

基于编码特征的二维码检测方法研究

刘阳 13307130167

复旦大学 计算机学院

I. 摘要

在当今智能设备大规模普及的情况下，由于二维码信息密度高，信息类型丰富，使得其应用场景越来越广阔。如何快速有效地检测和识别二维码，成了一个应用前景很广，实用性很强的课题。本报告先从二维码检测的研究历史和研究现状开始，接着介绍二维码的结构和编码。之后着重研究二维码检测方法和其实现，并讨论不同方法的优缺点。最后是实验和讨论总结部分。

II. 关键词

计算机视觉，图像处理，二维码，Max-Min算法，Canny算子，轮廓检测，透视变换，OpenCV

III. 引言

i. 问题陈述

本报告主要依据二维码的编码特点，研究二维码的检测方法。虽然二维码的编码规范有助于实现快速检测的算法，但是真实场景中，二维码取景时往往伴随着光照不足，摄像设备分辨率低，图像倾斜，变形，扭曲，甚至部分被遮挡的情况。而我們不单单满足于规范二维码的检测，而是要挑战复杂摄像条件下二维码的检测。

具体挑战分为：1. 光照不足引起的图像二值化效果不理想问题，2. 摄像设备分辨率低引起的图像边缘检测问题，3. 二维码取景时的噪声问题，4. 二维码定位模块的位置检测问题，5. 多二维码分割问题，6. 变形图像的变换问题。

ii. 应用场合

由于二维码具有信息编码密度高，信息类型丰富，版本众多，易于定制化等等特点，二维码在生活生产中具有广泛的应用场景。而实现在复杂场景下的二维码检测方法，则能够帮助二维码应用到更多场

合中，如光照不足的矿业，使用低端摄像设备的流水线生产行业等等。

iii. 研究历史

针对二维码检测中的主要环节，我们依次介绍相关方面的研究历史。

iii.1 二维码图像的二值化问题

应用场景中采集的二维码图像是RGB模式的三色图像，为了减少之后图像处理的计算量，加速二维码检测算法，我们首先需要将二维码图像转化为灰度图。摄像设备在不同光照条件下采集到的图像灰度水平不同。对于二维码来说，明确地分辨出对应像素点是黑还是白，是便于之后二维码定位的关键步骤。因此能否有效地将二维码二值化是一个很重要的问题。

对图像进行二值化分为全局阈值方法和局部阈值方法。

全局阈值方法中最简单最通用的方法是固定阈值，将某一灰度以下的像素全部设为黑色，而其他像素设为白色。而另一种效果较好的方法是找出图像中的最大值和最小值，然后将中点作为阈值。典型的方法是OTSU算法（最大类间方差法）。OTSU的中心思想是最好的阈值应该使得被阈值分开的两组的方差加权和达到最大[1]。而Kittler的实验[2]也表明，当图像灰度直方图呈现双峰时，使用最大类间方差法效果较好，而当图像中目标与背景的大小之比较小时，该方法的效果较差。

Bernsen是局部阈值方法的典型算法。它通过定义考察点的邻域，并由邻域计算模板来实现考察点灰度与邻域的比较，它较适合解决光照不均的问题，而影响二维码图像读取的最严重问题也正是光照不均问题，但Bernsen算法也有速度慢，对干扰比较敏感和易导致伪影的缺点，特别是伪影问题，如果伪影现象严重，不仅会给图像带来大量的干扰，

严重的还会导致后续译码的失败[3][4][5]。这是因为Bernsen算法以局部窗口内最大、最小值作为考察点的域值。当局部窗口内无目标经过时，个别噪声点就会引起阈值的剧烈变化。另外，背景灰度的非均匀性也可能影响到局部阈值。

