

**FUSION DE DONNEES MULTICAPTEURS GEOGRAPHIQUEMENT REPARTIS.
PROBLEME DES DONNEES MANQUANTES**

C. Schleiffer ^{†,‡}, R. Reynaud [†], G. Sella [‡]

[†] Institut d'Electronique Fondamentale, AXIS, Université d'Orsay, 91405 Orsay Cedex

[‡] Aérospatiale Missiles, 1 rue Pablo Picasso, 78114 Magny-Les-Hameaux

RÉSUMÉ

L'étude menée se décompose en trois parties. La première traite de la méthodologie conceptuelle de la fusion de données multicapteurs géographiquement répartis. Elle sera illustrée dans la seconde partie par les résultats de simulations. La troisième partie de l'étude a pour but de gérer les effets dus au réseau de communication qui engendrent des retards à l'arrivée des données au centre de fusion.

ABSTRACT

This paper deals with three parts of the methodology of multisensor data fusion: conceptual, applied and management of communication networks effects. In the first part we present an algorithm for tracking targets illustrated in the second part by an example of simulation. The third part explains the effects of transmission networks like delayed data, and the methods adopted to manage them.

I - INTRODUCTION

L'évolution technologique des engins de pénétration (discrétion, autonomie, manoeuvrabilité et brouillage de signal, ...) et des capteurs (tant au niveau de la qualité des senseurs qu'au niveau du traitement du signal associé), conduisent à la nécessité de concevoir une nouvelle génération de systèmes de surveillance qui peut assurer en temps réel et de manière automatique les fonctions de veille, de détection précoce des cibles et de poursuite simultanée de plusieurs cibles. La fusion de données a donc pour but de répondre à la nécessité d'augmenter la vitesse, la capacité et la précision du processus de pistage [1].

II - CONCEPTION DE L'ALGORITHME

2.1 - Architecture du système

On considère une configuration de capteurs géographiquement répartis. La figure 1 donne un exemple de scénario testé. Chaque capteur de type radar, infrarouge ou télémètre est connecté au centre de fusion permettant ainsi la transmission des informations dans le repère local. L'association des données peut se faire de façon centralisée ou décentralisée. Dans le cadre de cette étude, l'architecture adoptée est de type centralisée, chaque capteur transmettant ses mesures brutes au centre de fusion. Les différents aspects et problèmes étudiés dans la conception de ce système sont:

- ◊ optimisation de la performance de détection grâce à l'effet multicapteurs;
- ◊ optimisation du processus de pistage par utilisation de la complémentarité et redondance des données;
- ◊ gestion de données asynchrones par prise en compte de l'hétérogénéité des périodes d'échantillonnage;
- ◊ gestion des effets de transmission lors de l'acheminement des informations à la fonction de fusion.

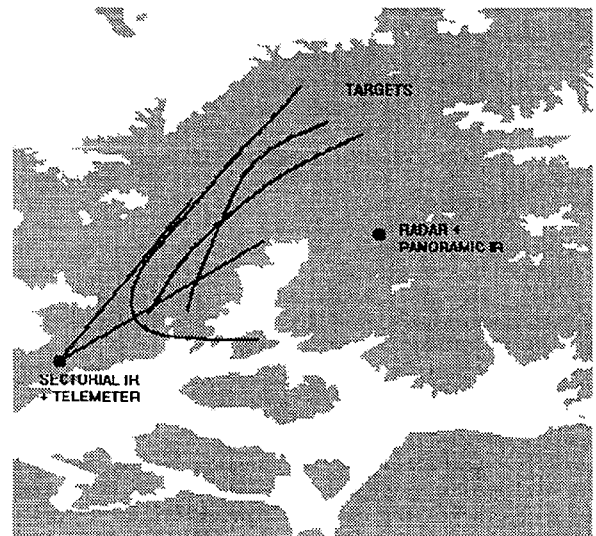


Figure 1: Exemple de scénario

2.2 - Présentation de l'algorithme de pistage et du processus de fusion

Les éléments de base qui composent l'algorithme de pistage sont classiques et sont représentés figure 2.

