

FRAMEWORK FOR ONLINE SUPERIMPOSED EVENT DETECTION BY SEQUENTIAL MONTE CARLO METHODS

O. Urfalioğlu¹, Ercan E. Kuruoğlu² and A. Enis Çetin¹

¹Department of Electrical and Electronics Engineering
Bilkent University, Ankara/Turkey

²ISTI/CNR, Pisa/Italy

ABSTRACT

In this paper, we consider online separation and detection of superimposed events by applying particle filtering. We concentrate on a model where a background process, represented by a 1D-signal, is superimposed by an Auto-Regressive (AR) 'event signal', but the proposed approach is applicable in a more general setting. The activation and deactivation times of the event-signal are assumed to be unknown. We solve the online detection problem of this superpositional event by extending the state space dimension by one. The additional parameter of the state represents the AR-signal, which is zero when deactivated. Numerical experiments demonstrate the effectiveness of our approach.

Index Terms— Event detection, Conditional Density, SIR, Importance Sampling, Bayesian Statistics

1. INTRODUCTION

Event detection is becoming an important and more frequently studied field in recent times. There are applications in intrusion detection, internet traffic analysis, bio-information processing, telecommunication, surveillance and more. In this paper, online model based event detection using sequential Monte Carlo methods, namely particle filtering [3, 5, 6, 8, 4, 2], is studied. The term model based emphasizes that the stochastic model of the event is known. On the other hand, the activation time of the event is unknown and the event is superpositional with respect to a background process. This stochastic event-process is modeled as an Auto Regressive (AR) process, which superimposes a background stochastic process. So, in this setting, only the result of this superposition is observable. The task of the proposed approach is to simulate and estimate the hidden background process, to detect the event activation/deactivation times and to estimate also the hidden event process.

In many event detection methods [11], the estimated state or a sequence of estimated states is undergone a secondary analysis by e.g. using a Hidden Markow Model (HMM). This

HMM represents the model of the event statistics. MCMC-based methods as in [9, 7, 10] are generally not applicable in an online approach due to high computational requirements. In [1], an overview of change point detection using particle filters is given. However, our approach not just attacks a change point detection problem, but enables also online source separation.

In the proposed approach, the detection method of the event is embedded in the particle filtering framework directly. By increasing the state space dimension by the number of additional event process parameters and appropriately choosing the importance functions, we are able to estimate both the hidden process and the superpositional event process simultaneously. This is accomplished by minimal modification of the particle filtering framework.

2. SEQUENTIAL MONTE CARLO METHODS - SIR

In the Sequential Monte Carlo (SMC) setting, the stochastic process consists of hidden state propagation and observation, represented as a HMM(1). The state propagation $x_t \rightarrow x_{t+1}$ at time t is modeled as

$$x_{t+1} = f(x_t) + v_t, \quad (1)$$

and the observation y_t is modeled as

$$y_t = g(x_t) + w_t, \quad (2)$$

where v_t and w_t are independent random variables. The sequential Bayesian inference consists of a prediction step and an update step via

$$p(x_{t+1}|y_{1:t}) = \int p(x_{t+1}|x_t)p(x_t|y_{1:t})dx_t \quad (3)$$

$$p(x_{t+1}|y_{1:t+1}) = \frac{p(y_{t+1}|x_{t+1})p(x_{t+1}|y_{1:t})}{p(y_{t+1}|y_{1:t})}, \quad (4)$$

with

$$y_{1:t} \triangleq \{y_1, \dots, y_t\}. \quad (5)$$

This work was carried out during the tenure of a MUSCLE Internal fellowship.

Due to generally intractable integrals, the sequential Bayesian inference is realized by approximation methods such as Sequential Importance Resampling (SIR) [3]. In the SIR framework, the posterior $p(x_t|y_{1:t})$ is represented by a particle set of N particles. A particle consists of a position vector $x_{n,t}$ and a weight-scalar $\omega_{n,t}$ with the approximation property

$$p(dx_t|y_{1:t}) \approx \sum_{n=1}^N \omega_{n,t} \delta(dx_{n,t}), \quad (6)$$

where

$$\sum_{n=1}^N \omega_{n,t} = 1. \quad (7)$$

The particle positions are sampled from an importance density

$$x_{n,t+1} \sim \pi(x_{t+1}|x_{n,t}, y_{1:t+1}) \quad (8)$$

at each time step. The weights are determined by

$$\tilde{\omega}_{n,t+1} = \omega_{n,t} \frac{p(y_{t+1}|x_{t+1})p(x_{t+1}|x_t)}{\pi(x_{t+1}|x_{n,t}, y_{1:t+1})} \quad (9)$$

and normalized afterwards

$$\omega_{n,t+1} = \frac{\tilde{\omega}_{n,t+1}}{\sum_{n=1}^N \tilde{\omega}_{n,t+1}}. \quad (10)$$

Due to degeneracy in this method regarding the importance weights, on which all but one particle has a weight of 1 and all others have zero weight, a resampling step is added after each iteration. The resampling is done by copying the particle $N\omega$ times in average by overwriting other particles, so particles with strong weights are reproduced more often, in average.

