

A Study of Peak Finding Algorithms for the Autocorrelation Function of Speech Signal

Shin-Ae So*, Kang-Hee Lee**, Kwang-Bock You***, Ha-Young Lim****, Ji Su Park*****

Abstract

In this paper, the peak finding algorithms corresponding to the Autocorrelation Function (ACF), which are widely exploited for detecting the pitch of voiced signal, are proposed.

According to various researchers, it is well known fact that the estimation of fundamental frequency (F0) in speech signal is not only very important task but quite difficult mission. The proposed algorithms, presented in this paper, are implemented by using many characteristics - such as monotonic increasing function - of ACF function. Thus, the proposed algorithms may be able to estimate both reliable and correct the fundamental frequency as long as the autocorrelation function of speech signal is accurate. Since the proposed algorithms may reduce the computational complexity it can be applied to the real-time processing.

The speech data, is composed of Korean emotion expressed words, is used for evaluation of their performance. The pitches are measured to compare the performance of proposed algorithms.

- ▶ Keyword: peak finding algorithms, autocorrelation function, fundamental frequency (F0), monotonic increasing function, speech signal

1. Introduction

공학 그리고 과학적 데이터 처리에서 많은 경우 신호의 peaks 들을 검출하고 이것들의 위치, 크기 등을 측정하는 것이 필요하다. 신호의 peak을 검출한다는 것은 해당 신호의 local 최대값들과 최소값들의 위치와 크기를 찾는 과정이라고 정의할 수 있다. 그래서 peak을 검출한다는 것은 신호의 상당한 정보를 찾는 것이다. Peak을 검출하는 일반적인 방법은 peak에서의 미분 (first derivative)이 그 최고점에서 영교차 (zero-crossing)를 만든다는 사실을 이용하는 것이다. 그러나, 실제의 신호 - 음성 신호 같은 - 에서는 random noise가 항상 존재하기에 이로 인한 많은 잘못된 영교차 (false 음성 처리 분야에서 피치 (pitch) 주기를 정확히

추정(estimation)하기 어렵다는 것은 잘 알려진 사실이다. 이는 음성 신호가 유성음 구간에서 피치 주기의 정확한 시작과 끝을 정의하기가 어렵다는 것과 준 주기적인 (quasi-periodic) 신호로 시간에 따라 변하기 때문이다. 그래도 피치의 주기성은 ACF의 기본 lag (지연)의 정수배마다 신호가 반복되는 것으로 알 수 있다. 피치에는 피치 주기에 나타나는 peak의 위치와 크기 정보 그리고 피치가 잘 검출되었음을 보여주는 증거가 되는 것으로 다른 peak들은 이 peak의 정수배에 위치한다는 정보가 있다 [14], [15]. 본 논문은 피치 검출 함수들에서 음성 신호의 피치를 정확하고 신뢰성 있게 추정할 수 있는 peak finding 알고리즘들을 연구하는 것을 목적으로 하였다. 일반적으로 peak 검출 알고리즘을 정확히 사용하려면 처리 할 신호에 대한 특징들을 알고 있어야 한다. 예를 들면, 주기적인지, Nyquist rate로 샘플 하였는지, 혹

• First Author: Shin-Ae So, Corresponding Author: Kang-Hee Lee

*Shin-Ae So (hetaira@ssu.ac.kr), Dept. Korean Language and Literature, Soongsil University

**Kang-Hee Lee (kanghee.lee@ssu.ac.kr), Dept. of Digital Media, Soongsil University

***Kwang-Bock You (kwangbockyou@ssu.ac.kr), School of Electronic Engineering, Soongsil University

****Ha-Young Lim (hayoung0631@naver.com), School of Electronic Engineering, Soongsil University

*****Ji Su Park (iwksuld@hanmail.net), School of Electronic Engineering, Soongsil University

• Received: 2016. 11. 01, Revised: 2016. 11. 30, Accepted: 2016. 12. 28.

• This work was supported by the Soongsil University Research Fund of 2013.

• This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A5A2A03927522)

은 신호가 noisy인지 등이 있다. 그래서, 본 논문에서 제시하는 알고리즘들은 피치 검출 함수의 특성들을 최대한 이용하였다.

본 논문은 다음과 같이 구성되어 있다.

2장에서는 널리 사용되고 있는 time-domain 방식에 기초한 피치 검출 함수인 Autocorrelation Function (ACF)에 대하여 알아보았다.

