

Channelized model observer for the detection and estimation of signals with unknown amplitude, orientation, and size.

Lu Zhang, Bart Goossens, Christine Cavaro-Ménard, Patrick Le Callet, Di Ge

▶ **To cite this version:**

Lu Zhang, Bart Goossens, Christine Cavaro-Ménard, Patrick Le Callet, Di Ge. Channelized model observer for the detection and estimation of signals with unknown amplitude, orientation, and size.. Journal of the Optical Society of America. A Optics, Image Science, and Vision, 2013, 30 (11), pp.2422-32. 10.1364/JOSAA.30.002422 . inserm-00954609

HAL Id: inserm-00954609

https://inserm.hal.science/inserm-00954609

Submitted on 3 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A channelized model observer for the detection and estimation of signals with unknown amplitude, orientation and size

Lu Zhang,[1,*] Bart Goossens,[2] Christine Cavaro-Ménard,[1] Patrick Le Callet,[3] and Di Ge[1]

[1] *LISA, EA 4094, University of Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France*

[2] *Ghent University, TELIN-IPI-iMinds, St. Pietersnieuwstraat 41, 9000 Ghent, Belgium*[†]

[3] *IRCCyN, UMR-6597, University of Nantes, Rue Christian Pauc, BP50609, 44306 Nantes cedex 3, France*

compiled: September 28, 2013

As a task-based approach for medical image quality assessment, model observers (MOs) have been proposed as surrogates for human observers. While most MOs treat only signal-known-exactly (SKE) tasks, there are few studies on the signal-known-statistically (SKS) MOs, which are more clinically relevant. In this paper, we present a new SKS MO named Channelized Joint detection and estimation Observer (CJO), capable of detecting and estimating signals with unknown amplitude, orientation and size. We evaluate its estimation and detection performances using both synthesized (correlated Gaussian) backgrounds and real clinical (MR) backgrounds. The results suggest that the CJO has good performances in the SKS detection-estimation task.

OCIS codes: 110.2960, 110.2970, 110.3000, 330.1880, 330.5510.

1. Introduction

Medical image quality assessment enables the evaluation and optimization of image acquisition systems, image-processing algorithms and image visualization systems. Task-based image quality assessment approaches [1, 2] characterize the image quality by measuring the performance of an observer, most oftenly a human observer, in completing a specific diagnostic task. Model observers (MOs) [3] have been proposed as surrogates for human observers to avoid the costly and time-consuming procedure involving human observers. The design of a model observer (MO) depends on both the task of interest and the image data.

In the detection task, the presence/absence of a signal (lesion) needs to be decided. A good review on MOs for this task is given in [4]. All these MOs compute a scalar test statistic via a discriminant function of the image data; and they differ by their discriminant functions [5, 6].

MOs can be divided into two categories based on the fact whether *exact* information is available about the signal: Signal-Known-Exactly (SKE) MOs [7, 8] where signal parameters (e.g. intensity amplitude, size, shape, orientation and location) are known exactly to the MOs; and Signal-Known-Statistically (SKS) MOs [9–11] where signal parameters are not known exactly *a priori*, but specified by a probability density function (PDF). SKS MOs are believed to be more clinically relevant, considering human observers' uncertainty and variability in signal parameters in routine diagnosis [12, 13]

In recent publications, we observed a gradual trend from SKE MOs to SKS MOs with the purpose of improving the clinical relevance of MOs. Eckstein *et al.* [14–16] proposed to establish different templates for different combinations of signal parameters and output the optimal template response (in a maximum likelihood sense) as the final test statistic. Zhang *et al.* [17] applied this method on the signal known exactly but variable (SKEV) task (signal parameters are selected from a set of values, then changes from trial to trial, but always known exactly by the observer in each trial), which could be considered as an intermediate step towards the SKS task. Lartizien *et al.* [9] extended this method to 3D (volumetric) images for signal location known exactly but variable. For the SKS task where the unknown signal parameter is its location, Park *et al.* [10] proposed a *scanning* CHO which exhaustively scans the image, then the location that gives the largest test statistic is chosen as the tentative location while that test statistic is the final test statistic. Gifford *et al.* [18–20] proposed a visual-search (VS) model which firstly identifies some candidate blobs guided by features of the test image and then applies the scanning CHO on each candidate blob.

Clarkson [21] proposed an ideal Estimation Receiver Operating Characteristic (EROC) observer, whose EROC curve lies above those of all other observers for the given joint detection estimation task. Whitaker *et al.* [11] proposed a scanning-linear estimator (SLE) for the pure estimation task. Here "scanning" means performing a global-extremum search to maximize a linear

* Corresponding author: lu.ge@insa-rennes.fr

† Bart Goossens is a postdoctoral researcher of Research Foundation Flanders (FWO) Belgium.

metric. The authors proved that the linear metric to be optimized can be derived as a special case of maximum a posteriori (MAP) estimation when the likelihood is Gaussian and a slowly varying covariance approximation is made. This approach is the most similar one to our proposed method. Recently, based on the joint detection and estimation (JDE) theory [22], Goossens *et al.* [23, 24] proposed a series of corresponding SKS CHOs, each treating a signal with varying amplitude, orientation or size separately, which deal with the estimation and the detection task.

We notice that until now SKS MOs suffers from several limitations:

1) The range of variable signal parameters of the existing SKS MOs needs to be widened. For example, Park *et al.* [10] and Gifford *et al.* [18–20] consider the SKS task only varying the signal position.

2) The method of Eckstein *et al.* [14–16] and that of Park *et al.* [10] share the same following limitations: the required number of training images is too large and the calculation amount increases rapidly with the number of variable signal parameters.

3) Although within the theoretical framework of Clarkson *et al.* [21], all signal parameters could be estimated along with the calculation of the test statistic, practically this has not been tackled yet.

4) The estimation performance of the SLE proposed by Whitaker *et al.* [11] has been evaluated when signal amplitude, size and location are variable. However they only focused on the estimation task and did not investigate the detection task performance.

To overcome the above problems of the existing SKS MOs, we propose here a new SKS MO that deals with both the detection and the estimation task, in which the signal amplitude, orientation and scale are variable. It is an extension of the SKS CHOs proposed in [23, 24]. We name it Channelized Joint detecting and estimating Observer (CJO).

The CJO has already been used in [25], by integrating with a Perceptual Difference Model for the joint detection-localization task, where the CJO performs the detection task and the Perceptual Difference Model performs the localization task. In this paper, more details (mathematical derivations, implementation procedure) of the CJO will be introduced, so that readers can implement completely the CJO themselves. In addition, we will give the performance evaluation results of the CJO using both the synthesized correlated Gaussian backgrounds and the real clinical backgrounds - magnetic resonance (MR) images.

The paper is organized as follows. Section 2 describes the mathematical background of the CJO. Section 3 introduces implementation procedures of the CJO, including the Tikhonov regularization technique for estimating the inverse of ill-conditioned covariance matrix. Section 4 presents and discusses its performance evaluation results. Finally, Section 5 gives some conclusions on this work.

2. Mathematical Background

We are interested in a signal detection task in a medical image, where the decision making is the validation of one of two exclusive hypotheses: $\mathcal{H}_0$ (signal absent) and $\mathcal{H}_1$ (signal present). The observed image data $\boldsymbol{g}$ is thus:

$$\mathcal{H}_k:\ \boldsymbol{g} = \boldsymbol{b} + \boldsymbol{n} + k\boldsymbol{x},\ k = 0, 1 \tag{1}$$

where the signal is denoted as $\boldsymbol{x}$, the background as $\boldsymbol{b}$, the noise as $\boldsymbol{n}$. The absence or presence of the signal is controlled by the binary variable k. Let the observed image be represented as column vectors, then $\boldsymbol{g}$ is an $M \times 1$ vector for an image with M pixels. In the literature, the amount of noise is usually assumed to be so small that it does not disturb the statistical properties of the background.

2.A. Background and Signal models

Firstly, we consider the background and signal models for our new SKS MO - CJO.

2.A.1. Background model

All artificial models (e.g., white Gaussian background, correlated Gaussian background, lumpy background [26], and clustered lumpy background [27]) are simplifications of real backgrounds encountered in clinical practice, but they allow different observer strategies to be easily assessed and compared.

