

P2P结题论文

姓名：章程

学号：2015010912010

专业：通信专业

基于OpenCV的基本图形识别

章程

(电子科技大学通信学院 学号: 2015010912010)

【摘要】

本文介绍了一个基于 OpenCV 的基本图形识别程序。实现过程包括利用 matlab 生成大量随机样本用以训练神经网络,通过边缘检测、傅里叶算子提取基本图形的特征向量,神经网络增量学习的简单实现,利用 QT 实现简单的 GUI 界面。

关键词: 图像处理; OpenCV; matlab; QT; 神经网络; 图像识别; 机器学习

1 形状特征描述

为了对图像中的基本形状进行识别,必须对形状进行描述,提取出形状的某些特征。为了能区分不同形状的物体,保证识别的正确性,我们希望这些特征相对不同的形状有较大的差异,而对同一物体,对其在图像的不同位置、大小和方向的变化保持不变。即该特征描述具有平移、旋转、尺度变换不变性。形状特征描述是图像识别的关键,特征提取的好坏直接决定了目标识别的效果^[1]。程序使用傅里叶描述子对形状的轮廓进行描述,介绍傅里叶描述子后,再介绍必要的预处理操作。

1.1 傅里叶算子

傅里叶描述符是描述闭合边界的一种方法,其基本思想是用物体边界的傅里叶变换作为形状描述,利用区域边界的封闭性和周期性,将二维问题转化为一维问题。该方法仅适用于单封闭曲线,而不能描述复合封闭曲线^[2]。

傅立叶描述子(FD)是物体形状边界曲线的傅里叶变换系数,它是物体边界曲线信号的频域分析的结果。根据傅里叶变换的性质,傅里叶描述子与形状的尺度、方向和曲线的起始点位置有关。因此需要对傅里叶描述子进行归一化,使用具有平移、旋转和尺度不变性的归一化傅里叶描述子来识别物体的形状^[1]。

《基于矩特征傅里叶描述的目标形状识别》^[1]已对傅里叶描述子的平移、旋转和尺度不变性进行了论证。在具体实现中,需要考虑两个细节:

- 离散傅里叶变换得到的频谱低频部分同时包含了正频率和负频率,频谱为离散周期函数,因此正频率和负频率的低频部分对应离散傅里叶变换所得序列的前几项和末几项。由于输入的序列为虚数,因此变换后的频谱不是偶函数,其正负频率对形状的描述同等重要,在实现中应该去除直流分量和归一化参考项,再截取频谱的前几项和末几项。
- 轮廓的点坐标有顺时针和逆时针的区别,方向不同会导致频谱的反转,若按照论文^[1]中的方法 $\|Z(k)\|/\|Z(1)\|$ 对频谱进行归一化,可能得不到正确的结果。因此在归一化前,还需要对轮廓的顺逆时针进行一次判断。

具体到基于OpenCV的实现过程,首先用findContours()提取图像中的闭合轮廓,并以x为实数,y为虚数,将轮廓点坐标转换为虚数序列,再用dft()得到轮廓的频谱,利用abs()计算得到幅频谱,再做顺逆时针判断、归一化操作,最后去除直流分量、归一化参考量,截取n项低频幅值得到最终的傅里叶描述子,封装成fourierDescriptor()函数。

长度为20的傅里叶描述子效果如图 1所示,可以看到长度为20的低频频谱已经可以很好地还原轮廓的形状,但是值得注意的是表示幅度的傅里叶描述子相对于复数频谱会损失相位信息。

