Journal of Computer and Communications, 2020, 8, 14-22
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Locality Preserving Discriminant Projection for Speaker Verification

Chunyan Liang*, Wei Cao, Shuxin Cao

College of Computer Science and Technology, Shandong University of Technology, Zibo, China
Email: *liangchunyan_sdut@163.com, 846102759@qq.com, 601933697@qq.com

How to cite this paper: Liang, C.Y., Cao, W. and Cao, S.X. (2020) Locality Preserving Discriminant Projection for Speaker Verification. *Journal of Computer and Communications*, **8**, 14-22.
https://doi.org/10.4236/jcc.2020.811002

Received: October 8, 2020
Accepted: November 8, 2020
Published: November 11, 2020

Abstract

In this paper, a manifold subspace learning algorithm based on locality preserving discriminant projection (LPDP) is used for speaker verification. LPDP can overcome the deficiency of the total variability factor analysis and locality preserving projection (LPP). LPDP can effectively use the speaker label information of speech data. Through optimization, LPDP can maintain the inherent manifold local structure of the speech data samples of the same speaker by reducing the distance between them. At the same time, LPDP can enhance the discriminability of the embedding space by expanding the distance between the speech data samples of different speakers. The proposed method is compared with LPP and total variability factor analysis on the NIST SRE 2010 telephone-telephone core condition. The experimental results indicate that the proposed LPDP can overcome the deficiency of LPP and total variability factor analysis and can further improve the system performance.

Keywords

Speaker Verification, Locality Preserving Discriminant Projection, Locality Preserving Projection, Manifold Learning, Total Variability Factor Analysis

1. Introduction

Speaker verification is a subtask of speaker recognition, whose purpose is to verify whether a segment of speech is spoken by a designated speaker [1] [2]. Total variability factor analysis has been widely used in speaker verification [3] [4] [5] [6]. In total variability factor analysis, the speaker and the channel variabilities are contained simultaneously in a low-dimensional space which is referred to as the total variability space. By the space mapping, the useful information can be obtained by reducing the dimensionality of the mean supervector of the Gaussian mixture model (GMM) and the latent variables can be estimated using li-

mited data. The low-dimensional variable characteristic of the speaker's identity is called the total variability factor vector, or i-vector. Support vector machine (SVM) can be used as a classifier for i-vector [7] [8].

As an application of probabilistic principal component analysis (PPCA), total variability factor analysis only analyzes the speech data from a global perspective [9] [10]. To compensate for the deficiency, we introduced locality preserving projection (LPP) [11], neighborhood preserving embedding (NPE) [12], and discriminant neighborhood embedding (DNE) [13] to speaker verification. By constructing a graph containing the neighborhood information of the speech data, the inherent local neighborhood relationship of the speech data is optimally preserved. Combined with total variability factor analysis, the performance of speaker verification is improved [14] [15]. Here, LPP is an unsupervised learning algorithm [11] [16] that is not concerned with the speaker label information in the dimensionality-reduction process and does not make use of the discriminative information between the speech data of different speakers. However, the speaker label information of the training data and the discriminative information of the speech data are of great importance in speaker verification.

In view of the above shortcomings of LPP, we apply the locality preserving discriminant projection (LPDP) algorithm in speaker verification. LPDP can bring in the speaker label information from the speech data and, through optimization, preserve the inherent local manifold structure of the speech data samples from the same speaker to reduce the distance between them. At the same time, the distance between the speech data samples from different speakers is enlarged to enhance the discriminative ability of the embedding space.

The remainder of this paper is organized as follows. The LPP algorithm based on i-vector is introduced in Section 2. The LPDP algorithm is proposed in Section 3. The experiment and results are presented in Section 4. The conclusion is given in Section 5.

