

# AMDF의 회전변환을 이용한 피치 주기 검출 알고리즘

서현수\* · 배상범\* · 김남호\*

\*부경대학교 공과대학 제어계측공학과

## Pitch Period Detection Algorithm Using Rotation Transform of AMDF

Hyun-Soo Seo\* · Sang-Bum Bae\* · Nam-Ho Kim\*

\*Dept. of Control and Instrumentation Eng., Pukyong National Univ.

E-mail : digitalshs@mail1.pknu.ac.kr

### 요약

최근 정보 통신 기술의 급속한 발전에 의해 음성 신호 처리에 관련된 많은 연구가 진행됨에 따라 피치 주기는 음성 인식, 화자 식별, 음성 분석 및 합성 등과 같은 많은 응용분야에서 중요한 요소로써 적용되고 있다. 이러한 피치 주기 검출에 관련된 시간 영역과 주파수 영역에서의 많은 알고리즘이 제안되었으며, 시간 영역의 피치 검출 알고리즘의 하나인 AMDF(average magnitude difference function)는 각 valley점의 거리를 피치 주기로 계산한다. 그러나 피치 주기 검출을 위한 valley점 선정에 있어서 알고리즘이 복잡해지는 문제점이 발생한다. 따라서 본 논문에서는 AMDF의 회전변환을 이용하여 전체 최소 valley점을 음성 신호의 피치 주기로 인식하는 간단한 알고리즘을 제안하였으며, 음성의 시작구간에 대해 경계값을 설정하여 피치 주기 선정에 대한 판단기준으로 사용하였다. 그리고 제안한 알고리즘을 시뮬레이션을 통해 기존의 방법들과 비교하였다.

### ABSTRACT

As recent information-communication technology is rapidly developed, a lot of researches related to speech signal processing have been processed. So pitch period is applied as important factor to many application fields such as speech recognition, speaker identification, speech analysis and synthesis. Therefore, many algorithms related to pitch detection have been proposed in time domain and frequency domain and AMDF(average magnitude difference function) which is one of pitch detection algorithms in time domain chooses time interval from valley to valley as pitch period. But, in selection of valley point to detect pitch period, complexity of the algorithm is increased. So in this paper we proposed pitch detection algorithm using rotation transform of AMDF, that taking the global minimum valley point as pitch period and established a threshold about the phoneme in beginning portion, to exclude pitch period selection. and compared existing methods with proposed method through simulation.

### 키워드

pitch period, AMDF, valley point, threshold

### 1. 서론

최근 정보통신기술의 급속한 발전에 의해 음성 신호 처리의 중요성이 증가하고 있다. 그리고 음성 신호의 기본 주파수인 피치 주기는 인간의 청각에 매우 민감하게 반응하는 파라미터로써, 음성 신호의 화자를 구분하는데 사용하며, 자연성에 큰 영향을 미친다. 그러므로 음성 합성과 음성 코딩에서 피치의 정확한 추출과 복원은 음질에 결정적

인 역할을 한다. 따라서 이러한 피치 검출을 위한 시간영역과 주파수영역에서의 많은 알고리즘이 제안되었다[1][2].

시간영역에서의 피치 검출 알고리즘은 파형의 주기성을 강조하여 피치를 검출하는 방법으로서 병렬처리법, ACF(autocorrelation function), AMDF 등이 있다[2]. 그러나 음성이 천이 구간 내에 존재하는 경우와 잡음이 중첩된 음성의 경우에는 피치 검출을 위한 알고리즘이 복잡해진다.

그리고 주파수영역에서의 피치 검출 알고리즘은 음성 스펙트럼의 고조파 간격을 측정하여 유성음의 기본주파수를 검출하는 방법으로 Cepstrum법, 고조파분석법 등이 있다. 그러나 처리과정에서 주파수영역으로의 변환과정이 필요하기 때문에 계산이 복잡하며, 기본 주파수의 정밀한 측정을 위해 FFT의 포인터 수를 증가 시킬 경우 처리시간이 길어지게 된다[3].

이에 따라, 적은 계산량과 정확성을 나타내는 시간 영역의 AMDF가 널리 적용되고 있으나, 피치 주기 검출을 위한 valley점 선정이 용이하지 않다. 따라서 본 논문에서는 AMDF의 회전변환을 이용한 피치 주기 검출 알고리즘을 제안하였으며, 음성의 시작구간에 대해 경계값을 설정하여 피치 주기 선정에 대한 판단기준으로 사용하였다. 그리고 제안한 알고리즘을 시뮬레이션을 통해 기존의 방법과 비교하였다.

