

# 음성 신호의 단구간 에너지 차를 이용한 새로운 프리엠퍼시스 방법에 관한 연구

論 文

50D-12-5

## A Study on a New Pre-emphasis Method Using the Short-Term Energy Difference of Speech Signal

金東浚\* · 金珠里\*\*  
(Dong-Jun Kim · Ju-Lee Kim)

**Abstract** - The pre-emphasis is an essential process for speech signal processing. Widely used two methods are the typical method using a fixed value near unity and the optimal method using the autocorrelation ratio of the signal.

This study proposes a new pre-emphasis method using the short-term energy difference of speech signal, which can effectively compensate the glottal source characteristics and lip radiation characteristics. Using the proposed pre-emphasis, speech analysis, such as spectrum estimation, formant detection, is performed and the results are compared with those of the conventional two pre-emphasis methods. The speech analysis with 5 single vowels showed that the proposed method enhanced the spectral shapes and gave nearly constant formant frequencies and could escape the overlapping of adjacent two formants. Comparison with FFT spectra had verified the above results and showed the accuracy of the proposed method. The computational complexity of the proposed method reduced to about 50% of the optimal method.

**Key Words** : Pre-emphasis, The short-term energy difference, Speech analysis

### 1. 서 론

음성 연구영역의 여러 분야에서 신호처리의 앞 단계에서 거의 필수적으로 적용되는 프리엠퍼시스(pre-emphasis)는 전체의 성능에 비교적 큰 영향을 미치는 과정임에도 불구하고 상대적으로 연구가 그다지 활발하지 않았다. 선형 음성 발생 모델로부터 성문의 부피 속도(glottal volume velocity) 파형과 입술의 방사 특성의 영향에 의해 유발되는 스펙트럼 기울기를 보상하기 위해서는 프리엠퍼시스가 필요하다. Wakita[1, 2]에 의해 행해진 실험은 적당한 프리엠퍼시스 함수가  $(1 - \mu z^{-1})$ 이라는 것을 보여주었다. 여기서  $\mu$ 는 1에 가까운 값이다. 일반적으로 널리 쓰이고 있는 프리엠퍼시스 방법은 프리엠퍼시스 계수를 0.9에서 1사이의 값으로 고정하는 방법과 음성 신호의 자기 상관 함수의 비를 이용하는 최적(optimal) 방법이다[3, 4]. 계수를 고정하는 경우에는 연구자에 따라 다르지만 0.94가 전형적인 값으로 알려져 있다. 그리고 자기 상관 함수의 비도 실제 구해보면 거의 0.9와 1사이의 값을 갖는 경우가 많고, 이는 자음이 섞여 있는 경우

에도 정도의 차이만 있을 뿐이지 마찬가지이다. 이와 같이 일반적으로 사용되는 두 가지 프리엠퍼시스 방법은 구현이 쉽다는 장점은 있지만, 프리엠퍼시스 필터의 전달 특성을 고려할 때 모음의 정상 상태에서만 정확한 결과를 낼 수 있고, 음성의 시작 부분과 천이 부분과 같이 신호가 안정(stationary)하다고 보기 어려운 부분에서는 부정확한 결과를 낼 가능성이 매우 크다.

보통 음성 신호 분석을 정확하고 효과적으로 수행하기 위해서는 음성 발생 메커니즘의 생리학적 연구로부터 나온 정보들을 잘 이용하는 것이 중요하다. 성문 면적과 부피 속도의 파형은 시작 부분에서 낮은 진폭을 가지다가 차츰 정상 상태에 이르며, 성문의 부피 속도는 입술 근처에서의 음압 파형과 거의 비례하게 나타난다[3]. 여기서 입술 근처에서의 음압은 곧 입술을 통해 방사되어 음성 신호를 형성하므로 음파의 단구간에서의 평균전력 또는 에너지는 성문의 크기와 거의 비례한다고 볼 수 있을 것이다. 따라서 음파의 단구간 에너지를 고려하여 프리엠퍼시스 계수를 설정할 수 있으면 음성 신호에서 음원 특성과 방사 특성을 더 효과적으로 제거할 수 있을 것이다.

본 연구에서는 성문의 음원 특성과 입술의 방사 특성을 더 효과적으로 제거하기 위하여, 성대 떨림이 시작되는 모음 신호의 초기에서 정상상태까지의 단구간 에너지와 그 차를 관찰하여 이를 통하여 음성 신호의 단구간 에너지 차(short-term energy difference : STED)를 이용한 프리엠퍼

\* 正 會 員 : 淸州大 理工大 電子·情報通信·半導體工學部  
副教授·工博

\*\* 準 會 員 : 淸州大 理工大 電子工學科 碩士  
接受日字 : 2001年 8月 29日  
最終完了 : 2001年 10月 23日

시스 기법을 제안한다. 제안한 프리엠퍼시스 방법을 이용하여 스펙트럼 추정 및 형성음 주파수 추출 실험을 수행하고, 기존에 널리 쓰이는 두 가지 프리엠퍼시스 방법과 그 성능을 비교·분석함으로써 제안된 방법의 정확성을 확인하고자 하였다. 이를 위하여 기존의 프리엠퍼시스 방법과 제안한 방법을 적용하여 각각 Burg 알고리즘을 이용한 선형 예측 분석을 하고, 스펙트럼 추정 및 형성음 주파수 추출 등의 음성 분석 실험을 수행하여 결과를 분석한다.