In this work, we use both real clinical backgrounds and Correlated Gaussian background (CGB) modeled by Eq. (2):

$$\boldsymbol{b} \sim \mathcal{N}(\mu_{\boldsymbol{b}}, \Sigma_{\boldsymbol{b}}). \tag{2}$$

where $\mu_{\boldsymbol{b}}$ is the mean and $\Sigma_{\boldsymbol{b}}$ is the covariance matrix. This model allows modeling the spatial correlation in clinical images while remains easy to be mathematically manipulated.

2.A.2. Signal model

As said in [28], for the detection task the abnormality of a suitable pathology should not be too conspicuous, and for this pathology the detection task should be the first and the most important step for its diagnosis and treatment. Considering the MRI modality, the multiple sclerosis (MS) was thus selected. We use the MS lesion as the signal in this paper. An example of MR image with multiple sclerosis (MS) lesions on T2 FLAIR sequence (most efficient sequence for the MS lesion detection) is shown in Figure. 1. From Figure. 1 we see clearly that 1) MS lesion appears as hyper-signal (having a higher intensity than its surrounding white matter) on T2 FLAIR sequence; 2) there are multiple lesions in the image and 3) the orientation, scale and amplitude of lesions are unknown. Note that the shape of the MS lesions may depend on the stage of the disease: in earlier stages, the lesion are more circularly symmetric; while in later stages, the asymmetric shape becomes more apparent. Thus we can well mimic the MS lesions on T2 FLAIR sequence using a 2D Gaussian function.

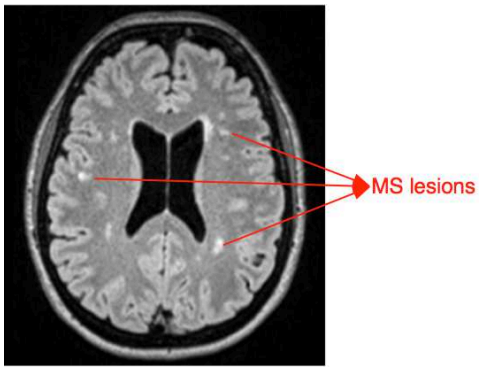


Fig. 1: Example of MR image with MS lesions.

The 2D Gaussian signal model can be written as:

$$[\boldsymbol{x}]_{\boldsymbol{p}} = a \exp\left(-\frac{1}{2}(\boldsymbol{p}-\boldsymbol{q})^{\mathrm{t}}\mathbf{B}^{\mathrm{t}}\mathbf{D}^{-1}\mathbf{B}(\boldsymbol{p}-\boldsymbol{q})\right) \quad (3)$$

where $[\boldsymbol{x}]_{\boldsymbol{p}}$ denotes the intensity value of the added signal at the 2D coordinate $\boldsymbol{p}$. The signal intensity attenuation is hereby modeled by a Gaussian function of peak amplitude a, centered at $\boldsymbol{q}$. The diagonal matrix $\mathbf{D}$ in Eq. (4) specifies the ellipse's scale σ and shape b ($b \geq 1$). Note that $\sqrt{b}\sigma$ and σ are one-half of the ellipse's major and minor axes respectively; when $b = 1$, the signal in Eq. (3) corresponds to a rotationally symmetric Gaussian function. The matrix $\mathbf{B}$ in Eq. (4) is a geometric rotation matrix, which rotates points in the xy-Cartesian plane counterclockwise through an angle θ about the origin of the Cartesian coordinate system.

$$\mathbf{D} = \begin{bmatrix} b\sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (4)$$

The signal parameters can be grouped in a vector:

$$\alpha = [a, \theta, b, \sigma, \boldsymbol{q}] \quad (5)$$

Note that other choices of signal models are also possible with our numerical observer. In general, we can consider a more general form of the signal:

$$[\boldsymbol{x}]_{\boldsymbol{p}} = f_\alpha(\boldsymbol{p}) \quad (6)$$

where $f_\alpha(\boldsymbol{p})$ is a function that represents a general parametrized signal with parameters α. According to the SKS principle, the signal parameters are not known in advance, but are rather specified by their respective PDFs. Note that the exact shape of the signal does not matter as long as it is the same in all the training/test images, except for the parameters.

2.B. Joint Detection and Estimation (JDE)

One possible solution for performing signal detection under signal uncertainty [29] is an approach based on the joint detection and estimation (JDE) theory [22], in which α and $\mathcal{H}_k$ estimates are chosen jointly to maximize the joint posterior probability $P(\alpha, \mathcal{H}_k \,|\, \boldsymbol{g})$. In this approach, we consider the signal parameters as a true characteristic of a signal that is either present or not present in an image. The signal parameter estimates only have a physical meaning when a signal is successfully detected ($\mathcal{H}_1$). This allows us to make a decision of signal presence using the following maximum a posteriori probability (MAP) estimation:

$$\begin{aligned}\widehat{(\alpha, \mathcal{H}_k)} &= \arg\max_{\alpha,\mathcal{H}_k} P(\alpha, \mathcal{H}_k \,|\, \boldsymbol{g}) \\ &= \arg\max_{\alpha,\mathcal{H}_k} \frac{P(\boldsymbol{g} \,|\, \alpha, \mathcal{H}_k)P(\alpha)P(\mathcal{H}_k)}{P(\boldsymbol{g})} \\ &= \arg\max_{\alpha,\mathcal{H}_k} P(\boldsymbol{g} \,|\, \alpha, \mathcal{H}_k)P(\alpha)P(\mathcal{H}_k) \end{aligned} \quad (7)$$

where we exploited the statistical independence of α and $\mathcal{H}_k$ (resulting in $P(\alpha, \mathcal{H}_k) = P(\alpha)P(\mathcal{H}_k)$). Under the assumption of a zero-mean correlated Gaussian background Eq. (2), the conditional probability density function $P(\boldsymbol{g} \,|\, \alpha, \mathcal{H}_k)$ can be written as:

$$\begin{aligned} P(\boldsymbol{g} \,|\, \alpha, \mathcal{H}_k) = &\frac{1}{\sqrt{(2\pi)^M \,|\, \Sigma_{\boldsymbol{b}} \,|}} \\ &\cdot \exp\left\{-\frac{1}{2}(\boldsymbol{g} - k\boldsymbol{x}_\alpha)^{\mathrm{t}}{\Sigma_b}^{-1}(\boldsymbol{g} - k\boldsymbol{x}_\alpha)\right\} \end{aligned} \quad (8)$$

In view of the monotonic logarithmic function, the maximization Eq. (7) is then equivalent to:

$$\begin{aligned}\widehat{(\alpha, \mathcal{H}_h)} = &\arg\max_{\alpha,\mathcal{H}_k} \{\ln P(\mathcal{H}_h) + \ln P(\alpha) \\ &- \frac{1}{2}(\boldsymbol{g} - h\boldsymbol{x}_\alpha)^{\mathrm{t}}\Sigma_{\boldsymbol{b}}^{-1}(\boldsymbol{g} - h\boldsymbol{x}_\alpha)\Big\} \\ = &\arg\max_{\alpha,\mathcal{H}_h} \Big\{\ln P(\mathcal{H}_h) + \ln P(\alpha) - \frac{1}{2}(\boldsymbol{g}^{\mathrm{t}}\Sigma_{\boldsymbol{b}}^{-1}\boldsymbol{g} \\ &- h\boldsymbol{x}_\alpha^{\mathrm{t}}\Sigma_{\boldsymbol{b}}^{-1}\boldsymbol{g} - h\boldsymbol{g}^{\mathrm{t}}\Sigma_{\boldsymbol{b}}^{-1}\boldsymbol{x}_\alpha + h^2\boldsymbol{x}_\alpha^{\mathrm{t}}\Sigma_{\boldsymbol{b}}^{-1}\boldsymbol{x}_\alpha)\Big\} \\ = &\arg\max_{\alpha,\mathcal{H}_h} \{\ln P(\mathcal{H}_h) + \ln P(\alpha) \\ &- \frac{1}{2}(-h\boldsymbol{x}_\alpha^{\mathrm{t}}\Sigma_{\boldsymbol{b}}^{-1}\boldsymbol{g} - h\boldsymbol{x}_\alpha^{\mathrm{t}}\Sigma_{\boldsymbol{b}}^{-1}\boldsymbol{g} + h^2\boldsymbol{x}_\alpha^{\mathrm{t}}\Sigma_{\boldsymbol{b}}^{-1}\boldsymbol{x}_\alpha)\Big\} \\ = &\arg\max_{\alpha,\mathcal{H}_h} \{\ln P(\mathcal{H}_h) \\ &+ \ln P(\alpha) + h\boldsymbol{x}_\alpha^{\mathrm{t}}\Sigma_{\boldsymbol{b}}^{-1}(\boldsymbol{g} - \frac{1}{2}h\boldsymbol{x}_\alpha)\Big\} \end{aligned} \quad (9)$$

It can be shown that this joint optimization problem can be solved by two sequential steps [24]: in the first step, the signal parameters are estimated irrespective of the hypothesis $\mathcal{H}_k$. In the second step, the estimated signal parameters are used to detect the signal. We will now describe these steps in more detail.

