

一种非规则卷积低密度校验码的构造和短环去除方法

马林华 常义林

(西安电子科技大学 综合业务网国家重点实验室 西安 710071)
(空军工程大学工程学院 航空电子工程系 西安 710038)

摘要 该文研究了使用为分组低密度校验(Low-Density Parity-Check, LDPC)码优化的度序列分布来构造卷积低密度(Low-Density Convolutional, LDC)码的方法, 详细讨论了LDC码的编码、译码和短环的消除算法, 实验结果说明用为分组LDPC码优化的非规则度序列分布所构造的LDC码, 其性能要优于目前文献上提出的规则(homogeneous)LDC码。

关键词 非规则低密度校验码, 卷积低密度校验码, 环路

中图分类号: TN911.22

文献标识码: A

文章编号: 1009-5896(2006)08-1408-04

Irregular Low-Density Convolutional Codes Construction and Short Cycles Elimination Method

Ma Lin-hua Chang Yi-lin

(State Key Lab. of Integrated Service Networks, Xidian Univ., Xi'an 710071, China)

(Aeronautical Electronics Engineering Dept., Air Force Univ. Engineering Institute, Xi'an 710038, China)

Abstract A method for constructing Low-Density Convolutional (LDC) codes with degree distributions optimized for block Low-Density Parity-Check (LDPC) codes is presented. The encoding, decoding, and short cycles elimination methods are discussed in detail. Experimental results show that the LDC codes constructed with the optimized irregular degree distributions for block LDPC codes are superior to homogeneous (i.e., regular) LDC codes.

Key words Irregular Low-Density Parity-Check (LDPC) codes, Low-Density Convolutional (LDC) codes, Cycles

1 引言

分组低密度校验(Low-Density Parity-Check, LDPC)码具有接近香农限的优良性质^[1]。但分组LDPC码需要把数据分成块进行编码后传送, 如果分组长度大, 将造成较大的时间延迟, 这在某些应用中是不利的。文献[2, 3]提出了规则(homogeneous)卷积低密度(Low-Density Convolutional, LDC)码, 它没有编码延迟, 可以进行连续数据传输, 并且其性能在相同运算复杂度下比分组LDPC码有较大的提高。本文给出一种LDC码的构造方法, 它基于为分组LDPC码优化的度序列分布, 如果度序列是非规则的, 则可构造出非规则LDC码。分组非规则LDPC码的性能被证明优于分组规则LDPC码^[1]; 我们发现, 非规则LDC码比规则LDC码的性能也有进一步的提高。本文组织如下, 第2节介绍LDC码的编码方法, 第3节介绍LDC码的译码方法, 第4节介绍LDC码短环的去除方法, 第5节是实验结果, 第6节是结束语。

2 LDC码的编码方法

给定规则或非规则的一个度序列分布, 分组LDPC码的校验矩阵可用随机方法生成^[1]。为方便说明问题, 我们用一

个简单的例子和图示来比较说明分组LDPC码和LDC码的编码方法。设 $H_{3 \times 6}$ 是码率 $R=1/2$ 的分组LDPC码的校验矩阵, 其统计特性符合一个优化的度序列分布(注意在实际应用中校验矩阵是很大的, 这里用一个小矩阵是为了方便说明问题)。用校验矩阵 $H_{3 \times 6}$ 对分组LDPC码进行编码, 则码的分组长度为6, 变量节点中分别有3个信息变量节点和3个冗余变量节点(这里我们沿用文献[1]中的叫法, 把一个分组中传输的数据叫变量节点(Variable nodes), 把变量节点参与的约束方程叫校验节点(Check nodes)或校验方程(Checks))。图1中, v_i^l 表示信息变量节点, v_j^r 表示编码时添加的冗余变量节点, 校验矩阵中 $h_{ij} \in (0,1)$, $h_{ij}=1$ 表示该分组的第 j 个变量节点参与了该分组的第 i 个校验方程, $h_{ij}=0$ 表示其对应的变量节点和校验方程没有关系。