Shih-Chang Hsia等人提出一种自适应的补偿背景光的方法[6]，可以将其用于二维码二值化前的预处理。

iii.2 倾斜纠正

倾斜的二维码需要倾斜纠正。根据图像边缘可以获得二维码的倾斜角度，然后将图像旋正。边缘是图像最基本的特征，所谓边缘是指图像周围像素灰度有阶跃变化或屋顶状变化的像素的集合。经典的边缘检测算法是构造对像素灰度级阶跃变化敏感的微分算子，常用的边缘提取算子有Sobel边缘算子，Laplacian算子，Canny算子等等[7]。边缘检测算法得到的往往是断续的、不完整的结构信息，对噪声较为敏感。为了有效抑制噪声，在作边缘检测前，必须对图像进行平滑处理。另外，使用微分算子进行边缘检测的计算量仍然很大。

在二维码中，检测倾斜角度应用较多的是Hough变换和Radon变换两种方法。

Hough变换：利用Hough变换计算倾斜图像角度是广泛使用的一种方法。Hough变换把直线上的点的坐标转换到过点的直线的系数域，对于直线信息明显且干扰小的图像，Hough变换效果比较明显[7]。但在实际使用中，会受到图像周边因素的干扰，或者条码线条比较粗，Hough变换会出现峰值扩散。对高噪声下的二维码，Hough变换效果不佳。并且Hough变换的计算量较大，不能满足实时译码的需要。

Radon变换：Radon变换是一种将数字图像变换为在某一指定角度射线方向上投影的方法[7]。在一定范围内对二维码边缘图像求出最大积分的值，然后求出对应于最大积分值的角度，即是条码的旋转角度。该算法能够提高识别的抗噪性，但是运算速度受到条码旋转角度范围和步进角度影响。

不论是Hough变换还是Radon变换，中心思想都是先计算图像倾斜的角度，然后使用点运算将图像旋转纠正。对于仅发生倾斜而没有发生形变的条码纠正效果较好，但对于采集图像过程中易发生的形变问题就束手无策了。

iii.3 条码定位

根据二维码的编码特点，大多数的方法都是首先找到二维码的四个顶点，然后采用几何变换处理图像的形变问题。几何变换的关键点是控制点的选择。控制点越准确，变换结果越准确。

Ohbuchi Eisaku等人提出对二维码区域的4个顶点定位采用多条直线试探扫描的方式[8]。但是这种扫描对图像中的噪声非常敏感，而且右下角顶点的获取精度不高。

在空间变换的函数选择上，大部分方法使用了双线性变换或者透视变换模型[9]。其中双线性变换较简单，但对非线性的扭曲却无能为力，不能解决图像的透视扭曲。当图像发生透视失真时，识别的准确率将在一定程度上降低。

在空间变换中的像素插值问题上，一般有最近邻插值、双线性插值、高阶插值等方法[7]。在这方面，大部分方法使用了双线性插值法。

iv. 现状

目前主流的算法大多基于纹理特征[10]，频域分析[11]和机器学习[12]等方法，来检测定位二维码。

IV. 二维码结构和编码

二维码可以分为行排式（堆叠式）二维码和矩阵式二维码。

行排式/堆叠式二维码形态上是由多行短截的一维条码堆叠而成。行排式/堆叠式二维码的编码原理是建立在一维条码基础之上的，按需要堆积成多行。它在编码设计、校验原理、识读方式等方面继承了一维条码的一些特点，识读设备与条码印刷和一维条码技术兼容。代表性的行排式二维码有Code49，Code 16k，PDF417等。

矩阵式二维码以矩阵的形式组成，在矩阵相应元素位置上用“点”表示二进制的“1”，用“空”表示二进制的“0”，有“点”和“空”的排列组成信息编码。矩阵式二维码是一种原理和方法与行排式二维码完全不同的条码系统。行排式二维码和一维条码的编码都是对条码黑白相间条空的宽度进行调制，而矩阵码是对条码整个编码区域内的点进行编码，所以矩阵码有比行排式二维码高得多的信息密度。行排式二维码只是在形式上像二维码，而本质上完全属于一维条码，所以有人也称行排

