

第四章

數位影像技術概述

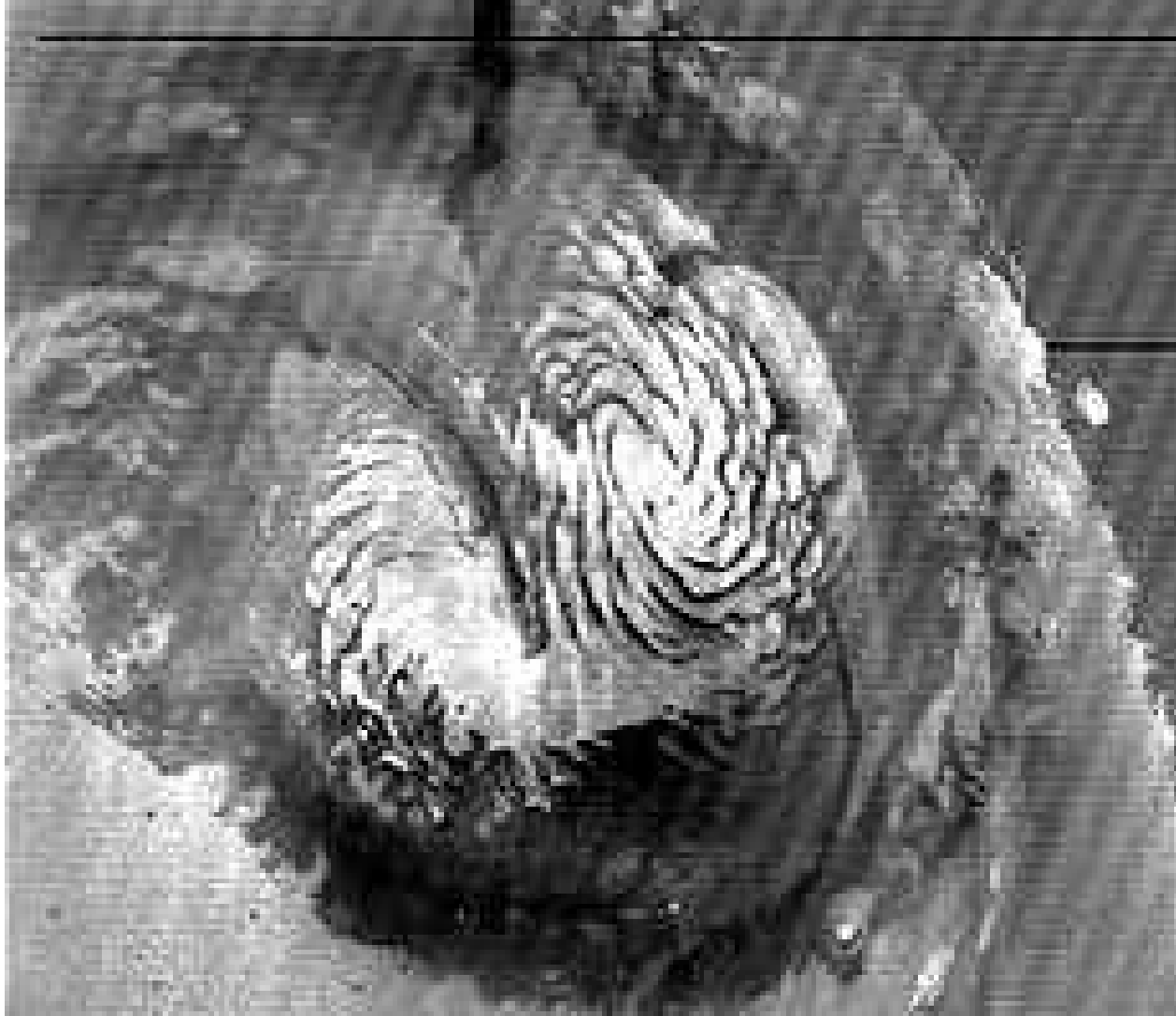
課程內容

- 數位影像表示法
 - 數位影像之形成
 - 取樣和量化
 - 專有名詞
- 影像壓縮與影像安全之目的
 - 影像壓縮
 - 數位影像浮水印
- 影像品質評估法

4.1 數位影像表示法

- 所謂影像表示模式，就是如何看此影像，從影像模式中可以分析此影像的特質，進而找出處理的方法。
- 一般影像可以空間領域和頻率領域表示，另外統計模式常用來分析此影像的特質。

- 影像空間表示直接代表此影像的特性，它是一個二維函數 $f(x,y)$ ，每一個 (x,y) 的灰階值 g 代表此影像在這點的物理特性，此物理特性決定於輸入感應器的特性。
- 每一點的特性可用此點和周圍鄰近的灰階值表示，此種表示方稱為空間領域表示。根據空間領域表示所衍伸出的處理方法稱為空間領域處理。



1972美國太空總署所發射的水手九號探測船所拍攝的火星北極冰帽(The North polar cap of Mars)的圖片

- 頻域表示將原信號看成是由不同正交信號所組成，如向量空間中的基底一樣。
- 常見的表示方法為 Fourier 表示，Fourier Transform 將原影像轉換成頻率成分，每一點代表一種基本頻率，其大小代表此頻率成分的大小。
- 一些影像頻率變化的特徵可以由此影像的頻率表示分析而得到，根據頻域表示的處理方式稱為頻域處理。

- 圖4-1 顯示一個 $N \times N$ 影像表示成 $N \times N$ 個基底頻率影像的線性組合， $B_{n,m}$ 代表不同頻率基底影像，其中 $B_{0,0}$ 代表直流影像，當 m, n 值越接近 $N/2$ 時期頻率越高，空間灰度變化越頻繁。

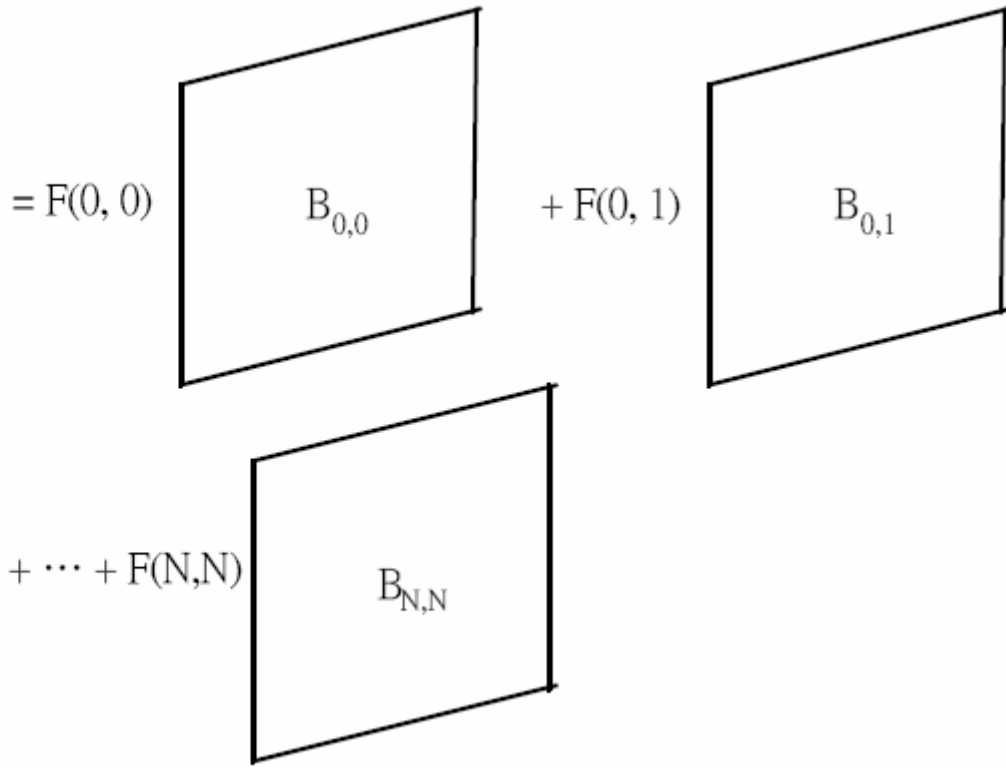
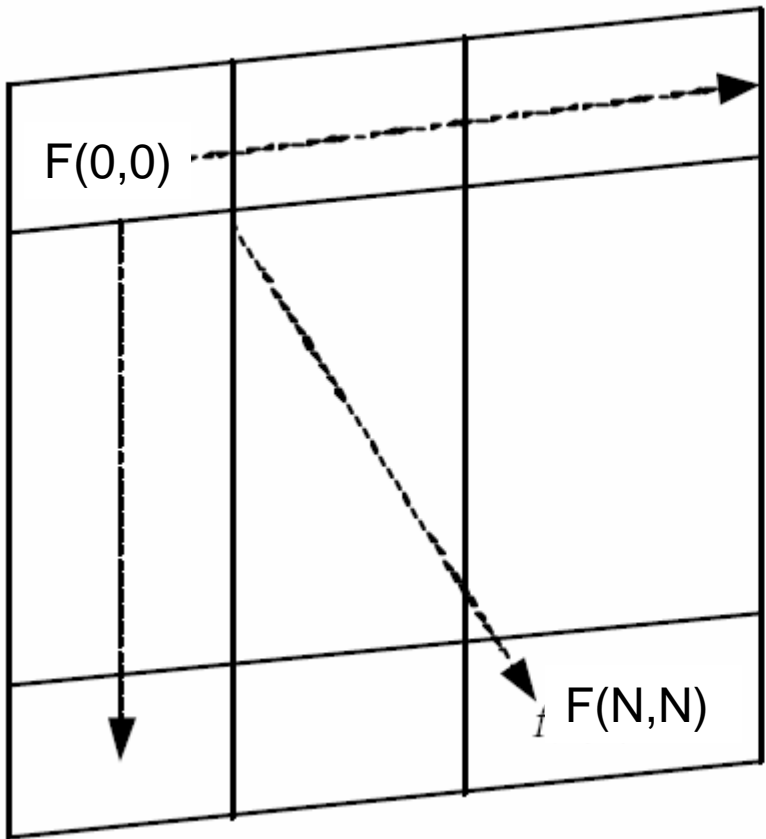


