Sur l'ajustement d'une fonction de croyance par une matrice de confusion

On the Ajustment of a Belief Function with a Confusion Matrix

D. Mercier[1] Z. Elouedi[2] É. Lefèvre[1]

[1] Univ. Lille Nord de France, UArtois, EA 3926 LGI2A, France
{david.mercier, eric.lefevre}@univ-artois.fr

[2] Institut Supérieur de Gestion de Tunis, LARODEC, Tunisie
zied.elouedi@isg.rnu.tn

Résumé :

Dans ce papier, une première étude est réalisée sur des mécanismes permettant l'ajustement de fonctions de croyance en se basant sur les résultats fournis par une matrice de confusion. Pour ce faire, trois approches basées sur des affaiblissements sont exposées.

Ces méthodes permettent l'estimation des coefficients d'affaiblissement attribués à une source, afin de corriger ses décisions brutes, et ce en tenant compte des décisions apprises et offertes par la matrice de confusion. Ces ajustements diffèrent selon qu'un affaiblissement classique ou contextuel est employé.

Ces méthodes sont illustrées sur un exemple mettant en avant leurs intérêts respectifs.

Mots-clés :

Théorie de Dempster-Shafer, Ajustement d'une fonction de croyance, Affaiblissements, Matrice de confusion.

Abstract:

In this paper, a pre-study is performed on mechanisms allowing the adjustment of belief functions based on the results provided by a confusion matrix. To this end, three different approaches based on discountings are displayed.

These methods allow the estimation of discounting rates assigned to a source to correct its raw data by taking into account learnt decisions given by the confusion matrix. These adjustments differ depending on whether a traditional or contextual discounting is realized.

These methods are illustrated with an example highlighting their respective interests.

Keywords:

Theory of Dempster-Shafer, Belief Function Ajustment, Discountings, Confusion Matrix.

1 Introduction

La théorie des fonctions de croyance de Dempster-Shafer [8, 24, 26] a déjà montré tout son intérêt pour la gestion de données imparfaites, notamment dans le domaine de la fusion d'informations [6, 7, 22, 29].

En pratique cependant, il est nécessaire de trouver un modèle de construction des fonctions de croyance qui vont représenter les informations disponibles, et il ne semble pas exister de solution universelle [7].

Généralement, cette phase de génération est très dépendante de l'application visée et de la nature des sources employées. Ainsi, dans de nombreuses applications en rapport direct avec un problème réel, les sources d'information (classifieurs, capteurs, ...) sont jugées suffisamment connues pour que les éléments focaux soient fixés et les masses déterminées à partir de seuils ou d'abaques caractéristiques de la source employée [5, 14, 21, 28, 23, 30]. En reconnaissance des formes, des méthodes d'affectation ont été définies à partir du vecteur forme de l'objet à identifier, par le calcul d'une distance [4, 9, 10] ou d'une vraisemblance [1, 2, 31] [15, chapitre 3]. Des méthodes ont aussi été développées pour l'élicitation d'avis d'experts [3].

Lorsque qu'une source ne fournit que des décisions brutes et que des données d'apprentissage sont disponibles, il est possible de convertir la décision fournie en une fonction de croyance à l'aide de sa matrice de confusion, qui regroupe toutes ses décisions passées au regard d'objets dont la vérité est connue [16, 19, 32].

Dans ce papier, nous nous intéressons au cas où une matrice de confusion est disponible en plus d'une méthode permettant à une source de fournir une fonction de croyance, cette méthode n'ayant pas été apprise à partir de cette matrice de confusion. Il s'agit alors de voir comment les données présentes dans la matrice peuvent venir corriger ou ajuster l'information fournie par la source.

Pour ce faire, plusieurs solutions sont proposées. En premier lieu, en se basant sur le taux de fiabilité ou taux de classifications correctes issu de la matrice de confusion, un affaiblissement [24, page 252] peut être appliqué. Une deuxième approche consiste à exploiter la fiabilité de la source pour chaque contexte en faisant appel à l'affaiblissement contextuel [18]. Enfin, une troisième proposition tiendra aussi compte des taux d'erreurs dans chaque contexte.

