

## 다중 형태 프레임 분할 방법을 이용한 효율적인 LSP 계수 부호화에 관한 연구

\* 이 윤근, 이 미숙, 김 기철, 이 황수  
한국과학기술원 정보 및 통신공학과

### An Efficient Coding of LSP Parameters Using Multiple Type Frame Segmentation

Yun Keun Lee, Mi Sook Lee, Kee Cheol Lee, Hwang Soo Lee  
Dept. of Information and Communications Eng. KAIST

#### 요약

본 논문에서는 음성 코딩에 있어서 스펙트럼 정보의 효율적인 전송 방법으로써 다중 형태 프레임 분할 방법을 제안하였다. 일반적인 보코더의 경우 엔코더에서 전송된 스펙트럼 정보를 인접 프레임 계수간의 선형보간에 의해 하부 프레임 단위의 스펙트럼 정보를 재생해 낸다. 이때 비선형적 스펙트럼 변화가 존재하는 경우에는 스펙트럼 왜곡이 발생하며 본 논문에서는 이를 intra frame spectral distortion (IFSD)이라 정의한다. 제안된 다중 형태 프레임 분할 방법은 한 프레임내에서 스펙트럼 계수의 변화를 잘 나타낼 수 있도록 차수마다 다른 분할 형태를 적용하여 IFSD를 최소화 시키는 방법이다.

스펙트럼 정보로서 line spectrum pair(LSP) 계수를 이용하였으며, 분할(segment) 갯수 및 분할 형태(type)의 갯수에 따른 IFSD를 분석하여 이를 기존의 방법과 비교하였다. 그 결과 기존의 방법보다 적은 정보량으로 왜곡이 적은 스펙트럼을 재생할 수 있었다.

#### 1. 서론

음성 코딩 방법이 발전함에 따라 저 전송율에서 고품질의 음성을 합성해 낼 수 있는 보코딩 기법이 많이 개발되고 있다[1]. 그런데 전송률 중 많은 부분이 스펙트럼 정보를 전송하는데 사용된다. 스펙트럼 정보는 복원된 음성신호의 음질에 많은 영향을 미치므로, 적은 양의 전송 비트를 이용해 정밀한 스펙트럼 정보를 전송하는 것이 중요하다[3, 4].

일반적으로 기존 보코더의 스펙트럼 분석 방법은 음성 신호를 창(window)을 이용해 프레임으로 나누고 창을 일정한 프레임 간격(interval)만큼 이동시켜 스펙트럼 정보를

추출하는 단시간 분석(short-time analysis) 방법을 주로 사용한다. 이동통신의 표준 보코딩 알고리즘으로 쓰이는 Qualcomm code excited linear prediction(QCELP) 보코더와 vector sum excited linear prediction(VSELP) 보코더 등에서는 20 msec 마다 한번씩 스펙트럼 계수를 추출하며 이때 20 msec의 창길이를 갖는 비중첩창(non-overlapped window)을 이용한다[1]. 또한, 디코더에서는 전송된 스펙트럼 계수를 선형보간하여 5 msec의 하부 프레임마다 스펙트럼 계수를 갱신한다. 이와같이 선형보간에 의해 하부 프레임 단위의 스펙트럼 계수를 생성하는 이유는 음성 신호의 스펙트럼이 단기간(20 msec)내에서는 선형적으로 변한다는 가정에 의해서이다. 그러나 실제로 음성 신호는 음소 환경에 따라 짧은 구간에서도 특성이 비선형적으로 급격히 변하는 경우가 있으며, 이때 선형보간에 의해 하부 프레임 단위의 스펙트럼을 재생하면 스펙트럼 왜곡이 발생하여 합성음질의 저하를 초래한다. 이러한 음질저하를 막기 위해서는 프레임 갱신율(frame update rate)을 증가시켜 매 프레임마다 추출된 스펙트럼 계수를 전송하여야 하는데, 이는 전송하여야 할 정보량이 프레임 갱신율에 비례하여 증가하므로 저 전송율의 보코더에는 부적합하다. 반면에 음성신호의 스펙트럼이 장구간동안 안정적(stationary)인 경우는 20 msec 이상의 간격마다 한번씩 스펙트럼 분석을 하고 이를 선형보간하여 하부 프레임의 스펙트럼을 재생하여도 많은 스펙트럼 왜곡이 발생하지 않는다. 즉 스펙트럼 분석을 위한 프레임 갱신율을 음소환경에 따라 변화시키는 것이 바람직하다[7]. 또한 스펙트럼 분석에 많이 사용하는 linear predictive coding(LPC), line spectrum pair(LSP) 분석 방법의 경우 각 스펙트럼 계수의 시간에 따른 변화 정도가 차수에 따라 일정하지 않는데 이를 동일한 갱신율(update rate) 즉, 20 msec 마다 한번씩 전송하는 것은 비효율적이라 할 수 있다.

