

16Kbps와 40Kbps의 Dual Rate G.723 ADPCM 음성 codec 구현

김재오, 한경호
 단국대학교 전기공학과

Implementation of Dual Rate G.723 ADPCM Speech codec

JaeOhe Kim, Kyongho Han
 Department of Electrical Engineering, Dankook University

Abstract - In this paper, the implementation of dual rate ADPCM using G.723 16Kbps and 40Kbps speech codec algorithm is handled.

For small signals, the low rate 16Kbps coding algorithm shows the same SNR as the high rate 40Kbps coding algorithm, while the low rate 16Kbps coding algorithm shows the lower SNR than the high rate 40Kbps 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 tested for good trade off data rate and speech quality.

Also the low pass filter effect of speech input and output devices is simulated at several cut-off frequencies. To simulation result shows the good speech quality at a low rate comparing with 16Kbps & 40Kbps.

1. 서 론

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 coding 방식으로 분류되어 진다. parameter coding 방식의 경우는 음성 압축률이 높아 전송 속도가 50[bps]에서 4.8[Kbps]로 waveform coding 방식보다 적지만 구현시의 복잡성으로 많은 비용과 노력이 필요하다. waveform coding 방식의 경우는 16[Kbps]에서 64[Kbps]로 상대적으로 높지만 음질이 우수하여 일반 음성통신에 많이 사용되고 있다.

본 논문에서는 이러한 waveform coding 방식의 하나인 Adaptive Differential Pulse Code Modulation(ADPCM)으로, PCM을 기반으로 하는 음성압축 기법을 이용하기로 한다.

이러한 ADPCM은 ITU(International Telecommunication Unit) 권고안 G.723에서 표준으로 규정하고 있으며, 전송율에 따라 16Kbps, 24Kbps, 32Kbps, 40Kbps 등으로 구분되어 진다. ADPCM은 μ -law, A-law, 또는 선형 PCM형식으로 입력되어지고, 동일한 형식으로 출력이 가능하다. 본 논문에서는 G.723의 16Kbps 알고리즘과 40Kbps 알고리즘을 이용하여 신호의 에너지를 비교하여 특정값 이상인 경우에는 40Kbps로, 그 이하인 경우에는 16Kbps로 압축을 하는 dual rate ADPCM을 제안하고자 한다.

본 논문에서는 앞으로 다루게 될 Texas Instruments사의 고성소수점 DSP인 TMS320C54를 사용하여 실시간으로 dual rate ADPCM을 구현하기에 앞서 IBM-PC 상에서의 시뮬레이션에 중점을 두었다.

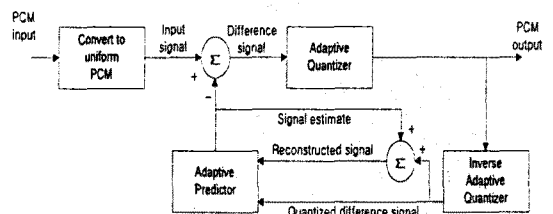
2. 본 론

2.1 ADPCM(Adaptive Differential PCM)

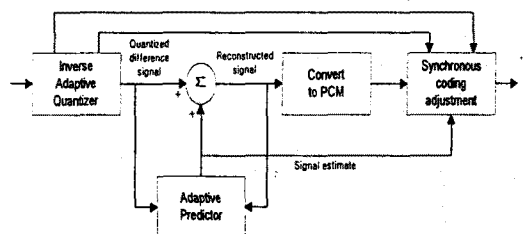
고정된 step size를 갖는 PCM이나 DPCM의 경우에는 그 레벨이 고정되어 있기 때문에 입력신호가 큰 폭으로 변화할 경우 상당한 왜곡이 생길 수 있다. 이러한 문제점은 양자화 노이즈의 범위를 줄일 수 있도록 적응 양자화기와 적응 예측기를 사용함으로써 가능하다.

