

DPCM 음성 부호화기의 부정합현상에 관한 연구

(On the Mismatch Phenomena in DPCM Coding of Speech)

柳得洙*, 趙東浩**, 殷鍾官**

(Deuk Su Yoo, Dong Ho Cho and Chong Kwan Un)

要 約

본 논문에서는 임의의 입력신호와 DPCM 음성 부호화기의 양자기 및 예측기 사이에서 발생하는 부정합 문제를 probability density function(pdf) 모양의 부정합, 분산의 부정합 및 상관의 부정합 등으로 대별하여 그 현상을 규명하였다. 32kbps 이상의 높은 전송률에서는 시스템 입력과 양자기간의 pdf 모양의 정합에 의해 성능이 개선되었으나, 낮은 전송률에서는 pdf 정합효과가 별로 나타나지 않았다. 또한 γ 양자기가 다른 양자기보다 pdf 모양과 분산의 변화에 대해서 비교적 덜 민감함을 알았다. 분산의 경우에는 전송속도가 높을수록 입력의 분산이 클 때가 작을 때보다 과부하 잡음양이 많아지기 때문에 부호화기의 성능이 떨어졌다. 결국 DPCM 부호화기에서는 양자기가 입력의 pdf 모양과 분산에 잘 따라가도록 하는 것이 예측기가 입력의 상관에 잘 따라가도록 하는 것보다 종래의 신호대잡음비(SNR) 척도로 볼 때 더 좋은 것으로 나타났다.

Abstract

This paper describes various mismatch phenomena in differential pulse code modulation (DPCM) coding, such as the mismatch effects of probability density function (pdf), signal variance, and correlation. At a high transmission rate (i.e., above 32 kbits/s), the performance of DPCM can be improved by matching the pdf shape between the input signal and the quantizer. However, the same gain cannot be obtained at a lower transmission rate. Also, it is shown that the γ quantizer is relatively robust to the variation of pdf shape and signal variance. Moreover, as the transmission rate increases, the performance of DPCM for the input signal with large variance is worse than that of DPCM for the signal with small variance due to the increase of overload noise. According to our simulation results, the mismatch effects of pdf shape and variance appear to yield more degradation than that of correlation in a DPCM system.

I. 序 論

DPCM은 크게 양자기와 예측기의 두개의 subsystem

으로 구성된다.^[1] 이 DPCM에도 PCM과 같이 고정시스템과 적응시스템이 있다.^[2] 비적용 DPCM은 시스템 입력이 stationary 성질을 갖고 있고, 그 통계적 성질을 미리 알 때, 입력신호의 redundancy를 효과적으로 제거할 수 있으나, 음성같이 nonstationary 성질을 갖고 있는 입력에 대해서는 시간에 따라 pitch 주기와 성도특성에 따른 신호의 진폭크기가 크게 변하기 때문에 그렇게 할 수 없다.^[3] 그와같은 nonsta-

*正會員, 韓國電子通信研究所
(Electronics and Telecoms. Research Institute)

**正會員, 韓國科學技術院 通信工學研究室
(Commun. Research Lab., KAIST)

接受日字: 1985年 12月 20日

tionary한 입력에 대해서는 양자기나 예측기 또는 두 subsystem을 모두 적용시키는 adaptive DPCM (ADPCM)을 사용함으로써 상당한 개선을 가져올 수 있었다.^{12,41}

그런데, 실제 상황에서는 입력음성의 통계적 성질이 시시각각 변화하므로 아무리 잘 설계된 시스템이라 할지라도 그 성능이 예측된 성능보다 상당히 떨어지게 된다. 이것은 입력의 통계적 특성과 두 subsystem인 양자기와 예측기사이에서 부정합 현상이 발생하기 때문이다. 따라서 음성의 pdf 및 분산 특성과 DPCM 부호화기내의 양자기의 pdf모형과 분산 특성을 정합시키고, 아울러 입력음성의 상관특성과 DPCM 부호화기내의 예측기의 상관특성을 정합 시킴으로써 DPCM의 성능을 개선하는 것이 필요하다.

일찌기 Mauersberger는 memoryless 양자기와 그 입력사이의 부정합, 즉 pdf모양의 부정합과 분산의 부정합 등에 관해서 연구했다.¹⁵¹ 판단척도로는 mean-squared error (MSE)를 사용했는데, 입력으로는 임의의 pdf모양과 분산값을 가진 sequence를 사용하여 simulation은 하지 않고 다만 memoryless 양자기에 대해서 pdf모양 및 분산의 부정합을 이론적으로만 고찰하였다. 이밖에 DPCM에서 입력과 예측기간의 상관의 정합 및 부정합에 관한 연구가 수행되었으며 예측기와 시스템입력간의 상관 부정합에 의해서도 시스템의 성능이 떨어짐이 밝혀졌다.^{16,71}

이와 같은 입력과 시스템간의 통계적 특성에 있어서의 부정합에 의한 성능저하의 구체적인 현상을 규명하여, 그 결과를 적절히 이용하면 음성의 특성에 정합된 DPCM 시스템을 구성할 수 있을 것이다. 또한 이 결과는 적응시스템에서 가능한 한 성능저하에 민감한 시스템 변수들을 먼저 신호의 특성에 따라 조절하고, 변수들을 고정시켜야 하는 경우에는 덜 민감한 시스템이 되도록 선정하여 ADPCM과 같은 적응시스템의 성능을 개선하는데 사용될 수 있다.

