BAYESIAN MAXIMUM A POSTERIOR DOA ESTIMATOR BASED ON GIBBS SAMPLING

Jianguo Huang, Xiong Li, Qunfei Zhang

College of Marine Engineering, Northwestern Polytechnical University
Friendship Road 127,710072,Xi'an, China
phone: +86-29-88495789, fax: +86-29-88495278, email: jghuang@nwpu.edu.cn

ABSTRACT

DOA estimation is an important research area in array signal processing. Bayesian maximum a posterior DOA estimator (BM DOA estimator) has been shown to possess excellent performance. However, the BM estimator requires a multi-dimensional search and the computation burden increases exponentially with the dimension. So it is difficult to be used in real time applications. In order to reduce the computation of BM DOA Estimator, Monte Carlo methods are applied and a novel Bayesian Maximum a posterior DOA Estimator based on Gibbs Sampling (GSBM) is proposed. GSBM does not need multidimensional search, and not only keeps the good performance of original BM, but also reduces the original computation complexity from $O(L^K)$ to $O(K \times J \times N_s)$ where L , K , J and N_s are the number of grid, sources, samples and iteration respectively. Simulation results show that GSBM performs better than Maximum Likelihood Estimator (MLE), MUSIC, and MiniNorm, especially in low SNRs.

1. INTRODUCTION

High-resolution DOA estimation is an important research area in array signal processing. It arises in many fields including sonar, radar, astronomy, radio communications and geophysics. DOA estimation has captured much attention in the past two decades, and many methods have been proposed for different applications. Eigen-decomposition based methods including MUSIC, MiniNorm and MLE are some well-known procedures, and their performances have been thoroughly studied. In recent years, Bayesian high-resolution techniques [1], [2] and [3], which apply Bayes theorem in frequency and DOA estimation, become attractive for their superior performance. However the Bayesian high-resolution DOA estimators require multidimensional grid search which are prohibitively expensive in the presence of large number of sources [2], [3]. In order to solve the computational problems in Bayesian estimation, Monte Carlo methods are introduced, which have been shown to be a very powerful numerical Bayesian method [4][5]. In this paper, an algorithm combining the Bayesian method and the Gibbs sampler for DOA estimation is proposed. The proposed method not only reduces the computational burden obviously, but also keeps the performance as good as the original Bayesian method.

Supported by National Natural Science Foundation of China and Specialized Research Fund for the Doctoral Program of Higher Education of China(60272067).

2. BAYESIAN MAXIMUM A POSTERIOR DOA ESTIMATOR (BM DOA ESTIMATOR)

Consider a linear equi-spaced array of M sensors. Multiple far-field sources emit narrow-band signals with the direction parameters θ_k and frequencies $f_k (k = 1,\ 2,\ \cdots,\ K)$, which impinge on the sensors. These signals can be coherent or incoherent. The additive noise is assumed to be Gaussian and white with zero mean and variance σ^2 . Let c denote the speed of the signal propagation in the medium, and $\tau_k = b \sin\theta_k / c$ where b is the inter-element spacing. Then the data collected from the m-th sensor at time t_n are

$$x_m(t_n) = \sum_{k=1}^{K} I_k(t_n) \exp\left[j\phi_k(t_n) \right] \exp\left[j2\pi f_k \left(t_n - (m-1)\tau_k \right) \right] + n_m(t_n) = \sum_{k=1}^{K} A_k(t_n) f_{mk}(t_n) + n_m(t_n) \qquad (1)$$

where $n = 1, 2, ..., N$ with N being the number of snapshots, $A_k(t_n) = I_k(t_n) \exp\left[j\phi_k(t_n) \right]$, $I_k(t_n)$ and $\phi_k(t_n)$ is the unknown amplitude and phase of the k-th signal at time t_n , $f_{mk}(t_n) = \exp\left[j2\pi f_k \left(t_n - (m-1)\tau_k \right) \right]$, and $n_m(t_n)$ is the noise at time t_n on the m-th sensor. Our main interest here is to estimate $\vec{\theta} = [\theta_1, \ldots, \theta_K]^{\mathrm{T}}$. The unknown complex amplitudes $\vec{A} = \left\{ A_k(t_n),\ \forall\ k,\ n \right\}$ and the noise variance σ^2 are considered as the nuisance parameters. From a Bayesian perspective, the main entity for estimation is the posterior distribution of θ which can be expressed as

$$p(\vec{\theta} \mid X) = \int p\left(X \mid \vec{\theta}, \vec{A}, \sigma^2 \right) p\left(\vec{\theta}, \vec{A}, \sigma \right) d\vec{A}\ d\sigma^2 \qquad (2)$$

