Preface

Speaker verification (SV) is the process of validating the claimed identity of an individual using his/her speech. State-of-the-art SV systems perform reasonably well when the spoken utterances are 'clean', i.e., free from any sort of distortions caused by external factors. However, the accuracy of such systems degrade severely when speech signals are distorted due to the presence of environmental or background noise. Robustness towards environmental noise is crucial for several SV applications, especially in hand-held devices where the background environments are uncontrolled, time-varying and unpredictable. The strategies used for handling background noise can be broadly categorized as (i) 'compensation' or 'adaptation' methods where features extracted in the test environment are denoised (compensated) to match with the training environment, or speaker model parameters estimated in the training environment are altered (adapted) to reflect the test environment, and (ii) developing speaker models or extracting features which are relatively immune towards the effect of background noise, by design.

This book explores novel methods from each of the aforementioned categories for robust SV in noisy background environments. A group of Gaussian mixture model (GMM) based stochastic feature compensation methods is proposed for SV in noisy environments. Alternatively, the robustness of GMM supervector-based speaker modeling approaches is explored for SV in noisy environments. A discriminative framework is used in which fixed-size vectors obtained by stacking GMM means (i.e., supervectors) or by total variability modeling (i.e., i-vectors) are in turn used for training speaker-specific support vector machines (SVMs). These SVMs are evaluated in noisy test environments. Training an ensemble of robust SVM classifiers using adaptive boosting is proposed for improving SV performance in noisy environments.

This book is mainly intended for researchers working on robust speaker recognition technologies. This book is also useful for the young researchers, who want to pursue research in speech processing with an emphasis on acoustic modeling and feature compensation. Hence, this may be recommended as the text or reference book for the postgraduate level advanced speech processing course. The book has been organized as follows:

Chapter 1 introduces the concept of speaker recognition (SR) and the various stages involved in the SR process. The application of robustness required for SR technologies has been emphasized. Chapter 2 provides a review of the diverse array of methods employed for robust speaker/speech recognition. Chapter 3 discusses the baseline SV systems developed using Gaussian mixture models. Chapter 4 explores stochastic feature compensation methods for robust SV in noisy environments. Chapter 5 explores robust speaker modeling methods for SV in noisy environments. Chapter 6 provides a brief summary and conclusion of the book with directions towards the scope for possible future work.

We would especially like to thank all professors of the School of Information and Technology, IIT Kharagpur, for their moral encouragement and technical discussions during the course of editing and organization of this book. Special thanks to our colleagues at Indian Institute of Technology, Kharagpur, India, for their cooperation to carry out the work. We are grateful to our parents and family members for their constant support and encouragement. Finally, we thank all our friends and well-wishers.

Kharagpur, India

K. Sreenivasa Rao
Sourjya Sarkar

http://www.springer.com/978-3-319-07129-9

Robust Speaker Recognition in Noisy Environments

Rao, K.S.; Sarkar, S.

2014, XII, 139 p. 31 illus., 25 illus. in color., Softcover
ISBN: 978-3-319-07129-9