

# 배경 잡음환경에서 가변 임계값에 의한 Dual Rate ADPCM 음성 부호화 기법

(Coding Method of Variable Threshold Dual Rate ADPCM Speech  
Considering the Background Noise)

한경호\*

(Kyong - Ho Han)

## 요 약

본 논문에서는 ITU G.726 규격을 만족하는 표준형 ADPCM 부호화 법을 이용하여 배경잡음의 크기에 따라 음성  
의 부호화율이 두가지로 가변하도록 함으로써, 낮은 데이터 전송률을 가지고도 단일 부호화율의 경우보다 개선된  
음질을 갖는 부호화 기법을 제안하였다. 이를 위하여 배경잡음보다 큰 음성신호에 대하여는 데이터의 양이 커지더  
라도 음질을 향상시키기 위하여 40[Kbps]로 압축하고, 작은 음성신호에 대하여는 16[Kbps]로 압축하여 데이터의  
양을 줄이도록 하여 전체적으로 압축데이터의 양을 줄이면서 음질을 개선하도록 하였다. 입력된 음성신호에 대하여  
두가지 압축율을 결정하기 위하여 영교차율(ZCR)을 사용하여 처리속도를 빠르도록 하였다.

## Abstract

In this papser, we proposed variable threshold dual rate ADPCM coding method which adapts two coding  
rates of the standard ADPCM of ITU G.726 for speech quality improvement at a comparably low coding rates.  
The ZCR(Zero Crossing Rate) is computed for speech data and under the noisy environment, noise data  
dominant region showed higher ZCR and speech data dominant region showed lower ZCR. The speech data  
with the higher ZCR is encoded by low coding rate for reduced coded data and the speech data with the lower  
ZCR is encoded by high coding rate for speech quality improvements. For coded data, 2 bits are assigned for  
low coding rate of 16[Kbps] and 5 bits are is assigned for high coding rate of 40[Kbps]. Through the  
simulation, the proposed idea is evaluated and shown that the variable dual rate ADPCM coding technique  
shows the good speech quality at low coding rate.

Key Words : Dual rate, ADPCM, ZCR, Speech coding

\* 주저자 : 단국대학교 전기전자컴퓨터공학부 교수

Tel : 02-709-2831, Fax : 02-793-1387

E-mail : kyonghan@dku.edu

접수일자 : 2003년 8월 12일

1차심사 : 2003년 8월 25일

심사완료 : 2003년 10월 15일

## 1. 서 론

ADPCM은, PCM을 기반으로 하는 음성 압축 기  
법으로 ITU 권고안 G.726에서 표준으로 규정하고 있  
으며, 전송률에 따라 16[Kbps], 24[Kbps], 32[Kbps],

40[Kbps] 등으로 구분되어 진다[1][2]. 음성을 압축하여 전송할 때, 이들 중 하나를 고정 전송율로 선택하여 PCM 음성정보를 각각 2, 3, 4, 5 비트로 코딩하여 전송하게 된다. 이 경우 모든 음성 입력 신호에 대하여 동일한 압축비를 적용하게 된다.

그런데 PCM 음성정보에는 대화를 하는 동안 음성정보가 존재하는 구간과 대화가 없음으로 인하여 음성정보가 존재하지 않는 묵음구간이 존재하며 묵음 구간에 대하여는 원하지 않는 배경 잡음에 대한 데이터까지도 전송하는 경우가 발생된다.

따라서 묵음 등의 구간에 대해서는 배경잡음의 영향이 큰 구간이므로 이 부분에 대하여는 굳이 좋은 음질의 재현이 필요 없으므로 음성정보가 존재하는 구간과 같은 압축율을 부여할 필요가 없다. 즉, 대화 등에 의하여 음성정보가 존재하여 잡음의 영향이 비교적 작은 구간에 대하여는 양질의 음질 재현을 위하여 높은 비트 전송율을 사용하고 묵음 등의 잡음의 영향이 큰 구간은 낮은 비트 전송율을 사용함으로써 전체적으로 음질의 저하는 거의 없으면서 전송비트 수를 줄이도록 하였다. 대부분의 음성 대화에서는 묵음의 구간이 전체 대화 구간에서 상당한 부분을 차지하므로 이 방법은 기존의 단일 압축율을 사용하는 방법보다 음성데이터 평균 전송비트수를 줄일 수 있다[3].

