

## 고차 통계를 이용한 잡음 환경에서의 음성신호의 피치 추출과 유,무성음 판별

신 태영, 김 재호, 손 경식, 김 형순  
부산 대학교 전자공학과

### Pitch Determination and Voiced/Unvoiced Decision of Noisy Speech Based on the Higher-Order Statistics

Tae Young Shin, Jae Ho Kim, Kyung Sik Son, Hyung Soon Kim  
Dept. of Electronic Engineering, Pusan National University

#### 요약

본 논문에서는 고차 통계(higher-order statistics)방법을 이용하여 잡음이 섞인 음성 신호에서 피치를 구하는 방법과 이를 이용하여 유성음 및 무성음 구간을 구분하는 방법을 구현하고 그 결과를 기술하였다. 고차 통계의 일종인 3차 cumulant 함수의 경우 Gaussian 또는 대칭적인 분포를 갖는 잡음 신호를 음성신호로부터 효과적으로 분리하여 제거시키는 특징을 가지고 있으며, 이러한 특징을 이용하면 잡음 환경에서 여러 가지 음성 특징 파라미터들을 보다 신뢰도 높게 추정할 수 있다.

본 논문에서는 음성 신호의 3차 cumulant 함수의 자기상관 함수로부터 음성의 피치 주기를 추정하였으며, 피치 위치에서의 normalized peak 크기에 의해 유성음과 무성음을 구분하였다. 또한 성능 비교를 위해 음성 신호 자체의 자기 상관 함수로부터 역시 피치 주기 및 유성음/무성음 구분을 수행하였다. 백색 및 유색 Gaussian 잡음 환경에서의 음성의 피치 주기 추정 실험 결과 SNR(Signal-to-Noise Ratio)가 낮은 경우에 3차 cumulant를 이용한 방법이 2차 통계에 비해 우수한 성능을 나타내었다. 또한 동일한 잡음 환경에서의 유성음/무성음 판별 시험에서도 3차 cumulant를 이용한 방법이 기존의 2차 통계를 이용한 방법에 비해 성능이 크게 향상된 결과를 얻었다.

#### 1. 서론

음성신호를 유성음과 무성음으로 구분하고 유성음의 피치(pitch)주기를 구하는 것은 음성신호 처리의 중요한 과제 중의 하나이다. 이러한 정보들은 저전송속도의 음성 부호화 기술을 비롯하여 음성 인식 및 합성 기술, 그리고 화자 인식기술 등의 응용 분야에 유용하게 이용될 수 있다. 이에 따라 음성신호를 유성음/무성음으로 구분하고 유성음의 피치주기를 추출하기 위한 다양한 방법들이 개발되어 왔다[1-4]. 이들 방법들 중에서

음성 신호의 2차 통계, 즉, 자기상관(autocorrelation)함수 및 그 Fourier 변환인 전력 스펙트럼(power spectrum)에 기반을 둔 방식들이 널리 사용되고 있다. 그러나, 이와 같은 2차 통계에 방식들은 잡음이 없는 환경에서는 매우 우수한 성능을 나타내지만, 잡음 환경, 특히 자동차 소음 등과 같이 어느 정도 주기적인 특성을 가지는 유색 잡음(colored noise)환경에서는 성능이 크게 저하된다.

최근 신호 처리 분야에서 많은 관심이 기울여지고 있는 고차 통계, 즉 cumulant 함수와 이와 Fourier 변환인 poly-spectrum을 이용한 방법은 Gaussian 신호를 억제하는 특성을 가지고 있어서, 음성 신호의 유성음과 같은 non-Gaussian 신호를 표현하는데 유용한 것으로 알려지고 있다[5-7]. 특히 3차 cumulant 함수의 경우, Gaussian 특성을 갖는 잡음 이외에도 대칭적인 확률 분포를 갖는 잡음 신호를 억제하는 특성을 가지고 있으므로, 잡음 환경의 음성 신호 처리에 효과적으로 사용될 수 있다. 이에 따라 음성 분석 및 음성 인식 등의 응용 분야에 고차 통계를 적용하려는 시도가 조금씩 이루어 지고 있다[8-12].

