

3DCV Final Project Report – Wear App

R09921089 黃賀伶、R09921051 林奕昀

I. Problem Definition

WEAR APP 希望利用深度相機捕捉場景深度資訊，利用這層資訊幫助我們更好的估測物體的 pose，並讓物體能夠移動到臉部相應位置。

目前的新冠疫情大大減少了大家出門購物的頻率，化妝品的試色以及配件的試戴更多了一層衛生考量，在這個情況下，虛擬試色試用可以為我們省掉許多不便，讓人們可以無時無刻，隨時隨地試用並購買。

在我們的調查下，目前業界只用 RGB 圖片建完臉模後，才將物體配戴至模型上，如此非常耗時，隨著深度相機的普及，我們想利用深度相機抓取場景深度資訊，希望藉由這層資訊幫助我們更好且更快地估測物體的 pose。

II. Implementation

Camera & Model We Used

1. 深度相機: INTEL REALSENSE D435i (Fig 1)

說明 :

拍攝過程可同時輸出 Depth map 以及 RGB image
Depth map 解析度為 1280 x 720、FPS 90 帖 / 秒
RGB image 解析度為 1920 x 1080、FPS 30 帖 / 秒



Fig 1: Intel Realsense D435i

2. 模型 : SLPT - Face alignment model (Fig 2)

說明 :

Face alignment 做的就是通過一定量的訓練集(人臉圖像和每個圖像上相對應的多個landmarks)，來得到一個 model，使得該 model 在輸入了一張

任意姿態下的人臉照片後，能夠對該照片中的關鍵點進行標記。此 Project 所使用的 model SLPT 可預測出 WLFW dataset 之 98 個 Landmarks (Fig 3)。

