

Superimposed event detection by particle filters

O. Urfalioglu¹ E.E. Kuruoglu² A.E. Cetin¹

¹Department of Electrical and Electronics Engineering, Bilkent University, 06800 Ankara, Turkey

²Istituto di Scienza e Tecnologie dell'Informazione 'Alessandro Faedo', Area della Ricerca CNR di Pisa, Via G. Moruzzi 1, 56124 Pisa, Italy

E-mail: onay.urfalioglu@gmail.com

Abstract: In this study, the authors consider online detection and separation of superimposed events by applying particle filtering. They observe only a single-channel superimposed signal, which consists of a background signal and one or more event signals in the discrete-time domain. It is assumed that the signals are statistically independent and can be described by random processes with known parametric models. The activation and deactivation times of event signals are assumed to be unknown. This problem can be described as a jump Markov system (JMS) in which all signals are estimated simultaneously. In a JMS, states contain additional parameters to identify models. However, for superimposed event detection, the authors show that the underlying JMS-based particle-filtering method can be reduced to a standard Markov chain method without additional parameters. Numerical experiments using real-world sound processing data demonstrate the effectiveness of their approach.

1 Introduction

Event and change detection using particle filtering have received increasing attention recently [1–8]. There are applications in speech and sound processing, intrusion detection, internet traffic analysis, bio-information processing, telecommunication, surveillance and more. In this paper, an online model-based event detection using sequential Monte Carlo methods (SMC), namely particle filtering [4–6], is studied. Some example applications are image separation [9], modelling of non-stationary auto-regressive alpha-stable processes [10], two-dimensional (2D) particle filter realisation in Markov Random Field (MRF)-modelled images [11] and astrophysical source separation [12].

In the problem setting, it is assumed that the stochastic models of the background and the event processes are known. The activation times of the event processes are unknown and the events are superimposed on a background process. The background and event processes can be modelled as auto-regressive (AR) processes, respectively. The AR-process model is widely used in speech and audio processing [10, 13, 14]. However, because of the inherent use of SMC methods, it should be pointed out that the methodology proposed is applicable for signals which have other parametric models. An important feature of the proposed method is that only single-channel data containing background and the event signals are assumed to be observed. The task of the proposed approach is to estimate the hidden background and event processes and to detect the event activation–deactivation times. In our previous work [1], a framework for superimposed event detection is introduced based on SMC methods, where the applicability is shown by synthetic data-based experiments.

In this paper, we generalise this framework and show that the problem can be modelled as a jump Markov system (JMS) [15–18]. JMSs can be regarded as a generalisation of

Hidden Markov Models (HMMs). JMSs introduce an additional model identification parameter represented by a discrete random variable. This additional parameter extends the state space by one. However, we show that the superimposed event detection problem can be transformed to a standard Markov Chain representation with state switch probabilities. As a result, there is no need for additional model identification parameters. This reduces the space and computational requirements of the particle filtering method. The novelty of our approach is to combine JMS as in [18] and particle filtering for superimposed event detection and to exploit the properties, which are specific to this type of problem.

We compare the proposed method to a change point detection method [2], which can be regarded as a simple event detection method. Change point detection determines the time point where the estimated states undergo a switch from the current model to another possible model [3]. In [2], a distinct particle filter is implemented for each model in parallel. Change point detection is then realised by computing logarithmic Likelihood ratios (LLR) for each model, using estimated states. In contrast, the proposed approach only uses a single filter which is computationally more efficient. Besides calculating the LLR, we also estimate the event probability using the posterior probability density approximation. In addition, our approach provides a solution to the online source separation problem [9, 19, 20] provided that good models of the background and the event signals are available.

We introduce experiments for superimposed event detection containing real audio signals and their superpositions.

2 Framework for superimposed event detection

The framework we consider is a generalisation of the framework presented in [1]. The background signal,

denoted by b_t , is superimposed by K event signals, denoted by $z_t^{(1)}, \dots, z_t^{(K)}$. All signals are assumed to be mutually independent. The background signal as well as the event signals are not observed directly. The only observation available is based on the superposition of the signals. We assume that the background signal as well as the event signals can be represented by stochastic models

$$b_{t+1} = f_b(b_t, v_t), \quad z_{t+1}^{(i)} = f_{z^{(i)}}(z_t^{(i)}, \mu_t^{(i)}), \quad i \in \{1, \dots, K\} \quad (1)$$

where v_t and $\mu_t^{(i)}$ are independent random variables. The corresponding transition densities are denoted by

$$p(b_{t+1}|b_t) \quad \text{and} \quad p(z_{t+1}^{(i)}|z_t^{(i)}), \quad i \in \{1, \dots, K\} \quad (2)$$

The event signals may occur in parallel, each with different start and stop times. As a result, each event signal $z_t^{(i)}$ is assumed to be only present for sometime window

$$T_E^{(i)} = [t_0^{(i)}, t_1^{(i)}] \quad (3)$$

The superposition of the signals yields a compound signal s_t

$$s_t = b_t + \sum_{i=1}^K \alpha_t^{(i)} z_t^{(i)}, \quad \alpha_t^{(i)} = \begin{cases} 1, & t \in T_E^{(i)} \\ 0, & \text{else} \end{cases} \quad (4)$$

