

TR-88-022

從兩張折紙狀物體的二維影像決定其
可見部份之運動參數

參 考 書
不 外 借

中研院資訊所圖書室



3 0330 03 000095 9

0095

題目：從兩張折紙狀物體的二維影像決定其可見部份之運動參數
(Determination of Motion Parameters of
Visible Portion Objects in Origami World
from Two 2-D Images)

研究生：蔡 鴻 旭

指導教授：鄭 國 揚 博士

中 華 民 國 七 十 七 年 五 月 日

目 錄

摘要	i
第一章 緒論	1
第二章 影像模式	9
第一節 成像原理與照相機模式	10
第二節 投影	14
第三節 照相機與實物間的幾何關係	17
第四節 數位影像之形成及前置處理	26
第五節 剛體運動	30
第六節 立體視覺	35
第三章 二維影像之解譯	40
第一節 名詞定義	41
第二節 斜量空間簡介	46
第三節 折紙狀物體的解譯影像之形成	54
第四節 部份被遮面發生的原因及限制	66
第四章 物體運動參數之決定	75
第一節 找出三維平面的法向量	77
第二節 運動參數的決定	87
第三節 模擬實驗及其結果	97
第四節 照相機光軸校正	107
第五章 結論	109
參考文獻	112

摘 要

「如何從兩張二維影像來決定出影像中物體之運動參數」一直是電腦視覺學之研究領域中令人值得去探討的有趣問題。然而，在解決此問題時，複雜的相對問題乃伴隨而至。所以在此論文中，我們將相對問題於全然未知的情況，簡化成，兩張影像間的可見相對接點不能消失太多，然後去分析出具有對稱面之折紙狀物體的三個運動參數。

首先，將兩張二維影像經數位影像的前置處理後，得到兩張線繪影像。接著，我們利用人工智慧學中的標線方法對這兩張影像標線，然後經由一些啓發式規則及斜量空間之性質，分別找出標線後的兩張影像中全部可見區域之法向量。而後，取兩張影像中所有的法向量來做匹配以測量出物體的旋轉角度。

其次，取出兩張影像中所有的接點來匹配後，又可決定出影像間的可見相對接點及這些相對接點與物體之二平移量所見有的關係式。以此關係式及物體所能平移之已知範圍來尋獲物體的二平移量。至此，物體的三個運動參數就全部被測量出了，並且亦找出了影像間的所有可見相對接點。

最後，我們依據相對點的觀念，提出一個照相機光軸校正的方法。以此方法，僅移動一次轉盤即能使照相機光軸通過轉盤中心而完成了照相機光軸的校正。

第一章 緒論

在電腦視覺學 (Computer Vision) 的研究領域中，如何從連續兩張二維影像中決定剛性物體 (rigid body) 的運動參數 (motion parameters)。這一直是很值得探討的重要問題。尤其是，在自動化工業環境中，如能讓有視覺系統的機械手 (robotic arms) 自動分析兩張影像中物體的運動參數，則機械手將可依據所算出的運動參數去抓取位置異動後的實際物體。如此，將可提高機械手的作業能力。但是在計算運動參數之過程中，總會遭遇到二張影像間的相對問題 (Corresponding problem)。一旦相對問題被決定後，運動參數就可以被計算出。然後再利用立體視覺 (Stereo Vision) 的技術，重建二維影像中的物體之立體形狀。能獲得實際物體的三維資料，不僅可以強化辨認物體的能力，亦有助於電腦繪圖學 (Computer Graphics) 和 CAD 的應用。

可是目前對二張影像間之相對問題的有效解法，仍舊是一項未解的問題。所謂相對問題，可以看成：如何找出兩張影像中的相對點 (Corresponding points)。也就是說，從觀察相同物體 Ω ，所獲得的兩張二維影像中，如果有一張影像中的點 p' 和另一張影像中的點 g' 。它們都是 Ω 上同一點 p 的投影，則 p' ， g' 互稱對方為自己的相對點。同理，決定相對問題亦包括相對線和相對面的尋找。為什麼由兩張二維影像來尋找出其間之相對關係 (Corresponding relation) 的問題仍是目前舉世公認的難題即所謂的公開問題 (open problem) 呢？原因是，決定相對點 (相對線

或相對面) 是否存在於二張影像之間的演繹法 (algorithm) 都具有 Combinatorial explosive 的特性。

由於相對問題是件複雜的問題。因此，對一些在求物體的運動參數之論文所提出的方法中，都儘量不涉及此問題。諸如，Ullman [18] 以三張正交投影 (orthographic projection) 的影像 $I_i, i = 1, 2, 3$ ，未計算剛體運動的六個未知參數 (分別對三個座標軸平移、旋轉)，其中 Ullman 假設三張影像中必須有不共平面的 4 點且已知其間之相對關係，又其中有一點必為座標原點。因此對未知的平移參數而言，僅用這座標原點的三個投影點即可被求出。所以只剩下六個旋轉參數 (因為各用三個不同軸之旋轉角度去獲得影像 I_2, I_3)，和影像 I_1 中其餘三個點之空間坐標 $p_i = (x_i, y_i, z_i)$ ， $i = 1, 2, 3$ 。總共有 15 個未知數。而依剛體運動的特性，每張影像將可有 6 條方程式，所以總共可有 18 條方程式。從這 18 條方程式中即能解出上述的 15 個未知數。此時，運動參數即被求出。而 Huang & Tsai [19] 提出：對透視投影所得的影像裏，如能給予至少 7 組相對點則他們可以估計出剛體運動的運動參數 ($\theta, \Delta x, \Delta y, \Delta z$) 而 θ 為物體針對空間中過座標原點的任一軸之旋轉角度，剩下三個參數為物體分別對三座標軸的平移量。在上述的方法中，都必須假設相對點已知，才來決定物體的運動參數。而 Silberberg [20] 乃事先去建好各物體的三維資料，並且將系統環境限制成，物體運動只能被針對實物坐標系 (World coordinate system) 中的 Z 軸旋轉並且其平移向量為 $D = (\Delta x, \Delta y, 0)$ ，相當於物體只能在平行 X - Y 平面

上平移。然後，對物體拍入單張二維影像，直接用物體已知的三維邊和影像中的二維邊線來匹配 (matching)，算出最有可能的旋轉角度 θ_z ，繼續再用已知的三維頂點 (vertex) 與影像中的二維接點 (junction) 來匹配，而求出可能的平移量， $\Delta x, \Delta y$ 。這是植基於一個 Generalized Hough Transform [7] 的方法，將物體的所有三維邊和三維頂點與影像中所有二維邊線和接點來匹配。最後從曾經算出的可能運動參數之中，選出一組出現次數最高者，此組運動參數即物體的運動參數。但是在這個方法中必須事先建造好物體的三維資料，此乃是唯一令人遺憾的事。

而我們的靈感乃出自於此，我們將不希望事先去建好物體的三維資料。於是乎我們在物體運動前拍入一張二維影像，待物體運動後再拍入另一張二維影像，然後就以這二張影像來分析出物體的運動參數 ($\theta_z, \Delta x, \Delta y$)。但在此篇論文中，我們假設物體被旋轉角度及 X, Y 軸方向的平移量都不超過某個範圍。並且只允許由對稱平面所構成的折紙狀物體，雖不限物體個數，但對不互相接觸物體之遮蔽現象，將限制於部份被遮的情況之下。我們整個系統的流程如圖 1—1 所示。首先，將輸入兩張二維的線繪影像 (line-drawing images)，利用前人在人工智慧學中所提到的標線方式 (line-labeling scheme)，將二維影像加以標線，而標線後的二維線繪影像就稱為解譯影像 (interpretative image)。

然後，再分別對兩張解譯影像，用幾個啟發式規則 (heuristic rules) 將影像中部分被遮區域 (partial occluded region) 找

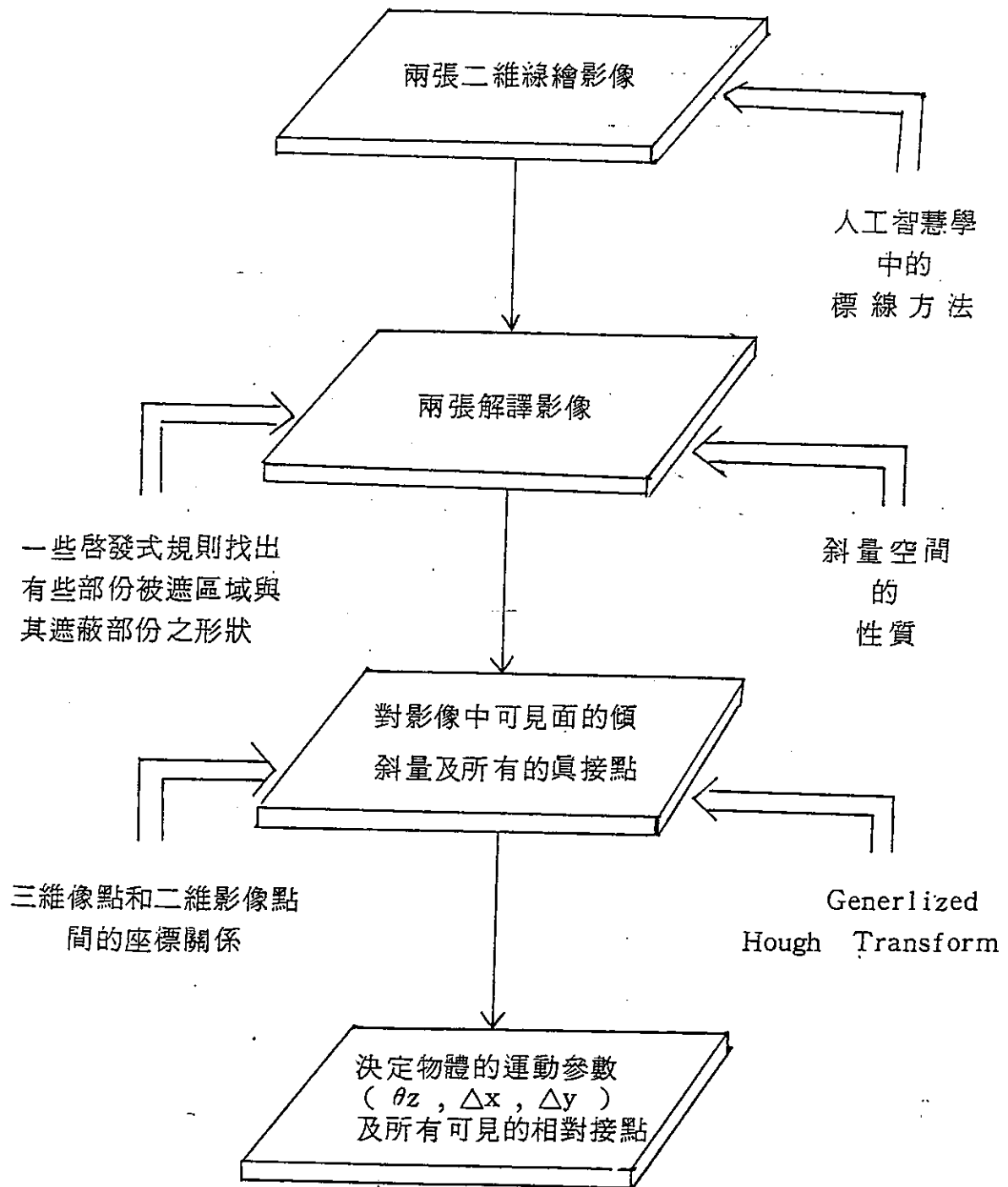


圖 1 — 1 系統流程

出，由此區域並可猜測出其可能的部份被遮形狀。其次，依兩張解譯影像中所有各全部可見區域之形狀與斜量空間（ Gradient Space ）上的特性，計算出這些區域所相對於三維面之傾斜量（ gradient ）。即這些三維面之法向量。

最後，以一個 Generalized Hough Transform 的方法，將第一張解譯影像所計算出各面之法向量，和第二張解譯影像所計算出各面之法向量來匹配。首先，將出現次數最高的旋轉角度挑出一 θ_z 。其次，再利用一次 Generalized Hough Transform 的方法，將第一張影像中所有的真接點和第二張影像中的所有真接點來匹配，而其匹配時，所利用的關係式為景像（ scene ）和影像平面（ image plane ）間，兩者的座標關係式。最後，將得到一條出現次數最高的二元一次線性方程式和所有可見的相對點。而此條方程式乃由二個未知的平移量 Δx ， Δy 所構成。而在此論文中假設物體的平移量在某範圍之內，因此我們將可在此範圍之內，找出一組平移量來符合該線性方程式而其誤差最小者，即是我們所要的。

一般在做影像分析與三維物體的重建或描述的系統環境中。所採用的照相機和三維物體之間的運動關係可以分底下四種類型：

1. 靜止的照相機，靜止的物體（ Stationary Camera, Stationary Object ）。
2. 靜止的照相機，動態的物體（ Stationary Camera, Moving Object ）。
3. 動態的照相機，靜止的物體（ Moving Camera, Station-

ary Object) 。

4. 動態的照相機，動態的物體 (Moving Camera, Moving Object) 。

在這四種類型中，第一種類型及第四種類型比較少被採用。而第二種類型和第三種類型其實是一體兩面的意思。因為照相機和三維物體之間的相對關係不變。在這篇論文中，是採用第二種類型。

在這篇論文中其餘的內容將分成四章來討論。

首先在第二章中將介紹成像原理以及此論文中所用的照相機模式和兩者間的關係。由於我們將利用景像中的三維像點 (scene points) 和其投影於影像平面上的影像點 (image points) ，兩者之間的座標關係，來做為兩張二維影像間相對點之決定的工具。因而，我們要瞭解景像中所處的座標系與影像平面所處的座標系，兩者之間座標系變換的關係及以何種方式的投影。我們知道，對於數位電腦來說，它所要求的輸入資料必需要為數位資料才行。因此，從照相機拍攝到的二維影像必得先經過數位化處理後才能輸入電腦加以處理。在這一章中，我們亦簡單的敘述二維影像之數位化過程，此外並將略述二維影像的前置處理，也就是說，從二維影像中粹取出接點、邊線。在此文中乃是採用「靜止的照相機，動態的物體」模式。即固定照相機而旋轉、平移三維物體，就等於說三維物體在做剛性運動，因此在這一章中，亦將介紹剛體運動及其特性。最後我們將描述一般常用的立體視覺 (stereo vision) 系統之種類，及其應具備的已知條件之下，才能

將兩張二維影像還原成實際物體之立體影像。

其次在第三章中我們將討論折紙狀物體的特性。由於這類物體的形狀比三角型平面多面體 (trihedral planar surface object) 的形狀還多，因此，以人工智慧學的標線方法及技巧來解釋其所形成的二維影像時，還必須藉由斜量空間 (Gradient Space) 中的性質來過濾一些令人難以想像的奇異物體形狀之不合理解譯影像發生。因此在本章中，亦將簡介斜量空間中的一些常用的性質。而最後我們將利用幾個啓發式規則來找出物體間所發生的部分被遮區域，及僅可能的猜測出這些被遮部份的形狀。

在第四章中乃以兩張解譯影像中的全部可見區域之形狀來算出其傾斜量 (gradient)。也就是說，這些區域相對於影像中的三維面之傾斜量，相當於三維面的法向量。我們乃以一張影像中的所有法向量和另一張影像中的法向量來匹配 (matching)。然後挑出一個出現次數最多的旋轉角度 θ_z ，此即是我們所欲測量出的物體旋轉角度。繼續，我們取一張影像中的所有真接點 (T型接點除外) 和另一張影像中的所有真接點來匹配，如此可得到所有可見的相對點，以及這些相對點和物體的二個平移量所形成的一條關係式。在知道物體的二平移量應不超過某個範圍之內的假設條件下，我們即可找出一組最有可能的平移量 Δx ， Δy 。至此，我們即可找出物體的運動參數了。接著我們將說明一些模擬實驗的方式及其結果。最後我們知道系統校正 (system calibration) 是件重要的工作。所以在本章的最後一節，我們將提出一個可增快 Tsai & Chen [26] 所提出的照相機光軸校正方法。

最後，在第五章中我們將做一個結論。

第二章 影像模式

在此文中主要是用兩張二維數位影像來決定物體的運動參數。由於我們未知二張影像中的相對點，因此我們乃用景像中的三維像點與其投影於二維影像平面的影像點，兩者間的座標關係來尋獲相對點。而在此文中大都是粹取影像中物體的特徵點 (feature points) 來當做是我們所要決定出的相對點。因此如何以數位影像處理的技術來粹取出影像中物體的真正特徵點是非常重要的。

在這一章中共分六節來介紹。第一節中是以成像原理來說明此文中所採用的照相機模式。接著在第二節中我們定義了兩種投影方式。最先定義透視投影 (perspective projection)，然後乃根據透視投影和另一種平行投影間的基本差異，並以微積分學中的極限觀念來定義出正交平行投影 (orthographic parallel projection)。緊接著在第三節中首先介紹此文所假設的系統環境。在該系統中有(1)一個可控制的轉盤 (turntable)，(2)一台靜止照相機，(3)一座發光系統，(4)一部電腦。然後乃定義轉盤上物體所處的一個實物座標系 (world coordinate system) 和另一個照相機所處的照相機座標系 (camera coordinate system) 以及在照相機座標系中的影像座標系 (image coordinate system)，並敘述三者間的幾何位置關係及座標系變換關係。在此論文中我們假設整個系統校正 (system calibration) 已經完成。緊接著第四節是介紹數位影像之形成。一般由照相機所感測到的影像資料，必須經數位化後，始能成爲一張數位影像，此時方能輸入數位電

腦來做處理。但，通常影像形成過程中免不了會受到外界因素的影響，以致會有雜訊（noise）產生於影像中。因此我們必須藉助影像處理的技術來消除雜訊。在第五節中是介紹剛體運動及其特性。最後在第六節中介紹立體視覺（Stereo vision）系統中常用的兩種方式。有一種方式是由兩台照相機所構成的立體視覺系統。另一種方式是僅由一台照相機所構成的立體視覺系統。若在照相機本身的條件都一樣的情況下，兩者間的不同為：後者必須分析出兩張影像中物體的運動參數，才能在兩張影像的相對點，已知的條件之下，求出影像中物體的實際立體影像。

第一節 成像原理與照相機模式

一般的成像設備都必須要具有二個重要部份：第一部份為一組透鏡，另一部份為影像平面（image plane）。當光線照射於二維物體的表面後，其反射光在經過透鏡中心後，將三維物體的影像顯現於這類的影像平面上。通常這類的影像平面僅是一個二維座標系所構成的平面。其間的成像幾何關係恰可由照相機的模式來表示如圖 2—1 所示。因為一般照相機有兩個重要部份：第一部份為透鏡，另一部為感光軟片。當我們在拍照時必須調整照相機的焦距，以取得最清晰的影像，而所調整的焦距量，三維物體及二維影像平面三者間的關係，乃是根據著名的光學法則：

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$$

所得到的，如圖 2—2，其中 p 是透鏡到三維物體之間的距離，

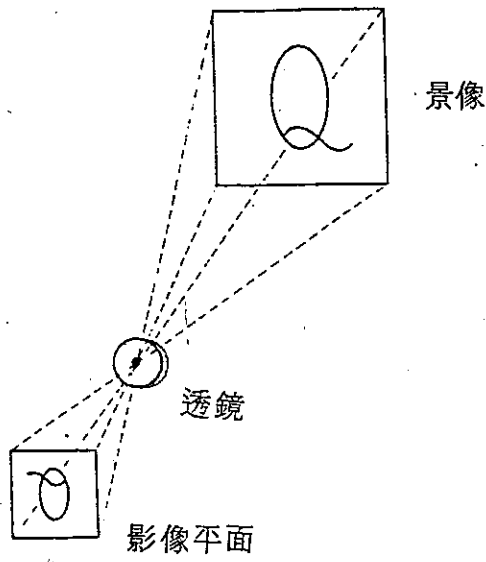


圖 2-1 照相機模式

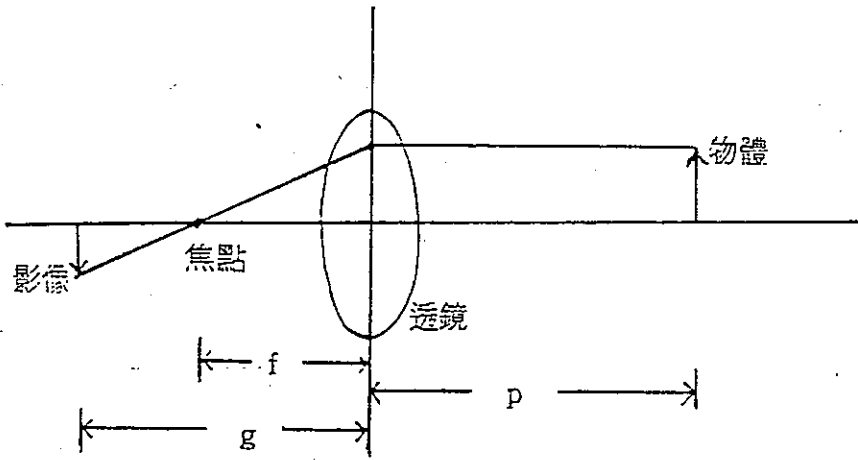


圖 2-2 三維物體和二維影像以及焦距間的幾何關係

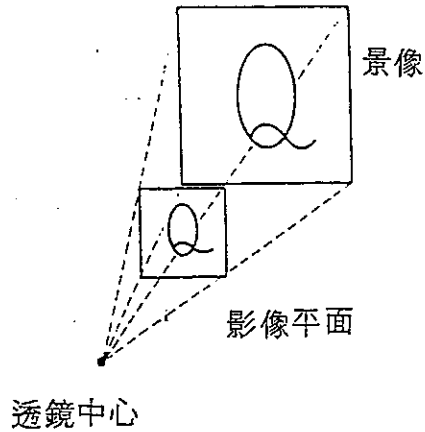


圖 2-3 修正過的照相機模式

q 為二維影像平面到透鏡的距離，而 f 為照相機的焦距。絕大部份的情況下， f 的值遠小於 q 的值。則由上式可以知道 p 幾乎等於 q ，在這種情況下就與剛才所描述的成像設備所必需具有的特性相符合。而針孔照相機在這種成像幾何的模式下，其影像平面所形成的二維影像和其三維物體的形狀上、下相反，且左、右顛倒。但是對我們的眼睛來說，我們可以想像二維影像平面可置於透鏡和三維物體之間，如圖 2—3 所示的情況，並且可將成像設備內的透鏡中心（或是照相機中心的針孔）當成是我們的視點（viewpoint）。

在此篇論文中，我們假設照相機是靜止於固定位置上，現在為了方便解釋起見，我們將影像平面假設和 $X - Y$ 平面平行，並置於原點 $O = (0, 0, 0)$ 上，換句話說，該影像平面為 $Z = 0$ ，而照相機的視點則定於 $(0, 0, f)$ 的坐標位置上，而 f 的絕對值就是照相機的焦距。如圖 2—4。

現在我們來看看，由照相機所拍攝的二維影像和三維影像間的幾何關係。如圖 2—5。假設在影像上有一點 $p = (x, y, z)$ ，一般稱 p 為像點（scene point），而 p 投影到二維影像平面中的影像點為 $p' = (x', y', 0)$

依三角形的相似原理，我們可得：

$$\frac{x'}{x} = \frac{\overline{AO}}{\overline{AB}} = \frac{f}{f - z}$$

所以我們會有：

$$x' = \frac{f \times x}{f - z}$$

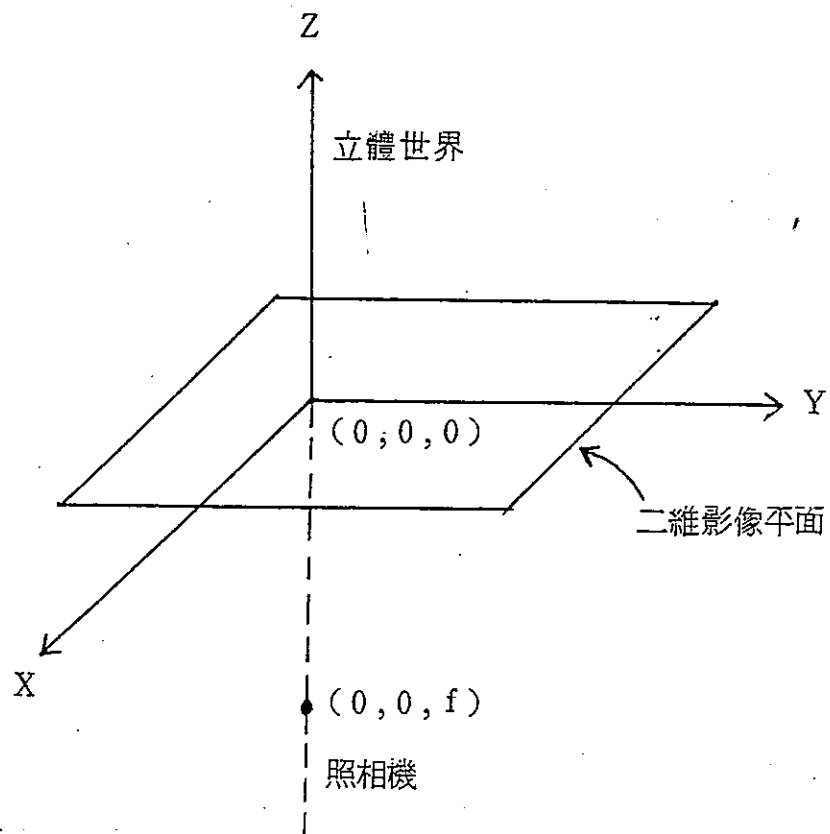


圖 2—4 影像平面圖

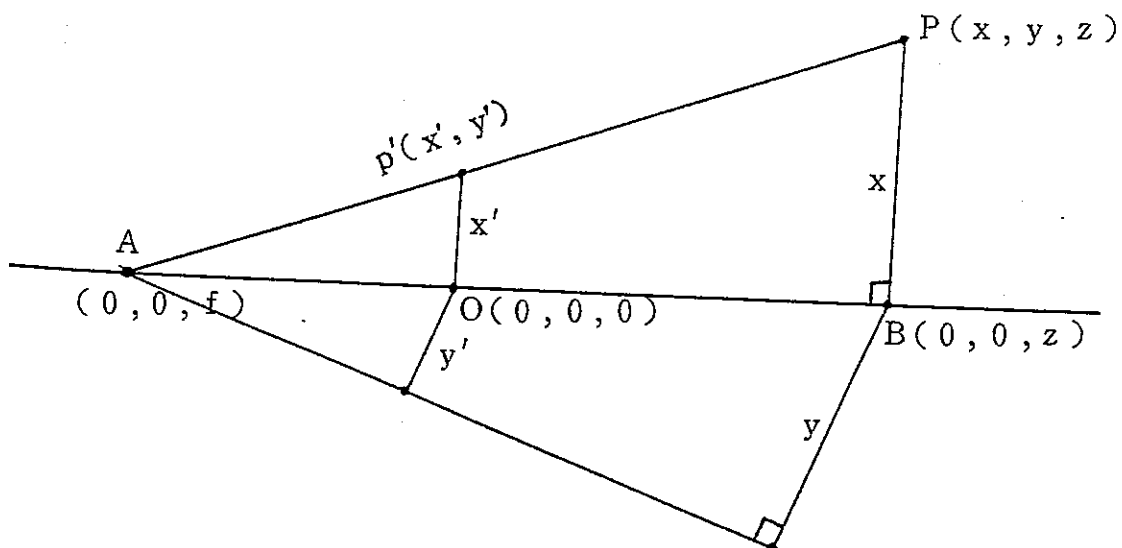


圖 2—5 三維像點和二維影像點的關係圖

$$y' = \frac{f \times y}{f - z} \quad (2-1-1)$$

換句話說，像點 $p = (x, y, z)$ 被投影到二維影像平面上的點 $p' = (x', y', 0)$ ，因 $z' = 0$ ，這種投影現象，我們稱之為中心投影 (central projection)，而這裏的 $\frac{f}{f - z}$ ，稱之為衰退係數。其意思為二維物體的實際形狀大小經中心投影後，在二維影像平面上所得到的影像會變形並且縮小其形狀大小。最重要的是喪失了深度值 (depth value)。

第二節 投影

我們知道，在景像中的三維物體都可以由一組三維像點的集合來描述。而當三維物體成像於影像平面時，相當於這組三維像點投影於影像平面上。一旦，我們欲從影像中來瞭解物體在三維空間中的資訊，那麼必須知道三維物體是透過那種投影模式而形成的二維影像。所以在本節中我們將討論兩種基本的投影模式：

第一種為透視投影 (perspective projection)。第二種為平行投影 (parallel projection)。上敘二種投影的主要不同點是：照相機的透鏡中心 (視點) 擺放的位置。於第一種投影模式中，照相機透鏡中心的位置和影像平面的距離是一個固定，有限遠的地方。所以當物體的反射光通過影像平面後，將集中於照相機的透鏡中心上。第二種投影模式，就是將照相機透鏡中心的位置移到無窮遠處，使得物體的反射光將平行通過影像平面，而相交於無窮遠處的照相機透鏡中心。雖然在本篇文章中，我們是採用

正交平行投影 (orthographic parallel projection)，但是我們仍須先介紹透視投影，然後再利用透視投影與平行投影之相異處來定義我們所用的正交平行投影。現在我們定義透視投影：

定義 2.3.1：假設影像平面為 $z = 0$ ，而照相機透鏡中心位於 $(0, 0, f)$ 的座標位置上，對於三維空間中的任意點 $v = (x, y, z)$ ，其幾何關係如圖 2—5。我們定義透視投影 p 如下式：

$$p(x, y, z) = \left(\frac{f \times x}{f - z}, \frac{f \times y}{f - z}, 0 \right)$$

