**基于Caltech101数据集的模板匹配方法图像检索**

院系：国家网络安全学院

作者：2019300003038 余斌

2019300003072 闵文韬

2019300003041 李潇楠

指导教师：蔡波

摘 要

本次技术报告的目的是采用模板匹配的方法，对Caltech101数据集中16个类别的图像进行特征相似度计算和检索，并对各个方法的查全率和查准率进行计算。

关键词：图像特征；颜色直方图；颜色矩

目录

[1 研究意义 2](#_Toc20759)

[2 项目环境介绍 3](#_Toc3655)

[3 研究内容及方法 4](#_Toc608)

[3.1 基于颜色直方图的图像检索 4](#_Toc516)

[3.2 基于颜色矩的图像检索 5](#_Toc4733)

[3.3 基于Hu矩的图像检索 5](#_Toc2394)

[3.4 基于SIFT描述子和BoW算法的图像检索 6](#_Toc26780)

[4 研究结论 8](#_Toc2361)

**1 研究意义**

近年来，随着通信及多媒体技术的迅速发展，数字图像的使用变得越来越广泛。在如此广泛的应用背景下，产生了大量各式各样的图像数据库，其中有海量的图片，如何更好地应用视觉数据，实现快速查询和检索用户所需的图像信息，成为人们迫切需要解决的问题。

较早时候的数据检索方式往往依赖关键字和文本信息，需要人工对图像进行描述，不能从图像本身的视觉特征来检索。因此本文采用了模板匹配方法，通过计算图像的特征向量，取相似度高的图像作为检索结果来解决该问题。

**2 项目环境介绍**

Python3.7,opencv-python,scipy,sklean

IDE: Pycharm,Spider

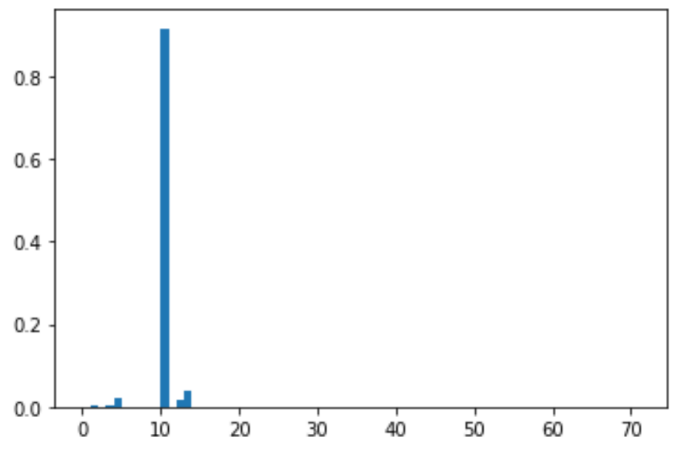
**3 研究内容及方法**

**3.1 基于颜色直方图的图像检索**

由于RGB颜色空间与人眼的感知差异很大，因此使用更适合的HSV颜色模型。先将图像的RGB值转换为HSV，再进行量化。

之后构造一维特征矢量G=9H’+3S’+V，其中整数G的取值范围为[0,71]，对图像绘制72bin一维直方图。



对所有得到的直方图数据建立KDTree，查询目标图像对应直方图矢量的近邻，即可得到检索结果。

（AK-47 总图像数98）

查全率+查准率：10/97 %

前5个检索结果：



**3.2 基于颜色矩的图像检索**

将图像的颜色矩作为特征向量。

查全率+查准率：2/97 %



**3.3 基于Hu矩的图像检索**

将图像的Hu矩作为特征向量。

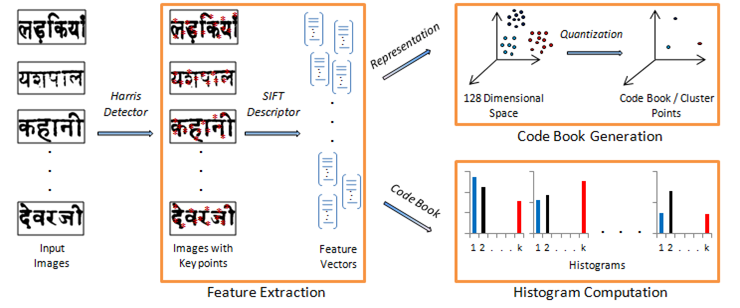
查全率+查准率：12/97 %



**3.4 基于SIFT描述子和BoW算法的图像检索**

实现步骤如下：

1. 对每幅图像计算SIFT描述子，得到x个128维向量（每幅图像的描述子数量不同）
2. 将所有向量按行合并为一个矩阵，做k-means，得到C个聚类中心。这些聚类中心也叫视觉词汇，它们的集合叫视觉词典。
3. 将描述子转换为特征向量。对每幅图像的x个SIFT描述子，计算它与聚类中心的欧几里得距离，寻找对应的聚类中心，统计得到含有C个bin的直方图，即该幅图像的特征向量。
4. 后续查询步骤同3.1-3.3。



|  |  |  |
| --- | --- | --- |
| C | 检索出的正确图像数量 | 检索结果 |
| 3 | 5 |  |
| 4 | 5 |  |
| 5 | 7 |  |
| 6 | 6 |  |
| 7 | 13 |  |
| 8 | 4 |  |
| 9 | 3 |  |
| 10 | 5 |  |

**4 研究结论**

本次研究比较了颜色直方图、颜色矩、Hu矩和SIFT算子这几种传统方法进行图像检索的结果，但实验效果并不理想。究其原因，与数据集的图像较为复杂有关，传统方法普遍无法准确定位图像中的物体，学习其中的语义，而被其他颜色和无关物体所干扰。