Moving target's detection performances in a sequence of infrared multispectral images

Christian Musso, Sidonie Lefebvre, Sophie Thetas
DOTA - Optics and Associated Techniques
ONERA – The French Aerospace Lab
Palaiseau, France
{christian.musso ; sidonie.lefebvre ; sophie.thetas}@onera.fr

***Abstract*—The paper deals with the detection performance of a moving target in multispectral IR image sequences with low-SNR. In this context, track-before-detect (TBD) is the generally used method, which consists in accumulating raw images over time to track before detect. For each target hypothesis (position, velocity, amplitude), the signal is integrated over time. In this way, the potential target with the best spatio-temporal correlation will be a candidate for a detection test. In this paper, we develop a minimum detection bound regardless of the TBD method used. Specifically, this bound gives the minimum average number of multispectral images required to detect the target.**

***Index Terms*—target detection, stopping time, sequential test, track-before-detect, multispectral Infrared images, particle filter**

I. Introduction

Detecting and tracking small targets in infrared images is an important topic, in medical and security fields for example, but also a challenging task in computer vision, especially when it comes to differentiating these targets from noisy or textured backgrounds. In recent years, multispectral sensors have been developed, combining visible and thermal wavelengths domains [1]–[3]. They can help improve the tracking accuracy and the detection for low contrast targets in single-band visible or infrared images, and solve some camouflage and decoy problems. In order to assess the contribution of the different spectral bands to the tracking and the detection, it is key to obtain performance bounds for different multispectral configurations.

We are specifically interested in estimating the minimum number of multispectral images required (called instant detection), in average, to detect the target regardless of the detection method. This bound will enable detection performance to be evaluated in terms of the target's IR signature on each multispectral band and in terms of SNR.

In Section II, we present a simplified model of the problem of target detection using multispectral images. In Section III, we propose an estimate of the stopping time of the sequential probability ratio test in the context of nonlinear filtering. From this estimate, we develop in section IV a lower bound of the number of images required for detection, regardless of the detection method used. Section V is devoted to simulations in which the lower bound is compared with the results of a particle filter for track-before-detect [20]. The TBD is an energy integration technique aimed at improving detectability of weak targets [4]–[9].

II. Problem modeling

We assume that the target, if present, follows approximatively a rectilinear motion (constant velocity) in the multispectral infrared images. For each band i, the target amplitude A_i is assumed to be constant. At time k, the state vector X_k of dimension $d = 4 + m$ is composed of the positions x_k, y_k, of the amplitudes and of the velocities $\dot{x}_k, \dot{y}_k$,

$$X_k = [\underbrace{x_k, y_k}_{x_{1,k}}, \underbrace{\{A_i\}_{i=1}^{m}, \dot{x}_k, \dot{y}_k}_{x_{2,k}}]^T \tag{1}$$

where m is the number of bands. We consider a noise-free dynamical target state,

$$X_k = F\, X_{k-1} \tag{2}$$

where F is the transition matrix,

$$F = \begin{bmatrix} 1 & 0 & \cdots & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \cdots & 0 & \Delta T \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

ΔT is the sampling period. The initial uncertainty of the position components of X_0 is the whole image. The method we propose can easily be generalized to dynamic noise.

At each sampling time k, we have an image composed of m spectral bands, each of them has M resolution cells. The intensity $Y_k^i(\tau)$ of the target in the band i into the cell τ is modeled as follows,

$$Y_k^i(\tau) = A_i\, \phi_{psf}^i(\tau, x_{1,k}) + \epsilon_k^i(\tau) \tag{4}$$

The point-spread function ϕ_{psf}^i (PSF) describes the contribution of the intensity of the target localized in $x_{1,k}$ to the cell τ. $\epsilon_k^i(\tau)$ is a centered Gaussian random variable. Generally the PSF is shift invariant, it depends only on the difference $\tau - x_{1,k}$, namely: $\phi_{psf}^i(\tau, x_{1,k}) = \phi_{psf}^i(\tau - x_{1,k})$. We assume that the noises $\epsilon_k^i(\tau)$ are independent from cell to cell and from band to band. The PSF is normalized such as the total

target's energy becomes one: $\int_{\mathbb{R}^2}(\phi^i_{psf})^2(\tau,x)\,d\tau = 1$. For convenience, we vectorize the equation (4) as follows,

$$\begin{cases}\mathbf{Y}^i_k & = \left[Y^i_k(\tau^1),\ldots,Y^i_k(\tau^M)\right]^T \\ \mathbf{\Phi}^i_k(x_{1,k}) & = \left[\phi^i_{psf}(\tau^1,x_{1,k}),\ldots,\phi^i_{psf}(\tau^M,x_{1,k})\right]^T\end{cases} \quad (5)$$

where $\tau^j = [\tau_x,\tau_y]^j$ for $j = 1,\ldots,M$ are the cartesian coordinates of the cell τ^j in the image plane. The measurements vector $\mathbf{Y}^i_k$ of dimension $M \times 1$ and the multispectral measurements $\mathbf{Y}_k$ can be expressed in a compact form as follows,

$$\begin{cases}\mathbf{Y}^i_k = A_i\,\mathbf{\Phi}^i(x_{1,k}) + \epsilon^i_k \overset{\text{def}}{=} h_i(X_k) + \epsilon^i_k \\ \mathbf{Y}_k \overset{\text{def}}{=} \{\mathbf{Y}^i_k\}_{i=1}^m\end{cases} \quad (6)$$

The noise ϵ^i_k is normally distributed with zero mean and with the diagonal spatial covariance matrix: $R_i = diag(\sigma^2_i)$. R_i is of dimension $M \times M$. The likelihood g_k is expressed as follows,

$$g_k(X_k) \overset{\text{def}}{=} \mathbb{P}(\mathbf{Y}_k|X_k) \propto \prod_{i=1}^m \exp\left(-\frac{1}{2}[\mathbf{Y}^i_k - A_i\,\mathbf{\Phi}^i(x_{1,k})]^T R_i^{-1}[\mathbf{Y}^i_k - A_i\,\mathbf{\Phi}^i(x_{1,k})]\right) \quad (7)$$

The computation of the likelihood is performed in a vicinity of $x_{1,k}$ where $\mathbf{\Phi}^i(\cdot)$ (6) does not vanish as suggested in [10]. For each band i, the signature of the target is characterized by the PSF ϕ^i_{psf} (4). The SNR, expressed in decibels, for each band i is defined as follows,

$$\text{SNR}(i) = 10\log_{10}\left[\frac{A_i^2}{\sigma_i^2}\right] \quad (8)$$

In the following, we seek to determine the minimum number of multispectral IR images required to detect a target, regardless of the detection method used. For this purpose, we rely on the stopping time of the sequential detection test.

