Boost Intelligibility of Speech in Noise - Real-time Implementation of SSDRC

Vassilis Tsiaras[1], Tudor-Cătălin Zorilă[2], Yannis Stylianou[3] and Masami Akamine[4]

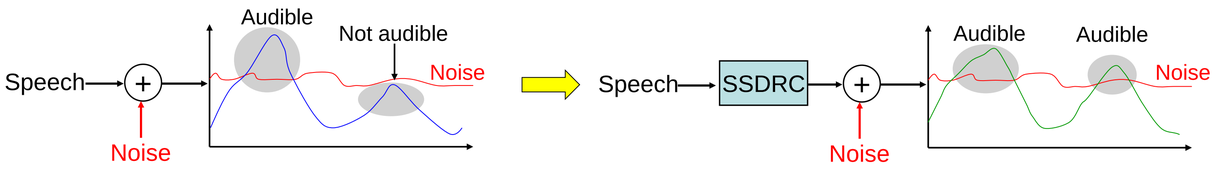
[1]Technical University of Crete, Chania, Greece, [2]University of Crete, Heraklion, Greece
[3]Toshiba Research Europe Limited, Cambridge Research Laboratory, Cambridge, UK, [4]Toshiba Research and Development Center, Kawasaki, Japan

Background

- **Effective Speech Communication in Noisy Environments**
  - Detect speech in Noise plays a significant role in our communication with others
  - Speech produced under real conditions is not always intelligible
- **Reduce Sound pollution**
  - Increase Intelligibility of Speech in Noise without Increasing the Volume.

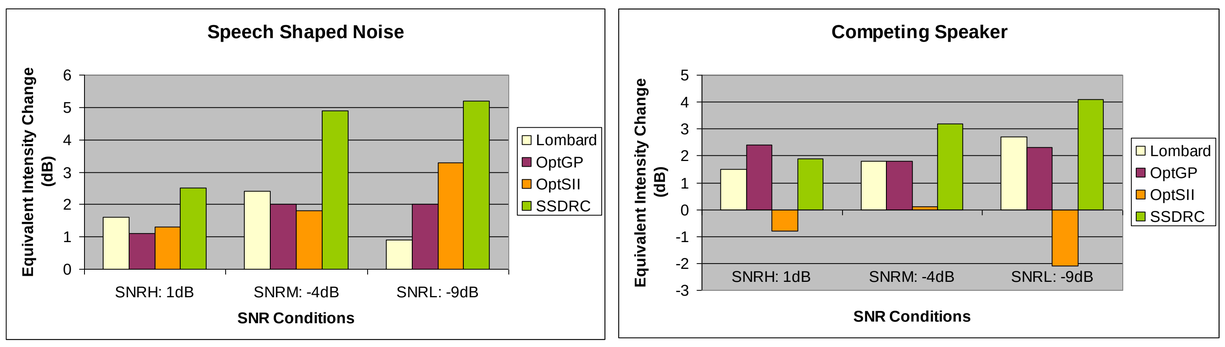
Our Suggestion: SSDRC

- **Spectral Shaping and Dynamic Range Compression (SSDRC):**
  - Protect speech information important for human perception, avoid picky signals.
  - Optimize energy redistribution according to properties of the human auditory system.



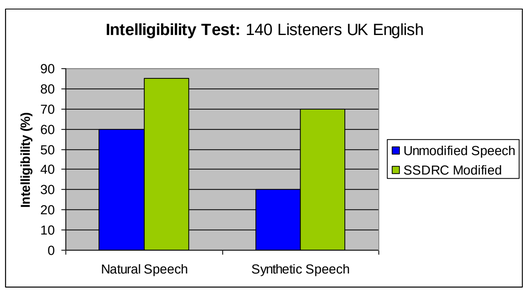
Benchmarking:

For Natural Speech



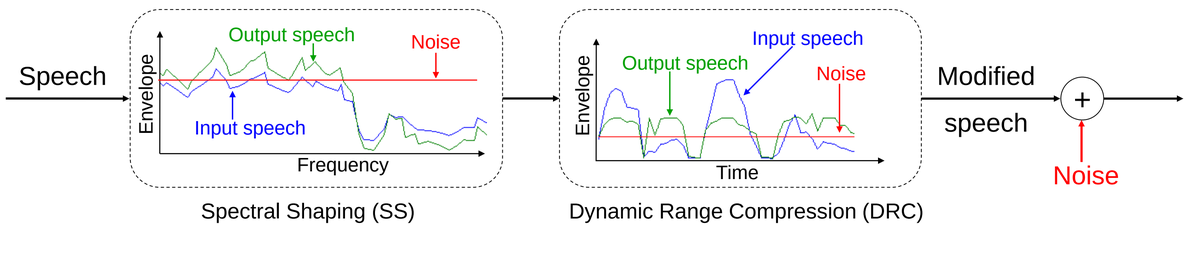
Equivalent intensity change for each modification type relative to natural plain speech

