# A WEIGHTED FEATURE REDUCTION METHOD FOR POWER SPECTRA OF RADAR HRRPS

*Lan Du, Hongwei Liu, Junying Zhang and Zheng Bao*

National Lab. of Radar Signal Processing, Xidian University
Xi'an, Shaanxi, China
phone: + (086) 02988202348-610, fax: + (086) 02988236159, email: dulan@mail.xidian.edu.cn
r

## ABSTRACT

Feature reduction is an important stage in pattern recognition. This paper deals with the feature reduction methods for a time-shift invariant feature, power spectrum, in radar automatic target recognition using high-resolution range profiles (HRRPs). Several existing feature reduction methods in pattern recognition are analyzed, and a weighted feature reduction method based on Fisher's discriminant ratio (FDR) is proposed. According to the characteristics of radar HRRP target recognition, the proposed weighted feature reduction method uses an iterative algorithm to search for the optimal weight vector for power spectra of HRRPs, and thus reduces feature dimensionality. Compared with the method of using the raw power spectra and some existing feature reduction methods, the weighted feature reduction method can not only reduce feature dimensionality, but also improve recognition performance with low computation complexity. In the recognition experiments based on measured data, the proposed method is robust to different test data and achieves good recognition results.

## 1. INTRODUCTION

Statistical pattern recognition may be subdivided into four phases, namely, data acquisition, data preprocessing, feature reduction and classification. Usually raw features obtained in the first two phases are in a higher dimensional feature space. Some salient features are extracted from the raw features in the feature reduction phase. Thereby feature reduction is an important approach to delete redundancy and decrease computation.

High-resolution range profiles (HRRPs) contain the target structure signatures, such as target size, scatterer distribution, etc., thereby radar HRRP target recognition has received intensive attention from the radar automatic target recognition community[1~3]. However, HRRPs are sensitive to target-aspect, amplitude-scale and time-shift[1~3]. The reference[3] discussed a set of time-shift invariant features, higher-order spectra, of HRRP in detail, which showed that power spectrum (the 1$^{st}$-order spectrum) has the best recognition performance among the higher-order spectra.

This paper deals with the feature reduction methods for the power spectra of HRRPs. Several existing feature reduction methods are analyzed, and a weighted feature reduction method based on Fisher's discriminant ratio (FDR) is proposed. Compared with the method of using the raw power spectra and some existing feature reduction methods, the proposed feature reduction method not only is robust to different test data, but also improves the recognition performance with low computation complexity in the recognition experiments based on measured data.

## 2. SOME EXISTING FEATURE REDUCTION METHODS

Fisher's linear discriminant is a class separability measure, which has two forms, FDR[4] and Fisher's optimal projection subspace[2,4]. FDR is applied to independent feature components, and Fisher's optimal projection subspace is applied to correlated feature components.

### 2.1 Direct Feature Reduction Method based on Fisher's Discriminant Ratio (FDR)

Let there be $c$ targets, and let $\{\boldsymbol{x}_{ik} \mid k = 1,2,\cdots,K\}$ represent all $N$ -dimensional features of target $\omega_i$ ( $i = 1,2,\cdots,c$ ), where $k$ denotes the sample number. FDR of the $n^{\text{th}}$ feature component is defined as

$$FDR(n) = \sum_{i=1}^{c}\sum_{j\neq i}^{c}\frac{\left(\mu_i(n)-\mu_j(n)\right)^2}{\sigma_i^2(n)+\sigma_j^2(n)}, \quad n = 1,2,\cdots,N \tag{1}$$

where $\mu_i(n)$ and $\sigma_i^2(n)$ represent the mean and the variance of the $n^{\text{th}}$ feature component of target $\omega_i$, respectively.

$$\mu_i(n) = \frac{1}{K}\sum_{k=1}^{K} x_{ik}(n),\ n = 1,2,\cdots,N\ ;\ i = 1,2,\cdots,c \tag{2}$$

$$\sigma_i^2(n) = \frac{1}{K-1}\sum_{k=1}^{K}\left(x_{ik}(n)-\mu_i(n)\right)^2,\ n = 1,2,\cdots,N\ ;\quad i = 1,2,\cdots,c \tag{3}$$

