

# 동의보감에 근거한 심장 소리의 음성 분석학적 특징값 추출

가민경\*, 김봉현\*, 이세환\*, 꺾지현\*\*, 조동욱\*\*

\*한밭대학교 컴퓨터공학과

\*\*충북도립대학 정보통신학과

e-mail : kplus@nate.com

## Analytical Voice Feature Values Extraction of Heart Sound Based on Donguibogam

Min-Kyoung Ka\*, Bong-Hyun Kim\*, Se-Hwan Lee\*, Ji-Hyun Kwak\*\*, Dong-Uk Cho\*\*

\*Dept. of Computer Engineering, Hanbat National University

\*\*Dept. of Information & Communications Science, Chungbuk Provincial University

### 요 약

현대 사회에서 건강을 해치는 요인으로 흡연, 당뇨, 비만 및 스트레스 등이 있다. 이와 같은 요인들로 순환기질환의 발병이 증가하고 있으며, 특히 심장 질환 사망률이 점차 증가하고 있는 실정이다. 이를 해결하기 위해 본 논문에서는 심장 질환에 대한 조기 진단을 위한 음성 분석학적 특징 요소를 분석하여 결과값을 추출하고자 한다. 이를 위해 본 논문에서는 대전 지역에 거주하고 있는 성인 남성중에서 심장 질환을 앓고 있는 환자와 심장에 이상이 없는 정상인들로 피실험자 집단을 구성하고 이들의 음성을 수집하여 음성 분석학적 특징 요소들을 추출하고자 한다. 특히 동의보감에서 제시한 심장의 소리를 음성 공학적으로 입증하기 위해 제 5 포먼트와 지터 등의 출력값을 비교, 분석하고자 한다.

### 1. 서론

현대 사회는 서구화된 식습관으로 과거와는 달리 심장에 관한 질병이 증가하고 있다. 특히, 관상동맥질환을 일으키는 주요 위험인자인 흡연을 비롯하여 심장이 펌프질할 때 동맥벽에 미치는 높은 압력인 고혈압, 동맥벽에 침전물을 형성하여 관상동맥이 막힐 수 있는 콜레스테롤, 혈당의 비정상적 상태로 발생하는 당뇨, 과도한 지방 섭취 및 불규칙적인 식사로 인한 비만, 그리고 일상 생활에서 피할 수 없는 스트레스 등은 심장 질환에 대한 발병율을 증가시키고 있는 실정이다[1]. 우리나라의 경우 심장 질환으로 인한 사망률이 3위에 해당하고 있다. 이와 같은 심장 질환에 의한 사망은 돌연사로 이어지고 있으며 이러한 돌연사 원인의 70%가 넘는 비율을 차지하고 있는 심근경색증은 후천성 심장 질환의 대표적인 질병이다. 심장 근육의 일부가 갑자기 죽게 되는 병으로서 발병 시 초기 사망률이 10~20%나 되는 무서운 질병이다.

이와 같이 사회적, 환경적인 변화로 인한 서구화된 식습관, 특히 동물성 지방의 섭취율이 높을수록 심장 질환의 유발이 증가되고 있다. 심장 질환은 음식의 서구화와 함께 과거에 비해 현재 급속도로 발병률이 증가하고 있다. 통계청의 2006년도 사망원인 통계결과에 따르면 암, 뇌혈관질환, 심장질환, 자살, 간질환 순으로 나타났다. 즉, 암, 뇌혈관질환 및 심장질환 등 3대 질환에 의한 사망률이 전체 사망자수의 절반에 가까운 47.6%를 차지하고 있으며 3대

질환 사망비율은 1997년 42.9%에서 해마다 증가하고 있는 추세이다[2].

