

基於深度學習進行人體姿態估計

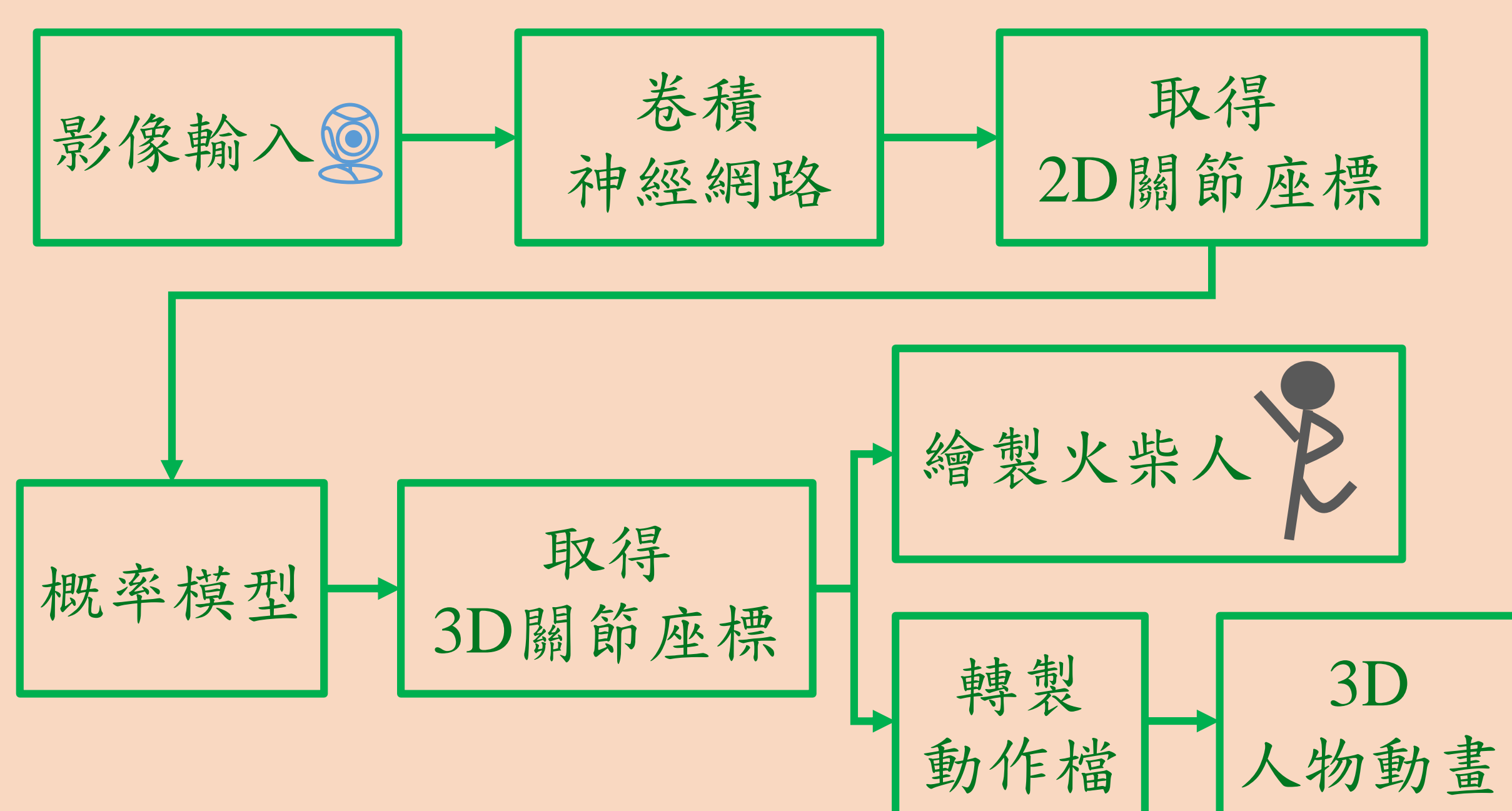
Human Pose Estimation via Deep Neural Networks

