

# 一種基於影像分區之有效特徵擷取與匹配方法

## An Efficient Method for Feature Extraction and Matching Based on Image Segmented Regions

黃登淵<sup>1</sup> 黃靖甯<sup>1</sup> 胡武誌<sup>2</sup> 孫宛琳<sup>1</sup> 劉立彬<sup>1</sup>

<sup>1</sup>大葉大學 電機工程系

E-mail: kevin@mail.dyu.edu.tw; daweimailbox@gmail.com;  
happyangel5200@gmail.com; cook8023@hotmail.com

<sup>2</sup>國立澎湖科技大學 資訊工程系

E-mail: wuchih.hu@gmail.com

### 摘要

本文探討如何改善傳統特徵擷取方法，例如 SIFT(Scale-Invariant Feature Transform)、SURF(Speeded Up Robust Features)與 ORB(Oriented BRIEF)等耗時的問題，並提出增進特徵點匹配準確性方法。在提升特徵擷取與匹配效能方面，本文提出分區影像特徵點計算與錯誤匹配對刪除法。首先將影像劃分成 25 個等分，取其中編號 1-8 的分區影像進行特徵點計算；接著再計算匹配特徵點間的角度關係，並做角度值方圖統計，然後將角度落在最大區間的匹配對視為正確的匹配對，並刪除其他的錯誤的匹配對。實驗結果顯示分區影像特徵計算比起整張影像(不分區)特徵計算可以節省 0.5~1 倍的時間；此外，本文所提錯誤匹配對刪除法可有效的降低錯誤的特徵匹配。

**關鍵詞：**SIFT、SURF、ORB、特徵擷取、特徵匹配

### Abstract

This paper investigates how to improve the time efficiency for traditional feature extraction methods, such as SIFT, SURF, and ORB, which are computationally intensive, and proposes an effective method to increase the accuracy of feature matching. To increase the time efficiency for feature extraction, we partition the test image into 25 equal regions and take only eight ones from them for feature extraction. Then, we calculate the angles of these candidate feature matching pairs to build a histogram. The angle with the maximum histogram is taken as the angle of correct feature matching pairs and thereby the matching pairs whose angles are not equal to the one with the maximum histogram are eliminated. Experimental results show the time efficiency can be improved by a factor of 0.5~1.0 using the proposed image segmented regions method. Moreover, the proposed method for error matching pair elimination can effectively reduce the error rate of matching pairs.

**Keywords:** SIFT, SURF, ORB, Feature extraction, Feature matching

### 1. 前言

特徵點的擷取在特徵匹配的計算上極為重要。通常物體影像特徵擷取的結果容易受到平移、旋轉、光線變化與尺度縮放之影響。因此，特徵點擷取的方法在於使擷取的特徵點能夠具有抗旋轉性、尺度不變性、不受光線變化影響及抗雜訊等。如此一來，有效的特徵點可以大幅降低影像匹配的錯誤率。因此，如何有效地提升特徵點擷取與匹配的準確性，進而提升計算的效能為本文研究的重點。

在現存的文獻中，對於特徵點擷取的方法有：尺度不變特徵轉換(Scale-Invariant Feature Transform; SIFT)[1]、加速強健特徵(Speeded Up Robust Features; SURF)[2]、加速分段特徵測試(Features from Accelerated Segment Test; FAST)[3]、二元強健獨立基礎特徵(Binary Robust Independent Elementary Features; BRIEF)[4]與具方向性 BRIEF(ORiented BRIEF; ORB)[5]等。Lowe[1]提出尺度不變特徵轉換方法(SIFT)，影像透過高斯差分生成尺度空間，進而搜尋尺度空間中的極值做為候選特徵點，然後針對候選特徵點刪除對比度低以及不穩定的邊緣點，接著再計算特徵點的方向及強度，藉以產生特徵描述。SIFT 方法擷取出來的特徵點具有尺度縮放不變性、抗旋轉性及抗光線變化等優點。Bay 等人[2]使用加速強健特徵(SURF)，利用積分影像(Integral image)加速 Hessian 矩陣行列式值(Determinant of Hessian; DoH)的運算速度，接著建立尺度空間，並在各尺度空間搜尋 Hessian 矩陣行列式值的極值，計算過程中利用盒子濾波器(Box filter)與積分影像來簡化原影像與二階高斯微分遮罩之運算。SURF 使用 Haar 小波作為特徵描述，並計算特徵點的方向及強度，藉以產生特徵描述。SURF 方法擷取出來的特徵點同樣具有尺度縮放不變性、抗旋轉性及抗亮度變化的優點。Rosten[3]等人提出 FAST 角點偵測方法，並以 7×7 的遮罩對影像進行掃描，同時比較遮罩中心點及其內最大圓圈上 16 個點灰階值的差異絕對值，當其中有 9 個以上的差異值大於設定的閾值，該中心點即視為特徵點。

