

## 음성 하모닉스 스펙트럼의 피크-피팅을 이용한 피치검출에 관한 연구\*

## A Study on the Pitch Detection of Speech Harmonics by the Peak-Fitting

김 종 국\*\* · 조 왕 래\*\* · 배 명 진\*\*

Jong-Kuk Kim · Wang-Rae Jo · Myung-Jin Bae

## ABSTRACT

In speech signal processing, it is very important to detect the pitch exactly in speech recognition, synthesis and analysis. If we exactly pitch detect in speech signal, in the analysis, we can use the pitch to obtain properly the vocal tract parameter. It can be used to easily change or to maintain the naturalness and intelligibility of quality in speech synthesis and to eliminate the personality for speaker-independence in speech recognition. In this paper, we proposed a new pitch detection algorithm. First, positive center clipping is process by using the incline of speech in order to emphasize pitch period with a glottal component of removed vocal tract characteristic in time domain. And rough formant envelope is computed through peak-fitting spectrum of original speech signal infrequency domain. Using the roughed formant envelope, obtain the smoothed formant envelope through calculate the linear interpolation. As well get the flattened harmonics waveform with the algebra difference between spectrum of original speech signal and smoothed formant envelope. Inverse fast fourier transform (IFFT) compute this flattened harmonics. After all, we obtain Residual signal which is removed vocal tract element. The performance was compared with LPC and Cepstrum, ACF. Owing to this algorithm, we have obtained the pitch information improved the accuracy of pitch detection and gross error rate is reduced in voice speech region and in transition region of changing the phoneme.

**Keywords:** Pitch Detection, Harmonics Spectrum, Formant Envelope, Residual Signal, Pitch Period

## 1. 서 론

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, 음성합성시에 자연성과 개성을 쉽게 변경하거나 유지할 수 있다. 또한 분석시 피치에 동기 시켜 분

\* 본 연구는 한국과학재단 특정기초연구(과제번호 R01-2002-000-00278-0)의 지원에 의하여 이루어졌음.

\*\* 숭실대학교 정보통신학과

석하면 성문의 영향이 제거된 정확한 성도 파라미터를 얻을 수 있다.

이러한 피치검출의 중요성 때문에 피치검출에 대한 방법들이 다양하게 제안되었으며 이러한 피치검출 방법은 시간영역법, 주파수영역법, 시간-주파수영역법으로 구분할 수 있다. 시간영역 검출법은 파형의 주기성을 강조한 후에 결정논리에 의해 피치를 검출하는 방법으로 병렬처리법, AMDF법, ACM법 등이 있다. 시간영역 검출법은 보통 시간영역에서 수행되므로 영역의 변환이 불필요하고 합, 차, 비교논리 등 간단한 연산만 필요하게 된다. 그러나, 음소가 천이구간에 걸쳐 있는 경우에는 프레임 내의 레벨변화가 심하고 피치주기가 변동하기 때문에 피치검출이 어렵다. 특히 잡음이 섞인 음성의 경우에는 피치검출을 위한 결정논리가 복잡해져서 검출 오류가 증가되는 단점이 있다[1][2]. 주파수영역의 피치검출법은 음성 스펙트럼의 고조파 간격을 측정하여 유성음의 기본주파수를 검출하는 방법으로 고조파분석법[3], Lifter법, Comb-filtering법 등이 제안되어 있다. 일반적으로 스펙트럼은 한 프레임(20-40 ms) 단위로 구해지므로, 이 구간에서 음소의 천이나 변동이 일어나거나 배경잡음이 발생하여도 평균화되므로 그 영향을 적게 받는다. 그러나 처리 과정상 주파수영역으로의 변환과정이 필요함으로 계산이 복잡하며, 기본주파수의 정밀성을 높이기 위해 FFT의 포인트 수를 늘리면 그만큼 처리시간이 길어지고 변화특성에 둔해지게 된다. 시간-주파수 혼성영역법은 시간영역법의 계산시간 절감과 피치의 정밀성, 그리고 주파수영역법의 배경잡음이나 음소변화에 대해서도 피치를 정확히 구할 수 있는 장점을 취한 것이다. 이러한 방법으로는 Cepstrum법, 스펙트럼비교법 등이 있고, 이 방법은 시간과 주파수영역을 왕복할 때 오차가 가중되어 나타나므로 피치추출의 영향을 받을 수 있고, 또한 시간과 주파수영역을 동시에 적용하기 때문에 계산과정이 복잡하다는 단점이 있다[3][4].

