donné de bonnes performances sur ces données [7]. Ces données sont issues d'un sonar permettant d'imager de vastes étendues sous-marines. La caractérisation des fonds à partir de telles données reste cependant difficile puisque les données sont souvent bruitées.

Dans la section 2 suivante nous rappelons l'approche classique du classifieur bayésien dans laquelle l'estimation des distributions de probabilités repose sur une hypothèse de mélange de gaussiennes. La section 3 rappelle les bases théoriques des fonctions de croyance en discret et en continu et en particulier l'approche proposée dans [1]. La section 4 présente une comparaison des approches en vue de caractériser les fonds marins, avant de conclure.

2 Classification probabiliste

L'objectif de cette partie est de rappeler une démarche classique permettant d'estimer des distributions continues à partir de données puis à partir du modèle construit d'effectuer la classification souvent nommée classifieur bayésien.

2.1 Estimation des distributions

L'estimation d'une distribution continue est un problème difficile en soi. Cependant, on peut approcher toute distribution de probabilité par un mélange de gaussiennes. Toute distribution d'une classe j peut alors s'écrire :

$$p(\mathbf{x}/C_j) = \sum_{k=1}^K W_k^j (2\pi)^{-\frac{d}{2}} \det(\Sigma_k^j)^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}(\mathbf{x} - \mu_k^j)^T (\Sigma_k^j)^{-1} (\mathbf{x} - \mu_k^j)\right\} \quad (1)$$

où \mathbf{x} correspond à une observation de dimension d , C_j représente une classe j possible, K représente le nombre de gaussiennes souhaitées,

W^j les fonctions poids, μ^j les moyennes et Σ^j les matrices de covariances. Dans le cadre de notre application, un nombre de $K=5$ gaussiennes suffit à bien modéliser les distributions.

Plusieurs approches d'optimisation sont envisageables afin d'estimer W^j , μ^j et Σ^j . L'algorithme EM (*Expectation-Maximization*) est une des approches couramment employée et efficace [2].

2.2 Approche bayésienne

Une fois les distributions *a priori* $p(\mathbf{x}/C_j)$ estimées pour chaque classe C_j , le classifieur bayésien nécessite l'estimation des probabilités *a priori* $p(C_j)$. Cette estimation peut se faire facilement à partir de la proportion de données de chaque classe dans la base d'apprentissage. En appliquant le théorème de Bayes, la probabilité *a posteriori* s'obtient par :

$$p(C_j/\mathbf{x}) = \frac{p(\mathbf{x}/C_j)p(C_j)}{\sum_{j=1}^n p(\mathbf{x}/C_j)p(C_j)} \quad (2)$$

où n est le nombre de classes et $p(C_j)$ représente la probabilité *a priori* de la classe C_j .

La décision à partir des probabilités *a posteriori* $p(C_j/\mathbf{x})$ peut se faire de différentes façons. En particulier, le maximum des probabilités *a posteriori* permet de décider la classe probable du vecteur \mathbf{x} .

3 Théorie des fonctions de croyance

La théorie des probabilités fait intervenir la notion de probabilités *a priori* ce qui est gênant lorsque nous ne pouvons pas l'évaluer. La théorie des fonctions de croyance nous permet d'éviter cette notion. De plus, elle permet de modéliser les imprécisions des mesures et elle prend en compte qu'une hypothèse peut être en dehors du cadre d'étude. Cependant, il est possible de faire face à ces problèmes en utilisant les travaux de Dempster [3], repris par Shafer [10], puis par Smets [12, 13] sur la théorie

des fonctions de croyance. Nous la présentons ici dans le domaine discret puis continu en vue de la classification.

