

Fusion de l'information : fusion de données et de modèles appliqués à la segmentation d'images écho-endoscopiques

Renaud DEBON¹, Bassel SOLAIMAN¹, Jean-Michel CAUVIN² et Christian ROUX¹

¹ENST de Bretagne, dépt. ITI, LaTIM (EA-2218)
BP 832, 29285 Brest Cedex, France

²Dépt. d'information médicale du
CHRU de Brest, LaTIM (EA-2218), 29200 Brest, France

Renaud.Debon@enst-bretagne.fr, Basel.Solaiman@enst-bretagne.fr
Cauvin@univ-brest.fr, Christian.Roux@enst-bretagne.fr

Résumé – De nos jours, la fusion de l'information est un sujet en pleine effervescence. Les applications de segmentation d'images médicales, auxquelles nous nous intéressons, nécessitent la plupart du temps d'effectuer la fusion de différentes sources de connaissances. Après une présentation générale de la fusion de l'information et des architectures correspondantes, cet article expose une méthode de détection de la paroi interne de l'œsophage à partir d'images écho-endoscopiques, basée sur la fusion de données et de modèles. La fusion de données repose sur la logique floue de telle sorte que le moteur de traitement peut être perçu comme un système monocapteur / multisource de fusion. La fusion de modèles est réalisée grâce à une utilisation des contours actifs adaptée aux spécificités des images traitées et à l'architecture du système de traitement choisie afin d'aboutir à une décision binaire, c'est-à-dire à la détection de la paroi œsophagienne.

Abstract – Nowadays, information fusion constitutes a challenging research topic. Our study proposes to achieve the fusion of several sources of knowledge, in order to detect the esophagus inner wall, from ultrasound medical images. After a brief description of information fusion concepts, we propose a system architecture including both model and data fusion. The data fusion is accomplished using fuzzy modeling, which can be seen as a mono-sensor / multiple sources data fusion system. The model fusion is performed using a full-adapted snake theory, which projects the fuzzy decision into the binary decision space.

1. Introduction

La fusion de l'information est un sujet relativement ancien qui trouve ses origines à partir du moment où les chercheurs ont fait leurs premières tentatives d'imitation de l'intelligence humaine. Cet article tourne autour de deux idées principales : la fusion de l'information et ses principaux concepts d'une part, les architectures de la fusion d'autre part, pour lesquels nous distinguons deux familles dominantes qui sont les architectures multicapteurs et les architectures monocapteurs / multisources d'information. L'ensemble de ces considérations et concepts généraux ont été pleinement utilisés dans ce qui constitue le cœur de nos travaux : le traitement d'images médicales et en particulier le traitement des images écho-endoscopiques de la région œsophagienne [1][2]. Cette technique d'imagerie relativement récente est connue pour être un outil extrêmement puissant, permettant d'obtenir une amélioration spectaculaire de la qualité des diagnostics des tumeurs de l'œsophage et d'augmenter la précision de leurs traitements.

Ce travail traite d'une méthode de détection automatique de la paroi interne de l'œsophage basée sur la fusion de l'information [3]. Cette détection présente un intérêt majeur : c'est la première étape de l'élaboration d'un outil d'analyse de la paroi et des sous-couches qui la constitue, le but étant à terme, de réaliser une reconstruction globale des éléments anatomiques utiles à l'élaboration d'un diagnostic médical.

Le corps de cet article est divisé en deux parties principales : la section 2 discute de généralités sur la fusion et des principes sur lesquels reposent nos travaux. La section 3 détaille l'architecture d'un système automatique de détection de la paroi interne de l'œsophage. Des résultats sur des cas cliniques sont présentés et commentés en section 4. Conclusions et perspectives suivent en section 5.

2. Concepts de la fusion d'information

2.1 Classification des types de fusion

Une formalisation possible de la fusion d'information introduit trois niveaux conceptuels correspondant à trois types d'information :

- **La fusion de données**- c'est le niveau conceptuel le plus bas. Elle consiste essentiellement à marier des informations de bas niveau comme par exemple des primitives, dans le but de rendre l'information moins bruitée que celle obtenue avec une seule source d'information.
- **La fusion de décisions**- ce type de fusion agit au niveau de l'espace de décision. Elle effectue l'association d'informations élaborées (numériques ou symboliques) qui peuvent être considérées comme des propositions de décision.
- **La fusion de modèles**- ce cas est celui dans lequel les aspects complémentaires de différentes approches sont mis à partie pour combler les imperfections dont souffrent chacune d'elles indépendamment. Dans [4], le problème de la

détection de contours est abordé par le biais d'un algorithme de Canny-Deriche combiné avec un réseau de neurones qui permet d'en régler les paramètres. De manière analogue, une modélisation markovienne coopère avec un réseau de neurones dans [5].

