

波形編碼技術

目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[慣例](#)

[脈衝編碼調制](#)

[篩選](#)

[取樣](#)

[數位化語音](#)

[量化和編碼](#)

[壓擴](#)

[A-law與u-law壓擴](#)

[差分脈衝編碼調制](#)

[自適應DPCM](#)

[特定32 KB/s步驟](#)

[相關資訊](#)

簡介

雖然人類已經做好了模擬通訊的準備，但模擬傳輸並不是特別有效。當模擬訊號由於傳輸損耗而變弱時，很難將複雜的模擬結構與隨機傳輸雜訊結構分離。如果放大模擬訊號，它也會放大雜訊，最終導致模擬連線噪音過大，無法使用。只有「1位」和「零位」狀態的數位訊號更容易與雜訊分離。它們可以無腐敗地被放大。數位編碼更能抵抗長距離連線上的噪音破壞。此外，全球的通訊系統已經轉換為一種稱為脈衝編碼調制(PCM)的數位傳輸格式。PCM是一種稱為「波形」編碼的編碼，因為它建立原始語音波形的編碼形式。本文概要地描述了模擬語音訊號到數位訊號的轉換過程。

必要條件

需求

本文件沒有特定需求。

採用元件

本文件所述內容不限於特定軟體和硬體版本。

慣例

如需文件慣例的詳細資訊，請參閱[思科技術提示慣例](#)。

脈衝編碼調制

PCM是在ITU-T G.711規範中定義的波形編碼方法。

篩選

將訊號從模擬轉換為數字的第一步是濾除訊號的高頻分量。這使得下游更易於轉換此訊號。口語的大部分能量在200或300赫茲和2700或2800赫茲之間。為標準語音和標準語音通訊建立了大約3000赫茲的頻寬。因此，它們不必有精確的過濾器（非常昂貴）。4000赫茲的頻寬是從裝置的角度出發（如果檢視）。該帶限濾波器用於防止混疊（消除混疊）。當輸入模擬語音訊號被欠取樣時（Nyquist標準定義為 $F_s < 2(BW)$ ）會發生這種情況。取樣頻率小於輸入模擬訊號的最高頻率。這樣在樣本的頻譜和輸入模擬訊號之間產生重疊。用於重構原始輸入訊號的低通輸出濾波器不夠智慧，無法檢測到這種重疊。因此，它會建立一個並非來自源的新訊號。這種在取樣時產生的虛假訊號稱為混疊。

取樣

將模擬語音訊號轉換為數字語音訊號的第二步是以恆定取樣頻率對濾波後的輸入訊號進行取樣。它通過使用稱為脈衝幅度調制(PAM)的過程來實現。此步驟使用原始模擬訊號調制具有恆定幅度和頻率的脈衝串的幅度。（請參見圖2。）

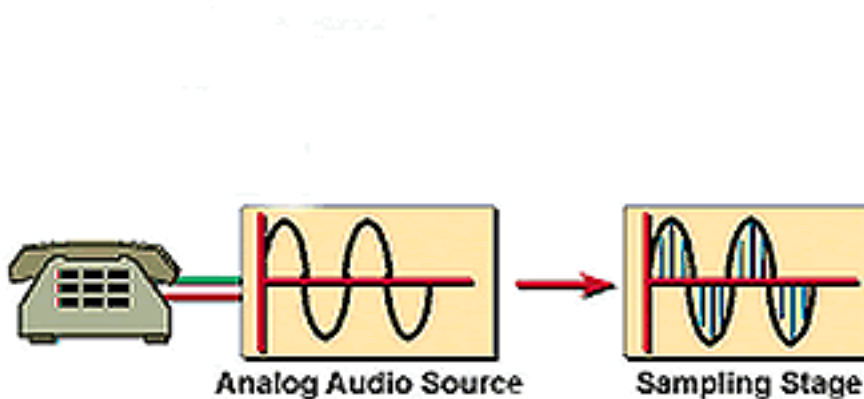
脈衝串以恆定的頻率移動，稱為取樣頻率。模擬語音訊號可以每秒100萬次或每秒2-3次取樣。如何確定取樣頻率？一位名叫哈里·奈奎斯特的科學家發現，如果採集到足夠的樣本，原始的模擬訊號可以重建。他確定，如果取樣頻率至少是原始輸入模擬語音訊號最高頻率的兩倍，則可在目的地通過低通濾波器重建該訊號。奈奎斯特的標準是這樣的：

