

연속모음에서의 Electroglottograph 신호해석에 의한 후두기능 평가

Laryngeal Function Assessment by Electroglottographic Signal Analysis during Sustained Vowel Phonation

· 송 철 규, 이 명 호
연세대학교 전기공학과

Song Chul Gyu, Lee Myoung Ho
Yonsei University Dept. of Electrical Engineering

-Abstract-

Perturbation in the fundamental frequency and in the peak amplitude of the EGG signal derived with a four-electrode EGG system were investigated for the purpose of developing useful measures for the detection of laryngeal pathology.

The data were compared to the degree of amplitude perturbation and frequency perturbation. There was a close relation between amplitude perturbation and frequency perturbation analysis of EGG signal and degree of laryngeal pathology.

1. 서론

후두(larynx)는 인두(pharynx)와 기관(organ)을 연결하는 인체 부위로서 소리를 내는 작용과 음식물이 호흡기도로 들어가지 않게 막는 반막과 같은 역할의 중요한 기능을 한다. 특히 후두는 이비인후과 영역 중에서 가장 진찰하기가 어려운 부위의 하나인데, 후두암이나 각종 질환의 의심이 있는 환자는 종래의 간접후두경 검사는 관찰이 효과적이기는 하지만 gag reflex가 심한 환자나 후두경의 모양과 위치로 인해 후두를 관찰하기가 어려운 경우가 매우 많다. 또한 후두질환의 조기발견을 위한 객관적인 변수를 얻기 위해 환자의 음성신호에서 주파수 변동과 진폭변화를 해석을 해 왔으나, 좀 더 정확한 검진을 위해서는 EGG를 이용하여 성도(vocal tract)의 음향적인 공진현상[1]에 영향을 받지 않는 성분파를 이용할 수가 있다. EGG신호에서 성문막(vocal folds) 진동의 불규칙성에 의해 서로 다른 주파수 및 진폭 변동은 서로 다른 형태를 나타내게 된다.

본 연구에서는 이러한 주파수와 진폭 변동의 임상적인 중요성을 평가하기 위하여 정상인 남녀와 다양한 후두질환을 갖는 환자의 데이터를 비교 및 평가함으로써 정상적인 후두기능의 상태를 분류하기 위한 새로운 판별기준을 제시하였다.

2. EGG 신호의 분석

EGG 신호의 분석을 위해서 연속으로 발음된 모음과 동기된 EGG 신호 100 주기씩을 피검자당 한 샘플로 하였다. 분석을 위한 각각의 샘플은 데이터 시작점으로 부터 300 ms 전에서 부터 100 주기가 분석을 하였다.

EGG 신호로 부터 후두의 각 질환들을 분류하기 위해서 주파수 변동(frequency perturbation)과 진폭변동(amplitude perturbation)의 해석에 적용한 수식은 두 변수의 비례적인 비교를 위해 보상값을 고려한 것으로 다음과 같다.

주파수 변동은 식 (1)에 의해 계산이 되었고, 이 값은 연속된 각 사이클들 간의 log값을 갖는 주파수의 평균차이이다. 각 주기의 결정은 기저선에 대한 zero crossing 횟수를 찾아 결정할 수가 있다.

$$\Delta F = (\sum_{i=1}^N |F_i - F_{i+1}|) / (N-1) \quad (1)$$

여기서, $F_i = \log(f_i/C_0)$ 이고 $f_i = 1/P_i$ 이 되며,

$F_i = i$ 번째 주기의 주파수(log scale),

$P_i = i$ 번째 주기의 시간,

$f_i = i$ 번째 주기의 주파수(Hz),

N 은 전체 주기의 수, $C_0 = 50$ 이다.

진폭변동은 식 (2)와 같이 표현되고, 연속된 주기간의 dB값을 갖는 진폭의 평균차이이다. 진폭의 크기는 각 주기당 최대값 피크에서 최소값 피크를 더한 값이다.

$$\Delta A (dB) = [\sum_{i=1}^N |20 \log(A_i + 1/A_i)|] / (N-1) \quad (2)$$

$A_i =$ 각 주기에서의 진폭의 크기

N 은 전체 주기의 수

그림 1은 단모음 '애'를 발음했을 때 음성신호와 이에 동기된 EGG 신호이다. 가로 축은 시간의 경과에 따른 샘플의 수이고, 세로 축은 신호의 크기이다. 그림에서 알 수 있듯이 음성신호에 EGG신호가 정확하게 동기화 됨을 볼 수 있다.

