**本科生科研训练与创新创业项目**

**（创新项目）**

**中期学术报告**

|  |  |
| --- | --- |
| **项　目　名　称** | 计算机视觉辅助校园海报设计 |
| **报告名称** | 计算机视觉辅助校园海报设计 |
| **项目负责人** | 张逸睿 |
| **项目组成员** | 袁宇昊 曹诗韵 |
| **申报人电话及手机** | 15342705585 |
| **导师姓名** | 常亮 副教授 |
| **所在院系** | 信息科学技术学院 |
| **导师电话及手机** | 13264298968 |
| **填　表　日　期** | 2019年3月4日 |

**计算机视觉辅助校园海报设计**

北京师范大学 张逸睿[[1]](#footnote-1)，袁宇昊, 曹诗韵

（北京师范大学信息科学技术学院，北京 100875）

常亮 副教授

**摘 要**：视觉文本展示布局(例如，数字杂志封面、海报、幻灯片和任何其他媒体)结合了漂亮的图像和叠加的可读文本，可以获得视觉上的吸引力，以吸引用户的注意力。因此，在商业印刷出版物和在线数字杂志中，视觉文本呈现布局的设计变得无处不在。然而，手工制作美观的布局仍然对业余用户具有挑战性。本文介绍了一种通过研究美学设计原则来自动对可视化文本评估及改进的系统。通过美学设计原则，普通用户可以在制作出不成熟的视觉文本后，运行该系统，可以得到对该可视化文本的评估以及更正。该系统由SVM分类器和显著性区域的裁剪与再加工组成，包括对显著性区域的提取，泊松融合以及填孔等技术。SVM分类器综合了原图数据和显著图数据，得到了较好的准确度。布局模板由网站综合排序，相关美学论文研究调研以及咨询设计师得出，显著性区域位置放置、空间布局、大小调整等。

**关键词**：SVM分类器；泊松融合；显著性区域提取；填孔技术

校园里经常举办各类活动，如何更好地向同学们传递活动信息是活动宣传者面对的主要问题。通过向在校大学生发放问卷，我们得知校园中的喷绘/易拉宝/海报是同学们最常接触各类活动信息的渠道并且这也是同学们心中最具影响力的宣传方式之一。校园喷绘不仅作为一种有效的传播媒介，而且也承载着一定的美学价值。调查显示同学们会对设计更好的喷绘有更多的关注，并且校园喷绘的好坏相当程度上影响了同学参加活动的意愿。但可惜的是，同学们普遍反映校园喷绘的质量一般，并对校园喷绘的创意、排版、画面和主题突出性提出了更高的要求。认为经过改善的喷绘能有超过百分之二十的效果提升自己参加活动意愿的同学占百分之八十。所以，校园喷绘的改善变得很有必要。

Ying Cao[1]等提供了面板模型设置眼部路线引导以研究漫画排版问题的方法，但海报和漫画不同，漫画中的一页分区，有相对固定一系列分区的模板，对于一个区域的漫画来说，是比较容易进行处理的，而海报只有一页，篇幅比漫画的一小格大，而且对于素材的要求也更多。XuYong Yang**[2]**等人提出了一个基于显著性区域检测以及视线引导的杂志封面排版方法，不过对于海报来说，杂志封面的元素更少，只有背景和文本，而且与杂志相比海报的重点更为突出。

因此，本研究将：（1）根据网站综合排序、下载量、上传时间，相关美学论文调研数据和专业的设计师的经验得出显著性区域位置放置、空间布局、大小调整方法以及优秀素材与不优秀素材。（2）优化FT[3]算法，经过二值化裁剪出图片的显著性区域。（3）将原图与显著性区域图一起放到SVM分类器中进行训练，得到准确度较高的评估参数。（4）根据美学规则，对于评为不优秀的图，利用泊松融合[4]以及填孔技术，将显著性区域提取裁剪放到适合的位置并给出评估。