式二维码为1.5维条码，而矩阵码才是真正的二维条码。代表性的矩阵式二维码有QR Code，Code One，Data Matrix等。

由于QR码在日常生活中应用最为广泛，本篇报告着重研究QR码的检测。

i. QR码的结构

QR码由正方形方格组成，组成正方形阵列。它由编码区域和包括定位模块，分隔线，修正模块，对齐模块在内的功能图形组成。功能图形不能用于数据编码。符号的四周由空白区域包围。QR码的编码示意如图1。



Figure 1: QR码结构

QR码的最小组成单元是方格，白色代表二进制“0”，黑色代表二进制“1”，如图2定义。

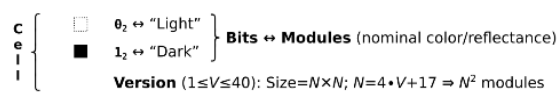


Figure 2: QR码的最小单元

图1中的功能图形分为分隔线，定位模块，对齐模块，修正模块等，如图3所示。

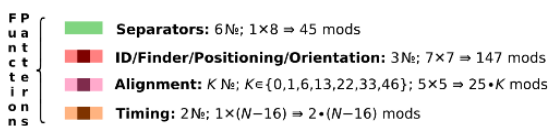


Figure 3: QR码的功能图形

图1中的编码区域分为格式信息，版本信息和数据内容，如图4所示。

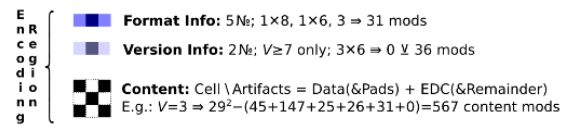


Figure 4: QR码的编码区域

各个模块的具体说明如下：

(1).符号版本和规格

QR码符号共有40种规格，分别是版本1 版本40。版本1的规格为21*21个方格，版本2的规格为25*25个方格。以此类推，版本N的规格为 $(21 + 4 * (N - 1))^2$ 个方格。

(2).定位模块

定位模块包括3个位置相同的位置探测图形，分别位于二维码图形的左上角、右上角和左下角。每个位置探测图形可以看作是由3个重叠的同心正方形组成，分别为7*7个黑色方格、5*5个白色方格，3*3个黑色方格。位置探测图形的模块宽度比为1 : 1 : 3 : 1 : 1。而在二维码其他地方遇到类似图形的可能性极小。因此，识别出组成定位模块的3个位置探测图形，就可以明确地确定二维码的位置和方向。这也是本次研究的核心思想：利用二维码的编码特点做二维码的检测。

(3).分隔线

每个定位模块和编码区域之间有宽度为1个方格的分隔线，全部由白色方格组成。

(4).修正模块

修正模块是垂直和水平方向一个方格宽的一列和一行，由黑色白色方格交替组成，其开始和结尾都是黑色方格。

(5).对齐模块

对齐模块可看作是3个重叠的同心正方形组成，由5*5个黑色方格、3*3个白色方格以及位于中心的一个黑色方格组成。对齐模块的数量由QR码的版本号决定，版本2及以上的二维码都有对齐模块。

(6).编码区域

编码区域包括数据内容，纠错码，版本信息和格式信息。

(7).空白区

空白区为环绕在二维码四周的4个方格宽的区域，一般为白色。

ii. QR码的编码

由于QR码的编码对于检测二维码几乎没有帮助，所以在这里只是简单介绍。

QR码通过RS码（Reed-Solomon）编码来实现纠错。在各块的纠错码字生成后，把数据码字和纠错码字排列成最终位流序列。

QR码编码还需要一个掩模的步骤。掩模的目的是均衡地安排黑色与白色模块，以及尽可能地避免位置探测图形的位图“1011101”出现在二维码的其他区域。掩模不能用于功能图形，用多个矩阵图形连续地对已知的编码区域的模块图形（格式信息和版本信息除外）进行异或操作。对不同掩模图形的结果计分，选择得分最低的掩模方案。

