TR－A－0162

TR－H－010

Dynamic Cepstrum Parameter Incorporating Time-Frequency Masking and Its Application Speech Recognition

Kiyoaki Aikawa, Hideki Kawahara, Yoh'ichi Tohkura

1993. 3. 4

ATR視聴覚機構研究所

〒619-02 京都府相楽郡精華町光台2-2 ☎07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1411
Facsimile: +81-7749-5-1408

Dynamic Cepstrum Parameter Incorporating Time-Frequency Masking and Its Application to Speech Recognition

Kiyoaki Aikawa†, Hideki Kawahara‡ and Yoh'ichi Tohkura‡
† ATR Auditory and Visual Perception Research Labs.
‡ ATR Human Information Processing Research Labs.
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

ABSTRACT

A Dynamic Cepstrum parameter is proposed that incorporates the time-frequency characteristics of forward masking. Psychological research reports that the forward masking pattern becomes more wide-spread over the frequency axis as the masker-signal time-interval increases. To simulate the masking characteristics, a novel lifter array operation is derived. The Dynamic Cepstrum can represent both the instantaneous and transitional aspects of speech spectra. The proposed parameter is superior to the conventional delta cepstrum in extracting high temporal resolution spectral dynamics, and outperforms the conventional cepstrum in phoneme recognition.

This technical report is the detailed version of the paper presented at the 124th Meeting of the Acoustical Society of America (J. Acoust. Soc. Am., Vol. 92, No.4, Pt. 2, pp.2476, 5pSP5, Oct. 1992).

[redacted]