由定義 2.3.1 可知道三維空間上的任一點 (x, y, z) ，投影到影像平面為 $z = 0$ 時，其投影點 $z' = (x', y')$ $= \left(\frac{f \times x}{f - z}, \frac{f \times y}{f - z} \right)$ 。所有投影點的深度值都是一樣為平面 z 的深度值 0。因此我們可以知道所有投影點將喪失了原先其在三維空間上的深度值。接著我們知道平行投影是將照相機透鏡中心移至無窮遠處。由此我們就可以來定義正交平行投影：

定義 3.2.2：假設照相機的視點位於 $(0, 0, f)$ ，影像平面為 $z = 0$ ，依據定義 3.2.1 我們有

$$x' = \frac{xf}{f - z} \quad \text{和} \quad y' = \frac{yf}{f - z} \quad (2-3-1)$$

如果 f 趨向於 ∞ 則利用微積分的極限定理，我們改寫式子 (2—3—1) 為：

$$\lim_{f \rightarrow \infty} x' = \lim_{f \rightarrow \infty} \frac{x}{1 - \frac{z}{f}} = x \quad \text{和}$$

$$\lim_{f \rightarrow \infty} y' = \lim_{f \rightarrow \infty} \frac{y}{1 - \frac{z}{f}} = y$$

所以我們定義正交平行投影 O 爲：

$$O(x, y, z) = (x, y, 0)$$

由定義 3.2.2 可知三維空間中的像點 $v = (x, y, z)$ ，經正交平行投影 O ，其投影點 $I = (x', y') = (x, y)$ 。在本論文中，我們是採用定義 3.2.2 這種正交平行投影。其投影方向是垂直於影像平面的。

現在我們已經瞭解所有三維像點被投影到影像平面時，都會喪失其深度值。而若給予二維影像平面上的一點 $v' = (x', y')$ ，其將可對應到三維空間上無數的像點， $v = \{ (x', y', z') \mid z' \in \mathbb{R} \}$ ，在 v 這個點集中的三維像點都會被投影到影像平面的 v' 點。如圖 2—6 所示，在 $\overrightarrow{v'v}$ 這條半線上的所有三維像點都投

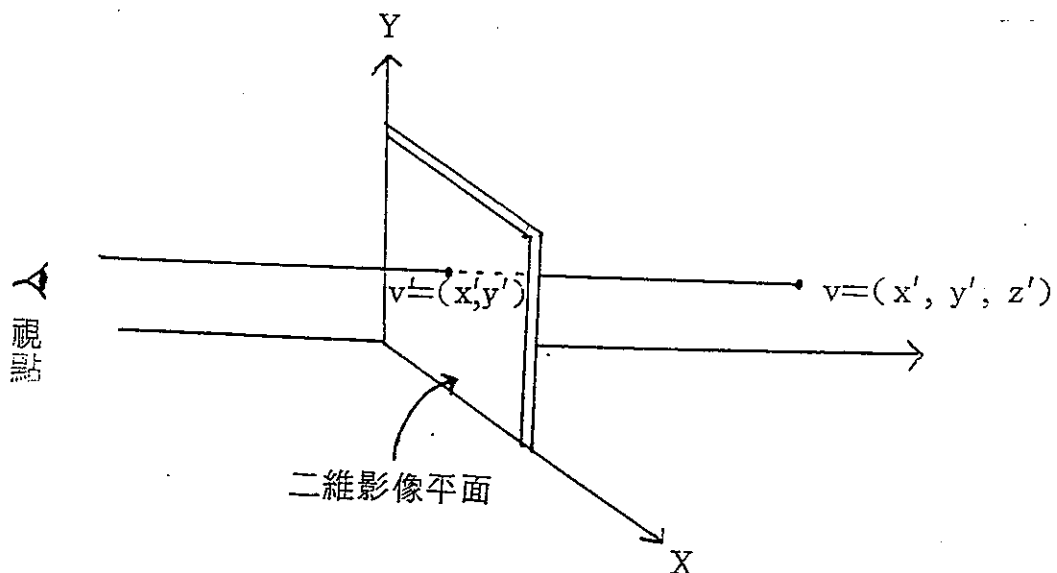


圖 2—6

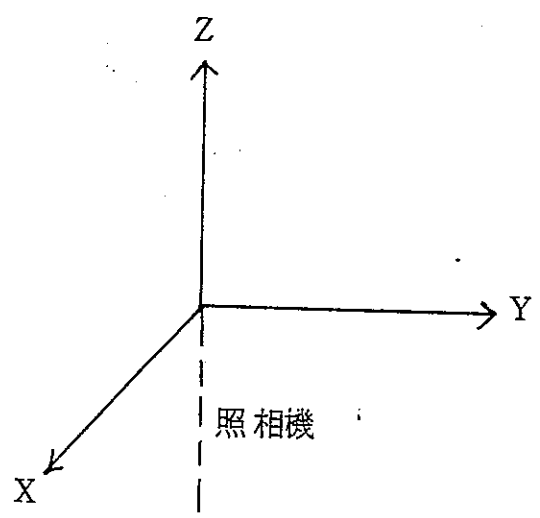
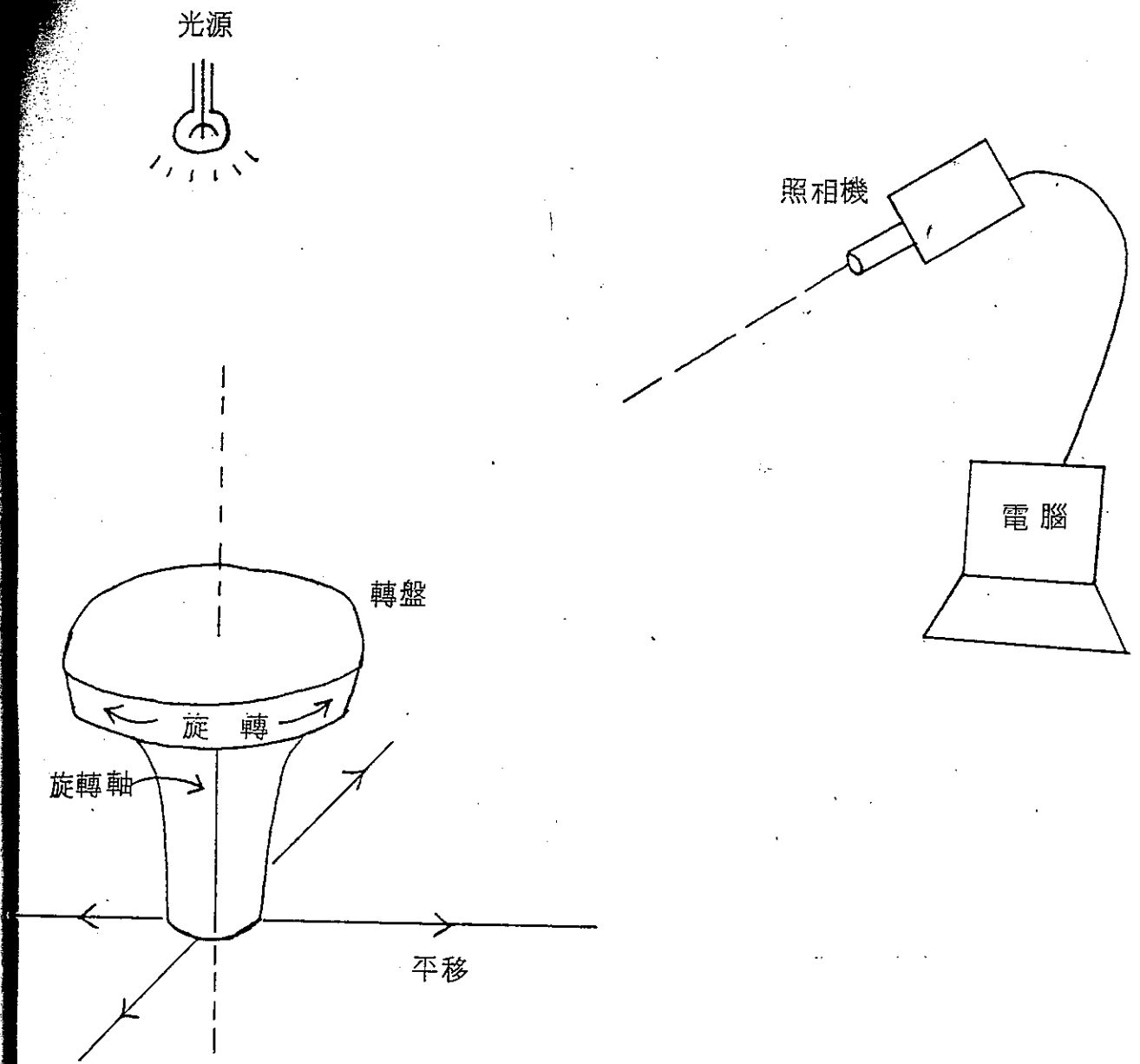
影至 v' 。因此在一般情況下如果吾人欲從單張二維影像去重建三

維物體，勢必造成混淆的現象。所以若想利用二維影像來重建三維物體時，必須利用多張二維影像。在本篇論文中我們將可利用兩張二維影像來重建出物體的可見部份之立體影像。由於我們所獲得的影像資料都是相對於影像平面所在的一個二維影像坐標系，然而除此之外，照相機、實際物體都各有一個不同的坐標系，所以我們要從影像資料來反求實際物體的三維資料及其運動參數，必須要先清楚這三個座標系統之間的幾何關係。

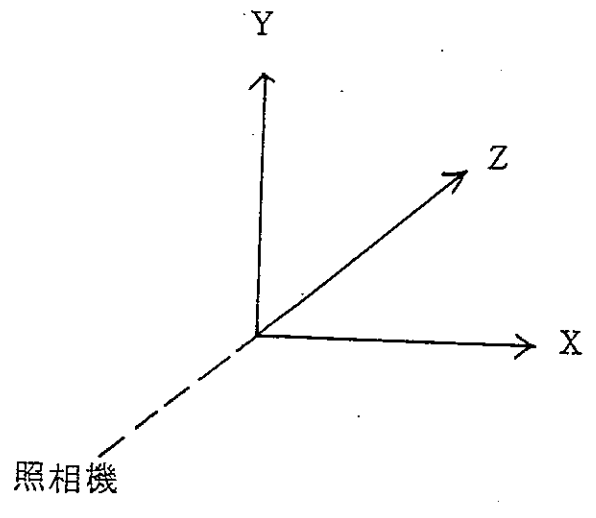
第三節 照相機與實物間的幾何關係

在此論文中，我們假設有一個控制系統可以來決定物體的運動參數，而在此系統中包括(1)置放物體的旋轉盤，轉盤只能針對其旋轉軸以逆時鐘或順時鐘方向旋轉，並且只能前後左右平移，(2)一部照相機，其必須不致使物體的形狀於成像後有發生扭曲現象，(3)一座發光系統，能使得物體本身和其背景之間有較明顯的亮暗對比，並且要能使得轉盤上的物體盡量沒有陰影產生，(4)一部電腦，其能接收數位影像資料加以處理且可發出控制訊號。其系統組織如圖 2—7 所示。而我們將在介紹一些常用的座標系及其座標位置表示法後，來定義圖 2—7 中三個主要的座標系，並敘述三個座標系間的幾何關係。

首先我們必須瞭解座標系中座標位置的表示法。假設有一個三度空間的座標系統是由三個互相垂直的軸 (orthogonal axes) 所構成。如果這三個軸分別以 X , Y , Z 表示。當照相機位於 Z 軸時，若座標系中的 Z 軸的正軸指向照相機，則這種座標系稱之



(a)



(b)

圖 2—8 座標系統

爲右手座標系，如圖 2—8 (a)。反之，稱之爲左手座標系。如圖 2—8 (b)。

利用互相垂直的三個軸 X ， Y ， Z 來代表的座標系稱之爲直角座標系 (rectangular coordinate system)。另外還有一種常見的球面座標系 (spherical coordinate system)。而後者當可由前者來說明之，因爲這二種座標系間的座標位置表示法是可以互換的。如圖 2—9 所示。在直角座標系中一點 $p = (x, y, z)$ ，其球面座標的表示法爲 (D, θ, ϕ) 。這裏的 D 是指點 P 到座標原點的距離，而角度 θ 則是線段 OP 在 XY 平面的投影，線段 OP' ，與 X 軸的夾角。這裏的 θ 之測量方式是從 X 軸之正軸方向以逆時鐘方向來測得。至於角度 ϕ 是線段 OP 與正 Z 軸的夾角。其關係式如下：

$$x = D \times \sin \phi \times \cos \theta$$

$$y = D \times \sin \phi \times \sin \theta$$

$$z = D \times \cos \phi$$

$$D = (x^2 + y^2 + z^2)^{\frac{1}{2}}$$

$$\theta = \tan^{-1} (y / x)$$

$$\phi = \cos^{-1} (z / D)$$

通常我們稱 P 點在 $X—Y—Z$ 座標系中的傾斜角 (tilt angle) 爲 ϕ 而泛角 (pan angle) 爲 θ 。

現在我們來定義圖 2—7 中實物和照相機二個座標系。首先我們對轉盤所處的座標系稱之爲實物座標系，如圖 2—10(a)，是

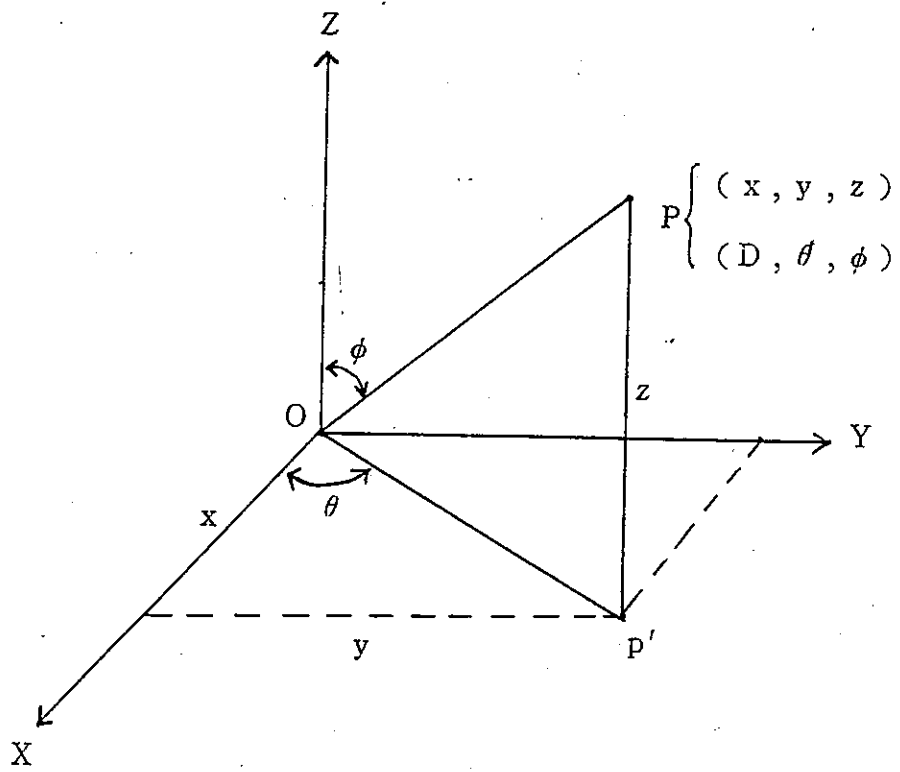


圖 2—9 直角座標系和球面座標系的關係

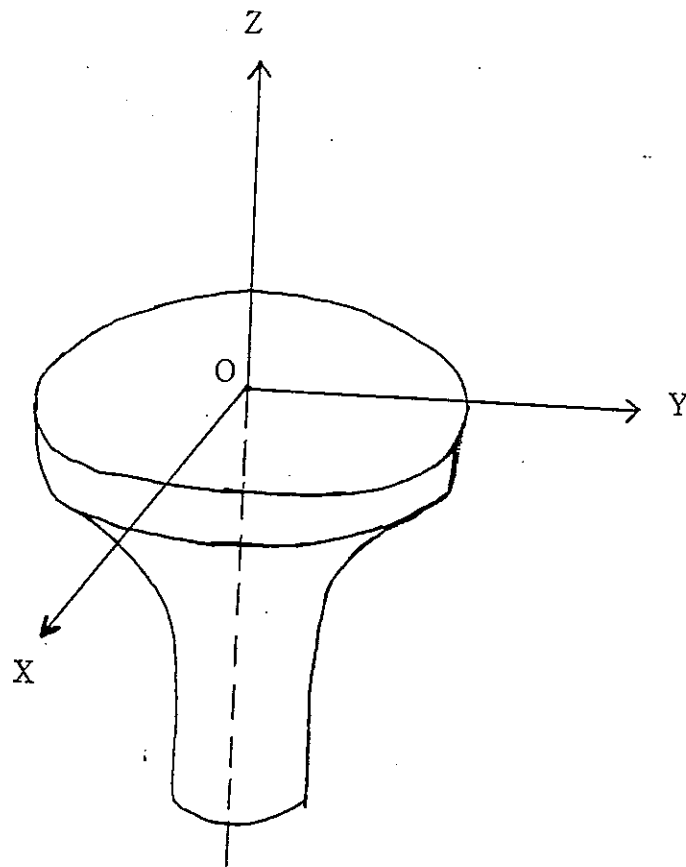


圖 2—10 定義實物座標系

由 X , Y , Z 三軸所構成的一個直角座標系。而此論文中，假設轉盤的平面將平行於 XY 平面，而其旋轉軸則平行於 Z 軸。而照相機上亦有一個座標系，我們稱之為照相機座標系。如圖 2—10(b)，是由 X_c , Y_c , Z_c 三軸所構成的直角座標系。另一個座標系是影像座標系是由 X_I , Y_I 二軸互相垂直所構成，並且 $X_I \parallel X_c$, $Y_I \parallel Y_c$ 。在此文中亦假設照相機的光軸係指正 Z_c 軸所指的方向，並且照相機的光軸將通過影像座標系的原點 O_I ，然後也要通過轉盤的中心點及實物座標系的原點。以上就是這三個座標系間的幾何關係。如圖 2—10(c) 所示。在此論文中，我們假設照相機是固定於實物座標系中的某個位置，並且系統校正已經完成。也就是說，照相機的光軸已經通過影像平面的中心和通過轉盤中心。

對於實物座標系中的一點 $p = (x, y, z)$ ，其必須透過一個座標軸變換 T_{wc} ，才能變換到照相機座標系中。假設 p 變換到照相機座標系上為 $p' = (x', y', z')$ ，那麼 p , p' 和 T_{wc} ，三者間有底下的關係：

$$p T_{wc} = p' \quad (2-4-1)$$

至於 T_{wc} 的獲得，我們將分成四個步驟來完成：

現在假設照相機透鏡中心位在實物座標系中的 (a, b, c) 。如圖 2—10(c)，i.e., $O_c = (a, b, c)$ 。相當於我們已經知道 $D = (a^2 + b^2 + c^2)^{\frac{1}{2}}$ ， $\theta = \tan^{-1} \left(\frac{b}{a} \right)$ ， $\phi = \cos^{-1} \left(\frac{c}{D} \right)$ 且 $a = D \sin \phi \cos \theta$ ， $b = D \sin \phi \sin \theta$ ， $c = D \cos \phi$ 。因此由下

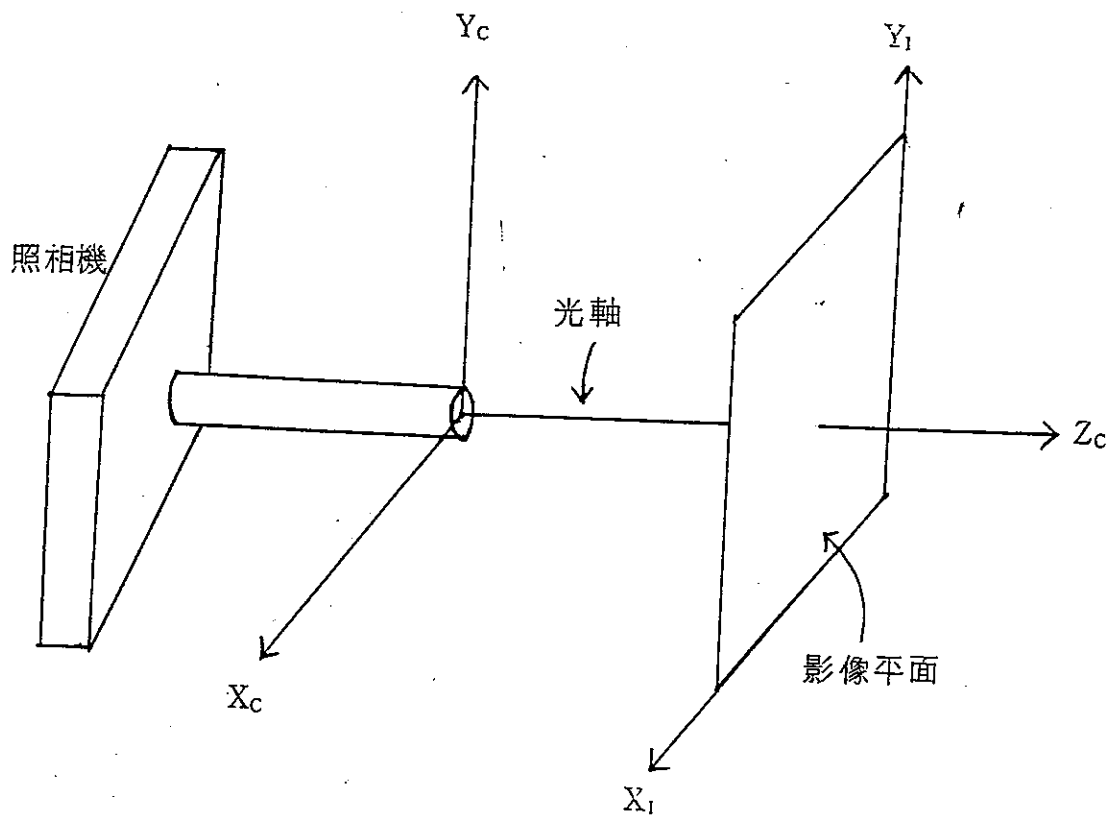


圖 2—10(b) 定義照相機座標系及影像座標系

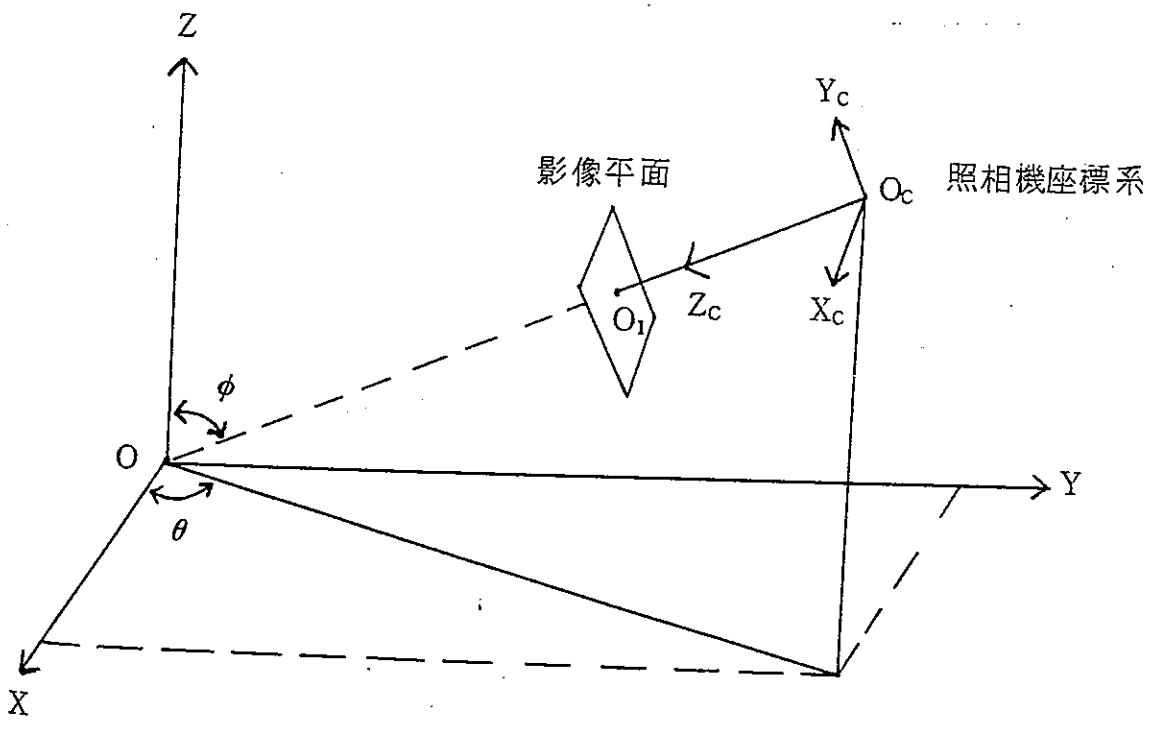


圖 2—10(c) 照相機座標系和實物座標系間的關係

列四個步驟就可以獲得 T_{wc} :

步驟 1 : 首先將實物座標系, 從原點 O , 平移至 $O_c = (a, b, c)$, 其透過平移矩陣為 :

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -a & -b & -c & 1 \end{bmatrix}$$

步驟 2 : 再以逆時鐘方向, 旋轉 Z 軸 $(90 - \theta)$ 度, 該旋轉矩陣為 :

$$T_2 = \begin{bmatrix} \sin \theta & \cos \theta & 0 & 0 \\ -\cos \theta & \sin \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

步驟 3 : 依順時鐘方向, 旋轉 X 軸 $(180 - \phi)$ 度, 該旋轉

$$T_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -\cos \phi & \sin \phi & 0 \\ 0 & \sin \phi & -\cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

步驟 4 : 將 X 軸的方向反轉, 利用底下 T_4 的矩陣 :

$$T_4 = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

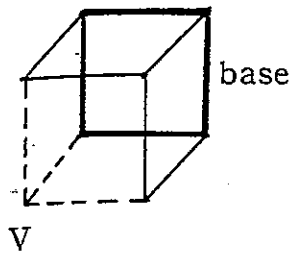
這個座標系變換 T_{wc} ，相當於這四個均質座標表示法的矩陣 T_1 ， T_2 ， T_3 ， T_4 的變換，因此我們可以將這四個矩陣整理成一個：

$$T_{wc} = \begin{pmatrix} -\sin \theta & -\cos \theta \cos \phi & -\cos \theta \sin \phi & 0 \\ \cos \theta & -\sin \theta \cos \phi & -\sin \theta \sin \phi & 0 \\ 0 & \sin \theta & -\cos \phi & 0 \\ 0 & 0 & D & 1 \end{pmatrix}$$

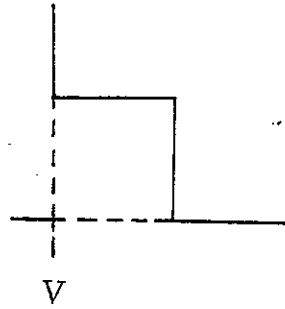
反之，若要從照相機座標系轉換到實物座標系則座標軸變換應為 T_{cw} ，而 $T_{cw} = T_4^{-1} T_3^{-1} T_2^{-1} T_1^{-1}$ ，i.e.，

$$T_{cw} = \begin{pmatrix} -\sin \theta & \cos \theta & 0 & 0 \\ \cos \phi \cos \theta & \cos \phi \sin \theta & -\sin \phi & 0 \\ -\sin \phi \cos \theta & -\sin \phi \sin \theta & -\cos \phi & 0 \\ D \sin \phi \cos \theta & D \sin \phi \sin \theta & D \cos \phi & 1 \end{pmatrix}$$

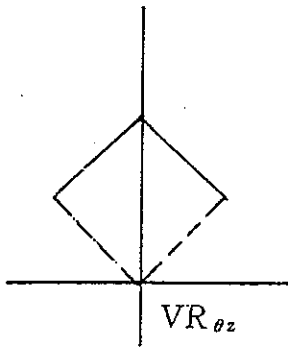
假如我們將 $p' = (x', y', z')$ ，以正交平行投影模式，投影到二維影像平面則 p' 的投影點之座標為 (x', y') ，若用透視投影模式則其投影點之座標為 $(\frac{fx'}{f-z'}, \frac{fy'}{f-z'})$ ， $|f|$ 為照相機的焦距。至此，我們知道，影像中的一點，實際上是實物座標系中的



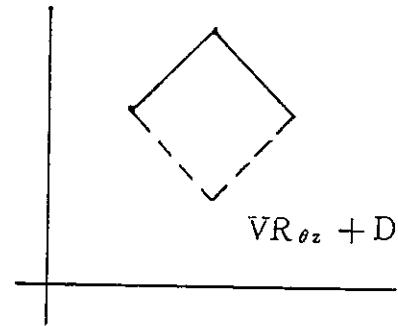
(a)



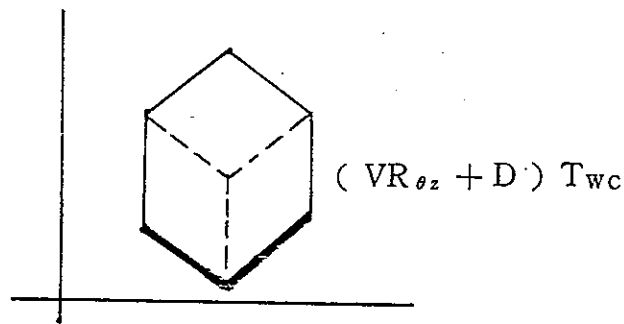
(b)



(c)



(d)



(e)

