

음성을 이용한 후두암의 집단선별검사

동아대학교 의과대학 이비인후과학교실

박헌수

= Abstract =

Acoustic screening test for laryngeal cancer

Heon-Soo Park, MD,

Department of Otolaryngology-Head and Neck Surgery, College of Medicine, Dong-A University, Pusan, Korea

Background and Objectives: Total laryngectomy is often required for advanced cases. But this operation induced the many inconvenience of basic daily life. Early diagnosis of laryngeal cancer is very important to prevent from this disastrous condition. In this point of view, mass screening test for early detection of laryngeal cancer is necessary. Screening test using voice has many advantages such as simple, less interventional. Voice collection by Automatic Response System(ARS) is comfortable and easy to get acoustic sample. Thus author tried to get the acoustic parameters which can differentiate normal, benign, and malignant laryngeal diseases and also checked the availability of parameters on neural network system.

Materials and Methods: Author has evaluated the voice from 17 laryngeal cancer patients and 45 benign laryngeal disease patients who visited at Department of Otolaryngology, Pusan National University Hospital from May 1998 to April 2001, and 15 normal control. Author chose the six parameters (Jitt, vFo, Shim, vAm, NHR, SPI) that was thought to be related with voice collected by ARS among thirty-three parameters analysed by a Multi-Dimensional Voice Program (MDVP). Two-step neural network was used for the availability of six parameters.

Results: The detection rate of normal voice by ARS voice analysis is 78.5% and detection rate of abnormal voice was 97.1%. Among abnormal voice, the detection rate of benign laryngeal diseases and laryngeal cancers were 82.4%, 70.6% respectively.

Conclusion: Author concluded that six parameters and Matlab based neural network software may be effective in development of acoustic screening system for laryngeal cancer and further study should be necessary for development of new acoustic parameters.

Key Words: Screening test · Laryngeal cancer, ARS

서 론

교신처 : 박헌수, 602-715 부산광역시 서구 동대신동3가
1번지 동아대학교 의과대학 이비인후과학교실
전화 : 051) 240-5428, 전송 : 051) 253-0712
E-mail : pbsolkor@unitel.co.kr

후두암은 최근 수술, 방사선 조사, 항암화학요법

등의 발전으로 5년 생존율이 70% 정도에 이르지만, 진행된 경우에는 후두의 광범위 절제 혹은 전 절제가 불가피하여 발성장애로 인한 환자의 삶의 질에 많은 저하를 초래한다. 따라서 후두암의 조기 진단은 생존율의 향상을 물론이고 발성 기능을 보존할 수 있어 후두암의 치료에 있어 매우 중요하다¹⁾.

후두암을 조기에 발견할 수 있는 집단 선별 검사(mass screening test)에 대한 필요성은 인식되어지고 있지만 임상에서 활용할 수 있는 시스템 및 기기의 개발은 거의 이루어지지 않고 있다. 1977년 Koike 등²⁾, 1987년 Mashima 등¹⁾이 몇 가지 변수를 조합하여 후두 질환을 진단할 수 있는 선별 검사를 개발하여 음성을 이용한 후두암의 진단 가능성을 제시하였다. 음성을 이용한 집단선별검사는 음성신호를 음향분석 기기를 통해 분석하므로 검사가 신속, 간편하고, 환자에게 전혀 고통이 없다는 장점이 있다¹⁾.

음성신호를 녹음하는 방법은 몇가지 있지만 그 중 방음실에서 녹음하는 방법이 가장 정확한 음성데이터를 얻을 수 있다. 왕 등³⁾은 방음실에서 녹음한 음성을 이용하여 후두암의 감별율이 87%라고 하였고, 조 등⁴⁾은 새로운 변수를 캡스터럼(Cepstrum) 방법을 이용하여 개발하여 후두질환의 식별을 위해 시행한 결과 정상과 비정상의 감별에 Jitter-Shimmer의 변수를 이용한 경우가 94.5%, 양성과 악성의 감별에 있어서는 NHR-Jitter가 91.7%이었다고 하였다. 그러나 방음실이 있는 곳은 한정되어 있으므로 현대인의 바쁜 생활 속에서 이런 곳을 찾는다는 것은 상당히 어렵다. 그러므로 보다 수월한 방법으로 전화기를 이용한 음성녹음 방법을 사용하기도 한다. 전화기를 이용하여 음성을 녹음한다면 방음실에서 녹음한 음성보다 음성의 질이 떨어지지만, 장소에 구애받지 않고 진단 받을 수 있다는 장점이 있다. 그러나 전화 음성데이터는 일반적인 음성 데이터와는 기본적인 샘플링 레이트나, 대역폭, 잡음성분의 포함정도, 또는 전화선 특유의 시간왜곡 등이 방음실에서 녹음한 음성과 많은 차이가 있다.