Calonder[4]等人提出 BRIEF 演算法，在特徵點

擷取之前先利用高斯濾波器對影像去除雜訊，接著計算特徵描述向量，但此方法得到的特徵點不具旋轉不變性、尺度不變性及對雜訊敏感。Rublee[5]等人針對 BRIEF 方法對旋轉及尺度變化敏感之缺點，提出具方向性的 BRIEF，稱為 ORB 演算法。該方法結合 FAST 特徵點檢測和 BRIEF 特徵描述的優點，首先對影像進行高斯模糊並產生尺度空間使特徵點具有尺度不變性，緊接著再對特徵點進行矩向量(Moment vector)運算，使得特徵點具備方向不變性，最後再對特徵點進行 BRIEF 特徵描述，讓所取得的特徵點具有旋轉不變性、尺度縮放不變性以及不易受雜訊干擾等優點。最後，ORB 搭配最佳相鄰點匹配法(Best of two nearest matching)[6]，以降低特徵點匹配錯誤率。

本文接下來的部分包括：第 2 節說明本文的研究方法，第 3 節為結果與討論，最後一節對本文的研究結果進行結論。

## 2. 研究方法

本文為提升特徵點擷取的效率，率先提出影像分區的做法，將原影像分割成 25 區，但僅用其中 8 區來做特徵擷取與匹配。本文所提分區影像特徵匹配方法流程如圖 1 所示。由圖 1 可知，本文所提方法包含分區影像特徵擷取、分區影像特徵匹配與錯誤匹配對刪除等，下文將針對這些方法逐一介紹。

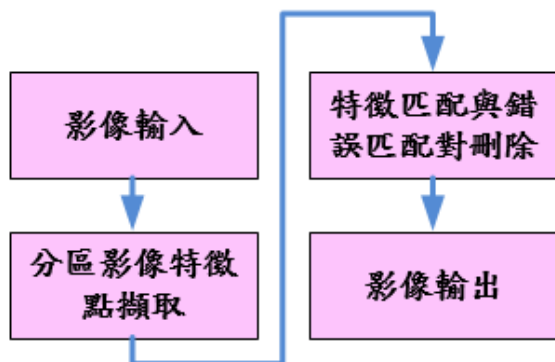


圖 1 本文所提基於分區影像特徵匹配流程圖

### 2.1 基於分區影像特徵點擷取

分區影像特徵點擷取之目的在於加速特徵點擷取的速度。本文首先將影像分割為 25 等分，但僅取其中編號 1-8 之分區影像進行特徵點擷取，本文所提影像分區示如圖 2。