ADPCM 알고리즘은 음성신호가 인접한 표본값사이의 상관관계가 높다는 점에 기인하여 예측신호와 실제 신호사이에서의 차이만을 양자화하므로 보다 정확하고 효과적으로 대역폭을 감소시킬 수 있다. 따라서 실제 신호표본들과 예측 신호표본들의 차이는 실제 신호표본들보다 작은 분산을 가지게 될 것이고, 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 제환 적응 양자화기(feedback adaptive quantizer)를 이용한다. 또한 시변특성을 가지는 음성신호의 본질상 예측기의 계수들도 지속적으로 변하는 신호의 특성들을 예측하기 위하여 주기적으로 변해야만 한다. 보다 정확한 예측값을 얻기 위한 적응예측기의 계수를 구하기 위한 방법으로는 covariance method나 autocorrelation method 등 여러 가지 방법이 있으나 여기서는 autocorrelation method를 이용하기로 한다.



(a) encoder



(b) decoder

그림1. ADPCM 부호기와 복호기의 block diagram

2.1.1 ADPCM encoder

인코더 부에서는 먼저 μ -law, A-law, 혹은 선형PCM의 입력 신호가 14bit 선형 PCM으로 변환된 후, 이 신호와 예측 신호를 뺀 차분신호가 얻어지게 된다(그림1. (a) 참조).

이러한 차분신호를 디코더로 전송함에 있어 5, 4, 3, 2 비트로의 압축에 대해 각각 31, 15, 7, 4레벨의 적응 양자화기가 사용된다. 역 적응 양자화기는 양자화 되었던 차분 신호를 만들어내는데, 이 신호는 입력신호의 복원된 형태를 만들어내기 위해 신호의 예측값과 더해지게 된다. 그 결과로 만들어지는 복원된 신호와, 양자화 되었던 차분신호는 적응예측기에 의해서 입력신호를 예측하게 되고 레환루프를 마치게 된다. 이 적응예측기는 2개의 극점과 6개의 영점을 가진 필터로, 각각의 필터계수들은 입력되는 샘플들에 의해 항상 업데이트된다.

2.1.2 ADPCM decoder

디코더는 인코더의 레환부분과 동일한 구조를 포함하며, 14bit 선형 PCM을 μ -law, A-law, 혹은 선형PCM으로 변환하는 부분 외에도 synchronous coding adjustment부분을 포함한다(그림1.(b)참조). 여기서 synchronous coding adjustment는 synchronous tandem coding시에 일어날 수 있는 점진적인 왜곡을 방지하게 되는데, 이것은 다음 ADPCM 인코딩단계에서 일어날 수 있는 양자화 왜곡을 제거하기 위해 PCM 출력값을 조절함으로써 얻어지게 된다.

2.2 16Kbps, 40Kbps dual rate ADPCM

음성신호의 에너지가 적은 영역은 원래의 신호와 복원된 신호를 비교할 때 상대적으로 적은 오차를 가지게 될 것이다. 마찬가지로 에너지가 큰 영역은 비교적 큰 오차를 가지게 된다. 이러한 점에 착안해서 에너지가 큰 영역에서는 보다 압축율을 낮게 하고 적은 영역에서는 압축율을 높게 주도록 하였다. 즉, 전자의 경우에는 40Kbps의 압축율을, 후자의 경우에는 16Kbps의 압축율을 적용하였다. 여기서 에너지의 높고 낮음에 대한 경계는 일반적인 음성신호에서 이 두 압축비를 동등하게 적용할 수 있도록, 즉 2bit 압축과 5비트 압축의 평균값인 3.5비트 수준으로 나올 수 있도록 설정하였다.