2. LPP Algorithm Based on I-Vector

2.1. Total Variability Factor Analysis

Based on the total variability space, the GMM mean supervector containing speaker and channel information in the speech data can be expressed as

$$\boldsymbol{M} = \boldsymbol{m} + \boldsymbol{T}\boldsymbol{w} \tag{1}$$

where $\boldsymbol{m}$ is the mean supervector of the universal background model (UBM) independent of the speaker and channel; T is the total variability space which is defined by the total variability matrix; and w is a low-dimensional latent variable that obeys the normal distribution, known as the total variability factor vector, or identity vector (i-vector). Total variability factor analysis can be regarded as a feature-extraction module. It projects the speech data into the low-rank total variability space $\boldsymbol{T}$ to obtain the i-vector $\boldsymbol{w}$. The training method of $\boldsymbol{T}$ and the extraction process of the i-vector have been described previously [4] [8].

The intersession compensation can be carried out in a low-dimensional space

where the i-vector lies. The linear discriminant analysis (LDA) approach [17] and within class covariance normalization (WCCN) approach [18] are often used for intersession compensation. After the intersession compensation, modeling and scoring are made using SVM.

2.2. LPP Algorithm

The speaker verification system framework, in which the LPP algorithm based on i-vector is used, is presented in **Figure 1**. The dashed boxes from left to right refer to Enrollment, Training and Testing, respectively.

On the basis of i-vector, the LPP algorithm is used to achieve an effective combination of the total variability factor analysis technique and the LPP algorithm that retains both the global and local neighborhood structures of the speech data, thereby significantly improving system performance [11]. However, the known speaker label information of the speech data is not used in the dimensionality-reduction process of the LPP algorithm. As a result, although the locality-preserving projection space matrix $\boldsymbol{P}$ has a strong descriptive ability, its discriminative ability is not strong, which to a certain degree affects the recognition

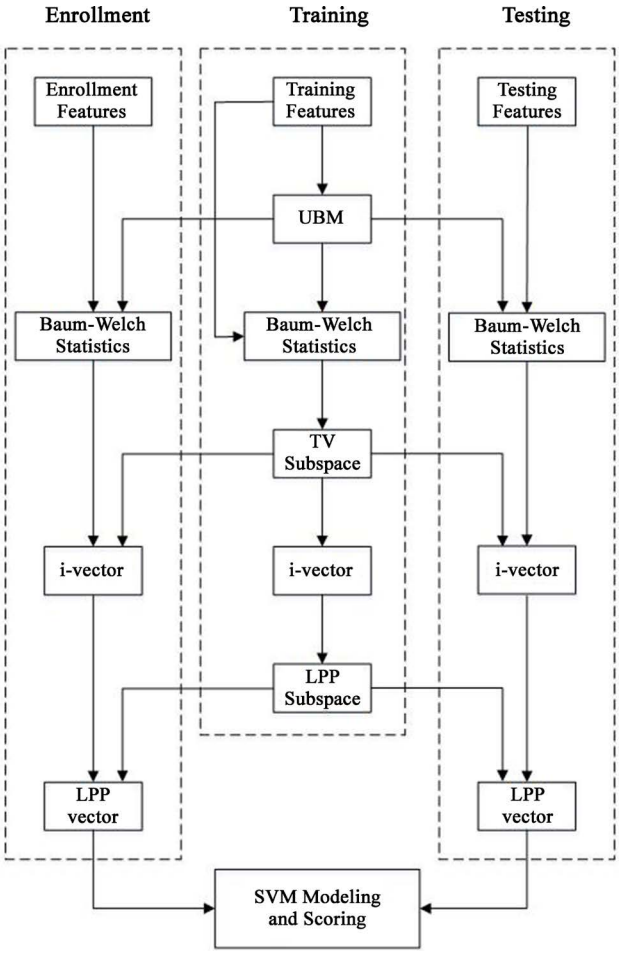

Figure 1. The framework of speaker verification system by using LPP algorithm based on i-vector.

performance of the system.

