

109-2電腦視覺期末報告

虛擬試衣

系級：統碩二
姓名：林澤慶
學號：108354015

背景與動機

現今隨著線上購物越來越興盛，越來越多的顧客線上購衣，然而線上購衣相比於實體店面，其無法直接試穿衣服，因此常常發生所購買的衣服與想像中的衣服有所落差，導致顧客對於購衣的體驗不佳，於是產生了退貨潮。

為了避免發生退貨潮，於是我對於虛擬試衣感到興趣，虛擬試衣可以解決顧客不必到實體店面試穿衣服，其所用的方式是給定人像與衣物的影像，即可利用深度學習的技術，模擬人試穿衣服的效果，其效果如圖(一)，此功能服務可以增加顧客與電商之間的信任感，讓退貨率降低。

輸入



輸出



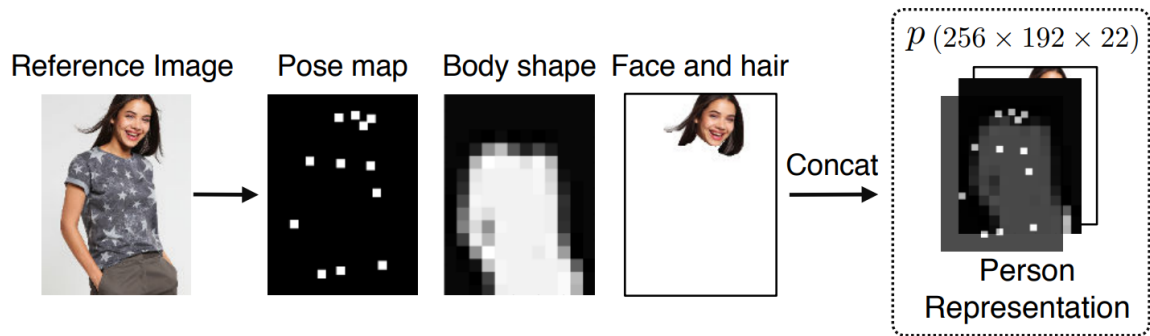
圖(一)

相關研究

本次參考文獻選自VITON(An image-based Virtual Try-on Network)
作者：Xintong Han, Zuxuan Wu, Zhe Wu, Ruichi Yu, Larry S
期刊：CVPR 2018

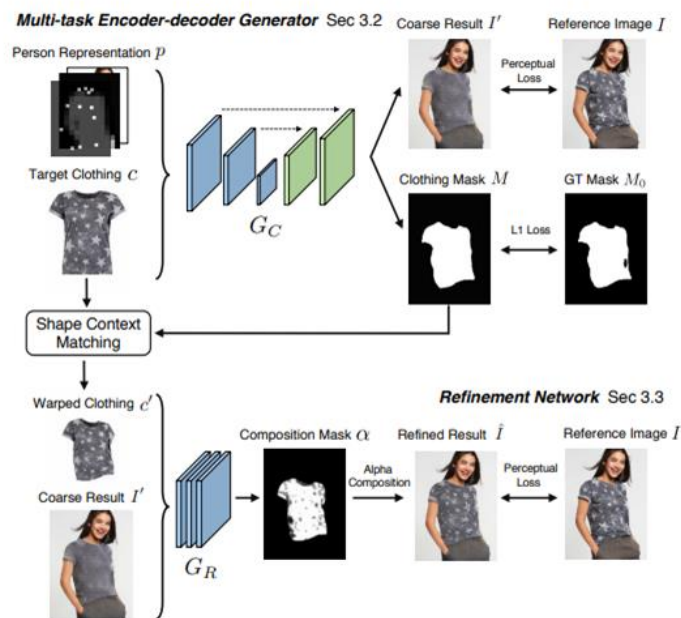
資料前處理

此篇的資料前處理如圖(二)，將人像的照片轉換出Pose map、Body shape與Face and hair三種圖片，其中Pose map代表人的姿勢，其可以透過open pose將原圖偵測出18個關鍵點，其關鍵點代表人的各個部位的位置，例如：頭、肩膀、手等等的位置，而每一個關鍵點會產生出大小為11*11的kernel，並透過kernel建立0-1 binary heatmap，Body shape代表人的身體形狀，Face and hair則代表人的臉部與頭髮，兩者圖皆可以透過JPP-NET來產生出人的語意分析圖所獲的。



圖(二)

模型架構：



圖(三)

此篇的模型架構如圖(三)，此模型分成兩種階段，分別為Encoder-decoder與Refinement階段。在encoder-decoder階段中，利用資料前處理所產生的Pose map、Body shape與Face and hair三種圖片與衣服圖片，透過U-NET產生出粗略的試穿衣服的人像 I' 與衣服的遮罩 M ，接續透過Shape Context Matching，將原本的衣服變形成欲試穿後的衣服的影像 c' ，在refinement階段中，將變形後的衣服影像與粗略的試穿衣服的人像，利用generator生成衣服的遮罩Composition Mask，並利用此遮罩產生出refine過後的試穿衣服後的人像。

