

A Method to Integrate GMM, SVM and DTW for Speaker Recognition

Ing-Jr Ding, Chih-Ta Yen* and Da-Cheng Ou

Department of Electrical Engineering, National Formosa University, Yunlin, Taiwan, ROC.

Received 05 August 2013; received in revised form 22 August 2013; accepted 24 September 2013

Abstract

This paper develops an effective and efficient scheme to integrate Gaussian mixture model (GMM), support vector machine (SVM), and dynamic time wrapping (DTW) for automatic speaker recognition. GMM and SVM are two popular classifiers for speaker recognition applications. DTW is a fast and simple template matching method, and it is frequently seen in applications of speech recognition. In this work, DTW does not play a role to perform speech recognition, and it will be employed to be a verifier for verification of valid speakers. The proposed combination scheme of GMM, SVM and DTW, called SVMGMM-DTW, for speaker recognition in this study is a two-phase verification process task including GMM-SVM verification of the first phase and DTW verification of the second phase. By providing a double check to verify the identity of a speaker, it will be difficult for imposters to try to pass the security protection; therefore, the safety degree of speaker recognition systems will be largely increased. A series of experiments designed on door access control applications demonstrated that the superiority of the developed SVMGMM-DTW on speaker recognition accuracy.

Keywords: speaker recognition, Gaussian mixture model, support vector machine, dynamic time wrapping, SVMGMM-DTW

References

- [1] B. K. Sy, "Secure computation for biometric data security — application to speaker verification," *IEEE Systems Journal*, vol. 3, no. 4, pp. 451–460, 2009.
- [2] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
- [3] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using Gaussian mixture models," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 1, pp. 72–83, 1995.
- [4] L. Burget, P. Matejka, P. Schwarz, O. Glembek, and J. Cernocky, "Analysis of feature extraction and channel compensation in a GMM speaker recognition system," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 7, pp. 1979–1986, 2007.
- [5] J. C. Wang, C. H. Yang, J. F. Wang and H. P. Lee, "Robust speaker identification and verification," *IEEE Computational Intelligence Magazine*, vol. 2, no. 2, pp. 52–59, 2007.
- [6] J. Louradour, K. Daoudi and F. Bach, "Feature space mahalanobis sequence kernels: Application to SVM speaker verification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 8, pp. 2465–2475, 2007.
- [7] W. M. Campbell, J. P. Campbell, T. P. Gleason, D. A. Reynolds and W. Shen, "Speaker verification using support vector machines and high-level features," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 7, pp. 2085–2094, 2007.
- [8] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 1, pp. 43–49, 1978.

* Corresponding author. E-mail address: chihtayen@gmail.com

Tel.: 886-5-6315630; Fax: 886-5-6315609

- [9] C. H. You, K. A. Lee and H. Li, "GMM-SVM kernel with a Bhattacharyya-based distance for speaker recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 6, pp. 1300–1312, 2010.
- [10] C. H. You, K. A. Lee and H. Li, "An SVM kernel with GMM-supervector based on the Bhattacharyya distance for speaker recognition," *IEEE Signal Processing Letters*, vol. 16, no. 1, pp. 49–52, 2009.
- [11] C. Longworth and M. J. F. Gales, "Combining derivative and parametric kernels for speaker verification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 4, pp. 748–757, 2009.