본 논문에서 제안한 피치검출 알고리즘에서 스펙트럼의 평탄화는 스펙트럼 신호로부터 포먼트의 영향이나 천이진폭의 영향을 제거하는 것이다. 따라서 정확한 피치검출과 포먼트검출에 적용할 수 있으며 스펙트럼 신호를 최대한 평탄화시킴으로써 포먼트의 영향을 제거하고 고조파 성분을 분리해 낸다. 따라서 본 논문에서는 시간영역파형에서 먼저 성도성분이 제거된 성문특성 즉 피치주기 성분을 강조하기 위해 음성신호의 기울기를 이용한 positive center clipping을 수행하고 주파수 영역에서는 원래의 음성신호의 스펙트럼과 센터클립된 음성신호의 스펙트럼과의 peak-fitting을 수행하고 선형 균등 인터플레이션(스무딩)을 통해 평탄화된 스펙트럼을 얻었다. 결과적으로 원 음성신호의 스펙트럼과 스무딩된 포먼트 포락선의 대수차에 의해서 평탄화된 하모닉스(Harmonics) 스펙트럼을 구할 수 있었으며 이 스펙트럼을 IFFT 수행하여 성도성분이 제거된 잔차신호(Residual signal)를 구할 수 있었다. 구해진 잔차신호를 ACF (Autocorrelation function)를 수행하여 피치(pitch)를 구하였다.

## 2. 피치 검출시의 문제점

음성신호의 피치주기를 정확하고 신뢰성 있게 측정하는 것은 여러 가지 이유 때문에 아주 어렵다. 첫 번째 이유는 성문의 여기파형이 완전히 주기적인 파형이 아니라는 것이다. 한 주기 내에 있는 파형의 세부구조와 함께 변화하는 음성파형의 주기를 측정하는 것은 어려운 작

업이다. 두 번째는 피치주기를 측정하는데 있어서 어떤 경우에는 성도의 포먼트가 성문과형의 구조를 완전히 바꿀 수 있기 때문에 피치주기를 검출하기가 어렵게 된다. 일반적인 상호작용은 조음기관이 빠르게 변화하는 경우이며 피치검출시에는 방해요소로 나타나게 된다[6]. 세 번째 문제점에는 유성음구간 동안 각 피치주기의 정확한 시작과 끝을 정의하기 어렵다는 것이다. 피치주기의 정확한 시작과 끝의 위치를 선택하는 것은 임의적일 수 있기 때문이다. 예를 들면, 음성과형에 근거하여 주기의 시작과 끝을 정의하기 위한 필요한 것은 각 피치주기의 시작과 끝점을 정의하기 위한 후보로는 주기내의 영교차 간격을 들 수 있다. 이러한 측정에서 필요한 것은 각 피치주기의 시작과 끝점의 정확한 위치를 정의하기 위해서는 주기에서 주기까지 일관성이 있어야 한다는 것이다. 이러한 일관성이 부족하면 잘못된 피치주기를 측정하게 된다. 음성과형의 주기성분을 영교차 단위로 검출하면 포먼트 구조에 민감하게 된다. 피치를 검출할 때 나타나는 네 번째 문제점은 무성음과 낮은 레벨의 유성음을 구별해야 하는 것이다. 많은 경우에 무성음 구간과 레벨이 낮은 유성음 구간의 변화는 별로 뚜렷하지 않기 때문에 정확한 위치를 찾기가 어렵다. 다섯 번째 문제점은 전화시스템을 통하여 전송되는 음성의 피치검출 문제에 직면할 때에 부가적인 복잡성이 발생하게 된다. 음성에 대한 전화시스템의 영향은 선형, 비선형처리를 포함하면서 음성신호에 잡음이 추가되는 경우이다. 전화시스템을 선형여파기로 간주하면 기본주파수와 고조파들의 성분을 약하게 만들 수 있는 대역통과여파기로 볼 수 있다. 이 때문에 전화음질에 대해 신호의 주기성을 검출하기가 더 어렵게 된다. 음성신호에 대한 전화시스템의 비선형특성은 전송시스템의 채널에 따라 좌우된다.

