

主動式雙眼視覺系統之精確校正

洪一平

資訊科學研究所副研究員

一、從固定式的電腦視覺到主動式的電腦視覺

電腦視覺(computer vision)是一門研究如何使用電腦由影像中抽取資訊(information)的學問。早期電腦視覺的研究大多著重於被動地分析靜態的影像。經過三十餘年的研究，雖然已有一些重要的成果以及某些成功的特定應用，但目前的電腦視覺系統相當不具彈性，對環境（如：背景、打光、物體的位置等）的要求相當嚴格，也導致一個電腦視覺系統如果是為某一特定應用而設計的，往往無法使用於另一種應用。例如：在工廠中使用的自動檢測系統即無法用來做門禁管制的臉孔辨識系統或汽車的自動駕駛系統。於是我們會問：為什麼現有的電腦視覺技術仍無法使機器人或智慧型系統具備人類所擁有的視覺能力？過去的研究是否缺少了那些重要的因素？

Gibson早在五十年代即指出：人類的視覺是主動的(active)而不是被動的(passive)。當我們在“看”的時候，眼球通常不自覺地或半自覺地跳動或轉動，並改變注視的焦點或觀測的位置。自1987年起，Aloimonos等人陸續對以往電腦視覺的研究方向提出一些批評，同時也約略勾勒出他們心目中電腦視覺的正確研究方向，其中一個相當受人矚目的觀念乃是主動視覺的概念。他們指出：有些視覺處理的問題，對被動觀測者(passive observer)而言，或許是不當的(ill-posed)、非線性的(nonlinear)、或不穩定的(unstable)，但對一

個主動觀測者(active observer)而言，那些問題可以變成適當的(well-posed)、線性的、或穩定的。他們並列舉五個著名的電腦視覺的問題以支持他們的論點，分別是：shape from shading, shape from contour, shape from texture, structure from motion，及optic flow。主動視覺的重要性與優越性從此廣為人知，並成為重要的研究課題。

二、主動視覺與立體視覺的結合—主動式雙眼視覺系統

這裡所謂的主動視覺(active vision)與傳統所謂的主動感測(active sensing)並不相同。傳統的主動感測指的是主動地傳送雷射光、超音波、雷達波等能量到被觀測物體上，然後藉由反射回來的信號，測量或分析該物體。至於主動視覺所要解決的問題則是：如何根據觀測者的目的（或所欲完成的工作）以及當前的狀態，決定並控制影像的攝取及處理方式。因此，為了從事主動視覺的研究與實驗，首先應具備一套電控可動式攝影系統。七十年代初期，Tenenbaum即在史丹佛大學構建了一個可以旋轉攝影機並變換鏡頭的取像系統。八十年代初期，Sanderson等人也在卡內基美崙大學裝配了一個旋轉式攝影裝置，可控制鏡頭之變焦(zooming)及對焦(focusing)，但仍僅使用一個攝影機。

八十年代後期，由於立體視覺(stereo vision)的技術逐漸發展成熟，開始有些研究人員結合主動視覺與立體視覺的特性，構建電控可動式立體攝影系統。這種主動式雙眼視覺系統較傳統的單眼系統更為接近人類的視覺系統。這種改變雖可簡化一些過去難以解決的問題，卻也產生了一些新的問題有待研究及解

決。其中一個重要的難題為主動式立體攝影機之精確校正，亦即如何在攝影機隨意轉動的狀況下，仍能得知任意組態的立體攝影機之內在與外在參數，以利進行立體視覺所需之運算。

三、新的問題—主動式立體攝影機之校正

簡單地說，立體視覺係使用兩個（或兩個以上）的攝影機對同一物體取像，再利用這兩張影像之間的視差（disparity），計算出三度空間中物體上的每一點的位置。而欲由二維影像視差計算三維物體點的位置之關鍵在於必須知道這兩個攝影機在取像時的參數。攝影機參數可以分為外在參數與內在參數。外在參數指的是攝影機取像時的方向與位置，包含旋轉與平移等六個參數；內在參數則包含鏡頭的焦距（或像距）與失真係數、CCD感測元件的像元大小、以及透鏡光軸與感測元件的交點等。對一個固定式的立體攝影機組而言，這兩個攝影機之參數可以經由對一個已知的校正物體取像而求得相當精確的估測值，此即所謂的“固定式攝影機之校正程序”。