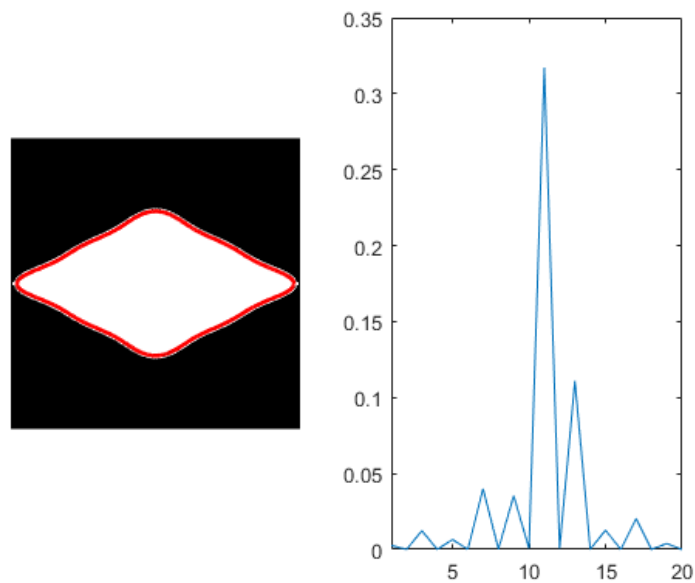


图 1 频谱经反变换所得轮廓（左）和其傅里叶描述子（右）

1.2 边缘检测

在利用OpenCV的findContours()函数取得轮廓坐标序列前，需要对图像进行预处理，得到闭合轮廓或连通区域的二值图像，如图 2 Canny边缘检测、图 3 OTSU阈值分割所示。

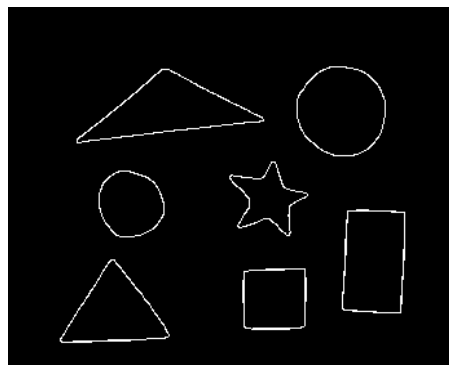


图 2 Canny 边缘检测

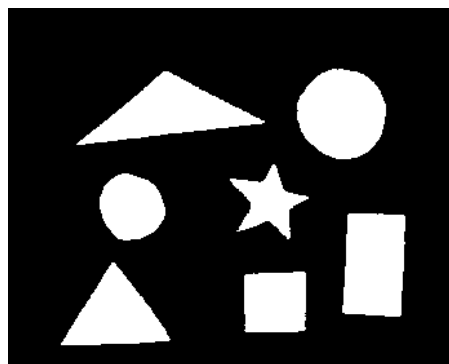


图 3 OTSU 阈值分割

程序尝试了OTSU算法来获得连通区域，局部自适应二值化、Canny边缘检测来获得闭合轮廓，三种方法在不同环境下各有所长，可供需要选取：

- Otsu是大津于1979年提出的一种阈值分割方法，它是建立在一幅图像的灰度直方图基础上的，依据类间距离极大准则来确定区域分割门限^[3]。在OpenCV中可以通过

threshold()调用，效果如图 3 OTSU阈值分割所示。

- 局部自适应阈值则是根据像素的邻域块的像素值分布来确定该像素位置上的二值化阈值。在OpenCV中可以通过adaptiveThreshold调用，适当调小窗口的大小，可以实现边缘检测的效果。
- 边缘检测的方法有很多，主要利用边缘处梯度较大的特性。常用的边缘检测算子有Roberts Cross算子、Canny算子、Laplacian算子，实验中使用Canny算子能取得较好的效果，如图 2 Canny边缘检测所示。

获取到轮廓的点坐标序列后，不仅可以利用傅里叶描述子对其进行描述，还可以利用OpenCV提供的moments()函数计算轮廓包围区域的矩特征，进而得到质心位置，并叠加文字标注。OpenCV还提供了contourArea()函数用以计算轮廓包围的面积，在提取轮廓后，可以遍历计算面积，并提出面积较小的连通区域，避免噪声、杂物的干扰。

2 神经网络

2.1 样本准备

神经网络的训练需要大量样本，因此利用matlab编写了自动生成基本形状图片的程序。实现过程中需要考虑两个细节：

- 一些具有偏差的基本形状可以近似认为符合某类的标准。如长宽比接近1的矩形、邻边不严格垂直但接近正方形，都可以归类为正方形。
- 除去归一化大小减少的一个自由度，同一类形状具有多个自由度的变化。如三角形具有高、顶点的水平位置两个自由度。