## 2. 프리엠퍼시스

### 2.1. 기존의 프리엠퍼시스

무성음과 유성음에 대한 음성 발생 과정의 주파수 특성은 그림 1과 같다. 그림에서 유성음의 경우 음원이 임펄스의 열로 표현되고, 무성음의 경우는 잡음원으로 표현된다.

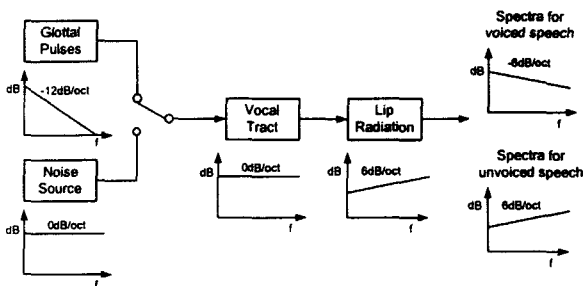


그림 1. 음성 발생 과정의 주파수 특성.

Fig. 1. Frequency characteristics of speech production process.

유성음의 경우 성문 부피 속도 파형은 약 -12[dB/oct] 기울기로 모델링되고, 입술의 방사 특성은 약 +6[dB/oct] 기울기로 모델링되므로 음원의 특성과 입술 방사 특성의 영향에 의해 유발되는 스펙트럼 기울기를 제거하기 위해서 프리엠퍼시스가 필요하다[3, 4].

일반적으로 널리 쓰이고 있는 프리엠퍼시스 방법은 프리엠퍼시스 계수를 0.9에서 1사이의 값으로 고정하는 방법과 자기 상관 함수의 비를 이용하는 최적 방법이다.

### 2.2. 단구간 에너지 차를 이용한 프리엠퍼시스

일반적으로 널리 사용되는 두 가지 프리엠퍼시스 방법은 구현하기 쉽고, 정상 상태에서 정확한 결과를 낼지라도, 음성의 시작 부분과 천이 부분에서는 부정확한 결과를 낼 가능성이 있다. 그러므로 본 연구에서는 음성의 정상 상태뿐만 아니라 시작 부분과 천이 부분에 대한 성문의 음원 특성과 입술에서의 방사 특성을 효과적으로 제거하기 위해 음성 신호의 단구간 에너지 차를 이용한 프리엠퍼시스 기법을 제안한다.

성문 면적과 부피 속도가 시작 부분에서 낮은 진폭을 가지다가 차츰 정상 상태에 이르는 것을 볼 수 있으며, 성문 부

피 속도가 입술에서의 음압 파형과 비례하게 나타난다[3]. 입술에서의 음압은 곧 입술로 방사되어 음성 신호를 형성하므로 음파의 에너지가 성문파의 크기와 거의 비례한다고 볼 수 있다. 따라서 음성의 단구간 에너지는 시작 부분에서는 작은 값을 가지다가 천이 구간에서 점점 증가하여 정상상태에 이르면 일정한 값으로 유지된다. 이에 따른 단구간 에너지 차 역시 대부분 시작 부분에서 큰 값을 가지고 천이 구간에서 점점 감소하다가 정상 상태에 이르면 거의 0에 가까운 값으로 감소하게 된다. 본 연구에서는 이러한 음성의 단구간 에너지 차를 이용하여, 단구간 에너지 차에 따른 프리엠퍼시스 계수  $\mu$ 를 설정함으로써 시작 부분과 천이 구간에 대하여 보다 정확한 보상을 해주게 될 것이다.

음성 신호  $s(n)$ 에 대한 단구간 에너지는 식(1)과 같고, 단구간 에너지 차는 식(2)와 같다.

$$E(i) = \sum_{n=iN}^{iN+N-1} s^2(n) \quad (1)$$

$$Ed(i) = E(i+1) - E(i) \quad (2)$$

여기서,  $i$ 는 프레임 인덱스를 나타내고,  $N$ 은 프레임의 길이를 의미한다.

본 연구에서는 적용 1차 역 필터를 토대로 STED 프리엠퍼시스 필터를 개발하였으며, 이는 식(3)과 같다.

$$F(z) = 1 - \mu_i z^{-1} \quad (3)$$

여기서,  $\mu_i$ 는  $i$ 번째 프레임의 역 필터 계수이다.

