사상체질의학회지 2011;23(4):458-469
J of Sasang Constitutional Medicine
Vol. 23, No. 4, 2011:458-469
http://www.esasang.com

음성 특성 및 음성 독립 변수의 사상체질 분류로의 적용 방법

김근호, 강남식, 구본초, 김종열*
한국한의학연구원*

Abstract

Application of Vocal Properties and Vocal Independent Features to Classifying Sasang Constitution

Keun-Ho Kim, Nam-Sik Kang, Bon-Cho Ku, Jong-Yeol Kim*
Korea Institute of Oriental Medicine

1. Objectives
Vocal characteristics are commonly considered as an important factor in determining the Sasang constitution and the health condition. We have tried to find out the classification procedure to distinguish the constitution objectively and quantitatively by analyzing the characteristics of subject's voice without noise and error.

2. Methods
In this study, we extract the vocal features from voice selected with prior information, remove outliers, minimize the correlated features, correct the features with normalization according to gender and age, and make the discriminant functions that are adaptive to gender and age from the features for improving diagnostic accuracy.

3. Results and Conclusions
Finally, the discriminant functions produced about 45% accuracy to classify the constitution for every age interval and every gender, and the diagnostic accuracy was meaningful as the result from only the voice.

Key Words : vocal feature, reduction of correlation, removal of outliers, normalization, diagnostic accuracy

공시사항: 본 연구는 지식경제부 및 한국산업기술평가관리원의 산업원천기술개발사업 (10028438, 오감형 한방 진단/치료 컨텐츠 개발) 및 교육과학기술부 특정연구개발사업 (20100020617)의 일환으로 수행되었음.

• 접수일 2011년 10월 04일; 심사일 2011년 10월 10일;
승인일 2011년 11월 21일
• 교신저자 : 김종열
대전광역시 유성구 전민동 461-24(한국한의학연구원)
Tel : 042-868-9489 Fax : 042-868-9480
한의사 면허 번호: 9456 E-mail : ssmed@kiom.re.kr

I. Introduction

Sasang constitution medicine (SCM) gives different treatment ways to the same symptom since it divides 4 types (TY, SY, TE, SE) of human according to one's appearance and personality.[1-2] In the traditional Korean medicine (TKM), voice inspection is one of four diagnosis methods including visual inspection, auscultation, palpation and a survey. Voice inspection has been an important theme for studying Sasang constitution.[3-6] The connection between Sasang constitution and voice is illustrated in 「Sasang-Inhaejinam Saseongron」 as follows: TY voice has high-pitch tone, derived from strong respiratory organs. Its voice is clean and smooth, which is also matched to sound of Shang (商音), in the traditional Chinese five sounds. TE has a large capacity in voice sound and therefore it feels like heavy, thick and gentle. As Yin is superior to Yang, the deepness and the capacity of sound are large enough. Since SY has weak respiratory organs, its sound is low-pitch tone and light. It sounds like pressing and can be spread out widely at ease. Since SE has rich voice sound, it sounds like lively, slow and easy. Yin is superior to Yang, and thus the deepness and the capacity of sound are large enough.[3]

Based on the TKM connection between Sasang constitution and voice, various researches that intend to interpret the relation scientifically and quantitatively have been tried so far. In 2004, Park and Kim acquired significant results with trying objective diagnosis by using pitch, shimmer, the difference in dB between h1 and h2 among harmonics as the difference of voice strength and formant bandwidth for personal identification.[7]

In 2005, Kim, et. al investigated the constitutional properties of the Korean adult women by using pitch, APQ, shimmer, energy and octave as vocal features.[8] In 2006, Kim, et. al researched the characteristics of Sasang constitutional type in the children from 6 to 12 years old with voice features.[9] In 2006, Choi, et. al studied the significant characteristics of diverse features from the adult male voice.[10]

Nonetheless, many researches including prior stud ies have proposed just the possibility by extracting a few vocal features and then applying them to statistical parametric estimation or have produced meaningless results in classifying constitution types even by adopting several pattern classification methods. In this research, application of various features, and the normalization of vocal features according to the vocal properties of gender and age for constitution classification are proposed. Moreover, it is provided to obtain independent variables by eliminating explanatory variables for quantitative and reliable discriminant analysis methods.

II. Methods and Materials

1. Voice acquisition environment

At first, for voice acquisition we used personal computers (PC) as hardware and the exterior sound card to avoid noise from the PC. In the experiment, we selected Sound Blaster Live 24-bit External as a soundcard and a Sennheiser e-835s model as a voice recording-only microphone. By using a stand the distance between a microphone and one's mouth should be 5 cm, and the main axis of cylinder of the microphone should be fixed in order to be parallel to ground and perpendicular to one's mouth. We used the GoldWave[11] as an audio recording software and saved the voice file as a WAV file. The sampling frequency is 44.1 kHz for high quality recording.

