

음성신호의 기본주파수 검출

배 명 진 이 상 효 권 흥 렬
승실대학교 정보통신공학과

On a Detection for the Fundamental Frequency of Speech Signals

Myungjin Bae Sanghyo Lee Hongyeol Jeon
Soongsil University

ABSTRACT

A pitch detector is an essential component in a variety of speech processing systems. Besides providing valuable insights into the nature of the excitation source for speech production, the pitch contour of an utterance is useful for recognizing speakers, aids-to-the handicapped, and is required in almost all speech analysis-synthesis system. Because of the importance of the pitch detection, a wide variety algorithms for pitch detection have been proposed in speech processing literature.

Thus, in this paper we discuss the various type of pitch detection algorithms which have been proposed until now. Then we provide the performance measurements for seven pitch detection algorithms(three time-domain, two frequency-domain and two time-frequency hybrid domain).

I. 서론

피치검출기는 여러 가지 음성 처리 시스템에 필수적인 요소이다. 피치검출기의 성능은 음성합성에 여기원의 특성을 나타내어 음질의 자연성을 좌우하고, 검출된 피치변화도는 화자인식용 및 발음 장애자를 위한 보조시스템용 파라미터로 널리 적용되는 등, 거의 모든 음성 분석-합성(보코더) 시스템에 널리 쓰이고 있다. 따라서, 이러한 피치검출에 대한 다양한 알고리즘들이 제안되어 왔다. 그렇지만 대부분의 피치검출기는 잡음이 있는 환경, 대역폭이 제한된 경우, 주파수의 진폭특성이 열화된 환경, 주파수의 위상특성이 변질되는 경우 등과 같은 복잡한 상황에서는 신뢰할만한 결과를 얻지 못하고 있다.

따라서 본 논문에서는 피치주기 검출에 관계되는 기존의 연구방법들을 검토하고, 피치검출시에 대두되는 제반 문제점들을 분석하여 그 한계성을 보완할 수 있는 새로운 검출법들에 관하여 우리 연구실에서 제안하였던 연구중인 내용에 대해 알아보고자 한다.

먼저, 제 II 절에서는 피치검출에 대한 일반적인 문제점들과 고려할 사항에 대해서 알아보고, 피치검출의 주된 처리영역에 따라 시간 영역법은 제 III 절에서, 주파수 영역 피치검출법에

대해서는 제 IV 절에서, 그리고, 시간-주파수 혼성영역법에 대해서는 제 V 절에서 다루었다. 이때 각 피치검출기들이 갖는 문제점들과 그 해결법에 대해 설명하면서 향후의 연구방향에 대해 고려했고 논문을 결론짓게 된다.

II. 피치검출시의 고려할 사항

음성신호의 피치주기를 정확하고 신뢰성 있게 측정하는 것은 다음과 같은 여러 가지 이유 때문에 매우 어렵다. 첫번째 이유는 성문의 여기파형이 완전히 주기적인 파형이 아니라는 것이다. 비록 주기적인 파형의 주기를 찾는 것이 간단하다고 할지라도 한 주기내에 있는 파형의 세부구조와 함께 변화하는 음성 파형의 주기를 측정하는 것은 어려운 작업이다.

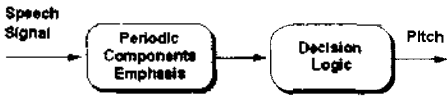
피치주기를 측정하는데 있어서 두번째 어려움은 성도와 성문간의 상호작용 때문이다. 어떤 경우에는 성도의 포먼트가 성문파형의 구조를 완전히 바꿀 수 있기 때문에 피치주기를 검출하기가 어렵게 된다. 일반적인 상호작용은 조음기관이 빠르게 변화하는 경우이며 피치검출시에는 방해요소로 나타나게 된다.

세번째 문제는 유성음구간 동안 각 피치주기의 정확한 시작과 끝을 정의하기 어렵다는 것이다. 피치주기의 정확한 시작과 끝의 위치를 선택하는 것은 임의적일 수 있기 때문이다. 예를 들면, 음성파형에 근거하여 주기의 시작과 끝을 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. 음성파형의 주기성분은 영교차 단위로 검출하면 포먼트, 잡음, 그리고, 음성파형의 어떤 DC 레벨에 민감하게 되는 반면 봉우리 측정을 통해 주기성을 검출하면 포먼트 구조에 민감하게 된다.

