

文章编号: 2096-1472(2017)-05-01-05

## 面向防伪认证的PDF417二维码水印技术的实现

李永强, 马雪娇, 杨 迈, 王家峰

(华东理工大学信息科学与工程学院, 上海 201424)

**摘要:** 利用PDF417二维条码以模组为原子的特点, 提出一种基于模组的PDF417条码水印的方法。该方法首先将水印转换成二进制数据, 再选取PDF417二维条码中符合特定条件的区域作为水印信息的嵌入点, 根据待嵌入的水印信息值修改二维码图像的像素值, 实现水印信息的嵌入和提取。实验结果表明, 相比同类算法, 根据本文提出的方法嵌入水印后的图像具有更高的质量和更大的水印容量。同时, 该方法具有较好的抗JPEG压缩和抗打印扫描的鲁棒性, 但是对设备分辨率要求较高。

**关键词:** 二维条码; PDF417; 数字水印; 容量; 顽健性

**中图分类号:** TP391 **文献标识码:** A

## Realization of PDF417 Two-Dimensional Code Watermarking Technology for Security Certification

LI Yongqiang, MA Xuejiao, YANG Mai, WANG Jiafeng

(College of Information Science and Engineering, East China University of Science and Technology, Shanghai 201424, China)

**Abstract:** Utilizing the characteristic that the PDF417 two-dimensional code takes the module as an atom, the paper proposes a digital watermark method based on the module of PDF417 two-dimensional code. The method first converts the watermark into binary data, and then selects the eligible area in the PDF417 two-dimensional code as the embedding point of the watermark. The pixel value of the two-dimensional image is modified according to the watermark information value to be embedded to realize the embedding and extraction of the watermark. The experimental results show that, compared with similar algorithms, the proposed method in this paper has higher image quality and larger watermark capacity. Additionally, the method has good robustness against JPEG compression, printing and scanning, but with higher resolution requirements for the device.

**Keywords:** two-dimensional code; PDF417; digital watermark; capacity; robustness

### 1 引言(Introduction)

二维条码本质上是一种特殊的二值图像, 像素点集中只有0和1, 即黑和白。结构简单, 易于生成和读取, 在商品识别、票证识别等方面有广泛运用。向二维条码中添加水印可以达到证件防伪和信息隐藏的目的。最常见的二维条码有QR(Quick Response)码和PDF417, 本文重点研究的是PDF417(Portable Data File 417)码。

文献[1]提出了基于QR码的盲数字图像水印技术, 使用密钥和抗攻击嵌入域提高的鲁棒性和视觉不可见性。文献[2]使用密钥技术对水印信息进行加密处理, 提升了水印信息的安全性。文献[3]提出了基于DWT和SVD的防伪QR码水印算法, 防伪性能和水印质量高。文献[4]提出了一种先用freeman链码

和B样条曲线误差预处理后再将水印信息内容的散列值嵌入到图像中的方法, 容量有所提高, 但其算法实现复杂, 信息受限。文献[5]使用小波变换在QR码中实现了水印嵌入。文献[6]利用变形技术, 适当的改变原有图像的宽度已达到嵌入水印的目的, 但水印容量较低。文献[7]在DCT(Discrete Cosine Transform)的基础上实现了一种利用了PDF417编码方法的水印嵌入算法, 其水印信息的容量和稳健性有较大提高。文献[8]提出了一种基于PDF417条码自身结构特点的水印信息隐藏方法, 将条空结构进行适量的微调, 在保证条码信息不丢失不畸变的前提下, 将信息隐藏到条码中, 但其抗水平畸变性能差。文献[9]在文献[8]的基础上修改了识别方法, 牺牲容量换取了较好的鲁棒性。本文提出了一种基于文献[8]的竖直

方向微调条空结构的水印明隐藏方法。

### 2 相关工作(Related work)

#### 2.1 PDF417二维条码的基本原理

PDF417二维条码是由王寅君博士发明的，其识别的原子单位是模组，数字图像中每个模组可有若干个像素构成，每个PDF417的码字包含了长度不等的4个黑条与4个白条，单条宽度不超过6个模组宽，总数为17个模组宽度，故得名417。如图1所示，完整的PDF417条码由起始区、左层指示器、数据区、右层指示区、结束区组成。起始区和结束区标记了条码的开始和结束位置，左右层指示区表明了解码所需的(Li, Ri)，数据区存放条码所携带的信息。



