

Fusion d'images basée sur la modélisation spatiale d'incertitudes et d'imprécisions par la théorie des fonctions de croyance

Benoit LELANDAIS¹, Isabelle GARDIN², Laurent MOUCHARD¹, Pierre VERA², Su RUAN¹

¹LITIS EA 4108 - QuantIF, Université de Rouen
22 boulevard Gambetta, 76183 Rouen Cedex, France

²Département de médecine nucléaire, centre Henri-Becquerel,
1 rue d'Amiens, 76038 Rouen Cedex 1, France
LITIS EA 4108 - QuantIF, Université de Rouen
22 boulevard Gambetta, 76183 Rouen Cedex, France

benoit.lelandais@univ-rouen.fr, isabelle.gardin@rouen.fnclcc.fr
laurent.mouchard@univ-rouen.fr, pierre.vera@rouen.fnclcc.fr
su.ruan@univ-rouen.fr

Résumé – En imagerie, les informations incertaines dues au bruit et imprécises dues à l'effet de volume partiel entraînent des ambiguïtés d'un point de vue spatial. Notre objectif est d'apporter une méthode de traitement d'image ayant un impact différencié entre les incertitudes et les imprécisions. Cette méthode utilise la théorie des fonctions de croyance en tirant partie de sa capacité à considérer non seulement des hypothèses singletons mais aussi des disjonctions. La méthode proposée, opérant sur des images mono-modales, comporte deux étapes de fusion tenant compte de la contribution du voisinage. Elle consiste en la combinaison disjonctive suivie d'une combinaison conjonctive des informations contenues dans les voxels voisins. Son évaluation sur une image simulée et sa comparaison à une approche récemment proposée dans la littérature montrent sa capacité à différencier les incertitudes des imprécisions. De plus, lorsque la fusion mono-modale telle que proposée est effectuée, cela a un effet positif sur la fusion multi-modale.

Abstract – In imaging, uncertain information due to noise and imprecise information due to partial volume effect lead to ambiguities from a spatial viewpoint. Our objective is to provide an image processing approach having a differentiate impact between uncertainties and imprecisions. This method uses belief function theory by taking advantage of its capability to consider not only singletons hypotheses but also disjunctions. The proposed method, processing mono-modal images, takes into account the neighborhood contribution. It consists in the disjunctive combination followed by a conjunctive combination of the information contained in neighboring voxels. Its assessment on a simulated image and its comparison to an approach recently proposed in the literature show its ability to differentiate uncertainties and imprecisions. Moreover, it is shown that when the mono-modal fusion as proposed is used, it has a positive effect on the multi-modal fusion.

1 Introduction

La Théorie des Fonctions de Croyance (TFC) comprend des outils permettant de fusionner plusieurs sources d'informations et est particulièrement bien adaptée à la fusion de données imparfaites par la manipulation de singletons (hypothèses simples) et de disjonctions (hypothèses multiples) [1]. En imagerie, les informations incertaines en raison du bruit et imprécises en raison de l'effet de volume partiel correspondent à des ambiguïtés d'un point de vue spatial. Elles peuvent facilement être identifiées, mais il est plus difficile d'en déterminer l'origine (incertitude et imprécision) pour éliminer le bruit et mettre en évidence les frontières entre régions, sièges de l'effet de volume partiel. Par ailleurs, l'une des difficultés réside dans la modélisation des disjonctions, alors qu'elles permettent de prendre en compte le manque de connaissance. Dans [2], Bloch propose de tirer partie des opérateurs de morphologie mathématique flous

afin de transférer pour chaque voxel, en fonction de son voisinage, une partie de la croyance sur les disjonctions. L'approche est intéressante mais est difficilement applicable lorsque la modélisation initiale porte sur de multiples hypothèses.

Notre but est d'apporter une méthode de traitement d'image ayant un impact différencié entre les incertitudes et les imprécisions à l'aide de la TFC. La méthode effectuée une combinaison disjonctive suivie d'une combinaison conjonctive de chaque voxel avec ses voisins. L'intérêt est, d'une part, qu'une grande partie des incertitudes dues au bruit sont levées, et d'autre part, que l'imprécision à la transition des régions est considérée au niveau des disjonctions. Lors de la combinaison disjonctive, au même titre que lors de la combinaison conjonctive, nous proposons de pondérer l'influence de chaque voisin par sa distance au voxel courant. Cette pondération permet aux voxels éloignés d'avoir une faible influence. Cette méthode est générique puisqu'elle peut être appliquée à partir de croyances réparties sur de

multiples hypothèses, et offre une modélisation pertinente des informations. Notre fusion mono-modale, apportant une modélisation optimisée des informations, permet d'effectuer, par la suite, une fusion multi-modale (fusion des voxels sur plusieurs images) également pertinente.

