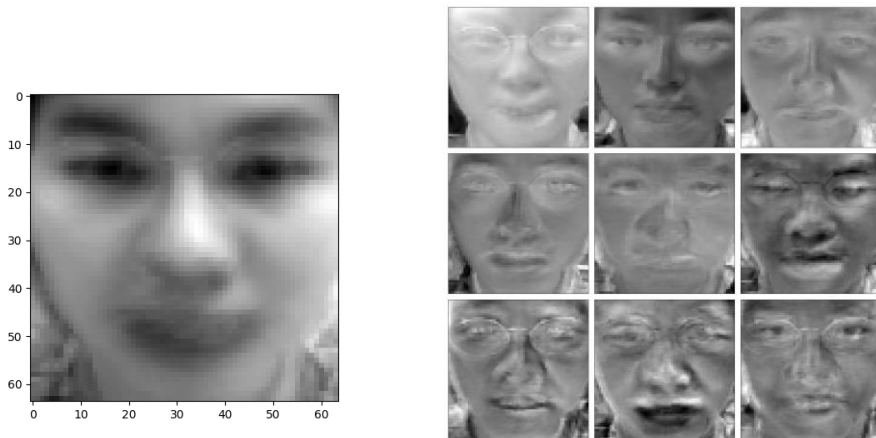


1.1. Dataset 中前 10 個人的前 10 張照片的平均臉和 PCA 得到的前 9 個 eigenfaces:

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)

60

2.1. 使用 word2vec toolkit 的各個參數的值與其意義:

答：

Train：train 的資料

Output：輸出的資料

Size：訓練出來向量的維度

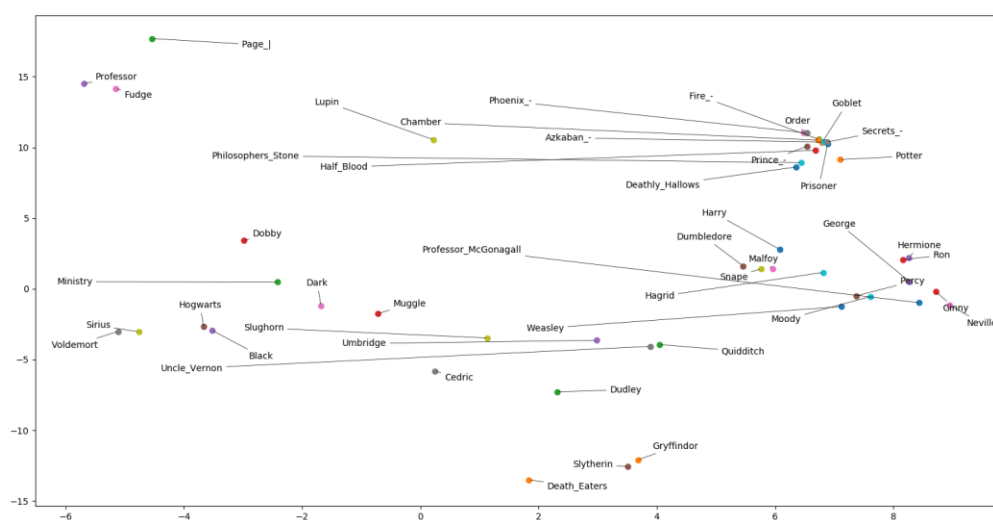
Min\_count：大於這個數量的詞才會被拿來 train

Window：能左右看幾個字的字數

Alpha：learning rate

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？

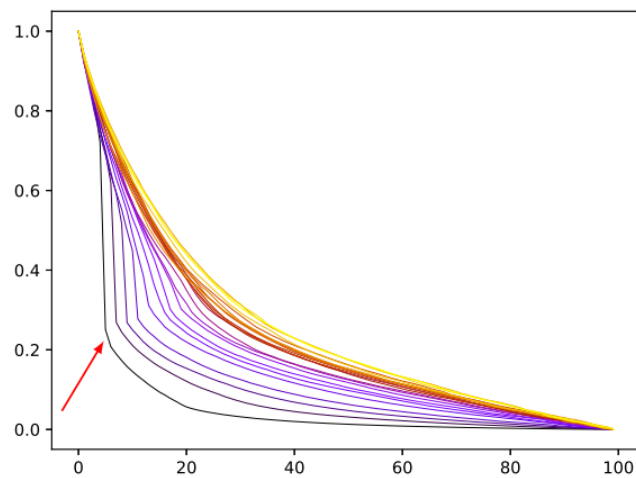
答：

我沒有看過哈利波特，不過大概知道史來折林和葛雷芬多是裡面兩個學院的名字，而在圖中這兩個字也是在附近。

而石內卜好像是鄧不利多的間諜?!在圖中兩者的關係也很近。

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：



利用助教課講到的，將每個 **eigenvalues** 除以 **eigenvalues** 的總和，會得到每個 **eigenvalue** 的重要性，再將其畫成曲線。因為 100 維中有許多維度是用來擴充的，所以多出來的維度會在圖上出現轉折，而轉折點可能就是原始的維度。但在越高維的時候，轉折點就越不明顯，這時再利用 **sklearn** 的 **SVR** 來學習預測轉折點。

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：

同樣利用上面的方法得到的維度是 10.8xxx。自己用肉眼對圖片分群的話，感覺分成了九群(明顯的轉動九次)，似乎是在合理的範圍內。