

周总结八

卢婧宇 13020022023

April 16, 2017

1 本周进展

1.1

SIFT 特征提取部分结束, 程序调试过程中具体解决的问题详见附件 2。BoW 部分在之前的学习过程中已经大体了解了它的概念, 在本周的学习中又对其进行进一步理解。BoW 模型的处理过程:

- a. SIFT 特征提取。SIFT 特征提取是求出图像的关键点信息, 包括角度, 大小以及强度。关键点, 也就是能够代表图像关键信息的部分, 这也是 Bag of words 中单词的组成。一个图像通常有很多的关键点。
- b. 聚类。将每幅图像中的关键点信息添加到词袋中, 并定义聚类中心数量 N 。然后将词袋中的关键点通过 Kmeans 算法聚类到 N 个类中。同时得到这 N 个类的中心点组成 $N \times 128$ 的 dictionary, 每个中心都可以代表这个类。
- c. 求图像的直方图。将图像的关键点信息重新放到词包中, 根据落在每个类中关键点的数量来得到图像的直方图, 大小为 $1 \times N$ 。将每幅图像进行处理, 得到图像在 BoW 模型下的特征。
- d. 图像匹配。将测试图像进行相同的处理, 同样也得到 $1 \times N$ 的特征。根据测试图像与训练图像特征之间的距离, 并将距离较小的图像作为检索的结果。其中, vlfeat 提供了 k-means 聚类和分层 k-means 聚类的轻量级的实现, 并且我已经找到 vlfeat 对它的说明。在通过从网上搜索, 我简单理解了一下 IKM 的概念。Integer K-means (IKM) 实现了整型数据的 K-means 聚类 (或者叫矢量量化)。在图像检索、识别领域, 经常会用到 Bag-of-words(BOW) 模型, 该模型对训练集图像提取特征并进行聚类, 得到固定数量的代表性特征集 (dictionary), 将测试集中提取出来的特征进行量化, 用之前得到代表性特征集中的特征 (word) 来表示, 这样每幅图像就可以表示成 bag-of-words。

1.2

对第一门课程计算导论与 C 语言基础进行了学习, 完成第六周的编程作业, 并已经将代码上传至 GitHub 上。由于操作不当我将其中一个仓库删除了, 导致之前的笔记也被清空无法恢复, 所以需要重新做笔记, 目前笔记内容很少。并且, 关于毕业论文的文档我也已经整理好上传至 GitHub 中。

1.3

本次总结是用 LaTeX 编写的，仅仅使用了一些基本的语句，用的还不好。

2 下周安排

2.1

对 BoW 模型部分进行学习，尽快可以得到程序并且进行调试。

2.2

空闲时间，继续听课进行编程学习完成作业。

2.3

继续对 LaTeX 进行学习。