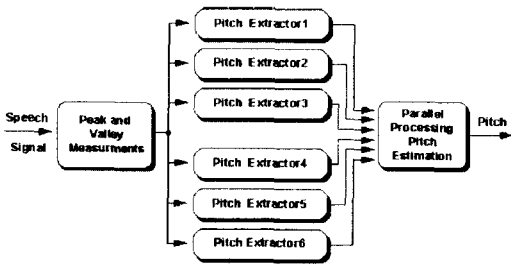
피치를 검출할 때 나타나는 네번째 문제점은 무성음과 낮은 레벨의 유성음을 구별해야 하는 것이다. 많은 경우에 무성음 구간과 레벨이 낮은 유성음 구간의 변화는 별로 뚜렷하지 않기 때문에 정확한 위치를 찾기가 어렵다.

다섯번째 문제점은 전화시스템을 통하여 전송되는 음성의 피치검출 문제에 직면할 때 부가적인 복잡성이 발생하게 된다. 음성에 대한 전화시스템의 영합은 선형, 비선형처리를 포함하면서 음성신호에 잡음이 추가되는 경우이다. 전화시스템을 선형여파기로 간주하면 기본주파수와 고조파들의 성분을 약하게 만들 수 있는 대역통과 여파기로 볼 수 있다. 이 때문에 전화음질에 대해 신호의 주기성을 검출하기가 더 어렵게 된다. 음성신호에 대한 전화시스템의 비선형특성은 전송시스템의 채널에 따라 좌우된다.

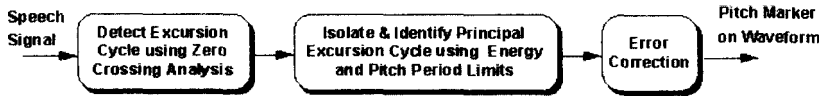
Time Domain Methods



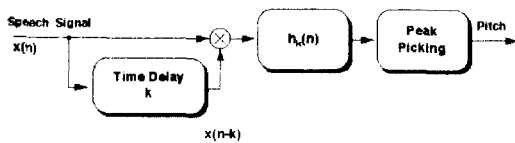
[그림 1] 전형적인 시간영역 피치검출법



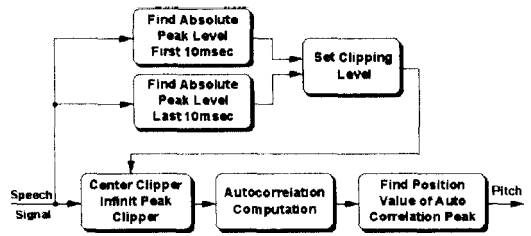
[그림 2] 병렬처리에 의한 피치검출법



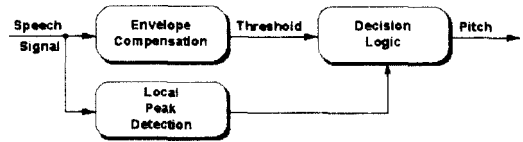
[그림 6] Data Reduction에 의한 피치검출법



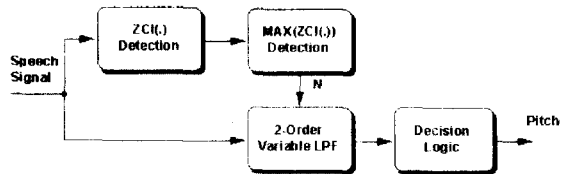
[그림 3] AMDF에 의한 피치검출법



[그림 4] 3-레벨 중앙 윈도우링 자기상관관계에 의한 피치검출



[그림 5] Harmonic Matching에 의한 피치검출법



[그림 7] 가변 지연용과 여러피기에 의한 피치검출

III. 시간영역 피치검출법

따라서 음성신호의 기본주파수 검출시에 고려할 사항은 다음과 같다[G1].

- 1) 화자의 남녀노소에 영향을 받지 않아야 한다.
- 2) 화자의 감정이나 개성을 잘 나타내도록 추정범위가 넓어야 한다.
- 3) 잡음이나 진송채널의 배경하에서도 실용적인 결과를 얻을 수 있어야 한다.
- 4) 음성신호의 내용에 무관한 검출능력을 가져야 한다. 즉, 유성자음으로 시작하는 구간이나 파열음이 연결된 유성음구간이나, 비음이 끼여 있는 유성음 구간에서도 검출이 잘 되어야 한다.