Ce papier est organisé de la manière suivante. Dans le paragraphe 2, les concepts de base dont nous avons besoin sur les fonctions de croyance sont rappelés. Dans le paragraphe 3, la définition d'une matrice de confusion exploitée dans ce papier est exposée et illustrée. Le paragraphe 4 propose alors trois méthodes pour ajuster une fonction de croyance avec la matrice de confusion, ces méthodes étant détaillées et illustrées sur un exemple. Enfin, le paragraphe 5 conclut sur les éléments apportés et ouvre une discussion sur ce travail.

2 Fonctions de croyance : concepts de base

2.1 Représentation de l'information

La connaissance d'un agent est modélisée par l'allocation d'une masse finie de croyance à des sous-ensembles de l'univers de discours ou cadre de discernement, noté $\Omega = \{\omega_1, \omega_2, \ldots, \omega_K\}$, et qui contient un ensemble fini de réponses à une certaine question Q d'intérêt.

Une *fonction de masse de croyance* sur Ω est une application $m^\Omega : 2^\Omega \to [0,1]$ telle que :

$$\sum_{A \subseteq \Omega} m^\Omega(A) = 1 \, . \tag{1}$$

En l'absence d'ambiguïté sur son cadre de discernement, une fonction de masse de croyance m^Ω pourra être notée plus simplement m.

La fonction de masse m représente l'état de connaissance l'agent Ag au regard de la question Q. Une masse $m(A)$ représente la part de croyance de l'agent Ag allouée exactement à la proposition : "la réponse à la question Q se trouve dans le sous-ensemble A de Ω, et dans aucun sous-ensemble strict". $m(\Omega)$ représente le degré d'ignorance de l'agent Ag. Ainsi, la fonction de masse définie par $m(\Omega) = 1$ représente l'ignorance totale. Cette fonction est appelée *fonction de masse vide* et est notée m_Ω. Le nombre $m(\emptyset)$ représente le degré de conflit [27]. Tout sous-ensemble A de Ω tel que $m(A) > 0$ est appelé *élément focal* (EF) de m.

La fonction de croyance et la fonction de plausibilité associées à une fonction de masse m sont définies respectivement par :

$$bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B) \, , \tag{2}$$

et

$$pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \, , \tag{3}$$

pour tout $A \subseteq \Omega$.

Les fonctions *bel*, *pl* et m sont en correspondance biunivoque, elles représentent la même information sous une forme différente.

2.2 Manipulation des informations

Combinaisons. Deux fonctions de masse m_1 et m_2 issues de deux sources d'informations fiables et distinctes peuvent être combinées en utilisant la *règle de combinaison conjonctive* définie pour tout $A \subseteq \Omega$ par [27] :

$$m_1 \bigcirc\!\!\!\!\cap\, m_2(A) = \sum_{B \cap C = A} m_1(B) m_2(C) \, . \tag{4}$$

Deux fonctions de masse m_1 et m_2 issues de deux sources distinctes d'informations dont l'une au moins est fiables peuvent être combinées en utilisant la *règle de combinaison disjonctive* définie pour tout $A \subseteq \Omega$ par [27] :

$$m_1 \circledcirc m_2(A) = \sum_{B \cup C = A} m_1(B) m_2(C) \, . \tag{5}$$

Affaiblissement classique d'une information. Un doute sur la fiabilité d'une source ayant fourni une information m est parfois possible. L'opération d'affaiblissement [24, page 252] de m par une constante $\alpha \in [0,1]$, appelé *taux d'affaiblissement*, permet de prendre en compte cette *métaconnaissance* sur l'information m. Cette opération de correction de m est définie par :

