

一种政府资源档案图像的二值化方法

杨 有^{1,2} 尚 晋³

(北京航空航天大学计算机学院数字媒体实验室 北京 100083)¹ (重庆师范大学数学与计算机
科学学院 重庆 400047)² (重庆电子职业技术学院计算机二系 重庆 400021)³

摘 要 针对政府资源档案灰度图像,应用线性变换增强图像对比度,利用空域滤波减少噪声对图像的影响,在此基础上,使用改进的大津率法计算图像的整体阈值 T_2 ,使用 Bernsen 算法计算图像的阈值曲面 T_1 ,进而定义笔划断裂阈值 T_3 、邻域平均阈值 T_4 、全局和局部相结合的阈值 T_5 ;然后综合利用 $T_1 \sim T_5$ 五个阈值,对灰度图像进行二值化。实验表明,这种基于全局和局部阈值相结合的方法对政府资源档案具有广泛适应性,对伪影和断笔划处理的效果令人满意。

关键词 档案图像,二值化,大津率法,Bernsen 算法

An Image Binarization Method of Government Document Resource

YANG You^{1,2} SHANG Jin³

(Digital Media Laboratory, Computer College of Beihang University, Peking 100083)¹

(College of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047)²

(The Second Department of Computer, Chongqing Electronic Profession College, Chongqing 400021)³

Abstract For the gray-level image of government document resource, a linear transform is employed to enhance the image contrast. A spatial filter is applied to reduce image noise. Based on this preprocessing, the global threshold T_2 is calculated by modified Otsu method. The threshold surface T_1 is computed by Bernsen algorithm. On the basis of T_1 and T_2 , some thresholds such as the stroke breaking value T_3 , the average value T_4 of local processing window and the union value T_5 between global and local are defined. Then the gray-level image is binarized through the combination of these five values. Our experiments show that the proposed method, based on the union of global and local threshold, is adaptive to various documents of government resource. It can achieve our purpose of processing of image ghost and broke rupture.

Keywords Document image, Binarization, Otsu threshold, Bernsen algorithm

1 前言

图像二值化是特征提取和模式识别的重要预处理过程^[1]。设 N 是自然数, (x, y) 是数值图像中某像素的空间坐标, G 表示灰度层的正整数集, $G = \{0, 1, \dots, L-1\}$, 图像函数可以定义为映射 $f: N \times N \rightarrow G$, (x, y) 处像素的灰度表示为 $f(x, y)$, 则对于门限 $t \in G$, $B = \{b_0, b_1\}$ 是一个二值灰度层, 对图像函数 f 二值化的过程, 就是将图像函数 f 变为图像函数 $f_t: N \times N \rightarrow B$ 的一个过程, 它使得

$$f_t(x, y) = \begin{cases} b_0, & f(x, y) < t \\ b_1, & f(x, y) \geq t \end{cases}$$

选择不同的门限值 t , 将会产生不同的二值图像 f , 为了获得正确的分割, 常常需要基于某种准则选择最佳门限值 t^* , 如果仅仅依靠像素本身的属性来决定 t^* , 就称为基于像素属性的门限方法; 如果由每个像素邻域的局部特性来决定 t^* , 就称为基于邻域的门限方法; 当考虑只用一个门限值去分割整体图像时就称为全局门限技术; 而将一幅图像先划分为若干个子图像, 然后分别确定每个子图像的门限值, 则称为局部门限技术。

大津率法, 也称最大类间方差法^[2], 它选取的阈值使得被分为黑白两类的类间方差最大, 是全局域值法中公认的最优算法。Bernsen 算法^[3]是一种典型的局部域值法。本文将两种方法结合在一起, 针对政府资源档案, 首先应用线性变换增强图像对比度, 利用空域滤波减少噪声对图像的影响; 然后利用改进的大津率法计算图像的整体阈值 T_2 , 利用 Bernsen 算

法计算图像的阈值曲面 T_1 , 并定义笔划断裂阈值 T_3 、邻域平均阈值 T_4 、全局和局部相结合的域值 T_5 ; 最后, 综合应用 $T_1 \sim T_5$ 域值, 对图像进行二值化。

2 原理及方法

2.1 图像预处理

在档案数码化过程中, 由于扫描系统或光电转换系统等多方面原因, 常常出现图像不均匀、对比度不足等弊端, 使图像的可视性降低; 另一方面, 由于光电敏感元件电荷粒子随机运动所产生的内部噪声、传输通道的干扰, 以及量化等原因, 数码化后的灰度图像含有一定的噪声, 因此, 在档案图像二值化之前, 有必要对图像进行预处理。这两方面的主要应对措施包括灰度变换和平滑滤波。前者对图像像素灰度值进行修正, 使整幅图像成像均匀; 后者主要平滑图像中的加性噪声。

