TRACKING OF TIME-FREQUENCY COMPONENTS USING PARTICLE FILTERING

Corentin Dubois[1], *Manuel Davy*[2] *and Jérôme Idier*[1]

[1] IRCCyN UMR 6597, BP 92101, 1 rue de la Noë, 44321 Nantes Cedex 03, France
[2] LAGIS UMR 8146, BP 48, Cité scientifique, 59651 Villeneuve d'Ascq Cedex, France
e-mail: corentin.dubois@irccyn.ec-nantes.fr

ABSTRACT

This paper investigates a new method for the extraction of spectral trajectories from nonstationary multicomponent narrow band signals. The main idea is to apply statistical filtering so as to track frequency trajectories in the Short Term Fourier Transform. A nonlinear observation model is defined and a special particle filtering algorithm is developed.

1. INTRODUCTION

A common Signal Processing task consists of extracting time-varying frequency information from signals. For example, in audio signal processing, we want to track the evolution of frequency components with respect to time so as to follow the *pitch* of a piece of music. An additional task can be to decompose a signal made of some time-varying 'narrow-band frequency components' into the individual components. Many methods have already been proposed to solve this problem, among which ARCAP based frequency tracking [1, 2] and Empirical Modal Decomposition [3].

In this paper, we propose to cast this problem into a statistical filtering form. Similar to a mobile object tracking problem, we define a model for the time evolution of the frequency and amplitude of each components. Additionally, we use as observations the Short Term Fourier Transform (STFT) of the signal: more precisely, assume the signal x, $t = 1, 2, \ldots$, can be locally written as the sum of k_t stationary sine waves embedded in additive noise. The observations model is made local by considering a windowed version x_t of x, $t = 1, 2, \ldots$:

$$x_t[\tau] = x[t + \tau - L_w/2]w[\tau]\,, \quad \tau = 1, \ldots, L_w \tag{1}$$

where $w[\tau]$ is the symmetric window of length L_w. We assume that, on the window w, the signal is locally stationary (this is the underlying assumption of the STFT) and that it writes as

$$x_t[\tau] = s_t[\tau] + b[\tau]\,, \quad \tau = 1, \ldots, L_w \tag{2}$$

with

$$s_t[\tau] = \left(\sum_{k=1}^{k_t} a_{t,k} \cos(2\pi f_{t,k}\tau) \right) w[\tau] \tag{3}$$

and b an additive noise.

The filtering problem consists of estimating online the unknown parameters k_t, $\mathbf{F}_t = [f_{t,1}, \ldots, f_{t,k_t}]$ and $\mathbf{A}_t$ (where $\mathbf{A}_t$ is built as $\mathbf{F}_t$). Let $\theta_t = \left[\mathbf{F}_t\,,\, \mathbf{A}_t\right]$. The problem consists of estimating (k_t, θ_t) sequentially from the observations, using a frequency/amplitude evolution model. To solve this filtering problem, we use sequential Monte Carlo methods (also referred to as particle filtering), which provides us with a particle approximation of the distribution of the state $p(k_{0:t}, \theta_{0:t}|y_{1:t})$ [4] where, e.g., $\theta_{0:t}$ denotes $[\theta_0, \theta_1, \cdots, \theta_t]$. It is easy to estimate (k_t, θ_t) from these particles.

This paper is organized as follows: in Section 2, we present the frequency/amplitude evolution model (transition probability) as well as the observation probability. In Section 3, we propose an efficient implementation of the sequential Monte Carlo algorithm samplers, specially designed for this problem. In Section 4, we present some results before conclusions and directions of future research are proposed.

2. A HIDDEN MARKOV MODEL FOR FREQUENCY/AMPLITUDE TRACKING

The frequency/amplitude evolution model is a simple random walk and transition equations are of general shape:

$$u_{t+1,k} = u_{t,k} + v_{t,k} \tag{4}$$

where $v_{t,k}$ is a noise and k the number of the trajectory at time t, $k = 1, \ldots, k_t$.

2.1. Transition probabilities

In this subsection, we give explicit expressions for the transition equation of the state parameters and we precise notations.

