

# Isometric Cost-Sensitive Laplacian Eigenmaps for Imbalance Radar Target Recognition

Xingjian Xu, Yuehua Li\* and Jianqiao Wang

*School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China*  
*flashxu@163.com, \*nlglyh2013@sina.cn*

## Abstract

*Traditional radar target recognition algorithms utilize balance data set to train the classifier and achieve a satisfactory result on a balance test data set. However, in the case of non-cooperative target recognition, we only obtain a small amount of non-cooperative target samples, while we can obtain a larger number of cooperative target samples easily, which leads to an imbalance training data set. In this paper, we consider the imbalance data classification problem in radar target recognition. We utilize the cost-sensitive approach and assume that different kinds of mistakes lead to different losses. Based on this assumption, a novel radar target recognition algorithm, called isometric cost-sensitive Laplacian eigenmaps (ICSLE), is presented. The basic idea of ICSLE is that the larger the misclassification cost is, the further the distance between two classes is, and vice versa. Moreover, in order to effectively utilize the cost information and local property of observation samples, we use the geodesic distance as the edge weight, instead of the local Euclidean distance. Experiments on millimeter wave radar high-resolution range profile (HRRP) demonstrate the effectiveness of our method.*

**Keywords:** *cost-sensitive, class imbalance problem, Laplacian eigenmaps, geodesic distance, high-resolution range profile*

## 1. Introduction

In the past decade, many manifold-based radar target recognition algorithms [1, 2] have been proposed. These methods attempt to achieve a low recognition error rate while preserving the local structure of observation samples and they achieve better performance than traditional methods. Both these novel methods and traditional methods train the classifier with a balance data set, in which the number of training samples from each class is the same. However, in the real world, the data set we obtained is often imbalanced. In other words, the number of training samples from each class is not equal. For example, we want to detect a military aircraft, but it is not common to obtain the military aircraft sample, most of the aircrafts we detected are the commercial ones. How to find out the small amount of the interested samples from the large number of irrelevant samples is a novel problem. Moreover, some binary-class classification algorithm will face up the imbalance data problem when they are generalized to multi-class classification algorithm by using one-vs-rest technique [3]. Actually, the imbalance data classification problem appears frequently in the field of medical diagnosis [4], financial fraud [5], identification [6], and so on. In this paper, we only consider the imbalance data classification problem in radar target recognition.

Cost-sensitive learning [4-13] is an effective method to deal with the imbalance data classification problem. In recent year, cost-sensitive learning has been studied widely and

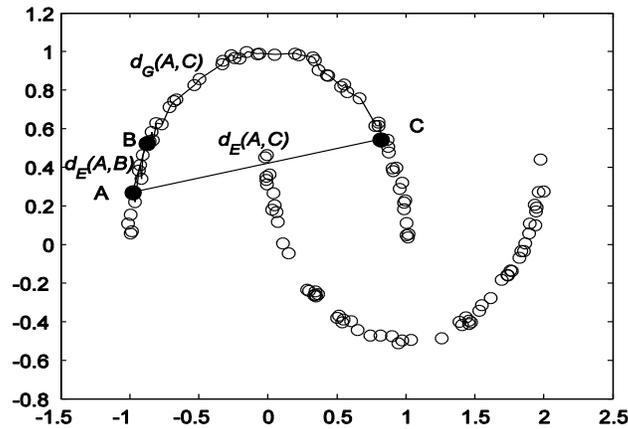
become one of the most important topics in the field of machine learning. Several cost-sensitive learning algorithms have been proposed, such as cost-sensitive boosting [7], cost-sensitive support vector machine [8], cost-sensitive multilayer perceptron [9] and cost-sensitive neural networks [10]. Many cost-blind methods have also evolved the cost-sensitive versions, such as cost-sensitive semi-supervised discriminant analysis [11], cost-sensitive Laplacian score and cost-sensitive constraint score [12, 13]. The basic assumption of cost-sensitive learning is that different kinds of mistakes lead to different losses. This assumption is more reasonable in many applications. The aim of cost-sensitive learning is to minimize the total cost rather than the total error. Generally, there are two kinds of misclassification cost. The first is class-dependent, where the costs of misclassifying any example in class A to B are the same. The second is example-dependent, where the costs of classifying examples in class A to class B are different [6]. We focus on the former one because radar target recognition is generally a class-dependent cost-sensitive problem.

