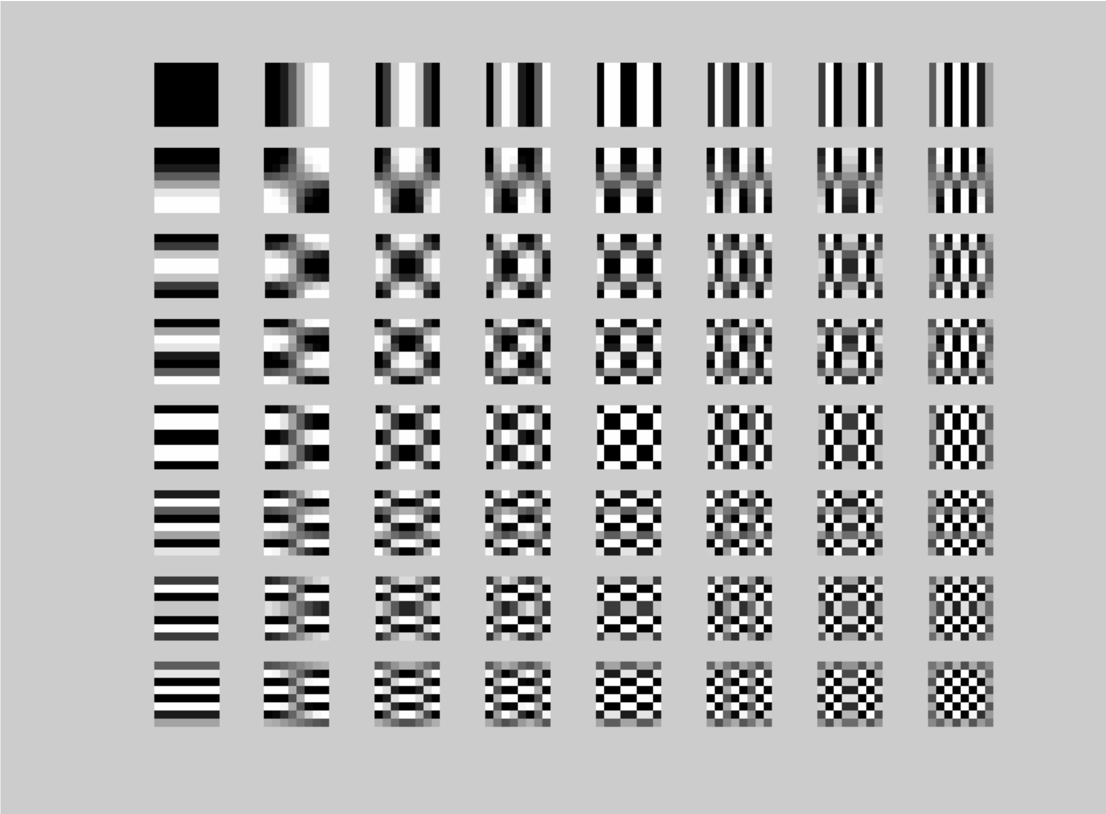
二維影像離散餘弦轉換原理（DCT, Discrete Cosine Transform）
（including Taylor Series, Fourier Series）

南台電機 趙春棠

```
clear
clc
pos=0;
a = zeros(8)
for i=1:8
   for j=1:8
      pos=pos+1;
      a(i,j)=-1024;
      id = idct2(a);               %「頻域頻譜係數」作「反轉換」
      ok = id +128*ones(8);
      subplot(8,8,pos);
%        imshow(ok/255)
      imshow(uint8(ok))
      a = zeros(8);
   end
end
```



