Neurocomputing 248 (2017) 67–75

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Classification of the trained and untrained emitter types based on class probability output networks

Lee Suk Kim[a], Han Bin Bae[a], Rhee Man Kil[a,*], Churl Hee Jo[b]

[a] Department of Electrical and Computer Engineering, Sungkyunkwan University, 2066, Seobu-ro, Jangan-gu, Suwon,Gyeonggi-do 16419, Republic of Korea
[b] Electronic Warfare R&D Laboratory, LIGNex1 Co., Ltd., 333, Pangyo-ro, Bundang-gu, Seongnam-City, Gyeonggi-do 13488, Republic of Korea

ARTICLE INFO

Article history:
Received 29 June 2016
Revised 19 January 2017
Accepted 22 January 2017
Available online 10 March 2017

Keywords:
Classification
Radar patterns
Trained and untrained emitter types
Conditional class probability

ABSTRACT

Modern airplanes and ships are equipped with radars emitting specific patterns of electromagnetic signals. The radar antennas are detecting these patterns which are required to identify the types of emitters. A conventional way of emitter identification is to categorize the radar patterns according to the sequences of radar frequencies, differences in time of arrivals, and pulse widths of emitting signals by human experts. In this respect, this paper proposes a method of classifying the radar patterns automatically using the network of calculating the *p*-values for testing the hypotheses of the types of emitters referred to as the class probability output network (CPON). The proposed method also provides a new way of identifying the trained and untrained emitter types. Through the simulation for radar pattern classification, the effectiveness of the proposed approach has been demonstrated.

1. Introduction

In modern days, radars are essential devices to detect objects such as airplanes or ships. For detecting objects emitting specific patterns of electromagnetic signals, the detected signal patterns should be analyzed and categorized according to the types of emitters. This emitter identification (or radar pattern classification) plays an important role, especially in the electronic warfare [1]. The robust performances of emitter identification become more important in complex environments of emitters and landscapes. In the conventional approach of emitter identification, the key features of radar patterns such as the sequences of radar frequencies (RFs), time of arrivals (TOAs), and pulse widths (PWs) are used to extract the emitter parameters and these parameters are compared with tabulated emitter parameters. However, this process usually requires high computational complexity and needs to be verified by human experts. As the methods of automatic classification of radar patterns, the neural network based models [2–6] are adopted. In these methods, the discriminant functions as the output functions of neural networks are used for the decision of emitter types and they do not provide the conditional probabilities of emitter types for the given radar patterns. Consequently, these methods are hard to provide the systematic procedure of discriminating the known (or trained) and unknown (or untrained) emitter types which plays an important role in the electronic warfare. In this respect, an approach of automatic classification of radar patterns is investigated to obtain the conditional class probability for the given radar pattern. There are various ways of implementing pattern classifiers. The most popular way is using the discriminant function whose value indicates the degree of confidence in the classification; that is, the decision of classification is made by selecting the class that has the greatest discriminant value. In this direction, the support vector machines (SVMs) [7] are widely used in many classification problems because they provide reliable performances by maximizing the margin between the positive and negative classes. However, more natural way of representing the degree of confidence for classification is using the conditional class probability for the given pattern. In this context, the class probability output network (CPON) [8] was proposed for the purpose of estimating the conditional class probability using the Beta distribution parameters. This method is implemented on the top of a classifier; that is, many-to-one nonlinear function such as the linear combination of kernel functions. Then, the classifier output is identified by Beta distribution parameters and the output of CPON; that is, the conditional class probability for the given pattern is calculated from the cumulative distribution function (CDF) of Beta distribution parameters. In this approach, the Beta distribution is selected for the identification of output distribution of a classifier because the class probability; that is, the normalized measure between 0 and 1, is estimated as an output of the classifier and the Beta distribution itself represents the proba-

* Corresponding author.
E-mail addresses: lktime@skku.edu (L.S. Kim), bhb0722@nate.com (H.B. Bae), rmkil@skku.edu (R.M. Kil), churlhee.jo@lignex1.com (C.H. Jo).

http://dx.doi.org/10.1016/j.neucom.2017.01.094
0925-2312/

bility (or conjugate prior) of Binomial distribution; that is, in our case, the probability that the given pattern belongs to the class. On the other hand, other density functions of exponential families such as Gaussian, Laplacian, exponential, Gamma, etc., have long tails on the positive side or both the positive and negative sides of data values, and these density functions are not well fitted to the probability distributions of classifiers. In this computation, the output of CPON represents the p-value for testing a specific class. For the final decision on the classification, the class which has the maximum conditional class probability is selected. As a result, the suggested CPON method is able to provide the consistent improvement of classification performances for the classifiers using discriminant functions alone. For the detailed descriptions of CPONs and CPON applications, refer to [8–10].

In this approach of emitter identification, the concept (or one-class) learning method is applied; that is, the selected features of radar patterns for a specific emitter type are used as the input to the classifier of many-to-one mapping nonlinear function and the output distribution is identified by Beta distribution parameters to obtain the p-value for testing a specific emitter type. As a result, the proposed method provides a normalized measure for testing a specific emitter type using the output of CPON. The proposed one-class learning using the CPON is favorable compared with multi-class learning methods from the viewpoint of computational complexity since each class only requires to train the data for the corresponding emitter type. In fact, the usual number of emitter types is very large (usually the order of 10^4) in many emitter identification problems. Furthermore, the proposed approach also provides a new way of identifying the trained and untrained (or unknown) emitter types which is usually not available in the conventional method of training algorithms. In this approach, an important process is a nonlinear mapping in such a way that the output data fit the Beta distribution. This is possible when the output data lie within the finite range and the data distribution is unimodal; that is, the distribution has one value that occurs with the greatest frequency. In many cases of classification problems, this is possible by controlling the kernel parameters of the nonlinear mapping function. For this purpose, the kernel width is adjusted in such a way that the output data fit the Beta distribution as much as possible.