## II. AMDF 알고리즘

### 1. Basic-AMDF 알고리즘

음성신호  $s(n)$ 은 길이  $N$ 을 가진 창함수  $w(n)$ 과의 연산을 통해 신호  $s_w(n)$ 이 된다. 이때, Basic-AMDF  $\gamma(l)$ 은 다음과 같이 정의된다.

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

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

식 (1)로부터,  $s_w(n)$ 이 유성음이라면  $\gamma(l)$ 은 주기적인 특징을 나타내며,  $\gamma(l)$ 의 이러한 특징은 유성음과 무성음을 분류하거나 유성음의 피치 주기를 결정하는 분야에 사용된다. 일반적으로 그림 1에서 보는 바와 같이 유성음은 국부 최소 valley점이 주기적으로 나타난다. 따라서 AMDF는 전체 표본에 대한 최초 국부 최소 valley점을 피치 주기로 계산한다.

### 2. 기존의 AMDF 알고리즘

Basic-AMDF의 특성을 개선시키기 위해, 선형 가중 AMDF(W-AMDF), 가변 길이 AMDF(LV-AMDF) 등이 소개되었다[4][5]. 여기서 W-AMDF는 다음과 같이 정의된다.

$$\gamma_w(l) = \frac{1}{N-l+1} \sum_{n=1}^{N-l+1} |s_w(n+l-1) - s_w(n)|, \quad (2)$$

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

식 (2)와 같은 W-AMDF는 low-pitched 잡음 환경에서 Basic-AMDF가 가지는 피치 검출 오류를 극복하기 위해 제안되었다. 즉,  $(N-l+1)^{-1}$ 은

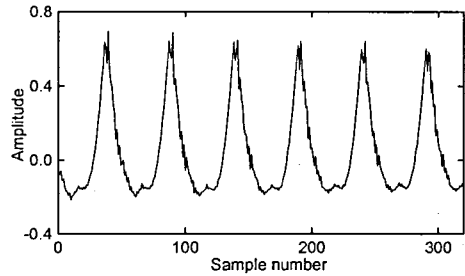
Basic-AMDF의 감소특성을 제거시켜준다.

그리고 LV-AMDF는 다음의 식 (3)과 같다.

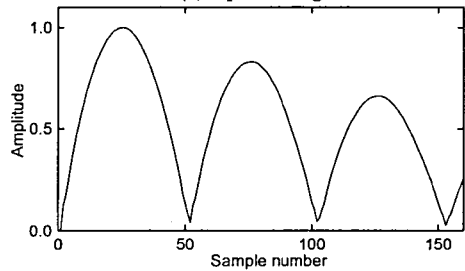
$$\gamma_{LV}(l) = \frac{1}{l} \sum_{n=1}^l |s_w(n+l-1) - s_w(n)|, \quad (3)$$

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

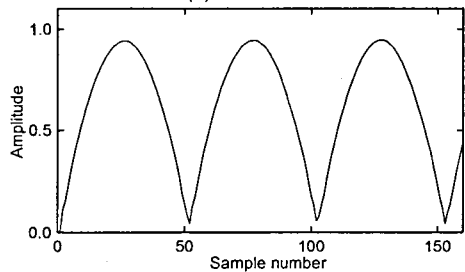
식 (3)과 같이 LV-AMDF는 합에 대한 연산량인 변수  $l$ 에 의해서 결정됨에 따라 Basic-AMDF와 W-AMDF에 비해 적은 연산량을 가진다.



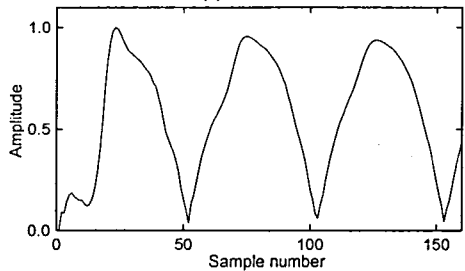
(a) Speech Signal



(b) Basic-AMDF



(c) W-AMDF



(d) LV-AMDF

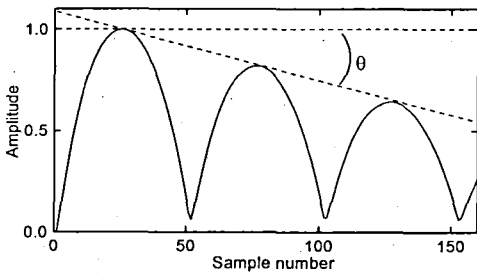
그림 1. 여러 가지 AMDF.

그러나 일반적인 유성음에서, 기존의 AMDF는 주기적인 valley점 검출에서 최초의 국부 최소 valley점이 피치 주기가 되지 않는 오류가 발생하고 있다.

