

Using Least Squares Support Vector Machines for Frequency Estimation

Xiaoyun Teng, Xiaoyi Zhang, Hongyi Yu

Information Science and Technology Institute, Zhengzhou, China

E-mail: ieukey@163.com

Received July 6, 2010; revised August 9, 2010; accepted September 11, 2010

Abstract

Frequency estimation is transformed to a pattern recognition problem, and a least squares support vector machine (LS-SVM) estimator is derived. The estimator can work efficiently without the need of statistics knowledge of the observations, and the estimation performance is insensitive to the carrier phase. Simulation results are presented showing that proposed estimators offer better performance than traditional Maximum Likelihood (ML) estimator at low SNR, since classification-based method does not have the threshold effect of nonlinear estimation.

Keywords: Carrier Recovery, LS-SVM, Pattern Recognition

1. Introduction

In digital communication systems with burst transmission, carrier recovery within each information burst is a critical issue. The estimation of carrier frequency in additive noise is one of the very important problems in the theory and applications of digital signal processing. Various techniques have been proposed for carrier frequency recovery [1-5]. The estimators in [1] are based on a maximum likelihood criterion, which is known to be an excellent estimator, but it suffers from the threshold effect of nonlinear estimators. Frequency estimation in colored noise is addressed in [6] and [7], which model the colored noise as an AR or MA process. However, most of above estimators require the statistics knowledge of the observations, such as, probability density function (pdf), autocorrelation, etc.

A Support Vector Machine (SVM) [8] uses training data as an integral element of the function estimation model as opposed to simply using training data to estimate parameters of an a priori model using maximum likelihood [9]. The SVM has the advantage over traditional learning approaches in terms of performance, complexity and convergence. SVMs have been widely used in solving classification and function estimation problems. Recently, SVM has been introduced to communication systems as a new method for channel equalization [9,10] and multiuser detector in CDMA communications [11]. The least squares support vector machines

(LS-SVM) involves solving linear equations instead of solving the quadratic programming problem required in the original SVM. In this paper, we view frequency estimation as a pattern recognition problem, and propose a different frequency estimator based on LS-SVM.

2. Signal Model

Using complex-envelope notation, the observed signal samples are expressed by

$$\begin{aligned} r_n &= a(n)e^{j(\omega n + \theta)} + v(n), \\ n &= 0, 1, \dots, N_p - 1, \dots, N - 1 \end{aligned} \quad (1)$$

where ω is the unknown carrier frequency normalized to sampling frequency f_s , for the sake of simplicity, $f_s = 1$, θ is an initial random phase, $v(n)$ are additive noise samples. a_n is the normalized transmitted BPSK symbol, *i.e.*, $a_n = \pm 1$. We consider scenarios where the signal a_n is known, *i.e.*, a training sequence is transmitted for carrier recovery.

Define the following vector

$$\mathbf{r} = [r_0, \dots, r_{N_p-1}, \dots, r_{N-1}]^T$$

$$\mathbf{A} = [a_0, \dots, a_{N_p-1}, \dots, a_{N-1}]^T$$

$$\Psi_N(\omega) = \text{diag} \{1, e^{j\omega}, \dots, e^{j(N_p-1)\omega}, \dots, e^{j(N-1)\omega}\}$$

$$\mathbf{v} = [v(0), v(1), \dots, v(N-1)]^T \tag{2}$$

The signal model can be arranged in matrix form as

$$\mathbf{r} = e^{j\theta} \Psi_N(\omega) \mathbf{A} + \mathbf{v} \tag{3}$$

3. SVM Based Frequency Estimation

3.1. SVM Introduction

SVM developed by Vapnik is a new type of learning machines which has a high generalization performance and sparse solution. It replaces empirical risk minimization by structural risk minimization (SRM). The goal of SVM is to find the hyperplane that maximizes the minimum distance between any point and the hyperplane. The idea of SVM can be expressed as follows.

