

聲音合成與電腦音樂環境之建構

黃志方

國立交通大學

音樂研究所助理教授

Email: jeffh@faculty.nctu.edu.tw

摘要

本文旨在闡述聲音合成的基礎法則，以及電腦音樂教育與創作的原理和方法，以建立數位藝術教育中，電腦音樂教育的環境。聲音合成法則，無論是加成式、相減式、調幅、調頻等，均為電腦音樂領域中重要的合成方法。本文透過簡明的數學表示，以及相關軟體的分析與實現，希冀可為有志於電腦音樂創作的人尋求清楚而有效之方法，並靈活運用聲學的基本知識與物理法則，將音樂與科技深刻地整合在一起，以建構出數位藝術教育中的電腦音樂創作環境。

關鍵字：電腦音樂教育環境、聲學、聲音合成

壹、序論

電腦音樂係利用音樂科技的方法來對聲音進行分析、合成、轉換與組合，可達成音樂創作與教育之目的。世界上主流之音樂創作已逐漸朝向電子化、數位化、與國際化的趨勢。以歐美國家為例，無論在資訊或是藝術兩種學院或系所下，擁有電子或電腦音樂相關之系所或是研究中心已是不勝枚舉。因此電子音樂或電腦音樂是現代音樂發展之重要方向與手段，包括了當代作曲大師如德國 Stockhausen、法國 Pierres Schaeffer、Pierre Boulez、希臘 Iannis Xenakis、美國 Edward Varese、John Cage 等都有重要作品問世。由於大量的藝術家投入電子數位影音藝術科技方面之發展，使得二十一世紀真正進入了數位藝術科技的時代。電子或電腦音樂對於聲音變換 (Transformation)、多軌並列、以及機遇或準則創作

(Algorithmic Composition) 上的卓越特質均可使得音樂創作更加豐富與便利,除了可為蓬勃發展中之新音樂創作注入新的動力外,更可透過不同的聲音合成法則與聲音之變換相結合,共創嶄新之數位聲音藝術創作形式與教育方法。

貳、聲學、物理與電腦音樂之關係

一、聲音與音樂的差異

許多人將聲音 (Audio) 與音樂 (Music) 混為一談。事實上,二者有其差異性,列舉如下:

(一) 聲音是物理性的,其參數僅包括了振幅、頻率、與時間三項。其中聲波之振幅與時間的對應關係為信號之時域 (Time Domain) 表示法,透過聲波處理軟體 Audition 載入後如圖一所示,橫軸是時間,縱軸是振幅;而頻率與振幅的對應關係為信號之頻域 (Frequency Domain) 表示法即為數位聲波之頻譜 (Spectrum),如圖二所示,橫軸是頻率,縱軸是振幅。

(二) 音樂是心理性的,其參數很多,包括了音高、時值 (Duration)、節奏、織度 (Texture)、演奏法 (Articulation)、力度 (Dynamics)、和聲、對位、速度 (Tempo) 等。

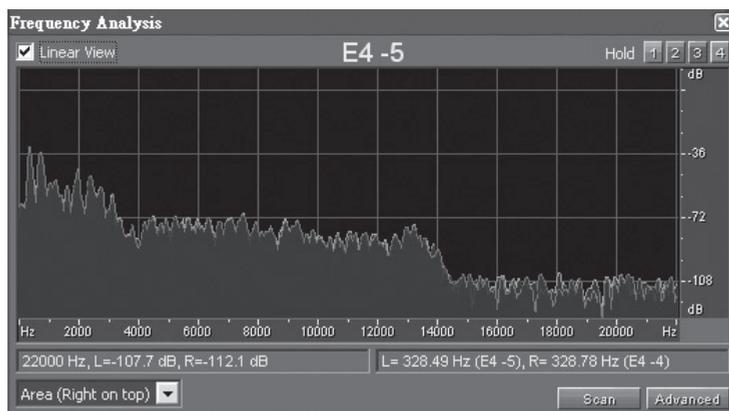
一般而言,使用電腦音樂環境的作曲家以心理性的音樂參數設定來創作音樂;而電腦音樂環境可將作曲家所設定之各種音樂參數轉化為物理性的聲音參數,透過信號處理來完成音樂作品。

