



1. INTRODUCTION

■ **THE ACT OF SPEAKING:** It usually occurs at **noisy acoustical environments**

- ❖ **HUMAN BEING:** It has a high capacity of recognition in such ones
- ❖ **AUTOMATIC SPEECH RECOGNITION (ASR) SYSTEMS:** We need to adapt ASR systems to such conditions

■ SOME APPROACHES TO IMPROVE MATCHING:

- ❖ Model adaptation
- ❖ Hidden Márkov Model (HMM) decomposition
- ❖ Feature enhancement or feature compensation

■ **OUR GOAL:** Improve the performance of an ASR system, trained with clean data, by feature enhancement

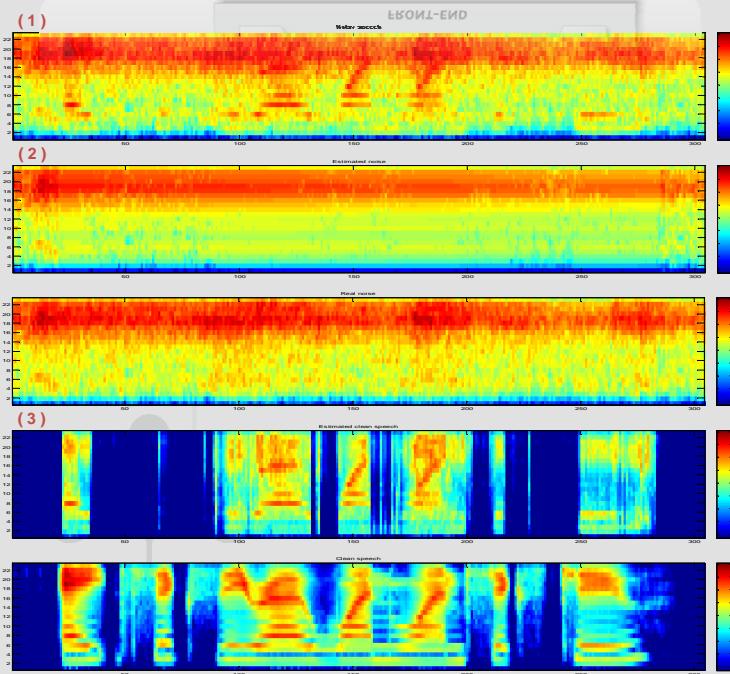
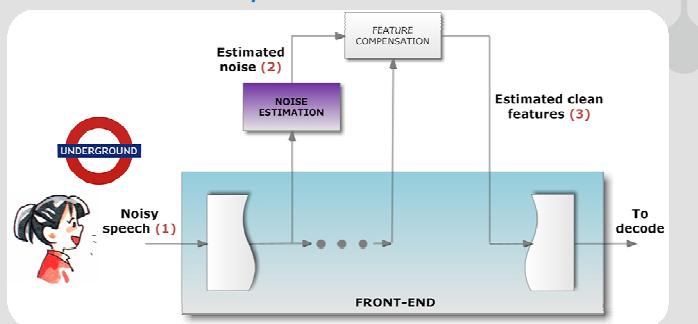
2. LINE OF WORK

■ Some of **feature enhancement algorithms** need an estimation of the **noise** present in the noisy utterance

■ **MAIN LINE:** Research and development of **new noise estimation algorithms** for feature enhancement

■ **OBJECTIVE:** Overcome the state-of-the-art in this topic

❖ Framework and example

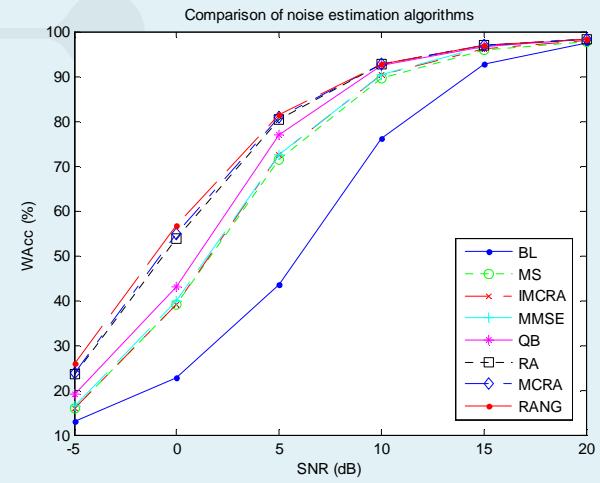


3. A COMPARISON OF SOME STATE-OF-THE-ART

■ We evaluate some noise estimation algorithms on test sets *A* and *B* from **Aurora-2** database:

- ❖ **Minimum statistics, IMCRA, MMSE (Hendriks), quantile-based, recursive averaging, MCRA and Rangachari's algorithm**

■ The **noise estimates** are used in **zeroth-order Vector Taylor Series expansion (VTS-0)** in order to enhance noisy log-Mel features



Method / SNR (dB)	-5	0	5	10	15	20	clean	Avg. 0-20
Base-Line	12.94%	22.88%	43.44%	76.18%	92.85%	97.40%	99.10%	66.55%
Minimum Statistics	15.83%	39.00%	71.38%	89.63%	95.93%	97.84%	98.95%	78.76%
IMCRA	16.03%	38.97%	72.58%	90.32%	96.13%	97.90%	99.08%	79.18%
MMSE Hendriks	16.43%	39.84%	72.46%	90.47%	96.72%	98.22%	99.08%	79.54%
Quantile-Based (Pow.)	18.98%	43.14%	76.89%	92.41%	96.80%	98.31%	99.07%	81.51%
Rec. Av. (Mel)	23.43%	53.85%	80.50%	92.67%	97.05%	98.22%	99.08%	84.46%
MCRA (Mel)	23.89%	54.89%	80.80%	92.67%	96.86%	98.17%	99.03%	84.68%
Rangachari (Mel)	25.82%	56.77%	81.46%	92.73%	96.87%	98.24%	99.13%	85.21%

4. CONCLUSIONS

- Feature enhancement techniques, that try to estimate the clean features, increase the performance of an ASR system that is trained with clean data
- It is important to have an appropriate noise estimation in order to improve the performance of the used feature enhancement method
- The obtained results give us some cues to get deep in the research and development of new noise estimation algorithms that are based in those aspects that produce a good performance

■ Contact information

Iván López Espejo

- ❖ Department of Signal Theory, Telematics and Communications
- ❖ Faculty of Sciences, Campus Fuentenueva
- ❖ University of Granada
- Tel: +34 958 240845
- E-mail: iloes@ugr.es
- Web: <http://www.ugr.es/~iloes>