

배경잡음을 고려한 4배 가변 압축률을 갖는 ADPCM의 C6000 DSP 실시간 구현

김대성, 한경호
단국대학교 전기공학과

Implementation of Quad Variable Rates ADPCM Speech CODEC on C6000 DSP considering the Environmental Noise

Dae-Sung Kim, Kyong-ho Han
Department of Electrical Engineering, Dankook University.

ABSTRACT

In this paper, we proposed quad variable rates ADPCM coding method and its implementation on C6000 DSP, which is modified from the standard ADPCM of ITU G.726 for speech quality improvement considering the environmental noise.

Four coding rates, 16Kbps, 24Kbps, 32Kbps and 40Kbps are used for speech window samples and the rate decision threshold is decided by the environmental noise level. The object of the proposed method is to reduce the coding rate while retaining the speech quality and the speech quality is considerably close to 40Kbps single rate coder with the coding rate close to 16Kbps single rate coder under the environmental noise.

The environmental noise level affects the coding rate and the noise level is calculated per every speech window samples. At high noise level, more samples are coded at higher rates to enhance the quality, but at low noise level, only the big speech signals are coded at higher rates and more speech samples are coded at lower coding rates to reduce the coding rates. The influence of the noise on the speech signal is considerably high for small signals and the small signal has the higher ZCR (zero crossing rate). The method is simulated in PC and to be implemented on C6000 floating point DSP board in real time operations.

1. 서론

음성의 디지털 부호화 기술은 크게 파형 부호화

(Waveform coding) 방식과 계수 부호화(Parameter coding) 방식으로 분류되어진다. 파형 부호화 방식은 계수 부호화 방식에 비해 전송 데이터 양이 크지만 음질이 우수하고 복잡성이 낮다는 장점을 가지고 있어서 일반 음성 통신에 많이 사용되고 있다. 이러한 파형 부호화의 한가지 방법인 ADPCM은 ITU G.726에서 표준화되어 있는 PCM을 기반으로 하는 음성 압축 기법이고 전송률에 따라 16Kbps, 24Kbps, 32Kbps, 40Kbps의 4가지의 전송률을 가진다.

실제 음성통신에서 실제 전송되는 데이터에는 순수한 음성뿐만 아니라 대화가 없는 동안의 데이터와 잡음이 섞이게 된 음성 데이터와 같이 우리가 원하지 않는 부분의 데이터까지 전송하게 된다. 본 논문에서는 잡음이 존재하는 환경에서 되도록 좋은 음질과 적은 데이터 전송률을 가지는 음성압축 기법에 대해 기술하였다. 그리고 음성데이터에 포함되어 있는 신호와 잡음의 특성을 분석할 때 사용하는 FFT, Filter 등의 계산량이 많은 루틴을 대신할 수 있는 영 교차율(Zero Crossing Rate)을 적용하여 각각의 구간별의 임계값을 바꿔서 각각의 전송률로 압축하는 4배 가변 압축률을 갖는 ADPCM을 제안하였다.

2. 본론

2.1 ADPCM

ADM 방식은 양자화기의 스텝 크기를 적응적으로 변화시키는 적응 양자화 방법을 사용하고 있으며, DPCM에서는 예측기를 사용한 예측 부호화 방법을 사용하므로 이들의 기본적인 개념을 조합하여, 적응 양자화와 예측 부호화 개념을 동시에 사용하는 방식으로, 음성신호의 경우 표본화된 펄스 간 서로 상관성이 크다. 현재 디지털 통신을 위해

사용중인 PCM의 전송속도는 64Kb/s로써, 이것은 대역폭의 사용 면에서 볼 때 아날로그의 SSB방식의 4KHz에 비해 비경제적이다. 이 같은 대역폭 문제를 해결하기 위해 음성대역폭 축소에 대한 예측

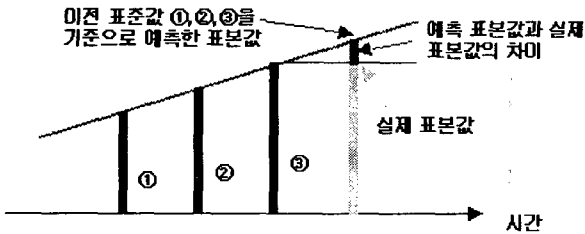


그림 1 ADPCM의 기본 원리

부호화(Predictive Coding)방식이 ADPCM 방식이다. ADPCM 예측 부호화 방식의 기본원리는 음성신호가 인접한 표본값 사이의 상관성이 큰 특성을 이용하여 음성신호를 직접 양자화하지 않고, 과거의 음성신호의 Sample을 기준으로 다음에 들어올 신호의 크기를 예측하고 실제의 입력 신호로부터 빼줌으로써 오차 신호를 발생시켜 이 오차 신호를 양자화해서 전송한다. 보통 오차 신호의 진폭은 입력 음성신호의 진폭에 비해 훨씬 더 작기 때문에 그만큼 양자화 레벨의 수도 감소되어 동일한 성능을 갖게될 경우, 전송속도를 PCM에 비해 약1/3 정도로 감소시킬 수 있다. (8비트 부호화 시 64Kbps 속도, 3비트 부호화 시 24Kbps)

