

# HYBRID BS-IS PARTICLE FILTER BASED ACOUSTIC SOURCE TRACKING ALGORITHM

T, Li and W, Ser

Center for Signal Processing, School of Electrical and Electronics Engineering  
Nanyang Technological University  
50 Nanyang Avenue, Singapore 639798, Singapore  
phone: + (65) 67905912, fax: + (65) 67912383, email: liti0001@ntu.edu.sg, ewser@ntu.edu.sg  
web: www.csp.eee.ntu.edu.sg

## ABSTRACT

*Tracking moving acoustic sources in a reverberant environment has been a challenging research goal for many years. The Bayesian Filtering approach, in particular the Bootstrap (BS) and the Importance Sampling (IS) based Particle Filter (PF) tracking algorithms, has been proposed recently and promising results have been obtained. While these two algorithms out-perform other algorithms, they can only implement the optimal importance function in respective sub-optimal form and suffer related drawbacks in tracking performances. A novel hybrid BS-IS Particle Filter based acoustic source tracking algorithm is proposed in this paper. The height of the target source is used to control a sampling switch importance function that allows the tracking algorithm to switch between using the BS based and the IS based algorithm. Numerical result shows that the proposed algorithm is able to out-perform the BS and the IS based tracking algorithms in reverberant environments.*

## 1. INTRODUCTION

Acoustic source localization and tracking (ASLT) has been an active research area over the past two decades. Indeed many such techniques have been reported in the literature and several of such techniques have been used in applications such as hand-free audio conferencing, conversational interface, sound based healthcare monitoring, etc. One main problem faced by such techniques in practice is the reverberation effect encountered in a closed room where a target signal will arrive as multiple signals at the sensor. This makes it difficult, if not impossible, for the basic form of conventional acoustic source localization (ASL) algorithms such as the Steered Beam Former (SBF), the Generalized Cross Correlation (GCC) based Time-Delay Estimator (TBE), the high resolution spectral estimation methods (e.g. Capon, MUSIC), etc., to function effectively. One reason for the basic form of such techniques to fail is that they do not take into consideration past estimations made on the source signals and the time-varying nature of the source signals. In order to tackle this problem, the Particle Filter (PF) based tracking approaches have been proposed recently [1-6].

The PF (Particle Filter) approach works by applying the sequential Monte Carlo (SMC) method and the hidden Markov chain (HMC) assumption. The source location information is estimated through a collection of particles iteratively sampled from the state space (the enclosed reverberant environment in ASLT case). The PF based acoustic source tracking (AST) approach then updates sequentially the source location by taking into consideration both current array observation and previously estimated locations. Several PF based algorithms have been proposed in the literature, one of them is the Bootstrap (BS) PF algorithm [5]. The BS-PF is able to minimize the reverberation effects for source tracking. However, the limitation of the BS-PF is that it does not make use of the current observations that may be useful in indicating the true source location. Another PF based tracking algorithm is the Importance Sampling (IS) PF algorithm [6], which is capable of locating the source rapidly. The limitation of the IS-PF algorithm is, however,

that it discards any previous source location estimations in the sampling of new particles.

This paper proposes a novel Particle Filter based tracking algorithm that is capable of overcoming the shortcomings of the BS-PF and IS-PF algorithms. The proposed algorithm overcomes the shortcomings by introducing a source height controlled sampling switch technique. The novel sampling switch technique is able to adaptively adopt either BS-PF or IS-PF at each iteration based on the accuracy of the corresponding source height estimation. The important assumption is that the source height is always a constant in the entire ASLT process, and that any inaccurate height estimation is caused by the reverberation effects. Therefore, if the estimated source height indicates strong reverberation influence at certain iteration, the sampling switch will automatically adopt the BS-PF to minimize the reverberation effects. Otherwise, if the estimated source height indicates weak reverberation influence, the IS-PF will be adopted to accelerate the locating speed. The resultant hybrid BS-IS PF tracking algorithm hence successfully combines the fast locating capability of the IS-PF and the robust tracking capability of the BS-PF.

The remaining part of this paper is organized as follows. Section 2 gives a description on the formulation of the AST problem. Section 3 discusses two well known Particle Filter based source tracking algorithms. The proposed modified Particle Filter algorithm is presented in Section 4. The performance of the proposed algorithm is shown in Section 5 through some numerical examples. Finally, Section 6 concludes the findings presented in this paper.

