Modelling and Tracking of Dynamic Spectra Using a Non-Parametric Bayesian Method

Dragana Carevic

Maritime Division, DSTO, HMAS Stirling, Rockingham, Australia

ABSTRACT

We consider a problem of modelling and tracking dynamic time-frequency spectra where the number of spectral components is allowed to vary in time. The proposed approach involves representing spectral densities as mixtures of normal distributions with an unknown number of components using a time-varying non-parametric Bayesian method. A multi-target tracking algorithm is applied to track the spectral mixture components, where the inference is done using a Rao-Blackwellised particle filter. The algorithm estimates the time-varying mean and variance of each spectral component using a specifically designed Bayesian filter, and allows for the inclusion of outliers or clutter measurements. The spectral features extracted in this way can be applied for signal smoothing and classification, or in conjunction with tracking algorithms to improve data association and filtering.

INTRODUCTION

Nonparametric Bayesian methods avoid often restrictive assumptions of parametric models by performing inference on infinite-dimensional spaces of functions or probability distributions. If suitably designed, these methods allow for efficient, data-driven posterior inference. Escobar and West (1995) introduced the hierarchical Dirichlet process mixture (DPM) model as a Bayesian non-parametric approach applicable to the problems of clustering and density estimation. The DPM uses the Dirichlet process (DP) (Ferguson 1974, Teh 2010) as a prior over the distribution of its parameters, allowing in this way for the uncertainty in its parametric form, as well as in the number of the mixed components.

In this paper a non-parametric mixture model based on the DPM is used to represent spectral densities as mixtures of normal distributions with variable number of components. Using this representation, an approach to tracking variable number of spectral features corresponding to the peaks in the frequency spectrum is presented. These spectral features are taken to be time-varying hidden states, and the estimation is done within a state-space estimation framework. The resulting spectral representation can be used for smoothing and recognition of speech signals (Ozkan, Ozbek & Demirekler 2009) or in conjunction with kinematic measurements with an aim to improve data association and state estimation (Pace, Mallick & Eldredge 2003, Wang, La Scala & Ellem 2008).

Markov chain Monte Carlo (MCMC), and more recently sequential Monte Carlo (SMC) methods, have provided efficient and accurate strategies for inference in the DPMs in batch and sequential processing of data (Neal 2000). A DPM-based Rao-Blakwellised Gibbs sampling scheme is used for clustering action potentials, or spikes, that occur in neurophysiological recordings (Wood & Black 2008) and in non-parametric belief propagation in graphical models (Sudderth 2006). A MCMC algorithm and a Rao-Blackwellised particle filter (RBPF) are proposed for state estimation in dynamic systems with unknown noise probability density (Caron et al. 2008). Recently, a DPM model-based RBPF for tracking unknown number of formants in the speech spectrum is proposed (Ozkan, Ozbek & Demirekler 2009). A drawback of this method is that it estimates only the means of the normal mixture components, and takes that the distribution variances are known and equal for all components. Moreover, it does not take into account the possibility of the presence of outliers in the data.

The approach proposed in this paper is motivated by, and improves on, the methods described above. First, an extension to the DPM is presented that allows for the inclusion of a known distribution (e.g. outlier distribution) into the mixture. Next, a Bayesian filter that simultaneously estimates the kinematic target state and the measurement error variance is appllied. We also present an outline of a non-parametric model-based RBPF that estimates time-varying means and variances of the spectral mixture components. It is conjectured that, with an appropriate choice of prior parameters, the proposed algorithm can be used to model and track both broadband and narrowband spectra. In this paper we present some results obtained by applying the algorithm to simulated and real narrowband spectra.

BAYESIAN NON-PARAMETRIC DENSITY ESTIMATION

The DP is a stochastic process that defines a distribution over probability measures on potentially infinite parameter spaces Θ. It is uniquely defined by two parameters: a base measure G_0 on the parameter space Θ, and a concentration parameter α, and is denoted by $DP(G_0,\alpha)$. It can be shown that realisations from a DP are discrete with probability one. Namely, using the so-called stick-breaking representation (Sethuraman 1994), a random probability measure (RPM) G drawn from a DP, $G \sim DP(G_0,\alpha)$, is equivalent to an infinite sum of atomic measures:

$$G(\theta) = \sum_{j=1}^{\infty} \pi_j \delta_{\theta_j} \qquad \theta_j \sim G_0(\cdot), \quad j = 1,2,\ldots \tag{1}$$

where δ_{θ_j} is a Dirac delta function located at θ_j, and $\pi_j = \beta_j \prod_{l=1}^{j-1}(1-\beta_l)$, $\beta_j \sim \text{Beta}(1,\alpha)$.

