

基於疊代霍夫轉換與反霍夫轉換的特徵點對應技術

陳心怡¹、林彥宇²、陳炳宇¹

¹ 國立台灣大學 ² 中央研究院資訊科技創新研究中心

Email: fensi@cmlab.csie.ntu.edu.tw; yulin@citi.sinica.edu.tw; robin@ntu.edu.tw

摘要

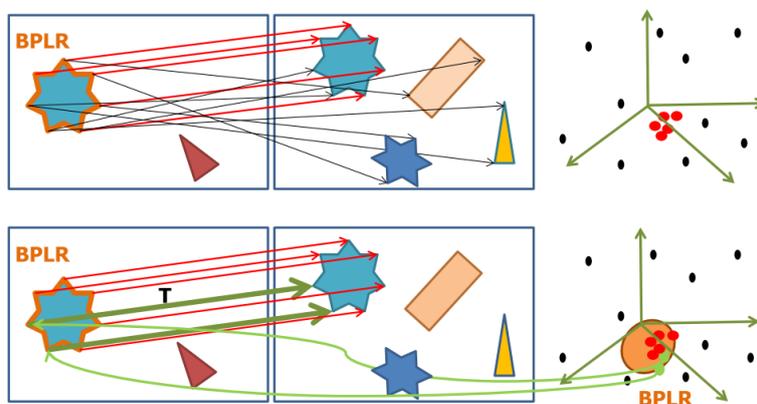
本論文旨在發展一個特徵點對應演算法，它可以同時提升對應之正確率與召回率。我們觀察到，在同一物件上鄰近的特徵點，在正確對應時，其對應轉換(如角度旋轉、大小縮放、位移)通常是十分相仿的；基於這項性質，我們將特徵點對應之問題，以霍夫空間(Hough space)中的密度估測來實現。具體來說，我們將每一個特徵點對應，依其對應轉換參數投影至霍夫空間，並在其中估測密度來判斷該特徵點對應的正確性。相關之特徵點對應將可依此方式進行交互幾何驗證，因而提高對應之正確率。另一方面，我們從區域成群特徵點(locally grouped features)推測一致性高的對應轉換，並依此動態地增加可能性高之特徵點對應，以便提高召回率。上述兩個過程彼此緊密結合，我們以迭代最佳化的方法連結霍夫轉換(Hough transform)與反霍夫轉換(inverted Hough transform)，逐步地擴增合理的對應轉換，進而檢測到更多正確的特徵點對應。最後在三組標竿(benchmark)資料集的實驗結果顯示我們提出的方法可以有效地得到特徵點對應。

關鍵詞：特徵點對應、霍夫轉換、反霍夫轉換

一、 介紹

研究目的

特徵點對應(feature correspondence)在電腦視覺領域中備受矚目，其目的在於識別兩張或多張影像之間像素的對應關係。這是一個十分基礎的影像內容分析技術，也是許多影像相關應用的基本元素，例如：目標辨識(object recognition) [1]、影像檢索(image retrieval) [2]、3D 重建(reconstruction) [3]、影像增強[4]與編輯[5], [6]。儘管具有高度之應用性，要建立高品質的特徵點對應仍然有至少兩個普遍存在的問題。第一，在特徵點匹配時的低正確性(low precision)。由於建立特徵點對應主要的方法是從區域特徵點開始產生可能性高之對應，在衡量相似度時，雖然區域描述子(local descriptors)的設計已經獲得顯著的進步，但描述子無法對物體在外觀或幾何上的變換維持不變性，且大量的背景雜訊更使得描述子無法成功分辨出正確的對應關係，所以這類方法的正確率會受到影響。第二，低召回率(low recall)的問題。許多改進的方法[7], [8], [9], [10]藉由確保幾何的一致性來提高正確率，但高階幾何檢查使得這些方法沒有很好地發展，另外，他們往往著重於如何在初始建構的特徵點組合中找尋正確的對應，而沒有利用已經得到的配對特徵點，去找出其他的潛在正確對應，因而限制住最終正確對應的數量(recall)。



[圖一](上)我們將特徵點對應投影至轉換空間，並根據集中程度分辨正確(紅)與錯誤(黑)的對應。
(下)潛在的(綠)對應可以透過探索由 BPLRs [11]組成之轉換的密度分佈逐步推斷而得。

本論文之內容主要參考[35]，在其中為了可以同時解決上述兩個問題，我們根據以下的觀察開發一個新的演算法。我們發現到，同一物件即使在外觀有許多不同的變異，但同一物件中相鄰的特徵點在對應時，必具有相仿之對應轉換(transformation or homography)，而這些正確對應組合(feature correspondence)的對應轉換往往會在其轉換空間(transformation space)中聚集在一起；相反的，錯誤的對應組合的對應轉換因不具有有一致性，因此在轉換空間中呈現零散分佈。這意味著，每一對應組

合的正確性可以經由其在轉換空間的密度來驗證，基於這個特性，我們將判斷對應正確與否的問題，轉化為一在轉換空間中密度測量(density estimation) 的問題。具體而言，我們透過比較轉換空間的資料密度，來判定每一特徵點對應組合是否正確；另一方面，透過探索轉換空間中高密度之處，我們可以動態地增加潛在的對應組合，參見圖一之說明。

文獻探討

特徵點對應之研究文獻相當廣泛，這裡我們僅針對與本論文較為相關的部分進行探討。

基於描述子(descriptor)建立特徵點對應之演算法

以描述子為基礎的特徵點匹配是處理特徵點對應這個問題的主要方式，一些值得注意的研究，像是[12], [13], [14], [15]，都為這個領域帶來了顯著的進步。雖然這些區域描述子(local descriptor)對一般形變有一定的穩健性，然而在現今視覺應用上，普遍的結論仍指出，沒有一種描述子足以用於處理由非剛性變形(nonrigid deformations)、亮度(illumination)與姿態(pose)改變等複雜的組合造成的變異。