圖 2 — 11 三維像點變換至影像平面

三維點，經座標轉換 T_{wc} ，再經由投影而形成影像點。底下是物體有經過旋轉、平移運動後，再成像於二維影像平面的過程。如圖 2—11。如果三維物體在實際座標系中的運動只限於在平行 XY 平面平移，和僅針對 Z 軸旋轉。假如物體的平移向量和旋轉矩陣分別為 D ， R_{θ_z} 。圖 2—11(a) 表示一個正立方體上有一點 V 和有一面是看不見的底部，圖 2—11(b) 表示此正立方體被於實物座標上的起始位置，圖 2—11(c) 表物體針對 Z 軸旋轉 θ_z 的角度後，物體的 V 點之新位置為 VR_{θ_z} 。圖 2—11(d) 是物體由圖 2—11(c) 中的位置，經由平移向量 D 平移到新的位置，所以點 V 的新位置為 $VR_{\theta_z} + D$ ，圖 2—11(e) 表物體從實物座標系變換到照相機座標系後，在影像平面上的投影所顯現的物體形狀。

第四節 數位影像的形成及其前置處理

在此文中假設照相機最先拍到的影像是一張明暗影像 (intensity image)，而這種明暗影像上各點所代表的光線亮暗變化乃是一種連續變化的訊號。如果要將這種影像輸入到數位電腦內時，必須先將這種由連續性的訊號所構成的影像，轉換成由數位電腦所能處理的符號來表示各點的光線亮暗程度之影像後，才能輸入數位電腦來處理及儲存。而上述所提到的轉換過程即數位化過程，所產生的影像稱之為數位影像 (digital image)。而在數位化過程中最常見的轉換器有影像數位化器 (image digitizer or scanner) 和電視照相機 (T.V. Camera)。前者是將拍好的照片全部掃描一遍，就可轉化成一張數位影像。而後者通常是直

接對景像 (scene) 做掃描，然後將景像中被掃描過的影像轉成數位影像。這種數位影像在電腦內的儲存方式為何？大部份的情形是以用矩形數值陣列來儲存。而怎樣來決定出陣列中元素的數值呢？這就須牽涉到取樣 (sampling) 和量化 (quantizing) 的問題。底下我們就來介紹這兩個問題。

實際上，以數學上的口吻來講，一張影像就是一個連續變化的函數，我們稱之為明暗函數 (intensity function)， $f(x, y)$ ，其中 x, y 代表在二維影像中的整數坐標值，而在 (x, y) 上光線的明暗程度為 $f(x, y)$ 。而這裏 $f(x, y)$ 的值就是所謂的灰度值 (gray level)，代表著光線亮暗的程度，而陣列上的各元素就是儲存了這種灰度值。

在數位化過程中，有兩個重要步驟：1. 取樣 (sampling)，和 2. 量化 (quantizing)，來決定出陣列中所儲存的灰度值。而所謂取樣就是在一張影像中取出一組樣本點 (sample points)，並且這組樣本點必須能夠再重建出原來連續的影像。所以在這個步驟中有兩個非常重要的工作：(一) 為取樣區間 (sampling interval) 的選定，(二) 為這組樣本點在空間上排列分佈的形狀。一旦上述的二個工作能順利完成，則這個取樣過程所取出的一組樣本點就能夠重建成原來連續的影像。接著我們在這組樣本點中，每一點給予一個數值。這就是所謂的量化步驟。而這組樣本點中各點被附予的數值就是我們所謂的灰度值。而由這組樣本點所組成的影像就是一張數位影像。所以構成一張數位影像的基本元素就是這些樣本點，它們是一張數位影像中最基本的元素，簡稱

爲圖元 (pixel, i.e., pictural element)。每個圖元的亮暗程度就以灰度值的大小來表示。有關於取樣和量化的技術可以在任何一本數位影像處理的書籍中找到更詳盡的探討，例如 Gonzalez [8]，Rosenfeld [9]，和 Duda [10]，等。

一張數位影像輸入電腦後，首先必需考慮如何去除雜訊 (noise)？以便能期使由雜訊所造成的誤差減低，然後下一步驟就是如何從二維數位影像中粹取出一些重要的特徵 (feature)？等等。當然在本文中所假設的物體都是由平面所組成，該類物體的重要特徵不外乎是頂點 (vertex)，邊 (edge) 及平面 (planar surface)，這些特徵被投影於二維影像平面上，分別會成爲接點 (junction)，邊線 (line-segment) 和區域 (Region) 因此在二維影像中對上述特徵的粹取是很重要的前置處理工作。底下就簡單來探討前置處理所需要的工作。

我們知道一張明暗影像中，在邊線附近的圖元其灰度值會產生急劇的變化。原因是在三維物體上，面與面相交的邊及其附近區域因光線照射在兩相鄰面上的亮度不同而發生的。於是乎抽取邊線成爲前置處理過程中最主要的工作。一般而言，有兩種抽取邊線的方法：第一種直接針對邊線而言，必須從整張影像中，去偵測出構成邊的點稱之爲邊點 (edge point)，然後將構成邊的一組組邊點偵測出，每組邊點所連成的直線就是一條邊線。這種方式稱之爲邊的方法 (edge method)。例如 Riseman [11] 就是用此種方式在二維影像中粹取出直線。在採用這種方法之前，必須要先了解影像中的物體形狀，可能出現的邊線型態。例如：等

高線 (contour line)，邊界線 (boundary line)，內部線 (internal line) 或者是接線 (crack line) 等等。因為以上這些型態的邊線，在二維影像中，它們鄰近區域圖元的灰度值變化情況各有緩急不同的類別。因此在採用此法之前還必須先訂定出邊的運算子 (edge operator)。這種運算子有一維和二維二種。而所謂一維的邊運算子最常用於粹取出一條已知方向的邊線上的邊點。例如： $D_{a(i)} = f(i+d) - f(i-d)$ ，上式在測試等個位置上的圖元是否為邊點，這裏的 d 是一個整數， $f(i+d)$ 和 $f(i-d)$ 分別表第 $i+d$ 個和第 $i-d$ 個位置上圖元的灰度值。假如 $D_{a(i)}$ 不為 0 則第 i 個位置上的圖元就是邊點，否則就不是邊點。而所謂二維的運算子，首先要設定一個窗戶 (window)，而窗戶的中心點位置的圖元就是要被決定其是否為邊點者，因此我們就必須計算出這個窗戶內的各圖元之灰度值的相互關係來偵測其中心點是否為邊點，如果這個窗戶的大小為 3×3 ，那我們稱此二維運算子為 3×3 的 sobel 運算子。利用這種運算子將整張影像的邊點粹取出後，接著就可利用 Nevatia [12] 所提出的演繹法，將所有粹取出的邊線點連接成邊線。第二種方法稱之為區域方法 (region method)，這種方法是先在一張二維影像上切分成幾個均勻的基本區域 (tiny atomic regions)，逐漸結合附近有相同灰度值的圖元，成為較大一連串相鄰的區域。而這裏所謂的基本區域可以是獨立的一個圖元或相鄰的一些相同性質的圖元所成的集合。而相同性質是指，諸如灰度值相同。若一張二維影像中的各區域已無相鄰的圖元可結合時，很容易就把圍成這

些區域之所有邊界線抽取出來，相當於已經把二維影像中的邊線粹取出了。利用上述二種粹取邊線的方式都能將一張二維的掃描式影像 (raster image) 轉化成線繪影像 (line-drawing or wire frame image)。在這種線繪影像上就只保留邊線和接點，而區域 (region) 則是指由邊線所圍成的封閉區域。接著我們還必須要知道邊線相交而成的接點為何種型態？這可以利用一個 5×5 的窗戶 (5×5 window)，將此窗戶的中心點移到一張影像中的各接點上，來檢驗出每個接點型態。假如恰有二條邊線經過該接點，此乃 V 型接點，若恰有三條邊線相交於一接點，那麼依據這三條邊線間的夾角，可區分成：T 型，Y 型和 W 型。如果檢驗出超過三條邊線相交時，這種接點為 M 型。當然在解譯線繪影像時，邊線和接點型態各有其實際上的幾何意義。而第三章中的第一節中，會加以詳盡的介紹。

第五節 剛體運動

當一個物體只做旋轉和平移運動時，若該物體原具有的面、邊、點的數目，不因物體運動而突然增加或減少，則我們稱這類物體為剛性物體，而其所做的運動為剛體運動 (rigid body motion)。在此論文中，我們假設所允許置放轉盤上的物體都是剛性物體，而物體只能被旋轉和平移運動，事實上就是在做剛體運動。而大家都知道，任意一個剛體運動 M，都可以被表示成如下的變換：

$$M(x, y, z) = (x, y, z)R + D$$

其中 R 為一個 3×3 的正規矩陣 (orthonormal matrix)，且 $\det(R) = 1$ ，稱為剛體運動 M 的旋轉矩陣。

$$R = \begin{bmatrix} n_1^2 + (1 - n_1^2) \cos \theta & n_1 n_2 (1 - \cos \theta) + n_1 \sin \theta & n_1 n_3 (1 - \cos \theta) - n_2 \sin \theta \\ n_1 n_2 (1 - \cos \theta) - n_1 \sin \theta & n_2^2 + (1 - n_2^2) \cos \theta & n_2 n_3 (1 - \cos \theta) + n_1 \sin \theta \\ n_1 n_3 (1 - \cos \theta) + n_2 \sin \theta & n_2 n_3 (1 - \cos \theta) - n_1 \sin \theta & n_3^2 + (1 - n_3^2) \cos \theta \end{bmatrix} \quad (2-5-1)$$

此處的 n_1 ， n_2 和 n_3 為任一過原點的旋轉軸分別對 X ， Y ， Z 三軸之方向餘弦，且 $n_1^2 + n_2^2 + n_3^2 = 1$ ， θ 為三維物體對該軸的旋轉角度。而 $D = (\Delta x, \Delta y, \Delta z)$ 稱為平移向量。所以參數 $n_1, n_2, n_3, \theta, \Delta x, \Delta y, \Delta z$ 稱為剛體運動 M 的運動參數 (motion parameters)。由於剛體運動 M 的旋轉矩陣 R 有 $\det(R) = 1$ 的特性，因此 M 可以保存原有三維物體之立體幾何架構中，面與面之間所夾的角度，頂點和頂點之間的距離，此外 M 亦可保存原有三維物體的體積。

對於一個剛體運動 M 而言，它也正是所謂的仿射變換 (affine transformation)，而對於一個仿射變換 T 來說，必定存在唯一的反仿射變換 (inverse affine transformation) T^{-1} ，使得 $TT^{-1} = T^{-1}T = I$ ，其中 I 為單位矩陣 (identity matrix)。換句話說，對於任一個剛體運動，必定存在唯一的反剛體運動 (inverse rigid body motion) M^{-1} ：

$$M^{-1}(x, y, z) = (x, y, z)R^{-1} - DR^{-1}$$

很顯然的， M^{-1} 也是一個剛體運動。此外，三度空間中的仿射變

換 T 還有一個非常美妙的特性： T 可以將一個三維平面依然對映到三維平面，而且可將三維線段對應到三維線段。當然，剛體運動 M 亦可將三維平面對映到三維平面，亦可將三維物體之邊線對映到另外一個位置的三維邊線。

底下我們舉一個例子來看看這些運動參數如何由照相機與旋轉盤間的幾何關係圖中求得？

範例 2.1：如圖 2—12(a) 中的三維物體 Ω ，圖中右邊為此三維物體在二維影像的形狀，而此三維物體 Ω 之面 s_1, s_2, s_3 分別投影到二維影像平面中的區域 R_1, R_2, R_3 ，而 Ω 的邊 $e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8$ 分別投影到二維影像平面中的邊線 $l_1, l_2, l_3, l_4, l_5, l_6, l_7, l_8$ ， Ω 的頂點 $v_1, v_2, v_3, v_4, v_5, v_6$ 則投影到 $J_1, J_2, J_3, J_4, J_5, J_6$ 。因為我們的旋轉盤上的轉軸是在實物坐標系統中的 Z 軸，並且在平行於 $X-Y$ 平面上平移。若我們訂定轉軸依逆時針方向旋轉時，其旋轉角度為正，反之，其角度為負。如果圖 2—12(a) 中左邊三維物體 Ω ，被轉軸依逆時針方向旋轉一角度 θ 後，可得圖 2—12(b) 中左邊三維物體 Ω' ，而 Ω' 的面 s'_1, s'_2, s'_3 ，邊 $e'_1, e'_2, e'_3, e'_4, e'_5, e'_6, e'_7, e'_8$ ，頂點 $v'_1, v'_2, v'_3, v'_4, v'_5, v'_6$ ，分別被投影在二維影像平面上的區域 R'_1, R'_2, R'_3 ，邊線 $l'_1, l'_2, l'_3, l'_4, l'_5, l'_6, l'_7, l'_8$ 和接點 $J'_1, J'_2, J'_3, J'_4, J'_5, J'_6$ 。這是因為轉盤的轉軸為 Z

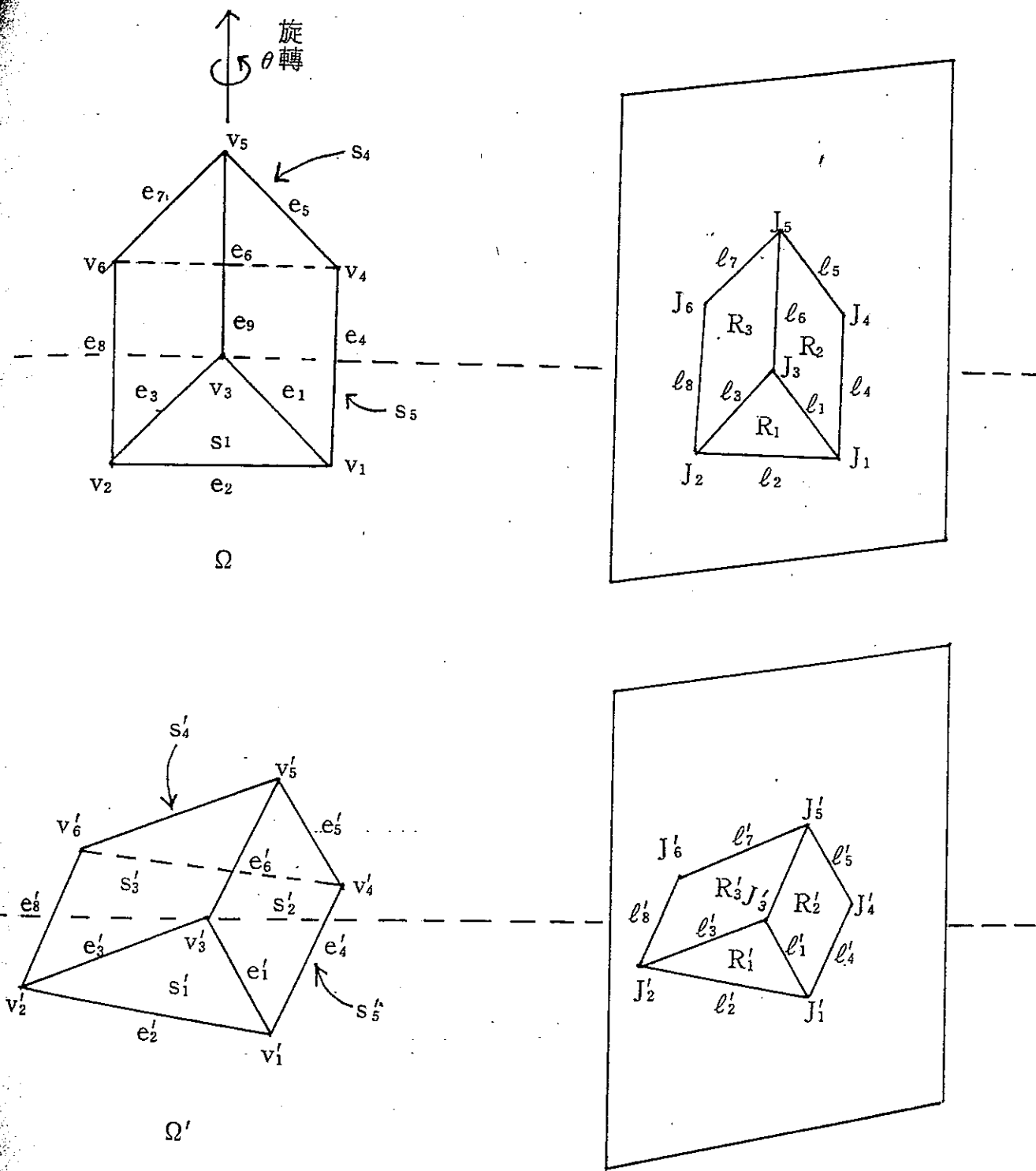


圖 2 — 12 三維物體旋轉圖

軸的緣故，而旋轉軸的方向餘弦， $n_1 = 0$ ， $n_2 = 0$ ，
且 $n_3 = 1$ ，因此由方向餘弦 n_1 ， n_2 ， n_3 以及旋轉角
度 θ 所求得的旋轉矩陣為：

$$R = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

因此，對於三維物體 Ω 中之可見像點 (visible scene
point) $v_i = (x_i, y_i, z_i)$ 和相對於旋轉盤的中心點
 $O' = (0, 0, c)$ 所成的向量 $O'v_i$ ，針對 Z 軸旋轉 θ
後，可表示為：

$$\begin{aligned} & [(x_i, y_i, z_i) - (0, 0, c)]R + (0, 0, c) \\ & = (x_i, y_i, z_i)R + D \end{aligned}$$

故我們可取 $D = (\Delta x, \Delta y, \Delta z)$ ，使得

$$\Delta x = 0 \cdot \cos \theta - 0 \cdot \sin \theta + c \cdot 0 = 0$$

$$\Delta y = 0 \cdot \sin \theta + 0 \cdot \cos \theta + c \cdot 0 = 0$$

$$\Delta z = 0 \cdot 0 + 0 \cdot 0 + c \cdot 1 = c$$

因此我們可以得到剛體運動， M 的所有運動參數，而

$$\text{且，} \quad M(x, y, z) = (x, y, z)R + D$$

由這個剛體運動 M ，我們就可以得到圖 2-12(b) 中左
邊的三維物體 Ω' 及其二維影像。

第六節 立體視覺

對於獲取景像中實際三維物體的資料，一般常用的方式有下列二類：

第一類：利用測距器（range finder）直接測得三維物體的距離影像（range image）而影像中的資料稱之為距離資料（range data）。這種影像資料中的各圖元（pixel）所含有的數值不再是明暗影像中各圖元的灰度值（gray level）所代表著光線的強弱，它已是代表一種深度值（或高度值）了。

第二類：可分成兩種。第一種是同時用二台照相機在兩個不同角度觀察同一物體，所拍得的兩張二維影像，再利用立體視差（stereo disparity）來求出三維物體的資料。第二種是僅用一台照相機，在固定位置拍得物體位於兩個不同位置的兩張二維影像，然後分析兩張影像中物體的運動參數，再反求出物體的三維資料。不管是用立體視差或分析物體的運動參數來求出物體的三維資料，上述的兩種方法，我們都稱之為立體視覺（stereo vision）或雙眼視覺（binocular vision）。

通常第二類比第一類來得常用，其原因有二：(1)後者不須特別昂貴的儀器及設備，僅需兩張二維明暗影像即可。(2)對後者的研究比前者的研究來得早些。但是目前在研究立體視覺的領域中，最主要的難題還是歸諸於相對問題。假如相對問題能有效的被解決則用兩張二維影像就可以很快的還原出影像中物體的實際立體形狀。

接著我們來介紹立體視覺系統中常用的兩種方式來獲取物體

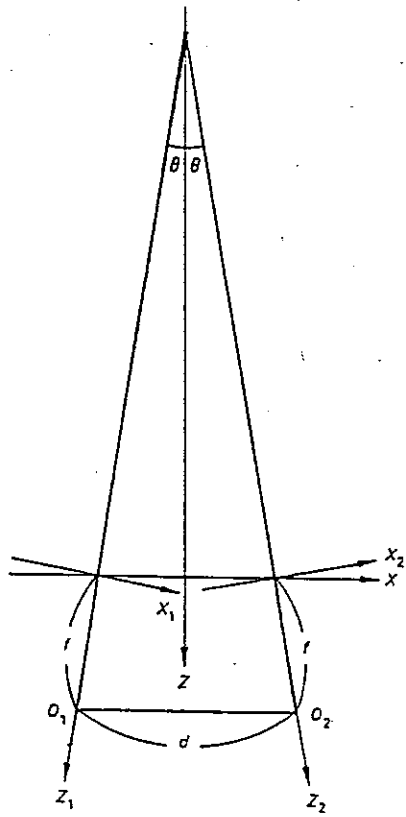


圖 2 — 13 立體視覺的幾何原理

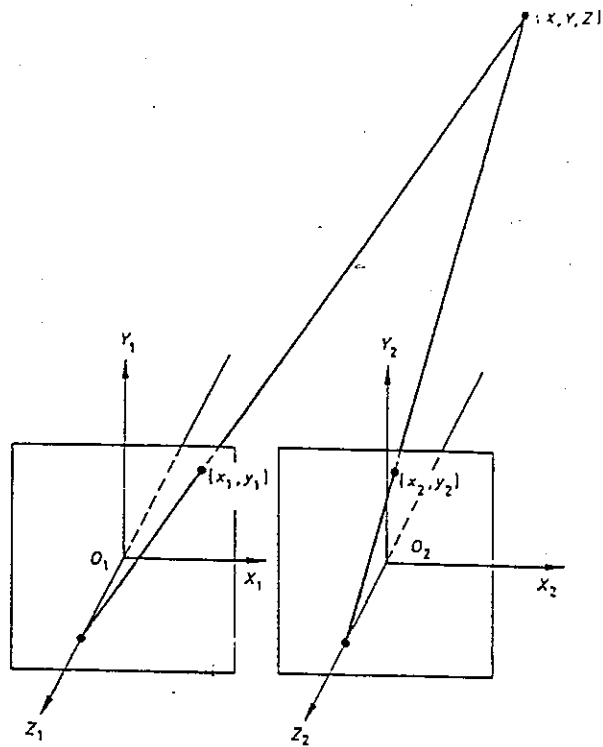


圖 2 — 14 立體影像系統

的三維資料。第一種：假設有兩台照相機，分別位於同一條水平線中，它們各有著相同的焦距 f ，而其透鏡中心點分別為 O_1 和 O_2 ，並且 $\overline{O_1O_2} = d$ ，如圖 2-13，在二個光軸間的夾角為 2θ ，這二個照相機各擁有二個不同的照相機座標系，分別以 $X_1 - Y_1 - Z_1$ 和 $X_2 - Y_2 - Z_2$ 來表示。對圖 2-13 而言， Y_1 和 Y_2 二個軸將是貫穿紙面。讓我們再定義一個座標系， $X - Y - Z$ ，使得 Z 軸能平分 Z_1 和 Z_2 軸所夾的角。對這三個座標的關係如下所示：

$$\begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix} = \left[x + \frac{d}{2} - f \sin \theta, Y, Z \right] R_y(\theta),$$

$$R_y(\theta) = \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix} \quad (2.6.1)$$

和

$$\begin{pmatrix} X_2 \\ Y_2 \\ Z_2 \end{pmatrix} = \left[X - \frac{d}{2} + f \sin \theta, Y, Z \right] R_y(-\theta),$$

$$R_y(-\theta) = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \quad (2.6.2)$$

如圖 2-14，假設某三維點 $p = (x, y, z)$ 被兩台照相機所

拍攝到，而點 p 在這兩台照相機的影像平面上的座標分別為 $I_1 = (x_1, y_1)$ 和 $I_2 = (x_2, y_2)$ 。則它們間的關係如下：

$$x_i = fx_i / (f - z_i) \quad (2.6.3)$$

和

$$y_i = fy_i / (f - z_i) \quad (2.6.4)$$

而 $i = 1, 2$ 。假設我們已知圖 2—14 中兩張影像中的點 I_1 和 I_2 是一組相對點，並且照相機的焦距 f 已知。那麼僅用兩張二維影像就能將點 p 在三維景像中的座標值求出。因為從 (2.6.1) 至 (2.6.4) 的四個式中，可獲得 10 條方程式，但卻僅有 9 個未知數待求。所以我們能解出這 9 個未知數。其中有三個即是點 p 的座標值 x, y, z 。如果兩張二維影像中的所有相對點已知，則依上述方法，將可求出影像中物體的實際三維立體影像。

另一種方式為：假設照相機只有一台，放於固定位置。若已分析出物體在兩個不同位置的運動參數 M 已知，則假設物體上有一點 $p = (x, y, z)$ ，而點 $p_1 = (x_1, y_1, z_1)$ 是物體在一位置上時，點 p 被變換到照相機座標系中的座標點，點 $p_2 = (x_2, y_2, z_2)$ 則是物體經運動 M 到另一位置後，點 p 出現於照相機座標系中的座標點。假設物體所在的實物座標系變換到照相機座標系間的座標系變換 T_{wc} 已知則上述關係可用如下表示：

$$(x, y, z) T_{wc} = (x_1, y_1, z_1) \quad (2.6.5)$$

$$(x, y, z) M T_{wc} = (x_2, y_2, z_2) \quad (2.6.6)$$

若 p_1 和 p_2 分別投影到影像平面時，其影像點為 $I_1 = (x'_1, y'_1)$ ， $I_2 = (x'_2, y'_2)$ 。（不限制何種投影）而 I_1 和 I_2 是一組相對點。所以現在對式子 (2.6.5) 和 (2.6.6) 而言，有 x' ， y ， z ， z_1 ， z_2 五個未知數，但卻有六條方程式。因此，將可解出上述五個未知數，然而點 p 的三維座標值 x ， y ， z 亦被求出了。所以我們知道在這一種立體視覺系統中只要知道兩張二維影像的相對點及其運動參數則必可重建影像中物體的實際立體影像。

在此論文中我們的系統環境是屬於第二種的立體視覺系統，但是我們將討論若在兩張二維影像之相對點未知的條件下，如何來決定物體之運動參數的問題。

第三章 二維影像之解譯

在此論文中，我們只考慮折紙狀 (Origami) 物體 [13]，並且也只允許折紙狀物體的形狀是包括三角平面多面體及上、下是透空的立體物。由於我們僅對合理的折紙狀物體來處理，並且我們在對這類物體的二維影像標線時，亦利用斜量空間的特性來過濾掉難以想像的解譯影像 (interpretative image) 之產生，因此一張二維影像在標線後將不致於得到兩種或兩種以上不同的解譯影像。然而，對這類折紙狀物體所形成的二維解譯影像，Kanade [14, 15] 可以從影像中各區域形狀就能還原物體的三維形狀 (非實際物體的立體形狀)。Kanade 乃利用各面在斜量空間上幾何關係及性質，及二維解譯影像中各全部可見區域的形狀來求出影像中物體各面之斜傾量 (gradient)，也就是法向量 (normal vector)。在此論文中，我們就利用 Kanade 所求出物體各面之法向量來測量出物體可能的旋轉角度。

在本章中共分四節來討論。第一節中是定義一些在此論文中常見的名詞及常用的符號及關鍵用語。第二節中我們是先定義一個面的傾斜量，和斜量空間，並且對本文中有用到各面在斜量空間上的性質做一簡介。第三節中首先描述折紙狀物體的特性，然後略述折紙狀物體如何形成解譯影像以及去除令人難以想像的不合理解譯影像。最後在第四節中，將限制多個折紙狀物體的遮蔽現象，並描述在這類遮蔽現象發生時，所形成的二維影像仍能用第三節的標線法則來標線而形成合理的解譯影像。最後我們利用

幾個啓發式規則將解譯影像中的部份被遮區域找出，並對一些區域猜測出其被遮部份之形狀。

第一節 名詞定義

在這一節中，我們將描述在這篇文章中比較常見的一些名詞，然後定義幾個較為重要的關鍵名詞及符號。

我們知道，對於一個三維空間中的平面物體 (Planar surface object) Ω 而言，它是由一些三維空間的點所構成。換句話說， Ω 相當於這些點所成的集合。一般而言，三維平面物體的主要特徵為面 (surface)，各面相鄰的邊 (edge) 及構成有限長度邊的頂點 (vertex)。而物體 Ω 在二維空間中的影像平面上的影像 $O (M (\Omega))$ 中，相對於 Ω 在三維空間中的主要特徵分別為區域 (region)，邊線 (line-segment) 和接點 (junction)。(這裏的 O ：表正交平行投影， M ：表某一方向看景像的運動參數。)

一個二維影像平面上之區域 R 乃是由邊線以及接點所構成，而邊線 L 則由兩個不同接點所構成。在影像平面上粹取一個區域 R 的過程中，首先我們找出一個邊線 L_1 ，然後依逆時鐘方向，取出和邊線 L_1 相鄰 (相交) 而且和邊線 L_1 間之逆時鐘交角最小的邊線 L_2 ，然後從邊線 L_2 出發，用同樣方法找出 $L_3, \dots, L_i, \dots, L_n$ ，直到找出邊線 $L_{n+1} = L_1$ 時，那麼區域 R 即由 $L_1, L_2, L_3, \dots, L_n$ 所構成。而每一條邊線 L_i ，都是由二接點 J_i, J_{i+1} 所構成，而整個區域 R 中所有接點 J_1, J_2, \dots, J_n 是以逆時鐘方向來排列。在這

篇文章中爲了處理上的方便以及符號上的統一起見，當我們說，某一個區域 R 由 n 個邊線 $L_1, L_2, \dots, L_1, \dots, L_n$ 和 n 個接點 $J_1, J_2, \dots, J_1, \dots, J_n$ 所構成時，這些邊線 L_1, L_2, \dots, L_n 以及接點 J_1, J_2, \dots, J_n 都是依逆時鐘方向排列的。在此值得注意的是：二維影像 $O(M(\Omega))$ 中的接點並不一定是全由三維物體上的頂點 V 在二維影像平面上的投影而得到。例如圖 3-1 中，接點 J_5 和 J_6 就不是由兩個三維物體上之任何頂點，在二維影像平面上之投影。

