

基于投影算子的图书索书号自动识别^①

曹海青¹, 王丹煜¹, 姚志英², 程 玥¹, 郑博一¹

¹(首都经济贸易大学 信息学院, 北京 100070)

²(北京物资学院 物流学院, 北京 101149)

通讯作者: 曹海青, E-mail: cxy8888@bit.edu.cn

摘 要: 图书书脊和索书号的识别是图书馆自动存取书机器人的关键技术, 提高书脊和索书号识别的效率是研究的热点. 研究一种简单的投影算子将其应用于书脊和索书号的识别以提高识别效率和速度; 以书架图书二值化图内容为基础, 构建图书列向量, 在列投影算子的作用下求得列投影向量, 根据列投影向量曲线的情况快速准确识别图书书脊; 以单本书脊二值化图内容为基础, 分别构建书脊行向量和列向量, 通过行投影算子得到行投影向量, 根据行投影向量曲线的变化情况识别索书号元素的行位置信息, 通过列投影算子得到列投影向量, 根据列投影向量曲线的变化情况识别出索书号元素的列位置信息, 根据行、列位置信息切分出索书号元素图片, 在归一化后计算待识别索书号元素图片与字符标准模板的匹配度, 将最大匹配度所对应的字符标准模板映射的字符作为识别结果; 经实验证明基于投影算子的书脊和索书号识别的速度快, 准确率高.

关键词: 图书书脊; 索书号; 行投影算子; 列投影算子; 识别

引用格式: 曹海青, 王丹煜, 姚志英, 程玥, 郑博一. 基于投影算子的图书索书号自动识别. 计算机系统应用, 2018, 27(3): 240–245. <http://www.c-s-a.org.cn/1003-3254/6261.html>

Auto Recognition of Call Number for Library Books Based on Projection Operator

CAO Hai-Qing¹, WANG Dan-Yu¹, YAO Zhi-ying², CHENG Yue¹, ZHENG Bo-Yi¹

¹(Information School, Capital University of Economics and Business, Beijing 100070, China)

²(Logistics School, Beijing Wuzi University, Beijing 101149, China)

Abstract: The recognition of the book spine and call number is crucial for a robot working in library, which has become a hot spot of present research. In this paper, a simple projection operator is proposed and studied to improve the detection of the book spine and call number. A column vector of book is constructed on the basis of the binary image of books in bookshelf, then a project column vector is obtained through a column projection operator. The book spine can be recognized correctly by analyzing the curve of the project column vector. Similarly, the row vector and the column vector of the book spine are constructed on the basis of the binary image of the book spine with a call number. Then, a project row vector is obtained through a row projection operator. The row position information of the call number can be recognized by analyzing the curve of the row projection vector. In a similar way, a project column vector is obtained through a column projection operator, and the column position information can be recognized by analyzing the curve of the column vector. On the basis above, each image element of the call number can be cut and the matching degree between the image element and standard char template can be calculated after normalization, through which the char corresponding to the maximal match degree will be regarded as the recognition result. Experiments show that the projection operator can quickly and correctly recognize the book back as well as the call number.

① 基金项目: 北京高等学校高水平人才交叉培养“实培计划”大学生科研训练计划深化项目; 北京市教育委员会科技一般项目(KM201710038001, SQKM20150037002); 首都经济贸易大学信息学院科研创新团队项目(数据化管理与决策)

收稿时间: 2017-06-15; 修改时间: 2017-07-12; 采用时间: 2017-07-20; csa 在线出版时间: 2018-02-09

Key words: book spine; call number; row projection operator; colomun projection operator; recognition