2.2 Les architectures de la fusion

Cette partie a pour objectif de rappeler les deux principaux schémas de la fusion de l'information. De par des raisons historiques, les premiers systèmes de fusion à avoir été mis en place ont été de types multicapteurs. C'est le cas intuitif dans lequel l'objet est observé par plusieurs capteurs physiques (ou bien le même mais selon différents angles d'observation comme c'est le cas en stéréovision). Dans de tels schémas, les capteurs physiques sont employés pour accéder à différentes informations issues d'un ou de plusieurs objets d'une scène réelle. Cependant en toute généralité, d'autres types d'architecture peuvent être envisagés, comme par exemple les systèmes de fusion monocapteurs. De tels systèmes sont basés sur l'utilisation d'un seul capteur physique. Des informations nouvelles sont extraites de l'information brute initiale, via des capteurs fictifs définis à partir de connaissances *a priori*. L'utilisation de la théorie des probabilités ou bien de la théorie des ensembles flous est généralement appliquée à ce niveau pour réaliser la fusion.

3. Détection de la paroi œsophagienne

Comme cela a déjà été mentionné, le but de cette étude est de réaliser la détection de la paroi interne de l'œsophage à partir de coupes acquises à l'aide d'un système d'imagerie écho-endoscopique. L'instrument utilisé est un écho-endoscope à vision latérale Olympus EU-M3 fonctionnant à 7.5Mhz et 12Mhz. Le capteur est composé d'un endoscope classique muni d'un transducteur ultrasonore. L'appareil est introduit par la bouche du patient puis progresse dans l'œsophage en direction du cardia, zone de confluence avec l'estomac. Le transducteur émet des trains d'ondes ultrasonores dans toutes les directions transversalement à sa direction de progression. La reconstruction d'une image se fait par analyse des échos reçus.

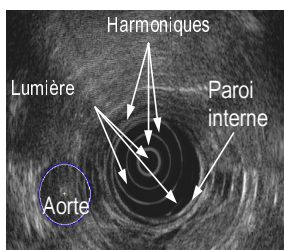


FIG. 1 : Exemple de coupe écho-endoscopique

L'observation des images écho-endoscopiques (FIG. 1) nous permet de conclure que les difficultés rencontrées dépendent principalement de deux phénomènes : le bruit speckle, du fait que ce mode d'imagerie est basé sur l'utilisation des ultrasons, et les réflexions multiples concentriques, appelées «harmoniques», qui proviennent de la gaine protectrice entourant le transducteur. D'autres difficultés peuvent apparaître, mais la plus notable est le

phénomène d'atténuation des parois œsophagiennes quand des tissus précancéreux ou cancéreux sont traversés.

3.1 Architecture générale

L'architecture proposée est composée de deux modèles radicalement différents - La modélisation floue et le modèle dynamique - qui apportent chacun des éléments permettant d'aboutir à la solution (FIG. 2). La modélisation floue est basée sur la fusion d'information monocapteur / multisource. Elle consiste en une extraction de sources d'information, en leur «fuzzification» et finalement en leur fusion. De cette dernière étape, qualifiée de raisonnement flou, résulte une décision floue. Le modèle dynamique est basé sur des efforts qui seront énoncés par la suite. Les spécificités et les avantages de chacun de ces modèles sont exploités via une fusion de modèle permettant d'aboutir à la solution.

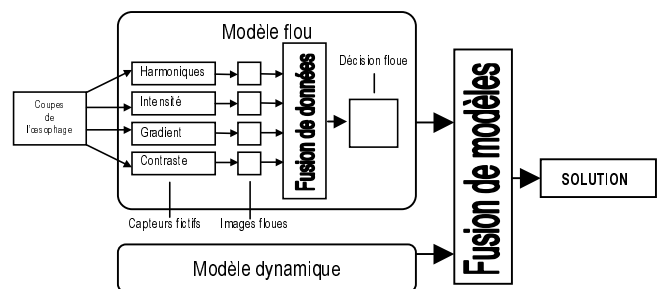


FIG. 2 : Architecture générale du système de traitement

3.2 Modélisation floue

La théorie des ensembles flous qui a été pour la première fois introduite par L. Zadeh nous procure un outil mathématique puissant pour modéliser la capacité humaine à fournir une décision dans le cas où les informations disponibles sont imprécises ou incomplètes [6][7]. Introduisons dans un premier temps le concept d'image floue. Une image floue est définie comme la transformation A d'une image en niveau de gris (c'est-à-dire une matrice $M \times N$) en une image de dimension égale mais dans laquelle chacun des niveaux de gris associé à un pixel P dénote le degré d'appartenance $\mu_A(P)$ à une propriété donnée, comme le traduit l'expression suivante :

$$A: \begin{matrix} M \times N & \rightarrow & [0,1] \\ P(x,y) & \rightarrow & \mu_A(P) \end{matrix} \quad (1)$$