### 3. 제안한 알고리즘

#### 3.1 시간영역처리

본 논문에서는 계산량을 줄이기 위해서 전처리과정으로 positive center clipping을 하였다. 피치검출과정을 수행하기 위해 먼저 식 (1)과 같이 중앙클리핑을 수행한다.[7]

$$\begin{aligned} y(n) &= x(n) - C_L & \text{if } x(n) \geq C_L \\ y(n) &= x(n) + C_L & \text{if } x(n) \leq -C_L \\ y(n) &= 0 & \text{otherwise} \end{aligned} \quad (1)$$

클리핑 레벨은 일반적으로 다음과 같이 식 (2)로 나타낸다.

$$C_L = 0.3A_{\max} \quad (2)$$

여기서  $A_{\max}$ 는 음성 세크먼트의 절대값의 최대값이다. 센터 클리핑 함수는 음성신호 중에서 pitch에 해당하는 성분은 크기가 크게 나타나는 특징을 이용해서 잔여성분을 제거하는 방법이다. 본 논문에서는 클리핑 레벨( $C_L$ )을 프레임 내의 음성신호의 레벨의 60% 이상을 기준으로

하였으며 먼저 시간영역처리에서 식 (3)과 같이 음성파형의 기울기를 구하였다.

$$\text{기울기} = \frac{y_2 - y_1}{x_2 - x_1} \quad (3)$$

$x_1$  = 한 프레임 내 첫 subframe 구간에서 최대값 위치

$x_2$  = 한 프레임 내 마지막 subframe 구간에서 최대값 위치

$y_1$  = 한 프레임 내 첫 subframe 구간에서 샘플의 최대값

$y_2$  = 한 프레임 내 마지막 subframe 구간에서 샘플의 최대값

제안한 논문에서 positive center clipping을 한 이유는 먼저 계산량 감소와 피치주기를 용이하게 찾기 위함이고 주파수 영역에서 좀 더 평탄한 스펙트럼을 구하기 위해 음성신호의 기울기를 측정하여 positive center clipping을 하였다.

### 3.2 주파수영역처리 방법

주파수영역처리에서 두 신호의 스펙트럼(원래의 음성신호와 센터 클리핑된 음성신호)을 peak-fitting을 하기 위하여 원래의 음성신호와 positive center clipping한 신호를 식 (4)와 같이 DFT하였다.

$$S(k) = \sum_{n=0}^{N-1} s(n) e^{-j\frac{2\pi}{N}kn} \quad 0 \leq n \leq N-1 \quad (4)$$

평탄화된 센터 클리핑한 신호의 스펙트럼을 원 신호의 스펙트럼과 Fitting하기 위하여 프레임별 구간에서 최대의 peak point과 그 점의 위치(position)를 찾아서 피치주기를 결정하였다. 검색구간 검출방법은 프레임내의 피치간격 즉 처음 두 샘플간의 피크점을 밴드사이즈로 정하고 4 샘플이나 5 샘플의 간격으로 전 프레임에 대하여 검색한다. 또한 밴드사이즈의 반으로 중첩해서 피치간격을 검색하였다. 밴드사이즈의 두 샘플의 피크를 찾아서 그 점의 최대값과 위치를 결정한다.

피크점들은 샘플들의 기울기가 증가하면 +1, 감소하면 -1로 결정을 하고 그 기울기를 더해서 0이 되면 피크점으로 결정하였다. 결정된 피크값과 그 위치를 원 음성신호의 스펙트럼과 Fitting을 한다. 그리고 포맷트 포락선을 만들기 위해 peak-fitting된 원 음성신호의 스펙트럼 상에서의 peak point들을 가지고 전체 프레임의 peak point들을 연결하고 원 음성신호의 포맷트 포락선과 동일하게 만들기 위하여 선형 균등 인터플레이션을 수행하였다. 결과적으로 스무딩된 포맷트 포락선을 구할 수 있었다. 주파수영역에서 LPC 역필터링의 결과와 같이 원 신호의 스펙트럼과 스무딩된 포맷트 포락선과의 대수차에 의하여 평탄화된 하모닉스 스펙트럼을 얻었다. 구해진 평탄화된 하모닉스 스펙트럼을 식 (5)와 같이 역푸리에변환(IFFT)을 수행하면 잔차신호를 구할 수 있으며 이 신호를 식 (6)을 이용하여 자기상관신호를 얻을 수 있다. 따라서 구해진 잔차신호를 ACF를 수행하여 피치(pitch)를 구하였다.

$$x(k) = \frac{1}{N} \sum_{n=-N/2}^{N/2-1} X(n) e^{j2\pi nk/N} \quad (k = k_0, \dots, k_0 + N - 1) \quad (5)$$

$$\phi(k) = \sum_{m=-\infty}^{\infty} x(m)x(m+k) \quad (6)$$

어떤 신호가 주기가 P인 주기신호라면  $\phi(k)$  역시 주기를 갖고,  $\phi(k)$ 의 최대값들은  $k=0, \pm P, \pm 2P, \dots$ 의 피치주기에서 발생한다.