The DPs are widely used in statistics in classification and mixture density estimation applications where the number of clusters or the mixture components is unknown a priori. In particular, for an unknown density $F(\cdot)$, we use the following non-parametric model that allows us to estimate F:

$$F(y) = \int_{\Theta} f(y|\theta) dG(\theta) \tag{2}$$

where $\theta \in \Theta$ is a latent variable (or cluster variable), $f(\cdot|\theta)$ is a mixed pdf, and $G(\cdot)$ is a mixing distribution. Under a Bayesian framework, $G(\cdot)$ is taken to be a RPM distributed according to a DP prior distribution, $G \sim DP(G_0,\alpha)$. Substituting (1) in (2) we can formulate F as:

$$F(y) = \sum_{j=1}^{\infty} \pi_j f(y|\theta_j) \tag{3}$$

This DPM model can be equivalently specified by the following equations:

$$G|G_0,\alpha \sim DP(G_0,\alpha), \quad \text{and for } k=1,2,... \tag{4}$$
$$\theta_k|G \sim G(\cdot) \tag{5}$$
$$y_k|\theta_k \sim f(y|\theta_k), \tag{6}$$

The DPM model is a mixture model with a countably infinite number of clusters. However, because π_j's decrease exponentially quickly, only a small number of clusters is used to model the data a priori. The actual number of clusters is not fixed and can be inferred from the data.

Composite Dirichlet Mixture Model

In some cases it may be more appropriate to use a model where the variates can be drawn either from a DPM model or from another specified (known) distribution. This situation occurs, for example, when the data that we want to model by the DPM is contaminated by outliers. Denote by ξ the prior probability that a variate is generated by a DPM. Then this composite DPM (CDPM) model can be defined as:

$$y_k \sim \xi F(\cdot) + (1-\xi)u(\cdot) \tag{7}$$

where F is the DPM model defined in (3) and u is the specified distribution. In the case of an outlier, u is usually modelled as an uniform pdf $u = 1/V$, where V is the volume of the measurement space.

Define by $\lambda_k \in \{0,1\}$ an auxiliary discrete random variable distributed according to a Bernoulli distribution with $\Pr(\lambda_k = 1) = \xi$. If the random draw $\lambda_k = 1$, we sample θ_k from G, where $G \sim DP(G_0,\alpha)$, and if $\lambda_k = 0$, we set $\theta_k = (0,0)$ (without loss of generality it is assumed that $\theta_k \sim G(\cdot)$ is always different from (0,0)). Supposing that the probability ξ is a random variable with a prior distribution $p(\xi) = \text{Beta}(\gamma,\tau)$, where γ and τ are known parameters, an alternative formulation for the CDPM model is then given by

$$G|G_0,\alpha \sim DP(G_0,\alpha), \quad \xi|\gamma,\tau \sim \text{Beta}(\gamma,\tau) \tag{8}$$
$$\lambda_k|\xi \sim B(\xi) \tag{9}$$
$$\theta_k|G,\lambda_k \sim \lambda_k G(\cdot) + (1-\lambda_k)\delta_{0,0} \tag{10}$$
$$y_k|\theta_k \sim f(y|\theta_k), \quad k=1,2,... \tag{11}$$

where B denotes Bernoulli distribution. In (11), if $\theta_k = (0,0)$, $f(y|\theta_k) = u$, otherwise $f(y|\theta_k)$ is one of the mixed pdfs related to the DPM model.

Similarly as the DPM model (Teh 2010), the CDPM model is characterised by the clustering property. Let $\theta_1,\ldots,\theta_n$ be n random i.i.d. samples from $\lambda_k G(\cdot) + (1-\lambda_k)\delta_{0,0}$ in (10) obtained as described above. Among $\theta_{1:n}$, there are θ_j^*, $j = 0,\ldots,m$ different (unique) values (corresponding to clusters). We denote by θ_0^* the dummy parameter representing the specified distribution u, and by θ_j^*, $j = 1,\ldots,m$, the parameters of the m_D clusters related to the DPM model. Next, we define the association variables $c_1,\ldots,c_n$, such that $c_k = j$ if $\theta_k = \theta_j^*$, $j = 0,\ldots,m$. Note that $c_k = 0$ denotes a realisation from u (e.g. an outlier), while $c_k = j$, $j > 0$, are related to the DPM model. The random variable ξ can be integrated out, and the predictive distribution of the association variable c_{n+1} conditional on the previous values $c_{1:n}$ can be defined as a function of the prior parameters α, γ and τ as follows

