

## 선형예측분석과 인근 프레임 유사도에 의한 음성 신호의 지속시간 변경

김명\*, 장경아, 홍성훈, 배명진  
승실대학교 \*컴퓨터학과, 정보통신공학과

### Time Scale Modification of the Speech Signal with Linear Prediction Analysis and Similar of the Neighboring Frames

Ming Jin\*, KyungA Jang, SungHun Hong, MyungJin Bae  
Department of Computer Science\* & Information and Telecom. Engr.  
Soongsil University  
mjjin1978@hanmail.net / mjbae@ssu.ac.kr

#### 요 약

본 논문에서는 선형예측 분석방법과 인근 프레임간의 유사도에 의한 음성신호의 지속시간 변경법을 제안하였다. 지속시간은 음성의 속도나 말의 리듬을 결정하며, 강세나 의미의 강조 등을 나타내는 중요한 정보를 포함하고 있다. 따라서 음성의 지속시간을 자유롭게 조절할 수 있다면 합성된 음성의 자연성 향상은 물론 언어장애인의 발음교정 및 어학학습 등에 사용할 수 있을 것이다. 지속시간 변경에서 흔히 사용되는 PSOLA합성법은 음성의 피치검출이 반드시 이루어져야 한다. 따라서 음성에 잡음이 섞인 경우 음성의 피치검출이 어렵게 된다. 따라서 피치 정보 오류로 인하여 PSOLA 방식을 이용하여 지속시간을 변경한 음성은 음질이 현저히 떨어지는 현상을 나타낸다. 따라서 제안한 방법에서는 선형예측 분석에 의한 프레임과 이전 프레임의 상관관계에 근거하여 지속시간을 변경하여 줄 수 있으며 지속시간이 변경된 음성의 평균 MOS는 3.8좌우의 음질을 가진다.

#### I. 서 론

지속시간은 음성의 속도나 말의 리듬을 결정하며, 강세나 의미의 강조 등을 나타내는 중요한 정보를 포함하고 있다. 음성의 지속시간을 자유롭게 조절할 수 있다면 합성된 음성의 자연성 향상은 물론 언어장애인의 발음교정 및 어학학습 등에 사용할 수 있을 것이다.

본 논문에서는 선형예측 분석과 인근 프레임간의 유사도에 의한 음성신호의 지속시간 변경법을 제안하고자 한다. 제안한 방법은 선형예측 계수와 에러 신호를 분리하고 에러 신호의 지속시간을 시간 영역에서 변경하여 준다. 지속시간을 확장할 경우 지속시간 변경율에 따라 현재 프레임의 따라오는 인근 프레임에서 프레임크기의 변경율만큼의 샘플을 현재 샘플과 유사도를 비교하여 제일 유사한 부분을 현재 프레임의 끝부분에 추가 시킨다. 지속시간을 줄일 경우에는 현재 프레임의 끝부분에서 변경에 따라 에러신호를 없

애주고 이어지는 부분에서 음성의 뒤는 효과를 제거하기 위하여 Median 필터링을 하여 주게 된다.

본 논문은 모두 6 장으로 구성되었으며 2 장에서는 선형예측 분석에 대하여 설명하였고 3 장에서는 인근 프레임의 유사도 측정 방법에 대하여 설명하였으며 4 장에서 본 논문에서 제안한 지속시간 변경법에 대하여 설명하였다. 그리고 5 장에서 실험 및 결과를 보여 주며 6 장에서 결론을 맺었다.

#### II. 선형예측분석

인근한 음성표본들은 높은 상관관계를 가지고 있다. 이러한 상관관계를 가정하여 간단한 선형예측을 식 (2.1)과 같이 표현할 수 있다.

$$y_n \cong a_1 y_{n-1} + a_2 y_{n-2} + \dots + a_p y_{n-p} \quad (2.1)$$

이 식에서 음성신호의 표본된 값( $y_n$ )은 상수  $a$ 가 곱해진 과거의  $p$  표본들에 의해서 예측할 수 있다는 가정