

사상 의학을 위한 음성 분석 시스템 설계 및 구현

허병도*, 손기락*, 전종원**

*한국의국어대학교 컴퓨터 및 정보통신공학과

**상지영서대학 컴퓨터 공학과

e-mail:byungdo@korea.com

Design and Implementation of Phonetic System for Sasang Constitution

Byung-Do Heo*, Ki-Rak, Sohn*, Jong-won Jeon**

*Dept of Computer Science & Telecommunication Engineering,
Han-Kook University for Foreign Studies

**Dept of Computer Science, Sangji Youngse College

요 약

우리나라에는 의사에서 그 유래를 찾아보기 어려운 독특한 체질의학이론인 '사상의학'이 있다. 이 이론을 제시한 학자는 이조말엽의 함경도 지방의 이재마라는 의학자에 의하여 창안된 이론이다. 본 논문의 연구는 음성을 이용하여 음성이 담고 있는 언어적 특징이나 음성요소들을 이용하여 사상 의학과 연관성을 찾는 데 목적을 두고 시작되었으며, 기존의 음성분석기의 요소들과 본 연구에서 개발된 요소를 병합하여 음성분석 모듈을 완성하였다.

1 서론

음성분석기는 인간의 청력보다 더 뛰어난 능력을 가지고 있고, 언어학이나 음성학은 물론 양방 ENT (이비인후과)에서도 많이 이용하고 있는 관계로 상당수 보급되어 있기 때문에 음성특성을 연구하기에는 환경이 좋은 편입니다. 단지 사상 체질 분야에 이용하지 않았기 때문에 생소하게 여겨졌을 뿐이다. 더구나 국립과학수사연구소 음성분석실에서는 오래 전부터 음성특성에 대한 연구가 상당수준 되어 있었고, 성문은 범죄수사에 있어서 증거로 채택되기도 한다.

음성에는 여러 가지 정보가 들어 있다. 마치 우리가 음성만 듣고 어떤 특정인을 떠올릴 수 있는 것처럼 5초 내외의 짧은 문장 속에서도 수십 개의 정보를 찾아낼 수 있습니다.

예를 들면 음성의 높낮이와 비유하면 첼로의 울림통이 바이올린에 비해 크기 때문에 소리가 웅장하고 (주파수가 낮다), 단소의 길이와 직경이 대금보다 짧

고 가늘기 때문에 애절하다는 것을 이해할 수 있을 것이다.

따라서 소리의 높낮이를 통해 대체적인 신체특성이 드러납니다. 마치 우리가 눈으로 보고 키가 크거나 작은지를 아는 것처럼 소리를 통해서도 어느 정도는 유추할 수 있다.

이런 음성 특성을 음성분석기로 수치화할 수 있고, 그 밖의 음성요소를 각 체질별로 귀납적으로 비교하면 일정한 체질별 음성특성을 구분해 낼 수 있다. 음성분석기의 장점은 소리를 귀로 해석하는 것이 아니라 시각적으로 이해할 수 있다는 것과 녹음된 자료를 통해 반연구적으로 불러낼 수 있다는 점이라 할 수 있을 것이다.

2. 음성분석기 개발 배경

일반적으로 음성(speech) 또는 언어(language)를 연구하는 음성과학자들이 연구의 해답을 찾을 때 주로 두 가지 방법에 의존하게 된다. 첫 번째 방법은

자료를 관찰하고 기록하는 기술적(descriptive) 연구이고, 두 번째 방법은 기술된 내용을 검증하는 실험적(experimental)연구이다. 후자의 경우에 연구자는 실험기기를 이용하여 언어음을 분석해야 하는데, 본 논문에서는 기존의 음성 분석용 기기들의 특징을 기반으로 하여 전자와 후자의 경우를 충족할 수 있는 사상체질 진단 진단용 음성분석기를 개발에 대한 필요요 개발을 하게 되었다.

