

파형 코딩 기법

목차

[소개](#)

[사전 요구 사항](#)

[요구 사항](#)

[사용되는 구성 요소](#)

[표기 규칙](#)

[펄스 코드 변조](#)

[필터링](#)

[샘플링](#)

[음성 디지털화](#)

[양자화 및 코딩](#)

[회사](#)

[A-law 및 U-law 회사](#)

[차등 펄스 코드 변조](#)

[적응형 DPCM](#)

[특정 32KB/s 단계](#)

[관련 정보](#)

소개

인간이 아날로그 통신을 잘 준비하긴 했지만 아날로그 전송은 그다지 효율적이지 않다. 아날로그 신호가 변속으로 약해질 때 복잡한 아날로그 구조를 무작위 전송노이즈 구조와 분리하기 어렵다. 아날로그 신호를 증폭하면 노이즈가 증폭되고 아날로그 연결이 사용하기에 너무 복잡해집니다. "1비트" 및 "0비트" 상태만 있는 디지털 신호는 잡음에서 더 쉽게 분리됩니다. 부패 없이 증폭될 수 있습니다. 디지털 코딩은 장거리 연결의 노이즈 손상에 대한 면역력이 더 높습니다. 또한 세계의 통신 시스템은 맥박 코드 변조(PCM)라고 불리는 디지털 전송 형식으로 변환되었다. PCM은 원래 음성 파형의 코딩된 형태를 만들기 때문에 "파형" 코딩이라고 하는 코딩 유형입니다. 이 문서에서는 아날로그 음성 신호의 디지털 신호에 대한 변환 프로세스를 개괄적으로 설명합니다.

사전 요구 사항

요구 사항

이 문서에 대한 특정 요건이 없습니다.

사용되는 구성 요소

이 문서는 특정 소프트웨어 및 하드웨어 버전으로 한정되지 않습니다.

표기 규칙

문서 규칙에 대한 자세한 내용은 [Cisco 기술 팁 표기 규칙](#)을 참조하십시오.

펄스 코드 변조

PCM은 ITU-T G.711 사양에 정의된 파형 코딩 방법입니다.

필터링

신호를 아날로그에서 디지털로 변환하는 첫 번째 단계는 신호 주파수의 상위 구성 요소를 필터링하는 것입니다. 이렇게 하면 다운스트림에서 이 신호를 쉽게 변환할 수 있습니다. 구어 에너지의 대부분은 200~300Hz~2700~2800Hz에 있다. 표준 음성 및 표준 음성 통신을 위한 약 3,000Hz 대역폭이 설정됩니다. 따라서 정밀한 필터를 사용할 필요가 없습니다(매우 비싼 경우). 4000Hz의 대역폭은 장비 지점에서 생성됩니다(볼 경우). 이 대역 제한 필터는 앨리어싱(엔티앨리어싱)을 방지하는 데 사용됩니다. 이는 Nyquist 기준에 의해 $F_s < 2(BW)$ 로 정의된 입력 아날로그 음성 신호가 언더샘플링될 때 발생합니다. 샘플링 빈도는 입력 아날로그 신호의 최고 빈도보다 작습니다. 이렇게 하면 샘플의 주파수 스펙트럼과 입력 아날로그 신호 간에 중첩이 생성됩니다. 원래 입력 신호를 재구성하는 데 사용되는 로우 패스 출력 필터는 이 중첩을 탐지할 만큼 똑똑하지 않습니다. 따라서 소스에서 시작되지 않는 새 신호를 생성합니다. 샘플링을 앨리어스라고 할 때 이러한 오탐이 생성됩니다.

샘플링

아날로그 음성 신호를 디지털 음성 신호로 변환하는 두 번째 단계는 지속적인 샘플링 빈도로 필터링된 입력 신호를 샘플링하는 것입니다. 이 작업은 PAM(Pulse Amplitude Modulation)이라는 프로세스를 사용하여 수행합니다. 이 단계에서는 원래 아날로그 신호를 사용하여 진폭 및 주파수가 일정한 펄스 열차의 진폭을 조절합니다.(그림 2 참조)

