

基于窗口 Hough 变换与阈值分割的矩形识别算法^①

贺 辉, 闫 明, 黄 静

(北京师范大学珠海分校 信息技术学院, 珠海 519087)

通讯作者: 贺 辉, E-mail: hehui@bnuz.edu.cn

摘 要: 提出一种基于窗口霍夫变换与阈值分割自动识别图像中的矩形策略: 通过图像窗口霍夫变换, 提取霍夫图像的峰值 (对应原始图像的线段), 当四个峰值满足某些几何条件时, 则检测出矩形; 对图像进行阈值分割, 将分割结果与霍夫变换的矩形做拟合修正. 对不同成像背景和光照环境下图像的集成测试结果表明, 本策略能够很好地抑制在多种自然光照不均和拍摄角度造成的干扰. 且采用了缩略图计算, 降低了逐像素运算的时间复杂度, 可满足实时性要求. 该技术可运用在实时准确裁剪银行票据目标等各个需要快速识别矩形的工程领域.

关键词: 窗口霍夫变换; 矩形检测; 自动裁剪; 阈值分割

引用格式: 贺辉, 闫明, 黄静. 基于窗口 Hough 变换与阈值分割的矩形识别算法. 计算机系统应用, 2018, 27(3): 131-135. <http://www.c-s-a.org.cn/1003-3254/6236.html>

Rectangle Detection Algorithm Based on Windowed Hough Transform and Threshold Segmentation

HE Hui, YAN Ming, HUANG Jing

(Information Technology Collage, Beijing Normal University, Zhuhai 519087, China)

Abstract: This study proposes a method for automatic recognition and cutting of bank bills using windowed Hough transform: by scanning each pixel to compute the Hough transform of the image and extract the peaks of the Hough transform (which correspond to line segments). A rectangle is detected when four extracted peaks satisfy certain geometric conditions (which correspond the border of bills). Threshold is used to segment the source image and perform fitting correction for the segment result and the extracted Hough transform rectangles. The integration test results of on different image backgrounds and illumination environment indicate that the proposed strategy has a good ability to suppress the interference caused by different natural illumination and shooting angles. In addition, thumbnails view is used to extract feature, reducing the time complexity of pixel-by-pixel operation.

Key words: windowed Hough transform; rectangle detection; auto-cropping; threshold segmentation

矩形结构的识别被广泛应用在各个领域, 例如低温电子显微镜下对矩形和圆形微粒的自动检测; 航拍图片中对矩形结构 (例如车辆、建筑物) 的自动或半自动检测; 或者检测图像或录像里的车牌辨识等等. 目前文献论述的大多数矩形检测方法是基于原始边缘和直线的检测^[1-3]以及基于图像阈值分割检测矩形^[4-8]. 例如 Lagunovsky 和 Ablameyko 提出了基于原始直线的

矩形检测技术^[1]. 首先, 提取出原始直线, 将这些直线分组聚合为线段. 对比其长度和方向来检测出四边形, 再进一步近似为矩形. Lin 和 Nevatia 提出了在航拍图像中检测矩形和平行四边形的技术^[2]. 他们的技术基于线检测, 之后选择某些在某些值范围内的线段 (取决于建筑物的最大和最小尺寸). 在给定的线段中搜索反平行线, 从而定义一个搜索区域, 再搜索矩形的余下两条边.

① 基金项目: 广东省高校重大科研项目-特色创新项目(自然科学)(2016KTSCX167); 广东省自然科学基金(2016A030313384); 海南省自然科学基金(20156227)

收稿时间: 2017-06-12; 修改时间: 2017-06-27; 采用时间: 2017-06-30; csa 在线出版时间: 2018-02-09

Jung 和 Schramm 提出了一种使用一个环形滑动窗口在图像中进行逐像素扫描,对当前环形窗口内的图像求 Hough 变换,通过检验 Hough 空间内的峰值的特性,来判断当前滑动环的圆心是否落在矩形的中心点^[3].这种方法可以有效精确地检测出任意矩形,但是逐行扫描的方式会导致大量像素点被重复计算,算法效率低. Mahnaz Shafii 和 Maher Sid-Ahmed 在近年来的研究中提出了一种基于图像中的平行轴边界框的最小面积来对文档中的结构进行倾斜检测和矫正.通过使用最小边界框的区域标准来增强垂直轮廓和平行轮廓.这种方法在多种倾斜角度中都可以有效地匹配^[4].而基于阈值分割的方法对输入图像目标和背景反差要求较高,容易受到噪声的干扰而不易准确的检测到目标矩形^[9-11].