The rest of this paper is organized as follows: in Section 2, the process of radar classification and the key features of emitter identification problems are described, Section 3 presents the radar classification system using the CPON, Section 4 presents a new method of identifying the trained and untrained emitter types, Section 5 shows simulation results for radar pattern classification, and finally, Section 6 presents the conclusion.

2. Key features for radar pattern classification

The proposed method is intended to identify radar patterns from various emitters. For the purpose of detecting reflected signals from the object, the radar patterns are composed of pulse packets and have many variations in the carrier frequencies of pulse packets called as the radar frequencies (RFs), the interarrival times of pulse packets called as the time of arrivals (TOAs), and the widths of pulse packets called as the pulse widths (PWs).

In this approach, it is assumed that the radar has the ability to monitor a region of the microwave spectrum and extract pulse patterns. The whole process of emitter identification (or radar pattern classification) is illustrated in Fig. 1. In this diagram, the feature extractor receives pulses from the microwave radar receiver and processes each pulse into feature values such as azimuth, elevation, intensity, frequency, and pulse width. These data are then stored and tagged with the time of arrival of the pulse. Then, the clustering block is grouping radar pulses into

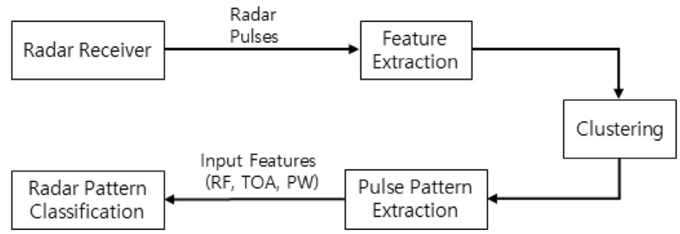


Fig. 1. The process of emitter identification: the key features of RFs, TOAs, and PWs are extracted from the received radar patterns and applied to the radar pattern classifier.

groups in which each group represents radar pulses from a single emitter. For each group of radar pulses, the pulse extraction block is analyzing the pulse repetition patterns of an emitter by using the information of time of arrivals. Finally, from the information of pulse repetition patterns, input features for the classifier are computed and the decision for the classification of emitter types is made by using extracted key features. In the proposed approach of emitter identification, the selected key features are the RFs, TOAs, and PWs. Then, for each sequence of m feature values $x_i,\ i = 1, 2, \ldots, m,$ the statistical measures such as the mean $\bar{x}$, variance s^2, skewness, and kurtosis are determined by

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i, \tag{1}$$

$$s^2 = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2, \tag{2}$$

$$\text{skewness} = \frac{\frac{1}{m}\sum_{i=1}^{m}(x_i - \bar{x})^3}{\left(\frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2\right)^{\frac{3}{2}}}, \quad \text{and} \tag{3}$$

$$\text{kurtosis} = \frac{\frac{1}{m}\sum_{i=1}^{m}(x_i - \bar{x})^4}{\left(\frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2\right)^2} - 3. \tag{4}$$

These statistical feature values are calculated for every sweep of receiving radar signals. Then, as a result, 12 feature values are used as the input to the classifier and the decision of emitter identification is made by using the CPON. For the distinction between less complicated feature distributions, the statistical measures up to the second moment; they are (1) and (2), might be favorable due to the required amount of data for the training of classifiers. However, in general, for the distinction between complicated feature distributions (for example, large numbers of emitter types with similar feature distributions), the statistical measures up to the fourth moment; they are (1) through (4), are more favorable for the classification of emitter types. In this approach, the distributions of these feature values are analyzed and the centroids of these feature values are used as the representation of emitter types.

3. Radar pattern classification system

In our problem of classifying radar patterns, the key features such as the RFs, TOAs, and PWs are extracted from a sweep of microwave radar receiver and four statistical measures for each feature are computed. For this process, twelve features are used as the input pattern to the radar pattern classifier. In our classification problem, the classification of radar patterns (or identification of emitter types) is assumed to be done at every radar sweep. The schematic diagram of the whole structure is illustrated in Fig. 2. Here, the automatic classification of radar patterns is constructed as follows: (1) from the features of 12 statistical measures, the distances between the estimated feature values and the previously learned centroids of feature values for emitter types are calculated, (2) the kernelized output for the calculated distances is passed to the input of the CPON, (3) the p-values for testing the hypotheses

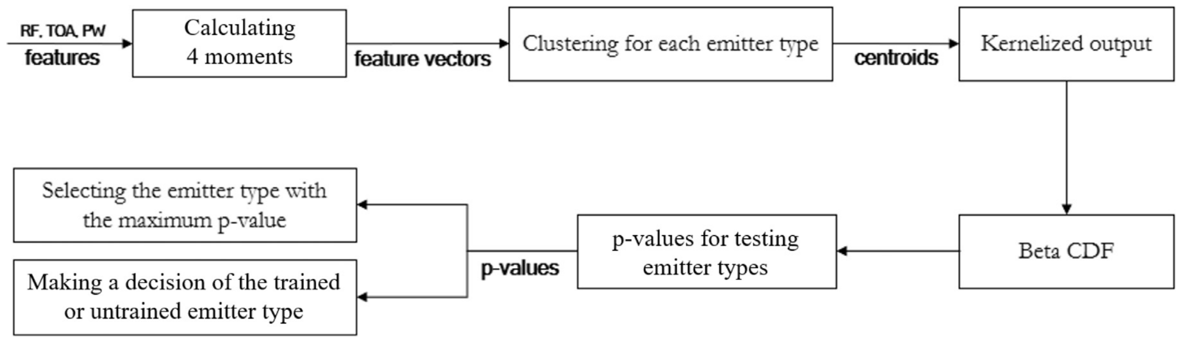


Fig. 2. Classification and identification of the trained and untrained emitter types using the CPON: from the kernelized output for the centroids of key feature vectors, Beta distribution parameters are estimated and used to identify the trained or untrained emitter types.

of emitter types are calculated using the Beta CDFs, (4) and then, the final decision is made by finding the emitter type with the maximum *p*-value among the *p*-values for testing emitter types.

