

論 文

중간 전송율에서 적응 포스트 필터링을 이용한 음성용 SBC의 성능 향상

正會員 金 元 九* 正會員 李 南 杰** 正會員 尹 大 熙* 正會員 車 日 煥*

Performance Enhancement of SBC for Voice Signal Using Adaptive Postfiltering at the Medium Bit Rate

Weon Goo KIM*, Nam Geol LEE**, Dae Hee YOUN*, Il Whan CHA* *Regular Members*

要 約

본 논문에서는 12 kbit/s와 16 kbit/s의 중간 전송율을 갖는 음성용 SBC(Sub Band Coding)의 성능을 향상시키기 위해 세가지 방법을 수행하였고, 인간의 청각 특성을 이용하는 적응 포스트 필터링(adaptive postfiltering)을 수신단의 마지막에서 수행하였다.

첫째는, SBC 시스템의 대역 필터를 QMF(Quadrature Mirror Filter) 대신 GQMF(Generalized QMF)를 사용하여 성능을 향상시켰고, 둘째는, 각 대역에 적응 비트 할당을 함으로써 음질의 향상뿐 아니라 variable rate 부호화할 수도 있었다. 세번째로는, APCM(Adaptive Pulse Code Modulation)과 ADPCM(Adaptive Differential Pulse Code Modulation)을 부호화기로 사용하여 각각의 성능을 평가할 결과, SB-APCM의 성능이 우수하였다.

또한, 수신단의 마지막에서 적응 포스트 필터링을 수행하여 부호화된 음성의 음질을 개선할 수 있었다. 본 논문에서는 두가지의 적응 포스트 필터링 기법을 제안하였는데 낮은 복잡성을 가지고도 부호화된 음성에서 상당량의 잡음 감쇄를 이룰 수 있었다.

ABSTRACT

In this paper, three methods are studied to enhance the performance of SBC(Sub-Band Coding) schemes for voice signal at the medium bit rate between 12 kbps and 16 kbps, and adaptive postfiltering using human auditory characteristics is done at the decoder output.

First, GQMF(Generalized Quadrature Mirror Filter) is used instead of QMF(Quadrature Mirror Filter) to have better performance. Second, by adaptive bit allocation to each sub-band, speech quality is enhanced and variable rate coding is possible. Third, comparison study of the coder performance using APCM(Adaptive Pulse Code Modulation) and ADPCM(Adaptive Differential Pulse Code Modulation), indicates that SB-APCM performs better than the other.

Adaptive postfiltering at the decoder output enhances the quality of the coded speech. The two

*延世大學校 電子工學科
Dept. of Electronics Eng. Yonsei University.

**金星中央研究所
論文番號：92-13(接受1991. 9. 13)

proposed postfiltering methods decrease the noise sufficiently at the expense of the low computational load.

I. 서 론

Subband 부호화(Sub-Band Coding)는 신호를 여러 대역으로 분할해서 각 대역을 독립적으로 부호화하는 파형 부호화(waveform coding)방식으로 다음과 같은 장점을 가지고 있다. 첫째, 보통의 양자화 오차(quantization error)는 백색 가우시안 분포(white gaussian distribution)를 갖는데, Subband 부호화에서는 이 양자화 오차가 각각의 대역에 한정됨으로써, 주관적인 잡음 차폐 효과(subjective noise masking effect)로써 이용할 수 있다. 둘째, 인간의 청각, 혹은 시각 인지 특성에 따라 각 대역에 효율적으로 비트를 할당 할 수 있다. 셋째, 에너지가 큰 대역에 많은 비트를 할당함으로써 데이터 압축(data compression)을 할 수 있다.

Subband 부호화기는 Crochiere 등에 의해 처음으로 제안되었다¹⁾. 초기에는 각 대역에 고정된 수의 비트를 할당했으나, 1982년에 Ramstad가 각 대역에 할당되는 비트수를 가변시키는 방법을 도입했으며, 최근에는 Honda, Itakura와 Soong, Cox, Jayant 등이 시간과 주파수 영역 모두에서 효율적으로 비트를 가변시키 할당하는 방법을 제안하였다²⁾. 또, 분할된 각 대역간에는 상관 관계가 작으므로 VQ(Vector Quantization)을 이용한 SBC도 제안되었다³⁾.

본 논문에서는 12/16kbps에서 Subband 부호화의 음질이 toll quality에 가까운 음질을 갖도록 하기 위하여 수신단의 마지막에 적응 포스트필터(adaptive postfilter)를 덧붙이는 방법을 수행하였다. Jayant에 의해 처음으로 제안된 적응 포스트 필터링(adaptive postfiltering)의 기본적인 원리는 인간의 청각 특성을 이용한 잡음 차폐 효과(noise masking effect)이다. 즉 noise shaping filter를 써서 백색 잡음인 양자화 오차의 분포를 음성의 스펙트럼과 비슷하게 분포시키 주관적이므로 청각적인 음질의 개선을 가져온다⁴⁾.

낮은 전송율에서는 부호화 잡음의 상닝량이 잡음을 차폐(mask)한 강한 레조넌스(resonance)가 없는 스펙트럼의 밸리(spectral valley) 영역으로 부터 온다. 포스트 필터는 부호화 잡음을 덜 인식하도록 하기 위해 스펙트럼의 밸리 영역에서 잡음 성분들을 일

하시킨다. 그러나, 이러한 필터링 동작은 필연적으로 음성 스펙트럼의 형태에 약간의 왜곡을 가져온다. 인간의 귀는 스펙트럼의 밸리영역에서의 왜곡에 대해서는 아주 민감하지는 않다⁵⁾. 그러므로, 적응 포스트 필터링은 단지 인식된 음성에서 매우 작은 왜곡만 일으키고, 인식된 잡음 레벨을 크게 줄인다.

