

음성신호의 예측부호화에 관하여 (On Predictive Coding of Speech Signals)

殷 鍾 官

韓國科學技術院 電氣 및 電子工學科 教授

요 약

본 논문은 디지털 음성통신에서 사용되는 예측부호화(predictive coding) 방식에 관하여 기술하고 있다. 특히 전송속도가 16~48kbit/s 대역에서 많이 사용되고 있는 adaptive differential pulse code modulation (ADPCM)과 adaptive delta modulation (ADM)에 관하여 중점적으로 토의한다. 또한 variable-rate ADPCM과 ADM에 관해서 기술하고, 이들 시스템의 noisy channel에서의 효과 및 성능개선방법, 그리고 PCM과의 transcoding에서의 문제점 등을 토의한다. ADPCM은 최근 CCITT에서의 표준화 결과로 앞으로 PCM과 함께 많이 쓰여질 전망이다, ADM은 시스템이 간단하고 또한 channel error에 강한 이유로 특수통신에 많이 쓰여질 것이다.

I. 서 론

지난 20여년 동안 computer 기술, VLSI 및 micro-processor 기술, 그리고 디지털 신호처리 기술이 급격히 발전함에 따라 디지털 통신기술도 괄목할만한 발전을 하여왔다. 지금까지 통신의 주종을 이루어 온 음성통신은 종래의 아날로그방식을 탈피하고 디지털 통신방식으로 대체되고 있고, 근래에는 음성과 데이터가 집적하여 같은 디지털 채널로 송수신할 수 있는 효과적인 방식들이 연구되고, 한걸음 더 나아가 음성과 데이터 뿐만아니라 영상신호등 모든 신호를 한 개의 통신망을 통해서, 송신할 수 있는 종합정보통신망(ISDN)의 구현에 박차를 가하고 있다.

음성을 디지털화 하는데는 크게 분류해서 세가지 방법이 있다. 첫째는 음성파형을 직접 양자화하는 방법이고, 둘째는 음성의 주기와 성도의 계수등 음성의 특성만 추출하여 전송해서 수신측에서 음성을 합성하는 vocoding 방법, 셋째는 파형 양자화 방법과 vocoding 방법의 잇점만 사용하는 혼합부호화(hybrid coding) 방

식을 들 수 있다.^{[1][2]}

파형 양자화(wave form quantization) 방식의 대표적인 예로 pulse code modulation(PCM) 방식을 들 수 있다. PCM은 지난 20년 동안 음성신호의 디지털 전송을 위하여 주로 쓰여져 전 세계적으로 많은 시설 투자를 하여왔다. PCM 방식은 타 음성부호화 방식보다 훨씬 간단하여 single chip IC로 codec을 구현할 수 있는 장점이 있으나 현재 사용되는 전송속도가 64 kbits/s로서 타 부호화 방식보다 비교적 높은 대역폭을 갖는 단점이 있다.

근래 전화의 이용자가 급격히 늘어남에 따라 통신의 한정된 채널 대역폭을 유용하게 사용하는 일이 극히 중요하게 되었다. 이러한 이유로서 지난 15년 동안 음성을 낮은 전송속도로 디지털화 하여 전송할 수 있는 부호화 방법들이 많이 연구되었는데 그중 상용으로 앞으로 많이 실용화될 방식이 예측부호화(predictive coding) 방식이다.^[3] 예측부호화 방식은 파형 양자화 방식의 한 부류인데 음성신호의 전송을 위한 전형적인 전송속도는 16~32kbits/s로서 대표적인 예로서는 adaptive differential PCM(ADPCM)과 adaptive delta modulation (ADM) 이 있다. 이들중 특히 ADPCM은 1984년 말 CCITT에서 표준화가 되었는데, 전송속도가 기존 PCM의 반인 32kbits/s 밖에 되지않은 반면 음질은 64kbits/s의 PCM과 거의 같아 앞으로 음성통신에 PCM보다 더 많이 쓰여질 전망이다.^[4]

본 논고에서는 ADPCM과 ADM을 중심으로한 음성신호의 예측부호화 방식에 관하여 설명하고 이들의 특성, noisy channel에서의 효과, PCM과 tandem시의 문제등 부수되는 여러가지 점들에 관하여 기술한다.

II. Adaptive Differential Pulse Code Modulation (ADPCM)

1. ADPCM의 기본 원리

ADPCM을 고찰하기 전에 그림 1에 그려진 기본

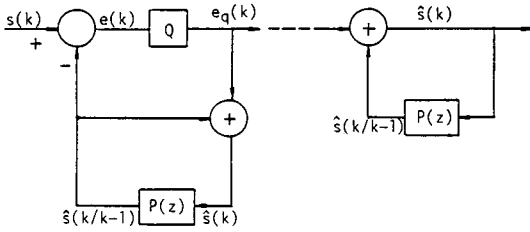


그림 1. 기본 DPCM시스템

DPCM을 간단히 설명하면 다음과 같다. DPCM에서는 대역제한후 Nyquist rate로 sample이 된 입력의 음성신호가 귀환회로(feedback path)의 예측기(predictor)에 의해서 예측이 되고 그 두 신호차 $e(k) = s(k) - \hat{s}(k/k-1)$ 를 L. level의 양자기(quantizer)에 의해서 양자화 하게 된다. DPCM의 양자화로 인한 잡음을 $n_q = e_q(k) - e(k)$ 라고 하자. 그러면,

$$\begin{aligned} S(k) &= \hat{S}(k/k-1) + e_q(k) \\ &= \hat{S}(k/k-1) + e(k) + n_q(k) \\ &= S(k) + n_q(k) \end{aligned} \quad (1)$$

를 얻는다. 위의 식(1)은 양자기나 예측기가 적응(adaptive) 또는 비적응(nonadaptive)함에 관계없이 언제나 성립한다. 그러므로 decode된 신호 $s(k)$ 를 원래의 입력신호 $s(k)$ 와 가능한한 같게 하기 위해서는 양자화로 인한 잡음 $n_q(k)$ 를 어떻게 최소로 만드는가에 따르게 되는데 여기에서 예측기와 양자기를 adaptive로 만들 필요성이 생기게 된다.

이를 위하여 양자기의 step의 크기를 양자기의 입력신호의 분산(variance)와 비례하게 만들어 주면 양자화로 인한 잡음전력은 입력신호의 전력에 거의 비례하게 된다. 그러므로 예측기를 adaptive로 만들어 입력신호를 가능한한 정확하게 예측해서, 양자기의 입력신호 $s(k) - \hat{s}(k-1)$ 의 분산을 줄이고, 더 나아가 양자기를 adaptive로 해서 이 분산과 match 시키면 DPCM을 최적 시스템으로 만들 수 있다.

ADPCM에서는 양자기와 예측기를 둘다 adaptive로 만들 경우 성능이 가장 좋으나 복잡도를 감안하여 경우에 따라서는 둘 중 하나만 adaptive로 만들고 다른 하나는 고정시킬 수 있다. 양자기를 adaptive로 만들 경우 예측기는 대개 tap이 하나인 누산기(accumulator)를 많이 쓴다.

예측기를 adaptive로 만들 경우에 어떻게 예측오차(prediction error)를 최소화 하는가에 따라 tap의 계수들을 구하는 방법이 여러가지 있는데, 크게 구분하면 tap계수의 수치를 정해진 시간마다 바꾸는 블록

(block) 해석 방법과, 매 sampling 시간마다 바꾸는 순차(sequential) 해석의 방법이 있다.¹⁴⁾

1) 적응 양자기(adaptive quantizer)

양자기의 step의 크기를 입력신호와 같이 만들기 위해서는 양자기의 입력신호(즉 오차신호)를 추정해야 하는데 그 방법에는 두 가지가 있다.

a) 전방추정(feedforward estimation)

이 경우는 음성신호의 진폭을 비교기(comparator) 앞에서 계산하는 것으로 이를 위해서는 신호를 저장하기 위한 buffer가 필요할 뿐더러, 또한 추측한 결과를 DPCM출력신호(즉 양자화된 오차신호)와 다중화해서 송신해야만 한다.

b) 귀환추정(feedback estimation)

양자기의 step의 크기를 양자화된 오차신호의 sample의 크기에 따라 귀환으로 결정하는 방법으로 물론 이 경우에는 추측된 결과를 송신할 필요가 없다.

