



第十五章

三維影像的彩現

內容

- 15.1 前言
- 15.2 Kinect 系統介紹
- 15.3 翹曲變形技術：DIBR 第一步驟
- 15.4 深度計算模型
- 15.5 缺空填補
- 15.6 結論

15.1 前言

介紹三維視訊的深度影像彩現(Depth-Image-Based Rendering, DIBR)技術。考慮的影像是成對的，一張為彩色影像 (Color Map)，另一張為深度影像 (DepthMap)。彩色影像中同位置的像素搭配同樣位置的深度像素可合成出該彩色像素對應的虛擬彩色像素。

15.2 Kinect 系統介紹

彩色影像是透過正中間的彩色感應鏡頭取得；3D 深度影像是透過兩側的深度感應鏡頭，利用紅外線的反射距離差算出像素的深度值。



圖15.2.1 體感操控器 Kinect



圖15.2.2 Kinect 拍得的彩色影像



圖15.2.3 Kinect 拍得的深度影像

15.3 翹曲變形技術：DIBR 第一步驟

有水平的兩台相機(第一台是真實的，第二台是虛擬的)，第一台相機中相片內有一像素，其像素位置為 $p_1 = (x_1, y_1, 1)^t$ ，我們可以透過該台相機的內部參數矩陣 K_1 、旋轉矩陣 $R_1 = I$ 、平移向量

$T_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ 及該像素之實際深度值，經由式 (15.3.1) 取得 p_1 與世界座標的位置 $P = (X_w, Y_w, Z_w)^t$ 的關係。

$$\begin{aligned}
 Z_w p_1 &= K_1 \left(R_1 \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + T_1 \right) \\
 &= K_1 \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + K_1 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = K_1 \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix}
 \end{aligned}
 \tag{15.3.1}$$

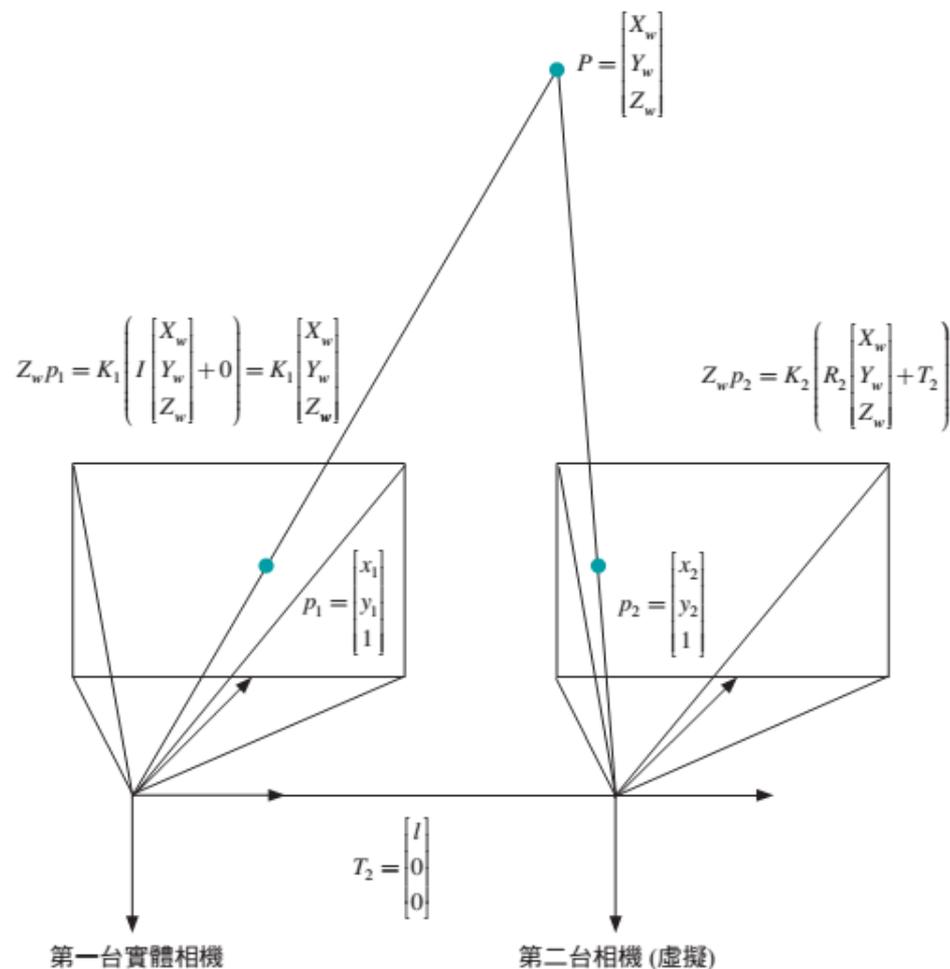


圖15.3.1
世界坐標系中的翹曲變形模型

同理，在第二台相機上，我們有下式：

$$Z_w p_2 = K_2 \left(R_2 \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + T_2 \right) \quad (15.3.2)$$

