

Semi-Parametric Statistic Model via Support Vector Regression for Radar Target HRRP Recognition

Zhu Jiehao, Li QiQin, Zhu Yu, and Fang Jian

Electronic Information Control Laboratory, No.29 Research Institute of CETC, Chengdu, China, 610036
zhujhms@gmail.com

Abstract

In radar target high-resolution range profile (HRRP) recognition, the target aspect sensitivity problem is mainly solved by angular domain division and template modeling for each angular sector. Semi-parametric statistic model has both advantages of parametric statistic model and non-parametric statistic model, and is proved to be a valid HRRP template. However, when the number of HRRP samples in an angular sector is large, the efficiency might be low because of the non-parametric correction factor used in semi-parametric statistic model. To solve this problem, a semi-parametric statistic model via support vector regression (SVR) is proposed in this paper. By SVR, the idea is to reduce all the samples to only support vectors to formulate the non-parametric correction factor. Experiment results using 5 aircraft HRRP dataset demonstrate the efficiency of the proposed model.

1. Introduction

As the increasing using of the wideband radar systems, detailed electromagnetic characteristics of radar target can be obtained. High-resolution range profile (HRRP) is one of the mostly used wideband radar target echo. It represents the target scattering distribution along the radar line-of-sight, and has the advantages of easy acquisition and processing. HRRP is now receiving more and more attention especially in radar automatic target recognition community.

Radar target has the aspect sensitivity when radar works in optical region. It is mainly solved by angular domain division and template modeling for each angular sector. Under Bayesian classification framework, the HRRP template model is established on the statistic distribution of each HRRP range bin [1-5]. To find proper description of HRRP statistic distribution, early studies mainly focused on single parametric statistic model, like Gaussian model [2, 3] and Gamma model [1, 5]. In [6], a compound statistic model is proposed, which improves the accuracy of single parametric statistic model. In [7], non-parametric statistic model is used. Non-parametric model is more flexible because it is formulated directly by the samples without knowing its distribution previously. In [8], a semi-parametric statistic model is proposed. It uses prior knowledge via parametric model and corrects it by non-parametric correction factor. The purpose is to combine both advantages of parametric statistic model and non-parametric statistic model.

Due to the target aspect sensitivity, in order to construct a complete HRRP template database, a large amount of samples are needed. Furthermore, radar target recognition is always a multi-classification problem. It leads to the multiplied samples when the number of class is increasing. The massive data brings the difficulties of template formulation and decreases the recognition efficiency. To solve the problem mentioned above, a novel and highly efficient HRRP template model via support vector regression (SVR) under the frame of semi-parametric radar target HRRP recognition method is proposed in this paper. Experiment results demonstrate that the efficiency is significantly improved but nearly without any recognition rate decreasing.

2. Brief review of SVR

Given a training dataset

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\} \in (\mathbf{R}^d \times \mathbf{R})^m \quad (1)$$

regression is to find a real-valued function $f(\mathbf{x})$ on \mathbf{R}^d , where $y = f(\mathbf{x})$ can be used to determine the value of y with respect to arbitrary input \mathbf{x} . If $f(\mathbf{x})$ is a linear function, it is also called the linear regression problem. The geometrical meaning of a linear regression problem is to find a hyper plane which deviates from the training dataset with least distance. Support vector regression is realized by converting the linear regression problem to linear classification problem [9] based on large margin criterion. It is formulated by the following optimization problem

$$\begin{aligned} & \min_{\omega, b, \xi, \xi^*} \quad \frac{1}{2} \|\omega\|^2 + C \cdot \frac{1}{m} \sum_{i=1}^m (\xi_i + \xi_i^*) \\ & \text{s.t.} \quad (\omega \cdot x_i + b) - y_i \leq \varepsilon + \xi_i, \quad i = 1, 2, \dots, m \\ & \quad y_i - (\omega \cdot x_i + b) \leq \varepsilon + \xi_i^*, \quad i = 1, 2, \dots, m \\ & \quad \xi_i \geq 0, \quad \xi_i^* \geq 0, \quad i = 1, 2, \dots, m \end{aligned} \quad (2)$$

where $\xi = (\xi_1, \xi_2, \dots, \xi_m)^T$ and $\xi^* = (\xi_1^*, \xi_2^*, \dots, \xi_m^*)^T$ are called relaxed variables, and C is the punish factor.

