



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

工學碩士 學位論文

음성신호의 피치주기 검출을 위한
변형된 AMDF에 관한 연구



2007年 2月

釜慶大學校 大學院

制御計測工學科

徐玄垂

工學碩士 學位論文

음성신호의 피치주기 검출을 위한
변형된 AMDF에 관한 연구

指導教授 金 南 虎

이 論文을 工學碩士 學位論文으로 提出함



2007年 2月

釜慶大學校 大學院

制御計測工學科

徐玄垂

徐玄垂의 工學碩士 學位論文을 認准함

2006年 12月



主 審 工學博士 黃 龍 淵 (印)

委 員 工學博士 安 永 珠 (印)

委 員 工學博士 金 南 虎 (印)

목 차

Abstract	ii
제1장 서 론	1
제2장 음성신호 분석	3
2.1 음성의 조음기관	3
2.2 음성의 분류	5
2.3 유·무성음의 구분	6
2.4 피치주기	8
2.5 음성 발생 모델	9
2.6 음성신호의 스펙트로그램	10
제3장 AMDF 알고리즘	12
3.1 Basic-AMDF	12
3.2 기존의 AMDF	13
3.3 추가의 알고리즘	16
제4장 변형된 AMDF	18
제5장 시뮬레이션 및 결과	23
5.1 최적의 α 와 β 설정을 위한 시뮬레이션 및 결과	23
5.2 AWGN이 중첩된 정현파 신호에 대한 시뮬레이션 및 결과	28
5.3 무소음 환경에서의 음성에 대한 시뮬레이션 및 결과	31
5.4 소음환경에서의 음성에 대한 시뮬레이션 및 결과	36
제6장 결 론	39
참고문헌	40

*A Study on Modified AMDF for Pitch Period Detection
of Speech Signal*

Hyun-Soo Seo

*Department of Control and Instrumentation Engineering
Graduate School of
Pukyong National University*

Abstract

As a lot of researches on the speech signal processing are performed due to the recent rapid development of the information-communication technology, pitch period which means the fundamental frequency of the speech signal is recognized as an essential and fundamental element for applications in various speech signal processing fields such as speech recognition, speaker identification, and speech synthesis, etc. Since it is very sensitive to the responses of human auditory sense, it can be used in creating diverse timbre through artificial transformation and pitch period serves as a useful parameter in analyzing the characteristics of a speaker. Also the accurate acquisition of the pitch period allows to easily extract useful information from the speech signals regardless of the speaker conditions, and in particular the accurate extraction and reconstruction of pitch period during speech synthesis and speech

coding play decisive roles on the sound quality. Therefore, many algorithms in time domain and frequency domain for the pitch period detection of the speech signal were proposed. In particular, though AMDF(average magnitude difference function) that can detect accurate pitch period with small amount of operations is generally being used. But if the differences between local minimum valley points are not large, it would be needed an additional algorithm for the detection of pitch period and the first local minimum valley point, the establishment of several thresholds for optimal detection conditions is not easy.

So in this paper, it is proposed modified-AMDF signifying adaptable rotation transformation by analyzing the decreasing characteristics of basic-AMDF within the frame, and in order to remove the detection errors of abnormal portion where repetition periods of voiced sound are not clear, we made it possible to detect accurate pitch period by setting up thresholds and determining normal portion. Also, since the speech signals can not remain free from the noises in a real environment where they are processed, in order to simulate the proposed method, AWGN(additive white gaussian noise) generated according to arbitrarily configured SNR(signal-to-noise ratio) and diverse actual noises in subway etc were superposed onto origin signal, and they were used as signals for test.

Lastly, simulation results proved the excellence of the proposed method in this paper by comparing it with other various existing methods.

제1장 서론

인간의 음성은 사고의 결과 및 감정 등을 포함한 다양한 정보를 가지고 있으며 의사전달을 위한 가장 기본적인 수단으로 사용되어 왔을 뿐만 아니라, 최근 인터넷과 휴대폰을 비롯한 멀티미디어 기술의 발전으로 인해 그 활용가치가 높아졌다. 따라서 현재 음성신호처리를 위한 여러 이론적 연구 뿐만 아니라 휴먼인터페이스시스템 개발이 활발히 진행되고 있으며, 이러한 성과가 현재와 미래에서의 인간의 생활에 지대한 영향을 미쳐, 보다 편리한 생활을 영위할 것이라는 예측은 지극히 당연하다.

그리고 음성신호의 기본 주파수를 의미하는 피치주기는 음성 인식, 화자 식별, 음성 합성 등과 같은 여러 음성신호처리 분야에서의 응용을 위해 필수적이고 기본적인 요소로 인식된다. 피치주기는 화자의 특성분석에 유용하게 이용되는 파라미터이며, 피치주기의 변화는 인간의 청각에 민감하게 반응하므로 인위적인 변환에 의해 다양한 음색을 만드는 것이 가능하다. 또한, 피치주기의 정확한 획득은 화자의 상태와는 상관없이 음성신호로부터 유용한 정보의 추출을 가능하게 하며, 특히 음성 합성과 음성 코딩에서 피치주기의 정확한 검출과 복원은 음질에 결정적인 역할을 한다. 따라서 음성신호의 피치주기 검출을 위한 시간영역과 주파수영역에서의 많은 알고리즘이 제안되었다[1]-[9].

시간영역의 피치주기 검출알고리즘은 음성신호에 대한 파형의 주기성을 강조하는 방법이 주로 이용되며, ACF(autocorrelation function), AMDF(average magnitude difference function) 등이 있다. 시간영역에서의 알고리즘은 음성신호를 분석하기 위한 시간영역에서 주파수영역으로의 변환이 불필요하며, 합과 차, 그리고 비교논리 등과 같은 비교적 간단한 연산을 사용함으로써 처리속도 측면에서 우수한 특성을 가진다. 그러나 잡음이 중첩된 음성이나 분석을 위한 프레임 내의 음소가 급격한 레벨변화를 일으킬 경우, 피치주기 검출을 위한 알고리즘이 복잡해지는 단점이 발생한다.

주파수영역에서 피치주기를 검출하기 위한 알고리즘은 획득된 음성 스펙트럼의 고조파 간격을 이용하여 기본주파수를 측정하는 방법이 사용되며, 대표적으로 Cepstrum법, 고조파분석법 등이 있다. 이러한 알고리즘은 주파

수대역별로 신호의 분리가 가능하므로 음소의 천이와 변동 그리고 소음과 같은 배경잡음에 대하여 그 영향을 적게 받는다. 그러나 처리과정에서 주파수영역으로의 변환과정이 필요하기 때문에 상대적으로 복잡한 연산이 요구되며, 기본주파수의 정밀한 측정을 위한 FFT(fast fourier transform) 포인트 수를 증가시킬 경우 처리시간이 길어지고 변화특성이 둔화된다.

이에 따라, 실시간 시스템에서는 처리속도가 빠른 시간영역 알고리즘이 많이 응용되어 사용된다. 특히, 적은 연산량과 정확한 피치주기 검출이 가능한 AMDF가 널리 적용되지만, 국부 최소 계곡점(local minimum valley point)들 간의 크기차가 크지 않음으로써, 피치주기 검출에 사용되는 최초 국부 최소 계곡점 검출을 위한 추가의 알고리즘이 필요하며, 최적의 검출 조건을 위한 여러 임계값의 설정이 용이하지 않다[10].

따라서, 본 논문에서는 basic-AMDF의 감소특성을 분석하여 적응적인 회전변환을 이용하는 변형된 AMDF를 제안하였으며, 음성음의 반복주기가 명확히 나타나지 않는 시작구간과 같은 비정상구간의 검출오류를 감소시키기 위해, 임계값을 설정하여 정상구간을 판단함으로써 정확한 피치주기 검출이 가능하도록 하였다. 또한 음성신호가 처리되는 실제 환경에서는 소음과 같은 잡음으로부터 자유로울 수 없으므로, 지하철 등을 비롯한 다양한 실제 소음 및 여러 값들로 설정한 SNR(signal-to-noise ratio)에 따라 발생된 AWGN(additive white gaussian noise)을 정현파 신호에 중첩시켜 테스트를 위한 신호로 사용하였다. 최종적으로 시뮬레이션 결과를 통해 기존의 여러 방법들과 비교하여 본 논문에서 제안한 방법의 우수성을 입증하였다.

제2장 음성신호 분석

2.1 음성의 조음기관

인간이 자의적으로 발생시킬 수 있는 신호인 음성신호의 발생은 첫 단계로써 허파(lung)에서 공기가 압축된다[11]. 이렇게 압축된 공기는 신호의 Gain, 즉 크기를 결정하는 역할을 하며, 공기의 흐름이 이루어지는 기관(trachea)을 통과해서 인간의 음성에서 고음과 저음 그리고 음성파형의 주기를 결정하는 성대(vocal cords)를 지나간다. 이에 해당하는 기관들의 명칭을 그림 1에 도식하였다.

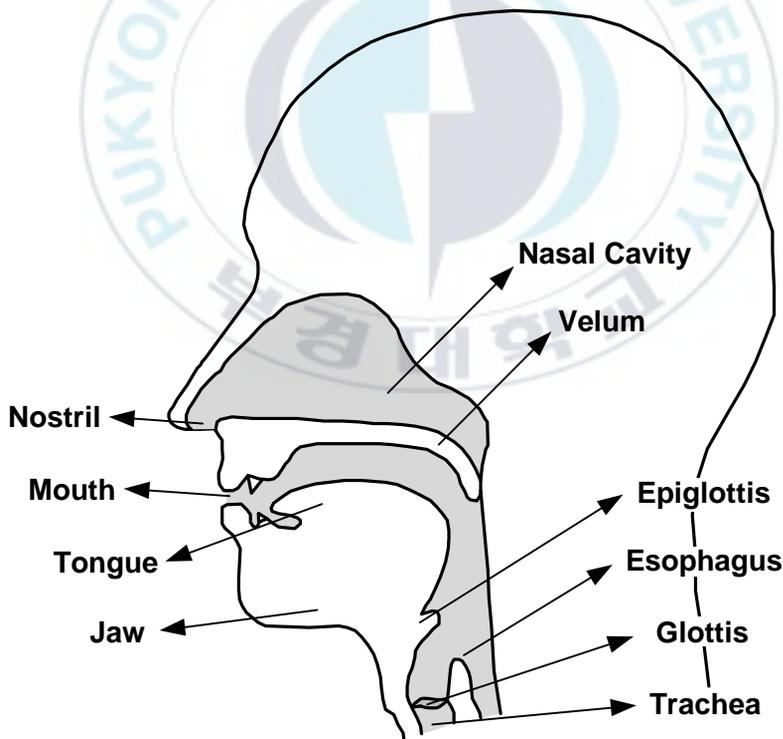


Fig. 1. The human vocal apparatus.

그 다음, 연속적으로 식도(esophagus)와 후두개(epiglottis) 및 연구개(velum)를 거쳐 혀(tongue)에서 인간의 발성이 완성된다. 그리고 이러한 과정의 블록 다이어그램을 그림 2에 나타내었다.

여기서 성도(vocal tract)는 성문(glottis)에서 입술까지의 구간으로 정의되며 인간의 발음 및 음성과형의 모양을 결정하는데, 남성의 경우 성도의 평균길이는 대략 17[cm] 정도이다. 또한 그림 1에서 볼 수 있는 연구개는 성도의 횡단면을 결정지으며, 그 넓이는 일반적으로 20[cm²] 정도이다[12].

또한 연구개부터 콧구멍까지의 구간을 의미하는 비도(nasal tract)는 비음을 발생시키며, 허파에 의해 공급되어진 공기는 연구개의 닫힘에 의해 코로 전송된다.

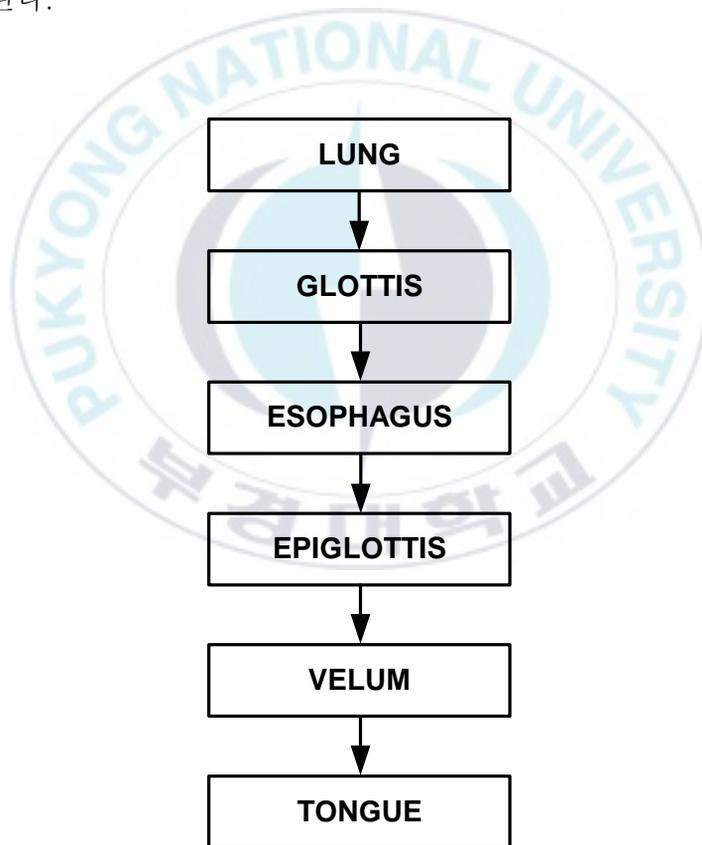
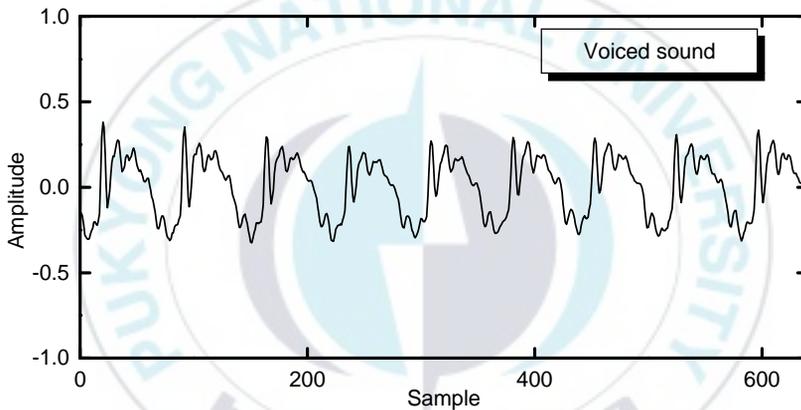


Fig. 2. Block diagram of speech-production.