1939년 R. Dudley에 의해 Vocoder의 이론이 제안된 후부터 피치검출에 관한 연구도 주된 처리영역에 따라 시간영역, 주파수영역, 시간-주파수 혼성영역으로 분리해서 다루어지고 있다.

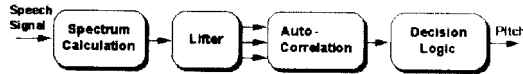
시간영역 피치검출법은 시간영역에서 직접 처리하기 때문에 다른 영역으로의 변환과정이 불필요하며 합, 차, 비교논리 등에 의해서만 처리가 가능해진다. 피치의 범위는 보통 2.5에서 25msec로 알려져 있고 음성을 8 kHz로 표본화하여도 그 범위가 20~200 표본 사이에 나타나기 때문에 시간영역에서의 검출은 분해능이 높다는 장점이 있다. 보통 음성의 분석은 프레임 단위로 처리되지만 파열음의 경우, 음소의 변화가 한 프레임안에서도 일어날 수 있다. 따라서 시간영역에서의 피치추정은 음소변화에 따른 특성을 검출할 수도 있다. 그렇지만, 음소의 종류나 결합위치에 따라 음소파형의 외곽선이 변화하는데, 이 변화를 측정하기 어렵기 때문에 이러한 파형부분에서 피치검출 오차가 많이 발생한다. 또한 특성전송 채널에 통과된 경우나 배경잡음이 부가된 경우에는 이와 영향이 커져서 피치검출 오차가 높아지는 단점이 있다.

시간영역에서의 피치검출법들은 일반적으로 음성파형의 주기성을 강조하고 나서 결정논리에 의해 주기성을 검출하게 된다. 주기성 강조는 성도의 공명현상에 나타나는 포먼트들의 영향을 제거시키고 여기원의 기본주파수만을 강조하는데 그 목적이

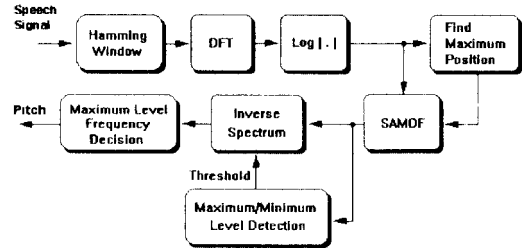
Frequency Domain Methods



[그림 8] 전형적인 주파수영역 피치검출법



[그림 9] Spectrum Autocorrelation에 의한 피치검출



[그림 10] SAMDF를 이용한 피치검출

있다. 따라서 주기성 강도가 복잡하면 할수록 결정능리가 간단해진다. 그림1은 전형적인 시간영역 피치검출법에 대한 블록도이다.

피치검출의 결정논리법에는 강조된 음성파형의 기본주기를 실험적인 문턱값이나 무게치를 적용하여 결정하는 것이 보통이다. 그 외에도 위상이나 진폭에 따라 뱃가지의 기준패턴을 만들어 입력된 음성의 파형에서 추출한 패턴과의 유사도를 측정하여 주기성을 결정하기도 한다. 결정능리가 실험적인 문턱값에 의존하면 검출오차가 확률적으로 커지기 때문에 경우의 수에 따라서 결정문턱값을 신중히 교차해야 하는 복잡성이 따른다. 따라서 최근에는 Fuzzy나 신경망의 결정논리를 적용한 사례가 계속해서 연구중에 있다. 시간영역 피치검출법으로는 그림 2, 3, 4, 5, 6에서 제시한 것처럼 Parallel Processing [T1,1969], Autocorrelation [T2,1977], AMDF [T3,1974], Data Reduction [T4], Harmonic Matching [G1]법 등이 종래의 기법으로 알려져 있다.

본 연구실에서 제안된 방법으로는 음성신호의 생성모델을 근거로 하여 주기성만을 검출하는 면적비교법[T5, 1985], 성문의 주파수특성만 통과시키는 역비례형 여파기법[T6], 원파형과 임의로 반복시킨 파형을 비교하여 피치를 검출하는 방법[T7, 1988], 개선된 Decimation Logic을 이용하는 피치검출법[T8, 1990], 창함수 적용에 따른 스펙트럼 누설현상을 이용한 피치검출법[T9, 1990], 가변 저역여파기를 이용한 피치검출법[T15, 1994] 등이 있다.

