



DÉTECTION RÉPARTIE ET QUANTIFICATION

M.-P. Boyer et B. Picinbono

Laboratoire des Sigaux et Systèmes
École Supérieure d'Électricité, Plateau du Moulon, 91192 Gif sur Yvette
France

RÉSUMÉ

Nous étudions des structures de détection répartie comprenant un réseau de capteurs et un processeur central de fusion des données. Cependant des considérations de fiabilité et de coût rendent impossible la centralisation complète de l'information. Il est donc important d'étudier la détection répartie lorsque les capteurs transmettent une information réduite au processeur central. De nombreux auteurs ont étudié des systèmes où une décision locale est prise sur chaque capteur et transmise à l'aide d'un seul bit. Afin de ne pas limiter autant le flux d'information, nous proposons une nouvelle approche dans le cadre de plus général de la quantification. L'optimisation globale de telles structures est difficile et nous considérons des systèmes sous-optimaux que nous optimisons à l'aide d'une approche heuristique basée sur l'interprétation graphique des domaines de décision. Nous comparons les performances de ces différentes structures à l'aide de l'exemple simple de la détection d'un signal noyé dans un bruit gaussien.

SUMMARY

In many distributed detection systems separated sensors and a central processor are used. Because of some constraints on communications all the information from the sensors cannot be sent to the central processor. This makes the perception of the system less accurate and induces a loss in performance. Previous studies have considered the case where a detection procedure is realized at each sensor. In this paper a new approach within the framework of quantization is presented. The two procedures are compared in Gaussian noise environment. It is shown that the quantized structure can reach better performance. The optimization of quantized systems according to the Neyman-Pearson criterion is shown to be difficult. Leaving the concept of absolute optimality, we determine the parameters of the quantized systems using heuristics arguments based on the geometrical interpretation of the decision domains.

1. INTRODUCTION

Les performances toujours croissantes demandées aux systèmes radar et sonar ont motivé l'étude de nouvelles structures de détection. Ces structures comprennent souvent plusieurs sources d'information et un processeur central de fusion des données. Il s'agit alors de déterminer la structure des récepteurs locaux, du processeur central et la règle de fusion des données afin d'optimiser les performances pour un objectif particulier. Dans cet article l'objectif en question est la détection qui consiste à choisir de façon optimale selon un certain critère entre deux hypothèses H_0 (bruit seul) et H_1 (signal plus bruit), à partir d'une observation. Au sein des systèmes étudiés, chaque source d'information délivre un vecteur x_i et l'observation globale X est la concaténation de tous ces vecteurs. Selon les cas les vecteurs x_i peuvent être les échantillons pris à la sortie d'un capteur à des instants successifs, les mesures d'un phénomène observé à l'aide de capteurs de nature différente (radar, rayons X, infra-rouge, ...), ou encore dans le cas de la détection répartie les observations de capteurs implantés sur différents sites géographiques. Chaque vecteur X est associé à un point de l'espace produit $E = E_1 \times \dots \times E_N$ où N est le nombre de capteurs et E_i représente l'espace d'observation du capteur i . Le rôle de la détection est d'attribuer à chacune des observations reçues une décision. Ceci revient à diviser l'espace E en deux domaines de décision disjoints D_0 et D_1 tels que si

l'observation tombe dans D_0 (respectivement D_1) la décision est H_0 (respectivement H_1).

La conception des systèmes de détection est liée au choix du critère d'optimisation. Les critères les plus souvent utilisés sont ceux de Bayes et de Neyman-Pearson (N.P.). Celui de Bayes fait appel aux probabilités a priori des deux hypothèses qui dans la réalité sont souvent inconnues. C'est pourquoi nous utilisons le critère de N.P. qui caractérise le problème de la détection à l'aide de la probabilité de fausse alarme Q_0 et la probabilité de détection Q_d . Ces probabilités s'expriment à l'aide des probabilités conditionnelles $p(x_i | H_j)$ que nous notons $p_j(x_i)$ $i = 1, \dots, N, j = 0, 1$ et qui sont connues à l'avance. La formulation de N.P. consiste à trouver le système qui maximise la probabilité de détection Q_d pour une valeur donnée de la probabilité de fausse alarme $Q_0 = \alpha$.

