

Estimation des Performances pour la Conception des Systèmes Multi-Capteurs

Jean-François GRANDIN
THALES SYSTEMES AEROPORTES
2, avenue Gay Lussac- 78851 Elancourt Cedex - FRANCE
Email : jean-francois.grandin@fr.thalesgroup.com

Résumé. Ce papier décrit les méthodes d'analyse de performances utilisables pour la conception d'un système multi-capteurs. Les fonctions évaluées sont les fonctions principales d'estimation, de classification et d'association (Pistage). L'analyse de performances a des implications importantes non seulement dans la conception du système, mais également dans l'allocation des ressources, la supervision des traitements et le contrôle du système.

1. Conception d'un Système de Fusion Multi-Capteurs

La conception et la mise au point d'un système de fusion Multi-Capteurs nécessitent d'analyser les performances d'ensemble et relatives à chaque composante. Ces analyses doivent permettre d'une part de déterminer quels sont les maillons faibles afin de rétroagir sur leur design, d'autre part de valider le système dans son domaine d'emploi et vérifier la tenue des performances hors domaine d'emploi. On s'intéresse plus particulièrement aux systèmes de surveillance, dont l'objectif est d'assurer la détection, la localisation et l'identification des objets transitant dans le domaine surveillé. Pour ce faire, on dispose d'un certain nombre de capteurs ou plus largement de sources d'information que l'on souhaite faire coopérer. La fusion de données [Grandin, 2004] comporte des fonctions de : estimation, classification (identification), association. On va donc présenter comment mesurer les performances de ces fonctions et comment les méthodes s'étendent à la fusion de données.

Toute conception d'un système de traitement d'information commence par un « bilan système ». Ce bilan système consiste à répertorier :

- les objets et leurs caractéristiques,
- les mesures physiques réalisables,
- les relations liant mesures physiques et caractéristiques des objets,
- les lois de bruit intervenant sur les mesures.

Il faut ensuite sélectionner ou réaliser les capteurs qui serviront de briques de base. Pour cela on dispose des performances et des caractéristiques des capteurs au sens large :

La fonction d'ambiguïté de l'instrument bien connue en traitement du signal (exemple : ambiguïté distance-doppler d'une forme d'onde radar) joue un rôle similaire à la matrice de confusion en association et en identification. Elle permet de connaître l'adéquation entre les mesures réalisées et les valeurs potentielles des caractéristiques. Dans certains cas, on parle

également de courbes iso-mesure ; les courbes iso-mesures étant le lieu des valeurs de caractéristiques qui engendrent une même valeur de mesure (exemple : hyperbole iso-délai de la triangulation hyperbolique). Généralement l'ambiguïté de mesure sera éliminée par l'intégration (ou la fusion) de mesures présentant des iso-mesures différentes.

La résolution de l'instrument indique dans quelles conditions deux objets proches risquent d'être non résolus c'est à dire pris pour un même et unique objet.

La précision de l'instrument est caractérisée par la matrice de covariance qui s'obtient à partir d'une dérivation locale de la fonction d'ambiguïté.

Le biais de l'instrument indique la valeur potentielle de l'erreur systématique de mesure. Un instrument à biais non nul est dit faux. Une mesure non biaisée est dite juste. Le biais n'est pas un défaut rédhibitoire car celui-ci peut souvent être corrigé par une stratégie d'acquisition des mesures (exemple : la double pesée), ou la calibration.

La fiabilité de l'instrument indique le pourcentage de mesures correctes réalisées. Soit un radar mesurant la SER d'une cible. Dans le cas d'une fausse alarme, la mesure de SER ne correspondra à aucune réalité vis à vis des cibles existantes. La PFA (Probabilité de Fausse Alarme) pourra alors être considérée comme la fiabilité de la mesure de SER. La fiabilité de mesure traduit la probabilité que la mesure effectuée par le capteur réponde au modèle de mesure nominal et ne soit pas par exemple une valeur aberrante. Cette fiabilité peut être définie pour chaque mesure (la PFA dépend du rapport S/B donc de la distance et de la SER effectives) ou pour le capteur.

La conception du système consiste à composer les fonctions du système de façon à maximiser des critères de performance comme :

- La précision et la justesse de localisation des cibles,

- La précision et la justesse d'identification,

- L'apparition de fausses pistes ou pistes fantômes, (sorte de Fausse Alarme système), leur durée de vie,

- La non détection ou le non pistage d'une cible d'intérêt, (sorte de Non Détection Système).

- La durée de vie des pistes correspondant à des vraies cibles.

Les délais d'obtention des détections, localisation, identification, mise en piste ; une information précise et juste n'a souvent plus d'intérêt si elle arrive trop tard.

Les questions auxquelles on doit répondre pour architecturer le système sont :

Localisation des cibles :

Sensibilité de l'estimation à la variation de chaque variable.

Complémentarité des capteurs (un capteur voit ou l'autre ne voit pas, un capteur voit quand l'autre ne voit pas).

Supplémentarité des capteurs (la concordance des positions estimées accroît la certitude de l'estimation).

Identification des cibles :

Pouvoir discriminant de chaque variable et contribution de chaque capteur.

Complémentarité et supplémentation des variables et des capteurs.

Complémentarité des capteurs (un capteur sépare des classes de cible que l'autre ne sépare pas).

Supplémentarité des capteurs (la concordance des décisions accroît la certitude de cette décision).

Association :

Importance des tests d'association.

Complémentarité des tests (un test sépare des pistes que l'autre ne sépare pas).

Supplémentarité des tests (la concordance des tests accroît la certitude de l'association)

Les concepts de redondance et de complémentarité de l'information sont à l'origine des bénéfices attendus de la Fusion de Données. Deux informations sont **complémentaires** si l'estimation d'une caractéristique d'un objet n'est possible que grâce à l'utilisation des deux informations ; par exemple, localisation d'un objet au croisement de 2 lignes de visée, surface = largeur x longueur. Les bénéfices apportés par la complémentarité de l'information sont les améliorations de la couverture spatiale et temporelle du système (les capteurs se partagent l'espace-temps), l'amélioration de la résolution spatio-temporelle (meilleure localisation des objets, capacité accrue de dénombrement), la réduction des ambiguïtés en Identification. Deux informations sont **redondantes** ou **supplémentaires** vis-à-vis du problème à résoudre si la connaissance de l'une d'entre elles élimine la nécessité de connaître l'autre. Cette redondance d'information est très utile car la confrontation des données redondantes va permettre de détecter les mesures erronées, de pallier les données manquantes (grâce à la gestion de tests suppléants) et de construire des indicateurs de confiance sur une information. Les bénéfices apportés par la redondance de l'information sont la robustesse (en particulier tolérance aux pannes), l'amélioration de la détection (l'intégration de mesures multiples d'un même événement accroît la certitude de sa détection) et de la confiance (confirmation du même événement), l'extension des missions (complémentarité des principes de mesure : par exemple imagerie radar versus imagerie IR ou visible).

2. Conception des Estimateurs

On pourra trouver un exposé complet sur la conception en estimation avec des exemples d'application dans [Grandin, 1991-2004].

2.1 Introduction à l'estimation

2.1.1 Introduction

Soit une variable X dont la densité $f(x, \theta)$ dépend du paramètre inconnu θ : Comment trouver la « meilleure » valeur possible de θ à partir d'un n échantillon d'observations $X : x_1, x_2, \dots, x_n$? On appellera T un estimateur de θ , et la valeur prise par T pour la réalisation particulière X l'estimation de θ , que l'on notera $\hat{\theta}$.