对于规则LDPC码, 校验矩阵中每列 h_{ij} 等于1的数目相同, 每行 h_{ij} 等于1的数目也相同; 对于非规则LDPC码, 校验矩阵中每列 h_{ij} 等于1的数目差异可能较大, 每行 h_{ij} 等于1的数目也可能不同。图1(b)是图1(a)中第2个分组的3个校验方程(其它分组的校验方程只需把对应分组的变量节点代入即可)。分组LDPC码的编码工作就是通过校验方程组来求解冗余变量节点的值。

2005-04-21 收到, 2006-05-08 改回
国家自然科学基金(60172030), 华为科技基金和 ISN 国家重点实验室资助课题

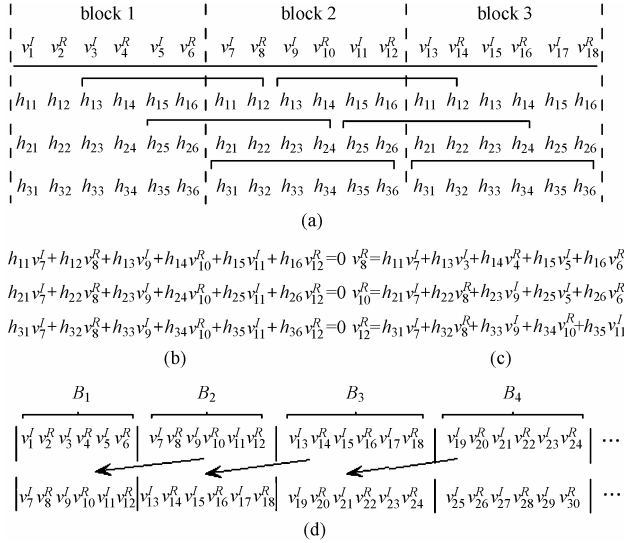


图 1 (a)为 LDC 码的结构
 (b)求解第 2 个分组中冗余变量节点的方程组
 (c)求解第 2 个分组位置上 LDC 码冗余变量节点
 (d) LDC 码的译码缓冲区

Fig.1 (a) Illustrates the LDC code structure

- (b) The equations to solve the redundant variable nodes of block 2
- (c) Computing the redundant variable nodes of block 2 for LDC codes
- (d) The decoding buffer for LDC codes

现在我们介绍使用上述分组LDPC码的校验矩阵 $H_{3 \times 6}$ 来构造LDC码的编码方法，它有 3 个条件：

(1) 在最初进行编码时，我们假定第 1 个分组的数据 $[v_1^r v_2^r v_3^r v_4^r v_5^r v_6^r]$ 对于发送和接收者都是已知的，我们从第 2 个分组 $[v_7^r v_8^r v_9^r v_{10}^r v_{11}^r v_{12}^r]$ 进行编码。

(2) 对于分组LDPC码，校验矩阵 $H_{3 \times 6}$ 可以进行随机生成^[1]，但对于LDC码，我们必须保证，一个分组中第 1(第 2、第 3)个冗余变量节点必须参与第 1(第 2、第 3)个校验方程，即， $h_{12} = 1 (h_{24} = 1, h_{36} = 1)$ 。满足了这个条件后，其它变量节点和校验方程的关系可以随机生成。

(3) 对于分组 LDPC 码，一个分组中所有校验方程中的变量节点的选取范围都被限制在该分组内，例如在图 1(a)中，第 2 个分组的校验方程的变量节点全部取之于第 2 个分组中的变量节点 $[v_7^r v_8^r v_9^r v_{10}^r v_{11}^r v_{12}^r]$ ，校验方程的形式如图 1(b)所示。为实现卷积码的连续输入输出编码，我们修改了分组 LDPC 码校验方程中变量节点的选择范围，其目的是使 1 个分组中第 1(第 2、第 3)个冗余变量节点成为该分组的第 1(第 2、第 3)个校验方程中的最后一个变量节点，如在图 1(a)中，横向方括号代表了 LDC 码校验方程中变量节点的选择范围，第 2 个分组中的第 1、第 2 和第 3 个校验方程，其变量节点的选择范围分别是 $[v_3^r v_4^r v_5^r v_6^r v_7^r v_8^r]$ ， $[v_5^r v_6^r v_7^r v_8^r v_9^r v_{10}^r]$ 和 $[v_7^r v_8^r v_9^r v_{10}^r v_{11}^r v_{12}^r]$ ，这样 v_8^r ， v_{10}^r 和 v_{12}^r 分别是第 2 个分组位置上 LDC 码的第 1，第 2 和第 3 个校验方程的最后一个变量节点；另外在修改 LDC 码中变量节点的选择范围时，必须满足一个条件，即 LDC 码和分组 LDPC 码的对应校验方程包含有相同的校验矩阵系数(比较式(1)一式(3)和图 1(b))，该条

