目录

[I. 绪论 2](#_Toc39594229)

[A. 选题背景 2](#_Toc39594230)

[B. 选题意义 2](#_Toc39594231)

[C. 本文工作简述 2](#_Toc39594232)

[II. 相关研究 3](#_Toc39594233)

[A. 综述 3](#_Toc39594234)

[B. 基于改进Faster RCNN的ROI提取与结算商品识别 3](#_Toc39594235)

[C. 结合基于模板匹配的图像分割和深度学习的货架商品识别 3](#_Toc39594236)

[D. 基于CNN的自然场景文本检测与识别 3](#_Toc39594237)

[III. 本工作：基于字符分割与模板匹配的方法 4](#_Toc39594238)

[A. 分析 4](#_Toc39594239)

[B. 数据集 4](#_Toc39594240)

[C. 预处理 5](#_Toc39594241)

[D. 字符分割 7](#_Toc39594242)

[E. 模板匹配 8](#_Toc39594243)

[F. 系统输出 8](#_Toc39594244)

[IV. 结果 9](#_Toc39594245)

[V. 总结与展望 9](#_Toc39594246)

[A. 总结 9](#_Toc39594247)

[B. 后续展望 9](#_Toc39594248)

[VI. 参考文献 10](#_Toc39594249)

# 绪论

## 选题背景

随着信息技术的快速发展和普及，新型的识别手段被应用到各种各样的场景中去，代替人获取环境信息，驱动后端的算法执行，提升行业的自动化水平。基于图像的识别由于与人类的感知特性相近，能够很好地融入各式生活和生产场景，目前已经成为领域的一大研究热点。

## 选题意义

超市作为线下零售的主要渠道，贩售的商品种类繁杂，数量较大，交易量也相当可观，由此产生了结余统计、查漏对账方面的需求，然而上述特征也给人工作业带去了较大的工作量，挤占了员工部分工作时间来进行统计；手工统计再转电脑输入存档也增加了产生错漏的可能，存在一定的重复劳动，进一步降低了工作的效率。

综合上述观察，考虑将图像识别技术引入超市结余场景，解放生产力，提高自动化与智能化程度，对于社会生产具有积极的意义。

## 本文工作简述

在这些考察之上，本文对图像处理技术在超市商品结余统计这一场景下的运用做了研究，针对“数量”、“距离过期的剩余时间”（以下简称剩余时间）两大主要的统计需求编写了程序。主要工作如下：

1. 使用基于最大极值稳定区域（MSER）的算法来提取含文字的ROI并将其分割出来。
2. 针对超市复杂的应用场景和商品包装的多样性，在HSV空间求取了图像背景光照并去除；通过检测包装的直线边缘进行了倾斜校正；编写了自适应的阈值化程序，能够处理浅色背景深色文字和深色背景浅色文字等不同情况。
3. 对前两步获得的图像进行进一步分割得到单个字符，基于GIST描述子进行识别。

经过能够通过图片识别商品二维码及批量地识别生产日期，进而得出所识别商品的数量和“剩余时间”并记录。经过检验，在保证输入图片清晰度的前提下，程序对于方形外包装的商品具有良好的识别效果，对环境光照、对象倾斜以及不同的字体和背景颜色等干扰因素具有较好的鲁棒性。

# 相关研究

## 综述

目前相关的研究更多的集中在购物结算方向上，通过引入图像识别技术来实现无人收银、自动结算，因此注重特征提取与对象识别，少有针对数量、日期统计的研究。因为背景、目标的复杂性和多样性，多使用神经网络等方法，需要先在具有一定规模的数据集上进行训练，且对运算能力有较高的要求。

## 基于改进Faster RCNN的ROI提取与结算商品识别

该方法(1)改进了Faster RCNN网络中用于提取ROI的RPNet，使其具有良好的非特异性，可以预先使用公开的数据集进行训练，学习到对象位置的先验知识，然后直接迁移到新建的数据集上进行商品位置的标注。将标注位置作为Grabcut算法的输入，能够切割出对象，用来制作和增强数据集以进行网络训练。整个算法在系统测试中具有较高的召回率和精度。