2.B.1. Estimation

The estimation problem here is actually to find the parameters maximizing the probability in a MAP sense, as given in Eq. (10), which is equivalent to maximizing separately $P(\alpha \,|\, \mathcal{H}_k, \boldsymbol{g})$ for $k = 0, 1$, then comparing the two results and choosing the bigger one. As already mentioned, the estimated parameters have no physical meaning in the case of hypothesis $\mathcal{H}_0$ (signal is absent).

$$\begin{aligned}\widehat{\alpha} &= \arg\max_{\alpha} P(\alpha \,|\, \boldsymbol{g}) \\ &= \arg\max_{\alpha} \max\{ \\ &\quad P(\boldsymbol{g} \,|\, \boldsymbol{x}_\alpha, \mathcal{H}_1)P(\alpha)P(\mathcal{H}_1), P(\boldsymbol{g} \,|\, \boldsymbol{x}_\alpha, \mathcal{H}_0)P(\alpha)P(\mathcal{H}_0)\} \\ &= \arg\max_{\alpha} \max\{ \\ &\quad P(\boldsymbol{g} \,|\, \boldsymbol{x}_\alpha, \mathcal{H}_1)P(\alpha)P(\mathcal{H}_1), P(\boldsymbol{g} \,|\, \mathcal{H}_0)P(\alpha)P(\mathcal{H}_0)\}\end{aligned} \tag{10}$$

Without loss of generality, two additional assumptions are often taken in practice to further simplify the expression:

(1) the parameters are uniformly distributed over the admissible space ($P(\alpha) \propto 1$);

(2) the two hypotheses $k = 0, 1$ are equiprobable: $P(\mathcal{H}_0) = P(\mathcal{H}_1)$.

These simplifications are adopted in this work to facilitate the validation of the model observer for SKS task, though an extension to a more generic model can be derived directly. Thus from Eq. (9) and Eq. (10), the estimation problem becomes:

$$\widehat{\alpha} = \arg\max_{\alpha} \left\{ \max \left\{ \boldsymbol{x}_\alpha^{\mathrm{t}} \Sigma_{\boldsymbol{b}}^{-1} (\boldsymbol{g} - \frac{1}{2}\boldsymbol{x}_\alpha), 0 \right\} \right\} \tag{11}$$

where 0 corresponds to the case $k = 0$ (signal absence). It follows that:

- if for all α, $\boldsymbol{x}_\alpha^{\mathrm{t}} \Sigma_{\boldsymbol{b}}^{-1} (\boldsymbol{g} - \frac{1}{2}\boldsymbol{x}_\alpha) \leq 0$, then the maximization in Eq. (11) is meaningless since the maximum is always zero;

- otherwise $\widehat{\alpha} = \arg\max_\alpha \left\{ \boldsymbol{x}_\alpha^{\mathrm{t}} \Sigma_{\boldsymbol{b}}^{-1} (\boldsymbol{g} - \frac{1}{2}\boldsymbol{x}_\alpha) \right\}$.

Consequently the estimation algorithm can be written as:

1. Calculate α_t:

$$\alpha_t = \arg\max_{\alpha} \left\{ \boldsymbol{x}_\alpha^{\mathrm{t}} \Sigma_{\boldsymbol{b}}^{-1} (\boldsymbol{g} - \frac{1}{2}\boldsymbol{x}_\alpha) \right\}; \tag{12}$$

2. Determine $\widehat{\alpha}$ according to the following rules:

$$\widehat{\alpha} = \begin{cases} \alpha_t & \text{if } \boldsymbol{x}_{\alpha_t}^{\mathrm{t}} \Sigma_{\boldsymbol{b}}^{-1} (\boldsymbol{g} - \frac{1}{2}\boldsymbol{x}_{\alpha_t}) > 0 \\ \text{any value in the space of } \alpha & \text{else} \end{cases} \tag{13}$$

It should be noted that in general there is no analytical solution for Eq. (12) because of the nonlinearity and non-concavity of the function. An alternative is to solve the maximization problem by an iterative method, such as Gauss-Newton approaches, non-linear conjugate gradient methods, etc. The iterative algorithm is then run over the parameter space of α, according to some updating rule which often involves matrix multiplications for each iteration.

2.B.2. Detection

Now we consider the posterior probabilities of the two hypotheses given the estimated signal parameters $\widehat{\alpha}$ and a particular observation $\boldsymbol{g}$:

$$\begin{aligned}P(\mathcal{H}_k \,|\, \widehat{\alpha}, \boldsymbol{g}) &= \frac{P(\mathcal{H}_k)P(\widehat{\alpha})P(\boldsymbol{g} \,|\, \mathcal{H}_k, \widehat{\alpha})}{P(\boldsymbol{g} \,|\, \widehat{\alpha})P(\widehat{\alpha})} \\ &\propto P(\mathcal{H}_k)P(\boldsymbol{g} \,|\, \mathcal{H}_k, \widehat{\alpha}) \quad k = 0, 1\end{aligned} \tag{14}$$

where $P(\boldsymbol{g} \,|\, \widehat{\alpha})$ in the denominator is omitted since it is not a function of $\mathcal{H}_k$.

The classical decision approach, based on signal detection theory, is to choose $\mathcal{H}_1$ when

$$P(\mathcal{H}_1 \,|\, \widehat{\alpha}, \boldsymbol{g}) > P(\mathcal{H}_0 \,|\, \widehat{\alpha}, \boldsymbol{g})$$

and $\mathcal{H}_0$ in the reverse case, which leads to the decision rule as below:

$$\begin{aligned}&\ln P(\mathcal{H}_1) + \ln P(\boldsymbol{g} \,|\, \mathcal{H}_1, \widehat{\alpha}) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \ln P(\mathcal{H}_0) + \ln P(\boldsymbol{g} \,|\, \mathcal{H}_0, \widehat{\alpha}) \\ \Leftrightarrow\ &-\frac{1}{2}(\boldsymbol{g} - \boldsymbol{x}_{\widehat{\alpha}})^{\mathrm{t}} \Sigma_{\boldsymbol{b}}^{-1} (\boldsymbol{g} - \boldsymbol{x}_{\widehat{\alpha}}) + \frac{1}{2}\boldsymbol{g}^{\mathrm{t}} \Sigma_{\boldsymbol{b}}^{-1} \boldsymbol{g} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \ln \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} \\ \Leftrightarrow\ &\lambda = \boldsymbol{x}_{\widehat{\alpha}}^{\mathrm{t}} \Sigma_{\boldsymbol{b}}^{-1} \left(\boldsymbol{g} - \frac{1}{2}\boldsymbol{x}_{\widehat{\alpha}} \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \ln \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)}\end{aligned} \tag{15}$$

It is easy to see that the test statistic λ is linear in $\boldsymbol{g}$, given the optimal estimated signal parameters, the background covariance matrix $\Sigma_{\boldsymbol{b}}$ and the prior probabilities of two hypotheses $P(\mathcal{H}_k)$ (often assumed to be equal without special knowledge of the data set).

Note that the test statistic is actually realized jointly with the estimation step. In other words, the comparison $\boldsymbol{x}_{\alpha_t}^{\mathrm{t}} \Sigma_{\boldsymbol{b}}^{-1} (\boldsymbol{g} - \frac{1}{2}\boldsymbol{x}_{\alpha_t}) > 0$ determines both the estimate of the signal parameters Eq. (13) and the validation of $\mathcal{H}_k$ (detection decision) Eq. (15), where the name JDE comes from.

2.C. Channelized Joint detection and estimation Observer (CJO)

Due to the high dimensionality of modern images, the direct calculation of the test statistic in Eq. (15) is practically infeasible: the inversion of the covariance matrix $\Sigma_{\boldsymbol{b}}^{-1}$ costs $\mathcal{O}(M^3)$, knowing that the dimension of $\Sigma_{\boldsymbol{b}}$ is $M \times M$ when the number of pixels in the image is M.

