

# 16Kbps와 40Kbps의 Dual Rate G.726 ADPCM 음성 codec 구현

## Implementation of G.726 ADPCM Dual Rate Speech Codec of 16Kbps and 40Kbps

金 載 沆\*, 韓 敬 浩\*

( Jae-Oh Kim\* and Kyong-Ho Han\* )

### 요 약

본 논문에서는 G.726 ADPCM 음성방식을 기존의 단일 압축율 대신 16Kbps 와 40Kbps의 두 가지 압축율을 사용한 가변 압축방식에 의한 음성 코딩 방식을 다루었다. 음성의 묵음 또는 소 신호 부분은 음질의 향상보다는 데이터 비트 수를 줄이기 위한 저 전송 16Kbps 압축율을 적용하였고 임계값 이상의 대 신호 부분은 음질을 향상하기 위하여 40Kbps의 압축율을 적용하여 모든 신호를 단일 압축율로 코딩하는 방식에 비하여 전체적으로 압축율을 높여 전송 비트 수를 줄이면서 음질을 저하시키지 않도록 하였다. 본 논문에서는 시뮬레이션을 통하여 여러 가지의 임계값에 의한 가변 압축 코딩 방식에 대하여 압축율과 음질의 관계를 다루었다. 또한 고정된 임계값에 대하여 입력 음성의 크기를 여러 가지로 변동하여 주변 배경잡음과 포화에 의한 음질의 저하를 고찰하여 가변 율에 의한 음성의 코딩방식에서 임계값과 입력의 크기가 음질 및 압축율에 미치는 영향을 다루었다. 각 시뮬레이션의 경우에 대하여 실제 음성의 원음에 대한 음질의 충실도를 임의의 집단에 대하여 비교하여 음질의 충실도를 확인하였다. 추후의 연구를 통하여 DSP에 의한 실시간 처리 시스템의 구현을 하고자 한다.

### Abstract

In this paper, the implementation of dual rate ADPCM using G.726 16Kbps and 40Kbps speech codec algorithm is handled. For small signals, the low rate 16Kbps coding algorithm shows almost the same SNR as the high rate 40Kbps coding algorithm, while the high rate 40Kbps coding algorithm shows the higher SNR than the low rate 16Kbps coding algorithm for large signal. To obtain the good trade-off between the data rate and synthesized speech quality, we applied low rate 16Kbps for the small signal and high rate 40Kbps for the large signal. Various threshold values determining the rate are applied for good trade-off between data rate and speech quality. The simulation result shows the good speech quality at a low rate comparing with 16Kbps & 40Kbps.

Key Word: ADPCM, Dual Rate, Speech Coding, Data Rate, Speech Quality

\* 檀國大學校 電氣工學科

(Department of Electrical Eng., Dankook Univ.)

※본 연구는 단국대학교 교내 연구비 지원에 의하여 수행되었습니다.

