Journal of Systems Engineering and Electronics
Vol. 28, No. 4, August 2017, pp.627 – 633

EXCLUSIVELY AVAILABLE VIA
IEEE Xplore®
DIGITAL LIBRARY

# Modulation recognition of communication signals based on SCHKS-SSVM

Xiaolin Zhang[1,2,*], Jian Chen[1], and Zhiguo Sun[1]

1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China;
2. Science and Technology on Information Transmission and Dissemination in Communication Networks Laboratory, Shijiazhuang 050000, China

**Abstract:** A novel modulation recognition algorithm is proposed by introducing a Chen-Harker-Kanzow-Smale (CHKS) smooth function into the C-support vector machine deformation algorithm. A set of seven characteristic parameters is selected from a range of parameters of communication signals including instantaneous amplitude, phase, and frequency. And the Newton-Armijo algorithm is utilized to train the proposed algorithm, namely, smooth CHKS smooth support vector machine (SCHKS-SSVM). Compared with the existing algorithms, the proposed algorithm not only solves the non-differentiable problem of the second order objective function, but also reduces the recognition error. It significantly improves the training speed and also saves a large amount of storage space through large-scale sorting problems. The simulation results show that the recognition rate of the algorithm can batch training. Therefore, the proposed algorithm is suitable for solving the problem of high dimension and its recognition can exceed 95% when the signal-to-noise ratio is no less than 10 dB.

**Keywords:** communication signal, modulation recognition, support vector machine, smooth function.

**DOI:** 10.21629/JSEE.2017.04.01

## 1. Introduction

Automatic recognition of the communication signal is of great significance in military and civil sectors [1,2]. It can be utilized in the communication countermeasures and the authentication and spectrum management for radio signals. The existing modulation recognition algorithms can be roughly divided into two categories: statistical pattern based recognition and decision theory based method [3,4]. Due to the simple structure, superior performance and huge potential applicability, the support vector machine (SVM) recognition algorithm has received widespread attention in recent years [5 – 9]. Lee et al. introduced the SVM method into modulation recognition in [10] for the first time. In [11], Valipour et al. combined the SVM with the particle swarm optimization algorithm to identify a variety of modulation signals. The higher-order cumulates were used together with the SVM for communication signals recognition, and a high recognition rate was reported in [12 – 14]. However, in the traditional SVM, the objective function of the quadratic programming problem does not have second-order smoothing property, especially in low sampling rate and small SNR scenarios, which leads to a very high recognition error. Based on the existing research results, this paper proposes a novel modulation recognition method for communication signals by introducing the Chen-Harker-Kanzow-Smale (CHKS) smooth function [15] into the C-support vector machine deformation algorithm [1]. The CHKS function will be used to approximate the non-differentiable term of the unconstrained SVM, which transforms the original non-differentiable model into a differentiable one. Moreover, the proposed algorithm will also jointly consider the statistical characteristics of signals to be recognized. Recognition simulations are carried out for seven types of digital communication signals by using a fast algorithm. When the signal-to-noise ratio is higher than 10 dB, its recognition rate can exceed 95%, which verifies the effectiveness of this algorithm. Meanwhile, the proposed method reduces the computational complexity and saves the storage space during the learning process. Last but not least, it significantly improves both the speed and quality of modulation recognition.

## 2. Principle of SVM-based modulation recognition

SVM originates from the statistical learning theory, and

Manuscript received January 08, 2016.
*Corresponding author.
This work was supported by the National Natural Science Foundation of China (61401196), the Jiangsu Provincial Natural Science Foundation of China (BK20140954), the Science and Technology on Information Transmission and Dissemination in Communication Networks Laboratory (KX152600015/ITD-U15006), and the Beijing Shengfeifan Electronic System Technology Development Co., Ltd (KY10800150036).

its basic principle is to perform a certain type of non-linear transform on the original feature space by defining a proper kernel function, which maps the original feature space into a higher-dimension one. Then, an optimal separating hyperplane is found in the new space, which correctly separates a set of patterns and guarantees the maximal separation gap. The SVM needs to be trained before performing any classification and recognition using the trained classifier. The SVM-based recognition process is shown in Fig. 1.

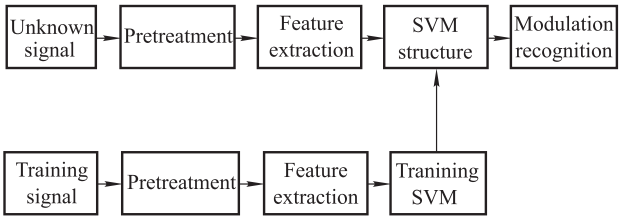


**Fig. 1 Block diagram of SVM-based modulation recognition**

The above mentioned SVM-based recognition is essentially a binary classifier. For linearly separable binary patterns, they can be correctly separated by constructing an optimal separating line. While, for the non-linearly separable counterpart, a certain type of non-linear mapping is needed to map the patterns into a linearly separable higher-dimension space before performing the separation.