本文面向银行票据自动裁剪应用需求,结合了基于直线的检测和基于图像阈值分割的两种策略的优点以及在特定环境下的局限性,提出了一种基于窗口霍夫变换与阈值分割的图像中矩形的自动识别策略:对全局图像的 Hough Space 峰值进行匹配,将匹配的结果与对图像阈值分割的结果做拟合,从而得到票据的目标区域.具有如下特点:1)能有效区分非目标区域的噪声结构干扰;2)无须设定阈值;3)算法性能优越,匹配结果快速准确.

1 算法原理

1.1 Hough 变换原理 (HT)

HT 是利用图像的全局特征将图像的形态学信息做变换与统计的方法,HT 用来检测一个图像的线性结构是很有效的. J.Princen 等提出了对 Hough 变换的正式的数学定义.广义的霍夫变换 (Hough transform) 可以表示为通过对目标形状的量化所得到的核函数 (Kernel Function) 在关于数据点集合之内的积分.其中, Kernel Function 为目标的形状和量化参数空间之间的转化关系^[12]. Duda 和 Hart^[13]的研究表明任何线在 xy 平面内都能被描述成 $\rho = x\cos\theta + y\sin\theta$. 其中, ρ 是垂直距离, θ 是直线的垂直角度.霍夫变换将一个二维图像的边缘点集合 (x_i, y_i) 使用二维函数转换为满足 $\rho = x\cos\theta + y\sin\theta$ 的线段的边缘点集合.而在实际应用中,倾角 θ 和垂直距离 ρ 可以被量化,得到一个数组 $C(\rho_i, \theta_i)$,这个数组的峰值可以被用来检测边缘点聚合成的线段^[13].

由于在霍夫空间的线段峰值特性明显,因此基于

霍夫空间的基线模式的检测被广泛应用. Abdelhak 所提出的基于随机霍夫变换的技术用于对阿拉伯语文件的倾斜校正和基线检测^[14].通过计算文本行中较低基线的斜率来识别和矫正文档的倾角. Trupti 的研究中也将霍夫变换应用于手写梵文文档的倾斜检测和矫正,通过提取文档的每个词,对每个词语单元做霍夫变换来检测歪斜^[4].我们早期的研究也通过霍夫空间的基线检测技术来对银行票据进行预处理^[15].

在图像中识别矩形包含多个对象,我们需要在给定的霍夫空间中检测出能够识别出矩形特征的模式.因此,我们记录了一些矩形所包含的特定几何联系,可以用来直接在霍夫空间中做检测.

1.2 在 Hough Space 里的矩形特征模式

假设一个矩形包含四个顶点 $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, $P_3 = (x_3, y_3)$ 和 $P_4 = (x_4, y_4)$, 其中平行边 P_1P_2 和 P_3P_4 长度设为 a , 平行边 P_2P_3 和 P_4P_1 长度设为 b , 如图 1.

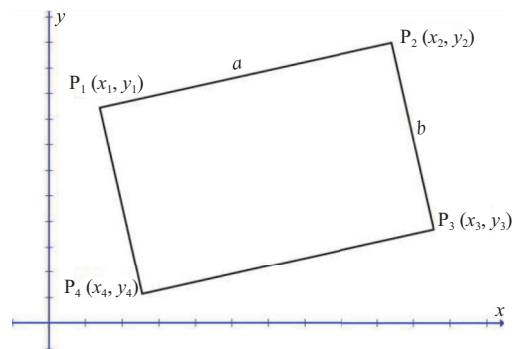


图 1 处在笛卡尔坐标系的矩形

这个矩形在霍夫空间的图像如图 2, 它有四个峰值点, 分别为 $H_1 = (\rho_1, \theta_1)$, $H_2 = (\rho_2, \theta_2)$, $H_3 = (\rho_3, \theta_3)$, $H_4 = (\rho_4, \theta_4)$. 分别对应矩形的 4 条边.

