

# 低冗余二维码汉字编码研究

杨康 袁海东 郭渊博

(中国人民解放军信息工程大学 郑州 450001)

**摘要** 随着二维码应用领域的不断扩大,二维码的优化和改进也势在必行。二维码在汉字编码过程中采用定长编码模式,忽略了汉字使用频率对二维码汉字编码效率的影响,存在着较大的编码冗余。采用变长编码算法可以减少高频汉字的编码长度,降低二维码汉字编码的平均编码长度,增加编码容量。首先,结合常用汉字的使用频率,对二维码的汉字编码进行分段处理,分析并制定基于汉字使用频率的变长汉字编码表;其次,在不破坏二维码原有编码结构的情况下,仅对二维码的汉字编码采用变长编码方式;最后,分析并比较二维码定长汉字编码算法和二维码变长编码算法的时间性能和空间性能,给出优化及改进思路。实验证明,二维码变长编码算法可降低二维码汉字编码 18.4% 的冗余。

**关键词** 低冗余,二维码编码,汉字编码

**中图法分类号** C      **文献标识码** A

## Research on Lower Redundancy of Two-dimension Barcode Encoding in Chinese Characters

YANG Kang YUAN Hai-dong GUO Yuan-bo

(PLA Information Engineering University, Zhengzhou 450001, China)

**Abstract** With the expansion of two-dimension barcode application field, optimization and improvement of two-dimension barcode are imperative. A fixed length coding algorithm has been used in Chinese character encoding, there is a large redundancy because the algorithm ignores the frequency of Chinese character which has an impact on the efficiency of two-dimension barcode encoding. A variable length coding algorithm can reduce the length of some commonly used Chinese characters. At the same time, it can reduce the average length of two-dimension barcode in Chinese characters, correspondingly, it can increase the maximum number of Chinese character encoded in the two-dimension barcode. First of all, according to the frequency of commonly used Chinese characters, all of the commonly used Chinese characters were divided into different segments and encoded independently, and a table for variable length coding algorithm was created. Then, all of the Chinese characters based on the variable length coding algorithm were encoded without breaking the standard two-dimension barcode encoding rules. At last, the time performance and spatial performance of fixed length and variable length coding algorithm were analyzed. Experiments show that the variable length coding algorithm can cut 18.4% redundancy in Chinese character encoding.

**Keywords** Lower redundancy, Two-dimension barcode encoding, Encoding in Chinese character

近年来,二维码以其优越的性能迅速取代一维码成为一种重要的信息传输载体,在信息传递、移动支付、数字凭证方面得到了广泛的应用。随着二维码相关应用的不断深化,人们对二维码的研究主要集中在编码、识别和加密这3个方面。在编码方面,标准二维码汉字编码采用定长编码,有利于降低程序的复杂度,提高程序的执行效率,降低编码的复杂度。采用定长编码算法就是将所有的汉字都作为一个孤立平等的个体,忽略了汉字在实际使用过程中的上下文关系及使用者的使用习惯和使用语境,人为地等值化了汉字的使用频率,存在着较大的编码冗余。结合前人对于汉字编码的优化思路和二维码编码的客观条件,本文提出一种基于汉字使用频率的变长二维码汉字编码算法,在考虑二维码编码效率的同时有效减少了二维码的汉字编码冗余,提高了二维码汉字的编码容量。

本文首先介绍前人在二维码汉字编码方面上的相关研究成果;其次描述变长二维码汉字编码算法的基本原理和架构;最后针对变长二维码汉字编码算法的有效性进行了实验和分析,客观地对二维码汉字编码的性能进行了评价。

## 1 二维码汉字编码的研究背景及相关成果

Two-dimension barcode 码是由日本 Denso 公司于 1994 年 9 月研制的一种矩阵二维码符号,它除具有一维条码及其他二维码所具有的信息容量大、可靠性高、可表示汉字及图像多种文字信息、保密防伪性强等优点外,还可高效地表示汉字,相同内容时其尺寸小于相同密度的 PDF417 条码。目前市场上的大部分条码打印机都支持 Two-dimension barcode 条码。我国对二维码技术的研究开始于 1993 年<sup>[1]</sup>。中国物

