文章编号：2095-4980(2015)06-0877-06

非收尾 Turbo 码交织器的识别方法

任亚博[1,2]，张　健[2]，刘以农[1]，张　伟[2]

(1.清华大学 工程物理系，北京 100084；2.中国工程物理研究院 电子工程研究所，四川 绵阳 621999)

摘　要：提出了一种对非归零 Turbo 码交织器的识别方法。二元域上某类有理式的级数展开具有周期性，从而可知，在同一个周期点上，来自信息序列上 2 个比特与相应交织位上的 2 个比特之和为 0，进而可得到一种识别方法：第 1 步恢复出 2 个交织位置，此后每步恢复 1 个位置。根据周期长度的不同，算法被分成若干条并行链路，链路数即为周期长度。

关键词：非收尾 Turbo 码；交织器；识别

中图分类号：TN762　　**文献标识码：**A　　**doi**：10.11805/TKYDA201506.0877

Recognition of unterminated Turbo-code interleaver

REN Yabo[1,2]，ZHANG Jian[2]，LIU Yinong[1]，ZHANG Wei[2]

(1.Department of Engineering Physics，Tsinghua University，Beijing 100084，China；
2.Institute of Electronic Engineering，China Academy of Engineering Physics，Mianyang Sichuan 621999，China)

Abstract：An algorithm to recover an unterminated turbo-code interleaver is proposed. It is found that for some rational functions, their series are periodic, and then the sum of two bits from the info sequence and two bits from the corresponding interleaver sequence is 0. There comes out a recognition algorithm: two interleaved positions are recovered first; other positions would be recovered step by step. The algorithm is sliced into several lists, and the number equals to the period.

Key words：unterminated turbo-code；interleaver；recognition

编码识别已成为一个研究热点[1-3]，其中 Turbo 码的识别主要在于其交织器的识别，目前集中在收尾 Turbo 码(也称归零 Turbo 码)的识别。Barbier 在文献[4]中提出了一种树状分枝的方法，如果已知交织前的数据和交织后的数据，可以有效恢复出交织器，算法没有提及如何获取交织后的数据；M Cluzeau 等人在文献[5]中提出了一种基于 Turbo 码译码的交织器恢复方法，该法可以有效恢复出归零的 Turbo 交织器；M Cˆote 等人在文献[6]中提出了一种特别的方法，通过将反馈卷积编码器的生成有理式(分子与分母均为多项式)作泰勒级数展开，可恢复出下一时刻交织后的数据，进而逐步恢复出整个交织器。国内方面，张永光老师在文献[7]中提出了一种无误码条件下的盲识别方法，该法将编码后的序列除以生成有理式，从而恢复出交织后的信息序列，进而恢复出交织器；李啸天等在文献[8-9]中提出了对 Turbo 码码长与交织器的识别。

收尾 Turbo 码编码时移位寄存器的初始状态为全 0(2 路分量码均收尾)，非收尾 Turbo 码的第 2 路卷积编码器未必收尾，即第 2 路移位寄存器的初始状态未知。目前对收尾 Turbo 交织器的恢复方法并不适用于非归零的情况。对非收尾码的分析表明，在同一个周期点上，来自信息序列上的 2 个比特与交织位上的 2 个比特之间具有相关性，其相关性不依赖于编码器的初始状态，据此本文提出了一种恢复交织器的方法。

1　Turbo 码识别模型

典型的并行级联 Turbo 码的结构如图 1 所示，其中信息序列为 x，经过编码器 1 后得到的

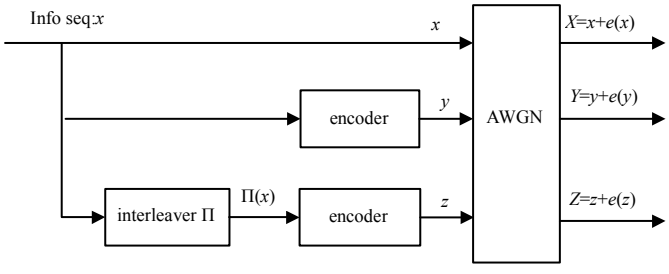

Fig.1 Turbo-code encoder
图 1 Turbo 码编码器

收稿日期：2014-08-17；修回日期：2014-10-05

序列为 y；信息序列经过交织后得到 $\prod(x)$，交织器的长度记为 L，编码后的序列为 z。3 路序列经过二进制对称信道后，接收方得到的序列为 X,Y,Z。图中 $e(\bullet)$ 表示噪声，为 1 个 0,1 序列，其中 1 的概率即为接收序列的误比特率。设 3 路的误比特率均为 τ。