10년 전과 비교한 지난해 한국인의 10대 사망원인

1996년	(단위:명, 인구 10만명당)	2006년
암 (110.1)	1	암 (134.8)
뇌혈관질환 (74.7)	2	뇌혈관질환 (61.4)
교통사고 (38.3)	3	심장질환 (41.5)
심장질환 (35.7)	4	당뇨병 (23.7)
간질환 (27.3)	5	자살 (23.0)
당뇨병 (17.4)	6	교통사고 (16.0)
자살 (14.1)	7	간질환 (15.6)
만성 하기도질환 (14.0)	8	만성 하기도질환 (14.5)
고혈압성질환 (13.8)	9	고혈압성질환 (9.5)
호흡기결핵 (7.3)	10	폐렴 (9.4)

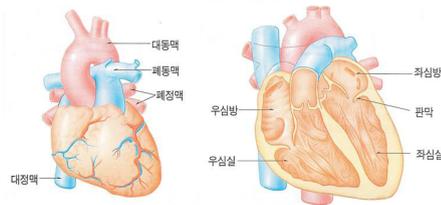
(그림 1) 한국인 10대 사망원인의 변화

이렇게 가장 중요한 기관인 심장 질환이 사망원인에서 큰 부분을 차지하는 것은 사람들의 무지와 무관심 때문이다[3]. 따라서 이와 같이 심장 질환에 대한 사회적 현상을 극복하기 위해 본 논문에서는 동의보감에서 제시하고 있는 심장의 소리에 대한 이론을 기반으로 심장 질환 진단에 필요한 시각적, 객관적 진단 결과를 표현하기 위한 IT 기술과의 연계를 실험을 통해 입증하고자 한다.

## 2. 동의보감에서의 심장 소리

동의보감은 1597년 임금이 병과 건강을 돌보는 어의였던 허준 선생이 선조의 명을 받아 중국과 우리나라의 의학서적을 하나로 모아 편집에 착수하여 1611년에 완성하고 1613년에 간행한 의학서적이다. 동의보감은 모두 23편으로 내과학인 '내경편', '외형편' 4편, 유행병·곽란·부인병·소아병 관계의 '자편' 11편, '탕액편' 3편, '침구편' 1편과 이외에 목록 2편으로 되어있고, 각 병마다 처방을 풀이한 체제정연한 서적이다. 동의보감은 중국과 일본에도 소개되었고, 현재까지 우리나라 최고의 한방의서로 인정받고 있다.

이러한 동의보감에서 목소리가 오장육부와 관련이 있다는 이론이 제시되고 있으며 특히 심장에 병이 들면 웅장한 소리가 나오고 심장은 소리를 받아들여 말을 만드는 인체 장기이기에 심장에 이상이 있을 때 발음이 부정확할 수 있다고 되어 있다. 또한 심장은 신(神)을 간직하는데 신이 실하면 계속 웃고 부족하면 슬퍼한다고 되어 있으며 기뻐하는 것과 웃는 것은 모두 심화(心火)에 속하며 심장에서 나오는 소리가 웃음이 된다고 나타나 있다[4].



(그림 2) 심장의 구조

이와 같이 한의학에서 말하는 심, 심장이라는 것은 물론 서양에서 말하는 가슴 한가운데 약간 왼쪽에서 뛰면서 온몸에 피를 공급하여 주는 펌프로서의 심장을 말하기도 하지만, 한의학에서는 인간과 인간 주위의 환경을 목. 화. 토. 금. 수의 오행으로 나누어 분류하는데, 심장은 화의 성질을 가지는 장기여서 정신적인 스트레스를 오랫동안 받게 되면 심장이 "열 받아" 맥이 빨라지고 혀가 바싹바싹 타면서 목이 마르고 얼굴이 붉어진다고 나타나 있다. 이렇게 겉으로 나타나는 화의 양상과 장기간의 스트레스가 있음을 근거로, 한의학에서는 심장에서 열이 있다고 표현한다. 이렇듯 사람의 마음을 주관하는 곳이 심장이라 정신적 충격, 특히 공포나 놀람 등을 겪게 되면 심장이 놀라서 가슴이 두근두근하고 펄떡펄떡 뛰게 되는데, 이러한 충격이 잦으면 작은 일에도 화들짝 놀라고, 일을 할 때도 대범하게 앞장서서 못하고 조마조마해하는데, 이런 증상을 가리켜 '심장이 약하다'라고 표현한다. 한의학적으로 심장이 약한 사람은 작은 일에도 잘 놀라고 겁이 많으며 소심한 성격을 가지고 있는 경우도 많다. 현대의학의 장기로는 뇌에서 행하는 기능과, 심장에서 행하는 기능을 모두 갖고

