

# 第九章 圖形識別、匹配與三維影像重建

# 內容

- 9.1 前言
- 9.2 統計圖形識別
- 9.3 影像間的匹配對應
- 9.4 匹配演算法原理
- 9.5 三維影像重建
- 9.6 二維影像的深度計算

# 9.1 前言

- 主要介紹統計式圖形識別、角點與關鍵點偵測及特徵表示、匹配和3D重建。

## 9.2 統計圖形識別

- 貝氏決策理論

假設有二類木頭， $A$ 和 $B$ ， $P(A) + P(B) = 1$ 。

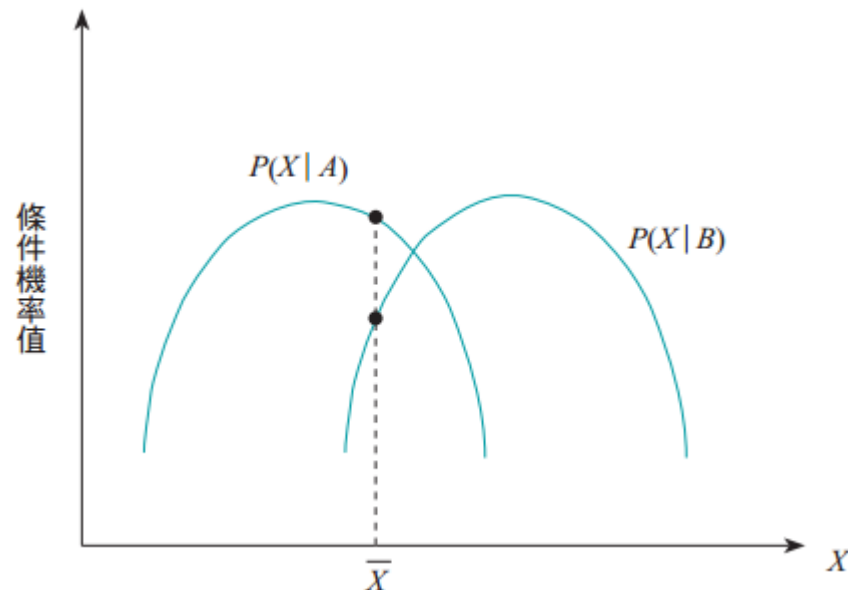


圖9.2.1  $P(X|A)$ 和 $P(X|B)$ 的分佈圖

給一個 $X$ 值，該木頭屬 $A$ 或 $B$ 的機率為何？依據貝氏法則，

$$P(A|X) = \frac{P(A \cap X)}{P(X)} = \frac{P(X|A)P(A)}{P(X)}$$

此處  $P(X) = P(X|A)P(A) + P(X|B)P(B)$ 。

透過貝氏法則， $P(A|X)$ 這個事後機率就可由事前機率 $P(A)$ 、 $P(B)$ 、 $P(X|A)$ 和 $P(X|B)$ 求得。

當  $X = \bar{X}$ ， $P(A|X) > P(B|X)$ ，這時可判斷該木頭為 $A$ ，畢竟冒的風險較低，也就是  $P(error|X) = P(B|X)$ 。簡化計算，當  $P(X|A)P(A) > P(X|B)P(B)$  時，我們判斷該木頭為 $A$ 。

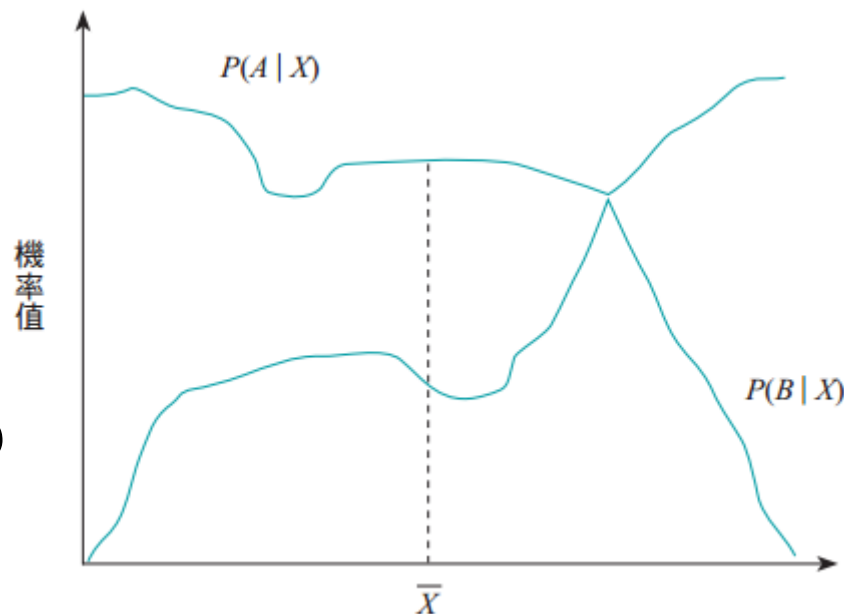


圖9.2.2  $P(A|X)$ 和 $P(B|X)$ 的分佈圖

## ■ 識別器

將紋理由一維擴充到  $d$  維而將木頭的種類由2種擴充到  $t$  種。

令第  $i$  類的識別器為  $g_i(X) = P(X | T_i)P(T_i)$ ，此處紋理向量

$X = (x_1, x_2, \dots, x_d)$  而  $T_i$  代表第  $i$  類木頭， $1 \leq i \leq t$ 。

$$\bar{g}_i(X) = \log P(X | T_i) + \log P(T_i)$$

如果  $\bar{g}_j(X)$  為最大值，  
該木頭分類為  $T_j$

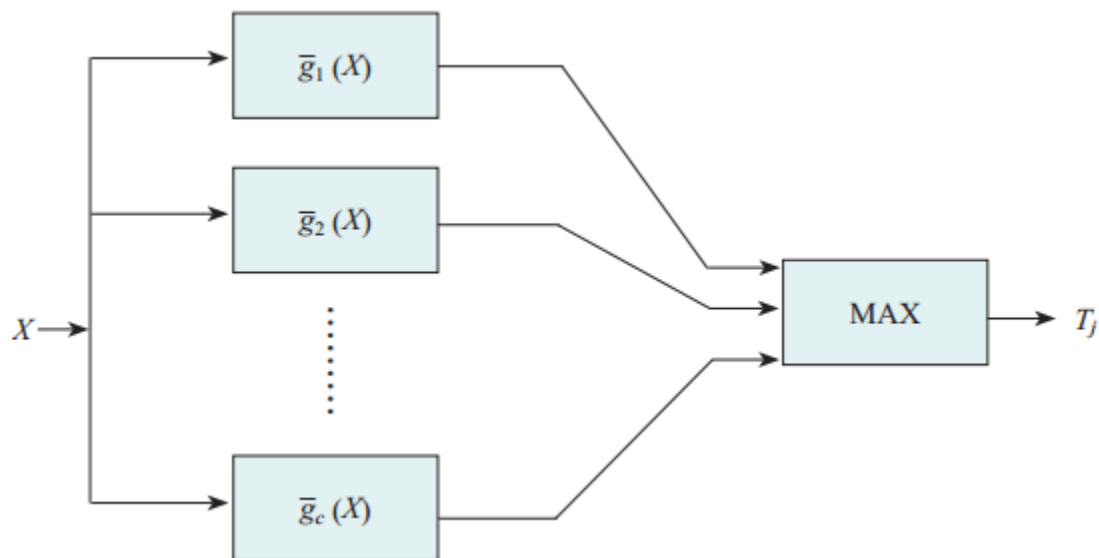
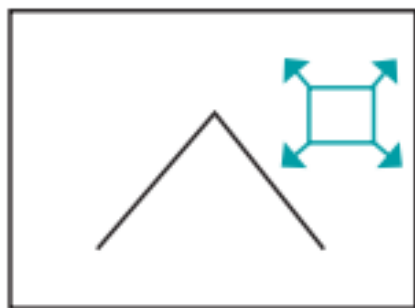


圖9.2.3 識別器示意圖

# 9.3 影像間的匹配對應

## 9.3.1 Harris 角點偵測法

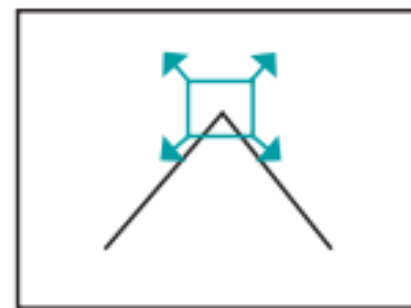
- (1)平面：往任何方向移動僅造成小變化
- (2)含一條邊：與邊平行的變化量小；反之則大
- (3)含角點或獨立點：往任何方向變化皆大



(1)



(2)



(3)

令

$$\nabla_x f = f * (-1, 0, 1) = \nabla_x$$

$$\nabla_y f = f * (-1, 0, 1)^t = \nabla_y$$

利用高斯函數來平滑雜訊的影響，令

$$A = \nabla_x^2 * G \quad B = \nabla_y^2 * G \quad C = \nabla_x \nabla_y * G$$

綜合灰階梯度變化之影響可表示為

$$\begin{aligned} Ax^2 + By^2 + 2Cxy &= (x \ y) \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= (x \ y) M (x \ y)^t = E \end{aligned}$$