2.1.1. Number of components k_t

The number of components k_t can either be kept constant ($k_{t+1} = k_t$), be increased ($k_{t+1} = k_t + 1$) or be decreased

	$k_t = k_{min}$	$k_{min} < k_t$ $k_t < k_{max}$	$k_t = k_{max}$
$\Pr(k_{t+1} = k_t - 1)$	0	1/4	1/3
$\Pr(k_{t+1} = k_t)$	2/3	1/2	2/3
$\Pr(k_{t+1} = k_t + 1)$	1/3	1/4	0

Table 1. Transition probabilities for the number of components k_t

$(k_{t+1} = k_t - 1)$ with the probabilities given in Table 1. The number of components ranges from k_{min} to k_{max}.

2.1.2. *Frequencies transition*

Using the shape of Eq. (4), the frequency transition equation is

$$f_{t+1,k} = f_{t,k} + v^f_{t,k} \tag{5}$$

where $v^f_{t,k}$ is a zero-mean white Gaussian noise of variance $(\sigma^f_{t,k})^2$. We allow this variance to change according to the following evolution model:

$$(\sigma^f_{t+1,k})^2 = (\sigma^f_{t,k})^2 \mathrm{e}^{\varphi_{t,k}} \text{ with } \varphi_{t,k} \sim \mathcal{N}(0, \sigma^2_\varphi) \tag{6}$$

2.1.3. *Amplitudes transition*

Similar to Eq. (4), the amplitude transition equation is

$$a_{t+1,k} = a_{t,k} + v^a_{t,k} \tag{7}$$

where $v^a_{t,k}$ is a zero-mean white Gaussian noise of variance $(\sigma^a_{t,k})^2$. Again, we allow the variance to change with time according to:

$$(\sigma^a_{t+1,k})^2 = (\sigma^a_{t,k})^2 \mathrm{e}^{\alpha_{t,k}} \text{ with } \alpha_{t,k} \sim \mathcal{N}(0, \sigma^2_\alpha) \tag{8}$$

2.2. Observations probability (likelihood)

The observation is given by:

$$y_t[\tau] = r\big(x_t[\tau]\big)\,, \quad \tau = 1, \ldots, L_w \tag{9}$$

where x_t comes from Eq. (1) and r is the magnitude of the STFT. The observation law is:

$$y_t[\tau] = r\big(s_t[\tau]\big) + v^y_t[\tau]\,, \quad \tau = 1, \ldots, L_w \tag{10}$$

where the signal s_t is computed from θ_t using Eq. (3). The additive noise $v^y_t[\tau]$, $\tau = 1, \ldots, L_w$, is assumed i.i.d Gaussian with variance $(\sigma^y_t)^2$. Similar to the evolution model variances, the observation model variance can change with time, according to:

$$(\sigma^y_{t+1})^2 = (\sigma^y_t)^2 \mathrm{e}^{\epsilon_t} \text{ with } \epsilon_t \sim \mathcal{N}(0, \sigma^2_\epsilon) \tag{11}$$

The observation probability, or likelihood, comes directly from Eq.'s (1)-(3) and is given by:

$$\begin{aligned} p(y_t|\theta_t) &= p(y_t|s_t) \\ &\propto \exp\left(-\frac{\big(y_t - r(s_t)\big)^{\mathtt{T}}\big(y_t - r(s_t)\big)}{2(\sigma^y_t)^2}\right) \end{aligned} \tag{12}$$

3. PARTICLE FILTERING ALGORITHM

Particle filtering methods are recursive algorithms which produce, at each time t, a set of particles $\{\theta^{(i)}_{0:t}; i = 1 \cdots N\}$ which give us an approximation of the posterior state distribution:

$$\hat{p}(\theta_{0:t}|y_{1:t}) = \frac{1}{N}\sum_{i=1}^{N} \delta_{(\theta^{(i)}_{0:t})}(\mathrm{d}\theta_{0:t}) \tag{13}$$

Particles are drawn from the proposal distribution $q(\theta_t|\theta_{0:t-1}, y_{1:t})$ and selected according to their weight:

$$\omega^{(i)}_t = \frac{p(y_t|\theta_t)p(\theta_t|\theta_{t-1})}{q(\theta_t|\theta_{0:t-1}, y_{1:t})} \tag{14}$$

The design of the proposal distribution is of paramount importance in sequential importance sampling algorithms. It have been shown [5] that particle filters with a proposal distribution obtained using the Unscented Kalman Filter (UKF) are particulary suitable if the model of observation is not linear. In our case, the observation is the absolute value of the STFT of the signal.

So, we propose the following implementation of the Unscented Particle Filter (UPF) [5], applied to our problem.