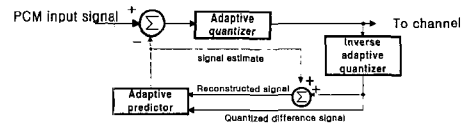
接受日: 1998年9月14日, 修正完了日: 1998年12月14日

### I. 서론

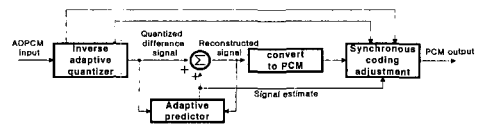
음성의 디지털부호화 기술은 그 해석방법에 따라

크게 음성 파형을 표본화하여 양자화 하는 파형 부호화(waveform coding)방식과 음성의 주기와 성도의 계수 등 음성의 특징만 추출하여 전송해서 수신 측에서 음성을 재생하는 parameter coding방식으로 분류되어진다. parameter coding 방식의 경우는 음성 압축율이 높아 전송속도가 50[BPS]에서 4.8[Kbps]로 waveform coding 방식보다 적지만 구현시의 복잡성으로 많은 비용이 필요하다. waveform coding의 방식의 경우는 전송속도가 16[Kbps]에서 64[Kbps]로 상대적으로 높지만 음질이 우수하며 복잡도가 상대적으로 낮아 일반 음성 통신에 많이 사용되고 있다.[1] 이러한 waveform coding 방식의 하나인 Adaptive Differential Pulse Code Modulation(ADPCM)은, PCM을 기반으로 하는 음성압축 기법으로 ITU(International Telecommunication Unit)권고안 G.726에서 표준으로 규정하고 있다. 전송률에 따라 16Kbps, 24Kbps, 32Kbps, 40Kbps 등으로 구분되어지며, ADPCM은 8bit  $\mu$ -law, A-law 또는 16bit 선형 PCM 형식의 음성신호가 입력되어지고, 동일한 형식으로 출력이 가능하다.[2,3] 본 논문에서는 음성의 특성상 silent인 부분이 70-80% 이상을 차지하는 것을 착안하여 G.726의 16Kbps 알고리즘과 40Kbps 알고리즘을 이용하여 신호의 에너지를 비교하여 특정값 이상인 경우에는 40Kbps로, 그 이하인 경우에는 16Kbps로 압축을 하는 dual rate ADPCM을 제안하였으며, IBM-PC상에서 시뮬레이션을 통하여 단일 비율의 ADPCM에 비해 원음에 대한 충실 도와 성능을 확인하였다.

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. 여기에 사용되는 적응양자화기는 양자화 제어 파라미터들을 전송하지 않고서도 부호화된 신호로부터 스텝 크기를 얻어낼 수 있는 레환 적응 양자화기(feedback adaptive quantizer)를 이용한다.[4] 또한 시변특성을 가지는 음성신호의 본질상 예측기의 계수들도 지속적으로 변하는 신호들의 특성들을 예측하기 위하여 주기적으로 변해야만 한다. 보다 정확한 예측값을 얻기 위한 적응예측기의 계수를 구하기 위한 방법으로는 autocorrelation method를 이용하기로 한다.[2]



(a) ADPCM encoder



(b) ADPCM decoder

그림 1. ADPCM의 block diagram

Fig. 1. The Block Diagram of ADPCM.

## II. 본 론

### 2.1 ADPCM (Adaptive Differential PCM)

고정된 step size를 갖는 PCM이나 DPCM의 경우에는 그 레벨이 고정되어 있기 때문에 입력신호가 큰 폭으로 변화할 경우 상당한 왜곡이 생길 수 있다. 이러한 문제점은 양자화 노이즈의 범위를 줄일 수 있도록 적응 양자화기와 적응 예측기를 사용함으로써 가능하다.[2],[3]

ADPCM 알고리즘은 음성신호가 인접한 표본값

그림 1.(a) 는 ADPCM의 encoder부분을 나타낸다.[5] 인코더 부에서는 먼저 8bit  $\mu$ -law, A-law 혹은 16bit 선형 PCM의 입력신호가 14비트 선형 PCM으로 변환된 후, 이 신호와 예측 신호를 뺀 차분신호를 디코더로 전송함에 있어 5,4,3,2비트로의 압축에 대해 각각 31, 15, 7, 4 개의 레벨을 갖는 적응 양자화기를 사용한다. 역 적응 양자화기는 양자화 되었던 차분 신호를 만들어 내는데, 이 신호는 입력신호의 복원된 형태

를 만들어내기 위해 신호의 예측값과 더해지게 된다. 그 결과로 만들어지는 복원된 신호와, 양자화 되었던 차분신호는 적응예측기에 의해서 입력신호를 예측하게 되고 케환루프를 마치게 된다. 이 적응예측기는 2개의 극점과 6개의 영점을 가진 필터로, 각각의 필터 계수들은 입력되는 샘플들에 의해 항상 업데이트된다. 그림 1.(b)는 ADPCM의 decoder부분을 나타낸다.[5] 디코더는 인코더의 케환부분과 동일한 구조를 포함하며, 14비트 선형 PCM을  $\mu$ -law, A-law, 혹은 선형 PCM으로 변환하는 부분 외에도 synchronous coding adjustment부분을 포함하는데 이것은 synchronous tandem coding시에 일어날 수 있는 점진적인 왜곡을 방지하여 다음 ADPCM 인코딩 단계에서 일어날 수 있는 양자화 왜곡을 제거하기 위해 PCM 출력을 조절함으로써 얻어지게 된다.