The original optimization problem (2) is often solved by its dual problem in practice. In dual problem formulation, kernel function can be easily implemented. Kernel function maps the original data space where data can be only non-linearly regressed into a higher dimension space where data can be linearly regressed. By introducing the kernel function, the original optimization problem (2) is now presented as follows

$$\begin{aligned} \min_{\alpha, \alpha^*} \quad & \frac{1}{2} \sum_{i,j=1}^m (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(x_i \cdot x_j) + \varepsilon \sum_{i=1}^m (\alpha_i^* + \alpha_i) - \sum_{i=1}^m y_i (\alpha_i^* - \alpha_i) \\ \text{s.t.} \quad & \sum_{i=1}^m (\alpha_i^* - \alpha_i) = 0, \quad i = 1, 2, \dots, m \\ & 0 \leq \alpha_i, \quad \alpha_i^* \leq \frac{C}{m}, \quad i = 1, 2, \dots, m \end{aligned} \quad (3)$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$, $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*)^T$, and $K(x_i, x_j)$ is kernel function.

By solving the optimization problem (3), and getting the optimal results α and α^* , final regression function can be formulated as follows

$$f(x) = \sum_{i=1}^m (\bar{\alpha}_i^* - \bar{\alpha}_i) K(x_i, x) + \bar{b} \quad (4)$$

where optimal solution \bar{b} is a constant. It can be obtained by choosing proper values of $\bar{\alpha}$ and $\bar{\alpha}^*$. The dual problem (3) is also called ε -SVR, and the inputs x_i corresponding to $\bar{\alpha}_i \neq 0$ or $\bar{\alpha}_i^* \neq 0$ are called support vectors (SV).

3. Support Vector Density Estimation

The probability density function (PDF) estimation can be solved by its distribution function regression. In fact, if the distribution function is correctly regressed by the training samples, the PDF is then easily calculated by finding the derivation of distribution function accordingly. As mentioned before, SVR is a learning machine based on the large margin criterion. Therefore, by using SVR, the regression of distribution function can be formulated only by the support vectors. Another advantage by using SVR to estimate PDF is that the distribution of data is not needed previously, so, SVR based PDF estimation is a non-parametric method.

The PDF estimation is in fact an ill problem actually. That is because minor fluctuation of distribution function $F(x)$ will cause a drastic change of its derivation. SVR can be used to solve this problem for that SVR is a regularization model, that is to say it solves the regression problem by a regularized objective function. The main idea of PDF estimation via SVR is to obtain the regressed distribution function, and then calculate the PDF by its derivation.

For i.i.d. training samples, an efficient kind of distribution function estimation is

$$F_m(x) = \frac{1}{m} \sum_{i=1}^m \theta(x - X_i) \quad (5)$$

where $\theta(\cdot)$ is the unit step function. $F_m(x)$ is also called the experienced distribution function. It can be seen that $F_m(x)$ strictly satisfies to the training samples. However, $F_m(x)$ should not only satisfy the training samples but also the arbitrary one. To achieve this object, regularization technique can be used. Utilizing SVR to regress $F_m(x)$ by data pairs $\{(x_1, F_m(x_1)), (x_2, F_m(x_2)), \dots, (x_m, F_m(x_m))\}$, we can get the regularized estimation of $F_m(x)$, and also the PDF $p(x)$.

Using “generalized square loss function” as the SVR objective function, the corresponding optimization function can be formulized again as follows [10]

$$\begin{aligned} \min_{\beta, \xi} \quad & \sum_{i=1}^m \xi_i^2 + C \sum_{i=1}^m \beta_i \\ \text{s.t.} \quad & \sum_{i=1}^m \beta_i K(X_i, X_j) + \xi_j = F(X_j), \quad j = 1, 2, \dots, m \\ & \sum_{i=1}^m \beta_i = 1 \\ & \beta_i \geq 0, \quad i = 1, 2, \dots, m \end{aligned} \quad (6)$$

The objective function of (6) has the meaning of minimal estimation error and least support vector number. By solving optimization problem (6), we can obtain the regression of the distribution function

$$F(x) = \sum_{i=1}^m \beta_i K(X_i, x) + \bar{b} \quad (7)$$

and its PDF is

$$p(x) = \sum_{i=1}^m \beta_i \tilde{K}(X_i, x) \quad (8)$$

where $\tilde{K}(\cdot)$ is the derivation of $K(\cdot)$.