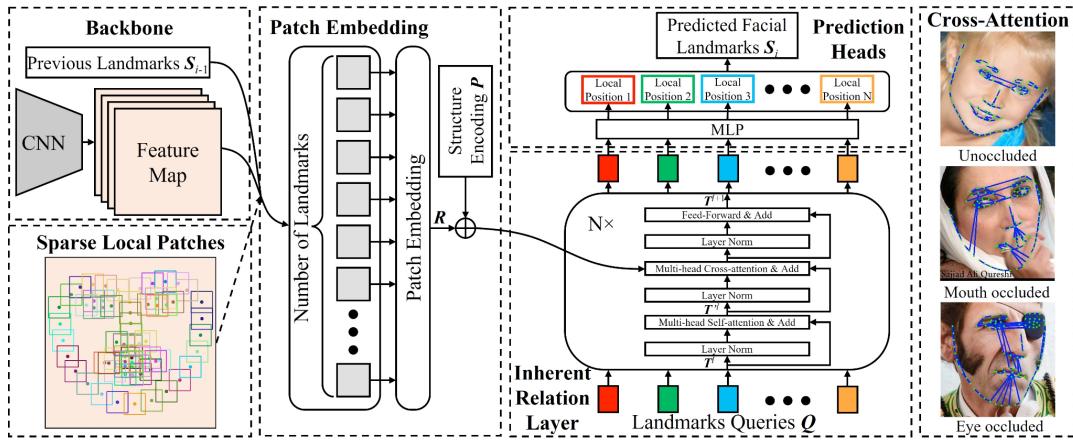


Fig 2: SLPT model

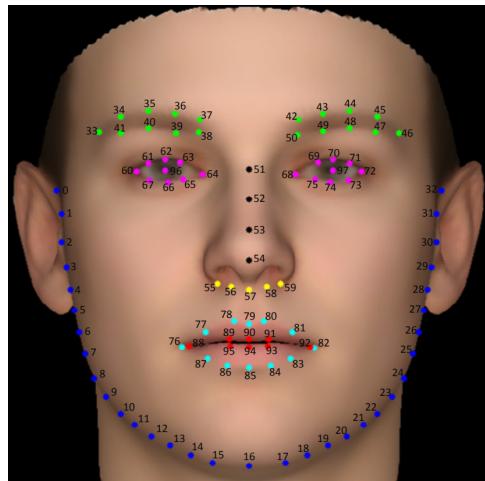


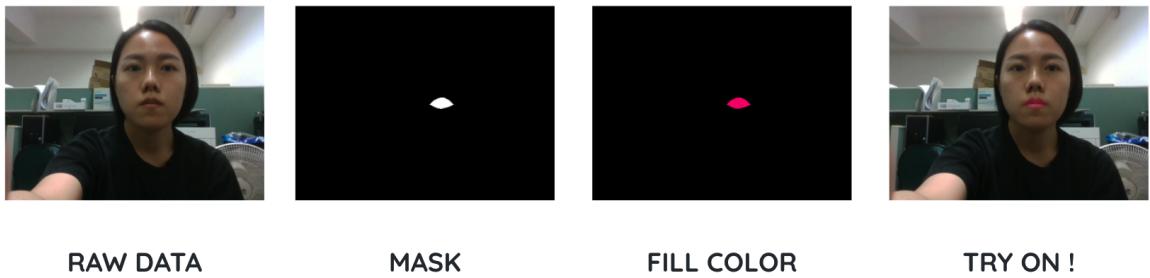
Fig 3: 98 Landmarks of WLFW dataset

Try On Cosmetic

在本案例，我們嘗試將試色應用在口紅與眉筆，實作細節如下：

- 由 SLPT model 偵測出來的臉部 Landmarks。
- 利用 Landmarks 框住所需部位並作遮罩，比如瞳孔、眉毛、嘴唇。
- 遮罩後將該區塊根據所指定之顏色進行上色。
- 將上色之遮罩圖調整顏色透明度、高斯模糊、明暗，嵌合到部位。

Fig 4, 5 之圖例分別為初始圖片、遮罩唇部(眉毛)、上色、試色後的樣子
Fig 6 為眉筆特殊色之上色近照



RAW DATA MASK FILL COLOR TRY ON !

Fig 4: 口紅試色



RAW DATA MASK FILL COLOR TRY ON !