본 논문에서 제안한 방법인 피치검출과정을 그림 1의 블록도로 나타내었다.

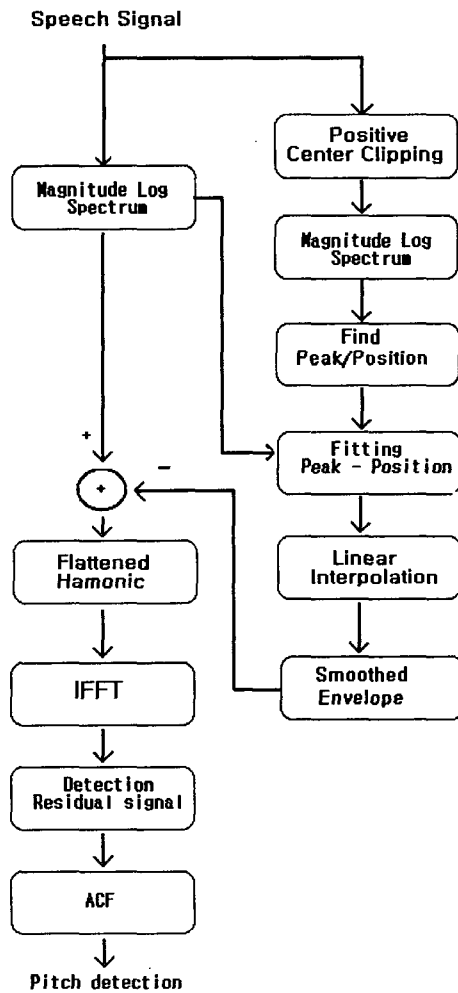


그림 1. 제안한 알고리즘의 블록도

그림 1의 피치검출과정에서 각 음성시료에 대해 한 프레임의 길이를 256 샘플로 하여 3개의 Subframe 나누어 첫 번째와 마지막 Subframe의 최대값을 찾아 음성파형의 기울기를 구하고 주기성을 강조하기 위해 positive 성분의 피크값을 찾아서 센터 클리핑을 수행하였다. 또

한 센터 클리핑된 음성신호를 주파수 영역으로 변환 후 센터 클리핑된 음성신호의 스펙트럼의 peak값과 위치(position)를 찾은 후 원래의 음성신호의 스펙트럼과의 peak-fitting을 통해 대략적인 포먼트 포락선을 구하였다. 구해진 대략적인 포먼트 포락선을 선형 균등 인터플레이션을 수행하여 스무딩된 포먼트 포락선을 구할 수 있었다[8]. 결과적으로 원래의 음성신호의 스펙트럼과 스무딩된 포먼트 포락선의 대수차에 의해서 평탄화된 하모닉스 스펙트럼을 구할 수 있었으며 이 스펙트럼을 IFFT를수행하여 성도성분이 제거된 잔차신호를 구할 수 있었다[9]. 구해진 잔차신호를 ACF를 수행하여 피치를 구하였다.

아래의 그림 2와 그림 3은 본 논문에서 제안한 알고리즘에서 평탄화된 하모닉스성분을 구하기 위해 먼저 원래의 음성신호와 센터 클리핑된 음성신호의 peak point를 찾은 출력 파형을 보여주고 있다.

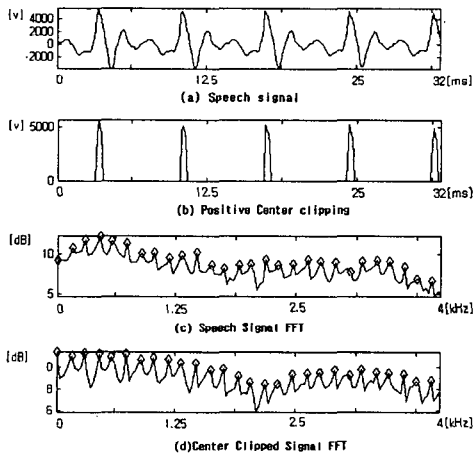


그림 2. 제안한 알고리즘의 출력(유성음구간)