3.1 Fonction de croyance discrète

Soit $\Theta = \{C_1, C_2, \dots, C_n\}$ correspondant à l'ensemble des classes possibles. Les fonctions de masse sont définies sur 2^Θ (ensemble de toutes les disjonctions de Θ) et à valeurs dans $[0, 1]$. Par construction, la fonction de masse m vérifie $\sum_{A \in 2^\Theta} m(A) = 1$. Lorsque $m(A) > 0$, on

dit que A est un élément focal. À partir de ces fonctions de masse, il est possible de construire différentes fonctions :

– la fonction de crédibilité, notée *bel*, rassemble les masses de toutes les propositions $B \subseteq A$. Elle indique la confiance totale accordée dans le sous-ensemble A . Elle est déterminée par :

$$bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B) \quad \forall A \in 2^\Theta \quad (3)$$

– la fonction de plausibilité, notée *pl*, indique la confiance maximale accordée à A . Elle est définie par :

$$pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \quad \forall A \in 2^\Theta \quad (4)$$

À partir des fonctions de masse, on peut également calculer la probabilité pignistique [11] utilisée lors de l'étape de décision. En effet, cet opérateur permet d'attribuer une probabilité sur chaque $C_j \in \Theta$. Elle est définie par :

$$BetP(C_j) = \sum_{A \subseteq \Theta, C_j \in A} \frac{m(A)}{|A|(1 - m(\emptyset))} \quad (5)$$

où $|A|$ représente la cardinalité de A . De nombreux travaux ont appliqué ce cadre théorique à des problèmes de classification et proposés des approches afin de définir les fonctions de masse [4, 7].

En particulier le classifieur k -plus proches voisins crédibilistes proposé par [4], est fondé

sur la définition de n fonctions de masse pour chaque observation \mathbf{x} . Ainsi pour chaque classe possible C_j de Θ , on peut définir les fonctions de masse par :

$$\begin{cases} m_i[\mathbf{x}_i, \mathbf{x}](C_j) = \alpha_j e^{\gamma_j d(\mathbf{x}, \mathbf{x}_i)} \\ m_i[\mathbf{x}_i, \mathbf{x}](\Theta) = 1 - \alpha_j e^{\gamma_j d(\mathbf{x}, \mathbf{x}_i)} \end{cases} \quad (6)$$

où $m[A](B)$ est la masse de B sachant A , C_j est la classe associée à \mathbf{x}_i et les vecteurs $\{\mathbf{x}_i\}_{i=1, \dots, k}$ sont les k plus proches voisins de \mathbf{x} . La distance $d(\mathbf{x}, \mathbf{x}_k)$ entre \mathbf{x} et \mathbf{x}_k peut être euclidienne, et $\alpha_j \in [0, 1]$, un coefficient d'affaiblissement, et γ_j , un coefficient de normalisation, peuvent être optimisés. La combinaison de ces fonctions de masse se fait ensuite par la règle orthogonale de Dempster, pour deux fonctions de masse m_1 et m_2 et pour tout $X \in 2^\Theta$:

$$m(X) = \frac{1}{1 - \sum_{Y_1 \cap Y_2 = \emptyset} m_1(Y_1)m_2(Y_2)} \times \sum_{Y_1 \cap Y_2 = X} m_1(Y_1)m_2(Y_2) \quad (7)$$

3.2 Fonction de croyance continue

Smets [15] a récemment proposé les fondements des fonctions de croyance continues. Les fonctions de masse deviennent des densités de masse définies sur des intervalles $[a, b]$ de \mathbb{R} par $m([a, b]) = f(a, b)$, où f est une densité de probabilité de $\{(x, y) \in \mathbb{R}^2 | x \leq y\}$ dans $[0, +\infty]$. Le degré de croyance sur un intervalle $[a, b]$, noté $bel([a, b])$, est déterminé par :

$$bel([a, b]) = \int_{x=a}^{x=b} \int_{y=x}^{y=b} f(x, y) dy dx \quad (8)$$

De même, le degré de plausibilité sur un intervalle $[a, b]$, noté $pl([a, b])$, peut être calculé par :

$$pl([a, b]) = \int_{x=-\infty}^{x=b} \int_{y=ma(x, a)}^{y=+\infty} f(x, y) dy dx \quad (9)$$