## 结合基于模板匹配的图像分割和深度学习的货架商品识别

此研究(2)注意到了超市结余统计的需求，使用基于SIFT特征算子的模板匹配算法，对包含多个商品的货架图像和单个商品的模板图像进行匹配，对匹配区域做分割得到单个商品图像，再使用深度学习进行识别，对于背景和重叠干扰较大的货架场景能够比较准确地定位商品。但该方法对不同商品，使用人工抠图制作针对性的掩模来确定特征区域，可迁移性相对较差，制作掩模的工作量也较大。

## 基于CNN的自然场景文本检测与识别

商品外包装对于时间变化相对稳定，故关于时间的信息可以通过对外包装上的文字的识别来实现。该文采用改进的MSER方法来辨识多变环境中文本的位置，通过非最大值抑制来过滤掉一部分干扰区域，最后训练CNN网络来进行文字识别(3)。

# 本工作：基于字符分割与模板匹配的方法

## 分析

结余统计的目标不同于购物结算，除了数量和种类，剩余时间也是一项重要的信息，方便及时挑出临近一定期限（以下称促销期限）的商品进行促销以及时售出，一来确保顾客不会选购到太过临近过期期限的商品，二来也避免超市必须处理过期商品造成的损失。

在获取图片的方式上，不同于结算消费场景下顾客所购商品一般稀疏的散在结算台上，为了节约空间，货架上的商品密集摆放，有比较严重的遮挡，商品外包装的形状颜色也和剩余时间相关性很小，如果直接对整个货架进行广角拍摄，基于视觉的方法很难获取数量和剩余时间的有效信息。进一步的观察分析，剩余时间的计算主要利用商品的生产日期和保质时间，种类和生产日期可唯一的与商品条形码对应，生产日期则一般在生产线的最后由自动化的器械直接印在外包装上。

绝大多数超市目前尚未形成自己的编码体系来标记每件商品，即所谓“一件一码”，而商品零售用条形码使用国际通行的无含义EAN编码，只与品种简单对应，不含有任何进一步的信息，而网上开放的条形码查询要么收费，要么使用验证来限制程序抓取，因此不使用调用外部接口来查询获得商品种类的方法，而是事先将商品条码和对应种类、保质期和促销期限信息作为系统变量录入，针对超市结余统计的应用场景，可以直接从商场的电子名录中导入。

基于如上分析，本方法决定采用采用比较窄的拍照视角，主要识别获取不固定的生产日期和商品条形码，先拍摄一张条码的图片来识别商品种类，接下来拍摄生产日期，直到下一次拍摄条码为止都认为识别的是那一种商品，通过记数图片内有效的日期来计算商品数量，直接输出包含统计信息的电子表格和是否应加入促销的提示信息。

## 数据集

原图片见附带文件夹pic,匹配模板图片保存在mask文件夹中。

说明：数据集为自制，1-19.jpg为超市实际光照条件下使用手机拍摄，12-14以及20-21五张使用一般画质拍摄，除此之外均使用高清晰度模式拍摄，20-21三张为在一般室内条件下使用一般画质拍摄，覆盖不同的商品数量、种类和角度以及光照条件。匹配模板图片为在Word中使用不同字体截图后处理制作。图片与商品的对应关系见下表。

|  |  |  |
| --- | --- | --- |
| 图片编号 | 对应品名 | 备注 |
| 1-3、20 | 百梦多咖喱（柔和醇香） | 生产日期 |
| 4 | 杨天嫩鱼晶 | 生产日期 |
| 5 | 毕阿婆桥牌腐乳 | 生产日期 |
| 6-8 | 潮香村意大利面番茄肉酱 | 生产日期 |
| 9-11 | 潮香村黑椒菲力牛排 | 生产日期 |
| 12-14 | 潮香村黑椒沙律牛排 | 生产日期 |
| 15 | 潮香村黑椒菲力牛排 | 条码 |
| 16 | 潮香村黑椒沙律牛排 | 条码 |
| 17 | 潮香村意大利面番茄肉酱 | 条码 |
| 18 | 毕阿婆桥牌腐乳 | 条码 |
| 19 | 杨天嫩鱼晶 | 条码 |
| 21 | 百梦多咖喱（柔和醇香） | 条码 |