线性变换是图像增强的常用手段。设原始图像为 $f(x, y)$, 图像灰度值 $D = f(x, y)$ 的可能值域为 $[D_{\min}, D_{\max}]$, 线性变换后的图像为 $g(x, y)$, 则线性变换的表达式为

$$g(x, y) = \begin{cases} 0, & af(x, y) + b < 0 \\ L-1, & af(x, y) + b > L-1 \\ af(x, y) + b, & \text{otherwise} \end{cases} \quad (1)$$

其中, $a = (D_{\max} - D_{\min}) / (L-1)$, $b = -a \cdot D_{\min}$, L 为灰度等级数, 一般情况下, $L = 256$ 。

空域滤波能有效减少随机噪声的影响。它在图像空间中借助模板对图像进行邻域操作, 让图像在傅立叶空间中某个范围内的分量得到抑制, 同时保持其它分量不变, 即改变图像

的频率分布,达到图像增强的目的。空域中的滤波器是通过模板卷积来实现的,其离散计算公式为:

$$h(x,y) = \sum_{k=-W}^{x+W} \sum_{l=-W}^{y+W} g(k,l)p(k,l) \quad (2)$$

其中 W 是一个与滤波模板大小相关的自然数, $g(x,y)$ 和 $h(x,y)$ 分别为滤波前和滤波后的图像。

2.2 图像二值化算法

(1) 采用 Bernsen 算法, 计算 $h(x,y)$ 各点的域值, 得到域值曲面 $T_1(x,y)$:

$$T_3(x,y) = \frac{1}{2} \left[\max_{-W \leq k, l \leq W} f(x+k, y+l) + \min_{-W \leq k, l \leq W} f(x+k, y+l) \right] \quad (3)$$

$$i(x,y) = \begin{cases} 0, & h(x,y) < (1-\alpha)T_2 \\ 1, & h(x,y) > (1+\alpha)T_2 \\ 0, h(x,y) < T_4(x,y) & T_3 > \alpha * T_2 \\ 1, h(x,y) \geq T_4(x,y) & T_3 > \alpha * T_2 \\ 0, h(x,y) < T_5(x,y) & T_3 \leq \alpha * T_2 \\ 1, h(x,y) \geq T_5(x,y) & T_3 \leq \alpha * T_2 \end{cases} \quad (1-\alpha)T_2 \leq h(x,y) \leq (1+\alpha)T_2 \quad (6)$$

其中 T_5 定义为:

$$T_5(x,y) = \frac{1}{2} [T_2 + T_4(x,y)] \quad (7)$$

2.3 算法特性分析

在上面的算法中, 用大津率法计算整体最佳阈值点 T_2 时, 不需要在整个灰度级上计算组间方差和组内方差, 可以将灰度直方图上像素点数为零的灰度级从候选阈值中排除, 然后再用大津判别准则来确定惟一阈值点, 以提高算法执行效率。事实上, 以这些灰度级作为门限, 在邻域内来区分目标点和背景点时, 会把邻域内所有点归类为同一组, 而另一组为空, 这显然是不合理的。

W 为笔划断裂宽度的最大估计值, 约为高度的 $1/10$ 。当邻域窗口增大时, 对时间耗费的影响与邻域尺寸平方关系成正比。

实验表明, α 的取值在 $[0.2, 0.4]$ 内比较合适。从公式 (6) 可以看出, α 取值越大, T_2 发挥的作用就越小, T_3 、 T_4 和 T_5 的作用就越大; 反之, α 取值越小, T_2 发挥的作用就越大, T_3 、 T_4 和 T_5 的作用就越小, 即 α 的取值与全局域值成反比,

其中邻域窗口大小为 $(2W+1) \times (2W+1)$ 。

(2) 采用最大类间方差法, 计算图像的整体域值 T_2 ;

(3) 计算邻域极值差 $T_3(x,y)$:

$$T_3(x,y) = \frac{1}{2} \left[\max_{-W \leq k, l \leq W} f(x+k, y+l) - \min_{-W \leq k, l \leq W} f(x+k, y+l) \right] \quad (4)$$

(4) 对域值曲面 $T_1(x,y)$ 进行中值滤波处理:

$$T_4(x,y) = \text{med}_{-W \leq k, l \leq W} T_1(x+k, y+l) \quad (5)$$

med 表示对大小为 $(2W+1) \times (2W+1)$ 窗口内的域值取中值。

(5) 取 $\alpha \in (0, 1)$, 按 (6) 式对图像逐点二值化:

与局部域值成正比。当 $\alpha \rightarrow 0$, 本文的算法就是大津率法; 当 $\alpha \rightarrow 1$ 时, 本文提出的算法就是一种改进的 Bernsen 算法。

邻域极值差 T_3 代表了邻域窗口内极大值和极小值的变化, 当其与 $\alpha * T_2$ 的值进行比较时, 其判断意义就是消除伪影。 T_4 是对域值曲面 $T_1(x,y)$ 的中值滤波, 它起到两个方面的作用, 一是消除光照不均等噪声所产生的域值突变, 二是避免伪影的强有力手段。

3 实验及分析

选用北京市某局的国土及房产资源档案进行实验, 原始图像 $f(x,y)$ 如图 1, 它是“京朝其 04 字 008 井井”号档案的第 1 页的截取子图, 具有 1068×1058 像素, 取 $\alpha = 0.25$, 经上面描述的算法处理后, 得到二值化的档案图像 $i(x,y)$, 如图 2 所示。原始灰度图像的字节量为 1126400 字节, 二值化后的图像字节量为 25906 字节。从中可以看出, 二值化后图像的可视性大大增强, 目标和背景明显分离, 字符及特征比较清楚。

北京市房屋所有权登记申请书

房产 登记种类: 转移登记 (✓) 变更登记 ()

1: 商品房预售 () 商品房预售 (✓) 经济适用房预售 () 经济适用房预售 ()

存量房买卖 (普通产、央产) 其他 () 收件日期: 2004 年 7 月 9 日

朝阳区 常营乡 鑫兆佳园 24 号楼 图号 I-4-2-

房屋状况

号	部位及房号	结构	总层数	竣工年份	建筑面积	备
4	2-102	钢筋混凝土	8	2004	99.5	
来源	购买 (✓)		合计		99.5	
用途	住宅 (✓) 商业 () 车库 ()					
价	451862.1 元	原房屋所有权证号			京朝其 04 字 008	

图 1

北京市房屋所有权登记申请书

房产 登记种类: 转移登记 (✓) 变更登记 ()

1: 商品房现售 () 商品房预售 (✓) 经济适用房现售 () 经济适用房预售 ()

存量房买卖 (普通产、央产) 其他 () 收件日期: 2004 年 7 月 9 日

朝阳区 常营乡 鑫兆佳园 24 号楼 图号 I-4-2-

房屋状况

号	部位及房号	结构	总层数	竣工年份	建筑面积	备
4	2-102	钢筋混凝土	8	2004	99.5	
来源	购买 (✓)		合计		99.5	
用途	住宅 (✓) 商业 () 车库 ()					
价	451862.1 元	原房屋所有权证号			京朝其 04 字 008	

图 2

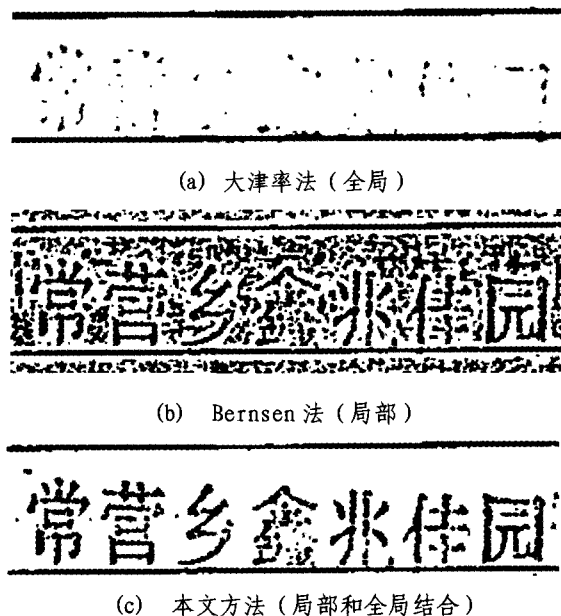


图 3

图 3 为三种方法二值化效果的比较。其中 a) 采用大津率法, b) 采用 Bernsen 算法, c) 是本文描述的方法。从 a) 可以看出, 由于区域中字符的灰度接近背景灰度, 因而全局阈值 T_2 将这些本属于目标点的像素误判为背景像素, 导致字符无法识别。在本例中, $T_2 = 95$, 被误判的像素灰度大多数在 $[101, 104]$ 区间内。从 b) 可以看出, 使用基于邻域的 Bernsen 算法进行二值化, 得到的字符笔划基本清楚, 但仍影现象严

(上接第 215 页)