III. Stopping time of the SPRT

The sequential probability ratio test (SPRT) minimizes the average number of measurements [11], [12]. Specifically, given a detection probability P_d and a false alarm probability P_{fa}, no other test can achieve the same P_d and P_{fa} with a smaller expected number of samples. In our case, the samples are the multispectral images. We recall the formalism of the SPRT applied to nonlinear filtering [13].

A. SPRT for nonlinear filtering

Consider the following detection framework where H_0 is the null hypothesis "no target" and where H_1 is the hypothesis "target present",

$$\begin{cases}H_0 : \mathbf{Y}_k = \epsilon_k \text{ (noise only)} \\ H_1 : \mathbf{Y}_k = \mathbf{A}\,\mathbf{\Phi}(x_{1,k}) + \epsilon_k \overset{\text{def}}{=} h(X_k) + \epsilon_k\end{cases} \quad (9)$$

where $\mathbf{Y}_k$ denotes the observations $\mathbf{Y}^i_k$ of all the m bands at time k. The SPRT is based on the likelihood ratio considered sequentially. The likelihood ratio (LR) is defined as,

$$\Lambda_k = \frac{\mathbb{P}(\mathbf{Y}_{1:k}|H_1)}{\mathbb{P}(\mathbf{Y}_{1:k}|H_0)} \overset{\text{def}}{=} \frac{\mathbb{P}_1(\mathbf{Y}_{1:k})}{\mathbb{P}_0(\mathbf{Y}_{1:k})} = \frac{\prod_{n=1}^k \mathbb{P}_1(\mathbf{Y}_n|\mathbf{Y}_{1:n-1})}{\prod_{n=1}^k \mathbb{P}_0(\mathbf{Y}_n|\mathbf{Y}_{1:n-1})}$$

where $\mathbf{Y}_{1:k}$ is the measurement vector up to time k. We express the LR in terms of random walk [14]:

$$Z_k = \log(\Lambda_k) = \sum_{n=1}^k \log\left[\frac{\mathbb{P}_1(\mathbf{Y}_n|\mathbf{Y}_{1:n-1})}{\mathbb{P}_0(\mathbf{Y}_n|\mathbf{Y}_{1:n-1})}\right] \overset{\text{def}}{=} \sum_{n=1}^k z_n \quad (10)$$

In average, the random walk Z_k increases under H_1 and decreases under H_0 ($\mathbb{E}_1(z_n) \geq 0$) [14]. The test accepts H_1 as soon as the trajectory $(Z_k, k \geq 1)$ exceeds a threshold and accepts H_0 as soon as Z_k falls below a threshold,

$$\begin{cases}\text{as soon as } Z_k \leq \log(\Gamma_0) \Rightarrow \text{noise only} \\ \text{as soon as } Z_k \geq \log(\Gamma_1) \Rightarrow \text{target present}\end{cases} \quad (11)$$

The threshold Γ_i depends on the false alarm probability P_{fa} and on the detection probability P_d we set. When P_d is close to one and P_{fa} close to 0, the following bounds are appropriate [14],

$$\Gamma_0 = \frac{1-P_d}{1-P_{fa}}, \qquad \Gamma_1 = \frac{P_d}{P_{fa}} \quad (12)$$

Indeed, one can prove that the risks (probabilities of wrong decisions) of the test (11) are controlled. Precisely, using these bounds (12), we have: $\mathbb{P}[\text{decide } H_1|H_0] \leq P_{fa}$ and $\mathbb{P}[\text{decide } H_0|H_1] \leq 1 - P_d$.
Let K be the first time such that Z_k exceeds the threshold,

$$K \overset{\text{def}}{=} \left\{\inf\, k : Z_k = \sum_{n=1}^k z_n \geq \log(\Gamma_1)\right\} \quad (13)$$

K is a (random) stopping time since it depends on the measurements $\mathbf{Y}_{1:k}$. When the variables z_n are i.i.d, thanks to the Wald identity, the expected stopping time, for each hypothesis, can be approximated as follows,

$$\mathbb{E}_0(K) \approx \frac{\log(\Gamma_0)}{\mathbb{E}_0(z_1)}, \qquad \mathbb{E}_1(K) \approx \frac{\log(\Gamma_1)}{\mathbb{E}_1(z_1)} \quad (14)$$

We are interested in estimating $\mathbb{E}_1(K)$ that is, the minimum number of multispectral images required, in average, to detect the target regardless of the estimation method used. In our case, the random variables z_n are independent but do not have the same expectation (10). Therefore, (14) cannot be applied. In this case of non-stationary observations, a lower bound of $\mathbb{E}_1(K)$ has been proposed [15]. But the latter is intractable.

B. Approximation of $\mathbb{E}_1(K)$

We propose the following approximation of $\mathbb{E}_1(K)$. In the following, we will call $\mathbb{E}_1(K)$ the detection instant.

Proposition 1. *If $\varphi(k) \overset{def}{=} \sum_{i=1}^k \mathbb{E}_1(z_i)$ is nearly linear, the the detection instant can be approximated as follows,*

$$\mathbb{E}_1(K) \approx K^* \overset{def}{=} \left\{\inf\, k : \sum_{n=1}^k \mathbb{E}_1[z_n] \geq \log(\Gamma_1)\right\} \quad (15)$$

where K^* is deterministic.

