

高误码率下归零 Turbo 码帧结构识别

钟兆根¹, 吴昭军², 刘 杰², 张立民²

(1. 海军航空大学电子基础教研室, 山东烟台 264001; 2. 海军航空大学信息融合研究所, 山东烟台 264001)

摘 要: 现有的高斯消元法在 Turbo 码帧结构识别中, 存在着容错性差且计算量大的缺点, 直接导致后续分量编码器以及交织器识别算法无法应用. 为此, 提出了一种基于矩阵模型列相关性的识别方法. 首先在待帧长下构建分析矩阵, 并对其前后相邻两列进行差分处理, 然后建立关于差分后矩阵列重的统计量, 根据码长及其整数倍处必有统计值能通过判决这一特点确定帧长. 最后, 通过分析正确码长下通过判决的统计值对应的相对位置关系, 归纳帧起点的同步方法. 仿真结果表明, 本文算法在误码率高达 0.1 条件下, Turbo 帧结构参数识别能够达到 100% 的识别率, 与现有方法相比, 本文算法计算复杂度更低且识别性能更好, 可应用于认知无线电等非协作通信领域.

关键词: 纠错编码; 归零 Turbo 码; 帧结构; 差分处理; 认知无线电

中图分类号: TN911.7 **文献标识码:** A **文章编号:** 0372-2112 (2019)10-2089-09

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2019.10.010

Blind Identification of Frame Structure of Turbo Codes on Trellis Termination at High Bit Error

ZHONG Zhao-gen¹, WU Zhao-jun², LIU Jie², ZHANG Li-min²

(1. Naval Aviation University Department of Electronic Basis, Yantai, Shandong 264001, China;

2. Naval Aviation University Research Institute of Information Fusion, Yantai, Shandong 264001, China)

Abstract: The Gauss elimination method to recognize Turbo code frame structure has the shortcomings of poor fault tolerance and high complex computation, which directly leads to the inapplicability of subsequent algorithms to recognize the component coder and interleaver. For this reason, a novel method based on correlation between columns of the matrix is proposed. Firstly, analysis matrix is constructed under the undetermined frame length, with the neighboring columns being differentiated. Then statistics about column weight is established, and frame length is confirmed by the statistical values, which can pass the decision threshold when the column is frame length and its multiple. Finally, frame starting point is synchronized according to the relative position of statistical values which pass the threshold. The simulation results show that the correct recognition rate of parameters of Turbo frame structure can reach 100%, when the bit error rate is as high as 0.1. Compared with the existing methods, the proposed algorithm has lower computational complexity and better recognition performance, and is very suitable for non-cooperative communication fields such as cognitive radio.

Key words: error-correcting codes; Turbo codes on trellis termination; the structure of frame; differential processing; cognitive radio

1 引言

随着现代通信系统的发展, 纠错编码技术被广泛应用于无线通信、卫星通信和深空通信等领域, 相应的, 纠错编码识别技术也成为认知无线电、自适应调制编码和信息截获等技术中的研究热点之一. Turbo 码是一类重要的纠错编码类型, 具有接近香农限的性能, 目前已被 CCSD、3GPP 等标准采用. 因此, 对 Turbo 码的识别技术进行研究具有非常重要的意义.

目前对 Turbo 码的识别研究主要包括递归系统卷积码(RSC, Recursive Systematic Convolutional code)和伪随机交织器的识别. 对于分量编码器的识别, Barbier 等^[4]采用高斯消元法直接计算各系数, 只能对无误码情况进行分析. 文献[5]采用基于剩余定理改进的欧几里得法进行求解, 识别性能同样受误比特位置的限制. 文献[6]采用秩准则法识别码长、信息位长度等信息, 以减小直接对上述参数进行迭代检验的计算复杂度,

但降低了识别性能. Debessu 等^[7,8]采用 EM 算法对生成多项式系数的概率模型进行迭代计算,但收敛的概率较低. 文献[9]将 RSC 的编码结构等效为一种 ARMA 系统模型,然后采用 MCMC 方法进行迭代计算,过程较复杂. 文献[10]采用最小二乘代价函数衡量校验关系式成立的正确性,但代价函数的计算涉及双曲正切函数,增加了计算复杂度. 对于交织器的识别,文献[11]给出了一种基于最大似然原理的树状分支方法,虽然都具有一定的容错性,但需要预先知道交织后的信息序列. 为解决这一问题,文献[12]、[13]采用基于软判决信息的 BCJR 译码方法对交织器进行恢复,但计算复杂度随交织长度的增加急剧增长. 文献[14]、[15]从校验关系式出发,分别利用硬判决比特和软判决比特对交织器进行逐位恢复,该方法性能较好,但均以 RSC 编码器的参数已知作为前提.

以上研究内容的前提均是已知 Turbo 码三路编码数据,这就要求对 Turbo 码的帧长进行识别并完成帧起点的同步,而现有相关文献较少,仅有文献[16~18]采用秩准则法进行了简单分析,但识别性能较差. 为此,本文根据归零 Turbo 码的结构特点,提出了一种基于矩阵模型列相关性的识别方法,根据编码数据帧构建的矩阵中某些列可能相同这一相关性,将分析矩阵前后相邻两列进行差分处理,进而根据其中低重量列的分布特点,提取帧长度并完成帧起点同步. 仿真实验证明了

$$C_1 = \begin{bmatrix} \overbrace{u_0^1 \ c_{1,0}^1 \ c_{2,0}^1 \ \cdots \ u_{L-1}^1 \ c_{1,L-1}^1 \ c_{2,L-1}^1}^{3L \text{ 位编码信息}} & \overbrace{u_L^1 \ c_{1,L}^1 \ \cdots \ u_{L+2m-1}^1 \ c_{2,L+m-1}^1}^{4m \text{ 位归零信息}} \\ u_0^2 \ c_{1,0}^2 \ c_{2,0}^2 \ \cdots \ u_{L-1}^2 \ c_{1,L-1}^2 \ c_{2,L-1}^2 & u_L^2 \ c_{1,L}^2 \ \cdots \ u_{L+2m-1}^2 \ c_{2,L+m-1}^2 \\ \vdots & \vdots \\ u_0^N \ c_{1,0}^N \ c_{2,0}^N \ \cdots \ u_{L-1}^N \ c_{1,L-1}^N \ c_{2,L-1}^N & u_L^N \ c_{1,L}^N \ \cdots \ u_{L+2m-1}^N \ c_{2,L+m-1}^N \end{bmatrix} = [u_0 \ c_{1,0} \ c_{2,0} \ \cdots \ u_{L-1} \ c_{1,L-1} \ c_{2,L-1} \ u_L \ c_{1,L} \ \cdots \ u_{L+2m-1} \ c_{2,L+m-1}] \quad (1)$$

其中 u_i 、 $c_{1,i}$ 和 $c_{2,i}$ 分别于每帧编码序列中的三路数据, m 为寄存器个数. 可以看出在该矩阵中每一路数据都周期性分布在相同的列中.

