Specific radar emitter identification based on two stage multiple kernel extreme learning machine

Yu Shi,1,2⁎ Ming-Dong Yuan,2 Junlin Ren,1 and Shengjun Xu1
1School of Information and Control Engineering, Xi’an University of Architecture and Technology, Xi’an, China
2R&D Department, The Smart City Research Institute of China Electronics Technology Group Corporation, Shenzhen, China

Email: shiyaworld@163.com

Introduction: Specific emitter identification (SEI) distinguishes each individual radio emitter of interest utilizing the extracted signal features [1]. We focus on the radar emitters here, i.e. specific radar emitter identification issue, of which the most challenging situation is that the radars may come from the same production line and thus have identical type [2, 3]. To meet this challenge, unintentional modulation on pulse (UMOP) has attracted great attention during the past thirty years or so [1–5]. The individual difference of radars originates from inevitable UMOP caused by the production technique of transmitter, providing a physical basis for feature extraction and individual identification [2, 3].

Pulse envelope [2, 6], frequency and phase modulation profiles [4, 5] were verified to be related to phase noise of transmitters, so they can represent the subtle difference of radars. Further, variational mode decomposition was used to decompose the envelope or instantaneous frequency into many modes to extract the secondary features [3]. In [7], Fourier spectrum was regarded as the primary feature, and then one-dimensional convolutional neural network was applied to obtain the deep feature and within-class scatter in a kernel-induced feature space, i.e. the discriminative adversarial network to magnify the unintentional modulation information. Based on cyclostationarity, [6] presented zero frequency slice of cyclic spectrum (CSO) feature. However, the above features are less comprehensive than the joint time-frequency ones [9–11], among which the ambiguity function (AF) performs well. In terms of accuracy, the optimal kernel combination weights and classifier parameter are solved by alternative optimization, so MKELMs need many iterations to converge. By contrast, two stage MKL finds the optimal kernel weights using independent criteria (e.g. kernel target alignment [12]) in the first stage and then train a standard kernel method using the combined kernel in the second stage, which is more efficient. Therefore, the KDR-based two stage MKELM (KDR-TSMKELM) algorithm is proposed in this paper. KDR is the ratio of between-class scatter and within-class scatter in a kernel-induced feature space, i.e. KDR can measure the class separability of training data [15]. Generally, good separability means high accuracy, so we can use KDR criterion to obtain the kernel weights in the first stage. Moreover, it has been proved that ELM’s minimal norm of output weight property is in line with the large margin theory [13]. Hence, our algorithm makes sense in theory.

Considering R different feature representations, the training data points are written as [xρ,u] = 1,…,U; ν = 1,…,Nuρ, where xρ,ν denotes the ρ-th data vector of the ν-th class in the ρ-th original feature space and νuρ is the amount of data in the υ-th class. Let Φ(·)ρ,ν, be implicit mappings, then their corresponding kernel functions and training kernel matrices are [Kρ,ν] and [Kρ,ν]T, respectively. Thus, the class mean vectors [μρ,u] and the global mean vectors are:

\[ \mu_ρ = \frac{1}{n_ρ} \sum_{ν=1}^{νu_ρ} \Phi(\rho, ν) / νu_ρ, u = 1,…,U \tag{3} \]

Just given the kernel functions, the between-class scatter [Sρ,u,v] and the within-class scatter [vwρ] are defined as:

\[ S_{ρ,u,v} = \sum_{ν=1}^{νu_ρ} (μ_ρ,u - μ_ρ,v)^T (μ_ρ,u - μ_ρ,v) \]

\[ W_{ρ,v} = \frac{1}{U} \sum_{u=1}^{U} \sum_{ν=1}^{νu_ρ} k(\rho, u, ν) \tag{4} \]

To make full use of the discriminative information containing in the whole ambiguity function (AF) plane, a novel two stage multiple kernel extreme learning machine (TSMKELM) method for specific radar emitter identification is proposed. Firstly, the AF plane is segmented into the non-overlapping Doppler shift stripes and each stripe is encoded as a kernel. Next, the discrimination of these stripes is evaluated and sorted according to the kernel discriminant ratio (KDR) criterion, which is in line with the large margin principle of KELM. Then, only the stripes with large KDRs are kept and the combined kernel is calculated by directly using the normalized KDRs as combination weights. At last, the proposed algorithm, named as KDR-TSMKELM, solves the kernel combination weights and kernel classifier parameter separately, bringing much efficiency in practice. Experiments on two real radar datasets validate the proposed algorithm.
shows the detailed segmentation strategy (taking part of the AF plane as feature representations. Hence, we can further segment the AF plane into time. However, there will be too many kernels if slices are used as features. Therefore, the discrimination and informativeness of AF slices can be well evaluated by KDRs, so Equation (6) is very reasonable. Figure 2 also shows that lots of AF slices are uninformative, so we sort the KDRs and keep the slices with large KDRs, which helps to save testing time. Of course, the stripe width should not be too large since the performance of adjacent slices in salient regions fluctuates greatly.

