LUND UNIVERSITY

A simultaneous maximum likelihood estimator based on a generalized matched filter

Gustavsson, Jan-Olof; Börjesson, Per Ola

Published in:
[Host publication title missing]

DOI:
10.1109/ICASSP.1994.389775

1994

Link to publication

Citation for published version (APA):
Gustavsson, J.-O., & Börjesson, P. O. (1994). A simultaneous maximum likelihood estimator based on a generalized matched filter. In *[Host publication title missing]* (pp. 481-484)
https://doi.org/10.1109/ICASSP.1994.389775

Total number of authors:
2

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 24. Apr. 2024

A SIMULTANEOUS MAXIMUM LIKELIHOOD ESTIMATOR BASED ON A GENERALIZED MATCHED FILTER

Jan-Olof Gustavsson[1] [2] *Per Ola Börjesson*[2]

[1]University College of Karlskrona/Ronneby, Department of Signal Processing, S-372 25 Ronneby, Sweden
[2]Luleå University of Technology, Division of Signal Processing, S-971 87 Luleå, Sweden

ABSTRACT

This paper discusses parameter estimation and detection in laplace distributed noise. The received signal is modeled as $r(\cdot) = As(\cdot,\theta) + n(\cdot)$, where A is an unknown amplitude, θ is the parameter vector to be estimated and $n(\cdot)$ is independent laplace distributed noise. The simultaneous maximum likelihood estimator of (A,θ) is derived. The derived estimator is based on a combination of a weighted median filter[1] and a generalized form of the ordinary matched filter[2].

Examples of performance for four different detectors are given for a case of binary detection, when the amplitude A or the signal shape $s(\cdot,\theta)$ are varied. Simulations indicate that the performance of detectors based on the generalized matched filter is not particularly dependent on either the estimate of the amplitude A or the signal shape.

INTRODUCTION

There are many problems where the task is to estimate a parameter in an observed signal[3]. Examples of such problems are those of estimating the numerical value of a parameter (e.g. arrival time estimation) or choosing one hypothesis out of several hypotheses (e.g. detection or classification of a signal). In many cases these problems can be dealt with by using a signal model of the form $r(\cdot) = As(\cdot,\theta) + v(\cdot)$, where A is a known amplitude and θ is the parameter vector to be estimated from the observed signal $r(\cdot)$.

There are a number of studies of how to estimate θ in an optimal way according to given optimization criteria assuming the signal model given above. Examples of such criteria are the maximum likelihood (ML), maximum a posteriori (MAP) and the minimum mean square error (MMSE) criterion[3]. It is usually assumed that the noise, $v(\cdot)$, is gaussian distributed[3, 4, 5]. There are at least two reasons why this assumption is so common. One reason is that the noise in many applications is at least approximately gaussian distributed according to the central limit theorem[6], and another reason is that the gaussian assumption is analytically tractable. However, situations also exist where the statistical distribution of the noise in the above model is non-gaussian[7, 8, 9]. Johnson and Rao[8] state that "Gaussian processes would seem to be imprecise representations of physical measurements.". During the last few years the problem of parameter estimation in non-gaussian noise has been an active field of research. and many results have been published, e.g. Kassam[10] including references.

When the noise is white gaussian, the optimal estimates of θ according to many different criteria can be derived from the output from ordinary matched filters[3] when $r(\cdot)$ is the input to the filters[1]. This follows from the fact that when the noise is white gaussian, the output from the ordinary matched filters is a sufficient statistic for calculating the a posteriori distribution for θ, given the observed signal $r(\cdot)$ and a known a priori distribution for θ. When the noise is independent non-gaussian it can be shown that the output from a modified form of the ordinary matched filter is a sufficient statistic for calculating the a posteriori distribution for θ. Gustavsson and Börjesson have derived and discussed this filter in [2]. The modified form of the ordinary matched filter can be seen as an ordinary matched filter where the multipliers have been replaced by non-linearities. The non-linearity is known from several works dealing with the problem of detection in non-gaussian noise[10].