Assume that there exists a set of elements $(x_1, y_1)$, $(x_2, y_2), \dots, (x_l, y_l), \boldsymbol{x} \in \mathbf{R}^n, y_i \in \{1, -1\}$, where $l$ stands for the number of elements, $n$ is the number of input dimensions. The goal of classification is to find a hyperplane which completely separates the elements set into two groups. The hyperplane $H$ can be expressed as follows:

$$\boldsymbol{\omega} \cdot \boldsymbol{x} + \boldsymbol{b} = 0. \tag{1}$$

A point $\boldsymbol{x}_i$ in the hyperplane satisfies that $\boldsymbol{\omega} \cdot \boldsymbol{x} + \boldsymbol{b} = 0$, where $\boldsymbol{\omega}$ is the normal vector of the hyperplane and also the weight vector of the corresponding classifier; $\boldsymbol{b}$ is the threshold for classification. The results of classification can be expressed as

$$\begin{cases} \boldsymbol{\omega} \cdot \boldsymbol{x}_i + \boldsymbol{b} \geqslant 0, & y_i = 1 \\ \boldsymbol{\omega} \cdot \boldsymbol{x}_i + \boldsymbol{b} < 0, & y_i = -1 \end{cases}. \tag{2}$$

In order to find the optimal separating hyperplane, solving the optimal classification problem eventually converts to a quadratic programming problem, which means to minimize the following functional under the constraint of (1):

$$\phi(\boldsymbol{\omega}) = \frac{\|\boldsymbol{\omega}\|^2}{2}. \tag{3}$$

However, for the traditional SVM, the objective function of the quadratic programming problem is not second-order smooth, which leads to relatively large errors for recognition of low sampling-rate and low signal-to-noise-ratio (SNR) communication signals. To address this problem, this paper proposes a novel modulation recognition method for communication signals by introducing the CHKS smooth function [15] into the C-support vector machine deformation algorithm [1]. The CHKS function is used to approximate the non-differentiable term of the unconstrained SVM, which transforms the original non-differentiable model into a differentiable one.

# 3. Design of SCHKS-SSVM-based classifier

## 3.1 Principle of SCHKS-SSVM

When both the linear separability and linear unseparability are taken into a unified consideration, the original problem of the deformation of the C-support vector classification machine becomes:

$$\min_{\boldsymbol{\omega},\boldsymbol{b},\boldsymbol{\xi}} \ \frac{1}{2}\|\boldsymbol{\omega}\|^2 + \frac{C}{2}\sum_{i=1}^{l}\boldsymbol{\xi}_i^2$$

$$\text{s.t. } y_i((\boldsymbol{\omega} \cdot x_i) + \boldsymbol{b}) \geqslant 1 - \boldsymbol{\xi}_i, \quad \boldsymbol{\xi}_i \geqslant 0, i = 1, 2, \dots, l \tag{4}$$

where $\boldsymbol{\omega}$ is the weight vector controlling the size of the classification interval. $\boldsymbol{b}$ is the threshold for classification. The larger the classification interval, the smaller the structural risk; $C$ is the error penalty parameter representing the punishment degree of a wrong classification. A bigger error penalty parameter $C$ means a severer punishment for the wrong classification; $\boldsymbol{\xi}_i$ is the relaxing factor maintaining a certain sort of balance between the empirical risk and the generalization performance, which allows the existence of a small number of the wrong classification.

It can be seen from (4) that the objective function considers both the empirical risk and the confidence range, which meets the structural risk minimization criteria. When $\boldsymbol{\xi}_i = 0$, it becomes a linearly separable SVM.

The optimization problem in (4) is not a strictly convex quadratic programming. For easy solution, (4) can be revised to a strictly convex one as shown below:

$$\min_{\boldsymbol{\omega},\boldsymbol{b},\boldsymbol{\xi}} \frac{1}{2}(\|\boldsymbol{\omega}\|^2 + \boldsymbol{b}^2) + \frac{C}{2}\sum_{i=1}^{l}\boldsymbol{\xi}_i^2$$

$$\text{s.t. } y_i((\boldsymbol{\omega} \cdot \boldsymbol{x}_i) + \boldsymbol{b}) \geqslant 1 - \boldsymbol{\xi}_i, \quad \boldsymbol{\xi}_i \geqslant 0, \quad i = 1, 2, \dots, l. \tag{5}$$

The relaxing factor $\xi_i$ in (5) should satisfy

$$\boldsymbol{\xi}_i = \phi(\boldsymbol{\omega}, \boldsymbol{b}) = (1 - y_i((\boldsymbol{\omega} \cdot \boldsymbol{x}_i) - \boldsymbol{b}))_+,$$

$$i = 1, 2, \dots, l. \tag{6}$$

$(\boldsymbol{v}_i)_+$ is a single variable function:

$$(\boldsymbol{v}_i)_+ = \begin{cases} \boldsymbol{v}_i, & \boldsymbol{v}_i \geqslant 0 \\ 0, & \boldsymbol{v}_i < 0 \end{cases}. \tag{7}$$

Substituting the relaxing factor $\xi_i$ in (5) with (6) yields an unconstrained optimization problem:

$$\min_{\boldsymbol{\omega},\boldsymbol{b},\boldsymbol{\xi}} \frac{1}{2}(\|\boldsymbol{\omega}\|^2+\boldsymbol{b}^2)+\frac{C}{2}\sum_{i=1}^{l}(1-y_i((\boldsymbol{\omega}\cdot\boldsymbol{x}_i)-\boldsymbol{b}))_+^2. \quad (8)$$