In order to deal with the cost-sensitive radar target recognition problem, we propose a novel algorithm, called isometric cost-sensitive Laplacian eigenmaps (ICSLE). The framework of ICSLE is inspired by Laplacian eigenmaps (LE) [14]. The basic idea of ICSLE is that the larger the misclassification cost, the further the distance between two classes. Since the edge weight in LE is sparse, some distance information may be discarded, and the corresponding cost information will also be discarded if we introduce the cost matrix directly. In ICSLE, we use geodesic distance [15] instead of local Euclidean distance. Since the geodesic distance matrix is not sparse, we can introduce the cost matrix directly. Using geodesic distance cannot only make full use of the cost information, but also utilize the local property of observation samples.

The remaining parts of this paper are organized as follows. Related work is briefly recalled in Section 2, followed by our proposed method in Section 3. The experimental results are presented in Section 4. Finally, contributions are concluded in Section 5.

## 2. Similarity Measurement and Geodesic Distance

In the field of pattern recognition, one of the basic problems is the measurement of the similarity between two samples. To overcome this problem, the cluster assumption is proposed. It assumes that: (1) nearby samples are likely to have the same label; (2) samples on the same structure are likely to have the same label [16]. A following problem is how to determine whether two samples are nearby samples or not. Euclidean distance which is utilized widely can be seen as the simplest method to address this shortcoming. Two samples which have the shortest Euclidean distance are nearby samples. Many classic algorithms utilize the Euclidean distance to measure the similarity and achieve good performance. However, in the nonlinear space, Euclidean distance does not perform well in globally. Take two-moon data set for example. As seen in Figure 1, there are two moons and each moon belongs to a class. For convenience, we define the Euclidean distance between sample  $x_i$  and  $x_j$  as  $d_E(x_i, x_j)$ . Since sample B is close to sample A, the similarity between A and B can be presented by  $d_E(A, B)$ . Sample C is far from A,  $d_E(A, C)$  is not suitable to present the similarity between A and C.



**Figure 1. Example of Euclidean Distance and Geodesic Distance**

The geodesic distance indicates the shortest path between two samples which lie on the manifold. It can be used to measure the distance in the nonlinear space if the samples are sufficiently sampled. A successful application of the geodesic distance is Isomap [15]. Isomap is a nonlinear dimensionality reduction algorithm. The framework of Isomap is as the same as the multidimensional scaling (MDS), just replace the Euclidean distance with the geodesic distance.

Similar with  $d_E(x_i, x_j)$ , we define the geodesic distance between sample  $x_i$  and  $x_j$  as  $d_G(x_i, x_j)$ . Given a data set  $X = [x_1, \dots, x_n]$ , the geodesic distance can be calculated by the following two steps.

1) Construct the neighborhood graph: Define the graph  $G$  over all samples by connecting samples  $x_i$  and  $x_j$  if  $x_i$  is one of the  $k$  nearest neighbors of  $x_j$ . Set the edge lengths as  $d_E(x_i, x_j)$ ;

2) Compute the geodesic distance: Initialize  $d_G(x_i, x_j) = d_E(x_i, x_j)$  if  $x_i$  and  $x_j$  are linked by an edge,  $d_G(x_i, x_j) = \infty$  otherwise. Then for each value of  $t = 1, \dots, n$  in turn, replace all entries  $d_G(x_i, x_j)$  by  $\min\{d_G(x_i, x_j), d_G(x_i, x_t) + d_G(x_t, x_j)\}$ .

From the above description, we can see that the geodesic distance is parameter sensitive. If we set different numbers of nearest neighbors, the geodesic distance may be different. The geodesic distance characterizes the nonlinear property of the samples. The nonlinear sample structure can not be well presented if only a few samples are known.

### 3. Isometric Cost-Sensitive Laplacian Eigenmaps

Given  $n$  points  $X = [x_1, \dots, x_n]$  in  $R^D$ , we want to find  $Y = [y_1, \dots, y_n]$  in  $R^d$  ( $d \ll D$ ), which have a minimum misclassification cost. We first provide a simple review on LE. LE is a cost-blind dimensionality reduction algorithm. It implicitly assumes that the losses of all misclassifications are the same when it is applied in the classification task. In LE, it is assume that if  $x_i$  and  $x_j$  are close, then  $y_i$  and  $y_j$  should also be close. Thus, a reasonable criterion [14] for choosing a good embedding is to solve the following objective function.

$$\min \sum_{ij} \|y_i - y_j\|^2 W_{ij} \quad (1)$$

where  $w_{ij}$  is the edge weight, and

$$W_{ij} = \begin{cases} \exp(-\|x_i - x_j\|^2 / t) & x_i \text{ and } x_j \text{ are neighbors} \\ 0 & \text{others} \end{cases} \quad (2)$$