Image floue des harmoniques : Avant de commencer tout traitement sur les images écho-endoscopiques, les caractéristiques des harmoniques, c'est-à-dire leur position aussi bien que leur niveau de gris, doivent impérativement être connus pour simplifier les traitements en aval. La position des harmoniques est estimée par l'utilisation d'un modèle prédéfini [1][8]. L'image floue peut alors être construite de la façon suivante : la valeur d'appartenance $\mu_h(P)$ est arbitrairement fixée à 0.3 pour les pixels de l'harmonique et à 1.0 pour les autres.

Image floue de brillance : La paroi de l'œsophage est en général hyper-échogène. De ce fait, les pixels qui la constituent ont un fort niveau de gris. Ainsi, un pixel reconnu brillant a des chances d'appartenir à la paroi œsophagienne même si le critère de brillance n'est pas absolu. Une fonction

S-shape [8] est appliquée sur l'image de niveau de gris, dans le but de construire l'image floue de brillance $\mu_b(P)$. L'emploi d'une telle fonction peut être considéré comme un processus de normalisation des niveaux de gris. Elle présente l'intérêt certain de rendre le critère indépendant des réglages de gain et de contraste du système d'acquisition. Les paramètres de la fonction S-shape sont définis comme suit, FIG. 3 :

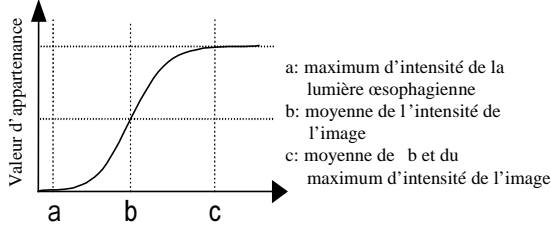


FIG. 3 : Paramètres de la fonction S-shape

Image floue de contours : L'information sur les contours est un élément prépondérant dans la détection de la paroi interne de l'œsophage. L'image floue de gradient, $\mu_g(P)$, représentant le degré d'appartenance de chaque pixel P à l'ambigu concept de contour est du plus grand intérêt. Un filtre «gradient» 5×5 similaire à un filtre de type «Sobel» est utilisé comme détecteur de contours. Après filtrage, l'étape de «fuzzification» est accomplie grâce à une fonction S-shape.

Image floue de contraste entre région : Lorsque l'on examine attentivement des images écho-endoscopiques, on note immédiatement le fort contraste qui existe entre la lumière œsophagienne (c'est-à-dire la région sombre du centre de l'image) et la paroi de l'œsophage (première région hyper-échogène rencontrée). Cette caractéristique intéressante est exploitée par la construction d'une image floue de contraste $\mu_r(P)$, dénotant qu'un pixel possède la propriété d'être à la frontière de la lumière œsophagienne.

Combinaison floue : L'étape de combinaison floue consiste à fusionner l'information utile contenue dans les images floues définies précédemment. Le résultat est une décision floue μ_p dans laquelle chaque niveau de gris correspond au degré d'appartenance d'un pixel à la paroi interne de l'œsophage. Dans la littérature ayant trait à la théorie des ensembles flous, existe un large champ d'opérateur de combinaison [9]. Dans notre étude, un simple opérateur d'intersection (c'est-à-dire un opérateur conjonctif) a été employé :

$$\mu_p(P) = \text{Min}(\mu_h(P), \mu_b(P), \mu_g(P), \mu_r(P)) \quad (2)$$

Dans l'expression (2), $\mu_p(P)$ est le degré d'appartenance du pixel P à la paroi interne de l'œsophage. Plusieurs arguments justifient le choix d'un opérateur conjonctif :

- Un opérateur de cette classe réduit, dans le résultat de la combinaison, le bruit ou d'éventuels artefacts présents dans une ou plusieurs des image floues.
- Une haute fiabilité de l'information est obtenue lorsque les sources à combiner ne sont pas contradictoires. En effet, une valeur d'appartenance élevée dénote qu'une information a été confirmée par tous les capteurs fictifs correspondant aux propriétés considérées. En conséquence, la certitude des informations est meilleure avec ce genre d'opérateur. Le risque d'obtenir des contours discontinus sera évité par l'utilisation conjointe de l'approche dynamique décrite dans le paragraphe suivant.

