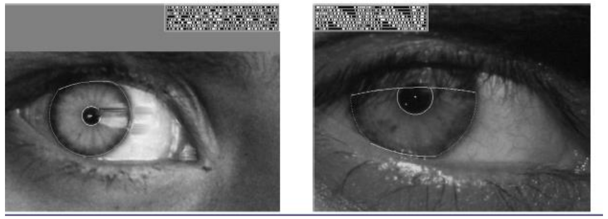
# 图像模式分类

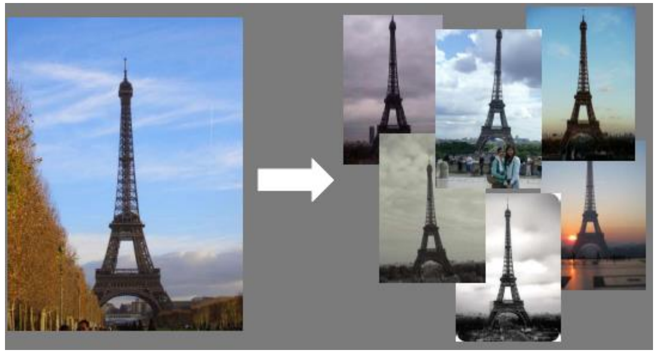
数字图像处理 第10讲

# 图像识别

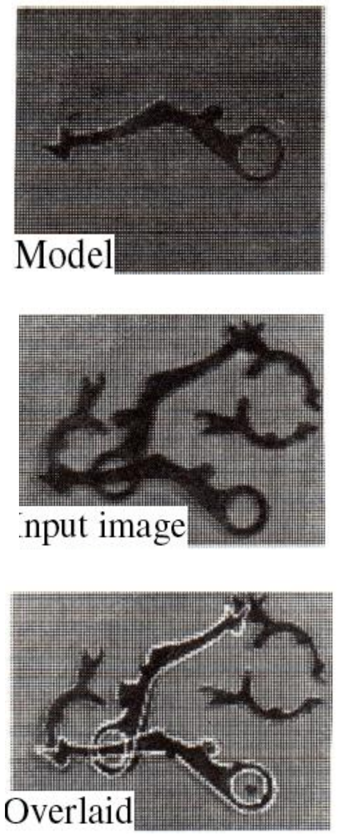
- 利用计算机对图像进行处理和分析，识别不同模式的目标的技术



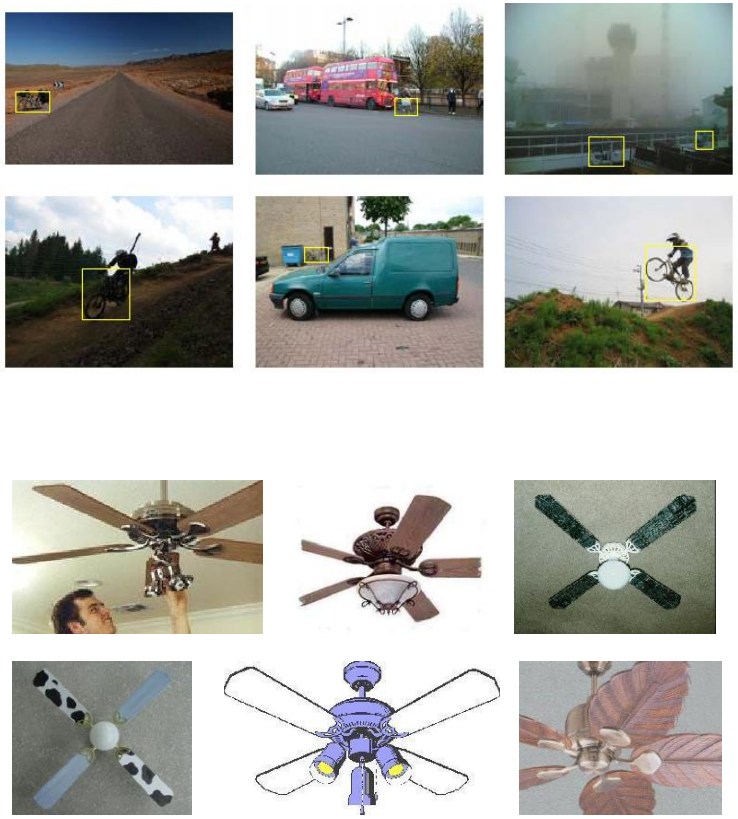


# 图像识别的发展趋势




1980s

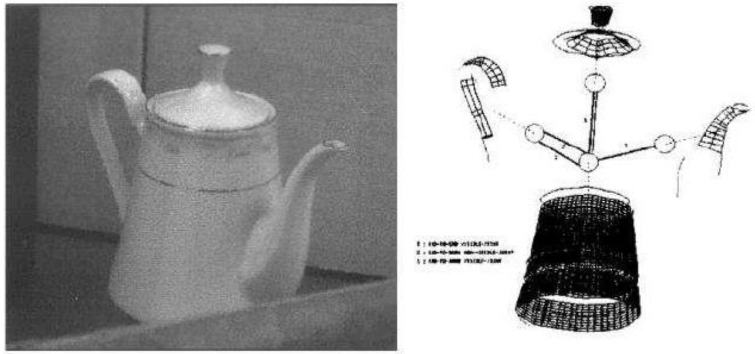
1990s to early 2000s

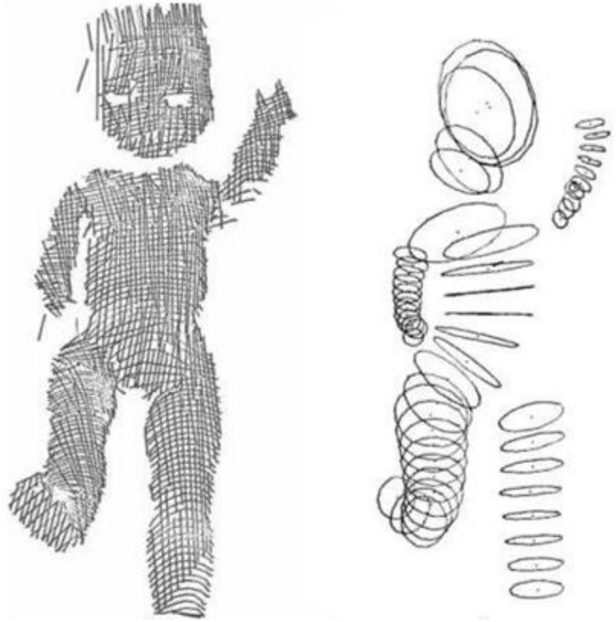


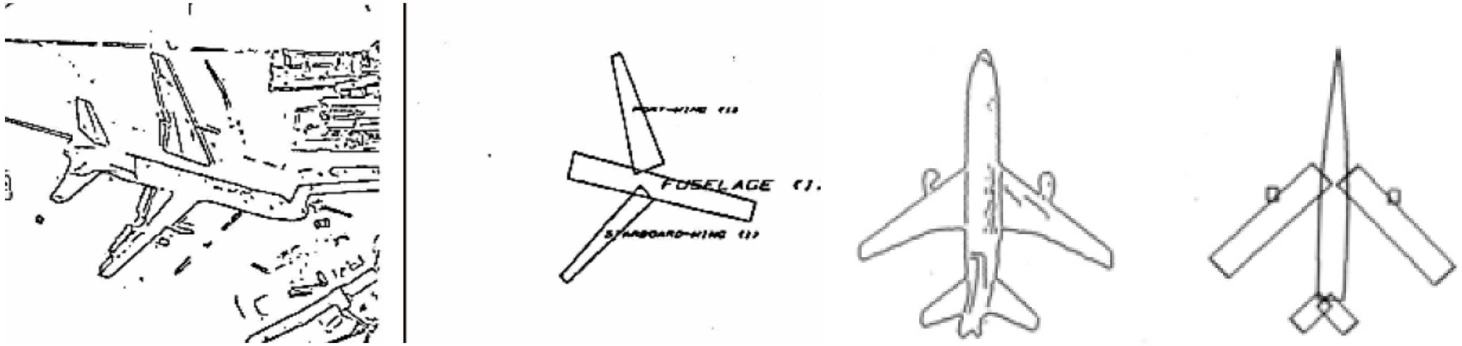
Currently

# 早期识别方法 - 基于几何模型

- Geometry based
  - 3维模型建模

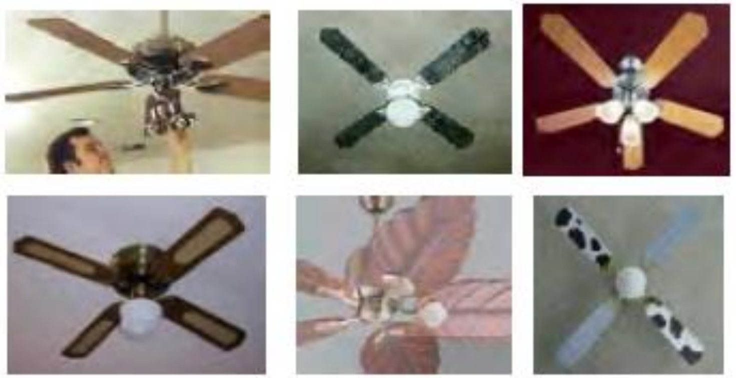






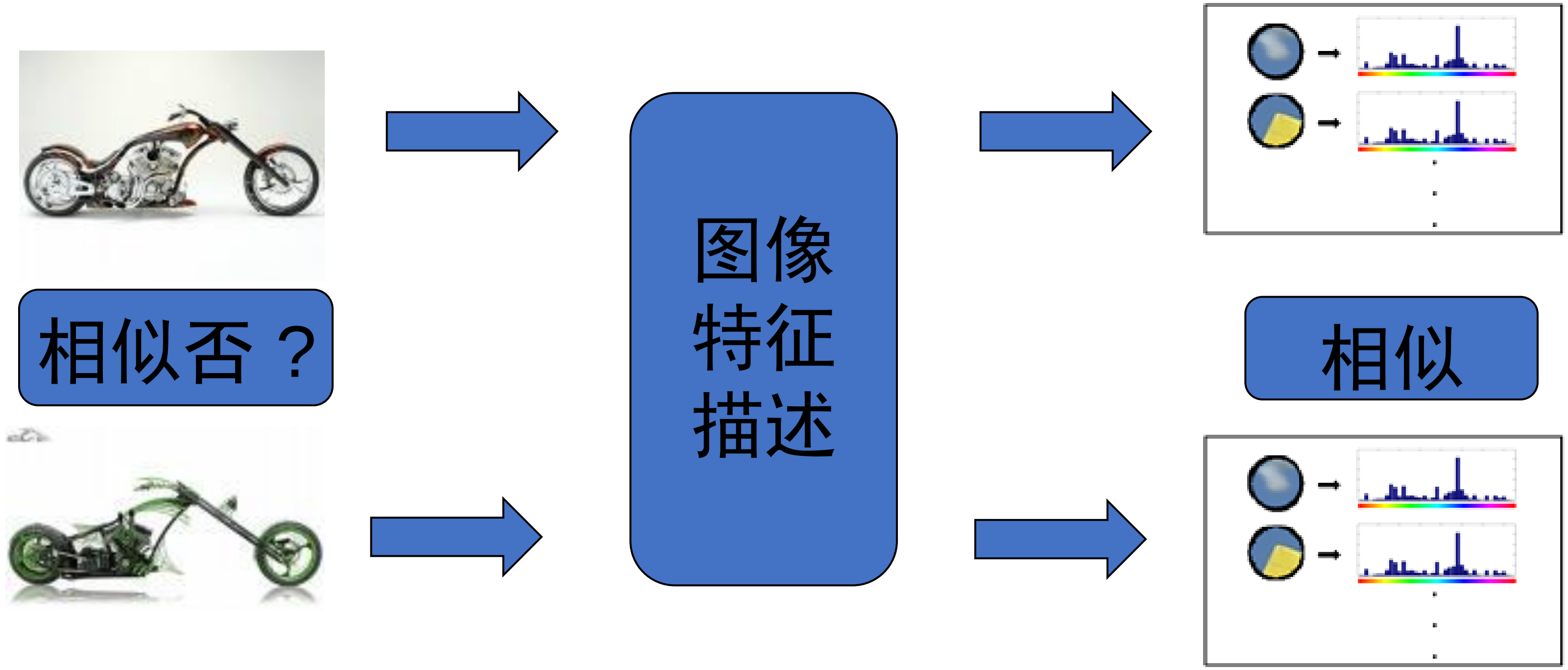
# 当前发展 - 基于表观模型

- Appearance based
  - 直接从图像出发



