

사상체질과 음성특징과의 상관관계 연구

권철홍¹ * · 김종열² · 김근호² · 한성만¹

A Study on Correlation between Sasang Constitution and Speech Features

Kwon Chul-Hong¹ · Kim Jong-Yeol² · Kim Keun-Ho² · Han Sung-Man¹

¹Dept. of Information & Communications Eng., Daejeon University

²Korea Institute of Oriental Medicine

Objective : Sasang constitution medicine utilizes voice characteristics to diagnose a person's constitution. In this paper we propose methods to analyze Sasang constitution using speech information technology. That is, this study aims at establishing the relationship between Sasang constitutions and their corresponding voice characteristics by investigating various speech variables.

Materials & Methods : Voice recordings of 1,406 speakers are obtained whose constitutions have been already diagnosed by the experts in the fields. A total of 144 speech features obtained from five vowels and a sentence are used. The features include pitch, intensity, formant, bandwidth, MDVP and MFCC related variables for each constitution. We analyze the speech variables and find whether there are statistically significant differences among three constitutions.

Results : The main speech variables classifying three constitutions are related to pitch and MFCCs for male, and formant and MFCCs for female. The correct decision rate is 73.7% for male Soeumin, 63.3% for male Soyangin, 57.3% for male Taeumin, 74.0% for female Soeumin, 75.6% for female Soyangin, 94.3% for female Taeumin, and 73.0% on the average.

Conclusion : Experimental results show that statistically significant correlation between some speech variables and the constitutions is observed.

Key words : speech technology, Sasang constitution, statistics

I. 서론

사상체질의학에서는 다양한 기준으로 체질을 분류하는 데, 그 기준에 체형, 얼굴 모양, 음성 등 사람의 외적인 특성이 포함되어 있다. 그 중에서

도 음성은 인체의 발성기관과 호흡기능을 중심으로 하여 인체 내부로부터 울어 나오는 중요한 인체 정보원이다. 사상체질의학 관련 문헌¹⁾에서는, “태양인은 호흡기가 크므로 소리가 높고, 맑고 등글다. 태음인은 성량이 풍부하여 소리가 무겁고, 탁하고 모가 난다. 소양인은 호흡기가 작으므로 소리가 가볍고 낮고, 급하고 물러간다. 소음인은 성량이 넓으므로 소리가 활발하고, 완만하고

* 교신저자 : 권철홍, 대전대학교 정보통신공학과

E-mail : chkwon@dju.kr

접수일 : 2011년 1월18일 게재확정일 : 2011년2월8일

평이하다.”라고 말하며 체질별로 음성이 특징 있게 구별되어 있음을 지적한다.

본 논문은 사람의 음성과 사상체질의 상관관계를 다룬다. 즉, 신체의 발성기관과 공명기관 등을 분석하여 얻은 음성학적 특성이 사람의 체질을 구별하는 요소인지를 찾아봄으로써, 음성기술 분야를 사상체질의학과 접목하는 융합연구이다. 여러 음성 특징 중에서 사상체질과 관련이 깊은 변수를 찾아냄으로써, 음성과 체질간의 상관관계를 규명하고자 사상체질 진단을 위한 음성신호처리 방법론을 제시한다.

문진(聞診)을 통해 체질을 진단하는 방법은 진단 기준의 모호성과 객관성의 부족으로 진단에 어려움이 있어, 사상체질 분류의 표준화 및 객관화에 대한 요구가 제기되고 있다. 이러한 이유로 사상체질과 음성공학 기술을 접목하여 체질을 객관적으로 진단하고자 하는 연구가 사상체질의학계에서 진행되어 왔다²⁻⁵⁾. 이러한 연구를 통해 성대의 기본 진동수인 피치의 평균과 분산 그리고 발성 지속 시간 등으로 체질 분류를 시도하여 태음인의 저음 사용이 소음인 보다 많다는 결과를 일부 확인하여 음성을 통한 체질분류에 대한 가능성을 보여 주었다. 그런데 평균 피치를 측정하는데 사용되는 문장이 통제되지 않는 등 음성학적인 관점에서 볼 때 방법상의 문제가 존재한다⁶⁾.

음성학을 전공하는 인문학자가 음성학적 분석을 바탕으로 사상체질을 분류하는 연구도 있다^{6,7)}. 이 연구에서는 사상체질의학 전문가들로부터 체질을 판정받은 23명의 음성 데이터를 수집하여, 체질별로 평균 기본주파수, 최고 기본주파수, 최저 기본주파수, 피치 범위, 발성 유형, Jitter, Shimmer, 발화속도 등에서 체질을 분류해 주는 변수를 찾고자 하였다. 이 연구는 여러 가지 음성 특성 중 음원과 직접적으로 관련이 되는 변수들을 주로 고려하였고, 발성유형이 체질별로 약간 일관된 패턴을 보인다는 결과를 도출하였으나 23명의 녹음자료라는 데이터의 부족으로 통계적으로 유의한 결과라고 보기는 어렵다.

