

基于 PDF417 二维条码的顽健数字水印方法

蒋翠玲, 黄帅, 吴安雯, 于文欣, 康周茂, 李永强

(华东理工大学信息科学与工程学院, 上海 200237)

摘要: 利用 PDF417 二维条码其条块边角的特点, 提出了一种基于条块边角的 PDF417 条码水印方法。首先对水印信息整体进行预判, 以选择合适的黑白交界区域作为水印嵌入位置, 根据水印内容对二维码图像的相关信息进行修改, 以实现水印的嵌入和提取。实验结果表明, 此算法相比同类空间域水印算法, 具有更大的水印嵌入容量和更高的图像品质。同时, 具有较好的抗 JPEG 压缩、高斯噪声和打印扫描功能, 具有较强的顽健性。

关键词: 二维条码; 数字水印; 容量; 顽健性

中图分类号: TP309

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2016.00095

Robust digital watermark method based on PDF417 two-dimension code

JIANG Cui-ling, HUANG Shuai, WU An-wen, YU Wen-xin, KANG Zhou-mao, LI Yong-qiang

(School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China)

Abstract: A digital watermark method based on the bar-block-margin characteristics of PDF417 two-dimension code was presented. The proposed method prejudices the watermark information to choose the suitable embedding positions. According to the watermark, the information of the original PDF417 image was modified for embedding and extracting. The experimental results showed that, in comparison with the other methods in the spatial domain, the proposed method has a higher capacity and better watermarked image quality, as well as better robustness against JPEG compression, Gaussian noise and printing and scanning.

Key words: two-dimension code, digital watermark, capacity, robustness

1 引言

二维条码是近年来在一维条码的基础上发展起来的一种新型条码, 从本质上来说是一种特殊的二值图像。二值图像即黑白图像, 其每个像素点只用一位表示, “0”代表黑, “1”代表白,

其存储简单、结构紧凑等特点使二值图像在数据传真、文字识别、条码识别中得到广泛应用^[1,2]。在二维条形码中嵌入数字水印信息, 能够实现如电子票据二维条形码的防伪等。PDF417 二维条码是二维码中的一种, 是一种高密度、高信息含量的便携式数据文件。

收稿日期: 2016-07-12; 修回日期: 2016-08-28。通信作者: 蒋翠玲, cuilingjiang@ecust.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61371150, No.61501187); 华东理工大学国家级大学生创新创业计划基金资助项目 (No.201510251055); 华东理工大学创新创业计划基金资助项目 (No.XY201505, No.2016S16065)

Foundation Items: The National Natural Science Foundation of China (No.61371150, No.61501187), National University Student Innovation Program of East China University of Science and Technology (No.201510251055), Innovation Program of East China University of Science and Technology (No.XY201505, No.2016S16065)

目前, 国内外关于二维码信息隐藏技术的文献不是很多, 研究对象主要是 PDF417 码(portable data file 417) 和 QR (quick response) 码。文献[3]提出结合 Freeman 链码和 B 样条曲线误差对 PDF417 二维码先进行预处理, 再将与该图像内容相关的散列值作为水印信息嵌入, 可进行篡改验证, 容量有所提高但算法实现繁琐, 而且水印信息受限。文献[4]利用变形技术对 PDF417 码符号字符组成单元的宽度进行适量的变动以实现隐藏信息的嵌入, 能有效地抗打印和扫描攻击, 但隐藏信息容量不大。文献[5]提出了一种基于 QR 二维码的证件认证技术。文献[6]在 DCT (discrete cosine transform) 基础上实现了一种结合 PDF417 码的水印嵌入算法, 采用 PDF417 编码将水印信号扩展成比信息带宽大得多的带宽信号, 提高了数字水印的稳健性, 该算法可有效满足水印的透明性和顽健性要求。文献[7]提出了一种改进的二维码水印算法, 在二维条码中嵌入水印信息, 并引入加密方法, 进一步提高了水印嵌入的安全性。文献[8]根据 PDF417 二维条码自身结构特点, 对其进行适量的结构微调, 提出明隐藏和暗隐藏方法, 将信息隐藏进条码当中。文献[9~11]将二维码技术应用于身份证管理和电子票证中以实现防伪目的认证。