본 논문에서 제안하는 적응 포스트 필터링은 필터된 음성에서 왜곡이 거의없이 인식된 부호화 잡음을 크게 줄여서 음질의 개선을 가져오도록 했다. 이 적응 포스트 필터링은 추가적 방법으로 수행하였다. 본 논문에서는 12kbps와 16kbps의 전송율에서 적응 포스트 필터를 이용하여 SB APCM(Sub Band Adaptive Pulse Code Modulation)과 SB ADPCM(Sub Band Adaptive Differential Pulse Code Modulation)의 음질을 개선하고, 이 개선된 음질을 평가하였다.

II. 대역 분할 및 합성 필터뱅크

일반적으로 SBC 시스템에서 대역 분할 및 합성시에 사용되는 필터뱅크는 다음과 같다.

QMF는 에일리어싱(aliasing) 및 위상 왜곡(phase distortion)이 없지만 크기 왜곡(magnitude distortion)이 있고, 나무(tree) 구조로 할 경우 지연(delay)이 커지고, 한 대역의 에러가 여러 대역에 퍼질된다.

PR QMF는 왜곡이 전혀 없지만, factorization 방법으로 설계할 경우 계수의 양자화시에 오차가 생기고, 무손실의 블록(lossless building block)으로 할 경우 계수의 양자화에는 영향을 없다⁶⁾. 두 방법 모두 대역 분할된 신호의 양자화시에는 완전 재생 특성이 없어진다.

GQMF는 나무 구조 QMF에 비해 지연이 작고, 한 대역의 에러는 인접 대역에만 영향을 끼친다. GQMF도 역시 왜곡을 있다.

PR GQMF는 PR QMF의 일반화된 형태로 왜곡은 없지만, 분석 필터가 선형 제한조건을 가질 때는 동래 대역 및 저역 대역의 특성이 나빠져서 필터의 댐 수를 증가시켜야 한다. PR-QMF처럼 대역 분할 신호의 양자화에 의해 완전 재생 특성이 없어지게 된

다. 또, 다른 QMF에 비해 설계 과정이 비교적 복잡한 편이다.

본 논문에서는 지연과 에러의 파급 효과를 감소시키고 임의의 수로 대역을 분할할 수 있는 GQMF를 대역 필터로 선택하였다. GQMF 디자인은 Hooke & Jeeves 알고리즘으로 Prototype LPF(low pass filter)를 설계했다⁵⁾.

Ⅲ. 부호화 알고리즘

이 장에서는 부호화기 부분을 APCM을 쓰는 경우와 ADPCM을 쓰는 경우로 나누어 본다. 또한, 각 밴드에 효율적인 비트 할당을 하기 위하여 적응 비트 할당(adaptive bit allocation) 방식을 사용한다.

1. SB-APCM(Sub Band Adaptive Pulse Code Modulation)

SB-APCM은 GQMF 대역 필터를 사용하고 부호화기로 APCM을 쓰는 방식을 가르킨다. SB-APCM의 대략적인 블록도는 그림 1에 있다.

위의 SB-APCM 시스템은 500 Hz의 대역폭을 갖는 8개의 대역으로 나누고 에너지 분포가 상대적으로 적은 7 번째와 8 번째 대역(3-4 kHz)은 전송하지 않는다. 또한, 각 프레임(frame)의 길이는 16 ms으로써 입력신호가 8 kHz 표본화율을 가지므로 128 samples/frame를 갖는다.

APCM 부호화기는 단위 분산(unit variance)를 갖도록 각 대역마다 분산을 예측한 후 표준 편차(standard deviation)의 값으로 정규화를 하며, 양자화기로는 Lloyd-Max Quantizer를 사용하여 정규화된 대역 신호를 1-5bit로 양자화 한다. 이 때, 각 대역

의 비트 수는 분산의 예측을 통한 적응 비트 할당(adaptive bit-allocation) 알고리즘을 이용하여 할당한다⁶⁾.

수신단에서의 역양자화시에도 비트 할당이 필요한데 이것을 위해 우리는 각 대역의 표준 편차를 송신단에서 부가 정보로 수신단에 보내는데, 이 때 μ -law PCM으로 양자화해서 보낸다. 또한 송신단과 수신단의 비트 할당에 필요한 표준 편차의 양자화 오차를 줄이기 위하여 역양자화 과정을 송신단에서 수행한다. 다음 비트 할당을 수행한다. 또한, 적응 비트 할당을 했을 경우에 embeded 양자화기를 사용한다면 수신단에서 부가 정보로 전송되어 온 표준 편차를 이용하여 전체 할당 되는 비트수를 조절하여 전송율을 변화시킬수도 있으므로 variable rate 부호화도 할 수가 있고, 주어진 비트 흐름에서 각 비트들의 중요도가 책정될 수 있으므로 embeded 부호화에도 이용할 수가 있다⁷⁾.

2. SB-ADPCM(Sub Band Adaptive Differential Pulse Code Modulation)

앞에서 제시한 SB-APCM과 거의 동일하나 부호화기 부분에서 APCM 대신 ADPCM을 사용하여 양자화시에 dynamic range를 줄여 더 높은 성능을 얻고자 한다.

SB-ADPCM의 블록도는 그림 2에 있다.

이 때, SB-ADPCM과 SB-APCM과의 차이는 MAX-Quantizer 대신 ADPCM을 쓴다는데 있다. 여기에서 사용한 ADPCM은 CCITT⁸⁾에 권고된 것을 사용하였다.

ADPCM은 다음의 두가지 조건을 만족할 때 정확히 동작한다. 첫째는 정확한 예측을 통하여 작은 값

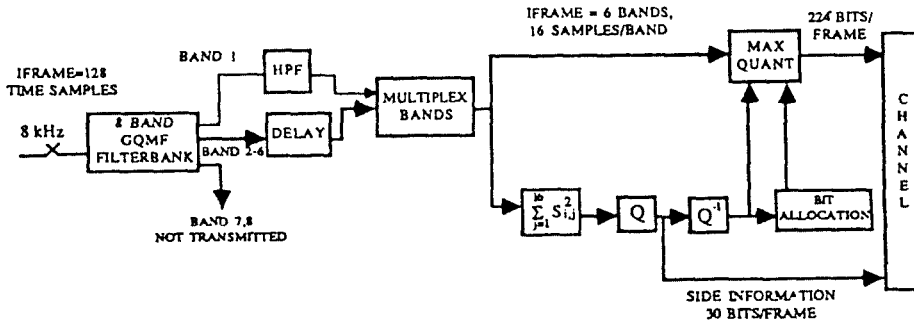


그림 1. SB-APCM의 블록도
Fig.1. Block diagram of SB-APCM.