본 논문에서는, 이러한 방법을 제시하여 음성 구간에 대하여 잡음의 영향에 따라 두가지 영역으로 분류하여 각각의 영역에 대하여 40[Kbps]와 16[Kbps] 의 서로 다른 두 가지 음성 압축율을 적용하는 음성압축 기법에 대하여 연구하였다. 또한 신호 및 잡음의 특성을 분석할 때 계산량이 상대적으로 많은 FFT 등을 대신하여 영교차율(Zero Crossing Rate)을 사용하여, 각 구간별로 구해지는 영교차율로 임계값을 가변시키고, 그 가변된 임계값 이상인 경우 40Kbps로 압축하고, 그 이하인 경우 16[Kbps]로 압축하는 가변임계값 dual rate ADPCM을 제안하였다.

제안한 방법은 PC 시뮬레이션을 통하여 단일 비율의 ADPCM과 고정임계값 dual rate ADPCM에 비해 주변 잡음특성에도 강하며 비교적 적은 데이터 전송률로도 충실한 음질을 얻을 수 있음을 보인다.

## 2. 본 론

### 2.1 ADPCM

ADPCM 알고리즘은 음성신호가 인접한 표본값 사이의 상관관계가 높다는 점에 기인하여 예측신호와 실제 신호사이에서의 차이만을 양자화 하므로 보다 정확하고 효과적으로 대역폭을 감소시킬 수 있다. 따라서 실제 신호 표본들과 예측 신호 표본들의 차이는 실제 신호 표본들을 양자화 하는데 필요한 비트 수보다 훨씬 더 적은 비트 수로 정확히 양자화 하게 될 것이다.

본 논문에서 사용된 적응 양자화는 양자화 제어 파라미터들을 전송하지 않고서도 부호화된 신호로부터 스텝 크기를 얻어낼 수 있는 궤환 적응 양자화기를 이용한다[4].

또한 시변 특성을 가지는 음성 신호의 본질상 예측기의 계수들도 지속적으로 변하는 신호들의 특성들을 예측하기 위하여 주기적으로 변해야만 한다. 보다 정확한 예측값을 얻기 위한 적응예측기의 계수를 구하기 위한 방법으로는 autocorrelation method 를 이용한다[4].

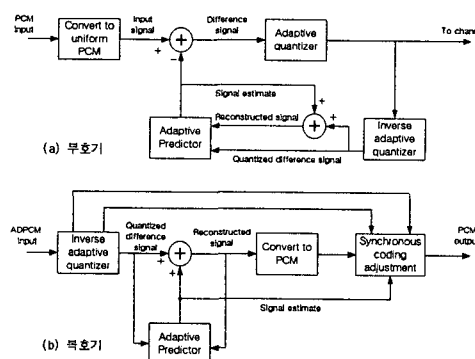


그림 1. ADPCM의 블록선도  
Fig. 1. ADPCM Block Diagram

그림1(a)는 ADPCM의 부호기를 나타낸다. 입력된 신호가 14비트 선형 PCM으로 변환된 후, 이 신호와 예측 신호를 뺀 차분 신호가 얻어지게 된다. 이러한 차분 신호를 디코더로 전송함에 있어 5, 4, 3, 2비트의 압축에 대해 각각 31, 15, 7, 4 레벨의 적응 양

자화기가 사용된다. 역 적응 양자화기는 양자화 되었던 차분 신호를 만들어 내는데, 이 신호는 입력 신호의 복원된 형태를 만들어내기 위해 신호의 예측값과 더해지게 된다. 그 결과로 만들어지는 복원된 신호와, 양자화 되었던 차분신호는 적응 예측기에 의해서 입력신호를 예측하게 되고 궤환 루프를 마치게 된다.