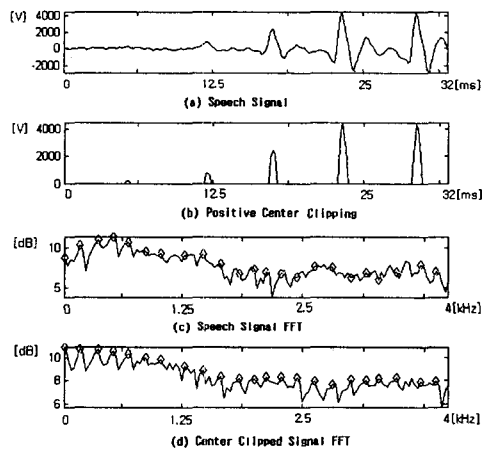


그림 3. 제안한 알고리즘의 출력(전이구간)

그림 2(b)와 3(b)는 positive center clipping을 보여주고 있으며 그림 2(c)와 2(d), 3(c)와 3(d)는 원래의 음성신호와 positive center clipping 신호의 log magnitude spectrum을 나타내었으며 여기에서 point들은 피치구간에서의 최대 peak를 표시한 것이다. 특히 그림 2(d)와 3(d)에서 스펙트럼이 원 음성신호의 스펙트럼보다 평탄화되었음을 보여준다. 그리고 본 논문에서 제안한 피치검출을 하기 위하여 그림 2(c)와 3(c)를, 그리고 그림 2(d)와 3(d)를 각각 peak-fitting하여 peak point와 그곳의 위치(position)를 연결하여 대략적인 포먼트 포락선을 구한다. 구해진 대략적인 포먼트 포락선을 선형 균등 인터플레이션을 수행하여 스무딩된 포먼트 포락선을 구한다.

#### 4. 실험 및 결과

컴퓨터 시뮬레이션에 이용한 장비는 IBM-PC (Pentium-III) 시스템이며 여기에 음성신호

를 입출력하기 위한 상용화된 16 비트 AD/DA변환기를 인터페이스하여 아래의 문장들을 남녀 각 10 명에게 발성시키면서 8 kHz의 표본율로 데이터를 입력하였다. 제안한 방법을 구현하기 위해서 C-언어와 MATLAB으로 구현하여 수행하였다.

발성 1: /인수네 꼬마는 천재소년을 좋아한다./

발성 2: /예수님께서 천지창조의 교훈을 말씀하셨다./

발성 3: /창공을 헤쳐 나가는 인간의 도전은 끝이 없다./

발성 4: /숭실대학교 정보통신과 음성통신 연구팀이다./

그림 4와 그림 5는 3절에서 보여준 그림 2와 그림 3에서 구해진 대략적인 포먼트 포락선을 선형 균등 인터플레이션을 수행하여 구한 평탄화된 하모닉스 스펙트럼과 잔차신호를 보여주고 있다.

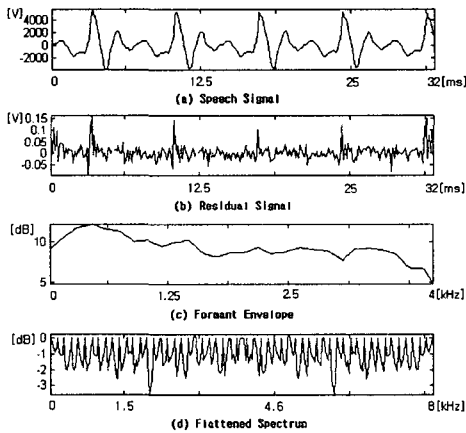


그림 4. 제안한 방법의 출력(유성음구간)

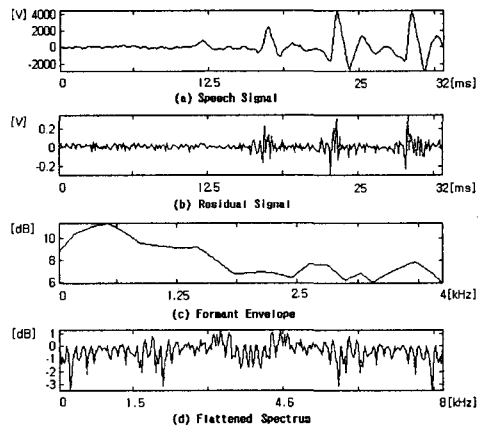


그림 5. 제안한 알고리즘의 출력(전이구간)