The above optimization problem is a strictly convex unconstrained problem which has a unique optimal solution. However, the function may be non-differentiable and nonsmooth, which is due to the fact that the function $(\boldsymbol{v}_i)_+$ is continuous at the point $\boldsymbol{v}_i=0$, but $\lim\limits_{u\to 0^-}(\boldsymbol{v}_+)^{\mathrm{T}}=0$, $\lim\limits_{u\to 0^+}(\boldsymbol{v}_+)^{\mathrm{T}}=\boldsymbol{v}_+$, which means the objective function at $u_i=0$ being non-differentiable. Therefore, the traditional optimization problem solving methods cannot be applied in this scenario, as they always require existence of the gradient of the objective function and Hesse matrix, etc.

To address the abovementioned problem, a novel approach called SCHKS-SSVM is proposed in this paper, where a CHKS smooth function expressed in (9) is introduced into the C-SVM deformation algorithm to bring smoothness and solve the second order non-differentiable problem in (8).

$$\phi(\boldsymbol{v};\boldsymbol{\varepsilon})=\frac{1}{2}\boldsymbol{v}+\frac{1}{2}\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}. \quad (9)$$

**Lemma 1** The following conclusions can be given for the CHKS smooth function (9):

(i) For any $\boldsymbol{v}\in\mathbf{R}^n$, $\phi(\boldsymbol{v},\boldsymbol{\varepsilon})$ is arbitrary-order differentiable with respect to $v$.

(ii) For any $\boldsymbol{v}\in\mathbf{R}^n$, $0<\boldsymbol{\varepsilon}^1<\boldsymbol{\varepsilon}^2$, we have $\phi(\boldsymbol{v};\boldsymbol{\varepsilon}^1)<\phi(\boldsymbol{v};\boldsymbol{\varepsilon}^2)$.

**Proof** For any $v\in\mathbf{R}^n$, $\boldsymbol{\varepsilon}>0$, the following equations can be easily obtained through simple derivation:

$$\frac{\partial(\phi(\boldsymbol{v};\boldsymbol{\varepsilon}))}{\partial(\boldsymbol{\varepsilon})}=\frac{2\boldsymbol{\varepsilon}}{\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}}>0. \quad (10)$$

$$\frac{\partial(\phi(\boldsymbol{v};\boldsymbol{\varepsilon}))}{\partial(\boldsymbol{v})}=\frac{1}{2}+\frac{\boldsymbol{v}}{2\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}}=\frac{\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}+\boldsymbol{v}}{2\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}}>0. \quad (11)$$

Therefore, the function $\phi(\boldsymbol{v};\boldsymbol{\varepsilon})$ is a monotone increasing function with respect to $\boldsymbol{\varepsilon}$. For any $\boldsymbol{\varepsilon}^1,\boldsymbol{\varepsilon}^2$, we get $\phi(\boldsymbol{v};\boldsymbol{\varepsilon}^1)<\phi(\boldsymbol{v};\boldsymbol{\varepsilon}^2)$ whenever $0<\boldsymbol{\varepsilon}^1<\boldsymbol{\varepsilon}^2$. □

(iii) For any $\boldsymbol{v}\in\mathbf{R}^n$, $\boldsymbol{\varepsilon}>0$, we have $\phi(\boldsymbol{v})\leqslant\phi(\boldsymbol{v};\boldsymbol{\varepsilon})\leqslant\phi(\boldsymbol{v})+\boldsymbol{\varepsilon}$.

**Proof** This lemma can be easily proved through the following:

$$\phi(\boldsymbol{v};\boldsymbol{\varepsilon})-\max(0,\boldsymbol{v})=$$
$$(\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2})/2-(\boldsymbol{v}+\sqrt{\boldsymbol{v}^2})/2=$$
$$(\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}-\sqrt{\boldsymbol{v}^2})/2. \quad (12)$$

So, for any $\boldsymbol{v}\in\mathbf{R}^n$, $\boldsymbol{\varepsilon}>0$, we obtain $0\leqslant\phi(\boldsymbol{v};\boldsymbol{\varepsilon})-\phi(\boldsymbol{v})\leqslant\boldsymbol{\varepsilon}$. □

(iv) For any $\boldsymbol{\varepsilon}>0$, $\phi(\boldsymbol{v};\boldsymbol{\varepsilon})$ is continuously differentiable and strictly convex.

**Proof** The following equation can be obtained from (8) and (9).

$$\nabla_\omega\varPhi(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})=$$
$$\boldsymbol{\omega}-\frac{C}{4}\sum_{i=1}^{l}y_i\boldsymbol{x}_i\left(2\boldsymbol{v}+\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}+\frac{\boldsymbol{v}^2}{\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}}\right). \quad (13)$$

$$\nabla_b\varPhi(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})=$$
$$b+\frac{C}{4}\sum_{i=1}^{l}y_i\left(2\boldsymbol{v}+\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}+\frac{\boldsymbol{v}^2}{\sqrt{\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2}}\right). \quad (14)$$