To solve the integration analytically, an orthogonalization on the data snapshots is performed [2], [3]. In particular, first the snapshots are divided into N_b blocks with each block having n_b snapshots. Then the orthogonalization of the data in the s-th block is accomplished by

$$\sum_{k=1}^{K} A_k(t_n) f_{mk}(t_n) = \sum_{k=1}^{K} B_k H_{mk}(t_n) \quad (3)$$

where

$H_{mk}(t_n) = \frac{1}{\sqrt{\lambda_k}} \sum_{l=1}^{K} e_{lk}^* f_{ml}(t_n)$, $B_k = \sqrt{\lambda_k} \sum_{l=1}^{K} A_l e_{lk}$, λ_k and $\vec{e}_k = [e_{1k}, e_{2k}, \cdots, e_{Kk}]^T$ are the eigenvalues and the eigenvectors of a $K \times K$ matrix F whose elements are defined as

$$F_{kl} = \sum_{n=1+(s-1)n_b}^{sn_b} \sum_{m=1}^{M} f_{mk}(t_n) f_{ml}^*(t_n) \quad (4)$$

Now, if the Jeffreys' priors are adopted, the desired posterior density function can be obtained as

$$p(\vec{\theta} \mid X) \propto \int \sigma^{2KN_b - 2MN - 1} \exp\left[-\frac{\left(\bar{d}^2 - \bar{h}^2\right)}{\sigma^2} \right] d\sigma$$

$$\propto \left[1 - \frac{\bar{h}^2}{\bar{d}^2} \right]^{(2KN_b - 2MN)/2} \quad (5)$$

where $\bar{d}^2 = \sum_{s=1}^{N_b} \bar{d}_s^2 = \sum_{s=1}^{N_b} \sum_{n=1+(s-1)n_b}^{sn_b} \sum_{m=1}^{M} | x_m(t_n) |^2$

$$= \sum_{n=1}^{N} \sum_{m=1}^{M} | x_m(t_n) |^2 \quad (6)$$

and $$\bar{h}^2 = \sum_{s=1}^{N_b} \sum_{k=1}^{K} \left| \sum_{n=1+(s-1)n_b}^{sn_b} \sum_{m=1}^{M} x_m(t_n) H_{mk}^*(t_n) \right|^2 \quad (7)$$

Notice that (5) is highly nonlinear and high dimensional with respect to $\vec{\theta}$. Thus calculations of the popular Bayesian estimators could be very intensive, especially when K is large. For instance, to obtain the maximum a posterior (MAP) estimator of $\vec{\theta}$, a K dimensional search is carried out to find the maximum peaks in the posterior distribution. The K angles corresponding to the peaks are the MAP estimate of the directions of the sources. Suppose that L grids are used for each dimension. The complexity of the K dimensional search is $O(L^K)$. Although the resolution ability of Bayesian method is rather high, the computational cost of the K dimensional computation and search could be prohibitively expensive for large K. To improve the real time computation of the Bayesian method, computational feasible solutions are demanded. Gibbs Sampling [5][6] is used to solve it, as described in the next section.

3. GIBBS SAMPLING

The Gibbs sampler is a Markov chain Monte Carlo (MCMC) sampling method for numerical evaluation of multidimensional integrations. The Gibbs sampling employs alternate conditional sampling, and may be considered as a special case of the Metropolis-Hastings algorithm [5]. Its popularity is gained from the facts that it is capable of carrying out many complex Bayesian computations. In the past decade, it has been intensively studied by statisticians and in recent years its applications in signal processing has been picked up.

The basic idea of the Gibbs sampler is to simulate a Markov chain in the state space of X so that the equilibrium of this chain is the target distribution $p(\vec{\theta} \mid X)$. So the Gibbs sampler algorithm is to first generate random samples from the joint posterior distribution $p(\vec{\theta} \mid X)$ by running Markov chains. Then the resulting samples are used by the Monte Carlo method to approximate the required high dimensional integrations. And the Gibbs sampler requires an initial transient period to converge to equilibrium. The initial period is known as the "burn-in" period, and the first n_0 samples in the period should always be discarded. Detection of convergence is usually done in some ad hoc way. For tutorials on the Gibbs sampler, see [5], [6].

4. BAYESIAN MAXIMUM A POSTERIOR DOA ESTIMATOR BASED ON GIBBS SAMPLING (GSBM)

It can be seen that the high dimensional integrations in (2) induce great computational difficulty and the K dimensional search for the DOA estimation. To solve the real time question, here we resort to the Gibbs sampler.