본 논문에서는 서론에 이어 II장에서는 음성 부호화기의 부정합 현상을 정의하고, 임의의 입력 sequence를 발생시키는 방법을 기술한다. 제III장에서 컴퓨터 simulation을 통해서 부정합 현상을 규명하며, 마지막으로 제IV장에서 결론을 맺는다.

II. 음성부호화기의 부정합 현상

1. 부정합 현상의 정의

DPCM 부호화기의 통계적인 부정합들을 정의하면 다음과 같다.

1) 확률밀도함수 모양의 부정합

양자기 모형과 양자기 입력의 pdf모양이 서로 다른 때를 말한다.

2) 분산의 부정합

양자기 모형의 pdf 분산과 그 입력의 분산이 서로 다른 경우를 일컫는다.

3) 상관의 부정합

예측기의 계수값들과 시스템 입력의 상관이 정확하게 대응되지 않을 때를 말한다.

본 논문에서는 음성 부호화기의 양자기를 gamma, Laplacian, Gaussian 및 uniform모형으로 modeling하고, 음성부호화기의 입력으로서 gamma, Laplacian, Gaussian 분포 등을 갖는 sequence들을 만들어서 사용하였다.

2. 임의의 확률밀도함수와 상관을 갖는 source 신호의 발생

일반적으로 임의의 확률밀도함수를 나타낼 수 있는 generalized Gaussian pdf는 다음식과 같이 표시된다.⁸¹

$$f(\epsilon; \sigma, c) = \frac{c \cdot \eta(\sigma, c)}{2 \cdot \Gamma(\frac{1}{c})} \exp[-|\eta(\sigma, c) | \epsilon|^c] \tag{1-a}$$

$$\eta(\sigma, c) = \frac{1}{\sigma} \left\{ \frac{\Gamma(\frac{3}{c})}{\Gamma(\frac{1}{c})} \right\}^{\frac{1}{2}} \tag{1-b}$$

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt \tag{1-c}$$

이다. 이 pdf는 그림 1에서 보듯이 σ^2 값은 그대로 둔 채 c값에 의하여 모양을 바꿀 수 있으며 c값이 1, 2 일 때 각각 Laplacian 및 Gaussian 분포를 나타낸다.

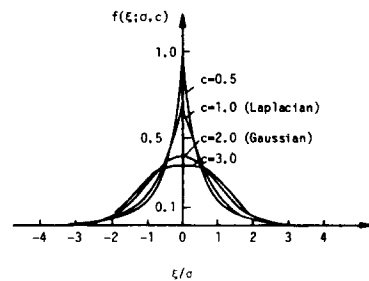


그림 1. 일반화된 Gaussian 확률밀도 함수
Fig. 1. Generalized Gaussian probability density.

이 c값이 양수이면서 작을수록 $\epsilon = 0$ 부근이 뾰족한 모양을 나타내며 c값이 무한히 커질 때는 uniform distribution이 된다. 한편 gamma분포는

$$f(\epsilon) = \frac{1}{2} \left(\frac{\sqrt{3}}{2\pi\sigma} \right) \exp\left(-\frac{\sqrt{3}|\epsilon|}{2\sigma} \right) \quad (2)$$

으로 표시되기 때문에 일반적인 Gaussian 분포로 표시할 수 없으나 $c \approx 0.75$ 일 때 근사한 gamma 분포를 얻을 수 있다.

인접 표본값사이에 상관관을 부여하는 방법에는 여러가지가 있으나 여기서는 분석의 용이성을 위해서 moving average (MA) process만 고려한다. MA process는 all-zero process라 하는데, M차의 all-zero process는 power spectral density (PSD), 함수가 σ_x^2 인 white noise sequence $\{z(n)\}$ 을 전달함수가 $A(z) = \sum_{j=0}^M a_j z^{-j}$ 인 그림 2와 같은 여파기에 통과시킴으로써

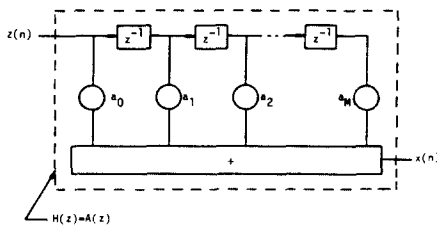


그림 2. MA(M) process의 발생
Fig. 2. Generation of MA(M) process.