Equations (13) and (14) can be written as

$$\nabla^2\varPhi(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})=$$
$$\begin{bmatrix} I_l+\dfrac{C}{4}\displaystyle\sum_{i=1}^{l}\boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}}\left[2+\dfrac{3\boldsymbol{v}}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)}}-\dfrac{\boldsymbol{v}^3}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)^3}}\right] & -\dfrac{C}{4}\displaystyle\sum_{i=1}^{l}\boldsymbol{x}_i^{\mathrm{T}}\left[2+\dfrac{3\boldsymbol{v}}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)}}-\dfrac{\boldsymbol{v}^3}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)^3}}\right] \\ -\dfrac{C}{4}\displaystyle\sum_{i=1}^{l}\boldsymbol{x}_i\left[2+\dfrac{3\boldsymbol{v}}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)}}-\dfrac{\boldsymbol{v}^3}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)^3}}\right] & 1+\dfrac{C}{4}\displaystyle\sum_{i=1}^{l}\left[2+\dfrac{3\boldsymbol{v}}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)}}-\dfrac{\boldsymbol{v}^3}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)^3}}\right] \end{bmatrix} \quad (15)$$

where $\boldsymbol{v}=1-y_i(\boldsymbol{\omega}\cdot\boldsymbol{x}_i-\boldsymbol{b})$. Assume $\lambda_i(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})=\left(2+\dfrac{3\boldsymbol{v}}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)}}-\dfrac{\boldsymbol{v}^3}{\sqrt{(\boldsymbol{v}^2+4\boldsymbol{\varepsilon}^2)^3}}\right)>0$. For any $\overline{\boldsymbol{\xi}}^{\mathrm{T}}=(\boldsymbol{\xi}^{\mathrm{T}},\boldsymbol{\xi}_0)\in\mathbf{R}^{n+1}$, $\overline{\boldsymbol{\xi}}\neq 0$, $\boldsymbol{\xi}\in\mathbf{R}^n$, the following equation can be obtained due to the fact that $\lambda_i(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})>0$.

$$\overline{\boldsymbol{\xi}}^{\mathrm{T}}\nabla^2\varPhi(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})\overline{\boldsymbol{\xi}}=\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi}+\sum_{i=1}^{l}\lambda_i(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})(x_i^{\mathrm{T}}\boldsymbol{\xi})^{\mathrm{T}}(x_i^{\mathrm{T}}\boldsymbol{\xi})+$$
$$2\sum_{i=1}^{l}\lambda_i(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})\boldsymbol{\xi}_0x_i^{\mathrm{T}}\boldsymbol{\xi}+\sum_{i=1}^{l}\lambda_i(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})\boldsymbol{\xi}_0^2=$$

$$\|\boldsymbol{\xi}\|^2+\sum_{i=1}^{l}\lambda_i(\boldsymbol{\omega},\boldsymbol{b};\boldsymbol{\varepsilon})(\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{\xi}+\boldsymbol{\xi}_0)^2>0. \tag{16}$$

Therefore, the lemma is proved. □

Substituting the function $(v)_+$ with (9) to smooth it, a smooth unconstrained problem similar to the non-smooth unconstrained one (8) can be established to approximate $(v_i)_+$. That is the principle of the proposed SCHKS-SSVM in this paper, which can be expressed as follows:

$$\min_{\boldsymbol{\omega},\boldsymbol{b},\boldsymbol{\xi}}\frac{1}{2}(\|\boldsymbol{\omega}\|^2+\boldsymbol{b}^2)+\frac{C}{2}\sum_{i=1}^{l}\left(\frac{1}{2}\boldsymbol{v}+\frac{1}{2}\sqrt{\boldsymbol{v}^2+4\varepsilon^2}\right)^2. \tag{17}$$

The proofs mentioned above demonstrate that the SCHKS-SSVM is a strictly convex unconstrained optimization problem having a unique optimal solution with arbitrary-order differentiability and smoothness. Therefore, it shows plenty of advantages on solving the classification problem for signals with a low sampling rate and SNR.

The proposed SCHKS-SSVM can not only solve linear classification problems, but also be applied to nonlinear problems when the previous conclusions are extended to nonlinear SCHKS-SSVM. The mapping from the input space $\boldsymbol{X}$ to the Hilbert space $\boldsymbol{H}$:

$$\varPhi:\left|\begin{array}{l}\boldsymbol{X}\to\boldsymbol{H}\\ \boldsymbol{x}\to\boldsymbol{x}=\phi(\boldsymbol{x})\end{array}\right.. \tag{18}$$

The kernel function:

$$K(\boldsymbol{x}\cdot\boldsymbol{x}')=(\phi(\boldsymbol{x})\cdot\phi(\boldsymbol{x}')). \tag{19}$$

Considering the dual problem of the following formula:

$$\min_{\boldsymbol{\omega},\boldsymbol{b},\boldsymbol{\xi}}\frac{1}{2}(\|\boldsymbol{\omega}\|^2+\boldsymbol{b}^2)+\frac{C}{2}\sum_{i=1}^{l}\boldsymbol{\xi}_i^2$$

$$\text{s.t. } y_i((\boldsymbol{\omega}\cdot x_i)+b)\geqslant 1-\xi_i,\quad \xi_i\geqslant 0,\quad i=1,2,\dots,l. \tag{20}$$