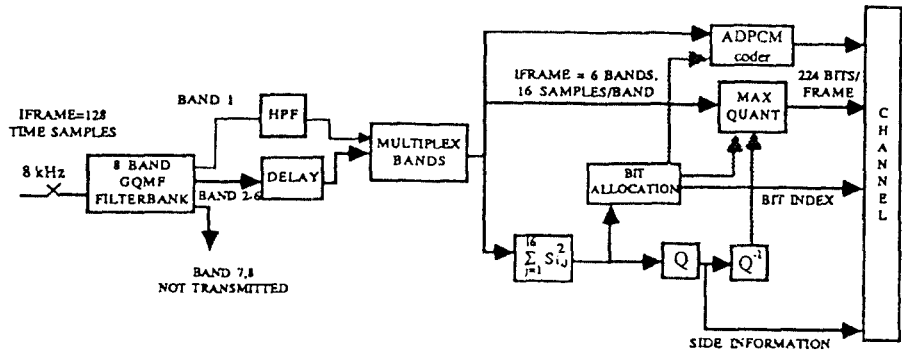


그림 2. SB-ADPCM의 블록도
Fig.2. Block diagram of SB ADPCM.

의 예측 오차를 가지므로써 전송을 감소시킬 수 있고, 둘째는 조절(adjustment)을 통하여 전송 에러가 있을 때에는 알고리즘이 송신단과 수신단을 잘 동기시켜서 좋은 재생 신호를 얻을 수 있다.

첫번째의 정확한 예측은 일반적인 LMS(Least Mean Square) 알고리즘을 쓰면 얻을 수 있고, 두번째의 조절은 예측기의 폴(pole)들이 단위원(unit circle)안에 있으므로 이루어지는 선형 안정도(uniform stability)에 의해 이루어진다.

위의 두 조건은 서로 상반되므로 적절히 조절해야 한다. CCITT G.721 알고리즘은 선형 안정도를 보장하기 위해 송신단의 폴들에 어떤 제한 조건(constraint)을 주어 불안정 상태를 피하도록 했다.

본 논문에서는 선형 안정도를 가지고 있고, 정확한 예측을 할 수 있도록 설계된 CCITT G.721의 ADPCM 알고리즘을 적용하여 SB ADPCM을 수행하였다.

1)후방 적응 예측기(Backward adaptive predictor)

본 논문에서 사용된 ADPCM의 예측기는 그림 3과 같이 2개의 폴과 6개의 제로들을 갖는 ARMA 모델로 구성된다. 후방 적응 예측기는 전방(forward) 적응 예측기에 비해 성능이 감소하지만 예측기의 정보를 전송하지 않아도 되므로 적응 양자화기의 성능이 우수한 경우에는 비트 효율면에서 장점이 있다. 더우기 순차(sequential) 후방 적응 예측기는 계산량이 많지 않도록 구현할 수 있다⁸⁾.

본 논문에서 적용된 알고리즘은 gradient 알고리즘으로써 입력 데이터에 대한 sign 함수만을 이용하여

gradient를 예측한 후 각 계수들을 조정하며, leakage factor와 안정도 조건(stability constraint)을 부여하여 선형 안정성을 보장하고, 전송상의 오차에 따른 예측기의 mistracking 현상을 방지하였다⁸⁾.

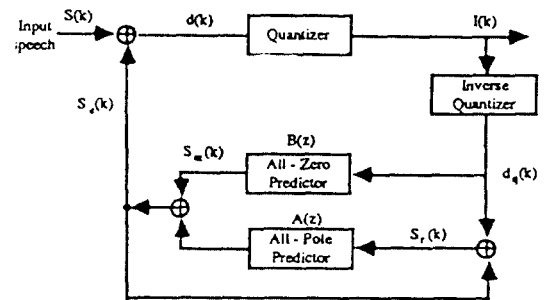


그림 3. 후방 적응 예측기의 블록도
Fig.3. Block diagram of backward adaptive predictor.

2)적응 양자화기(Adaptive quantizer)

본 논문에서는 그림 4와 같은 적응 양자화기를 사용하였다. SBC에서는 대역 분할 필터를 통과한 각 대역의 신호를 ADPCM을 통해 부호화할 때 수퍼수대역마다 적응 비트 할당에 따라 양자화 레벨을 다르게 하므로, 각각의 구간에서 각 대역의 신호를 효율적으로 양자화하기 위하여 다중 양자화(multi-quantization)기법을 적용하였던⁹⁾. 본 논문에서는 CCITT G.721의 양자화기 적용 방식을 적용하였는데, 예측 오차의 변화가 심한 음성에 대해서는 빠른 적응을 하고, 예측오차의 변화가 작은 톤이나 음성 대역 데

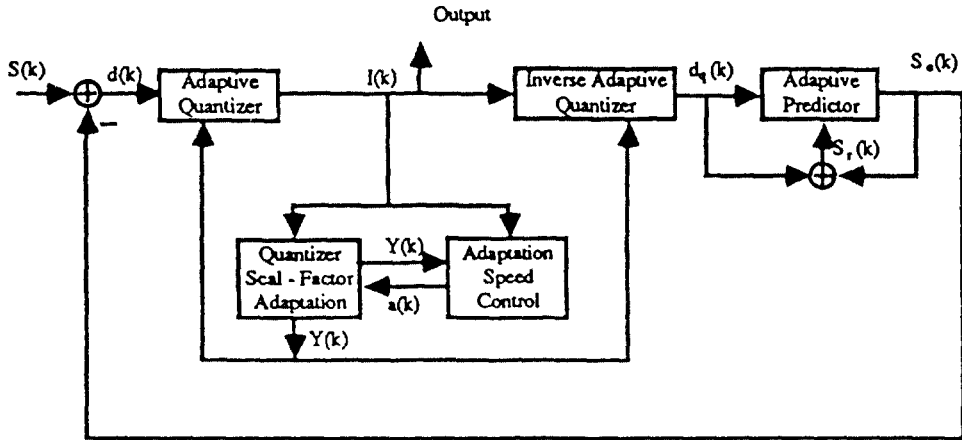


그림 4. 적응 양자화기의 블럭도
Fig.4. Block diagram of adaptive quantizer.