本文关于 Turbo 编码的运算均在 GF(2)或其扩域上进行，不再区分加法与减法。

为方便描述，本文中将序列 $x=x_0x_1x_2x_3\cdots$ 表示成多项式形式 $x=x_0+x_1D+x_2D^2+x_3D^3\cdots$，同样将 $\Pi(x),y,z$ 以及 X,Y,Z,e 均表示成多项式形式。考虑到 $\Pi(x)=x_{r_0}+x_{r_1}D+x_{r_2}D^2+x_{r_3}D^3+\cdots$，因此恢复交织器 Π 就是要计算出 $r_0,r_1,r_2,r_3\cdots$ 的取值。

假设反馈卷积编码器 2 的生成矩阵为 $\left[1,\ g_2(D)/g_1(D)\right]$，其中 $g_1(D)$ 与 $g_2(D)$ 为多项式，$g_j(D)=\sum_{k=0}^{d}g_{j,k}D^k$，$g_{j,0}=g_{j,d}=1$，$j=1,2$，对于归零 Turbo 码，每次编码前其移位寄存器初始的状态为 0，则有式(1)成立：

$$z=\Pi(x)\times\frac{g_2(D)}{g_1(D)} \tag{1}$$

2 识别方法

对于非归零 Turbo 码，其第 2 路卷积编码器的初始状态一般不为全 0，考虑移位寄存器长为 d，它的任一状态均能由全 0 状态输入 d 个比特得到，这相当于存在次数小于 d 的多项式 $q_x(D)$ 与 $q_z(D)$，满足

$$q_z(D)+D^d\times z(D)=\left(q_x(D)+D^d\times\Pi(x)\right)\times\frac{g_2(D)}{g_1(D)} \tag{2}$$

式(2)也可理解成：将信息序列 $\Pi(x)$ 的前面增加 d 个比特 $q_x(D)$，得到了序列 $q_x(D)+D^d\times\Pi(x)$，该序列经过原来的编码器(移位寄存器的初始状态为全 0)后得到一个序列 $q_z(D)+D^d\times z(D)$，再去除前 d 个比特 $q_z(D)$ 即得到原来的 $z(D)$。

对识别者而言，已知的内容为接收到的序列 X,Y,Z，与反馈卷积编码器 2 的生成矩阵为 $\left[1,\ g_2(D)/g_1(D)\right]$。而 x,y,z 及 $e(x),e(y),e(z),\Pi(x)$ 为未知。它们之间有如下关系：

$$X=x+e(x) \tag{3}$$

$$Y=y+e(y) \tag{4}$$

$$Z=z+e(z) \tag{5}$$

将式(3)~式(5)与式(2)结合有

$$Z\times\frac{g_1(D)}{g_2(D)}=\Pi(X)+D^{-d}\times q_z(D)\times\frac{g_1(D)}{g_2(D)}+\Pi(e(x))+D^{-d}\times q_x(D)+e(z)\times\frac{g_1(D)}{g_2(D)} \tag{6}$$

为方便描述，称级数 $\sum_{i=0}^{\infty}a_iD^i$ 为循环级数，如果存在常数 T，使得该级数能表示成 $\sum_{i=0}^{\infty}a_iD^i=\sum_{j=0}^{\infty}\left(D^{jT}\sum_{i=0}^{T-1}b_iD^i\right)$，则称该级数的周期为 T。

引理 1：将有理式 $D^{-d}\times q_z(D)\times\frac{g_1(D)}{g_2(D)}$ 展成级数 $\sum_{i=-d}^{-1}a_iD^i+\sum_{i=0}^{\infty}c_iD^i$，则 $\sum_{i=0}^{\infty}c_iD^i$ 为一个循环级数。

证明：考虑到 $q_z(D)$ 的次数小于 d，而 $g_1(D)$ 的次数恰为 d，故 $q_z(D)\times\frac{g_1(D)}{g_2(D)}$ 可表示成有理式 $h_1(D)+\frac{h_2(D)}{g_2(D)}$，其中 $h_1(D)$ 与 $h_2(D)$ 的次数均小于 d。从而 $D^{-d}\times h_1(D)$ 与 $\sum_{i=0}^{\infty}c_iD^i$ 无关。若 $h_2(D)$ 为 0，则显然 $\sum_{i=0}^{\infty}c_iD^i$ 也为 0，此为循环级数，下面证明当 $h_2(D)$ 不为 0 时的情况。

