CLUTTER ADAPTIVE MULTIFRAME DETECTION/TRACKING OF RANDOM SIGNATURE TARGETS

Marcelo G. S. Bruno

Electrical Engineering Depart., University of São Paulo
P.O. Box 61548, São Paulo SP 05424-970, Brazil
ph:(55-11) 3818-5290; email:bruno@lcs.poli.usp.br

José M. F. Moura

ECE Department, Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213, USA
ph: (412) 268-6341; email: moura@ece.cmu.edu

ABSTRACT

This paper develops the two-dimensional (2D) clutter *adaptive*, multiframe Bayes detector/tracker for targets with random signature. We model the background clutter and the target signature as samples of two independent, spatially correlated, 2D noncausal Gauss-Markov random fields (GMrfs). The target's motion is modeled by a 2D hidden Markov model (HMM). We study, through Monte Carlo simulations, the performance of the adaptive multiframe detector/tracker, and show that the performance of the adaptive tracker is very close to the performance of the tracker when the clutter model is perfectly known.

1. INTRODUCTION

The problem of automatic detection and tracking of moving targets in clutter has received increased attention in recent years, see e.g. [1, 2]. However, most references found in the literature [3] propose the suboptimal association of a single frame detector and a linearized tracking filter, or, alternatively, consider multiframe detection-only (no tracking) of moving targets [4]. By contrast, we developed in [5] a Bayes algorithm that optimally integrates detection and target position estimation. This detector/ tracker is a 2D nonlinear, recursive, multiframe, spatio-temporal algorithm that fully incorporates the models for target signature, target motion, and background clutter.

The algorithm in [5] assumed that the target signature was deterministic and perfectly known to the detector/tracker. In this paper, we extend this algorithm in two directions: random target signature modeled as a spatially correlated random field arising from stochastic fluctuations in the target's reflectivity and conditions of illumination; and clutter adaptation when the clutter statistical model parameters are unknown and need to be estimated. We detail the structure of the detector/ tracker, and evaluate its performance. Our studies show that there is little performance degradation of the adaptive multiframe detector/tracker versus the detector/tracker that has full knowledge of the clutter model.

We comment briefly on the remaining sections of the paper. Section 2 establishes the 2D target and clutter models, expanding on [5]. Section 3 describes the 2D nonlinear Bayes detector/tracker with random target signature. Section 4 details the structure of the detector/tracker, shows how the algorithm combines the statistical information from the target signature and clutter models, and discusses implementation issues. Section 5 examines the algorithm's performance and, finally, section 6 summarizes the contributions of the paper.

The first author's work was supported by FAPESP, Brazil.

2. REVIEW OF THE MODEL

We consider a 2D rigid target whose clutter-free image is contained in a bounded rectangular region of size $(r_i + r_s + 1) \times (l_i + l_s + 1)$. In this notation, r_i and r_s denote the maximum vertical pixel distances in the target image when we move away, respectively up and down, from the target centroid. Analogously, l_i and l_s are the maximum horizontal pixel distances in the target image when we move away, respectively left and right, from the target centroid. An imaging sensor periodically scans the surveillance space generating an L x M sensor image $\mathbf{Y}_n$ that contains the returns corresponding to the target plus the returns from the background clutter. The unknown quantity at each frame n is the target's centroid position, z_n, defined on an equivalent 1D extended lattice [5],

$$\widetilde{\mathcal{L}} = \{l:1 \leq l \leq (L + r_i + r_s)\,(M + l_i + l_s) + 1\} \quad (1)$$

where $L_1 = (L+r_i+r_s)\,(M+l_i+l_s)+1$ is a dummy target state that is used to indicate that the target is absent from the surveillance space at a given frame. The motion of a target that is present is described on the lattice $\widetilde{\mathcal{L}}$ by a first-order hidden Markov model (HMM) specified by the transition probabilities $T(i,j) = P(z_n = i \mid z_{n-1} = j)$, $(i, j) \in \widetilde{L} \times \widetilde{L}$.

