A RANDOM FINITE SET APPROACH FOR TRACKING TIME-VARYING NUMBER OF ACOUSTIC SOURCES USING A SINGLE ACOUSTIC VECTOR SENSOR

Xionghu Zhong, A. B. Premkumar and A. S. Madhukumar

Centre for Multimedia and Network Technology, School of Computer Engineering,
College of Engineering, Nanyang Technological University, Singapore, 639798.
{xhzhong, asannamalai and asmadhukumar}@ntu.edu.sg

ABSTRACT

Existing localization approaches developed using acoustic vector sensor (AVS) signals normally assume that the sources are static and the number of sources is known. In this paper, a novel approach is developed to estimate the 2-D direction of arrival (DOA) of an unknown and time-varying number of acoustic sources using a single AVS. A random finite set (RFS) is employed to characterize the randomness of the state process, i.e., the dynamics of source motion and the number of active sources. Also the measurement processes are modeled by RFS since we allow the AVS report undesired events by sending an empty set other than received signals under regular cases. We further employ particle filtering to approximate the posterior DOA distributions. The performance of the proposed approach is demonstrated by simulated experiments.

Index Terms— Acoustic vector sensor, random finite set, direction of arrival, particle filtering, multiple acoustic sources.

1. INTRODUCTION

Localization and tracking the 2-D (azimuth and elevation) direction of arrivals (DOA) of acoustic sources in a noisy environment is significantly important in many applications such as room speech enhancement, underwater target surveillance, sonar and acoustic radar signal processing. The tasks are traditionally achieved by using arrays equipped with several pressure sensors, together with estimation techniques based on the collected acoustic pressure measurements [1]. However, such techniques usually require either an array of sensors with a large aperture or multiple hybrid arrays. In recent years, acoustic vector sensor (AVS) [2] that is able to measure acoustic pressure as well as particle velocity at sensor position has been widely employed for acoustic source localization.

AVS was first introduced in signal processing and acoustic source localization problems in [2], in which an intensity based algorithm and a velocity covariance based algorithm are fully presented. A maximum likelihood based DOA estimation algorithm is developed in [3]. The conventional beamforming (Bartlett beamforming) and Capon beamforming for 2-D DOA estimation using acoustic vector sensors are investigated in [4]. It shows that both the azimuth and elevation can be unambiguously estimated by using an AVS array. Further, the subspace based approaches such as MUSIC [5] and ESPRIT [6] have been used in AVS localization problem. More practically, AVS localization in shallow ocean and room reverberant environments are investigated in [7] and [8]respectively.

The existing 2-D DOA estimation schemes assume that the sources are static and the number of sources are known, and extensively rely on localization approaches. These assumptions are often violated in real applications since the sources (e.g., submarines in underwater or speakers in the room environment) are in fact dynamic, and the number of sources is unknown and may be time-varying. In this paper, we consider such a scenario where the source motion and the number of sources are unknown and time-varying. A random finite set (RFS) framework is employed to characterize the randomness from the source dynamics as well as the measurement processes. Basically, RFS framework neglects the intrinsic data association between sources and measurements, and has been found promising for multi-object tracking problem [9–11]. In the state space, each element of an RFS is a random vector which can be employed to describe the state of source, and the cardinality of the set is also random and can be used to model the number of sources. The source motion is modeled by employing a constant velocity (CV) model, and the source appearance and disappearance are described by using birth and death processes respectively. For the measurement, the AVS signal is received under regular cases and an empty set is received if undesired event happens. Since the AVS measurement function is nonlinear, a particle filtering (PF) [12] implementation is employed to obtain the final DOA estimates. For rigorous mathematical discipline of RFS framework and its application in multi-object tracking problem, the reader is referred to [9, 11]. Particularly, RFS is employed for multiple room acoustic source tracking in [10]. Very recently, PF has been employed for 2-D DOA estimation of a single source by using a single AVS [13].