Assume that $\alpha=(a_1,a_2,\dots,a_l)$ is the solution of the dual problem, the following equation can be obtained

$$\boldsymbol{\omega}=\sum_{i=1}^{l}y_ia_i\boldsymbol{x}_i=\sum_{i=1}^{l}y_ia_i\phi(\boldsymbol{x}_i). \tag{21}$$

It can be used to transform the problem in (20), where $\|\boldsymbol{\omega}\|^2$ can be approximated by $\sum_{i=1}^{l}a_i^2$. Hence, (20) can be approximately expressed as

$$\min_{\boldsymbol{\omega},\boldsymbol{b},\boldsymbol{\xi}}\frac{1}{2}\left(\sum_{i=1}^{l}a_i^2+\boldsymbol{b}^2\right)+\frac{C}{2}\sum_{i=1}^{l}\boldsymbol{\xi}_i^2$$

$$\text{s.t. } y_i\left(\sum_{i=1}^{l}y_jaK(\boldsymbol{x}_i\cdot\boldsymbol{x}_j)+\boldsymbol{b}\right)\geqslant 1-\xi_i,$$

$$\xi_i\geqslant 0,\quad i=1,2,\dots,l. \tag{22}$$

Similar as linear problems, a smooth approximation function expressed (9) can be introduced for (22)

It eventually converts (22) into a smooth unconstrained optimization problem shown below:

$$\min_{\boldsymbol{\omega},\boldsymbol{b},\boldsymbol{\xi}}\frac{1}{2}(\sum_{i=1}^{l}a_i^2+\boldsymbol{b}^2)+$$

$$\frac{C}{2}\sum_{i=1}^{l}\left(\phi\left(1-y_i\left(\sum_{i=1}^{l}y_jaK(\boldsymbol{x}_i\cdot\boldsymbol{x}_j)+\boldsymbol{b}\right);\boldsymbol{\varepsilon}\right)\right)^2. \tag{23}$$

When $\boldsymbol{\varepsilon}$ is sufficiently large, (23) is similar to (22). After getting the optimal solution $(\boldsymbol{a}^*,\boldsymbol{b}^*)$, the decision function can be constructed as follows:

$$f(x)=\operatorname{sgn}\left(\sum_{i=1}^{l}y_ia_i^*K(\boldsymbol{x}_i\cdot\boldsymbol{x}_j)+\boldsymbol{b}^*\right). \tag{24}$$

The above objective function has continuous gradient and Hesse matrix, and is unconstrained. Therefore, the Newton-Armijo algorithm can be applied to solve the programming problem.

### 3.2 Newton-Armijo algorithm for SCHKS-SSVM

The objective function of the SCHKS-SSVM model is arbitrary-order differentiable. The Newton-Armijo algorithm can be utilized for the problem solving. Newton direction is used in the descent direction. The step length is determined by the Armijo criterion. The steps of the algorithm can be detailed as follows:

**Step 1** Set the initial points $(\boldsymbol{\omega}^0,\boldsymbol{b}^0)\in\mathbf{R}^{n+1}$, $\eta$ and $i=0$.

**Step 2** Calculate $\varPhi^i=\varPhi(\boldsymbol{\omega}^i,\boldsymbol{b}^i;\boldsymbol{\varepsilon})$ and $\boldsymbol{g}^i=\nabla\varPhi(\boldsymbol{\omega}^i,\boldsymbol{b}^i;\boldsymbol{\varepsilon})$.

**Step 3** If $\|\boldsymbol{g}^i\|\leqslant\eta$ or $\lambda_i<10^{-12}$, then $(\boldsymbol{\omega}^*,\boldsymbol{b}^*)=(\boldsymbol{\omega}^i,\boldsymbol{b}^i)$ is chosen and the calculation is terminated; otherwise, the descent direction $d^i$ is calculated by using the equation $\nabla^2\varPhi(\boldsymbol{\omega}^i,\boldsymbol{b}^i;\boldsymbol{\varepsilon})\boldsymbol{d}^i=-\boldsymbol{g}^i$.

**Step 4** Armijo step $\delta\in\left(0,\frac{1}{2}\right)$, $\lambda_i=\max\left\{1,\frac{1}{2},\frac{1}{4},\dots\right\}$ is chosen so that $\varPhi(\boldsymbol{\omega}^i,\boldsymbol{b}^i;\boldsymbol{\varepsilon})-\varPhi((\boldsymbol{\omega}^i,\boldsymbol{b}^i)+\lambda_i\boldsymbol{d}^i;\boldsymbol{\varepsilon})\geqslant-\delta\lambda_i\boldsymbol{g}^i\boldsymbol{d}^i$, where $(\boldsymbol{\omega}^{i+1},\boldsymbol{b}^{i+1})=(\boldsymbol{\omega}^i,\boldsymbol{b}^i)+\lambda_i\boldsymbol{d}^i$.

**Step 5** Increase $i$ by 1, i.e., $i=i+1$, and go to Step 2.

Unlike the traditional SVM, this algorithm does not need to solve the quadratic programming problem. Only the linear equation $\nabla^2 \Phi(\boldsymbol{\omega}^i, \boldsymbol{b}^i; \varepsilon)\boldsymbol{d}^i = -\boldsymbol{g}^i$ needs to be solved. Due to the strict convexity of the objective function, the Newton-Armijo algorithm which trains the proposed SCHKS-SSVM approach is globally convergent, and has a unique minimum.