Consider (x_i, y_i) , $i = 1, 2, \dots, N$ be a linearly separable training set, where $x \in R^d$ and $y \in \{-1, +1\}$, which can be separated by the hyperplane satisfying $w^T x + b = 0$, where w is the weight vector and b is the bias. If this hyperplane maximizes the margin, then we need to solve the following optimization problem.

$$\begin{aligned} &\text{minimize } \frac{1}{2} \|\mathbf{w}\|^2 \\ &\text{subject to } y_i (w \cdot x_i + b) \geq 1 \end{aligned} \tag{4}$$

For the inputs data that is not separable, SVM uses soft margins that can be expressed as follows, by introducing the non-negative slack variables $\xi_i, i = 1, \dots, N$:

$$\begin{aligned} &\text{minimize } \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l \xi_i^k \\ &\text{subject to } y_i (w^T x_i + b) \geq 1 - \xi_i \end{aligned} \tag{5}$$

Data points are penalized if they are misclassified. The parameter C controls tradeoff between the complexity of the model and the classification errors.

To construct nonlinear decision functions, SVM maps the training data nonlinearly into a higher-dimensional feature space via a kernel function, and constructs a separating hyperplane with maximum margin there. The kernel function is

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \tag{6}$$

The typical kernel functions include RBF $K(x, y) = \exp(-\|x - y\|^2 / 2\sigma^2)$ and polynomial $K(x, y) = (1 + x \cdot y)^d$.

We prefer LS-SVM over other models of SVM, for it offers a fast method for obtaining classifiers with good generalization performance in many applications. In LS-SVM, an equality constraint-based formulation is involved instead of the convex quadratic programming

(QP) problem in (5).

$$\begin{aligned} &\text{minimize } \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l \xi_i^2 \\ &\text{subject to } y_i (w^T x_i + b) = 1 - \xi_i \end{aligned} \tag{7}$$

To solve this problem Lagrange multipliers $(\alpha_i, i = 1, \dots, l; \alpha_i \geq 0)$ are used:

$$L_p = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l \xi_i^2 - \sum_{i=1}^N \alpha_i [y_i (w^T x_i + b) - 1 + \xi_i] \tag{8}$$

The KKT optimality conditions are given by

$$\begin{cases} \frac{\partial L_p}{\partial w} = 0 \rightarrow \left(w - \sum_{i=1}^l \alpha_i y_i \varphi(x_i) \right) = 0 \\ \frac{\partial L_p}{\partial b} = 0 \rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \\ \frac{\partial L_p}{\partial \xi_i} = 0 \rightarrow \alpha_i - C \xi_i = 0 \\ \frac{\partial L_p}{\partial \alpha_i} = 0 \rightarrow y_i (w^T \cdot \varphi(x_i)_i + b) - 1 + \xi_i = 0 \end{cases} \tag{9}$$

Optimal decision function (ODF) is then given by:

$$f(x) = \text{sign} \left[\sum_{i=1}^n \alpha_i y_i K(x_i \cdot x) + b \right] \tag{10}$$

3.2. SVM-FEA

To the best of our knowledge, SVM has not been implemented yet for parameter estimation in digital communication systems. We introduce SVM as a new method for frequency estimation, by building the following frequency estimator

$$\hat{\omega} = \arg \min_{\omega \in \Omega} \left\{ \sum_{n=1}^N \Lambda(n; \omega) \right\} \tag{11}$$

where the cost function

$$\Lambda(n; \omega) = |r(n) \cdot e^{-j(\omega n + \theta)} - s(n)| \tag{12}$$

In such a scenario, the frequency estimation problem can be transferred to a pattern recognition problem. The optimal estimate of ω can be attained by minimizing the classification error.

In a general way, the carrier phase θ existing in the received signal samples is unknown. Thus, the ideal ML detector is hard to handle the classification problem in (12). The powerful LS-SVM technique is applied in this paper.

To fit the support vector machine model, the output of the channel can be grouped into vectors

$$\mathbf{x}(n) = \left[\text{Re} \{ r(n) \cdot e^{-j\omega n} \}, \text{Im} \{ r(n) \cdot e^{-j\omega n} \} \right] \tag{14}$$

Where $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ mean the real and image part of $\{\cdot\}$, separately. For training purposes, taking $x(n)$ as the input sequence of SVM, and the transmitted symbol $a(n)$ to be the desired output sequence.