The key objective in a Gibbs sampling implementation is the generation of samples from the posterior distribution $p(\vec{\theta} \mid X)$. It is achieved through an iterative process. In a detail, given some initial values $\vec{\theta}^{(0)} = (\theta_1^{(0)}, ..., \theta_K^{(0)})$ of the K unknown directions, for $i = 1, 2, \cdots, N_s$, we proceed

1) Draw sample $\theta_1^{(i)}$ from $p\left(\theta_1 \mid \theta_2^{(i-1)}, \cdots, \theta_K^{(i-1)}, X\right)$,

2) Draw sample $\theta_2^{(i)}$ from $p\left(\theta_2 \mid \theta_1^{(i)}, \theta_3^{(i-1)} \cdots, \theta_K^{(i-1)}, X\right)$,

…

K) Draw sample $\theta_K^{(i)}$ from $p\left(\theta_K \mid \theta_1^{(i)}, \cdots, \theta_{K-1}^{(i)}, X\right)$.

Notice from (5) that $p(\theta_k \mid \theta_1^{(i)}, \cdots, \theta_{k-1}^{(i)}, \theta_{k+1}^{(i-1)}, \cdots, X)$ for $k = 1, 2, \cdots, K$ are not such distributions like the Gaussian or Gamma distributions. Therefore special care must be taken to achieve the required sampling objective. Next we proposed a procedure, which applies the sampling-resampling and kernel smoothing [7] techniques. In detail, the sampling of $\theta_k^{(i)}$ from $p(\theta_k \mid \theta_1^{(i)}, \cdots, \theta_{k-1}^{(i)}, \theta_{k+1}^{(i-1)}, \cdots, X)$ is carried out as follows:

1 Obtain J samples from the uniform distribution $U(-90, 90)$ and denoted them by $u(j), j = 1, \cdots, J$,

2 For each $u(j)$, form a new vector $\vec{\alpha}_j^T = [\theta_1^{(i)}, \theta_2^{(i)}, \cdots, \theta_{k-1}^{(i)}, u(j), \theta_{k+1}^{(i-1)}, \cdots, \theta_K^{(i-1)}]$, and then calculate from the distribution (5) the weights

$$\bar{w}_j \propto p(\vec{\alpha}_j \mid X)$$

Next obtain the normalized weights by

$$w_j = \bar{w}_j \Big/ \sum_{j=1}^{J} \bar{w}_j$$

3 Calculate the sample mean and variance according to

$$\bar{\mu} = \sum_{j=1}^{J} w_j u(j) \qquad \bar{\sigma}^2 = \sum_{j=1}^{J} w_j \left(u(j) - \bar{\mu}\right)^2$$

Then we approximate the conditional distribution $p(\theta_k \mid \theta_1^{(n)}, \cdots, \theta_{k-1}^{(n)}, \theta_{k+1}^{(n-1)}, \cdots, X)$ by a mixture density as

$$g\left(\theta_k\right) = \sum_{j=1}^{J} w_j TN\left(\theta_k \mid \bar{\mu}_j, h^2 \bar{\sigma}^2\right)$$

where $TN\left(\cdot \mid a, b\right)$ represents a truncated Gaussian with the mean a, the variance b, in particular, $h = \left(4/3\right)^{\frac{1}{5}} \Big/ J^{\frac{1}{5}}$, $\beta = \sqrt{1-h^2}$, and $\mu_j = \beta\, u\left(j\right) + \left(1-\beta\right)\bar{\mu}$.

4 Now our original objective of sampling from $p(\theta_k \mid \theta_1^{(n)}, \cdots, \theta_{k-1}^{(n)}, \theta_{k+1}^{(n-1)}, \cdots, X)$ becomes sampling from the mixture $g\left(\theta_k\right)$, $k = 1, \cdots, K$ which is implemented as

a) Sample an index j with probability w_j,

b) Sample $\theta_k^{(i)}$ from $TN\left(\theta_k \mid \bar{\mu}_j, h^2 \bar{\sigma}^2\right)$.

To ensure convergence, the above procedure is usually carried out for $\left(n_0 + N\right)$ iterations, and samples from the last N iterations are used to calculate the sources' directions. Finally, the directions of K sources can be obtained from the corresponding sample means as

$$\hat{\theta}_k = \frac{1}{N} \sum_{n=n_0+1}^{n_0+N} \theta_k^{(n)}, k = 1, 2, ..., K \tag{8}$$

It can be seen from above that the convergence of Gibbs sampling is very important because the method is iterative. Many parallel Markov chain are running simultaneously and when all these chains are stable, the iteration converges [5].

