Research Article

Estimating Frequency by Interpolation Using Least Squares Support Vector Regression

Changwei Ma

Department of Computer Science, ABA Teachers College, Wenchuan 623002, China

Correspondence should be addressed to Changwei Ma; changweim@126.com

Discrete Fourier transform- (DFT-) based maximum likelihood (ML) algorithm is an important part of single sinusoid frequency estimation. As signal to noise ratio (SNR) increases and is above the threshold value, it will lie very close to Cramer-Rao lower bound (CRLB), which is dependent on the number of DFT points. However, its mean square error (MSE) performance is directly proportional to its calculation cost. As a modified version of support vector regression (SVR), least squares SVR (LS-SVR) can not only still keep excellent capabilities for generalizing and fitting but also exhibit lower computational complexity. In this paper, therefore, LS-SVR is employed to interpolate on Fourier coefficients of received signals and attain high frequency estimation accuracy. Our results show that the proposed algorithm can make a good compromise between calculation cost and MSE performance under the assumption that the sample size, number of DFT points, and resampling points are already known.

1. Introduction

Estimating frequency of a single sinusoid based on discrete Fourier transform (DFT) has attracted considerable attentions for many decades. Rife and Boorstyn exploited the relationship of maximum likelihood estimator (MLE) to DFT and proposed a frequency domain periodogram (FDP) algorithm having two stages: coarse search and fine search [1]. In order to reduce the calculation cost, the frequency step of coarse search is appropriately broadened. As a result, the absolute frequency value is usually not in the DFT points. Hence, a fine search will be necessary for higher frequency estimation accuracy.

A great deal of fine search algorithms has erupted mainly from two sides: the dichotomous search and interpolation on Fourier coefficients. During the former, an iterative binary search for the true signal frequency has been presented, which is particularly suited for digital signal processing (DSP) implementation [2]. In [3], the same authors have proposed a number of hybrid estimators that combine the dichotomous search with various interpolation techniques in order to reduce the computational complexity at the expense of acquisition range; and other modified dichotomous search

frequency estimators have been addressed in [4–6]. In the latter, complex Fourier coefficients have been utilized to interpolate the true signal frequency between the maximum and the second highest bin [7]. However, it has been shown to have a frequency dependent performance [8]. Two improved estimators have been proposed, which were implemented iteratively [9, 10]. Rational combination of three spectrum lines (RCTSL) has been employed as the fine estimation because of its constant combinational weights in least square approximation [11]. Other methods used for interpolation include Lagrange interpolator [12], L-filter DFT [13], nonlinear filter [14], Kaiser window [15], trigonometric polynomial interpolator [16], and narrowband approximation [17]. However, all of the DFT-based algorithms attain better mean square error (MSE) performance at the expense of calculation cost.

As the concrete implementation of statistical learning theory (SLT) and structure risk minimization (SRM) principle, SVR overcomes the overfitting and local minimum problems currently existing in artificial neural network (ANN). Least squares support vector regression (LS-SVR) has the improvements of support vector regression (SVR): inequality constraint is substituted by equality one; a squared

loss function is taken for the error variable. Consequently, its good generalizing and learning capabilities are exploited for interpolation on Fourier coefficients, and a novel DFT-based frequency estimation algorithm adopting LS-SVR for fine search is proposed. Our results show that under the assumption that the sample size, number of DFT points, and resampling points are already known, the proposed algorithm can derive low calculation cost while still keeping good MSE performance.

The remainder of this paper is organized as follows. Section 2 concretely describes the process of our algorithm. In Section 3, interpolation by LS-SVR is analyzed, and its parameters are selected properly to improve generalization capability. Section 4 shows the results of simulations and experiments. The paper is finally concluded in Section 5.

2. The Proposed Algorithm

2.1. Signal Model. The sinusoid signal polluted by noise is modeled as

$$r_n = Ae^{j(2\pi fn + \theta)} + w_n, \quad n = 0, \ldots, N - 1. \tag{1}$$

Here, $A > 0$, $f \in [-0.5, 0.5)$, and $\theta \in [-\pi, \pi)$ are the amplitude, deterministic but unknown frequency, and initial phase, respectively; w_n is an independent complex additive white Gaussian noise (AWGN) with zero mean and variance σ^2; N is the sample size.

2.2. An Interpolation by LS-SVR in DFT-Based Frequency Estimation Algorithm. The key idea proposed in [1] is exploited and a two-step algorithm including coarse search and fine search is proposed. At first, FDP algorithm is employed and the number of DFT points is extended to K by appending $(K - N)$ zeros. Now, the frequency step is $1/K$, and

$$m = \arg \max_k \{|R(f_k)|\}, \quad k = 0, \ldots, K - 1, \tag{2}$$

where $R(f_k) = \sum_{n=0}^{N-1} r_n e^{-j2\pi f_k n}$, $f_k = k/K$. For the sake of convenience of fast Fourier transform (FFT) calculation, K is usually set to $2N, 4N, 8N, \ldots$.

Secondly, a training set $S = \{(x_n, y_n) \mid n = 1, \ldots, K\}$ is constructed and LS-SVR is utilized to fit it, where $x_n = n - 1$ and $y_n = R(n - 1)$. Because y_n is a complex series and LS-SVR is only suitable for real number, y_n is divided into real and image part.

(1) Assuming that $f_{\text{real}}(x) = (\mathbf{w}_{\text{real}} \cdot \phi_{\text{real}}(x)) + b_{\text{real}}$ insofar as for ε to completely fit all elements of $S_{\text{real}} = \{(x_n^{\text{real}}, y_n^{\text{real}}) \mid n = 1, \ldots, K\}$, where $x_n^{\text{real}} = x_n = n - 1$, $y_n^{\text{real}} = \text{Re}[y_n] = \text{Re}[R(n-1)]$, $\text{Re}[\cdot]$ is the operation of taking real part, $(\cdot)$ is an inner product operator, and $\phi_{\text{real}}(\cdot)$ is a nonlinear mapping

from low to high dimension feature space. ε-insensitive loss function is defined as

$$
\begin{aligned}
&L\left(y_i^{\text{real}}, f_{\text{real}}\left(x_i^{\text{real}}\right)\right) \\
&= \left|y_i^{\text{real}} - f_{\text{real}}\left(x_i^{\text{real}}\right)\right|_\varepsilon \\
&= \begin{cases} 0, & \left|y_i^{\text{real}} - f_{\text{real}}\left(x_i^{\text{real}}\right)\right| \le \varepsilon \\ \left|y_i^{\text{real}} - f_{\text{real}}\left(x_i^{\text{real}}\right)\right| - \varepsilon, & \text{else,} \end{cases} \\
&\hspace{5cm} i = 1, \ldots, K,
\end{aligned}
\tag{3}
$$

where ε is the coefficient of insensitive loss function. d_i^{real} is denoted as the distance from point $(x_i^{\text{real}}, y_i^{\text{real}}) \in S_{\text{real}}$ to $f_{\text{real}}(x)$:

$$
\begin{aligned}
d_i^{\text{real}} &= \frac{\left|\left(\mathbf{w}_{\text{real}} \cdot \phi_{\text{real}}\left(\mathbf{x}_i^{\text{real}}\right)\right) + b_{\text{real}} - y_i^{\text{real}}\right|}{\sqrt{1 + \|\mathbf{w}_{\text{real}}\|^2}} \\
&\le \frac{\varepsilon}{\sqrt{1 + \|\mathbf{w}_{\text{real}}\|^2}}, \quad i = 1, \ldots, K.
\end{aligned}
\tag{4}
$$

