
Clustering par fusion floue de données

Application à la segmentation d'images IRM

Abdelouahab Moussaoui *, Abdelmalik Taleb-Ahmed **, & Nabila Frahta ***

** Laboratoire LSI - Université Ferhat Abbas de Sétif
Département d'Informatique, UFAS Algérie*

*** Laboratoire LAMIH UMR CNRS-UVHC 8530 - Université de Valenciennes et du Hainaut
Cambrésis*

Le mont Houy 59313 Valenciennes Cedex 9 France

**** Laboratoire LSI - Université Ferhat Abbas de Sétif*

Département de tronc commun, UFAS Algérie

*moussaoui_abdel@yahoo.fr; abdelmalik.taleb-ahmed@univ-valenciennes.fr;
nabila.frahta@gmail.com*

Sections de rattachement : 27

RÉSUMÉ. Dans cet article nous proposons une nouvelle architecture de fusion de données basée sur la théorie possibiliste pour la segmentation des images médicales. Le processus de fusion est décomposé en trois phases fondamentales. Nous modélisons tout d'abord les informations dans un cadre théorique commun. Le formalisme retenu consiste à faire la coopération entre l'algorithme FCM (Fuzzy C-Means) dont la contrainte d'appartenance d'un individu à une classe est gérée d'une manière relative et l'algorithme possibiliste PCM (Possibilistic C-means) pour les points aberrants. Nous agrégeons ensuite ces différentes informations par un opérateur de fusion. Celui-ci doit affirmer les redondances, gérer les complémentarités et prendre en compte les conflits soulignant souvent la présence d'une pathologie. Nous construisons enfin une information synthétique permettant d'exploiter les résultats de la fusion.

MOTS-CLÉS : Imagerie médicale, Fusion de données, Segmentation, C-moyennes floues, C-moyennes possibiliste.

1. Introduction

Avec le développement des dossiers médicaux informatiques et la généralisation des techniques d'imagerie, il devient possible, pour une pathologie donnée, de disposer d'un grand nombre de données hétérogènes, complémentaires et parfois ambiguës. Le clinicien, analysant ces multiples informations, opère une agrégation de celles-ci, en fonction de jugements subjectifs et approximatifs fondés sur sa propre expérience. Le but de ce raisonnement est de synthétiser un état de la pathologie le plus complet possible, par exemple pour proposer un diagnostic, établir un pronostic ou même élaborer une aide à l'intervention chirurgicale [1].

Ces dernières années, des modélisations formelles de cette attitude ont été construites, fondées pour la plupart sur des approches prenant en compte les redondances, les complémentarités et les ambiguïtés inhérentes aux données médicales. Regroupées sous l'appellation "*fusion*", ces modèles ont pour but de gérer au mieux ces différents aspects pour faire converger les connaissances et proposer une information synthétique la plus exploitable possible [1].

Notre travail concerne essentiellement le développement de nouveaux outils issues des techniques de *data mining* pour l'extraction des connaissances par fusion floue de données. Il s'agit essentiellement de contribuer au développement de systèmes de classification guidés par les connaissances *a priori* où l'aspect flou et possibiliste sont pris en considération lors du processus de classification. Notre travail consiste à proposer une architecture de fusion de données guidée par ces connaissances *a priori*. Afin de valider les algorithmes développés, une application a été développée pour la segmentation des images IRM.

Le processus de fusion tel que nous l'envisageons ici est composé de trois étapes. Dans la première, les informations disponibles sont modélisées dans un cadre théorique commun, permettant de prendre en compte les connaissances vagues et ambiguës. Dans la seconde, les modèles d'informations sont agrégés, en tenant compte des redondances et des conflits exprimés. Dans la troisième, enfin, une décision est prise en fonction de toutes les informations précédemment fusionnées [2].

2. Approche proposée

La fusion d'images peut alors se décomposer en trois grandes étapes :

- *Modélisation des informations dans un cadre théorique commun.*
- *Fusion des informations issues de la modélisation précédente.*
- *Prise de décision.*

2.1. Modélisation

2.1.1. Choix des algorithmes utilisés

Nous devons, tout d'abord, préciser si l'algorithme doit être supervisé ou non. L'emploi d'un algorithme supervisé nécessite, comme nous l'avons vu précédemment, une base d'apprentissage pour chaque classe et pour chaque patient.

Cela constitue un premier inconvénient de ce type de méthodes. De plus, la littérature rapporte que les méthodes supervisées en segmentation d'images médicales peuvent être très dépendantes de la base d'apprentissage [3]. Pour toutes ces raisons, nous avons opté pour une méthode non supervisée.