二、泛音列與諧振

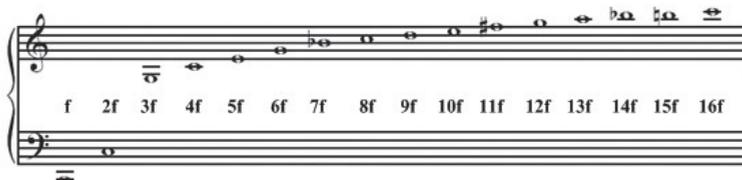
電腦音樂的發展過程中,聲學 (Acoustics) 扮演了極為重要的角色,特別是音樂領域上定義之「泛音」(Harmonics),實為物理或工程領域上所稱之「諧振」,而泛音分布的特性基本上決定了音色的變化。音樂與物理二領域若可確實掌握,則可透過工程了方法來對聲音信號加以處理。圖三為以譜表左方最低音的C音為基礎音 (Fundamental Note) 而產生之泛音列,其中基礎音的頻率為 f ,第二個泛音頻率為 $2f$,其餘類推至第十六個泛音,其頻率為 $16f$ 等。



圖一 時域上的聲波圖



圖二 聲波之頻域圖



圖三 泛音列與頻域關係圖

上圖中除了 $f, 2f, 4f, 8f$, 以及 $16f$ 外, 也就是所有低音C的八度音之外, 其餘的泛音項次均無法在鋼琴琴鍵中正確的表示, 因此僅為約略的音高, 這是因為鋼琴是十二平均律的調律系統, 而非泛音列上的純律系統。

上圖中基礎音低音C之音高為:



若以數學來表示泛音列的存在, 可假定聲波信號基礎音的表示式為:

$$y_i(t) = A_i(t) \cdot \sin(2\pi f_i t + \phi_i)$$

例如信號中其中第四個泛音為中央C:



其信號表示如下:

$$y_4(t) = A_4(t) \cdot \sin(2\pi f_4 t + \phi_4)$$

則完整以 f 為基礎音的聲波信號包含了基礎音與無限項次的泛音列, 其表示式如下:

$$y(t) = \sum_{n=1}^{\infty} A_n(t) \cdot \sin(2\pi f_n t + \phi_n)$$

此處為第 f_n 次項的泛音頻率, 亦即:

$$f_n = n f$$

$n=1, f_1=f$ 時即為基礎音的頻率。

ϕ_n 則為第 n 次項泛音的相位角度 (Phase Angle)。

$A_n(t)$ 為第 n 次項泛音的振幅。

透過各泛音項次的疊加 (Superposition), 聲波可成爲一個複合波 (Complex Wave), 並可透過數位信號處理執行各種聲音的變換。

三、杜普勒效應

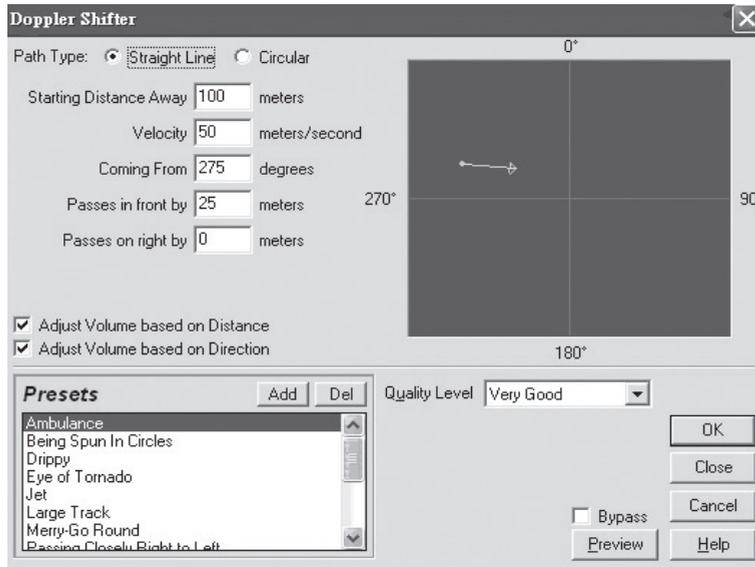
杜普勒效應 (Doppler Effect) 係利用音源與聽者的相對徑向 (Radial) 速度變化, 來產生頻率的變化, 在電腦音樂中可以因此而產生音高上的變化。

