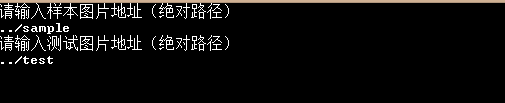
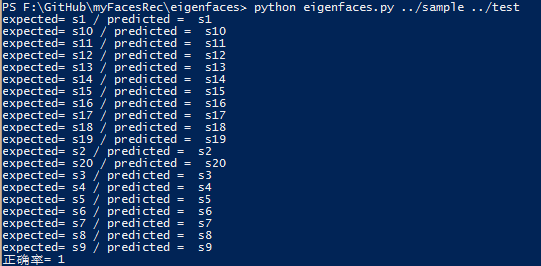
# 人脸识别程序

## 简介：

本程序是一个使用python2.7编写的简单人脸识别程序，主要使用了PCA方法，识别率能达到90%以上。

## 使用方法：

1. 直接启动。  
   本程序可以直接启动，启动后会进入命令行界面，按照提示依次输入所需参数即可。  
   
2. 使用命令行参数启动。  
   本程序还可以使用命令行传递参数的方式启动。  
   请将命令行工作路径切换至程序目录后用以下命令启动：  
   python eigenfaces.py 样本目录 测试目录。  
   例如：  
   
3. 启动程序后，本程序会自动读取样本目录下的图片信息，通过PCA方法计算出初始模型。然后通过此模型识别测试目录中的图片并输出识别结果。  
   

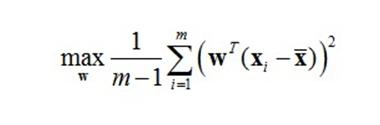
# 算法说明：

## PCA(Principal components analysis,主成分分析)算法

Principal Component Analysis(PCA)是最常用的线性降维方法，它的目标是通过某种线性投影，将高维的数据映射到低维的空间中表示，并期望在所投影的维度上数据的方差最大，以此使用较少的数据维度，同时保留住较多的原数据点的特性。

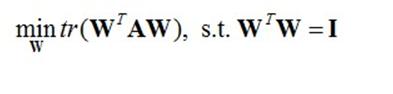
通俗的理解，如果把所有的点都映射到一起，那么几乎所有的信息（如点和点之间的距离关系）都丢失了，而如果映射后方差尽可能的大，那么数据点则会分散开来，以此来保留更多的信息。可以证明，PCA是丢失原始数据信息最少的一种线性降维方式。（实际上就是最接近原始数据，但是PCA并不试图去探索数据内在结构）

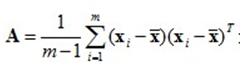
设n维向量w为目标子空间的一个坐标轴方向（称为映射向量），最大化数据映射后的方差，有：



其中m是数据实例的个数， xi是数据实例i的向量表达， x拔是所有数据实例的平均向量。定义W为包含所有映射向量为列向量的矩阵，经过线性代数变换，可以得到如下优化目标函数：

   W'W=I是说希望结果的每一个feature都正交，这样每一维度之间不会有冗余信息。



 其中tr表示矩阵的迹，A是数据协方差矩阵。

容易得到最优的W是由数据协方差矩阵前k个最大的特征值对应的特征向量作为列向量构成的。这些特征向量形成一组正交基并且最好地保留了数据中的信息。

# 源码说明：

## 综述：

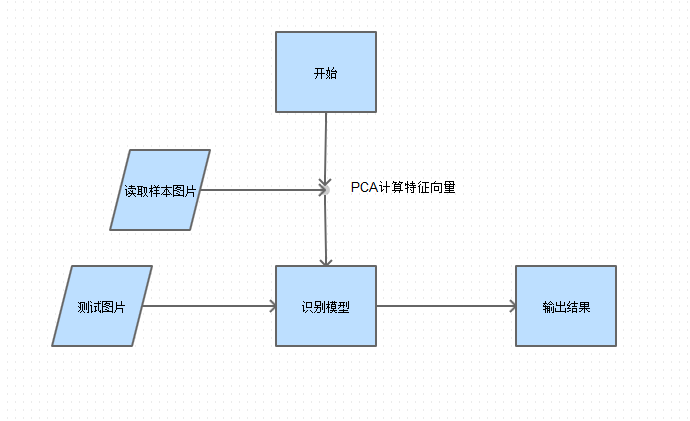
本程序参考了Philipp Wagner(bytefish)发布在github上的教程facerecognition\_guide[2]，并在原程序的基础上进行了精简与改动（主要是提取eigenface有关的部分并将其细节部分简化而将其余部分舍弃）。

本程序使用python2.7编写，使用了PIL、numpy、opencv以及python标准库，实现了自动图片读取与PCA特征向量的计算，并通过此特征向量对新图片进行识别。