또한 무성음, 목음일 경우에는 EGG의 진폭은 거의 0에 가깝게 되므로 음성 분석시에 음성만으로 하는 것 보다 EGG신호를 기준으로 하면 좀 더 정확한 분석을 할 수가 있는 것이다[2]. 그림 2는 전형적인 EGG 신호로서 정상인 단모음 '아'를 발음할 때의 신호이다. 본 연구에서 EGG신호를 분석하기 위해 주파수 변동과 진폭의 변동의 변수를 설정하였는데, P_i 는 주파수 변동을 계산하기 위해 기준이 되는 현재 주기에 대한 피치의 시간간격이고, A_i 는 진폭변동을 계산하기 위해 기준이 되는 현재 주기의 진폭

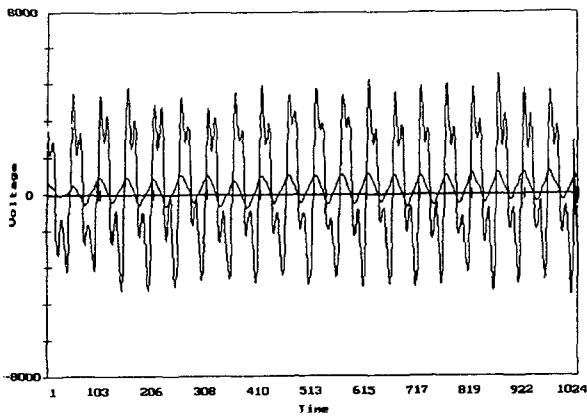


그림 1 음성신호와 동기된 EGG신호
Fig. 1 Speech signal and synchronized EGG signal

크기이다. 그림 2에서와 같이 P_i 는 해당 주기에서의 최대피크에서 다음 주기의 최대 피크까지의 거리를 나타낸다.

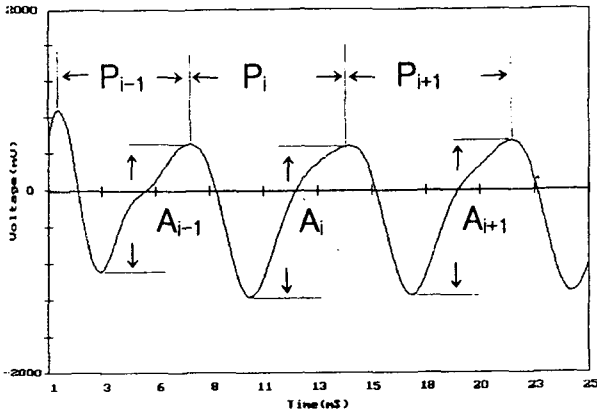


그림 2 단모음 '아'의 전형적 EGG 신호
Fig. 2 Typical EGG signal of single vowel 'a'

3. 데이터 수집

임상실험과 데이터 수집은 연세대학교 이비인후과에서 Laryngitis, vocal polyp, husky voice, laryngeal cancer 환자를 대상으로 하였으며, 각 질환에 따른 평가를 위해 어음재료로서 단모음 'ㄷ', 'ㄱ', 'ㄴ', 'ㅇ', 'ㅈ', 'ㅊ', 'ㅌ', 'ㄹ'를 일정한 톤(tone)으로 3 초 이상 발음하도록 하였고, '환경을 보호하자'라는 문장을 사용하였다.

측정한 임상 데이터들은 FOSTEX사의 4채널 멀티트랙커(X-28H)를 이용하여 바이어스(bias)가 동일한 FM tape에 기록하였다.

4. 결과

그림 3은 성대 폴립(vocal polyp) 환자의 '애' 발음시 EGG 신호에서, 각 사이클의 주기 P_i 와 진폭 A_i 를 표시하였다. 여기서 주기 P_i 는 첫번째 영교차(zero-crossing)후의 최대 피크값에서 두번째 영교차 후의 최소 피크값까지의 거리를 계산한 값이고 A_i 는 그때의 최대 피크값에 최소 피크치를 더한 값으로 계산된다. 성대 폴립환자의 경우 정상인과는 다르게 주기와 진폭의 변동이 크게 변화되고 있음을 알 수가 있다. 이 때 각 사이클당 피치 주기는 각각 15ms, 10ms, 14 ms로 큰 차이가 있고, 진폭 또한 A_{i-1} 은 A_{i+1} 보다 2배 이상 클 수 있다.