最后将格式信息和版本信息放入QR码的相应区域，完成QR码的编码。

整个过程的示意图5如下。

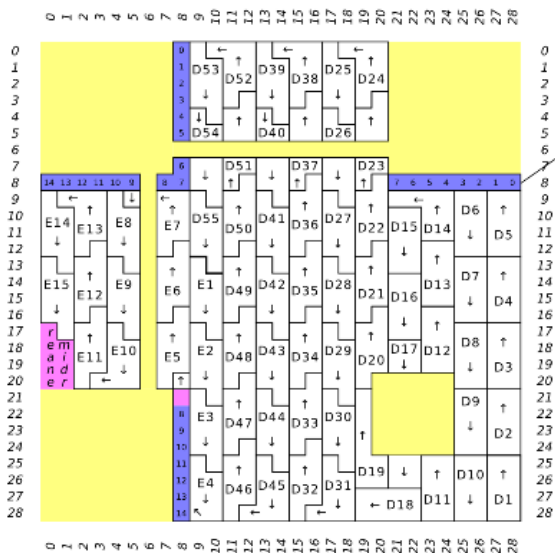


Figure 5: QR码的编码顺序

其中“D”开头的格子表示数据段，“E”开头的格子表示纠错段，浅紫色区域表示格式信息和版本信息。

V. 方法

有别于其他类型的检测问题，如车牌检测和人脸检测，二维码检测有其独特的检测特征——二维码的编码特征。

二维码在设计过程中引入的定位模块，修正模块，对齐模块等编码模块都有助于快速准确地检测到二

维码的位置。基于二维码的编码特征，我们实现了两个二维码检测的方法，第一个相对简单直接，第二个针对复杂摄像环境下，对第一个方法做了改进。

i. 第一种方法：基础方法

之所以称为基础方法，是因为这个方法直接利用二维码的编码特征，目的是为了从一幅简单的图片中定位二维码的位置。

基础方法的检测步骤如图6所示。

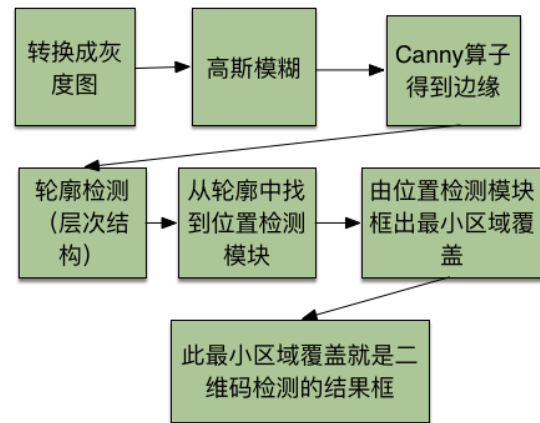


Figure 6: 基础方法的检测步骤

i.1 转换成灰度图

为了减小计算量，一般将彩色图像转换成灰度图像。灰度图像与彩色图像一样反映整幅图像的整体和局部的色度和亮度等级的分布和特征。

图像的灰度化处理可以用两种方法来实现：(1).以每个像素点的R,G,B三个分量的平均值作为该点的灰度值。(2).根据YUV（亮度，色度，色温）颜色空间和RGB空间的对应关系进行转换，其中Y分量的物理意义是亮度，可以代表图像的灰度等级。利用公式1可以将彩色图像转换成灰度图像，灰度级别为256，灰度图像中的像素为0 255的亮度值。

$$Gray = 0.299 * R + 0.5866 * G + 0.1145 * B \quad (1)$$

实验采用第二种方法来灰度化原始图像。

i.2 高斯模糊

对于灰度化的图像，我们采用高斯模糊来去除图像中的噪声，方便之后的轮廓检测。高斯滤波算法

中，分布不为零的像素组成的卷积矩阵与原始图像做变换。每个像素的值都是周围相邻像素值的加权平均。原始像素的值有最大的高斯分布值，所以有最大的权重，相邻像素随着距离原始像素越来越远，其权重也越来越小。这样进行模糊处理比其它的均衡模糊滤波器更高地保留了边缘效果，这样就有利于我们之后的轮廓检测。

