

行政院國家科學委員會專題研究計畫 成果報告

離散時間分數維度高斯雜訊的條件熵和共有訊息

計畫類別：個別型計畫

計畫編號：NSC94-2213-E-040-001-

執行期間：94年08月01日至95年07月31日

執行單位：中山醫學大學資訊管理學系

計畫主持人：張炎清

計畫參與人員：王仁昭、李明仁

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 95 年 10 月 11 日

中文摘要：

離散時間分數維度布朗運動(discrete-time fractional Brownian motion, DFBM)和它的增量程序(increment process)，離散時間分數維度高斯雜訊(discrete-time fractional Gaussian Noise, DFGN)經常廣泛地被運用在許多自然現象和醫學領域方面，這些訊號大都不是以前者就是以後者出現。此外，由於此類程序可由一個有限範圍(0,1)的赫斯特參數(Hurst parameter)來描述，使得原本看似複雜的程序顯得格外簡單與方便。在應用上，只要估測出赫斯特參數就能具體解析訊號間的差異。

目前已有許多估測此一參數的有效方法，然而在應用上缺乏一個評估訊號是否適合使用的檢驗方式。本計畫的貢獻就是利用條件熵(conditional entropy)和共有訊息(mutual information)兩種方法來評估所關心的訊號是否可以使用，亦即判別訊號是決定性(deterministic)還是隨機性(random)的訊號。理論上，當兩個訊號來自同一個系統時則它們間的條件熵會較小，然而共有訊息會較大。由於此類訊號是隨機雜訊，因此它們的條件熵會較大，然而共有訊息會較小。透過這種驗證方式可以有效區分訊號是決定性或隨機性。

關鍵字：離散時間分數維度布朗運動、離散時間分數維度高斯雜訊、赫斯特參數、條件熵、共有訊息、決定性的、隨機性的。

英文摘要：

The discrete-time fractional Brownian motion (DFBM) and its increment process, discrete-time fractional Gaussian noise (DFGN), are often widely used in natural phenomena and medical field. These signals largely appear either the prior or the rear. Furthermore, the fact that these processes can be described by one Hurst parameter with limited range (0,1) makes the seemingly complicated processes simple and convenient. In application, we can concretely analyze the difference between different signals by estimating the Hurst parameter.

At present there exist many effective methods to estimate this parameter. However, in application it lacks of an evaluation method for whether the interested signals are appropriate for use or not. In this project, we apply conditional entropy and mutual information to assess whether these signals are appropriate for use or not, i.e., judge these signals whether are deterministic or random signals. In theory, as two signals are coming from the same system, they have smaller conditional entropy, but larger mutual information. Since this kind of signals is random noise, they have larger conditional entropy, but smaller mutual information. We can effectively separate determinism from randomness on signals via this test and verification.

Keyword: discrete-time fractional Brownian motion, discrete-time fractional Gaussian Noise, Hurst parameter, conditional entropy, mutual information, deterministic, random.

一、前言

離散時間分數維度布朗運動(discrete-time fractional Brownian motion, DFBM)以及它的增量程序(increment process)，離散時間分數維度高斯雜訊(discrete-time fractional Gaussian noise, DFGN)，經常廣泛地用來描述自然現象和醫學訊號。這兩種程序只需一個參數，也就是赫斯特參數[1]，即可決定整個程序的特性。此外，此參數的區間是有界的，也就是在(0, 1)。自然界的現象和醫學訊號以時域的角度而言非常的複雜，然而利用這種簡單的參數來描述這類訊號卻非常的方便。

關於赫斯特參數的估測法有很多，值得注意的是最大相似估測器(maximum likelihood estimator, MLE)[2]、移動平均 (moving average, MA) 法[3]和自回歸模式(autoregressive, AR)法[4-5]。這些方法中最精確的是最大相似估測器，然而它的收斂速度非常的慢，不適合用在實際的應用上。而最接近 MLE 且執行效率最佳的是自回歸模式法，因此非常適合用在實際的赫斯特參數估測上。

以往使用離散時間分數維度布朗運動或者離散時間分數維度高斯雜訊來描述自然現象和醫學訊號大都根據視覺的判斷來決定是否可以使用此種模式。以往的經驗雖然沒有造成太大的困擾，但畢竟欠缺一個有規則的方式來決定。因此本計畫想研究是否可以找到適當的方式來判斷訊號是決定性或隨機性，因為離散時間分數維度布朗運動和離散時間分數維度高斯雜訊是隨機性的。