## 2.2 16Kbps와 40Kbps의 Dual Rate ADPCM

일반적인 대화시의 음성은 80% 이상이 무음(silent speech)에 가까운 소 신호에 해당한다. 이러한 부분에서는 16Kbps와 40Kbps가 서로 거의 같은 복원된 음성을 가지므로 40Kbps의 high bit rate로 압축할 필요가 없으며, 소 신호 영역은 low bit rate를 갖는 16Kbps의 2비트 압축율을 적용하고 그 이상의 값을 갖는 신호에 대해서는 고음질을 얻기 위해 40Kbps의 5비트 압축율을 적용하여 실제적인 음질은 40Kbps 수준을 유지하면서 압축 데이터의 양은 대폭 줄이는 방법을 제안하였다. 특히 본 논문은 궁극적으로 실시간 디지털 음성 전송을 구현함을 목적으로 하므로, dual rate coding에 대하여 단일 문장이 아닌 대화시의 음성은 음질과 압축율의 면에서 매우 뛰어난 특성을 가지게 될 것이다. 여기서 입력신호의 에너지의 높고 낮음에 대한 경계는 여러 가지 값으로 변화시키면서 압축율과 복호화된 음성의 음질의 변화를 조사하였다.

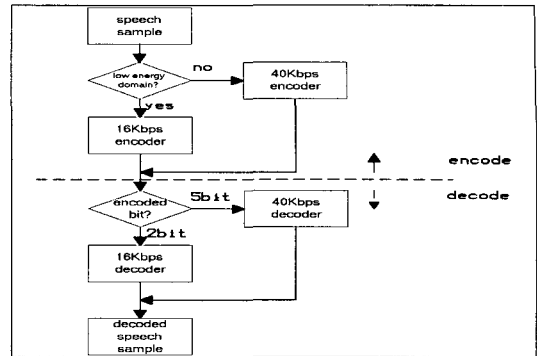


그림 2. Dual Rate ADPCM 처리 흐름도

Fig. 2. The Process Flow of Dual Rate ADPCM.

그림 2는 제안된 dual rate ADPCM의 처리흐름도를 나타낸다. 송신 단에서 입력되는 음성의 에너지가 경계영역의 밖일 경우에는 40Kbps로, 경계치 이내인 경우에는 16Kbps encoder로 압축하여 그 값을 수신 단으로 보낸다. 수신 단에서는 압축된 값이 5비트인 경우는 40Kbps로, 2비트인 경우는 16Kbps decoder로 복원하게 된다. 본 입력 신호(16bit linear PCM)에 대해서는  $\pm 150$ 을 임계값으로 정하여, 이 범위 내에서는 16Kbps로, 그 외 영역은 40Kbps로 정했다. 몇 개의 샘플로 선택된 음성 신호에 대해서도 경계영역을  $\pm 150$ 로 정하였을 경우에 평균적으로 3.4-3.8비트 수준으로 나타났다. 이러한 dual rate ADPCM은 16Kbps 단일 압축율을 가질 때보다 훨씬 더 충실한 음질을 얻을 수 있으며, 40Kbps 단일 압축율을 가질 때보다 더 나은 압축율을 가지게 된다.

