

피치 요소 추출 및 분석을 이용한 20-30대 남성의 사상체질 태음인 분류 기법

김봉현*, 가민경*, 이세환*, 광지현**, 조동욱**

*한밭대학교 컴퓨터공학과

**충북도립대학 정보통신과학과

e-mail : bhkim@hanbat.ac.kr

Classification Method of Sasang Constitution Taeumin 20-30 Years of Man Using Pitch Element Extraction and Analysis

Bong-Hyun Kim*, Min-Kyoung Ka*, 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

초고령화 사회로 진입하면서 건강의 유지 및 관리는 일상생활에서 매우 중요한 관심으로 대두되고 있다. 특히 우리나라 전통 의학에 대한 가치성이 향상되면서 이를 실생활에 접목하기 위한 노력이 활발하게 진행되고 있는 실정이다. 이를 위해 본 논문에서는 사상의학에 대한 체질 분류의 객관화를 통해 예방 차원의 건강 관리를 지속하기 위한 연구를 진행하고자 한다. 특히 사상의학에서 제안하고 있는 여러 사상체질 분류 방법 중에서 음성을 통한 사상체질 분류 방법을 공학적으로 구현하여 이를 객관적 진단 지표로 설계하고자 한다. 따라서 본 논문에서는 서울지역에 거주하고 있는 20-30대 남성의 음성을 수집하여 음성 분석학적 요소 중 피치 분석 요소들에 대한 추출 및 비교를 통해 사상의학의 사기론에서 제안하는 이론적 근거를 입증하고자 한다. 최종적으로 실험을 통해 추출된 결과를 사상의학 전문의의 진단 결과와 비교, 분석하여 제안하는 기법에 대한 검증을 수행하고자 한다.

1. 서론

최근 들어 초고령화 사회로 진입하면서 우리나라의 평균 수명은 현재 77.90세까지 증가하고 있는 추세이다. 그러나 평균 수명의 증가는 단순히 수명의 연장을 의미하고 있다. 즉 통계학적 수명 연장이 아닌 건강 수명의 연장이 현대 사회가 추구하고 있는 삶의 목표인 것이다. 현재의 건강 수명은 64.30세로 평균 13.60년을 각종 질병에 노출되어 힘든 노후를 보내야 하는 실정이다. 즉 고령화는 자연적인 현상이고 고령화에 의한 생리적, 신체적 요소는 강화보다 약화로 변화되어 간다. 특히 사회적 변화 요인에 대한 심리적 문제와 환경적 변화 요인에 대한 신체적 적응 문제 등이 심각하게 대두되고 있다. 이와 같은 사회적, 환경적 현상과 더불어 질병에 대한 저항의 약화는 각종 만성병과 노인병에 원인이 되고 있으며 이로 인해 노인 의료비 증가가 국민적 부담으로 가중될 것으로 전망된다. 특히 국내 실버 비즈니스 시장규모는 2000년에 17조원에서 2005년에는 약27조원으로 상승하였으며 2010년에는 약 41조원으로 높아질 것으로 예상하고 있다[1].

사상의학은 우리나라 고유의 전통적 체계를 보유하고

있는 의학으로 각 개인마다 체질별 맞춤 진단 및 처방이 행해지는 의학이다. 이와 같은 사상의학은 독창적이고 독보적인 의료체계로서 치료보다는 예방과 보건, 약품보다는 식품을 중요하게 여기는 이론을 기반으로 체질 위주의 처방 접근 방식을 택하고 있다. 이러한 사상의학에서 가장 중요한 것은 사상 체질의 분류이다. 따라서 본 논문에서는 임상과의 용모사기 방법을 계량화, 정량화하여 임상과의 직관을 객관적 기기로 개발하기 위한 방법론을 제안하고자 하며 이를 위해 음성 신호를 분석하여 태음인에 대한 유의성을 파악하여 태음인의 분류를 위한 방법을 제안하고자 한다.

