Radar Clutter Classification Using Expectation-Maximization Method

Sudan Han[1], Pia Addabbo[2], Danilo Orlando[3], Giuseppe Ricci[4]

[1] *Defense Innovation Institute, Beijing, China. (e-mail: xiaoxiaosu0626@163.com)*
[2] *Università degli studi Giustino Fortunato, Benevento, Italy. (e-mail: p.addabbo@unifortunato.eu)*
[3] *Engineering Department of Università degli Studi "Niccolò Cusano", Roma, Italy. (e-mail: danilo.orlando@unicusano.it)*
[4] *Dipartimento di Ingegneria dell'Innovazione, Università del Salento, Lecce, Italy. (email: giuseppe.ricci@unisalento.it)*

Paper ID:1262

1. Motivations

In the context of the design of **adaptive radar detection** architectures, it is usually assumed that a set of secondary data, which are target-free and share the same statistical properties of the interference as the primary data (**homogeneous assumption**) is available.

The homogeneous assumption might be **violated** due to the effect of clutter edges, clutter statistical property variation, clutter discretes, moving outliers and so on, leading to a **severe performance degradation** for the detection schemes devised under the homogeneous assumption.

In order to circumvent this drawback, we jointly exploit the **expectation-maximization (EM)** algorithm and **the latent variable model** to classify the collected data into homogeneous subsets assuming that a given number of clutter boundaries is present.

2. Problem Formulation

The considered radar system is equipped with $N \geq 2$ spatial and/or temporal channels and collects K samples from the operating area, each sample corresponds to a specific range bin. The N-dimensional complex vector corresponding to the kth range bin is denoted by $\mathbf{z}_k$, $k = 1, \ldots, K$.

Suppose that the illuminated scenario can be divided into L regions and the samples of a given region share the same statistical properties whereas those in different regions exhibit different statistical characterization. Then the set of samples can be partitioned into L subsets of statistically homogeneous data, whose lth component is given by

$$\Omega_l = \{\mathbf{z}_{i_{l,1}}, \ldots, \mathbf{z}_{i_{l,K_l}}\}$$

where K_l, $l = 1, \ldots, L$, denotes its cardinality.

In this paper, we assume that the samples in the lth region are statistically independent complex circular Gaussian random vectors with zero mean and covariance matrix $\mathbf{M}_l$ which is assumed unknown, namely,

$$[\mathbf{z}_{i_{l,1}} \cdots \mathbf{z}_{i_{l,K_l}}] \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l, \mathbf{I}),\ l = 1, \ldots, L.$$

The goal is to **estimate the subsets Ω_l along with the associated parameter $\mathbf{M}_l$, $l = 1, \ldots, L$.**

3. Classification Architecture Designs

Let us introduce K **independent and identically distributed discrete random variables**, c_ks say, which take on values in $\{1, \ldots, L\}$ with pmf

$$P(c_k = l) = p_l, \quad k = 1, \ldots, K,$$

and such that **when $c_k = l$, then $\mathbf{z}_k \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{M}_l)$**. As a consequence, the pdf of $\mathbf{z}_k$ can be written as

$$f(\mathbf{z}_k; \boldsymbol{\theta}) = \sum_{l=1}^{L} p_l f(\mathbf{z}_k | c_k = l; \mathbf{M}_l) = E_{c_k}[f(\mathbf{z}_k | c_k; \boldsymbol{\theta}))],$$

where $E_{c_k}[\cdot]$ denotes the statistical expectation with respect to c_k,

$$\boldsymbol{\theta} = \left[\mathbf{p}^T, \boldsymbol{\sigma}^T\right]^T$$

$\mathbf{p} = [p_1 \cdots p_L]^T$, $\boldsymbol{\sigma} = \left[\boldsymbol{\nu}^T(\mathbf{M}_1) \cdots \boldsymbol{\nu}^T(\mathbf{M}_L)\right]^T$, $\boldsymbol{\nu}(\cdot)$ a vector-valued function selecting the generally distinct entries of the matrix argument, and

$$f(\mathbf{z}_k | c_k = l; \mathbf{M}_l) = \frac{1}{\pi^N \det(\mathbf{M}_l)} \exp\{-\mathrm{Tr}\,[\mathbf{M}_l^{-1} \mathbf{z}_k \mathbf{z}_k^\dagger]\}.$$

According to the EM algorithm, assume that the $(h-1)$th estimate of the parameter $\boldsymbol{\theta}$ is available, **the E-step** leads to

$$q_k^{(h-1)}(l) = p(c_k = l | \mathbf{z}_k; \hat{\boldsymbol{\theta}}^{(h-1)}) = \frac{f(\mathbf{z}_k | c_k = l; \widehat{\mathbf{M}}_l^{(h-1)}) \hat{p}_l^{(h-1)}}{\sum_{l'=1}^{L} f(\mathbf{z}_k | c_k = l'; \widehat{\mathbf{M}}_{l'}^{(h-1)}) \hat{p}_{l'}^{(h-1)}},$$

whereas **the M-step** consists in the following maximization problem

$$\hat{\boldsymbol{\theta}}^{(h)} = \arg\max_{\boldsymbol{\theta}} \sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(l) \log \frac{f(\mathbf{z}_k | c_k = l; \mathbf{M}_l) p_l}{q_k^{(h-1)}(l)}$$