Each subject was asked to take a seat comfortably and speak naturally without tension and with feeling like I'm the only one in the room and maintaining ordinary volume and speed of voice. Uttering vowels and sentences, the subject keeps silence over 1 s at first, then speaks 'a', 'e', 'i', 'o' and 'u' for 3 s and keeps silence between the vowels for 1 s. Then one keeps 1 s silence and speaks twice the given sentence with keeping 1 s silence between the sentences and at the tip of the final sentence for finishing the recording. Before this series of experiments began, a standard operating procedure (SOP) like the previous acquisition procedure was established to work as a safeguard against gaps between individual operators. All procedures then followed this SOP.

The constitution of every subject was confirmed by SCM doctors in some Korean medicine clinics by recording reactions and observing one's improvements after administrating constitution-specific pharmaceuticals.

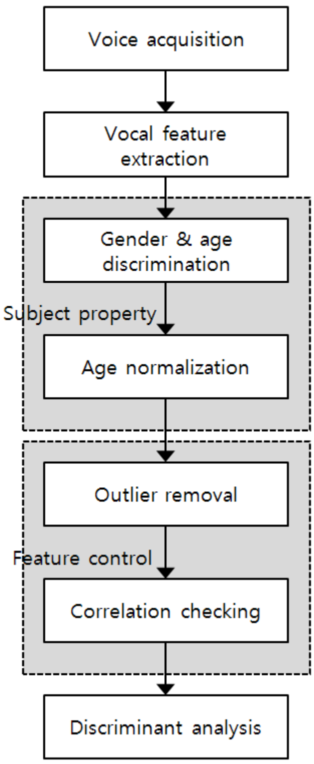

Figure 1. The flow of the overall procedure

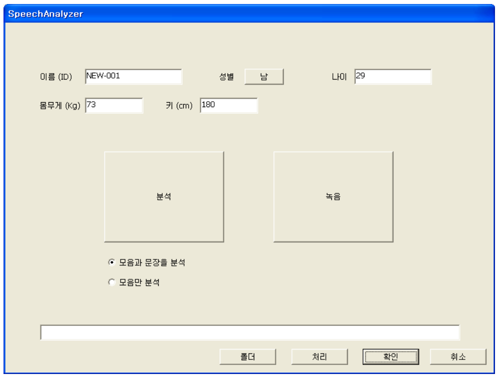

Figure 2. The vocal feature extraction program

2. Extraction of voice features

We implemented the program in Figure 2 combined with HTK[12] and PRAAT[13] through C++ for acquisition of the voice features. The voice features of vowels and a sentence are extracted from the voice wave file captured in the given environment as shown in Figure 1, which were selected with prior information since they express most of characteristics like frequency derivative features, variation of frequency, frequency band ratios, MFCCs and intensities of the vowels and the sentence. The window size of the feature extraction is 40 ms. The voice features are shown in Figure 3 and 4, respective for the vowels and the sentence. The features of the vowels are F0 (average pitch frequency), F1 and F2 (the first and second terms of formant frequencies)[14], JITA and PPQ (jitter and its average: variation of pitch frequency), MFCC4 (the fourth term of MFCC: mel-frequency cepstral coefficient) that can show the recognition of the voice pattern[15], iF0-aF0 (the difference between 'i' pitch frequency and 'a' one) and uF0-iF0 (the difference between 'u' pitch frequency and 'i' one) which show the relative difference of pitch frequencies, and the ratios of voice energies over fixed frequency bands such as 60~120Hz, 120~240Hz, 240~480Hz, 480~960Hz, 960~1920Hz and 1920~3840Hz. The features of the sentence are F50 (the 50^{th} percentile of pitch frequencies), I50 (the 50^{th} percentile of intensity), SF0 (average pitch fre-