일반적으로 단구간 에너지 차는 시작 부분에서 큰 값을 가지므로, 처음 세 개 프레임에서의 단구간 에너지 차 중에서 가장 큰 값을 최대값으로 간주하고, 이 최대값으로 모든 프레임의 에너지 차를 정규화시킨다. 따라서 최대 단구간 에너지 차와 각 프레임에서의 단구간 에너지 차의 비에 따라  $\mu_i$  값을 설정하게 되는데, 단구간 에너지 차의 비가 클수록 시작 부분에 해당하므로 작은  $\mu_i$  값을 설정하고, 단구간 에너지 차의 비가 작을수록 정상 상태에 해당하므로 큰  $\mu_i$  값을 설정함으로써 보다 정확한 프리엠퍼시스를 하고자 하였다.

## 3. 선형 예측 분석

음성의 음향적 특성을 가장 성공적으로 나타낸 모델 중의 하나는 Fant의 선형 음성 발생 모델이며[4], 식 (4)와 같이 나타낼 수 있다.

$$S(z) = E(z) G(z) V(z) R(z) \quad (4)$$

여기서,  $S(z)$  음성 신호 모델의  $z$ -변환이고,  $E(z)$ ,  $G(z)$ ,  $V(z)$ ,  $R(z)$ 는 각각 여기 모델(excitation model), 성문파

형성 모델(glottal shaping model), 성도 모델(vocal tract model), 입술 방사 모델(lip radiation model)을 나타낸다. 식 (5)의 역필터를 정의하면,

$$A(z) = \sum_{i=0}^M a_i z^{-i}, \quad a_0 = 1 \quad (5)$$

$$\cong \frac{1}{G(z)V(z)R(z)}$$

이고, 여기서  $M$ 은 필터의 차수를 나타낸다. 이 역필터를 이용하면 식 (4)는 다음과 같은 분석 모델로 나타낼 수 있다.

$$E(z) = S(z) A(z) \quad (6)$$

식 (6)을 샘플링된 데이터 영역에서 표현하면,

$$e(n) = \sum_{i=0}^M a_i s(n-i)$$

$$= s(n) + \sum_{i=1}^M a_i s(n-i) \quad (7)$$

$$= s(n) - \hat{s}(n)$$

이고, 여기서  $-a_i$ 는 예측기 계수(predictor coefficients),  $\hat{s}(n)$ 은 예측된 샘플이다.

본 연구에서는 선형 예측 분석을 위하여 격자(lattice) 방법인 Burg 알고리즘을 이용한다[5].

#### 4. 실험 및 결과 고찰

##### 4.1. 실험

본 연구에서 음성 신호의 스펙트럼 추정 및 형성음 주파수 추출 실험을 위하여 다음과 같은 음성 신호 분석 과정을 수행하였다.

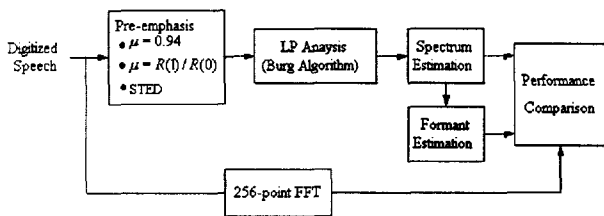


그림 2. 음성 신호 분석의 블록 선도.  
Fig. 2. Block diagram of speech analysis.

저잡음 환경에서 녹음된 음성 신호는 차단 주파수가 4.7[kHz]인 저역 통과 필터를 통과하여 10[kHz]로 샘플링되고, 프리엠퍼시스된 후 10차로 구성된 Burg 알고리즘에 의하여 반사계수가 계산된다. 반사계수는 Levinson recursion에 의하여 선형예측계수로 변환되어 스펙트럼 추정이 이루어진다.

추정된 스펙트럼으로부터 1차와 2차 미분을 이용한 침투치 검출에 의하여 형성음 주파수가 추출된다[6]. 실험에 사용한 음성은 5개(단모음) × 4명(화자) × 3번(발음) = 60개의 데이터를 사용하였으며, 음성 분석을 위한 프레임은 모음의 경우 일반적으로 25 또는 30[msec] 이내의 길이에서 WSS(wide sense stationary)하다고 가정하므로 20[msec]로 설정하였고, 이는 200 샘플에 해당한다. 본 연구에서 제안한 프리엠퍼시스 방법과 기존의 방법에 의해 추정된 스펙트럼 및 형성음 주파수의 정확성을 확인하기 위하여 256-point FFT 스펙트럼의 결과와 비교한다.

본 연구에서는 프리엠퍼시스 계수를 단구간 에너지 차를 이용하여 결정하게 된다. 단구간 에너지 차는 발음의 종류와 신호의 세기에 따라 크기가 다를 것이지만, 발음의 앞 부분에서는 차가 크다가 점점 작아져서 모음의 정상상태(steady-state)에 이르면 변도가 매우 작아질 것이다. 그림 3은 모음 /아/의 원파형, 단구간 에너지, 및 단구간 에너지 차를 나타내고 있다. 그림에서 알 수 있듯이, 단구간 에너지의 최대값은 발음의 초기에 크다가 정상상태에 이르면 작아진다. 따라서 본 연구에서는 처음 3프레임 이내에 최대값이 있다고 보고 최대값을 구한다. 그리고 나서 그 최대값으로 전체 단구간 에너지 차를 정규화하여 정규화된 단구간 에너지 차를 토대로  $\mu_i$  값을 결정한다. 즉, 정규화된 단구간 에너지 차가 크면 발음의 시작 부분에 해당되므로 작은  $\mu_i$  값을 설정하고, 값이 작아질수록  $\mu_i$  값을 증가시키게 된다.

