

CELP코더의 성능개선 및 실시간구현

정 창경* °임 상균**

*명지대학교 전기공학과 교수, 공학박사 **명지대학교 전기공학과 대학원 석사과정

Performance improvement and Realtime implementation in CELP Coder

Changkyung Chung* Sangkyun IM**

*Department of Electrical-Eng., Myong-Ji University **Department of Electrical-Eng., Myong-Ji University

ABSTRACT

In this paper, we researched about CELP(Code-Excited Linear Prediction) Speech coding algorithm using efficient pseudo-stochastic block codes, adaptive-codebook and improved fixed-gain codebook. The pseudo-stochastic block codes refer to stochastically populated block codes in which the adjacent codewords in an innovation codebook are non-independent. The adaptive-codebook was made with previous prediction speech data by storage-shift register. This CELP coding algorithm enables the coding of toll quality speech at bit rates from 4.8 kbits/s to 9.6 kbits/s. This algorithm was realized TMS320C30 microprocessor in realtime.

요 약

본 논문에서는 효과적인 pseudo-stochastic 블록코드와 적응코드(Adaptive-codebook), 개선된 고정이득코드북을 이용한 CELP(Code-Excited Linear Prediction)음성코딩 알고리즘에 관한 연구와 이 알고리즘의 실시간 구현을 본 논문의 목표로 한다. 이를 위해서 TMS320C30 마이크로 칩을 사용하여 구현하여 보았다. 이러한 CELP코딩 알고리즘은 중저의 전송율에서 Toll-quality를 얻을 수 있다.

제 1 장 서론

인간의 가장 기본적인 의사전달 수단인 음성으로 정보화 사회의 정보교환과 지역적인 불리함과 다양한 서비스의 제공을 전문적인 지식용 요함이 아니라 단지 본인의 의사를 전달하므로 이러한 요구사항을 충족하고자하는 음성통신 및 음성 관련된 응용분야는 무궁무진하다. 또한 음성의 통신만으로는 만족하지 않고 화상과 음성의 접목을 꾀하는 분야와 최근 관심의 초점이 되는 이동통신분야는 한정된 채널의 문제를 극복하고 저장 및 전송시 많은 비용이 요구되는 문제의 해결 방안으로 음성의 압축기술이 사용된다.

음성의 압축 기술은 음성 신호의 중복성(Redundancy) 제거와 그 음성의 특징이 되는 파라미터를 추출하여 전송하는 방법으로 음성 데이터량을 줄이고, toll-quality의 보존을 위해서 분석과 합성의 방법을 사용한다. 이런 방법으로 음성전송시 전송선의 효율을 높여주고, 저장시 필요한 기억용량을 줄여준다. 음성의 압축방법은 파형부호화 방법, 음원 부호화 방법, 복합부호화 방법으로 나누어진다. 파형부호화는 계산량이 적고, 32kbits/s - 64kbits/s에서는 양질의 음질을 얻을 수 있고, 단점으로는 많은 전송채널, 저장용량이 필요하고 전송율이 중저(16kbits/s이하)에서는 음질의 급격한 저하를 나타낸다. 음원부호화의 경우는 장점으로는 중저의 전송율에 적합함

으로 효율적인 전송채널의 사용과 적은 저장용량의 필요를 요하고, 단점으로는 많은 계산량을 요구하고 과형부호화보다는 음질 면에서 떨어진다. 이런 장단점을 고려하여 장점을 취하는 방법으로 복합부호화방법이 제시되었다. 복합부호화방법의 장점은 분석과 합성을 행함으로써 양질의 음질을 유지하고 음원부호화의 음성의 특징 파라미터만의 전송으로 전송율은 중저의 전송율로 만족한다.

