

# 数字媒体技术

## 1. PCA

### 1.1. 背景：

在多元统计分析中，主成分分析（英语：Principal components analysis，PCA）是一种分析、简化数据集的技术。主成分分析经常用于减少数据集的维数，同时保持数据集中的对方差贡献最大的特征。这是通过保留低阶主成分，忽略高阶主成分做到的。这样低阶成分往往能够保留住数据的最重要方面。但是，这也不是一定的，要视具体应用而定。由于主成分分析依赖所给数据，所以数据的准确性对分析结果影响很大。

### 1.2. 数学定义：

PCA的数学定义是：一个正交化线性变换，把数据变换到一个新的坐标系统中，使得这一数据的任何投影的第一大方差在第一个坐标（称为第一主成分）上，第二大方差在第二个坐标（第二主成分）上，依次类推[4]。定义一个 $n \times m$ 的矩阵 $X$ 为去平均值（以平均值为中心移动至原点）的数据，其行是数据样本，列为数据类别（注意，这里定义的是 $X$ 而不是 $X^T$ ）。则 $X$ 的奇异值分解为 $X = W\Sigma V^T$ ，其中 $m \times m$ 矩阵 $W$ 是 $XX^T$ 的本征矢量矩阵， $\Sigma$ 是 $m \times n$ 的非负矩形对角矩阵， $V$ 是 $n \times n$ 的 $X^T X$ 的本征矢量矩阵。据此，

$$\begin{aligned} Y^T &= X^T W \\ &= V \Sigma^T W^T W \\ &= V \Sigma^T \end{aligned}$$

当 $m < n - 1$ 时， $V$ 在通常情况下不是唯一定义的，而 $Y$ 则是唯一定义的。 $W$ 是一个正交矩阵， $Y^T$ 是 $X^T$ 的转置，且 $Y^T$ 的第一列由第一主成分组成，第二列由第二主成分组成，依此类推。为了得到一种降低数据维度的有效办法，我们可以利用 $W$ 把 $X$ 映射到一个只应用前面 $L$ 个向量的低维空间中去：

其中 $\Sigma_L = I_{L \times m} \Sigma$ ，且 $I_{L \times m}$ 为 $L \times m$ 的[单位矩阵](#)。

$X$ 的单向量矩阵 $W$ 相当于协方差矩阵的本征矢量 $C = X X^T$ ，

$$X X^T = W \Sigma \Sigma^T W^T$$

通常，为了确保第一主成分描述的是最大方差的方向，我们会使用平均减法进行主成分分析。如果不执行平均减法，第一主成分有可能或多或少的对应于数据的平均值。另外，为了找到近似数据的最小均方误差，我们必须选取一个零均值[\[5\]](#)。

假设零经验均值，数据集 $X$ 的主成分 $w_1$ 可以被定义为：

$$w_1 = \arg \max_{\|w\|=1} \text{Var}\{w^T X\} = \arg \max_{\|w\|=1} E\{(w^T X)^2\}$$

为了得到第 $k$ 个主成分，必须先从 $X$ 中减去前面的 $k - 1$ 个主成分：

$$\hat{X}_{k-1} = X - \sum_{i=1}^{k-1} w_i w_i^T X$$

然后把求得的第 $k$ 个主成分带入数据集，得到新的数据集，继续寻找主成分。

$$w_k = \arg \max_{\|w\|=1} E\{(w^T \hat{X}_{k-1})^2\}.$$

### 1.3. 算法实现一般过程：

1. 求平均值以及做normalization
2. 求协方差矩阵(Covariance Matrix)
3. 求协方差矩阵的特征根和特征向量
4. 选择主要成分（信息量）
5. 转化得到降维的数据