接著，我們定義一些重要的關鍵名詞及符號：

定義 3.1.1：假設 Ω 爲一個三維物體，且 Ω 由 m 個面 S_1, S_2, \dots, S_m 以及 n 個邊 e_1, e_2, \dots, e_n 和 P 個頂點 V_1, V_2, \dots, V_p 所構成。那麼我們定義 Ω 如下：

$$\Omega = (S, E, V)$$

其中 $S = \{s_1, s_2, \dots, s_m\}$ ，

且 $E = \{e_1, e_2, \dots, e_n\}$ ，

且 $V = \{v_1, v_2, \dots, v_p\}$ 。

在定義 3.1.1 中集合 E 爲一個佈於 V 中的二元關係 (binary relation)，也就是說，對於集合 E 中的任何一個元素 e_i 相當於一個次序對 (order pair) (V_x, V_y) ，其中 $V_x, V_y \in V$ 。而集合 F 中之每一個面 S_i 乃是由 V' 和 E' 二集合所構成。這裏的 V' 和 E' 分別爲集合 V, E 的一個子集合。底下我們將定義一個二維區域：

定義 3.1.2：假設 Ω 爲一未知三維物而 $O(M(\Omega))$ 爲對

應於 Ω 之二維影像。如果二維影像 $O(M(\Omega))$ 中存在某一個區域 R ，依逆時鐘方向，由 n 個接點 J_1, J_2, \dots, J_n 以及 n 條邊線 L_1, L_2, \dots, L_n 所構成， $n \geq 3$ 。那麼，我們定義區域 R 如下：

$$R = (J, L)$$

$$\text{而 } J = \{ J_1, J_2, \dots, J_n \},$$

$$L = \{ L_1, L_2, \dots, L_n \},$$

$$\text{其中 } L_i = (J_i, J_{i+1}), i = 1, 2, \dots, n-1,$$

$$\text{而 } L_n = (J_n, J_1)。$$

如圖 3—1 中在二維影像平面上之區域 R 依逆時鐘方向由 4 個接點 $J = \{ J_1, J_2, J_3, J_4 \}$ 以及 4 條邊線 $L = \{ L_1, L_2, L_3, L_4 \}$ 所構成。接著我們定義所謂的合理區域。

定義 3.1.3：假設 $\Omega = (S, E, V)$ 為一未知三維物體，對於一個二維影像 $O(M(\Omega))$ ，我們定義一個合理區域 (legal region) R 如下：假設集合 S' 存在一個面 s_i ，由 n 個邊 e_1, e_2, \dots, e_n 及 n 個頂點 v_1, v_2, \dots, v_n 所構成。而二維影像 $O(M(\Omega))$ 中存在某一個區域 $R = (J, L)$ ，其中 $J = \{ J_1, J_2, \dots, J_n \}$ 。且 $L = \{ L_1, L_2, \dots, L_n \}$ 。如果 $O(M(v_i)) = J_i$ 且 $O(M(e_i)) = L_i$ ， $i = 1, 2, \dots, n$ 。則區域 R 稱為合理區域。

如圖 3—1 中，左邊的三維物體有一個面 S 由 4 個頂點 $v_1, v_2,$

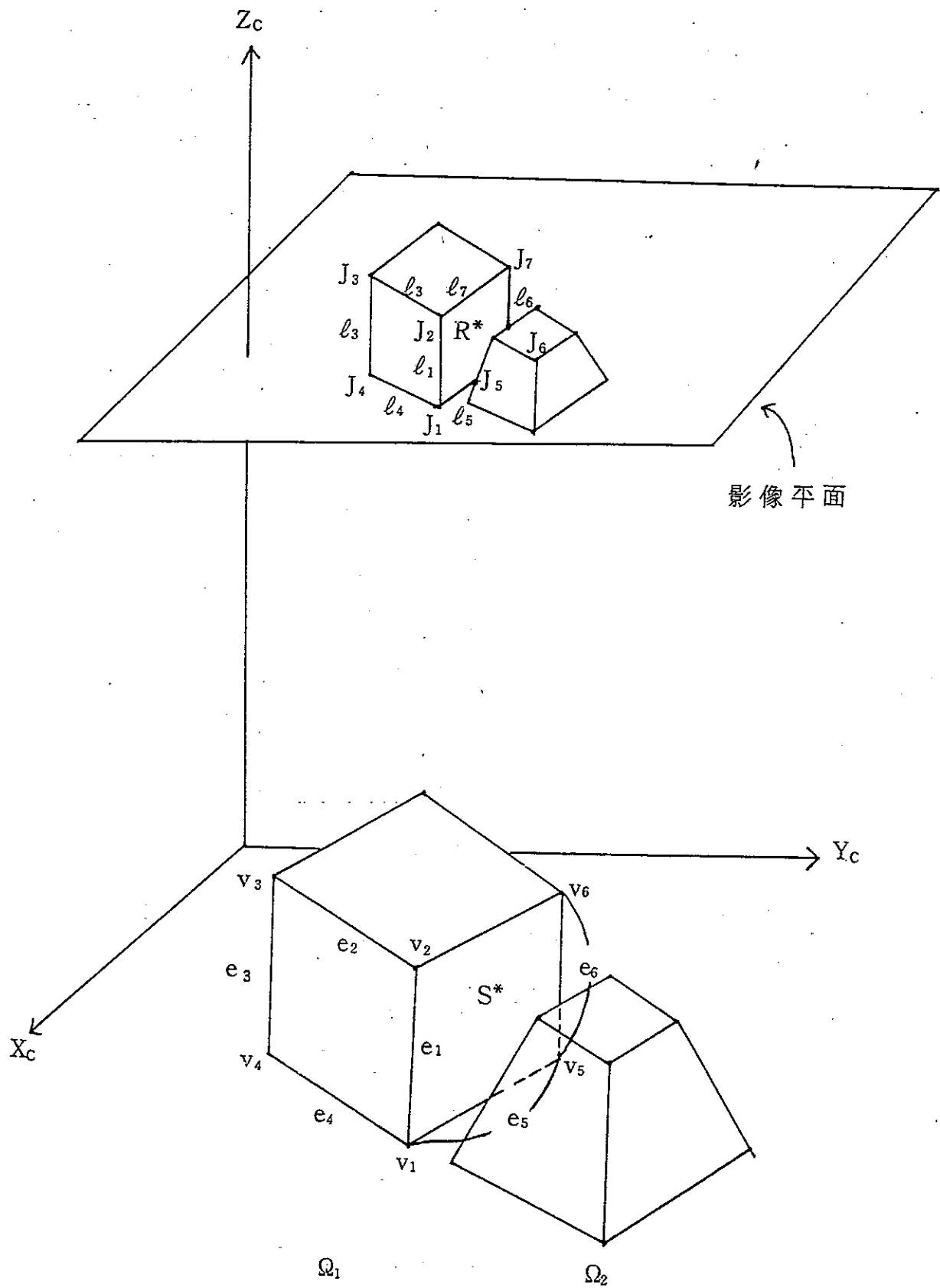


圖 3 — 1 三維物體之投影圖

v_3 及 v_4 和四個邊 e_1 , e_2 , e_3 和 e_4 所構成。而它對應到二維影像平面中的區域 R , R 是由四接點 J_1 , J_2 , J_3 , J_4 及四條邊線 L_1 , L_2 , L_3 , L_4 所構成。這時三維頂點 V_i , 和二維接點 J_i 及三維邊 e_i 和二維邊線 L_i 它們具有如下的關係：

$$J_i = O (M (V_i))$$

$$\text{且 } e_i = O (M (e_i)), i=1, 2, 3, 4$$

依定義 3.1.3 可以知道區域 R 為一個合理區域, 而其相對於三維空間上的面 S 是一個沒有被遮的平面 (not occluded planar surface)。

而當我們在觀察多個物體被置放於平坦的旋轉盤時, 極有可能發生物體 Ω_j 被 Ω_k 部份遮蔽, $j \neq k$, 也就是說 Ω_j 和 Ω_k 在二維影像平面上的投影, 極有可能出現一些不合理的區域 (illegal region), 而這就是部份被遮面 (partial occluding planar surface) 的形成。底下我們就來定義部份被遮面：

定義 3.1.4 : 假設景像中有 k 個物體,

$$\Omega = \{ \bigcup_{i=1}^k \Omega_i \}, \text{ 而對於其在二維影像平面的投影}$$

$O (M (\Omega))$, 我們定義一個部份被遮面 R 如下：

假設二維影像 $O (M (\Omega))$ 中存在某一區域 $R = (J, L)$, 其中 $J = \{ J_1, J_2, \dots, J_n \}$ 且 $L = \{ L_1, L_2, \dots, L_n \}$ 。如果區域 R 至少存在兩個相鄰的邊線 L_i 和 L_{i+1} , 而且對於 R 所屬的三

維物體 Ω_i 中之任意兩個相鄰邊 e_p 和 e_{p+1} ，它們滿足：

(1) $O(M(e_p)) \neq L_i$ 但 $O(M(e_{p+1})) \neq L_{i+1}$ ，
 或 (2) $O(M(e_p)) \neq L_{i+1}$ 但 $O(M(e_{p+1})) \neq L_i$ ，
 或 (3) $O(M(e_p)) = L_{i+1}$ 但 $O(M(e_{p+1})) \neq L_i$ ，
 或 (4) $O(M(e_p)) = L_i$ 但 $O(M(e_{p+1})) \neq L_{i+1}$ ，

則區域 R 即為部份被遮面。在本文中，我們以 R^* 來代表部份被遮面 S^* 投影於二維影像平面上的區域。

例如，圖 3-1 中，物體 Ω_1 上的面 S^* ，是由於 Ω_2 所造成的部份被遮面，而 $O(M(\Omega_1))$ 中的區域 $R^* = (J, L)$ ，其中 $J = \{ J_1, J_5, J_6, J_7, J_2 \}$ ，且 $L = \{ L_1, L_5, L_6, L_7 \}$ ，由圖可看出邊線 $L_5 \neq O(M(e_5))$ ，因為 L_5 實際上是 $O(M(e_5))$ 的部份邊線而已。同理邊線 $L_6 \neq O(M(e_6))$ 我們依定義 3.1.4 可確定 R^* 是一個部份被遮面所投影的區域。

第二節 斜量空間簡介

在這一節中，我們將詳細討論斜量空間 (Gradient Space) 上的一些常用性質。因為在此論文中，我們利用了這些性質來求出三維面的傾斜量 (gradient)，相當於三維面的法向量，並且還可知道一個三維物體其各平面在斜量空間上的幾何關係。然而，前者可用於測量物體旋轉角度，後者則用於下一節中過濾難以想像的解譯影像。由此可知，在此文中是相當倚重斜量空間這項

工具。由於本文假設三維物體是採正交平行投影的方式成像於二維影像平面上，所以底下介紹的斜量空間之性質及特性，也是在此種成像幾何之下來描述。並假設現在的照相機是位於景像 (scene) 中的座標原點，而影像平面乃位於 $Z = -1$ 的平面上。

首先我們先定義一些名詞來描述物體的幾何特性。在景像上的一點 $p = (x, y, z)$ 點 p 投影於影像平面上的點 $I = (x, y)$ 。在景像上的一向量 $V = (\Delta x, \Delta y, \Delta z)$ ，其相對於影像上的邊為 $E = (\Delta x, \Delta y)$ 。有了這些名詞及其幾何關係後，我們將定義一般面的傾斜量：

定義 3.2.1：假設某三維面 S ，若以一個 Monge Patch 表式則 S 被定義成 $-z = f(x, y)$ ，而 S 的傾斜量 $G = (p, q)$ 被定義為

$$G = (p, q) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = \left(-\frac{\partial z}{\partial x}, -\frac{\partial z}{\partial y} \right)$$

因此，全部的傾斜量所成的集合就稱為斜量空間。

對於定義 3.2.1 可得出底下的推論：

1. 如果有一個三維面 S ，包含了一個任意方向的向量 V ，則 S 的切線向量 (tangent vector) 為

$$\left(\frac{dx}{dV}, \frac{dy}{dV}, \frac{dz}{dV} \right) = \left(\frac{dx}{dV}, \frac{dy}{dV}, -p \frac{dx}{dV} - q \frac{dy}{dV} \right)$$

該切線向量在 $V = X$ 和 $V = Y$ 上的分向量分別為 $(1, 0, -p)$ 和 $(0, 1, -q)$ 。這二個分向量之外積 (cross product), $N = (p, q, 1)$ ，相當於 S 的法向量。

定理 3.2.2 : 如果有一個平面其方程式為

$$ax + by + cz + d = 0, \text{ 而 } c \neq 0$$

該平面的傾斜量為

$$(p, q) = \left(\frac{a}{c}, \frac{b}{c} \right)$$

證明：因為我們可以改寫其平面方程式為

$$z = \left(-\frac{a}{c} \right) x + \left(-\frac{b}{c} \right) y + \left(-\frac{d}{c} \right)$$

因此利用定義 3.2.1 可以得到

$$-\frac{\partial z}{\partial x} = \frac{a}{c} \quad \text{和} \quad -\frac{\partial z}{\partial y} = \frac{b}{c}, \text{ 得證。}$$

由圖 3-2 可知一個法向量在斜量空間的關係。也就是說，我們將一個面的傾斜量求出後相當於該面的法向量也被求出了。

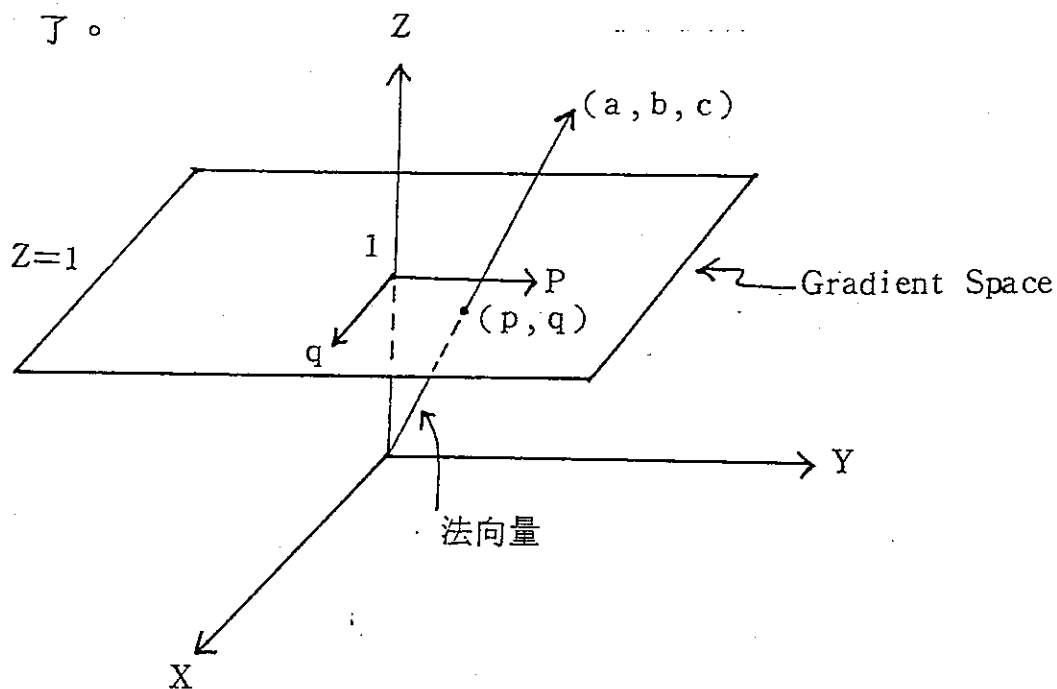


圖 3-2 法向量和斜量空間的關係

對定理 3.2.2 我們得出底下的一個推論：

1. 因為 d 對 p 和 q 沒有影響，所以全部平行的面，都具相同的傾斜量。在斜量空間上的一點 G ，代表一組平行的平面。

定理 3.2.3：假設三維空間中的一個向量 V 是 $(\Delta x, \Delta y, \Delta z)$ 則 V 的傾斜量被定義為：

$$G = (p, q) = \left(\frac{\Delta x}{\Delta z}, \frac{\Delta y}{\Delta z} \right)。$$

對於定義 3.2.3 所定義的傾斜量，其原理和定義 3.2.1 同，而底下有三個關於定義 3.2.3 的推論：

1. 三維空間中所有平行的向量都有相同的傾斜量。而三維空間中的直線之傾斜量為該線上任一向量的傾斜量。
2. 三維空間中的平面 S ，其傾斜量相同於 S 的法向量 V 的傾斜量，如圖 3-3。
3. 假設向量 $V = (\Delta x, \Delta y, \Delta z)$ 投影於影像平面上所成的邊 $E = (\Delta x, \Delta y)$ ，若 V 的傾斜量為 G 則，

$$G = \frac{E}{\Delta z}$$

而在斜量空間上的原點到 G 所連成的直線必會平行於影像平面中的 E 。如圖 3-4。

定理 3.2.4：假設一個平面 S 中包含有一向量 $V = (\Delta x, \Delta y, \Delta z)$ ，而 S 的傾斜量 $G_s = (p_s, q_s)$ 則 S 與 V 的關係如下：（“ \cdot ”表內積）

$$G_s \cdot (\Delta x, \Delta y) = -\Delta z \quad (3-2-1)$$

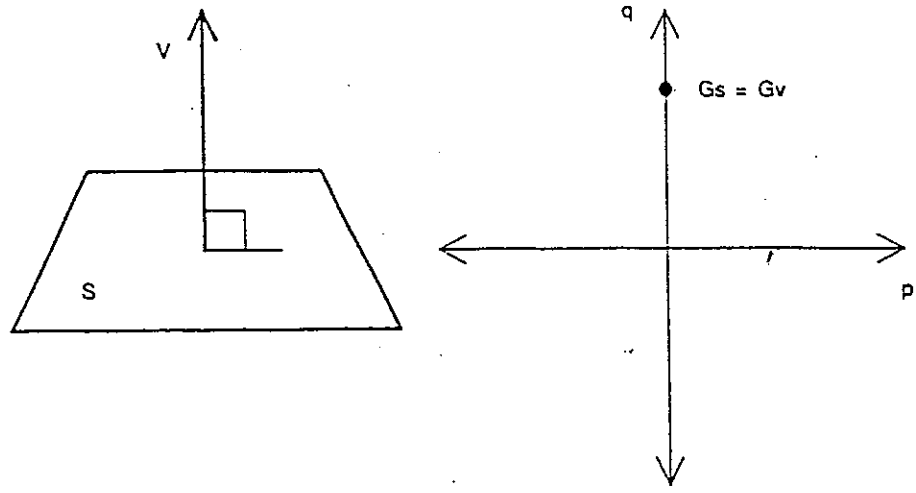


圖 3 - 3

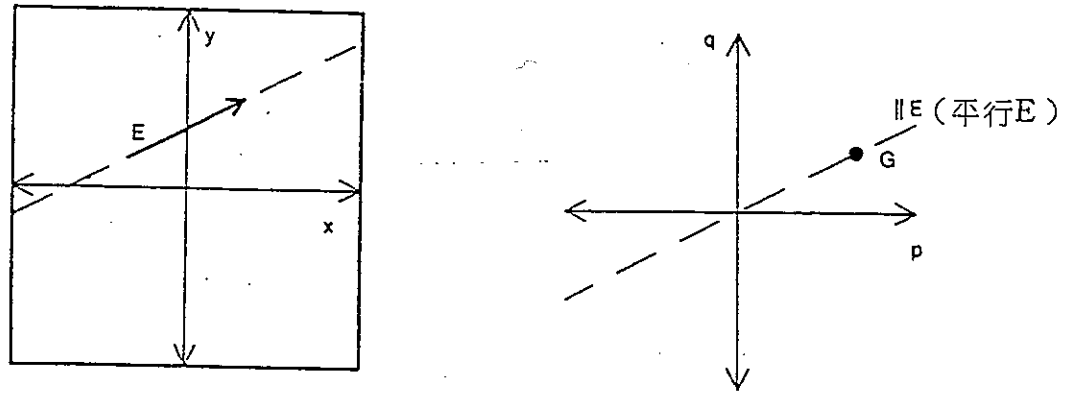


圖 3 - 4

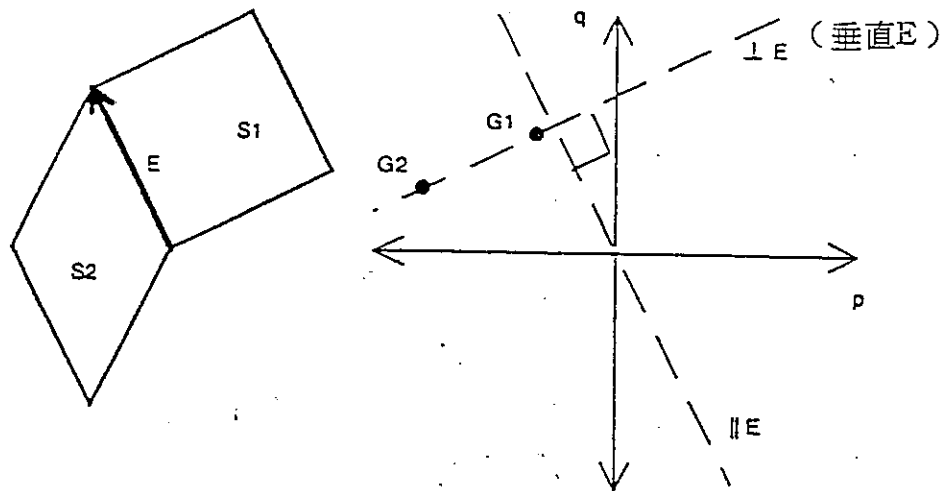


圖 3 - 5

證明：因爲 S 的法向量， $N_s = (p_s, q_s, 1)$ ，必須與向量 V 互相垂直，所以 N_s 和 V 的內積 (inner product) 會爲 0，如下式：

$$N_s \cdot V = (p_s, q_s, 1) \cdot (\Delta x, \Delta y, \Delta z) = 0$$

$$p_s \Delta x + q_s \Delta y + \Delta z = 0$$

因此， $G_s \cdot (\Delta x, \Delta y) = -\Delta z$ ，得證。

式子 (3-2-1) 在此論文中是相當有用的，因爲此文中所允許的物體其本身是由平面所構成的。而我們知道，每個平面都含有與相鄰面所共同擁有的相鄰邊。利用這種關係，我們可以了解相鄰的二個平面和其相鄰邊，三者在此斜量空間中的幾何性質。

定理 3.2.5：假設有二個相鄰平面 S_1 和 S_2 ，它們相鄰邊爲 $V = (\Delta x, \Delta y, \Delta z)$ ，而 S_1 和 S_2 的傾斜量分別爲 $G_1 = (p_1, q_1)$ 和 $G_2 = (p_2, q_2)$ 。如圖 3-5，若 V 投影於影像平面上所形成的邊線 $E = (\Delta x, \Delta y)$ 則 E 必須垂直於 G_1 和 G_2 在此斜量空間上的連線。

證明：由定理 3.2.4，我們可得到

$$-\Delta z = (p_1, q_1) \cdot (\Delta x, \Delta y) = G_1 \cdot E$$

$$= (p_2, q_2) \cdot (\Delta x, \Delta y) = G_2 \cdot E$$

$$0 = (G_1 - G_2) \cdot E$$

相當於 $(G_1 - G_2) \perp E$ ，

這裏的 $(G_1 - G_2)$ 是指斜量空間上，由 G_1 到 G_2 所連成的直線中的一個向量，得證。

對定理 3.2.5，我們可得出底下的推論：

1. G_1 和 G_2 在斜量空間上所連成的直線 $L = \overleftrightarrow{G_1G_2}$ ，必定會與影像平面上 S_1 和 S_2 所相鄰邊 E 互相垂直，i.e., $L \perp E$ ，而假如 S_1 和 S_2 相鄰邊 E 是凸形性質 (convexity) 則影像平面上若 S_1 在 S_2 的右邊，那麼在斜量空間上， G_1 是在 G_2 的右邊。但是相鄰邊 E 若是凹狀性質 (concavity)，那麼 G_1 就會在 G_2 的左邊。如圖 3—6。
2. 假如二平面 S_1 和 S_2 相鄰邊是凸狀性質，並且 S_1 和 S_2 在三維空間中的交角的形狀愈尖銳的情況下，則其相對於斜量空間上， G_1 和 G_2 的距離必然會愈遠。反之， S_1 和 S_2 的交角愈趨於鈍角的情況下， G_1 和 G_2 的距離將愈縮短，如圖 3—7 中最上方二組。假如 S_1 和 S_2 的相鄰邊 E 是凹狀的性質，則 S_1 和 S_2 在三維空間的交角大小與斜量空間上 G_1 和 G_2 的距離長短變化之關係恰和前者相反。如圖 3—7。

定理 3.2.6：假如有二個互相垂直的向量 $V_1 = (\Delta x_1, \Delta y_1, \Delta z_1)$ 和 $V_2 = (\Delta x_2, \Delta y_2, \Delta z_2)$ 。它們的傾斜量分別為 $G_1 = (p_1, q_1)$ 和 $G_2 = (p_2, q_2)$
則 $G_1 \cdot G_2 = -1$

證明：因為 $V_1 \perp V_2$ ，所有下列關係：

$$V_1 \cdot V_2 = \Delta x_1 \Delta x_2 + \Delta y_1 \Delta y_2 + \Delta z_1 \Delta z_2 = 0$$

對式中的每項除上 $\Delta z_1 \Delta z_2$ 可得

$$p_1 p_2 + q_1 q_2 + 1 = 0 \quad \text{相當於}$$

$$G_1 \cdot G_2 = -1, \quad \text{得證}$$

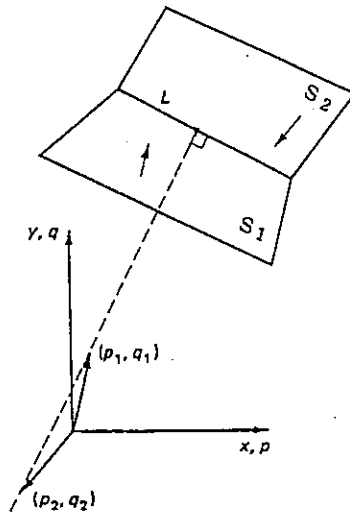


圖 3 — 6 相鄰邊為凹狀性質時兩平面之傾斜量的關係

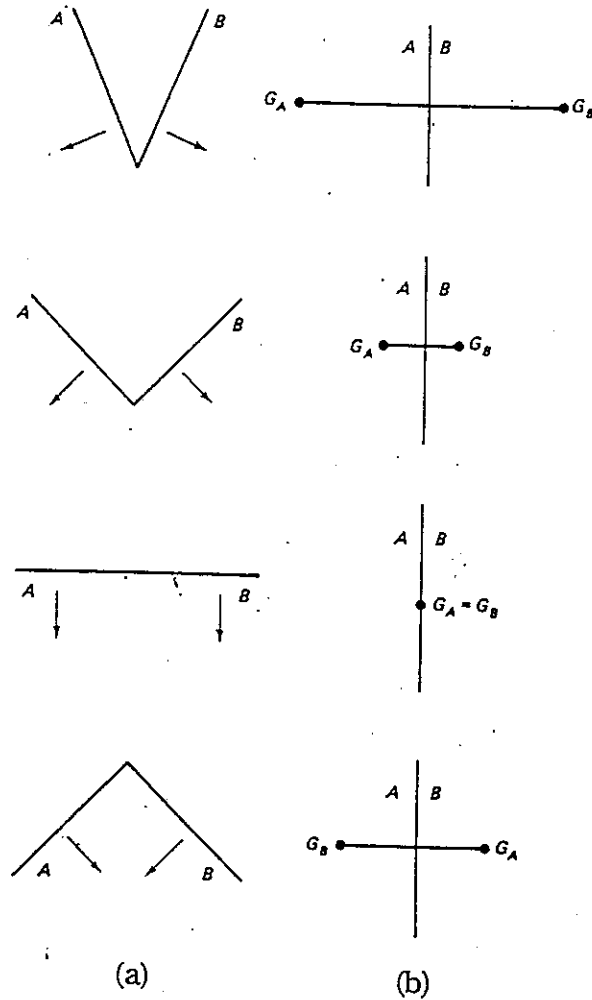


圖 3 — 7 兩平面的傾斜量與其夾角大小之關係圖

對定理 3.2.6，我們有底下的推論：

1. 假如二個平面互相垂直；則它們的傾斜量會有上述定理的結果。原因是這二個平面的法向量也是會互相垂直⁶
2. 假如一個平面包含有一個向量，則此平面和該向量兩者的傾斜量會有上述的結果。原因是平面上所包含的任一向量皆會與此平面的法向量互相垂直。

上述的定義和定理及其推論，往後在此論文將會引用到，而至於更詳盡，更多有關斜量空間的介紹可在 shafter [23] 所著一書上找到。

第三節 折紙狀物體的解譯影像之形成

在這節中，我們將先描述折紙狀物體所具有的特性，接著我們簡述依 A.I. 的標線法則來解譯這類物體所形成的二維影像，並且由於在標線期間將引用斜量空間上的性質來過濾掉一些奇異的解譯影像之形成，亦可加速標線程序的標線速度。

在此篇文章中所採用的物體是 Kanade [13] 在其論文中所提出的折紙狀物體 (Origami world)。所謂的 Origami 一詞乃源於古日本的一種折紙的手工藝品，它是由一張紙所折出的物體，如圖 3-9 (a)，我們很容易的知道圖中是一隻紙折成的鶴 (an Origami crane)。而這種折紙狀的物體是以平面紙張為主所構成的，並能單獨的站立著 (Stand-alone)。於此文中我們將不限制物體是必須由平面紙張所構成，我們知道，對於一般的三角平面多面體而言，其強調的物體是以立體實心物為主的模式 (