合併式 (15.3.1) 和式 (15.3.2)，可得

$$Z_w p_2 = Z_w K_2 R_2 K_1^{-1} p_1 + K_2 T_2 \quad (15.3.3)$$

由式 (15.3.3)，第一台相機所拍攝的相片像素 $p_1 = (x_1, y_1, 1)^t$ 轉換到第二台相機的像素 $p_2 = (x_2, y_2, 1)^t$ 上。

對二台相機而言，其拍攝的同一個物體點的深度資訊皆為 Z_w ，

旋轉矩陣為單位矩陣，即 $R_2 = I$ 。平移向量 $T = \begin{bmatrix} l \\ 0 \\ 0 \end{bmatrix}$ 。

平移向量表示只有 x 方向差距 l 的相機水平距離。兩台相機之內部參數矩陣相等，即 $K_1 = K_2$ 。

式 (15.3.3) 可轉換為 $P_2 = P_1 + \frac{K_2 T_2}{Z_w}$ 。

這裡 $K_2 = \begin{bmatrix} f & \tau & O_x \\ 0 & \eta f & O_y \\ 0 & 0 & 1 \end{bmatrix}$ ， f 代表焦距， (O_x, O_y) 代表相機座標系

與影像座標系之間原點的位移向量，如圖 15.3.2 所示， τ 與 η 為硬體感光元件的延遲係數。

由前述推導得 $P_2 = P_1 + \frac{f * l}{Z_w}$ 。

令 $d_x = \frac{f * l}{Z_w}$ 為水平位移，且其代表真實相片的某像素點經由水平翹曲變形技術映射到虛擬相片的水平位移。

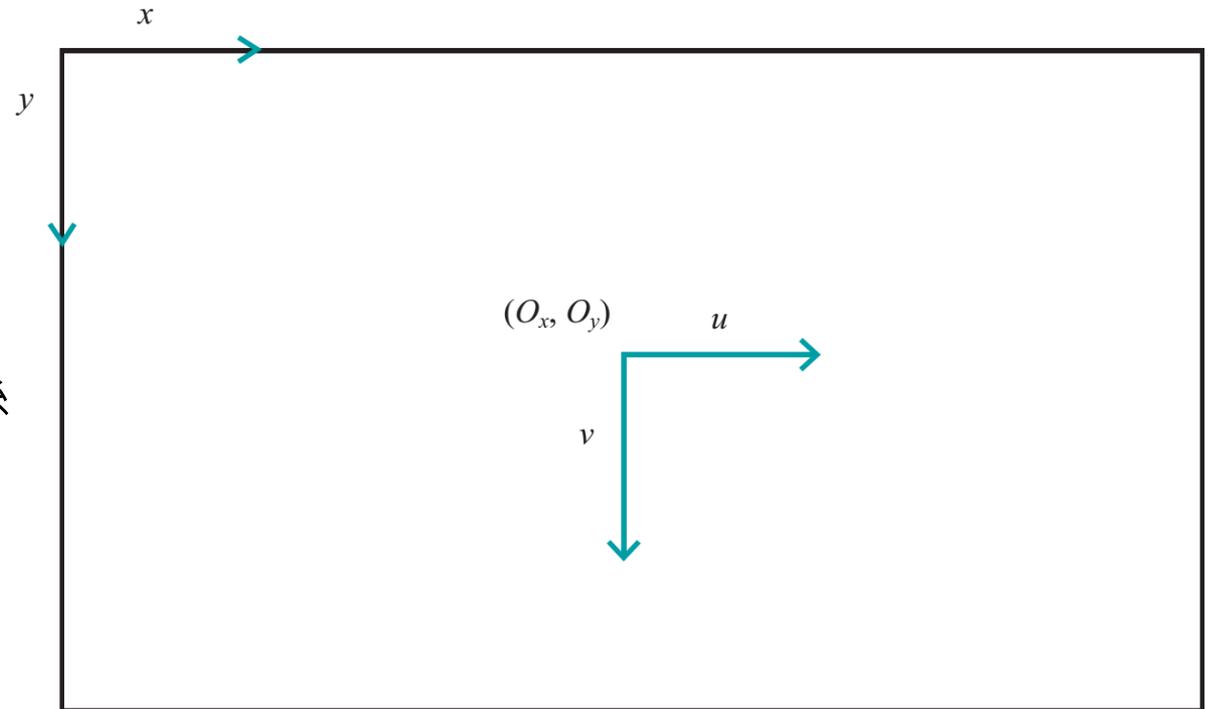


圖15.3.2
 O_x 和 O_y 的平移關係

15.4 深度計算模型

深度影像中的整數深度值 (z) 是由實際深度 (Z_w) 量化所產生的，其量化方法乃將實際深度利用最遠距離 (Z_{far}) 和最近距離 (Z_{near}) 以非線性的方式量化至 0 到 255 的整數值。

$$z = \text{Quant}(Z_w) = \left\lfloor 255 \cdot \frac{Z_{near}}{Z_w} \cdot \frac{Z_{far} - Z_w}{Z_{far} - Z_{near}} + 0.5 \right\rfloor$$

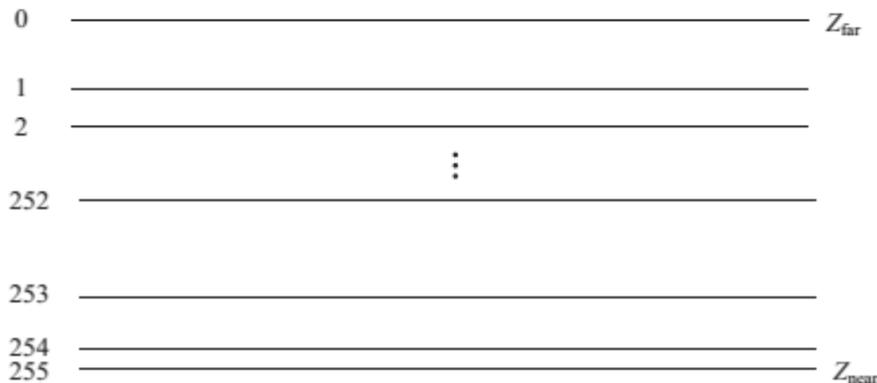