À partir d'une densité de probabilité pignistique, il est possible de déterminer une densité

de masse. Cependant, à une même densité de probabilité peut correspondre plusieurs densités de masse, elle est alors dite isopignistique. Pour simplifier le problème, nous cherchons la densité de masse consonante, c'est-à-dire ayant des éléments focaux emboîtés. Il existe alors un index u où les éléments focaux peuvent avoir pour étiquette I_u , vérifiant $I_u \subseteq I_{u'}$ avec $u' > u$ [1]. Cette densité de masse s'obtient par le principe de moindre engagement : plus l'intervalle considéré est petit et plus la confiance attribuée est faible. Smets a appliqué cette méthode pour déterminer la densité de masse pour une densité de probabilité pignistique unimodale. Cette approche a été reprise [1] pour établir une fonction de masse associée à un mélange de gaussiennes.

Ainsi, si l'on approxime une distribution de probabilité multimodale par un mélange de gaussiennes, on peut appliquer l'approche proposée par [1]. L'estimation du mélange de gaussiennes peut se faire comme exposé en section 2.1.

La fonction de plausibilité conditionnelle de la classe C_j au point $\mathbf{x} \in \mathbb{R}^d$, notée $pl[C_j](\mathbf{x}) = pl_j(x)$, est alors définie par [1] :

$$pl_j(\mathbf{x}) = 1 - \sum_{k=1}^K W_k^j \times F_{d+2} \left[(\mathbf{x} - \mu_k^j)^T (\Sigma_k^j)^{-1} (\mathbf{x} - \mu_k^j) \right] \quad (10)$$

avec $F_{d+2}(x) = \int_0^x \frac{u^{\frac{d+2}{2}-1}}{2^{\frac{d+2}{2}} \Gamma(\frac{d+2}{2})} \exp(-\frac{u}{2}) du$ la fonction de répartition de la loi du χ^2 avec $d+2$ degrés de liberté (d représentant la dimension du vecteur \mathbf{x}) et $\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t) dt$ la fonction Gamma.

La fonction de masse dans 2^Θ peut alors être calculée au point \mathbf{x} en utilisant le théorème généralisé de Bayes [14, 15] pour $A \in 2^\Theta$:

$$m[\mathbf{x}](A) = \prod_{C_j \in A} pl_j(\mathbf{x}) \prod_{C_j \in A^c} (1 - pl_j(\mathbf{x})) \quad (11)$$

où $m[\mathbf{x}](A)$ représente la fonction de masse connaissant l'observation \mathbf{x} .

La fonction de masse étant calculée, on peut alors la convertir en probabilité pignistique (5). Enfin, la dernière étape consiste à calculer le maximum de probabilité pignistique pour attribuer la classe la plus crédible à l'observation \mathbf{x} .

4 Application des fonctions de croyance continues à la caractérisation des sédiments marins

Dans cette partie, nous allons appliquer et comparer le classifieur bayésien décrit dans la section 2, le classifieur k -plus proches voisins crédibilistes [4] et le classifieur fondé sur les fonctions de croyance continues présentées dans la section 3.2 sur des données réelles issues d'images sonar en vue de caractériser les sédiments marins.

4.1 Présentation des données

Nous disposons d'une base de données de 42 images sonar fournies par le GESMA (Groupe d'Études Sous-Marines de l'Atlantique) (Figure 1). Ces images ont été acquises en utilisant un sonar latéral de type Klein 5400. Ces images ont subi un prétraitement afin de corriger la variation du gain et de réduire le bruit de chatolement donnant un aspect granulaire aux images sonar. Un expert a étiqueté ces images en indiquant le type de sédiment ainsi que son degré de confiance en distinguant sûr, moyennement sûr et pas sûr [8]. La ca-



Figure 1 – Exemple d'image sonar

ractérisation des sédiments s'effectuant à par-

tir des textures (Figure 1), on peut différencier 5 types de texture : roche, ride, vase, cailloutis et sable. En termes linguistiques, la tex-