The weights can be defined as

$$w(n) = \begin{cases} 1 \quad , \quad n_s = n & \dfrac{FDR(n)}{\max\limits_n\{FDR(n)\}} \geq \text{Th} \\ & , \\ 0 \quad , \quad n_u = n & \dfrac{FDR(n)}{\max\limits_n\{FDR(n)\}} < \text{Th} \end{cases}, \quad n = 1,2,\cdots,N \tag{4}$$

where Th denotes the threshold. The corresponding reduced feature vector is

$$\chi_{ik}(r) = w(n_s) \cdot x_{ik}(n_s)\,,\ r = 1, 2, \cdots, d\,;\quad i = 1, 2, \cdots, c\,;\ k = 1, 2, \cdots, K \tag{5}$$

where

$$d = \mathrm{L}\{w(n) \neq 0 \mid n = 1, 2, \cdots, N\} \tag{6}$$

$d$ denotes the number of nonzero elements of weight vector $\boldsymbol{w}$. Thus the reduced feature vector is a $d$-dimensional vector.

### 2.2 Feature Reduction Method based on Fisher's Optimal Projection Subspace

Fisher's optimal projection subspace is to seek a subspace $\boldsymbol{\Omega}$=Span$(\boldsymbol{U}) \in R^{N\times d}$ to maximize the class separability, which can satisfy

$$\max_{\boldsymbol{U}} J(\boldsymbol{U}) = \frac{\prod_{\text{diag}} \boldsymbol{U}^T \boldsymbol{S}_b \boldsymbol{U}}{\prod_{\text{diag}} \boldsymbol{U}^T \boldsymbol{S}_w \boldsymbol{U}} \tag{7}$$

where $\prod_{\text{diag}} \boldsymbol{A}$ denotes the product of all diagonal elements of the matrix $\boldsymbol{A}$, $\boldsymbol{S}_w$ and $\boldsymbol{S}_b$ represent the average within-class scatter matrix over all targets and the between-class scatter matrix, respectively.

$$\boldsymbol{S}_w = \sum_{i=1}^{c} P_i \boldsymbol{S}_i = \sum_{i=1}^{c} P_i \left[ \sum_{k=1}^{K} (\boldsymbol{x}_{ik} - \boldsymbol{m}_i)(\boldsymbol{x}_{ik} - \boldsymbol{m}_i)^T \right] \tag{8}$$

$$\begin{aligned} \boldsymbol{S}_b &= \sum_{i=1}^{c} P_i (\boldsymbol{m}_i - \boldsymbol{m})(\boldsymbol{m}_i - \boldsymbol{m})^T \\ &= \sum_{i=1}^{c} P_i \left( \boldsymbol{m}_i - \sum_{i=1}^{c} P_i \boldsymbol{m}_i \right) \left( \boldsymbol{m}_i - \sum_{i=1}^{c} P_i \boldsymbol{m}_i \right)^T \end{aligned} \tag{9}$$

where $\boldsymbol{m}_i$ denotes the clustering center of target $\omega_i$, $\boldsymbol{m}$ denotes the clustering center of all targets, and $P_i$ denotes the priori probability of target $\omega_i$. By derivation[2,4], this nonlinear optimization problem in Eq. (7) equals to a generalized eigenvalue decomposition problem of the matrix pencil $\{\boldsymbol{S}_b, \boldsymbol{S}_w\}$,

$$\boldsymbol{S}_b \boldsymbol{u}_j = \lambda_j \boldsymbol{S}_w \boldsymbol{u}_j\,,\quad j = 1, 2, \cdots, d\,,\quad d < N \tag{10}$$

where $\boldsymbol{u}_j$, the $j^{\text{th}}$ column vector of the matrix $\boldsymbol{U}$, is the eigenvector corresponding to the generalized eigenvalue $\lambda_j$ ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d \geq \cdots \geq \lambda_N$) of $\{\boldsymbol{S}_b, \boldsymbol{S}_w\}$. The corresponding $d$-dimensional feature vector is

$$\boldsymbol{\chi}_{ik} = \boldsymbol{U}^T \boldsymbol{x}_{ik}\,,\quad i = 1, 2, \cdots, c\,;\quad k = 1, 2, \cdots, K \tag{11}$$

Thus the reduced feature vector is acquired from linear change of the raw feature.