• At time $t = 0$
Initialization
For $i = 1 \cdots N$ set $R^{(i)}_0 = (\sigma^y_0)^2$ and

$$Q^{(i)}_0 = \begin{bmatrix} (\sigma^f_{0,k})^2 & 0 \\ 0 & (\sigma^a_{0,k})^2 \end{bmatrix}, \quad k = 1, \ldots, k_t$$

For $i = 1 \cdots N$ sample $\big(k^{(i)}_0, \theta^{(i)}_0\big) \sim p(k_0, \theta_0)$ and set $\bar{\theta}^{(i)}_0 = \mathbb{E}[\theta^{(i)}_0]$

• At time $t \geq 1$
For $i = 1 \cdots N$
Update the number of trajectories
$k^{(i)}_t = k^{(i)}_{t-1} + b$ with $b = -1, 0$ or 1 with the probabilities given in Table 1.
Update trajectories
i) compute sigma points

$$\mathcal{O}^{(i)}_{t-1} = \left[\bar{\theta}^{(i)}_{t-1}\;\bar{\theta}^{(i)}_{t-1} \pm \left(\sqrt{(n_\theta + \lambda)Q^{(i)}_{t-1}}\right)_l\right]$$

where $\left(\sqrt{(n_\theta+\lambda)Q_{t-1}^{(i)}}\right)_l$ is the lth column, $l = 1 \cdots n_\theta$, of the matrix square root of $(n_\theta + \lambda)Q_{t-1}^{(i)}$, n_θ is the dimension of the state and λ is the Unscented Kalman Filter parameter.

ii) propagate the particles using the unscented transform

$$\mathcal{O}_{t|t-1}^{(i)} = \mathcal{O}_{t-1}^{(i)}$$

$$\bar{\theta}_{t|t-1}^{(i)} = \sum_{j=0}^{2n_\theta} W_j^{(m)} \mathcal{O}_{j,t|t-1}^{(i)}$$

$$\mathcal{Y}_{t|t-1}^{(i)} = r\left(\mathcal{O}_{t|t-1}^{(i)}\right)$$

$$\bar{y}_{t|t-1}^{(i)} = \sum_{j=0}^{2n_\theta} W_j^{(m)} \mathcal{Y}_{j,t|t-1}^{(i)}$$

where the notation $r(\theta_t)$ hides the step where θ_t is transformed into s_t with Eq.(3) and $W^{(m)} = [2\lambda\, 1 \cdots 1]/2(n_\theta + \lambda)$.

iii) incorporate the new observation using the Kalman filter equations

$$\begin{aligned}\mathbf{P}_{y_t y_t} &= (\sigma_t^y)^2 \mathbf{I}_{L_w} + \\ &\sum_{j=0}^{2n_\theta} W_j^{(c)} [\mathcal{Y}_{j,t|t-1}^{(i)} - \bar{y}_{t|t-1}^{(i)}][\mathcal{Y}_{j,t|t-1}^{(i)} - \bar{y}_{t|t-1}^{(i)}]^{\mathrm{T}}\end{aligned}$$

$$\mathbf{P}_{\theta_t y_t} = \sum_{j=0}^{2n_\theta} W_j^{(c)} [\mathcal{O}_{j,t|t-1}^{(i)} - \bar{\theta}_{t|t-1}^{(i)}][\mathcal{Y}_{j,t|t-1}^{(i)} - \bar{y}_{t|t-1}^{(i)}]^{\mathrm{T}}$$

$$\mathbf{K}_t = \mathbf{P}_{\theta_t y_t} \mathbf{P}_{y_t y_t}^{-1}$$

$$\bar{\theta}_t^{(i)} = \bar{\theta}_{t|t-1}^{(i)} + \mathbf{K}_t (y_t - \bar{y}_{t|t-1}^{(i)})$$

$$\mathbf{P}_{tt}^{(i)} = Q_{t-1}^{(i)} - \mathbf{K}_t \mathbf{P}_{y_t y_t} \mathbf{K}_t^{\mathrm{T}}$$

iv) sample $\theta_t^{(i)} \sim q(\theta_t^{(i)}|\theta_{0:t-1}^{(i)}, y_{1:t}) = \mathcal{N}(\bar{\theta}_t^{(i)}, \mathbf{P}_{tt}^{(i)})$

