Tracking on Intensity-Modulated Sensor Data Streams*

Michael J. Walsh
Marcus L. Graham
Roy L. Streit
Tod E. Luginbuhl
Lauren A. Mathews
Naval Undersea Warfare Center Division
Newport, Rhode Island USA
401-832-4155, 3846, 4189, 8241, 7026
walshmj,grahamml,streitrl,luginbuhlte,mathewsla@npt.nuwc.navy.mil

Abstract—Conventional trackers are point trackers. Tracking energy on a field of sensor cells requires windowing, thresholding, and interpolating to arrive at data points to feed the tracker. This scheme poses problems when tracking energy that is distributed across many cells. Such signals are sometimes termed "over-resolved." It has been suggested that tracking could be improved by decreasing the resolution of the signal processor, so that the cells are large enough to encompass the bulk of the energy, and better match the point tracker assumptions. Larger arrays provide greater resolution at lower frequencies, with the potential for improved detection and classification performance, but in direct conflict with tracking "over-resolved" signals.

These issues are addressed by the histogram-based probabilistic multi- hypothesis tracking (PMHT) method discussed in this paper, which provides a means for modeling and tracking signals that may be spread across many sensor cells. This paper will focus on the initial development and testing of this algorithm for one-dimensional sensor data. Elements of the signal model, theory, and algorithm will be presented along with two frequency domain examples.

TABLE OF CONTENTS

1. INTRODUCTION

Tracking an unknown, nonstationary signal against a noisy, nonstationary background on an intensity-modulated sensor display is a difficult problem. Traditional techniques involve thresholding the sensor data and treating the exceedences as point measurements that are subsequently fed to a tracking algorithm. Choosing this threshold is a challenge in itself, and even then it is typically subject to a prescribed probability of detection and probability of false alarm. Linearizing the generalized frequency modulated (GFM) model in Luginbuhl [1, 2] and adding some of the Bayesian structure from Streit [3] results in a new tracking algorithm that uses all of the sensor data and avoids thresholding entirely. The fundamental premise of the proposed tracking algorithm is that losses due to thresholding can be eliminated completely if all of the sensor data are used by the tracking algorithm.

Section 2 describes this new tracking algorithm, referred to herein as histogram-PMHT, or H-PMHT; in particular, the key aspects of its derivation are discussed. A concise statement of the H-PMHT algorithm applied to one-dimensional sensor data is given in section 3. Two frequency domain examples, one involving a simulated linear frequency modulated (FM) chirp and one involving an at-sea recorded bowhead whale

call, are presented in section 4. A summary is given in section 5.

2. HISTOGRAM-PMHT

Signal and Measurement Models

The H-PMHT tracking algorithm is intrinsically a multi-signal tracking algorithm, and is based on a stochastic model of the signals and the noise background. As in [1, 2], it assumes that the signals, clutter, and background noise are described by a discrete mixture of continuous distributions in which each component PDF represents a target, clutter, or the background noise process. Furthermore, each component PDF may be a discrete mixture PDF. The sensor data are indirect realizations of this underlying distribution. The approach taken is first to quantize the real-valued senor data into a "pseudo-histogram," and then to use a multinomial distribution to model the counts in the histogram cells. The cell-level intensities of the sensor data are directly proportional to the cell counts of the histogram. The goal is to fit the underlying mixture distribution to the histogram at each scan; that is, to estimate the parameters of the mixture distribution that maximize the likelihood of the histogram at each scan. The theoretical framework of PMHT (see [4]) is used to "assign" the histogram samples to the mixture components, and to link these mixture distributions across scans with a dynamical signal model.

Expectation-Maximization

The expectation-maximization (EM) method formalized by Dempster, Laird, and Rubin (see [5]) is a powerful method for estimating the parameters of mixture distributions, and is the method used to solve the likelihood equation in H-PMHT. The method of EM is particularly well suited to so-called "missing" data problems; that is, problems in which the parameters of interest are comparatively straightforward to estimate if the observed data set is augmented with certain unobserved data. The basic strategy of EM is to include the missing data as random variables in the likelihood function, take the expectation of the log-likelihood with respect to the missing data conditioned on the observed data, and maximize the resulting expression, termed the auxiliary function. The EM method requires both the specification of the incomplete (observed) data likelihood function, and the complete data (observed data plus missing data) likelihood function. The incomplete- and complete-data likelihood functions for H-PMHT are described in the next two sections.

Incomplete-Data Likelihood Function

Let Z_t denote the sensor-data measurement vector at time t,

$$Z_t = \{z_{t1}, \ldots, z_{tL}\}, \qquad t = 1, \ldots, T, \tag{2.1}$$

where $z_{t\ell}$ is the output of the sensor at time t in display cell ℓ (a sequence of unaveraged, magnitude-squared DFT data versus frequency in the examples). Let $\hbar^2 > 0$ be a specified quantization level, and let

$$N_t = \{n_{t1}, \ldots, n_{tL}\}, \qquad t = 1, \ldots, T, \tag{2.2}$$

denote the quantized measurement vector corresponding to Z_t, where

$$n_{t\ell} = \left\lfloor \frac{z_{t\ell}}{\hbar^2} \right\rfloor \tag{2.3}$$

is the greatest integer less than or equal to $z_{t\ell}/\hbar^2$. Let

$$N_{t\Sigma} = \sum_{\ell=1}^{L} n_{t\ell} \tag{2.4}$$

denote the total cell count (sample size) at time t. Unlike in [1, 2], in [3] the quantized data vector N_t is used as an intermediate variable in the derivation of the H-PMHT algorithm; at an appropriate point in the derivation, the measurement vector Z_t is recovered in the limit as $\hbar \to 0$.