下面对该矩阵各列的性质进行分析. 为了便于分析,图 1 给出了 $(2,1,m)$ RSC 码编码结构,其输出对应编码的前两路数据.

根据图中的结构可知, c_k 与寄存器状态 s_k , ($0 \leq k \leq m$) 如下关系

$$c_k = \sum_{i=0}^m \oplus g_{2,i} s_{k-i} \quad (2)$$

其中 $g_{2,i}$ 为前向多项式系数. 在一帧数据编码完成后,归零开始,由于分量编码器带有反馈结构,单纯的增加 0 序列不能对寄存器状态归零,所以需要按照寄存器实际状态来添加比特序列,设 k 时刻,寄存器状态为 $s_k = (s_{k,1}, s_{k,2}, \dots, s_{k,m})$,以图 1 为例,在 k 时刻添加的归零比特为:

所提方法的有效性.

2 问题模型

并行级联 Turbo 码常常用在实际的通信系统中,如 WCDMA 协议,在该协议中,交织器采用伪随机交织方式,分量编码器为递归系统卷积 (Recursive Systematic Convolutional Code, RSC) 码,并且采用交叉复用方式串并转换. 设交织长度为 L , RSC 码存储级数为 m ,考虑 $1/3$ 码率的情况,则在第 i 帧编码序列中,信息序列 $u^i = (u_0^i, u_1^i, \dots, u_{L-1}^i)$ 经交织后得到 u_π^i ,然后 u^i 和 u_π^i 分别输入两个 RSC 编码器得到校验序列 $c_1^i = (c_{1,0}^i, c_{1,1}^i, \dots, c_{1,L-1}^i)$ 和 $c_2^i = (c_{2,0}^i, c_{2,1}^i, \dots, c_{2,L-1}^i)$,并对 u^i 、 c_1^i 和 c_2^i 交叉复用,最后分别对两个 RSC 编码器中的移位寄存器进行归零处理.

对于实际接收到的 Turbo 码编码序列,为了便于后续处理往往需要分离出三路数据,这就要求对帧长进行准确估计,并完成帧同步. 本文根据归零 Turbo 码编码序列所建立的矩阵中特有的列相关性,对上述问题进行解决.

3 基于矩阵列相关归零 Turbo 码盲识别

3.1 矩阵列相关模型

设接收的编码序列总帧数为 N ,以一帧数据为一行构建如下矩阵

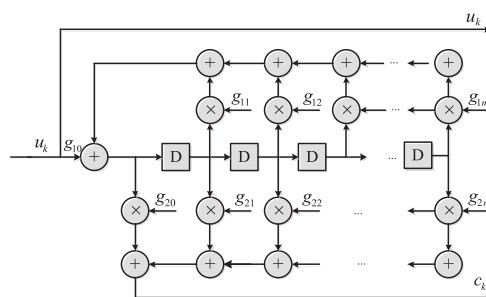


图1 $(2,1,m)$ RSC 码编码结构

$$u_k = \sum_{i=1}^m \oplus g_{1,i} \cdot s_{k,i} \quad (3)$$

故在 k 时刻,分量编码器输出的两路归零序列为 (u_k, c_k) . 其中

$$c_k = u_k \cdot g_{2,0} \oplus \sum_{i=1}^m s_{k,i} \cdot g_{2,i} \quad (4)$$

由此,当交织长度为 L ,码率为 n ,则实际的归零 Turbo 码码长 l 为:

$$l = nL + 2m(n - 1) \quad (5)$$

由于 $g_{1,0} = g_{2,0} = 1$,则第一路码元与第二路码元的差分值为:

$$u_k \oplus c_{1,k} = \left(\sum_{i=1}^m \oplus g_{2,i} u_{k-i} \right) \oplus \left(\sum_{i=1}^m \oplus g_{1,i} c_{1,k-i} \right) \quad (6)$$

式(6)给出了相邻两列 u_i 与 $c_{1,i}$ 对应位置元素模 2 加运算结果的递推关系. 由于前一帧归零码元的存在,使得每一帧都相当于 RSC 编码器重新开始编码. 在这种情况下,每一帧中对任意 $1 \leq i \leq m$,如果 $k - i \leq 0$,则 $u_{k-i} = c_{1,k-i} = 0$. 因此在每一帧的初始时刻,前两路编码数据的模 2 加结果与分量编码器系数 $\{g_{1i}\}$ 、 $\{g_{2i}\}$ 有关,如果两组系数中从第 1 位开始到第 p ($p \leq m$) 位, $g_{1,i} = g_{2,i}$ ($1 \leq i \leq p$) 均成立,故

$$u_k \oplus c_{1,k} = 0, \quad 0 \leq k \leq p \quad (7)$$

不难发现,在式(1)中列向量 u_0 和 $c_{1,0}$ 肯定相同,同时对 $1 \leq i \leq p$,如果 $g_{1,i} = g_{2,i}$ 成立,则 u_i 和 $c_{1,i}$ 也相同.