3장에서는 먼저 일반적인 의미에 기초하여 peak finding algorithm이 무엇인지 알아보고 2장에서 살펴 본 ACF 함수의 특성을 이용하여 효과적인 방법으로 peak (maxima)를 찾는 알고리즘들을 제시하였다. 또 동일한 음성 data를 사용하여 제시한 알고리즘들을 서로 비교해 보았다.

4장에서 본 논문에서 사용한 한국어 제시문의 구성과 그 과정을 설명하고, 그 음성 data를 이용하여 3장에서 제시한 알고리즘들에 대하여 수행한 Matlab을 이용한 시뮬레이션 결과들을 보이고 또 본 논문에서 제안한 알고리즘으로 피치를 추정하였다.

마지막으로 5장에서 본 논문의 결론으로 4장에서 측정한 피치들로서 본 논문이 제안한 알고리즘들의 성능을 비교하였고 향후에 더 연구가 필요한 것들에 대하여 설명하였다.

II. Autocorrelation Function (ACF)

Autocorrelation Function은 음성 신호의 피치를 구하는 방법으로 널리 사용되고 있다.

음성 신호는 Wide Sense Stationary (WSS) 이기 때문에 이의 Autocorrelation은 아래의 식으로 정의된다 [1], [2], [5], [6], [7].

$$R(m) = \lim_{N \rightarrow \infty} \frac{1}{(2n+1)} \sum_{n=-N}^N x(n)x(n+m), \quad (0 \leq m \leq M_0) \quad (1)$$

여기서, N 은 분석 할 sequence의 길이이며 M_0 는 계산 할 autocorrelation 점의 개수이다.

이 함수는 임의의 시간 n 에서의 신호 값 $x(n)$ 과 m 만큼의 시간 지연이 있을 때 시간 $n+m$ 에서의 신호 값 $x(n+m)$ 의 곱에 대한 평균으로 설명할 수 있다. 여기서, 우리는 피치 검출을 위해서 voiced 신호 $x(n)$ 이 주기적이라 할 수 있고, 즉, $x(n) = x(n+p)$ 모든 n 에 대해서, 따라서 이에 대한 autocorrelation function 또한 같은 주기를 갖게 된다, 즉, $R(m) = R(m+p)$ 이다. 이외에도 autocorrelation function은 아래와 같은 중요한 특성들이 있다.

1. $R(m) = R(-m)$, 즉 Even function이다.
2. $m=0$ 에서 최댓값을 가진다.

매우 짧은 구간동안에 Pitch Detection Algorithms (PDAs)을 적용하기에 위의 (1)식을 이에 맞춰 다시 쓰면 아래와 같다.

$$R(m) = \frac{1}{N} \sum_{n=0}^{N-1-m} x(n)x(n+m), \quad (0 \leq m \leq M_0) \quad (2)$$

여기서 m 은 lag 혹은 지연 (delay) 이라 부른다.

그림 1 은 한국어 문장 “우리들만 먼저 올라가자”를 여성 화자가 발성한 음성 신호를 16kHz로 샘플링 하였다. 위의 문장에서 “가자”를 “ㄱ” “ㄴ” “ㄷ” “ㄹ”로 분리하여 순서대로 시간 축에서 보여준 것이다.

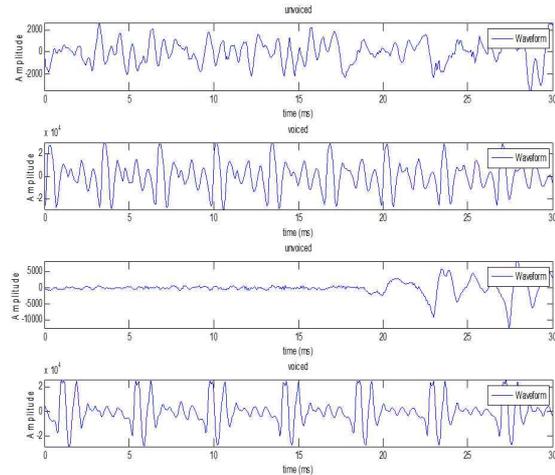


Fig. 1. original speech signal

이 신호들을 아래의 식들로 정규화 (normalizing) 하고 Hamming window로 전처리한 신호들을 그림 2에서 보여준다 [3], [4], [5].

$$\begin{aligned} \text{Sample Mean} : \mu_x &= \frac{\sum_{n=1}^N x(n)}{N} \\ \text{Sample Variance} : \sigma_x^2 &= \frac{\sum_{n=1}^N \{x(n) - \mu_x\}^2}{N} \\ \text{Normalized Data} : x(\hat{n}) &= \frac{x(n) - \mu_x}{\sigma_x} \end{aligned} \quad (3)$$

여기서, $x(n)$ 은 16 kHz로 샘플 된 음성 신호이고 N 은 한 프레임내의 총 샘플 수이다.