In the construction of CPON for radar pattern classification, first, a centroid as the representative of the radar patterns is assigned for each emitter type. In our case, one centroid is enough for the representation of an emitter type using the key features of the RFs, TOAs, and PWs. Here, the centroids are determined by a clustering algorithm such as the LBG algorithm [11]. Then, the kernel functions are located at the positions of centroids; that is, for a d dimensional feature vector $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_d)$ of K emitter types, the kth kernel function $\phi_k(\tilde{\mathbf{x}})$ is determined by

$$\phi_k(\tilde{\mathbf{x}}) = \exp\left(-\alpha_k \sum_{i=1}^{d} \frac{(\tilde{x}_i - \mu_{ik})^2}{2\sigma_{ik}^2}\right), \; k = 1, 2, \ldots, K, \tag{5}$$

where μ_{ik} and σ_{ik} represent the mean and standard deviation of the ith dimension of the kth kernel function, respectively, and α_k represents the dispersion of the kth kernel function to control the shape of output distribution (in our case, the Beta distribution).

Here, the output of (5) is normalized between 0 and 1 by using the linear scale and the normalized classifier's output distribution is approximated by the Beta distribution parameters; that is, the normalized output for the kth classifier y_k is given by

$$y_k(\tilde{\mathbf{x}}) = \frac{\phi_k(\tilde{\mathbf{x}}) - L_k}{U_k - L_k}, \tag{6}$$

where L_k and U_k represents the lower and upper limit of the kth kernel function $\phi_k(\tilde{\mathbf{x}})$.

In the proposed CPON, the probability model represents the conjugate prior for the Binomial distribution; that is, in our case, the conditional class probability in binary classification problems. In this context, we consider the following Beta probability density function (PDF) of a random variable Y_k as the kth normalized classifier output:

$$f_{Y_k}(y_k|a_k, b_k) = \frac{1}{B_k(a_k, b_k)} y_k^{a_k-1}(1 - y_k)^{b_k-1}, \;\; 0 \le y_k \le 1, \tag{7}$$

where a_k and b_k represents the parameters of the Beta distribution for the kth classifier, and $B_k(a_k,\, b_k)$ represents a Beta function for the kth classifier defined by

$$B_k(a_k, b_k) = \int_0^1 y_k^{a_k-1}(1 - y_k)^{b_k-1} dy. \tag{8}$$

Then, the Beta CDF for the kth classifier is determined by

$$F_{Y_k}(y_k|a_k, b_k) = \frac{1}{B_k(a_k, b_k)} \int_0^{y_k} x^{a_k-1}(1 - x)^{b_k-1} dx. \tag{9}$$

One of the advantages of the Beta distribution is that the distribution parameters can be easily guessed from the mean $E[Y_k]$ and variance $Var(Y_k)$ as follows:

$$a_k = E[Y_k]\left(\frac{E[Y_k](1 - E[Y_k])}{Var(Y_k)} - 1\right) \tag{10}$$

and

$$b_k = (1 - E[Y_k])\left(\frac{E[Y_k](1 - E[Y_k])}{Var(Y_k)} - 1\right). \tag{11}$$

Although this moment matching (MM) method is simple, these estimators usually do not provide accurate estimations especially for a smaller number of data. For more accurate estimation of Beta parameters, the maximum likelihood estimation (MLE) or the simplex method for searching parameters [12] can be applied. If the data distribution follows a Beta distribution and the optimal Beta parameters are obtained, then the ideal cumulative distribution function (CDF) values of the data $u = F_Y(y)$ follow a uniform distribution because

$$f_U(u) = \frac{f_Y(y)}{|dF_Y/dy|} = \frac{f_Y(y)}{|f_Y(y)|} = 1. \tag{12}$$

To check whether the data distribution fits with the proposed Beta distribution, the Kolmogorov–Smirnov (K–S) test [13] for data distributions can be considered as follows:

- First, determine the distance D_n between n empirical and ideal CDF values:

  $$D_n = \sup_u |F_U^*(u) - F_U(u)|, \tag{13}$$

  where $F_U^*(u)$ and $F_U(u)$ represent the empirical and theoretical CDFs of $u = F_Y(y)$; that is, the CDF values of the normalized output of a classifier. In this case, $F_U(u) = u$ since the data $u = F_Y(y)$ follow an uniform distribution if the data y follows the presumed (or ideal) Beta distribution.
- Determine the p-value for testing the hypothesis of the Beta distribution:

  $$p\text{-value} = P(D_n \ge t/\sqrt{n}) = 1 - H(t), \tag{14}$$

  where $t = \sqrt{n} d_n$ (the value of a random variable D_n) and the CDF of the K–S statistic $H(t)$ is given by

  $$H(t) = \frac{\sqrt{2\pi}}{t} \sum_{i=1}^{\infty} e^{-(2i-1)^2\pi^2/(8t^2)}. \tag{15}$$

- Make a decision of accepting the hypothesis of beta distribution H_0 using the p-value according to the level of significance α: accept H_0, if p-value $\ge \alpha$; reject H_0, otherwise.

In this training of classifiers, the Beta distribution parameters as well as the kernel parameters are adjusted in such a way that the classifier's output distributions become closer to the ideal Beta distributions. For this purpose, the value of α_k in (5) is selected when the p-value of K–S test of (14) has the maximum value; that is, the kernel width is adjusted in such a way that the output data of (5) fit the Beta distribution as much as possible. The algorithm of constructing the CPON for radar pattern classification is described as follows:

For $k = 1, 2, \dots, K$ (number of emitter types), apply the following learning procedure:

Step 1. For n sequences of feature vectors of RFs, TOAs, and PWs generated from the kth emitter type, four statistical measures of (1) – (4) are calculated and obtain a sequence of n vectors as

$$\tilde{\mathbf{x}}_j = (\tilde{x}_{1j}, \tilde{x}_{2j}, \dots, \tilde{x}_{dj}), \quad j = 1, 2, \dots, n,$$

where $d = 12$ (4 statistical measures for each feature).