Le bloc représenté est constitué d'un filtre de Kalman et d'un filtre JPDAF (Joint Probabilistic Data Association Filter) chacun prenant en compte trois modèles dynamiques de cibles [2]. Pour chaque piste, le vecteur d'état estimé par le filtre de Kalman est transmis au filtre JPDAF. L'estimée issue du JPDAF [3] est notée:

$$X_{k/k,S,T} = \sum_{j=0}^{n_{k,S}} \beta_{k,S,T,j} X_{k/k,S,T,j} \quad (1)$$

où $\beta_{k,S,T,j}$ est la probabilité conjointe que la i ème mesure $Z_{k,S,j}$ du capteur S à l'instant k provienne de la piste T (pour

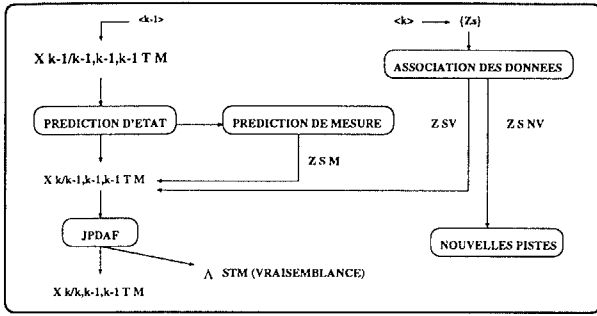


Figure 2: Eléments de base = 1 bloc

plus de détails, voir [4]). L'intérêt du JPDAF est de ne pas attribuer un plot à une piste, mais d'évaluer la probabilité de chaque plot d'appartenir à chaque piste et d'en déduire une position pondérée par ces probabilités.

L'ensemble du processus de pistage est composé de trois blocs dont l'un a été décrit précédemment, chaque bloc étant affecté à un capteur. L'articulation de ces blocs est représentée figure 3. Le détail de cet algorithme peut être consulté dans [5].

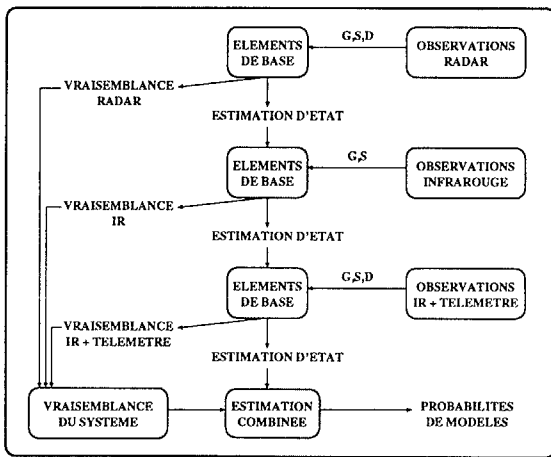


Figure 3: Processus de pistage

Rôles du centre de fusion:

Le centre de fusion contrôle l'origine et la date de chaque observation qu'il reçoit. Les données étant asynchrones, nous avons choisi de les traiter sous la forme de flots de données. L'écart maximum en temps entre deux données du même paquet est de 0.2 seconde et la période d'échantillonnage est calculée à chaque pas. Le centre de fusion sélectionne les données transmises et les oriente vers le bloc correspondant. Si un capteur ne transmet pas d'observation (cas de panne capteur, masquage ou non détection), le bloc de fonctions correspondant au capteur est contourné. On voit aisément que cette méthode peut être étendue à un nombre plus grand de capteurs en faisant varier le nombre de blocs de traitement des mesures.

Fusion de données:

Cet algorithme met en jeu deux types de fusion de données, le premier se situant à la jonction de chaque bloc par utilisation des estimations liées à un capteur, en entrée du traitement associé aux mesures du capteur suivant. Cette méthode permet d'affiner l'estimation d'une cible en prenant en compte

les données des autres capteurs. Le deuxième lieu de fusion de données se situe en fin de chaîne de traitement par le calcul d'une estimation globale de chaque cible. Elle est obtenue, pour chacune des pistes, par combinaison des vecteurs d'état $X_{k/k, k, k, T, M}$ pondérés par les probabilités de modèle ainsi que par les vraisemblances des modèles associés à chaque capteur. L'estimation résultante est donc indépendante des modèles et représente l'état de chaque cible à l'instant k.

