HAL
open science

Influence de l'estimation des paramètres de texture pour la classification de données complexes

Anthony Fiche, Jean-Christophe Cexus, Arnaud Martin, Ali Khenchaf

To cite this version:

Anthony Fiche, Jean-Christophe Cexus, Arnaud Martin, Ali Khenchaf. Influence de l'estimation des paramètres de texture pour la classification de données complexes. Extraction et Gestion des Connaissances, Jan 2011, Brest, France. pp.10h. hal-00657521

HAL Id: hal-00657521

https://hal.science/hal-00657521

Submitted on 6 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Influence de l'estimation des paramètres de texture pour la classification de données complexes

Anthony Fiche*, Jean-Christophe Cexus*
Arnaud Martin**, Ali Khenchaf*

*2 rue François Verny
29806 Brest Cedex 9
{Anthony.Fiche,Jean-Christophe.Cexus,Ali.Khenchaf}@ensieta.fr
**Université de Rennes 1 / IUT Lannion
UMR 6074 IRISA
Rue Edouard Branly BP 30219
22302 Lannion Cedex
Arnaud.Martin@univ-rennes1.fr

Résumé. Ce papier présente une classification de données basée sur la théorie des fonctions de croyance. La complexité de ce problème peut être vue de deux façons. Tout d'abord, ces données peuvent être imprécises et/ou incertaines. Ensuite, il est difficile de trouver le juste modèle permettant de représenter les données. Le modèle Gaussien est souvent utilisé mais reste limité lorsque les données sont complexes. Ce modèle n'est qu'un cas particulier des distributions α-stables qui permettent une plus grande souplesse dans la modélisation des données. La classification est divisée en deux étapes. La phase d'apprentissage permet de modéliser les données par un mélange de distributions α-stables et de Gaussiennes. La phase de test permet de classifier les données à partir de la théorie des fonctions de croyance et de comparer les deux modèles. La classification est d'abord réalisée sur des données générées puis réelles type images sonar.

1 Introduction

La classification de données réelles à partir d'images radar ou sonar est un problème complexe (Kernéis, 2007; Laanaya, 2007). Par exemple, des paramètres de texture calculés à partir des travaux d'Haralick (Haralick et al., 1973; Haralick, 1979) peuvent être extraits de ces images. La complexité se traduit par le fait que ces paramètres peuvent présenter une queue lourde, c'est à dire que la queue de la distribution décroît plus lentement que la queue de la Gaussienne, être asymétrique ou bien comporter plusieurs modes. Ces différentes contraintes entraînent des difficultés pour choisir un modèle permettant de représenter ces paramètres de texture sans perdre d'information. Le modèle Gaussien est très souvent utilisé du fait de sa simplicité d'utilisation. Cependant, ce modèle devient obsolète dès que les données sont complexes. Il est alors impossible de les représenter à partir d'une seule distribution mais plutôt avec un mélange de distributions. Le modèle Gaussien n'est qu'un cas particulier des distributions α-stables. Les distributions α-stables ont la particularité de modéliser des données non

symétriques ainsi que d'être pourvues d'une queue lourde. Ces distributions ont vu leur utilité s'accroître du fait qu'elles peuvent modéliser des bruits impulsifs en radar et en télécommunications.
L'objectif de cette contribution est de montrer l'intérêt de modéliser des distributions de données à partir d'un mélange de distributions α-stables par rapport à un mélange de Gaussiennes lors de classification de données. Les données issues de capteurs vont être modéliser à partir d'un mélange de distributions α-stables. Ces données sont supposées incertaines et/ou imprécises. La théorie des fonctions de croyance permet de prendre en compte ces considérations. Le papier se divise en trois parties. Tout d'abord, nous présentons les distributions α-stables. Ensuite, nous développons la théorie des fonctions de croyance. Enfin, nous effectuons une classification de données générées et réelles.

2 Les distributions α-stables

Les distributions α-stables ont été introduites par Paul Lévy (Lévy, 1924). Il existe plusieurs définitions permettant de caractériser une distribution α-stable. Dans cette partie, nous présentons tout d'abord la notion de stabilité, ensuite la fonction caractéristique et enfin la densité de probabilité.

