

研究成果

圖像語言的分析與辨識

張復（人社中心特聘副研究員）

「圖象語言」(Visual Language) 指的是出現在文件或其他媒介（如影視）上的影像文字。由於網際網路的普遍化，我們需要將文件轉換為可在網路搜索與呈現的數位資訊。當然，要做到這一點需要經過相當多的研究過程。基本上，文件內容對電腦而言，只是個畫面，要如何擷取部分文字，並加以切割與辨識，是這個研究領域所必須解決的問題。它不僅是一項繁雜的工程，而且是科學。

從事圖像語言分析，我們最初會想到的多半是人類在學習文字時被教導的方法。比如說，利用「永字八法」來拆解文字，利用某種影像分析的方法來尋找文字裡的筆畫等。在早期的研究裡，我們花了很多精力在這類的工作上，結果都事倍功半。從事後的檢討可看出，這些工作不會有令人滿意的成果，因為人類辨識文字大部分是在下意識裡進行的，我們未必能將該過程整理為有意識的程序。其次，所謂的「永字八法」或「筆畫分析」，是企圖在分類 (Classification) 之前先做一些子分類 (Sub-Classification) 的工作。然而，子分類的精確性通常比原先的分類差，因為前者所能夠處理的資訊遠比後者少。要建立高辨識率的成果，必須避免把分類工作建立在精確度不高的子分類成果上。

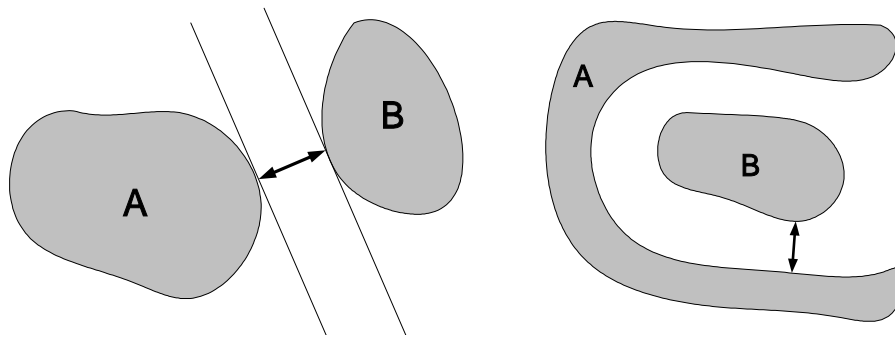
我們的研究工作開始獲得長足的進步，在於我們懂得使用機器學習的方法來從事各種工作。機器學習 (Machine Learning)，基本上是從樣本上來學習分類。然而，如何從樣本的特徵去做分類呢？人類在這問題上爭吵了一兩千年，足見這是一個十分擾人的問題。問題的關鍵在於，我們的老祖宗企圖以意識 (Consciousness) 所掌握的資訊來解釋下意識 (Sub-Consciousness) 的運作。所以，當我們被問到椅子與杯子的差異時，我們會說：「椅子有四條腿，而杯子沒有」，我們會這麼回答，是因為我們的意識在同一時間內只能聚焦於少數的特徵。所以，我們傾向於使用少數的特徵來解釋分類的原理，而且誤以為這就是我們的腦子所使用的方法。

要做好分類的工作，首要的認識是，特徵的種類愈多，愈能夠照顧單一特徵所不能照顧到的區別。比如說，「四條腿」只是某種椅子的特色，並不是所有椅子所必備的特徵，甚至不能將殷商的器皿排除在範疇之外，因為後者也有四條腿。為了提升分辨的精確度，我們必須收集更多的特徵，如「具有凹洞」、「可以儲存液體」等。多種特徵所帶來的問題是，它們不容易以邏輯的規則來處理，但可以經由機器學習的方法形成分類器 (Classifier)，來協助做分類的工作。

如果我們把一個物件所有的特徵數質放置在 D 度空間上，D 也是特徵的總量，那麼這個物件就成了該空間中的一個點。如果我們繼續將所有同類的物件放置在該空間中，接著又把另一個類別的物件也放置在同一空間裡，最後我們會看到甚麼呢？我們會發現，同類的物件雖然不必然集中於一點，卻會群聚在一起，而且與另一類物件所形成的群聚有明顯的分離。機器識別，就是企圖使用某些特定的數學方法，把兩個類別在空間上的分離狀態描述出來。

比如說，支持向量機 (Support Vector Machine) 會尋找兩個平行的分割線，將圖 1a 的兩個灰色色塊分離開來，這個方式很有效用，卻不容易分離圖 1b 的兩個色塊。此外，支持向量機只能應用在二元的分類上。在文字的分類應用上，如果文字的種類為 N，支持向量機就必須解決 N 個一對多的分類問題，或者 $N(N-1)/2$ 個一對一的問題。這個方法應用在文字類別很多的語言（如中文）上，極為不方便。為了因應後者的需要，我們發明了一種新的分類學習方法，稱做調適性的原型學習 (Adaptive Prototype Learning) 方法，可以同時建構多類別的原型，而且在分類的效應上，表現得比支持向量機要好，這是一個很令人振奮的研究成果。

圖 1 兩個灰色色塊代表兩組不同類別的樣本



(a) 這兩個色塊之間可以被兩個平行線分離開來。 (b) 這兩個色塊雖然有明顯的分離，卻無法以平行的分割線來分離。

除了將機器學習的方法應用在文字的辨識、排版分析、文件影像二色化、影視字幕辨識（圖 2）之外，我們還可以將同樣的方法應用在文章的分類、基因的族群分類、微陣列（microarray）基因分析等研究領域上，這是目前已完成或正進行的研究項目。

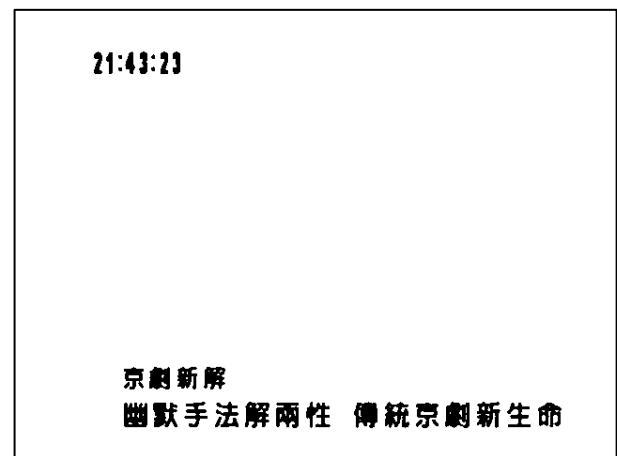


圖 2 左：影視畫面。 右：二色化的文字部分。

另一個我們即將完成的研究題目是「臉孔魅力的分析」，利用原型學習的方法，可讓機器來判別女性臉孔的魅力。其判別的能力並不輸給人類（但是機器需要人類幫忙測量數據，這就是不如人類的地方），我們提出充足的證據，駁斥了「平均就是美」的理論，並且提出了替代的看法。

我們正開始著手的一個有趣的研究，則是神經訊號的分析。大腦如何處理神經細胞所發出的電位訊號，是神經科學裡尚沒有定論的問題。我們認為，機器學習所企圖解決的問題，正是大腦在處理神經訊號所面對的問題。所以，我們期望原型的學習方法可以再度為我們找到更多有趣的發現。