
변형된 AMDF를 이용한 피치 주기 검출 알고리즘

서현수* · 배상범* · 김남호*

Pitch Period Detection Algorithm Using Modified AMDF

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

요 약

피치 주기는 음성 인식, 화자 식별, 음성 분석 및 합성 등과 같은 음성 신호 처리 분야에 있어서 중요한 요소이며, 이러한 피치 주기 검출에 관련된 다양한 알고리즘이 지금까지 연구되고 있다. 피치 검출에 사용되는 알고리즘의 하나인 AMDF(average magnitude difference function)는 각 계곡점의 거리를 피치 주기로 계산한다. 이때, 피치 주기 검출을 위한 계곡점 선정에 있어서 알고리즘이 복잡해지는 문제점이 발생한다. 따라서 본 논문에서는 AMDF의 회전 변환을 이용하여 전체 최소 계곡점을 음성 신호의 피치 주기로 검출하는 간단한 알고리즘을 제안하였으며, 시뮬레이션을 통해 기존의 방법들과 비교하였다.

ABSTRACT

Pitch period that is an important factor in speech signal processing is used in various applications such as speech recognition, speaker identification, speech analysis and synthesis. So many pitch detection algorithms have been studied until now. AMDF which is one of pitch period detection algorithms chooses the time interval from valley point to valley point as pitch period. In selection of valley point to detect pitch period, complexity of the algorithm is increased. So in this paper we proposed the simple algorithm using rotation transform of AMDF that detects global minimum valley point as pitch period of speech signal and compared it with existing methods through simulation.

키워드

pitch period, AMDF, valley point

I. 서 론

사회가 디지털 정보화 시대로 발달함에 따라, 음성 신호 처리에 관련된 다양한 연구들이 진행되고 있다. 그리고 음성 신호의 피치 주기는 음성 인식, 화자 식별, 음성 분석 및 합성 등의 여러 응용분야에서 필수적인 요소이다. 피치 주기는 인간의 청각에 매우 민감하게 반응하는 파라미터로써, 음성 신호의 화자를 구분하는데 사용하며, 자연성에 큰 영향을 미친다. 그러므로 음성 합성과 음

성 코딩에서 피치 주기의 정확한 추출과 복원은 음질에 중요한 역할을 한다[1]. 따라서 이러한 피치 검출을 위한 시간영역과 주파수영역에서의 많은 알고리즘이 제안되었다[2]-[3].

시간영역에서의 피치 검출 알고리즘은 파형의 주기적인 특징을 이용하여 피치를 검출하는 방법으로 병렬처리법, ACF(autocorrelation function), AMDF 등이 있다. 이러한 알고리즘은 분석을 위한 영역 변환이 불필요하며, 간단한 연산만을 사용하게 된다. 그러나 천이 구간내에 음

성이 존재하여 프레임 내에서 크기 변화가 크게 발생하거나 잡음이 중첩된 음성의 경우에는 피치 검출을 위한 알고리즘이 복잡해지는 단점을 가진다[7].

그리고 주파수영역에서의 피치 검출 알고리즘은 음성 스펙트럼의 고조파 간격을 측정하여 유성음의 기본주파수를 검출하는 방법으로 고조파분석법 등이 있다[4]. 일반적으로 프레임 내에서 주파수대역별로 신호의 분리가 가능하므로, 음성 주파수대역에 인접하여 잡음이 중첩되어 있거나 음성이 천이구간 내에 존재하는 경우에 강인한 특성을 가진다. 그러나 처리과정에서, 주파수영역으로 변환하기 위하여 복잡한 계산과정을 필요로 하며, 정확한 기본주파수의 검출을 위해 FFT의 포인트수를 증가시킬 경우 처리시간이 길어지고 변화특성이 둔해지게 된다. 그러므로 실시간 시스템에서는 적은 계산량과 정확한 피치 주기 검출의 장점을 가지는 AMDF가 널리 적용되고 있으나, 피치 주기 검출을 위한 계곡점 선정이 용이하지 않다.