說明：以上有 8*8=64 個方塊圖，其中每一個方塊圖，皆是由 8*8=64 個像素（pixel）所組成。

學理上告訴我們，任何一個 8*8=64 個像素所組成的圖案，都可以表示爲以上 64 個方塊圖的「線性組合」。

離散餘弦轉換是將「時域空間」轉換至「頻域」的一種轉換，以上 64 個方塊圖，其實就是每一個「頻域頻譜係數」作「反轉換」的結果。

您對以上的說明，感到懷疑嗎？以下再舉其他幾個有趣的現象。

■ Fourier Series

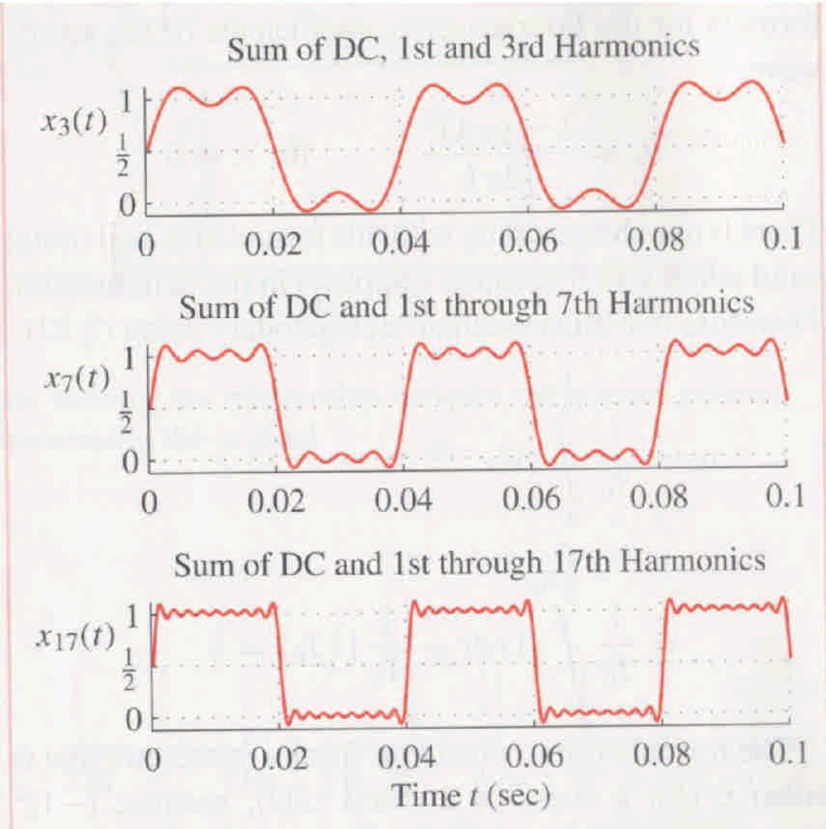
法國 傅立葉先生提出：幾乎所有的週期訊號，皆可以表示為「不同頻率的弦波訊號」的組合。

$$x(t)=a_0+\sum_{k=1}^{\infty}A_k\cos(k\omega_0 t+\phi_k)$$

$$=a_0+\sum_{k=1}^{\infty}\left(\frac{1}{2}A_k e^{j\phi_k}e^{jk\omega_0 t}+\frac{1}{2}A_k e^{-j\phi_k}e^{-jk\omega_0 t}\right) \quad （式 1）$$

$$x(t)=\sum_{k=-\infty}^{\infty}a_k e^{jk\omega_0 t} \quad ,\omega_0=\frac{2\pi}{T_0}$$

例如 1kHz 的方波，可以用 1kHz（基頻，即該週期訊號的頻率）、2kHz（二次諧波）、3kHz（三次諧波）、4kHz（四次諧波） …的弦波「線性組合」而成！這每個有倍頻關係的弦波，彼此之間是「線性獨立」的，構成空間的「基底」（Basis）。圖示如下：



【摘自 "Signal Processing First", McClellan, Schafer, and Yoder, Pearson Prentice Hall, 2003. 】

由上圖我們可以發現，若高次諧波取得越多，圖形就會更接近「方波」了。至於每個弦波的分量大小（相當於以上 a0, a1, a2 …係數），該如何求得？在學理上，是將此週期訊號與每個弦波分量去做「內積」（函數內積），就可以求得，公式如下：

$$a_k = \frac{1}{T_0}\int_0^{T_0} x(t)e^{-jkw_0}dt$$

$$\text{其中 } a_0 = \frac{1}{T_0}\int_0^{T_0} x(t)dt \quad \text{爲} DC\ Part,\ or \text{ 平均值}$$ （式 2）

簡單來說，就好比在三度空間中， (2,3,5) 這個座標，可以看作以下三個基底座標的線性組合：

(2,3,5)=2*(1,0,0)+3*(0,1,0)+5*(0,0,1)

而我們可以看到，(2,3,5) 在 (0,1,0) 的分量大小爲 3 ，而這個 3 ，正是 (2,3,5) 與 (0,1,0) 取向量內積（有取「投影」量的意涵）的結果：

(2,3,5) • (0,1,0) = 2*0+3*1+5*0 = 3

「向量內積」的運算如上，簡單得多，而「函數內積」比較複雜，是將兩個函數彼此相乘取積分（如（式 2））。

■ Taylor Series

泰勒展開式，意思是說許多函數，可以用冪級數（Power Series）展開而得。

您可以想像，像 e^x 、sinx，這樣的函數，可以經由泰勒展開，而用「多項式函數」來表示嗎？如下所示。

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

原理上，泰勒展開是在函數 f(x) 某一點 a，進行函數其他點的「預測」，利用該點「高階的微分」，就可以進行很精密的預測。公式如下：

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots$$

所以說前述 e^x 、sinx 的泰勒展開，就是在 a=0 處展開的結果，讀者也可以試試在其他點作展開。學理上告訴我們，只要 a 點是「可解析的」(Analytic)，那麼泰勒展開就可以存在。但是，泰勒展開式的成立，是有範圍限制的，只有當 $|x-a| < R$ 時才成立（其中 R 爲收斂半徑（Radius of Convergence））。上述 e^x 、sinx 的泰勒展開式中，R→∞。

最後不妨讓我們回憶過去中學所學的一次直線，以及二次曲線。如果是一條線，我們只要知道一階微分（即斜率），就可以預測其他點的函數值；如果是二次曲線，則除了一階微分外，只要再了解二階微分，就可以預測所有其他點的函數值了。

y=ax+b 在 x=0 的泰勒展開式爲

$$y=b+a(x-0)=b+ax$$

而 $y=ax^2+bx+c$ 在 $x=\frac{b}{-2a}$ 的泰勒展開式爲

$$y=\frac{b^2-4ac}{-4a}+(x+\frac{b}{2a})^2$$

看到以上的結果，不知是否讓您有印象？想起二次曲線的極大、極小値？

^_^