



ARCHITECTURES ET MÉTHODES DE FUSION POUR LA CLASSIFICATION MULTISOURCE[†]

P. De Maertelaere*, P. Ravazzola**, Ph. Ghesquière***, A. Beltrando****

* AERO, 3 avenue de l'Opéra, 75001 Paris, France

** AERO, allée Jean Giono, chemin de Léry, 83500 La Seyne, France

*** MATRA MS2i, Guyancourt, BP 235, 78052 Saint-Quentin-en-Yvelines, France

**** CTSN / LSA, BP 28, 83800 Toulon Naval, France

RÉSUMÉ

Cet article a pour objet la conception et l'évaluation d'architectures utilisant des méthodes de fusion multisource (Bayes, Dempster-Shafer) appliquées à un problème de classification dans un contexte de lutte air-surface.

On s'intéresse plus particulièrement ici à certains problèmes liés à la mise en œuvre du processus de classification qui sont assez peu abordés dans la littérature :

- procédure de transcodage dans le formalisme de la méthode de fusion,
- suivi temporel du processus de classification-fusion,
- utilisation de données contextuelles.

Les moyens de résoudre ces problèmes pour les différentes architectures sont considérés et on conclut par une analyse comparative des solutions proposées.

1- Introduction

Cet article a pour objet la conception et l'évaluation d'architectures utilisant des méthodes de fusion multisource appliquées à un problème de classification dans un contexte de lutte air-surface.

Les capteurs qui peuvent être utilisés pour la classification des cibles sont les suivants : veille panoramique IR, veille radar 2D et 3D, écoute radar, IFF.

Les données d'entrée du système de classification-fusion sont les pistes établies par un système de pistage multicapteurs.

A chacune des pistes est associée une liste d'attributs parmi lesquels on peut distinguer :

- . les attributs cinématiques (position, vitesse, accélération),
- . les attributs "signal" (IR, radar, ESM, IFF).

Ces attributs peuvent être extraits à partir des mesures brutes des capteurs puis regroupés, après association des pistes.

Le processus de classification-fusion est réalisé en deux étapes :

Première étape : pour chaque source d'information (nœud de fusion), on met en œuvre une méthode de classification directe (classification dans l'espace des attributs) pour évaluer le degré de certitude des différentes hypothèses de reconnaissance.

ABSTRACT

This paper describes the design and evaluation of architectures using multisource fusion methods (Bayes, Dempster-Shafer) applied to a classification problem in an air-surface warfare context.

Focus is brought on the way of operating classification scheme, and specially on problems which are seldom encountered in the literature :

- transformation step within the fusion method formalism,
- temporal evolution of the classification-fusion process,
- integration of contextual knowledge.

The way of solving these problems for various architectures are considered and different solutions are compared.

L'évaluation relative des hypothèses est alors transcrite dans le formalisme de la méthode de fusion utilisée, par une procédure appelée ici transcodage.

Deuxième étape : on met en œuvre une méthode de fusion pour combiner les résultats d'évaluation monosource : un critère de décision est alors appliqué sur le résultat de la fusion pour choisir l'hypothèse la plus probable.

On ne s'intéresse pas ici à la mise en œuvre des méthodes d'extraction d'attributs et de classification directe ; on s'intéresse plutôt à la mise en œuvre de méthodes de fusion numérique et en particulier à la résolution des problèmes suivants :

- mise en œuvre de la procédure de transcodage,
- suivi temporel du processus de classification-fusion,
- utilisation de données contextuelles.

Ces problèmes sont assez peu abordés dans la littérature : ils sont la plupart du temps "éludés".

Les méthodes de fusion symbolique (intelligence artificielle et logique floue), qui permettent de réaliser des inférences de plus haut niveau, ne sont pas étudiées dans le cadre de cet article, bien qu'elles puissent être utilisées de façon complémentaire avec les méthodes de fusion numérique proposées.