발생시킬 수 있다. 이때 발생된 process $\{x(n)\}$ 의 PSD function은

$$S_x(z) = \sigma_x^2 A(z) A(z^{-1}) \quad (3)$$

이고, 이와같은 MA(M) process는

$$x(n) = \sum_{j=0}^M a_j z(n-j) \quad (4)$$

에 의해서 발생된다고 할 수 있다. 이 all-zero 여파기의 impulse 응답은

$$h(n) = \begin{cases} a_n & \text{for } n=0, 1, 2, \dots, M \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

이므로 이 MA process의 autocorrelation function (acf)는

$$R_x(k) = \begin{cases} \sigma_x^2 \sum_{j=0}^{M-|k|} a_j a_{j+k} & \text{for } |k| \leq M \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

가 된다. 이 여파기가 안정하면 $\{x(n)\}$ 은 stationary 하다. 다음에는 실제로 사용될 1차 all-zero process에 대해서 고찰하면 $a_0 = 1$ 이고 평균이 0인 MA process는

$$x(n) = z(n) + a_1 z(n-1) \quad (7)$$

에 의하여 발생된다. 이것의 분산과 normalized acf는 다음과 같다.

$$\sigma_x^2 = R_x(0) = \sigma_z^2 (1 + a_1^2) \quad (8)$$

$$\rho_x(k) = \begin{cases} \frac{a_1}{1 + a_1^2} & \text{for } k=1 \\ 0 & \text{for } |k| > 1 \end{cases} \quad (9)$$

(9)식으로부터 $\rho_x(1)$ 값은 $-0.5 \leq \rho_x(1) \leq 0.5$ (등호는 $a_1 = \pm 1$ 일 때 성립)임을 알 수 있다.

III. 컴퓨터 Simulation결과와 고찰

본 논문에서 사용한 시스템 입력은 gamma, Laplacian, Gaussian 분포의 real number sequence와 실제 음성이었다. 실제 음성 sequence는 3400Hz의 대역폭을 갖는 음성을 8000Hz로 sampling한 것이다. 또 각 확률분포를 갖는 sequence는 IMSL package의 program을 이용하여 만든 뒤에 인접한 표본값사이에 상관관을 갖도록 1차의 all-zero 여파기를 거치게 했다. 이와 같이하여 발생시킨 real number sequence와 실제 음성의 통계적인 특성을 표 1과 그림 3에 나타내

표 1. Simulation에 사용된 입력 신호의 통계적 특성
Table 1. Statistical characteristics of input signal used for simulation.

Input Sequence	gamma	Laplacian	Gaussian	Real Speech
Mean	0.00	0.00	0.00	0.00
Variance	1.00	1.00	1.00	1.00
Normalized Auto-Correlation Function				
$\rho(1)$	+0.6324	+0.4981	+0.4946	0.9434
$\rho(2)$	-0.1302	-0.0005	+0.0063	0.8710

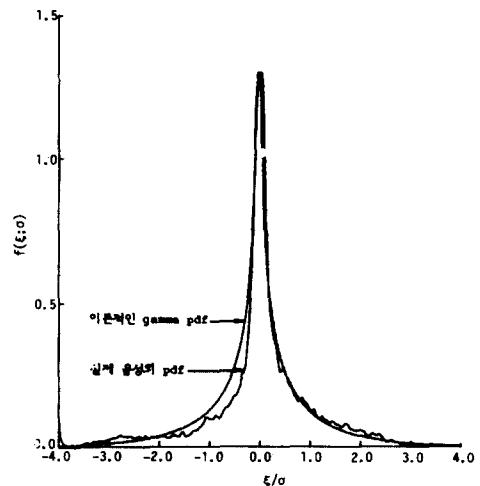


그림 3. 실제음성의 pdf와 이론적 gamma pdf($\mu=0, \sigma^2=1$)
Fig. 3. pdf's of gamma and real speech.

었다. 그런데 1 차의 all-zero process는 식 (7), (9) 에서 a_n 을 1로 했기 때문에 표 1에서 Laplacian sequence와 Gaussian sequence의 normalized acf값이 (9)식으로 구한 이론값과 일치하지만, gamma sequence의 경우에는 그렇지 않다. 이것은 1 차 all-zero process 입력의 pdf가 gamma 모양을 갖더라도 그 출력은 pdf가 달라지고 상관값이 작아져서 여러번의 시행착오 과정을 거쳤기 때문이다. 실제로 gamma sequence를 만드는 과정은 다음과 같다. 먼저 입력 gamma sequence를 부호는 그대로 두고 제곱한 뒤에 인접 표본값사이의 차를 구했다. 그 후에 인접 표본값사이의 차이를 3 제곱한 뒤에 그 sequence의 차를 구해서 상관이 0.63인 gamma분포의 sequence를 만들었다. 위의 gamma sequence를 만드는 과정을 식으로 나타내면 다음과 같다.

$$g_n^c = (g_{n+1} | g_{n+1} | - g_n | g_n |)^3 - (g_n | g_n | - g_{n-1} | g_{n-1} |)^3 \quad (10)$$

여기서 g_n^c 는 상관이 0.63인 gamma분포 sequence의 n번째 값이고, g_n 는 상관이 0인 원래 gamma 분포 sequence의 n번째 값이다. 위과정에서 제곱, 3제곱 등은 pdf가 gamma분포가 되게 하기 위하여 사용했고, 차를 이용한 것은 상관값을 증가시키기 위해서이다. 한편 그림 3에서는 실제음성의 pdf가 gamma pdf와 가장 유사함을 알 수 있다.