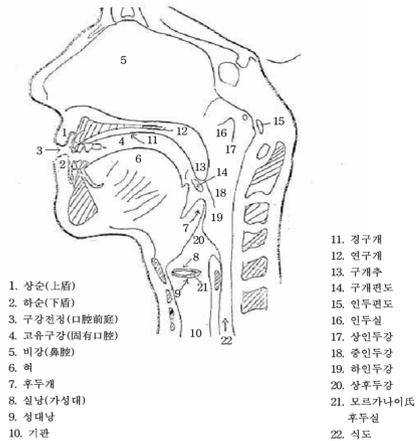
있는 장기라고 파악되고 있다[5].

따라서 본 논문에서는 동의보감에서 제시된 심장의 소리에서 심장에 병이 들면 목소리가 웅장하고 발음이 부정확할 수 있다는 이론적 근거를 바탕으로 목소리의 웅장함을 음성분석 요소 중 공명관과 관련있는 포먼트 분석 결과값과 상관성 분석을 행하였고 발음의 부정확함을 피치 분석 요소 중 지터 분석 결과값과 상관성 분석을 각각 행하였다.

## 3. 공명기관

사람들이 태어나면서부터 갖고 있는 음향관을 공명기관이라 하며 이것은 노래를 부르는 사람들이 공명의 도구로 사용할 수 있으며 공명강 또는 공명체라고도 한다. 공명기관은 호흡운동에 의해 성대에서 발생된 발음을 확대시키는 확장장치이며 이것들이 공명하는 정도의 차이와 공명하는 위치의 차이에 따라 음량과 음색이 변하기도 한다.

일반적으로 공명기관은 소리의 색깔로 변화시킨다. 말소리와 관련된 공명강은 후두강, 인두강, 구강, 비강, 동체, 부비강 등이 있다. 이러한 공명강의 구조는 매우 복잡하기 때문에 사람마다 각각의 차이가 있으며 이로 인해 음색이 다르게 나타나는 것이며 공명강의 역할로는 적절한 형태를 취함으로써 성대를 통해 발생된 소리의 증폭 또는 감속시키는 역할을 수행한다[6]. 즉, 공명에서는 음량과 발생기관의 균형에 의한 과장이 발생하며 이로 인해 소리의 전달력을 갖게 된다. 공명기관의 위치 및 명칭은 아래 (그림 3)과 같다.



(그림 3) 공명강의 위치

## 4. 연구과정 및 방법

### 4.1 연구과정

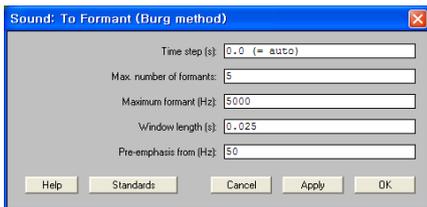
본 논문은 동의보감에서 제시된 심장에 이상이 있을

때 목소리의 음장합과 발음의 부정확성을 음성 분석 요소 추출을 통해 연구를 수행하였다. 이를 위해 심장 질환을 앓고 있는 성인 남성 15명과 동일한 연령대의 정상인 15명을 선정하여 피실험자 집단을 구성하였다. 심장 질환을 앓고 있는 피실험자는 현재 심장 질환 관련 약물 및 치료를 받고 있는 성인 남성들로 선정하였으며 이들은 동일한 음성 입력 장치와 환경에서 “아” 음성을 3초간 연장발성하여 음성을 수집하였다. 수집된 음성은 잡음 여부를 확인하여 분석이 어려운 경우 또는 발음상에 문제가 있는 언어 장애인 경우 음성을 재수집 및 실험 대상자에서 제외하였다.