Dans un premier temps, nous présentons notre méthode, à savoir la fusion mono-modale d'informations pour la prise en compte des incertitudes et des imprécisions. Ensuite, elle est évaluée et comparée à une autre approche [2] sur une image simulée. Enfin, nous l'avons appliquée à la fusion d'images médicales multi-modales.

2 Fusion mono-modale d'informations pour la prise en compte des incertitudes et des imprécisions

À partir de croyances initiales, nous proposons de tirer partie des opérateurs de combinaison issus de la TFC et des informations de voisinage pour réduire les incertitudes dues au bruit dans l'image et pour mettre en évidence les imprécisions dues à l'effet de volume partiel qui se situent à la frontière entre les régions.

La prise en compte de la contribution de voisinage a été réalisée de la façon suivante [3] : Soit $\Phi(V_c)$ un ensemble de voxels V_i voisins d'un voxel V_c , incluant V_c (avec $i = \{1, \dots, N\}$, où N est le nombre de voisins. V_i étant à une certaine distance de V_c , il a été choisi d'associer à chaque voxel de $\Phi(V_c)$ un coefficient α_i fonction de la distance le séparant de V_c .

Dans un premier temps, l'influence d'un voisin V_i de V_c est pondéré par le coefficient α_i . Les masses de croyance sur les hypothèses $A \neq \emptyset$, et \emptyset (l'ensemble vide) incluses dans l'ensemble Ω ont été calculées selon les expressions :

$$\begin{aligned} m'_{V_i}(A) &= \alpha_i m_{V_i}(A), & \forall A \neq \emptyset \\ m'_{V_i}(\emptyset) &= 1 - \alpha_i + \alpha_i m_{V_i}(\emptyset) \end{aligned} \quad (1)$$

Ainsi, la contribution des voxels V_i est d'autant plus faible qu'ils sont éloignés de V_c . Le transfert vers l'ensemble vide s'interprète comme un non-engagement envers l'ensemble des autres hypothèses.

À l'issue de cette étape, et dans le but de transférer les données incertaines et imprécises sur les disjonctions, la combinaison disjonctive [4] suivante est réalisée, pour chacun des voxels :

$$\mathcal{M}_{V_c}(\cdot) = \bigoplus_{V_i \in \Phi(V_c)} m'_{V_i}(\cdot) \quad (2)$$

Ainsi, des masses non nulles sont affectées aux disjonctions. Elles sont d'autant plus élevées que les informations portées par les voxels voisins diffèrent. C'est notamment le cas au niveau des contours des régions et pour les voxels situés dans un environnement fortement bruité.

Dans un second temps, afin de réduire l'incertitude, sans impacter l'ambiguïté portée par l'imprécision, l'opération inverse a été proposée. Il s'agit de la combinaison conjonctive

des voxels voisins V_i après affaiblissement sur Ω en fonction du coefficient α_i :

$$M_{V_c}(\cdot) = \bigoplus_{V_i \in \Phi(V_c)} \mathcal{M}''_{V_i}(\cdot) \quad (3)$$

où $\mathcal{M}''_{V_i}(\cdot)$ est la fonction de masse $\mathcal{M}_{V_i}(\cdot)$ ayant subi un affaiblissement. Il résulte de cette dernière opération une conservation de l'imprécision au niveau des contours (effet de volume partiel) et une réduction de l'incertitude due au bruit sur les voxels au sein des régions. Cette étape permet de lever les ambiguïtés dues au bruit, par transfert de leur croyance sur les singletons, alors que les voxels à la frontière entre régions restent représentés sur les disjonctions.

3 Évaluation

Les performances de notre méthode ont été évaluées sur une image simulée (FIGURE 1), définie dans [2]. Elle compte deux classes, un carré et le fond. Elle a été bruitée par un bruit gaussien d'écart type 0,14. L'objectif est de présenter l'apport de notre approche pour la réduction du bruit et la prise en compte de l'imprécision.