Proof. As the random variables $[\mathbb{1}_{K\leq n-1}]$ and z_n are independent, we have,

$$\mathbb{E}_1\left[\sum_{n=1}^{K} z_n\right] = \sum_{n=1}^{\infty} \mathbb{E}_1\left[\mathbb{1}_{K\geq n}\, z_n\right] = \sum_{n=1}^{\infty} \mathbb{E}_1[z_n]\, \mathbb{P}_1(K \geq n)$$

Now, by summation by parts, we obtain,

$$\begin{aligned}
&\sum_{n=1}^{N} \mathbb{E}_1[z_n]\, \mathbb{P}_1(K \geq n] = \mathbb{P}_1(K \geq N)\, \mathbb{E}_1(Z_N) \\
&+ \sum_{n=1}^{N-1} \varphi(n)\, [\mathbb{P}_1(K \geq n) - \mathbb{P}_1(K \geq n+1)] \\
&= \mathbb{P}_1(K \geq N)\, \mathbb{E}_1(Z_N) + \sum_{n=1}^{N-1} \varphi(n)\, \mathbb{P}_1(K = n)
\end{aligned}$$

Assuming $\lim_{N\to+\infty} \mathbb{P}_1(K \geq N)\, \mathbb{E}_1(Z_N) = 0$, we obtain,

$$\begin{aligned}
&\mathbb{E}_1\left[\sum_{n=1}^{K} z_n\right] = \sum_{n=1}^{\infty} \left[\sum_{k=1}^{n} \mathbb{E}_1(z_k)\right] \mathbb{P}_1(K = n) \\
&= \mathbb{E}_1\left[\sum_{n=1}^{K} \mathbb{E}_1(z_n)\right] = \mathbb{E}_1[\varphi(K)] \\
&\approx \varphi\left[\mathbb{E}_1(K)\right] = \sum_{n=1}^{\mathbb{E}_1(K)} \mathbb{E}_1(z_n)
\end{aligned} \tag{16}$$

The last approximation is valid because φ is assumed to be nearly linear.
When Z_k (10) exceeds $\log(\Gamma_1)$ for the first time, Z_k is close to $\log(\Gamma_1)$, as the increments z_n are relatively small. Therefore, by definition of K (13) and K^* (15), we have,

$$\begin{cases}
\mathbb{E}_1\left[\sum_{n=1}^{K} z_n\right] \approx \log(\Gamma_1) \\
\sum_{n=1}^{K^*} \mathbb{E}_1(z_n) \approx \log(\Gamma_1)
\end{cases} \tag{17}$$

Now, using (16), we obtain the desired result,

$$\sum_{n=1}^{\mathbb{E}_1(K)} \mathbb{E}_1(z_n) \approx \sum_{n=1}^{K^*} \mathbb{E}_1(z_n) \tag{18}$$

□

In the case where the variables z_n have the same mean, we retrieve the Wald identity (14) from (16) and (17).
Note that it is not necessary for the function φ to be approximately linear on the whole positive half-line to obtain $E_1[\varphi(K)] \approx \varphi[E_1(K)]$. In fact, it is sufficient for this function to be approximately linear on an interval containing K with a probability close to 1. This is the case for example if $\mathbb{V}(K)$ is small which occurs when the SNR (8) is low ($E_1(z_{K+1})$ close to $E_1(z_K)$). It remains to estimate $\mathbb{E}_1[z_n]$.

IV. Lower bound of the detection instant

A. Cramér-Rao bound

We recall briefly the Cramér-Rao lower bound (CRB). The CRB captures the information of the target signature for each IR band. As will be seen below, the CRB is involved in the estimation of the detection instant.
The CRB [16] is a useful tool in signal processing that applies in particular for evaluating the performances of tracking algorithms. This bound has been generalized to random dynamic process: the posterior Cramér-Rao lower bound (PCRB) [17]. This bound does not depend on the realizations on the observation process but depends only on the dynamic process model and on the observation process model. It can be calculated offline and reflects the performances of an ideal filter: the estimation errors of the unknown state obtained by any unbiased estimator are lower bounded by the PCRB,

$$\forall n \geq 0 \quad P_n \geq J_n^{-1} = PCRB(n) \tag{19}$$

where P_n is the covariance matrix of any filter. Asymptotically as $(n \to \infty)$ an efficient filter generally reaches this bound. Hereafter, we derive the CRB for the observation model (4). Since noises are independent from band to band, the global information matrix J_n at time n is the sum of the individual information matrices J_n^i for the band i. In the case in which the dynamic is linear and noise free (2), the PCRB reduces to the classical CRB. The CRB at time n is the inverse of J_n. The latter is expressed recursively as follows,

$$J_n = \sum_{i=1}^{m} \frac{\partial h_i^T(X_n)}{\partial X_n} R_i^{-1} \frac{\partial h_i(X_n)}{\partial X_n} + (F^T)^{-1} J_{n-1}\, F^{-1} \tag{20}$$

where F is the transition matrix (3) and where the measurement functions h_i are defined in (6). The recursion starts with $J_0 = P_0^{-1}$, the inverse of the initial covariance matrix of the initial state X_0. J_n is of dimension $(4+m, 4+m)$ where m is the number of bands. The matrices J_n^i are function of the PSF ϕ_{psf}^i in each band i and function of the SNR different from band to band. For example, we can consider an elliptically distributed target signature in the band i, characterized by the covariance matrix Σ_i. This yields,

$$\begin{aligned}
&Y_k^i(\tau) = \\
&\frac{A_i}{\sqrt{\pi}\det(\Sigma_i)^{\frac{1}{4}}} \exp\left(-\frac{1}{2}[\tau - x_{1,k}]^T\, \Sigma_i^{-1}[\tau - x_{1,k}]\right) + \epsilon_k^i(\tau) \\
&= A_i\, \phi_{psf}^i(\tau, x_{1,k}) + \epsilon_k^i(\tau) = h_i(X_k, \tau) + \epsilon_k^i(\tau)
\end{aligned} \tag{21}$$