2.1.2 Notion d'Information

On appelle quantité d'information de Fischer $I_n(\theta)$ apportée par un n échantillon sur le paramètre θ la quantité suivante positive ou nulle : $I_n(\theta) = E \left[\left(\frac{\partial \text{Log}(L(X, \theta))}{\partial \theta} \right)^2 \right]$ où

$$L(X, \theta) = \prod_{i=1}^n f(x_i, \theta)$$

Si le domaine de définition de X ne dépend pas de θ alors :

$$I_n(\theta) = -E \left[\frac{\partial^2 \text{Log}(L(X, \theta))}{\partial \theta^2} \right]$$

Exemple : Soit X une variable aléatoire de densité gaussienne $N(\theta, \sigma)$ où σ est connu.

$$f(x, \theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x-\theta}{\sigma} \right)^2 \right) \quad \text{On a :} \quad \frac{\partial^2 \text{Log}(f(x, \theta))}{\partial \theta^2} = -\frac{1}{\sigma^2} \quad \text{et}$$

$$I_1(\theta) = E \left(\frac{1}{\sigma^2} \right) = \frac{1}{\sigma^2}$$

L'information apportée par une observation sur la moyenne est d'autant plus grande que la dispersion est petite.

2.1.3 Qualités attendues d'un estimateur

Un estimateur doit être :

Précis : la variance d'estimation $E[(T - \theta)^2]$ doit être minimale.

Sans biais : l'erreur moyenne d'estimation $E[T] - \theta$ doit être nulle.

Convergent : lorsque le nombre n des observations augmente T doit converger vers θ .

On se contente de rechercher parmi les estimateurs sans biais, ceux ayant la plus faible variance.

2.1.4 Borne de Cramer-Rao

Si le domaine de définition de X ne dépend pas de θ , on a pour tout estimateur T sans biais de θ :

$$V(T) \geq \frac{1}{I_n(\theta)}$$

On se reportera à [Rao, 1965] pour la démonstration. L'inégalité de Cramer-Rao donne une borne minimale (appelée borne de Cramer-Rao) pour la matrice des covariances de tout estimateur $\hat{\theta}$ de θ . On pourra, sans réaliser l'estimateur, dire quelle sera la performance optimale pour un scénario et des hypothèses statistiques bien définies. La BCR donne les performances optimales, tout en sachant que, pour la plupart des estimateurs, elles ne seront pas atteintes. Un estimateur dont la matrice des covariances atteint la borne de Cramer-Rao est dit "efficace".

2.1.5 Localisation d'une cible par mesures indirectes

2.1.5.1 Modélisation

Sous forme matricielle, le problème d'estimation que nous nous posons est défini par:

Le vecteur des paramètres observés $\mu = (\mu_1, \dots, \mu_N)^T$ de dimension $(N \times 1)$,

Le vecteur des paramètres à estimer $\beta = (\beta_1, \dots, \beta_P)^T$ de dimension $(P \times 1)$,

Le vecteur des relations fonctionnelles $F(\mu, \beta) = (f_1(\mu, \beta), \dots, f_Q(\mu, \beta))^T = 0$ de dimension $(Q \times 1)$,

Le vecteur des mesures du vecteur μ , $m = (m_1, \dots, m_N)$ de dimension $(1 \times N)$,

La matrice des covariances des mesures Σ de dimension $(N \times N)$.

On suppose généralement que les mesures m suivent une densité multinormale de moyenne μ et de covariance Σ .

La densité de probabilité de m est alors :

$$f(m) = |\Sigma|^{-\frac{1}{2}} (2\pi)^{-\frac{N}{2}} \exp\left[-\frac{1}{2}(m - \mu)^T \Sigma^{-1} (m - \mu)\right]$$

Les estimés du maximum de vraisemblance sont les valeurs de μ et β qui maximise $f(m)$ sous la contrainte $F(\mu, \beta) = 0$.

$F(\mu, \beta)$ étant supposé à dérivées partielles continues, on définit les matrices :

$$F_\mu = \frac{\partial F}{\partial \mu} = \left(\frac{\partial f_i}{\partial \mu_j} \right) \text{ et } F_\beta = \frac{\partial F}{\partial \beta} = \left(\frac{\partial f_i}{\partial \beta_k} \right)$$

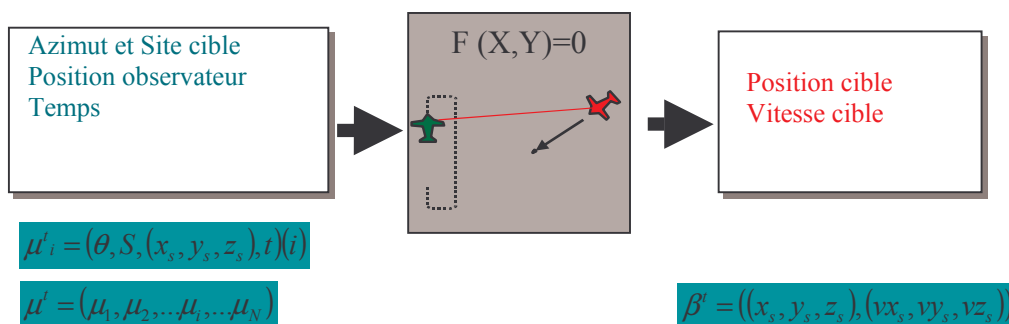
La matrice d'information de Fischer est alors donnée par : $\Phi = F_\beta^T (F_\mu \Sigma F_\mu^T)^{-1} F_\beta$

2.1.5.2 Borne de Cramer Rao:

Dans les conditions ci-dessus définies, la borne inférieure de la matrice des covariances de tout estimateur sans biais du vecteur β est donnée par : $\Sigma_{\hat{\beta}} \geq \Sigma_\beta = \Phi^{-1}$ ($A \geq B$ si $(A - B)$ est une matrice définie non négative)

Remarque: Le calcul de Σ_β se fait à partir de la valeur exacte (μ_0, β_0) qui annule la fonction F .

2.1.5.3 Exemple : Trajectographie air-air par mesures passives



Les relations fonctionnelles sont définies par :

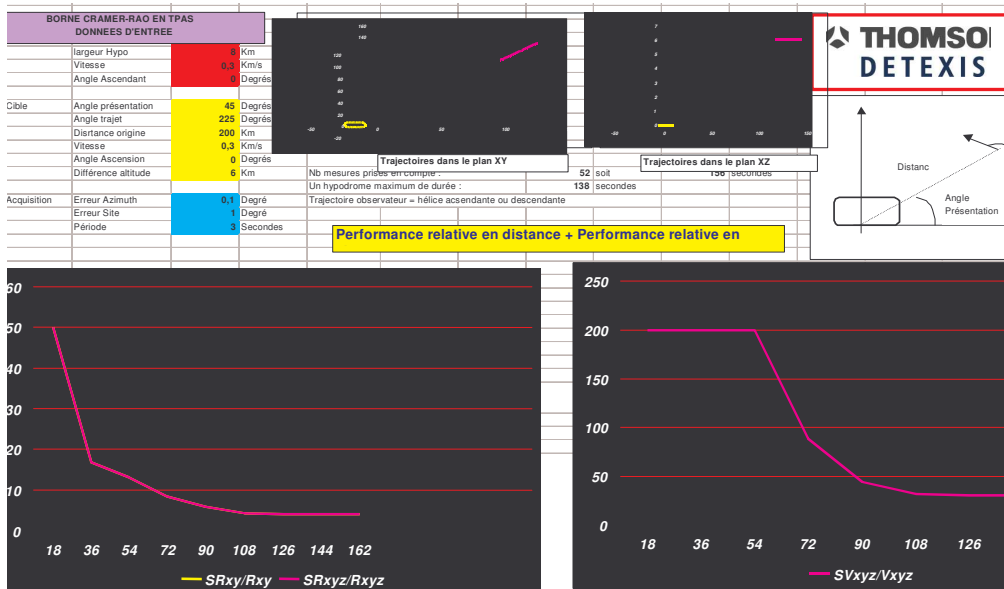
$$f_{i1} : [y_s(t_r) + (t_i - t_r)v_{ys} - y_p(t_i)] \sin \theta_i - [x_s(t_r) + (t_i - t_r)v_{xs} - x_p(t_i)] \cos \theta_i = 0$$

$$f_{i2} : \sqrt{[y_s(t_r) + (t_i - t_r)v_{ys} - y_p(t_i)]^2 + [x_s(t_r) + (t_i - t_r)v_{xs} - x_p(t_i)]^2} \sin S_i - [z_s(t_r) + (t_i - t_r)v_{zs} - z_p(t_i)] \cos S_i = 0$$

On pourra trouver dans [Grandin, 1991-2004] l'expression analytique de la borne, tout calcul fait.

