

運用電腦視覺自動辨識技術 輔助MABC測驗評分

作者：王芸若、鄭芝涵、許瀟予

指導教授：莊鈞翔教授

研究動機

MABC測驗用於評估發展協調障礙，進行過程皆是採取人工評分方式，這樣的評分方式受到評分者主觀判斷的影響，使得結果會存在差異，進而導致測驗結果產生誤差，即使是經過訓練的專業評估者，仍然難以完全消除這種主觀性的差異。因此我們希望能夠以電腦視覺自動辨識的技術來解決此問題，並且希望能將此測驗普及至每位學童皆能自行在家完成測驗。

因此我們透過顏色偵測、骨架辨識等方法，讓電腦能自動化評分。這樣的方式不僅可以精確的捕捉學童的運動行為，提供即時的数据分析，減少人為因素對測試結果的影響以及人工的耗費，還可以實現更多的數據記錄及長期追蹤，這些數據也可以為個別化的治療提供有價值的參考，因此本研究將探討如何使用電腦視覺自動辨識輔助MABC測驗的評分。

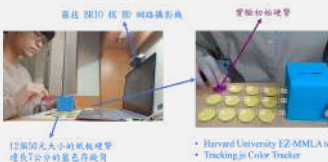
背景介紹

Movement Assessment Battery of Children-2 (MABC-2) 是一套用於評估發展協調障礙 (Developmental Coordination Disorder, DCD) 學童的評量工具，其中將年齡分成三個階段，3 到 6 歲、7 到 10 歲及 11 到 16 歲。評估的向度包含三大部份，分別是手部靈巧度 (Manual Dexterity)、平衡 (Balance) 以及瞄準和投往目標物 (Aiming and Catching)。一共包含 8 個子項目，並根據不同年齡子項目內容也會有所不同。

在本文的研究中，我們運用即時影片的方式來記錄實驗過程，並運用 EZ-MMLA Toolkit 來分析數據。EZ-MMLA Toolkit 是哈佛教育學院所開發的工具，用於分析影片或音源的數據，本研究將會使用工具包中的 Color Tracker 以及 PoseNet 來進行分析。

研究過程與方法

1. Manual Dexterity- Throwing coins



在 MABC 實驗中第一階段 (3 至 6 歲)，Manual Dexterity 項目中的 Throwing coins 實驗，以人工的方式來計算受測者將 12 枚硬幣投入盒子中的時間。然而這個方式難以界定實驗的開始與結束，透過計算總時長來界定手部靈巧度較不精確，因此我們想透過電腦來準確的定義何謂開始結束，並計算每個硬幣所花的时间，以進行更加精準的評分。

我們運用 EZ-MMLA toolkit 中的 Color Tracker 來輔助實驗。將硬幣設計成最容易辨識到的黃色，並把初始的硬幣做成粉色以方便界定實驗開始時間，將盒子設計成藍色與硬幣作區別並計算每個硬幣所需的時間。

在顏色偵測後會產生 Frame 數、物件長寬、位置座標及顏色等資料，運用 X,Y 座標判別硬幣是否移動，當粉色硬幣座標移動即為開始時間。另外運用每個 Frame 中被偵測到的物件數判斷硬幣投入盒子中的情況，當電腦偵測到只剩一個物件 (盒子) 時即為結束時間。

經過上述的改進後，我們會計算出整個實驗所耗費的時間以及硬幣個別花費的時間。

2. Aiming and Catching -Throwing beanbag onto mat



在 MABC 測驗中的第一階段 (3 至 6 歲) Aiming and Catching 項目中的 Throwing beanbag onto mat 實驗，採用人測評，受測者需於 1.8 公尺處將沙包投進目標區域十次。由於人工判斷在一些極端狀況容易出現主觀性的差異，例如沙包落在目標區域邊緣後滑走，沙包是否確實有接觸到目標區存在爭議，因此我們改善道具，並透過電腦視覺自動辨識來優化這個問題。

我們運用 EZ-MMLA toolkit 中的 Color Tracker 來偵測沙包顏色，將沙包做成三種不同顏色：黃色、粉色、藍色這三個較易偵測到的顏色。總共進行四次實驗，四次分別投擲三個、三個、兩個、兩個沙包。我們也將目標物改為白色方形，外圍用黑色色紙圍繞一圈，並且都貼上相同顏色的魔鬼氈，因為 color tracker 會用方形將偵測到的顏色標示出來，這樣在顏色偵測上較為精準。白色目標區的魔鬼氈目的是防止已投中目標物的沙包又滑出去；黑色區域魔鬼氈為防止未投中目標物的沙包滑進目標區的狀況。