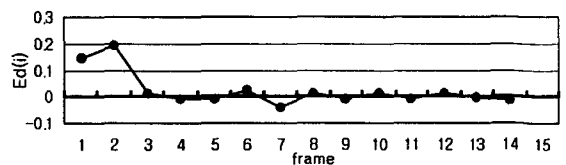
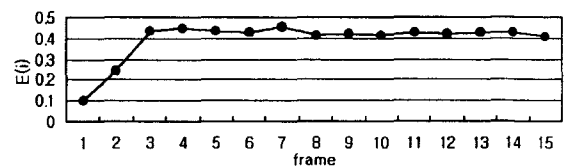
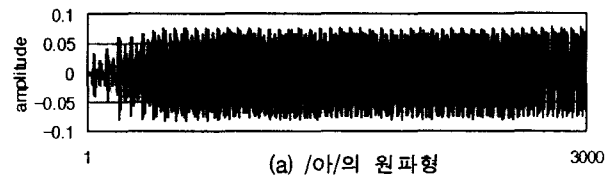


그림 3. /아/의 단구간 에너지 및 단구간 에너지 차.

Fig. 3. Short-term energy, and short-term energy difference for /a/.

정규화된 단구간 에너지 차에 따른  $\mu_i$  값은 표 1에서 보는 바와 같이, 에너지 차가 크면 발음의 시작부분으로 보고 작은 값을 설정하고, 실험적으로 설정한 문턱치인 0.4이하이면

정상상태에 이른 것으로 간주하여 1로 설정하였다.  $Ed_{max}$ 는 단구간 에너지 차의 최대값을 의미한다.

표 1. 정규화된 단구간 에너지 차에 따른  $\mu_i$  값 설정.

Table 1. Determination of  $\mu_i$  using normalized short-term energy difference.

$Ed/Ed_{max}$	1.0~0.8	0.8~0.6	0.6~0.4	0.4 이하
$\mu_i$	0.3	0.5	0.7	1.0

정규화된 단구간 에너지가 크면 발음의 앞 부분에 해당되어 작은 값을 설정하고, 점차 증가시키다가 특정 문턱값(0.4)보다 작으면 정상상태에 이른 것으로 보고 1로 설정한다.

그림 4는 표 1에 의해 설정된 계수를 갖는 STED 프리엠퍼시스 필터의 진폭 응답 특성을 나타낸 그림이다. 그림에서 보는 바와 같이, 계수에 따라 프리엠퍼시스 정도가 다르게 나타남을 알 수 있다. 따라서 시작 부분에서는 프리엠퍼시스 정도를 작게 해주고, 정상 상태에서는 프리엠퍼시스 정도를 크게 해주게 된다.

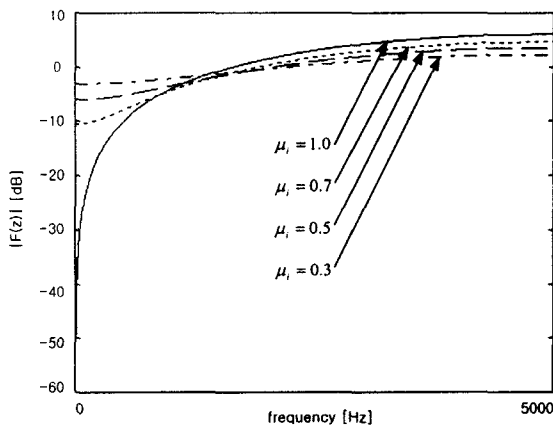
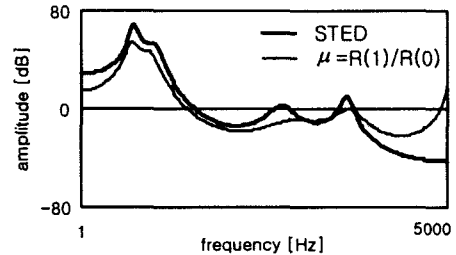
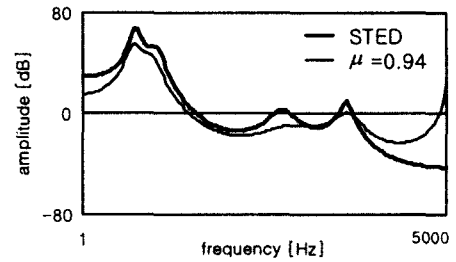


그림 4. STED 프리엠퍼시스 필터의 진폭 응답 특성.