近年来随着信息技术的迅猛发展,图书馆图书自动存取技术得到快速发展. Suthakorn J 等提出了通过机器人实现图书馆的自动化管理^[1],此后诸多学者开始研究图书馆自动存取书机器人及其相关技术,其中图书书脊和索书号的识别就是一个研究热点. Spemcer G. Fowers 等人利用书脊的颜色特征建立数据库,在使用时通过提取所拍摄到的图书书脊图像颜色特征,然后与数据库中颜色特征进行比对,从而进行书脊定位和索书号定位^[2]. 胡小锋等利用字符上下轮廓的凹凸特征检测字符的近似宽度形成索书号字符切分路径,并根据笔画宽度对切分后的笔画进行修复,实现对图书索书号粘连字符的切分^[3]. 方建军等应用 Canny 算子进行书脊边缘检测,采用霍夫变换法查找书脊边缘线段,实现索书号图像的提取与分割,后来他们又应用小波分析和概率 Hough 变换进行书脊视觉识别^[4,5]. 刘汝翠设计了一种结合深度学习中的 SoftMax 回归算法和书脊图像特征识别相结合的乱架检测算法,实现自动高效的图书乱架检测^[6]. XianZhi Du, Wael Abdalmageed 和 David Doermann 应用 M 超平面和局部敏感哈希算法实现签名的快速识别^[7]. Hong Shao, Tianshu Yu 和 Mengjia Xu 基于局部敏感哈希算法研究了一种快速的大规模签名识别算法^[8]. 综上可知基于图像和视觉处理技术的书脊检测和索书号识别是目前图书馆自动存取书机器人中应用较为广泛的技术,但这种技术的处理效果和效率受使用的环境、图像容量和图像质量的影响较大,影响了书脊和索书号识别的使用和推广. 本文将研究一种简单的投影算子,将其应用于书脊和索书号的识别,旨在改善书脊和索书号识别的效率和准确率,促进图书馆图书自动存取技术的应用和推广.

1 基于投影算子的书脊识别

图书馆自动存取书机器人实现从书架上自动存取图书的前提条件是能书架上的成排书中准确识别出每一个书. 本文将应用投影算子实现对图书书脊的识别.

1.1 投影算子

在高等代数中投影算子是从向量空间映射到自身的一种线性变换,投影变换将整个向量空间映射到它

的一个子空间,并且在这个子空间中是恒等变换. 用数学语言描述如下:

若空间 V 有一个子空间 W , 算子 P 将所有 V 中的元素都映射到 W 中, 且 P 在 W 上是恒等变换, 即 $\exists W$, 使得 $\forall u \in V, P(u) \in W$, 并且 $\forall u \in W, P(u) = u$, 则称算子 P 是投影算子.

目前投影算子已经被广泛应用于图像处理、几何图形识别和图形压缩等领域^[9-11]. 本文中将以二值化图像为研究对象, 以二值化图像的内容为基础构建一组 $n(n \geq 1)$ 维的向量, 定义投影算子 P 将 n 维向量的元素求和映射为一维向量. 用数学语言表述如下, 设一个长为 L 宽为 W 的二值化图像, 将其分别定义为 L 个 W 维的列向量 $C(i, j)(i=1, 2, \dots, L; j=1, 2, \dots, W)$ 和 W 个 L 维的行向量 $R(j, i)(i=1, 2, \dots, L; j=1, 2, \dots, W)$, 本文分别设计列投影算子 P_c 和行投影算子 P_r , 其中列投影算子对列向量求和, 行投影算子对行向量求和, 即:

$$P_c(C(i, j)) = \sum_{j=1}^W C(i, j) \quad (1)$$

$$P_r(R(j, i)) = \sum_{i=1}^L R(j, i) \quad (2)$$

1.2 基于列投影算法的图书书脊识别

图书馆自动存取书机器人利用视频设备抓图书彩色图片经二值化处理结果如图 1 所示, 本文将图 1 所示的二值化图其长为 130 像素, 宽为 170 像素, 将二值化图内容转化为 130 个 170 维列向量 $C(i, j)(i=1, 2, \dots, 130; j=1, 2, \dots, 170)$, 则列向量的元素由 0 和 1 组成, 具体取值情况如式 (3) 所示.

$$C(i, j) = \begin{cases} 1 & \text{source}(i, j) = 255 \\ 0 & \text{source}(i, j) = 0 \end{cases} \quad (3)$$

其中 $\text{source}(i, j)$ 表示二值化图中第 (i, j) 个像素点的像素值. 应用列投影算子 P_c 对列向量 $C(i, j)$ 进行运算后所得的列投影向量为 $A_c(i)(i=1, 2, \dots, 130)$, 列投影向量 $A_c(i)$ 的取值情况如图 2 所示, 由图可知曲线在 AB 段、BC 段和 CD 段列投影向量 $A_c(i)$ 的值也是经历了由大变小再变大的过程, 经研究分析可知造成列投影向量 $A_c(i)$ 如此变化的原因是由于在 AB、BC 和 CD