由于 $g_{1,0} = 1$,由图 1 的结构可知, u_k 与寄存器状态 $s_k = (s_{k,1}, s_{k,2}, \dots, s_{k,m})$ 之间存在约束关系为:

$$u_k = g_{1,0} s_k \oplus \sum_{i=1}^m \oplus g_{1,i} s_{k-i} = \sum_{i=0}^m \oplus g_{1,i} s_{k-i} \quad (8)$$

将式(2)和(8)左右两边相加,进一步得到

$$\begin{aligned} u_k \oplus c_k &= \sum_{i=0}^m \oplus g_{1,i} s_{k-i} \oplus \sum_{i=0}^m \oplus g_{2,i} s_{k-i} \\ &= \sum_{i=0}^m \oplus (g_{1,i} \oplus g_{2,i}) \cdot s_{k-i} \end{aligned} \quad (9)$$

从而对于第二路数据的归零部分有



图2 矩阵 D 列差分结果示意图

3.2 编码序列帧长识别

实际接收到的编码序列起始比特往往并不是帧起点,而帧长也未知,此时可以对帧长进行遍历检验,先设定一个假定的帧长值 n ,然后按式(1)建立矩阵 C_n . 当设定的 n 对应真实帧长时,即使矩阵 C_n 的第一列并不是每一帧的起点,其中仍存在相同的列,通过差分处理后矩阵 D_n 中仍存在特征列;反之,当设定的 n 不是真实帧长时,矩阵 C_n 中各列呈随机分布,从而矩阵 D_n 中存在特征列的概率很低.

设误比特率为 p_e ,当设定的 n 对应真实帧长时,特征列 d_i 中元素 $d_{i,j}$ ($1 \leq j \leq L$) 取值为 0 或 1 的概率分别为

$$c_{0,L+j} \oplus c_{1,L+j} = \sum_{i=0}^m \oplus (g_{1,i} \oplus g_{2,i}) \cdot s_{1,L+j-i} \quad (10)$$

同理,对第三路数据的归零部分有

$$c_{0,L+m+j} \oplus c_{2,L+j} = \sum_{i=0}^m \oplus (g_{1,i} \oplus g_{2,i}) \cdot s_{2,L+j-i} \quad (11)$$

在每一帧的结尾部分,通过输入归零码元,分量编码寄存器中的值顺次变为 0. 假设任意 $q \leq j \leq m$,分量编码器系数满足 $g_{1,j} = g_{2,j}$,则根据式(10)和(11)可得

$$c_{0,L+j} \oplus c_{1,L+j} = 0, \quad q \leq j \leq m \quad (12)$$

$$c_{0,L+m+j} \oplus c_{2,L+j} = 0, \quad q \leq j \leq m \quad (13)$$

因此,如果 $g_{1,j} = g_{2,j}$ 对任意 $q \leq j \leq m$ 成立,则式(1)中列向量 $c_{0,L+j}$ 和 $c_{1,L+j}$ 相同, $c_{2,L+m+j}$ 和 $c_{2,L+j}$ 也相同. 一般情况下,对于 RSC 分量编码器总有 $g_{1,0} = g_{2,0} = 1$ 和 $g_{1,m} = g_{2,m} = 1$,因此如果将式(1)的列从左至右依次两两做差分运算,生成 $W \times (3L + 4m - 1)$ 维矩阵

$$\begin{aligned} D &= [u_0 \oplus c_{1,0} \quad c_{1,0} \oplus c_{2,0} \quad \dots \quad u_{L-1} \oplus c_{1,L-1} \quad c_{1,L-1} \oplus c_{2,L-1} \\ &\quad \oplus c_{2,L-1} \quad c_{2,L-1} \oplus u_L \quad u_L \oplus c_{1,L} \\ &\quad \dots \quad u_{L+2m-1} \oplus c_{2,L+2m-1}] \\ &= [y_1 \quad y_2 \quad \dots \quad y_i \quad \dots \quad y_{3L+4m-1}] \end{aligned} \quad (14)$$

则理论情况下 D 中左右两端一定存在全零列. 但实际接收序列由于误码的存在,矩阵 C 中原本相同的列可能在部分对应位置的元素不同,从而差分后原本的全零列可能变为码重很小的列,为便于叙述,将矩阵 D 中的全零列或码重很低的列称为特征列. 经分析可知,特征列的分布如图 2 所示,其中 0 表示特征列, X 表示随机向量,这就给帧长的识别提供了依据. 为了便于后面的帧同步分析,将整个矩阵 D 分为如图 2 所示的 5 部分.

$$P(d_{i,j} = 0) = (1 - p_e)(1 - p_e) + p_e^2 = 1 - 2p_e + 2p_e^2 \quad (15)$$

$$P(d_{i,j} = 1) = 1 - \Pr(d_{i,j} = 0) = 2p_e - 2p_e^2 \quad (16)$$

对于 D_n 中的随机向量 y_i ,其中元素 0 和 1 呈随机分布,于是有

$$P(y_{i,j} = 0) = P(y_{i,j} = 1) = 0.5 \quad (17)$$

当 n 不是正确帧长值时,需要分以下 3 种情况进行考虑.

(1) 若分析矩阵 C_n 的列值 n 与真实帧长 n_0 满足 $n = \alpha n_0$,其中 $\alpha \in \mathbb{N}^+$ 且 $\alpha \neq 1$,则此时 C_n 的列长是帧长的整数倍,其中的每一列仍对应三路数据在每帧编码序列中相同的位置,因此差分处理后仍存在特征列,且 D_n 的特征列个数是真实帧长下的 α 倍,其中每个元素取值为 0 或 1 的概率同样满足式(15)、(16)和(17). 后面

可以利用这种整数倍处特征列的周期分布进一步确定帧长。

(2) 若分析矩阵 C_n 的列值 n 与真实帧长 n_0 满足 $n = \frac{\gamma}{\beta} n_0$, 其中 $\beta, \gamma, n \in \mathbb{N}^+$ 且 $\frac{\gamma}{\beta} \notin \mathbb{N}^+$, 不妨令

$$\varepsilon = \begin{cases} \frac{\gamma}{\beta} - \lfloor \gamma/\beta \rfloor, & \frac{\gamma}{\beta} - \lfloor \gamma/\beta \rfloor \leq \frac{1}{2} \\ 1 - \frac{\gamma}{\beta} - \lfloor \gamma/\beta \rfloor, & \frac{1}{2} < \frac{\gamma}{\beta} - \lfloor \gamma/\beta \rfloor < 1 \end{cases} \quad (18)$$

其中 $\lfloor \cdot \rfloor$ 表示向下取整, 则此时 C_n 中有 εL 的行依然对应各帧数据中相同的位置. 以列长 $\frac{n_0}{2}$ 为例进行说明, 建立的分析矩阵如图 2 所示, 可以看出图中奇数行为每一帧的前半部分, 偶数行为每一帧的后半部分, 因此奇数行和偶数行分别前后对应.