4. Semi-Parametric Statistic Model via SVR

According to [8], non-parametric correction factor $\hat{r}(x)$ and semi-parametrical PDF $\hat{p}(x)$ are given as follows

$$\hat{r}(x) = \frac{1}{m} \sum_{i=1}^m \frac{K_h(X_i - x)}{p(X_i, \hat{\theta})} \quad (9)$$

$$\hat{p}(x) = \hat{r}(x) p(x, \hat{\theta}) = \frac{1}{m} \sum_{i=1}^m K_h(X_i - x) \frac{p(x, \hat{\theta})}{p(X_i, \hat{\theta})} \quad (10)$$

In equation (9), $K_h(\cdot)$ is called window function, $p(x, \hat{\theta})$ is the parametric PDF estimation of variable x . Considering that non-parametric correction factor $\hat{r}(x)$ is formulated by all the training samples, when the number of training samples is large, it leads to more storage capacity and lower efficiency. However, if only support vectors are used to formulate the non-parametric correction factor, the above mentioned problems can be solved.

By using SVR, we can reformulate the non-parametric correction factor $\hat{r}_{SV}(x)$. Different from Parzen Window used in (9), we utilize a new formulation here^[11]

$$\hat{r}_{SV}(x) = \frac{P^1(x, r)}{P^0(x, r)} \quad (11)$$

where $p(x, r)$ stands for the probability that lies in a ball B where its center is sample x itself and its radius is r ,

$$P\{X \in B(x, r)\}, \quad B(x, r) = \{y : \|x - y\| \leq r\} \quad (12)$$

and P^1 and P^0 are the probability of random variable x estimated by non-parametric method and parametric method respectively. According to the redefined non-parametric correction factor $\hat{r}_{SV}(x)$, the semi-parametric statistic model is

$$\hat{p}(x) = \hat{r}_{SV}(x) p(x, \hat{\theta}) = \frac{P^1(x, r)}{P^0(x, r)} p(x, \hat{\theta}) \quad (13)$$

In equation (11), the radius r acts as a slide factor. When radius r increases, the non-parametric correction factor makes the curve smoother. In fact, when radius r limits to 0, equation (11) becomes to the ratio of non-parametric PDF and parametric PDF, and equation (13) degenerate to non-parametric PDF estimation. When radius r increases to include the whole sampling space, P^0 , P^1 , and equation (11) equal to 1, equation (13) then degenerates to parametric PDF estimation. So, when using semi-parametric PDF estimation, the radius r should be carefully tuned to satisfy the demand of the optimal PDF estimation.

5. Experiment and Analysis

In this section, HRRP dataset including 5 aircraft models, i.e. Su27, F16, M2000, J8II, and J6, is used. Each aircraft model has 1800 HRRPs, and all the HRRPs have been added Gaussian white noise on I/Q component respectively under different SNR. Meanwhile, all the HRRP data are obtained when the aircraft model is put on the turntable, so the HRRP shift sensitivity problem is not mentioned in this experiment. At last, all the HRRP data are preprocessed by L_2 normalization.

The designed experiment mainly compares the average recognition results on the HRRP dataset using three different methods, i.e., Parametric PDF (PPDF), SVR based Non-Parametric PDF (kerNPDF), and SVR based Semi-Parametric PDF (kerSPDF). When utilizing parametric PDF, Gamma model is chosen for parametric PDF, and SVR based PDF defined by (8) is used for non-parametric PDF. All the HRRP data is preprocessed by whiting operation to decrease the disadvantages brought by the difference of echo amplitude distribution in each range bin. Figure 1 illustrates the average recognition results of the 3 methods under different SNR. The detailed recognition results and the SV ratio under 20dB SNR are listed in table 1 and table 2 respectively. The support vector number is counted when the weighting coefficient satisfies $\beta_i \geq 0.0001$ otherwise it will be discarded.

