

한국어 단모음 [이, 에, 아, 오, 우, 으]에 대한 성도 3차원 모델링 및 모음 판별

성철재*, 박종원+, 김귀룡†

* 충남대학교 언어학과

+ 충남대학교 정보통신공학과

† 충북대학교 철학과

A 3D Vocal Tract Modeling and Vowel Discrimination of Korean Monophthongs [이, 에, 아, 오, 우, 으]

Cheoljae Seong*, Jongwon Park+, Guiryong Kim†

* Department of Linguistics, Chungnam National University

+Department of Information Communications Eng., Chungnam National University

† Department of Philosophy, Chungbuk National University

cjseong@cnu.ac.kr, jwpark@cnu.ac.kr, tragon@chungbuk.ac.kr

Abstract

We presents a new method for the measurement and analysis of the volume of the vocal tract using 3D magnetic resonance image. The relative ratios of volume A, B, and C, which are divided by the 2 constriction points formed on the horizontal and vertical plane in vocal tract, take a decisive role in discriminating Korean monophthong. Together with F1-F2 and the minimum cross sectional area in the vocal tract, the relative ratios of the regional volumes were proved to be significant parameter in statistic viewpoint.

1. 머리말

성도의 3차원 모델링 작업은 좀 더 정확한 조음-음향 모델을 얻기 위한 노력의 일환으로 발달되어 왔다. 60 - 70년대 X-ray 영상을 이용한 Fant[1]의 모델과 Liljencrants & Fant[2]에 소개된 성도 공명 주파수를 계산하기 위한 면적함수(area function) 등은 2차원적 평면에서 성도를 튜브 형태로 모델링한 뒤, 성도를 지나는 중앙선(midline)을 기초로 하여 이루어졌다.

Baer, Gore, Gracco, and Nye[3]; Narayanan, Alwan, and Haker[4]; Titze and Story[5], Engwall[6], Berends & Wismüller [7] 등으로 이어지는 자기공명 영상(MRI)에 바탕한 연구들은, 중앙 시상면(midsagittal plane)에 국한되었던 지난 연구들의 한계를 넘어서, 직접적으로 3차원 영상을 이용하여 면적 함수를 구할 수 있는 단계로 나아가는 모습을 보여준다. 조음 합성 시스템 구축을 지향하는 이와 같은 일련의 논문들은, 성도 전달함수(transfer function)를 구하기 위하여 다각형 모델(polygon model)과 같은 여러 가지 방법으로 성도를 모델링한 뒤, 면적 함수를 구하는 방식으로 연구를 진행한다(Foldvik et al.[8]).

이 논문은, 3차원 영상에 기반하여 얻을 수 있는 새로운 관점에서의 해부학적 조음 패러미터를 활용하여 모음을 기술하고 판별(discrimination)하는 일련의 과정을 제안하고자 한다. 이러한 작업은 한국어 조음 합성기(articulatory synthesizer)를 구현하기 위한 가장 기초적인 작업임과 동시에, 세계 언어들 간의 미묘한 조음적 차이를 해부학적 차원에서 기술할 수 있는 새로운 분류 기준을 세우는 작업이 될 수 있을 것이다.

모음 기술과 판별이라는 구체적인 과정으로 접근하기 위해, 능동 조음체와 수동 조음체가 이루는 전통적 관점에서의 좁힘점(constriction point)에 주목하고 있다.

자기공명영상을 이용하여 3차원 성도 모델을 구한 다음, 픽셀 숫자에 기반한 성도 면적 추적 기법을 활용하여 성도의 수평면과 수직면에서 각각 하나씩 총 2개의 좁힘점을 구한다. 이를 중심으로 얻을 수 있는 총 3개의 튜브에 대한 체적값을 구체적 목표값으로 삼는다. 성도 3차원 복원(reconstruction) 작업을 통해 구한 세 가지의 체적값을 더하고, 각 모음에 하나씩 할당할 수 있는 최소 좁힘점¹⁾의 단면적값(cross-sectional area)을 추가하면 총 4개의 패러미터가 설정되는 셈이다.