As in [1, 2], the quantized data vector N_t is assumed to have a multinomial distribution consisting of $N_{t\Sigma}$ draws (with replacement) on L "categories" with probabilities

$$\frac{P_\ell(X_t)}{P(X_t)}, \qquad \ell = 1, \ldots, L, \tag{2.5}$$

where

$$P_\ell(X_t) = \int_\ell f(\tau; X_t)\, d\tau, \tag{2.6}$$

and

$$P(X_t) = \sum_{\ell=1}^{L} P_\ell(X_t). \tag{2.7}$$

In equation (2.6), $f(\tau; X_t)$ is the underlying mixture density of the signals and the background noise, and X_t denotes the parameter vector of the mixture density at time t (minus the mixing proportions, which are implicit in all of the likelihood expressions in the sequel). The assumption of a multinomial distribution for the quantized data vector N_t implies that the counts N_t form a histogram with L cells and a sample size of $N_{t\Sigma}$, where the samples are independent and identically distributed with probability density function $f(\tau; X_t)$ (see [6] for more on the multinomial distribution).

Let $N = \{N_1, \ldots, N_T\}$ denote the collection of quantized measurement vectors, and let $X = \{X_1, \ldots, X_T\}$ denote the set of mixture parameters to be estimated. Assuming the scans are independent, the incomplete-data likelihood function for N is given by

$$p_{inc}(N; X) = \prod_{t=1}^{T} p_{inc}(N_t; X_t), \tag{2.8}$$

where

$$p_{inc}(N_t; X_t) = \frac{N_{t\Sigma}!}{n_{t1}! \cdots n_{tL}!} \prod_{\ell=1}^{L} \left[\frac{P_\ell(X_t)}{P(X_t)}\right]^{n_{t\ell}}. \tag{2.9}$$

In [1, 2, 3], a Bayesian model for the mixture parameters X is adopted. If $p_\Xi(X)$ denotes the *a priori* density of X, then the incomplete-data likelihood function is given by

$$p_{inc}(N, X) = p_\Xi(X)\, p_{inc}(N|X), \tag{2.10}$$

where the density $p_{inc}(N|X)$ is essentially identical to the density (2.8), and differs only in its statistical interpretation (Bayesian conditioning versus parametric dependence).

The derivation of a new prior density $p_\Xi(X)$ in [3] is an important contribution and theoretical development in H-PMHT. In short, in [3] it is shown that the prior needs to be sufficiently non-diffuse so that the synthetically generated histogram counts N, which depend on the arbitrary quantization level $\hbar^2$, do not overwhelm the prior as $\hbar^2 \to 0$. Under the usual Markov assumption on the signal states X_t, Bayes Theorem gives

$$p_\Xi(X) = p_{\Xi_0}(X_0) \prod_{t=1}^{T} [p_{\Xi_t|\Xi_{t-1}}(X_t|X_{t-1})]^{N_{t\Sigma}}, \tag{2.11}$$

where the power of $N_{t\Sigma}$ is necessary to account for the artificial abundance of quantized data N_t at each scan.

Complete-Data Likelihood Function

The missing data in H-PMHT are (1) the locations of the samples that make up the pseudo-histogram, and (2) their mixture component assignments; that is, the components in the underlying mixture distribution from which the samples are drawn.

Let $\zeta_{t\ell} = \{\zeta_{t\ell 1}, \ldots, \zeta_{t\ell n_{t\ell}}\}$ denote the locations of the samples within cell ℓ. The random variables in $\zeta_{t\ell}$ are assumed to be independent and identically distributed with probability density function $f(\tau|X_t)/P_\ell(X_t)$. Furthermore, let $\zeta_t = \{\zeta_{t1}, \ldots, \zeta_{tL}\}$ and $\zeta = \{\zeta_1, \ldots, \zeta_T\}$.

The sample density is assumed to be the mixture density

$$f(\tau|X_t) = \sum_{k=0}^{M} \pi_{tk}\, G_k(\tau|X_t), \tag{2.12}$$

where the mixing proportions $\pi_{tk} \geq 0$ for all t and k, $\sum_{k=0}^{M} \pi_{tk} = 1$ for all t, and the $G_k(\tau|X_t)$ are component probability density functions. For component $\pi_{t0}\, G_0(\tau|X_t)$, π_{t0} represents the fraction of the total power due to the background noise, and $G_0(\tau|X_t)$ models the cell-to-cell variation of the background noise. Likewise, for components $\pi_{t1}\, G_1(\tau|X_t), \ldots, \pi_{tM}\, G_M(\tau|X_t)$, π_{tk} is the fraction of the total power due to signal k, and $G_k(\tau|X_t)$ models the variation of signal k from cell to cell. The mixture model (2.12) assumes that the signal power levels may be spread across more than one cell of the sensor display.