## 2. PROBLEM FORMULATION FOR ACOUSTIC SOURCE TRACKING

The acoustic source tracking (AST) problem to be discussed in this paper assumes a single target source moving in a closed room where the reverberation effect is not negligible. Omnidirectional microphones are fixed at known locations on the surrounding walls.

The signals received by the array are organized into frames of equal length  $L$ . The objective of the AST problem here is to estimate the source locations based on the sampled signals,

$$y_m(k) = h_{totm}(k) * s(k) + n_m(k) \quad (1)$$

where  $y_m(k)$  is the signal received at the  $m$ th sensor and at the  $k$ th frame,  $h_{totm}(k)$  is the channel (direct-path together with multi-path) impulse response between the source and the  $m$ th sensor,  $s(k)$  is the source signal,  $*$  is the convolution operator of the source signal and the channel impulse response, and  $n_m(k)$  denotes the uncorrelated background noise.

The acoustic source's instantaneous location can be defined by its Cartesian coordinates  $(x_k, y_k)$ . Let the source's instantaneous velocity be  $\dot{x}_k$  and  $\dot{y}_k$  along the  $x$ -axis and  $y$ -axis respectively, the source's *state variable* at each time step  $k$  in the state space can be expressed as,

$$X_k = [x_k \ y_k \ \dot{x}_k \ \dot{y}_k]^T \quad (2)$$

With equation (2), the objective of source location estimation is now re-formulated to estimate the source's state variable (but the velocity components may not be emphasized in the general ASLT case). Let  $Y_k$ , be the (ASLT algorithms measured) observation variable. The AST problem can now be handled by the Bayesian filtering approach based on the HMC (Hidden Markov Chain) model. The state and observation variables, and their relationship can be described by the Anderson and Moore's equations [8].

$$X_k = g(X_{k-1}, u_k) \quad (3)$$

$$Y_k = h(X_k, v_k) \quad (4)$$

where  $g(\cdot)$  and  $h(\cdot)$  are the *transition function* and *observation function*. Both functions are not necessarily linear, and the noise terms  $u_k$  and  $v_k$  may also be non-Gaussian.

Let the posteriori probability density function,  $p(X_k|Y_{1:k})$ , be denoted as PDFp. The Bayesian filtering solution can now be applied to determine the PDFp, in two steps: "*prediction*" and "*update*", in a recursive manner,

$$p(X_k|Y_{1:k-1}) = \int p(X_k|X_{k-1})p(X_{k-1}|Y_{1:k-1})dX_{k-1} \quad (5)$$

$$p(X_k|Y_{1:k}) \propto p(Y_k|X_k)p(X_k|Y_{1:k-1}) \quad (6)$$

where in (5),  $p(X_k|Y_{1:k-1})$  is the *priori* PDF,  $p(X_k|X_{k-1})$  is the *transition density*, and  $p(X_{k-1}|Y_{1:k-1})$  is the PDFp that has been estimated at the previous  $(k-1)$ th time step; in (6),  $p(Y_k|X_k)$  denotes the *likelihood function*.

It is quite obvious that the two step approach effectively takes into consideration both the previous source location estimations and the current observations. The recursive manner of the two step approach allows the application of sequential Monte Carlo (SMC) technique, also known as the Particle Filtering (PF) technique, to sequentially estimate the desired PDFp. The source *state variable*,  $X_k$ , can be estimated from its PDFp as the Mean or the Mode. This constitutes the Bayesian filtering based AST solution.

### 3. ACOUSTIC SOURCE TRACKING USING PARTICLE FILTER

#### 3.1 General Concept of Particle Filter Theory

For AST problem, the state space is defined as the  $x$ - $y$  Cartesian plane of the entire enclosure inside which the acoustic source resides. The particle (state sample) is defined as the grid point that records a fixed  $(x, y)$  location inside the state space. In this way, at each iteration, new particle set of size  $N$ :  $\alpha_k^{(n)}$ , where  $n \in \{1, \dots, N\}$ , are sampled in the prediction step. The followed update step associate each particle with a weight  $\omega_k^{(n)}$  to denote its likelihood of representing true source location. The PDFp in the Bayesian filtering problem can be approximated by

$$p(X_k|Y_{1:k}) \approx \sum_{n=1}^N \omega_k^{(n)} \delta(X_k - \alpha_k^{(n)}) \quad (7)$$

where  $\delta(\cdot)$  denotes the Dirac delta function. The source *state variable* at  $k$ th time step can now be estimated by computing the Mean of the approximated PDFp, i.e.

