

行政院國家科學委員會專題研究計畫 成果報告

生物訊號之瞬時頻率頻譜分析

計畫類別：個別型計畫

計畫編號：NSC91-2213-E-002-059-

執行期間：91年08月01日至92年07月31日

執行單位：國立臺灣大學機械工程學系暨研究所

計畫主持人：伍次寅

共同主持人：王淑音

計畫參與人員：王傳禎、蘇致遠

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 92 年 11 月 14 日

生物訊號之瞬時頻率頻譜分析

Instantaneous Frequency-Time Analysis of Biological Signals

計畫編號：NSC 91-2213-E-002-059

執行期限：91年8月1日至92年7月31日

主持人：伍次寅 國立台灣大學機械工程學系

共同主持人：王淑音 私立中國文化大學畜產學系

計畫參與人員：王傳禎、蘇致遠 國立台灣大學機械工程學系

一、中文摘要

一般自然和物理界中所觀察到的振盪訊號多半呈現出非穩態(non-stationary)及非線性(nonlinear)之特性。傳統傅立葉(Fourier)分析法僅能求得整段訊號之平均振盪頻率；且傅立葉頻譜(Fourier spectrum)中之眾多諧和模態(harmonic modes)乃是傅立葉解析(Fourier decomposition)過程中將訊號展開成固定頻率諧和函數(harmonic functions)組合之產物，其與訊號本身之振動型態並無直接對應之物理關聯。對於振動頻率及振幅隨時都在改變之非穩態時間訊號而言，傅立葉頻譜無法提供訊號瞬時振動頻率變化及其相關訊息。

本計畫運用新近發展之瞬時頻率頻譜分析法(instantaneous frequency-time analysis)[1]來分析從生物體中所記錄的一些訊號。此法主要分成兩個步驟：首先將訊號分解成一系列非固定頻率之本質模態函數之組合(intrinsic mode functions, IMF)；接下來再將這些函數經 Hilbert 轉換以求得各個模態函數之瞬時頻率(instantaneous frequency)與振幅(amplitude)隨時間變化之關係。藉由此種方式所獲得之 Hilbert 頻譜代表著原始訊號所含能量於頻域(frequency domain)和時域(time domain)上的分佈狀態。相較於傅立葉頻譜僅能提供訊號平均能量之分佈，瞬時頻率頻譜更提供了訊號振盪頻率和能量隨時間變化之情形。由於此法是根據訊號局部時間尺度來將原始訊號分解成各個不同的模態函數，因此可用於分析處理非穩態之時間序列。本計畫初步分析了兩組時間序列；一為人體的脈搏訊號(blood pulse signal)，一為不同等級生物之 MHC (Myosin Heavy Chain)第二型基因的 DNA 序

列。前者為人體健康情形的一個重要指標，後者為生物體遺傳及生化反應指令訊息密碼之一；兩者皆呈現出明顯非穩態和出自於非線性系統之特性。本計畫的目的在於嘗試分析並解讀隱含於生物訊號內大量但尚未為人所知的訊息；同時藉由分析的過程來評估瞬時頻率頻譜法之優缺點作為日後改進的依據。

(關鍵詞：瞬時頻率頻譜、Hilbert轉換、本質模態函數、經驗模態分解、Hilbert頻譜、生物訊號)

英文摘要

Many processes appearing in physics and natural world are non-stationary and nonlinear in features. Conventional Fourier spectral analysis can only provide global (average) energy-frequency distributions. Validity of such an analysis relies crucially on the assumption that the data are linear and stationary. For data originating from nonlinear systems or representing non-stationary phenomena, Fourier analysis is incapable of extracting local (instantaneous) time-frequency variation information embedded in the data.

In the present project, we adopt recently developed instantaneous frequency-time spectral analysis method [1] to analyze two data sets recorded from biological systems. The method involves two major steps. The first step is to decompose the original signal into a series of intrinsic mode functions (IMF) spanning from high frequency to low frequency ranges via a sifting process. The second step is to apply the Hilbert transform to these IMF, yielding the instantaneous frequencies as functions of time

for each mode function. The final presentation of the results is the Hilbert spectrum running along the data string that represents the energy distribution over the frequency-time domain. Since the decomposition process (called the empirical mode decomposition, EMD in [1]) is based on the local time scale of the signal, the above method is applicable to nonlinear and non-stationary data. The two example data sets analyzed in this project are the blood pulse signals recorded from human being, and the DNA sequences of the MHC type II genes. Both data strings are apparently non-stationary and nonlinear in nature. The purpose of this project is to explore the intriguing yet unknown information possibly embedded in these biological signals, and to assess the merits as well as some potential problems or difficulties associated with the method.

(keywords: instantaneous frequency-time analysis, Hilbert transform, intrinsic mode functions, empirical mode decomposition, Hilbert spectrum, biological signals)

二、計畫緣由與目的

傅立葉頻譜分析(Fourier spectral analysis)在傳統的訊號及資料處理上一直佔有著一席重要的地位。其方法是將數據資料經由傅立葉轉換(Fourier transform)展開成一序列固定頻率諧和函數(harmonic functions)的線性組合，各個不同頻率諧和函數的大小與頻率間之關係則可視為數據訊號所含能量分佈於各個不同頻率的情形。對於一個振盪訊號而言，傅立葉頻譜(Fourier spectrum)中出現峰值之處即代表原始訊號之主要(平均)振盪頻率。一般而言，傅立葉頻譜分析法之主要缺失有下述兩方面：

(1) 由於建構原始資料訊號的基本函數為**固定頻率**之諧和函數，傅立葉頻譜僅適用於分析呈穩態分佈之時間序列(stationary time series)——亦即資料分佈的各個統計值(譬如數據平均值與自我關聯函數等)不隨時間變化的序列。對於一個非穩態形式的時間序列，傅立葉頻譜充其量僅提供了數據訊號於整段取樣時間內之平均振盪頻率和平均能量分佈，此平均資訊並未確切反應出訊號的特性。