透過圖形分割(graph partition)建立特徵點對應之演算法

一個以更多的幾何檢查來解決模糊配對的方法是將特徵點對應轉化為一個圖形匹配(graph matching)的問題。[16], [17], [18], [19]皆顯示，藉由定義一個基於光度相似(photometric similarity)和成對對應之間幾何相容性(geometric compatibility)的目標函數(objective function)，接著透過圖形匹配可以得到令人滿意結果。然而，這些方法通常在兩張影像只包含單一物件，簡單背景的拍攝條件下可以達成良好的匹配結果，若當兩張影像有多組共同物件出現時，這類方法是無法處理的。如同[20]這篇所提到的，此演算法容易受到錯誤的對應以及離群點(outliers)的影響。這個侷限性，使得圖形匹配演算法在許多現實的影像配對問題中無法被廣泛使用。

透過分群(clustering)建立特徵點對應之演算法

基於分群機制的研究已經用在處理無約束匹配(unconstrained matching)的情況。由下而上的群聚法(bottom-up clustering)可以整合局部適應約束(locally adaptive constraints)聚集一致的對應組合。ACC [21]透過聚合式階層分群法(hierarchical agglomerative clustering)，進行以物件為基礎之影像對應(object-based image clustering)。NRDC [4]採用一種粗略至精細(coarse-to-fine)的方式以及影像的連貫性來實現稠密特徵點對應(dense matching)。在另外一方面，CVP [20]提出一個由上而下的群聚法(top-down clustering)，在鄰近圖形(affinity graph)上偵測稠密的特徵點，並且在影像中找出多個共

同視覺樣板(common visual patterns)。儘管這類型演算法對於背景雜訊的穩健度高，但是仍然存在幾個缺點，最主要的就是此類演算法之時間複雜度高，較不適合在大量資料中進行影像特徵點配對，再來就是群集數量、群集合併的標準和相似條件(similarity thresholds)通常跟影像內容相關，很難找到一組群聚參數，可以適用於各種情況。

透過投票(voting)建立特徵點對應之演算法

RANSAC [22]是一種幾何驗證的模型，可以和區域描述子結合來增強其效能。[23]把每一組對應關係當成投票者(voter)，並且使用一個鄰近矩陣(affinity matrix)決定這些對應關係如何根據其相容性(compatibilities)給予其他對應關係相對的分數。就像 RANSAC，他們的方法都只適用於單一物件下之對應。[24]為解決多個物件的對應，提出一個霍夫轉換的變體，他們搭配金字塔匹配(pyramid match)[25]的方式，將對應組合排序。這個方法均勻量化轉換空間以達到快速配對，然而，正確的對應關係在經過轉換之後往往呈現不規則分佈，進而導致正確率下降。我們的方法也是一個基於投票的系統(voting-based system)，不一樣的地方在於每一個特徵點進行霍夫投票(Hough voting)的複雜度與對應組合的數量是互相獨立的；再者，其動態地增加可能之對應組合，以克服低召回率的問題。

對應組合擴增(correspondence enrichment)

大部分的特徵點對應演算法著重於從預先選定的小子集中找尋正確的對應，因此對應組合擴增成為一個相當重要的課題。匹配成長(match-growing)的方法[16], [26]以外觀為基礎，利用已知的對應關係擴增鄰近區域的像素(pixel)對應，但這類方法非常倚賴一開始建立對應關係的正確性。另外一方面，[27]發展一個區域成長(region-growing)的演算法來區分正確與不正確的對應。[28]提出一個漸進式的圖形匹配架構來擴增初始的對應點數目，然而在增加新的可能特徵點對應時，此方法會偏向選取密度較高的特徵點，而且結果可能會因為兩張影像的特徵點分布不同而顯得雜亂。相反的，我們的方法適用於由 BPLRs 指引的特徵束(feature bundles)，所以機率高的一致轉換會透過相關的特徵點轉移。事實證明，訊息可以被更有效率地傳播，且由此產生的候選對應組合(candidate correspondences)會更加有針對性。

二、 研究方法

本論文中，我們主要在於發展一個新的特徵點對應演算法，使其可以利用正確特徵組合在轉換空間中的分布一致性，提升特徵點比對之效能。以下分為兩部分進行說明，第一部分先定義我們要處理的問題，第二部分再詳細介紹我們所提出的方法。

定義特徵點對應

給定兩張影像 I^P 和 I^Q ，並個別提取一組特徵點 $V^P = \{v_i^P\}_{i=1}^{N^P}$ 和 $V^Q = \{v_i^Q\}_{i=1}^{N^Q}$ 。特徵點 $v_i \in V^P \cup V^Q$ 所在的區域和中心點分別以 S_i 和 \mathbf{x}_i 表示。 v_i 的外觀(appearance)用特徵向量 \mathbf{u}_i 描述，而其方向 θ_i 是根據在梯度直方圖(gradient histogram)中的一個主要方向估算而得[29]。積空間 $\mathcal{C} = V^P \times V^Q$ 代表所有可能的特徵點對應，而我們的目標則是要在 \mathcal{C} 中找到盡可能多的正確對應。

轉換空間

特徵點 v_i 的區域形狀和位置可以用一個 3×3 的矩陣 $T(v_i)$ 描述，其表示 v_i 的一個對於正規化修補(normalized patch)[29]的仿射變換(affine transform)：

$$T(v_i) = \begin{bmatrix} A(v_i) & \mathbf{x}_i \\ \mathbf{0}^\top & 1 \end{bmatrix} \quad (1)$$