본 논문에서는 엔코더에서 전송된 스펙트럼 계수를 디코더에서 선형보간하여 스펙트럼을 재생할때 발생하는 스펙

트럼 왜곡을 줄이며 동시에 스펙트럼 정보의 전송에 필요한 비트수를 감소시킬 수 있는 다중 형태 프레임 분할(multiple type frame segmentation : MTFS) 방법을 제안하였다. 편의상, 기존의 방법처럼 고정된 프레임 갱신율에 따라 스펙트럼 계수를 전송하는 방법을 단일 형태 프레임 분할(single type frame segmentation : STFS) 방법이라 부르기로 한다. 또한 디코더에서 재생한 스펙트럼의 왜곡 정도를 나타내는 척도로써 intra frame spectral distortion(IFSD)을 정의하여 사용한다.

제안된 MTFS 방법에서는 스펙트럼 계수의 차수에 따라 여러 가지 형태로 프레임을 분할하여 각 분할단위로 스펙트럼 계수를 전송한다. 이때 프레임의 분할 형태는 선형보간에 의해 복원된 스펙트럼의 왜곡(IFSD)을 최소화 시키도록 선택된다. 스펙트럼 정보로는 선형보간 특성이 좋은 LSP 계수를 이용하였다. 녹음된 남,녀 음성 시료를 이용하여 실험한 결과 기존의 방법에 비해 적은 정보량으로 IFSD를 줄일 수 있었다.

제 2장에서는 IFSD를 정의하고 실험에 의해 이를 분석하였으며, 제 3장에서는 제안된 MTFS 방법에 대해 설명하였다. 제 4장에서는 제안된 방법에 의한 시뮬레이션 결과를 기존의 방법과 비교 평가하였으며 제 5장에서는 결론을 맺었다.

## 2. 프레임 내의 스펙트럼 왜곡 (IFSD)

보코더에서는 음성 스펙트럼을 분석하기 위하여 음성 신호를 창(window)을 이용해 프레임으로 나누고 창을 일정한 프레임 간격(interval)만큼 이동시키며 스펙트럼 정보를 추출하는 단시간 분석(short-time analysis) 방법을 주로 사용한다. 보통 20 ~ 25 msec 마다 한연속 스펙트럼 분석을 하여 추출된 계수를 엔코더에서 전송하면, 디코더에서는 이를 선형보간하여 하부 프레임 단위의 스펙트럼 계수를 생성한다. 그러나 스펙트럼 특성이 비선형적으로 급격히 변화하는 음소환경에서는 선형보간에 의해 하부 프레임 단위의 스펙트럼을 재생하면 실제 음성 스펙트럼과 다른 스펙트럼이 생성되며 이를 IFSD이라 정의한다. 그러나 실제 음성신호의 스펙트럼을 정확히 추출해내는 것은 불가능하므로 본 논문에서는 스펙트럼 분석 프레임 갱신율을 충분히 높여 추출한 스펙트럼 정보로써 실제의 음성 스펙트럼을 근사시켰다. 편의상 이를 기준 스펙트럼이라고 하자. 프레임 갱신율을 높이면 낮은 프레임 갱신율에 의해 추출된 스펙트럼 정보의 선형보간으로 생성한 스펙트럼보다 실제 음성의 비 선형적 스펙트럼 변화를 정확히 나타낼 수 있다.