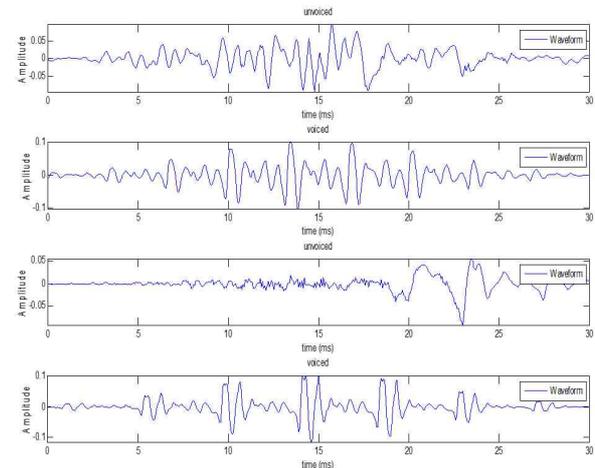


Fig. 2. Preprocessed speech signal

그림 2의 음성 신호에 식 (2)의 autocorrelation function을 적용한 것을 그림3에 나타냈다. 첫 번째와 세 번째 행의 무성음은 어떤 주기성도 없음을 알 수 있고 반면에 두 번째와 네 번째의 유성음은 상대적으로 주기성이 있음을 알 수 있다.

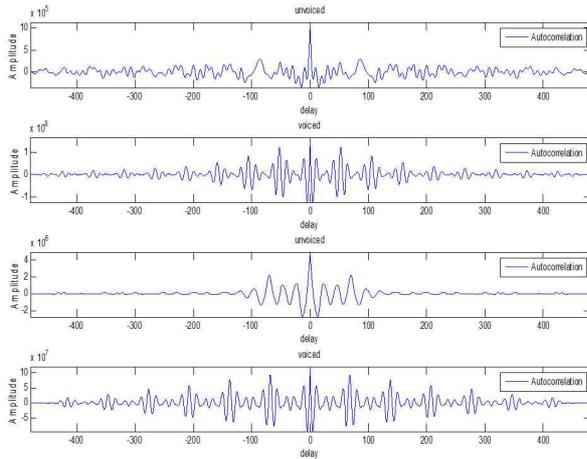


Fig. 3. Autocorrelated speech signal

III. Peak Finding Algorithms

1. Peak Finding Algorithms 현황

n 개의 elements를 가진 어떤 data set의 array $A[0 \dots n-1]$ 에서 어떤 element, $A[i]$ 가 그 주변에서 그것보다 큰 elements가 없는, 즉, $A[i] \geq A[i-1]$ 이고 $A[i] \geq A[i+1]$ 가 되는 peak $A[i]$ 의 index i 와 그 크기 혹은 관계된 정보를 구할 수 있는 어떤 알고리즘들을 peak finding algorithm이라고 정의 할 수 있다. 이런 형태의 알고리즘들은 주어진 임의의 조건들을 만족하는 모든 peaks들을 구하게 된다. 주어진 임의의 조건들만 충족되면 peaks를 찾는 알고리즘들은 무차별 대입 방법(Brute Force)으로 전 영역의 모든 데이터 집합의 요소들의 peak 정도 여부를 테스트한다. 즉,

```

for  $i$  in range ( $n$ ) :
    if  $A[i-1] \leq A[i] \geq A[i+1]$ :
        then return  $i$ 
    
```

이에 비해, MatLab의 Signal Processing Toolbox에 있는 “findpeaks”와 같은 알고리즘은 좀 더 진화된 형태로 일정 크기 이상 그리고 어느 정도의 최소 거리의 유지 등 조건들을 만족하는 data set의 모든 peaks들의 value (값) 와 indices를 구한다.

이외에도 2차 함수에서 꼭지점을 찾는 것과 같이 concave down에서 어떤 특별한 영역의 평균 곡률 (average curvature)

로 구하는 방식들 여러 알고리즘들이 있다. 이런 알고리즘들이 정확히 응용되려면 처리 할 신호에 대한 특징 (특성)들을 이용해야 한다는 것이다.

