

國立政治大學
數位內容碩士學位學程

技術導論期中報告

信號和基於語意的圖像檔案和檢索技術
Signal and Semantic-Based Image Archives and
Retrieval Technologies

研究生：古健樺

指導教授：郭正佩

中華民國一〇六年十二月

第一章 研究動機與目的

網路以及資訊科技快速發展之下，人們逐漸習慣於網路的生活，網路已經是什麼主要分享多媒體資料的主要媒介之一，在眾多地多媒體資料之中，圖片在多媒體網際網路中，有著非常廣泛的運用，所以在網路之中有非常大量的圖片，若我們想要找尋到適合的圖片，必須搜尋如此大量的圖片，必定會花費大量的時間，所以提高搜尋的效率以及準確度為圖像檢索中的一大重要研究方向。目前不論是以文字為基礎或是以內容為基礎的圖像檢索系統，技術都仍未完善，於是本文專注在研究語意化分析 (Semantic Analysis) 的圖像搜尋 (Image Archives)，提高圖像檢索的效率以及準確度。

第二章 背景介紹與相關研究

本章節將介紹何謂資訊檢索和其細分之三種模式，以及圖像檢索之三種方式。

2.1 資訊檢索(Information Retrieval)

所謂的資訊檢索，就是從大量非結構的資料，例如網頁，根據某些關鍵字，找出具有此關鍵字的文件。例如，搜尋引擎，即是一種資訊檢索的應用，其中又可以分成三種檢索模式，分別為二元模式、向量空間模式、機率模式。

2.1.1 二元模式 (Boolean Model)

二元模式是一種以集合理論(set theory)以及布林代數(Boolean algebra)為基礎的簡易檢索模式[2]。其主要的概念，即考慮關鍵字集合(index terms)中的關鍵字(index term)是否存在於文件(documents)中；換句話說就是利用權重值來代表查詢關鍵字是否出現。

二元模式的優點在於呈現方式非常的簡單、明白，缺點則在於二元模式是以完全比對(exact matching)的方式進行資料比，而此方式往往導致檢索結果過多與過少的問題產生。

2.1.2 向量空間模式 (Vector Space Model)

有鑑於改良二元模式中關鍵字權重的指定方式過於限制(只考慮相關與非相關)。因此在向量空間模式中，對於查詢要求(q)與文件(dj)所組成的關鍵字集合，則是利用非二元(non-binary)值來表示。所使用的方法，是將使用者所下的查詢要求與系統文件分別取得其各自擁有的關鍵字權重，並利用所得到的權重值做相似程度(similarity degree)的計算，最後依照其相似程度的大小將檢索結果做排序(sorting)。而主要的概念是為了實現資料部分比對(partial matching)的方式來增加檢索精確度。

向量空間模式的優點在於:

1. 透過權重的組合來改善檢索料率。
2. 利用部分比對的方式，讓檢索所下的條件更為接近被查詢文件。
3. 利用餘弦定理(cosine formula)，將查詢結果依其相似程度的大小做排序的動作。

向量空間模式的缺點在於:

1. 關鍵字集合中的關鍵字，被假設為彼此互相獨立的。

2.1.3 機率模式(Probabilistic Model)

機率模式，於西元 1976 年由 Roberston 以及 Sparck Jones 所提出的檢索模式。其主要的概念，就是嘗試著利用機率的方式來解決資訊檢索的問題。由於此模式的最終目的，是藉由使用者所下的查詢條件，來找到其心目中的理想結果(idea answer set)，因此使用者若能在查詢的過程中，盡可能的具體說明目標屬性，透過詳細的說明，來取得檢索資料。此模式的基本運作流程如下:

1. 於接到使用者所下的查詢後，系統會假設一初始的機率值，此機率值是

用來描述理想的結果集合(idea answer set)。

2. 透過機率值的描述，找到與查詢要求相關(relevant)以及非相關 (non-relevant)的文件集合(document set)，並透過與使用者互動的機制以改善檢索精確度。