The target energy is normalized such that $\int_{\mathbb{R}^2} (\phi_{psf}^i)^2(\tau, x)\, d\tau = 1$. To assess the information matrix (20) we calculate the derivatives of the observation function

h_i. For each cell τ and for each band i, we have,

$$\begin{cases} \dfrac{\partial h_i}{\partial x_{1,k}}(X_k,\tau) & = - A_i\, \Sigma_i^{-1}(x_{1,k}-\tau)\, \phi^i_{psf}(\tau, x_{1,k}) \\ \dfrac{\partial h_i}{\partial A_i}(X_k,\tau) & = \phi^i_{psf}(\tau, x_{1,k}) \\ \dfrac{\partial h_i}{\partial \dot{x}_k}(X_k,\tau) & = 0 \\ \dfrac{\partial h_i}{\partial \dot{y}_k}(X_k,\tau) & = 0 \end{cases} \tag{22}$$

B. Lower bound of the detection instant

We aim to estimate a tight lower bound of K^* (15) which approximates $\mathbb{E}_1(K)$ (15), the minimal number, in average, of images to detect a target. To do this (15), we first approximate the following jump of the random walk (10),

$$z_n = \log\left[\frac{\mathbb{P}_1(\mathbf{Y}_n|\mathbf{Y}_{1:n-1})}{\mathbb{P}_0(\mathbf{Y}_n|\mathbf{Y}_{1:n-1})}\right] \tag{23}$$

The next proposition provides an approximation of z_n as function of the SNR and of the information matrix J_n.

Proposition 2. *Assuming that the predicted density $\mathbb{P}_1(X_n|\mathbf{Y}_{1:n-1})$ is Gaussian, the jump of the random walk z_n can be approximated as follows,*

$$\begin{aligned} z_n \approx & -\frac{1}{2}\log\,\det(\hat{J}_n) + \frac{1}{2}\log\,\det(\hat{J}_{n|n-1}) \\ & - \frac{1}{2}\sum_{i=1}^{m}\left[\mathbf{Y}^i_n - h_i(\hat{x}_n)\right]^T R_i^{-1}\left[\mathbf{Y}^i_n - h_i(\hat{x}_n)\right] \\ & - \frac{1}{2}\left[\hat{x}_n - \hat{x}_{n|n-1}\right]^T \hat{J}_{n|n-1}\left[\hat{x}_n - \hat{x}_{n|n-1}\right] \\ & + \frac{1}{2}\sum_{i=1}^{m}(\mathbf{Y}^i_n)^T R_i^{-1}\mathbf{Y}^i_n \end{aligned} \tag{24}$$

where $\hat{J}_n = J_n(\hat{x}_n)$ and $\hat{J}_{n|n-1} = (F^T)^{-1}\hat{J}_{n-1}(\hat{x}_n)\,F^{-1}$ are the information matrices (20) evaluated at the maximum a posteriori (MAP) $\hat{x}_n$ and where $\hat{x}_{n|n-1} = F\,\hat{x}_{n-1}$ is the predicted MAP.

Proof. For the sake of clarity the approximation is proved for the one band case ($m = 1$). The predicted pdf of $\mathbf{Y}_n$ given the past observations $\mathbf{Y}_{1:n-1}$ can be expressed as follows,

$$\mathbb{P}_1(\mathbf{Y}_n|\mathbf{Y}_{1:n-1}) = \int_{\mathbb{R}^d} \underbrace{\mathbb{P}_1(\mathbf{Y}_n|X_n)}_{\text{Likelihood}} \underbrace{\mathbb{P}_1(X_n|\mathbf{Y}_{1:n-1})}_{\text{Predicted density}}\, dX_n \tag{25}$$

We approximate this integral by applying the Laplace method which states that,

$$\int_{\mathbb{R}^d} \mathrm{e}^{-\psi_n(x)}dx \approx (2\pi)^{d/2}\mathrm{e}^{-\psi_n(\hat{x}_n)}\det\left[\,\psi''_n(\hat{x}_n)\right]^{-1/2} \tag{26}$$

with

$$\psi_n(x) \stackrel{\text{def}}{=} -\log\mathbb{P}_1(\mathbf{Y}_n|X_n = x) - \log\mathbb{P}_1(X_n = x|\mathbf{Y}_{1:n-1}) \tag{27}$$

$\hat{x}_n$ is the global minimum of ψ_n namely the MAP,

$$\hat{x}_n = \arg\max_{x\in\mathbb{R}^d}\ \mathbb{P}_1(\mathbf{Y}_n|X_n = x)\,\mathbb{P}_1(X_n = x|\mathbf{Y}_{1:n-1}) \tag{28}$$

The approximation (26) is generally very accurate particularly where the integrand has an exponential decay. We assume that the predicted density $\mathbb{P}_1(X_n = \hat{x}_n|\mathbf{Y}_{1:n-1})$ is a Gaussian pdf with mean $\hat{x}_{n|n-1}$ and with covariance matrix $\hat{J}^{-1}_{n|n-1}$. The likelihood $\mathbb{P}_1(\mathbf{Y}_n|X_n = \hat{x}_n)$ is a Gaussian pdf with mean $h(\hat{x}_n)$ (9) and with covariance matrix R (6). Note that $\psi''(\hat{x}_n) = J_n(\hat{x}_n)$ is the observed (or posterior) information matrix. Thanks to (26), we obtain after some calculations,