段分别属于不同图书的书脊内容的投影,由于书与书之间以及书架和书之间的相互影响造成明显的边界效应,即在图像进行二值化处理时书与书的边界、书与书架的边界处容易造成黑色集中的情况发生,而且在这些边界点处列投影向量 A 的值呈现出由中间的最大值分别向其左右两边减少,在曲线中列投影向量的局部最大值点一般就是书与书之间的分界点,根据上述分析思路进行算法设计并对图1所示的图书图片进行处理后定位到书脊的切分线如图3所示,按照图3所示的切分线间的区域即为图书书脊区域,按照切分线进行书籍切分即可切分出图中所包含的每一本图书。

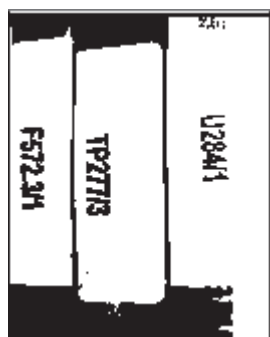


图1 书架上图书的二值化图片

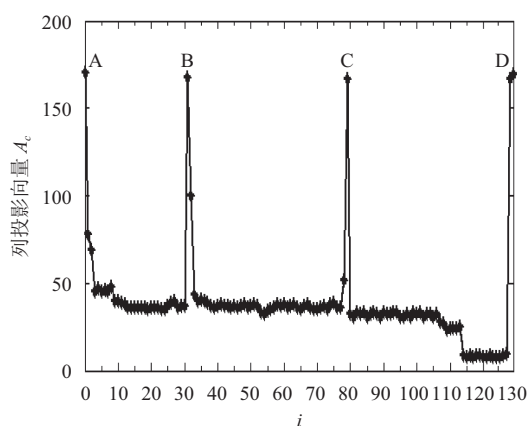


图2 列投影向量 A_c 的变化情况

2 基于投影算子的索书号识别

通过上述研究我们应用列投影算子成功地识别出图书的书脊,但在图书管自动存取书机器人应用中,不仅要准确识别书脊,而且还需要精确识别出书籍所对应的索书号才能实现存取书的自动化。接着我们将通过行投影算子和列投影算子进行图书索书号的定位和切分,然后应用模板匹配的方法进行索书号的识别。

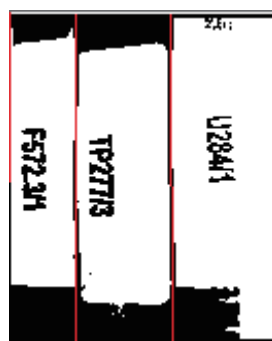


图3 基于列投影的书脊检测结果

2.1 基于投影算子的索书号定位

为了确定书脊图像中索书号的位置,需分别确定索书号在图像中行位置信息和列位置信息,接着我们将应用投影算子分别研究图像内容中索书号的行位置和列位置信息。

2.1.1 基于行投影算子的索书号行位置信息定位

图4为某一本图书经过书脊定位后切分、二值化和区域去噪后的结果,图4所示的图像的长为42像素,宽为170像素,以此二值化图内容为基础构建170个42维的行向量 $R(j, i) (i=1, 2, \dots, 42; j=1, 2, \dots, 170)$, 向量的取值情况如式4所示,

$$R(j, i) = \begin{cases} 1 & \text{source}(j, i) = 255 \\ 0 & \text{source}(j, i) = 0 \end{cases} \quad (4)$$

其中 $\text{source}(j, i)$ 表示二值化图中第 (j, i) 个像素点的像素值,则行投影向量 $B_r(j)$ 为:

$$B_r(j) = P_r(R(j, i)) = \sum_{i=1}^{42} R(j, i), j = 1, 2, \dots, 170 \quad (5)$$

图4 含索书号的书脊二值化图

行投影向量 $B_r(j)$ 的变化情况如图5所示,在图5曲线中BC、DE、FG、HI、JK、LM段内行投影向量 $B_r(j)$ 的值变化均从0增加到一定的值后又减少到0,经分析研究发现造成曲线如此变化的原因是由于索书号中内容在二值化图中为黑色,不同内容间有明显的白色间隙所造成的,为此可以根据行投影向量 $B_r(j)$ 的取值情况确定索书号在图像中行位置信息。图6所示的图书图像中索书号的行信息定位情况如图6所示,由图6可知行投影算子可以很好地实现索书号行位置信息的定位。