2. 사상의학과 사기론

사상의학은 체질에 따라 치료 및 약물의 처방을 달리 하는 우리나라 전통 의학으로 사상의학에서의 체질은 기본적으로 사람이 4가지 체질, 즉 태양인(太陽人), 소양인(少陽人), 태음인(太陰人), 소음인(少陰人)의 네 가지 체질을 설정하고 있다. 이와 같은 사상체질 분류 이론은 인간을 그 중 하나의 속성을 갖는 매개체로 인식하고 각 체질

별 생리, 병리, 진단, 변증, 치료와 약물에 이르기까지 서로간의 연계를 갖고서 입상에 응용할 수 있는 새로운 방향을 제시한 의학 이론이다. 사상의학은 용약(用藥)보다는 우선적으로 사상 체질을 분류하여 음식물 섭취에 치중하고 이를 바탕으로 질병을 바로 알고 처방을 하자는 접근 방식을 택하고 있다[2]. 따라서 사상의학을 보편화하고 활성화하기 위해서는 가장 중요한 것이 정확한 체질의 분류이다.

사상의학에서 제시하고 있는 사상체질의 분류 방법에는 여러 가지가 있으나 본 논문에서는 음성을 기반으로 한 사기론(詞氣論)을 연구의 배경으로 하였다. 사기론은 사상의학에서 체질을 구별하는 방법 중 용모에서 나오는 기운을 보고 체질을 판단하는 용모사기론(容貌詞氣論)에서 발생한 이론이다. 즉, 얼굴의 형태를 보고 체질을 판단하는 방법과 말하는 기운을 보고 판단하는 방법이 결합된 것을 용모사기론(容貌詞氣論)이라 하며 말하는 기운을 보고 체질을 감별하는 것이 사기론(詞氣論)이다. 이와 같이 목소리를 통해 나타나는 기운을 감지하여 사상체질간 차이를 구분하는 사기론은 사람의 목소리가 개인마다 독특한 특징을 반영하고 있다는 것으로 이러한 사실을 본 논문의 이론적 배경으로 선정하였다.

사상의학 원전에서의 사기론은 태양인이 길고 낮은 상음이며 태음인은 낮고 탁한 공음이고 소양인은 짧고 높은 치음이며 소음인은 높고 맑은 우음이라 명시되어 있다. 그러나 사상의학에서 말하는 각 체질별 음성에 대한 판단 기준은 주관적인 형태로서 이를 객관화하고 정량화하는 것이 필요하다. 즉 청각적 특징은 듣는 이의 직관에 따라 달라진다는 문제점을 극복하기 위해 음성학적 요소에 의한 실험을 통해 음성 정보값을 추출하고 이를 사상의학적 체질 분류 이론과 융합하는 것이 필요하다[3][4].

3. 음성 신호 분석 방법

본 논문에서는 음성 신호를 분석하여 사상체질 진단을 위한 객관적인 변수를 출력하기 위한 연구의 수행으로 동일한 음성 분석 환경에서 결과값을 추출하기 위해 환경과 조건을 통일시켰다. 즉 경희대학교 동서신의학병원 사상체질과에서 일괄적으로 같은 장소와 장비를 가지고 녹음을 하였으며 녹음시에 외부 잡음은 철저히 배제하였다.

이와 같이 본 논문에서 사용된 실험 자료는 사상체질 전문병원에서 수집하였으며 사상체질의학 분야에서 20년 이상 입상에 종사한 전문의로 하여금 각 체질별로 뚜렷한 특징을 보이는 30명의 피실험자 자료를 추출하여 각 체질별 집단군을 10명씩 구성하였다. 또한 녹음 내용은 입상

현장임을 고려하여 최대한 피실험자에게 불편함을 주지 않기 위해 간소화하였다. 실험에 사용된 녹음 자료는 [아]모음을 편안한 상태에서 3초간 발화한 것을 사용하였으며 음의 정확도를 위해 앞과 뒷부분을 제외한 중간부분으로 실험 분석을 행하였으며 이를 위해 수집된 음성에서 1.0초에서 2.0초 사이의 1.0초간을 사용하였다.