3. LPDP Algorithm

LPDP is an effective manifold learning method that has been successfully applied in face recognition [19]. The basic idea of LPDP is to divide the nearest neighbor graph in the LPP algorithm into intra-class and out-of-class graphs. LPDP can maintain the local neighborhood relationship of the same speaker's speech data samples and reduce the distance between them. At the same time, LPDP emphasizes the discrimination information between speakers and expands the distance between their speech data. Combined with total variability factor analysis, the algorithm can globally and locally analyze the feature structure of speech data more comprehensively, and at the same time reflects the between-speaker difference and enhances the discriminatory ability of the embedding space.

The idea of applying LPDP to speaker verification is similar to that of LPP as shown in **Figure 1**. The corresponding i-vectors of given N items of training speech data with speaker labels constitute a vector set $\boldsymbol{W}=\{\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_N\}$, where $\boldsymbol{w}_i\in \mathrm{R}^D$, $i=1,2,\cdots,N$. The purpose of LPDP is to find an optimal locality preserving discriminant projection space matrix $\boldsymbol{A}=[\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_K]$ and embed the i-vector of the speech in space R^D in the feature-space R^K ($K<D$). In the R^K space, the speech data point $\boldsymbol{x}_i$ is transformed to $\boldsymbol{y}_i=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{w}_i$. The steps to train the locality preserving discriminant projection space matrix $\boldsymbol{A}$ are as follows.

Step 1: Determine the neighborhood of the i-vector $\boldsymbol{w}_i$, which consists of all the i-vectors whose similarity with $\boldsymbol{w}_i$ is less than its average similarity, *i.e.*,

$$MS(w_i)=\frac{1}{N}\sum_{j=1}^{N}\frac{w_i^{\mathrm{T}}w_j}{\|w_i\|_2\|w_j\|_2} \tag{2}$$

$$NB(w_i)=\left\{w_j\left|\frac{w_i^{\mathrm{T}}w_j}{\|w_i\|_2\|w_j\|_2}>MS(w_i)\right.\right\} \tag{3}$$

where $MS(\boldsymbol{w}_i)$ is the average similarity of all the N i-vectors for the training speech data with i-vector $\boldsymbol{w}_i$, and $NB(\boldsymbol{w}_i)$ represents the neighborhood i-vectors of $\boldsymbol{w}_i$.

Step 2: Construct two subgraphs of the neighborhood graph: the in-class graph $\boldsymbol{G}^{\mathrm{in}}$ and out-of-class graph $\boldsymbol{G}^{\mathrm{out}}$. In both the in-class graph $\boldsymbol{G}^{\mathrm{in}}$ and the out-of-class graph $\boldsymbol{G}^{\mathrm{out}}$, the i-th node corresponds to the i-vector $\boldsymbol{w}_i$. For the in-class graph $\boldsymbol{G}^{\mathrm{in}}$, we put a directed edge from node i to j if i-vector $\boldsymbol{w}_j$ is in the neighborhood of i-vector $\boldsymbol{w}_i$ and is from the same class as i-vector $\boldsymbol{w}_i$. For the out-class graph $\boldsymbol{G}^{\mathrm{out}}$, we put a directed edge from node i to j if i-vector $\boldsymbol{w}_j$ is in the neighborhood of i-vector $\boldsymbol{w}_i$ but is from the different class of $\boldsymbol{w}_i$.

Step 3: Calculate the weights of the edges in $\boldsymbol{G}^{\mathrm{in}}$ and $\boldsymbol{G}^{\mathrm{out}}$, and obtain their respective weight matrices, $\boldsymbol{W}^{\mathrm{in}}$ and $\boldsymbol{W}^{\mathrm{out}}$.