$$F_s > 2(BW)$$

F_s = Sampling frequency

BW = Bandwidth of original analog voice signal

圖1:類比取樣



數位化語音

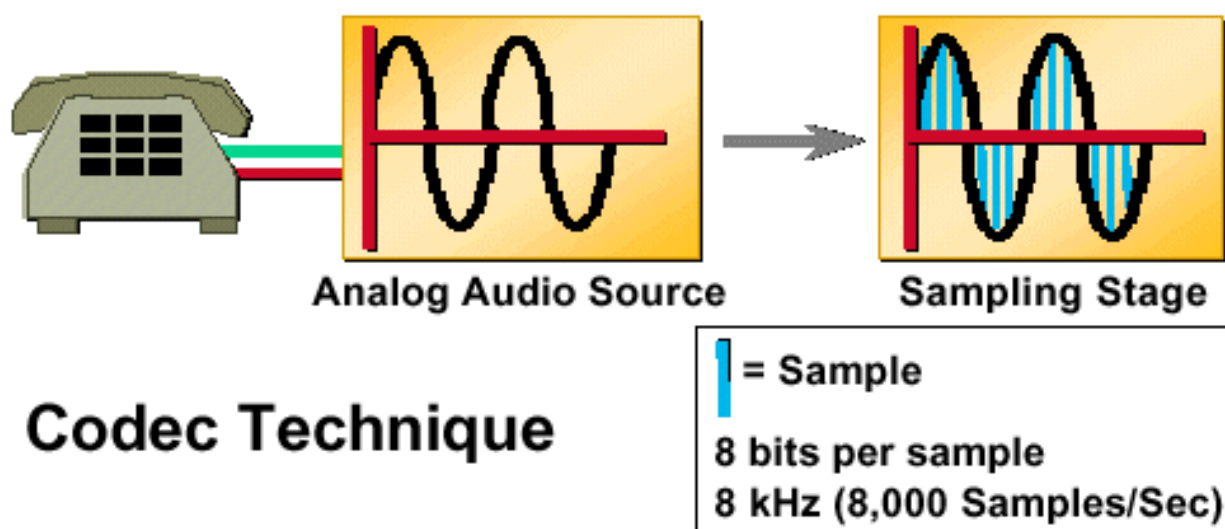
對輸入的模擬語音訊號進行過濾和取樣後（使用PAM），下一步是將這些取樣數位化，為通過電話網路傳輸做準備。將模擬語音訊號數位化的過程稱為PCM。PAM和PCM之間的唯一區別是PCM使流程更進一步。PCM使用二進位制碼字對每個模擬樣本進行解碼。PCM在源端具有模數轉換器，在目的端具有數模轉換器。PCM使用一種稱為量化的技術來編碼這些樣本。