杜普勒效應的表示式如下:

$$f_L = f_s \left(\frac{a + V_L}{a + V_s} \right)$$

其中 f_s 為音源發射的頻率, f_L 為聽者收到的頻率, a 為聲波的速度, 一般常溫下為 340 meters / sec, V_s 為音源的運動速度, V_L 為聽者的運動速度。

聽者所感受到 f_L 的音頻將隨著 V_L 和 V_s 的變化而產生變化，因此若是聽者與音源在徑向上為接近 (Approaching) 時，聽者所感受到的頻率會上升；反之若聽者與音源在徑向上為離開時，則聽者所感受到的頻率會下降，因此可利用杜普勒效應產生音頻的變化。



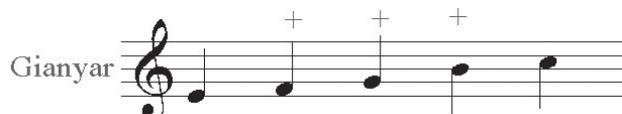
圖四 聲波杜普勒效應之設定方式

四、「拍」效應 (Beating Effect)

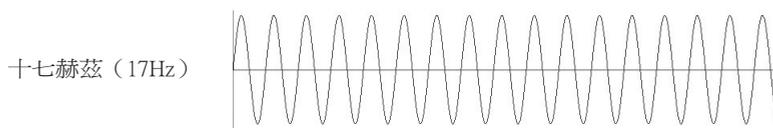
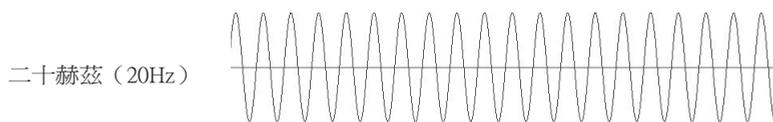
「拍」效應 (Beating Effect) 是指利用兩個接近的音高一起發聲時，其頻率與振幅均有忽大忽小變化的絢爛效果，在物理上稱之為「拍」。例如巴里島之甘美朗青銅樂器具有振動變化 (Shimmery) 之特質，原因是將成對之樂器在音高上故意稍微調成不同，造成兩個接近但不同之頻率一起發聲，產生「拍」效應。

吾人可透過電腦分析甘美朗音樂上拍的效應，亦可利用數學分析軟體 Matlab 來模擬產生此效應。例如以 Sulaan 與 Gianyar 兩種巴里島之律制將一對 Selisir Gong 分別調律成如下之系統，則會產生上述「拍」之效果。譜表中「+」表示比記譜的音高一些，而「-」表示比記譜的音低一些。