二、研究目的

本計畫的研究目的在於找出一個有規則的方法來判別訊號是隨機性的或決定性的。在以往的做法中，都將重心放在區分模式的赫斯特參數估測上，只要估測出參數落在(0, 1)間即表示可以使用。在應用上大都以目測的方式評估訊號是否符合此類模式的範疇，沒有一個正式的判別準則。雖然大都可以運用，但畢竟欠缺一個具體的方法。為此本計畫的目的即是提出一個可行的方法來驗證並評估真實訊號是否符合該模式。

三、文獻探討

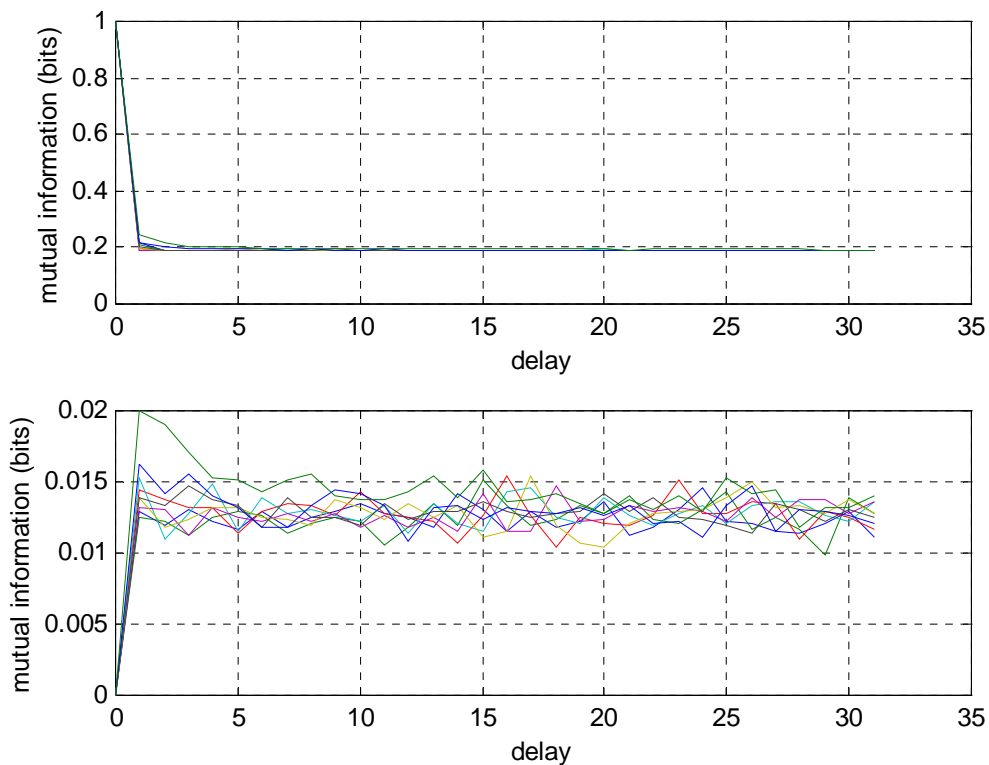
目前沒有直接的方法可以應用在此領域，或是以此目的為主題的研究探討，但在物理界的應用上有幾篇論文可以被適當地引用到本計劃中，其中一篇發表在 Physical Review Letters[6]，這方法主要是將訊號符號化(symbolization)，再去計算來自同一個系統中的兩個訊號相對延遲的條件熵(conditional entropy)，如果相對延遲為零則會有最小的條件熵。根據此一觀念，我們可以推論，來自不同系統的兩組訊號其條件熵會比來在相同系統的兩組訊號的條件熵還大，特別是兩組隨機的訊號，它們的條件熵會更大。根據條件熵的數值我們可以推斷訊號是否為隨機性或決定性，若是隨機性再去估測它們的赫斯特參數值。此外，在 IEEE Trans. Information Theory[7]有一篇利用共有訊息(mutual information)去尋找第一個最小的共有訊息，以便決定它的時間延遲作為重建時間序列(time-series)的相位圖像(phase-portrait)。因為第一個最小的共有訊息可以產生相對獨立的(independent)兩個時間序列。此法比傳統利用自我相關(autocorrelation)法更有說服力，因為共有訊息法所得到的結果是非線性獨立，而自我相關法是線性獨立。根據此一邏輯，我們可以推斷，如果該時間序列是來自像離散時間分數維度高斯雜訊的隨機訊號，則它們對時間延遲而言應該不會有明顯變化的共有訊息。如果現象是如此則可以推論此訊號來自隨機的訊號。

