OPTIMIZATION OF THE OBSERVER MOTION USING DYNAMIC PROGRAMMING.

J.P. LE CADRE, Olivier TRÉMOIS.

IRISA/CNRS, Campus de Beaulieu, 35042 RENNES cedex

ABSTRACT

Classical bearings-only target motion analysis (TMA) are restricted to constant motion parameters (usually : position and velocity). However most of the interesting sources have maneuvering abilities, degrading, thus, dramatically the TMA performances. A basic idea consists in modelling the states of the source by an hidden Markov model (HMM for the sequel). The main point is then to optimize the observer trajectory using methods derived from the general theory of dynamic programming.

1. INTRODUCTION.

The basic problem of target motion analysis (TMA) is to estimate the trajectory of an object (i.e. position and velocity for a rectilinear movement) from noise corrupted sensor data. However, for numerous practical applications and especially for long time scenarios, the source is maneuvering. A first approch consists to detect the source maneuvers. This is interesting for the detection of abrupt changes and is efficient when the data have the right statistical properties. These required properties include a correct modelling of the source maneuvers as well as a sufficient signal to noise ratio.

Since none of these requirements is generally valid in the passive sonar context, this advocates for the modelling of the whole source trajectory including maneuver uncertainity. For that purpose, a promising framework is those of hidden Markov model, widely used in other contexts like speech processing [6], frequency line tracking [8] and recently appeared in the active sonar context [3, 4]. In order to apply them to TMA, a basic idea consists in a (two-level) discretization of the state-space (position and velocity). Obviously, in the TMA context, the source state is only partially observed though noisy measurements (the estimated bearings).

An estimation of the source trajectory is then estimated by maximizing a likelihood functional. This task is devoted to a classical Viterbi algorithm [8]. This approach is an elegant solution to the maneuvering target tracking problem since it does not require any prior information on the maneuvers, so that its performance does not depend on some accurate criterion that hardly occurs in real scenarios. If the estimation of the source state is rather classical, the main problem comes from the optimization of the observer

trajectory. This problem is immersed in the general framework of dynamic programming. A main difficulty arises from the partial observation of the source state. This leads us to model the problem by means of POMDP (Partially Observable Markov Decision Process) for which efficient solutions exist [7]. But, once again, another type of difficulties emerges due to the number of states and decisions (observer maneuvers in our context).

2. BEARINGS ONLY TMA GENERAL FRAMEWORK.

Consider the source-observer encounter depicted in figure 1. The source located at the coordinates (r_{xs}, r_{ys}) moves with a constant velocity (v_{xs}, v_{ys}). The state vectors of the source and the observer are [5]:

$$\mathbf{X}_s \triangleq [r_{xs}, r_{ys}, v_{xs}, v_{ys}]^* \quad \mathbf{X}_o \triangleq [r_{xo}, r_{yo}, v_{xo}, v_{yo}]^*. \quad (1)$$

where the symbol '*' denotes transposition. In terms of relative state vector X, defined by $\mathbf{X} = \mathbf{X}_s - \mathbf{X}_o \triangleq [r_x, r_y, v_x, v_y]^*$, the discrete time equation takes the following form:

$$\mathbf{X}(t_k) = \Phi(t_k, t_{k-1})\mathbf{X}(t_{k-1}) + \mathbf{U}(t_k), \quad (2)$$

where:

$$\Phi(t_k, t_{k-1}) = \begin{pmatrix} \mathrm{Id} & (t_k - t_{k-1}))\mathrm{Id} \\ 0 & \mathrm{Id} \end{pmatrix}, \quad \mathrm{Id} \triangleq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In the above formula t_k is the time at the k-th sample while
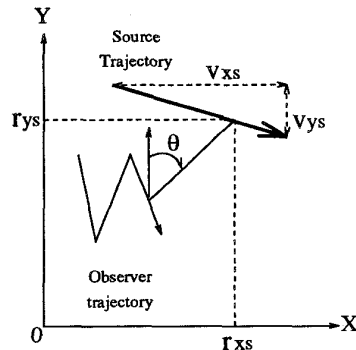

Figure 1: Source-Observer encounter.

the vector $U(t_k)$ accounts for the effects of the observer accelerations.