[아]모음은 음성 분석에서 일반적으로 널리 사용되고 있는 모음으로 3초간 발화한 것을 기반으로 피치 분석 요소의 최대, 최소 및 평균 추출값 및 피치 변화 범위 등의 음성 분석 결과값을 추출하기 위해 녹음하였다. 또한, 본 논문에서는 피실험자의 음성 분석 결과값을 추출하기 위해 피치 분석 요소를 선정하여 다양한 환경 설정에 따라 분석을 행하였다.

3.1 피치(Pitch) 분석 방법

사람의 목소리의 높이는 음향적으로 f0값으로 나타낸다. 보통 청각적으로 느끼는 음의 높이를 염두에 두고 피치가 높다 낮다라고 표현한다. f0는 기본주파수(fundamental frequency)라고 부르기도 하는데 말하는 사람의 감정과 정서의 변화에 따라 달라진다. 피치 값은 성대의 진동이 1초에 몇 번 있는가를 나타내는 것으로 성대의 크기와 길이, 질량 등에 영향을 받는다[5].

음성 파형의 정보는 일정 구간에서 유사한 모양이 계속적으로 반복되는 형태를 나타낸다. 이러한 반복 주기는 음성 파형을 특징화하는데 가장 중요한 정보를 제공하며 이를 음성 정보의 기본 주기인 피치(pitch)라 한다. 이와 같은 음성 분석 요소를 사용하기 위해 반복주기를 추출하였다. 즉 이산적인 샘플신호 $x(k)$ 가 정제적인 한 신호라고 할 때 샘플간의 유사도는 아래 (식 1)과 같다.

$$R(k) = \sum_{n=-\infty}^{n=\infty} x(n)x(n+k) \quad (\text{식 1})$$

$x(n)$ 과 $x(n+k)$ 는 단지 샘플링 위치만 다를 뿐 통계적인 성질이 동일하다고 가정할 수 있다. 이 때 샘플간의 거리가 k 만큼 떨어진 샘플들간의 유사도는 각 샘플값을 서로 곱하여 이를 무한대에 걸쳐 합을 구한 것으로 측정한다.

$$x(n)x(n+k) = x(n)^2 + x(n+k)^2 - 2x(n)x(n+k) \geq 0 \quad (\text{식 2})$$

$$x(n)^2 + x(n+k)^2 \geq 2x(n)x(n+k) \quad (\text{식 3})$$

3.2 지터(Jitter) 분석 방법

지터는 단위시간 안의 발음에서 성대의 진동인 피치의 변화가 얼마나 많은지를 나타내준다. 즉, 음성파형 1회 진동에 대한 음성파형 3회 진동 길이의 비율을 의미한다. 기본주파수 내 진동시간을 백분율로 표시하며, Relative Average Perturbation 측정을 목적으로 Koike공식으로 산출한다. 정상적인 음성에서는 변화율이 높지 않지만 성대에 결절이 있거나 압조적이 있으면 변화가 많게 된다. 성대의 질병여부 또는 음성장애의 평가로 Jitter가 1.0% 이상이면 병적인 음성으로 판단하는데 흔히 사용한다.

Jitter의 경우 연령의 증가와 상관이 있다는 연구와 연령의 증가와는 상관이 없다는 연구가 모두 공존한다. 또한 Jitter는 진동주파수의 한 주기가 얼마나 변동 적인지를 말하여 주는데, 청지각적으로 목 원소리와 거친 소리에 관련이 많다[6].

$$jitter = \frac{\sum_{i=2}^{N-1} |2T_i - T_{i-1} - T_{i+1}|}{\sum_{i=2}^{N-1} T_i} \quad (\text{식 4})$$

3.3 짐머(Shimmer) 분석 방법

음성파형에서 각 지점의 진폭 값의 변화가 얼마나 규칙적인지를 나타내준다. 성대진폭의 변이 양상을 보여주어 진폭의 정확도를 말하여 준다. 진동 주기당 진폭의 변화를 dB로 평가하며 지속 모음시 음폭의 불규칙한 변화를 백분율로 나타낸다. 정상 성인의 Shimmer는 약 3.8%이다. 성악가들은 음의 크기를 훈련에 의해 서서히 증가시키거나 감소시키는 조절을 할 수 있다. 보통 후두암과 같은 환자의 경우에는 성대의 진동이 불규칙적이면 각 시점의 진폭 값도 일정하지 않게 된다[7].