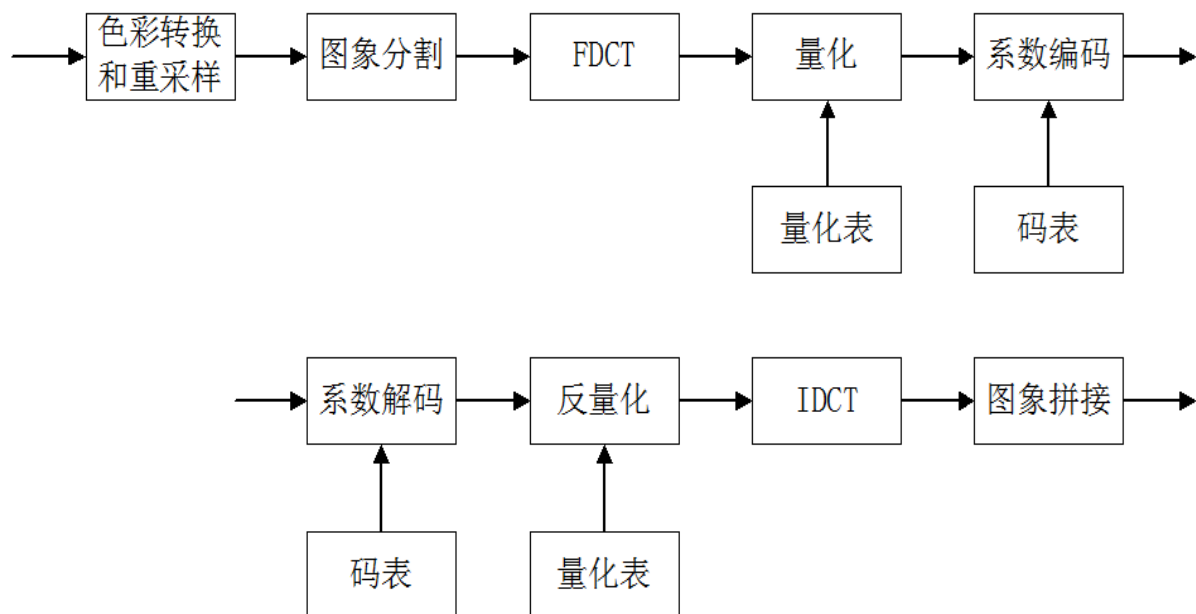
### 1.4. 实验结果

```
Interrupt: Press ENTER or type command to continue
请输入信息量 : 0.5
k = 5
请输入信息量 : 0.6
k = 9
请输入信息量 : 0.7
k = 16
请输入信息量 : 0.8
k = 28
请输入信息量 : 0.95
k = 73
请输入信息量 : 0.99
k = 108
请输入信息量 : 0.999
k = 125
请输入信息量 : 0.99999
k = 128
请输入信息量 : 0
```

## 2. JPEG压缩原理与DCT离散余弦变换

### 2.1. 背景：

DCT变换的全称是离散余弦变换(Discrete Cosine Transform)，主要用于将数据或图像的压缩，能够将空域的信号转换到频域上，具有良好的去相关性的性能。DCT变换本身是无损的，但是在图像编码等领域给接下来的量化、哈弗曼编码等创造了很好的条件，同时，由于DCT变换时对称的，所以，我们可以在量化编码后利用DCT反变换，在接收端恢复原始的图像信息。DCT变换在当前的图像分析已经压缩领域有着极为广大的用途，我们常见的JPEG静态图像编码以及MJPEG、MPEG动态编码等标准中都使用了DCT变换。



### 2.2. 数学定义：

#### 2.2.1. 一维DCT变换

一维DCT变换时二维DCT变换的基础，所以我们先来讨论下一维DCT变换。一维DCT变换共有8种形式，其中最常用的是第二种形式，由于其运算简单、适用范围广。我们在这里只讨论这种形式，其表达式如下：

$$F(u) = c(u) \sum_{i=0}^{N-1} f(i) \cos \left[ \frac{(i+0.5)\pi}{N} u \right]$$

$$c(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & u \neq 0 \end{cases}$$

其中， $f(i)$ 为原始的信号， $F(u)$ 是DCT变换后的系数， $N$ 为原始信号的点可以认为是一个补偿系数，可以使DCT变换矩阵为正交矩阵。

### 2.2.2. 二维DCT变换

二维DCT变换其实是在一维DCT变换的基础上在做了一次DCT变换，其公式如下：

$$F(u, v) = c(u)c(v) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) \cos \left[ \frac{(i+0.5)\pi}{N} u \right] \cos \left[ \frac{(j+0.5)\pi}{N} v \right]$$

$$c(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & u \neq 0 \end{cases}$$

由公式我们可以看出，上面只讨论了二维图像数据为方阵的情况，在实际应用中，如果不是方阵的数据一般都是补齐之后再作变换的，重构之后可以去掉补齐的部分，得到原始的图像信息，这个尝试一下，应该比较容易理解。

另外，由于DCT变换高度的对称性，在使用Matlab进行相关的运算时，我们可以使用更简单的矩阵处理方式：

$$F = AfA^T$$

$$A(i, j) = c(i) \cos \left[ \frac{(j+0.5)\pi}{N} i \right]$$

### 2.3. JPEG压缩流程：

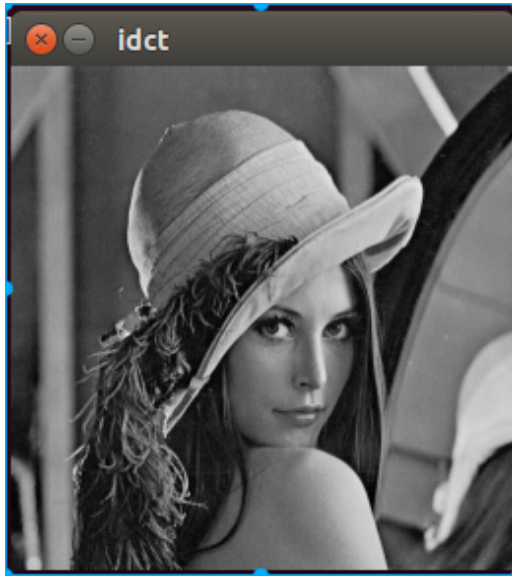
1. 以8x8的图像块为基本单位进行编码
2. 将RGB转换为亮度-色调-饱和度系统（YUV），并重新采样
3. FDCT
4. 量化
5. 编码
6. 解码
7. 反量化
8. IDCT
9. 图像拼接