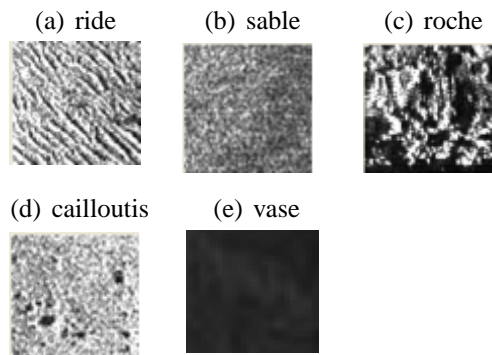


Tableau 1 – Exemples d’imassettes

ture représente la rugosité, le contraste, la finesse, la régularité... Plusieurs approches sont envisageables pour extraire des caractéristiques de texture [7], nous avons retenu ici celle reposant sur les matrices de cooccurrence proposée par [6] permettant d’extraire entre autre le *moment différentiel inverse*, la *corrélation*, le *contraste*, la *moyenne des sommes*, la *moyenne des variances*, l’*entropie des sommes*, l’*entropie des différences* et la *mesure de corrélation*. Pour le calcul de ces caractéristiques, il est nécessaire de découper les images en imassettes de taille 32×32 pixels, équivalent à $6,40m^2$ [7]. Chaque imasette est donc représentée par un vecteur de 8 dimensions. Il est possible de tracer les caractéristiques pour chaque sédiment (ne considérant que les imassettes ne contenant qu’un seul sédiment) en utilisant la méthode des histogrammes. En effet, cette méthode consiste à balayer à pas fixé toutes les données comprises entre une valeur minimale et maximale et d’y compter le nombre d’éléments contenus dans chaque intervalle. Nous présentons un exemple de distributions pour la caractéristique dite de *corrélation* pour tous les sédiments en figure 2. Les différents descripteurs d’Haralick ont été sélectionnés lorsqu’il était possible de discriminer plusieurs classes sur chaque courbe. Une texture décrite par ces 8 caractéristiques suit donc une distribution dans un espace de dimension 8. Chaque distribution

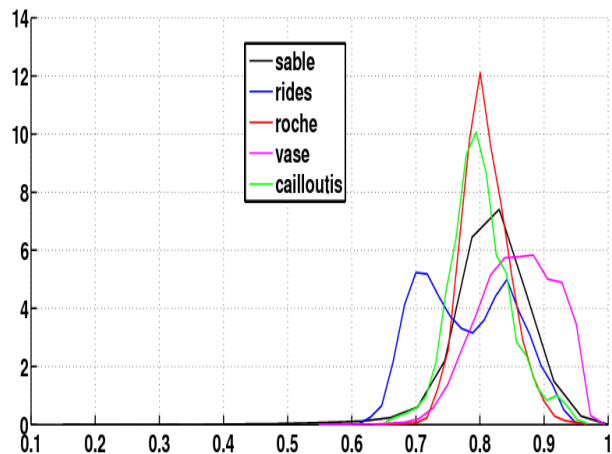


Figure 2 – Distributions de la caractéristique dite de *corrélation* par type de sédiment

peut être approchée par 5 gaussiennes multidimensionnelles suivant l’approche présentée en section 2.

4.2 Résultats et discussion

Afin d’évaluer et comparer les approches nous n’avons considéré que les imassettes avec un seul type de sédiment. À la suite d’un tirage aléatoire nous en avons conservé 5000. Sur ces 5000 imassettes la moitié est destinée à la base d’apprentissage et le reste à la base de test, ce qui fait 2500 vecteurs test. Nous effectuons dix fois cette procédure pour estimer au mieux la matrice de confusion et le taux de bonne classification moyenné. Cette matrice permet d’observer la qualité de la classification : chaque colonne représente la proportion de la classe estimée tandis que chaque ligne indique la classe réelle. Nous définissons donc les 5 classes correspondant aux 5 textures citées : C_1 =roche, C_2 =ride, C_3 =vase, C_4 =cailloutis, C_5 =sable.