$$p(c_{n+1}=0|c_{1:n}) = \frac{\tau+n_0}{\gamma+\tau+n}, \quad j=0 \tag{12}$$
$$p(c_{n+1}=j|c_{1:n}) = \frac{\gamma+n_D}{\gamma+\tau+n}\cdot\frac{n_j}{\alpha+n_D}, j\in\{1,\ldots,m\} \tag{13}$$
$$p(c_{n+1}=c_{new}|c_{1:n}) = \frac{\gamma+n_D}{\gamma+\tau+n}\cdot\frac{\alpha}{\alpha+n_D}, j\notin\{0,\ldots,m\} \tag{14}$$

where $n_0 = \sum_{k=1}^{n}\delta_{0,c_k}$ is the number of outliers among n samples, $n_j = \sum_{k=1}^{n}\delta_{j,c_k}$ is the number of repeats of θ_j^*, $n_D = \sum_{j=1}^{m} n_j$ is the total number of samples associated with the DPM model, and $n = n_0 + n_D$.

The choice of the concentration parameter α in (14) is important for the performance of the DPM-based models, since this parameter tunes the number of clusters m_D. For a large n_D (Teh 2010),

$$E(m_D|\alpha,n_D) = \alpha\log(1+n_D/\alpha). \tag{15}$$

So, α controls the number of clusters directly, with larger α implying a larger m_D a priory, while m_D grows only logarithmically with the number of observations.

BAYESIAN TRACKING OF SINGLE TARGET WITH UNKNOWN MEASUREMENT ERROR VARIANCE

We describe a single target tracking algorithm that simultaneously estimates the kinematic target state and the measurement error variance that is assumed unknown. The measurements are taken to be one-dimensional, $y_t \in \mathsf{R}^1$, $t = 1,2,\ldots$, and the state vector is updated after one measurement is received. A method for tracking an extended object or cluster, where the state vector is updated using several simultaneously received measurements and where the measurements are multidimensional, is proposed by Koch (2008).

Let $\mathbf{x}_t$ and R denote, respectively, the state vector and the unknown error variance, or extent, of the measurements to be estimated. The tracking algorithm is derived as a recursive updating scheme for the joint conditional probability density

$$\begin{aligned} p(\mathbf{x}_t,R|Y^t) &= p(\mathbf{x}_t|R,Y^t)p(R|Y^t) \\ &= N(\mathbf{x}_t;\mathbf{x}_{t|t},\mathbf{P}_{t|t}R)IG(R;\nu_t,\Sigma_t) \end{aligned} \tag{16}$$

at each time t, where $N(\cdot;m,C)$ is normal pdf with mean m and covariance matrix C, and $IG(\cdot;a,b)$ is inverse-gamma distribution with shape parameter $a > 0$ and scale parameter $b > 0$. $\mathbf{P}_{t|t}$ is the estimation error covariance, and $Y^t = \{y_i\}_{i=1}^t$ denotes the set of measurements up to time t.

The time dynamics of the system is described using a linear state-space representation defined as follows:

$$\mathbf{x}_t = \mathbf{F}_t\mathbf{x}_{t-1} + \mathbf{v}_t \quad \text{with} \quad \mathbf{v}_t \sim N(\cdot;0,\tilde{\mathbf{Q}}_t) \tag{17}$$
$$y_t = \mathbf{H}_t\mathbf{x}_t + w_t \quad \text{with} \quad w_t \sim N(\cdot;0,R) \tag{18}$$

and

$$\mathbf{x}_0 \sim N(\mathbf{x};\mathbf{x}_{0|0},\mathbf{P}_{0|0}R), \quad R \sim IG(R;\nu_0,\Sigma_0) \tag{19}$$

where $\mathbf{F}_t$ is the state transition matrix, $\mathbf{H}_t$ is the observation matrix, and $\tilde{\mathbf{Q}}_t = \mathbf{Q}_t R$ is the process noise covariance that is controlled by the parameter $\mathbf{Q}_t$ and the variance R.