[†] Travail réalisé sous contrat DCN Toulon-CESDA n° A 92 50 164



Les deux méthodes étudiées sont donc des méthodes classiques de fusion numérique :

- . méthodes bayésiennes (cf [1], [3] et [5]),
- . méthodes utilisant la théorie de Dempster-Shafer (cf [1], [2], [3] et [5]).

2. Définition des architectures

Du point de vue de la décision finale de classification, on peut distinguer les trois types d'architecture suivants :

- 1) les architectures complètement centralisées : tous les attributs sont transmis au centre de fusion qui réalise le processus de classification-décision,
- 2) les architectures partiellement centralisées : un processeur local traite une partie des attributs et réalise un processus de préclassification (évaluation relative des différentes hypothèses), puis transmet ses résultats au centre de fusion qui prend une décision globale après fusion des résultats de préclassification (structure de décision centralisée),
- 3) les architectures décentralisées : un processeur local traite une partie des attributs et réalise un processus de classification-décision puis transmet son résultat au centre de fusion qui prend une décision globale après fusion des décisions (structure de décision décentralisée).

On s'intéresse ici à la mise en œuvre des méthodes de fusion pour les architectures de type 2) et 3) qui seront appelées par la suite architectures centralisées et architectures décentralisées.

3. Mise en œuvre des méthodes de fusion

Les hypothèses du cadre de discernement E (ensemble des classes de cibles recherchées) sont notées : $H_1, \dots, H_i, \dots, H_N$.

Les attributs regroupés au niveau de chaque nœud de fusion (ou critère de classification local) sont notés : $A_1, \dots, A_j, \dots, A_M$.

Les décisions locales éventuelles sont notées $d_{k_1}^j, \dots, d_{k_j}^j, \dots, d_{k_M}^j$ (la notation $d_{k_j}^j$ correspond au choix de l'hypothèse H_{k_j}). La matrice de confusion du processus de décision locale, pour le critère j, est notée C^j :

$$C^j = \begin{bmatrix} p(d_1^j / H_1), & \dots, & p(d_N^j / H_1), & p(d_{N+1}^j / H_1) \\ & \dots & & \\ & & \dots & \\ p(d_1^j / H_N), & \dots, & p(d_N^j / H_N), & p(d_{N+1}^j / H_N) \end{bmatrix}$$

La décision notée ici d_{N+1}^j correspond à une situation éventuelle d'indétermination par rapport au cadre de discernement (non décision) : l'objet à reconnaître est alors de classe "inconnue".

3.1 Formalisme bayésien

La règle de fusion utilisée est la règle de Bayes et le critère de décision utilisé est celui du maximum a posteriori ; on fait généralement l'hypothèse de l'indépendance des différentes sources.

Architectures centralisées

Les probabilités a posteriori $p(H_i/A_j)$ sont évaluées localement par un processus de préclassification/transcodage. La règle de fusion est la suivante :

$$p(H_i / A_1, \dots, A_M) = p(H_i) \prod_{j=1}^M \left(\frac{p(H_i / A_j)}{p(H_i)} \right)$$

Architectures décentralisées

Les décisions $d_{k_j}^j$ sont établies localement par un processus de classification-décision. La règle de fusion est la suivante :

$$p(H_i / d_{k_1}^j, \dots, d_{k_M}^j) = p(H_i) \frac{\prod_{j=1}^M p(d_{k_j}^j / H_i)}{\prod_{j=1}^M p(d_{k_j}^j)}$$

Les probabilités $p(d_{k_j}^j / H_i)$ sont obtenues par utilisation de la matrice de confusion C^j , supposée connue a priori.

3.2 Formalisme de Dempster-Shafer

La règle de fusion utilisée est la règle de Dempster et le critère de décision utilisé est le maximum de plausibilité.