Fig 5: 眉筆試色



Fig 6: 眉筆試色近照

Try On Accessories

我們在物件的試戴，實作了兩個飾品，分別是眼鏡與耳環，在這邊的實作我們會加入深度資訊，以利於我們做 pose estimation，並讓物體成功配戴到對應的位置。

以下會先簡單介紹整個流程架構，再講解我們對物體的 CAD model 做了什麼前處理，最後，會詳細的介紹眼鏡與耳環分別實作的細節。

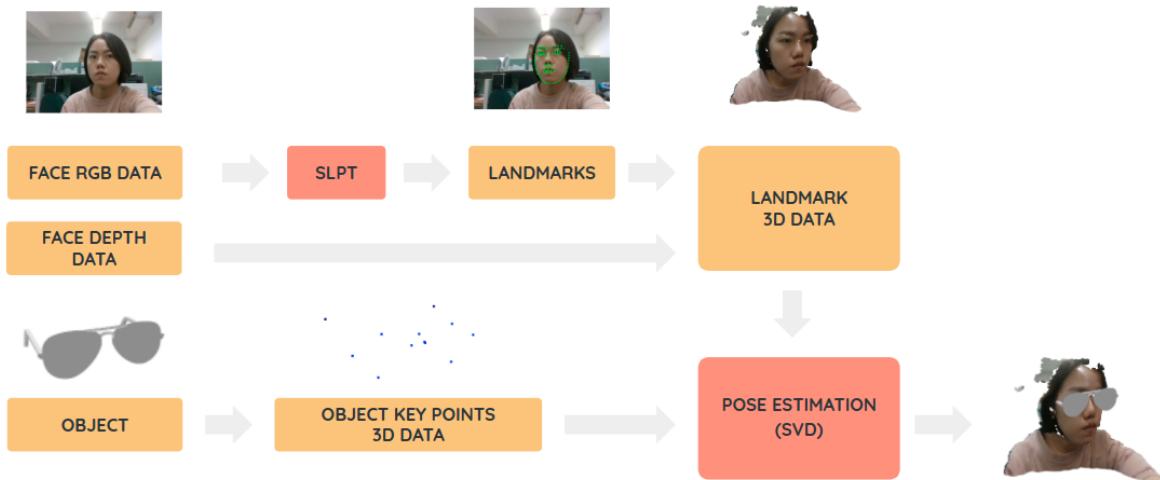


Fig 7: WorkFlow

WorkFlow Summary

1. 人臉的 Image 處理
 - 將RGB image 經過SLPT model, 取得 Landmarks 的 2D 座標
 - 從深度相片, 取得 Landmarks 座標點對應的深度
 - 將2D座標與深度結合, 得到3D的座標
 - 根據不同飾品, 找到臉部的key points 的 3D 座標
2. 物體 CAD model 處理
 - 在已知的CAD model, 標記並取得 key points 的 3D 座標
3. 將 step1. 與 step2. 分別取得的座標, 利用SVD 來做 pose estimation, 得到 R,T
4. 根據 R, T, 將物體移動至對應的位置
5. 將物體畫上, 並輸出 RGB image

CAD Model Pre-processing

眼鏡與耳環, 我們分別上網找了兩組免費的 CAD model, CAD Model 呈現於 Fig 8。



Fig 8: CAD models

我們使用Blender, 一款 3D 繪圖軟體, 來做 CAD model的前處理, 包含:

1. 統一座標:

- 正面朝向 +Z 軸, 水平平行 Y 軸
- 眼鏡的鼻樑位置, 置於座標中心 (0,0,0)
- 耳環的中心, 至於座標中心 (0,0,0)

2. 調整大小

3. 標記 Key points

4. 儲存格式

- 使用Blender 以外, 一開始我們使用 obj 檔來存取CAD model, 並 load 進open3d 之後再讀取圖片貼上texture, 但由於我們不熟悉如何取得mesh vertex的顏色, 所以最後我們選擇在mesh 上sample 六萬個點, 將 obj 檔轉成點雲的方式存取。

Try on Glasses

在臉部, 眼鏡的 Key points 不會完全貼合在眼睛上, key points 會約莫在眼睛前方一小段距離的位置, 所以我們會利用雙眼瞳孔連線的向量, 以及鼻尖位置, 找到與臉部平行的平面, 來計算其他關鍵點的位置與深度, 找出來key points (Fig.9)。



Fig 9: Key Points 在 3D 呈現的結果

在眼鏡上, 我們標記了多個 key points 嘗試做 SVD(Fig 10), 最終我們只選用了三個key points, 也就是鼻梁、左右眉毛下緣來做 pose estimation。

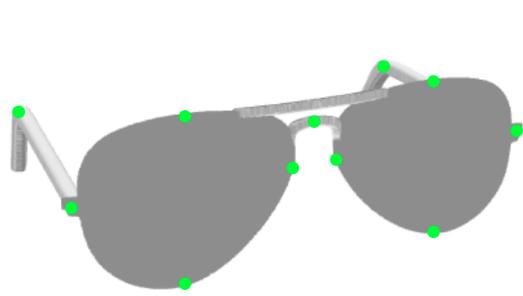


Fig 10: Key Points on Glass



Fig 11: Key Points on Earring

Try on Earrings

在耳環上，我們標記了三個 key points，分別是耳環的中心、耳針的頂點、耳環的最低點，如 Fig 11，標記這三個點的原因為：耳針的頂點可以控制旋轉，耳環的最低點可以控制垂掛的方向。

在臉部的部分，由於我們使用的model 並沒有偵測耳朵，所以我們使用臉部輪廓的landmark 來推算耳朵的位置。

並且，我們利用鼻尖到臉頰兩側的距離，來判斷人臉是否轉動，Fig 12為示意圖。以右耳為例，當人臉為正面時，耳環應朝向右邊；當人臉面向左側時，則耳環則面向正前方。我們利用上述耳環朝向的方向，以及左右臉距離的比例，來調整耳針要對應到的landmark，讓耳環有辦法跟著臉部一起旋轉。