- 函數 $E$ 是一種局部自我關聯的函數，矩陣 $M$ 就是函數 $E$ 的代表。矩陣 $M$ 的兩個特徵值代表：
  - $\lambda_1$ 和 $\lambda_2$ 皆很小：代表視窗內為平滑區
  - $\lambda_1$ 和 $\lambda_2$ 中，一大一小：代表含一邊的區域
  - $\lambda_1$ 和 $\lambda_2$ 皆很大：代表含角點的區域

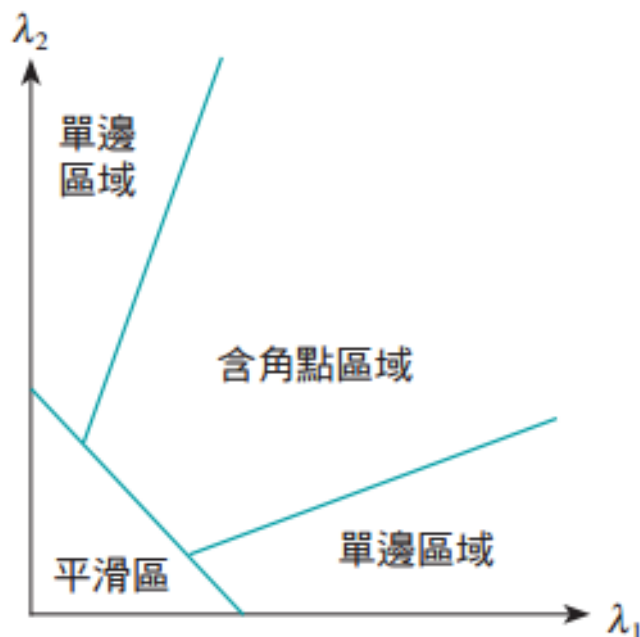


圖9.3.1.1 矩陣 $M$ 的特徵值所代表的意義

- 影響值  $R = \det(M) - k * (\text{trace}(M))^2$  決定是否小區域內有角點，判斷的準則為
  - $R > 0$  且  $R \approx 0$ : 代表平滑區
  - $R < 0$ : 代表含一邊的區域
  - $R \gg 0$ : 代表有角點的區域

利用以上方法可以將所有角點找出來。



(a)原始影像



(b)找出的角點集

## 9.3.2 SIFT關鍵點偵測法

- 將影像進行縮減取樣  $n$  次，例如  $n = 2$  可分成3個影像尺度。
- 利用不同標準差的高斯函數。例如  $m = 4$ ，利用  $\sigma, k\sigma, k^2\sigma, k^3\sigma$  將不同高斯函數與3個影像尺度作迴積運算：

$$L(x, y, k^t \sigma) = G(x, y, k^t \sigma) * I(x, y)$$

接著對同尺度且相鄰的影像做**DOG(Difference-of-Gaussian)**:

$$\begin{aligned} D(x, y, k^t \sigma) &= (G(x, y, k^{t+1} \sigma) - G(x, y, k^t \sigma)) * I(x, y) \\ &= L(x, y, k^{t+1} \sigma) - L(x, y, k^t \sigma) \end{aligned}$$

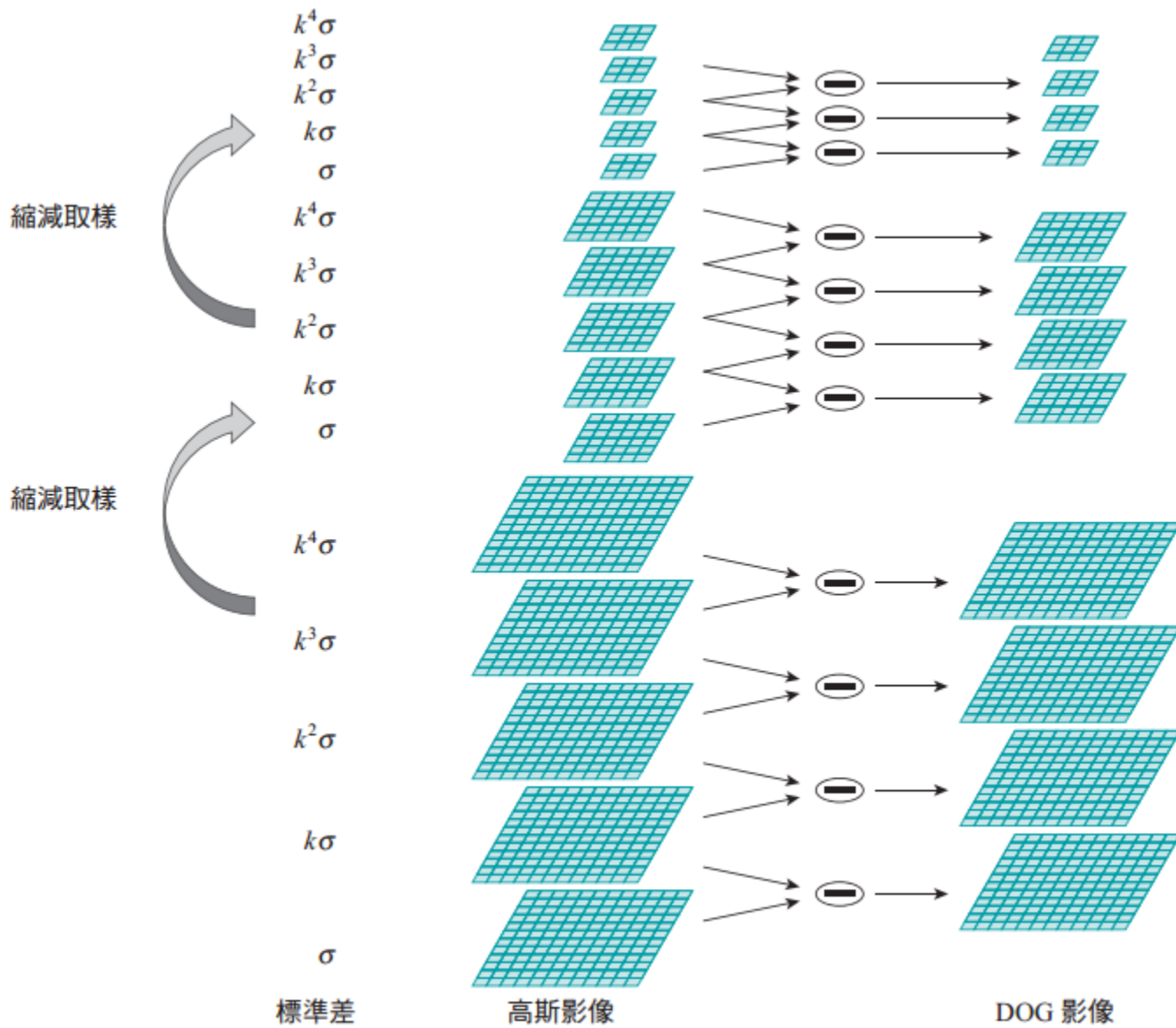


圖9.3.2.1 DOG金字塔

考慮每個像素與周圍8個像素及上下層同位置周圍9個像素做比較，如果該像素為極值，則設為候選關鍵點。

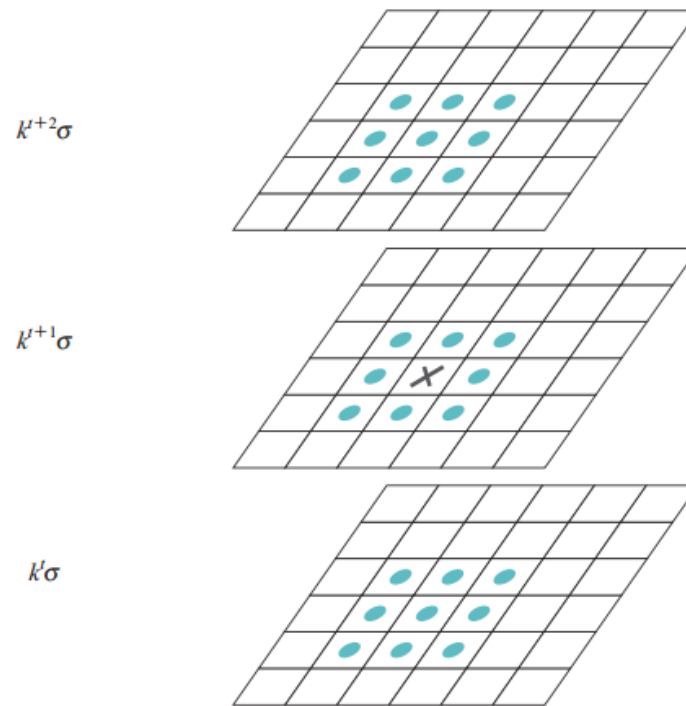


圖9.3.2.2 DOG金字塔找極值

以每個候選關鍵  $X$  點當原點，算出

$$D(X) = D(0) + \frac{\partial D(0)}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D(0)}{\partial X^2} X$$

其中  $X = (x, y, \sigma)^T$ 。令座標  $\mathbb{X}$  時有極值點  $D(\mathbb{X})$ 。

$$\frac{\partial D(0)}{\partial X} + \frac{\partial^2 D(0)}{\partial X^2} \mathbb{X} = 0$$

$$\mathbb{X} = - \left( \frac{\partial^2 D(0)}{\partial X^2} \right)^{-1} \frac{\partial D(0)}{\partial X}$$

如果  $\mathbb{X}$  內的座標值都小於 0.5，則我們將此點視為極值點並檢查  $|D(\mathbb{X})| < 0.03$  是否成立；否則將原點繼續移到新的整數座標點，並且繼續前述步驟，直到收斂。若  $|D(\mathbb{X})| < 0.03$ ，代表此區域的對比度較低，應移除此關鍵點。