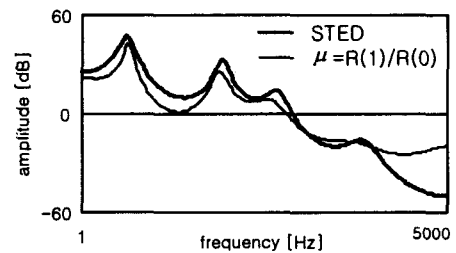
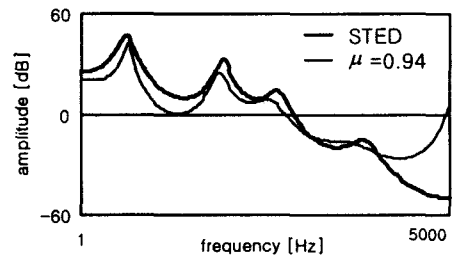
Fig. 4. Characteristic of amplitude response for STED pre-emphasis filter.

#### 4.2. 결과 고찰

세 가지 프리엠퍼시스 방법을 비교하기 위해서 각 방법에 대한 음성 신호의 스펙트럼 추정 실험을 수행하였다. 그림 5는 단모음 /아/, /에/에 대한  $\mu=0.94$ ,  $\mu=R(1)/R(0)$ , 및 STED 방법을 적용하여 분석 과정을 거친 후의 스펙트럼 추정 결과를 시작 부분에 대해 보여준다.



(a) /아/의 시작 부분



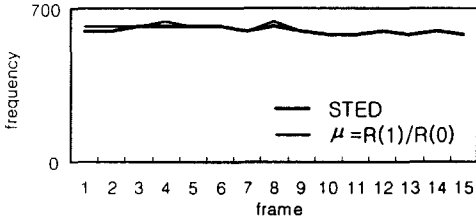
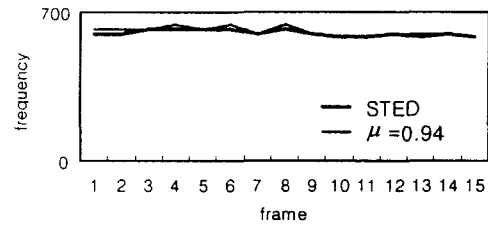
(b) /에/의 시작 부분

그림 5. 세 가지 프리엠퍼시스 방법에 의한 시작 부분 스펙트럼.

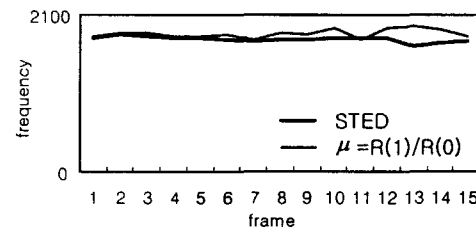
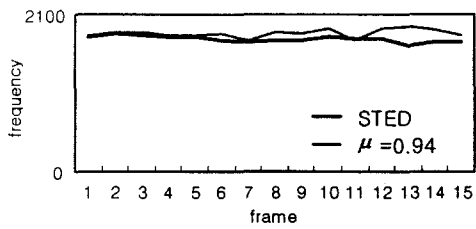
Fig. 5. Spectra of 3 pre-emphasis methods for starting part.

그림에서 보는 바와 같이, STED 방법과 기존의 프리엠퍼시스 방법의 스펙트럼 추정 결과에서 각 형성음 주파수의 위치가 다르게 나타나며, STED 방법이 기존의 방법에 비해 형성음 주파수의 피크가 더 뚜렷이 나타남을 알 수 있다.

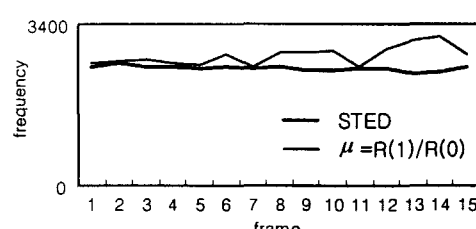
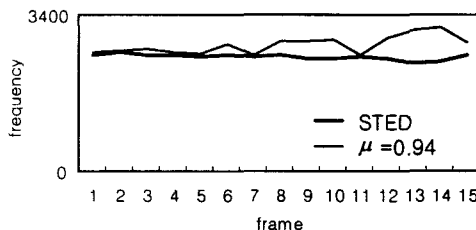
그림 4는 세 가지 프리엠퍼시스 방법에 대한 /에/의 시작 부분부터 정상 상태까지의 형성음 주파수 궤적을 나타낸다. 여기서 형성음 주파수는 스펙트럼 추정 결과로부터 1차와 2차 미분을 이용한 침두치 검출 방법으로 추출된다.