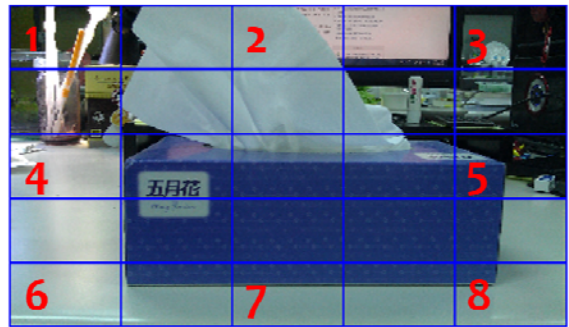


圖 2 基於分區影像特徵點擷取之影像分區示意圖

## 2.2 特徵點擷取

本文採用 3 種常用的特徵點擷取方法，分別為 SIFT、SURF 與 ORB 來進行分區影像之特徵擷取。下文將針對這 3 種方法進行簡述。

### 2.2.1 SIFT 特徵擷取

SIFT 演算法，其做法主要分成特徵點擷取與特徵描述兩部分，如圖 3 所示：

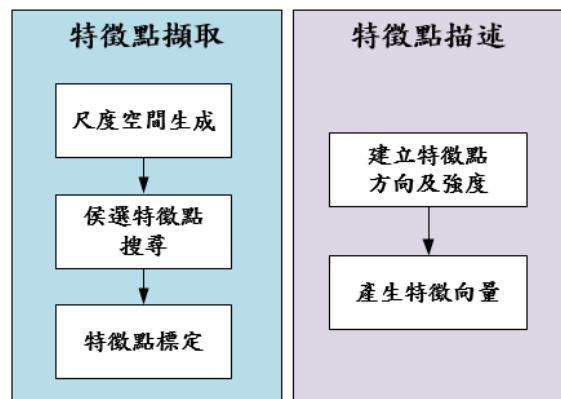


圖 3 SIFT 特徵擷取演算法流程圖

SIFT 演算法在特徵點擷取的部分使用 DoG (Difference of Gaussian)來生成尺度空間，並從中找出極值，這些極值又稱為候選特徵點，用 SIFT 找到的候選特徵點具有尺度不變的特性，接著刪除不穩定的候選特徵點，例如低對比度跟不穩定的邊緣點，如圖 4 所示：

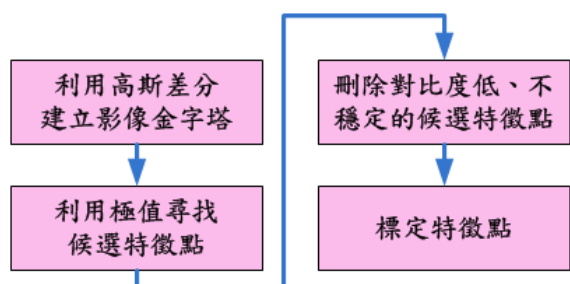


圖 4 SIFT 特徵點擷取流程

當特徵點被標定後，緊接著進行 SIFT 特徵描述。首先計算特徵點的方向，以特徵點為中心，取大小為 $16 \times 16$ 的遮罩區塊，並對區塊內所有像素進行梯度強度與方向的計算，如式(1)與(2)所示：

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

$$\theta(x, y) = \tan^{-1}((L(x+1, y) - L(x-1, y)) / (L(x, y+1) - L(x, y-1))) \quad (2)$$

式中  $m(x, y)$  與  $\theta(x, y)$  分別表示在座標位置  $(x, y)$  的梯度強度與方向， $L$  表示所選區塊內原影像與高斯模糊遮罩的迴旋積值。緊接著可利用梯度方向進行直方圖統計，最大值即為特徵點的方向。當求出所有特徵點的方向之後，接著再對特徵點進行特徵描述，即為特徵向量，圖 5 為 SIFT 特徵描述示意圖。

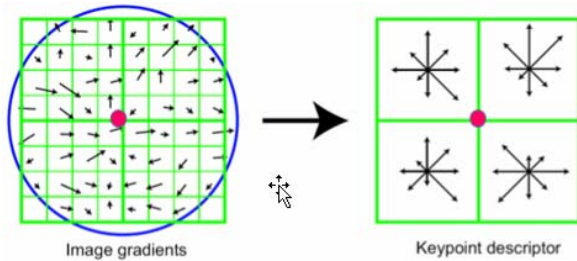


圖 5 SIFT 特徵點描述示意圖

### 2.2.2 SURF 特徵擷取

SURF 演算法流程如圖 6 所示，其包含積分影像的計算、快速 Hessian 特徵檢測、尺度空間生成，特徵點方向標定，最後利用特徵點的大小及方向來做為特徵描述，並藉以產生特徵向量。