Similar to the standard Kalman filter, the filtering equations for the state vector $\mathbf{x}_t$ are given by

$$\mathbf{x}_{t|t-1} = \mathbf{F}_t\mathbf{x}_{t-1|t-1} \tag{20}$$
$$\mathbf{P}_{t|t-1} = \mathbf{F}_t\mathbf{P}_{t-1|t-1}\mathbf{F}_t^T + \mathbf{Q}_t \tag{21}$$
$$\mathbf{x}_{t|t} = \mathbf{x}_{t|t-1} + \mathbf{W}_t\left(y_t - \mathbf{H}_t\mathbf{x}_{t|t-1}\right) \tag{22}$$
$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{W}_t S_t \mathbf{W}_t^T \tag{23}$$

where

$$S_t = \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T + 1, \tag{24}$$

is the innovation factor, $\tilde{S}_t = S_t R$ is the variance of the innovation term $(y_t - \mathbf{H}_t\mathbf{x}_{t|t-1})$, and

$$\mathbf{W}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^T S_t^{-1} \tag{25}$$

is the Kalman gain. Note that,

$$p(\mathbf{x}_t|R,Y^t) = N(\mathbf{x}_t;\mathbf{x}_{t|t},\mathbf{P}_{t|t}R)$$

where the actual uncertainty $\tilde{\mathbf{P}}_{t|t} = \mathbf{P}_{t|t}R$ is influenced by R.

The update for the variance R is given by

$$\nu_t = \nu_{t-1} + \frac{1}{2} \quad (26)$$

$$\Sigma_t = \Sigma_{t-1} + \frac{(y_t - \mathbf{H}_t\mathbf{x}_{t|t-1})^2}{2S_t}. \quad (27)$$

For $\nu_t > 1$, the posterior expectation and mode of R can be computed as

$$E(R|Y^t) = \frac{\Sigma_t}{\nu_t - 1}, \quad \text{mode}(R|Y^t) = \frac{\Sigma_t}{\nu_t + 1}. \quad (28)$$

The predictive distribution of a new measurement y_t given the previous measurements up to time $t-1$, Y^{t-1}, is the Student-t distribution with $2\nu_{t-1}$ degrees of freedom and with the parameters μ and σ^2 defined by

$$\mu = \mathbf{H}_t\mathbf{x}_{t|t-1} \quad \text{and} \quad \sigma^2 = \frac{S_t\Sigma_{t-1}}{\nu_{t-1}}, \quad (29)$$

so that,

$$p(y_t|Y^{t-1}) = t_{2\nu_{t-1}}\big(y_t;\, \mathbf{H}_t\mathbf{x}_{t|t-1}, \frac{S_t\Sigma_{t-1}}{\nu_{t-1}}\big). \quad (30)$$

MODELLING AND TRACKING OF DYNAMIC SPECTRA

Our aim is to model frequency spectra and track their dominant features as they evolve in time. The approach is to represent spectral densities as mixtures with an unknown number of components and where the number of components can vary over time. We use the non-parametric Bayesian model based on a DP defined by (8)-(11), where the mixed pdfs of the DPM model are taken to be normal distributions characterised by two parameters, the mean and the variance.

At each time step t we have a set of N_t measurements $y_{k,t}, k = 1,\ldots,N_t$. The measurements are obtained by sampling from the pdf that is proportional to the magnitude of the spectrum; in this way, the spectral magnitude becomes the target density to be approximated by a mixture. The problems to be solved are: 1) the estimation of the number of mixed pdfs, 2) the association of the measurements with the mixed pdfs, and 3) the recursive estimation of the mixture parameters.

We attempt to solve these problems by using a multi-target tracking approach, where the inference is done via a Rao-Blackwellised particle filter motivated by (Sarkka, Vehtari & Lampinen 2007), and where the dependance of the association variables on their past values is given by (12)-(14). The motions of individual mixed pdfs (or targets) are assumed to be independent from all other pdfs. The time dynamics of each component is modelled using a constant velocity linear state-space model based on (17)-(19). The state vector is defined as $\mathbf{x}_{j,t} = [x_{j,t}\ \dot{x}_{j,t}]$, where $x_{j,t}$ and $\dot{x}_{j,t}$ are, respectively, the mean of the jth normal mixed pdf at time t, and its first derivative.