According to (4), $f_{\text{real}}(x)$ is optimized through maximizing $\varepsilon / \sqrt{1 + \|\mathbf{w}_{\text{real}}\|^2}$, that is minimizing $\|\mathbf{w}_{\text{real}}\|^2$. Thereby, LS-SVR is presented as

$$
\begin{aligned}
\min \quad & J(\mathbf{w}_{\text{real}}, b_{\text{real}}) = \frac{1}{2} \|\mathbf{w}_{\text{real}}\|^2, \\
\text{s.t.} \quad & \left|\left(\mathbf{w}_{\text{real}} \cdot \phi_{\text{real}}\left(x_i^{\text{real}}\right)\right) + b_{\text{real}} - y_i^{\text{real}}\right| \le \varepsilon, \\
& \hspace{3cm} i = 1, \ldots, K.
\end{aligned}
\tag{5}
$$

In fact, fitting errors larger than ε always exist. By introducing error variables e_i^{real} and least squares (LS) method, (5) is converted into

$$
\begin{aligned}
\min \quad & J(\mathbf{w}_{\text{real}}, b_{\text{real}}) = \frac{1}{2} \|\mathbf{w}_{\text{real}}\|^2 + \frac{C}{2} \sum_{i=1}^{K} \left(e_i^{\text{real}}\right)^2 \\
\text{s.t.} \quad & y_i^{\text{real}} = \left(\mathbf{w}_{\text{real}} \cdot \phi_{\text{real}}\left(x_i^{\text{real}}\right)\right) + b_{\text{real}} + e_i^{\text{real}}, \\
& \hspace{3cm} i = 1, \ldots, K,
\end{aligned}
\tag{6}
$$

where penalty factor C is a positive constant to take compromise in LS-SVR's generalization capability and fitting errors which are denoted by the first and second item of $J(\mathbf{w}_{\text{real}}, b_{\text{real}})$, respectively.

Next step, using Lagrange multiplier method and replacing $(\phi_{\text{real}}(x_i^{\text{real}}) \cdot \phi_{\text{real}}(x_j^{\text{real}}))$ with kernel function $K(x_i^{\text{real}}, x_j^{\text{real}})$,

$$
\begin{bmatrix} 0 & \mathbf{1}^{\text{T}} \\ \mathbf{1} & \mathbf{Q} + \dfrac{1}{C}\mathbf{I} \end{bmatrix} \begin{bmatrix} b_{\text{real}} \\ \boldsymbol{\alpha}_{\text{real}} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{Y}_{\text{real}} \end{bmatrix},
\tag{7}
$$

where $\mathbf{1} = \underbrace{(1, \ldots, 1)}_{K}^{\text{T}}$, $\mathbf{Y}_{\text{real}} = (y_1^{\text{real}}, \ldots, y_K^{\text{real}})^{\text{T}}$, $\boldsymbol{\alpha}_{\text{real}} = (\alpha_1^{\text{real}}, \ldots, \alpha_K^{\text{real}})^{\text{T}}$, $\mathbf{I}$ is an unit matrix with K orders, $[\cdot]^{\text{T}}$ is

a transpose operator, $\alpha_1^{\text{real}}, \ldots, \alpha_K^{\text{real}}$ are Lagrange multipliers, and $\mathbf{Q}$ is a kernel function matrix. Radius basis function (RBF) is adopted in this study, so

$$
\begin{aligned}
\mathbf{Q}_{ij} = K\left(x_i^{\text{real}}, x_j^{\text{real}}\right) &= \exp\left(\frac{-\left\|x_i^{\text{real}} - x_j^{\text{real}}\right\|^2}{h^2}\right) \\
&= \exp\left(\frac{-\|i-j\|^2}{h^2}\right),
\end{aligned}
\tag{8}
$$

where $\mathbf{Q}_{ij}$ is the (i,j)th element of $\mathbf{Q}$ and positive constant h is the width of RBF.

Solving (7) and deriving the regression as

$$
\begin{aligned}
f_{\text{real}}(x) &= \left(\mathbf{w}_{\text{real}} \cdot \phi_{\text{real}}(x)\right) + b_{\text{real}} \\
&= \sum_{i=1}^{K} \alpha_i^{\text{real}} K\left(x_i^{\text{real}}, x\right) + b_{\text{real}}.
\end{aligned}
\tag{9}
$$

(2) Assuming that $f_{\text{imag}}(x) = \left(\mathbf{w}_{\text{imag}} \cdot \phi_{\text{imag}}(x)\right) + b_{\text{imag}}$ insofar as for ε to completely fit all elements of $S_{\text{imag}} = \{(x_n^{\text{imag}}, y_n^{\text{imag}}) \mid n = 1, \ldots, K\}$, where $x_n^{\text{imag}} = x_n = n - 1$, $y_n^{\text{imag}} =$ and $\text{Im}[y_n] = \text{Im}[R(n-1)]$, $\text{Im}[\cdot]$ is the operation of taking image part, deriving $f_{\text{imag}}(x) = \left(\mathbf{w}_{\text{imag}} \cdot \phi_{\text{imag}}(x)\right) + b_{\text{imag}} = \sum_{i=1}^{K} \alpha_i^{\text{imag}} K(x_i^{\text{imag}}, x) + b_{\text{imag}}$ as (1).

(3) $f(x) = f_{\text{real}}(x) + j \times f_{\text{imag}}(x)$.

Then, $f(x)$ is interpolated by $1/L$ interval between $m-1$ and $m+1$, where L is usually set to $2, 4, \ldots$. It means that $z_l = m - 1 + l/L, l = 1, \ldots, 2L-1$ are substituted into $f(x)$.

In the end, searching the series $f(z_l)$ and deriving $\widehat{f}_{\text{LS-SVR}} = ((L(m-1) + \arg\max_{l=1,\ldots,2L-1} f(z_l))/KL)_{-0.5}^{0.5}$ as the ultimate frequency estimation value, where the modulo-1 operation $(\cdot)_{-0.5}^{0.5}$ means that the value is reduced to the interval $[-0.5, 0.5]$.

3. Analyses

Taking the real part of y_n, for example, now $\mathbf{Y}_{\text{real}} = (y_1^{\text{real}}, \ldots, y_K^{\text{real}})^{\text{T}} = (\text{Re}[R(0)], \ldots, \text{Re}[R(K-1)])$. And, from (7),

$$
\begin{aligned}
b_{\text{real}} &= \frac{\mathbf{1}_K^{\text{T}} \left(\mathbf{Q} + (1/C)\,\mathbf{I}\right)^{-1} \mathbf{Y}_{\text{real}}}{\mathbf{1}_K^{\text{T}} \left(\mathbf{Q} + (1/C)\,\mathbf{I}\right)^{-1} \mathbf{1}_{K+1}}, \\
\boldsymbol{\alpha}_{\text{real}} &= \left(\mathbf{Q} + \frac{1}{C}\mathbf{I}\right)^{-1} \left(\mathbf{Y}_{\text{real}} - \mathbf{1}_K b_{\text{real}}\right),
\end{aligned}
\tag{10}
$$

where $(\cdot)^{-1}$ is the inverse operation.