이와 같은 신호에 대해서는 느린 적응을 하도록 했다.

IV. 적응 포스트 필터링(Adaptive Postfiltering)에 의한 음질 개선

이 절에서는 필터된 음성에서 큰 왜곡이 없이 인식된 부호화 잡음(coding noise)을 크게 줄이는 새로운 적응 포스트 필터링(Adaptive Postfiltering)을 제안한다. 적응 포스트 필터링의 기본적인 개념은 인간의 청각 특성을 이용한 잡음 차폐 효과(noise masking)이다. 즉, noise shaping filter를 써서 백색 잡음인 양자화 오차의 분포를 음성의 스펙트럼과 비슷하게 분포시켜 주관적이고도 청각적인 음질의 개선을 가져온다^[2,3].

본 논문의 적응 포스트 필터링 기술은 Ramamoorthy & Jayant의 연구에 기초를 두고 있다^[2,16]. 그들이 제안한 포스트 필터는 16 kbps이상의 전송율에서 오로지 ADPCM에서만 사용되는 2 pole-6 zero ADPCM 합성 필터로부터 유도된다. 그들의 연구는 매우 중요하고 진보적인 개념이지만, low bit rate coding에 직접 응용할 수 없고, 바라지 않는 머플링(muffling) 효과, 즉, 입력 SNR이 낮을 때 잡음 스펙트럼이 음성 스펙트럼을 억누르는(swamp) 경향을 갖는다^[3].

본 논문에서는 두가지 방법의 적응 포스트 필터링을 수행하였는데 첫번째는 각 대역의 에너지 분포와

각 대역의 필터 계수를 이용한 방법이고 두번째는 LMS(Least Mean Square) 알고리즘을 이용한 귀환 적응 LPC(backward adaptive linear predictive coding) 해석 방법이다.

지금까지 제시되어 온 적응 포스트 필터링 기술은 효과적으로 잡음을 제거할 수 있음을 보여 주었다. 그러나, 부수적으로 머플링 효과를 갖는다. 본 논문의 새로운 포스트 필터링은 필터된 음성에서 많은 머플링없이 실제적인 잡음을 제거한다.

1. 음질 향상을 위한 적응 포스트 필터링 기법

본 논문에서는 합성음의 음질을 향상시키기 위해 수신단의 마지막에서 두가지의 적응 포스트 필터링 기법을 수행한다.

첫번째는 부가 정보를 이용한 것으로써, SBC에서 부가 정보는 각 대역 신호의 양자화 잡음에 영향을 받지 않으므로, 이 부가 정보를 사용하여 적응 포스트 필터를 설계할 수 있다. 이 적응 포스트 필터는 더 큰 에너지를 갖는 대역은 강조하고 작은 에너지를 갖는 대역은 상대적으로 감소시킨다. 이 적응 포스트 필터는 수신단에서 각 대역의 양자화된 rms 에너지와 각 구간에서 각 대역의 필터계수를 이용하여 설계할 수 있다. 또한 이 필터는 각 구간마다 다르게 설계한다.

$$g(n) = \sum_{k=1}^6 (A + (1+A)\gamma E(k)/S) h_k(n) \quad (1)$$

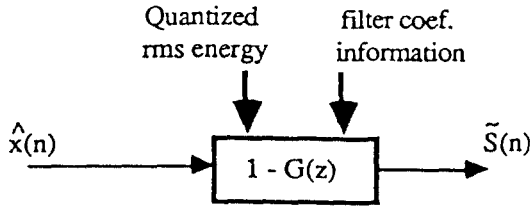


그림 5. 부가 정보와 각 대역의 필터 계수값을 이용한 포스트 필터

Fig.5. Postfilter using extra information and filter coef. of each band.

이 식 (1)에서 $E(k)$ 는 각 대역의 양자화된 rms 에너지이고, $S(k)$ 는 각 대역의 양자화된 rms 에너지의 합이며, $h_k(n)$ 은 각 대역의 필터 계수이다. 또한, 상수 A 와 γ 는 포스트 필터링의 상대적인 가중치를 주기위한 상수값으로 2.5와 0.8의 값을 가지며 $g(n)$ 은 설계된 적응 포스트 필터의 계수값이 된다. 이것의 블록도는 그림 5에 있다.

두번째는 backward adaptive LPC 해석 방법으로써 합성음으로 부터 LPC 계수를 구한 다음 적응 포스트 필터링을 수행한다. 이 적응 포스트 필터에서 예측 계수는 LMS 알고리즘을 이용하여 구했다.

$$\hat{x}(n) = h(1)x(n-1) + \dots + h(6)x(n-6) \quad (2)$$

$$e(n) = x(n) - \hat{x}(n) \quad (3)$$

$$h(k) = Bh(k) + usgn[e(n)]sgn[x(n-k)], \quad 1 \leq k \leq 6 \quad (4)$$

식 (4)에서 보듯이 전 샘플의 신호와 에러 신호의 sign 비트만 적응에 사용하므로 음성에서 양자화나 비트 에러에 의한 잡음을 감소시킬 수 있다. $\hat{s}(n)$ 이 적응 포스트 필터의 출력값이라고 할 때 식 (5)과 같이 나타낼 수 있다.

$$\hat{s}(n) = G\hat{x}(n) + g^h(1)\hat{s}(n-1) + \dots + g^h(6)\hat{s}(n-6) \quad (5)$$

식 (5)에서 $\hat{x}(n)$ 은 수신단에서 포스트 필터링되기 전의 합성음이고 $\hat{s}(n)$ 은 포스트 필터된 출력을 나타낸다. 또한 G 값은 필터의 전반적인 출력값이 원음과 비슷하도록 전력 이득(power gain)을 조절하는 값으로 시뮬레이션을 통해 0.5-0.75 사이의 값을 취하고 g 의 값은 포스트 필터 효과에 비례하는 값으로

12kbps에서는 0.7으로 16kbps에서는 0.7으로 16kbps에서는 0.6의 값을 사용한다. 이 적응 포스트 필터의 블록도는 그림 6에 나타내었다.

