

배경잡음을 고려한 가변임계값 Dual Rate ADPCM 음성 CODEC 구현

양재석, 한경호
단국대학교 전기공학과

Implementation of Variable Threshold Dual Rate ADPCM Speech CODEC Considering the Background Noise

JaeSeok Yang, Kyongho Han
Department of Electrical Engineering, DanKook University

Abstract - This paper proposed variable threshold dual rate ADPCM coding method which is modified from the standard ADPCM of ITU G.726 for speech quality improvement.

The speech quality of variable threshold dual rate ADPCM is better than single rate ADPCM at noisy environment without increasing the complexity by using ZCR(Zero Crossing Rate). In this case, ZCR is used to divide input signal samples into two categories(noisy & speech). The samples with higher ZCR is categorized as the noisy region and the samples with lower ZCR is categorized as the speech region. Noisy region uses higher threshold value to be compressed by 16Kbps for reduced bit rates and the speech region uses lower threshold value to be compressed by 40Kbps for improved speech quality.

Comparing with the conventional ADPCM, which adapts the fixed coding rate, the proposed variable threshold dual rate ADPCM coding method improves noise character without increasing the bit rate. For real time applications, ZCR calculation was considered as a simple method to obtain the background noise information for preprocess of speech analysis such as FFT and the experiment showed that the simple calculation of ZCR can be used without complexity increase.

Dual rate ADPCM can decrease the amount of transferred data efficiently without increasing complexity nor reducing speech quality. Therefore result of this paper can be applied for real-time speech application such as the internet phone or VoIP.

1. 서 론

ADPCM은, PCM을 기반으로 하는 음성 압축 기법으로 ITU 권고안 G.726에서 표준으로 규정하고 있으며, 전송률에 따라 16kbps, 24kbps, 32kbps, 40kbps 등으로 구분되어 진다. 하지만 실제 음성통신에서는 순수한 음성뿐 아니라 목음(대화가 없는 구간), 잡음 등이 섞이게 되는데 이런 경우엔 우리가 원하지 않는 부분의 데이터까지도 전송하는 경우가 발생된다.

본 논문에서는 잡음이 존재하는 환경에서 되도록 좋은 음질과 적은 데이터 전송률을 가지는 음성압축 기법에 대하여 기술하였다. 또한 신호 및 잡음의 특성을 분석할 때 사용하는 FFT, Filter 등의 계산량이 많은 루틴을 대신할 수 있는 영교차율(Zero Crossing Rate)을 적용하여, 각 구간별로 구해지는 영교차율로 임계값을 가변시키고, 그 가변된 임계값 이상인 경우 40kbps, 그 이하인 경우

16kbps로 압축하는 가변임계값 dual rate ADPCM을 제안하였다.

제안한 가변임계값 dual rate ADPCM을 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 실제신호사이에서의 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.

ADPCM의 부호기에 입력된 신호가 14비트 선형 PCM으로 변환된 후, 이 신호와 예측신호를 뺀 차분신호가 얻어지게 된다. 이러한 차분신호를 디코더로 전송함에 있어 5, 4, 3, 2비트로의 압축에 대해 각각 31, 15, 7, 4 레벨의 적용 양자화기가 사용된다. 역 적용 양자화기는 양자화 되었던 차분신호를 만들어 내는데, 이 신호는 입력신호의 복원된 형태를 만들어내기 위해 신호의 예측값과 더해지게 된다. 그 결과로 만들어지는 복원된 신호와, 양자화 되었던 차분신호는 적용 예측기에 의해서 입력신호를 예측하게 되고 궤환 루프를 마치게 된다.

복호기는 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(16kbps, 24kbps, 32kbps, 40kbps)로 음성신호를 부호화할 경우, 우리가 원하던 원치 않던 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.

제안한 가변임계값 dual rate ADPCM은 좋은 음질이 필요없는 구간(목음, 잡음)은 16kbps(2비트)로 압축하여 데이터를 줄이고, 좋은 음질이 필요한 구간(음성)은 40kbps(5비트)로 압축하여 음질의 충실도를 높이고, 그 결과인 음성을 청취했을 때 40kbps로 압축한

음성과 거의 비슷한 수준의 음질을 들을 수 있다. 따라서 입력되는 음성신호의 크기로 16kbps와 40kbps로 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 주파수대에 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.

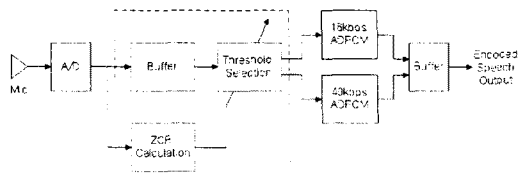


그림1 가변임계값 Dual Rate ADPCM의 구성도

그림1은 가변임계값 dual rate ADPCM의 시스템 구성도이고 점선부분이 영교차율을 계산하여 임계값(threshold value)을 가변시키는 부분을 나타낸다.

