矩阵计算

李宇峰
liyf@nju.edu.cn
人工智能学院

4.3 SVD的应用-图像压缩 (1)

1. 问题的背景

随着科学技术的高速发展和网络应用的普及，有大量的数字信息需要存储、处理和传送。图像信息作为重要的多媒体资源，其数据量很大。

例如，一幅1024×768的24位BMP图像，其数据量约为2.25*MB*。大数据量的图像信息会给存储器的存储容量、通信干线信道的带宽以及计算机的处理速度增加极大的压力。单纯靠增加存储器容量、提高计算机处理速度等方法来解决这个问题是不现实的，因此图像压缩就显得十分必要。

图像压缩之所以可行是因为图像数据是高度相关的，表现在：

1. 大多数图像内相邻像素之间有较大的相关性，存在很大的冗余度，这是空间冗余度；

2. 序列图像前后帧之间有较大的相关性，这是时间冗余度；

3. 用相同的码长表示不同出现概率的符号也会造成比特数的浪费，这是符号冗余度；

4. 允许图像编码有一定的失真也是图像可压缩的一个重要原因。

图像压缩编码的方法很多：

根据编码过程中是否存在信息的损耗将编码分成有损压缩编码和无损压缩编码两大类。

有损压缩编码就是有些信息丢失，不能完全恢复原始图像，存在一定程度失真；

无损压缩编码就是无信息损失，解压时可以从压缩数据精确恢复到原始图像。

根据编码原理，又可以将图像编码分为熵编码、预测编码、变换编码和混合编码等。

变换编码是常用的一种有损编码方法，这个方法不是直接对空域图像信号进行编码，而是首先将空域图像信号映射变换到变换域(比如频域)，使得数据的冗余度在变换域中大幅减少，然后对这些变换系数进行编码处理，获得较大的压缩比。

2. 问题的提出

奇异值分解是一种基于特征向量的矩阵变换方法，在信号处理、模式识别、数字水印技术等方面得到了广泛的应用。由于图像具有矩阵结构，因此可以将奇异值分解应用于图像压缩。

3. 问题的分析与处理

奇异值分解的一个重要作用是可以降维。

如果A表示n个m维向量，即$A \in R^{m\times n}$，则可以用$(m+n+1)$个r维的向量来表示, 即$U_r \in R^{m\times r}, \Sigma_r = diag(\sigma_1, \sigma_2, \cdots, \sigma_r), V_r \in R^{n\times r}$。如果$A$的秩$r$远远小于$m$和$n$，则通过奇异值分解可以大大降低$A$的维数。

用奇异值分解来压缩图像的基本思想是对图像矩阵进行奇异值分解，选取部分奇异值和对应的左、右奇异向量来重构图像矩阵。

1. 对于$n \times n$图像矩阵A，如果$rank(A) = r$, 将其中奇异值按照从大到小的次序排列。

2. 按奇异值从大到小取k个奇异值和这些奇异值对应的左奇异向量和右奇异向量来重构原图像矩阵A.

3. 如果选择$k \geq r$, 这是无损的压缩，实际上这就是没有压缩；

基于奇异值分解的图像压缩对应的总是$k < r$, 也就是有损压缩的情况。

这种情况，可以只使用$k(2n+1)$个数据来代替 n^2 个图像数据。

这$k(2n+1)$个数据分别是矩阵A的前k个奇异值，$n \times n$左奇异向量矩阵U的前k列向量$U_k = [u_1, u_2, \cdots, u_k]$, 和$n \times n$右奇异向量矩阵$V$的前$k$列向量$V_k = [v_1, v_2, \cdots, v_k]$. 因此图像的压缩比应该是

$$\rho = \frac{n^2}{k(2n+1)}$$

一般情况下，被选择的奇异值的个数k应该满足条件

$$k(2n+1) \ll n^2$$

只有在这样的条件下，在传输图像的过程中，只需要传$k(2n+1)$个有关奇异值和奇异向量的数据即可。在接收端，接收到奇异值$\sigma_1, \sigma_2, \cdots, \sigma_k$, 就可以通过

