Automatic Modulation Recognition by Support Vector Machines Using Wavelet Kernel

To cite this article: X Z Feng et al 2006 J. Phys.: Conf. Ser. 48 1264

View the article online for updates and enhancements.

Related content
- Activity Recognition in Egocentric video using SVM, kNN and Combined SVM&kNN Classifiers
  K.P. Sanal Kumar and R. Bhavani Dr.
- The Automation System Censor Speech for the Indonesian Rude Swear Words Based on Support Vector Machine and Pitch Analysis
  S N Endah, D M K Nugraheni, S Adhy et al.
- Support vector machines and generalisation in HEP
  Adrian Bevan, Rodrigo Gamboa Goñi, Jon Hays et al.

Recent citations
- Ka Mun Ho et al
- Xin Zhou et al
Automatic Modulation Recognition by Support Vector Machines Using Wavelet Kernel

X Z Feng, J Yang, F L Luo, J Y Chen and X P Zhong
College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha, China
E-mail: kd805fxz@sina.com

Abstract. Automatic modulation identification plays a significant role in electronic warfare, electronic surveillance systems and electronic counter measure. The task of modulation recognition of communication signals is to determine the modulation type and signal parameters. In fact, automatic modulation identification can be range to an application of pattern recognition in communication field. The support vector machines (SVM) is a new universal learning machine which is widely used in the fields of pattern recognition, regression estimation and probability density. In this paper, a new method using wavelet kernel function was proposed, which maps the input vector $x_i$ into a high dimensional feature space $F$. In this feature space $F$, we can construct the optimal hyperplane that realizes the maximal margin in this space. That is to say, we can use SVM to classify the communication signals into two groups, namely analogue modulated signals and digitally modulated signals. In addition, computer simulation results are given at last, which show good performance of the method.

1. Introduction

Automatic modulation identification plays an important role in electronic warfare (EW), electronic surveillance systems, and electronic counter measures (ECM). The task of modulation recognition of communication signals is to determine the modulation type and signal parameters. In fact, automatic modulation identification can be range to an application of pattern recognition in communication field.

The support vector machines (SVM) is a new universal learning machine proposed by Vapnik in 1992, which is widely used in the fields of pattern recognition, regression estimation and probability density function estimation. After the first preliminary studies, SVM have shown a remarkable efficiency, especially when compared with traditional artificial neural networks (ANN), which have the multilayer perceptron. The main advantage of SVM, with respect to ANN, consists in the structure of the learning algorithm, characterized by the resolution of a constrained quadratic programming problem (CQP), where the drawback of local minima is completely avoided. A SVM finds the hyper-plane that separates the largest possible fraction of points of the same class on the same side, while minimizing the distance from either class to the hyper-plane. This hyper-plane is called Optimal Separating Hyperplane (OSH) which minimizes the risk of misclassifying not only the samples in the training set but also the unknown samples of the test sets.

The wavelet transform (WT) is a powerful tool for analyzing non-stationary signals, which include digital communication signals. The WT magnitude of communication signals vary with modulation types.
2. The Principle of SVM
In pattern recognition, we try to estimate a function of $f: \mathbb{R}^N \rightarrow \{+1\}$ using training data, that is, $N$-dimensional patterns $x_i$ and class labels $y_i$,

$$(x_i, y_i) \in \mathbb{R}^N \times \{\pm 1\} \ i = 1, 2, \ldots, N$$

(1)

Figure 1. Optimal hyperplane.

For new examples $\{x, y\}$, which were generated from the same probability distribution $P(x, y)$ as the training data, the $f$ will correctly classify $\{x, y\}$, namely $f(x) = y$. Statistical theory shows that it is crucial to restrict the class of the functions that the learning machine can implement to one with a capacity that is suitable for the amount of available training data. Otherwise even a function does well on the training data, it will not have the correct results with the new examples. SV classifiers are based on the class of hyperplanes,

$$(w \cdot x) + b = 0 \quad w \in \mathbb{R}^N, b \in \mathbb{R}$$

(2)

Where decision functions are:

$$f(x) = \text{sign}((w \cdot x) + b)$$

(3)

As was shown in Figure 1, the optimal hyperplane can be defined as the one with the maximal margin of separation between two classes. Note that the margin is $2/\|W\|$.

Figure 2. The idea of Support vector machines.