$$\begin{cases} {}^{\alpha}m(A) & = & (1-\alpha)m(A), \quad \forall A \subset \Omega, \\ {}^{\alpha}m(\Omega) & = & (1-\alpha)m(\Omega) + \alpha \, , \end{cases} \tag{6}$$

ou, plus simplement :

$${}^{\alpha}m = (1-\alpha)m + \alpha \, m_{\Omega} \, . \tag{7}$$

Le coefficient $\beta = (1-\alpha)$ représente le degré de fiabilité de la source d'information. Si la source n'est pas fiable, ce degré de fiabilité β est égal à 0, le taux d'affaiblissement α est égal à 1, et ${}^{\alpha}m$ est égale à la fonction de masse vide m_{Ω}. Au contraire, si la source est fiable, le taux d'affaiblissement α est nul, et m n'est aucunement affaiblie.

D'autres mécanismes de correction sont exposés dans [13, 17, 33] dont l'affaiblissement contextuel [18, 20].

Affaiblissement contextuel d'une information. L'idée principale de l'affaiblissement contextuel est que la fiabilité d'une source peut varier en fonction de la vérité de l'objet à reconnaître (le contexte). Par exemple, un capteur en charge de reconnaître des cibles volantes peut être plus ou moins capable de discerner certains types d'avion.

Une méthode simple pour calculer un affaiblissement contextuel consiste à utiliser son expression par la règle de combinaison disjonctive (5).

Supposons qu'une source possède un degré de fiabilité β_A sachant que la vérité se trouve dans A, pour tout sous-ensemble A de Ω appartenant à un ensemble $\mathcal{A}$ d'éléments de Ω. L'affaiblissement contextuel d'une information m fournie par cette source est alors donnée par [20] :

$${}^{\boldsymbol{\alpha}}m = m \circledcirc_{A \in \mathcal{A}} A_{\beta_A} \tag{8}$$

où, pour tout $A \neq \emptyset$, $v \in [0,1]$, A_v est une autre notation [11] pour écrire une fonction de masse telle que :

$$\begin{array}{rlcl} A_v : & 2^{\Omega} & \to & [0,1] \\ & \emptyset & \mapsto & v \\ & A & \mapsto & 1-v \\ & B & \mapsto & 0 \, , \; \forall B \in 2^{\Omega} \setminus \{\emptyset, A\} \, . \end{array} \tag{9}$$

La fonction de masse A_v a donc deux éléments focaux : $\emptyset$ et A. De plus, on a $A_{v_1} \circledcirc A_{v_2} = A_{v_1 v_2}$.

Remarque 1 *Comme l'illustre l'équation (8), lors d'un affaiblissement contextuel, chaque masse $m(A)$ est transférée sur $B \supseteq A$ proportionnellement à la non fiabilité de la source à reconnaître les éléments situés dans $A \setminus B$ et la fiabilité de la source à détecter les éléments situés dans $\Omega \setminus A = \overline{A}$.*

Remarque 2 *Si on sait seulement que la source est fiable avec un degré β_{Ω}, l'affaiblissement classique est retrouvé :*

$${}^{\boldsymbol{\alpha}}m \; = \; m \circledcirc \Omega_{\beta_{\Omega}} \; = \; {}^{\alpha_{\Omega}}m \, . \tag{10}$$

En effet, $m \circledcirc \Omega_{\beta_{\Omega}}(A) = m(A) \, \Omega_{\beta_{\Omega}}(\emptyset) = \beta_{\Omega} m(A) = {}^{\alpha_{\Omega}}m(A)$ *pour tout* $A \subset \Omega$, *et* $m \circledcirc \Omega_{\beta_{\Omega}}(\Omega) = m(\Omega) \, \Omega_{\beta_{\Omega}}(\emptyset) + \Omega_{\beta_{\Omega}}(\Omega) \sum_{A \subseteq \Omega} m(A) = \beta_{\Omega} m(\Omega) + \alpha_{\Omega} = {}^{\alpha_{\Omega}}m(\Omega)$.