위의 두 방법 중 전자를 쓰면 후자의 방법보다 정확하게 입력신호의 크기를 추정할 수 있으나 다중화 등 시스템이 복잡하게 되는 관계로 잘 안쓰고 후자의 방법을 많이 쓰고 있다.

양자기의 각 step level의 adaptation은 입력신호 진폭의 확률분포함수 probability density function(pdf)에 따라서 설계할 수 있다. 음성신호의 pdf는 실험적으로 얻은 바에 의하면 Gamma pdf,

$$p(x) = \frac{\sqrt{k}}{2\sqrt{\pi}} \cdot \frac{e^{-k|x|}}{\sqrt{x}}$$

또는 Laplacian pdf,

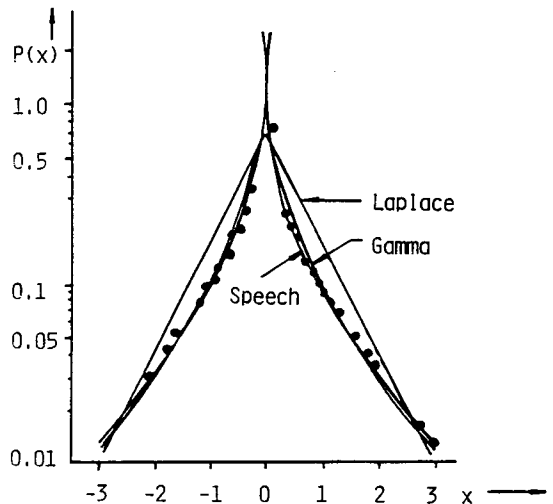


그림 2. 실제 음성의 진폭과 Gamma, Laplace 확률분포의 비교

$$p(x) = \frac{\alpha}{2} e^{-\alpha|x|}$$

에 근사하다. 여기서 k 와 α 는 입력신호의 rms값 σ_x 와 반비례 되는데 Gamma pdf에 대해서는 $k = \sqrt{0.75} / \sigma_x$ 이고 Laplacian pdf는 $\alpha = \sqrt{2} / \sigma_x$ 이다. 음성신호의 실제 진폭의 분포를 Gamma pdf, Laplacian pdf와 비교한 것이 그림 2에 그려져 있다. PCM에서 μ th law나 A1n law의 Log 압신(companding)을 쓰는 이유도 위와 같은 pdf를 가졌기 때문이다.

2) 적응 예측기(Adaptive Predictor)

a) 블록해석(Block Analysis)

DPCM에서 오차신호의 분산을 최소화하기 위해서는 예측기를 입력신호에 적용하도록 만들 필요가 있다는 것을 앞서 이야기한 바 있다. 적응 예측기를 쓸 경우에는 tap을 약 4개 정도 쓰게 되며 tap의 계수는 블록해석을 할때 다음과 같이 구하게 된다. 그림 1에서 보는 바와 같이 DPCM의 예측 오차신호는,

$$e(k) = s(k) - \hat{s}(k) \tag{2}$$

이다. 예측기의 tap이 p 개인 경우에 예측된 신호 $\hat{s}(k)$ 는 전의 N 개의 음성 sample을 가중(weighting)해서 가합으로서 얻을 수 있다. 즉,

$$\hat{s}(k) = \sum_{i=1}^p a(i) s(k-i) \tag{3}$$

위 식에서 $a(i)$ 를 예측계수라 한다. 따라서 식(2)를 다시 쓰면

$$e(k) = s(k) - \sum_{i=1}^p a(i) s(k-i) \tag{4}$$

가 된다. 여기서 예측계수는 잘 알려진 minimum-mean-square-error 과정에 의해서 구해지는 데 이를 위해서 N 개의 음성 sample에 대한 오차 에너지는

$$E = \sum_{k=0}^{N-1} e^2(k) = \sum_{k=0}^{N-1} [s(k) - \sum_{i=1}^p a(i) s(k-i)]^2 \tag{5}$$

가 된다. 위의 오차 에너지를 계수 $a(i)$ 에 관해서 최소화 하면 다음과 같은 식을 얻는다.

$$\bar{\Phi} \cdot \bar{a} = \bar{\Psi} \quad \text{또는} \quad \bar{a} = \bar{\Phi}^{-1} \cdot \bar{\Psi} \tag{6}$$

식(6)에서의 vector들은,

$$\bar{\Phi} \Delta = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \dots & \Phi_{1p} \\ \Phi_{21} & \Phi_{22} & \dots & \Phi_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{p1} & \Phi_{p2} & \dots & \Phi_{pp} \end{bmatrix}, \quad \bar{a} \Delta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}, \quad \bar{\Psi} \Delta = \begin{bmatrix} \Phi_{10} \\ \Phi_{20} \\ \vdots \\ \Phi_{p0} \end{bmatrix} \tag{7}$$

이다. 여기서,

$$\Phi_{ij} = \sum_{k=0}^{N-1} s(k-i) s(k-j) \tag{8}$$

이다.

위의 블록해석의 경우에는 입력음성 sample을 약 20ms로 나누어 기억기에 넣은 후 이 sample들을 분석해서 계수를 구하게 된다. 그렇게 함으로써 계수들의 수치를 매 20ms마다 다시 계산해서 송신하게 되는 데, 물론 이 계수의 수치는 양자화된 오차신호와 함께 다중화된 후 송신한다. 계수를 송신하기 위해서는 보통 선형으로 양자화한다. 이 경우에 coding의 정확도가 음질을 비교적 많이 좌우하지 않기 때문에 계수를 coding하기 위한 bit수는 4개의 계수를 송신하는 경우 약 400 bits/s에 불과하다.

b) 순차해석(sequential analysis)

적응 예측기에서 계수를 계산하는 또 한 가지 방법으로는 순차해석 방법이 있는데 이를 세분하면 adaptive gradient 방법과 kalman방법이 있다.^[3] 이 순차해석의 특징은 계수들의 스치가 일정한 시간마다 변경되는 것이 아니라 매 sampling 시간마다 계산되는 점이다. 또한 이 경우에는 예측계수를 직접 송신하지 않고, 수신기에서 decode된 음성 sample에 의해서 추정되는 것이 블록해석과 크게 다른 점이다.

순차해석에서 계수는 매 입력 sample마다 다음과 같이 추정하게 된다.

$$\hat{a}(k+1) = \hat{a}(k) + k(k+1) e_q(k+1) \tag{9}$$

여기에서

$$\hat{a}^T(k) = [a_1(k), a_2(k), \dots, a_p(k)]$$

$$\hat{a}^T(0) = [0, 0, \dots, 0]$$

이고, $e_q(k+1)$ 은 양자화된 $k+1$ 째의 오차신호이다. 식(9)에서 Gain $k(k+1)$ 은 순차해석의 방법을 adaptive gradient 방법으로 쓸 경우에는^[3]

$$K_{AG}(K+1) = \frac{0.08 \hat{S}_N(k)}{100 + S_N^T(k) S_N(k)} \tag{10}$$

이다. 여기에서

$$\hat{S}_N^T = [\hat{S}(k), \hat{S}(k-1), \dots, \hat{S}(K-N+1)]^T \tag{11}$$

이다.