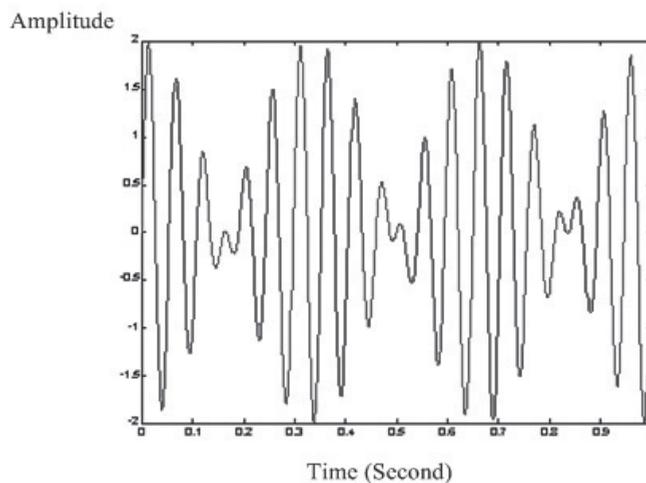
聲音合成與電腦音樂環境之建構



下面是兩個頻率接近之信號，一為二十赫茲 (20Hz)，另一為十七赫茲 (17Hz)：



則其一秒鐘長度之「拍」的效果如下：

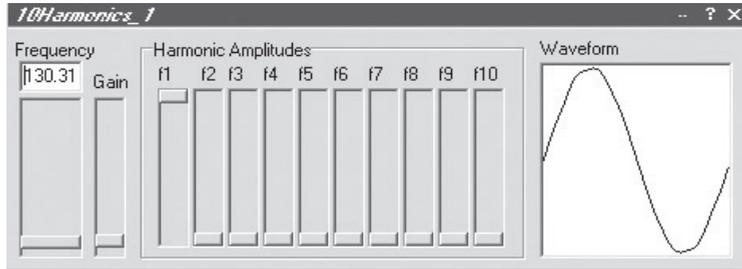


圖五 聲波「拍」的效果

參、聲音合成法則於電腦音樂之應用

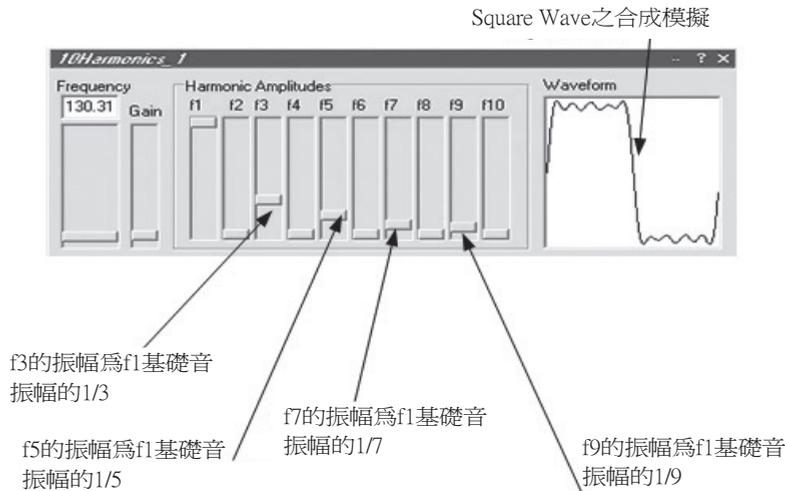
一、加成式聲音合成法 (Additive Sound Synthesis)

所謂加成式聲音合成法，可透過基礎音 $f = f_1$ 的頻率來與其各項泛音列的頻率分量來疊加，如圖六所示為基礎音的正弦波形。



圖六 基礎音的正弦波形

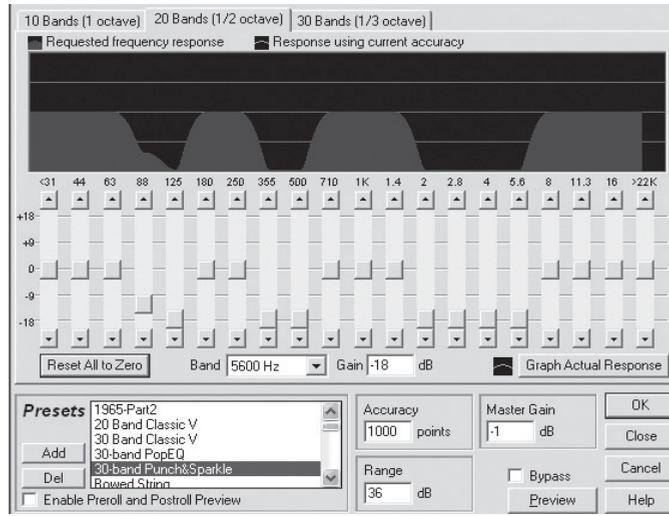
下圖是透過AudioMulch聲波處理軟體之「10 Harmonics」物件，將基礎音頻率設定為 $f = f_1 = 130.31\text{Hz}$ ，則二次泛音頻率為 $f_2 = 2f$ ，三次泛音頻率為 $f_3 = 3f$ 等。而利用加成式的聲音合成法，吾人可利用「只設定奇數次的泛音振幅，且其振幅大小反比於其奇數項次數目」的方式來執行方波 (Square Wave) 的聲音合成。例如 f_3 的 $= (1/3)f$ ， $f_5 = (1/5)$ 等。



圖七 方波 (Square Wave) 的聲音合成

二、相減式聲音合成法 (Subtractive Sound Synthesis)

聲波可透過各種濾波器的設計，去除掉某些特定頻段的聲訊，已達到所謂「相減式聲音合成法」的目的。例如可以產生一段噪音，無論是白噪音 (White Noise) 或是其他特性分布的噪音，均可利用「溝槽式濾波器」(Notch Filter) 來針對某些頻段及其諧振頻段加以減抑或是去除，如圖八所示，2K~5.6K Hz 頻段的聲訊會為濾波器所衰減或甚至消除，這就是所謂的相減式聲音合成法之基礎。