In the second stage, the optimal classifier parameter can be directly calculated by $\alpha^* = (K_{comb}^T + \lambda I)^{-1} Y$. In the testing period, the kernel matrices between any testing data point $|x_{te}|_{r=1}^R$ and all the training data are $(K^r = [k(x_{te}^r, x_u^1), \ldots, k(x_{te}^r, x_u^R)]) \in \mathbb{R}^{1 \times R}$. Consequently, the final output of KDR-TSMKELM is $\{\sum_{r=1}^R \mu^*_r K_{te}^r \}_{1 \times R}$. Since $\alpha^*$ and $\{\mu^*_r\}_{r=1}^R$ are closed-form solutions, our proposed algorithm is faster than MKELM and more suitable for specific radar emitter identification. Generally, to mine the whole AF plane, all the Doppler shift slices or Doppler shift stripes (see below) are regarded as the multiple homogenous feature representations of radar signals and then encoded by multiple kernels; next, the informativeness of AF slices or stripes is automatically sorted according to the KDRs and thus the kernels with large KDRs are chosen to calculate the combined kernel; finally, the standard KELM is used.

To better illustrate our idea, a real single-tone radar dataset ($U = 20$) is utilized. Figure 1 shows a typical sample and its corresponding AF representation. Using stripes, the total number of kernels is reduced and encoding time of different stripe widths. (b) Testing time of different stripe widths. (c) The sorted KDR curve of Data2. (d) The recognition accuracy curve with varying $\alpha$ on Data2.

As can be seen from Figure 1b, the AF plane of single-tone signals has three salient regions (i.e. bright regions). Figure 2a shows that the AF slices located in the salient regions perform better than those in other regions, which is in line with our cognitive. Comparing Figure 2a to b, the KDRs of AF slices are consistent with the recognition accuracy of AF slices. Therefore, the discrimination and informativeness of AF slices can be well evaluated by KDRs, so Equation (6) is very reasonable. Figure 2 also shows that lots of AF slices are uninformative, so we sort the KDRs and keep the slices with large KDRs, which helps to save testing time. However, there will be too many kernels if slices are used as feature representations. Hence, we can further segment the AF plane into stripes. For symmetry, a quarter of AF plane is considered and Figure 3 shows the detailed segmentation strategy (taking part of the AF plane as an example). Using stripes, the total number of kernels is reduced and the training time can be accelerated when the training proportion is large. Of course, the stripe width should not be too large since the performance of adjacent slices in salient regions fluctuates greatly.

Experiments: Data1 ($U = 20$, $N = 1000$) and Data2 ($U = 30$, $N = 3000$) are real radar datasets for experiments (see Figure 1a for waveform). The laptop with Intel core i7 CPU, 20G memory and MATLAB R2020b platform is used. The compared methods are pulse envelope, CS0, AFR with KELM classifier and ten AF slices [11] with KDR-TSMKELM. The Gaussian kernel parameter is empirically set as the mean of pairwise distances of training samples. Let $C = 100$. We randomly select 6%, 8%, 10%, 20% and 50% of data for training and all methods are repeated 10 times. For KDR-TSMKELM, AF slices (stripe width = 1), AF stripes of widths 2 and 3 are tested. To determine that how many kernels should be kept (denoted by $d$), Figure 4 plots the curve of sorted KDRs and the recognition accuracy with different $d$ when the training proportion is 6%. We can see that the best performance can be obtained around the points where the KDR curves start to change more and more slowly. Thus, $d = 120, 60$ and 40 for each stripe width on Data1 and $d = 50, 20$ and 10 on Data2. Table 1 gives the recognition results of Data1. Training and testing time of Data1 are shown in Figure 5.
As shown in Table 1, among the single features, AFR performs best and the envelope is almost useless when there are many classes. The fusion methods outperform the single-feature-based methods and mining the whole AF plane is indeed better than fusing only several AF slices, especially when the training proportion is small. Because of the fixed amount of information, the recognition accuracy of different stripe widths has no significant difference. In detail, KDR-TSMKELM with width 1 has best accuracy; the reason is that the adjacent AF slices have very different discriminative ability, thus the minimum width can eliminate the interference of redundant information to the maximum extent. In terms of time efficiency, the width 2 is preferable. To further verify these observations, Table 2 shows the results of Data2. We can see that KDR-TSMKELM with width 1 still performs best. The training time is 17.2, 11.0 and 8.8 s (50% proportion) for widths 1, 2 and 3 respectively. In summary, from the point view of recognition accuracy and time efficiency, the best choice is KDR-TSMKELM with width 2.