1) Denote the weight of the edge between i-vector $\boldsymbol{w}_i$ and i-vector $\boldsymbol{w}_j$ in $\boldsymbol{G}^{\mathrm{in}}$ as

W_{ij}^{in} and choose its value as

$$W_{ij}^{\text{in}} = \begin{cases} \exp\left(-\dfrac{\left\|w_i - w_j\right\|^2}{t}\right) & spk\left(w_i\right) = spk\left(w_j\right), w_j \in N\left(w_i\right) \text{ or } w_i \in N\left(w_j\right) \\ 0 & \text{other} \end{cases} \tag{4}$$

2) Denote the weight of the edge between i-vector $\boldsymbol{w}_i$ and i-vector $\boldsymbol{w}_j$ in $\boldsymbol{G}^{\text{out}}$ as W_{ij}^{out} and choose its value as

$$W_{ij}^{\text{out}} = \begin{cases} 1 & spk\left(w_i\right) \neq spk\left(w_j\right), w_j \in N\left(w_i\right) \text{ or } w_i \in N\left(w_j\right) \\ 0 & \text{other} \end{cases} \tag{5}$$

Here, $spk\left(w_i\right)$ represents the speaker label information of i-vector $\boldsymbol{w}_i$, and t is the mean distance of all the i-vectors for the training speech data.

Step 4: Calculate the locality preserving discriminant projection matrix $\boldsymbol{A}$. The idea of LPDP is that, in the embedding space, the i-vectors from the same speaker have the smallest in-class divergence after projection, *i.e.*, the distance between the same speaker's i-vectors is as small as possible. Conversely, the i-vectors from different speakers have the largest between-class divergence after projection, *i.e.*, they are as far from each other as possible. To achieve these goals, they are integrated into the following two optimization problems [20]:

$$\min \sum_{i,j} \left\|y_i - y_j\right\|^2 W_{ij}^{\text{in}} = \min tr\left(\boldsymbol{A}^{\text{T}} \boldsymbol{X} \boldsymbol{L}^{\text{in}} \boldsymbol{X}^{\text{T}} \boldsymbol{A}\right) \tag{6}$$

$$\max \sum_{i,j} \left\|y_i - y_j\right\|^2 W_{ij}^{\text{out}} = \max tr\left(\boldsymbol{A}^{\text{T}} \boldsymbol{X} \boldsymbol{L}^{\text{out}} \boldsymbol{X}^{\text{T}} \boldsymbol{A}\right) \tag{7}$$

where $L^{\text{in}} = D^{\text{in}} - W^{\text{in}}$ is a Laplace operator for the in-class graph, D^{in} is a diagonal matrix, $D_{ii}^{\text{in}} = \sum_j W_{ij}^{\text{in}}$, $L^{\text{out}} = D^{\text{out}} - W^{\text{out}}$ is a Laplace operator for the out-of-class graph, D^{out} is a diagonal matrix, and $D_{ii}^{\text{out}} = \sum_j W_{ij}^{\text{out}}$.

Using the constraint condition $\boldsymbol{A}^{\text{T}} \boldsymbol{X} D^{\text{out}} \boldsymbol{X}^{\text{T}} \boldsymbol{A} = \boldsymbol{I}$, (6) and (7) can be integrated into one optimization problem,

$$\left.\begin{array}{l} \min\left[\alpha tr\left(\boldsymbol{A}^{\text{T}} \boldsymbol{X} \boldsymbol{L}^{\text{in}} \boldsymbol{X}^{\text{T}} \boldsymbol{A}\right) - \beta tr\left(\boldsymbol{A}^{\text{T}} \boldsymbol{X} \boldsymbol{L}^{\text{out}} \boldsymbol{X}^{\text{T}} \boldsymbol{A}\right)\right] \\ \boldsymbol{H} = \alpha \boldsymbol{L}^{\text{in}} - \beta \boldsymbol{L}^{\text{out}} \end{array}\right\} \Updownarrow \min tr\left(\boldsymbol{A}^{\text{T}} \boldsymbol{X} \boldsymbol{H} \boldsymbol{X}^{\text{T}} \boldsymbol{A}\right) \tag{8}$$

which can be further transformed to a generalized eigenvalue problem,

$$\boldsymbol{X} \boldsymbol{H} \boldsymbol{X}^{\text{T}} \boldsymbol{A} = \lambda \boldsymbol{X} D^{\text{out}} \boldsymbol{X}^{\text{T}} \boldsymbol{A} \tag{9}$$

By solving Equation (9), the locality-preserving discriminant projection space matrix $\boldsymbol{A} = \left[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_K\right]$ can be obtained, where $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_K$ are the eigenvectors corresponding to the largest K eigenvalues of the above problem.

