

## 가변 대역폭 필터를 이용한 음성신호의 AM-FM 성분 분리에 관한 연구

## Decomposition of Speech Signal into AM-FM Components Using Variable Bandwidth Filter

송 민\* · 이 회 영\*\*  
 Min Song · Heyoung Lee

## ABSTRACT

Modulated components of a speech signal are frequently used for speech coding, speech recognition, and speech synthesis. Time-frequency representation (TFR) reveals some information about instantaneous frequency, instantaneous bandwidth and boundary of each component of the considering speech signal. In many cases, the extraction of AM-FM components corresponding to instantaneous frequencies is difficult since the Fourier spectra of the components with time-varying instantaneous frequency are overlapped each other in Fourier frequency domain.

In this paper, an efficient method decomposing speech signal into AM-FM components is proposed. A variable bandwidth filter is developed for the decomposition of speech signals with time-varying instantaneous frequencies. The variable bandwidth filter can extract AM-FM components of a speech signal whose TFRs are not overlapped in time-frequency domain. Also, amplitude and instantaneous frequency of the decomposed components are estimated by using Hilbert transform.

**Keywords:** Time-varying Harmonic Structure, Instantaneous Frequency, De-composition, Speech Signal, Variable Bandwidth Filter

## 1. 서 론

음성신호의 스펙트럼은 시간적으로 변하며, AM-FM 성분들의 합으로 모델링할 수 있다 [12, 19, 20]. 음성신호에서 흔히 시변(time-varying) 포먼트 주파수(formant frequency) 또는 시변 순시 주파수(instantaneous frequency)를 갖는 AM-FM 신호 성분들을 관찰할 수 있다. 특히, 음의 높낮이가 심하게 변하는 노래나 성악의 경우 음성신호의 순시 주파수는 시간에 따라 크게 변화한다. 음성신호를 각 AM-FM 성분별로 분리하거나 또는 각 성분별 스펙트럼을 추정할 경우 음성합성, 음성인식 및 음성코딩 등에 유용하게 사용할 수 있다[1, 10, 11, 13, 14]. 음성신호의 스펙트럼을 해석하기 위하여 선형 예측(linear prediction), 켈스트럼 해석(cepstral

---

\* 전남대학교 공과대학 전자공학과

\*\* 전남대학교 공과대학 정보통신공학부

analysis), 멜-켄스트럼(Mel-cepstrum) 방법 등을 사용한다[1]. 그러나 잡음을 동반한 음성 신호의 경우 간섭 효과로 인하여 상기와 같은 방법으로는 신빙성 있는 유용한 정보를 추출하기가 매우 어렵다[1].

AM-FM 신호  $s(t) = a(t)\cos(\phi(t))$ 에서 순시 주파수  $\omega(t)$ 는 다음과 같이 정의된다[2].

$$\omega(t) = D\phi(t) \quad \text{rad/s} \quad (1)$$

여기서  $D = d/dt$ 이다. 시변 순시 주파수를 갖는 신호를 시간-주파수 영역에서 관찰할 경우 순시 주파수의 궤적을 따라 신호의 에너지가 집중되는 것을 관찰할 수 있다[3]. 음성신호를 시간-주파수 영역에서 관찰해 보면 음성신호의 각 성분별로 시변 순시 주파수의 궤적을 따라 에너지가 집중되는 것을 알 수 있다. 흔히 신호의 시간-주파수 해석에는 스펙트로그램(spectrogram), Wigner-Ville 분포(distribution), Choi-Williams 분포(distribution), Cone-Shape 분포(distribution) 등이 사용된다[4].

일반적으로 신호를 시간-주파수 영역에서 해석할 경우 신호에 대한 순시 주파수, 순시 대역폭(instantaneous bandwidth) 및 각 성분별 경계에 대한 정보 등을 얻을 수 있다. Short-time Fourier 변환 등과 같은 시간-주파수 해석 방법을 사용하여 얻어진 시간-주파수 표현(time-frequency representation)에서 순시 주파수의 궤적을 따라 에너지가 집중된 부분을 배경과 분리할 경우 분리된 시간-주파수 표현에 대응하는 시간 영역의 실제 신호(real signal)가 항상 존재하는 것을 보장할 수 없다[6]. 즉, 변형된 시간-주파수 표현(masked time-frequency representation)을 역 Short-time Fourier 변환할 경우 일반적으로 시간 영역에서 허수부분(imaginary part)이 나타난다. 이와 같은 문제점을 극복하기 위하여 최소 자승 오차 방법(least-square error method)이 사용된다[6]. 즉, 변형된 시간-주파수 표현과 추정된 신호의 시간-주파수 표현 사이의 오차가 최소가 되도록 함으로써 신호를 복구할 수 있다[3]. 시간-주파수 표현을 바탕으로 한 최소 자승 오차 방법에 대한 연구가 광범위하게 수행되었다[4, 6-8]. 그러나 최소 자승 오차 방법을 사용하여 신호를 복구할 경우 매우 많은 계산이 요구되는 단점이 있다[9].

한편, Ashwin Rao와 Ramdas Kumaresan 등은[1] 음성신호의 각 성분별 순시 주파수들이 천천히 변한다는 가정 하에(slowly time-varying instantaneous frequency) 시변 Front-end 필터 뱅크(filter bank)를 사용하여 음성신호의 각 성분들을 분리하였다. 그러나 순시 주파수가 급격히 변할 경우 각 성분들이 정확히 분리되지 않는 단점이 있다[1].