In the bearings-only tracking (BOT for the sequel) context, the measurements are the estimated source bearings $\hat{\theta}_t$ relatively to the north axis :

$$\hat{\theta}_t = \theta_t + \omega_t \tag{3}$$

$$\text{with} : \theta_t = \tan^{-1}\frac{r_x(t)}{r_y(t)} \text{ and } \omega_t \text{ AWGN } \mathcal{N}(0, \sigma_t^2). \tag{4}$$

The variance σ_t^2 of the estimation noise depends on the relative positions of the source and the observer (array), as well as on the array axis. It is given by the Woodward's formula [2].

If the rectilinear and uniform motion assumption is made, the BOT problem reduces to the estimation of the initial state vector X_0. Even if the main difficulty comes from the non linear nature of the estimation problem, convenient solutions have been presented and analysed in the reference paper [5].

Assume now that the source can change its velocity. In fact interesting sources are those which maneuver for tactical reasons. The decisive advantage of Markov modelling is that one do not have to define different types of maneuvers with the risk that the source does not follow any of them. It is just assumed that the source velocity is not radically changed between consecutive instants. More precisely a two level Markov chain is considered in order to model the source trajectory. Let E_t be the source sate, with :

$$E_t = [x_t, y_t, v_{xt}, v_{yt}] = [X_t, V_t].$$

Using the elementary lemma :

$$\Pr(A, B|C) = \Pr(A|B, C)\Pr(B|C),$$

it comes :

$$\Pr(E_{t+1}|E_t) = \Pr(X_{t+1}|V_{t+1}, X_t, V_t)\Pr(V_{t+1}|X_t, V_t).$$

It is furthermore, quite reasonable to assume that the transition over the velocities is independent from the transitions over position space, so that :

$$\Pr(E_{t+1}|E_t) = \Pr(X_{t+1}|X_t, V_t)\Pr(V_{t+1}|V_t). \tag{5}$$

The Markov chain modelling the source trajectory and defined by (5) may be considered as a two-level one. A first level is devoted to the transition upon the position while the second one deals with the velocity. It is stressed that this model is quite general. The observations are defined by the following equation (Baye's formula) :

$$\Pr(\hat{\theta} = \theta_j|E = E_i) = \frac{\Pr(E = E_i|\hat{\theta} = \theta_j)\Pr(\hat{\theta} = \theta_j)}{\Pr(E = E_i)}. \tag{6}$$

Note that in the above equation the space of the observations (i.e. bearings) is discretized. Practically, this discretization corresponds to the beamwidth [2]. In the abscence of any prior information about the source state or the measurements $\hat{\theta}$, it is worth to consider them as constant, so that :

$$\Pr(\hat{\theta} = \theta_j|E = E_i) = \text{Cst}.\Pr(E = E_i|\hat{\theta} = \theta_j), \tag{7}$$

where 'Cst' is a normalizing factor.

The calculation of the probability $\Pr(E = E_i|\theta = \theta_j)$ is achieved by considering the position cells (x, y) associated with the line of sight $(\theta = \theta_i)$ and its neighborhood. It is now necessary to consider the statistical criteria.

3. STRUCTURE OF THE FISHER INFORMATION MATRIX (FIM).

The partial derivatives of θ_t w.r.t. X_0 the initial state vector (eq. 2) are easily obtained (rectilinear and uniform motion), yielding the following gradient vector G_t :

$$G_t = \begin{vmatrix} \frac{\partial \theta_t}{\partial r_x(0)} &= \frac{\cos \theta_t}{r_t} \\ \frac{\partial \theta_t}{\partial r_y(0)} &= -\frac{\sin \theta_t}{r_t} \\ \frac{\partial \theta_t}{\partial v_x} &= \frac{t \cos \theta_t}{r_t} \\ \frac{\partial \theta_t}{\partial v_y} &= -\frac{t \sin \theta_t}{r_t} \end{vmatrix}. \tag{8}$$