The rest of this paper is organized as follows. In Section 2, the AVS signal model is introduced. Section 3 and 4 present the RFS formulation for source dynamics and the measurement process respectively. Simulated experiments are organized in Section 5. Conclusions and future work are described in Section 6.

2. SIGNAL MODEL

Assume that m_t acoustic source signals impinge on an AVS at discrete time t. The source signals $\mathbf{s}(t)$ can be written as

$$\mathbf{s}(t) = [s_1(t), \dots, s_{m_t}(t)]^T \in \mathbb{C}^{m_t \times 1}, \quad (1)$$

where $\mathbb{C}$ denotes the complex domain and superscript T is the matrix transpose operation. Further assume that the mth source signal is emitted at a 2-D direction $\boldsymbol{\theta}_t^m$ given by

$$\boldsymbol{\theta}_t^m = [\phi_t^m,\ \psi_t^m]^T, \quad m = 1, \dots, m_t, \quad (2)$$

with $\phi_t^m \in (-\pi,\ \pi]$ and $\psi_t^m \in [-\pi/2,\ \pi/2]$ denoting the azimuth and the elevation angle respectively. AVS measures the acoustic pressure as well as three component particle velocities. Let $\boldsymbol{u}_t^m$ be the unit direction vector pointing from the origin toward the source position, and normalized by a constant term $\rho_0 c$, given as

$$\boldsymbol{u}_t^m = -\frac{1}{\rho_0 c_0} \left[\cos\psi_t^m \cos\phi_t^m, \cos\psi_t^m \sin\phi_t^m, \sin\psi_t^m\right]^T, \quad (3)$$

where ρ_0 and c_0 represent the ambient density and the propagation speed of the acoustic wave in the medium respectively. The received signal model for an AVS located at $\boldsymbol{r}$ can be written as

$$\mathbf{y}(t) = \sum_{m=1}^{m_t} \left[\begin{array}{c} 1 \\ \boldsymbol{u}_t^m \end{array} \right] s_m(t - \tau_t^m) + \left[\begin{array}{c} \mathrm{n}_p(t) \\ \mathbf{n}_v(t) \end{array} \right], \qquad (4)$$

where $\mathrm{n}_p(t) \in \mathbb{C}$ and $\mathbf{n}_v(t) \in \mathbb{C}^{3\times 1}$ represent the corresponding pressure and velocity noise terms separately. τ_t^m is the time delay of the mth wave between the sensor and the origin of the coordinate system, i.e., $\tau_t^m = -\boldsymbol{r}^T \boldsymbol{u}_t^m / c$.

For an acoustic source that moves relatively slowly, the DOA $\boldsymbol{\theta}_t^m$ can be assumed to be stable if a small number of snapshots are processed at each time step. Assume that N snapshots are taken into account at time step k, the number of sources is thus m_k and the snapshots of the source signal can be written as

$$\mathbf{S}_k = \left[\mathbf{s}(kN+1), \ldots, \mathbf{s}(kN+N)\right], \qquad (5)$$

where $\mathbf{S}(k) \in \mathbb{C}^{m_k \times N}$. The noise and received data matrices can be expressed as $\mathbf{N}_k = [\mathbf{n}(kN+1), \ldots, \mathbf{n}(kN+N)] \in \mathbb{C}^{4\times N}$ and $\mathbf{Y}_k = [\mathbf{y}(kN+1), \ldots, \mathbf{y}(kN+N)] \in \mathbb{C}^{4\times N}$ respectively. Accordingly, $\boldsymbol{\theta}_k$ is used to express the DOA at time step k. Equation (4) can thus be written as

$$\mathbf{Y}_k = \mathbf{A}(\boldsymbol{\theta}_k)\mathbf{S}_k + \mathbf{N}_k, \qquad (6)$$

where $\mathbf{A}(\boldsymbol{\theta}_k) = [\mathbf{a}(\boldsymbol{\theta}_k^1), \ldots, \mathbf{a}(\boldsymbol{\theta}_k^{m_k})]$, with $\mathbf{a}(\boldsymbol{\theta}_k^m) = \left[1, (\boldsymbol{u}_k^m)^T\right]^T$ denoting the steering vector. Both the azimuth and elevation information are thus included and 2-D DOA can be estimated.