The clutter-free pixel intensity of an observed target is described by a set of spatial signature parameters, $\{A_n(k, l)\}$, $-r_i \leq k \leq r_s$, $-l_i \leq l \leq l_s$. Let $\overline{A}_n(k, l) = E[A_n(k, l)]$ where $E[.]$ stands for the expected value. We model the spatial target signature as

$$A_n(k, l) = \overline{A}_n(k,l) + \phi_n(k, l) \tag{2}$$

where $\phi_n(k, l)$ is a zero-mean, noncausal, spatially homogeneous, 2D Gauss-Markov random field (GMrf) model [6] that describes the target's spatial correlation. The field $\{\phi_n(k, l)\}$ is generated by the 2D finite difference equation

$$\begin{aligned}\phi_n(k, l) &= \beta_v^\phi\left[\phi_n(k-1, l) + \phi_n(k+1, l)\right] \\ &+ \beta_h^\phi\left[\phi_n(k, l-1) + \phi_n(k, l+1)\right] + \varepsilon_n(k, l)\end{aligned} \tag{3}$$

where $\varepsilon_n(k, l)$ is a zero-mean Gaussian prediction error such that $E[\phi(k, l)\,\varepsilon(p, r)] = \sigma_\phi^2\delta(k-p, l-r)$, with δ denoting the 2D delta Kronecker function. Analogously, the clutter returns at frame n, $V_n(i, j)$, $1 \leq i \leq L$, $1 \leq j \leq M$, are described by the 2D zero-mean GMrf model

$$\begin{aligned}V_n(i, j) &= \beta_v^c\left[V_n(i-1, j) + V_n(i+1, j)\right] \\ &+ \beta_h^c\left[V_n(i, j-1) + V_n(i, j+1)\right] + U_n(i, j)\end{aligned} \tag{4}$$

where $E[V_n(i, j)\,U_n(p, r)] = \sigma_c^2\,\delta(i-p, j-r)$. The assumption of zero-mean clutter implies a pre-processing of the data that subtracts the mean of the background. A nonzero clutter mean could be accounted for trivially. We use Dirichlet (zero) boundary conditions to specify equations (3) and (4) near the borders of the respective lattice. We also assume that $L >> r_i + r_s$ and $M >> l_i + l_s$.

3. THE ALGORITHM

Define the *vec* operator that converts a P x Q matrix into a PQ x 1 long vector by sequentially stacking the rows of the matrix, and let $\mathbf{y}_n = vec(\mathbf{Y}_n)$, where $\mathbf{Y}_n$ is the observed L x M sensor frame at instant n. Define also the vector $\mathbf{a}_n = vec(\mathbf{A}_n)$ where $A_n(k, l)$, $-r_i \leq k \leq r_s$, $-l_i \leq l \leq l_s$, are the target's clutter-free spatial signature coefficients at frame n, see section 2.

The optimal multiframe Bayes detector/tracker involves the recursive computation at each frame n of the posterior probability mass function $P(z_n \mid \mathbf{Y}_0^n)$, where $z_n \in \widetilde{\mathcal{L}}$ is the unknown target state at frame n and $\mathbf{Y}_0^n = \{\mathbf{y}_0, \ldots, \mathbf{y}_n\}$ is the collection of all observations from instant zero up to instant n. The algorithm is divided into 4 steps.

Filtering Step Using Bayes law and combining assumptions that the sequence of clutter frames $\{\mathbf{V}_k\}$, $k \geq 0$, and the sequence of targets signatures $\{\mathbf{a}_k\}$ are both independent, identically distributed (i.i.d) in time, mutually independent of each other, and independent of the sequence of real target states $\{z_k\}$, $k \geq 0$, it can be shown that, for $n \geq 0$, away from the borders of the sensor lattice, we have

$$\begin{aligned}P(z_n \mid \mathbf{Y}_0^n) &= C_n\left[\int p(\mathbf{y}_n \mid \mathbf{a}_n, z_n)p(\mathbf{a}_n)\, d\mathbf{a}_n\right] \\ &\times P(z_n \mid \mathbf{Y}_0^{n-1})\end{aligned} \tag{5}$$

where C_n is a normalization constant. The recursion is initialized with $P(z_0 \mid \mathbf{Y}_0^{-1}) = P(z_0)$, i.e., the prior distribution of the initial state. Introducing the vectors $\mathbf{p}_{n|n}$ and $\mathbf{p}_{n|n-1}$ such that $p_{n|r}(l) = P(z_n = l \mid \mathbf{Y}_0^r)$ and defining the observations kernel, $\mathbf{S}_n$, such that

$$S_n(l) = \int p(\mathbf{y}_n \mid \mathbf{a}_n,\, z_n = l)p(\mathbf{a}_n)d\mathbf{a}_n\ , \tag{6}$$

we can rewrite equation (5) in matrix notation as

$$\mathbf{p}_{n|n} = C_n\mathbf{S}_n \odot \mathbf{p}_{n|n-1} \tag{7}$$

where the symbol $\odot$ denotes pointwise product.

Remark Equation (5) is still valid near the boundaries of the sensor lattice provided that the target signature vector $\mathbf{a}_n$ is properly defined to take into account portions of the target that fall outside the surveillance space and are no longer visible in the sensor image.