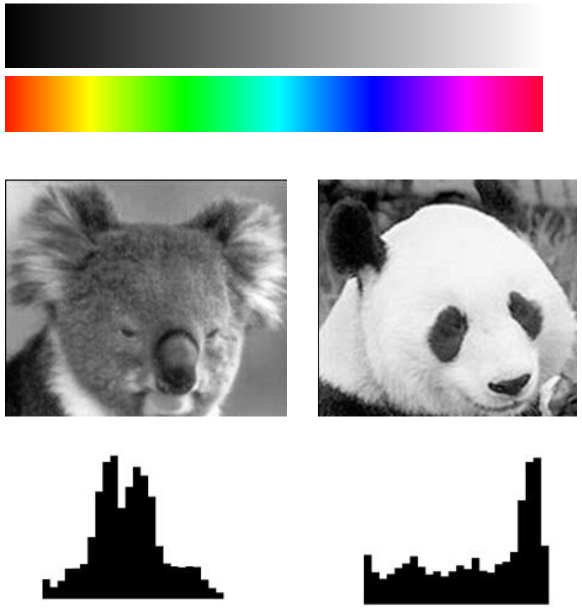
# 图像特征

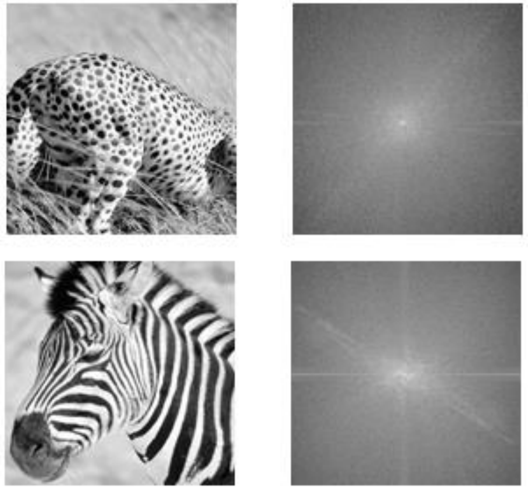
- 目的：对图像进行紧凑和有效地描述
  - 从图像到一个数值或向量的变换
  - 加入主观认识、抓住目标本质
  - 效率提高，处理速度更快

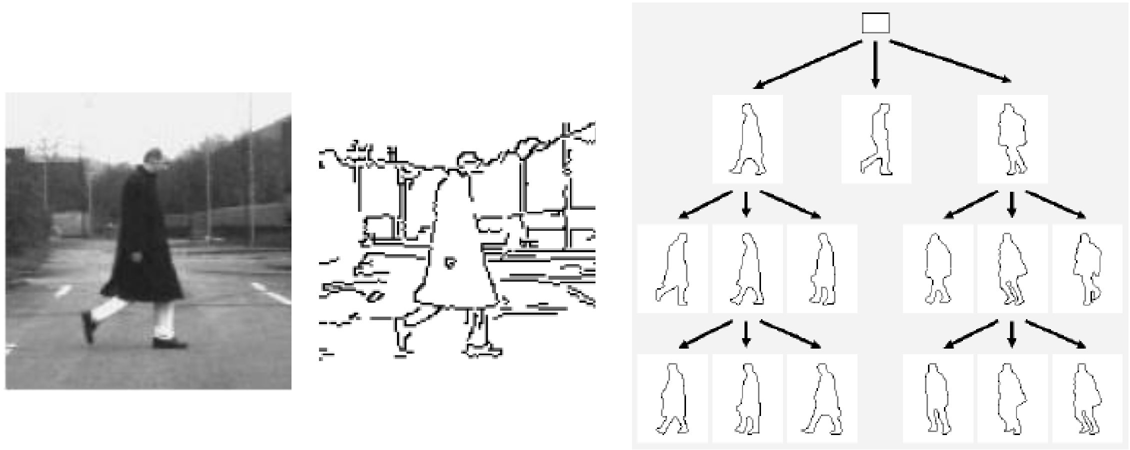


# 图像特征类型

- 颜色： 直方图
- 形状： 边缘、轮廓
- 频域： 傅里叶变换
- 特征点： 角点、SIFT







# 图像分类器

- 构造一个识别函数(分类器)
  - 输入是图片中提取的特征，输出是类别标签
  - 该函数通过已知类别标签的图片集训练得到

$$y = f(\mathbf{x})$$

output　prediction function　Image feature

f(  ) = “apple”

f(  ) = “tomato”

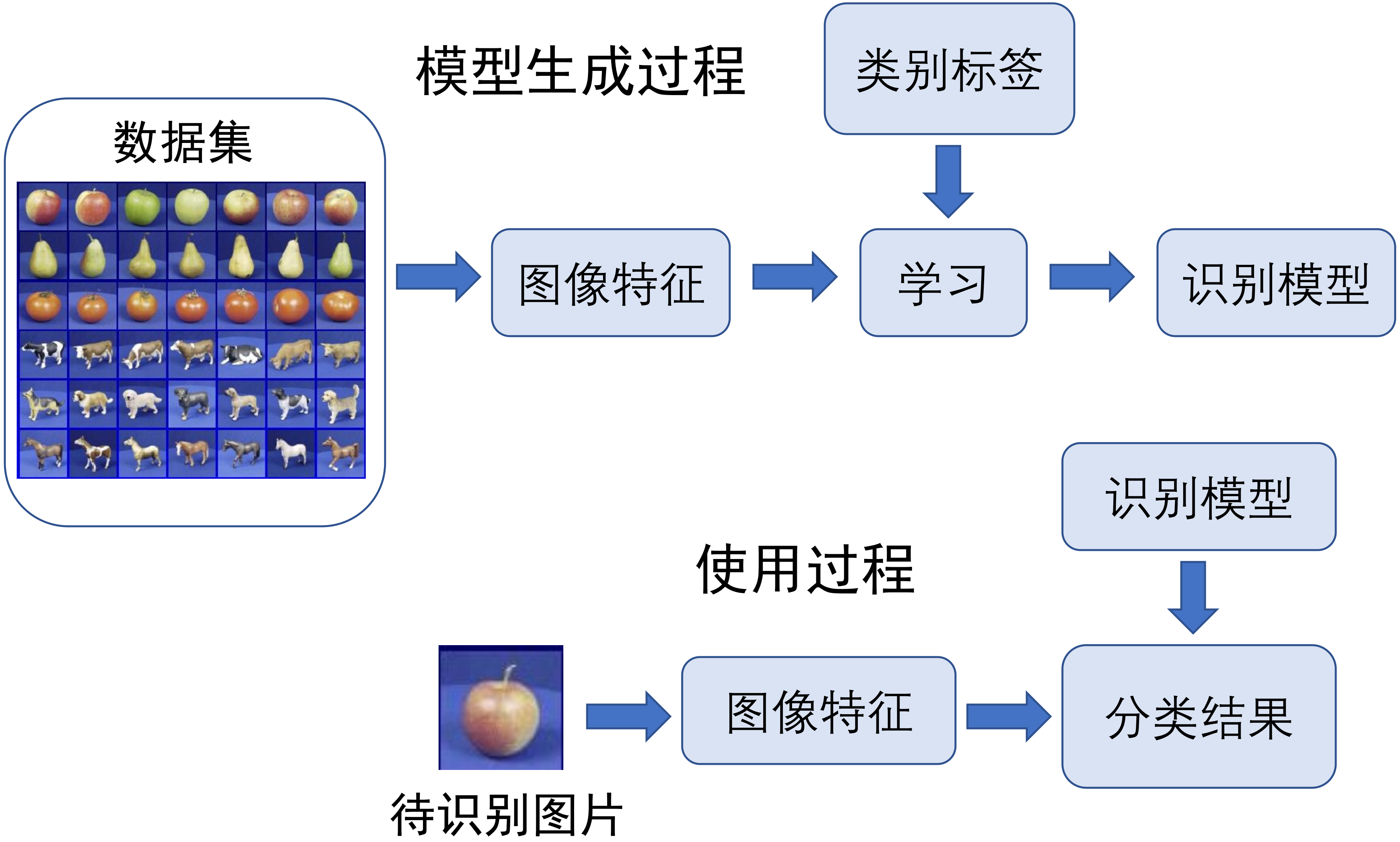
f(  ) = “cow”

# 训练集与测试集

训练集
(类别标签用于分类器的构建)

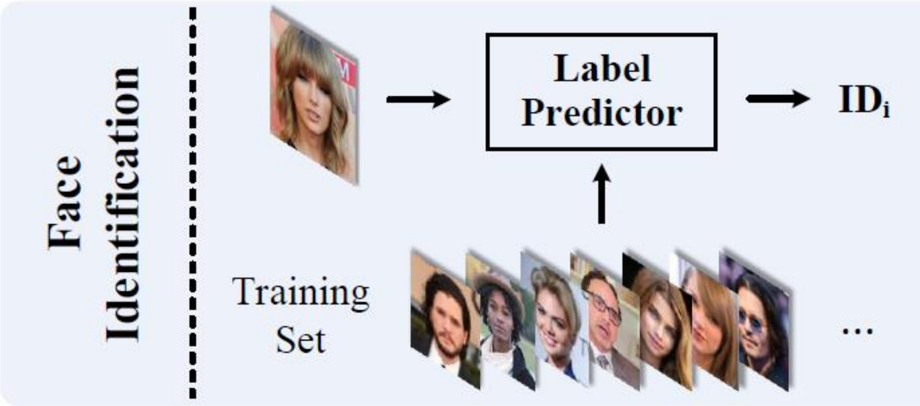
测试集
(类别标签仅仅用于分类器
测试结果的比对)

