

基于受控形态学的无限制票据分类

何秀玲^{1,2}, 杨 扬¹, 陈增照^{1,2}, 董才林²

(1. 北京科技大学信息工程学院, 北京 100083; 2. 华中师范大学最优控制与离散数学实验室, 武汉 430079)

摘 要: 目前的票据类型识别要求票据中必须存在框线、分类码等明确的分类标志, 而现有的实际票据往往不满足这些条件, 基于受控形态学操作的二次分类方法选取图像本身的结构特征进行了分级受控击中击不中变换, 对票据样式没有任何限制。分析和实验表明, 该方法计算复杂度低、速度快、准确率高。

关键词: 票据分类; 数学形态学; 受控击中击不中变换; 模板匹配

Unrestricted Check Type Identification Based on Regulated Morphology

HE Xiuling^{1,2}, YANG Yang¹, CHEN Zengzhao^{1,2}, DONG Cailin²

(1. School of Information Engineering, Beijing University of Science & Technology, Beijing 100083;

2. Optimized Control and Dispersed Mathematics Lab, Central China Normal University, Wuhan 430079)

【Abstract】 At present, the identification of check types need some marks on the checks such as frame lines or classified code. However, the classified marks are not shown in all the actual checks. An identification approach of twice classification based on regulated morphological operations which is transformed hierarchically with regulated hit-or-miss transformation by the structural feature of the image, is proposed without any restriction to the check style. Analysis and experiments show that the approach is with low computational complexity, high speed and high correct rate.

【Key words】 Check type identify; Mathematical morphology; Regulated hit-or-miss; Template matching

1 概述

虽然电子银行、电子货币、电子商务等概念现在已经非常普及, 但是邮政、电信、铁路、银行、工商、税务、财政、保险、证券、流通等相关行业使用的纸质票据数量还在以每年 15%~20% 的速度增长, 预计这种增长速度至少会延续到 2010 年。每天要处理大量的票据, 给这些部门造成了很大的负担, 储存和利用也很不方便。将这些任务全部或部分地交给计算机来处理势在必行。

票据处理的目标是提取票据中感兴趣的区域(region of interest, ROI)中的内容并进行识别。不同的行业, 甚至同一行业使用的票据有很多种类, 不同种类的票据 ROI 不同, 因此在提取票据内容之前的基本任务是要进行票据分类, 即票据类型的识别, 这也是票据处理不同于一般的文档识别的地方。

票据处理技术广阔的应用前景吸引了很多专家学者进行这方面的研究, 近年来出现了很多票据处理方面的文章。但这些文章多为关于单一或少量类型票据(如支票、飞机票)内容的分析和票据印刷体、手写体字符的识别^[1,2]。关于票据类型识别的文章较少^[3~5]。文献[3]根据原始票据上的票种代码进行分类, 必须要求票据上有分类代码, 而现在使用的绝大多数票据是没有这种专门的分类码的, 必须重新设计印刷票据, 这无疑给实用带来很大的障碍。文献[4]根据识别票据中关键词字符串分类。这种方法要预先识别出票据中的所有字符串, 再到关键词库中进行匹配, 匹配正确的类型被认为是票据的类型。这样分类前对票据的版面分析和文字识别是相当耗时的, 对于极为相近的类型(例如多联票据)则无法分类。

实验结论为正确率 97%, 拒识率为 3%。文献[5]通过提取特征线分类, 虽然其实验结论为正确率 96%, 但这种方法对非线段结构的票据无法处理, 对框线结构相同的相近的类型(例如多联票据)无法区分。

近年来, 数学形态学^[6]逐渐发展成为数字图像处理的一个主要研究领域, 其基本理论和方法在生物医学、遥感遥测、公路交通等诸多领域取得了许多成功的应用。数学形态学是一种非线性变换, 与传统的FFT变换、统计处理及卷积等相比, 由于它是根据图像的几何特点, 通过选取结构元素做变换的, 因此能较好地提取反映图像结构方面的一些特征。普通的数学形态学操作方法严格, 对噪声和图像边缘细节极为敏感。受控形态学^[7]对普通二值形态学算法进行了改进, 降低了对噪声的敏感性, 但是直接利用受控形态学操作非常耗时。