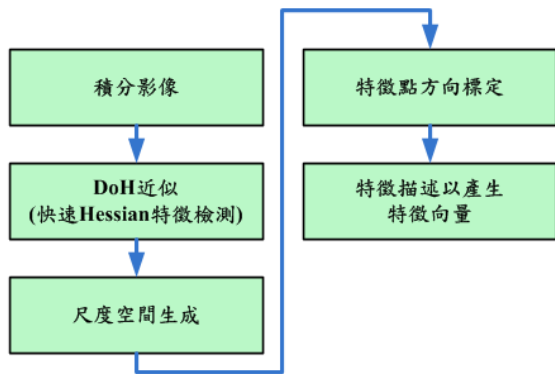


圖 6 SURF 特徵擷取演算法流程圖

積分影像之目的是為加速矩形特徵的計算，將影像與高斯二階微分遮罩的迴旋積計算轉化為對積分影像做加減運算。SURF 的特徵點檢測是利用 Hessian 矩陣行列式值的極大值來確定影像中特徵點的位置。建立影像金字塔可獲得不同尺度的特徵點，SURF 採用盒子濾波器(Box filter)和積分影像，再透過不同尺寸的盒子濾波器與積分影像求得

Hessian 矩陣行列式值的響應圖像，然後在響應圖像上採用 3D 非最大抑制 (Non-maximum suppression)，以取得不同尺度的特徵點。

SURF 對每個特徵點訂定一個主方向，以達到旋轉不變的特性，對影像使用 Haar 小波濾波器，並利用其響應值來計算特徵點的方向，Haar 小波濾波器如圖 7 所示。

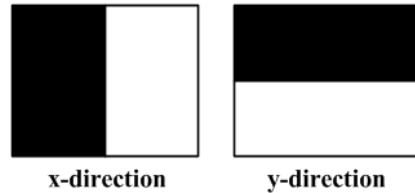


圖 7 Haar 小波濾波器

SURF 以特徵點為圓心， $6s$  為半徑的圓形鄰近區域內進行水平與垂直方向 Haar 小波響應 (Haar wavelet response) 之運算，其中  $s$  為特徵點可以被偵測到的尺度 ( $s$  is the scale at which the interest point was detected)，並在圓心角為  $\pi/3$  的扇形區域內繞著特徵點掃描，藉以統計扇形區域內所有特徵點的水平與垂直方向 Haar 小波響應值，並以扇形區域內最大加總值的掃描角度作為特徵點的主要方向，其做法如圖 8 所示：

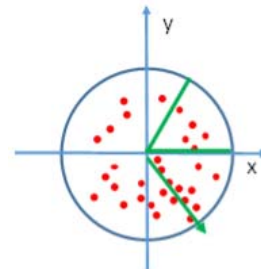


圖 8 SURF 特徵點方向計算

SURF 特徵向量是根據計算區域 Haar 小波響應值來進行描述，其方式是針對 Haar 小波響應值進行  $\sum dx$ 、 $\sum |dx|$ 、 $\sum dy$ 、 $\sum |dy|$  統計，藉以形成特徵向量，如(3)式所示。由於總共有  $4 \times 4$  個子區域，因此 SURF 特徵描述向量總共有  $4 \times 4 \times 4 = 64$  維，其特徵描述如圖 9 所示。

$$V = [\sum dx, \sum |dx|, \sum dy, \sum |dy|] \quad (3)$$

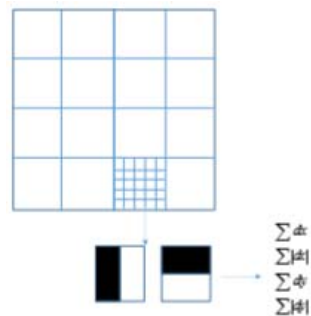


圖 9 SURF 特徵向量描述

## 2.2.3 ORB 特徵擷取

ORB 是改良 FAST 特徵點檢測和 BRIEF 特徵描述的方法，主要是針對 BRIEF 特徵描述演算法不具旋轉不變性及易受雜訊影響的問題做改良，ORB 演算法流程如圖 10 所示：