그림 4(b)와 그림 5(b)는 그림 4(d)와 그림 5(d)를 평탄화된 하모닉스 스펙트럼을 IFFT수행하여 구한 잔차신호이며 피치구간별 최대 peak를 잘 나타내고 있다. 그림 4(c)와 그림 5(c)는 원 음성신호의 스펙트럼과 positive center clipping을 한 스펙트럼과의 peak-fitting을 하여 대략적인 포먼트 포락선을 선형 인터플레이션을 수행하여 구한 스무딩된 포먼트 포락선이다. 또한 그림 4(d)와 그림 5(d)는 원 음성신호의 스펙트럼과 본 논문에서 구한 스무딩된 포먼트 포락선인 그림 4(c)와 그림 5(c)와의 대수차에 의해 구해진 평탄화된 하모닉스 스펙트럼을 나타내고 있다[5].

그림 6과 그림 7은 본 논문에서 제안한 알고리즘의 피치검출 성능을 ACF와 비교하였다.

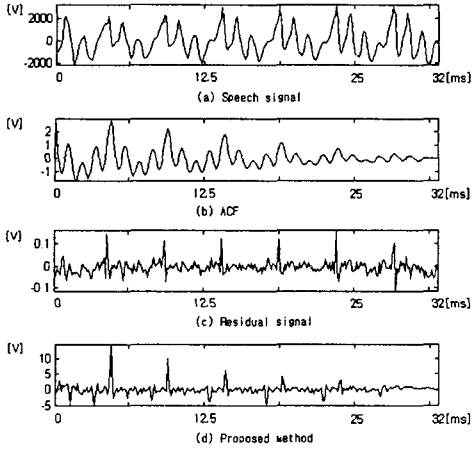


그림 6. 피치검출 성능비교(유성음구간)

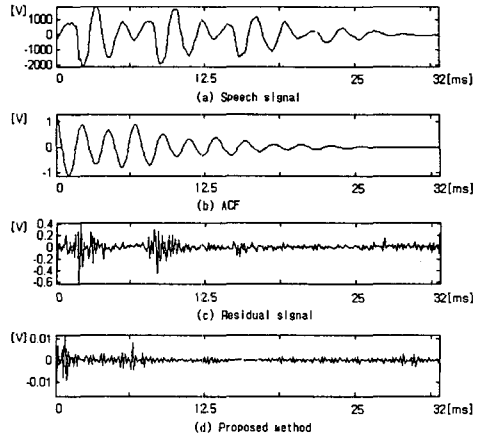


그림 7. 피치검출 성능비교(전이구간)

그림 6(a)와 7(a)는 전체 음성신호에서 한 프레임에 대한 음성신호를 나타내었으며 그림 6(b)와 7(b)는 원 음성신호에 대한 기존의 방법인 ACF를 한 파형이다. 그림 6(c)와 7(c)는 평탄화된 하모닉스 스펙트럼을 IFFT를 수행하여 구해진 잔차신호이다. 그림 6(d)와 7(d)는 구해진 잔차신호를 ACF한 것이며 이것은 그림 6(c)와 7(c)보다 정확하게 원 음성신호의 피치주기에 일치하는 피크위치를 보여주고 있다. 그림에서 알 수 있듯이 기존의 ACF방법보다 제안한 알고리즘의 피치검출 방법이 성능이 우수하다는 것을 알 수 있으며 정확하게 피치주기를 찾기 위해 구해진 잔차신호를 ACF를 수행하여 계산하였다.

그림 8과 그림 9는 본 논문에서 제안한 피치검출 성능을 평가하기 위하여 발성 1에 대하여 ACF, LPC, Cepstrum과 비교하여 피치변화도를 나타내었다.