高斯模糊之后的图像如图7(b)所示。

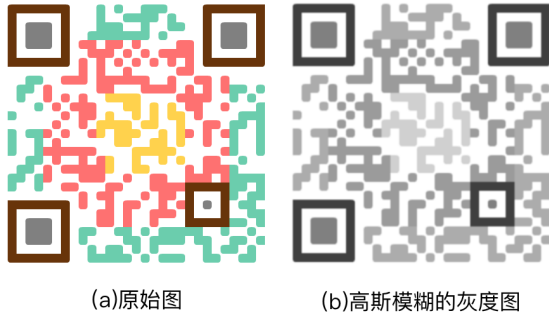


Figure 7: QR码的原始图和模糊灰度图

i.3 边缘增强

Max-Min算法 通常输入的图像可能包含大量的边缘纹理和噪声信息，Max-Min差分操作[13]能有效减少一些较小的噪声，同时目标区域的效果。Max-Min差分操作的基本表达式2如下：

$$D(x, y) = \max\{f(x+i, y+j)\} - \min\{f(x+i, y+j)\} \quad (2)$$

其中， $0 \leq x \leq w, 0 \leq y \leq h$, w 和 h 为输入图像的高和宽。

$$\max = \max\{f(x+i, y+j)\} \quad (3)$$

$$\min = \min\{f(x+i, y+j)\} \quad (4)$$

其中， $0 \leq i \leq 3, 0 \leq j \leq 3$, \max 和 \min 分别表示输入图像中一个窗口大小为 4×4 的区域内像素值的最大值和最小值。

经Max-Min操作后二维码区域被较好地突显出来，同时部分噪声也能够得到抑制。这种方法使提取的目标区域，即黑白相间的二维码区域的相邻像素间显示出较高的差异性，这种特性为后续二维码的分割提取提供了很好的保证。

Canny算子 边缘提取以及边缘增强是图像处理中的基本方法，Canny算子[14]增强效果明显。在

二维码检测定位的应用中，目标的边缘信息是非常重要的特征之一。图像边缘保留了原始图像中相当重要的部分信息，同时又使总的数量减少，这正符合二维码特征提取的要求。

Canny算子能够尽可能多地标识出图像中的实际边缘，漏检真实边缘的概率和误检非边缘的概率都比较小。其检测到的边缘点的位置距离实际边缘点的位置最近。因此，引入Canny算子作为本实验边缘提取的方法，用于找到二维码的最优边缘。经过Canny算子增强边缘之后的图像如图8所示。