1. 技术与方法

**1.显著性区域的提取**

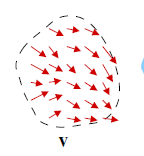
Achant[5]等人提出了基于频率的显著性区域提取的方法。图像在频率域可以分成低频部分和高频部分。低频部分反映了图像的整体信息，如物体的轮廓，基本的组成区域。高频部分反映了图像的细节信息，如物体的纹理。显著性区域检测用到的更多的是低频部分的信息，通过高斯平滑来实现对最高频的舍去。计算每个像素l,a,b的值同三个l,a,b均值的欧式距离，归一化后得到最终的显著图，之后我们对灰度图进行二值化，找到最大的连通分量，便于确定显著性区域的边界。我们爬取了超过400张包图网以及谷歌的海报，并对其进行加工。像素显著性可由公式1计算，其中，Iu为图像的平均特征，使用Lab颜色特征，后一项为像素p在高斯平滑后的Lab颜色特征，||.||为L2范式，即计算前一项和后一项在了Lab颜色空间的欧氏距离。

https://img-blog.csdn.net/20170606162423146?watermark/2/text/aHR0cDovL2Jsb2cuY3Nkbi5uZXQvcXFfMjIyMzgwMjE=/font/5a6L5L2T/fontsize/400/fill/I0JBQkFCMA==/dissolve/70/gravity/Center

公式1 像素显著性公式

**2.泊松融合的实现**[4]

通过差分方法，对待克隆图像进行梯度计算，接着计算背景图片的梯度，接着，按照规定位置，对融合图像的梯度和散度。计算待克隆区域的梯度场，利用卷积，计算梯度模长。得到融合图像的梯度后，也就是待重建图像的梯度场，接着对梯度求偏导，从而获得散度。梯度场的求取用一阶算子卷积就可以得到（比如sobel算子）。对背景图像dst做mask操作，来将Ω区域的像素置0。最后求解系数矩阵，分别求解三个通道的方程，就可以获得每个点的像素R,G,B的值。



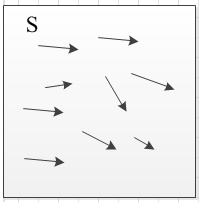
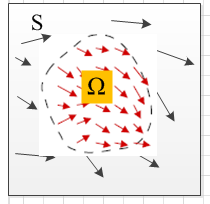


图3 待重建图像的梯度场

图2 背景图片的梯度场

图1 待克隆区域的梯度场

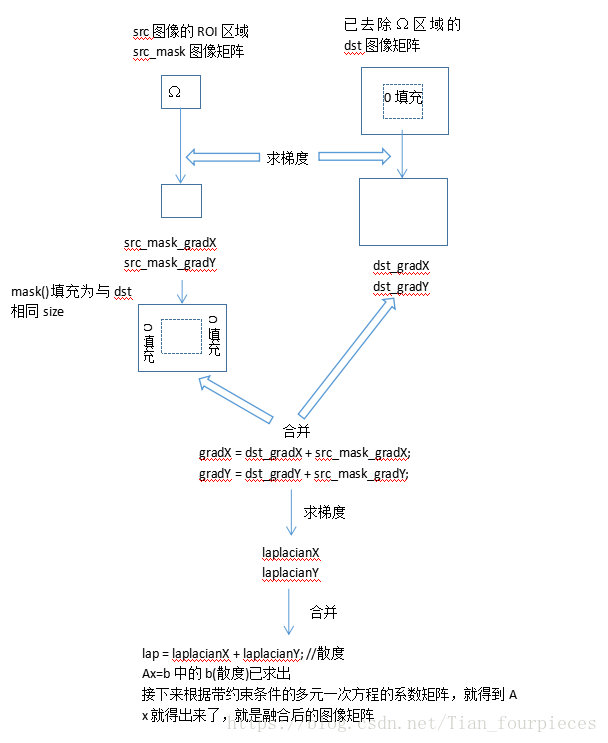


图4 算法流程

**2.SVM的调试**

（1）**训练、测试素材的标准化：**对于图像和其显著性区域图来说，svm训练时要求图像的格式、尺寸必须一致，同时形式统一的图像名字容易区别。使用python中的PIL库对图像进行格式、尺寸处理，使用os库对图像重命名、保存。

