



國立交通大學
National Chiao Tung University

出國報告（出國類別：**出國短期研究**）

題目：電腦視覺在影像辨識及識別上之
理論及應用

服務機關：**應用數學系**
姓名職稱：**張洛賓 助理教授**
前往國家：**美國**
出國期間：**101/06/13-101/9/14**
報告日期：**101/10/09**

一、摘要 (200-300 字)

近期與布朗大學美國國家科學院院士 **Stuart Geman** 教授有一個重要的研究計畫，這次移地研究三個月主要目的就是延續之前的理論研究並應用在影像辨識及影像識別上，人類辨識或識別影像的視覺功能遠遠超過電腦視覺，對人類而言許多輕而易舉甚至根本不需思考就能解答的問題(例如照片裡是否有摩托車、是否有美國歐巴馬總統、或是影片中是否有人在騎腳踏車等等)，對現今的電腦科技確似乎遙不可及，這類的問題我們稱為電腦視覺問題。

我們整個大型研究框架是貝氏影像生成模型。主要分成兩個部分，第一部分是影像隱藏涵義的先驗分布，這個部分主要目的要把各個影像裡的物體可能隱藏的訊息分析出來，包含物體間的相關性與其機率分析。第二部分是物體影像形成的影像資料模型，這個部分主要是連結數學模型與實際影像資料的介面。這個暑假最主要的研究工作是針對第一部分的數學模型局部化理論發展與計算，並且針對第二部分資料模型中的共享點格問題提出一個初步的解法。在接下來的章節中，將會有進一步對此次出國研究的進展報告。

二、目次

三、本文

(一) 目的.....	P.4
(二) 過程.....	P.4
(三) 心得及建議.....	P.6

三、本文

(一) 目的

我從 2007 年 9 月開始著手電腦視覺的研究，我的博士論文第二部分正是電腦視覺的理論分析與建模計算，在 2010 年與 2011 年間，延續博士班的研究，我們發明了敏感形隱式系統的數學分析，是一個全新的數學分支並且有很大的應用發展，尤其在電腦視覺問題中更是提供了一個可能突破的研究方向，正因為是一個創新的研究，我與我的博士班指導教授 Stuart Geman 決定繼續合作並朝著此方向發展，雖然電腦視覺問題在這 40 年來經過各個領域的研究學者努力(其中包含電腦資訊工程學家、應用數學家、理論數學家、統計學家、物理學家等等)始終都沒有辦法解決，我們依然深信我們的研究將會有更重要的貢獻與影響。上一年十月 Stuart Geman 教授才來台灣繼續我們的合作長達一個月，今年我也利用暑假與他們那邊的研究團隊繼續合作，很感謝布朗大學在這段期間提供研究設備與研究資源，這次出國短期研究最主要有兩個部分：

先驗模型的局部化理論分析:

我們知道建構一個合理的先驗機率模型，我們需要考慮的先驗知識或許遠遠超過手邊的影像資料，模型的複雜度就可想而知，利用多層次的建模方法可以將複雜的模型分層分析，但是各層間的敏感度、相關性也需要考慮，因此，當我們在進行局部最優分析時，就必須計算局部機率分布，但往往各層敏感度與相關性過於複雜導致計算上無法有效的進行，因此我們就必須先進行局部化的工作。

共享點格問題:

共享點格問題是一直困擾我們的一個問題，為了加速計算的速度與計算的效能，先計算粗略模組再計算精細模組已經成為一個重要的過程。粗略模組可以計算得非常快，所以可以很快排除非目標的情況，因此省下很多計算時間，通過粗略模組的測試後再去花時間做精細模組的計算。但問題是這兩模組之間的相關性是很難以分析的，因為兩個模組間會有共享點格的情況發生，共享點格的位置並非固定，而是隨機的，這使得建模上更為困難。所以這個暑假的目標之一，就是找出一個解決的方法。

(二) 過程

這次訪問布朗大學，第一個月剛好遇到 CVPR 的年度大會在普羅維登斯舉辦，因此很多做影像辨識的學者也順便訪問布朗大學，其中這屆 CVPR 主辦人 UCLA 的朱松純教授就跟我們有深入的學術交流，就他發展的紋理組織模型進行研討並分享雙方未來研究方向，我們發現他的模型實際上也可以從我們的建模架構中得出，紋理影像是我們將來必定會處理的一種影像，所以此次的研究

交流對於我們日後處理紋理影像應有一定的幫助。除了朱教授外還有從瑞士 ETH Zurich 大學來的 **Buhmann Joachim**，他是訊息理論的專家，他花了一個下午的時間與我們解說他的複雜度分析理論，藉由他的研究我們想要了解我們處理的計算問題是否在複雜度分析理論上是可以辦到的。