圖4-1、影像頻域表示

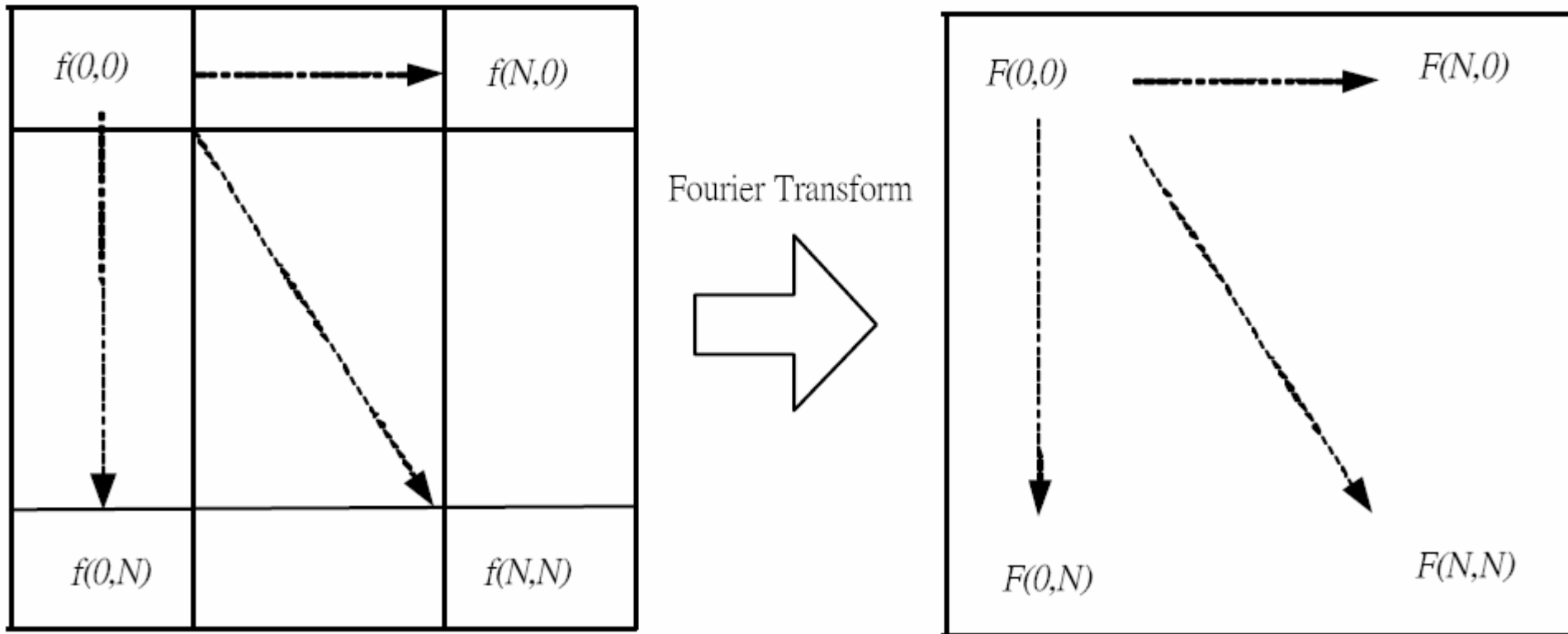


圖4-2、影像頻域表示

- 數位影像依其儲存方式可以分成兩大類：
 - 一、**向量影像(vector-based image)**：

影像中的圖案是由許多個物件所組成，每個物件都可以由數學方程式來表達、紀錄及處理，因此又稱為物件導向式(Object-Oriented)的繪圖格式。目前市面上常見屬於這類型的商業軟體有CorelDraw、AutoCad、Illustrator、MS-Visio 和3D Studio 等。

優點：

- 儲存空間較小；影像放大、縮小等轉換過程中，品質較不受影響；每個物件都可以隨意移動、修改等。

缺點：

- 是對於一般的自然影像(nature image)較難加以描述，並較難正確地取出物件。

二、點陣式影像(bit-mapped image)：

- 數位影像由像素排列組成，以彩色影像為例，每一個像素：計有X座標、Y座標、紅顏色(R)、綠顏色(G)和藍顏色(B)。
- 每一種顏色以八位元來表示，也就是每種顏色有256種色階，因此組成每一個像素有 $256 \times 256 \times 256 = 16,777,216$ (24位元)種顏色。
- 為了節省儲存空間，一般都會透過調色盤只使用其中的256種顏色而已。

- 屬於這類型的軟體，較常見的有 Photoshop、Painter、PhotoImpact、LviewPro、Paintshop Pro 等。
- 優點：色彩自然且富變化、逼真、對影像顏色之處理比較容易，如使用掃描器、數位相機等，就直接取得數位影像資料。
- 缺點：影像放大過程中產生失真、記憶體空間需求較大以及無法呈現3D 立體影像等。

4.1-1 數位影像之形成

- 黑白照片是由黑到白的連續淡濃所構成，由不可分割的明暗帶組成的。但數位影像處理卻只能對數位化，亦即是非連續性資料作運算及處理。
- 轉換連續性資料成為數位化訊號，是經由切取(chopping)來完成的，亦即所謂的數位化或稱取樣。如圖4-3所示。