La base d’apprentissage va donc permettre l’apprentissage des poids, moyennes et variances des mélanges de gaussiennes de chaque classe à partir de l’algorithme EM pour l’approche bayésienne et les fonctions de croyance continues, ainsi que l’estimation des probabilités *a priori* de chaque classe pour l’approche bayésienne et pour le voisinage des k -plus

proches voisins crédibilistes où $k = 3$ et α et γ sont optimisés [18].

Tout d'abord, nous travaillons de manière mono-dimensionnelle, c'est-à-dire que nous considérons chaque composante (*i.e.* les caractéristiques de texture) du vecteur x individuellement comme une source distincte. Nous cherchons donc à estimer un mélange de gaussiennes dans \mathbb{R} pour chaque descripteur que nous combinons ensuite, comme par exemple pour la *corrélacion* (Figure 2).

Pour l'approche bayésienne, il suffit de combiner les probabilités obtenues pour chaque caractéristique et chaque classe par la règle de Bayes. La matrice de confusion sur les 10 tirages est donnée dans le tableau 2.

		Classe estimée				
		C_1	C_2	C_3	C_4	C_5
Classe réelle	C_1	2583	3	9	1	2
	C_2	0	1791	741	151	211
	C_3	0	279	8096	550	2410
	C_4	0	120	189	508	336
	C_5	0	434	2621	492	3473

Tableau 2 – Matrice de confusion pour l'approche bayésienne.

Le taux de bonne classification (somme de la diagonale) est de 65.80 % avec un intervalle de confiance à 95 % de [63.94;67.66].

Dans le cas de l'approche fondée sur les fonctions de croyance continues, une plausibilité est calculée pour chaque classe et chaque caractéristique. Ces fonctions de plausibilité sont ensuite transformées en fonctions de masse pour effectuer la combinaison de Dempster. La décision est ensuite prise sur la probabilité pignistique du résultat de la combinaison. La matrice de confusion obtenue est donnée dans le tableau 3.

Le taux de bonne classification est de 64.49 % avec un intervalle de confiance à 95 % de [62.61;66.37]. Ces résultats sont moins bons

		Classe estimée				
		C_1	C_2	C_3	C_4	C_5
Classe réelle	C_1	2581	7	7	1	2
	C_2	0	1976	561	166	191
	C_3	0	346	7375	707	2907
	C_4	0	134	134	556	329
	C_5	0	485	2249	650	3636

Tableau 3 – Matrice de confusion pour l'approche fondée sur les fonctions de croyance continues.

qu'avec le classifieur bayésien.

		Classe estimée				
		C_1	C_2	C_3	C_4	C_5
Classe réelle	C_1	2598	0	0	0	0
	C_2	0	2578	238	17	61
	C_3	0	61	9373	45	1856
	C_4	0	44	317	487	305
	C_5	0	67	2684	32	4237

Tableau 4 – Matrice de confusion obtenue à partir du k -plus proches voisins crédibilistes.

Il est intéressant de comparer ces résultats avec le classifieur k -plus proches voisins crédibilistes [4]. La matrice de confusion est donnée dans le tableau 4. Nous obtenons un taux de bonne classification de 77.09 % avec un intervalle de confiance à 95 % de [75.44;78.74]. Ces résultats sont significativement meilleurs que pour le classifieur bayésien et les fonctions de croyance continues.

Nous pouvons souligner que la classe C_1 de roche est bien classifiée pour les 3 classifieurs. Nous observons de plus qu'il y a une confusion entre la classe C_3 et C_5 . Sur une image sonar, l'amplitude des ondes acoustiques réfléchies par les fonds sableux est légèrement supérieure à celle des fonds vaseux, mais sûrement pas assez pour influencer sur les propriétés texturales. Ensuite, la classe C_4 ressort du lot par son faible taux de bonne classification. En effet, la proportion des imagerie classées cailloutis est très inférieure aux autres et donc favorise les

erreurs.

Nous remarquons que la classification reste peu satisfaisante avec les approches modélisant les distributions des descripteurs de texture individuellement. Par conséquent, nous allons les modéliser de manière multi-dimensionnelle, c'est-à-dire que nous allons estimer un mélange de gaussiennes dans l'espace des $d = 8$ descripteurs de texture.