Now, we consider the cost-sensitive problem. Assume that there are  $c$  classes in  $X$  and each class is denoted by  $l_i (i = 1, \dots, c)$ . We define the cost matrix  $C$  as follow

$$C = \begin{bmatrix} 0 & C_{12} & \dots & C_{1c} \\ C_{21} & 0 & \dots & C_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ C_{c1} & C_{c2} & \dots & 0 \end{bmatrix} \quad (3)$$

where  $C_{ij}$  indicates the cost of misclassifying the samples from the  $i$ th class to the  $j$ th class. The diagonal elements in the cost matrix are zero because there is no loss for a correct classification. In order to minimize the total cost of classification, a naturally idea is that the larger the misclassification cost is, the further the distance between two classes is, and vice versa. We define the cost edge weight matrix  $w^c$  by introducing the cost information. The cost weight can be expressed as  $\exp(-C_{ij})$ . Then, the cost edge weight can be expressed as

$$W_{ij}^c = W_{ij} \exp(-C_{ij}) \quad (4)$$

It is noteworthy that the edge weight matrix  $w$  in LE is sparse. If  $x_i$  is not one of the  $k$  nearest neighbors of  $x_j$ , the corresponding edge weight  $w_{ij}$  is 0. In this case, the cost edge weight  $w_{ij}^c$  equals to 0, which means the cost information between  $x_i$  and  $x_j$  is lost. In order to overcome this shortcoming, we discard the neighborhood information in  $w$  and redefine the edge weight as

$$W_{ij} = \exp(-\|x_i - x_j\|^2 / t) \quad (5)$$

In the field of pattern recognition, the similarity between two samples is usually presented by the distance between these two samples. The edge weight defined by Eq. (5) is suitable for linear sample because the Euclidean distance can truly reflect the similarity between two samples in linear space. However, in most occasions of pattern recognition, such as radar target recognition, the samples are nonlinear. In order to present the similarity between two nonlinear samples, we substitute geodesic distance [15] for Euclidean distance. The geodesic distance contains the neighborhood information, thus the local structure of observation samples can be well preserved. Let  $d_G(x_i, x_j)$  be the geodesic distance between  $x_i$  and  $x_j$ , then the new cost edge weight can be expressed as

$$W_{ij}^{IC} = \exp(-d_G(x_i, x_j) / t) \exp(-C_{ij}) \quad (6)$$

Finally, we obtain the objective function of ICSLE as follow

$$\min \sum_{ij} \|y_i - y_j\|^2 W_{ij}^{IC} \quad (7)$$

The optimization scheme of Eq. (7) is similar with LE. We rewrite Eq. (7) as

$$\begin{aligned}
 & \sum_{ij} \|y_i - y_j\|^2 W_{ij}^{IC} \\
 &= \sum_{ij} (y_i^T y_i - y_i^T y_j - y_j^T y_i + y_j^T y_j) W_{ij}^{IC} \\
 &= \sum_{ij} y_i^T y_i W_{ij}^{IC} + \sum_{ij} y_j^T y_j W_{ij}^{IC} - 2 \sum_{ij} y_i^T y_j W_{ij}^{IC} \\
 &= 2 \sum_i y_i^T y_i D_{ii}^{IC} - 2 \sum_{ij} y_i y_j W_{ij}^{IC}
 \end{aligned} \tag{8}$$

where

$$D_{ii}^{IC} = \sum_j W_{ij}^{IC} \tag{9}$$

In the matrix form, the objective function can be write as

$$\min Y^T L^{IC} Y \tag{10}$$

where  $L^{IC}$  is the cost Laplacian matrix, and

$$L_{ij}^{IC} = D_{ii}^{IC} - W_{ij}^{IC} \tag{11}$$

In order to remove the arbitrary scaling factor in  $Y$ , we introduce a constraint

$$y_i^T D_{ii}^{IC} y_i = 1 \tag{12}$$

Then, the optimal solution of Eq. (7) can be obtained by solving the following generalized eigenvalue problem

$$L^{IC} Y = \lambda D^{IC} Y \tag{13}$$

The low dimensional embedding  $Y$  is given by the eigenvectors corresponding to the  $d$  smallest non-zero eigenvalues of Eq. (13).

We summarize the ICSLE algorithm in Table 1.

**Table 1. ICSLE Algorithm**

---

**Input:** the original sample matrix  $X$ , the label of  $X$ , the number of nearest neighbor  $k$ , the dimension of the low dimensional embedding  $d$ , the cost matrix  $C$ , the distance weight parameter  $t$  in Eq. (6).