본 연구에서는 음성신호에서 시변 순시 주파수를 갖는 각 성분들을 효과적으로 분리하기 위하여 선형 시변 가변 대역폭 필터(linear time-varying variable bandwidth filter)를 제안하였다. 제안된 방법을 사용할 경우 시변 순시 주파수를 갖는 신호들로 이루어진 다중 신호를 각 성분별로 효과적으로 분리해낼 수 있다. 제안된 가변 대역폭 필터는 확장 푸리에 주파수 영역에서 분수 형태로 표현된다. 본 논문은 다음과 같이 구성된다. 2장에서 가변 대역폭 필터 이론이 다루어졌으며, 3장에서 가변 대역폭 필터의 설계 방법 및 제안된 필터의 특성을 시간-주파수 영역에서 분석하였다. 4장에서는 제안된 필터가 시변 순시 주파수를 갖는 신호를 정확하게 복구할 수 있음을 검증하였다. 그리고 5장에서는 실제 음성 신호의 각 성분들을 분리

하고 각 성분별 순시 주파수와 진폭을 추출한 실험 예가 주어졌으며, 6장에서 결론을 맺는다.

## 2. 가변 대역폭 필터

다음과 같이 미분 연산자  $G$ 를 정의하자.

$$Gx(t) = g(t)Dx(t) \quad (2)$$

여기서  $g(t)$ 는 임의의 시간 함수로써  $0 < g(t) < \infty$ 이다. 미분 연산자  $G$ 를 사용하여 구성된 선형 시변 미분방정식을 고려하자.

$$a_n G^n y + a_{n-1} G^{n-1} y + \dots + a_0 y = b_m G^m x + \dots + b_0 x \quad (3)$$

여기서  $a_i, b_i$ 는 상수이고,  $x$ 는 입력,  $y$ 는 출력이다. 연산자  $G$ 는 선형이므로,  $Gx = \lambda x$ 의 관계로부터  $Gx^m = \lambda^m x^m$ 임을 알 수 있다. 여기서  $m$ 은 양의 정수이고,  $\lambda$ 는 상수이다. 그러므로 연산자  $G$ 에 대한 식 (3)의 특성 방정식(characteristic equation)은 다음과 같다[3, 5].

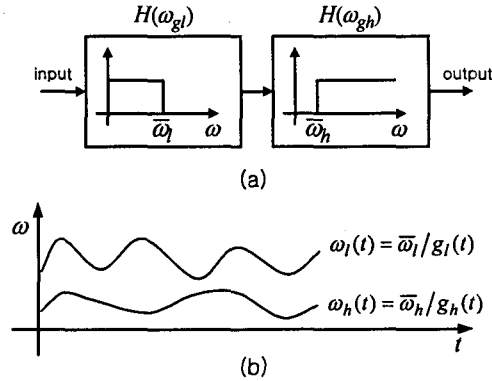
$$a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0 = 0 \quad (4)$$

식 (4)의 특성근(characteristic root)이 서로 다를 경우, 특성근  $\lambda_i = \sigma_i + j\omega_i$ 에 대응하는 고유 함수는  $x = c_i \exp(\sigma_i t + j\omega_i t)$ 이다. 식 (4)의 모든 특성근의 실수부가 음수일 경우, 식 (3)의 시스템은 안정하다[5]. 식 (3)에 표시된 시스템의 주파수 해석을 위하여 확장 푸리에 변환(EFT: extended Fourier transform) 및 역 변환을 다음과 같이 정의하자.

$$\begin{aligned} F[x(t), g(t)] &= \int_{-\infty}^{\infty} x(t) \frac{1}{g(t)} e^{\int_0^t \frac{-j\omega_x}{g(t)} dt} dt \\ &= X(\omega_g) \end{aligned} \quad (5)$$

$$\begin{aligned} F^{-1}[X(\omega_g), g(t)] &= \int_{-\infty}^{\infty} X(\omega_g) e^{\int_0^t \frac{j\omega}{g(t)} dt} d\omega_g \\ &= x(t) \end{aligned} \quad (6)$$

여기서  $\omega_g$ 는 주어진 함수  $g$ 에 대한 확장 푸리에 변수이다.



(a) 이상적인 대역통과 가변 대역폭 필터의 EFT 주파수 표현  
 (b) 시간-주파수 영역에서 표시된 통과대역

그림 1. 임의의 모양의 통과대역을 갖는 이상적인 대역통과 가변 대역폭 필터

정리: 임의의 신호  $x(t)$ 와 주어진 함수  $g(t)$ 에 대하여

$$F[G^n x(t), g(t)] = (j\omega_g)^n X(\omega_g) \tag{7}$$

이 성립한다. 여기서  $n$ 은 양의 정수이다.

증명:  $n=1$ 일 경우를 고려하자. 변수 분리법에 의하여 다음과 같은 관계식을 얻을 수 있다.

$$\begin{aligned} F[Gx(t), g(t)] &= \int_{-\infty}^{\infty} Dx(t) e^{-\int_0^t \frac{j\omega_g}{g(\tau)} d\tau} dt \\ &= j\omega_g \int_{-\infty}^{\infty} x(t) \frac{1}{g(t)} e^{-\int_0^t \frac{j\omega_g}{g(\tau)} d\tau} dt + (x(t) e^{-\int_0^t \frac{j\omega_g}{g(\tau)} d\tau}) \Big|_{-\infty}^{+\infty} \end{aligned}$$

만일,  $\lim_{t \rightarrow \pm\infty} x(t) = 0$ 이면,  $F[Gx(t), g(t)] = j\omega_g X(\omega_g)$ 이다. 그러므로 연속적으로 상기의 관계식을 적용하면,  $F[G^n x(t), g(t)] = (j\omega_g)^n X(\omega_g)$ 를 얻을 수 있다.