2.1.1.1. FCM et degrés d'appartenance relatifs

La modélisation de l'imprécision s'effectue en considérant des frontières graduelles au lieu de frontières nettes entre les classes [4][5][6]. L'incertitude s'exprime par le fait qu'un pixel possède aussi bien des attributs qui l'assignent à une classe qu'à une autre :

$$\forall i \in \{1..C\}, \forall j \in \{1..N\} \begin{cases} u_{ij} \in [0,1] \\ 0 < \sum_{j=1}^N u_{ij} < N \end{cases} \quad [1]$$

$$\forall j \in \{1..N\} \quad \sum_{i=1}^C u_{ij} = 1 \quad [2]$$

La contrainte de normalisation (2) utilisée pour la minimisation de la fonctionnelle est source d'erreur dans l'interprétation des degrés d'appartenance issus du FCM [3]. Krishnapuram et Keller [7] donnent une série d'exemples simples qui illustrent les problèmes associés à cette contrainte.

2.1.1.2. PCM et degrés d'appartenance absolus

La valeur de u_{ij} dans (1) ne dépend maintenant que de la distance du vecteur x_j à la classe i , contrairement à (2). Les degrés d'appartenance générés par PCM ne sont ainsi plus des degrés relatifs ou de partage, ils deviennent des valeurs absolues reflétant la force avec laquelle chaque vecteur appartient à toutes les classes.

Les images IRM sont affectées par des artefacts que l'algorithme PCM les gère mieux [6]. Pour toutes ces raisons, nous nous sommes orientés vers l'algorithme de classification possibiliste pour caractériser les tissus cérébraux.

2.1.2. Initialisation de l'algorithme

Les résultats de l'algorithme PCM vont évidemment dépendre de l'étape d'initialisation [3][4]. Ainsi, l'initialisation, si elle doit exister, ne doit pas être aussi précise que dans d'autres algorithmes comme FCM. Tout algorithme (flou ou non) peut

donc être utilisé, et l'algorithme FCM constitue une manière adéquate d'initialiser les données [5].

2.1.3. *Détermination du nombre de classes*

Notre objectif consiste à segmenter l'encéphale, ce qui nous pousse à fixer à 3 le nombre de classe à identifier ($C=3$) correspondant aux trois tissus cérébraux présents dans l'encéphale à savoir la matière blanche (MB), la matière grise (MG) et le liquide céphalo-rachidien(LCR).

2.1.4. *Choix des vecteurs forme*

Le choix des vecteurs forme est fondamental puisque leur pertinence va permettre de discriminer les pixels entre eux. Ce choix est défini suivant le type de modalité. L'image anatomique que nous utilisons est une IRM. L'imagerie par résonance magnétique est une modalité d'imagerie multispectrale (multimodale) donnant accès à un grand nombre de paramètres et donc de vecteurs forme. La première caractéristique qui peut être exploitée est le signal lui-même, principalement par l'intermédiaire d'images pondérées en T_1 , T_2 et en densité de protons. Le vecteur forme x_j d'un pixel j est alors formé des niveaux de gris de ce pixel dans toutes les modalités. Cette information est très largement utilisée en segmentation d'images [6].

2.2. *Fusion*

De nombreux auteurs ont comparé les théories des probabilités, des possibilités et des croyances, et ont détaillé les transformations permettant de passer d'un formalisme à l'autre [6][8][9].

2.2.1. *Limitations de la fusion probabiliste*

Notre choix ici est exclusivement orienté vers le domaine de l'imagerie. La fusion quant à elle est réalisée à l'aide de la règle de Bayes, et la décision est prise en fonction du maximum de vraisemblance. Cependant, cette théorie présente certains inconvénients qui limitent son utilisation dans le cas qui nous intéresse [10][11]. Ces limites sont résumées ci-dessous : dans les images traitées dans la partie suivante, les informations sont à la fois incertaines et imprécises. Or la théorie des probabilités ne peut pas bien prendre en compte l'imprécision des données, le formalisme, notamment introduit dans la règle de Bayes, requiert des connaissances a priori sur l'occurrence de chaque phénomène par l'intermédiaire des probabilités conditionnelles et des probabilités a priori des événements[12][13][14][15]. Des modèles peuvent représenter ces connaissances mais imposent alors des hypothèses fortes sur les étapes de modélisation et de fusion.

Nous avons comparé les théories des croyances et des possibilités par rapport au problème spécifique de la fusion d'images médicales, cette comparaison permet de se rendre compte que des différences importantes existent, ces différences sont :

La théorie de l'évidence prend en compte les ensembles composés de plusieurs classes, ce qui permet de considérer le doute entre les classes.

L'analyse de ces modèles pour les fusions envisagées a révélé tout d'abord que la théorie des probabilités était mal adaptée, principalement en raison des effectifs d'expérimentation faibles dont nous disposons. La relative simplicité des informations à extraire dans les images et la pauvreté des modes de combinaison en théorie des croyances ont été ensuite pour nous des éléments décisifs dans le choix d'un cadre formel et nous ont fait préférer l'approche possibiliste à la théorie des croyances.