---

**Output:** the low dimensional embedding  $Y$ .

---

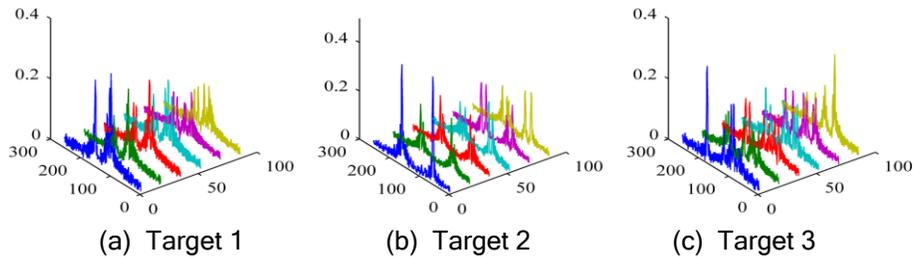
**Algorithm:**

- 1: Construct neighborhood graph: Define the graph  $G$  over all data points by connecting samples  $x_i$  and  $x_j$  if  $x_i$  is one of the  $k$  nearest neighbors of  $x_j$ . Set edge lengths equal to  $d_E(x_i, x_j)$ ;
  - 2: Compute geodesic distance: Initialize  $d_G(x_i, x_j) = d_E(x_i, x_j)$  if  $x_i$  and  $x_j$  are linked by an edge,  $d_G(x_i, x_j) = \infty$  otherwise. Then for each value of  $m = 1, \dots, n$  in turn, replace all entries  $d_G(x_i, x_j)$  by  $\min\{d_G(x_i, x_j), d_G(x_i, x_m) + d_G(x_m, x_j)\}$ ;
  - 3: Calculate the cost Laplacian matrix: Calculate the matrix  $D^{IC}$ , the cost Laplacian matrix  $L^{IC}$  by using (9) and (11), respectively;
  - 4: Calculate the low dimensional embedding: Solve the generalized eigenvalue equation in (13). The low dimensional embedding  $Y$  is given by the eigenvectors corresponding to the  $d$  smallest non-zero eigenvalues.
-

## 4. Experiment

### 4.1. Experiment Settings

We utilize millimeter wave radar high-resolution range profile (HRRP) test our method. For high-resolution range radar, in a resolution cell, the target can be seen as an assembly of many scatterers, which usually locate on the position where the curvature is small, such as corner and vertex. In this experiment, we use corner reflectors with different position to simulate three different kinds of targets. The experimental environment is as follows. The background is cement floor. The bandwidth of transmit signal is 400MHz, thus the range resolution is 0.375m. For each simulate target, we obtain a 256-point HRRP every 1° from azimuth 0° to 89°. Thus each target has 90 HRRP samples, and each HRRP sample is a 256-dimensional vector. Figure 2 shows 6 HRRP samples (azimuth 0°, 17°, 34°, 51°, 68° and 85°) for each target.



**Figure 2. HRRP Samples (azimuth 0°, 17°, 34°, 51°, 68° and 85°) for Three Targets**

In order to test the influence of the cost and the imbalance degree of samples, we perform experiments on four settings with different cost matrix and different number of samples. These four settings are listed in Table 2 in detail. The cost matrix is specified as follow

$$C_1 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 10 & 10 & 0 \end{bmatrix} \quad C_2 = \begin{bmatrix} 0 & 1 & 5 \\ 1 & 0 & 5 \\ 10 & 10 & 0 \end{bmatrix} \quad (14)$$

**Table 2. Experiment Settings**

setting	Cost matrix	Number of training samples			Number of test samples		
		Target 1	Target 2	Target 3	Target 1	Target 2	Target 3
C1N1	C1	45	45	45	45	45	45
C1N2	C1	45	35	25	45	35	25
C2N1	C2	45	45	45	45	45	45
C2N2	C2	45	35	25	45	35	25

From the cost matrix, we see that the cost of misclassifying samples in Target 3 to Target 1 or Target 2 is the largest; the cost of misclassifying samples in Target 1 to Target 2 or Target 2 to Target 1 is the smallest. The difference between cost matrix C1 and C2 is that the cost of misclassifying samples in Target 1 or Target 2 to Target 3 in C2 is much larger than that in C1. Both the training samples and test samples are obtained by random selection from the whole HRRP samples.