따라서 본 논문에서는 AMDF의 회전변환을 이용한 새로운 알고리즘을 제시하여, 우수한 피치 주기 검출 특성을 나타내도록 하였으며, 시뮬레이션을 통해 기존의 방법들과 비교하였다.

II. AMDF 알고리즘

1. Basic-AMDF 알고리즘

음성신호 $s(n)$ 은 길이 N 의 창함수 $w(n)$ 과 연산을 통해 신호 $s_w(n)$ 을 만들 수 있다. 이때, Basic-AMDF $\gamma(l)$ 은 다음과 같이 정의된다.

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

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

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 유성음은 그림 1에서 보는 바와 같이, Basic-AMDF 알고리즘에 의해 국부 최소 계곡점이 주기적으로 A, B, C 등에서 나타나며, 전체 표본에 대해서 최초의 계곡점을 피치 주기로 계산한다. 따라서 그림 1에서, 최초 국부 최소 계곡점 A와 시작점 사이의 시간 간격이 피치 주기이다. 그러나 최초 국부 최소 계곡점이 전체 최소 계곡점이 되지 않으므로써 피치 주기 검출을 위한 최초 계곡점 선정에 어려움이 있다.

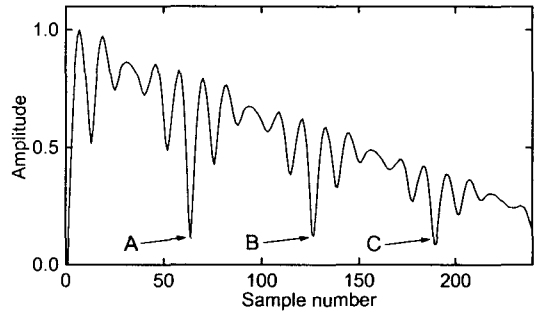


그림 1. 유성음에서의 Basic-AMDF
Fig. 1. Basic-AMDF in a voiced speech.

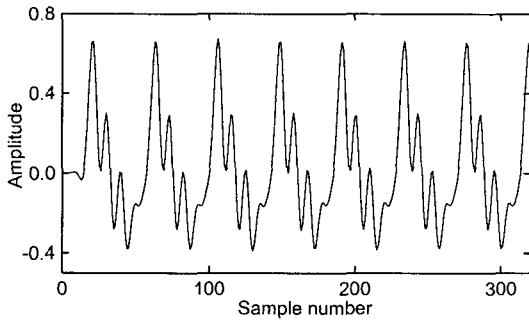
2. 기존의 AMDF 알고리즘

Basic-AMDF의 특성을 개선시키기 위해 다음과 같이, 선형 가중 AMDF(W-AMDF)와 가변 길 AMDF(LV-AMDF) 등이 소개되었다[5],[6]. 여기서, W-AMDF는 다음의 식 (2)와 같이 정의된다.

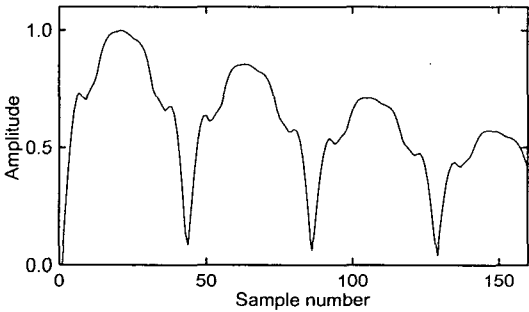
$$\gamma_w(l) = \frac{1}{N-l+1} \sum_{n=1}^{N-l+1} |s_w(n+l-1) - s_w(n)|, \quad (2)$$