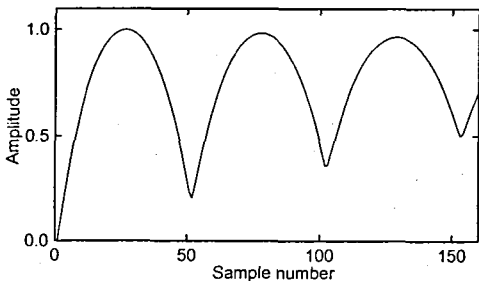
3. 제안한 변형된 AMDF

본 논문에서는 Basic-AMDF  $y(l)$ 의 특징을 분석하여, 그림 2의 (a)와 (b)에서 보여지는 바와 같이 AMDF의 최초 피크 지점과  $N$ 지점을 잇는 직선에 의해 수평선과의 각  $\theta$ 를 설정한다. 그리고 이러한 각  $\theta$ 를  $0^\circ$ 로 만들기 위해 AMDF를 양의 방향으로 회전시킨다. 다음의 식 (4)는 각  $\theta$ 를 설정하기 위해 유도된 식이다.

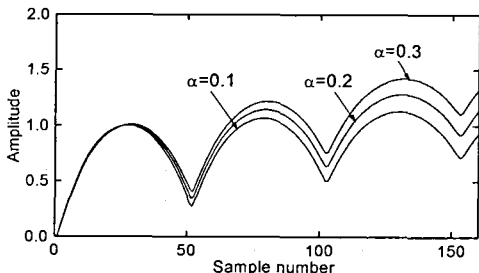
$$\theta = \arctan \frac{R_{\max}}{N} \tag{4}$$



(a) 변형된 AMDF의 각  $\theta$  설정



(b) 각  $\theta$  설정 후의 변형된 AMDF



(c) 파라미터  $\alpha$ 에 따른 변형된 AMDF

그림 2. 변형된 AMDF

여기서  $R_{\max} = \max\{y(l)\}$ 이며,  $N$ 은  $y(l)$ 의 분석을 위한 신호의 길이이다. 그리고 다음의 식 (5)는 Basic-AMDF를 각  $\theta$ 만큼 회전시킨 변형된 AMDF를 나타낸다.

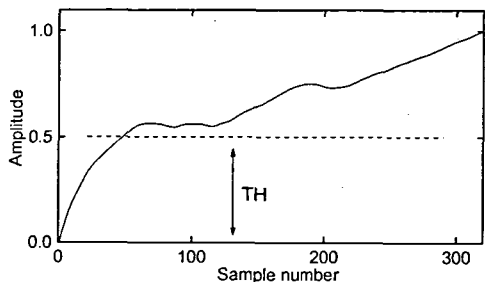
$$y_M(l) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \times y(l) \tag{5}$$

식 (5)에서 제안된 변형된 AMDF는 각  $\theta$ 를 이용하여 AMDF  $y(l)$ 를 변환시켜 주는 것에 의해, 피치 주기의 첫 번째 valley점에 대한 검출을 용이하게 한다.

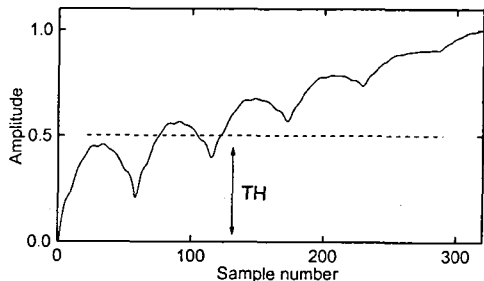
또한, 이러한 변형된 AMDF는 다음의 식 (6)과 같이 파라미터  $\alpha$ 를 추가하여, 회전각에 대한 유연성을 갖도록 하였으며, 그림 2의 (c)에서 그 영향을 나타내었다.

$$\theta = \arctan \frac{R_{\max}}{N} + \alpha \tag{6}$$

또한 음성의 시작구간에 대해서는 그림 3의 (a)와 같이 피치 주기 검출을 위한 최초 국부 최소 valley점이 명확하게 나타나지 않으므로써 에러가 발생하게 된다. 따라서 각 프레임에 threshold TH를 적용하여 TH보다 작은 크기를 가지는 valley점만을 피치 주기로 선정하였다. 이 때 TH는  $y(last) \times \beta$  ( $0 < \beta < 1$ )로 선택된다.