Simulation을 위해서 선택한 음성부호화기는 예측기의 기억소자가 하나인 그림 4의 1 차 DPCM이다.

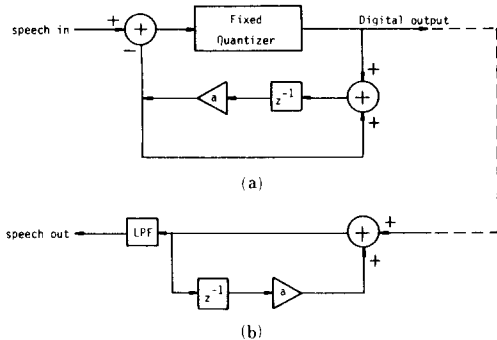


그림 4. DPCM 부호화기의 블록도
(a) 부호화기 (b) 복호화기

Fig. 4. Block diagram of DPCM coder.
(a) Encoder. (b) Decoder.

양자기의 모형을 nonuniform model인 gamma, Laplacian, Gaussian 모형과 uniform model인 optimum uniform gamma, Laplacian, Gaussian모형 등으로 구분하여 사용하였다.^{(8),(10)} 음질 평가척도로는 종래의

SNR척도를 사용하였으며, 본 simulation에서는 16, 32 kbits/s에서 1 차 DPCM의 pdf모양 부정합, 분산 부정합 및 상관 부정합 현상을 살펴보았다.

1. 확률밀도함수 모양의 부정합

어떤 pdf모양의 시스템의 입력에 대해서 일치되는 양자기의 모형을 찾이란 상당히 어려운 일이다. 왜냐하면 입력의 pdf 모양과 residual신호의 pdf 모양은 전송속도와 DPCM의 특성에 따라 약간 다르기 때문이다. 그러나 본 논문에서는 시스템 입력과 residual 신호의 pdf모양이 일치한다고 가정하고(simulation에 의하면 크게 차이나지 않았음), 시스템 입력의 pdf 모양을 양자기의 모형으로 하고 분산을 맞추기 위해 양자기 level의 최적 승수값을 구했다. 예측기 계수값도 최대 SNR값을 얻는 조건에서 찾았으며 시스템 입력의 분산은 1로 하였다.

한 시스템 입력에 대해서 양자기의 pdf모양을 바꿔가며 구한 SNR값들이 전송률에 따라 표 2에 나타나 있다. 이 결과를 보면 변수 c값이 편중되어 있어서 pdf모양의 변화에 따른 시스템 성능의 향상 또는 저하정향을 잘 알 수 없다. 그러나, 입력이 Gaussian분포의 sequence인 경우에는 양자기 모형의 변수 c가 그보다 작을 때(즉 Gaussian모양보다 뾰족한 모양의

표 2. First-order DPCM에서 확률밀도 함수 shape의 부정합 효과(단위는 dB임)

Table 2. Mismatch effect of probability density function in a first-order DPCM system (unit:decibels).

(a) 32kbits/s일 경우.
(a) At 32kbits/s.

Quantizer model (shape parameter, c)	Gamma (c≈0.75)	Laplacian (c=1.0)	Gaussian (c=2.0)	Uniform (c→∞)
Input sequence				
Real speech	24.24	23.09	21.20	19.59
Gaussian	20.02	20.76	21.46	20.66
Gamma	19.89	19.91	18.87	17.17
Laplacian	18.80	19.10	18.34	16.80

(b) 16kbits/s일 경우
(b) At 16 kbits/s.

Quantizer model (shape parameter, c)	Gamma (c≈0.75)	Laplacian (c=1.0)	Gaussian (c=2.0)	Uniform (c→∞)
Input sequence				
Real speech	13.57	13.75	13.21	12.90
Gaussian	9.35	9.97	10.25	10.17
Gamma	8.47	8.70	8.17	7.86
Laplacian	7.87	8.40	8.27	8.08

pdf모형 양자기를 사용할 때)가 클 때 보다 시스템 성능을 더욱 크게 떨어뜨리는 것을 볼 수 있다. 이 결과는 많은 입력 pdf모양에 대해서 이론적으로 MSE를 구해 이 현상을 연구한 Mauersberger의 결과에서도 마찬가지로 관찰되었다.

한편 전송률에 따른 pdf모양의 부정합에 의한 시스템 성능의 변화는 다른 부정합 효과와 관련되어 있으므로 나중에 종합적으로 논하기로 한다.