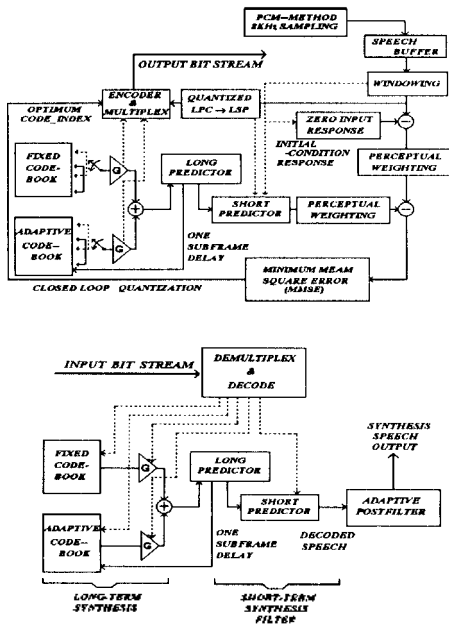
이 복합부호화방법의 예는 RPE(Regular Pulse Excitation), MPE(Multi Pulse Excitation), CELP(Code Excited Linear Prediction), LDCELP(Low-Delay CELP) 등이 있다. 위의 복합부호화 방법은 최근 이동통신분야에서 사용되고 있으며, 또한 비디오콘등에서도 사용되고 있다.

본 논문에서는 기본적인 CELP알고리즘에 계산량의 감축, 전송량의 감축, 적응코드북, 개선된 고정이득코드북, LPC를 사용한 방법을 통해 먼저 그 성능을 알아보고 그 후는 LSP 방법의 사용으로 보다 적은 왜곡의 방지를 사용했으며, 실시간의 구현을 목적으로 병렬처리와 독립적인 DMA의 사용과 처리속도는 대형기종에 버금가는 TMS320C30 마이크로 칩을 사용하여 실시간 처리를 행하였다.

제 2 장 CELP (Code Excited Linear Prediction)

다음은 본 논문에서 개선된 부호화기와 복호기의 블록도이다.

아래의 그림과 같다.



구성블럭으로는 고정코드북(Pseudo-stochastic codebook), 장기예측기(Long-term predictor), 단기예측기(Short-term predictor), 가중필터(Perceptual-weighting filter), 적응코드북(Adaptive-codebook), 고정이득코드북(gain-index-codebook), 양자화된 파라미터 코드북(quantized-LSP-parameter codebook) 등으로 구성되어 있다. 또한 가중필터의 분리로써 단기예측기의 계, 초기값에 의한 출력을 원신호로부터 제거함으로써 분석과 합성방법에서 더욱 간단한 수학적 복잡함을 구현한다.

2.1 CELP 시스템의 구성요소

개선된 CELP 코딩에서 코드북(Pseudo-stochastic codebook)은 음성처리의 부분프레임 크기인 5ms(40 샘플)을 하나의 디멘전으로 형성하고 고정코드북 크기는 9bit(512)를 사용하였고, 고정코드북의 설계는 random white Gaussian 분포를 갖는 잉여신호를 중앙에서 확률적으로 발생빈도가 높은 부분을 뽑아냄으로써 코드북을 구성하였다. 이 근거는 이상적으로 모든 음성에 대해서 분석을 시행했을 때 근사적으로 잉여신호의 분포는 Gaussian 분포에 근사화 된다는 것이다. 단기예측기(Spectrum envelope)는 음성의 처리프레임인 부분프레임(5ms)에 대해서 음성의 율곽을 재생하고, 음성의 중복성을 재생한다. 전형적으로 10차 LPC 파라미터를 사용한다. 또한 적은 저장량을 필요로 한다(2ms이하). 장기예측기(fine structure)는 음성의 처리 부분프레임에 대해서 피치를 찾아내는 역할을 한다. 정확한 의미에서는 유사주기를 찾는 형태가 된다. 일반적인 필터 차수는 3차이나 1차의 필터를 사용하여도 성능의 차이는 거의 없으므로 전형적으로 1차 필터를 사용한다.

또한 큰 저장량을 원한다(2 - 16ms). 적응코드북(Adaptive codebook)은 음성의 유사주기를 찾기 위해서 형성되는 코드북으로 고정되어 있는 것이 아니고 음성의 분석과 합성시에 생성되는 코드북으로 전형적인 크기는 128개의 정수 값과 128개의 비정수형 값을 가져 전체 256개(8 bits)의 코드북 인덱스를 갖는다. 가중필터(Perceptual-weighting filter)는 합성음과 원음사이의 오차에 대해서 다른 가중값을 줌으로써 인간의 청각에 보다 정확한 인식을 행할 수 있게 하는 역할을 행한다.