# 图像识别系统



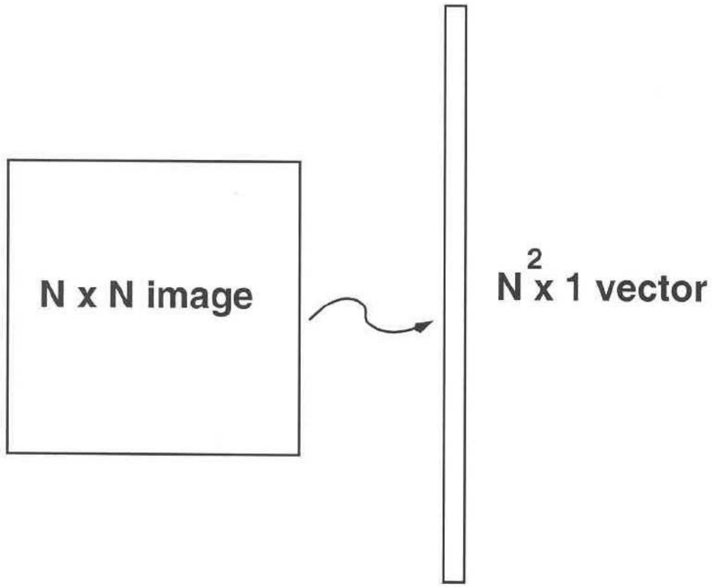
# 人脸识别问题描述

- 训练：给定一组人脸图像及其类别标号
- 测试：对于一幅新图片，判断其标号

# 特征数量往往是高维的

- 图像通常首先被向量化

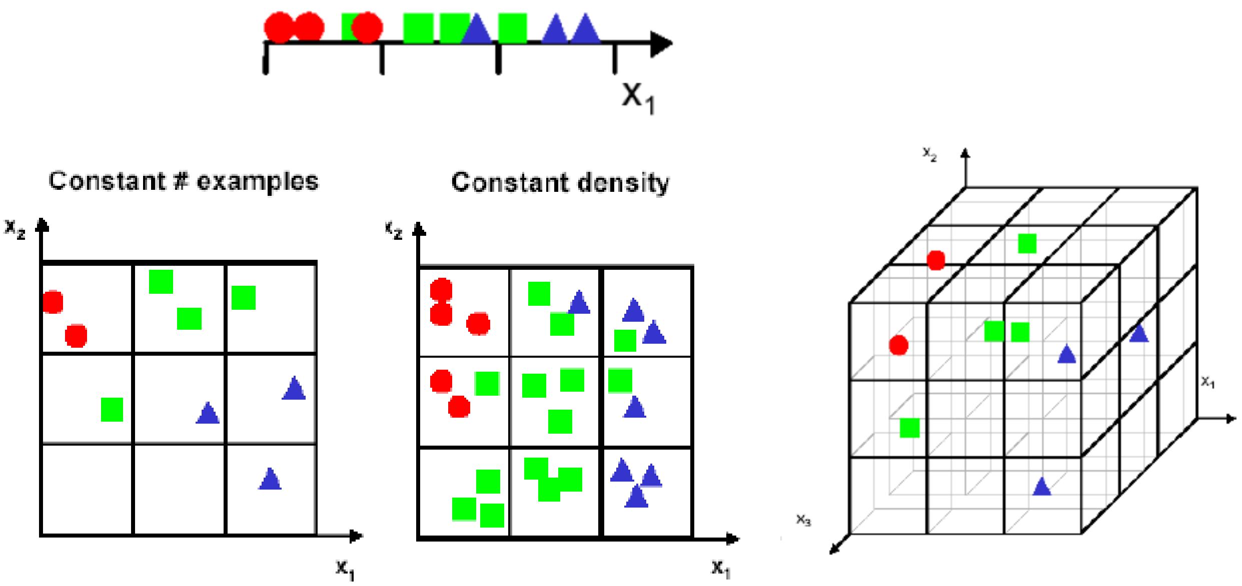


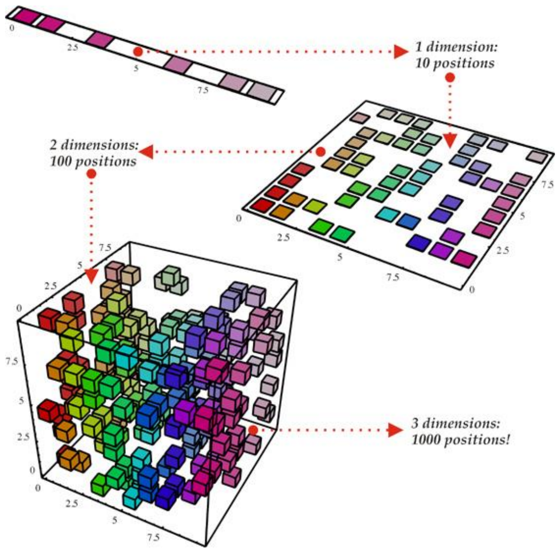

- 向量维度往往过高
  - 256 x 256 图像，亮度特征向量 65636维
- 带来的弊端
  - 计算量大增
  - 维度灾难

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

# 维度灾难

- Curse of Dimensionality
- 随着维度增加，样本空间的体积指数增长
- 数据稀疏，统计描述没有意义

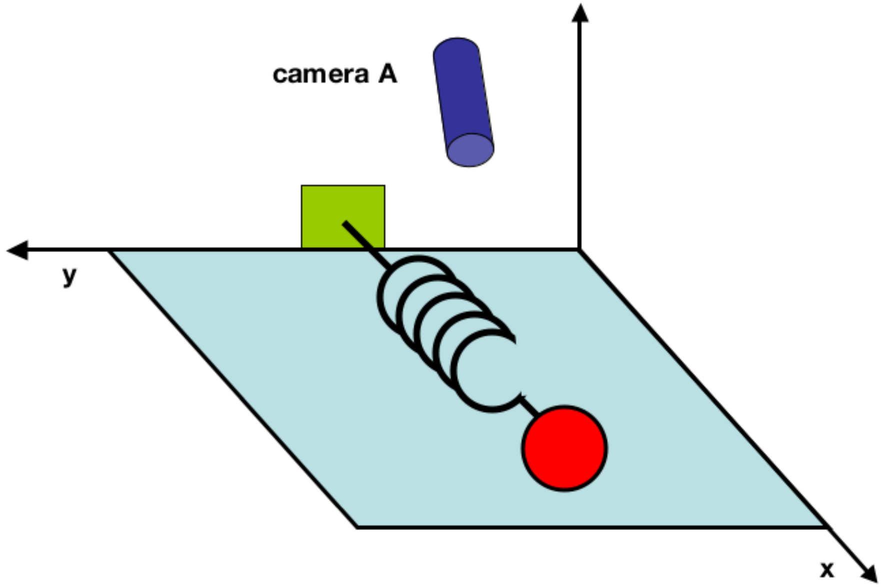


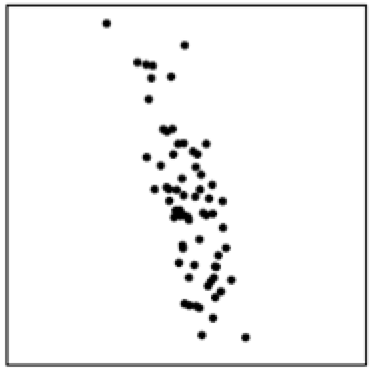



# 数据存在本征维度

- Intrinsic Dimensionality

$$x = Acos\left(\sqrt{\frac{k}{m}}t + t_0\right)$$



# 降低维度的方法

- 特征选择 Feature Selection
  - 选择一个子集
  - 排列组合多，寻优慢

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_N \end{bmatrix} \xrightarrow{\text{feature selection}} \begin{bmatrix} x_{i_1} \\ x_{i_2} \\ \\ x_{i_M} \end{bmatrix}$$

- 特征抽取 Feature Extraction
  - 将原有特征进行变换

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_N \end{bmatrix} \xrightarrow{\text{feature extraction}} \begin{bmatrix} y_1 \\ y_2 \\ \\ y_M \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M1} & w_{M2} & & w_{MN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_N \end{bmatrix}$$

# 主成分分析

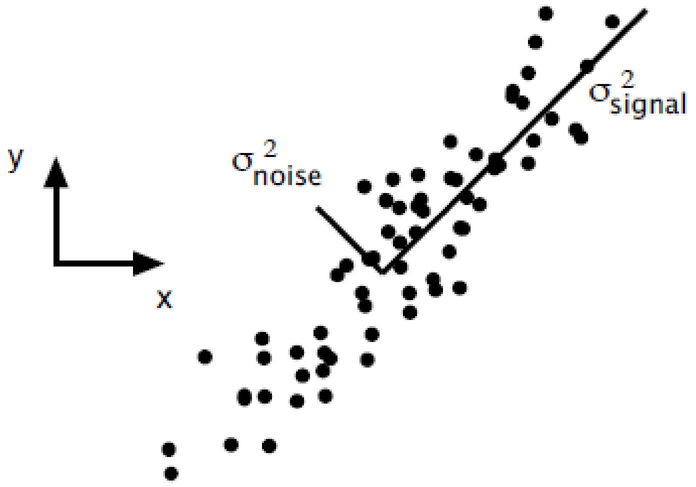
- Principal Component Analysis (PCA)
- Karl Pearson，1901年提出

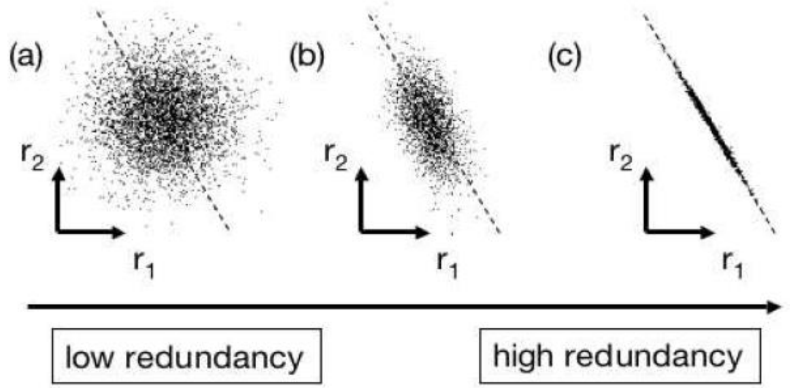
- 特点
  - 对原特征的进行线性变换
  - 特征之间互不相关，减少原有特征之间的冗余
  - 得到一组按重要性从大到小排列的新特征

- 保留其中最重要的那些特征

# PCA的主要思想

- 数据的特征空间中
  - 方差大的方向包含了信号本身
  - 方差小的方向包含了噪声
- 如果两个维度是相关性大
  - 相近的观测信息，有冗余
- 目标
  - 依次寻找方差大，且互不相关的维度（方向）
  - 一组新的基来描述数据

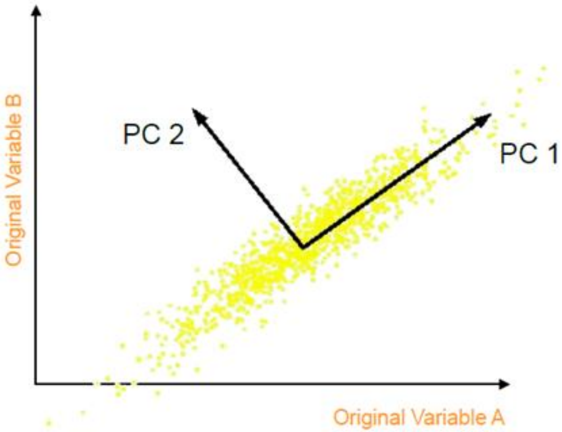


# PCA的求解思路

- 找到一组新的正交基向量来描述数据
  - 方差最大的方向作为第1个主方向
  - 和第1个方向正交的方向中，方差最大的，为第2个
  - 这两个方向的相关系数为0
  - 第3个方向和第1，2方向正交

$$\mathbf{y} = \mathbf{W}^T\mathbf{x}$$




$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \rightarrow \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N|2} & & w_{NN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