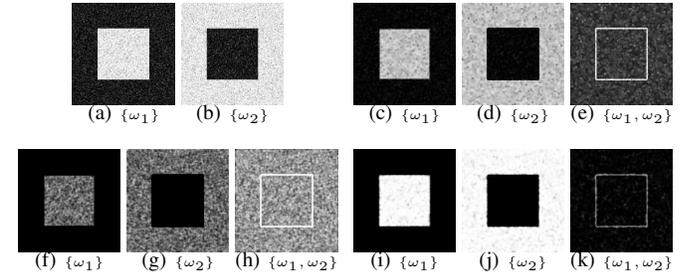


FIGURE 1 – Résultats de la fusion mono-modale sur une image simulée. (a) et (b) présentent les croyances initiales. (c), (d) et (e) présentent la répartition des croyances par la méthode proposée dans [2]. (f), (g) et (h) présentent les croyances affectées aux hypothèses à l'issue de la combinaison disjonctive. Enfin, (i), (j) et (k) présentent les croyances affectées aux hypothèses à l'issue de la combinaison conjonctive (notre méthode).

À partir de la modélisation initiale de l'image effectuée sur deux singletons $\{\omega_1\}$ et $\{\omega_2\}$ (FIGURE 1(a) et (b)), notre méthode a été appliquée. À l'issue de la combinaison disjonctive de chaque pixel avec son voisinage, les masses de croyance sont réparties sur les hypothèses $\{\omega_1\}$, $\{\omega_2\}$ et $\{\omega_1, \omega_2\}$ (FIGURE 1(f), (g) et (h)). La croyance de chaque pixel dont l'information contenue dans son voisinage est ambiguë est principalement représentée sur l'hypothèse $\{\omega_1, \omega_2\}$. Le résultat de la combinaison conjonctive de chaque pixel avec son voisinage est présenté FIGURE 1(i), (j) et (k). Au sein des régions, la méthode fournit des croyances élevées en faveur de $\{\omega_1\}$ et $\{\omega_2\}$. Les incertitudes dues au bruit ont donc bien été réduites. À la transition des régions, la croyance reste principalement représentée sur l'hypothèse $\{\omega_1, \omega_2\}$, permettant de mettre en évidence l'imprécision.

À titre comparatif, nous présentons FIGURE 1(c), (d) et (e) le résultat de la modélisation selon la méthode proposée dans [2]. Cette méthode considère davantage les informations bruitées comme imprécises.

TABLE 1 – Comparaison des performances de bonne classification des pixels, en pourcentages, sur l’image simulée en fonction de la quantité de bruit qu’elle présente (bruit gaussien) et de la taille du voisinage considéré.

rapport signal sur bruit		$\frac{1}{0,1}$	$\frac{1}{0,3}$	$\frac{1}{0,5}$	$\frac{1}{0,7}$
Sans voisinage		100	90,4	71,3	62,3
Taille du voisinage	3 × 3	100	99,7	90,2	75,3
	5 × 5	99,9	99,7	95,8	83,2
	7 × 7	99,9	99,7	99,1	90,3
	9 × 9	99,9	99,2	98,7	95,8

Afin d’évaluer la robustesse de notre méthode, nous avons évalué le pourcentage de pixels correctement étiquetés en fonction du bruit pour différentes tailles de voisinage (TABLE 1). À titre de comparaison, la robustesse de la classification sans prise en compte du voisinage est également donnée. Pour la prise de décision, notons que nous avons maximisé sur les singletons, les croyances converties en plausibilités. Lorsque le bruit devient important, notre méthode donne de bien meilleures performances qu’en l’absence de prise en compte du voisinage. En outre, nous pouvons constater qu’il peut être préférable d’adapter la taille du voisinage en fonction de la quantité de bruit.

4 Application à la fusion multi-modale d’images médicales

La méthode proposée a été à présent appliquée à la fusion multi-modale d’images médicales fonctionnelles TEP (Tomographie d’Émission de Positons) (FIGURE 2). Ces images sont obtenues à l’issue de l’injection au patient d’un traceur spécifique de la fonction que l’on souhaite étudier. À partir des traceurs FDG, FLT et FMISO, trois volumes ont été obtenus chez un patient atteint d’un cancer broncho pulmonaire non à petites cellules. Ils sont respectivement caractéristiques du métabolisme glucidique, de la prolifération cellulaire et de l’hypoxie (manque d’oxygénation des cellules). L’étude de ces trois traceurs est d’un intérêt majeur pour le traitement du cancer bronchique par radiothérapie externe. Le FDG permet une bonne définition des volumes cible tumoraux, notamment ganglionnaires [5]. La FLT présente une meilleure spécificité tumorale que le FDG [6] et permet d’envisager une augmentation de la fréquence des séances d’irradiation par radiothérapie des lésions hyper-prolifératives. Enfin, le FMISO permet de définir les tumeurs hypoxiques pour lesquelles une escalade de dose est envisageable en vue de réaliser ce que l’on appelle le « dose painting » [7].