The Rao-Blackwellisation is accomplished by decomposing the joint density of the state vectors $\mathbf{x}_{0:t}$, the variances $\mathbf{R}$ and the association variables $c_{1:t}$ related to the multitarget system up to time t, in the following manner:

$$p(\mathbf{x}_{0:t},\mathbf{R},c_{1:t}|y_{1:t}) = p(\mathbf{x}_{0:t},\mathbf{R}|c_{1:t},y_{1:t})p(c_{1:t}|y_{1:t}) \quad (31)$$

where $y_{1:t}$ denotes the set of all related measurements. Conditioned on the association variables $c_{1:t}$, the sufficient statistics of the parameters of the normal mixed pdfs in $\mathbf{x}_{0:t}$ and $\mathbf{R}$ can be estimated using the Bayesian filter described in the previous section. The marginal density of the association variables $p(c_{1:t}|y_{1:t})$ can be approximated using a particle set $\{\omega_t^{(i)}, c_{1:t}^{(i)}\}_{i=1}^{Q}$, as

$$p(c_{1:t}|y_{1:t}) \approx \sum_{i=1}^{Q} \omega_t^{(i)} \delta_{c_{1:t}^{(i)}}(c_{1:t}) \quad (32)$$

where the weights $\omega_t^{(i)}$ and the variables $c_{1:t}^{(i)}$ depend on $y_{1:t}$. Each particle represents a hypothesis on the association between the measurements and the mixture pdfs. In addition to the weights and the past values of the association variables, a particle also keeps the sufficient statistics of the parametric representation of the mixed pdfs.

Let there be Q particles. We give an outline of the processing related to the ith particle; all other particles are processed in the same way. A detailed description of the algorithm will be presented in a separate document.

At time t we receive N_t measurements, $y_{k,t}$, $k = 1,\ldots,N_t$. The measurements are processed sequentially. For each $y_{t,k}$ an optimal importance distribution is computed. This is a discrete distribution which gives the (normalised) probabilities of relating the variable $c_{k,t}$ to each of the existing mixed pdfs, to a new pdf, and to clutter, conditioned on the current measurement $y_{k,t}$, the previous measurements, and the previous values of the association variables. For each value of $c_{k,t}$ this probability is computed using the appropriate prior in (12)-(14) and a posterior predictive probability of $y_{k,t}$ (given by the general formula in (30), for $c_{k,t}$ related to a pdf, and being equal to u for clutter assignment).

The association of the measurement $y_{k,t}$ is sampled from the importance distribution. If the measurement is assigned to one of the existing mixed pdfs, the sufficient statistics of this pdf are updated using the Bayesian filter described in the previous section. If the association is to the new pdf, its sufficient statistics are similarly updated, and this pdf is added to the list of the established mixed pdfs.

After the measurement $y_{k,t}$ is processed for all Q particles, a weight corresponding to each particle is computed. In this paper, a particle with the highest weight at a time t is used to represent the modelled dynamic spectrum up to that time.

We note that it is possible to model short-term changes in the prior distributions of $c_{k,t}$. Based on this model we can decide which of the mixed pdf ceased to exist, and should be removed from the list. We also use a similar model to estimate nonstationary measurement variances.

EXPERIMENTAL RESULTS

This section presents results obtained by applying the proposed algorithm to simulated and real narrowband spectra. A simulated spectrum is generated in the time domain and is subsequently transformed to the spectral domain using the Fast Fourier Transform (FFT). Each narrowband spectral component at time t is represented as a complex exponential as

$$s_{j,t}(n) = A_{j,t}\exp\left[i(2\pi f_{j,t}n + \phi_{j,t})\right], \quad n = 1,\ldots,N \quad (33)$$

where $f_{j,t}$ is the normalised frequency related to the jth spectral component. $A_{j,t}$ is the random amplitude $A_{j,t} \sim U_{A_j}$, where U_{A_j} is the uniform distribution in the interval $[0, A_j]$, $\phi_{j,t}$ is the random phase $\phi_{j,t} \sim N(0,\sigma_{\phi_j}^2)$, and $i = \sqrt{-1}$. The noise is modelled as a complex white Gaussian process with zero mean and

the variance σ_v^2. The expected average signal-to-noise (SNR) ratio for each spectral component can be estimated as

$$SNR_j = 10\log_{10}\frac{E(A_{j,t}^2)}{2\sigma_v^2} \tag{34}$$

where $E(A_{j,t}^2) = A_j^2/3$.