2.2 CELP시스템의 분석과 합성

2.2.1 장기예측기 (Long-term Predictor)

장기예측기는 피치를 검색하는 예측필터이고, 정확한 의미는 유사음에 대한 유사주기를 찾는 것이다. 피치의 검색은 과거의 합성음에서 구해야 하므로 큰 메모리(2 - 13.375ms)를 필요로 하고, 이 필터차수는 일반적인 형태는 3차의 필터로 구성되나 전형적으로는 1차 필터를 사용한다. 그 이유는 결과의 누드러진 차이가 없고, 계산량의 감소를 위해서 1차 필터를 사용한다. 장기예측기를 식으로 표현하면 아래와 같다.

$$1/P(z) = 1/(1 - \sum_{i=0}^r b_i z^{-(D+i)})$$

단: $q = 1$ and $r = 3$: 3차필터

$q = 0$ and $r = 0$: 1차필터

여기서 자연을 나타내는 D가 피치 주기를 나타내고, $b(i)$ 는 장기예측기의 필터계수를 나타낸다. 이 필터계수는 처리 부분프레임에 대해서 매번 계산되어 전송되어지고, 피치의 주기를 나타내는 D도 처리 부분프레임에 대해서 매번 계산되어지고 전송된다. 즉 프레임에 부분프레임이 4개가 포함되므로 각과 4개의 필터계수와 피치주기가 전송되어진다. 필터의 피치주기와 필터계수를 구하는 식은 아래와 같은 식으로 표현된다.

$$Y_D(n) = V(n) + bY_D(n - D)$$

영입력에대해서 ; $V(n) = 0$

$$Y_D(n) = bY_D(n - D)$$

최소의오차를구하는것은 ;

$$E(D, b) = \sum_n [e^{(0)}(n) - bY_D(n)]^2$$

단 : $e^{(0)} = Z - W_0$

Z = 가중되진 원음성프레임

W_0 = 가중되진 합성음

$d_D(n)$ = 지연된 과거의합성음

단 : 20샘플 ≤ D ≤ 147샘플

2.5 $ms \leq D \leq 18.375 ms$

E(D,b)은 모든 가능한 구간에 대해서 모두 적용되어지고 그 방법은 개루우프 구조와 페루우프 구조가 있으며 본 방법은 페루우프 구조를 사용했고, 그 때의 최적 값을 선택하여 피치 주기인 D의 값과 필터계수인 b(i)값을 결정하고, 이 값들을 채널에 전송하므로 장기예측기의 역할은 끝난다. 그리고 다음에 처리할 처리 부분프레임을 위해서 하나의 처리 부분프레임은 자연 시킨다. 위에서의 개루우프 구조와 페루우프 구조의 차이점은 개루우프 구조는 선택된 결과에 대해서 부분적으로 최적이라는 보장을 할 수 없으나, 속도는 페루우프 구조보다는 빠른 장점이 있다. 반면에 페루우프 구조는 순환적인 방법이므로 선택되어진 값이 부분적으로 최적임을 보장할 수 있으나, 계산량이 많고 고정코드북과 적응코드북의 크기에 의해서 계산량이 결정된다. 본 논문에서는 페루우프 구조를 사용하였으며, 이 때의 오차의 계산은 MMSE(Minimum Mean Square Error Method)를 사용하였고, 매 부분프레임에서 전체가 가능한 피치주기를 매번 적용시켜 최적의 지연인덱스와 최적의 필터 계수를 구했다. 이러한 선형예측기는 시간축 상에서는 음성의 주기적인 특성을 결정하는 필터의 설정을 의미하지만 주파수 영역에서의 의미는 음성화자의 음색을 나타내는 주파수의 공진을 나타냄으로 화자특성의 주파수 특성을 나타냄을 의미한다.

