

말하기 학습을 위한 음성 비교 및 점수화 알고리즘

이영준¹ · 박춘소¹ · 장문수^{1*} · 장영은²

¹서경대학교 컴퓨터과학과

²중앙대학교 사회복지학부

ABSTRACT

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 비교해 점수화하는 알고리즘을 제안한다. 제안하는 알고리즘은 두 음성을 정규화하고 중첩된 면적을 계산해 점수화한다. 파형의 면적 비교는 파형의 형태에 따른 변화를 반영하지 못하기 때문에 파형의 기하학적 모양을 비교하여 면적에 의한 점수를 보정한다.

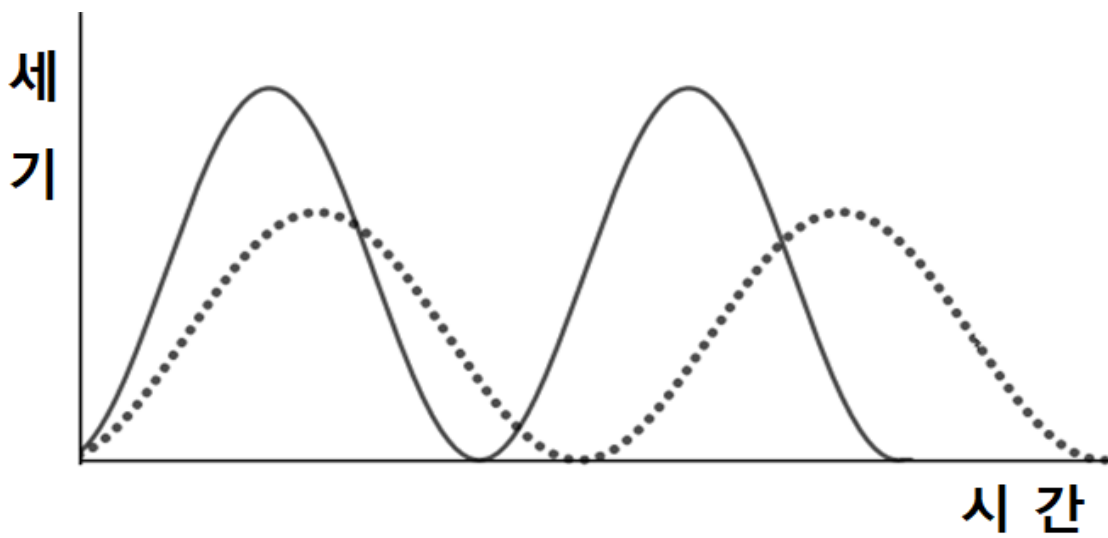
INTRODUCTION

한국어를 배우는 외국인의 수가 늘면서 한국어 학습의 중요성이 증가하고 있다. 외국인들은 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 **두 음성의 소리 세기를 기준으로 정규화하고 면적과 모양을 비교해 점수화하는 방법**이다.

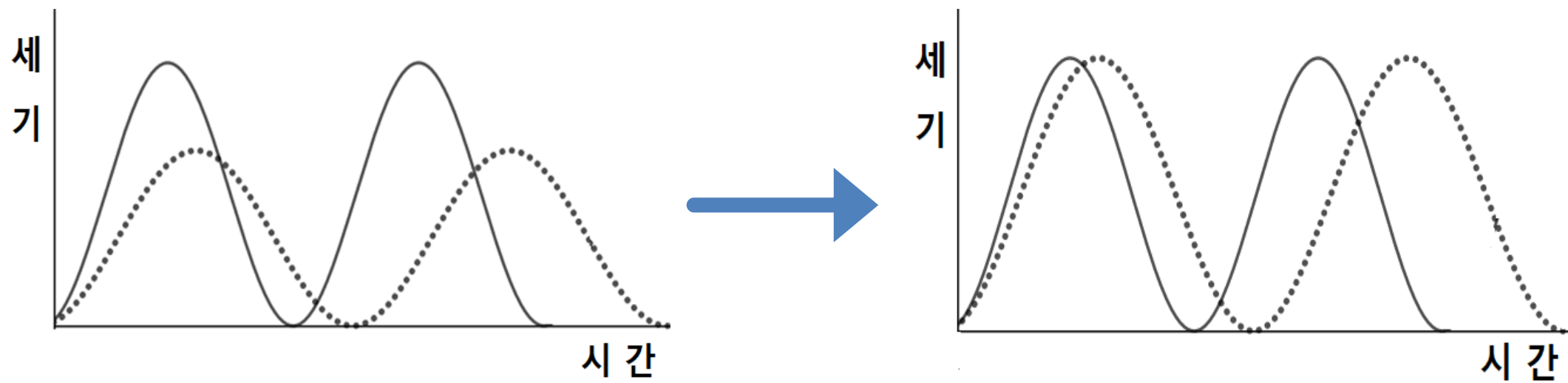
MAIN IDEA

음성을 시간 흐름에 따른 소리세기로 나타내면 파형의 면적과 모양을 알 수 있다. 본 논문에서는 두 파형을 비교하기 위해서 소리의 세기를 기준으로 정규화 한 후에 두 파형의 중첩된 면적과 모양을 비교한다. 그리고 면적 점수와 모양 점수를 종합하여 최종 점수로 나타낸다. 아래 그림은 두 음성의 소리 세기를 시간의 흐름에 따라 나타낸 것이다.



