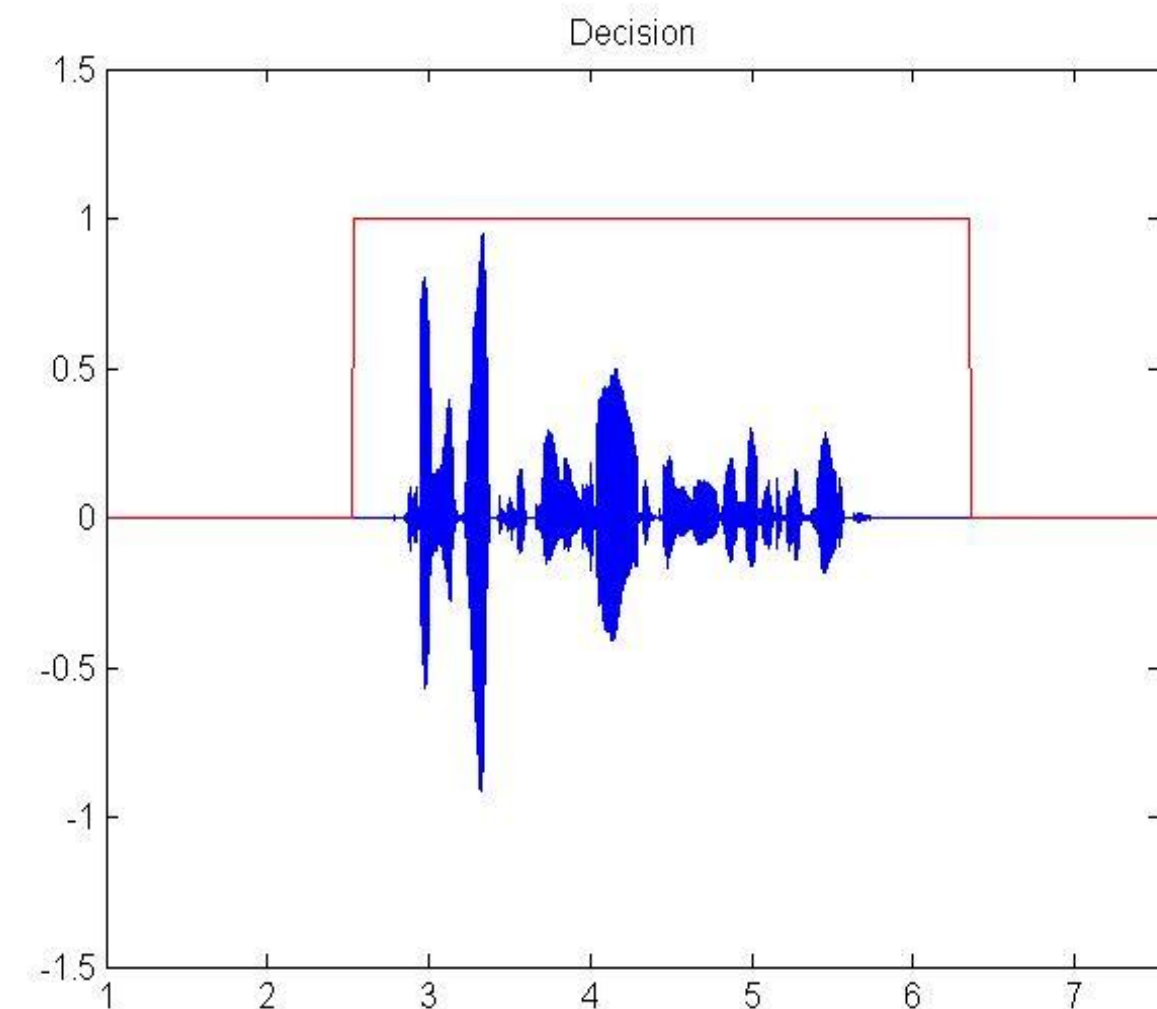


Long-term speech variability voice activity detection

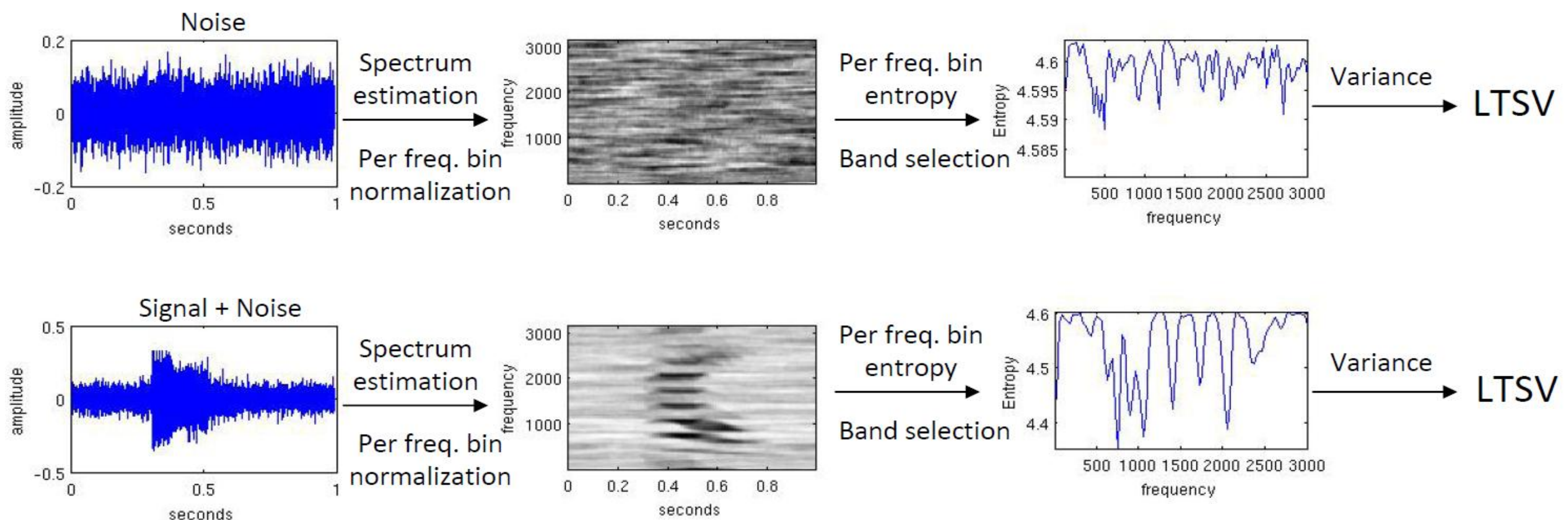
Andreas Tsiartas, EE/SAIL Lab

Motivation & Introduction

- **Goal:**
 - Detect the presence of speech or noise.
- **Importance:**
 - Machines focus on audio only when necessary.
- **Innovation:**
 - Exploit long-term speech properties
 - Separate speech and noise based on variability
 - Speech recognizers memory limitations.
 - Efficient implementations

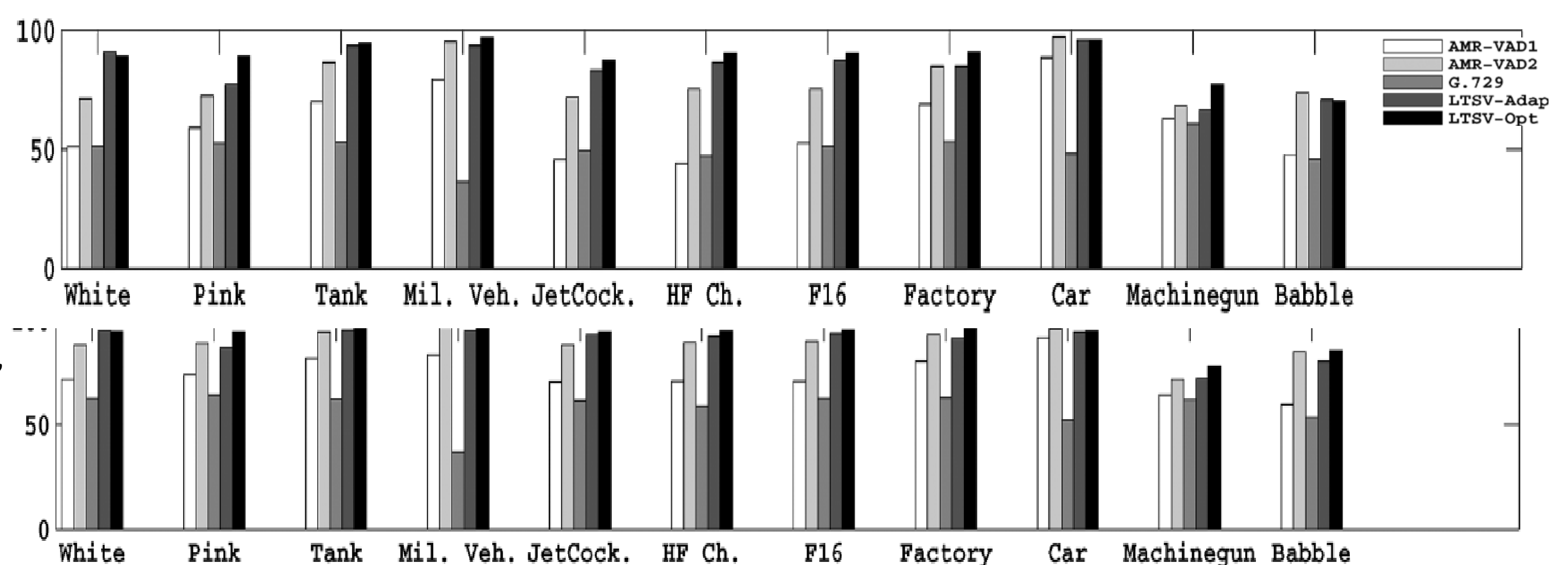


Approach



Experiments

- Tested on NOISX-92 database
- Against popular AMR-VAD2
- Per frame accuracy of eleven noises of five VAD schemes
 - Top figure: Accuracy at -10dB
 - Bottom figure: Accuracy averaged over -10dB, -5dB, 0dB, 5dB, 10 dB



Discussion

- Robust to SNR variations (theoretical results)
- Good discriminative power between Speech and non-Speech
- Shown multi-resolution analysis improves performance

Future work

- Feature to capture energy information
- Test various machine learning models
- DARPA RATS competition