

청지각적 음성장애평가에서의 객관적인 파라미터 추출

Objective parameter extraction in perceptual dysphonia assessment

장승진¹, 최예린², 김은연³, 김원식¹¹한국표준과학연구원 뇌인지융합기술연구단²명지대학교 언어치료학과³삼성서울병원 언어치료실

ABSTRACT

GRBAS(G : grade, R : rough, B : breathy, S : strained, A : asthenic) 음성장애평가는 성대의 이상 또는 말마비장애 등의 환자들을 평가하는 척도로 널리 사용된다. 하지만 사람에 의해 주관적인 평가로 이루어지는 방식의 문제점이 많이 제기되어, 자동화 알고리즘에 의한 객관적인 청지각적 음성장애 평가도구를 개발하려는 시도가 많이 연구되어왔다. 이러한 개발에 있어 보편적으로 선행되어야 하는 음소 분류 및 일치성 판단을 위한 객관적인 파라미터를 구하고자 함이 본 연구의 목적이다.

Keyword: GRBAS, dysphonia assessment, objective parameters

1. 서론

음성신호에서 특징을 추출하는 것은 기존의 특징추출방법 중 하나이다. 개발된 기법 중 가장 많이 쓰이는 것은 멜-주파수 cepstrum 계수(MFCC: Mel-Frequency Cepstral Coefficient) [1]이다. MFCC 는 사람의 청각모델을 바탕으로 하고 있으며 잡음이 없는 경우 및 있는 경우 모두 우수한 인식 성능을 내는 것으로 알려져 있다.

GRBAS 음성장애평가[2]는 특정한 모음 자음 및 문장에 대한 피검자의 발성 정확성 및 음성의 청지각적 등급평가에 의하여 평가를 하게 된다.

하지만 선천성 조음장애인의 경우 의식적/무의식적 상관없이 동일한 음성 및 음소에 대한 일치성이 떨어지고 본인의 청지각 능력 또한 정상인에 비해 떨어지게 된다.

본 연구에서는 정확한 음소 발성 및 일치성 평가에 필수적인 객관적인 파라미터를 개발하여 청지각적인 음성분석보다 객관적인 성능을 보이려한다.

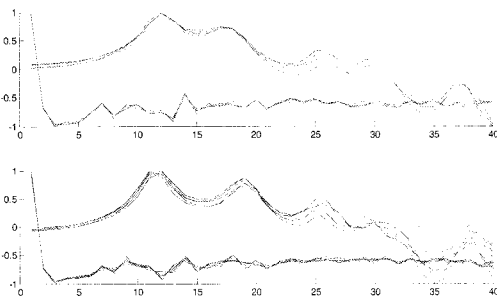
2. 실험 방법

입력된 신호를 한번에 15ms 씩 움직이며 30ms 크기의 프레임으로 분할하고 각

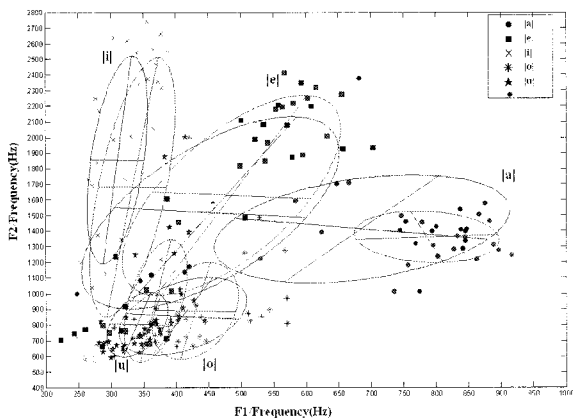
프레임별로 푸리에 변환, 필터뱅크 에너지 계산, 로그 변환, 이산 코사인 변환을 거쳐 40 차 MFCC 와 정규화된 프레임 에너지, 그리고 그것들의 시간 미분인 동적 특징(dynamic feature)을 청각특징으로 사용하였다.

멜주파수 캡스트럼계수(MFCC)로 분류되는 각 음소(모음, 자음: 파열음, 마찰음 등)들의 feature 를 추출한 후 PCA(Principle Component Analysis)로 supervised/unsupervised 형태로 clustering 하여 음소들을 생성한다. 또한, 환경음(공항, 시내, 자동차 등)에 대한 각 음소의 청각인지를 평가하기 위하여 각 음소들에 대하여 환경음의 주파수 및 강도를 변화하여 주파수대역/SNR(Signal-Noise-Ratio) 별 환경음을 제작하여 청각인지 데이터로 사용한다.

3. 실험 결과



<그림 1. 40 차 Mahalanobis distance in MFCC&LPA>



<그림 2. Loci of the mean and S.D. of 1 of F1 and F2 formant frequencies >

MFCC 와 LPA 를 병행하여 분석한 Mahalanobis distance 파라미터는 그림 1 에서 보는 것과 같이 일치성 평가에서 보다 더 정확하게 formant frequency 뿐만 아니라 각각의 주파수 밴드에서 생성되는 대표값으로 조금 더 정밀하게 음색 등을 감지할 수 있다. 이를 바탕으로 /a/, /e/, /i/, /o/, /u/ 모음에 대하여 분석한 결과 파란색과 같이 음성의 불일치성이 넓게 퍼져 있는 것들이 빨간색과 같이 조밀하게 모일수 있게 됨을 알 수 있다.

4. 논의

정확한 조음 및 음성, 음색과 같이 주관적인 청지각적 판단으로 음소를 평가해야 하는 경우 기존의 단일 파라미터 분석 방법보다는 주파수 도메인에서의 다차원 벡터를 기반으로 하여 차원을 늘려가면서 분석하는 것이 보다 정밀하게 음소를 판단할 수 있는 객관적인 파라미터로 작용할 수 있을것이다.

참고문헌

- [1] Zheng, Guoliang Zhang and Zhanjiang Song (2001), "Comparison of Different Implementations of MFCC," J. Computer Science & Technology, 16(6): 582-589.
- [2] Fecha de aceptación: PERCEPTUAL EVALUATION OF DYSPHONIA: CORRELATION WITH ACOUSTIC PARAMETERS AND RELIABILITY, Acta Otorrinolaringol Esp 2004; 55: 282-287