2.1 Notion de stabilité

Paul Lévy définit la notion de stabilité par le fait que la somme de deux variables aléatoires indépendantes, chacune suivant une loi stable, suit aussi une loi stable. Cette définition se traduit mathématiquement par : Une variable aléatoire X est dite stable si $\forall(a,b) \in (\mathbb{R}^+)^2$, il existe $c \in \mathbb{R}^+$ et $d \in \mathbb{R}$ tel que :

$$aX_1 + bX_2 = cX + d \tag{1}$$

avec X_1 et X_2 2 variables aléatoires stables indépendantes.
Dans la suite de l'article, nous travaillons avec la densité de probabilité. Or, cette définition ne nous permet pas de la représenter. Par la suite, nous présentons la définition de la fonction caractéristique.

2.2 Fonction caractéristique d'une α-stable

Il n'existe pas qu'une seule définition pour la fonction caractéristique d'une distribution α-stable, notée $S_\alpha(\beta,\gamma,\delta)$. La définition usuelle est celle proposée par Samorodnitsky et Taqqu (1994). Cependant, la fonction caractéristique n'est pas continue pour les valeurs de x où $\alpha = 1$ et $\beta = 0$. On préfère alors celle définie par Zolotarev (1986). Une variable aléatoire est dite stable si sa fonction caractéristique $\phi(t)$ vérifie :

$$\phi(t) = \begin{cases} \exp(it\delta - |\gamma t|^\alpha[1 + i\beta\tan(\frac{\pi\alpha}{2})sign(t)(|t|^{1-\alpha} - 1)]) & \text{si } \alpha \neq 1 \\ \exp(it\delta - |\gamma t|[1 + i\beta\frac{2}{\pi}sign(t)\log|t|]) & \text{si } \alpha = 1 \end{cases} \tag{2}$$

avec $\alpha \in]0,2]$, $\beta \in [-1,1]$, $\gamma \in \mathbb{R}^{+*}$ et $\delta \in \mathbb{R}$.

Ces quatre paramètres sont :
- α est appelé l'exposant caractéristique.
- β est le paramètre de d'asymétrie.
- γ représente le paramètre d'échelle.
- δ indique le paramètre de localisation.

La fonction de densité de probabilité (*fdp*) est obtenue en effectuant une transformée de Fourier de la fonction caractéristique :

$$fdp(x) = \int_{-\infty}^{+\infty} \phi(t) \exp(-itx) dt \tag{3}$$

Plusieurs problèmes rendent difficile la représentation de cette densité de probabilité. Tout, d'abord, l'expression de la fonction caractéristique est complexe. Ensuite, les bornes d'intégration sont infinies. Cependant, Nolan (1997) permet de résoudre ce dernier point en effectuant des changements de variable pour se ramener à des bornes d'intégration finies. Un programme Matlab suivant cette démarche a été développé [1].

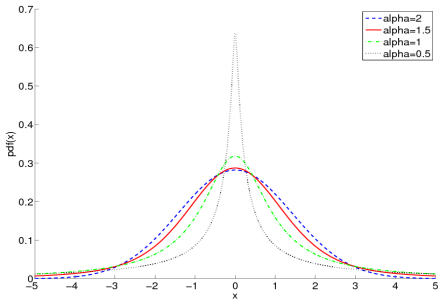


FIG. 1 – *Influence du paramètre* α *avec* $\beta = 0$, $\gamma = 1$ *et* $\delta = 0$.

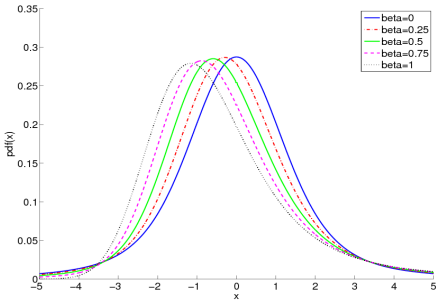


FIG. 2 – *Influence du paramètre* β *avec* $\alpha = 1.5$, $\gamma = 1$ *et* $\delta = 0$.

