**Continuous Speech Recognition
Using Linear Dynamic Models1**

Tao Ma2, Sundararajan Srinivasan3, Georgios Lazarou4 and Joseph Picone5

*Abstract*— Hidden Markov models (HMMs) with Gaussian mixture distributions rely on an assumption that speech features are temporally uncorrelated, and often assume a diagonal covariance matrix where correlations between feature vectors for adjacent frames are ignored. A Linear Dynamic Model (LDM) is a Markovian state-space model that also relies on hidden state modeling, but explicitly models the evolution of these hidden states using an autoregressive process. An LDM is capable of modeling higher order statistics and can exploit correlations of features in an efficient and parsimonious manner. In this paper, we present a hybrid LDM/HMM decoder architecture that postprocesses segmentations derived from the first pass of an HMM-based recognition. This smoothed trajectory model is complementary to existing HMM systems. An Expectation-Maximization (EM) approach for parameter estimation is presented. We demonstrate a 13% relative WER reduction on the Aurora-4 clean evaluation set, and a 13% relative WER reduction on the babble noise condition.

*Keywords*— Linear dynamic models, nonlinear statistical modeling, speech recognition, acoustic modeling

Manuscript submitted February 1, 2013. Revised May 20, 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. 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).
3. S. Srinivasan is with Nuance Communications Inc., 1198 East Arques Avenue  Sunnyvale, CA 94085, USA (phone: 408-992-6243; email: sundararajan.srinivasan@gmail.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).