본 논문은 음성기술을 이용하여 사상체질 분류

가 가능한지를 밝히려고 하는 연구이다. 기존 연구에서는 주로 음원에 대한 변수만 다룬 반면, 본 논문에서는 지금까지 음성공학계에서 음성의 특성을 연구하는데 사용된 성도 필터에 관계된 변수와, 임상분야에서 널리 사용하고 있는 MDVP(Multi-Dimensional Voice Program) 변수들을 모두 포함하여 연구함으로써, 사상체질을 구분하는 음성 변수를 찾아보고자 한다.

II. 본론

1. 실험 방법

1) 실험 대상 및 음성 데이터

음성데이터 수집 대상 피험자 나이는 10대 ~ 80대 사이에 분포하며, 성별과 체질별로는 남자 511명(소음 118명, 소양 174명, 태음 219명)과 여자 895명(소음 247명, 소양 318명, 태음 330명) 등 총 1,406명을 대상으로 하였다. 회귀 체질로 분류되는 태양인 샘플은 본 연구에서 제외하였다.

체질이 명확한 사람을 대상으로 하기 위해, 5년 이상 사상체질의학 임상분야에 경험이 있는 전문가에 의해 동일 체질 처방을 60첩 이상 투여 받고 최소 4회 이상 방문하여 경과를 관찰한 사람 중 부작용이 없으며 주증이 일정 정도 이상 호전된 사람, 혹은 동일 체질 처방을 20첩 이상 60첩 미만 투여 받고 부작용이 없으며 주증과 소증이 일정 정도 이상 호전된 사람을 피험자로 선정하였다⁸⁾.

음성 데이터는 Sennheiser e-835s 마이크를 이용하여 조용한 공간에서 수집하였다. 피험자는 편안히 앉은 상태에서 5개의 모음(아, 에, 이, 오, 우)을 각각 2초 이상 자연스럽게 발성하고, 하나의 문장(“우리는 높은 산에 올라가 맑은 공기를 마시고 왔습니다.”)을 두 번 발성하였다. 음성 데이터는 PCM signed 16bits, mono 형식으로 샘플링 주파수 16kHz로 수집하였다⁸⁾.

2) 음성 변수

수집된 1,406명의 음성 데이터로부터 5개의 모음과 2개의 문장에서 총 144개의 음성 변수를 추출하였다⁸⁾. 사용된 음성 변수는 피치(pitch), 세기(intensity), 포먼트(formant), MDVP 등이 다.

5개의 모음 구간에 대해 피치, 세기, 1,2차 포먼트 및 3dB 대역폭과, 임상분야 연구에서 많이 활용하는 MDVP 변수를 추출하였다. MDVP 변수로는, 발성 시 피치의 변화정도를 나타내는 Jitter 계열 변수들(Jitta, Jittp, RAP3, PPQ5)과, 진폭의 변화 정도를 나타내는 Shimmer 계열 변수들(Shimdb, Shimp, APQ5)을 다음과 같이 구하였다.

(식 1)

$$\begin{aligned}
 Jitta &= \frac{1}{N-1} \sum_{i=1}^{N-1} |T_{o,i} - T_{o,i+1}| \\
 Jittp &= \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_{o,i} - T_{o,i+1}|}{\frac{1}{N} \sum_{i=1}^N T_{o,i}} \times 100 \\
 RAP3 &= \frac{\frac{1}{N-2} \sum_{i=1}^{N-2} \left| \frac{T_{o,i} + T_{o,i+1} + T_{o,i+2}}{3} - T_{o,i} \right|}{\frac{1}{N} \sum_{i=1}^N T_{o,i}} \times 100 \\
 PPQ5 &= \frac{\frac{1}{N-4} \sum_{i=1}^{N-4} \left| \frac{T_{o,i} + T_{o,i+1} + T_{o,i+2} + T_{o,i+3} + T_{o,i+4}}{5} - T_{o,i} \right|}{\frac{1}{N} \sum_{i=1}^N T_{o,i}} \times 100 \\
 Shimdb &= \frac{1}{N-1} \sum_{i=1}^{N-1} |20 \log(A_{i+1}/A_i)| \\
 Shimp &= \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_i - A_{i+1}|}{\frac{1}{N} \sum_{i=1}^N A_i} \times 100 \\
 APQ5 &= \frac{\frac{1}{N-4} \sum_{i=1}^{N-4} \left| \frac{A_i + A_{i+1} + A_{i+2} + A_{i+3} + A_{i+4}}{5} - A_i \right|}{\frac{1}{N} \sum_{i=1}^N A_i} \times 100
 \end{aligned}$$

여기서 T_o 는 피치, A 는 세기를 나타낸다.

음성기술 분야에서 우수한 성능을 보이며 널리 사용되고 있는 특징 파라미터인 MFCC (Mel-Frequency Cepstral Coefficients)도 사

용하였다. 이 변수는 각 모음별로 12차 MFCC 변수와 1개의 에너지 변수를 추출하였다

문장 구간에 대해서는 피치와 세기를 구하여, 두 변수의 10, 50, 90번째 백분위수와 이들 간의 비(Pratio, Iratio), 피치와 세기의 상관도 (CorrPD)를 구하였다.

<표 1>에 본 연구에서 사용한 음성 변수에 대해 정리되어 있다⁸⁾.

