Bayes-Optimal Detectors Design Using Relevant Second-Order Criteria

Cédric Richard, Régis Lengellé, and Fahed Abdallah

***Abstract*—Statistical detection theories lead to the fundamental result that the optimum test consists in comparing any strictly monotone function of the likelihood ratio with a threshold value. In many applications, implementing such a test may be impossible. Therefore, we are often led to consider a simpler procedure for designing detectors. In particular, we can use alternative design criteria such as second-order measures of quality. In this paper, a necessary and sufficient condition is given for such criteria to guarantee the best solution in the sense of classical detection theories. This result is illustrated by discussing the relevance of well-known criteria.**

***Index Terms*—Bayes-optimal detector, deflection, detector design, Fisher criterion, generalized signal-to-noise ratio, second-order criterion.**

I. Introduction

The purpose of detection is to determine to which of two classes ω_0 or ω_1 a given observation belongs. Let the vector $x \in \mathbb{R}^n$ be an observation, and let $y \in \{0, 1\}$ be its class. Detection is accomplished with a discriminant function $h(x)$: $\mathbb{R}^n \longrightarrow \{0, 1\}$, which errs on x if $h(x) \neq y$. In this letter, the observations are assumed to be generated in accordance with the two known probability densities $p(X|Y = 0)$ and $p(X|Y = 1)$, denoted by the standard notations $p_0(X)$ and $p_1(X)$, respectively. Classical statistical detection theories (see e.g., [1]) such as Bayes, Neyman–Pearson and minimax lead to the following result of major importance that the optimum test consists in comparing the *likelihood ratio* (LR) defined as $L(X) \triangleq p_1(X)/p_0(X)$ with a threshold value ν in order to make a decision:

$$h(X) = \begin{cases} 1, & \text{if } L(X) \triangleq \dfrac{p_1(X)}{p_0(X)} > \nu \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Note, any strictly monotone function of $L(X)$ leads to an equivalent decision rule in the sense that the *receiver operating characteristic* (ROC) is the same.

In many practical applications, implementing the LR test may be an intractable problem due to excessive time and storage requirements, or may be impossible because of incomplete specification of the conditional probability densities $p_0(X)$ and $p_1(X)$. Therefore, we are often led to consider a simpler procedure for designing detectors defined as:

$$h(X) = \begin{cases} 1, & \text{if } S(X) > \nu \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

In particular, we can use alternative design criteria and performance measures. Perhaps the most popular alternatives are second-order measures of quality (see e.g., [2]–[4]). These criteria are defined in terms of second-order moments of $S(X)$, namely

$$m_i \triangleq \mathrm{E}\{S|\omega_i\}, \qquad \sigma_i^2 \triangleq \mathrm{Var}\{S|\omega_i\} \quad (3)$$

with $i \in \{0, 1\}$. There have been many contributions to justify individual second-order criteria, including convergence to optimal detectors of classical detection theories (see [3] and references therein). In [5, pp. 141–143], the objective of the author is to unify these results stating that the use of any function $\Psi(m_0, m_1, \sigma_0^2, \sigma_1^2)$ as a design criterion guarantees the best solution in the Bayes sense[1] for general nonlinear detector design. The main contribution of our letter is to show that in fact $\Psi(m_0, m_1, \sigma_0^2, \sigma_1^2)$ must satisfy a nontrivial condition to provide Bayes-optimal detectors, and thus, to be considered as a *relevant second-order criterion* for detector design.

The rest of this letter is organized as follows. In the next section, we give statistics $S(X)$ which optimize second-order criteria $\Psi(m_0, m_1, \sigma_0^2, \sigma_1^2)$ as a function of the LR. We then derive a necessary and sufficient condition for the existence of relevant second-order criteria. As shown in Section III, this condition can easily be used to test the relevance of well-known design criteria. Section IV contains some concluding remarks.

II. Characterization of Relevant Second-Order Criteria

Let Ψ be any function of $m_i \triangleq \int S(X)p_i(X)\,dX$ and $\sigma_i^2 \triangleq \int (S(X) - m_i)^2 p_i(X)\,dX$, with $i \in \{0, 1\}$. We first have to characterize statistics $S(X)$ which optimize Ψ. Operating on Ψ with a variational calculus, we obtain

$$\delta\Psi = \frac{\partial\Psi}{\partial m_0}\,\delta m_0 + \frac{\partial\Psi}{\partial m_1}\,\delta m_1 + \frac{\partial\Psi}{\partial \sigma_0^2}\,\delta\sigma_0^2 + \frac{\partial\Psi}{\partial \sigma_1^2}\,\delta\sigma_1^2. \quad (4)$$

Manuscript received April 20, 2001; revised November 15, 2001. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. R. L. de Queiroz.