In order to demonstrate the effectiveness of our proposed algorithm, we compare ICSLE with supervised LE (SLE) [17], cost sensitive principal component analysis (CSPCA) [18,

19] and cost sensitive locality preserving projections (CSLPP) [18, 19]. We applied the nearest neighbor classifier in the Euclidean space to perform recognition. All experiments repeat for 20 times. The average cost, average recognition error rate and three kinds of misclassification errors (Err1, Err2 and Err3) are compared. Let  $N_{Err_1}$  indicates the number of samples which are classified from Target 1 to Target 2 or Target 2 to Target 1,  $N_{Err_2}$  indicates the number of samples which are classified from Target 1 or Target 2 to Target 3,  $N_{Err_3}$  indicates the number of samples which are classified from Target 3 to Target 1 or Target 2,  $N_{t_i} (i = 1, 2, 3)$  indicates the number of test samples. Then, the three kinds of misclassification errors can be expressed as

$$Err1 = N_{Err_1} / N_{t1} \times 100 \% \quad (15)$$

$$Err2 = N_{Err_2} / N_{t2} \times 100 \% \quad (16)$$

$$Err3 = N_{Err_3} / N_{t3} \times 100 \% \quad (17)$$

#### 4.2. Experimental Results and Discussion

Figure 3 to Figure 6 shows the cost versus different feature dimension of SLE, CSPCA, CSLPP and ICSLE on the 4 different experiment settings, respectively, as well as the error. Table 3 to Table 6 lists the comparison of average cost, average recognition error and three kinds of misclassification error on the 4 different experiment settings in detail, respectively.

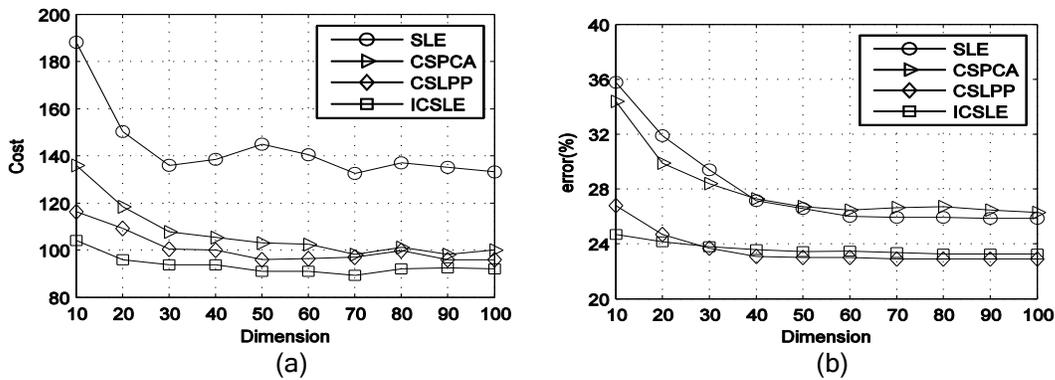


Figure 3. Performance Comparison of SLE, CSPCA, CSLPP and ICSLE on Experiment Setting C1N1: (a) Cost vs. Dimension, (b) Error (%) vs. Dimension

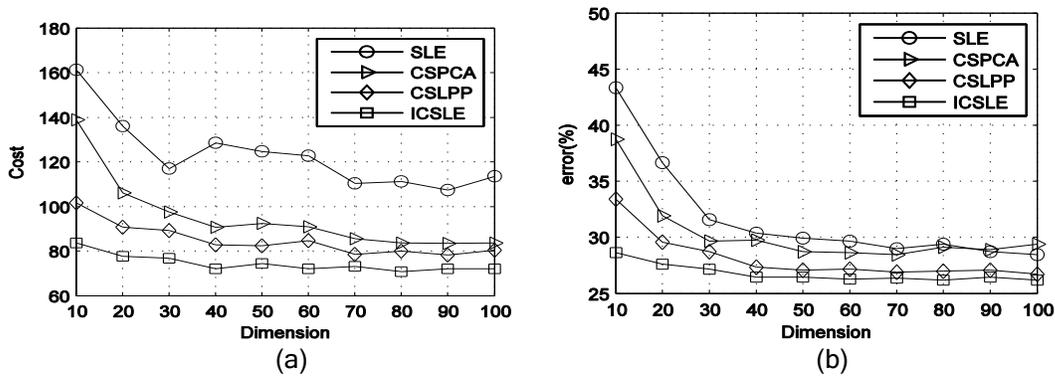


Figure 4. Performance comparison of SLE, CSPCA, CSLPP and ICSLE on Experiment Setting C1N2: (a) Cost vs. Dimension, (b) Error (%) vs. Dimension

