

Fusion de classifieurs en utilisant la théorie de l'évidence pour l'amélioration de la classification d'image

Salim CHITROUB, Amrane HOUACINE, et Boualem SANSAL

Faculté de Génie Electrique, USTHB

B. P. 32, El-Alia, Bab-Ezzouar, 16111, Alger, ALGERIE

s_chitroub@hotmail.com

Résumé - Le problème traité dans cet article concerne l'amélioration de la classification d'image dans les conditions d'insuffisance d'informations *a priori* déterministes et fiables sur l'état et la nature de la formation de l'image à l'instant de la prise de vue. Une méthode de fusion des classifieurs d'imagerie pour une meilleure classification des scènes imagées est proposée. La méthode est basée sur la théorie de l'évidence. Elle est générale et applicable à tout type de classifieur. On utilise le taux de fiabilité de la classification comme critère d'évaluation, les résultats obtenus, en utilisant des images simulées et réelles de télédétection, montrent que la méthode proposée donne de meilleurs résultats en comparaison avec les résultats des classifieurs considérés séparément.

Abstract - The problem addressed in this paper concerns the improvement of the image classification in conditions of insufficiency of deterministic and reliable prior information about the state and the nature of the image formation at the moment of acquisition. A fusion method of classifiers for a better classification of the observed scenes is proposed. The method is based on the theory of evidence. It is general and applicable to all types of classifiers. Using the rate of reliability of classification as criteria of evaluation, the obtained results, using the simulated and real remotely sensed data, show that the proposed method gives the best results compared with those of classifiers separately considered.

1. Introduction

La classification d'image, obtenue dans des conditions de prise de vue difficiles à contrôler, est un problème de reconnaissance mal posé de la scène imagée [1]. En effet, la non disponibilité d'informations *a priori* déterministes et fiables sur l'état et les conditions de fonctionnement du system d'acquisition et sur la nature de la scène imagée au moment de la prise de vue rend la classification difficile à évaluer d'une manière précise. Plusieurs classifieurs peuvent donner des réponses différentes sur l'attribution d'un pixel de l'image à la classe qui lui correspond le plus. Ceci est dû d'une part à la qualité du contenu du pixel et d'autre part à l'erreur spécifique au classifieur. Cette erreur dépend du modèle de décision du classifieur et de la base de données utilisée pour son apprentissage. Le comportement de chaque classifieur vis-à-vis de l'image à classifier est déterminé en fournissant des informations différentes sur l'état de la scène. Exploiter les différentes informations que les classifieurs fournissent en utilisant les méthodes de fusion de classifieurs peut conduire à obtenir une image classifiée plus précise. Même si le classifieur est moins performant, la connaissance de son comportement apporte une certaine information utilisable à propos de la vraie classe du pixel traité [2, 3]. Ainsi, la fusion peut éventuellement permettre d'en cumuler les avantages sans en cumuler les inconvénients, en réduisant l'erreur et en augmentant la fiabilité de la classification.

2. Méthode proposée

En considérant la fusion des classifieurs comme une deuxième classification qui succède à celles des classifieurs,

une méthode de fusion de classifieurs d'imagerie est proposée. Le principe est de fusionner les réponses des classifieurs sans tenir compte de leurs caractéristiques internes. De ce fait, la méthode proposée se veut générale et applicable à tout type de classifieur. Le schéma de fusion est parallèle, il est constitué de plusieurs classifieurs et d'un fusionneur. Les classifieurs utilisent les mêmes données mais n'interagissent pas entre eux. Le fusionneur n'a pas accès aux données d'entrée. Il traite les réponses des classifieurs et donne une réponse finale grâce à un processus de fusion basé sur la théorie de l'évidence [4].