4. Experiments

4.1. Experimental Setup

Experiments were carried out on the core test set of the NIST SRE 2010 tele-

phone training and telephone testing dataset. Equal error rate (EER) and minimum detect cost function (minDCF) were used as metrics for system evaluation [21] [22].

In the experiments, 36-dimensional Mel Frequency Cepstral Coefficient (MFCC) including 18 MFCC coefficients and their first order derivatives were utilized. Each frame of a speech utterance was processed by a 20 ms Hamming window with 10 ms shift. To mitigate channel effects, feature warping, cepstral mean subtraction (CMN) and cepstral variance normalization (CVN) were applied to the features.

Two gender dependent universal background models (UBM) with a Gauss number of 1024 were trained using the NIST SRE 2004 1-side dataset. The gender related total variability matrix T, LPP matrix, LPDP matrix, WCCN, and LDA matrix were trained by the NIST SRE 2004, 2005, and 2006 corpus. The background data for SVM were also selected from the data of NIST SRE 2004, 2005 and 2006 datasets. The SVM Light toolkit was used for SVM modeling [23].

4.2. Experimental Results

To verify the performance of the proposed LPDP algorithm, we experimentally compared it with the traditional total variability factor analysis and LPP algorithms.

Table 1 shows the performance comparison of the three algorithms without channel compensation. It is observed that applying the LPDP algorithm to i-vector is equivalent to effectively combining total variability factor analysis technology with the LPP algorithm. This combination can maintain the global and local neighborhood structures of the speech data. Compared to total variability factor analysis, which can only preserve the global structure of speech data, LPP and LPDP can significantly improve system performance. LPDP can also make effective use of the speaker label information of the speech data and, through optimization, maintain the intrinsic local manifold structure of the same speaker's speech data. As well, the distance between the speech data of different speakers is expanded in LPDP and the discrimination performance of the embedding space is enhanced to further improve system performance. Compared with LPP, LPDP leads to a relative improvement of 16.36% in EER and 13.04% in minDCF for male testing dataset, and 29.33% in EER and 8.67% in minDCF for female testing dataset.

Table 1. Comparison of EER and minDCF of LPDP, LPP, and total variability factor analysis (without channel compensation).

System	Male		Female	
	EER (%)	minDCF	EER (%)	minDCF
Total variability factor analysis	8.42	0.0672	9.84	0.0832
LPP	5.99	0.0606	8.66	0.0738
LPDP	5.01	0.0527	6.12	0.0674

Table 2 shows the experimental results of the three algorithms with LDA intersession compensation. The table shows that, with LDA channel compensation, LPDP performs better than LPP. For male and female testing dataset, EER of the LPDP system was relatively improved by 23.78% and 26.67%, respectively, and minDCF was relatively improved by 11.18% and 5.95%, respectively.

Table 3 shows the experimental results of the three algorithms with WCCN intersession compensation. When compared to the performance of LPP with WCCN channel compensation, Table 3 shows that LPDP system outperforms the LPP system, yielding 11.81% relative improvement in EER and 6.85% in minDCF for male testing dataset, as well as 8.2% relative improvement in EER and 5.19% in minDCF for female testing dataset.

Table 4 shows the experimental results of the three algorithms after performing both LDA and WCCN. The table shows that LPDP still outperformed LPP when channel compensation was provided by both LDA and WCCN. Compared to the performance of LPP, the LPDP system gives additional gains of 9.16% and 11.47% respectively in EER and minDCF for male testing dataset, as well as 10.94% and 10.40% respectively in EER and minDCF for female testing dataset.