在误比特率为 p_e 下, 对矩阵 C_n 邻列前后差分后, 对原本特征列 d_i 所在的列, 有

$$P(d_{i,j} = 0) = \varepsilon[(1-p_e)(1-p_e) + p_e^2] + \frac{1}{2}(1-\varepsilon) \quad (19)$$

$$P(d_{i,j} = 1) = -\varepsilon[(1-p_e)(1-p_e) + p_e^2] + \frac{1}{2}(1+\varepsilon) \quad (20)$$

而对其他的随机向量 y_i , 其 0, 1 分布概率与式(17)相同.

(3) 若分析矩阵 C_n 的列值 n 与真实帧长 n_0 不满足上述任何一种情况, 则 D_n 所有列中元素均呈随机分布, 此时所有列向量中元素取值为 0 或 1 的概率均满足式(17).

从以上分析可知, 可以建立假设检验, 对帧长进行判决. 为了建立关于特征列的统计量, 首先将矩阵 D_n 中 0 元素变为 1, 1 元素变为 -1, 即

$$Y_n = 1 - 2 \cdot D_n \quad (21)$$

并设 $y_i = (y_{i,1}, y_{i,2}, \dots, y_{i,N})^T$ 表示矩阵 Y_n 第 i 列. 取

$$w_i = \frac{1}{N} \sum_{j=0}^N y_{i,j} \quad (22)$$

不难发现, 当 n 为真实帧长或满足上述第(1)种情况且 y_i 对应特征列 d_i 时, w_i 取得极大值. 下面对不同情况下 w_i 的分布特点进行具体分析, 并确定帧长的判决门限.

根据中心极限定理, 当 N 足够大时, w_i 分布趋近于高斯分布 $w_i \sim N(\mu, \sigma^2/N)$, 因此当 n 为真实帧长或满足上述第(1)种情况且 d_i 为特征列时, $w_i \sim N(u_1, \sigma_1^2)$, 其中

$$u_1 = 1 - 4p_e + 4p_e^2 \quad (23)$$

$$\sigma_1^2 = 8(1-2p_e+2p_e^2)(p_e-p_e^2) \quad (24)$$

当 n 满足第(2)种情况且对原本特征列 d_i 所在的

列有 $w_i \sim N(u_2, \sigma_2^2)$, 其中:

$$u_2 = \varepsilon - 4\varepsilon p_e(1-p_e) \quad (25)$$

$$\sigma_2^2 = 1 - \varepsilon^2 + 4\varepsilon[(1-p_e)(1-p_e) + p_e^2] - 4\varepsilon^2[(1-p_e)(1-p_e) + p_e^2]^2 \quad (26)$$

在其他所有情况下, 有 $w_i \sim N(0, 1/N)$.

不难看出, 按照不同情况下的概率分布, 通过设定合适的判决门限, 即可实现参数的识别. 设虚警概率为 P_{fa} , 漏检概率为 P_{nd} , 根据最小错误概率准则, 定义判决门限 λ 为

$$\lambda = \arg \min_{\lambda} (P_{fa} + P_{nd}) \quad (27)$$

其中,

$$P_{fa} + P_{nd} = \int_{\lambda}^{\infty} \frac{\sqrt{N}}{\sqrt{2\pi}\sigma_2} e^{-\frac{N(\lambda-\mu_2)^2}{\sigma_2^2}} + \int_{-\infty}^{\lambda} \frac{\sqrt{N}}{\sqrt{2\pi}\sigma_1} e^{-\frac{N(\lambda-\mu_1)^2}{\sigma_1^2}} \quad (28)$$

在式(28)两端同时对 λ 求偏导并令右侧 0, 化简为:

$$\frac{N(\lambda-\mu_1)^2}{\sigma_1^2} - \frac{N(\lambda-\mu_2)^2}{\sigma_2^2} - \ln\left(\frac{\sigma_2}{\sigma_1}\right) = 0 \quad (29)$$

式(29)是一个关于 λ 的一元二次方程, 将其化为标准形式

$$a\lambda^2 + b\lambda + c = 0 \quad (30)$$

其中, $a = N\sigma_2^2 - N\sigma_1^2$, $b = -2N\mu_1\sigma_2^2 + 2N\mu_2\sigma_1^2$, $c = N\mu_1^2\sigma_2^2 - N\mu_2^2\sigma_1^2 - \sigma_1^2\sigma_2^2 \ln\left(\frac{\sigma_2}{\sigma_1}\right)$. 实际情况下, 一般满足 $0 < \varepsilon \leq 0.5$, $0 \leq p_e = 0.5$, 不妨取极限情况下的 $\varepsilon = 0.5$, 经验证此时 $b^2 - 4ac \geq 0$ 恒成立, 且方程应取的实根为 $\lambda = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$.

综上, 归零 Turbo 码的帧长识别过程可总结如下: 首先设定一个帧长初值 n , 利用接收数据建立 C_n , 然后根据式(21)、(22)计算 w_i 值, 并用判决门限 λ 依次进行判决. 如果 w_i 中所有值均无法通过判决, 则 n 加 1 并重复上面的步骤; 否则, 取分析矩阵列长为 an , 重复上述步骤. 最后, 根据通过判决的 w_i 值所呈现的周期分布, 确定帧长的真实值.

3.3 编码序列帧起点同步

下面对帧起点的同步进行分析. 图 2 中包含特征列的区域为 1、3 和 5 三部分, 可以发现前端第 1 部分中相邻特征列的间隔为 2, 末端第 3 和第 5 两部分特征列的分布完全相同, 且相邻特征列之间的间隔为 1, 而最后 1 个特征列位于第 5 部分末端, 因此根据最后一个特征列的相对位置即可确定帧起点. 在识别出帧长 n 后, 当构建的分析矩阵 C_n 起点不是编码序列帧起点时, 矩阵 D_n 的起点也可能在图 2 中五部分的任意位置, 下面分为以下 3 种情况进行分析.