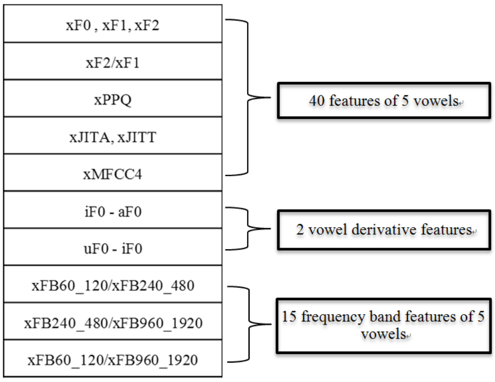

Figure 3. Voice features of 5 vowels

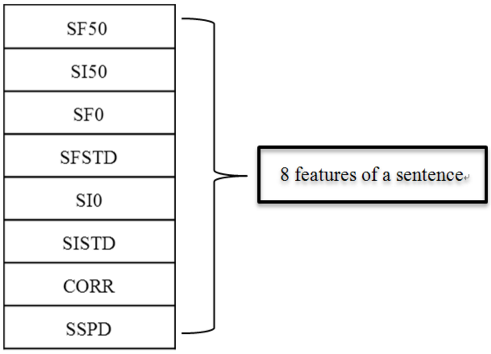

Figure 4. Voice feature of a sentence

quency), SFSTD (variation of pitch frequency), SI0 (average intensity), SISTD (variation of intensity), and CORR (Pearson correlation coefficient[16] between SF0 and SI0).

3. Division on gender and age and age correction

As the next step, Figure 5 shows the F0 and the F1 of females are higher than those of males and the differences are large significantly. In addition, Figure 6 reveals F0 of 'a' (aF0) increases linearly with age axis from the early 20s to the 70s in case of males, where aF0s of SY are larger than those of SE. In case of females, aF0 decreases linearly with age from the teenage years to the 70s, where aF0s of SE are larger than those of SY before the 60s, but vice versa after the 60s. The figure does not illustrate the aF0s of TY types since they cannot be analyzed statistically by reason of the small number of TY types (less than 1%). Due to the reason, we should analyze the dataset by excluding the teenager and TY, and dividing them into the age interval such as the 20s, the 30s ~ the 40s and above the 50s according to gender and correct each voice feature with its mean value over the age interval, which is age correction.

4. Removal of outliers

In the next step, if even one among features from a subject's voice data is outside 3×IQR (inter-quartile range[17]), the voice data is considered as an outlier and excluded for analysis, and the outlier interferes to make an accurate discriminant. The remaining data without the

