《数字图像处理基础》课程实践

**结题报告**

|  |  |  |
| --- | --- | --- |
| **项目名称** | ： | **用数字图像处理玩2048游戏** |
| **姓名** | ： | **MOU YUAN YAP** |
| **学号** | ： | **517030990013** |
| **联系电话** | ： | **18621807623** |
| **电子信箱** | ： | **mouyuan4598@sjtu.edu.cn** |
| **合作人姓名** | ： | **游竣皓、ZHICHENG LEE** |
|  |  |  |

2020年1月

**请参照下列提纲填报**

（一）研究内容综述

1. 对比开题报告，说明各个研究内容的完成情况

本次项目主要分为3大研究内容，即针对截取后图像的预处理（感兴趣区域提取）、判别各个区块的数字（目标检测）、以及游戏的执行（游戏策略算法和机械臂操作）。

感兴趣区域提取：在没有利用UP\_CUT变量的情况下，可以直接获取屏幕界面并且利用透射变换将其变换成固定的尺寸以及正立的角度。

目标检测：在成功获取感兴趣区域的前提下，各个区块的数字可以精准被读取并且将它们存在二位数组中（下称“游戏矩阵”）。

游戏的执行：解决项目开始时机械臂容易出现的“暴走”和误触问题。此外，游戏算法也成功实现可以至少达到512，并且很大机会可以出现1024的数字。

1. 未完成的研究指标应说明原因

感兴趣区域提取：以上所述完成情况必须是在摄像头没有完全拍到空白处的前提下才得以实现。原因是采用的方案是寻找画面中最大的面积轮廓。

目标检测：结果能否成功取决于环境的光照程度，以及平板的屏幕上是否会出现反光。如果出现反光，数字的读取结果将会不稳定。主要原因在于我们的目标检测是基于HOG+SVM。训练样本不足够，没有包含大量各种环境的数字样本。

游戏的执行：游戏无法达到2048主要原因在于我们人类本身玩游戏时都不能确保每一次都可以取得2048的成绩，因此无法设计出可以取得2048成绩的算法。是否可以达到1024也存在少许运气成分，如果情况不理想将只能取得512的成绩。

（二）研究方案

请详细撰写各个研究内容的研究过程、原理、实验设计与实验分析

研究过程中，我们以上述三大研究方向分工合作，同一时间进行各自的工作项目。

本人主要负责感兴趣区域提取这个模块。

一开始的想法是猜想为了更加稳定以及完整地获取轮廓可能需要进行大量的图像处理，例如形态学中的开运算、闭运算、形态梯度、顶帽和黑帽等处理。不过经过多次的调试之后得出的结果是只要利用Canny边缘检测和高斯滤波（Gaussian Blur）就能完美地利用findContours获取摄像头画面中的各种轮廓。

得到轮廓后可以利用contourArea计算各个轮廓所圈起来的面积。显然，平板电脑屏幕将会是拍摄画面中面积最大的部分。从中确定了该轮廓四个角的坐标。

之后就可以利用透射变换将随机角度拍摄到的游戏画面变换成一个固定长宽的正立图像。只要有了四个角变换前以及变换后期望的坐标就可以计算出旋转矩阵。

我们的创新点就是通过透射变换固定游戏画面，进而达到更加轻松地分析各个数字的坐标以及特征。我们改善了老师在课程介绍时所提到的利用盒子解决画面固定的方案。由于游戏感兴趣部分的提取在项目早期就得到了接近完美的成果，因此我也参与了部分目标检测的工作。

最初我们尝试使用模板匹配和SIFT关键点匹配对比游戏画面上的数字和模板的数字。对于彩色图，透视变换之后会造成像素的丢失，即数值为3的像素点经过1/2的缩小之后其像素值会成为被丢失的1.5。此时需要进行周围像素的插值处理。不过在实验过程当中，这个操作无法让问题得到很好的解决，仿射变换后的SIFT匹配效果仍然非常不理想。