Assume that: 1) the noise terms in (6) are independent identically distributed (i.i.d.), zero-mean complex circular Gaussian processes and are independent from different channels; and 2) the source signal $\mathbf{S}_k$ and the noise $\mathbf{N}_k$ are independent. The PDF of the measurements can be addressed as $\mathbf{Y}_k \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma}_k)$ where $\mathcal{CN}(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ represents a multivariate complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The noise process is characterized by the covariance matrix given by

$$\boldsymbol{\Gamma}_k = \mathbb{E}\{\mathbf{N}_k\mathbf{N}_k^H\} = \left[\begin{array}{cc} \sigma_p^2 & \mathbf{0} \\ \mathbf{0} & \sigma_v^2\mathbf{I}_3 \end{array} \right], \qquad (7)$$

where $\mathbf{I}_q$ is an qth order identity matrix, and σ_p^2 and σ_v^2 are the noise variances for the pressure and velocity components respectively.

3. RFS STATE DYNAMICS FORMULATION

Assume that the sources move with a velocity $\dot{\boldsymbol{\theta}}_k^m$ (in rad/s), for $m = 1, \ldots, m_k$. The source state $\mathbf{x}_{m,k}$ can be constructed by cascading the DOA $\boldsymbol{\theta}_k^m$ and the velocity $\dot{\boldsymbol{\theta}}_k^m$, i.e., $\mathbf{x}_{m,k} = [\boldsymbol{\theta}_k^m, \dot{\boldsymbol{\theta}}_k^m]^T$. The CV model [13] is employed to model the source dynamics given as

$$\mathbf{x}_{m,k} = \mathbf{F}\mathbf{x}_{m,k-1} + \mathbf{G}\mathbf{v}_k, \qquad (8)$$

where the coefficient matrix $\mathbf{F}$ and $\mathbf{G}$ are defined respectively by

$$\mathbf{F} = \left[\begin{array}{cc} 1 & \Delta T \\ 0 & 1 \end{array} \right] \otimes \mathbf{I}_2; \quad \mathbf{G} = \left[\begin{array}{c} \Delta\frac{T^2}{2} \\ \Delta T \end{array} \right] \otimes \mathbf{I}_2, \qquad (9)$$

where ΔT represents the time period in seconds between the previous and current time step, and $\otimes$ denotes the Kronecker product, and $\mathbf{v}(k)$ is a zero-mean real Gaussian process. For tracking unknown number of multiple acoustic sources, the parameters of interest will be the 2-D DOA of each source and the number of sources. We characterize these unknown parameters by using an RFS, given as

$$\mathcal{X}_k = \{\mathbf{x}_{1,k}, \ldots, \mathbf{x}_{m_k,k}\}, \qquad (10)$$

where $m_k = |\mathcal{X}_k|$ is the number of sources, with $|\cdot|$ representing the cardinality. For the source dynamics, we have following assumptions: 1) each active source follows the CV motion model; 2) the source motions are independent of each other; and 3) the maximum number of sources in the tracking scene is bounded by $M_{\max}$. It is observed in [3] that for a single AVS, only up to two sources can be uniquely identified. Hence, $M_{\max} = 2$ is chosen in this work.