<표 1> 피실험자 집단 구성표

심장질환자 집단			정상인 집단		
구분	나이	병명	구분	나이	심장상태
H-01	49	관막증	N-01	40	정상
H-02	52	협심증	N-02	54	정상
H-03	43	협심증	N-03	58	정상
H-04	58	부정맥	N-04	51	정상
H-05	59	협심증	N-05	56	정상
H-06	52	심근경색	N-06	47	정상
H-07	48	부정맥	N-07	45	정상
H-08	56	부정맥	N-08	52	정상
H-09	52	협심증	N-09	57	정상
H-10	47	심근경색	N-10	53	정상
H-11	43	심근경색	N-11	56	정상
H-12	59	협심증	N-12	58	정상
H-13	53	관막증	N-13	52	정상
H-14	57	부정맥	N-14	47	정상
H-15	50	협심증	N-15	56	정상
평균	51.87		평균	52.13	

이와 같은 방법에 의해 수집된 음성자료 30개는 음성 분석 프로그램인 Praat를 사용하여 제 5 포먼트와 지터 분석을 수행하였다. 먼저 제 5 포먼트는 Praat에서 Max. number of formants 및 Maximum formant를 각각 5, 5000으로 지정하여 추출하였다. 추출된 포먼트 분석 결과 과형에서 제 5 포먼트에 대한 평균값을 도출하기 위해 전체 60개의 프레임 중에서 10, 20, 30, 40, 50 번째 프레임에 해당하는 5개의 제 5 포먼트값을 출력하여 평균값을 산출하였다.



(그림 4) 제 5 포먼트 추출 요소값 선정

또한 지터값은 Praat의 피치 분석 요소를 이용하여 피치 결과 과형을 추출하고 이를 통해 음성 과형 중 포먼트

와 동일한 구간에 대한 지터값을 출력하였다.

## 4.2 연구방법

인체 내부에는 소리가 진동할 수 있는 관처럼 생긴 모양의 성도(Vocal track)가 있어 공기가 이 부분을 지날 때 소리가 생성되며 말단(성대, 혀, 이빨, 입술) 등에 의해서 최종 음이 발생한다. 이와 같은 성도의 공명을 포먼트(Formant)라 하며 이곳에서 생성되는 과형을 포먼트 주파수라 한다. 또한, 포먼트는 음형대라는 말로도 사용되는데 이는 특정 주파수대에서 에너지가 응집되어 나타나는 것으로 주로 광역 스펙트로그램에서 살펴본다. 즉, 모음은 각각 고유한 포먼트 분포를 가지고 있으며 이 분포를 가지고 해당 모음을 눈으로 추측할 수 있다. 이와 같이 음성 분석에 대한 파라미터 중에 에너지가 집중적으로 나타나는 영역에 대한 분석을 수행할 때 포먼트 분석 요소를 사용한다.

포먼트 주파수는 음성 신호의 LPC(Linear Predictive Coefficient)를 이용해서 추출할 수 있다[7]. LPC는 현재의 출력음성 신호를 과거의 입력 신호와 과거의 출력 신호와의 선형적 결합에 의해 예측할 수 있다[8]. 이것은 음성발생모델과 연관이 있어서 음성에 관한 특징을 적은 수의 파라미터만으로 표현할 수 있고 정확도와 계산속도면에서도 좋은 성능을 보이고 있다. 즉, 음성의 일정구간을 N개의 표본으로 나누면 음성 신호  $s(1), \dots, s(N)$ 에서 한 시점의 음성 신호  $s(n)$ 을 그 이전  $M(M < N)$ 개의 신호  $s(n-1), \dots, s(n-M)$ 에 의해 다음과 같이 표현할 수 있다.

$$s(n) = \sum_{i=1}^M a_i s(n-i) + e(n) \quad M+1 \leq n \leq N \quad (\text{식 1})$$

$$e(n) = \sum_{i=0}^M a_i s(n-i) = s(n) + \sum_{i=1}^M a_i s(n-i) \quad (\text{식 2})$$