### 2.4. 实验结果：

DCT:



IDCT:



压缩率：10



压缩率：50



JPEG压缩比例，就是通过控制量化的多少来控制。比如，上面的量化矩阵 $Q$ ，如果我把矩阵的每个数都double一下，那是不是会出现更多的0？！说不定都只有 $G(0, 0)$ 非0，其他都是0，如果这样，那编码时就可以更省空间啦，N个0只要一个游程编码搞定，数据量超小。

但也意味着，恢复时，会带来更多的误差，图像质量也会变差了。

### 3. Kmeans

#### 3.1. 背景

k-平均算法源于信号处理中的一种向量量化方法，现在则更多地作为一种聚类分析方法流行于数据挖掘领域。k-平均聚类的目的是：把n个点（可以是样本的一次观察或一个实例）划分到k个聚类中，使得每个点都属于离他最近的均值（此即聚类中心）对应的聚类，以之作为聚类的标准。这个问题将归结为一个把数据空间划分为Voronoi cells的问题。这个问题在计算上是困难的（NP困难），不过存在高效的启发式算法。一般情况下，都使用效率比较高的启发式算法，它们能够快速收敛于一个局部最优解。这些算法通常类似于通过迭代优化方法处理高斯混合分布的最大期望算法（EM算法）。而且，它们都使用聚类中心来为数据建模；然而k-平均聚类倾向于在可比较的空间范围内寻找聚类，期望-最大化技术却允许聚类有不同的形状。

#### 3.2. 数学定义

设计的目的：使各个样本与所在簇的质心的均值的误差平方和达到最小（这也是评价K-means算法最后聚类效果的评价标准）。

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|_2^2$$

#### 3.3. 一般步骤

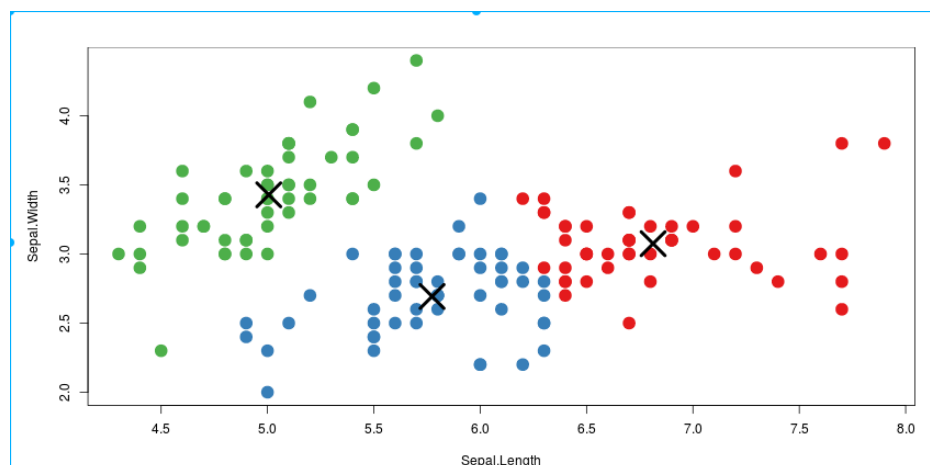
1. 创建k个点作为k个簇的起始质心（经常随机选择）。
2. 分别计算剩下的元素到k个簇中心的相异度（距离），将这些元素分别划归到相异度最低的簇。
3. 根据聚类结果，重新计算k个簇各自的中心，计算方法是取簇中所有元素各自维度的算术平均值。
4. 将D中全部元素按照新的中心重新聚类。
5. 重复第4步，直到聚类结果不再变化。
6. 最后，输出聚类结果。

#### 3.4. 伪代码

```
创建k个点作为K个簇的起始质心（经常随机选择）
当任意一个点的簇分配结果发生变化时（初始化为True）
    对数据集中的每个数据点，重新分配质心
        对每个质心
            计算质心到数据点之间的距离
        将数据点分配到距其最近的簇
    对每个簇，计算簇中所有点的均值并将均值作为新的质心
```

#### 3.5. 实验结果

由于数据作业所给数据集维数较高没法进行可视化，因此这里仅展示算法在二维数据下的可视化结果 数据集采用SIRI数据前两维



## 4. 实验心得

在本学期的数字媒体基础上机实验中，通过编写相应的程序，理解了 PCA 降维、DCT 变换 以及高维数据检索中的相关技术和其背后的原理。在信息爆炸性增长的时代，在这些信息中，多媒体数据（图片、视频）占据了绝大部分，掌握如何对这些媒体信息进行检索、处理的技能是一项必备的技能。熟悉了运用Python进行相关代码的编写，大大简化了工作量。