

음성 신호의 피치 검출에 관한 알고리즘의 성능 비교

84302

김 미현, 유 광복, 이 광형
 송전대학교 전자공학과

Performance Comparison of Several Pitch
 Detection Algorithms in Speech Signal

D. H. Kim, K. B. You, K. H. Lee
 Department of Electronic Engineering, Soong Jun University

ABSTRACT

Several pitch detection algorithms are studied and compared with the standard pitch detector in a term of some kinds of errors and each of speakers. Various types of errors are defined, and rank the performance of pitch detectors.

I 서 론

피치검출기(PD)는 여러 음성처리 시스템에서 중요한 요소로서 대개의 VOCODER 시스템이 필요하다.

시간적지각 같은 피치검출기 알고리즘이 제안되어 왔으나, 상호간의 성능의 우열을 가려다 보면 동의할일이 아니다.

본 논문은 총 7개의 피치검출 알고리즘의 성능을 비교(1) 검토하여, 각종의 오차를 평가 기준으로 정의하고, 이를 확자어 이하의 성능을 비교하였다.

피치검출기는 5가지 난점이 있다.

- 1) glottal 특성의 주기적인 흡수율이 양산점.
- 2) 성모와 glottal 영각과의 상호 작용.
- 3) 무성음 기간중, 성음검의 결영의 난점. (Fig. 7)
- 4) 무성음과 낮은 레벨의 무성음 사이의 구분이 난점.

5) 전화시스템에서의 피치검출의 난점

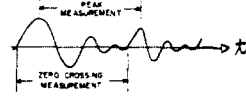


Fig. 1. Two waveform measurements which can be used to define pitch markers.

피치검출기는 다음 세종류로 분류할 수 있다.

가) 시간 영역상의 해석

a) 직접에서 피치 주기 측정

DaRD, PPROC

b) 자기 상관함수를 이용

AUTOC, AMDF

나) 주파수 영역상의 해석 DEP

다) 시간 및 주파수 영역상의 해석

SIFT, LPC

II AUTOC 방법(Modified autocorrelation method)

AUTOC 방법은 Sondhi의 center clipping

방법이 기본을 둔 수정 자기 상관 피치검출기이다.

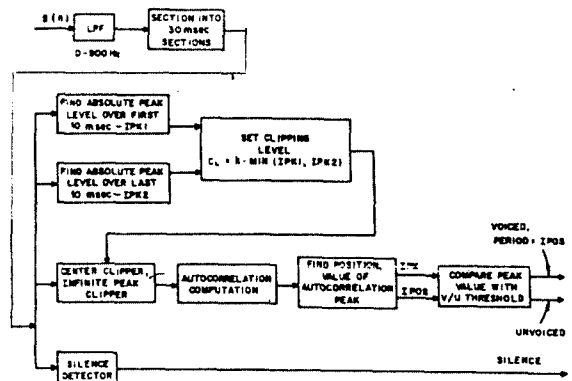


Fig. 2. Block diagram of the AUTOC pitch detector.

FIR LPF는 aliasing error를 방지한다. 30ms / 3 sample 구간에서 약 1ms는 서로 overlapping 시켜서 처리한다. Clipping level C_L 은 신호의 진폭과 관계없이 10ms 내에 있는 peak의 절대치의 0.4값으로 정한다.

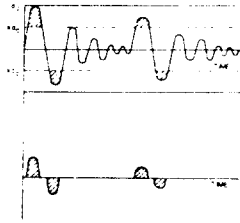


Fig 3

자기 상관함수를 계산하여, 그 최대값의 0.3을 threshold 값으로 결정한다. 무음 검출부는 최대 피크의 1/15 이하로 검출된다.