As the observation, denoted by the scalar y_t , is based only on this superposition, it can be written as a function of s_t

$$y_t = g(s_t, w_t) \quad (5)$$

where w_t is a random variable, independent of v_t and $\mu_t^{(i)}$. The superpositional signal s_t contains all the information about the hidden states. The event switch parameters $\alpha_t^{(i)}$ are random variables obeying a stationary, discrete and finite first-order Markov chain taking values in $\{0, 1\}$. The transition probabilities can be denoted by

$$P_{m,n}^{(i)} = P(\alpha_{t+1}^{(i)} = m | \alpha_t^{(i)} = n), \quad m, n \in \{0, 1\} \quad (6)$$

The task is to detect the event, that is,

1. detect the presence of superpositional event signals $z_t^{(i)}$, and
2. estimate the event signals $z_t^{(i)}$ and the background signal b_t .

It is assumed that the state transition models of the background signal as well as the event signals are known.

According to the JMS methodology found in [18], the following state vector \mathbf{x}_t can be used

$$\mathbf{x}_t = (b_t, z_t^{(1)}, \dots, z_t^{(K)}, \alpha_t^{(1)}, \dots, \alpha_t^{(K)}) \quad (7)$$

However, the special JMS structure of the superimposed event detection problem allows the reduction to a standard Markov chain by discarding the model identification parameters $\alpha_t^{(i)}$ and modifying the event state transition

density $p(z_{t+1}^{(i)}|z_t^{(i)}) \rightarrow \tilde{p}(z_{t+1}^{(i)}|z_t^{(i)})$ as follows

$$\begin{aligned} \tilde{p}(z_{t+1}^{(i)}|z_t^{(i)}) &= \delta[z_t^{(i)}](P_{0,0}^{(i)} \delta(z_{t+1}^{(i)}) + P_{1,0}^{(i)} p(z_{t+1}^{(i)}|z_t^{(i)})) \\ &\quad + (1 - \delta[z_t^{(i)}])(P_{0,1}^{(i)} \delta(z_{t+1}^{(i)}) + P_{1,1}^{(i)} p(z_{t+1}^{(i)}|z_t^{(i)})) \end{aligned} \quad (8)$$

where $\delta(\cdot)$ is the Dirac substitution and $\delta[\cdot]$ is the Kronecker delta defined by

$$\delta[x] = \begin{cases} 1, & x = 0 \\ 0, & \text{else} \end{cases} \quad (9)$$

The Dirac substitution $\delta(z_{t+1}^{(i)})$ produces exact zeros as ‘no-event’ samples. In this way, the information about the model, that is, whether the state describing an event is present or not, can be directly determined from the corresponding state vector component. Therefore there is no need for the parameters $\alpha_t^{(i)}$. As a result, the dimension of the state vector decreases from $2K + 1$ to $K + 1$. The state vector is then given by

$$\mathbf{x}_t = (b_t, z_t^{(1)}, \dots, z_t^{(K)}) \quad (10)$$

From the assumption about the statistical independence of the background signal and the event signal follows for the state transition density

$$p(\mathbf{x}_{t+1}|\mathbf{x}_t) = p(b_{t+1}, z_{t+1}^{(1)}, \dots, z_{t+1}^{(K)}|b_t, z_t^{(1)}, \dots, z_t^{(K)}) \quad (11)$$

$$= p(b_{t+1}|b_t) \prod_{i=1}^K \tilde{p}(z_{t+1}^{(i)}|z_t^{(i)}) \quad (12)$$

where $\tilde{p}(z_{t+1}^{(i)}|z_t^{(i)})$ is the modified conditional density in (8).

According to sequential importance resampling (SIR) [4] method for particle filtering, the particle positions are sampled from an importance density

$$\mathbf{x}_{n,t+1} \sim \pi(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{y}_{1:t+1}) \quad (13)$$

at each time step t . The unnormalised weights $\omega_{n,t+1}$ are determined by

$$\omega_{n,t+1} = \omega_{n,t} \frac{p(\mathbf{y}_{t+1}|\mathbf{x}_{n,t+1})p(\mathbf{x}_{n,t+1}|\mathbf{x}_{n,t})}{\pi(\mathbf{x}_{n,t+1}|\mathbf{x}_{n,t}, \mathbf{y}_{1:t+1})} \quad (14)$$

Normalised weights $\tilde{\omega}_{n,t+1}$ are defined by

$$\tilde{\omega}_{n,t+1} = \frac{\omega_{n,t+1}}{\sum_{n=1}^N \omega_{n,t+1}} \quad (15)$$

