

# 用SVD进行人脸数据分析

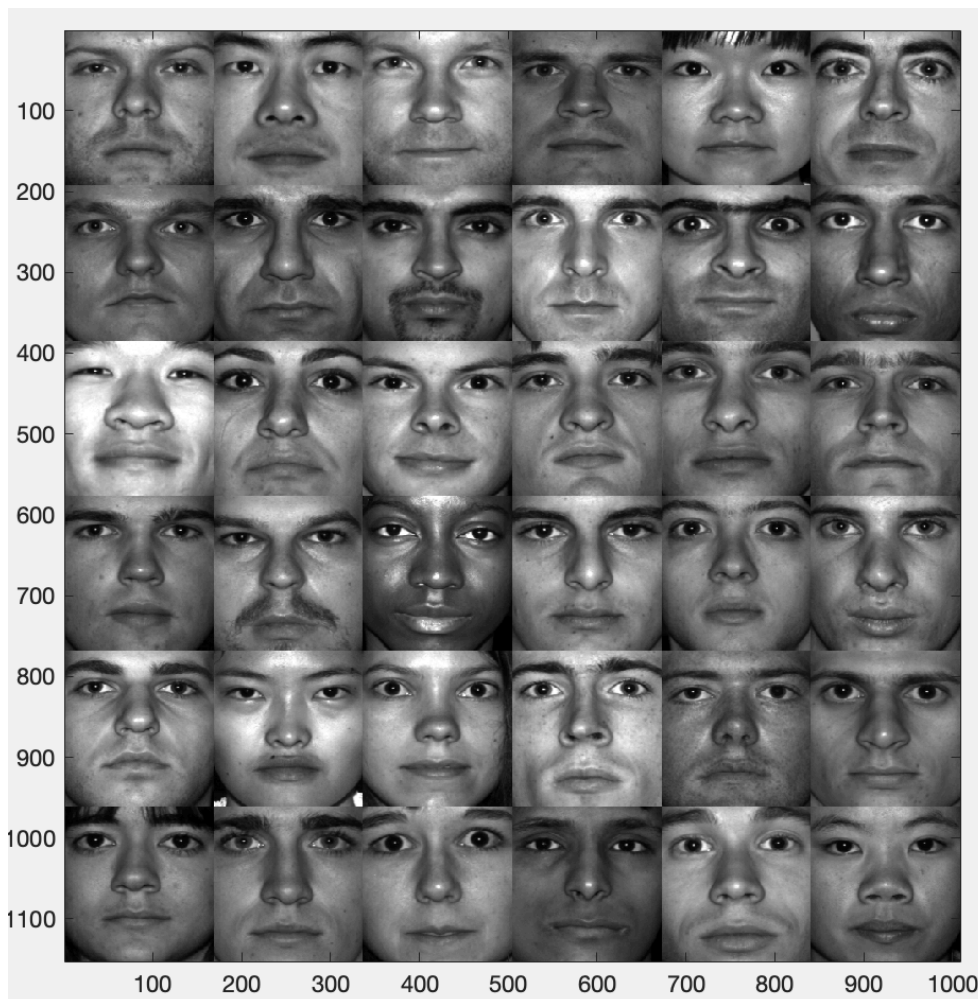
## # 问题描述

结合课上学习到的SVD相关的知识，学习了解PCA(Principle Component Analysis)方法，并利用给定的Extend Yale B Faces数据集进行人脸的特征分析和预处理。具体任务和示例如后面各小节所示，要求提交一份代码以及报告，说明各任务是如何实现的，以及自己的理解。

## 在matlab中加载并画出数据集

任务一：熟悉matlab加载数据与绘图的函数，画出人脸数据集中的数据。

提示：数据集包含38个人的脸，每张图片的大小为(192, 168)。画出的预期结果如下图所示。



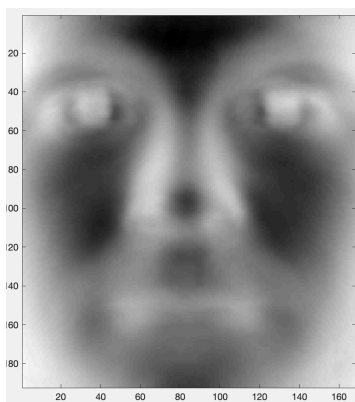
## 用SVD处理数据，计算特征脸

任务二：使用SVD对数据进行分析，提取并可视化前4张特征脸（左奇异向量）。

提示：首先将人脸数据去均值，方便后面数据的分析，画出的平均脸如下图所示。这里我们将数据集划分为包含前36个人的所有数据的训练集，和包含最后2个人所有数据的测试集。