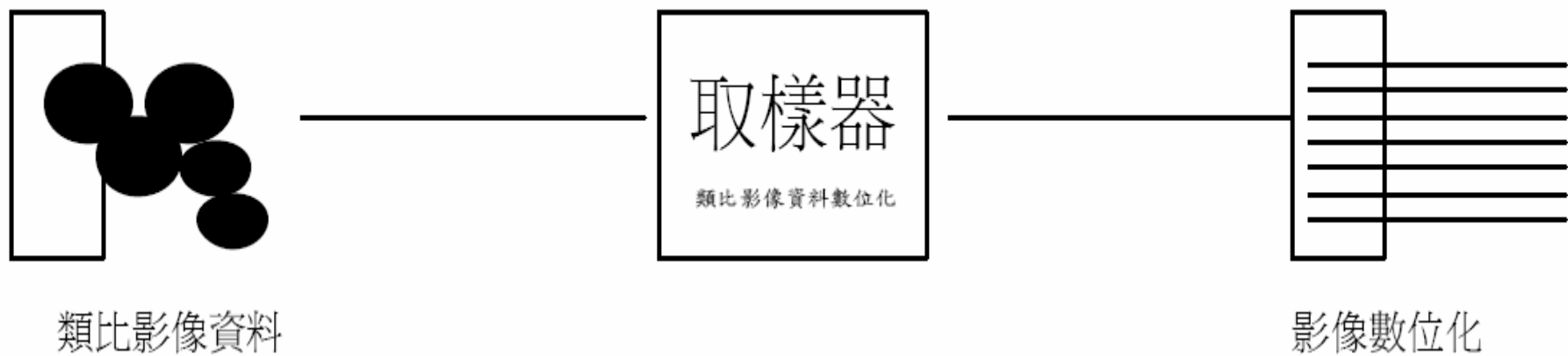


圖4-3、類比影像資料數位化

像素(pixel)

- 對每一個攝取到的資料，賦予一個數值代表它的強度及其對應位置，此一最小單位為影像元素、圖像元素或像素(pixel)。

座標

- 一張影像劃成方塊格子所組成的像素，每一個格子中都標有一對座標，X軸表行值(Cols)，Y軸表列值(Row)。

- 行值從這個影像的最左邊開始標誌從 0 到 N（N 表行值中的最大值）。同樣的，列值從最上方起，定為 0 往下移至 M 值，M 表示影像的全部列數。
- 例如：某一個像素它的座標式(200, 150)，表示其位置於第200 行和第150 列的交接處，可參考圖4-4 所示。

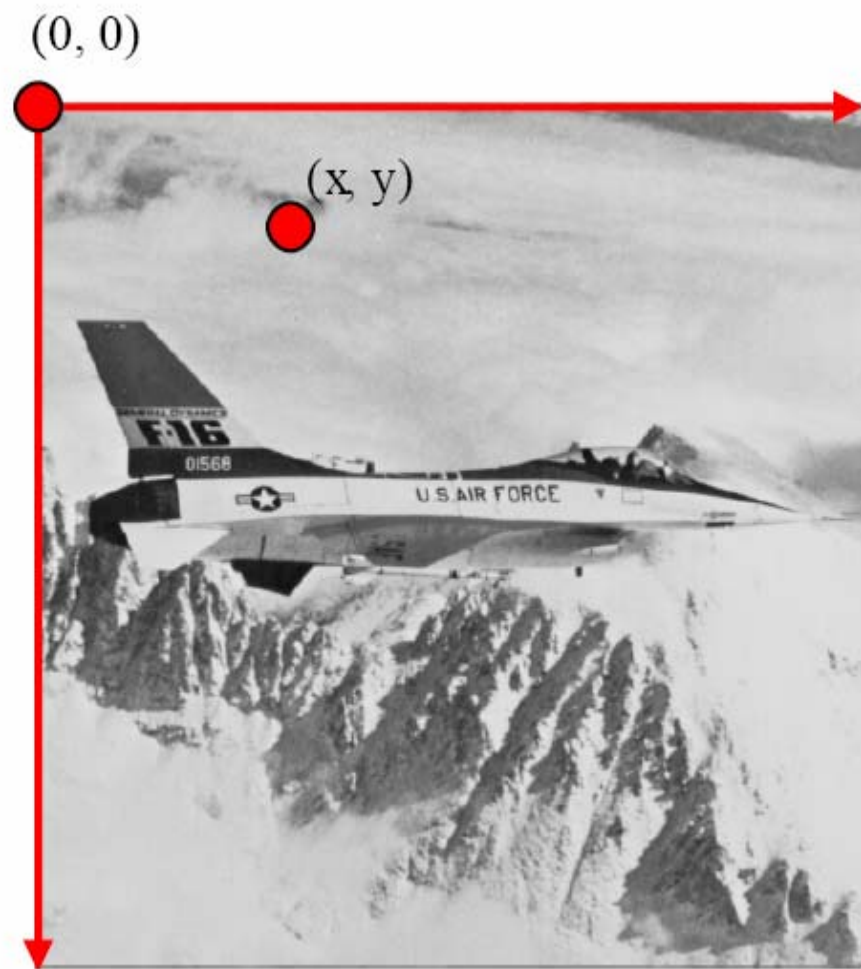
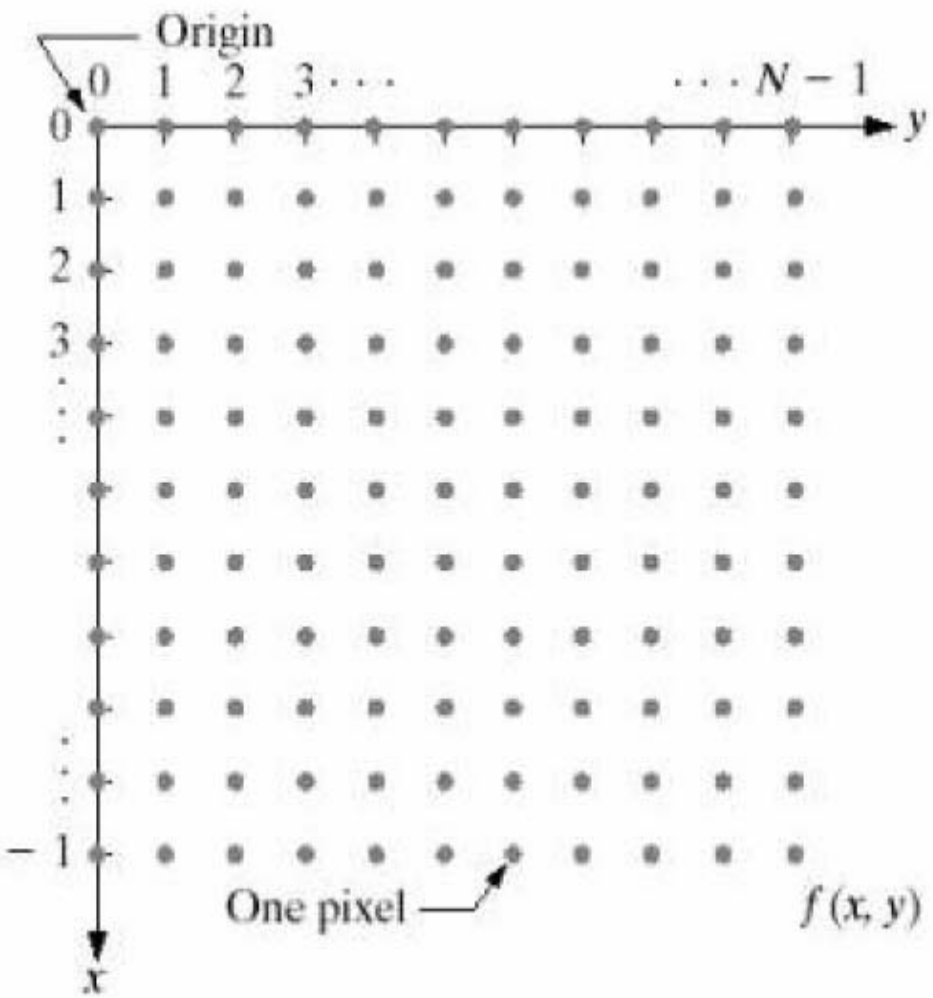


圖4-4、離散像素座標定義

解析度(相似度)：

- 處理影像的數位化，會產生一個問題，就是此一代表影像和原始影像比較，會達到何種的相似度(解析度)。
- 影像處理中一般解析度可以分為兩種形式，一種是空間的解析度，另一種則是亮度的解析度。

