

LSP를 이용한 적응 밴드 스펙트럼 평탄화에 의한 피치 검색 방법에 관한 연구

서지호, 배명진
송실대학교 정보통신공학과

A Study On the Pitch Extraction by the Spectrum Flattening in an Adaptive Sub-band using LSP.

JiHo Seo, MyungJin Bae
jihoseo@hotmail.com, mjbae@ssu.ac.kr

요약

음성인식, 합성 및 분석과 같은 음성신호처리 분야에 있어서 피치검출이나 포먼트검출은 매우 중요하다. 주파수 영역의 스펙트럼 신호는 잡음이 부가되는 경우에도 고조파정보와 포먼트 포락선 정보를 유지하기 때문에 음성신호처리분야에서 매우 유용하다고 할 수 있다. 고조파 정보나 포먼트 포락선 정보는 피치검출과 포먼트 주파수 검출에 직접 이용된다. 하지만 두 성분을 분리하는 방법에 따라 피치검출이나 포먼트 주파수 검출에 영향을 미칠 수 있으므로 기존의 방법보다 두 성분을 더 잘 분리할 수 있는 방법이 필요한 것이다.

본 논문에서는 스펙트럼 신호를 최대한 평탄화시킴으로써 포먼트의 영향을 제거하고 고조파 성분을 분리해 내어 이를 피치검출에 사용한다. LSP를 이용하여 적응적 밴드에서 평탄화를 시도하고 이를 피치 검출에 이용하였다.

1. 서론

음성은 사람이 다른 도구 없이 사용하는 정보 전달 매체로서 가장 많이 이용될 뿐 아니라 가장 간편한 수단이다. 이렇게 손쉬운 통신 수단인 음성을 인간과 기계 사이의 대화의 한 방법으로 사용하면 효과적일 것이다. 이러한 생각을 가지고 인간과 기계 사이의 통신수단으로 음성을 이용하고자 하는 시도가 오래 전부터 있어 왔다. 이를 위해서는 기계가 인간의 말을 알아듣도록 하는 음성인식기술과 기계가 음성을 통하여 인간에게 정보를 줄 수 있도록 하는 음성 합성기술이 동시에 개발되어야 한다.

본 논문에서는 주파수 영역에서 LSP를 이용하여 스펙트럼을 평탄화함으로써 포먼트의 영향을 제거하고 FFT 포인터의 수를 늘리지 않고도 주파수 해상도를 높여 피치 검출의 정확성을 높이는 피치검출법을 제안하고자 한다.

2. 피치 검색 시 고려사항

음성신호의 피치주기를 정확하고 신뢰성 있게 측정하는 것은 다음과 같은 여러 가지 이유 때문에 매우 어렵다. 첫 번째 이유는 성문의 여기파형이 완전히 주기적인 파형이 아니라는 것이고, 두 번째 어려움은 성도의 성문간의 상호작용 때문이다. 세 번째 문제점은 유성음구간 동안 각 피치주기의 정확한 시작과 끝을 정의하기 어렵다는 것이고, 네 번째 문제점은 무성음과 낮은 레벨의 유성음을 구별해야 하는 것이며, 다섯 번째 문제점은 전화시스템을 통하여 전송되는 음성의 피치검출 문제에 직면할 때에 부가적인 복잡성이 발생하게 된다는 것이다.

3. 제안한 피치검색방법

LSP를 구하는 과정을 보면 입력 음성신호를 받고 각 프레임마다 hamming window를 취한다. 윈도우를 취한 데이터를 이용해 LPC 계수를 찾아내어 LSP 파라미터를 얻어낼 수 있다. 10차의 LSP 파라미터는 각각 주파수 위치를 나타내고 있으며 두 파라미터의 위치간격으로 포먼트가 위치한 주파수 대역을 알아낼 수 있다. 이를 이용하여 선형보간하는데 있어 좀더 세밀한 선형적인 기술기를 추정할 수 있다. 또한 선형보간된 신호로부터

고조파 간격을 구하기 위해 자기 상관법을 사용하였다.