면적비교법[T5, 1985]은 영-레벨 이상의 면적들을 서로 비교함으로써 한 피치구간에서 성문의 영향이 지배적인 구간을 구하여 유성음 피치를 결정하는 방법이다. 이러한 면적비교법은 간단한 프로세서에 의해 실시간으로 처리할 수 있으며, 시간영역에서 처리하기 때문에 분해능이 우수하고 위상특성이 보존되는 장점이 있으며 결정능리가 간단하다.

스펙트럼 누설현상을 이용한 피치검출기는 창함수를 적용할 때 발생하는 스펙트럼 누설현상을 이용하는 방법이며 [T10,1990], 여러 창함수들 중에서 누설현상이 가장 심한 방형창함수를 적용하였다. 이 피치검출법은 프레임 내에서 음소가 현이 중이거나 피치주기가 변화하고 있어도 평균티 피치주기를 검출하고 프레임내의 모든 주기 성분을 참조하기 때문에 0db 정도의 낮은 신호대 잡음비(SNR)에서도 피치검출을 우수하게 수행하는 장점을 갖는다.

가변 저역여파기를 이용한 피치검출기는 그림 7에서처럼 네 프레임마다 G-peak의 폭을 강조하고 포먼트의 성분비가 다르기 때문에 이를 이용한 피치검출기는 G-peak검출과 동시에 피치를 검출하므로 피치시점이 동시에 얻어지어서 피치동기 분석

에 용이하고, 저역여파기를 사용하기 때문에 잡음에도 강인하며 실시간처리가 가능하다는 장점을 가진다.

앞으로는 계산시간의 절감을 위해 시간영역법을 선택하기보다는 음소의 변화를 추종하는 점에 착안해야 한다. 따라서 주기성 강조나 결정논리가 이것을 극복할 수 있도록 새로운 방법을 계속 연구중에 있다.

IV. 주파수영역 피치검출법

주파수영역에서의 피치검출은 그 처리과정상 주파수영역으로의 변환이 필요해서 계산시간이 많이 걸리고 복잡하지만, DSP칩의 시판으로 인해 계산시간의 부담이 제거되어 최근에 다시 관심의 대상이 되고 있다. 주파수영역 피치검출기는 보통 음성스펙트럼에서 고조파들의 간격을 측정하여 피치를 구한다.

고조파를 파라미터로 하여 피치를 검출하는 방법으로는 그림9,10에 제시한 것처럼 Harmonic Spectrum[F1, 1969], Peak-Valley[F2, 1978], Comb Filtering[F4, 1974], Spectrum Autocorrelation[F5, 1987] 등이 제안되어져 있다.

No1에 의해 제안된 스펙트럼 고조파 음의 합법[F1]은 다음 식으로 고조파를 강조한다.

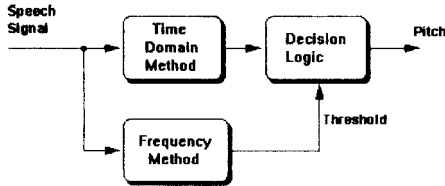
$$P(f) = 2 \times \sum_{n=1}^{\infty} \log |X(f, e^{j2\pi n T})|$$

이 식에 의해 고조파의 봉우리는 봉우리끼리 골은 골끼리 더해져서 고조파가 강조된다. 이 방법은 음성이 진화선로들 통해 음 때 기본 고조파가 감쇠되거나 잡음이 많은 경우에도 얼마간 성공적이지만 포먼트에 의해 평탄하지 못한 스펙트럼의 영향을 많이 받아서 피치결정시에 배주기(doubling)나 반주기(halving)가 많이 발생한다.

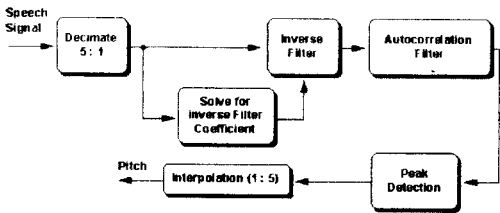
Seneff에 의해 제안된 고조파의 봉우리-골 측정법[F2]은 1100Hz이하의 스펙트럼상에 존재하는 고조파에 대해 복잡한 봉우리-골 검출[F3]에 대한 알고리즘을 제안하였다. 그렇지만 이러한 방법들은 배경잡음이 없는 경우에만 우수한 결과가 얻어진다.