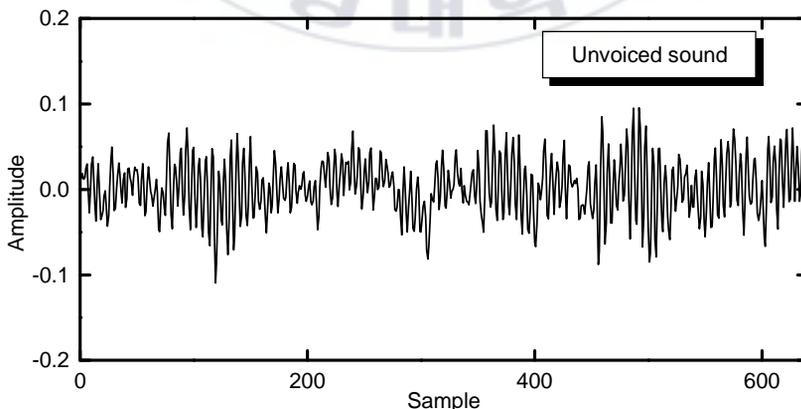
2.2 음성의 분류

일반적으로 인간에 의해 발생되는 음성은 크게 유성음(voiced sound)과 무성음(unvoiced sound)으로 나누어지며, 이러한 유성음과 무성음을 분류하기 위한 기준은 성대의 진동 유·무이다.

허파에서 압축된 공기가 성문을 지나갈 때, 성문이 닫혀 있는 상태에서 두 개의 얇은 막의 성대가 진동하여 발생하는 음을 유성음으로 정의한다. 즉, 성대에 의해 발생한 공진이 성대를 준주기적으로 움직이게 하여 발생되는 것이다. 따라서, 이렇게 반복된 규칙성이 인간에게 감지되어 진행되고



(a) A voiced sound /a/



(b) A unvoiced sound /k/

Fig. 3. Examples for voiced sound and unvoiced sound.

있는 음성신호의 의미를 인지하게 된다. 이러한 유성음의 대표적인 경우로는 /아/, /이/, /우/, /에/, /오/와 같은 모음이 있으며, 실제로 발성한 경우 목 부분이 떨리는 것이 확인된다.

무성음은 열려진 성대에 의해, 진동하지 않으면서 발생되는 음성으로 정의된다. 즉, 무성음은 빠르게 공기가 압축되어 공진이 발생하기가 힘들며, 또한 성대의 일부분을 좁히면서 재압축하여 난기류를 만들어내는 소리를 말한다. 일반적으로 자음은 성대의 떨림이 없으며, 상대적으로 많은 공기가 입에서 나온다.

그리고 그림 3은 유성음과 무성음의 예시이다. 그림 3(a)의 유성음은 성대의 떨림을 알아 볼 수 있을 정도로 파형이 주기성을 띄며, 3(b)의 무성음은 잡음과 같은 형태로 나타난다.

2.3 유·무성음의 구분

음성신호처리와 관련한 연구에서 유성음과 무성음을 구분하는 것은 기본적인 중요한 과제로 인식된다. 그리고 일반적으로 음성신호의 에너지가 유·무성음의 구분에 주로 이용된다. 이러한 에너지는 음성신호의 샘플값을 제곱한 후, 합연산을 통해 얻어지며, 식 (1)과 같이 정의된다.

$$E(m) = \sum_{n=-\infty}^{\infty} s(n)^2 \quad (1)$$

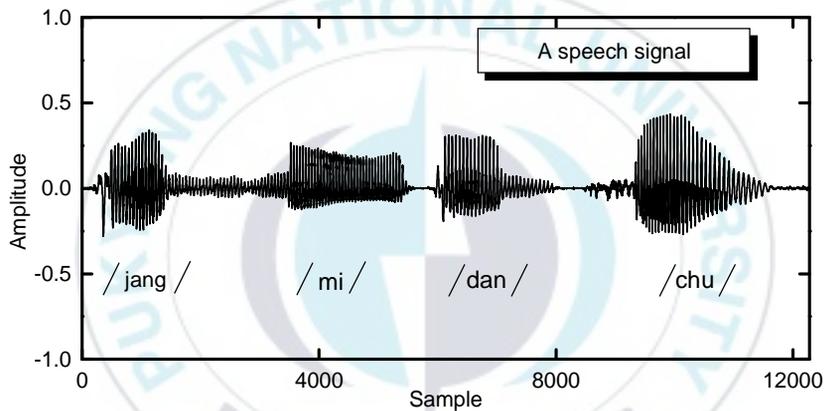
식 (1)에서 보인 음성신호의 에너지는 전체 신호에 대한 값이며, 처리를 위한 N 개의 샘플길이를 갖는 분석 프레임으로 음성신호를 나눈 후, 각 분석 프레임 내의 샘플값을 제곱하여 합연산을 할 경우 단구간 에너지로 구해진다.

그러므로 획득된 입력신호와 창함수를 연산한 후, $n=m-N+1$ 에서 $n=m$ 의 제곱에 대해 합연산으로 구한 단구간 에너지 $E_s(m)$ 은 식 (2)와 같이 표현하며, 여기서 $h(n) = w^2(n)$ 이다.

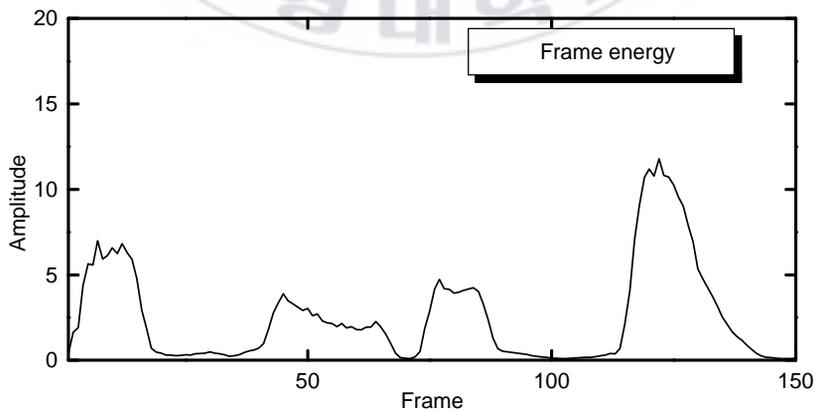
$$\begin{aligned}
 E_s(m) &= \sum_{n=m-N+1}^m \{s(n)w(m-n)\}^2 \\
 &= \sum_{n=m-N+1}^m s^2(n)h(m-n)
 \end{aligned}
 \tag{2}$$

식 (2)의 결과로써 얻어지는 단구간 에너지를 이용하여 시간적으로 변화하는 음성신호에서 성대의 진동에 의한 유성음과 그렇지 않은 무성음을 구분한다.

그림 4에서 단구간 에너지의 예를 나타내었다. 그림 4(a)는 /장미단추/로



(a) A speech signal pronounced as /jangmidanchu/



(b) Frame energy for speech signal

Fig. 4. Short-term frame energy.

발성되고 표본화 주파수 8[kHz]로 획득된 테스트 음성신호이며, 그림 4(b)는 10[ms]를 한 프레임으로 하고, 분석 프레임으로 네 프레임을 사용하여, 분석 프레임별로 구한 단구간 에너지이다.

그림 4(b)에서 확인할 수 있듯이, 발생되는 각 부분에 대해서 에너지가 프레임별로 변화하지만, 유성음에 해당하는 모음의 프레임 에너지가 무성음에 비해서 상대적으로 큰 값을 나타낸다.

또한, 영교차율은 분석을 위한 프레임 내에서 이산신호가 영점축을 기준으로 교차하는 횟수이다. 즉, 획득된 음성신호에서 이웃한 샘플값의 부호가 서로 다른 경우에 발생하는데, 이는 음성 분석, 음성 인식 등에 유용한 특징 벡터로써 활용된다. 영교차율은 화자의 성량에 대한 의존도가 낮은 특성을 나타내며, 유성음은 낮은 주파수에 집중되기 때문에 영교차율에서 작은 값을 가진다. 또한, 묵음에서는 주위환경에 따라 다소 차이가 있으나, 대체로 무성음보다는 큰 값을 갖는다. 그러므로 영교차율은 유·무성음의 구분 및 묵음 추출에 사용된다.

그리고 N 개의 샘플을 갖는 음성신호에 대해서, $n=m$ 에서 끝나는 프레임의 단구간 영교차율과 시그넘 함수 sgn 에 대한 정의식은 다음과 같다.

$$Z(m) = \frac{1}{N} \sum_{n=m-N+1}^m \frac{|sgn\{s(n)\}| - |sgn\{s(n-1)\}|}{2} w(m-n) \quad (3)$$

$$sgn|s(n)| = \begin{cases} +1, & \text{if } s(n) \geq 0 \\ -1, & \text{otherwise} \end{cases} \quad (4)$$

2.4 피치주기

음성신호처리의 특징벡터로 사용되는 피치주기는 기본주파수를 의미한다. 기본주파수는 DC 성분을 제외한 주기성을 갖는 가장 낮은 주파수로 정의되며, 많은 주파수가 중첩된 음성신호에서의 가장 낮은 주파수, 즉 시

간영역에서 가장 시간간격이 넓은 피크값들의 주기성을 의미하고, 이러한 주기성은 앞에서 설명한 바와 같이 성대의 떨림에 의한 것이다[12].

인간은 피치주기의 변화에 매우 민감하게 반응하며, 자연성에 큰 영향을 미치므로 피치주기의 인위적인 변화에 의해 여러 가지 음색을 생성해 낼 수 있다. 또한, 피치주기의 정확한 획득으로 인해 화자의 상태와는 상관없이 음성신호로부터 유용한 정보의 추출이 용이하며, 특히 음성 합성과 음성 코딩에서 피치주기의 정확한 추출과 복원은 음질에 결정적인 역할을 한다. 또한 유성음에서 발생하는 피치주기를 이용하여 유성음과 무성음을 구분할 수 있다. 피치주기는 사람마다 차이가 있지만 일반적으로 남성의 경우 50~250[Hz], 여성의 경우 120~500[Hz]의 주파수를 가지며, 발성시의 감정이나 억양 등에 영향을 받는다. 그리고 그림 5에 피치주기의 물리적 의미를 도식하였다.

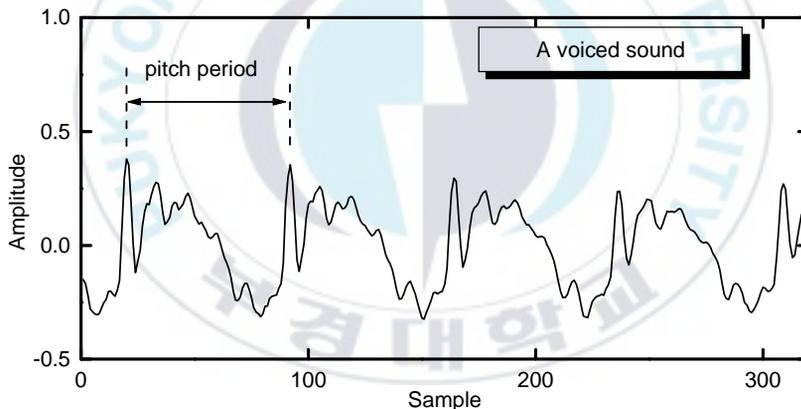


Fig. 5. Pitch period.

2.5 음성 발생 모델

인간의 음성 발생 시스템은 다음의 그림 6과 같이 모델링된다[12]. 음성 신호를 크게 유성음과 무성음으로 구분하고, 성대의 진동없이 발생하는 무성음의 경우는 랜덤 잡음으로, 성대의 떨림에 의해 발생하는 유성음의 경우는 주기성을 나타내므로 주기적인 임펄스 열로써 표현된다. 다음으로, 성

량에 해당하는 일정한 크기를 곱한 후 시간-가변 필터를 거쳐 최종적인 음성신호가 발생하는 것으로 모델링한다. 여기서, 유성음의 피치와 성도의 모델링을 위해 각각 임펄스 열과 시간-가변 필터가 사용된다.

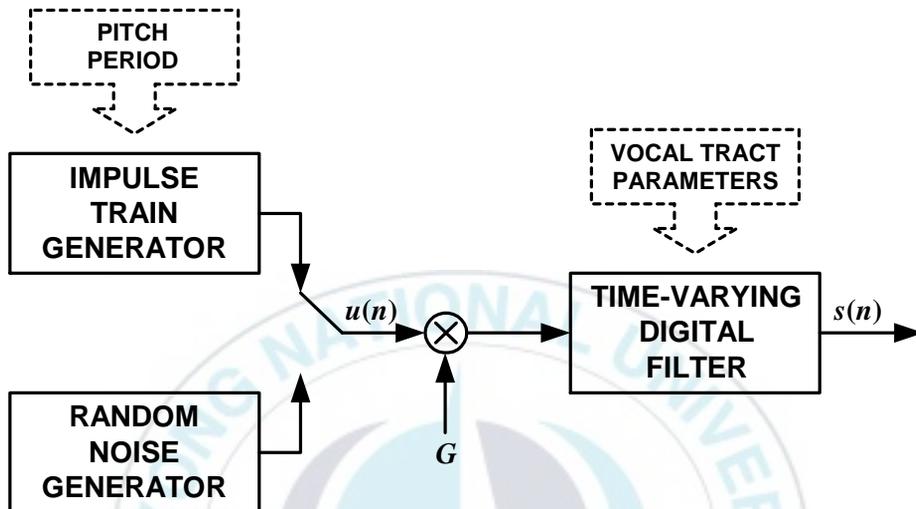
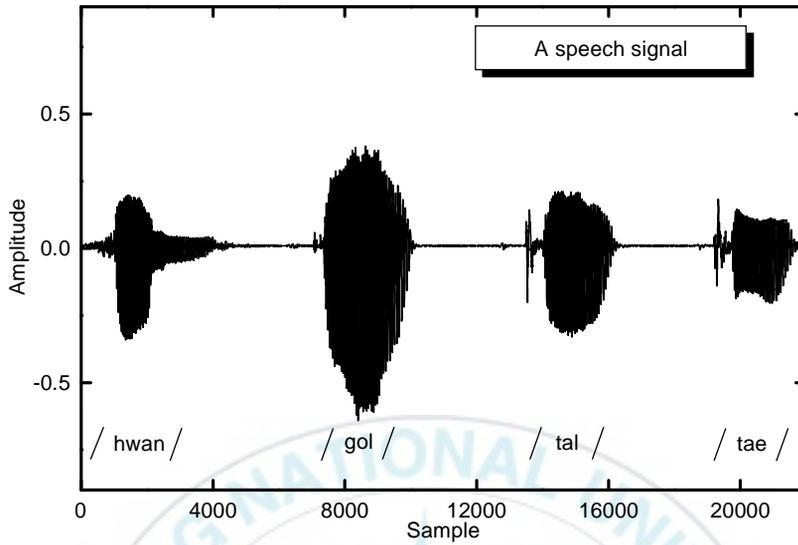


Fig. 6. Speech generator model.