$\widehat{\mathbf{Y}}_{\text{real}} = (\widehat{y}_1^{\text{real}}, \ldots, \widehat{y}_{2L-1}^{\text{real}})^{\text{T}} = (f_{\text{real}}(z_1), \ldots, f_{\text{real}}(z_{2L-1}))^{\text{T}}$ is defined as LS-SVR's output of $\mathbf{Y}_{\text{real}}$, and by substituting (10) into (9), we have

$$
\begin{aligned}
\widehat{\mathbf{Y}}_{\text{real}} &= \mathbf{Q}' \boldsymbol{\alpha}_{\text{real}} + \mathbf{1}_{2L-1} b_{\text{real}} = \mathbf{Q}' \left(\mathbf{Q} + \frac{1}{C}\mathbf{I}\right)^{-1} \mathbf{Y}_{\text{real}} \\
&\quad + \left[\mathbf{I}' - \mathbf{Q}'\left(\mathbf{Q} + \frac{1}{C}\mathbf{I}\right)^{-1}\right]\mathbf{1}_K b_{\text{real}} = \left\{\mathbf{Q}'\left(\mathbf{Q} + \frac{1}{C}\mathbf{I}\right)^{-1}\right. \\
&\quad \left. + \left[\mathbf{I}' - \mathbf{Q}'\left(\mathbf{Q} + \frac{1}{C}\mathbf{I}\right)^{-1}\right]\frac{\mathbf{1}_K \mathbf{1}_K^{\text{T}} \left(\mathbf{Q} + (1/C)\,\mathbf{I}\right)^{-1}}{\mathbf{1}_K^{\text{T}} \left(\mathbf{Q} + (1/C)\,\mathbf{I}\right)^{-1} \mathbf{1}_K}\right\} \\
&\quad \cdot \mathbf{Y}_{\text{real}} = \boldsymbol{\beta}\mathbf{Y}_{\text{real}},
\end{aligned}
\tag{11}
$$

where $\mathbf{1}_{2L-1} = \underbrace{(1, \ldots, 1)^{\text{T}}}_{2L-1}$, $Q'_{ij} = K(z_i, x_j^{\text{real}})$ is the (i,j)th element of matrix $\mathbf{Q}'$ with $2L-1$ rows and K columns, $\mathbf{I}' = \begin{bmatrix} \mathbf{1}_{2L-1} & \mathbf{0}_{(2L-1)\times(K-1)} \end{bmatrix}$, $\mathbf{1}_{2L/K-1} = \underbrace{(1, \ldots, 1)^{\text{T}}}_{2L/K-1}$, and $\mathbf{0}_{(2L-1)\times(K-1)}$ is an all-zero matrix with $2L-1$ rows and $K-1$ columns.

From (11), the LS-SVR process can be regarded as a linear interpolator whose output is the linear weighted sum of all inputs, where the weighted coefficients are dependent on $\boldsymbol{\beta} = \{\mathbf{Q}'(\mathbf{Q} + (1/C)\mathbf{I})^{-1} + [\mathbf{I}' - \mathbf{Q}'(\mathbf{Q} + (1/C)\mathbf{I})^{-1}](\mathbf{1}_K \mathbf{1}_K^{\text{T}}(\mathbf{Q} + (1/C)\mathbf{I})^{-1}/\mathbf{1}_K^{\text{T}}(\mathbf{Q} + (1/C)\mathbf{I})^{-1}\mathbf{1}_K)\}$, which means it is totally related with $\mathbf{Q}$, $\mathbf{Q}'$ and C.

When C decreases rapidly, $\mathbf{Q} + (1/C)\mathbf{I} \approx (1/C)\mathbf{I}$, $\boldsymbol{\beta} \approx \mathbf{1}_K \mathbf{1}_K^{\text{T}}/K$. From (11), $\widehat{y}_n^{\text{real}}$ at the different moments keep the same value obviously, which means the LS-SVR process can not fit $\mathbf{Y}_{\text{real}}$ accurately. Hence, C is set as large as possible. $f = 0.1$, $\theta = 0$, $N = 32$, $K = 2N$, $L = 8$, and $h = 1$ are set and the number of Monte Carlo experiments is 10000. Figure 1 illustrates the impact of C on the estimation performance, which is consistent with all analyses above. When C is larger than 100, its MSE performance will almost not increase anymore, so $C = 100$ is set.

At the same time, according to $K(x_i^{\text{real}}, x_j^{\text{real}}) = \exp[-(i-j)^2/h^2]$ and the definitions of Q_{ij} and Q'_{ij} in (8) and (11), it is concluded that

$$
\begin{aligned}
\lim_{h \to 0} Q_{ij} &= \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise,} \end{cases} \\
\lim_{h \to 0} Q'_{ij} &= \begin{cases} 1, & \text{if } i = L, \ j = 1 \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}
\tag{12}
$$

So when

$$
\begin{aligned}
C &= 100, \\
N &= 32, \\
\lim_{h \to 0} \beta_{ij} &= \begin{cases} 1, & \text{if } i = L, j = 1 \\ 0, & \text{if } i = L, \text{ except } j = 1 \\ \dfrac{1}{K}, & \text{otherwise,} \end{cases}
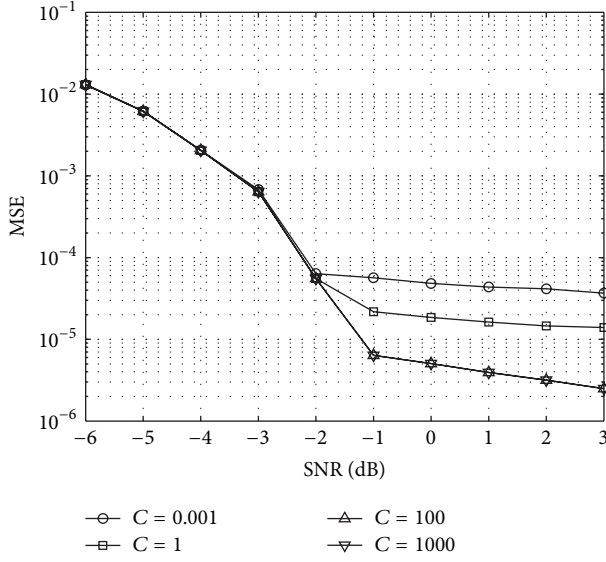\end{aligned}
\tag{13}
$$

FIGURE 1: Impact of C on MSE with $f = 0.1, \theta = 0, N = 32, K = 2N$, $L = 8$, and $h = 1$.
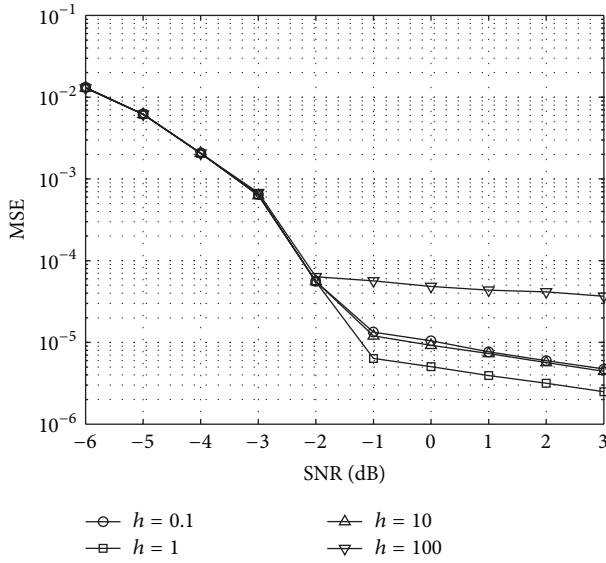

FIGURE 2: Impact of h on MSE with $f = 0.1, \theta = 0, N = 32, K = 2N$, $L = 8$, and $C = 100$.

where β_{ij} is (i, j)th element of matrix $\boldsymbol{\beta}$ with $2L - 1$ rows and K columns. It means now the LS-SVR process can only make ensure the accuracy of spectrum of original signals but not interpolate effectively. On the contrary, as a result of $\lim_{h \to +\infty} Q_{ij} = 1$, $\lim_{h \to +\infty} Q'_{ij} = 1$, now $\lim_{h \to +\infty} \beta_{ij} = 1/K$, which means that now the LS-SVR process can also not fit $\mathbf{Y}_{\text{real}}$ accurately. Everything is as in Figure 1 except that $C = 100$; Figure 2 illustrates the impact of h on the estimation performance when h is 0.1, 1, 10, and 100, respectively, which is consistent with all analyses above. So $h = 1$ is set.