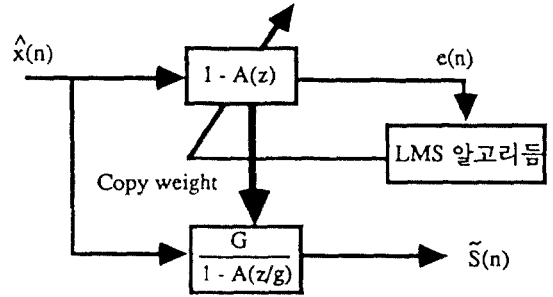


그림 6. 후방 적응 LPC 해석 방법을 이용한 포스트 필터
Fig.6. Postfilter using backward adaptive LPC analysis method.

V. 실험 및 결과 고찰

본 논문에서는 남녀가 각각 두 문장씩 발음한 네 개의 음성 데이터를 이용하여 SB-APCM, SB-ADPCM 등 각각에 대하여 12kbit/s, 16kbit/s의 중간 전송률에서 음질을 평가하였고, 각 부호화기에서 음질을 향상시키기 위해 수신단의 마지막에 적응 포스트 필터링(adaptive postfiltering)을 수행하였다. 각각의 부호화기의 음질을 주관적인 평가와 객관적인 평가를 통해 고찰했다.

실험 대상으로 사용한 데이터는 잡음이 없는 녹음실에서 남녀 하나운시가 발음한 4초 정도의 문장을 녹음하였다. 녹음한 이 데이터를 50-3400 Hz의 대역 통과 필터를 거친 후 표본화 주파수를 8 kHz로 하여 16비트로 A/D 변환하였다.

실험에서 사용한 문장은 다음과 같다.

- 1) 우리의 마음은 몸의 응용이다.
- 2) 불특정 다수에 대한 통계적 특성

위의 두 음성 데이터는 첫번째는 유성음이 많고 두번째는 무성음이 많은 데이터로써 각각 유성음과 무성음의 특성을 반영하도록 했다.

본 논문에서는 SBC 시스템에서 사용되는 대역 필터의 특성을 살펴보았는데, 지연과 에러의 파급 효과를 감소시키고 임의의 수로 대역을 분할할 수 있는 GQMF를 대역 필터로 선택하였다.

본 논문에서는 세가지 경우에 대하여 실험을 수행하였다. 첫째, SBC 시스템에서 부호화기 부분을 APCM을 사용한 경우, ADPCM을 사용한 경우와 APCM과 ADPCM을 함께 사용한 경우(ADPCM+APCM)에 대해서 성능을 평가하였다. APCM과 ADPCM을 함께 사용하는 방법은 비트를 할당할 때 2 비트 이상이면 ADPCM을 사용하고 1 비트이면 APCM을 사용하여, ADPCM 방식에 1 비트가 할당될 때 발생하는 큰 오차를 줄인다. 둘째는 각 시스템에서 비트 할당에 따른 성능을 평가하였다. 마지막으로 이 모든 SBC 시스템에서 청각적이고 주관적인 음질의 향상을 위하여 수신단의 마지막에 적응포스트 필터링을 수행한 다음에 성능을 평가하였다.

1. 비트 할당(bit allocation)에 따른 성능 평가

기존의 SBC 시스템에서는 각 대역마다 고정된 비트를 할당했으나, 본 논문에서는 각 대역에서 각 대역 신호들의 표준 편차를 통해 비트를 할당한 뒤 고정된 비트를 할당했을 때와 성능을 SEGSNR로 비교하였다. SEGSNR은 시간 영역에서 신호의 왜곡을 측정하는 방법으로, 전체 문장을 16ms의 구간으로 나누는 뒤에 각각의 SNR(Signal-to-Noise Ratio)을 구하고 이 각각의 SNR을 산술 평균한 값이다. 본 논문에서는 편이상 고정된 비트를 할당 했을 때에는 FB로 표시하고 적응 비트 할당을 했을 경우에는 AB로 표시한다. FB일 경우에 12 kbit/s에서는 SB-ADPCM은 (3,3,2,2,2,0), SB-APCM은 (4,3,1,1,1,0), SB-(APCM+ADPCM)은 (4,2,2,1,1,1)로 비트를 할당했고, 16 kbit/s에서는 (5,3,2,2,2,2), (5,3,2,2,1,1), (5,3,3,2,1,1)로 비트를 할당했다.

표 1과 표 2는 12 kbit/s과 16 kbit/s의 전송률에서 비트 할당에 따른 SEGSNR 값을 나타내었다. 표 1과 표 2를 통해 알 수 있듯이 고정된 비트 할당을 했을 때보다 적응 비트 할당을 했을 때 객관적인 성능이 향상됨을 알 수 있었다. 또한, 적응 비트 할당은 모든 비트에 상대적인 중요성을 정할 수 있으며 채널 코딩 시 효과적인 비트 에러 보호를 수행할 수 있고, embedded SBC에도 이용할 수 있다.

2. 적응 포스트 필터링에 의한 성능 개선

본 논문에서는 인간의 청각 특성을 이용한 적응 포스트 필터링을 수신단의 마지막에서 수행하여 음질을 개선하고자 했다.