Tableau 1 – Illustration d'une matrice de confusion.

Vérité / Décision	ω_1	$\ldots$	ω_K
ω_1	n_{11}	$\ldots$	n_{1K}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
ω_K	n_{K1}	$\ldots$	n_{KK}

3 Matrice de confusion et taux de fiabilité

Dans ce papier, une matrice de confusion sans décision de rejet est considérée.

Une *matrice de confusion* $\mathcal{M} = (n_{k\ell})_{k\in\{1,\ldots,K\}\ \ell\in\{1,\ldots,K\}}$ associée à une source est une table décrivant les performances de cette source sur un ensemble de test (Tableau 1). Chaque ligne k correspond à une décision en faveur de ω_k. Chaque colonne ℓ correspond au cas où la vérité est ω_ℓ. Le terme général $n_{k\ell}$ est égal au nombre d'objets de classe ω_ℓ ayant été classés dans la classe ω_k par la source.

Nous noterons $n_{k\cdot} = \sum_{\ell=1}^{K} n_{k\ell}$ le nombre d'objets classés dans ω_k, où $k \in \{1,\ldots,K\}$, et $n_{\cdot\ell} = \sum_{k=1}^{K} n_{k\ell}$ le nombre d'objets de type ω_ℓ, où $\ell \in \{1,\ldots,K\}$. Le nombre total d'objets classés dans la matrice est donné par $n = \sum_{k=1}^{K}\sum_{\ell=1}^{K} n_{k\ell}$.

Le *taux de fiabilité* ou *taux de classifications correctes* $\mathcal{T}_f$ d'une matrice est alors définie par $\mathcal{T}_f = \frac{\sum_{k=1}^{K} n_{kk}}{n}$.

4 Ajustement d'une fonction de croyance par une matrice de confusion

4.1 Ajustement par affaiblissement classique

Une première approche consiste à se servir du taux de fiabilité $\mathcal{T}_f$ d'une source d'information pour réaliser un affaiblissement de l'information m fournie par cette source :

$$^{\alpha}m \;=\; m \circledcup \Omega_{\mathcal{T}_f} \,. \tag{11}$$

4.2 Ajustements par affaiblissement contextuel

Utilisation du pourcentage de classifications correctes pour chaque contexte. Une deuxième approche consiste à exploiter la fiabilité de la source pour chaque contexte ω_ℓ, $\ell \in \{1,\ldots,K\}$, c'est-à-dire sachant que la vérité est ω_ℓ. Pour déterminer le taux de fiabilité de la source sachant que la vérité est ω_ℓ, on regardera la colonne ℓ de la matrice (Tableau 1). Ce taux, noté $\mathcal{T}_f[\omega_\ell]$, peut alors être défini comme le pourcentage de classifications correctes sachant que la vérité est ω_ℓ. Formellement :

$$\mathcal{T}_f[\omega_\ell] = \frac{n_{\ell\ell}}{n_{\cdot\ell}} \,. \tag{12}$$

À partir de ces informations, l'information fournie par la source peut être affaiblie de la manière suivante :

$$^{\alpha}m \;=\; m \circledcup_{\ell=1}^{K} \{\omega_\ell\}_{\mathcal{T}_f[\omega_\ell]} \,. \tag{13}$$

Utilisation d'une distance. Une troisième approche consiste à déterminer le taux de fiabilité de la source en tenant compte pour chaque contexte ω_ℓ, $\ell \in \{1,\ldots,K\}$ de toutes les décisions présentées par la source. Pour ce faire, une distance [12] est proposée permettant de comparer pour chaque contexte ω_ℓ, les différentes décisions ω_k présentées par la source d'information. Le taux de fiabilité $\mathcal{T}_{f_d}[\omega_\ell]$ est alors défini de la manière suivante :

$$\mathcal{T}_{f_d}[\omega_\ell] = 1 - \sqrt{\frac{1}{2}\sum_{k=1}^{K}\left(\frac{n_{k\ell}}{n_{\cdot\ell}} - \delta_{k,\ell}\right)^2} \,, \tag{14}$$

où $\delta_{k,\ell} = 1$ si $k = \ell$, et 0 sinon.

