1. 数据库说明

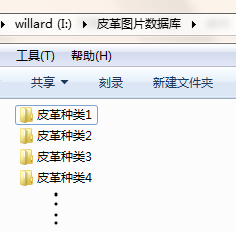


图1

算法开发时对皮革数据库的构建需做如下的说明：

1. 皮革数据库在什么量级，比如大概几万张图片；
2. 皮革种类大概有多少类；
3. 为便于后面对检索精度指标作评价，构建皮革数据库时采用图1的方式存放，即同一类皮革放一个文件夹；

说明：对于3中对皮革要求做的分类，只是为了便于后面做检索精度的指标评价，实际投入生产中的时候，无需做上面的要求。

1. 检索精度指标评价

检索精度指标按文档给出的说明是“该算法需要在给定的海量皮革或布料软体图库中精准的搜索出最相似的几张皮革或皮料图片（不大于5），…，算法精准度要求在90%以上”，这里需要对下确认：

1. “（不大于5）”是指每次给用户呈现的搜索检索时，显示的返回图片不大于5，是不是这样理解的？

图片5

图片4

图片3

图片2

图片1

查询图片

图2 用户界面

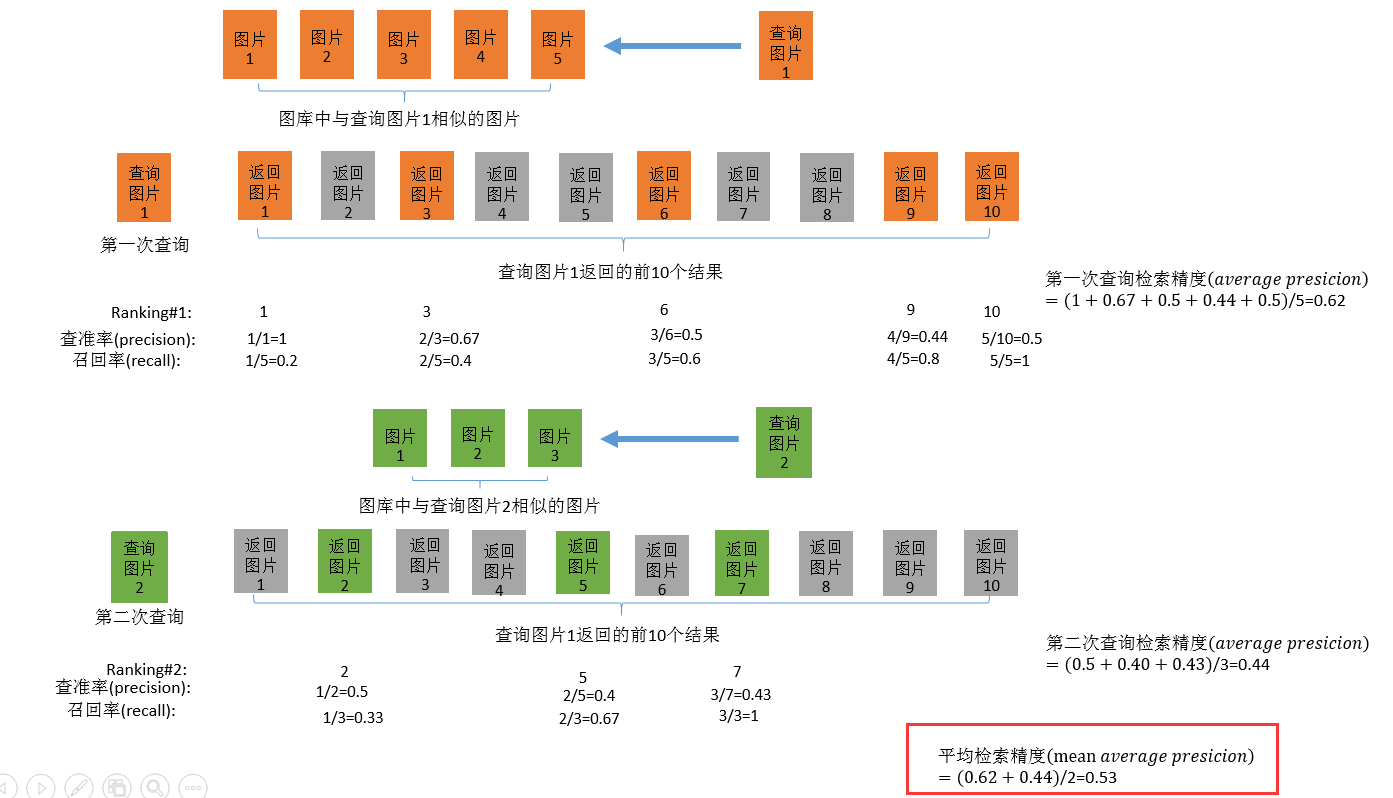
1. “算法精确度要求在90%以上”，这里的算法精确度采用的是哪种评价指标？在图像检索领域，通常采用平均检索精度MAP(mean average precision)来评价检索性能的好坏，MAP的具体示例计算如下: 

图3 平均检索精度(MAP)计算过程示意图

图3为图像检索最常用也最有效的评价指标MAP的计算过程。据此，我们来推算一下比较符合条件即用户查询可接受的MAP为多少。在5张返回的图片中，比较满意的情况是这5张图片中有3张是跟查询图片相似的，而这三种我们假设它们都排在第3、4、5这三个位置，按照图三的步骤，单次查询的AP是（1/3+2/4+3/5）/3=0.48，假设N次查询都是这种情况，则MAP=0.48\*N/N=0.48，而最理想的MAP是1，据此，在保证每次能够在返回的5张图片里如果能够保证有3张是同类型的皮革，则MAP可以达到在0.48以上。