### 3.3 Modulation recognition step

The communication signal modulation identification steps based on SCHKS-SSVM are as follows.

**Step 1** Extract the time-frequency characteristics of the signals to be detected and construct the training set.

**Step 2** Train the SCHKE-SVM classifier using the extracted eigenvectors.

**Step 3** Extract the features of the received unrecognized signal as the input of the trained SCHKS-SSVM classifier.

**Step 4** Output types of modulation identification.

## 4. Experimental results

In order to assess the performance of the proposed SCHKS-SSVM based modulation recognition algorithm. A Matlab-based simulation platform is built, on which SVC- and SSVM-based recognition algorithms are also simulated. Three types of classifiers including SVM, SSVM and SCHKS-SSVM modulation classifiers are implemented in the simulation platform. They are used to carry out recognition experiments for seven types of commonly used digital modulation signals including 2 amplitude shift keying (ASK), 4ASK, 2 frequency shift keying (FSK), 4 FSK, 2 phase-shift keying (PSK), 4 PSK and 16 quadrature amplitude modulation (QAM), respectively. Digital signal source is generated randomly, then modulated into narrow-band signals corrupted by white Gaussian noise. Its carrier frequency is $f_c = 20$ kHz, with symbol rate being $f_d = 1.8$ kHz and sampling rate being $f_s = 60$ kHz.

Six key features are extracted from instantaneous amplitude, phase and frequency, etc., as the characteristics of modulation recognition [16,17], including:

**Feature 1** Power spectral density of central normalized instantaneous amplitude to distinguish multiple ASK (MASK), multiple PSK (MPSK), 16 QAM signals containing amplitude information from MFSK signals not containing any amplitude information.

**Feature 2** Standard deviation for the absolute value of the central normalized instantaneous amplitude (AVCNIA) to distinguish the 2ASK signal with its AVCNIA being a constant from 4ASK, 16QAM with their AVCNIA not being a constant.

**Feature 3** Standard deviation for the nonlinear components of the central normalized instantaneous phase in non-weak signal segments to distinguish MPSK, MFSK, 16QAM from MASK signals, where the former contain phase information, while the latter does not.

**Feature 4** Standard deviation for the absolute value of the nonlinear components of the central normalized instantaneous phase (AVCNIP) in non-weak signal segments to distinguish BPSK signals with its AVCNIP being a constant from QPSK, 16QAM ones with their AVCNIPs not being a constant.

**Feature 5** Standard deviation for the absolute value of central normalized instantaneous frequency (AVCNIF) to distinguish the 2FSK signal with its AVCNIF being a constant from 4ASK ones with its AVCNIF not being a constant.

**Feature 6** Deviation of instantaneous amplitude (DIA) to distinguish 16QAM with a large DIA from 4PSK with a small DIA.

Fig. 2 demonstrates the detailed recognition process of the SCHKS-SSVM classifier.

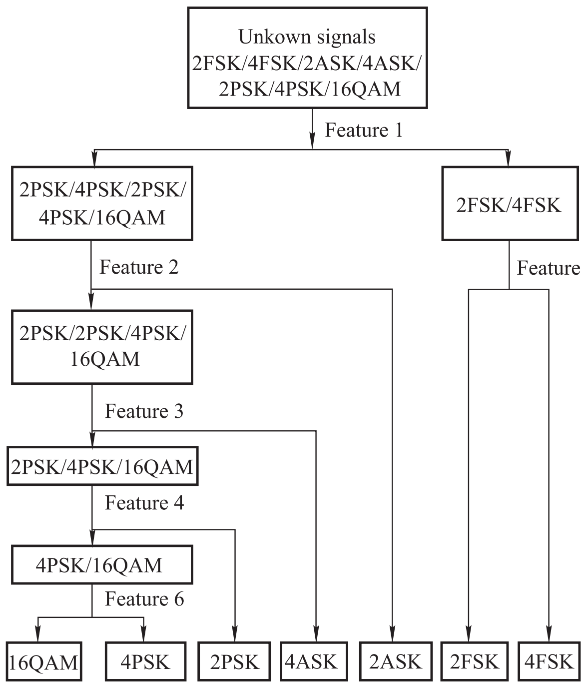


**Fig. 2 Recognition process of the SCHKS-SSVM classifier**

All experiments are carried out on a PC with its CPU running at 3.07 GHz and a 1 GB RAM. The Matlab source code for SSVM is provided by Mangasarian [18,19]. The source code for the other two algorithms, i.e., SSVM and SCHKS-SSVM, are programmed by using Matlab7.0. In this simulation, the SNR of the signal to be recognized ranges from −10 dB to 20 dB with a 10 dB interval. Two hundred signals each from the aforementioned seven types are chosen for training. After training, each of the three recognition algorithms is tested with 10 000 signals. The experimental results, which are averaged over 10 experiments, are shown in the following table.

Simulation results in Table 1 suggest that the SCHKS-SSVM algorithm has very high training accuracy and recognition rate.