Let $K_{t\ell} = \{k_{t\ell 1}, \ldots, k_{t\ell n_{t\ell}}\}$ denote the components of the mixture that generated the missing variables $\zeta_{t\ell} = \{\zeta_{t\ell 1}, \ldots, \zeta_{t\ell n_{t\ell}}\}$. It is assumed that the random variables in $K_{t\ell}$ are independent and identically distributed. Furthermore, let $K_t = \{K_{t1}, \ldots, K_{tL}\}$ and $K = \{K_1, \ldots, K_T\}$.

Including the missing sample locations and their mixture component assignments, the complete-data likelihood function is given by

$$p_{com}(N, \zeta, K, X) = p_\Xi(X) \prod_{t=1}^{T} \prod_{\ell=1}^{L} \prod_{r=1}^{n_{t\ell}} f_{k_{t\ell r}}(\zeta_{t\ell r}|X_t), \tag{2.13}$$

where

$$f_k(\tau|X_t) = \pi_{tk}\, G_k(\tau|X_t); \tag{2.14}$$

that is, the complete-data likelihood function is the product of all the histogram sample densities across all L cells and all T scans, scaled by the prior for the signal mixture parameters.

E-Step

The auxiliary function of the EM method, denoted here by Ψ, is defined as the conditional expectation with respect to the missing data of the logarithm of the complete-data likelihood function given the observed data and the current values of the parameters X, denoted X':

$$\Psi(X, X') = E_{\zeta K}[\log p_{com}(N, \zeta, K, X)\,|\,N, X'], \tag{2.15}$$

where $E_{\zeta K}$ denotes expectation with respect to the missing data. The mechanics of the E-step for H-PMHT are tedious but straightforward, and are well documented in [1, 2, 3]. The

final result in terms of the pseudo-histogram counts N is

$$\Psi(X, X') = \log p_{\Xi_0}(X_0) + \sum_{t=1}^{T} \frac{N_{t\Sigma}}{P(X'_t)} \log p_{\Xi_t|\Xi_{t-1}}(X_t|X_{t-1}) + \sum_{t=1}^{T}\sum_{\ell=1}^{L}\sum_{k=0}^{M} \frac{n_{t\ell}}{P_\ell(X'_t)} \int_\ell f_k(\tau|X'_t) \log f_k(\tau|X_t)\, d\tau. \quad (2.16)$$

In [3] it is shown that by taking the limit

$$\Psi^\sharp = \lim_{\hbar^2 \to 0} \hbar^2 \Psi, \quad (2.17)$$

and using definition (2.3) for the quantization, the EM auxiliary function Ψ can be replaced by $\Psi^\sharp$, a function of the unquantized sensor data $Z = \{Z_1, \ldots, Z_t\}$:

$$\Psi^\sharp(X, X') = \sum_{t=1}^{T} \frac{\|Z_t\|}{P(X'_t)} \log p_{\Xi_t|\Xi_{t-1}}(X_t|X_{t-1}) + \sum_{t=1}^{T}\sum_{\ell=1}^{L}\sum_{k=0}^{M} \frac{z_{t\ell}}{P_\ell(X'_t)} \int_\ell f_k(\tau|X'_t) \log f_k(\tau|X_t)\, d\tau, \quad (2.18)$$

where $\|Z_t\| = \sum_{\ell=1}^{L} z_{t\ell}$ is the L_1-norm of Z_t. This limit is not taken in [1, 2].

M-Step

The objective of the M-step is to maximize the auxiliary function $\Psi^\sharp$ with respect to the unknown signal parameters X. To proceed, application-specific terms in the auxiliary function (2.18) must be defined.

Let

$$X_t = \{x_{t0}, x_{t1}, \ldots, x_{tM}\}, \qquad t = 0, 1, \ldots, T, \quad (2.19)$$

where x_{t0} is the background noise parameter at time t, and x_{tk} is the parameter of signal k at time t. Assuming the signals are independent at all times,

$$p_{\Xi_t|\Xi_{t-1}}(X_t|X_{t-1}) = \prod_{k=0}^{M} p_{\Xi_{tk}|\Xi_{t-1,k}}(x_{tk}|x_{t-1,k}). \quad (2.20)$$

It is readily shown that $\Psi^\sharp$ can be separated into two terms, one involving the unknown mixing proportions $\pi = \{\pi_{tk}\}$, and one involving the unknown signal parameters X,

$$\Psi^\sharp = \sum_{t=1}^{T} \Psi_{t\pi} + \sum_{k=0}^{M} \Psi_{kX}, \quad (2.21)$$

where

$$\Psi_{t\pi} = \sum_{k=0}^{M} \left[\sum_{\ell=1}^{L} \frac{z_{t\ell}}{P_\ell(X'_t)} \int_\ell G_k(\tau|x'_{tk})\, d\tau \right] \pi'_{tk} \log \pi_{tk}, \quad (2.22)$$

and

$$\Psi_{kX} = \sum_{t=1}^{T} \frac{\|Z_t\|}{P(X'_t)} \log p_{\Xi_{tk}|\Xi_{t-1,k}}(x_{tk}|x_{t-1,k}) + \sum_{t=1}^{T}\sum_{\ell=1}^{L} \frac{\pi'_{tk}\, z_{t\ell}}{P_\ell(X'_t)} \int_\ell G_k(\tau|x'_{tk}) \log G(\tau|x_{tk})\, d\tau. \quad (2.23)$$

The updated mixing proportions π_t are obtained at each time t by maximizing with respect to π_t the Lagrangian equation involving $\Psi_{t\pi}$ and the constraint $\pi_{t0} + \pi_{t1} + \cdots + \pi_{tM} = 1$.