식 (7)로부터 식 (3)은 확장 푸리에 주파수 영역에서 다음과 같이 표현됨을 알 수 있다.

$$\begin{aligned} Y(\omega_g) &= \left( \sum_{k=0}^m b_k (j\omega_g)^k \right) / \left( \sum_{k=0}^n a_k (j\omega_g)^k \right) X(\omega_g) \\ &= H(\omega_g) X(\omega_g) \end{aligned} \tag{8}$$

여기서  $H(\omega_g)$ 는 식 (3) 시스템의 전달함수이다.

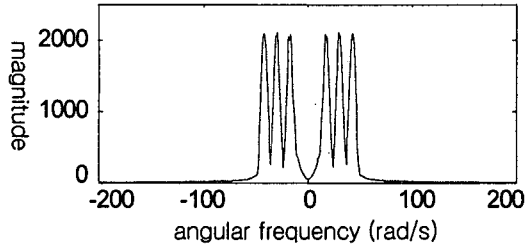


그림 2. 식 (13) 신호의 푸리에 변환

식 (8)에서 주파수 변수  $\omega_g$ 의 물리적 의미는 함수  $g(t)$ 에 따라 달라진다[3, 5].  $H(\omega_g)$ 의 차단 주파수는 계수  $a_i, b_i$ 에 의해 결정된다. 식 (6)으로부터 디락 임펄스 함수(Dirac delta function)의 역 확장 푸리에 변환은 다음과 같다.

$$F^{-1}[\delta(\omega_g - \omega_0), g(t)] = e^{\int_0^t \frac{j\omega_0}{g(t)} dt}$$

그러므로 식 (3)과 식 (8)로부터 차단 주파수가  $\bar{\omega}_g$ 인 저역통과 가변 대역폭 필터의 경우  $\omega_0 < \bar{\omega}_g$ 일 때, 상기 신호가 통과함을 알 수 있다.

### 3. 가변 대역폭 필터의 시간-주파수 영역 특성 분석 및 설계 방법

다음에 가변 대역폭 필터의 설계 방법을 살펴보자. 식 (3)의 계수  $a_i, b_i$ 에 따라 여러 가지 특성을 갖는 가변대역폭 필터를 구성할 수 있다. 예를 들어, Butterworth 다항식을 사용할 경우 Butterworth 선형 시불변 필터와 같은 주파수 특성을 갖는 가변대역폭 필터를 구현할 수 있다. 계수의 선택에 따라 저역통과 가변 대역폭 필터, 고역통과 가변 대역폭 필터, 대역통과 가변 대역폭 필터 및 대역저지 가변 대역폭 필터 등을 설계할 수 있다. 각각의 가변 대역폭 필터들의 차단 주파수는 계수  $a_i, b_i$ 에 의해 결정된다. 대역통과 필터 및 대역저지 필터는 각각 두 가지 방법으로 구현할 수 있다. 첫째, 그림 1과 같이 저역통과 가변 대역폭 필터와 고역통과 가변 대역폭 필터를 직렬 연결하여 대역통과 또는 대역저지 필터를 구현할 수 있다. 둘째, 대역통과 또는 대역저지 특성을 갖는 다항식을 사용하여 단일 형태로 구현할 수 있다. 첫 번째의 방법으로 구현할 경우 시간-주파수 영역에서 임의의 모양의 통과 대역을 갖는 대역통과 필터나 대역저지 필터를 얻을 수 있다. 두 번째 방법으로 구현할 경우 시간-주파수 영역에서 통과 대역의 상한과 하한의 비는 일정하다. 그러므로 첫 번째 방법으로 구현한 경우가 두 번째 방법으로 구현한 경우보다 잡음 제거 성능이 더 우수하다.

그림 1의 (a)는 저역통과 가변 대역폭 필터와 고역통과 가변 대역폭 필터를 직렬 연결하여 구성된 대역통과 가변 대역폭 필터를 나타낸다. 그림 1의 (b)는 통과 대역을 나타낸

다. 이와 같은 방법으로 구현한 대역통과 가변 대역폭 필터는 시간-주파수 영역에서 임의의 통과 대역을 갖음을 알 수 있다. 시간-주파수 영역에서 통과 대역의 위쪽 경계는 저역 통과 가변 대역폭 필터의 차단 주파수  $\bar{\omega}_l$ 와 함수  $1/g_l(t)$ 에 의해 결정된다. 통과 대역의 아래쪽 경계는 고역통과 가변 대역폭 필터의 차단 주파수  $\bar{\omega}_h$ 와 함수  $1/g_h(t)$ 에 의해 결정된다.

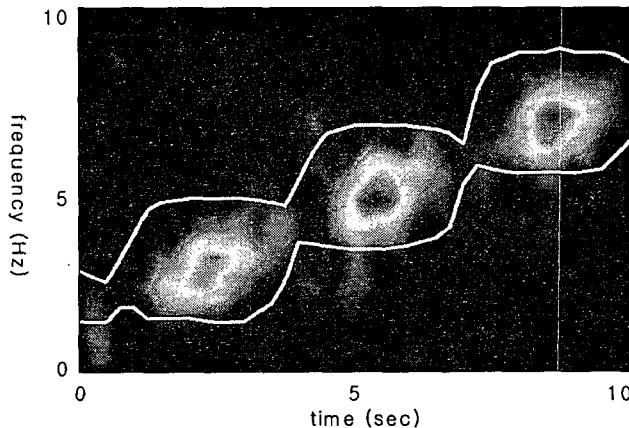
가변 대역폭 필터를 시간-주파수 영역에서 신호의 순시 대역폭의 바깥 부분에 있는 신호 또는 잡음을 제거하도록 설계할 수 있다. 시간-주파수 영역에서 통과 대역은 차단 주파수 및  $1/g(t)$ 에 의해 결정된다. 저역통과 필터의 경우 차단 주파수가  $\bar{\omega}_l$ 일 때, 통과 대역과 저지 대역의 경계는  $\bar{\omega}_l/g_l(t)$ 이다. 시간-주파수 평면  $t-\omega$ 에서 저역통과 가변 대역폭 필터의 통과 영역  $R_l$ 은 다음과 같다.