III CEP 방법 (Cepstral Method)

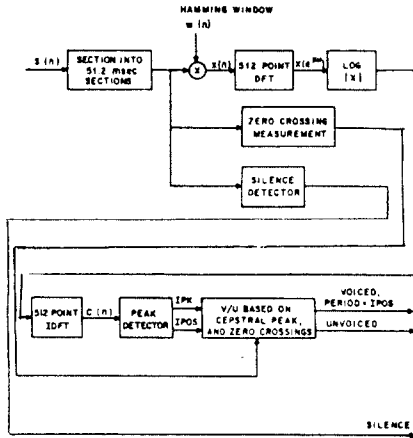


Fig 4 Block diagram of the CEP pitch detector.

512의 샘플(51.2ms)로 나누어진 구간은 해밍 윈도우로 weighting 되어 있다.

cepstral 피크 값에서 피크 주기가 결정되어지고, threshold 값을 초과하면 유성 음이 된다. threshold 의 값을 초과하지 못하면 zero crossing count로 동작이 옮겨진다. 관음 zero crossing count가 4이던 threshold 을 초과하면 그 구간은 무성음이 된다.

그렇지 않으면 무성음이 되고, 여기는 cepstrum의 최대값이 위치한다.

IV SIFT 방법 (Simplified Inverse Filter Tracking)

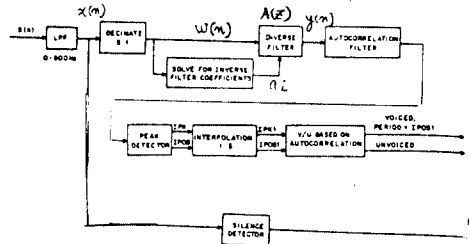


Fig 5. Block diagram of the SIFT pitch detector.

SIFT는 피치 검출의 기본은 Inverse Filter의 사용이 있다. 10KHz의 샘플링 주파수에서 5:1로 decimation 시켜서 2KHz 샘플링으로 바꾼다. 새로운 기열 $w(n)$ 은 관 입력 기열 $x(n)$ 의 피크치 샘플로 이루어진다.

inverse filter는 최소 지층 이층 이층을 반복 정확한다.

$$A(z) = 1 + \sum_{i=1}^M a_i \cdot z^{-i} \quad (1)$$

inverse filter $A(z)$ 의 이니셜 출력은 최소로 시키는 조건에서, LPC 이층 계수 a_i 는 Levinson의 자기상관 함수법으로 구해진다.

inverse filter의 spectrum

$$|A(e^{j\omega T})|^2 = \left| 1 + \sum_{i=1}^M a_i e^{-j\omega T} \right|^2 \quad (2)$$

의 값으로 정의되고, $|D(z)|^2 = \left| \frac{1}{A(z)} \right|^2$ 은

resonance의 동작을 나타낸다.

$y(n)$ 에서 피치 주기가 확실하게 잡은 것이 있으므로 다시 자기 상관함수로 계산하여 피치 주기를 검출하고, 1/5 보간(interpolation)하여, 정확한 피치 주기를 구하여 오·무성음을 판단한다.

V 개선된 신호 처리에 의한 성능 비교

V=0.7기의 알고리즘

- AUTOC, CEP 및 SIFT의 의미와 비교법과 비교한다.
- AMDF(Average Magnitude Different Function)
- DARD(Data Reduction Method)
- PPROC(Parallel Processing Method)

LPC(Linear Predictive Coding)

V-1 화자의 과치 및 과음 조건

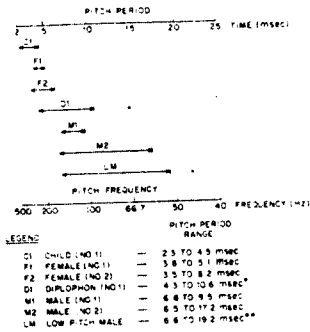


Fig. 6 Pitch variation for each of the speakers used in this study.

실험에 참가한 화자들의 과치 주기는 데이터베이스로부터 Fig. 6 과 같이 나타났다.

LM, M2, M1 은 남성

F1, F2 는 여성

C1 은 아동, D1 은 이중 음성 화자이다.