Fig. 1 shows the original simulated time-frequency spectrum (or spectrogram), where the narrowband spectral components are buried in noise at the average SNR of -10 dB. In this example all spectral components are linear. The spectrum reconstructed using the sufficient statistics of the spectral components corresponding to the particle with the highest weight is shown in Fig. 2, where the total number of particles is $Q = 40$. Fig. 3 shows the true frequencies (tracks) of the spectral components, and Fig. 4 shows the tracked frequencies estimated using the proposed RBPF algorithm corresponding to Fig. 2 (note that the y axis in Figs. 3 and 4 is reversed as compared to the y axis in Figs. 1 and 2). It can be seen that there are several false tracks. The pruning of the false tracks can be done based on their length and their relative energy, which is expected to be much smaller than that of the true tracks.

The real data is recorded using an underwater sensor array. The signal is first beamformed, and the bearings in the beamformed data related to a target are tracked manually. For each bearing/time corresponding to the target detection, the underlying frequency spectrum is extracted. Such spectra are stacked together to form the target time-frequency spectrum. Each spectrum is next processed using the 'order-truncate-average' (OTA) algorithm (Nielsen 1991). This algorithm applies a robust processing based on median filtering, and separates the spectrum into two components: the smoothed (broadband) spectrum, and the narrowband spectrum. In this paper we apply the proposed RBPF algorithm to the resulting narrowband time-frequency spectrum shown in Fig. 5. Fig. 6 shows the reconstructed real data spectrum obtained using the results of the proposed RBPF algorithm. Fig. 7 shows the estimated frequency tracks related to Fig. 6, and Figs. 8 and 9 show the original and the estimated real data spectrum, respectively, plotted in 3-D. Note that the y axis in Fig. 7 is reversed as compared to the y axis in Figs. 5 and 6.

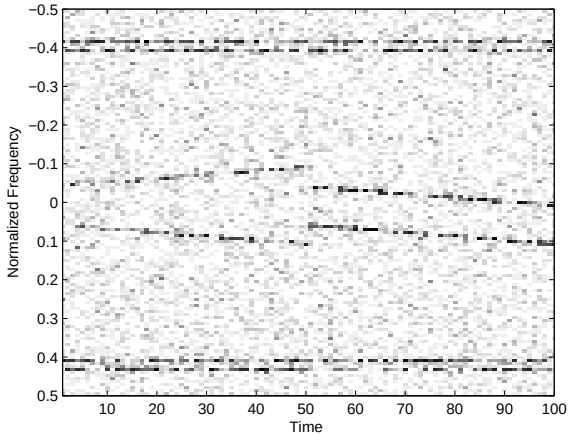

Figure 1: The original simulated time-frequency spectrum.

CONCLUSION

In this paper we proposed a multi-target tracking approach for the estimation and tracking of dynamic spectral components, where the inference is done via a Rao-Blackwellised particle filter. The dependance of the association variables on

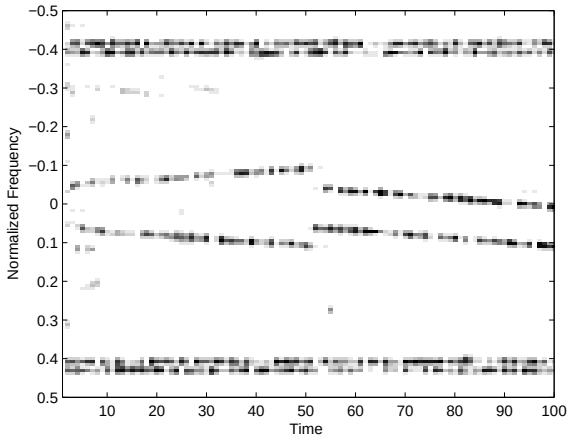

Figure 2: The spectrum in Fig. 1 reconstructed using the sufficient statistics of the spectral components corresponding to the particle with the highest weight.

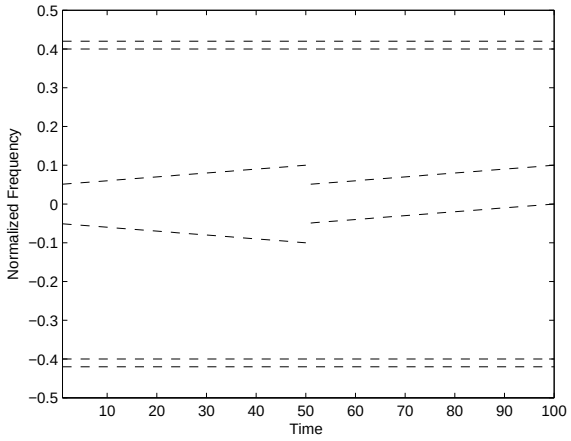

Figure 3: True tracks of the spectral components related to the simulated data in Fig. 1.