3. FRAMEWORK FOR SUPERIMPOSED EVENT DETECTION

The type of events we consider can be modeled as follows. The background signal, denoted by x_t , is superimposed by a second signal, denoted by z_t , which is independent of x_t

$$x_{t+1} = f(x_t) + v_t + \alpha_t z_{t+1}. \quad (11)$$

The event signal is assumed to be only present for some time window T_E

$$\alpha_t = \begin{cases} 1 & t \in T_E \\ 0 & \text{else.} \end{cases} \quad (12)$$

Since there is no 'pure' observation available from the signal z_t , it can only be estimated together with x_t . We assume that a parametric description of the signal z_t , specified by

$$z_{t+1} = h(z_t) + u_t, \quad (13)$$

is available

$$z_t = z_t(\theta_t). \quad (14)$$

The task is to detect the event, in this case to tell whether there is a superpositional z_t present and to estimate z_t . The proposed approach consists of using an SIR-particle filter, whose state space dimension is extended by the number of the required additional hidden parameters, having the state vector s_t

$$s_t = (x_t, z_t, \alpha_t, \theta_t). \quad (15)$$

Alternatively, the parameter α_t can be discarded by adapting the conditional probability density $p(z_{t+1}|z_t)$ of z_t by

$$p'(z_{t+1}|z_t) = \frac{1}{2}(\delta(0) + p(z_{t+1}|z_t)), \quad (16)$$

where $\delta(\cdot)$ is the Dirac substitution and $\delta(0)$ produces exact zeros as 'no-event' samples.

The state propagation density for the superimposed signal can be written as

$$p(x_{t+1}, z_{t+1}|x_t, z_t) = p(x_{t+1}|x_t)p(z_{t+1}|z_t). \quad (17)$$

3.1. The choice of importance functions

The choice of the importance function is crucial in the SIR-framework, since it has a great impact on the efficiency and even feasibility of the simulations. One of the most common methods is to use the state propagation density as the importance density function, as in [6]. Though this choice does not take the current observation into account, it is sufficient for many simulation problems.

The importance function for α_t can be chosen as

$$\pi_\alpha(\alpha_{t+1}|\alpha_t) = \frac{1}{2}\delta(0) + \frac{1}{2}\delta(1). \quad (18)$$

In the spirit of [6], a possible choice for the importance function of z_t is its propagation density

$$\pi_z(z_{t+1}|z_t, x_t, y_t, \alpha_t) = p(z_{t+1}|z_t). \quad (19)$$

or, in the case of discarding the parameter α_t , we may choose

$$\pi_z(z_{t+1}|z_t, x_t, y_t) = p'(z_{t+1}|z_t), \quad (20)$$

and so for the background process

$$\pi_x(x_{t+1}|x_t, y_t) = p(x_{t+1}|x_t). \quad (21)$$

For the joint importance density follows

$$\pi_x(x_{t+1}, z_{t+1}|x_t, z_t, y_t) = p(x_{t+1}|x_t)p'(z_{t+1}|z_t). \quad (22)$$

The importance densities may have higher variances than their corresponding propagation densities in order to 'capture' the additional uncertainty influenced from the observation model. The importance function of the parameters θ_t is highly dependent on the model dynamics and should be chosen accordingly.

3.2. Detection

The indicator I_t for the event is easily calculated by counting the number Z of exact zeros over all N particles at time step t , i.e.

$$Z_t = \sum_{n=1}^N \delta_{z_{n,t}}, \quad (23)$$

where $\delta_{z_{n,t}}$ is the so called Kronecker-Delta, the discrete version of the Dirac substitution, with the property

$$\delta_{z_{n,t}} = \begin{cases} 0 & z_{n,t} \neq 0 \\ 1 & z_{n,t} = 0. \end{cases} \quad (24)$$

Having calculated the number Z_t of exact zeros, the event indicator I_t is calculated by

$$I_t = \begin{cases} 0 & \frac{Z_t}{N} < \frac{1}{2} \\ 1 & \text{else.} \end{cases} \quad (25)$$

