

초보 학습자를 위한 주파수 분석을 통한 성악 연습

박소현, 임선영, 장우인, 박영호
숙명여자대학교 멀티미디어과학과
e-mail : {sohyun1218, sunnyihm, leader0710, yhpark}@sookmyung.ac.kr

A Study on Vocal Training by Analyzing Frequencies

So-Hyun Park, Sun-Young Ihm, Wu-In Zang, Young-Ho Park
Dept. of Multimedia Science, Sookmyung University

요 약

음악을 구성하는 중요한 3 가지 분류는 가락, 리듬, 박자이다. 성악을 학습하기 위하여 가락의 대표적인 속성인 음정을 분석한다. 서로 다른 음정은 각기 고유한 주파수를 갖고 있으므로, 주파수를 분석 함으로써 음정의 옳고 그름을 판단 할 수 있고, 이를 토대로 사용자를 교육할 수 있다. 본 논문에서는 안드로이드 플랫폼 상에서 사용자의 음성을 인식 하고 주파수로 변환하는 알고리즘과 주파수와 음정 매칭 알고리즘을 제안한다.

1. 서론

음악을 구성하는 중요한 3 가지 요소는 멜로디, 리듬, 박자이다. 멜로디, 리듬, 박자는 각각 음표의 높낮이, 음표의 길이의 다양성, 음악적인 시간을 구성하는 기본 단위를 의미 한다. 음악에는 성악, 기악, 관현악 등의 장르가 있는데, 음악의 3 요소는 세계의 장르 모두에서 학습자들이 기본적으로 습득해야 할 중요한 요소이다. 특히, 성악 장르에서는 정확한 음정으로 노래를 부르는 것은 중요하다. 음정이란, 멜로디의 구성 요소로써, 음의 높이에 따른 음표의 분류를 의미한다. 그리고, 멜로디를 구성하는 음표들을 서로 구분 하는 근본적인 요소는 주파수의 변화이다. 즉, 주파수의 변화를 분석 함으로써 사용자로부터 입력받은 음성의 음정이 옳고 그름을 확인 할 수 있다.

본 논문에서는, 사람 목소리의 음역 대에 대한 이론을 설명 한 후, 이론 적인 기반 위에 사용자 음성을 주파수로 변환하는 알고리즘을 제안한다.

본 논문의 구성을 다음과 같다. 2 장에서는 관련 연구를 비교, 분석하고 3 장에서는 음성의 주파수 변환 알고리즘과 주파수와 음정 매칭 알고리즘을 제안 하며 4 장에서는 결론 및 향후 연구에 대해 서술한다.

2. 관련연구

본 장에서는 관련 연구를 분석 및 설명한다. 음악 교육의 전자화는 성악, 바이올린, 피아노 등 많은 악기들을 대상으로 음정, 리듬 그리고 소리에 대한 구체적인 교육을 시행하고 있으며[1], [3] 장애인들을 위한 음악 교육 학습 방법도 제안한다[2].

Wang et al. [1]은 바이올리니스트를 위해서 음정이 옳고 그름을 확인하는 음정 학습 시스템을 설계하였다. 음정 학습 시스템은 바이올리니스트들이 학습 지

침으로 사용할수 있도록 주요 스케일과 아르페지오 스케일을 제공하였다. 그리고 그들의 학습 단계에 따라 속도를 다르게 설정할 수 있다. Sakajiri et al. [2]은 청각 장애인과 시각 장애인을 위하여 촉각을 이용한 디스플레이를 사용하여 음정 연습 시스템을 만들었다. 촉각을 이용한 음정 연습 시스템을 사용하여, 선천적인 시각 장애나 청각 장애를 앓았던 사람도 들을 수 있는 근육이 생긴다는 것을 발견하였다. Ferguson et al. [3]은 음악 중급 학습자의 들을 위한 소리의 시각화를 통한 소리 학습 시스템을 제안하였다. 소리 학습 시스템을 통하여, 이전의 음성 교육 학습 방법보다 학습자들에게 시각 자료를 제공하므로써 전달력을 높여 연습의 효율성을 높인다.

3. 알고리즘

본 장에서는 사람의 음역 대에 대한 이론을 설명한 후에, 음성을 주파수로 변환하는 알고리즘을 제안한다.