본 논문에서는 지금까지 제안된 방법들[8-12]을 토대로 백색 잡음 및 유색 잡음 환경에서 고차 통계를 이용하여 유성음의 피치 주기를 추정하는 방법과 이러한 피치주기 추정 과정을 통해 유성음과 무성음을 구분하는 방법을 구현해 보고, 실험을 통해 2차 통계를 이용한 방법과의 성능을 비교하였다. 실험 결과 여러 경우의 잡음 환경에 대해 고차 통계를 이용한 방법이 2차 통계를 이용한 방법보다 향상된 결과를 나타내었다.

본 논문의 구성은 다음과 같다. 2절에서는 고차 통계에서의 cumulant 함수의 정의와 성질에 대해 설명하고, 3절에서는 2차 통계 및 고차 통계를 이용하여 피치를 추정하는 방법과 유성음과 무성음을 구분하는 방법을 기술한다. 4절에서는 수행한 실험 내용 및 그 결과를 나타내었으며, 마지막으로 5절에서 결론을 맺는다.

2. 고차 통계 방법 개요

고차 통계 방법이란 3차 이상의 통계적 특성을 이용하여 신호를 처리하고자 하는 방법이다. 특히, 고차 통계에서 정의하고 있는 cumulant 함수는 non-Gaussian 특성을 가지는 신호를 분석하는데 아주 유용하게 사용될 수 있다. 일반적으로 고차 통계를 이용하는 목적은 여러 가지가 있으나[5] 고차 통계를 통하여 특성을 알수없는 Gaussian noise process를 제거하거나 대칭적인 분포를 가진 신호들을 제거할 수 있다[6]. k 개의 랜덤 변수로 이루어진 랜덤 벡터  $X = (x_1, x_2, \dots, x_n)$  에 대해 cumulant는 다음과 같은 특성 함수 (characteristic function)  $\Phi(W)$  의 정의로부터 유도된다.

$$\Phi(W) = E[\exp(jWX^T)] \quad (1)$$

여기서  $W = (\omega_1, \omega_2, \dots, \omega_n)$  는 특성 함수의 계수 벡터 이며,  $X^T$  는  $X$  의 transpose 를 의미한다. 이와 같은 특성 함수로부터 k차 cumulant는 다음과 같이 정의된다[6].

$$\text{Cum}[x_1^{k_1}, x_2^{k_2}, \dots, x_n^{k_n}] = (-j)^{k_1+k_2+\dots+k_n} \frac{\partial^{k_1+k_2+\dots+k_n} \ln|\Phi(W)|}{\partial \omega_1^{k_1} \partial \omega_2^{k_2} \dots \partial \omega_n^{k_n}} \Big|_{\omega_1=\omega_2=\dots=\omega_n=0} \quad (2)$$

여기서,  $k = k_1 + k_2 + \dots + k_n$  이다

$X(t)$  가 zero-mean을 가지는 stationary한 신호일 경우  $X(t)$  의 2차 및 3차 cumulant 함수는 각각

$$C_{2,X}(t_1, t_2) = \text{Cum}[X(t), X(t + \tau)] = E[X(t)X(t + \tau)] \quad (3)$$

및

$$C_{3,X}(t_1, t_2) = \text{Cum}[X(t), X(t + \tau_1), X(t + \tau_2)] = E[X(t)X(t + \tau_1)X(t + \tau_2)] \quad (4)$$

로 표현할 수 있다. 즉, 이 경우 2차 cumulant 함수는 자기상관 함수와 동일하다. Cumulant 의 성질들 중에서 잡음 환경에서의 음성신호 처리에서 유용하게 이용될 수 있는 성질은 다음과 같다[7]. 첫째로, 두 랜덤 변수 집합  $\{X(k);$  와  $\{Y(k);$  가 서로 독립적일 경우, 다음과 같은 관계가 성립 한다.

$$\text{Cum}[X(k) + Y(k)] = \text{Cum}[X(k)] + \text{Cum}[Y(k)] \quad (5)$$

둘째로, 랜덤변수집합  $\{Y(k);$  가 jointly Gaussian 분포를 가질 경

우 3차 이상의 cumulant는 0이고, 대칭적인 분포를 가지고 있을 경우 홀수차의 cumulant는 0이다.