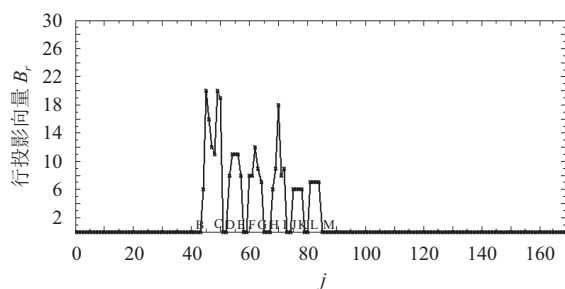
图5 行投影向量 B_r 的变化曲线

图6 索书号行位置信息的定位结果

2.1.2 基于列投影算子的索书数列位置信息定位

以图4所示的图书书脊二值化图为基础构建为42个170维的列向量 $C(i, j) (i=1, 2, \dots, 42; j=1, 2, \dots, 170)$, 向量的取值情况如式6所示,

$$C(i, j) = \begin{cases} 1 & \text{source}(i, j) = 255 \\ 0 & \text{source}(i, j) = 0 \end{cases} \quad (6)$$

其中 $\text{source}(i, j)$ 表示二值化图中第 (i, j) 个像素点的像素值, 则列投影向量 $A_c(i)$ 为:

$$A_c(i) = P_c(C(i, j)) = \sum_{j=1}^{170} C(i, j), i = 1, 2, \dots, 42 \quad (7)$$

列向量 $A_c(i)$ 的变化情况如图7所示, 由图可知曲线在AB段和CD段列向量 $A_c(i)$ 的值为零, 在BC段列向量 $A_c(i)$ 的值不为零, 造成曲线产生上述变化的原因是由于BC段是由于图书索书号的内容在二值化图中为黑色才导致列投影向量 $A_c(i)$ 的值才从零变化到大于零的某一个值, 又从大于这个值变化到零, 为此可以根据列投影向量 $A_c(i)$ 的取值变化情况确定索书号的列位置信息. 图4所示的图书图像中索书号的列信息定位情况如图8所示, 由图8可知列投影算子可以很好地实现索书数列位置信息的定位.

2.2 索书号切分

在确定索书号的行位置信息和列位置信息后, 根据行位置信息和列位置信息画出包含索书号每一个元素的矩形框, 具体如图9所示, 按照图9种所示的矩形框个数, 计算出索书号中所包含的元素的个数, 并按照矩形框的位置将索书号内容切分成一个一个图片实现索书号内容的切分.

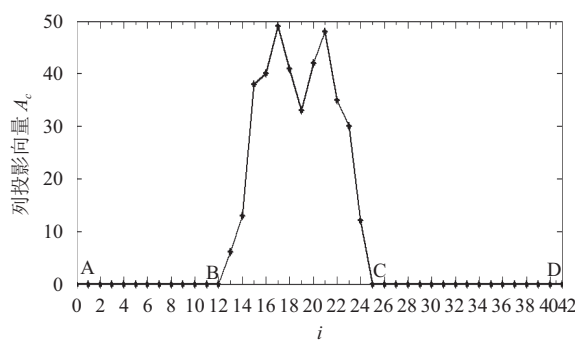
图7 列投影向量 A_c 的变化情况

图8 索书数列位置信息的定位结果

图9 索书号内容划分结果

2.3 基于模板匹配的索书号识别

2.3.1 字符模板库的构建

通过研究图书索书号的内容发现索书号的组成元素包括26大写英文字母、0到9共10个数字和诸如“.”、“/”和“-”等共20个符号. 由于本文采用的是基于图像处理的方法进行索书号的识别, 为此在研究中为这些字符构建了长为32个像素宽为16个像素的二值化图像, 作为识别中的标准模板字符, 图10中有元素F、元素U、元素3、元素1、元素8、元素\、元素4、元素5和元素2的标准模板字符样例.