2 PDF417 条码的原理及相关工作

2.1 PDF417 条码的原理

PDF417 二维条码的每一个条码字符由 4 个条和 4 个空共 17 个模组构成, 称为一个字码模块, 每个条或空的宽度都必须是单位模组宽度的整数倍, 故称为 PDF417 条码, 如图 1 所示。条码图形上下以及左右各有一个被称为空白区的区域 (图 1 中未显示), 除空白区外, 条码图的组成包括起始符号区、左层指示符、数据符号区、右层指示符、终止符号区。左右空白区的目的是为了将条形码和其他图形符号区别开来; 起始符号区和终止符号区是为了标识条码的起始和终止位置, 在所有的 PDF417 条码中它们都是不变的; 左右层指示符内容相同, 都是为了指示每层图形码的层号; 数据符号区由多个数据符号组成, 编码了图形码所携带的信息。

基于 PDF417 二维条码的数字水印技术一般包括 2 个阶段: 嵌入水印信息的过程和检测提取水印信息的过程。嵌入阶段就是在 PDF417 二维条码中嵌入水印信息, 通常包含预处理和嵌入 2 个过程。提取阶段是相对于嵌入阶段设计合理的提取算法, 将水印信息从 PDF417 二维条码中提取出来。

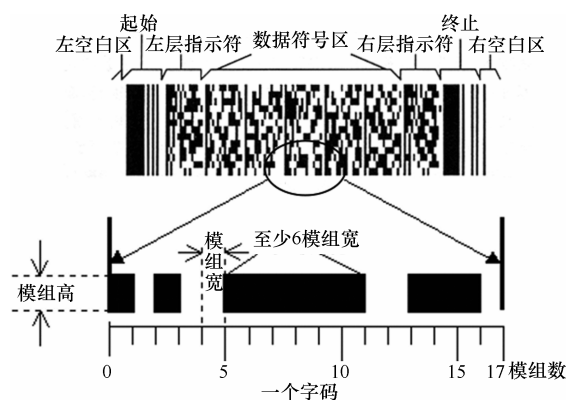


图1 PDF417 二维条码

2.2 相关工作

二维条码 PDF417 图像是一种特殊的二值图像, 在二值图像中嵌入信息时, 必须考虑该像素的邻域状况。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, 则会引起明显的图像失真; 而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。

文献[7]分析了基于二维条码图像图形空间结构的水印嵌入方法, 针对该类方法存在算法略显繁冗和算法结束条件的判断不够清晰等问题进行了改进。文献[8]根据 PDF417 二维条码自身结构特点, 对其进行适量的结构微调, 提出明隐藏和暗隐藏方法, 将信息隐藏进条码当中。相比文献[4,6,7], 文献[8]具有更大的水印容量, 并具有一定的顽健性。本文在文献[8]算法的基础上, 提出了一种改进的基于条码块边角的 PDF417 条码的水印方法。实验结果表明, 该方法不仅提高了水印容量, 兼顾了顽健性, 保证了二维条码的可读性, 并具有抗扫描打印功能。

3 基于条块边角的 PDF417 条码水印方法

本文改进的条块边角空域二维码水印算法在每一个 PDF417 字码的边界嵌入水印, 通过找出每一个 PDF417 字符内模组宽度在 2~6 之间的黑条(或白空), 改变每一个这样的黑条块(或白空)的右下角(或左下角)像素, 以实现水印信息的嵌入和提取。改进算法的二维码水印嵌入流程如图 2 所示。

3.1 嵌入算法

本文改进算法的嵌入思想是将水印信息编码成特定的比特流, 首先对水印比特信息进行预判, 选择嵌入水印到“黑条”还是“白空”, 然后嵌入到相应条块符合要求的边角像素点。