取最小的正整数 T，使得 $1+D^T$ 能被 $g_2(D)$ 整除。T 的存在性证明如下：假设 T 不存在，即对任意 T，$1+D^T$ 总不能被 $g_2(D)$ 整除，考虑 2^d 个多项式：$1,1+D,1+D^2,\cdots,1+D^{2^d-1}$，对 $g_2(D)$ 取模，必存在 2 个同模的，不妨取为 $1+D^i,1+D^j\ (i<j)$，则 D^i+D^j 能被 $g_2(D)$ 整除，即 $1+D^{j-i}$ 能被 $g_2(D)$ 整除，矛盾。

取上面的 T 值，令 $\frac{h_2(D)}{g_2(D)}\times\left(1+D^T\right)=h_3(D)$，则 $h_3(D)$ 的次数小于 T，进而 $\frac{h_2(D)}{g_2(D)}=\sum_{j=0}^{\infty}D^{jT}h_3(D)$，这表明 $\frac{h_2(D)}{g_2(D)}$ 可表示成循环级数，将 $\frac{h_2(D)}{g_2(D)}$ 向左移 d 次，得到 $\sum_{i=0}^{\infty}c_iD^i$，也为循环级数。

从而式(6)可转化为：

$$Z\times\frac{g_1(D)}{g_2(D)}=\Pi\left(X\right)+\sum_{i=0}^{\infty}c_iD^i+\Pi\left(e(x)\right)+e(z)\times\frac{g_1(D)}{g_2(D)} \tag{7}$$

记：$Z\times\frac{g_1(D)}{g_2(D)}=\sum_{i=0}^{\infty}u_iD^i$，$\frac{g_1(D)}{g_2(D)}=\sum_{i=0}^{\infty}v_iD^i$，$e=\Pi\left(e(x)\right)$，$e'$=$e(z)$，可得 $\sum_{i=0}^{\infty}c_iD^i$ 是周期为 T 的级数，而 $\sum_{i=0}^{\infty}v_iD^i-1$ 亦为周期为 T 的级数，考虑式(7)中各个级数的 0 次项与 T 次项的系数，有式(8)成立：

$$u_0+u_T=X_{r_0}+X_{r_T}+e_0+e_T+e'_0v_0+\sum_{i=0}^{T}e'_iv_{T-i} \tag{8}$$

由于 $\frac{g_1(D)}{g_2(D)}$ 中有 $v_0=1,v_T=0$，因此式(8)可简化为：

$$u_0+u_T+X_{r_0}+X_{r_T}=e_0+e_T+e'_0+e'_T+\sum_{i=1}^{T-1}e'_iv_{T-i} \tag{9}$$

更有式(10)成立：

$$u_i+u_{i+T}+X_{r_i}+X_{r_{i+T}}=e_i+e_{i+T}+e'_i+e'_{i+T}+\sum_{j=i+1}^{i+T-1}e'_jv_{i+T-j} \tag{10}$$

当无误码时，式(10)的右端为 0，当有误码时，其右端为 1，当且仅当右端有奇数项取 1。设右端为 1 的概率为 P_e，由于共有 $4+\sum_{i=1}^{T-1}v_i$ 项，可知：

$$P_e=\frac{1-\left(1-2\tau\right)^{4+\sum_{i=1}^{T-1}v_i}}{2} \tag{11}$$

如果 $\{r_0,r_T\}$ 取了错误的值 $\{r'_0,r'_T\}$，则 $u_0+u_T+X_{r'_0}+X_{r'_T}$ 取 1 的概率为 0.5，从而可得到一种求解 $\{r_0,r_T\}$ 的方法，虽然此时 r_0 和 r_T 彼此之间无法区分。方法如下：

1) 取 N 个码字，分别求出其 u_0+u_T。

2) 遍历集合 $\{r_0,r_T\}$，对每个取值 $\{r'_0,r'_T\}$，求 N 个 $u_0+u_T+X_{r'_0}+X_{r'_T}$，统计其为 1 的频率 P_f，当该频率最小时(此时约为 P_e)所遍历到的 $\{r'_0,r'_T\}$ 即为正确的 $\{r_0,r_T\}$，记为 $\{m_0,m_1\}$。

同样可求解任意的 $\{r_i,r_{i+T}\}$，当求解出全部的 $\{r_i,r_{i+T}\}$，$0\leqslant i\leqslant L-1-T$，即可恢复出 $r_0,r_1,r_2,r_3\cdots$。

