

第四章 影像融合相關研究

文獻中有關影像融合的相關研究已有許多，典型的影像融合方法分為尋找拍攝裝置轉換函數與多重解析度的影像融合方法兩種。尋找拍攝裝置轉換函數的原因是使用者不知道拍攝裝置的轉換函數，因此假設轉換函數為非線性，若要還原拍攝情境的動態範圍，一個根本的方法是直接求得情境的光強度，這得先還原轉換函數才能辦到。多重解析度的影像融合方法將影像分解為不同解析度的子影像，針對相同解析度但不同曝光的子影像作特徵抽取，計算出融合後影像所需的特徵，即可得到融合影像各個多重解析度的子影像，重新還原成單一解析度即為融合影像。以下詳細介紹幾個前人所提出的典型影像融合演算法。

4.1 尋找數位相機非線性轉換函數的影像融合

1997 年 Paul E. Debevec 等人提出一種方法可從多張照片回復高動態範圍情境的光強度映射圖(radiance maps)，光強度映射圖指的是情境反射光的光強度集合。當我們在拍照時，不是使用類比的軟片(film)感光，就是使用影像感測器來感測並將照片數位化取得二維陣列的亮度值，這些值都是情境中相對輻射強度的真實量測，如果有某一個像素的數值是另一個像素的兩倍，那麼其輻射強度不可能是另一像素兩倍，也就是說像素的數值與輻射強度是呈非線性關係。

此非線性的關係事先很難得知，因為事實上此非線性關係是由各種取像處理的非線性關係融合出來的，如圖 4-1 所示，在傳統的類比相機中，軟片首先曝光一段時間並且形成一個潛在的影像，之後軟片被發展成可將潛在的影像轉變為軟片上密度的變化，此軟片可用軟片掃描器(film scanner)將之數位化，也就是讓光線經由軟片投射到一個電子光感測陣列，將影像轉換成電壓，這些電壓被數位化成數值，經過處理後使用儲存裝置保存。而軟片掃描器本身也會造成非線性的轉換關係。在軟片剛開始的處理階段，軟片對曝光量 X (也就是軟片接受的光強度

E 乘上曝光時間 Δt) 的反應是非線性的函數，此函數稱為軟片的特性曲線，典型的特性曲線在沒有曝光時呈現小的反應，在高度曝光時呈現飽和。軟片掃描與數位化處理會造成自身的非線性特性，使得像素曝光量 X 與像素數值 Z 呈非線性。

數位相機使用電荷耦合裝置(charge coupled device, CCD)陣列來對情境取像，也是有同樣的問題，雖然 CCD 收集的電荷與其光強度成正比，但是大部分的數位相機在還未將影像寫入儲存裝置前對 CCD 輸出應用非線性的轉換，這非線性轉換使用很多種方法去模仿軟片的反應特性與事先知道的顯示裝置非線性反應，然後將 CCD 的類比到數位信號轉換器的 12 位元輸出轉換成通常用來儲存影像的 8 位元資料。這非線性曲線中最特別的地方在於它的飽和點，光強度超過一個特定階層的任何像素都會對應到相同的最大影像數值。

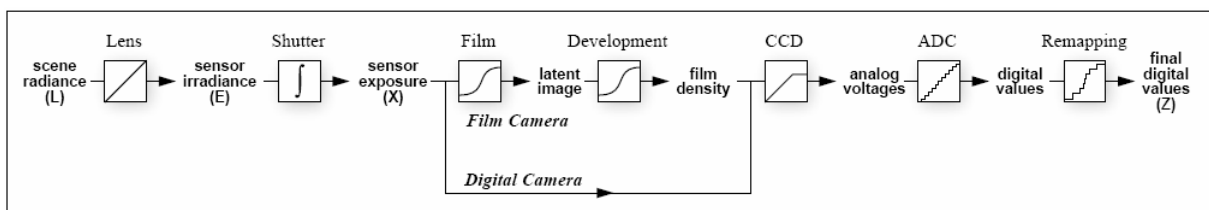


圖 4-1 影像擷取流程圖

為了得到情境中完整的動態範圍，可以使用一系列不同曝光的彩色照片，這篇提出一個簡單的技術可以還原這個非線性反應函數，並且使用這些照片中的像素數值去建立一個精確的光強度映射圖(radiance map)，此光強度映射圖將會包含由原有照片擷取下來的整個動態範圍。其演算法分為還原軟片反應函數與建立高動態範圍光強度映射圖兩個步驟，分別說明如下。