定义 *row*: 像素行数/模组高 *high*, *row* 也称之为层; *column*: 每层中的字码模块个数。

具体步骤如下。

1) 对水印比特信息进行预判, 若所含“0”多于“1”, 则选择“黑条”嵌入; 反之, 则选“白空”嵌入。

2) 略去二维条码图像的上下左右 4 个静空区后, 连续读取每两行(第 *i* 行、第 *i*+1 行)像素值。

3) 比较每 2 个像素行的像素值, 若相同, 则转到步骤 1), 继续往下读取, 若不相同, 则可确定这两行像素为 PDF417 字码的边界处, 可以在此处嵌入水印信息, 同时能够依据 PDF417 二维条码的特性得到每层字码的像素高 *high*, 方便后续寻找下一边界的判断。

4) 通过判断边界紧邻的上一行像素值, 找到边界紧邻的 2 个像素行: *m* 和 *m*+1, 寻找 *m* 行中连续为黑(或白)的像素点达 2~6 个的黑条(或白空), 水印信息将在黑条(或白空)的末(或首)

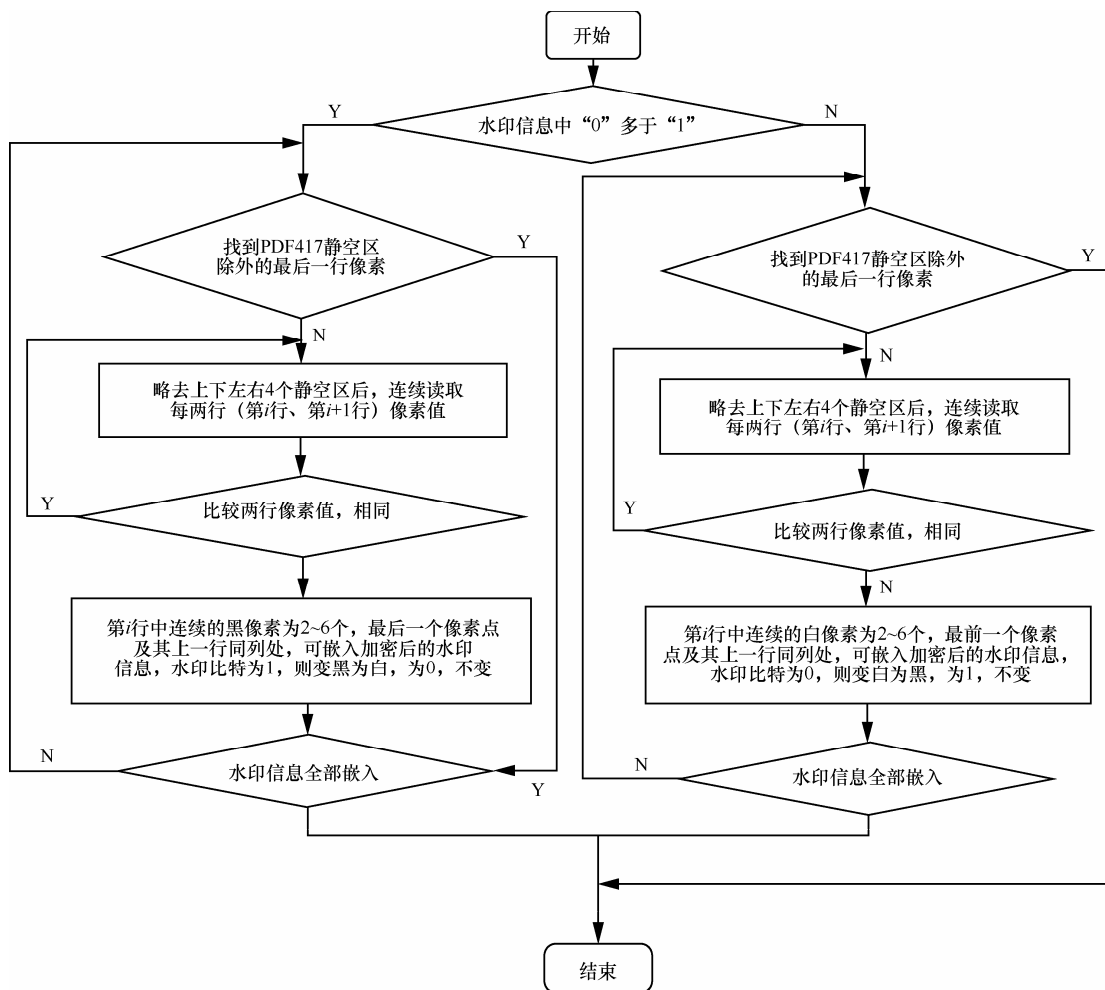


图2 本文改进算法的二维码水印嵌入流程

个像素点处嵌入，及其同列的上一行（ $m-1$ 行）像素点处嵌入。

5) 在加密编码成比特流的待嵌入水印信息中取一个比特，若比特信息为 0，则继续往下进行，不做改动；若比特信息为 1，则修改连续的黑像素点为白像素点；反之，同理进行。

6) 若水印信息已嵌入完毕，结束；否则，若连续点已用完，则转到步骤 2)，继续往下寻找下一边界，若已经无边界，则说明该二维条码图像不足以嵌入全部水印信息，退出。

3.2 提取算法

提取算法的主要思想：将水印信息从嵌入水印后的 PDF417 二维条码图像中提取出来，是水印嵌入的逆过程，由于嵌入水印时只用到了边界紧邻的上方一行像素的信息，并改变了该行某些符合要求的像素点。相应地，在提取的过程中就不需要用到多行像素的信息，只需要根据 PDF417 码的特性，以嵌入过程中得到的每行堆叠字码的像素高为依据判断边界，找到边界紧邻的上方一行，即为水印信息嵌入行。另外，由于嵌入过程中对水印信息做出了预判，相应地，提取过程需要以嵌入过程的预判结果为依据进行水印信息的提取。水印提取流程是嵌入的相反过程。具体算法步骤如下。

1) 以嵌入过程的预判结果为依据，判断提取方法。

2) 忽略二维条码图像的上下左右 4 个静空区后，依据二维码生成规则，连续读取每层字码中的倒数 3 行（该 3 行描述为：第 k 行，第 $k+1$ 行，第 $k+2$ 行）的像素值。

