SUPERVISED SIGNAL-TO-NOISE RATIO ESTIMATION OF SPEECH SIGNALS

Pavlos Papadopoulos, Andreas Tsiartas, James Gibson, Shrikanth Narayanan
Speech Analysis and Interpretation Laboratory (SAIL), Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089
{ppapadop,tsiartas,jjgibson}@usc.edu, shri@sipi.usc.edu

Introduction

- Supervised global SNR Estimation
- Different kinds of noise
- Assumptions:
  - Independence of speech and noise signals
  - Both Speech and Noise are zero-mean
  - Additive Noise
- Features that capture speech in a signal
- Create Estimators based on features
- Regression model on the estimators to estimate SNR

Speech SNR

$$\mathrm{SNR} = 20 \cdot \log_{10} \sqrt{\frac{\frac{1}{N}\sum_{i=1}^{N} s^2(i)}{\frac{1}{N}\sum_{i=1}^{N} n^2(i)}}$$

$$= 10 \cdot \log_{10} \frac{\mathbf{P}(S)}{\mathbf{P}(N)}$$

$$= 10 \cdot \log_{10} \frac{\mathbf{P}(X) - \mathbf{P}(N)}{\mathbf{P}(N)}$$

Features

- Energy
- Long-Term Signal Variability (LTSV)
- Pitch
- Voicing Probability

Estimators

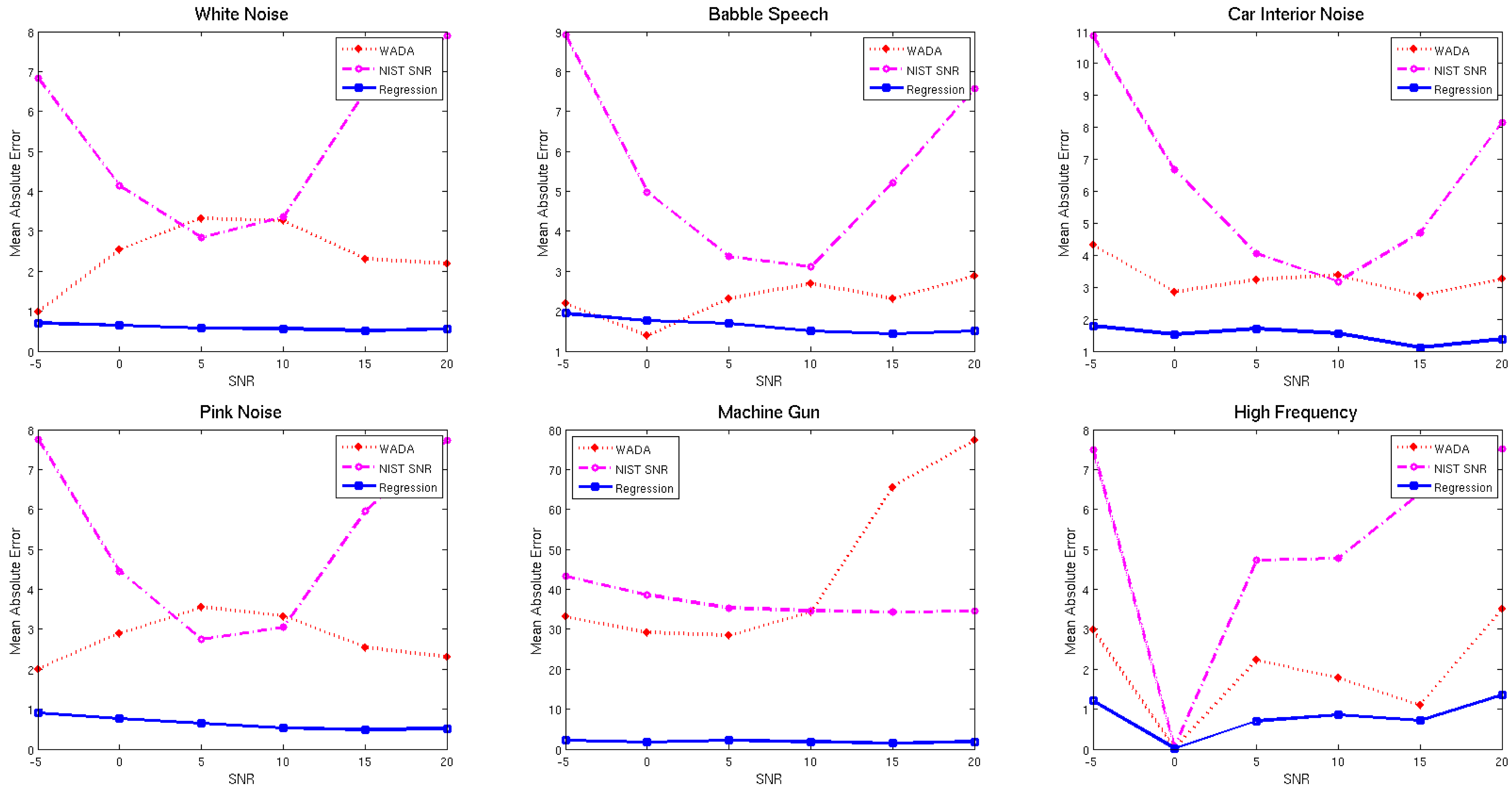
$$E_{a-b}^{c-d} = 10 \cdot \log_{10} \frac{\mathbf{E}(X_c^d) - \mathbf{E}(X_a^b)}{\mathbf{E}(X_a^b)}$$