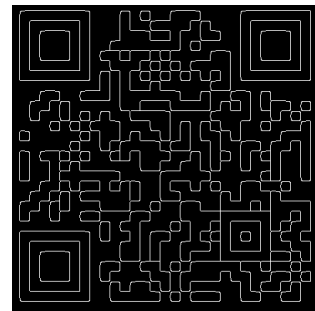


Figure 8: Canny算子增强边缘后的图像

i.4 轮廓检测

在上一步得到边缘信息之后，接下来我们通过检测轮廓来找到二维码中的定位模块。

在这里我们使用了OpenCV提供的轮廓检测方法：“findContours”。OpenCV中对轮廓采用层次结构的表示方式，如图9示意。由于这里一条

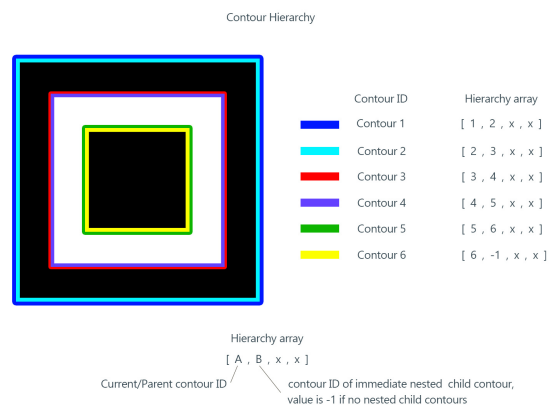


Figure 9: OpenCV中轮廓的表示形式

边缘会有内外两条轮廓，所以对于简单的图像，

只要遍历轮廓的层次关系，找到嵌套层数大于等于5的轮廓，就是我们要找的定位模块。图10(a)-(c)即为从图像中找到的定位模块。

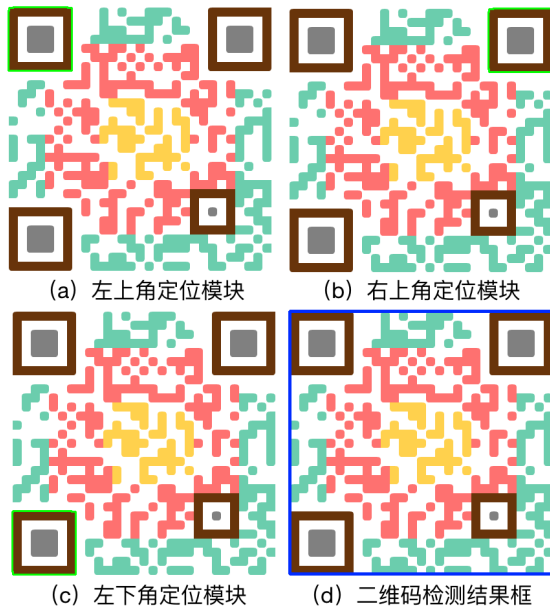


Figure 10: 检测到的QR码定位模块和检测结果框(绿色标记为定位模块，蓝色标记为检测结果框)

i.5 最小区域覆盖

在规范的二维码中，一旦检测到了三个定位模块，那么整个二维码的结果框就自然而然地形成了。在这里我们使用了OpenCV提供的最小矩形区域方法：“minAreaRect”和包围盒顶点方法：“BoxPoints”。最小矩形区域方法用来得到覆盖上述三个定位模块的最小矩形，从而将这个最小矩形设置为二维码检测的结果框。包围盒顶点方法用来得到最小矩形的四个顶点坐标，便于之后将检测结果框绘制出来。利用这两个便利的方法，我们可以得到原始二维码的检测结果框如图10(d)所示。