Table 1. Definition of Speech Variables

	변수 이름	정의
문장	CorrPI	피치와 세기의 상관도
	P10	피치의 10% 백분위수
	P50	피치의 50% 백분위수
	P90	피치의 90% 백분위수
	Pratio	(P90-P50)/(P50-P10)
	I10	세기의 10% 백분위수
	I50	세기의 50% 백분위수
	I90	세기의 90% 백분위수
Iratio	(I90-I50)/(I50-I10)	

	변수 이름	정의
모음	xFo	평균 기본주파수
	xTo	평균 피치
	xSTD	기본주파수의 표준편차
	xJitta	Jitter 절대값
	xJittp	Jitter 백분율 값
	xRAP3	Relative Average Perturbation
	xPPQ5	Pitch Perturbation Quotient
	xShimdb	Shimmer 데시벨 값
	xShim	Shimmer 백분율 값
	xAPQ5	Amplitude Perturbation Quotient
	xF1	첫번째 포먼트
	xBw1	첫번째 포먼트 대역폭
	xF2	두번째 포먼트
	xBw2	두번째 포먼트 대역폭
	xMFCC1~12	MFCC 1~12차
	xC0	에너지

※ 변수 이름 앞의 x는 5개 모음(a, e, i, o, u) 중 하나를 지칭한다.

2. 연구 내용 및 실험 결과

1) ANOVA 분석

앞에서 기술한 144개의 음성 변수를 이용하여 ANOVA(Analysis of Variance) 분석을 시행하여, 어느 변수가 집단을 구분하는데 통계적으로 유의한 차이가 있는지를 검출하여 체질을 분류하는데 사용한다. ANOVA는 세 개 이상의 모집단 평균 간의 차이를 검증하는 데 이용하는 분석방법이다⁹⁾. 본 연구에서 모집단(사상체질)이 3개여서 ANOVA 분석이 가능하므로, ANOVA 분석을 실시하여 집단 간에 차이를 보여 주는 음성 변수를 구한다.

ANOVA 분석을 시행하여 구한 통계적으로 유의미한($p < 0.05$) 변수와 평균값은 <표 2, 3>에 보인다. 표에서 보듯이, 남자의 경우 Fo 계열 변수와 MFCC 변수, 여자의 경우 포먼트 계열 변수와 MFCC 변수가 세 체질을 구분하는 주요 변수임을 알 수 있다.

Table 2. Constitution Classification Variables and Their Mean Values for Male

음성변수	SE	SY	TE
aFo	128.6	138.4	134.5
eFo	132.0	143.5	139.7
iFo	135.5	149.9	144.6
oFo	135.6	150.7	145.0
uFo	138.0	152.9	147.0
iTo	7.68	7.08	7.35
oTo	7.67	7.04	7.33
uTo	7.55	6.96	7.27
aSTD	1.40	2.66	2.19
eSTD	1.26	1.87	2.11
uSTD	1.37	2.18	2.85
uJitta	23.7	25.5	30.0
uJittp	0.31	0.36	0.42
uPPQ5	0.20	0.23	0.27
eF2	1825.0	1770.7	1796.8
oF2	934.9	1004.6	1103.9
uF2	1165.9	1111.5	1206.1
aBw2	244.0	241.2	299.7
CorrPI	0.134	0.167	0.291
P10	107.6	116.7	112.9

P50	127.0	139.6	135.5
P90	145.4	160.1	153.6
iMFCC2	-6.08	-7.20	-7.93
aMFCC3	0.32	2.88	2.11
eMFCC3	4.53	6.33	6.06
aMFCC4	-6.42	-10.52	-9.03
eMFCC4	2.94	-2.45	-0.43
iMFCC4	18.0	13.67	15.37
oMFCC4	1.74	-2.83	-0.41
uMFCC4	4.39	0.23	2.20
aMFCC8	13.62	11.17	10.41
eMFCC8	-6.34	-9.97	-10.30
iMFCC8	-1.72	-6.14	-5.78
oMFCC8	2.47	0.13	-0.63
uMFCC8	0.66	-2.36	-1.66
iMFCC11	-5.71	-7.97	-7.43
uMFCC11	-4.73	-7.01	-4.99
aMFCC12	0.48	-1.79	-0.97

Table 3. Constitution Classification Variables and Their Mean Values for Female

음성변수	SE	SY	TE
uTo	4.64	4.91	4.73
aBw1	289.5	323.9	285.6
uBw1	161.6	147.8	182.8
aF2	1428.4	1396.4	1389.6
eF2	2245.5	2219.1	2193.4
iF2	2626.6	2619.2	2580.5
iMFCC2	-6.92	-5.87	-7.61
uMFCC2	7.95	8.80	7.55
eMFCC6	-22.86	-20.64	-19.73
uMFCC6	-17.62	-15.62	-17.99
aMFCC8	2.96	4.05	5.06
aMFCC9	-9.73	-8.79	-7.77
eMFCC9	-7.42	-6.65	-5.67
iMFCC9	-15.16	-14.73	-13.20
oMFCC9	-8.32	-7.65	-6.57
eMFCC11	-5.84	-4.84	-3.85
aMFCC12	-4.18	-3.89	-2.71
eMFCC12	-7.72	-7.31	-6.17
oMFCC12	-6.23	-5.76	-4.63
uMFCC12	-11.18	-10.15	-9.06

2) 요인 분석(Factor Analysis)

<표 2, 3>에서 보듯이 체질 분류 변수가 다수 이므로 요인분석을 시행하여 요인변수를 구한다. 요인 분석은 변수들 간의 상관관계를 이용하여 변수의 차원을 축약하거나 그룹으로 묶는 방법이다⁹⁾.

