

行政院國家科學委員會專題研究計畫 成果報告

多張影像間幾何投射分析與其運用於強健式特徵點抽取及 相機投射矩陣估測

計畫類別：個別型計畫

計畫編號：NSC94-2213-E-009-127-

執行期間：94年08月01日至95年07月31日

執行單位：國立交通大學資訊工程學系(所)

計畫主持人：陳稔

報告類型：精簡報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 95 年 10 月 31 日

多張影像間幾何投射分析與其運用於強健式特徵點抽取及相機投射矩陣估測

Multiple-View Image Projection Analysis and Its Use for Robust Feature Point Extraction and Camera Projection Matrix Estimation

計畫編號：NSC-94-2213-E-009-127-

執行期限：94 年 8 月 1 日至 95 年 7 月 31 日

主持人：陳稔教授 / 國立交通大學資工系

E-mail: zchen@csie.nctu.edu.tw Tel: 03-5731875

1. 摘要

在電腦視覺研究中，我們要針對二張或更多張不同視度下所拍攝的景物影像，進行三維景物幾何形狀以及相機相互之間的幾何位置的推測典型的視覺重建步驟包括：

- (1) 抽取影像中具有區辨能力的特徵點或特徵線。
- (2) 尋找不同影像間的對應特徵點或線。
- (3) 估算影像的 epipole 及該相機的等效投影矩陣。
- (4) 估算景物的三維投影幾何結構(projective structure)。
- (5) 另一方面也可以運用多於二張的影像中的對應像點進行像點矩陣(data matrix D)的矩陣分解(matrix factorization)以獲得三維景物的幾何結構及相機內外參數資料。

我們將先建立影像間幾何轉換矩陣 H 。其次，我們將矩陣 H 分解為三個基本型矩陣 $H = H_t H_a H_p$ ，並推導 H_p ， H_a ， H_t 對影像作用的幾何特性，以瞭解 H_p ， H_a ， H_t 對影像所產生的影響。我們再針對矩陣 H 的幾何影響，考慮如何設計影像上特徵點(feature points) 的抽取，以使得部份特徵點在矩陣 H 的作用下，仍然能維持大致不變，以做為兩張影像間的點對應之用。我們將分析採用 Gabor filtering 技術做特徵點擷取(feature extraction)的適當性及濾波器的參數如何設計，以便所抽取的特徵點向量在前述 H_p ， H_a 及 H_t 的轉換下不致全部被破壞掉，再用這些對應像點估算

Fundamental matrix 及相機的等效投影矩陣，並求三維幾何結構。若能拍攝到足夠張的影像及獲得足夠的對應點數，用對應點的 data matrix 做 matrix factorization，求取相機投射矩陣及三維幾何結構，以觀察估測的準確度是否可提高？我們將做電腦模擬，以印證本計劃所發展的理論及瞭解其的實用性。

關鍵詞：相機投射矩陣，影像轉換，投射變形，旋轉，位移，多尺寸，多角位，嘉實濾波器，強健特徵點，基本矩陣，矩陣分解

Abstract:

In computer vision, two or more images taken under different viewpoints are used to estimate 3D projective structure and camera models. Typical steps in such a reconstruction task include:

- (1) Extraction of feature points with discriminative power
- (2) Point matching between image pairs
- (3) Computation of eipoles and canonical projection matrices
- (4) 3D projective structure determination
- (5) Factorization method for obtaining projective structure and motion from multiple views

We start with the computation of transformation matrix between the two given images. The transformation matrix is then decomposed into three elementary projection matrices: H_p ， H_a ， H_t . We

shall apply these matrices in order to any image to analyze their geometrical effects on the image projection. Based on the above analysis, we shall design a Gabor filtering technique for feature point extraction, aiming to sustain the attacks of the projection matrices so that some feature points will survive to be used in a latter process of point correspondence finding. We use point correspondences thus obtained to compute the fundamental matrices and eipoles, and also find the canonical camera projection matrices to further recover the 3D projective structure. We shall also attempt to use multiple images to apply the factorization method to derive the 3D projective structure and camera models simultaneously and evaluate the performance of the result.

Keywords: Camera projection matrix, Image transformation, Perspective distortion, Rotation, Translation, Multi-scale, Multi-orientation, Gabor filter, Robust feature point, Fundamental matrix, Matrix factorization.

2. Introduction

在電腦視覺研究中，我們要從二張或更多張所拍攝的影像，進行三維物體幾何模型以及相機內外參數的推測。標準的處理程序包括：

- (1) 抽取影像中的特徵點或特徵線。
- (2) 尋找不同影像間對應的特徵點或線。
- (3) 估算相機間的 Fundamental matrix。
- (4) 估算影像的 epipole 及該相機的投影矩陣 (projection matrices)。
- (5) 估算景物的三維投影幾何結構 (projective structure)。
- (6) 另一方面也可以運用多張影像中夠多的對應點，進行像點矩陣 (data matrix) 的矩陣分解 (matrix factorization) 獲得三維景物的幾何模型及相機內外參數資料。

目前雖然有不少方法可以抽取影像中的特徵

點[1]-[13]，但是它們分別都在某些假設下才能正常工作。若在一般情況下，同時受到不勻稱尺度變化、任意旋轉、以及投影變形的多種因素干擾，則上述方法所抽取的特徵點會隨拍攝而改變，而無法拿來做點對應。Gabor filter 或 Gabor wavelet 近來常用做 object 及 edge 偵測、image browsing、image compression、texture segmentation 等用途[14]-[26]。它是一種同時具有空間與頻率兩種參數的函數，故含有 multi-scale 及 multi-orientation 的功能。然而，Gabor filters 的參數設計在不同的應用下會有不同考慮。在本計畫我們不強調影像搜尋用的描述功能 (description power)，也不強調針對特定 object、edge 或 texture 的 segmentation power，而是要強調在 projective transformation matrix 下的特徵點的強健性 (feature robustness)。

特徵點之間的比對，由於有些特徵點會因拍攝時被遮擋或受影像轉換破壞而消失掉，故一般在兩張影像中尋找對應點並不是件容易的事。最常見的方法是用 correlation 方法，配上 ordering constraint，以及 epipolar constraint 等條件協助，再用 dynamic programming 技巧來做最佳的比對[27]。但若相機拍攝時移動過大，則不適合在兩張影像上直接套用 correlation 做比對[28]。我們將用特徵點本身局部的結構特性來做為比對依據，故要探討轉換矩陣對特徵點局部結構的影響，看要如何使特徵點不會全面性受破壞，而有一些特徵點能比對成功。