4. EXPERIMENTS

4.1. Setup

For the background process, we use the following state propagation function

$$x_{t+1} = 12 + 0.5x_t \sin(t/5) + \mathcal{N}(0, \sigma_x^2) \quad (26)$$

and the following observation function

$$y_t = 0.5x_t^2 - 2 + \mathcal{N}(0, \sigma_o^2), \quad (27)$$

where $\mathcal{N}(m, \sigma^2)$ denotes a normal probability density function with mean m and variance σ^2 . The resulting state propagation function including the superpositional AR(1) process z_t is given by

$$x_{t+1} = 12 + 0.5x_t \sin(t/5) + \mathcal{N}(0, \sigma_x^2) + z_{t+1}, \quad (28)$$

where the propagation of the AR(1) process is given by

$$z_{t+1} = az_t + \mathcal{N}(0, \sigma_z^2). \quad (29)$$

The extended state vector is determined by

$$s_t = (x_t, z_t). \quad (30)$$

The importance functions for the states x_t, z_t are chosen as

$$\begin{aligned} \pi(x_{t+1}|x_{n,t}, y_{1:t+1}) &= p(x_{t+1}|x_{n,t}) \\ \pi(z_{t+1}|z_{n,t}, y_{1:t+1}) &= 0.5\delta(0) + 0.5\mathcal{N}(az_{n,t}, \sigma_z^2), \end{aligned} \quad (31)$$

with variances of

$$\sigma_{\pi(x)}^2 = 6\sigma_x^2, \quad \sigma_{\pi(z)}^2 = 6\sigma_z^2. \quad (32)$$

The compound importance function is then defined by

$$\begin{aligned} &\pi(x_{t+1}, z_{t+1}|x_{n,t}, z_{n,t}, y_{1:t+1}) \\ &= \pi(x_{t+1}|x_{n,t}, y_{1:t+1})\pi(z_{t+1}|z_{n,t}, y_{1:t+1}) \end{aligned} \quad (33)$$

$$= p(x_{t+1}|x_{n,t})(0.5\delta(0) + 0.5\mathcal{N}(az_{n,t}, \sigma_z^2)). \quad (34)$$

It is obvious that the detection success probability depends on the variances σ_x^2 and σ_z^2 . With different values for σ_x^2 , an observation noise of $\sigma_o^2 = 0.001$ and an AR(1) process noise of $\sigma_z^2 = 0.2$, we performed simulations of the hidden states x_t and z_t .

The event was activated within $T_E = [50, 70[$. We calculated the detection rates, including the false positive alarm probability e^+ , and false negative alarm probability e^- by repeating the state sequence estimations 50 times each. The number of particles was set to $N = 500$.

4.2. Results

Figures 1, 2, 3 and 4 show the results of the filter estimates vs. true values of both the x_t and the z_t signals and the event detection indicator bars for the state-variances of $\sigma_x^2 = 10^{-2}$, $\sigma_x^2 = 10^{-3}$, $\sigma_x^2 = 10^{-4}$ and $\sigma_x^2 = 10^{-5}$. It is known from detection theory, that the success of correct detection depends on the noise of the signals. As expected, the detection error decreases for smaller variances of the noise of the x -signal. The detection errors are shown in table 1.

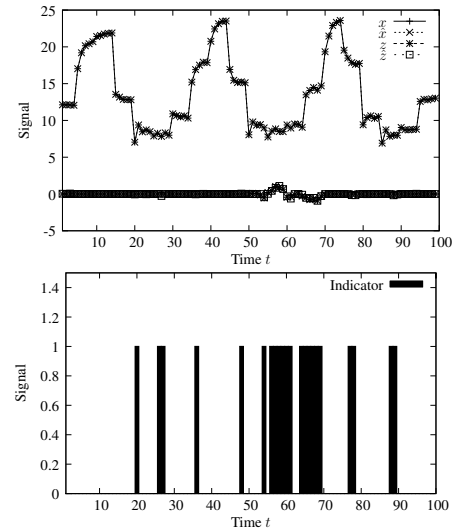


Fig. 1. True and estimated signals x_t, z_t at $\sigma_z^2 = 2 \cdot 10^{-1}$, $\sigma_x^2 = 10^{-2}$

5. CONCLUSIONS

In case where events can be described by superpositional stochastic processes and the state propagation densities, a.k.a. the 'models' are known, the proposed framework can be used for online separation and detection of 2 or more simultaneous events. Results can be further improved when the minimum activation time interval of the event signal is known and