## 3. A WEIGHTED FEATURE REDUCTION METHOD BASED ON FISHER'S LINEAR DISCRIMINANT

Compared with ordinary pattern recognition, several issues must be considered for radar HRRP target recognition. Firstly, time-shift invariant features should be extracted as the raw features due to the time-shift sensitivity of HRRP. Secondly, time-shift invariant features (such as higher-order spectra, central moments and so on) of HRRP are also sensitive to target-aspect variation. Thereby, the HRRPs' features of every target can be divided into many subsets according to different angular sectors, which are defined as feature sets of HRRP frames. Although every target has many feature sets in radar HRRP target recognition, the recognition result mainly depends on comparison between the between-class nearest patterns. Therefore, when computing weight vector or optimal projection subspace by means of various feature reduction methods, we need to use the between-class nearest sample sets to calculate the related weight vector or optimal projection subspace, and then average them to get the final result.

The direct feature reduction method in Section 2.1 is to select the feature components with larger FDRs to reduce feature dimensionality. In other words, the weights used in this method are 1 or 0. Considering that most classification algorithms use the similarity measure based on Euclidean distances, the Euclidean distances between samples will change, if raw feature components are multiplied by different weights. Therefore, a weighted feature reduction method based on FDR is proposed in this paper, in which feature components are multiplied by weights $w(n) \in [0,1]$ ($n = 1, 2, \cdots, N$) to increase rationality of the contribution of feature components to Euclidean distance calculation. The optimal weight vector is sought by an iterative algorithm to make the FDRs of feature components converge. The algorithm is as follows.

(1) Initialization: Let $p = 0$. Here the initialized feature vectors are the power spectra. Let $\boldsymbol{\chi}_{ikl}^{(0)} = \boldsymbol{x}_{ikl}$ ($i = 1, 2, \cdots, c$; $k = 1, 2, \cdots, K$; $l = 1, 2, \cdots, L$), where $i$ denotes target class, $k$ denotes frame number and $l$ denotes feature vector number in a frame;

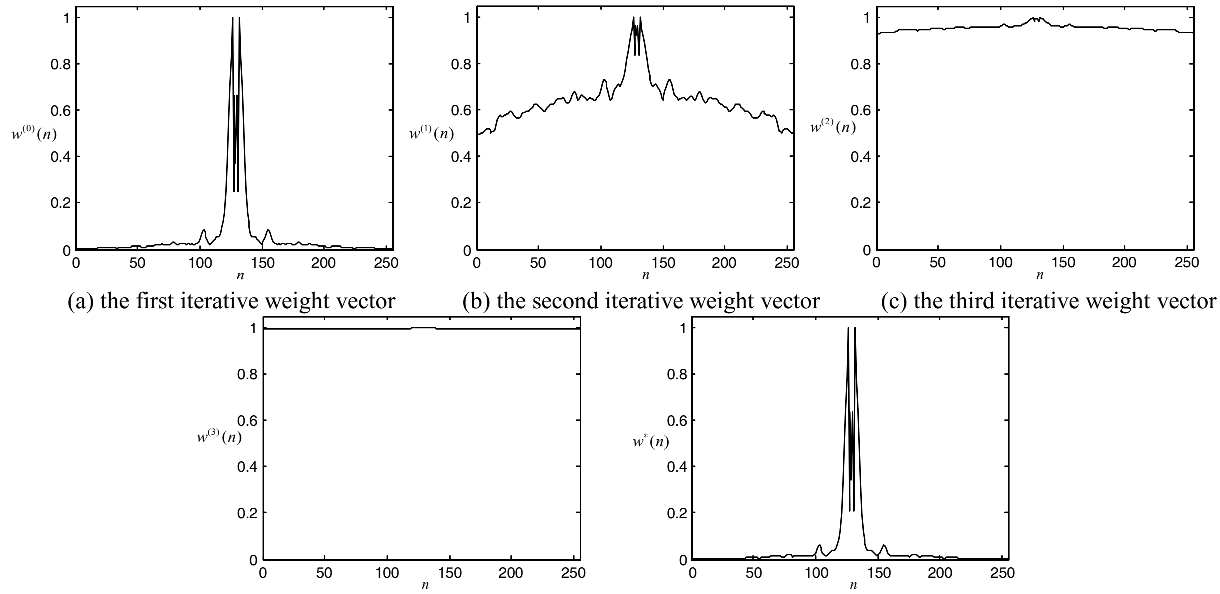
(2) Iterative Process: The between-class nearest feature sets of HRRPs frames are searched among the clustering centers. Let $\{\boldsymbol{\chi}_{ikl}^{(p)} \mid l = 1, 2, \cdots, L\} \to \{\boldsymbol{\chi}_{jm_j(i,k)l}^{(p)} \mid l = 1, 2, \cdots, L\}$ ($i, j = 1, 2, \cdots, c$, $i \neq j$, $k = 1, 2, \cdots, K$), where the $m_j(i,k)$-th frame of the $j$-th target is nearest to the $k$-th frame of the $i$-th target. Thus there are $c \cdot K$ between-class nearest feature sets in all. If Eq. (1) is used to calculate the FDR vectors of all between-class nearest feature sets $\{\boldsymbol{FDR}_q \mid q = 1, 2, \cdots, c\cdot K\}$, the average FDR vector is