當查詢與文件之間較為相似時，代表文件與查詢要求之相關機率值大於非相關機率值。反之，當查詢與文件之間的相似度越小時，即代表文件與查詢要求的非相關機率值大於相關機率值，所以主要概念就是透過條件機率(conditional probability)，利用相關與非相關之機率值所得到的比例取得文件，來判斷文件與查詢之相似程度。

機率模式的優點在於：

1. 可以藉由相關程度的機率值，將查詢結果依其相似程度的大小做排序的動作。

機率模式的缺點在於：

1. 必須猜測一初始化的機率值，將相關與非相關的文件做分類。
2. 假設每一個關鍵字集合(index terms)皆為獨立。

2.2 圖像檢索(Image Retrieval)

在多媒體眾多分類中，屬圖像 (image)的應用最為廣泛，例如：教學、研究、娛樂...等。現在使用網路已較以往普遍，使用者重視的不只是資料存取的速度，更重視是否方便且易於使用，唯有兩者並重的服務才會得到使用者的青睞。

目前在圖像檢索上分為三種：以文字為基礎的圖像檢索 (text-base image retrieval)、以內容為基礎的圖像檢索 (content-base image retrieval) 和混合文字及內容的圖像檢索 (hybrid content and text base image retrieval)，本節將分別介紹這三種檢索技術的相關研究。

2.2.1 文字為基礎的圖像檢索

以文字為基礎的圖像檢索屬於高階查詢，簡稱 TBIR，也稱做 KBIR，現今已有許多提供圖像檢索服務的網站，例如:Google image、Picsearch、...等，使用的就是以文字為基礎的圖像檢索技術。透過文字對圖像的描述來檢索，而所描述的關鍵字將儲存於圖像的後設資料庫中 (metadata)。

2.2.2 內容為基礎的圖像檢索

以內容為基礎的圖像檢索，簡稱 CBIR，與 TBIR 相反，屬於低階查詢，就是利用圖像的各種特徵值 (feature)來辨識圖像，例如:顏色分佈 (color distribution)、紋路 (texture)及輪廓形狀 (shape)、...等。在前置作業，事先建立圖像庫，經過不斷的訓練和修正，才能建立查詢機制，除了前置作業需由人們控制及調整外，規則建置完成後，便可直接提供查詢。

2.2.3 混合文字與內容的圖像檢索

混合文字及內容的圖像檢索，先利用圖像低階特徵值自動分類 (CBIR)，再輔以人類輸入對圖像描述 (TBIR)，將兩種技術的優點加以融合，而發展出的新技術，可以讓使用者輸入圖像、輸入關鍵字來查詢圖像，彈性的查詢，成為目

前最熱門的圖像檢索技術。

第三章 語意化圖像搜尋

本章節將從本體論開始介紹，然後說明語意網與本體論的關係，接著說明電子詞彙資料庫 (Word Net)，最後說明語意化圖像搜尋之方法，。

3.1 本體論 (ontology)

本體論(Ontology)源自哲學用語，意義上是用來解釋描述這個世界的現象，並且對應到生活中的實體。爾後，人工智慧的研究者將本體論應用在描述特定領域的知識，所以本體論已逐漸變成為相當實用的知識管理工具，用來描述一個領域中的基本概念及定義彼此之間的關係，完善的正規化表達某領域的知識。我們可以透過本體論，定義領域內的物體類型、屬性與關係，這樣的方式便於不同領域之間的分享，而且能夠讓機器更容易了解使用者的需求。

3.2 語意網 (Semantic Web)

有人說：「網路就像散了一地沒有頁碼的書」，缺乏系統化與結構化的知識，端賴使用者的搜尋技巧，自行發覺其中的內容與關聯。目前網路上的搜尋依賴關鍵字的比對，電腦只做比對的動作，它無法更進一步瞭解文件的內容，造成搜尋的資料筆數很多，但是效率卻不高。因此，全球資訊網發明人 Tim Berners-Lee 在 2001 年 Scientific American 提出語意網 (Semantic Web) 的概念，在現有的網路架構下，提供一個共同的框架，讓資料可以跨越不同的系統與程式、打破企業的界線分享與再利用。簡單地說，語意網主要就 做兩件事情，第一個是整合不同來源的資料變成共同的格式，第二個是將資料和真實世界連接。目前的網路只能說是交換文件的平台，相關電腦系統無法了解文件的