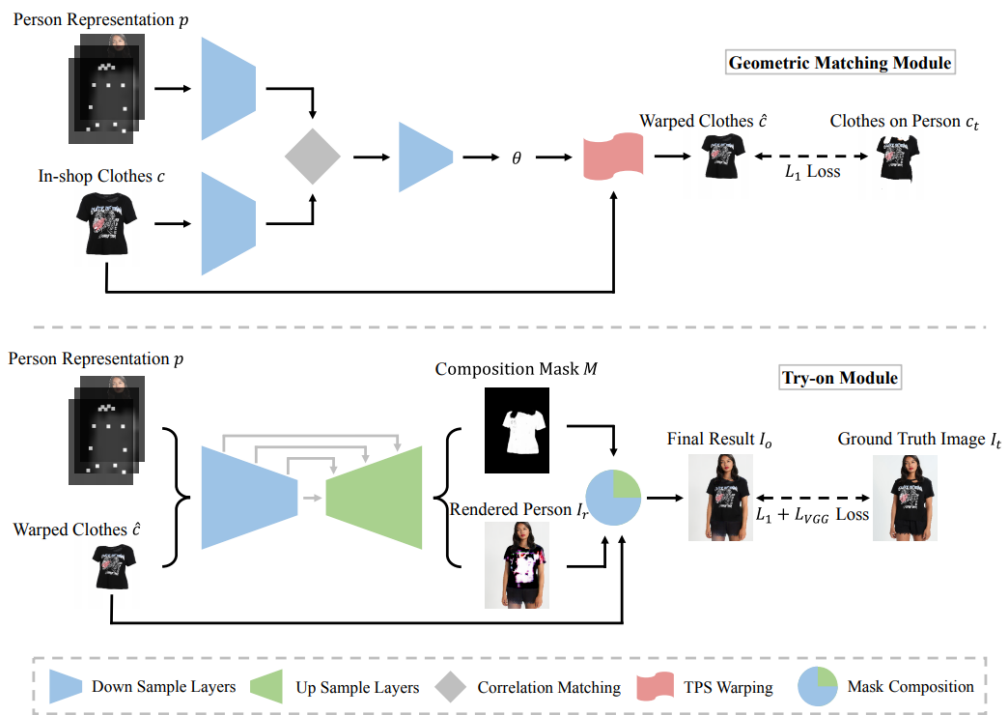
CP-VTON介紹

本篇論文CP-VTON(Toward Characteristic-Preserving Image-based Virtual Try-On Network)

作者：Bochao Wang, Huabin Zheng, Xiaodan Liang, Yimin Chen, Liang Lin, Meng Yang

期刊：ECCV 2018

模型架構：



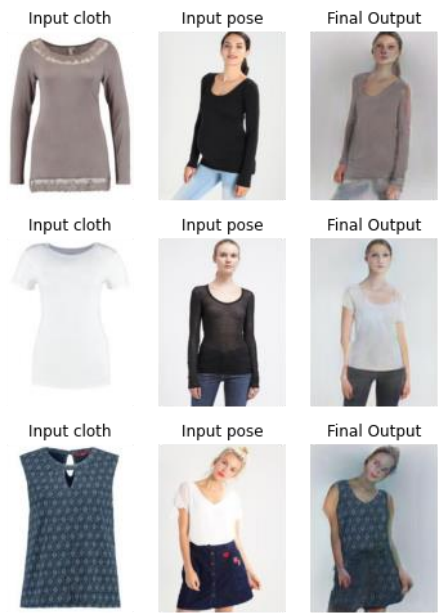
圖(四)

本篇的模型架構如圖(四)，此模型分成兩種階段，分別為Geometric Matching Module(GMM)與Try-on Module(TOM)階段，在GMM階段中，利用資料前處理所產生的Pose map、Body shape與Face and hair三種圖片與衣服圖片進行feature提取，接續進行correlation matching將兩組tensor合併為一組tensor，此目的用來配對衣服與身體的位置，接續將此資訊提供給變形的的方法，將原本的衣服利用Thin plate spline的轉換方法，變形成欲試穿人身上的衣服，在TOM階段中，將欲試穿人身上的衣服與Pose map、Body shape與Face and hair三種圖片透過U-NET產生出衣服的遮罩，最後將此遮罩與人的影像與變形後的衣服透過Mask Composition的技術，產生出試穿衣服後的影像。

資料集

本次資料集選自Zalando Dataset，在Train資料集中有14221筆資料，在Test資料集中有2032筆資料，每一筆資料中分別有人像與對應的衣服圖片，在人像中除了一般拍照的人像圖，還有人像的語意分析圖，在衣服圖片中，除了有一般拍照的衣服圖片，還有衣服的遮罩圖片。

實驗結果



圖(五)

本次實驗結果如圖(五)，在圖(五)中，左方代表欲試穿衣服的圖像，中間代表人的圖像，右方代表模擬出試穿衣服的結果，由實驗結果可以推論，基本上有看出試穿衣服的效果，但在衣服上細小細節的無法呈現出效果，因此在選擇素T會有比較好的效果。

結論



圖(六)

圖(六)為CP-VTON與VITON的比較圖，由圖可以發現CP-VTON比VITON表現還要好，在細節上CP-VTON能呈現出比較多，在兩者的技術上，CP-VTON比VITON多了 correlation matching 的技術，透過correlation matching與Thin plate spline的轉換方式，此模型在試穿衣服上，考慮了衣服如何變形以外，還多考慮衣服應該穿在人身上的位置，因此在結果呈現上，CP-VTON比VITON還要佳。

參考文獻：

1. Toward Characteristic-Preserving Image-based Virtual Try-On Network
2. VITON : An image-based Virtual Try-on Network
3. Look into Person: Joint Body Parsing & Pose Estimation Network and A New Benchmark