

NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

NOUVELLE APPROCHE DE LA DETECTION PAR SEUIL

B. PICINBONO

P. DUVAUT

LABORATOIRE DES SIGNAUX ET SYSTEMES - E.S.E. - Plateau du Moulon, 91190 - GIF-SUR-YVETTE.

RESUME

Il est bien connu que le rapport de vraisemblance est une statistique suffisante dans le problème de la détection de deux hypothèses simples (signal déterministe dans un bruit de loi de probabilité connue). De plus toute fonction monotone du rapport de vraisemblance est aussi une statistique suffisante, fournissant en particulier les mêmes courbes COR. On peut donc parler de la classe des récepteurs équivalents déterminée par toutes les fonctions monotones possibles. Dans cette classe on peut cependant tenter d'introduire une hiérarchie fondée sur un autre critère. On montre en particulier que si le critère choisi est celui du contraste, les récepteurs utilisant une fonction affine du rapport de vraisemblance sont les éléments optimaux.

L'équivalence des récepteurs de la classe précédemment introduite est fondée sur l'existence d'un seuil au sens mathématique du terme. On montre alors que si on prend en considération une imprécision sur la valeur du seuil les courbes COR des récepteurs de cette classe sont très différentes. Ceci est d'ailleurs illustré sur quelques exemples montrant qu'en présence d'imprécision sur le seuil les récepteurs à contraste maximum sont les meilleurs. La généralité de ce résultat est un problème ouvert qu'il serait intéressant de résoudre.

SUMMARY

It is wellknown that the likelihood ratio is a sufficient statistic in the detection problem between two simple hypotheses (deterministic signal in a noise with known probability distribution). Moreover any monotonic function of the likelihood ratio is also a sufficient statistic, giving in particular the same R O C curves. Then it is possible to introduce the class of equivalent receivers obtained by all the possible monotonic functions. Nevertheless it is possible to introduce in this class a hierarchy with another criterion. In particular it is shown that if the criterion is the contrast criterion, the optimal element is obtained for a linear function of the likelihood ratio.

If, instead of a mathematical perfect threshold, we use a threshold with a finite precision it is shown that the previous receivers are no longer equivalent. After introducing the decision rule we calculate the R O C curves for some examples of monotonic functions which show also that the best receivers have also the contrast maximum. It would be interesting to check if this property is general.



1. INTRODUCTION.

Il est bien connu que le récepteur optimal pour la détection de deux hypothèses simples est un *récepteur à seuil*. D'une manière plus précise ce récepteur doit calculer le rapport de vraisemblance, qui est une statistique suffisante dans ce problème, et le comparer à un seuil déterminé, par exemple, par la probabilité de fausse alarme. On en déduit immédiatement que tout récepteur utilisant une fonction monotone du rapport de vraisemblance est aussi optimal, une telle opération ne modifiant pas les performances décrites par les courbes COR (Caractéristiques Opérationnelles de Réception). On peut donc introduire la classe des récepteurs optimaux formée de tous les récepteurs équivalents qui calculent une fonction monotone du rapport de vraisemblance et la comparent à un seuil.

On peut alors légitimement se demander s'il n'est pas possible d'introduire une certaine hiérarchie dans cette classe. On peut par exemple introduire un autre critère et chercher l'élément optimal dans cette classe. C'est ce que nous ferons dans un premier temps en introduisant la notion de contraste et en cherchant le récepteur qui est à la fois optimal au sens de la détection et qui de plus donne le contraste maximum. Mais on peut aussi remarquer que la classe des récepteurs équivalents est fondée sur l'idée que le seuil est un seuil mathématique défini avec une *précision infinie*. Or cette précision infinie n'existe pas dans la pratique et en particulier la numérisation des dispositifs introduit toujours une précision finie. Si l'on admet ce fait il apparaît alors clair que les récepteurs théoriquement équivalents peuvent alors avoir des comportements fort différents. Nous mettrons ceci en évidence sur quelques exemples précis où apparaît le filtre adapté, récepteur optimal pour la détection d'un signal déterministe dans un bruit gaussien. En dernière partie nous poserons les jalons d'une théorie plus générale de la détection prenant explicitement en compte le fait que tous les seuils ont une précision finie.

2. DETECTION PAR SEUIL ET CONTRASTE.

2.1 - Définition et Optimisation du Contraste.