2.1 Event detection

We introduce an event indicator I_t , which determines the presence of a new event occurring at time t . The event indicator is calculated by evaluating the event probability from the approximation $\sum_{n=1}^N \tilde{\omega}_{n,t} \delta(\mathbf{x}_t - \mathbf{x}_{n,t})$ of the posterior density $p(\mathbf{x}_t|y_t)$, where $\tilde{\omega}_{n,t}$ is normalised importance weight and n is particle number. Particles representing the posterior can be divided into $K + 1$ groups consisting of one ‘no event’ class and K ‘event’ classes for each event signal $z_t^{(i)}$. The detections are made by checking each event state $z_t^{(i)}$. If $z_t^{(i)} = 0$, this means there is no event. If it is different from zero, then there must be an additive signal. At time index t ,

the event probability $P_{Et}^{(i)}$ for event $z^{(i)}$ is then determined by

$$P_{Et}^{(i)} = \sum_{n=1}^N (1 - \delta[z_{n,t}^{(i)}]) \cdot \tilde{\omega}_{n,t} \quad (16)$$

The event indicator $I_t^{(i)}$ for event $z^{(i)}$ is defined by

$$I_t^{(i)} = \begin{cases} 0, & P_{Et}^{(i)} < \frac{1}{2} \\ 1, & \text{else} \end{cases} \quad (17)$$

This means, an event is declared ‘on’ with $I_t^{(i)} = 1$ iff the event probability is greater than 0.5, otherwise the event is declared ‘off’ with $I_t^{(i)} = 0$.

2.2 Signal estimation

The separation of the background signal and the event signals, respectively, is done by using the posterior probability density function (PDF) approximation. After each filter step, we simply calculate the corresponding expectations using the particle set. The estimated background signal \hat{b}_t is determined by

$$\hat{b}_t = E[b_t] = \sum_{n=1}^N b_{n,t} \cdot \tilde{\omega}_{n,t} \quad (18)$$

Similarly, the estimated event signal $\hat{z}_t^{(i)}$ is determined by

$$\hat{z}_t^{(i)} = E[z_t^{(i)}] = \sum_{n=1}^N z_{n,t}^{(i)} \cdot \tilde{\omega}_{n,t} \quad (19)$$

In the literature, multi-channel source separation by particle filtering can be found in [9]. In contrast, the technique presented in this paper can achieve separation from a single-channel observation by exploiting the knowledge about the parametric models of the background and event signals.

2.3 Possible applications

The proposed framework is applicable wherever two or more signals are combined into one single signal and the corresponding models are known or can be trained. As an example, several audio signals are superimposed and the task is to detect whether a certain audio signal is present. In another example, in image processing, several textures can be superimposed, for example, background and smoke, and the task is to detect the smoke.

3 Experiments

We present experiments using audio signals involving the detection of audio (including speech) events and the separation of corresponding signals. The event signals are

superimposed on a background audio signal. In this experiment, we use ‘Flute’ and ‘Pad’ signals as background signals. The signals ‘Piano’, ‘Speech’ and ‘Trumpet-Wah’ are considered as event signals. The speech signal is generated by the spoken word ‘tram’. All combinations yield six different superimposed signals.

Detection error rates (false-positive and false-negative alarms) as well as signal estimation errors for both the estimated background and event signals are determined. The measurements are taken at three different observation noise variances σ_y^2 . At each setting, the filter run is independently repeated 50 times and the results are averaged.

In the following, the implementation of audio event detection based on the proposed framework is described. The observation y_t is modelled as a superposition of the hidden states

$$y_t = b_t + z_t + v_t \quad (20)$$

where v_t is an i.i.d. normal random variable. Thus, the 2D hidden state is determined by

$$\mathbf{x}_t = (b_t, z_t) \quad (21)$$

The state transitions of both the background and the event processes are modelled by an AR(M) process with additive normal distributed noise. Therefore the according distributions are conditional on M previous states

$$\begin{aligned} p(b_{n,t+1}|b_{n,1:t}) &= p(b_{n,t+1}|b_{n,t-M:t}) \\ p(z_{n,t+1}|z_{n,1:t}) &= p(z_{n,t+1}|z_{n,t-M:t}) \end{aligned} \quad (22)$$

where n is the sample number and $\mathcal{N}(\cdot)$ represents a normal PDF of an i.i.d. random variable. The AR-coefficients β_i, ζ_i are trained from corresponding clean signals using the linear least squares method [21]. The AR-process order is chosen to be $M = 60$. Table 1 shows the variances of the AR(60) representations of corresponding signals.

We assume equal model switching probabilities $P_{0,1} = P_{1,0} = 10^{-4}$. This means that a switch of the model (e.g. event ‘off’ \rightarrow ‘on’) is very unlikely. On the other hand, it is very likely to keep the previous state as the current state (e.g. event ‘off’), which is reasonable for such signals with a great number of samples.

The observation model PDF is assumed to be the Gaussian $\mathcal{N}(0, \sigma_y^2)$. In this particular case with Gaussian observation and transition PDFs and a linear observation model (20), an optimal importance density [4] given by

$$\pi(\mathbf{x}_{t+1}|\mathbf{x}_t, y_{t+1}) = \pi(b_{t+1}, z_{t+1}|b_t, z_t, y_{t+1}) \quad (23)$$

can be determined from the ML-estimation of the parameters b_t and z_t , which exploits the information about the current observation. This has the advantage that the same filter performance can be achieved with a smaller number of samples.