2. 분산의 부정합

분산의 부정합 추세는 시스템 입력과 양자기 모형의 모양이 같은 상황에서 알 수 있다. 여기서는 정합조건(한 시스템 입력의 분산을 1로 하고 최대 SNR을 얻는 최적 송수값과 예측기 계수값을 사용한 상태)에서 한 시스템 입력에 대한 분산을 0.2에서 4.0까지 변화시키면서 양자기 모형을 바꿔가며 시스템의 SNR값을 구했으며 그 결과가 그림 5와 그림 6에 도시되어 있다. 단, 여기서는 실제 음성의 pdf가 그림 3에서 알 수 있듯이 gamma 및 Laplacian pdf에 가까우므로 Gaussian 분포를 갖는 입력의 분산의 변화에 따른 영향은 고찰하지 않았으며 실제 음성의 분산의 부정합 현상은 4절에서 종합적으로 고찰한다.

전송속도가 높을수록 시스템 입력의 분산이 클 때가 같은 비율로 작을 때보다(예를 들어 입력의 분산이 4일 때와 0.25일 때의 SNR값을 비교해 보면 알 수 있음)시스템의 성능이 떨어진다. 이런 이유는 다음과 같이 말할 수 있다. 입력의 분산이 커지면, 과부하 잡음양이 많아지는데, 이 과부하 잡음성분의 크기는 양자기의 level값으로 제한될 수 없는 반면에, 분산이 작아지면 과부하 잡음양은 작아지고 granular noise

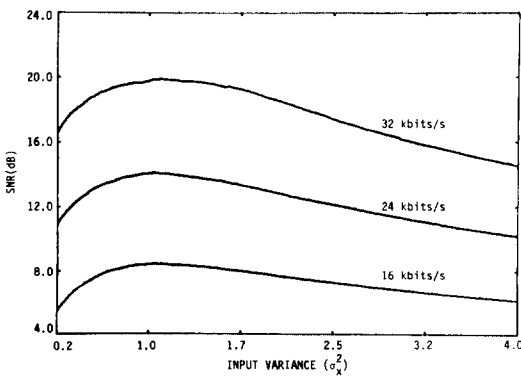


그림 5. Gamma입력의 분산의 변화에 따른 gamma 양자기를 사용한 DPCM의 성능

Fig. 5. Performance of DPCM system with gamma quantizer vs. variance of gamma-distributed system input.

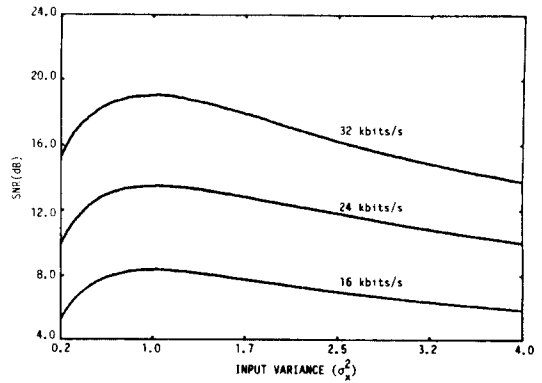


그림 6. Laplacian 입력의 분산의 변화에 따른 Laplacian 양자기를 사용한 DPCM의 성능

Fig. 6. Performance of DPCM system with Laplacian quantizer vs. variance of Laplacian distributed system input.

양이 상대적으로 많아지는데 이 granular noise 성분의 크기는 양자기의 level값에 의해 제한될 수 있다.

3. 상관의 부정합

정합조건에서 시스템 입력의 상관과 예측기의 계수값이 현저하게 다른 경우에 시스템의 성능변화를 관찰하는 것은 음성부호화기에서 큰 의의가 없으므로 본 연구에서는 실제음성을 입력으로 하고 예측기 계수의 갯수를 증가시켜 실제음성의 상관과 예측기의 상관이 정합상태에 도달함에 따라 시스템의 성능이 어떻게 변화하는가를 관찰하였다.

Simulation에서는 양자기는 고정시키고 예측기를 적응시키는 적응예측기(adaptive predictor)를 갖는 DPCM(DPCM-AP)를 사용하였으며, 적응방법에는 forward방식(DPCM-APF)과 backward방식(DPCM-APB) 두가지 방법을 택하였다. DPCM-APF는 10ms(80표본값)마다 예측기 계수를 갱신하여 수신단에 전송하는데, 부호화되기 직전의 80표본값에 rectangular window를 씌운 후에 autocorrelation방법을 사용하여 예측기의 계수값을 구했다. Backward 적응 예측은 gradient적응방법을 사용하였는데 이득을 0.77로 하였다. 이렇게 하여 얻어진 결과가 그림 7에 도시되어 있다.

DPCM-APF에 대한 simulation결과는 Noll의 연구결과와 거의 같은 경향을 보인다.¹⁾ 즉, 예측기 계수의 수가 많을수록 SNR 값이 증가하나 예측기 계수의 수가 10이상이면 정합상태가 더 이상 개선되지 않는다고 말할 수 있다.