Update the hyperparameters

We define

$$\beta_{t-1}^{(i)} = 2\big[\cdots \log(\sigma_{t-1,k}^f) \cdots \log(\sigma_{t-1,k}^a) \cdots \log(\sigma_{t-1}^y)\big]^{\mathrm{T}}$$

and

$$\varepsilon_t^{(i)} = \theta_t^{(i)} - \theta_{t-1}^{(i)}$$

Then, sample the set of log-variances

$$\beta_t^{(i)} \sim p(\beta_t^{(i)}|\beta_{t-1}^{(i)}, \varepsilon_t^{(i)})$$

with

$$p(\beta_t^{(i)}|\beta_{t-1}^{(i)}, \varepsilon_t^{(i)}) = \frac{p(\varepsilon_t^{(i)}|\beta_t^{(i)}) p(\beta_t^{(i)}|\beta_{t-1}^{(i)})}{p(\varepsilon_t^{(i)}|\beta_{t-1}^{(i)})}$$

and update $Q_{t-1}^{(i)}$ and $R_{t-1}^{(i)}$ into $Q_t^{(i)}$ and $R_t^{(i)}$ using $\beta_t^{(i)}$.

Evaluate the importance weights

$$\tilde{\omega}_t^{(i)} = p(y_t|\theta_t^{(i)}) \frac{p(\theta_t^{(i)}|\theta_{t-1}^{(i)})}{q(\theta_t^{(i)}|\theta_{0:t-1}^{(i)}, y_{1:t})} \frac{p(\beta_t^{(i)}|\beta_{t-1}^{(i)})}{q(\beta_t^{(i)}|\beta_{t-1}^{(i)}, \varepsilon_t^{(i)})}$$

Normalize the importance weights so that

$$\omega_t^{(i)} \propto \tilde{\omega}_t^{(i)} \ ; \ \sum_{i=1}^{N} \omega_t^{(i)} = 1$$

Selection step

Multiply/Suppress particles with high/low importance weight respectively, to obtain N random particles with importance weight $1/N$

Output

The algorithm gives us an approximation of the posterior distribution from which it is easy to compute statistical estimates such as:

$$\mathbb{E}[g_t(\theta_{0:t})] \approx \frac{1}{N} \sum_{i=1}^{N} g_t(\theta_{0:t}^{(i)})$$

In our method, the function we are interested in is the marginal conditional mean of $\theta_{0:t}$ i.e. $g_t(\theta_{0:t}) = \theta_t$

---

4. RESULTS

We apply this particle filtering algorithm on synthetic signals and we compare the estimated frequency/amplitude components with the ones used to synthesize the signal. An example of this study is represented on Fig.'s 1-2-3. The simulated signal is composed of one stationary tone at normalized frequency 0.18, a first transient tone at frequency 0.3 from time sample 1 to time sample 219 and a second one at frequency 0.27 from time sample 569 to time sample 1000, a sine-modulated component with mean normalized frequency 0.4 and a linear chirp. Note that the number of spectral components is time-varying. A Gaussian white noise is added to the data such that the SNR is about 17 dB. The following set of parameters was chosen: $N = 250$ (the number of particles), $k_{min} = 1$, $k_{max} = 15$, $\sigma_\varphi = \sigma_\alpha = \sigma_\epsilon = 0.1$ and w is a Hamming window of length 63.

In Fig. 1, we can see that both the number of components and the frequency trajectories are estimated with good accuracy. In part (c) of this figure, we circled three events where the estimation of the number of trajectories seems to be erroneous. The events in circles 1 and 2 are due to the windowing effect, which affects both the time and frequency resolutions of the spectrogram: the time events are well localized, up to the uncertainty caused by the window length. Event in circle 3 is not a mistake: when two trajectories cross, there exists effectively only one trajectory.

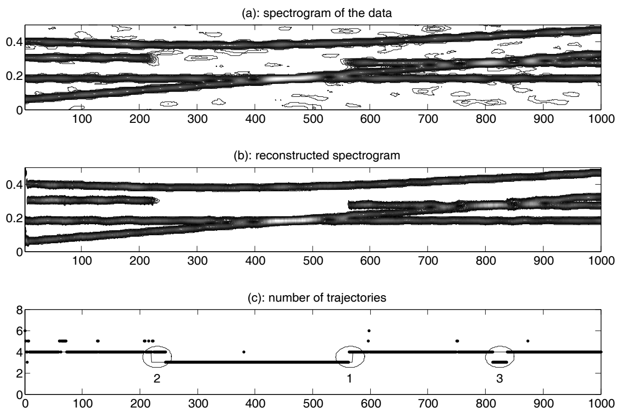

Fig. 1. (a): Spectrogram of the data. The column of this spectrogram at time t is the observation vector y_t. (b): Reconstructed spectrogram. The column of this spectrogram at time t is given by $r(\hat{\theta}_t)$ with $\hat{\theta}_t = \sum_{i=1}^{N} \theta_t^{(i)} r(\theta_t^{(i)})$. (c): the doted line is the estimated number of trajectories and the solid line indicates the simulated number of trajectories.