Enfin nous nous plaçons dans l'hypothèse où les vecteurs bruit sont indépendants d'un capteur à l'autre et identiquement distribués sur chaque capteur (hypothèse I.I.D.). Sans cette hypothèse le problème se trouverait très compliqué [1].

2. DÉTECTION RÉPARTIE ET QUANTIFICATION

Le problème de la détection répartie tel que nous venons de le présenter est entièrement résolu par la théorie classique de la



détection. Il est bien connu que le système optimum pour la détection de deux hypothèses doit calculer le rapport de vraisemblance qui est une statistique suffisante pour ce problème, et le comparer à un seuil déterminé par la probabilité de fausse alarme. Ce qui peut s'écrire grâce à l'hypothèse d'indépendance

$$\frac{p_1(X)}{p_0(X)} = L(X) = L_1(x_1) \dots L_N(x_N) \gtrless \eta. \quad (1)$$

Cette formule suggère de calculer localement les rapports de vraisemblance $L_i(x_i)$ et de les transmettre intégralement au processeur central de fusion des données. Des considérations de fiabilité, de coût, ou de largeur de bande passante rendent souvent impossible la centralisation de l'information et imposent de faibles flux d'information au sein des systèmes de détection répartie. Il est donc nécessaire d'étudier le problème lorsque les sources d'information transmettent une information réduite au processeur central.

De nombreux auteurs l'ont abordé sous l'aspect de la détection décentralisée. Cette dénomination caractérise la structure de détection pour laquelle un détecteur est associé à chaque capteur. Les décisions locales sont transmises à l'aide d'un seul bit d'information au processeur central qui évalue la décision globale en fusionnant les informations binaires qu'il reçoit. Tenney et Sandell [2] ont étudié la formulation bayésienne du problème et ont montré que l'optimisation des détecteurs locaux indépendamment du processeur central aboutit à des détecteurs par rapport de vraisemblance. De même le processeur central peut être optimisé indépendamment des détecteurs locaux [3]. Enfin l'optimisation globale des systèmes de détection décentralisée doit prendre en compte l'optimisation simultanée des détecteurs locaux et de la loi de fusion du processeur central [4] [5].

Considérant qu'il n'est pas nécessaire de limiter autant le flux d'information, nous avons proposé une nouvelle approche de la détection répartie dans le cadre de la *quantification*. En effet on peut améliorer la qualité des systèmes de détection répartie tout en respectant la contrainte de flux limité. Pour cela nous allons affiner la sensibilité des systèmes, c'est à dire les rendre capables de discerner un plus grand nombre de situations. Lorsque des détections locales sont transmises, l'information reçue par le processeur central fait partie d'un ensemble de 2^N éléments. Accroître ce nombre lorsque N est fixé nécessite de faire appel à un plus grand nombre de niveaux pour coder les informations locales. En conséquence si nous remplaçons les détections locales par des quantifications nous améliorons la sensibilité ce qui devrait améliorer les performances du système [6]. Toutefois il est clair que le système doit être conçu de façon à exploiter la dynamique offerte par la quantification. C'est ce que nous nous proposons d'étudier dans la suite.

La détection répartie avec quantification s'articule autour de deux niveaux de traitement des informations. Un premier niveau local, de compression des données où la quantification adapte les observations aux flux d'information du système. Un second niveau central, élabore la décision H_0 ou H_1 émise par le système en combinant les informations quantifiées à l'aide d'une loi de fusion. Avant d'aborder l'optimisation du système, identifions ses paramètres. Par la quantification nous voulons réduire l'information contenue dans le vecteur x_i à q valeurs. La quantification vectorielle consiste à diviser l'espace d'observation E_i en q domaines disjoints et à associer à chacun de ces domaines une valeur. Très souvent, pour des raisons de simplicité de réalisation, la quantification vectorielle s'opère par une quantification scalaire de chacune des composantes. Ici aucune contrainte de cet ordre n'est imposée à la structure des quantificateurs. La quantification est donc le résultat d'un

traitement global du vecteur d'entrée. De ce fait l'une des modélisations possibles est d'introduire une fonction scalaire $C_k(x_k)$ du vecteur d'entrée, et $q + 1$ seuils. La partition de l'espace d'observation E_k est alors définie par,

$$x_k \in \Delta_i \Leftrightarrow s_i \leq C_k(x_k) < s_{i+1} \quad 1 \leq i \leq q + 1, 1 \leq k \leq N. \quad (2)$$