欲除掉邊上的候選關鍵點。先求出赫斯矩陣

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

$H$ 的特徵值  $\lambda_1$  和  $\lambda_2$  可計算如下:

$$\begin{aligned} \det(H - \lambda I) &= \det \left( \begin{bmatrix} D_{xx} - \lambda & D_{xy} \\ D_{xy} & D_{yy} - \lambda \end{bmatrix} \right) \\ &= \lambda^2 - (D_{xx} + D_{yy})\lambda + (D_{xx}D_{yy} - D_{xy}^2) = 0 \end{aligned}$$

可得

$$\lambda_1 + \lambda_2 = D_{xx} + D_{yy} = \text{trace}(H)$$

$$\lambda_1\lambda_2 = D_{xx}D_{yy} - D_{xy}^2 = \det(H)$$

令  $\lambda_1 \geq \lambda_2$  且  $\lambda_1 = \gamma\lambda_2$ , 得到

$$\frac{(\text{trace}(H))^2}{\det(H)} = \frac{(\lambda_1 + \lambda_2)^2}{\lambda_1\lambda_2} = \frac{(\gamma\lambda_2 + \lambda_2)^2}{\gamma\lambda_2^2} = \frac{(\gamma + 1)^2}{\gamma}$$

得判斷式

$$\frac{(\text{trace}(H))^2}{\det(H)} < \frac{(\gamma' + 1)^2}{\gamma'}$$

令  $\gamma' = 10$ . 若判斷式成立, 代表主曲率 (Principal Curvature) 的比值小於10, 則此候選關鍵點不在邊上。反之, 則為邊上的候選關鍵點, 我們將其移除。



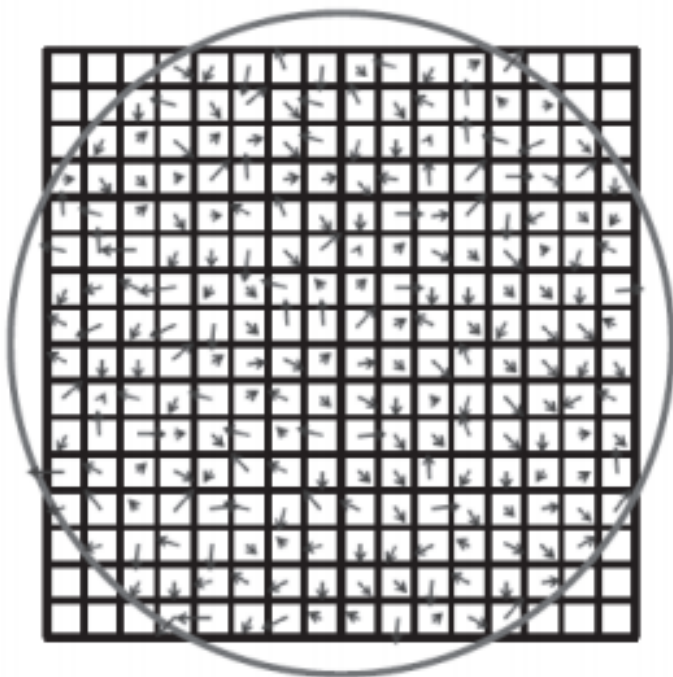
以剩餘關鍵點為中心的區塊，計算區塊內每個像素的梯度值

$m(x, y)$  與梯度方向  $\theta(x, y)$

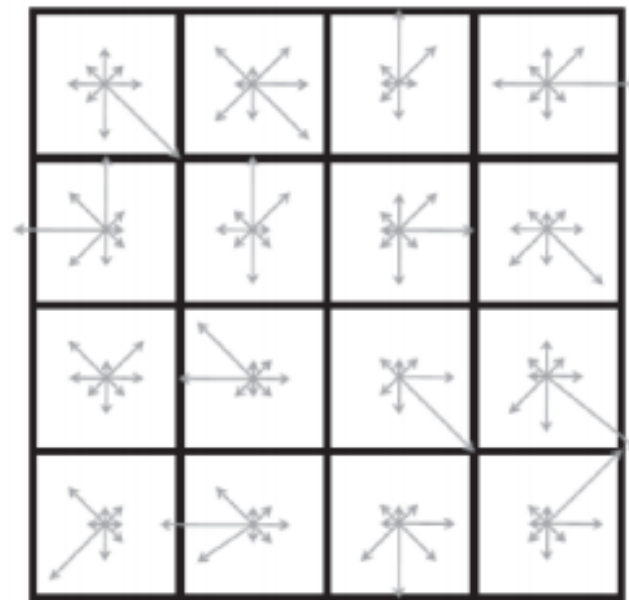
$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2}$$

$$\theta(x, y) = \tan^{-1} [(L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))]$$

關鍵點的旋轉不變性(Rotation Invariant)。



梯度向量



128 維的特徵向量

圖9.3.2.3 決定關鍵點的特徵向量

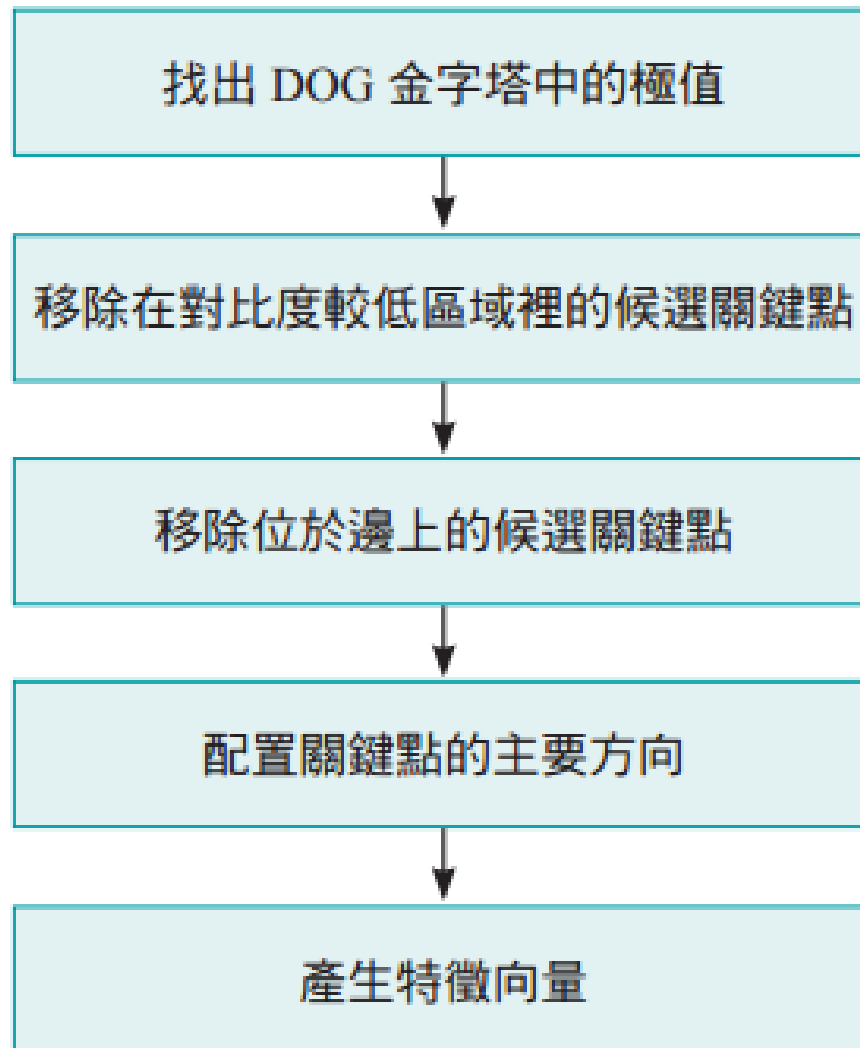
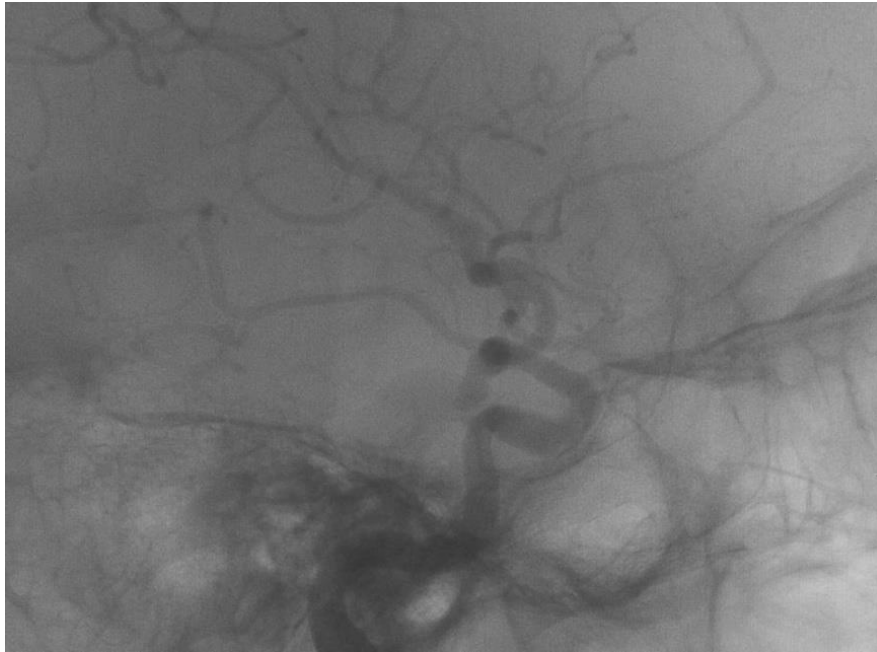
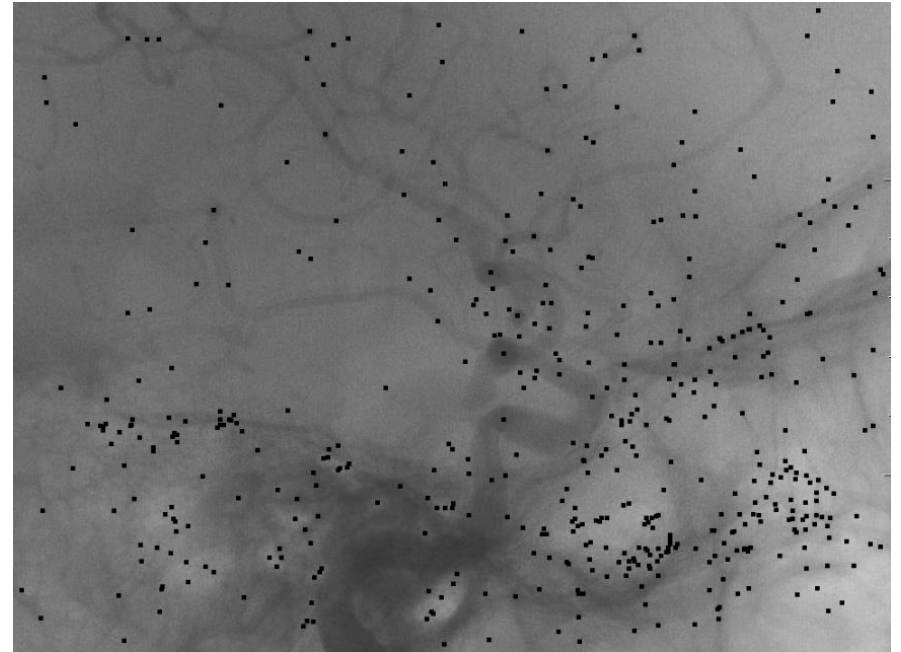


