

改进的二进制循环码盲识别方法

朱联祥, 李 荔*

(重庆邮电大学 信号与信息处理重庆市重点实验室, 重庆 400065)

(*通信作者电子邮箱 labidax123@163.com)

摘 要: 目前已有的循环码盲识别方法在低码率编码条件下效果较好,但在高误码率及高码率条件下不能高效识别,或者只针对循环码中某一子类。为有效解决高误码率以及高码率编码下的循环码盲识别问题,提出一种基于矩阵变换和码重分布的方法,首先对接收序列按估计码长构造矩阵,并对矩阵进行初等变换;然后利用改进的码重分布距离公式对循环码进行盲识别。仿真结果表明该方法在高误码率以及高码率编码时可实现高效的循环码盲识别。

关键词: 矩阵变换;循环码;盲识别;码重分布;码长识别

中图分类号: TP911 **文献标志码:** A

Improved blind recognition method for binary cyclic code

ZHU Lianxiang, LI Li*

(Chongqing Key Laboratory of Signal and Information Processing, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: The existing blind recognition methods of cyclic code have poor effects in the high or only Bit Error Ratio (BER) better in low code rate conditions, or the method is only for a subclass of the cyclic codes. In order to solve the blind identification for cyclic code with high BER or high code rate effectively, a method based on code weight distribution and matrix transformation was proposed. First of all the article structured the receiving sequence to matrix according to the estimated code length, and then realized the blind recognition using the improved weight distribution distance formula. The simulation results show that the method can realize the blind recognition for cyclic code with high BER and high code rate, and the results are better.

Key words: matrix transform; cyclic code; blind recognition; code weight distribution; code length identification

0 引言

循环码是线性分组码中重要的一类码,被广泛应用于各种无线通信中,其中由 Bose、Chaudhuri 和 Hocquenghem 共同提出的循环码(简称为 BCH 码)应用最为广泛。但目前有关循环码盲识别的文献不多,文献[1]提出了综合矩阵分析法得到码长,但计算量较大,且需要知道码字起点;文献[2]利用码字的重量分布距离得到码长,但只适用于低码率编码的盲识别且需要知道码字起点;文献[3]基于欧几里德方法算出二进制 BCH 码生成多项式,但先验信息码字起点已知;文献[4]是基于码根信息差熵和码根统计的 BCH 码识别方法,需要统计出各个多项式的生成码根,且只能识别 BCH 码;文献[5]通过汉明距离得到码长,但需要已知码字起点,且对码字间两两求距离,计算量较大;文献[6]虽然能同时识别出码字起点和码长,但所需数据量极大;文献[7]基于码重信息熵实现线性分组码的盲识别,但也只适用于低码率分组码。因此采用何种方法能在较少数据量下实现高误码率和高码率编码下的盲识别^[8]尤为重要。基于此问题,本文在已有循环码盲识别基础上,用一种基于矩阵变换和码重分布的方法来实现循环码的全盲识别,仿真结果表明该方法可以实现无误码和有误码时的循环码的全盲识别。

1 识别模型

定义 1 一个 n 重子空间,若对任何一个 $V = (a_{n-1},$

$a_{n-2}, \dots, a_0) \in V_{n,k}$, 恒有 $v_1 = (a_{n-2}, a_{n-3}, \dots, a_0, a_{n-1}) \in V$, 则称 $V_{n,k}$ 为循环子空间或循环码。

(n, k) 循环码中 n 是码字长度, k 是信息位, $r = n - k$ 是校验位,循环码的生成多项式表示为 $g(x) = g_{n-k}x^{n-k} + g_{n-k-1}x^{n-k-1} + \dots + g_1x + g_0$ 。

生成的码多项式为

$$c(x) = m(x)g(x) \quad (1)$$

其中 $c(x) = c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \dots + c_1x + c_0$, 则循环码的全盲识别就是在仅已知 $c(x)$ 的情况下,识别出循环码接收序列的码字起点、码长以及生成多项式 $g(x)$ 。

2 识别方法

2.1 传统码重分布识别法

传统的码重分布识别法^[9]中提到对于 (n, k) 分组码,如果估计码长不等于真实码长,则码字的信息位与校验位不存在约束关系,可认为 0 或者 1 是等概率出现的,则假设此时数据的码重分布是等概率的,可认为不同码重的码字出现的概率为 $1/(n+1)$,令 P_i 为实际重量为 i 的码组所出现的概率,则定义码重分布距离公式为

$$D = \frac{n}{n+1} \sum_{j=0}^n \left| P_j - \frac{1}{n+1} \right|^2 \quad (2)$$