DPCM-APB는 DPCM-APF와 달리 예측기 계수

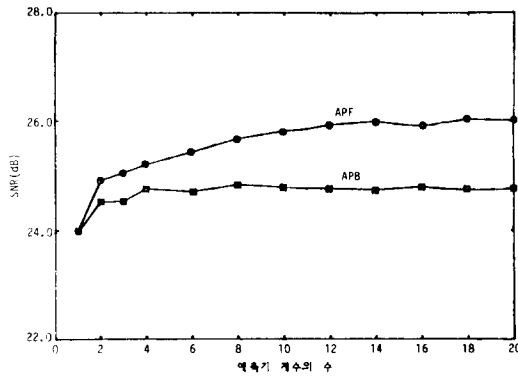


그림 7. 입력이 실제음성일 때 적응 예측기 계수의 수 변화에 따른 ADPCM의 성능
 Fig. 7. Performance of ADPCM system with an adaptive predictor vs. number of predictor coefficients when the input signal is real speech.

의 수가 4일 때 정합상태를 이룰을 알 수 있다. 왜냐하면 backward 방식인 경우에는 재생음성값에 양자화 잡음이 포함되어 있어서 예측계수의 수가 증가할수록 이것이 누적되어 예측된 값과 시스템 입력의 편차를 줄이지 못하기 때문이다. 한편 backward 방식에서는 예측기 계수의 수가 4 이상일 경우에는 정합상태가 더 이상 개선되지 않을 뿐만 아니라 APF의 경우에 비하여 SNR값도 약 1dB 만큼 작다.

4. 종합적인 부정합 현상에 대한 고찰

다음은 시스템 입력의 분산이 변할 때 pdf모형과 전송율과 관련하여 DPCM 시스템의 성능변화에 대해서 살펴본다.

입력이 gamma 분포를 가질 때는, 전송율이 높은 경우인 표 2-(a)와 그림 8에서 보듯이 gamma 양자기를 사용할 때와 Laplacian 양자기를 사용할 때의 최대 SNR 값이 거의 차이가 없음을 알 수 있다. 한편 전송율이 낮은 경우인 표 2-(b)와 그림 9에서 보면 Laplacian 양자기를 사용할 때의 최대 SNR 값이 다른 양자기를 사용할 때 보다 더 높음을 알 수 있다. 입력이 gamma 모양과 흡사한 실제 음성에 대해서는 낮은 전송률에서 Laplacian 양자기를 사용할 때의 성능이 다른 어떤 것 보다도 좋다. 이 사실들로부터 높은 전송률에서는 시스템 입력과 양자기간의 pdf모양의 정합에 의한 시스템 성능의 증가를 기대할 수 있으나, 낮은 전송률에서는 그와같은 pdf정합효과를 얻을 수 없음을 알 수 있다. 이런 이유는 다음과 같이 생각할 수 있다. 양자기의 입력인 residual 신호는 시스템 입력과 예측된 표본값과의 차이이며, 입력 sequence가 gamma 분포일 때 이들의 차는

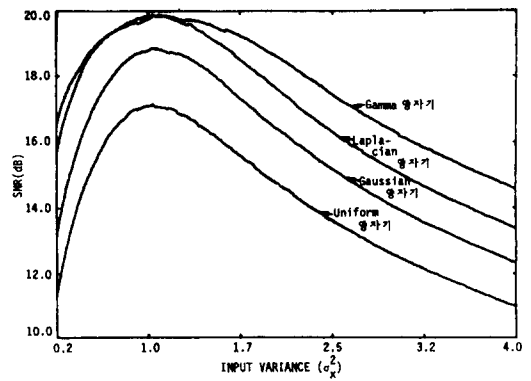


그림 8. 32kbts/s에서 gamma 입력의 분산을 변화시킬 때 여러가지 양자기의 성능
 Fig. 8. Performance of DPCM systems with various quantizers vs. input variance with gamma distribution at 32kbts/s.

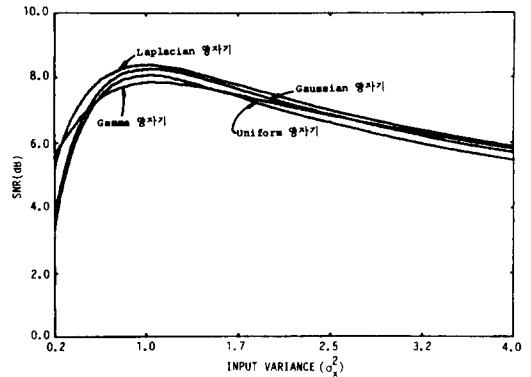


그림 9. 16kbts/s에서 gamma 입력의 분산을 변화시킬 때 여러가지 양자기의 성능
 Fig. 9. Performance of DPCM systems with various quantizers vs. input variance with gamma distribution at 16kbts/s.