圖9.3.2.4 SIFT流程圖



(a) 輸入的腦血管影像



(b) SIFT演算法選取到的關鍵點

圖9.3.2.5

## 9.3.3 點集合匹配法

- 另一種作法是尋找轉移矩陣 $T$ ，使  $ToF$  與  $\bar{F}$  有最小誤差。

令  $d_i(T)$ ， $1 \leq i \leq m$ ，代表  $F$  中的  $f_i$  與  $\bar{F}$  中的  $\bar{f}_i$  之距離

$$p(d_1(T), d_2(T), \dots, d_m(T) | T) = \prod_{i=1}^m p(d_i(T))$$

- 引入最大可能(Maximum-likelihood)的觀念，得

$$L(T) = \ln p(d_1(T), d_2(T), \dots, d_m(T) | T) = \sum_{i=1}^m \ln p(d_i(T))$$

其中  $p(d)$  為一常態分佈(Normal Distribution)。

- 匹配的精神就是找一個  $T$  使得上式有最大值。

# 9.4 匹配演算法原理

## 9.4.1 動態規劃式的BSSC解法

- **BSSC**(Banded String-to-String Correction)之目的在樣本和正本之間找出最匹配的對應點。

例子：

樣本 =  $P_1 P_2 P_3 P_4 P_5 \dots P_8 P_9 P_{10}$

正本 =  $T_1 T_2 T_3 T_4 T_5 \dots T_8 T_9 T_{10}$

將  $P_j$  和  $T_i$  想像成特徵值。因為視差  $\Delta$  的關係， $P_j$  只能與  $T_{i \pm \Delta}$  範圍內的特徵進行匹配。

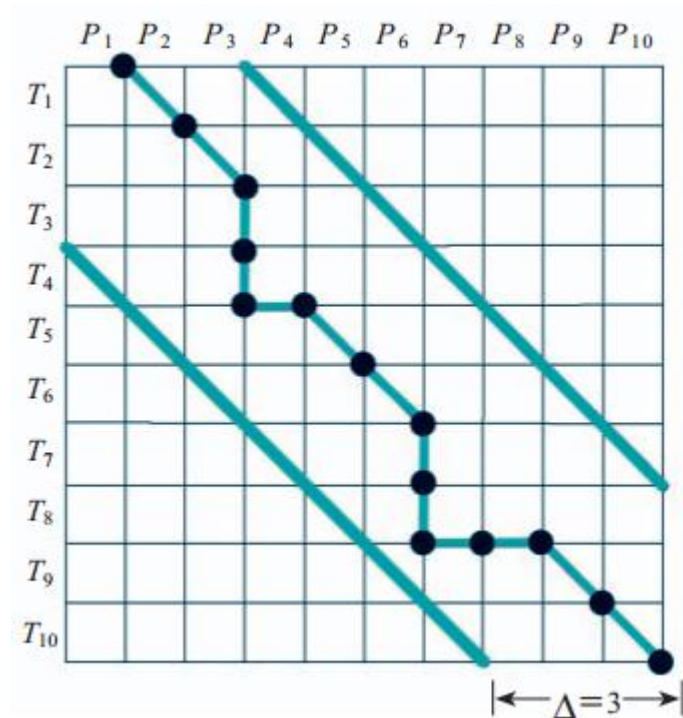


圖9.4.1.1 匹配結果

## ■ 匹配算子

取代算子  $R \rightarrow \begin{cases} a \neq b, R(a)=b \text{ 的花費定為 } 1 \\ a = b, R(a)=b \text{ 的花費定為 } 0 \end{cases}$

刪除算子  $D \rightarrow D(a)=\wedge$  的花費定為 1

插入算子  $I \rightarrow I(\wedge)=a$  的花費定為 1

圖9.4.1.1 一共含有下列十四個運算

- (1)  $I(\wedge)=P_1$    (2)  $R(T_1)=P_2$    (3)  $R(T_2)=P_3$    (4)  $D(T_3)=\wedge$    (5)  $D(T_4)=\wedge$   
(6)  $I(\wedge)=P_4$    (7)  $R(T_5)=P_5$    (8)  $R(T_6)=P_6$    (9)  $D(T_7)=\wedge$    (10)  $D(T_8)=\wedge$   
(11)  $I(\wedge)=P_7$    (12)  $I(\wedge)=P_8$    (13)  $R(T_9)=P_9$    (14)  $R(T_{10})=P_{10}$

- 動態規劃核心式子

$$\text{Edit}[i, j] = \min(\text{Edit}[i-1, j] + \text{edit}(T_i, \wedge), \\ \text{Edit}[i, j-1] + \text{edit}(\wedge, P_j), \text{Edit}[i-1, j-1] + \text{edit}(T_i, P_j))$$

表示將  $T[1 \dots i]$  轉換成  $P[1 \dots j]$  的花費。

$$T[1 \dots i] = T_1 T_2 \dots T_i \quad P[1 \dots j] = P_1 P_2 \dots P_j \quad \text{Edit}[0, k] = k \quad \text{Edit}[k, 0] = k$$

$\text{edit}(T_i, P_j)$  表示  $R(T_i) = P_j$  的花費

$\text{edit}(T_i, \wedge)$  表示  $D(T_i) = \wedge$  的花費

$\text{edit}(\wedge, P_j)$  表示  $I(\wedge) = P_j$  的花費

- BSSC 的時間複雜度為  $O(n\Delta)$

$n$  為正本長度， $\Delta$  為視差



## 9.4.2 KMP演算法

- 樣本先進行事前處理

利用樣本中子字串(Substring)與前置字串(Prefix String)的吻合度，並記錄其吻合的長度於陣列  $J[]$  中。

$$P[3]=a=P[1] \rightarrow J[3]=1$$

$$P[3\dots5]=aca=P[1\dots3] \rightarrow J[5]=3$$

$$P[7\dots10]=acac=P[1\dots4] \rightarrow J[10]=4$$

$i$	1	2	3	4	5	6	7	8	9	10
$P[i]$	$a$	$c$	$a$	$c$	$a$	$a$	$a$	$c$	$a$	$c$
$J[i]$	0	0	1	2	3	1	1	2	3	4

- KMP字串匹配演算法

$$T[] = cccacacaaacaccaa$$

$T[i] \neq P[i]$ ， $1 \leq i \leq 3$ 。  $T[4\dots13] = P[1\dots10]$ 。

$T[5] \neq P[1]$ ， $T[6] = P[1] \rightarrow T[6\dots8] = P[1\dots3]$

$T[9] \neq P[4]$ ， $J[4]=2 \rightarrow P[1\dots J[4]-1]=P[1]=T[9-J[4]+1\dots8]=T[8]$