上面的方法改进后可以减少 1 次遍历。假设求解出 $\{r_0,r_T\}=\{m_0,m_1\}$，但 r_0 和 r_T 之间无法区分，此时先假设 $r_T=m_0$，遍历 r_{2T}，求出 $\{r_T,r_{2T}\}$，再假设 $r_T=m_1$，也可求出 1 组 $\{r_T,r_{2T}\}$，取 2 组中较优的 1 组为解，此时即求出了 r_0,r_T,r_{2T}。下一步，令式(10)中 $i=2T$，可求得 r_{3T}，进而得到 $r_0,r_T,r_{2T},r_{3T},\cdots$。依次类推可以求解出 $r_1,r_{1+T},r_{1+2T},r_{1+3T},\cdots$，$r_2,r_{2+T},r_{2+2T},r_{2+3T},\cdots$，$\cdots$，从而恢复出 $r_0,r_1,r_2,r_3\cdots$，该过程可以表示成图 2。整个算法被分成了 T 条链路，且彼此之间不受影响，这为算法的并行处理提供了可能。同时也提供了一种对算法是否成功的自觉验证：如果各条链路恢复出的数据完全不同，可判断识别结果正确。

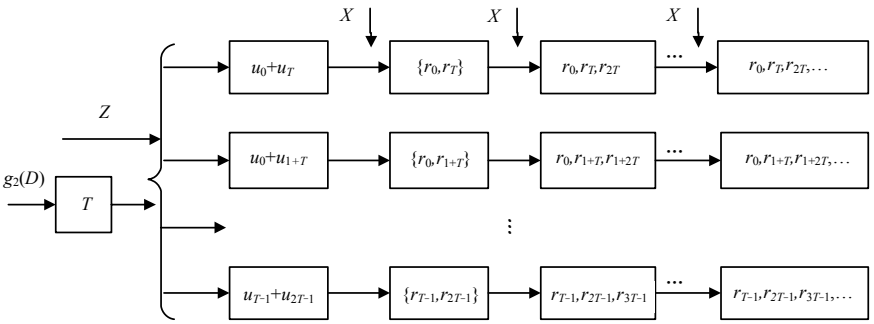

Fig.2 Diagram of the algorithm
图 2 算法的处理框图

3 算法的性能分析与参数取值

算法中的运算，主要来自于步骤 1)，全部的计算量约为 $TdNL^2$。Turbo 码中反馈卷积编码器的约束长度一般为 d=4 或更小，而 T 的值取决于 $g_2(D)$。正如引理 1 所证明的那样，T 的值小于等于 2^d-1，事实上仅当 $g_2(D)$ 为本原多项式时，等号成立。

步骤 2)中 P_f 服从的分布如下：

$$P_f \sim \begin{cases} \mathbf{N}\left(\frac{1}{2},\ \frac{1}{2\sqrt{N}}\right) & \text{非}\{r_0,r_T\} \\ \mathbf{N}\left(P_e,\ \frac{1}{\sqrt{N}}\sqrt{P_e(1-P_e)}\right) & \{r_0,r_T\} \end{cases} \tag{12}$$

记 $\lambda=\dfrac{\sqrt{N}\left(0.5-P_e\right)}{0.5-\sqrt{P_e\left(1-P_e\right)}}$，$\{r_0,r_T\}$ 被正确求解的概率为 $P\left(\{r_0,r_T\}\right)=\left(1-\Phi(\lambda)\right)^{L(L-1)/2-1}$，其中 $\Phi(\lambda)=\int_{-\infty}^{\lambda}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\mathrm{d}t$。而 r_{2T} 被正确求解的概率为 $P\left(r_{2T}\mid\{r_0,r_T\}\right)=\left(1-\Phi(\lambda)\right)^{2L-1}$，$r_{3T}$ 及 r_{iT}（$i>3$）被正确求解的概率为同一值，为：$P\left(r_{3T}\mid r_0,r_T,r_{2T}\right)=\left(1-\Phi(\lambda)\right)^{L-1}$。同样对于 $r_1,r_{1+T},r_{1+2T},r_{1+3T},\cdots$，$r_2,r_{2+T},r_{2+2T},r_{2+3T},\cdots$，…，整个交织器被完全求解的概率为：$P(r_0,r_1,r_2,\cdots,r_{L-1})\approx\left(1-\Phi(-\lambda)\right)^{(T/2+1)L\times L}$。