Architectures centralisées

Il existe deux approches pour réaliser le transcodage :

- 1) l'approche globale : on définit un jeu de masse $m^j()$ qui traduit une évaluation relative du degré de certitude des différentes propositions discernées :

$$m^j(P_1^j), \dots, m^j(P_i^j), \dots, m^j(P_N^j), m^j(E)$$

- 2) l'approche séparable (utilisée notamment dans [1] et [3]) : on considère chacune des hypothèses H_i séparément et on traduit les résultats de préclassification, sous la forme d'un jeu de masse $m_i^j()$ ayant les éléments focaux suivants :

$$m_i^j(H_i), m_i^j(\bar{H}_i), m_i^j(E)$$

Dans le premier cas, on applique directement la règle de Dempster pour réaliser la fusion inter-critères : $m^g = m^1 \oplus m^2 \oplus \dots \oplus m^M$.

Dans le deuxième cas, on réalise d'abord une fusion inter-critères ($m_i^g = m_i^1 \oplus m_i^2 \oplus \dots \oplus m_i^M$), pour chacune des hypothèses H_i , puis on réalise la fusion inter-hypothèses : $m^g = m_1^g \oplus m_2^g \oplus \dots \oplus m_N^g$.

Architectures décentralisées

S'il est possible de modéliser l'incertitude associée au processus de décision (choix de l'hypothèse H_{k_j}), on définit un jeu de masse $m^j()$ ayant les éléments focaux suivants :

$$m^j(H_{k_j}), m^j(\bar{H}_{k_j}), m^j(E)$$



Un exemple de ce type de transcodage, utilisant la matrice de confusion C^j , est présenté dans [4].

On applique alors la règle de Dempster pour réaliser la fusion inter-critères : $m^s = m^1 \oplus m^2 \oplus \dots \oplus m^M$.

Dans tous les cas, on évalue, à partir de la distribution de masses résultante, l'intervalle $[Spt(H_i), Pls(H_i)]$ pour chacune des hypothèses H_i et on choisit l'hypothèse H_k qui maximise la plausibilité $Pl(H_k)$.

4. Mise en œuvre de la procédure de transcodage

4.1 Position du problème

Le transcodage est la transcription de résultats de préclassification ou de décision sous la forme de grandeurs compatibles avec l'application du formalisme de fusion :

- probabilités a posteriori pour le formalisme bayésien,
- distribution de masses pour le formalisme de Dempster-Shafer,

Cette transcription doit tenir compte, autant que possible, du caractère imprécis et incertain de la connaissance qui a permis au traitement de réaliser son évaluation ou sa décision.

De façon générale, l'incertitude associée à un processus de classification peut être due aux causes suivantes :

- mauvaise séparabilité "intrinsèque" des classes (faible pouvoir discriminant des attributs),
- mauvaise définition du cadre de discernement,
- erreurs de modélisation des classes (inadéquation entre le modèle de référence et la réalité observée),
- imprécision de la mesure.

Il n'existe pas de méthode générale pour évaluer de façon optimale cette incertitude : on peut utiliser une approche théorique (mesure de perte d'information), empirique (apprentissage des performances), heuristique ou purement subjective (définition a priori d'un coefficient de confiance).

4.2 Formalisme bayésien

On s'intéresse ici aux architectures centralisées puisque le problème de transcodage ne se pose pas dans le cas des architectures décentralisées si on connaît la matrice de confusion (l'incertitude du processus de décision est alors modélisée par la donnée des probabilités de confusion).

Le transcodage dans le formalisme bayésien, pour les architectures centralisées, consiste à évaluer les probabilités a posteriori de chacune des hypothèses du cadre de discernement et à tenir compte de l'incertitude associée à cette évaluation dans l'application de la règle de fusion.

On suppose que l'on peut définir un coefficient $q^j \in [0,1]$ qui permet de quantifier la qualité du processus de préclassification, compte-tenu de la mauvaise connaissance des distributions de probabilité conditionnelle, de l'imprécision de la mesure, d'un mauvais rapport signal à bruit, etc...

La règle de fusion bayésienne inter-critères est alors modifiée de la façon suivante :

$$p(H_i / A_1, \dots, A_M) = p(H_i) \prod_{j=1}^M \left(\frac{p(H_i / A_j)}{p(H_i)} \right)^{q^j}$$

Si on interprète cette formule (utilisée notamment dans [5]), on peut dire que l'influence de l'évaluation locale $p(H_i/A_j)$ est "atténuée" par le facteur q^j .