## 预处理

1. 倾斜校正:

[校正后的灰度图，校正后的彩色图] = tilt\_corrc(待处理图片在系统路径的字符串)

将彩色图像灰度化后，使用自适应二维维纳滤波（wiener2）去噪，该滤波器具有选择性，可以更好的保留图像边缘，使用edge函数‘horizontal’模式求取图像与水平方向夹角小于45度的边缘，求出最长边

图1. 倾斜校正

缘的水平夹角作为图中商品的歪斜角度，旋转原图和灰度图完成校正。（参考自(4)，有改动）。

1. ROI提取:

保留ROI区域其余取零的掩模图 = MSERmain(灰度图)

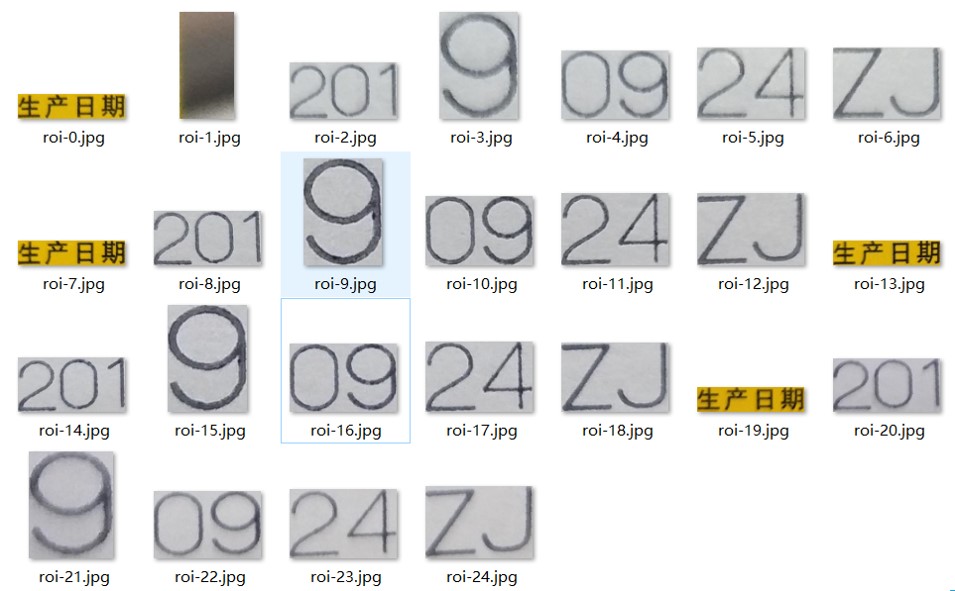
文本内部灰度变化较小，而其与背景间灰度对比较大，符合MSER区域定义(3, 5)。该函数对图像使用MATLAB内置的detectMSERFeatures()来检测灰度图的MSER区域以定位文本区域，conComp\_analysis()和f\_conComp\_analysis()两函数对二值化的MSER模板进行连通域分析和面积检测，滤除明显的非文本区域。最后将MSER区域置一，其他区域置零，和灰度图相乘得到掩模后的图像。（参考自(5)，有改动）。

1. ROI切割：

[切割出的行数，行灰度图元胞，对应彩色图元胞] = roiRowcut(掩模，对应彩色图)

[ROI区域总数，ROI彩色图元胞] = roicut(切割出的行数，行灰度图元胞，对应彩色图元胞)

包装上的文字分布比较规整，可以按行、列切分，去除背景，将图片划分的更小也可以减小后续处理的运算量。此两个函数搜索掩模后灰度图中的非零值检测ROI的坐标，roiRowcut()对图片按行进行搜索，roicut()对行搜索结果再进行按列的搜索，最后按得到的坐标到对应的彩色图像中去进行切割。

