Signal Processing 96 (2014) 315–320

Contents lists available at ScienceDirect

Signal Processing

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

Fast communication

Unknown signal detection by one-class detector based on Gaussian copula

Antonio Soriano, Luis Vergara*, Jorge Moragues, Ramón Miralles

Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain

ARTICLE INFO

Article history:
Received 1 March 2013
Received in revised form
18 September 2013
Accepted 22 September 2013
Available online 10 October 2013

Keywords:
Signal detection
One-class detector
Copula
Energy detector

ABSTRACT

One-class detector is an option to deal with the problem of detecting an unknown signal in a background noise, as it is only necessary to know the noise distribution. Thus a Gaussian copula is proposed to capture the dependence among the noise samples, meanwhile the marginals can be estimated using well-known methods. We show that classical energy detectors are particular cases of the proposed one-class detector, when Gaussian noise distribution is assumed, but are inappropriate in other cases. Experiments combining simulated noise and real acoustic events have confirmed the superiority of the proposed detectors when noise is non-Gaussian. An interpretation of the methods in terms of the Edgeworth expansion is also included.

1. Introduction

Signal detection in a random background noise is a classical problem in detection theory. Starting from the popular matched filter [1], which requires perfect knowledge of the signal waveform, different methods exist to deal with the practical problem of partial or null knowledge about the signal. Thus, energy detector [2] has been shown to be optimal if both the noise and the signal are independent zero-mean Gaussian random processes. On the other hand, subspace matched filter is optimal if noise is independent Gaussian and the signal lies in a known subspace. Different extensions of the energy detector [3,4] and the subspace matched filter [2,5] exist for the non-independent and/or non-Gaussian cases.

A different approach comes from recognizing that unknown signal detection is conceptually a one-class classifier problem [6], where the "noise" class may be learned (both in the Gaussian and non-Gaussian scenarios), but there is no possibility to learn the "signal" class. This can be also considered inside the so called novelty detection problem [7]. Both approaches converge under the likelihood ratio framework. It is well-known [1] that the optimum test is given by (hypothesis H_1 indicates the presence of signal and noise, hypothesis H_0 that only noise is present)

$$\Lambda(x) = \frac{f(x/H_1)}{f(x/H_0)} \mathop{\gtrless}_{H_0}^{H_1} \lambda, \tag{1}$$

where $f(\mathbf{x}/H_i)$ is the multidimensional probability density function (PDF) of the observation vector $\mathbf{x} = [x_1 x_2 \ldots x_N]^T$ conditioned to hypothesis H_i, and λ is a threshold selected to fit an acceptable probability of false alarm (Neyman–Pearson criterion) or to minimize a defined cost (Bayes approach). $\Lambda(\mathbf{x})$ is the likelihood ratio (*LR*) whose computation obviously requires knowledge of both $f(x/H_1)$ and $f(\mathbf{x}/H_0)$. However, if the signal is totally unknown, a simpler option is to assume that $f(x/H_1)$ is constant [6] leading to the one-class test:

$$f(\mathbf{x}/H_0)^{-1} \mathop{\gtrless}_{H_0}^{H_1} \lambda \Leftrightarrow -\ln f(\mathbf{x}/H_0) \mathop{\gtrless}_{H_0}^{H_1} \ln \lambda \tag{2}$$

Notice that $f(\mathbf{x}/H_0)^{-1}$ is a measure of the degree of departure of $\mathbf{x}$ from the distribution of the "noise" class.

* Correspondence to. Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain. Tel.: +34 963877308; fax: +34 968777919.
E-mail addresses: ansoto@iteam.upv.es (A. Soriano), lvergara@dcom.upv.es (L. Vergara), jormoes@upvnet.upv.es (J. Moragues), rmiralle@dcom.upv.es (R. Miralles).

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.09.011

In this paper we will focus in the one-class test of Eq. (2). Thus, the basic problem will be the estimation of the multidimensional PDF $f(\mathbf{x}/H_0)$. In case of statistical independence among the components of the noise, $f(\mathbf{x}/H_0)$ will simply be the product of the marginals, so that available parametric and nonparametric unidimensional methods are applicable. However, the most difficult aspect regarding estimation of $f(\mathbf{x}/H_0)$ is capturing the possible statistical dependence of the noise. There exist multidimensional nonparametric methods [8], but parametric extensions are not so obvious, except in the multivariate Gaussian model, where a correlation matrix parameterizes the dependence. A more flexible possibility is based on the use of copulas [9]. Although they have been a matter of research in the financial area since long time ago [10,11], they are recently being applied in signal detection problems [12–14], mainly in the context of fusion of heterogeneous detectors. The copula model assumes that $f(\mathbf{x}/H_0)$ can be factorized in two terms: the marginals and a multidimensional PDF of uniformly distributed variables (copula density). This later term captures the statistical dependence. Copulas allow defining a variety of parametric dependence models. On the other hand copula densities may be combined with both parametric and nonparametric estimation of the marginals.

The main contribution of this paper is to propose a new detector for unknown signal detection having general applicability in non-Gaussian and non-independent noise scenarios. It is based on approaching the problem as a one-class or novelty detection problem so that only the multivariate noise distribution is to be required. This latter is estimated assuming a Gaussian copula model. This particular type of copula has been selected due to its simplicity of implementation and its general applicability. Moreover, it leads to a natural extension of classical methods, based upon the energy computation, which only are optimum in Gaussian scenarios. As far as we know, this is the first time that a one-class copula approach has been proposed for unknown signal detection.