The inversion of Σ_b can be implemented more efficiently by using iterative solvers or by including an assumption of spatial stationarity. However, this approach still suffers from problems due to the "curse of dimensionality" [30] and very large numbers of images are typically required in order to obtain reliable estimates of the covariance.

Instead, as in other model observers [1, 7], we employ a linear dimension reduction technique that is called channelization: we project each input image onto a relatively small set of pre-defined channels. The channels are then designed to preserve the information that is relevant for the model observer, so that the ideal performance of the observer (e.g. in an asymptotic sense) is not affected by the channel selection. For more details, the reader may refer to [7].

The maximization in Eq. (12) based on the channel-

ized image $\boldsymbol{g}'$ is then formulated as follows:

$$\alpha_t = \arg\max_{\alpha} \frac{\boldsymbol{x}_\alpha^{\mathrm{t}}}{\|\mathbf{U}_\alpha\|_F^2} \left(\mathbf{U}_\alpha \left(\Sigma_{\boldsymbol{b}}'\right)^{-1} \mathbf{U}_\alpha^{\mathrm{t}}\right) \left(\boldsymbol{g} - \frac{1}{2}\boldsymbol{x}_\alpha\right) \tag{16}$$

where $\|\mathbf{U}_\alpha\|_F^2$ is a channel matrix energy normalization factor and $\|\cdot\|_F$ denotes the Frobenius norm of the matrix [24]. Ideally, we want the function in Eq. (16) to be close enough to that in Eq. (12), so that the impact of the dimentionality reduction is minimal. Intuitively, this implies that the covariance matrix $\Sigma_{\boldsymbol{b}}'$ and the channel matrix $\mathbf{U}_\alpha$ should be calibrated properly, namely $\Sigma_{\boldsymbol{b}}^{-1} \approx \frac{1}{\|\mathbf{U}_\alpha\|_F^2} \mathbf{U}_\alpha \left(\Sigma_{\boldsymbol{b}}'\right)^{-1} \mathbf{U}_\alpha^{\mathrm{t}}$. This is a sufficient condition but not necessary for that the α_t obtained by Eq. (16) is approximate to that given by Eq. (12).

The maximization problem in Eq. (16) is hardly realizable since practically it needs a whole set of channel matrices $\mathbf{U}_\alpha$ with varying α (the search space of α is huge). Our solution in [24] is to solve this problem by searching the optimal parameters in the channel domain instead of the spatial domain, without loss of accuracy. This requires that a transform on the signal in the spatial domain can be expressed as an equivalent transform on the signal in the channel domain, therefore an extra requirement for designing the channel matrix is added:

$$\mathbf{U}_\alpha = \mathbf{A}_\alpha^{\mathrm{t}} \mathbf{U} = \mathbf{U} \left(\mathbf{A}_\alpha'\right)^{\mathrm{t}} \tag{17}$$

where $\mathbf{U}$ is a fixed channel matrix which does not depend on α and conduce to reduce the data dimensionality; $\mathbf{A}_\alpha$ serve to map the parametric signal $\boldsymbol{x}_\alpha$ onto a reference signal $\boldsymbol{x}_0$ (with a known orientation/scale): $\mathbf{A}_\alpha \boldsymbol{x}_\alpha = \boldsymbol{x}_0$; $\mathbf{A}_\alpha'$ map the channelized parametric signal $\boldsymbol{x}_\alpha'$ onto the channelized reference signal $\boldsymbol{x}_0'$: $\mathbf{A}_\alpha' \boldsymbol{x}_\alpha' = \boldsymbol{x}_0'$, and the size of $\mathbf{A}_\alpha'$ is normally much smaller than that of $\mathbf{A}_\alpha$.

In channel domain, the estimation formulas Eq. (12), Eq. (13) and Eq. (15) could be rewritten as:

$$\alpha_t = \arg\max_{\alpha} \frac{1}{\left\|\mathbf{U}\left(\mathbf{A}_\alpha'\right)^{\mathrm{t}}\right\|_F^2} (\boldsymbol{x}_0')^{\mathrm{t}} \left(\Sigma_{\boldsymbol{b}}'\right)^{-1} \left(\mathbf{A}_\alpha' \boldsymbol{g}' - \frac{1}{2}\boldsymbol{x}_0'\right) \tag{18}$$

$$\text{such that} \quad \widehat{\alpha} = f(\alpha_t) \quad \text{as defined in } Eq.\ (13) \tag{19}$$

$$\text{and } \lambda = \frac{1}{\left\|\mathbf{U}\left(\mathbf{A}_{\widehat{\alpha}}'\right)^{\mathrm{t}}\right\|_F^2} (\boldsymbol{x}_0')^{\mathrm{t}} \left(\Sigma_{\boldsymbol{b}}'\right)^{-1} \left(\mathbf{A}_{\widehat{\alpha}}' \boldsymbol{g}' - \frac{1}{2}\boldsymbol{x}_0'\right) \tag{20}$$

which are used by the Channelized Joint detection and estimation Observer (CJO).

2.D. CJO for the amplitude-orientation-scale-unknown task

In this subsection, we will apply the CJO to the detection task in which the amplitude, orientation and scale of the signal are unknown. In Goossens et al. [24], suitable channels for different specific tasks of detection have been derived. The simplest task is the amplitude-unknown case for which:

$$\mathbf{A}_a' = a^{-1}\mathbf{I} \tag{21}$$

Here, the matrix $\mathbf{A}_a'$ scales the signal $\boldsymbol{x}_a = a\boldsymbol{x}_0$ with the reciprocal of its amplitude, resulting in the reference signal $\boldsymbol{x}_0$. The scaling itself does not impose any further restrictions with respect to the channel design. In fact, when only the signal amplitude is unknown and to be estimated, any type of channels, such as the Laguerre-Gauss channels [7] or the Difference of Gaussians (DOG) channels [8] can be used.

For more complicated detection tasks, in which next to the signal amplitude, also other parameters needs to be estimated, optimal estimation is obtained by choosing the channel matrix suitably. In case of the orientation-unknown detection task, a set of *orientation-steerable* channels needs to be used, while in the scale-unknown detection task *scale-shiftable* channels are required. In polar-frequency coordinates, let ω be the radial frequency and φ be the angular orientation. The orientation-steerable channels are given by:

$$f_{\theta_k}(\varphi) = \frac{(K-1)!2^{K-1}}{\sqrt{K(2K-2)!}} \left(\cos(\varphi - \theta_k)\right)^{K-1} \tag{22}$$

where $\theta_k = \frac{(k-1)\pi}{K}$ for $k = 1, \ldots, K$ are evenly spaced analysis angles and K is the number of steerable channels. Note that other choices than Eq. (22) are possible that are steerable in orientation.

Steerability in orientation means that any rotated version of $f_{\theta_k}(\varphi)$ can be generated by computing a linear combination of a fixed number of channels $f_{\theta_k}(\varphi)$, with $k = 1, \ldots, K$, or $f_{\theta'}(\varphi) = \sum_{k=1}^{K} \left[\mathbf{A}_{\theta'}'\right]_{1k} f_{\theta_k}(\varphi)$. Instead of computing the channel responses for channels with a continuous parameter θ_k, the same response can be obtained by calculating a linear combination of the responses for a fixed number of channels. Furthermore, this operation also preserves the energy of the channelized image $\boldsymbol{g}'$. This is reflected in the fact that the channels themselves are normalized in terms of energy:

$$\sum_{k=1}^{K} |f_{\theta_k}(\varphi)|^2 = 1 \tag{23}$$

Moreover, remark that the channels Eq. (22) are well-localized in orientation: the maximum is reached for $\varphi = \theta_k$, and we further have that $f_{\theta_k}(\varphi) < f_{\theta_k}(\theta_k)$.