建立至少八組對應點，則可以計算拍攝相機之間空間關係。此乃透過 fundamental matrix 估測，建立 epipolar geometry[29]-[31]。而 fundamental matrix 的建立可以反過來協助尋找更多特徵點，可將特徵點搜尋空間降為一維直線。下一步的三維重建工作，屬於電腦視覺最重要的一項課題。如何建出準確、快速、強健的結果是研究的重點[29]-[31]。本實驗室曾用未做過相機校正過的影像，進行用多個平面逼近三維景物[32]。近年來用多張影像對應點的 data matrix 做 matrix factorization，以重建 affine structure 或 projective structure 是一項突破

[33]-[35]。我們也將利用所獲得的精確對應點群，做此種方式之重建，看是否能獲得更好的結果？本計畫主要工作內容包括對特徵點抽取及比對兩項問題提出更好辦法，再用這些對應點來估算 Fundamental matrix、相機投影矩陣及三維重建，最後評估所得結果。工作將由影像間轉換矩陣的分析下手。轉換矩陣可依其幾何複雜度分為 projection matrix \rightarrow affine matrix \rightarrow similarity matrix 三層 [30]。我們將轉換矩陣分解為上述三種基本型矩陣 (elementary matrices) 的乘積。然後，分析各個基本型矩陣對影像圖形變形的影響。再由這些分析去設計特徵點抽取及比對方法。我們將用實際影像做電腦模擬，以印證所發展的理论及實用性。

3. Geometric properties of projective matrix

我們將先建立影像間之幾何轉換矩陣，若 scene 距離相機遠大於景物本身尺寸時，兩張景物的投射可近似為 weak-perspective projection，則

$$\begin{cases} \bar{u} = [A | \bar{b}] \begin{bmatrix} \bar{P}_3 \\ 1 \end{bmatrix} = A\bar{P}_3 + \bar{b} \\ \bar{u}' = A'\bar{P}_3 + \bar{b}' \end{cases}$$

令 $\bar{q} = \bar{u} - \bar{u}_0 = A(\bar{P}_3 - \bar{P}_{3,0}) = A\bar{Q}_3$; $\bar{P}_{3,0}$ is a specified point
 $\bar{q}' = \bar{u}' - \bar{u}'_0 = A'(\bar{P}_3 - \bar{P}_{3,0}) = A'\bar{Q}_3$

$$\bar{Q}_3 = A^T (AA^T)^{-1} \bar{q}$$

$$\bar{q}' = A' A^T (AA^T)^{-1} \bar{q}$$

$$\begin{pmatrix} \bar{q}' \\ 1 \end{pmatrix}_{3 \times 1} = \begin{pmatrix} A' A^T (AA^T)^{-1} \bar{q} \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} A' A^T (AA^T)^{-1} \bar{0} \\ \bar{0}^T \end{pmatrix}_{H_{3 \times 3}}}_{H_{3 \times 3}} \begin{pmatrix} \bar{q} \\ 1 \end{pmatrix}_{3 \times 1}$$

或

$$\bar{u}' = H_{3 \times 3} \bar{u}$$

所以上述兩張影像之間的轉換 H 是一個 affine matrix。若再考慮 perspective 投影因素，上述 H 矩陣的 h_{31} 及 h_{32} 不再為零，唯值很小，故成為一個通常的 projective matrix。

當有了兩張圖存在一個 projective 轉換矩陣

$\bar{u}' = H\bar{u}$ 時，我們將探討如何尋找強健的特徵點？我們探討的方法是將 projective matrix H 分解成三個矩陣： $H = H_t H_a H_p$

$$\text{假設 } H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix}$$

$$\text{定義 } H = \begin{bmatrix} 1 & 0 & h_{13} \\ 0 & 1 & h_{23} \\ 0 & 0 & 1 \end{bmatrix} \left[\begin{array}{c|c} A_{2 \times 2} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \hline 0 & 1 \end{array} \right] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_{31} & h_{32} & 1 \end{bmatrix}$$

then

$$\begin{aligned} A &= \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} - \begin{bmatrix} h_{13} \\ h_{23} \end{bmatrix} \begin{bmatrix} h_{31} & h_{32} \end{bmatrix} \\ &= \begin{bmatrix} h_{11} - h_{13}h_{31} & h_{12} - h_{13}h_{32} \\ h_{21} - h_{23}h_{31} & h_{22} - h_{23}h_{32} \end{bmatrix} \end{aligned}$$

因此，可獲得此 3 個基本 projective matrices 如下：

$$H_t = \begin{bmatrix} 1 & 0 & h_{13} \\ 0 & 1 & h_{23} \\ 0 & 0 & 1 \end{bmatrix},$$

$$H_a = \begin{bmatrix} h_{11} - h_{13}h_{31} & h_{12} - h_{13}h_{32} & 0 \\ h_{21} - h_{23}h_{31} & h_{22} - h_{23}h_{32} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and}$$

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_{31} & h_{32} & 1 \end{bmatrix}.$$

以下我們逐一分析此 3 個基本 projective matrices 對影像變形所產生之幾何影響。

3.1 H_t 對影像變形所產生之幾何影響

$$H_t: \begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} h_{31} \\ h_{32} \end{bmatrix}.$$

矩陣 H_t 的影響則是平移性的 (translation)，故不會造成影像尺寸或形狀的改變。

3.2 H_a 對影像變形所產生之幾何影響

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} h_{11} - h_{13}h_{31} & h_{12} - h_{13}h_{32} \\ h_{21} - h_{23}h_{31} & h_{22} - h_{23}h_{32} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$= A \begin{bmatrix} u \\ v \end{bmatrix} = Q_1 \Sigma Q_2^T \begin{bmatrix} u \\ v \end{bmatrix}$$

The transformation is also linear and spatially invariant, so it is a constant transformation. The geometric effect on the view transformation comprises two positive (or negative) rotations (i.e., the two orthogonal matrices Q_1 and Q_2) plus a scaling (i.e., the diagonal matrix Σ). Here $Q_1 \Sigma Q_2^T$ is the singular value decomposition of the affine matrix A . Let A' be the inverse of A . Under H_a a set of parallel lines: $\{\alpha u + \beta v = \gamma_i, i=1,2,3,\dots\}$ in the reference image becomes a set of parallel lines of the form $\{\gamma_i = [\alpha \ \beta] \begin{bmatrix} u \\ v \end{bmatrix} = [\alpha \ \beta] A' \begin{bmatrix} u' \\ v' \end{bmatrix}; i=1,2,3,\dots\}$ in the sensed image.