图 4 中, I、II、III、IV 分别为基于一阶、二阶一阶、二阶 div-curl 样条约束下非刚体光流法和 Horn and Schunck 光流法的面部表情识别结果。由图可以看出, 3 种方法对悲伤、惊奇两类识别率较高, 这主要是因为这两类表情有着较之于其它类别更明显的几何特征, 如惊奇表情的眼睛睁得比较大, 上眼皮被抬高, 下眼皮下落, 其嘴巴也大幅张开。相比之下, 高兴和恐惧这两类识别率普遍较低, 只有 60% 左右, 主要是因为高兴和恐惧在表现时脸部器官有较多相似运动特征, 如两者的嘴巴部分都处于相似程度张开状态。在实验中发现, 影响高兴识别率的主要是恐惧, 而影响恐惧类表情识别率的有高兴、厌恶和愤怒 3 类表情。由图也可看出, 由于本文光流算法相对于 Horn and Schunck 光流法能更准确提取面部表情变化的运动信息, 计算的光流场干扰信息更少, 因此基于一阶、二阶一阶、二阶 div-curl 样条约束下的光流法对各类表情的识别效果均比 Horn and Schunck 光流法好。

表 2 不同方法面部表情识别结果比较

	一阶 div-curl 样条	二阶 div-curl 样条	Horn-Schunck 法	文[4]
识别率	84.5%	85.3%	75.4%	65%

表 2 给出了不同方法对所有表情类的识别结果。引入文 [4] 作为对比, 文 [4] 采用弹性图匹配法, 使用了 9 个人的表情图像训练分类器, 对陌生人脸表情的测试结果是 65% 左右。由于提取的只是表情图像的几何特征, 文 [4] 实质上是一种静态的表情识别方法。而本文方法是一种动态表情识别方法,

重, 这是由于下面两方面原因, 使得算法将部分背景区量化为目标。一是 Bernsen 算法以局部窗口内最大、最小值作为考察点的域值, 当考察点内无目标经过时, 个别噪声点就会引起域值的剧烈变化; 另外, 背景灰度的非均匀性也可能影响到局部域值。从 c) 可以看出, 采用全局和局部结合的方法能综合前面两种方法的优点, 克服两者的缺点, 使得二值化后的图像特征十分明显。

总结 图像二值化是档案图像分析 (DIA) 的第一步, 它为图像几何结构和逻辑结构的提取创造条件。一方面, 二值化后的图像必须能够忠实地再现原字符, 即笔划中不出现空白, 且笔划基本保持原来文字的特征。另一方面, 政府资源档案数量庞大、种类繁多, 在数码化过程中受到主观的和客观的因素限制, 使得灰度图像参差不齐, 这为政府资源档案图像二值化带来挑战。如何尽最大可能提取样本信息和先验知识, 如何做好图像的前处理和后处理, 是我们今后需要继续努力做的工作。

参考文献

- 1 Nagy G. Twenty years of document image analysis in PAMI [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(1): 38~62
- 2 Otsu N. A threshold selection method from gray-level histogram [J]. IEEE Trans on System Man and Cybernet, 1989(8): 62~66
- 3 Bernsen J. Dynamic thresholding of gray-level images [A]. In: Proc. of 8th International Conference Pattern Recognition. Paris, France. IEEE Computer Society Press, 1986. 1251~1255

因此识别效果有明显提高。同时, 由于测试图像序列为陌生人脸 (训练分类器时未出现), 测试对象与人脸对象无关, 由此可知本方法具有较好的鲁棒性。

结论 本文详细讨论了非刚体光流法, 通过引入 div-curl 样条函数作为扩展光流约束方程的附加约束条件, 给出了一阶和二阶 div-curl 样条约束下光流的数值解。同时, 为有效提高光流场计算的精度, 自适应选取了散度和旋度平滑参数, 并且通过计算光流求解方程中系数矩阵 A 的条件数, 进一步判断光流解的稳定性。最后将该算法用于面部表情特征提取。实验结果表明面部表情识别率较常规方法得到显著提高。

参考文献

- 1 Barron J L, Fleet D J. Performance of optical flow techniques [J]. Int. Journal on Computer Vision, 1994, 12 (1): 43~77
- 2 Wildes R P, Michael J, et al. Recovering estimates fluid flow from image sequence data [J]. Computer Vision and Image Understanding, 2000, 80: 246~266
- 3 Bérézat D, Herlin I, Younes L. A generalized optical flow constraint and its physical interpretation [C]. In: Proceedings of Computer Vision and Pattern Recognition, Hilton Head Island, USA, 2000. 487~492
- 4 Hong Hai, Neven H, von der Malsburg C. Online Facial Expression Recognition Based on Personalized Galleries. In: Proceedings of Third International Conference on Automatic Face and Gesture Recognition, April 1998