4. 실험 및 고찰

4.1 음성 분석 실험 및 결과

본 논문에서 음성 신호에 대한 분석을 위해 사용한 실험 환경은 Microsoft Windows XP Professional 을 사용하였으며 음성 데이터 수집 장치로는 스테레오 고품질 녹음 모드인 STHQ 방식의 SONY ICD-SX67을 사용하였고 음성 신호의 비교·분석을 위해 사용한 음성 분석 도구는 Praat 4.2.07의 스펙트럼 분석 요소를 사용하였다.

아래 <표 1>에서 <표 3>은 각 체질별 집단군의 피치 평균값, 피치 최소 및 최대값, 피치값의 변화 등을 나타낸 것으로 태음인이 107.41Hz로 소양인 114.39Hz, 소음인

113.95Hz보다 상대적으로 피치 분석 요소의 평균값이 낮게 형성되고 있는 것을 알 수 있다. 또한 <표 4>에서 <표 6>은 각 체질별 집단군의 지터 및 짐머값에 대한 결과를 나타낸 것으로 뚜렷한 차이는 없지만 태음인의 지터값이 0.401%로 소양인 0.380%, 소음인 0.389%보다 상대적으로 높게 형성되고 있는 것을 알 수 있다.

<표 1> 태음인 집단의 피치 분석 결과값

구분	나이	피치평균	피치최소	피치최대
태음1	27	114.24	107.27	120.81
태음2	34	108.12	99.28	115.27
태음3	38	124.83	108.24	132.85
태음4	32	112.09	106.53	121.57
태음5	25	97.21	93.68	110.54
태음6	36	99.81	92.57	107.95
태음7	37	106.32	99.86	115.01
태음8	29	108.21	100.46	118.94
태음9	34	98.43	93.47	106.86
태음10	33	104.87	96.72	117.91
평균값	32.50	107.41	99.81	116.77

<표 2> 소양인 집단의 피치 분석 결과값

구분	나이	피치평균	피치최소	피치최대
소양1	33	108.95	102.34	117.64
소양2	38	124.34	114.39	131.54
소양3	37	119.20	107.58	127.32
소양4	30	107.52	101.24	112.04
소양5	29	98.27	93.57	109.86
소양6	26	105.85	99.87	110.39
소양7	33	112.46	103.85	128.47
소양8	38	122.41	113.89	139.83
소양9	27	118.04	112.74	125.24
소양10	38	126.84	109.84	131.41
평균값	32.90	114.39	105.93	123.37

<표 3> 소음인 집단의 피치 분석 결과값

구분	나이	피치평균	피치최소	피치최대
소음1	37	116.84	104.25	121.87
소음2	32	107.53	100.67	113.05
소음3	29	116.48	108.49	121.90
소음4	38	126.38	116.74	131.34
소음5	31	109.63	101.27	114.28
소음6	34	111.62	105.68	120.33
소음7	28	98.49	93.68	103.14
소음8	36	108.67	99.84	120.25
소음9	39	127.27	116.94	140.28
소음10	29	116.58	110.68	132.65
평균값	33.30	113.95	105.82	121.91

<표 4> 태음인 집단의 지터, 짐머 결과값

구분	지터(%)	짐머(dB)
태음1	0.295	0.268
태음2	0.342	0.103
태음3	0.351	0.148
태음4	0.511	0.162
태음5	0.590	0.241
태음6	0.355	0.168
태음7	0.396	0.176
태음8	0.493	0.229
태음9	0.263	0.139
태음10	0.413	0.251
평균값	0.401	0.189