$$\boldsymbol{FDR}_m = \frac{1}{c \cdot K} \sum_{q=1}^{c\cdot K} \boldsymbol{FDR}_q \tag{12}$$

Define the $p^{\text{th}}$ iterative weight vector as

$$\boldsymbol{w}^{(p)} = \frac{\boldsymbol{FDR}_m}{\max_n \{FDR_m(n)\}} \tag{13}$$

(3) Stop Condition: If the weights of feature components converge, the iterative process ends. Here variance is used to measure the convergence. Let $\varepsilon > 0$. The variance of weight vector $\boldsymbol{w}^{(p)}$ is



(a) the first iterative weight vector (b) the second iterative weight vector (c) the third iterative weight vector

(d) the fourth iterative weight vector (e) the optimal weight vector

Fig. 1 the iterative weight vectors and the optimal weight vector for power spectra of measured HRRPs using the iterative algorithm based on FDR

$$\sigma^2_{\boldsymbol{w}^{(p)}} = \frac{1}{N-1}\sum_{n=1}^{N}\left( w^{(p)}(n) - \frac{1}{N}\sum_{n=1}^{N} w^{(p)}(n) \right)^2 \tag{14}$$

If $\sigma^2_{\boldsymbol{w}^{(p)}} > \varepsilon$, then

$$\boldsymbol{\chi}_{ikl}^{(p+1)} = \boldsymbol{w}^{(p)} \bullet \boldsymbol{\chi}_{ikl}^{(p)} \tag{15}$$

where " $\bullet$ " denotes dot production between vectors. Let $p = p+1$, and the iterative process continue. Otherwise, the weights converge, thus let the iterative process end.

(4) Feature Reduction: Let the optimal weight vector be

$$\boldsymbol{w}^* = \boldsymbol{w}^{(0)} \bullet \boldsymbol{w}^{(1)} \bullet \cdots \bullet \boldsymbol{w}^{(p)} \tag{16}$$

The weights can be defined as

$$w(n) = \begin{cases} w^*(n) \quad , \quad n_s = n \, , \; w^*(n) \geq \mathrm{Th} \\ \;\; 0 \qquad , \quad n_u = n \, , \; w^*(n) < \mathrm{Th} \end{cases} , \qquad n = 1, 2, \cdots, N \tag{17}$$

Substituting $\boldsymbol{w}^*$ into Eq. (5) gives the reduced feature vectors $\{ \chi_{ikl} \mid i = 1, 2, \cdots, c; k = 1, 2, \cdots, K; l = 1, 2, \cdots, L \}$.

A fact worth emphasizing is that the FDR remains unchanged whether the feature components are weighted or not, according to Eq. (1). However, the between-class nearest sets will change when the features are weighted, which leads to different FDR. This is one of the reasons that why the feature reduction method proposed in this paper has better recognition performance.

## 4. RECOGNITION EXPERIMENTS BASED ON MEASURED DATA

The recognition experiments performed here are based on airplane data measured by radar with the center frequency of 5520 $MHz$ and the bandwidth of 400 $MHz$. Training data cover almost all of the target-aspect angles of test data, but their elevation angles are different.