1. http ://math.bu.edu./people/mveillet/research.html.

Chaque paramètre influe de manière différente sur la représentation de la densité de probabilité d'une loi stable. On remarque que si α est petit, la distribution présente un pic très important (*cf.* figure 1). Lorsque $\beta \to 1$, la distribution a une queue lourde à droite et inversement lorsque $\beta \to -1$ (*cf.* figure 2). Le paramètre γ permet de dilater ou de compresser les distributions (*cf.* figure 3). Enfin, δ permet de positionner le mode de la distribution suivant l'axe des abscisses (*cf.* figure 4).

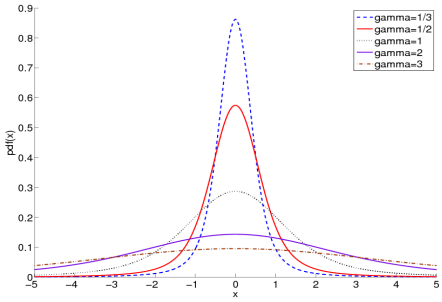


FIG. 3 – *Influence du paramètre* γ *avec* $\alpha = 1.5$, $\beta = 0$ *et* $\delta = 0$.

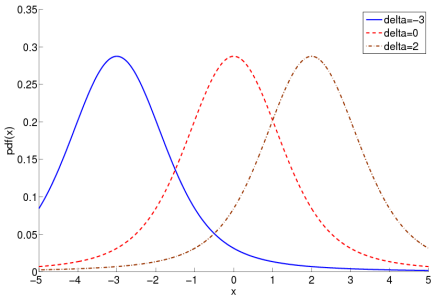


FIG. 4 – *Influence du paramètre* δ *avec* $\alpha = 1.5$, $\beta = 0$ *et* $\gamma = 1$.

2.3 Exemples de *fdp*

Malgré la difficulté à représenter les distributions α-stables, il est possible de décrire quelques lois connues. Lorsque $\alpha = 2$ et $\beta = 0$, on retrouve l'expression d'une distribution Gaussienne :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\delta)^2}{2\sigma^2}\right) \tag{4}$$

avec δ représentant la moyenne et σ^2 la variance. À partir de la définition de la fonction caractéristique, il faut que $\sigma^2 = 2\gamma^2$. Lorsque $\alpha = 1$ et $\beta = 0$, on définit une loi de Cauchy :

$$f(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x-\delta)^2} \tag{5}$$

Plus généralement, lorsque $\beta = 0$, on dit que la distribution α-stable est symétrique.
Les données issues des capteurs vont être modélisées par un mélange de distributions α-stables. Ces données sont considérées comme imprécise et/ou incertaine. Des travaux ont été proposés pour classifier des données modélisées à partir d'un mélange de Gaussiennes en s'appuyant sur la théorie des probabilités (voir (Williams, 2009)). La probabilité de chaque sédiment est calculée à partir du mélange de Gaussiennes. Cependant, cette théorie est limitée pour prendre en compte l'incertitude. La théorie des fonctions de croyance permet de prendre en considération cette notion. Par conséquent, la partie suivante a pour but de définir les concepts de cette théorie.

3 La théorie des fonctions de croyance

Les travaux de Dempster (1967) sont à l'origine de la théorie des fonctions de croyance. Shafer (1976) a d'ailleurs repris ces travaux pour formaliser cette théorie. Par conséquent, nous exposons les concepts de base de la théorie des fonctions de croyance dans le cadre discret puis dans le cadre continu.

3.1 Les fonctions de croyance dans le cadre discret

Dans cette section, nous définissons la fonction de masse m, la règle de combinaison permettant de combiner plusieurs fonctions de masses entre elles ainsi que la probabilité pignistique, permettant de prendre une décision.