2. Peak Finding Algorithms for ACF

2.1 Three Samples' Conventional Method with Certain Threshold

그림4는 남성 화자의 유성음 “아”에 hamming windowing을 한 480 샘플들의 autocorrelation function이다. 앞에서 설명하였듯이 ACF는 $R(0)$ 을 기준으로 even함수가 되며 $R(0)$ 이 최대값이다. 이 최대값을 중심으로 어느 정도 일정한 시간 지연 (time delay) 마다 $R(m)$ 의 값이 상대적인 (relatively) peak들이 보여진다. 이런 peak들 사이의 거리를 음성 신호의 피치 주기 (pitch period)라 한다.

신호를 그 최대값으로 나누는 간단한 정규화를 한 후, 3개의 샘플, x_{n-1}, x_n, x_{n+1} ,을 비교하여 가운데 샘플이 양 옆의 샘플보다 큰것들을 찾고 이를 peak라 한다. 즉, $x_{n-1} < x_n > x_{n+1}$ 가 되는 샘플 x_n 을 peak으로 정의하고 이를 찾는 것이다. 이의 과정은 위에서 설명한 MatLab의 함수 “findpeaks” 알고리즘을 이용하면 된다.

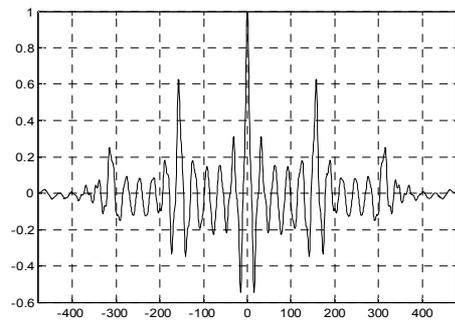


Fig. 4. Autocorrelation function for voiced signal "아" of a male speaker

이렇게 하여 찾은 모든 peaks들을 그림 5에 보였다. 보는 것과 같이 음성 신호의 피치를 구할 수 있는 지배적인(dominant) 5개의 peaks이외에도 상당히 많은 peaks들이 표시 되어 있다. 이는 이 알고리즘에서 사용한 MatLab의 “findpeaks” 함수가 3개의 샘플들의 관계를 만족하는 모든 peaks들을 찾게 되기 때문이다. 그래서, 분석할 음성 신호의 ACF를 구해서 peaks들 사이의 거리로 이 신호의 피치를 구하면 그 오차가 크게 된다는 것을 알 수 있다. 이 문제를 해결하는 한 가지 방법은 일정한 임계값(threshold)을 잡는 것으로, 보통은 최대치의 일정 비율로 하거나 혹은 신호의 통계적 특성 그리고 그 성질 - 기대치, 분산, standard deviation, 등 혹은 이들의 조합 - 을 사용한다.

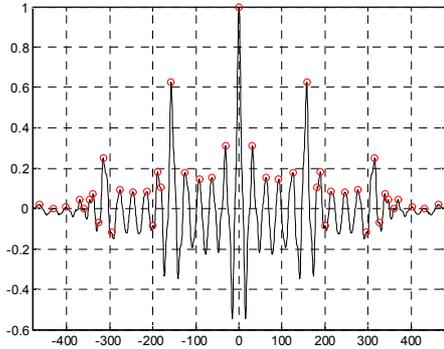


Fig. 5. All peaks of ACF

본 논문에서 제안하는 알고리즘은 Threshold를 활용하기 전에 그림 5에서 보인 모든 peaks들에 대해 한번 더 “findpeaks”를 수행해서 그 peaks들의 수를 줄인다. 제안하는 알고리즘은 위에서 언급한 신호의 성질을 이용한 어떤 상수 값으로 그 임계값을 정하는 대신에 ACF의 특성을 이용하였다. ACF의 peaks들은 $R(0)$, 즉, 중앙에 가장 큰 값이 나타나고 양 옆으로 갈수록 그 값들이 감소하는 단조 감소하는 특성을 갖는다. 따라서 false peaks들을 제거하기 위해 첫번째 샘플부터 최댓값인 $R(0)$ 까지 증가하는 값들만 찾는다. 그 결과를 그림 6에 보여 준다. 그림 6에서 보이는 것과 같이 ACF에서 신호의 피치를 매우 정확히 검출 할 수 있도록 중요한 5개의 peaks를 찾았다. 물론, 양쪽 끝에 여전히 오차를 유발 할 수 있는 peaks들이 있다. 그러나, 이런 peaks들은 피치를 측정하는 데는 사용하지 않으므로 제안 하는 알고리즘의 성능에는 영향이 없다. 지금까지 살펴 본 알고리즘은 conventional한 방법에 threshold 효과가 나타나도록 ACF의 특성을 이용하였다.