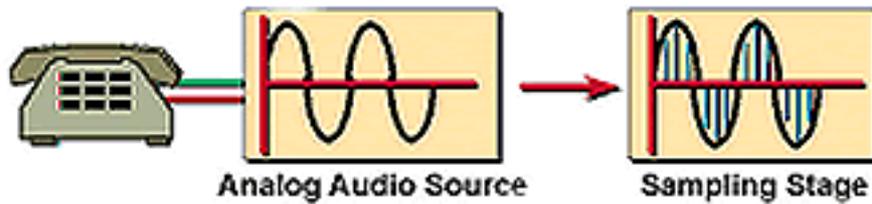
맥박 기차는 샘플링 빈도라고 불리는 일정한 주파수로 움직인다. 아날로그 음성 신호는 초당 100만 번 또는 초당 2~3회 샘플링할 수 있습니다. 샘플링 빈도는 어떻게 결정됩니까? 해리 나이퀴스트의 이름을 가진 한 과학자는 충분한 샘플을 가져가면 원래의 아날로그 신호를 재구성할 수 있다는 것을 발견했습니다. 그는 샘플링 주파수가 원래 입력 아날로그 음성 신호의 최소 주파수의 두 배 이상이면 목적지의 저통과 필터에 의해 이 신호를 재구성할 수 있음을 확인했습니다. 나이퀴스트 기준은 다음과 같습니다.

$$F_s > 2(BW)$$

F_s = Sampling frequency

BW = Bandwidth of original analog voice signal

그림 1: 아날로그 샘플링



음성 디지털화

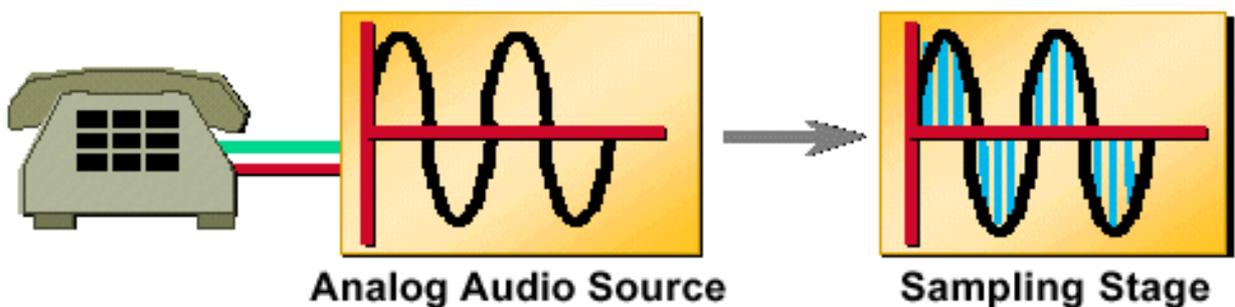
PAM을 사용하여 입력 아날로그 음성 신호를 필터링하고 샘플링한 후, 다음 단계는 텔레포니 네트워크를 통한 전송을 준비하기 위해 이러한 샘플을 디지털화하는 것입니다. 아날로그 음성 신호의 디지털화 과정을 PCM이라고 한다. PAM과 PCM의 유일한 차이점은 PCM이 한 단계 더 프로세스를 진행한다는 것입니다. PCM은 이진 코드 단어를 사용하여 각 아날로그 샘플을 디코딩합니다. PCM은 소스 쪽에 아날로그-디지털 변환기와 대상 쪽에 디지털-아날로그 변환기를 가지고 있습니다. PCM은 이러한 샘플을 인코딩하기 위해 양자화라는 기법을 사용합니다.

양자화 및 코딩

그림 2: 펄스 코드 변조 - 니퀴스트 정리

Pulse Code Modulation—Nyquist Theorem

Voice Bandwidth =
200 Hz to 3400 Hz



Codec Technique

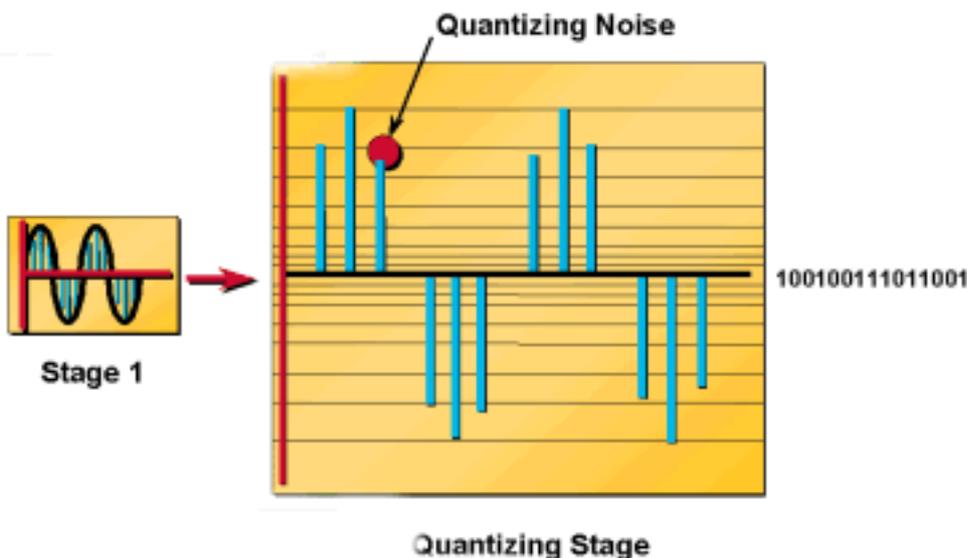
양자화는 각 아날로그 샘플 값을 고유한 디지털 코드 단어를 할당할 수 있는 개별 값으로 변환하는 프로세스입니다.