则当码重分布距离 D 最大时的估计码长 n 为真实码长。

该方法在高误码率时可有效识别低码率的分组码,但是当码率较大时,识别概率几乎为 0。本文主要针对线性分组

收稿日期: 2013-04-15; 修回日期: 2013-06-05。

作者简介: 朱联祥(1971-),男,陕西户县人,教授,博士,主要研究方向:无线定位通信信号处理、数字信号处理、信息论与编码; 李荔(1988-),女,贵州织金人,硕士研究生,主要研究方向:信道编码的盲识别。

码中的一类——循环码进行识别,为解决循环码高码率时的识别问题,结合矩阵变换,并改进传统码重分布距离识别算法,实现高码率下的循环码盲识别,仿真效果表明本文方法在高码率时能有效实现识别。

2.2 循环码码长识别

本文在利用码重分布识别法前先对接收序列构造矩阵,然后对矩阵进行变换,根据码重分布特性,改进码重分布距离算法,避免了原有码重分布距离识别法在高码率时识别效果不佳的问题,在高码率编码及高误码率下提升了循环码的识别能力。

假设接收序列长度为 N , 则循环码码长识别步骤如下:

1) 设估计码长为 n , 将接受序列按估计码长放入矩阵 A , 定义矩阵 A 的列数为 n , 行数为 m , 且行数 $m = \lfloor N/n \rfloor$, 其中 $m > 4n$;

2) 从列数 i 开始(初始化 $i = 1$), 若元素 $A(i, i)$ 为 0, 则将 i 列与 i 行右边第一个出现元素为 1 的列进行交换;

3) 若 i 列右边不存在元素为 1 的列, 则从上到下将 i 列中第一个出现元素 1 的行与 i 行交换;

4) 将 i 行 i 列右边所有元素不为 0 的列与 i 列进行伽罗华域的模 2 加;

5) 列数 i 加 1, 返回步骤 2) 进行, 直到总列数 n 结束, 得到化简后的新矩阵 B ;

6) 记矩阵 B 各列出现的元素为“1”的次数为 sum , 则各列出现 1 的概率

$$p_i = sum_i / row; i = 1, 2, \dots, n \quad (3)$$

7) 令 $D1$ 为改进后的码重距离, 则改进的码重分布距离公式如下:

$$D1 = \frac{1}{n} \sum_{i=1}^n (p_i - 0.5)^2 \quad (4)$$

8) 估计码长 n 加 1, 返回步骤 1) 进行, 选择 $D1$ 出现最大值时的估计码长为循环码真实码长。

循环码是线性分组码的重要一种子类, 校验位与信息位具有线性相关性。在无误码时, 如果估计码长等于真实码长, 则接收码字的校验位构成的列与信息位构成的列是线性相关的, 通过如上步骤的变换, 则某些列必为全 0 列, 且全 0 列的数目就是校验位长度 r , 从而可知信息编码码率 k/n ; 但如果估计码长不等于真实码长, 码字的信息位和校验位没有相关性, 则列数为 n 的矩阵不会出现全 0 列。同样在有误码时, 生成的码字中有误码, 即使估计码长等于真实码长, 也会导致部分码字的信息位和校验位没有相关性, 进而使矩阵出现全 0 列的概率降低, 但是通过上面矩阵的变换, 可以使列中的“1”元素出现的概率降低。

随机序列中 0、1 出现的次数相近, 所以设每列 0、1 出现的概率为 0.5, 当估计码长不等于真实码长时, 则构成矩阵的每一列中 0、1 的分布类似随机序列, 反之, 估计码长等于真实码长时, 通过上述的矩阵变换, 会使某些列中的 0、1 元素分布不均匀, 且 1 出现的概率远小于 0.5^[10-11], 所以得到式(4)的码重分布距离公式。

设接收序列构造的矩阵列数为 n , 行数为 m , 则文献[1]中的传统综合矩阵分析需要进行 $n(n-1)/2$ 次列的模 2 加, 同时还要进行 $m(m-1)/2$ 次行的模 2 加, 而通常构造的矩阵中行数较大, 所以效率较低, 而改进后的算法只需要进行 $n(n-1)/2$ 次列的模 2 加, 只有当列变换不满足码长识别条件时才进行行变换, 所以相比传统的综合矩阵分析法, 算法