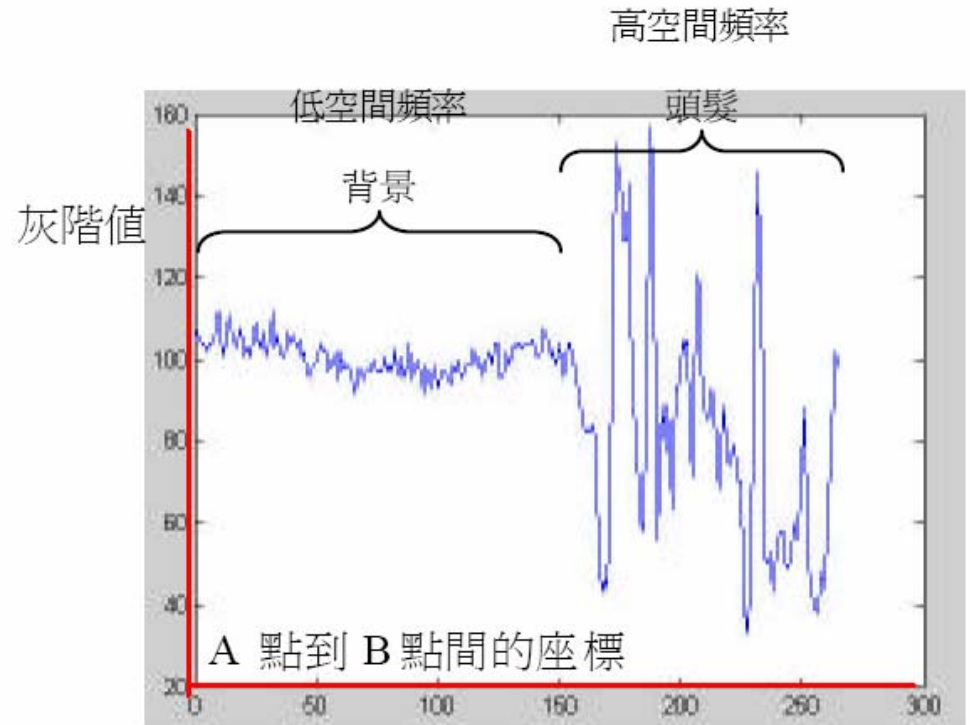
一、空間解析度

- 空間於此是指所謂的二維空間(2D space)。空間解析度是描述一張圖片可劃分為多少點，簡言之，解析度劃分的越細，可以越接近原始影像。
- 最佳的狀況：經過數位化後，沒有任何資訊遺失於轉換過程中，換言之，它和原始的影像是一致的。

- 空間頻率(spatial frequency):影像的亮度從暗轉為亮，或亮轉為暗的變化頻率。
- 例如:頭髮的精細度到人的輪廓陰影，以數學方式量化影像的明細變化。



(a)



(b)

圖4-5 (a)空間頻率的變化 (b) 亮度變化

- 在圖4-5 中，所呈現的例子說明空間頻率的內容。當我們觀察單一線條，它橫跨過此影像之時，這個影像的明暗可強調出此不規則而醒目的線條。
- 在頭髮部分有相當高的空間頻率，而背景部分則相當的平坦。即是在頭髮與背景部分分別具有高低不同的空間頻率。

- 對影像而言，空間解析度即是對一影像所能分割到最小元素的程度。
- 換言之，解析度越高，和原始影像相似度越高，亦更能鑑別出最細緻的濃淡低空間頻率背景高空間頻率頭髮灰階值A點到B點間的座標AB變化。

空間解析度：

- 即影像中明暗度的變化頻率，明暗度變化頻率計算是由左至右及由上至下。

抽樣頻率：

- 為計算此明暗度的變化頻率的抽樣頻率時，必須使用 Nyquist Criterion(亦稱為抽樣理論)。此取樣頻率必須比待測訊號的頻率至少快上兩倍以上，才有意義的資料。

Aliasing :

- Aliasing現象是原始影像的高頻部分發生錯誤。Aliasing 發生在Nyquist Criterion的影像原則被破壞時才會產生。
- 當空間頻率比取樣頻率的二分之一還要大時才會發生，這稱之為低取樣。這種高頻被轉換為較低頻率，許多的資料會在取樣過程中被遺失。

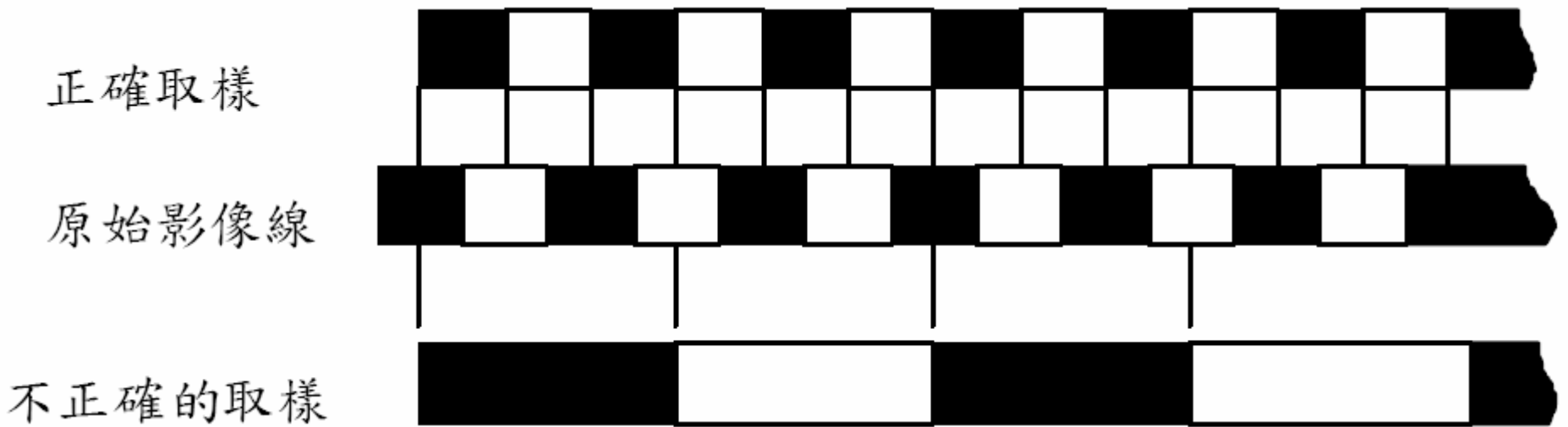


圖4-6、Aliasing 現象

二、亮度解析度(灰度解析度)

- 影像的第二種解析度，就是所謂的亮度。亮度解析度的觀念即是討論像素對於原影像在同一位置亮度的正確性。
- 影像數位化過程中，事先預定的方格位置中取樣，隨後此取樣值就被轉換成一個相等的整數值，即是所謂的量化(Quantization)。

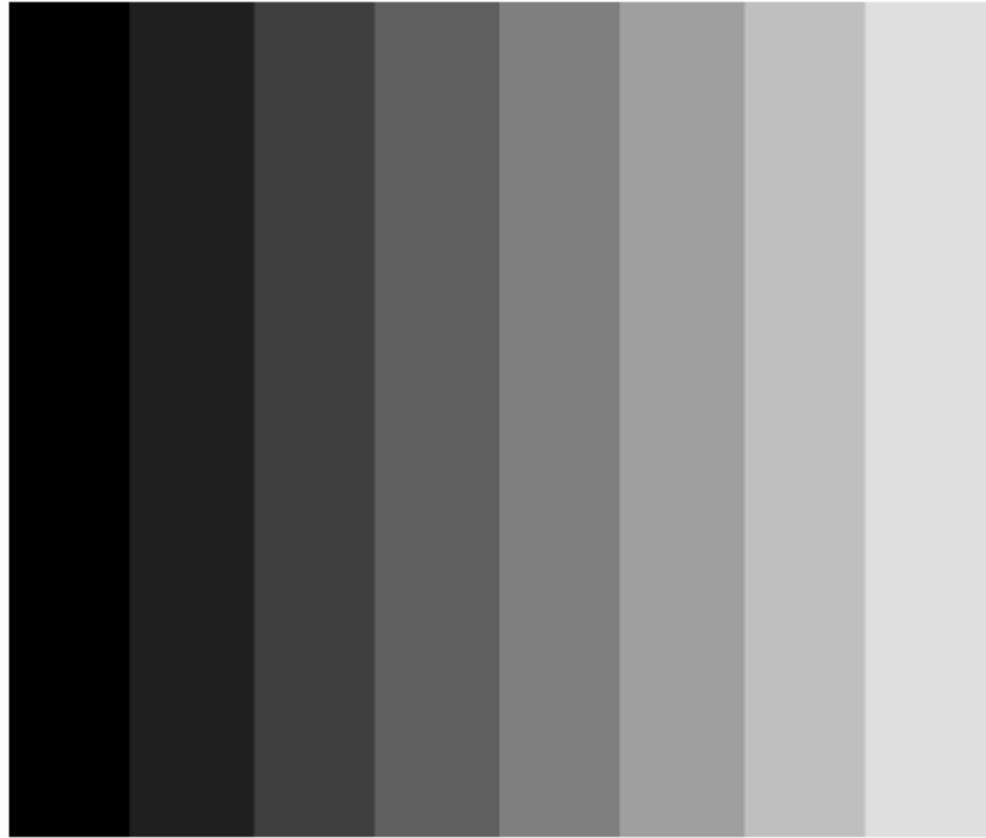


圖4-7、三位元亮度範圍(八個階層:000
、001、010、011、100、101、110 和
111)