$$R_l = \{(t, \omega) \mid \omega < \bar{\omega}_l/g_l(t), \forall t\} \quad (9)$$

여기서  $\omega$ 는 푸리에 주파수 변수이다. 즉,  $g(t)=1$ 인 경우의 확장 푸리에 주파수 변수이다. 마찬가지로 고역통과 필터의 경우 차단 주파수가  $\bar{\omega}_h$ 일 때, 통과 대역과 저지 대역의 경계는  $\bar{\omega}_h/g_h(t)$ 이다. 시간-주파수 평면  $t-\omega$ 에서 고역통과 가변 대역폭 필터의 통과 영역  $R_h$ 는 다음과 같다.

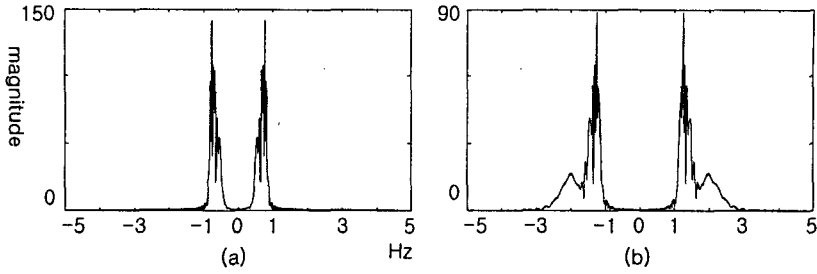
$$R_h = \{(t, \omega) \mid \omega > \bar{\omega}_h/g_h(t), \forall t\} \quad (10)$$

함수  $g_b(t)$ 를 가지고 구성된 대역통과 필터에서 차단 주파수가  $\bar{\omega}_{b1}$ ,  $\bar{\omega}_{b2}$ 일 때 ( $\bar{\omega}_{b1} < \bar{\omega}_{b2}$ ), 통과 영역  $R_b$ 는 다음과 같다.



백색 실선 : 추정된 순시 대역폭의 상한과 하한

그림 3. 백색 잡음을 포함하는 식 (13) 신호의 시간-주파수 표현



(a) 신호  $s(t)$ 의 순시 대역폭의 상한을  $1/g$ 로 택한 경우

(b) 신호  $s(t)$ 의 순시 대역폭의 하한을  $1/g$ 로 택한 경우

그림 4. 식 (8) 신호의 EFT, x축은 확장 푸리에 변수임

$$R_b = \{ (t, \omega) \mid \bar{\omega}_{b1}/g_b(t) < |\omega| < \bar{\omega}_{b2}/g_b(t), \forall t \} \quad (11)$$

식 (9)와 같은 통과 대역을 갖는 저역통과 필터와 식(10)과 같은 통과 대역을 갖는 고역통과 필터를 직렬 연결하여 구성된 대역통과 필터의 통과 영역  $R_{b1}$ 은 다음과 같다.

$$R_{b1} = \{ (t, \omega) \mid \bar{\omega}_h/g_h(t) < |\omega| < \bar{\omega}_l/g_l(t), \forall t \} \quad (12)$$

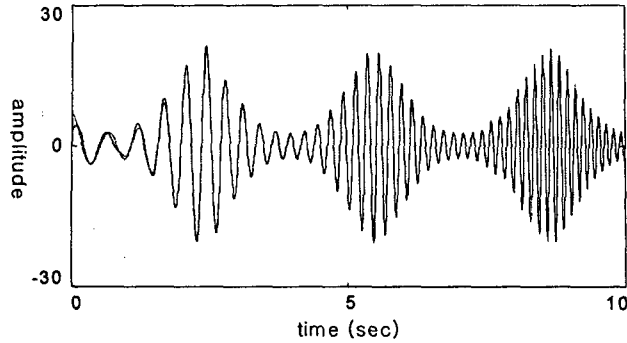
저역통과 및 고역통과 필터를 직렬 연결하여 대역통과 필터를 구성할 경우 통과 대역의 위쪽 경계 및 아래쪽 경계를 임의로 선택할 수 있다. 그러므로 직렬 연결하여 구성된 대역통과 필터의 경우 필터의 잡음 제거 성능이 매우 우수하다. 일반적으로 가변 대역폭 필터 설계를 위하여 Butterworth 다항식뿐만 아니라 Chebyshev 다항식, Bessel 다항식 등 여러 가지가 사용될 수 있다.

#### 4. 시변 순시 주파수 및 시변 순시 대역폭을 갖는 신호의 복구

가변 대역폭 필터의 시간-주파수 영역에서의 특성과 확장 푸리에 주파수 특성을 알아보기 위해 다음과 같이 시간적으로 변하는 진폭을 갖는 선형 칩(chirp) 신호를 고려하자.

