**10.6 试使用MATLAB中的PCA函数对Yale人脸数据集进行降维，并观察前20个特征向量所对应的图像。**

**①问题描述：**

本题要求使用Matlab中的PCA函数对Yale人脸数据集进行降维，并使用前20个特征向量将原始数据的维度降至20，并观察降维后所对应的图像分析其特点。PCA的主要思想是将n维特征映射到k维上，这k维是全新的正交特征也被称为主成分，是在原有n维特征的基础上重新构造出来的k维特征。PCA的工作就是从原始的空间中顺序地找一组相互正交的坐标轴，新的坐标轴的选择与数据本身是密切相关的。其中：第一个新坐标轴选择是原始数据中方差最大的方向，第二个新坐标轴选取是与第一个坐标轴正交的平面中使得方差最大的，第三个轴是与第1,2个轴正交的平面中方差最大的。依次类推，可以得到n个这样的坐标轴，即n维空间内的坐标轴。本题中系将原始的Yale人脸数据集降维至20维。

**②实现过程：**

a.读取YaleFaceDatabase文件夹内的数据集，并将165张图片合并为一个243\*320行，165列的数据集M。

b.使用Matlab的pca函数处理数据集M，得到主成分系数F。

c.因为需要降维至20维，因此仅提取主成分系数中的前20列，其他列置为0。

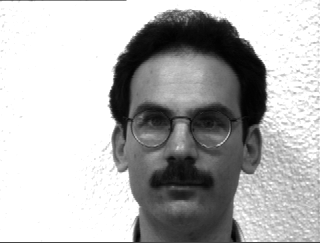
d.使用前20个特征还原数据集M得到20维下的图像数据。

e.将20维提取后的图像数据存入result文件夹。

**③实现结果：**

165维下的图像：







20维下的图像：







**④结果分析：**

降维后的图像比原始图像更加模糊，仅能突出其部分特征。人脸的眼睛、鼻子、嘴巴、头发等部分不再具有鲜明的特点，而是仅表征出不同人所共有的特征。戴着眼镜的人的图像数据在降维后，眼镜的特征较为弱化，在部分样本中甚至无法在降维后的数据中肉眼识别出眼镜的存在。若增加降维后的维数，则降维后的图像将与原始图像越来越接近，通过实验发现当维数增加值150左右时，所选特征的累计方差占比已经很接近于1了，可以用大约150维的向量来描述一张原本维数巨大的多维人脸图像。