3.1.1 Définitions

Tout d'abord, la théorie des fonctions de croyance permet de travailler sur l'ensemble $\Theta = \{C_1, \ldots, C_n\}$, appelé cadre de discernement. Θ s'interprète comme toutes les hypothèses possibles d'un problème. Les fonctions de croyance sont définies de 2^Θ dans $[0, 1]$, c'est à dire qu'il est possible d'attribuer une croyance sur des disjonctions de Θ. La quantité m appelée fonction de masse vérifie :

$$\sum_{A \in 2^\Theta} m(A) = 1 \tag{6}$$

À partir de cette fonction de masse, il est possible de définir d'autres fonctions :

$$bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B) \tag{7}$$

$$pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \tag{8}$$

$$q(A) = \sum_{B \subset \Theta, B \supseteq A} m(B) \tag{9}$$

Ces fonctions représentent la même information que m mais sous des formes différentes. La fonction de crédibilité, notée bel, correspond à la croyance minimum que l'on peut avoir en A.

La fonction de plausibilité, notée pl, correspond à la croyance maximale que l'on peut avoir en A. Enfin, la fonction de communalité, notée q, représente la somme de toutes les masses allouées à un sur ensemble de A et est très utilisée dans l'étape de combinaison.

3.1.2 La règle de combinaison

L'étape de combinaison permet de combiner plusieurs fonctions de masse entre elles. En effet, des experts peuvent avoir des opinions différentes sur un élément $A \subseteq \Theta$ en attribuant une masse m_j à A. Il existe plusieurs règles de combinaison répartissant différemment le conflit entre les sources. La plus employée est la règle de combinaison conjonctive. On obtient comme masse résultante :

$$m(A) = \sum_{C_1 \cap \cdots \cap C_n = A \neq \emptyset} \prod_{j=1}^{M} m_j(B_j) \qquad \forall A \in 2^{\Theta} \tag{10}$$

En pratique, il est difficile de programmer cette formule. On préfère utiliser les fonctions de communalité. Chaque masse m_j est convertit en sa fonction de communalité q_j. Il est possible de calculer la fonction de communalité résultante par :

$$q(A) = \prod_{j=1}^{M} q_j(A) \tag{11}$$

La masse résultante est ensuite obtenue en transformant la fonction de communalité en fonction de masse. Une fois la masse finale de chaque sous ensemble de A, il est nécessaire de prendre une décision.

3.1.3 La transformation pignistique

Plusieurs opérateurs, tels que le maximum de crédibilité et le maximum de plausibilité permettent de prendre une décision. Cependant, l'opérateur le plus utilisé est basé sur la transformée pignistique (Smets, 1990). Le principe est de répartir uniformément la masse des sous-ensembles de Θ sur les singletons qui les composent. La formule est la suivante :

$$betP(C_i) = \sum_{A \in 2^{\Theta}, C_i \in A} \frac{m(A)}{|A|(1 - m(\emptyset))} \tag{12}$$

où $|A|$ représente la cardinalité de A.
On choisit ensuite la décision C_i pour l'observation $\mathbf{x}$ en évaluant $\max_{1 \leq k \leq n} betP(C_k)(x)$.

3.2 Les fonctions de croyance continues

Il est possible de définir les notions définies précédemment dans le cadre continu. Les notions de bases ont été introduites par Shafer (1976), puis par Strat (1984). Récemment, Smets (2005) étend la théorie des fonctions de croyance sur l'ensemble $\mathbb{R} = \mathbb{R} \cup [-\infty, +\infty]$ où il attribue une masse sur des intervalles de $\mathbb{R}$.

3.2.1 Définitions

Considérons $\mathcal{I} = \{[x,y],(x,y],[x,y),(x,y); x,y \in \mathbb{R}\}$ un ensemble d'intervalles fermés, semi-ouverts et ouverts de $\mathbb{R}$. Les éléments focaux sont des intervalles fermés de $\mathbb{R}$. La quantité $m^{\mathcal{I}}(x,y)$ est appelée densité de masse et est reliée à une densité de probabilité. La densité de masse est nulle lorsque x est supérieur à y. Soit $[a,b]$ un intervalle de $\mathbb{R}$. Les fonctions définies dans le cadre discret deviennent :

$$bel([a,b]) = \int_{x=a}^{x=b} \int_{y=x}^{y=b} m^{\mathcal{I}}(x,y)dydx \tag{13}$$

$$pl([a,b]) = \int_{x=-\infty}^{x=b} \int_{y=max(a,x)}^{y=+\infty} m^{\mathcal{I}}(x,y)dydx \tag{14}$$

$$q([a,b]) = \int_{x=-\infty}^{x=a} \int_{y=b}^{y=+\infty} m^{\mathcal{I}}(x,y)dydx \tag{15}$$