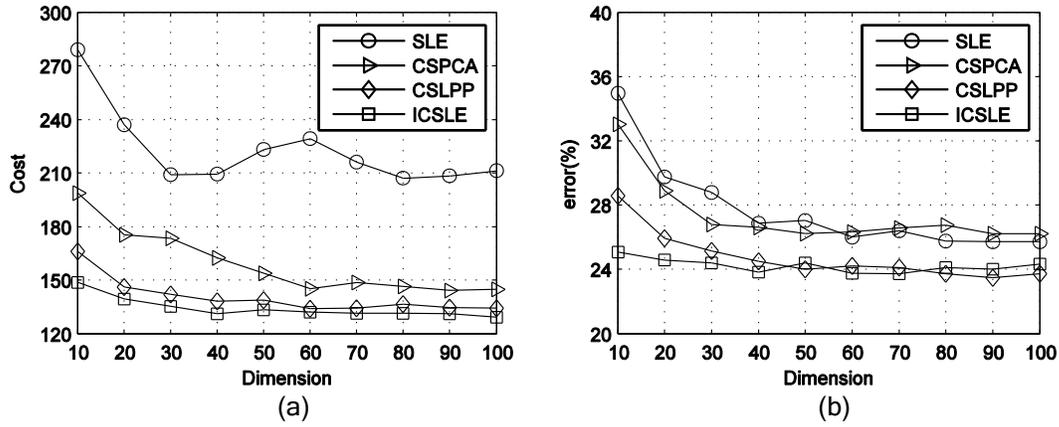


Figure 5. Performance Comparison of SLE, CSPCA, CSLPP and ICSLE on Experiment Setting C2N1: (a) Cost vs. Dimension, (b) Error (%) vs. Dimension

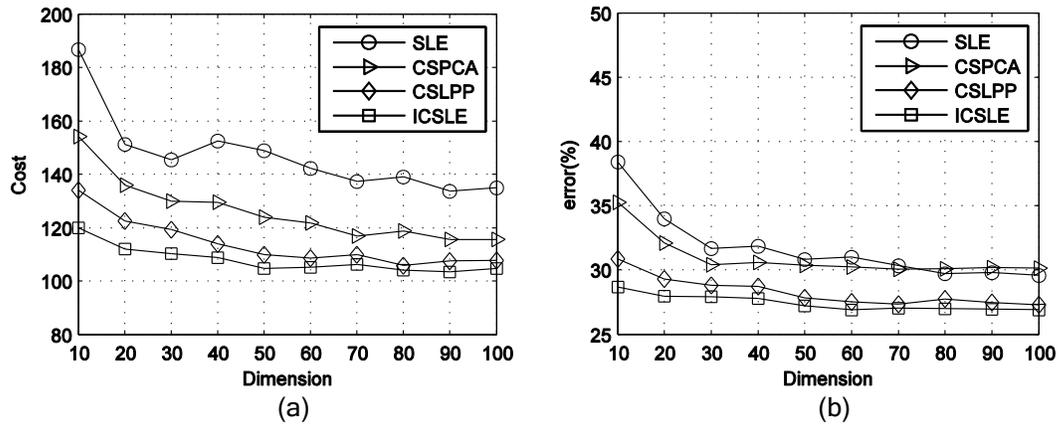


Figure 6. Performance Comparison of SLE, CSPCA, CSLPP and ICSLE on Experiment Setting C2N2: (a) Cost vs. Dimension, (b) Error (%) vs. Dimension

Table 3. Comparison of Average Cost, Average Recognition Error and Three Kinds of Misclassification Error on Experiment Setting C1N1

Algorithm	cost	Error(%)	Err1(%)	Err2(%)	Err3(%)
SLE	143.60	28.04	28.22	29.78	26.11
CSPCA	107.00	27.93	34.89	31.78	17.11
CSLPP	100.70	23.59	26.67	27.11	17.00
ICSLE	93.55	23.63	29.33	26.33	15.22

Table 4. Comparison of Average Cost, Average Recognition Error and Three Kinds of Misclassification Error on Experiment Setting C1N2

Algorithm	cost	Error(%)	Err1(%)	Err2(%)	Err3(%)
SLE	123.3	31.71	29.22	29.00	40.00
CSPCA	95.3	30.33	32.22	29.43	28.20
CSLPP	84.85	28.10	30.67	27.29	24.60
ICSLE	74.45	26.76	29.11	28.14	20.60

**Table 5. Comparison of Average Cost, Average Recognition Error and Three Kinds of Misclassification Error on Experiment Setting C2N1**