Sachant que les N vecteurs x_k sont quantifiés, l'espace E se trouve divisé en q^N domaines $\Delta_{i_1 \dots i_N}$

$$\Delta_{i_1 \dots i_N} = \bigcap_{j=1}^N \{X \mid s_{i_j} \leq C_j(x_j) < s_{i_j+1}\}. \quad (3)$$

Au niveau du processeur central l'observation est réduite à un vecteur de dimension N dont les composantes sont choisies parmi un alphabet de taille q . Il existe donc q^N vecteurs différents représentant chacun un domaine $\Delta_{i_1 \dots i_N}$. Le but ultime du système étant la détection, le processeur central associé à chaque message reçu une décision H_0 ou H_1 . Le processeur central est donc entièrement défini par un ensemble de q^N valeurs $v_{i_1 \dots i_N}$ valant 0 ou 1. Ces valeurs représentent la décision adoptée par le système si l'observation tombe dans $\Delta_{i_1 \dots i_N}$. L'ensemble des $v_{i_1 \dots i_N}$ constitue la loi de fusion du processeur central. Le système complet est donc entièrement déterminé par l'ensemble de paramètres θ défini par

$$\theta = [C_1(\cdot), \dots, C_N(\cdot); s_{11}, \dots, s_{1q+1}, \dots, s_{N1}, \dots, s_{Nq+1}; \{v_{i_1 \dots i_N}\}]. \quad (4)$$

Les probabilités de fausse alarme et de détection nécessaires pour optimiser le système ont donc pour expression

$$Q_0 = \sum_{i_1=1}^{q+1} \dots \sum_{i_N=1}^{q+1} v_{i_1 \dots i_N} \alpha_{i_1 \dots i_N}$$

$$Q_d = \sum_{i_1=1}^{q+1} \dots \sum_{i_N=1}^{q+1} v_{i_1 \dots i_N} \beta_{i_1 \dots i_N}$$

$$\alpha_{i_1 \dots i_N} = \int_{\Delta_{i_1 \dots i_N}} p_0(X) dX$$

$$\beta_{i_1 \dots i_N} = \int_{\Delta_{i_1 \dots i_N}} p_1(X) dX. \quad (5)$$

La formulation la plus générale du problème s'exprime de la façon suivante: déterminer θ tel que Q_d soit maximum et $Q_0 < \alpha$. On peut montrer que la fonction optimale C_k est le rapport de vraisemblance [7] [8]. La structure optimale des quantificateurs locaux est donc la quantification par rapport de vraisemblance pour laquelle on quantifie non pas l'observation mais son rapport de vraisemblance calculé au préalable. Malgré l'hypothèse I.I.D. le choix qui consiste à prendre les mêmes quantificateurs, c'est-à-dire les mêmes seuils sur tous les capteurs, n'est pas en général optimum. Ceci nous oblige à optimiser non pas les $q + 1$ seuils d'un quantificateur, mais les seuils d'une famille de quantificateurs. En termes de dimension l'optimisation des seuils de quantification à elle seule représente $N(q + 1)$ degrés de liberté.

La structure des quantificateurs connue, le système est entièrement défini par un ensemble de paramètres scalaires P défini par

$$P = [s_{1_1}, \dots, s_{1_{q+1}}, \dots, s_{N_1}, \dots, s_{N_{q+1}}; \{v_{i_1}, \dots, i_N\}]. \quad (6)$$