또한 지터는 단위시간 안의 발음에서 성대의 진동인 피치의 변화가 얼마나 많은지를 나타내주는 분석 요소로 Momentary pitch period(음성과형 1회 진동)에 대한 Short-term average pitch period(음성과형 3회 진동)길이의 비율을 의미한다. 이는 기본주파수 내 진동시간을 백분율로 표시하며, Relative Average Perturbation 측정용 목적으로 Koike 공식으로 산출한다. 정상적인 음성에서는 변화율이 높지 않지만, 성대에 결절이 있거나 암조적이 있으면 변화가 많게 된다. 성대의 질병여부 또는 음성장애의 평가로 지터가 1.0% 이상이면 병적인 음성으로 판단하는데 흔히 사용한다. 지터의 경우 연령의 증가와 상관이 있다는 연구와 연령의 증가와는 상관이 없다는 연구가 모두 공존한다. 또한 지터는 진동주파수의 한 주기가 얼마나 변동 적인지를 말하여 주는데 청각적으로 큰 목소리와 거친 목소리에 관련이 많다[9].

$$jitter = \frac{\sum_{i=2}^{N-1} |2T_i - T_{i-1} - T_{i+1}|}{\sum_{i=2}^{N-1} T_i} \quad (\text{식 } 3)$$

5. 실험 및 고찰

본 논문에서는 동의보감에서 제시하고 있는 심장의 소리를 공학적으로 분석하기 위해 1차적으로 목소리의 음성합과 공명기관의 부피와의 연관성 분석을 위한 음성 분석 요소인 제 5 포먼트값을 추출하여 연구를 수행하였으며 2차적으로 발음의 부정확성과 음성 분석 요소인 지터와의 연관성을 분석하는 연구를 수행하였다.

아래 <표 2>는 심장 질환자 피실험자 집단의 제 5 포먼트 분석 결과값을 나타낸 것이며 <표 3>은 정상인 피실험자 집단의 제 5 포먼트 분석 결과값을 나타낸 것이다. 표에서 알 수 있듯이 심장 질환자 집단의 제 5 포먼트 분석 평균 결과값이 4,903.96Hz로 정상인 집단의 평균 결과값 4,668.66Hz보다 상대적으로 높게 나타나고 있다.

<표 2> 심장질환자 집단의 제 5 포먼트 분석 결과표

구분	10프레임	20프레임	30프레임	40프레임	50프레임	평균
H-01	5031.54	5034.27	5035.61	5030.24	5027.24	5,031.78
H-02	4892.20	4892.16	4894.57	4895.27	4891.07	4,893.05
H-03	5094.25	5097.83	5097.64	5096.94	5094.27	5,096.19
H-04	4546.38	4544.92	4548.57	4549.62	4547.59	4,547.42
H-05	5083.41	5086.54	5087.06	5084.67	5083.97	5,085.13
H-06	4986.52	4987.05	4989.68	4986.58	4984.37	4,986.84
H-07	4897.62	4896.58	4897.58	4990.18	4898.83	4,916.16
H-08	4894.51	4894.59	4895.28	4894.92	4893.61	4,894.58
H-09	5051.24	5051.38	5052.48	5050.17	5050.34	5,051.12
H-10	4861.52	4861.83	4863.84	4867.72	4864.45	4,863.87
H-11	4497.28	4498.86	4498.31	4498.92	4496.83	4,498.04
H-12	4593.28	4596.52	4594.76	4594.92	4594.31	4,594.76
H-13	5180.31	5180.67	5180.94	5182.46	5182.10	5,181.30
H-14	5022.54	5022.64	5022.97	5021.06	5022.49	5,022.34
H-15	4896.37	4896.82	4896.94	4897.06	4896.62	4,896.76
평균	4,901.93	4,902.84	4,903.75	4,909.38	4,901.87	<b>4,903.96</b>

