Lecture 32

一、基的选取

1. JPEG 压缩

假设我们有一张图片，长和宽皆为 512 像素，用x_i表示第 i 个元素，如果是灰度图片，$0 \le x_i \le 255$，是 8 bits，对于承载图片的 x 来说，维度是 512×512。而如果是彩色照片，通常需要三个量来表示一个像素，则向量长度也会变为现在的三倍。

这么大的数据需要压缩后才能传送，教学录像采用的压缩方法就是 JPEG，该方法采用的就是基变换的方式压缩图像。比如说一块黑板，其附近的像素值应该非常接近，此时如果一个像素一个像素的描述黑白灰度值就太浪费空间了，所以标准基在这种情况下并不能很好的利用图片的特性。

标准基是：$\begin{bmatrix}1\\0\\0\\\vdots\\0\end{bmatrix}\begin{bmatrix}0\\1\\0\\\vdots\\0\end{bmatrix}\cdots\begin{bmatrix}0\\0\\0\\\vdots\\1\end{bmatrix}$，我们希望选取一组新的基。

其他的基：

a) 基中含有的一个向量$[1 \quad \dots \quad 1]^T$，分量全为 1 的向量，一个向量就可以给出像素一致的图片的信息。
b) $[1 \quad -1 \quad \dots \quad -1 \quad 1]^T$，正负 1 交替的向量，比如描述国际象棋棋盘
c) $[1 \quad 1 \quad \dots \quad -1 \quad -1]^T$，一半正一半负的向量，比如描述一半明亮，一半阴暗的图片。

2. 傅里叶基

以 8×8 大小来介绍傅里叶基，即将原图片分解为 8×8 这么大小的块，分开处理。

$$\begin{bmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{bmatrix}、\begin{bmatrix}1\\\omega^1\\\omega^2\\\omega^3\\\omega^4\\\omega^5\\\omega^6\\\omega^7\end{bmatrix}、\cdots、\begin{bmatrix}1\\\omega^7\\\omega^{14}\\\omega^{21}\\\omega^{28}\\\omega^{35}\\\omega^{42}\\\omega^{49}\end{bmatrix}$$

每次处理一块图像，就会得到 64 个像素，64 个基向量，64 个系数，在 64 维空间做基变换：

1) 输入信号 x 为 64 维向量 $\xrightarrow{\text{基变换}}$ 输出信号 c 为 x 在傅里叶基下的 64 个系数。注意前面做的都是无损的步骤，我们只是选了 R64 的一组基，接着把信号用这组基表达出来。
2) 一种方法是扔掉较小的系数，这叫做阈值量化，我们设定一个阈值，任何小于阈值的系数和基变量都被丢弃，系数 c 边为$\bar{c}$，肉眼很难区别图像的区别。通常$[1 \quad \dots \quad 1]^T$向量很难被丢弃，它通常具有较大的系数。但是$[1 \quad -1 \quad \dots \quad -1 \quad 1]^T$向量在平滑信号中的可能性就很小了。前一个的向量称作低频信号，频率为 0，后一个向量称作高频信号，也是我们能够得到的最高频率的信号，如果是噪音或抖动输出的就是它。
3) 用系数$\bar{c}$重构信号，$\bar{x} = \sum \bar{c}_i v_i$，求和项不是 64 项，因为有许多零存在。

3. 小波基

依旧以 8 维向量来介绍小波基 W。

$$\begin{bmatrix}1\\1\\1\\1\\1\\1\\1\\1\end{bmatrix}、\begin{bmatrix}1\\1\\1\\1\\-1\\-1\\-1\\-1\end{bmatrix}、\begin{bmatrix}1\\1\\-1\\-1\\0\\0\\0\\0\end{bmatrix}、\begin{bmatrix}0\\0\\0\\0\\1\\1\\-1\\-1\end{bmatrix}、\begin{bmatrix}1\\-1\\0\\0\\0\\0\\0\\0\end{bmatrix}、\cdots、\begin{bmatrix}0\\0\\0\\0\\0\\0\\1\\-1\end{bmatrix}$$

可以看出傅里叶基中频率最高的向量为小波后四个基向量之和。

一个像素$\mathbf{P}=\begin{bmatrix}p_1\\p_2\\p_3\\p_4\\p_5\\p_6\\p_7\\p_8\end{bmatrix}=c_1w_1+\cdots+c_8w_8=\mathbf{Wc}$，所以 c=$W^{-1}\mathbf{P}$

可以发现，在求解系数的过程中，需要计算基的逆，如果基选取的好，会使得求解过程变得简单。

a) 计算快

需要大量计算 P=Wc 计算图像在另一个基下的表示，傅里叶可以采用 FFT，小波基可以使用 FWT。

b) 少量向量可以重构图像

图形压缩时会丢弃较小的系数，重构仅使用系数较大的基，选取好的基会使重构使图像损失最小。

二、基变换

前面介绍小波基的时候我们就已经做了一次基变换。

将目标基的向量按列组成矩阵 W，则基变换就是$x=Wc$

加入我们有线性变换 T，$R^8\to R^8$，对于基$v_1\cdots v_8$，有转换矩阵 A，对于基$w_1\ldots w_8$，有转换矩阵 B，则矩阵 A 与 B 是相似矩阵。

对于第一组基$v_1\cdots v_8$，要完全了解线性变换 T，只需要知道 T 在每个基向量的结果即可。

$\mathrm{T}(v_1)=a_{11}v_1+a_{81}v_8$，基向量前的系数，就构成了矩阵 A 的第一列。即矩阵 A：

$$\begin{bmatrix}a_{11}&\cdots&a_{1n}\\\vdots&\ddots&\vdots\\a_{m1}&\cdots&a_{mn}\end{bmatrix}$$

设$v_1\cdots v_m$是一组特征向量，也就是$\mathrm{T}(v_i)=\lambda_i v_i$，矩阵 A 是多少。

$$\mathbf{A}=\begin{bmatrix}\lambda_1&\cdots&0\\\vdots&\ddots&\vdots\\0&\cdots&\lambda_m\end{bmatrix}$$