图2. 提取的ROI

1. 光照补偿：

光照补偿后图像 = lightcomp(待处理图像)

超市照明条件良好，但过强的光照反而容易造成照片部分区域过曝，影响二值化效果。该函数将图片转到HSV空间表示，对亮度分量进行多尺度高斯平滑再加权相加，求出图像的亮度背景，使用二维gamma函数进行亮度矫正。（代码参考自(6)，原文献为(7)）

1. 二值化：

二值化后图像 = mythre(待处理图像)

商品外包装设计较为自由，故存在文字和背景深浅色不定的情况，mythre()在使用graythresh()函数和去噪的基础上增加了一个判断环节，若二值化后白色部分像素绝对数量大于黑色部分，则认为现在的白色部分应为背景，即认为像素占比大的为背景，将二值化后图像全部统一为黑底白字的形式。

## 字符分割

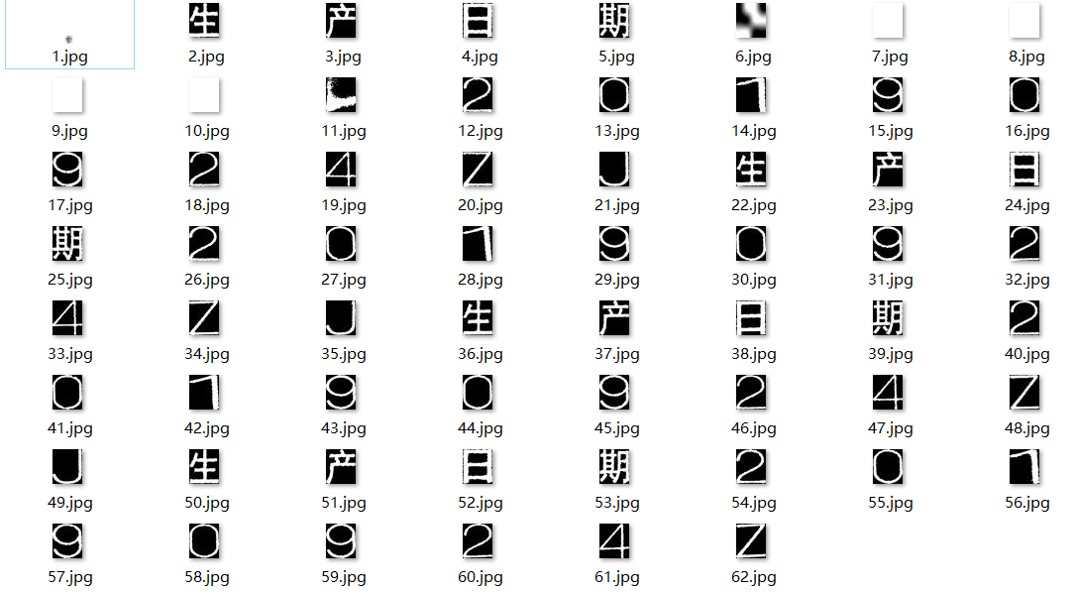
[行数, 行元胞] = rowcut(二值图)

[字符总数，字符元胞] = refine\_wordcut(行数, 行元胞)

[字符总数，字符元胞] = mser\_wordcut(待处理数, 待处理图元胞)

[字符总数，字符元胞] = myshape(待处理数, 待处理图元胞)

同前述ROI切割相似，通过分别按行按列搜索非零值并取最大值得到字符的方形外廓，进行切割和图像大小归一化，得到单个字符的等大图片。mser\_wordcut则主要针对字符在水平和垂直上都有重叠、无法分开的情况，使用MSER算法划定出不同字符的区域进行分割。myshape根据图片的形状特征，如长宽面积等去除明显不含数字的图片。

图3. 二值化与分割结果

## 模板匹配

matched = GISTmatch( shape\_num, shaped )