2.3. *Décision*

La dernière étape consiste à prendre la décision quant à l'appartenance d'un voxel de l'image I à une classe C_i . La règle de décision adoptée consiste à prendre une coupe de la carte de fusion en choisissant un seuil d'étiquetage.

$$\forall v \in I, \quad v \in C_i \quad \text{si} \quad \pi_{\text{Fusionné}}(v) \geq \text{seuil} \quad [3]$$

3. **Résultats**

Afin de prouver l'efficacité de l'approche développée, nous avons implémenté les algorithmes FCM, PCM et l'approche développée. Les images réelles sur lesquelles nous avons travaillé ont été acquises dans le cadre de la collaboration entre le laboratoire LSI (Laboratoire Systèmes Intelligents : équipe image et signaux) de l'université Ferhat Abbas de Sétif et le l'hôpital universitaire de la ville de Sétif. Il s'agit d'images pondérées en T1, T2 et en densité de proton (Dp) pour des patients de différents âges (taille pixel =1mm, taille de matrice 512 X 512). Les images sont en format DICOM (Digital Imaging and Communications in Medecine).

La figure 1 illustre les résultats obtenus après segmentation des différents tissus cérébraux d'images (a) et (b) pondérées en T1 et en T2 sur des coupes axiales. Les images (c), (d), (e) et (f) sont les résultats de la segmentation par FCM et PCM successivement.

Dans ces images (c, d, e, f) la distinction entre les différentes classes segmentées ne s'exprime pas complètement (problèmes d'initialisations des centres des classes), contrairement aux images (g), (h) issues de la segmentation par le système développé où on remarque la distinction entre les 3 classes du cerveau et la visibilité de la classe LCS, (invisible dans les deux approches précédentes). Une comparaison entre les images (c),

(d), (e), (f) et les images (g), (h) permet de voir l'apport de l'approche proposée par rapport aux algorithmes FCM et PCM. Les images issues de l'approche proposée et le PCM caractérisent bien les différentes classes cérébrales MB, MG, LCR. La zone entourée par le contour bleu représente une nouvelle classe non détectée par FCM et PCM.

Les performances de nos algorithmes nous ont conduits à réfléchir sur la validité de la segmentation obtenue. Il n'existe pas de "bonne" segmentation [16]. Seule l'appréciation de l'utilisateur (qualité visuelle) et le but recherché permettent de définir une bonne segmentation pour un type de données. Il nous a paru utile de mesurer et de quantifier les performances de notre segmentation de l'ensemble de l'encéphale.

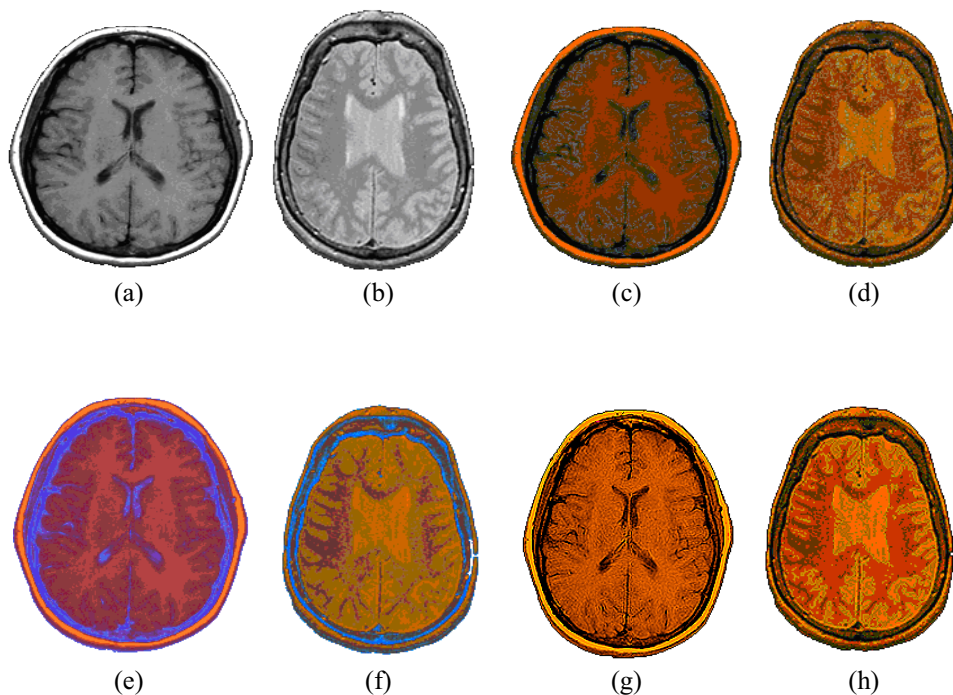


Figure 1. Résultats de segmentation

Nous avons préféré utiliser l'un des coefficients décrits dans [16], définis pour valider des systèmes d'imagerie. Le coefficient proposé est le taux de recouvrement (TR). Les résultats du tableau 1 confirment l'intérêt de l'approche proposée (floue possibiliste) par rapport aux approches prises indépendamment.

