

基于 PDF417 条码的信息隐藏方法

赵 博, 黄 进

(西安电子科技大学 机电工程学院, 陕西 西安 710071)

摘 要: 信息隐藏技术是信息安全研究领域最新热点之一, PDF417 条码是中应用十分广泛的二维条码之一, 将信息隐藏技术与二维条码技术相结合就产生了基于二维条码的信息隐藏技术。现有技术目前存在着信息嵌入算法复杂和隐藏信息量较小等问题。根据 PDF417 二维条码自身结构特点, 对其组成条空进行适量的微调, 将信息隐藏进条码当中。理论分析和实验结果表明该方法可以有效抵御打印扫描攻击, 具有嵌入算法简单、安全性好、鲁棒性高、信息隐藏量大等优点。

关键词: 信息隐藏; PDF417 条码; 条码结构微调; 嵌入算法; 安全性; 鲁棒性; 信息量

中图分类号: TN918 文献标识码: A 文章编号: 1000-7024(2007)19-4806-04

Information hiding technique based on PDF417 barcode

ZHAO Bo, HUANG Jin

(School of Electro-Mechanical Engineering, Xidian University, Xi'an 710071, China)

Abstract: Information hiding technique is one hotspot of the field of information security, PDF417 is one of two dimension barcode which widely used. They are complex arithmetic and low capacity of information that the existent problems of information hiding technique based on PDF417 barcode. An information hiding technique is proposed. By modifying the configuration of bar and space slightly, information is hidden into PDF417 barcode. Theory analysis and experimental results show that the method features simpleness, security, robustness, and high capacity of information hiding etc.

Key words: information hiding; PDF417 barcode; modifying the configuration of barcode; arithmetic of hiding; security; robustness; contents of information

0 引 言

信息隐藏技术是信息安全研究领域最新热点之一, 在版权保护和状态监视方面有着重要应用, 信息隐藏在民生安全和军事国防等方面有其特殊的意义。二维条码技术是在计算机与信息技术基础上发展起来的一种标准化信息存储、传递和自动识别技术, 具有成本低、容量大、可靠性高、编码方式灵活、便于携带等优点, 在国内外使用十分广泛^[1]。

文献[2]将需隐藏的信息利用二维条码技术编码获得二值图像, 再在密钥控制下对其进行混沌置乱, 最后将水印嵌入到载体图像的DCT(离散余弦变换)域中的办法来实现信息隐藏。文献[3]先对信息进行压缩及混沌加密处理, 然后将其编制成二维条码来实现信息的隐藏。这两种方法虽然隐藏信息的顺序不同, 但是都是将二维条码技术与混沌理论相结合来达到信息隐藏的目的, 其隐藏信息嵌入算法相对复杂。文献[4]提出了利用条码误差的累积实现信息的嵌入和提取, 利用每个符号字符中的 4 个单元条来隐藏一位信息, 没有充分利用二维条码的允许误差, 其可隐藏的信息量仍然是比较有限的。

本文针对二维条码 PDF417 自身的结构, 利用结构微调法对二维条码的组成模块进行适当的变形, 来实现信息隐藏, 具有信息嵌入算法简单、安全性好、鲁棒性高、信息隐藏量大等优点。仿真实验验证了本文方法的可行性。

1 PDF417 条码

PDF 是取英文 Portable Data File 这 3 个单词的首字母的缩写, 意为“便携数据文件”。因为组成条码的每一符号字符都是由 4 个条和 4 个空构成, 如果将组成条码的最窄条或空称为一个模块, 则上述的 4 个条和 4 个空的总模块数为 17, 所以称四一七条码或 PDF417 码^[5], 如图 1 所示。其每行结构如下: a 表示左空白区, b 表示起始符, c 表示左行指示符号字符, d 表示 1~30 个数据符号字符, e 表示右行指示符号字符, f 表示终止符, g 表示右空白区。PDF417 条码的符号尺寸是可变的, 高度范围为 3~90 行, 宽度范围为 $90X \sim 583X$, X 是符号的模块宽度。条码的最大数据量(纠错等级为 0 时)为每个符号表示 1 850 个大写字母或 2 710 个数字或 1 108 个字节。

PDF417 条码可以以多种密度印刷, 以适应不同的印刷及

收稿日期: 2006-10-19 E-mail: zhaobo_xidian@163.com

基金项目: 教育部留学回国基金项目(K40604040114)。

作者简介: 赵博(1980-), 男, 陕西西安人, 硕士研究生, 研究方向为精密测试技术等; 黄进(1968-), 男, 博士, 副教授, 研究方向为智能测控技术、计算机外部设备等。