✓ 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. 아래 그림은 점선 파형을 실선 파형에 맞게 정규화한 결과이다.



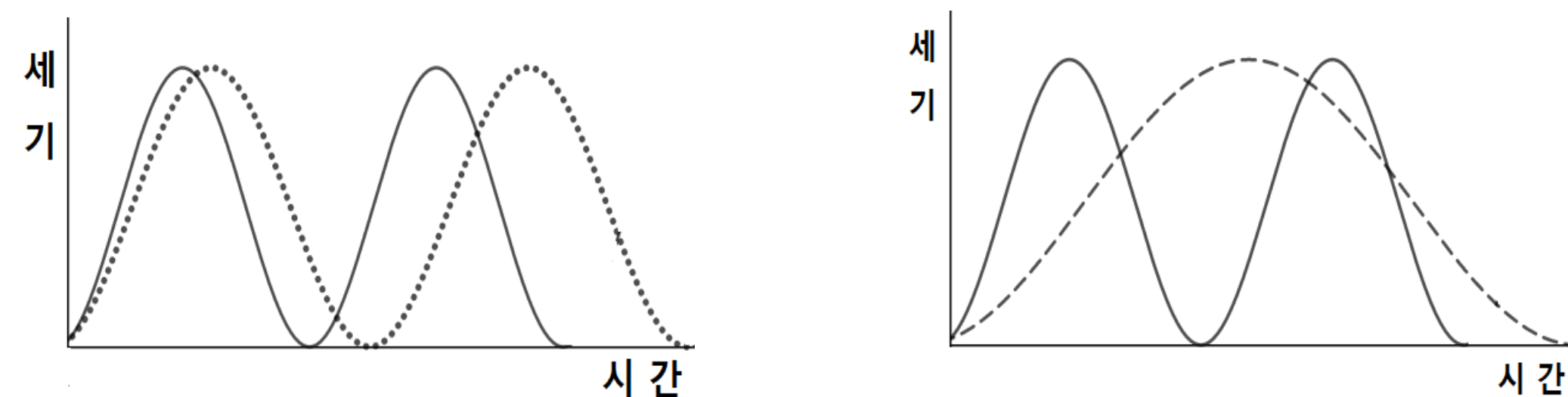
✓ 파형의 면적 비교

소리 세기의 변화는 위의 그래프처럼 파형으로 나타난다. 그러므로 두 음성의 세기 비교는 파형과 x축으로 둘러싸인 부분의 면적으로 나타낼 수 있다. **파형의 면적을 비교하여 점수화하기 위해 두 파형의 중첩되는 면적과 차이 나는 면적을 구하여 아래와 같은 비례식으로 나타낸다.** 중첩되는 면적(Common Part)은 두 파형의 시간 흐름에 따른 소리 세기가 겹쳤다는 것을 의미하며 차이나는 면적(Difference Part)은 두 파형의 세기가 다른 부분을 의미한다.

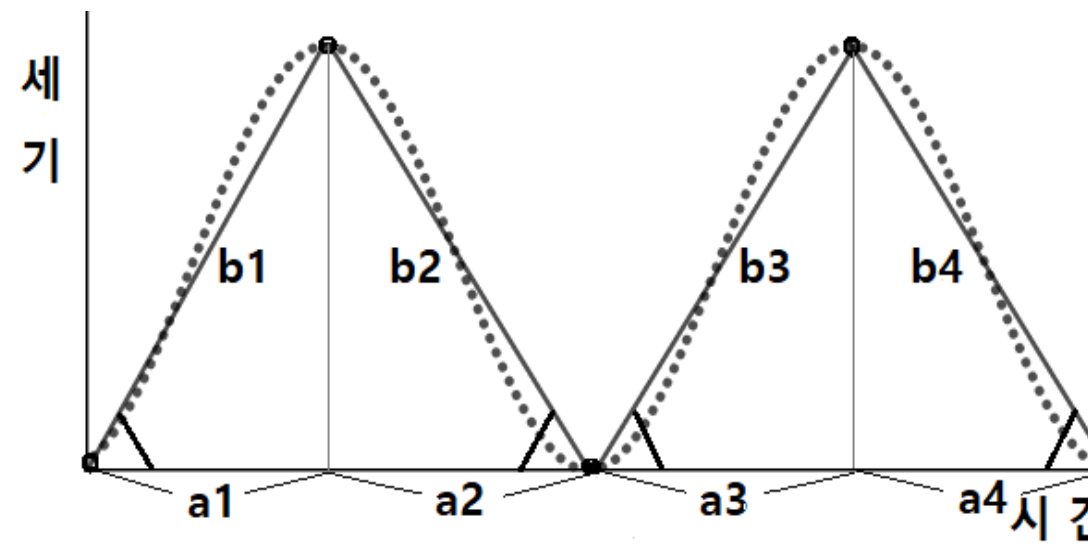
$$\text{areaScore} = \frac{\text{Common Part}}{\text{Common Part} + \text{Difference Part}} \times 100$$