四、研究方法

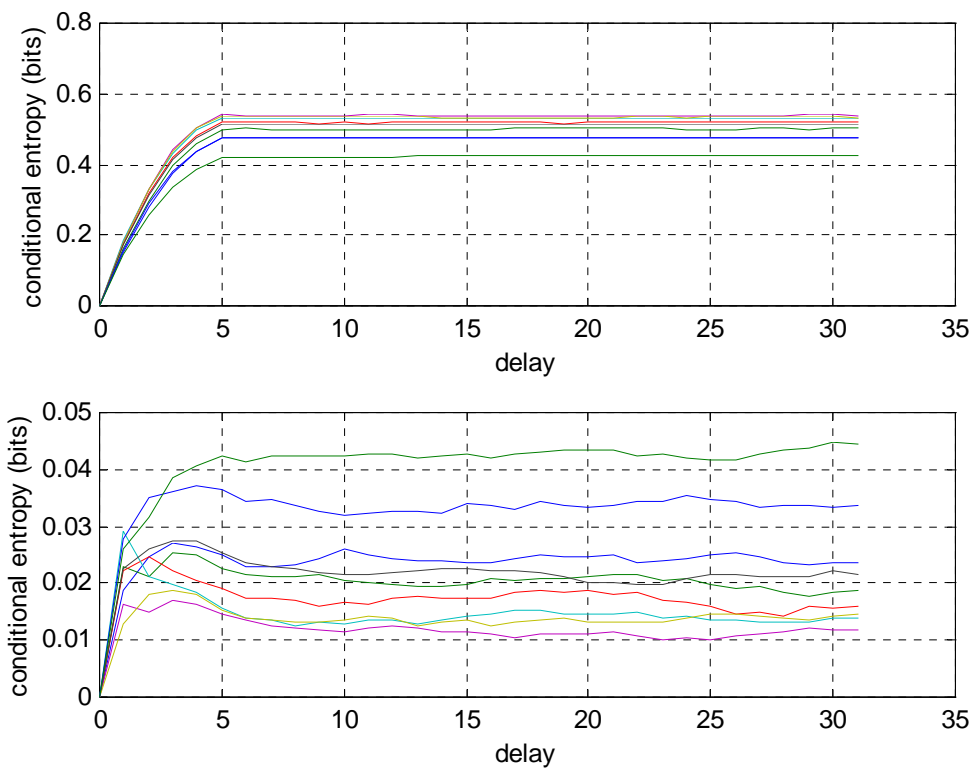
首先寫出計算條件熵和共有訊息的程式，接著產生數組決定性的訊號，例如弦波函數、勞倫茲方程式等，然後計算出它們的條件熵和共有訊息。其次，產生具代表性的離散時間分數維度高斯雜訊的訊號，這些訊號是隨機性的(random)，並計算出它們的條件熵和共有訊息。一般預估這兩類訊號所產生的條件熵和共有訊息會有明顯的差別，根據差異建立可行的參考準則，以作為判別的依據。最後將此方法應用在成鼠的尿道外括約肌 external urethral sphincter (EUS)肌電圖 electromyography (EMG)[8-10]的訊號上。

五、結果與討論

首先產生大小為 $N+M$ 的離散時間分數維度高斯雜訊的訊號，其中 N 代表所要處理的資料量， M 代表最大的延遲單位，例如 $M=31$ ，這些訊號的赫斯特參數分別是 $H=0.1, 0.2, \dots, 0.9$ ，針對每個 H 值產生 100 組的訊號，計算每組的前 N 個數據和其延遲(由 0 到 M ，每次遞增一個單位)所產生的 N 個數據的共有訊息和條件熵，其結果如下：

