

基于图像处理的档案盒标签的检测技术研究

李 婧, 苏 叶, 徐寅林

(南京师范大学计算机与电子信息学院, 江苏 南京 210023)

[摘要] 研究了一种基于图像处理的档案盒标签检测技术,通过摄像头拍摄存放在档案柜窗口区的一排档案以采集图像,利用 U-Net 网络对档案盒标签图片进行分割,经仿射变换矫正图片裁剪得到单本档案盒标签图片,最后利用 surf 匹配算法对库中档案盒标签进行检测。实验结果表明,该方法测得的档案盒图像匹配正确率可达 99.60%。

[关键词] 档案检测,图像分割,图像匹配

[中图分类号] TP391 **[文献标志码]** A **[文章编号]** 1672-1292(2021)02-0060-05

Research on Detection Technology of Archive Box Labels Based on Image Processing

Li Jing, Su Ye, Xu Yinlin

(School of Computer and Electronic Information, Nanjing Normal University, Nanjing 210023, China)

Abstract: This article studies a detection technology of archive box labels based on image processing, which captures an image of labels of archive boxes in the window area of the cabinet by a camera in front of the desk of the cabinet, segments the image of labels by U-Net network, and then obtains a single archive box label picture corrected by affine transformation, and finally, uses the surf algorithm to match the picture with the database. The experimental results show that the matching accuracy rate of archive boxes measured by this method reaches up to 99.60%.

Key words: archive detection, image segmentation, image matching

目前国内外常见的档案检测识别方法是基于条形码、QR 码或 RFID 技术,通过扫码或读取档案盒电子标签进行识别。基于条形码或 QR 码技术需用机械手夹持条码枪移动读取,实时性差。RFID 技术是将射频读卡器排列成行,与档案盒标签一一对应,该方法成本高且易受干扰。由于档案盒的检测识别与图书的检测识别具有一定的相似性,因此可借鉴图书的检测识别技术来研究档案盒检测识别的相关技术。

近年来,由于人工智能技术的快速发展,出现了通过拍摄图书书脊进行检测识别的技术。该方法利用摄像头或其他成像设备采集到书架上的书脊图像,通过识别书脊上的文字信息(如书名、作者、出版社等)来识别图书^[1-2],准确率高、实时性好且方便快捷。由于不同图书的书脊大小、书名及颜色等特征差异较大,书脊检测识别可有一定的容错度而不影响图书的区分。本文研究的档案盒标签往往差异不大,如人事档案标签一般单位、部门、籍贯等都一样,只在具体姓名或编号部分略有不同,因此仅通过文字识别技术很难达到较高的检测准确率。本文提出一种基于图像处理的档案盒标签检测方案,利用摄像头实时采集回转库档案柜窗口区域的整排档案盒的标签图像,通过深度学习 U-Net^[3]网络将整排档案盒标签图像进行分割,对一些形变的图像进行仿射变换矫正,裁剪得到独立的标签图片,最后使用 surf^[4]匹配算法对标签图像进行检测。实验结果表明,本文提出的算法具有较高的匹配正确率。

1 图像预处理

1.1 标签图像分割

将摄像头固定在距离档案柜窗口区前方约 50 cm 处,利用摄像头视角范围大、拍摄清晰且不易变形的

特点,只需要设置两个摄像头即可以获取回转库档案柜窗口区全部档案盒的标签图像.摄像头得到的是窗口区的全部档案盒图片,首先需要分割得到独立的一本本档案盒图片才能进行后续的检测识别.由于传统的分割方法易受噪声等干扰^[5],本文采用深度学习的 U-Net 网络进行分割. U-Net 模型是一个语义分割模型,可实现端到端的分割.具体方法是利用卷积进行下采样,提取出一层又一层的特征,利用这些特征再进行上采样,最终得到类别数量的特征图^[6]. U-Net 网络通过上采样过程中的级联,使得浅层特征和深层特征结合起来,在图像分割效果上有着不错的表现.

首先对 U-Net 网络进行训练,以计算出最佳网络参数. U-Net 分割模型的训练需要原图和与其对应的标签图片两个文件,其中标签图片通过 labelme 软件对原图进行手动标注可得.图 1 所示为具体的训练与测试过程.本文的训练集包含 107 张样本图片,测试集包含 37 张图片.训练过程中损失函数 loss 在迭代 20 次后趋于稳定.