(a) 시작구간에서의 변형된 AMDF



(b) 정상구간에서의 변형된 AMDF

그림 3. 변형된 AMDF의 threshold TH

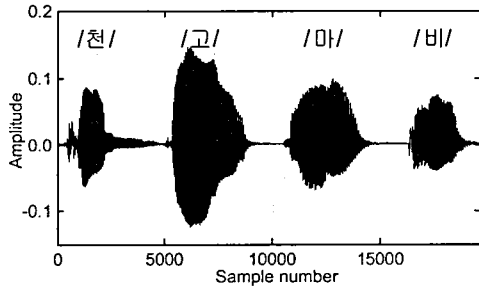
### III. 시뮬레이션 및 결과

본 논문에서 제안한 변형된 AMDF를 이용한 피치 주기 검출 알고리즘을 시뮬레이션하기 위해, IBM-PC의 사운드카드를 사용하였으며, 분해능은 16비트이고, 표본화 주파수는 8[kHz]로 설정하였다.

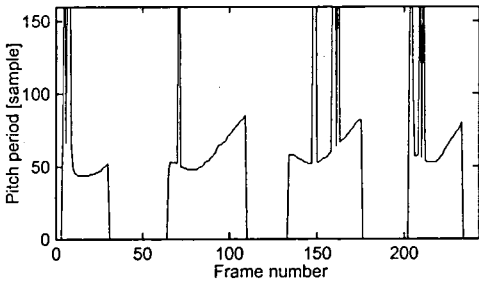
그리고 한 프레임의 길이는 10[ms](80표본)이며, 분석 프레임의 길이는 4프레임(320표본)으로 하였

고, 한 프레임씩 오버랩시켜 피치 주기를 검출하였다. 그림 4로부터 (a)는 /천고마비/로 발성된 테스트 음성 신호이며 (b)와 (c)는 각각 W-AMDF와 LV-AMDF에 의한 결과이고, (d)는 본 논문에서 제안한 변형된 AMDF( $\alpha = 0.2$ )에 의해 검출된 피치 주기를 나타낸 것이다.

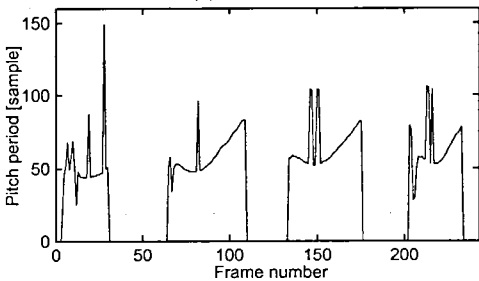
시뮬레이션 결과로부터, 기존의 방법들은 음성 신호의 피치 주기 검출을 위한 최초의 valley점 선정에서 많은 오류를 나타내었다. 그러나 본 논문에서 제안한 변형된 AMDF는 정확한 피치 검출 결과를 나타내었으며, 특히 음성 신호의 에너지만으로 구분하기 힘든 음성의 시작구간에 대한 오류를 threshold TH의 설정에 의해 제거함으로써 우수한 피치 주기 검출 결과를 나타내었다.



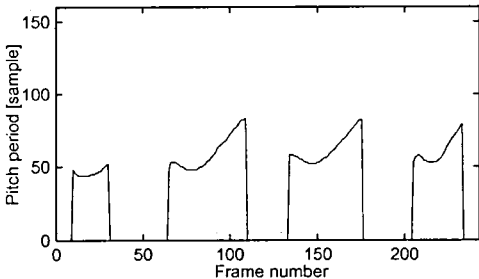
(a) Speech signal



(b) W-AMDF



(c) LV-AMDF



(d) Modified AMDF ( $\alpha = 0.2$ )

그림 4. 시뮬레이션 결과

### IV. 결론

본 논문에서는 음성의 피치 주기를 검출하기 위해, Basic-AMDF의 회전변환을 이용하여 전체 최소 valley점을 피치 주기로 선정하는 간단한 알고리즘을 제안하였으며, threshold TH를 설정하여 음성의 시작구간에 대한 오류를 제거시켰다.

그리고 시뮬레이션을 통해 변형된 AMDF 알고리즘을 기존의 W-AMDF 및 LV-AMDF와 비교하였으며, 결과로부터 제안한 방법은 기존의 방법들보다 우수한 성능을 나타내었다. 따라서 변형된 AMDF를 이용한 피치 주기 검출 알고리즘은 음성 합성, 음성 인식, 화자 식별등과 같은 다양한 음성 신호 처리 분야에 유용하게 적용되리라 사료된다.

### 참고문헌

- [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] S. Sneff, "Real Time harmonic pitch detection", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-26, pp. 358-365, 1978.
- [4] C. Manfredi, "Adaptive noise energy estimation in pathological speech signals", IEEE Trans. Biomedical Engineering, 47(11), pp. 1538-1543, 2000.
- [5] X. mei, J. Pan and S. Sun, "Efficient algorithms for speech pitch estimation", Proc. of ISIMVSP-2001, pp. 421-424, 2001.