

음원분석을 통한 장애음성의 음향적 특성분석에 관한 연구

창원대학교 제어계측공학과

조철우, 김대현

Analysis of Acoustical Characteristics of Pathological Voice Using Source Analysis

Dept. of Control & Instrumentation Eng., Changwon National University

Cheol-Woo Jo, Dae-Hyun Kim

cwjo@sarim.changwon.ac.kr

요 약

본 논문에서는 장애음성들의 분석을 위하여 기존의 파라미터들인 jitter, shimmer 및 NHR과 함께 음원의 추정에 의한 파라미터를 이용하여 장애음성의 음향적 특성분석을 위한 실험을 행하고 정상음성과 장애음성을 이들 파라미터에 의해 식별하고자 하였다.

1. 서 론

최근들어 인간의 건강에 대한 관심이 점점 증가하고 있다. 이비인후과 영역에서는 성대의 질환을 음성의 음향적 특성을 분석하여 질병의 징후를 발견하려는 시도가 여러 곳에서 이루어 지고 있다. 이러한 시도는 질병의 정확한 진단이 목적이라기 보다는 질병의 가능성을 사전에 발견하여 질병의 상태가 깊어 지지 않도록 예방한다는 차원에서 필요하다. 일반적으로 성대의 질환을 진단하기 위해서는 성대를 직접 들여다 볼 수 있는 내시경과 같은 기구를 이용하여 직접 보는 것이 가장 효과적이라고 한다. 실제로 의사들도 이러한 방법을 많이 사용하고 있다. 그러나 이 방법은 정확한 진단이 가능한 반면 숙달된 전문의사와 전문적인 기구가 필요하여 일반의사나 환자 자신이 질병의 유무를 판단하기에는 적합한 방법이 아니다. 이러한 목적으로는 환자의 음성만에 의하여 질병의 유무, 종류를 진단하려는 시도가 곳곳에서 이루어 지고 있다. 성대의 질환은 많은 경우 성질(Voice Quality)의 변화를 수반하지만 그렇지 않은

경우도 있다. 목소리만으로 질병을 진단할 수 있는 경우는 질병으로 인해 음성의 음향적 특징이 변화된 경우에 한한다.

본 논문에서는 이와 같은 음향적 분석법에 의한 성대질환의 진단을 목표 하는 파라미터를 개선하기 위하여 음원분석에 의한 장애음성식별을 시도하였다.

2. 장애음성의 음향적 특성

성대장애환자의 음성은 성대의 장애로 인하여 발성기관의 음원부분이 변형되면서 음의 특성이 변화하게 된다. 특성의 변화는 성대의 장애부위에 따라서 나타나는데 대개의 변화는 성대부위의 종양, 경직화등이 원인이 되어 불규칙적인 변화를 가져온다.

성대장애 환자의 음성의 음향적 특성은 정상인에 비하여 피치의 불규칙성, 잡음성분의 증가, 고주파성분의 증가, 등으로 나타난다. 이러한 변화를 측정하기 위하여 기존의 파라미터로 제안된 것들은 jitter, shimmer, NHR 등이 있다. 그러나 이러한 파라미터들중 어느 것도 장애음성의 특성을 명확히 구분해 주지는 못하고 있기 때문에 여러 가지 파라미터를 복합적으로 사용하여 진단에 사용하고 있다.

특히 기존 파라미터들의 경우 피치주기를 먼저 측정하여 주기의 불규칙성등을 계산하고 있기 때문에 피치의 측정이 잘못된 경우 파라미터의 값은 신뢰성이 떨어지게 된다. 장애음성의 경우는 피치가 두 배가 되든지 반이 되는 경우가 많이 일어나며 육안으로도

구분이 되지 않을 정도로 불규칙적인 경우가 많다. 장애음성을 진단하는 방법중의 하나로 음질을 이용하여 평가하는 방법이 있는데 이는 전문의사가 환자의 음성을 직접 듣고 R(rough), B(breathy), A(asthenic), S(strained)의 등급으로 구분하는 방법이다. 이 방법은 실제로 일본에서 많이 진단에 이용되고 있는 방법이다.