방음실에서 녹음한 음성은 잡음을 최소화한 음성 데이터로서 인간의 가청능력인 20Hz-20KHz까지의 대역폭을 모두 포함하는 50KHz의 샘플링 주파수로 샘플링되어 있다. 이에 반해 전화선을 통한 음성은

전화선의 대역폭의 한계인 3,400Hz이내의 음성정보만을 함유하고 있으며 11,025Hz로 샘플링하게 된다. 이런 샘플링 주파수의 차이나 녹음환경의 차이는 차후 파라미터에 상당한 영향을 미치게 된다.

방음실에서 채집한 음성데이터를 이용한 연구는 다소 있으나^{3,4)} 전화선을 이용하여 음성 데이터에 의한 후두암의 선별검사에 대한 연구는 미미한 실정이다. 이에 저자는 전화선을 이용한 정상인, 양성후두질환자, 후두암환자의 음성을 분석하여 어떤 변수가 후두암의 선별검사로서 유용한지를 알아보고자 본 연구를 시행하였다.

대상 및 방법

1. 대상

1998년 5월부터 2001년 4월까지 부산대학교 이비인후과에 내원하여 후두조직검사상 편평상피세포암으로 확진된 후두암 환자 17례, 양성후두질환으로 확인된 48례를 대상으로 하였으며 후두암환자군과 양성질환군을 합한 것을 비정상군으로 구분하였다. 정상대조군으로는 조음에 이상이 없으며 후두 병변이 없고 정상적인 발성이 가능한 성인 15례를 대상으로 하였다.

2. 음성의 수집

가정집에서 전화를 사용한다고 가정 하여 6평정도 되는 방안에서 전화기를 사용하여 발성하였다. 상대적 변수인 비율값을 주로 하는 MDVP의 여러 변수는 동일한 피험자에 의해 발성된 동일한 모음은 같은 음성분석결과가 나온다는 조 등⁴⁾의 연구결과에 의해 평상시와 같은 높이와 크기로 3초간 /아/를 발성하게 하여 녹음하였다. 음성은 전화선을 통해 내부 전화망을 타고 automatic response system(ARS) 보드에 와서 컴퓨터의 wav파일 형식으로 녹음하였다. 이 경우는 wav파일은 ARS 보드에서 지원하는 11,025Hz의 샘플링 레이트에 8bit의 샘플링 비트를 갖고 있다.

3. 음성의 변수

녹음된 음성을 Kay사(Kay Elemetrics Corp. USA)의 Computerized Speech Lab 4300B(CSL) 기

종의 음성분석 소프트웨어인 Multi-Dimensional Voice Program(MDVP)에서 사용되는 33개의 변수 중 6개의 변수를 선택하였다⁵⁾. 이러한 여러 변수중 전화선의 대역폭인 300-3400Hz에 영향이 없을 것으로 생각되는 변수를 제거한 후 주파수 동요, 피치 동요, 소음 등에 연관성이 있는 Jitt, vFo, Shim, vAm, NHR, SPI를 선택하였다⁶⁾.

4. 두 단계 신경회로망(two step neural network)에 의한 정상과 비정상, 양성질환과 악성질환의 감별역전파법(Back-propagation)의 학습방법을 이용한 신경회로망을 사용하였다. 일 단계 신경회로망에서 정상과 비정상을 감별한 후 다시 비정상에서 신경회로망에 의해 양성질환과 악성질환을 감별하는 두 단계의 신경회로망(two step neural network)을 이용하여 감별하였다. 각 단계의 신경회로망은 6개의 변수에 의한 6개의 입력을 가지고 7개의 layer(7 hidden layer)를 갖고 2개의 출력을 갖도록 되어 있다. 출력벡터는 두 개의 값으로 나오는데 두 값 중 첫 번째 값이 크면 정상, 두 번째 값이 크면 비정상으로 식별하였다(Fig. 1).