For this application, linear Gauss-Markov processes are assumed for the signals, so that for $k = 1, \ldots, M$ and $t = 1, \ldots, T$, the signal process models are given by

$$p_{\Xi_{tk}|\Xi_{t-1,k}}(x_{tk}|x_{t-1,k}) = \mathcal{N}(x_{tk}; F_{t-1,k}\, x_{t-1,k}, Q_{t-1,k}), \quad (2.24)$$

where $\mathcal{N}(\tau; \mu, C)$ denotes the multivariate normal probability density function with mean μ and covariance matrix C, the $F_{t-1,k}$ are known state matrices, and the Q_{tk} are known process covariance matrices.

Additionally, it is assumed that the signal components in the mixture distribution are also Gaussian, and that the means of these Gaussians are linearly related to the states of the signals $k = 1, \ldots, M$ at times $t = 1, \ldots, T$, so that

$$G_k(\tau|x_{tk}) = \mathcal{N}(\tau; H_{tk}\, x_{tk}, R_{tk}), \quad (2.25)$$

where the H_{tk} are known measurement matrices, and the R_{tk} are known measurement covariance matrices.

Finally, the background noise distribution in this application is assumed to be uniform and known for all t, so that the $G_0(\tau|x_{t0})$ terms are all constants.

With these assumptions, it is shown in [3] that for $X(k) = \{x_{0k}, x_{1k}, \ldots, x_{Tk}\}$, the value of $X(k)$ that maximizes the auxiliary function Ψ_{kX} for each signal k is the solution to a symmetric, block-tridiagonal, linear system of equations, and that this system is most efficiently solved by a recursive Kalman smoothing filter. The details of this result are omitted here, but the filter steps are listed explicitly in the next section.

3. ONE-DIMENSIONAL TRACKING ALGORITHM STATEMENT

For one-dimensional tracking, the signal parameters of interest are typically the instantaneous location on the sensor display γ_t (e.g., angle or frequency) and the instantaneous rate $\dot{\gamma}_t$ at time t; therefore, for signal k,

$$x_{tk} = \begin{bmatrix} \gamma_{tk} \\ \dot{\gamma}_{tk} \end{bmatrix}. \tag{3.2}$$

For this two-state linear Markov model, the state matrices $F_{t-1,k}$ and the process covariance matrices $Q_{t-1,k}$ have simple forms:

$$F_{t-1,k} = \begin{bmatrix} 1 & \Delta_{t-1} \\ 0 & 1 \end{bmatrix}, \tag{3.3}$$

where Δ_t is the elapsed time between time t and time $t-1$, and

$$Q_{t-1,k} = q_{t-1,k} \begin{bmatrix} \frac{1}{3}\Delta_{t-1}^3 & \frac{1}{2}\Delta_{t-1}^2 \\ \frac{1}{2}\Delta_{t-1}^2 & \Delta_{t-1} \end{bmatrix}, \tag{3.4}$$

where the $q_{t-1,k}$ are scale factors (see [7, p. 84]).

For two-dimensional time-intensity data, the measurement matrices H_{tk} and the measurement covariance matrices R_{tk} also have simple forms:

$$H_{tk} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \tag{3.5}$$

and

$$R_{tk} = \rho_{tk}^2, \tag{3.6}$$

where ρ_{tk}^2 is the variance of signal k at time t.

Let $\{\pi_{tk}^{(i)}\}$ be the set of estimated mixing proportions and $\{x_{tk}^{(i)}\}$ the set of estimated signal states at the i-th EM iteration. At the beginning of the algorithm (the 0-th iteration), the mixing proportions $\{\pi_{tk}^{(0)}\}$ are initialized so that $\pi_{tk}^{(0)} > 0$ and $\pi_{t0}^{(0)} + \pi_{t1}^{(0)} + \cdots + \pi_{tM}^{(0)} = 1$ for $t = 1, \ldots, T$. The signal state sequences $\{x_{0k}^{(0)}, x_{1k}^{(0)}, \ldots, x_{tM}^{(0)}\}$ are initialized with nominal values for $t = 1, \ldots, T$. For iterations $i = 0, 1, 2, \ldots$, the following nine quantities are computed (see [3] for further details):

1. Component bin probabilities for $t = 1, \ldots, T$, $\ell = 1, \ldots, L$, and $k = 0, 1, \ldots, M$:

$$P_{tk\ell}^{(i+1)} = \begin{cases} 1/L, & k = 0, \\ \int_\ell \mathcal{N}(\tau; H_{tk}\, x_{tk}^{(i)}, R_{tk})\, d\tau, & k = 1, \ldots, M. \end{cases} \tag{3.7}$$

2. Total bin probabilities for $t = 1, \ldots, T$ and $\ell = 1, \ldots, L$:

$$P_{t\ell}^{(i+1)} = \sum_{k=0}^{M} \pi_{tk}^{(i)}\, P_{tk\ell}^{(i+1)}. \tag{3.8}$$

3. Total scan probabilities for $t = 1, \ldots, T$:

$$P_t^{(i+1)} = \sum_{\ell=1}^{L} P_{t\ell}^{(i+1)}. \tag{3.9}$$

4. Bin centroids for $t = 1, \ldots, T$, $\ell = 1, \ldots, L$, and $k = 1, \ldots, M$:

$$\tilde{z}_{tk\ell}^{(i+1)} = \frac{1}{P_{tk\ell}^{(i+1)}} \int_\ell \tau\, \mathcal{N}(\tau; H_{tk}\, x_{tk}^{(i)}, R_{tk})\, d\tau. \tag{3.10}$$