The estimation problem generally becomes much more difficult if the amplitude A in the model $r(\cdot) = As(\cdot,\theta)+v(\cdot)$ is both continuous valued and unknown. A special case when the ML estimate of θ can still be calculated exactly is when the noise is gaussian distributed, since the estimator then becomes independent of the amplitude A. Another case where this can be done is discussed in this paper. The simultaneous ML estimator of A and θ is derived assuming independent laplace distributed noise. The ML estimator is based on a combination of a matched median filter[1], which gives an estimate of the amplitude A, and a modified form of the ordinary matched filter[2]. The derivation can easily be modified by using Bayes theorem to give a simultaneous ML estimator of A and a MAP or MMSE estimator of θ. In a case of binary detection discussed in this paper, the performance of the estimator is calculated by means of simulations, and the results are compared to the performance of three other estimators.

THE MAXIMUM LIKELIHOOD ESTIMATOR

Consider the problem of how to estimate a parameter vector θ from an observed signal $r(\cdot)$, where $r(\cdot)$ is modeled as

$$r(k) = As(k,\theta) + v(k), \quad k \in I. \tag{1}$$

In the model, $s(\cdot,\theta)$ is a signal dependent on the unknown parameter vector θ, A is an unknown amplitude factor, $A > 0, A \in \mathbf{R}$, $v(\cdot)$ is independent laplace distributed noise with probability density function $f_v(\cdot)$, k is a discrete 'time' variable and I the observation interval. The probability density function $f_v(\cdot)$ is given by

$$f_v(x) = \frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}|x|/\sigma}, \quad x \in \mathbf{R}, \tag{2}$$

where σ^2 is the variance of the noise. For all θ, the signal $s(k,\theta) = 0$ for all k outside the interval $I_\theta : [m(\theta), m(\theta) + n(\theta)]$ and outside the observation interval I, i.e. the signal $s(\cdot,\theta)$ is of finite length, $n(\theta)+1$, and completley contained in the observation interval I. The parameter θ, which can

[1]The output from filters matched with $s(\cdot,\theta)$ for all possible values of θ are necessary.

0-7803-1775-0/94 $3.00 © 1994 IEEE

only assume a countable number of values, is either a random or non-random vector independent of the noise $v(\cdot)$. $\hat{\theta}$ and $\hat{A}$ denote estimates of the unknown parameters θ and A.

A structure of a filter for estimating θ according to the ML, MAP or the MMSE criterion is proposed in [2] and shown in figure 1. This filter is built up of two parts. The first part, the pre-processor, processes the observed signal $r(\cdot)$ and gives as output $L(\theta|r(\cdot))$, for all possible values of θ. Ideally, the pre-processor should be chosen so that these outputs are a (minimal) sufficient statistic for calculating the a posteriori distribution of θ, provided that the a priori distribution of θ is known. The second part, the estimator, uses this information about θ to make an estimate of θ according to a desired criterion.

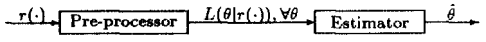

Figure 1. *A general block-structure of a two-step process for estimating θ. The output from the pre-processor is $L(\theta|r(\cdot))$, for all possible values of θ.*

The ML estimate of the parameters (A, θ) is achieved by maximizing the likelihood function for the parameters (A, θ). The likelihood function is

$$\begin{aligned} L_1(A,\theta) &= \prod_{k\in I} f_v(r(k) - As(k,\theta)) \\ &= \exp\left\{\sum_{k\in I} ln\{f_v(r(k) - As(k,\theta))\}\right\}. \end{aligned} \tag{3}$$

Maximizing $L_1(A, \theta)$ is equivalent to maximizing the exponent in (3). The exponent can be rewritten as[2]

$$\begin{aligned} \sum_{k\in I} ln\{f_v(r(k) - As(k,\theta))\} &= \sum_{k\in I} ln\{f_v(r(k))\} \\ +\sum_{k\in I_\theta} \{ln\{f_v(r(k) - As(k,\theta))\} - ln\{f_v(r(k))\}\}. \end{aligned} \tag{4}$$

Denote the second sum on the right hand side of (4) by $L_2(A, \theta)$ and note that the first sum is independent of both A and θ. Consequently, maximizing the likelihood function $L_1(A, \theta)$ is equivalent to maximizing $L_2(A, \theta)$. By using $s(k, \theta) = 0, k \notin [m(\theta), m(\theta) + n(\theta)]$ and by defining

$$b_k(A,\theta) = As(m(\theta) + n(\theta) - k, \theta) \tag{5}$$

and

$$l(b,x) = \ln\left\{\frac{f_v(x-b)}{f_v(x)}\right\} = \frac{\sqrt{2}}{\sigma}(|x| - |x-b|), \tag{6}$$

