



DIA 第四次作业： 基于BoW和VLAD的图象检索

2020年04月29日



任务说明-I

- 给定一个图象数据库，从每幅图像中提取SIFT 特征
 - 共1000张图片，每四张图片为一组相关的图象
 - ✓ 下载地址：staff.ustc.edu.cn/~zhwg/download/homework_4.zip
 - 图象编号从0开始，如果两幅图象的编号除以4后取整的值相同，则它们为相关图象
 - 利用提供的提取SIFT特征的批处理程序，得到数据库中每幅图象的SIFT特征
 - 检索精度计算方法：
 - ✓ 以一幅图象作为查询，计算查询图象与数据库所有图象特征的距离，将距离值按降序排列
 - ✓ 查看排序结果中前四个图像（与查询图象的距离最小）的编号，基于编号判断其是否与查询图象相关；统计前四个图象中相关图象的数量（最大为4），作为此次查询结果的检索精度值
 - ✓ 分别将数据库每幅图象作为查询，计算其检索精度的平均值



任务说明-II

□ 基于BoW的图象检索

- 对数据库中图象的SIFT特征进行随机采样（数量至少为码本大小的10倍），利用k-means训练视觉码本，码本大小为K。
- 利用视觉码本对每幅图象的局部SIFT特征进行量化，基于量化结果将每幅图象表达为视觉单词直方图，并进行L1或L2归一化。
- 基于BoW特征的L1或L2距离测度，计算查询图象与数据库图象相关度，并进一步计算检索精度。
- 把每一幅图像分别作为查询图象，计算平均精度。

□ 测试K=1000、2000、5000、10000、20000时相应的检索精度。

□ 测试不同的归一化方式（L1或L2）对检索精度的影响。



任务说明-III

□ 基于VLAD的图象检索

- 对数据库中图象的SIFT特征进行采样，利用k-means训练视觉码本，码本大小为K。
- 利用视觉码本对每幅图象的局部SIFT特征进行量化，并计算量化残差。通过拼接各个视觉单词下的累积量化残差，得到每幅图象的VLAD的特征表达，并进行L2归一化。
- 基于VLAD特征的L2-距离测度，计算查询图象与数据库各图象的距离，并对距离值进行降序排序；基于排序结果计算检索精度。
- 把每一幅图像分别作为查询图象，计算平均精度。

□ 测试K=8、16、32、64时，相应的检索精度。



提交时间和方式

- 提交截止时间：5月10日晚上12点前
- 提交方式：
 - 将**实验报告**和**代码文件**放到一个文件夹中，生成一个压缩文件，文件名**命名规则为：“DIA第四次作业_姓名_学号”**
 - 将以上压缩文件发到如下邮箱：ustcdia@163.com。
- 作业迟交处理
 - 如迟交 n 天，本次作业分数乘以 0.98^n