Supposons qu'il y ait N classifieurs Cl_k . Chaque classifieur k reçoit la même forme x (pixel de l'image à classifier). Soit Ω l'ensemble fini de toutes les hypothèses (évidences ou classes w) possibles et 2^Ω l'ensemble des sous-ensembles de Ω . La théorie de l'évidence est complètement caractérisée par une fonction appelée fonction de masse : $m : 2^\Omega \rightarrow [0, 1]$ telle que : $m(\emptyset) = 0$ et $\sum_{A \subseteq \Omega} m(A) = 1$. Sa valeur exprime le degré de l'évidence associée à l'ensemble A , appelé aussi élément focal, celui ci ne peut être associé à aucun sous-ensemble particulier de A . A chaque fonction de masse est associée une mesure de croyance : $Bel(A) = \sum_{B \subseteq A} m(B)$. Le fusionneur considère la réponse d'un classifieur comme une évidence sur la vraie classe du pixel x . Si la réponse est un rejet, on dira que cette évidence soutient complètement l'ensemble Ω , mais ne soutient aucun de ses sous-ensembles, c.-à-d., $m_k(\Omega) = 1$ et $m_k(A) = 0, \forall A \subset \Omega$. Si la réponse n'est pas un rejet, c.-à-d., w_{jk} , le classifieur soutient la proposition $x \in w_{ij}$. Soient τ_{rec}^k , τ_{err}^k , et τ_{rej}^k les taux de reconnaissance, d'erreur, et de rejet du classifieur k , respectivement. Supposons que la réponse du

classifieur k est correcte avec une confiance τ_{rec}^k , qu'elle est fautive avec une confiance τ_{err}^k , et comme la réponse de ce classifieur n'indique rien sur les autres classes, tout le reste de la confiance est donné à l'ensemble Ω . Soit alors :

$$mk(\{w_{jk}\}) = \tau_{rec}^k, \quad (1)$$

$$mk(\overline{\{w_{jk}\}}) = \tau_{err}^k, \text{ avec } \overline{\{w_{jk}\}} = \Omega - \{w_{jk}\} \quad (2)$$

$$mk(\Omega) = \tau_{rej}^k \text{ et } mk(A) = 0, \forall A \subset \Omega. \quad (3)$$

On estime les valeurs de ces taux à partir de la matrice de confusion C^k du classifieur k . Celle-ci étant calculée à partir des résultats de classification obtenus par le classifieur sur une base de test. Un élément de cette matrice c_{ij}^k représente le nombre de pixels dans la base qui appartiennent à la classe w_i et pour lesquels le classifieur k a répondu w_j :

$$c_{ij}^k = \left| \left\{ x \in w_i, Cl_k(x) = w_j \right\} \right|, \quad i = 0, 1, \dots, M \quad (4)$$

M représente le nombre de classes. Lorsque l'indice j prend la valeur 0 cela signifie que le classifieur a la possibilité de donner au pixel l'étiquette de la classe de rejet w_0 s'il considère que les informations mises à sa disposition ne lui permettent pas de donner un résultat fiable. Dans le cas où le résultat du classifieur n'aurait pas une classe de rejet, les éléments de la colonne $j = 0$ de la matrice C^k seront nuls. L'estimation des différents taux est alors donnée par :

$$\tau_{rec}^k = \frac{\sum_{i=1}^M c_{ii}^k}{\sum_{i=1}^M \sum_{j=0}^M c_{ij}^k}; \quad \tau_{err}^k = \frac{\sum_{i=1}^M \sum_{j=0, j \neq i}^M c_{ij}^k}{\sum_{i=1}^M \sum_{j=0}^M c_{ij}^k}; \quad (5)$$

$$\tau_{rej}^k = \frac{\sum_{i=1}^M c_{i0}^k}{\sum_{i=1}^M \sum_{j=0}^M c_{ij}^k}$$

Pour obtenir une attribution de masse qui contient toutes les évidences, il faut fusionner les N attributions de masse m_k associées aux classifieurs Cl_k . Pour cela, on utilise la règle de combinaison de Dempster [4] qui exige que les évidences soient indépendantes. Pour cela il faut que les classifieurs soient statistiquement indépendants conditionnellement aux classes [5]. En fait, le nombre de classes dans notre cas est un paramètre connu pour tous les classifieurs. Ce qui différencie les classifieurs ce sont les classes affectées aux pixels de l'image. Par conséquent, il est logique de considérer que les classifieurs sont indépendants par rapport aux classes.

Le processus de fusion des évidences en utilisant la règle de Dempster s'effectue en deux étapes. Chaque étape fusionne un nombre restreint d'évidences. La fusion dans une deuxième étape est effectuée sur les évidences résultantes.