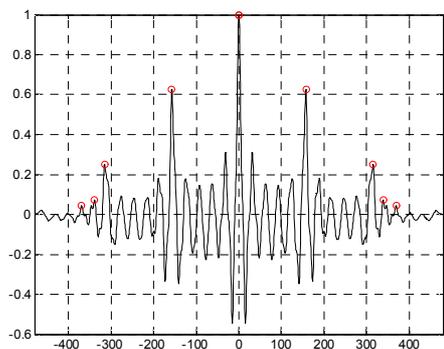


Fig. 6. Major peaks after performed the proposed algorithm

2.2 First-Rising-Slope Points (FRSP) Algorithm

그림7에서 첫 번째 행은 여성화자의 음성신호 “아”의 ACF를 보였고, 두 번째는 그것의 일부분을 확대한 것이다. 여기서 빨간 점으로 표시한 위치를 First-Rising-Slope Points (FRSP)로 정의하였다. 이 점들은 local minimum에서 다음 local minimum 사이의 m개의 점들로 구성된 cluster에서 - 분

석할 ACF에서 n개의 이런 clusters가 있다 - 그 크기 (Amplitude)가 0 (zero) 보다 크고 그것의 기울기가 positive 인 첫 번째 point이다. 즉, i번째 cluster, $S_i\{x_i(k)\}$, where $i = 1, \dots, n$ and $k = 1, \dots, m$ 에서 $x_i(k) > 0$ 이며 $\frac{dx_i(k)}{dt} > 0$ 이 되는 최초의 샘플들이다.

먼저, FSRP들을 모두 구한 후에 각 $S_i\{x_i(k)\}$ 에서 peaks (즉, $x_i(k_{peak})$, where $i = 1, \dots, n$ and $k = 1, \dots, m$)들을 구한다. ACF는 even function이므로 양단의 어느 방향에서든 시작해서 피치를 구할 수 있다. 각 clusters들에서 구한 peaks, $x_i(k_{peak})$ 들에서 $x_1(k_{peak})$ 과 $x_2(k_{peak})$ 를 비교해서 만일 $x_1(k_{peak}) > x_2(k_{peak})$ 이면 $x_2(k_{peak})$ 는 false peak으로 취급하고, 만일 그 반대라면 이 두 개의 peaks들 모두 피치를 구성한다. 이 계산을 ACF의 $R(0)$ 까지 해서 단조 증가하는 peaks들을 선별해서,

$$x_1(k_{peak}) < x_2(k_{peak}) < x_3(k_{peak}) < \dots < R(0)$$

되는 피치를 구한다.

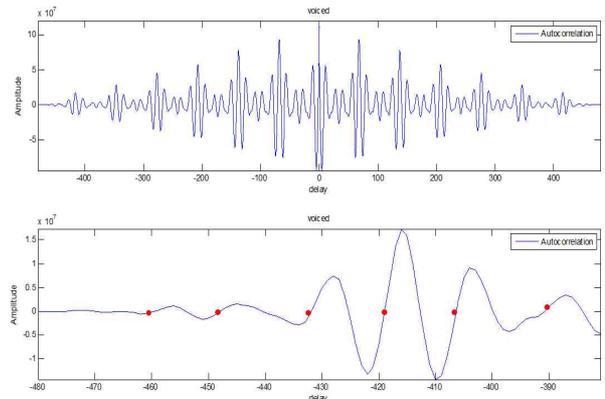


Fig. 7. Autocorrelation function for voiced signal “아” of a female speaker

그림 8은 그림 7에서 보인 ACF에 본 논문에서 제안한 FRSP 알고리즘을 적용하여 피치를 측정하는데 필요한 peaks들을 구한 것을 보여준다.