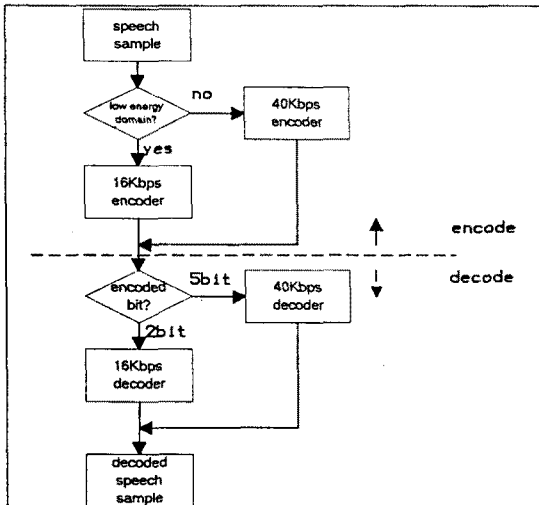


그림 2. dual rate ADPCM 처리 흐름도

그림2는 하나의 표본값에 대한 dual rate 처리 흐름도를 나타내며, 실시간 구현이나 off-line에서의 음성파일은 관계없이 처리될 수 있다.

2.3 시뮬레이션

아래 그림3은 입력 음성신호의 파형을 나타낸다. 본 신호의 입력으로는 PCM의 여러 가지 포맷 중에서 16bit linear PCM 방식으로 하였다.

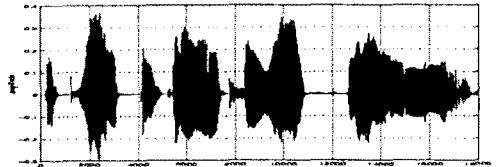
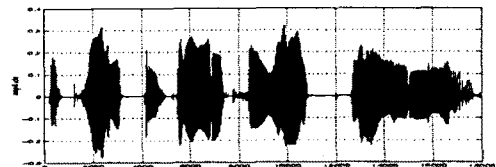


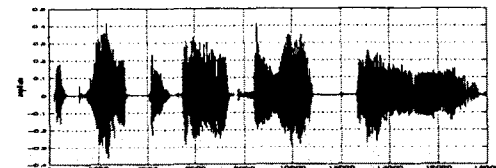
그림 3. 원래의 음성신호

아래 그림은 위에서 본 음성신호를 cut-off frequency가 1500Hz인 저역통과필터를 통과한 신호를 입력신호로 주었을 때의 파형을 나타낸다(a).

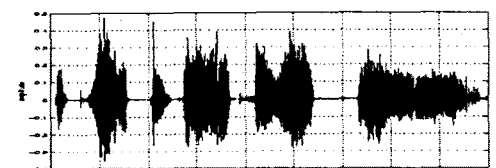
일반적으로 음성신호가 16Kbps 대 40Kbps의 비율이 5: 5의 압축비를 갖는 에너지 경계영역을 주었을 때의 dual rate codec 과(b)와, 4:6의 압축비를 갖는 에너지 경계영역을 주었을 경우의 dual rate codec결과(c)를 예로 들어 그 파형을 나타내었다.



(a) $f_c=1500\text{Hz}$ 의 저역통과필터를 통과한 입력신호



(b) 압축비 5:5의 경계영역을 준 경우의 결과파형



(c) 압축비 4:6의 경계영역을 준 경우의 결과파형

그림4. $f_c=1500\text{Hz}$ 인 경우 에너지 경계치값에 따른 결과

입력신호(a)에 대한 결과파형(b), (c)가 거의 흡사함을 알 수 있으며 그 중에서도 (b)보다는 (c)가 더 흡사함을 알 수 있다.

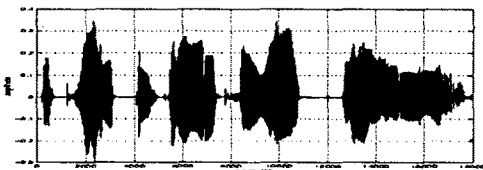
표1은 $f_c=1500\text{Hz}$ 인 경우에 위 그림에 나타난 각각의 신호에 대한 SNR(신호대 잡음비)과 평균 비트수, 그리고 16Kbps와 40Kbps의 단일 압축비를 가질 때의 결과