经过上述过程切割和二值化后的图像几乎只包含数字的形体信息，因而图像的GIST（空间包络）特征即能够充分地表达数字的特征。模板图片保存在文件夹mask下，paragener脚本可以读取该文件夹下的图片、生成模板特征向量组并以“gist2.mat”形式保存在data文件夹下，供匹配函数运行时调用。匹配结果保存在matched变量中。（参考自(5)和(8)，有改动）

## 系统输出

codecheck；checkandwrite；

两个函数分别用于检查EAN码和生产日期以及剩余时间，checkandwrite会生成消息窗口提醒哪一件商品需要加入促销，并且把统计信息累计的写入到data文件夹下的统计表.xlsx文件中。

# 结果

使用数据集中的图片来模拟了一次结余统计的状况，以下是统计后得到的表格，由于搜集数据时，上次结余已过去了一段时间，没有得到促销中商品的样本，故通过调整表格中第四种商品的剩余时间界限来模拟了需要促销的状况。可以看到由于倾斜校正算法的限制，程序对于方形外包装的商品（即下图右表中前四种）具有良好的识别效果，但对于柱状和圆形包装的商品（后两种）则没有效果，。下图左侧是促销相关的提示信息。



# 总结与展望

## 总结

以项目为核心实践了学习到的各种理论和算法，使所学更加牢固。对于图像处理的原理和流程有了一定的认识，除此之外也这个过程也是一种对于编程和调试能力的训练，对于这方面的提升也有很大的益处。

## 后续展望

初步做了将基于图像的识别引入超市结算统计场景的研究，但在以下方面还存在问题及待改进之处：

1. 整个程序目前还是基于MATLAB平台的，并没有实际地配置到应用场景中去进行测试，也没有讨论程序跨平台迁移的可行性，如执行时间和算法实现等。未来可以尝试在手机平台上进行配置，实现实用化。

2. 基于MSER的ROI检测无法区别包装上其他文字、日期数字和条码，提取出的ROI存在较多无用区域，增加了后续识别算法的复杂度和执行时间，整个程序对于计算能力的要求较高，目前尚不适用于手机等移动平台。

3. 基于边缘检测的倾斜校正对于非矩形包装的商品几乎没有效果，部分商品上印刷的生产日期比较模糊或使用断点格式，程序对这部分商品识别效果不是很好，适用范围受限。

2和3都可以考虑使用更加智能的算法来解决，如使用神经网络等。

# 参考文献

1. 胡正委. 基于深度学习的超市商品图像识别方法研究 [硕士]: 中国科学技术大学; 2018.

2. 白洁明 耿朱, inventor一种结合模板匹配与深度学习的商品种类检测方法. 中国2017.

3. 刘业鑫. 基于 CNN 的任意分布的场景文本检测与识别方法研究 [硕士]: 哈尔滨工业大学; 2019.

4. 桂哥317. 车牌图像倾斜校正算法的MATLAB实现 [Internet]. CSDN博客2018. [cited 2020]. Available from: <https://blog.csdn.net/qq_15971883/article/details/80537133?depth_1-utm_source=distribute.pc_relevant.none-task-blog-BlogCommendFromBaidu-2&utm_source=distribute.pc_relevant.none-task-blog-BlogCommendFromBaidu-2>.

5. dawnminghuang. matlab实现MSER（最大极值稳定区域）来进行文本定位 [Internet]. 博客园2015. [cited 2020]. Available from: <https://www.cnblogs.com/dawnminghuang/p/4738108.html>.

6. baolinq. 去光照不均匀matlab 算法和代码放出 [Internet]. CSDN博客2017. [cited 2020]. Available from: <https://blog.csdn.net/baolinq/article/details/78756322?depth_1-utm_source=distribute.pc_relevant.none-task-blog-BlogCommendFromBaidu-3&utm_source=distribute.pc_relevant.none-task-blog-BlogCommendFromBaidu-3>.

7. 刘学杰. 基于二维伽马函数的光照不均匀图像自适应校正算法. 北京理工大学学报.v.36;No.252(02):85-90+108.

8. Oliva A, Antonio T. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. International Journal of Computer Vision.42(3):145-75.