Table 1 Error variances of corresponding AR(60) process models, determined from the prediction errors on clean signals

	Flute (F), dB	Pad (P), dB	Piano (Pi), dB	Trumpet-Wah (Tr), dB	Speech (Sp), dB
PSNR	62.97	49.35	56.29	53.7	35.8

First two signals ‘Flute’ (F) and ‘Pad’ (P) are used as background signals, whereas the last three signals ‘Piano’ (Pi), ‘Trumpet-Wah’ (Tr) and ‘Speech’ (Sp) are used as event signals

For the case of no event, the optimal importance density for the background signal is deduced from maximising the following term

$$\bar{b}_{n,t+1} = \arg \max_{b'_{n,t+1}} \mathcal{N}(y_{t+1} - b'_{n,t+1}, \sigma_y^2) \mathcal{N}(\tilde{b}_{n,t+1} - b'_{n,t+1}, \sigma_b^2) \quad (24)$$

where $\tilde{b}_{n,t+1} = \sum_j^M \beta_j b_{n,t-M+j}$. The term (24) is maximal for

$$\bar{b}_{n,t+1} = \frac{\sigma_y^{-2} y_{t+1} + \sigma_b^{-2} \tilde{b}_{n,t+1}}{\sigma_y^{-2} + \sigma_b^{-2}} \quad (25)$$

The corresponding variance σ_b^2 of $\bar{b}_{n,t+1}$ is

$$\sigma_b^2 = \frac{1}{\sigma_y^{-2} + \sigma_b^{-2}} \quad (26)$$

In case event signals are present, we consider only one single event signal for simplicity. The generalisation to many event signals is straightforward. The following Likelihood has to be maximised by solving for $\bar{b}_{n,t+1}$ and $\bar{z}_{n,t+1}$ simultaneously

$$(\bar{b}_{n,t+1}, \bar{z}_{n,t+1}) = \arg \max_{(b'_{n,t+1}, z'_{n,t+1})} \{ \mathcal{N}(y_{t+1} - b'_{n,t+1} - z'_{n,t+1}, \sigma_y^2) \cdot \mathcal{N}(\tilde{b}_{n,t+1} - b'_{n,t+1}, \sigma_b^2) \cdot \mathcal{N}(\tilde{z}_{n,t+1} - z'_{n,t+1}, \sigma_z^2) \} \quad (27)$$

where $\tilde{z}_{n,t+1} = \sum_j^M \xi_j z_{n,t-M+j}$. The solution for $(\bar{b}_{n,t+1}, \bar{z}_{n,t+1})$ is given by

$$\begin{pmatrix} \bar{b}_{n,t+1} \\ \bar{z}_{n,t+1} \end{pmatrix} = A \begin{pmatrix} \sigma_y^{-2} y_{t+1} + \sigma_b^{-2} \tilde{b}_{n,t+1} \\ \sigma_y^{-2} y_{t+1} + \sigma_z^{-2} \tilde{z}_{n,t+1} \end{pmatrix} \quad (28)$$