2.2.2 단기 예측기 (Short-term predictor)

이 필터의 역할은 제거된 중첩성을 재생(Spectral envelope)시키는 역할을 한다. 이 말은 음성의 특징파라미터에 여기신호를 공진시킴으로서 원음성의 정보를 재생하는 의미를 나타내고, 본 논문에서는 10차 LPC 파라미터를 사용하였으며, 그 이유는 실험상으로 그 이상의 파라미터에 대해서 크게 음질의 변화가 없었으며 응용의 예가 실시간의 구현을 목적으로 함에따른 전송율의 제한에 의해서 10차의 AR(Autoregressive filter = all-pole model filter)를 사용하였다. CELP 코딩의 단점이자 장점인 LPC방법에서는 무성음과 유성음의 주기를 잡아주어야 하고, 유성음에 대해서는 피치주기를 찾아야 하고, 각각의 부분에 대해서 에너지를 측정해야한다. 이 방법에서 유성음과 무성음의 구간을 찾는 방법으로 가장 쉬운 방법은 음성의 에너지 분포와 영점 교차의 빈도 수를 계산함으로써 그 교점을 추출하는 방법으로 유성음과 무성음의 구분을 행하나, 본 CELP의 경우는 유성음과 무성음의 구분없이 모든 부분프레임에 대해서 피치주기를 추출하고 모든 처리 부분프레임에 대해서 맥색자음의 여기신호를 가한다. 즉 이 이유가 가장 큰 특징이 될 수 있고, LPC(Linear Prediction Code)에서는 preemphasis 와 deemphasis 을 행하여줘야 고주파로 갈수록의 음성의 방사전향에 의한 스펙트럼의 왜곡을 방지할 수 있으나, 본 방법은 이러한 선처리과정을 하지 않고 LPC(Linear Prediction Code)파라미터를 구한다. 또한 양자화(Quantization)의 비트수가 10비트 이하로 양자화(Quan-

tization)가 될 때는 왜곡(disturbance)이 커진다. 이를 보완하기 위해서 양자화시 LSP(Line Spectrum Pair)파라미터를 이용한다.

2.2.3 장기예측기와 단기예측기의 전송

장기예측기의 전송 데이터는 피치지연과 피치지연에 따르는 이득을 전송하고, 피치지연은 각 처리 부분프레임마다 1.3프레임은 8비트씩 전송하고 2.4프레임은 1.3프레임에 대한 상대적인 오프셋으로 전송하고, 각각 6비트씩, (단:프레임에는 처리 부분프레임이 4개 있다) 또한 장기예측기는 처리 부분프레임마다 갱신됨을 의미하므로 시간변화형 필터이다. 단기예측기의 경우는 프레임 길이 동안에 필터 파라미터는 한번 갱신되고, 전송 데이터는 10차의 LPC(Linear Prediction code)의 파라미터이고, 비트 수는 1프레임당 최대 10비트이다.

2.2.4 가중필터 (Perceptual Weighting filter)

S(n)이 원음성을 나타내고, S'(n)이 합성음을 나타낼 때, 이 두 신호간의 오차에 대한 에너지를 최소화하는 필터를 설계하는 것이 목적이다.

이 목적에 부합하여 오차에너지의 분포를 인간적인 관점에 조화시키는 것이 오차에너지 분포에 대해서 동일한 가중값을 주는 것이 아니고 각각에 다른 가중으로서 오차에너지를 최소화하는 방안이 가중필터의 목적이다. 이 말은 음성의 에너지 분포가 밀집한 부분에서의 오차는 인간청각에서는 큰 영향을 미치지 못하고 인간청각의 민감한 주파수 영역에서는 상대적으로 작은 오차에너지 일지라도 청각적으로 큰 오차를 나타냄을 이용한 것이 가중필터의 기본적인 접근방식이다. 오차에너지는 아래의 식으로 주어진다.

$$E(n) = [S(n) - S'(n)] * W(n)$$

위식에서 W(n)은 가중함수를 나타내고, 그와의 오차의 대수합(convolution)으로 오차를 표현할 수 있다. 위식을 2 영역으로 혹은 주파수 영역으로 변환하면 아래의 식으로 변환된다.

$$E(z) = [S(z) - S'(z)]W(z)$$

전형적인 CELP에서의 가중필터는 아래의 식으로 표현이 가능하다.

$$W(z) = \frac{1 - Q(z)}{1 - Q(z/r)}$$

단; $Q(z) = \sum_{i=1}^p a_i z^{-i}$

$$Q(z/r) = \sum_{i=1}^p \gamma^i a_i z^{-i}$$

단 : a(i) = LPC필터의 계수

p = LPC필터의 차수

r = 가중필터의 대역을 결정하는 파라미터

전형적으로 $0 < r \leq 1$

위식에서의 파라미터 변수(r)는 하나가 있고, 이 변수(r)의 전형적인 값으로는 0.8 값을 사용한다. 그러나 본 논문에

서는 이 변수를 변화시키면서 성능을 실험해 본 결과 최적의 변수 값은 0.9라는 실험 결과를 얻을 수 있었다.