# 协方差矩阵

- 方差 $$\sigma_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$ 协方差 $$\sigma(x, y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$
- 对于d维随机向量 $$\sigma(x_k, x_k) = \frac{1}{n-1}\sum_{i=1}^{n}(x_{ki} - \bar{x}_k)^2, k = 1, 2, \ldots, d$$
  $x \in R^{\mathrm{d}}$ $$\sigma(x_m, x_k) = \frac{1}{n-1}\sum_{i=1}^{n}(x_{mi} - \bar{x}_m)(x_{ki} - \bar{x}_k)$$
- d维随机向量x的协方差矩阵
  $$\Sigma = \begin{bmatrix} \sigma(x_1, x_1) & \cdots & \sigma(x_1, x_d) \\ \vdots & \ddots & \vdots \\ \sigma(x_d, x_1) & \cdots & \sigma(x_d, x_d) \end{bmatrix} \in \mathbb{R}^{d \times d}$$
- 需要找一个变换，让新坐标系下的协方差矩阵对角化

# 通过特征值分解求解

- 对于一个0均值化数据集 $\{x_1, x_2, \cdots x_n\}, x_i \in R^D$
  - 数据矩阵写为

$$X = [x_1, \ldots, x_n] \subset R^{D \times n}$$

  - 对应的协方差矩阵

$$\Sigma = \frac{1}{n-1} XX^T \in R^{D \times D}$$

  - 计算其特征值分解

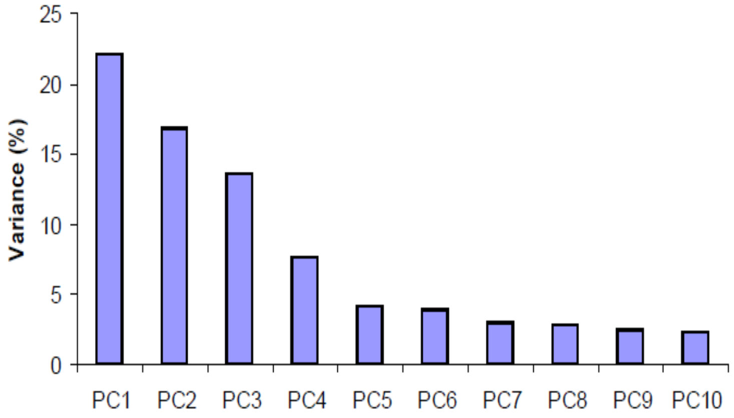
$$\sum = U \Lambda U^{-1} = U \Lambda U^T$$

  - 则其主成分方向为

$$U^T = [u_1, \ldots, u_D]^T \in R^{D \times D}$$

# 保留主成分来得到最有效特征

- 保留主成分 – 大特征值（方差）对应的方向
- 忽略那些次要的成分 - 特征值小的那些方向




$$U^T = [u_1, \ldots, u_D]^T \in R^{D \times D}$$

$$\tilde{U} = [u_1, \ldots, u_d]^T \in R^{d \times D} \quad d < D$$

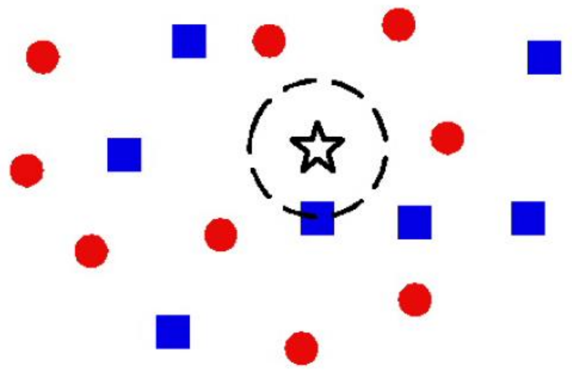
- 将数据投影到主成分上，得到样本的低维描述

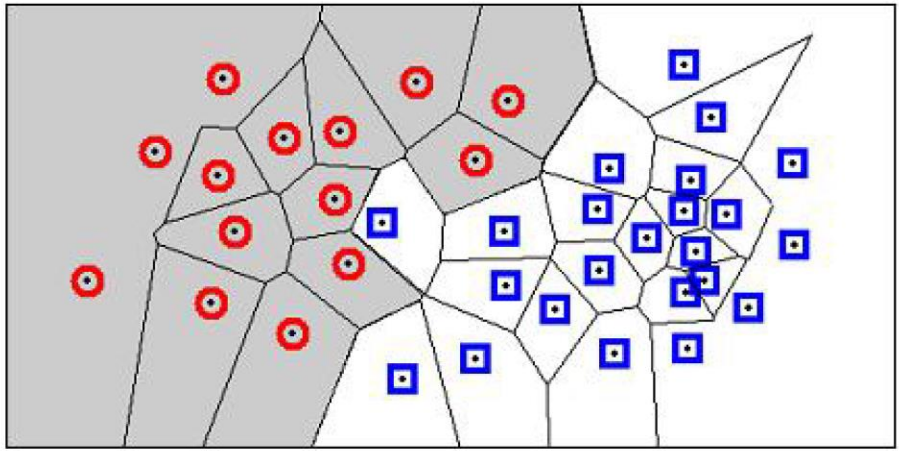
$$\tilde{Y} = \tilde{U} X$$

# 最近邻分类器

- 直观：离谁最近，就和谁的类别一样
- 数学描述：

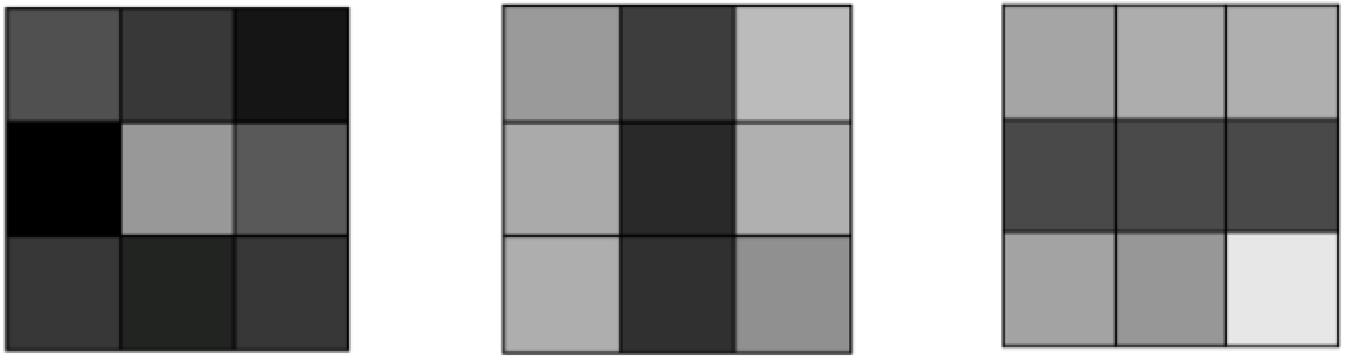
$$\mathbf{x} \to \theta(\mathbf{x}') \qquad \mathbf{x}' = \arg\min_{\mathbf{x}_i \in \mathcal{D}_n} \ \mathtt{dist}(\mathbf{x}, \mathbf{x}_i)$$





# PCA用于图像分类

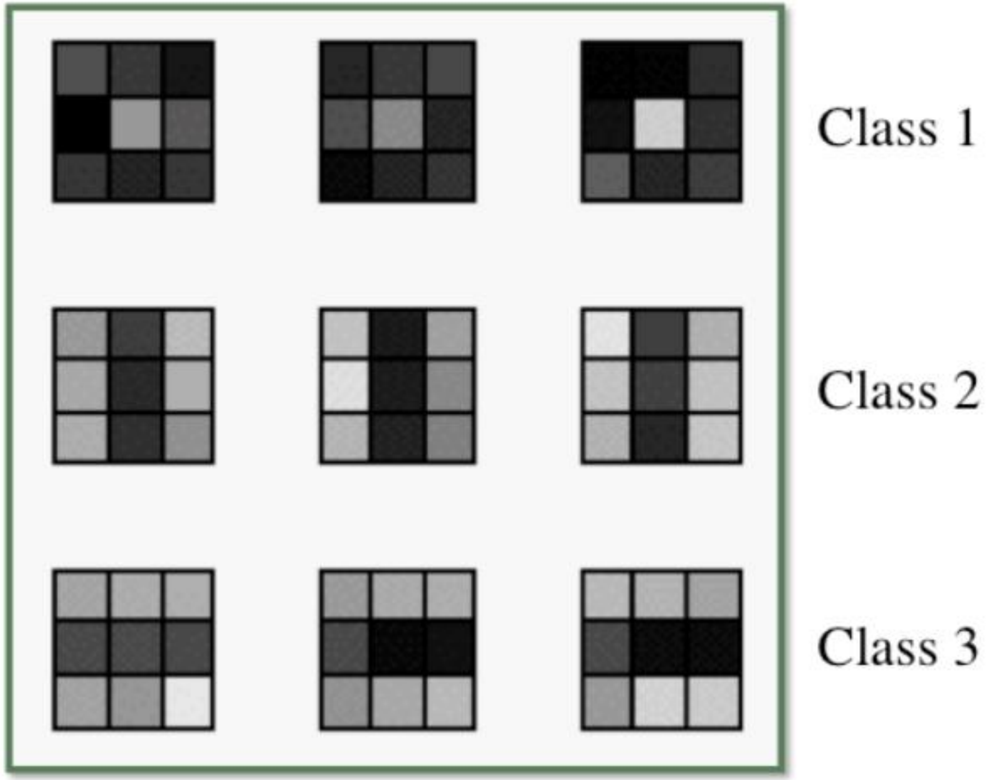
- 图像尺寸 3X3 = 9维
- 三类图像
  - Class1 中心浅，四周深
  - Class2 有竖直深条纹
  - Class3 有水平深条纹



Class 1 Class 2 Class 3

# PCA用于图像降维

- 每类3个样本



# PCA用于图像降维

- 9幅图像对应的像素值

$$\left[\begin{bmatrix}1.65 & 3.11 & 2.25\\ 3.22 & 5.79 & 3.09\\ 1.10 & 2.47 & 2.96\end{bmatrix},\begin{bmatrix}1.55 & 3.29 & 1.62\\ 2.91 & 3.88 & .71\\ 2.35 & 3.60 & 2.46\end{bmatrix},\begin{bmatrix}.80 & 2.43 & 2.04\\ 1.59 & 8.17 & .79\\ .69 & 1.96 & 4.34\end{bmatrix},\right.$$

$$\begin{bmatrix}6.36 & 2.39 & 9.36\\ 6.05 & .55 & 6.60\\ 5.97 & 3.49 & 7.33\end{bmatrix},\begin{bmatrix}6.43 & 1.43 & 7.01\\ 7.66 & 3.20 & 6.66\\ 6.96 & 1.82 & 7.52\end{bmatrix},\begin{bmatrix}6.52 & .89 & 7.74\\ 4.80 & 1.97 & 7.58\\ 5.75 & 1.06 & 7.24\end{bmatrix},$$

$$\left.\begin{bmatrix}8.11 & 8.94 & 5.85\\ 2.63 & 2.60 & 5.16\\ 7.20 & 6.09 & 6.12\end{bmatrix},\begin{bmatrix}6.94 & 6.68 & 5.99\\ 3.63 & 3.15 & 1.37\\ 8.50 & 6.89 & 6.49\end{bmatrix},\begin{bmatrix}7.02 & 7.73 & 7.08\\ 2.75 & 2.10 & 1.91\\ 5.92 & 6.85 & 7.16\end{bmatrix}\right]$$