通过matlab随机生成的7类形状，每类样本达2000个，充分保证了神经网络的泛化能力。部分生成的样本如图 4 部分生成的样本 所示。

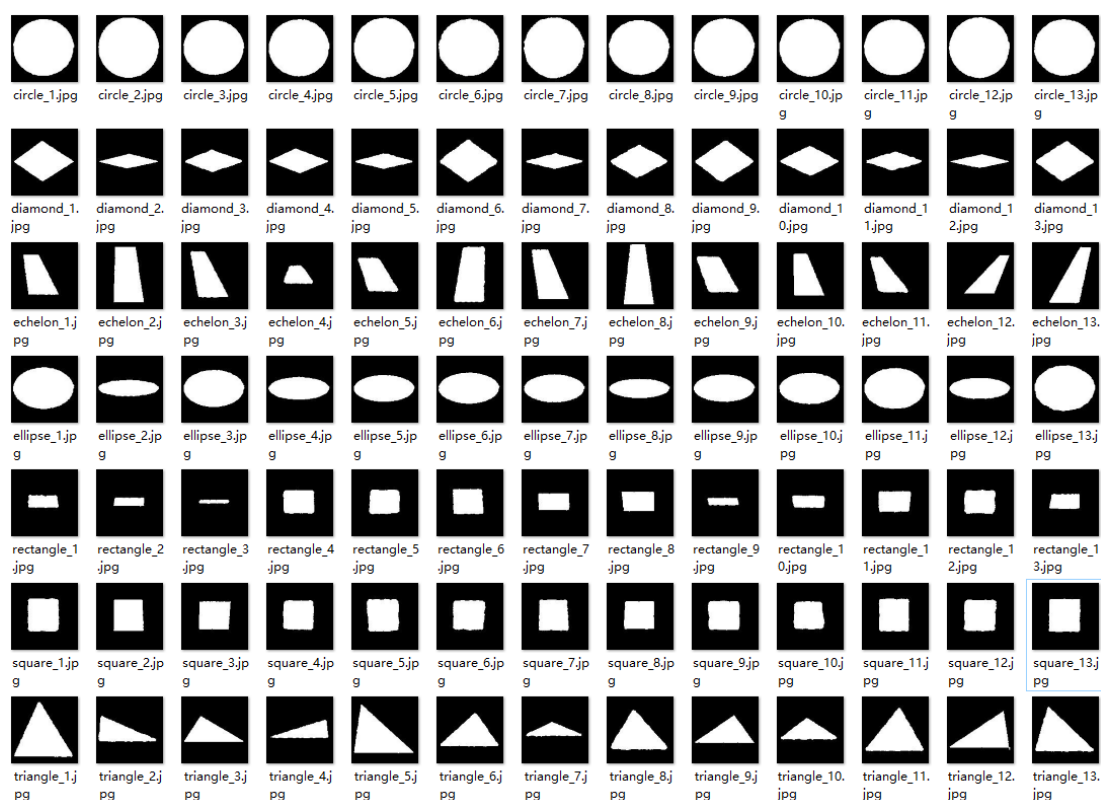


图 4 部分生成的样本

2.2 神经网络训练

OpenCV中神经网络的实现依赖于ANN_MLP类。但是该类包含有大量纯虚函数，不能直接实例化为对象，只能通过指针和ANN_MLP::create()方法使用该类。在实现过程中，我们希望神经网络的样本和结构能方便地保存读取，新增样本时能方便地做出更改和进行微调，因此程序实现了一个MyANN_MLP类，扩充了ANN_MLP类所不能实现的上述功能。

对于神经网络的训练，依赖于ANN_MLP类的train()方法，训练样本以Mat类型传入，Mat类可以用OpenCV提供的FileStorage类存储为xml文件和再读取。

2.3 增量学习

增量学习是对人类学习过程的模仿，其优势主要体现在当新数据集到来时，系统能在已有知识的基础上对新数据集进行学习，获得新知识的同时也不会遗忘有用的旧知识。由于增量学习的这一特性，目前主要是通过增量学习这个技术手段来完成在线学习任务。^[4]

目前OpenCV上没有增量学习的实现，但是直接将新的特征向量添加到训练样本中，再微调神经网络，也可以作为一种简单的实现方法。具体实现中需要考虑一下四个问题：

- 在增加已有类别样本时，新添加的样本相对于原有的数量众多的样本太少，无法达到纠错的效果。这时可以提升样本的权重或新增样本的数量。
- 大量重复的新增样本可能导致神经网络的训练过拟合，降低其泛化能力，易受噪声的影响。这时可以在新增同一样本时，向样本加入适量噪声。
- 新增类别的情况不仅要求增加样本的数量，还需要改变样本目标输出的结构，以及神经网络的输出层节点数。
- 对于新增的样本，其标签可以认为是准确无误的。在微调神经网络的过程中，固定的新增样本权重和固定的BP迭代次数可能不够使神经网络正确分类新增样本。在设计纠错代码时，应在每次训练结束验证新增样本是否分类正确，否则要增加样本数量，并继续微调神经网络。