결과적으로 non-Gaussian 분포를 가지는 음성신호에 Gaussian 분포를 가지거나 대칭적인 분포를 가지는 잡음이 섞인 경우 cumulant 는 음성 신호와 잡음 신호를 분리해 내고 잡음 신호를 제거할 수 있게 된다

3 피치 추정 및 유성음/무성음 구분 알고리즘

3.1. 피치 추정 알고리즘

먼저 음성신호의 자기상관 함수를 이용하여 피치주기를 구하는 방법은 다음과 같이 요약될 수 있다[13]. 한 프레임의 음성신호  $X(1), X(2), \dots, X(N)$  으로부터 자기상관 함수를 다음과 같이 구한다

$$R_x(k) = \sum_{t=0}^{N-k} X(t)X(t+k) \quad k = 0, \dots, N-1 \quad (7)$$

이러한 자기상관 함수는  $k = 0$  에서 최대점을 가지며, 음성신호가 주기적인 경우 준 주기성을 띄기 때문에 음성의 피치주기가 있을 것으로 예상되는 영역에서의 최대치를 곧 피치주기로 판정할 수 있다. 본 논문에서 cumulant를 이용하여 음성의 피치주기를 추정하는 방법은 Moreno 등이 제안한 방법을 기반으로 하고 있다[7]. 우선 음성신호  $X(1), X(2), \dots, X(N)$  으로부터 다음과 같이 cumulant 함수를 추정한다

$$C(0, k) = \sum_{t=0}^{\min(N-k, N-k)} X(t)X^2(t+k) \quad k = -(N-1), \dots, (N-1) \quad (8)$$

이러한 cumulant 함수도 음성신호가 주기적인 경우 준 주기적인 성질을 가지지만,  $k = 0$  에서 최대가 되지는 않는다. 따라서, cumulant 함수의 최대값으로부터 직접 피치주기를 추정할 수는 없으며, cumulant 함수의 주기성을 이용하여 cumulant들의 자기상관 함수를 다음과 같이 구한다[7]

$$R_{xx}(n) = \sum_{k=-(N-1)}^{N-1} C(0, k)C(0, k+n) \quad n = 0, \dots, N-1 \quad (9)$$

$R_{xx}(n)$  의 경우 피치주기에서 peak를 가지므로, 음성 신호 자체의 자기상관 함수와 마찬가지로 음성의 피치주기가 있을 것으로 예

상되는 영역에서 최대값을 찾음으로써 피치주기를 추정할 수 있다

3.2. 유성음/무성음 구분 알고리즘

일반적으로 유성음과 무성음을 구분하는 방법은 주어진 음성 신호에서 유성음과 무성음들이 가지는 특징 파라미터들을 이용한다[11][13]. 본 논문에서 사용한 2차 통계를 이용한 유/무성음 판별은 주어진 음성 신호의 자기상관 함수에서 피치를 구한 후 음성 신호의 에너지와 피치에서의 자기상관 함수 값의 비, 즉 normalized peak를 적절한 문턱값과 비교함으로써 이루어진다. 이때 normalized peak값  $NP_A$ 는 음성 신호의 자기상관 함수  $R_A(k)$ 로 부터 다음과 같이 구한다[13]

$$NP_A = \frac{R_A(P_A)}{R_A(0)} \tag{10}$$

여기서,  $P_A$ 는 음성 신호의 자기상관 함수로 부터 구한 피치값을 나타낸다. 3차 통계를 이용한 방법은 음성 신호의 cumulant 함수로 부터 구한 자기상관 함수  $R_C(p)$ 에서 피치를 결정한다. 이로부터 구한 normalized peak를 역시 적절한 문턱값과 비교함으로써 수행한다[14]. 3차 통계 방법에서의 normalized peak는 cumulant의 에너지와 피치에서의 자기상관 함수의 값의 비로서 다음과 같이 구한다.

$$NP_C = \frac{R_C(P_C)}{R_C(0)} \tag{11}$$

여기서,  $P_C$ 는 cumulant함수의 자기상관 함수로 부터 구한 피치값이다. 이러한 normalized peak를 이용하여 유성음과 무성음을 구분하기 위해서는 적절한 문턱값을 설정하는 것이 필요한데 본 논문에서는 실험을 통해 각각의 방법에서 가장 우수한 성능을 나타내는 문턱값을 선정하였다. 실제로는 SNR의 변화에 따라 normalized peak의 분포가 변하기 때문에 문턱값도 SNR에 따라 적절히 변화되어야 한다. 이를 위해서는 주어진 음성신호의 SNR를 추정하는 작업이 필요하므로 본 논문에서는 SNR에 따른 문턱값의 조정은 고려하지 않았다.