本文受国家自然科学基金(61501515)资助。

杨康(1992—),男,硕士生,主要研究方向为信息安全,E-mail:1275081820@qq.com;袁海东(1972—),男,博士,副教授,主要研究方向为图像处理;郭渊博(1975—),男,博士,教授,主要研究方向为大数据分析。

品编码中心对几种常用的二维码(PDF417, Two-dimension barcode, DataMatrix, MaxiCode, Code49, Code16K, CodeOne)的技术规范进行了翻译和跟踪研究。在消化国外相关技术资料的基础上,制定了两个二维码的国家标准:GB/T17172-1997《四一七条码》和 GB/T18284-2000《快速响应矩阵码》。我国原有的二维码国家标准是从美国的 PDF417 码和日本的 QR 码(Quick Response, 快速响应)翻译过来的。随着国内技术的不断创新,自主二维码技术开始出现,并得到了国家的认同和推荐。以国际自动识别制造商协会(AIMI)发布的《PDF417 规范》为基础,1997 年 12 月正式颁布了由中国物品编码中心负责编制的国家标准《四一七条码》(GB/T17172-1997)。2000 年 12 月颁布了基于日本 Two-dimension barcode 的国标 GB/T18284-2000《快速响应矩阵码》<sup>[2]</sup>。

由于标准的二维码汉字编码规则采用 13bit 定长编码方式,而二维码的编码位流数是唯一确定的,限制了二维码编码容量的提升。Subpratsavee P<sup>[3]</sup>对用户信息的格式进行了压缩处理,实现了二维码编码在数据容量上的提升。Nutchana Taveerad<sup>[4]</sup>通过将黑白图像转换为 HSV 图像,实现了二维码彩色图像的编码与应用,有效地提高了二维码编码的容量。Prathibha. N. Pillai<sup>[5]</sup>通过使用不同的信道共用同一块二维码图像空间,使用 3 种不同的颜色来表示相应信道上的信息,提升了二维码的编码容量。Max E. Vizcarra Melgar<sup>[6]</sup>通过分析二维码的编码密度,进一步证明了彩色图像对于二维码数据容量提升的有效性。Umara M M<sup>[7-8]</sup>采用多路复用技

术将多个数据位用新的码字表示,减少了二维码数据信息的存储空间,提高了二维码的编码数据量。陈元枝<sup>[9]</sup>通过使用彩色编码算法和 Gzip 压缩算法,在彩色图像的基础上进一步地提高了二维码的编码容量。Jen-Shiun Chiang<sup>[10]</sup>通过使用三维立体图像,分别用立体图的正面、侧面、顶面来放置二维码图像,在仅使用黑白两色的情况下有效提高了二维码的存储容量。夏春蕾<sup>[11]</sup>提出了一种将二维码汉字与其他字符混合编码的编码算法,实现了二维码编码数据量的压缩,提高了二维码的汉字编码容量。邹敏<sup>[12]</sup>使用 Huffman 编码算法来处理十六进制编码中间数据,有效地增加了二维码汉字编码的容量。

## 2 二维码变长汉字编码算法原理

标准二维码编码结构如图 1 所示,由于二维码变长编码算法仅对二维码的汉字编码部分加以处理,并不涉及字母、数字、标点符号类型的字符,为了保证二维码平台的通用性,对于二维码的非汉字字符依然采用原有的编码方式。对于汉字部分采用与定长编码相似的编码结构:在输入的数据字符的二进制队列前加上模式指示符(1101)、中国汉字子集指示符(4 位,对应 GB2312 的子集指示符为 0001)和字符计数指示符的二进制表示  $C(8, 10 \text{ 或 } 12 \text{ 位})$ ,在数据字段,将每一个汉字的编码用其在变长编码算法中所在的块号和块内偏移来代替,设输入数据共  $D$  个汉字。其基本结构如图 2 所示。

|                 |                   |                        |                         |
|-----------------|-------------------|------------------------|-------------------------|
| 模式指示符<br>(4bit) | 汉字子集指示符<br>(4bit) | 字符计数指示符位数<br>(8~12bit) | 输入的数据串 13bit/字,共 130bit |
|-----------------|-------------------|------------------------|-------------------------|

