

# 스펙트럼 보상에 의한 피치 검출에 관한 연구

안중현, 배명진  
송실대학교 정보통신공학과

## On a pitch detection with spectrum compensation technique of speech signal

Jung Hyun Ahn and Myung Jin Bae  
Dept. of Information and Telecommunication Engr.  
Soongsil University  
mjbae@ssu.ac.kr

### 요약

음성인식, 합성 및 분석과 같은 음성신호처리 분야에 있어서 기본주파수 즉, 피치를 정확히 검출하는 것은 중요하다. 그러나 포먼트의 영향과 천이진폭의 영향 때문에 음성신호에서 피치를 정확히 구하는 것은 매우 어렵다. 더구나 노이즈가 포함된 신호에서는 더 더욱 어려워진다. 따라서 본 논문에서는 켈스트럼영역에서 포먼트의 영향을 제거하고 밴드 필터링 한 후 기본 피치 정보를 강조 보상하여 피치주기를 검출하는 방법을 새로이 제안한다.

### 1. 서론

음성인식, 합성 및 분석과 같은 음성신호처리 분야에 있어서 기본주파수 즉, 피치를 정확히 검출하는 것은 중요하다. 만일 음성신호의 기본주파수를 정확히 검출할 수 있다면 음성인식에 있어서 화자에 따른 영향을 줄일 수 있기 때문에 인식의 정확도를 높일 수 있고, 음성합성 시에 자연성과 개성을 쉽게 변경하거나 유지할 수 있다. 또한 음성분석 시에 피치에 동기 시켜서 분석하면 윈도우 길이의 영향을 제거할 수가 있고 또한 성문의 영향이 제거된 정확한 성도파라미터 특성을 구할 수 있게 된다.[1]

이러한 피치검출의 중요성 때문에 피치검출에 대한 방법들이 다양하게 제안되어졌는데, 시간영역법, 주파수영역법, 시간-주파수 영역법 등으로 구분할 수 있다. 시간영역 피치검출법은 병렬처리법, AMDF법, ACM법 등이 있는데, 시간단위의

분해능은 높일 수 있으나 잡음의 영향이나 천이 구간에서 추출 오차가 큰 단점이 있다.

본 논문에서는 켈스트럼영역에서 켈스트럼을 보상함으로써 포먼트의 영향을 제거하고, 잡음의 영향을 줄일 수 있는 새로운 시간-주파수 피치검출법을 제안하고자 한다.

### 2. 켈스트럼 분석법

스펙트럼 신호에 로그를 취한 신호를 켈스트럼이라한다. 켈스트럼상에서는 여기정보와 여파기 정보를 쉽게 분리할 수 있다.

음성 신호는 시간영역에서 시간에 따라 느리게 변화하는 여파기 성분 (포먼트)과 상대적으로 높은 주파수 성분인 준 주기적 펄스인 여기성분의 컨벌루션으로 나타낼 수 있으며, 주파수 영역에서 음성 스펙트럼은 여기 스펙트럼과 여파기 스펙트럼의 곱으로 나타내어진다. 이러한 스펙트럼에 로그를 취하면 곱의 형태에서 합의 형태로 변환되는데 이 신호에 역 FFT를 취하면 0 근처의 낮은 주파수쪽에는 여파기 정보가 몰리고 상대적으로 높은 주파수 쪽에는 여기 정보가 있게 된다.[2][3]

### 3. 켈스펙트럼 보상에 의한 피치검출

음성신호는 켈스트럼 영역에서 스펙트럼 분석이 이루어진다. 그림1은 본 논문에서 사용한 블록도이다. 먼저 window leakage 영향을 제거하기 위해 여기서는 hamming window를 사용한다. 켈스트럼 블록에서는 포먼트 성분과 여기 성분으로

분리된 켈스트럼 신호를 얻고 Lifter에 의해 포만트 성분만을 제거하여 여기성분만을 남긴다.

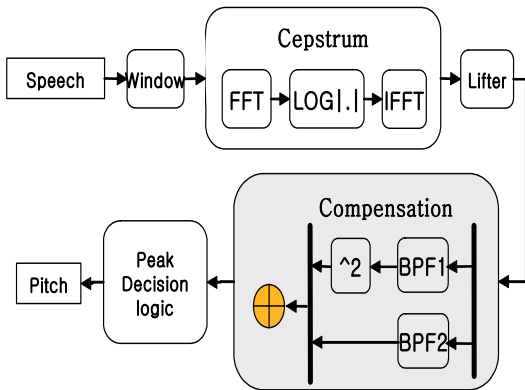


그림 1. 제안한 블록다이어그램

이 켈스트럼 신호는 두개의 밴드패스 필터를 통해 10~500Hz (BP1) 와 500Hz~4Khz (BP2) 성분으로 나누고 피치 성분이 많이 포함된 BP1의 신호는 자승을 취해 강조한 후 BP2 신호와 더한다. BP2 신호는 피치 신호의 섬세한 정보를 담고 있으므로 BP1에서 제거된 섬세 정보를 보상하여 예리한 피크 위치 정보를 유지한다. 따라서, 노이즈가 제거고 (500Hz 이상 신호) 주 피치 정보가 강조된 BP1 신호와 피크 섬세 정보를 담고 있는 BP2 신호와의 보상을 통해 노이즈에 강한 피치 검출이 가능해진다 (그림2).

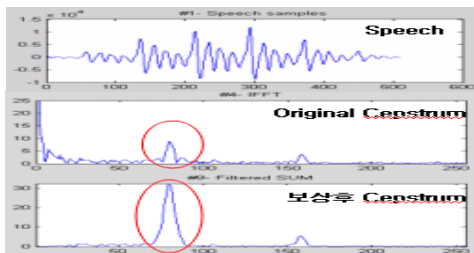


그림 2. 보상 전/후 켈스트럼

실험에서는 깨끗한 스피치 신호에 랜덤 노이즈를 더하여 S/N비를 낮추면서 피치 피크의 변화를 살펴 보았다. 보상 알고리즘에 의한 방법에서 노이즈에 1dB 더 강한 결과를 얻었다 (표1).

표1. 보상 전/후 S/N 비 비교

음성시료	피치검출가능최소S/N비(dB)	
	기존방법	제안한 방법
시료 #1	3.5	1
시료 #2	1.2	1.2
시료 #3	3	2.2
평균	2.57	1.47

#### 4. 결론

본 논문에서 제안한 방법은 시간-주파수영역으로 켈스트럼을 보상하여 기본하모닉스를 강조하고, 피치의 측정 분해능을 높이기 위해 시간영역에서 결정하는 방법이다. 제안한 방법은 노이즈가 섞인 신호에서 좋은 결과를 보였으며 1dB 이상의 더 낮은 S/N비를 얻었다. 그러나 밴드 필터링이 추가되어 연산량이 올라가므로 이러한 처리 과정을 단순화하는 노력이 필요하다.

#### 5. 참고문헌

- [1] 조왕래, 김종국, 배명진. "FFT 켈스트럼 처리시간 단축에 관한 연구", 음성과학 제10권 26호 pp.57-64. 2003년 6월
- [2] Alan V. Oppenheim, Ronald W. Schaffer, "Homomorphic Analysis of Speech, IEEE Trans. Audio Electroacoust., vol. AU-16, pp.221-226, June 1968
- [3] 정찬중, 배명진, 심태보, "대역 제한된 신호에서 피치 검출에 관한 연구", 대한전자공학회, 추계 학술 발표 논문집