

5聲의 分析과 分類

조병서*, 이용동, 허 응, 박영배**

*동원대학 정보통신과

E-mail : bscho@tongwon.ac.kr

명지대학교 대학원 전자공학과

**경희대학교 한의과대학

A study on the 5-Tone Analysis and Classification

B. S. Cho*, Y. D. Lee, J. K. Kim, W. Hur, Y. B. Pak

~Tongwon Colledge Dept. of Information & Communication Eng.

Dept. of Electronics Eng. Myongji Univ.

38-2 Nam-Dong, Yong-In city, Kyunggi-Do, Korea

School of oriental Medicine. Kyung Hee Univ.

요 약

The human speech sounds are use to diagnosis in oriental medicine with 'O-sung' theory. In general, human voice are sound waves which generated by phonation. Two major parts of phonation are vocal cords and vocal tract. The uniqueness of individual vocal sound depend on structure and usage of their vocal cords and tract.

In the oriental medicine, "O-sung (5-tones)" has been used to classify constitution of human body. In order to characterize the "O-sung", their frequency characteristics are investigated, and a principal frequency component is extracted. Then, the principal component is applied to classify sounds into "O-sung".

1. 서 론

음양오행론은 우리 나라 고대의 철학이론으로서 물질의 속성 및 그 상호 관계를 인식하는 방법론이었다. 의학중의 5행학설은 고대의 소박한 유물론과 자연발생변증법의 사상방법을 의학실천과 서로 결부한 산물이다. <5>란 木, 火, 土, 金, 水 5가지의 사물을 말하고 <행>이란 운동을 말한다. 이 학설은 5행의 속성으로 인체의 장부기관을 연계하며 따라서 5장을 중심으로 <相生>, <相克>, <相乘>, <相侮>의 이론을 운용하여 일부 생리적 현상과 병리적 변화를 설명한 것인데 임상경험을 총화하는데 쓰인다.

5음은 궁(宮)·상(商)·각(角)·치(徵)·우(羽)로 5음의 변화는 곧 물체의 질과 기능의 변동상태가 다르기 때문이며, 물질은 자발적이거나 외부의 힘에 의하여 진동이 발생하고, 진동할 때에 진폭의 대소와 진동수가 다르므로 인하여 강약이 다른 음조가 발생하니, 진동수가 높으면 음조가 높고 반대가 되면 음조가 낮으며, 진폭이 크면 소리의 울림이 강하고 반대가 되면 울림이 약하다. 5장은 생

리·병리의 상황에 따라 질과 기능이 다르게 변동하고, 오장의 자발적인 변동량의 차이에 따라 기가 유동하는 양에 영향을 주어 형성된 변동상황이 발음기관에 충격을 주는 동력에도 다르게 작용하므로 음성도 다섯가지로 변화하여 다르게 발현되는 것이다. 그러므로, 5성의 변화를 들어서 장부와의 상관관계를 변별할 수 있다.

본 연구에서는 5성의 변화와 장부와의 상관관계가 어떻게 존재하는지와 5성의 음향음성의 통계적 방법을 이용하여 각각의 특성 파라미터를 구하고 이 파라미터를 이용하여 분류를 시도하고자 한다.

2. 5성의 분류 방법

표 1.은 오장에 대응한 음성에 관한 정보를 카테고리별로 정리한 것이다. 간(肝)의 음은 角(牙音)·음조는 쌍조(雙調)·술(G)이고, 심(心)의 음은 徵(舌音)·음조는 황종(黃鐘)·라(A)이고, 비(脾)의 음은 宮(喉音)·음조는 일월(一越)·레(D)이고, 폐(肺)의 음은 商(齒音)·음조는 평조(平調)·미(E)이고, 신(腎)의 음은 羽(脣音)·음조는 맹사(魴沙)·시(B)이다.

| | | | | | |
|-----|----|----|----|----|----|
| 五 音 | 角 | 徵 | 宮 | 商 | 羽 |
| 五 處 | 牙音 | 舌音 | 喉音 | 齒音 | 脣音 |
| 五音調 | 雙調 | 黃鐘 | 一越 | 平調 | 魴沙 |
| 五 行 | 木 | 火 | 土 | 金 | 水 |
| 五 臟 | 肝 | 心 | 脾 | 肺 | 腎 |

표 1. 음성패턴에 관한 원전의 카테고리표

음령오행은 소리로서 음양오행을 변별하는 법이다. 각각의 소리를 발음기관에 따라서 각 오행을 배속시켜 오음으로 나타낸 분류이다. 이들의 소리를 어금니소리·혀소리·목구멍소리·잇소리·입

술소리의 오음으로 분류하여 소리오행으로 나타낸 표가 표 2이다[1].