件的意义将在下面进一步说明。

通过上面的 3 个条件，就可利用分组 LDPC 码的校验矩阵进行 LDC 码的编码，即实现卷积码的连续输入输出编码功能。编码的任务就是计算冗余变量节点值，根据构造 LDC 码的第 1 个条件，第 1 个分组的数据是已知的，所以不需要编码，我们从第 2 个分组开始编码；根据构造 LDC 码的第 2 和第 3 个条件，第 2 个分组的第 1，第 2 和第 3 个校验方程分别是：

$$h_{13}v_3^r + h_{14}v_4^r + h_{15}v_5^r + h_{16}v_6^r + h_{11}v_7^r + h_{12}v_8^r = 0 \quad (1)$$

$$h_{25}v_5^r + h_{26}v_6^r + h_{21}v_7^r + h_{22}v_8^r + h_{23}v_9^r + h_{24}v_{10}^r = 0 \quad (2)$$

$$h_{31}v_7^r + h_{32}v_8^r + h_{33}v_9^r + h_{34}v_{10}^r + h_{35}v_{11}^r + h_{36}v_{12}^r = 0 \quad (3)$$

根据构造 LDC 码的第 1 个条件，参看图 1(a)容易得出，在校验方程式(1)中，只有 v_8^r 是未知的，又根据构造 LDC 码的第 2 个条件， $h_{12} = 1$ ，即 v_8^r 参与校验方程式(1)，所以 v_8^r 可以被计算出来； v_8^r 被计算出来后，在校验方程式(2)中，同理分析可知只有 v_{10}^r 是未知的，因此可被立即计算出来； v_8^r 和 v_{10}^r 计算出来后，在校验方程式(3)中，同理可计算出 v_{12}^r ，在二进制编码中，其计算公式如图 1(c)，后续分组的冗余变量节点计算完全类似。

从以上 LDC 码冗余变量节点的计算可以看出，每个冗余变量节点的计算只与以前已出现的变量节点有关系，而与后续变量节点无关，因此可以被立即计算出来，从而实现了卷积码的连续输入输出编码。

虽然 LDC 码与 LDPC 码在编码时变量节点的选择范围不同，但是比较 LDC 码校验方程式(1)一式(3)与分组 LDPC 码的校验方程图 1(b)，可以看出校验方程中所使用的校验矩阵的系数完全相同，因此 LDC 码并没有改变分组 LDPC 码度序列分布的统计特性，第 5 节的实验结果说明，LDC 码同样具有较高的性能。

从 LDC 码的编码方法可以看出，该方法与给定的度序列是规则或非规则没有关系，即该方法既适应规则度序列分布也适应于非规则度序列分布。

对于其它度序列分布，完全可以用上述类似方法构造 LDC 码。

3 LDC 码的译码方法

LDC 码译码算法完全可以按照分组 LDPC 码译码算法进行。但由于 LDC 码校验方程中变量节点的选择范围不同，求解一个变量节点所需要的其它变量节点可能会分布在多个分组中，例如，图 1(a)中，若 v_9^r 参与第 1 和第 2 个校验方程，即 $h_{13} = 1$ 和 $h_{23} = 1$ ，那么求解 v_9^r 可能需要 $[v_5^r v_6^r \dots v_{14}^r]$ 中的变量节点，因此，在 LDC 码的译码过程中，需要一个译码缓冲区来存储多个分组长度的数据才能进行译码。