$$\begin{aligned} \hat{X}_k &= \int X_k p(X_k|Y_{1:k}) dX_k \\ &\approx \int X_k \sum_{n=1}^N \omega_k^{(n)} \delta(X_k - \alpha_k^{(n)}) dX_k \\ &= \sum_{n=1}^N \omega_k^{(n)} \alpha_k^{(n)} \end{aligned} \quad (8)$$

This constitutes the general concept of PF, and several PF based AST applications can be found in the literature, hereby only two typical ones will be discussed.

#### 3.2 The Bootstrap Particle Filter Algorithm

The bootstrap Particle Filter (BS-PF) algorithm is originally developed by Gordon *et al.* in [5], and is famous for its low demands of computational power and reduced difficulties of implementation.

The general work flow of BS-PF follows the standard two steps Bayesian filtering approach in (5) and (6). However, the prediction step of BS-PF is simplified in the way that only the previous state *source state variable* estimation is considered. Thus, current state particles are sampled only according to the pre-defined source state *transition function*:

$$\alpha_k^{(n)} \propto g(\alpha_{k-1}^{(n)}, u_k) \quad (9)$$

In practical implementation, the Langevin process [1] was suggested by [6] to serve as the source state *transition function*.

$$X_k = \begin{bmatrix} 1 & 0 & aT_U & 0 \\ 0 & 1 & 0 & aT_U \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{bmatrix} \cdot X_{k-1} + \begin{bmatrix} bT_U & 0 \\ 0 & bT_U \\ b & 0 \\ 0 & b \end{bmatrix} \cdot \mathbf{u}_k \quad (10)$$

where  $\mathbf{u}_k$  is the Gaussian noise variable

$$\mathbf{u}_k \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \quad (11)$$

in which  $\mathcal{N}(\mu, \Sigma)$  represents the density of a multi-dimensional Gaussian random variable with mean vector  $\mu$  and covariance matrix  $\Sigma$ ,  $T_U$  is the time interval between particle filter's two adjacent updates, and the other model parameters are defined as

$$a = \exp(-\beta T_U) \quad (12)$$

$$b = \bar{v} \sqrt{1-a^2} \quad (13)$$

where  $\bar{v}$  denotes the steady-state velocity and  $\beta$  indicates the rate constant.

While in the update step, the weights of the newly sampled particles will be updated according to the *likelihood function* and be normalized so that the summation of all the weights equals to one:

$$\tilde{\omega}_k^{(n)} = p(Y_k|\alpha_k^{(n)}) \quad (14)$$

$$\omega_k^{(n)} = \tilde{\omega}_k^{(n)} / \sum_{n=1}^N \tilde{\omega}_k^{(n)} \quad (15)$$

The choice of the *likelihood function* in practical implementation is also widely open. For the AST purpose, the output of the ASL algorithm can be used directly as the *pseudo likelihood function*. For example, we may use the SRP-PHAT [7] to approximate the *likelihood function*:

$$p(Y_k|\alpha_k^{(n)}) = \int \sum_{m=1}^M \sum_{n=1}^M \frac{S_m(\omega) S_n^*(\omega)}{|S_m(\omega) S_n^*(\omega)|} \exp(j\omega(\tau_m - \tau_n)) d\omega \quad (16)$$

where  $S_m(\omega)$  and  $S_n(\omega)$  denote the Fourier transform of the signal at  $m$ th and  $n$ th sensor, the symbol  $*$  denotes the complex conjugate of the fourier transformed signal, and  $\tau_m, \tau_n$  are the respective direct path delays.

The main limitation of the BS-PF based AST algorithm is that during the prediction step, the *transition function* sampled particles are drawn from the entire state space instead of from some high potential regions indicated by the current observation  $Y_k$ . As a result, the BS-PF based AST algorithm may take extra time steps to locate the source at the initial stage of the tracking process. Despite this limitation, the *transition function* based particle sampling is able to minimize the reverberation effects, making the BS-PF based AST algorithm competitively reverberation-robust.