然而如果攝影機的參數一經變動（例如：移動或轉動攝影機、或改變鏡頭之焦距），則以上繁複的攝影機校正程序必須重新再做一次。此時最大的問題為，如何在實際應用中，隨時找到一個精確的已知校正物體置於取像視野中，以校正攝影機在該組態的參數。即使我們可以有這麼一個校正物體，該物體最好能大到主動式雙眼視覺系統無論轉動到任何一個視角，都能夠看到該物體，也就是該校正物體最好能固定不動，或者能作精密的已知運動。但是，這麼一個龐大的校正物體很難製作得很精密，更何況一個智慧型機器人會帶著他的視覺系統到世界各個角落，因此，我們必須尋找其

他的方法。其中一個可行的方法即是設法事先校準主動式雙眼視覺系統的各项內部運動參數（kinematic parameter），例如：相鄰兩個轉軸之間的運動關係、攝影機與運動機構的幾何轉換參數等，然後再於實際應用時，使用這些內部參數自動計算任意組態下的攝影機參數。

四、研究方法與成果

為了研究主動式立體視覺，我們自1990年起即開始著手設計與構建國內第一套主動式雙眼視覺實驗系統。當時世界上僅有屈指可數的幾所國外著名大學（如：伊利諾大學、賓州大學、哈佛大學等）的實驗室中有早期的主動式立體視覺實驗系統。我們設計的主動式雙眼視覺系統具有八軸的自由度：最前端的兩軸控制左右兩攝影機之鏡頭對焦(focusing)，其次兩軸控制兩攝影機之聚合(verge)，再其次兩軸控制立體攝影機組的上下轉動(tilt)與左右轉動(pan)，最後兩軸則控制整個雙眼機械頭(bi-nocular head)之前後左右移動，以模擬自走機器人之水平運動—參見圖一所示之“楊戩一號”(即IIS head)。最近這兩年，雖然已有廠商開始銷售給實驗室使用的電控式立體攝影機之運動機構，但仍無法提供足夠精確的系統參數與校正程序。因此，主動式攝影機校正之研究仍有其重要性。

在完成硬體結構及初步實驗後，我們即發現，系統內部運動參數的估測（或校正）必須非常地精確，才能確保最後所計算出來的攝影機參數足夠準確，也才能將現有的固定式立體視覺技術運用在主動式的立體視覺系統上。由於這個精密校正問題的困難度相當高，近來國外有些學者轉而尋求線上(on-line)的自我校正(self-calibration)方法，或尋找一些對攝影機



圖一 我們所設計的主動式雙眼視覺系統——“楊戩一號”（即IIS head）。兩個攝影機之間配置有一個固態雷射，可投射結構光源以提昇系統之感測能力。

參數的精密度要求較低的應用，如：視訊追蹤、視訊導航等。但現實中仍有許多重要的應用是無法避免校正問題的，例如：景物模型之建立與學習、三維式量測等。而採用線上的自我校正也有一個先天上的問題——即一般性的應用中，不易在影像中自動找到足夠多且足夠精確的影像特徵，以提供給自我校正模組使用；更何況事先的精確校正至少可以提供一個良好的啓始值給線上的自我校正，因而可簡化該問題的困難度。因此，主動式攝影機的校正一直是電腦視覺領域中亟待解決的一個重要問題。

在初期的研究中，我們對攝影機的校正問題做了一個完整的分析，並完成了兩項研究成果。第一項研究係於考慮鏡頭失真(lens distortion)的前提下，將攝影機校正問題轉換成一個特徵值問題而得到一精確的線性解答。接著，我們又發現，如果所要求的精確度許