1. 還原軟片的反應函數

如圖 4-1，光線在經過相機開發的演算法(Development)、掃描與數位化處理之後，可以得到數值 Z ， Z 為原本曝光 X 的非線性函數，稱此函數為 f ，此函數是軟片特性曲線與所有後處理步驟造成的非線性關係組合而成，首先需還原此函數 f ，一旦有了此函數，可以計算每個像素的曝光量 X ，而

$X = f^{-1}(Z)$ 。假設函數 f 是單調遞增，因此其反函數 f^{-1} 有定義，曝光量 X 求得後以及已知曝光時間為 Δt ，可以求得光強度 $E = X/\Delta t$ 。

演算法的輸入是一連串不同曝光時間 Δt_j 的數位影像，假設情境為靜止不動的，軟片上每個像素 i 的光強度 E_i 是常數，像素數值稱為 Z_{ij} ，其中 i 是一個空間上的像素指標而 j 是曝光時間指標，可以寫成如公式(1)所示：

$$Z_{ij} = f(E_i \Delta t_j) \quad (1)$$

假設 f 是單調且可逆的，因此可重寫公式(1)為 $f^{-1}(Z_{ij}) = E_i \Delta t_j$ ，將兩邊取自然對數可以得到 $\ln f^{-1}(Z_{ij}) = \ln E_i + \ln \Delta t_j$ ，為了簡化表示，定義函數 $g = \ln f^{-1}$ ，因此可得到公式(2)方程式的集合：

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j \quad (2)$$

公式(2)的方程式集合中 Z_{ij} 與 Δt_j 是已知，未知的是 E_i 與函數 g ，不過可以假設 g 是平滑且單調的。我們希望可以還原函數 g 與光強度 E_i ，以最小平方誤差(least-squared error)的方式，使得這兩者能最滿足公式(2)的方程式集合。需注意的是還原函數 g 只需要還原有限數目的數值 $g(z)$ ，因為像素數值 Z 是有限的。令 Z_{\min} 與 Z_{\max} 是最小與最大的像素數值， N 是像素的個數， P 是照片的張數，我們將此還原的問題公式化，轉變為發現 $(Z_{\max} - Z_{\min} + 1)$ 個 $g(z)$ 與 N 個 $\ln E_i$ ，使能最小化如公式(3)的二次目標函數 O ：

$$O = \sum_{i=1}^N \sum_{j=1}^P [g(Z_{ij}) - \ln E_i - \ln \Delta t_j]^2 + \lambda \sum_{z=Z_{\min}+1}^{Z_{\max}-1} g''(z)^2 \quad (3)$$

公式(3)等號右邊第一項確保得到的解能以最小平方誤差來滿足公式(2)的方程式集合。第二項是一個平滑項以確保函數 g 是平滑的，我們設定

$g''(z) = g(z-1) - 2g(z) + g(z+1)$ ，此平滑項是必要的，因為在最小化的過程中它可以考慮到相鄰的 $g(z)$ ，而 λ 可以對平滑項進行加權計算，必須審慎選擇。因為目標函數 O 中的 E_i 多項式與 $g(z)$ 多項式均為二次多項式，因此最小化 O 是一個直接的 linear least squares problem，可以使用奇異值分解(singular value decomposition)的方法來求得最佳解。

相同的像素值 Z_{ij} 可能對應很多個 $g(z)$ ，可發現是因為曝光時間的不一致造成的，如圖 4-2(a)所示，不同曝光的點形成的曲線相差一個位移量 α ，我們可將各個曲線往上或往下位移 α ，公式(2)左右各加 α 之後，對於公式(3)來說仍不受影響，因此使用一個邊界條件 $g(Z_{mid})=0$ ，其中 $Z_{mid} = \frac{1}{2}(Z_{max} + Z_{min})$ ，可將不同曝光時間正規化為相同的曝光時間，並得到一個完整的反應曲線，如圖 4-2(b)所示。