Labat에 의해 제안된 스펙트럼의 자기상관 합수법은 스펙트럼을 한 신호로 취급하여 특정 리프터들을 통과시키는 방법[F5]이다. 고조파가 기본주파수의 정수배로 나타난다고 볼 수 있기 때문에, 이 방법은 스펙트럼을 리프터에 통과시켜서 최대값이 찾아지는 리프터의 위치를 피치주기로 결정하게 된다. 이러한 방법 역시 포먼트의 영향을 많이 받고 여성이나 어린이 화자인 경

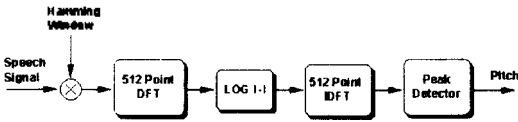
Hybrid Domain Methods



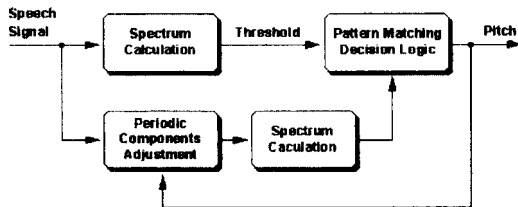
[그림 11] 전형적인 시간-주파수 혼성형 피치검출법



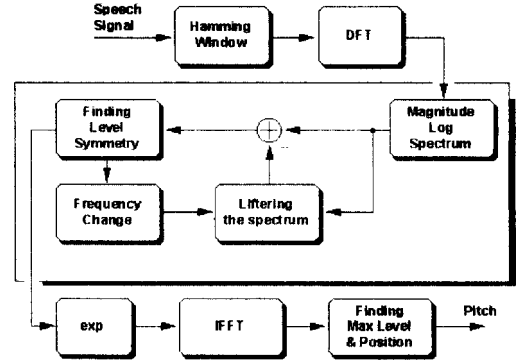
[그림 12] SIFT법에 의한 피치검출



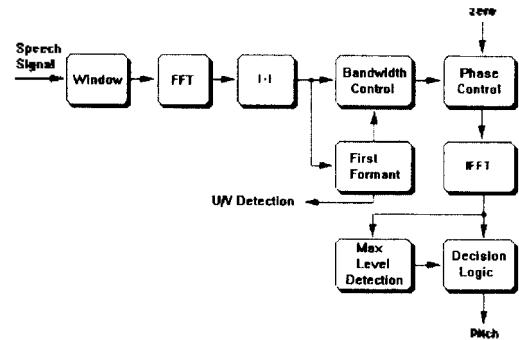
[그림 13] 웨스트럼에 의한 피치검출



[그림 14] 스펙트럼 비교에 의한 피치검출



[그림 15] 고조파의 평판화에 의한 피치검출



[그림 16] 위상조절에 의한 피치검출

우의 스펙트럼은 단순한 정현파의 스펙트럼처럼 단일 펄스형 스펙트럼이 되기 때문에 이러한 방법으로는 피치를 결정하기 어렵게 된다.

이상의 문제점들을 해결하기 위한 방법의 하나로 본 연구실에서는 고조파 류의 합 방법처럼 반복적인 봉우리와 골을 갖는 고조파들을 강조하여 결정논리가 용이하면서도 기본고조파의 최대나 최소값이 구해지도록 하는 함수인 SAMDF를 다음과 같이 제안하였다[TG, 1986].

$$SAMDF(.) = \sum_{k=1}^N |S_p(k) - S_p(k-d)|, \quad d=1, 2, \dots, F_m$$

이 식에 의해 스펙트럼의 고조파들이 주기성을 유지할 때마다 최소값을 형성하게 되고, 이에 역수를 취하면 기본 주파수들이 강조되고, 첫번째 고조파가 더욱 강조되며 그 첫번째 고조파거리의 간격이 바로 기본주파수가 된다. 이렇게 구해진 기본주파수에 역수를 취하여 피치를 검출하였다.