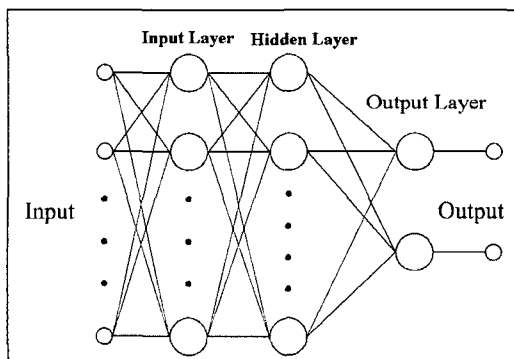


Fig 1. Neural network.

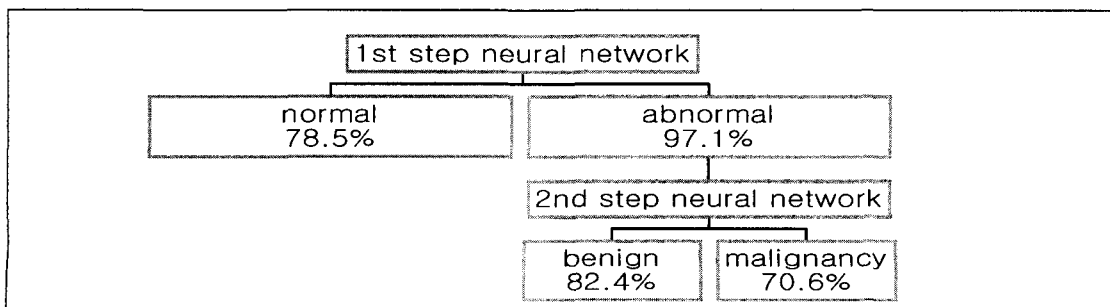


Fig 2. Two step neural network system and sensitivity of each group

결 과

1. 정상과 비정상의 감별

일 단계 신경회로망에서 정상 대조군을 음성 검사상 정상인으로 감별하는 비율은 78.5%이었고, 비정상군을 후두질환으로 감별하는 비율은 97.1%이었다(Fig. 2).

2. 양성과 악성질환의 감별

이 단계 신경회로망에서 양성과 양성질환을 합한 비정상질환군 중 양성질환을 양성질환으로 감별하는 비율은 82.4%이었고, 악성질환을 악성질환으로 감별하는 비율은 70.6%이었다(Fig. 2).

3. 전체적인 후두암의 감별을

음성으로 후두암을 진단할 수 있는 감별율은 68.5%이었다.

고 찰

후두질환에 있어서 가장 흔한 증상은 음성변화이다. 숙련된 이비인후과 의사들은 환자의 음성을 단순히 듣기만 하여도 어느 정도 후두의 병변을 진단할 수 있다. 또한 음성의 변화가 비교적 초기에 나타나기 때문에 후두질환의 조기진단에 음성분석을 이용하기도 한다. 이러한 사실에 입각하여 음성을 이용한 후두질환의 조기진단을 위한 집단 선별검사의 가능성이 제시되었다. 음성분석을 통한 집단선별검사는 환자에게 불편을 주지 않고 단순히 환자의 음성을 녹음하여 음향분석기기로 분석하기 때문에 검사가

간편하여 피검자의 시간적, 공간적 제약을 최소화할 수 있는 장점을 지니고 있기 때문이다.

초기의 연구들은 후두질환에 관련된 병적 음성을 Isshiki 등⁷⁾은 청각학적인 관점에서 분류하였으며, 이러한 분류를 이용하여 Murry 등⁸⁾과 Hammarberg 등⁹⁾은 후두질환 진단에 활용하고자 하였다. 그러나 이 방법은 검사자의 주관이 개입될 여지가 많고 정량화하기가 어려워 검사자간에 차이를 보일 수 있으며 검사 때마다 다를 수도 있다. 이러한 문제는 음성분석 장비의 도입으로 극복할 수 있게 되었고 또한 장비의 비약적인 발달로 그 정확성이 매우 높아졌다.