3.3 Contour actif

Le modèle de type «contours actifs» qui vient compléter avantageusement l'approche floue décrite précédemment permet l'obtention d'une décision binaire. Faire coopérer deux modélisations pour aboutir à la résolution d'un problème donnée, peut être considéré comme une fusion de modèles. L'approche employée dérive de celle déjà considérée par Kass, Witkin et Terzopoulos [10]. Mathématiquement, un contour dynamique peut être défini comme une courbe paramétrée :

$$\begin{aligned} \Omega : [0;1] &\rightarrow \mathfrak{R} \\ s &\rightarrow v(s) = (x(s), y(s)) \end{aligned} \quad (3)$$

Cette courbe évolue dans un champ de potentiel défini par l'image étudiée. La position stable du «snake» dans le champs de potentiel correspond au minimum de l'énergie totale qui est définie comme suit :

$$E = \int_0^1 E_{\text{interne}}(v(s)) + E_{\text{image}}(v(s)) ds \quad (4)$$

Dans l'expression (4), E_{interne} est l'énergie due aux efforts appliqués au contour dynamique, et E_{image} est le champ de potentiel crée par l'image étudiée.

Avant tout traitement, une transformation cartésien-polaire est opérée sur les coupes écho-endoscopiques. Cette opération présente l'intérêt de simplifier considérablement les traitements (par exemple, les cercles concentriques des harmoniques deviennent des droites). Le modèle dynamique proposé peut être considéré comme un ensemble de N nœuds interconnectés de façon élastique. Les nœuds M_i sont astreint à circuler sur des rails parallèles, sur une image en coordonnées polaires. De ce fait, un seul degré de liberté est autorisé comme le montre la FIG. 4. Le contour dynamique, se déplace grâce à l'action de forces externes et de forces internes appliquées aux nœuds M_i . Ces forces permettent respectivement de positionner le modèle pour les premières, et d'imposer des contraintes de régularité définies à partir de connaissances *a priori* [11] pour les secondes.

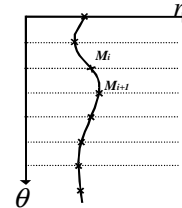


FIG. 4 : Schéma général du modèle dynamique

Les forces externes simulent l'espace physique. Ces efforts contraignent les déplacements du «snake». Elles sont définies comme suit :

- Une force de pression F_p qui pousse le modèle dans la direction du contour recherché, et supprime en même temps le problème d'initialisation généralement rencontré avec les modèles dynamiques.
- Une force de réaction R qui simule le fait que chaque nœud est contraint de se déplacer sur un rail.
- Une force de frottement fluide F_v , dans laquelle le coefficient de frottement γ dépend des variations locales du champ de «potentiel image».
- Une force F_{im} dérivant du «potentiel image», c'est-à-dire de la décision floue définie précédemment.

Les forces internes permettent le contrôle de la rigidité du «snake» et de sa forme générale. Elles sont les suivantes :

- a) Une force F_{el} représentant l'élasticité du modèle.
- b) Une force F_c dont le but est de rendre le «snake» résistant à la courbure.

Considérant le système régi par les équations de la dynamique et les nœuds dotés d'une masse nulle, une discrétisation utilisant un pas temporel τ permet d'estimer la position d'un nœud M_i de la façon suivante:

$$r_i(t + \tau) = r_i(t) + \frac{\tau}{\gamma_i} (F_{pr} + F_{imr} + F_{elr} + F_{cr}) \quad (5)$$

4. Résultats sur des images réelles

Les résultats présentés en FIG. 5 et FIG. 6 concernent une coupe typique d'une région affectée par une pathologie d'origine cancéreuse. Sur l'image originale, FIG. 6-a, nous pouvons immédiatement remarquer que le contour de la paroi interne disparaît complètement à plusieurs endroits. Sur ces images et en particulier sur l'image originale, on prend immédiatement conscience de la gêne qu'occasionne la présence des harmoniques. Sur la FIG. 5, sont présentées les quatre images floues correspondantes aux nouvelles sources d'information extraites à partir de l'image initiale, et de connaissances *a priori*.

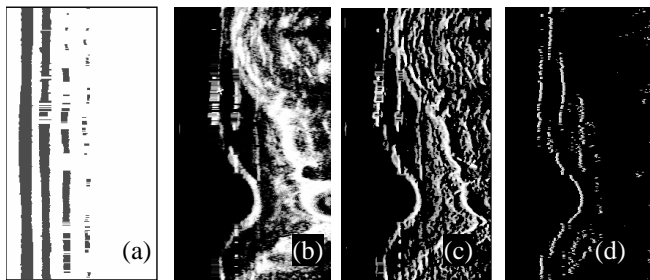


FIG. 5 : Images floues calculées à partir de la FIG. 6-a : harmoniques (a), brillance (b), gradient (c) et contraste entre régions (d).