內容。因此，要將不同的領域文件結合，除了利用共同的框架描述以外，尚須利用本體論定義領域的知識、類別與屬性的關係，因為定義清楚，才能夠分享、交流與再利用。

3.2 電子詞彙資料庫 (Word Net)

電子英文詞彙資料庫 WordNet 是由 Miller 於 1990 年所提出。WordNet 主要提供兩大類的語義資源，分別為「同義詞群的資料」以及「詞群與詞群之間的關聯性」。

3.2.1 同義詞群資料

何謂同義詞群呢？WordNet 將英文單字中詞義相同的詞群組成一個同義詞集合，如 {Human, Person, Someone, Somebody}，這樣的集合在 WordNet 中稱為同義詞群(synonym set; synsets)，也就是 WordNet 中的最小資料單位，因此我們可以將 WordNet 看成一個由許許多多同義詞群所組成的集合。

3.2.1 詞群與詞群之間的關聯性

在 WordNet 除了有同義詞群的關聯性外，還存在其他字群與字群之間的關係，其完整的關聯性，如表 1。在表 1 中每個欄位從左至右分別代表：「詞群之間的關係」、「關聯性的定義描述」、「詞性的對應型態」以及「範例」，因此藉由表 1，我們可以詳細的瞭解 WordNet 所提供的語意關聯性以及相對應的關係，主要的關聯性。

詞群關係	關係描述	對應詞群型態	例子
Synonym	A concept that means nearly the same as another.	{n,v,adj,adv} to {n,v,adj,adv}	{sea} is the synonym of {ocean}
Gloss	Gloss of synset.	{n,v,adj,adv}	
Hypernym	A concept that meaning denote superordinate.	{n,v} to {n,v}	{water} is the hypernym of {ocean}
Entailment	A concept that meaning denote inheritance.	{v} to {v}	{snore} is entails {sleep}
Similar	A concept that means a cluster head to the other concept.	{adj} to {adj}	{tiny,lesser} is the satellite of {small}
Member Meronym	A concept that is a member of another concept.	{n} to {n}	{people} is the member of {world}
Substance Meronym	A concept that is a substance of another concept.	{n} to {n}	{water} is the substance of {ice}
Part Meronym	A concept that is a part of another concept.	{n} to {n}	{arm} is the part of {body}
Cause of	A concept that is the cause of the other concept.	{v} to {v}	{show} is the cause of a result {see}

表 1 存在於 WordNet 中詞群與詞群之間的關係與範例

3.3 語意化圖像搜尋

為了讓檢索結果更為精確，有效的分析圖像註解資訊，顯的非常重要。因此在使用者完成圖像上傳與加註步驟後，系統便會將使用者對於圖像所做的描述先經過文字語意索引，取得足以描述圖像的關鍵字，然後將擷取出來的關鍵字透過 WSD 字義分析，並利用 WordNet Database 的語意關聯性，將取得的字義做關鍵字語意的延伸，最後將所得到的註解資訊存入語意資料庫中。其處理過程主要分為三個步驟，分別為：