다음은 기존에 대표적으로 사용되는 음성분석기에 대한 간략한 설명이다.

2.1 컴퓨터 음성분석기(CSL)

CSL은 컴퓨터를 기반으로 가장 종합적인 음성분석기로서 손쉽게 목소리를 분석하고 편집 및 분석된 결과를 그래프와 수치로 볼 수 있는 음성분석 시스템이다. 이 기기는 목소리를 구성하는 모든 요소에 대한 정확한 분석과 상태를 파악할 수 있게 제작되어 목소리의 변화에 대한 정확한 진단과 치료 후 환자상태에 대한 객관적인 측정 및 질환 예방, 임상연구와 교육 프로그램에 유용한 의료장비이다.

사용가능한 영역은 다음과 같다.

- MDVP(multidimensional voice program) : 종합 음성분석검사
- Spectrogram : 스펙트로그램
- EGG(electroglottography) : 전기성문파형검사
- VRP(voice range profile) : 성역검사
- DAF(delay auditory feedback) : 지연청각피드백
- ASL : 음성합성기

2.2 Dr. Speech (version 4)

컴퓨터 음성분석기로서 다른 부수적인 장치가 필요 없이 일반 컴퓨터에서 실제 사용할 수 있으며 노트북에 넣어 가지고 다닐 수 있는 편리하고 간편한 프로그램 중의 하나이다. Tiger DRS, Inc. 제품으로 음성분석기로서 여러 가지 음성에 대한 정보를 알 수 있으며 음성치료나 언어치료에서도 시각적인 자료를 제시해 주고 환자들이 보다 쉽게 이해할 수 있도록 도와주는 프로그램이다. 내용으로는 언어훈련 프로그램, 음성합성과 치료, 언어·음성치료, 음성분석, 음성평가 등으로 나눌 수 있다.

- Speech training program
- Voice synthesis & therapy
- Speech & voice therapy

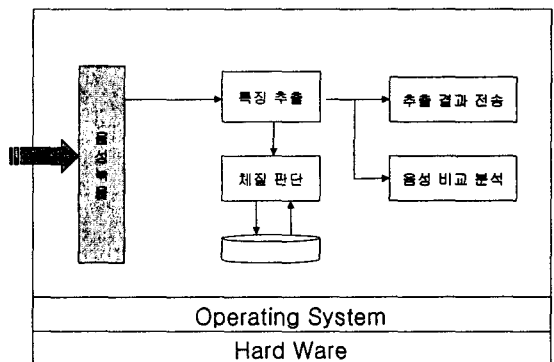
2.3 Phonotic Simulation (version 1.0)

언어치료적인 면에서 필요한 프로그램으로서 발성의 음성학적인 동영상을 보여주는 측면에서 시각적인 효과를 높일 수 있는 프로그램이다. 운동역학적인 측면, 공기역학적인 측면, 음향학적인 측면으로 세 가지의 특성을 동영상으로 제시해 준다.

3. 시스템 구현 내용

3.1 시스템 구성

본 논문에서 제안한 시스템의 구성은 [그림 1]의 음성녹음, 특징추출, 체질판단, 추출 결과 전송 그리고 음성비교분석 부분 중 체질판단 부분을 제외한 나머지 부분으로 구성 된다. 우선 음성 녹음 후 특징을 추출하는 부분에서는 기존에 사용되고 있는 음성분석기들의 알고리즘들 중 본 연구와 연관 있는 알고리즘들로 특징을 추출한다. 추출된 결과는 텍스트 파일로 저장되며 저장된 파일의 데이터는 “, ”로 구분되어 있어 CSV 파일로 변환하여 엑셀에서 사용이 가능하도록 구성되었고, 음성비교 분석 부분은 여러 번에 걸쳐 녹음된 음성간의 비교 및 분석을 할 수 있도록 설계되었다.