Given a realization $\mathcal{X}_{k-1}$ of the RFS state at previous time step $k-1$, the source state $\mathcal{X}_k$ at current step k is modeled by

$$\mathcal{X}_k = \mathcal{B}_k(\mathbf{b}_k) \cup \mathcal{S}_k(\mathbf{X}_{k-1}), \qquad (11)$$

where $\mathcal{B}_k(\mathbf{b}_k)$ is the state vector of sources born at time step k, and $\mathcal{S}_k(\mathcal{X}_{k-1})$ denotes the RFS of states that have survived at time step k. The source birth process can be formulated as

$$\mathcal{B}_k(\mathbf{b}_k) = \begin{cases} \emptyset, & \hbar_{\text{birth}}; \text{ or } |\mathcal{X}_{k-1}| = M_{\max}; \\ \{\mathbf{b}_k\}, & \bar{\hbar}_{\text{birth}}. \end{cases} \qquad (12)$$

where $\hbar_{\text{birth}}$ and $\bar{\hbar}_{\text{birth}}$ are the hypotheses for birth process and non-birth processes respectively, and $\mathbf{b}_k$ is the initial state vector under the birth hypothesis given as $\mathbf{b}_k = \mathbf{x}_0 \sim (\boldsymbol{\theta}_0, \dot{\boldsymbol{\theta}}_0)$. The survived state set $\mathcal{S}_k(\mathcal{X}_{k-1})$ can be formulated by considering a death process. When a death process happens, the corresponding state set will be empty, and the remaining states will evolve following the motion dynamics equation (8). $\mathcal{S}_k(\mathcal{X}_{k-1})$ can thus be given as

$$\mathcal{S}_k(\mathcal{X}_{k-1}) = \begin{cases} \mathcal{X}_{k-1} \setminus \{\mathbf{x}_{m,k}\}, & \hbar_{\text{death}} \text{ for } m\text{th source}; \\ \bigcup_{m=1}^{|\mathcal{X}_{k-1}|} \{\mathbf{F}\mathbf{x}_{m,k-1} + \mathbf{G}\mathbf{v}_k\}, & \bar{\hbar}_{\text{death}}. \end{cases} \qquad (13)$$

with $\setminus$ denoting the set minus, and $\hbar_{\text{death}}$ and $\bar{\hbar}_{\text{death}}$ are the hypotheses for death and non-death processes respectively. The birth and death processes happen with prior probability P_{birth} and P_{death} respectively.

The RFS state transition density can thus be expressed by taking the product of birth PDF and survival PDF, given as

$$p(\mathcal{X}_k|\mathcal{X}_{k-1}) = p(\mathcal{B}_k|\mathcal{X}_{k-1})p(\mathcal{S}_k|\mathcal{X}_{k-1}). \qquad (14)$$

The PDF of birth process can be formulated as

$$p(\mathcal{B}_k|\mathcal{X}_{k-1}) = \begin{cases} 1 - P_{\text{birth}}, & \mathcal{B}_k = \emptyset; \\ P_{\text{birth}} p(\mathbf{x}_0), & \mathcal{B}_k = \{\mathbf{x}_0\}; \\ 0, & \text{otherwise}. \end{cases} \qquad (15)$$

To formulate the PDF of death process $p(\mathcal{S}_k|\mathcal{X}_{k-1})$, we firstly consider a single source death process. Source dies with following PDF

$$p(\mathcal{S}_k(\mathbf{x}_{m,k-1})|\mathcal{X}_{k-1}) = \begin{cases} P_{\text{death}}, & \mathcal{S}_k(\mathbf{x}_{m,k-1}) = \emptyset; \\ (1 - P_{\text{death}})p(\mathbf{x}_{m,k}|\mathbf{x}_{i,k-1}), & \mathcal{S}_k(\mathbf{x}_{m,k-1}) = \{\mathbf{x}_{m,k}\}; \\ 0, & \text{otherwise}. \end{cases} \qquad (16)$$