其中， $A(v_i)$ 是一個 2×2 的非奇異矩陣(non-singular matrix)。

給定一對特徵點 $v_i^P \in V^P$ 和 $v_{i'}^Q \in V^Q$ ，從 v_i^P 到 $v_{i'}^Q$ 的相對變換(relative transformation) $H_{ii'}$ 可以衍生為

$$H_{ii'} = T(v_{i'}^Q) * T(v_i^P)^{-1} \quad (2)$$

在本論文中，我們以一個三件組(triplet) $m_{ii'} = (v_i^P, v_{i'}^Q, H_{ii'})$ 表示一個特徵點對應組合，即在相對影像上的兩個特徵點與其相對變換。如同 $H_{ii'}$ 是一個六自由度仿射變換(6-dof affine homography)， $m_{ii'}$ 可以視為在六維轉換空間中的一點。

轉換空間中的距離度量

給定兩個對應組合 $m_{ii'} = (v_i^P, v_{i'}^Q, H_{ii'})$ 和 $m_{jj'} = (v_j^P, v_{j'}^Q, H_{jj'})$ ， $m_{jj'}$ 相對於 $m_{ii'}$ 的投影誤差(projection error)可以定義為

$$d_{jj'|ii'} = \left\| \mathbf{x}_{j'}^Q - \rho \left(H_{ii'} \begin{bmatrix} \mathbf{x}_j^P \\ 1 \end{bmatrix} \right) \right\| \quad (3)$$

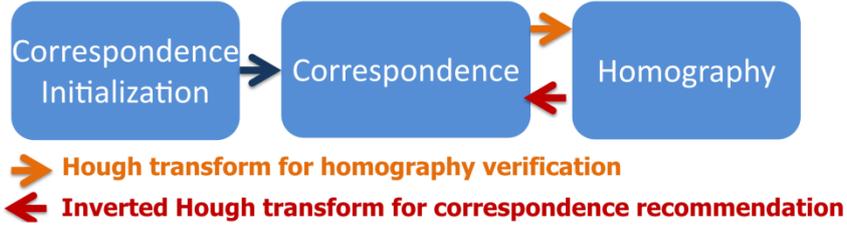
$$\text{where } \rho([a \ b \ c]^\top) = [a/c \ b/c]^\top \quad (4)$$

它檢查 $H_{ii'}$ 是否將 \mathbf{x}_j^P 投影到 $\mathbf{x}_{j'}^Q$ 周圍。

對於一對對應組合 $m_{ii'}$ 和 $m_{jj'}$ ，如果它們有相似的對應轉換(corresponding homography)，代表它們是相容的，因此我們採用重投影誤差(re-projection error)作為測量其相異程度的方式，即：

$$d(m_{ii'}, m_{jj'}) = \frac{1}{4} (d_{jj'|ii'} + d_{ii'|jj'} + d_{j'j|i'i} + d_{i'i|j'j'}) \quad (5)$$

需要注意的是，它是對稱的並且被用來計算在轉換空間中不同對應組合之間的距離。



[圖二] 迭代霍夫與反霍夫轉換之流程圖

迭代霍夫與反霍夫轉換

具有相容幾何組成(geometric configuration)的特徵點在匹配時是相互依存的。我們透過 BPLR 檢測器(detector) [11]偵測特徵依存度(feature dependence)，並將特徵點對應問題轉化為一密度測量問題。本論文以迭代霍夫與反霍夫投票的方法來實現上述概念，前者藉由將對應組合投影至轉換空間檢測一致的轉換；後者根據一致的轉換逐步增加可能的對應組合，整個方法的步驟如圖二所示。以下，我們首先說明如何建構初始對應關係，接著再分別介紹用於特徵點對應的霍夫與反霍夫轉換。

初始候選特徵配對

我們的方法從建構初始的候選特徵配對開始。對於每一特徵點 $v_i^P \in I^P$ ，我們根據外觀的相似程度，在 I^Q 中找尋 r 個可能的配對 $\{v_{i_k}^Q\}_{k=1}^r$ ，並且限制這 r 個配對不能有高度重疊。要達成這個目的，可以將所有可能配對的相似度做分數排序後，依序把那些與其前面有強重疊的(strong overlapping)選擇去除。具體而言，我們利用 SIFT 描述子估算兩個特徵點之間的相似度，即 $\|u_i^P - u_{i_k}^Q\|$ ，而其重疊程度則是由兩者區域的交集除以兩者區域的聯集估測而得，即 $S_i^P \cap S_{i_k}^Q / S_i^P \cup S_{i_k}^Q$ 。與 v_i^P 關聯的初始對應集合為

$$\mathcal{M}_i = \left\{ m_{ii_k} = (v_i^P, v_{i_k}^Q, H_{ii_k}) \right\}_{k=1}^r \quad (6)$$

其中， H_{ii_k} 是從 v_i^P 到 $v_{i_k}^Q$ 的相對轉換。對 I^P 中每一個特徵點重複上述步驟之後，以

$$\mathcal{M} = \bigcup_{i=1}^{N^P} \mathcal{M}_i \quad (7)$$

建構初始對應關係的集合。

初始之集合 \mathcal{M} 的大小是 $|\mathcal{M}| = r \times N^P$ ，其中包含許多錯誤的配對，因為在每一個 \mathcal{M}_i 裡面只會存在一個正確的對應組合。在複雜的配對任務中，常常發生的情況是，在 \mathcal{C} 中正確的對應組合，只有一小部分會被列入 \mathcal{M} 中。根據經驗，我們在所有的實驗中設定 $r = 5$ ，因為當 r 大於 5 的時