본 연구에서는 이와 같은 진단법을 자동으로 수행할 수 있게 하기 위한 연구의 일환으로 음원분석과 캡스트럼에 의한 장애음성신호의 분석을 시도하고 기존 파라미터들과의 관련성을 비교 검토하였다.

3. 음원분석법

음원의 분석을 위하여는 선형예측분석으로부터 구해진 예측오차신호로부터 Cepstrum을 구한 뒤 성도필터 성분과 주기성분, 잡음성분을 분리해 낸다. 이 결과로부터 Harmonic to Noise Ratio를 구한다. 이렇게 하여 구한 파라미터를 HNRR(Harmonic-to-Noise Ratio: Residual)이라고 한다.

또한 이와는 별개로 원래 음성에서 선형예측

분석을 행하지 않은 원음성신호를 캡스트럼 분석하여 하모닉성분과 잡음성분을 분리하여 비율 구한 것을 HNRS(Harmonic-to-Noise Ratio: Speech)라고 한다.

이들 파라미터들로부터 정확하지는 않지만 대략 다음과 같은 성질을 유추할 수 있다. HNRR은 음성의 음원에서의 하모닉성분과 잡음성분의 비율 추정된 것으로 볼 수 있고 HNRS는 원 신호에서의 하모닉성분과 잡음성분의 비율 추정된 것으로 볼 수 있다. 이러한 점에서 이들 파라미터는 음원의 변화를 대략적으로 나타내주는 파라미터로 볼 수 있다. 기존에 제안된 유사한 파라미터로는 NHR이 있는데 이는 원음성에서 70-4500Hz 사이의 하모닉성분과 1500-4500Hz사이의 잡음성분의 비율 FFT결과로부터 구한 것으로 성도의 성분을 제거하지 않고 구하였기 때문에 여기서 제안한 파라미터들과는 약간 차이가 있다.

제안한 파라미터들의 특성은 다음과 같다. 정상음성의 경우는 선형예측오차신호에서 주기성분이 더욱 강조되는 것이 일반적이기 때문에 HNRS에 비해서 HNRR이 더 커질 것이 예상된다. 그러나 환자음성의 경우는 불규칙성이나 잡음성분이 부가되므로 선형예측

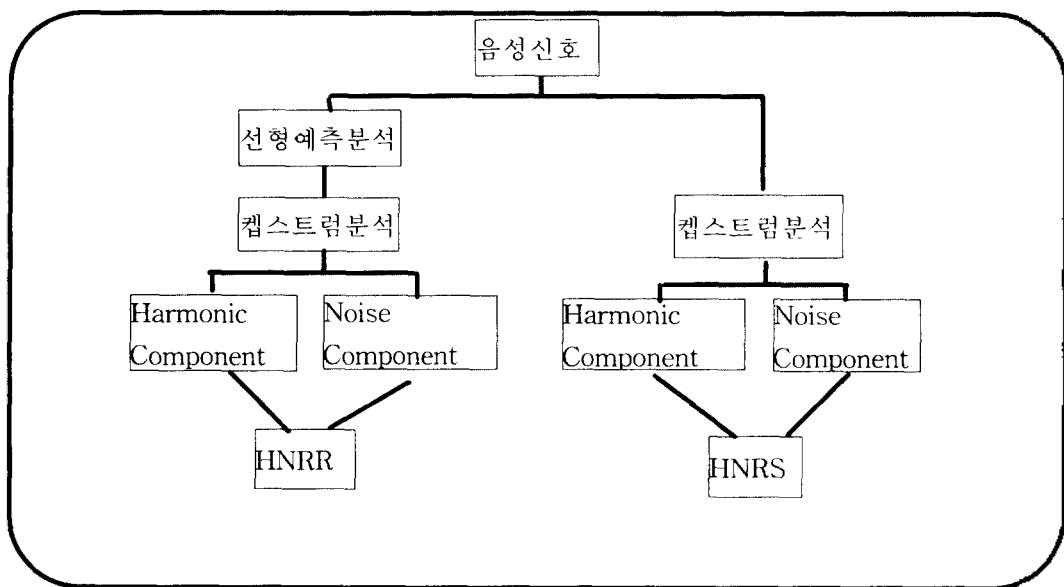


그림1. 신호처리방법

오차신호에서 하모닉성분이 감소할 것이 예상되어 HNRR의 경우 HNRS보다 작아질 것이 예상된다.

