

2020

Machine Learning

Fifth Assignment

姓名	班级	学号
王成航	统计 71	2176122248
张申铎	统计 71	2176112379
王泽昊	统计 71	2176112782

k-means 及其改进方法的应用

目录

1 K-means 用于图片压缩	1
1.1 颜色量化算法 (Color quantization algorithm)	1
1.2 k-means 压缩算法具体步骤	1
1.3 实际表现	2
2 FCM 和 KFCM 算法在图像分割中的应用及对比	5
2.1 k-means 与模糊聚类算法的对比	5
2.2 图像分割的原理	6
2.3 FCM 与 KFCM 结果对比	6

1 K-means 用于图片压缩

在日常生活中,由于图像和视频本身的数据量非常大,这就给存储和传输带来了很多不便,所以图像压缩和视频压缩得到了非常广泛的应用. 比如数码相机、USB 摄像头、可视电话视频点播、视频会议系统、数字监控系统等等,都使用到了图像或视频的压缩技术. 在此,下文将介绍一种使用 k-means 算法进行图像压缩的算法.

1.1 颜色量化算法 (Color quantization algorithm)

实际上, k-means 进行图像压缩是一种颜色量化算法的实际应用. 颜色量化是利用人眼对颜色的惰性,将原图像中不太重要的相似颜色合并为一种颜色,减少图像中的颜色,而使量化前后的图像对于人眼的认识误差最小,即量化误差最小. 在彩色图像中,每个像素的大小为 3 字节 (RGB),其中每种颜色的强度值都可以在 0 到 255 之间. 按照组合运算,可以表示的颜色总数为 $256 \times 256 \times 256$. 因此,图像压缩的主要目标是使用颜色量化来压缩图像,并且在图像的失真不太明显的情况下减小图像的储存空间.

1.2 k-means 压缩算法具体步骤

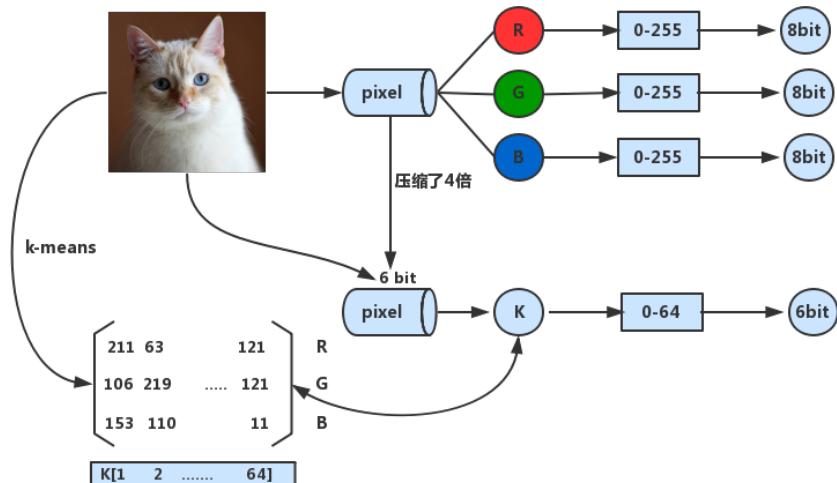


图 1: 算法流程图.

首先, 考虑一个大小为 $128 \times 128 \times 3$ 的图像. 如果对图片进行矢量化处理, 将得到一个大小为 16384×3 的数组, 每一个元素代表原图像的 R,G,B 三个值组成的一个三维向量. 从而可以将图像视为许多个三维向量的数据集. 此时, 将 K-Means 应用于此数据集, 通过选择聚类中心的簇数 n , 将 n 个最重要的质心选取出来. 用这个点的颜色取值来代替属于这一簇的元素的颜色. 这样, 只需保留少数质心的颜色数值, 剩下的点储存其属于哪一个质心便可. 从而对储存空间进行压缩.

对于 $128 \times 128 \times 3$ 的图像, 如果取 k-means 中的聚类簇数 $k = 16 = 2^4$, 即只使用 16 种颜色来对原图像进行颜色量化. 则新图像的总大小为 $128 \times 128 \times 4 = 65536$, 这样, 储存空间在直观上得到了压

缩.