陣列  $J[]$  提供了一個跳躍的機制，讓匹配動作可一直往右前進。

- KMP的時間複雜度為  $O(m + n)$

# 9.5 三維影像重建

## 9.5.1 稠密式視差估測

### (Dense Disparity Estimation)

- 問題定義

稠密式視差估測是在L上的所有像素中和R上的所有像素中找到一個對應。

- 二階段式的分割與克服方法

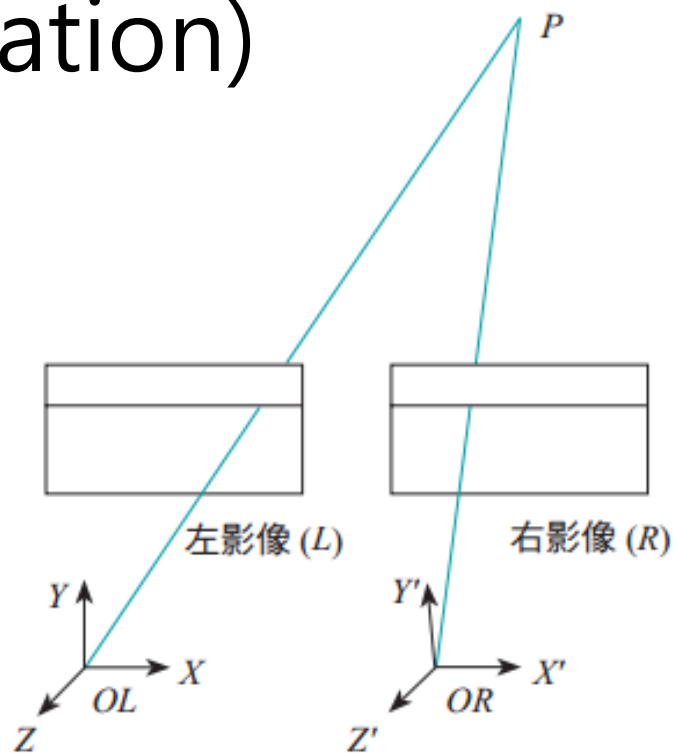


圖 9.5.1.1 稠密視差估測示意圖

## ■ 第一階段的分割與克服方法

1. 利用測邊法得到該列的主要邊點集或較具特徵的點集合：  
L上第  $N/2$  列中的主要特徵集可表示為：

$$S_L^{N/2} = S_L^{N/2}(1)S_L^{N/2}(2)\cdots S_L^{N/2}(N_{N/2}) \quad , \quad 1 \leq N_{N/2} \leq N$$

R上第  $N/2$  列中的主要特徵集可表示為：

$$S_R^{N/2} = S_R^{N/2}(1)S_R^{N/2}(2)\cdots S_R^{N/2}(N_{N/2}) \quad , \quad 1 \leq N_{N/2} \leq N$$

2.  $S_L^{N/2}$  和  $S_R^{N/2}$  的匹配工作

$$C(S_L^{N/2}(v), S_R^{N/2}(w)) = \min\left(C_{match}(S_L^{N/2}(v), S_R^{N/2}(w)), C_{rightocc}(S_L^{N/2}(v), S_R^{N/2}(w)), C_{leftocc}(S_L^{N/2}(v), S_R^{N/2}(w))\right)$$

$$\text{其中 } C_{leftocc}(S_L^{N/2}(v), S_R^{N/2}(w)) = C(S_L^{N/2}(v), S_R^{N/2}(w-1)) + C_{Occ}$$

$$C_{rightocc}(S_L^{N/2}(v), S_R^{N/2}(w)) = C(S_L^{N/2}(v-1), S_R^{N/2}(w)) + C_{Occ}$$

$$C_{match}(S_L^{N/2}(v), S_R^{N/2}(w)) = C(S_L^{N/2}(v-1), S_R^{N/2}(w-1)) + C_{S_L^{N/2}(v), S_R^{N/2}(w)}$$

取代算子：

$$C_{S_L^{N/2}(v), S_R^{N/2}(w)} = \sigma^2 * |l_1 - l_2|$$

其中

$$l_1 = S_L^{N/2}(v) - S_L^{N/2}(v-1)$$

$$l_2 = S_R^{N/2}(w) - S_R^{N/2}(w-1)$$

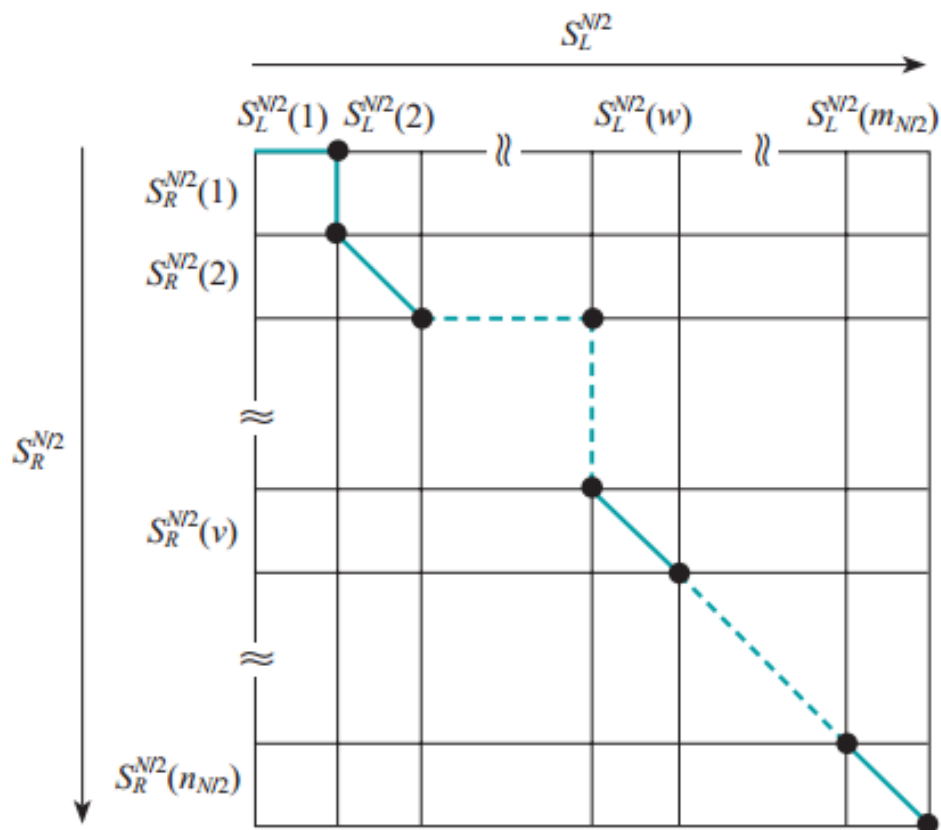
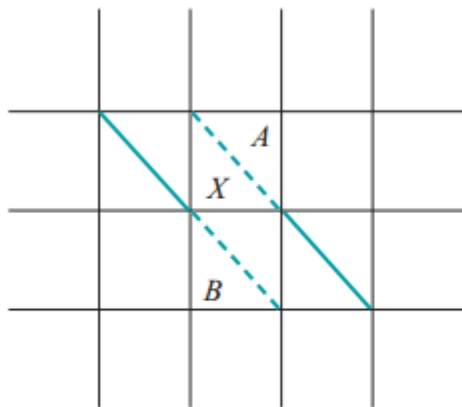


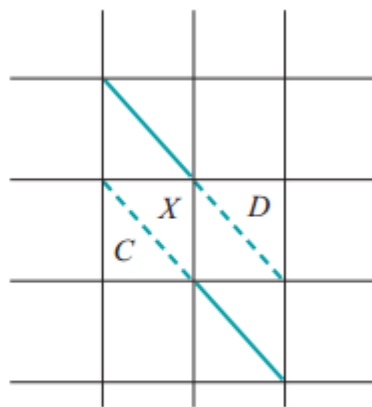
圖 9.5.1.2  $S_L^{N/2}$  和  $S_R^{N/2}$  的匹配

$$\text{mean} = \frac{1}{2} \left( \frac{1}{l_1} \sum_{p=1}^{l_1} I_L(S_L^{N/2}(v-1) + p) + \frac{1}{l_2} \sum_{q=1}^{l_2} I_R(S_R^{N/2}(w-1) + q) \right)$$

$$\sigma^2 = \frac{1}{2} \left( \frac{1}{l_1} \sum_{p=1}^{l_1} \left( I_L(S_L^{N/2}(v-1) + p) - \text{mean} \right)^2 + \frac{1}{l_2} \sum_{q=1}^{l_2} \left( I_R(S_R^{N/2}(w-1) + q) - \text{mean} \right)^2 \right)$$

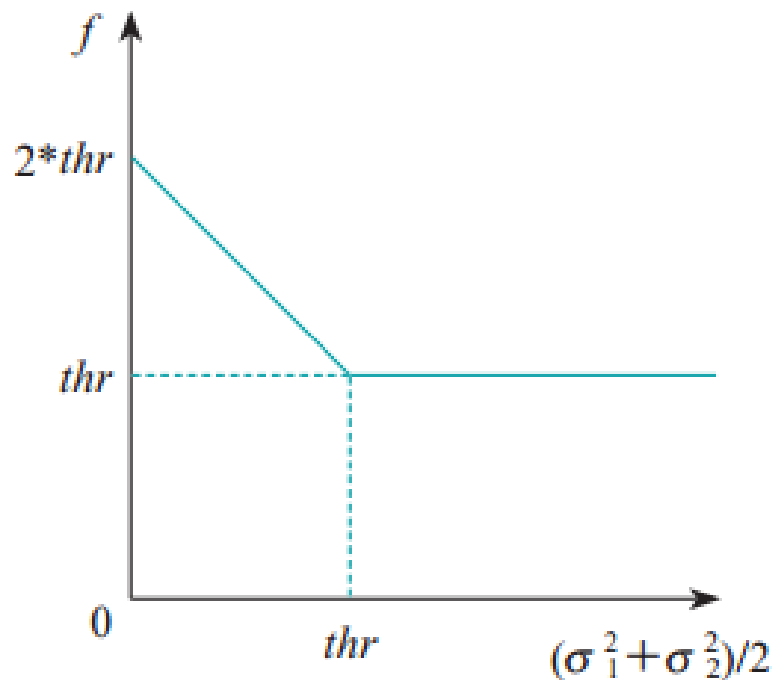


(a) 右遮蔽花費



(b) 左遮蔽花費

圖 9.5.1.3 右遮蔽花費和左遮蔽花費示意圖



$$C_{occ} = k \times f\left(\left(\sigma_1^2 + \sigma_2^2\right) / 2; thr\right)$$