본 논문의 최적의 인덱스와 최적의 이득을 찾는 기준은 이 가중필터를 거친 오차의 에너지 값을 가지고 가장 적은 오차에너지를 갖는 인덱스와 이득을 구하는 것이며, 그 값을 전송하였다.

제 3 장 적응코드북 탐색과 설계 및 계산량 감축
(Adaptive codebook design algorithm and Adaptive codebook searching method)

적용코드북은 고정코드북과는 다른 형태로 매 처리 부분 프레임별로 적응코드북은 갱신된다. 이러한 적응코드북의 주요 목적은 음성의 유사주기를 찾는 것으로 화자의 음색을 결정하는 중요한 역할을 한다.

3.1 설계

적용코드북의 크기는 8비트로 할당하고 있어, 256가지의 경우의 수가 발생하고, 여기에서 128개는 정수 피치지연을 갖는 부분이고, 나머지의 128개는 비정수 피치지연을 갖는 부분으로 구분이 된다. 먼저 정수 피치지연은 과거의 여기신호를 저장하는 형태에서 현재 처리할 처리 부분프레임의 가장 유사한 피치를 나타내는 피치지연을 찾는 부분이다. 기본적인 CELP 코딩에서는 이 정수 피치지연만을 가지고 현재 화자의 음색을 결정하는 피치를 결정했으나, 본 논문은 더욱 정확한 피치를 찾기위한 방법으로 비정수 피치지연까지를 고려하고, 이 방법은 정수형의 지연사이에 보간을 사용하여 5개의 새로운 비정수형의 지연을 생성했으며, 이 비정수형의 여기신호는 과거의 여기신호에서 주위의 여기신호에 Hamming 윈도우적인 가중을 곱해서 비정수적인 여기신호를 생성하고있다.

이때의 비정수적인 지연의 값은 0.25, 0.33, 0.5, 0.66, 0.75 와 같은 값으로 표시된다. 이 비정수 피치지연은 정수형의 여기신호로 적응코드북이 갱신된 후에 가중값에 의해서 발생된다. 적응코드북의 갱신은 매 부분처리 프레임마다 행해지며, 실제적으로는 Queue의 형태로 가장 오래된 여기신호는 제거된다. 피치지연의 범위는 여성화자의 경우는 음성이 높으므로 피치의 간격이 남성화자에 비해서 짧은 반복주기를 갖기 때문에 원음성의 8 (kHz) 샘플링시에 가장 짧은 피치주기를 20샘플로 가정하고, 남성화자의 경우는 이와는 반대로 피치주기가 여성화자보다는 긴 편이므로 가장 긴 피치주기를 147샘플로 가정하여 본 논문의 피치 탐색구간은 20샘플에서 147샘플까지로 피치탐색을 행했다. 단 전송은 매 처리 부분프레임마다 적응코드북의 인덱스(즉: 8비트)를 전송하면 되고, 전송 비트를 더욱 줄이는 방법으로 2.4 처리 부분프레임에서는 이전의 처리 부분프레임의 인덱스에 대한 오묘셋 값으로 그 인덱스를 전송한다.

3.2 계산량 감축을 위한 알고리즘

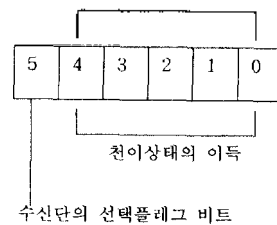
계산량의 감축은 회박중점형의 고정코드북 사용으로 메모리의 절감과 탐색의 과정에서 위에서 언급했듯이 매 부

분처리 프레임에서 40개의 여기신호중에 7개의 영이 아닌 값을 가지고 있으므로 이것을 먼저 판단하므로써 계산량을 줄일 수 있고, 적응코드북의 경우에서도 코드북의 형태가 중점형을 사용하고 있으므로 메모리의 절감을 꾀하고있다. 연산속도의 증가는 본 논문에서 사용한 TMS320C30 DSP칩은 곱셈을 빠르게 처리할 수 있도록 설계되어 있어, 본 논문의 모든 필터의 설계를 매트릭스 형태로 하여 곱셈만으로 필터의 출력을 얻을 수 있게 하였다.