L'information fournie par la source est alors affaiblie en utilisant l'équation (13) avec les taux de fiabilité $\mathcal{T}_{f_d}$.

Tableau 2 – Matrice de confusion d'une source S. Une absence de valeur indique un zéro.

Vérité / Décision	a	b	c	d
a	8	1		
b		6		
c		2	6	1
d		1	4	9

Remarque 3 *Pour chaque vérité ω_ℓ, si seulement un seul nombre de décision $n_{k\ell,\, k\neq\ell}$ est différent de zéro, n_{kk} pouvant être ou non égal à zéro (i.e. $n_{k'\ell} \neq 0$ si seulement si $k' = \ell$ ou $k' = k$), alors les équations (12) et (14) sont équivalentes. En effet, dans ce cas $n_{k\ell} + n_{\ell\ell} = n_{\cdot\ell}$, d'où :*

$$\begin{aligned} \mathcal{T}_{f_d}[\omega_\ell] &= 1 - \sqrt{\tfrac{1}{2}((\tfrac{n_{k\ell}}{n_{\cdot\ell}})^2 + (\tfrac{n_{\ell\ell}}{n_{\cdot\ell}} - 1)^2)} \\ &= 1 - \sqrt{\tfrac{1}{2}((\tfrac{n_{k\ell}}{n_{\cdot\ell}})^2 + (-\tfrac{n_{k\ell}}{n_{\cdot\ell}})^2)} \\ &= 1 - \tfrac{n_{k\ell}}{n_{\cdot\ell}} = \mathcal{T}_f[\omega_\ell] \end{aligned} \tag{15}$$

Exemple 1 *Considérons une source d'information S en charge de reconnaître 4 types d'objets a, b, c et d, dont la matrice de confusion est donnée par le Tableau 2.*

À la lecture de cette matrice, la source sait ainsi bien reconnaître les objets de type a (quand l'objet est de type a la source décide a, ce qui ne veut pas dire que quand la source décide a l'objet soit de type a). À l'inverse, quand l'objet est de type b la source ne le reconnaît que 6 fois, tout comme quand l'objet est de type c. Enfin, la source reconnaît 9 fois sur 10 les objets de type d.

Sur cet exemple, on a $n = 38$, $\mathcal{T}_f = \frac{8+6+6+9}{38} = .76$, et :

$$\begin{array}{ll} \mathcal{T}_f[a] = \frac{8}{8} = 1 & \mathcal{T}_{f_d}[a] = 1 - \sqrt{\frac{1}{2}\ 0} = 1 \\ \mathcal{T}_f[b] = \frac{6}{10} = .6 & \mathcal{T}_{f_d}[b] = 1 - \sqrt{\frac{1}{2}\ .22} = .67 \\ \mathcal{T}_f[c] = \frac{6}{10} = .6 & \mathcal{T}_{f_d}[c] = 1 - \sqrt{\frac{1}{2}\ .32} = .6 \\ \mathcal{T}_f[d] = \frac{9}{10} = .9 & \mathcal{T}_{f_d}[d] = 1 - \sqrt{\frac{1}{2}\ .02} = .9 \end{array}$$

Avec le taux $\mathcal{T}_f$, seul le pourcentage de classifications correctes compte, ainsi la source possède un même degré de fiabilité $.6$ pour les deux types d'objet b et c ($\mathcal{T}_f[b] = \mathcal{T}_f[c] = .6$).

Avec le taux $\mathcal{T}_{f_d}$, la distribution des décisions a une influence, ainsi la source est jugée plus fiable pour les objets de type b par rapport aux objets de type c ($\mathcal{T}_{f_d}[b] = .67$ et $\mathcal{T}_{f_d}[c] = .6$). Intuitivement, la dispersion des décisions entraîne moins d'ambiguïté sur la véritable nature de l'objet.