측정이 사용된 데이터베이스는 모의기한 11개의 단어와 음성음만으로 이루어진 단어장 2개, 그리고 음성음을 포함한 단어장 두 개를 사용한다.

Fig. 7 는 화자들이 이러한 과치 주기의 발생 빈도도이다.

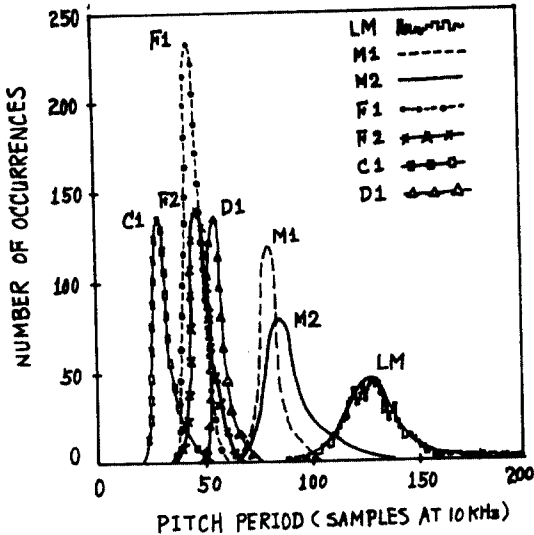


Fig. 7 PITCH PERIOD DISTRIBUTION

V-2 오차의 종류

오차의 종류는 과치 검출기와 비교하여 오차를 증가한다. 어떤 음성음의 과치 주기가 표준 과치검출로 P1 이고, 비교 과치검출에서 P2 일때

$$\text{오차 } e(m) = P1 - P2 \quad (3)$$

$$|e(m)| \geq 1ms \text{ (10 samples) 이면}$$

대오차 (Gross Pitch Period Error, G.P.E.)

라 한다.

소오차 (Fine Pitch Period Error, F.P.E.) 는

$$|ms| \geq |e(m)| \text{ 인 때이다.}$$

과치검출 성능 비교에서 다섯가지 오차를 사용한다.

1) 대오차수 (Gross Error Count)

복음된 과치 값에서 대오차의 빈도수

2) 소오차 (F.P.E.) 의 평균치

$$\text{평균: } \bar{e} = \frac{1}{N_i} \sum_{m=1}^{N_i} e(m_j) \quad (4)$$

3) 소오차의 표준 편차

$$\sigma_e = \sqrt{\frac{1}{N_i} \sum_{m=1}^{N_i} e^2(m_j) - \bar{e}^2} \quad (5)$$

4) U/V Error Rate

음성음으로 나올 것이 음성음으로 나올 오차를

5) V/U Error Rate

V-3 오차검출기의 성능 비교

전반적으로 AUTOC 와 LPC 방법의 오차가 제일 적다. (Table 1)

CEP 는 U/V 오차가 큰 것 외에는, 그 외의 오차는 적을 것이다.

AMDF 는 U/V 오차에서 제일 적은 오차를 나타내므로 CEP 와 보완하면 좋은 PD 알고리즘을 얻을 것이다.

SIFT 는 AUTOC 와 CEP 의 보완 방법으로 기본적으로 LPC 와 같은 상변이므로 거의 비슷하게 적은 오차를 나타낼 것이지만,

여기서 사용된 inverse filter 의 예측차수는 4차이지만, LPC 는 41 pole 을 사용하고 있으므로 미우적인 오차를 나타내고 있다고 하겠다.