III - RESULTATS DE LA SIMULATION

Pour tester cet algorithme, nous avons choisi un scénario composé de situations critiques pour un processus de pistage, c'est-à-dire composé de cibles qui se croisent, de cibles très proches ou de cibles manoeuvrantes. Les performances d'un algorithme de pistage sont directement liées à sa capacité à gérer et à suivre de telles situations, sans interruption dans l'estimation des trajectoires et avec une bonne précision. Les simulations sont faites sur des données synthétisées et calculées à l'aide de modèles fins des capteurs, considérant la météorologie et la situation géographique avec possibilités de masquage des cibles. Les plots de détection obtenus sont donc réalistes. Dans un premier temps, nous ne prenons pas en compte les moyens de transmission et nous considérons que les mesures capteurs sont directement transmises au centre de fusion, sans retard ni perte des informations.

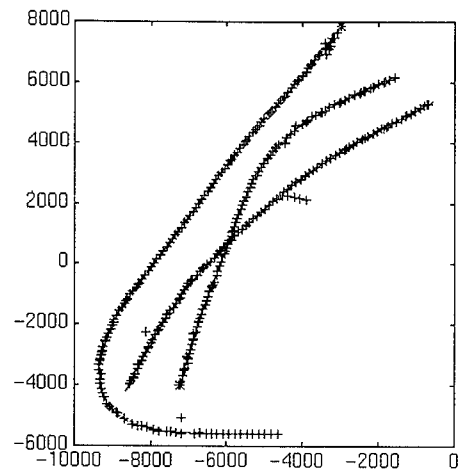


Figure 4: Résultat du pistage

La figure 4 représente la position en X et Y de chaque cible dans le référentiel du centre de fusion. On peut noter que les résultats sont satisfaisants. En effet, chaque cible est pistée sur l'ensemble de sa trajectoire avec une bonne précision. On remarque après le croisement, la détection d'une quatrième cible qui est amorcée puis abandonnée car elle n'est pas réactualisée après un certain nombre de pas. On peut expliquer sa présence par le fait que la période d'échantillonnage est de 0.2 secondes, ce qui entraîne le traitement simultanée de mesures représentatives d'états différents. Il a donc été envisagé pour la suite de l'étude, de traiter à chaque échantillonnage, les mesures correspondant au même instant de détection. Lors de la manoeuvre, on peut noter que le processus assimile le mouvement de la cible en augmentant les probabilités des modèles dynamiques 2 et 3. Les estimations selon ces deux modèles ont donc un poids fort dans l'équation de pondération permettant de calculer l'estimation globale.

De façon similaire, lorsque la cible retrouve une trajectoire rectiligne, le processus de pistage détecte ce changement de mouvement et attribue un poids fort au premier modèle dynamique. De même, lors du croisement des deux cibles, on peut constater le bon suivi des trajectoires, sans échange ni ambiguïté.

IV - RESEAU DE TRANSMISSION DES INFORMATIONS :

Le réseau de communication étant composé de canaux lents et de canaux rapides, le temps de transmission des informations n'est pas constant: les observations peuvent arriver au centre de fusion dans un ordre non chronologique et sont souvent asynchrones entre les différents capteurs. Le problème qui se pose alors est la gestion des données manquantes et retardées, avec comme questions principales:

- ◊ que doit on faire lorsque l'ensemble des données capteurs sont manquantes?
- ◊ doit on les attendre?
- ◊ doit on à leur arrivée les intégrer au processus?

Nous essayons de répondre à ces questions en faisant une estimation des données attendues et en quantifiant leur pertinence. Nous utilisons la théorie de l'information pour estimer ces observations et pour définir des critères de représentation de la pertinence dans différentes situations.

4.1 - Effets du réseau de transmission

Le but de l'étude est de montrer l'influence de la prise en compte ou non de données retardées sur la qualité du processus de pistage, c'est-à-dire sur le pouvoir de séparation entre les hypothèses d'association plot-piste, et l'influence sur la précision et continuité des pistes. Nous introduisons la notion de mesure de pertinence de l'information pour déterminer l'intérêt d'un flot de données et décider de son exploitation par un retour en arrière dans le processus de pistage, ou de son abandon pour accélérer le processus. La figure 5 présente un type de situation rencontrée.