但由于需要储存 16 个质心的颜色值, 每个质心需要 24 位来储存, 因此, 总的储存空间为 $65536 + 24 \times 16 = 65920$. 与原始图像进行比较, 原始图像所需的储存空间为 $128 \times 128 \times 24 = 393216$ 位. 显然, 基于这样的方法, 理论上在 $k = 16$ 时可以将图片压缩至原图片大小的 16.76%.

更一般的, 若原图像尺寸为 $m \times n$, 选取聚类簇数为 k , 则可将一般的 RGB 图像压缩为原图像的

$$\frac{m \times n \times \log_2 k + 24 \times k}{24 \times m \times n} \times 100\%.$$

1.3 实际表现

基于以上的想法, 我们对于两张动物图片进行了压缩, 聚类簇数分别选为 $k = 8, 16, 32, 64, 128$. 首先, 对于第一张图片进行了压缩, 得到的结果分别如图 2.

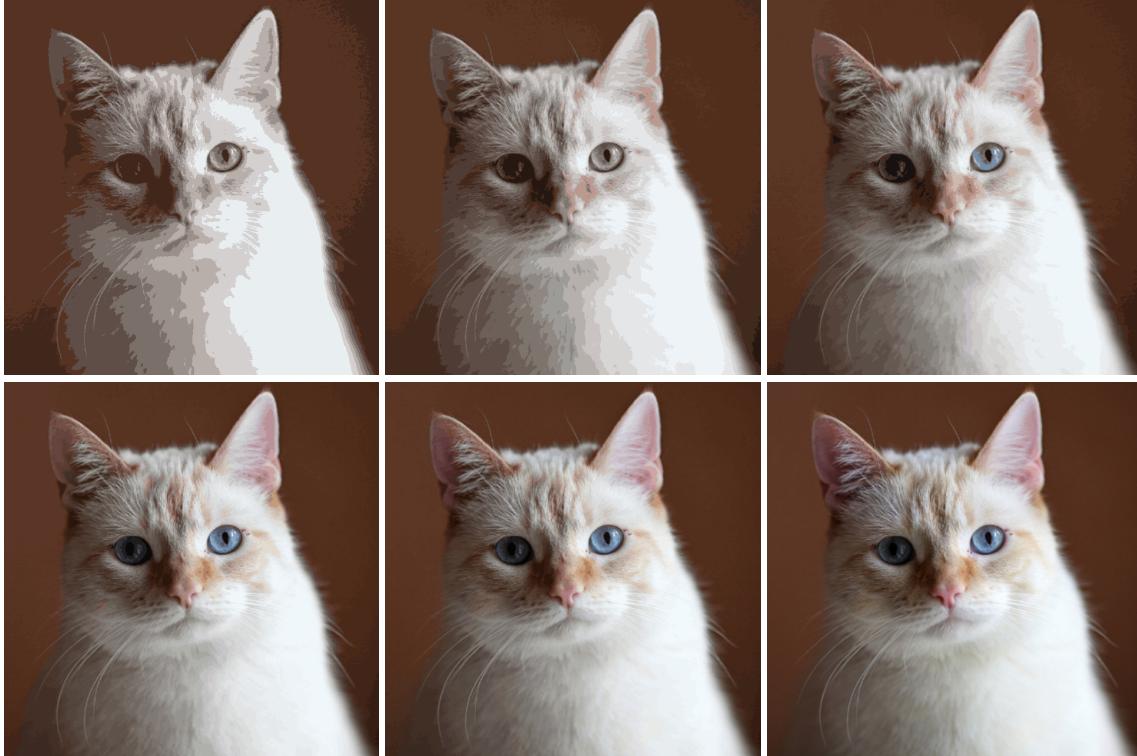


图 2: 从左到右、从上到下 k 的值依次增加, 最后一张图为原图.

每张图片的大小如下表所示

图片大小 (字节)	$k = 8$	$k = 16$	$k = 32$	$k = 64$	$k = 128$
682078	45197	75804	123399	165238	237626
压缩比 (%)	6.626	11.114	18.092	24.226	34.839

从图中可以看出, 实际上 k 取 64 甚至 32 时, 压缩过的图片就已经大致和原图相似了. 第 1 至 3 张图片中因为颜色数较少, 未能表现出猫眼的真实颜色, 但对猫毛的颜色刻画逐渐生动, 到第 4 张时

已经非常接近原图给人的视觉观感.