Property 1: 對任一 Affine transformation

$$H_a = \left[\begin{array}{c|c} A_{2 \times 2} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \hline \begin{bmatrix} 0 & 0 \end{bmatrix} & 1 \end{array} \right], A_{2 \times 2} \text{ 可分解成 } Q_1 \Sigma Q_2^T, \text{ 其物理意義}$$

為經由 Q_2^T 旋轉後，經由 Σ 之 anisotropic scaling 之壓縮，再做 Q_1 之旋轉，其所造成之幾何變形為全面性與立即性之變化，且由 $\det(Q_1)$ 或 $\det(Q_2^T)$ (二者 sign 會相同) 的正負號可以知道其壓縮之方向：

$$\det(Q_2^T) > 0 \text{ 壓縮方向在 } y \text{ 軸}$$

$$\det(Q_2^T) < 0 \text{ 壓縮方向在 } x \text{ 軸}$$

Property 2: 對任一 $H_a = \left[\begin{array}{c|c} A_{2 \times 2} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \hline \begin{bmatrix} 0 & 0 \end{bmatrix} & 1 \end{array} \right]$, s.t.

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = A \begin{bmatrix} u \\ v \end{bmatrix} = Q_1 \Sigma Q_2^T \begin{bmatrix} u \\ v \end{bmatrix} = Q_1 \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} Q_2^T \begin{bmatrix} u \\ v \end{bmatrix};$$

$\lambda_1 \geq \lambda_2$, 則存在一組正方形的四個點 $U_{i,i=1,2,\dots,4}$, 在原座標軸上以此四個點之中心作旋轉後獲得之新

座標 $U'_i = R U_i$, 則 $H_a U'_{i,i=1,2,\dots,4}$ 會成為一矩形

$$\text{Where } R = \begin{cases} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} Q_2^T & \text{if } \det(Q_2^T) < 0 \\ (Q_2^T)^{-1} & \text{if } \det(Q_2^T) > 0 \end{cases}$$

我們以實例來實際說明上述 property，圖 1 (a) 中之草綠色為以原 image 邊界做正方形切割之 grid，其經 H_a 轉換後如圖 1 (b) 之草綠色，此轉換之 grid 呈現為菱形，在此情況下，由於受 2 個旋轉矩陣影響，並不容易判斷其變形壓縮之方向。

但若將圖 1 (a) 草綠色圖形經 R_a 旋轉 $U' = R_a U$ (如圖 1 (a) 紅色圖形)

$$\text{where } R_a = \begin{cases} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} Q_2^T & \text{if } \det(Q_2^T) < 0 \\ (Q_2^T)^{-1} & \text{if } \det(Q_2^T) > 0 \end{cases}$$

則 $H_a U'$ 會如圖 1 (b) 紅色圖形所示呈現為一矩形，我們可以很清楚的判斷變形壓縮之方向。

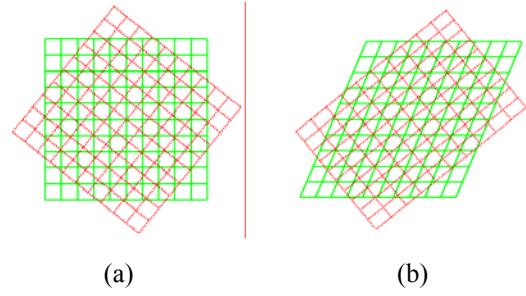


圖 1 H_a 對影像變形所產生之幾何影響

3.3 H_p 對影像變形所產生之幾何影響：

$$\text{Under } H_p: \begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} u / (h_{31}u + h_{32}v + 1) \\ v / (h_{31}u + h_{32}v + 1) \end{bmatrix}.$$

This transformation is nonlinear and spatially variant, so it is not a constant transformation. In the following we shall derive the properties of this transformation and its geometric effect. Then we point out that under the practical constraint on the values of h_{31} and h_{32} , the perspective transformation is locally equivalent to an affine transformation.

Lemma 1. Under the projective transformation

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_{31} & h_{32} & 1 \end{bmatrix} \text{ a line of the form } au + bv = c \text{ is}$$

transformed into a line of the form $(a + ch_{31})u' + (b + ch_{32})v' = c$.

Lemma 2: Under the projective transformation

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_{31} & h_{32} & 1 \end{bmatrix} \text{ the vanishing point of a set of}$$

parallel lines: $\{au + bv = c_i; i = 1, 2, 3, \dots\}$ is given by

$$\begin{bmatrix} u'_\infty \\ v'_\infty \end{bmatrix} = \begin{bmatrix} b/h_{31}b - h_{32}a \\ -a/h_{31}b - h_{32}a \end{bmatrix}$$

Lemma 3: Under the projective transformation

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_{31} & h_{32} & 1 \end{bmatrix} \text{ the set of parallel lines of the form}$$

$\{h_{31}u + h_{32}v = c_i; i = 1, 2, 3, \dots\}$ is transformed into a set of parallel lines of the form $\{h_{31}u' + h_{32}v' = c_i / (1+c_i); i = 1, 2, 3, \dots\}$. This is the only set of parallel lines that remains parallel after the transformation.

Lemma 4:

The set of parallel lines perpendicular to a set of parallel lines: $\left\{ \frac{h_{31}}{\sqrt{h_{31}^2 + h_{32}^2}}u + \frac{h_{32}}{\sqrt{h_{31}^2 + h_{32}^2}}v = c_i; i = 1, 2, 3, \dots \right\}$, where c_i is the inter-distance between two neighboring lines, is of the form

$$\left\{ \frac{h_{32}}{\sqrt{h_{31}^2 + h_{32}^2}}u - \frac{h_{31}}{\sqrt{h_{31}^2 + h_{32}^2}}v = d_i; i = 1, 2, 3, \dots \right\}$$

Lemma 5:

Define a new orthogonal coordinate system (a, b) in the (u, v) orthogonal coordinate system according to H_p :

$$\begin{bmatrix} a \\ b \\ 1 \end{bmatrix} = \begin{bmatrix} (h_{31}u + h_{32}v) / \sqrt{h_{31}^2 + h_{32}^2} \\ (-h_{32}u + h_{31}v) / \sqrt{h_{31}^2 + h_{32}^2} \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} h_{31} / \sqrt{h_{31}^2 + h_{32}^2} & h_{32} / \sqrt{h_{31}^2 + h_{32}^2} & 0 \\ -h_{32} / \sqrt{h_{31}^2 + h_{32}^2} & h_{31} / \sqrt{h_{31}^2 + h_{32}^2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