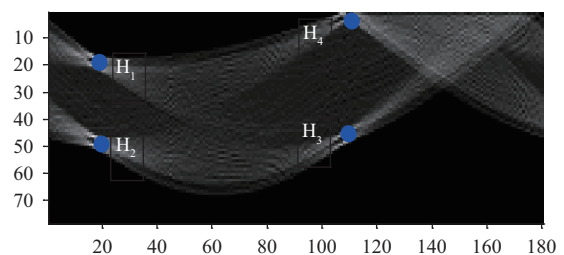


图 2 对矩形做 Hough 变换的 Hough space

不难观察到这 4 个峰值满足下面的几何关系:

1) 令 $\theta = \alpha_1$ 的两个峰值点 H_1 和 H_2 组成一对, $\theta = \alpha_2$ 的两个峰值点 H_3 和 H_4 组成一对. 这两个峰值对

关于 θ 轴的相对角度为 $\Delta\theta = |\alpha_1 - \alpha_2| = 90^\circ$.

2) 属于同一对峰值点的两个峰值高度是相等的, 对应到各自的线段的长度. 例如 $C(\rho_1, \theta_1) = C(\rho_2, \theta_2) = b$ 和 $C(\rho_3, \theta_3) = C(\rho_4, \theta_4) = a$.

3) 每对峰值之间的垂直距离正好是矩形的边, 即 $|\rho_1 - \rho_2| = a$ 和 $|\rho_3 - \rho_4| = b$.

若在当前图像中有其他结构, 这些边缘会和干扰信息和其他结构相关联, 也许也会匹配这些几何关系. 因此, 对干扰信息的去除也是不可缺少的步骤.

1.3 构建并分析峰值点集合

接下来是通过在所得的离散化 Hough 空间里寻找峰值来检测线段. 由于 $C(\rho_i, \theta_i)$ 表示满足线性方程 $\rho = x\cos\theta + y\sin\theta$ 的边缘点的数量, 因此找到霍夫图像的峰值的简单方法是提取满足 $C(\rho, \theta) > T_c$ 的所有点 (即检索像素点大于等于 T_c 的所有直线), 得到一个离散点聚合 $\{H_m \in A_{HT} | H = (\rho_m, \theta_m)\}$. 但是, 噪声和其他结构会降低这种估计峰值的精度^[7]. 为此, 使用 butterfly 模式去分析峰值附近区域可以有效地增强区域拟合度^[6]. Butterfly 模式在此不做太多解释, 此方面, Furukawa 和 Shinagawa 提出了一个简化版本的 butterfly 计算用来增强霍夫图像^[7]. 对于给定的图像, 对应的增强公式为:

$$C_{enh}(\rho, \theta) = hw \frac{C(\rho, \theta)^2}{\int_{-h/2}^{h/2} \int_{-w/2}^{w/2} C(\rho + y, \theta + x) dx dy} \quad (1)$$

其中 h 和 w 表示增强过的矩形区域的长和高. 由于 ρ 和 θ 已经被量化, 所以通过矩形遮罩的卷积来求上式的积分. 最终, 将满足 $C(\rho, \theta) > T_c$ 的增强图像 $C_{enh}(\rho, \theta)$ 的局部最大值存储为峰值.

对当前的点集 $\{H_m \in A_{HT} | H = (\rho_m, \theta_m)\}$ 求两两排列运算, 找到匹配符合下列两个条件的两点, 并标记为 $\text{pair}(H_i, H_j)$:

$$1) \Delta\theta = |\theta_i - \theta_j| < T_\theta, \text{ 即 } \theta_i \approx \theta_j = \alpha_0.$$

$$2) |C(\rho_i, \theta_i) - C(\rho_j, \theta_j)| < T_L \frac{C(\rho_i, \theta_i) + C(\rho_j, \theta_j)}{2},$$

即 $C(\rho_i, \theta_i) \approx C(\rho_j, \theta_j)$.

上式中, T_θ 是最小容错角度阈值, T_L 是最小容错归一化阈值, $\Delta\theta = |\theta_i - \theta_j| < T_\theta$ 所映射的原图关系是线段 H_i 和 H_j 互相平行, $|C(\rho_i, \theta_i) - C(\rho_j, \theta_j)| < T_L \frac{C(\rho_i, \theta_i) + C(\rho_j, \theta_j)}{2}$ 所映射的原图关系为线段 H_i 和 H_j 长度相等. 而所找到的 $\text{pair}(H_i, H_j)$ 即为具有平行特征的线段.