(a) F1 궤적



(b) F2 궤적



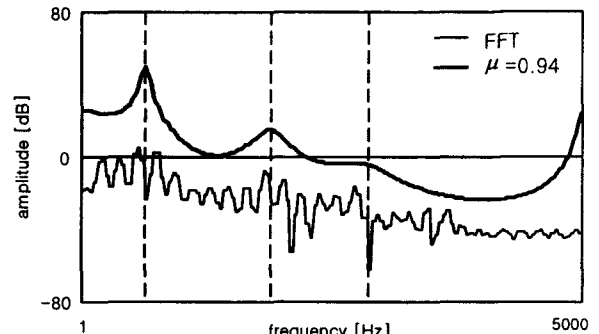
(c) F3 궤적

그림 6. 세 가지 프리엠퍼시스 방법에 의한 /에/의 형성음 주파수 궤적.

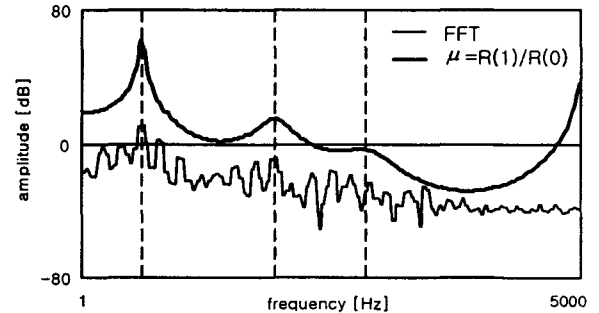
Fig. 6. Formant trajectories using 3 pre-emphasis methods for /에/.

그림 6에서 F1은 STED 방법이나 기존에 널리 쓰이는 두 가지 방법이나 큰 차이가 없지만, F2와 F3가 정상 상태로 갈수록 상당한 차이가 나타남을 알 수 있다. 그림에서 알 수 있듯이, 기존의 방법인  $\mu=0.94$ 와  $\mu=R(1)/R(0)$  방법은 F2와 F3가 전구간에 걸쳐 일정하지 않고 변동이 매우 심하고, 그에 반해 STED 방법에 의해 추출된 F2와 F3는 시작 부분에서 정상 상태까지 거의 일정한 값을 유지함을 알 수 있다.

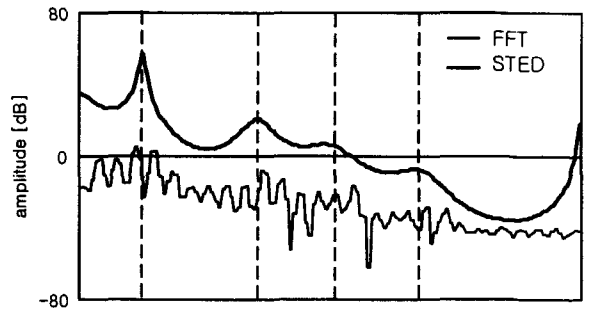
그림 6에서 나타난 형성음 주파수의 큰 차이에 대하여 정확성을 검증할 위하여 세 가지 프리엠퍼시스 방법에 의해 추정된 스펙트럼과 FFT 스펙트럼을 비교·분석하였다. FFT 스펙트럼과의 비교 실험은 그림 6에서 형성음 주파수의 차가 특히 크게 나타나는 프레임에 대하여 수행하였다. 그림 7은 발음 /에/의 8번째, 12번째 프레임에 대한 세 가지 프리엠퍼시스 방법의 결과를 보여준 것이다.



(a)  $\mu=0.94$



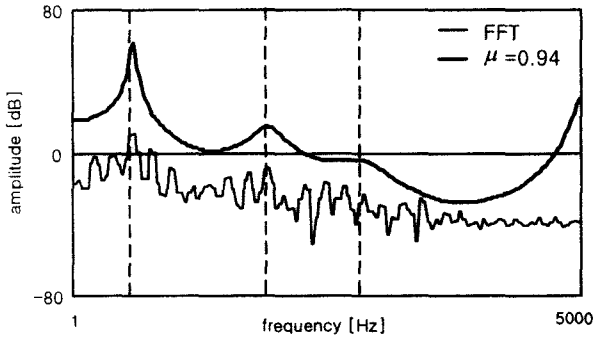
(b)  $\mu=R(1)/R(0)$



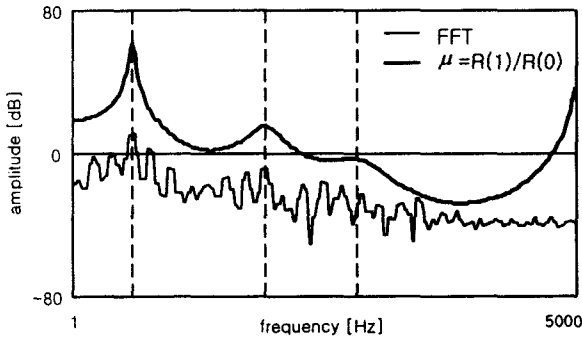
(c) STED

그림 7. /에/에 대한 세 가지 프리엠퍼시스와 FFT 스펙트럼의 비교 (8번째 프레임).