圖4-8、兩位元影像

- 參考圖4-8 的兩位元影像，這種現象的效應稱之為假輪廓(contouring)。
- 輪廓線的發生是由於量化位元的限制。在圖4-8中，所有像素僅可使用這四種亮度中的一個，因此原始影像即受到限制，所以，在數位化的影像中會產生突然間改變其灰暗度的現象。
- 選擇合適的亮度解析度，其輪廓線的效應將會被減少。

灰暗度之設定：

- 如何設定灰暗度位元數量呢？

- 眼睛的反應：

- 黑暗地區較敏感，亮度地方較遲鈍。輪廓線的現象通常發生於黑暗區域。

- 好的數量化亮度，即是採用對數的形式，而不採用線性形式。

- 黑暗區域(較低亮度)的範圍內，用較多的灰暗度來表示，其他較明亮的地方，則用較少的灰暗度來表達。
- 此種對數形式的灰暗度，於低亮度的部分有更精細的解析度配合眼睛的敏感度。



4.1-2 取樣和量化

- 影像函數 $f(x,y)$ 在空間和振幅上都需要進行數位化。
- 空間座標 (x,y) 的數位化稱為影像抽樣 (image sampling)，振幅的數位化則稱為灰階量化 (gray-level quantization)。

- 設一幅連續影像 $f(x,y)$ 可用等間隔取樣並被排成 $N \times M$ 的矩陣形式來近似，整個矩陣中的每個元素稱之為像素。

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, M-1) \\ f(1,0) & f(1,1) & \cdots & f(1, M-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(N-1,0) & f(N-1,1) & \cdots & f(N-1, M-1) \end{bmatrix}$$

- 數位化過程需要確定 N 與 M 的值和每個像素所允許的離散灰階值 G 。在數位影像處理中，通常取這些量為 2 的整數冪次，即

$$N = 2^n, \quad M = 2^k, \quad G = 2^m$$

- 儲存數位影像所需的位元數 b

$$b = M \times N \times m$$

- 為得到良好的近似，需要多少個取樣和多少階灰度。影像解析度(resolution)是指能辨別的細節程度，它主要取決於 N 、 M 和 m 這三個參數。它們越大，數位化矩陣越接近原來影像，數位影像所需的位元數 b 也相對大幅的增加。

- 圖4-9(a)中顯示了一幅玫瑰花 1024×1024 、256 灰階的數位影像。
- 圖4-9(b)~(f)分別是減少空間解析度 ($N=1024, 512, 256, 128, 64, 32$)後，所得到的結果，最大灰階數均為256。
- 低解析度影像中的像素要重複複製才能填滿整個顯示區域，這種像素複製(pixel replication)會產生馬賽克效果，在低解析度影像中尤其可見。

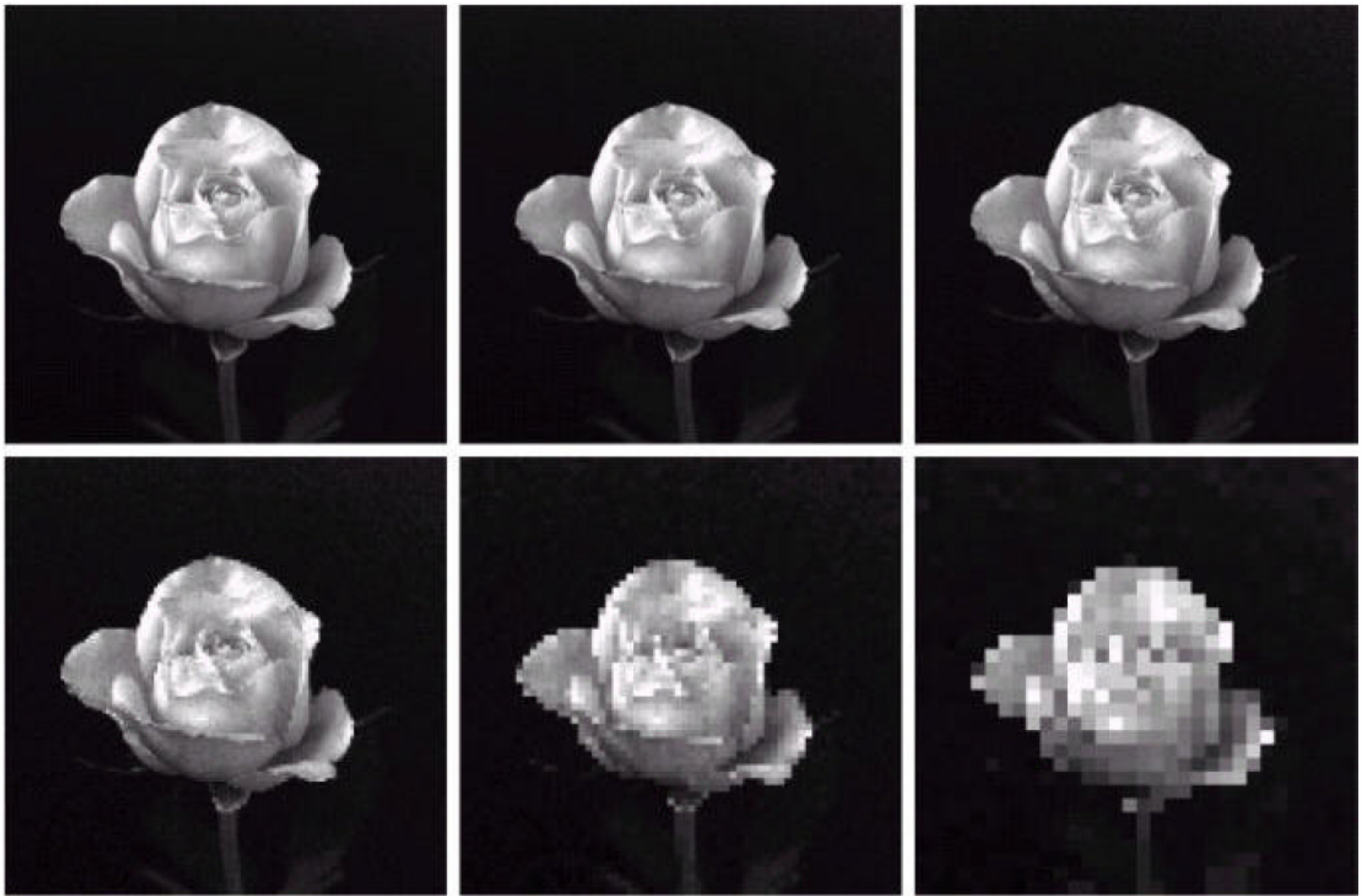


圖4-9、減少空間解析度的結果





4:1 取樣
42

- 圖4-10(a)為 1024×1024 的8位元影像。圖4-10(b)~(h)是保持空間解析度為 1024×1024 ，但量度位元數減少(m 從7到2)，其中256, 128, 64灰階的影像實際視覺效果是相同的。
- 圖4-10(d)所示的32灰階影像在平滑的灰度區域中產生一系列幾乎不能察覺的很細小的狹長隆起結構，這個效果是由數位影像中平滑灰階區域上的灰階數不足所引起，即為之前所述的假輪廓(false contouring)，它在16或更少的均勻間隔的灰階影像顯示中很常見。