$$A_k = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_k u_k v_k^T$$

重构出原图像矩阵。

重构图像A_k与原图像A的误差为

$$\|A-A_k\|_F^2=\sigma_{k+1}^2+\sigma_{k+2}^2+\cdots+\sigma_r^2$$

事实上第i个奇异值σ_i对整个图像的贡献率

可以定义为

$$\varepsilon_i=\frac{\sigma_i^2}{\sum_{j=1}^{r}\sigma_j^2}$$

对一幅图像来说，较大的奇异值对图像信息的贡献率比较大，较小的奇异值对图像信息的贡献率比较小。如果

$$\sum_{i=1}^{k} \varepsilon_i = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{j=1}^{r} \sigma_j^2}$$

接近于1，该图像的主要信息就包含在A_k中。

在满足视觉要求的基础上，按奇异值大小选择合适的奇异值个数$k(k<r)$, 就可以通过A_k将图像A恢复出来。k越小, 用于表示A的数据量越小，压缩比就越大；而k越接近于$r = rank(A)$，则A_k和A就越相似。

按照上面的流程得到了下面的$Lena$图像压缩结果：

原始图像是200×200，即原始奇异值$r = 200$；

这里显示200个奇异值的前50个，最大的σ_1= 50462.6522,

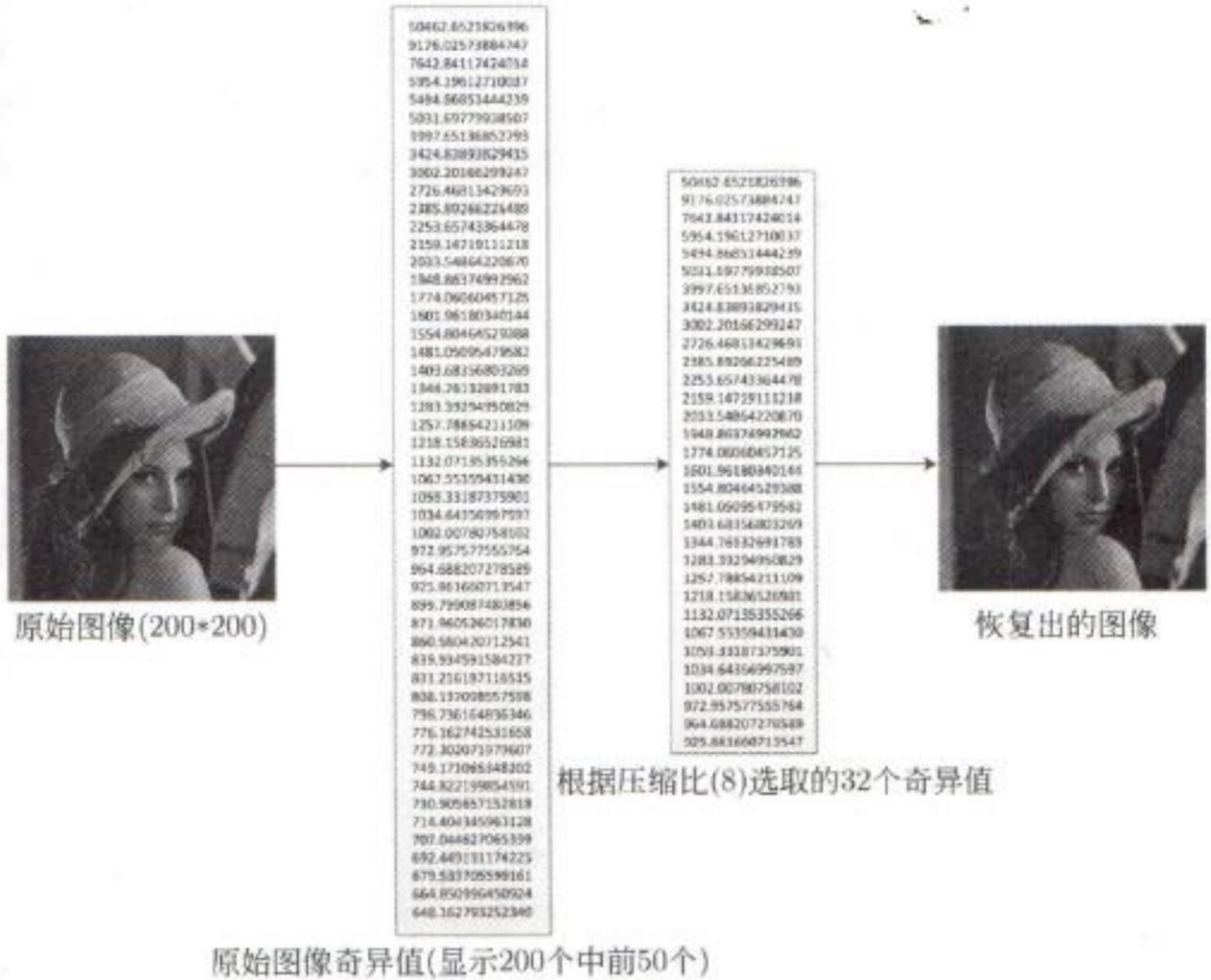
σ_{50}= 648.1628；

根据压缩比控制在$\rho \approx 3$左右，所以我们选取前面的32个奇异值，即$k = 32$.

最后根据

$$A_{32} = \sum_{i=1}^{32} \sigma_i u_i v_i^T$$

重构恢复出图像。


原始图像(200*200)
根据压缩比(8)选取的32个奇异值
恢复出的图像
原始图像奇异值(显示200个中前50个)

从上面的介绍可以看出，奇异值分解中所取的奇异值个数，对于图像的压缩以及质量有着至关重要的影响：

如果所取奇异值个数过小，压缩比会很大，比如上面的图形，我们取$k = 32$, 压缩比达到$\rho = 9.97$，但是图像质量会下降较多。