Le problème consiste maintenant à déterminer le vecteur P optimum. Le calcul des paramètres optimaux au sens de Neyman-Pearson conduit à un système d'équations du type

$$s_{ij} = g_i(s_{1_1}, \dots, s_{ij-1}, s_{ij+1}, \dots, s_{N_{q+1}}; \{v_{i_1}, \dots, i_N\}) \quad (7)$$

montrant que les optimisations des traitements locaux et du traitement central sont liées et doivent donc être menées simultanément. Or de par la nature du traitement central, le nombre de lois de fusion possibles est fini et vaut 2^{qN} . De plus aucune de ces lois n'est optimale pour toutes les valeurs de Q_0 . Ainsi pour optimiser globalement le système, on est conduit à procéder de la façon suivante. Il faut tout d'abord se fixer une loi de fusion, puis résoudre le système de $N(q + 1)$ équations déterminant les seuils de quantification, ce qui est lourd en raison de la taille du système. Ensuite à l'aide des équations (5) on calcule les performances du système pour la loi que l'on s'est fixée. Enfin il faut renouveler l'opération pour les 2^{qN} lois de fusion et choisir le système le plus performant. En théorie il existe 2^{qN} lois de fusion possibles, mais on peut montrer à l'aide, entre autre d'un argument de monotonie, que certaines lois ne peuvent être optimales [5] [8]. Supposons qu'une loi de fusion $\{v_{i_1}, \dots, i_N\}$ associe l'hypothèse H_1 à une entrée particulière, cette loi sera monotone si pour toute entrée déduite de la première en remplaçant une ou plusieurs de ses composantes par un niveau de quantification supérieur, la décision prise est également H_1 . Par exemple, dans le cas de la détection décentralisée avec quatre capteurs ($N = 4$ et $q = 2$), différents auteurs font passer le nombre de lois candidates à l'optimalité de 65536 à 114 [5] [8]. Il existe peut-être d'autres arguments permettant de diminuer ce nombre. Dans le cas de la détection décentralisée à deux capteurs, il est connu que seules deux règles de fusion sont optimales. Malgré la mise à l'écart d'une partie d'entre elles, le nombre de lois qui restent candidates à l'optimalité demeure important. Ceci nous conduit à abandonner l'idée d'optimalité absolue et à considérer des systèmes sous-optimaux que nous optimisons à l'aide d'une approche heuristique.

Pour simplifier le problème et en raison de l'hypothèse I.I.D., nous prenons des quantificateurs identiques sur tous les capteurs. Dans ce cas le nombre théorique de lois de fusion est de 2^M avec $M = C_N^{N+q-1} = \binom{N}{N+q-1}$

3. APPROCHE HEURISTIQUE

3.1 Interprétation graphique

L'optimisation au sens de N.P. peut se résoudre par la méthode classique du multiplicateur de Lagrange, ce qui revient à trouver le domaine D_0 tel que la quantité R soit minimum.

$$R = \int_{D_0} \{p_1(X) - \lambda p_0(X)\} dX = \int_{D_0} K_\alpha(X) dX \quad (8)$$

où λ est le multiplicateur choisi de façon à vérifier la contrainte $Q_0 < \alpha$. La quantité R sera minimale (R^*), si on choisit le domaine D_0^* où $K_\alpha(X)$ est négatif, c'est-à-dire si la frontière optimale entre les deux domaines de décision a pour équation,

$$(S^*) \quad K_\alpha(X) = 0 \Leftrightarrow L(X) = \lambda. \quad (9)$$

Par définition toute modification du domaine de décision $D_0 = D_0^* + \Delta D_0$ par rapport à l'optimum D_0^* entraîne une augmentation ΔR de R^* s'écrivant

$$\Delta R = \int_{\Delta D_0} K_\alpha(X) dX \quad (10)$$

La division optimale de E se fait selon (S^*) c'est-à-dire le long d'une surface en général courbe. Dans le cas de la détection répartie avec quantification, les quantifications locales divisent E en domaines rectangulaires Δ_{i_1, \dots, i_N} . Le processeur central associe chacun de ces domaines à l'une ou l'autre des régions de décision. La frontière n'est plus courbe mais polygonale. Comme nous le voyons sur la Figure 1, trois cas peuvent se présenter.