On the other hand, for the absent target state $z_n = L_1$, the filtering step reduces to

$$P(L_1 \mid \mathbf{Y}_0^n) = C_n p(\mathbf{y}_n \mid z_n = L_1)\, P(L_1 \mid \mathbf{Y}_0^{n-1}) \tag{8}$$

where, in this particular case, the kernel $p(\mathbf{y}_n \mid z_n = L_1)$ reduces to the clutter statistics, $p(\mathbf{V}_n)$.

Prediction Step Using the total probability theorem and the same assumptions made before on the independence of $\{z_k\}$, $\{\mathbf{V}_k\}$, and $\{\mathbf{a}_k\}$, we write for $n \geq 1$

$$P(z_n \mid \mathbf{Y}_0^{n-1}) = \sum_{z_{n-1}} P(z_n \mid z_{n-1})P(z_{n-1} \mid \mathbf{Y}_0^{n-1})\ . \tag{9}$$

Recalling from section 2 the transition probability matrix $T(i, j) = P(z_n = i \mid z_{n-1} = j)$, equation (9) is rewritten in matrix format as

$$\mathbf{p}_{n|n-1} = \mathbf{T}\,\mathbf{p}_{n-1|n-1}\ . \tag{10}$$

Detection Step Recalling that $z_n = L_1 = (L + r_i + r_s) \times (M + l_i + l_s) + 1$ indicates that the target is absent at frame n, then the minimum probability of error Bayes detector is the test

$$P(z_n = L_1 \mid \mathbf{Y}_0^n) \underset{H_1}{\overset{H_0}{\gtrless}} 1 - P(z_n = L_1 \mid \mathbf{Y}_0^n) \tag{11}$$

where H_1 denotes the hypothesis "target is present", and H_0 denotes the hypothesis "target is absent."
Tracking Step If hypothesis H_1 is declared true, compute the conditional probability vector

$$Q_l^f[n] = \frac{P(z_n = l \mid \mathbf{Y}_0^n)}{1 - P(z_n = L_1 \mid \mathbf{Y}_0^n)}, \qquad l \in \overline{\mathcal{L}} \tag{12}$$

where $\overline{\mathcal{L}}$ denotes the extended lattice $\widetilde{\mathcal{L}}$ minus the dummy state L_1. The MAP estimate of target's centroid position is

$$\hat{z}\text{map}[n] = \arg\max_{l \in \overline{\mathcal{L}}} Q_l^f[n] \ . \tag{13}$$

4. IMPLEMENTATION: FILTERING STEP

The implementation of the filtering step in equation (5) requires the computation of the observations kernel $\mathbf{S}_n$ defined in (6). Let $\bar{r} = r_i + r_s + 1$ and $\bar{l} = l_i + l_s + 1$ and introduce the symmetric, block-Toeplitz, block-tridiagonal matrices, $\mathbf{A}_c$ and $\mathbf{A}_\phi$ of size $\bar{r}\,\bar{l} \times \bar{r}\,\bar{l}$ such that

$$\mathbf{A}_{c,\phi} = \mathbf{I}_{\bar{r}} \otimes (\mathbf{I}_{\bar{l}} - \beta_h^{c,\phi}\mathbf{H}_{\bar{l}}) - \beta_v^{c,\phi}\,\mathbf{H}_{\bar{r}} \otimes \mathbf{I}_{\bar{l}} \tag{14}$$

where $\otimes$ is the Kronecker or tensor product, $\mathbf{I}_p$ is the p x p identity matrix, and $\mathbf{H}_l$ is an l x l matrix such that $H_l(i,j) = 1$ if $\mid i - j \mid = 1$ or zero otherwise. We also define the prediction error image at frame n, $\Upsilon_n(i,j)$, such that

$$\begin{aligned}\Upsilon_n(i,j) &= Y_n(i,j) - \beta_h^c\,[Y_n(i,j-1) + Y_n(i,j+1)] \\ &- \beta_v^c\,[Y_n(i+1,j) + Y_n(i-1,j)]\end{aligned} \tag{15}$$

with zero (null) boundary conditions. For $l = (p-1)(M+l_i+l_s)+q$, $r_i+r_s+1 \le p \le L$, $l_i+l_s+1 \le q \le M$, introduce now the $\bar{r}\,\bar{l}$ x 1 vector ν_n^l such that

$$\nu_n^l = vec\,[\Upsilon(p - r_i - r_s \colon p, q - l_i - l_s \colon q)] \ . \tag{16}$$