### 3.3 The Importance Sampling Particle Filter Algorithm

The importance sampling Particle Filter (IS-PF) is generally considered as an upgraded PF implementation of the rather simplified BS-PF [6].

In the IS-PF's prediction step, the specially introduced *importance function*  $q(\cdot)$  serves to sample particles  $\alpha_k^{(n)}$  from high potential state space regions indicated by the current *observation variable*  $Y_k$ . For the AST purpose, the state space regions associated with peak ASL algorithm output power can be denoted as the importance regions. Therefore, in practical implementation of the IS-PF, the *importance function* can be approximated by:

$$q_{is}(\alpha_k|Y_{1:k}) = p(\alpha_k|Y_k) \quad (17)$$

from which it is quite obvious that the simplified *importance function* only takes into consideration the current state observations, which can be the output of the SRP-PHAT algorithm.

The non-normalized importance weights can be calculated according to the same *pseudo likelihood function* as in (14) and (16), but with some update to avoid the degeneracy problem [8] (same problem also encountered by the BS-PF, which solved it by introducing an additional resampling step), i.e.

$$\tilde{\omega}_k^{(n)} = p(Y_k|\alpha_k^{(n)}) \cdot \min\left\{\frac{p(\alpha_k^{(n)}|Y_{1:k-1})}{q_{is}(\alpha_k^{(n)}|Y_{1:k})}, 1\right\} \quad (18)$$

The weights are then normalized in the same way as in (15).

The *importance function* based particle sampling makes the algorithm able to locate source rapidly, and this constitute the main advantage of the IS-PF based AST algorithm. However, due to the sub-optimal nature of the simplified *importance function*, the IS-PF algorithm's tracking performance is highly dependent on the accuracy of the current observations  $Y_k$ . If strong reverberation effects present, the observations may be affected and not be able to indicate potential particle sampling regions accurately. The IS-PF based AST algorithm may suffer from the inaccurate observation resulted low reverberation robustness during the tracking process.

## 4. PROPOSED HYBRID BS-IS PARTICLE FILTER ALGORITHM

### 4.1 Concept of the Proposed Hybrid Algorithm

In [9], to solve the weight variance caused degeneracy problem, the optimal *importance function* was derived as,

$$q_{opt}(\alpha_k|Y_{1:k}) = p(\alpha_k|\alpha_{k-1}, Y_k) \quad (19)$$

which clearly indicates that the sampling of current state particles should take into consideration of both previous state estimation and current state observation. The dual state consideration based sampling complies with the Bayesian filtering concept described in section 2.

The optimal *importance function*, however, suffers from two drawbacks: one is the ability to sample from both previous estimation and current observation simultaneously; the other one is the need to evaluate  $p(Y_k|\alpha_{k-1}) = \int p(Y_k|\alpha_k)p(\alpha_k|\alpha_{k-1})d\alpha_k$ , which generally has no analytic form. Therefore, only the sub-optimal form of the *importance function* may be applied in the IS-PF implementation, which may suffer from the reverberation effects if applied in the AST problem as discussed in section 3.3.

Besides the "current state observation only" *importance function* used in the IS-PF, if we treat the *transition function* used in the BS-PF prediction step as another simplified form of the optimal *importance function*, i.e.

$$q_{bs}(\alpha_k|Y_{1:k}) = p(\alpha_k|\alpha_{k-1}) \quad (20)$$

The "previous state estimation only" sub-optimal nature of the *importance function* used by the BS-PF may now contribute to the

limitation of the BS-PF based AST algorithm discussed in section 3.2.

Therefore, to develop fast and reverberation robust PF based AST algorithm that overcomes the shortcomings of the BS-PF and the IS-PF algorithms, we should sample the particles in the manner closely approximate the optimal *importance function*:

$$q(\alpha_k|Y_{1:k}) = p(\alpha_k|\alpha_{1:k-1})p(Y_k) \quad (21)$$

which can be interpreted as the sampling of current state particles should always take into consideration of previous estimations to guarantee the tracking algorithm's reverberation robustness; meanwhile, the sampling of new particles should also be subjected to the accuracy of the current state observation to effectively accelerate the locating speed.