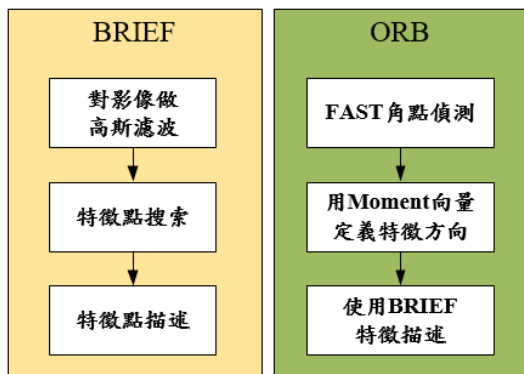


圖 10 ORB 演算法流程圖

BRIEF 特徵描述首先要對原始影像進行高斯濾波藉以降低雜訊干擾，然後在影像上選擇一個區塊 patch，用  $p$  來表示，其大小為  $S \times S$ ，隨後在  $p$  上提取 BRIEF 特徵。BRIEF 特徵描述  $\tau$  如式(4)所定義：

$$\tau(p; \mathbf{x}, \mathbf{y}) := \begin{cases} 1 & \text{if } p(\mathbf{x}) < p(\mathbf{y}) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

在(4)式中， $\mathbf{x}, \mathbf{y}$  是  $p$  內的兩個像素的位置向量， $p(\mathbf{x})$  和  $p(\mathbf{y})$  是像素位置  $\mathbf{x}$  和  $\mathbf{y}$  的亮度。一個 BRIEF 特徵是由若干個  $\tau$  所組成的二進位數值來描述。BRIEF 特徵首先對所求得的  $\mathbf{x}, \mathbf{y}$  對進行  $\tau$  測試，接著再用(5)式計算，即可得到 BRIEF 特徵描述。

$$f_{nd}(p) := \sum_{1 \leq i < nd} 2^{i-1} \tau(p; \mathbf{x}_i, \mathbf{y}_i) \quad (5)$$

從上式可看出要構造一個 512 個 bit 的 BRIEF 特徵描述，需要 512 對  $(\mathbf{x}, \mathbf{y})$ ，BRIEF 演算法的優點在於速度很快，但是它不具備旋轉不變性、尺度不變性及對雜訊敏感。

ORB 特徵擷取就是針對 BRIEF 的缺點進行改善。ORB 首先建立影像金字塔，並對影像金字塔的每一層進行 FAST 特徵點檢測，以使特徵點具有尺度不變性。接著再對特徵點進行 BRIEF 特徵描述，但所選的 patch 由 BRIEF 的  $31 \times 31$  改為  $5 \times 5$  的小區塊，藉以降低雜訊對特徵點的影響。同時，ORB 採用積分影像來加快特徵擷取的速度。為決定特徵點方向，ORB 採用矩向量(moment vector)來計算圓形重心的，並用特徵點與重心的夾角做為 FAST 特徵點的方向，其如式(6)~(8)所示：

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \quad (6)$$

$$c_x = \frac{M_{10}}{M_{00}} ; c_y = \frac{M_{01}}{M_{00}} \quad (7)$$

$$C_{ori} = \tan^{-1}\left(\frac{c_y}{c_x}\right) \quad (8)$$

(6)式中  $I(x, y)$  為輸入影像在點  $(x, y)$  的灰階值， $(c_x, c_y)$  分別為計算區塊質心的位置， $C_{ori}$  為特徵點的方向。

## 2.3 特徵點匹配與錯誤匹配對刪除

本文使用最佳相鄰點匹配法(Best of two nearest matching; BTNM)加上匹配對角度計算統計來刪除錯誤的匹配對，其流程如圖 11 所示：