남자의 경우 요인 분석한 결과 (식 2)와 같이 8 개의 요인변수로 축약되었다. 요인변수를 보면 같은 유형의 변수들이 묶이는 경향이 있음을 알 수 있다. 즉, MFA1은 Fo 계열 변수들, MFA2는 MFCC 8차, MFA3는 Jitter 계열 변수, MFA4는 MFCC 4차, MFA5는 MFCC 3차, MFA6은 포먼트와 그 대역폭 등이 묶여 있다.

(식 2)

$$\begin{aligned} \text{MFA1} &= (0.943 * \text{ZoFo} + 0.934 * \text{ZiFo} \\ &+ 0.933 * \text{ZeFo} - 0.932 * \text{ZoTo} \\ &+ 0.927 * \text{ZuFo} + 0.922 * \text{ZP9o} \\ &+ 0.921 * \text{ZP5o} - 0.918 * \text{ZiTo} \\ &+ 0.910 * \text{ZaFo} - 0.900 * \text{ZuTo} \\ &+ 0.875 * \text{ZP1o}) \\ \text{MFA2} &= (0.814 * \text{ZoMFCC8} + 0.774 * \text{ZuMFCC8} \\ &+ 0.761 * \text{ZeMFCC8} + 0.723 * \text{ZiMFCC8} \\ &+ 0.707 * \text{ZaMFCC8}) \\ \text{MFA3} &= (0.976 * \text{ZuJittp} + 0.972 * \text{ZuJitta} \\ &+ 0.870 * \text{ZuPPQ5} + 0.742 * \text{ZuSTD}) \\ \text{MFA4} &= (0.805 * \text{ZeMFCC4} + 0.773 * \text{ZaMFCC4} \\ &+ 0.707 * \text{ZoMFCC4} + 0.667 * \text{ZiMFCC4} \\ &+ 0.563 * \text{ZuMFCC4}) \\ \text{MFA5} &= (0.791 * \text{ZeMFCC3} + 0.736 * \text{ZaMFCC3}) \\ \text{MFA6} &= (0.731 * \text{ZuF2} + 0.721 * \text{ZoF2} \\ &+ 0.336 * \text{ZaBw2}) \\ \text{MFA7} &= (0.867 * \text{ZeF2}) \\ \text{MFA8} &= (0.487 * \text{ZuMFCC11} + .373 * \text{ZiMFCC11} \\ &- 0.371 * \text{ZaMFCC12}) \end{aligned}$$

여자의 경우 요인 분석한 결과 (식 3)과 같이 5개의 요인변수로 축약되었다.

(식 3)

$$\begin{aligned} \text{FFA1} &= (0.784 * \text{ZaMFCC9} + 0.781 * \text{ZoMFCC9} \\ &+ 0.780 * \text{ZeMFCC6} + 0.771 * \text{ZiMFCC9} \\ &+ 0.743 * \text{ZeMFCC9} + 0.685 * \text{ZeMFCC12} \\ &+ 0.578 * \text{ZuMFCC12} + 0.503 * \text{ZaMFCC12} \\ &- 0.362 * \text{ZiF2}) \\ \text{FFA2} &= (0.882 * \text{ZiMFCC2} + 0.833 * \text{ZuMFCC2} \\ &+ 0.587 * \text{ZuMFCC6} + 0.582 * \text{ZuT0} \\ &+ 0.533 * \text{ZaBw1}) \\ \text{FFA3} &= (0.688 * \text{ZaMFCC8} + 0.601 * \text{ZeMFCC11} \\ &- 0.494 * \text{ZeF2} - 0.367 * \text{ZaF2}) \\ \text{FFA4} &= (0.553 * \text{ZoMFCC12}) \\ \text{FF5} &= (0.345 * \text{uBW1}) \end{aligned}$$

3) 판별 분석(Discriminant Analysis)

앞에서 구한 요인변수(요인점수)로 판별분석을 시행하여 남자의 경우는 (식 4), 여자의 경우는 (식 5)와 같은 분류함수를 구하였다. 분류함수를 구하는데 사용된 음성변수들은 남자의 경우 주로 MFCC 변수이고 일부 포먼트와 대역폭, Jitter 계열 변수를, 여자의 경우는 주로 MFCC 변수이고 일부 포먼트와 대역폭 변수를 포함하고 있다.