(2) 由於傅立葉轉換是將原始資料訊號展開成特定諧和函數的組合，因此用之分析非線性系統中常見之非諧和振盪訊號時，傅立葉頻譜會出現許多額外的峰值於諧和頻率(harmonic frequencies)處，在線性波分析中稱之為波形的諧和扭曲(harmonic distortion)現象。此額外的頻率與原訊號之間並無直接對應的物理意義，純粹只是因為傅立葉展開過程中需疊加許多高頻之諧和函數來模擬被扭曲之波形。

根據上述的論點，以傅立葉頻譜分析非穩態或非線性系統所產生之數據易導致錯誤或無意義的解讀。於是，就有了所謂之時間—頻率分析法(time-frequency analysis)的發展；亦即將能量頻譜隨時間之變化也納入考量的一種分析法。

現行主要之時間—頻率分析法大致有三種：Spectrogram [2]，Wavelet Transform [3-4]，以及 Wigner-Ville distribution [5]。Spectrogram 又稱為 Short-Time Fourier Analysis 或 Windowed Fourier Analysis；其方法是將訊號 $X(t)$ 先乘上一中心位於時間 τ 之 window 函數 $w(t-\tau)$ 予以局部化後，再作傅立葉轉換，移動時間 τ 則得到一隨時間變化之局部傅立葉頻譜，稱之為 spectrogram。Wavelet 轉換是由 Morlet 等學者於 1980 年代初期所提出；其概念是將訊號 $X(t)$ 與一稱之為 wavelet 的函數 $\psi(\frac{t-b}{a})$ 作內積運算(inner

product)。若將 wavelet 函數 $\psi(\frac{t-b}{a})$ 視為一頻通脈衝響應(bandpass impulse response)的話，則 wavelet 分析可視為一 bandpass 分析。改變 wavelet 函數的尺度 a 及其中心時間位置 b 則可得到各個時間訊號能量於不同尺度(頻率)上之分佈狀況。由於各個不同 a 情形下之子函數與原來之母 wavelet 函數 $\psi(t-b)$ 相似，僅有尺度上之差異；因此，wavelet 轉換極適合應用於分析具碎形(fractal)特性之時間序列。至於 Wigner-Ville distribution 基本上是先定義訊號 $X(t)$ 之時間—頻率自我相關函數(time-frequency autocorrelation function)，然後再將此函數作時間和頻率雙重之傅立葉轉換(2-D Fourier transform)，所得出之時間—頻

率分佈函數稱之為 Wigner-Ville distribution。此法之優點在於時域以及頻域上之局部解析度較 wavelet 法來得高；其缺點為 Wigner-Ville 分佈函數並不保證處處為正值，因此無法以訊號能量分佈的觀點來解釋此函數。

上述三種方法有一共同缺點，亦即在頻譜解析上皆是以固定頻率與振幅之函數作為基本模態函數。例如 Spectrogram 及 Wigner-Ville distribution 是以固定頻率與振幅之諧和函數作傅立葉轉換；Wavelet Transform 中一系列代表不同尺度(頻率)的子波(wavelets)亦是由預先選定的母函數來決定；因此，用之於分析非線性時間序列上仍有著上述(2)之缺失。除此之外，上述三種方法皆無法明確定義訊號之瞬時振盪頻率。Huang *et al.* [1]認為應將訊號展開成不定頻率與振幅之模態函數的組合，且此模態函數必須能反應出此訊號非線性及非穩態之特性。於是 Huang 等人提出了一稱之為經驗模態分解(empirical mode decomposition)的方法，自訊號中篩選出一序列代表不同尺度之本質模態函數(intrinsic mode function, IMF)，將訊號表達成這些模態函數的組合，再運用 Hilbert 轉換求出這些本質模態函數之瞬時振盪頻率與振幅。不同於前述預先選定模態函數來展開訊號的形式，此處之本質模態函數是根據訊號本身局部尺度所篩選出，因此頗能反應出訊號局部振盪的訊息。藉由 Hilbert 轉換所得到之時間-頻率函數則代表著訊號在各個瞬間分佈於各頻率的情形。

Huang *et al.* [1]將其所提出之經驗模態分解法應用於一些特定的訊號例子上，結果驗證了此方法確實較傳統的傅立葉及 wavelet 分析法更具物理意義。本計畫即是運用 Huang 的方法來分析兩組訊號，一為記錄自人體的脈搏生理訊號，另一組是由 DNA 鹽基序列所轉換之生物訊號。脈搏是人體健康與否的重要指標之一，包括心、肺、肝、腎、腸胃等內臟之功能和運作多少皆會反應於脈搏中。若能分解整理出隱含於脈搏中之訊息，則對未來疾病之診斷必有所助益。同樣的，生物體 DNA 上之鹽基(僅 A、G、C、T 四種)排列順序蘊藏著大量的遺傳訊息與指令。到底 DNA 上之鹽基序列是否有模態可循？亦或是

一隨機(random)的模式？DNA 長鏈上帶有遺傳指令的片段(exon，亦即會轉譯成蛋白質的片段)與不帶遺傳訊息的插入片段(intron)其鹽基排列型態有何不同？相同類型之基因在不同演化等級生物中之 DNA 序列有何差異？這些問題目前或許沒有明確的答案。本計畫嘗試從一個不同於傳統分子生物學的觀點出發，希望能提供研究此類問題的另一個具發展潛力的選擇途徑。

三、研究方法

1. 經驗模態分解法

傳統上有許多定義訊號瞬時頻率(instantaneous frequency)的方法，其中以 Hilbert 轉換最適用。針對任一時間序列 $X(t)$ ，求取其 Hilbert 轉換得到另一序列 $Y(t)$ ，

$$Y(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{X(\tau)}{t-\tau} d\tau ; \quad (1)$$