Comparing Natural and Synthetic Speech



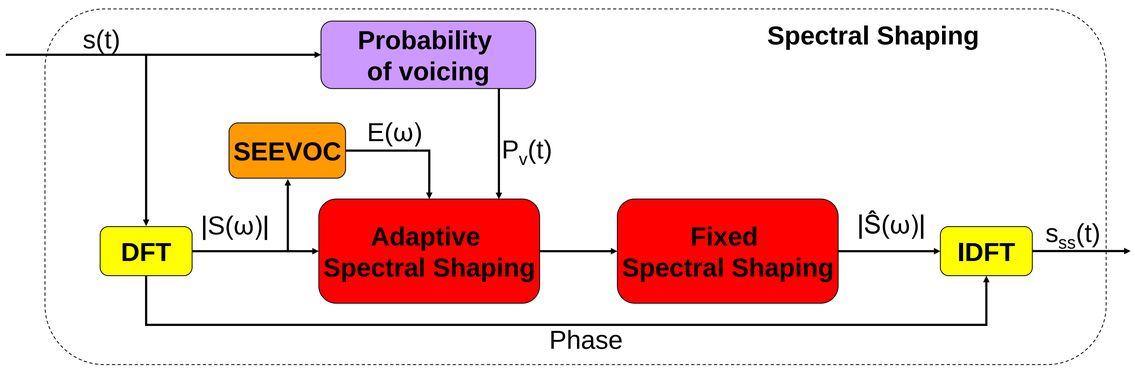
SSDRC: Technical Details

- The SSDRC algorithm consists of two sub-systems connected in cascade form.

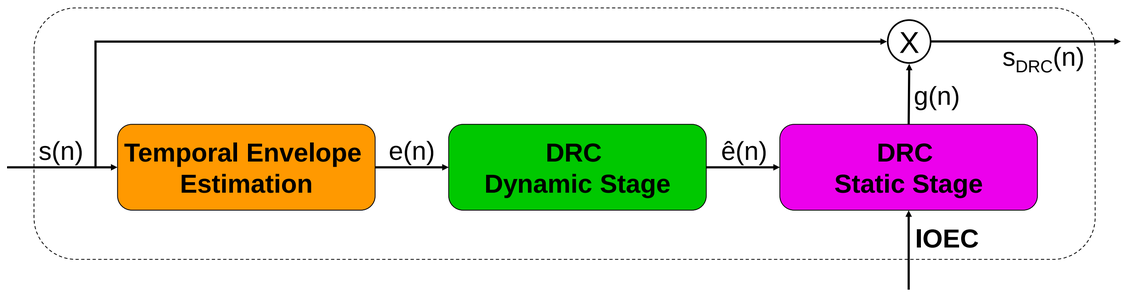


Spectral Shaping (SS) — Dynamic Range Compression (DRC)

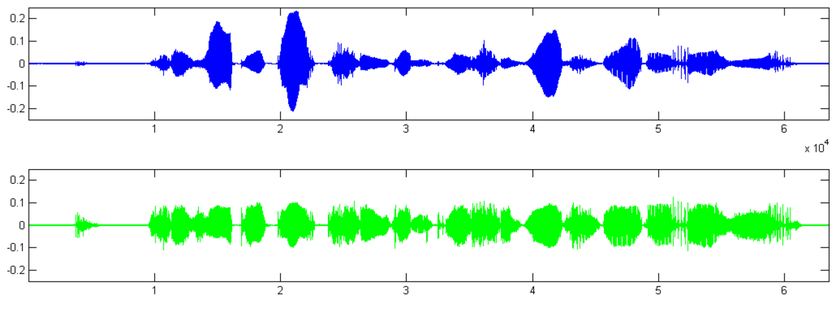
- **Spectral Shaping** is designed to re-adjust signal's spectral energy following results from clear and Lombard speech studies.



- **Dynamic Range Compression** is aimed at re-arranging the energy of speech waveform over time such that low energy segments (e.g., nasal, onsets and offsets) are amplified, while more energetic areas (sonorant sounds) are attenuated.