3) 根据读取的每层字码的第 k 行像素值，若连续的黑（或白）像素小于 2 个，则转到步骤 1)，依据步骤 2) 的规则继续读取下一层像素值，若连续的黑（或白）像素为 2~6 个，则可确定这一层字码的第 $k+1$ 行和第 $k+2$ 行像素包含水印信息，且位于 PDF417 字码边界的上方，紧邻边界。

4) 寻找当前层的第 k 行中连续的黑（或白）像素为 2~6 个的黑条（或白空）块，找出相应的第 $k+1$ 行和第 $k+2$ 行中列所对应的黑条（或白空）块的最后（或最前）一个像素点，记录位置。

5) 由水印嵌入算法可知，在记录下的连续像素点中，水印信息就包含在第 $k+1$ 行和第 $k+2$ 行里。若这 2 行像素为黑，则对应水印比特信息为 0；否则，若这 2 行像素为白，则对应水印比特信息为 1；反之也成立。

重复步骤 2)~步骤 5)，直到处理完全部二维条码图像。

4 实验结果与分析

本实验用 Matlab 编程实现了改进的基于条块边角空域二维码水印算法，实验样本选取随机生成的 50~100 幅二维码图像进行测试。以生成携带信息“华东理工大学信息科学与工程学院信息工程专业”的 PDF417 条码图像 ecust.bmp 为例，大小为 43×153 ，水印样本大小为 8×16 的二值图像“E-white.bmp”、超链接“http://www.ecust.edu.cn”等。分别进行常见的 3 种攻击测试：JPEG 压缩、高斯噪声（均值为 0，方差为 0.01）和打印扫描攻击（打印机、扫描仪的分辨率均为 600 dpi）。本实验中，将改进的算法与文献[7]和文献[8]中明隐藏算法进行比较，采用 PSNR（峰值信噪比）和 NC（相似度）性能指标作参数对比。PSNR 反映图像的品质，NC 表示两幅图像的相关程度。

$$PSNR = 10 \lg \frac{255^2}{MSE} \text{ dB} \quad (1)$$

$$MSE = \left(\frac{1}{N} \right)^2 \sum_{i=1}^M \sum_{j=1}^N (X_{ij} - \overline{X}_{ij})^2 \quad (2)$$

$$NC = \left(\frac{\sum_{i=1}^M \sum_{j=1}^N X_{ij} \overline{X}_{ij}}{\sum_{i=1}^M \sum_{j=1}^N X_{ij} X_{ij}} \right) \quad (3)$$