2. 성도 자기 공명 영상 수집

성도 자기 공명 영상은 20대 남성을 대상으로 하여 얻었다. 키 172 cm, 몸무게 74 Kg의 성악가로 표준어 화자이다. 한국어 단모음 6개 [이, 에, 아, 오, 우, 으]에 대해서 다음 표 1과 같은 수치로, 충남대학병원에서 촬영되었다.

| 패러미터 명 | 패러미터 수치 |
|---------------|--|
| 촬영 방향 | 1. Axial : 60장 2. Coronal : 36장 |
| 이미지 두께/ 간격 | 3mm/ 3mm |
| 이미지 크기 | 256 픽셀 x 256 픽셀 |
| 픽셀의 크기 | 1. Axial : 0.987500mm x 0.987500mm 2. Coronal : 0.703135mm x 0.703125mm |

표 1. 자기공명 영상 촬영을 위한 패러미터

3. 성도 분할(Segmentation) 및 3차원 복원(reconstruction)

본 연구에 사용된 자기공명 영상은 1차적으로, 병원의 의료 영상 저장 전달 시스템(PACS)에서 사용되는 다이콤(DICOM) 파일 포맷으로 저장된다. 다이콤 파일은 환자의 신상 정보 및 촬영에 관련된 각종 파라미터와 영상 자료가 모두 포함된 형태의 파일 포맷으로서, 본 연구팀은 이들 파일을 영상처리에 이용하기 위하여 영상 자료 부분만을 재처리하여 분할(segmentation) 작업을 수행하였다. 성도 분할을 위한 첫 번째 단계로, 다이콤 파일로부터 영상 자료 부분만을 획득한 후 이를 흑백 영상을 위한 무 손실 압축 파일 포맷인 PGM 파일

1) 예컨대, [이]의 경우는 육안으로 관찰되지 않는 혀 뒷면 아래쪽과 후인두(laryngo-pharynx)벽 사이에서 최소 좁힘점이 형성되고 [이]의 경우는 전설과 경구개 근처에서 최소 좁힘점이 형성된다. 성도를 따라 단면적값을 추적하면 최소값을 자동적으로 구할 수 있다.

포맷으로 변환하였으며, 두 번째 단계로, 변환된 PGM 파일을 대상으로 일반적인 영상 편집 응용프로그램과 도구(Tablet)를 활용하여 매뉴얼 방식으로 분할작업을 수행하였다[9, 10]. 최종적으로 얻은 영상은 흰색 픽셀과 검은색 픽셀로만 이루어진 2가(binary) 영상 자료로 만들어진다.

성도의 3차원 복원은, 분할작업 결과 영상을 이용하여 이루어지는데 이를 위해 보간(interpolation) 작업이 필요하다. 보간은 분할 작업된 영상을 3차원으로 재구성하기 위한 첫 단계로서, 각각의 슬라이스와 바로 뒤에 따라오는 슬라이스 사이의 연속된 변화를 표현해 주기 위해 임의의 슬라이스를 삽입하는 작업이다. 3차원 재구성을 위한 응용프로그램 중의 하나인 ImagJ를 이용하여 작업을 수행하였다.

3차원 복원 작업은 마이크로소프트사의 Visual C++ 6.0과 OpenGL 라이브러리를 이용하여 수행하였으며, 작업의 결과 3차원으로 재구성된 성도의 동영상 파일(상영 시간 약 21초)이 제작되었다. 그림 1, 2는 [이]모음 동영상 결과물의 방향별 캡처 영상을 보여주고 있다.

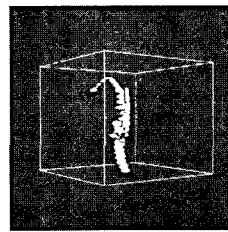


그림 1. [이] 모음, Axial 방향 영상

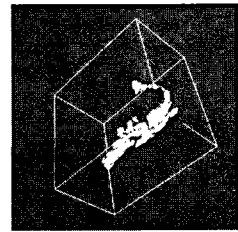


그림 2. [이] 모음, Coronal 방향 영상

4. 성도 체적 및 단면적 계산

자기공명 영상 한 슬라이스의 분할된 성도 이미지 부분 픽셀의 숫자를 활용하여 체적을 구하였다. 머리말에서 언급한 다각 모델링(polygon modeling)의 경우, 근사치 접근이라는 한계가 이미 노정되어 있기 때문에 본 논문에서 활용하고 있는 픽셀과 슬라이스 숫자, 그리고 슬라이스간의 간격을 이용한 체적 계산법이 좀 더 실측치에 가깝다고 평가한다. 이미 이 방법은 [9, 10]을 통하여 충분히 검증되었다. 단면적 또한 한 픽셀 크기와 숫자를 응용하면 고전적인 면적 함수(area function)를 고려하지 않고도 어렵지 않게 구할 수 있다. 체적을 구하는 공식은 다음 표 2와 같이 표현될 수 있다.