1. 關鍵字擷取(Keyword Extraction)。
2. WSD 處理(Word Sense Disambiguation Processing)。

3. 關鍵字延伸(Keyword Expansion)。

第四章 語意化圖像搜尋應用

本章節將介紹數個應用語意化圖像搜尋的案例。

4.1 根據多張圖片搜尋

能夠根據多張圖片之間的關係，去取得圖像搜尋的結果，給使用者有更多的彈性。



(a) Query Images



(b) Retrieved Images

4.1 對圖片進行語義切割

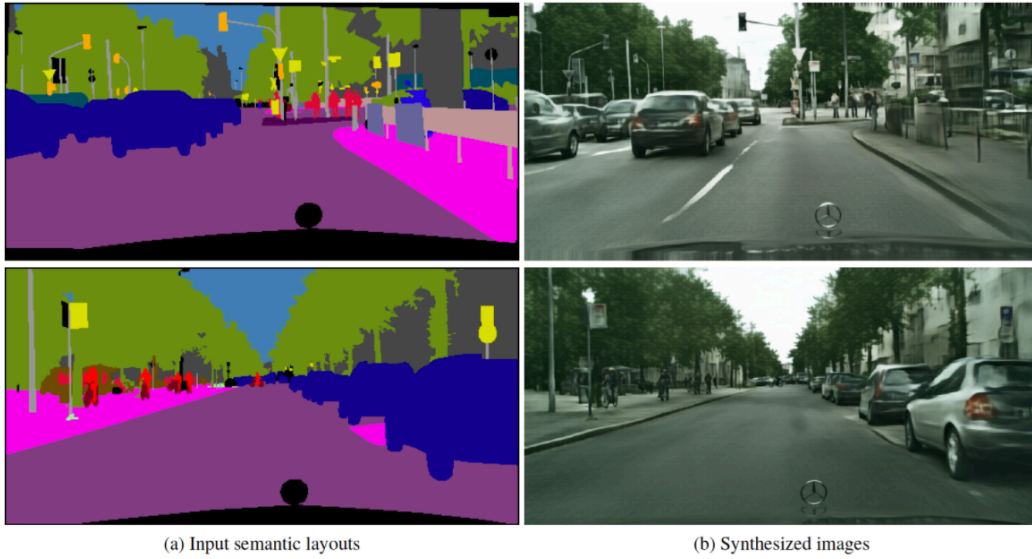


Figure 1. Given a pixelwise semantic layout, the presented model synthesizes an image that conforms to this layout. (a) Semantic layouts from the Cityscapes dataset of urban scenes; semantic classes are coded by color. (b) Images synthesized by our model for these layouts. The layouts shown here and throughout the paper are from the validation set and depict scenes from new cities that were never seen during training. Best viewed on the screen.

第五章 結論

由於圖像也是多媒體中重要的元素，而且相較於過去，網路頻寬逐漸加大，網路速度也開始飛快地提升，傳輸圖片已經是非常地快速了，在各類雲端的資料庫中充斥著大量的圖像，若是使用一般的文字搜尋圖片檔案名稱，抑或是根據人工建立 metadata，都需要花費比較高的成本來換取精準度，於是出現的語意化的搜尋方式，能夠更準確的判斷圖像中所包含的物品，更以貼近使用者的思維去搜尋，增加便利性，目前廣泛應用在許多需要精準辨識物品種類的裝置或平台中，目前語意搜尋已經算是發展中期的技術，未來將會更貼近生活中，目前深度學習的崛起，相信在語意網以及深度學習的交互碰撞之中，能夠讓圖像搜尋的效率以及準確度有更上一層的突破。

相關文獻

網路資料

[https://data-](https://data-sci.info/2017/08/08/%E5%88%A9%E7%94%A8%E5%B7%B2%E7%9F%A5%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%94%A2%E7%94%9F%E5%90%88%E6%88%90%E5%9C%96%E7%89%87-photographic-image-synthesis-cascaded-refinement-networks/)

[sci.info/2017/08/08/%E5%88%A9%E7%94%A8%E5%B7%B2%E7%9F%A5%E8%](https://data-sci.info/2017/08/08/%E5%88%A9%E7%94%A8%E5%B7%B2%E7%9F%A5%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%94%A2%E7%94%9F%E5%90%88%E6%88%90%E5%9C%96%E7%89%87-photographic-image-synthesis-cascaded-refinement-networks/)

[AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%94%A2%E7%94%9F%E5](https://data-sci.info/2017/08/08/%E5%88%A9%E7%94%A8%E5%B7%B2%E7%9F%A5%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%94%A2%E7%94%9F%E5%90%88%E6%88%90%E5%9C%96%E7%89%87-photographic-image-synthesis-cascaded-refinement-networks/)

[%90%88%E6%88%90%E5%9C%96%E7%89%87-photographic-image-synthesis-](https://data-sci.info/2017/08/08/%E5%88%A9%E7%94%A8%E5%B7%B2%E7%9F%A5%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%94%A2%E7%94%9F%E5%90%88%E6%88%90%E5%9C%96%E7%89%87-photographic-image-synthesis-cascaded-refinement-networks/)