$L_2(A, \theta)$ can be written

$$\begin{aligned} L_2(A,\theta) &= \sum_{k=m(\theta)}^{m(\theta)+n(\theta)} \ln\left\{\frac{f_v(r(k) - As(k,\theta))}{f_v(r(k))}\right\} \\ &= \sum_{k=0}^{n(\theta)} l(b_k(A,\theta), r(m(\theta) + n(\theta) - k)). \end{aligned} \tag{7}$$

A filter structure for calculating $L_2(A, \theta)$ for a given amplitude A is discussed in [2]. Now for each θ, maximize $L_2(A, \theta)$ with respect to A. Since the noise is independent laplace distributed, the maximum likelihood estimate $\hat{A}_\theta$ of A given θ is the weighted median of the observations, $r(\cdot)$, normalized by the signal $s(\cdot, \theta)$, where the weights are given by the absolute values of the signal $s(\cdot, \theta)$[1]. That is, $\hat{A}_\theta$ is given by

$$\hat{A}_\theta = \arg\left\{\min_{A\in\mathbf{R}} \sum_{k:s(k,\theta)\neq 0} |s(k,\theta)| \left|A - \frac{r(k)}{s(k,\theta)}\right|\right\}. \tag{8}$$

Since the parameter vector θ can assume only a finite number of values, $\hat{A}_\theta$ can be calculated for all possible values of θ. Denote the value $L_2(\hat{A}_\theta, \theta)$ with $L(\theta)$. A block structure illustrating the calculation procedure is shown in figure 2. In the figure the weighted median filter is based on equation (8) and the generalized matched filter is based on equation (7) with A replaced by $\hat{A}_\theta$. This structure has to be repeated for all possible values of θ to give $L(\theta), \forall\theta$. The simultaneous maximum likelihood estimate of (A, θ) is then $(\hat{A}_{\hat{\theta}}, \hat{\theta})$, where

$$\hat{\theta} = \arg\left\{\max_\theta L(\theta)\right\}, \tag{9}$$

c.f. the generalized likelihood ratio test in [3]. With the aid of Bayes theorem the estimation procedure can easily be modified to give a simultaneous ML estimate of A and a maximum a posteriori or minimum mean square error estimate of θ given that θ is a stocastic vector with known a priori distribution.

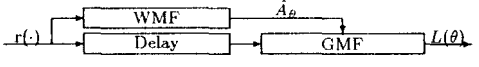

Figure 2. *A block structure for calculating $L(\theta)$ when the signal amplitude is unknown and the noise is independent laplace distributed. WMF: Weighted median filter, GMF: Generalized matched filter*

EXAMPLES OF PERFORMANCE

Consider the classical problem of deciding if a signal is present or not, to exemplify the performance of the simultaneous maximum likelihood estimator. The noise is independent laplace distributed with zero mean and unit variance. In this example, the problem consists of choosing between the hypotheses

$H_0 : \theta = 0$
$H_1 : \theta = 1$

where

$s(k, 0) = 0, \forall k$
$s(k, 1) = 1, k \in [1, 5]$ och $s(k, 1) = 0, k \notin [1, 5]$

For this case the performance of four different detectors, D1 - D4, has been calculated by means of simulations. In the different detectors, the pre-processor in figure 1 is based on:

D1 the generalized matched filter assuming that the amplitude A is known[2], that is, $\hat{A}_\theta = A$.

D2 the proposed simultaneous ML estimator of A and θ, which is the generalized matched filter with the amplitude A estimated using the weighted median[1], that is, $\hat{A}_\theta$ is given by (8). Henceforth, this detector will be referred to as either D2 or the simultaneous maximum likelihood detector.

D3 the generalized matched filter with $\hat{A}_\theta = 1$.

D4 the ordinary matched filter[3].

Note that the detectors D1 – D3 have in common that they all are based on the generalized matched filter. D1 is assumed to know the amplitude A, while D2 tries to estimate the amplitude A from the received signal $r(\cdot)$ using (8). Since $A > 0$ in (1), an estimate $\hat{A}_\theta < 0$ of A is changed to $\hat{A}_\theta = 0$.