Figure 2 shows the basic idea of SVM, which map the original data into a feature space $F$ via a nonlinear map. Consequently we can perform linear algorithm in feature space $F$ by the use of kernels. We define kernel function as:

$$k(x, y) = (\Phi(x) \cdot \Phi(y))$$

(4)
Where $\Phi$ is the nonlinear map from input space $\mathbb{R}^N$ to feature space $F$. Support Vector Machines realize the following idea. Suppose we are given two classes of samples $(x_i, y_i), x_i \in \mathbb{R}^N, y_i \in \{-1, 1\}, i=1,\ldots, l$, as a training set. We first map the input vector $x_i$ into a high dimensional feature space, $F$, through a nonlinear mapping function (4); then construct the optimal hyperplane that realizes the maximal margin in this space. With the so called kernel trick, the mapping (4) is implicitly implemented by some kernel function $K(x, y)$, which defines an inner product in the feature space. The decision function given by an SVM is thus in a linear form with $\Phi(x)$ as:

$$f(x) = \langle w, \Phi(x) \rangle + b$$

(5)

Where $w$ is the normal vector of the decision hyperplane in $F$, and $b$ is the bias. A sample $x$ with $f(x) > 0$ is assigned to class $\{1\}$, otherwise it is assigned to class $\{-1\}$.

3. Characteristic of communication signals

From the above discussions, we find the key issue is to choose a proper kernel function, or in other words, to choose a proper map $\Phi$. In this case, we can use the simple linear algorithm.

![Graph](image)

**Figure 3.** WT magnitudes.

The communication signals can be divided into two groups, namely analogue modulated signals and digitally modulated signals. In digitally modulated signals, symbol changes give rise to transients in the modulated signals. However these transients will not occur in analogue modulated signals. The time-frequency technique is suitable for these transients’ detection and analysis. The wavelet transform (WT) is a powerful tool for analyzing non-stationary signals, which include communication signals.

$$CWT(a, \tau) = \int s(t)\psi^*(t)dt = \frac{1}{\sqrt{a}} \int s\left(\frac{t - \tau}{a}\right)dt$$

(6)

The continuous wavelet transform (CWT) of a signal $s(t)$ is defined as (6), Where $a$ is the scale, $\tau$ is the translation, and the superscript * denotes complex conjugate. The function $\psi(t)$ is the mother wavelet, and the baby wavelet $\varphi(t)$ come from time-scaling and translation of $\psi(t)$.

The discrete-time Haar wavelets are

$$\frac{1}{\sqrt{a}} \psi\left(\frac{k}{a}\right) = \begin{cases} 
\sqrt{a} & k = -a/2, -a/2 + 1, \ldots, -1 \\
\sqrt{a} & k = 0, 1, \ldots, a/2 - 1 \\
0 & \text{otherwise}
\end{cases}$$

(7)

If we use the Harr wavelets as the nonlinear map $\Phi$ with a proper scale $a$, the magnitude of the wavelet transform of the analogue modulated signals and digitally modulated signals were shown in
Figure 3. From the characteristics of the communication signals, we find if we select a proper WT function, the communication signals can be classified into two groups in terms of the WT magnitude. These two groups will be analogue modulated signals and digitally modulated signals exactly, that is to say, we can use the WT kernel function SVM to classify the communication signals.

4. Simulation
The numerical results that we obtained with the method discussed above are presented in this section. The mother wavelet was the Haar function in (7). The test signals contained three modulation types: BPSK, QPSK, and AM, corrupted with a band-limited Gaussian noise. The sampling frequency was normalized to unity. The intermediate frequency was 0.25. Here we just classify these signals into two groups, which is digitally modulated signals (including BPSK and QPSK), and analogue modulated signals (including AM). The results are shown in table 1. In the table, $Train$ stands for the training set size, $Test$ for test set size. Recognition rates on the test set are shown at last column.

5. Conclusion
In this paper a new algorithm, which use WT as the kernel functions in SVM to classify analogue modulated signals and digitally modulated signals, was proposed. In addition, computer simulation results which show good performance of the method are given at last.

References

### Table 1. Recognition results.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>$Train$</th>
<th>$Test$</th>
<th>Recognition rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPSK</td>
<td>50</td>
<td>1000</td>
<td>0.844</td>
</tr>
<tr>
<td>QPSK</td>
<td>50</td>
<td>1000</td>
<td>0.832</td>
</tr>
<tr>
<td>AM</td>
<td>50</td>
<td>1000</td>
<td>0.865</td>
</tr>
</tbody>
</table>