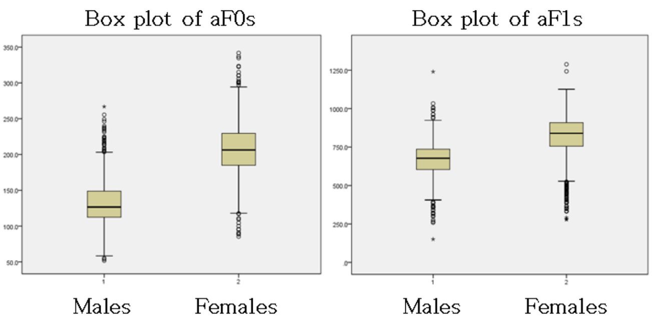

Figure 5. Analysis of the means in voice features (aF0 and aF1) between genders

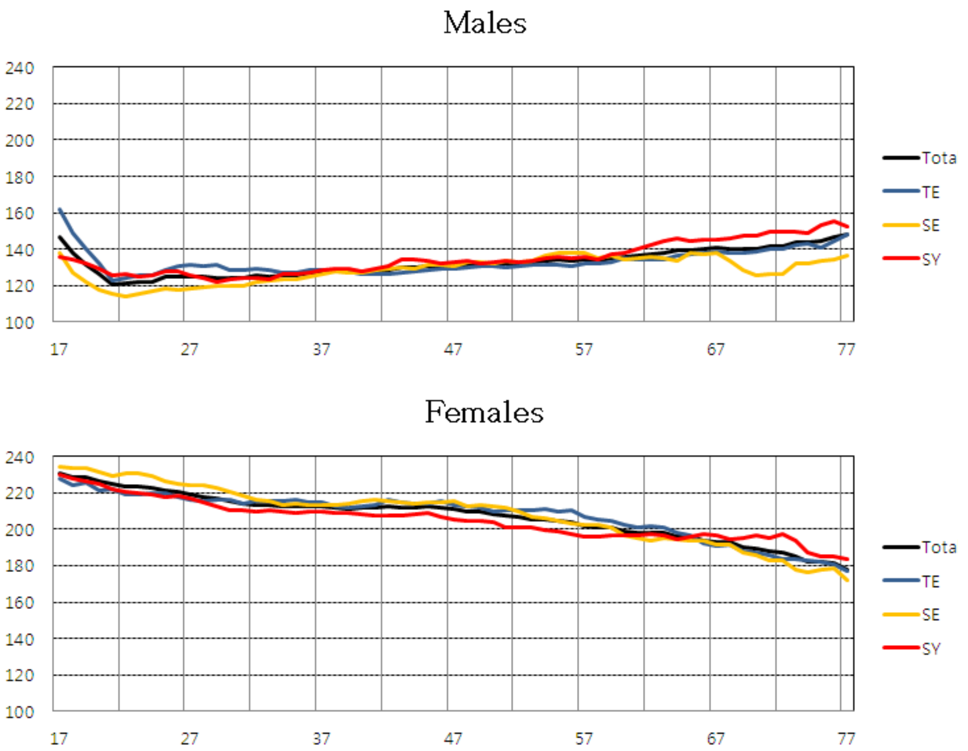

Figure 6. Means of aF0 along the age axis according to constitutional types

Table 1. Distribution of Dataset after Removal of Outliers

	TE	SE	SY	Total
Males	314 (42.84)	178 (24.28%)	241 (32.88%)	733
Females	399 (35.15%)	316 (27.84%)	420 (37.00%)	1135
	713	494	661	1868

outliers are used to obtain significant features for classifying the constitution. The outlier is likely to have the noise and the error in recording.

5. Checking the correlation between features

Usually, the features from the data without the outliers have correlation between the features each other. If there are strong correlation between two or more predictive variables in the dataset, those variables necessarily include the same information about their response. This phenomenon is named multicollinearity. It causes to augment standard deviations of the estimated regression coefficients and often makes confusing and misleading results; the model we build fits well even if all of the individual features are not statistically significant. The best solution to evade multicollinearity is to remove explanatory variables that seem logically inappropriate to the model by checking correlational structure of feature set. A popular way to detect multicollinearity is to compute variation inflation factor (VIF) for each explanatory variable x_j.

$$VIF_j = \frac{1}{1 - R_j^2} \qquad (1)$$

where R_j^2 is the coefficient to determine the model that includes all predictors except the jth predictor. If $VIF_j \geq 10$, a problem of multicollinearity exists. In case of our data, some voice features show a very high VIF, and thus the dimension of the data matrix is reduced before constructing a classification model.

6. Discriminant Analysis

This study aims at determining vocal features of frequency-derivative variables, intensities and the correlation having statistical significance related to Sasang constitutions. As mentioned before, since the number of TY type is not enough, its discrimination is not performed.

The statistical analysis result will reveal how four constitution types are related with the quantitative vocal features. The data are presented as mean and standard deviation. Student's t test is used to test for significance of differences of the gender. As there are three factors (TE, SE and SY types), it is possible to use a one-way ANOVA to test the null hypothesis that there is no difference among the three constitution types.

III. Experimental Results

1. Data Acquisition and Removal of Outliers

We collected 2040 subjects from 24 medical centers, which consisted of 799 and 1241 for males and females, respectively. The numbers of TE, SE, TY and SY types were 749, 526, 68 and 697, respectively. Among them, we analyzed 1972 data after excluding 68 TYs.

We extracted 65 features per subject, but since if only one among features from a subject is outside 3×IQR, the subject is considered as an outlier and removed, the number of subjects was reduced to 1868 as shown in Table 1. The constitutional ratios of males were 31.24%, 33.83% and 34.92% for TE, SE and SY, and those of females were 29.87%, 34.98% and 35.15%, respectively.

2. The Correlation between Features and the Representative Features

We calculated the VIF between features, some of which were strongly correlated, namely a large value.(Table 2(a)) For the accurate discriminant analysis, we had to extract only representative features among them. F0s, the difference between F0s, SF50 and CORR from vowels and a sentence were highly correlated each other, which means the similar frequency property about a subject and thus we chose aF0, iF0-aF0 and uF0-oF0 only from them. F1s, F2s and F2/F1s from vowels were highly correlated, and JITA, JITTs, and PPQ were. xFB60_120/xFB240_480 and xFB60_120/xFB960_1920 from all the vowels and the sentence were highly correlated and there was the close relation between SI50 and SI0 from the sentence, but MFCCs were lowly correlated. As a result, we could obtain the representative features like Table 2(b).

3. Classification of Constitution Types

Due to different feature properties according to the gender and age after age correction, we were able to split the subjects into 6 groups, which were combined with genders and three age intervals.

First, we extracted the significant features in each group statistically. Second, the discriminant function was produced from the significant features. Third, the classification of constitution by using each discriminant function was carried out. The procedure was iterated for 6 groups.