针对现有的票据分类方法^[3~5]对票据样式的限制, 以及直接利用受控形态学操作速度慢的问题, 本文提出根据图像本身的特征, 基于受控数学形态学中的击中击不中(Hit-miss transform, HMT)操作对票据做二次分类得到票据的类型, 不仅实现了无限制票据类型识别, 而且降低了计算复杂度, 使得票据识别达到实时处理的要求, 同时也避免了普通HMT变

基金项目: 湖北省重点新产品计划基金资助项目(2003BDST004)

作者简介: 何秀玲(1971—), 女, 副研究员、博士生, 主研方向: 模式识别与图像处理; 杨 扬, 教授、博导; 陈增照、董才林, 副研究员

收稿日期: 2005-11-11 **E-mail:** xlhe@mail.ccnu.edu.cn

换匹配过于严格，影响分类正确率的问题。实验表明本文方法的票据分类正确率达到 99.5%，分类速度为 0.1s/张。

2 原理与方法

法国数学家 G. Matheron 和 J. Serra 等人基于积分几何和几何概率提出数学形态学以来，数学形态学^[6]在图像处理中的应用便日趋广泛。利用形态学原理可以构造出许多有效的图像处理与分析方法，如基于击中击不中变换的目标识别。为了解决普通的形态学操作对噪声敏感的问题，Gady Agam 等人提出受控形态学^[7]对普通形态学算法进行了改进，但受控形态学操作计算复杂度较高。利用受控击中击不中变换进行二次分类识别票据类型，大大降低了计算复杂度。为证明该方法的可行性，先给出相关的定义及性质。

腐蚀、膨胀是二值数学形态学的基本运算。集合 A 被集合 B 腐蚀定义为

$$A \ominus B = \{x : B + x \subset A\} \quad (1)$$

其中 A 为输入的二值图像， B 为结构元素。当原点在结构元素内部时，腐蚀具有收缩输入图像的作用。腐蚀具有性质：

$$A \ominus (B \cup C) = (A \ominus B) \cap (A \ominus C) \quad (2)$$

集合 A 被集合 B 膨胀定义为

$$A \oplus B = [A^c \ominus (-B)]^c \quad (3)$$

其中 A^c 表示 A 的补集。当原点在结构元素内部时，具有扩大图像的效果。

击中击不中变换是基于形态学基本运算的复合运算，该变换需要一个由 2 个结构基元组成的结构元素对 $B = (B1, B2)$ ，定义为

$$A \otimes B = (A \ominus B1) \cap (A^c \ominus B2) \quad (4)$$

受控形态学的基本思想是使用一个约束量来控制形态变换，如膨胀、腐蚀等的程度大小。受控形态变换包括受控腐蚀、受控膨胀、受控开、受控闭、受控击中击不中变换操作。

受控膨胀、受控腐蚀及受控击中击不中变换的定义如下：

设 A 表示一幅二值图像， B 为结构元素，则具有约束量 s 的 A 关于 B 的受控膨胀变换定义为

$$A \oplus B = \left\{ x \mid \#(A \cap (-B)_x) \geq s \right\}, s \in [1, \min(\#A, \#B)] \quad (5)$$

这里， $\#$ 表示集合的势。

具有约束量 s 的 A 关于 B 的受控腐蚀变换定义为

$$A \ominus B = \left\{ x \mid \#(A^c \cap (-B)_x) < s \right\}, s \in [1, \#B] \quad (6)$$