$$L_{a-b}^{c-d} = 10 \cdot \log_{10} \frac{\mathbf{E}(L(X)_c^d) - \mathbf{E}(L(X)_a^b)}{\mathbf{E}(L(X)_a^b)}$$

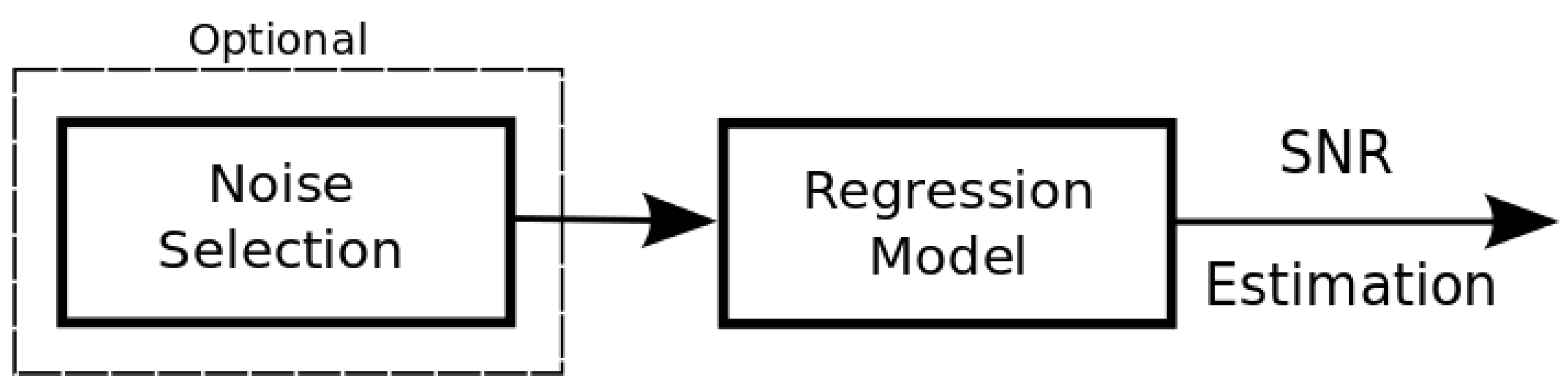
$$P_{a-b}^{c-d} = 10 \cdot \log_{10} \frac{\mathbf{E}(P(X)_c^d) - \mathbf{E}(P(X)_a^b)}{\mathbf{E}(P(X)_a^b)}$$

$$C_{a-b}^{c-d} = 10 \cdot \log_{10} \frac{\mathbf{E}(C(X)_c^d) - \mathbf{E}(C(X)_a^b)}{\mathbf{E}(C(X)_a^b)}$$

SNR Estimation Results for different kinds of noise



System Overview



- 13 MFCC features
- KNN Classifier (K=20)
- Good Performance when the noise is included in the training set
- Poor Generalization
- Regression Model for each noise type
- White, Babble speech, Machine Gun, Car Interior, Pink
- Independent variables are individual SNR estimators

Conclusions and Future Work

- Accurate SNR Estimation
- Independent of noise type
- Better Performance compared to NIST SNR and WADA
- Select more robust features for noise classification
- Test different classifiers for noise model selection