4. 실험 및 결과

지금까지 설명한 피치주기 추정 방법들의 성능 비교를 위해 실제 음성 데이터로부터 피치주기를 추정하는 실험을 수행하였다. 실험을 위한 음성 데이터는 3명의 화자가 발음한 한국어 단

음 /아/, /어/, /오/, /우/, /으/, /이/, /에/의 7가지로서 전체 21개의 음절을 사용했으며 음성 데이터의 sampling 주파수는 20 kHz이다. 음성 신호는 매 10 ms마다 40 ms를 한 프레임으로 하여 rectangular window를 씌운 다음 음성신호의 자기상관 함수 및 cumulant의 자기상관 함수를 구하였다. 그림 1은 SNR이 0dB가 되도록 백색 Gaussian 잡음을 섞은 음성 신호 /아/로 부터 피치주기를 추정하는 예를 보여주고 있다. 이때 잡음이 섞이지 않은 음성 신호의 파형을 통해 관측된 피치주기는 180sample(9 ms)이며, 그림에서 수직 실선으로 표시되어 있다. 그림 1의 (a)는 음성신호 자체의 자기상관 함수를 구한 경우로서, SNR이 0dB임에도 불구하고 그림에서 보는 바와 같이 피치주기에 해당하는 위치에 peak가 비교적 현저하게 나타남을 알 수 있다. 이는 백색 Gaussian 잡음의 경우 자기상관 함수는 델타 함수의 형태를 가지며 피치 위치에는 그다지 큰 영향을 주지 못하기 때문인 것으로 풀이된다. 그림 1의 (b)는 음성신호로부터 구한 cumulant 함수를 대상으로 다시 자기상관 함수를 구한 경우로서, 피치주기 위치에서의 peak가 그림 (a)에 비해 보다 현저하게 나타남을 알 수 있다. 이는 cumulant 함수가 백색 Gaussian 잡음을 억제하여 신호의 주기성을 향상시킨 때문으로 해석된다.

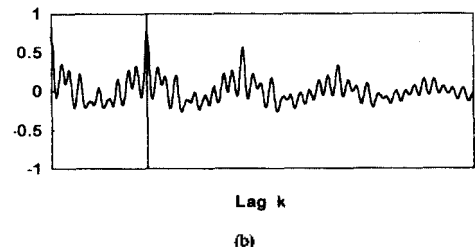
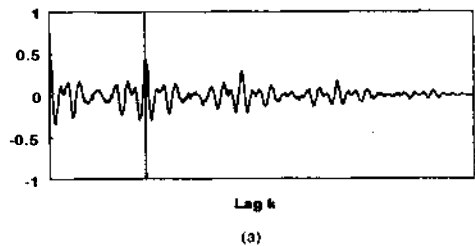
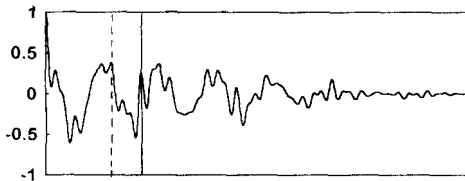


그림 1. SNR 0 dB인 백색 잡음 환경에서의 음성신호 /아/의 피치 추정 예  
(a) 음성 신호의 자기상관 함수를 이용한 경우  
(b) cumulant의 자기상관 함수를 이용한 경우

그림 2는 동일한 음성 신호 /아/에 SNR이 0 dB가 되도록 유

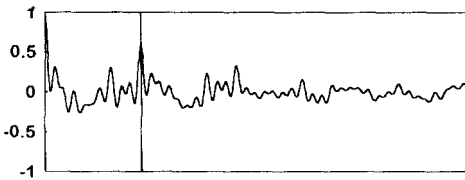
고차통계를 이용한 잡음환경에서의 음성신호의 Pitch 추종과 유,무성음 판별