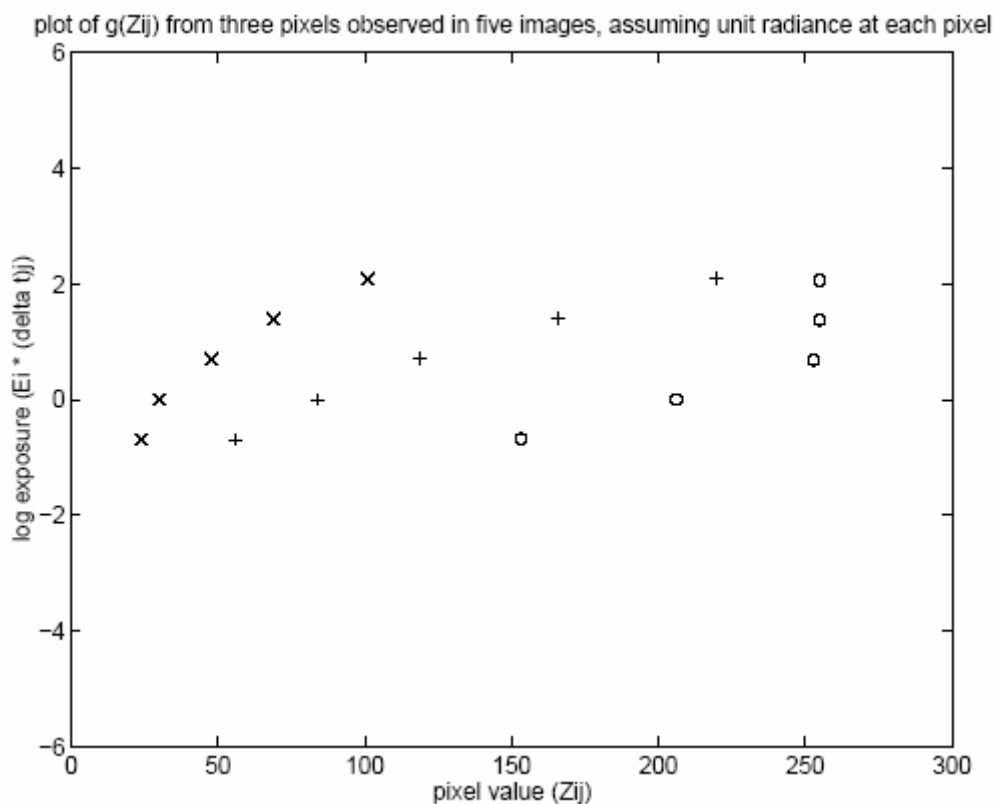
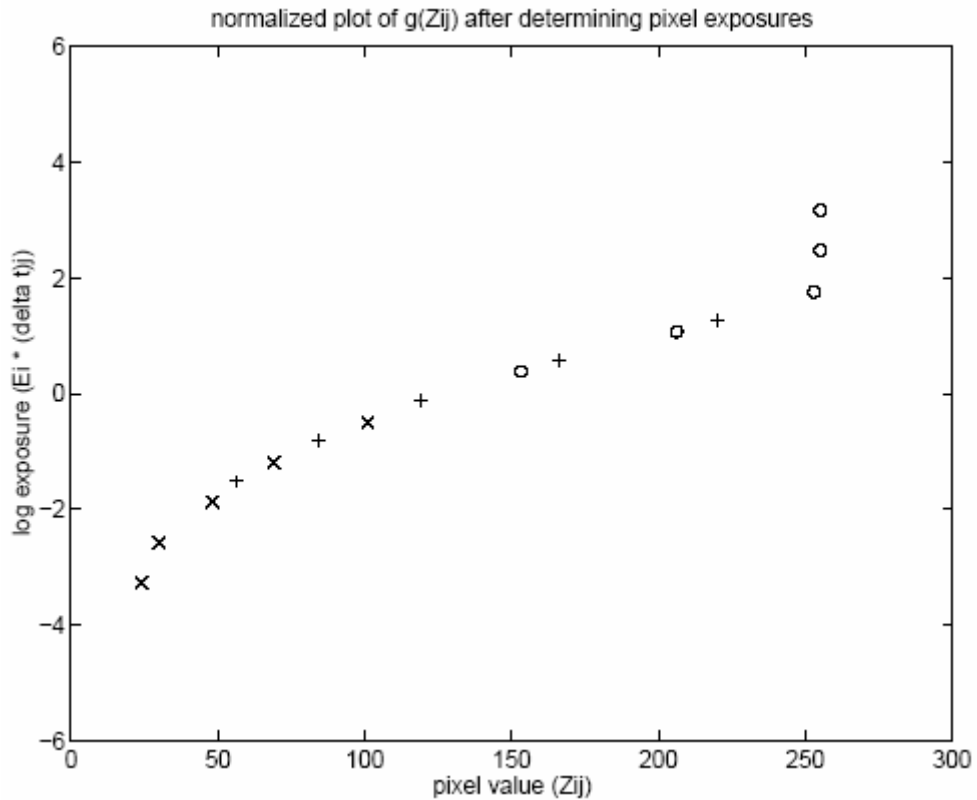