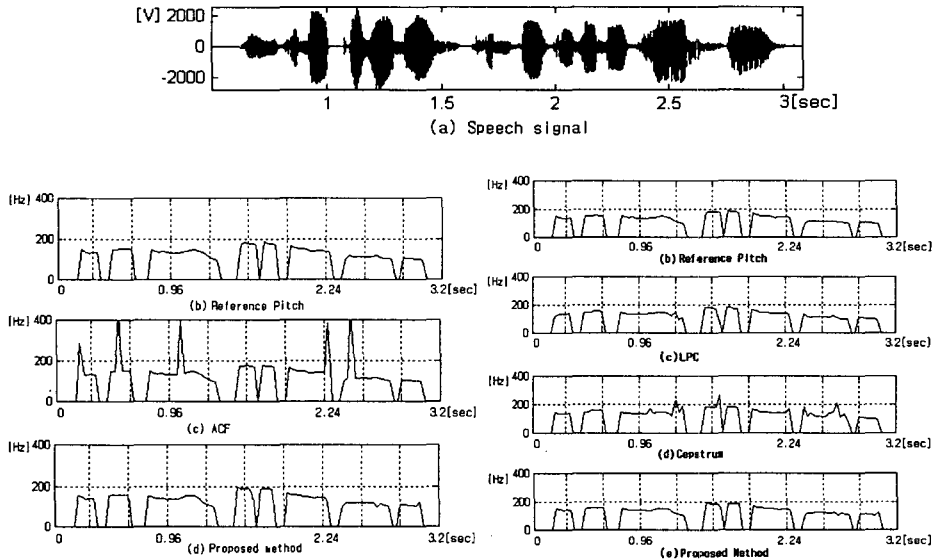


그림 8. ACF에 대한 발성1의 피치변화도

그림 9. LPC/Cepstrum에 대한 발성1의 피치변화도



전체 음성신호의 각 프레임에 대하여 기준피치를 정한 다음 제안한 방법과 ACF, LPC, Cepstrum과의 프레임별 피치변화도를 나타내었다. 여기에서 LPC 분석은 LPC 역필터링을 하여 잔차신호를 구하여 ACF를 통하여 피치를 계산하였으며 Cepstrum은 분석결과 성도성분 즉 포먼트 성분을 분리하여 원 음성신호의 스펙트럼과의 대수차에 의하여 평탄화된 스펙트럼을 구하여 피치를 계산하였다. 그림에서 알 수 있듯이 제안한 방법이 ACF, LPC, Cepstrum보다 성능이 우수하였다.

표 1은 각 발성에 대하여 제안한 방법과 ACF와의 gross error rate를 나타내었다. gross error rate에서 음성신호는 clean speech이며 전체 프레임 에러율은 한 프레임의 평균피치에서 1 ms의 오차가 발생하면 피치에러로 구하여 전체 프레임에 대하여 계산하여 구하였다. 즉 전체 프레임 에러율은 프레임별 피치에러를 찾아서 계산하였다. 표 1에서 알 수 있듯이 제안한 방법이 기존의 피치검출 방법인 ACF보다 성능이 향상된 것을 알 수 있다. 즉 ACF는 유성음구간에서 비교적 정확하였으나 전이구간에서는 많은 오차가 발생하므로 제안한 방법이 유성음은 물론이고 전이구간에서 좋은 성능을 나타냄을 알 수 있다. 표 1에서의 ACF의 에러는 대부분 전이구간에서 발생하므로 gross error rate가 1.91% 감소한 것은 전이구간에서 발생한 에러율이다.

또한 표 2는 발성 1, 2에 대하여 제안한 방법과 LPC, Cepstrum과의 gross error rate를 나타내었다. 제안한 방법은 잔차신호로부터 프레임 단위로 하나의 피치를 구하여 기준피치와 비교하여 에러율을 측정하였다. 기존의 피치검출 방법인 LPC, Cepstrum보다 성능이 향상되었음을 알 수 있으며 gross error rate가 LPC는 1.54%, Cepstrum은 1.93%가 감소하였다.

표 1. ACF에 대한 gross error rate

| 음성<br>시료 | gross error rate(%) |          |
|----------|---------------------|----------|
|          | ACF                 | Proposed |
| 발성1      | 2.91                | 0.58     |
| 발성2      | 0.80                | 0.55     |
| 발성3      | 2.91                | 0.557    |
| 발성4      | 3.39                | 0.66     |
| 평균       | 2.50                | 0.59     |

표 2. LPC/Cepstrum에 대한 gross error rate

| 음성<br>시료 | gross error rate(%) |          |          |
|----------|---------------------|----------|----------|
|          | LPC                 | Cepstrum | Proposed |
| 발성1      | 2.35                | 2.65     | 0.58     |
| 발성2      | 1.85                | 2.34     | 0.55     |
| 평균       | 2.1                 | 2.49     | 0.56     |

## 5. 결 론

음성신호처리분야에서 피치를 정확히 검출하면 음성인식에 화자에 따른 영향을 줄일 수 있기 때문에 인식의 정확도를 높일 수 있고, 음성합성시에 자연성과 개성을 유지하거나 쉽게 변경할 수 있다. 또한 분석시 피치에 동기시켜 분석하면 성문의 영향이 제거된 정확한 성도 파라미터를 얻을 수 있게 된다. 대부분의 피치검색기는 잡음이 있는 환경, 대역폭이 제한된 경우, 주파수의 진폭특성이 열화된 환경, 주파수의 위상특성이 변질되는 경우 등과 같은 복잡한 상황에서는 신뢰할 만한 결과를 얻지 못하고 있다.