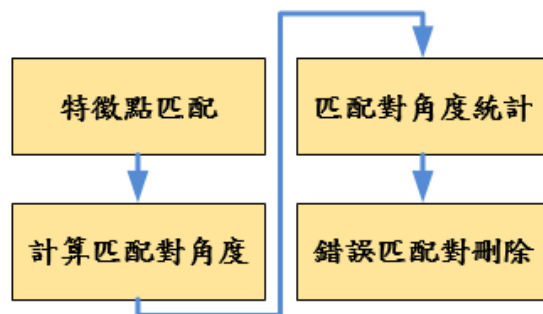


圖 11 特徵匹配與錯誤匹配對刪除流程圖

### 2.3.1 特徵點匹配

本文的匹配採用最佳相鄰點匹配法(Best of two nearest matching)，首先是用 RANSAC (Random Sample Consensus)找出一組內群(Inliers)影像間相容的單應性矩陣(Homography)，接著採用概率模型(Probabilistic Model)來驗證匹配。RANSAC 是使用一組隨機對應最小估計函數的影像變換參數，反覆提取  $n$  次最小點集估計該函數中參數的初始值，初始值將所有數據分為內群及離群(Outliers)，最後反過來用所有內群重新計算和估計函數的參數，如式(9)：

$$p(\text{H is correct}) = 1 - (1 - (p_i)^r)^n \quad (9)$$

式中  $p$  為正確的機率， $n$  為迭代的次數， $p_i$  為數據中內群出現之機率， $r$  為特徵對應點的數量。

接著使用概率模型來驗證匹配。對可能匹配的影像而言，會存在一組幾何形狀一致的特徵匹配，其稱為 RANSAC 內群，和一組位於重疊區域內但並不一致的特徵匹配，其稱為 RANSAC 離群，模型驗證的方法是比較內群/離群的機率。特徵的總數為在重疊的區域  $n_f$  中內群  $n_i$  的數目，影像匹配正確/錯誤由二進位變量來表示，如  $m \in \{0, 1\}$ ，第  $i$  個特徵匹配  $f^{(i)} \in \{0, 1\}$  是內群/離群值，被假設為獨立的



Bernoulli，以至於內群的總數為 Binomial 分佈，如式(10)-(11)所示：

$$p(f^{(ln_f)} | m=1) = B(n_i; n_f, p_1) \quad (10)$$

$$p(f^{(ln_f)} | m=0) = B(n_i; n_f, p_0) \quad (11)$$

其中  $p_0$  是一個特徵在正確的影像匹配中是內群的概率， $p_1$  是一個特徵在不正確的影像匹配中是內群的概率，設定特徵匹配的變數  $\{f^{(i)}, i=1, 2, \dots, n_f\}$ ，其表示為  $f^{(ln_f)}$ ，內群的數量  $n_i = \sum_{i=1}^{n_f} f^{(i)}$  和  $B(\cdot)$  是 Binomial 分佈，如式(12)所示：

$$B(x; n, p) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \quad (12)$$

接著再用貝氏準則(Bayes rule)計算影像匹配是正確的後驗概率(Posterior probability)，如式(13)-(14)所示：

$$p(m=1 | f^{(ln_f)}) = \frac{p(f^{(ln_f)} | m=1)p(m=1)}{p(f^{(ln_f)})} \quad (13)$$

$$= \frac{1}{\frac{p(f^{(ln_f)} | m=0)p(m=0)}{p(f^{(ln_f)} | m=1)p(m=1)}} \quad (14)$$