The authors are with the Laboratoire de Modélization et Sûreté des Systèmes (LM2S), Université de Technologie de Troyes (UTT), 10010 Troyes cedex, France (e-mail: cedric.richard@univ-troyes.fr).

Publisher Item Identifier S 1070-9908(02)02415-X.

[1]Throughout this letter, this refers to a detector in which the test statistic $S(X)$ is a strictly monotone function of the LR. Note that such detectors are also optimal in the Neyman–Pearson and minimax sense when the threshold is properly chosen.

Since $\delta m_i = \int \delta S(X) p_i(X)\, dX$ and $\delta\sigma_i^2 = 2\int (S(X) - m_i)\delta S(X) p_i(X)\, dX$ with $i \in \{0, 1\}$, we obtain

$$\delta\Psi = \int \Bigg[\frac{\partial\Psi}{\partial m_0}\, p_0(X) + \frac{\partial\Psi}{\partial m_1}\, p_1(X) + 2(S(X) - m_0)\,\frac{\partial\Psi}{\partial\sigma_0^2}\, p_0(X) + 2(S(X) - m_1)\,\frac{\partial\Psi}{\partial\sigma_1^2}\, p_1(X)\Bigg]\, \delta S(X)\, dX. \quad (5)$$

To make $\delta\Psi = 0$ regardless of $\delta S(X)$, the $[\cdot]$ term given above must be equal to 0. Using $L(X) = p_1(X)/p_0(X)$, we finally get the expression of the statistic $S(X)$ optimizing Ψ as a function of the LR

$$S(X) = -\frac{1}{2}\,\frac{\dfrac{\partial\Psi}{\partial m_0} + \dfrac{\partial\Psi}{\partial m_1}\, L(X)}{\dfrac{\partial\Psi}{\partial\sigma_0^2} + \dfrac{\partial\Psi}{\partial\sigma_1^2}\, L(X)} + \frac{m_0\,\dfrac{\partial\Psi}{\partial\sigma_0^2} + m_1\,\dfrac{\partial\Psi}{\partial\sigma_1^2}\, L(X)}{\dfrac{\partial\Psi}{\partial\sigma_0^2} + \dfrac{\partial\Psi}{\partial\sigma_1^2}\, L(X)}. \quad (6)$$

The above statistic is completely equivalent to the LR iff it is a strictly monotone function of $L(X)$. Evaluating the first order derivative of $S(X)$ with respect to $L(X)$, we obtain

$$\frac{dS}{dL}(X) = \frac{(m_1 - m_0)\,\dfrac{\partial\Psi}{\partial\sigma_0^2}\,\dfrac{\partial\Psi}{\partial\sigma_1^2} + \dfrac{1}{2}\left(\dfrac{\partial\Psi}{\partial\sigma_1^2}\,\dfrac{\partial\Psi}{\partial m_0} - \dfrac{\partial\Psi}{\partial\sigma_0^2}\,\dfrac{\partial\Psi}{\partial m_1}\right)}{\left(\dfrac{\partial\Psi}{\partial\sigma_0^2} + \dfrac{\partial\Psi}{\partial\sigma_1^2}\, L(X)\right)^2}. \quad (7)$$

Since the denominator of (7) is strictly positive for all X, we thus note that $S(X)$ defined by (6) is a strictly monotone function of $L(X)$ iff the numerator of (7) is not equal to 0. This result leads directly to the following proposition.

Proposition 1: $\Psi(m_0, m_1, \sigma_0^2, \sigma_1^2)$ is a relevant second-order criterion, i.e., it guarantees the best solution in the Bayes sense for detector design, iff

$$(m_1 - m_0)\,\frac{\partial\Psi}{\partial\sigma_0^2}\,\frac{\partial\Psi}{\partial\sigma_1^2} + \frac{1}{2}\left(\frac{\partial\Psi}{\partial\sigma_1^2}\,\frac{\partial\Psi}{\partial m_0} - \frac{\partial\Psi}{\partial\sigma_0^2}\,\frac{\partial\Psi}{\partial m_1}\right) \neq 0. \quad (8)$$

Since it is very difficult, if not impossible, to find the solutions of (8), our discussion will be limited to the relevance of well-known design criteria in the next section.