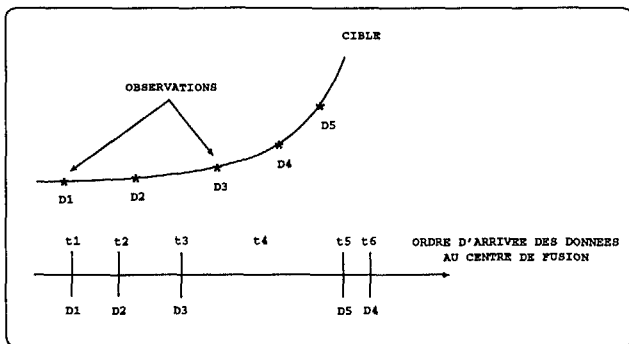


Figure 5: Exemple de données retardées

4.2 - Gestion des données manquantes

Le raisonnement adopté est le suivant:

- ◊ première étape: au temps t4, le centre de fusion peut détecter un retard de transmission à partir des

fréquences de balayage des capteurs, de la période d'échantillonnage et d'une évaluation du temps moyen de transmission;

- ◊ deuxième étape: au temps t5, après réception du flot de données D5 et examen de la date de détection, le centre de fusion est capable d'estimer les données D4 associées à une variance raisonnable;
- ◊ au temps t6, après arrivée des données retardées D5 au centre de fusion, la comparaison avec leur évaluation permet de décider de l'exploitation ou de l'abandon. A ce niveau, on parle de pertinence des informations.

4.3 - Définition de la méthode

Au temps t5, estimation des données manquantes: D4

La première idée consiste à prédire la position de la cible en utilisant les équations de Kalman. On obtient un vecteur d'état et une variance associée qui définit un volume dans lequel la donnée manquante est présumée se trouver. On définit de façon aléatoire un ensemble de vecteurs d'état appartenant à ce volume. Le but est de choisir le vecteur d'état le plus représentatif de l'observation manquante. Pour cela, on utilise la théorie générale de l'évidence [6], et spécialement une règle de décision qui permet de minimiser le risque total pour le choix de l'estimation de la donnée manquante.

On définit les notions suivantes:

"H0" est la probabilité que le vecteur d'état choisi représente bien la donnée retardée et P0 est la probabilité associée. "H1" est l'hypothèse que le vecteur d'état choisi n'est pas représentatif de la donnée manquante et P1 est la probabilité associée. "H0 ou H1" est l'hypothèse de non décision. Par définition:

$$Pr(H0 + H1) = Pr(H0) + Pr(H1) = 1 \tag{2}$$

Soient z la transformation de l'espace des événements O à l'espace des estimées des données retardées E, i.e.,

$$z : O \longrightarrow E$$

et d la transformation de l'espace des estimées à l'espace de décision D, i.e.,

$$d : E \longrightarrow D$$

avec $d_2 = d_o$ or d_1 .

Soit C_{ij} le coût associé à la décision i quand la bonne hypothèse est Hj. Le coût associé à une décision correcte est égal à zero, i.e., $C_{11} = C_{00} = 0$, tandis que le coût associé à une décision incorrecte est supérieur à une non prise de décision, i.e., $C_{ij} > C_{2j}$ pour tout $i \neq 2$ et $j = 0$ ou 1 . Sous ces conditions, la définition du risque total et sa minimisation conduisent à la loi de décision composée des expressions suivantes de la vraisemblance $\Lambda(z)$:

$$\Lambda(z) \begin{matrix} > & \frac{P_0 C_{10}}{P_1 C_{01}} \\ < & \end{matrix} \begin{matrix} d_1 \text{ or } d_2 \\ d_0 \text{ or } d_2 \end{matrix} \tag{3}$$



$$\Lambda(z) \begin{array}{l} d_2 \text{ or } d_1 \\ > \\ < \\ d_0 \text{ or } d_1 \end{array} \frac{P_0}{P_1} \frac{C_{20}}{C_{01} - C_{21}} \quad (4)$$

$$\Lambda(z) \begin{array}{l} d_1 \text{ or } d_2 \\ > \\ < \\ d_2 \text{ or } d_0 \end{array} \frac{P_0}{P_1} \frac{C_{10} - C_{20}}{C_{21}} \quad (5)$$

De façon évidente, la partition de l'espace de décision dépend du choix des valeurs de C_{ij} , et ces tests nous conduisent à sélectionner le vecteur d'état $\hat{D}4$ permettant un risque minimum.

La deuxième méthode d'évaluation de la donnée manquante $\hat{D}4$ consiste à interpoler la trajectoire de la cible entre $D3$ et $D5$. Cette méthode paraît plus simple et diminue le nombre de calculs. On peut de même calculer un risque associé à cette interpolation.