, then

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} h_{31} / \sqrt{h_{31}^2 + h_{32}^2} & -h_{32} / \sqrt{h_{31}^2 + h_{32}^2} & 0 \\ h_{32} / \sqrt{h_{31}^2 + h_{32}^2} & h_{31} / \sqrt{h_{31}^2 + h_{32}^2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ 1 \end{bmatrix}$$

And

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = H_p \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} u / h_{31}u + h_{32}v + 1 \\ v / h_{31}u + h_{32}v + 1 \\ 1 \end{bmatrix} =$$

$$\begin{bmatrix} (h_{31}a - h_{32}b) / \sqrt{h_{31}^2 + h_{32}^2} \left(a\sqrt{h_{31}^2 + h_{32}^2} + 1 \right) \\ (h_{32}a + h_{31}b) / \sqrt{h_{31}^2 + h_{32}^2} \left(a\sqrt{h_{31}^2 + h_{32}^2} + 1 \right) \\ 1 \end{bmatrix}$$

Property 3: 對於任何一個 H_p , reference image 上平行於消失線方向 $\begin{pmatrix} h_{31} \\ h_{32} \end{pmatrix}$ 的線，在 transformed image

上仍然會與消失線平行，而原先垂直於消失線方向 $\begin{pmatrix} -h_{32} \\ h_{31} \end{pmatrix}$ 的線，在 transformed image 上則會造成 fore-

shortening 現象，而交於一個消失點。

圖 2(a)中之草綠色為以原 image 邊界做正方形切割之 grid，其經 H_p 轉換後如 Fig. 2(b) 之草綠色，此轉換之 grid，變化沒有規則性，並不利於分析。依 Property 3，我們依消失線方向 $\begin{pmatrix} h_{31} \\ h_{32} \end{pmatrix}$ 、 $\begin{pmatrix} -h_{32} \\ h_{31} \end{pmatrix}$

之座標軸重新切割 grid 上，如圖 2(a) 紅色虛線所示，此 grid 經 H_p 轉換後如圖 2(b) 之紅色虛線。

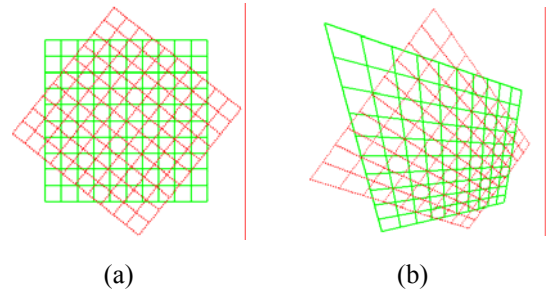


圖 2 H_p 對影像變形所產生之幾何影響

我們可以很明顯的觀察到如 Property 3 中說明之 H_p 幾何特性：

- (1) 垂直於消失線主軸方向之平行線間距會持續

漸進性的變化,造成因位置而異之 scaling 變化(接近消失點呈現縮小,其 scaling <1,遠離消失點則呈現放大現象,其 scaling >1)。

(2) 在成像光軸之中心點,grid 會呈現 uniform scaling 變化,隨著遠離光軸中心點,各地方均會呈現與位置有關之不同量 anisotropic scaling。

一般而言,為簡化 Perspective 變形,許多研究會想到將整張 image 之 Perspective deformation approximate 成單一之 Affine transformation,但由上圖中我們可以觀察到,僅在投影中心點為一長方形,而其餘地方為梯形,且各地方之梯形形狀及角度均因位置而異,故我們無法找到一個 global 一致的 transform,使得所有地方之梯形均變成長方形,而是要依因位置不同所造成之不同梯形個別去找其相近的平行四邊形。

Property 4: 影像上任一位置之 perspective deformation 程度不一,愈偏離消失線主軸愈大,但 local 來看均不嚴重,可以 approximate 成平行四邊形,這些位置不同之平行四邊形可以看成矩形(類似 H_a ,此平行四邊形如以另一角度看,可以變成矩形),再經一個不同旋轉角度做 anisotropic scaling 而變形,此旋轉角度非 global 的一致量,而是因位置而異。

圖 3(a)(b)說明依 Property 3,將 reference image 上之正方形 grid $U_{i,j=1,2,3,4}$ 經 R_p 旋轉使其一組平行邊與消失線方向平行 $U'_{i,j=1,2,3,4} = R_p U_{i,j=1,2,3,4}$,此 $U'_{i,j=1,2,3,4}$ 再經 $H_p U'_{i,j=1,2,3,4}$ 轉換後會呈現梯形,但此梯形可近似為平行四邊形 ($H_p U' \rightarrow H_a'_{(x,y)} U'$)。圖 3 (c)(d)則依 Property 2,再將 $H_a'_{(x,y)}$ decompose to $Q_1 \Sigma Q_2^T$,找出 rotation matrix $R_{2(x,y)}$,使得圖 3 (c) $U'' = R_{2(x,y)} U'$ 投影成圖 3 (d) $H_p U'' (\approx H_a'_{(x,y)} U'')$

$$\text{where } H_a'_{(x,y)} U'' = Q_1'_{(x,y)} \Sigma'_{(x,y)} Q_2'_{(x,y)}{}^T R_{2(x,y)} U'$$

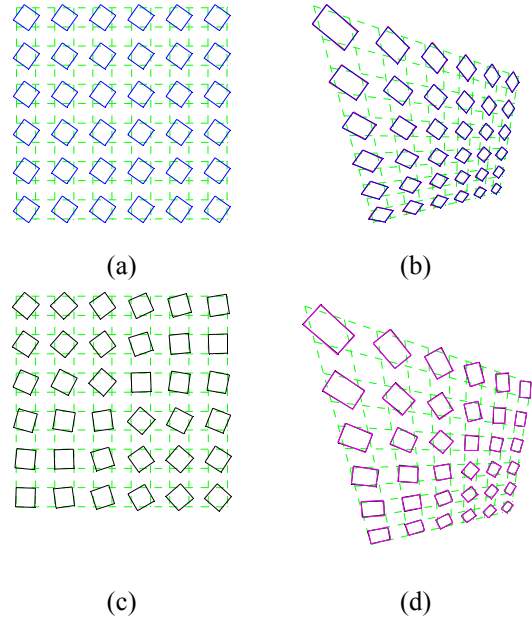


圖 3 H_p 對影像變形所產生之幾何影響

3.4 綜合考慮 $H_t H_a H_p$ 之影響

Property 5: 對 image 上任一組正方形的四個點 $U_{i,j=1,2,3,4}$,其經 $HU = H_t H_a H_p U$ 轉換,可 approximate 成 $HU = H_t H_a H_p U$,令 $H_a'_{(x,y)} = H_a H_p U_{i,j=1,2,3,4}$,則可套用 Property 2,找出 rotation matrix,使得轉換後會是一矩形(H 應用在正方形的 $U'' = R_{2(x,y)} R_1 U$,非 U)。