其中， X_{ij} 载体图像像素值； $\overline{X}_{ij}$ 为含水印信息图像像素值； M 、 N 为图像的宽度和高度。

JPEG 压缩是一种常见的图像处理操作，表 1 给出了改进的条块边角空域二维码水印算法与文献[8]中的算法经过 JPEG 压缩时，嵌入水印图像 E-white 时，2 种算法的 PSNR 和 NC 参数的比较。由表 1 可以看出，改进算法具有相对较大的 PSNR，说明在相同条件的 JPEG 压缩下，改进算

法的图像失真度要小于文献[8]中的算法, 即其抗 JPEG 压缩性能更佳。表 2 表明, 改进算法的含水印二维条码进行品质因子为 20~70 的 JPEG 压缩后, 与原二维条码样本图像仍具有较高的相似性。图 3 给出了在不同 JPEG 压缩时的含水印二维码图像, 从视觉上看, 不同 JPEG 压缩因子时的二维码与原始二维码及含水印的二维码图像非常相似。图 4 给出了在不同 JPEG 压缩因子时提取的水印信息, 当压缩因子低于 70 时, 提取的水印会有不同程度的失真, 但仍不影响原有二维条码内容的识读性。

表 1 不同 JPEG 品质因子时算法的 PSNR 性能比较

JPEG 品质因子	PSNR/dB		
	文献[7]方法	文献[8]方法	本文方法
20	18.922 5	18.149 3	20.193 1
30	20.323 2	18.583 9	21.525 6
40	21.367 4	19.019 6	22.513
50	21.983 1	19.320 4	23.139 3
60	22.350 5	19.426 2	23.466
70	22.753 3	19.606 5	24.019 5

表 2 不同 JPEG 品质因子时算法的 NC 性能比较

JPEG 品质因子	NC		
	文献[7]方法	文献[8]方法	本文方法
20	0.945 06	0.957 68	0.938 52
30	0.957 82	0.967 3	0.948 49
40	0.965 58	0.975 3	0.956 5
50	0.971 39	0.980 31	0.961 74
60	0.973 98	0.982 4	0.964 1
70	0.978 63	0.986 79	0.968 19

表 3 给出了 3 种算法的嵌入容量比较。文献[7]的容量公式 $Capacity = 4(row-1)column$, 其中, $row \cdot column$ 是二维码中字符个数, 每个字符由 4 个条和 4 个空组成, 条空交界处可以嵌入 1 bit 信息, 所以理论上一个字符可以嵌入 4 bit 信息, 因为是在行边界嵌入水印, 所以是 $row-1$ 。文献[8]中算法理论最大隐藏容量公式 $Capacity = 4row \cdot column$ ($row \cdot column$ 是二维码中字符个数, 每个字符由 4 个条 4 个空组成, 条空交界处可以嵌入 1 bit 信息, 所以理论上一个字符可以嵌入 4 bit 信息), 本文改进算法的理论最大隐藏容量公式可表示为 $Capacity = 4row \cdot column \cdot 2$ (因为采用明隐藏方式, 每个黑白交界处可以嵌入 2 bit 信息, 所