Assuming that the received signal is $r_n = g_n + w_n$, where $g_n = A\exp[j(2\pi fn + \theta)]$, w_n are signal and noise components

of r_n, the true interpolation value of spectrum of r_n at point z_l is

$$
\begin{aligned}
R(z_l) &= \frac{1}{K} \sum_{n=0}^{N-1} r_n \exp\left(-\frac{j2\pi nz_l}{K}\right) \\
&= \frac{1}{K} \sum_{n=0}^{N-1} (g_n + w_n) \exp\left(-\frac{j2\pi nz_l}{K}\right) \\
&= \frac{1}{K} \sum_{n=0}^{N-1} g_n \exp\left(-\frac{j2\pi nz_l}{K}\right) \\
&\quad + \frac{1}{K} \sum_{n=0}^{N-1} w_n \exp\left(-\frac{j2\pi nz_l}{K}\right).
\end{aligned}
\tag{14}
$$

At the same time, the estimating interpolation value of spectrum of r_n at point z_l is attained through LS-SVR. Consider

$$
\begin{aligned}
f(z_l) &= f_{\text{real}}(z_l) + jf_{\text{imag}}(z_l) = \sum_{i=1}^{K} \beta_{li} \text{Re}\left[R(i-1)\right] \\
&\quad + j \sum_{i=1}^{K} \beta_{li} \text{Im}\left[R(i-1)\right] = \sum_{i=1}^{K} \beta_{li} R(i-1) \\
&= \sum_{i=1}^{K} \beta_{li} \left\{ \frac{1}{K} \sum_{n=0}^{N-1} g_n \exp\left[-\frac{j2\pi n(i-1)}{K}\right] \right. \\
&\quad \left. + \frac{1}{K} \sum_{n=0}^{N-1} w_n \exp\left[-\frac{j2\pi n(i-1)}{K}\right] \right\} = \frac{1}{K} \\
&\quad \cdot \sum_{n=0}^{N-1} g_n \sum_{i=1}^{K} \beta_{li} \exp\left[-\frac{j2\pi n(i-1)}{K}\right] + \frac{1}{K} \\
&\quad \cdot \sum_{n=0}^{N-1} w_n \sum_{i=1}^{K} \beta_{li} \exp\left[-\frac{j2\pi n(i-1)}{K}\right].
\end{aligned}
\tag{15}
$$

As a result, the error between (14) and (15) is

$$
\begin{aligned}
e(z_l) &= R(z_l) - f(z_l) = \frac{1}{K} \sum_{n=0}^{N-1} g_n \left\{ \exp\left(-\frac{j2\pi nz_l}{K}\right) \right. \\
&\quad \left. - \sum_{i=1}^{K} \beta_{li} \exp\left[-\frac{j2\pi n(i-1)}{K}\right] \right\} + \frac{1}{K} \\
&\quad \cdot \sum_{n=0}^{N-1} w_n \left\{ \exp\left(-\frac{j2\pi nz_l}{K}\right) \right. \\
&\quad \left. - \sum_{i=1}^{K} \beta_{li} \exp\left[-\frac{j2\pi n(i-1)}{K}\right] \right\},
\end{aligned}
\tag{16}
$$

where the noise item $w(z_l) = (1/K) \sum_{n=0}^{N-1} w_n \{\exp(-j2\pi nz_l/K) - \sum_{i=1}^{K} \beta_{li} \exp[-j2\pi n(i -$

1)$/K]\}$ is still an independent complex AWGN. Its mean and variance are listed as follows:

$$
\xi_w(z_l) = E\{w(z_l)\} = E\left\{\frac{1}{K}\sum_{n=0}^{N-1} w_n\left\{\exp\left(-\frac{j2\pi nz_l}{K}\right)\right.\right.
$$

$$
\left.\left. -\sum_{i=1}^{K}\beta_{li}\exp\left[-\frac{j2\pi n(i-1)}{K}\right]\right\}\right\} = 0,
$$

$$
\sigma_w^2(z_l) = E\left\{|w(z_l) - E\{w(z_l)\}|^2\right\} = E\left\{|w_l|^2\right\} = \frac{\sigma^2}{K^2} \quad (17)
$$

$$
\cdot \sum_{n=0}^{N-1}\left|\exp\left(-\frac{j2\pi nz_l}{K}\right)\right.
$$

$$
\left. -\sum_{i=1}^{K}\beta_{li}\exp\left[-\frac{j2\pi n(i-1)}{K}\right]\right|^2.
$$

Then, the mean and variance of $e(z_l)$ are

$$
\xi_e(z_l) = E\{e(z_l)\} = \frac{1}{K^2}\sum_{n=0}^{N-1}g_n\left\{\exp\left(-\frac{j2\pi nz_l}{K}\right)\right.
$$

$$
\left. -\sum_{i=1}^{K}\beta_{li}\exp\left[-\frac{j2\pi n(i-1)}{K}\right]\right\},
$$

$$
\sigma_e^2(z_l) = E\left\{|e(z_l) - E\{e(z_l)\}|^2\right\} = E\{|w(z_l) \quad (18)
$$

$$
- E\{w(z_l)\}|^2\} = \frac{\sigma^2}{K^2}\sum_{n=0}^{N-1}\left|\exp\left(-\frac{j2\pi nz_l}{K}\right)\right.
$$

$$
\left. -\sum_{i=1}^{K}\beta_{li}\exp\left[-\frac{j2\pi n(i-1)}{K}\right]\right|^2.
$$

Define

$$
\frac{E\{\sigma_e^2(z_l)\}}{E\{|R(z_l)|^2\}} = \sigma^2\frac{E\left\{\sum_{n=0}^{N-1}\left|\exp(-j2\pi nz_l/K) - \sum_{i=1}^{K}\beta_{li}\exp[-j2\pi n(i-1)/K]\right|^2\right\}}{E\left\{\left|\sum_{n=0}^{N-1}r_n\exp(-j2\pi nz_l/K)\right|^2\right\}},
$$

$$
\frac{E\{|\xi_e(z_l)|^2\}}{E\{|R(z_l)|^2\}} = \frac{E\left\{\left|\sum_{n=0}^{N-1}g_n\left\{\exp(-j2\pi nz_l/K) - \sum_{i=1}^{K}\beta_{li}\exp[-j2\pi n(i-1)/K]\right\}\right|^2\right\}}{E\left\{\left|\sum_{n=0}^{N-1}r_n\exp(-j2\pi nz_l/K)\right|^2\right\}}.
$$

$$(19)$$

Everything is as in Figure 1 except that $C = 100$, $L = 4$, and SNR $= 0$ dB; the values of (19) with different z_l are calculated:

$$
\left[\frac{E\{\sigma_e^2(z_1)\}}{E\{|\xi_e(z_1)|^2\}},\ldots,\frac{E\{\sigma_e^2(z_7)\}}{E\{|\xi_e(z_7)|^2\}}\right]^{\mathrm{T}} = \left[7.59\right.
$$

$$
\times 10^{-5}, 1.04\times 10^{-4}, 7.43\times 10^{-5}, 4.51\times 10^{-5}, 7.39
$$

$$
\times 10^{-5}, 1.03\times 10^{-4}, 7.48\times 10^{-5}\right]^{\mathrm{T}},
$$

$$
\left[\frac{E\{|\xi_e(z_1)|^2\}}{E\{|R(z_1)|^2\}},\ldots,\frac{E\{|\xi_e(z_7)|^2\}}{E\{|R(z_7)|^2\}}\right]^{\mathrm{T}} = \left[2.32\right.
$$

$$
\times 10^{-6}, 4.27\times 10^{-6}, 2.39\times 10^{-6}, 1.23\times 10^{-6}, 4.52
$$

$$
\times 10^{-6}, 7.15\times 10^{-6}, 4.69\times 10^{-6}\right]^{\mathrm{T}}.
$$

$$(20)$$

The values of (20) are very small, which means when SNR is higher than threshold, the estimating value through LS-SVR and true value of spectrum of r_n at point z_l are almost the same. It validates the validity of interpolation in our proposed algorithm.