표 2 체적 계산 수식

```

if ( [Slice Number] == ([Start Slice Number]
or [End Slice Number]))
    Volume = ([Object Pixel #] *
    [Pixel Space] * ([Thick] * 1/2[Intv]))
else
    Volume = ([Object Pixel #] *
    [Pixel Space] * ([Thick] * [Intv]))
    
```

Object Pixel #: 성도에 해당하는 영역의 픽셀 수
 Pixel Space : 한 픽셀의 실제 크기
 Thick : 한 슬라이스의 두께
 Intv : 슬라이스간의 간격

성도 수평/수직면의 중첩점을 중심으로 하여 형성되는 3개의 튜브 각 영역의 체적을 구강 앞쪽 부분부터 차례로 체적 A, 체적 B, 체적 C로 명명하였다(그림 3).

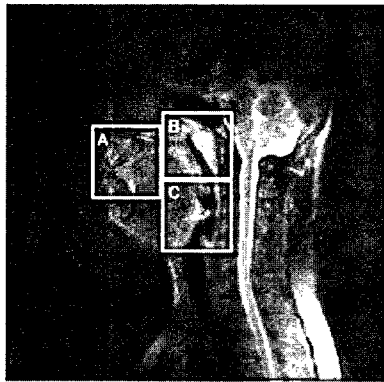


그림 3 수평/수직 중첩점을 중심으로 구분한 3 영역(A, B, C)

중첩점 두 군데를 중심으로 구한 A, B, C 영역의 체적값과 그 비율이 표 3에 제시되어 있다.

| 모음 | 영역 A | 영역 B | 영역 C | 전체 영역 |
|----|------------------|------------------|------------------|------------------|
| 이 | 3.524 (5.2) | 27.262 (40.2) | 36.967 (54.6) | 67.753 (100) |
| 에 | 14.942 (26.1) | 14.916 (26.0) | 27.417 (47.9) | 57.275 (100) |
| 아 | 48.082 (44.9) | 32.317 (30.2) | 26.715 (24.9) | 107.114 (100) |
| 오 | 37.587 (42.5) | 21.159 (24.0) | 29.647 (33.5) | 88.393 (100) |
| 우 | 29.945 (42.6) | 12.848 (18.3) | 27.469 (39.1) | 70.262 (100) |
| 으 | 18.092 (27.4) | 14.876 (22.5) | 33.001 (50.1) | 65.969 (100) |

표 3 성도의 영역 분할에 따른 체적(ml)/ ()는 비율값

각 영역 별로 부여된 비율값은 음향공간의 F1, F2 값과 마찬가지로 그 모음 고유의 특성을 나타내주는 패러미터로 설정할 수 있다. 어떤 언어의 모음이든 자기공명 영상 정보만 있다면 그 조음적 특성을 상호 비교 분석할 수 있는 패러미터라 할 만 하다. 표 3의 내용을 3차원 산점도로 나타내 보았다(그림 4).