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

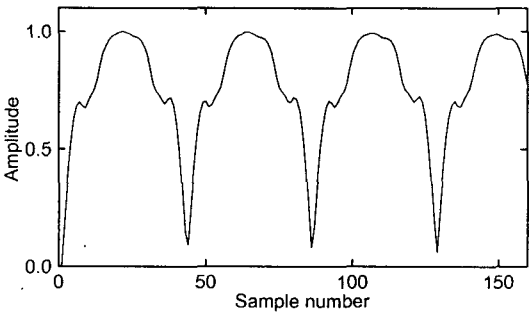
W-AMDF는 저음역의 잡음 환경에서 Basic-AMDF에서 나타나는 피치 검출 오류를 감소시키기 위해 제시되었다. 즉, 위의 식으로부터 스케일 인자 $(N-l+1)^{-1}$ 은 Basic-AMDF의 감소특성을 제거시켜준다. 일반적으로 AMDF의 최소값은 0에 가까운 값을 가지므로 기본주파수 즉, 피치를 검출하기 위해서, AMDF의 최소값으로 주어지는 적절한 l 을 선택해야 한다.



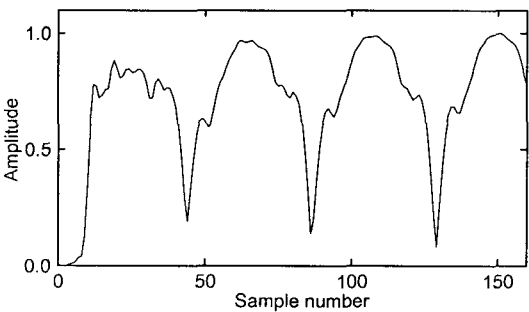
(a) Speech signal



(b) Basic-AMDF



(c) W-AMDF



(d) LV-AMDF

그림 2. 기존의 AMDF
Fig. 2. Existing AMDF.

이 때 주파수축 상에서의 기본주파수 f_0 의 위치가 표본화 주파수 F_s 에 의존하므로, l 은 F_s/f_0 의 관계로부터 얻을 수 있다.

그리고 LV-AMDF는 다음의 식 (3)과 같다.

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

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

식 (3)과 같이 LV-AMDF는 기존의 AMDF와 같은 주기성을 나타내며 합에 대한 연산량이 변수 l 에 의해서 결정됨에 따라, Basic-AMDF와 W-AMDF에 비해 적은 연산량을 가진다.

그림 2는 임의의 유성음으로부터 Basic-AMDF와 기존의 AMDF 알고리즘을 적용시킨 결과이다. 그러나 여전히 일반적인 유성음에서, 기존의 AMDF는 피치 주기 검출을 위한 최초 국부 최소 계곡점 검출에서 오류가 발생하고 있다. 따라서 기존의 AMDF는 피치 주기를 검출하기 위해 다음과 같은 추가적인 알고리즘을 필요로 하며, 각 임계값에 대한 도식적 표현은 그림 3과 같다[6].

- Step 1. 전체 최소점 Min+Th1을 만족하는 국부 최소 계곡점을 검출한다.
- Step 2. 시간 간격 임계값 Th2에 의해 만족하지 않는 값들을 제거해 나간다.
- Step 3. 계곡점의 깊이를 결정하는 임계값 Th3을 만족하는 가장 좌측의 값을 피치 주기로 선정한다.

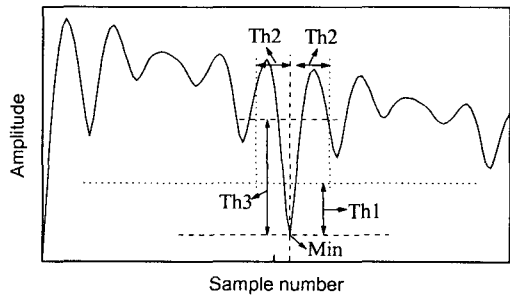


그림 3. AMDF의 최소점 검출
Fig. 3. Null picking of AMDF.