圖15.4.1 深度值的非線性量化

若整數深度值已知，我們可以利用反量化的深度值公式

$$Z_w = \text{Quant}^{-1}(z) = \frac{1}{\frac{z}{255} \left(\frac{1}{Z_{\text{near}}} - \frac{1}{Z_{\text{far}}} \right) + \frac{1}{Z_{\text{far}}}}$$

計算出實際深度值。

在 15.3 節中，已知 $d_x = \frac{f \cdot l}{Z_w}$ 。

水平位移量 dx 與實際深度 Z_w 的倒數有關係，故

$$d_x = \text{Disp}_x(Z_w) = \text{Disp}_x(\text{Quant}^{-1}(z)) = D_x(z)$$

我們稱 $D_x(z)$ 為水平位移函數 (Disparity Function)，此函數是一個嚴格遞增函數。深度值 z 愈大，水平位移就愈大。我們可以將水平位移函數改寫為

$$D_x(z) = \frac{f \cdot l}{\text{Quant}^{-1}(z)} = f \cdot l \cdot (C_1 \cdot z + C_2)$$

其中 $C_1 = \frac{1}{255} \left(\frac{1}{Z_{\text{near}}} - \frac{1}{Z_{\text{far}}} \right)$ 和 $C_2 = \frac{1}{Z_{\text{far}}}$ 。

15.5 空缺填補

處理兩種類型的影像缺空問題：

1. 深度攝影機在擷取深度資訊時產生的缺空
2. 彩現時所產生的缺空

為了使擷取的深度影像具有完整的資訊，及增加彩現後的完整影像呈現，我們會透過缺空填補的演算法，將這些缺空修補起來。

第一種類型的缺空是深度攝影機擷取深度影像時所產生的。有兩個因素導致缺空。



圖15.5.1
黑白區域代表深度影像中的缺空

第一類型缺空的填補方法

- 步驟 1：以每一個缺空像素點為中心在彩色影像相同位置開啟一個搜尋範圍。
- 步驟 2：在搜尋範圍內找尋與缺空像素之顏色相近的點，並將顏色相近的點作為候選點。
- 步驟 3：若沒有找到候選點，則跳到步驟 5；否則，跳到步驟 4。
- 步驟 4：從候選點對應的深度值中，取出其深度的中間值填補缺空的像素點。隨後，往下一個缺空像素進行填補。
- 步驟 5：在搜尋範圍中統計出現頻率最高的深度值，以此深度值作為缺空待修

