

图 10-7 中的第一张图片是第一个特征脸，第二张图片是用了 40 个特征脸的线性组合，之后依次是用了 80、120、160、…、360 个特征脸的线性组合，最后一张图片用到了全部 400 个特征脸的线性组合。

你会发现，其实不必把 400 个特征脸全部用上，只需要 80 ~ 120 个特征脸，也就是在第 3、4 轮时就已经能够识别出这个人了。

通过这个例子，我们直观地体会到了 PCA 在数据降维、消除特征相关性上的强大。接下来就从原理上了解一下 PCA 背后的数学思想。

10.4 专家解读

PCA 是应用层面的叫法，它背后依赖的数学原理是线性代数中的矩阵分解。

线性代数相关教程中的矩阵分解方法非常多，比如 LU、LDU 等。在工程中最常用的主要有两个，一个是对称方阵的正交分解，还有一个是一般矩阵的 SVD 分解。PCA 依赖的是前者，SVD 是研究生阶段的学习内容。

在线性代数中，一个实对称矩阵能够做很好的分解——正交分解：

$$A = V \Lambda V^T$$

其中：

- V 是单位正交矩阵，即 $VV^T = V^TV = I$ ， $V^T = V^{-1}$ ；
- Λ 是个对角矩阵，对角线上的元素是矩阵 A 的特征值。

单位正交矩阵 V 有以下几个特点：

- 它肯定是个方阵，行数等于列数；