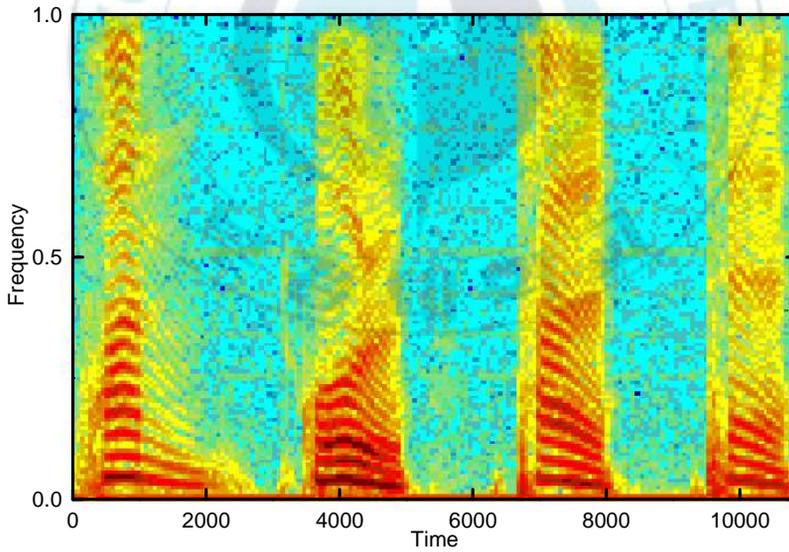
2.6 음성신호의 스펙트로그램

음성신호의 주파수 성분을 알 수 있는 스펙트로그램을 그림 7에 나타내었다. 그림 7(a)는 /환골탈태/로 발생된 음성신호이며, 그림 7(b)는 이러한 음성신호를 10[kHz]의 분해능으로 FFT한 스펙트로그램 결과이다. 그림으로부터, 음성신호는 대부분 유성음을 나타내는 저주파의 신호로 구성되며, 파열음과 같은 무성음의 고주파 성분이 중첩되었음을 확인할 수 있다. 주파수영역에서의 피치주기 검출알고리즘에서, 주파수분석은 기본주파수를 검출하는 데 이용된다.

스펙트로그램 분석으로부터, 음성신호는 여러 주파수성분이 복잡하게 중첩된 신호임을 알 수 있으며, 피치주기 검출알고리즘은 이러한 음성신호로부터 DC성분을 제외한 기본주파수를 정확히 검출해야 한다.



(a) A speech signal



(b) Spectrogram

Fig. 7. Spectrogram result about /hwangoltalae/.

제3장 AMDF 알고리즘

3.1 Basic-AMDF

외부에서 입력된 음성신호 $s(n)$ 은 길이 N 을 가지는 창함수 $w(n)$ 과의 연산을 통해 신호 $s_w(n)$ 이 된다. 이때, basic-AMDF $\gamma(l)$ 은 식 (5)와 같이 정의된다.

$$\gamma(l) = \sum_{n=1}^{N-l+1} |s_w(n+l-1) - s_w(n)|, \quad (5)$$

$l = 1, 2, \dots, N$

위의 식으로부터, 파라미터 l 이 1부터 프레임의 크기인 N 까지 변화함에 따라 $s_w(n)$ 과 $l-1$ 만큼 떨어진 $s_w(n+l-1)$ 과의 차에 대한 절대값을 임의의 구간 내에서 합연산을 수행한다. 그러므로 basic-AMDF에서는 주기성을 나타내는 부분에서 0에 가까운 최소값이 된다. 식 (5)로부터, $s_w(n)$ 이 유성음일 경우 $\gamma(l)$ 은 주기적인 특징을 나타내며, $\gamma(l)$ 의 이러한 특징을 이용하여 음성신호의 유성음과 무성음을 구분할 수 있다. 특히 basic-AMDF는 유성음의 피치주기를 결정하는 음성신호처리의 다양한 분야에 유용하게 적용된다.

일반적으로 유성음이 입력신호로 사용되었을 경우, 그림 8에서와 같이, basic-AMDF 알고리즘에 의해 국부 최소 계곡점들이 A, B, C 등에서 주기적으로 나타나며, 전체 표본에 대해서 최초의 계곡점과 시작점과의 시간간격을 피치주기로 계산한다. 따라서 그림 8에서는 최초 국부 최소 계곡점 A와 시작점 사이의 시간간격이 피치주기이다.

그러나 입력된 신호와 basic-AMDF를 연산한 후, 결과로써 나타난 최초 국부 최소 계곡점이 주어진 표본에 대해서 다른 국부 계곡점에 비해 크기가 작을 경우에는 전체 최소 계곡점을 피치주기로 선정하는 간단한 알고리즘이 적용되지만, 그렇지 않은 경우에는 최소점 검출이라는 간단한 알고리

즘 대신, 여러 임계값을 이용한 최초 국부 최소 계곡점 검출을 위한 추가의 알고리즘이 필요하다.

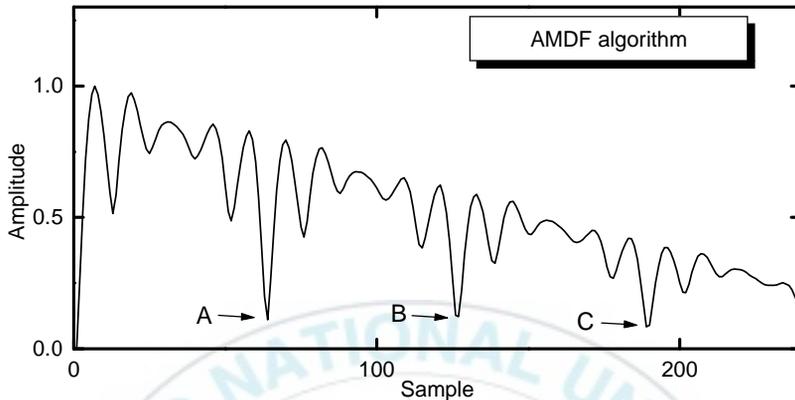


Fig. 8. AMDF in a voiced sound.

3.2 기존의 AMDF

Basic-AMDF의 특성을 개선시키기 위해, basic-AMDF를 응용한 선형가중 AMDF(W-AMDF)와 가변 길이 AMDF(LV-AMDF) 등이 소개되었다[12],[13].

여기서, W-AMDF는 식 (6)과 같이 정의된다.

$$\gamma_W(l) = \frac{1}{N-l+1} \sum_{n=1}^{N-l+1} |s_w(n+l-1) - s_w(n)|, \quad (6)$$

$$l = 1, 2, \dots, N$$

W-AMDF는 저음역의 잡음 환경에서 basic-AMDF에서 나타나는 피치주기 검출오류를 감소시키기 위해 제시되었다. 즉, 위의 식으로부터 스케일인자 $(N-l+1)^{-1}$ 은 basic-AMDF의 감소특성을 제거한다. 일반적으로 AMDF의 국부 최소값은 0에 가까운 값을 가지므로 기본주파수 즉, 피치주기를 검출하기 위해서, AMDF의 최초 국부 최소 계곡점으로 주어지는 적

절한 l 의 선택이 요구된다.

이때 주파수축 상에서의 기본주파수 f_0 의 위치가 표본화 주파수 F_s 에 의존하므로, l 은 F_s/f_0 의 관계로부터 획득된다. 이러한 W-AMDF를 그림 9에 나타내었다. 그림에서 볼 수 있듯이, $(N-l+1)^{-1}$ 의 영향으로 가중치가 부가되어 basic-AMDF의 감소특성이 제거된다.

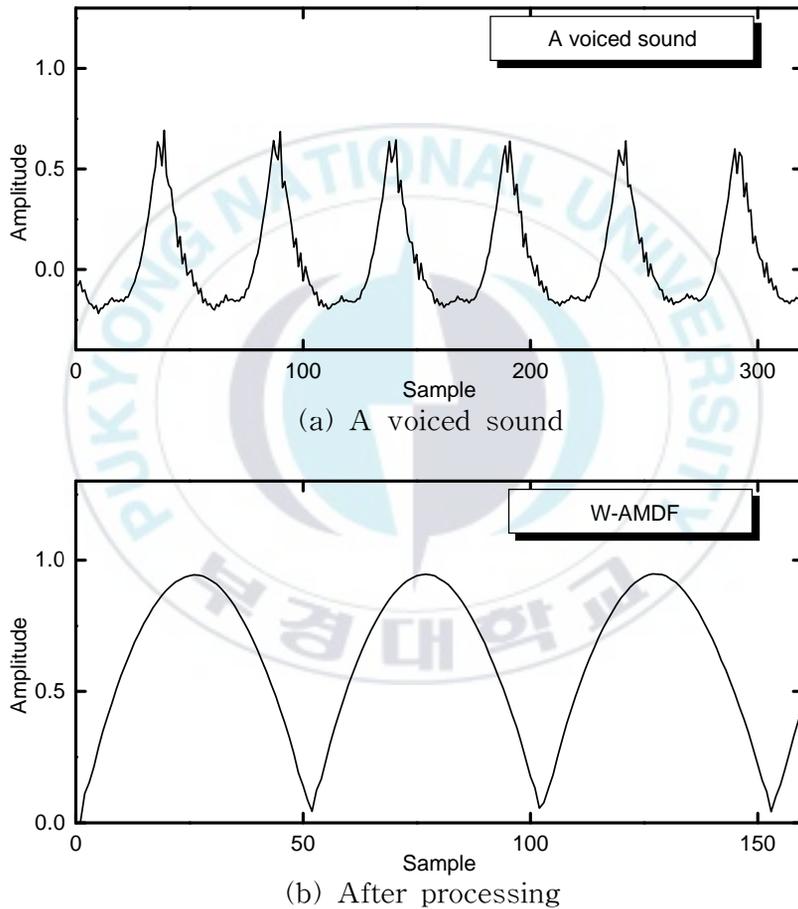


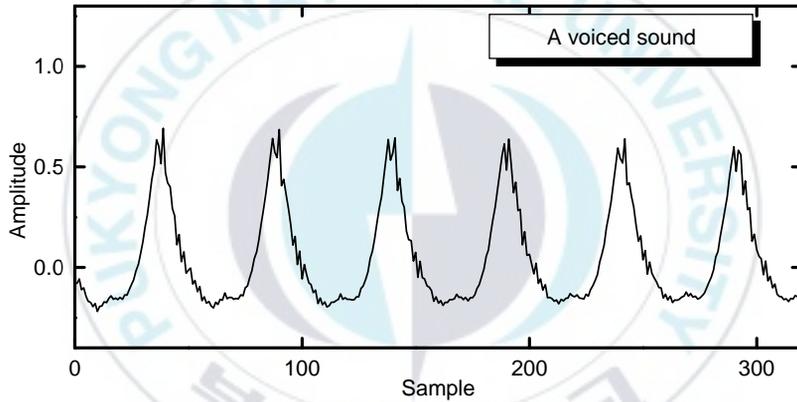
Fig. 9. W-AMDF.

그리고 LV-AMDF는 식 (7)과 같다.

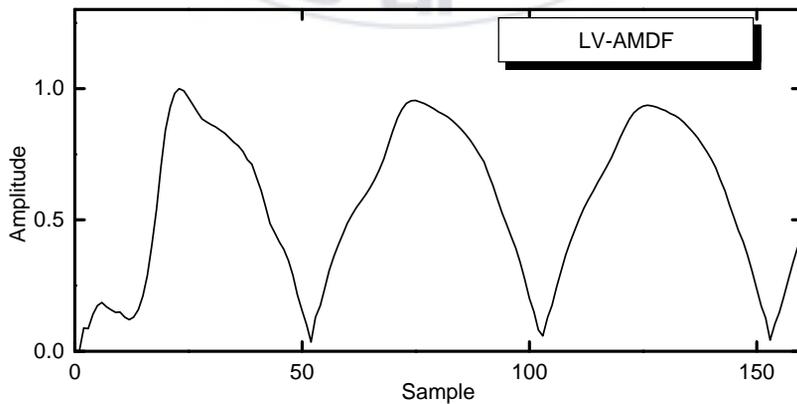
$$\gamma_{LV}(l) = \frac{1}{l} \sum_{n=1}^l |s_w(n+l-1) - s_w(n)|, \quad (7)$$

$$l = 1, 2, \dots, N$$

식 (7)로부터, LV-AMDF는 기존의 AMDF와 같은 주기성을 나타내며, 합에 대한 연산량이 변수 l 에 의해서 결정됨에 따라 basic-AMDF와 W-AMDF에 비해 적은 연산량을 가진다. 그림 10은 임의의 유성음에 LV-AMDF 알고리즘을 적용시킨 결과이다.



(a) A voiced sound



(b) After processing

Fig. 10. LV-AMDF.