Although Sasang constitution is composed of 4 types,

Table 2. Analytic Results from the Variation Inflation Factors of Voice Features (a) The 1st Stage

Original	VIF	Original	VIF	Original	VIF
aF0	Inf	oJITT	17.113	eF2/eF1	21.524
eF0	53.677	uJITT	5.827	iF2/iF1	9.908
iF0	Inf	aPPQ	15.003	oF2/oF1	28.567
oF0	Inf	ePPQ	16.194	uF2/uF1	30.256
uF0	Inf	iPPQ	20.412	iF0-aF0	Inf
aF1	13.895	oPPQ	12.017	uF0-oF0	Inf
eF1	17.509	uPPQ	10.022	aMFCC4	4.534
iF1	4.916	aFB60_120/aFB240_480	32.361	eMFCC4	7.245
oF1	14.032	eFB60_120/eFB240_480	99.738	iMFCC4	5.551
uF1	15.710	iFB60_120/iFB240_480	860.081	oMFCC4	5.415
aF2	8.230	oFB60_120/oFB240_480	115.773	uMFCC4	4.026
eF2	11.403	uFB60_120/uFB240_480	474.452	CORR	1.386
iF2	6.438	aFB240_480/aFB960_1920	11.485	SF50	150.863
oF2	35.874	eFB240_480/eFB960_1920	9.642	SI50	327.287
uF2	46.520	iFB240_480/iFB960_1920	3.824	SF0	156.383
aJITA	48.625	oFB240_480/oFB960_1920	6.719	SFSTD	2.385
eJITA	64.899	uFB240_480/uFB960_1920	5.160	SI0	326.268
iJITA	72.943	aFB60_120/aFB960_1920	49.499	SISTD	1.557
oJITA	35.517	eFB60_120/eFB960_1920	127.602	SSPD	1.198
uJITA	13.907	iFB60_120/iFB960_1920	905.114	SFB60_120/SFB240_480	98.090
aJITT	32.443	oFB60_120/oFB960_1920	129.791	SFB240_480/SFB960_1920	5.587
eJITT	45.511	uFB60_120/uFB960_1920	506.179	SFB60_120/SFB960_1920	115.018
iJITT	38.472	aF2/aF1	8.039		

(b) The Last Stage

Reduced	VIF	Reduced	VIF	Reduced	VIF
aF0	3.229	iJITA	1.612	eMFCC4	5.404
iF1	1.656	oJITA	1.653	iMFCC4	4.989
oF1	2.794	uJITA	1.445	oMFCC4	5.203
uF1	2.190	aFB60_120/aFB240_480	2.135	uMFCC4	3.769
aF2	1.896	iFB240_480/iFB960_1920	1.740	SI0	1.079
eF2	1.961	oFB240_480/oFB960_1920	2.459	SISTD	1.265
iF2	1.431	aF2/aF1	1.391	SSPD	1.092
oF2	1.567	iF0-aF0	1.240	SFB60_120/SFB240_480	2.752
uF2	1.916	uF0-oF0	1.193	SFB240_480/SFB960_1920	2.448
aJITA	1.627	aMFCC4	4.129		

the number of TY types is too small and thus they cannot be analyzed stochastically. Therefore after we excluded TY for analysis, we classified into 3 types as shown in Table 3, where SC means Sasang constitution. For males in their 20s, the accuracies of constitution types were 40.4, 53.5 and 43.9 % for TE, SE and SY, respectively, where the average was 44.9 %. For males in their 30s and 40s, the accuracies of constitution types were 37.4, 50.8 and 62.5 % for TE, SE and SY, respectively, where the average was 49.6 %. For males of more than 50 years of age, the accuracies of constitution types were 41.3, 47.7 and 41.8 % for TE, SE and SY, respectively, where the average was 42.7 %.

For female in their 20s, the accuracies of constitution types were 40.0, 56.8 and 44.1 % for TE, SE and SY, respectively, where the average was 46.2 %. For female

Table 3. Male and Female Discriminant Accuracy by Constitution Types

(a) The Coefficients of the Linear Discriminant Function and the Discriminant Results for Subjects in Their 20s

Males	SC			Females	SC		
	TE	SE	SY		TE	SE	SY
				aF0	-0.179	0.332	-0.073
uF1	0.121	0.311	0.1	uF1	-0.006	-0.155	0.145
aF2	0.094	0.076	-0.14	aF2	-0.117	0.148	-0.049
				oF2	-0.097	0.037	0.06
				oJITA	0.114	0.028	0.066
uF0-oF0	-0.046	-0.017	0.017	uF0-oF0	0.011	0.028	-0.015
eMFCC4	-0.154	0.296	0.029				
oMFCC4	0.097	-0.143	-0.192	iMFCC4	-0.09	-0.044	0.192
SFB60_120/SFB240_480	0.025	0.393	0.002	SFB240_480/SFB960_1920	-0.278	0.072	-0.097
(constant)	-1.159	-1.341	-1.146	(constant)	-1.248	-1.224	-1.176