음성분석을 통하여 여러 가지 후두질환을 감별진단하기 위한 여러 가지 시도가 있었다. 1967년 Yanagihara¹⁰⁾, 1982년 Yumoto 등¹¹⁾, 1986년 Kasuya 등¹²⁾은 병적 음성의 잡음성분을 이용하여 후두질환을 진단하고자 하였다. 1970년 Iwata 등¹³⁾은 발성지수(phonation quotient)를 이용하여 여러 가지 후두질환별 기준치를 제시한 후 이들을 이용하여 후두질환 감별진단에 관한 연구를 수행하였다. 또한 1972년 Iwata¹⁴⁾는 정상인과 후두질환 환자로부터 피치동요에 연관된 변수들의 상관계수를 구하여 이를 도식화한 후 나타나는 특징적인 형태를 기준으로 후두질환 감별진단에 관한 연구를 수행하였다.

피치동요에 연관된 변수를 이용한 연구와는 별도로 여러 가지 변수를 조합하여 후두질환을 감별 진단하려는 연구도 시행되었는데, 1977년 Koike 등²⁾은 여러 가지 후두질환별 기본주파수와 음성강도의 떨림을 측정하고 이들을 도식화한 후 각 질환별 분포곡선을 구하여 후두질환을 감별진단하는 방법을 제안하였다. 그리고 1980년 Horii¹⁵⁾는 음성의 떨림을 객관적으로 측정하기 위하여 음성 Jitt와 Shim를 이용한 연구를 수행하였으며, 1987년 Mashima 등¹⁾은 APQ, PPQ, 음성소음 등을 이용하여 선별검사를 수행할 수 있는 시스템 개발에 관한 예비 연구를 수행하였다. 또한 1989년 Klingholz 등¹⁶⁾은 병적 음성 강도의 분포를 이용한 연구에서 음성 강도의 분포는 후두기능을 잘 반영한다고 결과를 제시하였고, 1990년 Pinto 등¹⁷⁾은 기존의 연구자들이 사용하던 음성신호의 기본 주기와 음성 강도의 떨림을 정의하여 음성신호의 측정시의 문제점들을 규격화하고, 표준화된

측정 기법 등을 제안하였다.

1990년도에 접어들면서 음향학적인 변수 즉, 음성의 특징 파라미터를 이용하여 후두질환 감별진단을 수행하는 연구가 활발히 진행되어 왔다. 1995년 Koizumi 등¹⁸⁾은 성대 결절의 특징을 추정하기 위하여 분석·합성법을 사용하여 효과적인 파라미터 추정에 관한 연구를 하였으며, 1995년 Alwan 등¹⁹⁾은 음성 장애 정도를 측정하기 위해서 시·주파수 합성법에 의한 변수에 관한 연구를 하였다. 1997년 Ciocea 등²⁰⁾은 장애 음성을 감별 진단하기 위해 포먼트-면적 사상(formant-to-area mapping)을 이용하였다.

그러나 많은 연구자에 의해 후두질환 감별진단에 이용할 수 있는 파라미터들이 개발되었지만 파라미터들이 나타낼 수 있는 범위의 한계성 때문에 새로운 파라미터의 개발에 관한 연구가 지속적으로 진행되고 있다. 아울러 기존의 파라미터들을 이용한 개선된 측정 방법도 제안되고 있다.

왕 등³⁾은 방음실에서 녹음된 음성을 MDVP로 분석된 33개의 변수를 이용한 누적도수곡선(Cumulative frequency curve)에 의한 후두암의 집단선별검사를 시행하여, 후두암을 진단하고자 하였다. 또한 조 등⁴⁾은 역시 방음실에서 녹음된 음성을 MDVP로 분석하고 또한 캡스트럼(Cepstrum)을 이용한 새로운 변수를 개발하여 집단선별검사를 시행하여 80%이상의 선별율을 발표하였다. 본 연구는 전화기를 통해 음성을 녹음하여 후두암에 대한 집단선별검사를 시행한 것으로 정상을 정상으로 판단하는 경우는 78.5%이었고, 비정상을 비정상으로 판단하는 경우는 97.1%이었다. 후두질환에서 양성질환을 양성질환으로 감별하는 비율은 82.4%이었고, 악성질환을 악성질환으로 감별하는 비율은 70.6%이었다. 방음실에서 녹음한 음성데이터를 사용한 다른 보고에 비해 결과가 저조하여 추가적인 새로운 변수의 개발이 필요할 것으로 사료된다.