✓ 파형의 모양 비교

면적을 통해 점수화하면 아래 두 그림의 점수가 비슷하다. 그러나 눈으로 봤을 때 **왼쪽 그림의 두 파형의 모양이 더 유사하므로** 더 높은 점수를 받는 것이 타당하다. 따라서 **이를 보완하기 위하여 파형의 형태에 따른 기하학적 모양을 점수화하여 면적에 의한 점수를 보정한다.**



아래 그림은 파형의 기하학적 모양을 분석한 것이다. 파형의 극점의 개수를 구한 후 (n = 극점 개수) 파형을 시작점, 극점, 끝점을 기준으로 구간을 나누고 (a = 구간), 100점을 구간별로 동일하게 나눠 점수를 부여한다 (각 구간별 점수 = 100 / 구간개수). 그리고 비교하는 두 파형을 각각 구간별로 $\cos\theta$ 값을 구한 후 ($\cos\theta = a/b$) **각 구간에 대해서 두 파형의 $\cos\theta$ 값을 비교한다** ($\cos\theta_{\text{Big}}$ = 큰 값, $\cos\theta_{\text{Small}}$ = 작은 값).



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

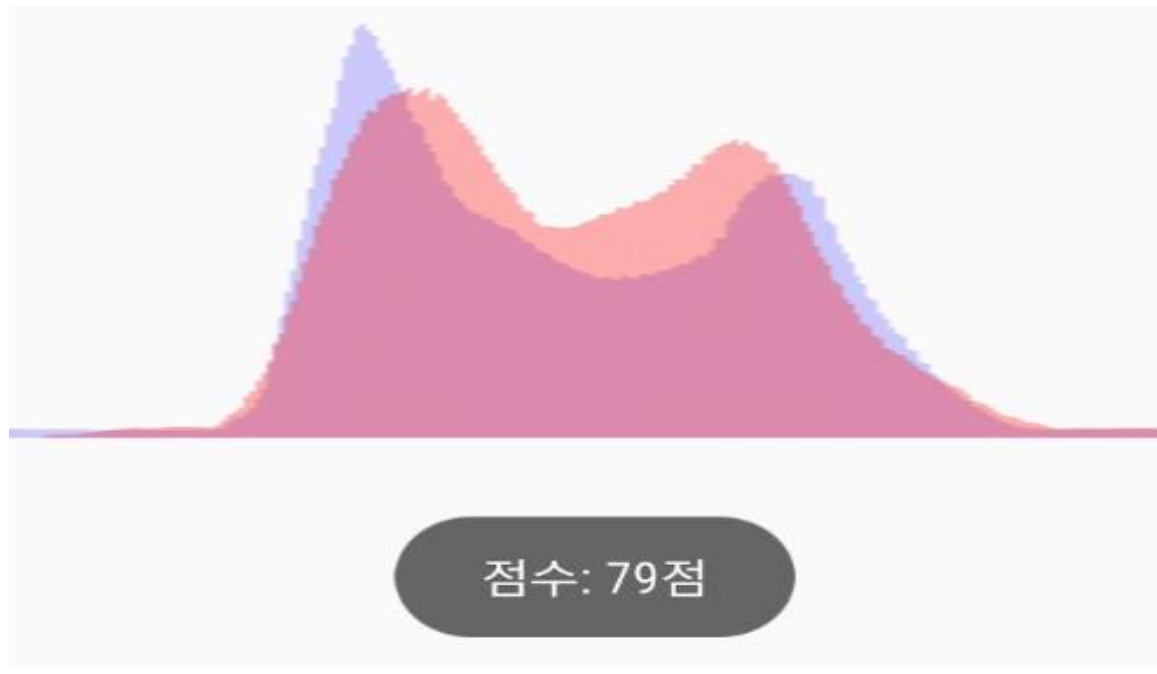
✓ 점수화

면적점수와 모양 점수를 반영해 최종 점수로 나타낸다. 실험을 통해 모양 점수들의 평균인 기준 모양 점수와 모양 점수에 부여할 가중치를 구한다.

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

RESULT

제안한 알고리즘을 바탕으로 Android Studio 3.0.1 개발환경에서 Java언어를 사용하여 프로그램을 구현한다.



CONCLUSION

본 논문에서는 참조 음성을 듣고 사용자가 따라 말한 음성과 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.**