표 1. 12 kbit/s의 전송률에서 비트 할당에 따른 성능 평가(SEGSNR : dB)

Table 1. Performance according to bit allocation with 12 kbit/s

	여자1	여자2	남자1	남자2
SB-ADPCM-FB	10.69	9.57	10.38	8.70
SB-APCM-FB	16.19	14.90	15.54	13.70
SB (ADPCM+APCM) FB	17.17	14.95	14.18	12.76
SB-APCM-AB	19.04	17.02	17.66	15.63
SB (ADPCM+APCM) AB	18.91	16.48	15.08	13.40

표 2. 16 kbit/s의 전송률에서 비트 할당에 따른 성능 평가(SEGSNR : dB)

Table 1. Performance according to bit allocation with 16 kbit/s

	여자1	여자2	남자1	남자2
SB-ADPCM-FB	12.24	10.62	12.13	10.36
SB-APCM-FB	19.62	18.00	18.56	16.93
SB (ADPCM+APCM) FB	20.43	17.89	16.31	15.19
SB-APCM-AB	20.93	19.19	19.81	18.19
SB (ADPCM+APCM) AB	20.77	18.04	16.38	14.82

본 논문에서 사용한 적응 포스트 필터링은 두가지 방법으로 수행하였다. 첫번째는 적응 비트 할당시에 사용되었던 각 대역의 에너지 분포를 나타내는 표준 편차와, 각 대역의 필터 계수를 결합하여 하나의 새로운 복합 필터를 설계하여 수행하였다. 이 방법을 편이상 P1이라 부르기로 한다. 두번째는 LMS 알고리즘을 통한 후방 적응 LPC (backward adaptive Linear Predictive Coding) 해석에 의해서 새로운 필터를 설계하여 수행하였고, 이 방법을 P2라고 부르기로 한다.

그림 7은 '우리의 마음은 몸의 응용이다.'라는 음성 신호에서 추출한 적당한 유성음 구간에 대한 포스트

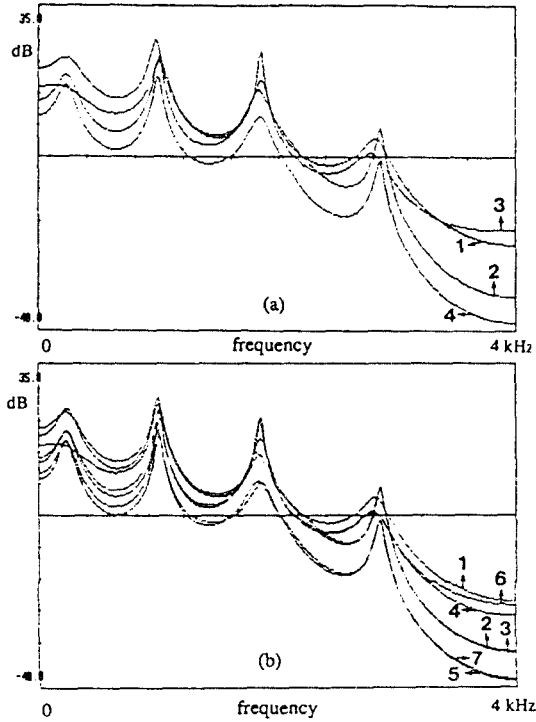


그림 7. 임의의 유성음 구간의 포스트 필터링을 한 뒤의 AR 스펙트럼

- (a) (1)원음 (2)SB-APCM-AB
- (3)SB-APCM-AB-P1 (4)SB-APCM-AB-P2
- (b) (1)원음
- (2)SB-(ADPCM+APCM)-FB
- (3)SB-(ADPCM+APCM)-AB
- (4)SB-(ADPCM+APCM)-FB-P1
- (5)SB-(APCM+ADPCM)-FB-P2
- (6)SB-(ADPCM+APCM)-AB-P1
- (7)SB-(ADPCM+APCM)-AB-P2

Fig.7. Postfiltered AR spectrum of arbitrary voiced speech samples.

- (a) (1)original speech
- (2)SB-APCM-AB
- (3)SB-APCM-AB-P1
- (4)SB-APCM-AB-P2
- (b) (1)original speech
- (2)SB-(ADPCM+APCM)-FB
- (3)SB-(ADPCM+APCM)-AB
- (4)SB-(ADPCM+APCM)-FB-P1
- (5)SB-(APCM+ADPCM)-FB-P2
- (6)SB-(ADPCM+APCM)-AB-P1
- (7)SB-(ADPCM+APCM)-AB-P2

필터링을 수행했을 때와 안했을 때의 AR 스펙트럼을 구해본 결과이다. 그림 7(a)는 각각 입력 데이터와 SB-APCM-AB와 P1과 P2로 포스트 필터링된 SB-APCM-AB 합성음의 AR 스펙트럼을 나타내고, 그림 7(b)는 각각 SB-(ADPCM+APCM)-AB의 P1, P2의 포스트 필터링했을 때의 AR 스펙트럼을 원음과 포스트 필터링 안했을 때의 AR 스펙트럼을 나타낸 것이다. AR 스펙트럼을 비교해 본 결과 P2의 경우가 포먼트 영역에서의 왜곡이 가장 적고 벨리 영역의 잡음도 적당히 억눌려 준다.

본 논문에서는 포스트 필터링에 의한 음질의 개선 정도를 측정하기 위하여 필터된 신호대 잡음비 SNRF('Filtered' Signal-to-Noise Ratio)를 구해본다. SNRF는 아래 그림 8과 같이 측정한다.

이 SNRF는 SNR 계산식에서 $x(n)$ 대신 $\hat{x}(n)$ 으로 바꾸어 계산하면 된다. 표 3과 표 4는 12 kbps와 16 kbps에서 P1 적용 포스트 필터링을 수행했을 때의 SNRF를 나타내고 표 5와 표 6은 P2 적용 포스트 필터링 SNRF를 나타낸다.