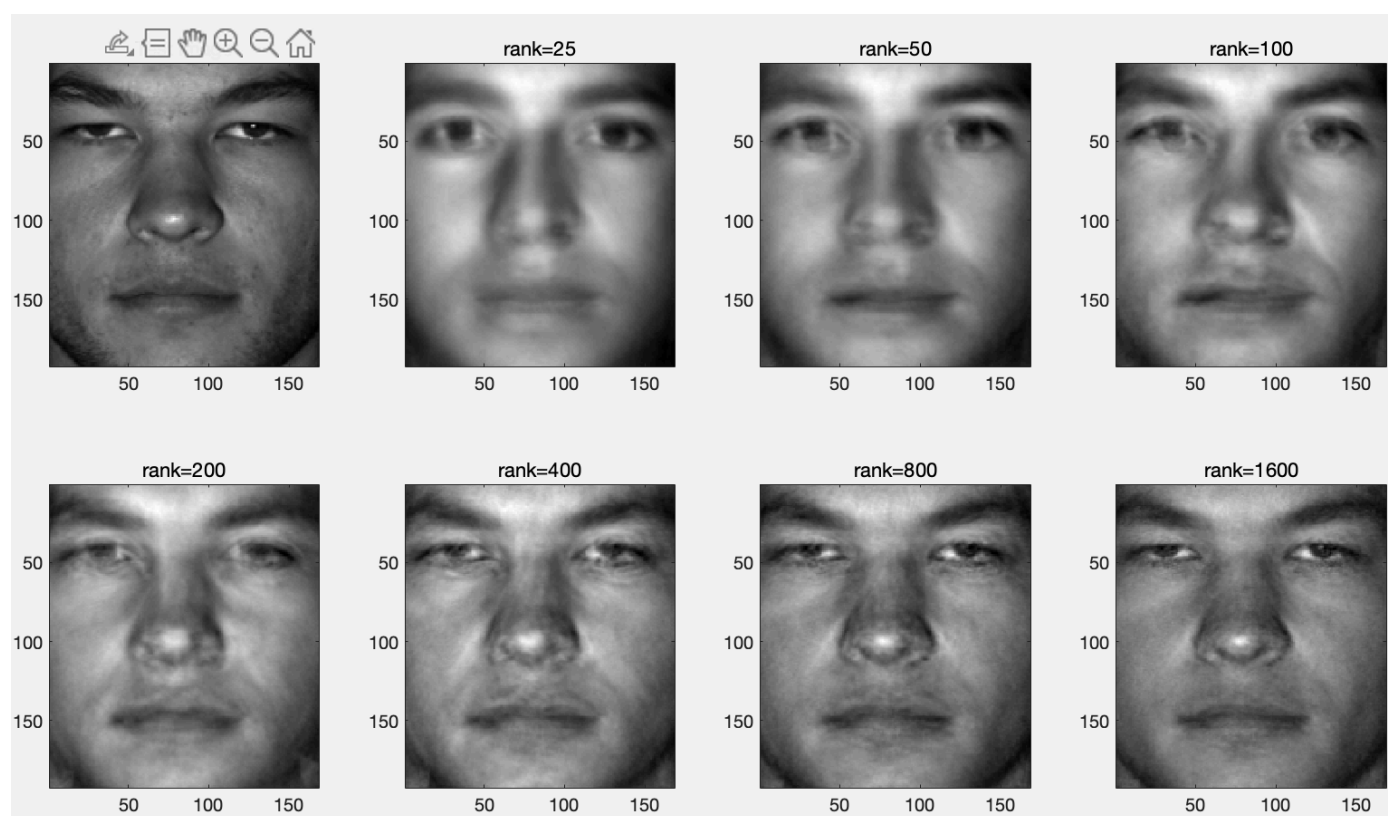
再用SVD对去均值后的人脸数据进行处理，画出特征脸。下图是提取的第一张特征脸。



## 人脸近似

任务三：利用SVD得到的奇异向量进行低秩矩阵近似(包括一张测试集中的人脸和一张你自己挑选的图片)，直观理解秩的大小对矩阵近似效果的影响。

提示：利用投影  $\hat{\mathbf{x}}_{test} = \mathbf{U}\mathbf{U}^T\mathbf{x}_{test}$  . 用学习到的特征脸来估计测试集中的数据，结果如图所示。注意这里的测试图片是不属于训练集的。