Soit \underline{x} l'observation à partir de laquelle une décision doit être prise. On se place dans le cas de deux hypothèses simples et on admet donc que les densités de probabilité de \underline{x} sous H_0 et H_1 , $P_0(\underline{x})$ et $P_1(\underline{x})$, sont connues. Tout système qui au vecteur \underline{x} associe le scalaire $y = S(\underline{x})$ calcule une *statistique* de l'observation. Nous allons définir le contraste associé

à toute statistique, puis nous déterminerons la statistique optimale. Par définition nous appellerons *contraste associé à S* la quantité

$$C(S) \triangleq \frac{[E_1(S) - E_0(S)]^2}{E_0(S^2) - [E_0(S)]^2} \quad (2-1)$$

où E_0 et E_1 représentent les espérances mathématiques sous H_0 et sous H_1 . Cette définition reprend en fait le *critère de déflexion* usuellement utilisé en théorie de la détection [1], [2], [3]. Il est facile de voir que

$$C(S) = C(\lambda S + \mu) \quad (2-2)$$

ce qui signifie que le contraste est invariant par homothétie et translation (ou dans toute opération linéaire affine).

Le contraste $C(S)$ peut se mettre sous la forme

$$C(S) = \frac{[\int S(\underline{x}) [P_1(\underline{x}) - P_0(\underline{x})] \cdot d\underline{x}]^2}{\int [S(\underline{x}) - E_0(S)]^2 P_0(\underline{x}) \cdot d\underline{x}} \quad (2-3)$$

qui en raison de l'inégalité de Schwarz est maximum pour

$$P_0(\underline{x}) [S(\underline{x}) - E_0(S)]^2 = k S(\underline{x}) [P_1(\underline{x}) - P_0(\underline{x})]. \quad (2-4)$$

Cette égalité est satisfaite si

$$\begin{aligned} S(\underline{x}) &= k [P_1(\underline{x}) - P_0(\underline{x})] / P_0(\underline{x}) \\ &= k [L(\underline{x}) - 1] \end{aligned} \quad (2-5)$$

où $L(\underline{x})$ est le rapport de vraisemblance. Comme le contraste est invariant par homothétie et translation, la statistique qui maximise le contraste est $S(\underline{x}) = L(\underline{x})$. En conclusion le récepteur dont la sortie a le contraste maximum est celui qui calcule le rapport de vraisemblance, c'est à dire le récepteur optimal dans la détection à seuil. La valeur du contraste maximum est

$$C(L) = 1 - 1 \quad (2-6)$$

$$1 \triangleq E_1[L(\underline{x})] = E_0[L^2(\underline{x})]. \quad (2-7)$$

On peut noter que, comme $E_0[L(\underline{x})] = 1$, $C(L)$ est la variance de $L(\underline{x})$ sous H_0 .

On peut maintenant revenir à la question posée dans l'introduction. Si l'on considère la classe des récepteurs optimaux obtenus par une fonction monotone du rapport de vraisemblance la sous-classe des récepteurs rendant maximum le contraste est définie par

$$S(\underline{x}) = \lambda L(\underline{x}) + \mu \quad (2-8)$$

qui est la classe d'équivalence des récepteurs équivalents à $S(\underline{x}) = L(\underline{x})$. Donc toutes les fonctions monotones de $L(\underline{x})$ non affines à $L(\underline{x})$ diminuent le contraste à la sortie.

2.2 - Exemple dans le cas Gaussien.

Supposons que sous H_0 et sous H_1 l'observation soit du type $N(\underline{0}, R)$ et $N(\underline{s}, R)$. Le rapport de vraisemblance s'écrit alors

$$L(\underline{x}) = c \exp T(\underline{x}) \tag{2-9}$$

où $T(\underline{x})$ est le filtre adapté [4]

$$T(\underline{x}) = \underline{s}^T R^{-1} \underline{x}. \tag{2-10}$$

On obtient aisément les contrastes liés à L et à T qui valent

$$C(L) = \exp d^2 - 1 \tag{2-11}$$

$$C(T) = d^2 \tag{2-12}$$

où d^2 est le rapport (S/B) à la sortie du filtre adapté soit $\underline{s}^T R^{-1} \underline{s}$. On voit que pour $d^2 < 1$ les contrastes sont pratiquement égaux, mais ce n'est plus vrai pour $d > 1$. Or même pour les signaux faibles le rapport d^2 peut être grand à la sortie du filtre adapté.