This model of SVM-based frequency estimator, that we call SVM-FEA, is illustrated in **Figure 1**.

The optimal estimator cannot be found in a single step because the input data has the unknown term $e^{-j\omega n}$. So an effective searching process is needed to achieve the frequency estimation. The procedure, which is similar to the coarse search and fine search routine in ML algorithm, is particularized as follows.

1) Choose a set of ω values according to a appropriate interval λ_1 , *i.e.*,

$$\omega_1 = 0, \omega_2 = \omega_1 + \lambda_1, \omega_3 = \omega_2 + \lambda_1, \dots$$

2) Construct the input sequence \mathbf{x}_w using each ω value; for example,

$$\mathbf{x}_{w_1}(n) = \left[\text{Re}\{r(n) \cdot e^{-j\omega_1 n}\}, \text{Im}\{r(n) \cdot e^{-j\omega_1 n}\} \right]$$

solve the QP problem and obtain the decision function.

3) Classify \mathbf{x}_w and identify the ω that minimizes the classification error, get a approximate estimate $\hat{\omega}$.

4) Set a refined interval λ_2 , get a new set of ω values between $\hat{\omega} \pm \lambda_1$, do 2) and 3) and get a fine estimate of ω .

4. Simulation Results

Computer simulations have been run to check the analytical results of the previous sections. We will observe the average estimate

$$E[\hat{f}] = \frac{1}{N_m} \sum_{i=1}^{N_m} \hat{f}_i$$

and the mean square error (MSE) of the estimate

$$MSE[\hat{f}] = \frac{1}{N_m} \sum_{i=1}^{N_m} (\hat{f}_i - f)^2$$

Linear LS-SVMs are used with 15 to 50 data samples. In such a scenario, C is the only parameter to be chosen by the user during LS-SVM training, C is the upper

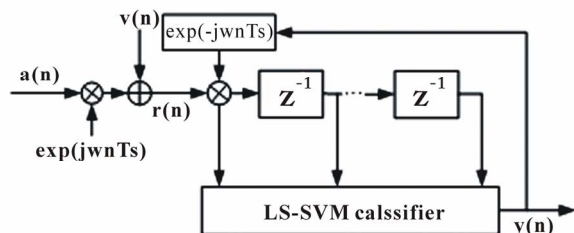


Figure 1. SVM based frequency estimator structure.

responds to assigning a higher penalty to classification errors. Simulation results showed that it has been more robust to set C between 0.1 and 10. Specifically, we fix $C = 5$.

Figure 2 illustrates the average estimates versus f when $\text{SNR} = 7$ dB, $N = 20$. The ideal line $E[\hat{f}] = f$ is also shown for comparison. The curves show that the range over which the estimates are unbiased is about $(-0.4, 0.4)$.

We compare the performance of the proposed SVM-FEA and the typical ML algorithm in AWGN channel. MSE of the estimates are compared with the CRLB as follows

$$CRLB_{\hat{\omega}} = \begin{cases} \frac{3}{SNR \cdot N(N-1)(2N-1)}, & \text{phase is known} \\ \frac{6}{SNR \cdot N(N^2-1)}, & \text{phase is unknown} \end{cases}$$

We first consider the case that the carrier phase is known, so ideal ML detector can be directly used to handle the classification problem in (13). Thus a ML detector based frequency estimator can be derived by (12). The curves of MSE versus SNR are shown in **Figure 3**. The simulation results are attained when $f = 0.34$, $N = 18$. The performance curves of the ML single tone, *i.e.* DA, frequency estimation algorithms with different N proposed in [1] are also shown for comparison. The likelihood functions are computed through FFT. For the sake of simplicity, only the coarse search is made and the FFT length is 32768. It can be seen in **Figure 2** that the MSE performance of the SVM-FEA is close to ML detector based frequency estimator in this case. And both classification based frequency estimation algorithm outperform the traditional ML estimator at low SNR (< -1 dB).