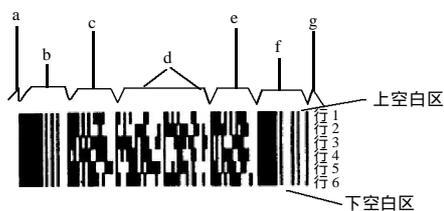


图1 PDF417 条码结构

扫描要求,但符号的模块宽度(X)不得小于 0.191 mm;在 PDF417 条码中,每一个符号字符的误差包括符号字符允许误差(Δp),条或空的允许误差(Δb),边缘到相似边缘的允许误差(Δe),单元的高度允许误差(Δh),它们的允许误差由式(1)~(4)确定;行间水平错位允许值 a 由式(5)确定。其具体含义如图2、图3所示。下列式子中, X 为符号的模块宽,单位为 mm。

$$\Delta b = \pm(0.40X - 0.0127) \quad (1)$$

$$\Delta e = \pm 0.20X \quad (2)$$

$$\Delta p = \pm 0.20X \quad (3)$$

$$\Delta h = \pm 0.20X \quad (4)$$

$$a = \pm 0.12X \quad (5)$$

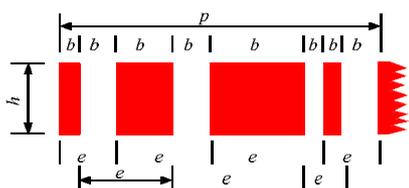


图2 PDF417 条码允许的误差

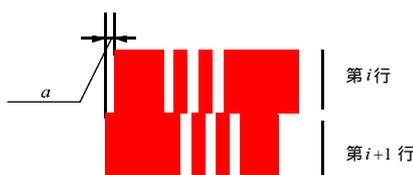


图3 行间水平错位允许值

2 信息隐藏的结构微调法

对条码符号字符的组成单元宽度引入适量的误差,利用误差实现隐藏信息^[6-7]的嵌入与提取。条码结构微调法与PDF417条码本身识读有着密切的关系,而与人的视觉无关,它涉及隐藏信息的嵌入、检测和提取、对隐藏信息的攻击和隐藏信息数量等。

2.1 信息的嵌入

二维条码 PDF417 图像是一种特殊的二值图像^[8],在二值图像中嵌入信息时,必须考虑该像素的邻阈状况。在二值图像中,如果在周围全是 0 的区域中将一个像素由 0 修改成 1,即在一片黑色区域中出现一个白点,则引入了明显的图像失真,如果在黑白区域交接处将一个像素由 0 修改为 1,就不会引人注意,因此,在二值图像中嵌入信息最好在黑白区域交接处进行。

在 PDF417 条码中,隐藏信息选择嵌入的区域从每行的左行指示符号到右行指示符。起始符和终止符不做任何的改变,

这种做法可以解决一般变形技术在解码时需要原始图像信息的缺陷,即实现盲提取。

暗隐藏:我们把在条码黑白模块交接处,嵌入信息(一列像素)的隐藏方法,称之为暗隐藏。明隐藏:我们把在条码黑白模块交界处,嵌入信息(几位像素)的隐藏方法,称之为明隐藏。其隐藏效果如图4所示。这里我们设加上黑点或者黑条为加入信息“0”,加上白点或者白条为加入信息“1”。

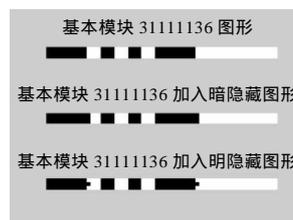


图4 符号字符“31111136”隐藏效果

2.2 嵌入算法分析

从上面的描述可以看出,在信息嵌入算法中引入的误差基本上不影响条码的尺寸规范。只要单元条中引入的误差值 $\Delta w < \Delta b$,边缘到相似边缘的允许误差(Δe)在允许范围之内,其它的参数:符号字符允许误差(Δp)、单元的高度允许误差(Δh)以及行间水平错位允许值 a 等并未发生变化。那么,条码变形产生的误差将符合 PDF417 条码的技术要求,不会影响 PDF417 条码的识读。

2.3 信息的检测和提取

(1)首先将经过打印扫描的 PDF417 条码图像保存为 Windows 位图;

(2)把保存好的条码图像进行放大,以便于我们观察检测;

