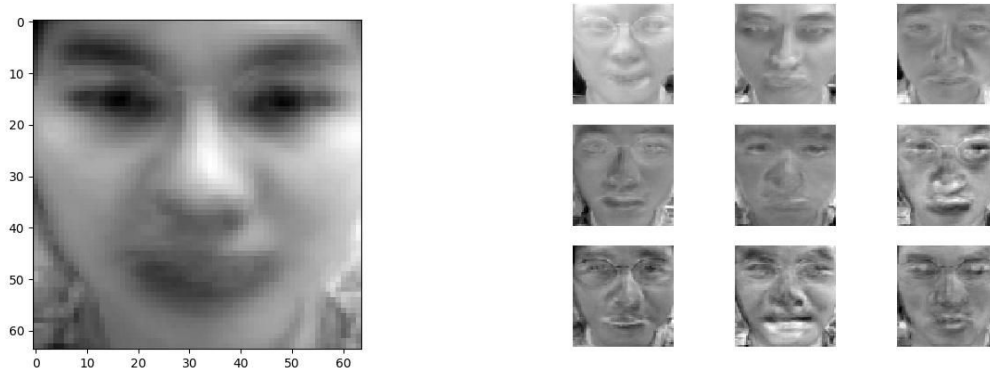


學號：R05921064 系級：電機所碩一 姓名：羅嵩詠

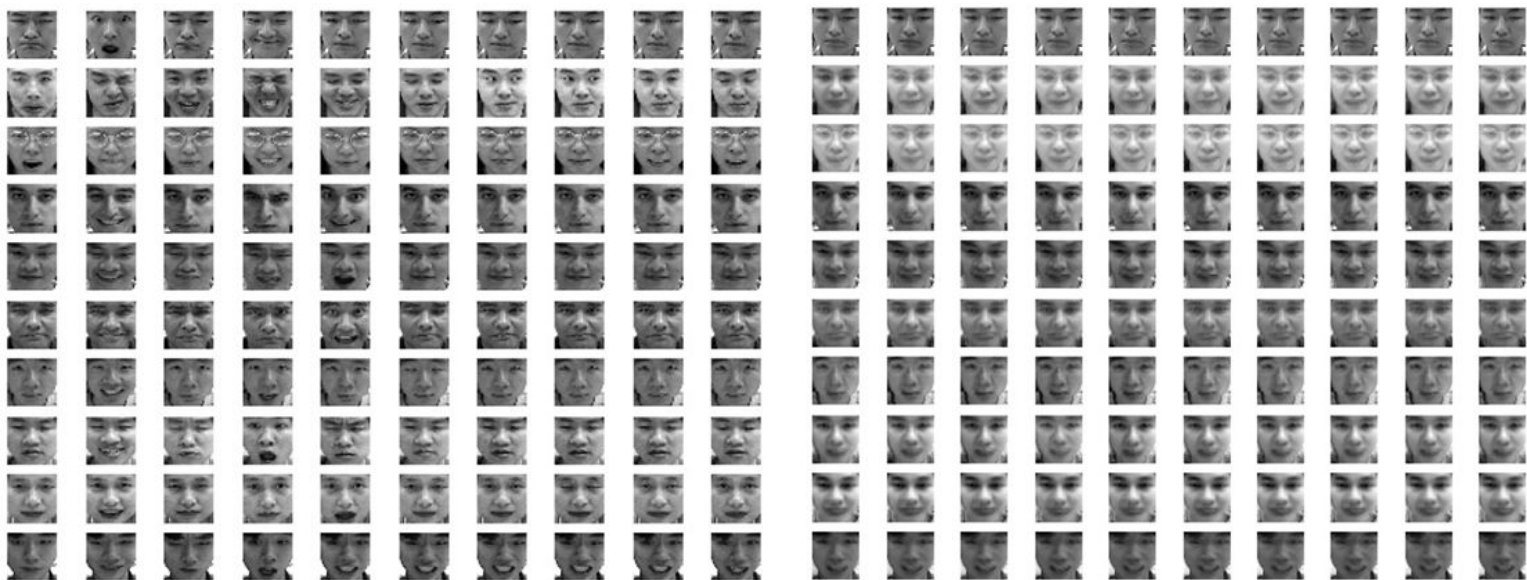
1.1. Dataset 中前 10 個人的前 10 張照片的平均臉和 PCA 得到的前 9 個 eigenfaces:

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

## 2.1. 使用 word2vec toolkit 的各個參數的值與其意義：

答：

size:表示訓練出的詞向量的維度。

alpha:初始的學習率。

cbow:要用哪種模型。0表示選用cbow，1表示選用skip\_gram。

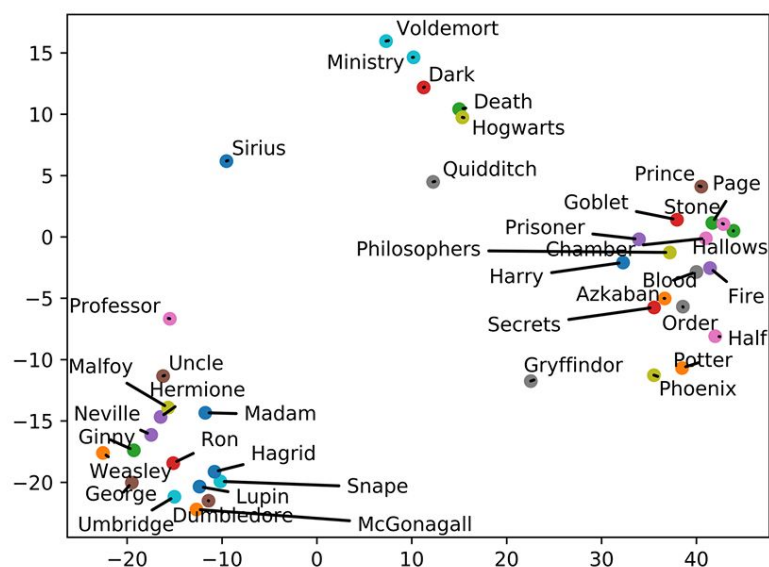
window:往前往後看，目前字詞與預測字詞的最大間距。

min\_count:若一個詞出現的次數小於此參數，則不會被視為訓練對象。

iter\_epoch跑的次數。

## 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：

中上的位置出現佛地魔、黑暗、死亡等，和反派角色相近的形容詞。左下角的位置是書籍中的角色與相關稱謂。中間偏右的位置為主角哈利波特以及和書名有關的詞語。

## 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：

使用助教課提供的方法，是為先生成所需的data，之後再根據其特徵值對應到的dim做SVR，稍微細調裡面的參數C，便可以得到不錯的結果。由於我們預先知道待預測data生成的方式，因此可以製造與其性質相近的訓練data。不過在大部分的實務中，我們對於待預測data的了解其實很少，因此這種方法在其他情況下，通用性並不高。

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：