[그림 1] 사상체질 진단용 음성분석기 시스템구성도

3.2. 구현 요소

3.2.1 APQ : Pitch period 내의 음성 강도 변이의 정도에 대해 상대적으로 평가하는 변인이다. 대체로 기식성 음성(breathy voice)이나 애성(hoarse voice)이 있는 경우 그 수치가 증가하며 그 수치가 증가함은 성대의 주기적 진동이 어렵다는 것을 뜻한다.

$$APQ = \frac{\frac{1}{N-4} \sum_{i=1}^{N-4} \left| \frac{1}{5} \sum_{r=0}^4 A^{(i+r)} - A^{(i+2)} \right|}{\frac{1}{N} \sum_{i=1}^N A^{(i)}} \quad (\text{수식 1})$$

3.2.2 PPQ : Pitch period 5개 길이 내에서의 음도 변이 정도에 대해 상대적으로 평가인이다. period의 수가 늘어남으로써, 바로 인접해 있는 period간의 변이 정도에 대해서 설명할 때는 효율성이 떨어지거나, 짧은 기간동안의 음성의 변이 상태에 대해서 설명할 때는 용이하게 사용될 수 있다.

$$PPQ = \frac{\frac{1}{N-4} \sum_{i=1}^{N-4} \left| \frac{1}{5} \sum_{r=0}^4 A^{(i+r)} - T_0^{(i+2)} \right|}{\frac{1}{N} \sum_{i=1}^N T_0^{(i)}} \quad (\text{수식 2})$$

3.2.3 Shim : pitch period 와 period 간 음성 강도 변이의 규칙성, 불규칙성을 상대적으로 평가하는 변인이다.

$$Shim = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} \left| A^{(i)} - T_0^{(i+1)} \right|}{\frac{1}{N} \sum_{i=1}^N A^{(i)}} \quad (\text{수식 3})$$

3.2.4 Pitch : 1초에 평균적으로 성대가 진동하는 횟수를 보여준다. 대개 정상 성인 남성은 100 ~ 130 Hz, 여성은 190 ~ 230 Hz에 분포하며, 아들은 대개 여성과 통계적으로 유의한 차이가 없는 분포를 보이는데 대개는 좀더 높을 때가 자주 관찰된다.

3.2.5 Jita : Pitch period와 period간 음도 변이의 규칙성·불규칙성을 절대적으로 평가하는 변인이다. 이는 절대적인 평가이므로 시간 단위로 계산되며 기본 주파수로 좌우되는 경우가 많아 남·여 차이가 나타날 수 있는데 대개 주파수가 높을수록 Jita의 수치가 낮아진다.

$$Jita = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| T_0^{(i)} - T_0^{(i+1)} \right| \quad (\text{수식 4})$$

3.2.6 CF : 중심주파수를 나타내며 일정 주파수 대역의 중간 주파수를 의미하거나, 옥타브 대역

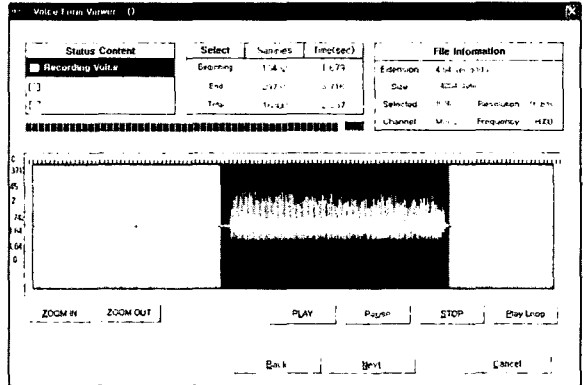
해석시 기준이 되는 주파수이며 중심 주파수를 각 대역별로 나누어 특정 주파수 대역의 상대적인 비율을 볼 수 있다.

3.2.7 ZOD : Time Domain의 Peak에 대한 Deviation을 볼 수 있으며, PPQ와 유사하나 ZOD는 절대 값에 대한 값으로 나타낸다.

각 알고리즘을 텍스트 파일로 저장되며 특정 알고리즘에 대한 세분화된 데이터를 구할 수 있다.