为此我们采取了另一个方案，就是在进行CLAHE图像增强和去噪之后，再进行二值化处理。 对比OTSU和 Adaptive Threshold(自适应二值化)，OTSU更加容易被方块的背景颜色和数字的颜色所影响，如果两者颜色相似将会出现两种颜色混为一体的情况，不过它的表现严重依赖于周围光照以及摄像头的角度所影响。Adaptive Threshold则可以弥补OTSU以上的缺陷，不过得到的数字的粗细程度却始终不一。最终，OTSU可以得到更好的二值化后的效果并且在SIFT匹配中取得更好的表现。经过长时间的调试，我们发现OTSU的鲁棒性非常弱，我们无法很好地通过调试匹配点数量的参数确定各个数字，最后选择放弃使用SIFT匹配。

我们转向进行助教推荐的HOG+SVM的方法进行数字的检测。在这个环节我的工作范围是数字样本的提取。就之前所提到的游戏画面的坐标不会由于摄像头变化而受到影响，第一个左上角数字的坐标是(45,155),每个方块的长宽分别是84和96，方块之间的间隙为5。通过这个已有的参数就不需要手动截取样本，样本的提取变得非常方便。我们以20个空格样本，10个“2”至“512”样本，5个“1024”样本，构建特征向量矩阵，行数等于样本个数，列数等于HOG描述子维数。训练SVM后保存模型利用其对于所读取的特征进行分类，判断数字的类型。

对于游戏的操作算法，最大值趋向角落，将数值按序排列。最大值移到角落可以避免其干扰到较小数的合成，并将数值按序排列可以使合成具有连贯性。其缺点是算法基于最大值在某个角落的基础上进行编写，一定程度上缺少健壮性。按序排列的实现（假定最大数向右下角趋向），判断操作是否有效（划动操作后游戏内数值是否会发生改变）。其基本原则为优先右划，次优先下划，再次为上划，最后考虑左划。无需判断具体数值的具体位置变化，只需判断较大数的个数变化。

对于机械臂的控制，自写点击函数的延迟避免了暴走和误触的情况。2048的操作简单只需要编写向上向下向左和向右的操作即可。只需让触笔去到游戏界面内在复位回到初始位置。由于userGameController的函数时每一帧都在运行的，这会导致笔还没有完成复位就截取了错误的图像。为此，我们的解决方案是让这个函数延迟120个单位。

（三）研究成果

对比开题报告中的研究指标，详细论述研究成果

经过了本次课程实践得到了以下总结：

一、如果要得到画面中的轮廓，Canny边缘检测的效果最好

二、仿射变换会造成像素丢失，尽管可以利用插值的方式解决，可是对于高精度的模板匹配仍然有所不足。

三、SIFT匹配对于二值化后的效果更加好

四、OTSU的二值化效果很好，可是过于受光照影响且对于画面中颜色相似的区域难以区分；Adaptive Threshold弥补了OTSU的缺陷，却因为是自适应所以无法调节二值化的程度

五、检测某种数字利用SVM的方法鲁棒性最好，可是却缺乏利用图像处理的知识。

（四）研究总结

1. 研究中存在问题、建议及需要说明的问题

本次课程实践的结果最大的问题是周围的光照状况;环境过于亮或者暗、顶灯位于摄像头上方，都会严重影响实验结果。此外，为了让摄像头可以在任何角度捕获游戏界面而使用的寻找画面中最大的轮廓也有一个问题：必须将摄像头调整到屏幕部分的面积是最大的。

1. 在本次课程实践中的感想与体会

《数字图像处理基础》这门课程是我本学期最喜欢的课程。非常感谢老师愿意提供机会让我们以实践的方式更加了解课程的内容并加以运用。我发现我比起单单读理论，更加了解了各种数字图像处理的原理。

1. 对本次课程实践的意见与建议

我觉得老师可以在上课时更加着重于指导与实现实践内容的课程。同时，也希望助教可以及时纠正我们研究的方向，避免我们浪费时间在不正确的道路上，毕竟这只是个2学分的课程，感觉上要花上的时间有一点太多了。