(3)将放大后的图像按像素读入,并转化为二值矩阵^[9],用检测隐藏信息相对应的矩阵的方法得出隐藏信息。

2.4 打印扫描攻击^[10]

对于二维条码结构微调技术来说,主要的攻击来自打印和扫描,经过打印和扫描后的条码图像会产生轻微的变形,因为在打印过程中黑色条中会产生因为着墨不均产生的白点,在扫描时黑色条的边缘还会产生毛刺,这种变形可能会影响到隐藏信息的检测和提取,同时还会影响到条码的正常识读。

为了衡量条码图像对抗打印、扫描、传真等转换攻击的能力,故而提出条码图像信息隐藏技术的鲁棒性指标,其定义为

$$R = 20 \lg \frac{P_m}{Q_m} \quad (6)$$

式中: P_m ——隐藏信息 $m \times n$ 矩阵的元素数量; Q_m ——噪声 $m \times v$ 矩阵的平均元素数量,这里把图像中隐藏信息与噪声都近似的看作像素点矩阵排列形式。

对于 R : 当 $R < 0$ 时,此时隐藏信息矩阵比噪声矩阵维数小,不能区分出隐藏信息与噪声;当 $R > 0$ 时,此时隐藏信息矩阵比噪声矩阵维数大,可以检测出隐藏信息;其中值得我们关注的是最小的隐藏信息矩阵与最大噪声矩阵元素数的比值 R_{\min} 的值,当 $R_{\min} > 0$ 时,我们可以区分出所有的隐藏信息与噪声。

2.5 隐藏信息的数量

由以上的隐藏过程我们知道二维条码可隐藏的信息量是

非常大的。隐藏信息选择嵌入的区域从每行的左行指示符到右行指示符,如果每一个单元条或空嵌入一位隐藏信息,则其最大嵌入信息量为 I

$$I=4j \times (k-2) \quad (7)$$

式中: j ——二位条码的行数 k ——二维条码的列数。

对暗隐藏来讲,每一个单元条或空恰好嵌入一位信息,设暗隐藏信息量为 I_a ,则

$$I_a=I \quad (8)$$

对于明隐藏来讲,设明隐藏信息量为 I_b ,我们从它的隐藏方式可以知道,假设其每一个单元条或空可以有 t 个位置供信息明隐藏,由其隐藏方式可得出

$$I_b=(C_1^t+C_2^t+\dots+C_{t-1}^t)I_a \quad (9)$$

这对于具有几十行、几十列,而尺寸不过几厘米的条码来讲,其嵌入的信息量是很大的。

3 仿真实验

3.1 仿真实验过程

用所含信息为“PDF417”的PDF417二维条码为例来进行实验。表示“PDF417”的PDF417条码条空组合形式如表1所示。

将其转化为二值数据矩阵,其中“0”代表黑色像素,“1”代表白色像素。取条码的模块宽度 X 为1 mm,对隐藏信息嵌入区域中每一个条或空引入的误差为0.20 mm,对第一行的符号字符嵌入10位隐藏信息“0000111100”绘制出实验所用的条码图形,如图5所示。

从图5我们可以观察到通常情况下,人眼是无法分辨得出隐藏在二维条码PDF417里面的细微变形的,所以结构微调法可以达到比较好隐藏效果。即使在条码图像被放大的情况下,如图5右侧所示,人眼也无法从辨别出暗隐藏信息;可以勉强辨别出明隐藏信息位的情况。

3.2 对实验条码图像的定性分析

经过打印扫描攻击后的条码图像经过放大后(这里用300 dpi的分辨率进行扫描),我们按像素将图像读入,然后将其转化为二值矩阵,发现在矩阵的边缘会有不规则的“0”、“1”出现,这与我们所观察到的图像相符合,如图6所示。但是隐藏信息位矩阵要比噪声位(主要是孤立的黑白点和边缘毛刺)明显得多,这样,我们可以通过检测隐藏信息矩阵的来读出隐藏信息。

3.3 对实验条码图像的定量分析

由实验条码所对应的二值矩阵我们发现基本模块 X 为 9×15 的像素矩阵,我们给PDF417条码的第一行分别加入暗、明隐藏信息,从矩阵中观察到明隐藏信息为 3×3 矩阵,暗隐藏信息为 9×3 矩阵;毛刺表现为矩阵边缘的一行或列不规则的“0”、“1”变化,我们只关心条码边缘毛刺的横向形变,从二值矩阵,我们观察到毛刺的平均形变只有一列。这里的毛刺矩阵最大为 3×1 矩阵;对于明隐藏我们只检测 3×3 矩阵(包括近