Away from the borders of the sensor grid and using the GMrf models for target and clutter from section 2, it is possible to show that, after absorbing constants that are independent of the state $z_n = l$, and using appropriate normalization factors, then

$$S_n(l) = \exp\left[\frac{(\mathbf{m}_n^l)^T\,\boldsymbol{\Sigma}\,(\mathbf{m}_n^l)}{2}\right] \tag{17}$$

where

$$\begin{aligned}\boldsymbol{\Sigma} &= (\sigma_c^{-2}\,\mathbf{A}_c + \sigma_\phi^{-2}\,\mathbf{A}_\phi)^{-1} &(18)\\ \mathbf{m}_n^l &= \sigma_c^{-2}\,\nu_n^l + \sigma_\phi^{-2}\,\mathbf{A}_\phi\mathbf{m}_a \ . &(19)\end{aligned}$$

In (19), $\mathbf{m}_a = E\,[\mathbf{a}_n]$, which, without loss of generality, we make independent of n. Using the results in [6], we note that the matrix $\boldsymbol{\Sigma}$ in (18) has a very well-defined eigenstructure, factoring as $\boldsymbol{\Sigma} = \boldsymbol{\Psi}\mathbf{D}^{-1}\boldsymbol{\Psi}^T$ where where $\boldsymbol{\Psi} = \mathbf{S}_{\bar{r}} \otimes \mathbf{S}_{\bar{l}}$, with $\mathbf{S}_p$ denoting the p x p 1D orthogonal discrete sine transform (DST) matrix such that, for $0 \le k, n \le p-1$,

$$S(k+1, n+1) = \sqrt{\frac{2}{p+1}}\sin\frac{\pi(k+1)(n+1)}{p+1}. \tag{20}$$

On the other hand, $\mathbf{D}$ is an $\bar{r}\,\bar{l}$ x $\bar{r}\,\bar{l}$ diagonal matrix whose nonzero diagonal entries are given, for $1 \le i \le \bar{r}$ and $1 \le j \le \bar{l}$, by

$$d\left[(i-1)\,\bar{l} + j\right] = c_1 - 2c_2\cos(\frac{\pi\,j}{\bar{l}+1}) - 2c_3\cos(\frac{\pi\,i}{\bar{r}+1}) \tag{21}$$

where $c_1 = \sigma_c^{-2} + \sigma_\phi^{-2}$, $c_2 = \sigma_c^{-2}\beta_h^c + \sigma_\phi^{-2}\beta_h^\phi$, and $c_3 = \sigma_c^{-2}\,\beta_v^c + \sigma_\phi^{-2}\,\beta_v^\phi$. The special eigenstructure of matrix $\boldsymbol{\Sigma}$ indicates that the quadratic form $(\mathbf{m}_n^l)^T\boldsymbol{\Sigma}(\mathbf{m}_n^l)$ in (17) can be computed efficiently by applying a 2D fast discrete sine transform to the matrix equivalent of the vector $\mathbf{m}_n^l$ and then adding the squares of each transform component using as weights the inverse of the nonzero entries of matrix $\mathbf{D}$ given by equation (21) and computed off-line.

Remark The general structure of the algorithm remains the same for the states $z_n = l$ that are located at the boundaries of the sensor lattice, except that the definitions of Υ_n, $\mathbf{A}_c$, and $\mathbf{A}_\phi$ must be modified to account for portions of the target that no longer fall inside the sensor grid. Some correction factors must be also introduced in the analytical expression for $S_n(l)$ to guarantee that all entries in the vector $\mathbf{S}_n$ have a consistent normalization. Similarly, a separate expression has to be derived for the entry $S_n(L_1)$ that corresponds to the absent target state. We omit further discussion on these technicalities for lack of space.

5. TRACKING PERFORMANCE

We present next tracking performance results for the proposed algorithm. We simulate 2D Gauss-Markov targets with correlation parameters $\beta_v^\phi = \beta_h^\phi = 0.16$ and an average pixel intensity $m_a = 1$ that is invariant in both space and time. The sensor image consists of the target image added to synthetic zero-mean Gauss-Markov clutter with correlation parameters $\beta_v^c = \beta_h^c = 0.24$. The simulated targets have size 9 x 9 and move inside a 120 x 120 sensor grid according to a 2D random walk model superimposed to a time-invariant, deterministic drift. The nominal target velocity in both dimensions is equal to 2 pixels/frame and the probability of fluctuation of one pixel around the nominal position in one of the two dimensions is equal to 20 %.