(1) 矩阵 D_n 的起点处于图 2 中第 1 部分区域, 如

图 3(a)所示.此时,末端第 3 和第 5 两部分保持完整,且其相对位置保持不变,而在第 5 部分后面间隔 1 位后会出现原第 1 部分的第 1 个特征列,因此可以很容易区

分出最末位处的特征列 $d_f (1 \leq f \leq 3L + 4m - 1)$, 则矩阵 C_n 中第 $(d_f + 2) \bmod (3L + 4m)$ 位对应帧起点.

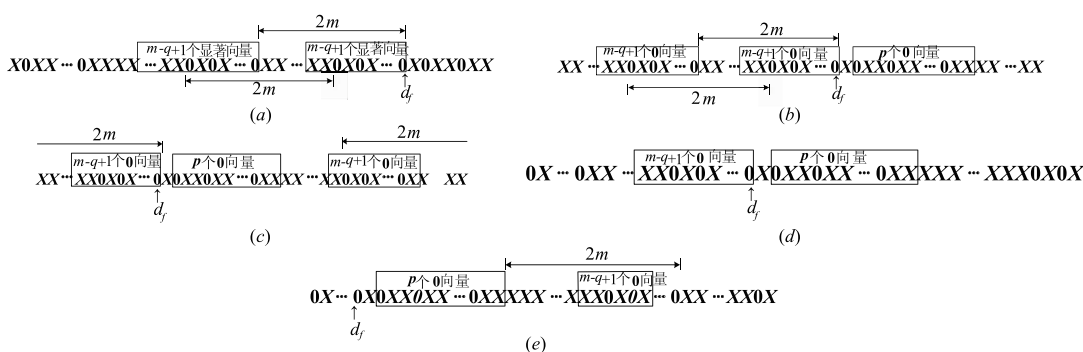


图3 末端两部分保持完整时的全零列分布示意图

(2) 矩阵 D_n 的起点处于图 2 中第 2 和第 4 部分区域,如图 3(b)、(c)所示.此时包含特征列的第 1、3 和第 5 三部分区域均保持完整,第 3 和第 5 部分区域在矩阵中的相对位置可能发生改变,但第 1 部分总是紧跟在第 5 部分的后面,据此可很容易按上面第 (1) 种情况相同的方式确定 d_f 的位置,进而得到矩阵 C_n 中帧起点的位置.

比特以模拟实际接收环境,然后采用本文所提的方法进行识别.

(3) 矩阵 D_n 的起点处于图 2 中第 3 和第 5 部分区域,如图 3(d)、(e)所示.此时第 1、2 和第 4 三部分区域均保持完整,可以看出,第 2 部分包含的随机列个数比第 4 部分的多,而每一帧的最后一个全零列应分布在第 4 部分之后,第 1 部分之前,根据这一特性,也很容易确定矩阵 C_n 中帧起点的位置.

首先对帧长进行分析,结果如图 4 所示,其中虚线表示判决门限 T_1 .可以看出, $n = 108$ 时最先有 w_i 最大值通过判决门限,后续又在 216 和 324 处通过判决,因此帧长为 108,与实际相符,识别正确.然后对帧起点点进行同步,取 $n = 108$,然后对所有 107 个 w_i 值进行判决,结果如图 5 所示,图中有三个位置能通过判决,分别是 71、77 和 79.与 3.3 节中两种情况对照不难看出,该情形与图 3(b)相似,因此原编码序列中第 79bit 对应帧起点,而前面的 78bit 数据属于上一帧中的不完整数据,该帧共缺失 $108 - 78 = 30$ bit 数据,与设定条件相符识别正确.

综合上面的三种情况,可以归纳得到归零 Turbo 码帧起点的同步步骤如下:首先判断差分矩阵 D_n 中是否存在特征列间隔为 1、数量相同的两个区域(即第 3 和第 5 部分),如果存在则根据间隔 1 位后是否仍出现特征列,确定最末位处的特征列 d_f 的位置;如果 D_n 中不存在上述两部分,则首先确定包含特征列且间隔为 2 的区域(即第 1 部分),然后确定包含随机列的两个区域(即第 2 和第 4 部分),选择随机列较少的那部分,并在其与第 1 部分之间的区域确定最末位处的特征列 d_f 的位置;最后,原矩阵 C_n 中第 $(d_f + 2) \bmod (3L + 4m)$ 位即对应帧起点.

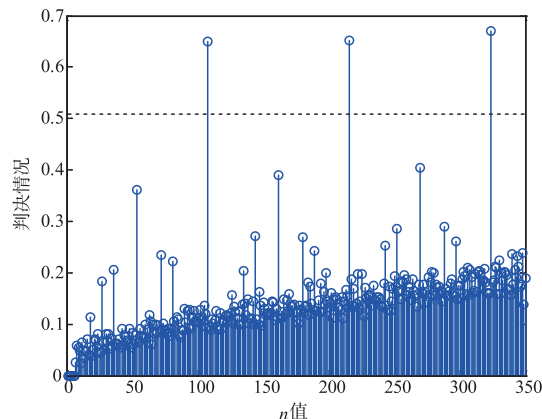


图4 帧长识别结果

4 仿真与分析

4.1 方法可行性验证

取交织长度为 32, 交织关 $a_L = (31, 11, 26, 14, 27, 10, 30, 32, 3, 1, 24, 22, 16, 23, 18, 9, 25, 17, 13, 5, 8, 20, 21, 6, 19, 12, 7, 2, 15, 4, 29, 28)$, 分量编码器生成多项式为 (13, 15), 即 $g_1(x) = 1 + x^2 + x^3$, $g_2(x) = 1 + x + x^3$ 的 1/3 归零 Turbo 码进行仿真.首先在误码率 $p_e = 0.01$ 条件下生成 1.08×10^5 bit 数据,并去掉序列最前面的 30

4.2 性能分析

根据前面方法介绍可知,影响帧长识别和帧起点同步的主要因素包括误比特率、数据帧数、交织长度和分量编码器参数(包括 m 值和生成多项式系数分布),下面将分别这些因素进行研究.