結合 $X(t)$ 和 $Y(t)$ 可定義出一複數時間序列 $Z(t)$ ，

$$Z(t) = X(t) + iY(t) = a(t) e^{i\theta(t)}, \quad (2)$$

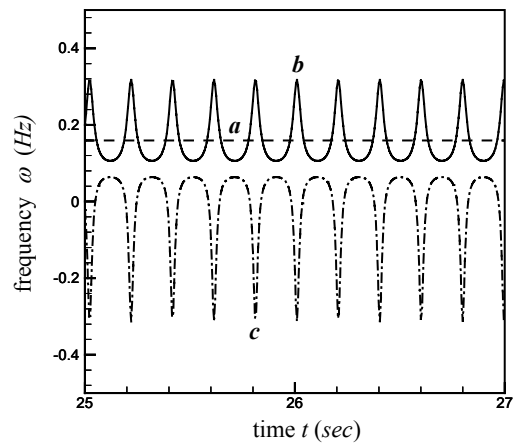
其中 $a(t) = [X^2(t) + Y^2(t)]^{1/2}$ 為此複數訊號之振幅， $\theta(t) = \tan^{-1}[Y(t)/X(t)]$ 為訊號振盪之相位，兩者皆為時間之函數。數學上可證明由 Hilbert 轉換(1)與原時間序列所組成之複數時間序列 $Z(t)$ 為一解析複變函數(analytic function)。由第(2)式中之相位角 $\theta(t)$ ，我們可定義時間序列 $X(t)$ 之瞬時振盪頻率 $\omega(t)$ 為

$$\omega(t) = \frac{d\theta(t)}{dt}. \quad (3)$$

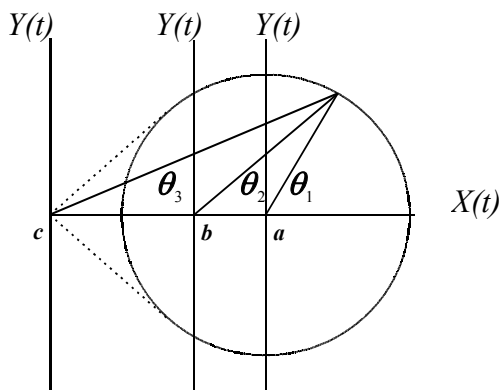
然而，以此種方式得出之瞬時頻率仍存有一些問題[1]。其中之一為第(3)式為單值函數(single-valued function)；亦即在任一時刻僅定義了唯一的一個頻率值，此與一般複雜訊號是由多個不同振盪頻率所組成之概念不符。第二個問題為倘若將原始訊號 $X(t)$ 直接作 Hilbert 轉換，所得到之瞬時頻率 $\omega(t)$ 在很多情況下不具任何物理意義。關於此點可以下例來說明。

茲考慮一簡單且有著固定振盪頻率的時間函數 $X(t) = \sin t$ ，經 Hilbert 轉換後得

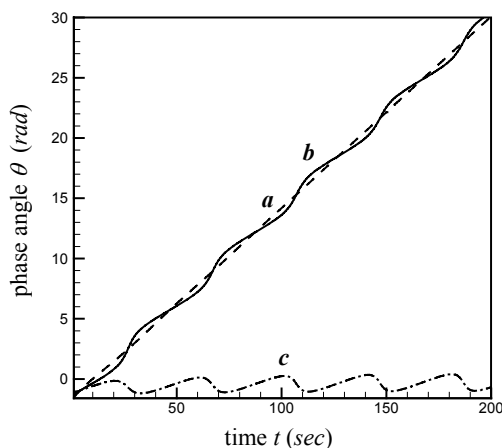
$Y(t) = -\cos t$ ，兩者所組成之複數函數 $Z(t)$ 在複數平面上為一中心位於座標原點之圓形軌跡(參見圖一)；其相角 $\theta(t)$ 隨時間呈線性遞增，根據(3)式之定義頻率 $\omega(t)$ 為一定值，與原函數之振盪頻率相符(圖二與三中之 **a** 例)。然而相同的函數若加上一常數值， $X(t) = \alpha + \sin t$ ；亦即頻率不變但振盪平均值不為零時，由定義(3)所得出之頻率 $\omega(t)$ 則不再是一個定值且甚至不具物理意義。譬如當 $0 < \alpha < 1$ 時 $Z(t)$ 的軌跡為一偏心圓，相角 $\theta(t)$ 不再是一條直線，由此相角所推得之瞬時頻率 $\omega(t)$ 也不是應有的定值。當 $\alpha > 1$ 時，瞬時頻率甚至形成了負值(圖二與三中 **b** 與 **c** 之例)。為了讓(3)式所定義之瞬時頻率符合物理意義，原函數 $X(t)$ 在作 Hilbert 轉換前需先行經過適當的處理。



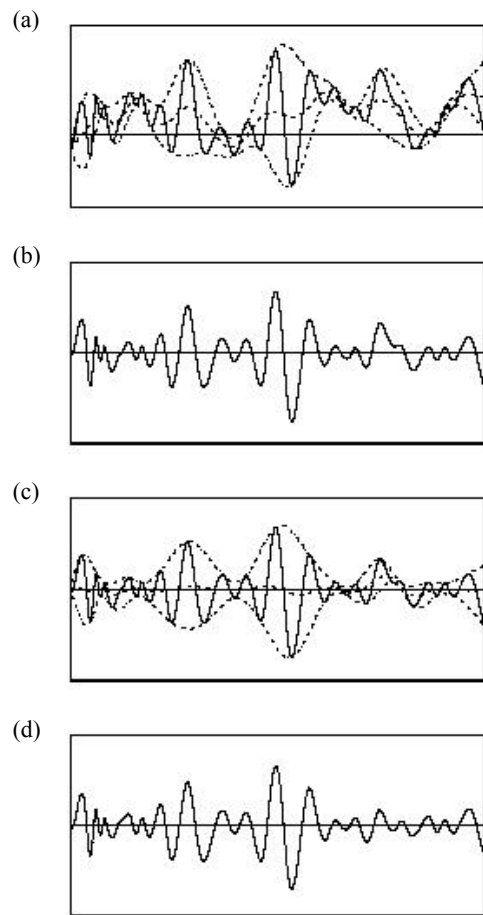
圖(三)、複數函數 $Z(t)$ 之瞬時頻率。



圖(一)、複數函數 $Z(t)$ 之軌跡。



圖(二)、複數函數 $Z(t)$ 之相位角。



圖(四)、經驗模態分解步驟圖示。

從上面的例子裡，Huang *et al.* [1] 獲致一個結論。一個具備符合瞬時頻率物理意義的時間序列必須滿足兩個條件：(1) 此時間序列於任兩相鄰局部極值(local extremes)之間必須通過一零點(zero point)；(2) 此時間序列的局部平均值(local mean)必須為零。為了將任一