候，對應的正確率會急速下降。



[圖三] 以我們的方法找到之特徵點對應。(a)輸入影像 I^P 和 I^Q ，加上 I^P 中所有的特徵點和一些 BPLRs (輪廓)。(b)霍夫投票以及它跟 SIFT 的比較。透過霍夫投票，在 \mathcal{M} 裡面 222 個正確的對應中辨認出 207 個。白色線條表示兩個方法都有檢測到的正確對應，紅色與青色的線條則分別代表只有霍夫投票和只有最近 SIFT 搜尋法有找到的正確對應。(c)反霍夫投票。它提供 147(= 369 - 222) 個正確的候選對應，並且再經過成功的霍夫投票之後，得到額外 130(= 337 - 207) 個正確的對應(綠色線條)。

霍夫轉換於轉換驗證

這個階段的目標是要檢測 \mathcal{M} 中正確的對應組合，而這個 \mathcal{M} 可以是初始的對應集合或是經過下一階段擴增的對應集合。我們發現到，正確的對應會有一致的轉換，而那些不正確的對應則會有它們各自不同的轉換，基於這個特性，在這裡採用霍夫投票進行轉換驗證，因為它可以處理具有高比例不正確對應的情況，並透過密度測量來檢測正確的對應。具體來說，每一對應組合的相對轉換都被視為在霍夫空間(Hough space)中的一點，且被認為是一個關於注意潛在轉換(underlying homography of interest)的假設。

儘管霍夫轉換有其穩健性，但它被開發在一個假設上，即假設空間是個獨立投票之總和，從而忽視特徵點中的空間相依。正如[30]所指出，選擇適當的投票者(voter)是霍夫轉換中的關鍵，特別是當投票者相互有關聯的時候。我們觀察到在同一物件中鄰近之特徵點對應會彼此相關，並透過可以注意到物件邊界以及擷取物件局部形狀的 BPLR 檢測器將相關的對應集合起來。事實證明，霍夫投票的效能有顯著地提高。此外，在密度測量時只使用相關且小型的對應組合，而不是整個 \mathcal{M} ，如此一來便可加速投票過程。

將上述過程公式化，以 $\mathcal{B} = \{b_\ell\}$ 代表在影像 I^P 中偵測到的 BPLRs 集合。對於每一特徵點 $v_i^P \in I^P$ ，我們用 $\pi(v_i^P) \subseteq \mathcal{B}$ 表示覆蓋 v_i^P 中心點的 BPLRs 集合。對於沒有被任何 BPLR 覆蓋到的 v_i^P ，即 $\pi(v_i^P) = \emptyset$ ，我們就簡單地將其分配至最近的 BPLR，不過因為 BPLRs 是密集取樣的，所以這樣的情況並不常發生。接著我們將 v_i^P 相關的特徵點做分群，而分群的依據是藉由檢查它們是否共同處於至少一個 BPLR 中，即：

$$G(v_i^P) = \{v_j^P | \pi(v_i^P) \cap \pi(v_j^P) \neq \emptyset\} \quad (8)$$

我們假設成群機率高的特徵點在匹配時會有相似的轉換，因此在霍夫投票中，與 v_i^P 相關的對應組合可以經由下列的式子被集結在一起：

$$R(v_i^P) = \cup_{v_j^P \in G(v_i^P)} \mathcal{M}_j \quad (9)$$

根據公式 $\mathcal{M}_i = \left\{ m_{ii_k} = \left(v_i^P, v_{i_k}^Q, H_{ii_k} \right) \right\}_{k=1}^r$ (6)，在 \mathcal{M}_i 中最多存在一個正確的對應，霍夫投票以及投票者 $R(v_i^P)$ 就是用來選擇與特徵點 v_i^P 相關最合理的對應。具體來說，我們是採用 KDE (kernel density estimation) 之技術來實現霍夫空間中的投票機制，即密度測量：

$$m_{ii'}^* = \arg \max_{m_{ii'} \in \mathcal{M}_i} \frac{1}{|R(v_i^P)|} \sum_{m \in R(v_i^P)} \exp\left(-\frac{d(m_{ii'}, m)}{\sigma}\right) \quad (10)$$

其中， σ 是一個正數，其值為現有對應中距離的平均。需要注意的是，正規化項目 $1/|R(v_i^P)|$ 並不會影響公式 $m_{ii'}^* = \arg \max_{m_{ii'} \in \mathcal{M}_i} \frac{1}{|R(v_i^P)|} \sum_{m \in R(v_i^P)} \exp\left(-\frac{d(m_{ii'}, m)}{\sigma}\right)$ (10) 的結果，但是在特徵點之間比較密度時，這個項目是必須存在的。

選擇合理的對應組合的這個步驟會在影像 I^P 中每一個特徵點上重複執行，最後產生 N^P 個選定的對應組合 $\mathcal{M}^* = \{m_{ii'}^*\}_{i=1}^{N^P}$ 。我們再根據公式 $m_{ii'}^* = \arg \max_{m_{ii'} \in \mathcal{M}_i} \frac{1}{|R(v_i^P)|} \sum_{m \in R(v_i^P)} \exp\left(-\frac{d(m_{ii'}, m)}{\sigma}\right)$

(10)，將它們的相關密度做排序，並留下順位在一個適當門檻之前的那些對應組合。圖三(b)為此步驟之結果範例。

反霍夫轉換於對應組合推薦

在霍夫轉換辨認出正確的對應組合 $\mathcal{M}^* \subseteq \mathcal{M}$ ，並且提升配對的正確率之後，反霍夫轉換的目標則是要擴增 \mathcal{M} 以提高召回率。經 BPLRs 做過區域分群的特徵點，有著一致的轉換，而且可以在找合理的對應時互相協助。我們探討這個性質並開發反霍夫轉換，使成群的點將它們的轉換傳播給彼此，並透過探索這些傳播的轉換來提供每一特徵點一致的對應組合。

