Autor :

Prof. Dr.-Ing. Hans-Günter Hirsch

Lehrgebiet Nachrichtenübertragung, Lineare Systeme und Netzwerke

Fachbereich Elektrotechnik und Informatik

Hochschule Niederrhein

Schriftenreihe des Fachbereichs Elektrotechnik und Informatik

herausgegeben von

Prof. Dr.-Ing. Hans Dieter Beims
Fachbereich Elektrotechnik und Informatik
Hochschule Niederrhein

Band 6/2008

Hans-Günter Hirsch

Automatic Speech Recognition in Adverse Acoustic Conditions

Shaker Verlag
Aachen 2008

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Printed in Germany.

ISBN 978-3-8322-7226-5
ISSN 1610-9392

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • e-mail: info@shaker.de

Abstract

Most of today's speech recognition systems are based on a cepstral analysis scheme for representing short-term speech features. Furthermore, the statistical approach of modeling speech units by Hidden Markov Models (HMMs) and the application of the Viterbi algorithm are state-of-the-art for the pattern recognition in this field. These analysis and modeling techniques are presented as an introduction to the problem of recognizing speech in adverse acoustic conditions.

Analyzing the speech input in application scenarios of recognition systems the speech signal is affected by the acoustic environment. An overview about the combination of possible distortion effects is given. A simulation tool is presented that has been developed to allow the generation of distorted speech data for a lot of different acoustic environments. A new database is described that has been created with this tool. It is publicly available for comparative investigations on the recognition of distorted speech data.

After an introduction to the principal approaches for improving the recognition in adverse conditions a few details of a standardized method are presented that aims at the extraction of robust features. The alternative approach of adapting certain HMM parameters in adverse conditions is introduced by presenting details about a new method to adapt the HMM parameters in case of a hands-free speech input in rooms. Furthermore, it is shown that this technique can be combined with an adaptation to stationary background noise and unknown frequency characteristics. We demonstrate the efficiency of the new approaches by evaluating the results for recognizing artificially distorted data as well as speech data that have been recorded in noisy environments.

Contents