2. ADPCM System의 예

1) Bell ADPCM System

ADPCM은 지난 10여년 동안 여러방식이 제안되었으나 그중 비교적 시스템이 간단한 반면 성능이 좋은 것으로 그림 3에 도시한 Bell연구소의 연구원들이 제안한 system을 들 수 있겠다.^[5] 이 ADPCM 시스템은 예측기는 고정되어 있고 양자기만 적용하도록 되어있

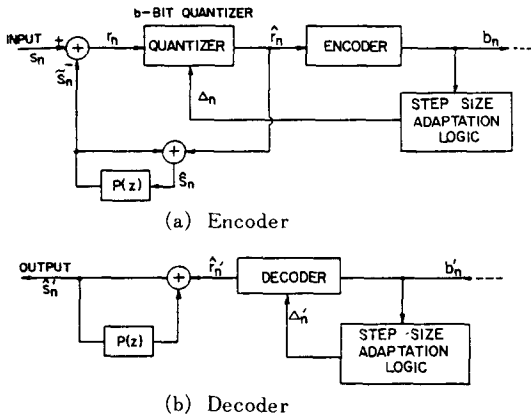


그림 3. ADPCM 블록도

다. 이 양자기는 step의 크기를 입력신호의 분산에 맞추기 위해서 1 word의 memory를 쓴다. 예를 들어 B-bit의 균일(uniform) 양자기의 출력률,

$$Y_r = H_r \frac{\Delta_r}{2}$$

$$\pm H_r = 1, 3, 5, \dots, 2^B - 1 \quad (\Delta_r > 0, B \geq 2)$$

라 하자. 그러면 현 step의 크기 Δ_r 은 전 step의 크기 Δ_{r-1} 와 고정된 multiplier 함수 $M(H_{r-1})$ 를 곱해서 얻는다;

$$\Delta_r = \Delta_{r-1} M(H_{r-1})$$

양자기의 level의 수가 4, 8, 16인 경우, 컴퓨터 simulation으로 얻은 multiplier 함수의 최적치[SQNR(signal-to-quantization noise ratio)을 최대화 하는 점에서]들이 도표 1에 나타나 있다. 위의 한 예로서 그림 4에 3-bit(또는 8 levels)의 적응 양자기가 그려져 있다.

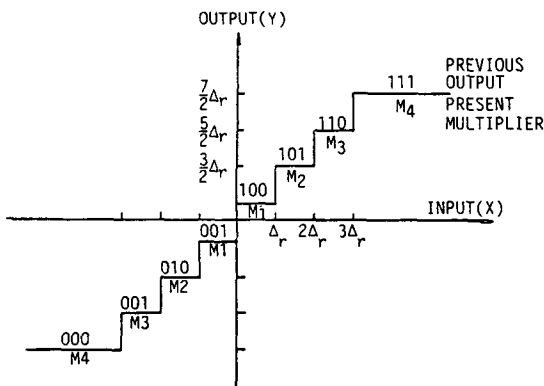


그림 4. ADPCM 3-bit quantization

표 1. ADPCM quantizer multiplication factors.

	M ₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₇	M ₈
4 levels	0.8	1.6						
8 levels	0.9	0.9	1.25	1.75				
16 levels	0.9	0.9	0.9	0.9	1.2	1.6	2	2.4

2) CCITT ADPCM System

최근 CCITT의 음성통신 연구 group에 의해서 제안된 ADPCM 시스템은 앞으로 디지털 음성통신에 큰 영향을 줄 것으로 기대된다. 이 CCITT 표준 ADPCM의 특성을 몇가지 열거하면 다음과 같다.¹⁴⁾

- Codec이 전송속도가 32kbps에서 운용되도록 설계되고
- 음성신호 뿐만 아니라 전화선을 사용하는 데이터 신호 그리고 tone 신호까지도 부호화할 수 있도록 설계되고
- 현재 64kbps PCM과 직접 연결이 가능하고
- 예측기와 양자기를 둘다 입력신호에 적응하도록 설계되고
- 채널 에러(channel error)에 비교적 강한 성능을 갖고 있다.

이 ADPCM 시스템은 올해 안에 VLSI화 되어 대량 생산이 예상되는 바 전송속도가 기존 PCM의 반밖에 되지 않는 반면 음질이 좋아 현재 설치된 PCM 시스템을 모두 대체하기는 경제적인 면을 고려할 때 어렵겠지만 앞으로 디지털 음성통신에서 주된 역할을 할 것임은 틀림없는 사실이다.

CCITT 표준 ADPCM 시스템의 coder와 decoder가 그림5,6에 각각 그려져 있다. 이 codec의 각 sub-system에 관하여 설명하면 다음과 같다. 먼저 이 ADPCM은 기존 64kbps PCM과 직접 연결이 되기 때문에 ADPCM으로 부호화되기 전에 μ -law로 양자화된 신호는 선형으로 바뀌어야 된다. 이 신호가 바로 ADPCM의 입력신호가 되고 예측기로 부터 예측된 신호와 빼져서 오차신호를 발생하게 된다. 예측기에서 사용되는 filter계수의 update 방식은 기본적으로 간략화된 gradient 방식으로 매 sampling 시간마다 변화시키는 순차 적응방식을 사용함으로써 예측기의 계수들을 별도로 송신하지 않도록 하였다. 예측 filter의 구조는 입력신호를 보다 정확히 할 수 있도록 6차의 zero와 2차의 pole을 modeling하는 pole 예측기와 zero 예측기의 두 section의 filter로 구성되어 있다. 한편 적응양자기는 level수를 2⁴(즉 4bit)개 갖고 log 성격을 갖는 순간압신방식을 사용하고 있는데 입력신호의 종류에

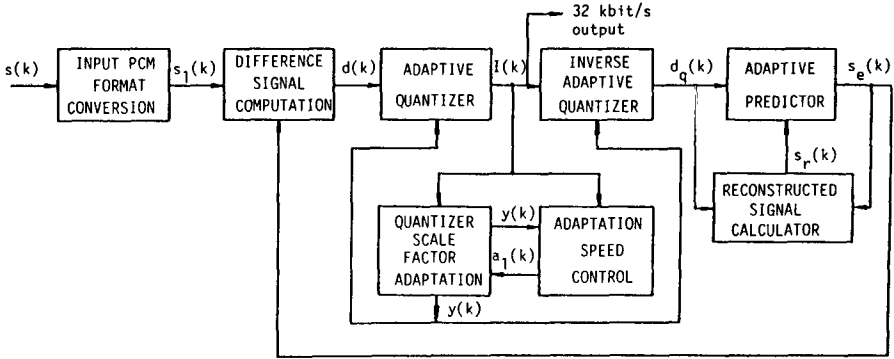


그림 5. CCITT ADPCM 송신기

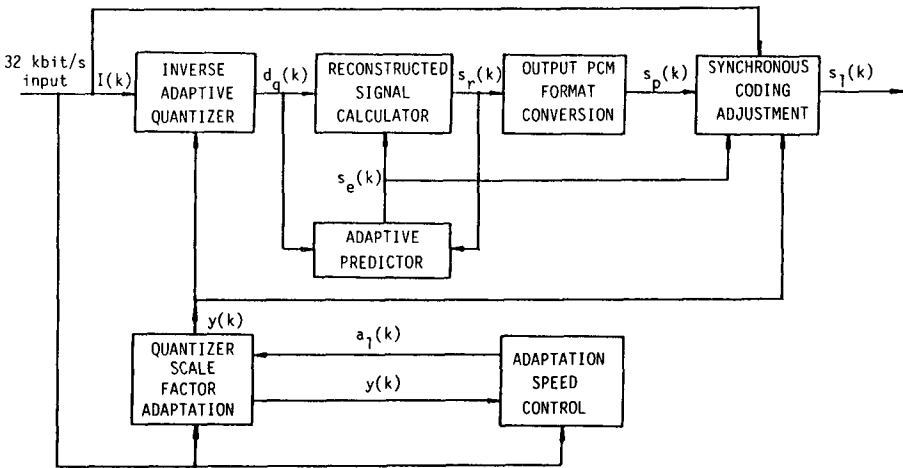


그림 6. CCITT ADPCM 수신기

따라 음성의 같이 급히 진폭이 변하는 경우에는 “fast mode”로 양자기의 model을 적응시키고 PSK등 변조된 데이터 신호와 같이 비교적 진폭의 변화가 작은 신호에 대해서는 “slow mode”로 작동하도록 설계되어 있다.

복호기(decoder)의 구조는 기본적으로 부호기의 feedback 부분과 같고 여기에 μ -law 압입의 PCM 형태 신호로 바꾸는 신호 변환기가 추가되어 있다.

위에서 설명한 CCITT 표준 ADPCM은 지금까지 연구개발된 ADPCM중 가장 복잡한 시스템의 하나이지만 hardware를 구현하는데 있어서는 multiplier를 거의 사용하지 않고도 구현할 수 있도록 설계되어 있다. 시스템의 성능면에서도 예측기와 양자기를 모두 입력 신호에 적응되도록 만들어져 음성신호를 부호화하는 경우 신호대 잡음비가 28dB로서 타 ADPCM보다 약 4dB가 좋으며 voiceband 데이터 신호를 전송하는 데

도 사용할 수 있다는 것이 큰 장점이라 하겠다.