À l’initialisation, nous avons choisi d’utiliser des modèles

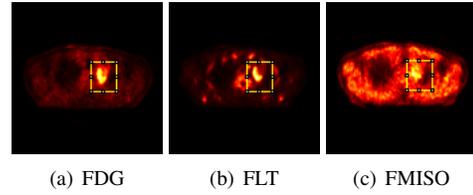


FIGURE 2 – Coupes transverses chez un même patient présentant un cancer pulmonaire. (a) Image TEP du métabolisme du glucose, (b) Image TEP de la prolifération cellulaire, (c) Image TEP de l’hypoxie. La zone d’intérêt (lésion tumorale) est localisée dans le rectangle.

TABLE 2 – Redéfinition des hypothèses afin que les images puissent être fusionnées de manière cohérente.

Image	Hypothèses		
	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_1, \omega_2\}$
FDG	$\{N\}$	$\{M, P, H\}$	$\{N, M, P, H\}$
FLT	$\{N, M, H\}$	$\{M, P, H\}$	$\{N, M, P, H\}$
FMISO	$\{N, M, P\}$	$\{M, P, H\}$	$\{N, M, P, H\}$

flous afin d’affecter à chaque voxel de chaque image des croyances relatives à différentes hypothèses. L’algorithme Fuzzy C-Means [8] a été appliqué au FDG et à la FLT. Concernant le FMISO, présentant un faible contraste, ainsi qu’un faible rapport signal sur bruit, un Z-Score [9] a été calculé en fonction des intensités mesurées dans une région musculaire. Afin que cette mesure soit convertie en masses de croyance, une fonction de répartition gaussienne a ensuite été utilisée. Ainsi, ces modèles permettent l’attribution de masses de croyance à chaque voxel de chaque image selon deux hypothèses $\{\omega_1\}$ et $\{\omega_2\}$. En apportant à la modélisation la contribution du voisinage 3D, les masses de croyance sont définies sur les hypothèses $\{\omega_1\}$, $\{\omega_2\}$ et $\{\omega_1, \omega_2\}$.

De manière à réaliser une fusion pertinente des images, celles-ci doivent être représentées sur un cadre de discernement commun et cohérent, c’est pourquoi nous redéfinissons nos hypothèses à l’issue de la modélisation. Les trois images TEP permettent la distinction de zones distinctes que l’on peut représenter selon quatre singletons, à savoir les tissus sains $\{N\}$ (Normal); ceux présentant un Métabolisme du glucose important $\{M\}$; une Prolifération cellulaire importante $\{P\}$; une Hypoxie importante $\{H\}$. Conformément à l’interprétation médicale des différentes images, leurs hypothèses ont été redéfinies de la manière spécifique présentée TABLE 2.

À l’issue de la redéfinition des hypothèses, les différentes images ont été fusionnées par l’opérateur de combinaison conjonctif [10]. Les masses de croyance sont alors réparties sur de multiples hypothèses. Afin de visualiser le résultat de la fusion sur les singletons qui nous intéressent, les masses de croyance ont été converties en plausibilités, ce qui permet de ne pas tenir compte du fait que les croyances ont été définies sur des hypothèses multiples.

Le résultat de la fusion de ces différentes images médicales

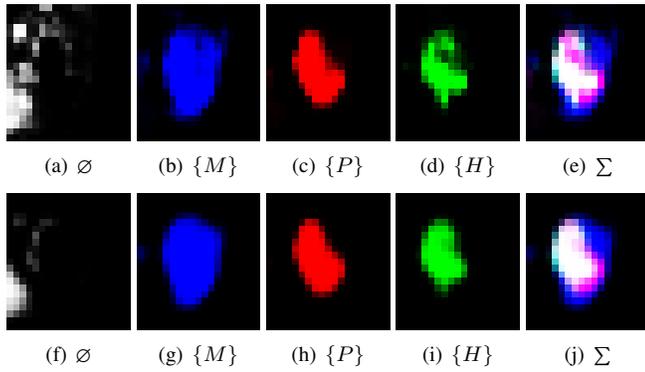


FIGURE 3 – Images présentant le résultat de la fusion des images à l’issue de leur modélisation. (a), (b), (c), (d) et (e) présentent le résultat de la fusion affranchie de la contribution voisine. (f), (g), (h), (i) et (j) présentent le résultat de la fusion tenant compte du voisinage.