3.2.2 La probabilité pignistique

La probabilité pignistique, avec $a < b$, est définie par la formule :

$$BetP([a,b]) = \int_{x=-\infty}^{x=+\infty} \int_{y=x}^{y=+\infty} \frac{|[a,b] \cap [x,y]|}{|[x,y]|} \times m^{\mathcal{I}}(x,y)dxdy \tag{16}$$

Il est possible d'obtenir une densité de masse à partir de la probabilité pignistique. Cependant, il existe plusieurs densités de masse associées à une même probabilité pignistique. Pour simplifier le problème, on choisit la densité de masse consonante, c'est à dire que les éléments focaux sont emboîtés. Les éléments focaux, notés I_u peuvent être rangés dans un index u tel que $I_u \subseteq I_{u'}\ avec\ u' > u$. Cette définition est utilisée pour appliquer le principe de moindre engagement. Le problème avec les fonctions de croyance est qu'elles ne sont pas totalement définies. La seule chose que l'on connaît est son appartenance à une famille de fonctions. Le principe de moindre engagement consiste à choisir la fonction de croyance qui est la moins informative.

3.2.3 Théorème de Bayes généralisé

Supposons que $x \in \mathbb{R}$ un vecteur d'observations. Il est possible de calculer la fonction de masse de chaque $A \in 2^{\Theta}$ connaissant l'observation x à partir du théorème de Bayes généralisé (Smets, 1993; Delmotte et Smets, 2004). Il s'écrit de la manière suivante :

$$m(A/x) = \prod_{C_j \in A} pl_j(x) \prod_{C_j \in A^c} (1 - pl_j(x)) \tag{17}$$

Les fonctions de plausibilité ont été calculées dans le cas symétrique unimodale, qui est vraie pour le cas Gaussian (Smets, 2005) et α-stable symétrique (Fiche et al., 2010a). Il est possible de généraliser le calcul des fonctions de plausibilité dans le cas d'un mélange de Gaussiennes (Caron et al., 2006). Nous reprenons cette démarche pour l'étendre à un mélange de distributions α-stables.

	α			γ			δ			W		
	α_1	α_2	α_3	γ_1	γ_2	γ_3	δ_1	δ_2	δ_3	W_1	W_2	W_3
Premier mélange	1.6	1.4	1.2	1.26	1.5	0.8	-4.5	0	10.1	1/2	1/3	1/6
Second mélange	1.2	1.4	1.6	1.2	3.2	1.5	-13.5	10.1	3	1/2	1/4	1/4
Troisième mélange	1.56	1.24	1.78	0.5	2	4	-8.7	1.3	5.5	1/6	1/6	2/3

TAB. 1 – *Valeurs des paramètres de chaque mélange de distributions α-stables.*

4 Application

Dans cette partie, nous allons classifier des données générées et réelles. Tout d'abord, nous présentons et classifions les données générés. Puis nous effectuons la même démarche que précédemment mais avec des données réelles type images sonar.

4.1 Classification de données générées

4.1.1 Présentation des données

Nous simulons 3 mélanges de distributions α-stables dont les valeurs sont indiquées dans le tableau 1. Chaque mélange est constitué de 3 α-stables générées (voir (Chambers et al., 1976)). Les densités de probabilité peuvent être vues comme des attributs correspondant à différentes classes. Dans la suite, nous allons appliquer la théorie des fonctions de croyance développée précédemment pour classifier nos données.