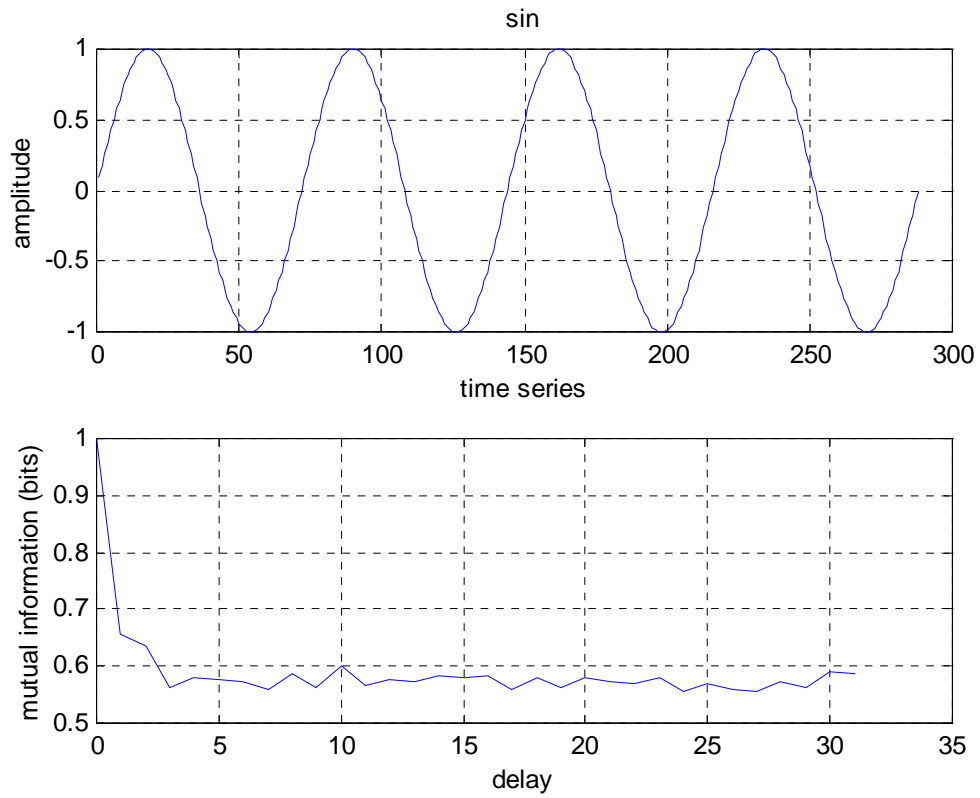


圖表 1: 上圖是 9 個 H 的平均共有訊息；下圖是 9 個 H 的共有訊息的標準變異量。

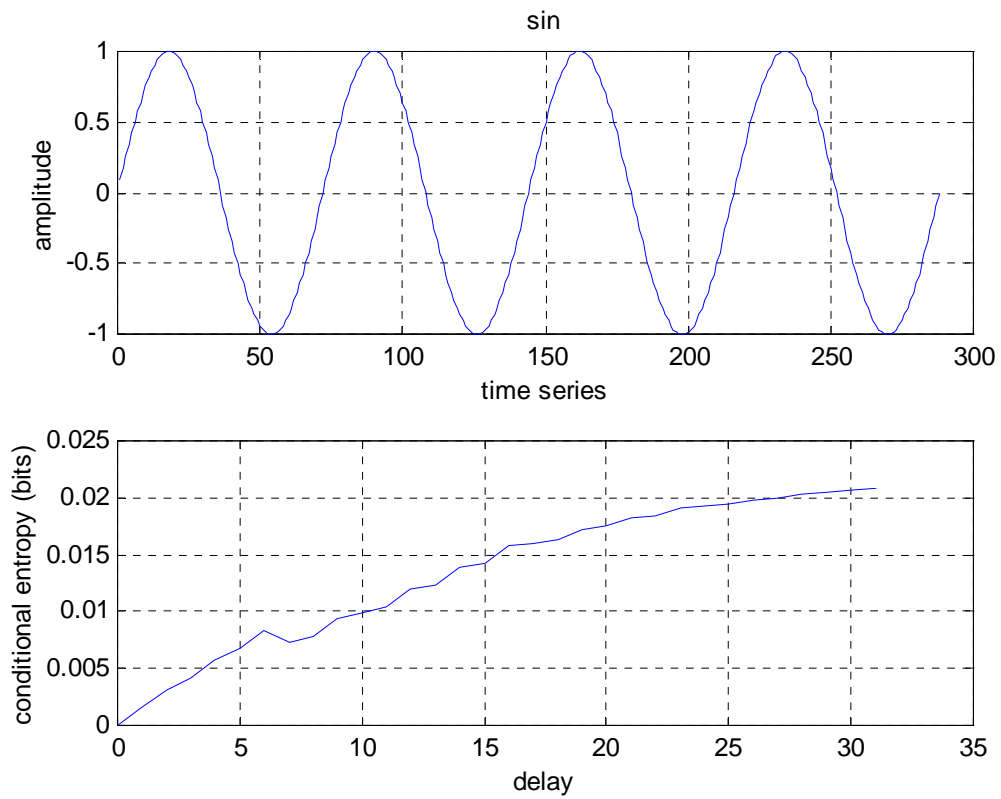


圖表 2: 上圖是 9 個 H 的平均條件熵；下圖是 9 個 H 的條件熵的標準變異量。