III. Examples of Second-Order Criteria

We are concerned with a type of second-order criteria that is frequently used in practice. These criteria are of the form $\Psi(u, v)$, where $u \triangleq (m_1 - m_0)^2$ is a measure of *between-class scatter* and $v \triangleq \alpha\sigma_1^2 + (1 - \alpha)\sigma_0^2$ with $\alpha \in [0, 1]$ is the *within-class scatter* (see e.g., [2]–[4]). To discuss the relevance of such criteria using (8), we need to compute the derivatives of Ψ with respect to m_i and σ_i^2. Thus

$$\frac{\partial\Psi}{\partial m_0} = -2\sqrt{u}\,\frac{\partial\Psi}{\partial u} \qquad \frac{\partial\Psi}{\partial m_1} = 2\sqrt{u}\,\frac{\partial\Psi}{\partial u} \quad (9)$$

$$\frac{\partial\Psi}{\partial\sigma_0^2} = (1 - \alpha)\,\frac{\partial\Psi}{\partial v} \qquad \frac{\partial\Psi}{\partial\sigma_1^2} = \alpha\,\frac{\partial\Psi}{\partial v}. \quad (10)$$

Substituting (9) and (10) into (8), we note that $\Psi(u, v)$ is a relevant criterion iff

$$\sqrt{u}\,\frac{\partial\Psi}{\partial v}\left[\alpha(1 - \alpha)\,\frac{\partial\Psi}{\partial v} - \frac{\partial\Psi}{\partial u}\right] \neq 0. \quad (11)$$

The existence of the optimal statistic (6) indicates that $\partial\Psi/\partial\sigma_0^2$ and $\partial\Psi/\partial\sigma_1^2$ cannot be both equal to 0. With (10), we then obtain $\partial\Psi/\partial v \neq 0$ so that (11) becomes

$$\alpha(1 - \alpha)\,\frac{\partial\Psi}{\partial v} \neq \frac{\partial\Psi}{\partial u}. \quad (12)$$

To solve (12) in the case $\alpha \in \{0, 1\}$ presents no difficulty. We directly obtain that Ψ must depend on $(m_1 - m_0)^2$. For any $\alpha \in]0, 1[$, (12) can be solved simply by posing $u = u' + v'$ and $v = \alpha(1 - \alpha)(u' - v')$; this leads to $\Psi(u, v) \neq \psi(u + (v/\alpha(1 - \alpha)))$, where ψ is any real-valued function. Therefore, criteria depending on a weighted sum of the between-class scatter and the within-class scatter are nonrelevant, e.g., $\Psi(m_0, m_1, \sigma_0^2, \sigma_1^2) = \alpha(1-\alpha)(m_1-m_0)^2 + \alpha\sigma_1^2 + (1-\alpha)\sigma_0^2$. In the present case, we verify that $S(X)$ given by (6) does not depend on $L(X)$, which conforms to Condition (8). Intuitively, note that such criteria are not satisfactory: they combine antagonistic quantities, the between-class scatter and the within-class scatter, via a sum.

The *generalized signal-to-noise ratio* (GSNR) postulates the performance measure of $S(X)$ to be the quantity (see e.g., [3])

$$\Psi(u, v) = \frac{u}{v}. \quad (13)$$

The value of Ψ is a measure of clustering for the two competing classes ω_0 and ω_1: it tends to be large when $u \triangleq (m_1 - m_0)^2$ increases and $v \triangleq \alpha\sigma_1^2 + (1-\alpha)\sigma_0^2$ decreases. This design criterion includes as particular cases various second-order measures used in practice such as Fisher criterion ($\alpha = 1/2$) [5], deflection ($\alpha = 1$) and complementary deflection ($\alpha = 0$) [3]. Observing that (13) satisfies (12), we immediately conclude that the GSNR is a relevant second-order criterion.

IV. Concluding Remarks

The theoretical results reported in this paper are concerned with the relevance of second-order criteria used for general nonlinear detector design. We have given a necessary and sufficient condition for these widely used measures of quality to guarantee the best detector in the sense of classical decision theories. We have illustrated this result by discussing the relevance of well-know design criteria. From this study, it can be concluded that there is a broad class of relevant second-order criteria. However, it should be noted that obtaining the optimum detector remains an unsolved problem.

References

[1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968, vol. I.
[2] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
[3] W. A. Gardner, "A unifying view of second-order measures of quality for signal classification," *IEEE Trans. Commun.*, vol. COM–28, pp. 807–816, June 1980.
[4] B. Picinbono and P. Duvaut, "Detection and contrast," in *Stochastic Processes in Underwater Acoustics*, C. Baker, Ed. New York: Springer-Verlag, 1986.
[5] K. Fukunaga, *Statistical Pattern Recognition*. New York: Academic, 1990.