時間序列 $X(t)$ 分解成符合上述二要求之基本模態函數的組合，Huang *et al.* [1] 提出了下述的分解步驟。

(I) 首先將序列 $X(t)$ 之局部極值(包括極大值與極小值)以 cubic spline 法連結成上下二封包線(envelopes)，再求取此二封包線之平均曲線，稱其為 m_{11} (見圖四之 a)。

(II) 求出序列 $X(t)$ 與 m_{11} 之差值，稱其為 d_{11} (圖四之 b)，

$$d_{11} = X(t) - m_{11}。 \quad (4)$$

理論上，此 d_{11} 序列應大致滿足上述二條件。否則，重複(I)之步驟於 d_{11} 序列；亦即求出 d_{11} 序列上下封包線之平均曲線 m_{12} (圖四之 c)，再求出 d_{11} 與 m_{12} 之差值(圖四之 d)，

$$d_{12} = d_{11} - m_{12}。 \quad (5)$$

通常經過再次篩選的序列 d_{12} 會較先前之序列 d_{11} 對稱(參考圖四之 d)，如此重複上述步驟 k 次後所得之序列 d_{1k}

$$d_{1k} = d_{1k-1} - m_{1k}； \quad (6)$$

若 d_{1k} 序列已滿足上述二條件，則將 d_{1k} 視為一模態函數，並以符號 M_1 表達之。由於此函數是由訊號本身所分離出來的，因此 Huang *et al.* [1] 將其定名為此訊號之本質模態函數(intrinsic mode function, IMF)，以區分傳統之傅立葉譜和模態函數。由分解的過程可得知 M_1 代表著訊號所含最小時間尺度(time scale)，亦即最高頻部分的訊息。

(III) 將 M_1 自訊號 $X(t)$ 中分離出；即求出序列 $X(t)$ 與 M_1 之差值，稱其為 r_1 ：

$$X(t) - M_1 = r_1。 \quad (7)$$

由於 r_1 序列仍含有許多中低頻之資訊待分解出，因此多次重複上述(I)至(III)步驟之後，可將代表訊號各個不同時間尺度訊息之模態函數 M_2 、 M_3 ... 篩選出：

$$r_1 - M_2 = r_2, \quad r_2 - M_3 = r_3, \quad \dots \dots \quad (8)$$

$$r_{n-1} - M_n = r_n。$$

此篩選程序一直進行至最終之殘餘序列 r_n 不再含任何振盪訊息為止，此殘餘序列 r_n 即為非穩態訊號 $X(t)$ 之趨勢(trend)。將所有本質模態函數以及最後之趨勢迭加，則可得回原訊號 $X(t)$ ：

$$X(t) = \sum_{i=1}^n M_i + r_n。 \quad (9)$$

如同傅立葉分解，Huang 等學者將一訊號展開成 n 個代表不同局部時間尺度訊息之經驗模態函數的組合。

由於模態函數 $M_1 \cdots M_n$ 基本上皆已滿足定義瞬時頻率所需具備的二條件，因此經由(1)式之 Hilbert 轉換後可得到各模態函數 M_j 之瞬時振盪頻率 $\omega_j(t)$ 和振幅 $a_j(t)$ ，而原始訊號 $X(t)$ 可表成：

$$X(t) = \text{Re} \left\{ \sum_{j=1}^n a_j(t) e^{i \int \omega_j(t)} \right\}。 \quad (10)$$

在式(10)中，若將振幅 a_j 表為時間 t 與頻率 ω_j 的函數，則可得到一瞬時頻率頻譜(instantaneous frequency-time spectrum)，或簡稱為 Hilbert 頻譜 $H(\omega, t)$ 。由於振幅的平方值 a_j^2 正比於能量，因此 Hilbert 頻譜 $H(\omega, t)$ 代表著訊號 $X(t)$ 之能量在任一時刻 t 分佈於各個不同頻率上的情形。

2. 應用於生物訊號分析

本計畫中分析了下列兩組訊號：

(一) 人體脈搏訊號

利用一個壓電材料的感應器(sensor) 連接上 PowerLab 公司的 ADInstrument 多功能生理訊號資料擷取儀，以每秒 200 個資料的擷取速率持續量測 100 秒，我們記錄了一位女性受測者在其懷孕 36 周以及產後恢復正常時的脈搏訊號；然後利用前述之經驗模態分解法配合 Hilbert 轉換之理論來進行訊號的分析。本計畫除了分析解讀測試者懷孕與未懷孕時脈象之差異，所得之結果並與應用 Morlet wavelet 轉換所得出之頻譜做比較。

(二) MHC type II 基因家族之 DNA 序列

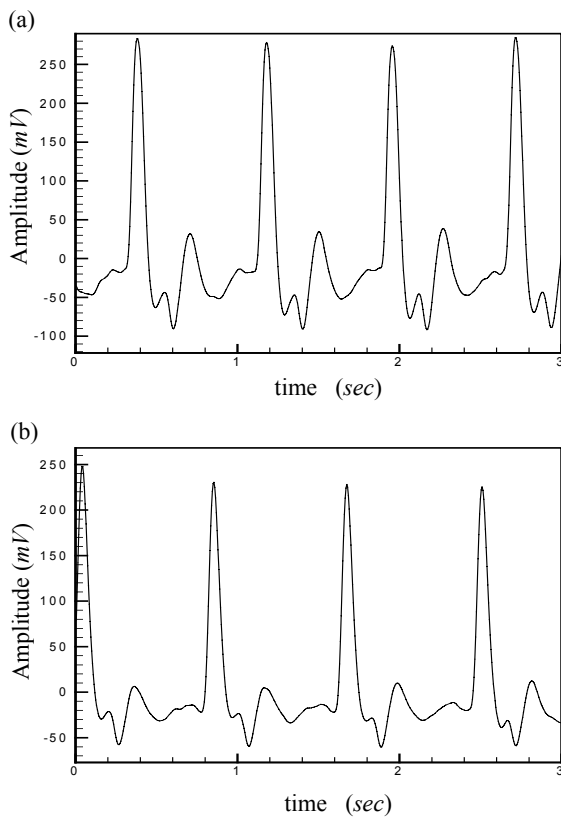
MHC (Myosin Heavy Chain) type II 基因為常見於動物肌肉內的一種收縮蛋白，掌控肌肉的收縮運動。由於 DNA 分子鏈上的鹽基只有嘌呤(purine A、G)與嘧啶(pyrimidine C、T)兩類，我們首先運用碎形(fractal)的觀念將 MHC 基因之 DNA 序列中代表嘌呤鹽基(A、G)局部分布疏密程度之 Hölder 指數 α 求出[6-7]，如此可得出類似時間振盪之 α 值訊