Supposons que la source fournisse une information m telle que $m(\{a, c\}) = .8$ et $m(\Omega) = .2$. On obtient alors :

- *par l'affaiblissement ${}^\alpha m = m \circledcup \Omega_{.76}$:*

$${}^\alpha m(\{a, c\}) = .61\ ,\ {}^\alpha m(\Omega) = .39$$

- *par l'affaiblissement contextuel (13) avec les taux $\mathcal{T}_f$ (12) ${}^{\boldsymbol{\alpha}} m = m \circledcup \{a\}_1 \circledcup \{b\}_{.6} \circledcup \{c\}_{.6} \circledcup \{d\}_{.9}$:*

$$\begin{aligned} {}^{\boldsymbol{\alpha}} m(\{a, c\}) &= .8(.6.9) = .432 \\ {}^{\boldsymbol{\alpha}} m(\{a, b, c\}) &= .8(.4.9) = .288 \\ {}^{\boldsymbol{\alpha}} m(\{a, c, d\}) &= .8(.6.1) = .048 \\ {}^{\boldsymbol{\alpha}} m(\Omega) &= .8(.4.1) + .2 = .232 \end{aligned}$$

- *par l'affaiblissement contextuel (13) avec les taux $\mathcal{T}_{f_d}$ (14) ${}^{\boldsymbol{\alpha}} m = m \circledcup \{a\}_1 \circledcup \{b\}_{.67} \circledcup \{c\}_{.6} \circledcup \{d\}_{.9}$:*

$$\begin{aligned} {}^{\boldsymbol{\alpha}} m(\{a, c\}) &= .8(.67.9) = .481 \\ {}^{\boldsymbol{\alpha}} m(\{a, b, c\}) &= .8(.33.9) = .239 \\ {}^{\boldsymbol{\alpha}} m(\{a, c, d\}) &= .8(.67.1) = .0535 \\ {}^{\boldsymbol{\alpha}} m(\Omega) &= .8(.33.1) + .2 = .2265 \end{aligned}$$

Sur cet exemple, l'affaiblissement contextuel permet de discriminer les hypothèses b et d : après cet affaiblissement la plausibilité de b

devient plus grande que la plausibilité de d *:* $pl(\{b\}) = .52$ *et* $pl(\{d\}) = .28$ *avec les taux* $\mathcal{T}_f$*, et* $pl(\{b\}) = .47$ *et* $pl(\{d\}) = .28$ *pour les taux* $\mathcal{T}_{f_d}$.

5 Discussion conclusion

Dans ce papier, trois méthodes ont été présentées pour ajuster une information à partir d'une matrice de confusion, ces méthodes étant basées sur la notion d'affaiblissement avec des taux de fiabilité calculés à partir de cette matrice.

On rappelle que la méthode avec laquelle la source fournit une fonction de croyance n'a pas été apprise à partir de la matrice de confusion qui constitue une deuxième information distincte. Dans le cas contraire, il faudrait utiliser des mécanismes plus prudents, par exemple inspirés de la règle prudente de Denœux [11] afin de ne pas compter plusieurs fois la même information.

Pour affaiblir l'information, il a été employé soit un taux de fiabilité globale (un affaiblissement classique), soit un taux de fiabilité pour chaque contexte (un affaiblissement contextuel fin). Or, il serait aussi possible d'employer des taux de fiabilité intermédiaires. Par exemple, en considérant un taux de fiabilité sur $\{a, b\}$ et $\{c, d\}$ dans l'exemple 1. Le choix de cette modélisation dépendra sûrement du nombre de classes d'objets, de la nature des liens existants entre ces classes c'est-à-dire de l'application, et d'une étude d'optimisation qui n'est pas encore réalisée.