Algorithm	cost	Error(%)	Err1(%)	Err2(%)	Err3(%)
SLE	222.95	27.70	21.56	28.33	33.22
CSPCA	159.35	27.37	39.67	22.00	20.44
CSLPP	140.45	24.74	35.44	22.22	16.56
ICSLE	134.30	24.22	36.22	20.44	16.00

**Table 6. Comparison of Average Cost, Average Recognition Error and Three Kinds of Misclassification Error on Experiment Setting C2N2**

Algorithm	cost	Error(%)	Err1(%)	Err2(%)	Err3(%)
SLE	147.10	31.71	31.89	32.43	30.40
CSPCA	126.25	30.95	37.22	27.43	24.60
CSLPP	113.85	28.29	34.67	24.43	22.20
ICSLE	107.90	27.43	34.22	23.71	20.40

From the Figures, we see that our method always obtains the smallest cost in the 4 settings. The cost of CSLPP is slightly larger than that of ICSLE and is smaller than that of CSPCA. The cost of SLE is always the largest. This result is reasonable, because SLE is cost blind, the other three methods are cost sensitive. Moreover, CSPCA is a linear cost sensitive algorithm and it can be seen as the cost sensitive version of PCA. CSLPP is the cost sensitive generalization of LPP [20]. Although LPP is a linear dimensionality reduction algorithm, the weight matrix utilized by LPP contains the nonlinear structure information of the samples. Our method can be seen as the cost sensitive version of LE. To enhance performance of our method, we utilize the geodesic distance to measure the similarity. This approach is significant. The geodesic distance is not sparse so that the cost information contained in the training samples can be utilized. Although LPP is the linear version of LE, the weight matrix is sparse, which leads to the loss of cost information.

Then, we discuss the recognition error. Although the average recognition error is not the primary performance indicator in cost sensitive learning, the recognition error of a specific target which has a small number of samples is what we concern. From Figure 3 to Figure 6, we see that in the case of C1N2 and C2N2, the average recognition error of our method is the smallest. While in the case of C1N1 and C2N1, the average recognition error of our method is similar with that of CSLPP. Table 3 to Table 6 gives a comparison of each kind of error in detail. Err3 of the cost sensitive methods in each table is the smallest in most cases (the exception is the Err3 of CSPCA in the case of C2N1), while Err3 of SLE does not satisfy this tendency. That means the error has some relationship with the cost. Then, we compare Err2 in Table 3 and Table 5. The difference between C1N1 (the experiment settings of Table 3) and C2N1 (the experiment settings of Table 5) is that the cost of Err2 on C1N1 is 1 while the cost of Err2 on C2N1 is 5. From Table 3 and Table 5, we see that Err2 in Table 5 is smaller than Err2 in Table 3. The similar result can be obtained from Table 4 and Table 6. From these comparisons, we find an empirical law: for the cost sensitive learning, the larger the cost, the smaller the recognition error.

Finally, we discuss the influence of the imbalance degree of samples. We compare Table 3 and Table 4, in which the cost is the same while the number of samples is different. The number of samples from target 3 and target 2 in Table 4 is smaller than those in Table 3. Err3 and Err2 in Table 4 are larger than those in Table 3 in most cases (the exception is the Err2 of CSPCA). The similar result can be obtained from Table 5 and Table 6. From these

comparisons, we find the second empirical law: for the cost sensitive learning, the larger the number of samples, the smaller the recognition error.

From the above experiments and discussion, we demonstrate the effectiveness of our method on HRRP data set. Moreover, we obtain two empirical law of the cost sensitive learning: the larger the cost, the smaller the recognition error; the larger the number of samples, the smaller the recognition error.

## 5. Conclusion

In this paper, we propose a novel cost-sensitive learning method, called ICSLE, for radar target recognition. The basic idea of ICSLE is that the larger the misclassification cost is, the further the distance between two classes is, and vice versa. Since the feature of HRRP is usually sampled from the nonlinear space, we use the geodesic distance to measure the similarity between two samples. The edge weight matrix in LE is sparse, while the geodesic distance is not sparse, so we can introduce the cost information directly. Moreover, the local property can be well preserved by using the geodesic distance. Experimental results on millimeter wave radar HRRP demonstrate the effectiveness of our method. In addition, we obtain some empirical law of the cost sensitive learning: the larger the cost, the smaller the recognition error; the larger the number of samples, the smaller the recognition error.

## Acknowledgements

The authors would like to thank the anonymous reviewers and editors for their helpful suggestions. This work is supported by National (Jiangsu) Natural Science Foundation of China under Grants 60901008 and 61001010 (BK2010490), National Ministry Foundation of China under Grants 51305050102.