3 GUI实现

3.1 HighGUI下的轮廓叠加与高亮

要实现纠错功能除了需要实现添加样本并微调神经网络的方法，还需要实现鼠标选取轮廓的方法，获取被选取的轮廓信息。

OpenCV为多边形轮廓提供了pointPolygonTest()函数以判断某一点是否在轮廓内，结合HighGUI提供的鼠标事件监测setMouseCallback()，为鼠标设置CV_EVENT_MOUSEMOVE事件的响应，获取鼠标的坐标并遍历所有闭合轮廓，找到鼠标所在的最小轮廓，以醒目的颜色绘制出来，即可达到高亮鼠标下轮廓的效果。再结合CV_EVENT_LBUTTONDOWN事件的响应，即可得到被选中的轮廓。高亮轮廓的效果如图 5 所示。

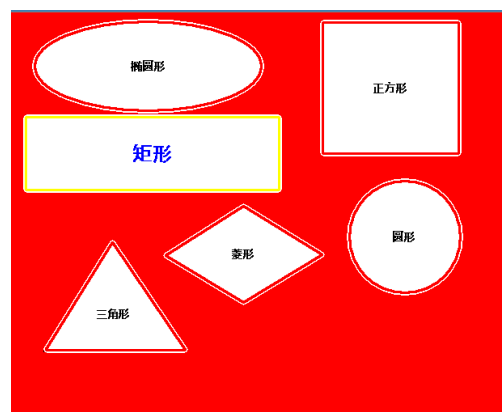


图 5 高亮鼠标下的轮廓

3.2 QT下的功能布局

至此已经实现了形状识别、神经网络训练、形状纠错三个功能,每个功能区的参数设置、功能选取多而杂,如果不使用GUI而依赖于命令行,就会影响使用体验,增加程序复杂度。为此,本程序包含了一个基于QT的简单GUI实现。

“形状识别”功能区包含了选取图片、摄像头输入功能按钮,预处理方法、摄像头ID、面积阈值三个参数设置。“重新训练神经网络”功能区包含了重置按钮,隐含层层数、层节点数两个参数设置。“形状纠错”功能区包含了添加类别、微调神经网络功能按钮,类别选择、新增类别名称、微调神经网络时每次增加样本的步长以及向轮廓添加噪声的强度几种参数设置。同时为了方便用户保存更改和还原更改,还增加了“保存或还原神经网络和样本”功能区,为了显示形状统计信息、操作提示,还增加了testview控件。最终的UI布局效果如图 6 UI布局效果所示。