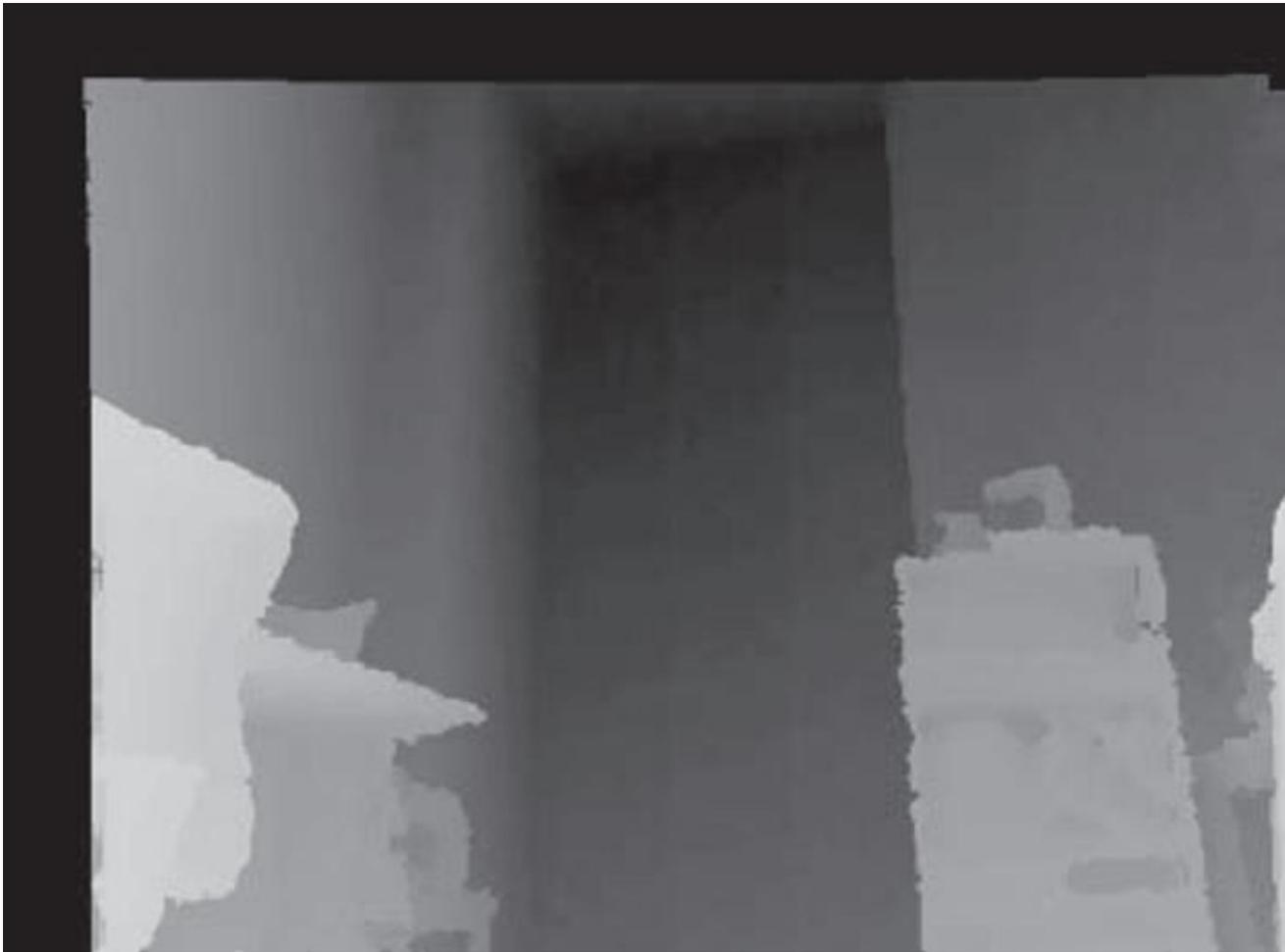


圖15.5.2
圖15.5.1的缺空填補結果

第二種類型是以深度影像搭配彩色影像進行彩現時所產生的缺空。考慮有些距離較遠的背景部分會被距離較近的前景物件所遮蔽，或是攝影機沒有拍攝到的物件都會導致彩現的缺空，這兩種現象我們都稱作遮蔽效應，以圖 15.5.3 為例。