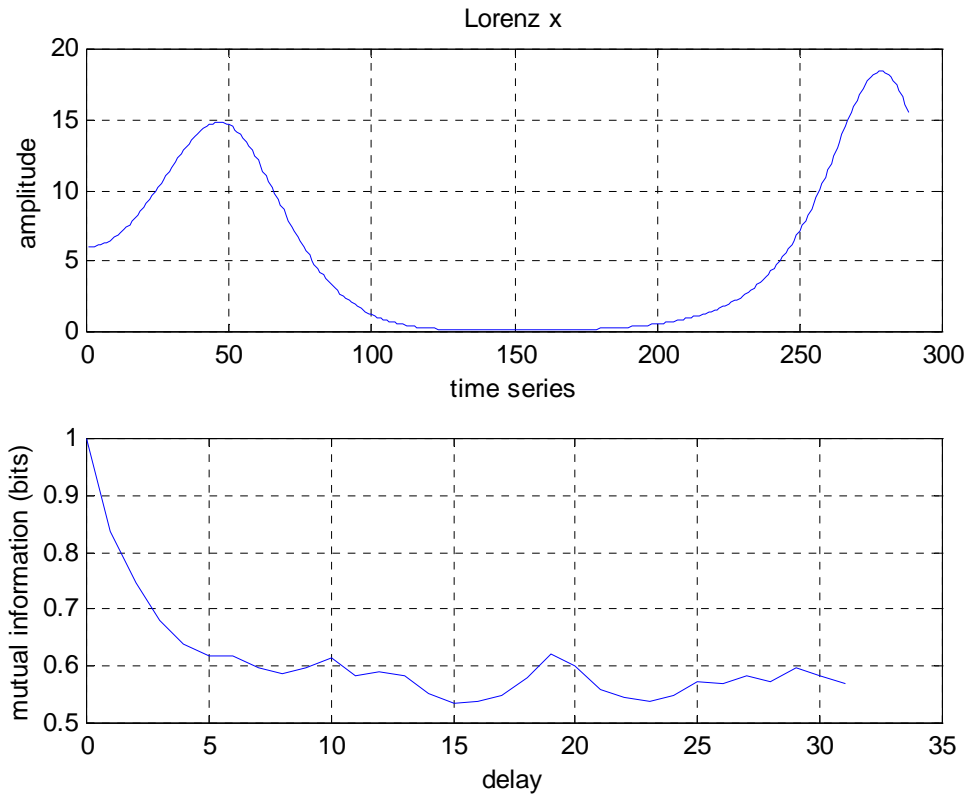
很顯然地離散時間分數維度高斯雜訊的共有訊息約為 0.2 (bits)，符合理論預測的狀況，訊號的隨機性愈高，共有訊息的值愈小，然而它的條件熵大致落在 0.42 (bits)到 0.54 (bits) 間。另外，當訊號來自決定性的系統時，例如弦波函數和勞倫茲方程式，則它們的共有訊息和條件熵如下圖：