III. Adaptive Delta Modulation(ADM)

1. ADM의 기본원리

선형 예측부호기의 또 한 부류로는 ADM을 들 수 있다. ADM도 입력음성 신호를 직접 양자화하지 않고 오차신호를 양자화하는 점에서는 ADPCM과 같으나, 다른 점은 입력신호의 sampling rate를 PCM이나 ADPCM에서 사용하는 nyquist rate 보다 훨씬 높은 (보통 2배 내지 4배) rate를 쓰고 대신 양자기의 level수는 단 2개(즉 1bit)를 갖는다. 다시 이야기하면 PCM이나 ADPCM은 입력음성이 word 단위로 부호화되지만 ADM에서는 bit단위로 부호화된다. 따라서 ADM의 hardware의 구현은 ADPCM보다 훨씬 간단하며, 입력 및 출력단에서 사용하는 filter도 sampling을 nyquist rate 보다 훨씬 높게 하기 때문에 ADPCM

에서 사용되는 filter 보다 훨씬 간단하다. ADM을 토의하기 전에 먼저 양자기의 step의 크기가 고정된 선형 DM(LDM)을 고찰하기로 하자. LDM이 그림 7에 그려져 있다. 먼저 입력신호가 들어오면 이는 예측기에 의해서 예측이 된 후, 이 두신호가 비교기에 의해서 비교된다. 입력신호가 예측된 신호보다 클 경우에는 비교기에서 binary 신호 "1"이 나오고 반대일 경우에는 "0"이 나온다. 물론 이 binary 신호의 signaling rate(전송속도)는 양자기의 clock rate와 동일하다. DM의 decoder는 encoder의 귀환회로(feedback path)와 동일하며 양자화로 인한 잡음을 줄이기 위해서 입력 filter와 같은 filter가 decode된 후에 쓰여진다.

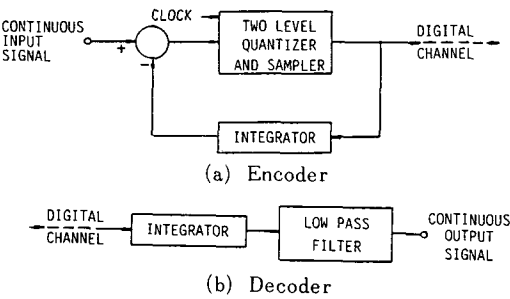


그림 7. 선형 delta modulation

ADM이 LDM과 근본적으로 다른 것은 양자기 step의 크기가 LDM은 항상 고정된 반면 ADM은 입력신호의 변화에 따라 커지고 작아지는 점이다. 양자기 step의 크기가 입력신호에 적응(adaptive)하도록 하기 위해서는 전방적응압신(feedforward adaptive companding) 방법과 귀환적응압신(feedback adaptive companding) 방법이 있는데, 주로 후자의 방법을 쓰고 있다. 귀환방법으로서 압신을 하는 경우에는 coding된 디지털신호에 따라서 같은 신호가 계속나오게 되면 step의 크기를 크게 만들어 slope overload 잡음을 적게 하고, binary 신호가 매 sampling 시간마다 계속 바뀌게 되면 step의 크기를 작게 해 양자화 잡음을 적게 한다. 따라서 step의 크기를 입력신호에 따라 조정함으로써 양자화잡음을 적게 할 뿐만 아니라, 또한 LDM에서는 불가능했던 큰 dynamic range를 얻을 수 있다. ADM에서 양자화잡음이 step의 크기에 따라 어떻게 변화되는가가 그림 8에 나타나 있다. 이 그림에서 양자기의 step크기의 최적범위가 ADM이 LDM보다 넓고, 또한 양자화잡음도 적은 점을 주시할 수 있다.

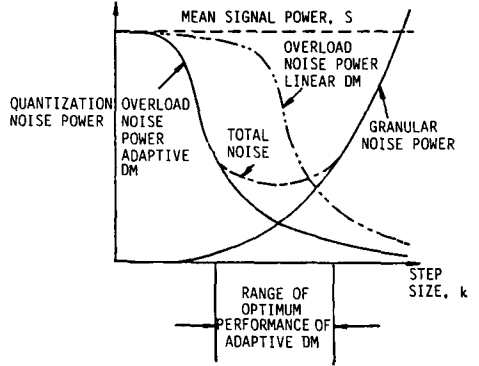


그림 8. ADM 양자화 잡음 전력

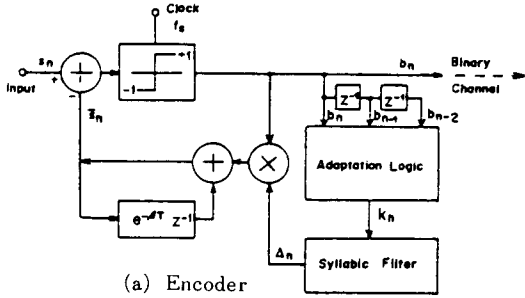
ADM은 대개 양자기의 기본 step의 크기를 어떻게 압신하는가에 따라 다음과 같이 구분할 수 있다.

- (1) 양자기의 step의 크기가 syllabic rate로 서서히 변하는 경우, 이를 음절압신 (syllabic companding) 이라고 하고 그 예로서는 continuously variable slope delta modulation (CVSD) 이나 digitally controlled delta modulation (DCDM) 을 할 수 있다.^[6]
- (2) Step의 크기가 매 sampling 시간마다 변하는 경우 이를 순시압신 (instantaneous companding) 이라고 하는데 그 예로서는 winkler의 high information delta modulation과 Jayant의 constant factor DM (CFDM) 방법이 있다.^[6]
- (3) 마지막으로 위의 두 방법을 겸용한 양자기를 사용하는 Hybrid companding 방법이 있다.^[7]

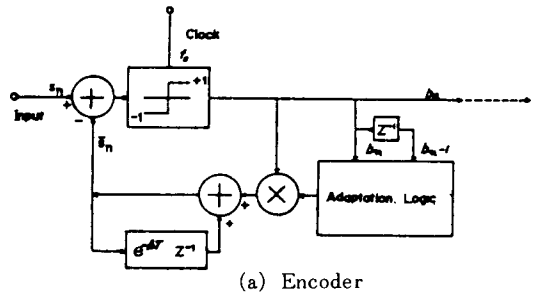
2. ADM System의 대표적 예

1) Continuously Variable Slope DM (CVSD)

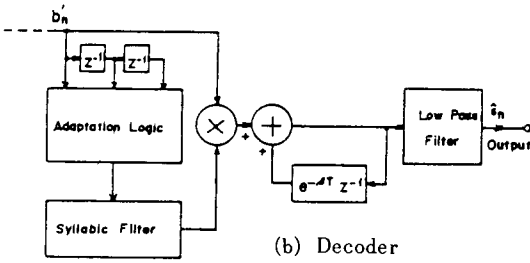
ADM의 여러방식중 현재 가장 많이 쓰여지고 있고 syllabic 압신 방법을 사용하는 codec으로서 CVSD를 들 수 있다. 그림 9에 그려진 CVSD의 원리를 간단히 기술하면 다음과 같다. Encoder는 주로 비교기(comparator)와 양자기, 압신기, 그리고 예측 filter의 세 부분으로 나눌 수 있다. 음성신호가 대역제한(보통 0.3-3.4kHz)되어 encoder에 들어오면 비교기는 이를 예측한 신호와 비교하여 오차신호를 발생시키게 되고, 양자기는 1bit 양자기로 이 오차신호를 매 sampling 시간마다 sample해서 그 극성을 송신하게 된다. 양자기의 step의 크기는 CVSD가 adaptive인고로 binary bit에 따라 조정이 되는데 이 조정을 바로 압신기가 하게 된다. 압신기는 3-bit 또는 4-bit의 Shift register와 비교기, 그리고 3dB점이 약 40Hz 정도의 간단한 RC low-pass filter로 구성된다. 압신기의 논리(logic)



(a) Encoder

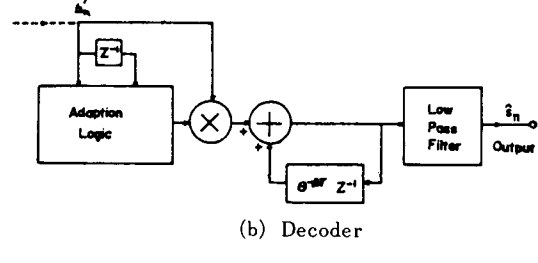


(a) Encoder



(b) Decoder

그림 9. CVSD 블록도



(b) Decoder

그림 10. HCDM 블록도

는 shift register에서 나오는 셋 또는 네개의 binary 신호를 비교기에서 동시에 비교해서 그들의 극성이 모두 같은 경우에는 pulse를 발생시키고, 다른 경우에는 안시키게 된다. 예측 filter는 RC 적분기에 불과하며, 입력신호는 압신기의 law-pass filter된 출력과 송신되는 binary 신호에 의해서 정해진다. CVSD의 decoder는 그림10의 encoder의 feedback 부분과 같다.