La BCR permet d'établir des abaques de performance optimale en fonction des scénarios envisagés :

Poursuivant l'exemple il s'agit ici de déterminer les performances relatives en distance et en vitesse pour l'interception d'une cible en Mouvement Rectiligne Uniforme (MRU) à partir de relevés angulaires site et gisement. Les graphiques ci-contre montrent la décroissance de la covariance d'estimation en fonction du temps, dans le scénario considéré.



2.1.5.4 Importance de chaque paramètre :

Le but de l'analyse de sensibilité [Beck et Arnold, 1977] est de déterminer pour chaque paramètre la contribution de l'erreur sur la connaissance de la valeur vraie de ce paramètre à l'erreur globale commise sur les estimés. On pourra ainsi isoler les paramètres jouant un rôle prépondérant et dualement identifier les paramètres qui ne posent pas de problème.

L'influence de μ_j sur β_i est caractérisée par le produit : $C_{\beta_i/\mu_j} \cdot \sigma_{\mu_j}$ Où

$C_{\beta_i/\mu_j} = \frac{\partial \beta_i}{\partial \mu_j}$ est le coefficient d'influence.

Ce coefficient est le terme général de la matrice de sensibilité donnée par :

$$C_{\beta/\mu} = \left(\frac{\partial \beta_i}{\partial \mu_j} \right) = - \left(F_{\beta}^T F_{\beta} \right)^{-1} F_{\beta}^T F_{\mu}$$

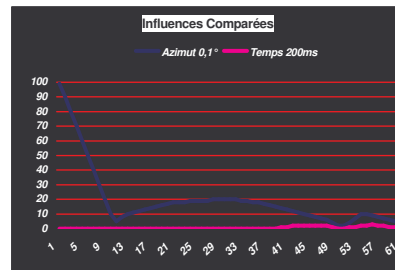
Poursuivant l'exemple de la trajectographie passive air-air on peut tracer l'influence comparée de l'azimut (supposé connu avec un écart type de mesure de $0,1^\circ$) et de la datation (le temps étant mesuré avec un écart type de mesure de 200ms).

Influence comparée de :

Azimuth $\sigma = 0,1^\circ$

Temps $\sigma = 200$ ms

On en déduit que vu les précisions fournies par des horloges classiques qui sont très inférieures, la mesure du temps peut ici être considérée comme sans erreur.



En suivant cette démarche on pourra donc établir les exigences sur les paramètres mesurés, dont en particulier, la datation des informations, la position et l'attitude du système de mesure.

2.1.6 Application à la fusion de données

2.1.6.1 Additivité de l'information

La méthode présentée s'applique à tout problème pouvant être formalisé sous la forme définie plus haut, en pratique tous les problèmes où l'on cherche à estimer des paramètres inconnus à partir d'informations délivrées par un ou plusieurs capteurs. On peut considérer un nombre quelconque de relations fonctionnelles, certaines relations étant définies pour un paramètre et un groupe d'observations particulier.

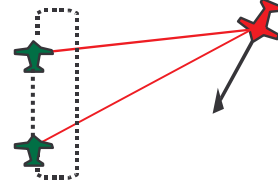
Dans le cadre de la fusion des mesures de plusieurs capteurs on reconduit donc la même méthode. Si les mesures sont indépendantes, les matrices d'informations sont additives, ce qui simplifie la mise en œuvre des calculs. Soit un vecteur de mesure m composé de 2 vecteurs conditionnellement indépendants m_1 et m_2 de vecteurs moyenne μ_1 et μ_2 inconnus et de matrices de covariances Σ_1 et Σ_2 connues, avec les relations de contraintes : $F(\mu_1, \beta)=0$ et $F(\mu_2, \beta)=0$.

Alors [Rao, 1965] : $\Phi = \Phi_1 + \Phi_2$

2.1.6.2 Exemple : Trajectographie passive par 2 observateurs

A titre d'exemple, on étend l'exemple de la trajectographie air-air en considérant 2 observateurs.

Chaque observateur réalise une mesure de site et gisement sur l'observé, et dispose de mesures de position et temps d'observation.



On applique alors la démarche :

avec les relations fonctionnelles liant observateur1 et paramètres inconnus, soit les 2n équations :

$$f_{i1}^{C1} : [y_s(t_r) + (t_i - t_r)v_{ys} - y_p^{C1}(t_i)] \sin \theta_i^{C1} - [x_s(t_r) + (t_i - t_r)v_{xs} - x_p^{C1}(t_i)] \cos \theta_i^{C1} = 0$$

$$f_{i2}^{C1} : \sqrt{[y_s(t_r) + (t_i - t_r)v_{ys} - y_p^{C1}(t_i)]^2 + [x_s(t_r) + (t_i - t_r)v_{xs} - x_p^{C1}(t_i)]^2} \sin S_i^{C1} - [z_s(t_r) + (t_i - t_r)v_{zs} - z_p^{C1}(t_i)] \cos S_i^{C1} = 0$$

avec les relations fonctionnelles liant observateur2 et paramètres inconnus, soit les 2n équations :

$$f_{i1}^{C2} : [y_s(t_r) + (t_i - t_r)v_{ys} - y_p^{C2}(t_i)] \sin \theta_i^{C2} - [x_s(t_r) + (t_i - t_r)v_{xs} - x_p^{C2}(t_i)] \cos \theta_i^{C2} = 0$$

$$f_{i2}^{C2} : \sqrt{[y_s(t_r) + (t_i - t_r)v_{ys} - y_p^{C2}(t_i)]^2 + [x_s(t_r) + (t_i - t_r)v_{xs} - x_p^{C2}(t_i)]^2} \sin S_i^{C2} - [z_s(t_r) + (t_i - t_r)v_{zs} - z_p^{C2}(t_i)] \cos S_i^{C2} = 0$$

2.1.6.3 Analyses paramétrique des performances

On pourra alors connaître les performances pour des géométries d'observations variées. Pour aller une étape plus loin dans la démarche analytique, il est possible moyennant certaines hypothèses sur la régularité de l'échantillonnage des observations d'estimer la BCR sous une forme analytique explicitant les dépendances aux paramètres définissant la géométrie d'observation [Grandin, 1991-2004].

On obtient alors la géométrie optimale d'observation par dérivation de la BCR analytique. On se reportera à [Grandin, 1991-2004] pour des exemples.

Disposant de plusieurs observateurs, on pourra déterminer par simulation la géométrie d'observation optimale soit la disposition statique ou dynamique optimale des capteurs. Dans le cadre de la Fusion de données, la BCR et le calcul d'influence permettent :

- de quantifier l'apport des différents capteurs, donc de les sélectionner,
- de quantifier les exigences sur la datation et les positionnements absolus et relatifs,
- d'établir une carte des performances sur les volumes de surveillance.

Les procédures de calibration qui s'avèreraient nécessaires sont également basées sur des estimateurs et les exigences relatives sont extraites selon la même démarche.

Dans le cadre de la conception rationnelle d'un système de Fusion, notamment « militaire » il faut raisonner en déterminant les exigences relatives au système, sous-systèmes et capteurs en partant du besoin exprimé.

3. Conception des Classifieurs

On pourra trouver un exposé complet sur la conception des classifieurs avec des exemples d'application dans [Grandin, 1997-2004].

3.1 Introduction à la classification

Par techniques de discrimination on désigne l'ensemble des techniques statistiques qui ont pour but d'identifier des classes d'objets, définies a priori, à l'aide de plusieurs variables descriptives (paramètres). Ces méthodes travaillent sur des représentations vectorielles d'une forme. L'ensemble des formes à classer est caractérisé par un ensemble de points dans R^p , où chacune des P variables représente un « trait caractéristique » (Feature en anglais) de la forme à reconnaître.

On nomme variable à expliquer, la variable codant le numéro de classe réelle des exemples d'apprentissage et variables explicatives les variables constituant les tests décisionnels.