對於每一特徵點 $v_i^P \in I^P$ ，我們用公式 $G(v_i^P) = \{v_j^P | \pi(v_i^P) \cap \pi(v_j^P) \neq \emptyset\}$ (8) 的 $G(v_i^P)$ 搜尋其相關的特徵點，在這些特徵點中，每一特徵點都會提供一個關於 v_i^P 之轉換的假設，而這些假設會被集結在

$$\tilde{\mathcal{M}}_i = \{m_{jj'}^* | v_j^P \in G(v_i^P)\} \quad (11)$$

其中 $m_{jj'}^*$ 是 v_j^P 經過霍夫投票後所選擇的對應關係。在 $\tilde{\mathcal{M}}_i$ 這個集合當中，可能會包含由錯誤的配對造成的離群點，因此我們挑選大部分合理對應 $\tilde{m}_{jj'} \in \tilde{\mathcal{M}}_i$ 的轉換作為建議，其中

$$\tilde{m}_{jj'} = \arg \max_{m_{jj'}^* \in \tilde{\mathcal{M}}_i} \sum_{m \in \tilde{\mathcal{M}}_i} \exp\left(-\frac{d(m_{jj'}^*, m)}{\sigma}\right) \quad (12)$$

假設， $\tilde{m}_{jj'}$ 相對的轉換是 $H_{jj'}$ ，而 v_j^P 透過 $H_{jj'}$ 從 I^P 投影到 I^Q 的區域用 S 表示，那麼在 I^Q 中配對的特徵點則是由下列方式決定：

$$v_k^Q = \arg \max_{v_k^Q \in V^Q} \frac{S \cap S_k^Q}{S \cup S_k^Q} \quad (13)$$

接著， $m_{ik} = (v_i^P, v_k^Q, H_{ik})$ 會被建議成為可能的對應關係，即 $\mathcal{M}_i \leftarrow \mathcal{M}_i \cup m_{ik}$ ，對於 I^P 中每一特徵點都會進行這個步驟，而最後在公式 $\mathcal{M} = \bigcup_{i=1}^{N^P} \mathcal{M}_i$ (7) 得到的對應集合 \mathcal{M} 就會逐步擴增。反霍夫轉換的範例展示於圖三(c)。

霍夫轉換與反霍夫轉換的功能是相輔相成的，前者於擴增的候選對應中稠密檢測正確的對應；後者由於有較好的檢測結果可以提供調和擴增(harmonic enrichment)。我們可以保證迭代投票這個程序的收斂性，因為對應組合(\mathcal{C})的總數，是有限且固定的，在每一次的迭代，候選對應(\mathcal{M})的數量會呈現嚴格遞增，再加上 \mathcal{M} 是 \mathcal{C} 的子集，所以迭代程序必定會收斂。從經驗上來看，整個過程在經過幾次迭代之後會很快地收斂，在我們所有的實驗中通常是 2 到 4 次。在本節的最後，將我們的方法總結在演算法 1。

演算法 1 特徵點對應演算法

- 1: **Input:** Feature sets V^P and V^Q ; Max iteration T
 - 2: **Output:** Matched correspondences \mathcal{M}^*
 $\mathcal{M}_{ii} = 1NP$ via $\mathcal{M}_i = m_{ik} = v_i^P, v_k^Q, H_{ik} = 1r$ (6)
 - 4: **while** $t < T$ **do**
 - 5: $\mathcal{M}^* \leftarrow \emptyset$
 - 6: **for all** $v_i^P \in V^P$ **do**
 - 7: Detect correspondence $m_{ii'}^* \in \mathcal{M}_i$ via $m_{ii'}^* =$
 $\arg \max_{m_{ii'} \in \mathcal{M}_i} \frac{1}{|R(v_i^P)|} \sum_{m \in R(v_i^P)} \exp\left(-\frac{d(m_{ii'}, m)}{\sigma}\right)$ (10);
 - 8: $\mathcal{M}^* \leftarrow \mathcal{M}^* \cup m_{ii'}^*$;
 - 9: **for all** $v_i^P \in V^P$ **do**
 v_k^Q via $v_k^Q = \arg \max_{v_k^Q \in V^Q} \frac{S \cap S_k^Q}{S \cup S_k^Q}$ (13);
 - 11: Construct $m_{ik} = (v_i^P, v_k^Q, H_{ik})$;
 - 12: $\mathcal{M}_i \leftarrow \mathcal{M}_i \cup m_{ik}$;
 - 13: Sort elements in \mathcal{M}^* with thresholding;
-