The complexity of the proposed Gibbs Sampling DOA estimator based on Bayesian method is $O(K \times J \times N_s)$. To compare with a K dimensional search, we observe that, first, due to the use of kernel smoothing techniques, J is smaller than L, the number of grid used in the K dimensional search. Secondly, as we will show next that the proposed algorithm converges very fast. Hence, N_s grows with K much slower than exponentially. As a result, for a large K, the computational demand of the GSBM is tremendously reduced with respect to that of the K dimensional search.

5. SIMULATION RESULTS

In this section, several experiments are conducted to show the performance of the GSBM.

Consider the case of two sources. The true DOAs of the sources are $\pm 2^\circ$. To determinate the convergence, in the first experiment we constructed two Markov chains and observe when two chains both are stable. The initial values of the two chains are different, $\vec{\theta}^{(1)} = (\theta_1^{(1)}, \theta_2^{(1)}) = (-5^\circ, -5^\circ)$ and $\vec{\theta}^{(2)} = (\theta_1^{(2)}, \theta_2^{(2)}) = (5^\circ, 5^\circ)$ respectively. In Figure 1,2,3,4, we plotted two trajectories of the samples of the two chains collected in the 50 iterations, in the cases of SNR=-5dB and SNR=15dB. It is clear that the samples of two chains converge very fast, fluctuate around the true values. And in high SNR the iteration converges faster than in low SNR. It can be seen from Figure 1,2,3,4, both Markov chains converge after about 5 iterations, so the Gibbs sampling converges after 5 iterations.

In the second experiment, the performance of the GSBM was compared with MLE, MUSIC, MiniNorm, and original BM method. These results are shown in Figure 5 and Figure 6. These results are based on large amount of computer simulations and the statistical analysis indicates that the GSBM keeps the excellent performance of original BM method, and it is much more robust under low SNRs. From Figure 5 we notice that under different SNRs that the resolution probability of GSBM is just same as original BM and MLE, always 100%, obviously superior to MUSIC and MiniNorm, especially in low SNRs . The resolution probability is the ratio of the times to be resolved to the whole times. Only when the SNR attains to 5dB, MUSIC and MiniNorm are able to distinguish the two sources completely. It can be seen from Figure 6 that estimation accuracy of GSBM is a little lower than original BM method, but obviously higher than MLE, MUSIC and MiniNorm, especially in low SNRs. An explanation why the Bayesian method performs better than the MLE is that the posterior distribution include the information of prior distribution and samples, while the MLE only utilizes the information of samples.

Furthermore, the calculation of GSBM is much less than the original BM method. The original BM method requires multidimensional grid search, but the new GSBM just needs to generate samples and iterate. As we have shown above, the computation complexity of the original Bayesian method is $O(L^K)$. The accuracy could be improved as L increases while the computation complexity becomes also higher. And when K becomes bigger, the complexity will be increased exponentially. But the complexity of GSBM, $O(K \times J \times N_s)$, is increased only linearly as K increases with also keeping the original good performance. For example, for three sources $K = 3$, L, the number of grid used in the K dimensional search, is 400, then the computation of original BM method is about $O(L^K) = O(400^3) = O(64 \times 10^6)$, while, let $J = 200$, $N_s = 50$, then the computation of GSBM is only about $O(K \times J \times N_s) = O(3 \times 200 \times 50) = O(3 \times 10^4)$. The reduction of computation is very obvious. The comparison of computation complexity between the original BM method and GSBM as K increases is shown in Figure 7. It is clear that the GSBM is a very efficient DOA high-resolution estimator for multiple source localization.

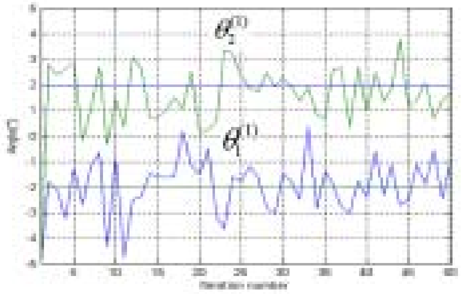

Figure 1 First Markov chain (SNR=-5dB)