그림1(b)는 ADPCM의 복호기를 나타낸다. 부호기의 궤환 부분과 동일한 구조를 포함하며, 14비트 선형 PCM을 음성출력으로 변환하는 부분 외에도 동기 부호화 보정기(synchronous coding adjustment)부분을 포함한다.

이 부분은 PCM-ADPCM-PCM으로의 변환 시에 일어날 수 있는 점진적인 왜곡을 방지하게 되는데, 이것은 다음 ADPCM 부호화 단계에서 일어날 수 있는 양자화 왜곡을 제거하기 위해 PCM 출력을 조절함으로써 얻어지게 된다.

## 2.2 가변임계값 Dual Rate ADPCM

종래의 일정한 압축비(16[Kbps], 24[Kbps], 32[Kbps], 40[Kbps])로 음성신호를 부호화할 경우, 우리가 원하는 원치 않는 일정한 비율의 데이터를 전송하여야 한다. 하지만 2가지의 압축비율을 사용하게 되면 각 압축비율이 가지는 비트 수의 평균값만을 전송하게 되므로 전송되는 데이터의 양을 많이 줄일 수 있게된다.

제한한 가변임계값 dual rate ADPCM은 좋은 음질이 필요없는 구간(목음, 잡음)은 16[Kbps](2비트)로 압축하여 데이터를 줄이고, 좋은 음질이 필요한 구간(음성)은 40[Kbps](5비트)로 압축하여 음질의 충실도를 높이고, 그 결과인 음성을 청취했을 때 40[Kbps]로 압축한 음성과 거의 비슷한 수준의 음질을 들을 수 있다. 따라서 입력되는 음성신호의 크기로 16[Kbps]와 40[Kbps]로 모드를 변환해주는 임계값이 필요하고, 또한 임계값의 크기를 적절하게 가변시킬 수 있는 기준값이 필요한데 이 값을 영교차율로 제안하였다.

### 2.2.1 영교차율(Zero Crossing Rate)

음성신호는 주위 잡음보다 낮은 영역의 주파수대

에 분포하고 있으므로 음성신호에 잡음이 섞이게 되면 주파수가 증가하고 주파수가 증가한다는 것은 상대적으로 영교차율이 증가하는 결과를 가지고 음성 및 잡음이 존재하는 구간을 구분해 낼 수 있다. 따라서 음성통화 시에 음성이 존재하는 구간에서는 저주파의 큰 신호이기 때문에 잡음이 섞이더라도 영교차율이 낮고, 목음이 존재하는 구간에서는 음성이 존재하는 구간보다 영교차율이 높아지게 된다. 이런 특성을 이용하여 영교차율로 음성신호에 섞인 잡음의 정도와 어느 부분에 잡음이 섞였는지를 알 수 있고, 이 값을 통하여 dual rate ADPCM의 임계값을 가변시켜 잡음이 섞인 환경에서도 충실한 음질의 음성을 전송할 수 있게 된다.

이러한 영교차율을 이용할 때의 장점은 신호의 특성을 분석 및 가공할 때 기존의 FFT 및 각종 필터를 사용하는 것이 아니기 때문에 계산량을 현저하게 줄일 수 있다. 하지만 영교차율을 적용하여 계산을 수행할 경우 신호의 샘플과 샘플사이의 부호 변화만 감지하기 때문에 잡음의 크기나 정확한 특성은 고려할 수 없는 단점이 있다. 하지만 입력되는 신호의 주파수가 높은지 낮은지를 실시간으로 검출하는데 있어서 영교차율은 빠른 계산속도를 보인다.

대다수 모든 디지털 신호처리는 A/D변환을 하면서 버퍼를 사용하기 때문에 입력되는 데이터를 버퍼에 저장하는 과정에서 영교차율을 계산하므로 거의 시간이 소모되지 않는다.