图 1 标准二维码编码结构图

|                 |                   |                        |                            |                              |     |                                |                                  |
|-----------------|-------------------|------------------------|----------------------------|------------------------------|-----|--------------------------------|----------------------------------|
| 模式指示符<br>(4bit) | 汉字子集指示符<br>(4bit) | 字符计数指示符位数<br>(8~12bit) | 变长编码汉字<br>所在块块号<br>(第一个汉字) | 变长编码汉字<br>所在块块内地址<br>(第一个汉字) | ... | 变长编码汉字<br>所在块块号<br>(第 $D$ 个汉字) | 变长编码汉字<br>所在块块内地址<br>(第 $D$ 个汉字) |
|-----------------|-------------------|------------------------|----------------------------|------------------------------|-----|--------------------------------|----------------------------------|

图 2 二维码变长汉字编码结构图

假设常用汉字使用频率表的汉字总数为  $N$ ;为了减小高频汉字的编码长度,将常用汉字使用频率表进行分块编码。设块数为  $n$ ,每一块的块号为  $K_i$ ,每一块的汉字总数为  $N_i$ ,其中  $i$  在 1 到  $n$  之间取值,则:

$$N = \sum_{i=1}^n N_i, 1 \leq i \leq n \quad (1)$$

编码过程中,每一块的块内编码从 1 到  $N_i$  依次编码,编码的位数  $S_i = \log_2 N_i$ ,为减少编码的冗余,除第  $n$  块外,将每一块的汉字总数  $N_i$  均选择 2 的  $S_i$  次幂,即  $N_i = 2^{S_i}$ ,则式(2)可转化为:

$$N = \sum_{i=1}^{n-1} 2^{S_i} + N_n, 1 \leq i \leq n-1 \quad (2)$$

对于任意的输入汉字  $A$ ,设  $A$  在汉字使用频率表中的偏移量为  $X$ ,在分块后的块号为  $K_A$ ,块内的偏移地址为  $X_A$ ,则:

$$X_A = X - \sum_{i=1}^{K_A-1} 2^{S_i}, 1 \leq i \leq K_A - 1 \quad (3)$$

$$S_{K_A} = \log_2 N_{K_A} \quad (4)$$

根据图 2 中所使用的变长汉字编码结构图,变长汉字编码的位流长度计算公式为:

$$B = 4 + 4 + C + \sum_{j=1}^D (\log_2 n + \log_2 N_{K_A}), 1 \leq j \leq D \quad (5)$$

其中,  $B$  为位流的位数,  $C$  为字符计数指示符的位数,  $D$  为输入的数据字符数。

为了分析比较变长汉字编码算法和定长编码算法在空间性能上的优劣,需比较变长编码算法的编码总长度与定长编码算法总长度之间的差值  $Diff$ 。

$$Diff = 13D - \sum_{j=1}^D (\log_2 n + \log_2 N_{K_A}), 1 \leq j \leq D \quad (6)$$

当  $Diff > 0$  时,变长汉字编码算法的空间性能较优;当  $Diff \leq 0$  时,定长汉字编码算法的空间性能较优;为了使  $Diff$  的值趋向于最大值,而输入数据字符数  $D$  已知,则必须根据汉字使用频率表,实验分析分块块数  $n$  与每块大小  $N_i$  的取值。

## 3 二维码变长编码算法生成树的基本结构

### 3.1 常用汉字使用频率表

对于任意的汉字  $A$ ,其使用频率  $f(A)$  是一个通过大量实验数据总结出来的统计值,受测试数据的影响。对于不同的使用环境,可以得出相应的汉字使用频率表,同时也可以在实际使用中不断优化频率表,获得与使用环境  $\lambda$  相关的使用频率表  $f_\lambda(A)$ 。在设计变长编码算法时,忽略了  $\lambda$  对于使用频率的影响,可通过大量的实验数据减少汉字使用环境对其使用频率的影响。二维码汉字编码算法在使用中需要提前确定汉字使用频率表,并将其固化在二维码应用程序中,以保证应