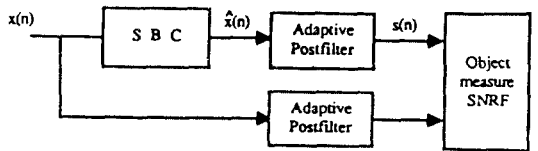


그림 8. 포스트 필터링된 신호대 잡음비(SNRF)의 측정
Fig.8. Postfiltered signal to noise ratio (SNRF) measure

표 3. 12 kbit/s에서 P1포스트 필터링된 신호대 잡음비 (SNRF : dB)의 평가

Table 3. Evaluation P1 postfiltered SNRF with 12 kbps.

	여자1	여자2	남자1	남자2
SB ADPCM-FB	12.89	11.84	13.15	10.87
SB-APCM FB	17.93	16.77	18.44	16.38
SB-(ADPCM+APCM)-FB	19.34	17.31	17.95	15.39
SB-APCM-AB	20.65	18.94	21.06	18.54
SB-(ADPCM+APCM)-AB	21.44	18.94	19.33	16.69

여기서 SNRF가 제일 높은 것만이 좋은 것이 아님을 주관적인 성능 평가인 표 7의 MOS(Mean Opinion Score) 테스트를 통하여 확인할 수 있었다.

표 4. 16 kbit/s에서 P1 포스트 필터된 신호대 잡음비 (SNRF : dB)의 평가

Table 4. Evaluation P1 postfiltered SNRF with 16 kbps.

	여자1	여자2	남자1	남자2
SB-ADPCM-FB	15.76	13.86	16.92	14.06
SB-APCM-FB	21.17	19.45	21.92	19.50
SB-(ADPCM+APCM)-FB	22.40	19.71	20.74	17.73
SB-APCM-AB	21.88	20.17	22.25	20.09
SB-(ADPCM+APCM)-AB	22.78	19.99	20.35	17.78

표 5. 12 kbps에서 P2 포스트 필터된 신호대 잡음비 (SMRF : dB)의 평가

Table 5. Evaluation P2 postfiltered SNRF with 12 kbps.

	여자1	여자2	남자1	남자2
SB-ADPCM-FB	11.24	10.30	11.57	9.61
SB-APCM-FB	17.23	16.14	17.08	15.34
SB-(ADPCM+APCM)-FB	18.05	15.95	15.62	14.09
SB-APCM-AB	19.90	18.16	19.04	17.17
SB-(ADPCM+APCM)-AB	19.52	17.54	16.38	14.91

3. MOS(Mean Opinion Score)Test

본 논문에서는 전송율이 16 kbit/s에서 SB-APCM-FB, SB-(APCM+ADPCM)-FB, SB-APCM-AB, SB-(APCM+ADPCM)-AB, SB-APCM-AB-P2, SB-(APCM+ADPCM)-AB-P2를 각각 방법1, 방법2, 방법3, 방법4, 방법5, 방법6으로 정하고 청취자가 다섯단계의 평가 등급을 매기는 주관적인 음질 평가인 MOS 테스트를 수행하였다.

표 6. 16 kbps에서 P2 포스트 필터된 신호대 잡음비 (SNRF : dB)의 평가

Table 6. Evaluation P2 postfiltered SNRF with 16 kbps

	여자1	여자2	남자1	남자2
SB-ADPCM-FB	12.94	11.44	13.77	11.56
SB-APCM-FB	20.36	19.14	19.76	18.57
SB-(ADPCM+APCM)-FB	21.01	18.95	17.76	16.59
SB-APCM-AB	21.53	20.12	20.64	19.46
SB-(ADPCM+APCM)-AB	21.36	18.93	17.50	16.15

표 7. MOS 테스트 결과표

Table 1. Result table of MOS test

	여자1	여자2	남자1	남자2
방법1	3.69	3.45	3.21	2.64
방법2	3.80	3.43	3.22	3.01
방법3	4.11	3.87	3.82	3.79
방법4	4.11	3.83	3.75	3.97
방법5	4.29	3.96	3.79	3.97
방법6	4.27	3.92	4.01	4.10

각 문장의 경우는 유성음이 많은 '우리의 마음은 몸의 응용이다.'의 문장을 여자1, 무성음의 특성을 반영하는 '불특정 다수에 대한 통계적 특성'의 문장은 여자2, 그리고 남자1, 남자2로 명명했을 때, 20대 남자 대학생 10명을 대상으로 시행한 MOS 결과는 표 7과 같다. 여기에서 '5'는 Excellent를 가르키고 '4'는 Good, '3'은 Fair, '2'는 Poor, '1'은 Bad를 나타낸다.

이 MOS 테스트 결과로써 합성음의 객관적 성능이 20 dB 가까이 되는 경우는 포스트 필터링을 통한 음질의 개선을 구별하기 힘들었으나, 원음에서의 잡음 제거를 수행하고, 또한 음질을 개선할 수 있었다.

본 논문에서 GQMF 대역 필터와 적응 비트 할당을 사용했을 때 음질이 개선되었고, 부호화기 부분을 APCM을 쓴 경우와 ADPCM을 쓴 경우를 비교해

본 결과, ADPCM의 경우에는 8개의 많은 대역으로 나누었기 때문에 각각의 대역에서 인접 샘플들간의 상관 관계가 약화되었고 ARMA 모델의 예측기를 쓰는 ADPCM의 장점이 약화되었기 때문에 APCM을 사용했을 때의 성능이 더 우수하였다.

VI. 결 론

본 논문에서는 크게 두가지 방법의 SBC 부호화 방법을 고찰했다. 첫번째는, Ramstad에 의해 제안된 개념을 적용한 것으로 적응 비트 할당을 한 SB-APCM이고, 두번째는, CCITT G.721 ADPCM 알고리즘을 적용한 적응 비트 할당의 SB-(ADPCM+APCM)이다.