Nous obtenons, pour l'approche bayésienne, la matrice de confusion donnée dans le tableau 5.

		Classe estimée				
		C_1	C_2	C_3	C_4	C_5
Classe réelle	C_1	2088	16	23	485	3
	C_2	0	2626	4	320	2
	C_3	0	1	9731	1104	526
	C_4	0	26	41	948	196
	C_5	0	0	650	710	5500

Tableau 5 – Matrice de confusion pour l'approche bayésienne.

Le taux de bonne classification obtenu est de 83.57 % avec un intervalle de confiance à 95 % de [82.12 ;85.02]. Ces résultats sont significativement meilleurs qu'avec le classifieur k -plus proches voisins crédibilistes.

Enfin la matrice de confusion avec l'approche fondée sur les fonctions de croyance continues est représentée dans le tableau 6.

		Classe estimée				
		C_1	C_2	C_3	C_4	C_5
Classe réelle	C_1	2436	0	7	172	0
	C_2	35	2661	5	247	4
	C_3	61	1	9717	460	1123
	C_4	36	5	9	1132	29
	C_5	35	0	514	261	6050

Tableau 6 – Matrice de confusion pour l'approche fondée sur les fonctions de croyance continues.

Le taux de bonne classification est de 87.98 % avec un intervalle de confiance à 95 % de

[86.90 ;89.26]. Les fonctions de croyance continues donnent ainsi des résultats significativement meilleurs que le classifieur bayésien et les k -ppv crédibilistes.

Ces derniers résultats montrent moins de confusion entre les imagerie vase C_3 et celles de sable C_5 . De plus les imagerie cailloutis C_4 sont mieux classifiées.

L'approche bayésienne est connue pour être optimale si les probabilités *a priori* sont connues. Afin d'observer l'influence de l'apprentissage nous allons faire varier la proportion du nombre de vecteurs d'apprentissage.

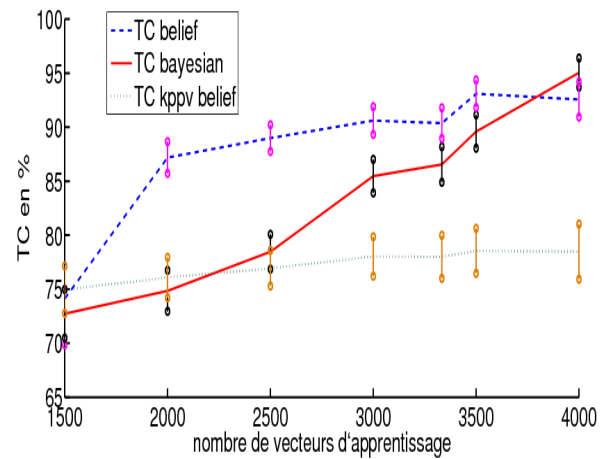


Figure 3 – Représentation des taux de classification en fonction du nombre de vecteurs d'apprentissage

Nous remarquons sur la figure 3 que, même avec un mauvais apprentissage entre 2000 et 3500 vecteurs d'apprentissage, l'approche avec les fonctions de croyance continues est significativement meilleure que les 2 autres approches. De plus, il faut noter que, lorsque les probabilités *a priori* sont bien estimées, les performances du classifieur bayésien sont supérieures à celles du classifieur utilisant les fonctions de croyance continues.

La différence majeure avec le classifieur k -plus proches voisins crédibilistes vient du fait que les approches bayésiennes et les fonctions de croyance continues estiment les corrélations entre les différents descripteurs de

texture. L'approche fondée sur les fonctions de croyance continues a le mérite de ne pas avoir besoin d'estimer les probabilités *a priori* de chaque classe. Ces derniers résultats valident de plus l'intérêt de cette approche en termes de performances dans des conditions éloignées de celles de l'apprentissage.