[cascaded-refinement-networks/](https://data-sci.info/2017/08/08/%E5%88%A9%E7%94%A8%E5%B7%B2%E7%9F%A5%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%94%A2%E7%94%9F%E5%90%88%E6%88%90%E5%9C%96%E7%89%87-photographic-image-synthesis-cascaded-refinement-networks/)

[https://data-sci.info/2017/06/23/%E4%BE%86%E8%87%AAgoogle-](https://data-sci.info/2017/06/23/%E4%BE%86%E8%87%AAgoogle-research%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%9A%84%E6%9C%80%E6%96%B0%E9%80%B2%E5%B1%95%EF%BC%9Arethinking-atrous-convolution-semantic-image-segmentation/)

[research%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%9A%84%E](https://data-sci.info/2017/06/23/%E4%BE%86%E8%87%AAgoogle-research%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%9A%84%E6%9C%80%E6%96%B0%E9%80%B2%E5%B1%95%EF%BC%9Arethinking-atrous-convolution-semantic-image-segmentation/)

[6%9C%80%E6%96%B0%E9%80%B2%E5%B1%95%EF%BC%9Arethinking-](https://data-sci.info/2017/06/23/%E4%BE%86%E8%87%AAgoogle-research%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%9A%84%E6%9C%80%E6%96%B0%E9%80%B2%E5%B1%95%EF%BC%9Arethinking-atrous-convolution-semantic-image-segmentation/)

[atrous-convolution-semantic-image-segmentation/](https://data-sci.info/2017/06/23/%E4%BE%86%E8%87%AAgoogle-research%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E7%9A%84%E6%9C%80%E6%96%B0%E9%80%B2%E5%B1%95%EF%BC%9Arethinking-atrous-convolution-semantic-image-segmentation/)

[https://data-sci.info/2017/07/13/awesome-image-semantic-segmentation-](https://data-sci.info/2017/07/13/awesome-image-semantic-segmentation-%E5%9C%96%E5%83%8F%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E4%B9%8B%E5%B0%8E%E8%AE%80%E8%AB%96%E6%96%87%E8%88%87%E5%B7%A5%E5%85%B7%E5%8C%85%E7%B8%BD%E5%8C%AF/)

[%E5%9C%96%E5%83%8F%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%](https://data-sci.info/2017/07/13/awesome-image-semantic-segmentation-%E5%9C%96%E5%83%8F%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E4%B9%8B%E5%B0%8E%E8%AE%80%E8%AB%96%E6%96%87%E8%88%87%E5%B7%A5%E5%85%B7%E5%8C%85%E7%B8%BD%E5%8C%AF/)

[B2%E4%B9%8B%E5%B0%8E%E8%AE%80%E8%AB%96%E6%96%87%E8%88](https://data-sci.info/2017/07/13/awesome-image-semantic-segmentation-%E5%9C%96%E5%83%8F%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E4%B9%8B%E5%B0%8E%E8%AE%80%E8%AB%96%E6%96%87%E8%88%87%E5%B7%A5%E5%85%B7%E5%8C%85%E7%B8%BD%E5%8C%AF/)

[%87%E5%B7%A5%E5%85%B7%E5%8C%85%E7%B8%BD%E5%8C%AF/](https://data-sci.info/2017/07/13/awesome-image-semantic-segmentation-%E5%9C%96%E5%83%8F%E8%AA%9E%E7%BE%A9%E5%88%86%E5%89%B2%E4%B9%8B%E5%B0%8E%E8%AE%80%E8%AB%96%E6%96%87%E8%88%87%E5%B7%A5%E5%85%B7%E5%8C%85%E7%B8%BD%E5%8C%AF/)

論文資料

Semantic Image Search from Multiple Query Images, G Vaca-Castano, 2015

Semantic Image Search Based on Spatial Relationships, Shian-Jiun Meng, 2005

Semantic image search using queries

應用本體論與語意網技術於國民中學資訊 課程之研究, 羅秀瑜, 2012

以中文語意概念為基礎之文件圖像檢索系統, 李政均, 2006

建置語意式索引於圖像檢索系統, 藍永孝, 2003