圖4-10、分別用256、128、64、32、16、8
個灰階顯示的1024x1024 影像

- 圖4-11 顯示了其中的三張影像。Lena 面部表情的圖片代表具有相當少細節的影像，Camera-Man 的圖片包含了中等數量的細節，相比之下，crowd 影像卻包含了大量的細節。



圖4-11、評價主觀影像品質的測試影像

- 由圖4-12的結果可以得到幾個經驗結論：
 - (1)正如所預見：隨著 N 和 m 的增加，影像品質會提高。在個別情況下，對於固定的 N ，可以透過減少 m 使影像品質得到改善，產生這種結果最大可能的原因是減少 m ，一般會增加影像所表露出的對比度。

(2) 隨著影像中的細節增加，這些曲線變得更加陡直。這意味著對於具有大量細節的影像，只需要少量的灰階就可以表示。

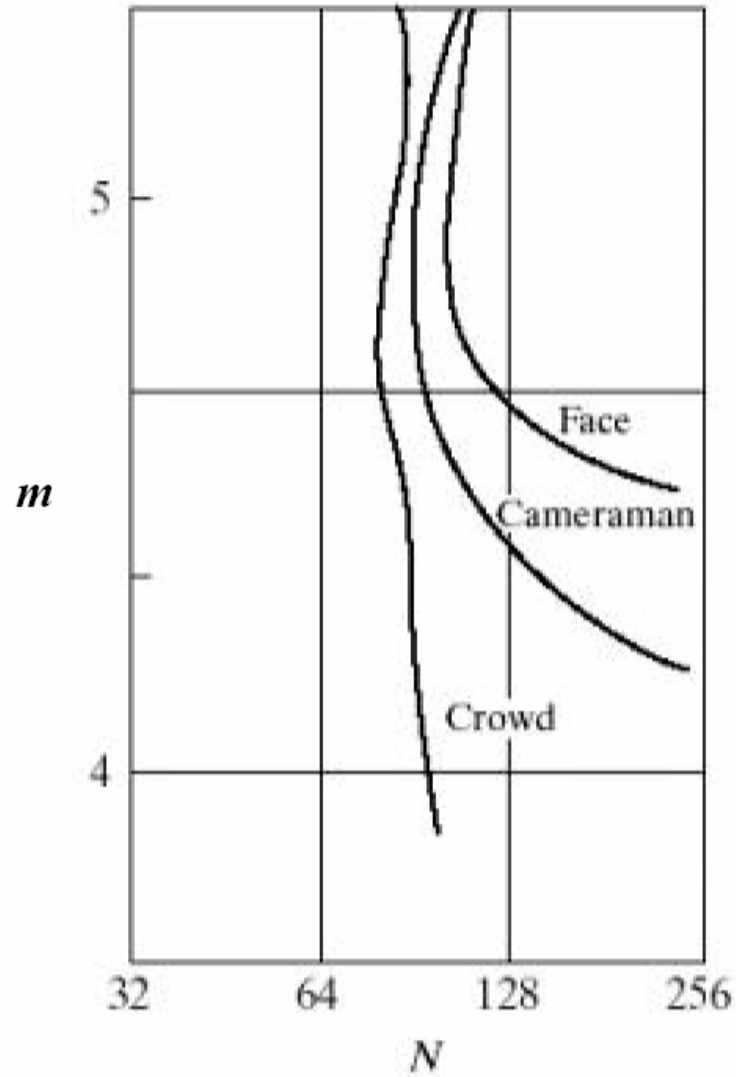


圖4-12、三幅影像的等優曲線

4.1-3 專有名詞

1. 影像大小(image size)

- 影像大小指的是影像中所含有的像素個數，通常以“寬×高”來表示影像大小，其中也說明了影像的幾何尺寸。
- 例如：一張影像的大小為“640×480”，表示該影像的寬度有640個像素，高度有480個像素，總共包含了307,200個像素，座標原點位於影像的左上角，如圖4-4。

2. 檔案大小

- 檔案大小是指儲存一張影像總共所需要的記憶體空間，通常以KB(千位元組)或MB(百萬位元組)為計算單位。
- 例如: 1024×768 的全彩影像需要的記憶體空間為

$$\begin{aligned}1024 \times 768 \times (8+8+8) &= 18,874,368 \text{ bits} \\ &= 2,304 \text{ KB} \\ &= 2.304 \text{ MB}\end{aligned}$$

3. 影像模式









RGB	CMYK	Grayscale	Index Color
			
Duotone	Lab MODE	Bitmap	Multichannel
			

圖4-12、不同影像模式下的影像

(1)RGB

- 絕大部分的可見光譜(visible spectrum)都可以用紅、藍、綠三種顏色光依不同的比例來加以呈現，因此稱紅、藍、綠為三原色。
- 須注意的是，這裡的紅、綠、藍是指光的顏色而非染料的顏色；換句話說，當你將紅、藍、綠三種水彩加在一起是不會產生白色的。

- 因此RGB 通常適用於監視器(Monitor)、彩色投影機等具發光的設備上。

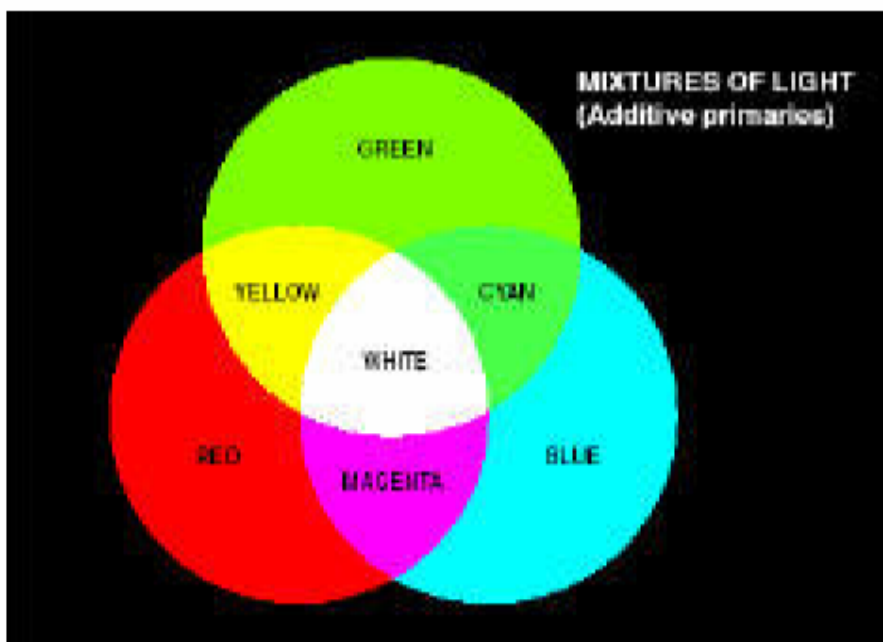
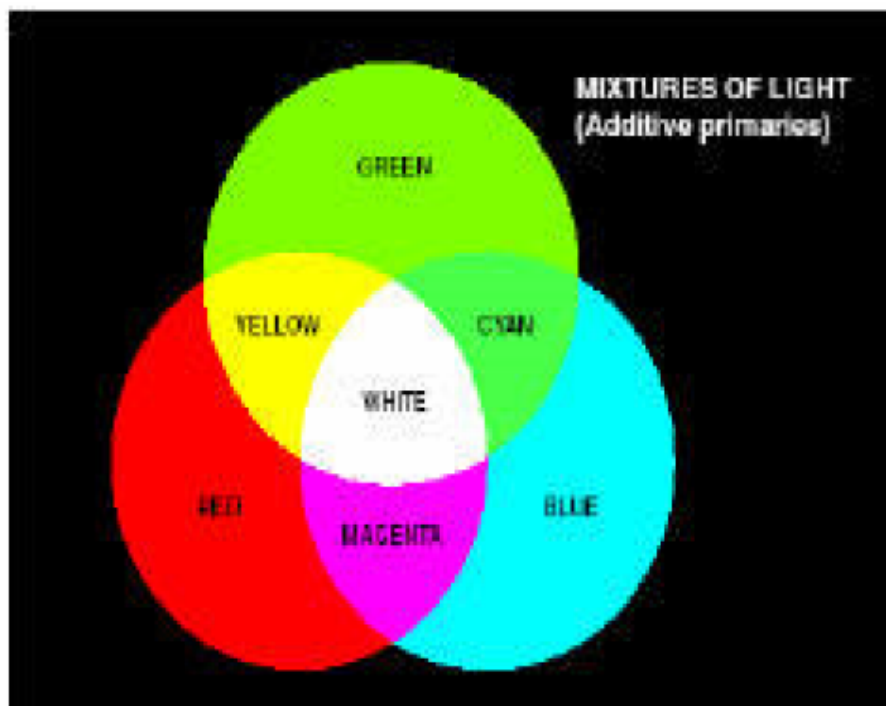