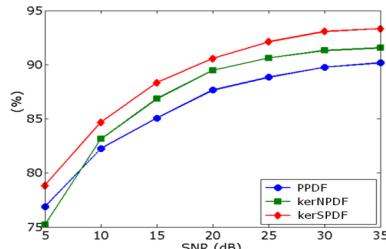


Figure 1 Recognition results of PPDF, kerNPDF, and kerSPDF under different SNR

Table 1 Recognition results of PPDF, kerNPDF, and kerSPDF (%), SNR=20dB

	Su27	F16	M2000	J8II	J6	Average
PPDF	92.6	89.3	94.7	69.9	91.8	87.6
kerNPDF	95.8	86.4	91.9	81.6	91.8	89.5
kerSPDF	94.9	90.0	95.1	80.2	92.7	90.6

Table 2 SV ratio of HRRP under kerSPDF (%), SNR=20dB

	Su27	F16	M2000	J8II	J6	Average
SV ratio	14.1	10.8	12.3	13.6	10.6	12.3

From the experiment results, it can be seen that the proposed method significantly reduced the model formulation samples, and even have a pretty better average recognition performance. Meanwhile, although the individual recognition performance may not be always the best, the worst recognition ratio is still between parametric method and non-parametric method. This can be explained that semi-parametric method merges the advantages of parametric method and non-parametric method. So, semi-parametric method is more adaptive for HRRP dataset and available to radar HRRP recognition.

6. Conclusion

Semi-parametric statistic model increases the HRRP template accuracy by using prior knowledge and correcting parametric statistic model by non-parametric correction factors. Semi-parametric statistic model merges both advantages of parametric method and non-parametric method. It forms a unified and simple statistic model for the distribution of HRRP range bins. However, the Parzen window based semi-parametric method has low efficiency especially when the number of training samples is large because the non-parametric correction factor is formulated by all the training samples. Semi-parametric statistic model via SVR proposed in this paper easily solves the above mentioned problems. Firstly, the non-parametric PDF is estimated by SVR to form the non-parametric correction factor. Secondly, by using this factor, semi-parametric statistic model is formulated by correcting the Gamma parametric statistic model. Because only support vectors are used to formulate the model, the efficiency problem is solved. HRRP dataset of 5 aircraft models is tested, and proves the availability and efficiency of the proposed method.

7. References

1. A. R. Webb, "Gamma Mixture Models for Target Recognition". Pattern Recognition, 2000, 33(12): 2045-2054.
2. S. P. Jacobs, "Automatic Target Recognition Using High-Resolution Radar Range Profiles". Washington: Washington University, 1999.
3. L. Du, "Study on Radar HRRP Target Recognition". Xi'an: Xidian University, 2007.
4. R. van der Heiden, F. C. A. Groen, "Box-Cox Metric for Nearest Neighbour Classification Improvement". Pattern Recognition, 1997, 30(2): 273-279.
5. K. Copsey, A. Webb, "Bayesian Gamma Mixture Model Approach to Radar Target Recognition". IEEE Transactions on Aerospace and Electronic Systems, 2003, 39(4): 1201-1217.
6. L. Du, H. W. Liu, Z. Bao, et al., "A Two-Distribution Compounded Statistical Model for Radar HRRP Target Recognition". IEEE Transactions on Signal Processing, 2006, 54(6): 2226-2238.
7. F. Zhao, J. Y. Zhang, and J. Liu, "Radar Target Recognition Based on Nonparametric Density Estimation". Journal of Electronics & Information Technology, 2008, 30(7): 1740-1743.
8. J. Zhu, J. J. Zhou, and J. Wu, "Radar Target Recognition Based on Semi-parametric Density Estimation". Journal of Electronics & Information Technology, 2010, 32(9): 2161-2166.
9. N. Y. Deng and Y. J. Tian, "A New Method of Data Mining - Support Vector Machines". Beijing: Science Press, 2004.
10. J. Weston, A. Gammerman, et al., "Support Vector Density Estimation". Cambridge, MA: MIT Press, 1999.
11. P. Chaudhuri, A. K. Ghosh, and H. Oja, "Classification Based on Hybridization of Parametric and Nonparametric Classifiers". IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(7): 1153-1164.