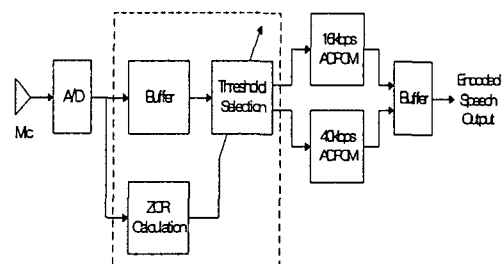


그림 2. 가변임계값 Dual Rate ADPCM의 구성도  
Fig. 2. Variable Dual Rate ADPCM Schem

그림2은 가변임계값 dual rate ADPCM의 시스템 구성도이고 점선부분이 영교차율을 계산하여 임계값(threshold value)을 가변시키는 부분을 나타낸다.

그림3은 가변임계값 dual rate ADPCM의 처리 순서도를 나타낸 것이다. 입력된 신호는 A/D 변환을 거쳐 버퍼에 저장되면서 영교차율을 계산한다. 계산된 영교차율이 기준값보다 크면 임계값을 높이고 그렇지 않으면 임계값을 낮추어 설정한다. 마지막으로 설정된 임계값을 가지고 입력된 음성신호가 임계값보다 낮으면 16[Kbps]로, 높으면 40[Kbps]로 음성을 전송한다. 그렇게 하므로 잡음구간은 16[Kbps]로, 음성구간은 40[Kbps]로 전송하여 음질의 저하없이 전송되는 데이터의 양을 줄인다.

수신단에서는 압축비에 따라 16[Kbps]와 40[Kbps] 가운데 해당되는 비트 수의 복호기로 복원하게 된다.

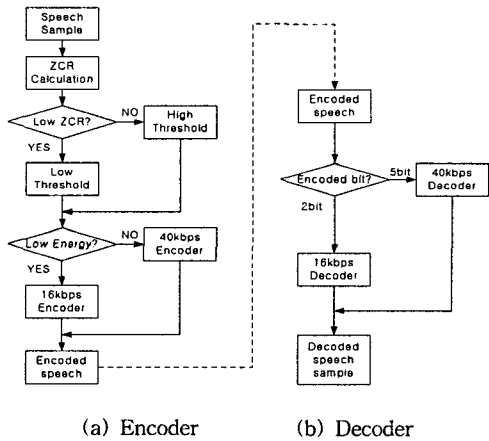


그림 3. 가변임계값 Dual Rate ADPCM의 순서도  
 Fig. 3. Variable Dual Rate ADPCM flowchart

2.2.2 영교차율 계산

그림4는 영교차율을 계산하는 부분을 나타낸 것이다.

```

for(i=0; i<BUFFER_SIZE; i++)
{
    temp = old_in_buffer ^ in_buffer[i];
    if(temp < 0)
    {
        zero_crossing_rate++;
    }
    old_in_buffer = in_buffer[i];
    k++; //Increase Sample Count(1~18,000)
}
    
```

그림 4. 영교차율 계산을 위한 프로그램  
 Fig. 4. Program for ZCR calculation

영교차율을 계산할 때 부호의 변화를 알기 위하여 곱셈을 한번 해야하는데 정확한 데이터 값을 알기 위함이 아니고 단지 부호의 변화 값만 필요하기 때문에 그 부분을 C-언어의 Exclusive-OR(^) 연산자로 대체하였다[3][5].

2.3 시뮬레이션

제안한 가변임계값 dual rate ADPCM과 기존의 16Kbps, 40Kbps ADPCM을 PC상에서 비교, 분석하였다. 입력신호로는 PCM의 여러 가지 포맷 중에서 음성통신의 표준인 8kHz의 샘플링 주파수를 가지는 16비트 선형 PCM방식으로 하였다[6][7].

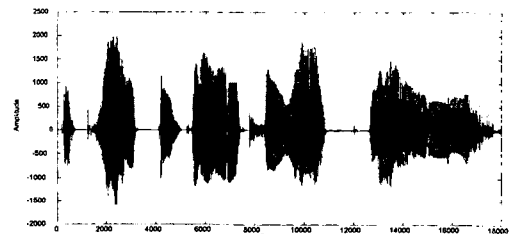


그림 5 입력에 사용된 음성신호  
 Fig. 5. Speech input signal

그림5는 입력 음성 신호의 파형을 나타낸다. 또한 잡음이 포함된 환경을 설정하기 위해 원음의 음성에 백색잡음과 도로나 지하철역에서 녹음한 잡음들을 합성하여 표본을 만들었다. 녹음된 PCM 음성데이터는 약 18,000개의 샘플들을 80개 샘플마다 버퍼에 저장하면서 영교차율을 계산하였다[8].