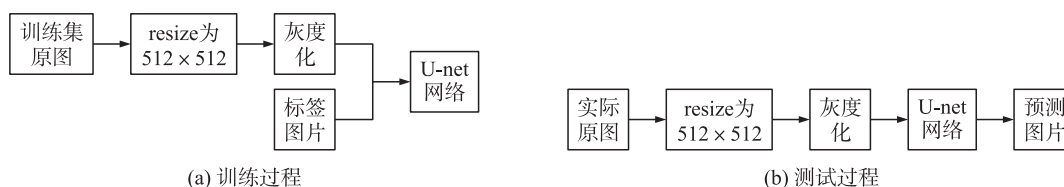


图 1 分割流程图

Fig. 1 Segmentation flow chart

1.2 标签图像的“平直性”纠正

由于摄像头固定在桌面上,必须以一定的仰角才能对准窗口区档案盒背脊,因而拍摄出来的档案盒图片会有一定的角度倾斜,进行矩形裁剪可能会得到残缺或冗余的档案盒图片.为了提高后续检测的正确率,需对图片进行矫正.摄像头仰角拍摄造成的图片畸变以剪切形变为主,本文采用仿射变换对其进行角度调整.

对分割预测图片进行轮廓检测得到每本档案盒轮廓的点的合集,通过最小外接矩形方法即可得到每本档案盒的倾斜角度 θ ,角度范围为 $-90^\circ \sim 0^\circ$.实际计算时,若倾斜角度小于 -45° ,则在原来角度上增加 90° ;若倾斜角度大于 -45° ,则直接对角度反转,最终实现图片的旋转.最后,对每本档案盒旋转后的图片进行裁剪,即得到单独的档案盒图片.

2 图像配准

由于档案盒图片大多背景相同,只有少数细节存在差异,因此图书识别中常用的文字识别技术并不适用于档案盒标签的检测识别.考虑到档案盒的局部差异性,本文采用 surf 匹配算法,主要包括特征点提取、特征点描述和特征点匹配.首先构建 Hessian 矩阵产生兴趣点,然后构建尺度空间,对特征点进行非极大值抑制得到最终稳定的特征点,再利用特征点的 Harr 小波特征确定特征点主方向,生成特征描述子,最后计算出特征点之间的欧氏距离以确定匹配好坏^[7].

2.1 图像特征提取

surf 算法的关键是利用 Hessian 矩阵来得到稳定的特征点,为特征点提取做准备. Hessian 矩阵为:

$$\mathbf{H}(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix}. \quad (1)$$

surf 算法使用盒式滤波器,从而用积分图像加速卷积.图 2 从左到右分别为 y 方向和 xy 方向的高斯二阶偏导数模板及其近似表示,记为 D_{yy} 、 D_{xy} ,此时 Hessian 矩阵的判别式为:

$$\det(\mathbf{H}) = D_{xx}D_{yy} - (\omega D_{xy})^2, \quad (2)$$

式中, D_{xx} 为 x 方向的高斯二阶偏导数近似表示; ω 一般取 0.9,可修正盒式滤波器的近似误差.

surf 算法在不改变目标图像大小的情况下,逐渐增大盒式滤波器的模板尺寸,从而获得图像金字塔,该方法提高了运算效率.

通过 Hessian 矩阵判别式的值得到极值点后,对极值点进行非极大值抑制,将极值点与其图像域和尺度域的所有相邻点进行比较,当其大于所有相邻点时,确定该点为特征点^[8-9].