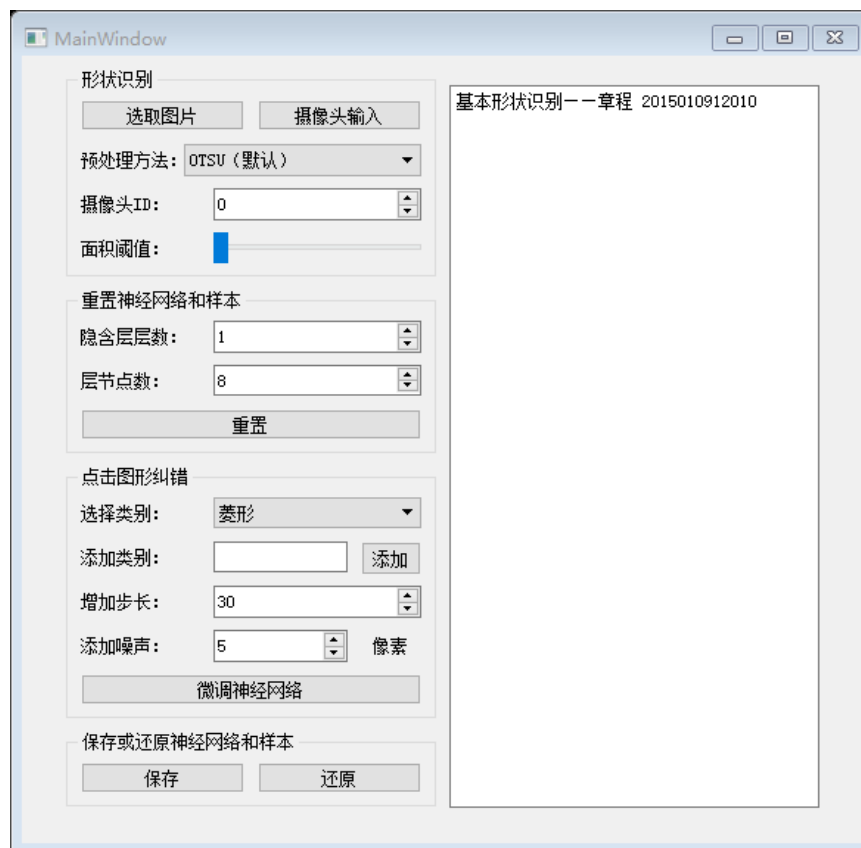


图 6 UI 布局效果

4 总结

本文介绍了一个基于OpenCV的基本图形识别程序。实现过程包括利用matlab生成大量随机样本用以训练神经网络,通过边缘检测、傅里叶算子提取基本图形的特征向量。经过试验证明傅里叶描述子是描述轮廓形状的一种较好的特征,利用pca算法对20维的傅里叶描述子进行降维后得到的3维散点图(图 7 可视化三维散点图)也显示,傅里叶描述子描述的各类形状具有良好的区分度。

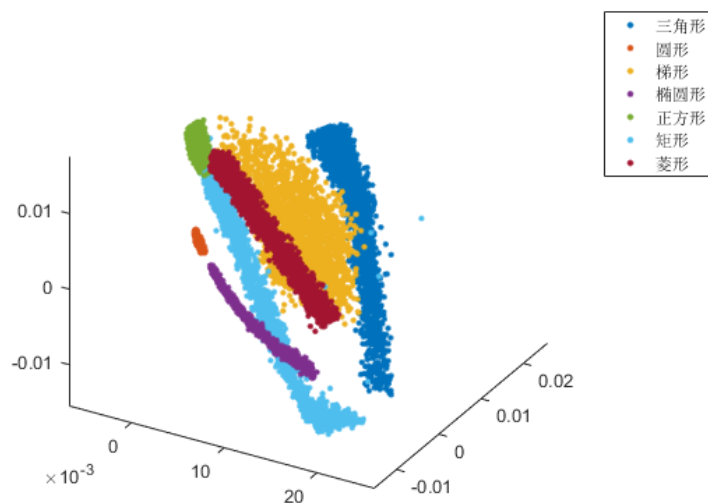


图 7 可视化三维散点图

本文还介绍了神经网络增量学习的简单实现，利用 QT 实现简单的 GUI 界面，最终完成的基本形状识别程序具有效率高、功能全面、准确度高等特点，不仅可以识别图片（），还可以对摄像头输入的视频进行识别，效果如图 9 摄像头输入图像识别所示。



图 8 图片形状识别

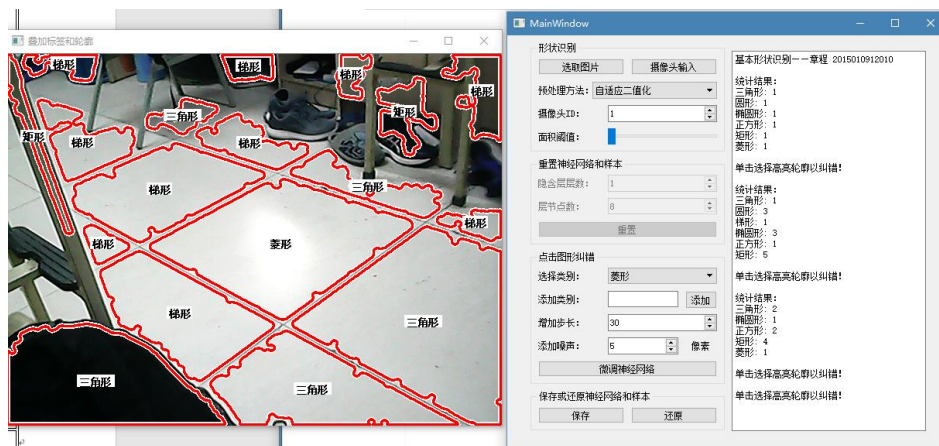


图 9 摄像头输入图像识别

参 考 文 献

- [1] 张金华. 基于矩特征傅里叶描述的目标形状识别[D]. 上海交通大学, 2009.
- [2] Zahn C T, Roskies R Z. Fourier Descriptors for Plane Closed Curves[J]. IEEE Transactions on Computers, 1972, 21(3):269-281.
- [3] 胡颖. Otsu算法的研究及改进[J]. 枣庄学院学报, 2009, 26(5):68-72.
- [4] 卞桂龙. 在线学习的集成分类器研究[D]. 浙江大学, 2014.