On peut aussi remarquer que la notion de confusion entre types d'objets est importante. Par exemple, au regard de la matrice illustrée par le Tableau 2, lorsque la vérité est c ou d, la source commet des erreurs restreintes à des confusions entre c et d. Un affaiblissement contextuel (8) tel que $\{d\} \in \mathcal{A}$ pourra transférer de la masse de c vers $\{c, d\}$ mais il transférera aussi le même pourcentage de masse de a vers $\{a, d\}$, et b vers $\{b, d\}$, ce qu'on ne souhaiterait pas au vue de la confusion limitée à c et d. Pour modéliser cette confusion, il faudrait une autre opération.

Enfin, les affaiblissements proposés consiste en une combinaison disjonctive de la masse fournie par la source avec une masse initialisée au regard de la matrice de confusion. Cette dernière pourrait être initialisée par d'autres méthodes [16, 19] puis fusionnée à l'information fournie par la source par l'intermédiaire d'autres combinaisons. Ainsi l'information pourrait par exemple être renforcée au lieu d'être affaiblie. Ces pistes, de même que l'utilisation de matrices de confusion plus développées (comportant des décisions de rejet, des décisions en faveur de sous-ensembles de $\Omega, \ldots$), sont laissées en perspective.

Références

[1] A. Appriou. Probabilités et incertitude en fusion de données multi-senseurs. *Revue Scientifique et Technique de la Défense*, 11 :27–40, 1991.

[2] A. Appriou. Multisensor signal processing in the framework of the theory of evidence. *NATO/RTA, SCI Lecture Series 216 on Application of Mathematical Signal Processing Techniques to Mission Systems*, RTO EN-7 :(5-1)–(5-31), 1999.

[3] A. Ben Yaghlane, T. Denœux, et K. Mellouli. Elicitation of expert opinions for constructing belief functions. *Proceedings of IPMU'2006*, Paris, France, Juillet 2–7, pages 403–411, 2006.

[4] S. Ben Hariz, Z. Elouedi, et K. Mellouli. Clustering approach using belief function theory, *Proceedings The Twelveth International Conference on Artificial Intelligence : Methodology, Systems, Applications, AIMSA2006*, Varna, Bulgarie, 13-15 Septembre pages 162-171, 2006.

[5] I. Bloch. Some aspects of dempster-shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account. *Pattern Recognition Letters*, 17(8) :905–919, 1996.

[6] I. Bloch. *Fusion d'informations en traitement du signal et des images*. Hermès, Paris, France, 2003.

[7] I. Bloch. Fusion d'informations numériques : panorama méthodologique. *Journées Nationales de la Recherche en Robotique*, Guidel, France, Octobre, pages 79–88, 2005.

[8] A.P. Dempster. Upper and lower probabilities induced by a multiple valued mapping. *Annals of Mathematical Statistics*, 38 :325–339, 1967.

[9] T. Denœux. A k-nearest neighbor classification rule based on dempster-shafer theory. *IEEE Transactions on Systems, Man and Cybernetics*, 25(5) :804–813, 1995.

[10] T. Denœux. A neural network classifier based on Dempster-Shafer theory. *IEEE transactions on Sys-*

tems, Man and Cybernetics A, 30(2) :131–150, 2000.

[11] T. Denœux. Conjunctive and Disjunctive Combination of Belief Functions Induced by Non Distinct Bodies of Evidence. *Artificial Intelligence*, 172 :234–264, 2008.

[12] Z. Elouedi, K. Mellouli et P. Smets. Assessing sensor reliability for multisensor data fusion with the transferable belief model. *IEEE Transactions on System Man and Cybernatics - Part B*, 34, 782-787, 2004.

[13] S. Fabre, A. Appriou et X. Briottet. Presentation and Description of Two Classidication Methods Using Data Fusion Based on Sensor Management. *Information Fusion*, 2 :49–71, 2001.

[14] Z. Hammal, A. Caplier et M. Rombaut. A fusion process based on belief theory for classification of facial basic emotions. *Proceedings of FUSION'2005*, Philadelphie, É-U, Juillet 25–29, paper C1–3, 2005.