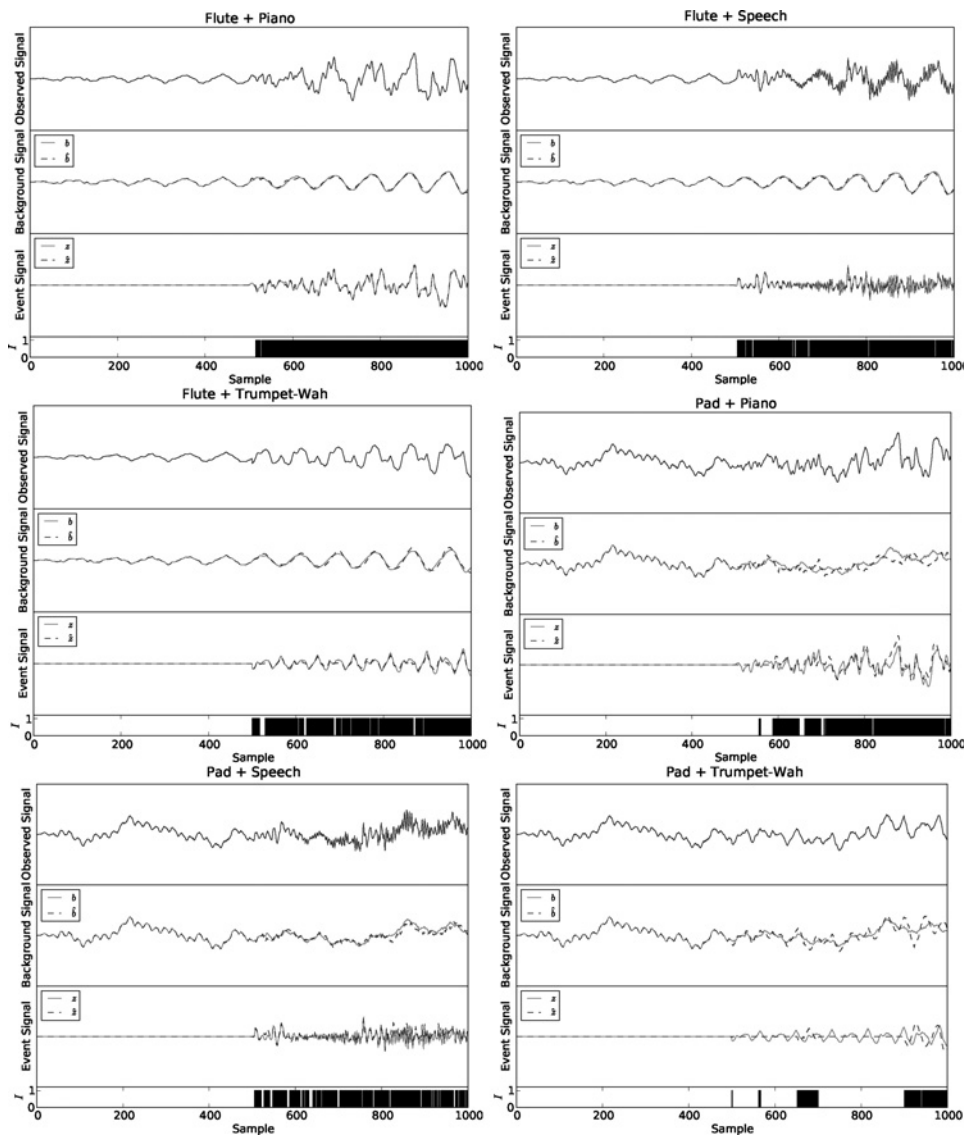


Fig. 1 Typical estimation results: the observed signal (top of each window), true b_t and estimated \hat{b}_t background signals (middle of each window) and true z_t and estimated \hat{z}_t event signals (bottom of each window)

Bar chart (very bottom of each window) represents the event indicator parameter I_t defined in (17). These graphs show that the events are always clearly detected, starting from sample no. 500

where

$$A = \begin{pmatrix} \sigma_y^{-2} + \sigma_b^{-2} & \sigma_y^{-2} \\ \sigma_y^{-2} & \sigma_y^{-2} + \sigma_z^{-2} \end{pmatrix}^{-1} \quad (29)$$

The corresponding covariance matrix is $C = AA^{-1}A^T = A$. The utilised importance distribution for the ‘event off’ case is specified by

$$\pi_0(\mathbf{x}_{n,t+1} | \mathbf{x}_{n,t}, y_{n,t}) = \mathcal{N}(\bar{b}_{n,t+1}, \sigma_b^2) \delta(z_{n,t}) \quad (30)$$

and for the ‘event on’ case

$$\pi_1(\mathbf{x}_{n,t+1} | \mathbf{x}_{n,t}, y_{n,t}) = \mathcal{N}((\bar{b}_{n,t+1}, \bar{z}_{n,t+1})^T, C) \quad (31)$$

Importance sampling is done by using a uniformly distributed discrete random variable $u_{n,t} \sim U(\{0, 1\})$, deciding whether the event is ‘on’ or ‘off’

$$\pi(\mathbf{x}_{n,t+1} | \mathbf{x}_{n,t}, y_{n,t}) = \delta[u_{n,t}] \pi_1(\mathbf{x}_{n,t+1} | \mathbf{x}_{n,t}, y_{n,t}) + (1 - \delta[u_{n,t}]) \pi_0(\mathbf{x}_{n,t+1} | \mathbf{x}_{n,t}, y_{n,t}) \quad (32)$$

In order to compare the proposed method with the parallel filter approach using the logarithmic Likelihood ratio (LLR) presented in [2], the sum of LLRs are determined from the state estimates in two parallel filter runs. The first run corresponds to the ‘event off’ model, yielding states $\mathbf{x}_t^{(\alpha=0)}$. The second filter run corresponds to the ‘event on’ model, which yields the states $\mathbf{x}_t^{(\alpha=1)}$.

The LLR is defined by

$$S_t^L = \sum_{i=t}^L \log \left(\frac{p(y_i | \hat{\mathbf{x}}_i^{(\alpha_i=1)})}{p(y_i | \hat{\mathbf{x}}_i^{(\alpha_i=0)})} \right) \quad (33)$$

where L is a positive integer and $p(y_i | \hat{\mathbf{x}}_i^{\alpha_i})$ is the Likelihood. The event is declared ‘on’ for

$$S_t^L > \tau_{LLR} \quad (34)$$

otherwise it is declared ‘off’. The threshold τ_{LLR} must be determined a priori for each case. The two filters are run with $N/2$ particles each, so that the computational expense is comparable to the proposed single-filter approach with N particles.

Results of the instrumental and natural sound mixing experiments with one background and one event signal are shown below. All audio samples are 16-bit wide and sampled at 44.1 kHz. For convenience, the sample value range is scaled to $[-1, 1]$. All experiments are done using audio data containing 1000 samples. The event signal is always activated from sample number 500 and lasts for 500 samples. In Fig. 1, typical detection and signal estimation results are shown for all combinations of the background signals ‘Flute’ and ‘Pad’ with the event signals ‘Piano’, ‘Speech’ and ‘Trumpet-Wah’. Corresponding signal estimation error variances as well as false-positive and false-negative event detection probabilities at various observation noise levels and a sample count of $N = 100$ are shown in Tables 2 and 3.