2.1 Première étape de fusion

Dans notre application, l'espace des attributions de masse m_k est restreint aux trois éléments focaux seulement, à savoir : $\{w_j\}$, $\overline{\{w_j\}}$ et Ω . Toute autre hypothèse composée des hypothèses $\{w_j\}$ n'a aucune signification, puisqu'on ne peut pas considérer une classe comme un ensemble de classes de caractéristiques différentes. La première étape de fusion consiste donc à répartir les N classifieurs en N_G groupes selon

leurs réponses. Les classifieurs de chaque groupe ont tous la réponse w_{gi} , donc leurs fonctions d'attribution de base ont toutes les mêmes focaux, mais avec des valeurs différentes de m . Soit G_i un groupe de p classifieurs. Leurs attributions de masse $m_{k1}, m_{k2}, \dots, m_{kp}$ ont toutes les trois focaux $\{w_j\}$, $\overline{\{w_j\}}$ et Ω . L'attribution de la masse totale du groupe G_i est telle que : $m_{G_i} = m_{k1} \oplus m_{k2} \oplus \dots \oplus m_{kp}$, où \oplus est la règle de combinaison de Dempster. Le calcul de m_{G_i} est récursif en utilisant la règle de Dempster à deux évidences :

$$\begin{aligned} m_l &= m_{k_l}, \\ m_r &= m_{r-1} \oplus m_{k_r}, \quad r = 2, 3, \dots, p, \\ m_{G_i} &= m_p \end{aligned} \quad (6)$$

Notons que pour combiner deux évidences la règle de Dempster est telle que :

$$\begin{aligned} m(A) &= (m_1 \oplus m_2)(A) = K \sum_{w_1 \cap w_2 = A} m_1(w_1) m_2(w_2) \\ K &= \left(\sum_{w_1 \cap w_2 \neq \Phi} m_1(w_1) m_2(w_2) \right)^{-1} \end{aligned} \quad (7)$$

Une évidence de la forme $m_i(\Omega) = 1$ n'a aucun effet sur le résultat de la combinaison. En outre, nous considérons que : $0 < \tau_{rec}^k < 1$ et $0 \leq \tau_{err}^k < 1$, quel que soit le classifieur. Cette condition nous assure que les évidences fournies par les classifieurs ne seront pas contradictoires, c.-à-d., que le dénominateur de (7) n'est pas nul. Une preuve simple est donnée dans [6]. Les cas de $\tau_{rec}^k = 1$, $\tau_{err}^k = 1$, et $\tau_{rej}^k = 0$ sont des cas dégénérés qui peuvent être traités à part.

L'espace des attributions de masse m_{G_i} est restreint aux trois éléments focaux seulement à savoir : $\{w_j\}$, $\overline{\{w_j\}}$ et Ω . Le groupe G_i est considéré comme un classifieur qui a répondu w_{gi} avec $\tau_{rec} = m_{G_i}(\{w_{gi}\})$, $\tau_{err} = m_{G_i}(\overline{\{w_{gi}\}})$ et $\tau_{rejet} = m_{G_i}(\Omega)$. L'étape 1 change les N_{Cl} classifieurs en N_G classifieurs qui ont tous des réponses différentes, telles que : $w_{g1} \neq w_{g2} \neq \dots \neq w_{gN_G}$.

2.2 Deuxième étape de fusion

La deuxième étape du processus de fusion consiste à fusionner les évidences des N_G groupes telles que : $m = m_{G_1} \oplus m_{G_2} \oplus \dots \oplus m_{G_{N_G}}$. Comme les N_G groupes donnent des classes différentes les unes des autres, le calcul de m devient très lourd avec la règle de Dempster. Afin de réduire le temps de calcul des fonctions de croyance, les simplifications présentées dans [6] et [7] ont été utilisées. Sachant que $N_G \leq M$ et $N_G \leq N$, nous obtenons des relations simplifiées pour calculer les fonctions de croyance en chacune des classes :

$$\begin{aligned} a &= \sum_{i=1}^{N_G} \frac{m_{G_i}(\{w_{gi}\})}{1 - m_{G_i}(\{w_{gi}\})}, \quad b = \prod_{i=1}^{N_G} (1 - m_{G_i}(\{w_{gi}\})), \\ c &= \prod_{i=1}^{N_G} m_{G_i}(\overline{\{w_{gi}\}}), \quad F = \begin{cases} (1+a)b - c & \text{si } N_G = M \\ (1+a)b & \text{si } N_G < M \end{cases} \end{aligned}$$

