

學號：B04705014 系級：資管二 姓名：梁涓程

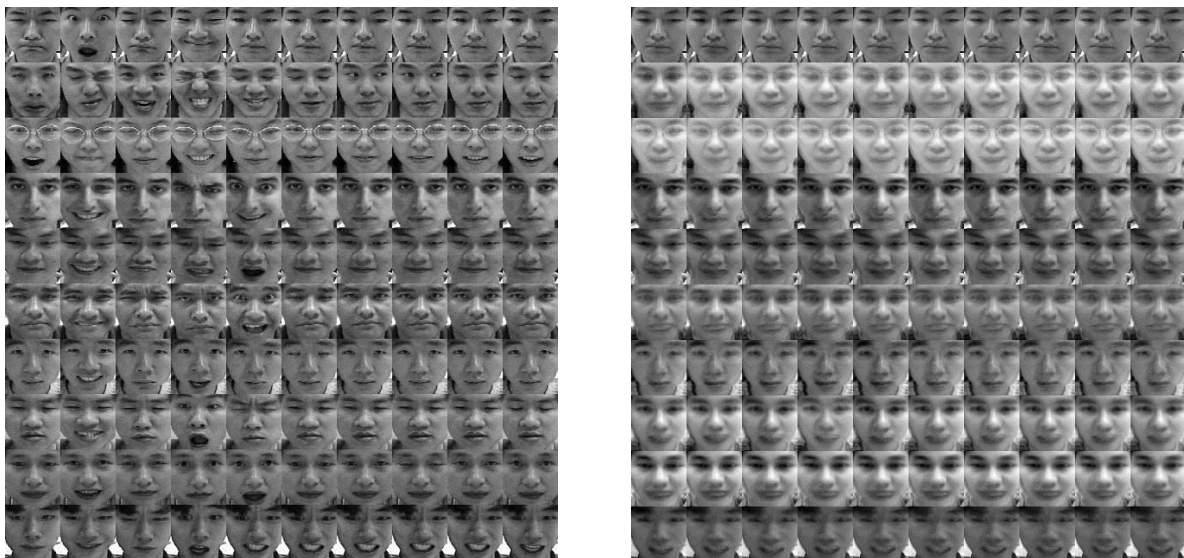
1.1. Dataset 中前 10 個人的前 10 張照片的平均臉和 PCA 得到的前 9 個 eigenfaces:

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 = 59 (測試結果如下圖)



### 2.3. 從上題視覺化的圖中觀察到了什麼？

答：

人名(左下)以及書名(右下)很容易就全部聚再一起，至於右上角感覺是想要從 dark lord 延伸出一些東西，但是結果貌似不太好(dark lord 跟 gryffindor 很靠近)。

### 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：

首先生產出原始維度是  $1 \sim 60$  的 60 類 data(處理後的維度是 100)，然後針對每一類，把處理後的 data 降維到  $1 \sim 60$ ，再 reconstruct 回 100，計算 reconstruction loss，得到 60 組機率分佈(組共  $60 \times 60$  個數字)，分別代表原始維度是  $n(n=1 \sim 60)$  降維到  $1 \sim 60$  再 reconstruct 回 100 的機率分佈(通常離原始維度越近，loss 就會越小)，接著再對 test data 計算新的機率分佈，看看跟哪個分佈比較接近，就猜那個維度。我認為這個方法對於依照助教的 code 生產出來的 test data，是非常合理的，畢竟 data 的性質都很相近，適用同樣的結論(loss 的機率分佈)。對於其他類型的 data，應該就沒辦法通用了。

### 3.2. 將你的方法做在 hand rotation sequence dataset 上得到什麼結果？合理嗎？請討論之。

答：

我認為我用上一題的方法估計出來的維度很有可能不太準確，原因是因為上一題我用的方法是直接計算 train data 的 reconstruction loss 分佈，理論上需要具有類似性質的 data 才能正常估計。