这里，假定 $\#A < \infty$ 。

具有约束量 s 的 A 关于 B 的反扩展受控腐蚀变换定义为

$$A \ominus B = (A \ominus B) \cap A, s \in [1, \#B] \quad (7)$$

A 关于 $(B1, B2)$ 的约束量为 $(s1, s2)$ 受控击中击不中变换定义为

$$A \otimes (B1, B2) = (A \ominus B1) \cap (A^c \ominus B2) \quad (8)$$

这里， $B1$ 是前景集合， $B2$ 是背景集合， $s1$ 和 $s2$ 分别是前景集合和背景集合的约束量参数。

受控操作具有以下性质：

约束量 s 相对于集合 B 的补定义为 $\overline{s_B} = \#B - s + 1$ ，则具有约束量 s 的 A 关于 B 的受控腐蚀操作可由式(9)得到

$$A \ominus B = \left\{ x \mid \#(A \cap (B)_x) \geq \overline{s_B} \right\}, s \in [1, \#B] \quad (9)$$

以下本文将证明：当选择的结构元素子集对应区域的信噪比高于结构元素对应区域的信噪比时，选取适当的约束量，可使图像关于结构元素的受控 HMT 操作结果集包含于此图像关于该结构元素子集的受控 HMT 操作结果集中。

对二值图像 A ，设有 2 个结构元素对 $T = (T1, T2)$ 和 $TS = (TS1, TS2)$ 满足关系 $TS1 \subseteq T1, TS2 \subseteq T2$ ，即 $TS \subseteq T$ ， P_T 表示 T 的像素数， P_{TS} 表示 TS 的像素数，将相对于图像大小

的约束量分为 N 个级别，分别取约束量的补为 $\overline{s_{T1}} = \frac{P_{T1}}{N} \times n$ ，

$\overline{s_{T2}} = \frac{P_{T2}}{N} \times n$ ， $\overline{s_{TS1}} = \frac{P_{TS1}}{N} \times n$ ， $\overline{s_{TS2}} = \frac{P_{TS2}}{N} \times n$ ，其中 n 表示相

对于图像大小的约束量， n/N 为约束比例。可以证明：当选择的结构元素子集 TS 对应 A 的区域信噪比高于结构元素 T 对

应 A 区域的信噪比时， $A \otimes TS \supseteq A \otimes T$ 成立。

证明 对任意 $x \in A \otimes T1$ ，设 $\#(A \cap (TS1)_x) = M_{TS1}$ ，

$\#(A \cap (T1)_x) = M_{T1}$ ，根据式(10)可得 $M_{T1} \geq \overline{s_{T1}}$ 。如果平移 x 后结构元素子集 TS 所覆盖 A 中的区域信噪比高于结构元素 T 所覆盖 A 中的区域的信噪比，那么 TS 中覆盖的 A 中区域像

素的匹配比例将比 T 中的大，即 $\frac{M_{TS1}}{P_{TS1}} \geq \frac{M_{T1}}{P_{T1}}$ ，这时根据约束量

的定义有 $\frac{s_{TS1}}{s_{T1}} \leq \frac{M_{TS1}}{M_{T1}}$ ，前面已证 $M_{T1} \geq \overline{s_{T1}}$ ，所以 $M_{TS1} \geq \overline{s_{TS1}}$ ，这样

$x \in A \otimes TS1$ 。因此 $A \otimes TS1 \supseteq A \otimes T1$ 。同理可证 $A^c \otimes TS2 \supseteq A^c \otimes T2$ ，

从而得到 $(A \otimes TS1) \cap (A^c \otimes TS2) \cap A \supseteq (A \otimes T1) \cap (A^c \otimes T2) \cap A$ ，

即 $A \otimes (TS1, TS2) \supseteq A \otimes (T1, T2)$ 。证毕。

根据证明所得的结论，通过适当选取约束量的表示形式及结构元素，对元素 T 的受控 HMT 操作的结果包含在对其子集 TS 的操作结果中。因此在利用受控 HMT 进行模板匹配时，可以分两级进行：先用结构元素子集参与运算，与图像进行受控 HMT 操作得到匹配的候选集合即粗分类结果；再在匹配的候选集合中利用结构元素参与运算，做进一步匹配，即细分类。结构元素取自模板票据图像本身，不必含有分类码、框线等特殊标记。