주파수영역에서 기본주파수를 검출하면 분해능이 떨어

지게 되는데 본 연구실에서는 분해능을 향상시키는 방법에 대해 계속 연구중에 있다.

V. 시간-주파수영역 피치검출법

피치검출시에는 각 영역별로 처리가 용이하고 특징적인 장점을 취할 수 있다. 시간-주파수 혼성법에서는 시간영역에서의 장점과 주파수영역에서의 특징을 병합특성을 이용하여 처리하는 방법이며, 웨스트럼[H1, 1969], SIFT 법[H2, 1972], LPC inverse filtering and AMDF[H3, 1977] 등이 제안되어져 있다.

시간-주파수 혼성영역법은 시간영역에서의 계산시간 절감과 검출된 피치의 정밀성, 그리고 주파수영역에서 평균화 효과에 의해 배경잡음이나 음소간의 천이, 음소의 변동 등에서 나타나는 배주기와 반주기 문제를 극복할 수 있다. 이러한 장점들 때문에 혼성형 피치 검출기의 개발이 절실히 요구되고 있지만 아직도 강력한 혼성형 알고리즘이 제안되지 않고 있는 실정이다.

본 연구실에서 혼성형 피치검출기에 대한 새로운 알고리즘을 창안하기 위해 처리과정을 두부분으로 나누어 처리하고 있다. 먼저 음성파형의 주된 봉우리를 검출하는 과정은 시간영역에서 처리하게된다. 이 때 예비 피치들의 집합이 얻어지게 되며, 이것은 주파수영역에서 구한 결과에 의해 검증되고 실제의 피치를 결정하게 된다. 이렇게 하면 시간영역법이나 주파수영역법이 갖는 장점을 취할 수 있다. 이러한 결과로 시간영역의 변화와 주파수 영역 관계를 이용하여 두 영역의 장점을 취하는 스펙트럼 비교에 의한 기본주파수 검출법[H4, 1987]과 음성 고조파의 평탄화에 의한 피치검출법[H5, 1990]을 제안하였다.

본 연구실에서 시간-주파수 혼성으로 제안한 다른 방법은 시간-주파수 변환특성을 이용하는 방법이다. 이 방법은 주파수영역에서 음성신호의 위상 특성을 동위상으로 조절하여 시간영역으로 변환하게 되면, 시간영역에서 음성신호는 위상이 시작되는 영시간 근방에 에너지가 집중되는 특징이 나타난다. 이 경우 대부분의 포먼트 변화는 음레벨에 나타나게 되어, 2.5msec 이상에서 주된 양레벨을 검출하게 되면 피치가 검출된다. 이러한 원리로 피치를 검출하는 방법을 음성신호의 위상동기에 의한 피치검출법[H6, 1992]이라 명명하였다.

최근에는 신호처리 전용칩인 DSP칩의 발달로 인해 계산시간이 절감됨에 따라 새로운 관심을 가지고 기존 방법들의 문제점과 장점들을 분석하면서 새로운 혼성형 알고리즘을 계속해서 연구중에 있다.

VI. 결론

피치검출기는 여러 가지 음성처리 시스템에 필수적인 요소이다. 음성합성시에 여기원의 특징을 나타내는 것 외에도 피치변화도는 화자인식, 장애자를 위한 보조시스템에 유용하며, 거의 모든 음성 분석-합성(보코더) 시스템에 쓰이고 있다. 이러한 피치검출의 필요성 때문에 여러 가지 피치검출 알고리즘이 제안되어 왔다.

본 논문에서는 피치주기 검출에 관계되는 기존의 연구 방법들에 대해 다시 한번 검토해보고, 기존의 피치검출법들이 갖고 있는 문제점들을 해결하기 위해 본 연구실에서 제안하였거나 연구중인 내용에 대해서도 간단히 서술하였다.

피치검출법에 대한 앞으로의 연구방향을 영역별로 살펴 보면 먼저, 시간영역에서는 음성의 원이구간을 검출할 수 있는 진폭변화의 보상과 피막인 방법을 적용하며 시간영역법상의 문제점을 해결하는 연구가 진행되어야 하며, 주파수영역에서는 분해능을 높이는 연구가 계속되어야 하고, 시간-주파수 혼성영역에서는 스펙트럼에서 나타나는 포먼트의 영향과 시간영역의 진폭변화를 보상할 수 있는 연구가 계속 진행되어야 한다.