3.3 추가의 알고리즘

앞에서 보인 바와 같이, 주기성을 띄는 입력 신호에 대해 기존의 AMDF 알고리즘에서는 국부 최소 계곡점들이 주기적으로 잘 나타난다. 그리고 정확한 피치주기 검출을 위한 중요한 요소는 최초 국부 최소 계곡점 검출이며, 일반적으로 AMDF 알고리즘에서는 전체 최소점을 검출하여 그것을 피치주기로 선택한다. 그러나, 최초 국부 최소 계곡점이 다른 계곡점들의 크기에 비해 큰 값을 나타내어 전체 최소점이 되지 않는 경우에는 피치주기 검출이 용이하지 않다. 또한 높은 주파수에 대해서는 국부 최소 계곡점들이 더욱 많이 존재하므로, 최초 국부 최소 계곡점이 전체 최소점이 될 확률은 감소한다. 따라서 이러한 AMDF 알고리즘의 단점을 개선하기 위해 다음과 같이, 최초 국부 최소 계곡점 검출을 위해 여러 임계값들을 적용한 추가의 알고리즘이 필요하다[10].

추가의 알고리즘에 대한 임계값들을 그림 11에 도식하였으며, 최초 국부 최소 계곡점이 다수의 국부 최소 계곡점 중에서 최소가 되지 않는 경우이다. 피치주기를 검출하기 위해, 먼저 최소값 Min을 검출하며, 다수의 국부 최소 계곡점을 구하기 위해 $Min+Th1$ 과 $Th2$ 를 적용하여 다수의 국부 최소 값을 구한다.

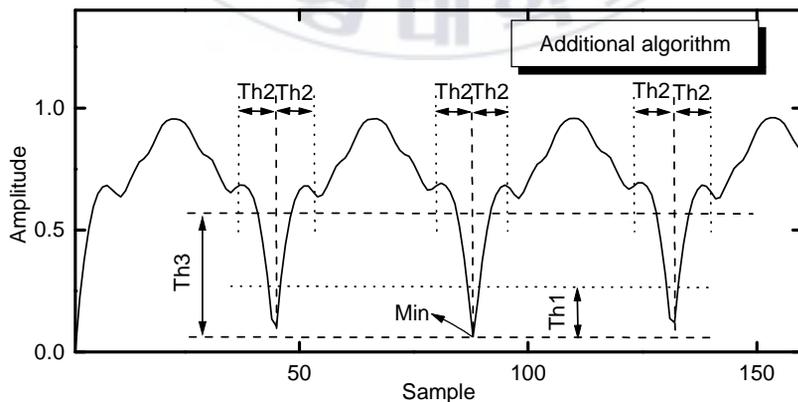


Fig. 11. Null picking of AMDF.

그리고 Th_2 을 적용하여 국부 계곡점의 깊이를 결정하였으며, 이러한 국부 계곡점의 깊이가 Th_3 보다 적은 경우 그 지점은 제거된다. 그 결과 국부 최소 계곡점이 한 개 존재할 경우에는 이 값이 피치주기가 되며, 다수인 경우에는 가장 좌측의 국부 최소 계곡점이 피치주기가 된다. 그림 12는 이러한 추가의 알고리즘에 대한 블록 다이어그램이다.

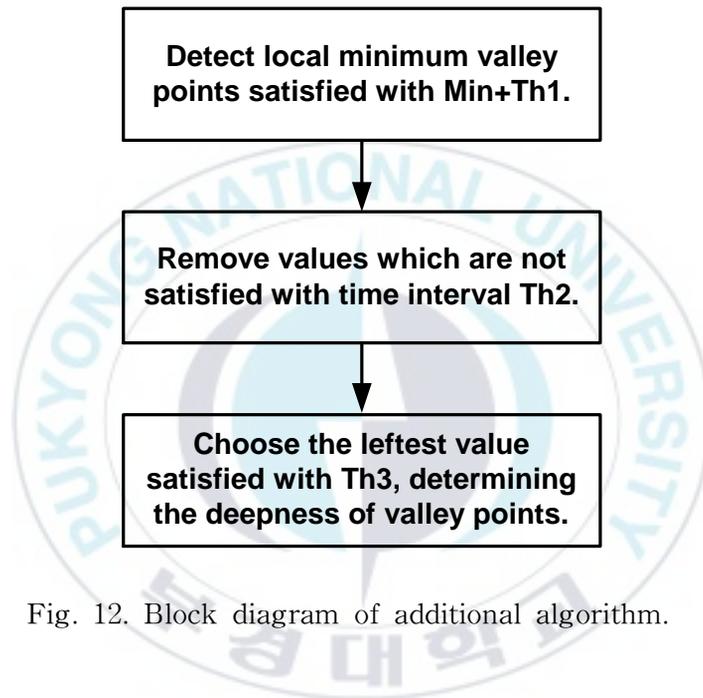


Fig. 12. Block diagram of additional algorithm.

제4장 변형된 AMDF

기존의 AMDF 알고리즘에서 피치주기 검출을 위한 최초 국부 최소 계곡점이 전체 최소 계곡점으로 되지 않는 경우에는 피치주기 검출이 용이하지 않다. 이러한 단점을 보완하기 위해, 기존의 AMDF 알고리즘에서는 최초 국부 최소 계곡점 검출을 위한 추가의 알고리즘을 적용해야 하며, 이러한 과정에서 최적의 임계값을 설정해야 하는 어려움이 존재한다. 따라서 본 논문에서는 피치주기 선정에 이용되는 최초 국부 최소 계곡점의 정확한 검출을 위해 변형된 AMDF 알고리즘을 제안하였으며, 이때 basic-AMDF $\gamma(l)$ 의 물리적 특징으로부터, 샘플이 커질수록 basic-AMDF의 크기가 감소하는 특성을 분석하였다.

이러한 basic-AMDF의 감소특성에 따라, 그림 13에서 볼 수 있듯이, 먼저 최초 피크 지점과 N 지점을 연결하는 가상의 직선을 설정한다. 그리고 이러한 가상의 직선과 수평선과의 각 θ 를 구하며, 이때 각 θ 는 각각의 프레임에서 적응적으로 구해진다. 아래의 식 (8)은 각 θ 를 설정하기 위해 유도된 식이다.

$$\theta = \arctan \left\{ \frac{R_{\max}}{N - l_{R_{\max}}} \right\} \quad (8)$$

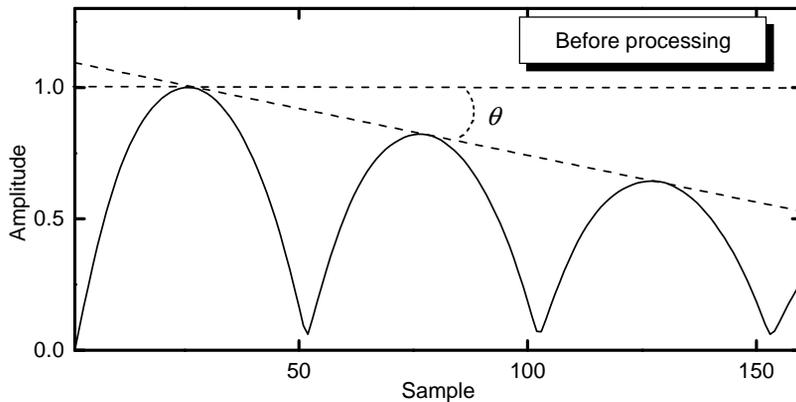


Fig. 13. The establishment of angle θ .

여기서 R_{\max} 는 각 프레임에서의 피크값이며, N 은 분석을 위한 프레임의 길이이고 $l_{R_{\max}}$ 는 피크지점의 인덱스이다.

그리고 식 (9)는 본 논문에서 제안한 변형된 AMDF이다.

$$\gamma_M(l) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \gamma(l) \quad (9)$$

식 (9)에서 제시된 변형된 AMDF는 설정된 각 θ 를 이용하여 basic-AMDF $\gamma(l)$ 을 회전변환함으로써, 피치주기 검출을 위한 최초 국부 최소 계곡점의 선정을 용이하게 한다. 이러한 회전변환 후의 변형된 AMDF를 그림 14에 나타내었다. 그림으로부터, 최초 국부 최소 계곡점과 나머지 계곡점들 사이의 크기차가 커짐에 따라, 최초 국부 최소 계곡점이 전체 최소 계곡점으로 검출된다. 따라서 변형된 AMDF는 피크값 이후의 최소값을 피치주기로 선정하는 간단한 알고리즘을 가진다.

또한 회전각에 대한 유연성을 갖도록 하기 위하여 변형된 AMDF에 다음의 식 (10)과 같이 파라미터 α 를 추가하였다.

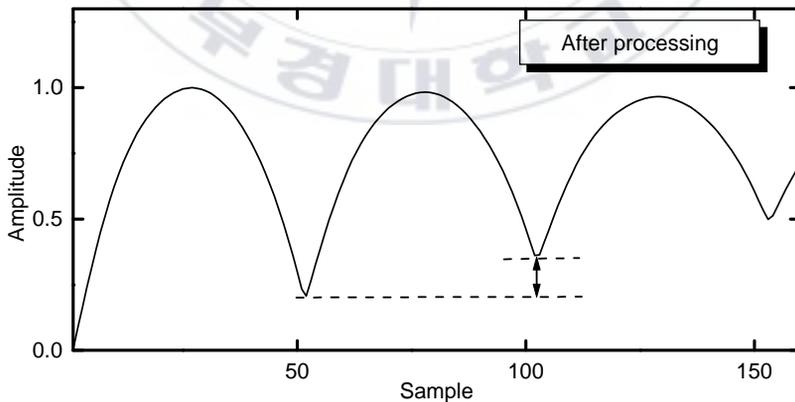


Fig. 14. Modified AMDF after processing.

따라서, 적절한 α 를 설정함으로써 피치주기인 최초 국부 최소 계곡점과 나머지 계곡점과의 크기차가 더욱 크게 나타난다.

$$\theta = \arctan \left\{ \frac{R_{\max}}{N - l_{R_{\max}}} \right\} + \alpha \quad (10)$$

그림 15는 식 (10)에서 제시된 파라미터 α 를 0.1, 0.2, 0.3으로 설정하였을 경우에 대한 영향이다. 그림에서 볼 수 있듯이, α 의 크기가 증가함에 따라 국부 계곡점들 간의 크기차가 더욱 커지므로, 적절한 α 의 선정으로 인해 피치주기가 보다 용이하게 검출된다.

또한, 음성의 시작구간에 대해서는 그림 16(a)와 같이 피치주기 검출을 위한 최초 국부 최소 계곡점이 명확하지 않으므로 검출오류가 발생한다. 따라서 그림 16(b)와 같이 각 프레임에 임계값 TH 를 적용하여 최초 피크 값 이후, TH 보다 작은 크기를 가지는 프레임의 계곡점만을 피치주기로 선정하였다. 이때, TH 는 식 (11)과 같이 정의된다.

$$TH = \beta \cdot \gamma_M(\text{last}) \quad (11)$$

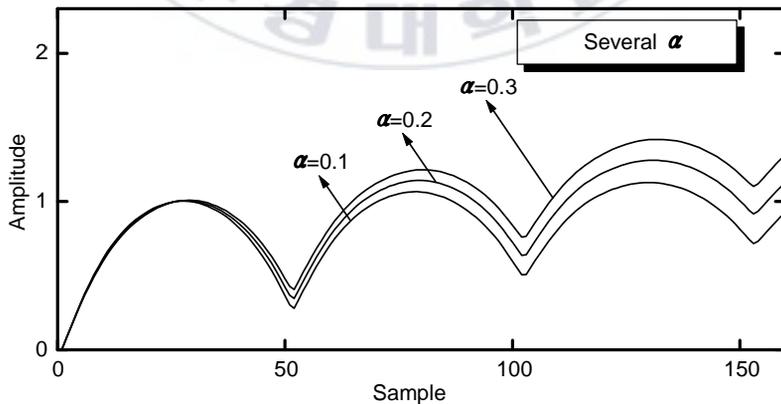
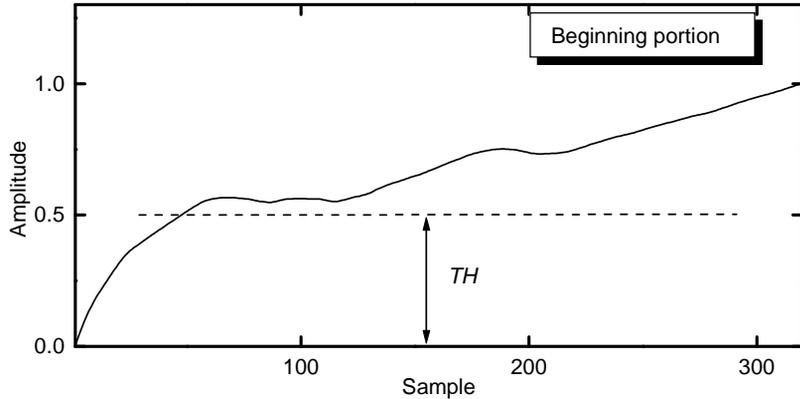
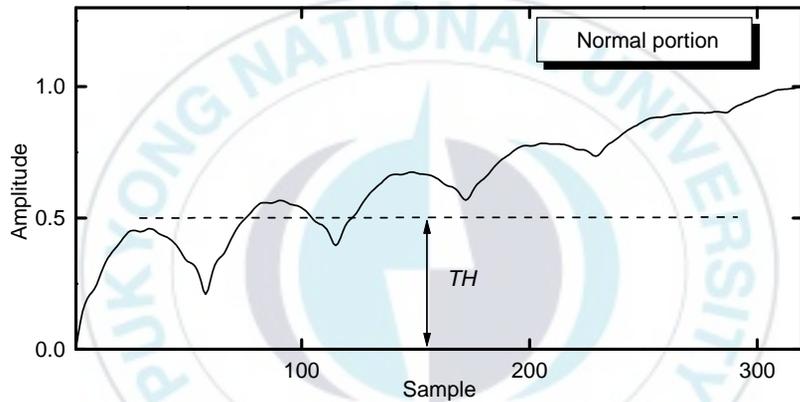


Fig. 15. Modified AMDF using several α .



(a) Beginning portion



(b) Normal portion

Fig. 16. Threshold TH of modified AMDF.

그리고 제안한 알고리즘의 블록 다이어그램을 그림 17에 나타내었다. 먼저, 음성획득 장치로부터 얻어지는 신호를 $s(n)$ 으로 설정한다. 이러한 $s(n)$ 으로부터, 변화하는 음성에 대해서 최적의 음성특징을 검출하기 위하여 프레임별로 나누어 분석하며, 한 프레임씩 중첩하여 분석을 진행한다. 그런 다음, 창함수와의 연산을 통해서 처리를 위한 신호 $s_w(n)$ 을 획득한다. 그리고 각 프레임별로 단구간 에너지를 구하여 큰 값의 에너지를 가지는 유성음과 그렇지 않은 무성음을 구분한다.