4 算法仿真示例

反馈卷积编码器中取 $g_1(D)=1+D^3+D^4$，$g_2(D)=1+D+D^3+D^4$，取交织长度为 L=1 000，交织图样由计算机随机产生，如图 3 所示。

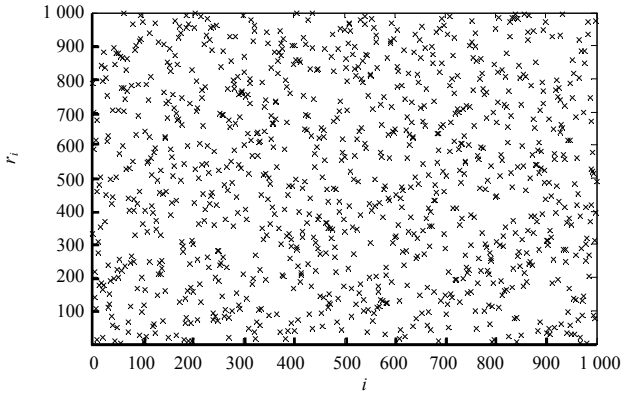

Fig.3 Preset interleaved pattern
图 3 预设的交织图样

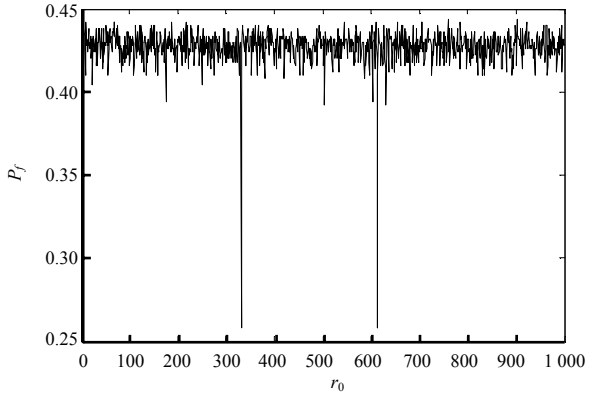

Fig.4 P_f at every point
图 4 各点的 P_f取值

对接收到的数据添加 5%的误码，取 N=500 个码字进行识别，计算求得 T=6。为计算 $\{r_0,r_6\}$，需对 r_0,r_6 进行 2 层遍历，产生 $L(L-1)/2$ 个 P_f，仿真中为了便于绘图，对 r_0,r_6 的遍历采用了 2 层互不干扰的遍历，这样共产生了 L^2 个 P_f，这相当于将数据重复了一遍。图 4 给出了当 r_0 取某一值时，r_6 遍历 1 次求得的所有 P_f 的最小值。

由图可知，当 r_0 取 332 或 612 时，P_f 有 2 个明显的最小值，且这 2 个值相等，此时对应的 r_6 分别为 612 或 332。注意到 P_f 在其他位置的取值约为 0.42~0.43，而不是 0.5，这是因为此时的 P_f 为 L 个服从 $\mathbf{N}\left(\frac{1}{2},\ \frac{1}{2\sqrt{N}}\right)$ 分布的最小值，经计算机模拟验证，该最小值的期望约为 0.423。另一方面，图 4 中的最低点理论均值应为 P_e，依式(11)，P_e 可精确计算出来为 0.260，与图 4 几乎一致。继续算法，分别求出 T=6 条链路的初始状态，如表 1 所示。

表 1 6 条链路的初始状态
Table1 The initial states of the 6 lists

$\{r_0,r_6\}$	$\{r_1,r_7\}$	$\{r_2,r_8\}$	$\{r_3,r_9\}$	$\{r_4,r_{10}\}$	$\{r_5,r_{11}\}$
{332,613}	{104,784}	{588,675}	{698,459}	{10,216}	{140,308}