張俞方 施宣妤

如阿凡達、猩球崛起等賣座電影，其特效的製作過程常需使用動作捕捉裝置協助取得動作資訊。專題中我們希望透過攝影機獲取的二維影像，結合卷積神經網路學習的技術，轉換成肢體關節點在三維空間中的位置，進而捕捉人體的姿態及動作，並應用到 3D 動畫中。專題中實作的系統可以較低的成本與操作門檻，達到昂貴的動作捕捉裝置之相似效果，讓此類型的應用在日常中也能簡易的實現。

系統流程

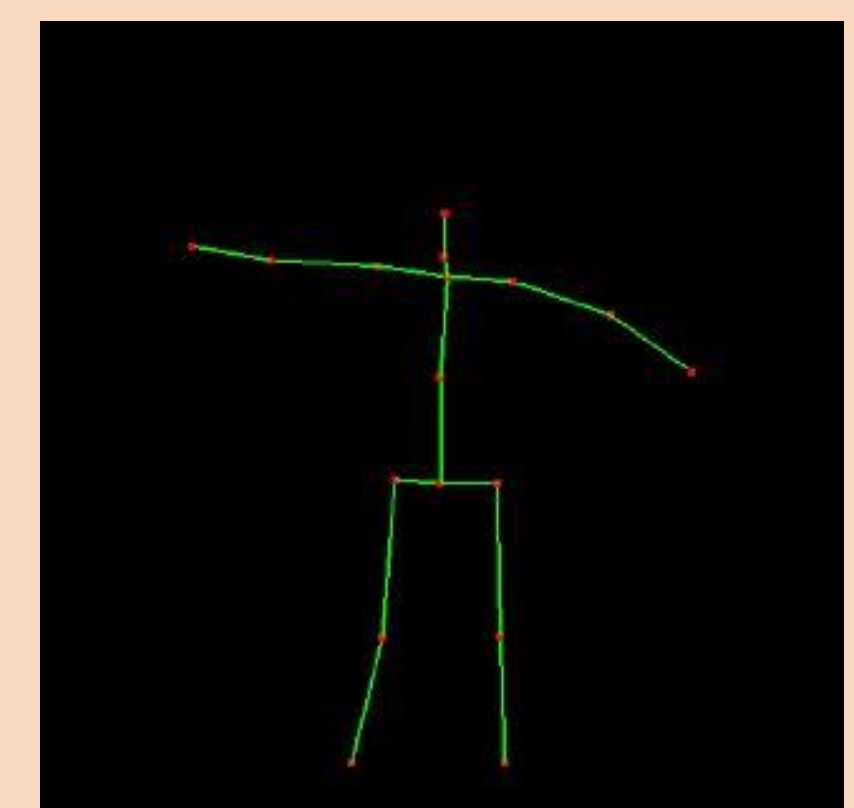
1. 藉由 Webcam 拍攝影像作為卷積神經網路的輸入
2. 卷積神經網路會捕捉關節特徵與關節間的向量特徵，最後得到二維的關節座標
3. 將二維關節座標透過概率模型提升至三維，得到三維關節座標
4. 以 OpenGL 繪製成火柴人展示
5. 將三維關節座標轉至成 3D 模型動作檔，輸入 3D 動畫製作軟體，完成人物動畫