接下来, 对当前所有 $\text{pair}(H_i, H_j)$ 求归一化期望特征

$P_k = (\pm\xi_k, \alpha_k)$, 其中: $\alpha_k = \frac{1}{2}(\theta_i + \theta_j)$, $\xi_k = \frac{1}{2}|\rho_i - \rho_j|$, $P_k = (\pm\xi_k, \alpha_k)$ 即为每对 $\text{pair}(H_i, H_j)$ 的平均期望特征.

最后, 对 $P_k = (\pm\xi_k, \alpha_k)$ 求两两排列运算, 匹配符合此条件的 pair :

$$\Delta\alpha = |\alpha_k - \alpha_l| - 90^\circ < T_\alpha$$

所找到的两对 $\text{pair}(H_i, H_j)$ 即为具有矩形特征的 4 条线段, 对应图像里的矩形.

2 关键算法实现

2.1 Canny 算子^[7]对票据做边缘检测预处理

由于光照的干扰, 将彩色票据图像转化为灰度图像会有可能丢失边缘细节特征, 考虑到接下来的工作需要图像做阈值分割, 所以本文直接对彩色图像的 RGB 三通道进行处理.

2.2 对边缘图像做 Hough 变换

在 $M*N$ 目标图像中, 将 ρ 离散化为 $p*\rho$ 个参数空间, 将 θ 离散化为 $K*\theta$ 个参数空间. 对于 p 和 K 的选取, Furukawa 和 Shinagawa 所提出的方法具有借鉴意义, 对于一个 $M*N$ 图像来说, 计算出的霍夫图像长为 $4M/3$, 宽为 $4N/3$. 在这个情况下, 可以设定 $M=N=D_{\max}$, 即可得离散步长: $d_\theta = \frac{3\pi}{4D_{\max}}$, $d_\rho = \frac{3}{4}$ ^[12]. 对于在本例的票据实验用例中, 为了简化运算, 我们取 $p = \text{Sqrt}(M^2 * N^2)$, $K=180$, 步长 $\text{step} = 1$. 因此计算得到的 Hough 变换结果图像宽度和高度分别为 $\text{Sqrt}(M^2 * N^2)$ 和 180.

2.3 构建并分析峰值点集合检测矩形

在实际应用中, 由于银行票据通常具有固定不变的长宽比, 该约束条件可以用来在当前所找到的两对 $\text{pair}(H_{i1}, H_{j1})$ 、 $\text{pair}(H_{i2}, H_{j2})$ 的集合中再一次搜寻, 寻找符合以下条件的 Pair, 即进一步完成了对目标区域的约束:

$$T_{\xi_{\min}} \leq \xi_1 / \xi_2 \leq T_{\xi_{\max}} \quad (2)$$

式中, $T_{\xi_{\min}}$ $T_{\xi_{\max}}$ 分别表示实际票据的长宽比.

2.4 对图像做阈值分割

此处采用我们早前提出的自适应直方图阈值二值化的目标分割算法^[15]. 分割结果往往包含噪声, 如孤立点噪声或呈块状的噪声, 可以分别通过中值滤波和对形态学操作来消除.

2.5 将分割结果与霍夫变换的矩形做拟合修正

最后, 我们得到了一个存储图像矩形信息的集合 $List<R>$ 和一个二值化的图像. 遍历 $List<R>$ 的元素, 将

每一个矩形元素映射到二值化图像中,对目标矩形区域内的像素做采样,记录矩形元素和二值图像的拟合值,选择最大值的矩形元素.此矩形即为目标矩形.

3 算法测试

为了验证本文所提出的方案的有效性,本文面向银行票据自动裁剪需求,选取了50组真实拍摄银行票据图像进行识别率测试.银行票据具有较为完整的矩形结构,但是真实拍摄的银行票据图像存在的多种自然光照不均和拍摄角度造成的干扰对票据的准确识别带来了难度.

