

# TIDigits corpus classification based on time series modeling

Bonny Banerjee

Soundararajan Srinivasan

# Speech Clustering/Recognition

- Treat every word in the corpus as a class that is modeled as a time series by using HMM.
- Continuous speech is given by a sequence of discrete parametric representations.
- Bayes' rule is used to formulate the problem as an diagnostic inference one.
- The model is found as the MLE (normally distributed) from the training data.

# Comparisons

---

- **If time permits we plan to do the same by using ANN to do the same clustering and compare the performance.**
- **Comparison can also be made with the performance of human subjects.**

**FOR MORE INFO...**

<http://www ldc.upenn.edu/Catalog/LDC93S10.html>