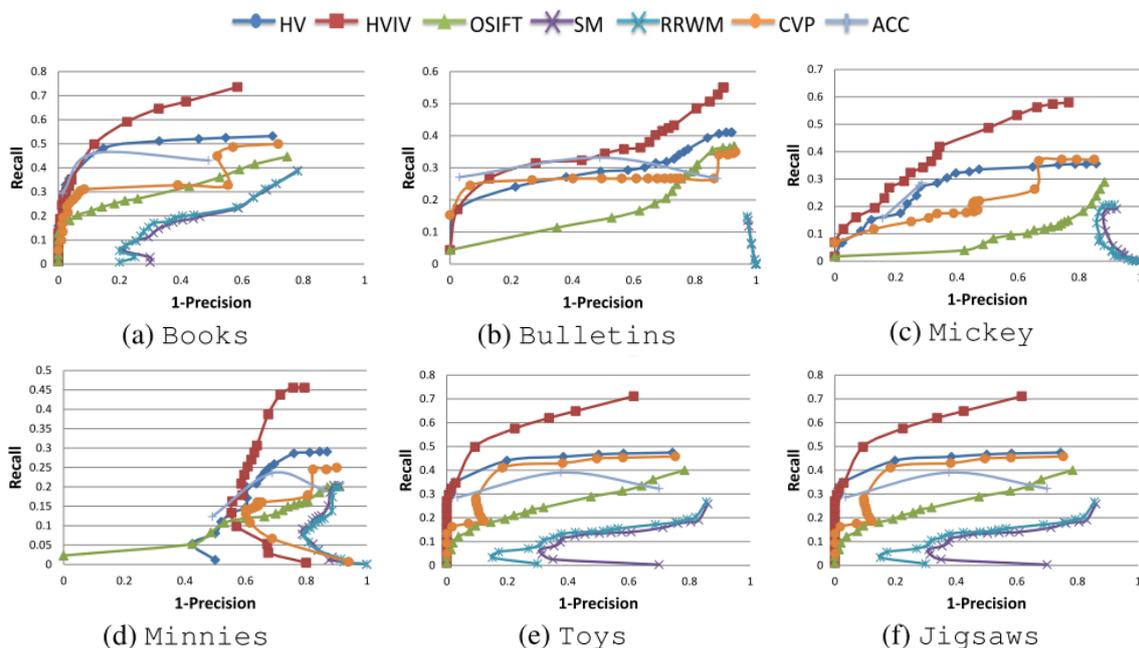
三、 結果與討論

我們進行三組實驗來評估本論文提出之方法表現。第一組實驗，針對辨認多個共同物件方面，我們的方法將與其他現有的技術做比較。第二組實驗，展示了我們的方法在逐步擴增候選對應的好處。第三組實驗，表明我們的方法可以和不同的檢測器與描述子一起使用，並在光照有強烈變化的影像上進行特徵點配對。

與多個共同物件進行配對

在這個實驗中使用的是 SNU 資料集[16]，其中包含 6 組影像對，影像中有數個共同物件、遮蔽現象以及複雜的背景雜訊，且物件出現的大小、方向與位置都有所不同，這些都是在處理特徵點對應時經常遭遇的挑戰。然而，它提供一個良好的測試平台，體現幾何驗證與對應推薦的重要性。

我們參考[31]，認為如果預測區域與真實區域之交集範圍除以兩區域之聯集範圍大於 40%，則代表這組對應是正確的。我們使用正確率與召回率評估特徵點對應之表現，其中正確率指的是在檢測到的對應中是正確對應之比例，而召回率是指在所有正確對應中有被檢測到的對應所占之比例，最後藉 precision-recall 曲線來呈現每種方法之結果表現。



[圖四] 在 SNU 資料集中每一組測試影像上之不同方法的比較。

將我們的方法 *Hough voting and inverted Hough voting* (HVIV) 與其他現有的方法做比較，包含基於描述子的方法：*opponent SIFT* (OSIFT) [32]；基於分群的方法：*common visual pattern discovery* (CVP) [20]、*agglomerative correspondence clustering* (ACC) [21]；基於圖形的方法：*spectral matching*

(SM) [18]、*reweighted random walks* (RRWM) [33]。此外，我們附加一個 baseline，*Hough voting* (HV)，這個 baseline 只執行我們的方法中霍夫轉換的部分。為求有公平的比較，全部的方法都使用相同的特徵點，每一點皆由 Hessian affine detector [29] 檢測到，並用 opponent SIFT descriptor [32] 來描述，所有方法的初始對應集合則是根據 opponent SIFT 的最近搜尋所選定。

量化的結果彙整於圖四中。基於圖形匹配的方法，例如：SM 和 RRWM，因為容易受到離群值的影響，所以表現並不穩定。在這個資料集中建構初始對應關係時，錯誤的對應常遠多於正確的對應。Baseline OSIFT 因為使用一元區域特徵點(unary local feature)，沒有足夠的能力處理雜亂的背景以及複雜的物體變形與轉換，所以沒有好的表現。相反的，CVP、ACC、HV 和 HVIV 藉由探討幾何的一致性來提高性能；其中，我們的 HVIV 在大部分情況下的表現又遠優於 CVP、ACC 和 HV，這是因為我們的方法將相關之特徵點分群用於對應驗證與推薦，它避免離群值的負面影響，解決過去基於初始候選特徵配對的限制，並因此達到高正確率與高召回率。各個方法的一些特徵點對應結果顯示於圖五，可以觀察到我們的方法找到更正確且密集的對應關係。