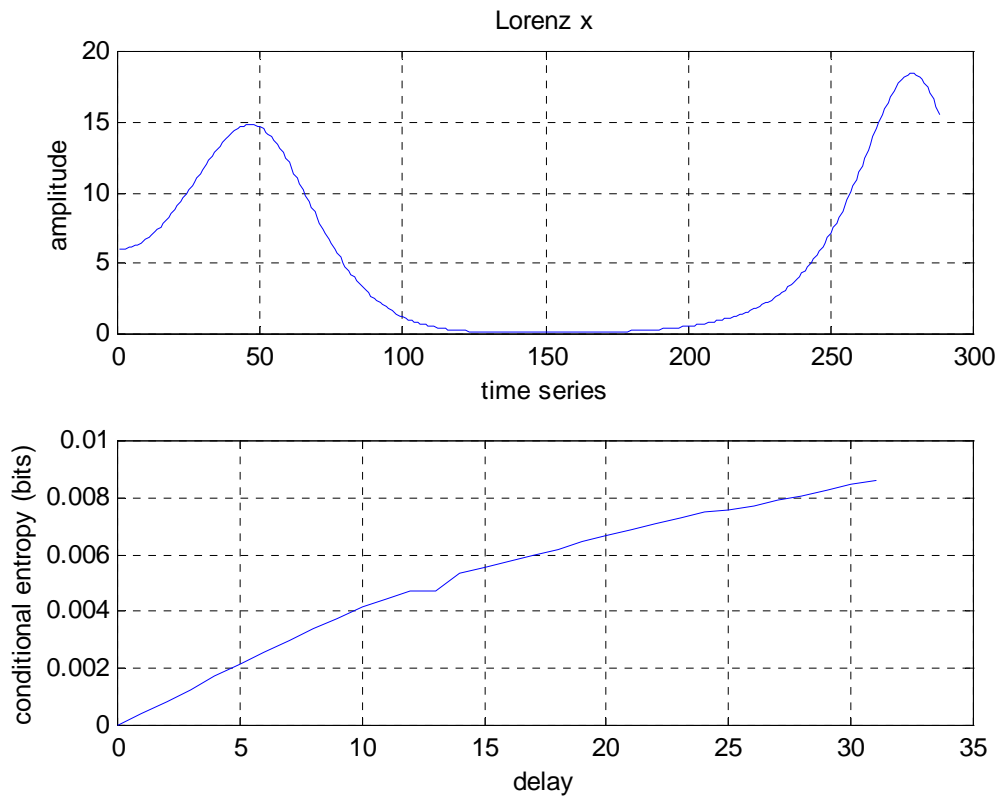
圖表 3: 上圖是弦波訊號；下圖是弦波的共有訊息。



圖表 4: 上圖是弦波訊號；下圖是弦波的條件熵。



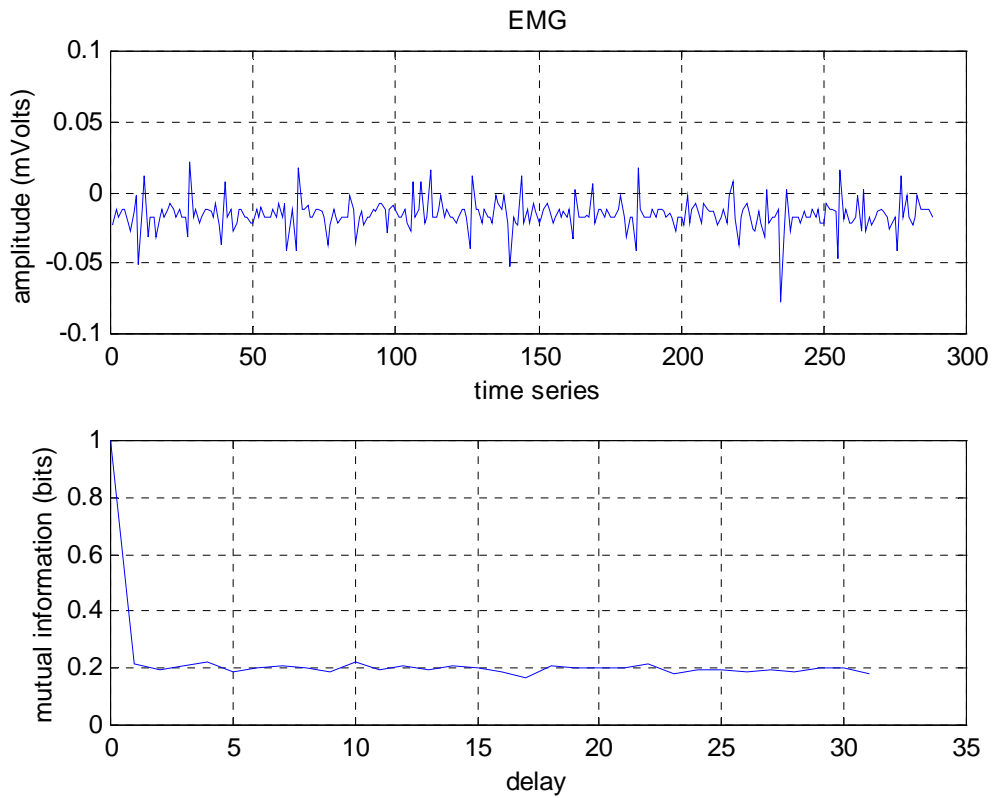
圖表 5: 上圖是勞倫茲方程式的訊號；下圖是勞倫茲方程式的共有訊息。



