

말하기 학습을 위한 음성 비교 및 점수화 알고리즘 Speech Comparison and Scoring Algorithm for Speaking Learning

이영준¹ · 박춘소¹ · 장문수^{1*} · 장영은²

Young Jun Lee, Chun So Park, Moon Soo Chang^{*} and Young Eun Chang

¹서경대학교 컴퓨터과학과

Department of Computer Science, Seokyeong University

²중앙대학교 사회복지학부

School of Social Welfare, Chungang University

요 약

한국어 듣기, 말하기 학습을 하는 외국인 또는 다문화가정 유아들에게 듣기 자료와 본인의 말을 비교해주면 학습효과가 크게 나타난다. 본 논문에서는 참조 음성을 듣고 사용자가 말한 음성과 비교하여 점수화하는 알고리즘을 제안한다. 제안하는 알고리즘은 두 음성을 정규화하고 중첩된 면적을 계산하여 점수화한다. 파형의 면적 비교는 파형의 형태에 따른 변화를 반영하지 못하기 때문에 본 논문에서는 파형의 기하학적 모양을 비교해 점수화하는 방법을 제안한다.

키워드 : 말하기 학습, 음성 비교, 점수화 알고리즘, 음성 파형

1. 서 론

최근 한국어를 배우는 외국인의 수가 늘고 있다. 한국어교육기관대표자협의회에 따르면 한국어를 배우는 외국인이 2007년 740명에서 2015년 4만3천여명으로 8년 만에 58배로 늘어났다. 이에 따라 한국어 학습의 중요성이 증가하고 있다.

한국어를 학습하는 방법은 일반적으로 말하기, 듣기, 쓰기, 읽기가 있는데 외국인들이 한국어 학습에 있어 듣기, 말하기와 같은 회화 부분에 가장 큰 어려움을 느끼고 있다고 한다.

본 논문에서는 한국어 듣기, 말하기 학습을 하는 외국인 또는 다문화가정 유아들을 위해 참조 음성을 듣고 사용자가 말한 음성과 비교하여 점수화하는 알고리즘을 제안한다. 제안하는 알고리즘은 노래방기계에서 소리 세기를 반영하여 점수화하는 것에 기초하여[1] 두 음성의 소리 세기를 기준으로 정규화하고 면적과 모양을 비교하여 점수화하는 방법이다. 그리고 이 알고리즘을 바탕으로 프로그램을 구현하고자 한다.

2. 음성 비교와 점수화

시간의 흐름에 따른 소리 세기를 음성 파형으로 나타내면 파형의 면적과 모양을 알 수 있다. 본 논

문에서는 두 파형을 비교하기 위해서 소리의 세기를 기준으로 정규화한 후에 두 파형의 중첩된 면적과 모양을 비교한다. 그리고 면적 점수와 모양 점수를 종합하여 최종 점수로 나타낸다. 그림 1은 두 음성의 소리 세기를 시간의 흐름에 따라 파형으로 나타낸 것이다.

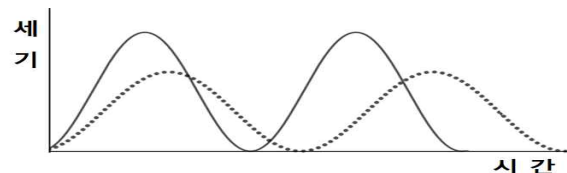


그림 1. 시간의 흐름에 따른 두 파형

2.1 파형의 정규화

사람마다 가지고 있는 소리의 세기가 다르므로 소리 세기를 정규화하는 과정이 필요하다. 정밀한 정규화를 위해 음성을 스무딩(Smoothing)하여 잡음을 제거한 후 사용자 음성과 참조 음성의 소리 세기 상위 15% 값에 대한 평균의 비율을 구한다. 그리고 사용자 음성을 비율에 맞춰 정규화한다. 그림 2는 그림 1의 점선 파형을 실선 파형에 맞게 정규화한 것이다.

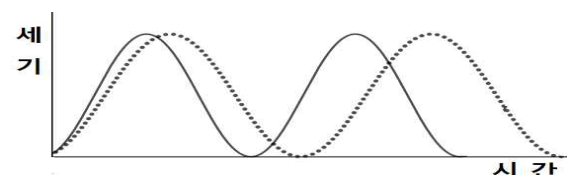


그림 2. 정규화한 점선 파형

* 교신저자

이 논문은 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임
(NRF-2017S1A5A2A03068038)