제 4 장 개선된 이득코드북의 설계

본 논문에서는 이득코드북의 VQ(Vector Quantization)방법에서 잘 알려진 LBG알고리즘과 K-means알고리즘을 사용하였다. 그러나 음성의 처리에서 이 알고리즘으로 양자화된 코드북은 무성음에서 유성음으로의 천이에서 발생하는 큰 이득 값은 제거되어 양자화 됨으로 SNR(dB)의 큰 저하 요인으로 작용하기 때문에 정상상태에서의 고정이득코드북과 천이상태의 고정이득코드북으로 구성하여 SNR(dB)의 저하를 방지할 수 있었다. 고정이득 코드북의 크기는 정상상태의 이득코드북 5비트와 천이상태의 이득코드북 5비트와 수신단에서 이 두개의 코드북의 선택을 하는 플래그 비트인 1비트를 갖는다. VQ 알고리즘의 실험적인 한계값을 줄이므로써 최적의 대표값을 선택할 수 있었다.

정상상태의 이득



제 5 장 CELP알고리즘의 PC simulation 및 TMS320C30을 이용한 실시간 구현

5.1 CELP알고리즘의 PC상에서의 Simulation방법

이 방법은 시간적인 제약은 없고, 음성 데이터는 8 (kHz) 샘플링의 PCM형식으로 코딩화된 것을 사용한다. 그러나 디코더에서는 전송된 비트열로 실시간 복원은 가능했다.

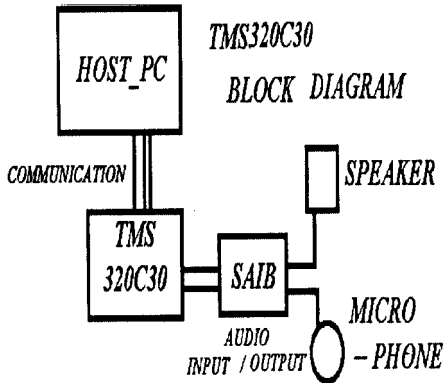
5.2 TMS320C30을 이용한 하드웨어의 구성

5.2.1 Host와 통신으로 구성

시스템의 음성입력은 SAIB(Stereo Audio Interface Box)를 사용하여 음성을 8 (kHz) 샘플링과 12비트의 ADC(Analog to Digital Converter)를 사용하여 PCM코딩을 한 음성을 입력으로 사용하여 TMS320C30 보드상에서 본 논문의 알고리즘을 사용하여 입력음성을 비트열로 압축한 후 디코더에서 이 비트열을 가지고 원음성을 복원한 후 SAIB(Stereo Audio Interface box)의 출력단으로 복원음을 출력하였다. 단 이방법은 계속적인 Host와의 통신으로 현재의 처리를 알림으로 최

적의 처리성능을 나타내지는 못했다. 이것은 DSP프로세서에 대한 계속적인 기다림이 발생됨을 의미한다.

아래의 그림은 대략적인 시스템 블록도이다.



5.2.1.1 처리시간 개선

첫째, DSP보드상의 처리시간을 개선하기 위해서 프로그램의 함수호출이 잦은 함수는 TMS320C30 cache메모리에 함수를 호출한 후 계속적으로 반복 사용하였다. 이 방법으로는 메모리의 호출시간을 크게 줄일 수 있었다.

둘째, HOST와의 통신에서는 계속적인 기다림이 아닌 가로채기(interrupt)를 사용하여 HOST와 DSP보드상의 통신조건을 크게 개선할 수 있어서 처리시간을 줄일 수 있었다.