测试中,本文将使用我们早期的研究结果自适应阈值分割方法(即对纠偏图像进行自适应二值分割,确定裁剪框的方法)^[15]和本文所提出的方法进行识别率的比对.

3.1 测试数据

本研究工作的测试数据为高清摄像机采集的照片图像,分为2种分辨率,分别是2592*1944和1600*1200.本文将给出其中的2个典型数据,如图3所示.其中,图3(a)的票据整体呈矩形,但是四条边均存在褶皱,图3(b)里包含一张黄色的小矩形作为匹配的非目标区域.两组数据的背景和票据的灰度差异并不能完全地拉开,这无疑降低了阈值分割方法的匹配度.

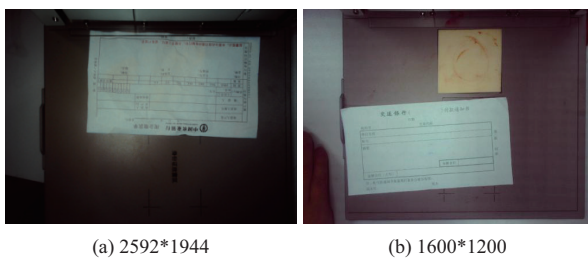


图3 待测试的票据图像

3.2 实验结果及分析

使用自适应阈值分割裁剪的结果如图4所示.实验结果表明,在没有其他矩形干扰情况下,该方案能有效准确地识别出票据,并且不会留下明显的黑边,如图4(a)所示.然而,如果背景和票据的灰度反差降低,会导致在二值图像内仍然存在除了目标区域之外的背景区域未被分割.易将包括亮背景区域的矩形当成目标矩形,从而导致裁剪结果不当,如图4(b)所示.

本文方法裁剪结果如图5所示.结果表明,在存在

严重背景干扰情况下,此裁剪结果仍能有效地识别出票据.这个结果主要是得益于Hough变换带来的所有可能性的矩形匹配.

由对比结果可以看出,自适应二值化分割算法一定程度上依赖于背景和主题目标的高对比度,因此在低对比度的环境下,可能会发生错误匹配的情况.而本方案依赖于矩形特征和对比度两个方面,能有效地去除低对比度环境下的干扰.

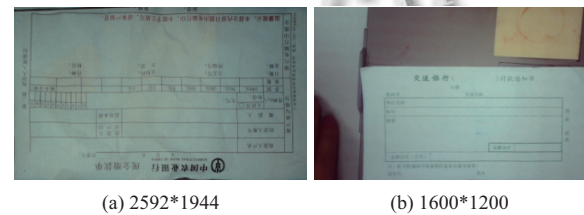


图4 阈值分割法的裁剪结果

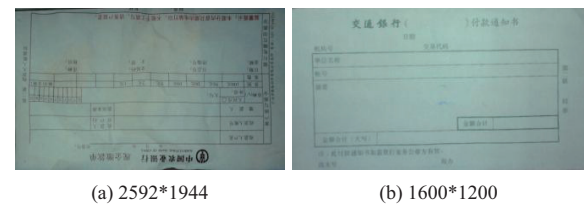


图5 本文方法的裁剪结果

3.3 性能改进

由于对一张图像做Hough变换涉及逐像素进行浮点运算,其运算过程所消耗的时间占处理图像时间的极大比重.而对图像做Hough变换是为了得到图像的量化结构信息,因此Hough变换的结果只与图像本身结构有关,与图像分辨率无关.对于待处理图像,本文将以其以固定宽为100像素等比例压缩.对压缩之后的缩略图做Hough变换,所得到的量化结构信息做矩形识别.对识别出的裁剪框按照比例还原在原图的位置,再对原图进行裁剪和倾斜校正.表1和表2统计了对原图以及缩略图做识别所用的平均时间.

表1 处理原图所用平均时间(单位:ms)

	平均(2592*1944)	平均(1600*1200)
测试1	14774.5	5457.25
测试2	15 570	6390.75
测试3	15 316.5	6394.5
测试4	15 558.5	6374.75
测试5	15 546.75	6418
平均	15 353.25	6207.05

从表1和表2分析可以看出,使用缩略图进行处理,其算法的速度优化明显.大量的逐像素运算已经不再成为性能瓶颈.尤其是对较高分辨率的图像,由于Hough变换与图像大小无关,因此处理时间减幅更大,使得实现实时票据裁剪和远程存储成为可能.