It follows from (7) that under the hypothesis $H_0 : \theta = 0$, the output from the pre-processor $L(\theta|r(\cdot)) = 0, \forall r(\cdot)$. Consequently, the estimator must base the decision between H_0 and H_1 solely on the output from the pre-processor in figure 2, $L(\theta)$, with $\theta = 1$, that is $L(1)$.

In the four different detectors, the estimator used is based on a Neyman-Pearson test[3]. The estimator compares $L(1)$ to a threshold λ, where λ is chosen so that a given probability of false alarm, p_f (that is, to decide H_1 when H_0 is true), is achieved. The decision rule is:

$$L(1) \begin{cases} \leq \lambda : & \text{decide } H_0. \\ > \lambda : & \text{decide } H_1. \end{cases} \qquad (10)$$

Let p_d denote the probability of detection, that is, to decide H_1 when H_1 is true.

In figures 3 and 4, the simulated results of the probability of detection, p_d, respectively the probability of miss, $p_m = 1 - p_d$, are shown as a function of the amplitude A in (1) for the detectors D1 – D4, when the probability of false alarm, p_f, is 0.1 and 0.01. To facilitate comparison of the different detectors, figures 5 and 6 show the quotient between p_d for D2 – D4 and p_d for the detector D1, when $p_f = 0.01$ and $p_f = 0.1$ respectively. In figures 3 – 6 the following can be observed:

- the probability of detection is not particularly dependent on the estimate $\hat{A}_\theta$ of A.
- for small values of A, the detector based on the ordinary matched filter, D4, has a lower probability of detection than the other detectors used.
- for large values of A, the detector based on the ordinary matched filter, D4, has a higher probability of detection than the other detectors, i.e. D2 and D3, when the true value of the amplitude A is unknown.

In figure 7 the pulse shape of $s(k,1), k \in [1,5]$ varies instead of the amplitude. The pulse shapes are given by

$$s(\cdot,1) = \left[x, \frac{1+x}{2}, 1, \frac{1+x}{2}, x\right] \qquad (11)$$

where x varies from 0.1 to 3, c.f. [1]. All signals are scaled so that the signal energy becomes 5. This signal energy corresponds to $A = 1$ in figures 3 – 6. In figure 7 the following can be observed:

- the probability of detection is not particularly dependent on the signal shape.
- the detector based on the ordinary matched filter, D4, has a lower probability of detection than the other detectors simulated for all signal shapes used.

CONCLUSIONS

The simultaneous maximum likelihood estimator has been derived for estimating a parameter vector θ and an unknown signal amplitude A in independent laplace distributed noise. This estimator consists of a combination of a weighted median filter and a generalized form of the ordinary matched filter. The estimator can easily be modified to become an ML estimator of the amplitude and a MAP or MMSE estimator of the parameter vector.

Simulations have been made for a case of binary detection discussed in this paper, where the performance of four different detectors has been compared. The simulations indicate that, in general, the performance of the simultaneous maximum likelihood detector is better than the performance of other detectors used, when the received signal amplitude is unknown.

The simulations also indicate that the performance of detectors based on the generalized matched filter are not particularly dependent on either the estimate of the amplitude A or the signal shape.

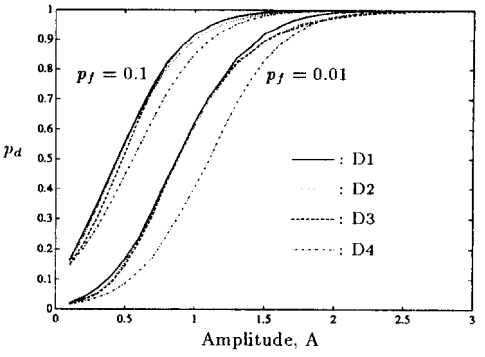

Figure 3. *The probability of detection, p_d, for four different detectors when the probability of false alarm is 0.1 and 0.01 as a function of the signal amplitude. The noise is laplace distributed.*

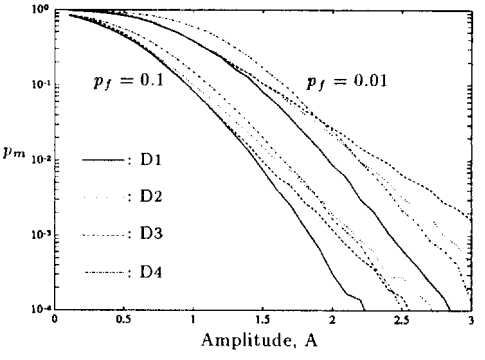

Figure 4. *The probability of miss, $p_m = 1 - p_d$, for four different detectors when the probability of false alarm is 0.1 and 0.01 as a function of the signal amplitude. The noise is laplace distributed.*

REFERENCES

[1] J. Astola and Y. Neuvo. Matched median filtering. *IEEE transactions on communications*, 40(4):722–729, April 1992.