在图 1(d)，译码缓冲区的大小是 4 个分组长度，分别表示为 B_1, B_2, B_3 和 B_4 ，在最初译码开始时，根据构造 LDC 码的第 1 个条件， B_1 中的数据已知，因此不需要传送和译码，

当 B_2 , B_3 和 B_4 中的数据接收到后, 译码过程开始。由于 B_4 中的变量节点可能参与的校验方程所包含的数据还未接收到, 因此对 B_4 中的变量节点不进行译码, 它的存在只是为译码 B_3 中的数据, 这样只对 B_2 和 B_3 中的数据进行译码。知道每个校验方程中的变量节点和每个变量节点参与的校验方程, 就可以用与分组LDPC码一样的算法译码LDC码, 如和积译码算法^[1]。

当 B_2 中的数据被正确译码后(所有校验方程都满足), B_2 中的数据就输出, 同时被移入 B_1 , B_3 中的数据移入 B_2 , B_4 中的数据移入 B_3 , 新接收到的数据存入 B_4 , 然后重复上述的译码过程。

由于我们没有译码 B_4 中的数据, 则 B_4 中的错误数据会影响 B_3 中数据的正确译码, 如果因为 B_4 中的错误数据, 使得 B_3 中的数据没有正确译码, 则该错误数据也会影响 B_2 中数据的正确译码, 因此 B_4 中的错误数据有一个传播效应。当信道误码率较高时, 就必须增加译码缓冲区的大小, 使第 2 个分组的译码性能不会受到最后一个分组位置的错误数据传播效应的影响。

在实际中, 我们使用一个简单的方法来控制译码缓冲区的大小。最小译码缓存区设置为 4 个分组长度。设 N_{B_2} 和 N_{B_3} 分别表示 B_2 和 B_3 中译码后错误数据的数量, 在迭代译码次数超过预设的最大迭代次数后, 如果 $N_{B_2} > N_{B_3} > 0$, 我们认为 B_2 受到缓冲区中最后一个分组位置中错误数据扩散效应的影响, 因此给缓冲区增加一个分组长度; 如果 $N_{B_2} \approx N_{B_3}$, 则认为 B_2 没有受到缓冲区中最后一个分组中错误数据扩散效应的影响, 缓冲区的大小不变; 如果 $N_{B_2} = N_{B_3} = 0$, 则缓冲区可减小一个分组长度。

4 LDC 码中去除短环的方法

文献[4]提出了在构造分组 LDPC 码时去除短环的 PEG (Progressive Edge-Growth)算法, 它是根据变量节点和校验节点之间的连接关系来建立一个局部连接关系树来检测短环的。该算法应用于 LDC 码, 需要进行以下两点修正。(1) 变量节点和校验方程的连接关系要符合 LDC 码的构造条件(2)和条件(3); (2) 对于分组 LDPC 码, 从一个变量节点或校验节点开始的局部连接关系树总是限制在一个分组内, 在这个分组内随着局部连接关系树的生长, 所有长度的环路都可以被检测出来; 但对于 LDC 码, 由于其卷积结构, 某些局部连接关系树, 可能会横跨几个分组长度, 如果在建立局部关系树时, 对展开局部关系树的区域设得太小, 则局部关系树就不完整, 导致某些环路可能检测不到, 因此这里需要确定, 针对所分析的环路长度, 局部关系树需要在多大分组区域来展开建立局部关系树。

下面我们研究, 对于给定环路长度, 环路中的节点所构成的关系树在最大可能情况下所占据的分组区域。图 2 中, 粗横线代表一个校验方程中变量节点所占的最大区域, 即校

验节点中最左边的一个变量节点和最右边的一个变量节点之间的区域, 实心圆点代表一个变量节点, 通过实心圆点的垂直线连接两个校验方程代表这两个方程拥有一个共同的变量节点。由于LDC码各分组长度内的结构相同, 因此只需研究分组 B_N 的情况。图 2 中, C_1 代表 B_N 中第 1 个校验方程(最左边的校验方程), 其大部分变量节点在前一个分组区间 B_{N-1} ; C_2 代表 B_N 中最后一个校验方程(最右边的校验方程), 其所有变量节点都在分组区间 B_N 。因此从 C_1 生长出的局部关系树, 其包含的环路可以决定我们要确定的分组区域的左边界, 而从 C_2 生长出的局部关系树, 其包含的环路可以决定我们要确定的分组区域的右边界。为确定可能的最大分组区域, 在图 2 中, 假定所有校验方程的拥有最大长度, 即一个分组长度。