3. 제안한 변형된 AMDF

Basic-AMDF $y(l)$ 의 감소 특성을 분석함으로써, 다음의 그림 4에서 보여 지는 바와 같이 최초 피크 지점과 N 지점을 연결하는 직선에 의해 수평선과의 각 θ 를 설정할 수 있으며, 이러한 각 θ 는 각각의 프레임에서 적용적으로 구해진다. 그리고 이러한 각 θ 를 0° 로 만들기 위해, 그림 4의 (b)와 같이 양의 방향으로 Basic-AMDF를 회전시킨다.

다음의 식 (4)는 각 θ 를 설정하기 위해 유도된 식이다.

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

여기서 R_{max} 는 각 프레임에서, 분석을 위한 구간내의 최대값을 나타내며, N 은 프레임의 길이이다.

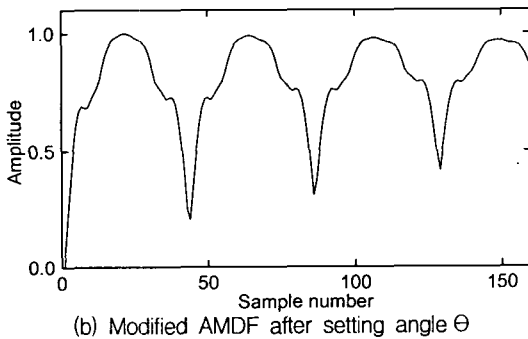
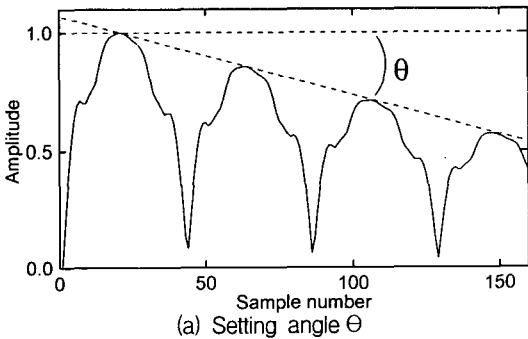


그림 4. 변형된 AMDF
Fig. 4. Modified AMDF.

그리고 다음의 식 (5)는 본 논문에서 제안한 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 제시된 변형된 AMDF는 각 θ 를 이용하여 Basic-AMDF $y(l)$ 를 회전 변환시켜 주는 것에 의해, 피치 주기 검출을 위한 첫 번째 계곡점에 대한 선정을 용이하게 한다. 따라서 피크값 이후의 최소값을 피치 주기로 선정할 수 있는 장점을 가질 수 있다.

또한 이러한 변형된 AMDF는 다음의 식 (6)과 같이 회전각에 대한 유연성을 갖도록 하기 위하여 파라미터 α 를 추가할 수 있다. 따라서, 적절한 α 값을 선정하는 것에 의해 피치 주기인 최초 계곡점과 나머지 계곡점과의 크기를 더욱 크게 설정할 수 있다.

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

다음의 그림 5는 식 (6)에서 제시된 파라미터 α 값을 0.1, 0.2, 0.3으로 설정하였을 경우에 대한 영향을 나타내고 있다. 그림으로부터 α 값이 증가함에 따라 계곡점 간의 크기 차이가 더욱 증가함을 알 수 있으며, 본 논문에서는 α 값을 실험값으로 0.3을 선택하였다.

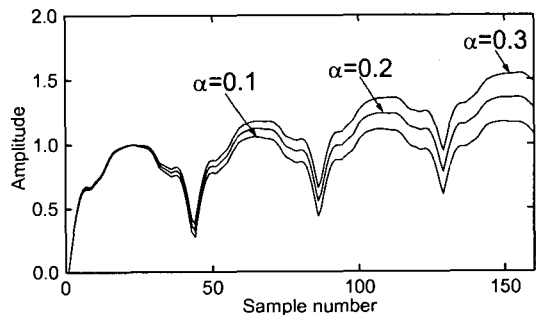


그림 5. 파라미터 alpha에 따른 변형된 AMDF
Fig. 5. Modified AMDF of parameter alpha.