为了度量压缩图片的失真度, 我们将原图片与压缩后的图片逐像素点做差, 将误差绝对值之和作为对图片失真度的一个量化.

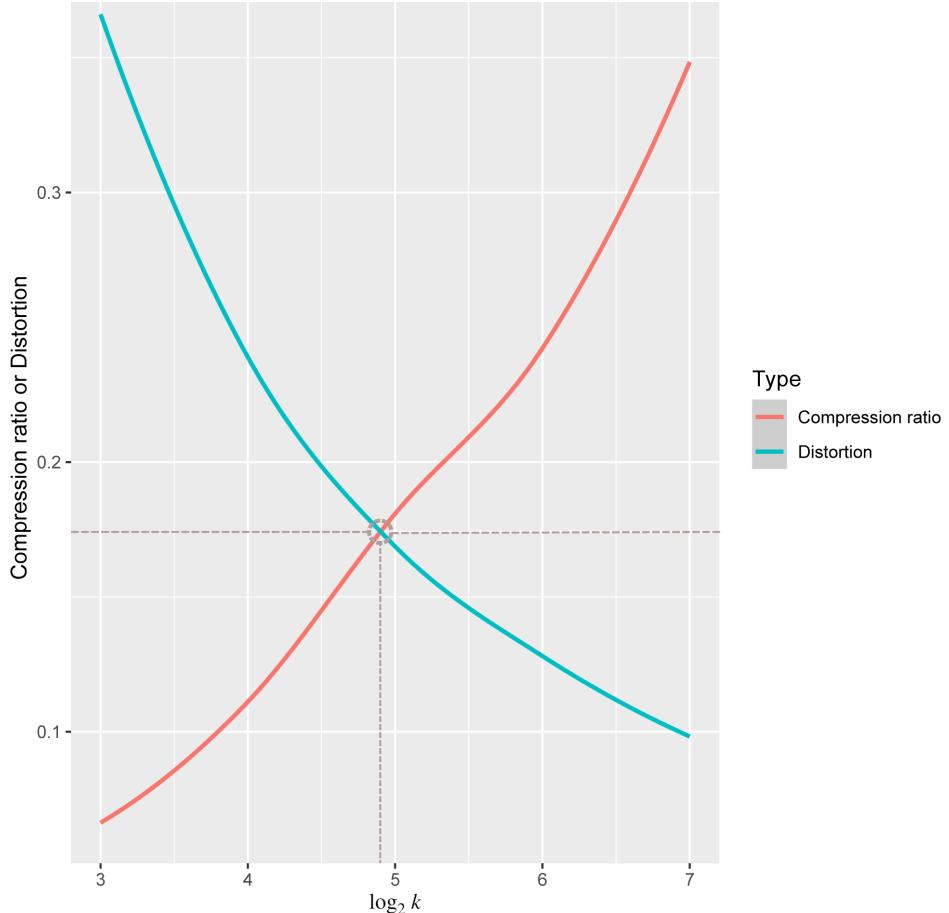


图 3

从中, 我们可以看出, 在 $k = 5$ 时, 压缩比和失真度达到了一个比较均衡的状态, 所以在此时选用 32 位色彩进行压缩时, 可以保证失真度不是很高的情况下达到较好的压缩比.

之后, 对于第二张图片进行了压缩, 得到的结果分别如下:



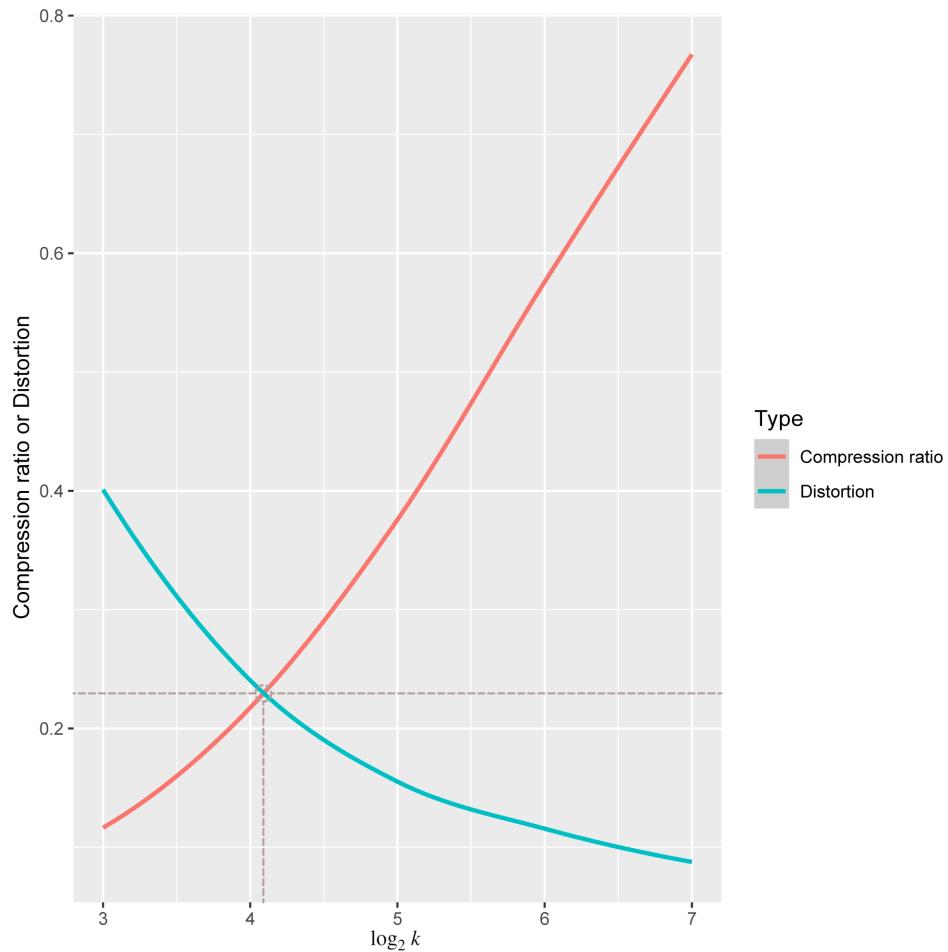


图 4: 从左到右、从上到下 k 的值依次增加, 最后一张图为原图.

每张图片的大小如下表所示

图片大小(字节)	$k = 8$	$k = 16$	$k = 32$	$k = 64$	$k = 128$
1140062	132802	248046	428423	656589	874825
压缩比 (%)	11.647	21.757	37.579	57.592	76.735

同样的, 我们绘制了它的失真度与压缩比曲线



相比于图2,图4的颜色构成较为简单,主色即为蓝、白、黑,所以可以看出在第2张图中,只使用8种颜色时,便已经可以基本表现出猫毛、猫眼等主要特征,虽然在色彩交界处出现了比较明显的断层。因此,此时选用16位颜色便能够在大小和质量上做到很好地平衡。

通过这两张图片的k-means压缩的对比,可以发现,k值的选取与图片本身十分相关。当图片内容比较多时,需要选取较大的k值才能更好地保留原图像的信息。

但k-means图像压缩仍是一种有损的图像压缩方式,即以牺牲一些图像信息为代价减少图片大小,目前也已经有了各种无损的压缩图片技术,但多半都有压缩比率不高的问题存在,但我们相信随着科学技术的发展,会有更加完善的无损压缩技术出现。

2 FCM 和 KFCM 算法在图像分割中的应用及对比

K-means算法可以看作高斯混合聚类在混合成分方差相等、且每个样本仅指派给一个混合成分时的特例。而基于K-means的思想也继而衍生了很多变体,如k-medoids算法[1]强制原型向量必为训练样本,k-mosed算法[[2]可处理离散属性。FCM[3]则是“软聚类”算法,允许每个样本以不同程度属于多个原型。引入核技巧则可得KFCM算法。

2.1 k-means与模糊聚类算法的对比

在k-means中,目标函数是最小化此公式

$$J(\theta) = \sum_{i=1}^k \sum_{k=1}^N d^2(x_k, \theta_i)$$

而模糊聚类算法的目标函数则为

$$J(U, \theta) = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m d^2(x_k, \theta_i)$$

