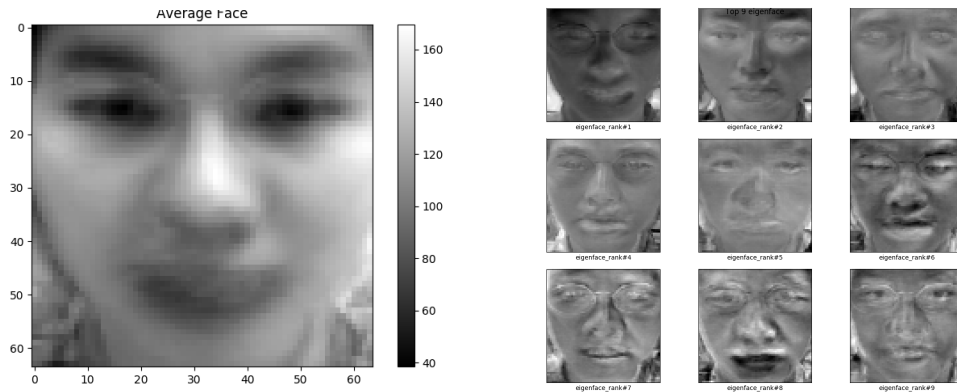


1.1. Dataset 中前 10 個人的前 10 張照片的平均臉和 PCA 得到的前 9 個 eigenfaces:

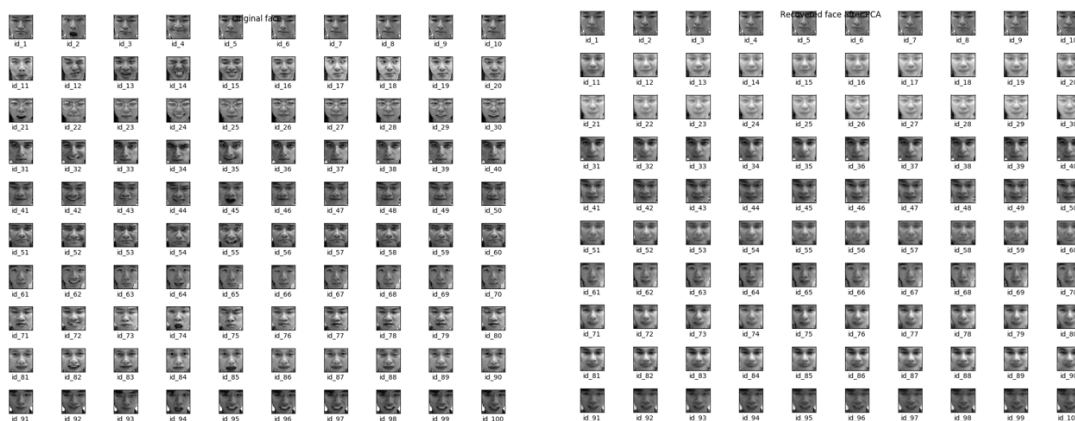
答：(左圖平均臉，右圖為 3x3 格狀 eigenfaces, 順序為 左到右再上到下)



可以發現最主要 eigenfaces 大致上為不同人的臉，我想這也符合預期，畢竟是拿十位不同人的十張照片來做 PCA，所以最後最主要 eigenfaces 便應該會是這些人的基本表情。

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)



將 reconstruct 的圖和原圖作比較可以發現大致上還是可認出這個人的長相，但是同一個人的不同張照片最後看起來的結果卻會差不多，無法還原出原本的細微表情差異。我想是因為原本各自的表情資訊是保存在非前五張的其他 eigenfaces 上，因此

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)

根據以下公式：

```
def est_error(e_face):  
    return np.sqrt(((e_face-pic)**2).sum()/(e_face.shape[0]*e_face.shape[1]))/256
```

當 $k=60$ ，會使得 reconstruction error $< 1\%$

在調參數的過程中，我觀察到當 vector 的向量維度調得愈大的時候可以把字分得愈明顯，因此我將維度調成 250 維。在上圖中可以發現，這些點大概被分做三群，而且可以發現關係密切的名詞大部分會靠在一起，例如榮恩、妙麗那一群等等互相有愛恨情仇的人的名字會聚在一起、鳳凰會和阿茲卡班這些比較偏機構類或者物品類的名字會聚在一起，另外一群則大概都是跟學校有關的東西，像魁地奇、葛來分多之類的。

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：

在這個題目中我嘗試了三種方法：

(1) 直接對測資做 SVD，觀察回傳的 singular value 的大小直接設一個門檻(我設 2300)來決定應該要選幾維當作答案，上傳到 Kaggle 後 public 的正確率大概都在 0.4 左右，後來發現，隨著真實維度的增加，門檻值也會隨之降低，直接寫死不是個好方法。

(2) 聽 windQAQ 分析說，因為所有測資的產生都是從較低維度的高斯分佈經過一些拉拉扯扯的處理後變換到高維，因此在高維的某些點他們的值會很接近，因此利用作業說明中所提供的 gen.py 所產生的 1~60 維不同的資料，計算這些資料中每個點和其最近的那個點的距離，並將這些距離的標準差當作這個維度的特色記錄下來，最後共有 60 筆標準差的資料，最後再用同樣的方法處理測資，比較其標準差和哪個維度最接近來當作答案。不過缺點是這個方法的計算量有點大，所以紀錄標準差時，每個維度不能產生太多筆資料，導致有時候如果取樣點不足，便會受隨機的部分影響產生爛結果，而最後在判別測資時也一樣，由於計算量的關係只能大概取個幾千筆資料來判斷，無法全部都下去算，因此最後丟到 Kaggle 上 public 部分的結果時好時壞，大概在 0.33 ~ 0.14 之間。

(3) 採用助教的方法，在資料集中隨便找一些參考點（找了十個），並尋找和其最近的幾個點（我設 200）當作是該參考點上切平面附近的點，再做 PCA，將得到的 singular value 的值簡單處理一下（除以當中的最大值，並再扣掉其平均），便將其當作該維度的特色，最後再經過 Linear SVM 找出最符合每個維度 singular value 趨勢的線。而在檢驗測試資料的維度時也經過一樣的處理，最後丟進去之前經過 SVR 跑出來的模型，便可以得到預測的維度。在不改 sample code 的情況下，直接丟到 Kaggle 上，public 的部分便有 0.14，而且運算相較方法二所需時間也少了許多，最後再稍微調一下參數（例如將 sample 的點提高，或者將其 penalty parameter 調高一點）便能夠輕鬆地跨過 baseline。附帶一提，為了降低每次隨機的結果導致最後預測的差異，因此我一開始便有先手動設定 random seed。

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：

因為想不到其他更好的方式，所以我便直接暴力使用 PCA，將資料（481 張照片）從原本的 481×245760 維降成 481×100 維，並將此作為一個 data set，然後再利用助教提供的方式最後估算得到的維度為 11 維。而我想這樣的結果或許也稱得上合理，因為每張照片的資訊其實都一樣，差別只在於手選轉的角度不一樣而已，因此若我們已經有了手的資訊以及其他背景資訊的話，只再需要一些其他資訊便能夠完整表達這個 data set 中的一張圖片。