圖八 相減式的聲音合成法

三、調幅聲音合成法

調幅 (Amplitude Modulation, AM) 是利用調變信號 (Modulator Signal) 將載波信號 (Carrier Signal) 之振幅加以變化。例如可利用一個低頻振盪器 (Low Frequency Oscillator, LFO) 代表調變信號，而以電壓控制振盪器 (Voltage Controlled Oscillator, VCO) 代表載波，則其信號之表示式如下：

$$y(t) = \sin(2\pi f_m t) \cdot \sin(2\pi f_c t)$$

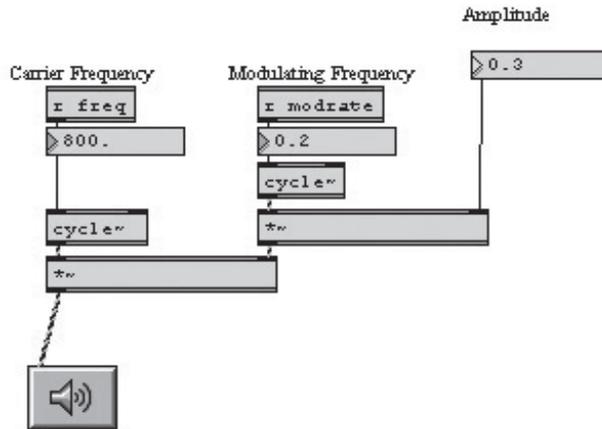
其中 f_m 與 f_c 分別代表調變信號與載波之頻率。

經過三角函數運算後，可得調變後之信號為

$$y(t) = \frac{1}{2} [\cos(2\pi [f_c + f_m] t) - \cos(2\pi [f_c - f_m] t)]$$

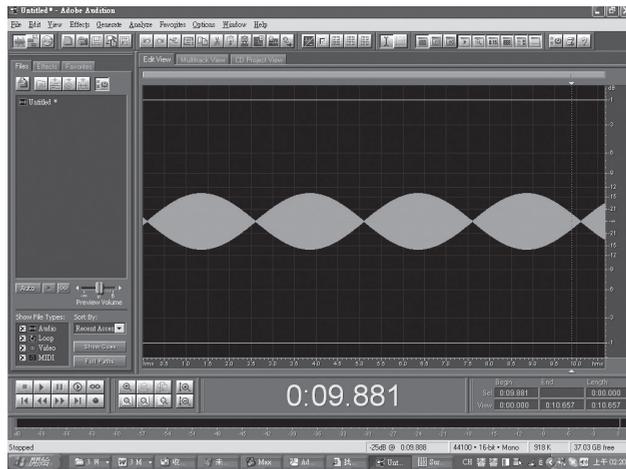
在頻譜上可獲得 $f_c + f_m$ 與 $f_c - f_m$ 兩個頻率分量。

吾人可透過Max/MSP電腦音樂程式來產生此種「Ring Modulation」，亦即將兩個「cycle~」物件所產生之正弦波信號相乘，其 $f_c = 800$ Hz，而 $f_m = 0.2$ Hz，如下圖所示：



圖九 以Max/MSP執行AM聲音合成

此時我們可以利用Audition聲波處理軟體將Max/MSP所產生的AM調變聲波錄下，如下圖所示：



圖十 AM的Ring Modulation波形

此時我們可聽到的聲波（載波頻率）本身是800 Hz之音頻，但是其振幅上調變的頻率（調變頻率）如上圖波形所示，為0.2 Hz。這種AM聲波調變用於電腦音樂上，一般是作為顫音（Tremolo）之用。

四、調頻聲音合成法

調頻（Frequency Modulation, FM）應用於聲音合成領域上是源自1973年史丹佛大

聲音合成與電腦音樂環境之建構

學的John Chowning。類似AM調變方式，FM也使用調變信號來改變載波的頻率特性（AM為改變載波振幅特性）。其信號之表示式如下：

$$y(t) = \sin [2\pi f_c t + I \sin (2\pi f_m t)]$$

此處I為調變指數 (Modulation Index)，用以控制載波之扭曲程度。

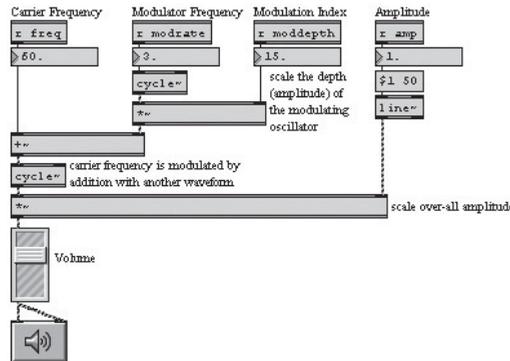
FM信號經展開後，可得下列表示式：

$$y(t) = \sum_{n=0} J_n(I) [\sin(2\pi[f_c + n f_m] t) + \sin(2\pi[f_c - n f_m] t)]$$