$H_a H_p U$ 仍會有一組線會與消失線平行,但方向與 $H_p U$ 之平行線方向不同 ($H_a H_p X = Q_1 \Sigma Q_2^T H_p X$,也就是 $H_p X$ 經過旋轉後再壓縮,因此, $H_p X$ 之平行線方向與 $H_a H_p X$ 之平行線方向不一樣),其性質與 H_p 類似,但 anisotropic scaling 會產生變化。

4. Robust Feature Extraction Using Scale

/Orientation - adaptive Gabor Filtering Technique

依據上面的分析,影像在轉換矩陣作用之下,會產生影像不勻稱的尺寸變化,位置旋轉,漸近的形狀扭曲。如想在這些狀況下尋找不易被破壞的影

像特徵點，則所要的特徵點必須是由局部性的結構 (local structure pattern) 上出發。其次，全域中有不同的尺寸變化，故特徵點要不受 scaling 的影響。除此之外，還要考慮在同一張影像先天具有大小不同的特徵結構，故特徵點的抽取不能只針對固定一種視窗做運算。它必須是 multi-scale、multi-orientation 的。

- (1) 為抓出各種不同尺寸物件結構之特徵，及因應上述 projective transformation 下 scaling 及 rotation 變化會因位置不同而不同之現象，我們提出一個 scale/orientation-adaptive Gabor filtering technique, 此方法係基於 multi-scale, multi-orientation 之觀念，且每一 local pattern 具有 dominant scaling 及 dominant orientation 的特性。

4.1 Important Properties of the Gabor-filtered Image

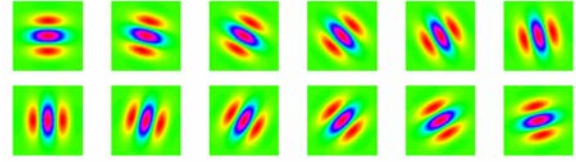
Gabor function :

$$g^{s,N,l}(x,y) = \frac{1}{2\pi\alpha\sigma_s^2} \exp\left\{-\left[\frac{x'}{\sigma_s}\right]^2 + \left[\frac{y'}{\alpha\sigma_s}\right]^2\right\} \exp(jw_s x')$$

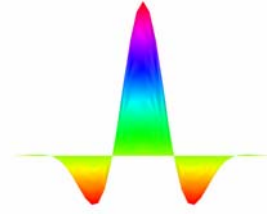
where

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta_l & \sin\theta_l \\ -\sin\theta_l & \cos\theta_l \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

with θ_l being the orientation parameter; σ_s and $\alpha\sigma_s$ are the Gaussian window size parameters ($\alpha = 1$ is assumed here); w_s is the spatial frequency parameter. A normalization condition is usually imposed on the parameters σ_s and w_s such that $\sigma_s w_s = N\pi/4$ (Here $4\sigma_s$ is used to approximate the Gaussian window size, which is also called the filter size) for all scale indices s .



(a)



(b)

Fig. 4 (a) A set of even-symmetric Gabor filters. **(b)** The plot of a typical even-symmetric Gabor filter with $N = 3$.

Let $I(x,y)$ be the input image function. For multiple scales s , $s \in \{1, 2, 3, \dots, S\}$, and multiple orientations θ_l , $\theta_l = l \times \Delta\theta$, $l = 1, 2, \dots, L$ (π is a multiple of $\Delta\theta$), the filter responses or outputs are given by the convolution operations:

$$R^{s,N,l}(x,y) = I(x,y) * g_{even}^{s,N,l}(x,y)$$

In the following, we give the important properties of Gabor-filtered image that lead to the robust feature points.

Property 6: Let $\hat{R}^{s,N,l}(x,y)$ and $R^{s,N,l}(x,y)$ be the filter responses to the images $\hat{I}(x,y)$ and $I(x,y)$. If $\hat{I}(x,y)$ is obtained from $I(x,y)$ by a rotation through an angle ϕ in the counter-clock direction, i.e., $\hat{I}(x,y) = I(x',y')$ for points (x,y) and (x',y') that are related by

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Then $\hat{R}^{s,N,l}(x,y) = R^{s,N,\theta_l+\phi}(\bar{x},\bar{y})$

$$\text{where } \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Property 7: Let $\hat{R}^{s,N,l}(x,y)$ and $R^{s,N,l}(x,y)$ be the Gabor filter responses to input images $\hat{I}(x,y)$ and $I(x,y)$. If the two input images are related by a scale factor such that $\hat{I}(x,y) = I(kx,ky)$, then there is a correspondence between their responses given by $\hat{R}^{s,N,l}(x,y) = R^{s',N,l}(kx,ky)$ with $\sigma_{s'} = k\sigma_s$

We compute the responses $R^{s,N,l}(x,y)$ at each point (x,y) for $s = 1, 2, \dots, S$ and $l = 1, 2, \dots, L$ (N is fixed).

Then the total energy value at point (x,y) at a scale s is defined as

$$E^{s,N}(x,y) = \sum_{l=1}^L |R^{s,N,l}(x,y)|$$

Then the maximum energy value is calculated at point (x,y) with respect to all scales $\{1,2,\dots,S\}$ (N is fixed)

$$E^N(x,y) := E^{s^*,N}(x,y) = \max_s \{E^{s,N}(x,y)\}$$

Property 8. Consider the Gabor filter output energy $E^{s,N}(x,y)$ for a binary rectangular object image with a fixed filter size parameter σ_s and a varying lobe parameter N . Along a longer rectangle side, the maximum energy value occurs at consecutive points near the middle point of the side for $N \leq 1$; while for $N > 1$, the maximum energy value occurs at a single point near the vertex of the side. (Please refer to Fig. 5 for an illustration)

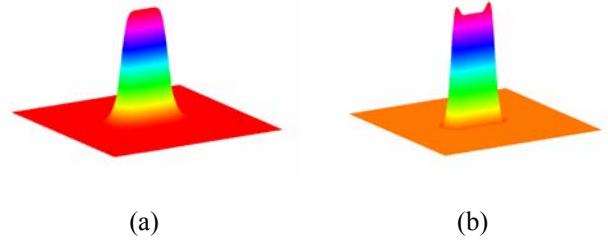


Fig. 5 (a) There are no dominant points in the maximum energy map for a rectangular object image of size 8×33 for $N=1$. (b) There are dominant points in the maximum energy map for $N=3$.