如果  $p(m=1 | f^{(ln_f)}) > p_{\min}$ ，則接受該影像匹配，如式(15)-(16)所示：

$$\frac{B(n_i; n_f, p_1)p(m=1) > \text{accept}}{B(n_i; n_f, p_0)p(m=0) < \text{reject}} \frac{1}{p_{\min} - 1} \quad (15)$$

$$n_i > \alpha + \beta n_f \quad (16)$$

其中正確的影像匹配其  $\alpha=8.0$  與  $\beta=0.3$ 。

### 2.3.2 匹配角度計算及統計

假設  $(x_1, y_1)$  與  $(x_2, y_2)$  為特徵匹配對，首先計算該匹配對的斜率，其如式(17)所示：

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad (17)$$

在式(17)中， $(x_1, y_1)$  為第一張影像的匹配點，而  $(x_2, y_2)$  為第二張影像的匹配點，接著對  $m$  值做  $\tan^{-1}$  轉換，得到的  $m$  值為範圍是  $0^\circ \sim 360^\circ$  的角度，然後對所有的角度進行角度直方圖統計。在本文中，以  $5^\circ$  或  $10^\circ$  為一個直方圖統計區間。

### 2.3.3 錯誤匹配對刪除(Error matching pair elimination; EMPE)

當取得角度直方圖後，本文將特徵點角度落在最大區間的匹配對設為正確的匹配對，其餘的匹配對則判定為錯誤，並將之刪除，其中角度直方圖如圖 12 所示。在圖 12 中，橫軸表示角度區間，例如數值 0 表示特徵點角度落於區間  $0^\circ \leq \theta < 5^\circ$ ，數值 1 表示  $5^\circ \leq \theta < 10^\circ$ ，其餘依此類推。

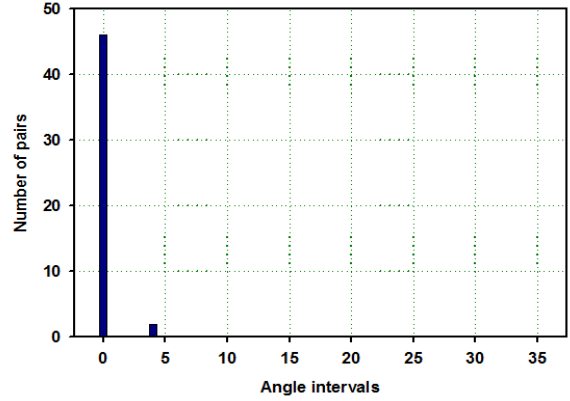


圖 12 角度直方圖統計

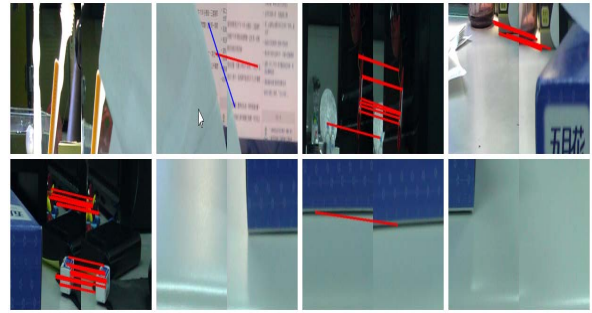


圖 13 分區影像特徵匹配結果

## 3. 結果與討論

本實驗為了測試不同特徵擷取方法與影像分區及不分區之差異，總共採用 2 張測試圖片進行 6 個不同測試，其結果如表 1 與表 2 所示，其中表 1 與表 2 分別採用圖 14 與圖 15 做為測試圖片。在表 1 與表 2 中，**SIFT** 方法表示採用 DoG 特徵點檢測與 SIFT 描述子(SIFT descriptor)；**SURF** 方法採用 Fast-Hessian 特徵點檢測與 SURF 描述子(SURF descriptor)；**ORB** 方法採用 FAST 特徵點檢測與 BRIEF 描述子(BRIEF descriptor)。至於特徵點匹配，表 1 與表 2 均採用最佳相鄰點匹配法(Best of two nearest matching; **BTNM**)。

在表 1 中，首先比較分區與整張(不分區)影像在計算時間效能上之差異，其中計算時間包含特徵點檢測、特徵向量描述與特徵點匹配。由於本文所提錯誤匹配對刪除法(**EMPE**)與前述特徵點檢測、特徵向量描述與特徵點匹配所費時間相比均低於 1 個 order，因此可忽略不計。由表 1 可知，當採用本文所提分區影像法來進行特徵檢測、特徵描述與特徵匹配時，在時間效能上約比整張(不分區)影像節省約 0.5~1 倍的時間。更進一步地，在不採用本文所

