GRP Progress Report

Student: Ying Dongwen Supervisor: Dang Jianwu

1. The aim of my research

In recent years, ASR (automation speech recognition) technology is widely used in our daily life. Commercial systems have been developed to handle small to medium vocabularies with moderate performance. Large vocabulary systems have been developed and demonstrated under laboratory conditions. Although these systems are demonstrated well in laboratory environment, their systems' performance suffer substantial degradations in recognition accuracy when there is any kind of difference between the conditions in which the system is trained and the conditions in which the systems is finally tested. Our research purpose is to develop a noise-robust algorithm to make ASR system robust to noise. Although the system runs in adverse environment, the ASR system still works well.

2. My approach

In our research, we try two ways to resolve this problem. One is to build a noise reduction system; the other is to catch the noise-robust feature for ASR system. These two ways are based on the online estimation of noise.

In the first way, we try to find out how noise affect speech and to which extent the speech is affected. We provide a concept noise sensitivity function to quantify this effect. Noise sensitivity function describes how noise affects speech under different SNR conditions. For a given type of additive noise and clean speech, the intensity mixture ratio of speech to noise is changed based on the given SNR. In cepstral domain, the cepstrum is also the mixture of speech and noise. The noise sensitivity function is defined as the ratio of cepstral coefficient increment to SNR increment. According to noise sensitivity curves, it can be expected to estimate the clean speech cepstrum from any noisy speech cepstrum by integrating the noise sensitivity curves from the current SNR to the highest SNR, at which the speech can be approximately considered as a clean speech.
In the Second way, we try to catch the periodic component in speech,

3. The progress of this year

a. Find the robust parameter to describe noise sensitivity

Noise sensitivity function describes how noise affects speech under different SNR conditions. For a given type of additive noise and clean speech, the intensity mixture ratio of speech to noise is changed based on the given SNR. In cepstral domain, the cepstrum is also the mixture of speech and noise. The noise sensitivity function is defined as the ratio of cepstral coefficient increment to SNR increment.

b. Finding the analytic formula of NSF

We found the relationship between spectrum and noise sensitivity function. Our formula show that the noise sensitivity function is mainly decided by sub-band signal-noise ratio. It is also decided by correlation coefficients.

c. Using the general noise sensitivity function for noise compensation.

According to noise sensitivity curves, it can be expected to estimate the clean speech cepstrum from any noisy speech cepstrum by integrating the noise sensitivity curves from the current SNR to the highest SNR, at which the

speech can be approximately considered as a clean speech. Our compensation method is compared with the conventional spectral subtraction algorithm. Our method is tested by TIMIT database. The experiments showed that the proposed method works better than a conventional method of spectrum subtraction.

4. Future direction

The mentioned method only considered the average information of NSF. But each frame has its own noise sensitivity function. If its own NSF is substituted by the average NSF, some errors will be brought to compensation result. So, it is important to estimate each frame's NSF based on the observation value. So, we use the gain factor to adjust the average NSF for each frame. Each frame's gain factor can be estimated by EM algorithm.

5. Publication

D.W. Ying, X.G. Lu, J. Dang, A Measurement of Noise Sensitivity and Its Application in Robust Speech Recognition, 2005 Spring Meeting of ASJ (Acoustic Society of Japan).