SC	Males					Females				
	TE	SE	SY	Total	Accuracy	TE	SE	SY	Total	Accuracy
TE	46	33	35	114	0.404	48	36	36	120	0.4
SE	18	38	15	71	0.535	23	54	18	95	0.568
SY	20	26	36	82	0.439	35	36	56	127	0.441
	84	97	86	267	0.449	106	126	110	342	0.462

(b) The Coefficients of the Linear Discriminant Function and the Discriminant Results for Subjects in Their 30s/40s

Males	SC			Females	SC		
	TE	SE	SY		TE	SE	SY
iF1	-0.015	0.22	-0.216	aF2	-0.235	0.073	-0.115
uF1	0.212	0.148	-0.062	iF2	-0.056	-0.087	-0.2
eF2	-0.05	0.362	-0.182				
iJITA	-0.072	0.096	0.165	oJITA	0.12	0.028	0.095
uJITA	0.184	-0.067	-0.044				
iFB240_480/iFB960_1920	-0.073	-0.18	-0.074	oFB240_480/oFB960_1920	0.012	-0.192	-0.238
aF2/aF1	0.172	0.327	0.08	aF2/aF1	0.187	0.128	0.027
uMFCC4	0.018	0.113	-0.375	eMFCC4	-0.095	0.029	0.246
				SI0	-0.165	0.15	-0.144
				SFB240_480/SFB960_1920	-0.384	-0.092	-0.021
(constant)	-1.263	-1.481	-1.263	(constant)	-1.313	-1.209	-1.234

SC	Males					Females				
	TE	SE	SY	Total	Accuracy	TE	SE	SY	Total	Accuracy
TE	34	29	28	91	0.374	59	41	43	143	0.413
SE	14	32	17	63	0.508	31	59	45	135	0.437
SY	14	16	50	80	0.625	46	46	86	178	0.483
	62	77	95	234	0.496	136	146	174	456	0.447

in their 30s and 40s, the accuracies of constitution types were 41.3, 43.7 and 48.3 % for TE, SE and SY, respectively, where the average was 44.7 %. For females of more than 50 years of age, the accuracies of constitution types were 38.2, 44.2 and 48.7 % for TE, SE and SY, respectively, where the average was 43.3 %.

Both of males and female in all the age intervals were

(c) The Coefficients of the Linear Discriminant Function and the Discriminant Results in Subjects over 50 Years Old

Males	SC			Females	SC		
	TE	SE	SY		TE	SE	SY
				aF0	0.162	-0.056	-0.025
oF2	0.099	0.027	0.048	oF1	0.053	0.054	0.349
				eF2	-0.182	-0.114	0.046
				iF2	-0.133	0.039	-0.25
iFB240_480/iFB960_1920	-0.046	-0.088	-0.155	aFB60_120/aFB240_480	0.026	0.012	0.341
oMFCC4	0.011	-0.028	-0.427				
SSPD	0.023	-0.283	0.094				
SFB240_480/SFB960_1920	-0.067	0.384	0.3				
(constant)	-1.238	-1.242	-1.364	(constant)	-1.189	-1.116	-1.258

SC	Males					Females				
	TE	SE	SY	Total	Accuracy	TE	SE	SY	Total	Accuracy
TE	45	33	31	109	0.413	52	50	34	136	0.382
SE	12	21	11	44	0.477	23	38	25	86	0.442
SY	26	20	33	79	0.418	22	37	56	115	0.487
	83	74	75	232	0.427	97	125	115	337	0.433

the same inclination that the rate of TE was lower than those of SE and SY. The 30s and the 40s had higher rate than the others in males and the 20s had higher rate than the others in females. The average rates of males and females were similar each other.

IV. Discussion and Conclusions

The goal of this study was to utilize scientific methods to detect significant features to help determine the Sasang constitution of a patient. SCM doctors refer to various standards for the judgment of the Sasang constitution, and vocal characteristics are commonly considered as an important factor in determining the Sasang constitution and the health condition. The scientific methods used in this study borrowed datamining techniques such as data preprocessing and significant feature selection.

We tried to find out the classification procedure to distinguish the constitution types objectively and quantitatively by analyzing the characteristics of subject's voice without noise and error.

Originally, we collected the 2669 voice data from multiple clinic centers, 629 of which were excluded manually by operators in the following analysis since some were short, little, noisy superficially or had errors in recording time even if they were acquired strictly with SOP and others were teenagers in order to avoid a break of voice. Since due to the small number, TYs would produce statistically meaningless results, they were also excluded for the analysis. The data were distributed diversely from the 20s to the 70s according to each gender.