提錯誤匹配對刪除(EPME)的情形下,分區影像造成的錯誤匹配對相較於整張影像來得高。但在採用本文所提 EPME 的條件下,不管分區影像或整張影像均可將錯誤匹配對降低至零的水平。表 1 的結果,在表 2 中也呈現相同的趨勢。由此可知,本文所提分區影像與錯誤匹配對刪除法具有提升運算效能與增進特徵點匹配正確率之結果。

表 1 不同特徵點檢測方法與影像分區/不分區特徵匹配之結果(使用圖 2 之圖片)

		時間(S)	錯誤對/ 不使用 EPME	錯誤對/ 使用 EPME
整張影像	SIFT+BTNM	4.169	2	0
	SURF+BTNM	4.858	3	0
	ORB+BTNM	5.269	1	0
分區影像	SIFT+BTNM	2.513	2	0
	SURF+BTNM	3.134	3	0
	ORB+BTNM	2.953	3	0

表 2 不同特徵點檢測方法與影像不分區/分區特徵匹配之結果(使用圖 13 之圖片)

		時間(S)	錯誤對/ 不使用 EPME	錯誤對/ 使用 EPME
整張影像	SIFT+BTNM	9.723	3	0
	SURF+BTNM	6.244	2	0
	ORB+BTNM	9.979	2	0
分區影像	SIFT+BTNM	6.027	2	0
	SURF+BTNM	3.857	2	0
	ORB+BTNM	4.759	1	0



圖 14 表 1 特徵匹配所使用之影像

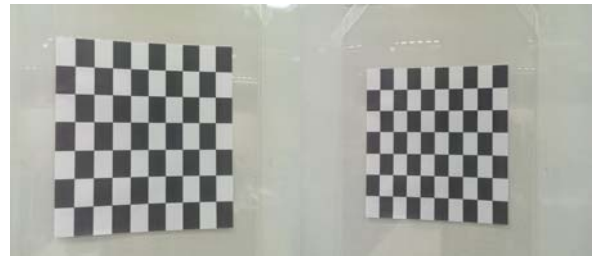


圖 15 表 2 特徵匹配所使用之影像

#### 4. 結論

特徵點檢測與匹配容易受影像的縮放、旋轉與雜訊影響,本實驗採用 SIFT、SURF 及 ORB 三種特徵檢測與描述方法均具有抗旋轉性、尺度不變性及抗雜訊的能力,搭配本文提出的影像分區特徵點計算與錯誤匹配對刪除法可有效的提升運算速度,同時降低特徵點匹配的錯誤率。

實驗結果顯示:本實驗所使用的影像特徵點分區計算方法,將影像劃分成 25 個等分,本文僅取其中編號 1-8 的分區影像進行特徵點計算,在時間效能上約比整張(不分區)影像節省約 0.5~1.0 倍的時間;當引入錯誤匹配對刪除法可使錯誤的特徵匹配對下降至 0,由此證明本文所提方法的有效性。

#### 誌謝

本論文承國科會 NSC-102-2221-E-212-015-經費補助,在此向國科會表達誠摯感謝之意。

#### 參考文獻

- [1] D. G. Lowe, "Distinctive image features form scale-invariant keypoints," *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110, January 2004.
- [2] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features", *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346-359, 2008.
- [3] E. Rosten, and T. Drummond, "Machine learning for high-speed corner detection", in: *Proceedings of the 9th European Conference on Computer Vision*, Graz, Austria, Part I, pp. 430-443, May 2006.
- [4] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: Binary robust independent elementary features," in: *Proceedings of the 11th European Conference on Computer Vision*, Crete, Greece, Part IV, pp. 778-792, September 2010.
- [5] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," in: *Proceedings of IEEE International Conference on Computer Vision*, Barcelona, Spain, pp. 2564-2571, November 2011.
- [6] M. Brown, and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision*, Vol. 74, No. 1, pp. 59-73, August 2007.