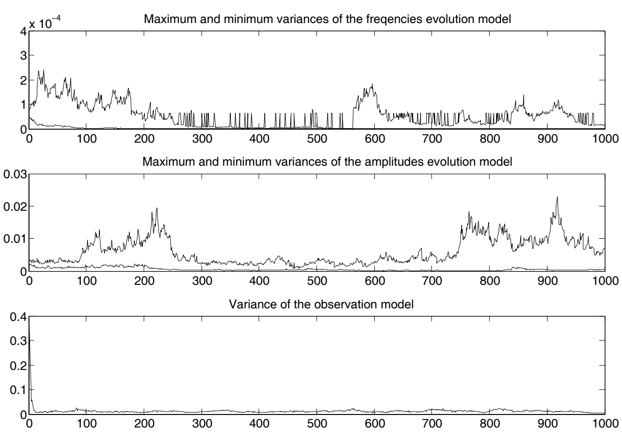

Fig. 2. Estimated variances increase when there is a change of number of components but the main behaviour is the decrease, what means that the estimation of the variance is meaningful.

5. CONCLUSION

Estimation of instantaneous frequencies and tracking of time-frequency components can be performed from Short Term Fourier Transform and particle filtering. This method allows birth or death of spectral trajectories. This enables detection of events of limited duration, in audio signal processing, for instance.

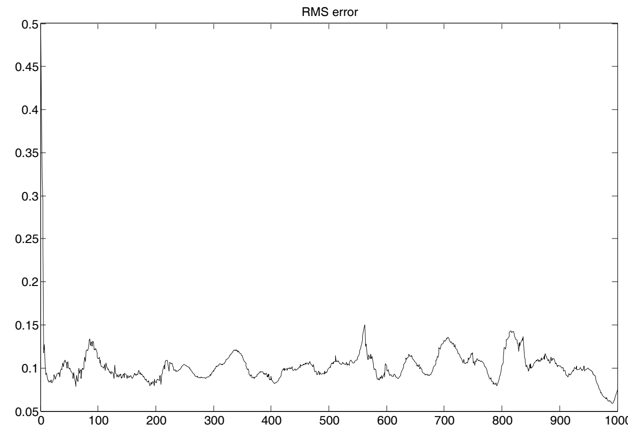

Fig. 3. At time t, the Root Mean Square error is computed by: $\sqrt{\frac{\left(y_t - r(\hat{\theta}_t)\right)^{\mathrm{T}}\left(y_t - r(\hat{\theta}_t)\right)}{L_w}}$

6. REFERENCES

[1] M. Davy, B. Leprettre, C. Doncarli, and N. Martin, "Tracking of spectral lines in an ARCAP time-frequency representation," in *EUSIPCO-1998, 9th European Signal Processing Conference*, Island of Rhodes, Greece, 1998, pp. 633–636.

[2] H. Cottereau, M. Davy, C. Doncarli, and N. Martin, "Using arcap time-frequency representations for decisions," in *EUSIPCO-2000, 10th European Signal Processing Conference*, Tampere, Finlande, Sept. 2000.

[3] N.E. Huang, Z. Shen, S.R. Long, M.L. Wu, H.H. Shih, Q. Zhen, N.C. Yen, C.C. Tung, and H.H. Liu, "The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis," *Porc. Roy. Soc. London A*, vol. 454, pp. 903–995, 1998.

[4] A. Doucet, "On sequential simulation-based methods for bayesian filtering," Tech. Rep., Signal Processing Group, Department of engineering, University of Cambridge CB2 1PZ Cambridge, 1998.

[5] R. van der Merwe, A. Doucet, N. de Freitas, and E. Wan, "The unscented particle filter," Tech. Rep., Department of engineering, University of Cambridge CB2 1PZ Cambridge, 2000.

[6] C. Andrieu, M. Davy, and A. Doucet, "Efficient Particle Filtering for Jump Markov Systems. Application to Time-Varying Autoregressions," *IEEE trans. on Signal Processing*, July 2003.