음성의 음향적 특성은 완전히 주기적인 파형이 아닌 비정상파로 구성되어있으므로, 10~30ms의 단시간 해석에 의한 정상파로 가정하여 해석하고 있다. 그러나, 한의학에서 사용하는 5성은 단순히 하나의 음성을 대상으로 5행분류를 시도하지 않고, 대화중에 발생된 소리를 한의사의 직관적이고 주관적인 사고로 5행분류 및 판단을 하고있다. 이러한 5성의 객관적 분석을 위하여, 음성의 통계적 특성을 분석하였다. 실험을 위해, 본 논문에서 사용한 데이터는 훈련받지 않은 6명(남자 3명, 여자 3명)이 10초 동안 의미없이 발생한 5개의 문장을 자연스럽게 발생한 것을 채집하였다.

| 오행 | 자 음 | 발음기관 | 오음 | 비 고 | 훈민정음 소리오행 |
|----|---------|-----------------|----|------------|-----------|
| 木 | 가,카 | 어금니소리 아음(牙音) | 角 | ㄱ, ㅋ | 가,카 |
| 火 | 나,다,라,타 | 혀소리 설음(舌音) | 徵 | ㄴ, ㄷ, ㄹ, ㄷ | 나,다,라,타 |
| 土 | 아,하 | 목구멍소리 후음(喉音) | 宮 | ㅇ, ㅎ | ㅁ, ㅂ, ㅍ |
| 金 | 사,자,차 | 잇소리 치음(齒音) | 商 | ㅅ, ㅈ, ㅊ | ㅅ, ㅈ, ㅊ |
| 水 | 마,바,파 | 입술소리 순음(脣音) | 羽 | ㅁ, ㅂ, ㅍ | ㅇ, ㅎ |

표 2. 소리 오행 도표

3. 음성의 통계적 특성

음성은 여러 발성기관에 의해 발생되는데, 인간의 귀에 들릴 때, 의미와 음색의 정보 즉, 음운과 음질을 판정지워주는 여러 특징적인 성질이 있다.

일반적으로, 음파의 진폭은 음의 세기에, 진동수는 음의 높이에 그리고 진동파형 포락선, 진동 주파수 스펙트럼등은 음색과 관련이 있다. 그러므로, 음성의 장시간 통계량으로서 취급 가능한 2가지

5聲의 分析과 分類

방법은 음성의 순시치의 통계적 분포이고 또 하나는 음성 에너지의 주파수 스펙트럼의 통계값이다.

비교적 장시간 동안 발생된 음성의 순시 진폭은 정상과정으로 가정할 수 있어, 음성파의 순시 진폭을 확률변수로 하는 분포함수를 나타낼 수 있다.[2]

음성의 이러한 확률적 통계적 성질은, 음성파에 대한 여러 가지 정보 이론적인 취급의 한 기초를 제시하는 것이 된다.[3]

또한, 장시간 평균 스펙트럼에서, 음성의 물리적 성질은 시간 파형으로도 알 수 있지만, 음색 등에 대한 경우에는, 주파수 영역에서의 구조 즉, 음의 스펙트럼을 고려하는 것이 바람직하다.

장시간 평균 스펙트럼(long-term averaged spectrum) 측정시에는, 일반적으로 분석정도가 $\frac{1}{8}$ Oct. 대역폭의 필터를 사용하며, 각 주파수 대역내에서의 확률분포로부터 계산된 장시간 실효치를 각각의 대역폭에 대해 계산하여, 측정시에 이용한다. 그리고, 장시간 평균 스펙트럼을 구할 때, 발생자가 발생한 문장중 어느 구간 이상 동안의 목음 구간은 제외한다.[4]

일반적으로, 남성과 여성 사이에는 성대의 기본 주파수의 차이가 있기 때문에, 성별을 구분하여, 결과를 구했다.

또한, 파워 포락 파형의 스펙트럼을 이용하여 그 분산을 조사하였다. 이 방법으로 연속 발생한 음성 중 음절의 길이를 비교하였다.

4. 실험 방법

실험의 방법을 위한 흐름도는 그림 1과 같다.

실험을 위해 음성 데이터의 초·중성을 분리한다. 초성은 자음으로 구성되므로 자음의 판정으로 소리오행 배속을 시도한다. 모음부인 중성을 처리하기 위하여 기본 주파수(F_0)와 포만트의 추출을 시도하였다. 유성음인 모음은 기본 주파수의 하모닉스들로 구성된다. 이들의 특징은 성도의 공진 특성의 영향을 받아 공진 주파수의 봉우리 부분이 관

찰되는데, 이들 봉우리들을 낮은 주파수로부터 제 1 포만트, 제 2포만트등으로 나타낸다.