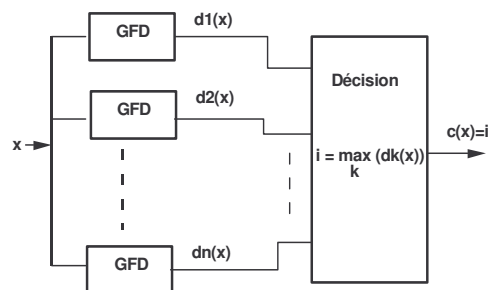
La discrimination consiste donc à configurer un classifieur qui, à toute observation x , associe une classe d'appartenance $C(x)$. L'attribution d'une observation à une classe se fait grâce à des fonctions dites discriminantes d_i .

Problème :

Configurer un classifieur C , qui à toute entrée x , associe une décision de classe $C(x) = 1, 2, \dots, n$ quand n classes sont possibles.

Résolution :

La résolution consiste généralement à calculer des fonctions discriminantes d_i et à les appliquer au pattern à tester suivant le critère du maximum.



Lors de la phase inférencielle, les règles de classement d_i sont construites à partir d'un ensemble d'apprentissage constitué par des couples $(x, C(x))$ où $C(x)$ donne la classe d'appartenance de x . Cet ensemble d'apprentissage est une donnée d'entrée du problème de discrimination.

La phase décisionnelle ou généralisation consistera à attribuer une classe à toute nouvelle observation.

La construction du discriminateur est toujours réalisée à partir d'échantillons de taille limitée et la question du bien-fondé de ce discriminateur pour une population plus vaste est une question cruciale. Ainsi l'analyse discriminante donne-t-elle une place très importante à l'évaluation de la pertinence et de la stabilité des résultats proposés par les différentes techniques.

D'après ce que nous avons pu voir, seule l'attribution d'une classe à une observation compte. En fait, il s'avère très utile, pour une observation, de connaître non seulement la classe à laquelle elle appartient "majoritairement", mais d'avoir ses probabilités d'appartenance aux différentes classes. Dans ce cas, les fonctions discriminantes s'expriment en termes de probabilités, ce qui permet d'introduire des notions de base (probabilité d'erreur de classement, seuil de rejet...) qui sont nécessaires pour intégrer (fusionner) des modules de discrimination dans un système complexe de décision.

3.2 L'analyse de la réalité des classes

Une hypothèse de base de la reconnaissance statistique est que chaque classe forme dans R^p un nuage compact. Si une classe comporte plusieurs modalités il est essentiel de le détecter et de séparer cette classe d'objet en autant de sous classes que de modalités ; sinon cela pourrait perturber certaines procédures d'analyse discriminante.

Pour cela on procède généralement comme suit (pour chaque classe):

1/ recherche sur-numéraire de centroïdes représentatifs par la méthode des moyennes mobiles.

2/ construction de l'arbre hiérarchique mettant en évidence les distances entre sous-classes centroïdes.

3/ mise en évidence sur l'arbre hiérarchique de « gaps » significatifs entre sous-classes. Découpe ou regroupement en autant de sous classes que nécessaire.

La même procédure sera reproduite sur l'ensemble des classes afin de mettre en évidence les « classes réelles » et mesurer la difficulté du problème de discrimination posée (atteinte des « classes désirées »). On a en effet accès grâce à l'arbre hiérarchique à une représentation planaire des classes, ce qui permet d'appréhender le problème de discrimination en voyant quelles classes sont bien séparées et quelles autres sont trop rapprochées.

On mesure par ailleurs la corrélation entre les classes naturelles et les classes désirées. On peut à cette fin synthétiser le résultat par une matrice de confusion qui croise les deux types de classe. Ce score de corrélation indique si la représentation paramétrique du problème permet ou non sa résolution efficace.

Toute bonne analyse est poursuivie par la vérification d'un certain nombre d'hypothèses ; ainsi on mesurera pour chaque sous-classe :

La dépendance linéaire des variables 2 à 2, un test discriminant ne devant être construit qu'avec des variables peu ou pas redondantes.

La dispersion des valeurs d'une classe d'objet donnée sur chaque variable (moyenne et covariance, classe par classe).

La normalité ou la log-normalité des bruits perturbateurs selon ces variables.

3.3 Estimation des performances asymptotiques

On doit avant tout estimer les performances asymptotiques liées au problème et non à la méthode de classement. L'estimation des performances asymptotiques revient à une estimation de l'erreur de Bayes notée EB. L'estimation de EB est intéressante car elle indique si les mesures effectuées (comprendre les paramètres de représentation) sont suffisantes pour permettre la résolution du problème avec des performances acceptables. Si ce n'était pas le cas, hors problème de prétraitement, il convient d'améliorer ou de rechercher d'autres types de mesure. Cette estimée de EB servira également de référence en terme d'objectif à atteindre pour les méthodes de classement proposées (rôle similaire à la BCR en Estimation). Il existe plusieurs estimateurs ou techniques d'encadrement de EB : coefficient de Battacharya, Kppv techniques [Fukunaga, 1990], estimation de EB par les « 2 PPV techniques polarisés ». Cette détermination de EB est une des fonctions du logiciel IDH [IDH étude Dret].

3.4 Construction du classifieur

Les méthodes numériques de classification effectuent une modélisation numérique à partir d'exemples qui leur sont fournis dans la phase d'apprentissage. La phase de test dite de "généralisation" consiste à mesurer les performances sur des exemples non appris.

3.4.1 Le problème de la généralisation

Le terme de généralisation se réfère à la réponse du classifieur pour un exemple non appris. On distingue l'interpolation et l'extrapolation. Les méthodes de classification sont capables d'interpoler : elles peuvent attribuer à une entrée qu'elles n'ont pas apprise la bonne valeur de sortie pourvu que des exemples suffisamment proches de cette entrée aient été appris lors de l'apprentissage et qu'il n'y ait pas de surapprentissage (nombre trop grand de paramètres de la décisionnelle par rapport au nombre d'exemples de la base d'apprentissage). En revanche, l'extrapolation n'est pas possible : la réponse à un exemple éloigné de tous les prototypes de la base d'apprentissage sera fausse a priori.

3.4.2 Stratégies d'apprentissage

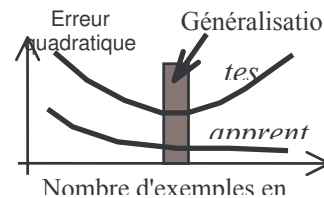
La stratégie la plus employée pour contrôler la généralisation consiste à découper en trois parties l'ensemble des exemples dont on dispose :

La première partie sert de **base d'apprentissage**, c'est elle que le classifieur est prié d'apprendre,

La deuxième partie sert de **base d'entraînement (ou de test de généralisation)**: périodiquement, on regarde les performances de la décisionnelle sur cet ensemble de test. Si elles sont suffisamment bonnes, on décide de figer l'algorithme. Sinon on itère en sélectionnant les tests les plus généralisants.

La troisième partie, la **base de validation (ou généralisation)** permet d'estimer les performances du classifieur en utilisation sur des exemples qu'il n'aura jamais vu.

Lorsque l'on augmente la taille de la base d'apprentissage, l'erreur quadratique moyenne sur cette base diminue toujours, elle tend vers 0 lorsque les données sont non-bruitées, et vers la variance du bruit lorsque les données sont bruitées. L'erreur quadratique moyenne sur la base de test, elle, passe, par un minimum; si l'apprentissage se prolonge au-delà, les performances en test diminuent.



En général, on admet que l'erreur en apprentissage et l'erreur en généralisation bornent la valeur des performances réelles et on se contente de cet encadrement.

Lorsque la complexité du classifieur est bien adaptée au problème posé, l'écart entre apprentissage et généralisation au point de meilleur fonctionnement atteint lui aussi un minimum. Ainsi dans [Ledoux et Grandin, 1994] on montre qu'une procédure de simplification sélective d'un classifieur consistant à éliminer les tests de variance trop élevée amène progressivement à réduire cet écart, permettant d'obtenir une performance de généralisation maximale.