4. Simulations and Experiments

The proposed algorithm entitled LS-SVR is compared with other fives: FDP proposed in [1]; dichotomous search (DS) proposed in [2]; modify dichotomous search (Modify DS) proposed in [4]; iterative interpolation on Fourier coefficients (IIFC) proposed in [9]; RCTSL proposed in [11]. Analyzing these five algorithms, it is known that the MSE performances of FDP and RCTSL algorithms will be improved as their DFT points increase, and the MSE performances of DS, Modified DS, and IIFC algorithms will be improved as their iterative numbers increase. In order to compare these algorithms' calculation cost under nearly the same precondition, $K = 32N$ in FDP, $K = 2N$, $Q_{DS} = 20$ in DS, $Q_{M-DS} = 20$ in Modified DS, $Q_{IIFC} = 10$ in IIFC, $K = 8N$ in RCTSL, $K = 2N$, and $L = 8$ in LS-SVR are set to keep their MSE performance almost the same.

4.1. Mean Performance. Everything is as in Figure 1 except that $C = 100$. Figure 3 illustrates the mean of these six algorithms, while SNR is -4 dB and 0 dB, respectively. Obviously, when SNR is low, the unbiased performances of Modified DS and IIFC algorithms are worse than the other four.

4.2. Frequency Estimation Performance. As in Figure 3, Figure 4 illustrates the MSE curves of these six algorithms versus SNR. From Figure 4, the threshold values of DS,
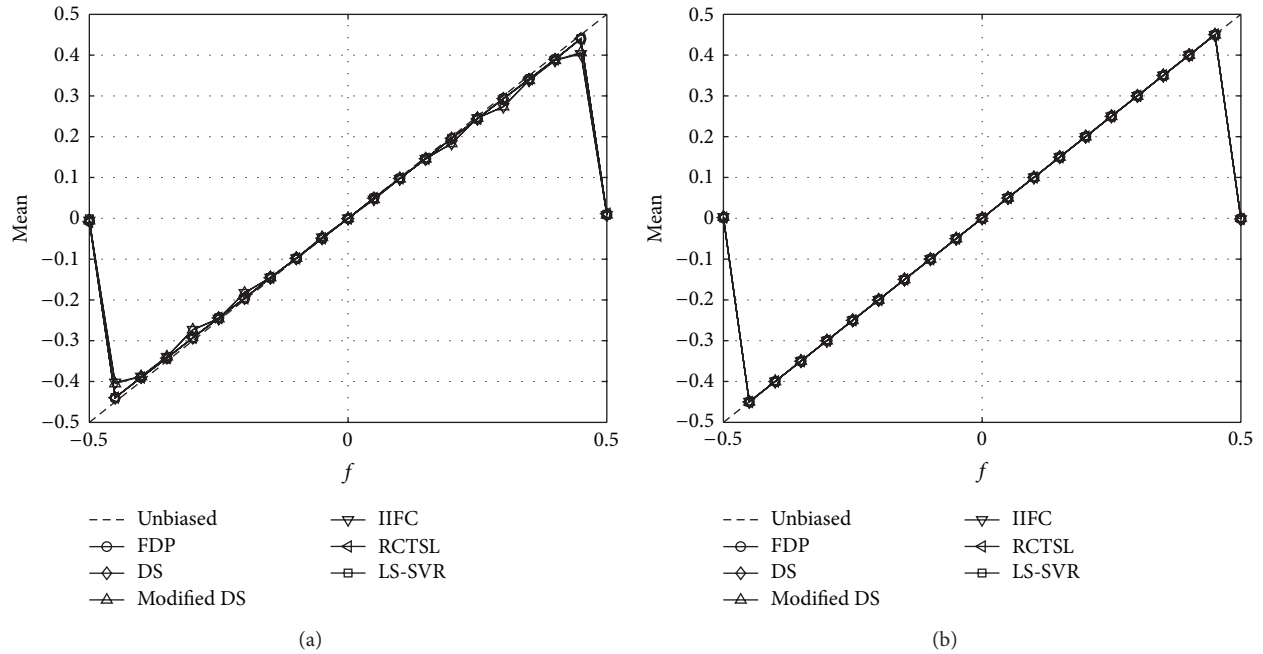
(a)

(b)

FIGURE 3: (a) Mean with $\theta = 0$, $N = 32$, $K = 2N$, $L = 8$, $C = 100$, $h = 1$, and SNR $= -4$ dB. (b) Mean with $\theta = 0$, $N = 32$, $K = 2N$, $L = 8$, $C = 100$, $h = 1$, and SNR $= 0$ dB.
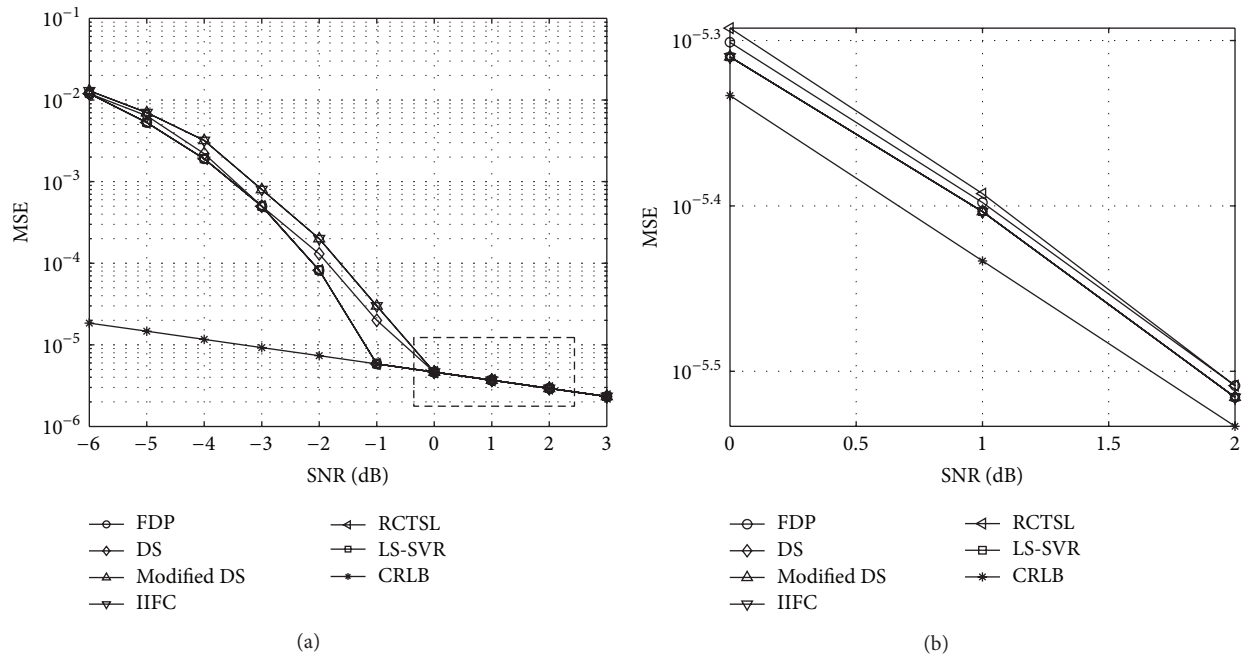

(a)

(b)

FIGURE 4: (a) MSE of frequency estimation versus SNR with $f = 0.1$, $\theta = 0$, $N = 32$, $K = 2N$, $L = 8$, $C = 100$, $h = 1$, and SNR $\in [-6, 3]$ dB. (b) MSE of frequency estimation versus SNR with $f = 0.1$, $\theta = 0$, $N = 32$, $K = 2N$, $L = 8$, $C = 100$, $h = 1$, and SNR $\in [0, 2]$ dB.