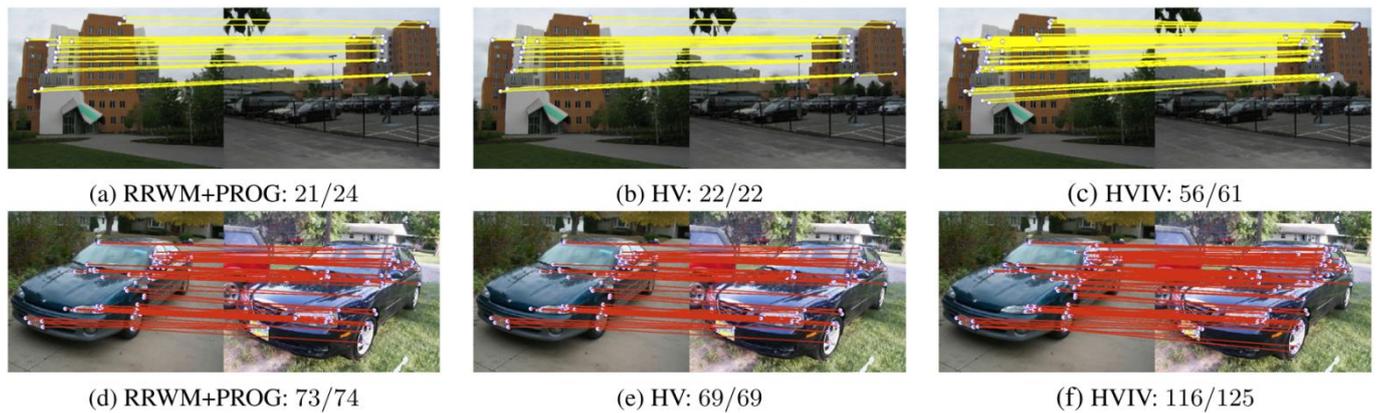


[圖五] 各種方法在一些影像對上的特徵點對應結果。(a)~(c) Books；(d)~(f) Bulletins；(g)~(i) Toys。每張圖皆有標示所使用的方法以及其表現(\mathcal{M} 中，正確檢測到的對應/正確的候選對應)。

對應組合逐步擴增

在這個實驗中，要展示我們的方法在對應組合逐步擴增的好處，並與 baseline *Hough voting* (HV)、*progressive graph matching framework* (PROG) [28] 比較，其中 PROG 是在擴增對應組合方面最好的方法之一。

這裡我們採用由[28]收集的標竿資料集來做效能評估，其中包含 30 組影像對，每一成對影像只包含一個物件。這個實驗使用的設定如同上述，包括使用相同的區域特徵檢測器、描述子及評估標準。PROG 為漸進式的圖形匹配提供一個通用機制，所以它可以跟任何一種基於圖形的方法一起使用，在這裡採用 RRWM 作為 PROG 之圖形匹配的模組(module)，關於 PROG 的參數 k_1 和 k_2 分別設定為 10 和 5。



[圖六] 使用 RRWM+PROG、HV 和我們的 HVIV 得到之特徵點對應結果。RRWM+PROG 和 HV 幾乎檢測到全部正確對應，然而我們的方法藉由逐步擴增對應組合進一步改善結果。

使用相同的檢測器、描述子、匹配標準和初始對應關係，圖六顯示分別用 RRWM+PROG、HV 和 HVIV 檢測到的正確對應。注意到，RRWM+PROG 得到的結果和原始論文的有所不同，這可能是因為對應結果容易受到採用的檢測器、描述子和匹配標準影響。

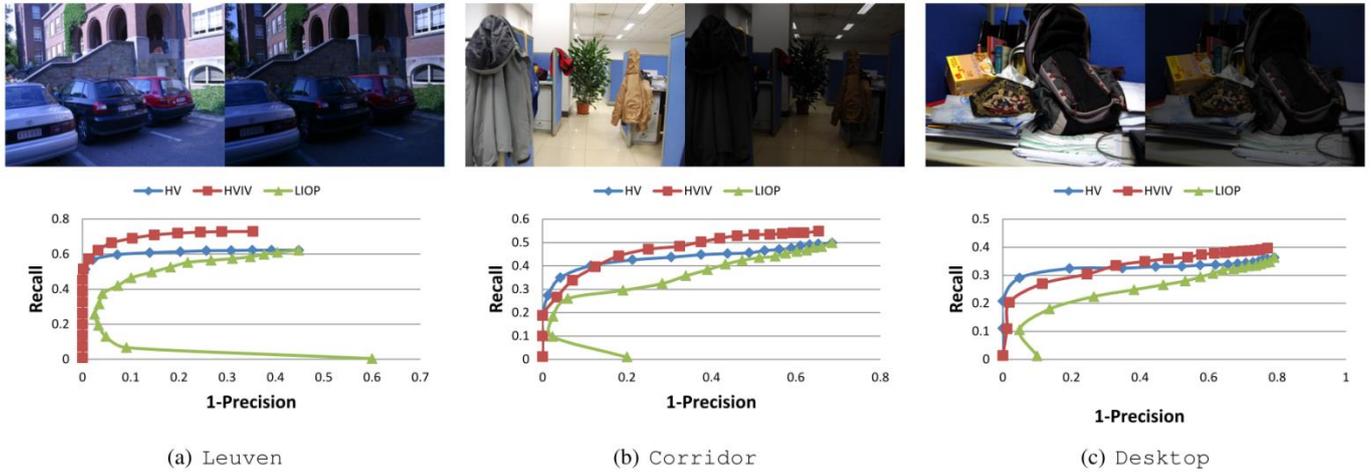
可以觀察到，HVIV 大幅增加真實候選配對的數量，而有更好的配對結果，這是因為在我們的方法中，對應關係驗證與推薦兩步驟的目標函數皆在密度上導出，有其一致性，這兩個步驟相輔相成，共同帶來更好的結果。除此之外，由 BPLRs 集結成群的特徵點，在霍夫與反霍夫轉換過程中，忠實地辨認相關的投票者(voter)，藉此將分散的投票者排除並提供高品質的建議。在相同的正確率下，我們的方法在真實的對應關係(true correspondences)中，跟 HV 比較，平均達到 54.0% 的成長率，而跟 PROG 比較也有 44.4% 的成長率。

用其他特徵描述子的 plug-in

我們的方法可以被視為一種幾何濾鏡(geometric filter)，它透過幾何檢查去除錯誤的對應組合，然後藉由在相關的特徵點之間傳播一致的轉換來提高配對，我們可以將它應用在不同類型的特徵描述子並提升效能。在這個實驗中，我們的方法與 LIOP 描述子[34]一起使用，在光照有強烈變化的影像之間建立對應關係。

我們在複雜的 illumination 資料集[34]上進行定量分析，它包含兩組影像對 Desktop 和

Corridor。另外，也採用在 Oxford 資料集中有曝光改變的 Leuven 影像對。對於每一組影像對，使用 Hessian-affine 檢測器定位特徵點位置，而初始候選配對則是利用具有光照不變性的 LIOP 來建造。