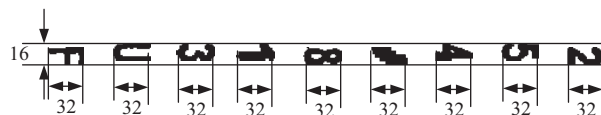


图10 字符模板样例

2.3.2 索书号元素图像归一化处理

根据上述研究所切分的索书号元素的图片因各个图书馆索书号采用的印刷字体、字体大小、是否加

粗、印刷是否倾斜以及时间流逝印刷字的老化等原因,导致切分后所得索书号元素的图片大小各不相同,为此在进行模板匹配之前需对切分所得索书号元素图片进行归一化处理,使归一化后元素图片的几何尺寸为长 32 像素,宽 16 像素. 图 11 为切分所得未进行归一化处理的元素图片和归一化后元素图片.

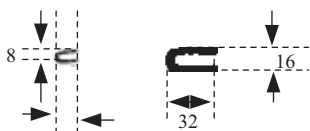


图 11 切分所得索书号图片归一化处理结果

2.3.3 基于欧氏距离的索书号识别

论文在研究过程中构建了 56 个组成索书号内容的元素字符模板二值化图作为研究的字符标准模板库,在图书索书号识别时通过计算切分且归一化处理的图书索书号元素二值化图与字符模板库中所有模板的匹配度,取其中匹配度最大的字符模板对应的字符为最终的识别结果. 具体数学解释如下: 首先在索书号识别系统中将根据每一个字符标准模板二值化图内容构建 16 个 32 维的向量 $M(i, j, k)$ ($i=1, 2, \dots, 56; j=1, 2, \dots, 16; k=1, 2, \dots, 32$); 接着将以每个切分且归一化处理待识别的索书号元素对应的二值化图构建 16 个 32 维的向量

$N(j, k)$ ($j=1, 2, \dots, 16; k=1, 2, \dots, 32$), 设 $d(i)$ 为待识别的索书号元素二值化图与第 i 个字符标准模块的欧氏距离, 则:

$$d(i) = \sqrt{\sum_{j=1}^{16} \sum_{k=1}^{32} (M(i, j, k) - N(j, k))^2} \quad (i = 1, 2, \dots, 56) \quad (8)$$

然后计算带识别索书号元素二值化图与字符标准模板的匹配度 $e(i)$, 其中 $e(i)$ 表示与第 i 个字符标准模板的匹配度, 且:

$$e(i) = \frac{1}{d(i)} \quad (9)$$

当 $d(i) = 0$ 时, $e(i) = +\infty$; 当 $\max(e(i)) = e(j)$, $i = 1, 2, \dots, 56$ 时, 待识别的索书号元素即为第 j 个字符标准模板所对应的字符; 根据如上所述的方法即可识别出索书号的所有元素.

2.3.4 索书号识别的实例化验证

为了检验本文的研究成果, 我们规范如图 12 所示的索书号识别流程. 通过进行大量案例测试发现基于投影算子的图书书脊定位准确率达 99.86%, 索书号内容识别的准确率达 99.73%, 对于成排书架上单本图书书脊的定位时间仅需 0.1 s, 索书号内容单个元素识别时间仅为 0.02 s, 与其他方法相比基于投影算子的书脊识别准确率、索书号识别准确性和性能均有极大的提高.