Conclusion: We propose the KDR-TSMKELM algorithm under the framework of MKL. To identify individual radars, the AF plane of radar signals is transformed to AF stripes by non-overlapping segmentation and KDR is used as the kernel combination weights to fuse the stripes. The validity of our method is tested on two real single-tone radar datasets and more complex signals will be tested in future work.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (61803293) and Natural Science Basic Research Plan in Shaanxi Province of China (2019JQ760).

References
1 Liu, M.W., Doherty, J.F.: Nonlinearity estimation for specific emitter identification in multipath channels. IEEE Trans. Inf. Forensics Secur. 6(3), 1076–1085 (2011). https://doi.org/10.1109/TIFS.2011.2134848
2 Zhao, S.Q., Zeng D.G., Wang, W.H.: Mutation grey wolf elite PSO balanced XGBoost for radar emitter individual identification based on measured signals. Measurement 159, 107777 (2020). https://doi.org/10.1016/j.measurement.2020.107777
3 Gok, G., Alp, Y.K., Arikam, O.: A new method for specific emitter identification with results on real radar measurements. IEEE Trans. Inf. Forensics Secur. 15, 3335–3346 (2020). https://doi.org/10.1109/TIFS.2020.2988558
4 Langley, L.E.: Specific emitter identification (SEI) and classical parameter fusion technology. In: Proceedings of WESC0N ’93. San Francisco, pp. 377–381 (1993). https://doi.org/10.1109/WESC0N.1993.488465
5 Ye, H., Liu, Z., Jiang, W.: Comparison of unintentional frequency and phase modulation features for specific emitter identification. Electron. Lett. 48(14), 875–877 (2012). https://doi.org/10.1049/el.2012.0831
6 Li, L., Ji, H.B.: Radar emitter recognition based on cyclostationary signatures and sequential iterative least-square estimation. Expert Syst. Appl. 38, 2140–2147 (2011). https://doi.org/10.1016/j.eswa.2010.07.155
7 Xiao, Y., Wei, X.Z.: Specific emitter identification of radar based on one dimensional convolution neural network. J. Phys.: Conf. Ser. 1550, 032114 (2020). https://doi.org/10.1088/1742-6596/1550/3/032114
8 Chen, P.B., Guo, Y.L., Li, G.: Discriminative adversarial networks for specific emitter identification. Electron. Lett. 56(9), 438–441 (2020). https://doi.org/10.1049/el.2020.0116
9 Liu, S.K., Yan, X.P., Li, P.: Radar emitter recognition based on SIFT position and scale features. IEEE Trans. Circuits Syst. II, Exp. Briefs 65(12), 2062–2066 (2018). https://doi.org/10.1109/TCSII.2018.2819666
10 Zhu, M.Z., Zhang, X.L., Qi, Y.: Compressed sensing mask feature in time-frequency domain for civil flight radar emitter recognition. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Algary, Canada, pp. 2146–2150. https://doi.org/10.1109/ICASSP.2018.8462139
11 Shi, Y., Ji, H.B.: Kernel canonical correlation analysis for specific radar emitter identification. Electron. Lett. 50(18), 1318–1320 (2014). https://doi.org/10.1049/el.2014.1459
12 Wang, T.H., Zhao, D.Y., Feng, Y.S.: Two-stage multiple kernel learning with multiclass kernel polarization. Knowledge Based Syst. 48(2), 10–16 (2013). https://doi.org/10.1016/j.knosys.2013.04.006
13 Huang, G.B., Zhou, H.M., Ding, X.J.: Extreme learning machine for regression and multiclass classification. IEEE Trans. Syst. Man, Cybern. B, Cybern. 42(2), 513–520 (2012). https://doi.org/10.1109/TSMCB.2011.2168604
14 Liu, X.W., Wang, L., Huang, G.B.: Multiple kernel extreme learning machine. Neurocomputing 149, 253–264 (2015). https://doi.org/10.1016/j.neucom.2013.09.072
15 Mika, S., Ratsch, G., Weston, J.: Fisher discriminant analysis with kernels. In: IEEE Signal Processing Society Workshop. Madison, WI, pp. 41–48 (1999). https://doi.org/10.1109/NNSP.1999.788121