其中 $J_n(I)$ 為Bessel函數的係數，代表各頻率分布的振幅大小。

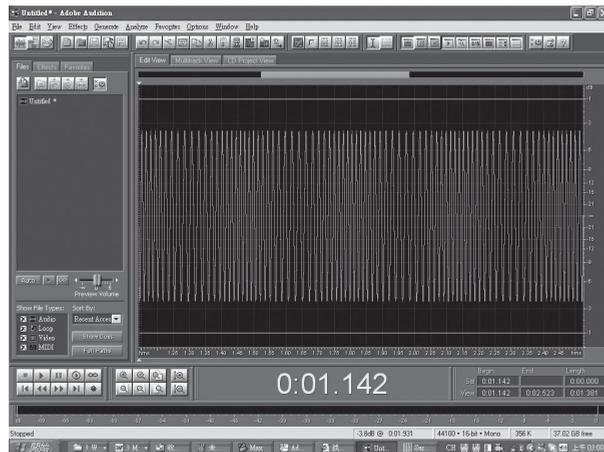
在頻譜上可獲得， $f_c \pm f_m, f_c \pm 2f_m, f_c \pm 3f_m$ ，等各個頻率分量。

吾人可透過Max/MSP電腦音樂程式來產生FM Modulation，亦即將兩個「cycle~」物件所產生之正弦波信號帶入FM調變的公式中，其 $f_c = 60$ Hz，而 $f_m = 3$ Hz， $I = 15$ ，如下圖所示：



圖十一 以Max/MSP執行FM聲音合成

經聲波處理軟體Audition錄下後，得到的時域聲波圖形為：



圖十二 FM聲音合成的波形

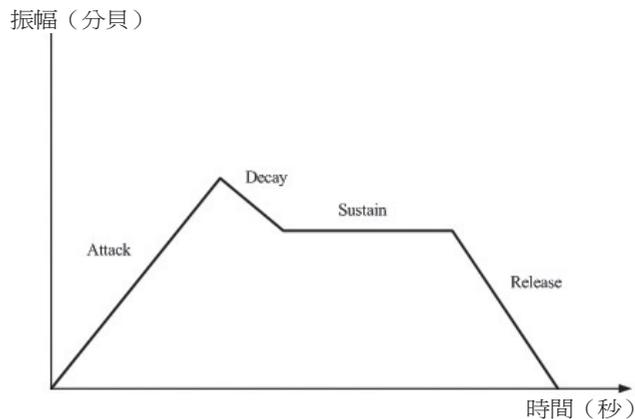
由上圖可觀察到FM聲波信號的振幅沒有變化，但是頻率卻有疏密之週期性變化。

肆、電腦音樂教育與創作方法討論

一、聲波之直接合成(Direct Sound Synthesis)

聲波可透過各式振盪器(Oscillator)組合而直接產生，例如電壓控制振盪器(Voltage Controlled Oscillator, VCO)可以產生正弦波、方波、鋸齒波、三角波等，再透過調變信號所需的低頻振盪器(Low Frequency Oscillator, LFO)，再加上振幅包絡線(Amplitude Envelope)設計，可以完成所需之調頻與調幅等各式聲音合成。

振幅包絡線可決定樂器音色之演奏法(Articulation)特色，例如擊樂特性之音色具有較短的包絡線，上升(Attack)與衰減(Decay)均很快速；而一般管絃樂器特性之音色具有較長之延長(Sustain)與釋放(Release)時間。



圖十三 ADSR振幅包絡線

二、聲波之間接合成(Indirect Sound Synthesis)

係利用電腦音樂軟體載入聲波，透過各式的變換(Transform)來將聲波特性改變。在聲波處理軟體Audition中有許多的聲音變換方式，包括時間 / 音高(Time / Pitch)變換、殘響(Reverberation)、濾波器(Filters)等。下圖為所選取聲波片段對時間 / 音高(Time / Pitch)作變換的方法。