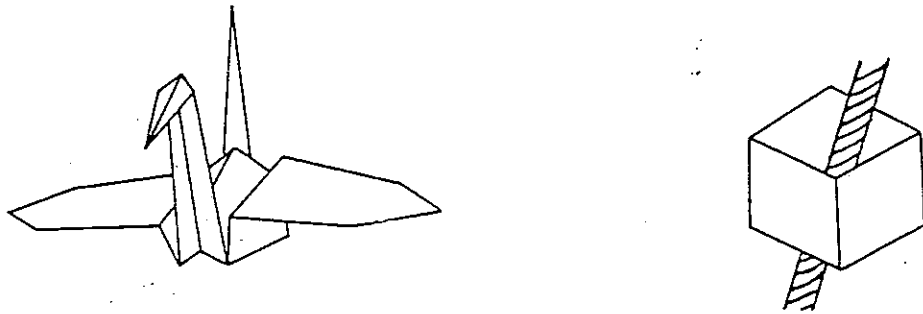


圖 3 — 9 (a) Origami Crane , (b)上、下透空之折紙物

solid-object Oriented model)，這類物體的每一個頂點必須恰由三個平面相交而成。但是折紙狀物體的各頂點之形成條件如下：

- (1)每個頂點都是由不超過三個不同方向的平面所相交而成。
- (2)每個頂點也是由不超過三條不同方向的邊所相交而成。

我們稱這類頂點為 up - to - 3 - surface 頂點。由上述的條件，我們可以知道，up - to - 3 - surface 頂點將包括，由一面所形成的頂點 (1 - surface vertices)，二面相交所構成的頂點 (2 - surface vertices) 和由三面相交所形成的頂點。因此，由 up - to - 3 - surface 頂點所衍出來的物體，其形狀將比三角頂點 (trihedral vertices) 所衍生出來的物體之形狀有更豐富的物體形狀。如圖 3 — 10 所示，其三個線繪影像中，圖 (a) 表擬似一個無蓋

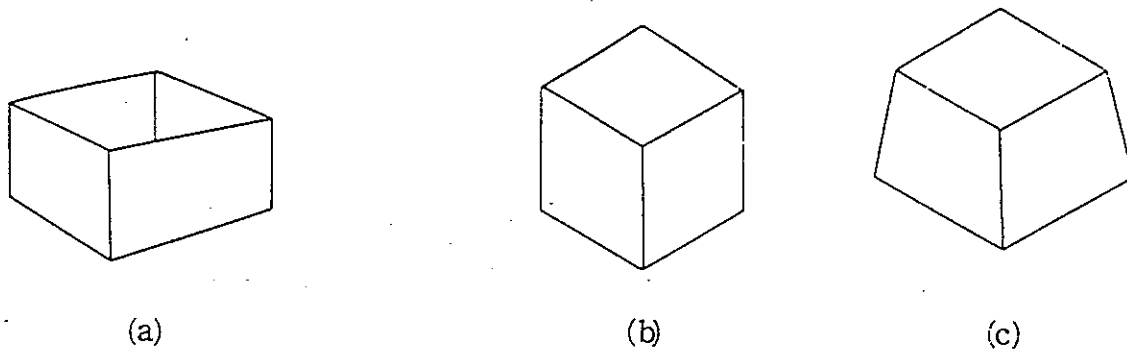
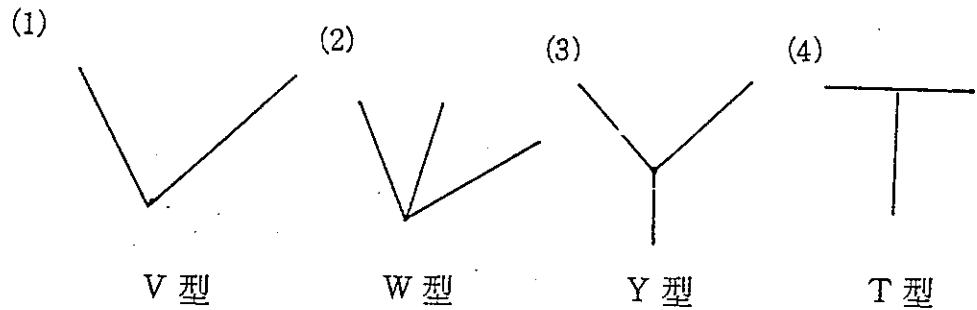


圖 3 — 10

、中空的箱子形狀，(b)為擬似正立方體的形狀，(c)為擬似梯形體的形狀。

既然我們已經知道折紙狀物體的特性，接著就是要對其所形成的二維影像來標線，以瞭解這影像中物體上各面在三維空間上的立體幾何關係。對於這種人工智慧學上的標線方法，最早是由 Huffman [2] 和 Clowes [3] 在 1971 年提出。本來 Guzman [1] 定義了一些接點型態（例如 L，W，和 Y 型），並且 Guzman 依據這些接點型態提出很多啓發式規則，將一張線繪影像中的各物體所屬區域逐一分開，使各物體只能擁有屬於自己的區域。而這些啓發式規則大多缺乏理論根據及缺乏物體在三維空間中實際的立體幾何意義。所以 Huffman 和 Clowes 就改進 Guzman 的方法。

他們實際觀察線型三角多面體 (line-drawing trihedral scene) 在三維空間中，物體的實際可能組合，歸納出四種不同的接點型態和三種邊線型態。此四種接點型態如下：



上圖中，V型接點的產生在於二維影像中存在某一個接點J，恰巧只由兩條相鄰邊線 L_i 和 L_{i+1} 所相交而成，並且接點J必須是邊線 L_i 和 L_{i+1} 中的一個端點。另外三種型態的接點均恰由三條邊線相交而成。如果接點J恰由三條邊線所相交而成，並且這三條邊線之中，存在有兩條邊線，它們之間的夾角恰好等於180度，則接點J即為T型接點。如果三條相交邊線之中任兩條邊線的夾角小於180度，則由此三條相交而成的接點J即為Y型接點。假如三條相交的邊線之中，存在兩條邊線它們之間的夾角小於90度，則由這三條邊線相交而成的接點為W型接點。

至於 Huffman 和 Clowes 所歸納出之三種邊線型態及其相對應的標線型態為：

- (1) 凸型邊線 (convex line-segment)，表示，共同擁有這個邊線的兩個區域，在二維影像中都是可見的 (visible)，而且這邊線所相對應於三維空間上的邊 (edge) 是凸型邊 (convex edge)。以符號“+”來表示此邊線的標線

型態。

(2)凹型邊線 (concave line-segment)，表示共同擁有這個邊線的兩個區域在二維影像中也都是可見的，而且這邊線相對應於三維空間的邊是凹型邊 (concave edge)。以符號“—”來表示其標線型態。

(3)邊界線 (occluding line-segment)，表示共同擁有這個邊線的兩個區域在二維影像中僅有一個區域是可見的，而另外一個區域被完全遮蔽，而其標線符號用箭號“→”來代表，當觀察者順著箭頭方向時，箭頭的右邊的區域代表可見的，箭頭的左邊是指被完全遮蔽的區域。但是也有可能發生另一種邊界線的型態為“←”，剛好與第一種邊界線型態有相反情況，換句話說，後者的左邊是指可見區域，右邊為被完全遮蔽的區域。

所以對於二維線繪影像中，邊線的標線型態總共有“+”，“—”，“→”和“←”四種情形。故對於V型接點來說，其相鄰邊線的標線型態有 $4^2 = 16$ 種可能組合。同理，對Y型，T型，和W型接點做標線時有 $4^3 = 64$ 種標線組合。因此，當我們在對某一接點的相鄰邊線做標線時，僅有208種可能的情形。而Winston在[4]的書中列出這208種可能的組合當中，只僅有18種存在於真實世界中。

Waltz [4]於1975年延伸了Huffman和Clowes所歸納出的接點型態和邊線型態。Waltz擴充了接點型態和邊線型態的種類，並且對三角多面體提出一個很有效率的演繹法，

來對這類物體所形成的二維影像標線。這個演繹法首先在二維線繪影像中，任意找出一個接點 J_1 ，然後考慮接點 J_1 所屬接點型態的所有標線組合 (line-label combinations)。對接點 J_1 的所有相鄰接點 J_i, J_{i+1}, \dots, J_R 的標線組合將受制於接點 J_1 的標線組合理態，這將減少接點 J_i, J_{i+1}, \dots, J_R 的許多可能標線組合理態，接著我們再考慮 J_i, J_{i+1}, \dots, J_R 相鄰接點間的標線組合，這時候 $J_1, J_{i+1}, J_{i+2}, \dots, J_R$ 之可能標線組合理態又可以更加減少。依此類推，重覆上述步驟。最後即可將整張二維影像標線完畢。

Kanade [13] 依據 Huffman-Clowes-Waltz 的理論，對折紙狀物體所產生四種接點型態和三種邊線型態，建造出可能實際存在的所有接點標線組合情況，如表 3 — 1。其與三角多面體的比較如表 3 — 2 所示。至於 Kanade 對折紙狀物體的標線程序是採用 Waltz 的標線演繹法和由表 3 — 1 的接點標線組合所衍生的擴增接點字典 (Augmented Junction Dictionary)，來完成一張折紙狀物體的二維影像之標線。而所謂的擴增接點字典就是對表 3 — 1 中的每個接點標線組合中，再加上一些額外的資訊。這些資訊就是接點標線組合中，各相鄰面的傾斜量和其相鄰邊線的立體實際意義，二者在斜量空間中的限制條件。如圖 3 — 11 (a) 中最左邊的圖代表 Y 型接點標線組合中的一種，而在圖的最右邊表示其額外需加入的限制條件。在圖中間表示每個邊線型態連結一對區域，而各區域所相對的限制條件就在圖的最右邊所示。這裏的區域 R_1, R_2 和 R_3 都屬可見的區域。如果圖 3 — 11 (b) 中的

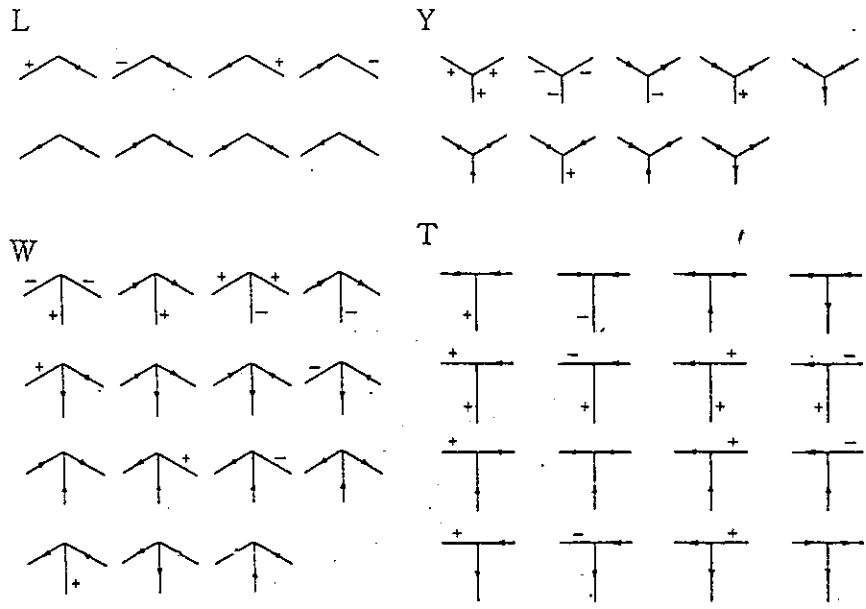


表 3 - 1

Junction type	Huffman-Clowes dictionary	Origami world dictionary
L	6	8
W	3	15
Y	3	9
T	4	16

表 3 - 2

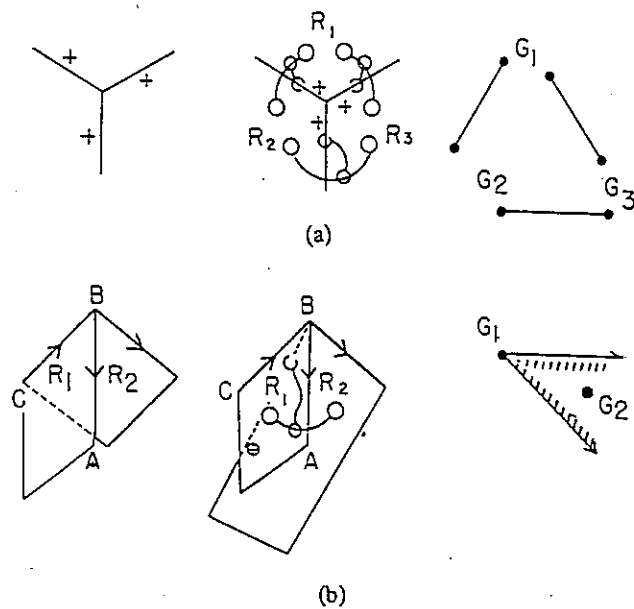
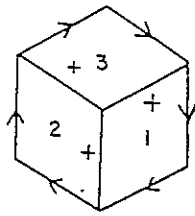


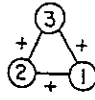
圖 3 - 11

最左圖 R_1 是一完全可見面，但 R_2 卻是屬於不合理區域（同第一節中所描述的），也就是說 R_2 相當於一個部份被遮面。而且 R_1 和 R_2 相接的邊線也被遮蔽了。在此圖中我們可以想像成，有一張紙其中的區域 R_1 是沿著邊線 BC 折向觀察者所形成，並且 R_1 將 R_2 的部份區域遮蔽了。如果圖 3—11(b) 中間的圖是表示一般接點 B 的標線組合，而區域 R_1 和 R_2 的相鄰邊線（被遮蔽）會落於 $\angle ABC$ 的範圍之內。假設 R_1 和 R_2 相交的邊線型態為凸型性質則接點 B 的標線組合所必須加入 R_1 和 R_2 在斜量空間上的限制條件即圖 3—11(b) 中最右圖所示。為什麼我們在對折紙狀物體的二維影像做標線時，還必須去考慮物體上各面的傾斜量在斜量空間中的位置幾何關係來做為限制條件呢？原因是：折紙狀物體的構成之條件較三角多面物體的形成條件還來得不嚴格使然。所以在解譯折紙狀物體的線繪影像時，乃利用這種擴增接點字典，來完成標線工作，將使其不致產生不合理的解譯影像。

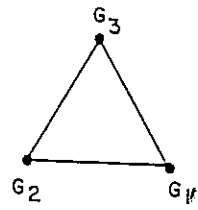
當一張解譯影像完成時，將會伴隨產生一個限制條件，這個限制條件乃表示物體上各面的傾斜量在斜量空間上的位置幾何關係。而 Kanade [13] 在標線過程中乃以一個稱之為面連接圖（surface connection graph），簡稱 SCG，來代表。該限制條件，並代表物體上各面間相鄰關係。SCG 是一個標線圖（label graph），也就是說，圖裏的節點（node）代表面，節點間的連結線（link）代表二相鄰面之共同邊線的實際意義。如圖 3—12(b) 所示。而圖 3—12(a) 乃對圖 3—10(b) 標線後的一種解譯影像，而圖 3—12(c) 為標線後，物體各面的傾斜量在斜量空間中的位置



(a)

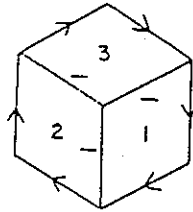


(b)

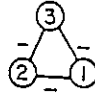


(c)

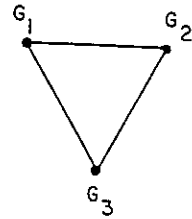
图 3-12



(a)

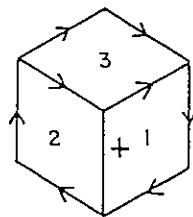


(b)

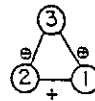


(c)

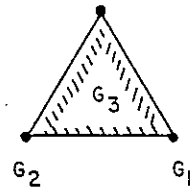
图 3-13



(a)

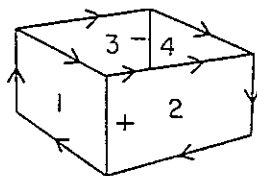


(b)

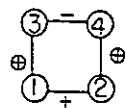


(c)

图 3-14



(a)



(b)

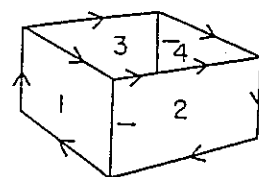


图 3-15

幾何關係，而圖中三角形的大小和位置皆未定。其實際位置及大小，我們將在第四章的第一節中來決定。

用上述的標線演繹法及擴增接點字典來對同一張折紙狀物體的二維影像標線後，仍然可能產生不只一種的解譯影像。如圖 3—12(a) 至 3—14(a) 都可能是圖 3—10 (b) 的合理解譯影像，又如圖 3—15 (a) , (b) 也可能是圖 3—10(a) 的合理解譯影像。由於在此論文中所允許置於轉盤的折紙狀物體，其上、下可以是透空的立體物，但其周圍的平面必須要構成一個封閉空間。如圖 3—9 (b)。所以圖 3—13(a) 和 3—15(b)，在此篇論文中將不予考慮。而圖 3—14(a) 是屬奇異的解譯影像，此類解譯影像所代表的三維意義就無法通過 Kanade [15] 的平行線啟發式規則。因為 Kanade 的平行線啟發式規則是：「如果在影像中的兩條線是平行者，則在不失去一般性的情況下，此二線在三維空間中亦是兩條平行線。」

而在正交投影的模式下，上述規則的反命題也同樣是可以成立的。現在我們來解釋為什麼圖 3—14(a) 失敗於這條規則上，首先重劃圖 3—14(a)，如圖 3—16，在圖 3—16(a) 中 S_1 和 S_2 是兩個全部見面，並且它們共有一條凸形性質的相鄰邊，卻是一個部份被遮面。這三個面 S_1 , S_2 和 S_3 的傾斜量分別為 G_1 , G_2 和 G_3 , 其在斜量空間中的位置幾何關係如圖 3—16 (b) 所示。 G_3 應該要落於三角形內部或是在 G_1 和 G_2 所連成的線段上，而這將不合乎平行線的啟發式規則，因為 G_3 平面有兩個邊（在線段 V_6V_7 和 V_2V_7 的正下面）相當於嵌入到 S_1 和 S_2 所圍成的內部區域。可是

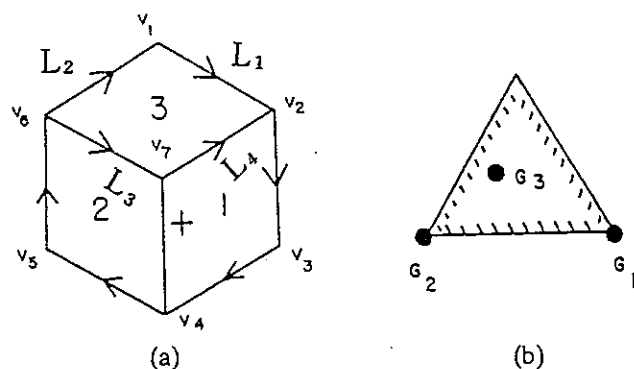


圖 3—16

邊線 V_1V_2 和 V_6V_7 是互相平行的，這將要求 G_2 和 G_3 的連線必須垂直於邊線 V_6V_7 。而另一方面，邊線 V_1V_6 和 V_2V_7 也是互相平行，這也要求 G_1 和 G_3 的連線須垂直於邊線 V_2V_7 。此時，我們發現同時要讓這兩種情況成立是不可能的，除非 G_3 的位置能落於頂部的三角形點上，但是這種情況就相當於圖 3—12(a) 的解譯影像才可能發生。所以圖 3—14(a) 的解譯影像將無法通過平行線規則。因此在此文中將去除此類奇異解譯影像。而圖 3—15(b) 也是一種奇異解譯影像。

在此我們如何知道區域 S_3 是一個部份被遮面呢？底下將介紹一個由 Liaw [27] 所提出的法則來判斷之。首先，我們將定義所謂的解譯區域 (interpretative region)：

定義 3.3.1：假設存在某個經上述標線程序後的二維影像 $O (M$

(Ω)), 而 $O(M(\Omega))$ 中存在某個區域 $R_i = (J, L)$, 其中 $J = \{J_1, J_2, \dots, J_n\}$ 且 $L = \{L_1, L_2, \dots, L_n\}$, $n \geq 3$ 。我們定義一個區域 R_i 相對應的解譯區域, IR_i , 如下:

$$IR_i = (J, L, h),$$

其中 $J = \{J_1, J_2, \dots, J_n\}$,

且 $L = \{L_1, L_2, \dots, L_n\}$,

而 h 是一個函數, 從邊線集合 L 對映到標線型態集合 $\{“+”, “-”, “\rightarrow”, “\leftarrow”\}$ 上, 且

$$h(L_i) = \begin{cases} “+”, & \text{若邊線 } L_i \text{ 的標線型態爲 “+” 的凸形邊線} \\ “-”, & \text{若邊線 } L_i \text{ 的標線型態爲 “-” 的凹形邊線} \\ “\rightarrow”, & \text{若邊線 } L_i \text{ 的標線型態爲 “\rightarrow” 的邊界線,} \\ & \text{且箭頭的方向從接點 } J_i \text{ 到接點 } J_{i+1}。 \\ “\leftarrow”, & \text{若邊線 } L_i \text{ 的標線型態爲 “\leftarrow” 的邊界線,} \\ & \text{且箭頭的方向從接點 } J_{i+1} \text{ 到接點 } J_i。 \end{cases}$$

法則 3.3.1: 假設某一張經標線後的二維影像 $O(M(\Omega))$ 中存在一個解譯區域 $IR_i = (J, L, h)$, 其中 $J = \{J_1, J_2, \dots, J_n\}$ 且 $L = \{L_1, L_2, \dots, L_n\}$, $n \geq 3$ 。如果存在某個邊線 L_i , $1 \leq i \leq n$, 使得 $h(L_i) = “\leftarrow”$, 則, 與解譯區域 IR_i 相對應的區域 $R_i = (J, L)$ 必爲一個部份被遮面。

現在舉一個例子來說明上述的法則, 如圖 3-16(a) 中的面 S_3 而言, 其相對應的解譯區域 $IR_3 = \{J, L, h\}$ 且 $L = \{L_1, L_2$

, L_3 , L_4 } 和 $J = \{ J_1 = V_6, J_2 = V_1, J_3 = V_2, J_4 = V_7 \}$ 。我們發現其中邊線 L_3 和 L_4 的標線型態為“←”，所以依據上述法則， S_3 相對於一個部份被遮面。 S_3 是同時給面 S_1 和 S_2 所部份遮蔽。至於在多面體中找出被遮面的法則在 Liaw [27] 的論文中有詳盡的討論。

第四節 部份被遮面發生的原因及限制

由於此文中我們允許轉盤上所置放的物體數量，不再僅僅是單一物體，也可以置放不互相接觸的多個物體。因此對固定於某一位置的照相機而言，在其所拍入的影像中，物體間的遮蔽現象勢將無法避免。在本節中，我們首先將描述兩類部份被遮面形成的原因，然後再對此文中所能處理的部份遮蔽現象加以限制，最後依據上述的限制條件來討論多折紙狀物體的解譯影像之形成，並且利用一個啓發式規則去猜測出有些部份被遮面其被遮部份之形狀。

在此文對折紙狀物體的假設條件下，部份被遮面發生的原因有二：(一)由物體本身的構造而發生的。例如圖 3—15(a) 中，一個無頂蓋的箱狀物體，其部份被遮面 S_3 和 S_4 的形成就是因全部可見面 S_1 和 S_2 所遮而產生的。而且 S_1 和 S_2 與 S_3 和 S_4 皆屬於同一物體的四個可見面。因此對這種原因所產生的部份被遮面，在此論文中就稱之為自發性的部份被遮面。另外，(二)：這類部份被遮面發生的原因乃是由於周圍相鄰近而不互相接觸之不同物體所產生的遮蔽現象。如圖 3—17 中有兩個物體 Ω_1 和 Ω_2 ，對圖左的

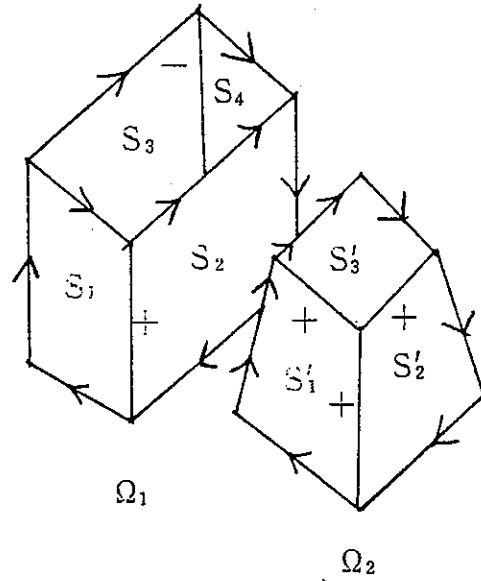


圖 3 — 17

Ω_1 而言 $\Omega_1 = \{ S_1, S_2, S_3, S_4 \}$; 另一物體 $\Omega_2 = \{ S'_1, S'_2, S'_3 \}$, 其中 Ω_1 的 S_2 被 Ω_2 給部份遮蔽了, 而形成一個部份被遮面。這種非由物體本身所造成的遮蔽現象乃因鄰近物體的遮蔽所產生的部份被遮面, 我們在此稱之為非自發性的部份被遮面。

對這類部份被遮面而言, 我們將以物體所出現的全部可見面數 (包括假想的平面, 如圖 3 — 17 中的 Ω_1 可想像其頂蓋為一個全部可見平面) N , 來歸類, 進而去限制物體被遮蔽的程度。因為折紙狀物體都是由 up — to — 3 — surface 頂點所衍生出的物體形狀, 所以影像中, 若發生有由 4 條或 4 條以上的接線相交而成的接點型態, 如 X 型, K 型, …… , 等, 則這類接點都將視為假接點, 且其將形成一張不真實的線繪影像 (not realizable line-drawing image)。Sugihara [17] 已經有證明出。然而, 假如我們有獲取到這類不真實的線繪影像, 則我們將稍微變換轉盤位

置，再重拍入影像，使影像中不會再出現有上述的不真實影像發生。現在我們就可以來討論物體被遮蔽現象的限制情況。在此文中主要分成兩種情況。第一種情況是當 $N = 2$ 或 3 時：只允許物體本身周圍的全部可見面 S_i 可產生非自發性的部份被遮面， S'_i ， $1 \leq i \leq N$ 。並且，必須能夠猜測出每個 S'_i 其被遮部份的形狀。如圖 3—18(a) 中，二個部份被遮面的形狀可以被預測出。圖 (b) 物

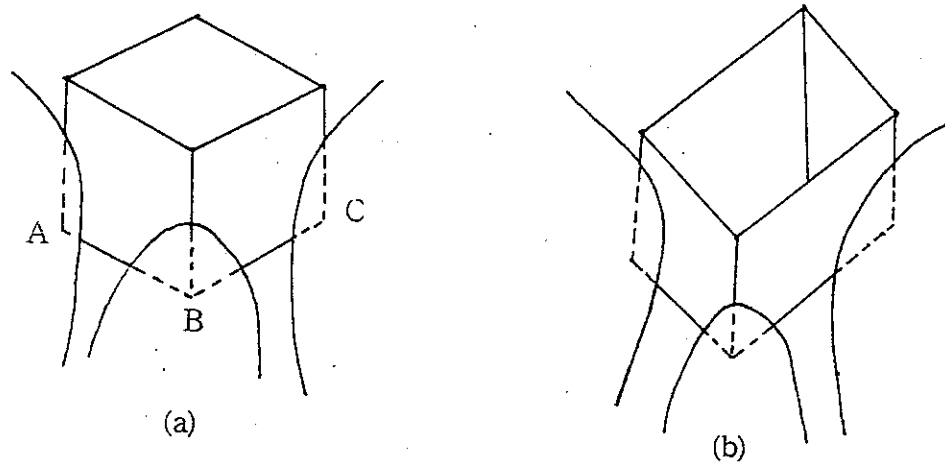


圖 3—18

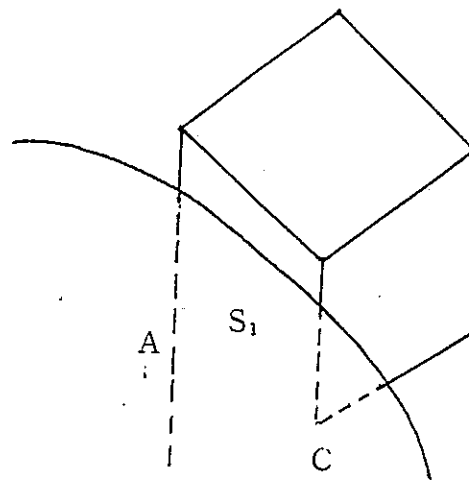


圖 3—19

體被遮之限制情況與圖(a)相同。也就是我們不允許如圖 3—19 中的被遮情況產生。因面 S_1 的二頂點 A, C 同時被遮, 這將無法猜測出這個部份被遮面的被遮部份之形狀。若發生這種被遮現象, 其將無法有足夠的條件來計算出物體上所有全部可見面的傾斜量。另外, 第二種情況是 $N \geq 4$: 對物體上全部可見面 S_i , $1 \leq i \leq N$, 可能產生非自發性的部份被遮面, S'_i , 但是其中必須有三個相鄰的部份可見面 S'_{kr} , $1 \leq k_r \leq N$, $1 \leq r \leq 3$, 能被猜測出其被遮部份的形狀, 而剩餘的 S'_i , $\forall i \neq k_r$, 則不在此限。這裏還須再注意的就是對所有的 S'_{kr} , 其中有一個部份被遮面 S'_{kr_0} , $1 \leq k_{r_0} \leq N$, $1 \leq r_0 \leq r$, 必須是物體上與最多面相鄰者。而 S'_{kr_0} 的選擇函數如下:

$$\text{若 } g_2(S'_i, S'_j) = \delta(g_1(S'_i), g_1(S'_j)), \\ 1 \leq i, j \leq N,$$