III. 시뮬레이션 및 결과

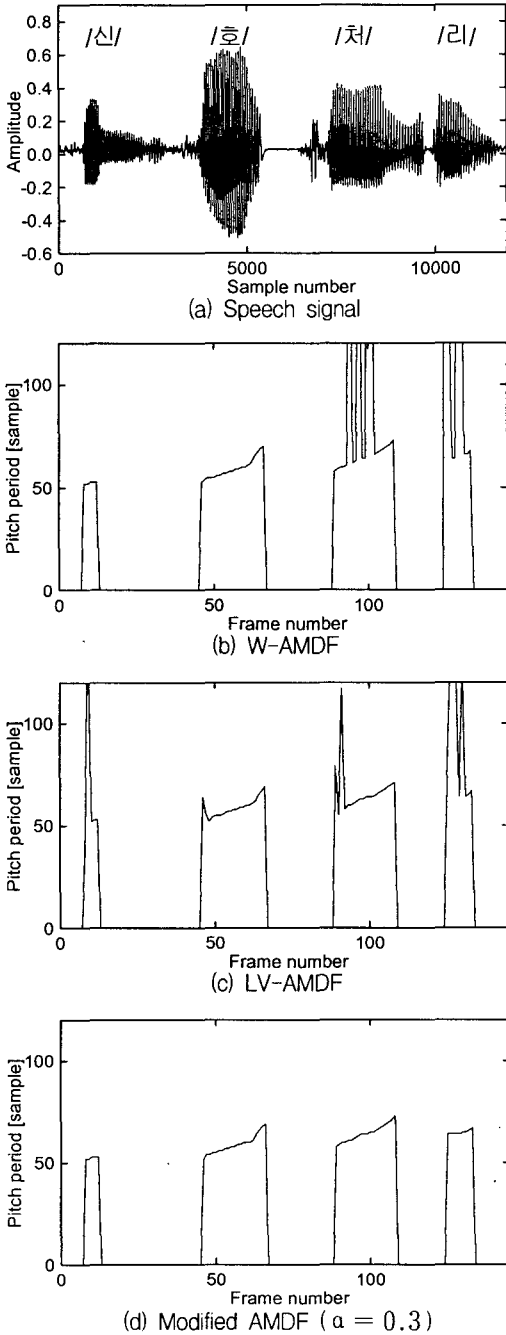


그림 6. 시뮬레이션 결과
Fig. 6. Simulation results.

본 논문에서 제안한 변형된 AMDF를 이용한 음성 신호의 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, 분석프레임의 길이는 4프레임(320표본)으로 하였고, 한 프레임씩 오버랩시켜 피치 주기를 검출하였다.

그림 6은 /신호처리/로 발생된 음성 신호에 대한 시뮬레이션 결과이다. 그림에서, (a)는 테스트 음성 신호를 나타내고 있으며, (b)는 W-AMDF, (c)는 LV-AMDF에 의한 결과이고, (d)는 변형된 AMDF ($\alpha = 0.3$)에 의해 검출된 피치 주기를 나타낸 것이다. 여러 AMDF에서 피치 주기 검출을 위한 계곡점 선정기준으로 전체 최소 계곡점을 피치 주기로 계산하였다.

시뮬레이션 결과로부터, 기존의 방법들은 음성 신호의 피치 주기 검출을 위한 최초의 계곡점 선정에서 많은 오류를 나타내었다. 그러나 본 논문에서 제안한 변형된 AMDF는 대부분의 프레임에서 우수한 피치 주기 검출 결과를 나타내었다.

표 1은 전체 최소 계곡점을 피치 주기로 설정하는 알고리즘 A와 피치 주기 검출을 위한 추가의 계곡점 선정 알고리즘 B를 각각 사용하여, 변형된 AMDF와 기존의 AMDF의 성능을 비교한 것이다. 결과의 수치는 세 가지 음성의 수동 검출 결과와 비교한 프레임 단위의 에러이다. 표로부터, 제안한 방법은 추가적인 알고리즘을 사용한 기존의 AMDF보다 우수한 결과를 나타내었다.