입력 신호 샘플이 양자화 단계를 시작하면 양자화 간격에 할당됩니다. 모든 양자화 간격은 입력 아날로그 신호의 동적 범위 전체에서 균등하게 간격(균일한 양자화)입니다. 각 양자화 간격에는 이진 코드 단어의 형태로 불연속 값이 할당됩니다. 사용되는 표준 단어 크기는 8비트입니다. 입력 아날로그 신호가 초당 8000회 샘플링되고 각 샘플에 8비트 길이의 코드 단어가 주어지면 PCM을 사용하는 텔레포니 시스템의 최대 전송 비트 속도는 초당 64,000비트입니다. 그림 2는 PCM 시스템에 대한 비트 전송률이 어떻게 파생되는지 보여줍니다.

각 입력 샘플에는 진폭 높이에 가장 가까운 양자화 간격이 할당됩니다. 입력 샘플에 실제 높이와 일치하는 양자화 간격이 할당되지 않은 경우 PCM 프로세스에 오류가 발생합니다. 이 오류를 양자화 소음이라고 합니다. 양자화 노이즈는 음성 신호의 SNR(Signal-to-Noise Ratio)에 영향을 주는 무작위 잡음에 해당합니다. SNR은 배경 노이즈를 기준으로 신호 강도를 측정합니다. 이 비율은 일반적으로 데시벨(dB)으로 측정됩니다. 마이크로볼트의 수신 신호 강도가 V_s 이고 노이즈 수준(마이크로볼트)도 V_n 이면 $S/N = 20 \log_{10}(V_s/V_n)$ 에서 데시벨 단위의 신호 대 노이즈 비율(S/N)이 지정됩니다. SNR은 데시벨(dB)으로 측정됩니다. SNR이 높을수록 음성 품질이 향상됩니다. 양자화 노이즈는 신호의 SNR을 줄입니다. 따라서 양자화 노이즈의 증가는 음성 신호의 품질을 떨어뜨립니다. 그림 3은 양자화 노이즈가 어떻게 생성되는지 보여줍니다. 코딩을 위해 N 비트 단어는 2N 양자화 레이블을 생성합니다.

그림 3: 아날로그-디지털 변환

Pulse Code Modulation— Analog to Digital Conversion



양자화 노이즈를 줄이는 한 가지 방법은 양자화 간격의 양을 늘리는 것입니다. 입력 신호 진폭 높이와 양자화 간격의 차이는 양자화 간격이 증가함에 따라 감소합니다(간격이 증가하면 양자화 노이즈가 감소함). 그러나 양자화 간격의 증가에 비례하여 코드 단어도 늘려야 합니다. 이 프로세스에서는 더 많은 코드 단어를 처리하기 위해 PCM 시스템의 용량을 처리하는 추가적인 문제가 발생합니다.

SNR(양자화 노이즈 포함)은 단일화된 양자화의 음성 품질에 영향을 주는 가장 중요한 단일 요소입니다. 균일 양자화는 입력 아날로그 신호의 전체 동적 범위에 걸쳐 동일한 양자화 레벨을 사용합니다. 따라서 낮은 신호에 작은 SNR(낮은 신호 레벨 음성 품질)이 있고 높은 신호에 큰 SNR(높은 신호 레벨 음성 품질)이 있습니다. 대부분의 음성 신호는 낮은 종류이기 때문에 더 높은 신호 레벨에서 더 나은 음성 품질을 갖는 것은 음성 신호를 디지털화하는 매우 비효율적인 방법입니다. 낮은 신호 레벨에서 음성 품질을 개선하기 위해 PCM(Uniform Quantization)은 companding이라는 비균일 양자

화 프로세스로 대체됩니다.

