

多媒體導論與應用-第三版

第 2 章 音訊媒體

- 2.1 聲音的基本原理
- 2.2 類比與數位的轉換
- 2.3 取樣頻率
- 2.4 量化一位元深度(解析度)
- 2.5 修剪
- 2.6 音訊檔案容量
- 2.7 音訊壓縮原理
- 2.8 多聲道音訊
- 2.9 常見音訊檔案格式
- 2.10 課後練習



F7842A 多媒體導論與應用-第三版

1

2.1 聲音的基本原理

- 介質震動，造成壓力，而此壓力會以波的形式藉由介質向外擴散，傳到人的耳朵且頻率範圍在人耳可感應的範圍內(通常是20Hz~20kHz)，耳膜會因感應而聽見聲音。
- 在真空中是聽不見聲音的。
- 聲音在物理中用來表示音量的單位為分貝(dB)，分貝是使用對數來訂定的：

$$N_{AB} \text{ dB} = 10 \log\left(\frac{A}{B}\right)$$

F7842A 多媒體導論與應用-第三版

2

2.1 聲音的基本原理

- 對一般的多媒體音訊工作者來說，類比與數位的轉換、取樣頻率、量化、修剪、解析度、壓縮原理、檔案格式、音訊播放...等內容較為重要。

表 2-1 各種分貝與功率之範例

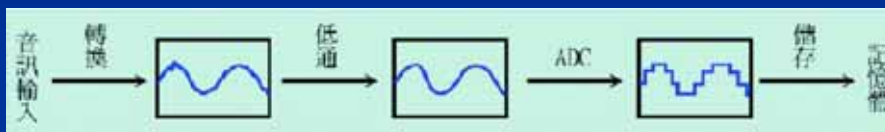
分貝 (dB)	功率 (W)	範例
10	0.000000001	農村的靜夜
20	0.00000001	樹葉落地
30	0.0000001	耳邊輕聲低語
50	0.000001	普通說話聲
60	0.00001	百貨公司內
80	0.0001	公共汽車上
90	0.001	捷運呼嘯聲
100	0.01	高速公路汽車聲
110	0.1	電燈旁邊
140-150	100-1000	飛機旁邊

F7842A 多媒體導論與應用-第三版

3

2.2 類比與數位的轉換

- 類比轉換為數位



- 數位轉換為類比

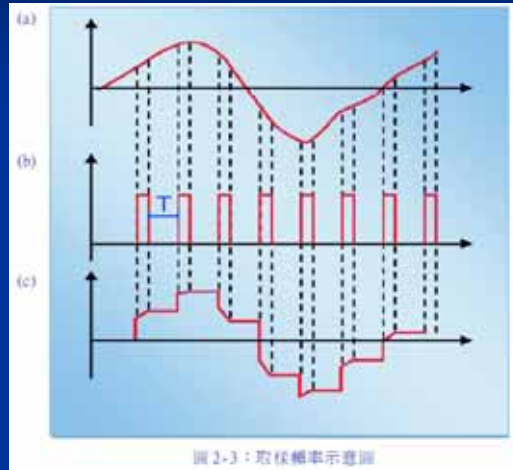


F7842A 多媒體導論與應用-第三版

4

2.3 取樣頻率

- 聲音數位化最重要的就是將類比訊號取樣。

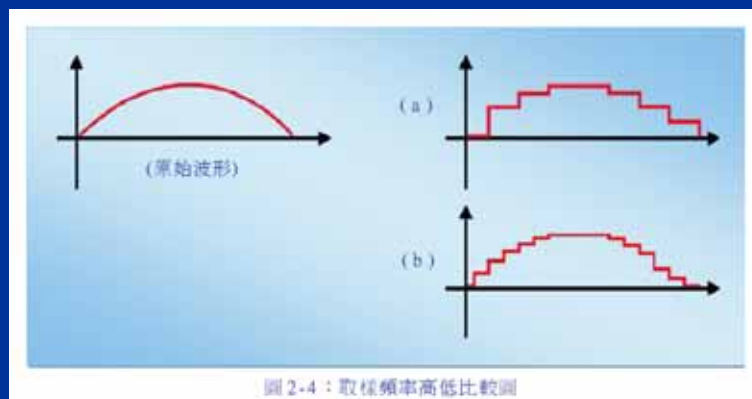


F7842A 多媒體導論與應用-第三版

5

2.3 取樣頻率

- 取樣頻率越高，亦即取樣間隔時間越短，所擷取後的數位音訊資料也就越準。



F7842A 多媒體導論與應用-第三版

6

2.4 量化－位元深度(解析度)