2) Hybrid Compand DM (HCDM)

HCDM은 지금까지 제안한 여러 ADM 시스템 중에서 성능이 가장 좋고 음성신호와 같이 non-stationary 하고 진폭이 많이 변화하는 신호에 대하여 특히 효과적인 것으로 알려져 있다.¹⁷⁾ HCDM은 음절압신과 순시압신을 둘 다 사용하는데 codec의 블록도가 그림10에 그려져 있다.

먼저 band-limit된 입력신호가 들어오면 first-order의 예측기에 의해서 예측된 신호와 비교되어 오차신호를 발생시키고 이 오차신호의 부호가 출력신호로 나가게 된다. 양자기의 step의 크기는 두가지 방법으로 조정되는데 먼저 예측된 신호로서 step의 기본 크기를 정하여 음성의 변화에 따라 매 5ms 정도의 시간마다 변화를 시키는 동시에 HCDM의 출력 bit pattern을 보고 매 sampling 시간마다 순시압신을 시킨다. 순시압신은 여러가지 방법이 있지만 가장 최적방법으로 알려진 것은 표2에 보는 바와 같이 3개의 출력 bit를 보고 계속 같은 bit이 나가면 step의 크기를 증가시키고 bit pattern이 바뀌게 될 때는 크기를 감소시

표 2. HCDM companding logic.

b_n	b_{n-1}	b_{n-2}	Multiplication Factor
+	+	+	1.5
-	-	-	1.5
-	-	+	1
+	+	-	1
-	+	+	0.66
+	-	-	0.66
-	+	-	0.66
+	-	+	0.66

키거나 먼저 크기 그대로 두는 방법이 효과적인 것으로 알려져 있다. HCDM의 binary 출력신호 b_n 이 만들어지는 식을 표현해 보면

$$b_n = \text{sgn}(s_n - \bar{s}_n)$$

이다. 여기서

$$\bar{s}_n = e^{-\delta T} \bar{s}_{n-1} + b_{n-1} \Delta_{n-1}$$

$$\Delta_n = \Lambda \gamma_n$$

$$\gamma_n = k_n \gamma_{n-1}$$

$$k_n = f(b_n, b_{n-1}, b_{n-2}) \text{ 이다.}$$

HCDM의 decoder는 coder의 feedback 부분과 똑같으며 decode된 신호는 최종적으로 law-pass filtering을 함으로써 음성이 재생된다.

IV. ADPCM 과 ADM의 성능 비교

앞서 토의한 바와 같이 ADPCM과 ADM은 둘 다 예

측부호기인 점에서는 공통점이 있으나, 그들의 특성이나 실제적으로 회로를 구성하는데 있어서는 여러가지로 다른점이 많다. 예를 들면 다음과 같다.

- 낮은 전송속도에서는 ADM의 SQNR이 ADPCM보다 좋으나 높은 전송속도에서는 ADPCM이 더 좋다 (그림11 참조). 따라서 높은 질이 요구되는 음성이나 영상신호의 부호화 방법으로는 ADPCM의 사용이 보다 더 바람직하다.

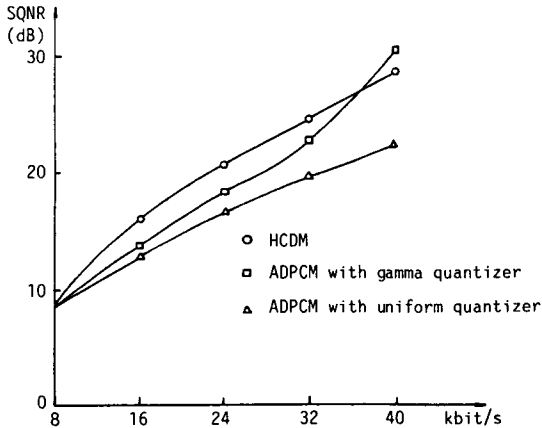


그림11. Bell ADPCM과 HCDM의 성능비교

- ADPCM은 ADM보다 입력신호의 phase를 보다 정확히 재생할 수 있다. 이 점은 psk등 변조된 data 신호를 coding하는데 큰 장점이라 하겠다.
- ADPCM이나 ADM을 PCM과 같이 tandem으로 사용할 때 부호변환(Transcoding)의 문제가 있는데 (제11장 참조) ADPCM은 PCM과 같은 sampling rate을 사용하기 때문에 ADM보다 용이하게 변환할 수 있다. 이점은 실제적으로 중요한 문제로서 CCITT에서 32kbits/s의 coder로서 ADM 대신 ADPCM을 택한 주 이유의 하나가 바로 trans coding에 있다. 위의 예들은 ADM에 대한 ADPCM의 대표적인 장점인데, 한편 ADM의 장점은 다음과 같은것을 들 수 있다.
- ADPCM이나 PCM은 multi-level의 양자화를 요하기 때문에 word의 동기(synchronization)가 필요하나 ADM은 필요가 없다.
- ADPCM은 bit를 잃어버렸을 때 경우에 따라 영향이 크나 ADM은 비교적 영향이 적다.
- ADPCM은 channel 오차의 영향이 크고 오차도가 최고 10^{-3} 정도 허용되지만 ADM은 영향이 적을뿐만

아니라 허용 오차가 최고 10^{-1} 까지 올라갈 수 있다.

- ADPCM은 sampling할때 aliasing문제 때문에 입력 filter의 주파수 특성의 경사가 급변해야 하므로 filter의 구성이 복잡해지나 ADM의 경우는 over-sampling의 이유로 복잡한 filter를 요하지 않는다. 앞서 기술한 Bell 연구소의 ADPCM 시스템과 ADM중 성능이 가장 좋은 HCDM 시스템의 SQNR이 channel에 잡음이 없을 경우 그림11에 나타나 있다. 그림에서 보는 바와 같이 전송속도가 35kbits/s 이하에서는 HCDM의 성능이 ADPCM보다 좋으나 그 이상에서는 ADPCM이 좋음을 알 수 있다. 또한 ADPCM의 양자기를 uniform quantizer 보다 gamma quantizer를 사용하는 것이 좋은 것으로 나타나 있다. 그 이유는 앞에서 설명한 바와 같이 음성신호나 오차신호의 진폭의 분포함수가 uniform 분포함수 보다는 gamma 분포함수에 가까우기 때문이다.

V. Variable-Rate ADPCM 및 ADM 시스템

앞장에서 기술한 ADPCM과 ADM은 전송속도가 주로 32kbits/s로 쓰여지는데 전송속도를 16kbits/s로 줄여 channel의 수를 늘리거나 또는 전송속도를 늘리고 성능을 보다 더 좋게하기 위해서는 시스템의 개선이 필요하다. 본장에서는 Bell ADPCM과 HCDM을 근거로한 variable rate coding에 관하여 기술한다.

일반적으로 음성을 전송하는데 있어서 variable rate coding을 하는 방법은 두가지로 나눌 수 있다. 하나는 coder와 decoder에서 전송속도를 직접 변화시키는 방법이고 또하나의 codec과는 관계없이 전송시스템에서 channel의 상태에 직접 전송속도를 변화시키는 방법이다. 전자의 경우의 예로는 speech interpolation이나 음성의 유성음, 무성음의 변화에 따라 coder의 전송속도를 변화시키는 variable rate coder를 들 수 있고,¹¹⁰⁻¹¹¹ 후자의 예로는 embedded coding 방식을 들 수 있다.¹¹¹ 아래에 variable-rate sampling HCDM, embedded ADPCM과 HCDM에 관해서 기술하겠다.