在图 2(a), 变量节点(V_1, V_3) 参与校验方程(C_1, C_3), 组成环路 4, 容易看出(V_1, V_3)在最左端不可能超过 B_{N-1} ; 同样地, (V_2, V_4) 组成环路 4, 在最右端不可能超过 B_N 。所以, 给定环路长度为 4, 环路中节点构成的局部关系树所占的最大分组区域是 $[B_{N-1}, B_N]$, 即把局部关系树在 $[B_{N-1}, B_N]$ 中展开, 则所有长度为 4 的环路都可以被检测到。

在图 2(b), 变量节点(V_1, V_3, V_5) 构成长度为 6 的环, 由于(V_1, V_3)参与 C_1 , 因此(V_1, V_5) 在最左端不可能超过 B_{N-1} , 又 V_1 和 V_3 的距离小于一个分组长度, 因此 V_5 在最左端不超过 B_{N-2} ; 同样地, (V_2, V_4, V_6) 构成长度为 6 的环, 在最右端不可能超过 B_{N+1} 。给定环路长度为 6, 环路中节点构成的局部关系树所占的最大分组区域是 $[B_{N-2}, B_{N+1}]$ 。

在图 2(c), (V_1, V_3, V_5, V_7)和(V_2, V_4, V_6, V_8) 构成长度为 8 的环, 类似于分析图 2(b), 可得, 给定环路长度为 8, 环路中节点构成的局部关系树所占的最大分组区域是 $[B_{N-2}, B_{N+1}]$ 。

使用上面相同的方法, 可得到更大环路长度下, 其对应所覆盖的最大分组区域。我们已经得到如果要检测环路长度

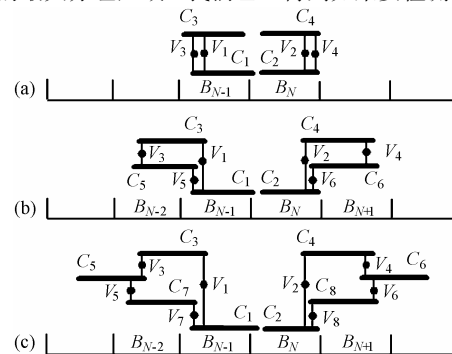


图 2 环中的节点所构成的局部关系树与其可能占据的最大分组区域的关系 (a), (b)和(c)是环长分别为 4, 6 和 8 时,

环中节点所形成的关系树可能占据的最大分组区域

Fig.2 The relation between the length of cycles and the maximum block range that the cycles may spread over (a), (b), and (c) are the maximum block ranges for the cycles of length 4, 6, and 8, respectively

为 4((6,8), (10,12), (14,16)), 局部关系树的展开分组区域最小要设为 $[B_{N-1}, B_N]([B_{N-2}, B_{N+1}], [B_{N-3}, B_{N+1}], [B_{N-4}, B_{N+3}])$ 。且局部关系树生长到 L 层(这里校验节点和其包含的变量节点合起来为一“层”), 长度为 $2 \times L$ 的环路可以被检测到。

考虑到 LDC 码的结构和环路长度与局部关系树展开区域的关系, 则 PEG 算法也可以用于 LDC 码去除短环, 并且这种消除短环的算法对于文献[2, 3]提出的规则 LDC 码同样适用。

5 实验结果

$$\left. \begin{aligned} \lambda(x) &= 0.27684x + 0.28342x^2 + 0.43974x^8 \\ \rho(x) &= 0.01568x^5 + 0.85244x^6 + 0.13188x^7 \end{aligned} \right\} \quad (4)$$