음성 시료를 이용해 기존의 STFS 방법에 의한 IFSD를

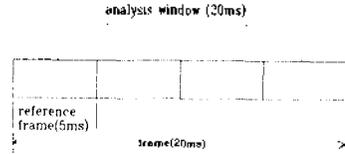


그림 1. 프레임의 구성

측정해 보았다. STFS 방법은 1장에서 설명한 바와 같이 20 msec의 프레임 간격을 갖고 20 msec의 비중첩창을 이용해 LSP 계수를 추출하여 전송하고 이를 디코더에서 선형보간하여 5 msec의 하부 프레임 스펙트럼을 재생하는 방법이다. 기준 스펙트럼 추출을 위하여 그림 1과 같이 기준 프레임 간격을 하부 프레임 길이와 동일하게 5 msec로 하고 창길이 20 msec의 중첩창(overlapped window)을 이용하여 스펙트럼 분석을 하였다. 이는 1장에서 설명한 바와 같이 기존의 표준 보코더에서 많이 사용하는 창길이인 20 msec로 통일시킴으로써 기존의 방법과의 성능 비교를 용이하게 하기 위함이다. 기준 프레임 간격도 5 msec보다 짧게 하여도 무방하나 실험 결과에는 큰 영향을 미치지 못하였다.

먼저 각 기준 프레임 간격 5 msec마다 LSP 계수 ( $LSP_{5ms}$ )를 구한다. 또한 20 msec의 프레임 간격으로 LSP 계수 ( $LSP_{20ms}$ )를 구한 후 이를 식 (1)과 같이 인접 프레임 간 1차 선형보간을 하여 기준 프레임 간격에서의 LSP 추정 계수 ( $\widehat{LSP}_{5ms}$ )를 계산한다.

$$\widehat{LSP}_{5ms}[i \times 4 + k] = LSP_{20ms}[i] + slope \times k \quad (1)$$

$$slope = (LSP_{20ms}[i+1] - LSP_{20ms}[i]) / 4$$

$$0 \leq i < N, \quad 0 \leq k < 4$$

여기서  $i$ 는 프레임 인덱스이며  $N$ 은 전체 프레임의 갯수,  $k$ 는 프레임내의 기준 프레임 인덱스이다. IFSD은 한 프레임 구간내에 존재하는 스펙트럼 왜곡으로써, 여기서는  $LSP_{5ms}$ 에 의한 스펙트럼과  $\widehat{LSP}_{5ms}$ 에 의한 스펙트럼의 스펙트럼 거리(spectral distance)로 정의되며 식 (2)와 같이 계산할 수 있다[2, 5].

$$IFSD^2 = \sum_{\omega=0}^{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi} (10 \log \frac{S_{5ms}[\omega, i]}{S_{5ms}[\omega, i]})^2 d\omega \quad (2)$$

여기서  $S_{5ms}[\omega, i]$ 와  $\widehat{S}_{5ms}[\omega, i]$ 는 각각  $i$ 번째 기준

프레임의  $LSP_{3ms}$  및  $LSP_{3ms}$ 에 의한 기준 프레임 단위의 스펙트럼이다. 약 2초 정도의 길이를 갖는 서로 다른 256개의 단문을 남자 8명 여자 8명이 발음하여 얻은 음성 시료를 이용하여 위의 실험을 한 결과 평균 IFSD은 1.50 dB로 나타났다.

