

## 「眼見為憑」的背後

陳建中\*



現在回想起來，會一頭栽進視覺科學這個領域，似乎是個天意。

小學畢業的那個暑假，家人在一次出遊的時候，把相機交給我，在簡單的說明使用方法後（當時還沒有傻瓜相機），要我幫忙拍照。家人本來用意只是要留個到此一遊的紀念，但接下來的旅程中，這個相機就被我占著，成了我的筆記本，記錄我所看到的事事物物。從此以後，我開始對影像的產生與鑑賞有了莫大的興趣。高中時，看到攝影社招募社員，不假思考就加入了，也讓它伴了我高中三年的時光。大學聯考過後，會選擇心理系就讀的諸多因素中，其中之一就是當時認為對一個理組學生而言，心理學大概是最能幫忙我理解視覺藝術

\* 國立臺灣大學心理學系暨研究所特聘教授

的科系。在大三、四修習了黃榮村老師的知覺心理學和視覺消息處理之後，對未來的規劃就清楚了。再加上對數學的興趣，赴美念博士班的時候，選了以視覺為主題的計算神經科學。一直到今天都還在這個領域。

那我研究的課題是什麼呢？

生物視覺系統的功能是平穩、快速而正確的從周遭環境中擷取有用的訊息並加以解釋，使得個體能即時對外環境的變化做出反應，以增進自身生存的機會及福祉。由於視覺系統的高效能，我們往往沒有意識到這個時時刻刻在使用的功能是有多複雜。舉例來說，我們日常生活的每一瞬間所看到的影像都包含了成千上萬個物件。比如說走在街頭，我們會看到路上的行人及他們的臉孔、體態及穿著，路邊商店的招牌及上面的文字和店內的商品，路樹及其枝葉和花朵，建築物及其門窗和牆上的裝飾等等。而每一個物件都由許多像線段、顏色、質地等的影像元素所組成。在瞬間處理如此龐大的訊息，辨識成千上萬個物件遠超過人類現有的科技的能力。但對擁有上百億個神經元的視覺系統而言還不算上太大的難題，真正的挑戰還在於影像元素之間的組合可以有無限多種，也沒有任何先驗的理由可以指出哪種組合比較好，但不同的組合卻會導致完全不同的圖像及解釋。比如說圖形 X，可以是兩條對角線交叉，也可以是兩個左右箭頭 $\langle$ 交會，或是上下箭頭 $\vee$ 交會，也可以是兩個斜 T 對接。如果要考慮到我們眼睛所看到的影像，其實是三度空間的世界在一個平面上的投影，把線段的相對深度和斜率加以考量，即使是像 X 這樣簡單的圖形都會有無限多種解釋。然而，我們的視覺系統卻可以從這麼多的、可能的影像元素的組合中，在每個影像中瞬間取得穩定的唯一解。這個在一片混沌中，迅速而正確的組合影像元素的能力，就是我在過去十多年的研究主題：知覺群聚。

在影像進入眼睛後，早期視覺的功能是把影像分解成個別的影像元素。比如說，在大腦的初階視覺皮層中，每一個神經只會對視野中的一個小區域（接受域）內的特定指向、斜率、寬度的線段產生反應。換言之，初階視覺皮層的處理是把一個影像分解成一堆線段。當然，我們看到的世界是完整的物件和場景所構成的，而非一堆分離的線段。那視覺系統是怎麼把這些影像元素組織起來的呢？

我在博士班和博後的研究工作，主要是研究影像呈現的脈絡是如何影響人類的視覺表現。我們以單細胞電生理量測初階視覺皮層的細胞反應，發現初階視覺皮層細胞雖然不會對接受域外的視覺刺激反應，但它對在接受域內的刺激反應，卻會受到這些域外刺激的影響。這樣的遠距互動會根據域外與域內刺激的影像特性的差異而由所變化。比如說域外與域內刺激如果是共線的話，域外

刺激的出現會促進細胞對域內刺激的反應，但如果是互相平行就不會。這樣的現象，不光是在單一細胞層次，在細胞群的層次甚至是行為的層次，都可以由功能性核磁共振造影和心理物理學方法量測到。其實，要產生這種促進作用，不只域外與域內刺激要共線，還要共面，它們之間也要有一定的相似度。把各種因素加起來看，似乎遠距促進作用產生的條件，是在域內外的刺激合併起來可以表徵一個立體物件。

從這裡我開始思考，這種需要不同刺激之間要有特定空間關係才會產生的遠距促進作用，可能會是知覺群聚的內在機制。知覺群聚其實是一個早期心理學家就注意到的現象。尤其是完形心理學家所歸納的諸多法則，不僅是大一普通心理學必教，認知科學科普讀物必載，也往往是人們除了心理治療之外對心理學的第一印象。然而，由於欠缺適當的量測方法與數理模型，早期對知覺群聚的了解往往限於對現象的質性描述，難以探討其內在機制。我們從遠距促進作用切入正好可以彌補這樣的缺憾。

比如說，物體的對稱性知覺是一個重要的知覺群聚現象，人類可以十幾毫秒內偵測視野中對稱的部分。畢竟自然環境中，對我們生存重要的食物、朋友和敵人都是對稱的物體，而不重要的背景像是石頭或山川就不是，所以對稱性知覺在演化上有重大意義。然而要知道一個圖形是否對稱，我們先要把不同影像元素配對起來，如果各個不同的配對的中點都共線，那這個圖形就是對稱的，而中點的連線就是對稱軸。然而，由於對稱物體有可能出現在影像中的任何一個角落，任兩個影像元素的配對都可能是對。所以一個有  $n$  個元素的影像，就有  $m = n(n-1)/2$  個可能的配對，光是倆倆比較每一個配對又可能要和  $m(m-1)/2$  配對比較。這麼龐大的任務是如何達成的？在對稱性的計算中，影像特性一致的元素之間才會形成配對，這個符合遠距促進作用的特性，而可依此達成。不同的配對之間的比對則是第二階的遠距促進作用，中點共線的會整合起來，而不共線的則會互相抑制。因此，就會造成某些神經元的反應會隨著對稱的強度和對稱軸的位置和指向改變。這些神經的反應就可以讓視覺系統用來偵測影像中的對稱性。以這樣的進路，我們這幾年找到了許多知覺群聚現象的心理及神經機制。隨著累積的資料越來越多，我們開始看到不同知覺群聚現象的共通性，我們未來會試著釐清這些現象的關係，並試著用一個更一般化的模型統整解釋這些現象的內在機制。希望未來的幾年會有更令人興奮的發現和突破。