Signal Processing 89 (2009) 1902–1909

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

A low complexity robust detector in impulsive noise ☆

Puskal P. Pokharel *, Weifeng Liu, Jose C. Principe

Computational NeuroEngineering Laboratory, ECE Department, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:
Received 30 November 2008
Received in revised form
10 March 2009
Accepted 21 March 2009
Available online 1 April 2009

Keywords:
Matched filter
Correntropy
Information theoretic learning
Alpha-stable
Kernel method

ABSTRACT

This paper demonstrates the effectiveness of a nonlinear extension to the matched filter for signal detection in certain kinds of non-Gaussian noise. The decision statistic is based on a new measure of similarity that can be considered as an extension of the correlation statistic used in the matched filter. The optimality of the matched filter is predicated on second order statistics and hence leaves room for improvement, especially when the assumption of Gaussianity is not applicable. The proposed method incorporates higher order moments in the decision statistic and shows an improvement in the receiver operating characteristics (ROC) for non-Gaussian noise, in particular, those that are impulsive distributed. The performance of the proposed method is demonstrated for detection in two types of widely used impulsive noise models, the alpha-stable model and the two-term Gaussian mixture model. Moreover, unlike other kernel based approaches, and those using the characteristic functions directly, this method is still computationally tractable and can easily be implemented in real-time.

1. Introduction

The linear correlation filter or matched filter (MF) has been the basic building block for a wide range of applications requiring detection of known signals [1,2]. The limitations of the matched filter though are already defined by the assumptions under which it is optimal. It is well known that for the detection of a known signal in additive white Gaussian noise (AWGN) the matched filter maximizes the signal to noise ratio (SNR) among all linear filters [3]. Theoretically, it also uses the maximum likelihood statistic for hypothesis testing under the assumptions of linearity and Gaussianity. Optimal detection in non-Gaussian noise (or nonlinear) environments usually requires the use of the characteristic function and is much more complex due to the need of higher order statistics to accurately model the noise [4]. This motivates the recent interest in nonlinear filters (kernel matched filters) [5,6] or non-quadratic cost functions [7], but the computational complexity of such solutions outweighs their usefulness in applications where high processing delay cannot be tolerated such as in radar and mobile communication systems.

Kernel methods [8–10] transform the data points x and y from the input space to a higher dimensional feature space of vectors $\Phi(x)$ and $\Phi(y)$, where the inner products can be computed using a positive definite kernel function, $\kappa(x,y) = \langle \Phi(x), \Phi(y) \rangle$ satisfying Mercer's conditions [11]. This can also be employed to obtain a nonlinear solution to the template matching problem [5]. But the correlation matrix formed in this infinite dimensional feature space is also infinitely large and the resulting formulation is complex using a large set of training data. Alternatively, it can be formulated as a discriminant function in kernel space [6], but still suffers from the need to train the system before hand and store the training data. The matched filter based on quadratic mutual information (MI) is another recently introduced nonlinear filter that maximizes the mutual information between the template and the output of the filter [7]. This method does not

☆ This work was supported by NSF Grants ECS-0300340 and ECS-0601271.

* Corresponding author. Tel.: +1 352 392 2682; fax: +1 352 392 0044.
E-mail addresses: pokharel@cnel.ufl.edu (P.P. Pokharel), weifeng@cnel.ufl.edu (W. Liu), principe@cnel.ufl.edu (J.C. Principe).

0165-1684/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2009.03.027

require an initial training step since it is non-parametric. However, the method requires the estimation of the quadratic mutual information with kernels and the computational load is still $O(N^2)$ at best.

The method in this paper uses a recently introduced positive definite function called correntropy [12], which quantifies higher order moments of the noise distribution with a low computational complexity of $O(N)$. The performance of the proposed correntropy matched filter (CMF) is compared with that of the traditional matched filter (linear correlation filter) and the matched filter using mutual information in a variety of scenarios using an additive channel, with Gaussian and impulsive noise distributions. For α-stable distributed impulsive noise, the locally suboptimal (LSO) detector is also presented for comparison, which is designed for optimal performance exclusively for this type of noise. The comparison is illustrated using the receiver operating characteristics (ROC), which shows the trade off between the probability of detection and the probability of false alarms [13].