图1 PDF417条码结构

Fig.1 The structure of PDF417 bar code

#### 2.2 相关文献方法

面向防伪认证的PDF417水印技术主要包含嵌入和提取两个步骤。嵌入阶段需要计算容量和水印长度并进行一定的预处理，如添加校验信息、加密和编码压缩，将水印信息按一定的算法嵌入条码中。而提取阶段则要提取水印信息并进行相对应的解压、解密和校验。

文献[8]提出了一种基于条空的PDF417码的水印方法，其本质是在PDF417码允许误差范围内调整每个模块的宽度，其与标准值的差表明是否为1，连起来即为比特流形式的水印信息。其原理图如图所示。这种方法可以顺利嵌入水印，并且暗隐藏模式下在视觉上不易被发现。但是由于利用了PDF417码允许误差的特点，导致实际图像发生误差允许内变形后水印信息误读，数据读取异常，鲁棒性较差。

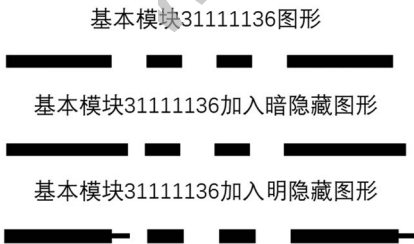


图2 基于条空的水印方法

Fig.2 The watermark method based on bar blanks

文献[9]在文献[8]的基础下提出了一种新的基于码字条空误差的水印判断方法，根据一个码字中嵌入水印导致误差的正负来判断1和0，其原理图如图3所示。这种方法大大提升了

水印信息的鲁棒性，但是容量也大打折扣。

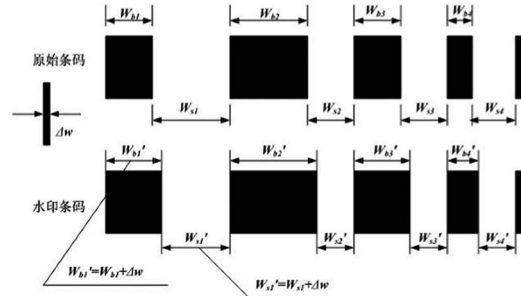


图3 基于码字条空误差的判断方法

Fig.3 The judgment method based on error of the bar blanks of code words

### 3 本文提出的算法(The algorithm proposed in this paper)

本文提出了一种基于文献[8]的改进算法，在竖直方向上调条空结构，几乎不影响水平方向的条空结构。使用本方法可以增加鲁棒性又不会缺失大量的容量。

#### 3.1 预处理

##### 3.1.1 图像二值化

图像在传播的过程中经常会被统一转换成彩色图像，如在安卓系统中图片在显示时会统一转换成bitmap位图，其色彩方案是ARGB。为了降低程序运行的难度和处理的数据量，需要先将其二值化处理。

首先分别读取R、G、B的数值，然后根据式(1)将图像转换成灰度图像，再根据式(2)的阈值将其二值化，其中A参数不影响结果，因此忽略。

$$X = 0.3 * R + 0.59 * G + 0.11 * B \tag{1}$$

$$C = \begin{cases} 0, & X < \text{阈值} \\ 255, & X \geq \text{阈值} \end{cases} \tag{2}$$

另外一种获得二值化图像的方法是将像素与0x00808080按位与(80H=1000000B=128D)，获得的值与0进行比较从而快速判断像素的黑白值。由于二维码图像的直方图会有明显且较宽的分界，此方法不会有太大的误差，处理速度更快。

表1 两种颜色二值化后的结果

Tab.1 The results of binaryzation for two colors

ARGB	A	R	G	B	ARGB	A	R	G	B
原始值(灰色)	FF	6F	6F	6F	原始值(乳白色)	FF	C9	DF	8A
按位与后	00	00	00	00	按位与后	00	80	80	80
判断结果			黑色		判断结果			白色	

##### 3.1.2 角度矫正

二维码在扫描后的图片会产生不同程度空间的倾斜，因

此在处理前需要先将其矫正。利用Hough变换 $\rho = x \cos \theta + y \sin \theta$ 求出其边界再将其旋转到水平位置<sup>[4]</sup>。

### 3.1.3 水印预处理

若对水印本身有加密需要，可以再进行一定的加密算法。为了防止水印被篡改，还可以添加校验位，校验失败即认为水印是伪造的。如使用BASE64进行预编码。

### 3.1.4 图像定位和码字检测

图像经过预处理后可认为是理想情况，可以进行水印操作。首先确定二维码的四个顶点位置和模块大小。从头和尾进行定位两个对角点即可得到二维码有效区域。PDF417码拥有固定的起始区和结束区，因此可以快速得到模块宽度。对于未加工过的模块，其模块高度和条码高度必然满足式(3)的关系：

$$H = n * h, \quad n \text{ 为正整数} \quad (3)$$

因此可以遍历数据区每个模块的高度，排除不满足式(3)后选取最小的值即为模块高度。由于PDF417码中的每个码字都有4个黑条和4个白条，数据区的每一行都不会出现全白或全黑的情况，此法具有较高的可靠性。