다음으로, 실험값으로 구해진 α 를 이용하여 basic-AMDF를 회전변환함으로써 변형된 AMDF인 $\gamma_M(l)$ 를 구한다. 또한, 실험값으로 설정된 β 를 적

용한 임계값을 사용하여 비정상구간을 판별하고 최종적으로 정상구간에 대해서만 피치주기를 결정한다.

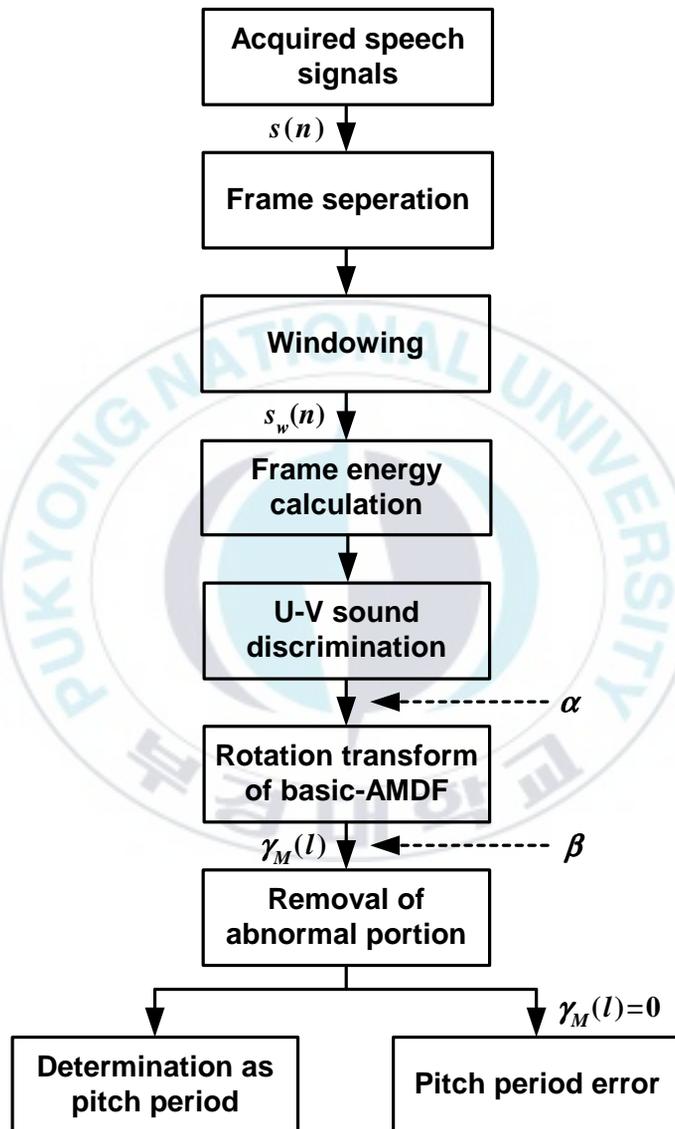


Fig. 17. Block diagram of modified AMDF.

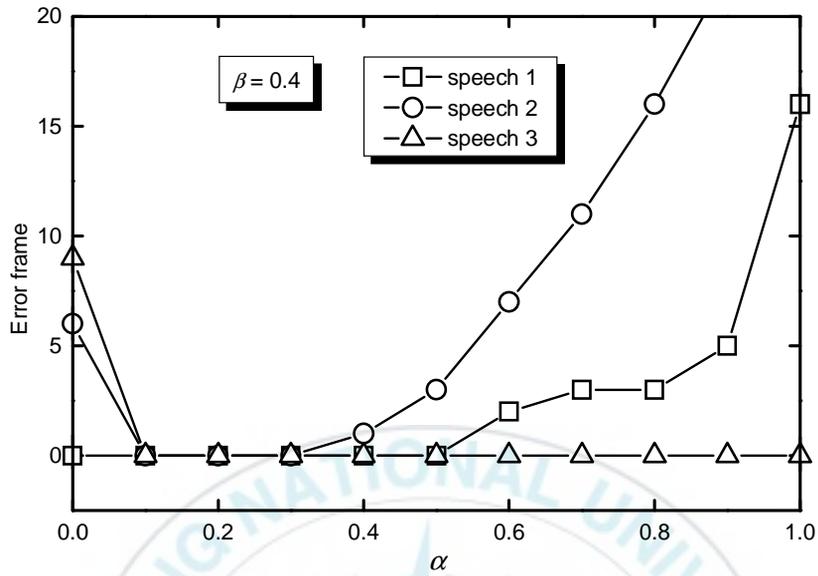
제5장 시뮬레이션 및 결과

본 논문에서 제안한 변형된 AMDF 알고리즘의 피치주기 검출성능을 테스트하기 위한 시뮬레이션에서, 사운드 카드는 P4i4SGV 모델을 사용하였으며 마이크로폰은 YMC-2511 모델과 콘덴서 마이크로폰을 병행하여 사용하였다. 획득한 음성신호 및 소음신호의 분해능은 16비트이고, 표본화 주파수는 8[kHz]로 설정하였다. 그리고 한 프레임의 길이는 10[ms](80표본)이며, 분석 프레임의 길이는 네 프레임(320표본)으로 하였고, 한 프레임씩 중첩시켜 피치주기를 검출하였다. 또한 ‘method A’는 AMDF 알고리즘에서 전체 최소값을 피치주기로 선정한 것이며, ‘method B’는 최초 국부 최소 계곡점 검출을 위한 추가의 알고리즘을 적용한 것이다. 그리고 추가의 알고리즘에서 적용된 임계값 Th1, Th2, Th3는 각각 0.2, 8, 0.5이다.

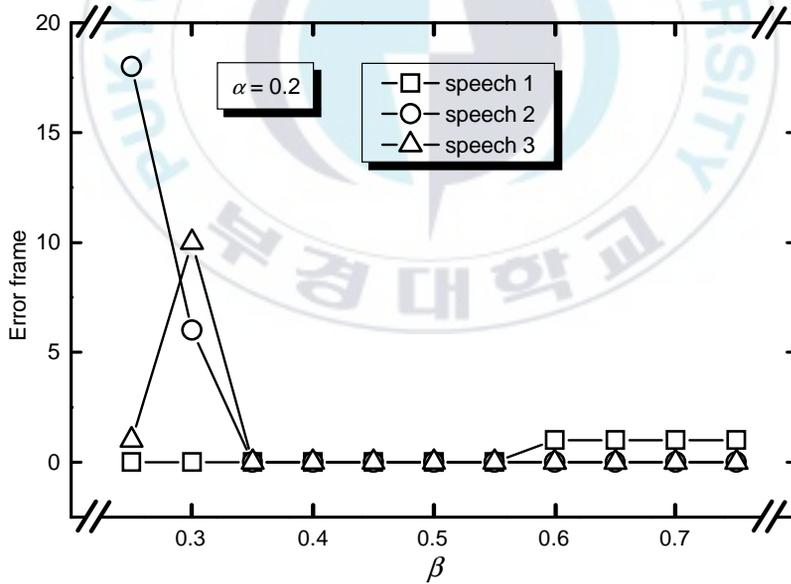
5.1 최적의 α 와 β 설정을 위한 시뮬레이션 및 결과

본 논문에서 정의한 α 와 β 를 최적의 값으로 설정하기 위해 그 크기를 가변하여 테스트하였다. 이러한 테스트에 /환골탈태/, /일사천리/, /천고마비/로 발생된 세 가지의 음성이 사용되었으며, 그림 18에 그 결과를 나타내었다.

그림 18(a)에서 확인할 수 있는 바와 같이, 세 가지의 테스트 음성에 대해 공통적으로 α 가 0.1, 0.2, 0.3일 경우, 가장 적은 에러 프레임을 나타내어 α 의 설정값으로 0.2를 선택하였다. 또한 그림 18(b)에서 β 는 세 가지의 테스트 음성에 대해 0.4와 0.5구간에서 가장 적은 에러 프레임을 나타내어 β 의 설정값으로 0.4를 선택하였다. 이러한 실험결과를 바탕으로 최적의 설정값을 선택하였다. 그리고 각 음성에 대한 α 와 β 의 변화에 따른 결과를 그림 19부터 그림 24까지 나타내었다.



(a) Comparison of error frames with α



(b) Comparison of error frames with β

Fig. 18. Simulation results for setting α, β .

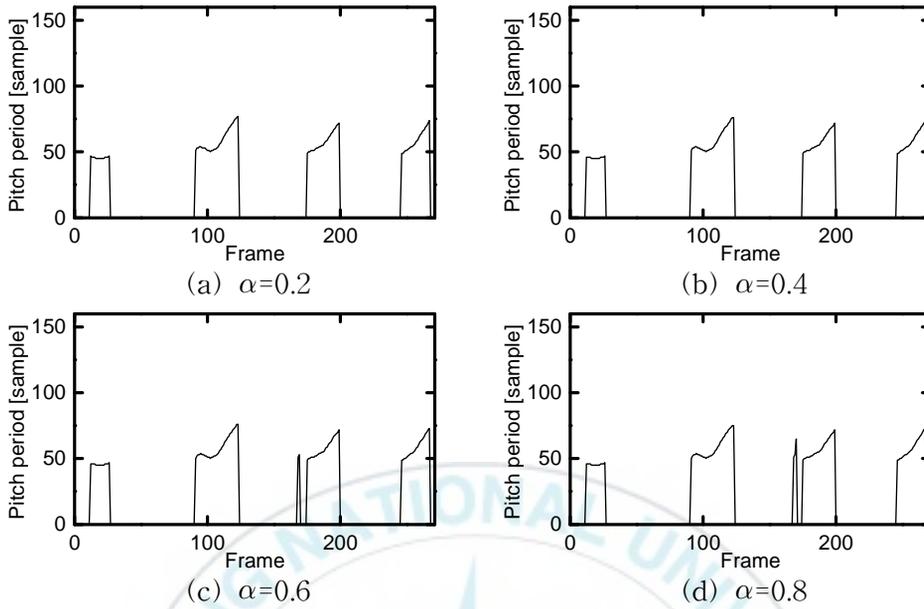


Fig. 19. The influence of variation of α ($\beta=0.4$) on pitch detection /hwangoltalae/.

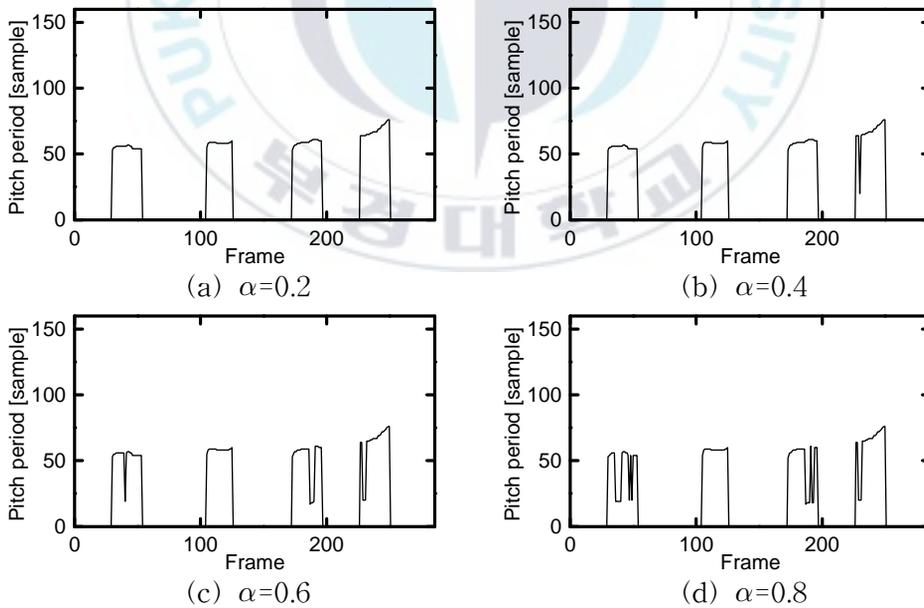


Fig. 20. The influence of variation of α ($\beta=0.4$) on pitch detection /ilsacheonli/.

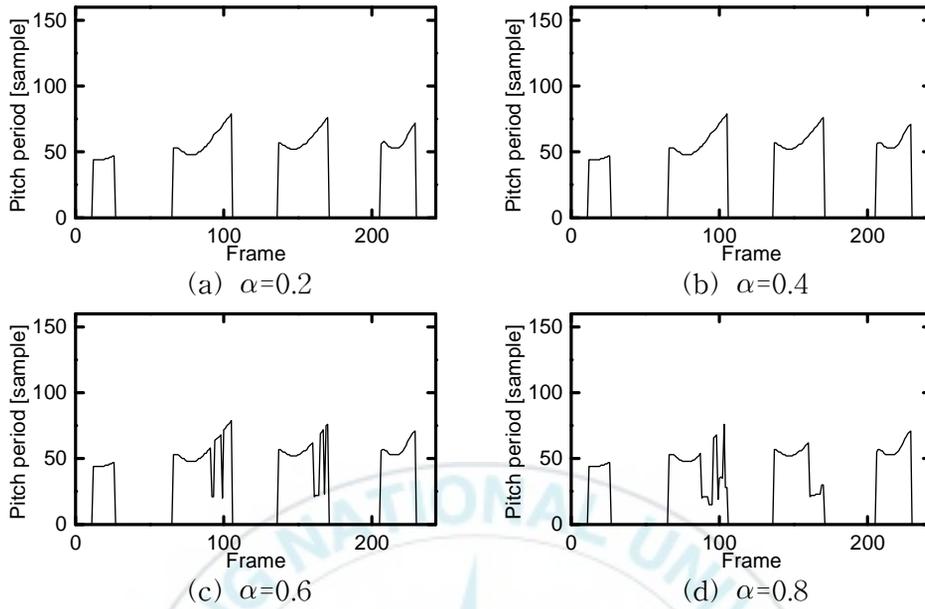


Fig. 21. The influence of variation of α ($\beta=0.4$) on pitch detection /cheongomabi/.