<표 5> 소양인 집단의 지터, 짐머 결과값

구분	지터(%)	짐머(dB)
소양1	0.277	0.087
소양2	0.345	0.182
소양3	0.326	0.127
소양4	0.424	0.109
소양5	0.341	0.098
소양6	0.357	0.102
소양7	0.268	0.121
소양8	0.448	0.141
소양9	0.370	0.093
소양10	0.640	0.151
평균값	0.380	0.121

<표 6> 소음인 집단의 지터, 짐머 결과값

구분	지터(%)	짐머(dB)
소음1	0.382	0.154
소음2	0.439	0.157
소음3	0.293	0.245
소음4	0.440	0.206
소음5	0.299	0.141
소음6	0.412	0.144
소음7	0.327	0.180
소음8	0.515	0.248
소음9	0.309	0.354
소음10	0.476	0.124
평균값	0.389	0.195

4.2 재현성 실험

본 논문에서는 제안한 방법의 정확성 및 재현성을 측정하기 위해 방법론에 대한 임의의 실험을 행한 것으로 사상의학 전문의에 의해 뚜렷한 체질적 특징을 보이는 임의의 피실험자 10명을 선정하여 연구 결과와 동일한 출력값을 나타내는지 실험하였으며 아래 <표 7>과 같은 결과를 출력하였다. 5명의 임의의 태음인 피실험자 중에서 4명의 결과가 연구 내용과 일치하는 것으로 나타났다. 이는 80%의 진단 결과가 연구 결과값과 동일하게 나타났음을

알 수 있다. 물론 전문의의 진단이 100% 정확하다는 것도 보장되지 않은 것이 사상의학의 현실이며 객관적 진단 지표에 의한 명확한 분류 방법의 개발이 시급하게 해결되어야 할 과제이다.

<표 7> 태음인 분류를 위한 재현성 실험 결과값

구분	피치평균	지터	음성판단	전문의판단
1	102.51	0.437	태음인	태음인
2	99.82	0.389	태음인	태음인
3	114.27	0.267	소음인	소음인
4	123.96	0.448	소양인	소음인
5	116.24	0.392	소음인	소양인
6	98.35	0.365	태음인	태음인
7	126.83	0.640	소양인	소양인
8	106.82	0.435	태음인	소양인
9	110.24	0.268	소음인	소양인
10	96.72	0.381	태음인	태음인

5. 결론

사상의학은 개인별 체질 기반의 맞춤형 의학 형태를 표현하는 것으로 체질 분류를 기반으로 적절한 건강관리 및 유지 방법을 제안하는 예방, 보건의학이다. 이와 같은 사상의학에서 가장 중요한 것이 사상체질의 정확한 분류이다. 그러나 기존의 사상체질 분류 방법들이 임상외의 직관에 의해 행해지고 있어 신뢰성 및 정확성 측면에서 문제점을 발생하고 있다. 따라서 본 논문에서는 이러한 오류를 극복하기 위해 음성 분석을 통한 정보값을 출력하여 사상체질 집단간 비교·분석을 통해 분류하는 방법을 제안하였다. 끝으로 본 논문은 지식경제부 차세대신기술 사업인 “고령친화형 사상체질 기반의 진단 및 치료기기 개발” 사업비로 수행된 연구 결과임을 부기하는 바이다.

참고문헌

- [1] 한국보건사회연구원, <http://www.kihasa.re.kr>, 2006.
- [2] 이명복, 태양인 이체마 사상의학, 선영사, 2002.
- [3] 손병욱, 사상의학의 이해, 행림출판, 1997.
- [4] 김봉현, 이세환, 조동욱, “20-30대 여성의 사상체질 분류 시스템 구현을 위한 피치 분석 요소 정보 추출 및 분석”, 한국통신학회2008추계종합학술대회발표집, 2008.
- [5] 정옥란, “음성총괄평가”, 대한음성언어의학회지, 1994.
- [6] Boone, D. M. & McFarlane, S. C. The Voice and Voice Therapy(4th ed.), Englewood Cliffs, NJ:Prentice Hall, 1988.
- [7] 이형석 외 5명, “Dr. Speech Science를 이용한 정상 및 후두질환 환자의 음향분석”, 대한음성언어의학회지, 1997.