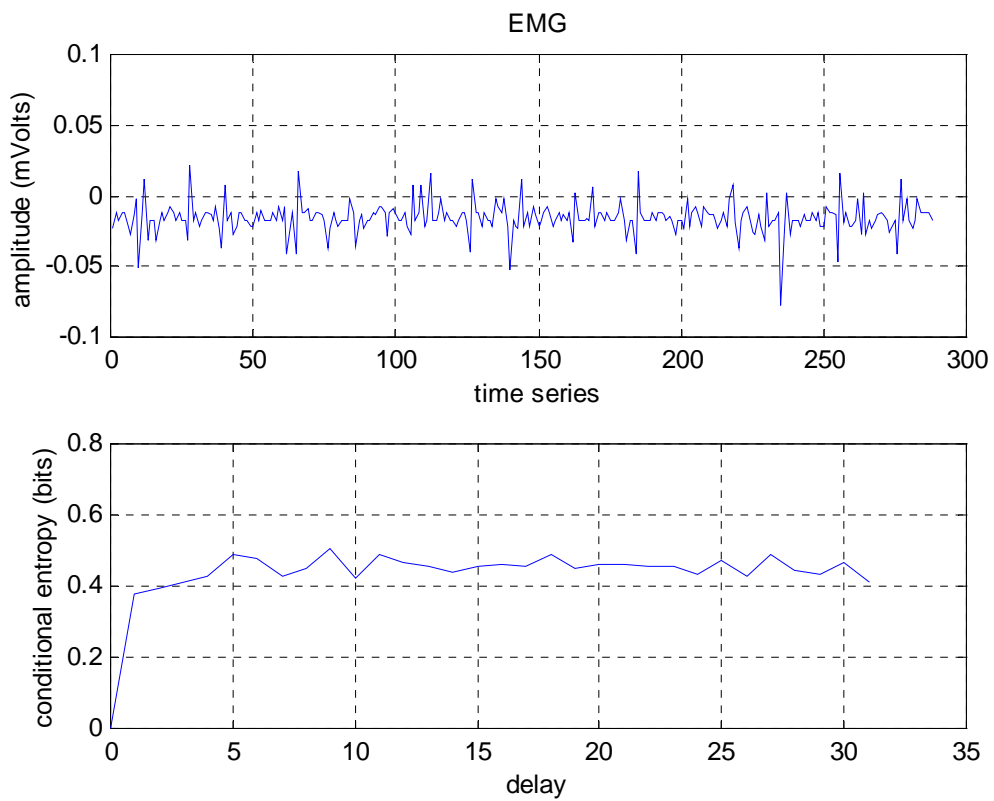
圖表 6: 上圖是勞倫茲方程式的訊號；下圖是勞倫茲方程式的條件熵。

由圖 3-6 可知，決定性的訊號的共有訊息約在 0.5 (bits)至 0.6 (bits)間，而條件熵小於 0.025 (bits)。因此我們可以利用計算共有訊息和條件熵的方式判別訊號到底來自隨機性的還