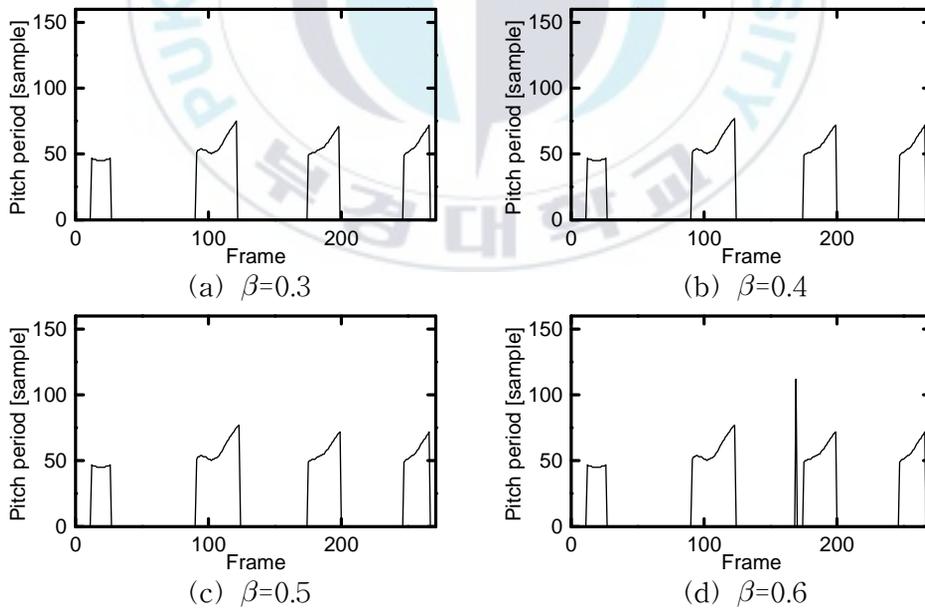


Fig. 22. The influence of variation of β ($\alpha=0.2$) on pitch detection /hwangoltalae/.

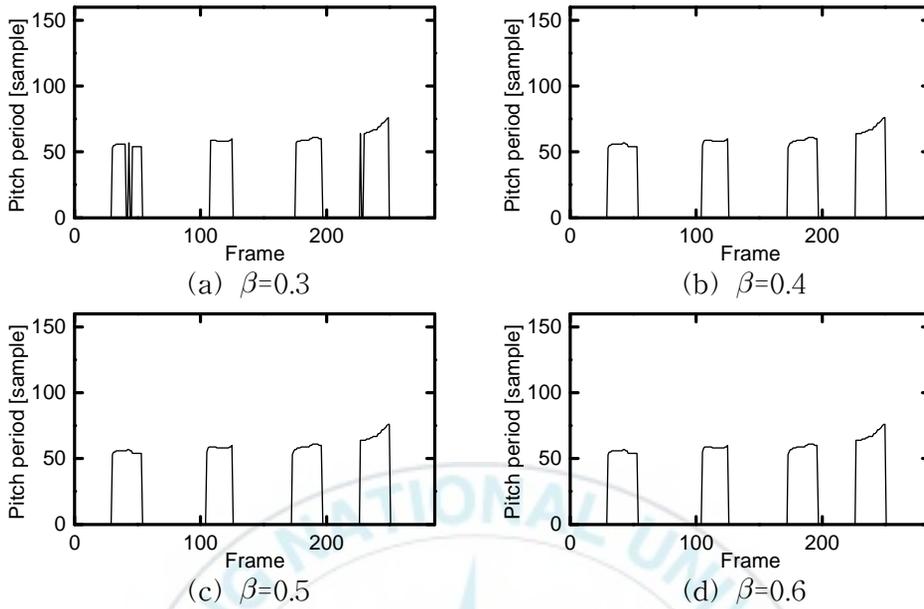


Fig. 23. The influence of variation of $\beta(\alpha=0.2)$ on pitch detection /ilsacheonli/.

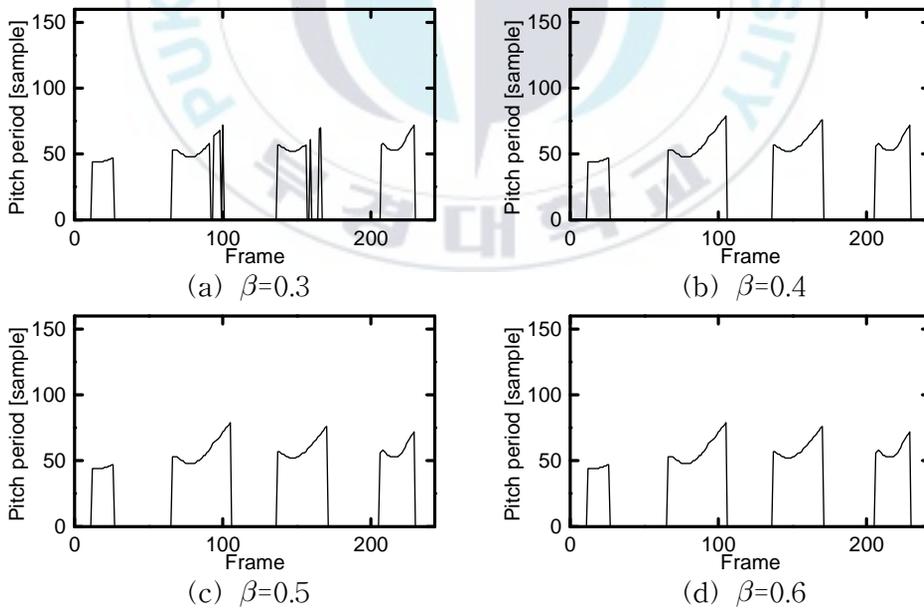


Fig. 24. The influence of variation of $\beta(\alpha=0.2)$ on pitch detection /cheongomabi/.

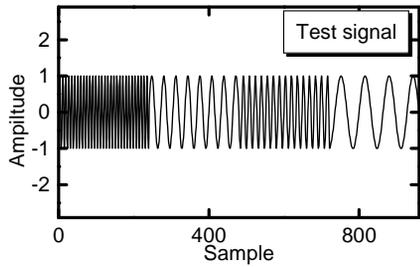
5.2 AWGN이 중첩된 정현파 신호에 대한 시뮬레이션 및 결과

피치주기는 음성신호의 기본주파수이며, 피치주기가 발생하는 유성음은 저주파의 정현파로 모델링될 수 있다. 또한, 피치주기 검출알고리즘은 준주기성을 띠는 파형에 대해서 기본주파수를 정확하게 추출해야 한다. 그리고 음성신호에는 기본주파수만이 존재하는 것이 아니라 여러 주파수가 중첩되어 있는 형태이다. 그러므로 다양한 주파수를 포함하는 입력신호에 대해, 정확히 기본주파수를 검출하는 것은 피치주기 검출알고리즘의 중요한 목적이다.

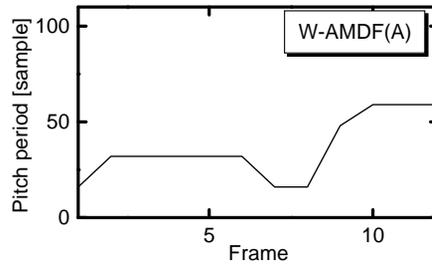
따라서, 주파수가 1[kHz], 0.25[kHz], 0.5[kHz], 0.125[kHz]로 변화하는 정현파 신호에 대해서 AMDF 알고리즘을 사용하여 80샘플씩 이동하면서 기본주파수를 검출하였으며, 이때 표본화 주파수는 8[kHz]이다. 또한 실제 잡음 환경을 고려하여 SNR 5[dB]의 AWGN을 정현파 신호에 중첩시켜 테스트하였다.

그림 25는 AWGN을 중첩시키지 않은 테스트 신호에 대한 기본주파수 검출결과이다. 그림으로부터 (a)는 주기가 변화하는 정현파 신호이며, (b), (c), (d), (e), (f)는 각각 W-AMDF(method A), W-AMDF(method B), LV-AMDF(method A), LV-AMDF(method B), 변형된 AMDF에 의한 기본주파수 검출결과이다. 결과로부터, 제안한 방법의 알고리즘은 주파수가 가변되는 정현파 신호에 대해서 우수한 검출결과를 나타내었다.

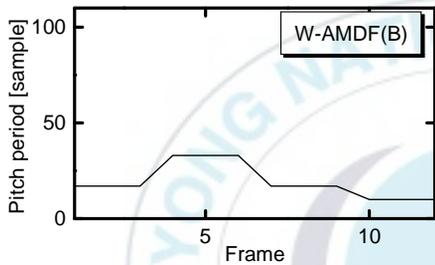
그림 26은 SNR 5[dB]의 AWGN을 정현파 신호에 중첩하여 시뮬레이션한 결과이다. 그림으로부터, 잡음의 영향으로 인해 기존의 W-AMDF와 LV-AMDF에 추가의 알고리즘을 적용한 경우에서도 많은 검출오류가 발생하였다. 그러나 제안한 방법에서는 잡음에 대한 강인한 특성으로 인해 테스트 신호의 기본주파수를 정확히 검출하였다.



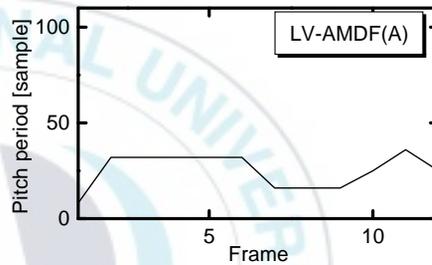
(a) A test signal



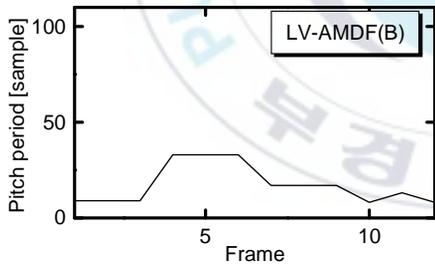
(b) W-AMDF(method A)



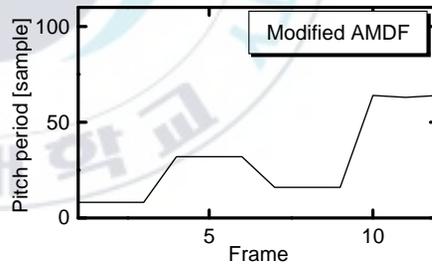
(c) W-AMDF(method B)



(d) LV-AMDF(method A)



(e) LV-AMDF(method B)



(f) Modified AMDF

Fig. 25. Simulation results for sinusoidal signals.

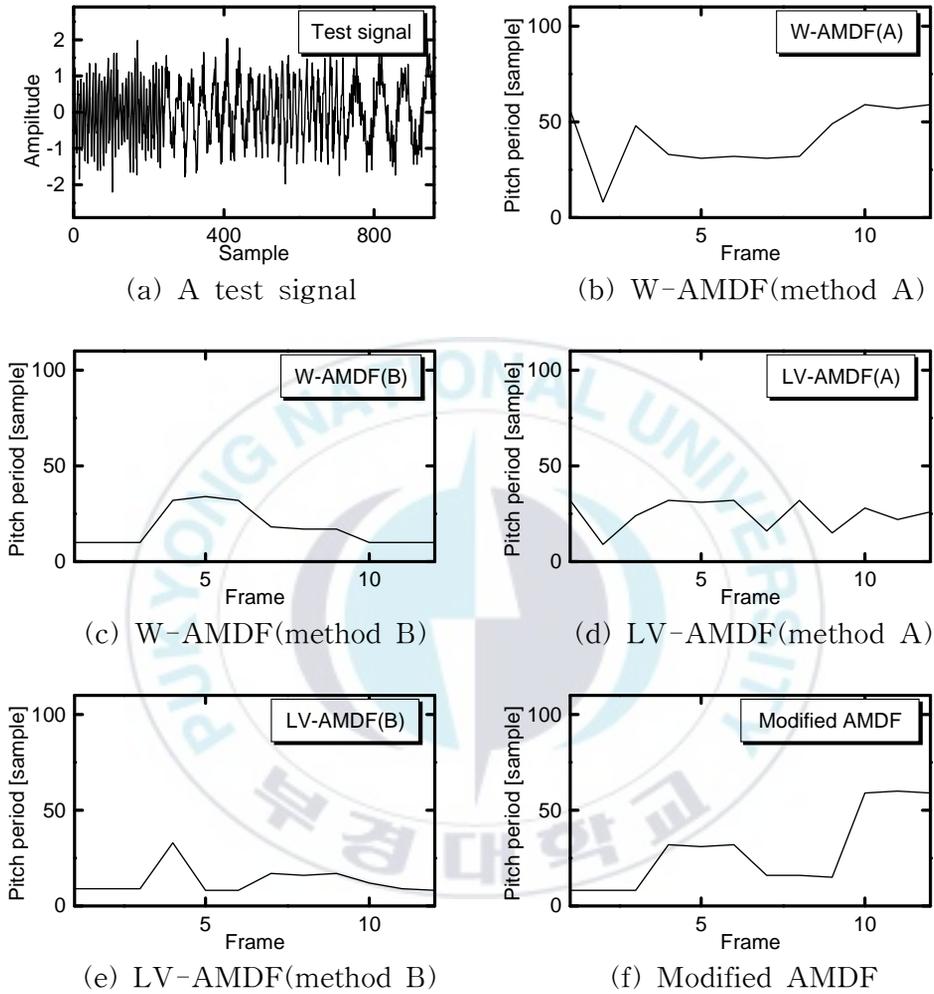
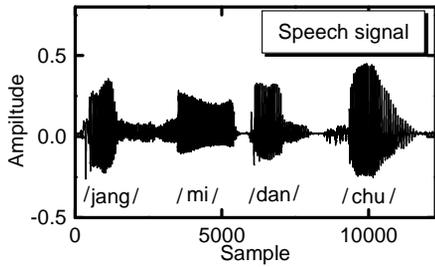


Fig. 26. Simulation results for sinusoidal signals superposed on AWGN(SNR=5[dB]).