표 1. 세 가지 음성에 대한 에러 프레임 비교
Table 1. Comparison of error frames about three speeches.

| | | Speech 1 (75 frame) | Speech 2 (96 frame) | Speech 3 (65 frame) | Total (236 frame) |
|---------|---|------------------------|------------------------|------------------------|----------------------|
| W-AMDF | A | 14 | 11 | 11 | 36 |
| | B | 2 | 0 | 1 | 3 |
| LV-AMDF | A | 20 | 17 | 11 | 48 |
| | B | 10 | 8 | 4 | 22 |
| M-AMDF | | 1 | 0 | 1 | 2 |

IV. 결 론

본 논문에서는 음성의 피치 주기를 검출하기 위해, Basic-AMDF의 회전변환을 이용한 변형된AMDF 알고리즘을 제시하여 기존의 W-AMDF 및 LV-AMDF와 비교하였다.

기존의 AMDF에서는 국부 최소 계곡점이 주기적으로 잘 나타나지만 피치주기 검출을 위한 최초 국부 계곡점 선정이 쉽지 않은 단점을 가진다. 그러나 제안한 변형된 AMDF은 전체 최소 계곡점을 피치 주기로 계산하는 간단한 알고리즘으로 피치 주기를 검출하는 장점을 가진다.

시뮬레이션 결과, 기존의 방법들은 피치 주기 검출을 위한 최초 계곡점 선정에서 많은 오류를 보였으나, 본 논문에서 제안한 새로운 알고리즘은 우수한 검출 결과를 나타내었다.

따라서 제안한 방법은 향후 수행되는 유·무성음 판별 알고리즘 등과의 접목을 통해 음성 합성, 음성 인식, 화자 식별 등과 같은 음성 신호 처리의 다양한 분야의 유용하게 적용되리라 사료된다.

참고문헌

[1] S. Shukla, A. E. Ertan and T. P. Barnwell III, "Circular LPC modeling and constant pitch transform for accurate speech analysis and high quality speech synthesis", in Proc. IEEE International Conference on Acoust., Speech and Signal Processing, pp. I269-I272, 2002.

[2] 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.

[3] 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.

[4] S. Sneff, "Real Time harmonic pitch detection", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-26, pp. 358-365, 1978.

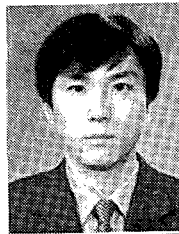
[5] X. mei, J. Pan and S. Sun, "Efficient algorithms for speech pitch estimation", Proc. of ISIMVSP-2001, pp. 421-424, 2001.

[6] W Hess, "Pitch determination of speech signals: algorithms and devices", Berlin-Verlag, 1983.

[7] 배상범, 김남호, "웨이브렛 변환쌍과 적응-길이 매디안 필터를 이용한 임펄스 노이즈 제거에 관한 연구", 한국해양정보통신학회논문지, 제7권 제7호, pp. 1575-1581, 2003.

저자소개

서현수(Hyun-SooSeo)



2005년 2월 부경대학교 공과대학 제어계측공학과 공학사
2005년 3월~현재 부경대학교 대학원 제어계측공학과 석사과정

※ 관심분야: 통신 및 신호처리

배상범(Sang-BumBae)



1999년 8월 부경대학교 공과대학 제어계측공학과 공학사
2002년 2월~현재 부경대학교 대학원 제어계측공학과 공학석사

2002년 3월~현재 부경대학교 대학원 제어계측공학과 박사과정

※ 관심분야: 통신 및 신호처리

김남호(Nam-HoKim)



제8권 제3호 참조
1992년 3월~현재 부경대학교 공과대학 전기제어계측공학부 정교수

※ 관심분야: 신호처리, 회로 및 시스템