## 2.3 컴퓨터 시뮬레이션

본 장에서는 컴퓨터 시뮬레이션을 통해 앞 절에서 제안한 dual rate ADPCM과 기존의 16Kbps와 40Kbps의 파형을 PC상에서 비교, 분석하였다. 입력신호로는 PCM의 여러 가지 포맷 중에서 16bit linear PCM방식으로 하였다. 아래 그림 3은 입력 음성 신호의 파형을 나타낸다. 이 파형은 마이크로폰으로 입력 받은 음성 신호를 PC내부의 사운드카드(sound blaster

16WE)에 내장된 ADC를 통하여 16비트 선형 PCM데이터로 입력받은 신호이다. 그림 4 는 위의 입력신호를 각각 16Kbps와 40Kbps의 단일 압축 codec시의 결과 파형을 나타낸다. 사운드카드를 통해서 받아들인 파일로 저장한 후에 16Kbps와 40Kbps로 각각 ADPCM encoder로 부호화한 다음 ADPCM decoder에서 다시 선형 PCM 데이터로 복원시킨 후에 그 결과를 파일로 저장하여 다시 사운드카드를 통하여 다시 출력시킨 파형 들이다.

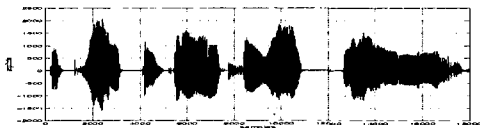
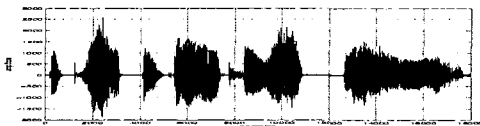
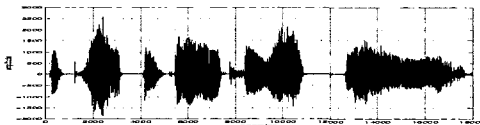


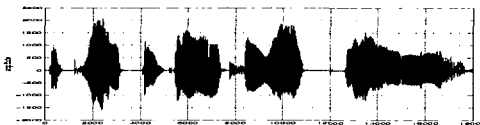
그림 3. 입력에 사용될 음성신호  
Fig. 3. Original Speech Signal.



(a) 16Kbps 단일 압축율인 경우의 복원된 음성



(b) 40Kbps 단일 압축율인 경우의 복원된 음성



(c) 16/40Kbps dual rate인 경우의 복원된 음성

그림 4. 각각의 압축율에 따른 codec 결과

Fig. 4. The Codec Result Signals for Various Coding Rate.

그림 4.(c)는 본 논문에서 제안된 16/40Kbps의 dual rate 압축율을 가질 때의 codec결과를 나타낸 파형이다. 표 1 은 이 세 가지 경우의 압축율에 대한 신호 대 잡음비를 비교한 것이다.

표 1. 각각의 압축율에 따른 신호 대 잡음비 비교  
Table 1. The Comparison of SNR for Various Coding Rate.

	16Kbps	dual rate	40Kbps
신호 대 잡음비	14.866	25.723	33.069
평균 비트 수	2	3.508	5

dual rate 인 경우를 40Kbps와 비교할 때 훨씬 적은 비트 수를 가지고도 16Kbps 보다 월등히 뛰어난 음질을 얻을 수 있었다.

2.4 경계영역과 입력신호의 크기의 변화에 따른 codec 결과

표 2 는 경계영역을 변화시켰을 때 평균비트율과 신호 대 잡음비에 미치는 영향을 나타낸 표이다. 경계영역을 ±50부터 100단위로 ±650까지 변화시켜 codec한 결과를 평균 비트 율과 신호 대 잡음비로 비교하였다. 그림 5 는 표 2 의 데이터를 도표 상으로 나타낸 것이다. 경계영역이 커질수록 2비트의 압축영역이 커지고 5비트의 압축영역이 줄어들게 되므로 전체적인 평균 비트 율은 아래로 완만한 곡선형태를 그리며 떨어지게 되며, 따라서 신호 대 잡음비도 떨어지게 된다.