Afterwards, various features from the voices were extracted in order to express most of characteristics like frequency derivative features, variation of frequency, frequency band ratios, MFCCs and intensities of vowels and sentences, which a few researchers had agreed with using prior information. To avoid internally existent noise in vocal data, only the data were chosen to have all the features within 3×IQR. Then we tried to minimize correlation between features of each vowel and sentence by using VIFs, and used 7~8 as the threshold for VIF empirically since the VIF of one feature more than 7 or 8 is closely correlated with the others. Only one was selected from the correlated features and analyzed for the classification. We found out the close correlation between the pitches and their differences of vowels and sentences, the formant frequencies and their ratios of each vowel, and the ratios of frequency band for each vowel or each sentence, which were the property of frequency. In addition, the average and the median intensities of the sentence were closely correlated. Through the avoidance of the explanatory variables, the accurate discriminant function was able to be obtained.

We found out that the average pitch and the average first formant of males were lower than those of females, which showed the different property of the frequency according to gender. Especially, the average pitch of males increased with the increment of age, and that of females decreased almost linearly with it. When we observed the pitch by the constitution, the average pitches of males had large differences between the types before the 30s and had the uniform differences and linearly increased from the 30s to the 40s. The magnitude of the types reversed trends from around the 50s onwards. The average pitches of females had the same trend as those of males before the 30s and from the 50s. They had the uniform differences and linearly decreased from the 30s to the 40s.

Therefore, the interval of age was divided into 3 parts. Here in order to normalize each feature in each part, we calculated the normalized feature by subtracting the mean of each feature and then dividing its standard deviation like (2).

$$z = \frac{x - m_x}{\sigma_x} \qquad (2)$$

In fact, since the mean value of a feature varies even in each part as shown in Figure 6, the part needed to be departmentalized for more accurate analysis. However, since we required more than the given number of samples for a relevant statistical analysis, the interval was divided into only three parts as a trade-off.

The discriminant function was made with such normalized features to reduce correlation according to gender and age interval.

The discriminant function was made using vocal features to reduce the correlation according to the gender and age interval as follows. Generally, among the vowels, a is a low tone, e is an intermediate tone, and i, o, and u are relatively high tones. For males in their 20s, TE gave a larger value than SY on the F2 of the low pitched a, and had a low pitch property. SY had a larger value than TE on the uF0-oF0, which indicated that u was larger than o. In particular, SE on the SFB240_480/SFB240_480 showed a low tone as a high value.

For females in their 20s, SY types were larger than TE on the F1 of the high pitched u. In addition, since SY showed a larger F2 and a smaller pitch variation for the letter o, this implies that SY types pronounced this letter higher and more clearly, compared to TE.

For males in their 30s and 40s, SE was larger on the F2 of the intermediate pitched e, and the pitch variations of the high pitched u for SY and SE were smaller, which showed that their voices were clearer.

For females in their 30s and 40s, the strength (sI0) of SE on the sentence was significant, which was an important factor distinguishing SE types. In addition, TE was larger on the pitch variation of the high pitched o, which showed that TE types had rough voices and TE on the oFB240_480/oFB960_1920 showed a low tone as a high value.

For males above the age of 50, SY was faster than other types on the reading speed of a sentence. TE on the iFB240_480/iFB960_1920 showed a low tone as a high value and SE was the intermediate tone on it.

For females above the age of 50, SY was larger on the F1 of the high pitched o, and TE was larger on the F0 of the low pitched a. It is known that factors distinguishing SE were not large in this age interval.

From the preceding comments, we can deduce that the variables contributing to the determination of each constitution type differ according to gender and age.

The MFCCs are often used in voice recognition, but these were a little utilized in the discriminant functions for distinguishing the four constitution types. However, the formant frequencies were utilized in every age interval and gender, and these parameters reflect the structure of the vocal tract through the oral and nasal cavities for articulation—the velum, jaw, tongue, and lips.

Finally, the discriminant functions produced about 45% accuracy to classify the constitution types for every age interval and every gender, and the accuracy was meaningful as the result from only the voice. If there are three types of constitution, the probability of one type being chosen at random is only 1/3.

In this study, we extracted the vocal features from voice data selected with prior information, removed outliers, minimized the correlated features, corrected the features with normalization according to gender and age, and made the discriminant functions that were adaptive to gender and age from the features to show the accuracy. The accuracy was about 45% but if the integrative algorithm to contain the shapes of a face and a body as well as voice is developed for the classification of constitution, the voice can contribute to the improvement of accuracy more.