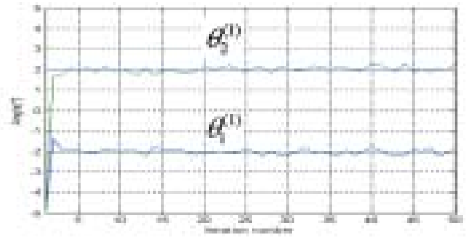

Figure 2 First Markov chain (SNR=15dB)

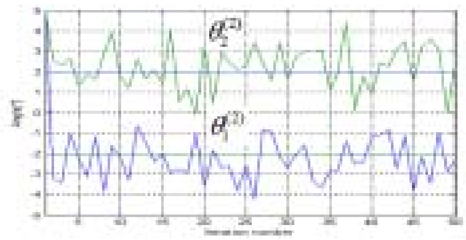

Figure 3 Second Markov chain (SNR=-5dB)

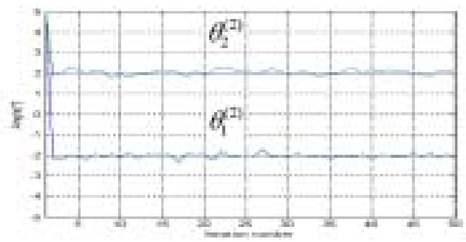

Figure 4 Second Markov chain (SNR=15dB)

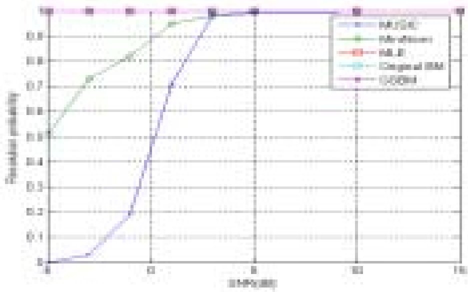

Figure 5 Resolution probability

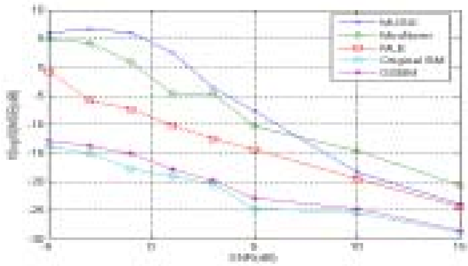

Figure 6 Mean square error

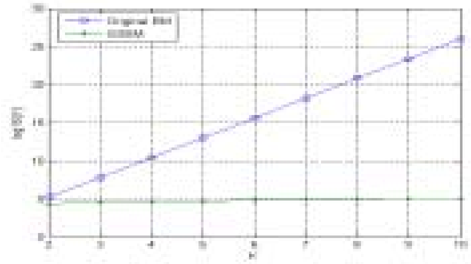

Figure 7 Computation comparison

6. SUMMARY

In this paper a new method (GSBM) is presented which combines the Bayesian maximum a posterior high-resolution DOA estimator with the Gibbs sampling. The formulation of the new method has been deduced and its promising performance has also been investigated. It has been shown that the new estimator possesses not only good performance as original BM method, but also obviously reduces original computation complexity from $O(L^K)$ to $O(K \times J \times N_s)$. The simulations also demonstrate that the performance of GSBM is better than MLE, MUSIC and MiniNorm, especially in low SNRs.

REFERENCES

[1] C. Cho and P. M. Djurić, "Detection and Estimation of DOA's of Signals via Bayesian Predictive Densities", IEEE Transactions on Signal Processing, vol.42, no.11, pp.3051-3060, Nov.1994.

[2] J. Huang, J.Chen, C.Liu and P. M. Djurić, "A Bayesian approach to high resolution direction-of-arrival estimation", in Proc. ICSP'98, Beijing, China, October 12-16, 1998, pp.377-379

[3] J. Huang, P. Xu, Y. Lu and Y. Sun, "A Novel Bayesian High-Resolution Direction-of-Arrival Estimator", in Proc. Oceans'2001 MTS/IEEE, Honolulu, U.S.A, November 5-8, 2001, pp.1697-1702

[4] J.J.K.O'Ruanaidh and W.J.Fitzgerald, Numerical Bayesian Methods Applied to Signal Processing, New York: Springer Verlag, 1996

[5] M.Chen, Q.Shao and J.G.Ibrahim, Monte Carlo Methods in Bayesian Computation, New York: Springer Verlag, 2000

[6] G. Casella and E. I. George, "Explaining the Gibbs sampler", American Statistician, vol.46, pp. 167-174, Aug.1992.

[7] Smith, A.F.M., Gelfand, A.E., "Bayesian statistics without tears: a sampling-resampling perspective", American Statistician, Vol. 46, pp.84-88, May.1992