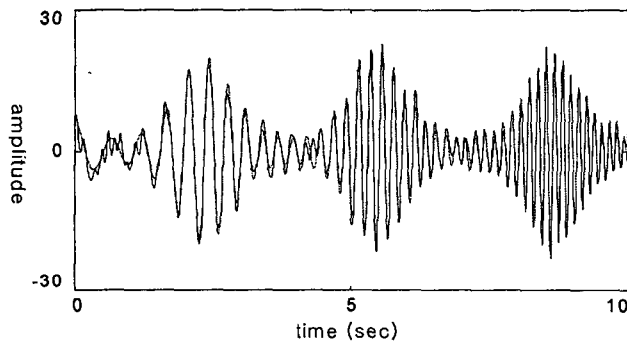
$$s(t) = \exp(2 - 2\sin t \cos t) \cos(8t + 2t^2) \quad (13)$$

식 (13)의 AM-FM 신호의 위상은  $\phi(t) = 8t + 2t^2$ 이므로, 순시 주파수는  $\omega(t) = 8 + 4t$  rad/s이다. 가변 대역폭 필터의 특성을 효과적으로 관찰하기 위하여 시간에 따라 변하는 순시 주파수와 순시 대역폭을 갖는 단일 성분 신호를 대상으로 시뮬레이션을 수행하였다.



실선 : 원 신호, 점선 : 복구한 신호

그림 5. 백색 잡음이 있는 신호  $s(t)$ 를 그림 1과 같은 대역통과 가변 대역폭 필터를 사용하여 복구한 신호



실선 : 원 신호, 점선 : 복구한 신호

그림 6. 백색 잡음이 있는 신호  $s(t)$ 를 선형 시불변 필터를 사용하여 복구한 신호

그림 2는 신호  $s(t)$ 의 푸리에 변환을 나타낸다. 그림 3은 백색잡음을 포함하는 신호  $s(t)$ 의 스펙트로그램이다. 여기서 순시 주파수가 시간에 따라 선형적으로 증가함을 알 수 있다. 또한 신호의 순시 대역폭이 시간에 따라 변함을 알 수 있다. 그림 3의 백색 실선은 추정된 순시 대역폭의 위쪽 경계와 아래쪽 경계를 나타낸다. 여기서 순시 대역폭의 경계는 시간의 함수임을 알 수 있다.

그림 4의 (a)는  $1/g(t)$ 를 그림 3에 나타낸 순시 대역폭의 상한으로 택하였을 때, 신호  $s(t)$ 의 EFT이다. 또한 그림 4의 (b)는  $1/g(t)$ 를 그림 3에 나타낸 순시 대역폭의 하한으로 택하였을 때, 신호  $s(t)$ 의 EFT이다. 그림 4의 (a)에서 대부분의 스펙트럼이 주파수 범위  $|\omega_d| < 1$  내에 있음을 알 수 있다. 또한 그림 4의 (b)에서 대부분의 스펙트럼이 주파수 범위  $|\omega_d| > 1$  내에 있음을 알 수 있다. 그러므로 그림 4의 (a)로부터 차단 주파수가  $\bar{\omega}_1 = \pm 1$ 이고 함수  $1/g(t)$ 를 그림 3에 나타낸 순시 대역폭의 상한으로 선택한 저역통과 가변 대역폭 필터에 신호  $s(t)$ 를 적용하면 신호의 순시 대역폭의 위쪽 경계의 바깥 부분에 있는 잡음을 제



거할 수 있음을 알 수 있다. 또한 그림 4의 (b)로부터 차단 주파수가  $\bar{\omega}_h = \pm 1$ 이고 함수  $1/g_h(t)$ 를 그림 3에 나타낸 순시 대역폭의 하한으로 선택한 고역통과 가변 대역폭 필터에 신호  $s(t)$ 를 적용하면 신호의 순시 대역폭의 아래쪽 경계의 바깥부분에 있는 잡음을 제거할 수 있음을 알 수 있다.

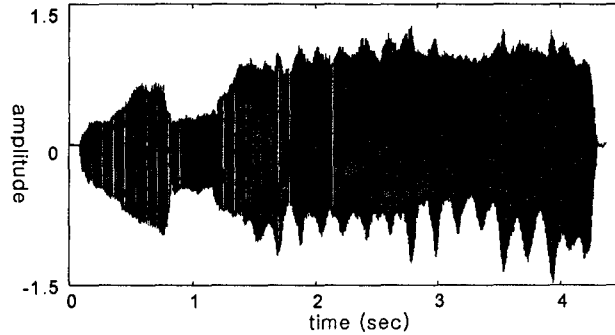
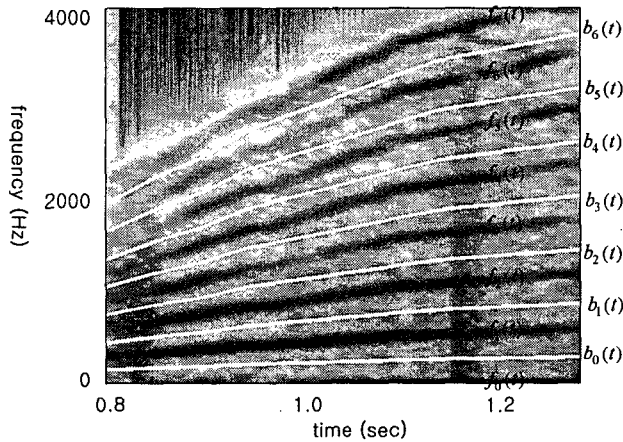


그림 7. 30세 성인 남성의 음성신호(모음 “아”)