색 Gaussian 잡음을 섞은 상태에서 피치 주기를 추정하는 예를 보여 준다. 실험에 사용한 유색 Gaussian 잡음은 백색 Gaussian 잡음을 passband가 100 Hz에서 150 Hz인 band pass filter를 통과시켜 구하였으며 피치주기 추정에 영향을 줄 수 있는 주기성을 가지고 있다. 그림 2의 (a)는 유색 잡음이 섞인 음성신호 자체의 자기상관 함수를 구한 것이다. 그림에서 실선으로 표시된 수직선이 실제 음성의 피치 위치를 나타낸 것이며 점선으로 표시된 것은 2차 통계 방법에 의해 잘못 찾아낸 피치 위치이다. 이는 잡음 섞인 음성 신호 자체의 자기상관 함수를 이용한 방법으로는 음성 신호의 주기성과 유색 잡음의 주기성을 구분할 수 없기 때문에 비롯된 결과이다. 한편, 그림 2의 (b)는 잡음 섞인 음성 신호의 cumulant 함수를 구한 다음 cumulant의 자기상관 함수를 이용하여 피치를 구하는 경우이다. 그림에서 볼 수 있듯이 이 방법을 이용하면 유색 잡음 환경에서도 피치 위치를 잘 찾아낼 수 있음을 알 수 있다. 그리고 이 결과는 cumulant 함수가 유색 Gaussian 잡음을 효과적으로 제거할 수 있다는 사실을 나타낸다.



Lag k

(a)



Lag k

(b)

그림 2. SNR 0 dB인 유색 잡음 환경에서의 음성신호  $x(k)$ 의

피치 추정 예

(a) 음성 신호의 자기상관 함수를 이용한 경우

(b) cumulant의 자기상관 함수를 이용한 경우

호에서 구한 피치와의 차이가 1ms를 초과했을 경우를 error라고 판정하였다(6). 표 1의 (a)와 (b)는 각각 백색 및 유색 잡음의 경우 SNR의 변화에 따른 피치 추정 error rate를 나타낸 것이다. 표 1에서 보는 바와 같이 SNR이 높은 경우에는 cumulant에 의한 방법의 성능이 2차 통계에 의한 방법보다 약간 뒤떨어지거나 비슷한 수준을 나타내며, SNR이 0dB로 떨어지면 cumulant 방법이 상대적으로 우수한 성능을 보이고 있다. 그러나, 유색 잡음의 경우 예상과는 달리 고차 통계의 장점이 크게 부각되지 못하였는데, 피치 추정 오류의 원인을 살펴 본 결과 간헐적으로 피치 주기의 2배(doubling)의 위치에서 피치를 잘못 찾아내는 경우가 많았다. 실제로 cumulant 함수를 이용한 피치 추정 error 가운데 50%가 doubling error였으며, 음성 신호의 자기상관 함수를 이용한 방법에서는 28%가 doubling error 때문에 발생한 것이었다. 이와 같은 error는 smoothing을 통해서 제거될 수 있는데 표 2는 유색 잡음 환경에서의 피치 주기 추정 결과를 5-point median filter에 의해 smoothing한 다음 피치 추정 error rate를 나타낸 것이다. 표 2로 부터 SNR이 0dB인 경우 3차 통계를 사용하면 2차 통계를 사용한 경우에 비해 피치 추정 error가 절반 이하로 줄어들게 됨을 알 수 있다.

표 1 SNR의 변화에 따른 피치 추정 error rate (%)

(a) 백색 잡음의 경우

	20 dB	10 dB	5 dB	0 dB
Autocorrelation에 의한 방법	0.2	0.7	0.5	2.8
Cumulant에 의한 방법	0.7	0.7	0.8	1.3

(b) 유색 잡음의 경우

	20 dB	10 dB	5 dB	0 dB
Autocorrelation에 의한 방법	0.2	0.7	2.5	26.5
Cumulant에 의한 방법	0.5	0.7	2.2	18.8

표 1은 다양한 잡음 환경에서 각각의 방법으로 구한 피치와 원래의 음성신호에서 구한 피치와의 차이를 비교한 것이다. 본 논문에서는 각각의 방법으로 구한 피치 값과 원래의 음성신

표 2. 유색 잡음 환경에서 피치 추정 결과를 smoothing 한 후의 SNR 에 따른 피치 추정 error rate(%)

	20 dB	10 dB	5 dB	0 dB
Autocorrelation 에 의한 방법	0.2	0.2	1.2	19.2
Cumulant 에 의한 방법	0.2	0.2	1.2	8.9