La notion de contraste est une notion intuitive qui essaye d'opposer la situation H_0 et la situation H_1 afin de mieux les discriminer. Pour un récepteur à seuil mathématique on peut fortement modifier le contraste sans affecter les performances. Mais on sent intuitivement que ceci n'a plus de raison d'être vrai si le seuil n'est infiniment fin, et l'on va donc maintenant envisager ce point.

3. DETECTION AVEC IMPRECISION SUR LE SEUIL.

Nous voulons mettre en évidence que si l'on prend en compte une imprécision sur le seuil deux récepteurs déduits l'un de l'autre en faisant subir à la grandeur décisionnelle une transformation strictement monotone ne sont plus équivalents.

Commençons par rappeler que si le seuil est défini sans erreur les récepteurs sont équivalents, ce qui signifie que le réseau de courbes COR est le même. Pour le voir considérons une statistique de l'observation $S(\underline{x})$ et soit $F(s)$ sa fonction de répartition

$$F(s) = \Pr (S < s) \tag{3-1}$$

définie sous H_0 comme sous H_1 . Soit $R = g(S)$ une autre statistique déduite de S au moyen d'une fonction strictement monotone $g(\cdot)$. On a alors évidemment

$$\begin{aligned} H(r) &= \Pr [R < r] = \Pr [S < g^{-1}(r)] \\ &= F [g^{-1}(r)]. \end{aligned} \tag{3-2}$$

Dans un récepteur à seuil il y a alarme dès que le seuil est dépassé. Les probabilités d'alarme pour les statistiques S et R avec les seuils σ et ρ sont donc

$$P_S(\sigma) = 1 - F[\sigma] \tag{3-3}$$

$$P_R(\rho) = 1 - F[g^{-1}(\rho)] \tag{3-4}$$

On voit donc que si $\rho = g(\sigma)$, on a *quelque soit* F

$$P_R[g(\sigma)] = P_S(\sigma). \tag{3-5}$$

En appliquant ceci à F_0 et F_1 on obtient respectivement les probabilités de fausse alarme (α) et de détection (β) ce qui prouve que les courbes COR donnant β en fonction de α sont confondues.

Reprenons le même problème en présence d'une imprécision sur le seuil. Il faut tout d'abord définir l'événement alarme. Pour ceci si la grandeur décisionnelle est $D(\underline{x})$ et le seuil δ avec une erreur ϵ on décide H_1 si $D(\underline{x}) \geq \delta + \epsilon$, H_0 si $D(\underline{x}) < \delta - \epsilon$ et on choisit au hasard, par exemple avec des probabilités 1/2 entre H_0 et H_1 si $\delta - \epsilon \leq D(\underline{x}) < \delta + \epsilon$.

Ceci étant il faut bien distinguer le cas de l'imprécision absolue ou de l'imprécision relative.

3.1- Seuils à imprécision absolue.

Reprenant les mêmes notations que ci-dessus les probabilités d'alarme correspondant aux statistiques S et R sont

$$P_S(\sigma, \epsilon) = 1 - \frac{1}{2}\{F(\sigma + \epsilon) + F(\sigma - \epsilon)\} \tag{3-6}$$

$$P_R(\rho, \epsilon) = 1 - \frac{1}{2}\{F[g^{-1}(\rho + \epsilon)] + F[g^{-1}(\rho - \epsilon)]\}. \tag{3-7}$$

On retrouve évidemment (3-3) et (3-4) si $\epsilon = 0$.

3.2- Seuils à imprécision relative.

Il suffit de remplacer dans les relations précédentes $\rho \pm \epsilon$ par $\rho(1 \pm \eta)$ et $\sigma \pm \epsilon$ par $\sigma(1 \pm \eta)$ ce qui donne

$$P_S(\sigma, \eta) = 1 - \frac{1}{2}\{F[\sigma(1 + \eta)] + F[\sigma(1 - \eta)]\} \tag{3-8}$$

$$P_R(\rho, \eta) = 1 - \frac{1}{2}\{F[g^{-1}[\rho(1 + \eta)]] + F[g^{-1}[\rho(1 - \eta)]]\}. \tag{3-9}$$