표 2. 경계영역 변화시의 평균 비트 율과 신호 대 잡음비의 관계

Table 2. The Average Bit Rate and SNR for Various Threshold Values.

경계영역	평균 비트 율	신호 대 잡음비[dB]
50	3.9565	26.386
150	3.5080	25.723
250	3.2227	22.649
350	2.9803	21.585
450	2.7622	20.212
550	2.5588	19.546
650	2.3925	18.336

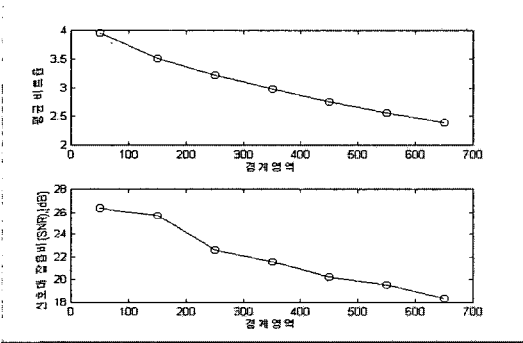


그림 5. 경계영역 변화시의 평균 비트 율과 신호 대 잡음비의 관계

Fig. 5. The Average Bit Rate and SNR for Various Threshold Values.

표 6은 입력신호의 크기를 변화시킨 경우의 결과를 나타낸 표이다. 입력신호의 크기를 0.1배부터 10배까지 변화시켜 평균 비트 율과 신호 대 잡음비를 비교하였으며, 그림 6은 표 3의 값을 도표 상으로 나타낸 것이다. 입력신호의 전체적인 크기가 줄어들면 상대적으로 2비트의 압축영역이 커지게 되고 5비트 압축영역은 적어지므로 평균비트율은 떨어지게 된다. 한편 입력신호의 크기에 따른 신호 대 잡음비는 그림에서 보듯이 1배에서 4배 일 때가 가장 높은 값을 보이게 되는데 이것은 그 이하의 배율인 경우는 상대적으로 약한 음성신호에 대한 배경잡음의 비중이 커져 신

호 대 잡음비가 떨어지게 되었으며, 4배 이상으로 증폭될 경우는 점차적으로 16비트 음성의 포화 (saturation)에 의한 왜곡이 발생하여 왜곡에 의한 신호 대 잡음비의 작아짐을 보였다.

표 3. 입력신호 크기 변화시의 평균 비트 율과 신호 대 잡음비의 관계

Table 3. The Average Bit Rate and SNR for Various Input Signal Amplitude.

입력신호의 크기	평균 비트 율	신호 대 잡음비 [dB]
0.1배	2.0125	10.517
0.2배	2.2815	14.699
0.4배	2.9205	19.659
0.6배	3.2235	22.088
0.8배	3.3947	22.330
1배	3.5080	25.723
2배	3.8060	25.256
4배	4.0415	25.708
6배	4.1500	20.487
8배	4.2132	17.275
10배	4.2605	15.396

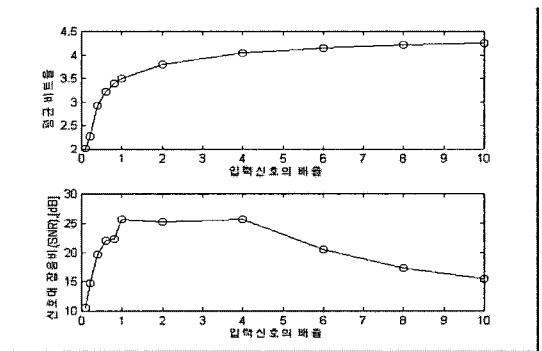


그림 6. 입력신호 크기 변화시의 평균 비트 율과 신호 대 잡음비의 관계

Fig. 6. The Average Bit Rate and SNR for Various Input Signal Amplitude.