유성음과 무성음을 구분하는 방법을 실험하기 위해서는 피치 추정 시 사용했던 화자 중 2명이 발음한 한국어 단음절의 음성 데이터가 사용되었다. 이 때 사용된 단음절은 13가지 초성 자음 [기/나/디/리/미/비/시/지/치/키/티/피/히] 과 3 가지 모음 [아/이/우]로 구성된 전체 78개의 단음절들이다. 잡음 신호 역시 피치 추정 시 사용했던 것과 동일한 백색 및 유색 Gaussian 잡음들을 사용했다. 음성 신호의 유성음/무성음 구분은 수작업에 의해 구분한 결과와 비교하도록 하였으며, 이 때 유성음 또는 무성음의 어느 한 쪽으로 분류할 수 없는 프레임들은 제외시켰다. 그림 3은 SNR이 10 dB인 백색 잡음의 경우 유성음과 무성음에 대한 normalized peak의 분포를 나타낸 것이다. (a)는 음성 신호의 자기상관 함수로 부터 구한 것이고 (b)는 cumulant 함수의 자기상관 함수로 부터 구한 것이다. 그리고 그림 4의 (a),(b)는 그림 3과 동일한 조건에서 SNR 10 dB인 유색 잡음의 경우 유성음과 무성음에 대한 normalized peak의 분포를 그린 것이다. 그림에서 알 수 있듯이 cumulant를 이용한 방법이 신호 자체의 자기상관 함수를 이용한 방법보다 백색 잡음과 유색 잡음 모두의 경우에 변별력이 우수하다는 것을 알 수 있다. 표 3의 (a)는 백색 잡음의 경우에 SNR의 변화에 따른 유성음/무성음 판별 error rate를 각각의 경우에 대해 구한 결과이다. 그리고 (b)는 유색 잡음의 경우에 대해 구한 결과이다. 음성 신호의 자기상관 함수를 이용한 방법에서는 문턱값을 0.35로 하였고 cumulant의 자기상관 함수를 이용한 방법에서는 문턱값을 0.6으로 정하였으며, 이들 값은 각각의 방법에 대해 가장 우수한 성능을 나타내는 값들이다. 표 3에서 알 수 있듯이 cumulant의 자기상관 함수를 이용한 방법이 백색 잡음과 유색 잡음의 경우 우수한 성능을 나타내었다. 또한 그림 3과 4에서 보는 바와 같이 cumulant의 자기상관 함수를 이용한 방법이 음성 신호의 자기상관 함수를 이용한 방법보다 SNR의 변화에 민감하지 않다는 것을 알 수 있으며, 문턱값의 변화에 따른 error rate의 변화 역시 음성 신호 자체의 자기상관 함수를 이용한 경우에 비해 둔감한 특성을 나타내었다.

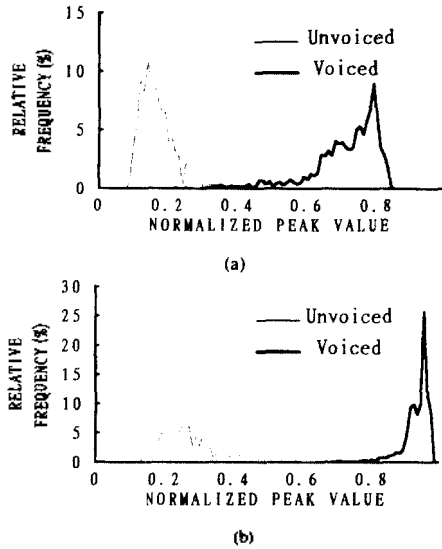


그림 3. SNR 10 dB인 백색 잡음 환경에서의 유성음과 무성음의 normalized peak 값의 분포

- (a) 음성신호의 자기상관 함수를 이용한 경우
- (b) Cumulant의 자기상관 함수를 이용한 경우

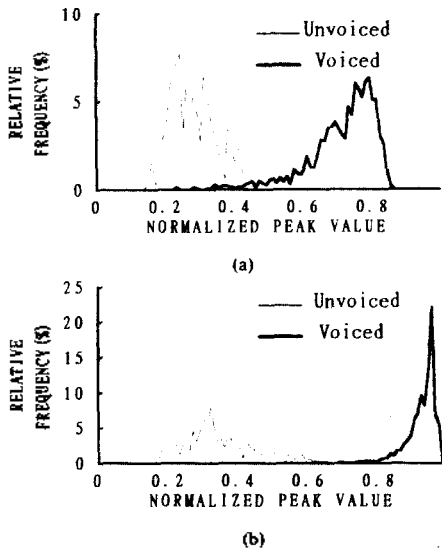


그림 4. SNR 10 dB인 유색 잡음 환경에서의 유성음과 무성음의 normalized peak 값의 분포

- (a) 음성 신호의 자기상관 함수를 이용한 경우
- (b) Cumulant의 자기상관 함수를 이용한 경우

표 3. SNR의 변화에 따른 유성음/무성음 판별 error rate (%)