4.1.2 Résultats

Chaque mélange est constitué de 6000 échantillons. Ces échantillons sont divisés en deux parties : l'une sert à la base d'apprentissage et l'autre à la base de test. Lors de la phase d'apprentissage, il est difficile de choisir un modèle permettant d'estimer les distributions sans perdre d'information. Nous estimons chaque attribut à partir d'un mélange de 3 distributions α-stables. L'outil que nous utilisons pour l'estimation du mélange est l'algorithme Expectation-Maximization. Initialement, il a été développé pour estimer un mélange de Gaussiennes (voir (Dempster et al., 1977)). Nous l'avons étendu dans le cas d'un mélange de distributions α-stables (Fiche et al., 2010b). Une estimation à partir d'un mélange de Gaussiennes est effectuée pour avoir des éléments de comparaisons. Comme chaque mélange est estimé à partir de 3 α-stables, soit 12 paramètres à estimer, nous prenons 3 Gaussiennes pour modéliser chaque mélange, soit 12 paramètres à estimer. À partir de l'estimation de chaque mélange, nous allons classifier les éléments de notre base de test. Tout d'abord, nous calculons chaque fonction de plausibilité dans le cas d'un mélange de Gaussiennes et d'α-stables. Le théorème de Bayes généralisée permet de calculer la fonction de masse attribuée à chaque $A \in 2^{\Theta}$. La transformation pignistique permet de travailler sur les singletons. La décision finale est faite en utilisant le maximum de probabilité pignistique. Nous effectuons cette démarche 5 fois. On obtient alors un taux de bonne classification moyenné. Nous estimons aussi chaque mélange avec 3 et 5 Gaussiennes. Les résultats sont représentés dans le tableau 2. On remarque que le taux de classification sous l'hypothèse d'un mélange de distributions α-stables est sensiblement meilleur que les taux de classification sous l'hypothèse de mélange de Gaussiennes. Cependant, on

	Taux de classification	Intervalles de confiance
Mélange de 3 α-stables	60.51	[58.77 ;62.27]
Mélange de 4 Gaussiennes	53.96	[52.19 ;55.75]
Mélange de 3 Gaussiennes	50.79	[49.76 ;52.58]
Mélange de 5 Gaussiennes	57.83	[56.07 ;59.60]

TAB. 2 – *Taux de classification et intervalles de confiance associés à chaque mélange.*

remarque que plus le nombre de Gaussiennes augmente et plus les taux de classification se rapprochent de ceux du mélange de 3 α-stables. L'inconvénient est qu'on augmente le nombre de paramètres à estimer, par exemple 15 paramètres dans le cas d'un mélange de 5 Gaussiennes. On note qu'il est difficile de choisir le bon modèle lors de la phase d'apprentissage. Le choix d'un mélange de distributions α-stables peut être pertinent pour modéliser des données. Dans la partie suivante, nous allons classifier des données réelles en comparant les taux de classification sous l'hypothèse mélange de Gaussiennes et d'α-stables.

4.2 Classification de données réelles

Nous classifions des images sonar en extrayant des paramètres d'Haraclick. Tout d'abord, nous présentons les données et ensuite nous exposons les résultats obtenus.

4.2.1 Présentation des données

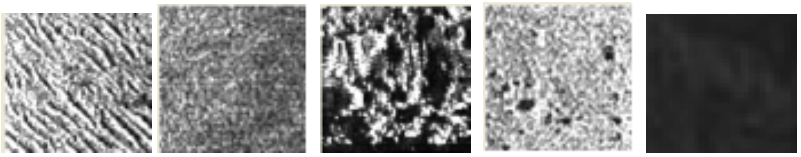
Nous disposons d'une base de données de 42 images sonar (Exemple figure 5) fournie par le GESMA (groupes d'Études Sous-Marines de l'Atlantique). Ces images ont subi un prétraitement afin de corriger la variation de gain et réduire le bruit de chatoiement responsable de l'aspect granulaire des images sonar. Des experts ont attribués à chaque pixel des images un type de sédiments. On en retrouve 5 : sable, rides de sable, vase, cailloutis et roche (Exemple figure 6). Chaque image est découpée en imagettes de taille 32×32 pixels. Une imagette est dite "roche" si tous ses pixels sont classifiés "roche". On extrait des paramètres de texture de chaque imagette grâce aux travaux d'Haralick. Nous en choisissons 8 en particulier : moment des différences inverses, corrélation, contraste, moyenne des sommes, moyenne des variances, entropie des sommes, entropie des différences et mesure de corrélation.