$$Bel(\{w_{gi}\}) = \begin{cases} F\left(\frac{m_{G_i}(\{w_{gi}\})}{1 - m_{G_i}(\{w_{gi}\})} \cdot b + \frac{m_{G_i}(\Omega)}{m_{G_i}(\{w_{gi}\})} \cdot c\right) & \text{si } N_G = M \\ & \text{ou } N_G = M - 1, i = M \\ F\left(\frac{m_{G_i}(\{w_{gi}\})}{1 - m_{G_i}(\{w_{gi}\})} \cdot b\right) & \text{sinon} \end{cases} \quad (8)$$

$$Bel(\overline{\{w_{gi}\}}) = \begin{cases} F\left(\left(a - \frac{m_{G_i}(\{w_{gi}\}) - m_{G_i}(\overline{\{w_{gi}\}})}{1 - m_{G_i}(\{w_{gi}\})}\right) b - c\right) & \text{si } N_G = M \\ F\left(a - \frac{m_{G_i}(\{w_{gi}\}) - m_{G_i}(\overline{\{w_{gi}\}})}{1 - m_{G_i}(\{w_{gi}\})}\right) b & \text{si } N_G < M, i \leq N_G \\ F\left(a - \frac{m_{G_i}(\overline{\{w_{gi}\}})}{1 - m_{G_i}(\overline{\{w_{gi}\}})}\right) b & \text{sinon} \end{cases} \quad (9)$$

2.3 Règle de décision

Les fonctions de croyance des réponses des N_G classifieurs étant calculées, une règle de décision adéquate est utilisée pour reclassifier la scène imagée. La règle tient compte des informations portées par $Bel(\{w_{gi}\})$ et $Bel(\overline{\{w_{gi}\}})$. Soit : $Bel_{nette}(w_{gi}) = Bel(\{w_{gi}\}) - Bel(\overline{\{w_{gi}\}})$ la croyance nette en w_{gi} . La règle utilisée pour la classification finale est alors :

$$y = \begin{cases} w_j & \text{si } Bel_{nette}(\{w_j\}) = \max_{w \in \Omega} Bel_{nette}(\{w\}) \geq \alpha, 0 < \alpha < 1 \\ w_0 & \text{sinon} \end{cases} \quad (10)$$

3. Résultats expérimentaux

Notre première expérience, pour tester la méthode proposée, a été effectuée sur une image simulée (Fig. 1). Cette image est obtenue en ajoutant du bruit à l'image de réalité de terrain simulée (Fig. 2, 251x251 pixels) par un bruit additif gaussien de variance égale à 551. Elle a été classifiée par trois classifieurs différents, à savoir : le classifieur de maximum de vraisemblance (MV), classifieur de minimum de distance de Mahalanobis (MDM), et le classifieur de minimum de distance Euclidienne (MDE). Le résultat de MDE (Fig.3) est le plus précis et l'image correspondante est la moins bruitée. La même remarque peut être formulée pour le MDM (Fig. 4) par rapport au MV (Fig. 5). L'image de ce dernier est très dégradée et loin de la réalité terrain. La figure 6 représente l'image résultat de la fusion de trois classifieurs. Nous constatons une amélioration portée par l'image de fusion en comparaison avec les images des classifieurs. Il y a moins de pixels mal classifiés. Malgré que l'image la plus dégradée soit incluse dans le processus de fusion, l'image résultat de fusion est moins bruitée que le meilleur résultat des trois classifieurs. La figure 7 nous donne les taux de fiabilité des trois