[圖七] Plug-in 與 LIOP 在三組影像對上之比較。

將 HV 和我們的 HVIV 應用在 LIOP 找到的對應組合上，成果以 precision-recall 曲線的形式呈現在圖七中。儘管 LIOP 已具有在光照變化下的穩健性，但它的表現仍然可以透過確認轉換的一致性以及擴增候選對應來增強。

四、 結論

在本論文中，我們提出一個簡單卻有效率的方法，迭代進行霍夫投票與反霍夫投票，以在複雜的匹配任務中建立特徵點對應，並提高正確率和召回率。這個方法將特徵點對應轉化為一個密度測量的問題，經由迭代最佳化，從擴增的候選對應中檢測出更多正確的對應組合，而在相關之特徵點間傳播一致的轉換，逐步擴增合理的對應組合。此外，我們透過 BPLRs 將相互依存的特徵點集結成群，這不僅提升準確率也加快整體速度。在三個資料集使用不同的描述子，對於提出的方法進行綜合評估，從結果可以看到其實用性。未來，將會把我們的方法應用到處理共同切割、影像合成及重建，因為高品質且密集的配對通常有利於實現這些應用。

參考文獻

- [1] A. C. Berg, T. L. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondence. In *IEEE CVPR*, 2005.
- [2] Z. Wu, Q. Ke, M. Isard, and J. Sun. Bundling features for large scale partial-duplicate web image search. In *IEEE CVPR*, 2009.
- [3] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3D. *ACM TOG*, 2006.
- [4] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski. Non-rigid dense correspondence with application for image enhancement. *ACM TOG*, 2011.
- [5] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM TOG*, 2009.
- [6] D. Mahajan, F.-C. Huang, W. Matusik, R. Ramamoorthi, and P. Belhumeur. Moving gradients: A path-based method for plausible image interpolation. *ACM TOG*, 2009.
- [7] A. Albarelli, E. Rodol`a, and A. Torsello. Imposing semi-local geometric constraints for accurate correspondences selection in structure from motion: A game-theoretic perspective. *IJCV*, 2012.
- [8] M. Leordeanu, M. Hebert, and R. Sukthankar. An integer projected fixed point method for graph matching and map inference. In *NIPS*, 2009.
- [9] L. Torresani, V. Kolmogorov, and C. Rother. Feature correspondence via graph matching: Models and global optimization. In *ECCV*, 2008.
- [10] F. Zhou and F. D. Torre. Factorized graph matching. In *IEEE CVPR*, 2012.
- [11] J. Kim and K. Grauman. Boundary preserving dense local regions. In *IEEE CVPR*, 2011.
- [12] H. Bay, T. Tuytelaars, and L. V. Gool. SURF: Speeded up robust features. In *ECCV*, 2006.
- [13] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE TPAMI*, 2002.
- [14] A. C. Berg and J. Malik. Geometric blur for template matching. In *IEEE CVPR*, 2001.
- [15] D. Lowe. Object recognition from local scale-invariant features. In *IEEE ICCV*, 1999.
- [16] M. Cho, Y. M. Shin, and K. M. Lee. Co-recognition of image pairs by data-driven monte carlo image exploration. In *ECCV*, 2008.
- [17] T. Cour, P. Srinivasan, and J. Shi. Balanced graph matching. In *NIPS*, 2006.
- [18] M. Leordeanu and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. In *IEEE ICCV*, 2005.
- [19] R. Zass and A. Shashua. Probabilistic graph and hypergraph matching. In *IEEE CVPR*, 2008.
- [20] H. Liu and S. Yan. Common visual pattern discovery via spatially coherent correspondences. In *IEEE CVPR*, 2010.
- [21] M. Cho, J. Lee, and K. M. Lee. Feature correspondence and deformable object matching via agglomerative correspondence clustering. In *IEEE ICCV*, 2009.
- [22] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *CACM*, 1981.
- [23] Y. Yuan, Y. Pang, K. Wang, and M. Shang. Efficient image matching using weighted voting. *PRL*,

2012.

- [24] G. Toliás and Y. Avrithis. Speeded-up, relaxed spatial matching. In *IEEE ICCV*, 2011.
- [25] K. Grauman and T. Darrell. The pyramid match kernel: Discriminative classification with sets of image features. In *IEEE ICCV*, 2005.
- [26] V. Ferrari, T. Tuytelaars, and L. van Gool. Simultaneous object recognition and segmentation by image exploration. In *ECCV*, 2004.
- [27] J. Cech, J. Matas, and M. Perdoch. Efficient sequential correspondence selection by cosegmentation. *IEEE PAMI*, 2010.
- [28] M. Cho and K. M. Lee. Progressive graph matching: Making a move of graphs via probabilistic voting. In *IEEE CVPR*, 2012.
- [29] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *IJCV*, 2004.
- [30] P. Yarlagadda, A. Monroy, and B. Ommer. Voting by grouping dependent parts. In *ECCV*, 2010.
- [31] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE TPAMI*, 2005.
- [32] K. van de Sande, T. Gevers, and C. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE TPAMI*, 2010.
- [33] M. Cho, J. Lee, and K. M. Lee. Reweighted random walks for graph matching. In *ECCV*, 2010.
- [34] Z. Wang, B. Fan, and F. Wu. Local intensity order pattern for feature description. In *IEEE ICCV*, 2011.
- [35] H.-Y. Chen, Y.-Y. Lin, and B.-Y. Chen, "Robust Feature Matching with Alternate Hough and Inverted Hough Transforms," In *IEEE CVPR*, 2013.