Le choix de l'espace de représentation est un compromis à effectuer entre la fidélité de la description des formes originales, et la performance de discrimination dont principalement la robustesse. La théorie de la minimisation du risque empirique de Vapnik [Vapnik, 1982] nous apprend qu'une représentation trop riche conduit à accroître la variance de l'estimation de la fonction discriminante tandis qu'une représentation trop réduite conduit à un biais d'estimation. Il est important pour la robustesse de la fonction discriminante de réaliser des tests de dimension aussi réduite que possible. Cela ne veut pas dire que les autres variables doivent être abandonnées ; bien au contraire elles serviront à construire des « tests suppléants » qui permettront d'une part de pallier à l'absence momentanée d'une des variables du test principal et d'autre part à mesurer l'incertitude sur le résultat du test.

3.4.3 Mesure des performances de classification

La matrice de confusion est donnée par :

$$p^c(\omega_j / \omega_i) = \sum_{m^c} p(\omega_j / m^c) p(m^c / \omega_i)$$

Pour en avoir une quantification précise il faut disposer d'un nombre suffisant de mesures m^c représentatives de l'ensemble des situations possibles (on parle d'exhaustivité de la base d'apprentissage). En pratique, il faut faire attention que la base de données utilisée pour l'analyse ne privilégie pas certaines classes (effectifs plus ou moins nombreux des classes).

Pour faciliter l'analyse de la matrice de confusion il est utile de restructurer celle-ci par échange de lignes et colonnes afin de faire apparaître une structure diagonale avec des blocs [Grandin et Souchay, 1993]. Il est alors possible de séparer :

L'ensemble des classes reconnues à 100%.

Les classes diffuses qui en empiétant sur de nombreuses autres classes perturbent la classification.

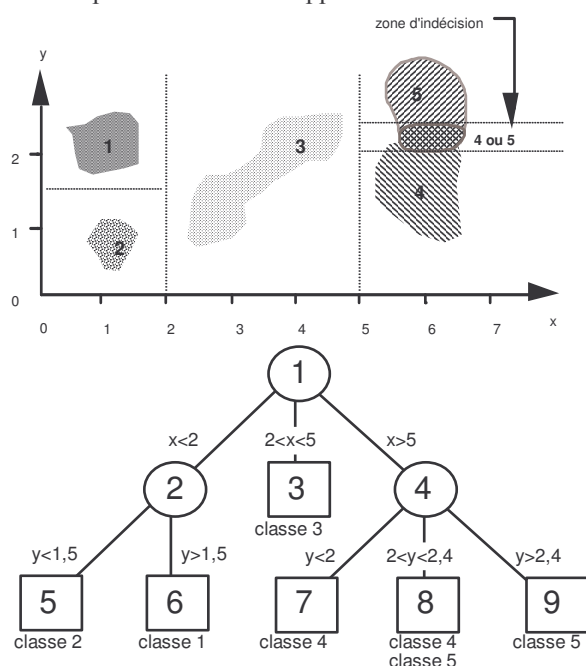
Les classes recouvrantes qui ne sont pas séparables par les paramètres de représentation choisis. Il sera possible d'apporter localement l'information discriminante.

Dans les endroits de l'espace de représentation où les classes se recouvrent il est nécessaire de modéliser statistiquement les lois de distribution sous-jacentes afin de calculer une probabilité d'appartenance pour chaque classe présente.

3.4.4 Intérêt des méthodes de segmentation

Il existe de nombreux types de classifieurs. Parmi ces procédures les méthodes segmentation soit de construction d'arbres de décision sont basées sur la construction hiérarchique d'un ensemble de tests de dimension réduite, associée à la composition logique des tests inhérente à la structure d'arbre et permettant ainsi de suivre des configurations non linéaires par composition logique de tests simples donc intelligibles.

On peut schématiser l'approche comme suit :



Les méthodes de construction d'arbres de décision offrent pour l'analyse des données les avantages suivants :

La complexité est faible, ce qui confère au classifieur un grand pouvoir de généralisation et une grande robustesse.

Les classes multi-modales sont isolées et traitées automatiquement sans problème particulier.

Le pouvoir discriminant des variables explicatives est mesuré.

La description est facilement interprétable.

L'utilisateur dispose à chaque nœud de décision de la représentation des données totales ou locales, ce qui lui permet d'appréhender localement le problème, de confirmer les tests sélectionnés automatiquement mais aussi de découvrir d'éventuels erreurs d'étiquetage toujours possible (erreur sur la réalité terrain), voire d'introduire localement des connaissances utiles à la décision.

Un des avantages, fondamental en pratique (basé sur la capacité d'évaluation du caractère discriminant des variables) est de construire des tests redondants (dits test suppléants) permettant d'accroître la robustesse du système de Fusion.

Les méthodes de ce type ont été popularisées par Breiman [Breiman, 1984]. THALES SA a amélioré ce type de méthode en lui adjoignant des capacités d'estimation locales de densité [Grandin et Jacq, 1993] et d'apprentissage en temps réel, puis a réalisé sur contrat DRET des extensions de ce type de principe, le logiciel IDH [IDH étude Dret] applicable à la plupart des problématiques de classification.

3.4.5 Synthèse des performances d'un système de capteurs

Dans le cas de 2 capteurs, on obtient :

$$p^F(\omega_j / \omega_i) = \sum_{m^1, m^2} p(\omega_j / m^1, m^2) p(m^1, m^2 / \omega_i)$$

Dans le cas où les deux capteurs délivrent chacun une décision hard : $d_i^c = \omega_i$ alors :

$$p^F(\omega_j / \omega_i) = \sum_{(d_m^1, d_n^2) / d^F = F(d_m^1, d_n^2)} p^1(d_m^1 / \omega_i) p^2(d_n^2 / \omega_i)$$

Ce qui dans le cas d'une fusion logique de type Intersection (et logique) conduit au produit des matrices de confusion :

$$p^F(\omega_j / \omega_i) = p^1(\omega_j / \omega_i) p^2(\omega_j / \omega_i)$$

La fusion par « et logique » constitue un extrême, recommandé lorsque les confusions sur chaque capteur sont importantes. Cette fusion conduit à un taux de rejet maximal. La matrice de confusion résultant du produit constitue une première vision des complémentarités attendues de la fusion des 2 capteurs. La matrice de confusion résultant de tests logiques élémentaires peut être calculée en suivant la méthode exposée dans [Hill et al., 2003]. Pour avoir une quantification précise, dans le cadre d'une fusion optimale, il sera nécessaire d'utiliser les procédés décrits dans le cadre mono-capteur et permettant d'apprécier la performance réalisable à partir des informations combinées.

3.4.5.1. Supplémentarité des variables (des capteurs)

Lorsque l'on dispose de nombreuses variables le test décisionnel recherché peut être établi par plusieurs jeux de variables quasi équivalents. Ainsi il sera possible grâce à plusieurs tests équivalents, de suppléer l'absence ponctuelle d'une mesure ou de détecter une valeur aberrante grâce au conflit engendré sur les tests. La construction de tests suppléants a été étudiée par Breiman [Breiman, 1984] dans le cadre des arbres de décision.

Un critère important pour un système décisionnel est de produire non seulement une décision contrastée mais également de vérifier grâce à l'ensemble de l'information disponible la cohérence de cette décision. Sous l'angle bayésien, le fait de disposer de plusieurs tests concordants (rapport de vraisemblance favorable à la même hypothèse) renforce la marge décisionnelle donc la robustesse de la décision.

3.4.5.2 Importance des variables (des capteurs)

Outre l'utilisation précédemment décrite, les tests suppléants offrent un moyen de mesurer l'importance d'une variable pour le problème posé. La procédure, proposée par Breiman [Breiman, 1984] consiste à sommer les indices de décroissance d'impureté lié dans chaque nœud de décision au test suppléant bâti sur cette variable. Il s'agit en quelque sorte de mesurer l'information apportée par la variable pour la résolution du problème de classification. Par ce moyen on peut classer les variables puis les capteurs par ordre d'importance.