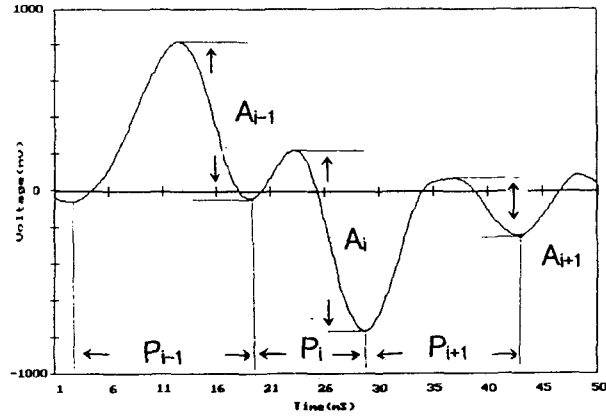


그림 3 성대 폴립(vocal polyp)의 EGG 신호
Fig. 3 EGG signal of vocal polyp

또한 성대가 열리고 닫히는 반복적인 운동에 의해서 유성음이 형성되는데, 이러한 음원발생의 함수를 EGG시스템을 이용하면 성문(vocal folds)에서의 시변 음향 임피던스로 구할 수가 있다 [3]. 그림 5는 정상인이 모음 '오'를 발성할 때 얻어진 EGG신호로부터 성대의 열림과 닫히는 구간을 표시하였다. 그림에서 열림과 닫힘 구간은 각각 4.4 msec와 2.8 msec로 계산되었다.

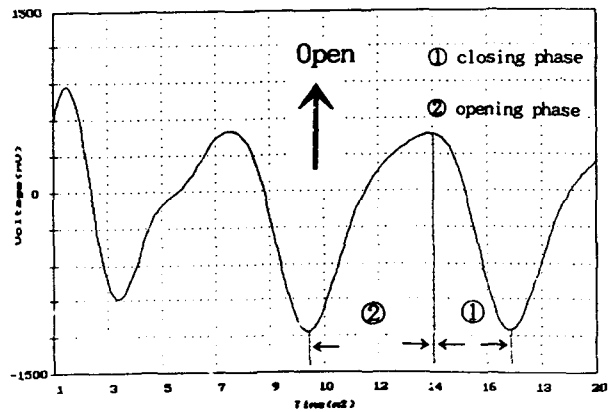


그림 4 성문의 열림과 닫힘구간
Fig. 4 The opening and closing phase of vocal folds

음성파형은 음원과 성도의 전달 임피던스가 결합된 함수이므로, 음원 즉 후두에서의 기동 만을 나타내는 함수를 얻기 위해서는 EGG를 이용하여야 한다. 그림 5는 정상인의 'ㅈ' 발음시 얻어진 EGG의 LPC(linear predicted coefficient)스펙트럼이다. 계산된 예측 차수는 12차로 하였다. 스펙트럼 상에서 최대값 피크가 120 Hz 지점에서 나타났으므로 본 피검자의 경우 'ㅈ' 발음시 피치는 평균적으로 120 Hz가 됨을 알 수가 있다.

그림 6은 통계 패키지인 MINITAB을 이용하여 후두의 질환이 없는 정상 남자와 여자에 대한 주파수 변동과 진폭의 변동을 나타낸 것인데 가로축이 주파수 변동이고 세로축이 진폭변동이다. 그림 7에서 알 수 있듯이 여성의 경우가 주파수 변동 및 진폭의 변동의 크기가 남자보다 뚜렷하게 크다는 것을 알 수가 있다. 이러한 값은 Monsen등[4]이 연구한 바와 같이 정상인 경우, 남자보다 여자의 경우가 피치의 범위가 더 넓다는 것을 설명하는 것이 다.

그림 7은 정상 남녀 그룹을 'A', 후두기능에 이상이 있는 남녀 환자 그룹을 'B'라고 표시 했을 때 8가지 단모음의 평균적인 주파수 변동과 진폭변동에 따른 분포도이다. 정상과 비정상 그룹을 구별할 수 있는 경계값을 정상인에 있어서의 주파수 변동과 진

폭변동 각각의 평균과 표준편차로 나타낼 수가 있었다. 즉 주파수 변동의 $m+2s$ 가 0.193이고, 진폭변동의 $m+2sd$ 가 0.285로서 그림 8에서와 같이 정상그룹과 비정상 그룹을 분류할 수가 있었다.

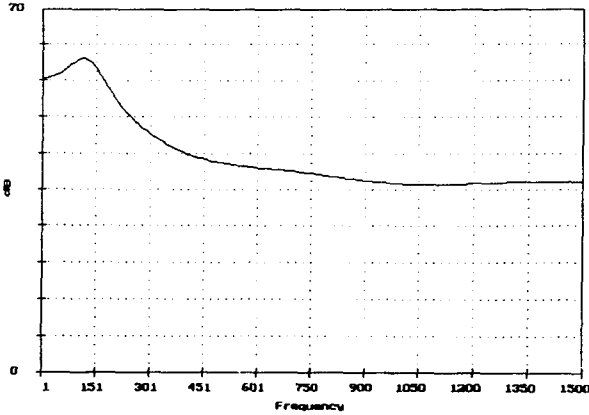


그림 5 EGG신호 'e'에 대한 LPC 스펙트럼
Fig. 5 LPC spectrum of EGG signal during 'e' phonation