		16Kbps	dual rate		40Kbps
			5:5 비율의 경계치	4:6 비율의 경계치	
S N	filtering을 거치지 않은 신호	22.09	22.12	22.11	22.14
	fc=1500Hz의LP F를 거친 신호	38.36	38.51	38.64	46.38
평균 bit 수		2	3.49	3.80	5

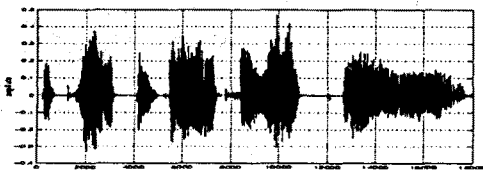
표 1. $f_c=1500\text{Hz}$ 일 때 각각의 신호에 따른 SNR

를 비교한 표이다.

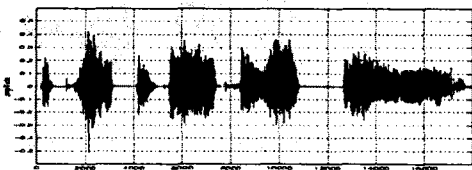
아래의 파형은 $f_c=2000\text{Hz}$ 로 한 경우에 대한 각각의 결과 파형들이다.



(a) $f_c=2000\text{Hz}$ 의 저역통과필터를 통과한 입력신호



(b) 압축비 5:5의 경계영역을 준 경우의 결과파형



(c) 압축비 4:6의 경계영역을 준 경우의 결과파형

그림5. $f_c=2000\text{Hz}$ 인 경우 에너지 경계치값에 따른 결과

표2는 $f_c=2000\text{Hz}$ 인 경우에 위 그림에 나타난 각각의 신호에 대한 SNR(신호대 잡음비)과 평균 비트수, 그리고 16Kbps와 40Kbps의 단일 압축비를 가질 때의 결과를 비교한 표이다.

		16Kbps	dual rate		40Kbps
			5:5 비율의 경계치	4:6 비율의 경계치	
S N	filtering을 거치지 않은 신호	24.27	24.33	24.33	24.37
	fc=2000Hz의LP F를 거친 신호	38.26	38.62	38.62	46.38
평균 bit 수		2	3.49	3.81	5

표 2. $f_c=2000\text{Hz}$ 일 때 각각의 신호에 따른 SNR

4. 결 론

4.1 결과 및 고찰

본 논문에서는 ITU 권고안 G.72x의 ADPCM 알고리즘을 통하여 16Kbps와 40Kbps의 dual rate ADPCM을 제안하였다.

제안된 dual rate ADPCM 구조를 소개하였고, 구체적으로 하나의 음성을 예를 들어 결과를 설명하였다. 예로 든 음성은 cut-off frequency가 1500Hz와 2000Hz인 저역통과필터를 거친 후의 신호이며 각각에 대해서 압축비가 16Kbps 대 40Kbps가 5:5를 갖는 에너지 경계 영역과 4:6의 비율을 갖는 에너지 경계 영역인 경우에 대한 결과와, 단일 비율인 경우의 결과와도 비교하였다.

filtering을 거치지 않은 신호와 각각의 single, dual rate codec 결과의 신호와의 SNR이 filtering 한 경우보다 비교적 더 적은 차이를 보였다.

향후 본 알고리즘으로 Texas Instruments(TI)사의 고정소수점 DSP인 TMS320C54 보드에서 실시간 구현할 예정이며, TI사에서 제공되는 C-compiler, 어셈블러와 링커를 사용하여 비교할 예정이다.

[참 고 문 헌]

- [1] L. R. Rabiner and R. W. Schafer "Digital Processing of Speech Signals", Prentice-Hall, 1978
- [2] Thomas P. Barnwell III, "Speech Coding A Computer Laboratory Textbook", John Willy & Sons Inc., 1996
- [3] Richard A. Jaddad, Thomas W. Parsons, "Digital Processing Theory, Applications, and Hardware", Computer Science Press, 1991
- [4] A. Nejat Ince, "Digital Speech Processing", Kluwer Academic Publishers, 1992
- [5] 진년강, "아날로그와 디지털 통신", 청문각, 1990