가변을 음성 압축에 의한 음질을 확인하기 위한 방법으로 20명을 대상으로 하여 16Kbps, 40Kbps 그리고 본 가변을 압축에 의하여 압축되고 재생된 음성신호를 대상으로 가장 좋은 신호를 선택하도록 하였다. 그 결과는 표 4 에서 보이며 실험 결과, 가변을 압축에 의한 재생 음의 음질은 40Kbps와 비교하여 거의 같음을 보여주고 있다.

표 4. 16Kbps, 40Kbps, 가변을 코딩에 의한 재생 음의 음질 비교

Table 4. The Comparison of Speech Quality of 16Kbps, 40Kbps and Dual Rate Coding.

	16Kbps (2비트)	40Kbps (5비트)	가변을 (3.5비트)
음성 1	0%	50%	50%
음성 2	0%	50%	50%
음성 3	0%	50%	50%
음성 1	제 이름은 김재오 입니다		
음성 2	The pipe begins to rust while new		
음성 3	이 음성은 dual rate ADPCM에 의한 음성입니다		

III. 결론

본 논문에서는 16Kbps와 40Kbps의 dual rate ADPCM에 의한 음성 코딩 방법을 제안하였다. 제안

한 알고리즘에 의하여 복호화된 음성은 16Kbps보다 훨씬 더 뛰어난 음질을 가지며 40Kbps 보다는 훨씬 더 적은 압축율을 가지며 대등한 음질을 가짐을 시물레이션을 통해서 확인하였다.

실험에 대한 대상으로 예제 문장을 입력 음성신호로 주었으나 실제 전화상에서 대화시 silent speech에 가까운 소신호가 60% 이상이므로 실제적인 데이터 전송량은 40Kbps보다 훨씬 더 작을 것으로 예상되며, 16Kbps보다는 훨씬 더 우수한 음질을 얻을 수 있을 것으로 보인다.

향후 본 알고리즘으로 Texas Instruments사의 고정소수점 DSP인 TMS320C54 보드에서 실시간 구현할 예정이다.

참 고 문 헌

[1] 진년강, 아날로그와 디지털 통신, 청문각, pp178, 1990  
 [2] L. R. Rabiner and R. W. Schafer, Digital Processing of Speech Signals, Prentice-Hall, pp200-211, 1978  
 [3] A. Nejat Ince, Digital Speech Processing, Kluwer Academic Publishers, pp 161-173 1992  
 [4] Thomas P. Barnwell III, Speech Coding A Computer Laboratory Textbook, John Willy & Sons Inc., pp61-64, 1996  
 [5] V.K.Ingle and J.G.Proakis, Digital Signal Processing Using Malab V.4, PWS Publishing Company, Boston, pp396, 1997

저 자 소 개



金 載 沄 (學生會員)  
1975년 2월 7일생. 1997년 서울 단국대 공대 전기공학과 졸업. 1999년 동 대학원 전기공학과 졸업예정. 주요관심분야: 음성 압축 구현, DSP 응용. 마이크로프로세서 응용



韓 敬 浩 (正會員)  
1959년 6월 25일생, 1982년 서울대학교 전자공학과 졸업(학사), 1984년 동 대학원 전자공학과 졸업(공학 석사), 1992년 미국 Texas A&M University, Electrical Engineering (컴퓨터공학 전공) 졸업 (Ph. D). 1984.2 - 1985.5 삼성전자 휴렛패커드 연구원, 1984.8 - 1985.2 육군제삼사관학교 (예비역소위). 1985.5.- 1987.12 한국통신 품질보증단 (전임연구원). 1989.1 - 1992.10 미국 Texas A&M University, Electrical Engineering (Unix & Network System Administrator) 1992.10 - 1993.8 한국전자통신연구소 이동통신연구단 (전임연구원) 1993. 8 - 1995.9 단국대학교 전기공학과 전임강사 1995.9 - 현재 단국대학교 전기공학과 조교수. 주요관심분야: 컴퓨터 구조, 마이크로프로세서 설계, 마이크로프로세서 응용, 컴퓨터 통신, 디지털 VLSI 설계, 음성 압축 구현, DSP 응용.