表2 处理缩略图所用平均时间(单位:ms)

	平均(2592*1944)	平均(1600*1200)
测试1	248.25	218.75
测试2	230.25	214.5
测试3	234.25	218.75
测试4	234.5	218.75
测试5	242.25	218.75
平均	237.9	217.9

4 结论与展望

本文针对银行票据自动裁剪应用需求,基于窗口Hough变换和阈值分割,提出了自适应Hough变换的矩形匹配和阈值二值分割算法.窗口Hough变换对目标的识别具有结构约束,二值分割对目标识别具有灰度对比约束,能够最大可能降低光照对分割的不利影响.在目前的50张实际拍摄银行票据图像的测试中,能通过97.5%的测试数据.反映出本算法的可靠性和稳定性,具有推广应用价值.

参考文献

- Lagunovsky D, Ablameyko S. Straight-line-based primitive extraction in grey-scale object recognition. *Pattern Recognition Letters*, 1999, 20(10): 1005–1014. [doi: 10.1016/S0167-8655(99)00067-7]
- Lin CG, Nevatia R. Building detection and description from a single intensity image. *Computer Vision and Image Understanding*, 1998, 72(2): 101–121. [doi: 10.1006/cviu.1998.0724]
- Jung CR, Schramm R. Rectangle detection based on a windowed Hough transform. *Proceedings of the 17th Brazilian Symposium on Computer Graphics and Image Processing*. Curitiba, Brazil. 2004. 113–120.
- Jundale TA, Hegadi RS. Skew detection and correction of Devanagari script using Hough transform. *Procedia Computer Science*, 2015, (45): 305–311. [doi: 10.1016/j.procs.2015.03.147]
- Illingworth J, Kittler J. A survey of the Hough transform. *Computer Vision Graphics & Image Processing*, 1988, 43(2): 280.
- Leavers VF. Survey: Which Hough transform? *CVGIP: Image Understanding*, 1993, 58(2): 250–264. [doi: 10.1006/ciun.1993.1041]
- Furukawa Y, Shinagawa Y. Accurate and robust line segment extraction by analyzing distribution around peaks in Hough space. *Computer Vision and Image Understanding*, 2003, 92(1): 1–25. [doi: 10.1016/j.cviu.2003.07.002]
- 李牧, 闫继红, 李戈, 等. 自适应Canny算子边缘检测技术. *哈尔滨工程大学学报*, 2007, 28(9): 1002–1007.
- 韩思奇, 王蕾. 图像分割的阈值法综述. *系统工程与电子技术*, 2002, 24(6): 91–94, 102.
- 刘欣欣, 李雪, 王琼. 基于灰度直方图的多阈值分割法. *计算机应用与软件*, 2013, 30(12): 28–30, 63. [doi: 10.3969/j.issn.1000-386x.2013.12.008]
- 陈果, 左洪福. 图像阈值分割的两种新技术. *模式识别与人工智能*, 2002, 15(4): 468–473.
- Princen J, Illingworth J, Kittler J. A formal definition of the Hough Transform: Properties and relationships. *Journal of Mathematical Imaging and Vision*, 1992, 1(2): 153–168. [doi: 10.1007/BF00122210]
- Duda RO, Hart PE. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 1972, 15(1): 11–15. [doi: 10.1145/361237.361242]
- Boukharouba A. A new algorithm for skew correction and baseline detection based on the randomized Hough transform. *Journal of King Saud University—Computer and Information Sciences*, 2017, 29(1): 29–38. [doi: 10.1016/j.jksuci.2016.02.002]
- 贺辉, 刘琨, 肖红玉. 银行票据自动裁剪方案设计与控件开发. *计算机与数字工程*, 2016, 45(7): 1327–1332.
- 陈强, 朱立新, 夏德深. 结合Canny算子的图像二值化. *计算机辅助设计与图形学学报*, 2005, 17(6): 1302–1306.
- Shafii M, Sid-Ahmed M. Skew detection and correction based on an axes-parallel bounding box. *International Journal on Document Analysis and Recognition*, 2015, 18(1): 59–71. [doi: 10.1007/s10032-014-0230-y]