### 3. 加入消去法則

若兩兩匹配位置的差，形成之序列為  $\langle 5, -2, 5, 6, 5 \rangle$ ，則  $-2$  的視差偏移不太正常，可以予以去除。假設在  $N/2 \pm k$  列時的平均視差偏移序列為  $\langle 5, 5, 6, 4 \rangle$ 。若新進來的視差偏移序列為  $\langle 5, 7, 3, 16, 4 \rangle$ ，則  $16$  的視差偏移可予以去除。

## ■ 第二階段的分割與克服方法

1. 在第二階段中進一步找出個別的像素配對

2. 利用平均視差偏移量  $d$  ,

可將第二階段的工作轉

成BSSC問題的解決上

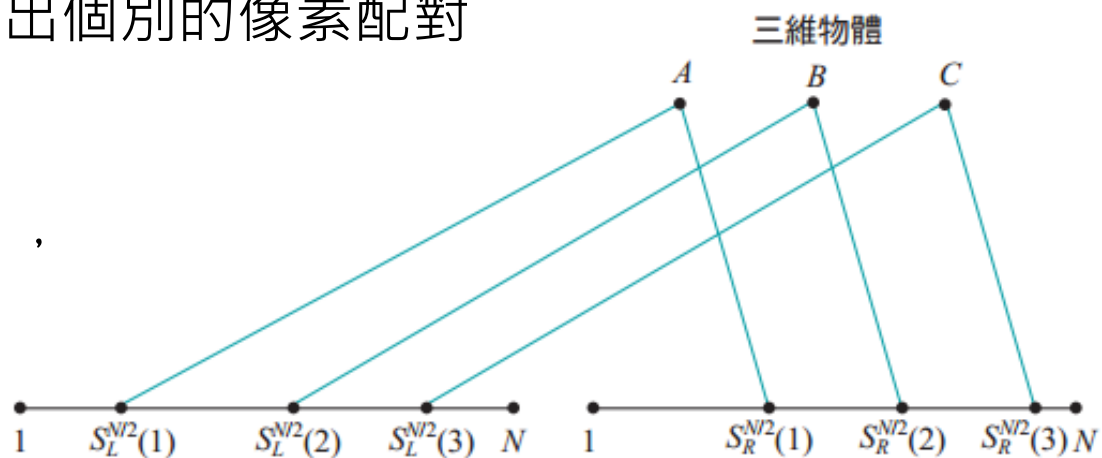


圖9.5.14 二子區間匹配

$R$ 算子： $C(S_L^{N/2}(v) + \hat{v}, S_R^{N/2}(w) + \hat{w}) =$

$$\sum_{x,y=-W/2}^{W/2} \left| I_L(N/2+x, S_L^{N/2}(v) + \hat{v} + y) - I_R(N/2+x, S_R^{N/2}(w) + \hat{w} + y) \right|$$

$W$ 為事前定義好的小視窗

$D$ 算子：類似於前面定義的左遮蔽花費

$I$ 算子：類似於前面定義的右遮蔽花費

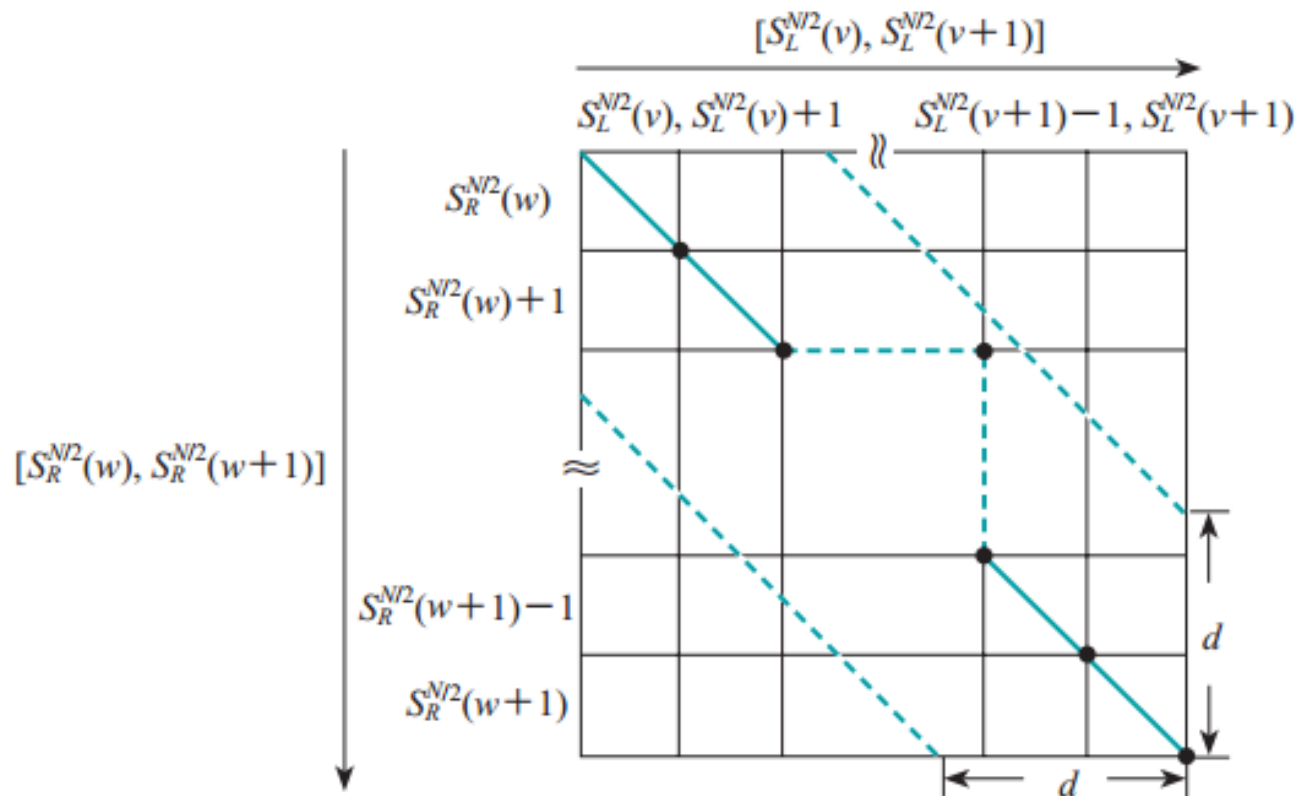


圖9.5.1.5 對應的BBSC搜尋空間



(a) 影像 $L$



(b) 影像 $R$

圖9.5.1.6 輸入的二張影像

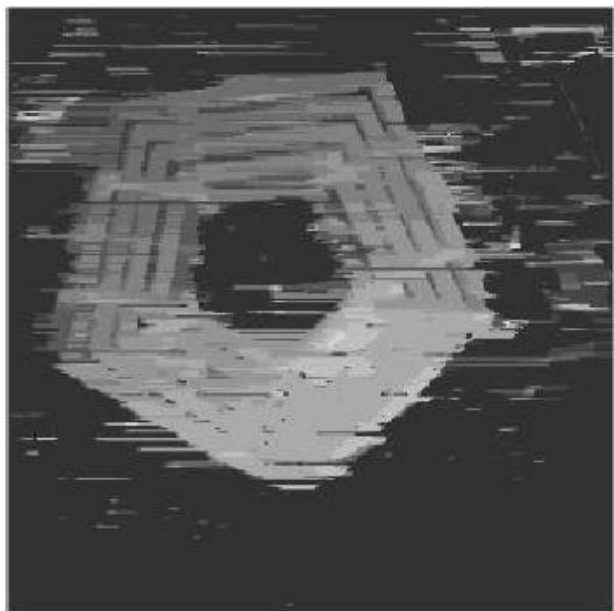


圖9.5.1.7  
得到的視差圖

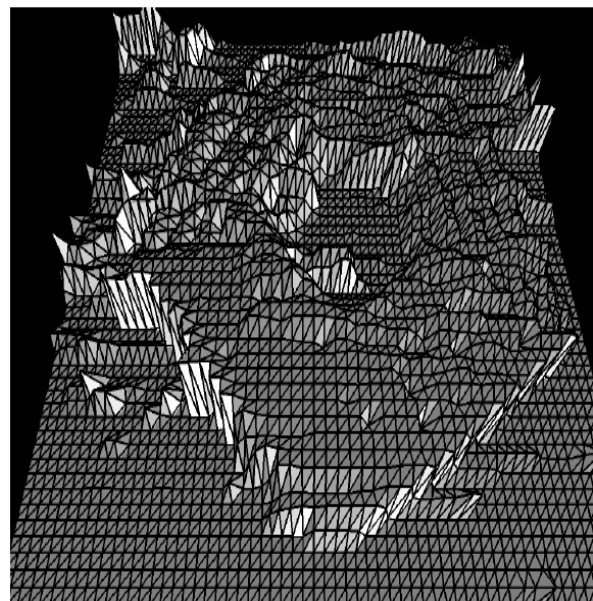


圖9.5.1.8  
重建後的三維  
五角大廈圖



## 9.5.2 相機校正

- 我們有三個座標系統，**世界座標系統**(World Coordinate System)、**相機座標系統**(Camera Coordinates System)和二維的**影像系統**(Image System)。