### 3. 다중 형태 프레임 분할 (MTFS) 방법

음성 신호의 스펙트럼 정보를 LSP 계수로 나타낼 경우 보통 10 차의 LSP 계수를 이용한다. 장구간 분석을 할 경우 LSP 계수는 한 프레임 구간내에서도 어느 정도 변화할 수 있으며 그 변화 정도는 차수에 따라 다르고 프레임 구간내의 위치에 따라서도 다르다. 그러나 일반적으로 음성 코딩에서는 식 (1)과 같이 각 프레임의 LSP 계수를 인접 프레임과 선행보간 하여 각 하부 프레임의 LSP 계수로 이용한다. 이는 분석구간 내에서 모든 차수의 계수가 동일하게 선행적으로 변화한다고 가정함을 의미한다. 그러나 실제로는 LSP 계수가 비선형적으로 변하는 구간이 존재하며 또한 각 차수마다 변화의 형태 및 정도가 다르다. 이로 인해 제 2장에서 살펴본 바와 같이 IFSD이 존재한다.

따라서 본 논문에서는 다중 형태 프레임 분할(MTFS)에 의한 스펙트럼 분석 방법을 제안 하였다. MTFS 방법에는 기존의 STFS 방법의 프레임 구간보다 긴 프레임 구간이 필요하다. 이를 편의상 상부 프레임 구간  $L$ 이라 한다. 상부 프레임 구간은 여러 개의 하부 프레임으로 나뉘는데 이때 하부 프레임 간격을  $L_s$ 라 하고  $L$ 은  $L_s$ 의 정수배라 하자.

먼저 각 하부 프레임마다 LSP 계수를 추출한다. 추출된 계수를 모두 전송하면 전송 비트가 증가하므로 계수를 선택적으로 전송할 필요가 있다. 이 경우, 전송된 계수를 이용하여 디코더에서 선행보간에 의해 각 하부 프레임 단위의 스펙트럼을 복원할 때 IFSD이 작아지도록 전송 계수를 선택하여야 한다. 즉, 각 차수의 LSP 계수 중 장구간 프레임 내에서 비선형적으로 변하여 왜곡이 발생하는 경우는 상부 프레임 구간을 여러 개의 프레임으로 분할하여 각 분할의 LSP 계수를 전송하여야 하며, 선형적으로 변하는 계수의 경우는 적은 수의 프레임으로 분할하여도 왜곡이 별로 발생하지 않는다. 이는 LSP 계수의 각 차수마다 독립적으로 분할되며  $i$  차수의 분할 개수를  $N_i$ 라 하면 전체 분할 개수

$N_{total}$ 은 식 (3)과 같다.

$$N_{total} = \sum_{i=1}^{order} N_i \quad (3)$$

모든 하부 프레임 경계를 분할할 경우  $N_{total}$ 은 하부 프레임의 개수에 차수를 곱한 값이 된다. 상부 프레임 구간내에  $N_{total}$ 이 너무 크면 전송해야 할 정보량이 많아져서 실용성이 없으므로 적당한 값으로 제한할 필요가 있다.  $N_{total}$ 을  $M$ 이라 고정하고  $N_i$ 의 최대치를  $M_i$ 라 하자. 여기서  $M_i$ 는 상부 프레임 구간내의 하부 프레임의 개수,  $L/L_s$ 와 같다. 이 경우 상부 프레임 구간을 분할하는 방법의 전체 가지수  $C$ 는 식 (4)와 같다.

$$C = \binom{(M_i-1) \times \rho}{M-\rho} \quad (4)$$

$$= \frac{((M_i-1) \times \rho)!}{(M-\rho)! \times ((M_i-1) \times \rho - (M-\rho))!}$$