聲音合成與電腦音樂環境之建構



圖十四 聲音的Stretch變換

三、聲納圖在電腦音樂創作與教育所扮演的角色

聲納圖 (Sonogram) 或稱動態頻譜圖 (Spectrogram) 係利用短時富立葉轉換 (Short-Time Fourier Transform, STFT) 持續地將時域上的聲音信號轉換至動態的頻譜 (Spectrum)，其公式如下所示。

$$STFT(\tau, \omega) = \int_{-\infty}^{\infty} x(t) \cdot w(t-\tau) e^{-j\omega t} dt$$

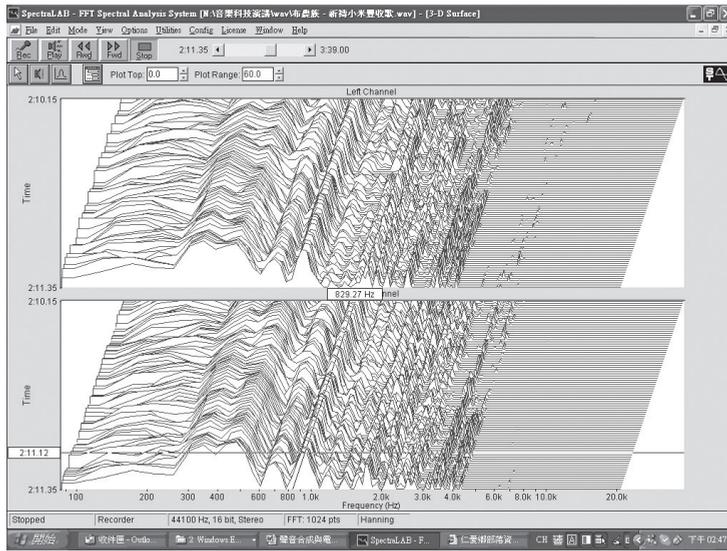
電腦軟體之聲納圖可將聲音樣本載入後呈現時間、振幅、以及頻率三者變化關係，對於電腦音樂創作、分析與教育等有極大的貢獻，特別是對於無法以傳統樂譜來呈現的電腦音樂創作，聲納圖提供了對於音色細節分析的良好方法。數位形式的短時富立葉轉換則如下所示：

$$STFT(m, \omega) = \sum_{n=-\infty}^{\infty} x[n] w[n-m] e^{-j\omega n}$$

其聲納圖之振幅為：

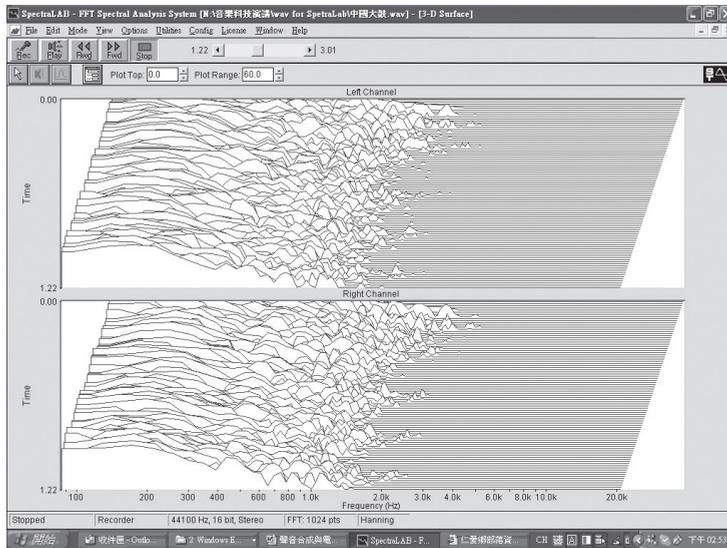
$$Mag = |STFT(\tau, \omega)|^2$$

下圖係以3D聲納圖來分析台灣布農族的八部和音 - 祈禱小米豐收歌，可以看出低音的頻譜分布是廣闊而連續，且領唱的低音部份聲頻有逐漸上昇之趨勢。而泛音分布的情形是顯著而連續的。