Next, the scale-shiftable channels are defined by a logarithmic warping of the radial frequency axis:

$$f_{\sigma_j}(\omega) = \operatorname{sinc}\left(\operatorname{sign}(\omega) \log_2\left(\frac{|\, 2^{\sigma_j}\omega \,|}{\pi}\right)\right) \qquad |\,\omega\,| < \pi; \tag{24}$$

where $\sigma_j = 2^{j-1}$ for $j = 1, \ldots, J$ and J is the number of scale-shiftable channels. It can easily be checked that the scale-shiftable channels are also normalized in energy, in

the following sense:

$$\sum_{j=-\infty}^{+\infty} \left|f_{\sigma_j}(\omega)\right|^2 = 1 \tag{25}$$

By exploiting the fact that $f_{\theta_k}(\varphi)$ only depends on the polar angle φ and that $f_{\sigma_j}(\omega)$ depends on the radial frequency ω, channel functions can be composed that are jointly steerable in orientation and shiftable in scale:

$$f_{\theta_k,\sigma_j}(\omega,\varphi) = f_{\theta_k}(\varphi) f_{\sigma_j}(\omega) \quad k = 1,\ldots,K; j = 1,\ldots,J \tag{26}$$

These channels are particularly useful to solve the amplitude-orientation-scale-unknown detection task, because with a suitable choice of $\mathbf{A}'_\alpha$, equation Eq. (17) can be satisfied. An example of channels ($K = 3$ and $J = 4$) is shown in Figure. 2.

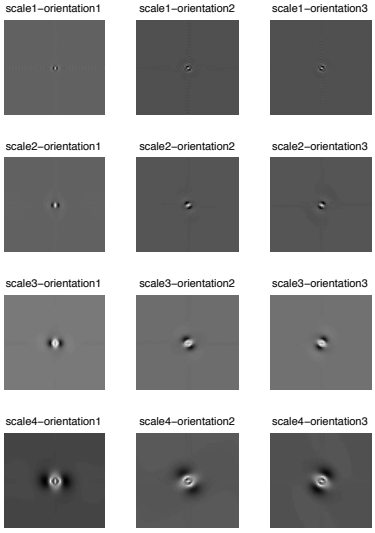


Fig. 2: An example of the filters for the amplitude-orientation-scale-unknown case, when the number of steerable channels $K = 3$ and the number of scale-shiftable channels $J = 4$

A channel matrix of dimensions $M^2 \times JK$ is constructed by computing the inverse polar Fourier transform of Eq. (26) and subsequently by storing each sampled channel function $[\mathbf{U}]_{m,n}$ in a column, as follows:

$$\begin{aligned}[\mathbf{U}]_{m,n} = &\int_0^{2\pi}\int_0^{+\infty} f_{\theta_k}(\varphi) f_{\sigma_j}(\omega) \\ &\cdot \exp\left\{2\pi i\left[\left(m_1 - \frac{M}{2}\right)\omega\cos\varphi \right.\right. \\ &\left.\left. + \left(m_2 - \frac{M}{2}\right)\omega\sin\varphi\right]\right\}\omega \mathrm{d}\varphi \mathrm{d}\omega \end{aligned} \tag{27}$$

with $m = (m_2 - 1)M + m_1$ and $n = (j-1)K + k$. Here $[\cdot]_{m,n}$ denotes the matrix element of the m-th line and n-th column. i is the imaginary unit.

Two corresponding transform matrices for the orientation-unknown case and scale-unknown case are given by Eq. (28) (a Dirichlet kernel) and Eq. (29) (a sinc interpolation, which is a method to reconstruct a continuous-time bandlimited signal from a set of equally spaced samples) respectively:

$$[\mathbf{A}'_\theta]_{m,n} = \frac{1}{K}\frac{\sin(\pi(m-n) - \theta K)}{\sin(\pi(m-n)/K - \theta)} \quad m,n = 1,\ldots,K; \tag{28}$$

$$[\mathbf{A}'_\sigma]_{m,n} = \mathrm{sinc}\left(((\mathrm{m}-\mathrm{n}) - (-\log_2\sigma))\right) \quad \mathrm{m},\mathrm{n} = 1,\ldots,\mathrm{J}; \tag{29}$$

where θ is the signal orientation and σ is the signal scale; and the sinc function is:

$$\mathrm{sinc(t)} = \begin{cases} 1 & t = 0 \\ \frac{\sin(\pi t)}{\pi t} & t \neq 0 \end{cases} \tag{30}$$

Interestingly, exploiting the separability in polar frequencies of Eq. (26), the joint transform matrix can be constructed using the Kronecker product '$\otimes$' of two matrices:

$$\mathbf{A}'_\alpha = \mathbf{A}'_{a,\theta,\sigma} = \frac{1}{a}\mathbf{A}'_\sigma \otimes \mathbf{A}'_\theta \tag{31}$$

Note that the order of appearance of $\mathbf{A}'_\sigma$ and $\mathbf{A}'_\theta$ in Eq. (31) depends on the sample storage ordering during the construction of channel matrix $\mathbf{U}$.

The matrix $\mathbf{A}'_\theta$ is orthonormal:

$$(\mathbf{A}'_\theta)^{\mathrm{t}}\mathbf{A}'_\theta = \mathbf{A}'_\theta(\mathbf{A}'_\theta)^{\mathrm{t}} = \mathbf{I}. \tag{32}$$

This is useful for reducing the calculation burden of the normalization factor

$$\left\|\mathbf{U}(\mathbf{A}'_\alpha)^{\mathrm{t}}\right\|_F^2 = \frac{1}{a^2}\left\|\mathbf{U}(\mathbf{A}'_\sigma \otimes \mathbf{A}'_\theta)^{\mathrm{t}}\right\|_F^2, \tag{33}$$

which costs $\mathcal{O}(M^2 \cdot (JK)^2)$ by direct evaluation. A simplified numerical method is provided here.

From the definition of the Frobenius norm, we have:

$$\begin{aligned} &\left\|\mathbf{U}(\mathbf{A}'_\sigma \otimes \mathbf{A}'_\theta)^{\mathrm{t}}\right\|_F^2 \\ =&\mathrm{trace}\left((\mathbf{A}'_\sigma \otimes \mathbf{A}'_\theta)\mathbf{U}^{\mathrm{t}}\mathbf{U}(\mathbf{A}'_\sigma \otimes \mathbf{A}'_\theta)^{\mathrm{t}}\right) \\ =&\mathrm{trace}\left(\mathbf{U}^{\mathrm{t}}\mathbf{U}(\mathbf{A}'_\sigma \otimes \mathbf{A}'_\theta)^{\mathrm{t}}(\mathbf{A}'_\sigma \otimes \mathbf{A}'_\theta)\right) \\ =&\mathrm{trace}\left(\mathbf{U}^{\mathrm{t}}\mathbf{U}\left((\mathbf{A}'_\sigma)^{\mathrm{t}} \otimes (\mathbf{A}'_\theta)^{\mathrm{t}}\right)(\mathbf{A}'_\sigma \otimes \mathbf{A}'_\theta)\right) \\ =&\mathrm{trace}\left(\mathbf{U}^{\mathrm{t}}\mathbf{U}\left[\left((\mathbf{A}'_\sigma)^{\mathrm{t}}\mathbf{A}'_\sigma\right) \otimes \left((\mathbf{A}'_\theta)^{\mathrm{t}}\mathbf{A}'_\theta\right)\right]\right) \\ =&\mathrm{trace}\left(\mathbf{U}^{\mathrm{t}}\mathbf{U}\left[\left((\mathbf{A}'_\sigma)^{\mathrm{t}}\mathbf{A}'_\sigma\right) \otimes \mathbf{I}\right]\right) \end{aligned} \tag{34}$$

Let $(\mathbf{A}'_\sigma)^{\mathrm{t}}\mathbf{A}'_\sigma = S = [s_{ij}]$ $(i,j = 1,\ldots,J)$, then divide the matrix $\mathbf{U}^{\mathrm{t}}\mathbf{U}$ (dimension $JK \times JK$) into $J \times J$ submatrices T_{ij} $(i,j = 1,\ldots,J)$, each of dimension $K \times K$, the right hand can be written as:

$$\begin{aligned} \sum_{i=1}^{J}\mathrm{trace}\left(\sum_{j=1}^{J} s_{ji}T_{ij}\right) &= \sum_{i=1}^{J}\mathrm{trace}\left(\sum_{j=1}^{J} s_{ij}T_{ij}\right) \\ &= \sum_{i=1}^{J}\sum_{j=1}^{J}\mathrm{trace}\left(s_{ji}T_{ij}\right) \\ &= \sum_{i=1}^{J}\sum_{j=1}^{J} s_{ij}\mathrm{trace}\left(T_{ij}\right) \end{aligned} \tag{35}$$

Therefore it costs only $\mathcal{O}(J^2)$ to calculate the normalization factor given an odd value of K. We also note that T_{ij} $(i,j=1,\ldots,J)$ is known a priori, independent of image data, thus trace (T_{ij}) can be pre-calculated.