Modified DS, and IIFC algorithms are higher than other three ones. And, these six algorithms' MSEs are kept almost the same when they are all above their threshold values, where LS-SVR one is the best one.

4.3. Impact of DFT Points K. As in Figure 3, the MSE curves of LS-SVR algorithm versus SNR are plotted in Figure 5,

while K is N, $2N$, $4N$, and $8N$, respectively. Now, the sample size and interpolation times are fixed. From Figure 5, when $K = N$, its MSE value will not approach Cramer-Rao lower bound (CRLB) as SNR increases, which means now the LS-SVR algorithm is ineffective; when $K = 2N$, $4N$, and $8N$, their MSE values are nearly the same, and the $K = 2N$ condition is a little better than the two others. The reason is
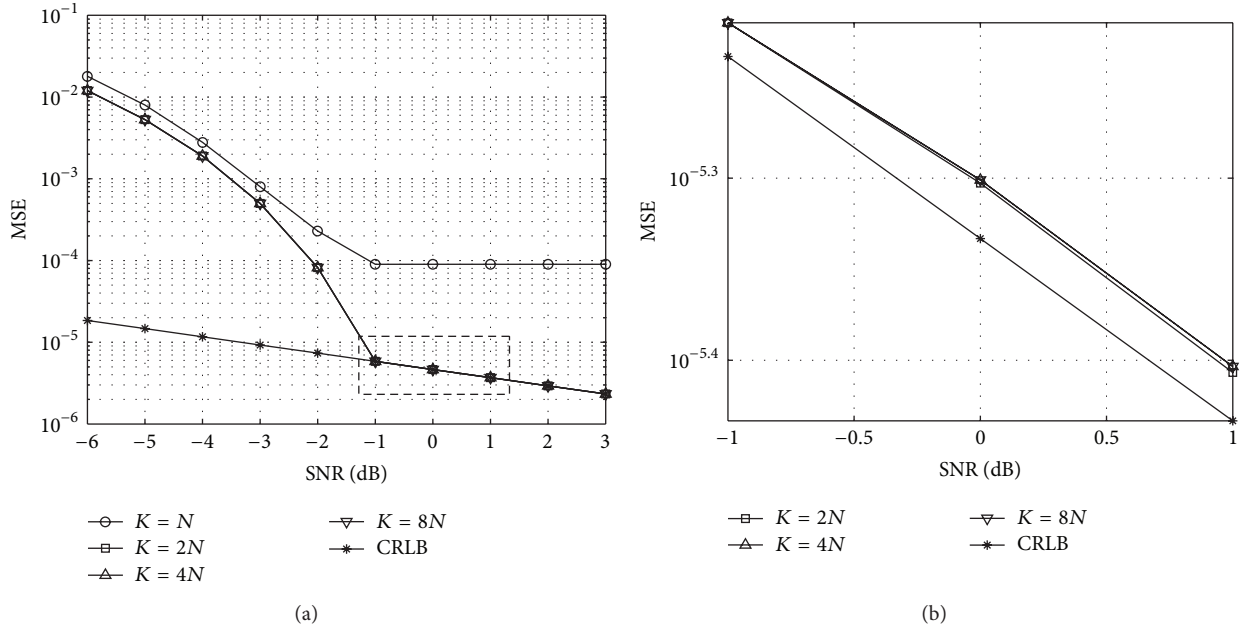
(a)

(b)

FIGURE 5: (a) Impact of K on MSE with $f = 0.1$, $\theta = 0$, $N = 32$, $L = 8$, $C = 100$, $h = 1$, and SNR $\in [-6, 3]$ dB. (b) Impact of K on MSE with $f = 0.1$, $\theta = 0$, $N = 32$, $L = 8$, $C = 100$, $h = 1$, and SNR $\in [-1, 1]$ dB.
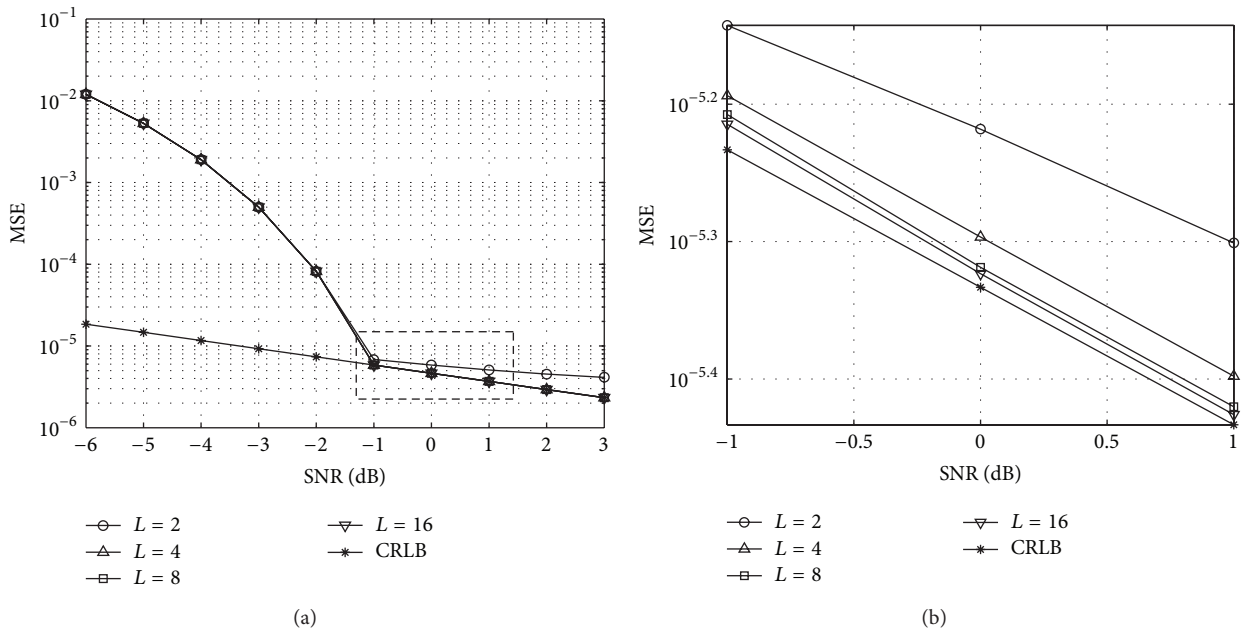

(a)

(b)

FIGURE 6: (a) Impact of L on MSE with $f = 0.1$, $\theta = 0$, $N = 32$, $K = 2N$, $C = 100$, $h = 1$, and SNR $\in [-6, 3]$ dB. (b) Impact of L on MSE with $f = 0.1$, $\theta = 0$, $N = 32$, $K = 2N$, $C = 100$, $h = 1$, and SNR $\in [-1, 1]$ dB.

the fact that the proposed interpolation process is fulfilled by LS-SVR fitting in the paper; too many interpolation numbers during the adjacent DFT points will result in overfitting. Besides, the calculation cost of $K = 2N$ is less than $K = 4N$ and $8N$; hence $K = 2N$ is set in this paper.