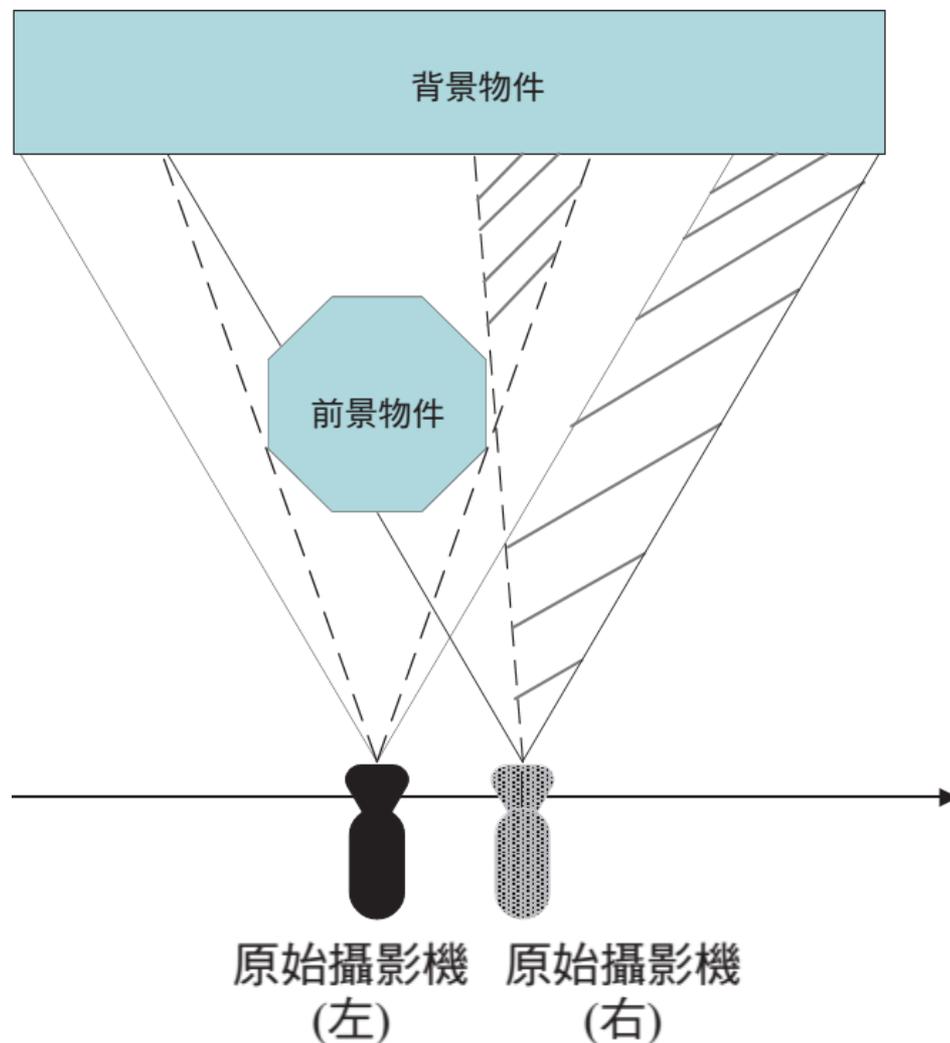


圖15.5.3 遮蔽效應示意圖

另外考慮，在翹曲變形時，以深度影像計算每個像素所對應的新位置可能會產生小數的座標值，經過四捨五入之後形成多個像素對應到相同位置，而造成某些位置沒有像素值可對應，也會形成缺空。



圖15.5.4
翹曲變形後的缺空結果

在此僅介紹一個較容易實作的水平內插法 [9]。基於水平翹曲變形時，得到的新像素位置只有向左或向右其中的一種移動。我們可利用空缺像素的左右方存在的像素，以內插的方法填補缺空。水平內插法以下式表示：

$$I[n] = (1 - w)I_l[n] + wI_r[n], \quad \forall n \in H$$

這裏

$$I_l[n] = I[m_1], \quad m_1 \in \partial H;$$

$$I_r[n] = I[m_r], \quad m_r \in \partial H;$$

$$w = \frac{n - m_1}{m_r - m_1}$$

其中 n 為目前待修補的缺空像素位置， H 為缺空像素的點集合， ∂H 為缺空像素點集合的邊界外緣像素點集合， m_r 和 m_1 分別為位置 n 往右和往左第一個碰到缺空的邊點像素。