L'avantage de cette seconde vision est que les probabilités d'alarme, donc les courbes COR, sont invariantes par homotétie ce qui est une propriété importante à conserver dans la comparaison avec la théorie du contraste. En effet si la fonction monotone est $g(x) = kx$ on voit immédiatement sur (3-8) et (3-9) que *quel que soit* $F(s)$,

$$P_S(\sigma, \eta) = P_R(k\sigma, \eta). \tag{3-10}$$

En appliquant ceci sous H_0 et sous H_1 on obtient les mêmes relations pour les probabilités de fausse alarme et de détection, ce qui assure que le réseau de courbe COR est le même pour les statistiques S et R , malgré l'imprécision du seuil. Ceci n'est en général pas vrai pour les seuils à imprécision absolue.

Par contre l'introduction d'une erreur sur le seuil ne permet plus d'écrire l'égalité (3-5). Ceci



indique que les courbes COR n'ont aucune raison de rester invariantes dans la transformation monotone $g(\underline{x})$. Nous allons illustrer ce fait sur des exemples particuliers.

4. EXEMPLES.

Nous allons reprendre plus en détail l'exemple abordé à la fin du paragraphe 2 et qui correspond au problème classique de la détection d'un signal déterministe dans un bruit centré gaussien de covariance R. Le rapport de vraisemblance est d'après (2-9) une fonction monotone de la sortie du filtre adapté donnée par (2-10). La statistique $S(\underline{x})$ dont la fonction de répartition $F(x)$ figure dans toutes les équations précédentes est donc $d^{-1}T(\underline{x})$. Cette quantité est une variable aléatoire $N(0,1)$ sous H_0 et $N(d,1)$ sous H_1 et la fonction de répartition $F(x)$ se déduit directement de la fonction erreur.

$$\operatorname{erf}(x) = (2\pi)^{-1/2} \int_x^{\infty} \exp(-t^2/2) \cdot dt \quad (4-1)$$

On obtient ainsi

$$F_0(x) = 1 - \operatorname{erf}(x) \quad (4-2)$$

$$F_1(x) = 1 - \operatorname{erf}(x-d) \quad (4-3)$$

Pour tracer le réseau de courbes COR avec imprécision sur le seuil nous choisissons trois fonctions $g(\cdot)$ déterminant une fonction monotone du filtre adapté $T(\underline{x})$. La première est évidemment l'identité, $g_0(x) = x$. La seconde est la fonction exponentielle, $g_1(\cdot) = \exp(\cdot)$, ce qui donne le rapport de vraisemblance défini par (2-9). La troisième est précisément choisie pour diminuer le contraste et ainsi mieux mettre en évidence l'influence d'une erreur sur le seuil. Pour avoir une forme analytique simple nous avons choisi la fonction

$$g_2(x) = \operatorname{Artg}\left(\frac{x}{a}\right) \quad (4-4)$$

qui est bien une fonction monotone de x mais possède un effet de saturation marqué pour les grandes valeurs de (x/a) . On peut par contre dire que pour les petites valeurs de x/a cette fonction est pratiquement une homothétie.

Toutes ces remarques sont confirmées par le graphe de la figure 1 représentant l'évolution des courbes COR suivant la fonction monotone choisie $g_1(\cdot)$, $g_2(\cdot)$ ou $g_0(\cdot)$, pour une probabilité de fausse alarme donnée ($\alpha=5 \times 10^{-2}$) et une incertitude sur le seuil η fixée à 2×10^{-1} . La valeur élevée de η permet de visualiser clairement l'influence d'un seuil fortement bruité sur les récepteurs étudiés.

D'une part, on perçoit très nettement la dégradation progressive du récepteur $\operatorname{Artg}\frac{x}{a}$ lorsque le

paramètre a diminue, il s'agit de l'effet de saturation évoqué ci-dessus.