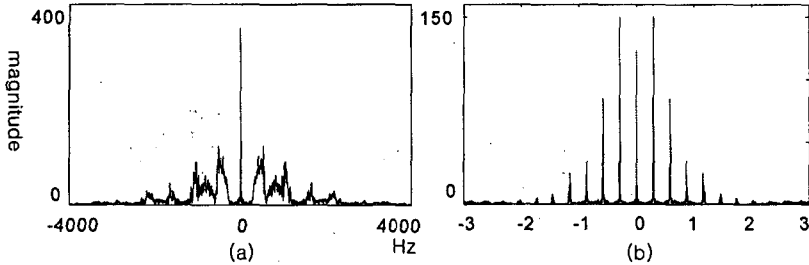


백색 실선 : 각 성분별 경계

그림 8. 그림 7의 신호의 일부분의 시간-주파수 표현(dB)

그림 5는 그림 1과 같은 대역통과 가변 대역폭 필터를 사용하여 백색 잡음이 섞인 신호  $s(t)$ (식 (13))의 잡음을 제거한 신호를 나타낸다. 여기서 저역통과 가변 대역폭 필터는 차단 주파수를  $\bar{\omega}_l = \pm 1$ 로 하고  $1/g_l(t)$ 를 그림 3에 표시한 순시 대역폭의 위쪽 경계로 선택하여 구현하였고, 고역통과 가변 대역폭 필터는 차단 주파수를  $\bar{\omega}_h = \pm 1$ 로 하고  $1/g_h(t)$ 를 그림 3에 표시한 순시 대역폭의 아래쪽 경계로 선택하여 구현하였다. 그림 6은 차단 주파수가  $\omega_c = 62.8$  rad/s인 선형 시분변 저역통과 필터를 이용하여 백색 잡음이 섞인 신호  $s(t)$ 의 잡음을 제거한 신호를 나타낸다. 실험에 의하면 가변 대역폭 필터를 이용하여 잡음을 제거한 경우

신호 대 잡음비는 21.4 dB이고, 선형 시불변 필터를 이용하여 잡음을 제거한 경우에는 신호 대 잡음비가 13.8 dB이다. 따라서 선형 시불변 필터보다 가변 대역폭 필터의 잡음 제거 성능이 매우 뛰어나다는 것을 알 수 있다.



(a) 푸리에 변환, (b) EFT,  $b_3(t)$ 를  $1/g(t)$ 로 택한 경우

그림 9. 그림 7의 음성신호의 일부분의 주파수 표현( $0.8 < t < 1.3$ )

## 5. 음성신호의 각 성분별 분리 및 진폭과 순시 주파수 추정

그림 7은 성인 남성의 노래의 일부분으로 모음 “아”의 시간 영역 파형이다. 샘플링 주파수가 8 kHz인 16-bit A/D 컨버터를 사용하였다. 그림 8은 그림 7의 시간  $t=0.8$  sec 및  $t=1.3$  sec 사이에 있는 신호의 시간-주파수 표현이다. 여기서 8 개의 시변 순시 주파수를 관찰할 수 있다. 각 순시 주파수를  $f_0(t), \dots, f_7(t)$ 라 하자(그림 8 참고). 그림 8에 표시된  $b_0(t), \dots, b_6(t)$ 는 각 시변 순시 주파수를 갖는 신호를 분리하기 위해 사용되는 대역통과 가변 대역폭 필터의 통과대역의 상한과 하한이다. 예를 들어, 두 번째 순시 주파수  $f_1(t)$ 를 분리하기 위해  $b_0(t)$ 를 아래쪽 경계로 하고,  $b_1(t)$ 를 위쪽 경계로 하는 대역통과 가변 대역폭 필터를 사용한다. 그림 9의 (a)는 그림 7의 시간  $t=0.8$  sec 및  $t=1.3$  sec 사이에 있는 신호의 푸리에 변환이다. 그림 9의 (b)는  $b_3(t)$ 를  $1/g(t)$ 로 택하여 음성신호를 확장 푸리에 변환한 것을 나타낸다. 그림 9의 (a)에서 각 성분별 푸리에 스펙트럼이 서로 겹치므로 음성신호의 각 성분을 분리하기가 매우 어려움을 알 수 있다. 그러나 그림 9의 (b)에서는 각 순시 주파수 성분이 서로 분리되어 있음을 확인할 수 있다. 그림 10은 대역통과 가변 대역폭 필터를 사용하여 음성신호를 각 성분별로 분리한 신호를 나타낸다.

AM-FM 신호  $x(t) = a(t)\cos(\Phi(t))$ 의 순시 주파수와 진폭을 추출하기 위하여 흔히 Hilbert 변환을 사용한다. Hilbert 변환은 다음과 같이 정의된다[15].

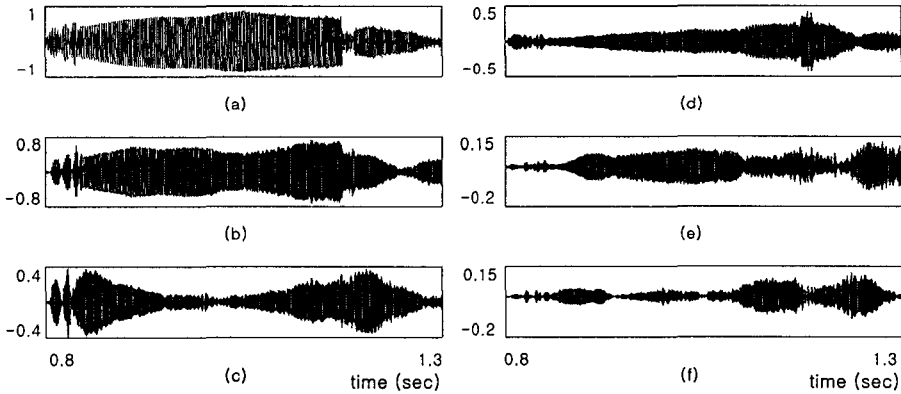
$$H[x(t)] = c.p.v. \int_{-\infty}^{\infty} (x(\tau) / \pi(t-\tau)) d\tau \quad (14)$$