이 두가지를 비교하기에 앞서 기존의 고정된 비트 할당의 SBC와 적응 비트 할당 SBC의 성능을 평가해 본 결과, 적응 비트 할당을 했을 경우에는 음질의 향상뿐 아니라, variable rate SBC에도 이용할 수 있을 것을 살펴보았다.

마지막으로, 각 방법에서 음질을 향상시키기 위해 수신단의 마지막에 적응 포스트 필터를 덧붙이는 방법을 적용했다. 두가지 기법의 적응 포스트 필터링을 수행하였는데 첫번째는 각 대역의 에너지 분포와 각 대역의 필터 계수를 이용한 방법이고 두번째는 LMS 알고리즘을 이용한 후방 적응 LPC(backward adaptive Linear Predictive Coding) 해석 방법이다. 그 결과, 상대적으로 낮은 복잡성과 음성에서 거의 무시할 수 있는 왜곡을 가진 채 상당한 잡음 감쇄를 이루는 것을 고찰할 수 있었다.

또한, SB-APCM과 SB-ADPCM을 비교해 본 결과, 많은 대역으로 나누었을 때에는 각 대역에서 샘플들 간의 상관 관계의 감소로 APCM이 성능이 우수한 것으로 나타났다.

본 논문에서 제안한 적응 포스트 필터링 기법은 잡음 있는 입력 음성의 음질을 개선하는데 사용할 수도 있다. 앞으로 주파수 영역뿐 아니라 시간 영역도 고려하여 비트 할당을 하고 각 대역 신호의 리던던시(redundancy)를 잘 제거하는 벡터 양자화 방법을 SBC에 적합하도록 개발한다면 작은 지연(delay)를 갖고 효율적인 더 좋은 성능의 SBC를 개발할 수 있을 것이다.

참 고 문 헌

1. F.K.Soong, R.V.Cox, and N.S.Jayant, "A High Quality subband speech coder with backward adaptive predictor and optimal time-frequency bit assignment," Proc. ICASSP, vol.4, no.3, pp.2387-2390, Apr. 1986
2. V.Ramamoorthy and N.S. Jayant, "Enhancement of ADPCM speech by Adaptive post-filtering," AT&T Bell Laboratory Technical Journal, vol.63, no.8, pp.1465-1475, Oct. 1984
3. M.R. Schroeber, B.S. Atal, and J.L. Hall, "optimizing digital speech coders by exploiting masking properties of human ear," JASA. vol. 66, no.6, pp.1647-1652, Dec. 1979
4. R.Hooke, T.Jeeves, "Direct Search Solution of Numerical and Statistical Problems", J.Ass. Comp. Mach, vol.8, no.3, pp.212-229, Apr. 1961
5. J.H.Rothweiler, "Polyphase Quadrature Filters-A New Sub-band Coding Technique", Proc. ICASSP, vol.3, no.3, pp.1280-1283, Mar.1983
6. T.A.Ramstad, "Consideration on Quantization and Dynamic Bit Allocation in Subband Coders", Proc ICASSP, vol.1, no.3, pp.841-844, Apr. 1986
7. H. H. Lee and C. H. Un, "Embedded waveform coding of speech", J. of KIEE., vol.21, no.3, pp.73-83, May. 1984
8. CCITT, "Recommendation G721 : 32kbts/s Adaptive Differential Pulse Code Modulation (ADPCM)", Red Book, vol. III, Oct. 1984
9. T. Taniguchi et al, "A high efficiency speech coding algorithm based on ADPCM with multiquantizer", IEEE. Proc. ICASSP86, vol.3, no.3, pp.1721-1724, Apr. 1986
10. N.S.Jayant and V.Ramamoorthy, "Adaptive postfiltering of 16 kbts/s ADPCM speech," Proc. ICASSP, vol.1, no.3, pp.829-832, Apr. 1986
11. R.E.Crochire, S.A.Webber, J.L.Flangan, "Digital Coding of Speech in Sub-bands", Bell Syst., Tech., J., Vol.55, No.8, pp.1069-1085, Oct. 1976

12. T.A.Ramstad, "Sub-band Coder with a Simple Adaptive Bit-Allocation Algorithm-A Possible Candidate for Digital Mobile Teleph-

ony?", Proc., ICASSP, Vol.1, pp.203-207, May. 1982

※본 연구는 한국과학재단의 연구비 지원에 의하여 이루어진 것임.



金元九(Weon Goo KIM) 정회원
 1964년 4월 3일생
 1987년 2월: 연세대학교 전자공학과 졸업.
 1989년 8월: 연세대학교 전자공학과 석사학위 취득.
 1989년 9월~현재: 연세대학교 전자공학과 박사과정 재학중.

• 주 관심분야는 디지털 신호처리 및 음성신호처리 등임.



李南杰(Nam Geol LEE) 정회원
 1965년 9월 29일생
 1989년 2월: 연세대학교 전자공학과 졸업.
 1991년 2월: 연세대학교 전자공학과 석사학위 취득.
 1991년 2월~현재: 삼성사중양연구소 연구원.

• 주 관심분야는 음성코딩, 음성합성 및 처리, 이동통신, audio 인터페이스 등임.



尹大熙(Dae Hee YOUN) 正會員
 1951년 5월 25日生
 1977년 2월: 延世大學校 電子工學科 卒業(工學士)
 1975년 5월: Department of Electrical Engineering, Kansas State University(M. S)

1982년 5월: Department of Electrical Engineering Kansas State University(Ph.D)

1982년 8월~1985년 6월: Department of Electrical and Computer Engineering, University of Iowa 助教授

1985년 9월~現在: 延世大學校 電子工學科 助教授



車日煥(Il Whan CHA) 正會員
 1935년 12월 25日生
 1959년 2월: 延世大學校 電氣工學科(工學士)
 1983년: 延世大學校 大學院 電子工學科(工學博士)
 1969년~1970년: 英國 University of Southampton Institute of Sound and Vibration 研究員

1986년 9월~1987년 7월: Department of Electrical and Computer Engineering, University of Iowa 客員教授

1973년~現在: 延世大學校 電子工學科 教授