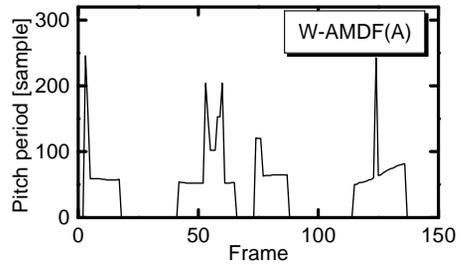
5.3 무소음 환경에서의 음성에 대한 시플레이션 및 결과

무소음 환경에서 획득한 음성신호에 대한 피치주기 검출결과를 그림 27부터 그림 29까지 나타내었다. 각각의 그림으로부터, (a)는 /장미단추/, /부경대학교/, /환골탈태/로 발성된 음성신호이다. 또한 (b), (c), (d), (e), (f)는 각각 W-AMDF(method A), W-AMDF(method B), LV-AMDF(method A), LV-AMDF(method B), 변형된 AMDF에 의한 피치트랙 검출결과이다. 각 ‘method A’에서 기존의 AMDF 알고리즘은 피치주기를 의미하는 최초 국부 최소 계곡점이 전체 최소 계곡점이 되지 않음에 따라, 많은 오류가 발생하여 정확한 피치트랙 검출이 불가능하였다. 그리고 ‘method B’에 의한 결과는 ‘method A’에 비해 개선된 성능을 보이지만 여전히 부분적으로 검출오류가 발생되었다. 그러나 제안한 변형된 AMDF 알고리즘에서는 회전변환으로 인해 최초 국부 최소 계곡점이 강조되어 기존의 AMDF 알고리즘에 추가의 알고리즘을 적용한 결과보다 우수한 피치주기 검출특성을 나타냈다.

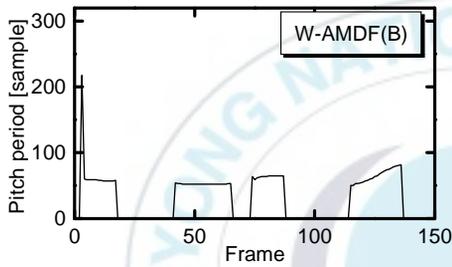
그리고 표 1은 여섯 개의 음성에 대한 결과 데이터이며, speech 1에서 speech 6은 각각 /장미단추/, /부경대학교/, /대한민국/, /환골탈태/, /일사천리/, /천고마비/로 발성된 음성이다. 표 1로부터, W-AMDF와 LV-AMDF의 ‘method A’의 경우에는 피치주기 검출오류가 많이 발생하였으며, 기존의 AMDF에 추가의 알고리즘을 적용한 ‘method B’에서는 ‘method A’에 비하여 검출 에러 프레임의 수가 상대적으로 감소하였다. 그러나 제안한 변형된 AMDF 알고리즘은 추가의 알고리즘 없이 ‘method B’에 비해서 상대적으로 적은 에러 프레임 수를 나타내었다. 이러한 결과로부터, 제안한 방법의 정확한 피치주기 검출특성을 확인하였다.



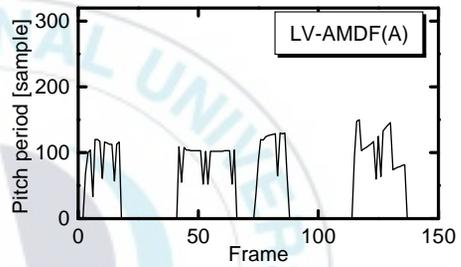
(a) A speech signal



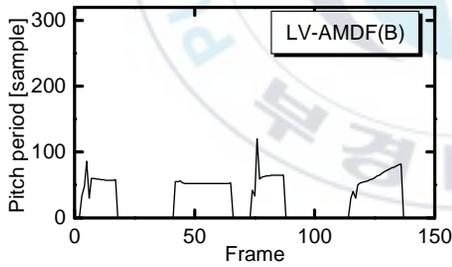
(b) W-AMDF(method A)



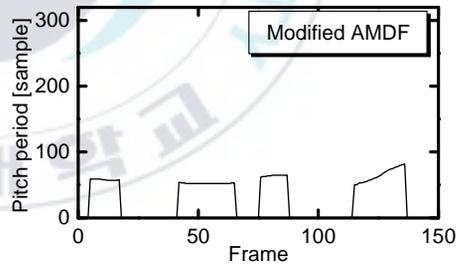
(c) W-AMDF(method B)



(d) LV-AMDF(method A)

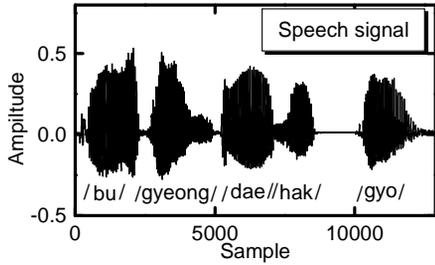


(e) LV-AMDF(method B)

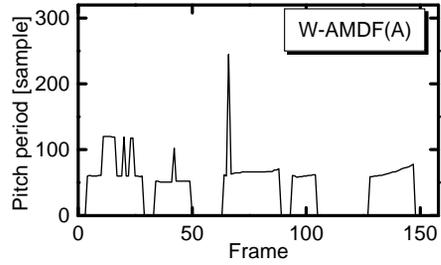


(f) Modified AMDF

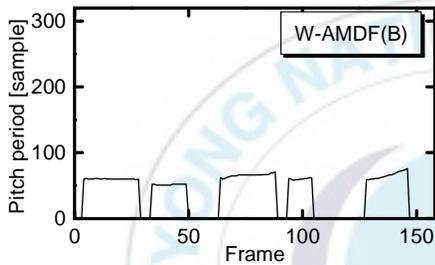
Fig. 27. /Jangmidanchu/ simulation results.



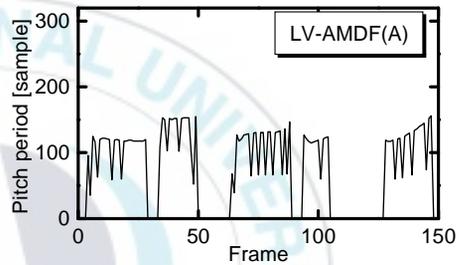
(a) A speech signal



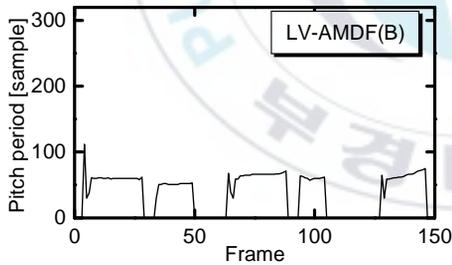
(b) W-AMDF(method A)



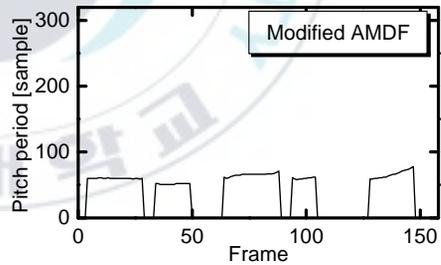
(c) W-AMDF(method B)



(d) LV-AMDF(method A)

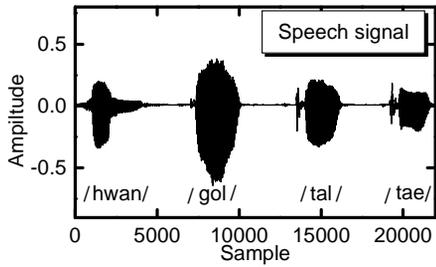


(e) LV-AMDF(method B)

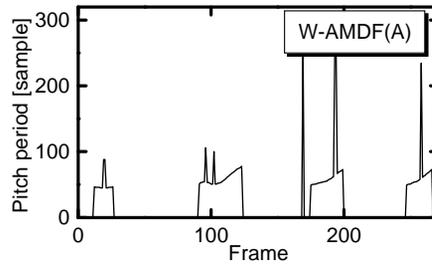


(f) Modified AMDF

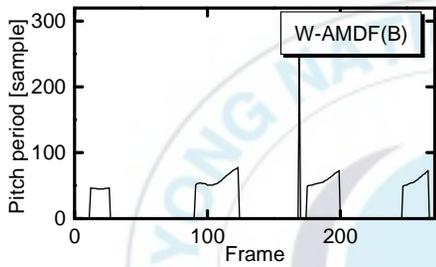
Fig. 28. /Bugeongdaehakgyo/ simulation results.



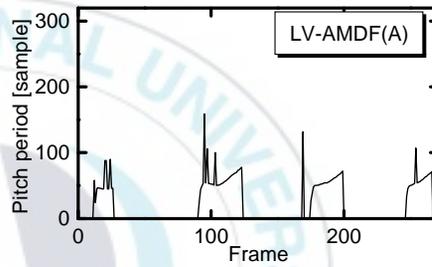
(a) A speech signal



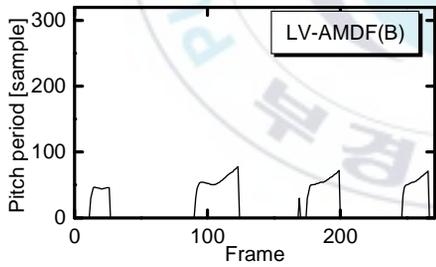
(b) W-AMDF(method A)



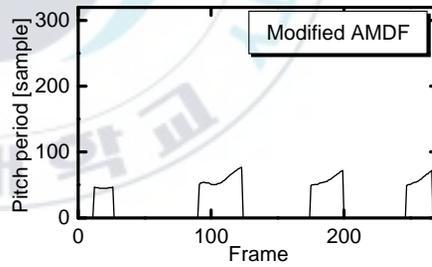
(c) W-AMDF(method B)



(d) LV-AMDF(method A)



(e) LV-AMDF(method B)



(f) Modified AMDF

Fig. 29. /Hwangoltaltae/ simulation results.

Table 1. Comparison of pitch period detection error frames for six speeches

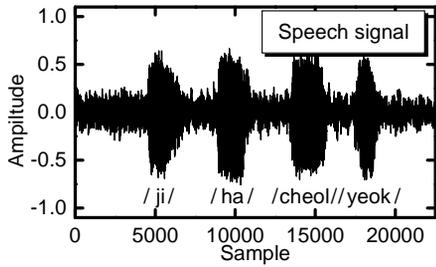
	Error frames				
	W-AMDF (method A)	W-AMDF (method B)	LV-AMDF (method A)	LV-AMDF (method B)	Modified AMDF
Speech 1 (75 frames)	14	1	20	10	0
Speech 2 (96 frames)	11	0	17	8	0
Speech 3 (65 frames)	11	1	11	4	1
Speech 4 (94 frames)	8	1	12	1	0
Speech 5 (93 frames)	14	0	18	4	0
Speech 6 (113 frames)	8	0	12	4	0
Total	66	3	90	31	1

5.4 소음환경에서의 음성에 대한 시뮬레이션 및 결과

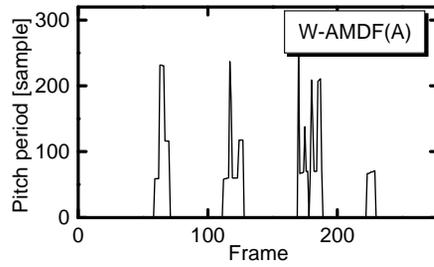
음성신호처리가 응용되는 많은 환경에서는 전자파 및 주변 소음과 같은 외란이 발생하므로, 이를 제거하거나 입력되는 음성신호를 복원하기 위한 방법이 추가되어야 한다. 또는 사용되는 알고리즘이 소음에 강인한 특성을 나타내어 이러한 실제 환경에서 발생하는 소음이나 잡음에 대처해야 한다. 따라서 본 논문에서 제안한 알고리즘에서는 소음환경에 대한 방안으로 임계값을 적용하였으며, 이러한 피치주기 검출 성능을 확인하기 위하여 지하철과 버스에서 획득한 소음을 음성신호에 중첩하여 시뮬레이션하였다.

그림 30은 지하철에서 발생한 소음에 대한 시뮬레이션 결과이다. (a)는 무소음 환경에서 발생한 /지하철역/ 음성과 부산 지하철 2호선에서 획득한 소음에 대한 합연산으로부터 얻어진 테스트 신호이다. 그리고 (b), (c), (d), (e), (f)는 각각 W-AMDF(method A), W-AMDF(method B), LV-AMDF(method A), LV-AMDF(method B), 변형된 AMDF의 피치트랙 검출결과이다. 그림으로부터, 기존의 AMDF 알고리즘의 ‘method A’는 지하철에서 발생한 소음에 의해 정확한 피치트랙 검출이 불가능하였으며, ‘method B’에서는 ‘method A’에 비해 개선된 결과를 나타내었다. 그러나 제안한 변형된 AMDF에 의한 시뮬레이션 결과인 (d)에서는 회전변환으로 인해 최초 국부 최소 계곡점이 강조되어 중첩된 소음에 대한 강인한 특성을 나타내었으며, 임계값의 적용으로 인해 비정상구간에 대한 오류를 제거시킴으로써 ‘method B’의 방법보다도 정확한 피치트랙 검출결과를 나타내었다.

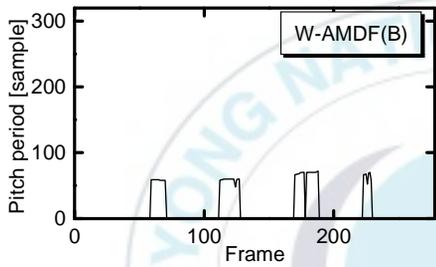
그림 31은 자동차 소음에 대한 시뮬레이션 결과이다. (a)는 무소음 환경에서 발생한 /일사천리/ 음성과 BM090 모델의 버스에서 획득한 소음의 합연산을 통해 생성된 테스트 신호이다. 그리고 (b), (d)는 각각 W-AMDF(method A), LV-AMDF(method A)의 피치트랙 검출결과이며, 소음에 의한 영향으로 정확한 검출이 불가능하였다. 또한 (c), (e), (f)는 각각 W-AMDF(method B), LV-AMDF(method B) 그리고 제안한 방법에 의한 시뮬레이션 결과이며, 제안한 방법이 보다 정확한 피치트랙 검출특성을 나타내었다.