In this case, the FIM is straightforwardly deduced from (8), giving :

$$F_t = \sum_{k=1}^{t} G_k G_k^* = \begin{pmatrix} \sum_{k=1}^{t} \Omega_k & \sum_{k=1}^{t} k\Omega_k \\ \sum_{k=1}^{t} k\Omega_k & \sum_{k=1}^{t} k^2 \Omega_k \end{pmatrix}, \tag{9}$$

with :

$$\Omega_k = \frac{1}{r_k^2 \sigma_{\theta_k}^2} \begin{pmatrix} \cos^2 \theta_k & -\sin \theta_k \cos \theta_k \\ -\sin \theta_k \cos \theta_k & \sin^2 \theta_k \end{pmatrix}. \tag{10}$$

The general structure (9) of the FIM is quite remarkable. It may be easily extended to the case of a maneuvering source. For that purpose, the source trajectory is modelled as a multileg one, which is quite coherent with our discrete modelling (5) of the source trajectory. Consequently, the dimension of the state vector is enlarged since it includes now the initial position of the source as well as its various velocity vectors. To be more precise, consider a source trajectory constituted of n legs, each one corresponding to J bearings, then the FIM (denoted $F_{1,nJ}$) takes the following form :

$$F_{1,nJ} = F_{1,J} + F_{J+1,2J} + \cdots + F_{(n-1)J,nJ}$$

$$= \sum_{k=1}^{J} \left[D_{0,n+1}(k) D_{0,n+1}^*(k) \right] \otimes \Omega_k$$

$$+ \sum_{k=J+1}^{2J} \left[D_{1,n+1}(k) D_{1,n+1}^*(k) \right] \otimes \Omega_k$$

$$+ \cdots$$

$$+ \sum_{k=(n-1)J+1}^{nJ} \left[D_{n-1,n+1}(k) D_{n-1,n+1}^*(k) \right] \otimes \Omega_k$$

with :

$$D_{p,q}(k) = (\underbrace{1 \ \overbrace{J\delta T \cdots J\delta T}^{p} \ (k-pJ)\delta T \ \overbrace{0 \cdots 0}^{q-p-2}}_{q})^* \tag{11}$$

Ω_k given by (10), $\otimes$: Kronecker product.

4. MARKOV DECISION PROCESS (MDP), GENERAL PRINCIPLES.

The general aim of the MDP is to determine a sequence of decisions, generally denoted d (here : observer maneuvers)

which maximizes a criterion related to the state i of the Markov chain. If the process is in state i at time t and an action d is chosen, then two things occur [1] :

1. A cost $C(i, d)$ is incurred,
2. The next state of the system is chosen according to the transition probability $P_{ij}(d)$.

If we let X_t denote the state of the process at time t, and d_t the decision chosen at t, then assumption (2) is equivalent to stating that :

$$\Pr[X_{t+1} = j | X_0, d_0, \dots, X_t = i, d_t = d] = P_{ij}(d). \quad (12)$$

Thus, both the elementary costs $C(i, d)$ and the transition probabilities $P_{ij}(d)$ are functions only of the last state and the subsequent decision. It is easily shown that if a stationary policy (i.e. π) is employed, then the sequence of states $\{X_T, t = 0, 1, 2, \dots\}$ forms a Markov chain with transition probabilities $P_{ij} \triangleq P_{ij}(\pi(i))$, giving thus the denomination MDP to the process.

The problem is to find the control policy $\pi = (\pi_0, \pi_1, \dots)$ such that the sequence of decisions $d_k = \pi_k(X_k, d_{k-1})$ minimizes a cost functional J defined classically by :

$$J(U) = \underset{\pi}{\mathrm{E}} \left\{ \sum_{k=0}^{N-1} C_{X_k, X_{k+1}}(U_k) \right\}. \quad (13)$$

The dynamic programming equation takes the general form :

$$J^*(k, i) = \min_{d \in C} \sum_{j \in S} [C_{i,j}(d) + J^*(k+1, j)] P_{ij}(d) \quad i \in S. \quad (14)$$

The value of the control which gives the minimum is the optimal decision rule at time k if the state is i at this instant. This equation is solved backward in time with the terminal condition $J^*(N, i) = 0 \quad (i \in S)$.