量化和編碼

圖2:脈衝編碼調制 — 奈奎斯特定理

Pulse Code Modulation—Nyquist Theorem

Voice Bandwidth =
200 Hz to 3400 Hz



Codec Technique

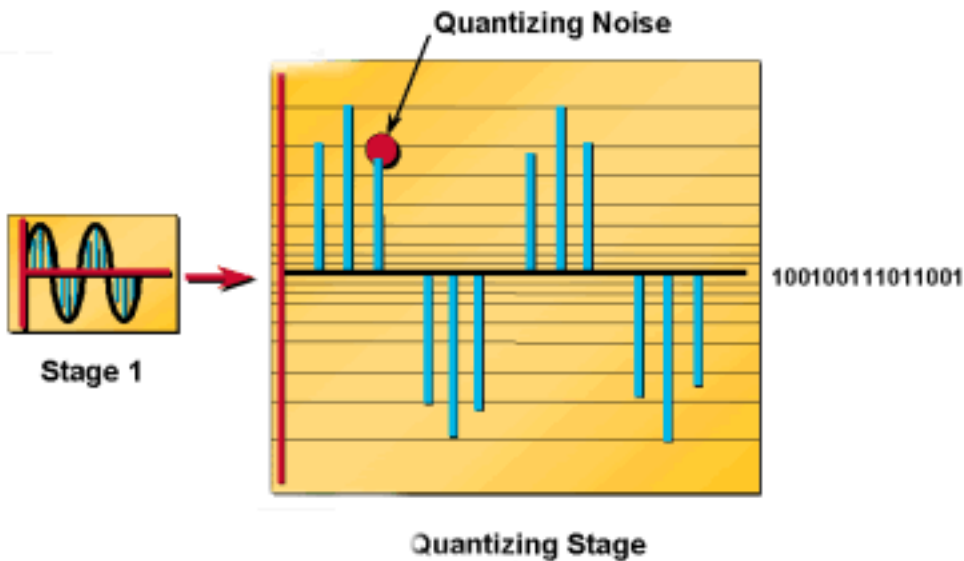
量化是將每個模擬樣本值轉換為離散值的過程，該離散值可以分配一個唯一的數字碼字。

當輸入訊號樣本進入量化相位時，它們被分配給量化間隔。在輸入模擬訊號的整個動態範圍內，所有量化間隔均等間隔（均勻量化）。每個量化間隔被分配一個二進位制碼字形式的離散值。使用的標準字長為8位。如果輸入模擬訊號每秒取樣8000次，並且每個取樣給定一個長度為8位的碼字，則使用PCM的電話系統的最大傳輸位元率為每秒64,000位。圖2說明了如何為PCM系統匯出位元率。

每個輸入樣本被分配一個最接近其幅度高度的量化間隔。如果輸入樣本未分配與其實際高度相匹配的量化間隔，則在PCM處理中引入錯誤。這個錯誤稱為量化雜訊。量化雜訊相當於影響語音訊號的訊雜比(SNR)的隨機雜訊。SNR是相對於背景雜訊的訊號強度的度量。該比率通常以分貝(dB)來衡量。如果輸入訊號強度（微伏）為 V_s ，雜訊水準（微伏）為 V_n ，則訊雜比以分貝形式由公式 $S/N = 20 \log_{10}(V_s/V_n)$ 給出。SNR以分貝(dB)為單位測量。訊雜比越高，語音品質越好。量化雜訊降低了訊號的SNR。因此，量化雜訊的增加降低了語音訊號的品質。圖3顯示量化雜訊是如何產生的。為了編碼目的，一個N位字產生 2^N 量化標籤。

圖3:類比到數位轉換

Pulse Code Modulation— Analog to Digital Conversion



減少量化雜訊的一種方法是增加量化間隔的量。輸入訊號幅度高度和量化間隔之間的差異隨著量化間隔的增加而減小（間隔的增加減少了量化雜訊）。然而，編碼字的數量也需要與量化間隔的增加成比例地增加。此過程引入了一些額外的問題，涉及PCM系統處理更多代碼字的容量。

SNR（包括量化雜訊）是影響均勻量化中語音品質的一個最重要的因素。均勻量化在輸入模擬訊號的整個動態範圍內使用相等的量化級別。因此，低訊號具有小SNR（低訊號級語音品質），而高訊號具有大SNR（高訊號級語音品質）。由於大多數語音訊號都是低類的，在較高訊號水準下具有更好的語音品質是一種非常低效的語音訊號數位化方法。為了改善較低訊號級別的語音品質，統一量化（統一PCM）被稱為壓擴的非統一量化過程所取代。

壓擴

壓擴是指首先在源頭壓縮模擬訊號，然後在訊號到達目標時將其擴展回原始大小的過程。術語壓縮是通過將兩個術語組合成一個單詞而建立的。在壓擴處理時，將輸入模擬訊號樣本壓縮成對數段。然後，使用統一量化對每個段進行量化並編碼。壓縮過程是對數的。壓縮隨著取樣訊號的增加而增加。換句話說，較大的取樣訊號比較小的取樣訊號被壓縮得更多。這導致量化雜訊隨著取樣訊號的增加而增加。在整個輸入取樣訊號的動態範圍內，量化雜訊的對數增加保持在整個動態範圍內的SNR恆定。ITU-T壓擴標準稱為A-law和u-law。

A-law與u-law壓擴

A-law和u-law是由國際電話電報諮詢委員會(CCITT)G.711定義的音訊壓縮方案（編解碼器），將16位線性PCM資料壓縮為8位對數資料。

A-law公司

將線性樣本值限制為12個大小位，A律壓縮由這個方程定義，其中A是壓縮引數（歐洲A=87.7），x是要壓縮的標準化整數。

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

u-law Compaider

將線性樣本值限制為13個量位，u律（本文中可互換使用u律和μ律）壓縮由這個方程定義，其中m是壓縮引數（美國和日本為m=255），x是要壓縮的標準化整數。

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

A-law標準主要被歐洲和世界其他地區使用。u-law被北美和日本使用。

A律與u律的相似性

- 兩者都是對數輸入/輸出關係的線性近似。
- 兩者均使用八位碼字（256級，每個量化間隔一個）實現。八位代碼字的位元率為每秒64千位元(kbps)。這是通過將取樣率（輸入頻率的兩倍）乘以代碼字的大小（2 x 4 kHz x 8位 = 64 kbps）來計算的。
- 兩者將動態範圍劃分為總共16個資料段：8個正片段和8個負片段。每個網段的長度是前一個網段的兩倍。在每個分段內使用統一量化。
- 兩者都使用類似的方法來編碼8位字：第一個(MSB)標識極性。二、三、四位標識資料段。最後四位將分段量化為比A-law更低的訊號電平。