Table 2. Comparison of EER and minDCF of LPDP, LPP, and total variability factor analysis (LDA channel compensation).

System	Male		Female	
	EER (%)	minDCF	EER (%)	minDCF
Total variability factor analysis + LDA	5.07	0.0516	7.40	0.0723
LPP + LDA	5.55	0.0492	7.65	0.0622
LPDP + LDA	4.23	0.0437	5.61	0.0585

Table 3. Comparison of EER and minDCF of LPDP, LPP, and total variability factor analysis (WCCN channel compensation).

System	Male		Female	
	EER (%)	minDCF	EER (%)	minDCF
Total variability factor analysis + WCCN	6.77	0.0532	9.32	0.0752
LPP + WCCN	5.08	0.0456	6.22	0.0540
LPDP + WCCN	4.48	0.0426	5.71	0.0512

Table 4. Comparison of EER and minDCF of LPDP, LPP, and total variability factor analysis (LDA + WCCN channel compensation).

System	Male		Female	
	EER (%)	minDCF	EER (%)	minDCF
Total variability factor analysis + LDA + WCCN	4.61	0.0502	6.14	0.0607
LPP + LDA + WCCN	4.43	0.0462	5.85	0.0577
LPDP + LDA + WCCN	4.02	0.0409	5.21	0.0517

5. Conclusion

On the basis of LPP, this paper introduced LPDP to speaker verification. LPDP makes full use of the speaker label information of the speech data to categorize and differentiate the neighborhood. It can overcome the shortcomings of the total variability factor analysis method and maintain the intrinsic local neighborhood relationship of in-class (same speaker) speech data and more comprehensively reflect the global and local structure of the speech data. It can also address the inadequacy of LPP and maximize the distance between out-of-class (different speakers) speech data to obtain the most discriminative feature vector and enhance the discriminative ability of the projection space, thereby improving the recognition performance of the system. Our future work will be devoted to enhance the discrimination of the embedding space and further improve the recognition performance of the system.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No.11704229).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Naika, R. (2018) An Overview of Automatic Speaker Verification System. In: *Intelligent Computing and Information and Communication*, Springer, Singapore, 603-610. https://doi.org/10.1007/978-981-10-7245-1_59

[2] Kinnunen, T. and Li, H. (2010) An Overview of Text-Independent Speaker Recognition: From Features to Supervectors. *Speech Communication*, **52**, 12-40. https://doi.org/10.1016/j.specom.2009.08.009

[3] Chen, C. and Han, J.Q. (2018) Partial Least Squares Based Total Variability Space Modeling for I-Vector Speaker Verification. *Chinese Journal of Electronics*, **27**, 1229-1233. https://doi.org/10.1049/cje.2018.06.001

[4] Luo, J., Leung, C.C., Ferras, M. and Barras, C. (2018) Parallelized Factor Analysis and Feature Normalization for Automatic Speaker Verification. *INTERSPEECH* 2018, Brisbane, 1409-1412.

[5] Su, H. and Wegmann, S. (2016) Factor Analysis Based Speaker Verification Using ASR. *INTERSPEECH* 2016, San Francisco, 2223-2227. https://doi.org/10.21437/Interspeech.2016-1157

[6] Zhang, X., Zou, X., Sun, M., Zheng, T.F., Jia, C. and Wang, Y. (2019) Noise Robust Speaker Recognition Based on Adaptive Frame Weighting in Gmm for I-Vector Extraction. *IEEE Access*, **7**, 27874-27882. https://doi.org/10.1109/ACCESS.2019.2901812

[7] Guo, W., Dai, L.R. and Wang, R.H. (2009) Speaker Verification Based on Factor Analysis and SVM. *Journal of Electronics & Information Technology*, **31**, 302-305.

