

1, 2, 3

1 2 3

{reraj¹, msh², khcho³}@etri.re.kr

Vocal-cord Signal Study based on Phonological Feature for Vocal-cord Signal Isolated-Word recognizer

Young Giu Jung¹, Mun Sung Han², Kwan Hyun Cho³
Smart Interface Research Team^{1 2 3}, ETRI

가 .
 가 .
 ,
 ,
 FFT , MFCC
 / 가
 ZCPA MFCC 16%
 ZCPA
 Keyword : , , ZCPA,

1. [3], [4],

Ubiquitous compute,
pervasive compute, wearable compute

가 가
 가 가 , ,
 가 가 가 가 ,
 ,
 가

[1], soft-
decision [2],

MFCC(Mel-Frequency Cepstral Coefficient)

가 + , + , 가

[5]

가 / ZCPA (voiced), (vocalic)/ (consonantal), (syllabic), (sonorant)/ (obstruent) [6].

2.

- : 가
- / : 가
- : 가
- / : 가 가

2-1

21 “ㅅ” (ㄱ, ㄷ, ㅂ, ㅈ, ㅎ)

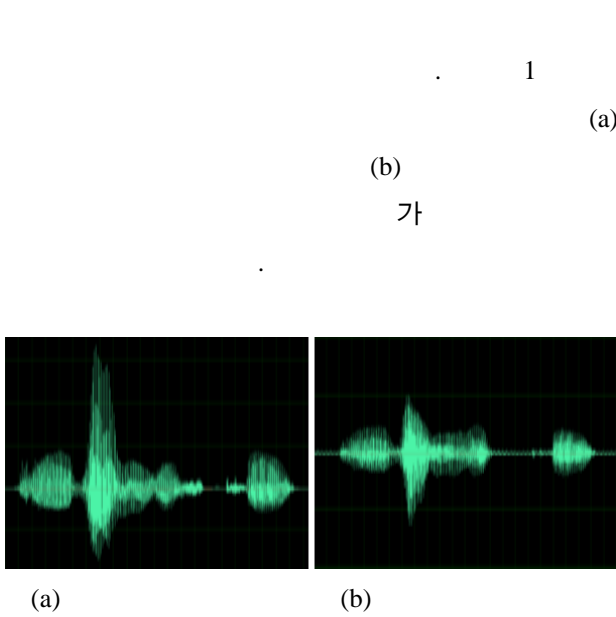
19

1 ()

1.

	ㅂ	ㄷ	ㅈ	ㅅ	ㄱ	
	ㅃ	ㄸ	ㅉ	ㅆ	ㄲ	
	ㅍ	ㅌ	ㅊ		ㅋ	ㅎ
	ㅁ	ㄴ			ㅇ	
		ㄹ				

2-2.



1.

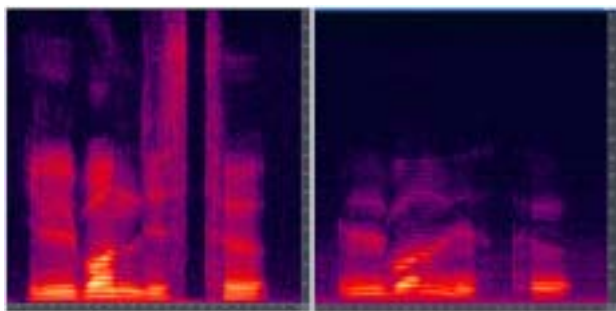
100 가 MFCC

30% 가
 가
 critical band pass filter
 MFCC 가 가,
 가

1

2

1

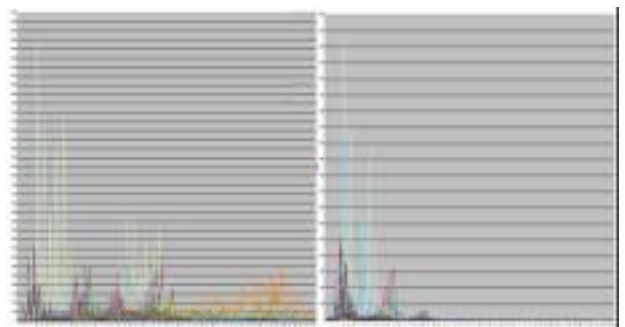


(a)
2.

(b)

가
 1
 (a) 가
 (b) 가
 2.b
 가
 4Khz

가
 MFCC(Mel-Frequency Cepstral Coefficient)
 FFT(Fast Fourier Transform)
 3 16K, 16bit wave
 Pre-emphasis, Hamming-Windowing
 FFT



(a)
3.