<표 1>은 성별에 따른 파트의 분류, 음역 대 그리고 해당 주파수 대역을 나타낸 것이다. 음역 대는 사람의 목소리가 낼 수 있는 주파수의 범위를 의미한다. 주파수는 초당 진동 수를 의미하고 음정마다 고유한 주파수를 갖고 있으며, 단위는 헤르츠(Hz)로 주파수가 클수록 높은 소리를 내고 주파수가 작으면 낮은 소리를 낸다. 여성의 파트는 음역 대의 종류에 따라 소프라노, 메조 소프라노, 알토로 분류하고 남성의 파트는 테너, 바리톤, 베이스로 분류한다. 음역 대의 인덱스는 글자와 숫자로 구성하는데, 글자는 도, 레, 미, 파, 솔, 라, 시를 알파벳 식인 C, D, E, F, G, A, B, C 로 표현한 것이고, 숫자는 옥타브의 단계를 의미하는데, 옥타브의 숫자가 클수록 높은 소리를 내고 옥타브의 숫자가 작을수록 낮은 소리를 낸다.

(그림 1)은 <표 1>을 오선지 상에 그림으로 표현한 것이다. <표 1>의 소프라노 음역 대인 C4 부터 A5 와 그에 해당하는 주파수 대역인 261Hz 부터 880Hz 까지를 (그림 1)에서 확인 할 수 있다. 또한, (그림 1)을 통하여 옥타브의 의미를 파악할 수 있는데, 한 개의 옥타브는 C 부터 B 까지 총 7 개의 음정을 포함하고, 그 보다 하나 위의 옥타브는 더 높은 소리를 낸다. 예를 들어 C4 부터 B4까지는 7 개의 음정을 포함하는 4 번째 옥타브를 의미한다.

<표 1> 성별에 따른 파트의 분류와 해당 음역 대

성별	파트	음역 대	주파수(Hz)
여	소프라노	C4 - A5	261 - 880
	메조 소프라노	G3 - G4	196 - 784
	알토	F3 - E5	174 - 659
남	테너	C3 - A4	130 - 440
	바리톤	G2 - G4	98 - 392
	베이스	F2 - E4	87 - 329



(그림 1) 오선지상의 음정 표현

음성의 주파수 변환 알고리즘과 주파수와 음정 매칭 알고리즘을 설명하기 이전에, 알고리즘을 실험하기 위한 구현 환경을 설명한다. 두 개의 알고리즘에서 사용하는 함수인 `Convert_Voice_To_Frequency` 는 TarsosDSP-Android-2.0 라이브러리에서 제공하는 함수이다. 본 논문에서는 TarsosDSP-Android-2.0 라이브러리의 정확도를 측정하기 위하여 `Pitch_Accuracy_Estimate` 라는 안드로이드 기반 어플리케이션을 사용한다.

`Convert_Voice_To_Frequency` 함수는 사용자의 목소리를 입력 받아 실수 형의 주파수로 바꾼다. TarsosDSP-Android-2.0의 주파수 정확도를 측정하기 위하여 고유한 주파수를 가진 특정 음을 출력하면 `Pitch_Accuracy_Estimate` 어플리케이션이 특정 음에 반응하는 실험을 하였다. 50 개의 음을 대상으로 실험한 결과, `Convert_Voice_To_Frequency` 함수는 50 개 음의 주파수를 정확하게 측정하는 것으로 보아 99.9%의 정확도를 보인다.

(그림 2)은 사용자의 음성(`UVoice`)을 주파수(`UFreq`)로 바꾸는 과정의 알고리즘이다. 음성의 주파수 변환 알고리즘의 입력 값(Input)은 실수 형의 데이터 형인 `UVoice` 이고 출력 값(Output)은 `UVoice` 를 주파수로 바꾼 실수 형의 데이터 형인 `UFreq` 이다. 첫 번째에서 두 번째 줄까지 `UVoice` 변수와 `UFreq` 변수를 실수형의 데이터 형으로 초기화 한다. 세 번째에서 여덟 번째

줄까지 다음의 과정을 반복한다. 세 번째 줄에서 `WHILE` 문을 사용하여 사용자의 음성의 입력이 끝날 때까지 세 번째에서 여덟 번째 줄을 반복하라고 명령한다. 네 번째 줄에서 다섯 번째 줄까지는 `IF` 문을 사용하여 `UVoice` 가 없다면 사용자에게 음성 재입력을 요청하는 에러 메시지를 전송한다. 여섯 번째부터 일곱 번째 줄까지는 `ELSE IF` 문을 사용하여, `UVoice` 가 존재한다면, 음성을 주파수로 바꾸는 함수인 `Convert_Voice_To_Frequency` 를 적용하여 `UFreq` 변수에 대입한다. 함수 `Convert_Voice_To_Frequency` 는 디지털화 된 사람의 음성을 실수 형의 주파수 형태로 변환하는 기능을 한다. 여덟 번째 줄에서는 `WHILE` 문을 종료 한다. 마지막으로 아홉 번째 줄에서는 음성을 실수 형으로 변환한 주파수 변수인 `UFreq` 를 주파수와 음정 매칭 알고리즘에 결과 값을 반환한다.