其中 m 为模糊指数, $m > 1$, U 为隶属度矩阵, μ_{ik} 为隶属度(即 U 的元素), $d(\cdot, \cdot)$ 为不相似测度函数, x_k 是输入样本特征, θ_i 是聚类中心, N 是输入样本(即图像像素)总数, c 为聚类数目。

聚类的过程是通过调整 U 和 θ 使上式的求和值最小化。可以理解这个最小化过程是一个聚类过程:要使该求和式最小,则要求其中的每一乘积项最小,就是说, $d(\cdot, \cdot)$ 大的项,要由小的 μ_{ik} 与其相乘。换言之,样本与聚类中心越不相似,则要求该样本对这个聚类中心的隶属度越小,反之,则隶属度越大。由此看出上式的收敛过程就是一个聚类过程。

对比K-means与FCM的目标函数,可以知道,K-means是排他性聚类算法,即一个数据点只能属于一个类别,而FCM只计算数据点与各个类别的相似度,即对于任意一个数据点,使用K-means算法,其属于某个类别非否即是,而对于FCM算法,其属于某个类别的相似度只是一个百分比。

基于 FCM, 又引入了核方法形成了 KFCM. 核方法的基本思想是通过非线性映射 ϕ 隐式地将输入空间 x 中的样本映射到一个高维的特征空间, 并在这个高维空间中应用线性算法, 使其等价于原空间中的非线性算法, 从而改善算法的分类能力.

2.2 图像分割的原理

FCM 与 KFCM 对图像进行分割的原理与对于普通数据进行聚类没有什么太大的区别, 就只是将普通的数据集换成了图像的灰度矩阵而已. 通过对不同灰度的像素点的聚类实现图像的分割.

2.3 FCM 与 KFCM 结果对比

我们使用磁共振成像 (Magnetic Resonance Imaging, 简称 MRI) 图像作为实验对象, 磁共振成像作为作为颅脑检查的有效手段可以形成多方向的图像, 其数据分辨率、软组织对比度高、可以提供其他成像技术难以匹敌的分辨率, 对病灶的定位更加准确, 可以提供更加准确的颅内病理和生理信息.

MRI 有不同事物序列, 最常用的是 T1 加权序列和 T2 加权序列, T1 序列产生的图与 T2 序列的图相配合, 更便于医生看清楚病变. 一般来说医学上常使用 6 张 MRI 图结合观察, 所以这里对于 FCM 和 KFCM 的聚类数设为 6, 且 KFCM 中核参数取 150. 原图为如图 5.

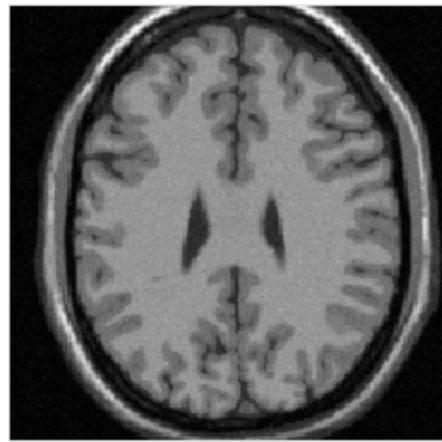


图 5

FCM 的聚类结果如图 6



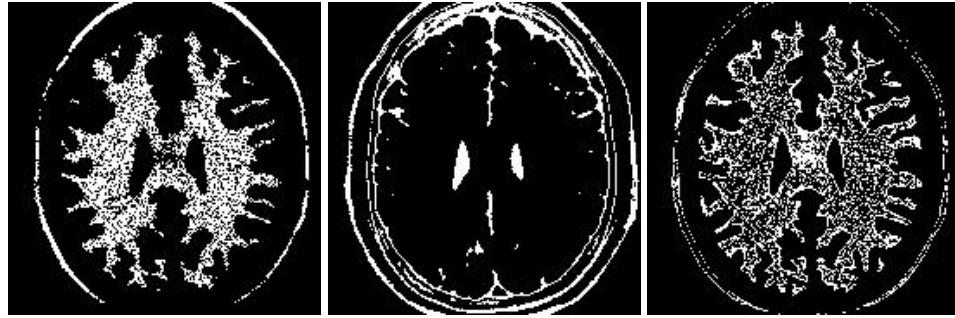


图 6: 从左到右、从上到下依次为第 1 到 6 类的聚类结果.