방음실에서 잡음을 최소화하여 녹음한 음성과 전화기를 이용하여 녹음한 음성은 녹음환경과 스펙에 상당한 차이를 보인다. 마이크를 이용하여 방음실에서 DAT로 녹음한 음성은 샘플링 레이트가 50000Hz로 대역폭이 25000Hz이고 비트 레이트가 16bit이다. 그러나 유선전화기로 일반적인 가정집에서 녹음한 음성은 샘플링 레이트가 11025Hz로 대역폭이 3400Hz

이고 비트 레이트가 8bit이다. 그러므로 전화기로 녹음한 음성은 샘플링 주파수의 차이나 녹음환경의 차이로 인한 여러 가지 음성 변수에 상당한 영향을 미친다. 전화기로 녹음한 음성에 대한 음향학적 분석에 대한 연구는 아직 미미하며, 전화기의 음성을 이용하여 질병을 진단하기 위한 변수에 대한 연구는 전무한 실정이다.

MDVP에 의해 분석되는 33개의 변수 중 기본 주파수 Fo [Average Fundamental Frequency]나 기본 진폭값 Ao [Average Fundamental Amplitude] 같은 변수는 음성의 특징을 나타내주는 변수지만, 성별, 나이 등에 따라 사람마다 상당한 편차를 보여준다. 그러므로 질병을 진단하기 위한 선별검사의 변수로는 부적합하다. 그래서 가능한 각 사람의 고유의 특성을 나타내지 않는 변수를 선택하여야 한다. 이러한 변수를 분류해 보면, 기본주파수의 정보에 관한 값 (STD [Standard Deviation of the Fundamental Frequency]), 주파수 동요에 관한 값 ($Jitt$ [Jitter Percent], RAP [Relative Average Perturbation], PPQ [Pitch Period Perturbation Quotient], $sPPQ$ [Smoothed Pitch Period Perturbation Quotient], vFo [Coefficient of Fundamental Frequency Variation]), 진폭 동요에 관한 값($Shim$ [Shimmer Percent], APQ [Amplitude Perturbation Quotient], $sAPQ$ [Smoothed Amplitude Perturbation Quotient], vAm [Coefficient of Amplitude Variation]), 소음에 관한 값(NHR [Noise-to-Harmonic Ratio], VTI [Voice Turbulence Index], SPI [Soft Phonation Index]) 등이 있다. 저자는 이러한 변수 중 전화선의 대역폭이 300-3400Hz이므로 이러한 주파수와 연관성이 있는 변수가 전화기로 녹음한 음성의 분석에 유의할 것으로 생각되는 $Jitt$, vFo , $Shim$, vAm , NHR , SPI 을 사용하였다. $Jitt$ 는 난류비율을 의미하는 변수로 음성표본의 구간별 피치의 변화를 의미하며, 단기간의 상대적인 변이를 나타낸다. 일반적으로 음성의 떨림(perturbation) 현상을 관측하기 위한 음성 분석 변수이다⁶⁾. vFo 는 기본주파수 변이를 의미하는 변수로 구간별로 계산한 기본주파수의 상대적인 표준편차를 나타낸다. 그러나 피치 변이의 형태와는 무관하다⁶⁾.

$Shim$ 은 구간별 진폭치의 상대적인 동요 비율을 나타낸다. 따라서 $Shim$ 를 이용하면 매우 짧은 구간에서 음성의 Peak-Peak 크기의 불규칙성을 측정할 수 있다. 즉 음성 신호에 얼마 정도의 잡음이 섞였는지 아니면 음성 신호의 진폭 변이가 얼마만큼 불규칙적인지를 알 수 있다⁶⁾. vAm 은 정점 진폭 동요를 나타내는 변수로 정점 진폭의 상대적인 표준편차를 의미하며, 장기간의 진폭 동요를 나타낸다. 즉 vAm 은 음성의 각 피치에서의 진폭의 변화를 보여준다⁶⁾. NHR 은 소음 대 배음의 비유로서 70-4500Hz 사이의 배음 성분의 에너지와 1500-4500Hz 범위의 비정상 배음 에너지의 평균 비율을 나타내는 지수이다. 이것은 분석 대상이 된 음성의 전반적인 잡음 특성에 관해 설명해 준다⁶⁾. SPI 는 소음의 측정치를 의미하지 않고 주파수대역상의 배음 구조를 나타낸다. 70-1600Hz의 저주파 성분음의 에너지와 1600-4500Hz의 고주파 성분음의 에너지의 평균비율을 나타낸다⁶⁾.