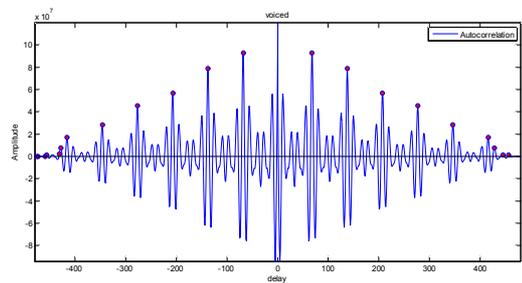


Fig. 8. The Results of applying FRSP algorithm to the ACF of Fig. 7

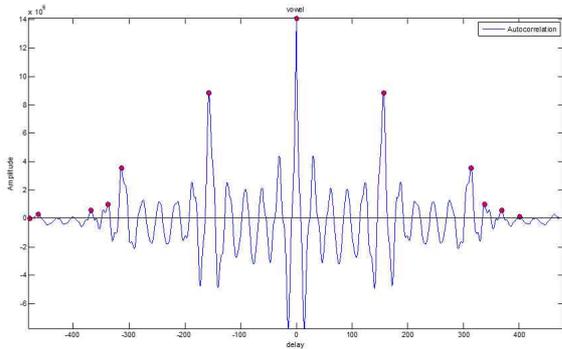


Fig. 9. Comparison with Fig. 6

그림 9는 앞에서 제안한 ‘Three Samples’ Conventional Method with Certain Threshold 알고리즘과 비교를 위해 그림 6과 같은 data를 사용해서 그 결과를 보여주고 있다.

IV. Simulation Results and Discussion

1. 녹음 환경과 데이터의 선택 (Recording Environments and its data)

본 논문에서 사용한 data는 한국어의 감정 표현 언어들로 기쁨, 슬픔, 화남, 놀람, 그리고 중립의 5가지 범주에서 녹음을 한 것이다. 스마트 폰이나 디지털 녹음기 같은 일반적인 기구(device)로 녹음을 하였다.

평서문, 의문문 (판정, 설명, 그리고 확인), 명령문, 그리고 청유문의 4가지 문형 - 특히, 의문문은 3가지로 세분해서 - 으로 구성하여 자연스러운 감정 표현을 할 수 있도록 대화문 양식으로 제시문을 구성하였다. 억양 곡선이 잘 드러나도록 공명음을 포함한 어휘로 구성하였고, 화자의 부자연스러움과 얼버무림을 배제하기 위해서 시작과 끝부분에 채움 문장을 두었다 [8-11].

각 제시문은 서로 다른 지역의 사람들로 - 여성과 남성, 연령 별 (주로 20대에서 40대까지로 하였다) - 녹음을 하였다.

이런 전체적인 구성을 가진 제시문에서 본 논문은 남성과 여성 화자의 중립, 기쁨, 그리고 슬픔의 세 가지 감정 표현 언어를 사용하여 본 논문이 제안하는 알고리즘들에 적용하였다. 우리가 분석에 사용한 문장은 “우리들만 먼저 내려가자.”으로 청유문이다.

2. Simulation Results

본 논문에서 제안한 peak finding 알고리즘들은 그 목적이 음성 신호의 피치를 좀 더 정확하고 효과적으로 검출하는 것이다. 이런 이유로 선택한 음성 신호들의 피치들을 시뮬레이션으로 측정한 결과들을 표1에 정리 하였다.

Table 1. The measured pitches by using the proposed algorithms

Gender	Emotion	Peak Finding Algorithm	
		Certain Threshold	First-Rising-Slope Points
Female	기쁨	4.375	4.375
	슬픔	6.5	6.5
	중립	5.625	5.625
Male	기쁨	3.375	3.375
	슬픔	6.875	6.875
	중립	5.8125	5.8125

표1은 전통적인 Certain Threshold 알고리즘(2.1절)과 비교하여 본 논문 (2.2절)에서 제안하는 First-Rising-Slope 알고리즘에서 측정된 값들이 그 차이가 없음을 보여주고 있다. 이 결과들은 제안된 연계 알고리즘들이 안정적이고, 정확한 peak 검출 성능을 가지고 있음을 보여준다.