4.3 Formalisme de Dempster-Shafer

Architectures centralisées

On présente ici une façon de réaliser le transcodage de probabilités $p(H_i/A_j)$ obtenues par classification bayésienne :

- on suppose que la distribution des probabilités a posteriori définit un jeu de masse bayésien $m^j()$ traduisant une évaluation "précise et certaine" :

$$\begin{cases} m^j(H_i) = p(H_i / A_j) \\ m^j(E) = 0 \end{cases}$$

- on suppose que la qualité du processus de classification peut être quantifiée par la donnée d'un coefficient $q^j \in [0,1]$,

- la modification du jeu de masse bayésien $m^j()$ est alors obtenue par une procédure d'affaiblissement (cf [1]) qui permet de définir un jeu de masse incertain $m_a^j()$ de la façon suivante :

$$\begin{cases} m_a^j(H_i) = q^j m^j(H_i) \\ m_a^j(E) = m^j(E) + \sum_i (1 - q^j) m^j(H_i) \end{cases}$$

5. Suivi temporel

5.1 Position du problème

L'obtention de données nouvelles sur une cible en cours de classification doit normalement permettre de réduire progressivement l'incertitude : il s'agit donc de trouver un mécanisme qui permette de "combiner" les données nouvelles avec la situation existante pour améliorer le pouvoir de classification au cours du temps.

5.2 Approches possibles

Les approches possibles pour prendre en compte de nouvelles mesures issues des capteurs sont les suivantes :

- Approche 1 : on mémorise les observations passées (sur un horizon de temps à définir), ce qui permet de constituer, un vecteur d'attributs séquentiel et on réalise une classification directe dans cet espace d'attributs multidimensionnel (on peut alors utiliser le test séquentiel de Wald si une décision locale doit être prise),

- Approche 2 : on utilise uniquement les attributs courants pour réaliser la classification directe (et son transcodage associé) et on réalise une fusion temporelle locale des résultats passés (évaluations des hypothèses pour les architectures centralisées ou décisions locales pour les architectures décentralisées) avant de mettre en œuvre la fusion multicritère.



On décrit ici la façon de mettre en œuvre l'approche 2.

5.3 Formalisme bayésien

On note $X_t^j = (X_{t-1}^j, A_j)$ le vecteur d'attributs séquentiel enrichi à l'instant t par la mesure de l'attribut courant A_j .

Si on fait l'hypothèse d'indépendance temporelle, la fusion temporelle par application de la règle de Bayes s'écrit :

$$p(H_i / X_t^j) = \frac{p(H_i / A_j)}{p(H_i)} p(H_i / X_{t-1}^j)$$

5.4 Formalisme de Dempster-Shafer

On note $m_t^j()$ le jeu de masses évalué à l'instant t et $M_t^j()$ le jeu de masses obtenue par fusion temporelle.

La fusion temporelle des résultats peut être réalisée (cf [6]) par application de la règle de Dempster : $M_t^j = M_{t-1}^j \oplus m_t^j$

6. Utilisation du contexte

6.1 Position du problème

L'obtention de données contextuelles doit normalement permettre d'améliorer la précision du processus de classification : il s'agit donc de trouver un mécanisme qui permette de "combiner" ces données avec l'évaluation de la situation multicapteurs.

On peut distinguer deux catégories de données contextuelles:

- les informations de type brouillage, météo : tout ce qui affecte l'environnement de mesure des capteurs,
- les informations de type renseignement : forces présentes...

6.2 Approches possibles

L'introduction du contexte peut se faire à différents niveaux des architectures de fusion :

- 1) au niveau du transcodage,
- 2) au niveau de la décision locale pour les architectures décentralisées,
- 3) au niveau de la décision globale.

On décrit ici la façon d'utiliser le contexte au niveau 1.