<표 3> 정상인 집단의 제 5 포먼트 분석 결과표

구분	10프레임	20프레임	30프레임	40프레임	50프레임	평균
N-01	4539.62	4539.83	4540.24	4540.34	4539.67	4,539.94
N-02	4505.31	4505.62	4505.82	4505.63	4505.29	4,505.53
N-03	5022.48	5022.67	5023.64	5026.42	5024.31	5,023.90
N-04	4543.61	4544.67	4544.92	4543.87	4543.37	4,544.09
N-05	4426.35	4426.84	4428.54	4427.64	4426.84	4,427.24
N-06	4764.30	4764.83	4765.34	4765.61	4764.92	4,765.00
N-07	5039.74	5039.62	5040.31	5040.82	5042.34	5,040.57
N-08	4425.19	4425.61	4425.86	4426.34	4426.58	4,425.92
N-09	4997.80	4998.63	4998.86	4990.67	4998.75	4,996.94
N-10	4486.53	4488.64	4488.82	4487.59	4486.05	4,487.53
N-11	4525.65	4526.83	4526.90	4527.61	4526.84	4,526.77
N-12	4600.31	4600.82	4602.94	4601.28	4601.02	4,601.27
N-13	4579.28	4579.68	4580.91	4582.64	4580.33	4,580.57
N-14	5022.54	5023.64	5025.74	5024.39	5024.08	5,024.08
N-15	4539.82	4540.06	4542.18	4540.97	4539.62	4,540.53
평균	4,667.90	4,668.53	4,669.40	4,668.79	4,668.67	<b>4,668.66</b>

또한 아래 <표 4>는 피실험자 집단의 지터 분석 결과값을 나타낸 것으로 심장 질환자 집단의 지터 평균 결과값이 0.569%로 정상인 집단의 지터 평균 결과값인 0.322%보다 상대적으로 높게 나타나고 있다.

<표 4> 피실험자 집단의 지터값 분석 결과표

심장질환자 집단		정상인 집단	
구분	지터(%)	구분	지터(%)
H-01	0.408	N-01	0.240
H-02	0.553	N-02	0.408
H-03	0.362	N-03	0.112
H-04	0.479	N-04	0.301
H-05	0.438	N-05	0.135
H-06	1.220	N-06	0.440
H-07	0.551	N-07	0.127
H-08	1.036	N-08	0.364
H-09	0.405	N-09	0.516
H-10	0.468	N-10	0.418
H-11	0.479	N-11	0.286
H-12	0.841	N-12	0.481
H-13	0.405	N-13	0.418
H-14	0.419	N-14	0.373
H-15	0.468	N-15	0.210
평균	<b>0.569</b>	평균	<b>0.322</b>

6. 결론

본 논문은 우리나라 사망률의 3위에 해당하며 돌연사의 70%를 차지하고 있는 심장 질환에 대한 조기 진단의 방법을 설계한 것으로 동의보감에 제시된 심장의 소리에서 심장에 이상이 있을 때 목소리의 음성합과 발음의 부정확성을 음성 분석학적 요소 추출을 통해 심장질환자의 음성과 상관성 연구를 진행하였다. 이를 위해 본 논문에서는 음성분석 요소 중 제 5 포먼트 분석 결과값과 지터값에서 심장질환자 집단이 정상인 집단보다 상대적으로 결과값이 높게 나타나고 있다는 것을 알 수 있다. 물론 예외에 해당하는 피실험자도 있었지만 이와 같은 방법론적 설계를 통해 심장의 이상 유무를 간단한 방법으로 조기 진단할 수 있는 기초를 마련 할 것으로 생각된다.

참고문헌

- [1] 이종구, 심장병 알면 이길 수 있다, 중앙생활사, 2007.
- [2] 통계청 온라인간행물, <http://www.nso.go.kr/>, 2006.
- [3] 대한심장학회, <http://www.circulation.or.kr/>, 2007.
- [4] 동의과학연구소, 동의보감, 휴머니스트, 2002.
- [5] 김성길, 전문 한의사 108인의 비법, 넥서스, 2003.
- [6] 조기봉, 인중 발성법과 두성, 한솜미디어, 2002.
- [7] Lutz Welling, Hermann Ney, "Formant Estimation for Speech Recognition", IEEE Trans. on Speech and Audio Processing, Vol.06, pp.1063-1076, 1998.
- [8] H. Wakita, "Direct Estimation of the vocal Track shape by Inverse Filtering of Acoustic Speech waveforms", IEEE Trans. A&E, Vol.50, No.02, pp. 637-655, 1971.
- [9] 양병근, 프라트를 이용한 음성분석의 이론과 실제, 만수출판사, 2003.