3. Practical implementation of the CJO

In this section, we give more details on how the CJO can be implemented and used in practice. We consider both the training phase and testing phase of the model and we discuss a regularization technique to deal with the ill-conditioning of the covariance matrix when the number of channels is large.

3.A. Stages of implementation

As SKE CHO's implementation, we can realize the CJO for the amplitude-orientation-scale-unknown task by training and test stages [7].

3.A.1. Training stage

In the training stage, the inputs are also two sets of images (with-signal and without-signal) whose ground truth is known for the CJO, including the signal profile x_α. Let $\boldsymbol{g}$ denote a particular image, $\boldsymbol{x}_\alpha$ a particular signal.

The outputs of this stage are the estimated channelized reference signal $\boldsymbol{x}_0'$ and a template $\boldsymbol{w}$.

$$\widehat{\boldsymbol{x}_0'} = \langle \mathbf{A}'_{a,\theta,\sigma}\boldsymbol{x}'_\alpha\rangle = \langle \mathbf{A}'_{a,\theta,\sigma}(\mathbf{U}^{\mathrm{t}}\boldsymbol{x}_\alpha)\rangle \tag{36}$$

$$\boldsymbol{w} = \left(\widehat{\Sigma'_{\boldsymbol{b}}}\right)^{-1}\widehat{\boldsymbol{x}_0'} \tag{37}$$

where $\langle\cdot\rangle$ denotes the sample average. $\widehat{\Sigma'_{\boldsymbol{b}}}$ is the estimated background covariance matrix:

$$\begin{aligned}\widehat{\Sigma'_{\boldsymbol{b}}} =& \frac{1}{2}\langle(\boldsymbol{g}'-\langle\boldsymbol{g}'\,|\,H_0\rangle)\,(\boldsymbol{g}'-\langle\boldsymbol{g}'\,|\,H_0\rangle)^{\mathrm{t}}\,|\,H_0\rangle + \\ & \frac{1}{2}\langle[(\boldsymbol{g}'-\boldsymbol{x}'_\alpha)-\langle(\boldsymbol{g}'-\boldsymbol{x}'_\alpha)\,|\,H_1\rangle] \\ & \cdot[(\boldsymbol{g}'-\boldsymbol{x}'_\alpha)-\langle(\boldsymbol{g}'-\boldsymbol{x}'_\alpha)\,|\,H_1\rangle]^{\mathrm{t}}\,|\,H_1\rangle\end{aligned} \tag{38}$$

where $\boldsymbol{g}'$ is the channelized image with the M^2-by-JK channel matrix $\mathbf{U}$ constructed from Eq. (26):

$$\boldsymbol{g}' = \mathbf{U}^{\mathrm{t}}\boldsymbol{g} \tag{39}$$

Recall that during the training stage, the exact values of α are assumed to be available (in contrast to the testing stage). Consequently, the training stage of the CJO is quite similar to the training stage of the CHO.

3.A.2. Test stage

In the test stage, $\boldsymbol{w}$ and $\boldsymbol{x}_0'$ are used to estimate the signal parameters, as well as to calculate the test statistic λ for each input test image $\boldsymbol{g}$, the ground truth is however unknown to the CJO.

To find the maximum in Eq. (18), as a first step, we have simply used a brute-force search for a fixed number (denoted by N) of equally spaced signal parameters. Then the maximum of the test statistics yielded by all the possible combinations of chosen parameters is chosen as the final test statistic:

$$\begin{aligned}\lambda =& \max_{a,\theta,\sigma}(\lambda_{a,\theta,\sigma}) \\ =& \max_{a,\theta,\sigma}\left(\frac{\boldsymbol{w}^{\mathrm{t}}}{\left\|\mathbf{U}\left(\mathbf{A}'_{a,\theta,\sigma}\right)^{\mathrm{t}}\right\|_F^2}\left(\mathbf{A}'_{a,\theta,\sigma}\boldsymbol{g}'-\frac{1}{2}\widehat{\boldsymbol{x}_0'}\right)\right)\end{aligned} \tag{40}$$

Because the optimization in Eq. (40) takes place in the channel domain, the maximization of Eq. (40) is still relatively fast, even when done using a brute-force search.

Finally, the detection performance of the model can be calculated using the area under the ROC Curve (AUC) (e.g., employing the Mann-Witney-Wilcoxon test) [31], or alternatively using the area under the EROC [21].

3.B. Tikhonov regularization

Note that the estimated covariance matrix $\widehat{\Sigma'_{\boldsymbol{b}}}$ in Eq. (38) becomes ill-conditioned when the number of channels is large ($(K,J)=(3,9),(5,8),(7,8),(9,7)$...). That causes the numerical instabilities (big errors) when evaluating the CJO's performance. Thus we use the Tikhonov regularization [32] method to estimate the covariance matrix here.

Instead of calculating Eq. (37) by direct inverse as the solution of a linear problem :

$$\widehat{\Sigma'_{\boldsymbol{b}}}\cdot\boldsymbol{w} = \widehat{\boldsymbol{x}_0'}, \tag{41}$$

a bias is introduced deliberately in the solution research by considering the following problem :

$$\min_{\boldsymbol{w}}\left\{\frac{1}{2}\|\widehat{\Sigma'_{\boldsymbol{b}}}\cdot\boldsymbol{w}-\widehat{\boldsymbol{x}_0'}\|^2+\frac{1}{2}\eta\|\boldsymbol{w}\|^2\right\}, \tag{42}$$

for which the positive coefficient η controls the regularization level of the expected solution.

Note that the empirical estimate matrix $\widehat{\Sigma'_{\boldsymbol{b}}}$ is by construction (cf. Eq. (38)) non-negative definite and allows the following singular value decomposition (SVD) :

$$\widehat{\Sigma'_{\boldsymbol{b}}} = \mathbf{V}\mathbf{D}\mathbf{V}^{\mathrm{t}}, \tag{43}$$

where $\mathbf{V}$ is an orthonormal matrix containing normalized eigenvectors ($\mathbf{V}\mathbf{V}^{\mathrm{t}}=\mathbf{I}$), and $\mathbf{D}=\mathrm{diag}(d_1,\ldots,d_p)$ is a diagonal matrix whose main diagonal contains eigenvalues in the descending order $d_1\geq\ldots\geq d_p>0$. We recall that the condition number is the ratio d_1/d_p. Its direct inverse writes thus $\widehat{\Sigma'}^{-1}=\mathbf{V}\mathbf{D}^{-1}\mathbf{V}^{\mathrm{t}}$.

In the SVD point of view, the regularization method can then be applied to replace $\left(\widehat{\Sigma'_{\boldsymbol{b}}}\right)^{-1}$ in Eq. (37) with

$$\boldsymbol{w}^\dagger = \Sigma^\dagger\widehat{\boldsymbol{x}_0'} \tag{44}$$

$$\Sigma^\dagger = \left(\widehat{\Sigma'}^{\mathrm{t}}\widehat{\Sigma'}+\eta\mathbf{I}\right)^{-1}\widehat{\Sigma'}^{\mathrm{t}} = \mathbf{V}\begin{bmatrix}\frac{d_1}{\eta+d_1^2} & & \\ & \ddots & \\ & & \frac{d_p}{\eta+d_p^2}\end{bmatrix}\mathbf{V}^{\mathrm{t}} \tag{45}$$

by choosing $d_1^2\gg\eta\gg d_p^2$ for the ill-conditioned cases. We note that for cases with relatively small condition

numbers, the regularization has negligible effects by imposing $d_1^2 \gg \eta$: $\Sigma^\dagger$ is close to the direct inverse of $\widehat{\Sigma'_{\boldsymbol{b}}}$ and thus the detection performances should not change significantly. In fact, for the limit case $\eta = 0$ (45) even reduces to the pseudo-inverse. On the contrary, for ill-conditioned cases the presence of η avoids the division by zero with the eigenvalues such as d_p.

4. Performance evaluation of the CJO

In this section, we will evaluate the CJO's estimation and detection performance using both the correlated Gaussian backgrounds and the real clinical backgrounds.

4.A. Parameter setup for the CJO

For both synthesized images and real clinical images, the image size is 65×65; 500 pairs of images are used for the training stage (500 with-signal and 500 without-signal images), and 100 pairs of images are used for the test stage (note that the test images are different from the training images). The simulated lesion is located in the center of the images. The signal orientation range is $[\theta_{\min},\ \theta_{\max}] = [0,\ \pi]$. The signal amplitude range $[a_{\min},\ a_{\max}]$ and the signal scale range $[\sigma_{\min},\ \sigma_{\max}]$ cover the corresponding maximum accessible parameter space according to the background type. For example, $[a_{\min},\ a_{\max}] = [1,\ 255]$ (pixel intensity value), and $[\sigma_{\min},\ \sigma_{\max}] = [1,\ 6]$ with $\sqrt{b} = 2$ for real clinical images. N used in the brute-force search is set to be 20. We use $\eta = 10^{-14} d_1^2$ for the Tikhonov regularization.