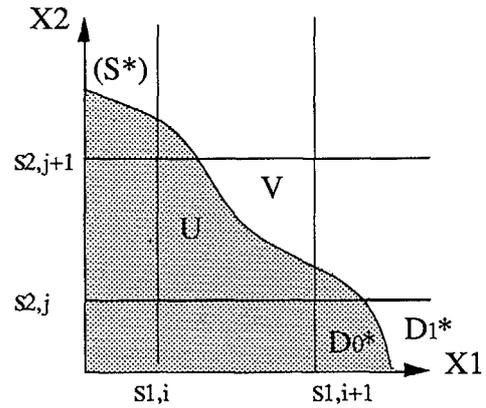


Figure 1. Surface de séparation et quantification

- (1) Le domaine Δ_{i_1, \dots, i_N} est tout entier inclus dans D_0^* . Le processeur central doit décider H_0 afin de ne rien changer par rapport au système optimum.
- (2) Le domaine Δ_{i_1, \dots, i_N} est tout entier inclus dans D_1^* . Le processeur central doit décider H_1 afin de ne rien changer par rapport au système optimum.
- (3) Le domaine Δ_{i_1, \dots, i_N} est traversé par la frontière (S^*) , ce que nous appelons domaine ambigu. Dans ce cas quelle que soit la décision prise, la division de E sera modifiée par rapport à l'optimum, ce qui entrainera des pertes de performance.

À de tels domaines nous associons deux intégrales.

$$R_u = \int_U K_\alpha(X) dX ; \quad R_v = \int_V K_\alpha(X) dX. \quad (11)$$

Où U et V représentent les parties de Δ_{i_1, \dots, i_N} appartenant à D_0^* ou D_1^* . Si le processeur central décide H_0 lorsque l'observation tombe dans un domaine ambigu, l'équation d'optimisation devient

$$R = R^* + R_v. \quad (12)$$

Si le processeur central décide H_1 , alors nous avons.

$$R = R^* + R_u. \quad (13)$$

3.2 Règles de dessin des systèmes sous-optimaux

L'interprétation graphique nous permet de conclure que seuls les domaines ambigus sont responsables des pertes de performance des systèmes quantifiés. Nous déduisons de cette

remarque deux enseignements que nous allons utiliser lors de la conception de systèmes sous-optimaux.

Règle A : Dans la mesure du possible, les seuils doivent être choisis de façon à définir des domaines non ambigus.

Règle B : Le choix de la décision associée à un domaine ambigu se fait de façon à minimiser la variation par rapport à l'optimum. Ce qui revient à mettre en oeuvre le test suivant,

$$R_u \underset{H_0}{\overset{H_1}{\geq}} R_v \tag{14}$$

3.3 Application à un cas particulier

Nous considérons un problème de détection répartie à deux capteurs. Le bruit est gaussien centré de variance unité. Nous voulons détecter un signal constant d'amplitude A . La forme optimale de la détection décentralisée ($q = 2, N = 2$) est connue et on a montré que dans le cas gaussien le choix de seuils égaux sur les deux détecteurs $s_1 = s_2 = s$ est optimum [7]. Nous allons améliorer les performances de ce système en utilisant un niveau supplémentaire ($q = 3$) pour coder les informations locales, c'est-à-dire en transformant les détections locales en quantifications. Nous voulons tout d'abord montrer qu'il est possible de concevoir des systèmes plus performants que la détection décentralisée en utilisant des règles simples pour déterminer le seuil supplémentaire (noté t) et la loi de fusion. Enfin nous allons comparer les performances du système quantifié simplifié aux performances du système quantifié optimum. Les règles que nous avons définies se réfèrent au système à centralisation complète de l'information qui est le système optimum absolu. Dans le cas gaussien l'équation de la frontière optimale absolue s'écrit,

$$(S^*) \quad x_1 + x_2 = \xi. \tag{15}$$

La Figure 2 représente les domaines de décision que nous obtenons en appliquant la règle A pour le choix de s et de t et la règle B pour le choix de la loi de fusion. La connaissance des domaines de décision du système optimum absolu D_0^* et D_1^* nous a permis de choisir t de façon à déterminer au sein de l'un des domaines ambigus du système à détection décentralisée un domaine rectangulaire entièrement inclus dans D_0^* . Nous avons ainsi réduit "l'ambiguïté" due aux détections locales. La Figure 3 représente les performances de quatre structures. Le système optimum absolu (optimum), le système quantifié optimisé (3 levels*), le système quantifié obtenu par l'approche heuristique (3 levels), et le système à détection décentralisée optimisé (and).