- 取樣在每一個上升邊緣時，ADC 會將當時的值紀錄下來，而此紀錄的值稱為樣本，單位為bit (或稱為解析度)，此動作即為量化。
- 舉例來說，當輸入波形範圍從 0.0V~10.0V 時，我們採用 4-bit 位元深度做為樣本，而 4-bit 為 0~15，扣除 0 有 15 種區間，所以：

$$\frac{10}{15} = \frac{2}{3}V$$

F7842A 多媒體導論與應用-第三版

7

2.4 量化－位元深度(解析度)

- 輸入的值為 $(2/3)V$ 時，樣本值為 00012，而 $(4/3)V$ 時，樣本值為 00102，可發現，如果當輸入值為 1.0V 時，將被四捨五入成為 00102= $(4/3)V$ ，因為位元深度造成的誤差，就造成了數位還原為類比時的不連續，雜訊就是這樣產生的。
- 總和來說，越高的位元深度，或稱為解析度，會有越真實的音質，但無論我們採用多高的位元深度來取樣，也不可能完全無誤差的記錄下輸入的類比訊號，這就是「量化失真」。

F7842A 多媒體導論與應用-第三版

8

2.4 量化－位元深度(解析度)

- 輸入音波 (紅色線段)，以 4-bit 量化後的樣本 (黑色梯線) 和 2-bit 量化後的樣本 (藍色梯線) 的比較，我們可看出解析度越高的樣本越接近原始波形，但相對的所需容量也較大，目前一般新 PC 使用之音效卡均可支援到 24-bit (HD Audio 標準) 解析度。

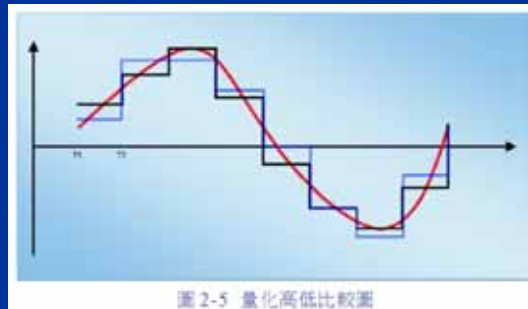


圖 2-5 量化高低比較圖

F7842A 多媒體導論與應用-第三版

9

2.5 修剪

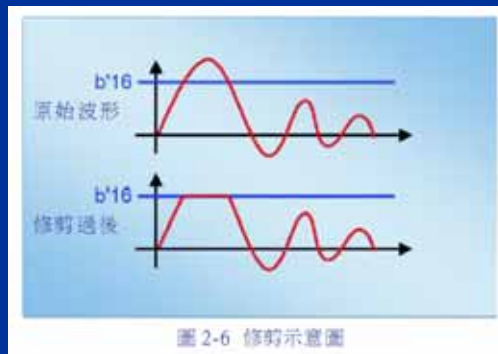
- 因量化過程中的位元深度不足，輸入的類比訊號最大值超過此位元深度可使用的區間時，就會將振幅超出最大值的部份修剪為可用區間的最大值，而被修剪掉的部份就成為了失真或是雜音，因此輸入的聲音過大時，錄製起來的聲音會「爆音」即是此原因。

F7842A 多媒體導論與應用-第三版

10

2.5 修剪

- 「b'16」代表 16-bit 位元深度的最高區間。16-bit 位元深度 (解析度) 最高可紀錄區間為 32767，當輸入波的峰值超過 32767 時，即會被修剪成爲 32767，此種轉換上的誤差，就稱爲修剪。



F7842A 多媒體導論與應用-第三版

11

2.6 音訊檔案容量

- 數位媒體工作者另一個必須了解的重點就是「檔案容量大小」，如何控制可接受的音質與傳輸速率，是相當重要的，以下公式爲音訊檔案在不壓縮的情況下的容量計算公式：

取樣率 (Hz) * 時間 (s) * 位元深度 (bit) / 8 * 聲道數量
其中位元深度 / 8 的 8 是因爲一個位元組 (byte) 有 8 個 bit

例：30 秒的 CD 音質錄製下的立體聲檔案大小爲 52920000bytes
 $44100 * 30 * 16 / 8 * 2 = 52920000$ bytes (位元組)