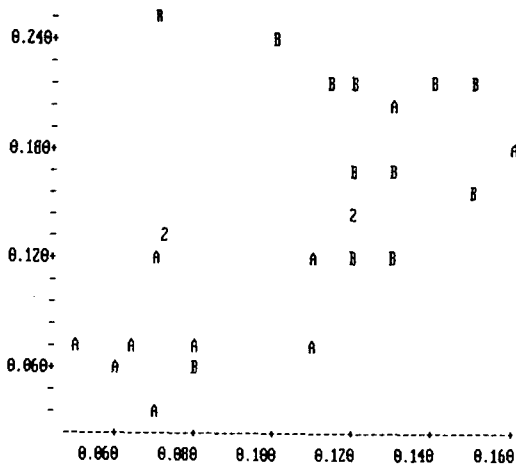


그림 6 정상 남자와 여자에 대한 주파수 변동과 진폭의 변동
Fig. 6 Frequency perturbation and Amplitude perturbation of normal male and female

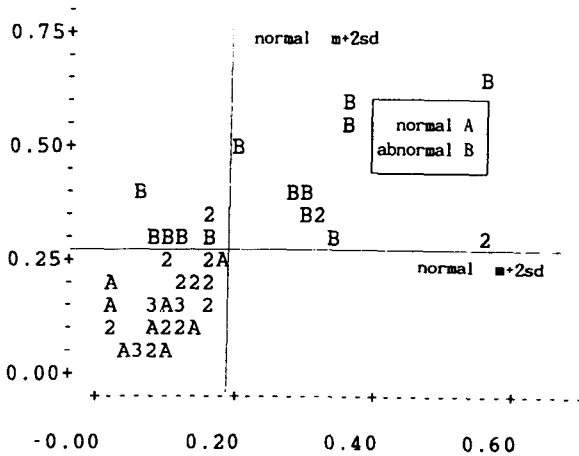


그림 7 정상과 비정상 그룹의 주파수와 진폭변동
Fig. 7 Frequency perturbation and Amplitude perturbation of normal and abnormal group

5. 결론

연속적인 모음의 발생시 EGG의 신호의 주파수 변동과 진폭변동을 분석함으로써 얻어진 결론은 다음과 같다.

1. EGG 시스템을 이용하여 Vocal tract의 전달함수에 영향을 받지 않는 후두에서 의 vocal folds에 의한 피치신호를 검출할 수가 있었다.
2. 주파수 변동과 진폭변동의 크기가 여성의 경우 남성보다 뚜렷하게 컸다.
3. 부드러운 소리 보다는 큰 소리일 때 EGG의 closing 구간이 짧고 급격하였다.
4. 정상음의 경우 기본주수의 변화간격은 여자가 남자보다 간격이 컸고, EGG 신호는 더 대칭적이었다.
5. 주파수 변동과 진폭변동을 이용하여 정상인 그룹 과 비정상인 그룹의 구분을 효과적으로 할 수 있었는데 그 값은 정상인의 주파수 변동 및 진폭 변동의 각각 $m+2sd$ 가 됨을 알 수가 있었다.

앞으로의 연구는 EGG 신호로 부터 vocal folds의 opening 구간과 closing 구간의 비교를 통한 각 질환별 분류와 각 질환에 따른 주파수 변동과 진폭변동의 계수를 이용하여 분포영역의 contour를 산출함으로써 보다 정확한 판별기준을 제시하는 것이다. 아울러 한국어에 대한 유성음원의 추정과 분석, 그리고 음원의 정확한 추정을 통해 합성음의 음질향상을 목적으로 하는 EGG 기반의 음성합성 시스템에 관한 연구가 계속 진행되어야 할 것이다.

6. 참고문헌

- [1] J. L. Flanagan, K. Ishizaka, and K. L. Shipley, "Synthesis of speech from a dynamic model of the vocal cord and vocal tract," B.S.T.J., vol.54, No.3, pp. 420-506, 1975.
- [2] A.K. Krishnamurthy and D.G. Childers, "Two-Channel Speech Analysis," IEEE Trans., Vol. ASSP-34, No. 4, Aug. 1986.
- [3] J.R. Deller, Jr. and D.J. Anderson, "Automatic Classification of Laryngeal Dysfunction Using the Roots of the Digital Inverse Filter," IEEE Trans., Vol. BME-27, No. 12, Dec. 1980.
- [4] R. B. Monsen, A. M. Engerbretson, " Study of variation in the male and female glottal wave," J. Acoust, Soc, Amer., Vol. 62, pp. 228-232, 1977.