USC University of Southern California

USC Viterbi School of Engineering

Ming Hsieh Department of Electrical Engineering

Novel Variations of Group Sparse Regularization Techniques with Applications to Noise Robust Automatic Speech Recognition

Qun Feng Tan and Shrikanth Narayanan, SAIL

Motivation & Introduction

- Sparse representation approach to speech in field of Missing Data Techniques (MDT)
- Application of regularization to spectral denoising for better MFCC reconstruction
- Grouping of dictionary to capitalize on structure and variability
- Demonstrate improved results over non-grouped version of algorithms

Methodology

- Baseline: Elastic Net formulation
- Choice of Elastic Net over LASSO: Ability of Elastic Net to handle collinear dictionaries, which happens with speech spectral elements due to energy concentration in similar regions of time-frequency plane of similar sounding utterances
- Proposal: Grouped Elastic Net
- New algorithm takes into account dictionary partitioning
- Consider 2 types of intuitive partitions: K-means and speaker identity groupings

Grouped Elastic Net

Step 1: Initialize with Least-Squares solution

Step 2: For each group i=1 to N, we define the residual as

$$r = f - \sum_{j \neq i} \phi_j x_j$$

We proceed to solve the following formulation:

$$\min ||r - \phi_i x_i||_2^2 + \lambda_1 s_i ||x_i||_2^2 + \lambda_2 r_i ||x_i||_1$$

Step 3: Iterate over all groups repeatedly

Experimental Results for SNR 5 corruption

Algorithm	*ASR Accuracies in percent*
Original Noisy	43.82
LASSO	64.24
Elastic Net	66.36
Grouped Elastic Net with Kmeans	**67.42**
Grouped Elastic Net with Speaker Identity	**67.78**

Discussion

- From our results, we see that the Grouped Elastic Net outperforms the Elastic Net algorithm and also existing grouped regularization algorithm Group Sparse LASSO.
- This shows that grouping the dictionary appropriately helps in the speech spectral denoising process, which ultimately leads to improved ASR performance.
- Due to the fact that the Grouped Elastic Net is based upon the Elastic Net formulation, we are also retaining the complexity advantages of the original algorithm, making the implementation efficient and having potential to be implemented in real-time systems.

qtan@usc.edu, shri@sipi.usc.edu

Ming Hsieh Institute
Ming Hsieh Department of Electrical Engineering