모양이 gamma보다 약간 완만한 분포가 되지만, 표 2-(b)에서 보듯이 전송률이 낮아지면서 residual 신호는 Laplacian 분포에 가까와짐을 SNR값으로 미루어 알 수 있다. 또한 32kbts/s DPCM이 그보다 낮은 전송률의 DPCM보다 비교적 정교한 부호화기이어서 예측된 신호가 시스템 입력과 비교적 높은 상관을 갖기 때문에 residual 신호의 분산이 작고 그 pdf 모양은 gamma 분포에 더 가깝다. 한편 낮은 전송률에서는 예측된 신호가 시스템 입력과 비교적 낮은 상관을 갖고 있어서 residual 신호의 분산도 크고 그만큼 시스템의 최대 SNR값이 낮다. 이때 residual 신호는 gamma pdf보다 훨씬 완만한 모양(즉, Laplacian pdf의 모양에 가까

우나 그보다 뾰족한 분포)를 가지게 되므로 γ 양자기를 사용할 때보다 Laplacian 양자기를 사용할 때가 최대 SNR 값이 더 높다.

그리고 γ 외의 다른 pdf 모양의 입력에 대해서는 그림 10과 11에서 보듯이 입력 분산의 변화에 대해 γ 양자기를 사용할 때의 최대 SNR 값이 비록 크지 않더라도 SNR 값의 변화가 비교적 작다는 사실로부터 γ 양자기가 pdf 모양과 분산의 변화에 대해서 비교적 덜 민감함을 알 수 있다. 그 현상은 γ 양자기의 모양이 다른 것에 비해서 가장 뾰족하기 때문이다.

표 2에서, 인접한 표본값사이의 상관성이 큰 실제 음성

에 대하여 전송률이 낮아질수록 양자기 모형에 관계없이 다른 sequence를 입력으로 할 때보다 DPCM의 성능이 더 좋을 수 있다. 그 이유는 전송률이 낮아짐에 따라 인접한 표본값사이의 상관성이 클수록 선형예측 효과가 크게 나타나기 때문이다. 그림 12와 13에서 보듯이 실제 음성의 분산 변화에 따라서 양자기의 각각의 pdf 모양에 대한 SNR 값이 다른 입력에 비해 기복있게 변하는 것은, 다른 입력들은 stationary한데 비해 음성은 nonstationary하기 때문이라고 말할 수 있다. 특히 입력음성의 분산이 1보다 클 때 그런 현상이 현저한 것은 양자기의 level값으로 제한하지 못하는 residual 신호가 더 많이 만들어지기 때문이다.

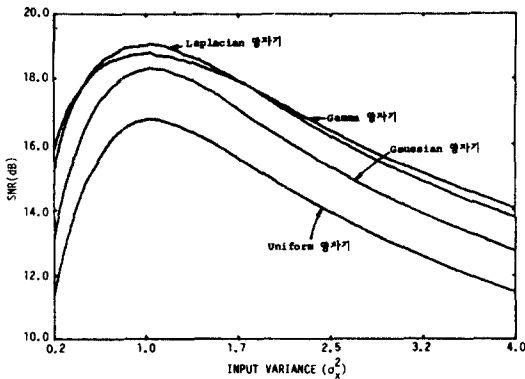


그림 10. 32kbits/s에서 Laplacian 입력의 분산을 변화시킬 때 여러가지 양자기의 성능

Fig. 10. Performance of DPCM systems with various quantizers vs. input variance with Laplacian distribution at 32 kbits/s.

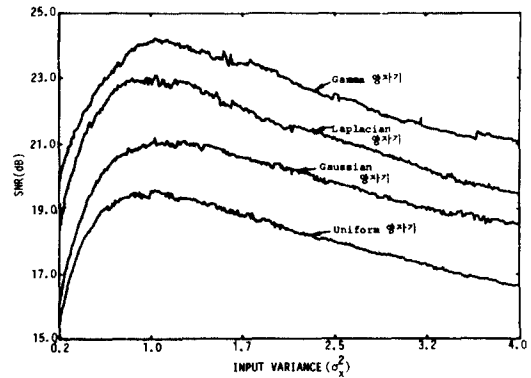


그림 12. 32kbits/s에서 실제음성 입력의 분산을 변화시킬 때 여러가지 양자기의 성능

Fig. 12. Performance of DPCM systems with various quantizers vs. variance of real speech input at 32 kbits/s.

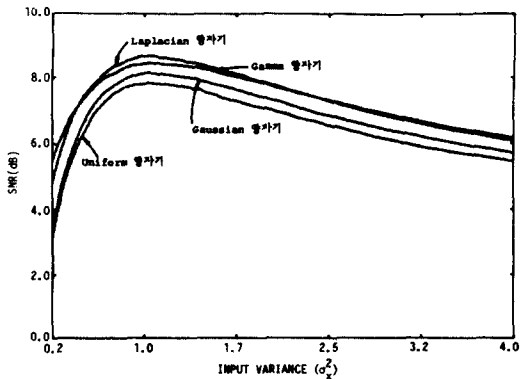


그림 11. 16kbits/s에서 Laplacian 입력의 분산을 변화시킬 때 여러가지 양자기의 성능

Fig. 11. Performance of DPCM systems with various quantizers vs. input variance with Laplacian distribution at 16 kbits/s.