5. Synthetic measurements for $t = 1, \ldots, T$ and $k = 1, \ldots, M$:

$$\tilde{z}_{tk}^{(i+1)} = \frac{\sum_{\ell=1}^{L} \left[z_{t\ell} \left(P_{tk\ell}^{(i+1)} / P_{t\ell}^{(i+1)} \right) \right] \tilde{z}_{tk\ell}^{(i+1)}}{\sum_{\ell=1}^{L} \left[z_{t\ell} \left(P_{tk\ell}^{(i+1)} / P_{t\ell}^{(i+1)} \right) \right]}. \tag{3.11}$$

6. Synthetic measurement covariance matrices for $t = 1, \ldots, T$ and $k = 1, \ldots, M$:

$$\tilde{R}_{tk}^{(i+1)} = \frac{R_{tk}}{\pi_{tk}^{(i)} \sum_{\ell=1}^{L} z_{t\ell}^{(i+1)} \left(P_{tk\ell}^{(i+1)} / P_{t\ell}^{(i+1)} \right)}. \tag{3.12}$$

7. Synthetic process covariance matrices for $t = 0, 1, \ldots, T-1$ and $k = 1, \ldots, M$:

$$\tilde{Q}_{tk}^{(i+1)} = \frac{P_t^{(i+1)}}{\|Z_t\|}\, Q_{tk}. \tag{3.13}$$

8. Estimated mixing proportions for $t = 1, \ldots, T$ and $k = 0, 1, \ldots, M$:

$$\pi_{tk}^{(i+1)} = \frac{\pi_{tk}^{(i)} \sum_{\ell=1}^{L} z_{t\ell}^{(i+1)} \left(P_{tk\ell}^{(i+1)} / P_{t\ell}^{(i+1)} \right)}{\sum_{k'=1}^{M} \pi_{tk'}^{(i)} \sum_{\ell=1}^{L} \left(P_{tk'\ell}^{(i+1)} / P_{t\ell}^{(i+1)} \right)}. \tag{3.14}$$

9. Estimated signal states for $t = 0, 1, \ldots, T$ and $k = 1, \ldots, M$, using (for computational efficiency) a recursive Kalman smoothing filter, which comprises a forward filter initialized at time $t = 0$ with $\tilde{y}_{0|0}^{(i+1)}(k) = x_{0k}^{(0)}$ and a large (diffuse) state covariance matrix $P_{0|0}^{(i+1)}(k)$, and

given, for $t = 0, 1, \ldots, T-1$, by the recursions

$$\begin{aligned}
P_{t+1|t}^{(i+1)}(k) &= F_{tk}\, P_{t|t}^{(i+1)}(k)\, F_{tk}^{*} + Q_{tk}, \\
W_{t+1}^{(i+1)}(k) &= P_{t+1|t}^{(i+1)}(k)\, H_{t+1,k}^{*} \\
&\quad \times \left[H_{t+1,k}\, P_{t+1|t}^{(i+1)}(k)\, H_{t+1,k}^{*} + R_{t+1,k} \right]^{-1}, \\
P_{t+1|t+1}^{(i+1)}(k) &= P_{t+1|t}^{(i+1)}(k) - \\
&\quad W_{t+1}^{(i+1)}(k)\, H_{t+1,k}\, P_{t+1|t}^{(i+1)}(k), \\
\tilde{y}_{t+1|t+1}^{(i+1)}(k) &= F_{tk}\, \tilde{y}_{t|t}^{(i+1)}(k) + W_{t+1}^{(i+1)}(k) \\
&\quad \times \left[\tilde{z}_{t+1,k}^{(i+1)} - H_{t+1,k}\, F_{tk}\, \tilde{y}_{t|t}^{(i+1)}(k) \right],
\end{aligned}$$

and a backward filter initialized at time $t = T$ with $x_{Tk}^{(i+1)} = \tilde{y}_{T|T}^{(i+1)}(k)$ and given, for $t = T-1, T-2, \ldots, 1$, by the recursion

$$\begin{aligned}
x_{tk}^{(i+1)} = \tilde{y}_{t|t}^{(i+1)}(k) + P_{t|t}^{(i+1)}(k)\, F_{tk}^{*} \left(P_{t+1|t}^{(i+1)}(k) \right)^{-1} \\
\times \left[x_{t+1,k}^{(i+1)} - F_{tk}\, \tilde{y}_{t|t}^{(i+1)}(k) \right],
\end{aligned}$$

and, for $t = 0$, by

$$x_{0k}^{(i+1)} = F_{0k}^{-1}\, x_{1k}^{(i+1)}, \tag{3.15}$$

where the asterisk denotes matrix transposition.

The most common convergence tests for termination of the algorithm are based on the rate of increase of the incomplete-data likelihood function. Other tests are based on the rate of change of the estimated parameters. In practice, usually some combination of these two tests is used.

4. EXAMPLES

Linear Chirp

In this section, a two-state H-PMHT frequency tracker is used to track the instantaneous frequency and frequency-rate of a low-frequency linear chirp.