Due to the fact that the accuracy of the observation is affected by the reverberations, which can be treated as the "observations" based on previous source states. We may therefore interpret the observation into two components, i.e.

$$p(Y_k) = p(Y_k|\alpha_{1:k})p(\alpha_{1:k}) + p(Y_k|\alpha_{1:k-1})p(\alpha_{1:k-1}) \quad (22)$$

where the first component represents accurate observation and includes valid information of current source state, sampling particles according to such observation component can be quite safe; the latter component represents inaccurate observation and only includes previous source state information, hence the sampling of particles should follow the *transition function* to avoid being misled by the spurious observation.

Because that at any discrete time step, the observation can only be accurate ( $p(Y_k|\alpha_{1:k})p(\alpha_{1:k}) \gg p(Y_k|\alpha_{1:k-1})p(\alpha_{1:k-1})$ ) or inaccurate ( $p(Y_k|\alpha_{1:k-1})p(\alpha_{1:k-1}) \gg p(Y_k|\alpha_{1:k})p(\alpha_{1:k})$ ). Intuitively, this would suggest a hybrid structure of the BS-PF and the IS-PF, and the accuracy of the current step observation is the key decision maker. For implementation purpose, an additional binary *sampling switch(SS) density*  $p_{ss}(Y_k)$  is needed to help decide whether the current step observation  $Y_k$  is accurate enough. The resulted *sampling switch importance function* is,

$$\begin{aligned} q_{ss}(\alpha_k|Y_{1:k}) &= p(\alpha_k|\alpha_{k-1}) \cdot (1 - p_{ss}(Y_k)) + p(\alpha_k|Y_k) \cdot p_{ss}(Y_k) \\ &= \begin{cases} p(\alpha_k|\alpha_{k-1}), & p_{ss}(Y_k) = 0 \\ p(\alpha_k|Y_k), & p_{ss}(Y_k) = 1 \end{cases} \end{aligned} \quad (23)$$

where  $p_{ss}(Y_k) = 0$  indicates inaccurate observation at  $k$ th time step, and  $p_{ss}(Y_k) = 1$  indicates accurate observation at  $k$ th time step.

### 4.2 Proposed Source Height Estimation based Hybrid BS-IS Algorithm

If we have two additional microphone arrays along the ceiling and one side wall of the room, the source height estimation (SHE), i.e. the source's ( $z$ ) Cartesian coordinate, can be obtained readily by applying the SRP-PHAT algorithm in the  $x$ - $z$  Cartesian plane and calculate the moving average of a series of reasonable estimations.

In most AST problems, the source's height is always a constant during the entire AST process, and the ideal SHE should be a horizontal line along the time axis. However, the accuracy of the SHE is also affected by the reverberation effects that have the same detrimental effects on the source location estimation, especially for small reverberant rooms. Unlike the inaccurate observation, the inaccurate SHE can be more easily detected, as any inaccurate SHE value may only fluctuate significantly along the constant source height level.

Assuming that the SHE is performed independently and repeatedly at each iterative step of the PF based AST process, and that the SHE accuracy is only affected by the reverberation effects, then any SHE that differs from the SHE steady state value to certain extent may indicate an inaccurate observation at the corresponding time step. So the SHE based *sampling switch decision* is defined as,

$$\begin{cases} p_{ss}(Y_k) = 0, & \text{if } |H_k - \bar{H}| > \delta_H \\ p_{ss}(Y_k) = 1, & \text{if } |H_k - \bar{H}| \leq \delta_H \end{cases} \quad (24)$$

where  $H_k$  is the SHE at  $k$ th step,  $\bar{H}$  is the Mean of all accurate SHE up to  $k$ th step, and  $\delta_H$  is the defined SHE error threshold that decides whether  $H_k$  is accurate.