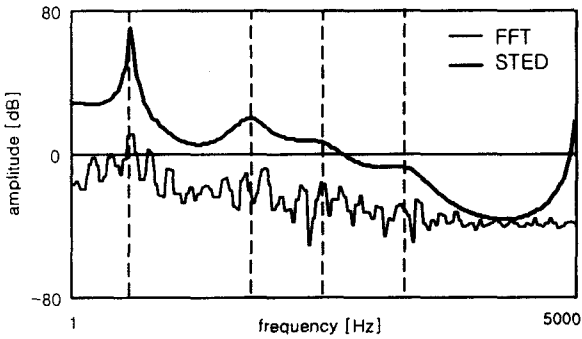
Fig. 7. Comparison of 3 pre-emphasis methods and FFT spectra for /에/ (8th frame)



(a)  $\mu=0.94$



(b)  $\mu=R(1)/R(0)$



(c) STED

그림 8. /에/에 대한 세 가지 프리엠퍼시스와 FFT 스펙트럼의 비교 (12번째 프레임).

Fig. 8. Comparison of 3 pre-emphasis methods and FFT spectra for /에/ (12th frame).

그림 7에서는 STED 방법이 F2, F3의 위치가 약간 벗어나 있지만, 기존에 방법에서는 나타나지 않던 F4가 STED 방법에서는 FFT 스펙트럼과 거의 일치되게 나타남을 알 수 있다. 그림 8에서 기존 방법의 결과에서는 F3와 F4가 중첩되어 하나의 형성음 주파수로 나타남을 볼 수 있고, STED 결과는 FFT 스펙트럼과 모든 형성음 주파수의 위치가 거의 일치하게 나타난다. 이상에서 살펴본 바와 같이, FFT 스펙트럼과 세 가지 프리엠퍼시스 방법을 비교한 결과, 본 연구에서 제안한 STED 프리엠퍼시스 방법이 기존에 널리 쓰이는 두 방법보다 발음의 시작 부분에서부터 정상 상태까지

더 안정되고 정확한 형성음 주파수를 검출할 수 있고, 스펙트럼의 형태도 더 정확하다는 것을 알 수 있다.

본 연구에서 제안한 STED 방법의 계산적 효율성을 확인하기 위해서 각 프레임에서의 연산량을  $\mu=R(1)/R(0)$ 를 이용하는 기존의 최적 방법과 비교하였다. 이 때  $\mu$ 를 0.94의 고정된 값으로 사용하는 경우는 연산이 필요없으므로 제외하였다. 비교 결과, 제안한 프리엠퍼시스 방법이 기존의 최적 방법이라 알려진 방법에 비해 약 50[%]의 연산량이 감소하였다. 표 2는 최적 방법과 STED 방법의 연산량을 비교해 놓은 것이다.

표 2. 최적 방법과 STED 방법에서의  $\mu$  연산량 비교.

Table 2. Comparison of  $\mu$  computations in the optimal and the STED methods.

연산량/frame	곱셈	나눗셈	계
최적 방법	400	1	401
STED	200	1	201

### 5. 결론

본 연구에서는 음성 신호 발생 과정에서 음원 특성과 방사 특성을 효과적으로 제거하기 위하여 음성 신호의 단구간 에너지 차를 이용한 프리엠퍼시스 기법을 제안하였다. 제안한 프리엠퍼시스 방법을 이용하여 한국어 5개 단모음에 대한 스펙트럼 추정 및 형성음 주파수 추출 실험을 수행하였고, 기존에 널리 쓰이는 두 가지 프리엠퍼시스 방법을 이용해서도 동일한 음성 신호 분석을 거쳐 그 결과를 비교·분석함으로써 제안된 방법의 정확성을 확인하고자 하였다. 실험을 통하여 얻어진 결론은 다음과 같다.

첫째, 선형 예측 분석을 통한 스펙트럼 추정 결과, 본 연구에서 제안한 STED 프리엠퍼시스 방법에 의한 결과가 기존에 널리 쓰이는 두 방법에 비해 형성음 주파수의 피크가 더 뚜렷이 나타났으며, 전체적으로 스펙트럼 형태가 개선됨을 알 수 있었다. 둘째, 스펙트럼 추정 결과로부터 형성음 주파수 추출 실험 결과, 기존의 널리 쓰이는 두 방법은 형성음 주파수가 전구간에 걸쳐 일정하지 않고 변동이 매우 심한 반면, 본 연구에서 제안한 방법은 형성음 주파수가 시작 부분에서 정상 상태까지 거의 일정한 값을 유지하여 안정된 형성음 주파수의 획득을 얻을 수 있었다. 셋째, 추출된 형성음 주파수의 정확성을 확인하기 위한 FFT 스펙트럼과 비교 결과, 본 연구에서 제안한 방법에 의한 형성음 주파수의 위치가 기존의 두 방법보다 FFT 스펙트럼과 더 잘 일치하였으며, 기존의 두 방법에서 F3와 F4가 중첩되어 하나의 주파수로 나타나던 형성음 주파수가 STED 방법에서는 FFT 스펙트럼과 거의 일치되게 F3와 F4가 모두 뚜렷이 나타났다. 마지막으로, 기존의 최적 방법과 연산량을 비교해 본 결과, 본 연구에서 제안한 STED 방법이 기존의 최적 방법에 비해 약 50[%]의 연산량이 감소하였다. 이는 제안된 방법이 더 정확하면서도 음성 관련 실시간 시스템에서도 더 유리하다는 것을 알 수 있다.