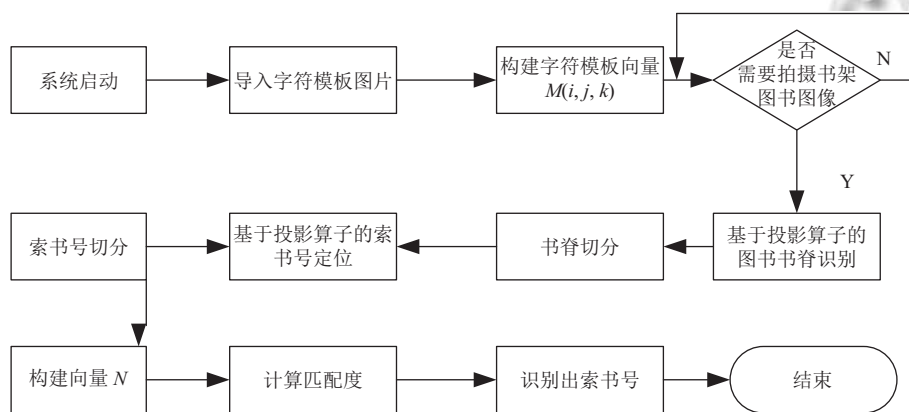


图 12 索书号识别流程

为了进一步验证本文的研究成功, 应用参考文献[5]中方法测试本文方法对书脊的定位情况, 其测试结果如表 1 所示, 由表可知本文所研究的基于投影的书脊定位方法的准确定位率高于文献[5]中的方法, 且耗时平均不达文献[5]中方法的 1/10, 值得进一步的研究和

探索.

3 结论与展望

本文研究了一种简单的投影算子, 将其应用于图书书脊识别和索书号的识别. 本文以书架图书的二值

化图像为基础构建了列向量,应用列投影算子对列向量进行处理,分析列投影向量曲线的变化情况进行图书书脊识别;以单本图书书脊二值化内容为基础分别构建行向量和列向量,应用行投影算子对行向量进行处理得到行投影向量,根据行投影向量曲线的变化情况识别出索书号的行位置信息,应用列投影算子对列向量进行处理得到投影列向量,根据列投影向量曲线的变化情况识别出索书号的列位置信息,根据行、列位置信息切分出索书号元素图片,对切分所得元素图片进行归一化处理;计算待识别元素图片和每一个字符标准模板的欧氏距离和匹配度,则匹配度最大的字符模板所对应的字符即为识别结果;经实验验证基于投影算子的图书书脊识别和索书号识别的准确率高,识别速度快,可以改善图书馆自动存取书机器人的效率,促进图书馆自动化水平的提高。

表1 书脊定位性能比较

厚度(d)类别	本文方法		文献[5]中方法	
	有效定位率 (%)	耗时 (s)	有效定位率 (%)	耗时(s)
厚($d>1000$ 页)	100	0.082	100	1.0
较厚($500\leq d\leq 1000$ 页)	100	0.097	99	1.2
中等厚度($300\leq d\leq 500$ 页)	99.93	0.102	95	1.8
较薄($100\leq d\leq 300$ 页)	99.71	0.102	90	1.9
薄($d\leq 100$ 页)	99.52	0.159	80	2.0

参考文献

1 Suthakorn J, Lee S, Zhou Y, *et al.* A robotic library system

for an off-site shelving facility. IEEE International Conference on Robotics and Automation. Washington, DC, USA. 2002. 3589–3594.

2 Fowers SG, Lee DJ. An effective color addition to feature detection and description for book spine image matching. ISRN Machine Vision, 2012, 2012: 945973.

3 胡小锋, 陈超, 叶庆泰. 图书索书号粘连字符的切分. 上海交通大学学报, 2005, 39(4): 553–556, 560.

4 方建军, 赵强强. 图书馆在架图书的索书号图像提取与分割. 北京联合大学学报, 2015, 29(1): 87–92.

5 方建军, 杜明芳, 庞睿. 基于小波分析和概率 Hough 变换的书脊视觉识别. 计算机工程与科学, 2014, 36(1): 126–131.

6 刘汝翠. 基于自动智能分类器的图书馆乱架图书检测. 现代商贸工业, 2016, 37(25): 172–174.

7 Du XZ, Abdalmageed W, Doermann D. Large-scale signature matching using multi-stage hashing. 12th International Conference on Document Analysis and Recognition (ICDAR). Washington, DC, USA. 2013. 976–980.

8 Shao H, Yu TS, Xu MJ, *et al.* Image region duplication detection based on circular window expansion and phase correlation. Forensic Science International, 2012, 222(1-3): 71–82. [doi: 10.1016/j.forsciint.2012.05.002]

9 赵峰, 杨健. 结合投影算子与小波变换的人脸识别方法. 计算机应用, 2013, 33(S1): 230–232.

10 段汕, 谢英华. 一种基于倾斜投影的图像分析方法. 中南民族大学学报 (自然科学版), 2015, 34(4): 103–108.

11 赵知劲, 张滇华. 基于斜投影算子的压缩采样匹配追踪算法. 杭州电子科技大学学报, 2014, 34(1): 30–33.