Consider the continuous-time, complex-sinusoidal signal $x(t)$ with constant amplitude A and constant phase ϕ,

$$x(t) = A \exp[i\, 2\pi\gamma(t)\, t + i\, \phi], \tag{4.2}$$

where

$$\gamma(t) = \gamma_0 + \dot{\gamma}_0\, t, \tag{4.3}$$

γ_0 is the nominal frequency in Hz, and $\dot{\gamma}_0$ is the chirp-rate in Hz/s. Converting to radians, equations (4.2) and (4.3) are rewritten as

$$x(t) = A \exp[i\, \Omega(t)\, t + i\, \phi], \tag{4.4}$$

and

$$\Omega(t) = \Omega_0 + \dot{\Omega}_0\, t, \tag{4.5}$$

where $\Omega(t)$ is the instantaneous frequency in rad/s, Ω_0 is the nominal frequency in rad/s, and $\dot{\Omega}_0$ is the chirp-rate in rad/s^2.

Without loss of generality, it is assumed that the signal is observed starting at time $t = 0$, and that estimates of the instantaneous frequency and frequency-rate of $x(t)$ are required every $\mathcal{T}$ seconds. To avoid aliasing, the signal is sampled well above the Nyquist rate for the observation period of interest $S\mathcal{T}$, where S is the number of scans; that is

$$\Omega_s = \frac{2\pi}{\mathcal{T}_s} \gg 2\, \Omega(S\mathcal{T}), \tag{4.6}$$

where Ω_s is the sampling rate in rad/s, and $\mathcal{T}_s$ is the sampling period in seconds. Using n to denote the sampling index, the discrete-time versions of equations (4.4) and (4.5) are given by

$$\begin{aligned}
x(n\mathcal{T}_s) &= A \exp[i\, \Omega(n\mathcal{T}_s)\, n\mathcal{T}_s + i\, \phi] \\
x[n] &= A \exp(i\, \omega[n]\, n + i\, \phi),
\end{aligned} \tag{4.7}$$

and

$$\begin{aligned}
\mathcal{T}_s\, \Omega(n\mathcal{T}_s) &= \mathcal{T}_s\, \Omega_0 + \mathcal{T}_s\, \dot{\Omega}_0\, n\mathcal{T}_s \\
\omega[n] &= \omega_0 + \dot{\omega}_0\, n,
\end{aligned} \tag{4.8}$$

for $n = 0, \ldots, N-1$, where $\omega[n]$ is the instantaneous frequency in rad/sample, ω_0 is the nominal frequency in rad/sample, and $\dot{\omega}_0$ is the chirp-rate in rad/sample^2 (see [8] for details on time-sampling).

For simplicity, it is assumed that the signal $x(t)$ is corrupted by additive, complex, zero-mean white Gaussian noise with variance σ^2; that is, the observed time-series is given by

$$y[n] = x[n] + v[n], \tag{4.9}$$

where

$$E\{v[n]\} = 0, \tag{4.10}$$

$$E\{v[n]\, v[n+k]\} = \sigma^2\, \delta[k], \tag{4.11}$$

for all n and k. An observation interval of length $\mathcal{T}$ seconds and a sampling period of length $\mathcal{T}_s$ seconds yields a time-series of length $\mathcal{T}/\mathcal{T}_s + 1$ samples every $\mathcal{T}$ seconds, for a total of $N = S(\mathcal{T}/\mathcal{T}_s + 1)$ samples.

For each observation interval, the signal-to-noise ratio (SNR) η is defined as the ratio of the average signal power P to the noise power σ_{bin}^2 in a bin,

$$\eta \equiv \frac{P}{\sigma_{bin}^2} \tag{4.12}$$

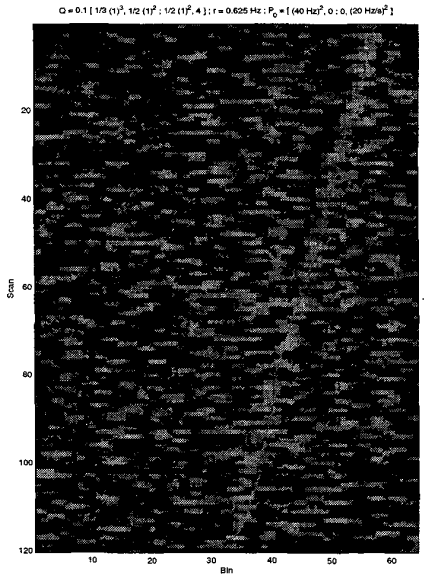


Figure 1. Magnitude-Squared STFT of Linear Chirp in White Gaussian Noise ($\eta = 4$ dB, $\rho = 1$ Bin (0.625 Hz), $q = 1 \times 10^{-3}$) and Estimated Track

For the signal of interest, the average signal power $P = A^2$, so $\eta = A^2/\sigma_{bin}^2$.