Step 2. Determine the centroid (sample mean) and sample variance for each emitter type:

$$\mu_{ik} = \frac{1}{n}\sum_{j=1}^{n} \tilde{x}_{ij}, \quad i = 1, 2, \dots, d, \quad \text{and}$$

$$\sigma_{ik}^2 = \frac{1}{n-1}\sum_{j=1}^{n} (\tilde{x}_{ij} - \mu_{ik})^2, \quad i = 1, 2, \dots, d.$$

Step 3. For $l = -2, -1, \dots, 2$, set $\alpha_k = 2^l$ and apply the following procedure:

(1) Determine the kernelized output of (5) for each emitter type; that is,

$$\phi_k(\tilde{\mathbf{x}}_j) = \exp\left(-\alpha_k \sum_{i=1}^{d} \frac{(\tilde{x}_{ij} - \mu_{ik})^2}{2\sigma_{ik}^2}\right), \quad j = 1, 2, \dots, n.$$

(2) Normalize the output value between 0 and 1 using the linear scale; that is,

$$y_k(\tilde{\mathbf{x}}_j) = \frac{\phi_k(\tilde{\mathbf{x}}_j) - L_k}{U_k - L_k}, \quad j = 1, 2, \dots, n,$$

where L_k and U_k represents the lower and upper limit of the kth kernel function $\phi_k(\tilde{\mathbf{x}})$.

(3) The distribution of normalized classifier output is identified by Beta distribution parameters using (10) and (11). For more accurate estimation method, the MLE method can be applied. For the detailed description of estimating parameters, refer to [8,9].

(4) Calculate the p-value of (14) for the output distribution of y_k.

Step 4. Determine the value of α_k corresponding to the maximum p-value in the step 3.

Step 5. Construct the CPON for the kth emitter type with the selected kernel function $\phi_k(\tilde{\mathbf{x}})$ and the corresponding Beta distribution parameters.

In this algorithm, the first process is to calculate statistical measures for each key feature using (1)–(4). The computational complexity of this process $O(mn)$ for n sequences. After the feature extraction process, the proposed algorithm is composed of finding centroids for a sequence of n feature vectors, determining the kernelized output data, and estimating Beta parameters. In this procedure, each process takes the computational complexity of $O(n)$ because one centroid is assigned for each class in this classification problem, the kernel center and width (variance) are determined by the centroid, and the simplest way of estimating Beta parameters is using the mean and variance of output data using (10) and (11). Even in the case of using the MLE method for Beta parameters, each iteration of searching the space of Beta parameters takes the computational complexity of $O(n)$ and the number of iterations is not proportional to n (usually much less than n); that is, the computational complexity of the MLE method is $O(n)$. In the case of K–S test for Beta distribution, the sorting of empirical Beta CDF values for kernelized output data is required and the computational complexity of this sorting is $O(n \log n)$ using the efficient sorting algorithm such as quick sort or heap sort. Therefore, the proposed algorithm takes the overall computational complexity of $O(mn)$; that is, the total number of data for n sequences since the value of m is usually much greater than the value of $\log n$. This fact is favorable for the fast learning of various emitter types since the computational complexity of the proposed algorithm only depends on the total number of data linearly.

After the CPON is trained, the classification of radar patterns can be determined by the Beta distribution for each class. First, for the received radar pattern, the normalized output y for each classifier is computed. Here, if the normalized output value is greater than 1, the output value is set as 1; on the other hand, if the value is less than 0, the output value is set as 0. Then, the conditional class probability is determined by the CPON output as the CDF value of the Beta distribution for the normalized classifier output.

For multi-class classification problems, the CPON can be constructed for each classifier output. Then, the following conditional probability for the kth class C_k; that is, the output of the kth CPON $F_k(y_k)$ for the kth normalized classifier output y_k is calculated as

$$F_k(y_k) = P(C_k | Y_k \le y_k) = F_{Y_k}(y_k), \tag{16}$$

where Y_k represents a random variable for the kth class C_k and $F_{Y_k}(y_k)$ represents its Beta CDF of (9). This output implies the p-value for testing hypotheses of the kth class C_k. Then, the final decision can be made by selecting the class with the maximum p-value; that is, for K classes, the selected class C_l is determined by

$$l = \arg\max_{1 \le k \le K} F_k(y_k). \tag{17}$$

From the above equation, the final decision of the type of emitter is made. This method provides an accurate estimation of conditional class probabilities since in this training of classifiers, the Beta distribution parameters as well as the kernel parameters are adjusted in such a way that the output distributions of classifiers become closer to the ideal Beta distributions. Furthermore, the suggested CPON method is able to provide the degree of uncertainty [9,10] for the final decision on classification by estimating the confidence intervals for the conditional class probabilities.

4. Identification of the trained and untrained emitter types

In the problems of emitter identification, the detection of untrained (or unknown) types of emitters is an important issue. For this issue of identifying untrained emitter types, the proposed classification model for emitter identification makes two cases of classification results after the decision making of (17):

1. The p-value of selected class of (17) is very small (for example, less than the usual level of significance = 0.05). In this case, it is highly probable that the received radar pattern belongs to the unknown emitter type.
2. The p-value of selected class of (17) is quite significant (for example, much larger than the usual level of significance = 0.05). In this case, it is not quite sure that the received radar pattern belongs to the known or unknown emitter type because the received radar pattern might be highly correlated with trained radar patterns.

In the first case, a new class can be easily found by checking the p-value of the selected class of (17). However, in the second case, the p-value does not provide a good measure of identifying a new class. From this context, a new method of analyzing the output distributions of CPONs is investigated. For this purpose, let the p-value of (16) be a random variable U. Then, U follows a uniform distribution over an interval (0, 1) due to (12). In this case, the mean of a random variable U is determined by 1/2. Here, the sample mean of n data U_i, $i = 1, \ldots, n$, is defined by

$$\bar{U} = \frac{1}{n}\sum_{i=1}^{n} U_i. \tag{18}$$

Then, this sample mean approximately follows a normal distribution; that is,

$$\bar{U} \dot{\sim} N\left(\frac{1}{2}, \frac{1}{12n}\right) \tag{19}$$

for large n by the central limit theorem (CLT). From this observation, with a probability of $1 - \alpha$,

$$|\bar{U} - \frac{1}{2}| \le z_{\alpha/2}\sqrt{\frac{1}{12n}}. \tag{20}$$

Another way of testing the uniform distribution can be performed by the K–S test. In this test, a sequence of p-values for the selected class U_i, $i = 1, 2, \ldots, n$ is collected and sorted in the ascending order; that is,

$$\tilde{U}_1 < \tilde{U}_2 < \ldots < \tilde{U}_n,$$

where $\tilde{U}_i$ s' are sorted random variables.