圖十五 以3D聲納圖來分析台灣布農族的八部合音 - 祈禱小米豐收歌

圖十六則以3D聲納圖來分析中國大鼓的音色變化，可以清楚地看到大鼓低頻頻譜分布情形，因無突出的基礎音與其高次項的泛音分布，因此自然無法產生固定音高 (Definite Pitch) 的音樂。



圖十六 以3D聲納圖來分析中國大鼓的音色變化

四、多軌聲波之蒙太吉 (Montage) 創作方法討論

利用上述電腦音樂的各式合成方法，吾人可利用聲波處理軟體Audition來載入各式聲音樣本，透過循環 (Looping)、反轉 (Reverse)、各式特效 (Effects) 等聲音變換，進而形成數位具象音樂 (Digital Concrete Music)。而透過MIDI介面使用音源器裝置亦可產生模擬樂器之聲響，惟吾人應留意的是此類音源缺乏流動之演奏法 (Articulation)，無法產生生動活潑之音色變換，因此宜透過Audition聲波處理軟體之聲音變換與多軌蒙泰吉 (Multi-track Montage) 能力，結合各式聲音合成法則，創作出優秀的作品。

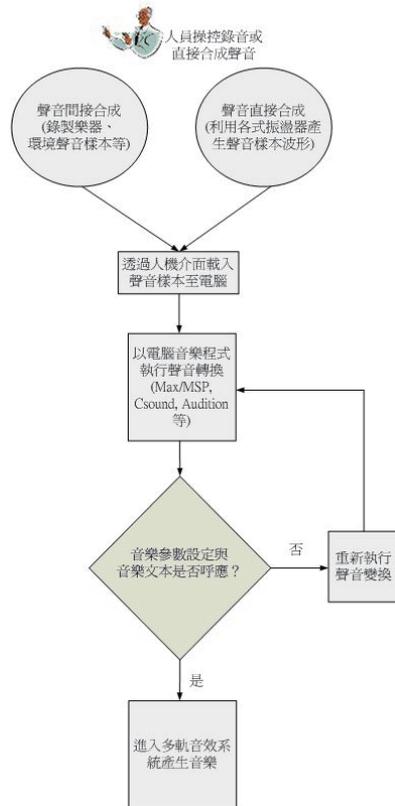


圖十七 數位聲波之多軌蒙泰吉

值得注意的是，電腦音樂之創作一如文章寫作般，需要具有結構思考與處理之基本能力，與音樂文本 (Text) 能有所呼應，否則將如文字之堆砌，失去音樂的張力。如圖十八所示，無論採用何種聲波合成與變換處理的方法，亦不論所使用的電腦程式軟體為何，所設定之音樂或聲波變換參數均需要考量到與音樂文本之呼應性，亦即起、承、轉、合的有機性，才可進入多軌並貼處理的最後程序。

伍、結論

由上述本文的探討，吾人可了解到基礎的聲學與物理原理，並學習以電腦音樂軟體來實作各種聲波合成方法，以及聲音變換的方式。唯有將音樂與科技深刻的研究並整合在一起，才有可能使聲波合成、電腦音樂創作，與數位藝術教育之環境建構成功。電腦音樂與藝術方面的研究和工程科技的研究發展若能合而為



圖十八 電腦音樂創作方法流程

一，將可充分展現出多元化、彈性化、整合化之科技整合特色，使科學研究可應用於數位音樂藝術創作，尤其是透過資訊研究與科技整合上之本質特色，使藝術與資訊科技之間搭建出一座穩固而極具整合效益的橋樑，希冀逐步建構完整之電腦音樂教育環境，使台灣發展成爲世界級藝術科技之跨領域研究重鎮。

參考文獻：

- Boulandger, R. (2000). *Csound Book, Perspectives in Software Synthesis, Sound Design, Signal Processing, and Programming*. Cambridge, Massachusetts: The MIT Press.
- McPhee, C. (1976) *Music in Bali-A Study in Form and Instrumental Organization in Balinese Orchestral Music*. New York: Da Capo Press.
- Morgan, R. P. (1991). *Twentieth-Century Music-A History of Musical Style in Modern Europe and America*. New York: W. W. Norton and Company.
- Sears, F. W. (1986). *University Physics*, Singapore: Addison-Wesley Publishing Company.
- Titon, J. T. (1996). *World of Music-An Introduction to the Music of the World's Peoples*. New York: Prentice Hall International.