	FCM	PCM	Approche proposée
MG	0,85	0,64	0,90
MB	0,9	0,52	0,92
LCR	0,63	0,65	0,90

Tableau 1. Comparaison des taux de recouvrement obtenus par différents algorithmes

Les résultats du tableau 1 confirment l'intérêt de l'approche de fusion par rapport aux approches prises indépendamment.

4. Conclusion

Dans cet article nous avons démontré l'importance et les difficultés de ces travaux de segmentation d'images IRM.

En premier lieu, nous avons modélisé des informations, numériques ou symboliques, dans un cadre commun permettant de prendre en compte les ambiguïtés, les imprécisions et les incertitudes. Nous nous sommes pour cela placés dans le cadre de la coopération entre l'algorithme flou FCM et l'algorithme possibiliste PCM afin de rendre l'algorithme plus robuste face aux imprécisions et aux données aberrantes.

En second lieu, nous avons fusionné différentes données. Cette agrégation a été réalisée par des opérateurs de fusion qui modélisent l'analyse quotidienne du médecin confronté à des données cliniques hétérogènes. Nous avons explicité l'opérateur qui nous a semblé le plus adapté, en fonction de propriétés mathématiques souhaitées, de considérations intuitives sur la nature des données et de tests expérimentaux.

En dernier lieu, nous avons présenté ces informations fusionnées au clinicien. Ici encore, nous avons proposé pour chaque type de fusion une solution, soit sous la forme d'une image, soit sous la forme d'une nouvelle image de synthèse.

Nous avons présenté les résultats de notre travail qui consiste à utiliser plus d'un algorithme pour segmenter des images médicales en vue d'améliorer la qualité de la segmentation.

Bibliographie

- [1] V Barra V, « Fusion d'Images 3D du Cerveau : Etude de Modèles et Applications », Ph.D. Thesis, Université d'Auvergne, Clermont-Ferrand (France), 2000.
- [2] W Dou, « Segmentation d'images multi spectrales basée sur la fusion d'informations : application aux images IRM », Thèse doctorat de l'université de Caen, 2006.
- [3] J Bezdek, L Hall, L Clarke, « Review of MR Image Segmentation Techniques using

- Pattern Recognition”, *Medical Physics*, 20:1033-1048, 1993.
- [4] Moussaoui A, «Aneuro-fazzy image segmentation», *International journal of soft computing* 1(3):232-238 Medwell Online. 2006
 - [5] Raff U, Scherzinger A, Vargas P, Simon J, «Quantitation of grey matter, white matter and cerebrospinal fluid from spin-echo magnetic resonance images using an artificial neural network Technique», *Medical Physics*, 1994
 - [6] Barra.V, «Segmentation floue des tissus cérébraux en IRM 3D: une approche possibiliste versus autres méthodes». Thèse doctorat. Universités Blaise Pascal 1999.
 - [7] Krishnapuram R., «The Possibilistic C-Means Algorithm: Insights and Recommendations», *IEEE transactions on Fuzzy Systems* 1996.
 - [8] I Bloch, «Fusion de Données en traitement d’images : modèles d’information et décision », *Traitement du signal*, 1994, vol.11, n°6, pp 435 - 446.
 - [9] I Bloch, «Information Combination Operators for Data Fusion: A Comparative Review with Classification», *IEEE Transactions on Systems, Man, and Cybernetics*, 1:52-67, 1996.
 - [10] Zadeh L, «Fuzzy Sets as a Basis for Theory of Possibility», *International Journal of Fuzzy Sets and Systems*, 1:3-28. 1978 [11]
 - [11] D Dubois D, H Prade , «Possibility Theory, an approach to the computerized processing of uncertainty», *Plenum Press*, 1988.
 - [12] G Shafer, « A Mathematical Theory of Evidence», *Princeton University Press*, 1976.
 - [13] A P Dempster, «Upper and lower probability function in a context of uncertainty», *Annals of math. Statistics*, vol 38, pp. 325-339, 1967.
 - [14] A P Dempster, «A generalization of bayesian inference», *Jour. of the Royal Statistical Society*, vol 30, pp. 205-247, 1968.
 - [15] L Zadeh , «Fuzzy Sets as a Basis for Theory of Possibility, *International Journal of Fuzzy Sets and Systems*», 1:3-28, 1978.
 - [16] Shufelt J.A. «Performance Evaluation and Analysis of Monocular Building Extraction from Aerial Imagery». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4), pp. 311–326. 1999.