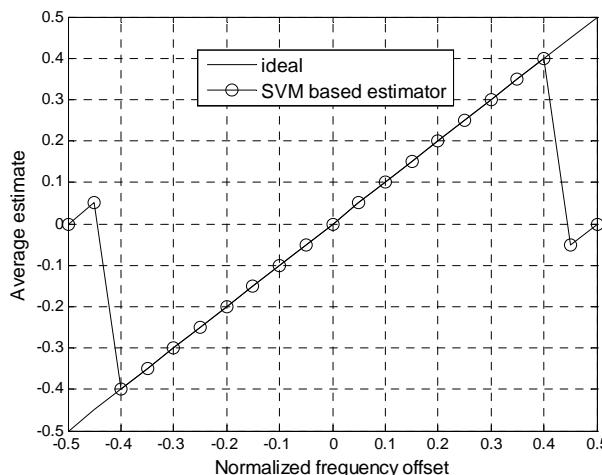


Figure 2. Average frequency estimate versus f .

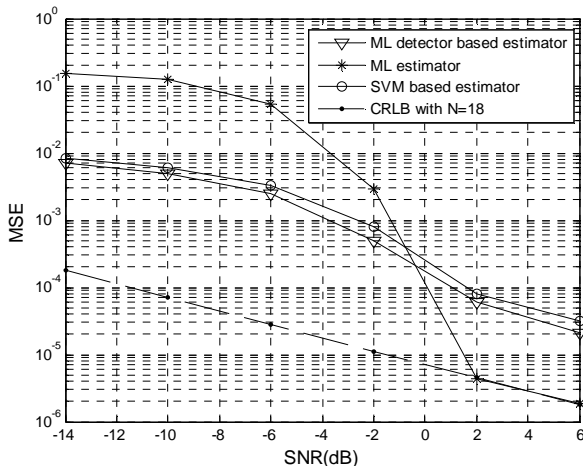


Figure 3. MSE versus SNR when phase is known.

Figure 4 illustrates the comparison of the MSE performance of the SVM-FEA and ML estimator when the phase is unknown. The simulation results are attained when $f = 0.34$, $N = 20$ and $N = 30$. Since all the data points in the input vectors have the same constant caused by the carrier phase, SVM-FEA shows almost the same performance in both cases. The performance of SVM-FEA improves with the increasing of the data length, which is not as remarkable as that of ML estimator (the cross of two curves change from 0 dB to -4 dB).

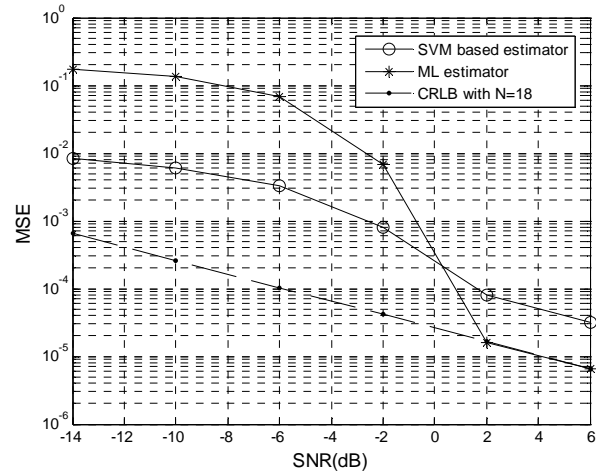
It is noticed that although SVM-FEA present a significant improvement over the ML estimator at low SNR, it can not reach the CRLB even at high SNR. The reason is that classification based frequency estimation algorithm identifies the estimate of ω corresponding to the classification error, which is insensitive to a very small frequency change to a certain extent.