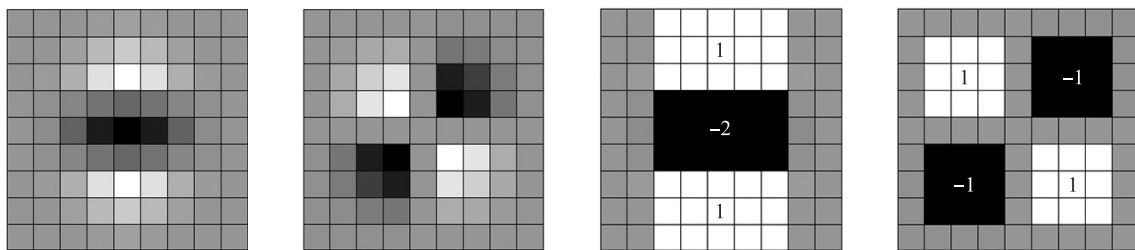


图 2 高斯二阶偏导数的近似表示

Fig. 2 Approximate representation of Gaussian second-order partial derivative

2.2 特征点描述

首先,为了确保特征点的旋转不变性,需给每个特征点分配一个主方向.在以特征点为中心的领域内,计算 60° 范围内所有点的 Harr 小波特征总和,再依次旋转 0.2 个弧度重复计算^[10],最终特征点的主方向为矢量值最大的方向.

提取特征点周围 4×4 个矩形区域块(每个区域边长为 $5s \times 5s$, s 为盒式滤波器的边长尺寸),方向沿着特征点主方向,使用尺寸为 $2s \times 2s$ 模板的 Harr 小波计算每个区域块内 25 个像素点的 Harr 小波特征之和以及 Harr 小波特征绝对值之和,最终 surf 特征描述子共有 $4 \times 4 \times 4 = 64$ 维^[11].

2.3 特征点匹配

surf 算法中常用欧式距离衡量特征点是否匹配,本文采用 Flann(快速最近邻搜索)匹配算法^[12-13].在 Flann 算法中,第一个匹配的是最近邻,第二个匹配的是次近邻,可通过查看两个特征点之间距离的不同来判断是否匹配成功.本文经实验对比,得到设定最近邻与次近邻距离比值小于 0.7 为最佳匹配;当设定阈值小于 0.7 时,随着阈值减少,匹配错误数增多;当设定阈值大于 0.7 时,随着阈值增大,误匹配数增多.

3 实验分析

论文设计了不同标签的 107 个档案盒作为实验数据集.首先,每本档案都要录入一个图片作为比对的对象,即底库照片,共有 107 张作为实验比对的底库标准集.然后将每本档案盒摆放在架子的不同位置拍摄,共采集到 $1\,498$ 张图片.数据集的制作方法有两种,第一种将实际拍摄的图片,按 U-Net 生成的预测图片模板大小进行调整,然后分割、仿射变换后作为测试集 1.该方法预处理裁剪出来的单本档案盒图片宽度小,拍摄照片调整尺寸为正方形后造成图片质量下降,档案盒图片文字也不清晰,影响算法的正确率.第二种方法是 U-Net 得到的掩模图片调回到原图尺寸,对原图进行掩模、分割、仿射变换、裁剪等一些列操作,可以得到高质量的测试集 2.表 1 为两种测试集在 surf 算法下的正确率,可以看出测试集图片的质量对算法有较大影响.

3.1 Hessian 阈值对 surf 算法的影响

由于在 surf 算法中 Hessian 阈值的设定对匹配结果的好坏有一定影响,阈值越高,检测到的特征点越少,特征点越稳定.因此,可用试探法来得到该档案盒匹配的最优检测.图 3 所示为不同 Hessian 阈值下对应的匹配错误数目,可以看出,在阈值 $590 \sim 610$ 之间,匹配的误差数最少,趋于稳定不再减少;而当阈值过高时,则可能因为检测到的特征点太少而无法匹配.因此,本文取阈值 600 为最佳值,匹配正确数目为 $1\,492$,正确率达到 99.60% .