4. 실험 및 고찰

본 실험에서는 앞서 제안한 NHRR과 HNRS를 Kay의 Disordered Speech Database를 대상으로 실험하였다. 이 데이터베이스에는 700여명의 정상인 및 각종 성대장애환자의 음성데이터가 수록되어 있으며 jitter, shimmer등 MDVP로 분석가능한 33개 파라미터가 구해져 같이 수록되어 있다. 이들중 vocal fold edema(성대부종)란 병명의 44개 데이터와 정상인의 음성 53개 데이터에 대하여 실험하였다.

우선 그림1의 신호처리방법에 의해 HNRR과 HNRS를 모든 음성에 대하여 구한 뒤 기존에 구해져 있는 jitter, shimmer, NHR과 비교하였다. 이 세 개의 파라미터와 비교한 이유는 데이터베이스의 모든 장애음성과 정상음성을 33개 파라미터에 대하여 통계적으로 분석한 결과 이 3개의 파라미터가 가장 변별력이 있는 것으로 측정되었기 때문이다.

그림1은 jitter, shimmer, NHR과 HNRS, HNRR의 상관도를 각각 비교한 그림이다.

그림에서 보면 jitter, shimmer, NHR과 HNRS, HNRR은 각각 반비례하는 형태를 보이고 있다. 각각을 보면 jitter나 NHR에 비해서는 분포가 조밀한 형태를 보이고 있으며 shimmer에 대하여는 흩어진 분포를 보이고 있다.

HNRS와 HNRR간의 차이는 별로 나타나지 않았는데 이는 분석자료로 삼은 데이터가 불규칙성을 많이 갖지 않고 있는 때문으로 볼 수 있다.

그러나 이 그래프들로부터 제안한 파라미터가 기존의 파라미터들에 비하여 식별파라미터로 유용함을 확인할 수 있다.

5. 결 론

본 논문에서는 선형예측오차신호와 캡스트럼 신호를 이용한 장애음성식별파라미터를 도출하고 실제 음성데이터에 적용한 뒤 기존의 파라미터와 비교하여 그 유효성을 검증하였다. 실험결과 제안한 파라미터는 제시된 범위의 질병에 대하여 정상음성과의 구분을 할 수 있는 파라미터로 유효함이 입증되었다. 차후의 연구에서는 보다 많은 종류의 질병에 대하여 정상음성과의 식별 및 각 질병간의 구분에 대한 실험이 진행될 예정이다.

감사의 글

본 연구는 한국학술진흥재단의 1997 학제간 연구비 지원에 의한 연구결과의 일부입니다. 지원에 감사드립니다.

참고문헌

1. Operations Manual, 'Multi-dimensional Voice Program(MDVP)' ,Model4305, Kay Elemetrics Corp, 1993
2. Operations Manual, 'Disordered Voice Database', Model 4337, version 1.03, Kay Elemetrics Corp, 1994
- 3.F.Plante, H.Kessler ,B.Cheetham ,J.Earis , "Speech Monitoring of Infective Laryngitis", Proceedings of ICSLP'96, pp.749-752 ,Philadelphia, 1996
- 4.B.Yegnanarayana, C.d'Alessandro, V.Darsinos,"An Iterative Algorithm for Decomposition of Speech Signals into Periodic and Aperiodic Components",IEEE trans. on Speech and Audio Processing, Vol.6, No.1, Jan. 1998