## References

- [1] X. L. Yu, X. G. Wang and B. Liu, "Supervised Kernel Neighborhood Preserving Projections for Radar Target Recognition", *Signal Processing*, vol. 9, no. 88, (2008), pp. 2335-2339.
- [2] X. L. Yu and X. G. Wang, "Kernel Uncorrelated Neighborhood Discriminative Embedding for Radar Target Recognition", *Electronics Letters*, vol. 2, no. 44, (2008), pp. 154-155.
- [3] K. Huang, Z. Xu, I. King and M. R. Lyu, "Semi-Supervised Learning from General Unlabeled Data", *Proc. 8th ICDM*, (2008) December 15-19, pp. 273-282, Pisa, Italy.
- [4] M. Cebe and C. G. Demir, "Qualitative Test-Cost Sensitive Classification", *Pattern Recognition Letters*, vol. 13, no. 31, (2010), pp. 2043-2051.
- [5] J. Kim, K. Choi, G. Kim and Y. Suh, "Classification Cost: An Empirical Comparison among Traditional Classifier, Cost-Sensitive, and Metacost", *Expert Systems with Applications*, vol. 4, no. 39, (2012), pp. 4013-4019.
- [6] Y. Zhang and Z. H. Zhou, "Cost-Sensitive Face Recognition", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 32, (2010), pp. 1758-1769.
- [7] H. M. Shirazi and N. Vasconcelos, "Cost-Sensitive Boosting", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 2, no. 33, (2011), pp. 294-309.
- [8] Y. F. Li, J. Kwok and Z. Zhou, "Cost-Sensitive Semi-supervised Support Vector Machine", *Proc. Nat. Conf. Artificial Intell.*, (2010) July 11-15, pp. 500-505, Atlanta, GA, United States.
- [9] C. L. Castro and A. P. Braga, "Novel Cost-Sensitive Approach to Improve the Multilayer Perceptron Performance on Imbalanced Data", *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 6, no. 24, (2013), pp. 888-899.
- [10] S. Raudys and A. Raudys, "Pairwise Costs in Multiclass Perceptrons", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 7, no. 32, (2010), pp. 1324-1328.
- [11] J. Lu, X. Zhou, Y. P. Tan, Y. Shang and J. Zhou, "Cost-Sensitive Semi-Supervised Discriminant Analysis for Face Recognition", *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 7, (2012), pp. 944-953.

- [12] L. Miao, M. Liu and D. Zhang, "Cost-Sensitive Feature Selection with Application in Software Defect Prediction", Proc. 21st ICPR, (2012) November 11-15, pp. 967-970, Tsukuba, Japan.
- [13] R. S. Rodriguez and D. G. Garcia, "Cost-Sensitive Feature Selection Based on the Set Covering Machine", Proc. Int. Conf. Data Mining Workshops, (2010) December 13-13, pp. 740-746, Sydney, Australia.
- [14] M. Belkin and P. Niyogi, "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation", Neural Computation, vol. 6, no. 15, (2003), pp. 1373-1396.
- [15] J. B. Tenenbaum, V. Silva and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction, Science, vol. 5500, no. 290, (2000), pp. 2319-2323.
- [16] O. Chapelle, J. Weston and B. Scholkopf, "Cluster Kernels for Semi-Supervised Learning", Proc. 15th NIPS, (2002) December 9-14, Vancouver, British Columbia, Canada.
- [17] T. Perry, H. Zha, P. Frias, D. Zeng and M. Braunstein, "Supervised Laplacian Eigenmaps with Applications in Clinical Diagnostics for Pediatric Cardiology", Proc. CoRR abs/1207.7035, Canada, (2012).
- [18] J. Lu and Y. P. Tan, "Cost-Sensitive Subspace Learning for Human Age Estimation", Proc. 17th ICIP, (2010) September 26-29, pp. 1593-1596, Hong Kong, China.
- [19] J. Lu and Y. P. Tan, "Cost-Sensitive Subspace Learning for Face Recognition", Proc. 23th CVPR, (2010) June 13-18, pp. 2661-2666, San Francisco, CA, USA.
- [20] X. He and P. Niyogi, "Locality Preserving Projections, Proc. Conf. Advances in Neural Information Processing Systems", (2003) December 8-13, Vancouver, British Columbia, Canada.

## Authors



**Xingjian Xu**, he received the Master Degree in Control Theory and Control Engineering from Nanjing University of Science and Technology in 2009. His current research interests on manifold learning, radar target recognition, electric circuit and system.