The total PDF of death process can be written as [10]

$$p(\mathcal{S}_k|\mathcal{X}_{k-1}) = P_{\text{death}}^{m_k - m_{k-1}}(1 - P_{\text{death}})^{m_{k-1}} \sum_{1 \le i_1 \ne i_m \le m_{k-1}} \prod_{j=1}^{m_k} p(\mathbf{x}_{m,k}|\mathbf{x}_{i_j,k-1}), \qquad (17)$$

where

$$\sum_{1\le i_1\neq i_m\le m_{k-1}} = \sum_{i_1=1}^{n}\sum_{i_2=1,i_2\neq i_1}^{n}\cdots = \sum_{i_1=1}^{n}\sum_{i_m=1,i_m\neq i_{m-1}\neq\cdots\neq i_1}^{n}. \tag{18}$$

Hence the RFS state transition PDF is formulated. In practice, P_{birth} and P_{death} are unknown and are obtained based on experimental studies. Usually, increasing P_{birth} or P_{death} will enable the algorithm easier to detect the birth and death respectively of a source. However, an very large P_{birth} or P_{death} will lead to an overestimation or underestimation of the number of sources.

4. MEASUREMENT AND LIKELIHOOD MODEL

In practice, the measurements can either be regular signals emitted by sources or be strange signals (which can be easily detected) due to transient interferences. The former is modeled by equation (6) and the later is modeled by using an empty set. This means that we allow the sensor to ignore the strange signals and simply report an undesired event by sending an empty set to enhance the robustness of the algorithm. Assume that the RFS measurement is $\mathcal{Z}_k$. It may be constructed by one of following cases: $\mathcal{Z}_k = \{\mathbf{Y}_k\}$ if the measurement is generated by a source/sources; $\mathcal{Z}_k = \{\mathbf{N}_k\}$ if it is a pure noise; and $\mathcal{Z}_k = \emptyset$ if it fails to give any report. Given the probability of false alarms P_{f} and the probability of detection P_{d}, the prior probabilities of the measurement presentation can be written as

$$P(\mathcal{Z}_k|\mathcal{X}_k) = \begin{cases} (1-P_{\text{f}})(1-(1-P_{\text{d}})^{|\mathcal{X}_k|}), & \mathcal{Z}_k = \{\mathbf{Y}_k\}; \\ P_{\text{f}}, & \mathcal{Z}_k = \{\mathbf{N}_k\}; \\ (1-P_{\text{f}})(1-P_{\text{d}})^{|\mathcal{X}_k|}, & \mathcal{Z}_k = \emptyset. \end{cases} \tag{19}$$

It is obvious that when $\mathcal{Z}_k = \{\mathbf{N}_k\}$, the measurement process is a false alarm. The prior is thus P_{f}. When $\mathcal{Z}_k = \emptyset$, the measurement process is not a false alarm and also we missed all the sources. Hence the prior for such a case is $(1-P_{\text{f}})(1-P_{\text{d}})^{|\mathcal{X}_k|}$. Since the total prior probability must sum to unity, the prior probability for $\mathcal{Z}_k = \{\mathbf{Y}_k\}$ is naturally the rest of it. When the measurement set is generated by a source or multiple sources, the density function is [2]

$$\begin{aligned} p(\mathcal{Z}_k = \{\mathbf{Y}_k\}\,|\mathcal{X}_k) &= p(\mathbf{Y}_k|\mathbf{x}_{1,k},\ldots,\mathbf{x}_{|\mathcal{X}_k|,k}) \\ &= (e\pi)^{-4N}\det(\mathbf{\Pi}_k\hat{\mathbf{R}}_k\mathbf{\Pi}_k + \hat{\sigma}^2\mathbf{\Pi}_k^0)^{-N}, \end{aligned} \tag{20}$$

where

$$\begin{aligned} &\mathbf{\Pi}_k = \mathbf{A}(\boldsymbol{\theta}_k)(\mathbf{A}^H(\boldsymbol{\theta}_k)\mathbf{A}(\boldsymbol{\theta}_k))^{-1}\mathbf{A}^H(\boldsymbol{\theta}_k);\;\; \mathbf{\Pi}_k^0 = \mathbf{I} - \mathbf{\Pi}_k; \\ &\hat{\sigma}^2 = \frac{1}{4-m_k}\text{tr}(\mathbf{\Pi}_k^0\hat{\mathbf{R}}_k);\quad \hat{\mathbf{R}}_k = \frac{1}{N}\mathbf{Y}_k\mathbf{Y}_k^H. \end{aligned} \tag{21}$$