3 计算复杂度

根据受控 HMT 操作的定义，可以观察到其执行复杂度主要依赖于受控腐蚀和交操作。一个 $P \times P$ 的结构元素对于一个 $M \times M$ 图像的直接腐蚀需要 $O(P^2 M^2)$ 操作， n 个图像的直接交需要 $O(n M^2)$ 次操作，利用基于区间编码的高速算法这个操作数可以减少到 $O(p P m \log p P)$ 和 $O(n m \log n)$ (p 是结构元素中区间数， m 是图像中的区间数)^[8]。这里 p 与 P 为同一数量级， m 与 M 为同一数量级， n 远远小于 M 。采用本文提出的二次分类方法后，粗分类时取的是结构元素的特征子集参加运算，使得 P 远远小于 M ，这时基于区间编码算法的复杂度减少为 $O(m)$ 。细分类是在粗分类的结果(设结果元素有 s 个)中再由这 s 个 $P \times P$ 大小的图像对原图像进行受控 HMT 操作，通过选择合适的结构元素特征子集使 s 远远小于 p ，受控腐蚀

的算法复杂度为 $O(p)$ 。几种方法的算法复杂度比较见表 1。

表 1 算法复杂度比较

采用的算法	受控腐蚀	交
直接操作	$O(P^2 M^2)$	$O(nM^2)$
基于区间编码的高速算法	$O(pPm \log pP)$	$O(nm \log n)$
二次分类	$O(m) + O(p)$	$O(m)$

4 实验与结论

在同样的软硬件环境下,对如图 1(a)中所示的 1334×648 二值 TIFF 图像,受控 HMT 操作的约束比例取为 180/255,当直接定义如图 1(b)所示子图像为结构元素时,分类时间为 5 735ms;采用如图 1(c)所示两级结构元素,分类时间为 94ms。

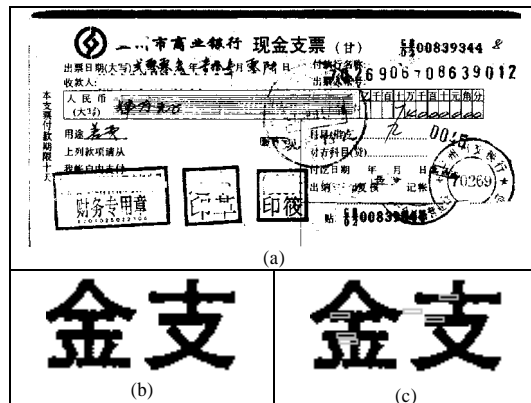


图 1 图像及结构元素示意图

利用本文提出的基于受控形态学的二次分类方法对 19 000 张实际银行票据(200DPI)做类型识别测试。其中票据的类型覆盖了存在框线和不存在框线的,有表头和无表头的,套打票据、手写票据和针打票据等,共 120 种;测试用机 CPU

为 P4-2.4GHz、512MB 内存。

根据前述分析,选择结构元素及其子集时尽量选择信噪比高的区域,且使得粗分类结果尽量少,类型识别平均速度达到 0.1s/张,识别正确率 99.5%,误识率 0%,拒识率 0.5%。其中拒识的主要原因是一些针打票据由于色带过浅导致图像质量很差。