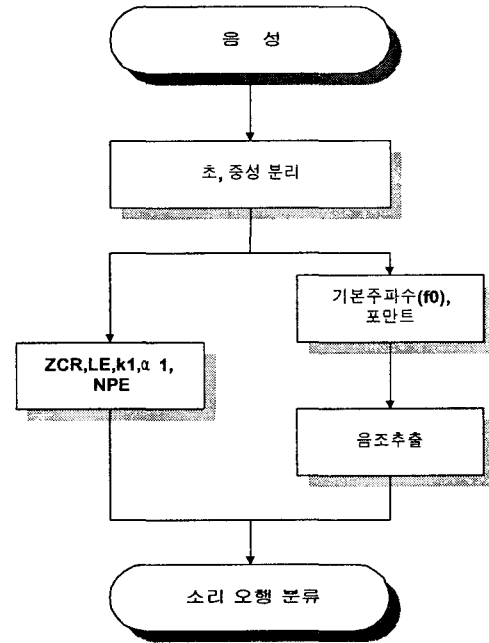


그림 1 오행 분류 흐름도

포만트 주파수 근처에서 에너지의 분포가 높고, 그 밖에 기본 주파수의 정수배에 해당되는 세부구조가 하모닉스로서 나타난다. 기본 주파수는 소리 높이의 지표로 사용되며 강세나 억양의 표현 매체로서 정보전달에 중요한 의미를 갖는다. 또한, 포만트(formant)는 성도의 모양에 의해 기본적으로 결정되는 것으로 조음점의 위치 변동에 따라 시간적으로 변화하게 된다. 따라서, 포만트는 음성의 특성을 규명하는데 중요 단서로서 유용하므로 초·중성을 분리시 판정 데이터로 필요하다.[5]

또한, 음성의 통계적 특성을 이용하여 순시 진폭 분포, 장시간 평균 스펙트럼 그리고 파워 포락 파형의 스펙트럼을 이용하여 전체 발생한 음성의 통계적 처리를 행한 후, 그 특징을 비교하였다.[6][7][8]

이러한 음향음성 특성의 경향(acoustic trends)을 5성에 판별에 파라미터로 사용하므로, 이 파라미터의 통계적 특성으로 특징지을 수 있고, 그 결과 카테고리화 할 수 있는 분류를 하였다.

그리고, 기본 주파수의 값들을 구한 결과, 남자는 111.7~146.7Hz, 여성은 202.3~242.3Hz의 분포로 나타났다. 그리고, 문장내의 유·무성음의 시간 비율로 유·무성음을 판별한 후, 구한 결과 20.5%로 나타났다. 이 결과, 유·무성음의 시간 비율이 약4:1로 지속됨을 알 수 있다.

또한, 청취시에 유용한 정보로 사용하는 음색은 제 4 포먼트에 나타나므로 이들의 분포특성과 레벨을 이용하여 파라미터로 이용하였다.

5. 결론

본 논문에서는, 한의학에서 5성의 객관적 분류를 위한 성의 분석을 주파수 영역과 시간 영역에서 음성의 통계적 특성으로 음색과 관련이 있는 장시간 평균 스펙트럼 그리고 기본 주파수를 추출하였고, 이것과 포먼트 주파수에 의한 음색정보를 파라미터로 사용하였고, 시간영역의 통계적 특성으로 순시 진폭 분포와 유·무성음의 시간 비율등을 구했다.

음향음성학적 관점에서 볼 때, 사용된 데이터의 양이 충분치 않으나, 이러한 파라미터들을 이용하여 5성의 객관적 분류를 시도한다면 타당성을 갖추는데 충분하다고 생각된다.

계속적인 연구의 수행을 위하여, 한국어의 통계적 특성을 음성 분석에 이용하고 더불어 언어의 구조와 문법의 차이등으로 인하여 짧은 구간내의 변화에도 언어의 따른 많은 차이가 존재한다. 이와 관련한 문제들의 해결을 위해서는 장시간분석과 더불어 효율 좋은 음성의 분석방법을 사용하여 세밀한 분류 작업을 수행해야 할 것이다.

참 고 문 헌

- [1]이봉교, 한방진단학, 정보사, pp.110~124, 1997
- [2]L.R.Rabiner and R.W.Schafer, Digital Processing of Speech Signals, Englewood Cliffs Prentice Hall, 1978
- [3]H.K Dunn and S.D White, "Statistical Measurements on Conversational Speech", J.Acoust. Soc. Amer., vol.11, pp.278~288, 1940
- [4]M.R.Sambur and L.R.Rabiner, "An Algorithm for Determining the End points of Isolated Utterances", Bell. Syst. Tech. J., vol.52, pp.297~315, Feb., 1975
- [5]Daniel J. France "Acoustical Properties of Speech as Indicators of Depression and Suicidal Risk", IEEE Trans. Biomedical Eng., Vol. 47, No.7, pp.829 ~ 837, July 2000
- [6]人和, "音聲瞬時 레벨 統計器", 通研月報 第7卷 2號, pp.81 ~ 86, 1986
- [7]M.D.Paez and T.H.Glisson, "Minimum Mean Squared Error Quantization in Speech", IEEE Trans. Commun. vol. Com-20, pp.225~230, April, 1972
- [8]C.K.Un and S.C. Yang, "A Pitch Extraction Algorithm Based on LPC Inverse Filtering and AMDF", IEEE Trans. Acoustic Speech Signal Processing, vol. ASSP-25, Dec., 1977