The proposed SHE based Hybrid BS-IS Algorithm can now be summarized in the followed procedure:

1. Resampling(optional): resampling is performed if and only if the BS-PF is adopted at the previous  $(k-1)$ th step.
2. Source Height Estimations: estimate source height  $H_k$  with the additional microphone arrays and the ASL algorithm applied in the  $x$ - $z$  Cartesian plane.
3. Sampling Switch Decision: decide the binary value of the *sampling switch density*  $p_{ss}(Y_k)$  based on (24).
4. Particle Prediction: sample  $\alpha_k$  according to BS-PF:  $p(\alpha_k|\alpha_{k-1})$ , if  $p_{ss}(Y_k) = 0$ ; sample  $\alpha_k$  according to IS-PF:  $p(\alpha_k|Y_k)$ , if  $p_{ss}(Y_k) = 1$ .
5. Weight Update: update the new sampled particles' weights according to the designated weight update function of either BS-PF ( $p_{ss}(Y_k) = 0$ ) or IS-PF ( $p_{ss}(Y_k) = 1$ )
6. Source Location Estimation: estimate the source  $(x, y)$  Cartesian coordinates with the weighted particles:  $\hat{X}_k = \sum_{n=1}^N \omega_k^{(n)} \alpha_k^{(n)}$

## 5. NUMERICAL EXAMPLES

### 5.1 Simulation Setup

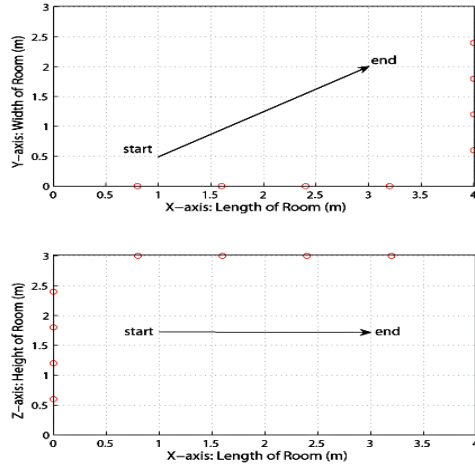


Figure 1: An Illustration of the Simulated Reverberant Room, Microphone Array ( $\circ$ ) Distributions and Acoustic Source Trajectory (arrow).

The enclosed reverberant environment in this simulation is defined as a typical rectangular room with dimensions of  $4m \times 3m \times 3m$ . Both the  $x$ - $y$  and  $x$ - $z$  Cartesian planes of the room have been divided into grid segmentations of  $0.01m \times 0.01m$ . The reverberations inside the room are simulated by the Image Method [10] for a range of reverberation times  $RT_{60} \in [0, 0.462]$  seconds. White Gaussian noises are also added at each microphone to approximate an average input SNR of 36.4dB.

There are a total of  $M = 16$  microphone, grouped into 4 linear microphone arrays. The distribution of the microphone arrays is illustrated in Figure 1. The two microphone arrays mounted in the  $x$ - $y$  Cartesian plane have a constant height of 1.5m and serve to estimate the acoustic source's  $(x, y)$  Cartesian coordinates. While the rest two microphone arrays are allocated along the vertical central axis of a side wall and the horizontal central axis of the ceiling respectively. The latter two arrays are used to estimate the acoustic source's height. The distance between two adjacent microphones in

the same microphone array is either 0.8 m (along  $x$ -axis) or 0.6 m (along  $y$ -axis or  $z$ -axis).

The acoustic source used in the simulation is a male speaker with a fixed height of 1.7 m, and the source is assumed to move straightly from (1, 0.5) to (3, 2) in  $(x, y)$  plane with constant speed. The acoustic signal used here is a pre-recorded clean male speech sample that lasts for about 7.86 seconds. The signal's transmission speed is set to be 343 m/sec. The microphone array received signals are sampled at a frequency of 8 kHz and decomposed into non-overlapping frames of  $L = 512$  samples each.

For all the PF approaches, the number of particles is  $N = 50$ , SRP-PHAT is applied as the ASL algorithm (*observation function*) to measure observation  $Y_k$ , and SRP-PHAT's output is used as the *pseudo likelihood*  $p(\alpha_k|Y_k)$ . While for BS-PF, the Langevin process is used to define the source state *transition function*, the particle resampling size threshold is  $N_{th} = 38$ . For the proposed hybrid BS-IS PF,  $\delta_H$  is set to be 0.05m.