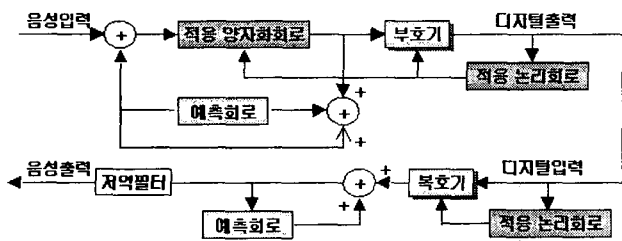


그림 2 ADPCM 시스템 구성도

ADPCM의 부호화기는 크게 양자화기/부호화기(Quantizer / Coder)와 예측기(Predictor)등 두 개의 Subsystem으로 분류할 수 있다. 본래 예측기(Predictor)는 예측필터의 형성을 위한 계수를 고정시킬 수도 있고, 입력 신호에 따라 변화시킬 수도 있는데, ADPCM방식에서는 고정 예측기를 선택하고 있다. 일반적으로 예측기를 형성하는 Filter의 계수와 양자화기의 양자화 레벨의 크기를 결정할 때 음성신호가 Stationary 상태라고 가정하고 있으

나, 음성신호의 특성으로 인해 실제로는 고정상태가 아니므로 양자화기에 입력되는 오차 신호가 급격히 변화하여 Slope Over Load현상이 발생하는 경우가 있다. ADPCM 방식은 Non-Stationary 상태에 있는 음성 신호를 국부적으로 Stationary상태로 간주할 수 있는 구간으로 분할해서 신호의 통계적 특성을 구한 후, 그에 따라 적응 예측방식이나, 적응 양자화 방식을 적용한다. 적응 양자화 방식은 양자화의 레벨의 크기를 변화시키는 것을 말하며, 적응 예측 방식은 양자화의 레벨의 크기를 고정시킨 상태에서 예측기의 Filter계수를 입력되는 신호에 대응해서 변화시키는 것을 말한다. 적응 양자화기에서 입력되는 신호의 Dynamic Range가 크면 양자화 레벨의 크기를 입력되는 신호의 크기에 비례해서 증가시킨다. 그리고 원래 신호와 예측신호 사이의 오차 신호의 분산은 원래 신호의 분산에 비례하는 성질을 이용하여 적응 예측기의 filter 계수를 입력 신호의 구간 변화 특성에 따라 대응시킴으로써 오차 신호의 분산을 작게 할 수 있다. 이와 같이 ADPCM은 32Kbps 디지털 음성 통신에 전송속도가 기존PCM의 1/2밖에 되지 않는 반면음질이 양호하여 경제적 이여서 앞으로 디지털 음성 통신에 중요한 역할을 할 것임은 틀림없는 사실이다.

2.2 가변 임계값 Quad Rate ADPCM

기존의 일정한 압축비(16Kbps, 24Kbps, 32Kbps, 40Kbps)로 음성신호를 부호화할 경우에는 낮은 압축비로 압축을 하면 전송 데이터는 작으나 음질이 낮아졌고 높은 압축비로 압축을 하게되면 음질은 좋으나 전송 데이터가 커져서 전송 데이터 양과 음질 두 가지를 만족시킬 수 없게 된다. 하지만 4가지의 압축비율을 사용하게 되면 좋은 음질을 필요로 하는 구간(실제 음성)에는 높은 압축비로 압축하고 좋은 음질이 필요하지 않은 구간(대화가 없는 동안, 잡음)에는 낮은 압축비로 압축을 하여 전송 데이터 양은 줄이고 음질은 40Kbps로 압축한 음성과 거의 유사한 음질을 얻을 수 있게 된다. 이러한 기능을 수행하기 위해서는 입력되는 음성신호의

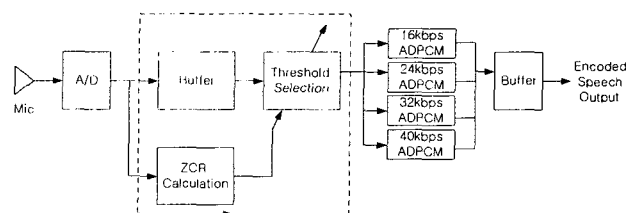


그림 3 가변임계값 Quad Rate ADPCM의 구성도

크기에 따라서 4가지 압축률로 모드를 변환해주는 임계값이 필요하고 또한 임계값의 크기를 적절하게 가변 할 수 있는 기준이 필요한데 이 값을 영 교차율을 이용하여 구한다.