Note that the maximization problem with respect to p_l and $\mathbf{M}_l$, $l = 1, \ldots, L$ **can be solved separately**. More precisely, the optimization with respect to $\mathbf{p}$ can be solved exploiting the method of **Lagrange multipliers**. It is not difficult to show that the optimizer is

$$\hat{p}_l^{(h)} = \frac{1}{K} \sum_{k=1}^{K} q_k^{(h-1)}(l).$$

The maximization problem over $\mathbf{M}_1, \ldots, \mathbf{M}_L$ is tantamount to the following optimization problem

$$\hat{\boldsymbol{\sigma}}^{(h)} = \arg\max_{\boldsymbol{\sigma}} \sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(l) \log f(\mathbf{z}_k | c_k = l; \mathbf{M}_l).$$

The above problem is addressed assuming **two different expressions for $\mathbf{M}_l$, $l = 1, \ldots, L$**, namely,

1. $\mathbf{M}_l$ is a positive definite Hermitian matrix;
2. $\mathbf{M}_l = \sigma_{c,l}^2 \mathbf{M}$, where $\sigma_{c,l}^2 > 0$ represents the clutter power of the lth class while $\mathbf{M}$ is the **common clutter structure** shared by all the K range bins.

Proposition 1 Assume that $K \geq N$ and form 1 for $\mathbf{M}_l$, then an approximation to the relative maximum point of

$$g_1(\mathbf{M}_1, \ldots, \mathbf{M}_L) = \sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(l) \log f(\mathbf{z}_k | c_k = l; \mathbf{M}_l)$$

has the following expression

$$\widehat{\mathbf{M}}_l^{(h)} = \frac{\sum_{k=1}^{K} q_k^{(h-1)}(l) \mathbf{z}_k \mathbf{z}_k^\dagger}{\sum_{k=1}^{K} q_k^{(h-1)}(l)}, \quad l = 1, \ldots, L.$$

Proposition 2 Assume that $K \geq N$ and form 2 for $\mathbf{M}_l$, then, given the function

$$g_2(\boldsymbol{\sigma}_c^2, \mathbf{M}) = \sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(l) \log f(\mathbf{z}_k | c_k = l; \sigma_{c,l}^2 \mathbf{M})$$

where $\boldsymbol{\sigma}_c^2 = [\sigma_{c,1}^2 \cdots \sigma_{c,L}^2]^T$, an approximation to the relative maximum point can be achieved by means of the following **cyclic procedure** with respect to the iteration index t, $t = 1, \ldots, t_{\max}$, (with $t_{\max}$ a proper design parameter)

$$(\hat{\sigma}_{c,l}^2)^{(1),(h)} = \frac{\sum_{k=1}^{K} q_k^{(h-1)}(l) \mathbf{z}_k^\dagger (\mathbf{M}^{(t_{\max}),(h-1)})^{-1} \mathbf{z}_k}{N \sum_{k=1}^{K} q_k^{(h-1)}(l)},$$

$$\widehat{\mathbf{M}}^{(t),(h)} = \frac{1}{K} \sum_{k=1}^{K} \sum_{l=1}^{L} q_k^{(h-1)}(l) \frac{\mathbf{z}_k \mathbf{z}_k^\dagger}{(\hat{\sigma}_{c,l}^2)^{(t),(h)}},$$

$t = 1, \ldots, t_{\max}$, and

$$(\hat{\sigma}_{c,l}^2)^{(t),(h)} = \frac{\sum_{k=1}^{K} q_k^{(h-1)}(l) \mathbf{z}_k^\dagger (\mathbf{M}^{(t-1),(h)})^{-1} \mathbf{z}_k}{N \sum_{k=1}^{K} q_k^{(h-1)}(l)},$$

$t = 2, \ldots, t_{\max}$, $l = 1, \ldots, L$.

Once the unknown quantities have been estimated, **data classification can be accomplished by exploiting the following rule**

$$\forall k = 1, \ldots, K : \mathbf{z}_k \sim \mathcal{CN}_N(\mathbf{0}, \widehat{\mathbf{M}}_{\hat{l}_k}),$$

where

$$\hat{l}_k = \arg\max_{l=1,\ldots,L} q_k^{(h_{\max})}(l).$$

4. Illustrative Examples

The considered covariance matrix model is given by

$$\mathbf{M}_l = \sigma_{c,l}^2 \mathbf{M}_c,$$

where $\mathbf{M}_c(i, j) = \rho^{|i-j|}$ with $\rho = 0.9$. In addition, we assume $N = 16$, $K = 96$, $L = 3$, $K_1 = 32$, $K_2 = 32$, $K_3 = 32$, and consider the following **three cases for the clutter power levels**: (1) [20,25,30] dB; (2) [20,30,40] dB; (3) [20,35,50] dB.