(a) 백색 잡음의 경우				
	20 dB	10 dB	5 dB	0 dB
Autocorrelation 에 의한 방법	0.9	4.1	1.0	2.2
Cumulant 에 의한 방법	0.2	0.2	0.2	1.2
(b) 유색 잡음의 경우				
	20 dB	10 dB	5 dB	0 dB
Autocorrelation 에 의한 방법	0.9	0.9	3.3	6.3
Cumulant 에 의한 방법	0.2	0.2	0.5	1.6

5 결론

본 논문에서는 잡음 환경에서 음성 신호의 피치 추출과 유성음과 무성음을 구분하는 데 있어 3차 통계를 적용하는 방법을 구현하였다. 실험 결과에서 볼 수 있는 바와 같이 피치 추정과 유/무성음 구분에 있어서 3차 통계를 이용한 방법은 기존의 2차 통계를 이용한 방법보다 잡음 환경에서 향상된 성능을 나타내었다. 앞으로 본 논문에서 사용된 피치 추출 및 유/무성음 판별 알고리즘을 이용하여 고차 통계 방법을 이용한 잡음 환경에서의 화자 인식 시스템의 구현이 계속 진행될 예정이다. 본 논문에서는 피치 주기의 크기 정보만을 이용하였으나 그 이외의 다양한 음성 파라미터들을 추가한다면 성능을 더욱 향상시킬 수 있으리라 판단되며 실제 잡음 환경에서 발음한 음성에 대해 고차 통계를 이용한 방법을 적용하는 것도 검토되어야 할 것이다.

참고 문헌

[1] Wolfgang J. Hess, "Pitch and Voicing Determination", in *Speech Signal Processing*, S. Furui and M. M. Sondri, Eds., Marcel Dekker, pp. 3-48, 1992.

[2] 배 명진, 이 상효, 전 홍렬, "음성 신호의 기본 주파수 검출", 제 11회 음성 통신 및 산 처리 워크샵, pp. 42-47, 1994.

[3] L. R. Rabiner, M. J. Cheng, A. E. Rosenberg and C. McGonegal, "A Comparative Performance Study of Several Pitch Detection Algorithms", *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 24, pp. 399-417, Oct. 1976.

[4] Bishnu S. Atal and Lawrence R. Rabiner, "A pattern Recognition Approach to Voiced-Unvoiced-Silence Classification with Application to Speech Recognition," in *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-24, pp. 201-202, June 1976.

[5] J. R. Mendel, "Tutorial on higher-order statistics(spectra) in signal processing and system theory: theoretical results and some applications," in *Proc. IEEE*, vol. 79, no. 3, pp. 869-891, July 1987.

[6] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectral Analysis*, Englewood Cliffs, NJ, Prentice-Hall, 1993.

[7] C. L. Nikias and M. R. Raghuveer, "Bispectrum estimation: a digital signal processing framework," in *Proc. IEEE*, vol. 75, no. 7, pp. 869-891, July 1987.

[8] A. Moreno and J. A. R. Fonollosa, "Pitch determination of noisy speech using higher order statistics," in *Proc. ICASSP*, pp. 1-133-136, 1992.

[9] A. Moreno, J. A. R. Fonollosa, "Cumulant based Voicing Decision in Noise Corrupted Speech" in *Proc. ICSLP'92 - Int. Conf. on Spoken Language Processing*, pp. 531-534, Oct. 1992.

[10] B. Wells, "Voiced/unvoiced decision based on the bispectrum", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 1589-1592, 1985.

[11] K. K. Palwal and M. M. Sondri, "Recognition of Noisy Speech Using Cumulant-Based Linear Prediction Analysis", in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 429-432, May 1991.

[12] 이 형근, 양 원영, 조 용수, "유색 잡음 환경에서 Cumulant를 이용한 한국어 단모음 인식", *한국음향학회지*, 제 13권, 22호, pp. 50-59, 1994.

[13] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Englewood Cliffs, NJ, Prentice-Hall, 1978.