4. Conception du Pistage

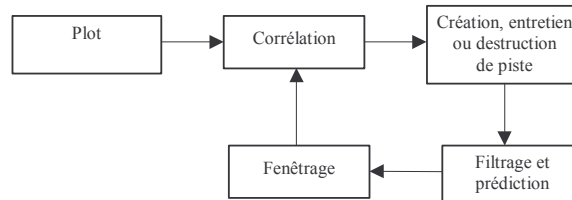
4.1 Rappel sur les méthodes de pistage

4.1.1 L'approche batch

L'approche optimale consiste à traiter globalement les plots délivrés par les capteurs. Le prix de cette approche est une complexité élevée. Les méthodes d'association plot-plot ou méthode de clustering (que nous appelleront méthode batch) cherchent à former une partition des plots où chaque ensemble de plots encore appelé cluster est une piste. Pour atteindre la partition solution, on réalise la maximisation de la fonction de vraisemblance sur l'ensemble des partitions plausibles. La fonction de vraisemblance étant minimale dans le cas où tous les plots formeraient chacun une piste, il est nécessaire soit de pénaliser la fonction de vraisemblance par un critère contradictoire basé sur une évaluation de la complexité du modèle soit de superviser l'optimisation par des critères ad hoc.

4.1.2 L'approche séquentielle

Pour des raisons de ressources limitées, en temps réel contraint, l'approche batch peut être abandonnée au profit d'une approche séquentielle.



Le fonctionnement général du pistage peut alors être résumé par le schéma ci-contre :

La corrélation réalise l'association plot - piste ; La gestion de piste considérera alors 3 cas:

Le plot est associé à une piste existante \Rightarrow entretien de piste,

Le plot n'est associable à aucune piste \Rightarrow initialisation d'une nouvelle piste,

Pour certaines pistes il n'y a pas de plot correspondant \Rightarrow destruction potentielle de piste.

Le filtrage réalise l'estimation à l'instant courant k et la prédiction à l'instant $k+1$.

Le fenêtrage a pour objectif de réduire le coût de calcul de la corrélation en générant grâce à un mécanisme peu cher la liste des pistes possibles pour la mesure.

4.1.2.1 Le critère d'association

Sauf cas trivial, l'association ne se résume pas au calcul d'un simple écart quadratique entre plot et piste, mais recherche une véritable « interprétation des données » de probabilité maximale. Afin de garantir une cohérence globale des différents critères d'association, il est indispensable d'utiliser au niveau de chaque paramètre une distance spécifique au type d'information considérée et fondée sur la « géométrie de l'information » puis de rendre ces critères commensurables. Le raisonnement probabiliste permet d'atteindre cet objectif.

4.1.2.2 La décision d'association

Un principe de base en traitement d'information est de toujours prendre les décisions « faciles » en premier et de différer les décisions risquées. En présence de pistes bien séparées, un calcul d'association immédiat à la plus proche piste peut assurer une décision sans risque. Dans les zones de confusion, il serait regrettable de prendre une décision hâtive dont les conséquences vont se propager dans les résultats de traitement futurs.

Si plus d'une piste passe le seuil d'association, on élabore la marge de sécurité de l'association, c'est à dire le ratio entre la compatibilité maximale obtenue et la compatibilité immédiatement inférieure. Si la marge de sécurité est suffisamment grande, on décide l'association et on réalise la combinaison des données. Si la marge de sécurité est inférieure au seuil choisi, il existe alors deux méthodes de traitement :

la première « simpliste » consiste à mettre les mesures dans un espace de rejet et à les réexaminer à la lumière de nouvelles informations ou à renoncer à les inclure dans le processus de décision,

la deuxième consiste à propager dans le temps les hypothèses correspondant aux différentes décisions en concurrence, en attendant que les informations acquises permettent de lever l'ambiguïté de décision.

4.1.2.3 L'importance de la qualité des pistes

A la suite du traitement des plots, on dispose d'un certain nombre de pistes. Une grande partie d'entre elles reflètent la réalité ; les autres sont des mélanges de plots de plusieurs origines, des résidus d'éclatement. Il est souhaitable que ces quelques fausses informations ne viennent pas perturber le fonctionnement du système en venant altérer sa vue de la situation. Un des buts de la gestion des pistes est d'éliminer ces pistes nuisibles ou au moins de différer leur prise en compte par des traitements décisionnels (une décision sur des données incertaines est risquée). En s'appuyant sur des facteurs de confiance relatifs aux pistes cette élimination est immédiate : toute piste de score (ou de probabilité) inférieur à un certain seuil est abandonnée ou différée.

4.1.2.4 Le fenêtrage

Le fenêtrage a pour but d'éviter une comparaison fine de chaque plot avec toutes les pistes. Cette fonction repose souvent sur un mécanisme d'indexation rapide des données en mémoire. L'index utilisé doit être construit afin de garantir la sélection de la bonne piste parmi les pistes candidates à la notation fine même en présence de fausses mesures. On parle d'« indexation tolérante ». On peut employer ici un vote majoritaire sur un ensemble de tests suppléants.

4.2 Calcul des performances de pistage

4.2.1 Méthode de calcul des performances

Dans un problème de pistage le calcul des performances est trivial lorsque les plots (description d'un objet à l'instant d'acquisition courant) sont largement espacés dans l'espace de représentation. Dans ce cas les plots sont toujours associés correctement et leur succession engendre une piste « correcte ». Les seules performances à vérifier sont les performances classiques des estimateurs soit principalement : le biais, la précision (matrice de covariance), l'efficacité (atteinte de la BCR) et la convergence de l'estimateur.

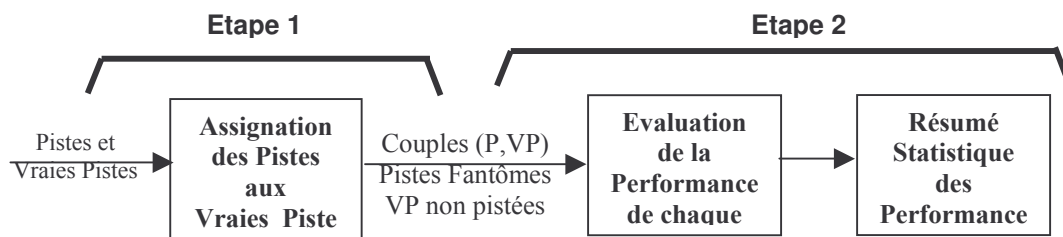
Juger des qualités d'un pistage devient beaucoup plus difficile en présence de mauvaises associations, liées principalement aux ambiguïtés d'association entre un plot et plusieurs pistes proches. Une possibilité est de se ramener au cas simple précédent par la méthode suivante :

Etape 1 : Il faut au préalable associer Pistes et Vraies_Pistes par paire. L'association (Piste-Vraie_Piste) (on dira aussi la mise en correspondance) sera réalisée comme proposée dans [Colegrove et al., 2003] par une assignation 2D optimale [Drummond et al., 1990]. L'association optimale (Piste-Vraie_Piste) utilise une matrice de distance (Piste-Vraie_Piste). La distance d'association (Piste-Vraie_Piste) doit être seuillée pour ne pas

polluer la perception des critères de performance. Deux types de distance peuvent être utilisés : D'une part la distance basée sur la matrice de confusion, d'autre part la distance basée sur les mesures (distance du Khi2). Cette deuxième distance présente l'intérêt d'être plus conforme à la nature des décisions dans un pistage opérationnel.

Etape 2 : Une fois la mise en correspondance réalisée il ne reste plus qu'à rendre compte des erreurs de pistage.

La méthode d'analyse de performance peut être schématisée comme suit :



4.2.2 Evaluation des Associations

Les erreurs résultant d'un algorithme d'association sont de deux types :

- mélange** : une piste contient des plots relatifs à plusieurs Vraies_Pistes,
- éclatement** : les plots d'une Vraie_Piste sont répartis sur plusieurs pistes.

4.2.2.1 Matrice de confusion en association

Une matrice de confusion permet de visualiser à la fois les mélanges et les éclatements. C'est une matrice dont les colonnes représentent les Vraies_Pistes, et dont les lignes représentent les Pistes. Chaque case (i,j) de la matrice contient le nombre de plots de la Vraie_piste j affectés à la piste i .