表1 所含信息为“PDF417”的二维条码条空组合形式

行数	bsbsbsbs						
第1行	81111113	31111136	51111251	31312223	51211142	51111152	71131121
第2行	81111113	41111315	41131115	42113231	22163111	51111125	71131121
第3行	81111113	11111246	12211532	32411132	12361121	11111345	71131121

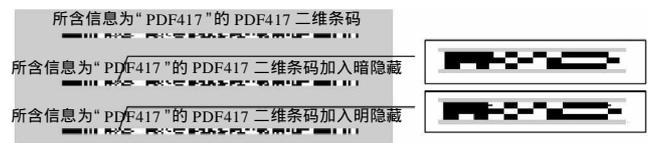


图5 所含信息为“PDF417”的二维条码加入隐藏信息效果对比



图6 经过打印扫描后的加入明隐藏的条码局部放大

似 3×3 的矩阵);对于暗隐藏我们只检测 9×3 矩阵(包括近似 9×3 的矩阵),即可。

3.3.1 计算鲁棒性指标 R

由式(6)可得:

$$\text{对于暗隐藏 } R_{\text{min}}=20 \lg \frac{9 \times 3}{3 \times 1} \approx 19.085 > 0。$$

$$\text{对于明隐藏 } R_{\text{min}}=20 \lg \frac{3 \times 3}{3 \times 1} \approx 9.542 > 0。$$

3.3.2 计算最大隐藏信息量 I

由式(8)、式(9)可得:

$$\text{对于暗隐藏 } I_a=4j \times (k-2)=4 \times 3 \times (7-2)=60(\text{位})。$$

$$\text{对于明隐藏 } I_b=(C_3^3+C_3^3)I_a=6I_a=360(\text{位})。$$

4 条码中明、暗信息隐藏方法的比较

从上面的仿真实验我们可以得出暗、明信息隐藏的比较情况,如表2所示

表2 暗、明信息隐藏方法比较

隐藏类型	最大嵌入信息量	隐藏效果	鲁棒性
暗隐藏	$I_a=4j \times (k-2)$	好	较强
明隐藏	$I_b=(C_3^3+C_3^3)I_a$	较好	较弱

在印有PDF417条码的物品运输过程中,条码有可能遭到污损,我们应该根据条码所在的环境来对,来对明、暗隐藏做出选择。同时应该注意到明隐藏信息矩阵并不是无限可小的,其所对应的像素点阵面积一定要比噪声大,同时必须满足PDF417条码的技术要求,这样的隐藏才有实际意义。

在信息隐藏技术中,鲁棒性和嵌入容量是最为重要的两个指标。从表2我们可以看出暗隐藏比明隐藏嵌入信息量小,但是鲁棒性比明隐藏强,所以,对于信息隐藏技术,增强鲁棒性会以牺牲隐藏容量作为代价,信息隐藏技术的嵌入容量与鲁棒性是一对矛盾。

5 结束语

从以上的实验以及分析过程,我们可以看到本文提出的结构微调法将信息隐藏于常见的PDF417二维条码之中,成功地实现了信息隐藏。本文提出的方法具有以下优点:隐藏信息嵌入算法简单;PDF417条码的十分常见,以条码为载体不易引起察觉,同时加入隐藏信息的PDF417条码图像具有良好的伪装品质,人眼无法分辨,故而其安全性较高;将隐藏信息加入PDF417条码之中,其能够抵御打印扫描攻击,

说明该方法隐藏信息鲁棒性较强；PDF417 条码隐藏信息能力大，而其尺寸较小，便于携带。

参考文献：

- [1] 中国物品编码中心. 条码技术与应用[M]. 北京:清华大学出版社,2003:121-162.
- [2] 李劲. 一种基于水印语义编码的数字图像水印算法[J]. 微型电脑应用,2005,21(11):10-12.
- [3] 田丽华,赵季中. 基于多重压缩与加密的信息隐藏方法[J]. 计算机工程与应用,2006,42(18):164-166.
- [4] 牛夏牧,黄文军. 基于二维条码的信息隐藏技术[J]. 中山大学学报(自然科学版),2004,43(增刊(2)):21-25.
- [5] 中华人民共和国国家标准四一七条码 GB/T 17172-1997 [S]. 国家技术监督局 1997-12-25 批准,1998-08-01.
- [6] 周琳娜,杨义先,郭云彪,等. 基于二值图像的信息隐藏研究综述[J]. 中山大学学报(自然科学版),2004,43(增刊(2)):71-75.
- [7] Stefan K F, Petitcolas A P. 信息隐藏技术——隐写术语数字水印[M]. 吴秋新,译. 北京:人民邮电出版社,2001:49-53.
- [8] William K Pratt. 数字图像处理[M]. 邓鲁华,张延恒,译. 北京:机械工业出版社,2005:269-297.
- [9] 王家文,曹宇. MATLAB6.5 图形图像处理[M]. 北京:国防工业出版社,2004:166-193.
- [10] 张全法,齐永奇,李力. 扫描仪误差分析[J]. 传感器与微系统,2006,25(5):21-24.