A-law與u-law的區別

- 不同的線性近似會導致不同的長度和斜率。
- 八位碼字中位位置到片段的數字賦值和片段的量化級別是不同的。
- A-law比u-law提供了更大的動態範圍。
- u-law比A-law為低電平訊號提供了更好的訊號/失真效能。
- A-law需要13位才能獲得統一PCM等效值。u-law需要14位才能獲得統一PCM等效值。
- 國際聯絡需要使用A-law，u到A的轉換是u-law國家的責任。

差分脈衝編碼調制

在PCM處理時，輸入取樣訊號之間的差異最小。差分PCM(DPCM)設計為計算此差值，然後傳送此小差值訊號，而不是整個輸入取樣訊號。由於輸入取樣之間的差值小於整個輸入取樣，所以減少了傳輸所需的位元數。這樣可以減少傳輸語音訊號所需的吞吐量。使用DPCM可以將語音傳輸的位元率降低到48 kbps。

DPCM如何計算當前取樣訊號與前一個取樣之間的差值？DPCM的第一部分工作方式與PCM完全相

同（這就是它稱為差異PCM的原因）。以恆定的取樣頻率（輸入頻率的兩倍）對輸入訊號進行取樣。然後用PAM過程對這些樣本進行調制。此時，DPCM流程將接管一切。取樣輸入訊號儲存在稱為預測器的中。預測器取得儲存的樣本訊號並通過微分器傳送該訊號。微分器將先前樣本訊號與當前樣本訊號進行比較，並將該差異傳送到PCM的量化和編碼階段（該階段可以是A-law或u-law的均勻量化和壓縮）。在量化編碼之後，差值訊號被傳送到其最終目的地。在網路的接收端，一切都是相反的。首先，對差分訊號進行去量化。然後，將該差分訊號加到儲存在預測器中的取樣訊號中，並傳送到重建原始輸入訊號的低通濾波器。

DPCM是降低語音傳輸位元率的一種有效方法。但是，它也會導致一些與語音品質相關的問題。DPCM量化並編碼先前樣本輸入訊號和當前樣本輸入訊號之間的差。DPCM使用均勻量化來量化差分訊號。均勻量化產生的SNR對於小的輸入取樣訊號是小的，對於大的輸入取樣訊號是大的。因此，在較高訊號時語音品質更好。這種場景效率非常低，因為人類語音產生的訊號大多很小。語音品質需要關注小訊號。針對這一問題，提出了自適應DPCM。

[自適應DPCM](#)

自適應DPCM(Adaptive DPCM)是在ITU-T G.726規範中定義的波形編碼方法。

ADPCM調整在DPCM處理時生成的差分訊號的量化級別。ADPCM如何調整這些量化級別？如果差訊號較低，則ADPCM增加量化級別的大小。如果差分訊號高，ADPCM減小量化級別的大小。因此，ADPCM根據輸入差分訊號的大小自適應量化等級。這產生在整個差訊號的動態範圍內均勻的SNR。使用ADPCM可將語音傳輸的位元率降低到32 kbps，是A-law或u-law PCM的位元率的一半。ADPCM可產生「付費品質」語音，與A-law或u-law PCM類似。編碼器必須有反饋環路，使用編碼器輸出位來重新校準量化器。

[特定32 KB/s步驟](#)

適用於ITU標準G.726。

- 將A-law或Mu-law PCM樣本轉換為線性PCM樣本。
- 計算下一個樣本的預測值。
- 度量實際樣本與預測值之間的差異。
- 用四位來表示代碼差異，傳送這些位。
- 將四個位反饋給預測器。
- 將4位反饋給量化器。

[相關資訊](#)

- [語音技術支援](#)
- [語音和整合通訊產品支援](#)
- [Cisco IP電話故障排除](#)
- [技術支援 - Cisco Systems](#)