4.B. Performance evaluation of the CJO

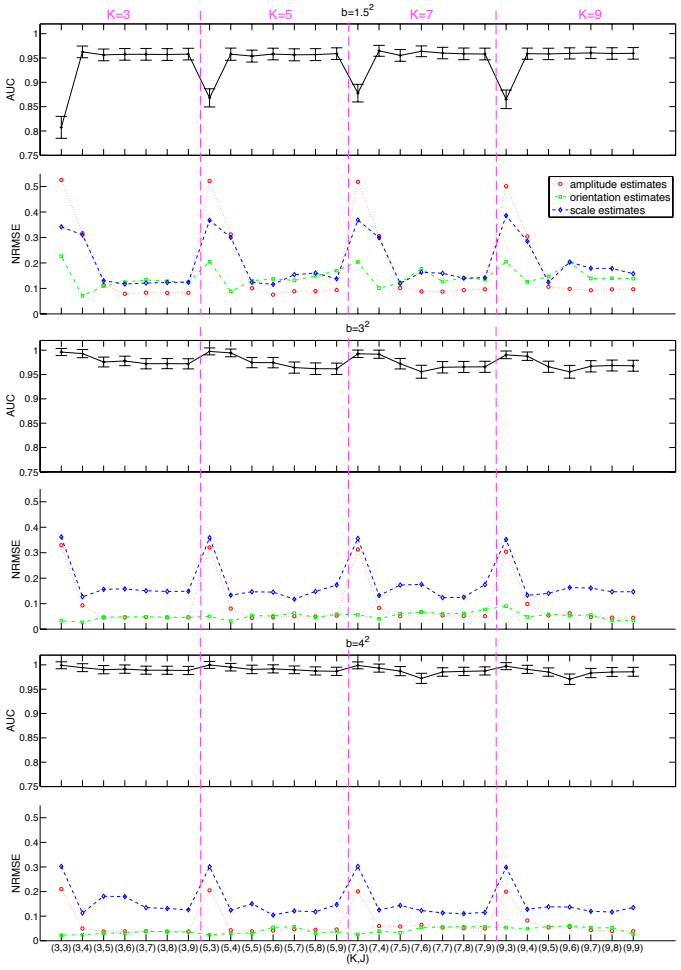
In this study, we used the AUC and the estimate of its variance [33] as the FOM for the detection performance evaluation. In general, a higher value of AUC means a better detection performance. Moreover, we used the normalized root-mean-square error (NRMSE) [34] to quantify the signal parameter estimation errors to facilitate their comparison under different setups. The NRMSE is the root-mean-square error (RMSE) [35] divided by the range of true values. Lower NRMSE value indicates that the difference between the estimated value and the true value is smaller. When computing the NRMSE for the error in orientation estimation, we took the wrapping effect into account.

4.B.1. CJO's performance on correlated Gaussian background

First, we tested CJO's performances using Correlated Gaussian Background (CGB). The CGB here was generated by convolving white noise following the distribution $\mathcal{N}$ (0,1) with a 2D Gaussian kernel characterized by $\sigma_{\boldsymbol{b}} = 10$.

Plots of AUC and NRMSE for the CJO with different combinations of number of *steerable* channels K and number of *scale-shiftable* channels J are shown in Figure. 3. Different signal shape is also considered: $\sqrt{b} = 1.5, 3, 4$ (note that the signal shape estimation is out of the scope of the CJO). A bigger value of b means the elliptical signal gets a more elongated shape.

Fig. 3: AUCs and NRMSEs as a function of (K, J) for correlated Gaussian background when $\sqrt{b} = 1.5, 3, 4$.



4.B.2. CJO's performance on real clinical background

Secondly, we tested CJO's performances using the real clinical background. The real clinical background here was a white matter region, extracted from healthy brain MR axial images of a T2 FLAIR sequence (acquired from a retrospective database of the University Hospital of Angers). Examples of white matter regions and those with simulated lesions are shown in Figure. 4. Plots of AUC and NRMSE are shown in Figure. 5.

4.C. Discussion

Throughout all the results, we observe that:

1) Among different parameters, the CJO estimates better the signal orientation than the signal amplitude and scale. And low orientation estimation error helps to gain a really good detection task performance ($AUC > 0.95$), even though amplitude and scale estimation errors remain higher (cf. Figure. 3 and Figure. 5 where $\sqrt{b} = 3, 4$). Ultimately the orientation estimation seems

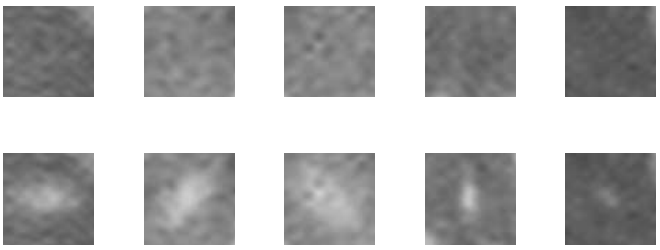
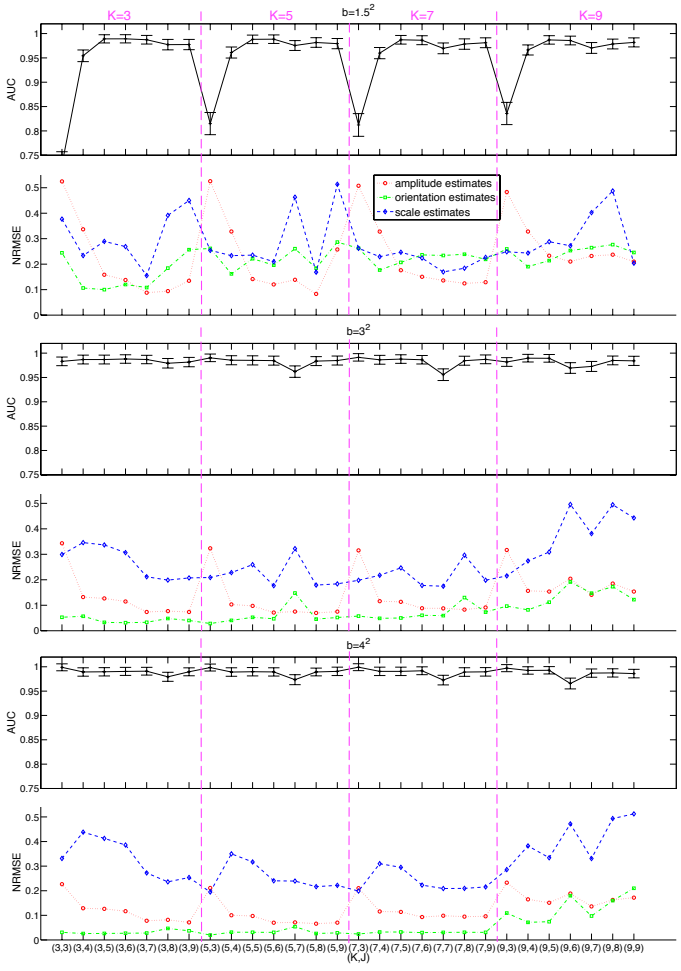


Fig. 4: First row: examples of white matter regions (WMR); second row: WMR plus simulated lesions.

Fig. 5: AUCs and NRMSEs as a function of K, J for real clinical background when $\sqrt{b} = 1.5, 3, 4$.



to have the greatest effect on the detection performance among the three unknown signal parameters.

2) Contrary to the number of *scale-shiftable* channels J, the number of *steerable* channels K does not influence a lot the detection and estimation performances. The detection performance tends to be stable when J reaches 4 or 5. Since fewer channel number requires fewer calculation complexities, the optimal channel numbers could be $(K, J) = (3, 4)$ or $(3, 5)$ here for this given dataset.

3) For certain (K, J) settings, e.g., $(K, J) = (3, 3), (5, 3), (7, 3), (9, 3)$, the estimation errors are more important (bigger NRMSE values), which lead to lower values of AUC, viz. worse detection task performance.