號。運用上述之經驗模態分解法與 Hilbert 轉換可求得此 α 訊號之 Hilbert 頻譜。各個不同等級之動物譬如老鼠、雞、人類皆有 MHC type II 基因，經由演化，其基因容或不盡相同。藉由研讀比對 Hilbert 頻譜，或許可幫助我們了解基因內 DNA 之鹽基排列型態與生物演化層次之間的關聯。

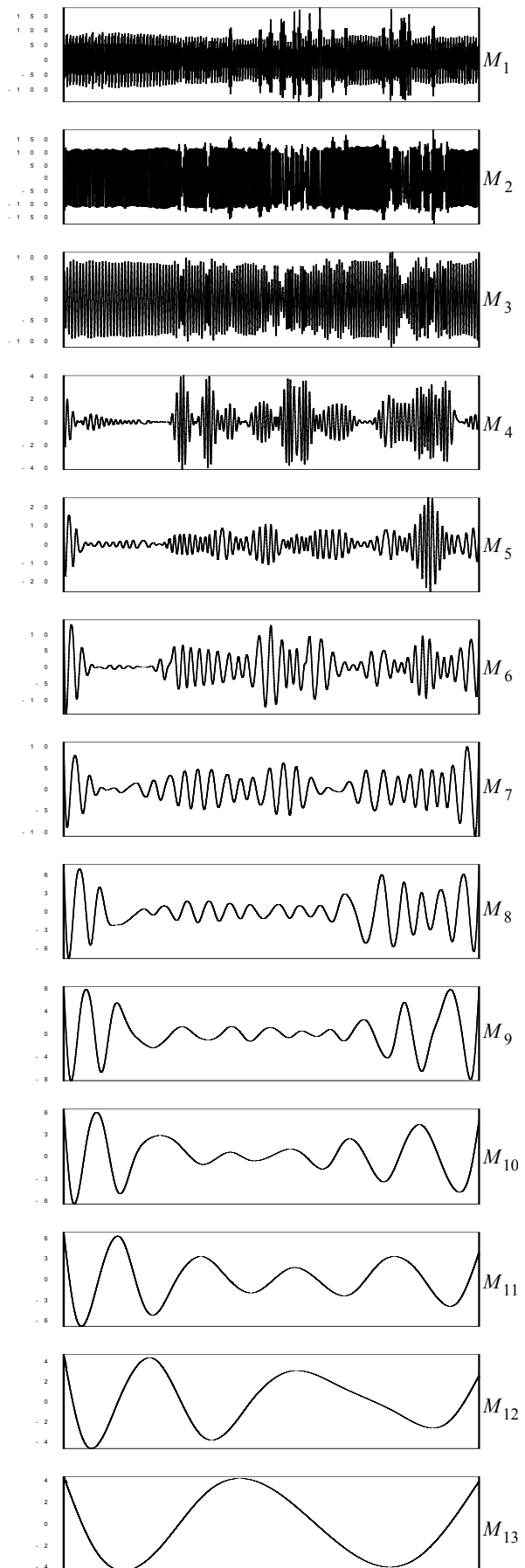
四、研究結果

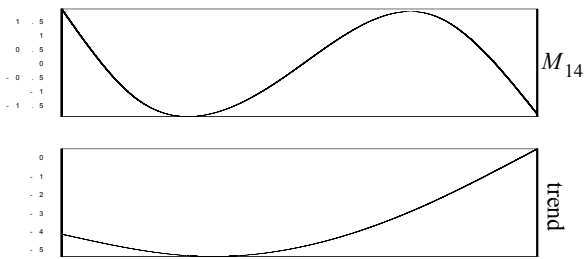
1. 脈搏訊號分析

圖五為記錄自受測者懷孕 36 周及未懷孕時橈動脈處之脈搏訊號，單從脈搏之波形無法分辨出兩者之差異。運用經驗模態分解法則可分離出如圖六(懷孕 36 周)與圖七(未懷孕)所示之由高頻至低頻的 14 個與 16 個本質模態函數(IMF)，和代表訊號非穩態特性的一個趨勢(trend)。比較二者，我們可發現懷孕 36 周時脈搏所分離之本質模態函數數目較未懷孕時少 2，且懷孕 36 周時心跳頻率模式出現在第三個模態 M_3 ，而未懷孕時心跳頻率模式則是出現在第四個模態 M_4 。