$$\begin{aligned} \log\mathbb{P}_1(\mathbf{Y}_n|\mathbf{Y}_{1:n-1}) = & -\frac{1}{2}\log\,\det(\hat{J}_n) - \frac{1}{2}\log\det(R) \\ & - \frac{1}{2}\log\,\det(\hat{J}_{n|n-1}) - \frac{1}{2}\left[\mathbf{Y}_n - h(\hat{x}_n)\right]^T R^{-1}\left[\mathbf{Y}_n - h_i(\hat{x}_n)\right] \\ & - \frac{1}{2}\left[\hat{x}_n - \hat{x}_{n|n-1}\right]^T \hat{J}_{n|n-1}\left[\hat{x}_n - \hat{x}_{n|n-1}\right] - \frac{M}{2}\log(2\pi) \end{aligned} \tag{29}$$

where M is the dimension of $\mathbf{Y}_n$ (5). We derive now $\mathbb{P}_0(\mathbf{Y}_n|\mathbf{Y}_{1:n-1})$ (under H_0), the pdf of the Gaussian noise distribution (6),

$$\begin{aligned} & \log\mathbb{P}_0(\mathbf{Y}_n|\mathbf{Y}_{1:n-1}) = \log\mathbb{P}_0(\mathbf{Y_n}) \\ & = -\frac{M}{2}\log(2\pi) - \frac{1}{2}\log\det(R) - \frac{1}{2}\,\mathbf{Y}_n^T R^{-1}\mathbf{Y}_n \end{aligned} \tag{30}$$

Putting together the equations (29) and (30), we obtain the desired result.

□

The proof can be extended simply in the case of m bands. Indeed, since the bands $(\mathbf{Y}^i_n\,,\, i = 1,\ldots,m)$ are independent given X_n, the cumulative likelihood is expressed as: $\mathbb{P}_1(\mathbf{Y}_n|X_n) = \prod_{i=1}^m \mathbb{P}_1(\mathbf{Y}^i_n|X_n)$. We apply then the Laplace method by taking ψ_n (27) in which $\mathbb{P}_1(\mathbf{Y}_n|X_n)$ is now the cumulative likelihood and in which $\mathbb{P}_1(X_n = x|\mathbf{Y}_{1:n-1})$ is the predicted pdf given the past of all multispectral images. We now need to estimate $\mathbb{E}_1(z_n)$ (15) to assess the detection instant (15). To be more precise, the next proposition provides an upper bound of $\mathbb{E}_1(z_n)$.

Proposition 3 (Upper bound of the jump of the random walk)**.** *Under the hypothesis H_1 (target present), the average jump of z_n (10) verifies,*

$$\mathbb{E}_1(z_n) \leq \frac{1}{2}\sum_{i=1}^{m}\frac{A_i^2}{\sigma_i^2} + \frac{1}{2}\,\log\left[\frac{\det\,(J_{n-1})}{\det\,(J_n)}\right] \stackrel{\text{def}}{=} \beta_n \tag{31}$$

J_n and J_{n-1} are evaluated at the true state respectively at X_n and X_{n-1}.

Proof. We calculate $\mathbb{E}_1(z_n)$ term by term using (24). The expectations under H_1 of the observed information matrices (24) are approximated by substituting the MAP with the true values of the state. That is, $\mathbb{E}_1[\log\det(\hat{J}_n)] \approx \log\det[J_n(X_n)]$ and the same applies for $\mathbb{E}_1[\log\det(\hat{J}_{n-1})]$. Note that $\det(\hat{J}_{n|n-1}) = \det(\hat{J}_{n-1})$ since $\det(F) = 1$ (3). We also have $\mathbb{E}_1[\left(\mathbf{Y}^i_n - h_i(\hat{x}_n)\right)^T R_i^{-1}\left(\mathbf{Y}^i_n - h_i(\hat{x}_n)\right)] \approx M$,

the degree of freedom of the $\chi 2$ distribution (M is the dimension of $\mathbf{Y}_n^i$ and R_i the covariance matrix of $\mathbf{Y}_n^i$). Under H_1 (target present), we have (6),

$$\mathbb{E}_1\left[(\mathbf{Y}_n^i)^T R_i^{-1}\mathbf{Y}_n^i\right] = \mathbb{E}_1\left[(\mathbf{Y}_n^i - A_i\mathbf{\Phi}^i)^T R_i^{-1}(\mathbf{Y}_n^i - A_i\mathbf{\Phi}^i)\right] + 2\,A_i\,\mathbb{E}_1(\mathbf{Y}_n^i)R_i^{-1}\mathbf{\Phi}^i - A_i^2\,(\mathbf{\Phi}^i)^T R_i^{-1}\mathbf{\Phi}^i$$

Recall that $R_i = diag(\sigma_i^2)$ and $(\mathbf{\Phi}^i)^T\mathbf{\Phi}^i = 1$ (see section II), we obtain: $\mathbb{E}_1\left[(\mathbf{Y}_n^i)^T R_i^{-1}\mathbf{Y}_n^i\right] = M + \frac{A_i^2}{\sigma_i^2}$. Denoting $\xi_n = -\frac{1}{2}\left[\hat{x}_n - \hat{x}_{n|n-1}\right]^T \hat{J}_{n|n-1}\left[\hat{x}_n - \hat{x}_{n|n-1}\right]$, we finally derive an approximation of $\mathbb{E}_1(z_n)$,

$$\mathbb{E}_1(z_n) \approx \frac{1}{2}\sum_{i=1}^{m}\frac{A_i^2}{\sigma_i^2} + \frac{1}{2}\log\left[\frac{\det\left(J_{n-1}\right)}{\det\left(J_n\right)}\right] - \frac{1}{2}\mathbb{E}_1(\xi_n) \tag{32}$$

The expectation of the positive random variable ξ_n is difficult to estimate. □

The mean of the random jump $\mathbb{E}_1(z_n)$ is thus upper bounded resulting in a lower bound of $\mathbb{E}_1(K)$ (15) subject to the following proposition. Indeed, with a larger positive increment β_n (31), fewer increments are required for the deterministic walk to exceed $\log(\Gamma_1)$ (15).