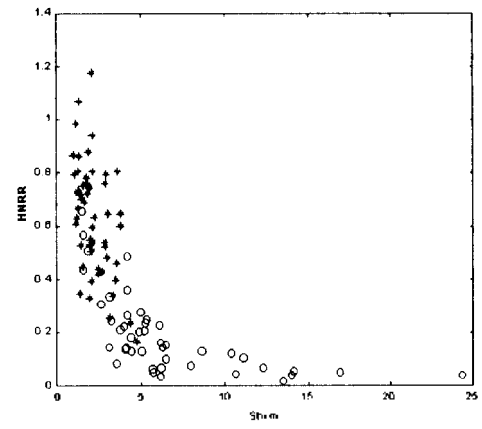
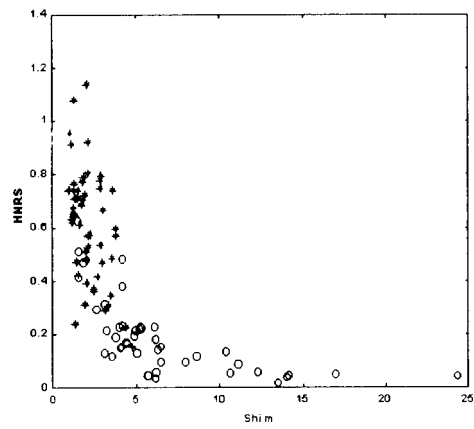
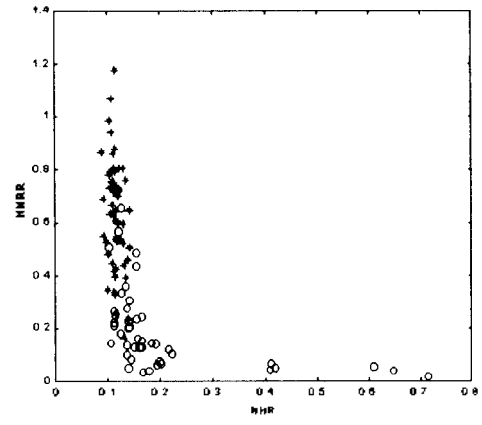
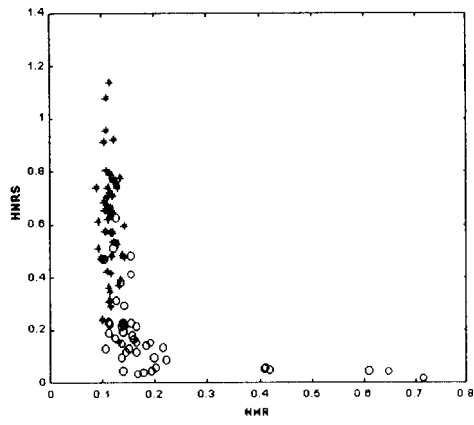
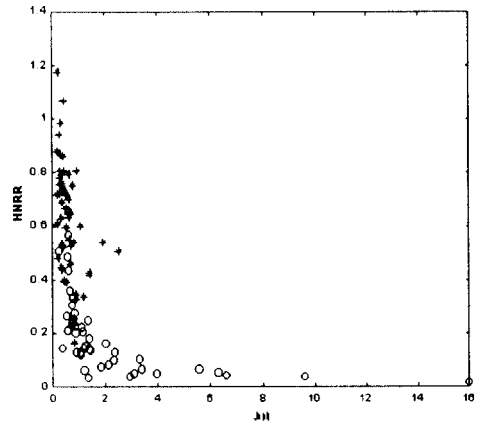
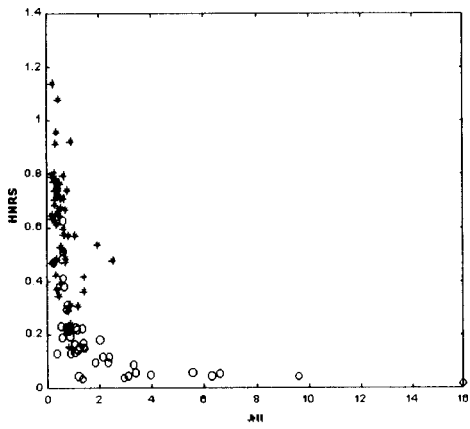


그림2. HNRS, HNRR과 jitter, shimmer, NIIR과의 상관도