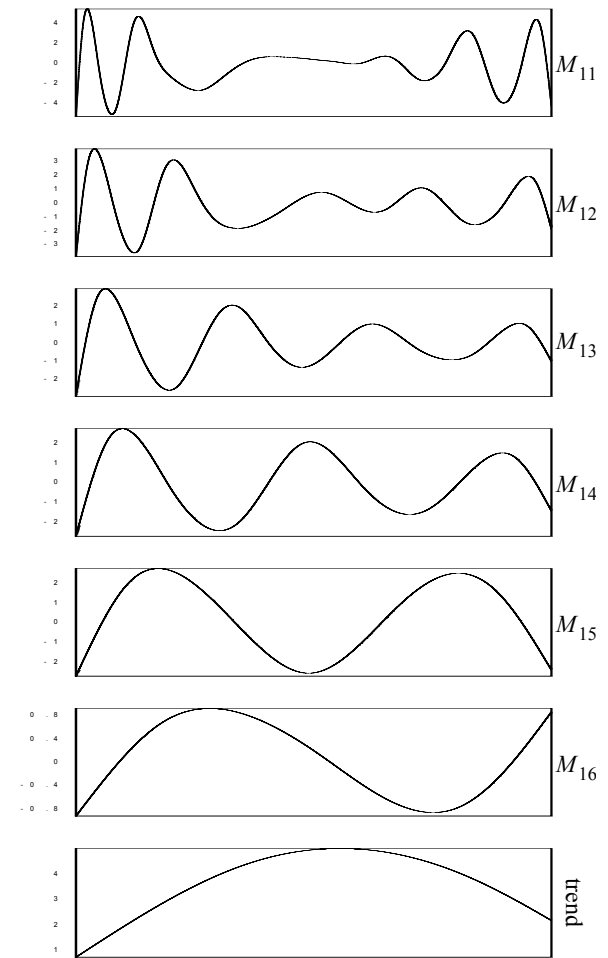
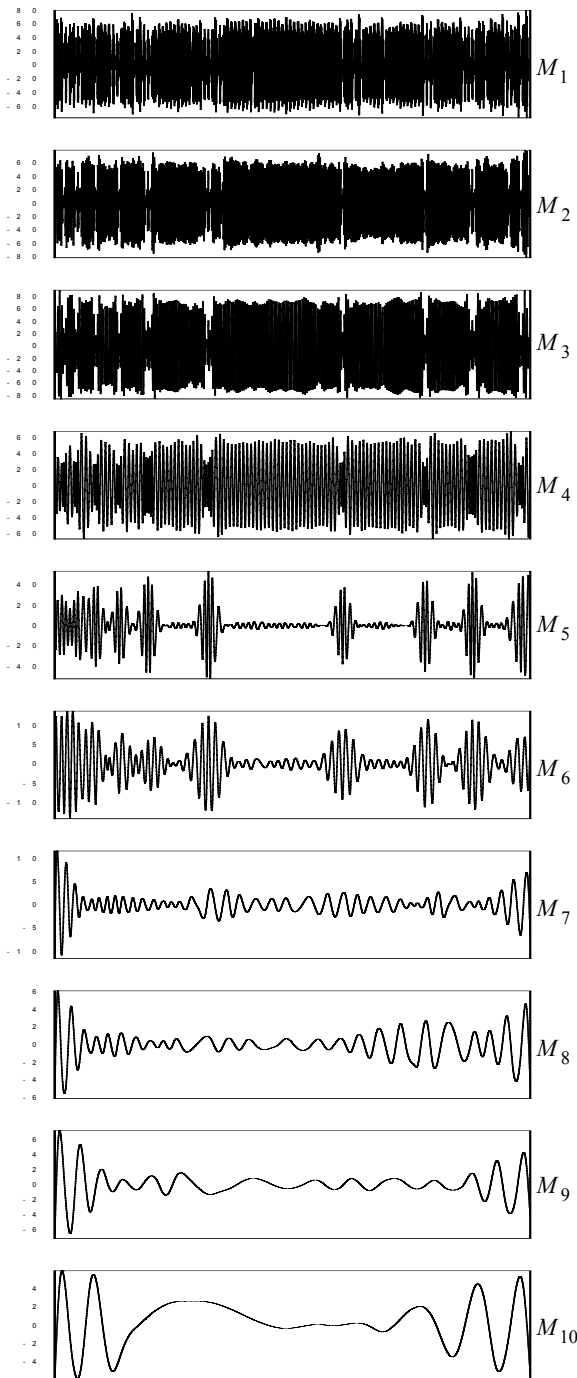


圖(五)、脈搏訊號(a)懷孕 36 周(b)未懷孕。





圖(六)、受測者懷孕 36 周時脈搏訊號所分解出之本質模態函數。

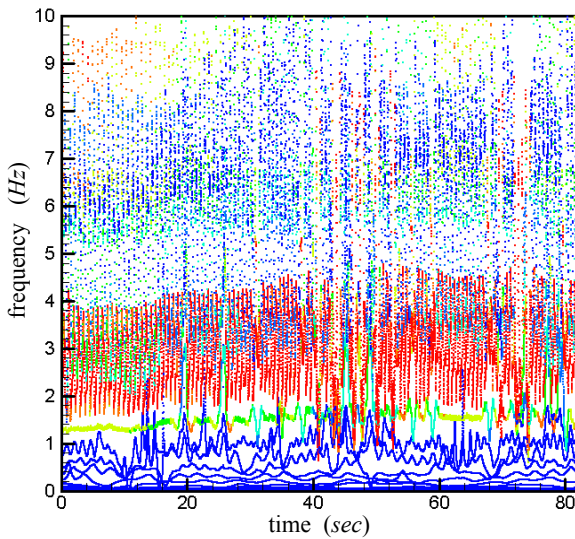


圖(七)、受測者未懷孕時脈搏訊號所分解出之本質模態函數。

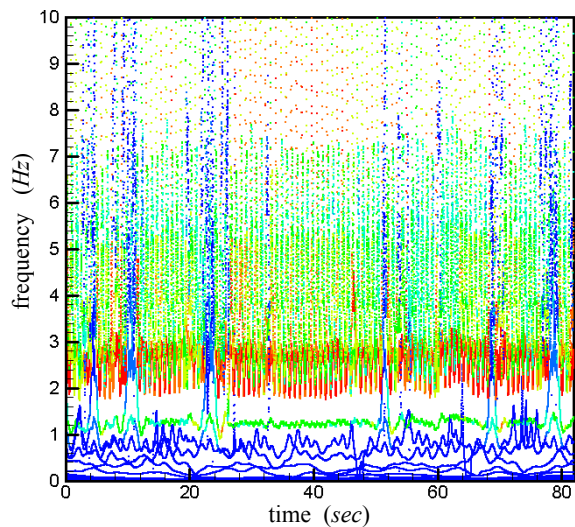
圖六與七中所分解出脈搏訊號之本質模態函數經由 Hilbert 轉換式(1)轉換後與原模態函數結合，可作出原始脈搏訊號之 Hilbert 頻譜圖，即圖八(懷孕 36 周)與圖九(未懷孕)。此 Hilbert 頻譜圖分別由 14 條(圖八)與 16 條(圖九)曲線所組成，每條曲線代表各個模態函數之瞬時頻率隨時間變化的情形。各個模態函數瞬時振幅 $a_j^2(t)$ 之大小以不同顏色標示出(顏色越紅處振幅值越大)，因此由顏色的變化可看出脈搏訊號能量於時域及頻域上之分布狀況。圖八中顯示了懷孕 36 周時脈搏訊號絕大部分能量是集中於頻率介於 2 至 4.5 Hz 之間的模態上(M_2)，而未懷孕時(圖九)脈搏訊號能量主要是分布在頻率介於 2 至 3 Hz 之間的模態上(M_3)。此意味著懷孕時脈搏訊號能量會轉移至較高的頻率上，且原先(未懷孕時)脈搏訊號的第二與第三個模態會合併形