Table 2 Superposition of background signal Flute with event signals Piano, Speech and Trumpet-Wah: estimation

Event signal	σ_y	MSE (\hat{b}_t) (PSNR, dB)	MSE (\hat{z}_t) (PSNR, dB)	e^+	e^-
Pi	5×10^{-4}	5.5×10^{-4} (39.2 ± 2.1)	5.5×10^{-4} (39.2 ± 2.1)	$4.0 \times 10^{-5} \pm 2.0 \times 10^{-4}$	$2.0 \times 10^{-2} \pm 5.0 \times 10^{-3}$
	1×10^{-3}	5.3×10^{-4} (39.2 ± 1.8)	5.3×10^{-4} (39.2 ± 1.8)	$1.0 \times 10^{-4} \pm 3.0 \times 10^{-4}$	$2.0 \times 10^{-2} \pm 1.0 \times 10^{-3}$
	2×10^{-3}	6.7×10^{-4} (38.6 ± 2.6)	6.7×10^{-4} (38.6 ± 2.6)	$1.0 \times 10^{-4} \pm 3.5 \times 10^{-4}$	$2.0 \times 10^{-2} \pm 1.0 \times 10^{-2}$
Sp	5×10^{-4}	2.3×10^{-4} (42.9 ± 2.0)	2.3×10^{-4} (42.9 ± 2.0)	0.0 ± 0.0	$9.0 \times 10^{-3} \pm 2.0 \times 10^{-3}$
	1×10^{-3}	2.3×10^{-4} (42.7 ± 1.7)	2.3×10^{-4} (42.7 ± 1.7)	$2.0 \times 10^{-4} \pm 6.0 \times 10^{-4}$	$5.0 \times 10^{-3} \pm 3.0 \times 10^{-12}$
	2×10^{-3}	2.5×10^{-4} (42.7 ± 2.4)	2.5×10^{-4} (42.7 ± 2.4)	$1.0 \times 10^{-4} \pm 3.0 \times 10^{-4}$	$5.0 \times 10^{-3} \pm 7.0 \times 10^{-12}$
Tr	5×10^{-4}	7.7×10^{-4} (37.8 ± 2.3)	7.7×10^{-4} (37.8 ± 2.3)	$2.0 \times 10^{-2} \pm 6.0 \times 10^{-2}$	$3.0 \times 10^{-2} \pm 1.0 \times 10^{-2}$
	1×10^{-3}	6.5×10^{-4} (38.3 ± 1.8)	6.5×10^{-4} (38.3 ± 1.8)	$3.4 \times 10^{-4} \pm 2.0 \times 10^{-3}$	$2.0 \times 10^{-2} \pm 1.0 \times 10^{-2}$
	2×10^{-3}	4.7×10^{-4} (39.7 ± 1.8)	4.7×10^{-4} (39.7 ± 1.8)	$2.0 \times 10^{-4} \pm 6.0 \times 10^{-4}$	$2.0 \times 10^{-3} \pm 5.0 \times 10^{-3}$

MSE’s and event detection false positive e^+ and false negative e^- alarm probabilities for several observation noise variances σ_y^2 and $N = 100$

Table 3 Superposition of background signal Pad with event signals Piano, Speech and Trumpet-Wah: estimation

Event signal	σ_y	MSE (\hat{b}_t) (PSNR, dB)	MSE (\hat{z}_t) (PSNR, dB)	e^+	e^-
Pi	5×10^{-4}	8.8×10^{-3} (27.2 ± 1.9)	8.8×10^{-3} (27.2 ± 1.9)	$2.0 \times 10^{-5} \pm 1.4 \times 10^{-4}$	$5.0 \cdot 10^{-2} \pm 4.0 \cdot 10^{-2}$
	1×10^{-3}	8.5×10^{-3} (27.5 ± 2.1)	8.5×10^{-3} (27.5 ± 2.1)	0.0 ± 0.0	$6.0 \times 10^{-2} \pm 3.0 \times 10^{-2}$
	2×10^{-3}	8.2×10^{-3} (27.2 ± 1.5)	8.2×10^{-3} (27.2 ± 1.5)	$1.0 \times 10^{-2} \pm 6.0 \times 10^{-2}$	$1.0 \times 10^{-1} \pm 4.0 \times 10^{-2}$
Sp	5×10^{-4}	8.6×10^{-4} (36.7 ± 0.8)	8.6×10^{-4} (36.7 ± 0.8)	$2.0 \times 10^{-5} \pm 2.0 \times 10^{-4}$	$5.0 \times 10^{-2} \pm 3.0 \times 10^{-12}$
	1×10^{-3}	8.6×10^{-4} (36.7 ± 0.7)	8.6×10^{-4} (36.7 ± 0.7)	0.0 ± 0.0	$5.0 \times 10^{-3} \pm 3.0 \times 10^{-12}$
	2×10^{-3}	8.6×10^{-4} (36.8 ± 0.9)	8.6×10^{-4} (36.8 ± 0.9)	$2.0 \times 10^{-5} \pm 1.4 \times 10^{-4}$	$5.0 \times 10^{-3} \pm 2.0 \times 10^{-4}$
Tr	5×10^{-4}	6.0×10^{-3} (30.5 ± 3.4)	6.0×10^{-3} (30.5 ± 3.4)	$1.0 \times 10^{-2} \pm 6.0 \times 10^{-2}$	$4.0 \times 10^{-2} \pm 3.0 \times 10^{-2}$
	1×10^{-3}	4.3×10^{-3} (31.1 ± 3.2)	4.3×10^{-3} (31.1 ± 3.2)	$2.0 \times 10^{-2} \pm 8.0 \times 10^{-2}$	$5.0 \times 10^{-2} \pm 4.0 \times 10^{-2}$
	2×10^{-3}	1.0×10^{-2} (29.3 ± 5.2)	1.0×10^{-2} (29.3 ± 5.2)	0.0 ± 0.0	$1.5 \times 10^{-1} \pm 1.0 \times 10^{-1}$