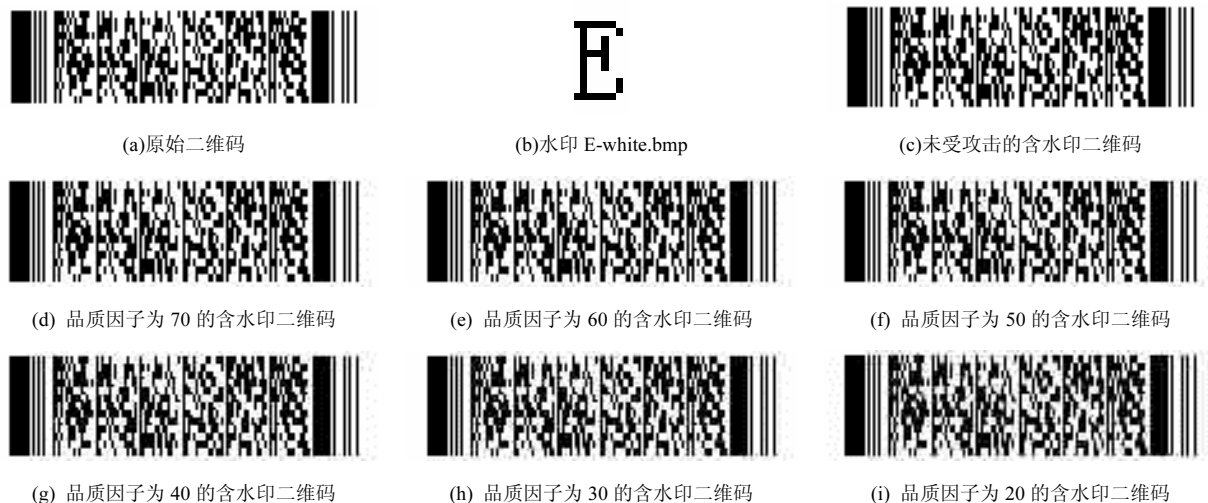


图 3 不同 JPEG 压缩 (品质因子 20%~70%) 后的含水印的二维条码



图 4 不同 JPEG 压缩 (压缩因子 20%~70%) 后提取出的水印

以嵌入信息总量为 2 倍)。相比文献[8], 本文改进算法容量扩大了 2 倍。当 $high \geq 3$ 时, $high$ 值越大, 本算法效果越好。表 4 给出了在均值为 0, 方差为 0.01 时抗高斯噪声的 3 种方法性能对比。由表 4 可以看出, 改进算法在同等高斯噪声攻击下, $PSNR$ 和 NC 值略小于其他 2 种方法, 具有可接受的抗 Gaussian 噪声性能。

表 3 最大嵌入容量比较

方法	嵌入容量/bit		
	http://www.ecust.edu.cn	E-black.bmp	E-white.bmp
文献[7]	288	288	288
文献[8]	312	312	312
本文方法	624	624	624

表 4 抗 Gaussian 噪声的性能对比

方法	参数	
	$PSNR/dB$	相似度 NC
文献[7]	20.310 8	0.952 86
文献[8]	20.314 2	0.953 91
本文方法	20.298 2	0.952 8

图 5 表明, 当加入高斯噪声攻击时, 由于图像有损, 提取出的水印不是二值图像, 因而直观显示与原始水印有一定差别, 但把提取出的水印转化为二值图像时, 得到的水印与原始水印一致。



图 5 高斯噪声攻击

在打印机、扫描仪的分辨率均为 600 dpi 的条件下, 嵌入水印后进行抗打印扫描测试的 PDF417 二维条码图像如图 6 所示, 由于打印扫描后图像有一定的失真度, 可由视觉直观看出, 得到的二维条码图像放大后可看出有一定的边缘毛刺。由于像素分辨率变化, 对应水印提取过程需要先对扫描图像进行二值化处理, 然后根据算法找到水印嵌入行, 逐行依次提取。扫描后提取的水印如图 7 所示,

可以看出, 提取出的水印与原始水印完全一致, 故该算法具有较好的抗打印扫描功能。



图 6 本文方法的嵌入水印 E-white.bmp 后进行打印扫描的二维条码和其局部放大图



图 7 打印扫描前后的水印图像对比

同样, 使用 PDF417 二维条码解码软件解码后得到相关二维条码携带的原始信息, 未加入攻击的嵌水印图像和加入 3 种典型攻击后得到的含水印图像均能够保持不失真解码, 尤其当 JPEG 压缩因子为 20 时, 二维条码仍然能够保持不失真解码, 具有较高的识读性, 保证了二维条码携带数据的完整性。