Algorithm 1: Transform_Voice_To_Frequency

```

Input: UVoice : user's voice
Output: UFreq : digitized voice frequency
Algorithm:
1: Initialize UVoice; //variable for saving user voice in float
2: Initialize UFreq; //variable for saving voice frequency in float

3: WHILE UVoice != End;
4:   IF UVoice == NULL THEN
5:     Print Error message to user to input voice again;
6:   ELSE IF UVoice != NULL THEN
7:     UFreq = Convert_Voice_To_Frequency(UVoice);
8:   END WHILE

9: RETURN UFreq;
    
```

(그림 2) 음성의 주파수 변환 알고리즘

(그림 3)은 `UVoice` 에 해당하는 음정(Note)를 출력해 주는 과정의 알고리즘이다. 주파수와 음정 매칭 알고리즘의 입력 값은 실수 형의 데이터 형인 `UFreq` 와 음정 주파수 배열(`NFreq`)이고, 출력 값은 문자 형의 데이터 형인 `Note` 이다. 첫 번째 줄에서 두 번째 줄까지 `Nfreq` 와 `Note` 를 실수 형의 데이터로 초기화 한다. 세 번째 줄에서는 `UVoice` 를 매개 변수로 받는 `Transform_Voice_To_Frequency` 함수의 결과 값을 `UFreq` 에 대입 한다. 네 번째 줄에서 아홉 번째 줄까지는 `UVoice` 가 끝날 때까지 다음의 과정이 반복 한다. 다섯 번째에서 여섯 번째 줄까지는 `IF` 문을 사용하여 `UFreq` 와 `NFreq` 가 같은지 확인하고, `UFreq` 와 `NFreq` 가 같다면 그에 맞는 음을 출력한다. 일곱 번째 줄에서 여덟 번째 줄까지는 마찬가지로 `ELSE IF` 문을 사용하여 `UFreq` 와 `NFreq` 가 같은지 확인하고, `UFreq` 와 `NFreq` 가 같지 않다면 사용자에게 음성 재입력을 요청하는 에러 메시지를 전송한다. 여덟 번째 줄에서는 `WHILE` 문이 종료 되고 열 번째 줄에서는 안드로이드 스크린 상에서 해당 음정을 결과값으로 출력한다.

Algorithm 2: Determine_Note_By_Voice_Frequency

Input: (1) *UFreq* - digitized voice frequency;
 (2) *NFreq* - array of note frequencies;
Output: *Note* - matched note;
Algorithm:

```

1: Initialize NFreq; //array for saving note frequency in float
2: Initialize Note; //variable for storing note in string

3: UFreq = Transform_Voice_To_Frequency(UVoice);

4: WHILE U Voice != End;
5:   IF UFreq == NFreq THEN
6:     Print corresponding Note;
7:   ELSE IF UFreq != NFreq THEN
8:     Print Error message to user to input voice again
9:   END WHILE

10: RETURN Note; //print matched note on android screen

```

(그림 3) 주파수와 음정 매칭 알고리즘

4. 결론 및 향후 연구

본 논문에서는 사람의 음역 대에 대한 이론을 설명한 후, 이론 적인 기반 위에 안드로이드 스마트 폰에서 사용자 음성의 주파수 분석하는 알고리즘과 주파수와 음정 매칭 알고리즘을 제안하였다. 추상적인 음성이라는 입력 값을 구체적인 실수 형의 데이터인 주파수로 변환함으로써 사용자의 음정을 정확하게 확인하고 교육 지침을 내릴 수 있다.

향후 연구에서는, 성악 연습을 위한 음정 연습에 더하여, 리듬 및 박자 연습의 구체적인 커리큘럼을 제안하고 자신의 연습 결과를 수치화 및 평균화하여 그래프화 하는 기능을 추가하고 성악 연습 시스템을 확장한다.

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2012003797)

참고문헌

- [1] J. H. Wang, S. A. Wang, W. C. Chen, K. N. Chang, "Real-Time Pitch Training System for Violin Learners," In Proceedings of 2012 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), pp. 163 - 168, 2012.
- [2] M. Sakajiri, K. Nakamura, S. Fukushima, S. Miyoshi, "Effect of voice pitch control training using a two-dimensional tactile feedback display system," In Proceedings of 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 2943 - 2947, 2012.
- [3] S. Ferguson, A. V. Moere, D. Cabrera, "Seeing Sound: Real-time Sound Visualization in Visual Feedback Loops Used for Training Musicians." In Proceedings of 9th International Conference on Information Visualization, pp.97 - 102, 2005.