3.2 surf 算法的鲁棒性研究

3.2.1 光照强度对图片匹配正确率的影响

实际应用场景下,可能存在各种各样的环境光强,光照不变性是 surf 算法的一个优点.图 4 所示为不同光照强度下的匹配正确数,其中光照强度等于

表 1 不同测试集的匹配正确率

Table 1 Matching accuracy of different test sets

	匹配正确数	测试总数	正确率/%
测试集 1	1 231	1 498	82.18
测试集 2	1 477	1 498	98.60

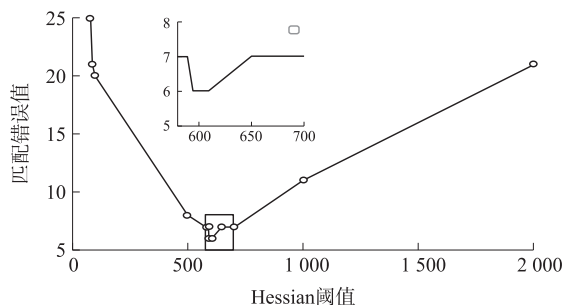


图 3 不同阈值下匹配错误数

Fig. 3 Number of matching errors under different thresholds

1.0 为原图. 从图 4 可以看出,当光照强度为 0.4~1.6, 匹配正确数变化不大,匹配结果较好;当图片过亮或过暗时,surf 算法性能逐渐下降,这充分反应了 surf 特征点检测具有很强的光照不变鲁棒性.

3.2.2 噪声对图片匹配正确率的影响

实际场景中采集得到的档案盒图片还可能存在着许多干扰,最典型的有角落位置档案盒光照明暗不均、档案盒存放时间过久纸张泛黄或者沾染上灰尘等. 为考察这些噪声对图片匹配正确率的影响,本文模拟测试了各种噪声情况下 surf 算法的稳定性,其匹配结果如表 2 所示. 其中明暗不均与泛黄图片受干扰影响较少,稳定的特征点较多,正确率较为稳定;灰尘图片受到较多噪声点干扰,匹配效果较差.

分析 surf 的工作原理可以发现,特征点的提取主要依赖于 Hessian 矩阵判别式,判别式为局部极大值时,该点是领域内最亮或最暗的点,从而确定特征点位置. 而矩阵判别式的值与高斯二阶偏导数及图像的卷积值有关. 局部图像灰度的缓慢变化对卷积值影响较小,因此 surf 算法对图片偏亮、偏暗、缓慢变化的明暗不均及泛黄具有很好的抗干扰能力. 对于灰尘干扰,其颗粒大小往往和图像斑点尺度相近,这严重影响了 surf 的特征点提取,从而影响了图片的匹配正确率.

对于实际场景下受到不同污染的档案盒图片,可以对其进行处理,以提高匹配的正确率. 对明暗不均和泛黄的档案盒图片,尽管 surf 算法抗干扰能力很强,但通过对图片进行适当处理,可以使匹配效果更好. 本文考虑直方图均衡算法和自动色彩均衡(ACE)算法进行处理. 由于直方图均衡化算法一般用于灰度图像,对于档案盒图片这种彩色图像易出现颜色不均、失真等问题. 对明暗不均的加噪测试集进行直方图均衡化,实验测得的匹配正确率约为 95.13%,反而降低了正确率. 自动色彩均衡(ACE)算法能够考虑到图像中亮度和颜色的空间位置关系,实现具有局部和非线性特征的图像亮度、色彩及对对比度调整,可有效提高匹配正确率.

对有灰尘的档案盒图片,本文考虑了中值滤波、非局部平均去噪等方法. 中值滤波是指在以某点像素为中心的邻域内,对所有像素进行排序,将排序中的中值作为该点的响应. 该算法对于去除椒盐噪声效果显著. 对灰尘图片进行中值滤波,可明显去除大部分噪声,此时测得的正确率约为 93.79%. 对于剩余的干扰可再采取非局部平均去噪的算法进行处理,该算法是通过将整幅图像中相似区域求平均进行去噪,需注意亮度权重衰减参数 h 的选择,该值设置高虽可更好地消除噪声点,但同时也会消除图片的细节.

图 5 显示了不同亮度权重衰减参数测得的匹配错误数,可以看出,当参数设为 10.2 时,匹配正确率最优. 由于灰尘档案盒图片本身对比度不高,经中值滤波后直接进行非局部平均去噪处理会丢失很多细节信息,因此可以对灰尘档案盒图片先进行自动色彩均衡,增强对比度后再进行非局部平均去噪处理,以提升匹配效果. 经处理后的档案盒匹配正确率如表 3 所示.

经去噪处理后的图片检测到的特征点数增多,匹配点对数增多,误匹配对数减少,匹配效果较未处理之前有显著提升. 可见,surf 算法对于一定程度下光照变化及实际情况下明暗不均、泛黄干扰均能保持着较好的匹配效果,鲁棒性较好.