5. Conclusions

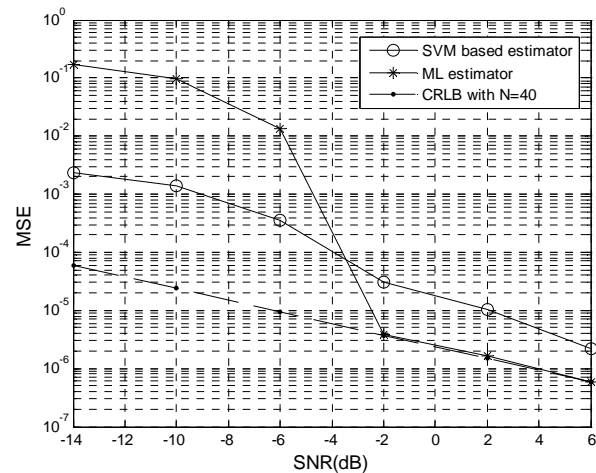
In this paper, we formulated frequency estimation of digital communication signals as a classification problem, and applied SVM technique to solve it. A primary searching routine has been proposed to find the optimal frequency estimate.

Simulations have shown that the SVM provides a robust method for frequency estimation with following attractive features: The estimator can work efficiently without the need of statistics knowledge of the observations, and the estimation performance is insensitive to the carrier phase; it shows a better performance than traditional ML estimator at low SNR, for SVM-FEA has not the threshold effect of nonlinear estimation.

A main drawback of the proposed algorithm is the



(a)



(b)

Figure 4. MSE versus SNR when phase is unknown.

high computational cost, which can be reduced by introducing faster optimization techniques and improving our searching routine. Future work will also be carried out on the generalization of the proposed procedure to multi-level modulations and other channel conditions, such as fading channel and colored noise conditions.

6. 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, 1974, pp. 591-598.
- [2] M. Morelli and U. Mengali, "Feedforward Frequency Estimation for PSK: A Tutorial Review," *European Tran-*

- sactions on Telecommunications*, Vol. 9, No. 2, 1998, pp. 103-116.
- [3] W. Y. Kuo and M. P. Fitz, "Frequency Offset Compensation of Pilot Symbol Assisted Modulation in Frequency Flat Fading," *IEEE Transactions on Communications*, Vol. 45, No. 11, 1997, pp. 1412-1416.
- [4] U. Mengali and M. Morelli, "Data-Aided Frequency Estimation for Burst Digital Transmission," *IEEE Transactions on Communications*, Vol. 45, No. 1, 1997, pp. 23-25.
- [5] Y. Wang, E. Serpedin and P. Ciblat, "Optimal Blind Carrier Recovery for MPSK Burst Transmissions," *IEEE Transactions on Communications*, Vol. 51, No. 9, 2003, pp. 1571-1581.
- [6] D. N. Swingler, "Approximate Bounds on Frequency Estimates for Short Cissoids in Colored Noise," *IEEE Transactions on Signal Processing*, Vol. 46, No. 5, 1998, pp. 1456-1458.
- [7] J. Viterbi and A. M. Viterbi, "Nonlinear Estimation of Psk-Modulated Carrier Phase with Application to Burst Digital Transmissions," *IEEE Transactions on Information Theory*, Vol. 29, No. 4, 1983, pp. 543-551.
- [8] V. N. Vapnik, "Statistical learning theory," John Wiley & Sons, Inc., New York, 1998.
- [9] D. J. Sebald, "Support Vector Machine Techniques for Nonlinear Equalization," *IEEE Transactions on Signal Processing*, Vol. 48, No. 11, 2000, pp. 3217-3226.
- [10] S. Chen, S. Gunn and C. Harris, "Decision Feedback Equalizer Design Using Support Vector Machines," *Vision, Image and Signal Processing*, Vol. 147, No. 3, 2000, pp. 213-219.
- [11] F. Perez-Cruz, A. Navia-Vazquez, P. Alarcon-Dianna and A. Artes-Rodriguez, "SVC-Based Equalizer for Burst TDMA Transmission," *Signal Processing*, Vol. 81, No. 6, 2001, pp. 1681-1693.
- [12] J. G. Proakis, "Digital Communications," 4th Edition, McGraw-Hill, New York, 2001.