회사

컴패닝은 소스에서 아날로그 신호를 먼저 압축한 다음 대상에 도달할 때 이 신호를 원래 크기로 다시 확장하는 프로세스를 말합니다. 압축과 확장이라는 두 단어를 한 단어로 결합하여 'companing'이라는 용어가 만들어집니다. 코딩 프로세스 시 입력 아날로그 신호 샘플은 로그 세그먼트로 압축됩니다. 그런 다음 각 세그먼트는 균일한 양자화를 사용하여 양자화 및 코딩됩니다. 압축 프로세스는 로그입니다. 샘플 신호가 증가하면 압축이 증가합니다. 즉, 더 큰 샘플 신호는 더 작은 샘플 신호보다 더 많이 압축됩니다. 이렇게 하면 샘플 신호가 증가함에 따라 양자화 노이즈가 증가합니다. 입력 샘플 신호의 동적 범위 전체에서 양자화 노이즈의 대수 증가는 이 동적 범위 전체에서 SNR 상수를 유지합니다. 기업의 ITU-T 표준은 A-law와 U-law라고 불립니다.

A-law 및 U-law 회사

a-law 및 u-law는 16비트 선형 PCM 데이터를 8비트의 로그 데이터까지 압축하는 CCITT(Consultative Committee for International Telephony and Telegraphy) G.711에 의해 정의된 오디오 압축 체계(codex)입니다.

A-law 기업

선형 샘플 값을 12진수로 제한하면 A-law 압축은 이 방정식에 의해 정의됩니다. 여기서 A는 압축 매개변수(A=87.7(유럽))이고 x는 압축할 표준화된 정수입니다.

$$F(x) = \begin{cases} \frac{A * |x|}{1 + \ln(A)} & 0 \leq |x| < \frac{1}{A} \\ \frac{\text{sgn}(x) * (1 + \ln(A|x|))}{1 + \ln(A)} & \frac{1}{A} \leq |x| \leq 1 \end{cases}$$

u-로 기업

선형 샘플 값을 13진수로 제한하면 u-law(u-law 및 Mu-law는 이 문서에서 동일하게 사용됨) 압축은 이 방정식에 의해 정의됩니다. 여기서 m은 압축 매개변수(미국과 일본의 m=255)이고 x는 압축할 정규화된 정수입니다.

$$F(x) = \frac{\text{sgn}(x) * \ln(1 + \mu|x|)}{\ln(1 + \mu)} \quad 0 \leq |x| \leq 1$$

A-law 표준은 주로 유럽과 다른 세계에 의해 사용된다. u-law는 북미 및 일본에서 사용됩니다.

A-law와 U-law의 유사성

- 둘 다 로그 입력/출력 관계의 선형 근사값입니다.
- 두 가지 모두 8비트 코드 단어(256 레벨, 각 양자화 간격마다 하나씩)를 사용하여 구현됩니다. 8비트 코드 단어를 사용하면 64kbps(킬로비트/초)의 비트 속도를 사용할 수 있습니다. 이는 샘플

플링 속도(입력 주파수의 2배)에 코드 단어의 크기($2 \times 4 \text{ kHz} \times 8 \text{ 비트} = 64 \text{ kbps}$)를 곱하여 계산됩니다.

- 둘 다 동적 범위를 총 16개의 세그먼트로 분할합니다. 8개의 양수 및 8개의 음수 세그먼트. 각 세그먼트의 길이는 이전 세그먼트의 두 배입니다. 각 세그먼트 내에서 균일 양자화가 사용됩니다.
- 두 단어 모두 8비트 단어를 코딩하는 데 비슷한 방식을 사용합니다. First(MSB)는 극성을 식별합니다. 비트 2, 3, 4는 세그먼트 식별마지막으로 세그먼트를 정량화하는 4비트는 A-law보다 낮은 신호 수준입니다.

A-law와 U-law의 차이점

- 선형 근사치가 다르면 길이와 슬로프가 달라집니다.
- 8비트 코드 단어에서 세그먼트에 대한 비트 위치의 숫자 지정과 세그먼트 내의 양자화 레벨은 서로 다릅니다.
- A-law는 u-law보다 더 큰 동적 범위를 제공합니다.
- u-law는 A-law보다 낮은 수준의 신호에 대해 더 나은 신호/왜곡 성능을 제공합니다.
- 법률은 균일한 PCM에 해당하는 13비트를 필요로 한다. u-law에는 동일한 PCM에 14비트가 필요합니다.
- A-law를 사용해야 하는 국제적인 연결에서는 A-law를 A-Conversion으로 전환하는 것이 U-Law 국가의 책임입니다.