### 3.1.5 色块识别

以一个16\*4的模块为例，读取其中一个色块即可获得模块的颜色信息，但是如果在这个色块出现严重噪声将会导致识别错误，因此可以读取3—5个点以排除误差。16\*4模块如果进行边缘取反处理，那么其高度应为2即其水印部分只有2\*4，选择3—5个点即可可靠获得模块颜色信息。

### 3.1.6 可用空间判断

本文提出的算法对嵌入空间有一定要求，上下同色的边界处改变色块会造成视觉显著变化，不适合隐藏，因此需要找出所有上下模组颜色不同的区域作为水印空间。这些水印空间有两种情况，即上黑下白和上白下黑。嵌入水印信息1时在边界白色区域添加一定高度黑条。

改变色块的高度影响分辨能力，越高越容易被发现，甚至影响PDF417码本身的识别，越低越容易导致程序本身无法识别。因此，本文中改变色块区域的高度均采用式(4)，其中H为码字高，[]表示取整数部分。

$$h = [H * 0.1] + 1, \quad H > 1 \quad (4)$$

## 3.2 本文提出的水印嵌入算法

具体嵌入步骤如下：

- (1)生成原始二维码和水印信息的比特流。
- (2)对原始二维码进行定位。

(3)遍历可用空间。

(4)读取一位水印信息，为0不修改，为1添加黑条。

(5)判断可用空间是否用完，或水印是否全嵌入完毕。若空间有剩余，水印未嵌入完毕，重复步骤(3)。

(6)输出嵌入水印后的二维码图像。

程序流程图如图4所示。

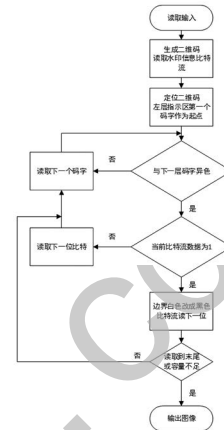


图4 本文提出的水印嵌入算法流程图

Fig.4 The flow chart of the algorithm for watermark embedding proposed in this paper

## 3.3 水印的提取

水印的提取算法与嵌入类似，具体步骤如下：

(1)图像预处理。

(2)对二维码图像进行定位。

(3)遍历可用空间。

(4)判断是否嵌入水印，嵌入水印比特流添加1，否则添加0。

(5)可用空间仍有空间则重复步骤(3)。

(6)输出比特流。

算法流程图如图5所示。

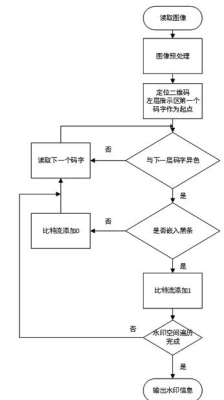


图5 本文提出的水印读取算法流程图

Fig.5 The flow chart of the algorithm for watermark extraction proposed in this paper

## 4 实验结果及性能分析(Experimental results and performance analysis)

### 4.1 鲁棒性分析

JPEG是一种对图像的有损压缩，量化质量因子Q为1—100，Q越高，信息损失越少。本文对一个嵌入水印的PDF417码进行了Q=10、Q=40、Q=70的压缩攻击，均顺利得到正确的条码信息和水印信息。

表2 不同Q下二维码读取结果

Tab.2 The results of reading two-dimensional code in different values of Q

质量因子	(a)原始	(b)Q=70	(c)Q=40	(d)Q=10
PDF417码图像				
条码信息	华东理工大学	华东理工大学	华东理工大学	华东理工大学
水印信息	信息学院	信息学院	信息学院	信息学院

通过放大表2中d的PDF417码图像，可以发现图像已经有很大程度的失真。

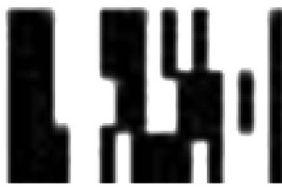


图6 Q=10的局部放大图

Fig.6 The partial enlarged drawing when Q is equals to 10

打印扫描会使图像损失信息、图形失真，在打印精度低的打印机和扫描机上尤为显著。



(a)打印扫描后图像整体图

(a)The global image after printing and scanning



(b)打印扫描后图像局部放大图

(b)The partial enlarged drawing after printing and scanning



(c)预处理后局部放大图

(c)The partial enlarged drawing after pretreatment

图7 打印扫描攻击结果分析

Fig.7 The analysis on results of printing and scanning attacks

从图7(a)可以看出打印扫描后图像有明显噪点，从图7(b)可以看出有些区域的噪点明显泛白，尤其是边缘的噪点会导致识别错误。但是图像本身的信息还很好的保留着。通过预处理可以将其转换成适合提取信息的理想图形。