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T \quad R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \quad T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

根據三角比例關係，可得

$$X_t = f \frac{x}{z} \quad Y_t = f \frac{y}{z}$$

其中  $f$  為焦距長。

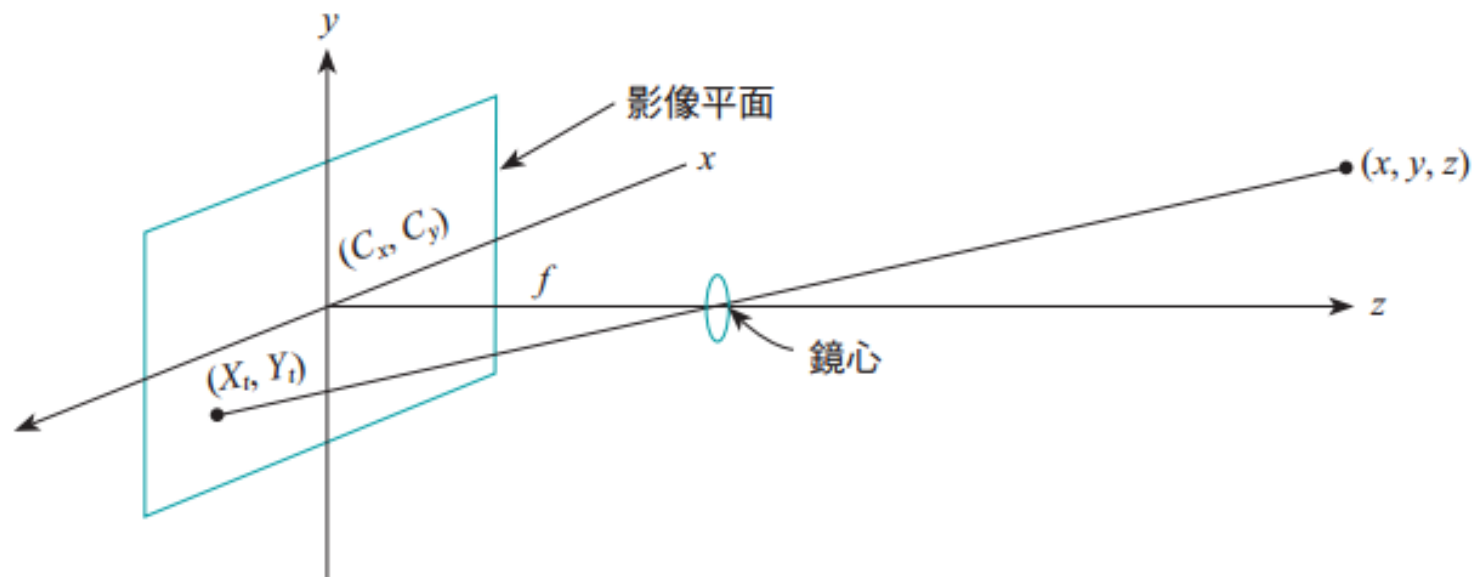


圖 9.5.2.1 相機座標系統和影像座標系統

理論上 $(X_t, Y_t)$ 為 $(x, y)$ 投射到影像座標系統的理想位置。  
 因為透鏡輻射效應， $(x, y)$ 實際投射到的影像座標系統的位置應  
 為 $(X_a, Y_a)$ 。滿足下式

$$X_a + E_x = X_t$$

$$Y_a + E_y = Y_t$$

上兩式中的誤差項可表示成

$$E_x = X_a(k_1 r^2 + k_2 r^4)$$

$$E_y = Y_a(k_1 r^2 + k_2 r^4)$$

$$r = \sqrt{X_a^2 + Y_a^2}$$

在光學透鏡的**輻射失真**(Radial Distortion)的影響下，內部參數滿足  $k_1 \ll k_2$ ，我們令  $k_2 = 0$ ，只關心  $k_1$  的求解。

前面所提的  $(X_a, Y_a)$  為實數，但在數位影像的座標系統一般皆為整數座標  $(X_I, Y_I)$ 。假設**感應器**(Sensor)的  $X(Y)$  方向共有  $S_x(S_y)$  個感應器。 $N_x$  是影像在  $x$  方向的解析度。令

$$d'_x = d_x \frac{S_x}{N_x}$$

而 $(C_x, C_y)$ 為影像的中心座標。 $(X_a, Y_a)$ 和 $(X_I, Y_I)$ 的關係可表示如下

$$X_I = C_x + s \frac{X_a}{d'_x} \quad Y_I = C_y + \frac{Y_a}{d_y}$$

其中  $s$  為待解的放大係數。

已知 $(x, y)$ 投影到影像座標系統的理想位置為 $(X_I, Y_I) = \left(f \frac{x}{z}, f \frac{y}{z}\right)$ 。

進而得到

$$\begin{aligned} f \frac{x}{z} &= X_a + E_x = \frac{(X_I - C_x)d'_x}{s} + E_x \\ &= \frac{(X_I - C_x)d'_x}{s} + \frac{(X_I - C_x)d'_x k_1 r^2}{s} \\ &= \frac{\bar{X}d'_x}{s} + \frac{\bar{X}d'_x k_1 r^2}{s} \end{aligned}$$

同理可得

$$d_y \bar{Y} + d_y \bar{Y} k_1 r^2 = f \frac{y}{z}$$

且

$$r = \sqrt{\left(\frac{\bar{X}d'_x}{s}\right)^2 + [d_y (Y_I - C_y)]^2} = \sqrt{\left(\frac{\bar{X}d'_x}{s}\right)^2 + (d_y \bar{Y})^2}$$

再利用三維世界座標系統和二維影像座標系統的關係可得

$$\frac{\bar{X}d'_x}{s} + \frac{\bar{X}d'_x k_1 r^2}{s} = f \frac{r_1 x_w + r_2 y_w + r_3 z_w + T_x}{r_7 x_w + r_8 y_w + r_9 z_w + T_z}$$

$$d_y \bar{Y} + d_y \bar{Y} k_1 r^2 = f \frac{r_4 x_w + r_5 y_w + r_6 z_w + T_y}{r_7 x_w + r_8 y_w + r_9 z_w + T_z}$$

旋轉矩陣  $R$  分別對  $x$  軸、 $y$  軸和  $z$  軸達到任一角度的旋轉，可寫成下式

$$R = \begin{bmatrix} \cos \psi \cos \theta & \sin \psi \cos \theta & -\sin \theta \\ -\sin \psi \cos \phi + \cos \psi \sin \theta \cos \phi & \cos \psi \cos \phi & \cos \theta \sin \phi \\ \sin \psi \sin \phi + \cos \psi \sin \theta \cos \phi & -\cos \psi \sin \phi + \sin \psi \sin \theta \cos \phi & \cos \theta \cos \phi \end{bmatrix}$$

至此，已經六個外部參數(Extrinsic Parameters)分別為 $\psi$ 、 $\theta$ 、 $\phi$ 、 $T_x$ 、 $T_y$ 和 $T_z$ 及五個內部參數(Intrinsic Parameters)分別為 $f$ 、 $s$ 、 $k_1$ 、 $C_x$ 和 $C_y$ 。接下來要利用數值的方法求解這些參數。

先解出五個外部參數，在二維影像座標上取一經過原點的向量  $(X_a, Y_a)$ ，且在相機座標上取一平行的向量  $(x, y)$ ，所以得

$$(X_a, Y_a) \times (x, y) = 0$$

$$X_a \cdot y - Y_a \cdot x = 0$$

$$X_a(r_4 x_w + r_5 y_w + r_6 z_w + T_y) = Y_a(r_1 x_w + r_2 y_w + r_3 z_w + T_x)$$

令  $Z_w = 0$ ，則我們得到

$$[Y_a x_w \quad Y_a y_w \quad Y_a \quad -X_a x_w \quad -X_a y_w] \begin{bmatrix} T_y^{-1} r_1 \\ T_y^{-1} r_2 \\ T_y^{-1} r_3 \\ T_y^{-1} r_4 \\ T_y^{-1} r_5 \end{bmatrix} = X_a$$

但是只要利用五組世界座標  $(x_w, y_w, z_w)$  和五組底片上的真實座標  $(X_a, Y_a)$  就可解出  $T_y^{-1} r_1$ 、 $T_y^{-1} r_2$ 、 $T_y^{-1} r_3$ 、 $T_y^{-1} r_4$ 、 $T_y^{-1} r_5$ 。

令

$$C = \begin{bmatrix} \bar{r}_1 & \bar{r}_2 \\ \bar{r}_4 & \bar{r}_5 \end{bmatrix} = \begin{bmatrix} T_y^{-1} r_1 & T_y^{-1} r_2 \\ T_y^{-1} r_4 & T_y^{-1} r_5 \end{bmatrix}$$

則旋轉矩陣可改寫成

$$R = \begin{bmatrix} \bar{r}_1 T_y & \bar{r}_2 T_y & r_3 \\ \bar{r}_4 T_y & \bar{r}_5 T_y & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix}$$