On définit la piste de rejet (pas d'affectation) en dernière ligne de la matrice et si besoin la Vraie_Piste 0 (Bruit perturbateur) en dernière colonne.

	1	2	3	4
1	4	80	0	0
2	74	0	8	2
3	0	0	61	35

4.2.2.2 .Comportement du pistage

On observe le comportement général suivant [Blackman et Popoli, 1999] :

Zone 1 : Corrélations non ambiguës

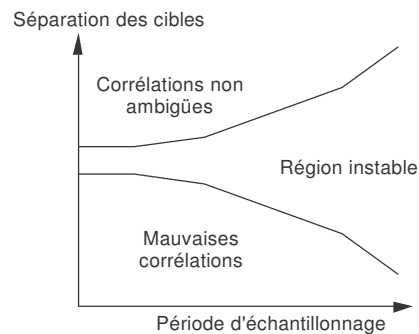
Cette zone représente les cibles très espacées. La zone est d'autant plus importante que T est faible.

Zone 2 : Région instable

Cette zone représente les cibles moyennement espacées. De mauvaises corrélations peuvent apparaître dans cette zone, de plus l'abandon de pistes est très fréquent.

Zone 3 : Mauvaises corrélations sans abandon de pistes.

Cette zone représente les cibles faiblement espacées. De mauvaises corrélations apparaissent dans cette zone, il n'y a pas d'abandon en général car les cibles sont très imbriquées les unes dans les autres (le nombre de pistes est en général inférieur au nombre de cibles).



En cas de mauvaises associations on peut observer différentes situations :

Une piste peut être composée en majorité de plots provenant d'un même émetteur ; Dans ce cas on pourra observer une piste « correcte » correspondant à un objet d'origine . On peut parler de détection de l'objet d'origine.

Une piste peut être un mélange de plots de diverses origines ; On peut parler de fausse alarme.

Une piste correspond à un doublon d'une piste déjà existante pour un objet d'origine.

4.2.2.3 Mesure du taux d'ambiguïté intrinsèque de la situation

La qualité du résultat du pistage dépend donc principalement de la séparation naturelle des objets, soit du pouvoir discriminant de chacun des paramètres des objets. Il est possible de mesurer les ambiguïtés potentielles existantes entre deux objets. Si la Vraie_Piste ne possède au cours du temps qu'un mode de description (sa représentation paramétrique est unimodale) l'analyse est très simple car il suffit de vérifier que les 2 valeurs diffèrent. Si par contre la Vraie_Piste possède au cours du temps plusieurs modes de description (l'objet prend pour certains paramètres des valeurs différentes, discrètes) il faut également vérifier que les caractéristiques diffèrent mais permettent de mesurer une similarité « collective » des différents plots de la même Vraie_Piste.

4.2.3 Calcul des performances

D'un point de vue « opérationnel » on s'intéresse généralement à la capacité à suivre chacune des Vraies_Pistes d'origine; Pour chaque objet on s'intéresse à :

La détection de la Vraie_Piste, soit avoir une piste dont les caractéristiques correspondent très exactement à la Vraie_Piste ; Pour cela il convient en général que la piste ne soit pas polluée par des plots provenant d'autres Vraies_Pistes : on parle de mélange et de façon duale de pureté de la piste.

L'éclatement de la Vraie_Piste soit la dispersion dans différentes pistes des plots correspondant à cette Vraie_Piste ;

Le rejet soit les plots de la Vraie_Piste qui sont placés par le pistage en rejet.

L'apparition de pistes fantômes: pistes de mélange ne correspondant à aucune des Vraies_Pistes.

La correspondance (Piste – Vraie_Piste d'origine) étant réalisée il est possible de mesurer les qualités de la piste soit sa justesse (biais), sa précision (covariance), la convergence et l'efficacité.

Cette vision statique ne suffit pas, et on s'intéresse :

au délai de mise en piste de chacun des Vraies_Pistes.

à la permanence de la détection des différentes Vraies_Pistes (la Vraie_Piste est elle toujours associée à la même piste ou autrement dit cette piste est-elle permanente et stable). Au cours du temps certaines pistes peuvent pister différentes Vraies_Pistes d'origine (on parle d' « échange »).

à la durée de vie des fausses pistes.

4.2.4 Performances de confusion

On réalise à partir de la matrice de confusion certains calculs classiques :

Terminologie	Code	Définition
Pureté piste	TP_P	Nombre d'impulsions de la Vraie_Piste majoritaire sur nombre total dans piste.
Distance Piste – Vraie_Piste	D_PVP	$D(VP, P) = [card(VP \cup P) - card(VP \cap P)] / card(VP \cup P)$
Rejet Vraie_Piste	TR_VP	Nombre de plots de la Vraie_Piste rejetés sur nombre total de plots de la Vraie_Piste.
Détection de la Vraie_Piste	TD_VP	Nombre de plots de la Vraie_Piste dans piste maximale où elle est majoritaire, sur nombre total de plots de la Vraie_Piste.
Pollution du à la Vraie_Piste	TP_VP	Nombre de plots de la Vraie_Piste diffusés dans des pistes où elle n'est pas majoritaire sur nombre total de plots de la Vraie_Piste.

On appelle :

Vraie_Piste majoritaire d'une piste : La vraie piste qui fournit à la piste le plus grand nombre de plots.

Piste majoritaire d'une Vraie_Piste : La piste qui récupère le plus grand nombre de plots de la vraie piste.

4.2.5 Performances basées sur la qualité des pistes

Une autre façon de comparer Pistes et Vraies_Pistes est de réaliser cette comparaison du point de vue de leurs caractéristiques respectives. On définit pour cela des mesures d'efficacité du pistage [Blackman et Popoli, 1999] [Rothrock et Drummond, 2000].

On distingue les performances statiques et les performances dynamiques.

4.2.5.1 Performances statiques

Les Pistes et les Vraies_Pistes sont tout d'abord caractérisées par leur statut au niveau du pistage :

Terminologie	Code	Définition
Piste Correcte Unique	Pcu	La piste a des valeurs de paramètres correctes et est l'unique piste pour cette Vraie_Piste.
Piste Correcte Multiple	Pcm	Idem mais il existe d'autres pistes pour cette Vraie_Piste.
Piste Multiple Correcte	Pmc	Piste additionnelle correcte d'une même Vraie_Piste.
Piste Mélange (Fantôme)	PFA	La piste recouvre des valeurs de paramètre de plus d'une Vraie_Piste.
VP non détectée	VP_x	Une Vraie_Piste est non représentée par une piste ou a été absorbée par la piste d'une autre Vraie_Piste.
VP non Observable	VP_no	La Vraie_Piste ne produit pas suffisamment de plots pour créer une piste.
VP mal caractérisée	VP_z	Une Vraie_Piste est représentée par une piste avec un ou plusieurs paramètres incorrects.

Le nombre total de Vraies_Pistes prises en compte pour l'évaluation doit pouvoir être filtré en éliminant les Vraies_Pistes de code VP_no.

Le tableau suivant résume les statistiques à renseigner :

Score Statistique	Interpretation
N_{VP}	Nombre total de Vraies_Pistes
$N_{VP_NO} = \text{Somme}(VP_no)$	Nombre de Vraies_Pistes non observable
$N_{VP_C} = \text{Somme}(Pcu + Pcm)$	Nombre de Vraies_Pistes avec au moins une piste correcte
$N_{VP_X} = \text{Somme}(VP_x)$	Nombre de Vraies_Pistes non pistés
$N_{VP_Z} = \text{Somme}(VP_z)$	Nombre total de Vraies_Pistes représentées par une piste avec au moins un paramètre erroné
NT_P	Nombre total de pistes
$N_P_FA = \text{Somme}(PFA)$	Nombre de FA additionnelles
$N_P_M = \text{Somme}(Pmc)$	Nombre de pistes additionnelles
$N_P_ND = \text{Somme}(Pmc + PFA)$	Nombre de pistes « indésirables »

Les pourcentages suivants sont obtenus :

Score Statistique	Interpretation
$N_{VP_C} / (N_{VP} - N_{VP_NO})$	Pourcentage de Vraies_Pistes avec au moins une piste correcte.
$N_{VP_X} / (N_{VP} - N_{VP_NO})$	Pourcentage de Vraies_Pistes non pistées.
$N_{VP_Z} / (N_{VP} - N_{VP_NO})$	Pourcentage de Vraies_Pistes mal caractérisées.
N_{PP_FA} / NT_P	Pourcentage de pistes Fausses Alarmes.
N_P_M / NT_P	Pourcentage de pistes multiples.
N_P_ND / NT_P	Pourcentage de pistes « indésirables ».