這裏的 $g_1(S'_k) = S_k$, $1 \leq k \leq N$, 和

$$\delta(g_1(S'_i), g_1(S'_j)) \\ = \begin{cases} 1, & \text{假如 } g_1(S'_i) \text{ 和 } g_1(S'_j) \text{ 互為相鄰面。} \\ 0, & \text{否則。} \end{cases}$$

則我們選擇 $S'_{kr_0} = \text{MAX } \Sigma(g_2(S'_i, S'_j))$ 。

$$1 \leq i, j \leq N, \quad i \neq j$$

現在舉一個例子, 如圖 3—20 中的物體 $\Omega = \{ S_i \mid i=1, 2, \dots, 9 \}$, $N=9$, 而每個 S_i 的傾斜量為 G_i , 其可能產生的非自發性部份被遮面 S'_i 。假如 $S'_{kr_0} = S_{k_1} = S_1$, $S_{k_2} = S_2$, $S_{k_3} = S_3$ 的傾斜量分別為 G_1 , G_2 , G_3 已知。而圖 3—20(a) 經標線程序後, 知

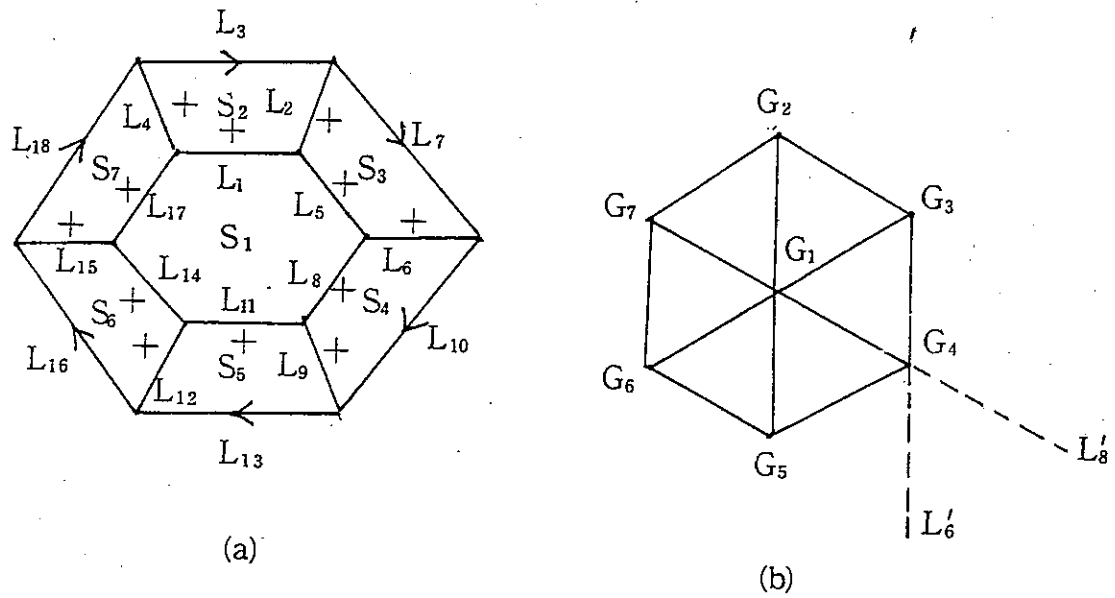


圖 3 — 20

道各面的傾斜量在斜量空間上的位置幾何關係如圖 3 — 20(b)。因此如能計算出 G_1 , G_2 和 G_3 在斜量空間上的實際位置，那麼雖無法猜測 S'_i , $i = 4, 5, 6, 7$, 其被遮部份形狀，以算出其 G_i , $i = 4, 5, 6, 7$ 。但是我們卻能從圖 3 — 20(b)中，引用定理 3.2.5 及其推論，來推算出 G_i , $i = 4, 5, 6, 7$ 。因為 G_1, G_2, G_3 已知，而 L_6 和 L_8 的方向可從影像（圖 3 — 20(a)）中測得，並知道 $h(L_6) = h(L_8) = “+”$ 。根據定理 3.2.5 及其推論，如果在斜量空間上，過 G_1 和 G_3 分別做一直線 L'_6 和 L'_8 , 並且 L'_6 和 L'_8 的方向分別垂直於 L_6 和 L_8 。如此一來， L'_6 和 L'_8 將相交於一點 G_4 , 該點就是 S'_4 的傾斜量，對於 G_5 , G_6 , G_7 的求法亦復如此。

如果物體間的被遮現象能符合上述的假設時，則對一張多物

體的線繪影像做標線時，依然可以用上述的標線程序來完成一張解譯影像。原因是：該擴增接點字典是實際去觀察折紙狀物體在三維空間中可能出現的所有接點標線組合而成的集合。所以將不會形一張奇異的解譯影像。如圖 3-21(a)中，如果現在標線程序

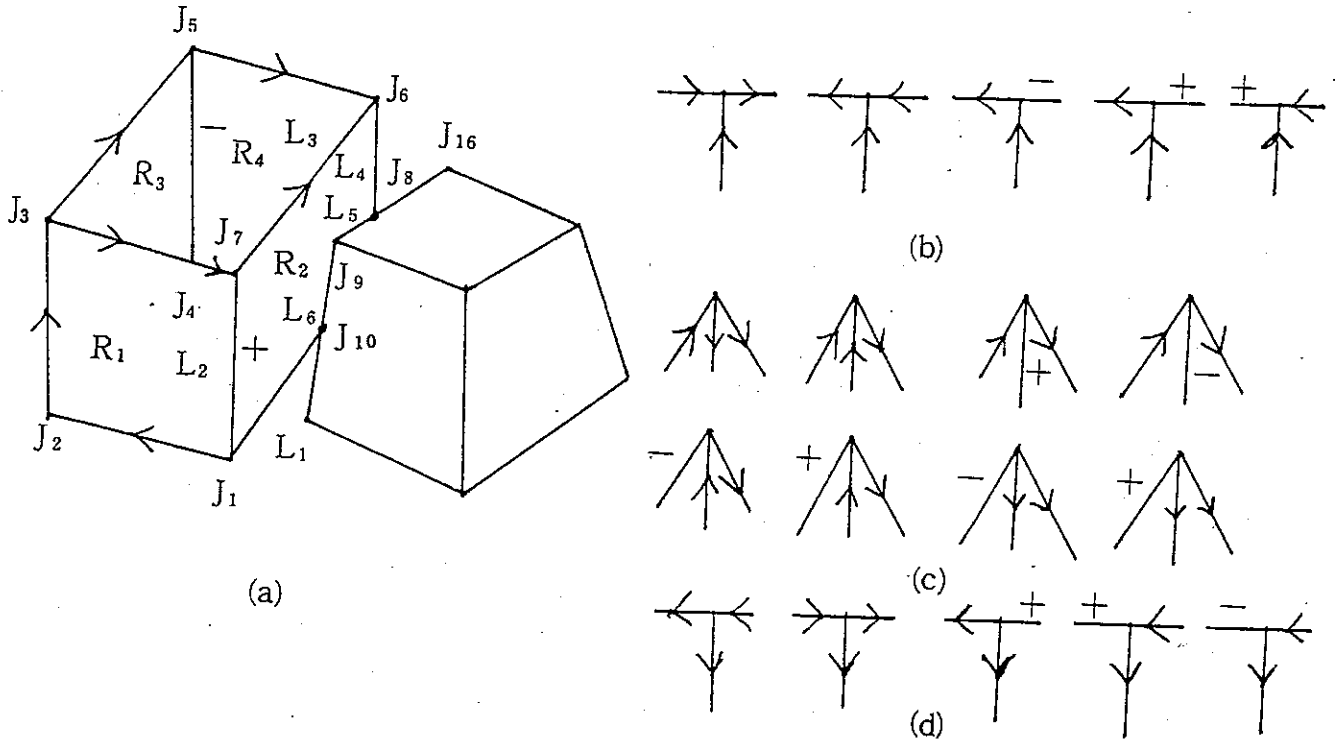
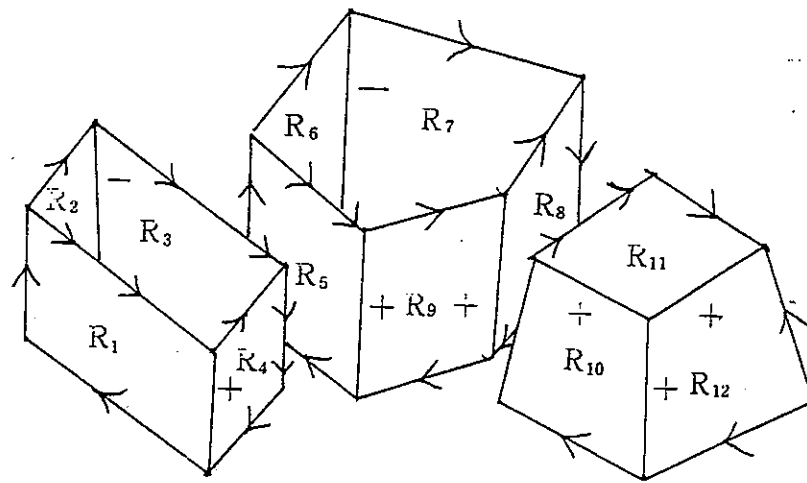


圖 3 - 21

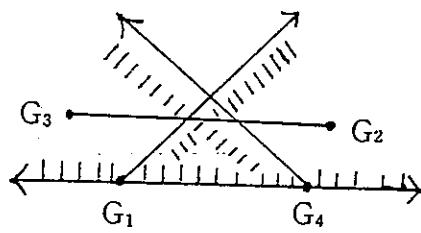
正對區域 R_2 中 J_8 的T型接點做標線，則目前有5種標線組合的選擇，如圖 3-21(b)，而我們很容易知道線段 J_8J_{16} 是邊界線，並且其接線型態為“ \rightarrow ”，因此 J_8 的接點標線組合只剩下圖 3-21(b)中的第一種，其次考慮 J_9 的W型接點有8種選擇，如圖 3

— 21 (c)，若我們再往下考慮 J_{10} 的 T 型接點時，知邊界線 $J_{10}J_{11}$ 的接線型態為 “→”，所以 J_{10} 的接點標線組合由 5 種剩至第一種合乎選擇，如圖 3—21 (d)。此時再返回過濾 J_9 的可能接點標線組合，我們發現只剩下圖 3—21(c) 中上列的四種接點標線組合，合乎選擇。至此，我們可以獲得區域 R_2 的解譯區域 $IR_2 = (J, L, h)$ ，而 $J = \{ J_1, J_{10}, J_9, J_8, J_6, J_7 \}$ ，並且 $L = \{ L_1, L_2, L_3, L_4, L_5, L_6 \}$ ，而 $h(L_2) = “+”$ ， $h(L_1) = h(L_3) = h(L_4) = “→”$ ，和 $h(L_5) = h(L_6) = “←”$ 。因此，由法則 3.3.1 可判斷出區域 R_2 是一個部份被遮區域。由上例可知，利用上述的標線程序將不致把發生遮蔽的區域 R_2 解譯成不合理的意義。

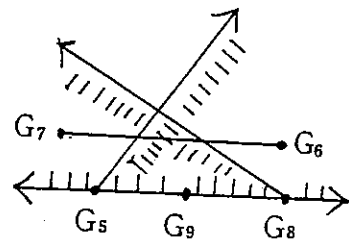
當一張多物體的解譯影像完成時，如圖 3—22(a)，我們很容易地由該張解譯影像中各面的傾斜量在斜量空間上的位置幾何關係，如圖 3—22(b)，(c) 和 (d)，可看出影像中那些區域是為那個物體所有。假設圖 3—22(a) 中的區域 R_i 其相對的傾斜量為 G_i ， $i = 1, 2, \dots, 12$ 。而圖 3—22 (a) 中所有的 G_i ， $i = 1, 2, 3, 4$ ，都沒出現於圖 3—22 (c) 和 (d) 中，由此可知圖 3—22 (a) 有一個物體 $\Omega_1 = \{ R_1, R_2, R_3, R_4 \}$ ，而同理可得， $\Omega_2 = \{ R_5, R_6, R_7, R_8, R_9 \}$ ，和 $\Omega_3 = \{ R_{10}, R_{11}, R_{12} \}$ 。最後我們知道圖 3—22 (a) 中是由三個物體 Ω_1 ， Ω_2 和 Ω_3 所構成。由於我們在下一章將利用影像中各全部可見區域的形狀來求出其傾斜量。但有些物體的全部可見區域數目不夠，以致無法直接算出該物體上各全部可見區域之傾斜量，但由於該物體上有些非自發性的部份被遮



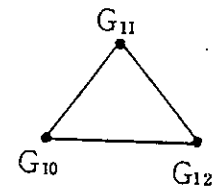
(a)



(b)



(c)



(d)

圖 3 - 22

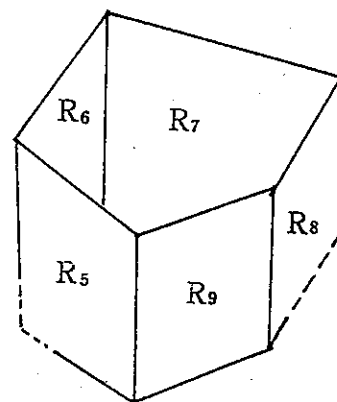


圖 3 - 23 虛線部份即被猜測出的形狀

面，其被遮部份之形狀可以被猜測出，然後即可有足夠的全部可見區域之數目，如此一來，就可以在下一章直接算出物體上這些全部可見區域之傾斜量了。所以現在底下有一個啓發式法則將可以完成有些非自發性的部份被遮面之被遮部份的形狀之猜測工作：

步驟 1：利用法則 3.3.1 將一張解譯影像中所有的部份被遮區域找出。並置於集合 IR^* 中，假設有 m 個部份被遮區域，則 $IR^* = \{ \bigcup_{i=1}^m IR_i^* \}$ 。而其傾斜量分別為 G_1^* ， G_2^* ， \dots ， G_m^* 。

步驟 2：檢驗每一個 IR_i^* ，若 IR_i^* 的傾斜量 G_i^* 在斜量空間上其位置未定則 IR_i^* 是為自發性的部份被遮面，那麼不做猜測工作。否則到步驟 3 做猜測工作。

步驟 3：找出 IR_i^* 中最鄰近的兩個 T 型接點， J_{k_0} ， J_{k_1} 。若 J_{k_0} 和 J_{k_1} 分別位於邊線 L_{k_0} 和 L_{k_1} 上，且 $h(L_{k_0}) = h(L_{k_1}) = \text{“}\rightarrow\text{”}$ ，則延長 L_{k_0} 和 L_{k_1} 。如 L_{k_0} 和 L_{k_1} 的延長線有相交於一點 J' 則連接兩線段 $J'J_{k_0}$ 和 $J'J_{k_1}$ 。重覆步驟 3，直到 IR_i^* 中所有最鄰近的兩個 T 型接點都被找過。若每次執行步驟 3 時，都有做連接兩線段的工作，則部份被遮區域 IR_i^* 的所有被遮部份之形狀都已被猜測出。

步驟 4：直到 IR^* 中所有的部份被遮區域皆被檢驗過，則猜測工作完畢。

若圖 3—22 (a)，經上述 4 個步驟後，其結果，如圖 3—23，其中 Ω_2 的二個部份被遮區域 R_5 和 R_8 的虛線部份即是所猜測出的被

遮部份之形狀。然而， R_5 和 R_8 現在就可以算是二個全部可見區域了，並且將可在下一章中求出其傾斜量。

第四章 物體運動參數之決定

我們知道，如果僅用一台位置固定的照相機，分別對運動前、後的同一靜止物體拍得兩張二維影像。如果物體的運動參數和兩張影像中的相對點都已知，則將可重建影像中物體的可見部份之實際立體影像。但是，在此論文中將在無上述兩個已知條件之下，我們提出一個測量方法來求出物體的運動參數及兩張二維影像中的相對點。然而，這個方法又可被引用到照相機光軸校正的技術上。

本章內容共分四節來敘述。在第一節中是介紹如何從二維影像中，以各全部可見面的形狀來計算出其傾斜量 (gradient)。因為當一個面的傾斜量被計算出，則其法向量 (normal vector) 就知道了。然後在第二節中，就是利用一個 Generalized Hough Transform [7] 將第一張影像中，所計算出的全部法向量和第二張影像中所計算出的全部法向量來匹配，將所有可能的旋轉角度算出，然後挑出一個出現次數最多者，此即是物體的旋轉角度。接著，再利用一次 Generalized Hough Transform，將第一張影像中所有的真接點 (T 型接點除外) 和第二張影像中所有的真接點，全部匹配過。此時我們得到兩張影像中真接點的相對點，並得到這些相對點和兩個未知平移量，所形成的一條線性方程式。在這條線性方程式內，我們將可以找出物體的兩個平移量。因為此論文限制物體的平移運動是在某已知的範圍內。至此，我們就已將物體的運動參數算出。底下在第三節中將介紹我們的模擬

實驗方式及所得的結果。最後，我們在 Tsai & chen [26] 的 3-D 測量系統中知道，系統校正 (system calibration) 是件重要但複雜又費時的工作。因此，在該節中我們將引用第二節中所求出的線性方程式，提出一個不同於 Tsai & chen 所提出的照相機光軸校正方法。

第一節 找出三維平面的法向量

通常對置於旋轉盤上的三維物體 Ω 而言，當旋轉軸旋轉了 θ 度時，在物體 Ω 上的每一個面之法向量，必然也會針對旋轉軸旋轉 θ 度。此篇論文就利用這個現象來求出物體被旋轉的角度 θ 。而我們的做法是：在物體運動前、後，分別拍得兩張影像，然後算出影像中的所有法向量。接著利用一個 Generalized Hough Transform [7] 來測量出物體所旋轉的角度 θ 。因此在此節中，將對一張合理的解譯影像，求出影像中各可見區域之傾斜量。也就是法向量。

在此文中所討論的物體，都是用一些平面所組成的折紙狀物體。而一個三維平面 S ，其本身在實際三維空間中若具有一個對稱軸 (symmetry axis)，且以該軸為中心， S 的左、右兩半部之形狀是互相對稱時，則三維平面 S 是一個對稱平面。上述的對稱現象乃沿著垂直於對稱軸的一些交叉橫線 (transverse lines) 上來觀察出的。我們知道，假如一個對稱三維面 S ，垂直於觀察者視線方向時，則此三維平面 S 的對稱軸和交叉橫線相交成直角。若於上述的位置上稍微傾斜 S 則在觀察者位置不動

之下，S 將被觀察成一個具斜對稱 (skewed symmetry) 性質的平面，也就是說，S 的對稱軸和交叉橫線不再相交成正直角，它們分別被觀察成，斜對稱軸 (skewed-symmetry axis) 和斜交叉橫軸 (skewed-transverse axis)。如圖 4—1 (a)~(c) 所示。所以一個斜對稱平面相當於在該平面上可以定義出斜對稱軸和斜交叉橫軸兩個方向角。如圖 4—1 (d) 所示。圖中， α 和 β 分別是斜

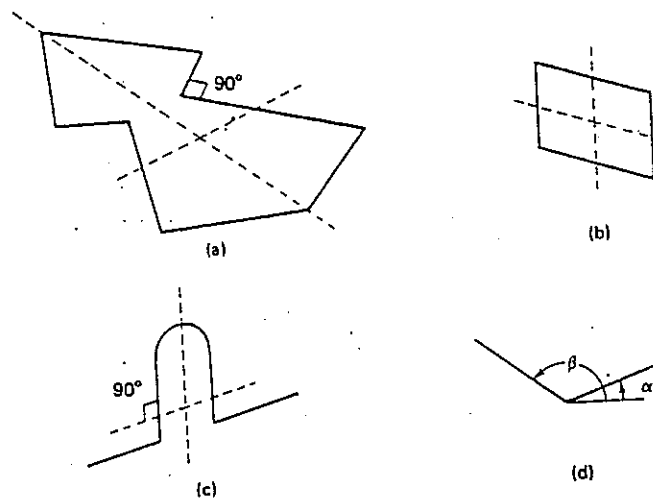


圖 4—1 (a),(b),(c)乃斜對稱的例子，而(d)是其兩斜軸的方向角

對稱軸和斜交叉橫軸的兩個方向角。由於我們是採用正交投影模式。因此在三維空間中，若有一個實際對稱的三維平面，平行投影至二維影像平面時若產生斜對稱的情形，那麼影像中的斜對稱平面所定義的斜對稱軸和斜交叉橫軸在實際三維空間上必須會真正互相垂直。

基於上述斜對稱平面在正交投影模式下的特定，我們可以對

二維影像中具斜對稱的全部可見區域，描述該區域的傾斜量在斜量空間的實際位置。讓我們假設某三維對稱平面 S 的傾斜量， $G = (p, q)$ ，而 S 投影到二維影像平面上後，若可形成一個具斜對稱區域 R，則在 R 中將可定義出二個斜對稱軸的方向角，分別為 α 和 β ，則三維空間中相對於 α 和 β 方向上的三維向量分別為：

$$A = (\cos \alpha, \sin \alpha, -G \cdot a) \quad (4-1-1)$$

和

$$B = (\cos \beta, \sin \beta, -G \cdot b) \quad (4-1-2)$$

而

$$a = (\cos \alpha, \sin \alpha) \quad \text{和} \quad b = (\cos \beta, \sin \beta)$$

這裏的“ \cdot ”代表兩向量的內積， a 和 b 為區域 R 中所定義的兩個斜對稱軸的單位向量。對於 $(-G \cdot a)$ 和 $(-G \cdot b)$ 是依據定理 3.2.4 及其推論。它們分別代表三維向量 A, B 的高度差。也就是說，區域 R 中的兩個斜對稱軸，向量 a 和 b ，是三維面 S 中的對稱軸和交叉橫軸，向量 A 和 B 在影像平面上的投影。而在實際的三維空間中，三維面 S 的對稱軸和交叉橫軸是互相垂直的。相當於 $A \cdot B = 0$ ，或

$$\cos(\alpha - \beta) + (G \cdot a)(G \cdot b) = 0 \quad (4-1-3)$$

現在若將斜量空間中的座標軸， $p - q$ ，旋轉 $\lambda = (\alpha + \beta) / 2$ 度，變成新的座標軸， $p' - q'$ ，則式子 (4-1-3) 可改寫成下式：

$$p' \cos^2((\alpha - \beta) / 2) - q' \sin^2((\alpha - \beta) / 2) = -\cos(\alpha - \beta) \quad (4-1-4)$$

而 $p' = p \cos \lambda + q \sin \lambda$ 和

$$q' = p \sin \lambda - q \cos \lambda$$

然而，式子 (4 - 1 - 4) 將在斜量空間上形成一條雙曲線，如圖 4 - 2。其意義為：凡三維對稱平面之傾斜量落於該雙曲線上者，其投影於影像平面上之區域中的兩斜對稱軸必能由方向角 α 和 β 來定義出。而，於圖 4 - 2 中雙曲線上的任一點 (p, q) 到

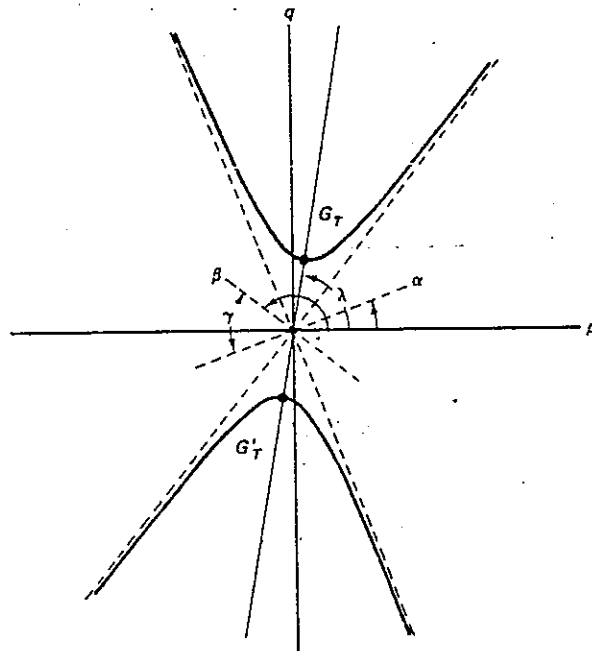


圖 4 - 2 斜對稱在斜量空間的限制

斜量空間之原點的距離表傾斜量為 (p, q) 之三維面的傾斜變化量 (magnitude of the surface slant) (相對於觀察者的視線方向)。而雙曲線上的二頂點 G_T 和 $G_{T'}$ 距離斜量空間的原點最近，所以具有傾斜量是 G_T 和 $G_{T'}$ 之三維面其傾斜變化量最少。對式子 (4 - 1 - 4) 而言，有一個特別的情形就是當：

$$|\alpha - \beta| = \frac{\pi}{2} \text{ 時,}$$

其意思乃為包含這個斜對稱之三維平面與觀察者視線方向垂直。因此，在這種情況下的斜對稱乃是真正的對稱。若由圖 4—2 來說明，就是圖中的雙曲線將退化成為兩條互相垂直且經過斜量空間的原點之直線，也就是當於， $G \cdot a = 0$ 且 $G \cdot b = 0$ 。

底下我們將描述如何從一張解譯影像中，求得組成物體之各區域的斜對稱軸與斜交叉橫軸及其方向角。然後，依據各區域所定義出的斜對稱軸和斜交叉橫軸在三維空間之關係式與兩相鄰區域的相鄰邊線在斜量空間上所具有的性質，來計算出這些區域的傾斜量。也就是說，我們可以實際決定這些區域在斜量空間上的位置。首先，我們求出影像中，具斜對稱之各區域所定義出的二個斜對稱軸。如圖 4—3 中，三個全部可見區域， $R_1 = \{ J_1, J_6, J_5, J_7 \}$ ， $R_2 = \{ J_1, J_7, J_3, J_2 \}$ 和 $R_3 = \{ J_7, J_5, J_4, J_3 \}$ 。對於各區域之斜對稱軸和斜交叉橫軸之取法是：連接各區域兩組相對邊線的中點。如圖 4—3 中，各區域內的兩條虛線即是。然而，這只是一個啟發式規則的取法，因為上述的取法將較具整體性 (globally)。對於此種取斜對稱的問題，Kanade 和 Kender [14] 在其提出的論文中有詳細的討論。而在此論文中，往後所採用的取法乃依據上述的啟發式取法為準則。接著，現在我們就可取得各區域的斜對稱軸和斜交叉橫軸之方向角，以定義出這兩個軸在影像平面上之單位向量。首先對區域 R_1 中斜對稱軸和斜交叉橫軸的兩個方向角， α_1 和 β_1 ，是以影像平面的水平座標軸方向為起始，依逆時鐘方向，分別到線段 $M_5 M_7$ 和 $M_6 M_8$ 。

間的角度。如圖 4—3。所以，區域 R_1 的斜對稱軸和斜交叉橫軸的單位向量 a_1 和 b_1 ，分別如下：

$$a_1 = (\cos \alpha_1, \sin \alpha_1) \quad \text{和} \quad b_1 = (\cos \beta_1, \sin \beta_1),$$

同樣方法可得區域 R_2 和 R_3 的斜對稱軸和斜交叉橫軸的兩個方向角分別為 α_2, β_2 和 α_3, β_3 。而其斜對稱軸和斜交叉橫軸的單位向量依次分別為 a, b 和 a_3, b_3 ，如下所示：

$$a_2 = (\cos \alpha_2, \sin \alpha_2) \quad \text{和} \quad b_2 = (\cos \beta_2, \sin \beta_2),$$

$$a_3 = (\cos \alpha_3, \sin \alpha_3) \quad \text{和} \quad b_3 = (\cos \beta_3, \sin \beta_3),$$

這裏須有一個限制條件就是：

$$\pi/2 > |\alpha_i - \beta_i| > \pi, \quad i=1, 2, 3,$$

原因是，如圖 4—4 中的菱形體 (a rhomboidal prism)，其中的各面實際上本身就是一個斜對稱面，而非對稱面。所以在此論文中將不考慮此類物體。

既然現在已經求得各區域之斜對稱軸和斜交叉橫軸在影像上的單位向量。而在正交投影模式下，依斜對稱的性質，將可用式子 (4—1—3) 的方式表示。若圖 4—3 中，區域 R_1, R_2 和 R_3 相對應的三維平面 S_1, S_2 和 S_3 之傾斜量分別為 G_1, G_2 和 G_3 ，則我們可得出下列三個關係式：

$$\cos(\alpha_1 - \beta_1) + (G_1 \cdot a_1)(G_1 \cdot b_1) = 0, \quad (4-1-5)$$

$$\cos(\alpha_2 - \beta_2) + (G_2 \cdot a_2)(G_2 \cdot b_2) = 0, \quad (4-1-6)$$

$$\cos(\alpha_3 - \beta_3) + (G_3 \cdot a_3)(G_3 \cdot b_3) = 0 \quad (4-1-7)$$

而 $G_i = (p_i, q_i)$ ， $i=1, 2, 3$ 。由上面三個式子可以知道現在有 $p_i, q_i, i=1, 2, 3$ ，六個未知數，但目前只有三個方