# PCA用于图像降维

- 所有图像矢量化
  - 数据归1化之后，再0均值化

$$X = \left[ \begin{bmatrix} -.133 \\ .0500 \\ -.102 \\ .0750 \\ .324 \\ .0910 \\ -.188 \\ .00100 \\ -.0630 \end{bmatrix}, \begin{bmatrix} -.117 \\ .127 \\ -.141 \\ .0930 \\ .186 \\ -.152 \\ -.0140 \\ .185 \\ -.0740 \end{bmatrix}, \begin{bmatrix} -.231 \\ -.0450 \\ -.143 \\ -.114 \\ .505 \\ -.163 \\ -.239 \\ -.0710 \\ .0450 \end{bmatrix}, \begin{bmatrix} .0480 \\ -.149 \\ .184 \\ .0710 \\ -.266 \\ .131 \\ .0300 \\ -.0670 \\ .0320 \end{bmatrix}, \begin{bmatrix} .0530 \\ -.202 \\ .0520 \\ .162 \\ -.116 \\ .135 \\ .0860 \\ -.161 \\ .0440 \end{bmatrix}, \begin{bmatrix} .0840 \\ -.229 \\ .124 \\ .0200 \\ -.178 \\ .217 \\ .0410 \\ -.200 \\ .0570 \end{bmatrix}, \begin{bmatrix} .126 \\ .197 \\ -.0290 \\ -.129 \\ -.157 \\ .0370 \\ .0800 \\ .0630 \\ -.0520 \end{bmatrix}, \begin{bmatrix} .0810 \\ .0930 \\ -.00600 \\ -.0660 \\ -.120 \\ -.163 \\ .172 \\ .124 \\ -.0150 \end{bmatrix}, \begin{bmatrix} .0900 \\ .157 \\ .0600 \\ -.113 \\ -.177 \\ -.132 \\ .0310 \\ .126 \\ .0270 \end{bmatrix} \right]$$

# PCA用于图像降维

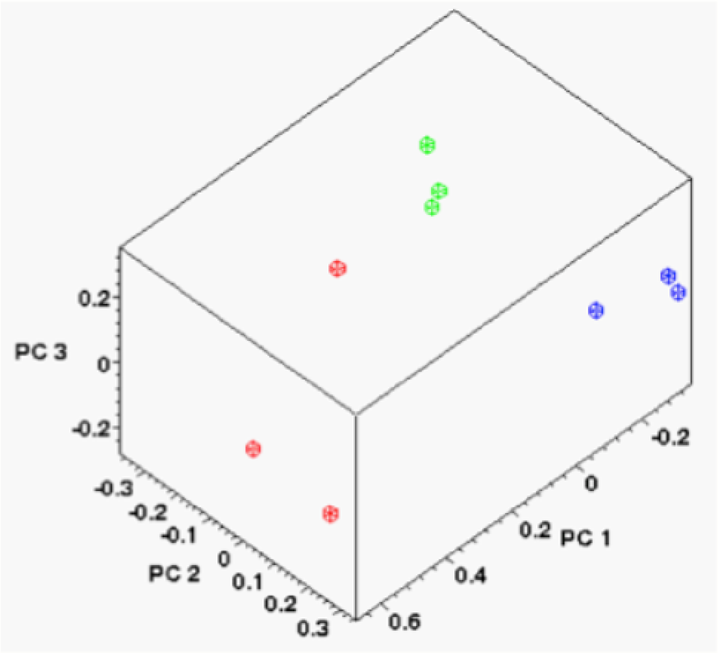
- 对协方差矩阵做特征值分解

$$\Sigma = UDU^T$$

$$\underset{U}{\begin{bmatrix} -.29 & -.080 & -.31 & -.51 & .40 & .30 & -.17 & -.15 & -.51 \\ -.14 & .46 & .10 & -.48 & -.27 & -.17 & -.46 & -.30 & .35 \\ -.23 & -.060 & .34 & .050 & -.49 & -.29 & -.18 & .16 & -.66 \\ -.23 & -.52 & .64 & -.060 & .12 & .16 & .050 & -.44 & .17 \\ .84 & .040 & .16 & -.10 & -.090 & .21 & -.12 & -.32 & -.30 \\ .080 & -.47 & -.42 & -.36 & -.45 & -.29 & .35 & -.22 & .080 \\ -.14 & -.10 & -.38 & .60 & -.080 & -.020 & -.40 & -.54 & -.050 \\ -.14 & .50 & .080 & .080 & .13 & -.21 & .62 & -.47 & -.23 \\ -.20 & .16 & -.050 & .050 & -.53 & .77 & .21 & 0 & .040 \end{bmatrix}} \underset{D}{\begin{bmatrix} 1.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .55 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .19 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .12 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .030 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .010 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}} \underset{U^T}{\begin{bmatrix} -.29 & -.14 & -.23 & -.23 & .84 & .080 & -.14 & -.14 & -.20 \\ -.080 & .46 & -.060 & -.52 & .040 & -.47 & -.10 & .50 & .16 \\ -.31 & .10 & .34 & .64 & .16 & -.42 & -.38 & .080 & -.050 \\ -.51 & -.48 & .050 & -.060 & -.10 & -.36 & .60 & .080 & .050 \\ .40 & -.27 & -.49 & .12 & -.090 & -.45 & -.080 & .13 & -.53 \\ .30 & -.17 & -.29 & .16 & .21 & -.29 & -.020 & -.21 & .77 \\ -.17 & -.46 & -.18 & .050 & -.12 & .35 & -.40 & .62 & .21 \\ -.15 & -.30 & .16 & -.44 & -.32 & -.22 & -.54 & -.47 & 0 \\ -.51 & .35 & -.66 & .17 & -.30 & .080 & -.050 & -.23 & .040 \end{bmatrix}}$$