In this example, a low-frequency linear FM chirp with amplitude $A = 1$, phase $\phi = 0$, nominal frequency $\gamma_0 = 20$ Hz, and chirp-rate $\dot{\gamma}_0 = 0.125$ Hz/s is sampled every $\mathcal{T}_s = 0.0125$ second, and its short-time Fourier transform (STFT) is taken every $\mathcal{T} = 1$ second for $S = 120$ consecutive, nonoverlapping (unaveraged) blocks of data (scans) to yield a spectrogram of length $S\mathcal{T} = 2$ minutes. This sampling period corresponds to a sampling rate of $\gamma_s = 1/\mathcal{T}_s = 80$ Hz, which is well above the Nyquist rate of $2\,\gamma(S\mathcal{T}) = 2\,[\gamma_0 + \dot{\gamma}_0(S\mathcal{T})] = 70$ Hz for the 2-minute observation period. Each 1-second observation interval yields $\mathcal{T}/\mathcal{T}_s + 1 = 81$ samples, for a total of $N = 9720$ samples over the whole data set. The spectrogram for a signal with $\sigma = 5A$ is shown in Fig. 1.

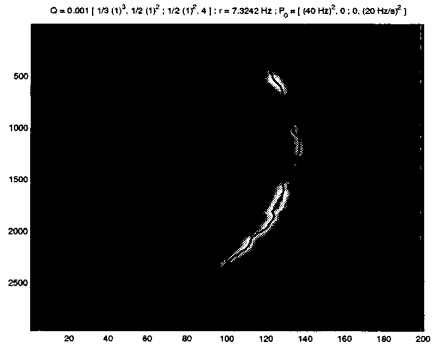
The STFT in this example is computed using a 128-point fast Fourier transform (FFT) and an 81-point Hamming window. The window sequence is normalized such that it sums to one; this scaling allows the magnitude-squared FFT output to be directly interpreted as power. Since for a P-point DFT the spacing between DFT frequencies is $2\pi/P$, and the relationship between the discrete-time frequency ω and the continuous-time frequency Ω is $\omega = \Omega\mathcal{T}_s$, where $\mathcal{T}_s$ is the sampling period, the frequency resolution is given by $2\pi/P\mathcal{T}_s$ rad/s, so the DFT bin-width for this example is $1/(128 \times 0.0125) = 0.625$ Hz. Only the positive half of the spectrum is shown in Fig. 1, and only the first $L = 64$ bins of sensor data are considered in the sequel.

The H-PMHT frequency tracker outlined in section 3 was implemented as a single-scan ($T = 1$), single-signal ($M = 1$) sequential filter and was applied to the data in Fig. 1 on a per-scan basis; that is, the scans were processed sequentially, where the estimate for the previous scan was used to initialize the estimate for the current scan (see [3]).

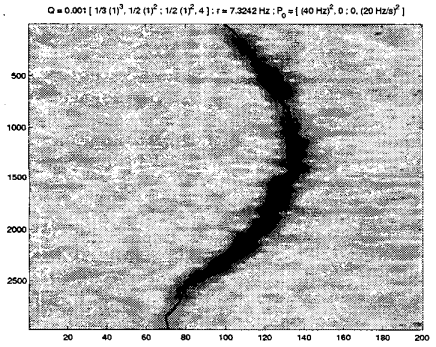
The estimated instantaneous frequencies are shown connected by a solid line the figure. The jaggedness of the track will increase as the SNR drops and the signal power varies more from cell to cell. The jaggedness of the track is also a function of the process noise; the stochastic signal model allows for further excursions from the deterministic, constant-rate model as the process noise scale factor q is increased, resulting in larger random accelerations. There is an intimate relationship between the values of the filter parameters ρ and q, the SNR, the track initialization, and the tracker behavior. High SNR signals will tend to draw the signal component of the mixture closer if the signal is within the effective "gate" determined by ρ. The amount by which the signal component will move to data outside the range of the deterministic process model depends on the SNR, the value of q, and the track's local estimate of its own quality (i.e., the size of the state covariance $P_{t|t}$ on x_t). In low SNR, the "inertia" of the tracker will allow it to coast according to the deterministic process model until there is good reason (i.e., supporting data) for it to change course.

Bowhead Whale Call

The spectrogram of a 1.2-second bowhead whale call recorded at sea is shown in Fig. 2. This signal was sampled at 2500 Hz and processed with a 421-point Chebyshev window and a 1024-point FFT with a bin resolution of 2.4414 Hz. Values of $\rho = 3$ bins (7.3242 Hz) and $q = 1 \times 10^{-3}$ were used to track this signal. The estimated instantaneous frequencies are shown connected by a solid line in Fig. 2. The intent of this example is to show the ability of the tracker to track complex, nonstationary signals, and the ability of the stochastic process model to accommodate dynamics of higher order than the deterministic process model. The benefit of including process noise is the ability to track high dynamics with a reduced parameter set. However, for too low a model order, the burden of tracking the signal dynamics falls squarely on the shoulders of the process noise, and the result for high process noise may be very jagged tracks or, for low processes noise, tracks whose dynamics lag those of the actual track.



(a) Linear-Scale



(b) Logarithmic-Scale

Figure 2. Spectrogram of Bowhead Whale Call

5. CONCLUDING REMARKS

The H-PMHT tracking methodology avoids the thorny issues of thresholding the sensor data to provide measurements for a point tracker by (1) using all of the sensor data, and (2) modeling the signals as potentially distributed over several cells. The H-PMHT signal model is a stochastic model in that the signal centers are modeled as the component means of a discrete mixture distribution; the signal "widths" are modeled by the variances of the mixture components.

The interpretation of the sensor output as a pseudo-histogram plays an important role in the derivation of H-PMHT. The real-valued cell outputs are quantized and treated as the cell counts of a pseudo-histogram whose distribution is multinomial; the underlying density of this multinomial distribution is the mixture density of the signals and the background noise. It is a remarkable fact [3] that the sensor data are recovered in the algorithm as the quantization level $\hbar^2$ is taken to zero.