When the measurement is due to noise, it is a circular complex white Gaussian process. The PDF is then

$$p(\mathcal{Z}_k = \{\mathbf{N}_k\}\,|\mathcal{X}_k) = (e\pi)^{-4N}\det(\hat{\mathbf{R}}_k)^{-N}. \tag{22}$$

When the measurement is an empty set, the likelihood is not available and only the prior will be considered. Like P_{birth} or P_{death}, P_{f} and P_{d} are decided based on some rough guess. Generally, reducing P_{f} and increasing P_{d} are expected to enhance the robustness of the algorithm and the capability of discovering new sources. However, overly large values will risk the accuracy of state estimation.

The above description gives an RFS presentation for AVS signal based detection and tracking problem. For PF implementation, we use a number of particles to approximate the posterior PDF of the interesting parameters. Assume that we have particles $\mathcal{X}_{k-1}^{(i)}$ for $i = 1,\ldots,L$ at previous time step $k-1$ and the corresponding importance weight $w_{k-1}^{(i)}$. Note that different from the PF in [13] where each particle is a random vector, the particle drawn here is a random set. The particles at time step k are generated according to the state dynamic process described in Section 3, given as

$$\mathcal{X}_k^{(i)} \sim p(\mathcal{X}_k^{(i)}|\mathcal{X}_{k-1}^{(i)}). \tag{23}$$

Since a prior importance [12] is used, the particles are weighted by

$$w_k^{(i)} = w_{k-1}^{(i)}p(\mathcal{Z}_k|\mathcal{X}_k^{(i)}). \tag{24}$$

After resampling, the posterior distribution is thus approximated by

$$p(\mathcal{X}_k^{(i)}|\mathcal{Z}_k) \approx \sum_{i=1}^{L}\tilde{w}_k^{(i)}\delta_{\mathcal{X}_k^{(i)}}(\mathcal{X}_k), \tag{25}$$

where $\tilde{w}_k^{(i)}$ is the normalized weight, and $\delta_{\mathcal{X}}(\mathcal{Y})$ is a set-valued Dirac delta function defined as 1 if $\mathcal{X} \subseteq \mathcal{Y}$ and 0 otherwise.

Due to an RFS presentation, each particle may differ from the others in dimension and the elements in it have an arbitrary order. Extracting the state estimation is thus not as straightforward as that in the single source scenario. Based on the particles and the corresponding weights $\{\mathcal{X}_k^{(i)}, w_k^{(i)}\}_{i=1}^{L}$, the number of sources can be approximated by

$$M_k \approx \sum_{i=1}^{L} w_k^{(i)}|\mathcal{X}_k^{(i)}|. \tag{26}$$

Since the number of sources should be an integer, we obtain the estimation of source number by using a rounding operation, i.e., $\hat{M}_k = \lceil M_k \rfloor$. Then a K-means algorithm is employed to cluster all the RFS particles. The centers of these clusters $\{\hat{\mathbf{x}}_{m,k}\}_{m=1}^{\hat{M}_k}$ are taken as the final state estimates.