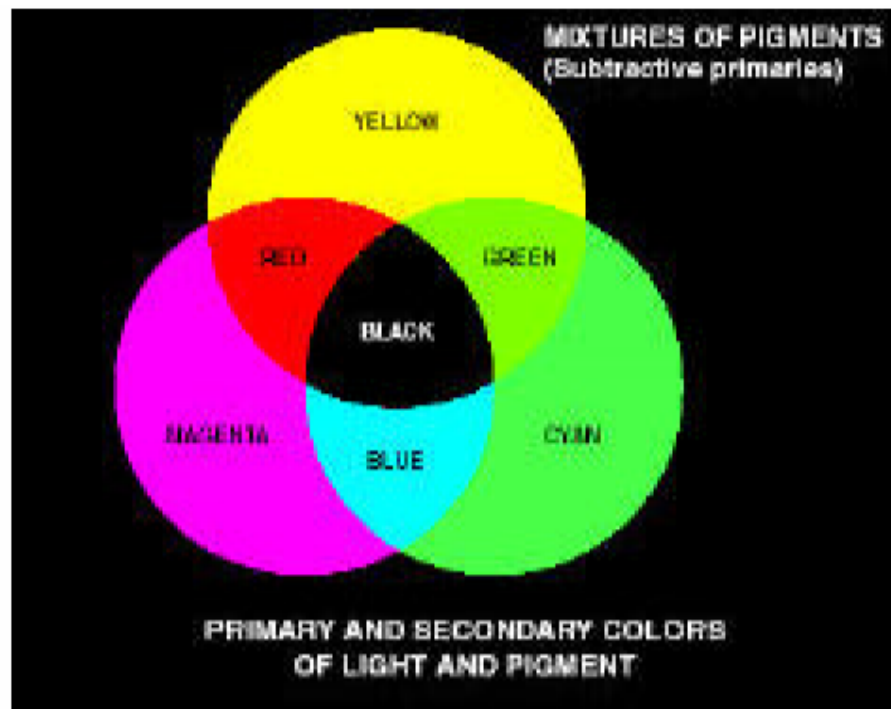
圖4-13、RGB 光三原色

(2)CMY & CMYK

- CMYK 指的是Cyan(C), Magenta(M), Yellow(Y)和Black(K)四種顏色而CMY則缺少Black 顏色。
- 與RGB 不同的是CMYK 主要是因顏料 (pigment)吸收光(light-absorbing)的質量不同而組成各種不同的顏色，因此CMYK 主要是用於印刷的顏色描述。



(a)



(b)

圖4-15、(a) CMY 影像模式 (b) CMYK 影像模式₅₆

- R、G、B 顏色值之值域常態化 (normalize) 成為介於0 與1 之間的實數，則RGB 與CMY 之間的轉換公式可以表示為

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

(3) Grayscale(灰度)

- Grayscale 又稱之為Graylevel，通常用256(八位元)個色階來表示該像素介於黑色(灰階等於0)與白色(灰階等於255)之間的程度。



圖4-16、Grayscale 影像模式

(4) Indexed color(索引色)

- Indexed Color 用來減少色彩數量，RGB 模式的像素值，共有16,777,216種顏色。
- 一般而言，256 種顏色對於一般常用的影像已經綽綽有餘了。Indexed color 所採用的方法就是建立一個具代表性的256色索引表格(Color LookUp Table, CLUT)，又稱之為調色盤(palette)或編碼簿(codebook)。

- Indexed color 中每一種顏色仍然包含RGB三種顏色值，每一個像素只用一個位元組(八位元)來指出其顏色是位於調色板中第幾個顏色，常見的影像格式GIF 即是採用Indexed Color影像模式。

(5)Bitmap(點陣)

- 當每一個像素只有黑色或白色兩種色階時，只需要用一個位元就可以表示一個像素值，此種影像模式稱之為Bitmap，而此種模式的影像稱之為二元影像(binary image)或黑白影像。
- 此種模式常使用於只包含純文字的影像，這是因為文字通常只有兩種顏色。



圖4-18、Bitmap 影像模式(黑白)



圖4-19、Bitmap 影像模式(變色調)

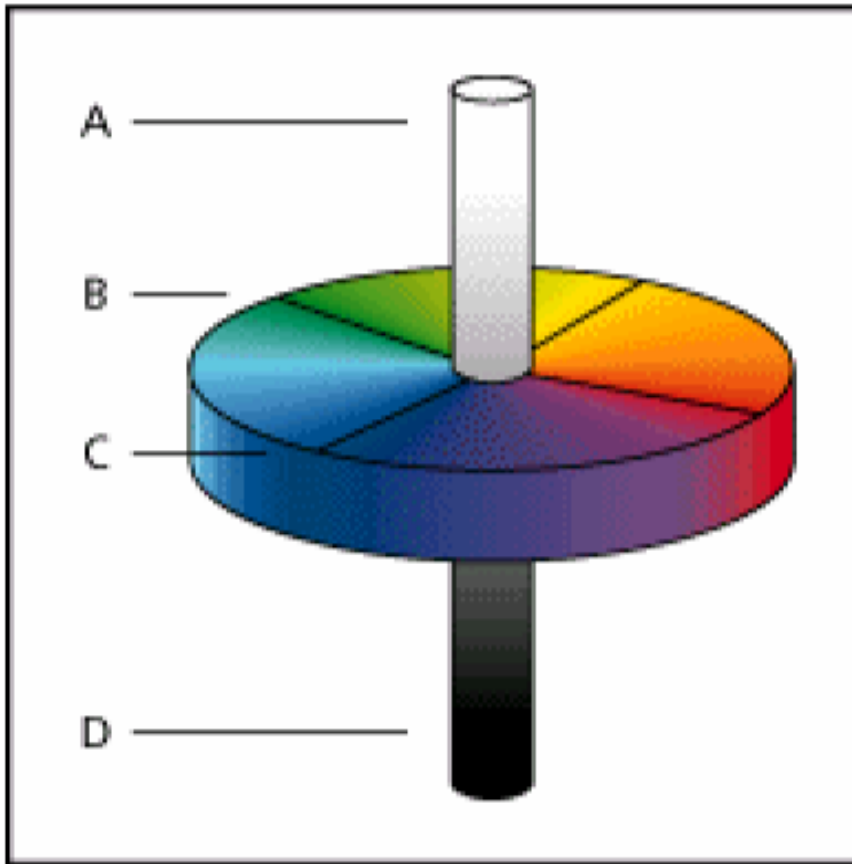
(6) Duotone(雙色調)

- Duotone 影像模式與Bitmap 相同都是只有兩種顏色的像素值，但不同的是Duotone 不一定是黑與白兩種顏色。
- 另外還有Tritone(three color)、Quadtone(four color)等模式，但這些影像模式比較少被使用。

(6) Lab(Lab 彩色)

- 在電腦螢幕上看到的影像顏色與同一張影像經由印表機輸出的顏色會有誤差。
- L*a*b(簡稱為Lab)彩色模型是由CIE (Commission Internationale d 'Eclairage) 於1976 年所制定的，主要是為與設備無關(deviceindependent)的顏色表示法，為的是讓影像不管在螢幕上、紙上或掃描器擷取的同一張影像都得到相同的顏色。