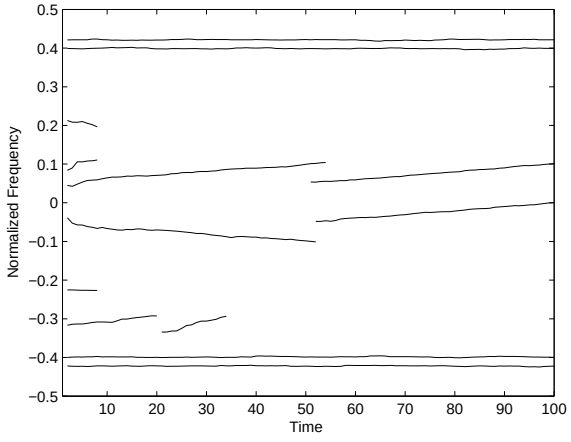

Figure 4: Estimated tracks of the spectral components related to the reconstructed spectrum in Fig. 2.

their past values is modelled using an extension to the standard Dirichlet process model that allows for an outlier distribution to be included in the mixture. The sufficient statistics of the unknown means and variances of the tracked probability density functions are simultaneously estimated using a specifically designed Bayesian filter. The results obtained by applying the algorithm to simulated and real narrowband spectra are also presented.

The results obtained to date are encouraging and show that there is merit in further investigation of this approach. Further work is required to characterise the performance of the algorithm in real scenarios. There is also a need to understand the computational load imposed by this approach compared to the more conventional processing in order to assess the cost effectiveness of using the approach in practical systems.

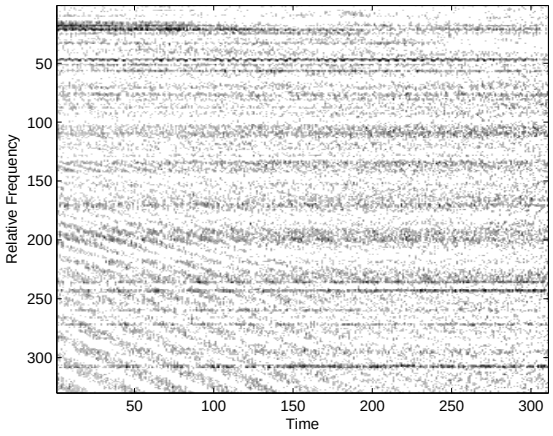

Figure 5: The original time-frequency spectrum obtained using real data.

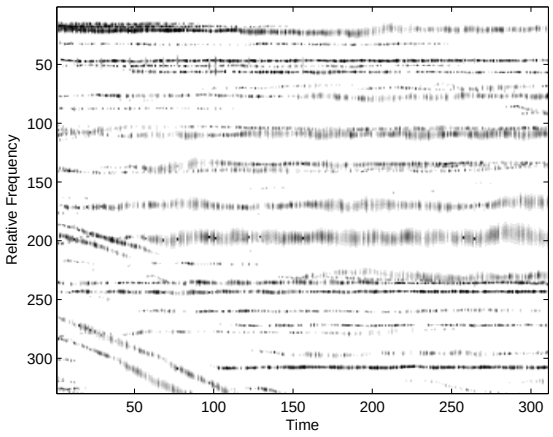

Figure 6: The real data spectrum reconstructed using the sufficient statistics of the spectral components corresponding to the particle with the highest weight.

REFERENCES

Caron, C, Davy, M, Doucet, A, Douflos, E & Vanheeghe, P 2006, 'Bayesian inference for linear dynamic models with Dirichlet process mixtures', *IEEE Trans. Signal Process.*, vol. 56, pp. 71–84.

Escobar, MD & West, M 1995, 'Bayesian density estimation and inference using mixtures', *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 577–588.

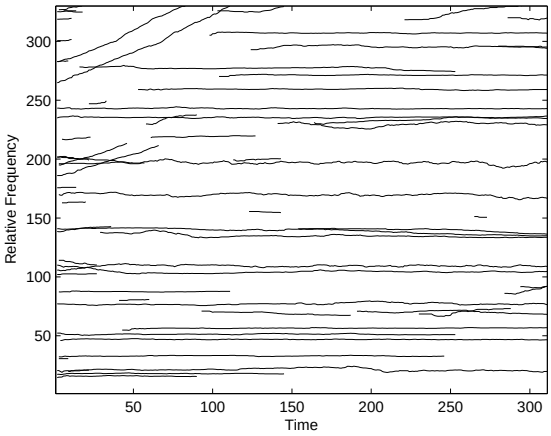

Figure 7: Estimated tracks of the spectral components related to the reconstructed spectrum in Fig. 6.