用程序的通用性。

通过测试 7705551 个汉字,得出的常用汉字的使用频率如图 3 所示。

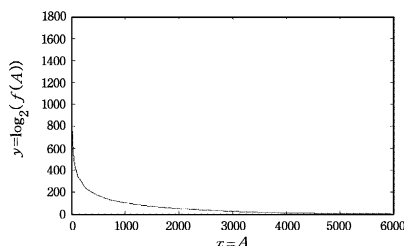


图 3 常用汉字使用频率图

图 3 中, X 轴为测试数据中出现的汉字数, Y 轴为测试汉字出现的频数。

由于当  $x$  很小时曲线的斜率很大,即汉字  $A$  使用的频率越高,通过对  $A$  进行变长编码,相对于定长编码算法,可以节省的空间就越大。由于二维码定长编码算法采用 13bit/字的方式,当二维码的存储空间大小确定时,减少二维码汉字编码的单字字长,增加二维码的汉字存储量,将进一步扩展二维码的应用前景。

### 3.2 基于汉字使用频率构建生成树

由于 GB2312 中定义的常用汉字共有 6763 个,则理论上常用汉字频率表分块的块数  $i$  的最大值为 6763 个,即每一个块内有且仅有一个汉字,可直接使用块号来唯一地表示一个汉字,可以将定长编码算法看作是分块数为 6763 的变长编码算法。由于汉字最终通过二进制流的形式进行编解码,可以用 6763 叉树  $T$  来表示所有的汉字。每一个叶子节点代表一个汉字,所有叶子节点相互之间都是兄弟关系,而其所在的位置代表着汉字的编码。当分块的块数  $i$  小于 6763 时,即需要将 6763 叉树  $T$  进行剪枝,用父节点代替两个孩子节点,形成新的分块后的树  $T_i$ 。按照  $T_i$  的分块进行汉字编码,通过实验分析,得出  $Diff$  的最优解。

由图 3 可知,当汉字  $A$  的使用频率越高,  $A$  可能出现的范围就越小。由于不同环境下的高频汉字可能不同,为保证较好的兼容性,高频字块的尺寸不宜太小。为了实验需要,令  $N_1 = 2^3$ ,即  $S_1 = 3$ ,当汉字  $A$  的使用频率越低,  $A$  可能出现的范围就越大。因此,令  $N_i \leq N_{i+1}$ ,由于任意  $A$  的使用频率为  $f(A)$ ,则  $A$  在测试中出现的概率  $p(A) = cf(A)$ 。  $c = 1/7705551$ ,任意汉字出现在第  $i$  块内的概率  $p(i) = \sum p(A) (K_A = i)$ ,式(7)可转化为:

$$Diff = (13 - \sum_{j=1}^n p(i) * S_i) D \quad (7)$$

根据实际编码需要,为保证测试结果具有较好的兼容性,不宜将分块长度设定得过小,令  $S_i \geq 6$ 。为保证解码器能通过块号唯一识别出对应数据块,块号长度必须定长。

当使用经典哈夫曼算法构建生成树时,分块的块长均相等,分块的块数为  $6763/2^{S_i}$ 。使用哈夫曼算法时,生成树构造较为复杂,且当  $S_i$  较小时,分块的块数较大,编码的效率受汉字使用频率的影响较大。当  $S_i = 10$  时,根据汉字使用频率表,构造生成树,如图 4 所示。生成树中所有的非叶子节点均没有右子树,即在编码过程中,编码的块号长为 6。当  $S_i = 11$  时,编码的块号长为 3。当  $S_i = 12$  时,编码的块号长为 1。则在编码过程中,实际单字编码长度均不小于 13,因此经典哈