F7842A 多媒體導論與應用-第三版

12

2.6 音訊檔案容量

- 由公式可看出，如果未壓縮的情形下，音訊大小與品質是成正比的斜率成長，且該成長速率非常可觀，因此對於聲音的壓縮就變得相當重要。

表 2-2 常用的音訊設定格式與容量

取樣率(Hz)	解析度 (bit)	聲道	容量 (位元組每秒)
44100	16	2	172kbytes
44100	16	1	86kbytes
44100	8	2	86kbytes
22050	16	2	86kbytes
22050	16	1	43kbytes
22050	8	2	43kbytes
11025	16	2	43kbytes
11025	16	1	21kbytes
11025	8	2	21kbytes

2.7 音訊壓縮原理

- 常見的音訊壓縮有 MP3、WMA 等，而 MP3 的全名為 MPEG Audio Layer 3，而 WMA 則為 Windows Media Audio；DVD 所使用的壓縮技術則為 LPCM (Linear Pulse Code Modulation)、Dolby Digital、DTS 等。
- 以上常見的壓縮技術除了 LPCM 之外，均為「破壞性壓縮」(但 LPCM 檔案相當大)。
- 以 MP3 為例，為什麼 MP3 的壓縮比可以到 1:10 甚至更高？

2.7 音訊壓縮原理

- 一般常用的 MP3 為 128kbps (16kBytes)，其壓縮後的容量僅為原始的十分之一不到，但聲音卻好的令一般人難以察覺，其主要技術為「最小聽覺門檻」與「遮蔽效應」，這些技術除了有好的編碼之外，尚需要依據人類心理學模型來決定該如何壓縮。

2.7 音訊壓縮原理

- 最小聽覺門檻：
是一種用來減少資料流的方法。人耳對於 2kHz ~ 5kHz 的敏銳度與察覺度最高，所以它利用人耳的這項特性，將其他頻率的紀錄容量減少，甚至將微小不易察覺的高頻與低頻訊號刪除，以達到資料容量減少的目的。

2.7 音訊壓縮原理

- 遮蔽效應：

是一種運用人類聽覺神經特性的技術。在心理學中這是一種聽覺模型，它是說當一個感覺非常強烈時，同時間人類不太容易覺察到其他的感覺。遮蔽效應就是利用刪除，或是分配較低的取樣頻率給這些被別的較突出的聲音遮蔽後變的較不易被察覺的聲音，來減少其容量大小。

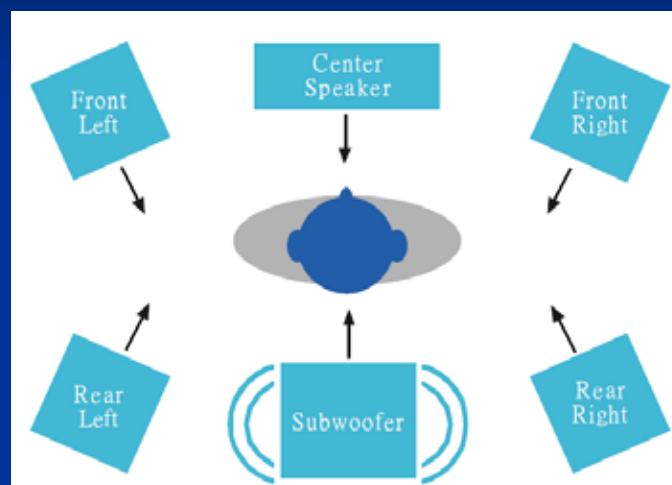
2.7 音訊壓縮原理

- 雖然 MP3 一直是數位音樂界的市場大宗，但 MP3 的破壞性壓縮方式所造成的音質損失還是有相當大的可改進空間，因此最近有許多新的壓縮技術如：MPEG2/4 - AAC、OGG。
- 新的壓縮技術除了運用更新的聲學模型之外也因為電腦處理器的進步得以使用更複雜的壓縮解壓縮演算法，因此可以達到更好的壓縮比，以 AAC 為例，AAC 可在 1:20 的壓縮比下達到與 MP3 的 1:10 的壓縮比有著相同的聲音品質。
- 但由於 MP3 具有容易實作且低成本與高效率的優勢，所以未來還是會繼續領導個人音樂市場。

2.8 多聲道音訊

- 廣泛應用於 DVD Movie 中的多聲道音訊編碼格式為 Dolby Digital (AC3) 與 Digital Theater Systems (DTS)。這兩種格式均為破壞性壓縮的 6 聲道系統，包含中央、前左、前右、後左、後右與重低音。
- AC3 的位元率為 448kbps，DTS 則為 1536 kbps，DTS 有著較佳的音質，但 AC3 則有較廣大的市場使用之，這兩類的音訊壓縮系統絕大多數使用於電影與遊戲工業，在一般的多媒體應用中較難以被實作出來。