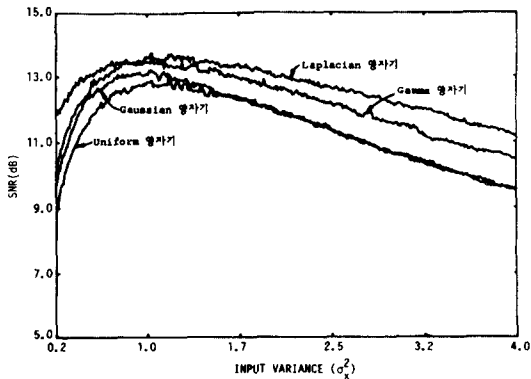


그림 13. 16kbits/s에서 실제음성 입력의 분산을 변화시킬 때 여러가지 양자기의 성능

Fig. 13. Performance of DPCM systems with various quantizers vs. variance of real speech input at 16 kbits/s.

IV. 結 論

본 논문에서는 입력신호와 DPCM 부호화기의 양자기 및 예측기 사이에서 발생하는 pdf모양의 부정합, 분산의 부정합, 및 상관의 부정합 현상 등을 규명하였다. pdf모양의 변화에 따른 시스템 성능의 향상 또는 저하 경향을 일반적으로 설명할 수는 없지만, Gaussian 분포의 sequence인 경우에는 Gaussian모양보다 뾰족한 모양의 pdf양자기를 사용할 때가 Gaussian 모양보다 완만한 형태의 pdf를 갖는 양자기를 사용할 때보다 시스템 성능을 더욱 크게 떨어뜨리는 것을 알 수 있었다. 또한 시스템 입력과 양자기간의 pdf모양의 정합에 의한 성능향상은 높은 전송률에서는 얻을 수 있지만 낮은 전송률에서는 얻을 수 없었으며, gamma양자기가 다른 양자기에 비해서 pdf 모양과 분산의 변화에 대해서 덜 민감함을 알았다. 분산의 경우에는 전송속도가 높을 수록 시스템 입력의 분산이 클 때가 같은 비율로 작을 때 보다 과부하 잡음 양이 많아지므로 시스템의 성능이 떨어졌다. 한편 입력과 DPCM 부호화기의 예측기사이의 상관관계에 있어서는 예측기의 계수의 개수가 증가함에 따라 성능이 향상되지만 forward예측기에서는 10이상이 되면 포화되고, backward 예측기에서는 4 이상이 되면 더 이상 성능이 개선되지 않음을 알았다. 이외의 본 논문의 결과로 DPCM 시스템에서는 양자기가 입력의 pdf 모양과 분산에 잘 따라가도록 하는 것이 예측기가 입력의 상관에 잘 따라가도록 하는 것보다, 종래의 SNR척도로 볼 때 더 좋은 것으로 나타났다.

參 考 文 獻

- [1] P. Cummsky, N.S. Jayant, and J.L. Flanagan, "Adaptive quantization in differential PCM Coding of Speech," *Bell Syst. Tech. J.*, vol. 52, pp. 1105-1118, Sept. 1973.
- [2] P. Noll, "A comparative study of various quantization schemes for speech encod-

ing," *Bell Syst. Tech. J.*, vol. 54, pp. 1597-1614, Nov. 1975.

- [3] J. Markhoul, "Linear prediction: a tutorial review," *Proc. IEEE* vol. 63, pp. 561-580, Apr. 1975.
- [4] J.D. Gibson, "Adaptive prediction in speech differential encoding systems," *Proc. IEEE*, vol. 68, pp. 488-525, Apr., 1980.
- [5] W. Mauersberger, "Experimental results on the performance of mismatched quantizers," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 381-386, July, 1979.
- [6] D.K. Purcell and J.D. Gibson, "Theoretical and experimental SNR of DPCM with prediction mismatch," in *Proc. IEEE Int. Conf. Commun.*, pp. 4G.2.1-4G.2.5, June 1982.
- [7] M.L. Sethia and J.B. Anderson, "Effects of mismatch and low rate for coding of autoregressive data by tree searched DPCM," in *Proc. IEEE Int. Conf. Commun.*, pp. 4H.2.1-4H.2.5., June 1982.
- [8] 김남철, "일반화된 공분산 모델에 근거한 영상 데이터 감축에 관한 연구" 한국과학기술원, 박사학위 논문, 1984년 2월.
- [9] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 7-12, Mar. 1960.
- [10] M.D. Paez and T.H. Glisson, "Minimum mean-squared error quantization in speech PCM and DPCM systems," *IEEE Trans. Commun.*, vol. COM-20, pp. 225-230, Apr. 1972.

謝 辭

본 논문의 개선을 위해서 많은 조언을 주신 심사위원님께 감사드립니다.