成一振盪能量較高之單一模態。由圖八及圖九亦可清楚看出懷孕與未懷孕時心跳主頻率平均值大約在 1.6 Hz (96 beat/min) 及 1.3 Hz (78 beat/min) (圖八與九中較平穩之綠色曲線)。

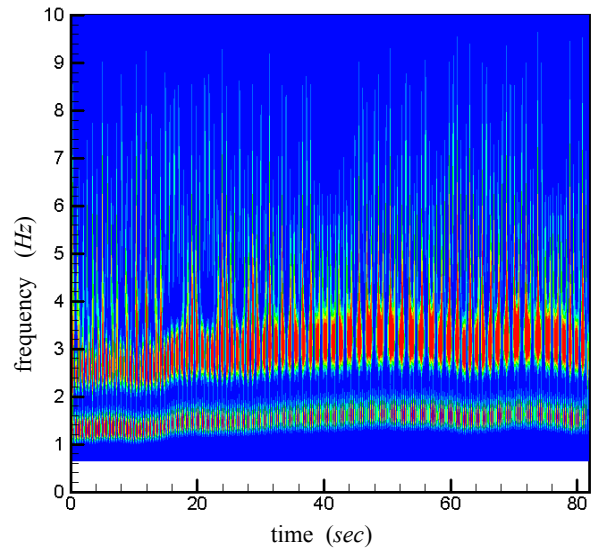
圖十與圖十一為應用 Morlet wavelet 對脈搏訊號作 Wavelet 轉換後所得出之頻譜圖。與圖八和圖九對照，Wavelet 頻譜圖上之能量分布趨勢大致與前述 Hilbert 頻譜相符；然而無論時域或是頻域上訊號的解析程度，Wavelet 頻譜皆不如 Hilbert 頻譜詳盡。尤其，Wavelet 頻譜無法提供訊號瞬時振盪頻率與振幅等資訊。