Then, the distance D_n between the empirical and ideal CDF values: that is,

$$D_n = \sup_i\left|\frac{i}{n+1} - \tilde{U}_i\right| \tag{21}$$

is calculated and the p-value for the K–S test of a uniform distribution is determined by (14).

In the case of the usual level of significance; that is, $\alpha = 0.05$, the value of z quantity is determined by $z_{\alpha/2} = z_{0.025} = 1.96$. From this property of p-values, the process of identifying new types of emitters can be described as follows:

Step 1. From a sequence of radar patterns, the CPON output values (p-values for testing emitter types) are calculated using (16) and determine the best matching emitter type using (17); that is, the emitter type corresponding to the maximum average p-value for the given sequence.

Step 2. From a sequence of p-values, determine the sample mean $\bar{U}$ and the distance D_n of (21).

Step 3. Determine whether the sample mean satisfies the condition of (20) and the p-value for the K–S test of a uniform distribution is greater than or equal to the level of significance α:

$$\bar{U} \in \left(\frac{1}{2} - \frac{z_{\alpha/2}}{\sqrt{12n}}, \frac{1}{2} + \frac{z_{\alpha/2}}{\sqrt{12n}}\right) \text{ and} \tag{22}$$

$$p\text{-value for the K-S test} \ge \alpha \tag{23}$$

Step 4. If the both of $\bar{U}$ and p-value for the K–S test do not satisfy the above condition, the given emitter can be treated as the untrained (or unknown) emitter type. Otherwise, there exists a high probability that the given emitter is the trained (or known) emitter type.

As an example of identifying new types of emitters, the sample means and the p-values for the K–S tests of a uniform distribution of the sequences of CPON output values were plotted for the trained and untrained emitter types which were generated from the emitter simulator developed by LIGNex1 as illustrated in Fig. 3. In this example, 20% out of 50 emitter types were untrained emitter types and the rest of them were trained using the CPON. As shown in this Figure, there are only trained (or known) emitter types in the region that the both conditions of (22) and (23) are satisfied. On the other hand, there are only untrained (or unknown) emitter types in the region that the both conditions of (22) and (23) are not satisfied. This implies that the sample mean and the p-value for the K–S test of a uniform distribution provide an effective guideline for identifying the trained and untrained emitter types. If one of two conditions is met, the confidence in identifying the trained or untrained emitter type might be reduced. In this work, the decision of the untrained emitter type is made when the given radar patterns do not satisfy both conditions to achieve the high confidence in the identification of untrained emitter types.

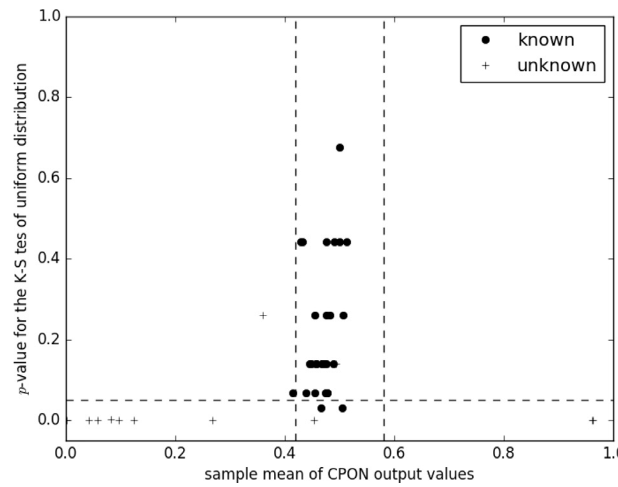


Fig. 3. Identification of the trained and untrained emitter types using the sample mean and the p-value for the K–S test of a uniform distribution: the trained emitter types surely lie within the region where the p-value for the K–S test is greater than 0.05 and the sample mean of CPON output values is within the confidence interval, and in the region of the opposite case, only the untrained emitter types exist.

5. Simulation

To demonstrate the effectiveness of the proposed method, the simulation for emitter identification (or radar pattern classification) was performed for the radar patterns generated from the emitter simulator which was developed by LIGNex1. This simulator was designed to accommodate the variety of key features such as the RFs, TOAs, and PWs of real emitter types. In this benchmark data set, there were 50 sets of emitter types containing the key features of the RFs, TOAs, and PWs and each data set included 100 sequences of radar patterns. For the evaluation of the proposed method, 10-fold evaluation method was used; that was, the given data set was split into ten parts, each one of ten classifiers was trained on permutations of nine out of ten parts, and the classifier was evaluated on the remaining tenth part as the test data. Then, for each emitter type, the following exact match ratio (EMR) was determined by

$$\text{EMR}_i = \frac{1}{100}\sum_{j=1}^{100} I(\mathbf{x}_{ij}), \quad i = 1, 2, \ldots, 50, \tag{24}$$

where $\mathbf{x}_{ij}$ represents the jth test pattern of the ith emitter type and $I(\mathbf{x}_{ij})$ represents the following indicator function:

$$I(\mathbf{x}_{ij}) = \begin{cases} 1 & \text{if } L(\mathbf{x}_{ij}) = D(\mathbf{x}_{ij}) \\ 0 & \text{otherwise.} \end{cases}$$

Here, $L(\mathbf{x}_{ij})$ and $D(\mathbf{x}_{ij})$ represent the class label and the decision label of the classifier for $\mathbf{x}_{ij}$, respectively; that is, the EMR represents the ratio of correct decision of the classifier. Then, the

Table 1
Simulation results for emitter identification when all types of emitters were trained.