3.3 구간 선택

아래 [그림 2]는 녹음된 음성의 파형, 파일 정보, 분석하기 위해 선택한 음성 구간 정보를 보여주는 화면으로, 음성분석은 실시간으로 녹음된 음성에 대한 분석과 저장되어 있는 음성 파일에 대한 일괄처리가 가능하도록 구현하였다.



[그림 2] 녹음된 음성의 분석 범위 설정 화면

3.4 결과 출력

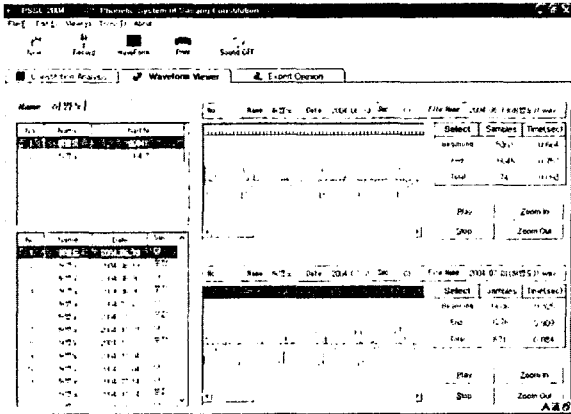
아래 [그림 3]은 선택된 범위 내에 추출된 음성 데이터 결과를 보여주는 화면으로 추출된 데이터는 5~6의 파일로 출력되며 각 파일들은 각 데이터들간의 비교를 쉽게 할 수 있도록 하기 위해 주파수 관련, 타임 도메인 관련, 옥타브 관련, 에너지 관련 등의 내용으로 구현하였다.

Pkts	C	D	E	F	G	H	I	J	K	L	M	N
1	113.281	570.312	400.993	2863.281	343.75	1242.187	26908920	18935362	0.718502	3148984	1120448	0.357183
3	128.906	773.437	769.531	3269.531	671.874	1078.124	1.65E+08	81491681	0.49405	13332183	7156738	0.536802
4	109.375	644.531	578.124	3339.843	542.968	1128.906	97823123	40529202	0.415403	10880622	9728830	0.526608
5	164.062	820.312	664.062	3003.124	698.093	1152.343	17792333	9738205	0.547996	2067250	1460769	0.65642
7	97.656	843.749	785.156	2863.281	667.968	1179.687	53808492	15889043	0.297147	5008482	3076296	0.614263
10	105.562	789.062	896.312	2874.999	398.437	1181.406	15306988	10068184	0.687751	2188994	1304411	0.801417
8	97.656	714.843	895.312	2953.124	675.781	1082.031	75408561	35654857	0.472296	4464653	2681064	0.593395
9	93.75	832.031	753.906	2937.489	570.312	1214.843	43202396	21376677	0.493646	5286002	2933219	0.554803
10	117.187	792.968	757.812	2531.249	664.062	1178.687	70580863	32480728	0.457828	4427874	2046567	0.596351
11	113.281	684.062	636.718	2984.531	696.093	1148.437	31757827	11202362	0.952747	1529215	8454014	0.556302
12	105.468	718.749	703.124	2730.468	613.281	1180.156	29785892	17810484	0.539796	2010791	1059678	0.526998
13	113.281	843.749	785.156	2839.843	664.062	1121.993	48168133	27429245	0.568994	3533380	1896226	0.664079
14	93.75	785.156	753.906	2932.812	579.124	1140.624	42010988	20257802	0.629505	4560348	2341337	0.512401
15	105.468	583.749	454.843	3170.687	312.5	1234.374	27390985	11699919	0.427134	2813239	1847367	0.585577
16	70.312	821.093	809.374	3136.718	589.843	1124.999	20480139	11701943	0.57138	2003705	1019893	0.508906
17	93.75	277.343	273.437	3035.156	281.718	1205.662	11833624	8331154	0.703179	4009557	362871	0.534122
18	101.562	710.937	691.406	2777.343	585.937	1121.093	30484327	17008802	0.557507	3461237	1876055	0.485233
19	160.825	433.563	402.343	2789.531	292.968	1230.488	87342547	31471640	0.358708	5509303	3023689	0.548933
20	109.375	781.718	757.812	2940.624	652.343	1199.374	46824811	26252815	0.624713	1659910	8339564	0.502031
21	113.281	675.781	667.968	2906.249	570.312	1132.812	49425688	32859283	0.844822	3925096	2040297	0.570752
22	74.218	852.343	636.718	3015.624	503.906	1181.406	14232316	6259198	0.437654	1732688	1006719	0.577353
23	171.875	879.687	864.062	2314.062	640.624	1095.937	1138408	37467581	0.331027	11400989	8302018	0.554109
24	117.187	589.843	582.031	3003.906	365.468	1070.312	20893267	12298356	0.594317	1278414	707088	0.553985
25	117.187	681.499	593.749	2946.312	488.749	1144.531	53889747	27680551	0.512279	4492253	2164307	0.482232
26	128.906	730.468	707.031	2950.156	639.906	1195.312	1326408	89598137	0.468989	1187392	6864373	0.594811
27	156.25	636.718	628.906	2964.843	617.187	1105.468	81604504	25938811	0.317873	7365731	2763385	0.378241
28	117.187	688.218	699.268	3089.843	656.249	1175.781	60504689	59163143	0.484575	4689519	2330231	0.498923
29	93.75	558.563	542.968	2955.781	382.812	1132.812	20522572	15376515	0.579752	1843616	9859829	0.534809
30	132.812	789.062	773.437	3097.656	656.249	1183.583	80480885	37641605	0.467709	8474745	3881793	0.388584
31	132.812	675.781	667.968	2939.843	604.624	1195.312	42554268	21460588	0.504311	4425244	2046972	0.543918