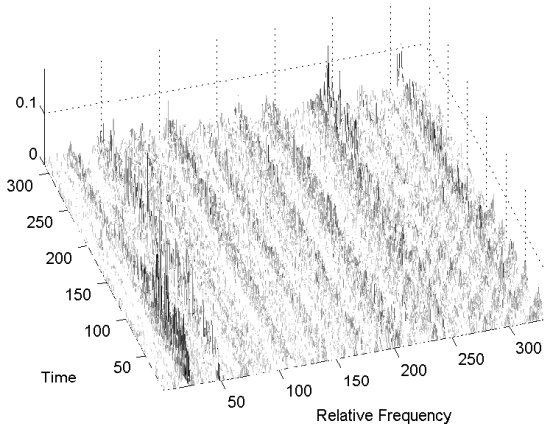

Figure 8: The original real data time-frequency spectrum plotted in 3-D.

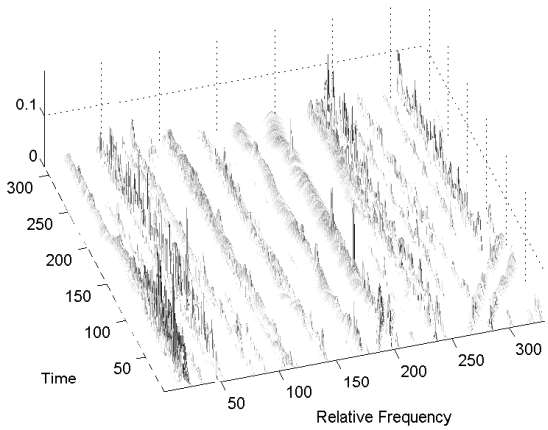

Figure 9: The reconstructed real data time-frequency spectrum in Fig. 6 plotted in 3-D.

Ferguson, TS 1974, 'A Bayesian analysis of some nonparametric problems', *Annals of Statistics*, vol. 1, pp. 209–230.

Koch, JW 2008, 'Bayesian approach to extended object and cluster tracking using random matrices', *IEEE Trans. Aerospace Electr. Syst.*, vol. 44, no. 3, pp. 1042–1058.

Neal, RM 2000, 'Markov chain sampling methods for Dirichlet process mixture models', *Journal Computational Graphical Statistics*, vol. 9, no. 2, pp. 249–265.

Nielsen, RO 1991, *Sonar signal procesing*, Artech House, Norwood MA.

Ozkan, E, Ozbek, IY & Demirekler, M 2009, 'Dynamic speech spectrum representation and tracking variable number of vocal tract resonance frequencies with time-varying Dirichlet process mixture models', *IEEE Trans. Audio, Speech, Language Process.*, vol. 17, no. 8, pp. 1518–1532.

Pace, DW, M. Mallick, M & Eldredge, W 2003, 'Spectral feature-aided multi-target multi-sensor passive sonar tracking', *Proceedings of OCEANS 2003*, vol. 4, San Diego, CA, pp. 2120 – 2126.

Sarkka, S, Vehtari, A & J. Lampinen, J 2007, 'Rao-Blackwellized Monte Carlo data association for multiple target tracking', *Special Issue of the Seventh International Conference on Information Fusion-Part II*, pp. 2–15.

Sethuraman, J 1994, 'A constructive definition of Dirichlet priors', *Statistica Sinica*, vol. 4, pp. 639–650.

Sudderth, EB 2006, 'Graphical models for visual object recognition and tracking', PhD thesis, Massachusetts Institute of Technology.

Teh, YW 2010, 'Dirichlet process', *Encyclopedia of Machine Learning*, New York, Springer.

Wang, X, La Scala, B & Ellem, R 2008, 'Feature aided probabilistic data association for multi-target tracking', *Proceedings of International Conference on Information Fusion*, pp. 1–7.

Wood, F & Black, MJ 2008, 'A nonparametric Bayesian alternative to spike sorting', *Journal Neurosci. Methods*, vol. 173, no. 1, pp. 1–12.