Table. 1 Kinds of errors vs. comparative performance of PDS

| 유차의 종류 | | 피치검출기 | | | | | | | SUM |
|-----------|----------|-------|-----|------|------|-------|-----|------|-----|
| | | AUTOC | CEP | SIFT | DARD | PPROC | LPC | AMDF | |
| SPE | UN. | 2 | 1 | 4 | 6 | 5 | 1 | 3 | 22 |
| | SM. | 1 | 1 | 4 | 2 | 3 | 2 | 3 | 16 |
| F. 표준편차 | UN. | 1 | 1 | 2 | 3 | 3 | 1 | 4 | 15 |
| | SM. | 2 | 4 | 3 | 5 | 6 | 1 | 7 | 28 |
| U/V ERROR | UN. | 2 | 6 | 5 | 4 | 1 | 3 | 1 | 22 |
| | W.B. UN. | 2 | 4 | 3 | 3 | 1 | 1 | 1 | 15 |
| | SM. | 3 | 6 | 4 | 5 | 1 | 2 | 1 | 22 |
| | W.B. SM. | 3 | 6 | 4 | 5 | 2 | 1 | 2 | 23 |
| | SUM | 26 | 33 | 45 | 39 | 33 | 28 | 34 | |
| Y/U ERROR | UN. | 2 | 1 | 5 | 3 | 4 | 6 | 4 | 25 |
| | W.B. UN. | 3 | 1 | 4 | 1 | 3 | 2 | 1 | 15 |
| | SM. | 2 | 1 | 3 | 1 | 2 | 5 | 4 | 18 |
| | W.B. SM. | 3 | 1 | 4 | 1 | 2 | 3 | 3 | 17 |
| | SUM | 10 | 3 | 6 | 5 | 3 | 2 | 4 | |

표의 숫자는 각종의 피치검출기가 어떠한 오차가 적은 순위를 나타낸다. (1)

UN 은 unsmooth 이고, SM 은 smoother 를 붙인 경우이고 WB 는 광미인 경우이다.

V-4 확장된 검출기 성능 비교

Table. 2 Each speakers vs. comparative performance of PDS

| 피치검출기 화자 | 피치검출기 | | | | | | | SUM |
|-------------|-------|-----|------|------|-------|-----|------|-----|
| | AUTOC | CEP | SIFT | DARD | PPROC | LPC | AMDF | |
| LM | 3 | 7 | 1 | 3 | 2 | 5 | 5 | 26 |
| M1 | 1 | 6 | 3 | 6 | 3 | 1 | 5 | 25 |
| M2 | 2 | 3 | 1 | 6 | 3 | 7 | 5 | 27 |
| F1 | 2 | 7 | 6 | 4 | 2 | 5 | 1 | 27 |
| F2 | 1 | 5 | 5 | 7 | 3 | 2 | 3 | 26 |
| C1 | 1 | 6 | 7 | 4 | 3 | 2 | 4 | 27 |
| SUM | 10 | 34 | 23 | 30 | 16 | 22 | 22 | |
| 평균 | 1 | 6 | 4 | 5 | 2 | 3 | 4 | |

안정적인 기본 피치 추적에 대하여 AUTOC 방법의 가장 우수한 결과들을 보인다.

피치검출기 LM 과 M2 확장에 대하여 SIFT 가 우수한 성능을 나타내고, 일면의 F1, F2, C1 확장은 추적에 미흡한 입증한다. 이것은 SIFT에서 Decimation 한 대역폭의 서미은 거스를 그하고, 자기 상관 함수 값을 그하고 있다.

VI 기본 피치 검출기

7개의 피치검출기 알고리즘의 성능을 피치와 확장된 범위로 비교 분석하였다. 그 결과 프라수 영인의

CEPSTRAL 방법과 소인 영인의 AMDF 방법을 보면 다음 기준에 비교해 주어야 된다.

확장의 범위에 따라 성능을 나타낸 것처럼 단 코도에서, 각각의 피치 검출기 알고리즘의 우열을 나타내고 있음을 알 수 있다.

AUTOC 방법의 일차 추적에 대하여 오차가 적은 일하고, 성능이 우수함을 나타내고 있다.

참고 문헌

- (1) L.R.RABINER, M.J.CHEUNG, A.E.ROSENBERG and C.L.McGONEGAL, "A Comparative Performance Study of Several Pitch Detection Algorithms", IEEE Trans. vol. ASSP-24 OCT. 1976
- (2) L.R.RABINER, R.J.Schafer, "Digital Processing of Speech Signals", PRENTICE HALL 1978.