表1的结果与预设的一致，对6条链路继续使用算法，直到恢复出整个交织图样，如图5所示。图5与图3一致，说明了恢复出的交织器完全正确。

对不同误码率及不同的码字数目的情况进行仿真，其识别成功的概率如图6所示。由图知增加码字数目可增大识别的误码率。

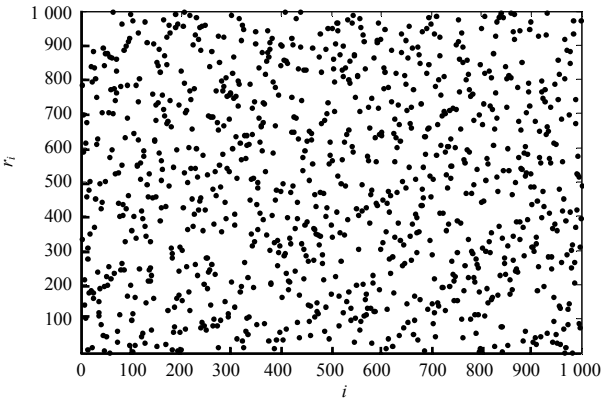

Fig.5 Recovered interleaved pattern
图5 恢复出的交织图样

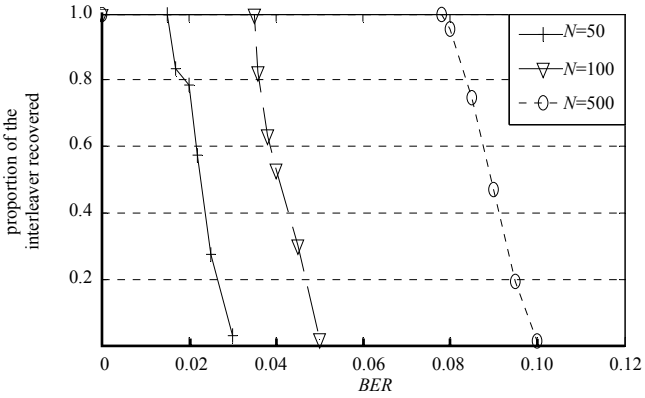

Fig.6 Proportion of the interleaver recovered at different BERs
图6 不同误码率下交织器的识别成功率

5 结论

本文通过将Turbo码中交织后的编码序列进行逆编码后可得到:同一周期点上的相邻2点与来自信息序列上的2点具有相关性，进而提出了一种针对非归零Turbo码交织器的识别方法。算法复杂度低，仿真显示，在较高的误码率下仍然有效，算法有一定的实用性。

参考文献:

[1] 解辉,黄知涛,王丰华. 信道编码盲识别技术研究进展[J]. 电子学报, 2013,41(6):1166-1176. (XIE Hui,HUANG Zhitao, WANG Fenghua. Research progress of blind recognition of channel coding[J]. Acta Electronica Sinica, 2013,41(6): 1166-1176.)

[2] 于沛东,李静,彭华. 一种利用软判决的信道编码识别新算法[J]. 电子学报, 2013,41(2):301-306. (YU Peidong,LI Jing, PENG Hua. A novel algorithm for channel coding recognition using soft decision[J]. Acta Electronica Sinica, 2013,41(2): 301-306.)

[3] 刘建成,杨晓静. 基于校验统计的(2,1,m)卷积码盲识别[J]. 电子信息对抗技术, 2013,28(1):1-4. (LIU Jiancheng,YANG Xiaojing. Blind recognition of(2,1,m) convolutional code based on parity-check statistics[J]. Electronic Information Warfare Technology, 2013,28(1):1-4.)

[4] Barbier J. Reconstruction of turbo-code encoders[C]// In SPIE Defense and Security Symp. Space Communications Technologies Conf., 2005:463-473.

[5] Cluzeau M,Finiasz M,Jean-Pierre Tillich. Methods for the reconstruction of parallel turbo codes[C]// In Proc. of the IEEE Int. Symp. Information Theory. Austin,Texas,USA:IEEE, 2010:2008-2012.

[6] Cˆote M,Sendrier N. Reconstruction of a turbo-code interleaver from noisy observation[C]// In Proc. of the IEEE Int. Symp. Information Theory. Austin,Texas,USA:IEEE, 2010:2003-2007.

[7] 张永光. 一种Trubo码编码参数的盲识别方法[J]. 西安电子科技大学学报, 2011(2):167-172. (ZHANG Yongguang. Blind recognition method for the Turbo coding parameter[J]. Journal of Xidian University, 2011(2):167-172.)

[8] 李啸天,李艳斌,咎俊军,等. 一种基于矩阵分析的 Turbo 码长识别算法[J]. 无线电工程, 2012,42(4):23-26. (LI Xiaotian, LI Yanbin,ZAN Junjun,et al. An algorithm for recognition of Turbo code length based on matrix analysis[J]. Radio Engineering of China, 2012,42(4):23-26.)

[9] 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, 2013,40(4):161-166. (LI Xiaotian,ZHANG Runsheng,LI Yanbin. Research on the recognition algorithm of Turbo codes on trellis termination[J]. Journal of Xidian University, 2013,40(4):161-166.)

(下转第 896 页)