**Table 1 Experimental results for comparison among the three algorithms ($C = 20$)**

| SNR/dB | Parameter | SVC | SSVM | SCHKS-SSVM |
|---|---|---|---|---|
| −10 | Training time/s | 4.65 | 0.47 | 0.31 |
| | Training accuracy/% | 68.08 | 83.99 | 85.15 |
| | Test accuracy/% | 59.47 | 80.01 | 83.22 |
| 0 | Training time/s | 4.99 | 0.52 | 0.28 |
| | Training accuracy/% | 88.75 | 92.14 | 92.52 |
| | Test accuracy/% | 85.92 | 89.20 | 90.05 |
| 10 | Training time/s | 5.01 | 0.33 | 0.26 |
| | Training accuracy/% | 95.38 | 100 | 100 |
| | Test accuracy/% | 87.65 | 99.04 | 100 |
| 20 | Training time/s | 8.09 | 0.38 | 0.21 |
| | Training accuracy/% | 100 | 100 | 100 |
| | Test accuracy/% | 90.45 | 100 | 100 |

At least, it has the same performance as that of SSVM. Moreover, the SCHKS-SSVM algorithm can process a large amount of input data very fast, due to the fact that it usually requires only a few iteration steps.

**Table 2 Comparison of average correct recognition rate based on SCHKS-SSVM**

| SNR/dB | Method | 2ASK | 4ASK | 2FSK | 4FSK | 2PSK | 4PSK | 16QAM |
|---|---|---|---|---|---|---|---|---|
| −10 | DT | 0 | 0 | 0 | 75.0 | 0 | 0 | 60.5 |
| | HOC | 22 | 0 | 0 | 0 | 65 | 2 | 41.2 |
| | CHKS-SSVM | 85.02 | 83.01 | 84.75 | 80.12 | 83.65 | 82.89 | 76.14 |
| 0 | DT | 21 | 1 | 0 | 82.5 | 0 | 0 | 87.0 |
| | HOC | 97.5 | 95.5 | 99.0 | 83.5 | 100 | 100 | 81.2 |
| | CHKS-SSVM | 94.55 | 95.02 | 93.88 | 95.09 | 94.34 | 93.65 | 89.18 |
| 10 | DT | 65.0 | 89.0 | 94.0 | 95.0 | 100 | 81.5 | 100 |
| | HOC | 98.5 | 99.0 | 100 | 100 | 100 | 100 | 87.2 |
| | CHKS-SSVM | 100 | 100 | 100 | 99.78 | 99.62 | 96.64 | 95.26 |
| 20 | DT | 99.0 | 98.0 | 97.5 | 96.0 | 100 | 100 | 100 |
| | HOC | 99.5 | 100 | 100 | 100 | 100 | 100 | 100 |
| | CHKS-SSVM | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Table 2 demonstrates the recognition rate of the SCHKS-SSVM algorithm detailed in Fig. 2 for seven types of digital modulation signals including 2ASK, 4ASK, 2FSK, 4FSK, 2PSK, 4PSK and 16QAM. To validate the effectiveness of the SCHKS-SSVM algorithm, another two types of recognition algorithms, namely, decision tree (DT) and higher-order cumulant (HOC) [20,21], are chosen as a benchmark in the simulation for comparison. The simulation results in Table 2 show that the proposed SCHKS-SSVM algorithm has a higher recognition rate, which demonstrates its validity.

## 5. Conclusions

In this paper, a novel SCHKS-SSVM-based modulation classification and recognition algorithm has been proposed. It transforms the convex quadratic programming problem of the traditional SVM algorithm into an unconstrained one, which simplifies the calculation process. The introduction of the proposed CHKS smooth function eventually smoothes the transformed unconstrained problem. This consequently solves the second-order undifferentiability of the unconstrained problem, and reduces its corresponding training errors as a result. The Newton-Armijo algorithm is chosen to train the SCHKS-SSVM. The algorithm improves the training speed and saves a large amount of memory through batch training. The simulation results show that the recognition rate of the algorithm can exceed 95% when SNR is no less than 10 dB. Compared with other algorithms, the proposed one can save the testing and training time for up to 10 times, while improving the training accuracy by 3%∼5%, which validates the effectiveness of the proposed algorithm. Theoretically, the proposed SCHKS-SSVM shows its advantages when dealing with classification problems with high dimensional space and massive data volume.

## References

[1] J. P. Wu, Y. Z. Han, J. M. Zhang, et al. Automatic modulation recognition of digital communication signals. *Proc. of the International Conference on Pervasive Computing Signal Processing & Applications*, 2010: 590 – 593.

[2] B. I. Dahap, H. S. Liao. Advanced algorithm for automatic modulation recognition for analogue & digital signals. *Proc. of the International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering*, 2015: 32 – 36.

[3] S. B. Sadkhan. A proposed digital modulated signal identification based on pattern recognition. *Proc. of the International Multi-Conference on Systems Signals and Devices*, 2010: 1 – 6.

[4] X. Liu, J. Su, W. Wu. A modulation recognition method based on carrier frequency estimation and decision theory. *Proc. of the 16th Asia-Pacific Conference on Communications*, 2010: 6 – 11.

[5] S. P. Li, F. C. Chen, L. Wang. Modulation recognition algorithm of digital signal based on support vector machine. *Control & Decision Conference*, 2012, 229(5): 3326 – 3330.