夫曼算法不适用于二维码汉字编码的优化改进,且哈夫曼编码的分块的长度均相等,不利于分块的合并。

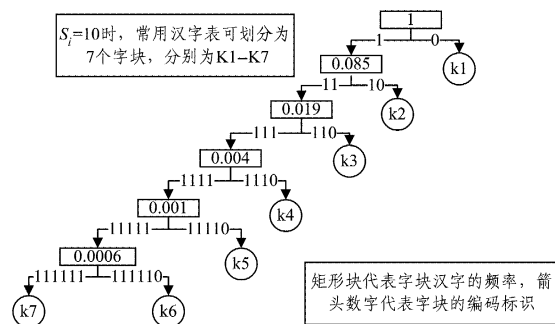


图 4  $S_i = 10$  时,哈夫曼算法生成的构造树结构图

当使用贪心算法时,首先将常用汉字按照使用频率分割为不同的子块,其中最后一块为使用频率最低的子块,统一按照 13bit/字的方式进行编码。对于使用频率相对较高的字块使用 6-12bit/字的编码字长进行编码。假定划分的块数为 7,每块的大小为 1024,其基本结构如图 5 所示。由于在分块时将最后一块固定为 13bit/字,保证了所有常用汉字在编码过程中均有唯一的代码与其对应。假设将图 5 中的  $K_4-K_7$  合并为  $K_4$ ,则块号的长度为 2,最小编码长度为 10,理论上满足对二维码汉字编码的优化和改进。

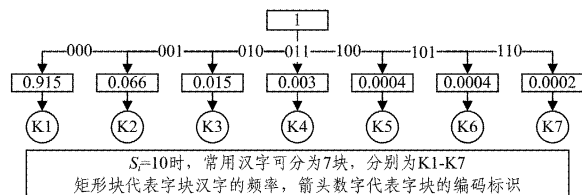


图 5  $S_i = 10$  时,贪心算法生成的构造树结构图

## 4 二维码变长编码算法的实验与分析

### 4.1 常用汉字使用频率表分块及测试

本文使用贪心算法。当  $S_i + 1 = S_i (n \geq 2)$  时,测试结果如图 6 所示。图 6 中所有曲线均有最小值点,即编码算法在对常用汉字进行分块时,若分块的长度均相同(最后一块除外),则编码的效率仍然可以优化。

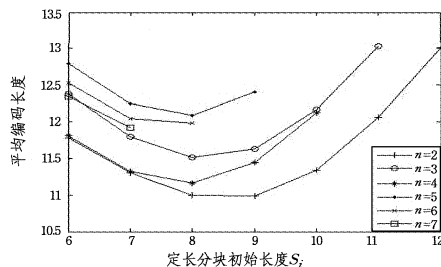


图 6 定长分块测试图

当  $S_i$  取定值时,由图 6 可知,汉字的平均编码长度随分块数  $n$  的增大出现震荡增加的趋势,且极小值点的取值  $n$  均为 2 的幂指数,同时,极小值点之间的间隔为 2 的幂指数且间隔越来越大。测试结果如图 7 所示。即当分块长度  $N_i$  等长,  $n = 2^1, 2^2, \dots$  时,变长编码算法的平均长度较小,即  $\sum_{j=1}^n p(i) * S_i$  的值较小,  $Diff$  的值较大。若取曲线的每一个极小值,则变长编码的算法的分块长度必须取  $2^{S_i}$ ,且  $S_{i+1} > S_i$ 。

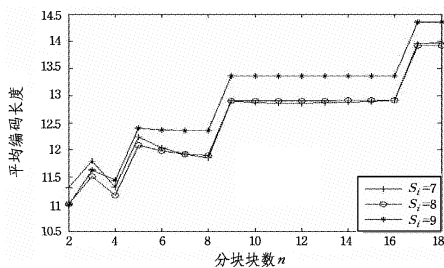
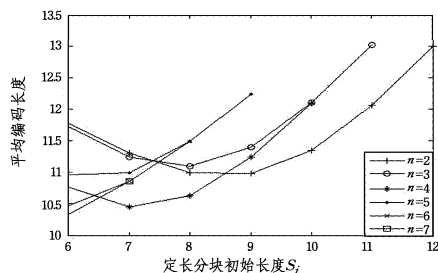


图7 平均编码长度与分块块数的关系图

当  $S_{i+1} = S_i + 1, n \geq 2$  时, 测试结果如图8所示。

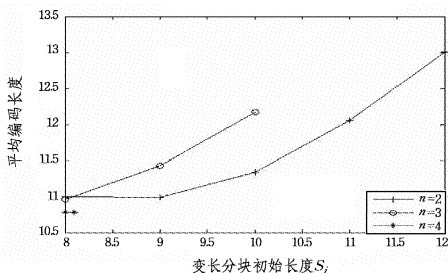
图8 变长分块测试图( $S_{i+1} = S_i + 1$ )