圖 4-2(a) 未正規化的反應曲線



(b) 正規化後的反應曲線

圖 4-2 正規化前與正規化後之反應曲線

由於實際上函數 $g(z)$ 並非真正平滑，在 Z 接近 Z_{min} 或 Z_{max} 時曲線是十分陡峭的，而公式(3)求得的 $g(z)$ 是較平滑的，因此我們需要加入權重的觀念，如公式(4)所示，越接近 Z_{mid} 的權重越高，越接近 Z_{min} 或 Z_{max} 的權重越低，因此公式(5)求得的 $g(z)$ 會比較接近真實的特性。

$$w(z) = \begin{cases} z - Z_{min} & \text{for } z \leq \frac{1}{2}(Z_{min} + Z_{max}) \\ Z_{max} - z & \text{for } z > \frac{1}{2}(Z_{min} + Z_{max}) \end{cases} \quad (4)$$

$$O = \sum_{i=1}^N \sum_{j=1}^P \{w(Z_{ij})[g(Z_{ij}) - \ln E_i - \ln \Delta t_j]\}^2 + \lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} [w(z)g''(z)]^2 \quad (5)$$

2. 建立高動態範圍的光強度映射圖

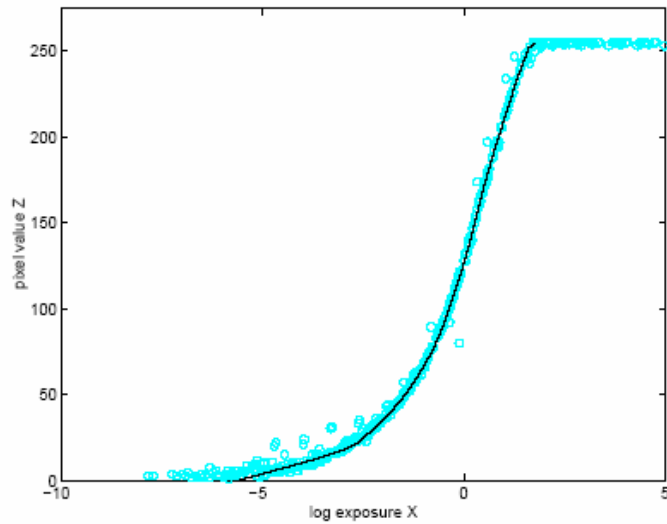
一旦反應曲線 $g(z)$ 求得之後，假設曝光時間 Δt_j 為已知，可以很快地將像素數值轉換成光強度值，由公式(2)可以知道 $\ln E_i = g(Z_{ij}) - \ln \Delta t_j$ ，由於有很多張影像，必須挑選較合理的像素來求得光強度值，因此我們可以用公式(6)來求得，其中 $w(Z_{ij})$ 如公式(4)所示。

$$\ln E_i = \frac{\sum_{j=1}^P w(Z_{ij})(g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^P w(Z_{ij})} \quad (6)$$

如圖 4-3，使用 11 張不同曝光的灰階圖首先最佳化出圖 4-5(a)的反應曲線，再還原出如圖 4-4(b)的高動態範圍光強度映射圖，此高動態範圍光強度映射圖是藉由對光強度取自然對數後映射成灰階圖的結果。