(b)

3 가 256

2Khz 가
 2Khz 4Khz

가
4Khz

[8].

가

MFCC

[9]

critical band filter

가 MFCC /

EIH

가 ZCPA(Zero-Crossings with Peak

가 Amplitudes)[10]

2-3.

가 /

3.

가 2

가 2

가 MFCC /

가 ZCPA

FFT

가 MFCC

가 Mel-

cepstrum critical band filters

가

Ghitza Ensemble Interval

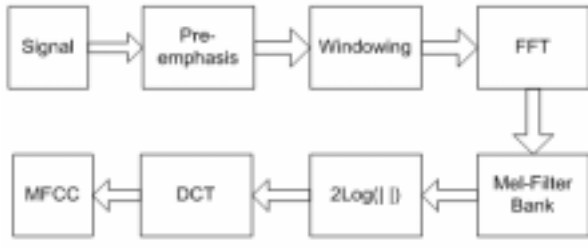
Histogram(EIH)[7] band pass

cochlear bank 1KHz filter

log

가 / 4 Mel-cepsturm MFCC

가 /



4. MFCC

Nonlinear stages band pass cochlear bank
 ZCPA ZERO-CROSSING
 PEAK amplitude

EIH

ZCPA zero-crossing
 , peak

intensity 1 t
 ZCPA

$$y(t:i) = \sum_{channel} \sum_{k=1} \delta_{ij_k} f(A_k), 1 \leq i \leq N \quad (1)$$

K channel upward zero-crossing
 , N frequency bin j_k k
 (k+1) zero crossing
 frequency bin A_k peak amplitude

δ_{ij} Kronecker delta

4. 가

100 50
 25 50

TDNN

2.

	MFCC	ZCPA
(%)	97.14	98.86

critical band filter
 MFCC 가 2

5Khz
 Mel-filer 5K
 24 Mel-filer bank filter
 Rectangle
 1KHz 1KHz
 shift 3

3. MFCC

	(%)
MFCC	67.8
MFCC+CMS	74.95
5KHz 24 filter(MFCC+CMS)	68.49
Rectangle Filter(MFCC+CMS)	71.547
Filter (MFCC+CMS)	70~73

5KHz 24
 Mel-filer MFCC CMS

가 5KHz

가

Rectangle filter
 triangle filter 2~3%
 [11]

1KHz
 1KHz shift

가
 band pass filter

2

Peak zero-crossing MFCC ZCPA

4. MFCC ZCPA

	(%)
ZCPA	83.63
MFCC	67.8

4 ZCPA
가 MFCC 16%
2

5.

Ubiquitous computing, wearable computing

가

가

가 /

가

band

peak zero-crossing ZCPA

MFCC CMS ZCPA

가 16%

83.63%

[1] S.F. Boll "Suppression of acoustic noise speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-27, 113-120, Apr., 1979;

[2] R.J. McAulay and M. L. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Trans. Acoust., Speech, Signal Processing*, 28, 137-145, Apr.1980.

[3] Y.Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Trans. Acoust.,Speech, Signal Processing*, 33, 443-445, Apr. 1985.

[4] Y. Ephraim, "Statistical-modal-based speech enhancement system," *IEEE Proceeding*, 80, 1526-1555, Oct. 1992.

[5] , , , 1999.

[6] , , , 2000.

[7] O. Ghitza, "Auditory models and human performances in tasks related to speech coding and speech recognition," *IEEE Trans. Speech and Audio Processing*, vol.2, no. 1, part II, pp. 115-132, 1994.

[8] C.K. Un and S. C. Yang, "A Pitch extraction algorithm based on LPC inverse filtering and AMDF," *IEEE Trans. Acoust., Speech Signal Processing*, ASSP-25, 565-572, Dec. 1977.

[9] R. P. Cohn, "Robust Voiced/Unvoiced Speech Classification using a Neural Net," *ICASSP-91*, 437-440, 1991.

[10] Doh-Suk Kim, Soo-Young Lee, Rhee M. Kil "Auditory Processing of Speech Signals for Robust Speech Recognition in Real-Word Noisy Environments," *IEEE Tran. Speech and Audio Processing*, vol., 7 No. 1, Jan., 1999.

[11] Donghoon Hyun, Chulhee Lee, "Optimization of mel-cepstrum for speech recognition," *IEEE SMC '99 Conference Proceedings Volume 1*, pp. 500-503, Oct. 1999.