(식 4)

$$\begin{aligned} \text{MD_SE} &= 0.346 * \text{MFA2} - 0.214 * \text{MFA3} \\ &+ 0.343 * \text{MFA4} - 0.254 * \text{MFA6} - 1.607 \\ \text{MD_SY} &= -0.074 * \text{MFA2} - 0.017 * \text{MFA3} \\ &- 0.240 * \text{MFA4} - 0.137 * \text{MFA6} - 1.112 \\ \text{MD_TE} &= -0.128 * \text{MFA2} + 0.129 * \text{MFA3} \\ &+ 0.006 * \text{MFA4} + 0.246 * \text{MFA6} - 0.884 \end{aligned}$$

(식 5)

$$\begin{aligned} \text{FD_SE} &= -0.091 * \text{FFA2} - 0.116 * \text{FFA3} \\ &- 0.222 * \text{FFA4} - 1.316 \\ \text{FD_SY} &= 0.168 * \text{FFA2} - 0.081 * \text{FFA3} - 1.049 \\ \text{FD_TE} &= -0.094 * \text{FFA2} + 0.165 * \text{FFA3} \\ &+ 0.166 * \text{FFA4} - 1.023 \end{aligned}$$

4) 체질 분류 결과

판별분석을 이용하여 체질 분류한 결과와 사상

체질의학 전문가의 체질 분류 결과와의 일치도가 <표 4, 5>에 보인다. 남자의 경우 소음인 28.0%, 소양인 46.0%, 태음인 60.7%, 평균적으로 48.1%의 일치도를 보인다. 여자의 경우 소음인 8.9%, 소양인 51.3%, 태음인 58.8%, 평균적으로 42.3%의 일치도를 보인다. 소음인의 일치도가 낮고 태음인의 일치도가 높은 결과를 보여 주었다.

Table 4. Classification Results for Male(%)

체질	SE	SY	TE	평균
SE	28.0	19.5	52.5	48.1
SY	11.5	46.0	42.5	
TE	12.8	26.5	60.7	

Table 5. Classification Results for Female(%)

체질	SE	SY	TE	평균
SE	8.9	47.8	43.3	42.3
SY	6.6	51.3	42.1	
TE	4.5	36.7	58.8	

3. 군집분석을 이용한 체질 분류

앞에서 기술한대로 음성 데이터 수집 대상자는 사상체질을 전문으로 하는 한의사가 약을 처방하여 호전된 사람을 대상으로 체질을 판정하였으며 올바르게 체질이 진단되었다고 말할 수 있다. 그런데 한의사는 체질을 판정하는 기초 데이터로 음성만 이용한 것이 아니라 얼굴 모양, 체형도 고려하고 또한 상담을 통하여 판정하였다. 즉 체질을 판정하는데 음성이 유일한 데이터가 아니다. 한의사는 그 사람의 체질을 음성만으로 판정하는데 때로는 어려움이 있다고 말한다. 따라서 음성이 체질 판정의 결정적 요소가 되는 사람들을 따로 분류하여 실험하는 것이 본 연구와 부합하다고 판단되어 군집분석이라는 방법을 사용하게 되었다. 즉, 동일한 체질을 갖는 모집단을 음성변수를 이용하여 군집분석을 시행하여 몇 개의 집단으로 분할하여 동질성을 갖는 집단을 대상으로

체질 판정을 시행하고자 한다.

1) 군집 분석(Clustering Analysis)

군집분석은 다양한 데이터나 실험의 결과로 얻어진 측정값들의 특성을 동질적인 몇 개의 군집으로 분류하는 방법이다⁹⁾. 군집분석은 무엇을 기준으로 해서 데이터를 분류하느냐에 따라 상이한 결과가 나올 수 있다. 대표적인 두 가지 방법은 첫째, 관측 대상 간에 정해지는 유사성을 기초로 해서 비슷한 것끼리 하나의 그룹으로 묶어 전체를 몇 개의 그룹으로 분할하는 것이다. 둘째, 데이터 및 그룹 내의 분산 개념을 기초로 해서 그룹 간의 분리 정도를 기준으로 분류하는 것이다. 여기서는 전자의 방법을 사용하였는데, 이 방법은 특성들의 유사성, 즉 특성 자료가 얼마나 비슷한 값을 가지고 있는지를 거리로 환산하여 거리가 가까운 대상들을 동일한 집단으로 묶는다.

본 연구에서는 K-평균 군집분석을 시행하였는데, 이 방법은 사전에 정한 군집 수 K에 기초하여 전체 데이터를 상대적으로 유사한 K개의 군집으로 분류하는 방법이다⁹⁾. 즉 데이터 사이의 거리를 이용해 거리가 가까운 데이터를 같은 군집에 포함시키는 최적 분리 군집 방법으로, 군집 간 변이에 비해 군집내 변이를 최소화하는 방법이다. 이 방법은 다음과 같은 과정을 거쳐 군집을 분류한다.

- ① 군집의 수 K와 초기 K개 군집의 중심을 선정한다.
- ② 각 데이터를 그 중심과 가장 가까운 거리에 있는 군집에 할당한다.
- ③ 각 군집별로 그에 속하는 데이터를 이용하여 새로운 군집중심을 계산한다.
- ④ ①과 ②의 과정을, 기존의 중심과 새로운 중심의 차이가 없을 때까지 반복한다.

각 체질별로 군집분석을 시행한 결과, 남자의 경우 소음인은 총 118명 중에서 76명, 소양인은 총 174명 중에서 89명, 태음인은 총 219명 중에서 75명이 동질성을 갖는 한 그룹으로 분류되었다. 여자의 경우 소음인은 총 247명 중에서 104명, 소양인은 총 318명 중에서 119명, 태음인은

총 330명 중에서 106명이 한 그룹으로 분류되었다. 이 때 각 체질별로 그룹으로 묶이는 데이터 수가 비슷하게 하였다.