(a) Average EMRs for the instances of radar patterns when two moments are calculated for each feature.				
Classifier	kNN	SVM_L	SVM_G	CPON
Average EMR	0.7670	0.9254	0.9310	**0.9432**
(b) Average EMRs for the instances of radar patterns when four moments are calculated for each feature.				
Classifier	kNN	SVM_L	SVM_G	CPON
Average EMR	0.6908	0.9376	0.9402	**0.9458**

average of EMR_i values for 50 emitter types were determined as the classification performance.

To compare the performances of the proposed CPON-based method, the k-nearest neighbor (kNN), SVM, and random forest (RF) classifiers were also trained for the same training data using the Scikit-learn package [14] and evaluated for the same test data. In this simulation, the same features of the RFs, TOAs, and PWs were also used for the training and testing the classifiers. For the multiclass classification of 50 emitter types, the kNN, SVM, and RF classifiers were trained as 50 binary (one-against-rest) classifiers whereas the proposed CPON classifier were trained as 50 one-class classifiers in which each CPON classifier was trained using the proposed algorithm. In this training of CPON classifiers, for each class, one centroid was assigned using the clustering algorithm in the feature space, one kernel function with the values of $\alpha_k = 1$ was assigned, and the corresponding Beta parameters were estimated using the MLE method. To compare with the effect of the number of features, 6 and 12 features were also used as the radar patterns in which 2 and 4 moments for each key feature were calculated, respectively.

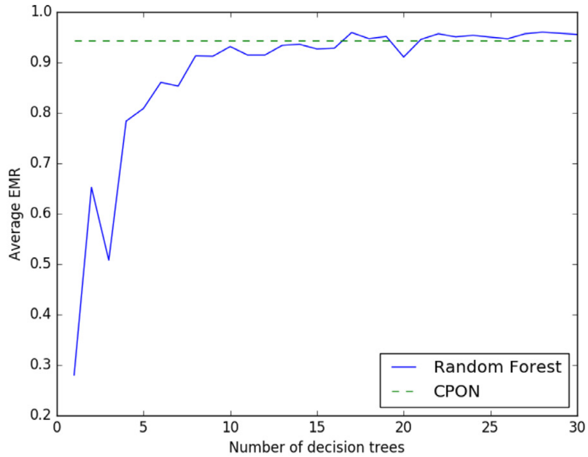
For the training of the kNN and SVM classifiers, the effective ranges of learning parameters were searched to find the best classification performance; that is, outside of this range, there was no effect on the performance improvement. In the case of kNN classifiers, the values of k nearest neighborhood were searched as $k = 2^i,\ i = 0, 1, \ldots, 5$ to find the best classification performance in the test data. For the neighborhood index, the Euclidean distance in the feature space was applied. As a result, the best classification performances were obtained as $k = 1$ and $k = 2^5$ for 6 and 12 features, respectively. For the training of SVM classifiers, two types of kernels such as the linear and Gaussian kernels were selected and proper learning parameters of C (the control parameter of the summation of slack variables) and γ (the kernel parameter of Gaussian function) were searched to find the best classification performance in the test data. In the case of SVM with linear kernels, the values of C were searched as $C = 10^i,\ i = 0, 1, \ldots, 4$ whereas in the case of SVM with Gaussian kernels, the values of C and γ were searched as $C = 10^i,\ i = 0, 1, \ldots, 4$ and $\gamma = 2^i,\ i = -4, -3, \ldots, 4$, respectively. As a result, the best classification performances were obtained as $C = 10^4$ in the case of SVM with linear kernels for both 6 and 12 features, and $(C = 10^4, \gamma = 2^2)$ and $(C = 10^4, \gamma = 2^{-1})$ in the case of SVM with Gaussian kernels for 6 and 12 features, respectively. The simulation results for emitter identification using the kNN, SVM with linear kernel (SVM_L), SVM with Gaussian kernel (SVM_G), and CPON classifiers were summarized in Table 1. These simulation results have shown that (1) the proposed CPON-based method provided the best performance in the average EMR measures compared with other classifiers in Table 1 and (2) in the case of 6 and 12 features, the SVM and CPON classifiers provided improved performances of average EMR measures for 12 features while the kNN classifier provided the improved performance for 6 features.

Table 2
The best classification performances of RFs when the number of decision trees is increased up to 30.

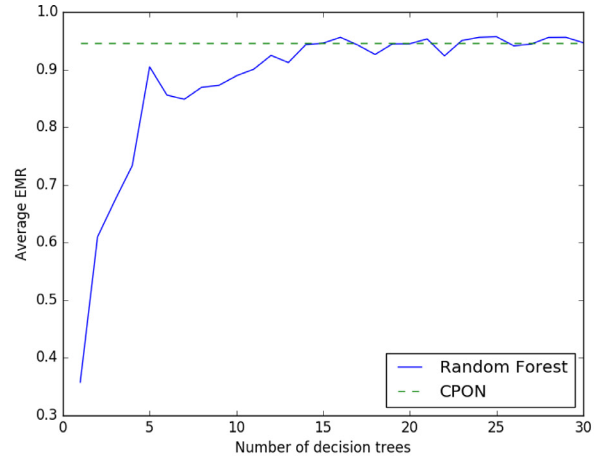
Classifier	RF		CPON	
Average EMR	**0.9600**	**0.9570**	0.9432	0.9458
No. of estimators for each class	28	25	1	1
No. of moments for each feature	2	4	2	4

In the case of RF classifiers, the depth of binary decision trees was set as 6; that was, up to 127 feature nodes and the number of binary decision trees was increased up to 30. The simulation results for emitter identification using the RF classifiers were illustrated in Fig. 4. The best classification performances of RF classifiers were also summarized in Table 2. These simulation results have shown that (1) the RFs provided saturated classification performances for more than or equal to 14 decision trees and (2) the performances of CPON classifiers with 6 and 12 features were equivalent to the performances of RF classifiers with 17 and 14 decision trees, respectively. In this simulation, the RF classifiers demonstrated the improved performance compared with CPON classifiers using the concept of committee machine in which the variance of estimation model was reduced by combining multiple decision trees. In the case of CPON classifiers, one-class learning with one kernel function for each class was applied. This is favorable for fast on-line learning which is one of the important requirements in many emitter identification problems. Furthermore, the CPON classifier provides the conditional class probability for the given pattern with high classification performance. In the proposed method, the improvement of classification performance is also possible using the deep structure of CPON classifiers [10]. From these simulation results, it was demonstrated that the proposed statistical features of the RFs, TOAs, and PWs were quite effective to identify the characteristics of the emitter types and the proposed CPON-based classification was also an effective approach to emitter identification problems.