(上接第 4777 页)

- [4] Lindstrom P, Pascucci V. Visualization of large terrains made easy[C]. San Diego, California, USA: Proceedings of IEEE Visualization,2001:363-370.
- [5] Balogh András. Real-time visualization of detailed terrain [D]. Budapest: Budapesti Muszaki-es Gazdasagtudományi Egyetem, Master'S Thesis,2003.
- [6] George van Venrooij. Planetary terrain rendering[EB/OL]. <http://www.organicvector.com>.
- [7] Thatcher Ulrich. Rendering massive terrains using chunked level of detail control [C]. Computer Graphics Proceedings, Annual Conference Series. San Antonio, Texas: ACM SIGGRAPH, 2002.
- [8] David Hill. An efficient, hardware-accelerated, level-of-detail rendering technique for large terrains[D]. Toronto: Graduate Department of Computer Science, University of Toronto,2002.
- [9] Jochen Winzen. Interactive visualization of a planetary system [Z]. Karlsruhe: Student Research Project,2003.
- [10] Liqiang Zhang, Chongjun Yanga, Donglin Liua, et al. A web-mapping system for real-time visualization of the global terrain [J]. Computers and Geosciences,2005,31:343-352.
- [11] 郑昊. 数字地图符号化处理技术的研究与实现 [D]. 北京:华北计算技术研究所硕士学位论文,2003.
- [12] Adam Szofran. Global terrain technology for flight simulation [R]. San Jose, California: GDC, 2006.
- [13] Claudio T Silva, Yi-Jen Chiang. Out-of-core algorithms for scientific visualization and computer graphics[C]. Boston, MA: Proceedings of IEEE Visualization, 2002.
- [14] Garret Potts. osgPlanet's docs and source code [CP/OL]. <http://www.ossim.org>,2006.
- [15] Robert Osfield. Open scene graph's introduction and DOCs[CP/OL]. <http://www.openscenegraph.org>.

(上接第 4787 页)

- [2] Hamadi R, Benatallah B. A petri net-based model for web service composition [C]. South Australia: Database Technologies, 14th Australian Database Conference,2003:115-121.
- [3] Gronmo R, Skogan D. Model-driven web services development [C]. Taipei, Taiwan: IEEE International Conference on e-Technology, e-Commerce and e-Service,2004:67-74.
- [4] Tony Andrews, Francisco Curbera. Business process execution language for web services [EB/OL]. <http://www.siebel.com/downloads/misc/bpel.pdf>, 2003.
- [5] Stephen A White. Business process modeling notation(BPMN) [S]. [http://www.bpmn.org/Documents/BPMN % 20V1-0% 20 May% 203% 202004.pdf](http://www.bpmn.org/Documents/BPMN%20V1-0%20May%203%202004.pdf), 2004.
- [6] Hugo Haas, Philippe Le Hégarret. Web service description language (WSDL) version 2.0 part 3: bindings[EB/OL]. <http://www.w3.org/TR/2004/WD-wsd120-bindings-20040803>, 2004.
- [7] Milanovic N, Malek M. Current solutions for web service composition [J]. IEEE Internet Computing,2004,8(6):51-59.
- [8] Sibli R, Mansour N. Testing web services [J]. IEEE Transaction on Software Engineering,2005(5):45-52.
- [9] 李艳华,邱锦伦,宋炯. 用 BPMN 辅助构建 Web Services [J]. 计算机工程,2006(7):43-47.
- [10] 张昌利,李暑瑜,胡正国,等. 一种基于 Web 服务的应用集成解决方案[J]. 微电子学与计算机,2005,22:55-61.
- [11] 张晓青,相春艳. 开放环境下 Web 服务组合研究评述[J]. 烟台师范学院学报(自然科学版),2005,21(3):192-196.