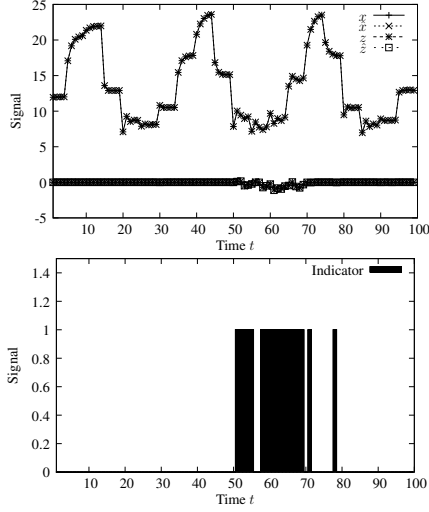


Fig. 2. True and estimated signals x_t , z_t at $\sigma_z^2 = 2 \cdot 10^{-1}$, $\sigma_x^2 = 10^{-3}$

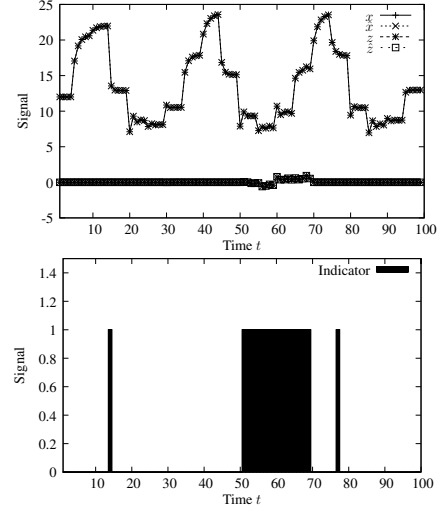


Fig. 4. True and estimated signals x_t , z_t at $\sigma_z^2 = 2 \cdot 10^{-1}$, $\sigma_x^2 = 10^{-5}$

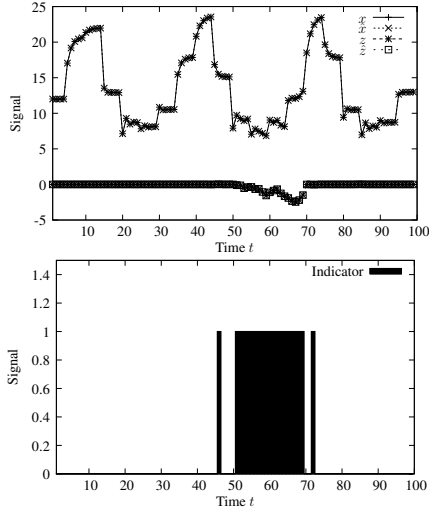


Fig. 3. True and estimated signals x_t , z_t at $\sigma_z^2 = 2 \cdot 10^{-1}$, $\sigma_x^2 = 10^{-4}$

σ_x^2	σ_z^2	e^+	e^-
10^{-5}	$2 \cdot 10^{-1}$	0.0057	0.061
10^{-4}	$2 \cdot 10^{-1}$	0.0252	0.074
10^{-3}	$2 \cdot 10^{-1}$	0.0582	0.105
10^{-2}	$2 \cdot 10^{-1}$	0.1115	0.186

Table 1. Event detection false positive alarm probabilities and false negative alarm probabilities for several background process noise variances

covers more than 1 sample. In this case, a modified event indicator would depend on the whole time interval and decide accordingly.

6. REFERENCES

- [1] C. Andrieu, A. Doucet, S. S. Singh, and V. B. Tadic. Particle methods for change detection, system identification, and control, 2004.
- [2] P. Del Moral. *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. Springer, 2004.
- [3] A. Doucet. On sequential simulation-based methods for bayesian filtering, 1998. Technical Report CUED/F-INFENG/TR.310.
- [4] A. Doucet, J. F. G. de Freitas, and N. J. Gordon. *Sequential Monte Carlo Methods in Practice*. Springer, 2001.
- [5] P. Fearnhead. *Sequential Monte Carlo methods in filter theory*. Dissertation, University of Oxford, 1998.
- [6] M. Isard and A. Blake. Condensation - conditional density propagation for visual tracking. *Int. J. Computer Vision*, 29:5–28, 1998.
- [7] M. Lavielle and E. Lebarbier. An application of mcmc methods for the multiple change-points problem. *Signal Processing*, 81(1):39–53, January 2001.
- [8] J. S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer, 2001.
- [9] S. Suparman, M. Doisy, Tourneret, and J.-Y. Changepoint detection using reversible jump mcmc methods. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 1569–1572, 2002.
- [10] J.-Y. Tourneret, M. Doisy, and M. Lavielle. Bayesian off-line detection of multiple change-points corrupted by multiplicative noise: application to sar image edge detection. *Signal Process.*, 83(9):1871–1887, 2003.
- [11] D. Zotkin, R. Duraiswami, and L. S. Davis. Multimodal 3-d tracking and event detection via the particle filter. In *Workshop on Event Detection in Video, International Conference on Computer Vision*, 2001.