An important theoretical development in the derivation of H-PMHT not discussed in this paper is the use of a negative multinomial distribution (see [6]) to model sensor cells in which no data are collected. The negative multinomial functions as an interpolator in this case, serving to restore the missing cell counts in the histogram. In this capacity, the negative multinomial may be useful to reduce "edge effects" that may bias estimates when the tails of signal components extend beyond the ends of the sensor display. The reader is referred to [3] for further details on the use of the negative multinomial in H-PMHT.

REFERENCES

[1] T. E. Luginbuhl, "Estimation of General Discrete-Time FM Processes," Ph.D. Thesis, University of Connecticut, Storrs, 1999.

[2] T. E. Luginbuhl and P. Willett, "Tracking a General, Frequency Modulated Signal in Noise," *Proceedings of the* 38^{th} *IEEE Conference on Decision and Control*, December 1999.

[3] R. L. Streit, "Tracking on Intensity-Modulated Data Streams," NUWC-NPT Technical Report 11,221, Naval Undersea Warfare Center Division Newport, Rhode Island, 1 May 2000.

[4] R. L. Streit and T. E. Luginbuhl, "Probabilistic Multi-Hypothesis Tracking," NUWC-NPT Technical Report 10,428, Naval Undersea Warfare Center Division, Newport, Rhode Island, 15 February 1995.

[5] A. P. Dempster, N. M. Laird, and D. B. Rubin,"Maximum Likelihood from Incomplete Data via the EM Algorithm,"

Journal of the Royal Statistical Society, Series B, vol. 39, 1977, pp. 1–38.

[6] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Discrete Multivariate Distributions*, Wiley & Sons, New York, 1997.

[7] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, Academic Press, New York, 1988.

[8] A. V. Oppenheim, R. W. Schafer, with J. R. Buck, *Discrete-Time Signal Processing*, Prentice-Hall, New Jersey, 1999.

Michael J. Walsh *received the B.A. degree in physics from Bard College and the B.S. degree in mechanical engineering from Washington University, both in 1996 as part of a dual-degree program. In the summer of 1994, he was a student fellow at the Woods Hole Oceanographic Institution (WHOI) in the Ocean Systems and Moorings Laboratory, where he worked on hydrodynamic modeling of surface moorings. He has been employed by the Naval Undersea Warfare Center Division, Newport, Rhode Island since 1996, where he works on target tracking applications for the Combat Systems Department. He is also working towards the M.S. degree in electrical engineering at the University of Massachusetts, Dartmouth. His current research interests include applications of estimation theory, probabilistic and statistical modeling, and numerical analysis to target tracking.*

Marcus L. Graham *received the B.S. and M.S. degrees in electrical engineering from Case Western Reserve University in 1983, and the Ph.D. degree in electrical engineering from the University of Rhode Island in 1992. He has been employed by the Naval Undersea Warfare Center Division, Newport, Rhode Island, from 1984 to present, where he works on various aspects of target motion analysis and contact management. His current interests include nonlinear filtering and regularization methods for ill-posed problems.*

Roy L. Streit *received the B.S. degree (Honors) in physics and mathematics in 1968 from East Texas State University, the M.A. degree in mathematics in 1970 from the University of Missouri (Columbia), and the Ph.D. in mathematics in 1978 from the University of Rhode Island. From 1981 to 1982 he was a Visiting Scholar in the Department of Operations Research at Stanford University, Stanford, California, and from 1987 to 1989 he was an Exchange Scientist at the Defence Science and Technology Organisation, Adelaide, Australia. He joined the U. S. Navy Underwater Sound Laboratory in New London, Connecticut, in 1970, and he is now with the Naval Undersea Warfare Center (NUWC) Division in Newport, Rhode Island. His previous research includes diverse investigations supporting NUWC acoustic array R&D programs, applications of hidden Markov models to frequency line detection and tracking, and probabilistic methods for neural network design. His current research interests include distributed and sequential detection problems, multi-target tracking using advanced Bayesian methods, and multi-source acoustic data fusion using statistically validated belief networks.*

Tod E. Luginbuhl *received the B.S. degree in Electrical Engineering in 1985 from the University of Rochester in New York, and the M.S. and Ph.D. degrees in Electrical Engineering in 1991 and 1999 respectively from the University of Connecticut. He has been employed by the Naval Undersea Warfare Center Division, Newport, Rhode Island, since 1985. His previous research includes acoustic data fusion, tracking frequency lines and frequency line families, probabilistic design of neural networks, multi-target tracking, signal detection and signal classification. His current research interests include sequential detection problems, tracking frequency line families, multi-target tracking, statistical signal classification and statistical methods for acoustic data fusion.*

Lauren A. Mathews *received the B.S. degree in 1992 and the M.S. degree in 1999 from the University of Massachusetts, Dartmouth, both in electrical engineering. She has been employed by the Naval Undersea Warfare Center Division, Newport, Rhode Island since 1992. During this time, she has been involved in developing, implementing and testing target tracking algorithms for submarine combat control systems. Her current efforts include the mathematical modeling and development of geometry-based target localization algorithms in a 3-D visualization environment. Her primary interests are estimation theory, modeling and simulation, and real-time implementation.*