首先分析帧块数目对算法的影响,设定交织长度

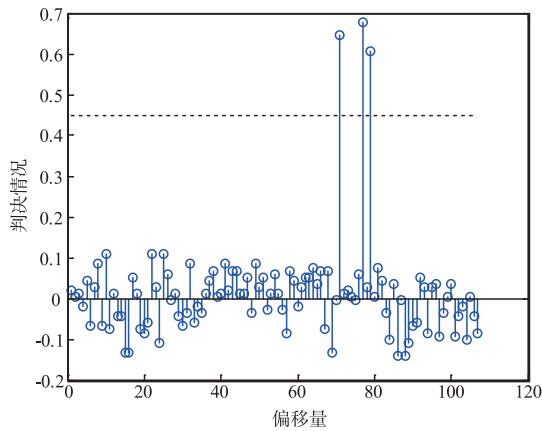
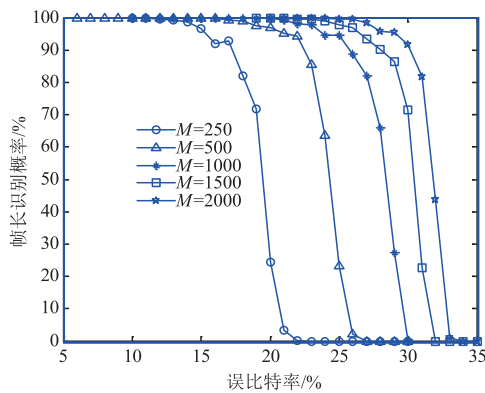
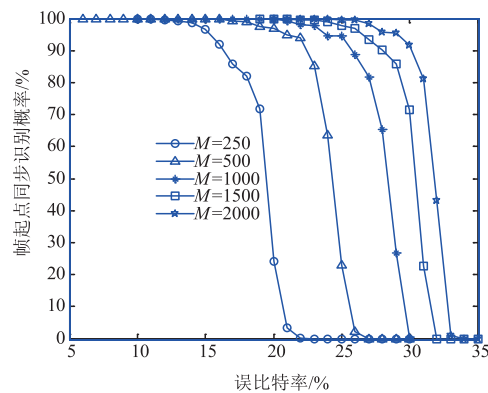


图5 帧起点同步结果

为 128, 取数据帧数 M 分别为 250, 500, 1000, 1500 和 2000, 5 种, 误比特率设定在 0.05 至 0.34 之间, 然后给出在不同误比特率, 1000 次蒙特卡洛仿真的结果, 分别如图 6(a)、(b) 所示.

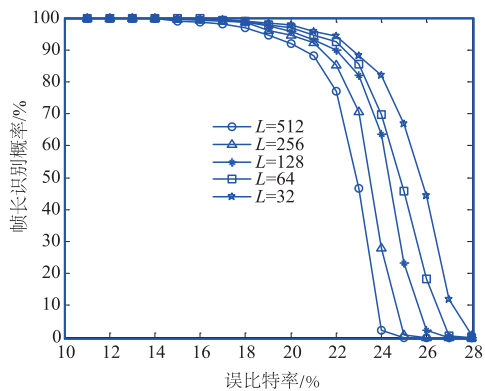


(a)对帧长识别的影响

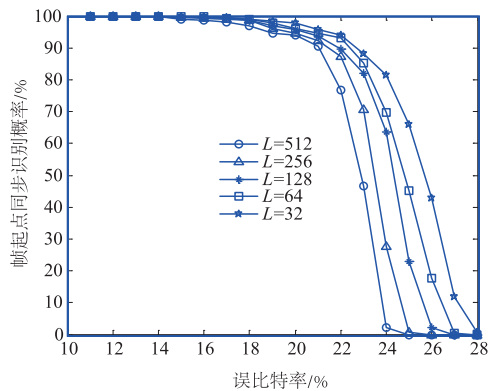


(b)对帧起点同步的影响

图6 数据帧数对识别的影响



(a)对帧长识别的影响



(b)对帧起点同步的影响

图7 交织长度对识别的影响

下面分析分量编码器参数对识别的影响, 主要从编码记忆长度和生成多项式系数分布两方面进行考虑. 选取 (13, 17)、(15, 17)、(25, 33) 和 (27, 37) 等四组

由图 6 可知, 在相同误比特率下帧数越多, 识别概率越高. 由于误比特率会破坏相邻列之间的相关性, 因此帧长识别和帧起点同步成功的概率随着误比特率的增加而逐渐降低. 此外, 帧起点同步识别率要略小于帧长识别率, 主要原因是起点识别需要准确标定所有特征列的位置, 而帧长识别只需要 1 个特征列, 因此帧起点识别对误比特率更敏感, 这也与实际相符. 同时可以看出, 在相同误比特率下数据帧数越多, 识别效果越好.

下面分析交织长度对识别的影响. 取交织长度分别为 32, 64, 128, 256 以及 512, 5 种; 分量编码器生成多项式为 (13, 15) 的 1/3 归零 Turbo 码进行研究, 数据帧数固定为 500, 误比特率设定在 0.10 至 0.28 之间, 然后在对应误比特率进行 1000 次蒙特卡洛仿真, 结果如图 7(a)、(b) 所示. 可以看出在相同误比特率下, 交织长度越大, 识别效果越差. 这是因为帧长随交织长度的增加而线性增长, 计算特征列时带来的干扰就越多.