最後偵測每一個 Frame 內有幾個相異顏色沙包，並判斷該沙包的 X,Y 座標是否在白色目標區域內。加總四次投擲裡投中目標區的沙包數，將其作為實驗的總成功次數。

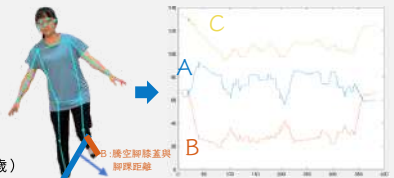
3. Balance-One leg balance



在 MABC 實驗中第一階段 (3 至 6 歲)，Balance 項目中的 One-leg balance 實驗，受測者需要單腳平衡站保持 30 秒，當一腳離開地板後即開始計算時間，過程中若騰空腳落地即代表失敗，並且騰空腳不可以勾著或靠著平衡腳，根據成功保持單腳平衡的時間換算分數。然而騰空腳抬起及放下的瞬間時間難以精準掌握，因此我們希望透過電腦視覺的輔助，判斷出客觀且精準的開始結束時間。

我們運用 EZ-MMLA toolkit 中的 PoseNet 來輔助實驗。在此骨架偵測下，可以得出身體各部位之 X,Y 座標資料，我們選擇在單腳平衡時，會改變最多的膝蓋與腳踝間距離與兩隻腳的腳跟距離做分析。首先，計算出騰空腳每個 frame 之膝蓋與腳踝間距離，以及兩隻腳的腳跟距離，分別將兩者取 2 倍 fps 之移動平均值，並將兩組移動平均值相加，得出最後可用來觀察移動距離變化之數據資料。

最後，將此筆資料的每個 frame 前後相減，取出第一個出現相差大於 0.7 的 frame，即為開始時間，以此類推找出結束時間，兩者相減即為單腳平衡測驗之實驗時間。



研究結果

我們用三種評分方式來檢測我們的實驗結果。分別是維持原本的人工肉眼評分、人工剪輯軟體評分、以及利用我們的演算法進行評分。其中人工肉眼的部分，我們將每個影片自行看碼表測試，每個影片測三次，並由三位人員評分，最後每支影片各 9 筆數據做平均。而人工剪輯軟體評分我們則是用剪輯軟體，以幀為單位，來計算確切的時間 (類似於慢動作觀看後評分)

- 硬幣：在硬幣實驗的結果可以從表格中看出人工測評比演算法平均多出了約 0.3719 秒，而剪輯軟體比演算法平均多出了 0.395 秒。我們推測會有這樣的誤差是因為，在人為測評時，每位觀測員認為「何謂投入」的定義不同，我們的演算法是定義整個硬幣的顏色無法再偵測到就算投入。
- 平衡：演算法計算的平衡時間大約比肉眼判讀時間多出約 1.48 秒。而我們認為在平衡測評中，較具爭議的是騰空腳是否有確實勾起，在一般實驗中皆以人工判斷和口頭提醒。因此利用騰空腳膝蓋到腳跟距離以及兩腳跟距離寫出的演算法，可以幫助此項目有更精確的標準。
- 沙包：由於結果都會是自然數，因此人工測評與電腦測評較無誤差。但從表格中的第一個實驗可以看到結果明顯不同。人工測評為 7 個沙包，而演算法測評為 7+2 個沙包。原因是在偵測顏色時均是以方格來界定顏色範圍，因此當沙包出界但靠近邊界時，方格仍會被偵測在範圍內 (如圖)，這也是之後實驗需要再進一步改變的地方。



項目	影片編號	人工剪輯測評三位評分委員取平均	人工剪輯軟體測評	電腦演算法
Throwing coins	影片1	14.295.3	54.8	14.295.8
	影片2	15.815.8	85.5	15.809.8
	影片3	12.979.6	12.934	12.997.6
Throwing beanbag onto mat	影片1	7	7	7+2
	影片2	9	9	9
	影片3	10	10	10
one leg balance	影片1	31.04	32.4	31.2081
	影片2	33.1289	33.79	31.1181
	影片3	33.63	33.834	31.0195