The proposed CPON-based method is also able to provide p-values for testing the types of emitters. In practice, this information of p-values helps us to make a decision whether the received radar pattern is a new type of emitter or one of the known types of emitters. For example, if the maximum p-value is less than some threshold value (the usual value is 0.05), then there is a high probability that the received radar pattern comes from a new type of emitter. Another method is using the criteria of (22) and (23) for the sequences of radar patterns. In this work, the given radar pattern was identified as the untrained emitter type when the both conditions of (22) and (23) were not satisfied. This ability of finding a new type of emitter is also an important issue in emitter identification problems. For the performance measures of detecting untrained (or new) types of emitters, the following accuracy, precision, recall (or detectability), and F_1 measures were calculated:



(a) The average EMR measures of RF classifiers with 6 features



(b) The average EMR measures of RF classifiers with 12 features

Fig. 4. The classification performances of RF classifiers as the number of decision trees is increasing: (a) and (b) represents the average EMR measures when two moments and four moments are calculated for each feature, respectively.

Table 3
Simulation results for emitter identification when there are unknown (or untrained) emitter types.

(a) Performances of detecting new types of emitters when an instance of radar pattern is applied to the CPON classifier.				
Percentage of unknowns	Acc	p_{new}	r_{new}	F_1
10	**0.9144**	**0.9480**	**0.9600**	**0.9512**
20	0.8996	0.9420	0.9339	0.9377
30	0.8604	0.9394	0.8720	0.9042
40	0.8514	0.9374	0.8409	0.8859
50	0.8460	0.9338	0.7948	0.8576
(b) Performances of detecting new types of emitters when a sequence of radar patterns is applied to the CPON classifier.				
Percentage of unknowns	Acc	p_{new}	r_{new}	F_1
10	**0.9580**	0.9618	**0.9910**	**0.9761**
20	0.9560	0.9625	0.9826	0.9723
30	0.9320	0.9629	0.9419	0.9521
40	0.9340	0.9671	0.9294	0.9473
50	0.9520	**0.9840**	0.9256	0.9529

- The accuracy measure Acc is defined by

$$Acc = \frac{t_u + t_k}{t_u + t_k + f_u + f_k}, \tag{25}$$

where t_u and t_k represent the numbers of correctly detected unknown and known emitter types, respectively, and f_u and f_k represent the numbers of falsely detected unknown and known emitter types, respectively.
- The precision measure p_{new} is defined by

$$p_{new} = \frac{t_u}{t_u + f_u}. \tag{26}$$

- The recall (or detectability) measure r_{new} is defined by

$$r_{new} = \frac{t_u}{t_u + f_k}. \tag{27}$$

- Finally, the F_1 measure, which is a trade-off between the precision and recall; that is,

$$F_1 = \frac{2 p_{new} r_{new}}{p_{new} + r_{new}}. \tag{28}$$

The simulation for identifying the trained and untrained emitter types was performed using the CPON-based method and the results of the accuracy, precision, recall, and F_1 measures were summarized in Table 3. In this simulation of identifying unknown types of emitters for the sequences of radar patterns, the differences in time of arrivals (dTOAs) were used instead of TOAs since the dTOAs provided the better fit of Beta distribution for classifier's output data. For the evaluation of this simulation, the first half of the data (50 sequences) of each data set were used as the training data and the remaining 50 sequences of data were used as the test data. Then, the train and test data were interchanged, and the performance measures were collected. In this data set, the untrained emitter types were randomly selected 10 times and the average performance measures were collected. These simulation results have shown that (1) the identification method of using a sequence of radar patterns provided the better performances in all performance measures than the identification method of using an instance of a radar pattern in which the maximum p-value was used for the decision of identifying new types of emitters and (2) the accuracy and F_1 measures in both methods provided the best performances when the number of untrained emitter types was the smallest; that was, 10. This is due to the fact that the discriminating power of identifying untrained emitter types has the tendency to be improved as the number of trained emitter types is increasing. To observe the effect of sample size for detecting new types of emitters, the first 10, 20, 30, 40, and 50 radar patterns in the test data were used as the sequences of radar patterns when there were 10% and 20% of emitter types were untrained. The results of average performance measures were summarized in Table 4. These simulation results have shown that average EMRs and F_1 measures provided the best performances when the sample size was 30. This implies that the sequence of

Table 4
The effect of sample size for detecting new types of emitters when a sequence of radar patterns is applied to the CPON classifier.

(a) Performances of detecting new types of emitters when 10% of emitter types are untrained.

Sample size	*Acc*	p_{new}	r_{new}	F_1
10	0.9280	0.9618	0.9578	0.9596
20	0.8860	0.8851	0.9853	0.9324
30	**0.9740**	**0.9822**	0.9890	**0.9855**
40	0.8660	0.8578	0.9899	0.9190
50	0.9580	0.9618	**0.9910**	0.9761

(b) Performances of detecting new types of emitters when 20% of emitter types are untrained.

Sample size	*Acc*	p_{new}	r_{new}	F_1
10	0.9180	0.9650	0.9353	0.9497
20	0.8900	0.8850	0.9757	0.9279
30	**0.9700**	**0.9850**	0.9781	**0.9814**
40	0.8740	0.8600	0.9805	0.9160
50	0.9560	0.9625	**0.9826**	0.9723

the first 30 radar patterns can be a good representative to identify the distribution of radar patterns of these emitter types.