est présenté FIGURE 3, sans et avec la prise en compte du voisinage. Nous pouvons remarquer que lorsque le voisinage contribue à la modélisation, les régions sont nettement mieux représentées. En outre, notre fusion permet la visualisation des zones d’intérêt (FIGURE 3(g), (h) et (i)) et la visualisation des zones conflictuelles entre les différentes sources (FIGURE 3(f)). Ces zones conflictuelles correspondent à des hyperfixations en FLT et FMISO et à une absence d’hyperfixation au FDG. Enfin, la FIGURE 3(j) correspond à une superposition des images 3(g), (h) et (i), et témoigne de l’activité totale des voxels. L’ensemble des images paramétriques est une aide précieuse pour le contourage du volume cible tumoral et l’identification de sous-volumes cible nécessitant une intensification du traitement par radiothérapie externe.

5 Discussion-Conclusion

Actuellement, en traitement d’image, très peu d’auteurs ont tenu compte à la fois de l’imprécision et de l’incertitude au niveau spatial dans la modélisation des informations par la TFC [2]. Nous proposons d’effectuer une fusion d’informations de voisinage par une combinaison disjonctive, suivie d’une combinaison conjonctive. Cette méthode permet, de ce fait, la prise en compte des deux types d’imperfection.

Pour la fusion d’informations de voisinage, nous proposons que celle-ci s’opère par une combinaison disjonctive, suivie d’une combinaison conjonctive. Comme nous avons pu le constater d’après les résultats obtenus sur l’image simulée, l’intérêt est que les incertitudes dues au bruit sont en grande partie levées, et que seule l’imprécision au niveau des contours des régions est prise en compte dans la modélisation. En outre, avant la combinaison disjonctive, un transfert des croyances sur l’ensemble vide a été proposé. Il permet, au même titre que l’affaiblissement avant la combinaison conjonctive, de réduire l’influence des voxels éloignés du voxel courant.

Cette méthode automatique a été validée sur des images si-

mulées mono-modales et médicales multi-modales, prouvant sa capacité d’adaptation. La méthode est générique puisqu’elle peut être appliquée quelque soit la répartition des croyances initiales. Elle a été proposée pour tenir compte de l’imprécision au niveau spatiale et pourrait également être appliquée au niveau de l’espace des caractéristiques dans le cadre d’un problème de classification.

Ces résultats très encourageants, doivent cependant être confirmés sur une base de données plus importante afin d’évaluer la robustesse de la méthode, et sur d’autres types d’images afin d’en confirmer le caractère générique.

Références

- [1] G. Shafer, *A Mathematical theory of Evidence*, Princeton university press, 1976.
- [2] I. Bloch, Defining belief functions using mathematical morphology – Application to image fusion under imprecision, *International Journal of Approximate Reasoning*, Vol. 48, No. 2, pp. 437-465, 2008.
- [3] P. Zhang *et al.*, Information Fusion Using Evidence Theory for Segmentation of Medical Images, *International Colloquium on Information Fusion*, Vol. 1, pp. 265-272, 2007.
- [4] P. Smets, Belief functions : The disjunctive rule of combination and the generalized Bayesian theorem, *International Journal of Approximate Reasoning*, Vol. 9, No. 1, pp. 1-35, 1993.
- [5] M.K. Gould *et al.*, Test performance of positron emission tomography and computed tomography for mediastinal staging with non small cell lung cancer : a meta-analysis, *Ann. Intern. Med.*, Vol. 139, No. 11, pp. 879-892, 2003.
- [6] B. Xu *et al.*, Can multimodality imaging using (18)F-FDG/ (18)F-FLT PET/CT benefit the diagnosis and management of patients with pulmonary lesions ?, *Eur. J. Nucl. Med. Mol. Imaging*, 2010.
- [7] W. Choi *et al.*, Planning study for available dose of hypoxic tumor volume using fluorine-18-labeled fluoromisonidazole positron emission tomography for treatment of the head and neck cancer., *Radiother. Oncol.*, Vol. 97, No. 2, pp. 176-182, 2010.
- [8] C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, 1981.
- [9] H. Matsuda *et al.*, Automated discrimination between very early Alzheimer disease and controls using an easy Z-score imaging system for multicenter brain perfusion single-photon emission tomography, *AJNR Am. J. Neuro-radiol.*, Vol. 28, pp. 731-736, 2007.
- [10] P. Smets, The combination of evidence in the transferable belief model., *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 12, No. 5, pp. 447-458, 1990.