여기서  $c.p.v.$ 는 Cauchy 주치(principle value)이다.  $a(t)$ 와  $\cos(\Phi(t))$ 의 푸리에 스펙트럼이 서

로 겹치지 않고,  $\cos(\phi(t))$ 의 푸리에 스펙트럼이 매우 좁은 영역에 분포할 경우 다음과 같은 식을 얻을 수 있다[15-18].

$$\begin{aligned} z(t) &= x(t) + jH[x(t)] \\ &\cong x(t) + ja(t)H[\cos(\phi(t))] \\ &\cong x(t) + ja(t)\sin(\phi(t)) \\ &= a(t)\exp(j\phi(t)) \end{aligned} \quad (15)$$

음성신호의 각 성분의 진폭과 각 성분의 위상부분의 푸리에 스펙트럼이 서로 겹치지 않고, 위상부분의 푸리에 스펙트럼이 매우 좁은 영역에 분포한다고 가정하자. 이와 같은 가정 하에 식 (15)를 이용하여 그림 10에 나타낸 각 신호들의 진폭 및 순시 주파수를 구할 수 있다. 그림 11은 Hilbert 변환을 사용하여 구한 각 성분별 진폭을 나타낸다. 그림 12는 각 성분별로 추정된 순시 주파수를 보여준다.



(a)  $1/g_h=b_0, 1/g_l=b_1$ , (b)  $1/g_h=b_1, 1/g_l=b_2$ , (c)  $1/g_h=b_2, 1/g_l=b_3$ ,  
(d)  $1/g_h=b_3, 1/g_l=b_4$ , (e)  $1/g_h=b_4, 1/g_l=b_5$ , (f)  $1/g_h=b_4, 1/g_l=b_6$

그림 10. 대역통과 가변 대역폭 필터를 사용하여 각 성분별로 분리한 신호

## 6. 결론 및 향후 연구

음성신호의 순시 주파수가 시간에 따라 변하는 경우 각 성분별 푸리에 스펙트럼이 서로 겹쳐지므로 각각의 순시 주파수에 해당되는 신호성분을 분리하기가 어렵다. 본 연구에서는 음성신호에서 시변 순시 주파수를 갖는 각 신호성분을 분리할 수 있는 가변 대역폭 필터를 제안하였으며, 이를 사용하여 각 성분별로 음성신호를 분리하였다. 또한 Hilbert 변환을 사용하여 각 성분별 진폭 및 시변 순시 주파수를 추정하였으며, 이를 스펙트로그램에서 나타난 시변 순시 주파수와 비교하였다. 음성신호에서 시변 순시 주파수를 갖는 각 신호 성분을 분

리할 경우 음성인식이나 음성코딩에 매우 유용하게 사용할 수 있다.