分量编码器进行对比研究, 交织长度设为 32, 数据帧数 N_1 固定为 500, 然后在对应误比特率进行 500 次蒙特卡洛仿真, 结果如图 8(a)、(b) 所示. 可以看出, 4 个分量

编码器下的帧长识别结果相近,但帧起点同步性能各有差异,其中(13,17)和(25,33)的识别结果最好,(15,17)次之,(27,37)表现最差.将四组分量编码器变换为多项式形式有

$$G_1(x) = [1, (1+x+x^3)/(1+x^2+x^3)],$$

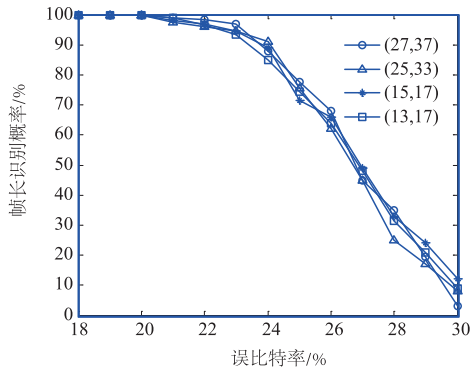
$$G_2(x) = [1, (1+x+x^2+x^3)/(1+x^2+x^3)],$$

$$G_3(x) = [1, (1+x+x^3+x^4)/(1+x^2+x^4)],$$

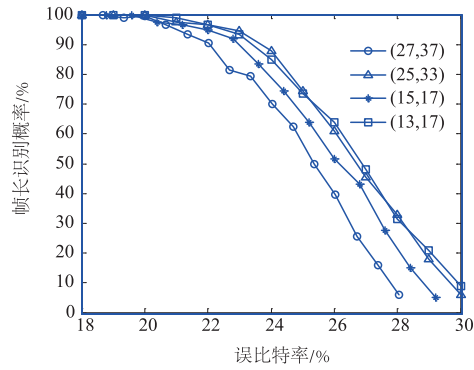
$$G_4(x) = [1, (1+x+x^2+x^3+x^4)/(1+x^2+x^3+x^4)].$$

不难发现,虽然(13,17)和(25,33)的编码记忆长

度分别为 3 和 4,但两者生成多项式矩阵中第二项分子分母系数中均只有首尾两项相同(即 $g_{1,0}^{(i)} = g_{2,0}^{(i)}, g_{1,m}^{(i)} = g_{2,m}^{(i)}, i = 1, 3$),而(15,17)和(27,37)中分别有连续 2 项和 3 项相同,四组分量编码器对应的特征列个数分别为 3、5、3 和 7.由于帧长识别只需检测特征列是否存在,而帧起点同步需要检测所有的特征列,因此分量编码器对识别的影响主要取决于生成多项式中系数分布情况,在相同误比特率下帧长识别性能相同而帧起点同步概率随特征列的增加而减小.



(a)对帧长识别的影响



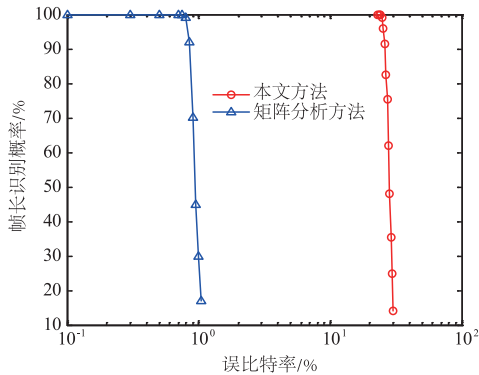
(b)对帧起点同步的影响

图8 分量编码器参数对识别的影响

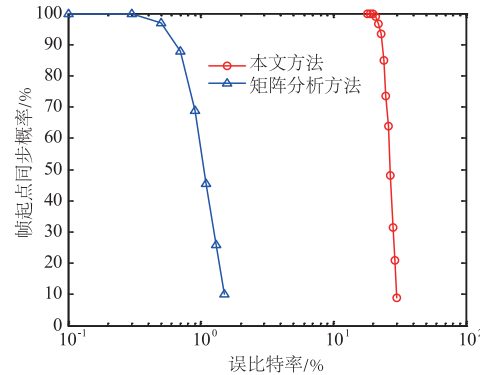
4.3 与现有方法的对比

下面将本文所提供方法与文献[17]中基于矩阵分析的方法进行对比.仍然以交织长度为 32,分量编码器为(13,15)的 1/3 码率归零 Turbo 码为例,在数据帧数为 500 的情况下两种方法在不同误比特率下的识别性

能如图 9(a)、9(b)所示.可以看出,对比方法仅能在误比特率小于 10^{-3} 时能完成识别,而本文所提方法识别性能明显优于前者,在误比特率为 10^{-2} 量级时仍然有效.



(a)帧长识别对比



(b)帧起点同步对比

图9 与文献[17]方法识别对比结果

下面对识别所需的数据量(帧数)进行分析和对比.文献[17]中构建分析矩阵时,需满足行数为列数的 2 倍,即 $S = 2(3L + 4m)$.对本文所提方法,根据统计学上“3 倍标准差准则”,取 $u_1 - 3\sqrt{\sigma_1^2/N} > u_2 + 3\sqrt{\sigma_2^2/N}$,解得 $N > 9(\sigma_1 + \sigma_2)^2$.选取分量编码器为(13,15)RSC 码,在误比特率分别为 0.001 和 0.01,交织长度分别为 32,64 和 128 的情况下,两种方法所需的

数据帧数如表 1 所示.可以看出,在相同仿真条件下,本文识别时所需的数据量远小于对比算法,具有更大的实用价值.

下面对识别过程中的计算复杂度进行分析和对比.根据文献[17],基于矩阵分析的方法计算复杂度为 $O((3L + 4m)^2)$.本文识别过程中,帧长识别和帧起点

同步模 2 加运算量为 $\sum_{i=7}^{3L+4m} [S(i-1)]$, 分量码识别的模 2 加运算量为 $[0.5S(L-m) + 2(m+1)]2^{2(m+1)}$, 交织关系的计算复杂度为 $\sum_{i=2}^L (w-1+i)S$. 由此可知, 尽管本文方法在交织长度较小时计算复杂度要大于对比方法, 但后者计算复杂度随交织长度的增长速率要远大于本文方法. 因此在交织长度较大时, 本文所提方法更具优势.

表 1 误码率为 0.001 时所需数据帧数对比

误比特率	交织长度	文献[17]所需数据量	本文方法所需数据量
0.001	32	216	113
	64	408	113
	128	792	113
0.01	32	216	131
	64	408	131
	128	792	131

5 结论

针对归零 Turbo 码的识别, 提出了一种基于矩阵模型相关性识别的方法. 根据归零结构与分量编码器参数特性, 总结出数据帧构建的矩阵中列之间的相关性, 然后对该矩阵前后相邻列进行连续差分并建立关于差分后矩阵列重的统计量, 通过其分布特点完成帧长识别和帧起点的同步. 从识别性能、所需数据量和计算复杂度多方面的对比结果表明, 与现有的基于矩阵分析的方法相比, 本文所提方法性能更优.