i.6 优缺点

优点 基本方法的优点在于实现简单明了，检测的速度也有保证，对于拍摄清晰端正，背景简单的图像有较好的检测效果。

缺点 另一方面，这类简单方法的缺点也很明显。(1).对于复杂图像场景下的二维码，检测框可

能会跑偏，检测到另外的二维码的定位模块。在实验中，如图11的二维码场景就会出现误检，将检测框覆盖了其他的二维码的定位模块。(2).当二维码



Figure 11: 基本方法的问题：检测框错误覆盖

图像在拍摄时因为角度问题产生透视时，基本方法的检测框显得过大，而且检测结果仍是透视的，不便于之后的解码识别操作。图12即是这个问题的具体例子。

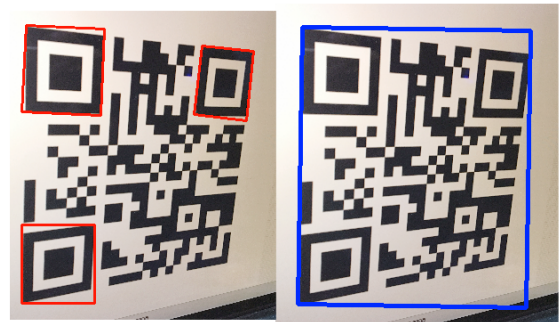


Figure 12: 图像透视时的检测问题

ii. 第二种方法：改进方法

针对于基本方法的缺点，我们再次基于二维码的编码特征，来改进我们的检测方法，作为实验的第二种方法。

ii.1 利用修正模块辅助定位

在介绍二维码结构的部分，我们介绍了二维码的修正模块，也就是定位模块间一横一竖的两条黑白相间的线。

对于类似于图11中的问题，引入修正模块之后就可以迎刃而解。只有某个二维码本身的定位模块之间，才会有两条修正模块，而其他二维码的定位

模块之间不会存在和修正模块一样的编码模式。利用这个编码特征，我们就可以使用如图13所示的方法来辅助定位。图14是所有定位模块两两配对之

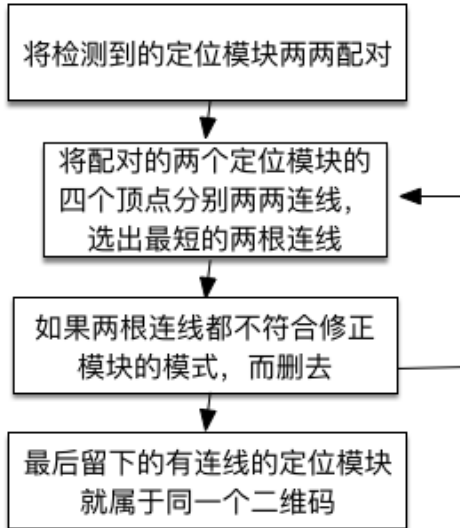


Figure 13: 利用修正模块辅助定位的方法

后选出的配对间最短的两根连线。由二维码的编

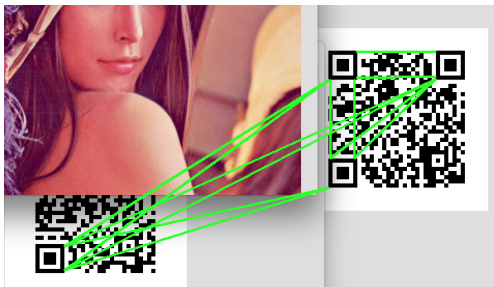


Figure 14: 定位模块配对间最短的两根连线

码可以知道，修正模块是黑白方格间隔分布形成的。由此，我们只要构造出直线遍历器，统计上图所有连线上的黑白像素个数，看统计结果中黑白方格是否均匀分布即可。

由统计学知识我们知道，判断分布是否均匀，可以计算数据的方差，方差小意味着数据相近或相同，即黑白方格均匀分布，连线处于修正模块上。反之，则是黑白方格不均匀分布，连线不在修正模块上。

一条典型的修正模块模式的连线上，黑白间隔像素数的统计结果可能是[5,5,6,6,5]，其方差是0.24。而不是修正模块模式的连线上，黑白间隔像素数的统计结果可能是[1,5,9,3,6]，其方差是7.36。

利用这种方法，我们可以筛选出如图15的两条位于修正模块上的连线。进一步地，修正模块连接的

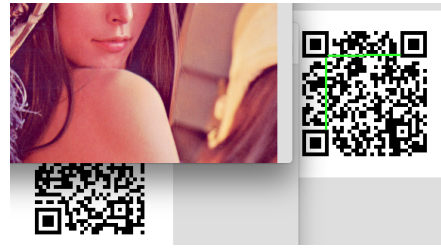


Figure 15: 筛选出的位于修正模块上的连线

三个定位模块就可以确定一个二维码的检测框，如图16所示。

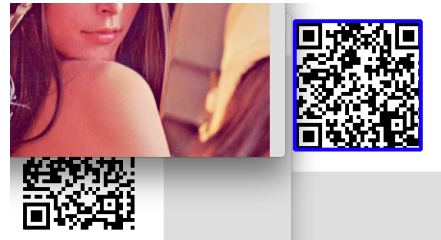


Figure 16: 修正模块辅助定位得到的检测框

ii.2 利用对齐模块进行反向透视变换

在设计二维码结构时也考虑到了图像透视问题，所以在其结构中添加了对齐模块。我们接下来的反向变换就是建立在标识出对齐模块的基础上的。一开始尝试了利用三个定位模块的外顶点作为反向变换的定位点，对变形的图像做反向仿射变换，结果如图17所示。从图中可以看出，反向仿射变换