classifieurs et celui du résultat de leur fusion. Ce taux mesure la fiabilité de la classification donnée par : $fiabilité = \tau_{rec} / (1 - \tau_{rejet})$. Le taux de la fusion est supérieur aux taux des trois classifieurs. La deuxième expérience a été effectuée sur une image multispectrale de télédétection fournie par le satellite SPOT. L'image représente une partie de la ville de Blida (Nord d'Algérie, 140x244 pixels). La figure 8 représente la troisième bande spectrale du satellite. Les trois bandes spectrales ont été utilisées comme entrée pour les trois classifieurs, MV, MDM, et MDE. La même conclusion que dans le cas de l'image simulée peut être formulée sur les précisions des images classifiées, figures 9, 10, et 11, respectivement. La fusion des trois classifieurs (Fig. 12) a amélioré le résultat final de la classification. Nous observons moins de points isolés dans les zones homogènes de la scène. Les frontières entre les classes sont aussi bien délimitées. Puisque dans la littérature il n'y a pas assez de critères d'évaluation de classification en absence de réalité terrain, nous avons utilisé la définition de Haralick d'une bonne segmentation [8]. Ceci consiste à calculer la moyenne des taux de fiabilité calculés dans plusieurs zones homogènes. Le taux moyen de la fusion est supérieur aux taux moyens des classifieurs (Fig. 13). Ainsi, la méthode proposée ici a permis d'obtenir une classification plus précise et prouver de ce fait son efficacité.

4. Conclusion

Dans cet article, une méthode de fusion de classifieurs pour l'amélioration de la classification des scènes imagées a été proposée. En considérant la fusion comme une deuxième classification qui succède à celle des classifieurs, la méthode cumule les avantages de chacun sans en cumuler les inconvénients. Elle est basée sur le principe de la fusion parallèle des systèmes distribués en utilisant la théorie de l'évidence. Les sorties des classifieurs sont traitées en ignorant complètement les caractéristiques internes de ces derniers. De ce fait, la méthode est générale et applicable à tout type de classifieur. Le développement théorique ainsi que la mise en œuvre de la méthode ont été donnés. La méthode a été appliquée pour classifier des images simulées et réelles de télédétection. On utilise le taux de fiabilité de la classification comme critères d'évaluation, les résultats obtenus montrent que la méthode proposée donne de meilleurs résultats en comparaison avec les résultats des classifieurs considérés séparément.

Références

- [1] Tikhonov A. N. and Arsenin V. Y., *Solutions of Ill-Posed Problems*, Winston & Sons, Washington D. C., 1977.
- [2] Bonnin P., Hoeltzner-Douarin, and Pissaloux E., A new way of image data fusion: the multi-spectral co-operative segmentation, *ICIP 95*, Vol. 3, pp. 572-575, Washington, USA, 1995.
- [3] Rosenberger C. and Chehdi K., Genetic Fusion: Application to multi-components image segmentation, *ICASSP 2000*, Vol. 3, pp. 2223-2226, Istanbul, Turkey, 2000.
- [4] Shafer G., *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

[5] Pearl J., Geiger D., and Verma T., Conditional independence and its representations", *Kybernetika*, Vol. 25, N° 2, 1989.

[6] Xu L., Krzyzak A., and Suen C., Methods of combining multiple classifiers and their applications to handwriting recognition, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 22, N° 3, pp. 418-435, 1992.

[7] Barnet J. A., Computation methods for a mathematical theory of evidence, In *17th Inter. Joint Conf. on Artificial Intelligence*, pp. 868-875, Vancouver, Canada, 1985.

[8] Harralick R. M., Statistical and structural approaches to texture, In *Proc. of IEEE*, Vol. 69, pp. 786-804, May, 1979.