[2] J-O. Gustavsson and P.O. Börjesson. A generalized matched filter. In *Proceedings of the third international symposium on signal processing and its applications*, pages 16–22, Gold Coast, Australia, August 1992.

[3] H.L. van Trees. *Detection, estimation and modulation theory*, volume 1. Wiley, USA, 1968.

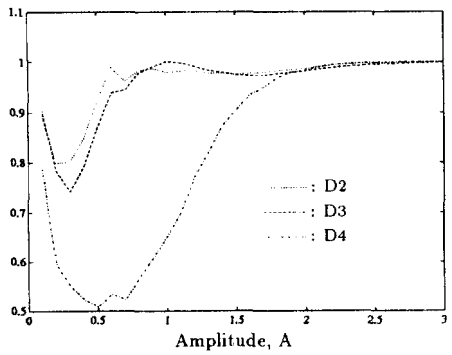

Figure 5. *The quotient between the probability of detection for three different detectors, D2 - D4, and the probability of detection for the detector, D1, when the probability of false alarm is 0.01 as a function of the signal amplitude. The noise is laplace distributed.*

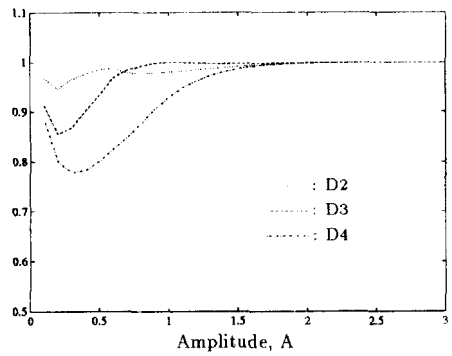

Figure 6. *The quotient between the probability of detection for three different detectors, D2 - D4, and the probability of detection for the detector, D1, when the probability of false alarm is 0.1 as a function of the signal amplitude. The noise is laplace distributed.*

[4] P.O. Börjesson, O. Pahlm, L. Sörnmo, and M-E. Nygårds. Adaptive QRS detection based on maximum a posteriori estimation. *IEEE Transactions on biomedical engineering*, BME-29(5), May 1982.

[5] J.M. Mendel. *Lessons in digital estimation theory.* Prentice-Hall, USA, 1987.

[6] A. Papoulis. *Probability, random variables and stochastic processes.* McGraw-Hill, USA, second edition, 1984.

[7] V.R. Algazi and R.M. Lerner. Binary detection in white non-gaussian noise. Report DS-2138, Lincoln Laboratory, M.I.T., USA, 1964.

[8] D.H. Johnson and P.S. Rao. *On the existence of* gaussian noise. In *Proceedings of ICASSP 91*, Toronto, Canada, May 1991.

[9] F.W. Machell and C.S. Penrod. Probability density functions of ocean acoustic noise processes. In E.J. Wegman and J.G. Smith, editors, *Statistical signal processing.* Marcel Dekker, New-York, USA, 1984.

[10] S.A. Kassam. *Signal detection in non-gaussian noise.* Springer Verlag, USA, 1988.

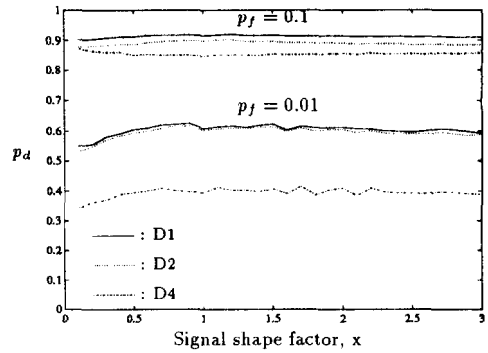

Figure 7. *The probability of detection, p_d, for three different detectors when the probability of false alarm, p_f, is 0.01 and 0.1 as a function of the signal shape. The noise is laplace distributed.*