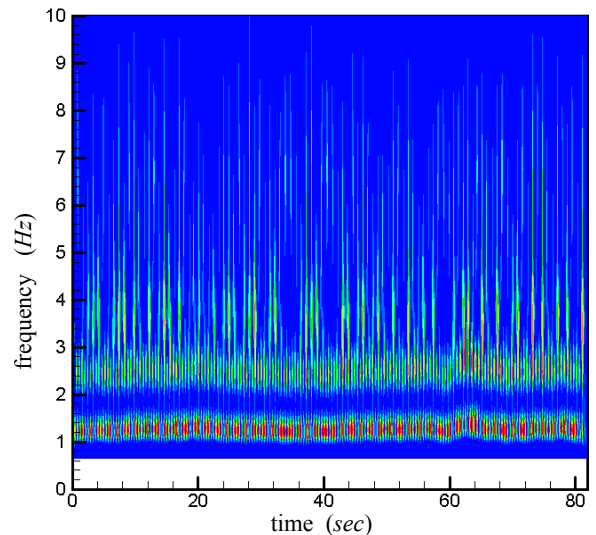
圖(八)、脈搏訊號(懷孕 36 周)之 Hilbert 能量頻譜圖。



圖(九)、脈搏訊號(未懷孕)之 Hilbert 能量頻譜圖。



圖(十)、脈搏訊號(懷孕 36 周)之 Morlet Wavelet 能量頻譜圖。

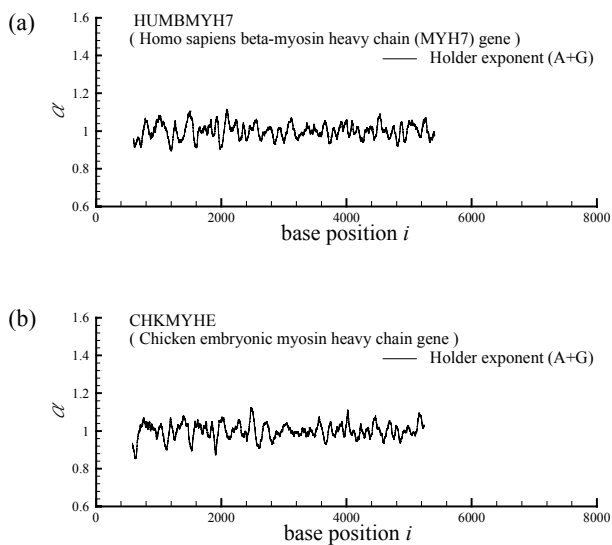


圖(十一)、脈搏訊號(未懷孕)之 Morlet Wavelet 能量頻譜圖。

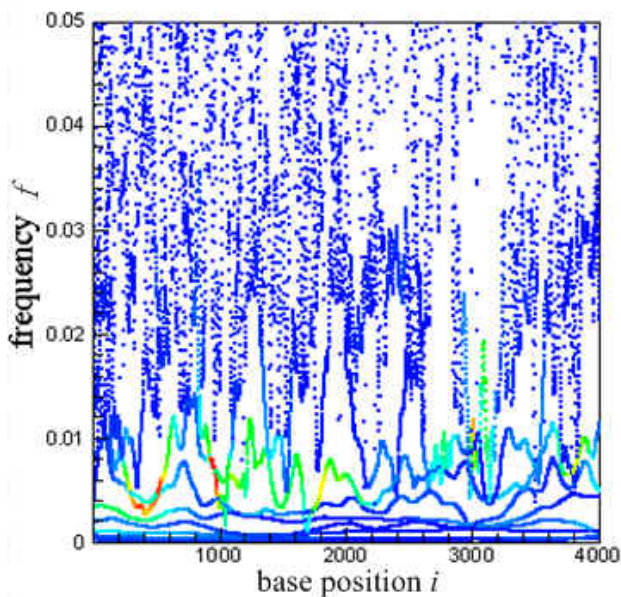
2. MHC type II 基因家族之 DNA 序列分析

圖十二為人類與雞 MHC 基因 exon 片段(即會轉譯成蛋白質的部分)鹽基序列之局部分布密度比例指數(Hölder exponent) α 曲線圖(有關碎形 Hölder 指數之定義與其求取方法請參考作者上一年度國科會計畫之結案報告[6])。 $\alpha = 1$ 代表嘌呤鹽基在 DNA 長鏈上為均勻分布； $\alpha < 1$ 代表嘌呤鹽基局部分布呈中心密，向外漸疏之形式；而 $\alpha > 1$ 則是以內疏外密之形式分布[7]；因此由 α 曲線圖可了解特定鹽基在 DNA 序列中分布疏密之變化情形。圖十三與圖十四作出了圖十二中兩條 α 曲線

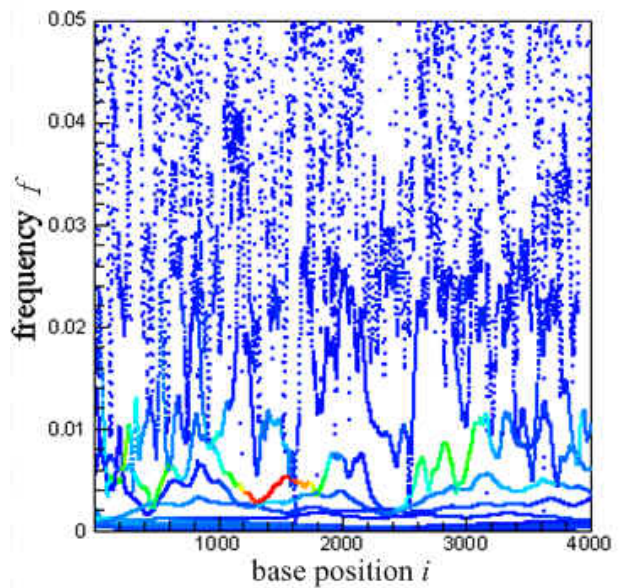
之 Hilbert 頻譜分布。由圖中可得知二者之能量頻譜非常類似，二者皆可分解出 12 個本質模態函數，且主要能量分布的位置亦相當接近。由於此處所分析的 DNA 序列是已將不會轉譯成蛋白質的部分(intron)剔除後的序列，因此同一家族系列之基因其鹽基分布亦應相近，Hilbert 頻譜證實了此論點。將來，我們或許可利用 Hilbert 頻譜來比對已知與未知功能基因間之相似性，藉以協助我們推測未知基因之功能。



圖(十二)、人類(a)與雞(b)MHC 基因 exon 片段嘌呤鹽基分布之 Hölder 指數曲線圖。



圖(十三)、人類 MHC 基因 exon 片段嘌呤鹽基分布之 Hilbert 能量頻譜分布圖。



圖(十四)、雞 MHC 基因 exon 片段嘌呤鹽基分布之 Hilbert 能量頻譜分布圖。

如同作者在本計畫之提案書中所言，本研究的一個困難處在於資料之取得。我們費了將近一年的時間才記錄了一位受測者懷孕前後的脈搏訊號來供我們測試程式以及改進分析方法。由於數據的缺乏，因此我們在本研究中所觀察到測試者懷孕前後模態函數的變化，其與懷孕這個生理現象之間到底存在一個什麼樣的關聯；各個不同頻率層次之模態函數所對應以及代表的生理狀況又是如何？以目前的認知，這些問題皆尚未有明確的答案。無論如何，初步的分析結果顯示了以 Hilbert 轉換所得出之時-頻域頻譜確可成功的應用在生理與生物訊號之解析上。在接續的計畫裡我們預計與醫學院合作以取得大量臨床之脈搏訊號，藉由長期收集統計及比對不同癥候患者的脈搏分析資料後，祈能整理出脈搏訊息與生理狀況之間定性及定量的關聯。同樣的，在 DNA 的訊號分析中亦須經過大量計算基因資料庫(GenBank)所收錄眾多不同等級生物體之基因序列後，方能得出具體的結論。

五、參考文獻

[1] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H. (1998) "The Empirical Mode Decomposition and the Hilbert

- Spectrum for Nonlinear and Non-stationary Time Series Analysis,” *Proc. R. Soc. Lond.. A*, Vol. 454, pp. 903-995.
- [2] Gabor, D. (1946) “Theory of Communication,” *J. Inst. for Elect. Eng.*, Vol. 93, pp. 429-439.
- [3] Morlet, J. (1983) “Sampling Theory and Wave Propagation,” in *NATO ASI series, Issues in acoustic signal/image processing and recognition*, Vol. 1, pp. 233-261.
- [4] Morlet, J., Arens, G., Fourgeau, I., and Giard, D. (1982) “Wave Propagation and Sampling Theory,” *Geophysics*, Vol. 47, pp. 203-236.
- [5] Cohen, L. (1995) *Time-Frequency Analysis*. Englewood Cliffs, Prentice-Hall.
- [6] 伍次寅、王淑音、蘇致遠 (2002) “DNA 序列之多重碎形特性分析”，國科會專題研究計畫成果報告，NSC90-2213-E002-082。
- [7] Wu, T., Su, Z.-Y., and Wang, S.-Y. (2004) “Multifractal and Local Scaling Analyses of Nucleotide Sequences in DNA,” in submission to *Phys. Rev. Lett.*