可，則忽略鏡頭失真不但可簡化攝影機校正的問題，更可以加速後續的三度空間資料之計算。因此，第二項研究乃對未考慮鏡頭失真的傳統線性方法加以分析，並推導出一誤差上限估計公式——此誤差估計公式除了能提供一套準則以協助選擇合適的攝影機校正配置外，更可用來決定是否必須考慮鏡頭失真的因素才能達到所要求的精確度。

接下來，當我們開始處理主動式攝影機的校正問題時，才真正體會到問題的困難度，也深刻瞭解為何這方面的研究進展始終停滯不前，其中幾個重要的原因包括：非常微小的轉動角度估測誤差即可以造成非常大的影像預測誤差；某些系統內部參數彼此間有非常敏感的相關性…等。過去幾年我們曾嘗試許多方法，但都無法有大幅進展。直到1995年底，研究工作才算有了真正的突破。首先，針對四種不同型態的攝影機校正問題，我們對攝影機參數之估測精確度作了一次徹底的理論分析，藉以瞭解在不同型態的校正問題下，各估測參數之間的相關性，以及各估測參數對給定參數之誤差的敏感性。然後，根據該理論分析的結果，我們進一步發展了一套新的四階段校正方式，有效地分割會相互影響的系統參數，而終於能夠獲得很精確的主動式雙眼視覺系統的校正結果。在此四階段中，第一階段先做攝影機與電控鏡頭的內部參數校正，第二階段則估測運動機構的各軸關係，第三階段估測攝影機與運動機構間的幾何轉換關係，第四階段再做一次總體的運動參數修正——在我們的雙眼系統中，共有39個運動參數必須修正。經過第四階段的自動學習後，我們實驗室中的八軸雙眼視覺系統即使在各軸同時運動的情況下，仍可精確地預測任何三維目標物的二維影像座標；實驗結果

證實該預測之平均誤差小於1個像素(pixel) (參見圖二)，而同軸線(epipolar line)上的誤差更可小至0.2像素(參見圖三)。這個結果為世界上現有同類的主動式雙眼視覺系統中最精確的校正成果。

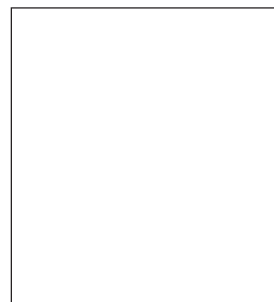
五、成果之重要性與未來研究展望

過去幾十年來，電腦視覺之研究已由最早的著重靜態影像分析，演進至後來的動態影像分析，乃至於最近的主動視覺之研究—欲使一智慧型系統可以主動地控制某些視覺參數的變動，以達到特定的目標，如：模型重建或物體

追蹤。我們的校正成果使得主動式雙眼視覺系統的攝影機參數可以隨時被精確地求得，進而使現有的固定式立體視覺技術能輕易地被運用到攝影機可以運動的狀況—目前我們正在執行的一項研究計劃即是要運用這項校正成果，整合由一個主動式雙眼視覺系統在不同觀測方位所得到的三維資訊，以重建四周的環境模型，供虛擬實境系統使用。對未來的各種智慧型系統而言，讓攝影機可以自主地轉動與移動乃是必然的趨勢。有朝一日，電腦要擁有像人類一樣的視覺能力可能不再是夢想，而其所能為人類社會帶來的便利與經濟效益更是無可限量。

圖二 根據我們的實驗結果，在第四階段的自動學習中，平均預測誤差可迅速降至1個像素以下。

圖三 在幾十次的自動學習之後，同軸線上的平均預測誤差即可降至0.2像素。



洪一平

學經歷：

國立台灣大學電機工程學系學士(1982)

美國布朗大學工程碩

士(1987)、應用數學碩士(1988)、以及工程博士學位(1990)

本院資訊科學研究所副研究員(1990-迄今)

國立台灣大學資訊工程研究所、國立中正大學電機工程研究所兼任副教授(1990-迄今)