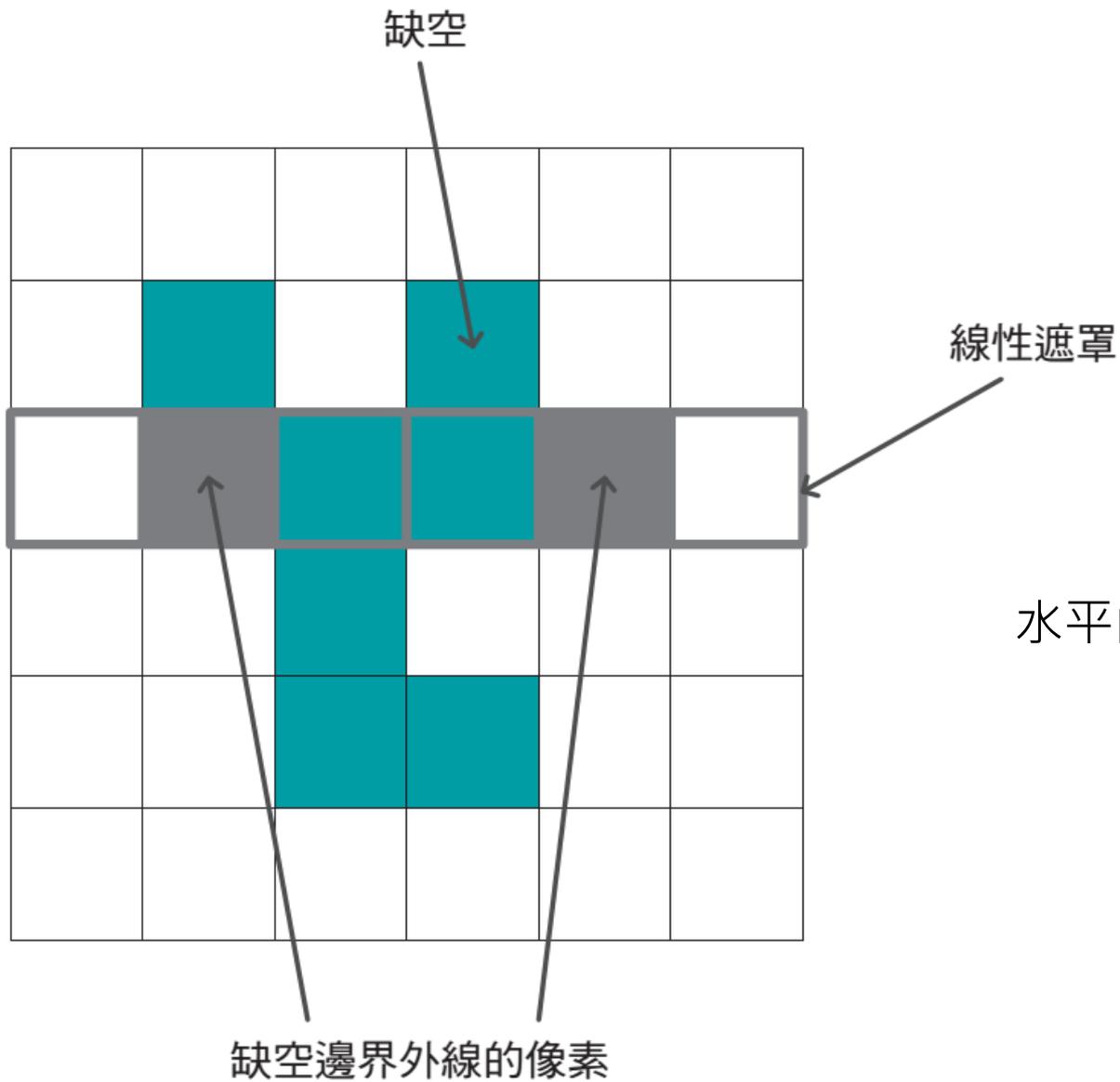


圖15.5.5
水平內插法示意圖



圖15.5.6
圖 15.5.4 以水平內插法填補缺
空的結果