效率提高。

2.3 循环码码字起点识别

在分析循环码码长时提到估计码长等于真实码长时, 会出现全 0 列, 导致 $D1$ 出现最大; 同理, 在估计的码字同步起点等于真实的码字同步起点时, 通过矩阵变换会使某些列中“1”元素最少, 也即使 $D1$ 最大。假设码字从 b 位开始进行识别, 具体码字同步起点识别步骤如下:

1) 将接收序列从 b 处按码长 n 放入矩阵 A , 矩阵 A 的行数 $m = \lfloor N/n \rfloor$, 列数为码长 n ;

2) 对矩阵 A 进行码长识别中的步骤 2) ~ 7);

3) 使 b 加 1, 返回步骤 1) 进行, 直到遍历起点数等于码长 n , 取码重距离 $D1$ 出现最大时的 b 为真实码字起点。

2.4 循环码生成多项式识别

在识别出码字同步及真实码长后, 生成多项式识别就较为简单, 由循环码以及有限域知识可知, 循环码的码多项式 $c(x)$ 均是生成多项式 $g(x)$ 的倍式, 因此在知道循环码真实码长和起点后, 对接收序列按码长分组, 则无误码的码字之间必然有最大公因式^[12]。

无误码时, 假设得到 X 组码字, 从第一个码字开始与下一个码字运用辗转相除法求两个多项式的最大公因式, 再将该最大公因式与下一个码字求最大公因式, 直至与最后一个码字求得最大公因式, 最后的公因式即为循环码的生成多项式 $g(x)$, 共需要进行 $X-1$ 次求多项式的最大公因式运算。

有误码时, 按照如下步骤进行生成多项式识别:

1) 将码长识别中构造的矩阵 A 循环左移一位, 形成新矩阵 A_1 ;

2) 利用欧几里德算法分别求矩阵 A 和 A_1 中同一行的最大公因式(对于二进制的循环码来说, 欧几里德算法可以用辗转相除法实现), 通过该处理得到一个最大公因式矩阵 H ;

3) 记 H 中所有能除尽 $x^n - 1$ 的多项式集合为 $g_i(x)$;

4) 将 $g_i(x)$ 中多项式分别除矩阵 A 中所有码多项式, 统计所有能除尽的码字概率, 概率最大的即为生成多项式 $g(x)$ 。

3 仿真分析

本文以下进行的仿真, 都是利用 Matlab 软件产生比特序列, 仿真模拟信道是二进制对称信道, 即信道导致的比特错误不会影响其他比特的传输。

3.1 循环码码长识别验证

对(31, 11)和(31, 26)二进制循环码分别产生 1000 组的序列, 在误码率 $p_e = 4 \times 10^{-3}$ 时分别对传统码重分布距离分析法和本文中方法进行码长识别仿真, 仿真结果如图 1 和图 2。

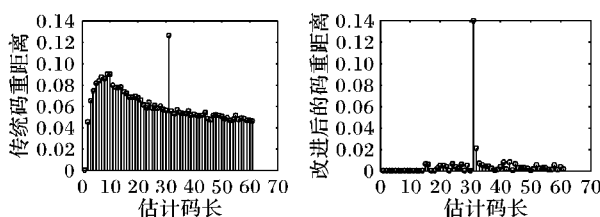


图 1 (31, 11) 循环码在 $p_e = 4 \times 10^{-3}$ 时两种方法码长识别图

从图 1(a) 和(b)中可以看到在 $p_e = 4 \times 10^{-3}$ 时, 传统码重识别方法和本文的识别方法均可以在估计码长为 31 时得到最大值, 即估计出的码长 31 为(31, 11)循环码真实码长。

但对于(31,26)的循环码,在 $p_e = 4 \times 10^{-3}$ 时,由图2(a)可以看出,传统码重识别法已不能正确识别出循环码码长,但根据图2(b)所示,本文方法仍能在估计码长为31时得到最大值,相比传统码重识别方法,本文对接收序列先进行矩阵变换的处理,能有效识别高码率的循环码码长。