利用  $R$  中每一行向量與列向量為單位長，可得下式

$$R = \begin{bmatrix} \bar{r}_1 T_y & \bar{r}_2 T_y & \pm \left[ 1 - T_y^2 (\bar{r}_1^2 + \bar{r}_2^2) \right]^{1/2} \\ \bar{r}_4 T_y & \bar{r}_5 T_y & \pm \left[ 1 - T_y^2 (\bar{r}_4^2 + \bar{r}_5^2) \right]^{1/2} \\ \pm \left[ 1 - T_y^2 (\bar{r}_1^2 + \bar{r}_4^2) \right]^{1/2} & \pm \left[ 1 - T_y^2 (\bar{r}_2^2 + \bar{r}_5^2) \right]^{1/2} & \pm (-1 + M T_y^2)^{1/2} \end{bmatrix}$$

其中  $M = \bar{r}_1^2 + \bar{r}_2^2 + \bar{r}_4^2 + \bar{r}_5^2$ 。

因為  $R$  為正交矩陣，任兩行的內積必為零。將  $R$  的第一行和第二行做內積，可以得到

$$(\bar{r}_1\bar{r}_5 + \bar{r}_2\bar{r}_4)^2 T_y^4 - M T_y^2 + 1 = 0$$

此等式可解得  $T_y^2$  的兩個解，如下

$$T_y^2 = \frac{M \pm [M^2 - 4(\bar{r}_1\bar{r}_5 + \bar{r}_2\bar{r}_4)^2]^{\frac{1}{2}}}{2(\bar{r}_1\bar{r}_5 + \bar{r}_2\bar{r}_4)^2}$$

前面曾推導過

$$d_y \bar{Y} + d_y \bar{Y} k_1 r^2 = f \frac{r_4 x_w + r_5 y_w + r_6 z_w + T_y}{r_7 x_w + r_8 y_w + r_9 z_w + T_z}$$

暫時將  $k_1$  和  $Z_w$  設為零，得

$$d_y \bar{Y} = f \frac{r_4 x_w + r_5 y_w + T_y}{r_7 x_w + r_8 y_w + T_z}$$



令  $t_1 = r_4 x_w + r_5 y_w + T_y$  且  $[t_1 - d_y \bar{Y}] \begin{bmatrix} f \\ T_z \end{bmatrix} = t_2 d_y \bar{Y}$  上式可改寫成

$$\begin{bmatrix} t_1 - d_y \bar{Y} \end{bmatrix} \begin{bmatrix} f \\ T_z \end{bmatrix} = t_2 d_y \bar{Y}$$

再利用兩組以上的世界座標  $t_1 = r_4 x_w + r_5 y_w + T_y$  和真實座標  $(X_a, X_a)$

解出  $t_1 = r_4 x_w + r_5 y_w + T_y$  和  $(x_w, y_w, z_w)$ 。然後以此  $t_1 = r_4 x_w + r_5 y_w + T_y$  和  $(x_w, y_w, z_w)$  及令  $f$  為初始值來解線性方程式

$$T_z$$

如此即可透過數值的解法求得  $\kappa_1 = 0$  的逼近解。

# 9.6 二維影像的深度計算

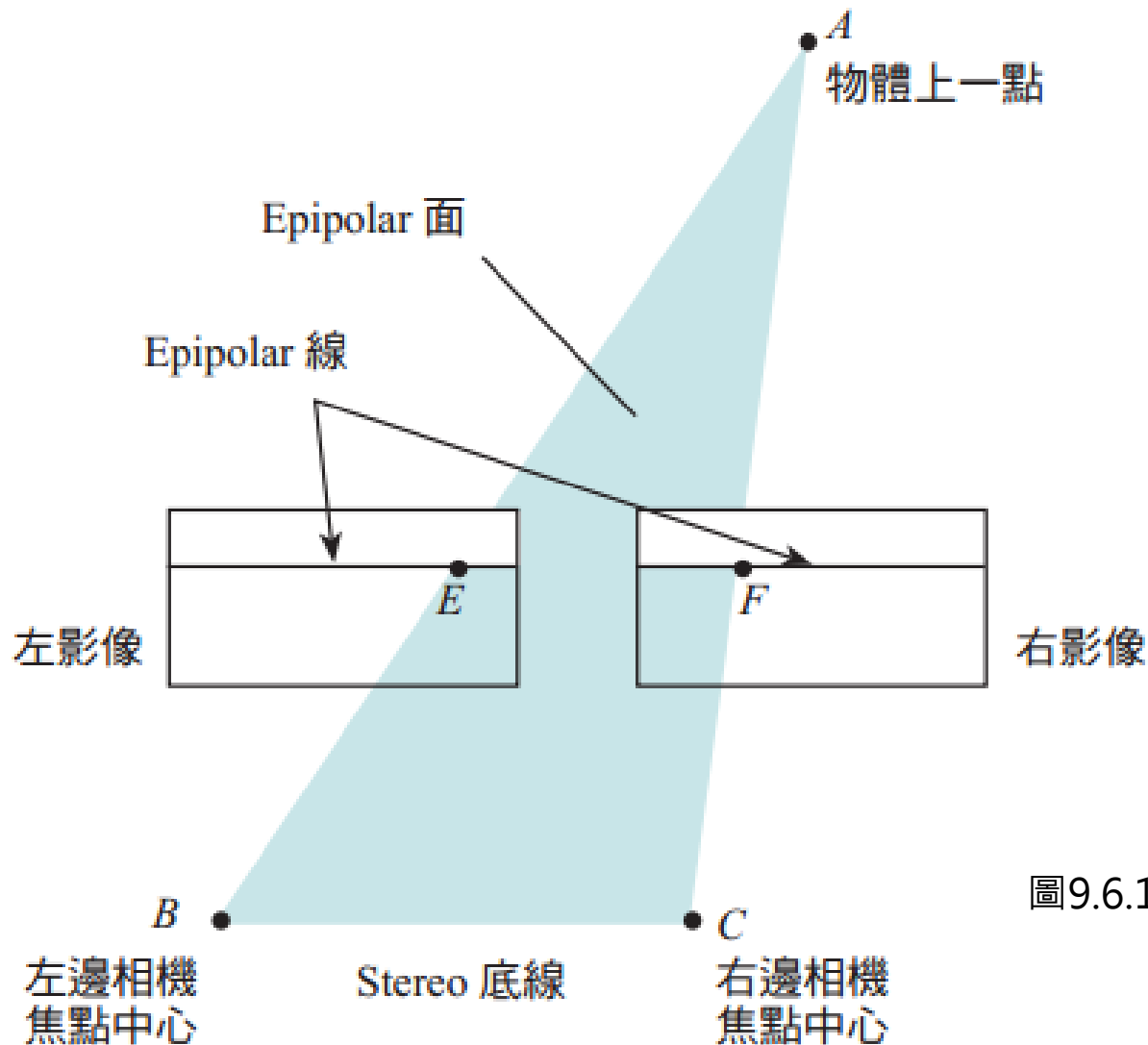


圖9.6.1 Epipolar面

範例1:已知兩部相機有相同的焦距 $f$ ，假設左邊的Epipolar線之長度為 $x_l$ 而右邊的Epipolar線之長度為 $x_r$ 。已知物件上的一點 $A(x, y, z)$ ，試問 $z$ 之值為何？

解答:

$\Delta BFG \approx \Delta BDA$  和  $\Delta CHI \approx \Delta CEA$ ，透過兩相似三角形的邊比例關係，可得到

$$\frac{x}{z} = \frac{x_l}{f}$$

$$\frac{x - b}{z} = \frac{x_r}{f}$$

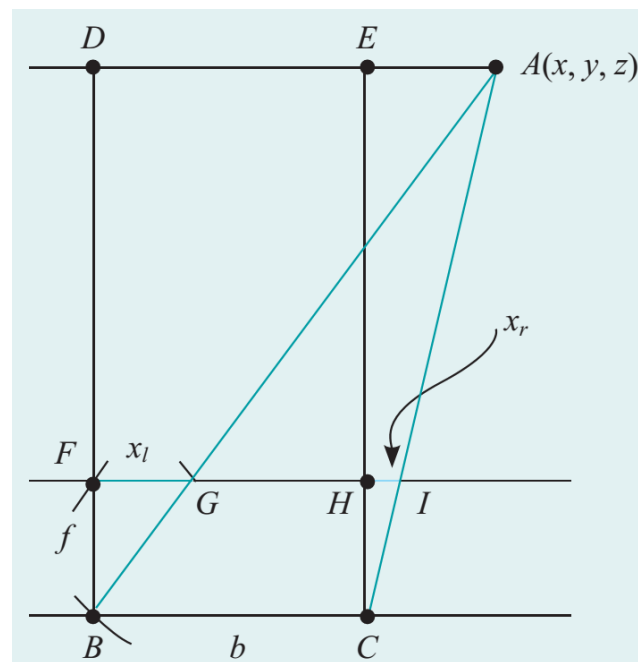


圖9.6.2 利用兩平行相機求 $z$ 值

在上式中，我們假設兩部相機相距 $b$ ，座標系統的原點設在 $B$ 點。可進一步推得

$$x = \frac{zx_l}{f}$$
$$x = (zx_r + bf)/f$$

再利用 $zx_l = zx_r + bf$ ，可得到深度為

$$z = \frac{bf}{x_l - x_r}$$

解答完畢