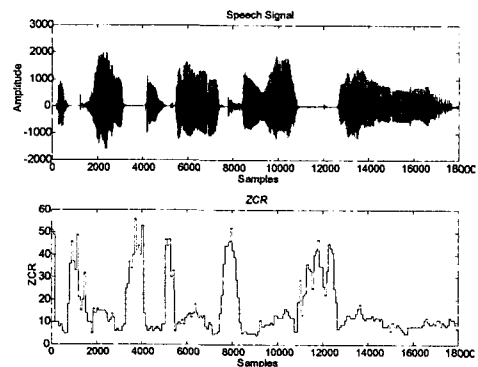


그림 6. 샘플음성의 영교차율  
 Fig. 6. ZCR for sample speech

배경 잡음환경에서 가변 임계값에 의한 Dual Rate ADPCM 음성 부호화 기법

그림6~8는 각 샘플의 영교차율을 구한 것이다. 그림들을 보면 알 수 있듯이 음성신호에 잡음이 섞이게 되면 영교차율이 상승하는데, 특히 음성신호의 크기가 작은 구간에서 영교차율이 급격히 상승한다. 이 구간은 음성 신호보다 잡음 신호가 더 지배적인 구간으로 구분된다.

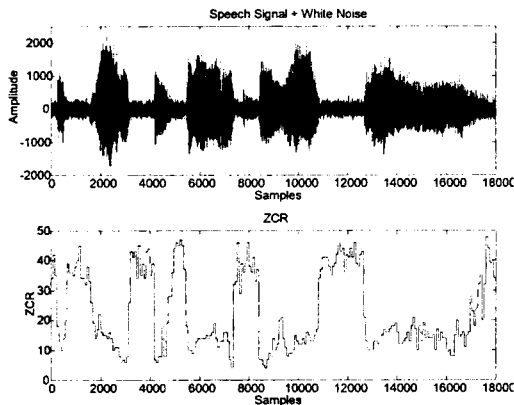


그림 7. 샘플음성+백색잡음의 영교차율  
Fig. 7. ZCR for sample speech with white noise

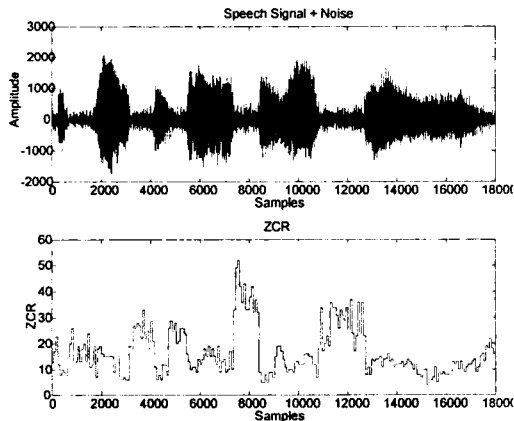


그림 8. 샘플음성+잡음의 영교차율  
Fig. 8. ZCR for sample speech with noise

음성신호가 잡음신호보다 큰 부분은 음성신호가 지배적인 구간으로 이 구간의 영교차율은 크게 상승하지 않는다. 따라서 제안한 영교차율로써 잡음이 지배적인 구간과 음성 신호가 지배적인 구간을 구분할 수 있다.

모든 설정값(영교차율 기준, 잡음구간일 때 임계값, 음성구간일 때 임계값)은 실험에 의하여 선정하였고, 최종적으로 영교차율의 기준은 버퍼크기(80)의 1/4로 설정하였고, 영교차율 기준보다 큰 잡음구간은 임계값을  $\pm 250$ 로, 음성구간은  $\pm 150$ 으로 설정하여 시뮬레이션 하였다.