REFERENCE

[General]

[G1] L. R. Rabiner and R. W. Schafer, "Digital Processing of Speech Signals", Englewood Cliffs, NJ:Prentice-Hall, 1978.
 [G2] R. W. Schafer and L. R. Rabiner, "Digital Representation of Speech Signals", Proc. IEEE, Vol. 63, pp. 662-667, Apr. 1975.
 [G3] L. R. Rabiner, M. J. Cheng, A. E. Rosenberg, and C. A. McGonegal, "A Comparative Performance Study of Several Pitch Detection Algorithms", IEEE Trans. Acoust.,

Speech, Signal processing, Vol. ASSP-24, pp. 399-417, Oct. 1976.

[G4] S. E. Stearns and R. A. David, "Signal Processing Algorithms", Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1988.
 [G5] Myungjin BAE and Souguil ANN, "Speech-Silence Detection using the Major Zero Crossing Interval Energy", Tencon87, Seoul IEEE Conf., pp. 1278-1281. Oct. 1987.
 [G6] J. D. Markel and A. H. Gray, Jr., "Liner Prediction of Speech Signals", Springer-Verlag, 1976.
 [G7] I. H. Witten, "Principles of computer Speech", Academic Press, 1982.
 [G8] G. Bristow, "Electronic Speech Synthesis", McGraw-Hill, 1984.
 [G9] R. D. Mori and C. Y. Suen, "New systems and architectures for Automatic Speech Recognition and Synthesis", Springer-Verlag, 1985.
 [G10] S. Saito and K. Nakata, "Fundamentals of speech Signal Processing", Academic Press, 1985.
 [G11] P. E. Papamichalis, "Practical Approaches To Speech Coding", Prentice-Hall, 1987.

[Time domain]

[T1] B. Gold and L. R. Rabiner, "Parallel Processing Technique for Estimation Pitch Periods of Speech in the Time Domain", J. Acoust. Soc. Am., Vol. 46, No. 2., pp. 442-448, August, 1969.
 [T2] L. R. Rabiner, "On the use of Autocorrelation Analysis for Pitch Detection", J. Acoust., Speech, Signal processing, Vol. ASSP-25, pp. 24-33, Feb. 1977.
 [T3] M. J. Ross, H. L. Lshaffer, A. Cohen, R. Freudberg, and H. J. Manley, "Average Magnitude Difference Function Pitch Extractor", IEEE Trans. Acoust. Speech and Signal Proc., Vol. ASSP 22, pp. 535-562, Oct. 1974.
 [T4] R. L. Miller, "Performance Characteristics of an Experimental Harmonic Identification Pitch Extraction(HIPEX) Systems", J. Acoust. Soc. Amer., Vol. 43, pp. 1593-1601, Dec. 1970.
 [T5] Myungjin BAE and Souguil ANN, "The High Speed Pitch Extraction of Speech Signals using the Area Comparison Method", KIEE, Vol. 22. No. 2, pp. 101-105. Feb. 1985.
 [T6] Myungjin BAE and Souguil ANN, "Inverse Rate Type Filtering for the Pitch Extraction", J. of the Acou. Soc of Korea, Vol. 5, No. 3. pp. 5-12, Sept. 1986.
 [T7] Myungjin BAE and Dongjin SHIN and Souguil ANN, "The Pitch Extraction of Voiced Speech by the Comparison between the Original and the Repeated Partial Waveform", J., Acoust. Soc. Korea. Vol. 7, No. 5. Nov. 1988.
 [T8] 이윤재, 강동규, 배명진, 안수길, "Second Order Decimation Logic에 의한 시간영역 피치검출기에 관한 연구", 대한전자공학회 하계 학술발표, Vol. 13, No. 1, pp. 656-659, 1990.
 [T9] D. KANG, M. BAE, S. ANN, "On Extraction the Center Pitch using The Spectrum Leakage Phenomenon for Noise corrupted Speech Signals", J. Acoust., Soc., Korea, Vol. 10, No. 1, pp. 37-46, Feb. 1991.

[T10] N.C. Geckinki, D. Yavuz, "Algorithm for Pitch Extraction using Zero Crossing Interval Sequence", IEEE Trans. Acoust. Speech and Signal Proc., Vol. ASSP-25, No. 6, pp. Dec. 1977.

