

22. 피치 정보를 이용한 모음의 특징 벡터 변별력 향상에 관한 연구

컴퓨터공학과 하 동 경
지도교수 신 옥 근

음성인식을 위한 대부분의 음성신호처리에서는 음성신호가 짧은 시간 간격 동안에는 통계적으로 정적이라는 특징을 이용한다. 먼저, 음성신호를 10~30ms 사이의 고정된 시간간격의 프레임 단위로 나눈 다음, 각 프레임에 대한 음성학, 혹은 음향학적인 특징을 추출하여 인식의 기본단위로 사용하고 있다.

대부분의 화자 독립 음성인식에서는 정규화된 음성 특징을 추출하기 전에 화자 종속적인 특성인 피치 정보는 제거할 대상이었다. 본 논문에서는 유성음이 피치주기를 단위로 준주기적이며, 한 주기 동안의 음성신호열은 성문에서 발생하는 각각의 임펄스 입력에 대한 성도 필터의 임펄스 응답으로 볼 수 있다는 점에 착안하여 피치 정보를 이용한 신호처리 방법을 제안한다.

피치 정보를 이용하기 위해 먼저 AMDF를 이용해 유성음의 주기인 피치를 실시간으로 검출하는 방법을 개발하였다. 제안한 방법으로 피치 검출 실험을 수행한 결과, 2.85%의 오검출률을 보여 기존 방법들과 비교해 비교적 우수한 성능을 가짐을 확인할 수 있었다.

이렇게 검출한 피치 정보를 음성신호 처리에 이용하기 위한 두 가지 방안을 검토하였는데,



먼저 음성 프레임의 길이를 피치 주기와 일치시킨 프레임의 특징 벡터들과 고정 길이 프레임의 특징 벡터들의 변별력을 조사해 보았다. LPC 계수, CEP 계수, WCEP 계수, 그리고 MFCC 등의 특징 벡터에 대해 실험해 본 결과, 고정 길이 프레임의 경우보다 제안한 방법이 3~10% 정도 변별력이 우수함을 알 수 있었다. 다음으로 MFCC의 멜 단위 필터뱅크에 피치 정보를 적용시킴으로써 각각의 필터가 화자의 피치와 포먼트 특성을 잘 반영할 수 있도록 하는 방법을 제안하였다. 이 방법을 검증하기 위한 실험 결과, 분산을 기준으로 한 변별력은 14% 증가하였으며, 특징 벡터들을 직접 비교한 판별력 실험에서는 8.5% 증가하여 제안한 방법이 종래의 필터뱅크보다 우수함을 확인 할 수 있었다. 제안한 방법으로 추출한 음성 특징 벡터를 HMM이나 신경회로망 등의 인식기에 사용할 경우 보다 좋은 인식률을 보일 것으로 기대된다.