5 Conclusion

Dans cet article, nous avons montré l'intérêt des fonctions de croyance continues face à une approche bayésienne et une approche fondée sur les fonctions de croyance dans un cadre discret, pour répondre à des problèmes de classification. Les résultats de classification pour la caractérisation des fonds marins à partir d'images sonar sont significativement meilleurs avec l'approche fondée sur les fonctions de croyance continues. Par rapport à l'approche discrète, l'estimation des distributions multi-dimensionnelles des caractéristiques de texture permet de considérer les corrélations entre ces caractéristiques. Lorsque les classes ont des propriétés similaires sur les textures, les résultats sont cependant sensiblement les mêmes. Nous remarquons que l'approche avec les fonctions de croyance continues est plus robuste à un mauvais apprentissage. Cette propriété est intéressante puisqu'en réalité, les conditions d'acquisition des données servant à la base d'apprentissage et à la base de test ne sont pas les mêmes. Ce type d'étude est très lié à la définition des caractéristiques de texture; ces caractéristiques sont toutefois toujours dans un espace continu, les fonctions de croyance continues sont donc bien adaptées à leur modélisation. De plus, ce cadre théorique permettra de tenir compte finement de données multi-vues telle proposée dans [17] à partir d'une approche bayésienne sur le même type d'application.

Références

- [1] F. Caron, B. Ristic, E. Duflos, P. Vanheeghe. Least Committed basic belief density induced by a multivariate Gaussian pdf. *International Conference on Information Fusion (FUSION'2006) Florence, Italie*, 2006.
- [2] A. Dempster, N. Laird, D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Statistical Society B*, 39 : 1-38, 1977.
- [3] A. Dempster. Upper and Lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38 : 325-339, 1967.
- [4] T. Denœux. A k-Nearest Neighbor Classification Rule Based on Dempster-Shafer Theory. *IEEE Transactions on Systems, Man, and Cybernetics-Part A : Systems and Humans*, 25(5) : 804-813, Mai 1995.
- [5] T. Denœux. Extending stochastic ordering to belief functions on the real line. *Information Sciences*, 179 : 1362-1376, 2009.
- [6] R. Haralick, K. Shanmugan, I. Dinstein. Textural features for image classification. *IEEE Transaction on Systems, Man, and Cybernetics* 3(6) : 610-621, 1973.
- [7] A. Martin. Comparative study of information fusion methods for sonar images classification. *International Conference on Information Fusion, Philadelphia, USA*, 25-29 juillet 2005.
- [8] A. Martin. Fusion for Evaluation of Image Classification in Uncertain Environments. *International Conference on Information Fusion, Florence, Italy*, 10-13 July 2006.
- [9] B. Ristic, Ph. Smets. Belief function theory on the continuous space with an application to model based classification. *Proceedings of Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU*, 4-9, 2004.
- [10] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.
- [11] Ph. Smets. Constructing the pignistic probability function in a context of uncertainty. *Uncertainty in Artificial Intelligence*, 29-39, 1990.
- [12] Ph. Smets. The Combination of evidence in the Transferable Belief Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5) : 447-458, 1990.
- [13] Ph. Smets, R. Kennes. The Transferable Belief Model. *Artificial Intelligent*, 66 : 191-234, 1994.
- [14] Ph. Smets. Belief functions : the disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, 9 : 1-35, 1993.
- [15] Ph. Smets. Belief functions on real numbers. *International journal of approximate reasoning*, 40(3) : 181-223, 2005.
- [16] T.M. Strat. Continuous belief functions for evidential reasoning. *Proceedings of the National Conference on Artificial Intelligence, University of Texas at Austin*, 1984.
- [17] D.P. Williams. Bayesian data fusion of multi-view Synthetic Aperture Sonar imagery for seabed classification. *IEEE Signal and Image Processing*, 8(4) : 1239-54, 2009.
- [18] L.M. Zouhal, T. Denœux. An Evidence-Theoretic k-NN Rule with Parameter Optimization. *IEEE Transactions on Systems, Man, and Cybernetics - Part C : Applications and Reviews*, 28(2) : 263-271, Mai 1998.