Nous avons pu constater qu'il existe d'autres méthodes d'estimation de données, telle que la celle utilisant l'algorithme EM, mais qui nécessitent un nombre trop important de calculs et qui de part leur complexité et leur lourdeur ne paraissent pas adaptées à l'application désirée dans le cadre de cette étude. La méthode d'interpolation mentionnée ci-dessus, nous apparaît moins robuste dans des situations où le nombre de cibles est important et les manœuvres fréquentes. En effet, le problème d'association plot-piste se pose alors, et il peut y avoir confusion et ambiguïté dans l'estimation des trajectoires.

Définition de deux tests d'aide à la décision

Lorsque les données $D4$ arrivent au centre de fusion, la question principale est alors de savoir si ces données doivent être introduites dans le processus de pistage ou abandonnées.

La définition du premier test d'aide à la décision est simple: si la donnée retardée $D4$ est située dans le volume défini par son estimation $\hat{D}4$ et la variance associée, la quantité d'information contenue dans $D4$ est alors faible de même que sa pertinence. $D4$ ne sera donc pas exploitée par le centre de fusion et sera abandonnée. Inversement, si la quantité d'information contenue dans $D4$ est importante sa pertinence sera élevée. Le centre de fusion intégrera $D4$ au processus de pistage par un retour en arrière sur l'échelle du temps puis réexploitera la donnée $D5$.

Le deuxième test d'aide à la décision utilise le test de Neyman-Pearson [7]. On considère le problème à deux hypothèses, $H0$ désignant l'hypothèse que la pertinence de $\hat{D}4$ est nulle et $H1$ l'hypothèse que la pertinence est égale à 1. U désigne la décision prise par le centre de fusion connaissant la donnée retardée $D4$ et son évaluation $\hat{D}4$. Si la décision du centre de fusion favorise $H1$, alors $U = +1$, sinon $U = -1$. Le centre de fusion utilise le test de N-P conduisant aux équations suivantes:

$$a = \begin{cases} \log \frac{P_0}{P_1} & \text{if } U = +1 \\ \log \frac{1-P_1}{1-P_0} & \text{if } U = -1 \end{cases} \quad (6)$$

avec:

$$\begin{array}{l} H1 \\ a > t \\ < \\ H0 \end{array} \quad (7)$$

Le seuil t doit être déterminé pour permettre une certaine incertitude sur l'évaluation $\hat{D}4$ de la donnée retardée $D4$.

V - CONCLUSION

La mise en place de l'algorithme défini est complexe, mais son caractère modulaire permet de pouvoir l'adapter à tout type de scénario par un simple changement du nombre de blocs de traitement. Il est nécessaire d'affecter un bloc à chacun des capteurs mis en jeu dans le système de détection. Les retards ou pertes d'informations dus aux moyens de communication engendrent des pertes de performances sur le processus de pistage qu'il est impératif de gérer dans un but de précision et de continuité des estimations des trajectoires des cibles.

REMERCIEMENTS

Nous tenons à remercier la DRET et Aérospatiale Missiles qui financent cette étude.

References

- [1] J. Dezert. Vers un nouveau concept de navigation d'engin autonome; un lien entre le filtrage à association probabiliste de données à la théorie de l'évidence. Technical report, Université d'Orsay, Université Paris Sud, 91405 Orsay, France, 1990.
- [2] K.C. Chang Y. Bar-Shalom and H.A.P. Blom. Tracking a maneuvering target using input estimation versus the interacting multiple model algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 25, March 1989.
- [3] Y. Bar-Shalom and K.C. Chang. Joint probabilistic data association in distributed sensor networks. *IEEE Transactions on Automatic Control*, 31:889-897, 1986.
- [4] Y. Bar-Shalom. *Multitarget-Multisensor Tracking: Applications and Advances*. Artech House, 1992.
- [5] R. Reynaud C. Schleiffer and G. Sella. High-performance multi-target multi-sensor tracking for a geographically distributed system. *10th IEEE International Symposium on Intelligent Control*, August 1995.
- [6] S.C.A. Thomopoulos. Sensor integration and data fusion. *Journal of Robotic Systems*, pages 337-372, 1990.
- [7] S.C.A. Thomopoulos and R. Viswanathan. Optimal decision fusion in multiple sensor systems. *IEEE Transactions on Aerospace and Electronic Systems*, 23:644-652, September 1987.