6.3 Formalisme bayésien

La connaissance externe de l'environnement doit permettre de choisir, si c'est possible, un modèle de connaissance adapté au contexte courant : distribution de probabilités conditionnelles (pour les architectures centralisées) ou matrice de confusion (pour les architectures décentralisées).

Si on sait que le modèle disponible n'est pas adapté au contexte, on peut agir sur l'estimation de la qualité d'évaluation des probabilités en modifiant la valeur nominale du coefficient q^j (dégradation par rapport aux conditions d'apprentissage).

Les données de renseignement (présence des forces) peuvent permettre d'estimer plus précisément les probabilités a priori des hypothèses.

6.4 Formalisme de Dempster-Shafer

La connaissance externe de l'environnement permet de modéliser l'incertitude de façon appropriée.

La dégradation de la précision du processus d'évaluation (pour les architectures centralisées) ou de décision (pour les architectures décentralisées) peut être simulée par affaiblissement du jeu de masse nominal (cf 4.3).

L'apport d'informations de type renseignement peut être introduit dans le formalisme de Dempster-Shafer sous la forme d'une source externe, représentée par un jeu de masses a priori : l'actualisation de l'évaluation multicapteurs, compte tenu des informations contextuelles, se fait alors par agrégation avec la source externe ($m_{\text{après actualisation}} = m_{\text{capteurs}} \oplus m_{\text{renseignement}}$).

7. Evaluation-comparaison

Pour comparer les performances des différentes architectures et méthodes proposées, il faut évaluer les critères suivants :

- **aspects théoriques** : aptitude à gérer l'incertain, pertinence du critère de décision,
- **aspects pratiques** : influence de la qualité des données de référence, difficulté du transcodage, aptitude à gérer le suivi temporel, à utiliser des données contextuelles, débit d'information, charge de calcul,
- **aspects technico-opérationnels** : robustesse, adaptabilité, modularité, évolutivité, rapidité, réactivité, fiabilité.

En résumé, on peut dire que les architectures décentralisées sont celles qui satisfont au mieux les exigences opérationnelles (modularité, évolutivité, adaptabilité) : elles sont cependant très liées à l'apprentissage des matrices de confusion.

Parmi les architectures centralisées, il semble que les méthodes basées sur la théorie de Dempster-Shafer soient plus robustes, plus aptes à représenter le niveau de discernement des capteurs et à gérer leur complémentarité : mais il y a sûrement un effort à faire pour adapter l'implémentation au problème traité (procédure de transcodage, choix du critère de décision).

L'efficacité des procédures de suivi temporel et d'utilisation du contexte est difficile à évaluer théoriquement ; or, elle conditionne les performances de rapidité, réactivité, fiabilité du système : des simulations comparatives des performances doivent donc être réalisées pour conclure à ce sujet.

Références

- [1] A. APPRIOU, "Probabilités et incertitude en fusion de données multi-senseurs", Revue scient. et tech. de la défense 1991.
- [2] P.L. BOGLER, "Shafer-Dempster reasoning with applications to multisensor target identification systems", IEEE Trans. on SMC, Nov./Dec.1987.
- [3] J. DEZERT, "Vers un nouveau concept de navigation autonome d'engin. Un lien entre le filtrage PDAF et la théorie de l'évidence", Th. de doctorat de l'Université de PARIS XI, 1990.
- [4] L. XU, A. KRZYSAK, C.Y. SUEN, "Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition", IEEE Trans. on SMC, May/June 1992.
- [5] T. LEE, J.A. RICHARDS, P.H. SWAIN, "Probabilistic and Evidential Approaches for Multisource Data Analysis", IEEE Trans. on Geoscience and Remote Sensing, May 1987.
- [6] L. HONG, A. LYNCH, "Recursive Temporal-Spatial Information Fusion with Applications to Target Identification", IEEE Trans. on AES, April 1993.
- [7] A. BELTRANDO, Ph. PEREZ, "Identification globale de porteurs : plan de travail de l'étude préalable", NT CESDA n°1145 ET/DA du 23.7.91.