圖 4-3 數位相機以不同曝光拍攝的灰階圖



(a) 曝光量與像素值之反應曲線



(b) 還原後的高動態範圍光輝映射圖

圖 4-4 反應曲線與光輝映射圖

4.2 採取金字塔結構的影像融合

1999 年 Luca Bogoni、Michael Hansen 與 Peter Burt 等人提出一個可以有效延展影像感測器的動態範圍與焦深(depth of focus)的圖樣選擇式(pattern-selective)彩色影像融合方法。圖樣選擇的方法提供了一個融合多重影像重要資訊的機制，此機制主要概念在於經由在多階子影像中選擇特徵較顯著的來源像素。此論文提出此影像融合方法的架構可以應用在灰階、彩色與多頻帶的影像。

融合多重影像成單一且品質提升的影像重點在於圖樣選擇的影像融合本身，為了簡化描述，只介紹融合兩張影像的演算架構。假設影像 C 是由影像 A 與 B 融合而來，整個系統架構如圖 4-5 所示，首先將輸入影像 A 與 B 分解為金字塔(pyramid)式影像集合，根據 A 與 B 的金字塔影像集合分別量測特徵的顯著性 $\sigma(A)$ 與 $\sigma(B)$ ，經由一個選擇的流程 δ ，此流程 δ 根據 A 與 B 金字塔測量出來的特徵明顯度大小，結合 A 與 B 的特徵明顯度形成融合後的選擇係數金字塔 M，根據影像 A 與 B 金字塔與選擇係數金字塔，融合出最後的影像 C。各個處理步驟說明如後。

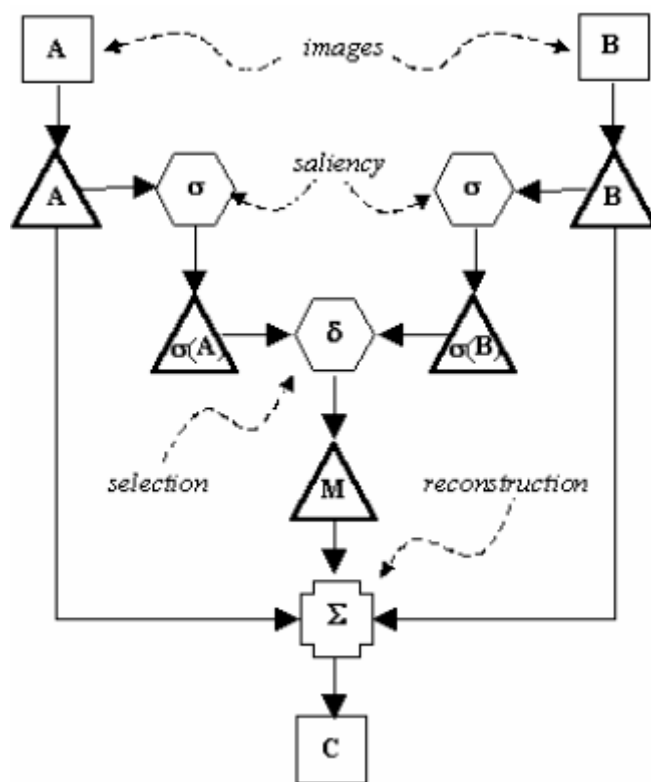


圖 4-5 圖樣選擇式影像融合系統架構

1. 金字塔的建立

金字塔結構的影像表示法主要用意在於估計來源影像特徵的明顯度以及最後融合影像的重建。最高階層解析度(level 0)意指影像原始解析度，而最低階層的解析度(level N)意指影像解析度的寬度與高度已經由原來的解析度減少了 2^N 倍。此影像融合架構主要使用的金字塔是 Laplacian pyramid 表示

法，Laplacian pyramid 主要依據 Filter Subtract Decimate(FSD)的方式建立的 [2]，第 k 個 Laplacian pyramid 階層 L_k 是從相對應的第 k 個 Gaussian pyramid 階層 G_k 依據如公式(7)的關係得到的，其中 w 代表一個 binomial Gaussian filter，通常以 5×5 像素的範圍作濾波。

$$\begin{aligned} L_k &= G_k - w * G_k \\ &= (1-w) * G_k \end{aligned} \quad (7)$$