由图8可知, 当  $S_i \leq 7, n = 2$  时, 汉字的平均编码长度较大。即当分块的大小  $N_i \leq 2^7$  时, 汉字的平均编码长度失去了随分块数  $n$  的增大震荡上升的趋势。编码长度受分块大小的影响较大, 因此当  $S_i \leq 7$  时, 不宜作为变长编码算法的初始分块长度。

当  $S_i = 8$  时, 汉字平均编码长度依然随分块数  $n$  的增大震荡上升, 且极小值点的分块数  $n$  依然为 2 的幂指数。即可以通过增加变长编码算法的步长降低汉字编码的平均编码长度。

令  $S_{i+1} = (S_i + 1) + 1, n \geq 2, S_1 = 8$ , 由于常用汉字总数  $6763 < 2^{13}$ , 则  $S_i \leq 13$ , 且  $n \leq 4$ 。

实际测试结果如图9所示。由图9可知, 当  $S_{i+1} = S_i + 2$  时, 仅当  $n = 4, S_1 = 8$  时平均编码长度出现了最小值, 但不满足图7所示的震荡上升趋势, 不存在继续优化的空间。因此  $S_{i+1} - S_i \leq 2$ 。

图9 变长分块测试图( $S_{i+1} = S_i + 2$ )

结合图8和图9可知, 变长编码算法的初始大小为  $2^{S_i}$ , 当  $S_1 = 8, n = 4$  时, 变长编码算法的平均编码长度最小。对比图8和图9中  $S_1 = 8, n = 4$  的汉字平均编码长度可知, 图8中的平均编码长度较小, 因此, 对常用汉字进行分块编码时采用步长为 1 的变长编码方式, 即令  $S_{i+1} = S_i + 1$ 。

#### 4.2 二维码变长汉字编码算法的实验验证及性能分析

首先代入 4.1 节中的测试结果, 对常用汉字使用频率表  $T$  进行分块处理, 即将原常用汉字表分割为 4 个子块, 如图4所示。每一子块可看作一个新的二叉树, 且除  $T_4$  外,  $T_1$ ,

$T_2, T_3$  均为满二叉树, 长度分别为  $2^8, 2^9, 2^{10}$ 。其次对每一个子节点添加块号, 并对每一子树进行内部编码。最后代入二维码平台实现二维码变长汉字编码。

对于任意的汉字  $A$ , 在汉字使用频率表中的偏移量为  $X$ , 根据式(4),  $X_A = X - \sum_{i=1}^{K_A-1} 2^{S_i}$ ,  $S_1 = 8$ , 由于  $X_A \geq 0$ , 则  $K_A, X_A$  可解。假设  $X = 613$ , 则  $K_A = 2, X_A = 357$ , 则变长汉字编码算法中汉字  $A$  的块号为 01, 块内编码为 101100101, 则汉字  $A$  的编码为 01101100101。

由于标准的二维码汉字编码采用图1中描述的结构, 二维码对于汉字编码的单字编码长度是确定的, 在纠错等级最低时, 二维码汉字编码上限为 1817 个。当输入数据为 1817 个汉字时, 测试定长编码算法和变长编码算法的实际性能。对于相同的输入, 分别代入二维码定长编码算法和二维码变长编码算法, 测试结果如表1所列。

表1 二维码编码算法的编码结果( $D=1817$ )

| 算法类别 | 编码数据量<br>(汉字) | 状态   | 纠错    | 版本 | 掩码 |
|------|---------------|------|-------|----|----|
| 定长编码 | 1817          | 处理完成 | L(%7) | 40 | 4  |
| 变长编码 | 1817          | 处理完成 | L(%7) | 38 | 4  |

由表1可知, 对于相同的输入数据, 定长编码算法使用的版本号为 40, 而变长编码算法使用的版本号为 38。定长编码模式下, 当汉字长度一定时, 编码使用的位数唯一确定, 且编码过程中使用的版本号也唯一确定。则对于相同的输入, 变长编码算法使用的数据位数小于定长编码算法。根据图5所示结果,  $Diff = (13 - 10.6)D = 2.4D$ , 则二维码变长汉字编码算法可平均减少 18.4% 的冗余。因此变长编码算法在空间性能上优于定长编码算法。