표 1. 표준 압축 ADPCM(16Kbps, 40Kbps)과 가변임계값 dual rate(16,40Kbps) ADPCM의 비교  
Table 1. Comparison of standard ADPCM with variable dual rate ADPCM.(16(Kbps), 40(Kbps))

		평균 비트수	SNR[dB]
16[Kbps]	원음	2	14.866
	원음+백색잡음	2	11.802
	원음+잡음	2	13.082
40[Kbps]	원음	5	33.069
	원음+백색잡음	5	28.364
	원음+잡음	5	30.114
가변임계값 dual rate ADPCM (16, 40[Kbps])	원음	3.502	25.409
	원음+백색잡음	3.492	21.496
	원음+잡음	3.550	22.604

가변임계값 dual rate ADPCM으로 압축하여 전송한 음성은 40[Kbps] 단일 압축으로 전송한 것보다는 SNR이 떨어지게 된다. 그 이유는 잡음 구간으로 판정된 부분을 16[Kbps]로, 음성구간은 40[Kbps]로 전송되기 때문에 전체적으로는 SNR이 내려가지만, 음질은 거의 동일한 효과가 있다.

### 3. 결 론

본 논문에서는 잡음이 있는 환경에서도 기존의 음질 상태를 유지하면서 데이터의 전송량을 줄일 수 있는 16[Kbps]와 40[Kbps]의 가변임계값 dual rate ADPCM 부호화기법을 제안하였다.

제안한 알고리즘이 16[Kbps]와 40[Kbps] 각각의 단일 ADPCM보다 압축률 및 부호화된 음질의 양면에서 우수함을 시뮬레이션 및 직접 들어봄으로써 확인하였다. 또한 영교차율을 제시함으로써 잡음구간을 분석해

낼 수 있는 복잡한 과정을 단순하게 줄임으로 계산량과 수행 속도를 현저하게 줄일 수 있었다.

실제 음성 통화시에는 전체 대화시간 중에 묵음이 차지하는 비율이 80% 정도이므로 조용할 때나 잡음이 있을 때를 구분하여 압축한다면 적은 양의 데이터를 전송하면서도 보다 좋은 음질의 통화를 할 수 있을 것이라 생각된다. 또한 제안된 영교차율을 이용하여 주파수의 높고 낮음을 분석하는데 사용한다면 FFT를 사용하는 것보다 훨씬 간단한 작업으로도 비슷한 성능을 낼 수 있고, 계산량을 줄일 수 있어 수행속도 또한 빠르게 할 수 있다.

이 연구는 2003학년도 단국대학교 대학연구비지원으로 연구되었음.

## References

- (1) 진년강, "아날로그와 디지털통신", 청문각, 1990.
- (2) PULSE CODE MODULATION(PCM) OF VOICE FREQUENCIES, ITU-T Recommendation G.711, 1988.
- (3) V.K.Ingle and J.G.Proakis, "Digital Signal Processing Using Matlab V.4", PWS Publishing Company, Boston, 1997.
- (4) Thomas P.Barnwell III, "Speech Coding A Computer Laboratory Textbook", John Willy & Sons Inc., 1996.
- (5) Richard A. Jaddad and Thomas W. Parsons, "Digital Processing Theory", Applications and Hardware, Computer Science Press, 1991.
- (6) L. R. Rabiner and R. W. Schafer, "Digital Processing of Speech Signals", Prentice-Hall, 1978.
- (7) 40, 32, 24, 16 kbit/s ADAPTIVE DIFFERENTIAL PULSE CODE MODULATION (ADPCM), CCITT Recommendation G.726, 1990이소베도시오, "C언어와 RS-232C/CP-IB", 다다미디어, 1993.
- (8) T.P. Barnwell, "Recursive windowing for generating auto-correlation analysis for lpc analysis," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-29, 1981.

## ◇ 저자소개 ◇

### 한경호 (韓敬浩)

1959년 6월 25일생. 1982년 서울대학교 공과대학 전자공학과 졸업. 1984년 서울대학교 대학원 전자공학과 졸업(석사). 1992년 미국 Texas A&M University 졸업(박사). 1993년 ~ 현재 단국대학교 교수.