（2）**读取素材：**对于训练素材，一次性读入所有的训练素材。对于测试的素材，考虑读取测试素材的路径，当用到当前素材时再对素材进行读入。这样比一次读入所有素材占用的内存小，并且随时可以添加新的测试样本。

（3**）训练函数：**使用交叉验证法进行训练。

（4）**训练函数的接口：**训练函数是重载函数，可以接受多种不同形式的输入信息。选用的是标准的样本库、标记库输入法。

（5）**结果的记录：**程序对于每一次训练的过程，包括训练的时间、训练中设定的参数、测试的结果等都作了详细的记录。

1. **对svm分类器参数调整的结果分析**

Svm分类器的分类效果对于参数十分敏感，需要特别注意对svm分类器参数的调整。

|  |  |
| --- | --- |
| 核函数 | 从低维到高维的映射，常用的有Linear、Poly、RBF |
| degree | 针对Poly核函数，多项式的最大指数 |
| gamma | 支持向量个数的倒数 |
| cost | 惩罚系数 |
| maxn | 最大迭代次数 |
| eps | 终止判断的依据 |

表1 SVM分类器的参数

最后选用了Linear或RBF作为核心函数进行训练。

对于Linear核函数来说，没有什么参数需要调整。对与RBF核函数来说，cost和gamma参数需要调整。对于两个核函数来说，maxn和eps两个参数对于最后的训练结果影响不大，所以采用默认值。

下面给出一些具体的参数训练的结果简要做出说明：



表2 参数训练结果

训练结果1是线性核函数的结果，它没有什么需要调整的参数。

从训练结果3、4、5可以看出，在相同且合适的cost参数下，不同gamma参数对分类器的影响：gamma参数越小，分类器对于反面测试样例的检测就越差，反过来说明分类器对于反面的认定过于狭小。结合训练结果6可以看到，当gamma小于1E-6时，分类器已经不具有分类的基本效果。

从训练结果2、3、7可以看出，在相同且合适的gamma参数下，不同cost参数对分类器的影响：cost参数越小，分类器对反面测试样例的检测就越差，反过来说明分类器对于反面的认定要求更严格。结合训练结果8可以看到，当cost小于1时，分类器已经不具有分类的基本效果。

反过来可以想到的是，当gamma和cost越来越大的时候，正面测试样例的准确度会下降。所以我们可以基本确定cost和gamma的范围：

于是，在这个范围基础之上，我们再来调整cost和gamma参数。

至此，这是一个二维的参数调整，cost和gamma互相影响，所以要综合考虑。考虑使用二维平面上的三分方法来找(cost,gamma)使得总准确率f(cost,gamma)最大。

二维三分既是在分别对cost和gamma进行一维三分的基础上，把三分cost得到的两个cost分点和三分gamma得到的两个gamma分点进行组合，从而得到2\*2共4个(cost,gamma)组合，再计算出4个组合的f(cost,gamma)选出最大的f作为保留点，在保留其最大组合位置的基础上更新范围即可。

这个方法的好处是每三分后的结果f都会大于等于上一次的结果f，从而是个不减序列。不足之处是不能证明最后得到的结果一定是最大的，但是可以证明一定是极大的。

和其他svm分类器相比较：我们的gamma参数远远小于其他的svm分类器，例如和Wang等人给出的划分平面点集的分类器来说，其gamma取值为0.1，比我们的gamma参数小了好几个数量级。原因在于我们使用的是高维的数据，是一副图像完整的RBF值，所以需要更多的支持向量去分类，否则就会出现没有完整应用图像数据的情况。

参考文献

1. Cao Y, Lau R W H, Chan A B. Look over here: Attention-directing composition of manga elements[J]. ACM Transactions on Graphics (TOG), 2014, 33(4): 94.
2. Yang X, Mei T, Xu Y Q, et al. Automatic generation of visual-textual presentation layout[J]. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 2016, 12(2): 33
3. Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection[J]. 2009.
4. Pérez P, Gangnet M, Blake A. Poisson image editing[J]. ACM Transactions on graphics (TOG), 2003, 22(3): 313-318.
5. Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection[J]. 2009.

1. **作者简介：**张逸睿（1998-），女，湖北十堰人，本科学生。 [↑](#footnote-ref-1)