참 고 문 헌

[1] H. Wakita, "Estimation of the Vocal Tract Shape by Optimal Inverse Filtering and Acoustic/Articulatory Conversion Methods", SCRL Monograph No.9, Speech Communications Research Laboratory, Santa Barbara, California, 1972.

[2] H. Wakita, "Direct Estimation of the Vocal Tract Shape by Inverse Filtering of Acoustic Speech Waveforms", IEEE Trans. Acoust., Speech, Signal Processing, Vol. AU-21, No. 5, Oct. 1973.

[3] S. Furui : *Digital Speech Processing, Synthesis, and Recognition*, 2nd ed., Marcel Dekker, INC., 2001.

[4] J. D. Markel, A. H. Gray : *Linear Prediction of Speech*, Springer-Verlag·Berlin·Heidelberg·New York, 1976.

[5] S. M. Kristensen, M. D. Sorensen, H. Gesmar, and J. J. Led, "Estimation of Signal Intensities in 2D NMR Spectra with Severe Baseline Distortion by Combined Linear Prediction and Least-Square Analysis", J. of Magnetic Resonance, Series B 112, pp. 193-196, 1996.

[6] S. J. Orfanidis : *Optimum Signal Processing : An Introduction, 2nd ed.*, Macmillan Publishing Company, 1988.

[7] L. R. Rabiner, R. W. Schafer : *Digital Processing of Speech Signals*, Prentice-Hall, 1978.

[8] S. Saito, K. Nakata : *Fundamentals of Speech Signal Processing, 2nd ed.*, Academic press, 1985.

[9] A. H. Gray, J. D. Markel, "A Spectral Flatness Measure for Studying the Autocorrelation Method of Linear Prediction of Speech Analysis", IEEE Trans., ASSP-22, pp. 207-217, 1974.

[10] D. O'Shaughnessy : *Speech Communication - Human and Machine*, IEEE Press, 2000.

[11] B. S. Atal, "Speech analysis and synthesis by linear prediction of speech wave", J. Acoust. Soc. Am, Vol. 41, pp. 65(A), 1970.

[12] K. S. Nathan, Y. T. Lee and H. F. Silverman, "A Time Varying Analysis Method for Rapid Transitions in Speech", IEEE Trans. Signal Processing, Vol. 39, No. 4, pp. 815-824, April 1991.

[13] G. Fant : *Acoustic Theory of Speech Production*, Mouton, 1970.

[14] H. Fujisaki, M. Ljungqvist, "Proposal and Evaluation of Models for the Glottal Source Waveform", IEEE Int. Conf. on Acoustics. Speech. and Signal Processing, pp. 1605-1608, 1986.

[15] J. Schroter, J. N. Larar, and M. M. Sondhi, "Speech Parameter Estimation using a Vocal Tract/Cord Model", IEEE Int. Conf. on Acoustics. Speech. and Signal Processing, pp. 308-311, 1987.

[16] E. P. Neuburg, W. R. Bauer, "On the Source-Filter Model of the Vocal Tract", IEEE Int. Conf. on Acoustics. Speech. and Signal Processing, pp. 1609-1612, 1986.

[17] Y. T. Lee, H. F. Silverman, "A Model for Non-stationary Analysis of Speech", IEEE Int. Conf. on Acoustics. Speech. and Signal Processing, pp. 1617-1620, 1986.

[18] H. Fusisaki, M. Ljungqvist, "Estimation of Voice Source and Vocal Tract Parameters Based on ARMA Analysis and a Model for the Glottal Source Waveform", IEEE Int. Conf. on Acoustics. Speech. and Signal Processing, pp. 637-640, 1987.

[19] A. M. de L. Araújo, F. Violaro, "Formant Frequency Estimation Using a MEL Scale LPC Algorithm", IEEE Int. Conf. on Acoustics. Speech. and Signal Processing, pp. 207-212, 1998.

저 자 소 개



김 동 준 (金 東 浚)  
 1963년 4월 14일 생. 1988년 연세대 전기공학과 졸업. 1990년 동 대학원 전기공학과 졸업(석사). 1994년 동 대학원 전기공학과 졸업(공학). 현재 청주대 이공대학 전자·정보통신·반도체공학부 부교수  
 Tel : 043-229-8460  
 E-mail : djkim@chongju.ac.kr



김 주 리 (金 珠 里)  
 1976년 10월 20일 생. 1999년 청주대 정보통신공학과 졸업. 2001년 동 대학원 전자공학과 졸업(석사).  
 Tel : 043-229-8460  
 E-mail : julee@spa.chongju.ac.kr