Proposition 4 (Lower bound of the detection instant)**.** *The minimum detection instant, in average, obtained by any detection test based on the dynamics model (2) and on the measurement model (6) is lower bounded as follows,*

$$\mathbb{E}_1(K) \geq \tilde{K}^* \stackrel{def}{=} \Big\{\inf\, k : \frac{k}{2}\sum_{i=1}^{m}\frac{A_i^2}{\sigma_i^2} + \frac{1}{2}\sum_{n=1}^{k}\log\left[\frac{\det\left(J_{n-1}\right)}{\det\left(J_n\right)}\right] \geq \log(\Gamma_1)\Big\} \tag{33}$$

where $\Gamma_1 = \frac{P_d}{P_{fa}}$ and where J_n and J_{n-1} are evaluated at the true state respectively at X_n and X_{n-1}. $\sum_{i=1}^{m} A_i^2/\sigma_i^2$ stands for the multispectral SNR.

Therefore, $\tilde{K}^*$ provides a lower bound of the minimum number of multispectral images, in average, required to detect the target. $\tilde{K}^*$ depends on all the parameters setting the model. It depends on the signature of the target (PSF) in each band through the information matrix J_n (20) and on the SNR of each band (8). J_n is based on the recursion (20) where P_0 is the initial covariance matrix of the state. Note that the computing cost of $\tilde{K}^*$ is negligible. Moreover, the matrices J_n can be computed offline as they depend entirely on the observation and dynamic process models. This bound allows the detection performance to be evaluated as function of the target's IR signature on each multispectral band and of SNR.

V. Experimental

In this section, we evaluate the relevance of the lower bound $\tilde{K}^*$ (33) by comparing it with the empirical instant detections provided by a track-before-detect (TBD) method. The TBD is an energy integration technique aimed at improving detectability of weak targets. A number of techniques for TBD design and implementation have been proposed, including particle filtering [4]–[7] that we will use. Useful surveys can be found in [18], [19].

The particle filter aims to estimate the posterior pdf at time k,

$$p_k(x_k) \stackrel{\text{def}}{=} p(X_k = x_k|\mathbf{Y}_{1:k}) \propto g_k(x_k)\,q_k(x_k) \tag{34}$$

where $q_k(x_k) = p(X_k = x_k|\mathbf{Y}_{1:k-1})$ is the prior (or predicted) pdf and where g_k is the likelihood (7). The posterior is approximated by a weighted sum of N Dirac functions,

$$p_k(x_k) \approx \sum_{i=1}^{N} w_k^i\,\delta(x_k = X_k^i) \tag{35}$$

where the particles $\{X_k^i\}_{i=1}^N$ are approximatively samples of the posterior,

$$X_k^i = [x_k^i, y_k^i, \left\{A_j^i\right\}_{j=1}^m, \dot{x}_k^i, \dot{y}_k^i]^T \tag{36}$$

We apply here the Laplace particle filter (LPF) described in [20]. We recall briefly this TBD algorithm based on the LPF.

A. Laplace particle filter

The principle behind this filter is based on a new resampling stage taking (see algorithm 1) into account the current measurement $\mathbf{Y}_k$. At this stage, new particles $\tilde{X}_k^i$ are sampled around local maxima a posteriori (MAP) $\hat{\xi}_k^j$ solutions of (28). Precisely, the particles $\tilde{X}_k^i$ are generated according to the following importance function (IF),

$$\tilde{q}_k(x) = \sum_{j=1}^{n_c}\rho_k^j\,\phi\left(x, \hat{\xi}_k^j, P_{k|k-1}\right) \tag{37}$$

with $\sum_{j=1}^{n_c}\rho_k^j = 1$. ϕ is a Gaussian pdf with mean $\hat{\xi}_k^j$ and predicted covariance matrix $P_{k|k-1}$ defined in (38). The maximum number of local MAP, n_c, is a parameter we set. The weights ρ_k^j and the way to obtain a fast estimation of the local MAP are described in [20].

The predicted pdf q_k is assumed to be Gaussian with mean $\hat{X}_{k|k-1}$ and covariance matrix $P_{k|k-1}$,

$$\begin{cases}\hat{X}_{k|k-1} = \sum_{i=1}^{N} w_{k-1}^i X_{k|k-1}^i \\ P_{k|k-1} = \sum_{i=1}^{N} w_{k-1}^i (X_{k|k-1}^i - \hat{X}_{k|k-1})(X_{k|k-1}^i - \hat{X}_{k|k-1})^T\end{cases} \tag{38}$$

B. Empirical detection instant

When a new IR multispectral image $\mathbf{Y}_n$ is available, we can update the random jump z_n (23). To this end, we first estimate the distribution $\mathbb{P}_1(\mathbf{Y}_n|\mathbf{Y}_{1:n-1})$ using the particles. The particle filter draws weighted samples $\{X_{n|n-1}^i, w_{n-1}^i\}_{i=1}^N$ (prediction step in algorithm 1) from the predicted distribution

Algorithm 1 The Laplace-based particle filter

1) **Initialisation** (k=1) For $i = 1, \ldots, N$. Generate the particles X^i_{k-1}. Put the particles $x^i_{1,k-1}$ (the position components) in each cell and generate $x^i_{2,k-1}$ (amplitude and velocities) randomly with the weights $w^i_{k-1} \equiv 1/N$

2) **Prediction** For $i = 1, \ldots, N$. Propagate the weighted particles by applying the target dynamics (2), $X^i_{k|k-1} = F\, X^i_{k-1}$

3) **Correction** For $i = 1, \ldots, N$. Compute the likelihood (7) $g_k(X^i_{k|k-1}) = P(\mathbf{Y}_k|X^i_{k|k-1})$ and the weights $w^i_k \propto g_k(X^i_{k|k-1})\, w^i_{k-1}$ such that $\sum_{i=1}^N w^i_k = 1$. Compute $N_{eff} = \frac{1}{\sum_{i=1}^N [w^i_k]^2}$

   * If $N_{eff} \geq N_{th}$. The corrected particles are $\left(X^i_k, w^i_k\right) = \left(X^i_{k|k-1}, w^i_k\right)$