5.2.2 독립적인 TMS320C30 하드웨어의 구성

이 방법은 위에서 사용한 방법과 동일하나 Host와의 통신은 시스템의 시작시에만 통신을 하고 그 후는 TMS320C30 보드상의 ROM(Read_Only_Memory)에 실린 명령어로 모든 처리를 행한다. 현재는 원보드의 복원에서 끝나지 않고 목적보드를 설계, 제작하여 실질적인 독립성을 유지하면서 두개의 TMS320C30보드 사이의 유선통신을 행하게 할 수 있다. 이는 실질적으로 두채널 사이의 전화통신을 구축하는 기본이 되며, 이것을 바탕으로 무선통신을 실시간으로 처리할 수 있을 것이다. 이 하드웨어의 구성은 각각의 HOST와 두개의 DSP보드를 제작하여 시리얼통신 방법을 선택하여 양쪽의 유선통신을 구축할 예정이다. 현재는 목적보드는 제작되었고, 두보드상의 처리방법을 실험중이다.

제 6 장 실험 및 고찰

6.1 전송 비트

전송비트는 4개의 부분처리 프레임에 대해서 적응코덱에 대한 이득코덱 5비트, 고정코덱에 대한 이득코덱 6비트(이때는 5비트의 정상상태의 이득코덱, 5비트의 천이상태에 대한 이득코덱 5비트, 수신단에서 이득코덱의 선택에 사용하는 플래그 1비트를 갖는다. 즉 플래그 비트의 1혹은 0인가에 상태에 따라서 정상상태의 고정이득코덱인가 아니면 천이상태의 고정이득코덱인가를 선택한다). 고정코덱 인덱스 10비트, 적응코덱 인덱스 8비트(이 때는 프레

임내에 4개의 처리부분 프레임이 존재하므로 첫번째와 세번째는 적응코덱 인덱스 8비트를 모두 전송하지만, 두번째와 세번째는 그전 처리부분 프레임의 상대차를 전송하므로 6비트만을 전송하면 된다.), 10차 LPC계수는 VQ를 사용하여 최대 10비트의 인덱스를 전송하면 되고, 이것을 개선하기 위해서 5비트의 LSP계수 인덱스를 전송하면 된다. 이것을 전체 비트율로 나타내면 5.9kbps의 코딩율을 갖는다. 실험적으로 5비트의 LSP방법은 성능이 좋지 않은 결과로 나타나 10비트의 LSP방법을 사용하였다.

6.2 객관적 성능평가 및 성능비교

아래의 성능평가는 본 연구실에서 행했던 학술진흥재단의 지원 음성코딩(LD-CELP ; 8kbps)의 실험결과를 참조한다.

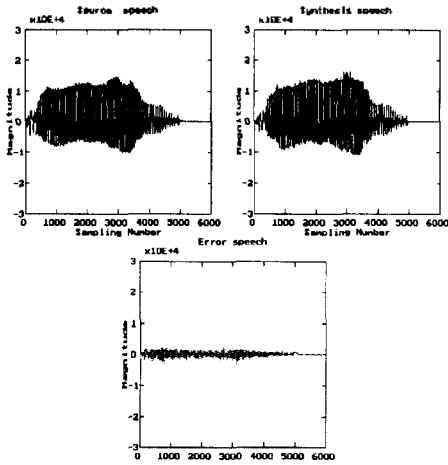
	원음성	PC Simulation	TMS320 DSP	LD_CELP
1	디지털 신호처리 연구실입니다	19.89	20	
2	서울	17.9	17.6	15.0
3	인천	17.2	17.3	15.2
4	용인	17.9	17.6	15.7
5	평택	13.0	13.2	12.3
6	문산	17.9	17.3	16.0
7	구리	17.5	16.8	16.5
8	대구	15.9	15.8	14.4
9	포천	13.3	12.7	10.5
10	안성	14.6	14.0	11.7
11	이천	14.6	14.3	
12	성남	13.2	13.5	
13	가평	12.0	12.3	10.5
	평균(dB)	15.76	15.57	13.79

(단:위 표의 값은 SNR(dB)이다)