In the next section of this paper we present the Copula-based One-Class Detector. In particular, a Gaussian copula is proposed to capture possible dependences. We show that classical energy detectors are particular cases of the proposed one-class detector for both independent and non-independent noise. Experimental results are presented in Section 3 to illustrate the improved performance of the new proposed detector. We have considered real acoustic events corrupted by simulated noises having different probability densities. An interpretation, in terms of the Edgeworth expansion is given in Section 4 about this superior performance, to reinforce the general interest of the Gaussian copula. Conclusions end the communication.

2. One-class detector based on Gaussian copula

Let us focus in the problem of detecting the presence of an unknown signal vector $\mathbf{s}=[s_1 \ldots s_N]^{\boldsymbol{T}}$ in a noise background vector $\mathbf{w}=[w_1 \ldots w_N]^{\boldsymbol{T}}$. Under H_1, $\mathbf{x}=\mathbf{s}+\mathbf{w}$, and under H_0, $\mathbf{x}=\mathbf{w}$. We are going to use a copula to model the multidimensional noise density $f(\mathbf{w})=f(\mathbf{x}/H_0)$. Let $F(\mathbf{w})$ be the corresponding multidimensional cumulative distribution function (CDF) i.e., $f(\mathbf{w})=\partial^N(F(\mathbf{w}))/\partial w_1 \partial w_2 \ldots \partial w_N$. The Sklar's theorem [15] stays that there exists a unique copula function such that:

$$F(\mathbf{w})=C(F_1(w_1),F_2(w_2),\ldots,F_N(w_N)) \tag{3}$$

where $F_n(w_n)$ is the marginal CDF of random variable w_n, so the random variable $u_n=F_n(w_n)$ is uniformly distributed in the interval [0,1]. Deriving (3) we may express $f(\mathbf{w})$ in the form:

$$\begin{aligned} f(\mathbf{w}) &= \frac{\partial^N}{\partial w_1 \partial w_2 \ldots \partial w_N} C(F_1(w_1),F_2(w_2),\ldots,F_N(w_N)) \\ &= f_1(w_1) f_2(w_2) \ldots f_N(w_N) \times c(F_1(w_1),F_2(w_2),\ldots,F_N(w_N)). \end{aligned} \tag{4}$$

where $c(\mathbf{u}/H_0)=\partial^N(C(\mathbf{u}/H_0))/\partial u_1 \partial u_2 \ldots \partial u_N$ is the copula density. There are a number of possible parametric copulas with its corresponding copula densities, but we will focus on the Gaussian copula [16] due to its simplicity, general applicability and straightforward connection with classical energy detectors. A Gaussian copula assumes that if the uniform random variables $u_n=F_n(w_n)$ are transformed into standard Gaussian variables (ones having zero mean and unit variance), then the multidimensional PDF in the transformed domain is multivariate Gaussian. Let us call v_n to the transformed variable, i.e., $v_n=\Phi^{-1}(u_n)$, where $\Phi(\cdot)$ is the CDF of the standard Gaussian random variable. Then, it is assumed that:

$$f_v(\mathbf{v})=\frac{1}{2\pi^{N/2}|\mathbf{R}_{\boldsymbol{v}}|^{1/2}} \cdot \exp\left(-\frac{\mathbf{v}^T\mathbf{R}_{\boldsymbol{v}}^{-1}\mathbf{v}}{2}\right) \tag{5}$$

where $\mathbf{R}_{\boldsymbol{v}}=E[\mathbf{v}\mathbf{v}^{\mathrm{T}}]$ is the standard correlation matrix:

$$\mathbf{R}_{\boldsymbol{v}}(n,m)=E(v_n v_m)=\begin{cases} 1 & n=m \\ <1 & n \neq m \end{cases}.$$

The multivariate Gaussian model assumed in (5) for $f_v(\mathbf{v})$ leads straightforwardly [16] to a particular copula model for $f(\mathbf{w})$, the so called Gaussian copula:

$$f(\mathbf{w})=\frac{1}{|\mathbf{R}_{\boldsymbol{v}}|^{1/2}} \cdot \exp\left(-\frac{\mathrm{g}(\mathbf{w})^T(\mathbf{R}_{\boldsymbol{v}}^{-1}-\mathbf{I})\mathrm{g}(\mathbf{w})}{2}\right) \prod_{n=1}^{N} f_n(w_n) \tag{6}$$

where $\mathrm{g}(\mathbf{w})=[\mathrm{g}(w_1)\mathrm{g}(w_2)\ldots\mathrm{g}(w_N)]^T$ and $\mathrm{g}(w_n)=\Phi^{-1}(F_n(w_n))$. Thus $\mathrm{g}(\cdot)$ is a nonlinear function that transforms the original noise components in standard Gaussian random variables.

Finally, considering in (2) that $f(\mathbf{x}/H_0)=f(\mathbf{w})$, we may define from (6) the new proposed (Gaussian) Copula-based One-Class Detector (COCD):

$$\mathrm{g}(\mathbf{x})^T(\mathbf{R}_{\boldsymbol{v}}^{-1}-\mathbf{I})\mathrm{g}(\mathbf{x})-2\sum_{n=1}^{N}\ln f_n(x_n/H_0) \gtrless_{H_0}^{H_1} \ln\left(\frac{\lambda^2}{|\mathbf{R}_{\boldsymbol{v}}|}\right) \tag{7}$$

Notice that COCD is a new detector which clearly separates the marginals from the joint statistical properties of the noise. The second term is the only one present in the case of independent components. The first term captures the possible dependences among the transformed noise components and it will be present only if $\mathbf{R}_{\boldsymbol{v}} \neq \mathbf{I}$.

It is straightforward to show that (7) leads to classical energy detectors if the noise is assumed multivariate Gaussian. So let us consider that the components of vector