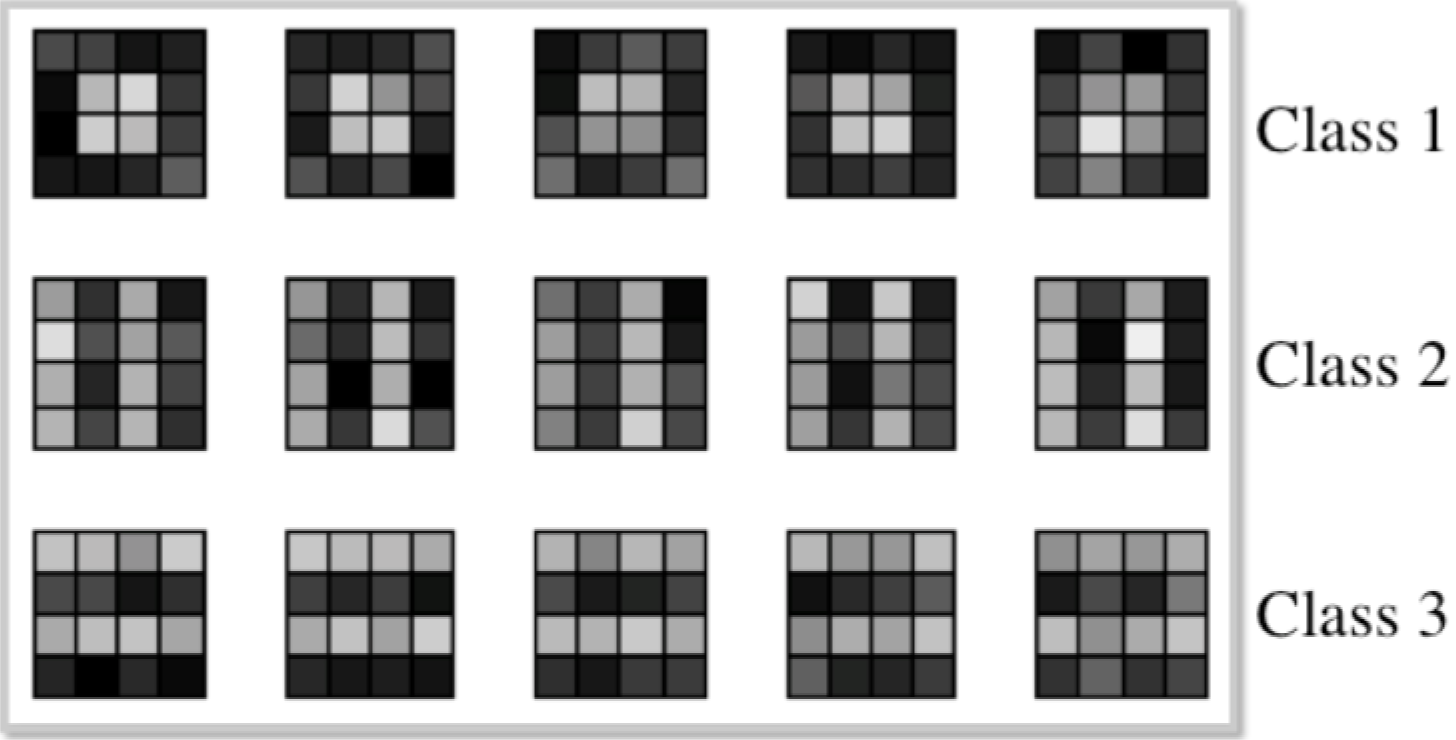
# PCA用于图像降维

- 将图像投影到前3维上
- 用最近邻法可以实现分类



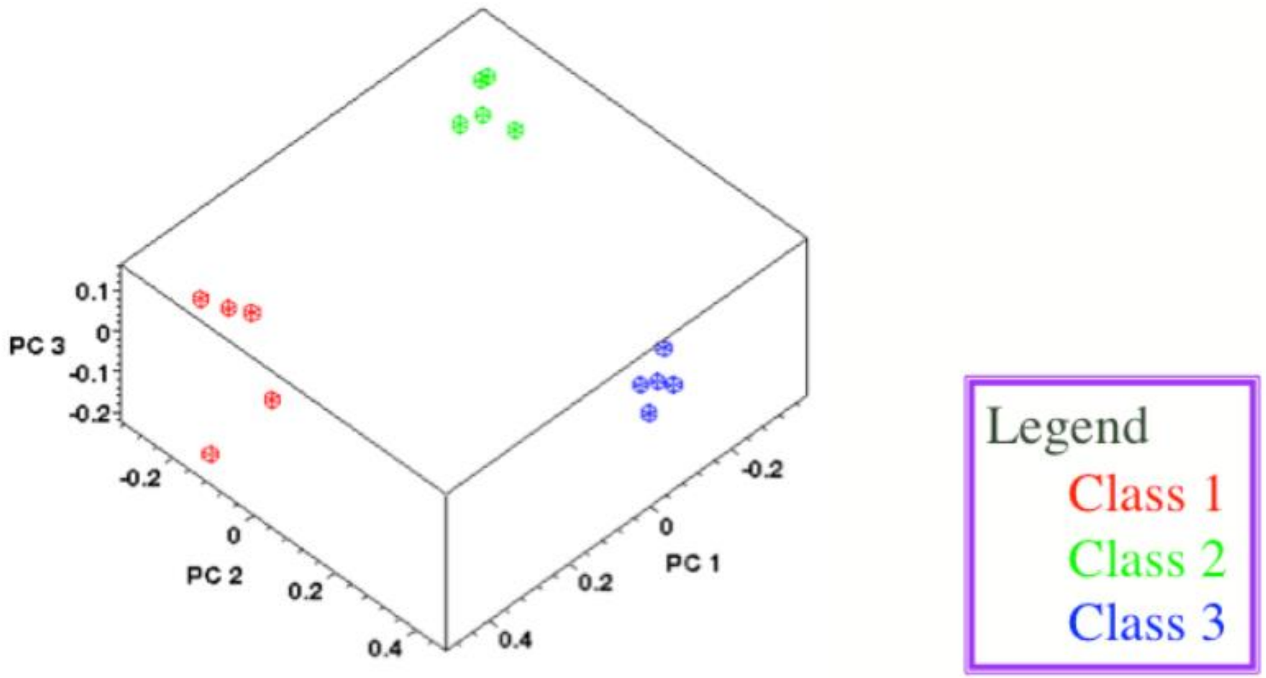

# PCA用于图像降维

- 每类5幅图片



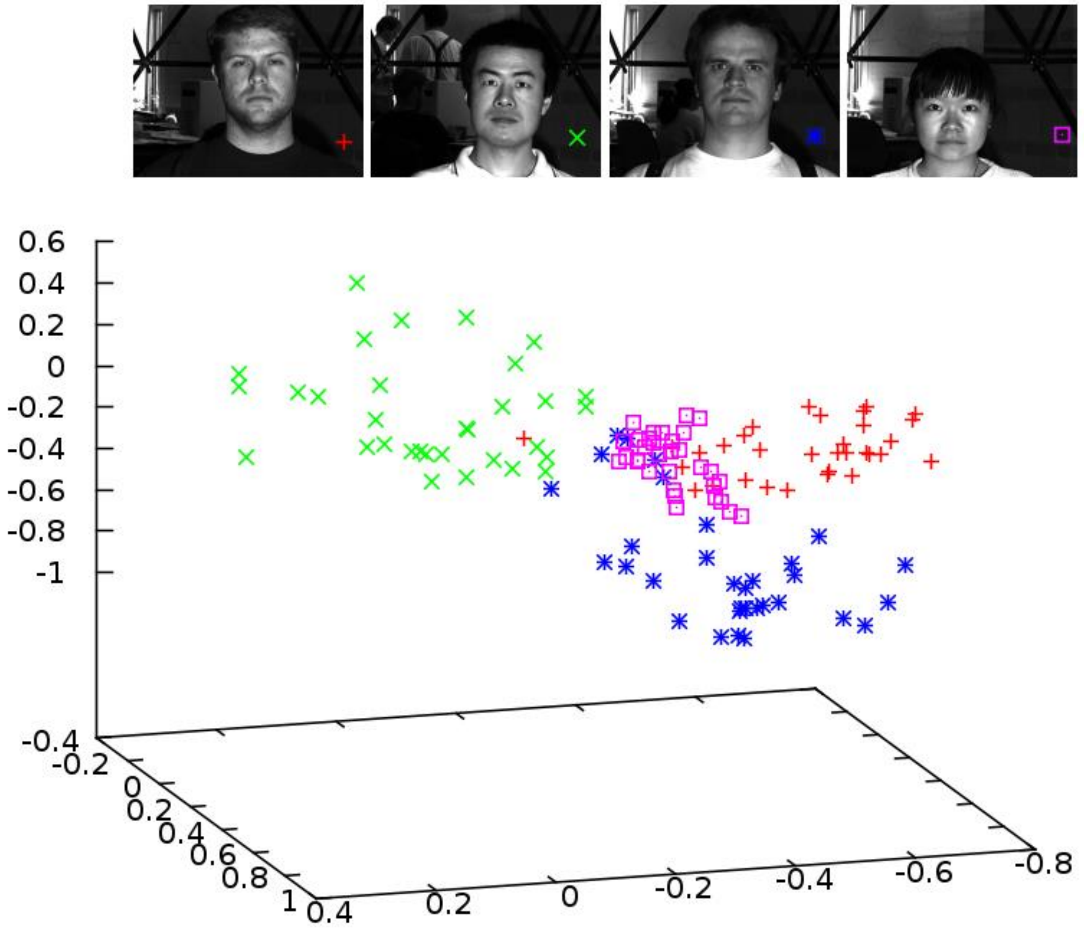
# PCA用于图像降维

- 将图像投影到前3维上
- 同样用最近邻法可以实现分类



# Face Space 人脸空间

- 每张人脸是高维空间中的一个点
  - 图像维度过高，直接处理困难
  - 图像数据各维之间有大量冗余
- 找出人脸最重要的特征
  - 用PCA方法
- 在低维子空间中比较



# 数据预处理

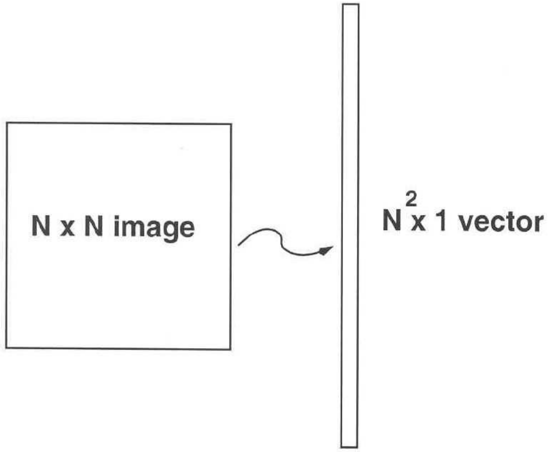
- 人脸数据标准化

$$x_i \in R^d, d = N \times N$$

- 图像数据转成向量化表示
- 数据每维0均值化

$$\hat{X} = \begin{bmatrix} \hat{x}_1 & \hat{x}_2 & \cdots & \hat{x}_n \end{bmatrix} \in R^{d \times n}$$

$$\hat{x}_i = x_i - \frac{1}{n}\sum_{i=1}^{n} x_i$$



# 人脸的主成分分析

- 计算协方差矩阵 $\Sigma \in R^{d\times d}$ 的特征值分解

$$\sum = U\Lambda U^{-1} = U\Lambda U^{T}$$

- 得到数据的前k个主成分

$$\tilde{U} = \left[u_1,\ldots,u_k\right]^T \in R^{k\times d} \quad \mathrm{k} < d$$

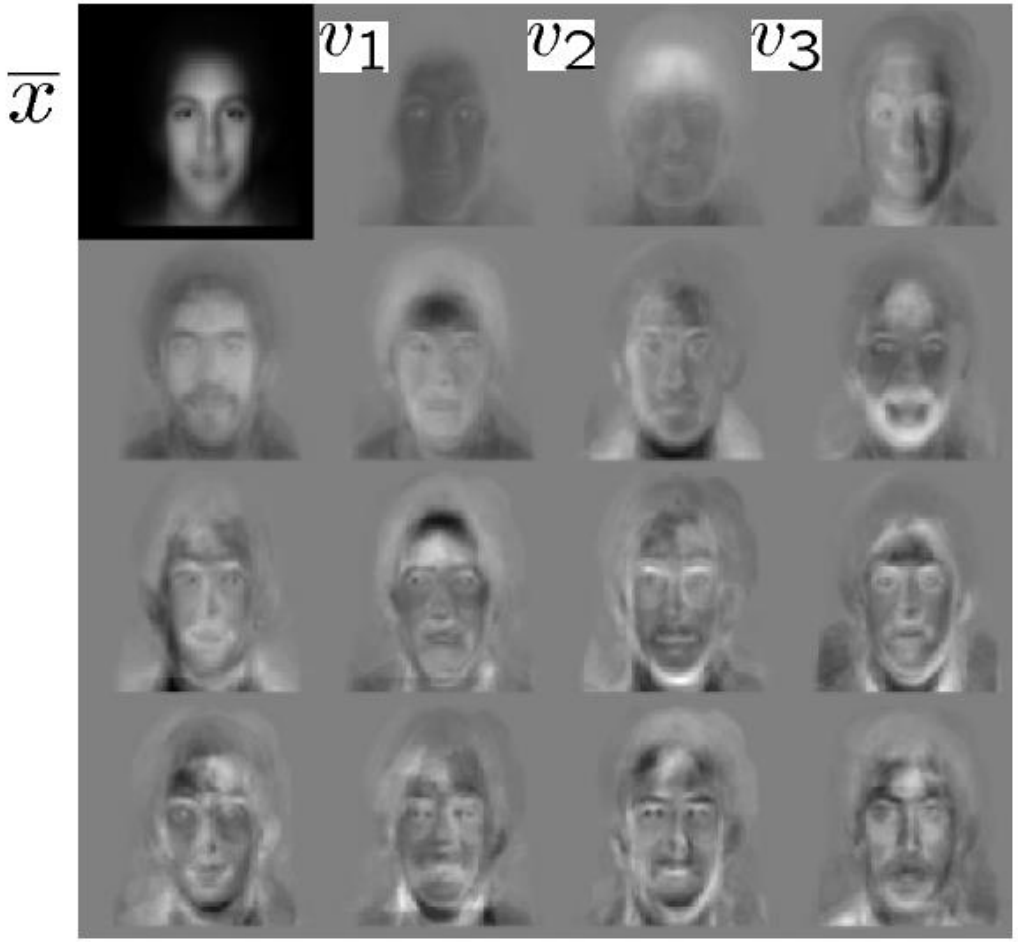
- 每个人脸数据在k个主成分上得到k个投影值

$$\tilde{U}\hat{X}$$

# 特征脸 Eigen face

- 每个特征向量（维度d=NxN）
- 用一个N x N的图像可视化出来

$$\tilde{U} = \left[u_1, \ldots, u_k\right]^T \in R^{k \times d} \quad \mathrm{k} < d$$

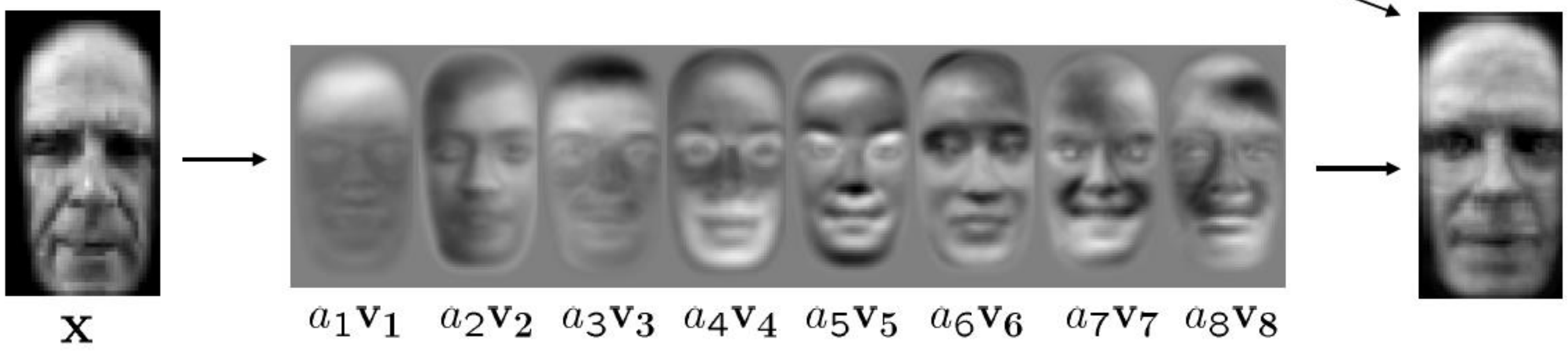


# 特征脸表示

- 每幅图片都可以由特征脸加权表示
- 权重就是投影系数

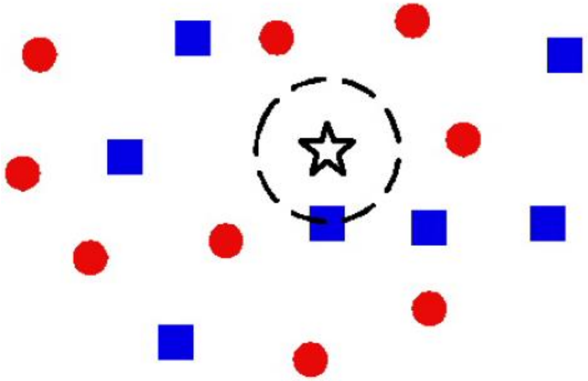
$$\mathbf{x} \rightarrow ((\underbrace{\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v}_1}_{a_1},\ \underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v}_2}_{a_2}, \ldots,\ \underbrace{(\mathbf{x}-\overline{\mathbf{x}}) \cdot \mathbf{v}_K}_{a_K})$$

$$\mathbf{x} \approx \overline{\mathbf{x}} + a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_K\mathbf{v}_K$$