5. SIMULATED EXPERIMENT

A single AVS located at the origin is used as the receiver, i.e., $\boldsymbol{r} = [0,0,0]^T$. Two source signals are generated by using i.i.d. complex circular Gaussian processes. The sampling frequency is 1kHz. We consider following source dynamics scenario: one source starts from time step 1 and goes on until time step 40, with the corresponding DOA $\boldsymbol{\theta}_{1,1} = (-3\pi/4, -\pi/12)$ and $\boldsymbol{\theta}_{1,40} = (-\pi/4, 5\pi/12)$, and the other is active from time step 21 to 60, with the corresponding DOAs $\boldsymbol{\theta}_{2,21} = (-2\pi/3, 3\pi/12)$ and $\boldsymbol{\theta}_{2,60} = (\pi/3, -5\pi/12)$ respectively. The source trajectories are depicted in Fig. 1. To simulate undesired events, we set the measurements at time step 10 and 35 as empty sets. The background noise level is evaluated by signal-to-noise ratio (SNR), and is simulated by adding a complex white Gaussian noise (WGN) into the received signal. In this paper, we assume that the noise variance for the pressure and velocity components are the same, i.e., $\sigma_p^2 = \sigma_v^2 = \sigma^2$. The SNR is thus the same across different channels. Some general parameters for PF are set as: initial velocity $\mathbf{v}_0 = [0.02\ 0.02]^T$, noise variance in CV model $\sigma_\phi^2 = \sigma_\psi^2 = 4\times10^{-4}$, $L = 1000$, $P_{\text{birth}} = P_{\text{death}} = 0.15$, $P_{\text{f}} = 0.2$, and $P_{\text{d}} = 0.9$. It is worth mentioning that slightly changing these parameters will not lead to significantly different tracking results.

Figure 1 gives the tracking results based on a single trial. To get a rough idea of the tracking performance, we also implement Capon beamforming method [4] here. At most two peaks are collected in the Capon response to obtain the DOA estimates. The results

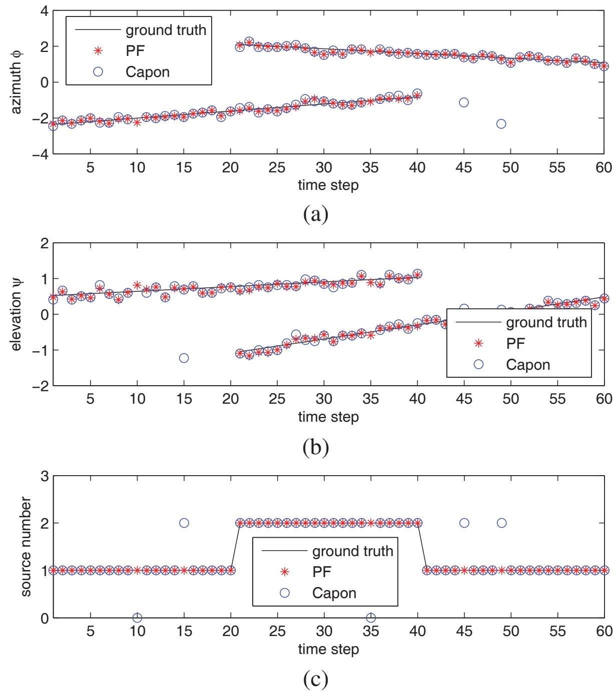

Fig. 1. Single trial under SNR $= -8$dB and $N = 128$; (a) azimuth; (b) elevation; and (c) source number.

show that the proposed approach is able to discover and track the sources accurately. However, at several time steps, spurious peaks are presented in the Capon response and result in incorrect number of sources estimation. Also, the proposed tracking approach is able to keep locking on the source even when the AVS fails to provide appropriate measurements (see results at time step 10 and 35 where measurements are empty sets).

To evaluate the average tracking performance, the optimal subpattern assignment (OSPA) metric [14] is employed here. Basically, OSPA metric transfers the cardinality estimation error into DOA error by employing a penalty value c. In this work, a moderate penalty value $c = \pi/4$ and an exponential factor $p = 2$ is used (the reader is referred to [14] for a detailed definition of c and p). Fig. 2 shows the OSPA error over 100 Monte Carlo trials under SNR $= -8$dB and $N = 128$, and SNR $= -6$dB and $N = 256$. Not only the proposed approach is able to present better number of sources estimates, but also the DOA estimation is more accurate than that of Capon approach. The proposed approach is able to discover the sources and estimate the DOA of the sources accurately only except at time step 21 when new source appears in the tracking scene. However, it is able to discover the new source and report its DOA quickly.