4.2 Dominant Point Extraction

Based on the above properties of the Gabor filtered image, we outline below an algorithm for extracting dominant points from a given image.

Algorithm for dominant point extraction:

- (1) Choose a set of Gabor filters with filter parameters $N = 3$ and $L = 12$.
- (2) Determine a set of filter size scales based on the structure pattern of the object under consideration. Apply the multi-scale Gabor filters to the image to obtain the energy maps at the multiple scales and merge them into a single maximum energy map by retaining the maximum value among them.
- (3) Construct the maximum energy histogram and choose a lower bound on the energy value for a candidate dominant point in the upper $T\%$. ($T = 10$ in our case)
- (4) Partition the image into non-overlapping blocks of $m \times m$ pixels each ($m=7$ in the experiments).
- (5) Find one candidate dominant point having the maximum energy in each block. Then check if

each candidate dominant point obtained so far is also a local maximum in a neighborhood of size $m \times m$ centered at the point.

4.3 The feature vector representation of Gabor-based feature point

We use a feature vector \vec{V}_{p_i} consisting of L filter responses to a set of Gabor filters with an incremental orientation step of π/L to characterize the local pattern around the feature point $p_i(x_i, y_i)$. The principal scale $s_{p_i}^d$ and the principal orientation $\theta_{p_i}^d$ are iteratively tuned to a high accuracy. The mathematical form of a Gabor-based feature point is given as

$$\vec{V}_{p_i} = [R^{(s_{p_i}^d, \theta_{p_i}^d)}(x_i, y_i), R^{(s_{p_i}^d, \theta_{p_i}^d + \frac{1}{L}\pi)}(x_i, y_i), \dots, R^{(s_{p_i}^d, \theta_{p_i}^d + \frac{L-1}{L}\pi)}(x_i, y_i)]^T$$

The similarity $S(\vec{V}_{p_i}, \vec{V}_{q_j})$ between feature points \vec{V}_{p_i} and \vec{V}_{q_j} can be measured by the normalized cross correlation:

$$S(\vec{V}_{p_i}, \vec{V}_{q_j}) = \frac{(\vec{V}_{p_i} - \text{mean}(\vec{V}_{p_i})) \cdot (\vec{V}_{q_j} - \text{mean}(\vec{V}_{q_j}))}{\|\vec{V}_{p_i} - \text{mean}(\vec{V}_{p_i})\| \cdot \|\vec{V}_{q_j} - \text{mean}(\vec{V}_{q_j})\|}$$

5. 實驗結果

首先，我們使用 synthetic data 來分析 image 上任一位置對變形之抵抗性，實驗是將圖 6(b) 之 pattern 經過特定角度旋轉後放在圖 6(a) background image 上之特定位置，產生一張 reference image(如圖 7(a))，此 reference image 經 H 轉換後產生一張 sensed image(如圖 7(b))，在此二張 image 上分別去抓特徵點，產生 feature vector 作比對。

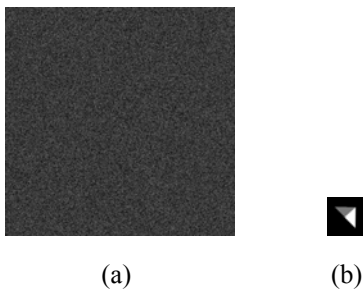


圖 6 (a)Background, (b)pattern

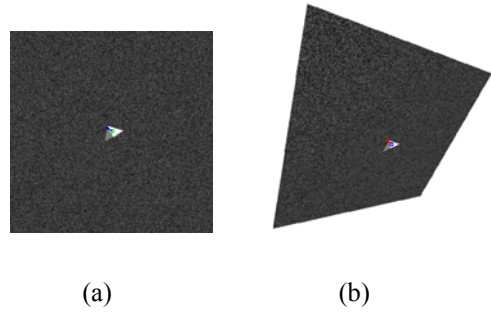


圖 7(a)Reference image, (b)sensed image

圖 8 為分別在各種位置上將 pattern 於 0~360 度間每隔 15 度作一次旋轉，reference 及 sensed 影像上每組對應的特徵點作特徵比對，對於比對係數大於 0.9 者，我們顯示其 dominant orientation，以觀察各種不同位置之影像變形對特徵點存活率之影響，實驗結果可歸納如下：

- (1) 對於趨近於 isotropic scaling 變形的地方(如左上角點編號 1)，所有方向特徵點均能存活。
- (2) anisotropic scaling 變形越嚴重可存活的方向越少，特徵點存活的方向與被壓縮變形的方向垂直。

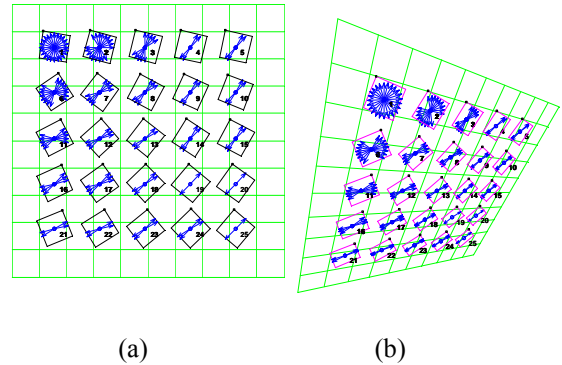


圖 8(a)Reference image, (b)sensed image，顯示比對係數大於 0.9 者之特徵點方向

圖 9 展示特徵點編號 1(影像中之左上角點)及編號 25(影像中之右下角點)各種不同方向之特徵向量比對係數。

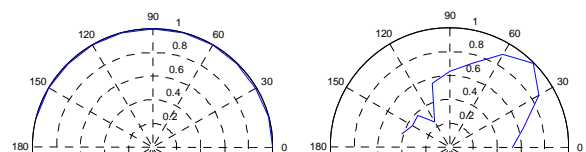


圖9 各種不同方向之特徵向量比對係數(a)點編號 1, (b)點編號 25

接下來，我們使用 real 空照影像對來作實驗(如圖 10)，Reference image 上的點與 sensed image 做比對，比對係數大於 0.9 以上的點其 principal orientation 以 blue solid line 來表示，否則以 red dashed line 來表示；Grid 線是將 reference image partition 成方格，以 transformation matrix 轉換到 sensed image 中以展現同樣範圍內的 pattern 在 sensed image 所對應的範圍，以觀察 sensed image 與 reference image 間成像的變形關係，我們可以觀察到比對成功的點，其 principal orientation 與變形的方向垂直，觀察 feature vector 變化，在 principal orientation 方向附近的 energy 變化不大，相反的，變化大的地方反而是在原本就較弱的方向，因此，他整體的 energy 變化不大，使得他能抵抗在此方向較大的變形。相反的，在平行於變形方向的點中，energy 變化大的地方是在主要方向，有些點會存活，有些點則不會，端看他們的 feature vector 被破壞的程度。