- Lab 包含了一個亮度值(Luminance) ，L(範圍從0 到100)及兩個彩色值(chromatic components) ，當 $L = 0$ 時，代表黑色，而 $L=100$ 時則表示白色。而a component 是用來表示從綠色到紅色的程度，令一個b component 則是用來表示由藍色到黃色的程度，參考圖4-20。



A. 明度 = 100 (白) **B.** 綠到紅成份 **C.** 藍到黃成份 **D.** 明度 = 0 (黑)

圖4-20、Lab 影像模式

(7) YUV & YIQ

- YUV 和 YIQ 是兩種常用於視訊影像中的的模式，YUV 主要是用於 PAL 視訊標準而 YIQ 則是用於 NTSC 視訊標準上。
- YUV 包含了三個組成元素 Y、U、V，而 YIQ 則是由 Y、I、Q 所組成，兩者均可從 RGB 轉換得來，其轉換公式定義如下：

$$Y = 0.299R + 0.587G + 0.114B$$

$$I = 0.596R - 0.274G - 0.322B$$

$$Q = 0.211R - 0.523G + 0.312B$$

$$U = -0.148R - 0.289G + 0.437B$$

$$V = 0.615R - 0.515G - 0.100B$$

在PAL 實作上，U 和V 是採用下列轉換公式：

$$U = 0.492 (B - Y)$$

$$V = 0.877 (R - Y)$$

- YIQ 則可以簡化為下列的轉換矩陣

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- 事實上，Y 即是由彩色轉換成灰階影像的灰階值或稱之為亮度值(luminance)。
- 其轉換公式主要是依據人類眼睛對於紅、綠、藍三原色不同敏感度而來，其中係數越大則表示人的眼睛對於該顏色較為敏感，所以三中顏色之敏感度依序為綠色(0.587)、紅色(0.299)、藍色(0.114)。

(8) 檔案格式

- 檔案格式(file format)指的是電腦系統如何把影像的像素值儲存在記憶體設備(磁碟、硬式磁碟、光碟、碟帶等)中，檔案格式與壓縮方式及檔頭(file head)的設計方式有關。
- 常見的檔案格式有PSD(Photoshop)、GIF、TIFF、TGA、JPEG、PNG、EPS(enhanced postscript)等等。

4.2 影像壓縮與影像安全之目的

• 4.2-1 影像壓縮

- 影像數位化是近年來影像資料常用之儲存及傳送方式，在儲存數位影像時往往佔用許多空間，在傳輸過程中更是耗費相當多的時間。
- 例如：每張低解析度(電視等級)品質、彩色視訊影像所含資料量：512×512 像素/顏色，8 位元/像素，3 顏色→約 6×10^6 位元。

- 若經由電話線傳輸此低解析度(512×512×8 位元/像素×3 顏色)的影像。使用bound 為9600(bound 即位元/秒)的modem，那麼傳輸時間為11 分鐘。
- 這對於大部分的系統是無法容忍的，因此影像壓縮最主要的目的是為了減少影像儲存空間及加速影像傳輸速度。

- 一般影像壓縮方式大多是失真性壓縮(lossy compress)，因為些許的失真並不會造成影像太大的問題，因為人的眼睛是相當不敏銳。目前常見的壓縮技術為JPEG、MPEG…等。

- 資料壓縮：減少所謂的“累贅資料”(data redundant)，這是資料被壓縮的主要原因。
- 對於數位資料，常討論到的資料累贅包含(但不是只有以下所列)：編碼累贅、取樣間累贅及畫面間累贅。藉由減少，乃至於去除這些累贅中的一樣或多樣，資料可以達到壓縮的目的。

4.2-2 數位影像浮水印

- 數位影像浮水印是在數位影像中加入有意義的圖形或文字(圖4-19)，保障此影像之著作權，以防止影像遭盜用的應用問題。
- 數位浮水印可以被看成影像的數位簽章(digital signature)。類似畫家在其所創作的作品上簽名，但是其中最大的不同點在於：數位媒體極易遭資訊技術來加以篡改，使得合法者的簽章很容易遭到破壞。

Lena 影像



藏有數位浮水印
之 Lena 影像



數位浮水印

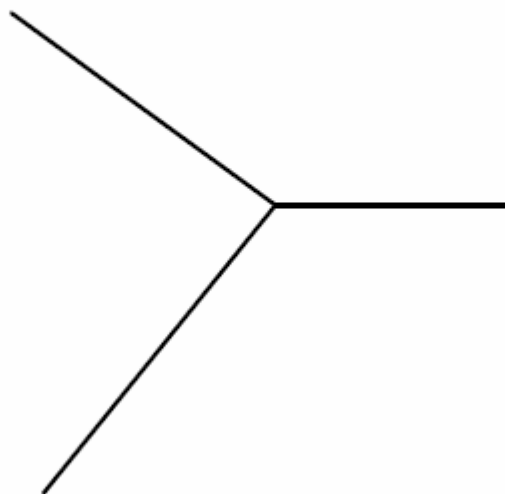
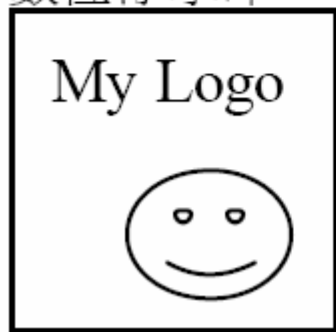


圖4-21 藏匿浮水印的影像

- 目前發展的浮水印技術均是由使用者自行執行藏入以及驗證的工作，因此Craver等人提出了，其他人也可以聲稱其具有該媒體之所有權，也就是任何人都可以在已被加入浮水印的媒體上在加入自己的浮水印，如此一來，同一份媒體上存在了兩份浮水印。
- 如前所述，同一份媒體藏有多個浮水印是可以被允許的，因此到底誰才是真正且合法的該媒體擁有者便無從得知。

- 如Voyatzis 及Pitas提出數位浮水印必須透過可信賴的第三者來協助，真正的解決智慧財產權保護的問題。
- 因為它必須面對廣大的使用大眾，這會使得此一可信賴的第三者，對於運算和儲存空間之負荷過大。
- 如何設計出一個有效率且具安全性的演算法，透過認證中心之協助來達成上述目的亦是數位浮水印技術相關研究中，相當重要的一環。

4.3 影像品質評估法

- 將一張數位影像經過影像壓縮和原始數位影像做比較，會發現此二張影像有所差異。表達其差異程度的方式有：均方根誤差(Root Mean Square Error, RMSE)、訊號雜訊比(Signal-to-Noise Ratio, SNR)和訊號峰雜訊比(Peak Signal-to-Noise Ratio, PSNR)。

- 令一維 f 表示原訊號， f' 表示壓縮-解壓縮後之訊號，訊號長度為 N ，則 RMSE 與 MSE 為：

$$\sigma = RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (f(i) - f'(i))^2}$$

$$\sigma^2 = MSE = \frac{1}{N} \sum_{i=1}^N (f(i) - f'(i))^2$$

- 若訊號為二維影像，且其大小為 $N \times N$ ，則則 RMSE 與 MSE 為：

$$\sigma = RMSE = \sqrt{\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (f(i, j) - f'(i, j))^2}$$

$$\sigma^2 = MSE = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (f(i, j) - f'(i, j))^2$$

- PSNR 之定義為：(Q 為取樣之最大值，如 8 位元之影像其 Q 值為 255)

$$PSNR(dB) = 10 \log_{10} \left(\frac{Q^2}{\sigma^2} \right)$$

$$SNR(dB) = 10 \log_{10} \left(\frac{\sigma_f^2}{\sigma^2} \right)$$

$$\sigma_f^2 = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N f(i, j)^2$$

- 兩張數位影像經 PSNR 計算後的結果值越大(RMSE 值越低)，則我們就認為這兩張數位影像越相似；換言之經處理過的影像越佳。
- 因此，除了 RMSE (或 PSNR) 這一類的條件來評估影像品質外，還是需要再由人眼輔以判斷才是比較客觀與公正的做法。