

以電腦視覺數數

江蔚文

國立台北護理學院資訊管理系

摘要

以人眼辨識人類伸出的手指數目，是一件簡單的事，然而對電腦而言卻需要縝密的程式才能正確判斷。不變矩(moment invariants)常被用來描述形狀，因為它能在影像縮放、旋轉與平移之下保持不變，所以相同形狀的物體，若在日後拍攝時因遠近與位置的改變形成不同的影像，其所計算出的不變矩，會與先前的值十分接近。然而不變矩所代表的是該形狀的面積、重心、離心率等等所轉換出來的幾何特徵，這些特徵未必能代表伸出的手指數目這樣的邏輯特徵。所以本研究另外以重疊影像取交集並配合鏈碼找出凸出區域，以凸出區域個數作為伸出的手指數目這樣的邏輯特徵。將兩種方法的辨識率加以比較，希望找出較為適用的方法。

在第一種方法中，先計算各樣本影像的7個不變矩量，當待辨識影像的7個不變矩量也被求出之後，採用歐氏距離的度量方法，找出最相似的樣本影像。以該樣本影像中伸出的手指數目，作為辨識結果。本研究共使用24張影像，其中左手與右手影像各佔一半，又依伸出的手指數目為0至5隻共六張作為一組，左右手各伸出相同手指數在不同位置拍攝兩張，故左右手影像各二組。用不變矩辨識時，先計算左手一組的六張樣本影像的不變矩，再分別計算其他18張測試影像的不變矩。辨識結果顯示，其平均正確辨識率約百分之五十六。若考慮左右手，也就是左手樣本僅測試左手影像，那麼平均正確辨識率為百分之八十三。為了怕左手六張樣本影像的不變矩不具代表性，作者也嘗試將各組相同手指數四張影像的不變矩取平均值作為該手指數的辨識樣本，然後測試本研究所使用的24張影像。這種方法得到的平均正確辨識率為百分之七十九，優於左手一組為樣本的百分之五十六，但低於左手樣本僅測試左手影像而右手樣本僅測試右手影像的百分之八十三。

第二種方法主要以鏈碼記錄圖形邊緣，並利用圖形重疊相交方式找出凸出的區域，其演算步驟的說明如後。二值化後之影像已先經過兩次形態學的侵蝕再用兩次膨脹來增加邊緣平滑度。接著依據方向碼記錄手形邊緣之鏈碼與所在座標。然後藉由平均值的計算，可以輕易求出手形的重心所在。假設手形的大小都經過正規化，可以經由固定直徑為180像素的圓形切出凸出的手指區域，因此下一步驟的工作在於找出直徑180像素的圓形邊緣的鏈碼。在找出手形邊緣鏈碼與圓形邊緣鏈碼之後，再將手形重心與圓形中心重疊在一起，找出手形與圓形邊緣重疊的段落，並記錄各段落的起迄座標與長短。根據各段落的起迄座標，在手形邊緣找尋相同的座標，即可計算凸出於圓形外之邊緣長度。若凸出物真的是手指的一部分且圓心的位置適當，其凸出於圓形外之邊緣長度A應遠大於和圓相交的段落長度B，因此本研究中設定 $A/B > 2$ 者才被辨別為手指。因為手形的重心未必是最適合作為圓心的位置，所以除了該點之外，也嘗試將圓心移至手形重心周圍 ± 20 個像素外的八個點，重新計算凸出區域。因為最適當的圓心位置會讓圓與最多的手指相交，也就會找出最多的手指數目，所以最後辨識結果是採計九個(重心與周圍八個點)不同圓心位置中，得出最多手指數目者當作辨識答案。用本研究的24張影像以鏈碼找出凸出區域去測試時，其正確辨識率為百分之百，沒有任何錯誤辨識。

從上面結果可以得知，用鏈碼找出凸出區域個數這種邏輯特徵，其辨識率優於用不變矩找出幾何特徵的辨識率。例如左右手各伸出五個手指時，兩者凸出區域個數這種邏輯特徵相同，然而它們的幾何特徵卻不一樣，所以容易辨識錯誤。也就是說，當辨識的標的有明顯的邏輯特徵時，應設法找出該邏輯特徵，而非盲目採用常見的不變矩計算方式。本研究專注於形狀的辨識，嘗試以具有明顯形狀差異的影像，比較兩種方法的辨識正確率，作為形狀辨識研究的初探。日後會將此兩方法應用到其他問題上，例如藥錠外觀形狀的辨識，或各種醫學影像的切割(segmentation)。

關鍵字：電腦視覺、不變矩、侵蝕(erosion)、膨脹(dilation)、鏈碼(chain code)