6. CONCLUSIONS AND FUTURE WORK

An RFS approach is developed for 2-D DOA tracking of a time-varying number of acoustic sources using a single AVS in this paper. RFS is employed to characterize the randomness of the state process, i.e., the dynamics of source motion and the number of active sources. Also the measurement processes are modeled by RFS since we allow the AVS report undesired events by sending an empty set other than received signals under regular cases. The performance of the proposed algorithm in estimating the number of sources as well

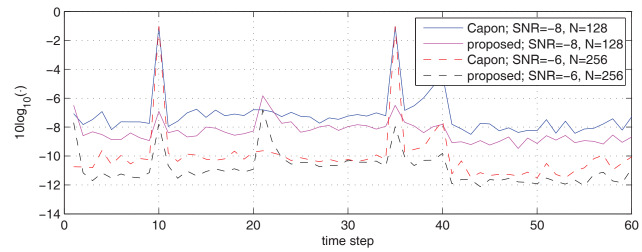

Fig. 2. OSPA error over 100 Monte Carlo trials.

as the source DOAs is much better than that of Capon beamforming method. However, only a single acoustic source is considered in this paper. Hence, future work includes developing a PF algorithm to track multiple acoustic source using an AVS array in different noisy environments. Also, comparing the DOA tracking performance with that based on other tracking approaches is an interesting direction.

7. REFERENCES

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, 1996.

[2] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2481–2491, 1994.

[3] B. Hochwald and A. Nehorai, "Identifiability in array processing models with vector-sensor applications," *IEEE Trans. Signal Process.*, vol. 44, no. 1, pp. 83–95, 1996.

[4] M. Hawkes and A. Nehorai, "Acoustic vector-sensor beamforming and capon direction estimation," *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2291–2304, 1998.

[5] K. T. Wong and M. D. Zoltowski, "Self-initiating MUSIC-based direction finding in underwater acoustic particle velocity-field beamspace," *IEEE J. Ocean. Eng.*, vol. 25, no. 2, pp. 262–273, 2000.

[6] P. Tichavsky, K. T. Wong, and M. D. Zoltowski, "Near-field/far-field azimuth and elevation angle estimation using a single vector hydrophone," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2498–2510, 2001.

[7] K. G. Nagananda and G. V. Anand, "Subspace intersection method of high-resolution bearing estimation in shallow ocean using acoustic vector sensors," *Signal Process.*, vol. 90, pp. 105–118, 2010.

[8] Dovid Levin, Emanul A. P. Habets, and Sharon Gannot, "On the angular error of intensity vector based direction of arrival estimation in reverberant sound fields," *J. Acoust. Soc. Amer.*, vol. 128, no. 4, pp. 1800–1811, Oct. 2010.

[9] B.-N. Vo, S. Singh, and A. Doucet, "Sequential monte carlo methods for multitarget filtering with random finite sets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 4, pp. 1224–1245, Oct. 2005.

[10] Wing-Kin Ma, Ba-Ngu Vo, Sumeetpal S. Singh, and Adrian Baddeley, "Tracking an unknown time-varying number of speakers using tdoa measurements: A random finite set approach," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3291–3304, 2006.

[11] Ronald P. S. Mahler, *Statistical Multisource-Multitarget information fusion.*, Artech House, 2007.

[12] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[13] X. Zhong, A. B. Premkumar, and A. S. Madhukumar, "Particle filtering and posterior Cramér-Rao bound for 2-D direction of arrival tracking using an acoustic vector sensor," *to appear on IEEE Sensors J.*

[14] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.