각 상부 프레임 구간은  $C$  개의 분할 형태중 선행보간에 의해 발생하는 IFSD을 최소화 하는 형태로 분할된다. 각 분할에는 한 개 이상의 하부 프레임이 포함되며 분할내의 첫번째 하부 프레임의 스펙트럼 정보가 전송되며 나머지 하부 프레임의 스펙트럼 계수는 선행보간에 의해 재생된다. 즉 먼저 하부 프레임 간격마다 구해진 LSP 계수중 주어진 분할의 형태에 따라 각 분할의 첫번째 하부 프레임 계수를 이용해 기준 프레임 단위의 LSP 계수를 선행보간에 의해 구한다. 이와 같이 생성된 스펙트럼과 기준 프레임 간격에서 추출한 기준 스펙트럼과의 거리(spectral distance), 즉 IFSD이 식 (2)에 의해 계산되어 가장 작은 스펙트럼 거리를 갖는 분할 형태를 선택한다. 이는 식 (5)와 같다. 기준이 되는 분석 간격  $L_s$ 은 2장에서 설명한 바와 같이 충분히 작은 값으로 선택되며 하부 프레임 간격보다 작거나 같다. 그림 2는 분할 개수( $N_{total}$ )가 9인 경우의 MTFS 방법에 의한 프레임 분할의 예이다.

$$seg\ type_{optimal} = \arg \min_{seg\ type} \sum_{i=0}^{L/L_s} SD_{seg\ type}[i]^2 \quad (5)$$

$SD$ : 스펙트럼 거리(spectral distance)

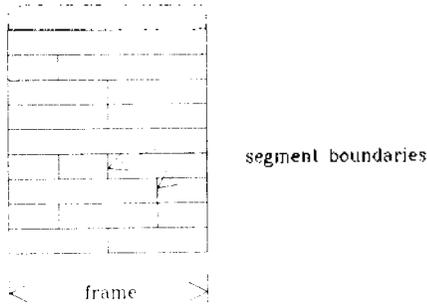


그림 2. 프레임 분할의 예

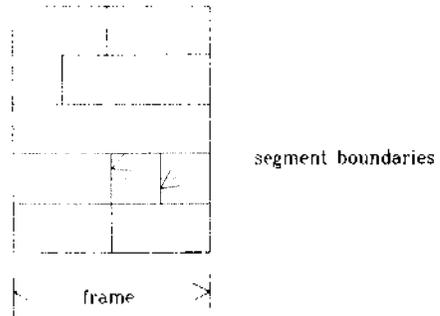


그림 3. 변형된 프레임 분할의 예

#### 4. 실험

본 실험에서는 10차 LSP 계수를 이용하여 MTFB 방법에 의해 스펙트럼 분석을 하였으며 이때의 IFSD를 기존의 방법인 STFS의 경우와 비교하였다. LSP 계수를 선택한 이유는 선형보간 특성이 좋기 때문이다. 상부 프레임 구간  $L$ 은 음성 신호의 안정구간이 충분히 유지될 수 있을 정도의 긴 구간으로 선택되어야 하는데, 너무 길면 분할 갯수가 증가하고 분할 형태의 가지수가 급격히 증가하여 이에 따라 검색 시간이 증가하므로 여기서는 40 msec로 하였으며, 하부 프레임 간격  $L_s$ 는 10 msec로 하였다. 또한 기준 프레임 간격  $L_r$ 은 2장에서와 마찬가지로 5 msec로 하였다. 모든 분석 창 의 길이는 20 msec로 하였으며 분석 간격만큼 이동하며 증결하여 사용하였다. 실험에 사용한 음성 시료는 제 2장에서 설명한 바와 같이 약 2초 정도의 길이를 갖는 서로 다른 256 개의 단문을 남자 8명, 여자 8명이 발음하여 무향실에서 녹음하였다.

분할의 갯수와 분할 형태의 가지수를 변화시키며 IFSD의 변화를 살펴보았으며, 이를 기존의 STFS 방법에 의한 IFSD와 비교하였다. 만약 각 계수의 분할 위치를 창구간 프레임의 중앙으로 고정시키고 분할갯수( $N_{tot}$ )를 20으로 하면 제 2장에서 설명한 20 msec의 고정 프레임 구간을 갖는 STFS의 경우와 동일해진다. 이는 기존의 많은 보고서에서 채택한 사항이므로 본 실험에서는 이를 제외한 방법과 비교하였다.