   * If $N_{eff} < N_{th}$. Perform Laplace-based resampling.
   Compute the local modes of the posterior $\{\hat{\xi}^j_k\}_{j=1}^{n_c}$ (28)
   Generate samples $\tilde{X}^i_k$ from the proposal $\tilde{q}_k$ (37) and compute the importance weights $\tilde{w}^i_k \propto \frac{g_k(\tilde{X}^i_k)\, q_k(\tilde{X}^i_k)}{\tilde{q}_k(\tilde{X}^i_k)}$ such that $\sum_{i=1}^N \tilde{w}^i_k = 1$.
   The corrected particles are $\left(X^i_k, w^i_k\right) = \left(\tilde{X}^i_k, \tilde{w}^i_k\right)$

4) **State estimation** The state is estimated by $\hat{X}_k = \sum_{i=1}^N w^i_k\, X^i_k$

   Go to the prediction step with $(k = k + 1)$

$\mathbb{P}_1(X_n|\mathbf{Y}_{n-1})$ which allows us to apply the following Monte Carlo integration,

$$\mathbb{P}_1(\mathbf{Y}_n|\mathbf{Y}_{1:n-1}) = \int_{\mathbb{R}^d} \mathbb{P}_1(\mathbf{Y}_n|\mathbf{Y}_{n-1}, X_n)\, \mathbb{P}_1(X_n|\mathbf{Y}_{n-1})\, dX_n \approx \sum_{i=1}^N w^i_{n-1}\; g_n(X^i_{n|n-1})$$

where g_n is the likelihood (7).
The distribution $\mathbb{P}_0(\mathbf{Y}_n|\mathbf{Y}_{n-1})$ is simply the distribution of the observation noise,

$$\mathbb{P}_0(\mathbf{Y}_n|\mathbf{Y}_{n-1}) = \mathbb{P}_0(\mathbf{Y}_n) \propto \prod_{i=1}^m \exp\left(-\frac{1}{2}\mathbf{Y}^i_n R_i^{-1} \mathbf{Y}^i_n\right)$$

We can then estimate Z_k (10) and evaluate $\hat{K}$, the empirical detection instant (13), i.e. the first time such that the empirical estimation of Z_k exceeds the threshold $\log(\Gamma_1)$.

C. Experimental data

We evaluate the relevance of the lower bound $\tilde{K}^*$ (33) by comparing it with the empirical instant detections provided by the LPF. The images are composed of 3 bands with different SNR and different PSF. We consider the case of an isotropic PSF different for each band (21),

$$\Sigma_i = \begin{bmatrix} \sigma^2_{i,psf} & 0 \\ 0 & \sigma^2_{i,psf} \end{bmatrix} \tag{39}$$

We have performed 100 Monte Carlo trials of the LPF for different global multispectral SNR ($\sum_{i=1}^m \frac{A^2}{\sigma_i^2}$). The noise Std σ_i and the PSF are described below. The tuning parameter μ is such that the global SNR varies from 0 dB to 10 dB.

Scenario data

- Number of spectral bands: $m = 3$
  Image size: $M = 100 \times 100$
  Sampling period : $\Delta T = 0.1\, s$
- Std of the PSF for the 3 bands:
  $\sigma_{1,psf} = 1.5, \sigma_{2,psf} = 2, \sigma_{3,psf} = 2.5$ (cells)
- Target amplitude: $A = 1$
  Covariance noise matrices: $R_i = diag(\sigma_i^2)$
  Noise Std for the 3 bands:
  $\sigma_1 = 1 \times \mu, \sigma_2 = 1.5 \times \mu, , \sigma_3 = 2 \times \mu$
- Number of particles: $N = 10000$
  Maximal number of local MAP: $n_c = 10$
  Threshold redistribution: $N_{th} = \frac{2}{500} N$
- Initial location of the target in the image:
  $x_0 = 28$, $y_0 = 35$ (cells)
  Velocities: $\dot{x} = 0.37$, $\dot{y} = -0.12$ (cell/s)
- Detection probability: $P_d = 0.99$
  False alarm probability: $P_{fa} = 10^{-6}$
- Initial uncertainty of the target position: one particle per resolution cell of the image
  Initial velocities uncertainty: $\sigma_{\dot{x}} = \sigma_{\dot{y}} = 2$ (cell/s)
  Initial amplitude uncertainty: $A \in [0, 5]$

For each Monte Carlo trial i, the particle filter provides the detection instant $\hat{K}^i$ (see section V-B). This instant is the first time such that the random walk Z_k exceeds the threshold (13). In the presence of a target, Fig. 1 shows a few random walks that exceed the threshold $\log(\Gamma_1)$ (10) for global SNR=6 dB.

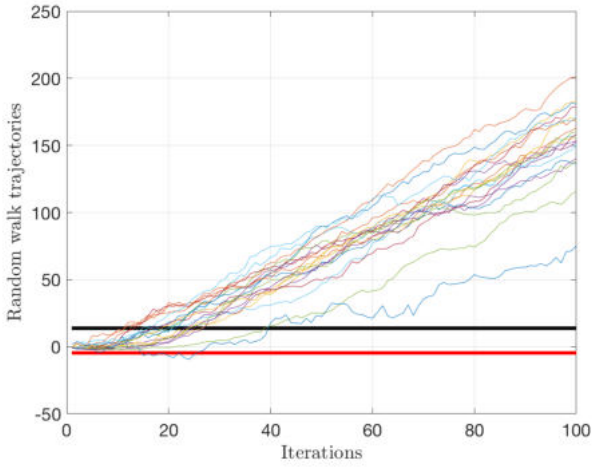

Fig. 1. Random walks Z_k provided by the particle filter. SNR = 6 dB. (–): threshold $\log(\Gamma_0)$, (–): threshold $\log(\Gamma_1)$