In summary, the proposed CPON-based method provides an effective way of classifying radar patterns by calculating the p-values for testing emitter types. As a result, the proposed method achieves higher classification performances than other methods using discriminant functions. For the training of the proposed CPON, one-class learning is applied for each emitter type and the computational complexity of the proposed learning algorithm is proportional to the number of total data. This fact is favorable for the fast on-line learning of each emitter type. Furthermore, the proposed method also provides an efficient way of detecting new types of emitters using the hypothesis test of a uniform distribution of CPON output.

6. Conclusion

A new method of radar pattern classification was proposed based on the class probability output network (CPON). In the proposed method, the sequences of key features such as the frequencies, time of arrivals, and pulse widths of emitting signals are analyzed and statistical measures of these features such as the mean, variance, skewness, and kurtosis are extracted, and used as the input to the CPON. Then, the CPON is used to construct a hypothesis of specific emitter type from the distributions of these features. As a result, the proposed CPON provides the p-values for testing the hypotheses of emitter types. Through the simulation for radar pattern classification, it has been demonstrated that the proposed method provides better performance of classification than the other classifiers using discriminant functions. For the training of the proposed CPON, one-class learning is applied for each emitter type and the computational complexity of the proposed learning algorithm is proportional to the number of total data. This fact is favorable for fast on-line learning of each emitter type. Furthermore, the proposed CPON-based method is able to provide the systematic procedure of making a decision whether the received radar pattern comes from a new type of emitter using the hypothesis test of a uniform distribution of CPON output. This ability of finding new types of emitters is also an important issue in emitter identification problems.

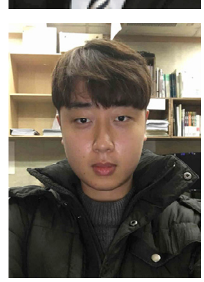
Acknowledgment

This work was supported in part by the LIGNex1 contract research program (Project No. LIGNEX1-2015-0079).

References

[1] D. Schleher, Introduction to electronic warfare, Report, Eaton Corp., AIL Div., Deer Park, NY, 1986.
[2] J. Anderson, M. Gately, P. Penz, D. Collins, Radar signal categorization using a neural network, Proc. IEEE 78 (10) (1990) 1646–1657.
[3] C.-S. Shieh, C.-T. Lin, A vector neural network for emitter identification, IEEE Trans. Antennas Propag. 50 (8) (2002) 1120–1127.
[4] Z. Wu, Z. Yang, Z. Yin, L. Zuo, H. Gao, A novel RBF neural network for radar emitter recognition based on rough sets, J. Chin. Inst. Eng. 35 (7) (2012) 901–907.
[5] N. Petrov, I. Jordanov, J. Roe, Radar emitter signals recognition and classification with feedforward networks, Procedia Comput. Sci. 22 (2013) 1192–1200.
[6] C.-M. Lin, Y.-M. Chen, C.-S. Hsueh, A self-organizing interval type-2 fuzzy neural network for radar emitter identification, Int. J. Fuzzy Syst. 16 (1) (2014) 20–30.
[7] V. Vapnik, Statistical Learning Theory, John Wiley and Sons, 1998.
[8] W. Park, R. Kil, Pattern classification with class probability output network, IEEE Trans. Neural Netw. 20 (10) (2009) 1659–1673.
[9] H. Rosas, R. Kil, S. Han, Automatic media data rating based on class probability output networks, IEEE Trans. Consum. Electron. 56 (4) (2010) 2296–2302.
[10] S. Kim, Z. Yu, R. Kil, M. Lee, Deep learning of support vector machines with class probability output networks, Neural Netw. 64 (2015) 19–28.
[11] Y. Linde, A. Buzo, R. Gray, An algorithm for vector quantizer design, IEEE Trans. Commun. 28 (1980) 84–95.
[12] S. AbouRizk, D. Halpin, J. Wilson, Fitting beta distributions based on sample data, J. Constr. Eng. Manag. 120 (2) (1994) 288–305.
[13] V. Rohatgi, A. Saleh, Nonparametric Statistical Inference, John Wiley and Sons, 2015, pp. 575–636.
[14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: machine learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830.

Lee Suk Kim received the B.S. Degree in Computer engineering from MyungJi University, Youngin, Korea in 2014, and the M.S. Degree in Electrical and Computer Engineering from Sungkyunkwan University, Suwon, Korea in 2016. He worked on a project for LIGnex1 during the M.S. Course. Currently, he plans to have a further study of research areas such as the machine learning, text recognition, and pattern recognition.



Han Bin Bae received the B.S. Degree in Computer engineering from Sungkyunkwan University, Suwon, Korea, 2016. He worked on a project for Electronics and Telecommunications Research Institute (ETRI) during the B.S. Course. Currently, he studies for the M.S. in Electrical and Computer Engineering, Sungkyunkwan University. He plans to have a further study of research areas such as the automatic speech recognition, pattern classification, machine learning, and deep learning.

Rhee Man Kil received the B.S. Degree in electrical engineering from Seoul National University, Seoul, Korea, in 1979 and the M.S. and Ph.D. Degrees in computer engineering from the University of Southern California, Los Angeles, in 1985 and 1991, respectively. He was with the Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea. Currently, he is a Professor at the Department of Electrical and Computer engineering, Sungkyunkwan University (SKKU), Suwon, Gyeonggi-do, 16419, Korea. His general research interests lie in the theories and applications of machine learning. His current interests focus on the areas of pattern classification, model selection in regression problems, time series prediction, data mining, and noise-robust speech recognition.

Churl Hee Jo received the M.S. Degree in Electronics Engineering from Kangwon National University, Korea in 1994. His major was the image processing by fractal theory. Since 1997, he has been with the radar signal processing projects in the R&D Center, LIGNex1. His current research interest focuses on the self-recognition of signal identification.