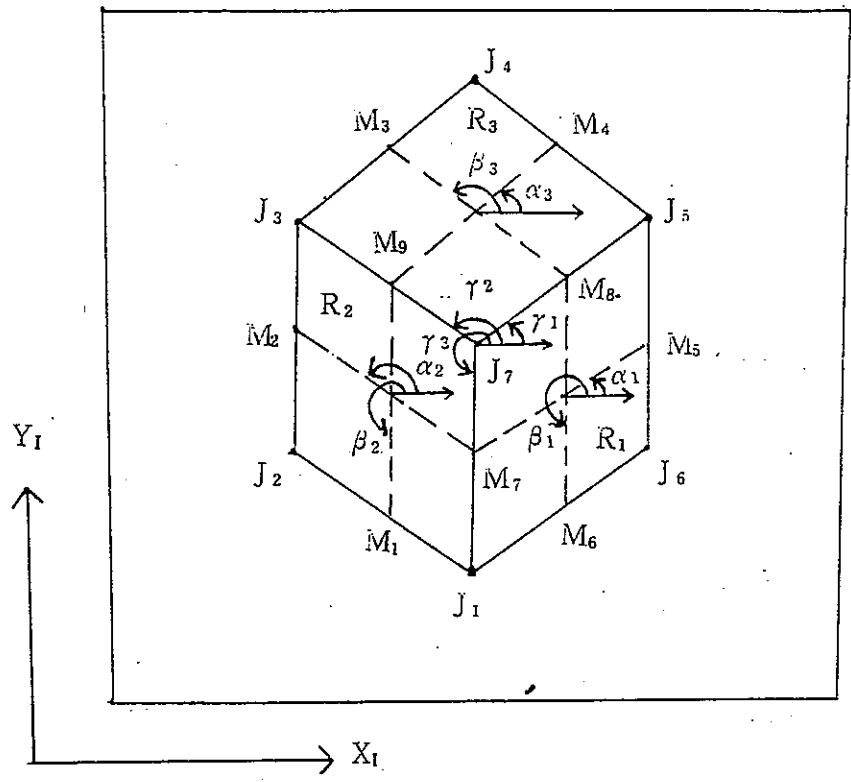


圖 4 - 3

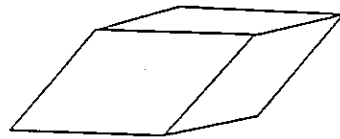


圖 4 - 4

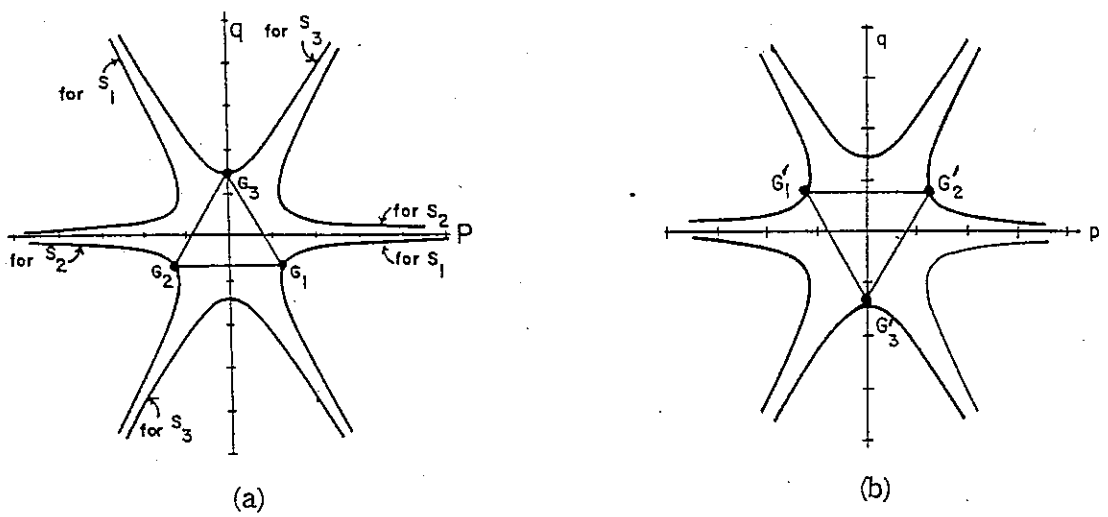


圖 4 - 5

程式而已。因此，我們再依定理 3.2.4 和定理 3.2.5 及其推論可知，兩個相鄰區域 R_1 和 R_2 ，與其相鄰邊線 $J_1 J_7$ 會有下列關係：

$$G_1 \cdot (\cos \gamma_3, \sin \gamma_3) = G_2 \cdot (\cos \gamma_3, \sin \gamma_3) \quad (4-1-8)$$

而 $(\cos \gamma_3, \sin \gamma_3)$ 表邊線 $J_1 J_7$ 的單位向量，而 γ_3 即邊線 $J_1 J_7$ 的方向角，其取法同前。同理，對圖 4-3 中的區域 R_2 ， R_3 與邊線 $J_3 J_7$ 和區域 R_1 ， R_3 與邊線 $J_5 J_7$ 分別有下列關係：

$$G_2 \cdot (\cos \gamma_2, \sin \gamma_2) = G_3 \cdot (\cos \gamma_2, \sin \gamma_2), \quad (4-1-9)$$

$$G_3 \cdot (\cos \gamma_1, \sin \gamma_1) = G_1 \cdot (\cos \gamma_1, \sin \gamma_1), \quad (4-1-10)$$

而邊線 $J_3 J_7$ 和 $J_5 J_7$ 在影像平面上之單位向量分別為 $(\cos \gamma_2, \sin \gamma_2)$ 和 $(\cos \gamma_1, \sin \gamma_1)$ 。由於我們利用二相鄰區域與其共同擁有之相鄰邊，三者在此斜量空間上的性質，所以又可增加三個方程式。因此，我們就可以對六個方程式來解出這六個未知數的兩組解。而這兩組解是互差一個負號。若以圖 4-3 為例，我們可測得：

$$\begin{aligned} \gamma_1 &= 30^\circ, \quad \gamma_2 = 150^\circ, \quad \gamma_3 = 270^\circ, \\ \alpha_1 &= 30^\circ, \quad \beta_1 = 270^\circ, \quad \alpha_2 = 150^\circ, \quad \beta_2 = 270^\circ, \quad \alpha_3 = 30^\circ, \\ \beta_3 &= 150^\circ. \end{aligned}$$

代入上述的六個方程式，然後解這組非線性聯立方程式，得出一組解：

$$G_1 = \left(\sqrt{\frac{3}{2}}, -\sqrt{\frac{1}{2}} \right), G_2 = \left(-\sqrt{\frac{3}{2}}, -\sqrt{\frac{1}{2}} \right) \text{ 和 } G_3 = (0, \sqrt{2})。$$

而另一組解為 $G_1' = -G_1$, $G_2' = -G_2$ 和 $G_3' = -G_3$ 。由於圖 4—3 中的解譯影像形成時，影像中各區域之傾斜量在斜量空間的限制條件如圖 3—12(c)。所以我們乃取第一組的解。若圖 4—3 中的解譯影像如圖 3—13(a)。則應取另一組解。但由於此文假設物體本身周圍各面必須構成一封閉空間，所以我們不會取得如圖 3—13(a) 的解譯影像。而這二組解和式子 (4—1—5)，(4—1—6) 及 (4—1—7) 在斜量空間的關係如圖 4—5。對於圖 3—15(a) 中的物體，我們將假想其上面有一個蓋子的假想平面，並且與其前面兩個平面相接。依上述的方法，亦可求出影像中物體各可見面的傾斜量。如圖 4—6 所示。

如果影像中同一物體的全部可見面超過四個（包括假想平面）的情況下，我們毋須將全部可見面的所有有關的方程式找出，然後直接求解。如圖 4—7(a) 中，若區域 R_1, R_2, R_3 和 R_4 所相對的三維平面 S_1, S_2, S_3 和 S_4 之傾斜量分別為 G_1, G_2, G_3 和 G_4 。那麼，我們可以只求出 G_1, G_2 和 G_4 或者是 G_2, G_3 和 G_4 ，就能把影像中物體的四個面之傾斜量求出。因為我們依據定理 3.2.5 及其推論，和影像中各面的傾斜量在斜量空間的幾何位置關係，如圖 4—7(b)，即可求出。其詳細過程在上一章的第四節中已有描述。

如果我們能從二維解譯影像中算出物體上某面之傾斜量， $G = (p, q)$ ，依據定義 3.2.1 和定理 3.2.2，則該面的法向量為

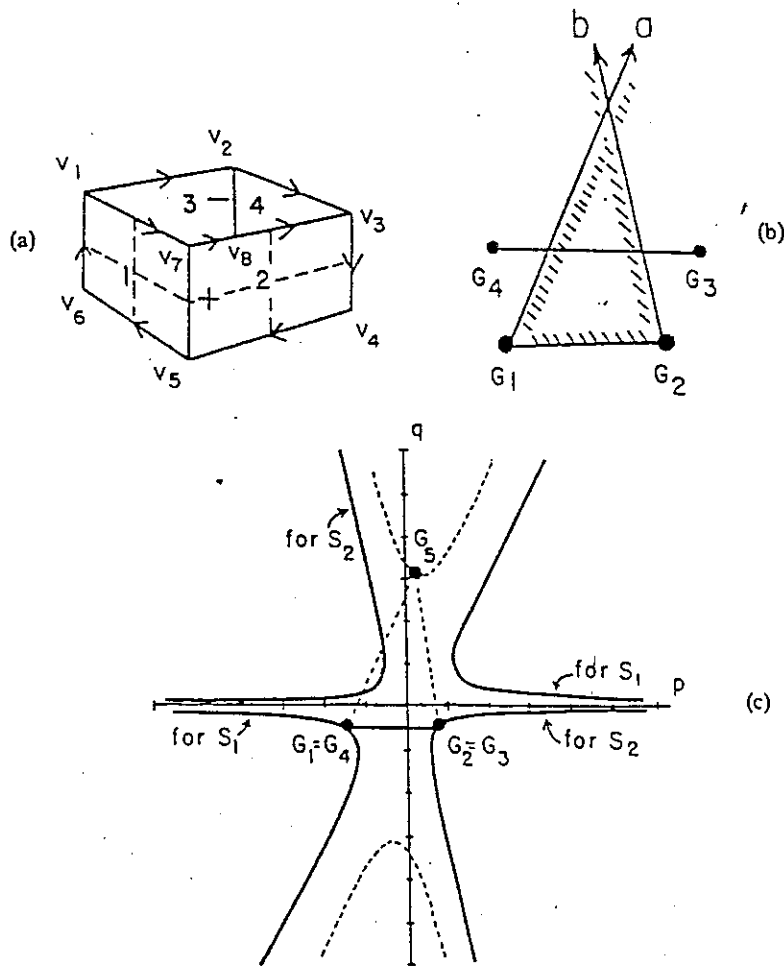


圖 4 — 6

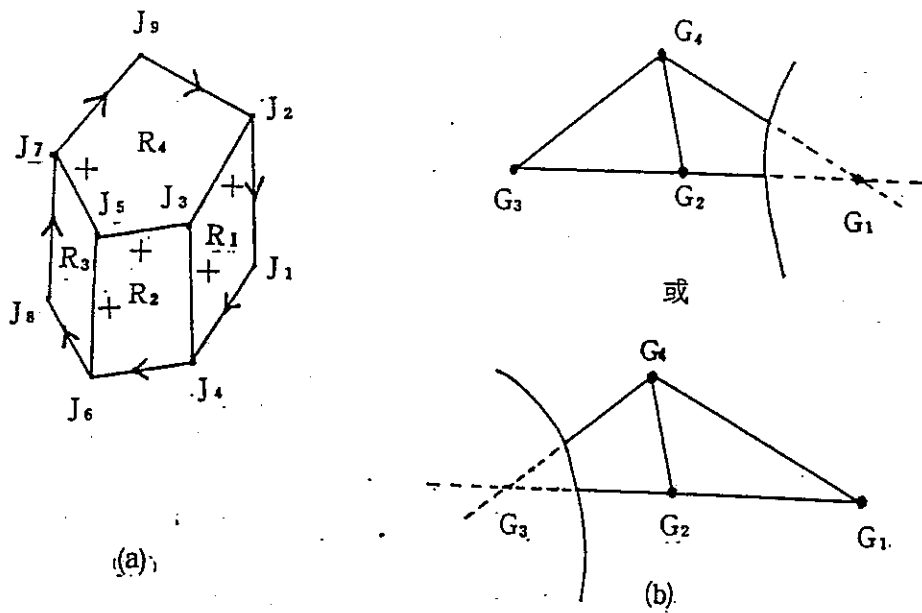


圖 4 — 7

(p , q , 1) 。

第二節 運動參數的決定

在本節中，首先介紹此論文中在決定運動參數時所引用的原理以及所使用的方法。接著就描述如何測量出物體的旋轉角度 θ_z ，然後才測量出平移向量中的 Δx ， Δy 。

由於我們假設物體運動前、後，分別所拍得的兩張影像中，其相對點未知。但其相對點不能消失太多，所以我們乃限制物體的旋轉角度 θ_z 及平移向量中的 Δx ， Δy 都必須在某一已知範圍內。在這種情況下，我們依據兩個原理，來從兩張影像中決定物體的運動參數。首先，我們乃利用實物所處的實物座標系 (world coordinate system) 與照相機座標系 (camera coordinate system) 二者間的座標系變換 T_{wc} 和 T_{cw} 及投影時的影像變換，如圖 2—10(c)，來導出實物的 3—D 座標與其影像上的座標及運動參數 (θ_z ， Δx ， Δy) 三者間的關係式。依據上述的關係式，我們可以來求兩張影像中可見的 (visible) 相對點及物體의 平移量 Δx ， Δy 。假設實物座標系中，轉盤上有一物體 Ω ，如果 Ω 上有一點 $v = (x, y, z)$ 。當照相機拍下第一張影像時， v 被 T_{wc} 變換至照相機座標系中的 $v' = (x', y', z')$ 且 v' 被以正交投影方式，變換到影像平面上的影像點 $I' = (x', y')$ 。現在將 Ω 旋轉 θ_z 且平移了 $D = (\Delta x, \Delta y, 0)$ 。此時，照相機再拍入第二張影像時， v 在照相機座標系的位置為 $v'' = (x'', y'', z'')$ ，再投影到影像平面上的影像點 $I'' = (x'', y'')$ 。此時， I'

和 I'' 就是一組可見的相對點。在均質座標系下，上述的整個關係可由下式表示：

$$VT_{wc} = V' \quad (4-2-1)$$

$$(VR_{\theta_z} + D)T_{wc} = V'$$

而實物座標系變換到照相機座標系的座標系變換 T_{wc} ，由於假設在系統校正 (system calibration) 已經完成，所以 T_{wc} 就已知了。如下表示之：

$$T_{wc} = \begin{pmatrix} T_{11} & T_{12} & T_{13} & T_{14} \\ T_{21} & T_{22} & T_{23} & T_{24} \\ T_{31} & T_{32} & T_{33} & T_{34} \\ T_{41} & T_{42} & T_{43} & T_{44} \end{pmatrix} \quad (4-2-2)$$

且 R_{θ_z} 為針對正 Z 軸方向旋轉 θ_z 度的旋轉矩陣，若以逆時鐘方向旋轉則 $\theta_z > 0$ ，反之， $\theta_z < 0$ ，如下表示之：

$$R_{\theta_z} = \begin{pmatrix} \cos(\theta_z) & \sin(\theta_z) & 0 & 0 \\ -\sin(\theta_z) & \cos(\theta_z) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4-2-3)$$

此時， $D = (\Delta x, \Delta y, 0, 0)$ 。因 x', y', x'', y'' ，於影像上可以取得，但 z', z'' 仍未知，所以我們改寫式子 (4-2-1) 為 (4-2-4)。

$$\left\{ \begin{array}{l} T_{11}x + T_{21}y + T_{31}z = x' - T_{41} \\ T_{12}x + T_{22}y + T_{32}z = y' - T_{42} \\ (R_{11}T_{11} + R_{12}T_{21})x + (R_{21}T_{11} + R_{22}T_{21})y + T_{31}z + T_{11}\Delta x + \\ T_{12}\Delta y = x'' - T_{41} \\ (R_{11}T_{12} + R_{12}T_{22})x + (R_{21}T_{12} + R_{22}T_{22})y + T_{32}z + T_{21}\Delta x + \\ T_{22}\Delta y = y'' - T_{42} \end{array} \right.$$

(4 - 2 - 4)

且 $R_{11} = \cos(\theta_z)$, $R_{12} = \sin(\theta_z)$, $R_{21} = -\sin(\theta_z)$ 和 $R_{22} = \cos(\theta_z)$ 。

由式子 (4 - 2 - 4) 可知，目前總共有 $x, y, z, \theta_z, \Delta x, \Delta y$ 六個未知數，而僅只有四條方程式。因此，我們無法直接解 (4 - 2 - 4) 來求出這六個未知數。所以，我們將從物體運動前、後所拍得的兩張影像先來測量出 θ_z ，其所引用的原理是：當轉盤上的物體被旋轉 θ_z 時，則該物體上的所有平面之法向量亦都隨之旋轉 θ_z 了。並且在植基於一個 Generalized Hough Transform [7] 的方法來測量 θ_z 。而其詳細的做法如下：假設現在所得到的二張影像已經過上述前二章及本章中前兩節的處理了。

我們拿到第一張影像所求的傾斜量有 r 個，若 $G' = \{ \bigcup_{i=1}^r G'_i = (p'_i,$

$q'_i) \}$ ， r 屬於自然數 }。而第二張影像中有 s 個傾斜量，若 $G'' =$

$\{ \bigcup_{j=1}^s G''_j = (p''_j, q''_j) \}$ ， s 屬於自然數 }。假設物體的旋轉角

度 θ_z 必須落於某範圍內，若 $\theta_z \in [\theta_{\min}, \theta_{\max}] = \text{Interval}(\theta_z)$

。所以當物體被逆時鐘旋轉 θ_z ，則 $0 \leq \theta_z \leq \theta_{\max}$ ，反之， $0 \leq$

$|\theta_z| \leq |\theta_{\min}|$ 。而 r 和 s 可以不相等。若此論文中假設旋轉角度的解析度 (resolution) 為轉盤的一個旋轉單位的 w 倍， w 屬於自然數。

至於如何來測量出 θ_z 。簡單地說：我們以一個 Generalized Hough Transform 將第一張影像中所有的法向量和第二張影像中所有的法向量來做匹配 (matching)。然後可得到一組旋轉角度，我們選擇出現次數最高的旋轉角度，此乃物體的旋轉角度 θ_z 。如演繹法 1。如果 G' 和 G'' 分別可以化成二個單位法向量的集合 (依定義 3.2.1 和定理 3.2.2 及其推論)， $N' = \{ \bigcup_{i=1}^r N'_i = \frac{(p'_i, q'_i, 1)}{\sqrt{p_i'^2 + q_i'^2 + 1}} \}$ 和 $N'' = \{ \bigcup_{j=1}^s N''_j = \frac{(p''_j, q''_j, 1)}{\sqrt{p_j''^2 + q_j''^2 + 1}} \}$ ，則若將 N' 中的一個單位法向量 N'_{i_0} ， $1 \leq i_0 \leq r$ ，分別與 N'' 中的每一個 N''_j 匹配後，得到一組可能的旋轉角度， $\{ \bigcup_{j=1}^s \theta_{i_0 j} \}$ 。假如在第一張影像中的一個平面 S'_{i_0} 其單位法向量為 N'_{i_0} ，而且在第二張影像中有一平面 S''_{j_0} ， $1 \leq j_0 \leq s$ ，並與 S'_{i_0} 互為相對面，那麼，必定有一個旋轉角度 $\theta_{i_0 j_0} \in \{ \bigcup_{j=1}^s \theta_{i_0 j} \}$ ，使得 $\theta_{i_0 j_0} = \theta_z$ 。相當於 $N'_{i_0} T_{cw} R_{\theta_z} = N''_{j_0} T_{cw}$ 。由於我們求出各面的法向量是相對於照相機座標系，因此在計算旋轉角度時，必須將這些法向量透過，照相機座標系對實物座標系的座標軸變換， $T_{cw} = T_1^{-1} T_3^{-1} T_2^{-1}$ ，因為法向量的方向是以實物座標系的原點 $(0, 0, 0)$ 為起始點，所以不需要平移矩陣 T_1^{-1} 。一旦，我們將 N' 中所有法向量 N'_i 全部和 N'' 中的所有法向量 N''_j 都一一匹配後，會得到一組旋轉角度，

an array ANGLE ($\theta_{\min} : \theta_{\max}$) initialized to zero.

for each N_i , $1 \leq i \leq r$

for each N'_j , $1 \leq j \leq s$

solve $N_i T'_{cw} R_\theta - N'_j T'_{cw} = 0$.

if $\theta \in [\theta_{\min}, \theta_{\max}]$ then

ANGLE(θ) = ANGLE(θ) + 1.

end.

end.

end.

θ'_z = the index are given by maxima in array ANGLE.

for each (ANGLE(θ) / ANGLE(θ'_z)) > threshold

for each $\theta'' = \theta \pm \delta_\theta$

ANGLE(θ) = ANGLE(θ) + ANGLE(θ'').

ANGLE(θ'') = 0, if $\theta'' \neq \theta$.

end.

end.

θ_z = the index are given by maxima in array ANGLE.

演繹法 1

我們用一個陣列 ANGLE 來記錄這些旋轉角度在匹配過程中所出現的次數。而 ANGLE 中的指標 (index) 就是這些旋轉角度。由於數值運算的誤差是不可避免的。因此，我們乃在幾個出現次數較高的旋轉角度之附近，做一次群集 (clustering) 的工作。最後，我們就選擇陣列 A 中具有最高值的元素其指標即物體的旋轉角度 θ_z 。然而，還須注意兩件事情是：(1) 實物座標系中的物體 Ω ，若 Ω 中的平面其法向量平行於 Z 軸，則雖 Ω 被旋轉 θ_z ，但此類的法向量仍然不變。所以我們在執行演繹法 1 的過程中，將拿掉這類法向量。(2) 儘管我們將最有可能的旋轉角度 θ_z 決定出，但我們不能保證：第一張影像中有一個面 S'_{k_1} 其單位法向量為 $N'_{k_1} \in N'$ ， $1 \leq k_1 \leq r$ ，且第二張影像中亦有一個面 S''_{k_2} ， $1 \leq k_2 \leq s$ ，其單位法向量為 $N''_{k_2} \in N''$ 。若 $N'_{k_1} T'_{cw} R_{\theta_z} = N''_{k_2} T'_{cw}$ 則 S'_{k_1} 和 S''_{k_2} 互為相對面 (corresponding planar surface)。因為第二張影像中具 N''_{k_2} 的平面可能不只 S''_{k_2} 一個平面而已。例如，在轉盤上放二個同形狀的物體，使其具同一方向。

既然 θ_z 已經被測量出了，我們將 (4-2-4) 整理成：

$$\begin{cases} f_1(x, y, z, \Delta x, \Delta y) = f'_1(x', y', x'', y'') \\ f_2(y, z, \Delta x, \Delta y) = f'_2(x', y', x'', y'') \\ f_3(z, \Delta x, \Delta y) = f'_3(x', y', x'', y'') \\ f_4(\Delta x, \Delta y) = f'_4(x', y', x'', y'') \end{cases} \quad (4-2-5)$$

在相對點未知的條件下，目前仍有五個未知數，依然無法直接解 (4-2-5)。可是我們觀察出一個現象：在 T_{cw} 和 θ_z 已知，對兩張影像 I_1, I_2 中的所有相對點 $I'_1 = (x'_1, y'_1)$ ， $I''_1 = (x''_1, y''_1)$

而言（假設有 m 組， $1 \leq i \leq m$ ），各組相對點和運動參數 $\theta_z, \Delta x, \Delta y$ 及其實物座標系的像點（scene points）將形成式子（4-2-4）的關係，並整理成式子（4-2-5）後，我們發現：

$$f'_4(I'_1, I''_1) = f'_4(I'_2, I''_2) = \dots = f'_4(I'_m, I''_m) = K \quad (4-2-6)$$

而 K 是一個固定的常數。一旦我們知道，任意兩點， $I'_{i_0} \in I_1, I''_{i_0} \in I_2$ 互為相對點，那麼 I'_{i_0}, I''_{i_0} 和 $\Delta x, \Delta y$ 的關係，是一條以 $\Delta x, \Delta y$ 為變數的二元一次線性方程式：

$$f_4(\Delta x, \Delta y) = f'_4(I'_{i_0}, I''_{i_0}) = K \quad (4-2-7)$$

所以說：如果我們能得到式子（4-2-7），然後就可以來估算出物體的二平移量 $\Delta x, \Delta y$ ，因為其被限制於某範圍內，假設 $\Delta x \in [\Delta x_{\min}, \Delta x_{\max}] = \text{Interval}(\Delta x), \Delta y \in [\Delta y_{\min}, \Delta y_{\max}] = \text{Interval}(\Delta y)$ ，並且其解析度為 x, y 軸上的一個單位長度的 h 倍， h 屬於自然數。在上述的假設條件之下，我們將 $\text{Interval}(\Delta x)$ 和 $\text{Interval}(\Delta y)$ 中，所有 $\Delta x_k, \Delta y_k$ 組合成一個序對 $(\Delta x_k, \Delta y_k)$ ，其中 $\Delta x_k \in \text{Interval}(\Delta x)$ 且 $\Delta y_k \in \text{Interval}(\Delta y)$ 。如果有 $t \times t$ 個序對， $1 \leq k \leq t^2$ ，然後將所有序對 $(\Delta x_k, \Delta y_k)$ 中的二元素代入式子（4-2-7），挑出一組 $(\Delta x_{k_0}, \Delta y_{k_0})$ ， $1 \leq k_0 \leq t^2$ ，其所產生的誤差值，

$$\delta_{k_0} = |f_4(\Delta x_{k_0}, \Delta y_{k_0}) - K| = \text{MIN}(|f_4(\Delta x_k, \Delta y_k) - K|), 1 \leq k \leq t^2,$$

是所有序對 $(\Delta x_k, \Delta y_k)$ 所得出的誤差值中，最小者。這組序對 $(\Delta x_{k_0}, \Delta y_{k_0})$ 中的兩元素即是物體的兩平移量 $\Delta x = \Delta x_{k_0}, \Delta y = \Delta y_{k_0}$ 。

現在的問題是我們如何從二張未知相對點的影像中，來尋找

出相對點及式子 (4 - 2 - 7) 中的 K 值。

假如我們有兩張影像 I_1, I_2 的特徵點，指真接點 (T 型接點除外) 而言。如果 I_1, I_2 中分別有 c_1, c_2 個特徵點， $I_1 = \{ \bigcup_{i=1}^{c_1} J_i = (x_i, y_i), c_1 \text{ 屬於自然數} \}$ ， $I_2 = \{ \bigcup_{j=1}^{c_2} J'_j = (x'_j, y'_j), c_2 \text{ 屬於自然數} \}$ ，則我們再用一次 Generalized Hough Transform 將 I_1 和 I_2 中互為相對接點者尋找出，且式子 (4 - 2 - 7) 中的 K 值亦一併找出。如演繹法 2。首先我們將 I_1 中的所有接點 J_i 分別和 I_2 中所有的接點 J'_j ，代入式子 (4 - 2 - 7)，全部匹配。在匹配過程中，我們使用一個陣列 ACC 來計數曾經出現過相同 K 值的次數，並用一陣列 MATCH 來記錄不同 K 值及得出 K 值的所有相對點。由於電腦內數值運算過程中，誤差的發生是無法避免。因此我們勢將無法得到如 (4 - 2 - 6) 式中的理想狀況，所以我們將幾個出現次數較高的 K 值拿出，在其附近的值做一次群集 (clustering) 的工作。最後，挑出陣列 ACC 中次數最高之元素的指標 K'' ，接著我們在陣列 MATCH 中取出其指標為 K'' 之元素 MATCH (K'')，其所記錄的全部 K 值之平均值就是我們所要測量的 K 值。一旦 K 值被測量出，則我們就可用演繹法 3 來估算物體的平移量 $\Delta x, \Delta y$ 了。

至此，我們已經知道相對點，運動參數，所以將其代入式子 (4 - 2 - 5) 即能求出形成這些相對點之三維像點的實際座標。亦求出這些相對點位於照相機座標系上的深度值， z_i, z'_i 。又我們已算出物體全部可見面之法向量，其相對的方程式即可被

$ACC(i) = 0, MATCH(i) = \phi$, for all i .

for each $J'_i, 1 \leq i \leq c_1$

for each $J'_j, 1 \leq j \leq c_2$

compute $K_1 = f'_i(J'_i, J'_j)$.

$ACC(K_1) = ACC(K_1) + 1$.

$MATCH(K_1) = MATCH(K_1) \cup \{(J'_i, J'_j, K_1)\}$.

end.

end.

K' = the index are given by maxima in array ACC .

for each $(ACC(K_1) / ACC(K')) > \text{threshold}$

for each $K_3 = K_1 + \delta_k$

$ACC(K_1) = ACC(K_1) + ACC(K_3)$.

$MATCH(K_1) = MATCH(K_1) \cup MATCH(K_3)$.

$ACC(K_3) = 0$, if $K \neq K_3$.

end.

end.