如果奇异值的个数取得过大，则压缩比明显接近1，图像质量损失较小，但时间、空间花费增加。所以选取合适的奇异值个数对于图像压缩的综合效果很重要。

4. 压缩结果与质量评价

图2给出了不同压缩比的情况下得到的*Lena*图片的压缩结果。对压缩图像的质量评价主要有：均方差（*MSE*）和信噪比（*SNR*）. 假设一幅压缩前的图像表示为一向量

$$a = (a_0, a_1, \cdots, a_{m-1})^T,$$

经过压缩后的图像为向量

$$b = (b_0, b_1, \cdots, b_{m-1})^T.$$

而$d(a,b)$表示向量a与b之间的差距。通常$d(a,b)$定义为

$$d(a,b) = \|a-b\|_2^2 = \sum_{i=0}^{m-1} (a_i - b_i)^2$$

而a和b的均方差D为

$$D = \frac{1}{m}\|a-b\|_2^2 = \frac{1}{m}\sum_{i=0}^{m-1} (a_i - b_i)^2$$

在图像压缩中，信噪比(SNR)通常采用峰值对峰值信噪比($PSNR$)

$$PSNR = 10 \cdot \log_{10}\frac{x_p^2}{D} = 20 \cdot \log_{10}\frac{x_p}{\sqrt{D}}$$

压缩比	压缩结果	标准均方差	PSNR(db)
1.2		1.9736	45.1782
1.4		2.5136	44.1278
2		3.9643	42.1491
4		7.2143	39.5489
8		10.5480	37.8991
16		14.2434	36.5947

由于采用上述两种评定方法所取得的结果与人眼评定结果并不总是一致的，因此主观评定也是成为不可缺少的方法。图像呈现给评定者，并让他们在1~5基础上打分。

诊断精确性评价主要应用于医疗等图像的压缩中，尤其应用于医院外科仿真中（如屏幕诊断）。

5. 结果分析

通过观测上述压缩结果和评价指标，随着压缩比的增大、信噪比的减小、压缩前与压缩后图像像素的标准均方差的增加，压缩后的图像清晰度越来越低，这与图像压缩的基本规律是一致的。

这说明，*SVD*分解应用于图像压缩可以得到比较理想的结果。