인간의 학습 능력과 인식 및 추론 능력을 컴퓨터에 부여하여 다양한 동적인 변화에 대응하는 특정 파라미터를 인식하고 추론하여 판단 및 결정을 내리도록 하는 알고리즘의 개발에 관한 연구가 여러 연구자에 의해 진행되고 있다. 그 중 가장 보편적으로 사용되는 알고리즘으로는 신경회로망, 퍼지(fuzzy), 유전자 알고리즘(genetic algorithm) 등이 있다²¹⁾.

신경회로망은 음성, 화상, 제어 분야에서 인간의 정보 처리 능력이 순차적 처리 구조를 갖는 일반 컴퓨터보다 탁월한 성능을 보인다는 사실에 근거했으며, 인간의 정보 처리 과정을 모델링해 간단하고 많은 처리 요소를 병렬로 상호 연결하여, 학습을 통해 입력 패턴에 내재된 정보를 스스로 찾아내고 처리할 수 있도록 하는 기법이다. 이러한 신경 회로망은 음소인식 또는 단어단위 고립단어인식 과 같은 정적인 패턴 분류에 성능이 우수하다고 알려져 있다. 이러한 신경회로망 중 가장 널리 사용되는 것은 다층 퍼셉트론(multilayer perception, MLP)으로 교사학습(supervised learning)에 의한 인식 및 분류, 합성 근사화 등에 사용되고 있다. MLP는 입력층과 출력층 사이에 하나 이상의 중간층(hidden layer)이 존재하는 층 구조이다. 이러한 신경회로망의 학습방법으로는 오차 역전파법(back propagation)을 사용하는데,

이 방법에서는 오차 표면에서 국부 최소점 (local minima)에 빠지게 되면 학습이 더 이상 진행되지 않게 하므로 네트워크의 크기 및 학습률 등을 잘 정해 주어야 한다. 본 연구에서도 back propagation에 의한 7 hidden layer을 가진 신경회로망을 사용하였다²²⁾²³⁾.

음성분석에 있어 이러한 신경회로망의 장점은 첫째, 기본적으로 분별 학습(discriminant learning)을 한다는 점이다. 둘째, 변이가 심한 음성 신호에 대해 일반화(generalization) 성이 뛰어나다. 셋째, 기존의 순차적인 방식과 달리 자체의 구조적인 특성으로 인해 병렬 수행이 가능하여 한번에 많은 양의 자료를 효율적으로 처리할 수 있으며 간단한 하드웨어 구현이 가능하다. 넷째, 일반적으로 제한된 학습 데이터에 대해서 좀 더 강한 특성을 보인다. 이러한 장점들에도 불구하고 신경회로망은 구조적 제약으로 연속 음성 인식으로의 확장이 용이하지 않은 단점이 있다²²⁾²³⁾. 이러한 신경회로망을 이용한 음성분석은 후두질환으로 인한 음성 장애시 음성신호만으로도 질병을 진단하는데 도움이 되며 또한 이러한 기술은 향후 원격 진단용 임상 기기의 개발에 도움을 줄 것으로 생각된다.

결 론

전화기를 통한 정상인, 양성후두질환, 악성후두질환 환자에서 음성을 수집한 후 MDVP의 33개의 변수 중 6개의 변수(Jitt, vFo, Shim, vAm, NHR, SPI)를 이용하여 음성을 분석하였다. 이러한 분석에는 두 단계 신경회로망(two step neural network)을 사용하였다. 식별에 사용된 신경회로망은 back-propagation을 사용한 7 hidden layer, 2출력을 가지는 신경회로망이다. 6개의 변수를 사용할 경우 정상을 정상으로 판단하는 경우는 78.5%이었고, 비정상을 비정상적으로 판단하는 경우는 97.1%이었다. 후두질환에서 양성질환을 양성질환으로 감별하는 비율은 82.4%이었고, 악성질환을 악성질환으로 감별하는 비율은 70.6%이었다. 최종적으로 전화기로 녹음된 후두암환자의 음성으로 후두암을 진단할 수 있을 가능성은 68.5%이었다.