除了跟其他相關領域的訪問學者交流外，其餘時間都投注在與 **Stuart Geman** 教授研究團隊的研究合作，值得一提的是 **Pedro Felzenszwalb** 教授也加入了我們很多研究討論，他是電腦影像辨識的專家，並在 2008 與 2009 年拿下 PASCAL 電腦影像辨識比賽的世界冠軍(雖然世界目前最好的結果還遠遠不及於人類的視覺辨識系統)，他也帶給我們很多新的想法與實務經驗，希望未來能有機會與他合作。這次出國對於主要兩部分的研究階段性成果與進展簡述如下:

先驗模型的局部化理論分析:

理論上如果有整個先驗機率模型，局部先驗機率模型就可以藉由多重積分得到，但實際應用上，往往需要處理多重求和的大量計算，根本無法達成，因此，我們就必須做一些逼近解。這次在美國主要任務之一就是研究這個逼近問題，主要的逼近方法有兩種，而我們的研究必須都用到，第一種方法是模型的逼近，也就是數學上找出可計算的一個局部先驗機率模型但又非常接近真實的局部先驗機率模型，**Pedro Felzenszwalb** 教授當時建議我們可以考慮現在熱門的置信傳播計算法去做局部化的逼近計算，但我們實際操作過後發現在某些情況並非逼近的很理想，並且到目前為止還沒有理論可以證明置信傳播的收斂性，這部分可以成為未來研究的一個方向，目前為止我們所發展的逼近公式是馬世概念的逼近，主要局部化後只跟多層次模型架構裡的相對母層、相對子層與兄弟層有關，目前除了還沒有簡易的表達式外，應用上還沒發現其他問題，數學理論將是將來的一個重點。

第二種方法就是大多數機率統計學家常用的蒙地卡羅法，簡單的說就是取樣本點再用樣本點估計積分，用到的原理雖然只有大數法則，但是如果考慮收斂速度，那就要牽扯到複雜的重要樣本點取法。這個暑假，我們將在局部化過程中的第一步所得到的逼近模型，用來當作重要樣本點取法的樣本隨機模型，理論上收斂性是可以證明的，但實際應用的誤差估計還有待研究中。

共享點格問題:

相鄰畫素點格間的相關性是不可少的，但過度詮釋相關性卻衍生共享點格問題，我們這個暑假就是想要克服這個問題，我們嘗試把相鄰畫素點格間的相關性減弱至只依賴平均值，當然如此一來，先驗機率模型就必須有更審密的參數估計已彌補這個相關性的弱化，我們這個嘗試主要的目的和好處是使得共享點問題變得更容易處理，只需要在共享點的位置替換平均值就可以處理了，而且對於任何相對位置，都可以如此操作。目前我們正在努力研究並了解這個方法的極限在哪裡?有那些問題是沒辦法處理的，就在我要回國前兩周，我們發現這個方法需要對影像進行所謂的標準化過程，而且這個過程並不是所有的影像都相同，他是跟照相機或錄影設備的有關，標準化主要的目的是要除去光線強弱

對影像辨識模型的影響，目前為止我們已經推導出合理的方程式，接下來就是需要利用數值分析方法實際將函數解出來，並用來做標準化測試。

（三）心得及建議

電腦視覺最主要的困難是許多人類認為理所當然的事要實際上把他用數學式表示出來並不容易，再加上即使能用數學表示出來，並非就能實際運用，因為往往面面俱到的數學模型在計算上需要用到指數增長的電腦運算，意思就是計算量過於龐大以至於電腦無法完成，再加上傳統的數學建模方法仰賴大量的參數，導致無法有效地估計參數。所以數學模型的設計需要考慮計算的可行性並且要能夠提出一套有效的參數估計方法，這正是我這幾年來投注大量的精力與 Stuart Geman 教授要共同研究的目標。

在此我要感謝交通大學在學術上給予我的支持與協助，可以讓我有充分的資源繼續與國外合作，交通大學在應用數學上的發展可以說是全台灣最多元的也是最有潛力的大學，電腦視覺問題在台灣數學界算是一個新的研究領域，再加上我們發展的數學又是新的分支，學校願意給我發展新領域的空間，我非常樂意在這個研究方向上繼續突破。這個領域的應用是非常廣泛的，在美國也是非常熱門的研究領域，許多高科技公司(像微軟公司、雅虎公司、Google 等等)都有設置的專門的研究部門進行研究，在未來，我希望能夠培養台灣學生對這方面的認知與興趣，並開設一些近代的應用數學課程，讓台灣的數學教育能夠與國際接軌。

另外在學術研究方面，為了擴展台灣與交通大學在國際上的知名度，計畫將在近期與美國政府部門 DARPA(Defense Advanced Research Projects Agency)簽署關於影像辨識的研究合作計畫成為計畫的主要研究員(Principal Investigator)，我將代表台灣交通大學與布朗大學和約翰霍普京斯大學負責此計畫的研究。因為這個研究，我也會需要出國做短期研究與合作，希望交通大學能夠繼續給我在學術上的支持與出國的彈性。另一方面，如果能培養出優秀的學生對我們的研究領域相當感興趣，我希望學校能盡可能協助爭取讓學生有機會短期出國進修、交換學生或甚至到國外一流大學攻讀博士。