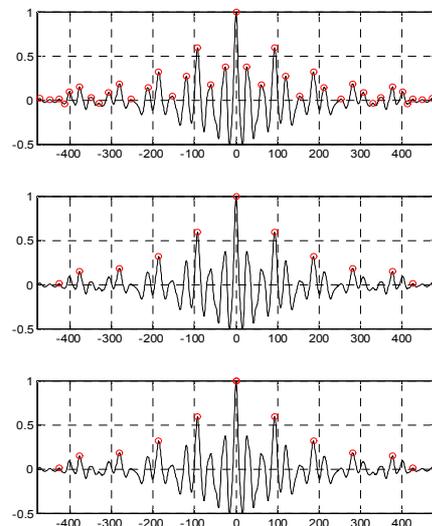


Fig. 10. Three Samples’ Conventional Method with Certain Threshold - Male

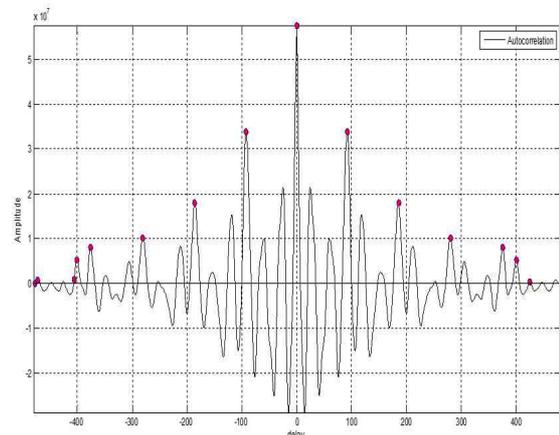


Fig. 11. First-Rising-Slope Points (FRSP) Algorithm - Male

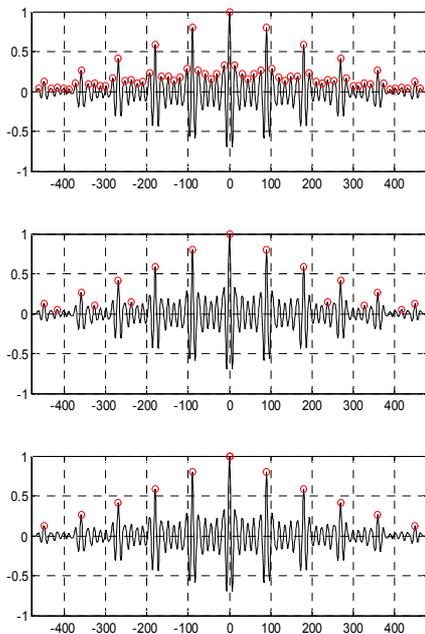


Fig. 12. Three Samples' Conventional Method with Certain Threshold - Female

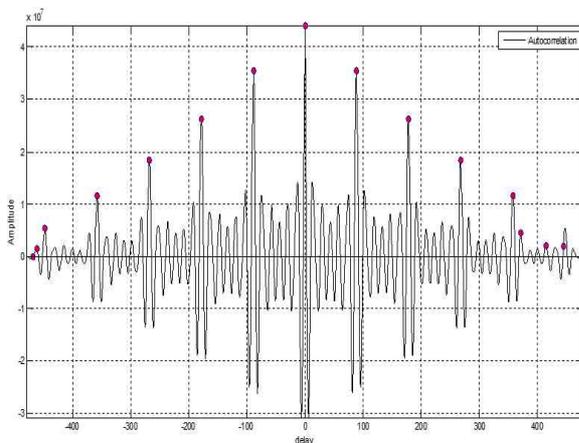


Fig. 13. First-Rising-Slope Points (FRSP) Algorithm - Female

그림 10에서 그림 13은 울산 지역에 사는 남자의 중립 감정과 인천 지역에 사는 여성의 중립 감정의 음성 신호에 본 논문이 제안하는 두 가지 알고리즘을 적용하는 과정과 그 결과들을 보여주고 있다.

V. Conclusion

본 논문은 time 도메인에서 피치 검출에 많이 사용하는 함수인 ACF에 대하여 서로 다른 두 종류의 peak finding algorithms들을 제안하였고 이들의 성능을 test하기 위해서 남

성과 여성의 서로 다른 감정 표현 언어에 대하여 피치를 구하여서 비교 하였다.

4장에서 보인 simulation의 결과들은 본 논문에서 제안하는 알고리즘들의 우수한 성능을 보여주고 있다. 피치 측정값들이 전통적인 방법과 차이가 없는 정확한 검출 성능을 보여준다. 특별히 본 simulation에서 이용한 유성음과 무성음을 구분하는 알고리즘들도 본 연구와 관계하여 개발한 것들을 사용하였다. 본 simulation의 결과는 이 유,무성음 구분 알고리즘들도 또한 우수한 성능을 가짐을 보여준다. 향후에 이와 관련한 내용을 설명하도록 하겠다.

제안하는 알고리즘들은 피치 검출 함수의 특징들을 잘 이용하여 하도록 설계를 하여서 그의 계산량이 줄어들어 real-time processing에 적용하기에 매우 유리하다. 본 논문에서 제안한 알고리즘들은 time domain의 피치 검출 함수에만 적용 가능하기에 향후에 주파수 도메인에서의 피치 검출 함수인 Cepstrum에 대해서도 적용 할 수 있는 알고리즘을 개발하여서 서로를 비교 분석 할 필요가 있다 [12-13].