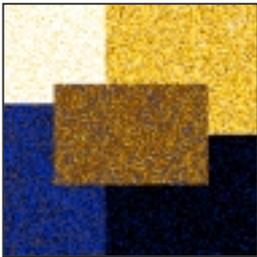


Fig. 1 : image simulée bruitée

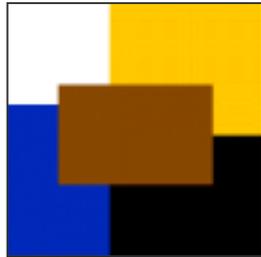


Fig. 2 : image réalité terrain

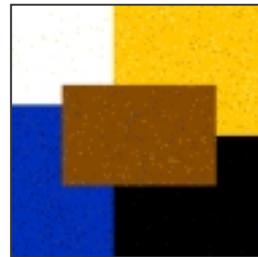


Fig. 3 : classification MDE

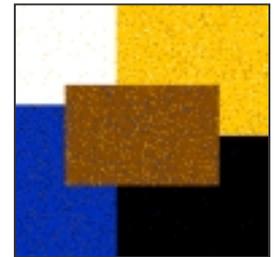


Fig. 4 : classification MDM

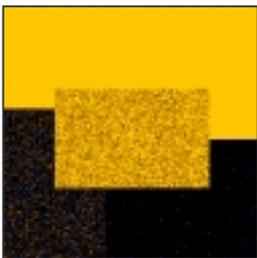


Fig. 5 : classification MV

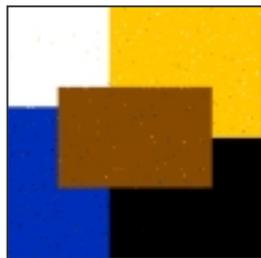


Fig. 6 : résultat de la fusion

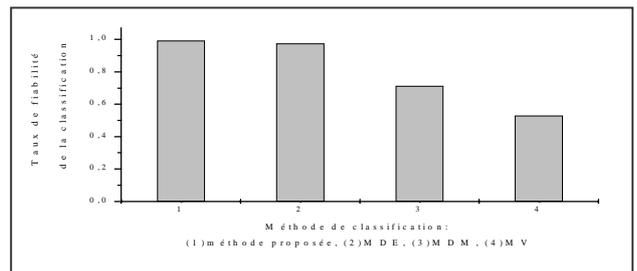


Fig. 7 : taux de fiabilité de la classification

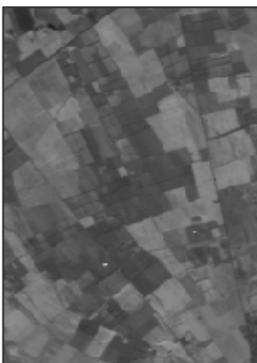


Fig. 8 : 3^{ème} bande spectral

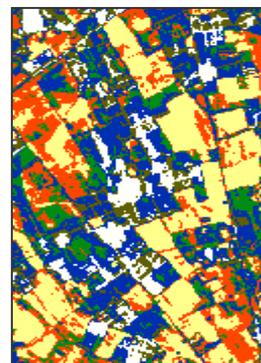


Fig. 9 : classification MDE

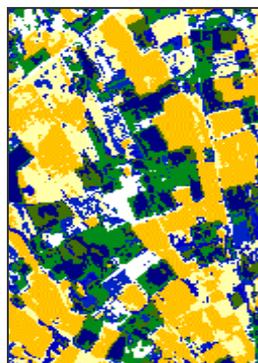


Fig. 10 : classification MDM

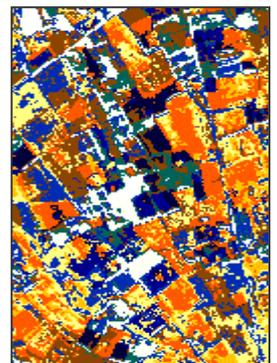


Fig. 11 : classification MV

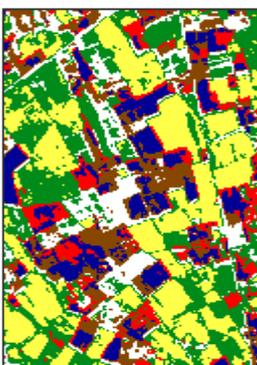


Fig. 12 : résultat de la fusion

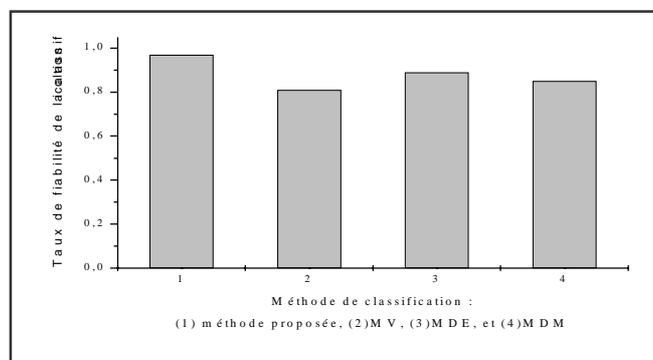


Fig. 13 : taux de fiabilité de la classification