[그림 3] 음성데이터 결과 출력 화면

3.5 음성 비교 분석

아래 [그림 4]는 동일한 또는 다른 사람의 음성을 검색하여 시간의 흐름에 대해 음성의 변화에 대한 차이나, 몸 상태, 기후에 관하여 음성에 대한 변화 유무에 대해 어떤 변화를 보이는지 비교·분석해 볼 수 있도록 구현 되었고, 저장된 음성파일을 비교·분석 할 경우 파일 리스트에서 원하는 항목을 선택 하면 선택된 음성파일에 대한 파형 및 음성 정보를 볼 수 있도록 구현하였다.



[그림 4] 음성 데이터 비교 분석 화면

4. 실험 환경

CPU	인텔 펜티엄 4 1.6Ghz
RAM	512MB
HDD	30GB
Sound Card	Creative JCHyun SB Live DE 5.1 Lite
MIC	AUDIO-TECHNICA ATM-75 헤드셋마이크
OS	Windows XP

5. 결론 및 향후 연구과제

본 논문에서는 사상체질 진단용으로 개발된 음성 분석기를 설계 및 구현 하였다. 본 음성분석기는 기존의 QSCC의 설문 방식에 비해 시간절약 및 정확도면에서 탁월한 성능을 보이고 있다고 본다. 진단 방법도 몇 가지의 단어 및 문장을 읽는 것으로 분석이 끝난다. 향후 음성특징과 체질과의 상관성 있는 요소를 추가하여 정확도 및 신뢰도를 높이도록 해야겠다.

참고문헌

- [1] 한진수, "음성신호처리", 2000
- [2] 고도홍·정옥란 외 공편 "음성 및 언어 분석기기 활용법", 2001
- [3] 문영일, "음성과 언어", 1991
- [4] 안희영, "음성검사법 기초편", 1996
- [5] 안희영, "음성검사법 임상편", 1996
- [6] 안희영, "음성검사법 일본음성언어학회 편", 1992