當建構一個 Laplacian pyramid 時，由於有消除(Decimate)這個過程，以及濾波器 w 不夠理想的因素，在使用公式(7)的金字塔重建影像時會有資訊的遺失，為了部分校正此效應，在公式(7)附加一個校正項，此校正項為原來 Laplacian pyramid 減去經過 w 濾波後的 Laplacian pyramid，因此形成如公式(8)的金字塔。

$$\begin{aligned} \tilde{L}_k &= L_k + (1-w) * L_k \\ &= (2-w) * (1-w) * G_k \end{aligned} \quad (8)$$

2. 特徵明顯度(Saliency)計算

特徵明顯度計算過程 σ 代表了一連串的數學函數，這些函數作用在影像 A 與影像 B 的金字塔影像集合以得到由特徵明顯度組成的金字塔。特徵明顯度函數主要在擷取融合資訊的重要性，例如當融合の影像有不同的焦距時，當某一個焦距の影像上明顯表現比較好時，必須要有一個量測方法可以判別出好壞，因此好的特徵明顯度量測是非常重要的。在此例中一個適當の量測方法是著重在邊緣或是頻帶能量の量測，因此可能是一個如公式(9)の第 k 階金字塔明顯度量測方法。公式(9)代表當第 k 階金字塔在融合時必須選擇來自邊緣能量較強の來源影像の資訊。

$$\sigma_k(L_{Ak}) = (L_{Ak})^2 \quad (9)$$

3. 選擇機制

選擇的機制 δ 也代表了一連串的數學函數，這些函數作用在影像 A 與 B 的特徵明顯度金字塔 $\sigma(A)$ 與 $\sigma(B)$ ，這兩個明顯度金字塔是經由特徵明顯度計算過程而來，經由選擇的機制得到一個選擇係數的金字塔 M，此金字塔定義了兩張影像的像素需加重多少權值來選取，特徵明顯度越高的像素權值越高，如公式(10)所示， δ_k 代表一個第 k 階的選擇函數，例如可以是取最大值函數。

$$M_k = \delta_k(\sigma_k(L_{Ak}), \sigma_k(L_{Bk})) \quad (10)$$

4. 重建融合影像

重建的過程 Σ 作用在影像 A 與影像 B 的金字塔影像集合，並配合選擇係數金字塔 M 以求得最後的融合影像 C。分別由最低階到最高階的 Laplacian pyramid 進行重建，重建時如公式(11)與(12)，融合後第 k 階的 Laplacian pyramid L_{Ck} 是由 L_{Ak} 與 L_{Bk} 配合係數金字塔 M_k 得來，而從最 N 階 Laplacian pyramid 一直重建到第 k 階 Laplacian pyramid 的重建影像為 C_k ，而重建到最後的 C_0 即為融合後影像 C。

$$L_{Ck} = M_k L_{Ak} + (1 - M_k) L_{Bk} \quad (11)$$

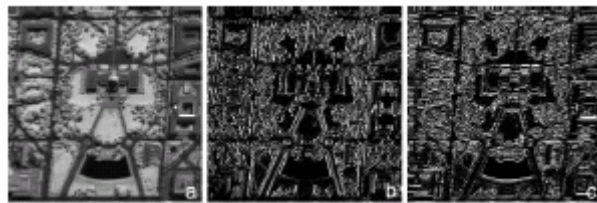
$$C_k = L_{Ck} + w * [C_{k+1}]^{\wedge 2} \quad (12)$$

換個方式來說，我們可將公式(11)表示為較易理解的映射函數，令 $a = L_{Ak}(i, j)$, $b = L_{Bk}(i, j)$, $c = L_{Ck}(i, j)$, $u = M_k(i, j)$ ，而 l 與 h 代表 $M_k(i, j)$ 的高臨界值與低臨界值，可形成如公式(13)的映射函數。當選擇係數大於或小於某個臨界值時，則完全選擇某一方的資訊，否則皆結合兩方不同的資訊。

$$c = \begin{cases} a & u < l \\ b & u > h \\ ua + (1 - u)b & \text{otherwise} \end{cases} \quad (13)$$