Deux conclusions s'imposent, les structures à quantification locale s'avèrent plus performantes que la détection décentralisée, en particulier le système obtenu par l'approche heuristique permet de combler en partie l'écart de performance par rapport à l'optimum absolu au prix d'une augmentation de la complexité des calculs réduite.

Références

[1] J.N. Tsitsiklis and M.Athans, "On the complexity of decentralized decision making and detection problems." *IEEE Trans. on Comm.*, vol. AC-30, no. 5, pp. 440-446, 1985.
 [2] R.R. Tenney and N.R. Sandell, "Detection with distributed sensors." *IEEE Trans. Aerospace & Electronic Systems*, vol. AES-17, no. 4, pp. 501-510, 1981.

[3] Z. Chair and P.K. Varshney, "Optimal data fusion in multiple sensor detection systems." *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-22, no.1, pp. 98-101, 1986.
 [4] R. Srinivasan, "Distributed radar detection theory." *IEE Proc., part F*, vol. 133, no. 1, pp. 55-60, 1986.
 [5] A.R. Reibman and L.W. Nolte, "Optimal detection and performance of distributed sensor system." *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-23, no. 1, pp. 24-30, 1987.
 [6] M-P. Boyer and B. Picinbono, "Quantization and distributed detection" *Proc IEEE ICASSP Glasgow 1989*.
 [7] M.-P.Boyer, Internal report L.S.S.
 [8] R. Viswanathan, A. Ansari and S.C.A. Thomopoulos, "Optimal partitioning of observations in distributed detection.", *Proceedings of the International Symposium on Information Theory*, June 1988, pp. 197.

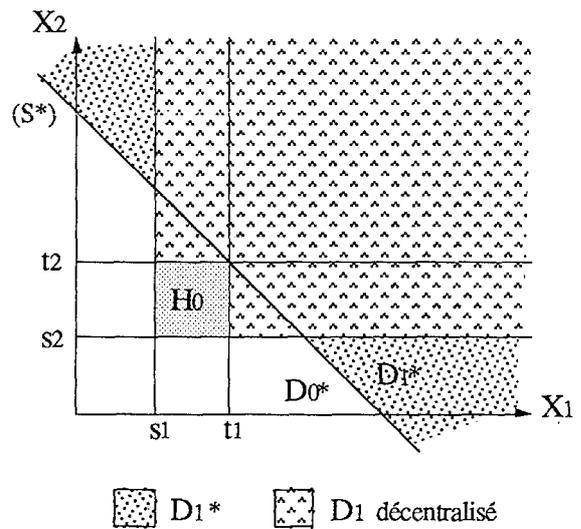


Figure 2. Domaines de décision avec quantification

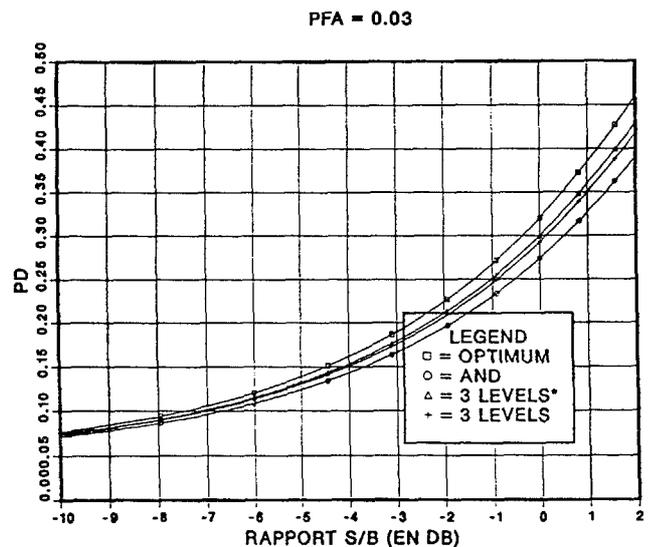


Figure 3. Performances comparées des systèmes