

WAV 聲音檔製作(錄音程式)與處理

引用自：<http://sparc2.nhltc.edu.tw/~honda/audio/wav/wav.htm>

聲音原理

人類對聲音最直接的認識來自聽覺感受，因此「聲音」的最原始定義就是「人耳所能聽到的」。很久以前，人們就發現任何振動著的物體都可以成為聲源。從聲源到人耳總會有一段距離，聲音通過這段距離就是聲音的傳播，而傳播必須依賴介質。最常見的介質就是空氣，如果沒有空氣，例如在真空中，這段距離就是再近人耳也聽不到聲音。

所以聲音是指任何一種壓力的變化（來自物體或分子的振動），經由水、空氣，或其他介質的傳播，而能被人耳所偵測到的機械能。聲音的波形基本上是一種正弦波的組合，因此它具有正弦波的各種數學與物理特性，如頻率、波幅、時間長短、與相位角等。

空氣固然是傳遞聲波到人耳的媒介，然而各式各樣的聲音卻是由兩個基本要素所構成，也就是有兩個成分會影響聲波的形狀：一是頻率(frequency)，二是響度(loudness)。

■ 聲音的頻率

頻率與物體所造成的振動[快慢]有關，影響所及是在某一時間單位內聲音波形的多寡。

頻率是以時間為基準，振動快則頻率高，音調較為高，振動慢則則頻率低，其衡量單位為赫茲(Hertz, Hz)，若以一秒鐘為基準單位，則一千赫茲(1000Hz，或稱 KHz)就等於每秒鐘振動一千次，也就是每秒鐘產生一千個音波。人耳的頻率響應範圍則大約在 20-20,000 赫(Hz)之間。

■ 聲音的響度

響度是指發聲的[振動幅度]，亦可稱為[力度]，影響所及是聲音波形的高低，音波的振幅愈大，則響度愈大，其衡量基準並不是以時間為準，而是以振幅的大小為準，以[分貝(decibel, dB)]來衡量。為了使聲音有一可供比較的基準，故設定 0 分貝是正常人剛好能百分之五十偵測到的聲音響度，相當於 0.0002 微巴 (μ bar) 的大氣壓力，這個基準壓力大概相當於將一小粒花生米，掉落一厘米 (cm) 所作的功，所以是非常微小的。分貝的計算是採取一種叫作「聲音音壓位準」(SPL)的方式，音壓則相當於波幅的大小。利用這種方式可以決定聲音的大小水準，故也稱之為 dB SPL。它的計算方式如下：

$$L=20\log(P/P_0)$$

P:聲音的音壓(或波幅大小)

P₀:基準音壓

L:聲音的分貝值。

取對數的理由除了可以避免使用太大的 P 值外，尚因為由過去的很多研究，認為人對外界物理量的感覺是一種對數形式。雖然這種認定現在看起來，並不見得正確，但並不妨礙將其當為一種約定俗成的定義方式。在 CGS 公制中，SPL 定義下的參考音壓(dBref)為 0.0002 微巴=0.0002 達因/平方厘米，每一達因表示大約可使一克質量產生 1 厘米/每秒每秒之加速度的力。若採用現在通行 SI(國際標準)制，則 SPL 中的 dBref 應為 20 μ Pa(1 μ Pa=1 牛頓/ 平方公尺)。

分貝與音壓(P)之間的關係並不是線性的，DB 的尺度是呈指數增長的，每隔 20 分貝，其響度或振幅則增加 10 倍，

例如:40 分貝比 20 分貝的響度，提高了 10 倍，而 60 分貝則比 20 分貝提高了一百倍，80 分貝提高一千倍之多。

一般人耳可聽見範圍 0-140dB。所謂零分貝，乃指健康耳朵所能聽見的最小音壓，國際標準是以 0.0002 微巴 (dyne/cm²) 為基準。

人不僅是在偵測聲音頻率上有所限制，在音量（以分貝表示）上也有不能偵測的部份，太低的音量人耳便不能測到，稱之為「聽覺閾」，太高則祇能產生痛覺，稱之為「痛覺閾」。

人的可聽聞範圍，受到頻率與音量範圍限制的情形，它同時標示出一般性語言與音樂在頻率與音量上的範圍，人耳大體上都可察覺到語言與音樂的存在。

用聲音音壓位準方式量出來的分貝值，與人耳所主觀感覺到的響度(loudness)值之間，存在有某種具規律性的關係。在低頻部份，人耳所感覺到的響度值(或稱主觀分貝值)，會比純粹由音壓所算出來的分貝值來的小；但在超過 1000 赫以後的高頻部份，則略有起伏。

由於人耳對不同頻率的聲音音量，會有衰減與增強等不同的加權方式，所以探討聲音對人的影響時，必需將這個因素考慮進來。現在的噪音量測方式，當考慮到是對人的影響時，就在量測儀器內設計一種「A 權線路」使儀器在接收與計算聲音音量時，就像人耳在接收一樣，使低頻部份的音量被大幅衰減。

現在國內所使用的環境噪音分貝值，由於目的是在評估其對人體的影響，所以也都採取這種 A 式加權的作法，所得的分貝值稱之 dB(A)，以與一般所稱的 dB 有所分別。外界的聲音常是由不同頻率帶上都有音壓的聲音混合組成，所以若該聲音以低頻為