4.3 以梯度為基礎之多重解析度影像融合

2004 年 Valdimir S. Petrovic 與 Costas S. Xydeas 提出一個以梯度為基礎的影像融合演算法[9]，此方法概念在於 1. 針對梯度圖(gradient map)來融合的方法以及 2. 先融合再作多重解析度分解的新概念。梯度圖涵蓋了影像所有變化的資訊，將影像表示為水平梯度圖與垂直梯度圖，如圖 4-6 所示，先各別對水平梯度圖與垂直梯度圖進行影像融合，融合後的梯度圖再以梯度濾波器(gradient filter)分解出各個頻帶的融合影像，經由各個頻帶的融合，形成最後的融合影像。



(a)輸入影像(b)水平梯度圖(c)垂直梯度圖

圖 4-6 梯度圖

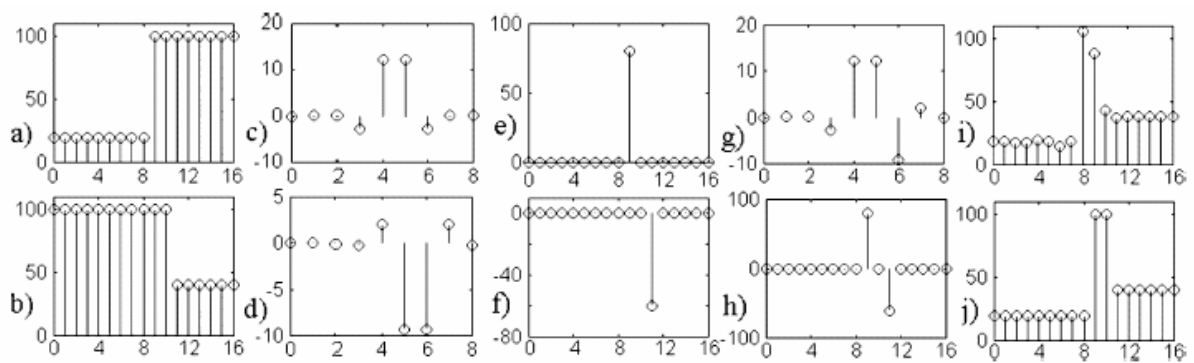


圖 4-7 QMF 與梯度範例

一般的多重解析度影像融合方法均直接對影像作離散小波轉換或是 QMF(Quadrature Mirror Filters) 分解，如此並不能保留來源影像最重要的資訊，如圖 4-7 所示，而梯度圖卻可將影像中所有變化均保留下來，圖 4-7(a)與(b)為一維輸入影像，(c)與(d)為輸入影像經由 QMF 的結果，(e)與(f)為輸入影像的梯度，(g)為(c)與(d)融合的結果，(h)為(e)與(f)融合的結果，(g)經由重建之後的影像為(i)，

而(h)經由重建之後的影像為(j)，從圖中可看出經由 QMF 分解之後影像融合的影像仍有小漣波現象，但梯度圖再經由梯度濾波器分解之後影像融合的影像保存較好的資訊。

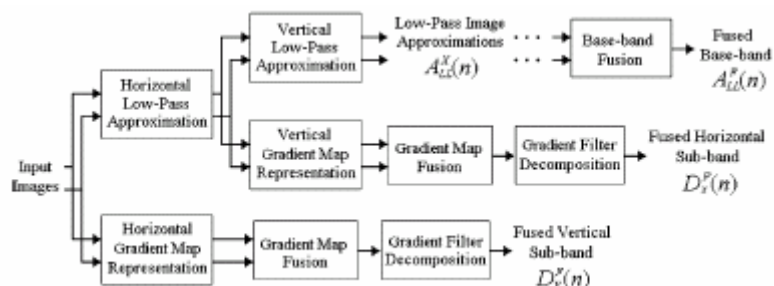


圖 4-8 以梯度為基礎的影像融合與分解架構

此篇論文提出如圖 4-8 的系統架構，輸入影像先經由水平方向的梯度圖作影像融合並用一個梯度濾波器分解取得融合後垂直高頻的頻帶，同時輸入影像經由水平方向的低通濾波器可濾出水平低通頻帶，水平低通頻帶經由垂直方向又可得到垂直梯度圖並且最後得到融合後水平低頻的頻帶，如此分解下去，最後再經由重建程序即可得到融合的影像。