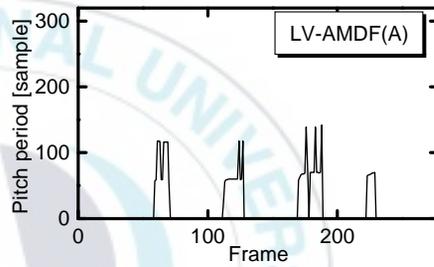
(a) A speech signal



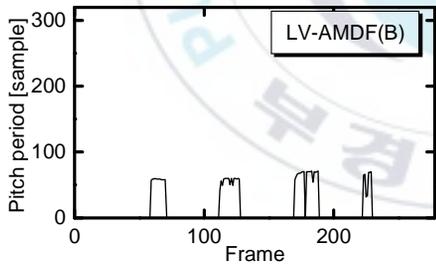
(b) W-AMDF(method A)



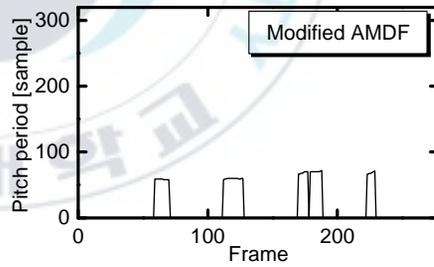
(c) W-AMDF(method B)



(d) LV-AMDF(method A)

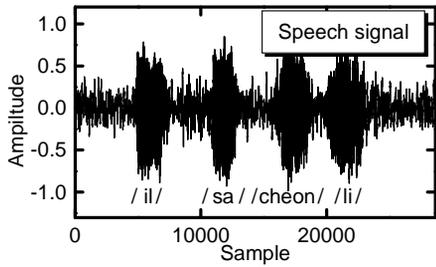


(e) LV-AMDF(method B)

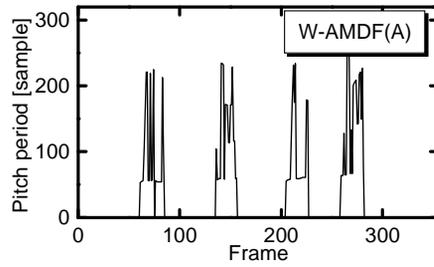


(f) Modified AMDF

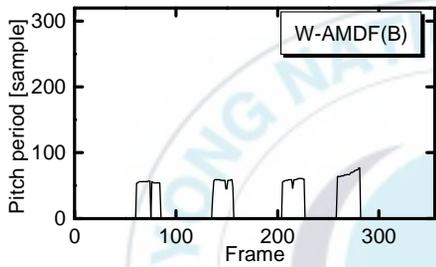
Fig. 30. /Jihacheolyeok/ simulation results for the subway environment.



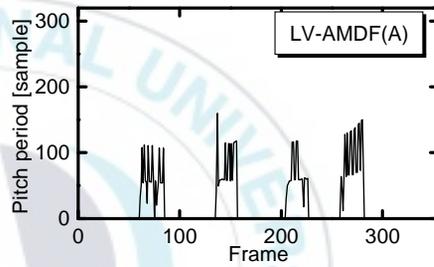
(a) A speech signal



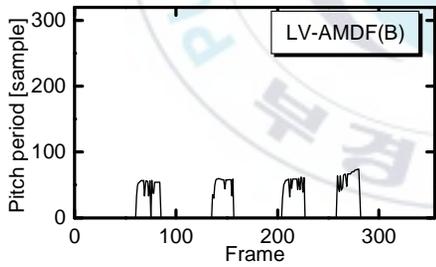
(b) W-AMDF(method A)



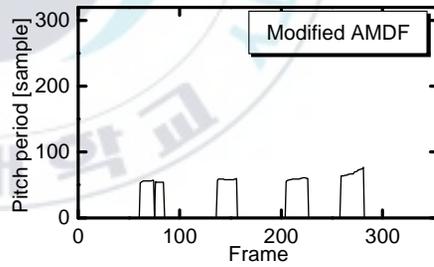
(c) W-AMDF(method B)



(d) LV-AMDF(method A)



(e) LV-AMDF(method B)



(f) Modified AMDF

Fig. 31. /ilsacheonli/ simulation results for the bus environment.

제6장 결 론

본 논문에서는 음성신호의 피치주기를 검출하기 위해, 기존의 AMDF 알고리즘을 응용한 변형된 AMDF를 제안하였으며, basic-AMDF를 회전변환함으로써 최초 국부 최소 계곡점이 전체 최소 계곡점이 되어 피치주기 검출이 용이하였다. 따라서 제안한 방법은 전체 최소 계곡점을 바로 피치주기로 선정하는 간단한 알고리즘이며, 임계값을 설정하여 음성의 시작구간과 같은 비정상 구간에 대한 오류를 감소시켰다.

그리고 시뮬레이션 및 결과에서, AWGN을 가변하여 중첩시킨 정현파 신호에 대해 정확한 기본주파수 검출 및 잡음에 강인한 특성을 확인하였다. 또한 무소음 환경에서 발생된 여섯 개의 음성신호에 대한 시뮬레이션에서도 우수한 특성을 나타내었으며, 기존의 AMDF에 최초 국부 최소 계곡점 검출을 위한 추가의 알고리즘을 적용한 경우와 비교하여도 제안한 방법에서 적은 검출 에러 프레임을 나타내어, 실제 음성신호에 대한 적용에서도 우수한 특성을 보였다. 그리고 음성신호처리가 이용되는 실제 환경에서는 주변 소음과 같은 외란에 자유로울 수 없으므로, 지하철과 자동차 등에서 획득한 소음과 음성의 합연산을 통해 만든 테스트 신호에 대해서도 시뮬레이션을 하여 소음에 강인한 특성을 나타내고자 하였으며, 결과로부터 기존의 AMDF 알고리즘에 비하여 제안한 변형된 AMDF 알고리즘에서 정확한 피치주기 검출 특성을 확인하였다.

따라서, 본 논문에서 제안한 변형된 AMDF 알고리즘은 음성 합성, 음성 인식, 화자 식별등과 같은 다양한 음성신호처리 분야에서 특징 벡터 추출을 위해 유용하게 적용되리라 사료된다.

참 고 문 헌

- [1] L. R. Rabiner et al., "A comparative performance study of several pitch detection algorithms", *IEEE Trans. Acoust., Speech, Signal Processing*, 24(5), pp. 399-417, 1976.
- [2] M. J. Ross et al., "Average magnitude difference function pitch extractor", *IEEE Trans. Acoust., Speech, Signal Processing*, 22(5), pp. 353-362, 1974.
- [3] B. Gold and L. R. Rabiner, "Parallel processing techniques for estimating pitch periods of speech in the time domain", *J. Acoust. Soc. Amer.*, vol. 46, pp. 443-448, Aug, 1976.
- [4] A. M. Noll, "Cepstrum pitch determination", *J. Acoust. Soc. Amer.*, 41(2): 293-309, 1967.
- [5] L. R. Rabiner, "On the use of autocorrelation analysis for pitch detection", *IEEE Trans. ASSP*, ASSP-25: 24-33, 1977.
- [6] S. Kadambe and G. F. Boudreaux-Bartels, "Application of the wavelet transform for pitch detection of speech signals", *IEEE Trans. Information Theory*, 38(2), pp. 917-924, 1992.
- [7] G. S. Ying, L. H. Jamieson and C. D. Michell, "A probabilistic approach to AMDF pitch detection", *IEEE ICSLP-96*. pp. 1201-1204, 1996.
- [8] C. Kwanun and S. C. Ying. "A pitch extraction algorithm based on LPC inverse filtering and AMDF", *IEEE Trans. Acoust., Speech, Signal Processing*. ASSP-25(6): 565-572, 1977.
- [9] Sneff, "Real time harmonic pitch detection", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, 1978.
- [10] W. Hess, "Pitch determination of speech signals: algorithms and devices", Berlin: Springer-Verlag, 1983.
- [11] L. Rabiner and B. W. Juang, "Fundamentals of speech recognition",

Prentice Hall PTR, 1993.

- [12] 한진수, “음성 신호 처리”, 오성미디어, 2003.
- [13] C. Manfredi, “Adaptive noise energy estimation in pathological speech signals”, IEEE Trans. Biomedical Engineering, 47(11), pp. 1538-1543, 2000.
- [14] X. mei, J. Pan and S. Sun, “Efficient algorithms for speech pitch estimation”, Proc. of ISIMVSP-2001, pp. 421-424, 2001.
- [15] 서현수, 김남호, “AMDF의 회전변환과 임계값에 기반한 피치주기 검출 알고리즘에 관한 연구”, 한국 신호처리·시스템 학회 논문지, 제7권, 제4호, pp. 178-183, 2006.
- [16] 서현수, 배상범, 김남호, “변형된 AMDF를 이용한 피치 주기 검출 알고리즘”, 한국해양정보통신학회 학회지 논문지, 제10권, 제1호, pp. 23-28, 2006.
- [17] Hyun-Soo Seo, Sang-Bum Bae, Nam-Ho Kim, “MODIFIED AMDF ALGORITHM FOR PITCH DETECTION OF SPEECH SIGNAL”, The International Symposium On Advanced Engineering, pp. 153-157, 2005. 12. 2.
- [18] 서현수, 류지구, 김남호, “지하철 환경에서의 피치주기 검출을 위한 변형된 AMDF 알고리즘”, 한국통신학회·대한전자공학회·제어자동화시스템공학회 부산·경남지부 추계합동학술논문발표회 논문집, 제15권, 제2호, pp. 51-54, 2006. 12. 2.
- [19] 서현수, 류지구, 김남호, “잡음환경에서 음성신호의 피치주기 검출을 위한 변형된 AMDF에 관한 연구”, 한국 신호처리·시스템 학회 추계 학술대회 논문집, 제7권, 제2호, pp. 265-268, 2006. 11. 18.
- [20] 서현수, 배상범, 김남호, “변형된 AMDF를 이용한 자동차 소음환경에서의 피치주기 검출알고리즘”, 한국해양정보통신학회 추계종합학술대회 논문집, 제10권, 제2호, pp. 415-418, 2006. 10. 27.
- [21] 서현수, 배상범, 김남호, “소음환경에 강인한 피치주기 검출 알고리즘에 관한 연구”, 한국해양정보통신학회 추계종합학술대회 논문집, 제10권, 제1호, pp. 481-484, 2006. 5. 26.

- [22] 서현수, 배상범, 김남호, “AMDF의 감소특성과 임계값을 이용한 피치 검출에 관한 연구”, 한국멀티미디어학회 추계학술발표대회 논문집, 제8권, 제2호, pp. 63, 2005. 11. 25.
- [23] 서현수, 배상범, 김남호, “AMDF의 회전변환을 이용한 피치 주기 검출 알고리즘”, 한국해양정보통신학회 추계종합학술대회 논문집, 제9권, 제2호, pp. 1019-1022, 2005. 10. 28.
- [24] 서현수, 배상범, 김남호, “AMDF의 이용한 새로운 피치 주기 검출 알고리즘”, 한국해양정보통신학회 인공지능 및 지능정보시스템 학술대회 논문집, pp. 140-143, 2005. 8. 9.
- [25] 서현수, 김옥, 노현용, 배상범, 김남호, “음성신호의 피치 주기 검출을 위한 새로운 AMDF에 관한 연구”, 한국통신학회·대한전자공학회·제어자동화시스템공학회 부산·경남지부 추계합동학술논문발표회 논문집, 제14권, 제2호, pp. 73-74, 2005. 6. 4.
- [26] 서현수, 배상범, 김남호, “AMDF의 회전변환을 이용한 음성신호의 피치 주기 검출에 관한 연구”, 한국 신호처리·시스템학회 춘계학술대회 논문집, 제6권 제1호, pp. 31-34, 2005. 7. 16.
- [27] 서현수, 배상범, 김남호, “변형된 AMDF를 이용한 음성신호의 피치 주기 검출에 관한 연구”, 한국해양정보통신학회 춘계종합학술대회 논문집, 제9권, 제1호, pp. 515-519, 2005. 5. 27.

감사의 글

학부시절부터 학위논문을 쓰는 이 순간까지 잠은 따뜻하게 자는지, 생활하는 것에 어려움은 없는지, 항상 세심하게 챙겨주시고 때로는 엄하심으로 제가 올바른 길을 갈 수 있도록 이끌어주신 지도교수님이신 김남호 교수님께 고개 숙여 감사드립니다. 교수님의 가르침을 항상 가슴에 새기고 교수님께서 보여주시는 학자로서의 기품과 정도를 닮아가도록 노력하겠습니다. 교수님 감사합니다. 아울러, 바쁘신 와중에도 미비한 저의 논문에 많은 신경을 써주시고 항상 인자하신 미소로 격려와 충고의 말씀을 아끼지 않으신 황용연 교수님, 안영주 교수님께 다시 한번 감사드립니다. 지금까지 수준 높은 수업으로 저의 시야를 넓혀주셨고, 항상 학생들을 위해 노력하시고 희생하시는 이형기 교수님, 최연욱 교수님, 변기식 교수님, 김만고 교수님, 이경창 교수님께 감사드립니다.

열정이 넘치시는 강경덕 선배님, 학문적 도움과 따뜻한 미소로 용기를 주신 배상범 선배님, 모범이 되시는 구본석 선배님을 비롯한 우리 M&DSP 연구실 선배님들께 감사합니다. 연구실 생활을 재미나게 해준 동기 노현용, 중국에서 온 후배 Jin Bo, 최덕천 그리고 최귀홍님께 고마움을 전합니다. 시스템제어 연구실의 친형님같은 김동원 선배님을 비롯한 김혜란 선배님, 이덕, 이창원 후배님, 후배 류나이에게 감사의 말을 전합니다. 힘이 되어준 친구 박성혁, 항상 배려를 해주는 정은미, 청심회 동기들, 학부 동기들, 친구들과 행정적인 도움을 준 박명규, 신혜숙 조교에게도 고마움을 전합니다.

저를 이 세상에 존재하게 해주시고 어떤 어려운 상황에서도 항상 당신보다 아들을 먼저 생각하시며 못난 자식을 희망으로 바라보시는 아버지, 어머니 정말 감사합니다. 그리고 하나밖에 없는 든직한 형님에게 감사드리며, 미처 이름을 다 밝히지 못한 많은 분들께도 감사의 말씀을 드립니다.

이러한 큰 도움을 주신 많은 분들께 저의 미약한 재능으로 완성한 학위 논문을 바치며, 부끄럽지 않은 모습으로 항상 노력하겠습니다. 감사합니다.

2006년 12월
서 현 수 올림