[6] J. Sun, T. Zhang, F. Liu. Automatic multi-class digital modulation recognition algorithms based on support vector machines. *Journal of Xi'an Jiaotong University*, 2004, 38(6): 619 – 622. (in Chinese)

[7] T. S. Tabatabaei, S. Krishnan, A. Anpalagan. SVM-based classification of digital modulation signals. *Proc. of the IEEE*

*International Conference on Systems Man and Cybernetics*, 2010: 277 – 280.

[8] Z. Zhao, Y. Zhou, F. Mei, et al. Automatic modulation classification by support vector machines. *Proc. of the Advances in Neural Networks — International Symposium on Neural Networks*, 2004: 654 – 659.

[9] N. Y. Deng, Y. J. Tian. *The new method in data mining-SVM*. Beijing: Science Press, 2004. (in Chinese)

[10] Y. J. Lee, O. L. Mangasarian. SSVM: a smooth support vector machine for classification. *Computational Optimization and Applications*, 2001, 20(1): 5 – 22.

[11] M. H. Valipour, M. M. Homayounpour, M. A. Mehralian. Automatic digital modulation recognition in presence of noise using SVM and PSO. *Proc. of the 6th International Symposium on Telecommunications*, 2012: 378 – 382.

[12] M. Liu, Y. Zhao, L. Shi, et al. Research on recognition algorithm of digital modulation by higher order cumulants. *Proc. of the International Conference on Instrumentation & Measurement*, 2014: 686 – 690.

[13] L. X. Wang, Y. J. Ren. Recognition of digital modulation signals based on high order cumulants and support vector machines. *Communications Technology*, 2009, 4: 271 – 274.

[14] P. H. Li, H. X. Zhang, X. Y. Wang, et al. Modulation recognition of communication signals based on high order cumulants and support vector machine. *Journal of China Universities of Posts & Telecommunications*, 2012, 19(11): 61 – 65.

[15] J. Z. Xiong, J. L. Hu, H. Q. Yuan, et al. Research on a new class of functions for smoothing support vector machines. *Acta Electronica Sinica*, 2007, 35(2): 366 – 370. (in Chinese)

[16] A. Ebrahimzadeh, S. A. Seyedin. Digital signal types identification using a hierarchical SVM-based classifier and efficient features. *Proc. of the International Conference on Computing: Theory & Applications*, 2007: 521 – 525.

[17] E. H. Keshk, M. A. El-Naby, R. M. Al-Makhlasawy, et al. Automatic modulation recognition in wireless multi-carrier wireless systems with cepstral features. *Wireless Personal Communications*, 2015, 81(3): 1243 – 1288.

[18] J. M. Su, S. Y. Ruan, Y. L. Wang. *Matlab engineering mathematics*. Beijing: Electronic Industry Press, 2005. (in Chinese)

[19] F. Zhang. Matrix analysis and applications. *International Journal of Information and Systems Sciences*, 2004, 309(1): 1 – 9.

[20] Q. Chi. Research on incremental decision tree algorithm. *Proc. of the International Conference on Electronic & Mechanical Engineering & Information Technology*, 2011: 303 – 306.

[21] S. Ghasemi, A. Gangal. An effective algorithm for automatic modulation recognition. *Proc. of the Signal Processing & Communications Applications Conference*, 2014: 903 – 906.

## Biographies

**Xiaolin Zhang** was born in 1971. He is an engineering doctor of information and communication engineering. He graduated from the College of Information and Communication Engineering in 1995 and received a Ph.D. degree from Harbin Engineering University, Harbin, China, in 2007. Now, he is an associate professor, a master graduate student tutor in Harbin Engineering University. He has published a number of academic papers, of which 10 pieces of EI retrieval. He successively presided over complete postdoctoral fund in China and postdoctoral fund in Heilongjiang Province, as a technical director successively participated in national, provincial and lateral projects. He is the member of China Communication Association. His research interests are communication signal detection and recognition, and broadband digital communication system.

E-mail: zhangxiaolin@hrbeu.edu.cn

**Jian Chen** was born in 1990. In 2014, he enrolled in Harbin Engineering University with a major of information and communication engineering. He is pursuing his engineering master's degree in Harbin Engineering University, now. Especially in the direct sequence spread spectrum communication, he has a very deep research experience. He has published two international EI conference papers. His research interests are spread spectrum communication, cognitive precision communication electronic warfare, communication electronic countermeasure, and non-cooperative wireless communication detection.

E-mail: chen_jian@hrbeu.edu.cn

**Zhiguo Sun** was born in 1977. He was an engineering doctor of information and communication engineering. He received a Ph.D. degree from Harbin Engineering University, Harbin, China, in 2005. Now, he is an associate professor, a doctor/master graduate student tutor in Harbin Engineering University. His research interests are cognitive precision communication electronic warfare, non-cooperative wireless communication detection, high efficiency electromagnetic interference guidance, and multi-dimensional feature covert communication technology. He has participated in National High Technology Research and Development Program, National Defense Pre-Research and over 20 National Natural Fund Transversal Period Research Projects. He has published more than 30 academic papers and more than 10 patents for national invention.

E-mail: sunzhiguo@hrbeu.edu.cn