[T11] J.D. Wise, J.R. Caproo, and T. W. Parks, "Maximum Likelihood Pitch Estimation", IEEE Trans. Acoust., Speech, Signal processing, Vol. ASSP-24, pp. 418-423, Oct. 1976.

[T12] D. Terzopoulos, "Co-occurrence Analysis of Speech Waveforms", IEEE Trans. Acoust. Speech, Signal processing, Vol. ASSP-33, No. 1, Feb. 1985.

[T13] K.K. Paliwal and P.V.S. Rao, "A Synthesis-based Method for Pitch Extraction", Speech Commun., Vol. 2, pp. 37-45, May. 1983.

[T14] H. Duifhuis, L. F. Willems, and R. J. sluyter, "Measurement of Pitch in Speech : An Implementation of goldstein's Theory of Pitch Perception", J. Acoust. Soc. Amer., Vol. 71, pp. 1568-1580, June., 1982.

[T15] H. KEUM, G.R. BAEK, M.J. BAE, " Pitch Detection Using Variable LPF ", WESTPRAC V Proc., Vol. 2, pp. 963-970, 1994.

[Frequency domain]

[F1] A. M. Noll, "Pitch Determination of Human Speech by the Harmonic Product Spectrum, the harmonic sums spectrum and a maximum likelihood estimate", In Proc. Symp. Comput. Processing Commun., Apr. 1969, pp. 779-798.

[F2] S. Seneff, "Real Time Pitch Detector", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-26, pp. 358-365, Aug. 1978.

[F3] T. V. Screenivas, P. V. S. Rao, " Pitch Extraction from Corrupted Harmonics of the Power Spectrum", J. Acoust. Soc. Amer., Vol. 65, pp. 223-228, Jan. 1979.

[F4] J. A. Moorer, "The Optimum Comb Method of Pitch Period Analysis of Continuous Digitized Speech",

IEEE. Trans. Acoust., Speech Signal Processing, Vol. ASSP-22, pp. 330-338, Oct. 1974.

[F5] M. Lahat, R. j. Niederjohn, and D. A. Krubsac, "A Spectral Autocorrelation Method for Measurement of the Fundamental Frequency for Noise-corrupted Speech", IEEE. Trans., Acoust., Speech, Signal processing, Vol. ASSP-35, No. 6, June. 1987.

[F6] 배명진, 박찬수, 안수길, "배경잡음하에서 주파수영역 피치 검출에 관한 연구-스펙트럼 AMDF에 의한 제1포먼트 영향 제거법", 한국음향학회, 한국음향학회 논문집, 10권, 4호, pp. 12-18, 8월 1991.

[Time-Frequency domain]

[H1] A. M. Noll, "Cepstrum Pitch Determination." J. Acous. Soc. Am., Vol. 41, pp. 293-309, Feb. 1967.

[H2] J. D. Markel, "The Sift Algorithm for Fundamental Frequency Estimation", IEEE Trans.on Audio and Electro acoustics, Vol. Au-20, No. 5, pp. 367-377, Dec. 1972.

[H3] C. K. Un and S. C. Yang, "A Pitch Extraction Algorithm Based on LPC inverse Filtering and AMDF", IEEE Trans. acoust., Speech, Signal Processing, Vol. ASSP-25, pp. 565-572, Dec. 1977.

[H4] Myunjin BAE and Souguil ANN, "Fundamental Frequency Estimation of Noise Corrupted Speech Signals using the Spectrum Comparison", J. Acoust. Soc. Korea, Vol. 8, No. 3, pp. 57-64, June. 1987.

[H5] M. BAE and S. Ann, "On the Time-Frequency hybrid Technique for Detecting the Pitch noise Corrupted Speech Signals(Time Domain Processing)", J. Acoust. Soc. Korea, Vol. 9, No. 1, 1990.

[H6] 이윤재, 박찬수, 배명진, 안수길, "시간-주파수 혼성형 피치 검출에 관한 결정논리", 대한전자공학회 추계학술발표논문집, Vol. 13, No. 2, pp. -. 1990.

[H7] B. LEE, M. BAE, S. ANN, "Pitch Detection by Synchronizing the Phase of Noise Corrupted Speech Signals", J. Acoust. Soc. Korea, Vol. 11, No. 1E, pp. 42-49, July 1992.