MSE’s and event detection false positive e^+ and false negative e^- alarm probabilities for several observation noise variances σ_y^2 and $N = 100$

Table 4 LLR-based parallel filter approach compared with the proposed approach

Signal	Mean PSNR, dB	LLR e^+	LLR e^-	τ_{LLR}
Parallel filter approach				
F + Pi	28.2 ± 2.0	$2.0 \times 10^{-4} \pm 1.0 \times 10^{-3}$	$1.0 \times 10^{-1} \pm 8.0 \times 10^{-4}$	20
F + Sp	41.2 ± 1.7	$5.0 \times 10^{-4} \pm 0.0 \times 10^{-0}$	$5.0 \times 10^{-2} \pm 0.0 \times 10^{-0}$	20
F + Tr	38.1 ± 1.8	0.0 ± 0.0	$1.5 \times 10^{-2} \pm 1.0 \times 10^{-2}$	3
P + Pi	26.6 ± 0.1	$1.5 \times 10^{-1} \pm 4.0 \times 10^{-2}$	$1.4 \times 10^{-1} \pm 4.0 \times 10^{-2}$	0.2
P + Sp	36.0 ± 0.7	$2.0 \times 10^{-2} \pm 1.6 \times 10^{-2}$	$1.1 \times 10^{-1} \pm 3.0 \times 10^{-2}$	0.7
P + Tr	31.7 ± 0.7	$1.4 \times 10^{-1} \pm 4.0 \times 10^{-2}$	$2.7 \times 10^{-1} \pm 4.0 \times 10^{-2}$	0.25
Proposed approach				
F + Pi	39.3 ± 2.5	0.0 ± 0.0	$1.6 \times 10^{-2} \pm 2.0 \times 10^{-3}$	20
F + Sp	42.9 ± 2.0	0.0 ± 0.0	$6.0 \times 10^{-3} \pm 4.0 \times 10^{-4}$	20
F + Tr	39.4 ± 1.3	0.0 ± 0.0	0.0 ± 0.0	20
P + Pi	27.4 ± 1.2	0.0 ± 0.0	$7.0 \times 10^{-2} \pm 3.0 \times 10^{-2}$	0.5
P + Sp	36.6 ± 0.8	$3.0 \times 10^{-3} \pm 5.0 \times 10^{-3}$	$9.0 \times 10^{-3} \pm 4.0 \times 10^{-3}$	1
P + Tr	30.8 ± 2.9	$6.0 \times 10^{-4} \pm 4.0 \times 10^{-3}$	$8.0 \times 10^{-2} \pm 3.0 \times 10^{-2}$	0.5

Better results are highlighted in boldface. In both approaches the LLRs are used as detection indicators. The results are achieved with a total number of $N = 100$ particles and an observation noise with $\sigma_v = 5 \times 10^{-4}$. The LLR-based filter pair is run with 2×50 particles

In Table 4, the proposed method is compared to the parallel filter approach. In nearly all parts of the experiments, the proposed approach yields improved results over the parallel filter solution, in both the detection accuracy and estimation accuracy. The main reason for this overall improvement is that in the proposed approach, only samples that match the current model best survive the resampling process. Therefore the sample history contains a higher density of ‘model-conform’ samples than in the parallel filter method. This leads to a more effective approximation of the posterior density in general.

4 Conclusions

The proposed framework presents event detection using particle filtering as a JMS. In the problem setting, a background signal and one or more superpositional event signals yield the observed single-channel signal. In other words, the observed signal is the result of the superposition of all signals. The event signals are only present for some time frames, independently of each other. The start and stop times of the events are unknown a priori. It is assumed that all signals are mutually statistically independent and that parametric models are available to represent the signals stochastically.

It is shown that the special structure of the problem at hand enables reduction of the JMS to a standard Markov chain system by transforming the event state transition densities. In this way, the state space dimension is decreased by the number of the recognised event signals, which leads to increased space and computational efficiency.