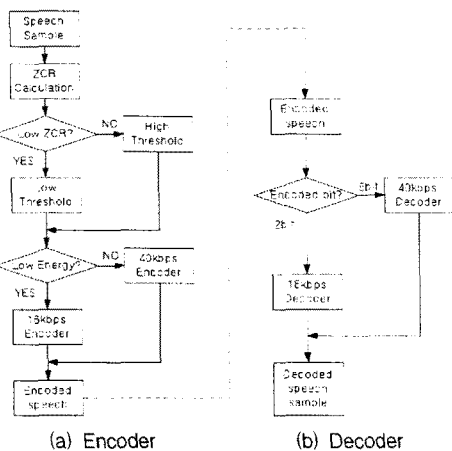


그림2 가변임계값 Dual Rate ADPCM의 순서도

그림2는 가변임계값 dual rate ADPCM의 처리 순서를 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 16kbps로, 높으면 40kbps로 음성을 전송한다. 그렇게 하므로 잡음구간은 16kbps로, 음성구간은 40kbps로 전송하여 음질의 저하없이 전송되는 데이터의 양을 줄인다.

수신단에서는 해당되는 비트 수의 복호기로 복원하게 된다.

2.2.2 영교차율 알고리즘

그림3은 영교차율을 계산하는 부분을 나타낸 것이다. 버퍼크기만큼 순환하면서 버퍼 내의 영교차율을 구하게 된다.

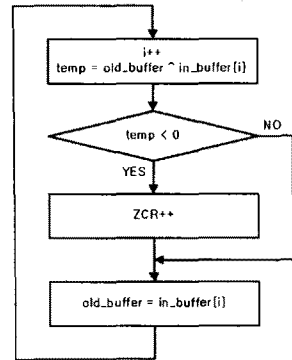


그림3 영교차율 계산 순서도

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.

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.

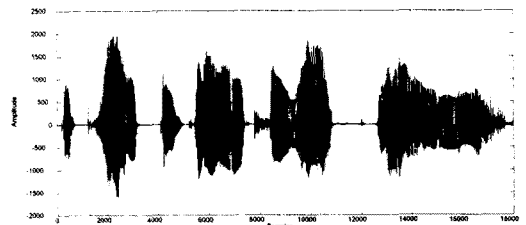


그림4 입력에 사용된 음성신호

그림4는 입력 음성 신호의 파형을 나타낸다. 또한 잡음이 포함된 환경을 설정하기 위해 원음의 음성에 백색 잡음과 도로나 지하철역에서 녹음한 잡음들을 합성하여 표본을 만들었다. 또한 음성신호의 크기보다 큰 잡음이 섞이면 분석이 불가능하기 때문에 음성신호 크기의 약 1/5로 설정하였고, 차후 DSP에 의한 실시간 처리를 위하여 18,000개의 샘플들을 80개씩 버퍼에 저장하면서 영교차율을 구하여 처리하는 방식으로 구현하였다.

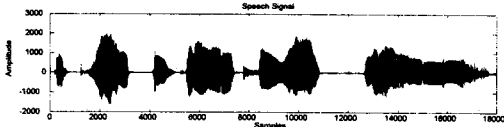


그림5 샘플음성의 영교차율

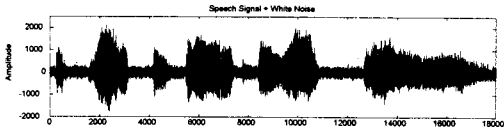


그림6 샘플음성 + 백색잡음의 영교차율

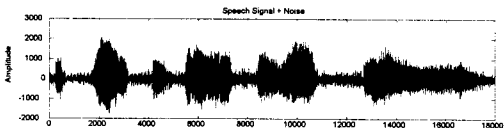


그림7 샘플음성 + 잡음의 영교차율

그림5~7은 각 샘플의 영교차율을 구한 것이다. 그림들을 보면 알 수 있듯이 음성신호에 잡음이 섞이게 되면 목음영역에서 영교차율이 급격히 상승한다. 따라서 제안한 영교차율로써 잡음구간과 음성구간을 구분할 수 있다.

표1에서 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 ± 250 로, 음성구간은 ± 150 으로 설정하여 시뮬레이션 하였다.

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

표1 표준압축 알고리즘과 가변임계값 dual rate ADPCM의 비교

가변임계값 dual rate ADPCM으로 압축하여 전송한 음성은 40kbps 단일 압축으로 전송한 것보다는 SNR이 떨어지게 된다. 그 이유는 잡음 구간으로 판정된 부분을 16kbps로, 음성구간은 40kbps로 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 16kbps와 40kbps의 가변임계값 dual rate ADPCM 부호화기법을 제안하였다.

제안한 알고리즘이 16kbps와 40kbps 각각의 단일 ADPCM보다 압축률 및 부호화된 음질의 양면에서 우수함을 시뮬레이션 및 직접 들어봄으로써 확인하였다. 또한 영교차율을 제시함으로써 잡음구간을 분석해낼 수 있는 복잡한 과정을 단순하게 줄임으로 계산량과 수행 속도를 현저하게 줄일 수 있었다.

실제 전화 통화 시에는 전체 대화시간 중에 목음이 차지하는 비율이 80% 정도이므로 조용할 때 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 줄임으로 수행속도 또한 빠르게 할 수 있다.

(참 고 문 헌)

- [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, 1990