먼저 분할 갯수를 변화시키며 성능 분석을 하였다. 분할 갯수가 20인 경우 한 프레임을 분할하는 방법의 검색 가지수  $C$ 는 식 (4)에 의해 다음과 같다.

$$C = \binom{(4-1) \times 10}{20-10} = 30045015$$

이를 음성 코딩에 적용할 경우 분할 형태에 관한 정보를 전송하여야 하는데 여기에 25 비트가 필요하다. 따라서 이는 실용적이지 않다. 그런데 LSP 계수 중 서로 인접한 차수의 변화 패턴은 유사할 것이므로 인접한 두 개의 계수를 하나의 그룹으로 처리하여 하나의 분할 패턴을 할당한다. 그림 3은 하나의 예를 보여준다. 이 경우 분할 방법의 가지수는 다음과 같다.

$$C = \binom{(4-1) \times 5}{10-5} = 3003$$

분할 갯수에 따라 전송하여야 할 정보량이 변하므로 분할 갯수를 20 개에서 12 개까지 변화시키며 이때 발생하는 IFSD를 측정하였다. 여기서 IFSD는 식 (2)와 같이 MTFB 방법에 의해 복원된 스펙트럼과 기준 프레임 간격 단위로 분석한 스펙트럼과의 스펙트럼 거리로 주어진다. 결과는 그림 4에 나타나 있다.

그림 4에서 보면 분할 갯수가 16 개 이상이면 STFS 방법 보다 좋은 성능을 나타낸다. STFS 방법의 경우 20개의 계수를 전송하여야 하는 반면, 분할 갯수가 16 개인 MTFB 방법의 경우 16 개의 계수만 전송하면 된다. 즉 적은 양의 계수로 왜곡이 더 적은 스펙트럼을 재생할 수 있음을 알 수 있다. 그러나 분할 갯수가 16 개인 경우 가능한 분할 형태의 가지 수가 455 개이므로 9 비트의 부가 정보의 전송이 필요하다. 그런데 분할 형태가 발생하는 빈도를 조사하면 분할 형태에 따라 발생 빈도의 차이가 심한 것을 알 수 있다. 그러므로 발생 빈도가 적은 분할 형태는 부가 전송 부담을 줄이며 검색 시간의 단축을 위해 제외시키는 것

차 검색 등이 포함된 전체 알고리즘의 계산량에 비해 매우 적으므로 구현에 문제가 되지 않는다.

### 5. 결론

음성 신호의 스펙트럼 정보를 추출함에 있어서 효율적인 프레임 분할에 의해 적은 정보량으로 IFSD를 최소화 하기 위해 MTFB 방법을 제안하였다. 하나의 상부 프레임 구간 내에서 LSP 계수의 변화를 잘 나타낼 수 있도록 차수마다 다른 분할 형태를 적용하여 프레임 내의 스펙트럼 왜곡을 최소화시켰다.

남녀 화자가 발성한 256 개의 단문으로 구성된 음성 시료를 이용하여 제안된 알고리즘의 성능 평가를 하였다. 40 msec의 상부 프레임 구간, 10 msec의 하부 프레임 간격을 갖는 경우에 대해 분할 갯수 및 분할 형태의 가지수를 변화시켜 가며 IFSD를 측정하였다. 그 결과 16 개의 분할을 가지며 32 가지의 분할 형태를 갖는 MTFB 방법의 경우 기존의 STFS 방법에 비해 더 작은 IFSD를 발생시킴을 알 수 있었다. 이 경우 4 개의 계수에 해당하는 전송 비트를 절약할 수 있다. 부가 정보로서 분할 형태를 지정하는 5 비트의 부가 전송 정보가 필요하나 이는 절약되는 비트에 비해 적은 양이다. 즉 MTFB 방법을 택할 경우 기존의 STFS 방법에 비해 적은 전송 비트를 이용하여 왜곡이 더 적은 스펙트럼을 재생할 수 있음을 알 수 있다.

