**General ideas;**

**IMCRA图**

**Maximum likelihood**

**Introduction**

Present the overview;

Introduce the background knowledge about the CI: function, limits…

Talk about the progress I’ve made so far, including…

**Importance**: the cochlear implant (CI) is a device that can help the patients with profound hearing loss to regain a part of the ability.

**Components**: it consists mainly of 1) microphone 2) audio processor 3) electrodes implanted directly on the cochlear to stimulate auditory nerves.

**Problem:** it works fine in quiet & stationary environments, however in non-stationary environment, the background noise becomes challenging for audio processors to distinguish speech from noise.

**SCS（EP intro）:** The process of transforming audio signals into excitation patterns (EP) in cochlea is called “sound coding strategy” (SCS), among, many SCSs, the one called Advanced Combinational Encoder (ACE) is choosed.

**Idea:** However, the received noise in two sides are different, which lead to the necessity of distributing a Master/Slave relationship between two sides based on the signal-to-noise ratio (SNR). That is, the side with better SNR obtain the dominance to decide the band selection for both side.

**Considerations:** This constant connection between two sides requires additional processing requires additional energy consumption, therefore, the efficiency of coding and comparing must be considered.

**Work so far:** There exists already an efficient codec for excitation patterns (EP) instead of audio signals.

**Motivation:** explore further possibility to reduce the energy consumption, it would be optimal that man can compare SNR by just EP instead of audio signals.

In short, this approach generates EP signals by: 1) perform short-time Fourier transformation (STFT) of each frame and group them into 22 bands; 2) select the information from certain bands with bigger DFT power, the so-called “N-of-M selection”; 3) mapping them to electric stimulations through a logarithmic growth function with thresholding.

**So far**

Until now I have found three algorithms for noise estimation with available code, namely

“Improved Minimum Controlled Recursive Averaging (IMCRA)” by Cohen, 2003,

|  |
| --- |
|  |
| - bayes’ rule  - what’s apriori/aposteriori SNR?  - how is recursive smoothing done?  \*  - why calculate sap first?  - SPP可使用MMSE-SSP里的定义 |

“MMSE-based Noise Tracking with Bias Compensation (MMSE\_BC)” by Hendricks, 2010

|  |
| --- |
|  |
| - maximum likelihood?  - MMSE? how defined to estimate noise psd?  - direct-decision?  - Safety-net? |

1. MMSE-estimation(baye’s rule): define the expected noise power by “conditional noise power estimate” and both noise & speech follows Gaussian distribution.

2. Bias is due to the DD approach’s understimation，DD使用了rough decision

“Unbiased MMSE-based Noise Power Estimation using Speech-Presence Probability (MMSE\_SSP)” by Gerkmann, 2012.

|  |
| --- |
|  |
| - why unbiased?  - how to replace safety net?  - difference of spp in [1] and [3]? (definition is the same (Bayes), but derivation is different: MMSE vs. ) |
| Backup  - Bayes’ rule: E(Y|N^2)、E(Y) 都是 |

All of these algorithms follow the similar procedure of “”:

- Perform STFT and get the frame-wise spectral information.

- compute a-priori and a-posteriori SNR (ζand ξ), maximum-likelihood estimation (MLE) is a usual method.

- estimate the prior knowledge of noise power E{N2|y, ξ}, using a-posteriori SNR and the noise power estimation of previous frame.

- Judging the presence of speech, which can be achieved by, e.g. “Direct-Decision” or “Soft-Decision/ Speech-Presence Probability” approach.

- Store the result for this recursive computation, and move on to the next frame till the end.

**How modified for EP and how EP is treated？**

As mentioned above, the EPs are computed from PSD whose first two frequency bins of PSD are ditched (only consider the 3th ~ 65th frequency-bins). They are firstly grouped into 22 bands as envelop amplitudes and select 8~12 biggest ones.

The idea of employing these envelops is simply treating them as a new kind of PSD without DC. So when they serve as input of those algorithms, the first procedure, namely STFT is not needed, since the envelops already provide the PSD.

The rest of the process is basically similar. And in order to get reasonable results, many pre-set parameters, e.g. smoothing factor, gain values or bias compensation will also need to be reconsidered.

**Preliminary Results**

CCITT-noise: the spectral characteristics of the CCITT-noise are to simulate a very large number of competing speakers comparable with situations found in train stations or football stadiums.

Testing

**SegSNR**

**LogErr**：over→speech attenuated；under→music noise

**STD**：n-1 → unbiased estimation, when actual **μ** is unknown.

**SE**：std/sqrt(n)

**Confidence interval**: 95% chance that the population mean lies within this range.

Central Limit Theorem

EP

- stimulation rates (8~12)

- DFT resolution

In the future

Testing Scenarios

In testing scenarios, the ground truth is known, since the noisy signals are produced by clean speech and specific noise/ mask. Therefore the instant SNRs can be calculated beforehand using STFT as the standard answer for evaluating.

1. 同文件同snr：imcra和imcraEP的瞬时SNR对比看哪个更明显

2. 同文件不同snr：是否在不同噪音等级中区别足够大

3. 同文件同snr同imcraEP：

Statistical Evaluation

0. In all those algorithms, noise and speech power distribution follows Gaussian distribution and independent.

-

- the one with the most stable outputs(smallest std).

2.

**Time plan**

**So far:**

**11.01~ 11.15**

Review the literature for acquaintance with the background knowledge and related studies, and search for candidate algorithms for noise estimation.

**11.16~ 11.30**

Work on the baseline algorithm, i.e., modify the available code respectively for the input form of normal sound-wave signals and EP signals, in order to get instant SNRs for each frame.

**12.01~ 12.15**

Continue to explore two other algorithms and perform the same modification for EP signals.

**Planned:**

**12.16~ 01.15**

Generate testing materials including: 1) simulate real-world stereo speeches using HRTR function 2) converting sound-wave files into EPs;

Design the test scenarios and statistical evaluation;

Improve the code of all algorithms correspondingly.

**01.16~ 01.31**

Perform preliminary tests using the three algorithms on respectively sound-wave signals for comparison.

**02.01~ 02.14**

Exploit the possibility of noise level estimation solely from the knowledge of the band selection.

If it is possible, then evaluating the reliability of this method.

**02.15~ 02.29**

Perform further tests on three algorithms as well as the band-selection-based methods.

**03.01~ 03.31**

Formulate and refine the thesis.