차등 펄스 코드 변조

PCM 프로세스 시 입력 샘플 신호 간의 차이는 최소화됩니다. DPCM(차등 PCM)은 이 차이를 계산한 다음 전체 입력 샘플 신호 대신 이 작은 차이 신호를 전송하도록 설계되었습니다. 입력 샘플의 차이가 전체 입력 샘플보다 작으므로 전송에 필요한 비트 수가 줄어듭니다. 따라서 음성 신호를 전송하는 데 필요한 처리량이 줄어듭니다. DPCM을 사용하면 음성 전송 비트 속도를 48kbps로 줄일 수 있습니다.

DPCM은 현재 샘플 신호와 이전 샘플 간의 차이를 어떻게 계산합니까? DPCM의 첫 번째 부분은 PCM과 정확히 유사하게 작동하는데, 이 때문에 PCM이 차등 PCM이라고 불립니다. 입력 신호는 일정한 샘플링 빈도(입력 주파수의 2배)로 샘플링됩니다. 그런 다음 PAM 프로세스를 사용하여 이러한 샘플을 모듈화합니다. 이 시점에서 DPCM 프로세스가 인계됩니다. 샘플링된 입력 신호는 조건자라고 하는 위치에 저장됩니다. 예측자는 저장된 샘플 신호를 가져와 차별화 요소를 통해 전송합니다. 차별화 요소는 이전 샘플 신호를 현재 샘플 신호와 비교하고 이 차이를 PCM의 양자화 및 코딩 단계로 보냅니다(이 단계는 동일한 양자화 또는 A-law 또는 u-law와 동일 수 있음). 양자화 및 코딩 후 차이 신호가 최종 목적지로 전송됩니다. 네트워크 수신기의 끝에서 모든 것이 역전됩니다. 먼저 차이점 신호는 미처리 상태입니다. 그런 다음 이 차이 신호는 조건자에 저장된 샘플 신호에 추가되고 원래 입력 신호를 다시 구성하는 로우 패스 필터로 전송됩니다.

DPCM은 음성 전송의 비트 속도를 줄이는 좋은 방법입니다. 그러나 음질에 관련된 다른 문제가 발생합니다. DPCM은 이전 샘플 입력 신호와 현재 샘플 입력 신호 간의 차이를 계산하여 계산합니다. DPCM은 균일 양자화를 사용하여 차이 신호를 수신합니다. 균일 양자화는 작은 입력 샘플 신호에는 작고 큰 입력 샘플 신호에는 큰 SNR을 생성합니다. 따라서, 음성은 더 높은 신호에서 더 좋다. 인간의 음성으로 생성되는 대부분의 신호가 작기 때문에 이 시나리오는 매우 비효율적입니다. 음성 품질은 작은 신호에 중점을 두어야 합니다. 이 문제를 해결하기 위해 적응형 DPCM이 개발되었습니다.

적응형 DPCM

적응형 DPCM(ADPCM)은 ITU-T G.726 사양에 정의된 파형 코딩 방법입니다.

ADPCM은 DPCM 프로세스 시 생성된 차이 신호의 양자화 레벨을 조정합니다. ADPCM은 이러한 양자화 레벨을 어떻게 조정합니까? 차이 신호가 낮으면 ADPCM은 양자화 레벨의 크기를 늘립니다. 차이 신호가 높으면 ADPCM은 양자화 레벨의 크기를 줄입니다. ADPCM은 양자화 레벨을 입력 차이 신호의 크기에 맞게 조정합니다. 이렇게 하면 차이 신호의 동적 범위 전체에서 균일한 SNR이 생성됩니다. ADPCM을 사용하면 음성 전송 비트 속도가 32kbps로 감소하여 A-law 또는 U-law PCM의 비트 전송률이 절반으로 줄어듭니다. ADPCM은 법률 또는 법률 PCM과 같이 '통행료 품질'의 목소리를 낸다. Coder에는 피드백 루프가 있어야 합니다. 인코더 출력 비트를 사용하여 양자자를 재지정합니다.

특정 32KB/s 단계

ITU 표준 G.726에 적용됩니다.

- a-law 또는 Mu-law PCM 샘플을 선형 PCM 샘플로 변환합니다.
- 다음 샘플의 예상 값을 계산합니다.
- 실제 샘플과 예상 값의 차이를 측정합니다.
- 코드 차이는 4비트로 전송하십시오.
- 4비트를 예측자에게 공급합니다.
- 4비트를 양자자에게 공급합니다.

관련 정보

- [음성 기술 지원](#)
- [음성 및 통합 커뮤니케이션 제품 지원](#)
- [Cisco IP 텔레포니 문제 해결](#)
- [Technical Support - Cisco Systems](#)