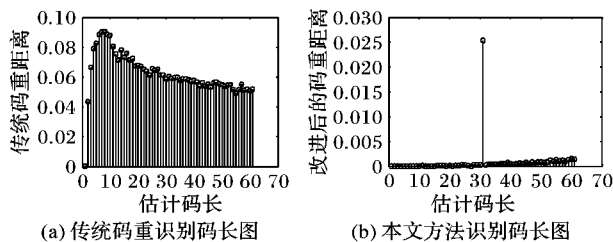


图2 (31,26)循环码在 $p_e = 4 \times 10^{-3}$ 时两种方法码长识别图

对(31,26)、(31,11)、(23,12)和(15,7)循环码在不同的误码率下进行码长识别概率仿真,得到如图3的识别概率图。

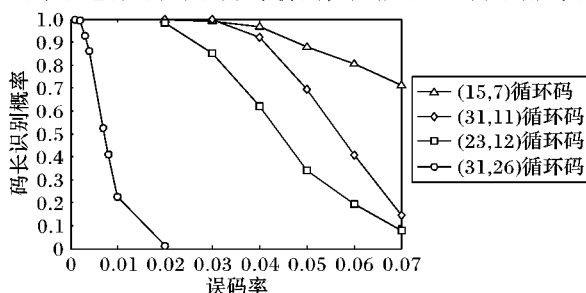


图3 不同误码率下循环码码长识别概率

图3中每个点均经过600轮的仿真,通过仿真图能看出不同码长的循环码随着误码率的增大识别概率降低,对于低码率的(31,11)、(23,12)、(15,7)循环码,在误码率为 10^{-2} 以内均能100%识别出码长,其中(15,7)循环码在误码率为 7×10^{-2} 时能达到70%的识别概率,(31,11)、(23,12)分别在 4×10^{-2} 和 3×10^{-2} 时达到90%和80%的识别概率。对于高码率的(31,26)循环码,在误码率为 5×10^{-3} 时,识别概率能达到80%以上。

通过该仿真实验可知在码长不是100%识别概率前提下,在相同误码率下,随着码长的增加,识别能力逐渐降低;对于同一码长的循环码,在相同误码率下,随着信息位的增加,识别能力降低。

3.2 循环码字起点识别验证

码字起点识别验证是建立在识别出码长的基础上,分别对低码率的(15,7)循环码和高码率的(31,26)循环码在误码率为 $p_e = 4 \times 10^{-3}$ 时进行仿真,同样产生1000组序列,并假设Matlab产生的循环码序列不是从一个完整码字的起点开始,码字起点识别的目的就是找到第一个完整码字的起点,仿真结果如图4。

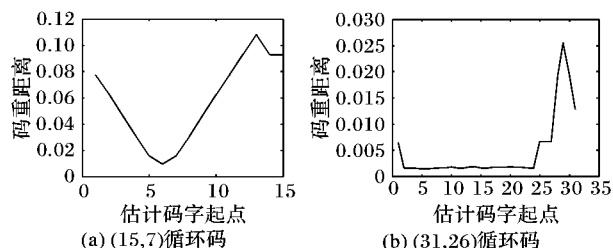


图4 不同码长循环码在 $p_e = 4 \times 10^{-3}$ 码字起点识别图

如图4(a)所示,(15,7)循环码在码重距离D1最大时,可以识别出它的码字起点,即在估计码字起点13时,循环码才开始出现一个完整码字;同样对于高码率编码的(31,26)循

环码,从图4(b)中看出在估计码字起点为28时为循环码完整码字起点。

对(31,11)、(31,26)和(15,7)二进制循环码在不同误码率下进行码字起点识别,得到如图5的识别概率图。

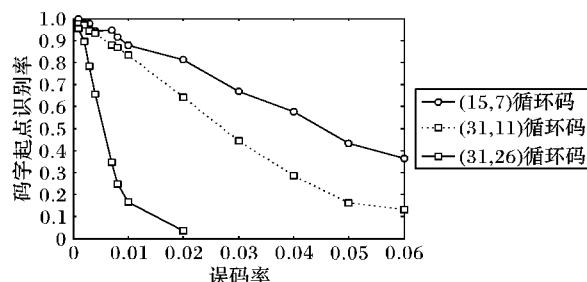


图5 不同误码率下循环码码字起点识别概率

从图5中可知对低码率的(15,7)和(31,11)的循环码,随着误码率的增大,码字起点识别概率下降,但在误码率0.01内,码字起点识别概率能达到80%以上。对于高码率的(31,26)循环码,识别能力有所降低,在误码率为 5×10^{-3} 以内,识别概率能达到50%以上。