5 结束语

本文提出了一种基于条块边角的 PDF417 条码水印方法, 结合 PDF417 二维条码自身的特点, 利用其条块边角特点来进行水印的嵌入和提取。实验结果表明, 本文提出的算法具有较强的顽健性, 同时具有较好的抗打印扫描功能, 具有一定的实用价值。二维条码与数字水印技术相结合, 并推广应用到具体相关领域是下一步深入研究的重点。

参考文献:

- [1] 贾助军. 基于二维条形码的数字水印技术[D]. 武汉: 武汉理工大学, 2012.
JIANG Z J. Digital watermarked technique based on two-dimension code[D]. Wuhan: Wuhan University of Science and Technology, 2012.
- [2] IANG X H, GANG W, JOHAN S. Fragment-type UHF RFID tag embedded in QR barcode label[J]. Electronics Letters, 2015, 51(4): 313-315.
- [3] 吴桐树. 基于二维码的数字水印技术研究[D]. 杭州: 浙江理工大学, 2015.
WU T S. Study on digital watermarked technique based on two-dimension code[D]. Hangzhou: Zhejiang University of Science and technology, 2015.

- [4] 牛夏牧, 黄文军, 吴迪, 等. 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(22): 21-25.
NIU X M, HUANG W J, WU D, et al. Information hiding technique based on 2D barcode[J]. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2004, 43(22): 21-25.
- [5] 潘璐. 基于二维码的证件认证技术研究[实现][D]. 北京: 北京邮电大学, 2014.
PAN L. Research and implementation of a technology of certificate authentication based on two-dimension code[D]. Beijing: Beijing University of Posts and Telecommunications, 2014.
- [6] 张基宏, 肖薇薇, 纪震. 基于二维条码 PDF417 的数字图像水印算法[J]. 深圳大学学报(理工版), 2002, 19(1): 1-6.
ZHANG J H, XIAO W W, JI Z. Digital image watermarked algorithm based on two-dimension code[J]. Journal of Shenzhen University (Science & Engineering), 2002, 19(1): 1-6.
- [7] 魏伟光. 二维条码空间域数字水印嵌入方法研究[J]. 北方工业大学学报, 2010, 22(3): 5-9.
WEI W G. Research of method of digital watermarking in 2-D barcode based on space Structure[J]. Journal of North China University of Technology, 2010, 22(3): 5-9.
- [8] 赵博. 二维条码研究[D]. 西安: 西安电子科技大学, 2007.
ZHAO B. Study on 2D barcode[D]. Xi'an: Xidian University, 2007.
- [9] 林海. 二维码技术在身份证管理中的应用研究[D]. 天津: 天津大学, 2012.
LIN H. Research and application of identification management base on two-dimension code[D]. Tianjin: Tianjin University, 2012.
- [10] 欧阳麒. 移动二维条码电子票证安全数字水印方案研究[D]. 上海: 上海交通大学, 2007.
OUYANG Q. A new security framework of mobile twodimensional bar code electronic ticket system-based on digital watermark[D]. Shanghai: Shanghai Jiaotong University, 2007.
- [11] 武莎莎. 基于 PDF417 二维条码的数字水印技术研究[D]. 南京: 南京航空航天大学, 2010.
WU S S. Study on digital watermarked technique based on PDF417 two-dimension code[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2010.

作者简介:



蒋翠玲(1976-), 女, 湖北随州人, 博士, 华东理工大学讲师, 主要研究方向为信息隐藏、图像处理。

黄帅(1994-), 男, 安徽安庆人, 华东理工大学本科生, 主要研究方向为图像处理。

吴安雯(1994-), 女, 江苏南通人, 华东理工大学本科生, 主要研究方向为图像处理。

于文欣(1995-), 女, 山东烟台人, 华东理工大学本科生, 主要研究方向为图像处理。

康周茂(1995-), 男, 甘肃天水人, 华东理工大学本科生, 主要研究方向为图像处理。

李永强(1995-), 男, 福建福鼎人, 华东理工大学本科生, 主要研究方向为图像处理。