본 논문에서 제안한 새로운 피치검출방법은 기존의 방법들에 비하여 시간-주파수 변환에 따른 계산량 증가와 알고리즘의 복잡성이 있지만 피치성능을 비교하면 기존의 방법들보다 성능면에서는 향상되었다. 특히 음성이 변하는 전이구간에서 피치를 더 잘 찾을 수 있었다. 기존의 피치검출 알고리즘인 ACF, LPC, Cestrum과의 비교에서 gross error rate가 ACF는 1.91%, LPC는 1.54%, Cestrum은 1.93%이 감소하였다. 결과적으로 본 논문에서의 주된 목적은 기존의 피치 검출 알고리즘에 관하여 처리시간과 복잡성보다는 피치검출 성능을 개선하고 향상하는데 목적을 두었다. 따라서 시간영역과 주파수 영역의 변환으로 인한 피치주기 검색 에러율을 최소화하였고 음성이 변하는 전이구간에서의 피치검출 방법을 개선하였다. 따라서 향후과제로는 피치검색 시간과 계산량 감소, 알고리즘의 단순화를 강구해야 하며, 또한 음성 합성에서 중요한 피치시점 검출, 피치변경, 지속시간 변경에 적용할 수 있는 피치검출기를 구현해야 할 것이다.

### 참 고 문 헌

- [1] Gold, B. & L. R. Rabiner. 1969. "Parallel processing technique for estimation pitch periods of speech in the time domain." *J. Acoust. Soc. Am.*, Vol. 46, No. 2., 442-448.
- [2] Rabiner, L. R. 1977. "On the use of autocorrelation analysis for pitch detection." *J. Acoust., Speech, Signal Processing*, Vol. ASSP-25, 24-33.
- [3] Ross, M. J., H. L. Lshaffer, A. Cohen, R. Freudberg & H. J. Manley. 1974. "Average magnitude difference function pitch extractor." *IEEE Trans. Acoust. Speech and Signal Proc.*, Vol. ASSP-22, 535-562.
- [4] Miller, R. L. "Performance characteristics of an experimental harmonic identification pitch extraction (HIPEX) systems." *J. Acoust. Soc. Amer.*, Vol. 43, 1593-1601.
- [5] Rabiner, L. R. & R. W. Schafer. 1978. *Digital Processing of Speech Signals*. Englewood Cliffs, NJ: Prentice-Hall.
- [6] 이을재, 강동규, 배명진, 안수길. 1990. "Second order decimation logic에 의한 시간영역 피치검출기에 관한 연구." *대한전자공학회 하계학술발표*, 제13권 제1호, 656-659.
- [7] 배명진, 이상호. 1998. *디지털 음성분석*. 동영출판사.
- [8] 김종국, 박원, 배명진. 2001. "스펙트럼상에서 하모닉스 파형의 피크-피팅을 이용한 정확한 피치 검출에 관한 연구." *한국통신학회 하계종합학술발표회 논문집*, 제23권 제 2호, 1308-1311.
- [9] 김종국, 최호진, 배명진. 2001. "시간-주파수영역에서의 새로운 피치검출방법." *한국통신학회 신호처리합동학술대회 논문집*, 제14권 제1호, 805-808.

접수일자: 2003. 4. 26.

게재결정: 2003. 6. 4.

▲ 김종국

서울특별시 동작구 상도5동 1-1 (우: 156-743)  
숭실대학교 정보통신공학과 음성통신연구실  
Tel: +82-2-824-0906  
E-mail: kokjk@hanmail.net

▲ 조왕래

서울특별시 동작구 상도5동 1-1 (우: 156-743)  
숭실대학교 정보통신공학과 음성통신연구실  
Tel: +82-2-824-0906  
E-mail: wrjo@unitel.co.kr

▲ 배명진

서울특별시 동작구 상도5동 1-1 (우: 156-743)  
숭실대학교 정보통신공학과 음성통신연구실  
Tel: +82-2-820-0902  
E-mail: mjbae@ssu.ac.kr