4.2.2 Classification des images sonar

On dispose d'une base de plus de 30 000 imagettes classifiées roche, sable, rides de sable, cailloutis et vase. On effectue un tirage aléatoire de 5000 imagettes. Une moitié est utilisée pour l'apprentissage des mélanges et l'autre utilisée comme base de test. La classification de ces données a déjà été étudiée dans Fiche et Martin (2009) où chaque paramètre de texture était estimé à partir d'un mélange de 5 Gaussiennes. Dans notre étude, nous procédons de la manière suivante : nous fixons le nombre de paramètres à estimer à 12, ce qui revient à estimer chaque paramètre de texture à partir d'un mélange de 4 Gaussiennes et d'un mélange de 3 α-stables. Puis, nous faisons varier le nombre de Gaussiennes du mélange : mélange de 3 Gaussiennes, soit 9 paramètres à estimer, et un mélange de 5 Gaussiennes, soit 15 paramètres

FIG. 5 – *Exemple d'images sonar.*

à estimer. On utilise la même démarche que dans le cas des données générées. On obtient alors des taux de classification moyennés sur 5 tirages.
L'estimation des paramètres à partir d'un mélange de Gaussiennes donne un taux de classification de 65.34 %, avec un intervalle de confiance à 95 % de [63.47 ;67.21], tandis que l'estimation des paramètres à partir d'un mélange de distributions α-stables offre un taux de classificatio de 64.58 %, avec un intervalle de confiance à 95 % de [62.70 ;66.65]. On remarque que l'hypothèse mélange de Gaussiennes permet d'avoir un meilleur taux de classification mais ce n'est pas significatif puisque les intervalles de confiance se chevauchent. Nous effectuons la même démarche avec un mélange de 3 Gaussiennes et avec un mélange de 5 Gaussiennes (*c.f.* Tableau 3). On remarque, en diminuant le nombre de Gaussiennes, qu'on obtient un taux de bonne classification meilleur qu'avec le mélange d'α-stables, mais pas significativement. Ceci peut s'expliquer par le fait que les paramètres de texture extraits des images ont une densité de probabilité de type Gaussienne.



(a) ride (b) sable (c) roche (d) cailloutis (e) vase

FIG. 6 – *Exemples d'imagettes.*

	Taux de classification	Intervalles de confiance
Mélange de 3 α-stables	64.58	[62.70 ;66.45]
Mélange de 4 Gaussiennes	65.19	[63.32 ;67.06]
Mélange de 3 Gaussiennes	64.66	[62.78 ;66.53]
Mélange de 5 Gaussiennes	65.34	[63.47 ;67.21]

TAB. 3 – *Taux de classification et intervalles de confiance associés à chaque mélange.*

5 Conclusion

La classification de données générées montre qu'il peut être intéressant d'utiliser un mélange de distributions α-stables par rapport à un mélange de Gaussiennes lors de la phase d'apprentissage, en considérant le même nombre de paramètres à estimer. Cependant, la limite est qu'en augmentant le nombre de Gaussiennes, il est toujours possible d'avoir une estimation correcte des données. Ensuite, il y a une différence entre la théorie et la pratique. En effet, la classification de données réelles montre que les résultats sont significativement les mêmes suivant l'hypothèse de mélange choisi. Dans ce cas, il est difficile de conclure quant à la pertinence du choix du modèle. Les paramètres extraits des images sonar ont une densité de probabilité plutôt type mélange de Gaussiennes ce qui peut expliquer le fait que les résultats soient significativement les mêmes. Il serait intéressant de travailler sur des données qui se prêtent plus aux distributions α-stables pour avoir des taux de classification meilleurs.