When $p = 3$, the weight vector $\boldsymbol{w}^{(3)}$ converges in the experiment. Fig. 1 shows the corresponding $\boldsymbol{w}^{(0)} \sim \boldsymbol{w}^{(3)}$ and the optimal weight vector $\boldsymbol{w}^*$, and they include all of the weights corresponding to positive and negative frequency components of power spectra. Actually only a half of feature components and weights, namely, the ones corresponding to positive or negative frequency components, are applied to recognition. The weight vectors of the direct feature reduction method based on FDR and of the proposed weighted feature reduction method are shown in Figs. 1 (a) and (e), respectively. From the waveforms of weight vectors, we can see the feature components corresponding to lower frequency components have larger class separability than those corresponding to higher frequency components. For a comparison in recognition performance between the feature reduction methods, 18-dimensional feature vectors (reduced from the 128-dimensional power spectra) are used in the recognition experiments.

Template matching method under the maximum correlation coefficient criterion (MCC-TMM)[1,3] is used to compare the recognition performances of the feature reduction methods. The recognition rates of the raw power spectra and the three kinds of reduced feature vectors are shown in Table I. Obviously, the weighted feature reduction method has the advantage over the method of directly using FDR to reduce feature dimensionality of better recognition performance with the same computation complexity. The recognition rate of An-26 by reduced feature vectors based on Fisher's optimal projection subspace is lower than that by raw power spectra. The reason is that An-26 is a propeller-driven aircraft and its HRRPs which are modulated by propellers scatter widely in the feature space. It is difficult to seek an optimal projection subspace which can divide all features of Yark-42, Cessna Citation S/ II and An-26 well. The result is that the recognition performance on An-26 is not as good as on the other two, and the average recognition rate by the subspace method is lower than that by the weighted method. Therefore, the weighted feature reduction method proposed in this paper is robust to different test data. Moreover, the computation burden of the subspace method is larger than that of the weighted method. According to Eq. (11), for a $N$-dimensional raw power spectrum (for our measured $N = 128$), the computation of the subspace method is 128 times larger than that of the weighted method.

Table I a comparison in recognition performance between the feature reduction methods

| features | targets | recognition rates | average recognition rates |
|---|---|---|---|
| raw power spectra | Yark-42 | 94.25% | 83.08% |
| | Cessna Citation S/II | 83.25% | |
| | An-26 | 71.75% | |
| reduced feature vectors based on FDR | Yark-42 | 92.50% | 81.08% |
| | Cessna Citation S/II | 78.75% | |
| | An-26 | 72.00% | |
| weighted feature vectors based on FDR | Yark-42 | 94.50% | 88.75% |
| | Cessna Citation S/II | 89.75% | |
| | An-26 | 82.00% | |
| reduced feature vectors based on Fisher's optimal projection subspace | Yark-42 | 96.00% | 86.25% |
| | Cessna Citation S/II | 92.00% | |
| | An-26 | 70.75% | |

## 5. CONCLUSION

A new feature reduction method, the weighted feature reduction method based on FDR is presented in this paper. Compared with the method of using the raw power spectra and some existing feature reduction methods, the weighted feature reduction method can not only reduce the feature dimensionality, but also improve the recognition performance with low computation complexity. In the recognition experiments based on measured data, the proposed method is robust to different test data and achieves good recognition results.

## REFERENCES


[1] M. D. Xing, Z. Bao and B. N. Pei, "The properties of high-resolution range profiles," Optical Engineering, Vol. 41 (2), pp. 493-504, 2002.

[2] X. D. Zhang, Y. Shi, Z. Bao, "A New Feature Vector Using Selected Bispectra for Signal Classification with Application in Radar Target Recognition," IEEE Trans. S. P., Vol. 49 (9), pp. 1875-1885, 2001.

[3] L. Du, H. W. Liu, Z. Bao, "Radar HRRP Target Recognition by the Higher-order Spectra Features", Proceeding of the IASTED International Conference on Artificial Intelligence and Applications, ACTA, Calgary, pp. 627-632, 2004.

[4] S. Theodoridis, K. Koutroumbas, Pattern Recognition (Second Edition), Elsevier Science, New York, 2003.