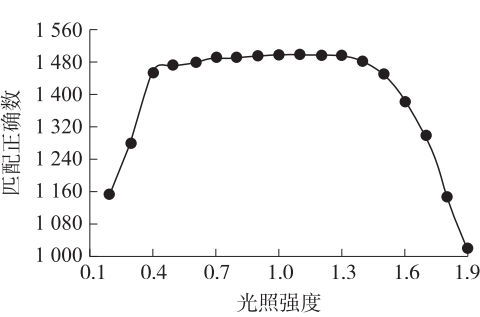


图 4 不同光照强度下匹配正确数
Fig. 4 Matching correct number under different light intensity

表 2 不同干扰下的匹配正确率

Table 2 Matching accuracy under different interference			
	匹配正确数	测试总数	正确率/%
明暗不均	1432	1498	95.59
泛黄	1465	1498	97.80
灰尘	667	1498	44.53

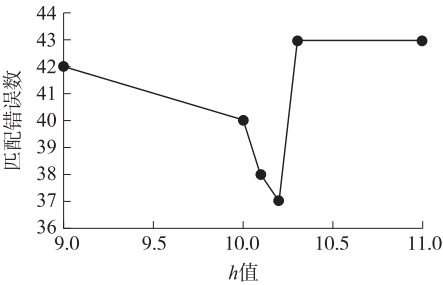


图 5 不同 h 值下匹配情况
Fig. 5 Matching under different h values

表 3 处理后的匹配正确率
Table 3 Matching accuracy after processing

	去噪算法	匹配正确数	测试总数	正确率/%
明暗不均	ACE	1458	1498	97.33
泛黄	ACE	1481	1498	98.87
灰尘	中值滤波,非局部平均去噪	1461	1498	97.53

4 结论

本文阐述了基于图像的档案盒标签检测技术,通过使用 U-Net 网络进行档案盒图片分割、对形变档案盒进行仿射变换等一系列预处理得到独立档案盒图片,使用 surf 算法进行档案盒图片的匹配,并测试了在不同光照强度及不同噪声干扰下 surf 算法的匹配性能,最后讨论了在噪声干扰下的应对方法. 由实验结果可知,该技术对回转库档案盒的检测有较好的效果.

[参考文献] (References)

[1] 殷策. 基于计算机视觉的书脊识别算法研究[D]. 赣州:江西理工大学,2016.

[2] 冯庸. 基于视觉的图书馆在架错序图书自动识别技术研究[D]. 沈阳:沈阳工业大学,2017.

[3] RONNEBERGER O,FISCHER P,BROX T. U-Net:convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin, Germany: Springer, 2015: 234-241.

[4] BAY H,TUYTELAARS T,GOOL L V. SURF: speeded up robust features [C]//9th European Conference on Computer Vision. Garz, Austria:Springer-Verlag,2006:404-417.

[5] 姜枫,顾庆,郝慧珍,等. 基于内容的图像分割方法综述[J]. 软件学报,2017,28(1):160-183.

[6] 柳小波,张育维. 基于 U-Net 和 Res_UNet 模型的传送带矿石图像分割方法[J]. 东北大学学报(自然科学版),2019,40(11):1623-1629.

[7] 胡蝶,侯俊,何晨航,等. 基于特征匹配与 CNN 的浮动验证码识别研究[J]. 软件导刊,2020,19(5):37-41.

[8] 贾迪,朱宁丹,杨宁华,等. 图像匹配方法研究综述[J]. 中国图象图形学报,2019,24(5):677-699.

[9] 崔晨. 基于图像的书籍检测与识别关键技术研究[D]. 南京:南京理工大学,2018.

[10] 李伯轩. 鬼成像目标识别算法研究[D]. 南京:南京理工大学,2017.

[11] 王永明,王贵锦. 图像局部不变性特征与描述[M]. 北京:国防工业出版社,2010.

[12] MUJA M,LOWE D G. Fast approximate nearest neighbors with automatic algorithm configuration[C]//Proceedings of IEEE Conference on Computer Vision Theory and Applications. Lisbon,Portugal:IEEE Computer Society,2009:331-340.

[13] 徐明,刁燕. 基于 SURF 算子与 FLANN 搜索的图像匹配方法研究[J]. 现代计算机,2020,14:49-52,57.

[责任编辑:严海琳]