# 基于PCA的人脸识别

- 算法步骤
  - 用PCA从训练集中得到主成分向量
  - 将训练集所有人脸投影到 *d* 个主成分上， *d* 小于原始数据维度
  - 将待测图像投影到 *d* 个主成分上
  - 在 *d* 维空间中做最近邻法分类



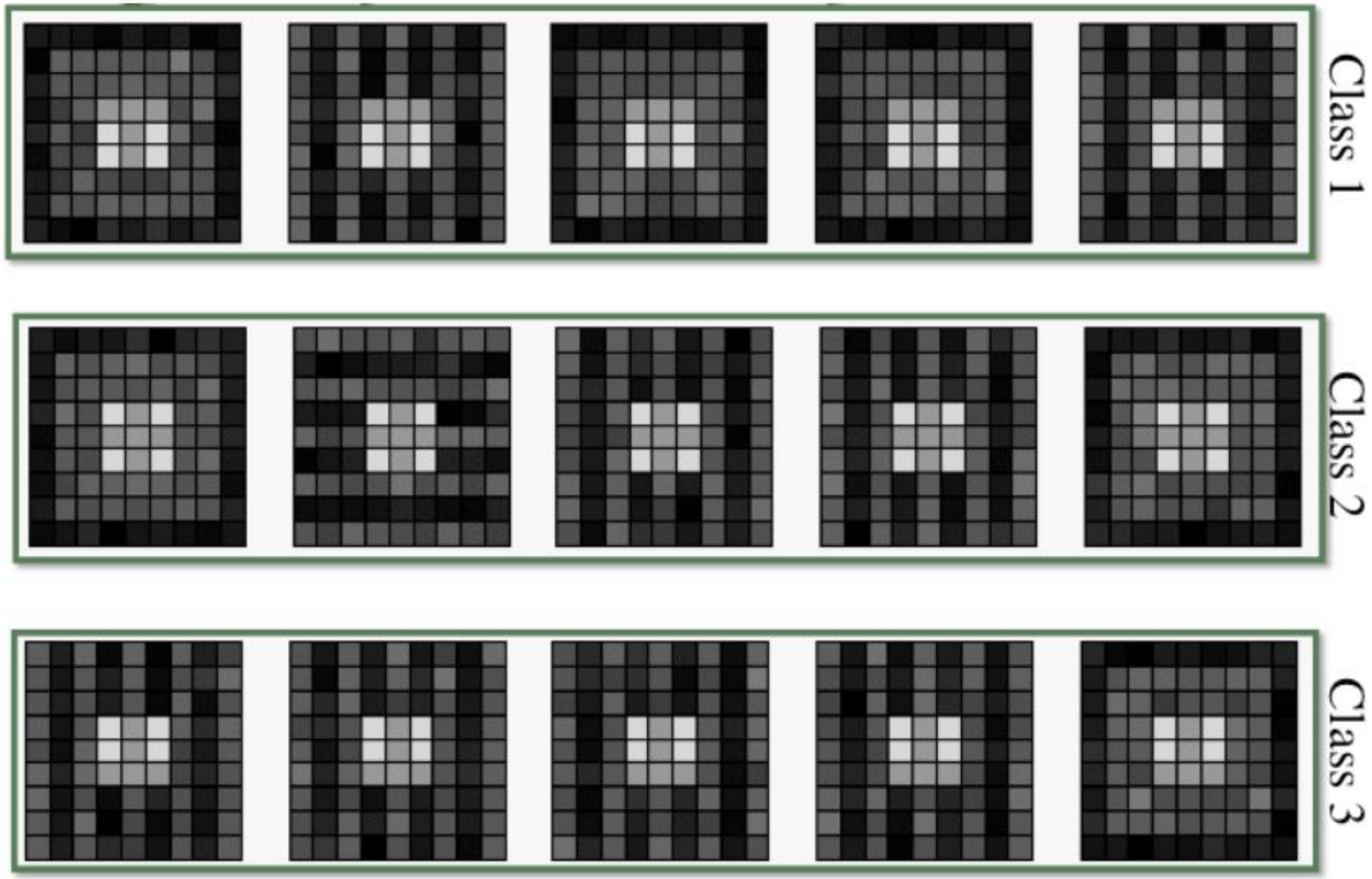
$$x' = \arg\min_{j=1,\cdots,N} D(x, x_j)$$

# 测试结果

- 包含16个人的图像集
- 每次随机选择16幅人脸图像用于训练
  - 每人1幅
- 保留7个主成分向量

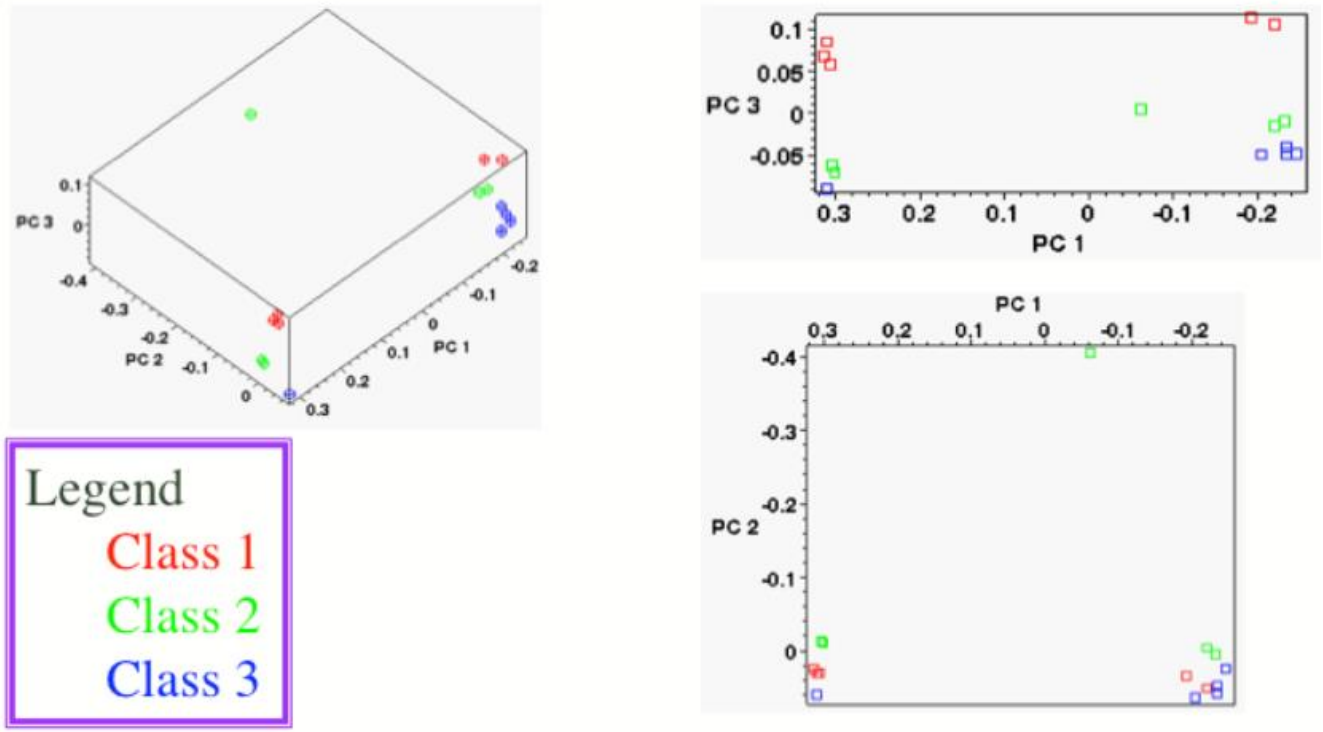
- 多次实验后的平均值
  - 光照变化情况下，准确率 96%
  - 人脸朝向变化情况下，准确率 85%
  - 人脸图像尺寸变化情况下，准确率 64%