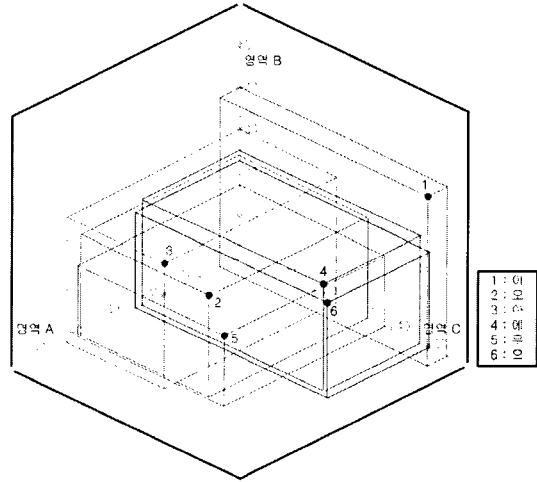


그림 4 성도의 영역 분할에 따른 체적 비율의 3차원 산점도

5. 모음 판별 함수

이 연구의 결과와 함께, 모음의 F1, F2, 그리고 각 모음에 대한 최소 중첩점의 단면적값까지 첨가하여 총 6개의 패러미터를 대상으로, 통계학적 판별분석을 수행해 보았다. 각 모음에 대한 변수 토큰이 하나 밖에 없으므로, 원 값을 기준으로 $\pm 5\%$ 차로 더하고 빼서 하나의 변수당 총 22개의 변수를 새롭게 생성하여 통계 작업을 하였다(최종적으로 $\pm 50\%$ 까지 진행). 각 모음별로 Fisher의 판별함수(discriminant function)와 정준(canonical) 판별 함수를 구하였고, 이 함수를 바탕으로 원래 모음 재배치를 해본 결과, 100 %의 결과를 얻어 내었다. 구해진 Fisher의 판별함수 중 [아] 모음을 예시하면 다음과 같다.

$$f([\text{아}]) = -32.796 - 0.017 * F1 - 0.011 * F2 - 2.667 * [\text{vol. a}] - 2.332 * [\text{vol. b}] + 2.59 * [\text{vol. c}] + 0.084 * [\text{cs_area_min}]$$

(vol. a/b/c: 성도의 수평/수직면에서 구해진 2개의 중첩점을 중심으로 구한 a/b/c 영역 체적; cs_area_min: 성도최소단면적)

6. 맺음말

성도(필터)의 특성을 가장 잘 나타내는 음향변수는, 성도의 공명특성인 F1, F2 값이라 볼 수 있다. 성도를 3차원으로 표현하여 각 모음에 따라 달라지는 좁힘점과 이에 따라 구분되는 세 가지 영역의 체적값을 비율값으로 나타내본 이번 연구의 접근법은 성도의 조음적 특성을 나타내는 새로운 물리적 접근법이라 할 만하다.

음향적 패러미터인 F1, F2 값, 조음적 패러미터인 체적값, 성도 단면적 값 등을 변수로 활용하여 설정된 모음 판별함수는, 다양한 모음의 음향적/조음적 특성을 동시에 기술하고 비교할 수 있게 하는 자질 특성(feature specification)이라 할 수 있다.

감사의 글

본 논문은 한국학술진흥재단의 인문사회 기초분야 연구지원(KRF-2002-074-AM 1021)의 도움으로 이루어졌다.

참고문헌

- [1] G. Fant, *Acoustic Theory of Speech Production*, Mouton, den Haag, 1960.
- [2] J. Liljencrants and G. Fant, "Computer program for VT-resonance frequency calculations," *KTH STL-QPSR vol. 4*, pp. 15-20, 1975.
- [3] T. Baer, J. Gore, R. Gracco, and P. Nye, "Analysis of Vocal Tract Shape and Dimension using Magnetic Resonance Imaging: Vowels," *JASA 90(2)*, pp. 799-828, 1991.
- [4] S. Narayanan, A. Alwan, and K. Haker, "Towards Articulatory-Acoustic Models for Liquid Approximants Based on MRI and EPG Data," *JASA 101(2)*, pp. 1064-1089, 1995.
- [5] I. Titze and B. Story, "Vocal Tract Area Functions from Magnetic Resonance Imaging," *JASA 100(1)*, pp. 537-554, 1996.
- [6] O. Engwall, "Modeling of the Vocal Tract in 3 Dimensions," *Proc. of Eurospeech 1999*, pp. 113-116, 1999.
- [7] Berends & Wismüller, "A Segmentation and Analysis Method for MRI Data of the Human Vocal Tract," *FIPKM 37*, pp. 179-189, 2001.
- [8] A. Foldvik, U. Kristiansen, J. Kvaerness, A. ,Torp,"3-Dimensional and Magnetic Resonance Imaging: A new Dimension in Phonetic Research," *Proc. of ICPHS 95*, pp. 46-49, 1995.
- [9] Y. C. Sung, C. J. Song, S. M. Noh and J.W. Park, "Tissue Segmentation and Volumetry on Brain MR Image of Coronal Section having the Partial Volume Artifact," *Proc. of 2000 IEEE International Symposium on Intelligent Signal Processing and Communication Systems, Vol.I*, pp. 369-372, 2000.
- [10] D .Y .Kim, J. W. Park, "Computerized quantification of carotid artery stenosis using MRA axial images," *Magenetic Resonance Imaging, Vol 22/3*, pp. 353-359, 2004.