2) ANOVA 분석

군집 분석한 남자 240명(=76+ 89+ 75)과 여자 329명(=104+ 119+ 106)의 데이터를 이용하여 ANOVA 분석을 시행한다.

세 체질을 구분하는 ANOVA 분석을 시행하여 구한 통계적으로 유의미한(p < 0.05) 변수는 <표 6, 7>과 같다. 표에서 보듯이, 남자, 여자 모두 Fo 계열 변수와 MDVP 변수, 그리고 MFCC 변수가 세 체질을 구분하는 주요 변수로 작용하였다.

Table 6. Constitution Classification Variables for Male

체질간	분류 변수
소음 vs 소양	aFo, eFo, iFo, oFo, uFo, aTo, eTo, iTTo, oTo, uTo, aJitta, eJitta, iJitta, oJitta, uJitta, aJittp, eJittp, iJittp, oJittp, uJittp, aRAP3,eRAP3,iRAP3,oRAP3,uRAP3, ePPQ5,iPPQ5,oPPQ5,uPPQ5,aShimdb, eShimdb,iShimdb,oShimdb,uShimdb, aShim,eShim,iShim,oShim,uShim, aAPQ5,eAPQ5,iAPQ5,oAPQ5,uAPQ5, aF1, eF1, iBw2, P50, P90, oMFCC1,uMFCC1,aMFCC2,eMFCC2, iMFCC2,oMFCC2,uMFCC2,aMFCC4, eMFCC4,iMFCC4,oMFCC4,uMFCC4, aMFCC5,eMFCC5,oMFCC5,uMFCC5,oMFCC7,uMFCC7,iMFCC8,uMFCC8, iMFCC11,uMFCC11,eMFCC12
소음 vs 태음	aJitta, eJitta, iJitta, oJitta, uJitta, aJittp, eJittp, iJittp, oJittp, uJittp, aRAP3,eRAP3,iRAP3,oRAP3,uRAP3, ePPQ5, iPPQ5, oPPQ5, uPPQ5, aShimdb, eShimdb, oShimdb, aShim, eShim, oShim,

	aAPQ5, eAPQ5, oAPQ5, eF1,iBw1,eMFCC2,iMFCC2,uMFCC2, eMFCC4,oMFCC7,uMFCC7,oMFCC10
소양 vs 태음	eFo, iFo, oFo, uFo, eTo, iTTo, oTo, uTo, P50, P90, uF2, oMFCC4, uMFCC4, aMFCC5, eMFCC5,iMFCC5,oMFCC5,uMFCC5, eMFCC8,iMFCC8,oMFCC8,uMFCC8, eMFCC11, uMFCC11, aMFCC12

Table 7. Constitution Classification Variables for Female

체질간	분류 변수
소음 vs 소양	aFo, aTo, eTo, oTo, uFo, uTo, aJitta, aJitt, aRAP3, aPPQ5, aBw1, aMFCC12, iMFCC3, iMFCC10, oMFCC1, oMFCC12, uMFCC12
소음 vs 태음	aFo, eFo, iFo, oFo, uFo, aTo, eTo, iTTo, oTo, uTo, aShimdb, aShim, uShimdb, uShim, aAPQ5, uAPQ5, iJitta, aC0, eC0, iC0, oC0, uC0 aF1, aBw1, aBw2, eF1, eF2, iF1, iBw1, oF1, oF2, uF1, uBw1, uBw2, CorrPI, P10, P50, P90, I10, I50, I90, MFCC1,aMFCC2,aMFCC3,aMFCC4, aMFCC6,aMFCC10,eMFCC1,eMFCC2, eMFCC3,eMFCC4,eMFCC6,eMFCC11, iMFCC2,iMFCC4,iMFCC6,iMFCC7, oMFCC2,oMFCC4,oMFCC5,oMFCC6, oMFCC8,oMFCC10,oMFCC12, uMFCC1,uMFCC2,uMFCC4,uMFCC5, uMFCC6,uMFCC7,uMFCC8,uMFCC9, uMFCC10,uMFCC11,uMFCC12
소양 vs 태음	eFo, iFo, oFo, uFo, eTo, iTTo, oTo, uTo, P50, P90, uF2, oMFCC4, uMFCC4, aMFCC5, eMFCC5,iMFCC5,oMFCC5,uMFCC5, eMFCC8,iMFCC8,oMFCC8,uMFCC8, eMFCC11, uMFCC11, aMFCC12

4) 군집 분석 체질 분류 결과

앞에서 구한 체질간 분류 변수로 판별분석을 시행하여 남자의 경우는 (식 6), 여자의 경우는 (식 7)과 같은 분류함수를 구하였다. 분류함수를 구하는데 사용된 변수들은 남자의 경우 주로 MFCC 변수이고 일부 MDVP 변수를, 여자의 경우는 주로 MFCC 변수이고 일부 MDVP와 포먼트 변수를 포함하고 있다.