2.8 多聲道音訊



2.8 多聲道音訊

■ 虛擬環繞音效

- 隨著 DVD 與環繞音效的普及，但多聲道喇叭系統在一般家庭中的付出成本較高，因此一些虛擬環繞音效系統也逐漸被發展出來，其中最著名的就是 Dolby Headphone 與 Dolby Virtual Speaker。
- 最著名的就是 Dolby Headphone 與 Dolby Virtual Speaker。兩個技術都是將 5.1 聲道解碼後透過聲學模型與頻域轉換等的演算法，使之在兩聲道中也能讓使用者有接近於 5.1 環繞音效的效果，前者為耳機的使用環境，後者則適用於一般的雙聲道喇叭。

2.8 多聲道音訊

■ 環繞音效的發展現況

- 在 2008 年塵埃落定的下一代 High Definition (HD) 影音儲存媒體之戰最終由 Blu-ray Disc 戰勝了 HD-DVD。但在規格戰的激烈過程中，許多新的環繞音效技術也隨之誕生，這些新的技術都有辦法支援到 8 聲道 (相較於 5.1 聲道增加了左、右兩聲道) 與 24-bit/96kHz。
- 杜比實驗室以 Dolby Digital 為基礎發展出了 Dolby TrueHD 技術，TrueHD 支援 24-bit/96kHz 的位元深度與取樣率和最高達至 18Mbit/s 和 8 聲道。雖然 HD DVD 和 Blu-ray Disc 標準現今限制於最大聲道位數為 8，但 Dolby TrueHD 的規格可輸出至 14 個聲道。

2.8 多聲道音訊

- 環繞音效的發展現況
 - DTS 的延續規格則是 DTS-HD Master Audio，DTS-HD 支援 8 聲道、96kHz/24bit 的 LPCM 非破壞性壓縮聲音資訊，最高位元率達 24.5Mbit/s，藉由下一代光碟的大容量，以無損耗的方式大幅提升多聲道音效的音質。

2.9 常見音訊檔案格式

未壓縮的波形音訊格式

- WAV
 - 由微軟制定，採 PCM 編碼的未壓縮波形格式，主要用於 Windows PC 中，符合 RIFF Resource Interchange File Format 規範。
 - 所有的 WAV 都有檔案標頭，其記載著此檔案的編碼參數。
 - WAV 並無對取樣的硬性規定值，可以在錄音時自行設定。
 - WAV 也可以像 AVI 那樣使用不同的 CODE 來壓縮，不過一般人對於 WAV 的定義都是使用 PCM 編碼，但其實 AVI 一樣是可以利用別的 CODE 來編碼的，甚至是 MP3。

2.9 常見音訊檔案格式

未壓縮的波形音訊格式

■ AU

- AU 是 UNIX 下一種常用的格式，為昇陽 (SUN) 所開發，其副檔名為 .au。
- AU 本身也支援多種壓縮模式，但本身的架構不如 AIFF 和 WAV。此格式現在已經鮮少人知了，目前大概只有在使用到 JAVA 的時候才會使用到 AU 格式。

2.9 常見音訊檔案格式

未壓縮的波形音訊格式

■ AIFF

- AIFF 是 Apple 的標準格式，其副檔名為 .aiff。平時我們熟知的 QuickTime 就是使用 AIFF 作為音訊的軟體。
- AIFF 本身是一種功能很強的格式，其支援了許多的壓縮技術，但是它為 MacOS 專用的格式，因此在 PC 上較為少見。但 Apple 在多媒體播放工具上的佔有率也不容小覷，正因為如此，AIFF 格式到現在還是有其固定的使用者。

2.9 常見音訊檔案格式

非破壞性壓縮格式

■ APE

- Monkey's Audio，是一種常見的非破壞性音訊壓縮格式，一個壓縮為Monkey's Audio 的音訊檔案聽起來與原音訊完全一樣，不會因解壓縮和壓縮而改變。
- Monkey's Audio文件常使用 .ape 的副檔名，目前被廣泛應用於音樂 CD 的數位化保存。

■ FLAC

- Free Lossless Audio Codec 的縮寫，FLAC 是一套著名的自由音訊壓縮編碼，其特點是非破壞性壓縮。不同於其他破壞性壓縮編碼如 MP3 及 AAC，它不會破壞任何原有的聲音資訊，所以可以還原音樂光碟音質。