[8] Ibrahim, N.S. and Ramli, D.A. (2018) I-Vector Extraction for Speaker Recognition Based on Dimensionality Reduction. *Procedia Computer Science*, **126**, 1534-1540.

https://doi.org/10.1016/j.procs.2018.08.126

[9] Mak, M., Pang, X. and Chien, J. (2016) Mixture of PLDA for Noise Robust I-Vector Speaker Verification. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, **24**, 130-142. https://doi.org/10.1109/TASLP.2015.2499038

[10] Lei, Y. and Hansen, J.H.L. (2010) Speaker Recognition Using Supervised Probabilistic Principal Component Analysis. *INTERSPEECH* 2010, Makuhari, 382-385.

[11] Yang, J., Liang, C., Yang, L., Suo, H., Wang, J. and Yan, Y. (2012) Factor Analysis of Laplacian Approach for Speaker Recognition. 2012 *IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, 25-30 March 2012, 4221-4224. https://doi.org/10.1109/ICASSP.2012.6288850

[12] Liang, C.Y., Yang, L., Zhao, Q.W. and Yang, Y.H. (2012) Factor Analysis of Neighborhood-Preserving Embedding for Speaker Verification. *IEICE Transactions on Information & Systems*, **95**, 2572-2576. https://doi.org/10.1587/transinf.E95.D.2572

[13] Liang, C.Y., Yuan, W.H., Li, Y.L., Xia, B. and Sun, W.Z. (2019) Speaker Recognition Using Discriminant Neighborhood Embedding. *Journal of Electronics & Information Technology*, **41**, 1774-1778.

[14] Chien, J. and Hsu, C. (2017) Variational Manifold Learning for Speaker Recognition. 2017 *IEEE International Conference on Acoustics, Speech and Signal Processing*, New Orleans, 5-9 March 2017, 4935-4939. https://doi.org/10.1109/ICASSP.2017.7953095

[15] Wu, D. (2015) Speaker Recognition Based on I-Vector and Improved Local Preserving Projection. In: *Proceedings of the* 2015 *Chinese Intelligent Automation Conference*, Springer, Heidelberg, 115-121. https://doi.org/10.1007/978-3-662-46469-4_12

[16] He, X.F., Yan, S.C., Hu, Y.X., Niyogi, P. and Zhang, H.-J. (2005) Face Recognition Using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**, 328-340. https://doi.org/10.1109/TPAMI.2005.55

[17] Hatch, A.O., Kajarekar, S. and Stolcke, A. (2006) Within-Class Covariance Normalization for SVM-Based Speaker Recognition. *INTERSPEECH* 2006, Pittsburgh, 1471-1474.

[18] Haeb-Umbach, R. and Ney, H. (1992) Linear Discriminant Analysis for Improved Large Vocabulary Continuous Speech Recognition. 1992 *IEEE International Conference on Acoustics, Speech, and Signal Processing*, San Francisco, 23-26 March 1992, 13-16. https://doi.org/10.1109/ICASSP.1992.225984

[19] Zhao, Z.H. and Hao, X.H. (2014) Linear Locality Preserving and Discriminating Projection for Face Recognition. *Journal of Electronics & Information Technology*, **35**, 463-467. https://doi.org/10.3724/SP.J.1146.2012.00601

[20] Wang, J.F. and Gao, Q. (2015) Discriminant Neighborhood Structure Embedding Using Trace Ratio Criterion for Image Recognition. *Journal of Computer & Communications*, **3**, 64-70. https://doi.org/10.4236/jcc.2015.311011

[21] The NIST Year 2010 Speaker Recognition Evaluation Plan. https://www.nist.gov/itl/iad/mig/speaker-recognition-evaluation-2010

[22] Scheffer, N., Ferrer, L., Graciarena, M., Kajarekar, S., Shriberg, E. and Stolcke, A. (2011) The SRI NIST 2010 Speaker Recognition Evaluation System. 2011 *IEEE International Conference on Acoustics, Speech and Signal Processing*, Prague, 22-27 May 2011, 5292-5295. https://doi.org/10.1109/ICASSP.2011.5947552

[23] Joacjims, T. (2018) SVM-Light Support Vector Machine. http://svmlight.joachims.org