1. Variable-rate sampling HCDM 시스템¹¹¹

Variable-rate sampling HCDM (VSHCDM)에서는 고정된 sampling rate을 쓰는 대신 입력음성의 분산 또는 slope energy의 변화에 따라 몇개의 다른 sampling rate를 사용한다. ADM에서 sampling rate는 전송속도와 같다는 점은 앞서 언급하였다. 따라서 ADM의 출력율을 동기식 고정속도로 전송하기 위해서는 VSHCDM의 끝단에 buffer가 필요하다.

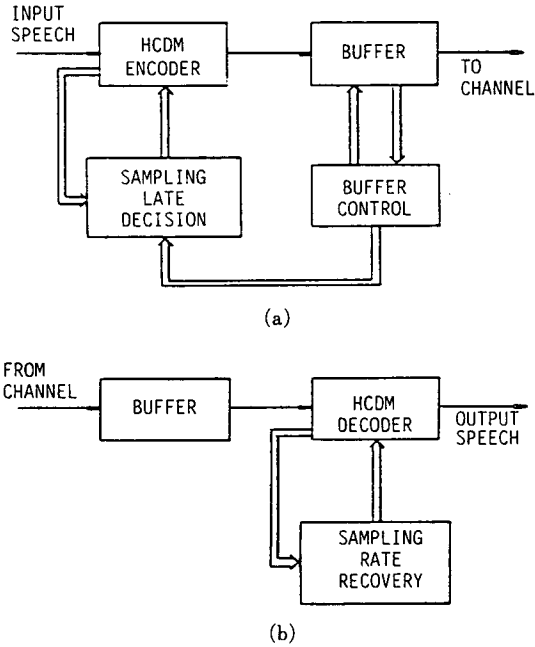


그림12. Variable-sampling rate HCDM 시스템

VSHCDM 시스템의 block diagram이 그림12에 그려져 있다. 여기서 HCDM은 제IV장에서 설명한 것과 같다. 이 시스템에서는 양자기를 입력신호의 변화에 따라 적응 압신시키는 외에 sampling rate도 입력신호의 slope 추정치의 변화에 따라 변경시켜 준다. 본 시스템에서는 다음과 같이 네개의 sampling frequency f_{s1} 를 사용한다.

- $f_{s1} = 6 \text{ KHz}, E_1 < E_1$
- $= 12 \text{ KHz}, E_1 \leq E_1 < E_2$
- $= 16 \text{ KHz}, E_2 \leq E_1 < E_3$
- $= 24 \text{ KHz}, E_1 \geq E_3$

여기서 E_i 는 i 번째 block의 음성 slope energy이고 E_1, E_2, E_3 는 threshold 값들이다. 위의 sampling frequency들은 48KHz의 공통인수들로서 출력단의 전송속도는 16kbits/s로 된다.

앞서 언급한 바와 같이 VSHCDM의 sampling rate은 변화하지만

$$f_{s1} \leq F_s, \quad b_1 \geq \frac{9}{10} B$$

$$f_{s1} \geq F_s, \quad b_1 < \frac{1}{10} B$$

여기서 F_s 는 전송속도이고 b_1 는 i 번째 frame에서 송신 buffer의 bit수이고 B 는 buffer의 크기이다.

VSHCDM과 기존 HCDM 및 CVSD의 성능을 ideal

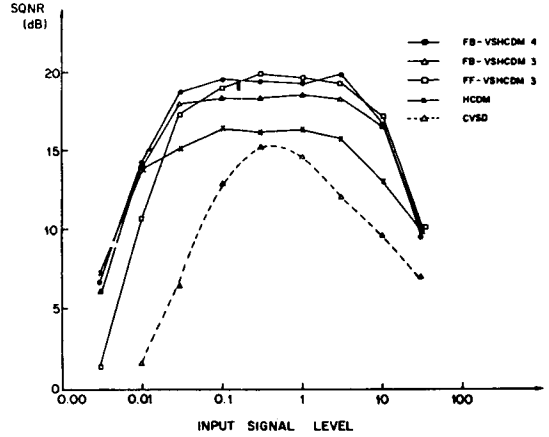


그림13. Ideal channel에서의 VSHCDM, HCDM 및 CVSD의 성능비교

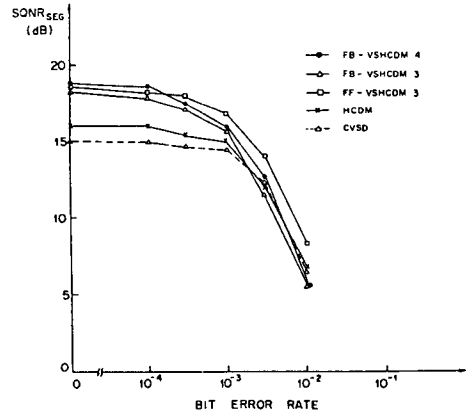


그림14. Noisy channel에서의 VSHCDM, HCDM 및 CVSD의 성능비교

및 noisy channel에서 비교한 것이 그림13과 그림14에 각각 그려져 있다. 이 결과에 의하면 VSHCDM이 보통 HCDM보다 약 4dB가량 좋은것으로 나타나 있다.

2. Embedded Predictive Coding Systems

앞서 기술한 바와 같이 embedded coding은 codec과는 관계없이 전송시스템에서 channel의 상태에 따라 coding된 bit stream의 일부를 priority 순서에 의해서 버려 전송속도를 조정한다. Embedded coding은 packet 교환 방식에 의해서 음성을 전송할 때 channel의 flow control을 효과적으로 할 수 있기 때문에 근래에 많은 연구가 되어 왔다. 여기서는 embedded ADPCM (EADPCM)과 embedded HCDM (EHCDM)

을 기술한다.^[11]

1) Embedded ADPCM System

EADPCM의 구조가 그림15에 그려져 있다. 이 시스템은 오차신호 r_n 의 coding을 위한 ADPCM과 양자화해서 생기는 error신호 q_n 의 coding을 또하나의 coder로 구성된다. 따라서 channel의 전송속도는 bit stripping과 insertion방식에 의해서 sampled B bit부터 $A(=B+R)$ bit까지 변화시킬 수 있다(그림15 참조).

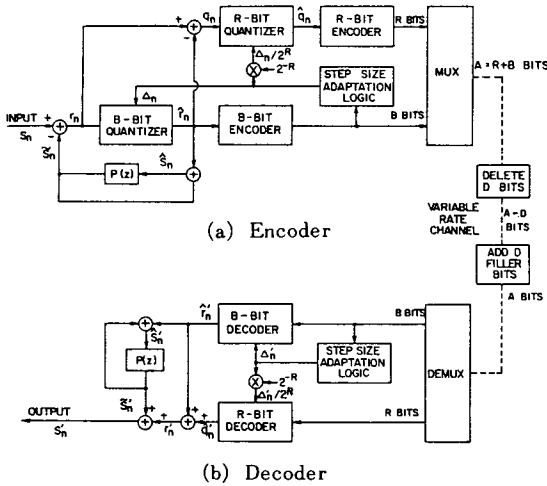


그림15. EADPCM 시스템의 블록도

여기서 ADPCM은 제 II 장에서 기술한 Bell ADPCM을 쓰고 양자기 error신호 coding을 위해서는 R-bit의 uniform quantizer의 PCM을 쓴다. Embedded coding을 위해서 error신호는 sign-magnitude dinary form으로 coding이 되는데 전송속도는 sample당 D개의 LSB를 버림으로서 R로부터 R-D bit으로 줄일 수 있다. 수신측에서는 (D+1)개의 bit를 (한개의 1과 D

개의 0) 수신된 R-D개의 bit에다 첨부를 시켜decoding을 하게 된다. 이 경우 물론 전송속도는 낮아지지만 음질은 저하가 될것이다. 그림16에 $\Delta_q/2$ 의 odd multiple을 4bit 이진수로 나타내는 8-level의 mid-riser quantizer와 4bit중 마지막 네번째와 세번째 bit를 drop하고 receiver에서 1과 0를 첨부하여 만든 4-level의 quantizer가 예시되어 있다. 그림17에 EADPCM의 성능을 16에서 64kbts/s까지 변화시켜 얻은 결과가 그려져 있다.