(식 6)

$$\begin{aligned}
 MD_SE &= 0.513*iRAP3 + 0.416*oRAP3 \\
 &- 0.219*P90 - 0.229*aMFCC12 \\
 &+ 0.484*eMFCC4 + 0.470*iMFCC2 \\
 &+ 0.189*iMFCC8 - 0.002*uMFCC11 \\
 &- 1.671 \\
 MD_SY &= -0.249*iRAP3 - 0.185*oRAP3 \\
 &+ 0.303*P90 - 0.152*aMFCC12 \\
 &- 0.210*eMFCC4 - 0.210*iMFCC2 \\
 &- 0.340*iMFCC8 - 0.279*uMFCC11 \\
 &- 1.282 \\
 MD_TE &= -0.222*iRAP3 - 0.200*oRAP3 \\
 &- 0.142*P90 + 0.415*aMFCC12 \\
 &- 0.239*eMFCC4 - 0.224*iMFCC2 \\
 &+ 0.217*iMFCC8 + 0.337*uMFCC11 \\
 &- 1.356
 \end{aligned}$$

(식 7)

$$\begin{aligned}
 FD_SE &= 0.137*aShim - 0.228*aBw1 \\
 &+ 0.124*iJitta - 0.819*oF0 \\
 &+ 0.384*oF2 - 0.151*uBw2 \\
 &+ 0.323*CorrPI+ 0.283*eMFCC4 \\
 &- 0.551*iMFCC10+ 0.878*oMFCC2 \\
 &+ 0.344*oMFCC10 - 0.328*oMFCC12 \\
 &+ 0.578*uMFCC4 + 0.082*uMFCC5 \\
 &+ 0.173*uMFCC6 - 0.329*uMFCC11 \\
 &- 1.981 \\
 FD_SY &= 0.430*aShim + 0.447*aBw1 \\
 &- 0.308*iJitta - 1.152oF0 - 0.166*oF2 \\
 &- 0.495*uBw2 + 0.131*CorrPI \\
 &+ 0.429*eMFCC4 + 0.05*iMFCC10
 \end{aligned}$$

$$\begin{aligned}
 &- 0.014*oMFCC2 + 0.5*oMFCC10 \\
 &+ 0.352*oMFCC12 + 0.512*uMFCC4 \\
 &- 0.435*uMFCC5 + 0.664*uMFCC6 \\
 &- 0.278*uMFCC11 - 2.074 \\
 FD_TE &= -0.617*aShim - 0.278*aBw1 \\
 &+ 0.224*iJitta + 2.097*oF0 - 0.191*oF2 \\
 &+ 0.704*uBw2 - 0.465*CorrPI \\
 &- 0.759*eMFCC4 + 0.484*iMFCC10 \\
 &- 0.846*oMFCC2 - 0.898*oMFCC10 \\
 &- 0.073*oMFCC12 - 1.142*uMFCC4 \\
 &- 0.569*uMFCC5 - 0.915*uMFCC6 \\
 &+ 0.634*uMFCC11 - 4.579
 \end{aligned}$$

군집분석을 적용한 뒤 판별분석을 이용하여 체질 분류한 결과와 사상체질의학 전문가의 체질 분류 결과와의 일치도가 <표 8, 9>에 보인다. 남자의 경우 소음인은 73.7%, 소양인은 63.3%, 태음인은 57.3%, 그리고 평균적으로 64.7%의 일치도를 보인다. 여자의 경우 소음인은 65.4%, 소양인은 77.3%, 태음인은 93.4%, 그리고 평균적으로 78.7%의 일치도를 보인다.

Table 8. Classification Results using Clustering Analysis for Male(%)

체질	SE	SY	TE	평균
SE	73.7	11.8	14.5	64.7
SY	15.6	63.3	21.1	
TE	14.7	28.0	57.3	

Table 9. Classification Results using Clustering Analysis for Female(%)

체질	SE	SY	TE	평균
SE	65.4	32.7	1.9	78.7
SY	21.8	77.3	0.8	
TE	3.8	2.8	93.4	

III. 고찰 및 결론

1. 피험자는 성별과 체질별로 남자 511명(소음 118명, 소양 174명, 태음 219명), 여자 895명(소음 247명, 소양 318명, 태음 330명)으로 총 1,406명이 다. 1,406명의 음성 데이터로부터 5개의 모음과 2개의 문장에서 피치, 세기, 포먼트, MDVP, MFCC 등 총 144개의 음성 변수를 추출하였다.

2. 144개의 음성 변수로 ANOVA 분석을 시행하여 세 집단을 구분하는데 통계적으로 유의한 차이가 있는 변수를 검출하였다. 남자의 경우 피치 계열 변수와 MFCC 변수, 여자의 경우 포먼트 계열 변수와 MFCC 변수가 세 체질을 구분하는 주요 변수임을 알 수 있다.

체질 분류 변수가 다수인 관계로 요인분석을 시행하여 요인변수를 구하였다. 요인변수를 보면 같은 유형의 변수들이 묶이는 경향이 있음을 알 수 있다.