2. Background

2.1. Linear matched filter

For the case of detecting one signal s_{1i} over another signal s_{0i} in the presence of AWGN, the maximum likelihood decision statistic is simply the difference between the correlation of the received signal r_i with the two signal templates [14] and can be expressed as

$$L(\mathbf{r}) = \frac{1}{N}\sum_{i=1}^{N} r_i(s_{1i} - s_{0i}). \tag{1}$$

This, in fact, is the output $y_k = r_k * h_k$ of the matched filter at the time instant $k = N+1$, N being the signal length, where $*$ denotes convolution and $h_k = s_{1,N+1-k} - s_{0,N+1-k}$ is the matched filter impulse response. The simplicity, as is apparent, makes the wide applicability of the matched filter principle possible.

2.2. Information theoretic learning and mutual information

Information theoretic learning (ITL) is an adaptive methodology (cost function and learning) of extracting information directly from data in a non-parametric manner [15]. In information theory, the mutual information between two random variables X and Y is traditionally as defined by Shannon [16]. The crucial property of mutual information for our purposes is the fact that the value of mutual information increases as the dependency (even nonlinear) of X on Y increases. On the other hand, if X and Y are independent, MI becomes zero. Essentially, we want a measure of divergence of the variables X and Y from independence. A qualitatively similar measure of independence called the Cauchy–Schwartz mutual information (CS-QMI) can be derived between X and Y using the Cauchy–Schwartz inequality for inner products in vector spaces and is given by [15]

$$I_s(X;Y) = \frac{1}{2}\log\frac{\iint p_{XY}^2(x,y)\,dx\,dy \iint p_X^2(x)p_Y^2(y)\,dx\,dy}{(\iint p_{XY}(x,y)p_X(x)p_Y(y)\,dx\,dy)^2} \tag{2}$$

with data available, $I_s(X;Y)$ can be estimated using Parzen window density estimation and can be used as a statistic for signal detection as in [7]. We shall use CS-QMI as a comparison against our proposed method as well since this is a direct template matching scheme that requires no training and shows improved performance in non-Gaussian and nonlinear situations [7].

2.3. Impulsive noise distributions

Since this paper aims to show the effectiveness of the proposed method in impulsive noise environments, we shall briefly introduce the most commonly used PDF models for such distributions. These distributions are commonly used to model noise observed in low-frequency atmospheric noise, fluorescent lighting systems, combustion engine ignition, radio and underwater acoustic channels, economic stock prices, and biomedical signals [17–19]. There are two main models used in literature that we present next.

2.3.1. Two-term Gaussian mixture model

The two-term Gaussian mixture model, which is an approximation to the more general Middleton Class A noise model [20] has been used to test various algorithms under an impulsive noise environment [6,19,21]. The noise is generated as a mixture of two Gaussian density functions such that the noise distribution $f_N(n) = (1-\varepsilon)N(0,\sigma_1^2) + \varepsilon N(0,\sigma_2^2)$, where ε is the percentage of noise spikes and usually $\sigma_2^2 \gg \sigma_1^2$.

2.3.2. α-Stable distribution

α-Stable distributions are also widely used to model impulsive noise behavior [17,18]. This distribution gradually deviates from Gaussianity as α decreases from 2 to 1 (when the distribution becomes Cauchy). This range is also appropriate because even though the higher moments may not exist, the mean is still defined. Here we shall only consider the symmetric α-stable noise. The characteristic function of such noise is given by

$$\Psi_\alpha(u) = e^{-\sigma^\alpha |u|^\alpha}, \tag{3}$$

where σ represents a scale parameter, similar to a standard deviation. Such a random variable has no moment greater than or equal to α, except for the case $\alpha = 2$ [22].

2.3.3. Locally suboptimal receiver

We shall compare the performance of the proposed CMF detector with the locally suboptimum detector, which gives an impressive performance with minimum complexity [23]. The LSO detector is derived directly from the locally optimum (LO) detector [24], whose test statistic is given by

$$T_{LO}(\mathbf{r}) = \sum_{k=1}^{N} s_k g_{LO}(r_k), \tag{4}$$