### 5.2 Simulation Results and Discussions

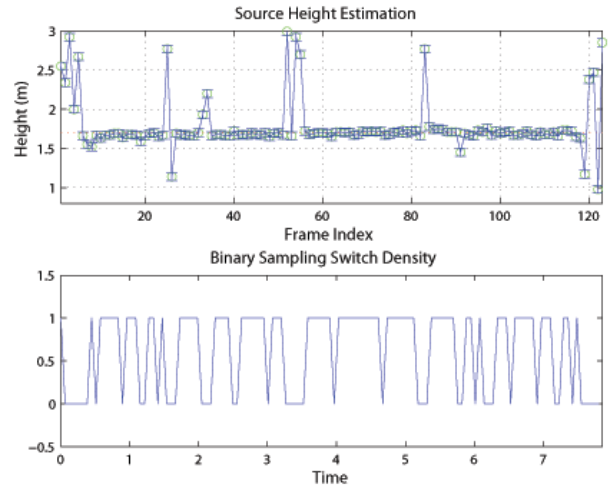


Figure 2: Source Height Estimation and corresponding Sampling Switch Density for  $RT_{60} \approx 0.24$  sec and  $SNR \approx 36.4$  dB.

Figure 2 shows the simulation resulted SHE and the corresponding *sampling switch density*. The SHE is marked with circle, the error bar indicates the predefined SHE error threshold (0.05m in this case), and the horizontal dotted line denotes the constant source height level (1.7m in this case). The finally estimated  $\bar{H} = 1.6852m$  is quite close to the actual source height 1.7m. By observing the SHE fluctuation, it is quite convincing to see that the reverberation resulted inaccurate SHE differs significantly from  $\bar{H}$  and from each other. This would help to verify the applicability of our proposed SHE based *sampling switch decision* in (24). It is also quite convincing to see that the binary *sampling switch density* could closely respond to the fluctuating SHE. Once there is large SHE error, the corresponding observation  $Y_k$  is also inaccurate, i.e.  $p_{ss}(Y_k) = 0$ . While for the other frames, the SHE fall within allowed error range, the observations  $Y_k$  at those corresponding frames are classified as accurate, i.e.  $p_{ss}(Y_k) = 1$ . The change of  $p_{ss}(Y_k)$  values and the SHE fluctuation has been shown to match with each other in fast and accurate manner.

Figure 3 gives a better insight with the comparison of the tracking performances achieved by the three PF based algorithms. The solid line presents the estimated source trajectory, the broken line denotes the estimation standard deviation, and the dotted line is the source actual trajectory. With randomized initial particles distribution, the BS-PF based algorithm encountered great difficulties in locating the acoustic source. From the middle plot, it is easy to see that the BS-PF based algorithm spent almost the first half of the

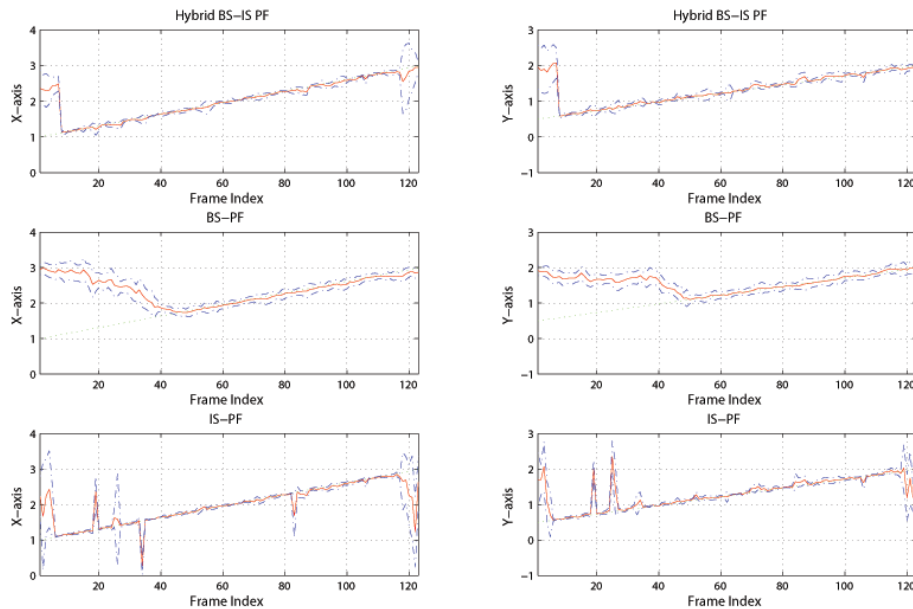


Figure 3: AST Results of Hybrid BS-IS PF, BS-PF and IS-PF along x-axis and y-axis respectively for  $RT_{60} \approx 0.24$  sec and  $SNR \approx 36.4$  dB.