# PCA的局限性



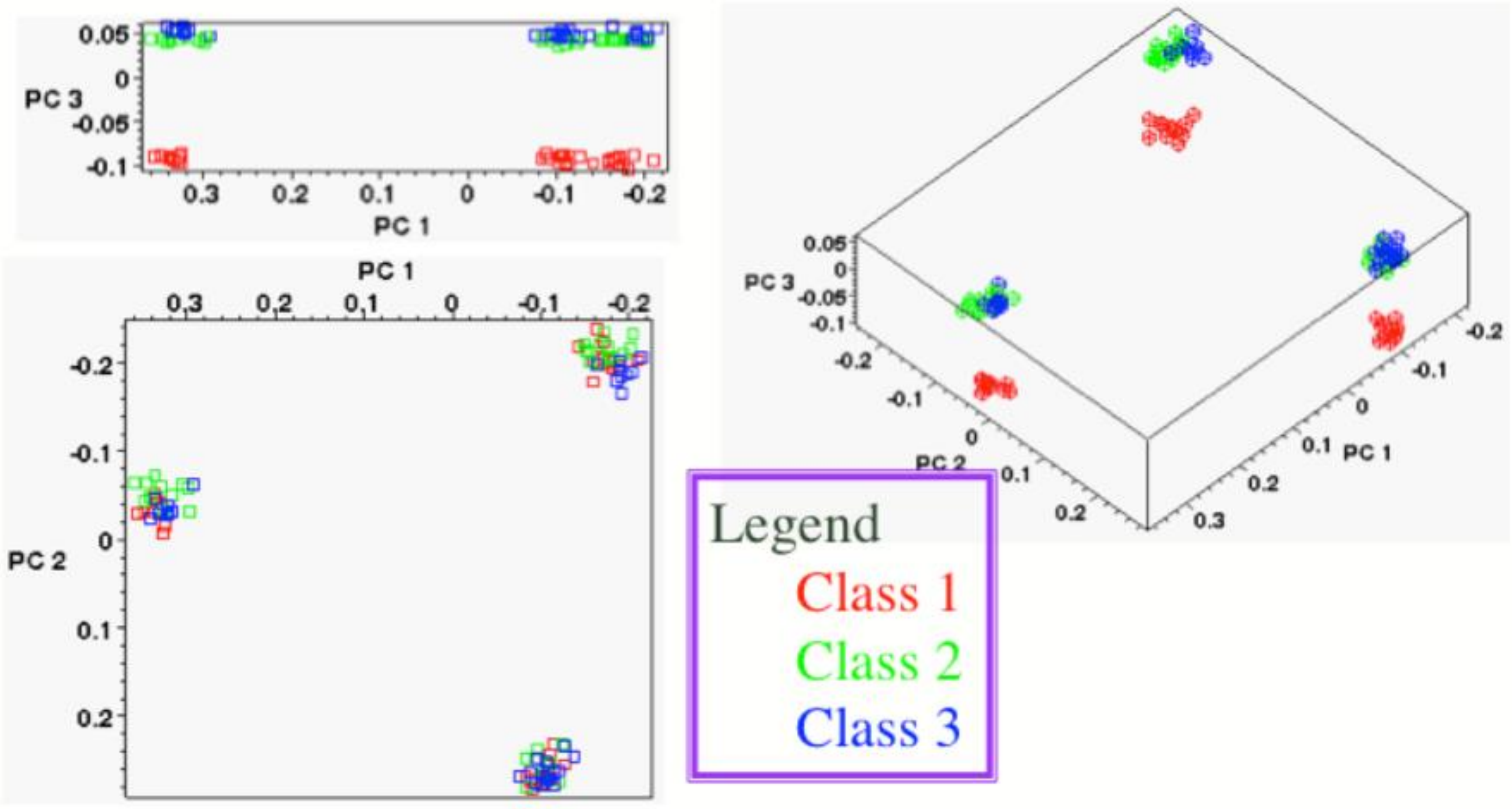
# PCA提取特征的局限性

- 前两个主成分并不包含分类所需的信息



# PCA提取特征的局限性

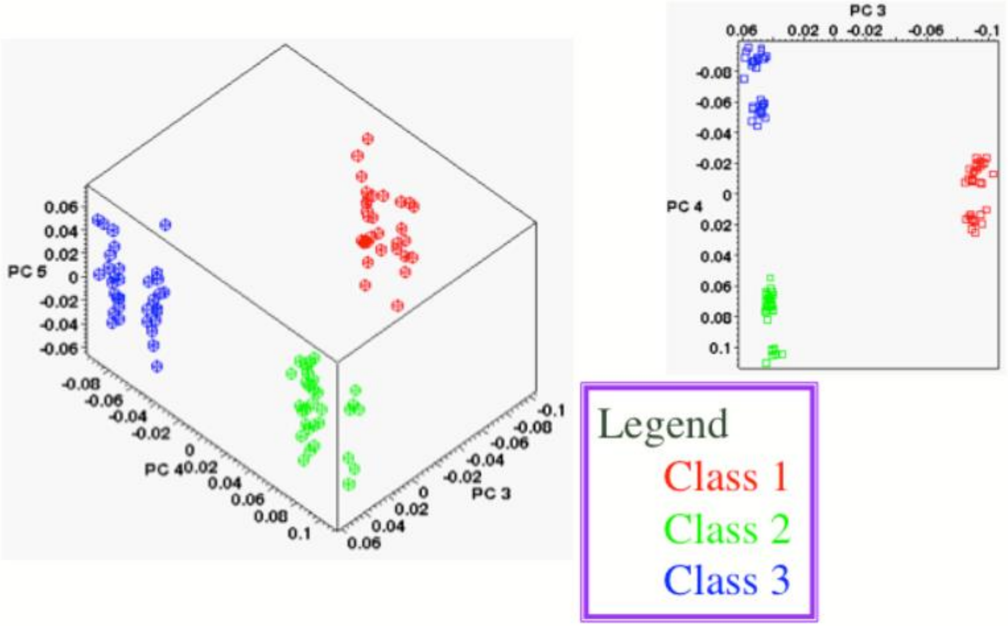
- 增加样本到99个



# PCA提取特征的局限性

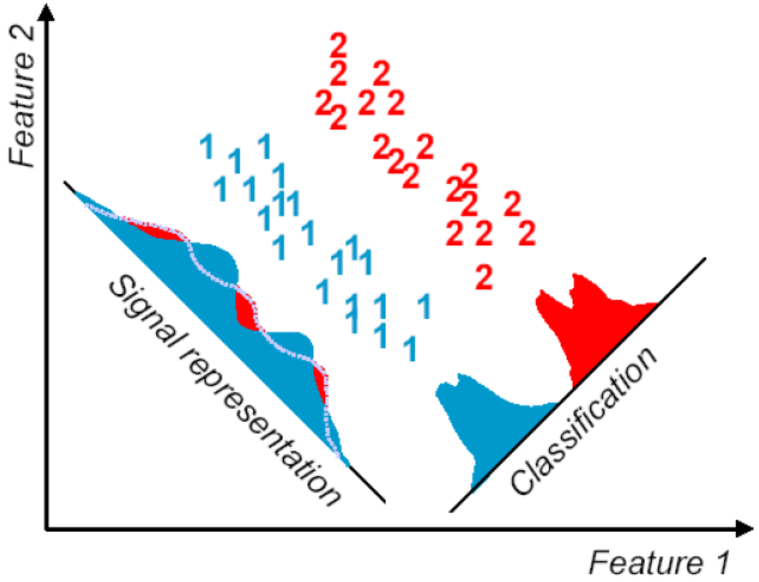
- 第3，4个主成分包含了足够的分类信息
- 最重要的主成分对分类问题并不一定是有效的

[4.775, 3.846, .4245, .3970, .07430 ]
1 2 3 4 5



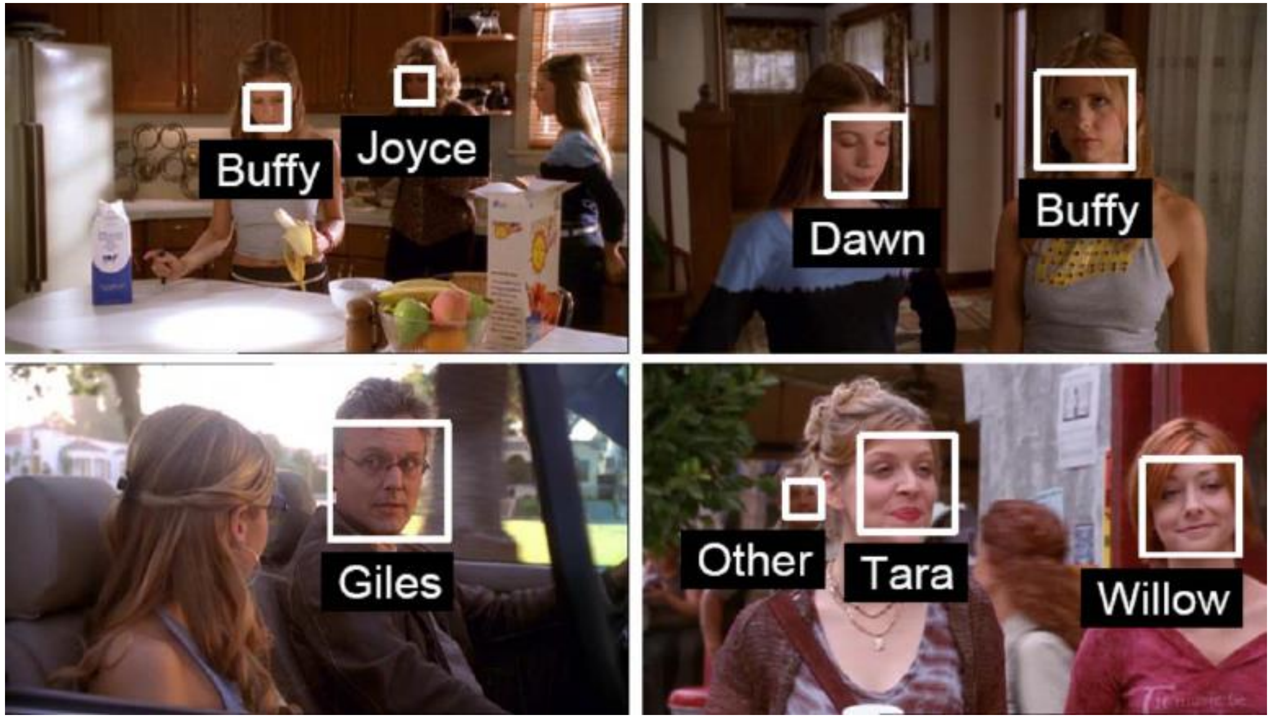
# PCA提取特征的局限性

- 方差最大的投影方向不一定是分类最优的
- 线性判别分析
  - 利用Label信息



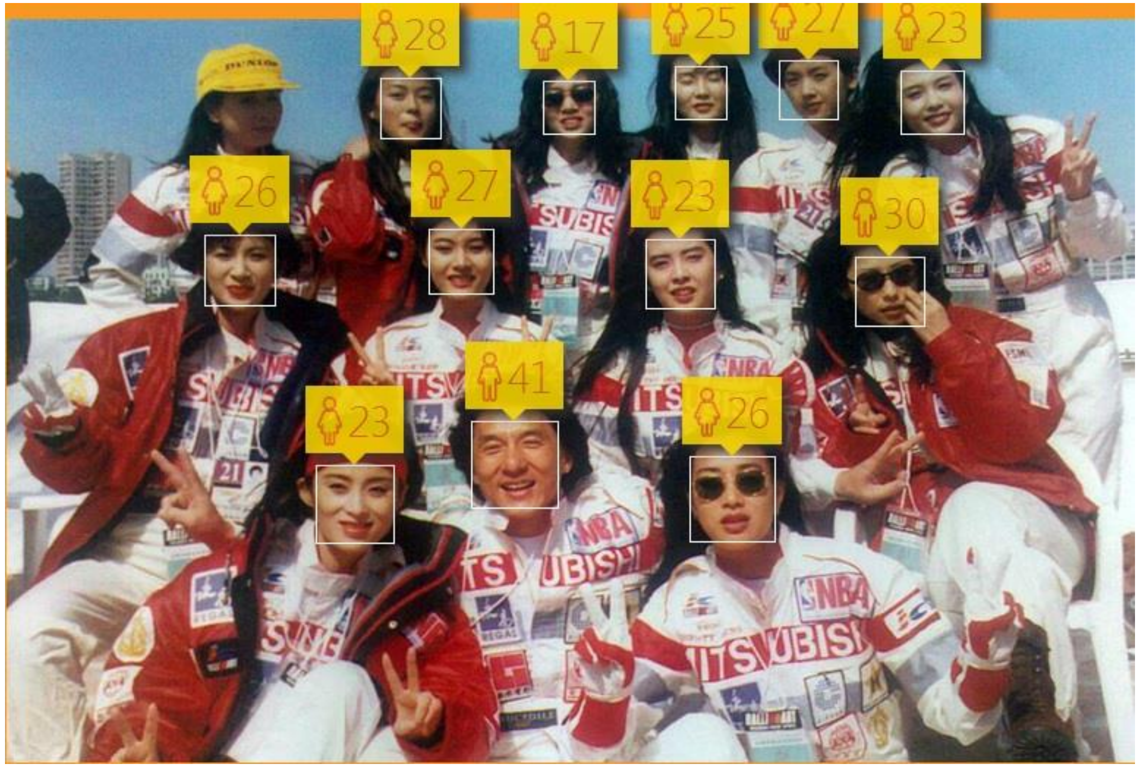
# 人脸识别问题描述

- 训练：给定一组人脸图像及其类别标号
- 测试：对于一幅新图片，判断是哪个人



# For fun

- http://www.myheritage.cn/
- http://how-old.net/

# PCA小结

- PCA将一组相关的坐标变成不相关的坐标
- 可以将数据转换到低维表达
- 非监督方法 （不用到类别标号）
- 可以看作原始坐标系的旋转
- 新坐标轴互相正交，并依次沿着最大方差方向
- 可以重建图像，因此可以用来做数据压缩
- 对分类问题并不一定最优