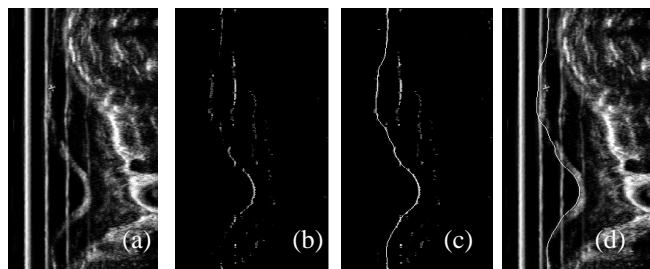


FIG. 6 : Cas pathologique en coordonnées polaires (a), décision floue (b), modèle stabilisé sur la décision floue (c), et modèle positionné sur l'image originale (d).

Le résultat de l'application de l'opérateur de combinaison employé peut être apprécié sur la FIG. 6-b. Comme nous pouvons le noter, les valeurs d'appartenance des pixels de la paroi interne sont très élevées. En conséquence, il est possible de conclure que les connaissances extraites constituent un ensemble d'information suffisant pour isoler les pixels de la paroi des pixels des autres tissus.

5. Conclusions

Cette méthode pourrait constituer une solution robuste à la détection de la paroi interne de l'œsophage, même dans des cas pathologiques. Il arrive parfois que des imprécisions de un ou deux pixels puissent être notées au voisinage des harmoniques, mais la méthode semble pleinement satisfaisante dans les cas non pathologiques. Ces résultats prometteurs s'expliquent par l'adaptation totale de méthodes génériques au cas particulier d'une paroi convexe et fermée. Même si l'algorithme est itératif, les temps de calculs sont relativement faibles et, après différents tests, la méthode s'avère très résistante au bruit.

Les études actuelles sont conduites dans le but d'étendre la méthode à la troisième dimension. L'ensemble du système devrait prochainement être intégré dans une architecture de type multi-agents. Des tests cliniques plus complets sont en cours.

Références

- [1] B. Solaiman, F. Pipelier, C. Roux and M. Robaszekiewicz, "Esophageal wall detection using endosonographic imaging systems", IEEE Engineering in Medicine and Biology Society Conference, EMBS 95 pp. 611-612, 1995.
- [2] P. Wilinski, B. Solaiman, C. Roux and M. Robaszekiewicz" Texture analysis of ultrasonographic endoscopy images using the Master Classifier Method", VII European Signal Processing Conference, EUSIPCO-94, 13-16 Sep, Edinburgh, Scotland, 1994.
- [3] F. Pipelier, P. Danduran, B. Solaiman, C. Roux and M. Robaszekiewicz, "Esophagus inner wall detection on echo-endoscopic images sequence", IEEE Workshop on Nonlinear Signal and Image Processing pp.799-802, 1994.
- [4] B. Solaiman, R. Koffi, M.C. Mouchot and A. Hillion, "An Information Fusion Method for Multispectral Image Classification Post-processing", IEEE Transaction. on Geoscience and Remote Sensing, Vol.36, No.2, pp. 395-406, March 1998.
- [5] P. Wilinski, B. Solaiman and A. Hillion, "Towards the borders between neural and Markovian paradigms", IEEE Transaction on System, Man, and Cybernetics, Vol.28, No.2, pp. 146-159, April 1998.
- [6] L.A. Zadeh. "Fuzzy sets. Inform", Inform. Contr., vol. 8, pp. 338-353, 1965.
- [7] L.A. Zadeh, "Fuzzy algorithm", Inform. Contr., vol. 12, pp. 94-102, 1968.
- [8] F. Pipelier, "Analyse d'images écho-endoscopiques de l'œsophage" PHD dissertation thesis, ENST-Bretagne – Université de Rennes I, 1997.
- [9] I. Bloch, " Information combination operators for data fusion: a comparative review with classification," IEEE Trans. Sys. Man Cyber. A 26 (1), pp. 52-67, 1996.
- [10] M. Kass, A. Witkin, and D. Terzopoulos. "Snakes: Active contour models", International Journal of Computer Vision, pp. 321-331, 1998.
- [11] F. Pipelier, B. Solaiman, S. Grassin, C. Roux. "A new dynamic model: application on ultrasound images", IEEE Engineering in Medicine and Biology Society Conference, EMBS 96, pp. 167-168, 1996.