之后, 还使用了 KFCM 对于原图像进行了分割, 得到的结果如 图 7.

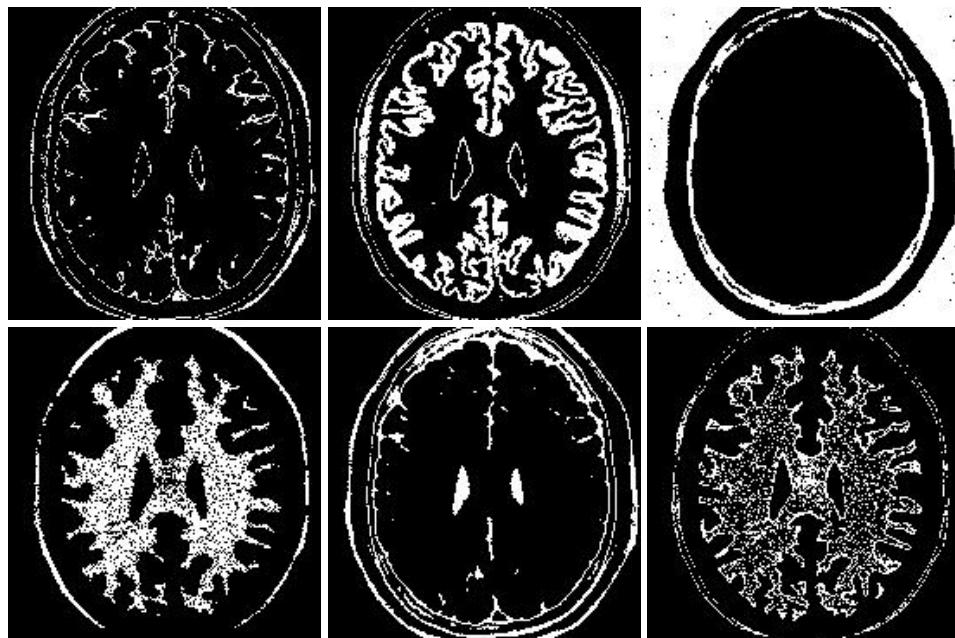


图 7: 从左到右、从上到下依次为第 1 到 6 类的聚类结果.

可以看到两个方法的聚类结果基本上无法用肉眼区别, 通过查阅资料得知需要通过比较每个像素是否正确分类来判断性能, 由于我们医学这使得匮乏和硬件限制, 所以我们的实验停止在这一步.

虽然无法切实比较两个算法的性能, 但可以看到通过 FCM 和 KFCM 我们能成功将原始的 MRI 图 “切片”, 即从原本的一张二维图片中提取出六个层次的信息, 为医生的分析提供了便利.

通过这样的一个应用, 可以看到, 输入的图片进入计算机后, 计算机识别每个像素灰度后进行聚类分析生成多层次图供以我们参考和进行医学诊断, 其实这也可以说做广义的“编码解码”过程, 在这个过程中计算机实现了“编码”、“储存”, 通过图像我们所进行决策判断的过程即为“解码”过程.

参考文献

- [1] Peter J. Rousseeuw, Marie Paule Derde, and Leonard Kaufman. Principal components of a fuzzy clustering. *Trac Trends in Analytical Chemistry*, 8(7):249–250.
- [2] Jing, Zhang, , Yuanyao, Xu, , Guohua, Huang, , and Huiju and. Palladium-catalyzed synthesis of chiral, nonracemic isopropyl arylmethylphosphinates.
- [3] Bezdek and C James. Pattern recognition with fuzzy objective function algorithms. 22(1171):203–239, 1981.
- [4] 周志华. 机器学习. 清华大学出版社, 2016.
- [5] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. The MIT Press, 2016.
- [6] 使用 k-means 算法将图像压缩 6 倍 - 腾讯云. <https://cloud.tencent.com/developer/article/1411114>.
- [7] Image compression using k-means clustering : Colour quantization. <https://appliedmachinelearning.blog/2017/03/08/image-compression-using-k-means-clustering/>.