AST process to locate the source. However, once the BS-PF based algorithm had correctly located the source, it performed very well in the followed tracking process. The BS-PF's competitive reverberation robustness has been well shown with its tracking performance.

In contrast, from the bottom plot, the IS-PF based algorithm showed its rapid source locating capability by finding the source location in very few frames. However, the tracking performance of the IS-PF based algorithm clearly illustrated its relative poor reverberation robustness. Even if the IS-PF based algorithm had followed the source trajectory for quite a while, it could still encounter temporary tracking difficulties by showing large estimation errors. This could be contributed to the inaccurate SRP-PHAT measurements at those corresponding frames.

Last but not the least, the hybrid BS-IS PF based algorithm's tracking performance from the top plot clearly indicates that the hybrid algorithm has successfully overcome the shortcomings of both the BS-PF and the IS-PF. At the initial stage, all the particles were randomly distributed, and the SHE at the very beginning frames were not very accurate. This suggested that at initial stage the SRP-PHAT based observations did not indicate any valid information for particles' sampling. Consequently, the hybrid BS-IS PF based algorithm should switch to BS-PF mode that follows the source state transition function to minimize the reverberation effects. After a few frames, the SHE had approached steady state, suggesting accurate observations at corresponding frames. The hybrid BS-IS PF based algorithm rapidly located the source by switching to the IS-PF mode. For the followed AST process, the hybrid BS-IS PF algorithm could firmly track the source trajectory even if there were some inaccurate observations caused by reverberations at certain frames. This shows that the hybrid algorithm is free of reverberation influences by switching from IS-PF to BS-PF whenever necessary.

## 6. CONCLUSIONS

The objective of this work is to track a moving acoustic source inside an enclosed reverberant and noisy environment with distributed acoustic sensor arrays. In this paper, we proposed a novel hybrid BS-IS PF algorithm with source height estimation based sampling switch technique. The height estimation can be readily obtained with two additional microphone arrays, and can be used to indicate whether the corresponding observation is accurate enough to adopt proper PF algorithm. Numerical results show that the hybrid BS-IS

PF based algorithm has improved AST performance over the BS-PF and the IS-PF based algorithms in reverberant environments.

## REFERENCES

- [1] J. Vermaak and A. Blake, "Nonlinear filtering for speaker tracking in noisy and reverberant environments," in *Proc. ICASSP 2001*, vol. 5, pp. 3021-3024.
- [2] D.B. Ward, E.A. Lehmann and R.C. Williamson, "Particle filtering algorithms for tracking an acoustic source in a reverberant environment," *IEEE Transactions on Speech and Audio Processing*, vol. 11, pp. 826-836, 2003.
- [3] E.A. Lehmann and R.C. Williamson, "Particle filter design using importance sampling for acoustic source localisation and tracking in reverberant environments," *EURASIP Journal on Applied Signal Processing, special issue on Advances in Multi-Microphone Speech Processing*, vol. 2006, pp. 168-176, 2006.
- [4] E.A. Lehmann and A.M. Johansson, "Particle filter with integrated voice activity detection for acoustic source tracking," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, pp. 28-38, 2007.
- [5] N.J. Gordon, D.J. Salmond, and A.F.M. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," in *Proc. Radar and Signal Processing*, vol. 140, pp. 107-133, 1993.
- [6] E.A. Lehmann, "Particle Filtering Methods for Acoustic Source Localisation and Tracking," *PhD Thesis, Australian National University, Australia*, 2004.
- [7] J.H. DiBiase, H.F. Silverman and M.S. Brandstein, "Robust localization in reverberant rooms," *Microphone Arrays Signal Processing Techniques and Applications*, Springer, Germany, 2001.
- [8] B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, N.J., 1979.
- [9] A. Doucet, S. Godsill, and C. Andrieu, "On sequential monte carlo sampling methods for bayesian filtering," *Statistics and Computing*, vol. 10, pp. 197-208, 2000.
- [10] J.B. Allen and D.A. Berkley, "Image method for efficiently simulating small-room acoustics," *Journal of the Acoustical Society of America*, vol. 107, pp. 384-391, 2000.