D'autre part, dès que a prend des valeurs plus élevées il y a convergence vers le récepteur "identité", $\operatorname{Artg}\frac{x}{a}$ tendant alors vers une homothétie. La quantité a n'a d'ailleurs pas besoin de croître considérablement pour que le récepteur $\operatorname{Artg}\frac{x}{a}$ se situe très près de sa limite supérieure (en effet pour $a=2$ et $d^2=9$ $\beta_S - \frac{\beta \operatorname{Artg} \frac{s/a}{\beta S}}{\beta S} < 6\%$)

Au regard des courbes, il est également possible d'écrire que :

$$\begin{aligned} & \forall d^2 \\ & \alpha = 5 \times 10^{-2} \quad \beta \exp S \geq \beta_S \quad (4-5) \\ & \eta = 0,2 \end{aligned}$$

ce qui montre la supériorité du récepteur $\exp(\cdot)$ devant le récepteur $\operatorname{id}(\cdot)$. Dans ces conditions, on peut résumer les conclusions tirées de l'étude graphique par la relation :

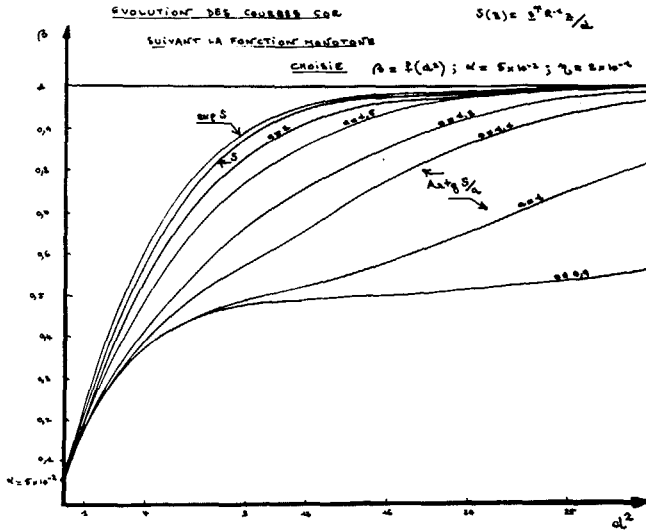
$$\begin{aligned} & \forall d^2 \\ & \alpha = 5 \times 10^{-2} \quad \beta \exp S \geq \beta_S \geq \beta \operatorname{Artg} \frac{S}{a} \geq \beta \operatorname{Artg} S \geq \beta \operatorname{Artg} \frac{S}{a''} \\ & \eta = 0,2 \quad \text{où } a'' < 1 < a' \quad (4-6) \end{aligned}$$

Cette inégalité est intéressante dans la mesure où elle permet de distinguer des récepteurs, qui, dans l'hypothèse d'un seuil mathématique auraient des courbes "COR" identiques, en les classant suivant un critère de détection. Cette hiérarchie introduite au sein de récepteurs appartenant à une même classe n'est pas sans rappeler, celle établie par le critère du contraste (2-2) entre $C_{\exp(\cdot)}^{g_1(\cdot)}$ et $C_{\operatorname{id}(\cdot)}^{g_0(\cdot)}$.

Ainsi, pour étayer l'idée selon laquelle, en présence d'un seuil "épais" à un meilleur contraste, correspondent de meilleures performances en détection il faudrait calculer le contraste associé à $g_2(\cdot) = \operatorname{Artg}\left(\frac{\cdot}{a}\right)$, $C_{g_2(\cdot)}$ et le trouver inférieur à $C_{g_0(\cdot)}$ quelle que soit la valeur du rapport signal sur bruit. C'est alors que la relation

$$C_{g_2(\cdot)} < C_{g_0(\cdot)} < C_{g_1(\cdot)} \quad \forall d^2 \quad (4-7)$$

et l'inégalité (4-6) cautionneraient l'existence d'un lien étroit entre le critère du contraste et les courbes COR.



5. CONCLUSION.

Les résultats précédents montrent que la prise en considération d'une imprécision sur le seuil remet en cause des résultats bien établis sur la détection par seuil. Les exemples étudiés montrent une parenté entre la théorie du contraste et celle de détection par seuil avec imprécision. La question de savoir si ce résultat est général reste encore ouverte. Elle est d'ailleurs liée à celle de la recherche du récepteur optimal en présence d'erreur sur le seuil actuellement en cours d'étude.

REFERENCES.

- [1] A. BLANC-LAPIERRE et B. PICINBONO, Propriétés Statistiques du Bruit de Fond, Masson, 1961.
- [2] P.Y. ARQUES, Décisions en traitement du signal, Masson, 1982, p. 79.
- [3] W.A. GARDNER, A Unifying view of second-order measures of quality for signal classification, IEEE Trans. on Comm. COM-28, p. 807-816, 1980.
- [4] B. PICINBONO, Eléments de théorie du signal, Dunod, 1980, p. 75.