This is the classical approach of dynamic programming. However, our problem may differ from it by the nature of the cost functional. To be more precise, the cost functional (13) is replaced by the following one :

$$J(U) = H \left[\underset{\pi}{\mathrm{E}} \left\{ \sum_{k=0}^{N-1} F_t(i, d_t, j) \right\} \right], \quad (15)$$

where :

- H is a matrix functional,
- $F_t(i, d_t, j)$ is the instantaneous FIM associated with the transition from state i to state j with the decision d_t.

The choice of the functional H is very critical. Actually, the principle of dynamic programming optimization (14) can be applied if the functional H satisfies the following monotony property :

Definition : Let A and B be two positive definite matrices and C positive semi-definite. Then the monotony property holds iff. :

$$H(A) < H(B) \Rightarrow H(A + C) < H(B + C). \quad (16)$$

Assume now that H is furthermore a differentiable functional on $\mathcal{M}_n^+$, the following property holds :

Property : If H is differentiable and satisfies (16), then :

$$H(A) = g(\mathrm{Tr}(AM)), \quad (17)$$

where g is a monotone (real) function, Tr is the matrix trace and M is a fixed matrix.

Actually it is obvious than the determinant functional does not satisfies (16). This seriously advocates for the use of the trace functional, despite the fundamental role of the determinant in information theory. Using (11) the trace of F_t takes the following simple form :

$$\mathrm{Tr}[F_t(i, d_t, j)] = \sum_{\tau=0}^{N_{sam}-1} \frac{1 + (l-1)J^2 \delta T^2 + (t + \frac{\tau}{N_{sam}} - (l-1)J)^2 \delta T^2}{r^2(t + \frac{\tau}{N_{sam}}) \sigma^2_{\theta(t + \frac{\tau}{N_{sam}})}} \quad (18)$$

with :

- δT : time interval between two consecutive steps of the MDP,
- l : index of the source leg,
- J : number of steps associated with a leg,
- N_{sam} : number of samplings during a tansition.

In the case of complete information (the source state is available) the observer trajectory may be optimized by using the DP algorithm (14) with a functional given by (18). For each time and system state, a decision is chosen. These computations have to be done backward in time. The decisions consist of the observer velocity changes. The following decision table (see fig. 2) has been obtained for the two last steps ($t = 19$, $t = 20$) of a 20 step scenario. The observer is always on the center of the position grid, and on each grid node there is an arrow representing the optimal decision (change of observer velocity) when the source is on this node.

However, in practical situations, the source state is not directly observable. The only available informations consist of estimated bearings and the MDP problem then becomes far more complicated. Actually, it will be immersed in the general framework of Partially Observable Markov Decision Process (POMDP).

5. A POMDP FRAMEWORK.

The central process (the Markov chain) X_t is not directly observable. An observation θ_t (here the bearing) is associated with X_t. Let $\Pi(X) = \{\pi \in \mathbb{R}^n | \pi \geq 0, \sum_{i=1}^n \pi_i = 1\}$ be the set of all the distribution on X and $H_t = \{\pi(1), d_1, \dots, \pi(t-1), d_{t-1}\}$ the "history" of the decisions and observations up to time t. We consider, the, the following system evolution :

1. state transition : $p_{ij}^d \triangleq \Pr[X_{t+1} = j | X_t = i, d_t = d]$,
2. observation probability : $r_{j\theta}^d \triangleq \Pr[\theta_t = \theta | X_{t+1} = j, d_t = d]$,
3. $H_{t+1} = H_t \cup \{d_t, \theta_t\}$,
4. elementary cost : $w_{ij\theta}^d$ associated with the following event : under the decision d the state goes from i to j and and observation θ is produced.
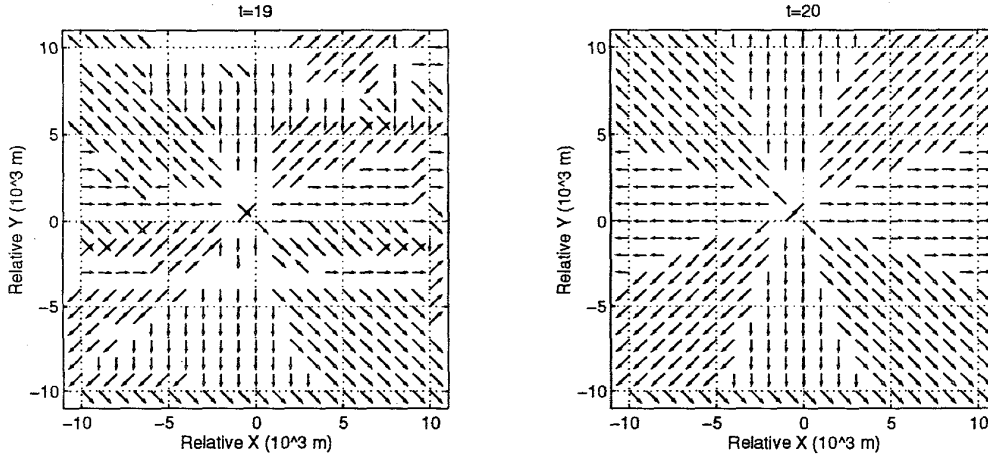
3569