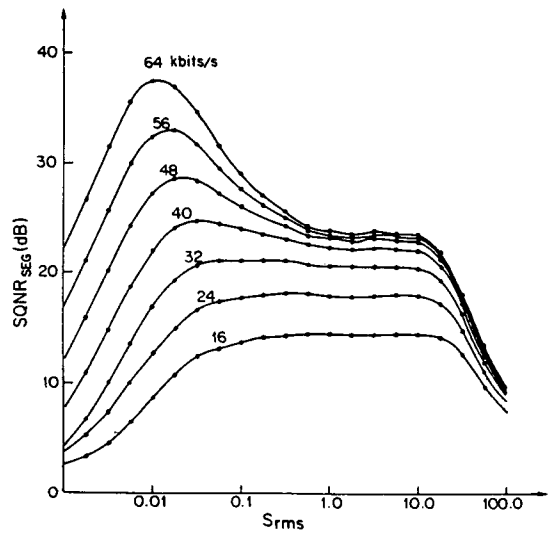


그림17. 16~64kbts/s 전송속도에서의 EADPCM의 성능

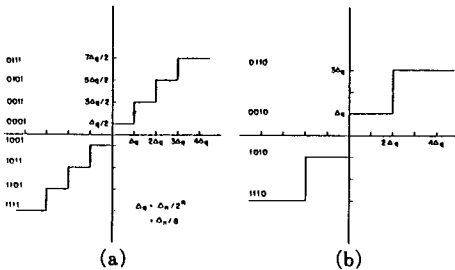


그림16. 양자화 잡음을 coding 하기 위한 EADPCM의 양자기
(a) Step size Δ_q 를 가진 3-bit midriser quantizer
(b) (a)에 해당하는 2-bit mid-riser quantizer

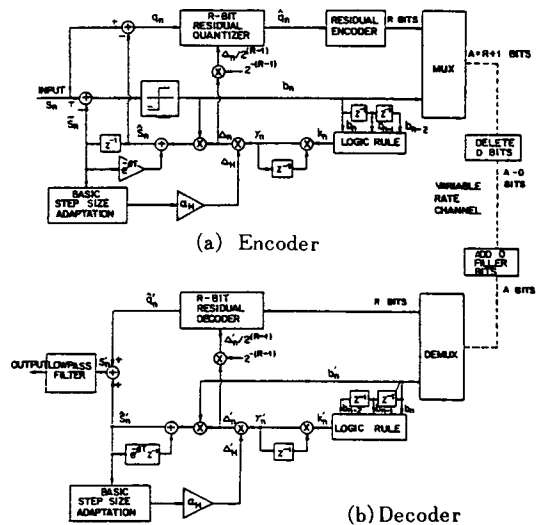


그림18. EHCDM 시스템의 블록도

2) Embedded HCDM System

EHCDM의 기본적인 coding 개념은 EADPCM과 같다. 즉 입력신호가 들어오면 전술한 바와 같이 HCDM으로 coding이 되고 양자화로 생기는 error(또는 residual) signal을 추가로 coding하여 두 신호를 다중화하여 channel로 보낸다. 이 시스템의 block diagram이 그림18에 그려져 있다. 여기서 residual encoder는 EADPCM과 같이 PCM이 사용된다. Channel에서의 bit stripping과 insertion은 EADPCM과 같이 한다. EHCDM의 성능이 그림19에 나타나 있다.

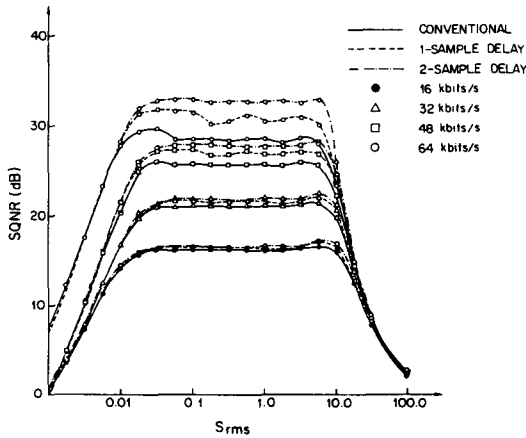


그림19. 16~64kbits/s 전송속도에서의 EHCDCM의 성능

VI. 부수되는 문제점들의 고찰

여기서는 predictive coding 시스템의 여러가지 문제점들 중에서 noisy channel에서의 bit errors가 decoding에 미치는 영향과, 현재 많이 쓰여지고 있는 PCM과 tandem으로 연결될 때의 성능저하에 관하여 기술한다. Noisy channel에서의 성능저하는 ADM보다 ADPCM이 훨씬 더 심각하므로 ADPCM에 관해서만 고려한다.

1. ADPCM의 Bit Error 효과와 Noisy Channel에서의 성능 개선

PCM이나 ADPCM은 noisy channel에 의해서 bit error가 생길 경우 양자기의 level뿐만 아니라 step의 크기도 달라지기 때문에 bit error에 관해서 성능이 아주 민감하다. 더우기 이 error 효과는 일정한 시간 동안 propagate이 되기 때문에 error rate이 10^{-3} 정도가 되면 음질이 아주 나빠진다.^[12]

ADPCM에서 이러한 error 효과를 줄이기 위해서 그동안 몇가지의 방법들이 제안되었다. 그 예로서는 Goodman과 Wilkinson은 error effect의 propagation을 줄이기 위해서 양자기의 step size factor가 sampling 시간에 따라 exponential decay를 할 수 있도록 설계할 것을 제안하였고,^[13] Jayant는 feedforward adaptation 방식(DPCM AQF)으로 양자기의 step size를 따로 coding하여 보내는 방법을 제시하였다.^[14] 이 두 방법들을 사용할 때 channel error가 생길때 성능저하의 효과는 개선할 수 있으나 전자의 경우는 codec의 dynamic range가 줄어들고 후자는 SNR이 ideal한 channel에서의 ADPCM보다 감소되는 단점이 있다.

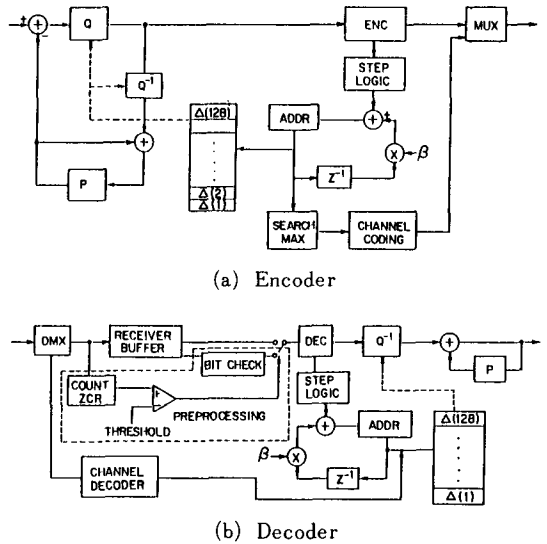


그림20. Channel error에 강한 새로운 ADPCM 시스템

이러한 문제점을 보완한 channel error에 강한 새로운 ADPCM 시스템이 근래에 제안되었다.^[12] 이 시스템의 block diagram이 그림20에 그려져 있다. 이 ADPCM의 step size adaptation logic은 Bell ADPCM과 같고, maximum step size를 따로 coding하여 ADPCM의 bit stream과 다중화하여 송신하므로서 channel error의 효과를 최소화 하도록 하였다. 또한 decoder에서는 유성음의 MSB가 error가 날 경우 그 error를 detect하여 correction까지 할 수 있도록 하였다. 이 시스템은 Bell ADPCM과 비교하여 좀 복잡한 것이 단점이나 ideal channel에서는 성능이 같고 noisy channel에서는 3~5dB 정도 좋은것으로 나타났

다. Noisy channel에서의 제안된 시스템과, Bell ADPCM 그리고 Jayant의 DPCM-AQF의 성능비교를 그림21에 도시하였다.

