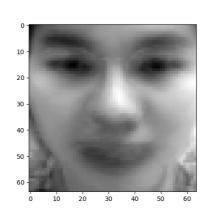
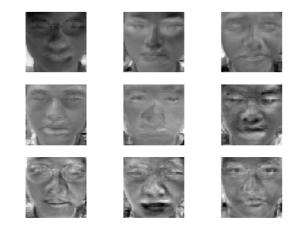
學號:B02404002 系級: 醫技四 姓名:葛竑志

1.1. Dataset 中前 10 個人的前 10 張照片的平均臉和 PCA 得到的前 9 個 eigenfaces:

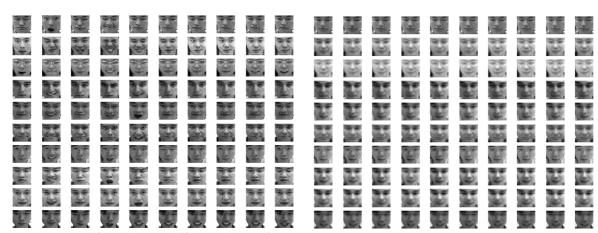
答: (左圖平均臉,右圖為 3x3 格狀 eigenfaces, 順序為 左到右再上到下)





1.2. Dataset 中前 10 個人的前 10 張照片的原始圖片和 reconstruct 圖 (用前 5 個 eigenfaces):

答:(左右各為 10x10 格狀的圖, 順序一樣是左到右再上到下)



1.3. Dataset 中前 10 個人的前 10 張照片投影

到 top k eigenfaces 時就可以達到 < 1% 的 reconstruction error.

答: (回答 k 是多少)

 $k=60_{\circ}$

2.1. 使用 word2vec toolkit 的各個參數的值與其意義:

答:

cbow = 1: Continuous bag of word model(CBOW),是透過指定位置周遭的 word 去 predict 出最可能為此 current word 的模型。Skip-gram 則是使用 current word 去 predict 周遭最有可能出現的 word。

size = 400:決定 word vector 的 dimensionality。

min count = 20: word 出現次數小於這個數字時會被忽略,而不進去 train。

window = 15: 決定在同一個句子內, Current word 與 Predicted word 的最遠距離。

negative = 1:要使用多少個 negative example。

iter = 300:決定 training 的迭代次數。

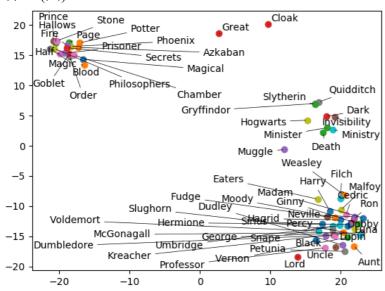
alpha = 0.025: Model 的初始 learning rate, 在訓練終點時將趨近於 0。

sample= 0(default): 對要處理的 word 設定其出現頻率的閾值。

hs= 1(default): 使用 Hierarchical softmax, 此法是將 output layer 做分類階層,可以改善傳統 softmax 輸出 neuron 的 dimension 必須要與對應字的總量相同,造成訓練沒效率的 窠臼。

2.2. 將 word2vec 的結果投影到 2 維的圖:

答:(圖)



2.3. 從上題視覺化的圖中觀察到了什麼?

答:

可以看到確實有相關性的詞會在分佈上有群聚的情形:

- (1) 左上角是關於 Harry Potter series 的書名,如 The Philosopher's Stone、The Prisoner of Azkaban、The Goblet of Fire 等。
- (2) 右中可以看到系列中出現的學院名稱 Gryffindor、Slytherin 與學院競技的 Quidditchn 所構成的群集。
- (3) 右下角則是系列中出現過的角色名。

- 3.1. 請詳加解釋你估計原始維度的原理、合理性, 這方法的通用性如何? 答:
- (1)原理:按照 handout 提供生出 data 的方法 generate 出一叢 data set 後,從各個 data set 中隨機 sample 幾個點與其各自的 neighbor,再去做 PCA 得到平均的 singular value,最後下去做 Linear SVR,得到一個 dimensionality(由於 Kaggle 評分方式,此處對其做 nature log)與 singular value 的回歸模型。當把 testing data 做 PCA 得到的 eigenvalue 送進 SVR 模型中即可 predict 出其 intrinsic dimension。
- (2)合理性:因為生出 data 的方法不會讓 data 在高維平面中被扭曲的很嚴重,因此可以假設在高維平面上取得的 neighborhood,在該低維平面上看其 neighborhood 也有相近的性質(可以想像成在一個曲面上的切平面)。從這些 neighborhoods 取得的eigenvector 理論上指向的方向不會差異太大,sampling 的參數若配合得宜,可以得到好的結果。
- (3)通用性:此方法有幾個前提:必須要了解其 data 的 generation 方式,否則無法 generate 出這些有 label 的 data 下去 train 出一個 model。此外,data 在高維平面上不可以被扭曲的太嚴重。
- 3.2. 將你的方法做在 hand rotation sequence datatset 上得到什麼結果? 合理嗎? 請討論之。

答:

- (1)結果:由於每張圖片的維度為 512x480 對記憶體來說過大無法計算,因此有使用 PIL 壓縮圖片。套用上題的模型後,預估的 nature log of dimension 是 0.96,大約 介於 2-3 維之間。
- (2) 合理性:不太合理,因為構築單一個黑白圖片的資訊需要兩維的資料,手部在轉動時會需要新的維度來構築其不同面向的長相,最後也需要再加上描述手部運動過程的維度。因此應該不會只有這麼少維度。