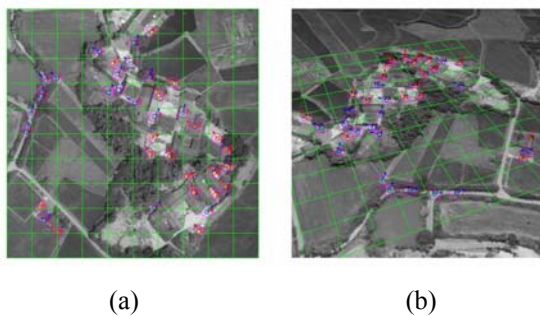


圖 10 空照影像對

圖 11 為利用比對成功的點求取轉換矩陣後套合結果。



圖 11 影像套合結果

在取得強健式特徵點之對應後，我們拍攝足夠張的影像，用對應點的 data matrix 做 matrix factorization，求取相機投射矩陣及三維幾何結構。結果展示在圖 12。在此，由於圖張數較多，未排成兩欄，以一欄放在報告最後。我們對特徵點座標加上 Gaussian noise: mean = 0, variance = 1, 2, 4 以觀察估測的準確度是否受溫和雜訊影響？結果發現多張影像可平衡雜訊影響。圖 12(b)的不同姿態的三維估測點的 wire-frame 圖顯示各張輸入二維影像所對應姿態。

6. 結論

本計畫針對影像間幾何轉換矩陣 H ，分解為轉換矩陣為三個基本型矩陣： $H = H_t H_a H_p$ ，漸次就分析各矩陣對影像產生幾何影響，並根據其結果設計一套強健式特徵點粹取方法，解決影像間特徵點對應比對之難題。在取得強健式特徵點之對應後，我們拍攝足夠張的影像，用對應點的 data matrix 做 matrix factorization，求取相機投射矩陣及三維幾何結構。結果發現估測的準確度不受溫和雜訊影響，多張影像可平衡雜訊影響。

7. 參考文獻

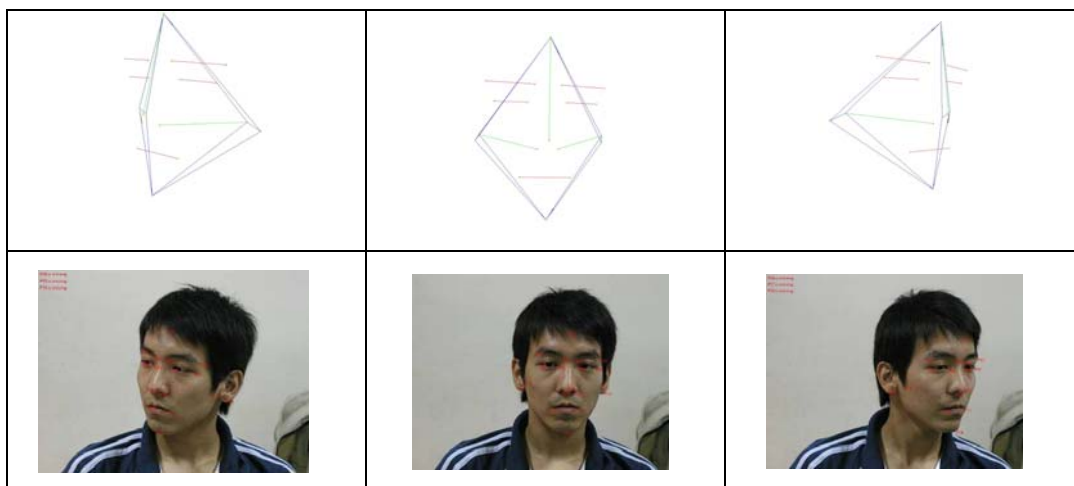
1. A. Rattarangsi and R. T. Chin, "Scale-based detection of corner of planar curves," *IEEE*