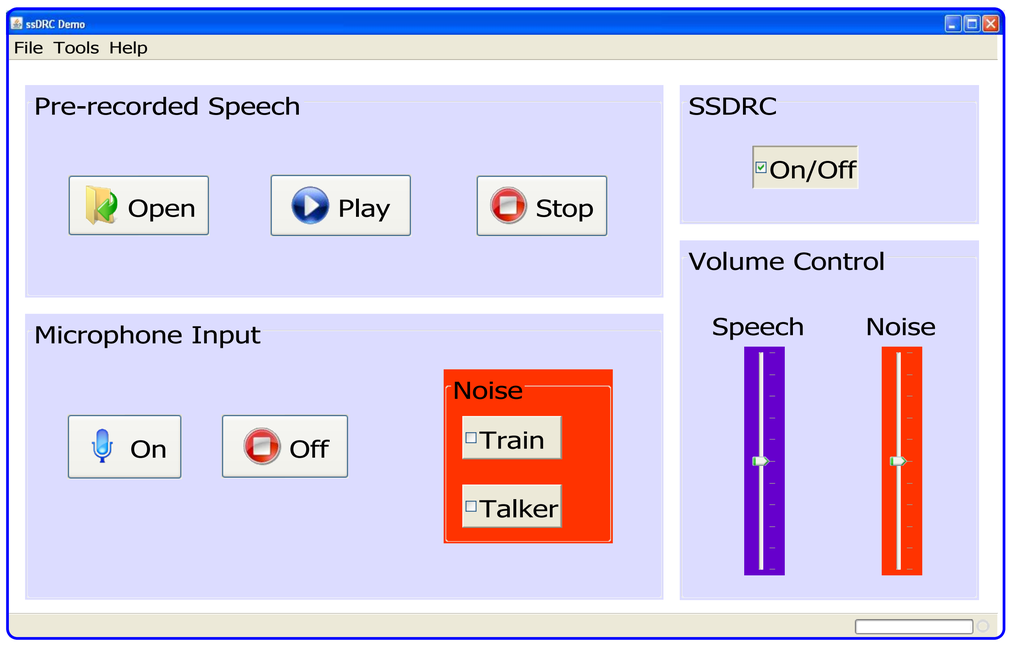


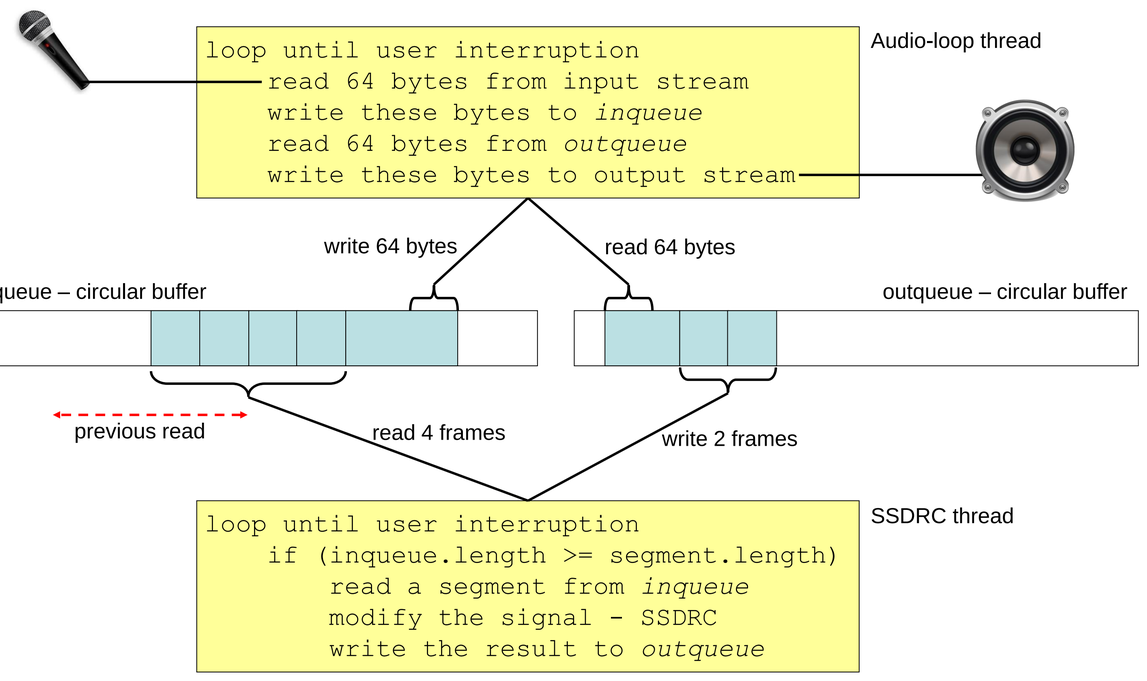
- Example



Real-time SSDRC Implementation

- The (rt)SSDRC library is written in C and is compiled into a shared library.
- The demo program is written in Java and calls the (rt)SSDRC library through the Java Native Interface.
- The Java program starts a thread that handles the input/output operations and a thread that runs the (rt)SSDRC.





Real-time SSDRC issues

➢Speech Segmentation

- Two problems must be addressed when selecting a segmentation strategy:
  - attenuated energy at the two ends of a segment
  - discontinuities between consecutive segments
- Both problems can be solved if **consecutive segments overlap**.

➢Non-causal Operations

- Baseline SSDRC contains three non-causal operations which require a prior knowledge of the whole speech signal:
  - computing the normalization constant for the probability of voicing
  - computing the maximum value of the signal's time envelope
  - computing the global multiplication constant that is used to preserve the energy of original and enhanced speech signals.
- Real-time SSDRC provides alternatives for all three non-causal operations
  - initialization values for all three parameters are computed from their statistics
  - the parameters are then updated as more frames of input speech become available

Applications