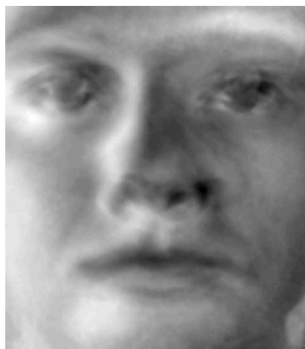


尝试用其他类型的图片进行测试，观察效果。

Test image



$r = 25$



$r = 50$



$r = 100$



$r = 200$



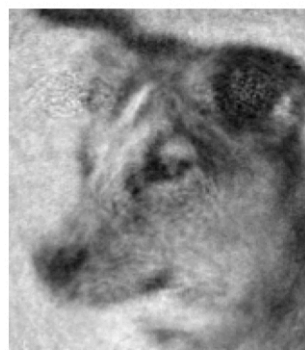
$r = 400$



$r = 800$



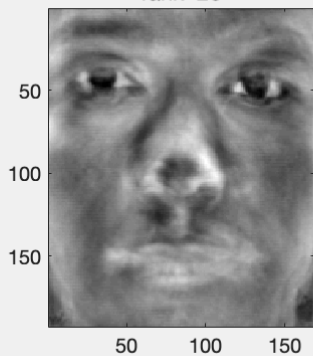
$r = 1600$



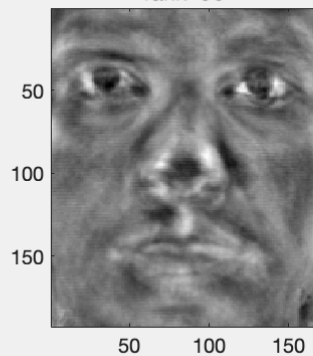
Test image



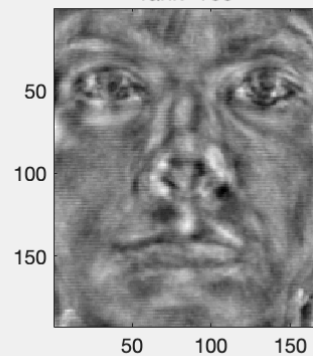
rank=25



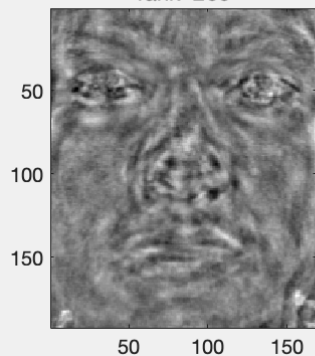
rank=50



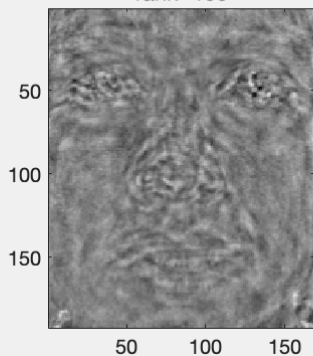
rank=100



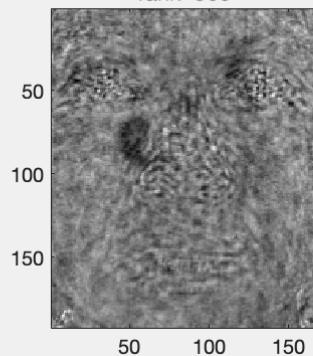
rank=200



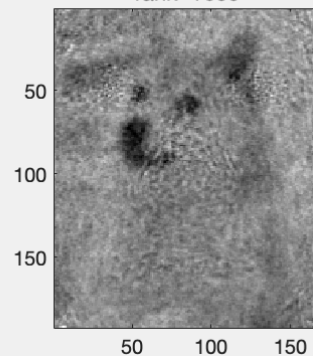
rank=400



rank=800



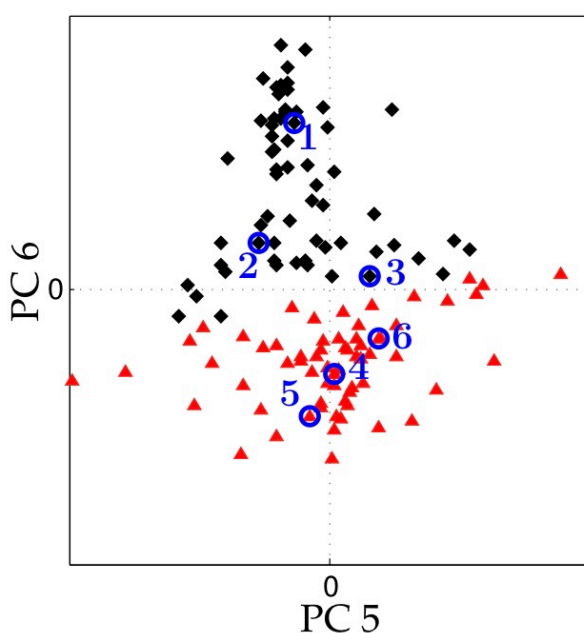
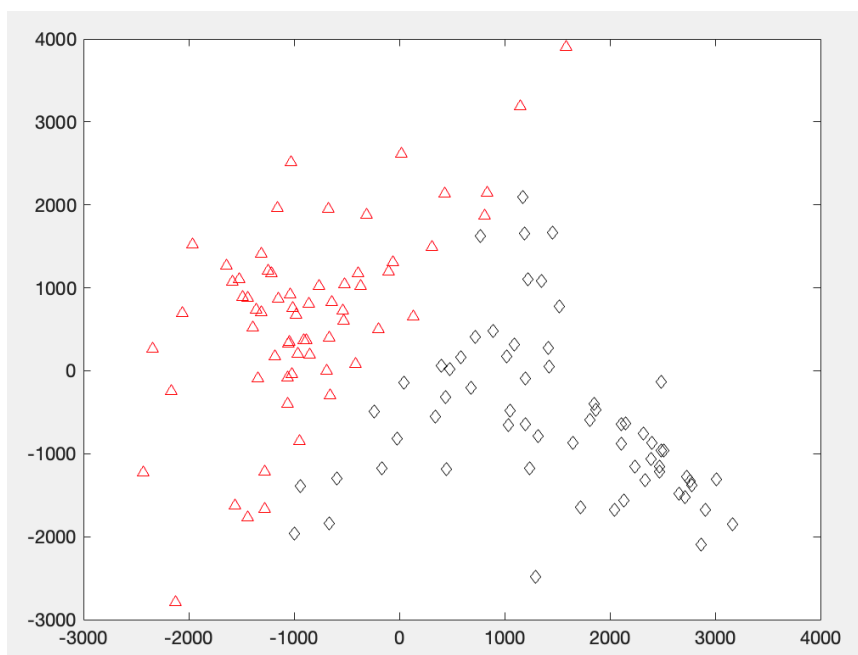
rank=1600