본 논문에서 제안하는 알고리즘들은 각 남성 화자와 여성 화자의 서로 다른 세 가지 감정 표현 언어에 적용하는 실험을 하였다. 여러 제시문들에서 한가지에만 적용하였기에 데이터가 적었다고 할 수 있다. 향후에 더 많은 데이터에 적용하여서 비교 분석을 할 필요가 있다.

REFERENCES

- [1] Savitha S Upadhy, "Pitch Detection in Time and Frequency Domain," ICCICT, 2012.
- [2] E. Verteletskaia and B. Simak, "Performance Evaluation of Pitch Detection Algorithms," <http://access.feld.cvut.cz/service.php>
- [3] A. Leon-Garcia, Probability, Statistics, and Random Processes for Electrical Engineering, 3rd Edition, Prentice Hall, 2009.
- [4] A.M. Kondoz, Digital Speech (Coding for Low Bit Rate Communications Systems), John Wiley & Sons Ltd., 1994.
- [5] L.R. Rabiner and R.W. Schafer, Theory and Applications of Digital Speech Processing, 1st edition, Prentice Hall, 2011
- [6] J.R. Deller, JR., J.G. Proakis, and J.H.L. Hansen, Discrete-Time Processing of Speech Signals, Macmillan, 1993
- [7] Myron J. Ross, Harry L. Shaffer, A. Cohen, R. Freudberg, and H. Manley, "Average Magnitude Difference Function Pitch Extractor," IEEE Trans. On ASSP, Vol., ASSP-22, No. 5, October 1974.

- [8] Lee, Ho-Young, The Structure of Korean Prosody, PhD dissertation, University of London, 1990.
- [9] Jun, Sun-Ah, The phonetics and phonology of Korean prosody, Ph.D. dissertation, Ohio State University, 1993.
- [10] Jun, Sun-Ah, K-ToBI labelling conventions(version3.1), UCLA Working Papers in Phonetics 99, pp.149-173, 2000.
- [11] Jin-Koo Ji and Sung-il Yun, Design and Implementation of Speaker Verification System Using Voice, Journal of the Korea Society of Computer & Information, Vol. 5, No. 3, pp. 91-98, 2000.
- [12] Celia Shahnaz, Wei-Ping Zhu, and M. Omair Ahmad, "Pitch Estimation Based on a Harmonic Sinusoidal Autocorrelation Model and a Time-Domain Matching Scheme," IEEE Transactions on Audio, Speech, and Language Processing, Vol. 20, No. 1, pp. 322-335, 2012.
- [13] Seo Weon Heo, Don Hur, and Hyungsuk Kim, "Optimization of the autocorrelation weighting function for the time-domain calculation of spectral centroids," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 62, No. 3, pp. 421-427, 2015.
- [14] Sebastian Kraft and Udo Zolzer, "Polyphonic Pitch Detection by Matching Spectral and Autocorrelation Peaks," 2015 EUSIPCO
- [15] Jun Gao and Dan Xu, "Noise-robust Pitch Detection Algorithm Based on AMDF with Clustering Analysis Picking Peaks," 2016 IEEE Information Technology, Networking, Electronic and Automation Control Conference

Authors



Shin-Ae So received the B.S., M.S. and Ph.D. degrees in Korean Language and Literature from Sogang University, Korea, in 1999, 2002 and 2006, respectively.

Dr. So is currently an Assistant Professor in the Department of Korean Language and Literature, Soongsil University. She is interested in Korean phonology, Korean dialectology, and linguistic variation and change.



Kang-Hee Lee received the B.S., M.S. and Ph.D. degrees in EECS from KAIST, Korea, in 1999, 2001 and 2006, respectively.

Dr Lee is currently an Associate Professor in the Department of Digital Media, Soongsil University. He is interested in robotic intelligence & emotion and media robotics.



Kwang-Bock You received the Ph.D., in ECE from Stevens Institute of Technology, Hoboken, NJ., USA, in 1998.

Dr. You joined the School of Electronic Engineering, Soongsil University as an Associate Professor in 2010. He is interested in speech signal processing, sampling theorem, and wireless communications.



Ha-young Lim received the B.S. degrees in Information Communication Electronic Engineering from Soongsil University, Korea, in 2016. Ms Lim joined FADU Corporation, Seoul, Korea, in December, 2016 as an FPGA engineer.



Ji Su Park received the B.S degrees in information Communication Electronic Engineering from Soongsil University, Korea, in 2016. Ms. Park joined FADU Corporation, Seoul, Korea, in January,

2016. Currently, she is working as a Firmware engineer.