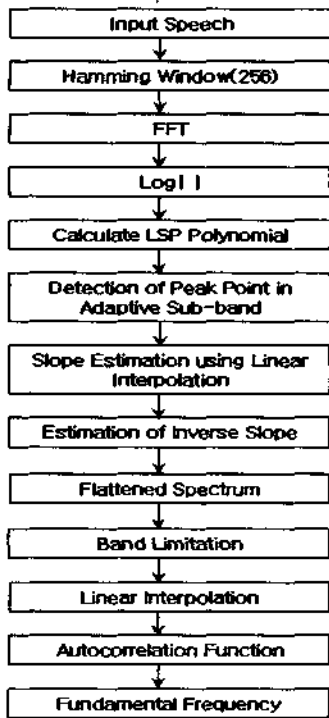


그림 1. 제안한 피치검출과정

4. 실험 및 결과

음성신호를 입출력하기 위한 상용화된 16비트 A/D 변환기를 인터페이스하여 8kHz의 표본율로 데이터를 입력하였다. 각 시료에 대해 한 프레임의 길이를 256표본 단위로 처리하였다. 처리결과와 성능을 측정하기 위해 다음의 대표적인 문장을 연령층이 다양한 남녀 5명의 화자가 각 5번씩 발성하여 시료로 사용하였다.

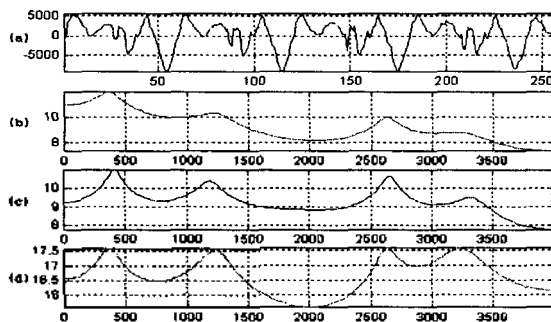


그림 2. 유성음의 기율기 측정

(a) 음성신호 (b) 원음성의 LPC분석 (c) 자기상관법을 이용한 LPC분석 (d) 제안한 방법을 이용한 LPC분석

또한 각 음성에 잡음이 추가되었을 때의 성능을 측정하기 위해 백색잡음을 사용하였고 4가지 발성의 SNR이 각각 30dB, 20dB, 10dB, 6dB, 3dB가 되도록 잡음을 부

가하였다.

표 1. 제안한 방법에 대한 Gross Error Rate(단위 %)

	30dB	20dB	10dB	6dB	3dB
발성1	0.33	1.36	3.43	3.93	4.63
발성2	0.96	1.12	2.12	2.56	3.71
발성3	1.31	2.74	3.44	4.32	5.92
발성4	0.72	1.62	3.56	5.16	7.21
Average	0.83	1.71	3.13	3.99	5.36

4. 결론

하모닉스 간격을 측정하여 피치를 검출할 때는 포만트 성분을 효과적으로 제거할 필요가 있는데 기존에 사용된 방법에는 LPC법과 Cepstrum법이 있다. 하지만 모두 정확히 포만트 성분을 모델링하고 있지 못하다. 따라서 더 정확히 포만트 성분을 제거할 수 있는 방법이 필요하다.

본 논문에서는 주파수영역에서 피치를 구하는 방법 중에서 고조파 분석법을 개선하여 피치 검색 방법의 성능을 향상시켰다. 피치정보는 음성 신호에 있어서 매우 중요하다. 앞으로 본 연구를 바탕으로 배경잡음에도 훨씬 강하고 더 낮은 에러율을 갖는 피치검출기를 구현하기 위해 계속 연구해나가야 할 것이다.

감사의 글

본 연구는 한국과학재단 특정기초연구 (과제번호 R01-2002-000-00278-0)의 지원으로 이루어 졌음

참고문헌

1. 배명진, "디지털 음성분석", 동영출판사, 1998. 4.
2. 민소연, 강은영, 배명진, "성문특성이 제거된 성도 특성에 관한 연구", 대한전자공학회, 추계 종합학술대회, 2000년 11월 25일.
3. 김영규, 민소연, 배명진, "부밴드 선형근사에 의한 여기 스펙트럼 검출에 관한 연구", 제15회 신호처리 합동학술대회, Vol.15, No.1, pp.219, 2002년 9월 28일.