## 人脸识别及分类预处理

任务四：将两个不同的人的脸投影到特征空间并可视化(可以参考下面的示例)，说明如何将PCA应用于人脸识别或分类任务。

提示：尝试用特征脸来作为坐标系统，构成一个特征脸空间。将不同的人脸投影到这个特征空间，观察它们的特点。这里将第2个人和第7个人的全部64张脸投影到第5和第6个主成分上，得到二维的特征空间，如图所示。



## 附加题

此任务是加分题，选做。在前面的基础上，结合试验结果观察分析PCA的特点。思考能否找到一种比PCA更适合于图像分类的降维方法，并说明原因。

## 说明

1. 数据集导入之后会得到一个元胞，一共有38个元素，对应38个人的所有数据。每个元素里面都是一个矩阵，对应这一个人的所有数据。矩阵的每一个列向量都是一幅人脸拉长之后形成的向量( $32256 = 168 \times 192$ )。
2. 每个人的数据数量可能会不一样，也就是矩阵的列的数目不一样(不是标准的Extended Yale B数据集)。使用的时候先数一数每个人有多少。按照数据集的划分方式(即前36个人的所有数据作为训练数据)，最后得到的训练集数据矩阵 $\mathbf{X}$ 大小为 $32256 \times 2286$ 。
3. 去均值的操作是对训练集的数据 $\mathbf{X}$ 进行的。后续测试时测试集中的数据也要减去同样的均值。另外因为数据矩阵比较大，所以SVD分解时要用'econ'。
4. 绘制人脸图像的函数用`imagesc`，设置`colormap`为 `gray`。示例：

```
imagesc(reshape(X(:,1),192,168)), colormap gray
```