当二维码汉字编码的输入大于 1817 个汉字时, 定长编码算法使用的位数大于二维码平台所能表示的最大长度, 软件报错。而变长编码算法使用的数据位数依然小于二维码平台所能表示的最大位数。表2为输入汉字总数为 1900 时, 定长编码算法和变长编码算法的测试结果。

表2 二维码编码算法的编码结果( $D=1900$ )

| 算法类别 | 编码数据量<br>(汉字) | 状态   | 纠错    | 版本 | 掩码 |
|------|---------------|------|-------|----|----|
| 定长编码 | 1900          | 错误   | —     | —  | —  |
| 变长编码 | 1900          | 处理完成 | L(%7) | 39 | 1  |

采用定长编码算法时, 中文汉字机内码转 13bit 二维码编码仅需要图1中描述的有限次数操作, 其算法复杂度为  $O(D)$ 。对于变长编码算法, 首先必须将汉字机内码转换为变长编码算法对应的编码, 其次对于每个汉字必须进行查找操作, 时间复杂度为  $O(6763/2)$ , 对于  $D$  个汉字输入, 变长编码算法的时间复杂度为  $O(6763D/2) = O(D)$ 。理论上变长编码算法的时间复杂度与定长编码算法相同, 但在实际运行中, 变长编码算法的查找时间不可忽略, 即定长编码算法在时间性能上略优于变长编码算法。

将变长编码算法和标准定长编码算法代入相应的编解码程序, 通过改变输入数据的大小, 实际测试两种算法的处理速率, 测试结果如表3所列。实验证明二维码汉字编码算法采用定长编码方式的时间性能优于变长编码方式。

表 3 二维码汉字编码算法性能对比表

| 字数 $D$ /个 | 标准定长编码算法耗时/s | 变长编码算法耗时/s |
|-----------|--------------|------------|
| 250       | 0.125        | 0.374      |
| 500       | 0.359        | 0.483      |
| 750       | 0.405        | 0.671      |
| 1000      | 0.562        | 0.702      |

### 4.3 二维码变长编码算法的优化和改进

(1)如 3.1 节中所述,汉字的使用频率  $f$  是一个统计值,受实际使用环境的影响。理论上,当测试数据量无限大时, $f$  将唯一确定,可以满足不同使用环境下的编码需求。文中忽略了使用环境  $\lambda$  对  $f$  的影响,可以在某一使用环境下,来测试生成  $f_{\lambda}$ ,并将其用于后续的编码中,从而提升算法的空间性能。

(2)由于变长编码算法的查找时间不可忽略,可以定义一个包含汉字机内码和汉字变长编码的编码对照表,并将所有的汉字按机内码和变长编码分别进行排序,使用折半查找算法查找相应的汉字编码,减少汉字机内码与汉字变长编码之间相互转换的时间,从而提升算法的时间性能。

**结束语** 二维码变长汉字编码算法充分考虑了汉字使用频率对于二维码汉字编码长度的影响,能够平均减少二维码汉字编码 18.4% 的冗余,增加了二维码汉字的编码容量,扩展了二维码的应用范围。本文采用了 GB2312 汉字编码标准,对于 GBK,GB18030,CJK 等编码标准,变长二维码汉字编码算法仍然可以满足实际的编码需求。变长二维码汉字编码的基本思路可拓展到其他字符类型的编码中,以提升二维码在实际应用中的性能。

## 参 考 文 献

- [1] 胡俊翹,李德群,叶苏丹.一种高信息密度的二维码—矩阵码的技术及应用[J].计算机工程,1995(4):46-48.
- [2] 国家质量技术监督局.快速响应矩阵码: B/T 18284-2000[S]. 2000.
- [3] SUBPRATATSAVEE P, KUACHAROEN P. An implementa-

tion of a high capacity 2D barcode[C]// International Conference on Advances in Information Technology. Springer Berlin Heidelberg, 2012: 159-169.