추후 많은 음성의 데이터베이스를 확장하여 더 많은 장애음성의 수집이 필요하며, 또한 전화기에 의해 녹음된 음향 신호의 음향적 특성을 고려한 새로운 변수의 개발에 대한 추가적인 연구가 필요할 것으로 생각된다.

References

- 1) Mashima K, Ebihara S, Kasuya H: Acoustic screening for laryngeal cancer. *Jpn. J. Clin. Oncol* 1987;17:41-7.
- 2) Koike Y, Takhashi H, Calcatera TC: Acoustic measurements for detecting laryngeal pathology. *Acta Otolaryngol* 1977;85:105-17.
- 3) Wang SG, Baek MJ, Yang BG, Jo CW, Park HM, Kweon SB, et al.: Acoustic parameters for the early detection and differential diagnosis of pathologic voice. *Korean J Otolaryngol* 1999;42:1561-7.
- 4) Jo CW, Wang SG, Yang BG.: A study on the diagnosis of laryngeal diseases by acoustic singal analysis. *Korean J Speech Science* 1999;5:151-65.
- 5) Operations manual. Multi-dimensional voice program (MDVP), model 4305, Kay electronic Corp, 1993.
- 6) Operations Manual. Disordered Voice Database, Model 4337, version 1.03, Kay Elemetrics Corp, 1994.
- 7) Isshiki N, Okamura H, Tanabe M, Morimoto M: Differential diagnosis of hoarseness. *Folia Phoniatr* 1969;21:9-19.
- 8) Murry T, Singh S, Sargent M: Multidimensional classification of abnormal voice qualities. *J Acoust Soc Am* 1977;61:1630-5.
- 9) Hammarberg B, Fritzell B, Gaffin J, Sundberg J, Wedin L: Perceptual and acoustic correlates of abnormal voice quality. *Acta Otolaryngol* 1980;90:41-51.
- 10) Yanagihara Y: Significance of harmonic changes

- and noise components in hoarseness. *J Speech Hear Res.* 1967;10:531-41.
- 11) Yumoto E, Gould WJ, Baer T: Harmonic-to-noise ratio as an index of the degree of hoarseness. *J Acoust Soc Am.* 1982;71:1544-50.
 - 12) Kasuya H, Ogawa S, Mashima K, Ebihara S: Normalized noise energy as an acoustic measure to evaluated pathologic voice. *J Acoust Soc Am.* 1986;80(5).
 - 13) Iwata S, Leden H: Phonation quotient in patient with laryngeal disease. *Folia Phoniatica.* 1970;22:117-28
 - 14) Iwata S: Periodicities of pitch perturbation in normal and pathologic larynges. *Laryngoscope.* 1972;82:87-96.
 - 15) Horii Y: Jitter and shimmer in sustained vocal fry phonation. *Folia Phonetica.* 1985;37:81-6
 - 16) Klingholz F, martin F: Distribution of the amplitude in the pathologic voice signal. *Folia Phonetica.* 1989;41:1278-88.
 - 17) Pinto NB, Titze IR: Unification of perturbation measures in speech signals. *J Acoust Soc Am.* 1990;87:1278-89.
 - 18) Koizumi T, Taniguchi S, Mori M, Imazawa A: An analysis-by-synthesis approach to the estimation of vocal cord nodule features. *Proceeding of ICPhS 95.* 1995;2:254-7.
 - 19) Alwan A, Bangayan P, Kreiman J, Long C: Time and frequency synthesis parameters of severely pathologic voice qualities. *Proceeding of ICPhS 95.* 1995;2:250-3.
 - 20) Ciocea S, Schoentgen J, Crevier-Buchman L: Analysis of dysarthric speech by means of formant-to-area mapping. *Proceeding of Eurospeech 97.* 1997;4:1799-1802.
 - 21) Mc Culloch WS, Pitts WH: A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics.* 1943;5:115-33.
 - 22) Martin TH, Howard BD, Mark B. *Neural network design.* PWS, 1996.
 - 23) Lippman RP: An introduction to computing with neural nets. *IEEE ASSP* 1987;4:4-20.