参考文献

- [1] BARBIER J. Reconstruction of Turbo-code encoders[J]. the International Society for Optical Engineer, 2005, 5819(5): 463-473.
- [2] NASERI A, AZMOON O, Samad Fazeli. Blind recognition algorithm of Turbo codes for communication intelligence systems[J]. International Journal of Computer Science Issues, 2011, 8(6): 68-72.
- [3] 钟阳晶, 梁茹冰. 基于多项式遍历的码字相关的 Turbo 码盲识别[J]. 电讯技术, 2017, 57(7): 819-824. ZHONG Yangjing, LIANG Rubing. Blind recognition of Turbo code based on polynomial traverse and codes correlation[J]. Telecommunication Engineering, 2017, 57(7): 819-824. (in Chinese)
- [4] BARBIER J. Reconstruction of Turbo-code encoders[J]. the International Society for Optical Engineer, 2005, 5819(5): 463-473.
- [5] GAUTIER R, MARAZIN M, and BUREL G. Blind recovery of the second convolutional encoder of a Turbo-code when its systematic outputs are punctured[A]. IEEE Communications[C]. Bucharest: IEEE Press, 2008. 345-348.
- [6] MARAZIN, M, GAUTIER R and BUREL G. Blind recovery of k/n rate convolutional encoders in a noisy environment[J]. EURASIP Journal on Wireless Communications and Networking, 2011, 2011(1): 168.
- [7] DEBESSU Y G, WU H, JIANG H. Novel blind encoder parameter estimation for Turbo codes [J]. IEEE Communications Letters, 2012, 16(12): 1917-1920.
- [8] DEBESSU Y G, WU H C, JIANG H, et al. Blind encoder parameter estimation for Turbo codes [A]. IEEE Global Communications Conference (GLOBECOM) [C]. Anaheim: IEEE Press, 2012. 4233-4237.
- [9] YU P, LI J, PENG H. A least square method for parameter estimation of RSC sub-codes of Turbo codes [J]. IEEE Communication Letters, 2014, 18(4): 644-647.
- [10] YU P D, LI J, PENG H. Gibbs sampling based parameter estimation for RSC sub-codes of Turbo codes [A]. Sixth International Conference on Wireless Communications and Signal Processing (WCSP) [C]. USA: IEEE Press, 2014. 1-5.
- [11] CØTE M, SENDRIER N. Reconstruction of a Turbo-code interleaver from noisy observation [A]. International Symposium on Information Theory [C]. Austin, Texas, USA: IEEE Press, 2010. 2003-2007.
- [12] CLUZEAU M, FINIASZ M, TILLICH J P. Methods for the reconstruction of parallel Turbo codes [A]. International Symposium on Information Theory [C]. Austin: IEEE Press, 2010. 2008-2012.
- [13] TILLICH J P, TIXIER A, SENDRIER N. Recovering the interleaver of an unknown Turbo-code [A]. IEEE International Symposium on Information Theory [C]. Honolulu: IEEE Press, 2014. 2784-2788.
- [14] 吴昭军, 张立民, 钟兆根. 低信噪比下随机交织器识别 [J]. 电讯技术, 2018(1): 58(01): 52-58. WU Zhaojun, ZHANG Limin, ZHONG Zhaogen. Blind recognition of random interleaver at low SNR [J]. Telecommunication Engineering, 2018, 58(1): 52-58. (in Chinese)
- [15] 刘俊, 李静, 彭华. 基于校验方程平均符合度的 Turbo 码交织器估计 [J]. 电子学报, 2016, 44(5): 1213-1217. LIU Jun, LI Jing, PENG Hua. Estimation of Turbo-code interleaver based on average conformity of parity-check equation [J]. Acta Electronica Sinica, 2016, 44(5): 1213-1217. (in Chinese)
- [16] 张永光. 一种 Turbo 码编码参数的盲识别方法 [J]. 西安电子科技大学学报 (自然科学版), 2011, 38(2): 167-172. ZHANG Yongguang. Blind recognition method for the

Turbo coding parameter[J]. Journal of Xidian University, 2011, 38(2):167-172. (in Chinese)

- [17] 李啸天,张润生,李艳斌. 归零 Turbo 码识别算法[J]. 西安电子科技大学学报(自然科学版),2013,40(4):161-166.

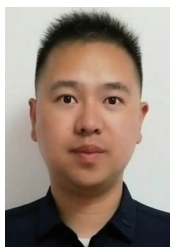
LI Xiaotian, ZHANG Runsheng and LI Yanbin. Research on the recognition algorithm of Turbo codes on trellis termination[J]. Journal of Xidian University, 2013, 40(4):

161-166. (in Chinese)

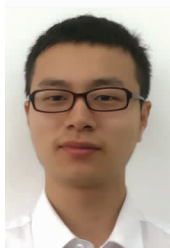
- [18] 张旻,陆凯,李歆昊,吕全通. 归零 Turbo 码的盲识别方法[J]. 系统工程与电子技术, 2016, 38(6):1424-1427.

ZHANG Min, LU Kai, LI Xinhao, Lü Quantong. Blind recognition method for the Turbo codes on trellis termination[J]. Journal of Systems Engineering and Electronics, 2016, 38(6):1424-1427. (in Chinese)

作者简介



钟兆根 男,1984年4月出生于江西南昌,博士,海军航空大学讲师,主要研究方向为通信信号盲分离与统计信号处理.
E-mail:zhongzhaogen@163.com



吴昭军(通信作者) 男,1992年9月出生于四川遂宁.现为海军航空大学博士研究生.主要研究方向信道编码盲识别.
E-mail:wuzhaojun1992@qq.com



刘杰 男,1990年出生于湖北宜昌.现为海军航空大学博士研究生.主要研究方向为现代信号处理以及应用.
E-mail:iamlIU1573@163.com



张立民 男,1966年出生于辽宁开原.现为教授、博士生导师,主要研究方向为卫星信号处理及应用.
E-mail:iamzlm@163.com