K'' = the index are given by maxima in array ACC .

$$K = \left(\frac{\sum K_1}{(J'_i, J'_j, K_1) \in MATCH(K'')} \right) / ACC(K'').$$

演繹法 2

求出。然而對兩張影像中，非相對點的接點而言，它們將被代入其所屬的平面方程式中，直接可計算出其深度值。如此一來，我們即可獲得影像中物體的可見部份之實際立體形狀。

```
for each  $\Delta x_k \in \text{Interval } (\Delta x)$ 
  for each  $\Delta y_k \in \text{Interval } (\Delta y)$ 
    compute  $\delta = | f'_z (\Delta x_k, \Delta y_k) - K |$ 
    if  $\delta < \delta_{\min}$  then
       $(\Delta x, \Delta y) \leftarrow (\Delta x_k, \Delta y_k)$  and  $\delta_{\min} \leftarrow \delta$ 
    end.
  end.
end.
```

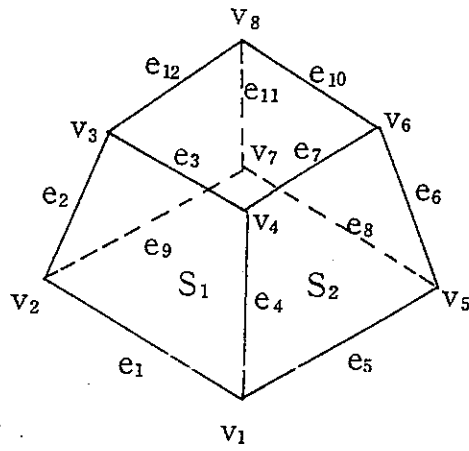
演繹法 3

第三節 模擬實驗及其結果

在此節中是介紹本章前兩節所提的方法之模擬實驗方式並討論所得的結果。第一個實驗是以第一節所描述的方法算出一張二維圖形中物體的可見面之傾斜量。第二個實驗乃是對已算出各可見面之傾斜量的兩張圖形，依第二節中的三個演繹法來分析出圖形中物體之運動參數。

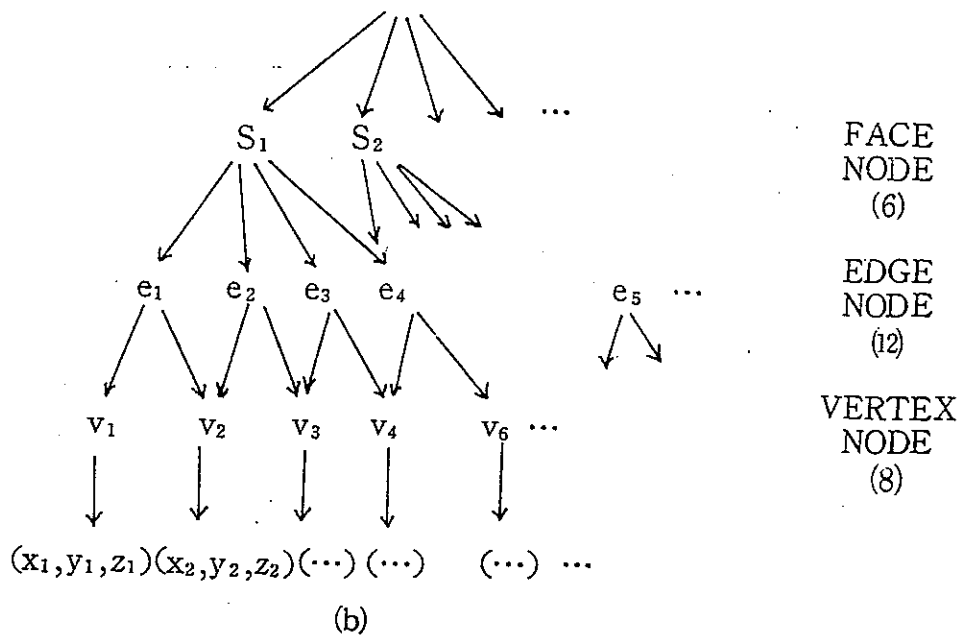
首先，我們以邊界表示法 (Boundary representation) [29] 來定義實驗所使用之物體。下圖 4—8 即是定義一個梯形體的例子。在物體定義完畢後，接著即定義兩座標系，實物座標系和照相機座標系，兩者之間的幾何關係，並決定影像平面的所在位置。而在本節模擬實驗中係採正交投影模式，且已知物體上，面與面之相鄰邊的實際性質 (convexity or concavity)，最後我們假想有一轉盤，並定義轉盤旋轉中心位於實物座標系之原點。至於轉盤之運動限制將與前述同。

在第一個模擬實驗中，首先將兩個已定義之物體 Ω_1 和 Ω_2 分別置放於實物座標系中的轉盤上，接著，旋轉轉盤以獲取四張二維圖形。如圖 4—9 (a) 中的第一張圖乃將 Ω_1 置於尚未旋轉過之轉盤上所獲得的圖形，且以此位置為轉盤之原始旋轉位置分別依順時鐘方向將轉盤旋轉 20° (其實是將 Ω_1 旋轉 -20°) 得到第二張圖，依逆時鐘方向將轉盤 30° ， 60° 獲得第三、第四張圖。在前三張圖中 Ω_1 的 S_i 皆為可見面，但在第四張圖就因旋轉角度稍大而消失了。如果 S_i 的傾斜量為 (p_i, q_i) ， $i=1, 2, 3, 4$ 。則我們從這四張圖形中計算出各可見面之傾斜量與實際各可見面



(a)

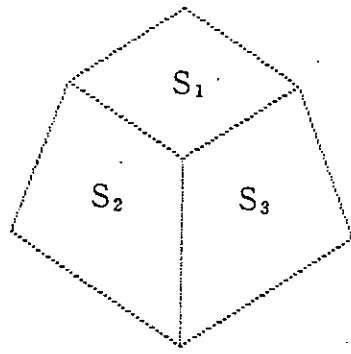
Object
boundary



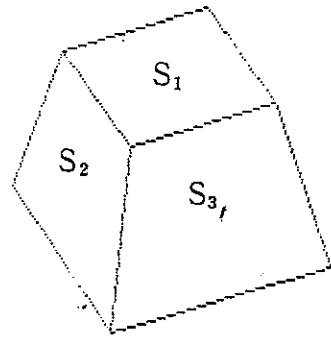
(b)

圖 4 — 8 梯形體的邊界表示法

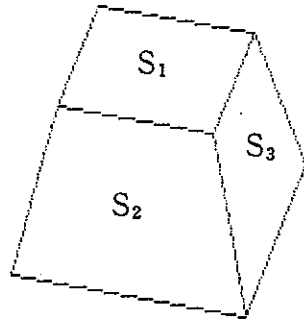
之傾斜量的比較如圖 4 — 9 (b) 所示。有關 Ω_2 的圖形及資料將列於圖 4 — 10 中。由上兩個圖中我們知道，以第一節所描述的方法計算出各可見面之傾斜量和實際之傾斜量其誤差均不超過 0.5 %。



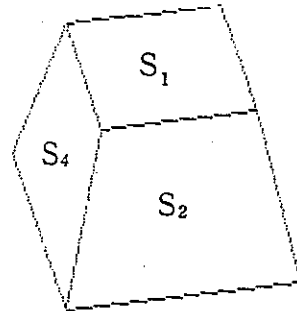
(1)



(2)

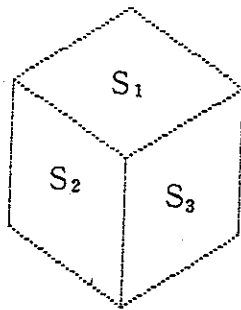


(3)

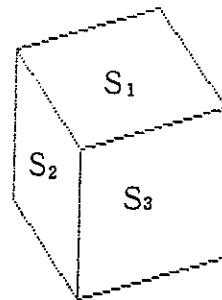


(4)

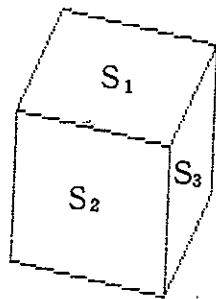
图 4 — 9 (a)



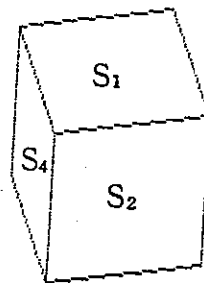
(1)



(2)



(3)



(4)

图 4 — 10 (a)

第一張圖形 (0°)

計算出的傾斜量	實際的傾斜量
p1 = 6.0120366E-06	p1 = 4.130488E-07
q1 = 1.4142092	q1 = 1.414705
p2 = -.97758249	p2 = -.9797511
q2 = -.27901509	q2 = -.282666
p3 = .97759302	p3 = .9797501
q3 = -.27902287	q3 = -.2826657

第二張圖形 (-20°)

計算出的傾斜量	實際的傾斜量
p1 = -.0041903481	p1 = 1.032622E-07
q1 = 1.4169531	q1 = 1.414705
p2 = -1.8558792	p2 = -1.851835
q2 = -.077165653	q2 = -8.131141E-02
p3 = .47621782	p3 = .4778657
q3 = -.36352034	q3 = -.3606882

第三張圖形 (30°)

計算出的傾斜量	實際的傾斜量
p1 = .0008061009	p1 = 2.065244E-07
q1 = 1.4160695	q1 = 1.414705
p2 = -.27689891	p2 = -2.773848
q2 = -.37776578	q2 = -.3787501
p3 = 2.7165847	p3 = 2.715914
q3 = .15544259	q3 = .1539498

第四張圖形 (60°)

計算出的傾斜量	實際的傾斜量
p1 = 6.4218580E-06	p1 = 1.544933E-07
q1 = 1.4146708	q1 = 1.414705
p2 = .2739037	p2 = .2773834
q2 = -.37877838	q2 = -.3787502
p4 = -2.7159037	p4 = -2.715926
p4 = .15393524	q4 = .1539523

(b)

圖 4-9

第一張圖形 (0°)

計算出的傾斜量	實際的傾斜量
p1 = 5.6290921E-06	p1 = -3.614177E-07
q1 = 1.4142092	q1 = 1.414705
p2 = -1.2247414	p2 = -1.224604
q2 = -.70710407	q2 = -.7068612
p3 = 1.2247508	p3 = 1.224603
q3 = -.70711382	q3 = -.7068613

第二張圖形 (-20°)

計算出的傾斜量	實際的傾斜量
p1 = 5.6156683E-06	p1 = 1.032622E-07
q1 = 1.4108388	q1 = 1.414705
p2 = -2.6268865	p2 = -2.626168
q2 = -.70878776	q2 = -.7068612
p3 = .57192767	p3 = .5710423
q3 = -.70880049	q3 = -.7068613

第三張圖形 (30°)

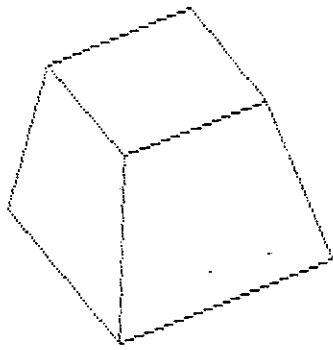
計算出的傾斜量	實際的傾斜量
p1 = 5.6269267E-06	p1 = 2.065244E-07
q1 = 1.4136672	q1 = 1.414705
p2 = -.32835512	p2 = -3.28132
q2 = -.70737877	q2 = -.7068607
p3 = 4.5694385	p3 = 4.570274
q3 = -.70739826	q3 = -.7068612

第四張圖形 (60°)

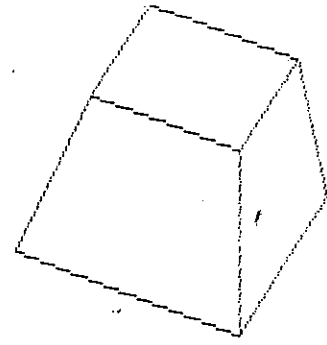
計算出的傾斜量	實際的傾斜量
p1 = 5.6269267E-06	p1 = 0
q1 = 1.4146731	q1 = 1.414705
p2 = .32813556	p2 = .3281302
q2 = -.70687838	q2 = -.7068613
p4 = -4.5702552	p4 = -4.570297
q4 = -.70685888	q4 = -.7068613

(b)

圖 4-10



(a)



(b)

轉盤旋轉中心位置：

運動前

(-20° , -2 , -2)

運動後

(30° , -4 , -7)

測量出之旋轉角度有：

旋轉角度

50°
-40°

出現次數

2 ✓
1

測量出之 K 值有：

K 值

-4.892
-4.656
-3.428
-2.200
-1.964

出現次數

2
2
7 ✓
2
2

測量出之二平移量有：

(Δ x , Δ y)

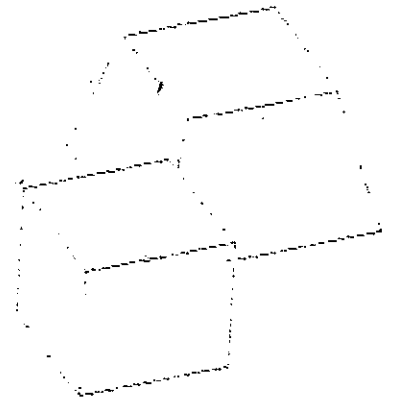
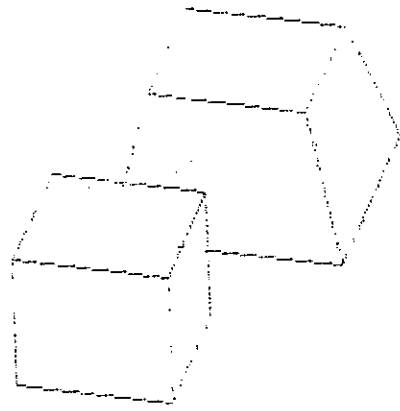
-2 -5
9 -9

誤差值

4.768372E-07 ✓
2.199888E-03

(c)

圖 4-11



(a)

(b)

轉盤旋轉中心位置：

運動前 運動後

(30° , -4 , -4) (-20° , -2 , 2)

測量出之旋轉角度有：

旋轉角度 出現次數

-50° 4 ✓
40° 2

測量出之 K 值有：

K 值 出現次數

1.528 4
1.687 6
2.345 4
2.504 14 ✓
2.663 4
3.321 6
3.480 4

測量出之二平移量有：

(Δ x , Δ y) 誤差值

2 6 3.004074E-05 ✓
6 -5 2.171517E-03

(c)

圖 4-12

第二個模擬實驗中，將計算轉盤運動前、後的三個運動參數。也就是轉盤上物體之運動參數。我們將以 (θ, x', y) 表示轉盤的位置， θ 表轉盤已由原始位置旋轉了 θ 度，假如 $\theta < 0$ 則表轉盤依順時鐘方向旋轉反之，依逆時鐘方向旋轉，而 x, y 表目前轉盤旋轉中心在 $X Y$ 平面上的座標。並且在此模擬實驗中我們所設定的旋轉解析度為 1 度，平移解析度為物體的 1 個單位長度（未做 scaling）。在圖 4—11 至 4—13 中各圖(a), (b)分別

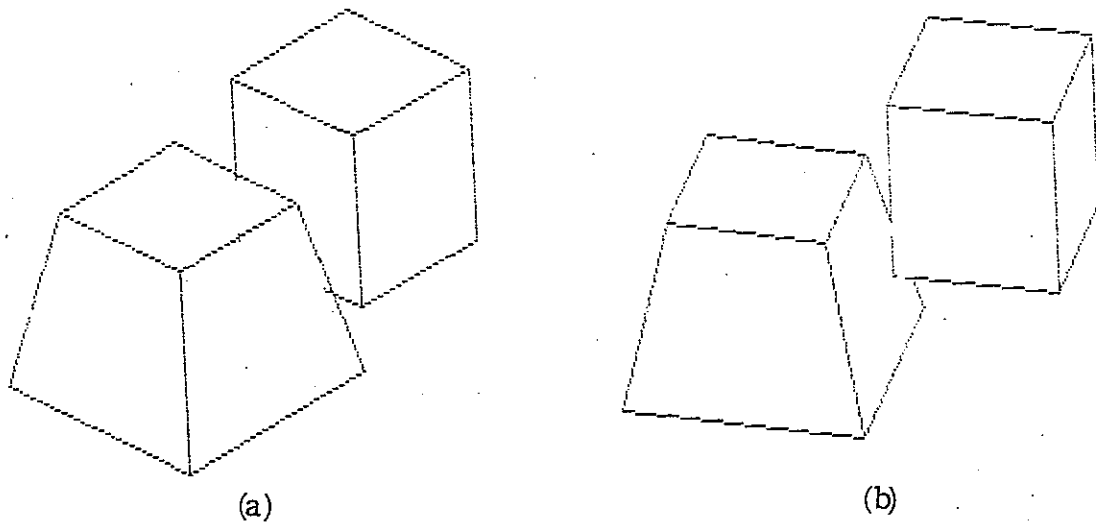


圖 4—13

是，物體被置放於轉盤後，轉盤運動前、運動後的二維圖形。由第一個實驗結果與前一節中的演繹法 1，將可測量出轉盤之旋轉角度，接著再以演繹法 2 和演繹法 3，依序求出式子（4—2—7）中的 K 值和轉盤的兩個平移量 $\Delta x, \Delta y$ 。其結果如各圖中的圖(c)所示。在圖 4—13 中，由於轉盤運動後，對圖(a), (b)而言，

轉 盤 旋 轉 中 心 位 置 :

運 動 前

運 動 後

(0° , 0 , 2)

(60° , -3 , 2)

測 量 出 之 旋 轉 角 度 有 :

旋 轉 角 度

出 現 次 數

-30°
60°

4
2 ✓

測 量 出 之 K 值 有 :

K 值

出 現 次 數

-3.484
-3.156
-2.697
-2.559
-2.369
-2.121
-2.099
-1.772
-1.662
-1.523
-1.334
-1.064
-.736
-.139
.299
.897

7
4
4
3
4
3
3
4
3
4
4
4
7
4
4
3

-1.806
-1.443
-.863
-.518
-.173
-.156
.189
.408
.770
1.115

6
5
7
12 ✓
6
4
4
4
7
5

測 量 出 之 二 平 移 量 有 :

(Δx , Δy)

誤 差 值

-7 1
-3 0
1 -1
8 -3

4.622323E-02
2.980232E-07 ✓
4.622382E-02
3.383541E-02

(c)

圖 4-13

有相對面及相對點消失。因此，在這種情況下，可能會暫時性地選出錯誤的旋轉角度，而這將可於測量K值的過程中來矯正。如圖4—13(c)中，我們首先將會選（ -30° ）當作是轉盤的可能旋轉角度，而接著計算出所有可能的K值（如圖中虛線以上所示）。可是因為所有可能的K值其出現次數太平均了以致於無法凸顯出哪個K值才是正確的。由此可得知，（ -30° ）應不是轉盤的旋轉角度，所以我們必須選擇次一個可能旋轉角度來測量出K值。當我們取旋轉角度， 60° 來計算出所有可能的K值時（如圖(c)虛線以下所示），將可很明顯的找到一個出現次數較高者。由此我們推測轉盤的旋轉角度應為 60° ，並繼續算出轉盤的兩個平移量。

由以上的實驗顯示，用前兩節所描述的方法來決定兩張二維圖形中物體之三個運動參數將可獲得令人滿意的結果。

第四節 照相機光軸校正

一般而言，凡利用實物座標系中之三維像點（scene points）與其投影在影像平面的影像點（image points）二者之間的座標關係式，去獲得該三維像點的實際座標，則系統校正（system calibration）必然是首先且重要的工作。例如：Tsai & chen [26] 所提出的 3-D 測量系統，及本文所用來決定物體運動參數的系統都必須事先做好系統校正。而目前系統校正仍是件複雜且費時的工作。Tsai & chen [26] 在其系統校正過程中，用 5 種校正技術來測量出系統校正所需的參數值。於本節中，我們將略述 Tsai & chen 所提出的系統校正過程及各過程所欲測量的參數，然後我們提出一種不同於 Tsai & chen 的照相機光軸校正中所使用的方法。

Tsai & chen 提出的 3-D 測量系統，與此文所假設的系統環境大致類似，只差於發光系統。其發光系統所發出的雷射光源可使其照相機拍入一張距離影像（range image），然後即可反求影像中物體的實際座標。而其系統校正如下所示：

(1) 照相機參數校正（camera parameters calibration）：

測量照相機之透鏡中心位於實物座標系中的座標 $O_c = (x_c, y_c, z_c)$ 和 O_c 的泛角（pan angle） θ 與傾斜角（tilt angle） ϕ 。如圖 2-10(c)。

(2) 影像中心校正（image center calibration）：

測量照相機座標系中影像平面的中心點 $O_I = (x_I, y_I)$ 。

(3) 照相機焦距長度校正 (camera focus length calibration) :

測量照相機的焦距長度 f 。

(4) 轉盤旋轉中心校正 (turntable rotation center calibration) :

測量轉盤旋轉中心 $O_r = (x_r, y_r)$ 。

(5) 光軸校正 (optical axis calibration) :

使照相機光軸通過轉盤旋轉中心。

而其光軸校正過程如下：(假設前四項已完成)

step 1 : 重覆 2 — 4 , 直到 step 3 所計算出的交點, 夠接近 (x_1, y_1) 。

step 2 : 獲取一張影像, 該影像乃是實物座標系中, x, y 二軸所形成的兩相交直線。

step 3 : 計算出影像中兩直線之交點。

step 4 : 假如 step 3 中所計算出的交點, 不夠接近 (x_1, y_1) 則利用線性內插技術 (linear interpolation technique), 移動轉盤使得 step 3 所計算出的交點能夠接近 (x_1, y_1) 。

由上述的光軸校正過程可知, 該方法將必須移動轉盤數次後, 才能將轉盤移動到正確位置使得照相機光軸可以通過轉盤中心。

而我們所提的方法是: 首先引用式子 (4 — 2 — 5) 計算出, 目前轉盤旋轉中心的位置離正確位置之平移向量 $D = (\Delta x,$

$\Delta y, 0$)，相當於計算出 x, y 軸兩方向的平移量 $\Delta x, \Delta y$ 。然後透過 D 直接移動目前轉盤旋轉中心的位置到正確位置。也就是說，僅移動一次轉盤即可完成照相機光軸校正。其更詳細的描述於後：

假設目前轉盤旋轉中心位於 $P_1 = (x_1, y_1, z_T)$ ，而轉盤高度 z_T 已知。現在我們也採用上述光軸校正過程中 step 2 和 step 3 的方法，計算出影像中兩直線的交點，其實該交點乃 P_1 在影像平面上的投影點 $P'_1 = (x'_1, y'_1)$ 。如果轉盤旋轉中心的正確位置是在 $P_2 = (x_2, y_2, z_T)$ 。現在的工作就是將目前轉盤旋轉中心 P_1 移到 P_2 ，使得 step 2 和 step 3 所計算出的交點 $P'_1 = (x'_1, y'_1)$ 能重合於影像中心點 $P'_2 = (x'_1, y'_1)$ ， P'_2 乃是 P_2 在影像平面上的投影點。

由上述的關係我們可以知道，當 P_1 平移到 P_2 時，兩張影像中的交點會是分別為 P'_1 和 P'_2 。此時我們將 P'_1 和 P'_2 看成是一組相對點，因為 P'_1 和 P'_2 都是由 P_1 的投影所得。所以現在就可引用式子 (4-2-5) 即可算出 $x_1, y_1, \Delta x, \Delta y$ 。此時我們就可移動目前轉盤旋轉中心 $P_1 = (x_1, y_1, z_T)$ ，透過平移向量 $D = (\Delta x, \Delta y, 0)$ ，移至 P_2 ，而 P_2 的座標 $(x_2, y_2, z_T) = (x_1 + \Delta x, y_1 + \Delta y, z_T)$ 。也就是說，將目前轉盤位置分別在 X, Y 兩軸上各平移 $\Delta x, \Delta y$ 倍的平移單位長度 (resolution) 後，就可到達正確位置使得照相機的光軸能通過轉盤旋轉中心。利用這個方法僅移動轉盤一次就可完成照相機光軸校正。

第五章 結論

在此篇論文中，我們從兩張折紙狀物體的二維影像測量出物體的三個運動參數。當然構成物體之各面必須具對稱性，且物體之旋轉角度不能過大而以致於使兩張影像中的相對面消失過多。在上述的情況下此論文之結果將有助於：

(1) 工業自動化環境中，具有視覺系統的機械手能依據所測量出的運動參數，自動抓取位置異動的三維物體。並且從所還原物體的三維資料亦能幫助機器人 (robot) 記住自己的位置以及提高其辨認能力。而對於深度資料之利用將可避免行動中的機器人有碰撞之情況發生。

(2) 一旦物體的運動參數和影像間的可見相對接點都已知後，再加上兩影像中大部份的可見面之法向量也已算出。由以上的資料，勢將可獲取三維物體可見部份之立體形狀。若能再拍入物體表面其餘的二維影像則亦可逐步將三維物體重建出。

(3) 我們利用，兩影像間的相對點會與影像中物體的二平移量形成一關係式，提出了一個改進的照相機光軸校正 [26] 方法。應用此方法，僅將移動一次轉盤即可到達正確位置，使得照相機光軸通過轉盤中心，完成照相機光軸校正的程序。

在此論文中，我們所得到之結果乃源自於一些理想之狀況。然而，我們應對可能會發生於實施 (implementation) 時的一些誤差 (error) 加以事前考慮，這類誤差可能是：(1) 實際所採用的照相機模式和此文中所假設的照相機模式之差異，(2) 不可靠的

數位影像之前置處理，(3)系統校正時所發生的誤差，(4)數值運算的過程。因此，若在實施時要獲得較正確的結果，就應該去分析何種誤差影響了所測量出之運動參數的精確度，然後去降低誤差的發生。

最後，我們知道對於，從兩張二維影像來決定影像中物體之運動參數，或從多張二維影像來重建三維物體，瞭解三維物體，……等，上述這類問題的探討。必定會伴隨著相對問題的發生，也就是說，它們與相對問題是一體兩面的關係。而偏偏相對問題又是個複雜且棘手的公開問題（open problem），所以我們必須去限制各種條件來簡化相對問題，然後再探討上述的問題。因此，若能找出相對問題的有效解則上述問題當可迎刃而解。

參 考 文 獻

1. Guzman, A., "Computer Recognition of Three Dimensional Objects in a Visual Scene," MAC-TR-59, MIT, Cambridge, MA, 1968.
2. Huffman, D. A., "Impossible Objects as Nonsense Sentences," in : Meltzer, B. and Michie, D. (ed.), Machine Intelligence , Edinburgh University, Edinburgh, 1971.
3. Clowes, M. B., "On Seeing Things," Artificial Intelligence 2 (1), pp. 79-112, 1971.
4. Waltz, D., "Understanding Line Drawings of Scenes with Shadows," in: Winston, P. H. (ed.), The Psychology of Computer Vision, McGraw-Hill, New York, 1975.
5. Sugihara, K., "Range-Data Analysis Guided by a Junction Dictionary," Artificial Intelligence 12, pp. 41-69, 1979.
6. Lee, S. J., Haralick, R. M. and Zhang, M. C. , "Understanding Objects with Curved Surfaces from a Single Perspective View of boundaries," Artificial Intelligence 20, pp. 145-169, 1985.
7. Ballard, D. H. and Brown, C. H., Computer Vision, Prentice-Hall, New Jersey, 1982.
8. Gonzalez, R. C. and Wintz, P., Digital Image Processing, Addison-Wesley, Taiwan, 1977.
9. Rosenfield, A. and Kak, A. C., Digital Image Processing, vol. 1, Academic Press, New York, 1982.
10. Duda, R. O. and Hart, P. E., Pattern Classification and Scene Analysis, Wiley, New York, 1973.
11. Riseman, E. M., Burns, J. B. and Hanson, A. R., "Extraction Straight Lines," IEEE Trans. Pattern Analysis Machine Intelligence, vol. PAMI-8, no. 4, pp. 425-455, 1986.
12. Nevatia, R. and Babu, K. R., "Linear Feature Extraction and Description," Computer Graphics Image Processing, vol. 13 , pp. 257-269, 1980.
13. Kanade, T., "A Theory of Origami World," Artificial Intelligence 13, pp. 279-311, 1980.
14. Kanade, T. and Kender, J., "Skewed Symmetry: Mapping Image Regularities into Shape," Tech. Rept. CMU-CS-80-133, Department of Computer Science, Carnegie-Mellon University.
15. Kanade, T., "Recovery of The Three-Dimensional Shape of An Object from a Single View," Artificial Intelligence 17 , pp. 409-460, 1981.

16. Machworth, A. K., "Interpreting Pictures of Polyhedral Scenes," *Artificial Intelligence* 4, 1974.
17. Sugihara, k., "Mathematical Structures of Line Drawings of Polededrons ; Toward Man-Machine Communication by Means of Line Drawings," *IEEE Trans. Pattern Analysis Machine Intelligece*, vol. PAMI-4, no. 5, pp. 458-469, 1982.
18. Ullman, S., *The Interpetation of Visual Motion*, Cambridge , Mass.: MIT Press, 1979.
19. Huang, T. H. and Tasi, R. Y., "Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Object with Curved Surfaces," *IEEE Trans. Pattern Analysis Machine Intelligece*, vol. PAMI-6, no. 1, pp. 13-27, 1984.
20. Silberberg, J. M. et al., "Three Dimensional Object Recognition Using Oriented Model Points," in: Rosenfield, A. (ed.) , *Techniques for 3-D Machine Perception*, North-Holland, Amsterdam, 1986.
21. Newman, W. M. and Sproull, R. F., *Principles of Interactive Computer Graphics*, McGraw-Hill, New York, 1979.
22. Park, C. S., *Interactive Microcomputer Graphics*, Addison-Wesley, Taiwan, 1985.
23. Shafer, S. A., *Shadows and Silhouettes in Computer Vision* , Kluwer Academic Publishers, Boston, 1985.
24. Marr, D., *Vision*, Freeman, San Francisco, 1982.
25. Shirai, Y., *Three-Dimensional Computer Vision*, Springer-Verlag, New York, 1987.
26. Chen, R. K. and Tasi, W. H., "3-D Object Surface Data Acquisition Using Range Data," MIST Technical Report no. TR-MIST-E76008, National Chiao Tung University, Hsinchu, Taiwan, Republic of China, 1987.
27. Liaw, C. D., "Understanding of 3-D Polyhedral Object from Mutiple Views," M.S. Thesis, Institute of Applied Mathematics, National Chung Hsing University, 1987.
28. Tsai, C. W., "Reconstruction of Single 3-D Object from Multiple Views," M.S. Thesis, Department of Computer Science and Information Engineering, National Taiwan University, 1986.
29. Requicha, A. A. G., "Representation for Rigid Solid : Theory, Method and System," *Computer Surveys*, vol. 12, no. 4, 1980.
30. EUREKA : THE SOLVER, Version 1.0, Borland, 1986.