根据度序列式(4), 我们随机构造分组 LDPC 码和 LDC 码(1024,512), 它们中长度为 4 的环路被去除。在比较环路数量时, 我们仅计算了由度 2 和度 3 变量节点所组成环路数量, 因为低度数的变量节点得到正确值的概率低, 仅由它们组成的环路危害性较大, 是影响码的性能的主要因素。表 1 对这两类码的不同长度环路数量进行了比较, 可以看出 LDC 码的环路数量少于分组 LDPC 码。

表 1 在分组 LDPC 码和 LDC 码(1024,512)中, 由度 2 和度 3 变量节点所构成环路的数量

Tab.1 The numbers of the cycles that pass through degree 2 and degree 3 nodes of the block LDPC codes (1024, 512) and the LDC codes (1024, 512)

环长度	4	6	8	10	12	14	16
分组码	0	25	96	338	1357	5786	25396
卷积码	0	14	62	216	941	4185	18397

图 3 显示了低密度校验码在二进制相移键控(BPSK)调制, 二进制输入加性高斯白噪声(BIAWGN)信道下的性能比较。码 I 是度序列式(4)所构造的分组 LDPC 码(1000,500)^[1]的性能; 码 II($m_s=1025, j=3(2), k=5$)^[2]和码 III($m_s=1025, j=3, k=6$)^[3]是规则 LDC 码的性能, 其中 m_s 是校验方程中变量节点的最大距离, j 是变量节点的度数, k 是校验节点的度数; 码 IV 是根据度序列式(4)所构造的非规则 LDC 码(1024,512)的性能。可以看出, 卷积低密度校验码的性能优于分组 LDPC 码, 而非规则 LDC 码的性能优于规则 LDC 码, 在 BER=10⁻⁵ 时, 非规则 LDC 码比规则 LDC 码 III 有 0.15dB 的增益。

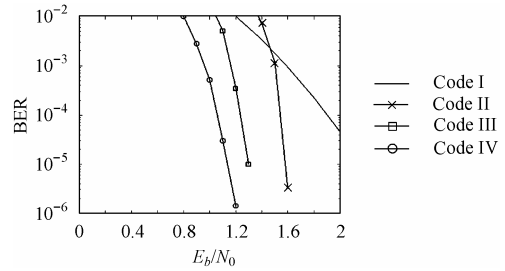


图 3 在 BIAWGN 信道 BPSK 调制下, 码 I、II、III 和 IV 的误比特率性能曲线。码 I 为分组 LDPC 码(1000,500); 码 II(1025,3(2),5)和码 III(1025,3,6)是规则 LDC 码; 码 IV(1024,512)是非规则 LDC 码
Fig.3 Bit error performances of code I, II, III and IV over BIAWGN channels with BPSK modulation. Code I(1000, 500) is the block LDPC codes. Code II(1025, 3(2), 5) and code III(1025, 3, 6) are the homogenous LDC codes. Code IV(1024, 512) is the irregular LDC codes

6 结束语

根据 LDC 码的结构, 即 LDC 码的构造条件(2)和条件(3), 使用为分组 LDPC 码优化的序列度分布, 可以构造出一类 LDC 码。实验结果显示, 非规则 LDC 码可进一步提高码的性能。它的缺点是带来了译码延迟, 但当信道 BER 较低时, 可通过动态调整译码缓冲区来尽可能地减少译码延迟。

参考文献

- [1] Richardson T, Shokrollahi M, Urbanke R. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans. on Info. Theory*, 2001, 47(2): 619-637.
- [2] Felstrom A J, Zigangirov K S. Time-varying periodic convolutional codes with low-density parity-check matrix. *IEEE Trans. on Info. Theory*, 1999, 45(6): 2181-2191.
- [3] Lentmaier M, Truhachev D, Zigangirov K S. On the theory of low-density convolutional codes II. *Problems of Information Transmission*, 2001, 37(4): 288-306.
- [4] Hu X, Eleftheriou E, Arnold D M. Progressive edge-growth tanner graphs. *GLOBECOM 2001- IEEE Global Telecommunications Conference, Texas, USA, IEEE, 2001, no. 1: 995-2001.*

马林华: 男, 1965 年生, 副教授, 主要研究方向为图像通信和编码调制。
常义林: 男, 1944 年生, 教授, 主要研究方向为无线通信、网络管理。