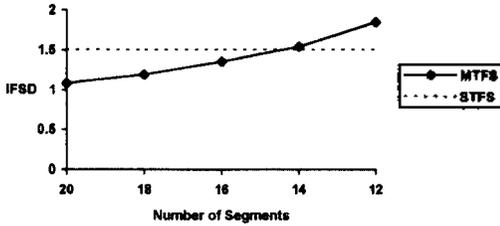


그림 4. 분할 갯수에 따른 평균 IFSD의 변화

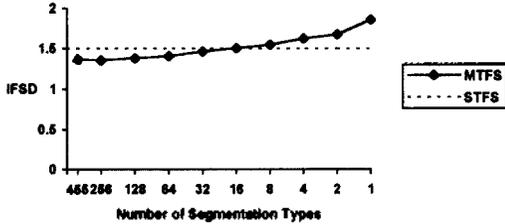


그림 5. 분할 형태의 갯수에 따른 평균 IFSD의 변화  
(분할 갯수는 16으로 고정)

이 유리하다.

적당한 갯수의 분할 형태를 선택하기 위해 다음과 같은 실험을 하였다. 먼저 456 개의 가능한 모든 분할 형태를 이용하여 음성 시료를 분석한다. 이때 분할 형태의 발생 빈도를 측정하여 발생 빈도가 높은 256 개의 분할 형태만을 선택한다. 이와 같이 하면 1 비트의 부가 정보를 절약할 수 있다. 다음은 256 개의 분할 형태만을 이용하여 같은 방법으로 128 개의 분할 형태만을 선택한다. 이와같은 방법으로 분할 형태의 가지수를 줄여가면서 IFSD를 계산하여 그 결과를 그림 5에 나타냈다. 결과에 의하면 32 가지의 분할 형태로 기존의 STFS 방법보다 우수한 성능을 나타냄을 볼 수 있다.

결과적으로, 제안된 MTFB 방법에 의해 스펙트럼 분석을 한 결과 기존의 STFS 방법에 비해 4 개의 전송 계수를 감소시키면서도 왜곡이 더 적은 스펙트럼을 재생할 수 있으며 이에 따른 부가정보는 5 비트로써 전송 계수의 감소에 따른 비트 절약에 비하여 적은 양임을 알 수 있다. MTFB 방법에 의해 기존의 STFS 방법보다 스펙트럼 분석시의 계산량이 2 배 이상 증가하는데 일반적으로 보코딩 알고리즘에서 스펙트럼 분석에 소요되는 계산량은 코드북 검색, 피

### 【 참고문헌 】

- [1] B.S.Atal, V.Cuperman and A.Gersho, Advances in Speech Coding, Kluwer Academic Publishers, 1991.
- [2] P.E.Papamichalis, Practical Approaches to Speech Coding, Prentice-Hall, Inc. 1987.
- [3] K.K.Paliwal and B.S.Atal, "Efficient Vector Quantization of LPC Parameters at 24 bits/frame," Proc. ICASSP'91, pp. 661-664, May, 1991.
- [4] F.K.Soong and B.H.Juang, "Optimal Quantization of LSP Parameters," IEEE Tran. on Speech and Audio Processing, pp. 15-24, Jan. 1993.
- [5] E.Erzin and A.E.Cetin, "Interframe Differential Vector Coding of Line Spectrum Frequencies," Proc. ICASSP'93, vol. 2, pp. II25-II28, April, 1993.
- [6] J.M.Lopez and N.Farvardin, "A Combined Quantization-Interpolation Scheme for Very Low Bit Rate Coding of Speech LSP Parameters," Proc. ICASSP'93, vol. 2, pp. II21-II24, April, 1993.
- [7] T. Svendsen, "Segmental Quantization of Speech Spectral Information," Proc. ICASSP'94, vol. 1, pp. 1517-1520, April, 1994.
- [8] P.A.Chou and T.Lookabaugh, "Variable Dimension Vector Quantization of Linear Predictive Coefficients of Speech," Proc. ICASSP'94, vol. 1, pp. 1517-1520, April, 1994.