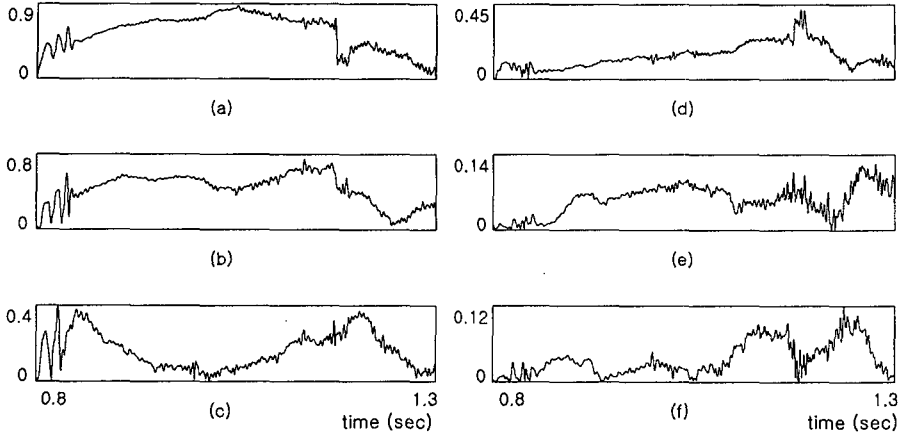


그림 11. Hilbert 변환을 사용하여 추정한 각 신호 성분별 진폭

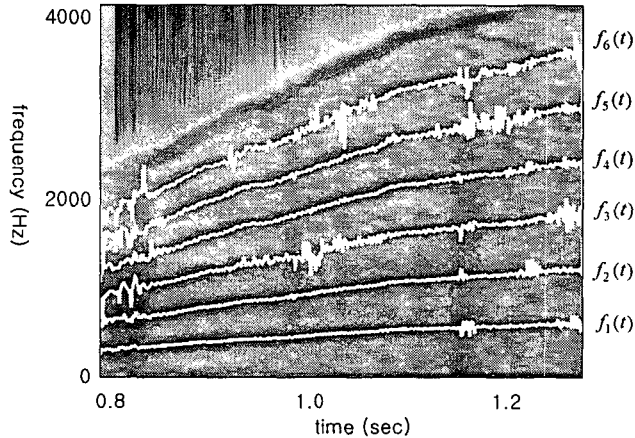


그림 12. Hilbert 변환을 사용하여 추정한 각 신호성분별 시변 순시 주파수

또한 본 연구에서는 가변 대역폭 필터를 설계하는 방법을 제안하였으며, 가변 대역폭 필터의 특성을 시간-주파수 영역에서 고찰하였다. 시간-주파수 영역에서 임의의 통과대역을 갖는 가변 대역폭 필터를 설계할 수 있다. 기존의 방법을 사용할 경우 다중 신호의 각 성분들이 시간-주파수 영역에서 겹쳐지지 않으나 푸리에 주파수 영역에서 겹쳐질 경우 다중 신호를 각 성분별로 분리해 내기가 매우 어렵다. 그러나 이와 같은 신호를 확장 푸리에 주파수 영역에서 관찰할 경우 각 신호성분의 확장 푸리에 스펙트럼은 서로 겹쳐지지 않으므로 각 성분별로 쉽게 분리할 수 있다. 신호의 시간-주파수 표현을 얻기 위해 사용되는 창 함수의 크기에 따라 신호의 스펙트로그램이 달라지므로 가변 대역폭 필터의 성능을 향상시키기 위하여 창 함수의 효과를 제거하기 위한 체계적인 연구가 필요하다.

## 참 고 문 헌

- [1] Rao, A. & R. Kumaresan, 2000, "On Decomposing Speech into Modulated Components." *IEEE Tr. on Speech and Audio Processing*, Vol. 8, No. 3, 240- 254.
- [2] Boashash, Boualem. 1992, *Time-frequency Signal Analysis*. published by Longman Cheshire.
- [3] Lee, H. & Z. Bien. 1998. "Reconstruction of Signals with Known Instantaneous Frequency using Linear Time-varying Filter." *Electronics Letters*, Vol. 34, No. 24, 2312-2313.
- [4] Qian, Shie. & D. Chen. 1996. *Joint Time-frequency Analysis: Methods and Applications*. published by Prentice-Hall.
- [5] Lee, H. & Z. Bien. 1998 "On the Eigenstructure of Linear Quasi-time-invariant Systems," *Int. J. Syst. Sci.*, Vol. 29, No. 8, 873-887.
- [6] Griffin, D. W. & J. S. Lim. 1984. "Signal Estimation from Modified Short-time Fourier Transform," *IEEE Tr. on ASSP*, Vol. 32, No. 2, 236-243.
- [7] Portnoff, M. R. 1980. "Time-frequency Representation of Digital Signals and Systems based on Short-time Fourier Analysis," *IEEE Tr. on ASSP*, Vol. 28, No. 1, 55-69.
- [8] Cunningham, G. S. & W. J. Willians. 1996. "Vector-valued Time-frequency Representations," *IEEE Tr. on SP*, Vol. 44, No. 7, 1642-1656.
- [9] Qian, Shie., M. E. Dunham. & M. J. Freeman. 1995. "Transionospheric Signal Recognition by Joint Time-frequency Representation." *Radio Science*, Vol. 30, No. 6, 1817-1829.
- [10] Rabiner L. R. & R. W. Schafer. 1978. *Digital Processing of Speech Signals*. published by Prentice-Hall.
- [11] O'Shaughnessy, Douglas. 2000 *Speech Communications Human and Machine*. published by Piscataway.
- [12] Quatieri, T. F., T. E. Hanna. & G. C. O'Leary. 1997. "AM-FM Separation using Auditory Motivated Filters." *IEEE Tr. on Speech Audio Processing*, Vol. 5, No. 4, 465-480.
- [13] Ingle, Vinay K. & John G. Proakis. 1997. *Digital Signal Processing* published by PWS.
- [14] H, James., R. McClellan., W. Schafer. & Mark A. Yoder. 1998. *DSP First, A Multimedia Approach* published by Prentice-Hall.
- [15] Boashash, Boualem. 1992. "Estimating and Interpreting the Instantaneous Frequency of a Signal-part I, II: Fundamentals." *Proc. IEEE*, Vol. 80, No. 4, 520-568.
- [16] Bedrosian, Edward. 1962. "The Analytic Signal Representation of Modulated Waveforms." *Proc. IRE*, Vol. 50, No. 10, 2071-2076.
- [17] Picinbono, Bernard. 1997. "On Instantaneous Amplitude and Phase of Signals." *IEEE Tr. on Signal Processing*, Vol. 45, No. 3, 552-560.
- [18] Vakman, David. 1996. "On the Analytic Signal, the Teager-Kaiser Energy Algorithm, and Other Methods for Defining Amplitude and Frequency." *IEEE Tr. on Signal Processing*, Vol. 44, No. 4, 791-797.
- [19] Almeida, L. B. & J. M. Tribolet. 1983. "Nonstationary Spectral Modeling of Voiced Speech." *IEEE Tr. on ASSP*, Vol. 31, No. 6, 664-678.
- [20] McAulay, R. J. & T. F. Quatieri. 1986, "Speech Analysis-synthesis based on a Sinusoidal Representation" *IEEE Tr. on ASSP*, Vol. 34, No. 7, 744-754.

접수일자: 2001. 10. 17.

게재결정: 2001. 11. 30.

▲ 송 민

광주광역시 북구 용봉동 300번지 (우: 500-757)

전남대학교 전자공학과

Tel: +82-62-530-0132 Fax: +82-62-530-1759

E-mail: u0020641@chonnam.ac.kr

▲ 이희영

광주광역시 북구 용봉동 300번지 (우: 500-757)

전남대학교 정보통신공학부

Tel: +82-62-530-1751 Fax: +82-62-530-1759

E-mail: leehy@chonnam.ac.kr