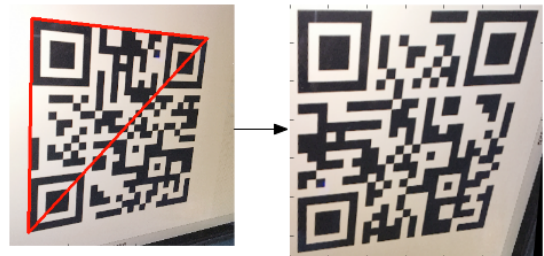


Figure 17: 反向仿射变换

并没有解决图像透视的问题，对于图像规范化的帮助不大。

于是自然而然地想到将二维码的对齐模块作为第四

个定位点，利用反向透视变换，来得到较为规范的二维码图像。

有了之前检测定位模块的经验，检测对齐模块也是同样的道理。不过添加了额外一步，通过检测到的矩形的面积大小来筛选对齐模块，以防止误检。对图像做反向透视变换的结果如图18所示。



Figure 18: 反向透视变换

ii.3 优缺点

优点 相较于基本方法，能够检测复杂背景下的多个二维码，同时也能很好的处理透视变形的二维码图像。

缺点 相较于基本方法，改进方法的计算量较大，实时性的表现较弱。

VI. 实验和讨论

VII. 结语

参考文献

- [1] Otsu N. *A threshold selection method from gray-level histograms*[J]. *Automatica*, 1975, 11(285-296): 23-27.
- [2] Kittler J, Illingworth J. *On threshold selection using clustering criteria*[J]. *IEEE transactions on systems, man, and cybernetics*, 1985 (5): 652-655.
- [3] Sahoo P K, Soltani S, Wong A K C. *A survey of thresholding techniques*[J]. *Computer vision, graphics, and image processing*, 1988, 41(2): 233-260.
- [4] Chen D, Zhang F, He G. *An Improved Binarization Algorithm for Document Image* [J][J]. *Computer Engineering*, 2003, 13: 032.
- [5] Bernsen J. *Dynamic thresholding of grey-level images*[C] *International conference on pattern recognition*. 1986, 2: 1251-1255.
- [6] Hsia S C, Chen M H, Chen Y M. *A cost-effective line-based light-balancing technique using adaptive processing*[J]. *IEEE transactions on image processing*, 2006, 15(9): 2719-2729.
- [7] Gonzalez R C, Woods R E. *Digital image processing*[J]. 2002.
- [8] Ohbuchi E, Hanaizumi H, Hock L A. *Barcode readers using the camera device in mobile phones*[C] *Cyberworlds, 2004 International Conference on*. IEEE, 2004: 260-265.
- [9] Sun A, Sun Y, Liu C. *The QR-code reorganization in illegible snapshots taken by mobile phones*[C] *Computational Science and its Applications, 2007. ICCSA 2007. International Conference on*. IEEE, 2007: 532-538.
- [10] Hu H, Xu W, Huang Q. *A 2D barcode extraction method based on texture direction analysis*[C] *Image and Graphics, 2009. ICIG'09. Fifth International Conference on*. IEEE, 2009: 759-762.
- [11] 刘宁钟, 杨静宇. 基于波形分析的二维条码识别[J]. *计算机研究与发展*, 2004, 41(3): 463-469.
- [12] Leong L K, Yue W. *Extraction of 2d barcode using keypoint selection and line detection*[C] *Pacific-Rim Conference on Multimedia*. Springer Berlin Heidelberg, 2009: 826-835.
- [13] Lin D T, Lin C L. *Multi-symbology and multiple 1d/2d barcodes extraction framework*[C] *International Conference on Multimedia Modeling*. Springer Berlin Heidelberg, 2011: 401-410.
- [14] Canny J. *A computational approach to edge detection*[J]. *IEEE Transactions on pattern analysis and machine intelligence*, 1986 (6): 679-698.