从图7(c)可以看到预处理后图形锐利了许多，但仍然有少量毛刺存在。毛刺所占比例很小，通过3.1.5中介绍的色块识别规则可以避免此毛刺影响。通过图8可以发现水印信息成功无损提取。



图8 水印信息的提取

Fig.8 The extraction of watermark information

### 4.2 水印容量

水印容量是根据图像本身的可利用空间计算的，其只能在二维码生成之后再计算空间并嵌入，不同的二维码可利用空间不同，可能有较大差异，但是可用通过一系列观测值推测其容量期望和水印内容大小的关系。例如6个汉字的“华东理工大学”生成的二维码水印容量为296bit，6个汉字的“信息科学技术”生成的二维码水印容量为315bit。

在常用汉字表中抽取880组含有6个汉字的组合进行生成容量的测算，其容量平均值为313.0057bit，方差273.1329，最大值379bit，最小值238bit，极差141bit，散点图如图9所示。相同情况下文献[8]的平均容量为192bit，文献[9]的容量为48bit。

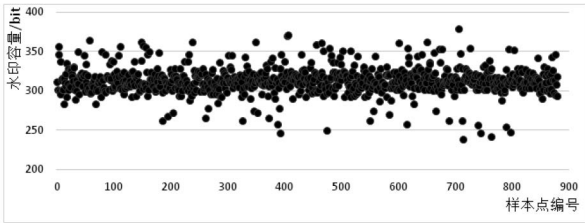


图9 不同的6个汉字组合下水印容量散点图

Fig.9 The scatter plot of watermark capacity in different combinations of 6 Chinese characters

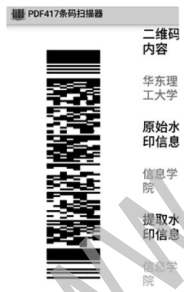
### 4.3 基于PDF417二维码水印的APP实现

在安卓环境下，使用开源项目zxing的二维码框架用于生成普通的PDF417条码，开发了一款能够读取、生成PDF417码，根据本文算法添加读取水印的安卓APP程序。如图10所示，该程序基本完成了展示成果的功能。



(a)PDF417码生成器

(a)The generator for PDF417 code



(b)条码生成结果

(b)The result of barcode generation



(c)(b)中条码的扫描结果

(c)The result of scanning the bar code of (b)

图10 APP主要演示界面

Fig.10 The main presentation interface of APP

## 5 结论(Conclusion)

本文提出了以一种垂直方向上的微调条空的水印方法。该水印方法实现了预期性能指标，由于其为明隐藏，放大后可能会引起视觉上的明显异常。要消除此异常，可以使用暗隐藏，即将水印空间集中去除左右两排码字与本身相同的子集，但是会减少水印容量。此方法对设备分辨率和图像矫正结果要求较高，否则会导致结果出错。

## 参考文献(References)

- [1] Thulasidharan P P,Nair M S.QR code based blind digital image watermarking with attack detection code[J].AEU-International Journal of Electronics and Communications, 2015,69(7):1074-1084.
- [2] Lin P Y,Chen Y H.High payload secret hiding technology for QR codes[J].Eurasip Journal on Image & Video Processing,2017(1):14.
- [3] Qin J H,et al.Anti-fake Digital Watermarking Algorithm Based on QR Codes and DWT[J].International Journal of Network Security,2016,18(6):1102-1108.
- [4] 吴桐树.基于二维条码的数字水印技术研究[D].浙江理工大学,2015.
- [5] 李智伟.基于小波变换域的双重数字水印算法研究[D].安徽理工大学,2016.
- [6] 牛夏牧,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(s2):21-25.
- [7] 张基宏,肖薇薇.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.
- [8] 赵博.二维条码研究[D].西安电子科技大学,2007.
- [9] 何利军.基于信息隐藏技术的防伪系统研究与实现[D].华中科技大学,2008.

## 作者简介:

李永强(1995-), 男, 本科生.研究领域: 图像处理.  
 马雪娇(1995-), 女, 本科生.研究领域: 软件开发.  
 杨 迈(1997-), 女, 本科生.研究领域: 软件开发.  
 王家峰(1995-), 男, 本科生.研究领域: 软件开发.