Figure 2: Commands for the two last steps of a scenario : $(v_{xo}, v_{yo}, v_{xs}, v_{ys}) = (0, 0, 10, -10)$.

The abscence of knowledge of the state leads to replace it by the information vector $\pi(t)$ defined by :

$$\pi(t) = (\pi_1(t), \ldots, \pi_n(t))^*, \quad \pi_i(t) \triangleq \Pr[X_t = i | H_t]. \quad (19)$$

Actually it may be shown that $\{\pi(t)\}$ is itself a Markov process. Moreover, for a decision d and an observation θ the following updating equation holds [7] :

$$\pi_j(t+1) = \frac{\sum_i \pi_i(t) p_{ij}^d r_{j\theta}^d}{\sum_{i,j} \pi_i(t) p_{ij}^d r_{j\theta}^d} = T(\pi(t)|d, \theta). \quad (20)$$

Then the DP algorithm stands as follows :

$$V_t(\pi) = \max_{d \in \mathcal{D}_t} \left[\sum_{i=1}^n \pi_i(t) \sum_{j=1}^n p_{ij}^d \sum_\theta r_{j\theta}^d \left\{ w_{ij\theta}^d + V_{t+1}\left(T(\pi|d, \theta)\right)\right\} \right]. \quad (21)$$

Using (20) and (21) the following fundamental result has been obtained by Smallwood and Sondik [7] :

$$V_t(\pi) = \max_k < \alpha^k(t), \pi >, \quad (22)$$

where $<,>$ represents a scalar product and the vectors $\alpha^k(t)$ are calculated by a recursion derived from (21). Eq. (22) represents the key for practical implementation of the POMDP in our context.

6. CONCLUSION

HMM constitutes a promising answer to maneuvering target tracking, especially for passive sonar applications. The problem of observer trajectory optimization appears fundamental for HMM BOT. A natural framework is the POMDP one.

7. REFERENCES

[1] Dimitri P. Bertsekas. *Dynamic programming: deterministic and stochastic models.* Prentice-Hall, Inc., 1987.

[2] W.S. Burdic. *Underwater acoustic system analysis.* Signal Processing series, Prentice-Hall, Englewood cliffs, 2nd edition, 1991.

[3] F. Martinerie and P. Forster. Data association and tracking from distributed sensors using hidden markov models and dynamic programming. *ICASSP*, March 1992.

[4] F. Martinerie and P. Forster. Data association and tracking from distributed sensors using hidden markov models and evidential reasoning. *CDC*, December 1992.

[5] Steven C. Nardone, Allen G. Lindgren, and Kai F. Gong. Fundamental properties and performance of conventional bearings-only target motion analysis. *IEEE Transactions on Automatic Control*, 29(9):775–787, September 1984.

[6] L. R. Rabiner and B. H. Juang. An introduction to hidden markov models. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 4, January 1986.

[7] Richard D. Smallwood and Edward J. Sondik. The optimal control of partially observable markov processes over a finite horizon. *Operations Research*, 21:1071–1088, 1973.

[8] Roy L. Streit and Ross F. Barrett. Frequency line tracking using hidden markov models. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(4):586, April 1990.

3570