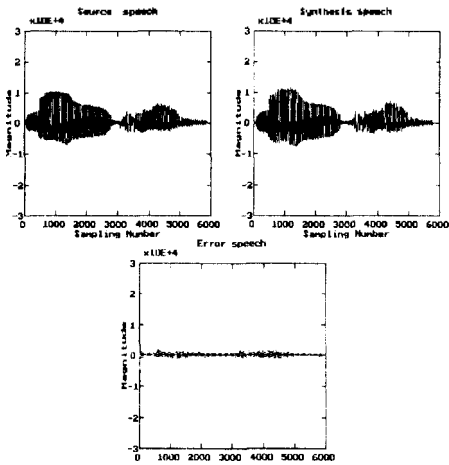
제 7 장 결과

본 논문에서 사용한 방법은 전방예측방법(Forward Method)을 사용하고 있기 때문에 처리한 원음성의 비퍼링지연, 알고리즘지연 수신단의 비퍼링지연 등에 의해서 실질적인 실시간처리에 많은 메모리를 필요로 하고 있으며 또한 코덱의 탐색알고리즘의 문제를 안고 있고, 이러한 문제는 처리량과 실시간 처리의 큰 문제점으로 대두된다. 실시간 처리를 행함에 정확한 실시간의 처리는 되지 않은 점이 아쉬움으로 남고, 위의 결과 표에서 나타나듯이 PC와 TMS320상의 차이는 TMS320C30의 부동소수점 연산이 일반적인 IEEE형태와 다르기 때문이다. 본 논문의 실험결과를 음질은 32kbps의 ADPCM정도의 음질을 나타내며 압축율은 6분의 1정도의 정보량을 필요로 하고, 객관적인 평가인 SNR(dB)의 평균값은 15.57로 비교적 좋은 음질로 평가할 수 있다. 앞으로의 과제는 실질적인 유무선망의 구성과 채널전송의 효율성, 채널전송시의 에러에 대한 대책 등이 있다.

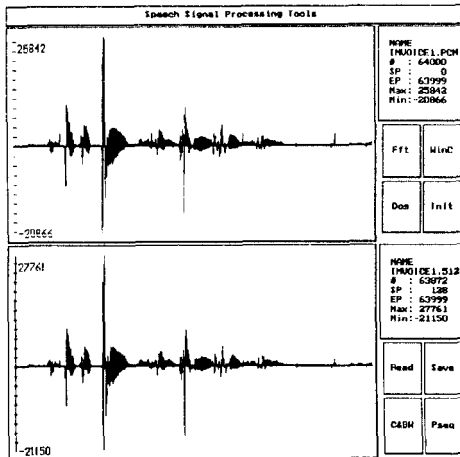
참고문헌



“서울”의 원음, 합성음과 오차파형 (남성)



“문산”의 원음, 합성음과 오차파형 (남성)



“디지털 실험실입니다”의 원음과 합성음 (남성)

1. “Quantization procedure for the excitation in CELP coders” Peter Kroon and Bishnu S. Atal 1987 IEEE 1649 - 1652 p
2. “The New 4800 bps Voice coding standard” Joseph P. Campbell, Jr 14 November 1989 (U.S Government)
3. “Speech coding using Efficient Pseudo-Stochastic Block codes” Daniel Lin 1987 IEEE 1354 - 1356p
4. “Efficient Procedure for finding the Optimum innovation in Stochastic coders” I.M.Trancoso, B.S.Atal 1986 IEEE 2375 - 2378p
5. “Efficient Search procedure for selecting the optimum innovation in stochastic coders” B.S.Atal 1990 IEEE 385 - 396p
6. “Fast Method for the CELP speech coding Algorithm” W.Bastian Kleiju 1990 IEEE 1330 - 1342p
7. Federal standard 1016
Analog to Digital converter Radio Voice By 4800 bit/second code excited linear prediction
1991 February 14
8. “Application of Line-Spectrum Pair to low-bit-rate speech encoders” George S.Kang and Lawrence J. Franssen
9. “Speech coding with Dynamic bit allocation” Tomohiko Taniguchi, Robert M. Gray
10. “The DOD 4.8 kbps standard” Joseph P. Campbell, Jr.
11. TMS320C30 C Source Debugger/ TMS320C30 Assembly Language Tools, Microprocessor Development System Digital Signal Products
12. Third-Generation TMS320 User’s Guide TI (Texas Instrument)
13. “LPC-Vocoder의 실시간 구현” 명지대학교 전기공학과 석사논문 조한신 1993
14. “지지연 CELP-Coder ” 명지대학교 전기공학과 석사논문 장태성 1993



임 상근(任 祥均)
1989년 3월 - 1993년 2월 명지대학교 전기공학과 (학사)
1993년 3월 - 현재 명지대학교 전기공학과 대학원 재학중