通过仿真实验可知,在码字起点不是100%识别概率前提下,在相同误码率下,随着码长的增大,码字起点识别能力降低;在相同误码率下,对同一码长的循环码,随着信息位的增加,码字起点识别能力降低。

3.3 生成多项式识别验证

以(31,26)二进制循环码为例,由Matlab产生1000组序列,误码率 $p_e = 0.005$,则通过2.4节所述的处理后,得到的所有的 $g_1(x)$ 的八进制数表示以及能除尽码组总数如表1。

表1 (31,26)二进制循环码生成多项式估计

$g_1(x)$ 对应的八进制表示	除尽码组的概率
45	0.848
140	0.491
157	0.419
336	0.419
450	0.848
2570	0.029
11566	0.014

从表1中可以看出在对应的八进制数为45和450时的概率最大,八进制数为45和450时候分别对应的二进制数为100101,100101000,又二进制循环码的生成多项式 $g(x)$ 常数项必须为1,则将100101000循环右移3位后得到100101,与八进制数45时候相同,所以(31,26)的生成多项式系数为100101,也即生成多项式 $g(x) = x^5 + x^2 + 1$ 。

4 结语

本文在运用码重分布识别前先对生成序列作矩阵变换的处理,再利用循环码码重分布和随机序列码重分布理论,改进了传统码重识别方法,解决了传统码重分析法在高误码率及高码率下不能有效识别的问题。通过仿真实验表明,该方法在有误码时能有效识别循环码,对于本文中仿真的低码率的(15,7)、(23,12)、(31,11)线性循环码在误码率0.03内码长识别概率能达到70%以上,高码率的(31,26)在误码率为0.005以内码长识别概率也能达到90%左右。此外,本文基于矩阵变换和改进的码重分布的识别码字起点和码长的方法可推广到线性分组码中。

(下转第2768页)

实验三 当信噪比为 10 dB 时,测试正确识别率随观察符号数变化的情况。符号个数从 200 增加到 1800,变化步长为 200,每一种信号在不同符号数下进行 200 次蒙特卡罗仿真,其识别结果如图 4。

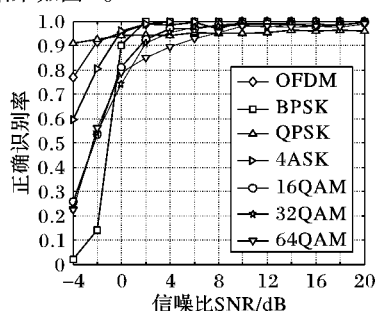


图3 正确识别率与信噪比变化的关系

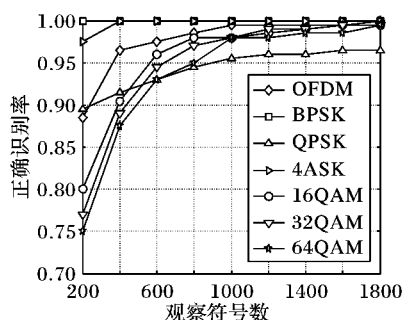


图4 正确识别率与符号个数变化的关系

由图4可以看出,正确识别率随着观察符号个数的增多而提高,当观察符号数多于1000个时,七种信号的正确识别率均达到95%以上,识别效果较好。由此可以得到一个结论:在信噪比不变的情况下,用于观察的符号个数越多,识别信号的正确率就越高,当观察符号个数达到一定数量后,识别率趋于稳定。其原因是由于只有当接收到的数据量足够大时,才能更好地体现信号的调制特性,并且高斯噪声的高阶累积量才趋于零,因此正确识别率会随着数据量的增加而提高。但是算法的计算复杂度会随着数据量的增加而提高,所以在保证识别率的前提下,为了减少运算量,选择长度合适的接收数据进行处理也很重要。

4 结语

本文在瑞利信道下提出了一种基于累积量组合的识别方法,仅用两个特征参数和决策树分类器就能够实现 OFDM 信

号与常见单载波信号的类型识别及单载波信号之间的类内识别。通过理论推导与分析,所提方法对瑞利衰落与噪声干扰不敏感。计算机仿真结果表明,当信噪比大于 4 dB 时,整体识别率都在 90% 以上,说明本文方法的可行性及有效性。文献[8]用四个特征参数仅实现了五种信号的识别,且没有考虑 OFDM 信号;文献[9]只是讨论了三种 QAM 信号的分类,识别类型有限,因此,本文方法与同类算法相比,具有一定的优越性。但当信噪比小于 0 dB 时,识别性能下降,如何能在低信噪比的情况下有效实现调制方式的识别是需要进一步研究的问题。