The performance of the proposed classification architectures is assessed resorting to standard **Monte Carlo counting techniques** over 1000 independent trials and the **performance metrics** are

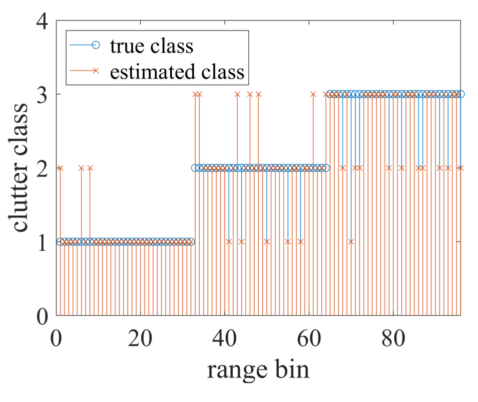
1. the **classification results** of one single Monte Carlo trial;
2. the **root mean square classification error (RMSCE)** with the classification error defined as the number of range bins whose class is not correctly identified.

Table 1: RMSCE for different clutter powers

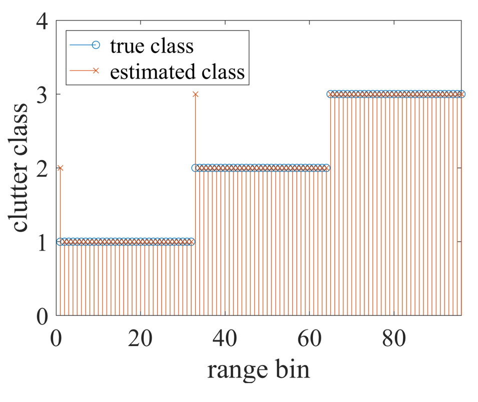
	case (1)	case (2)	case (3)
Proposition 1	19.85	2.87	0.29
Proposition 2	3.10	0.06	0

These simulation results confirm the superiority of the classification scheme based on Proposition 2, indicating that a better classification performance can be achieved **exploiting a priori information about the structure of the clutter covariance matrix** since the adopted clutter covariance matrix model is more compliant with Proposition 2 than Proposition 1.

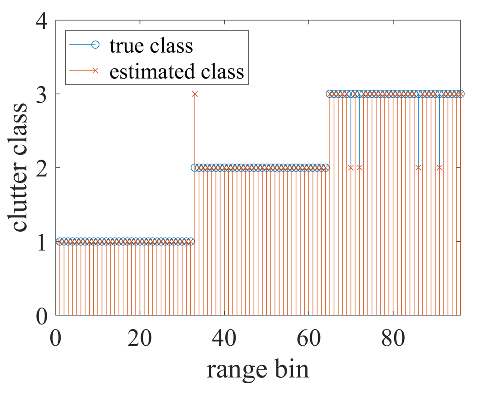
Classification Results



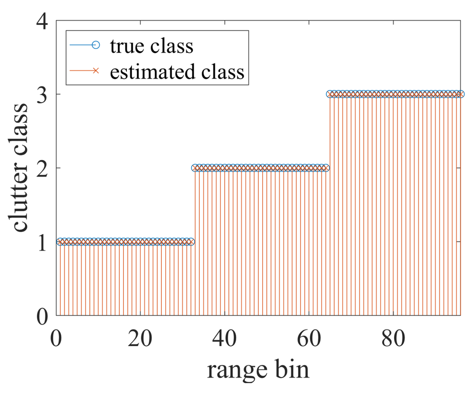
Case 1, Proposition 1



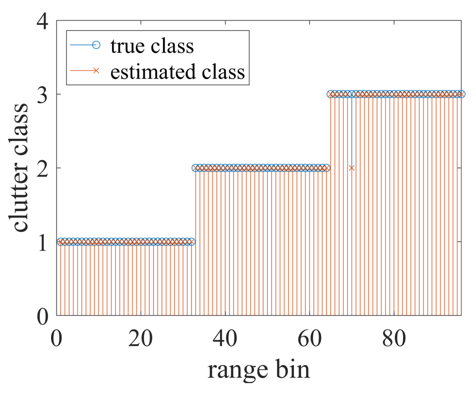
Case 1, Proposition 2



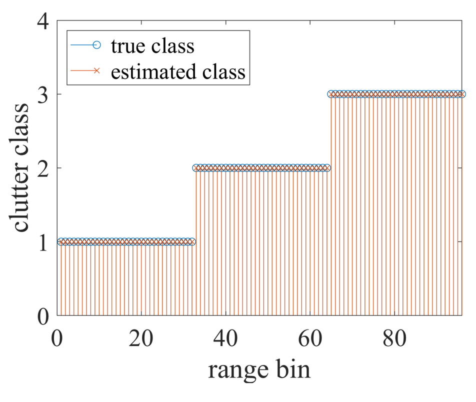
Case 2, Proposition 1



Case 2, Proposition 2



Case 3, Proposition 1



Case 3, Proposition 2