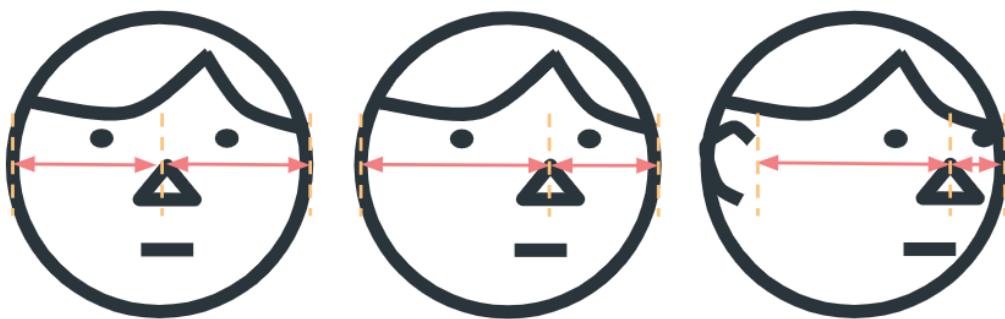


Fig 12: 臉部旋轉時，兩頰與鼻尖的距離變化

III. Result

[Video Link](#)

Results on 2D Images:



Results on 3D:



IV. Difficulties we have faced

1. 硬體：

REALSENSE 在拍攝過程中，由於 Depth map 跟 RGB image FPS 不一致，以至於最一開始生成的兩張圖常常是失準的，我們嘗試了 Github 上多方的解決方案才改善這個問題，但生成圖片過程仍有部分並未成功 Align，導致我們後處理上的困難。

2. 軟體：

- 在處理物體的 CAD model 時，我們使用的是 Blender，他的座標系與我們一般使用的坐標系不太一樣，為 -Z forward，而我們一般使用的是 Y forward，因此最初在處理 CAD model 時將物體投回 3D 空間時的樣子一直不如預期。
- 我們原本使用 triangle mesh 來讀取物體的 CAD model，但由於我們不熟悉如何使用，所以在貼上 texture 後，仍然無法取得 vertex 的顏色，後來我們選擇在 mesh 上 sample 多個點，將 mesh 轉為點雲來處理。

3. 準確度

- 最終的結果，在臉部轉動幅度太大時，我們計算的 RT 其實無法將飾品配戴在正確的位置，我們認為的原因是深度失準的問題，當臉部轉向側面，測量出的深度在建立成點雲後，左右眼與鼻樑的距離會有明顯的差異（在正面時距離相同），而這樣的差異也直接導致計算出的臉部 key points 無法與眼鏡的 key points 很好的匹配上，而這個現象就算調整 key points 的數量也無法解決。

V. Novelty

1. 在 Pose estimation 的環節加入深度資訊以作更精準的估測。
2. 以鼻尖之 landmark 為基準，作平行於臉部的平面，並將眼睛、眉毛投影至該平面作為與 Glasses key points 對映之 corresponding points。
3. 以鼻尖至臉頰兩側的距離，判斷臉部是否旋轉。

VI. What we have learned

1. 使用 SVD 計算點對點的 R, T 來匹配時，方法乍似容易，但在挑選 key points 仍有許多細節需注意，目前在選定 key points 時仍是以 demo 之 model 的特徵為考量，然而這樣的做法仍有通泛性不足的問題，如何考量臉型、眼距、鼻高等資訊來更有效更精準地配戴，會是未來我們得更深入考量的問題。
2. 在處理 3D model 相關的議題時，許多所面臨的問題已不純然聚焦在軟體實作本身，這個過程會需要硬體的協助(e.g. 深度相機)也需要第三方軟體的支援(e.g. Blender)——硬體面，如若對硬體規格不夠熟悉，那也無法有效地運用其生成的資源，其硬體限制也會是我們能否達到 real-time 的 bottleneck；軟體面，不同 3D 軟體都有其設計，座標系的差異、texture 的呈現、軟體如何操控，許多細節都考驗我們對線性變換或空間概念是否有通透的認知，並需再三確認，才能確保效果如預期。
3. 3D CV Project 每個環節幾乎都離不開矩陣運算或點雲處理的環節，本次 Project 我們在撰寫過程中仍稍嫌不夠有效率，如何減少時間與空間的浪費，並加速運算過程以達 real-time，是未來在實作相關專案時需深入考入的議題。

VII. Division of the work

R09921089 黃賀伶	R09921051 林奕昀
MODEL BUILDING TRY ON COSMETICS POSE ESTIMATION	OBJECT DATA PREPROCESS TRY ON ACCESSORIES POSE ESTIMATION