Study on change of urgency perception with manipulating temporal amplitude envelope of speech

2310140 Saki Fusano

Cochlear implant is one of the artificial organs used by people with hearing impairments. It converts input sound into electrical stimulation and stimulates the auditory nerve directly through electrodes. Insertion into the cochlea can significantly restore hearing. Treatment with cochlear implants is particularly effective for children with hearing loss. If patients restore hearing by early fitting of cochlear implant in infancy, they can communicate like that of normal hearing person. The number of cochlear implant users is expected to increase further, as the criteria for cochlear implantation in children are being revised to allow for earlier cochlear implantation. On the other hand, there are several issues in vocal communication using cochlear implant. Especially, Cochlear implant users find it difficult to get non-linguistic information. To realize rich communication for cochlear implant users, the transmission of non-linguistic information must be improved using speech synthesis processing for cochlear implants.

Noise-vocoded speech (NVS) is the simulator of cochlear implant speech that uses band-limited random noise (temporal fine structure) modulated with temporal amplitude envelope (TAE) information. Therefore, it can synthesize speech in which only the temporal information cues (information of TAE) are preserved, without any spectral information cues. If NVS can be used to reveal the acoustic features involved in the transmission of non-linguistic information, it can contribute to the improvement of speech synthesis for cochlear implants. In this study, the acoustic features involved in the urgency perception, one of the non-linguistic information are clarified.

It has been suggested by previous studies using NVS that TAE of speech includes not only linguistic information (intelligibility and comprehensibility), but also important cues related to non-linguistic information (emotion and individuality). Furthermore, it has revealed that the TAE of speech includes cues that are important for urgency perception. Furthermore, it has found that the temporal modulation frequencies for urgency perception important cues related to urgency perception were the temporal modulation frequency range from 4 Hz to 16 Hz. However, it has found that there was no difference in the urgency perception (i.e., it was difficult to distinguish the different of urgency perception)

in the NVS with the time-reversed TAE included the modulation frequency band important for the urgency perception. This result has suggested that not only the long-time averaged modulation frequency components in the TAE of but also the temporal features of modulation frequency components are important cues for urgency perception. Therefore, it is necessary to examine the relationship between the instantaneous modulation components (IMCs) of TAE and the urgency perception to clarify the cause of the change in the degree of urgency of NVS to the time reversal processing.

On the other hand, it has investigated whether the IMCs of TAE contribute to emotion perception. It suggested that the cause of the reduction in the emotion recognition rate in NVS with time-reversed TAE. The reduction was reproduced by manipulating the IMCs using temporal stretching and compression on the TAE. If this result is not limited to emotion perception, but is also related to the urgency perception, it may be possible to clarify the cause of the reduction in urgency perception of the NVS with time-reversed TAE by manipulating IMCs of TAE.

This study aims to investigate the change in the urgency perception of the NVS by manipulating IMCs in TAE. The following four point of issues were investigated for this purpose. The first point is whether TAE contains cues for urgency perception. The second point is the modulation frequency components that are important for the urgency perception by using Modulation Filterbank (MFB) to manipulate the modulation frequency band of TAE. Before using MFB, it was investigated that whether MFB processing of TAE affects the urgency perception of NVS. Then, the modulation frequency band that is important for the urgency perception was investigated. The third point is the change urgency perception of the NVS through time-reversal processing of TAE. Finally, the fourth point is the change urgency perception of the NVS through manipulating of IMCs using temporal stretching and compression on TAE.

The above points were investigated l by conducting listening experiments with NVS. The modulation frequency domain and time domain of TAE were processed according to the conditions, and changes in the urgency perception of NVS with that TAE were investigated. The results of the experiment revealed the following:

1. TAE contain cues to the urgency perception in speech.
2. The modulation frequency components that are important cues for the urgency perception are from 4 Hz to 16 Hz.
3. Time-reversal processing of TAE reduce the urgency perception of NVS.
4. Manipulating IMCs using temporal stretching and compression on TAE

can reproduce the change urgency perception of NVS caused by time-reversal processing on TAE, except for a few results.

These finding suggest that IMCs of TAE plays an important role in urgency perception in speech.