是決定性的。最後將共有訊息和條件熵的計算應用在成鼠的尿道外括約肌肌電圖的訊號上，其結果如下：



圖表 7: 上圖是 EMG 的訊號；下圖是 EMG 的共有訊息。



圖表 8: 上圖是 EMG 的訊號；下圖是 EMG 的條件熵。

從圖 7-8 中發現此訊號很接近隨機性的訊號，它的共有訊息很接近 0.2 (bits)，然而條件熵大約在 0.4 (bits) 以上，從此結果可以推斷它適合利用估測赫斯特參數來進一步判斷訊號間的差異。

六、參考文獻

1. B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Rev.*, vol. 10, pp. 422-437, Oct. 1968.
2. T. Lundahl, W. J. Ohley, S. M. Kay, and R. Siffert, "Fractional Brownian motion: A maximum likelihood estimator and its application to image texture," *IEEE Trans. Med. Imag.*, vol. MI-5, no. 3, pp. 152-161, Sep. 1986.
3. S. C. Liu and S. Chang, "Dimension estimation of discrete-time fractional Brownian motion with applications to image texture classification," *IEEE Trans. Image Processing*, vol. 6, no. 8, pp. 1176-1184, Aug. 1997.
4. Chang, Y. C., and Chang, S. "A fast estimation algorithm on the Hurst parameter of discrete-time fractional Brownian motion," *IEEE Trans. Signal Processing*, vol. 50, no. 3, pp. 554-559, March 2002.
5. 離散時間分數維度高斯雜訊之自回歸模式的階數決定(NSC91-2213-E-235-001)
6. M. Lehrman, A. B. Rechester, and R. B. White, "Symbolic analysis of chaotic signals and turbulent fluctuations," *Phys. Rev. Lett.*, vol. 78, no. 1, pp. 54-57, Jan. 1997.
7. A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A*, vol. 33, no. 2, pp. 1134-1140, Feb. 1986.
8. S. Chang, S. T. Mao, T. P. Kuo, S. J. Hu, W. C. Lin, and C. L. Cheng, "Fractal Geometry in Urodynamics of lower urinary tract," *The Chinese Journal of Physiology*, vol. 41, no. 4, pp. 25-31, 1999.
9. S. Chang, S. T. Mao, S. J. Hu, W. C. Lin, and C. L. Cheng, "Studies of detrusor-sphincter synergia and dyssynergia during micturition in rats via fractional Brownian motion," *IEEE Trans. Biomed. Eng.* vol. 47, no. 8, pp. 1066-1073, Aug. 2000.
10. S. Chang, S. J. Hu, and W. C. Lin, "Fractal dynamics and synchronization of rhythms in urodynamics of female Wistar rats," *Journal of Neuroscience Methods*, vol. 139, Issue: 2, pp. 271-279, 2004.

七、計劃成果自評

本計劃預估在執行上可能面臨到的三個問題都順利克服並完成，它們分別是：

1. 程式的撰寫，特別是共有訊息部分，因為它牽涉到許多複雜的程序。
解決方法：由於對訊息理論(information theory)的了解和長期的程式撰寫經驗，相信可以克服此一困難。
2. 所得到的結果沒有兩篇論文所宣稱的明顯效果。
解決方法：不論從直覺上或者理論上，我們可以預測結果的可行性，相信經過不斷的改良應該可以解決此一困難。
3. 在應用上沒有得到顯著的效果。
解決方法：再尋找其它配合的方法以突破這個僵局。

另外預期建立幾組決定性訊號的共有訊息和條件熵也已完成，例如弦波函數和勞倫茲方程式，接著也建立了離散時間分數維度高斯雜訊的共有訊息和條件熵的參考值，同時找出隨機性訊號和決定性訊號的差異值，最後將此差異性應用在成鼠的尿道外括約肌肌電圖訊號的共有訊息和條件熵。

此計畫完成後，對生物醫學領域的應用有極大的幫助，因為在此領域中，經常需要了解訊號的本質是決定性的還是隨機性的。如果是決定性的訊號則可以採取對應的方法來分析；如果是隨機性的訊號則採用適當的方式來處理，例如類似離散時間分數維度高斯雜訊的訊號則可以應用赫斯特參數的估測法，並進一步釐清該訊號的意義，因此非常適合發表在學術期刊上。