- [4] TAVEERAD N, VONGPRADHIP S. Development of Color QR Code for Increasing Capacity[C]// 2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). IEEE, 2015: 645-648.
- [5] PILLAI P N, NARESH K. Improving the Capacity of QR Code by Using Color Technique[J]. International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering, 2014, 3(7): 2320-3765.
- [6] MELGAR M E V, SANTANDER L M. Channel capacity analysis of 2D barcodes: QR Code and CQR Code-5[C]// 2016 IEEE Colombian Conference on Communications and Computing (COLCOM). IEEE, 2016: 1-5.
- [7] UMARIA M M, JETHAVA G B. Enhancing the Data Storage Capacity in QR Code Using Compression Algorithm and Achieving Security and Further Data Storage Capacity Improvement Using Multiplexing [C] // 2015 International Conference on Computational Intelligence and Communication Networks (CICN). IEEE, 2015: 1094-1096.
- [8] UMARIA M M, JETHAVA G B. A Novel Approach for Enhancing Data Storage Capacity in Quick Response Code Using Multiplexing and Data Compression Technique[C]// 2015 International Conference on Computational Intelligence and Communication Networks (CICN). IEEE, 2015: 1091-1093.
- [9] 陈元枝,邓艳,史绍亮,等.基于 Gzip 压缩算法的彩色 QR 码生成与识别方法[J].电子技术应用,2015,41(12):116-119.
- [10] CHIANG J S, HSIA C H, LI H T. High density QR code with multi-view scheme[J]. Electronics Letters, 2013, 49(22): 1381-1383.
- [11] 夏春蕾,戴曙光,张仁杰.汉字信息的二维条码拆分迭代编码算法[J].计算机应用,2007,27(9):2337-2339.
- [12] 邹敏,张瑞林,吴桐树,等.编码技术增加 QR 码的信息容量[J].工业控制计算机,2015(9):111-112.

(上接第 537 页)

视化。基于 SQL Server 数据库的数据存储功能,对风电场的有功出力实现了数据存储,并实现了对风电场有功出力数据的管理,成功地实现了数据的远程操作和历史数据的浏览及查询。同时,基于此软件平台,提出了一种系统构架方法。因此,本文工作对于完善和丰富“风电场有功出力仿真软件”的知识理论体系以及对未来风电系统的研究具有一定的积极意义。

## 参 考 文 献

- [1] 袁铁江,陈洁,刘沛汉,等.储能系统改善大规模风电场出力波动的策略[J].电力系统保护与控制,2014,42(4):47-53.
- [2] 王海超,鲁宗相,周双喜.风电场发电容量可信度研究[J].中国电机工程学报,2005,25(10):103-106.
- [3] 孙建峰.风电场建模和仿真研究[D].北京:清华大学,2004.
- [4] 姜文,严正,杨建林.计及风电场的发输电可靠性评估[J].电力系统保护与控制,2010,38(22):126-130.
- [5] 徐丽华,杨青斌.基于 VC++ 的风储一体化电站中储能容量优化配比[J].电力电容器与无功补偿,2015,36(4):80-86.
- [6] 黄海煜,于文娟.考虑风电出力概率分布的电力系统可靠性评估

[J]. 电网技术,2013,37(9):2585-2591.

- [7] 李辉,王荷生,史旭阳,等.基于遗传算法的风电场等值模型的研究[J].电力系统保护与控制,2011,39(11):1-8.
- [8] 袁铁江,晁勤,李义岩,等.大规模风电并网电力系统经济调度中风电场出力的短期预测模型[J].中国电机工程学报,2010,30(13):23-27.
- [9] 蒋程,刘先正.风电机组出力的概率性评估[J].现代电力,2013,30(1):51-54.
- [10] 张建华,王昕伟,蒋程,等.基于蒙特卡罗方法的风电场有功出力的概率性评估[J].电力系统保护与控制,2014,42(3):82-87.
- [11] 徐玉琴,张林浩,王娜.计及尾流效应的双馈机组风电场等值建模研究[J].电力系统保护与控制,2014,42(1):70-76.
- [12] 严干贵,李鸿博,穆钢,等.基于等效风速的风电场等值建模[J].东北电力大学学报,2011,31(3):13-19.
- [13] 夏欣. VC 与 MATLAB 混合编程实现方法及具体实例研究[J].青岛理工大学学报,2011,33(3):83-88.
- [14] 闫旭.浅谈 SQL Server 数据库的特点和基本功能[J].价值工程,2012(22):229-231.
- [15] JENSEN N O. A note on wind generator interaction; Risø-M-2411[R]. Risø National Laboratory, 1983.