[15] É. Lefèvre. *Fusion adaptée d'informations conflictuelles dans le cadre de la théorie de l'évidence, application au diagnostic médicale*. Thèse de doctorat, Institut National des Sciences Appliquées de Rouen, 2001.

[16] D. Mercier, G. Cron, T. Denœux et M.-H. Masson. Fusion of multi-level decision systems using the Transferable Belief Model. *Proceedings of FUSION'2005*, Philadelphie, É-U, Juillet 25–29, paper C8–2, 2005.

[17] D. Mercier, T. Denœux et M.-H Masson. A Parametrized Family of Belief Functions Correction Mechanisms. *Proceedings of IPMU'2008*, J. V.E. L. Magdalena and M. Ojeda-Aciego (Ed.), Malaga, Espagne, pages 306–313, 2008.

[18] D. Mercier, B. Quost et T. Denœux. Refined Modeling of Sensor Reliability in the Belief Function Framework Using Contextuel Discounting. *Information Fusion*, 9(2) :246–258, 2008.

[19] D. Mercier, G. Cron, T. Denœux et M.-H. Masson. Decision fusion for postal address recognition using belief functions. *Expert Systems with Applications*, 36(3) :5643–5653, 2009.

[20] D. Mercier. Extending the contextual discounting of a belief function thanks to its canonical disjunctive decomposition. *1st Workshop on Belief Functions*, Brest, France, avril, paper 61, 2010.

[21] N. Milisavljević et I. Bloch. Sensor fusion in antipersonnel mine detection using a two-level belief function model. *IEEE Transaction on Systems, Man and Cybernetics C*, 33(2) :269–283, 2003.

[22] M. Rombaut. Fusion : état de l'art et perspectives. *Rapport DGA DSP 99.60.078*, LIS UMR 5083 - INPG, 47 pages, 22 octobre 2001.

[23] F. Rottensteiner, J. Trinder, S. Clode, et K. Kubik. Using the Dempster-Shafer method for the fusion of LIDAR data and multi-spectral images for building detection. *Information Fusion*, 6(4) :283–300, 2005.

[24] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J., 1976.

[25] P. Smets and R. Kennes. The Transferable Belief Model. *Artificial Intelligence*, 66 :191–243, 1994.

[26] P. Smets. What is Dempster-Shafer's model ? *Advances in the Dempster-Shafer theory of evidence*, R. R. Yager, M. Fedrizzi and J. Kacprzyk (Ed.), Wiley, New-York, pages 5–34, 1994.

[27] P. Smets. Analyzing the Combination of Conflicting Belief Functions. *Information Fusion*, 8(4) :387–412, 2007.

[28] F. Tupin, I. Bloch et H. Maître. A First Step Towards Automatic Interpretation of SAR Images using Evidential Fusion of Several Structure Detectors. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3) :1327–1343, 1999.

[29] P. Vannoorenberghe. Un état de l'art sur les fonctions de croyance appliquées au traitement de l'information. *Revue I3*, 3(2) :9–45, 2003.

[30] A. Veremme, É. Lefèvre, G. Morvan et D. Jolly. Application of the belief function theory to validate multi-agent based simulations. *1st Workshop on Belief Functions*, Brest, France, avril, paper 123, 2010.

[31] P. Walley et S. Moral. Upper probabilities based only on the likelihood function. *Journal of the Royal Statistical Society*, Series B, 61 :831–847, 1999.

[32] L. Xu, A. Krzyzak et C.Y. Suen. Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Transactions on Systems, Man and Cybernetics*, 22(3) :418–435, 1992.

[33] H. Zhu et O. Basir. Extended discounting scheme for evidential reasoning as applied to MS lesion detection. *Proceedings of FUSION'2004*, Per Svensson and Johan Schubert (Ed.), pages 280–287, juin, 2004.