The proposed method is compared with the state-of-the-art change point detection method using one distinct filter in parallel for each possible background–event signal combination and LLR measures. Compared with the parallel filter approach, the proposed approach yields improved results, both in signal estimation accuracy and detection accuracy. Instead of the requirement of a distinct filter run for each possible background–event signal combination, the proposed approach only requires one single-particle filter run, enabling a simpler implementation and less overhead.

5 Acknowledgments

This work was partly published in IEEE Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP2008) and supported by the European Commission Seventh Framework Program with EU Grant: 244088 (FIRESENSE-Fire Detection and Management through a Multi-Sensor Network for the Protection of Cultural Heritage Areas from the Risk of Fire and Extreme Weather Conditions).

6 References

- Urfalioglu, O., Kuruoglu, E.E., Cetin, A.E.: ‘Framework for online superimposed event detection by sequential Monte Carlo methods’. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, 2008, pp. 2125–2128
- Li, P., Kadiramanathan, V.: ‘Particle filtering based likelihood ratio approach to fault diagnosis of nonlinear stochastic systems’, *IEEE Trans. Syst., Man, Cybern.*, 2001, **31**, pp. 337–343
- Andrieu, C., Doucet, A., Singh, S.S., Tadic, V.B.: ‘Particle methods for change detection, system identification, and control’, *Proc. IEEE*, 2004, **92**, pp. 423–438 [Online], available at: citeseer.ist.psu.edu/andrieu04particle.html
- Doucet, A.: ‘On sequential simulation-based methods for bayesian filtering’. Technical report CUED/F-INFENG/TR.310, 1998
- Liu, J.S.: ‘Monte Carlo strategies in scientific computing’ (Springer, 2001)
- Doucet, A., de Freitas, J.F.G., Gordon, N.J.: ‘Sequential Monte Carlo methods in practice’ (Springer, 2001)
- Zotkin, D., Duraiswami, R., Davis, L.S.: ‘Multimodal 3-d tracking and event detection via the particle filter’. IEEE Int. Conf. on Computer Vision, Workshop on Detection and Recognition of Events in Video (ICCV-EVENT), 2001
- Xiao, X., Zhang, X.: ‘Event detection based on hierarchical event fusion’. Int. Conf. on Environmental Science and Information Application Technology, 2009, vol. 2, pp. 483–486
- Costagli, M., Kuruoglu, E.E.: ‘Image separation using particle filters’, *Digit. Signal Process.*, 2007, **17**, (5), pp. 935–946
- Gencaga, D., Ertuzun, A., Kuruoglu, E.E.: ‘Modeling of non-stationary autoregressive alpha-stable processes by particle filters’, *Digit. Signal Process.*, 2008, **18**, (3), pp. 465–478
- Kayabol, K., Kuruoglu, E., Sankur, B.: ‘2d particle filter realization in MRF modelled images’. IEEE 16th Signal Processing, Communication and Applications Conf., 2008, SIU 2008, April 2008, pp. 1–4
- Costagli, M., Kuruoglu, E.E., Ahmed, A.: ‘Astrophysical source separation using particle filters’, *Indep. Compon. Anal. Blind Signal Separation*, 2004, **3195**, pp. 930–937
- Quatieri, T.F.: ‘Discrete-time speech signal processing: principles and practice’ (Prentice-Hall, 2001)

- 14 Gencaga, D., Kuruoglu, E.E., Ertuzun, A., Yildirim, S.: 'Estimation of time varying AR $S\alpha S$ processes using gibbs sampling', *Signal Process.*, 2008, **88**, (10), pp. 2564–2572
- 15 Doucet, A., Gordon, N., Krishnamurthy, V.: 'Particle filters for state estimation of jump markov linear systems', *IEEE Trans. Signal Process.*, 2001, **49**, (3), pp. 613–624
- 16 Nicoli, M., Morelli, C., Rampa, V.: 'A jump markov particle filter for localization of moving terminals in multipath indoor scenarios', *IEEE Trans. Signal Process.*, 2008, **56**, (8), pp. 3801–3809
- 17 Driessen, H., Boers, Y.: 'An efficient particle filter for jump Markov nonlinear systems'. *Target Tracking 2004: Algorithms and Applications*, 2004, pp. 19–22
- 18 Andrieu, C., Davy, M., Doucet, A.: 'Efficient particle filtering for jump markov systems'. *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, 2002 (ICASSP'02)*, 2002, vol. 2, pp. 1625–1628
- 19 Kuruoglu, E.E., Bedini, L., Paratore, M.T., Salerno, E., Tonazzini, A.: 'Source separation in astrophysical maps using independent factor analysis', *Neural Netw.*, 2003, **16**, (3–4), pp. 479–491
- 20 Wilson, S.P., Kuruoglu, E.E., Salerno, E.: 'Fully bayesian source separation of astrophysical images modelled by mixture of Gaussians', *IEEE J. Sel. Top. Signal Process.*, 2008, **2**, (5), pp. 685–696
- 21 Oppenheim, A.V., Schaffer, R.W., Buck, J.R.: 'Discrete-time signal processing' (Prentice-Hall, 1999)