4) For both types of background, the CJO's estimation and detection performances are better for more elongated signal (bigger value of b).

5) The CJO performs better on the CGB than on the real clinical background. This is understandable considering the background model used in the CJO is indeed the CGB and the complexity of the problem in real clinical situations. The values of AUC are close to 1, which signifies that the CJO performs the detection task very well while being capable of estimating signal parameters.

5. Conclusions

In this paper, we proposed a SKS MO (CJO) to address an important task involving medical images, i.e. the detection of a signal with subsequent estimation of para-meters (amplitude, orientation and size) that characterize the signal. Compared to few existing SKS MOs, the CJO extends the range of variable signal parameters and has a good practicability (with a reasonable number of training images and calculation burden). The performance study results present that the detection task performances of the CJO are desirable both for the CGB and MR image regions of the T2 FLAIR sequence. The limitation of the CJO is that the signal should be asymmetric, so that it has a dominant orientation, which is captured by the parameters θ and b.

Acknowledgment

This work was supported by the Pays de La Loire region in France. The authors are grateful for the funding of EQuIMOSe (Subjective and objective Evaluation of the Quality of Medical Images for an Optimal use of the display, archiving and transmission Systems).

References

[1] H. Barrett and K. Myerrs, *Foundations of Image Science* (John Wiley and Sons, Inc., Hoboken, New Jersey, USA, 2004).

[2] J. Brankov, Y. Yang, L. Wei, I. El Naqa, and M. Wernick, "Learning a channelized observer for image quality assessment," IEEE Transactions on Medical Imaging **28**, 991–999 (2009).

[3] H. H. Barrett, J. Yao, J. P. Rolland, and K. J. Myers, "Model observers for the assessment of image quality," in "Proceedings of the National Academy of Sciences," , vol. 90, D. C. Wilson and J. N. Wilson, eds. (SPIE, 1993), vol. 90, pp. 9758–9765.

[4] L. Platisa, B. Goossens, E. Vansteenkiste, S. Park, B. D. Gallas, A. Badano, and W. Philips, "Channelized hotelling observers for the assessment of volumetric imaging data sets," Journal of the Optical Society of America A **28**, 1145–1163 (2011).

[5] C. Abbey and J. Boone, "An ideal observer for a model of x-ray imaging in breast parenchymal tissue," Digital Mammography pp. 393–400 (2010).

[6] R. F. W. Harrison H. Barrett, J. L. Denny and K. J. Myers, "Objective assessment of image quality. II. fisher information, fourier crosstalk, and figures of merit for task performance," Journal of the Optical Society of America A **12**, 834–852 (1995).

[7] B. D. Gallas and H. H. Barrett, "Validating the use of channels to estimate the ideal linear observer." Journal of the Optical Society of America A **20**, 1725–1738 (2003).

[8] C. K. Abbey and H. H. Barrett, "Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability," Journal of the Optical Society of America A **18**, 473–488 (2001).

[9] C. Lartizien, P. Kinahan, and C. Comtat, "Volumetric model and human observer comparisons of tumor detection for whole-body positron emission tomography," Academic radiology **11**, 637–648 (2004).

[10] S. Park, E. Clarkson, M. A. Kupinski, and H. H. Barrett, "Efficiency of the human observer detecting random signals in random backgrounds," Journal of the Optical Society of America A **22**, 3–16 (2005).

[11] M. Whitaker, E. Clarkson, and H. Barrett, "Estimating random signal parameters from noisy images with nuisance parameters: linear and scanning-linear methods," Optics express **16**, 8150 (2008).

[12] R. M. Manjeshwar and D. L. Wilson, "Effect of inherent location uncertainty on detection of stationary targets in noisy image sequences," Journal of the Optical Society of America A **18**, 78–85 (2001).

[13] C. Castella, M. P. Eckstein, C. K. Abbey, K. Kinkel, F. R. Verdun, R. S. Saunders, E. Samei, and F. O. Bochud, "Mass detection on mammograms: influence of signal shape uncertainty on human and model observers," Journal of the Optical Society of America A **26**, 425–436 (2009).

[14] M. P. Eckstein and C. K. Abbey, "Model observers for signal-known-statistically tasks (SKS)," in "Proceedings of SPIE," , vol. 4324 (SPIE, 2001), vol. 4324, pp. 91–102.

[15] M. P. Eckstein, B. Pham, and C. K. Abbey, "Effect of image compression for model and human observers in signal-known-statistically tasks," in "Proceedings of SPIE," , vol. 4686 (SPIE, 2002), vol. 4686, pp. 91–102.

[16] M. P. Eckstein, Y. Zhang, B. Pham, and C. K. Abbey, "Optimization of model observer performance for signal known exactly but variable tasks leads to optimized performance in signal known statistically tasks," in "Proceedings of SPIE," , vol. 5034 (SPIE, 2003), vol. 5034, pp. 91–102.

[17] Y. Zhang, B. T. Pham, M. Eckstein, and S. Barbara, "Evaluation of JPEG 2000 encoder options: human and model observer detection of variable signals in x-ray coronary angiograms," IEEE Transactions on Medical Imaging **23**, 613–632 (2004).

[18] H. C. Gifford and M. A. King, "A comparison of human and model observers in multislice LROC studies," IEEE Transactions on Medical Imaging **24**, 160–169 (2005).

[19] H. Gifford and M. King, "Implementing visual search in human-model observers for emission tomography," in "Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE," (IEEE, 2009), pp. 2482–2485.

[20] H. Gifford and M. King, "Em clustering for holistic search in human-model observers," in "Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE," (IEEE, 2010), pp. 3584–3587.

[21] E. Clarkson, "Estimation receiver operating characteristic curve and ideal observers for combined detection/estimation tasks," JOSA A **24**, B91–B98 (2007).

[22] G. Olmo, E. Magli, and L. L. Presti, "Joint statistical signal detection and estimation. part I: theoretical aspects of the problem," Signal Processing **80**, 57–73 (2000).

[23] B. Goossens, L. Platiša, E. Vansteenkiste, and W. Philips, "The use of steerable channels for detecting asymmetrical signals with random orientations," in "Proceedings of SPIE," , vol. 7627 (SPIE, 2010), vol. 7627.

[24] B. Goossens, L. Platiša, and W. Vansteenkiste, E.and Philips, "Design of Model Observers for joint Detection and Estimation of Random Parametric Signals in Images," (2012). In preparation.

[25] L. Zhang, C. Cavaro-Ménard, P. Le Callet, and J. Tanguy, "A perceptually relevant channelized joint observer (PCJO) for the detection-localization of parametric signals," IEEE Transactions on Medical Imaging **31**, 1875–1888 (2012).

[26] S. Park, "Performance of a channelized-ideal observer using laguerre-gauss channels for detecting a gaussian signal at a known location in different lumpy backgrounds," in "Proceedings of SPIE," , vol. 6146 (Spie, 2006), vol. 6146, pp. 61460P–61460P–9.

[27] C. Castella, K. Kinkel, F. Descombes, M. P. Eckstein, P.-E. Sottas, F. R. Verdun, and F. O. Bochud, "Mammographic texture synthesis: second-generationclustered lumpy backgrounds using agenetic algorithm," Opt. Express **16**, 7595–7607 (2008).

[28] L. Zhang, C. Cavaro-Ménard, and P. L. Callet, "Key issues and specificities for the objective medical image quality assessment," in "Sixth International Workshop on Video Processing and Quality Metrics (VPQM)," , vol. 7966 (2012), vol. 7966.

[29] B. Goossens, "Multiresolution image models and estimation techniques," Ph.D. thesis, Ghent University, Ghent, Belgium (2010).

[30] R. Bellman, *Adaptive control processes: a guided tour*, vol. 4 (Princeton university press Princeton, 1961).

[31] J. A. Swets, *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers* (Lawrence Erlbaum Associates, Mahwah, New Jersey, 1996).

[32] A. Tihonov, "Solution of incorrectly formulated problems and the regularization method," Soviet Math **4**, 1035–1038 (1963).

[33] B. D. Gallas, "One-shot estimate of MRMC variance: AUC," Academic Radiology **13**, 353–362 (2006).

[34] M. Anderson and W. Woessner, *Applied Groundwater Modeling: Simulation of Flow and Advective Transport* (Academic Press, 1992), 2nd ed.

[35] J. Armstrong and F. Collopy, "Error measures for generalizing about forecasting methods: Empirical comparisons," International Journal of Forecasting **8**, 69–80 (1992).