

Speech recognition in a dialog system: from conventional to deep processing

A case study applied to Spanish

Aldonso Becerra¹  · J. Ismael de la Rosa¹ · Efrén González¹

Received: 24 January 2017 / Revised: 5 July 2017 / Accepted: 28 August 2017/

Published online: 6 September 2017

© Springer Science+Business Media, LLC 2017

Abstract The aim of this paper is to illustrate an overview of the automatic speech recognition (ASR) module in a spoken dialog system and how it has evolved from the conventional GMM-HMM (Gaussian mixture model - hidden Markov model) architecture toward the recent nonlinear DNN-HMM (deep neural network) scheme. GMMs have dominated for a long time the baseline of speech recognition, but in the past years with the resurgence of artificial neural networks (ANNs), the former models have been surpassed in most recognition tasks. An outstanding consideration for ANNs-based acoustic model is the fact that their weights can be adjusted in two training steps: i) initialization of the weights (with or without pre-training) and ii) fine-tuning. To exemplify these frameworks, a case study is realized by using the Kaldi toolkit, employing a mid-vocabulary with a personalized speaker-independent voice corpus on a connected-words phone dialing environment operated for recognition of digit strings and personal name lists in Spanish from Mexico. The obtained results show a reasonable accuracy in the speech recognition performance through the DNN acoustic modeling. A word error rate (WER) of 1.49% for context-dependent DNN-HMM is achieved, providing a 30% relative improvement with regard to the best GMM-HMM result in these experiments (2.12% WER).

Keywords Speech recognition · Neural networks · Gaussian mixture models · Hidden Markov models · Deep learning · Spoken dialog system

✉ Aldonso Becerra
a7donso@uaz.edu.mx

J. Ismael de la Rosa
ismaelrv@ieee.org

Efrén González
gonzalez_efren@hotmail.com

¹ Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Av. López Velarde No. 801, Col. Centro, C.P. 98068, Zacatecas, México