参考文献:

- [1] DONRE O A, ABDI A, BAR-NESS Y, *et al.* Survey of automatic modulation classification techniques: classical approaches and new trends[J]. *Institution of Engineering and Technology*, 2007, 1(2): 137-156.
- [2] 党月芳, 徐启建, 张杰. 数字调制识别的关键技术与发展趋势[J]. *传感器与微系统*, 2012, 31(10): 1-5.
- [3] SWAMI A, SADLER B M. Hierarchical digital modulation classification using cumulants[J]. *IEEE Transactions on Communication*, 2000, 48(3): 416-429.
- [4] 陈卫东, 杨绍全. 利用累量不变量对 MPSK 信号分类[J]. *西安电子科技大学学报*, 2002, 29(2): 229-232.
- [5] 李世平, 陈方超. 基于小波和高阶累积量的数字调制识别算法[J]. *计算机应用*, 2011, 31(11): 2926-2928.
- [6] LI P H, ZHANG H X, WANG X Y, *et al.* Modulation recognition of communication signals based on high order cumulants and support vector machine[J]. *The Journal of China Universities of Posts and Telecommunications*, 2012, 19(1): 61-65.
- [7] 李彦松, 罗明, 李霞. 基于高阶累积量的 OFDM 信号调制识别技术[J]. *电子信息对抗技术*, 2012, 27(4): 1-4.
- [8] 李鹏, 汪美平, 王赞基. 瑞利平坦衰落信道中的调制识别算法[J]. *电路与系统学报*, 2009, 14(2): 107-110.
- [9] 李艳玲, 李兵兵, 刘明骞. 瑞利衰落信道下 MQAM 信号的盲识别方法[J]. *华中科技大学学报: 自然科学版*, 2012, 40(4): 76-79.
- [10] PROAKIS J G. *Digital communication*[M]. 4th ed. New York: McGraw-Hill, 2001.
- [11] 杨琳. 数字通信信号调制方式自动识别技术研究[D]. 合肥: 中国科学技术大学, 2008.
- [12] 张贤达. *现代信号处理*[M]. 2 版. 北京: 清华大学出版社, 2002.

(上接第 2764 页)

参考文献:

- [1] 张永光, 楼才义, 王挺. 一种线性分组码编码参数的盲识别方法: 中国, 2010101311103. 3[P]. 2011-09-28.
- [2] 鲁俊军, 李艳斌. 低码率二进制线性分组码的盲识别[J]. *无线电工程*, 2009, 39(1): 19-22.
- [3] 王兰勋, 李丹芳, 汪洋. 二进制本原 BCH 码的参数盲识别[J]. *河北大学学报: 自然科学版*, 2012, 32(4): 416-420.
- [4] 杨晓静, 闻年成. 基于码根信息差熵和码根统计的 BCH 码识别方法[J]. *探测与控制学报*, 2010, 32(3): 69-73.
- [5] 闫郁翰, 汤建龙. 基于汉明距离的二进制线性分组码盲识别方法[J]. *通信对抗*, 2011(4): 20-23.
- [6] 丁宇. 分组码的矩阵识别分析[J]. *计算机安全*, 2012(7): 39-42.
- [7] 陈金杰, 杨俊安. 基于码重信息熵低码率线性分组码的盲识别[J]. *电路与系统学报*, 2012, 17(1): 41-46.
- [8] 宋镜业. 信道编码识别技术研究[D]. 西安: 西安电子科技大学, 2009.
- [9] 张永光, 楼才义. 信道编码及其识别分析[M]. 北京: 电子工业出版社, 2010.
- [10] CLUZEAU M, FINIASZ M. Recovering a code's length and synchronization from a noisy intercepted bitstream[C]// *Proceedings of the 2009 IEEE International Symposium on Information Theory*. Piscataway: IEEE, 2009: 2737-2741.
- [11] 陈金杰, 杨俊安. 基于比特频率检测低码率线性分组码的盲识别[J]. *电子测量与仪器学报*, 2011, 25(7): 642-647.
- [12] 杨晓静, 闻年成. 基于秩函数和 Euclidean 算法的循环码盲识别[J]. *电路与系统学报*, 2012, 17(5): 120-123.