2.2 파형의 면적 비교

소리 세기의 변화는 그림 2의 그래프에서 파형으로 나타난다. 그러므로 두 음성의 세기 비교는 파형과 x축으로 둘러싸인 부분의 면적으로 나타낼 수 있다. 파형의 면적을 비교하여 점수화하기 위하여 본 논문에서는 두 파형의 중첩되는 면적과 차이 나는 면적을 구하여 식(1)과 같은 비례식으로 나타낸다. 중첩되는 면적은 두 파형의 시간 흐름에 따른 소리 세기가 겹쳤다는 것을 의미하며 차이 나는 면적은 두 파형의 세기가 다른 부분을 의미한다.

$$areaScore = \frac{CommonPart}{CommonPart + DifferencePart} \times 100 \quad (1)$$

2.3 파형의 모양 비교

두 음성 파형을 면적으로 비교하여 점수화하면 소리 세기가 얼마나 비슷한지 알 수 있다. 그러나 면적을 통한 점수화에는 다음과 같은 문제가 있다.

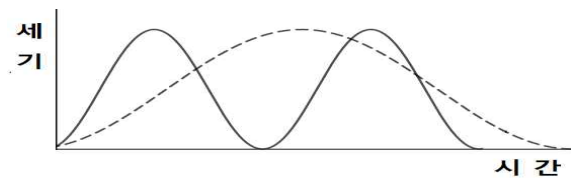


그림 3. 모양이 다른 두 파형

면적을 통해 점수화하면 그림 2와 그림 3은 점수가 비슷하다. 그러나 눈으로 봤을 때 그림 2의 두 파형의 모양이 그림 3의 두 파형의 모양보다 더 유사하므로 더 높은 점수를 받는 것이 타당하다. 본 논문에서는 이를 보완하기 위하여 파형의 형태에 따른 기하학적 모양을 점수화하여 면적에 의한 점수를 보정한다. 그림 4는 파형의 기하학적 모양을 분석한 것이다.

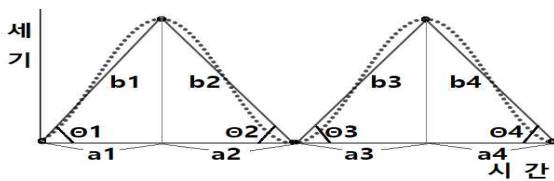


그림 4. 파형의 기하학적 모양 분석

그림 4에서 파형의 시작점, 극점, 끝점을 기준으로 구간을 나눠 x축의 길이를 구한 것은 a1, a2, a3, a4이며 각 점 사이의 길이를 구한 것은 b1, b2, b3, b4와 같다. 삼각비에 의해서 a를 b로 나눈 값은 $\cos\theta$ 와 같다. 파형의 모양을 통해 점수화한 것은 식 (2)와 같다.

$$shapeScore = \sum_{k=1}^{n+1} \left(\frac{100}{n+1} \times \frac{\cos\theta_{Small}}{\cos\theta_{Big}} \right) \quad (2)$$

식 (2)에서 n은 비교할 두 파형의 극점 개수 중 많은 쪽의 수를 의미하며 n+1은 극점을 기준으로 나뉘진 구간의 개수이다. 100을 n+1로 나눔으로써 구간마다 동일한 점수를 부여하게 된다. 그리고 비교하는 두 파형을 각각 구간별로 $\cos\theta$ 값을 구한다. $\cos\theta_{Big}$ 는 각 구간에서 두 파형의 $\cos\theta$ 값 중 큰 값, $\cos\theta_{Small}$ 는 작은 값을 의미한다. 이 모양 점수는 0에서 100 사이의 값을 가지게 된다.

2.4 점수화

두 음성 파형의 면적과 모양을 비교하여 점수화하기 위해서 일반인들을 대상으로 실험하여 기준 모양 점수와 가중치를 구한다. 그리고 면적 점수에다 모양 점수와 기준 모양 점수의 차이에 가중치를 부여한 값을 종합하여 최종 점수를 구한다.

$$Score = areaScore + \{\alpha \times (shapeScore - CriteriaShapeScore)\} \quad (3)$$

3. 구현 및 결과

제안한 알고리즘을 바탕으로 Android Studio 3.0.1 개발환경에서 Java언어를 사용하여 프로그램을 구현한다. 그 결과는 그림 5와 같다.



그림 5. 프로그램 구현 결과

4. 결론

본 논문에서는 참조 음성을 듣고 사용자가 따라 말한 음성으로 비교하여 점수화하는 알고리즘을 제안하고 이를 바탕으로 프로그램을 구현하였다.

그러나 본 논문에서는 알고리즘을 단순화하기 위하여 음의 높낮이를 고려하지 않고 크기만 비교하고 있다. 향후 연구에는 본 논문의 알고리즘에 사용되지 않은 음성 주파수까지 비교하는 점수화 알고리즘이 필요하다.

참고 문헌

- [1] Hung-Min Wang. "Scoring device and method for a karaoke system". US 6,326,536 B1, filed May 8, 2000, and issued Dec 4, 2001