该技术已经在银行票据识别、自动阅卷等系统中采用,取得了很好的效果。下一步的工作是利用机器学习的方法自动确定结构元素的特征子集。

参考文献

- Oliveira L S, Sabourin R, Bortolozzi F, et al. A Modular System to Recognize Numerical Amounts on Brazilian Bank Cheques[C]. ICDAR'01, Seattle: IEEE Computer Society, 2001: 389-394.
- 张重阳, 娄震, 杨静宇. 银行支票中小写金额图像的提取[J]. 中文信息学报, 2003, 17(2): 42-47.
- 陈强, 杨静宇, 娄震等. 基于模板匹配和差影处理的票据图像检测[J]. 计算机工程与应用, 2004, 40(26): 206-219.
- Sako H, Seki M, Furukawa N. Form Reading Based on Form-type Identification and Form-data Recognition[C]. ICDAR'03, Edinburgh: IEEE Computer Society, 2003: 926-930.
- 李琥. 银行单据类型识别研究及实现[D]. 杭州: 浙江大学, 2004.
- 崔屹. 图像处理与分析——数学形态学方法及应用[M]. 北京: 科学出版社, 2002.
- Agam G. Regulated Morphological Operations[J]. Pattern Recognition, 1999, 32(6): 947-971.
- Agam G. Compound Regulated Morphological Operations and Their Application to the Analysis of Line-drawings[C]. GREC'99. Jaipur: Springer-Verlag, 1999: 58-68.

(上接第 13 页)

是,容易看出,并不是所有可能的参数都不安全。比如,选择 $n=N$,这时,攻击者攻击成功的概率不大于找出一个 n 位输出的安全哈希函数的一对碰撞的概率,因此该参数的选择是安全的。然而,这样选择的参数不适合不可靠的通信环境,因为, $n=N$ 意味着所有的数据包不能有丢失现象,都必须收到才能进行签名验证。我们发现,如果合理选择参数且 $n>N$,既可以保证安全,也适合在不可靠通信环境中采用。方案的安全性可以根据 Chernoff 界来说明。

Chernoff 界^[6] X_1, X_2, \dots, X_n 为两两独立的 0-1 分布的随机变量,对所有 i , $\Pr[X_i=1]=1/2$ 。对所有 ε , $0 < \varepsilon \leq 1/4$, 有

$$\Pr\left[\left|\sum_{i=1}^n X_i - n/2\right| > n \cdot \varepsilon\right] < 2e^{-2\varepsilon^2 n}$$

在我们的攻击中,上述随机变量 X_1, X_2, \dots, X_n 可以这样定义:对在 a_{uv} 中第 i 位,如果它与 \tilde{a}_{uv} 的对应位相同,则 $X_i=1$;否则, $X_i=0$ 。如果取 $N=3n/4$,那么 $\varepsilon=1/4$ 。所以,根据 Chernoff 界,如果 n 足够大,伪造一个数据包 p 。使得它的认证码 \tilde{a}_{uv} 与原认证码 a_{uv} 具有的相同位数 $t \geq 3n/4$ 的可能性是可忽略的,即攻击算法成功的概率可以忽略。

4 结论

本文给出了对 EMAS 的一种特殊的攻击方法,该方法通

过选择性地截留部分通信数据包,从而成功地伪造数据流通过认证。虽然原方案并不安全,但是,根据 Chernoff 界,发现适当地选择参数,可以大大降低攻击成功的可能性。基于此,本文进一步讨论并给出了对原方案的参数选择的改进和限制,从而提高方案的安全性,避免选择性截留攻击。

参考文献

- Gennaro R, Rohatgi P. How to Sign Digital Streams[M]. Berlin: Springer-Verlag, 1997: 180-197.
- Canneti R, Garay J, Itkis G, et al. Multicast Security: A Taxonomy and Some Efficient Construction[C]. Proceedings of the 6th ACM Computer and Communications Security Conference, 1999: 93-100.
- Golle P, Modadugu N. Authenticating Streamed Data in the Presence of Random Packet Loss[C]. ISOC Network and Distributed System Security Symposium, 2001: 13-22.
- 李先贤, 怀进鹏. 高效的动态组播群通信认证签字方案[J]. 软件学报, 2001, 12(10): 1486-1494.
- Bellare M, Rogaway P. Random Oracles Are Practical: A Paradigm for Designing Efficient Protocols[C]. Proceedings of the 1st Conference on Computer and Communications Security, 1992: 62-73.
- Goldreich O. Foundations of Cryptography: Basic Tools[M]. Cambridge University Press, 2001.