主(如工廠內的機器)，則所測的 dB(A)值會比 dB 值來得低。

數位錄音原理

本節內容引用自：<http://sparc2.nhltc.edu.tw/~honda/audio/wav/wav.htm>

數位錄音的最大的目標就是解決記錄類比訊號(Analog)的傳統錄音不佳的問題。

數位錄音所記錄的都是數位(Digital)的訊號;必須先把類比式的聲音符號轉化為數位式的聲音符號，這種技術稱為[類比-數位轉換](Analog-to-Digital，D/A)，但是當欲將儲存於 CD 或磁碟上的數位聲音，透過揚聲器播出時，又要將數位式的聲音資料再轉換回類比式的聲音資料，其技術稱為[數位-類比轉換](Digital-to-Analog，D/A)。

所謂數位訊號指的正是 0 與 1 的訊號，除了能夠更精確地記錄各式各樣的聲音之外，也不容易因為媒材變質而造成訊號的衰減。轉換的過程尚牽涉到聲音的轉換品質，其品質則是由聲音取樣的方式來決定。

聲音的取樣是指將類比音訊轉化成數位音訊過程的準確度，它包含了三個重要的因素:一是音波取樣頻率，二是音波取樣大小，三是單立體音。

1.取樣頻率

取樣頻率(Sampling Rate)指音效卡在一秒之中對聲音(波形)做記錄的次數。

根據研究，聲音播出時的品質常常只能達到取樣頻率的一半，因此須採取雙倍樣率才能將原音準確重現。也就是只要取樣頻率大於原始訊號頻率的兩倍以上，即可減低錯誤，達到和原始聲音極真實的音訊。人的聽力極致約為 20KHz，所以高品質的取樣應為其兩倍以上，當聲音來源為音樂時，因位它所橫跨的頻率變化極為寬廣，通常以採 44.1KHz 的頻率為 CD 音樂取樣率的標準;但是若以語音為主，由於人說話的語音大約為 10KHz，因此加倍採樣，只取 22KHz 即可。取樣率越高，所記錄下來的音質就越清晰;當然，越高的取樣所記錄下來的檔案就會越大。

2.取樣大小

解析度決定了取樣的一音波是否能保持原先的形狀，愈接近原形則所需解析度愈高。

若以 8 位元來記錄取樣，則其所能表達的組合種類是 2 的 8 次方，即 256，表示用 8 位元的取樣大小能分辨出 256 個層次的聲音;若採 16 位元來取樣，則能分辨的差異將高達 2 的 16 次方，為 65536，其精確度自然大為提高。

16bit，8bit 取樣的差別在於動態範圍的寬窄;動態範圍寬，音量起伏的大小變化就能夠更精細地被記錄下來。如此一來不論是細微的聲音或是強烈的動感震撼，都可以表現得淋漓盡致;而 CD 音質的取樣規格正是 16 位元取樣的規格。

在取樣的過程中，不斷連續變化的類比訊號要用數位化的數值來表示，這樣的過程就會產生所謂的量化誤差(Quantization error)。所謂量化誤差指的是實際聲音訊號之震幅 (amplitude)和數位化後所得數字之間的差異。如果用將數位信號還原成類比訊號的角度來看，量化誤差就是失真(Distortion)。我們可以用增加取樣大小的方式來降低量化誤差，也就是利用更多的位元(bits)來表示一個取樣訊號，這樣便可以提高精確度。

所謂的量化(Quantization)就是將類比訊號所代表的連續範圍分成一段一段的區間 (Interval)，每一段區間我們定義一個數位化的值。區間的數目是跟取樣大小有關，舉例來說，有一種最簡單的量化法稱為"線性量化法"(Linear quantization)，這種量化法採用等距離的間隔空間，假設一個訊號它的最大值是 5.0，取樣大小為 3 個位元，則每個量化區間就是 $5.0/2^3$ ，也就是 0.625 單位。另外一種相反的量化法就是"非線性量化法"(Nonlinear quantization)，這種量化法採用不同的間隔空間。以"對數量化法"(Logarithm quantization)為例，低震幅範圍的量化區間就比高震幅範圍的區間較為接近，用這種量化法產生的結果就是在低震幅時我們會得到較好的效果。通常如果使用同樣的取樣大小，非線性量化法會比線性量化法得到較佳的音訊品質。但是如果是要對音訊做濾波(filtered)或一些運算時，使用線性量化法會比較容易處理。

目前已經發展了許多音訊編碼的方法用以減少儲存量或是傳輸的時間，以下所列為兩種較普遍的編碼方法：

- a. PCM(Pulse code modulation)
- b. ADPCM(Adaptive delta pulse modulation)

3.單音(mono)、立體音(stereo)

單音乃將聲音儲存在一個聲道(channel)，立體音則儲存成左、右兩個聲道，故立體音聲音品質較佳，檔案也大單音一倍。

聲音後製作(Sound Post-Production)

聲音處理(sound processing)可使數位聲音在僅些微失真的情況下作編輯，當聲音被錄製後可用特效(special effects)來作增強，如倒播(reverse)或迴音(echo)，最常見的編輯乃是將一長段的錄音依需要切取其中一小段。

音效(Sound Effects)

指表達某些事件發生所使用的特殊聲音，例如用翻書的聲音來表達換幕。通常用來表達音效的聲音並非真實的原音，例如街上的馬蹄聲，實際是木塊敲打桌面的聲音。

另一種聲音的處理則是聲音的結合，常用來營造特殊的氣氛，例如在一場有關戰役的演講中，將演講者的講詞內容配上戰爭的背景音樂。