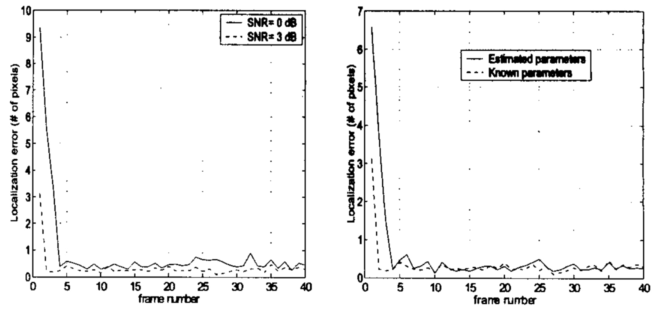
A simulated target departs from an *unknown*, random location in the 30 x 30 upper corner of the sensor image and is subsequently tracked over 40 consecutive frames.

Figure 1 shows the standard deviation of the position estimation error in the vertical dimension (measured in number of pixels) versus the frame number in the image sequence. The error curves for the horizontal dimension are qualitatively similar and are omitted here for lack of space. In Figure 1(a), the tracker has perfect knowledge of the parameters in the target and clutter models. The dashed line corresponds to a clutter standard deviation $\sigma_c = 0.7$, or average SNR$= 20 \log\,(m_a/\sigma_c) = 3$ dB. The solid line corresponds to $\sigma_c = 1$, or SNR$=$ 0 dB. In both plots of Figure 1(b), σ_c is equal to 0.7. The dashed line in Figure 1(b) also assumes perfect knowledge of the clutter parameters, while in the solid line the clutter parameters β_h^c, β_v^c and σ_c are estimated from the data at each frame. These parameters are estimated using a variation of the parameter estimation algorithm introduced in [7]. We refer the reader to that paper for further details.

All curves in Figure 1 are obtained by 50 Monte Carlo runs using a target standard deviation $\sigma_\phi = 0.2$. The four plots in Figure 1 show a similar behavior: there is an initial position error that is rapidly recovered as the tracker processes additional frames and the errors converge to a small steady state value. A comparison between the solid line and the dashed line in Figure 1(b) indicates that the error in the estimation of the clutter parameters causes an increase in the initial position estimate error and in the target acquisition time (i.e., the number of frames that are necessary for the error to decline to the steady state value). However, there is no statistically significant difference between the steady state behavior of the error with known clutter parameters or with estimated clutter parameters. In both cases, near-perfect tracking is achieved as the number of frames increases. This is a quite surprising result.

6. SUMMARY

We presented in this paper a new clutter *adaptive* Bayesian algorithm for integrated, *multiframe* detection and tracking of random signature targets that move randomly in 2D cluttered environments. The random components of the target and clutter returns are modeled by 2D *spatially correlated*, noncausal GMrfs. The motion of the target's centroid in the sensor grid is modeled by a 2D HMM. Experimental results show good tracking performance in scenarios of heavily cluttered targets. Our simulations also show that the long-term behavior of the proposed algorithm is similar in both cases when the clutter model parameters are perfectly known or, in the adaptive case, when the clutter model parameters are estimated from the data.



(a) (b)

Figure 1: Tracking performance, $\sigma_\phi = 0.2$: (a) $\sigma_c = 0.7$ (dashed), $\sigma_c = 1$ (solid); (b) $\sigma_c = 0.7$, estimated clutter parameters (solid), known clutter parameters (dashed)

7. REFERENCES

[1] H. Im and T. Kim, "Optimization of multiframe target detection schemes", *IEEE Transactions on Aerospace and Electronic Systems*, vol.35, n.1, pp 176-187, January 1999.

[2] R. Bethel and G. Paras, "A PDF multisensor, multitarget tracker," *IEEE Transactions on Aerospace and Electronic Systems*, vol.34, n.1, pp 153-168, January 1998.

[3] Y. Bar-Shalom and X. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. YBS, Storrs, CT, 1995.

[4] S. C. Pohlig, "An algorithm for detection of moving optical targets," *IEEE Transactions on Aerospace an Electronic Systems*, vol. 25, pp 56-63, January 1989.

[5] M. G. S. Bruno and J. M. F. Moura, "The optimal 2D multiframe detector/tracker," *AEÜ International Journal of Electronics and Communications*, 53(6), pp 346-355, December 1999.

[6] J. M. F. Moura and M. G. S. Bruno, "DCT/DST and Gauss-Markov fields: conditions for equivalence," *IEEE Transactions on Signal Processing*, vol. 46, n. 9, pp 2571-2574, September 1998.

[7] J. M. F. Moura and N. Balram, "Noncausal Gauss-Markov random fields: parameter structure and estimation," *IEEE Transactions on Information Theory*, IT-39(4), pp 1333-1355, July 1993.