2.9 常見音訊檔案格式

非破壞性壓縮格式

■ TTA

- True Audio (縮寫 TTA) 是一種簡單的非破壞性編解碼器。
- TTA 是一種基於自適應預測過濾的非破壞性音訊壓縮，與目前主要的其他各式相比，雖不能有更好的壓縮效果，但其重點是在演算法的簡單與高效率，使用者不須要有太好的設備即可即時解碼播放。

2.9 常見音訊檔案格式

破壞性壓縮格式

■ MP3

- MP3 (MPEG Audio Layer 3) 屬於 MPEG 標準的一環，其副檔名為 .mp3。
- 其高效率的資料壓縮與音質效果，再加上便利的播放程式以及編碼程式支援，使得 MP3 的使用者越來越多，目前絕對是多媒體音訊的標準。
- 它可以藉由調整位元傳輸速率來調整其壓縮大小，範圍從 16kbps~320kbps 都有，而在一般的標準 128kbps 下，可使得一分鐘的 MP3 檔案壓縮成不到 1MB，而且播放的效果幾乎聽不出與原來的差異。

2.9 常見音訊檔案格式

破壞性壓縮格式

■ WMA

- WMA 就是 Windows Media Audio，由微軟開發，其副檔名為 .wma。
- 最大特色就是比 MP3 容量還要小一半，且音質不輸 MP3 的特色。WMA 有著無失真、有失真、語音，全方面的需求支援。
- 它在網路上的串流品質較佳，再加上 WMA 的編碼與播放都是相當方便，直接在微軟的網站與 Windows XP/Vista 都有內建，不另外收費！

2.9 常見音訊檔案格式

破壞性壓縮格式

■ AAC

- 由 Fraunhofer IIS、Dolby、蘋果、AT&T、索尼等公司共同開發，以取代 mp3 格式。
- 2000 年，MPEG-4 標準增加了 AAC，並重新整合了其特性，故現又稱 MPEG-4 AAC，即 m4a。
- AAC 可在 1:20 的壓縮比下達到與 MP3 的 1:10 壓縮比有著相同的聲音品質，而目前因為 iPod 的盛行也增加了 AAC 在市場上的能見度。

2.9 常見音訊檔案格式

破壞性壓縮格式

■ OGG

- Ogg Vorbis 是一種類似於 Mp3 的破壞性音訊壓縮格式，但他有更佳的聲學模型，同時也與 AAC 相同，可以藉由更複雜的壓縮解壓縮演算法達到更佳的音質與壓縮比。
- OGG 格式目前還不算經常被使用。

2.9 常見音訊檔案格式

網路串流格式

串流格式就是將一般的音訊切割成很多帶有標記順序的小封包，並且一個一個的接收到之後，可以不用等待全部封包傳送完成就先播放，而播放的同時則繼續傳輸剩下的封包，因此達到邊傳邊播的目的。

2.9 常見音訊檔案格式

網路串流格式

■ WMA/ASF

- WMA 的競爭對手就是網路媒體著名的 Real Networks。
- 微軟聲稱，在只有64kbps的位元速率下，WMA 可以達到接近 CD 的音質 (相當於 128kbps 的MP3)。
- 且 WMA 支援串流技術可邊讀邊播，因此 WMA 可以很輕鬆的完成線上廣播。
- 基於種種優勢，使得 WMA 與Real Audio 並列為網路串流的標準。

2.9 常見音訊檔案格式

網路串流格式

■ Real Audio

- 普遍應用於網際網路上的聲音格式，必須安裝 Real Player 播放程式，才能播放此聲音檔。
- 早期網路頻寬低的時候 Real Audio 確實成功的創造了網路音訊串流的新紀元，但隨著網路頻寬的增加，以及 WMA 的出現，Real Audio 的音質與位元率的令人失望，儘管新的 Real Audio 有在做補強，但仍是慢了一步。
- 不過 Real Networks 在網路上打下的金字招牌可是相當硬的，到現在仍然是網路串流的主流之一。

2.10 課後練習

1. 請說明類比音訊轉換為數位音訊的流程為何。
2. 請計算 10 秒鐘的 CD 音質錄製下的立體聲檔案大小為何？
3. 請說明壓縮技術中「最小聽覺門檻」原理為何？
4. 常見的破壞性壓縮音訊格式有哪些？其特性為何？