
视频首帧识别补充实验报告

朱文韬 1600012785 - 2016年12月12日

方法一：

计算HSV色彩空间中的色彩直方图并比较

a. Correlation (CV_COMP_CORREL)

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \bar{H}_1)(H_2(I) - \bar{H}_2)}{\sqrt{\sum_I (H_1(I) - \bar{H}_1)^2 \sum_I (H_2(I) - \bar{H}_2)^2}}$$

其中

$$\bar{H}_k = \frac{1}{N} \sum_I H_k(I)$$

N是直方图中bin的数目。

采用较为均衡的比较方法（Correlation）

优势

运算速度相对理想

光照变化影响较小

讲课人移动影响较小，精度高

劣势

对版式完全相同的两帧纯文字ppt区分能力弱

改进1：将屏幕区域等分为16份依次计算相似度，去掉最大值最小值取平均。

改进2：将屏幕区域等分为16份依次计算相似度，去掉最大两个值、最小两个值取平均。

改进后分析：

可以认为相对削弱了PPT版式的影响，能够较好地地区别出微小差异，召回率显著上升，但是精度有所降低（局部光照变化等噪声影响被放大，教师头部移动影响放大）

延伸阅读：

[HSV色彩空间](#)

[Histograms色彩直方图](#)

方法二：

SIFT

优势

特征点具备尺寸旋转不变形

劣势

速度慢，不适用于视频场景

SIFT适用于实际物体（刚体），对PPT（光影、非实物）识别不理想，不稳定

延伸阅读：

[SIFT算法](#)

[SIFT匹配](#)

方法三：

感知哈希算法

分别采用了平均值哈希、差异值哈希

差异值哈希的效果略优于平均值哈希

优势

运算速度相当快

光照变化影响较小

对文字段落区别能力强

劣势

讲课人快速移动影响、精度偏低（但在实际操作中可以考虑以较小的工作量人工筛选，获得最大效率）

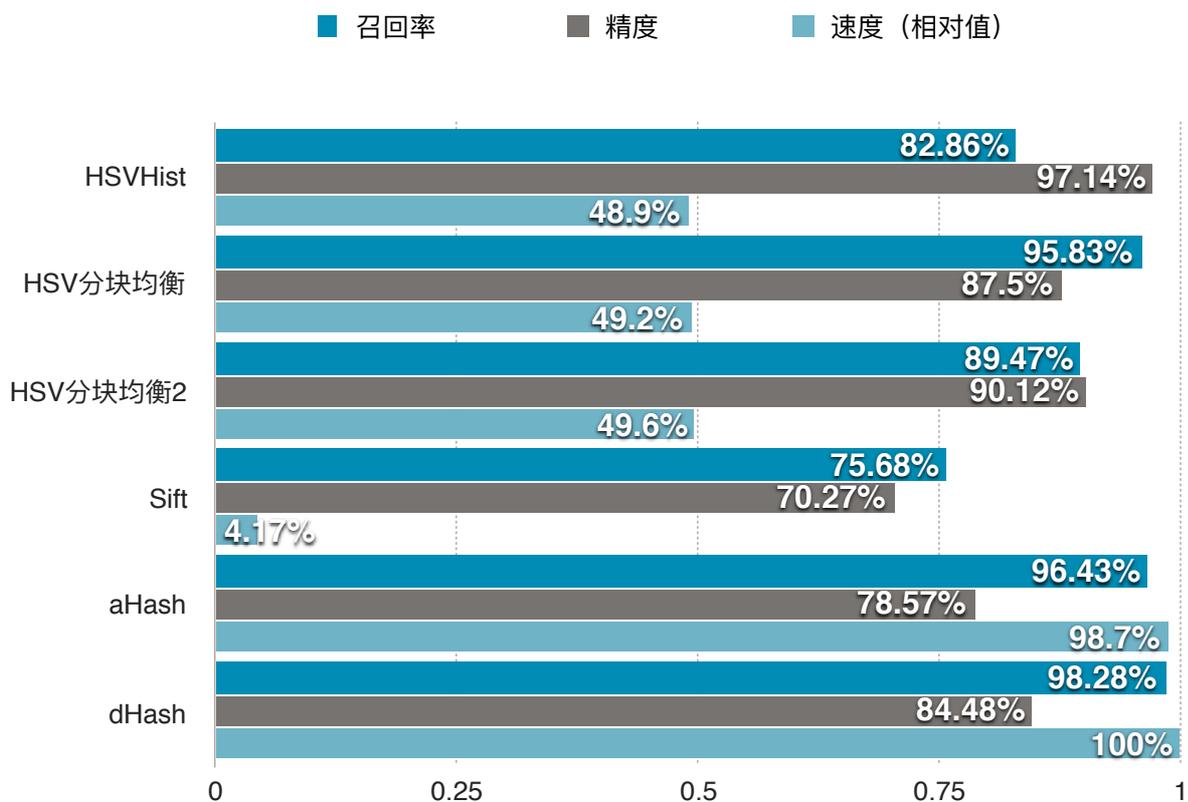
二维的图像可以分解成不同的频率成分。其中，低频成分描述大范围的信息，而高频成分描述具体的细节。在灰度图像中，亮度变化小的区域主要是低频成分，而亮度变化剧烈的区域（比如物体的边缘）主要是高频成分。感知哈希算法选择性地忽略了图像的高频信息，快速提取图像的主要信息，已被广泛应用于商标 / 图片侵权检索、以图搜图等许多功能，并有望应用于钓鱼网页自动识别、语音识别、物体运动跟踪等领域。

除PPT介绍的基础方法之外，实际应用时往往结合其他技术对图像进行预处理。比如对于人物可以先进行人脸检测，再在面部区域进行局部的哈希，或者背景是纯色的可以先过滤剪裁等等，最后在搜索的结果中还可以根据颜色、风景、产品等进行过滤。

延伸阅读：

[图像感知哈希算法及应用](#)

[基于感知哈希的图像内容鉴别性能分析](#)



***速度说明:**

dHash速度 (1) 约为: $\text{处理时间} / \text{视频长度} = 0.2$

如果仅输出切分时刻不输出图片速度还有一定提升

***更进一步的思考:**

结合dHash召回率高和HSVHist精度高的特点, 在比较函数中对两者取调和平均, 发现可以将召回率、精度均控制在95%以上。判断此讲课视频前后两段演示文稿的类型, 发现 (并通过了其他较小实例检验), dHash对于文本文档类讲课材料处理表现更好, HSVHist对于图片类讲课材料处理较好, 可以针对课程演示文稿类型自定义地选择比较函数以获得更好的适应性。