1.1.1.1 Performances dynamiques

Les performances dynamiques sont codifiées comme suit :

Terminologie	Code	Définition
Latence (délai de mise en piste)	VP_LB	Temps requis pour débiter la piste à partir du premier plot de la Vraie_Piste correspondante reçu.
Latence de destruction	VP_LE	Temps requis pour détruire la piste à partir de la disparition de la Vraie_Piste.
Taux de Perte de la Vraie_Piste	VP_R	Pourcentage de temps où la Vraie_Piste n'a pas de piste en correspondance.
Nombre de swap de la Vraie_Piste	VP_S	Nombre de fois où la Vraie_Piste change de piste correspondante.
Eclatement de la	VP_MT	Nombre total de pistes attribuées à une Vraie_Piste

Vraie Piste		
Nombre de Vraies_Pistes stables	N_VP_S	Nombre de Vraies_Pistes pistées continument par au moins une piste.
Nombre de Vraie_Piste instables	N_VP_U	Nombre de Vraies_Pistes pistées avec au moins une perte de continuité.
Vraie_Piste mal caractérisée	N_VP_m C	Pourcentage de fois où la Vraie_Piste est mal caractérisée.
Vraie_Piste mal localisée	N_VP_m L	Pourcentage de fois où la Vraie_Piste est mal localisée.
Vraie_Piste mal identifiée	N_VP_mI	Pourcentage de fois où la Vraie_Piste est mal identifiée.
Taux de swap de la piste.	TS_P	Nombre de fois où la piste change de Vraie_Piste.
Nombre de Fausses Pistes	N_P_FA	Une fausse piste est une piste qui n'est pas associée à une Vraie_Piste. Nombre de fausses pistes.
Durée de la Fausse Piste	P_FA_LD	Une fausse piste est créée. Temps mis pour la détruire.

4.2.6 Evaluation des marges

Dans les méthodes séquentielles, un test d'association est toujours réalisé à partir d'un calcul de similarité entre le plot et la piste suivi d'un seuillage de cette similarité. Afin de qualifier la qualité de la décision d'association on calcule la marge d'ambiguïté soit le rapport entre la similarité maximale et la similarité immédiatement inférieure parmi l'ensemble des similarités plot - pistes. Si la marge d'ambiguïté est faible, l'association du plot pourra être différée jusqu'à l'obtention d'une information discriminante, ou plus simplement placé en espace de rejet.

La mesure du degré de corrélation entre la marge de sécurité lié à la piste et les qualités résiduelles de la piste est un outil permettant de fixer la stratégie de seuillage en rejet qui sera utilisée dans l'algorithme de pistage.

5. Conclusion

La fusion multi-capteurs est en plein essor dans les domaines civils et militaires ; on ne parle plus d'ailleurs de système multi-capteurs mais de système de systèmes. La recherche dans le domaine s'est principalement focalisé sur les algorithmes de fusion. Hors le problème de l'industriel est avant tout de concevoir les systèmes multi capteurs rendant le service désiré au moindre coût.

Il existe grossièrement 2 types de système :

dans le premier type on fusionne des capteurs d'opportunité et il n'est pas rare d'employer des algorithmes très sophistiqués pour pallier les défauts de capteurs non conçus pour être fusionnés,

dans le second type on essaie de concevoir le système et d'en sélectionner les éléments avec une stratégie d'allocation hiérarchique des exigences.

Il est clair que l'industrie militaire s'intéresse principalement à ce second type de réalisation.

Pour réaliser un système multi-capteur efficace il faut au minimum :

- Des références communes soit des bases de temps et de positionnement relatif distribuées ce que réalise le GPS ou des protocoles de synchronisation. Les exigences sur ces références provenant des calculs de sensibilité.
- Un support de communication permettant de transmettre l'information même si celle-ci est limitée au juste nécessaire.
- Des capteurs conçus pour être « fusionable » (conçu pour être facilement intégrable à un système), c'est à dire délivrant au minimum une information et la qualité de cette information (précision et fiabilité).
- Un contrôle commun, qui peut aller jusqu'au contrôle distribué où chaque capteur est un acteur qui participe au contrôle collectif.

Références

[Grandin, 2004] JF Grandin « Fusion de Données : Théorie et Méthodes » Revue Techniques de l'Ingénieur. Parution prévue au 1^{er} semestre 2004.

[Grandin, 1991-2004] JF Grandin Cours sur la construction des estimateurs de Localisation dans le cadre du module de cours ELA 005 de la SAE, (1991-2004).

[Rao, 1965] Rao, « Linear Statistical Inference and its applications » John Wiley & Sons, New York, 1965.

[Beck et Arnold, 1977] J.W. Beck, K.J. Arnold « Parameter estimation in Engineering and Science » John Wiley & Sons 1977.

[Grandin, 1997-2004] JF Grandin, « Cours Réseaux de Neurones ; Ecole Centrale de PARIS 3^{ème} Année Option Mathématiques Appliquées ». (1997 - 2004).

[Fukunaga, 1990] K Fukunaga, « Introduction to Statistical Pattern Recognition - Second edition ». Academic Press Inc. 1990.

[IDH étude DRET] « Identification D'Hostiles » contrat DRET n° 94/396530.

[Ledoux et Grandin, 1994] C Ledoux et JF Grandin « Two original weight pruning methods based on statistical tests and rounding techniques ». IEE Proc. Image Signal Processing, Vol 141, N°4, August 1994.

[Vapnik, 1982] V.N. Vapnik « Estimation of Dependencies based on Empirical Data ». Springer-Verlag, New-York 1982.

[Grandin et Souchay, 1993] JF Grandin et Ph Souchay, « Traitements avancés en guerre Electronique ». AGARD meeting on 'Challenge of Future EW System Design' October 1993.

[Breiman, 1984] Leo Breiman « Classification and Regression trees » Wadsworth & Brooks Books & software 1984.

[Grandin et Jacq, 1993] JF Grandin et C Jacq, « Identification de Pistes Radars par l'Analyse de Données Symbolique-Numérique » JAD 1993.

[Hill et al., 2003] J.M. Hill, M.E. Oxley and K.W. Bauer, « Receiver Operating Characteristic Curves and Fusion of Multiple Classifiers » Sixth International Conference on Information Fusion, 2003 Queensland Australia, pp 815-822.

[Colegrove et al., 2003] S.B. Colegrove, B.Cheung and S.J. Davey, « Tracking System Performance Assessment » Sixth International Conference on Information Fusion, 2003 Queensland Australia, pp 926-933.

[Drummond et al., 1990] O.E. Drummond, D.A. Castanon and M.S. Bellovin “ Comparison of 2-D Assignment Algorithms for Sparse, Rectangular, Floating Point, Cost Matrices ” Journal of the SDI Panel on Tracking, IDA, Alexandria, VA, Issue No 4 pp 81-97 December 1990.

[Blackman et Popoli, 1999] Samuel Blackman, Robert Popoli – « Design and Analysis of Modern Tracking Systems». Artech House 1999.

[Rothrock et Drummond, 2000] R.L. Rothrock and O.E. Drummond, « Performance Metrics for Multiple-Sensor, Multiple-Target Tracking », Proceedings of the SPIE, Vol. 4048, 24-26 April 2000, Orlando USA, pp 521-531.

Summary

This paper describes the methods of performances analysis usable for the design of a system multi-sensors. The evaluated functions are the principal functions of estimate, classification and association (Tracking). The performances analysis has significant implications not only in the systems design, but also in the allowance of the resources, the supervision of the treatments and the control of the system