**A NONLINEAR AUTOREGRESSIVE MODEL
FOR SPEAKER VERFICATION1**

Sundararajan Srinivasan2, Tao Ma3, Georgios Lazarou4 and Joseph Picone5

*Abstract*— Gaussian Mixture Models (GMM) have been the most popular approach in speaker recognition and verification for over two decades. The inefficiencies of this model for signals such as speech are well documented and include an inability to model temporal dependencies that result from nonlinearities in the speech signal. The resulting models are often complex and overdetermined, which leads to a lack of generalization. In this paper, we present a nonlinear mixture autoregressive model (MixAR) that attempts to directly model nonlinearities in the trajectories of the speech features. We apply this model to the problem of speaker verification. Experiments with synthetic data demonstrate the viability of the model. Evaluations on standard speech databases, including TIMIT, NTIMIT, and NIST-2001, demonstrate that MixAR, using only half the number of parameters and only static features, can achieve a lower equal error rate when compared to GMMs, particularly in the presence of previously unseen noise. Performance as a function of the duration of both the training and evaluation utterances is also analyzed.

*Keywords*— Gaussian mixture models, mixture autoregressive model, nonlinear statistical models, speaker verification

Manuscript submitted February 15, 2013.

1. This material is based upon work supported by the National Science Foundation under Grant No. IIS-0414450. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.
2. S. Srinivasan is with Nuance Communications Inc., 1198 East Arques Avenue  Sunnyvale, CA 94085, USA (phone: 408-992-6243; email: sundararajan.srinivasan@gmail.com).
3. T. Ma is with Siri at Apple Inc., 2 Infinite Loop, mailstop 302-4APP, Cupertino, California 95014, USA (phone: 408-643-5909; email: tma@apple.com).
4. G. Lazarou is with The New York City Transit Authority, 30-74 38th Street, Apt 1A, Astoria, New York, New York, USA 11103 (phone: (662) 617-2064; email: glaz@ieee.org).
5. J. Picone is with the Department of Electrical and Computer Engineering at Temple University, 1947 North 12th Street, Philadelphia, Pennsylvania 19027 USA (phone: 215-204-4841; fax: 215-204-5960; email: joseph.picone@isip.piconepress.com).