4.4. Impact of Interpolation Times L. As in **Figure 3**, the MSE curves of LS-SVR algorithm versus SNR are plotted in

Figure 6, while L is 2, 4, 8, and 16, respectively. Now, the sample size and DFT points are fixed. From **Figure 6**, its MSE performance will be improved as L increases.

4.5. Impact of Sample Size N. As in **Figure 3**, the MSE curves of LS-SVR algorithm versus SNR are plotted in **Figure 7**, while N is 8, 16, 32, and 64, respectively. Now, the interpolation times are fixed, and DFT points are directly proportional
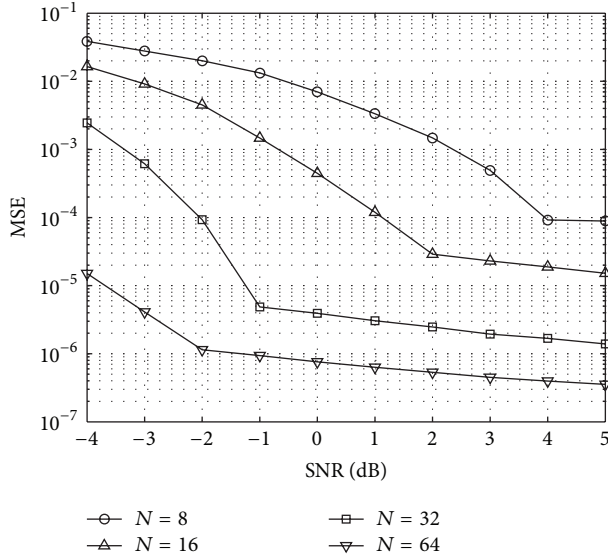
FIGURE 7: Impact of N on MSE with $f = 0.1$, $\theta = 0$, $K = 2N$, $L = 8$, $C = 100$, $h = 1$, and SNR $\in [-6, 3]$ dB.

to N with a fixed value. From Figure 7, its threshold value will decrease and MSE performance will be improved as N increases.

4.6. Calculation Cost. Calculation cost is an important factor to impact the real time of frequency estimation algorithm. The calculation cost of these six algorithms is analyzed through calculating the operation number of their real multiplication and real addition under the same parameter setting as Figure 3.

The DFT value with K points can be derived by fast Fourier transform (FFT), that is, $K\log_2 K/2$ complex multiplication and $K\log_2 K$ complex addition. Besides, 1 complex multiplication needs 4 real multiplications and 2 real additions; 1 complex addition needs 2 real additions. In order to decrease the calculation cost, FDP algorithm also integrates coarse and fine search, where the point of coarse search is $2N$. Contemporarily, 1 modular operation needs 2 real multiplications and 1 real addition; 1 comparing operation needs 1 real addition. Consequently, the operation numbers of real multiplication M_{FDP} and real addition A_{FDP} of FDP algorithm are

$$
\begin{aligned}
M_{\text{FDP}} &= \frac{32N \log_2 (32N)}{2} \times 4 + 2N \times 2 + (2 \times 16 - 2) \\
&\times 2 = 64N \log_2 N + 324N + 60 = 20668, \\
A_{\text{FDP}} &= \frac{32N \log_2 (32N)}{2} \times 2 + 32N \log_2 (32N) \times 2 \\
&+ 2N + (2N - 1) + (2 \times 16 - 2) \\
&+ (2 \times 16 - 2) = 96N \log_2 N + 484N + 59 \\
&= 30907.
\end{aligned}
\tag{21}
$$

The coarse search of DS algorithm adopts FDP one with $K = 2N$ points, so it needs $4N\log_2 N + 8N$ real multiplication and $6N\log_2 N + 10N - 1$ real addition. Also, 1 iteration needs N complex multiplication, $N - 1$ complex addition, 4 real additions, and 1 modular operation. Accordingly, the operation numbers of real multiplication M_{DS} and real addition A_{DS} of DS algorithm are

$$
\begin{aligned}
M_{\text{DS}} &= 4N \log_2 N + 8N + (N \times 4 + 2) Q_{\text{DS}} \\
&= 4N \log_2 N + 88N + 40 = 3496, \\
A_{\text{DS}} &= 6N \log_2 N + 10N - 1 \\
&\quad + [N \times 2 + (N - 1) \times 2 + 4 + 1] Q_{\text{DS}} \\
&= 6N \log_2 N + 90N + 59 = 3899.
\end{aligned}
\tag{22}
$$

The coarse search of Modified DS algorithm adopts FDP one with $K = N$ points, so it needs $2N\log_2 N + 2N$ real multiplication and $3N\log_2 N + 2N - 1$ real addition. Also, 1 iteration needs $2N$ complex multiplication, $2N - 2$ complex addition, 6 real additions, and 2 modular operations. Accordingly, the operation numbers of real multiplication $M_{\text{M-DS}}$ and real addition $A_{\text{M-DS}}$ of Modified DS algorithm are

$$
\begin{aligned}
M_{\text{M-DS}} &= 2N \log_2 N + 2N + (2N \times 4 + 2 \times 2) Q_{\text{M-DS}} \\
&= 2N \log_2 N + 162N + 80 = 5584, \\
A_{\text{M-DS}} &= 3N \log_2 N + 2N - 1 \\
&\quad + [2N \times 2 + (2N - 2) \times 2 + 6 + 1] Q_{\text{M-DS}} \\
&= 3N \log_2 N + 162N + 59 = 5723.
\end{aligned}
\tag{23}
$$

The coarse search of IIFC algorithm adopts FDP one with $K = N$ points, so it needs $2N\log_2 N + 2N$ real multiplication and $3N\log_2 N + 2N - 1$ real addition. Also, 1 iteration needs $2N$ complex multiplication, $2N - 2$ complex addition, 6 real additions, and 2 modular operations. Accordingly, the operation numbers of real multiplication M_{IIFC} and real addition A_{IIFC} of IIFC algorithm are

$$
\begin{aligned}
M_{\text{IIFC}} &= 2N \log_2 N + 2N + [(2N + 1) \times 4 + 1] Q_{\text{IIFC}} \\
&= 2N \log_2 N + 82N + 50 = 2994, \\
A_{\text{IIFC}} &= 3N \log_2 N + 2N - 1 \\
&\quad + [(2N + 1) \times 2 + 2N \times 2 + 6] Q_{\text{IIFC}} \\
&= 3N \log_2 N + 82N + 79 = 3183.
\end{aligned}
\tag{24}
$$

The coarse search of RCTSL algorithm adopts FDP one with $K = 4N$ points, so it needs $8N\log_2 N + 24N$ real multiplication and $12N\log_2 N + 32N - 1$ real addition. Also, it needs 7 real multiplications and 4 real additions. Accordingly,

TABLE 1: Calculation cost of all these six algorithms.