요인변수(요인점수)로 판별분석을 시행하여 분류함수를 구하였다. 분류함수를 구하는데 사용된 요인변수들은 남자의 경우 주로 MFCC 변수이고 일부 포먼트와 대역폭 변수 그리고 Jitter 계열 변수를, 여자의 경우는 주로 MFCC 변수이고 일부 포먼트와 대역폭 변수를 포함하고 있다.

판별함수를 이용하여 체질 분류한 결과와 사상체질의학 전문가의 체질 분류 결과와의 일치도는, 남자의 경우 소음인 28.0%, 소양인 46.0%, 태음인 60.7%, 평균적으로 48.1%의 일치도를 보이고, 여자의 경우 소음인 8.9%, 소양인 51.3%, 태음인 58.8%, 평균적으로 42.3%의 일치도를 보였다.

3. 동일 체질을 갖는 모집단을 음성변수를 이용하여 군집분석을 시행하여 몇 개의 집단으로 분할하여 동질성을 갖는 집단을 대상으로 체질 판정을 시행하였다.

K-평균 군집분석을 시행하여 선정된 군집에 대해, 144개의 음성 변수로 ANOVA 분석을 시

행하여 세 집단을 구분하는데 통계적으로 유의한 차이가 있는 변수를 검출하였다. 남자, 여자 모두 피치 계열 변수와 MDVP 변수, 그리고 MFCC 변수가 세 체질을 구분하는 주요 변수로 작용하였다.

이 변수들로 판별분석을 시행하여 분류함수를 구하였다. 분류함수를 구하는데 사용된 변수들은 남자의 경우 주로 MFCC 변수이고 일부 MDVP 변수를, 여자의 경우는 주로 MFCC 변수이고 일부 MDVP와 포먼트 변수를 포함하고 있다.

군집분석을 한 뒤 판별분석을 이용하여 체질 분류한 결과와 사상체질의학 전문가의 체질 분류 결과와의 일치도는, 남자의 경우, 소음인 73.7%, 소양인 63.3%, 태음인 57.3%, 평균적으로 64.7%의 일치도를 보이고, 여자의 경우 소음인 74.0%, 소양인 75.6%, 태음인 94.3%, 평균적으로 81.2%의 일치도를 보였다.

4. 앞에서 보듯이 모든 음성 데이터(1,406명)를 대상으로 체질 판정을 시도한 결과는 50% 이하의 분류 성능을 보여 주나, 군집분석을 통해 동질성을 갖는 집단을 대상으로 체질 판정을 시행한 결과 상당히 좋은 성능을 보여 주었다. 이 결과로부터 음성이 체질 판정의 결정적 요소가 되는 사람들을 따로 분류하는 절차가 선행되는 것이 필요함을 알 수 있다.

5. 종합적으로 살펴보면, 세 체질을 구분해 주는 주요 변수는 남자의 경우 MFCC 변수와 피치 계열 변수이고 여자의 경우 MFCC와 포먼트 계열 변수임을 알 수 있다.

6. 세 체질을 구분하는 음성 변수가 존재하고, 이 변수를 이용하여 판별 분석한 결과 좋은 성능을 얻은 결과로부터, 음성기술을 이용하여 사상체질을 분류하는 시도가 타당성이 있음이 입증되었다. 향후 연구로는 음성인식에서 사용되는 HMM(Hidden Markov Models) 기법이나 기계학습 분야에서 사용하는 SVM(Support Vector Machine)과 CART(Classification and

Regression Tree) 등을 적용하여 연구를 진행할 계획이다.

감사의 글

본 연구는 한국한의학연구원 기관고유 사업인 체질건강수준 표준개발 과제 (K10070)와 지식경제부 및 한국산업기술평가관리원의 산업원천기술개발사업 (10028438, 오감형 한방 진단/치료 콘텐츠 개발)의 지원으로 수행되었음.

참고문헌

1. 김달래. 동의수세보원초고. 정담출판사, p. 134, 1999.
2. 김달래, 박성식, 권기록. 성문분석법에 의한 사상체질 진단의 객관화 연구(I). 사상의학회지 10(1):65-80, 1998.
3. 신미란, 김달래. CSL을 통한 음향특성과 사상체질간의 상관성 연구. 사상체질의학회지, 11(1):137-157, 1999.
4. 양상목, 김선형, 유준상, 김형석, 이영훈, 김달래. Pitch range와 bandwidth를 이용한 음향 특성과 사상체질간의 상관성 연구. 사상체질의학회지, 13(3):31-39, 2001.
5. 박성진, 김달래. Harmonics와 formant bandwidth를 이용한 음성특성과 사상체질간의 상관성 연구. 사상체질의학회지, 16(1):61-73, 2004.
6. 문승재, 박종주, 황혜정. 음성과 사상체질: 음원을 중심으로. 대한음성학회, 말소리, 48:19-33, 2003.
7. 문승재, 탁지현, 황혜정. 음성학적으로 본 사상체질. 대한음성학회, 말소리, 55:1-14, 2005.
8. 강재환, 유종향, 이해정, 김종열. 음성을 이용한 사상체질 분류 알고리즘. 한국음성학회, 말소리와 음성과학, 1(3):155-161, 2009.
9. 장원경, 김태균. 한글 SPSS를 이용한 자료 분석의 이해와 적용. 도서출판 대경, 2005.