本程序所使用的人脸图片库来自于AT&T Laboratories, Cambridge.[3]感谢他们的无偿共享。

## 工作流程：

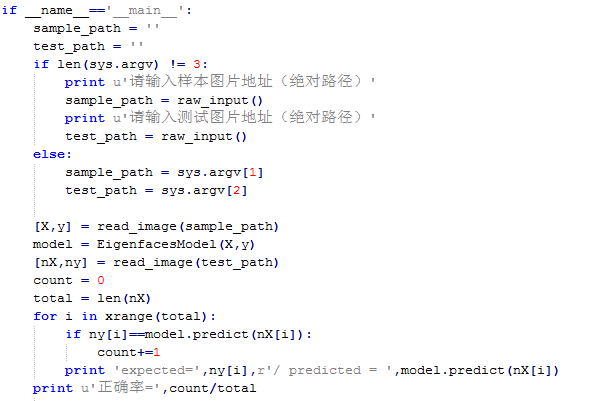
流程图



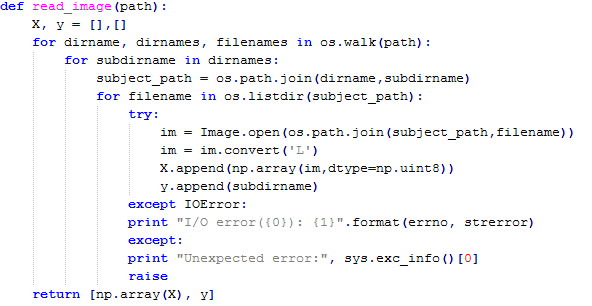
## 程序各部分说明：

本程序主要分为四个部分：eigenfaces.py model.py util.py distance.py.

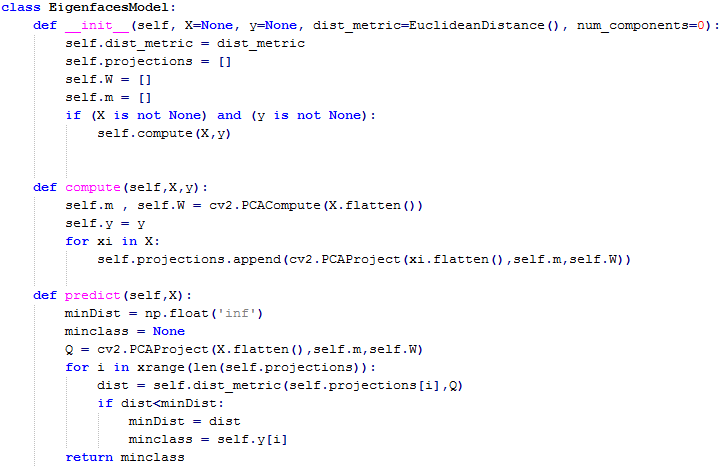
### Eigenfaces.py

这是程序的主要运行模块，定义了程序的运行逻辑，即：获取原始样本与测试样本，使用原始样本计算出特征向量，再用此特征向量识别测试样本，最后输出结果。

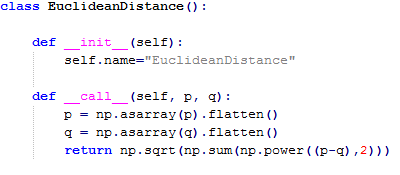
### Util.py

此模块定义了程序中的常用函数，精简后只留下了读图函数read\_image.

### Model.py

此模块由上面提到的教程中同名模块精简并将主要运算步骤改用opencv实现后获得。只包含了一个EigenfacesModel类，用处是：使用PCA方法计算并存储各图片的特征向量并定义新图片的测试方法：计算新图片的特征向量并计算与之欧氏距离最短的特征向量，取得判定结果。

### Distance.py



此模块定义了距离的计算方法，留下了欧氏距离。

# 总结：

此程序本质上其实是使用最近邻原则配合PCA计算特征向量的方法对图片进行分类，所识别的图片是不是人脸都无所谓。此外，此程序只适用于识别经过预处理提取出人脸的图片，还没有从图片中找出人脸进行识别的功能。

# 改进：

1. 此程序可以配合facedetect程序从图片中提取人脸进行识别。
2. 此程序中测试模型一旦计算完毕就不再改变了，针对此点可作如下改进：  
   将识别出的人脸进行标记后放入训练样本中重新计算模型，达到不断改进模型的目的。

注：

[1]本程序发布于https://github.com/zwy1135/myFacesRec

[2]所参考的facerecognition\_guide地址<https://github.com/bytefish/facerecognition_guide>

[3]图片库下载地址http://www.cl.cam.ac.uk/research/dtg/attarchive/pub/data/att\_faces.zip