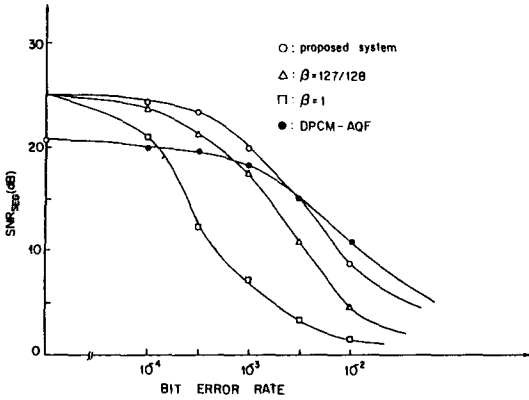


그림21. 제안된 ADPCM, BELL ADPCM, Jayant의 DPCM-AQF와의 성능비교

2. ADPCM과 PCM, Tandem 효과

앞서 언급한 바와 같이 예측부호화기, 특히 ADPCM이 32kbits/s의 codec으로서 PCM과 같이 쓰게됨에 따라 coding방식이 다른, 두개의 bit stream의 digital domain에서의 code 변환문제, 이에 따른 성능저하 문제등에 관심을 갖게되었다. Digital domain에서 code 변환문제는 ADPCM과 PCM뿐만 아니라 ADM과 PCM 사이의 code 변환도 그동안 연구되었다.¹¹⁵⁾ ADPCM과 PCM은 둘다 입력신호를 Nyquist Sampling을 하기 때문에 conversion이 간단하나, ADM은 over-sampling을 하는 이유로 PCM이나 ADPCM과 tandem을 하게 될 경우 conversion이 비교적 복잡하다.

ADPCM과 PCM을 tandem할 경우 생기는 성능저하 문제에 관하여 최근에 이론적으로 연구되었다.¹¹⁶⁾ 이 연구결과에 의하면 single tandem(PCM-ADPCM-PCM 또는 ADPCM-PCM-ADPCM)의 경우 transcoding 그 자체에 의한 성능저하는 무시할 정도이나 PCM의 전송속도는 64kbits/s이고 ADPCM은 32kbits/s 때문에 생기는 전송속도 감소에 의한 음질저하가 문제점으로 나타났다. 따라서 ACM/ADPCM transcoding 시스템의 성능향상을 위해서는 tandem시키는 ADPCM의 성능을 최대한도로 향상시키는 일이 중요하다. PCM과 ADPCM을 double, triple tandem을 시킬 경우에는 양자화 잡음이 누적됨으로서 code conversion의 효과가 점차로 나타나 triple tandem은 single tandem

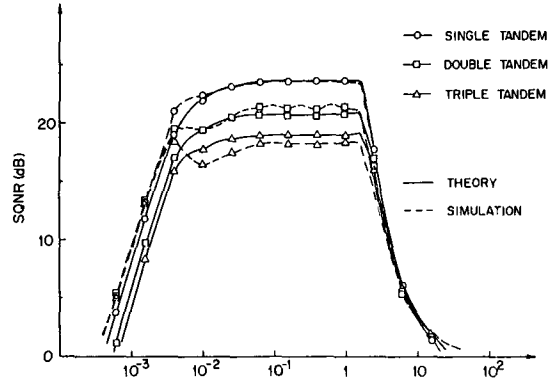


그림22. PCM과 ADPCM이 tandem 되었을때의 성능

보다 약 5dB 정도 성능이 저하되게 된다. 이 결과가 그림22에 나타나 있다.

Ⅶ. 결 론

지금까지 음성신호의 digital coding 방법중 대역폭 축소에 효과적이고 예측부호화 방법의 대표적인 ADPCM과 ADM에 관하여 자세히 논하였다. ADPCM과 ADM system은 16~32kbits/s의 전송속도에서 가장 효과적이고 system의 복잡도면에서도 아주 간단하여 앞으로 많이 실용화 될것이 기대된다. 특히 ADPCM은 32 kbits/s에서 표준화가 되었기 때문에 앞으로의 사용이 급격히 늘어날 전망이며, ADM은 codec이 극히 간단하고 noisy channel에서 성능이 어느 system보다 월등하기 때문에 mobile telephone 등 특수음성통신이 많이 사용될 것이다.

參 考 文 獻

- [1] C. K. Un, "Trends in digital speech communication technologies," *Review of KIEE*, vol. 11, pp. 30-40, Dec., 1984.
- [2] J. L. Flanagan, et al., "Speech coding" *IEEE Trans. Commun.*, COM-27, pp. 710-736, 1979.
- [3] J. D. Gibson, "Adaptive prediction in speech differential encoding systems," *Proc. of IEEE*, vol. 68, pp. 488-527, April, 1980.
- [4] CCITT study group XVIII, "Report on 32 kb/s ADPCM," CCITT document, Nov., 1983.
- [5] P. Cummiskey, et al., "Adaptive quantiza-

- tion in differential PCM coding of speech," *Bell Sys. Tech. J.*, vol. 52, pp. 1105-1108, Sept., 1973.
- [6] C. K. Un and H. S. Lee, "A study of the comparative performance of adaptive delta modulation systems," *IEEE Trans. on Commun.*, vol. COM-28, pp. 96-101, Jan., 1980.
- [7] C. K. Un, H. S. Lee and J. S. Song, "Hybrid companding delta modulation," *IEEE Trans. on Commun.*, vol. COM-29, pp. 1337-1344, Sept., 1981.
- [8] J. J. Dubnowski and R. E. Crochiere, "Variable rate coding of speech," *Bell Syst. Tech. J.*, vol. 58, pp. 577-600, Mar., 1979.
- [9] R. V. Cox and R. E. Crochiere, "Multiple user variable rate coding for TASI and packet transmission systems," *IEEE Trans. Commun.*, vol. COM-28, pp. 334-344, Mar. 1980.
- [10] C. K. Un and D. H. Cho, "Hybrid companding delta modulation with variable-rate sampling" *IEEE Trans. on Commun.*, vol. COM-30, pp. 593-599, April, 1982.
- [11] H. H. Lee and C. K. Un, "Embedded waveform coding of speech" *J. of KIEE*, vol. 21, No. 3, pp. 73-83, May, 1984.
- [12] H. D. Kim and C. K. Un, "An ADPCM system with improved error control," in Proc. of 1983 IEEE Global Commun. Conf., pp. 40.2.1-40.2.5, Dec., 1983.
- [13] D. J. Goodman and R. M. Wilkinson, "A robust adaptive quantizer," *IEEE Trans. on Commun.*, vol. COM-23, Nov., 1975.
- [14] N. J. Jayant, "Step-size transmitting differential coders for mobile telephony," *Bell Sys. Tech. J.*, vol. 54, pp. 1557-1581, Nov., 1975.
- [15] H. Jeong and C. K. Un, "A PCM/ADM and ADM/PCM code converter," *IEEE Trans. on Acow., Speech and Signal Processing*, vol. ASSP-27, pp. 762-768, Dec., 1979.
- [16] J. I. Lee and C. K. Un, "Performance analysis of PCM/ADPCM transcoding systems," *IEEE Trans. on Commun.*, vol. COM-33, Nov., 1985 (to appear). *

♣ 用 語 解 說 ♣

EBCDIC

Extended Binary Coded Decimal Interchange Code(확장된 이진화 십진 코드)의 약어. 8비트 코드로서 256개의 문자, 숫자, 특수 문자를 나타낸다.

CROM

Control Read Only Memory의 약어, 제어 논리를 해독하기 위해서 고안되고 마이크로 프로그램 기억용의 특수한 ROM. 여러 종류의 마이크로 프로세서의 주요 부분이다.

EPROM

시스템 프로그램을 개발하는데 빠른 회수 시간이 중요한 고성능 마이크로 컴퓨터 시스템에 적합한(전기적으로) 프로그램 할 수 있는 ROM을 말함.

PROM

Programmable Read Only Memory의 약어. 일반적으로 프로그램하기 위한 특별한 물리적 작용을 하지 않는 한 기록할 수 없는 기억 매체. 특별한 물리적 과정을 거쳐서 기억 내용을 지우고 다시 프로그램할 수 있는 PROM도 있다.

ROM

판독 전용 기억 장치(read-only-memory)로서 여러 종류가 있다. 제조 과정에서 프로그램되는 ROM 이 가장 기본적인 마스크 ROM이며, 사용자가 한 번 프로그램하는 PROM, 그리고 특수한 빛이나 전기적인 신호로써 내용을 지우고 다시 프로그램할 수 있는 EPROM이 있다. 정보는 영구히 보관되거나 반영구적으로 보관되며, 자료를 읽어내지만 동작 중 기억된 정보의 내용은 변환되지 않음.