	Real multiplication	$N = 32$	Real addition	$N = 32$
FDP	$64N\log_2 N + 324N + 60$	20668	$96N\log_2 N + 484N + 59$	30907
DS	$4N\log_2 N + 88N + 40$	3496	$6N\log_2 N + 90N + 59$	3899
Modified DS	$2N\log_2 N + 162N + 80$	5584	$3N\log_2 N + 162N + 59$	5723
IIFC	$2N\log_2 N + 82N + 50$	2994	$3N\log_2 N + 82N + 79$	3183
RCTSL	$8N\log_2 N + 24N + 7$	2055	$12N\log_2 N + 32N + 3$	2947
LS-SVR	$4N\log_2 N + 68N + 30$	2846	$6N\log_2 N + 70N - 1$	3199

the operation numbers of real multiplication M_{RCTSL} and real addition A_{RCTSL} of RCTSL algorithm are

$$M_{\mathrm{RCTSL}} = 8N\log_2 N + 24N + 7 = 2055,$$
$$A_{\mathrm{RCTSL}} = 12N\log_2 N + 32N - 1 + 4 \qquad (25)$$
$$= 12N\log_2 N + 32N + 3 = 2947.$$

The coarse search of LS-SVR algorithm adopts FDP one with $K = 2N$ points, so it needs $4N\log_2 N + 8N$ real multiplication and $6N\log_2 N + 10N - 1$ real addition. Also, if N, K, and L are fixed, matrices $\mathbf{Q}$ and $\mathbf{Q}'$ are already known, so the value of matrix $\boldsymbol{\beta}$ with $2L - 1$ rows and K columns can be calculated. As a result, it only has to add the calculation cost for (11) and comparison of $2L - 1$ real number. It is worth stressing that (11) is used for real and image part, respectively. Accordingly, the operation numbers of real multiplication $M_{\mathrm{LS\text{-}SVR}}$ and real addition $A_{\mathrm{LS\text{-}SVR}}$ of LS-SVR algorithm are

$$M_{\mathrm{LS\text{-}SVR}} = 4N\log_2 N + 8N + [2N \times (2 \times 8 - 1)] \times 2$$
$$+ (2 \times 8 - 1) \times 2$$
$$= 4N\log_2 N + 68N + 30 = 2846,$$
$$A_{\mathrm{LS\text{-}SVR}} = 6N\log_2 N + 10N - 1 \qquad (26)$$
$$+ [(2N - 1) \times (2 \times 8 - 1)] \times 2 + 2 \times 8 - 1$$
$$+ 2 \times 8 - 1 = 6N\log_2 N + 70N - 1$$
$$= 3199.$$

As in Figure 7, the operation numbers of real multiplication and real addition of these six algorithms are listed in Table 1. Additionally, $N = 32$ is taken, for example, to compare. From Table 1, the calculation cost of FDP one is much larger than the other five; second larger is Modified DS one and the next is DS one. The calculation cost of LS-SVR one is between IIFC and RCTSL one, and the distances between each of two are very small.

5. Conclusions

The MSE performance and the calculation cost in DFT-based frequency estimation algorithm are a pair relationship of interaction and interrestriction. The better its MSE performance is, the higher its calculation cost is. However, very high calculation cost will restrict the algorithm's real time capacity.

So, how to compromise between these two factors is a hot and key point.

This paper utilizes LS-SVR's good generalizing and fitting capabilities while keeping lower computational complexity and proposes a LS-SVR-based interpolation frequency estimation algorithm. Firstly, the sparser DFT points are learned and the continuous spectrum with less fitting errors is predicted. Then, because any output of LS-SVR is the linear weighted sum of all inputs, LS-SVR is used as a linear interpolator to resample the DFT points densely and decrease the calculation cost. At last, the resampling DFT points between the left and right of its maximum value are searched to derive the frequency estimation value.

In order to compare the calculation cost of the proposed and other DFT-based algorithms, the parameters of all these algorithms are set to ensure nearly the same precondition. Stressed that, the sample size, number of DFT points, and resampling points are already known. Our results show that the proposed algorithm can make a good compromise between MSE performance and calculation cost under the assumption mentioned above.

At the same time, the proper ranges of LS-SVR's parameters are obtained. As a next step, how to select them exactly is an important research point.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

[1] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Transactions on Information Theory*, vol. 20, no. 5, pp. 591–598, 1974.

[2] Y. V. Zakharov and T. C. Tozer, "Frequency estimator with dichotomous search of periodogram peak," *Electronics Letters*, vol. 35, no. 19, pp. 1608–1609, 1999.

[3] Y. V. Zakharov, V. M. Baronkin, and T. C. Tozer, "DFT-based frequency estimators with narrow acquisition range," *IEE Proceedings: Communications*, vol. 148, no. 1, pp. 1–7, 2001.

[4] E. Aboutanios, "A modified dichotomous search frequency estimator," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 186–188, 2004.

[5] H. Xu and D. M. Zhang, "A simple iterative carrier frequency estimation algorithm," in *Proceedings of the International Conference on Networks Security, Wireless Communications and*

Trusted Computing (NSWCTC '09), pp. 724–727, Wuhan, China, April 2009.

[6] G.-B. Zhang, "Novel algorithm for frequency estimation of sinusoid signal with random length," in *Proceedings of the International Conference on Electronic and Mechanical Engineering and Information Technology (EMEIT '11)*, pp. 523–526, IEEE, Harbin, China, August 2011.

[7] B. G. Quinn, "Estimating frequency by interpolation using Fourier coefficients," *IEEE Transactions on Signal Processing*, vol. 42, no. 5, pp. 1264–1268, 1994.

[8] B. G. Quinn, "Estimation of frequency, amplitude, and phase from the DFT of a time series," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 814–817, 1997.

[9] E. Aboutanios and B. Mulgrew, "Iterative frequency estimation by interpolation on Fourier coefficients," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1237–1242, 2005.

[10] E. Aboutanios, "Generalised DFT-based estimators of the frequency of a complex exponential in noise," in *Proceedings of the 3rd International Congress on Image and Signal Processing (CISP '10)*, pp. 2998–3002, Yantai, China, October 2010.

[11] C. Yang and G. Wei, "A noniterative frequency estimator with rational combination of three spectrum lines," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 5065–5070, 2011.

[12] S. R. Dooley and A. K. Nandi, "Fast frequency estimation and tracking using Lagrange interpolation," *Electronics Letters*, vol. 34, no. 20, pp. 1908–1910, 1998.

[13] I. Djurović and V. V. Lukin, "Estimation of single-tone signal frequency by using the L-DFT," *Signal Processing*, vol. 87, no. 6, pp. 1537–1544, 2007.

[14] G. Campobello, G. Cannatá, N. Donato, A. Famulari, and S. Serrano, "A novel low-complex and low-memory method for accurate single-tone frequency estimation," in *Proceedings of the 4th International Symposium on Communications, Control, and Signal Processing (ISCCSP '10)*, pp. 1–6, Limassol, Cyprus, March 2010.

[15] S. W. Chen, D. H. Li, and X. P. Wei, "Accurate frequency estimation of real sinusoid signal," in *Proceedings of the 2nd International Conference on Signal Processing Systems (ICSPS '10)*, pp. V3-370–V3-372, IEEE, Dalian, China, July 2010.

[16] Z. Ye, G. F. Xu, and D. X. Guo, "An accurate estimation algorithm of frequency and phase at low signal-noise ratio levels," in *Proceedings of the International Conference on Wireless Communications and Signal Processing (WCSP '10)*, pp. 1–5, Suzhou, China, October 2010.

[17] G. Wei, C. Yang, and F.-J. Chen, "Closed-form frequency estimator based on narrow-band approximation under noisy environment," *Signal Processing*, vol. 91, no. 4, pp. 841–851, 2011.

Advances in
Operations Research

Advances in
Decision Sciences

Journal of
Applied Mathematics

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Submit your manuscripts at
http://www.hindawi.com

Hindawi

Journal of
Complex Analysis

Journal of
Mathematics

Mathematical Problems
in Engineering

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

International Journal of
Stochastic Analysis

Journal of
Optimization