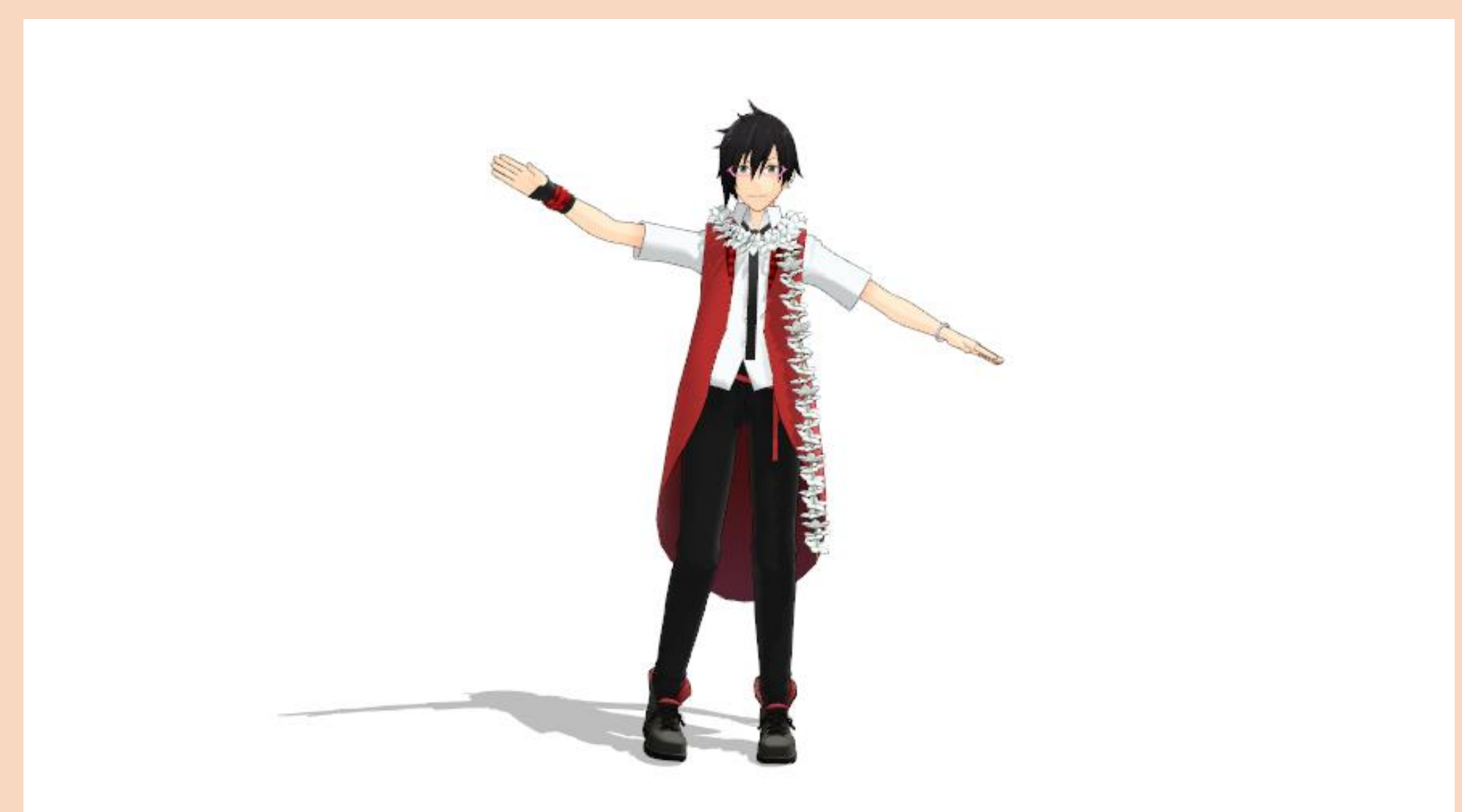
系統結果



▲ 原始影像輸入



▲ 火柴人繪製



▲ 3D人物模型輸出

優點

- 設備成本較低且方便取得
- 容易操作

結論

雖然專題中實作的系統無法進行臉部表情與手指精細動作偵測，但在估算人體姿態的正確性可達到相當不錯的水準。未來我們也將會繼續精進本系統的穩定度，使其可應用在更多需要人體姿態資訊的應用領域。