- Trans. PAMI*, vol. 14, no.4 (1992), pp.430-449.
2. A. Garrido , P. D. L. Blanca and M. Garcia Silvente , “Boundary simplification using a multiscale dominant-point detection algorithm ,” *Pattern Recognition* , vol.31 , No.6 (1998) , pp.791-804.
 3. J. Fayolle , L. Riou and C. Ducottet , Robustness of a multi-scale scheme of feature points detection , *Pattern Recognition* , vol.33(2000) , pp1437-1453.
 4. L. Kitchen and A. Rosenfeld. Gray-level corner detection. *Pattern Recognition Letter.* (1982) , pp.95-102.
 5. J. A. Noble , Finding corners , *Image Vision Comput.* 6(1988) , pp.121-128.
 6. C. H. Chen , J. S. Lee and Y. N. Sun , Wavelet transform for gray-level corner detection , *Pattern Recognition* , vol.28 , no.6(1995) , pp.853-861
 7. Q. Zheng and R. Chellappa , Automatic feature point extraction and tracking in image sequences for arbitrary camera motion. *International Journal of Computer Vision.* Vol.15(1995) , pp.31-76.
 8. B. S. Manjunath , C. Shekhar and R. Chellappa , A new approach to image feature detection with application , *Pattern Recognition* , vol.29 , no.4 (1996) , pp627-640.
 9. J. W. Hsieh , H. Y. Liao , K. C. Fan , M. T. Ko and Y. P. Hung , Image Registration Using a New Edge-based Approach , *Computer Vision and Image Understanding* , Vol.67 , No.2(1997) pp.112-130.
 10. B. Robbins and R. Owens , 2D feature detection via local energy. *Image Vision and Computing* 15(1997) , pp 353-368.
 11. M. Trajković and M. Hedley , Fast corner detection , *Image and Vision Computing* , vol.16 (1998) , pp.75-87.
 12. J. Chen , Y. Sato , and S. Tamura , Orientation space filtering for multiple orientation line segmentation. *IEEE Trans , PAMI* , vol. 22 , no. 5 (2000) , pp. 417-429.
 13. A.P. Witkin , Scale-space filtering. In proc 8th Int. Joint Conf. *Artificial Intell.* , 1983 , pp. 1019-1021.
 14. R. Mehrotra , K.R. Namuduri , and N. Ranganathsn , “Gabor filter –based edge detection ,” *Pattern Recognition* , Vol. 25 , pp. 1479-1494 , 1992.
 15. A.K. Jain , N.K. Ratha , and S. Lakshmanan , “Object detection using Gabor Filters ,” *Pattern Recognition* , Vol. 30 , pp. 295-309 , 1997.
 16. Z. Wang and M. Jenkin , Using complex Gabor filters to detect and localize edges and bars. In: C. Archibald and E. Petriu , (eds.): *Advanced in Machine Vision: Strategies and Applications* , vol. 32 , River Edge , NJ: World Scientific , 1992 , pp. 151-170.
 17. R.P. Wurtz and T. Lourens , Corner detection in color images through a multiscale combination of end-stopped cortical sells. *Image Vision and Computing* 18(2000) , pp.531-541.
 18. B.S. Mahjunath and W.Y. Ma , “Texture features for browsing and retrieval of image data” , *IEEE TPAMI* , pp. 837-842. 1996.
 19. J.G. Daugman , Two-dimensional spectral analysis of cortical receptive field profile. *Vision Research* , vol. 20(1980) , pp. 847-856.
 20. J.G. Daugman , Uncertainty relation for resolution in space , spatial frequency , and orientation optimized by two-dimensional visual cortical filters. *J. Optical Soc. Amer.* , vol. 2 , no.7(1988) , pp. 1169-1179.

21. J.G. Daugman , Computing discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Trans. Acous. , Speech , Signal Processing* , vol. 36 (1988) , pp. 1169-1179.
22. T. S. Lee , Image representation using Gabor wavelets. *IEEE Trans. PAMI* , vol 18 , no. 10(1996) , pp. 959-97.
23. X. Wu and B. Bhanu , Gabor wavelet representation for 3D object recognition. *IEEE Trans. Image Processing* , vol.6 , no. 1(1997) , pp. 47-64.
24. S.E. Grigorescu , N. Petkov , and P. Kruizinga , “Comparison of texture features based on Gabor filters” *IEEE Trans. on Image Processing* , Vol.11 , No. 10 , pp. 1160-1167 , 2002
25. A. Kumar , and G.K.H. Pang , “Defect detection in textured materials using Gabor filters” *IEEE Trans. on Industry Applications* , Vol. 38 , pp. 425-440 , 2002.
26. D. Dunn and W.E. Higgins , “Optimal Gabor filters for texture segmentation” , *IEEE Trans. on Image Processing* , Vol. 4 , No. 7 , pp. 947-964 , 1995.
27. Y. Ohta and T. Kanade , “Stereo by intra- and inter-scanline search , ” *IEEE TPAMI* , 7(2) , pp. 139-154 , 1985.
28. F. Devernay and O.D. Faugeras , “Computing differential properties of 3D shapes from stereopsis without 3D models , ” *Proc. IEEE Conf. on CVPR* , pp. 208-213 , 1994.
29. Olivier Faugeras , *THREE-DIMENSIONAL COMPUTER VISION* , MIT Press , Cambridge , Massachusetts , 1993.
30. Richard Hartey and A. Zisserman , *MULTIPLE VIEW GEOMETRY , IN COMPUTER VISION* , Cambridge University Press , London , England , 2000
31. David A. Forsyth and Jean Ponce , *COMPUTER VISION: A MODERN APPROACH* , Prentice Hall , New Jersey , 2003.
32. H. L. Chou and Z. Chen , “A novel 3D planar object reconstruction from multiple uncalibrated images using the plane-induced homographies” , To appear in *Pattern Recognition Letters*.
33. J. Costeria and T. Kanade , “A multi-body factorization method for motion analysis , ” *Int. J. Computer vision* , 29(3) , pp. 159-180 , 1998.
34. S. Mahamud and M. Hebert , “Iterative projective reconstruction from multiple views , ” *Proc. IEEE Conf. Comp. Vision and Patt. Recog.* , pp. II-430-437 , 2000.
35. S. Mahamud , M. Hebert , Y. Omori , and J. Ponce , “Provably-convergent iterative methods for projective structure from motion , ” *Proc. IEEE Conf. Comp. Vision and Patt. Recog.* , pp. 1018-1025 , 2001.



Figure 12(a). The multiple images of a human head with extracted feature points.



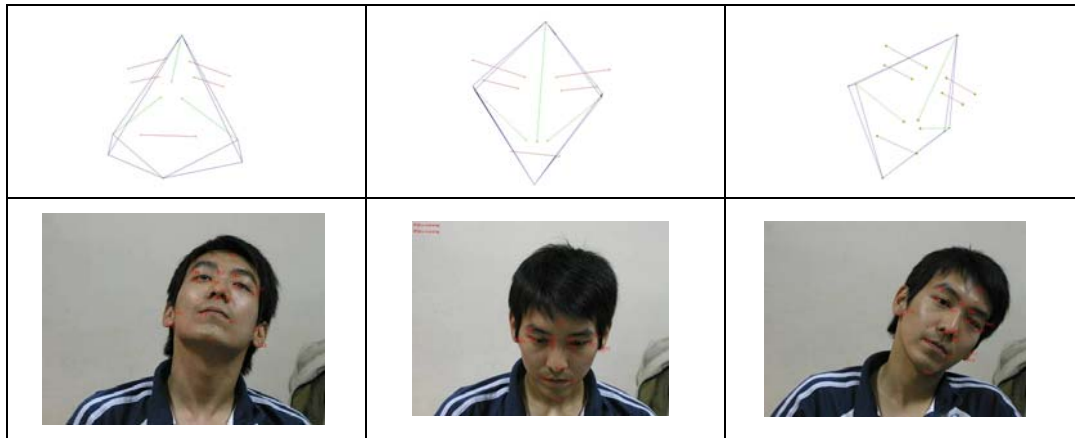


Figure 12(b). The corresponding wire-frame representations (upper) of the reconstructed 3D point structure for each input view (bottom).