2.2.1 영교차율(Zero Crossing Rate)

주위의 소리들은 다양한 주파수대에 분포하고 있는데 잡음은 대부분 음성신호 보다 높은 주파수에 분포하고 있다. 음성신호는 주위 잡음의 주파수보다 낮은 영역의 주파수대에 분포하고 있으므로 음성신호에 잡음이 섞이게 되면 주파수가 증가하게 된다. 결국 이것은 상대적으로 영교차율이 증가하는 결과를 얻게 되고 이 결과로 음성 및 잡음이 존재하는 구간을 구분해 낼 수 있다. 그러므로 음성 통화에서 사람의 목소리는 저주파의 큰 신호이므로 잡음이 섞이더라도 영교차율이 낮고, 대화가 없는 구간에서는 음성구간과 비교할 때 영교차율이 높아지게 된다. 이러한 영교차율의 값을 이용해서 ADPCM의 임계값을 변화시킨다면 잡음이 섞여있는 환경에서도 원음에 충실하고 전송되는 데이터의 양이 작은 신호를 생성할 수 있게 된다.

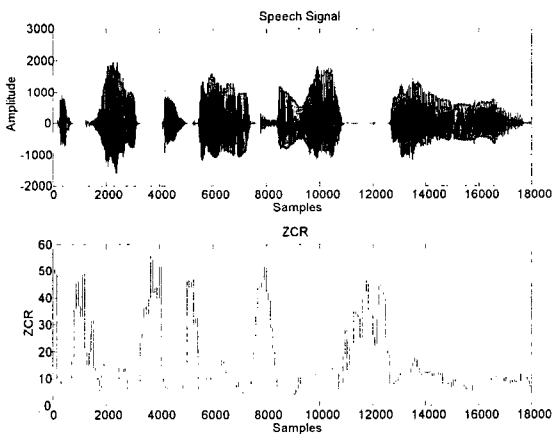


그림 4 음성 파형에 따른 ZCR값

이러한 영교차율을 이용하게 되면 신호의 특성을 분석할 때 기존의 FFT 및 각종 필터를 사용하는 것이 아니므로 계산량을 현저하게 줄일 수 있고 입력되는 신호의 주파수가 높고 낮음의 검출에 있어서 빠른 계산속도를 얻을 수 있어 실시간 구현이 가능해진다. 하지만 영교차율을 적용시켜서 계산을 수행하게 되면 신호 샘플사이의 부호 변화만 감지하기 때문에 잡음의 크기나 정확한 특성들은 알 수 없는 단점이 있다.

3. 결 론

본 논문에서는 잡음이 존재하는 환경에서 원음의

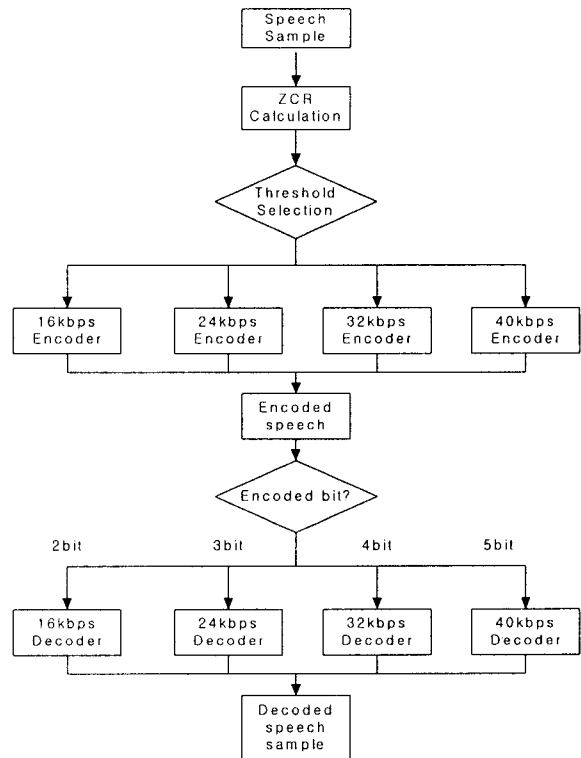


그림 5 가변 임계값 quad Rate ADPCM의 순서도

음질에 충실하면서 데이터 전송량을 줄일 수 있는 4배 가변압축률을 갖는 가변 ADPCM기법을 제안하였다. 또한 영교차율을 제시함으로써 잡음구간을 분석해내는 복잡한 과정을 단순하게 줄여 수행속도를 줄일 수 있는 방법을 제안하였다.

실제 전화 통화 시에는 전체 대화시간 중에 묵음이 차지하는 비율이 80% 정도이므로 조용할 때나 잡음이 있을 때를 구분하여 압축한다면 적은 양의 데이터를 전송하면서 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] Thomas P. Barnwell III, "Speech Coding A Computer Laboratory Textbook", John Wiley & Sons Inc., 1996
- [4] Richard A. Jaddad and Thomas W. Parsons, "Digital Processing Theory", Applications and Hardware, Computer Science Press, 1991
- [5] L. R. Rabiner and R. W. Schafer, "Digital Processing of Speech Signals", Prentice-Hall, 1978.