In future, we need to find out more effective vocal features for the classification, and suggest the discriminant method to distinguish the health condition as well as the constitution. Especially, the voice analysis method will play a critical and essential role, leading to the system for u-Healthcare and a smart phone.

V. References

1. Kim JY, Pham DD. Sasang constitutional medicine as a holistic tailored medicine. eCAM. 2009;6(sup. 1):11-20.
2. Han C, Park SH, Lee SJ, Kim M-G, Wedding D, Kwon Y-K. Psychological profile of Sasang typology: a systematic review. eCAM. 2009;6(sup. 1):21-30.
3. Lee J. Dongeuisusebowon-chogo. Seoul:Cheongdam (Translated by Kim D-R). 1999.(Korean)
4. Cho DU. Sasang Constitution Classification by Speech Signal Processing. J of Korea Information and Communications Society. 2006;31(5C):548-555.(Korean)
5. Moon SJ, Tak JH, Hwang HJ. A Phonetic study of 'Sasang Constitution'. Malsori. 2005;55:1-14.(Korean)
6. Kwon CH, Kim JY, Kim KH, Han SM. A Study on Correlation between Sasang Constitution and Speech Features. J of traditional Korean medicine. 2009;19(2): 219-228.(Korean)
7. Park S-J, Kim D-R. A study on the correlation between Sasang Constitution and sound characteristics used harmonics and formant bandwidth. J Sasang Constitut Med. 2004;16(1):61-73.(Korean)
8. Kim S-H, Han D-Y, Youn J-Y, Kim D-R, Jeon J-W. A study on the characteristics of the Korea adult women sound as by Sasang Constitution analysed with PSSC-2004. J Sasang Constitut Med. 2005;17(1): 84-102.(Korean)
9. Kim H, Yang S-M, Shim G-H, Yoo J-S, Kim D-R.

Sound characteristics of Sasang Constitutional type using PSSC-2004 in the Korean children. J Sasang Constitut Med. 2006;18(2):55-67.(Korean)

10. Choi J-W, Song H-S, Han D-Y, Cho S-E. A study on the characteristics of the Korean adult male sound according to Sasang Constitution using PSCC with a sentence. J Sasang Constitut Med. 2006;18(3):64-74. (Korean)
11. Goldwave Inc. Goldwave v5.58. 2010 Aug. Available from:URL:http://www.goldwave.com
12. Woodland P. Version 3.4.1 of HTK. Hidden Markov Model Toolkit (HTK). 2009. Available from:URL:http://htk.eng.cam.ac.uk
13. Boersma P, Weenink D. Version 5.2.11 of Praat. Praat: doing phonetics by computer. 2011 Jan. Available from:URL:http://www.fon.hum.uva.nl/praat/
14. Deller JR, Hansen JHL, Proakis JG. Discrete-Time Processing of Speech Signals, New York: Wiley-Interscience. 2000:104-115.
15. Rahim MG, Juang BH. Signal bias removal by maximum likelihood estimation for robust telephone speech recognition. IEEE Trans. Speech & Audio Processing. 1996;4(1):19-30.
16. Epina GmbH. Pearson's Correlation Coefficient. Virtual Institute of Applied Science. 2006 Jan. Available from:URL:http://www.vias.org/tmdatanaleng/cc_corr_coeff.html
17. Upton G, Cook I. Understanding Statistics. Oxford:Oxford University Press. 1996:55.

Appendix: Brief Description of Abbreviations

Brief description of vowels	x=a, e, i, o, u
Average pitch frequency	xF0
1st term of formant frequencies	xF1
2nd term of formant frequencies	xF2
F2/F1 of a vowel	xF2/xF1
Pitch Pertubation Quotient: wider average of jitter	xPPQ
Average Jitter	xJITA
Jitter	xJITT
Ratios of voice energies over fixed frequency bands between 60~120Hz and 240~480Hz	xFB60_120/xFB240_480
Ratios of voice energies over fixed frequency bands between 240~480Hz and 960~1920Hz	xFB240_480/xFB960_1920
Ratios of voice energies over fixed frequency bands between 60~120Hz and 960~1920Hz	xFB60_120/xFB960_1920
Difference between 'i' and 'a' pitch frequencies	iF0-aF0
Difference between 'u' and 'i' pitch frequencies	uF0-iF0

Brief description of a sentence	S
50th percentile of pitch frequency	SF50
50th percentile of intensity	SI50
Average pitch frequency	SF0
Variation of pitch frequency	SFSTD
Average intensity	SI0
Variation of intensity	SISTD
Pearson correlation coefficient between SF0 and SI0	CORR
Time to read one sentence	SSPD