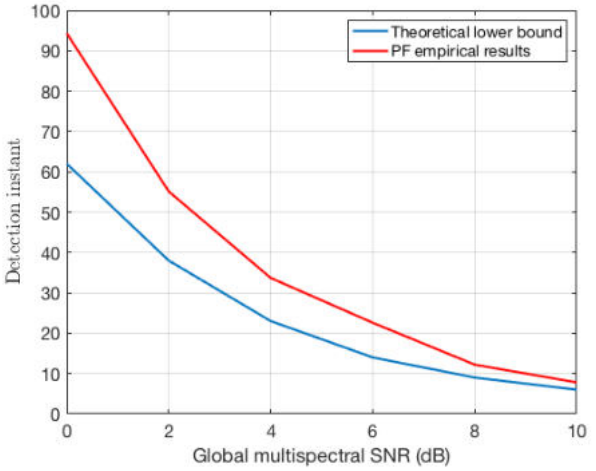

Fig. 2. Comparison of the lower bound $\tilde{K}^*$ with the empirical mean detection instant $\hat{K}$ provided by the particle filter.

By averaging $\hat{K}^i$, for each SNR, we derive the empirical mean detection to be compared with the lower bound $\tilde{K}^*$ (33). Fig. 2 compares the empirical detection instants with the theoretical lower bound $\tilde{K}^*$. The empirical results are consistent with the proposed lower bound. Performance detection is fairly predicted by the lower bound. However, for low SNR values, this bound is optimistic. This bound can be refined. For example, we will attempt to estimate the expectation $\mathbb{E}_1(\xi_n)$, thus avoiding majorizing the jump of the random walk (32). On the other hand, the particle filter used may not be optimal. These questions will be the subject of future work.

VI. Conclusion

A new method was proposed for estimating the minimum number of multispectral images required to detect a target, regardless of the detection method used. This lower bound is calculated from an estimate of the stopping time of the optimal sequential ratio test in the case of non-stationary observations. Simulations of a particle filter for track-before-detect show that detection performance is fairly well predicted by the proposed bound. Further work will aim to improve this bound.

References

[1] Negin Pourmomtaz and Manoochehr Nahvi, "Multispectral particle filter tracking using adaptive decision-based fusion of visible and thermal sequences," *Multimedia Tools and Applications*, vol. 79, no. 25, pp. 18405–18434, 2020.

[2] Zijian Cui and Honglin Ji, "Target tracking based on staring multispectral imager," in *2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA)*. IEEE, 2022, pp. 700–704.

[3] Olga Duran, Efstathios Onasoglou, and Maria Petrou, "Fusion of kalman filter and anomaly detection for multispectral and hyperspectral target tracking," in *2009 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2009, vol. 4, pp. IV–753.

[4] Yvo Boers and JN Driessen, "Multitarget particle filter track before detect application," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 151, no. 6, pp. 351–357, 2004.

[5] Mark G Rutten, Branko Ristic, and Neil J Gordon, "A comparison of particle filters for recursive track-before-detect," in *2005 7th International Conference on Information Fusion*. IEEE, 2005, vol. 1, pp. 7–pp.

[6] Samuel J Davey, Mark G Rutten, and Brian Cheung, "A comparison of detection performance for several track-before-detect algorithms," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, pp. 428036, 2007.

[7] Yaxin Gong, Hongwen Yang, Weidong Hu, and Wenxian Yu, "An efficient particle filter based distributed track-before-detect algorithm for weak targets," in *Radar Conference, 2009 IET International*. IET, 2009, pp. 1–6.

[8] Haichao Jiang, Wei Yi, Guolong Cui, Lingjiang Kong, and Xiaobo Yang, "Knowledge-based track-before-detect strategies for fluctuating targets in k-distributed clutter," *IEEE Sensors Journal*, vol. 16, no. 19, pp. 7124–7132, 2016.

[9] M Morelande and B Ristic, "Signal-to-noise ratio threshold effect in track before detect," *IET radar, sonar & navigation*, vol. 3, no. 6, pp. 601–608, 2009.

[10] DJ Salmond and H Birch, "A particle filter for track-before-detect," in *Proceedings of the American control conference*, 2001, vol. 5, pp. 3755–3760.

[11] Abraham Wald, "Sequential tests of statistical hypotheses," *The annals of mathematical statistics*, vol. 16, no. 2, pp. 117–186, 1945.

[12] Abraham Wald and Jacob Wolfowitz, "Optimum character of the sequential probability ratio test," *The Annals of Mathematical Statistics*, pp. 326–339, 1948.

[13] Alexandre Lepoutre, Olivier Rabaste, and François Le Gland, "A particle filter for target arrival detection and tracking in track-before-detect," in *Sensor Data Fusion: Trends, Solutions, Applications (SDF), 2012 Workshop on*. IEEE, 2012, pp. 13–18.

[14] AA Borovkov, "Mathematical statistics: a textbook," *Science*, 1984.

[15] Yong Liu and Steven D Blostein, "Optimality of the sequential probability ratio test for nonstationary observations," *IEEE Transactions on Information Theory*, vol. 38, no. 1, pp. 177–182, 1992.

[16] Harald Cramer, "Mathematical methods of statistics," *Princeton U. Press, Princeton*, p. 500, 1946.

[17] Petr Tichavsky, Carlos H Muravchik, and Arye Nehorai, "Posterior cramér-rao bounds for discrete-time nonlinear filtering," *IEEE Transactions on signal processing*, vol. 46, no. 5, pp. 1386–1396, 1998.

[18] M Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on signal processing*, vol. 50, no. 2, pp. 174–188, 2002.

[19] Arnaud Doucet and Adam M Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," *Handbook of nonlinear filtering*, vol. 12, no. 656-704, pp. 3, 2009.

[20] Christian Musso, Frédéric Champagnat, and Olivier Rabaste, "Improvement of the laplace-based particle filter for track-before-detect," in *2016 19th International Conference on Information Fusion (FUSION)*. IEEE, 2016, pp. 1095–1102.