Références

Caron, F., B. Ristic, E. Duflos, et P. Vanheeghe (2006). Least Committed basic belief density induced by a multivariate Gaussian pdf. In *9th International Conference on Information Fusion, Florence, Italie*.

Chambers, J., C. Mallows, et B. Stuck (1976). A method for simulating stable random variables. *Journal of the American Statistical Association 71*(354), 340–344.

Delmotte, F. et P. Smets (2004). Target identification based on the transferable belief model interpretation of Dempster-Shafer model. *Systems, Man and Cybernetics, Part A : Systems and Humans, IEEE Transactions on 34*(4), 457–471.

Dempster, A. (1967). Upper and lower probabilities generated by a random closed interval. *The Annals of Mathematical Statistics 38*, 325–339.

Dempster, A., N. Laird, D. Rubin, et al. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society 39*(1), 1–38.

Fiche, A. et A. Martin (2009). Approche bayésienne et fonctions de croyance continues pour la classification. In *Rencontre francophone sur la Logique Floue et ses Applications, Annecy, France*.

Fiche, A., A. Martin, J. Cexus, et A. Khenchaf (2010a). Continuous belief functions and alpha-stable distributions. In *13th International Conference on Information Fusion, Edinburgh, United-Kingdom*.

Fiche, A., A. Martin, J. Cexus, et A. Khenchaf (2010b). Estimation d'un mélange de distributions alpha-stables à partir de l'algorithme em. In *Rencontre francophone sur la Logique Floue et ses Applications, Lannion, France*.

Haralick, R. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE 67*(5), 786–804.

Haralick, R., K. Shanmugam, et I. Dinstein (1973). Textural features for image classification. *IEEE Transactions on systems, man and cybernetics 3*(6), 610–621.

Kernéis, D. (2007). *Amélioration de la classification automatique des fonds marins par la fusion multicapteurs acoustiques*. Ph. D. thesis, Thèse de doctorat de l'ENSTB.

Laanaya, I. (2007). *Classification en environnement incertain : application à la caractérisation de sédiments marins*. Ph. D. thesis, thèse de doctorat de l'UBO.

Lévy, P. (1924). Théorie des erreurs : La loi de gauss et les lois exponentielles. *Bulletin de la Société Mathématique de France 52*, 49–85.

Nolan, J. (1997). Numerical calculation of stable densities and distribution functions. *Communications in Statistics-Stochastic Models 13*(4), 759–774.

Samorodnitsky, G. et M. Taqqu (1994). *Stable non-Gaussian random processes : stochastic models with infinite variance*. Chapman & Hall.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton university press Princeton.

Smets, P. (1990). Constructing the pignistic probability function in a context of uncertainty. In *Uncertainty in artificial intelligence*, Volume 5, pp. 29–39.

Smets, P. (1993). Belief functions : The disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning 9*(1), 1–35.

Smets, P. (2005). Belief functions on real numbers. *International journal of approximate reasoning 40*(3), 181–223.

Strat, T. (1984). Continuous belief functions for evidential reasoning. In *Proceedings of the National Conference on Artificial Intelligence, University of Texas at Austin*.

Williams, D. (2009). Bayesian data fusion of multiview Synthetic Aperture Sonar imagery for seabed classification. *Image Processing, IEEE Transactions on 18*(6), 1239–1254.

Zolotarev, V. (1986). One-dimensional stable distributions, Translations of Mathematical Monographs, vol. 65. *American Mathematical Society*.

Summary

This paper shows a classification of data based on the theory of belief functions. The complexity of this problem can be seen as two ways. Firstly, data can be imprecise and/or uncertain. Then, it is difficult to choose the right model to represent data. Gaussian model is often used but is limited when data are complex. This model is a particular case of α-stable distributions. Classification is divided into two steps. Learning step allows to modelize data by a mixture of α-stable distributions and Gaussian distributions. Test step allows to classify data with the theory of belief functions and compare the two models. The classification is realized firstly on generated data and then on real data type sonar images.