

多阶段决策方法用于6:8优化映射等重调制码的设计

陈端荣^{1),2)} 谢长生²⁾ 裴先登²⁾ 肖卫军²⁾

¹⁾(湖北大学数学与计算机科学学院 武汉 430062)

²⁾(华中科技大学计算机科学与技术学院 武汉 430074)

摘 要 6:8等重调制码是全息存储记录通道中的一种有效通道码.然而,各种各样的噪声源和符号间串扰极易使检测到的一个等重码字与原等重码字相比发生一个0与1之间的2码元交换错误事件(2-BREE).对于6:8二进制顺序映射等重调制码(NCWMC)而言,由此导致在一个6比特数据字中平均产生约3比特误码.作者借助多阶段决策方法和贪心算法,设计出一种6:8优化映射等重调制码(OCWMC).6:8 OCWMC与6:8 NCWMC相比,一个2-BREE所致的平均误比特数下降了0.5844比特.文章以4:6 OCWMC的设计为例,详细介绍了其设计过程.误码性能的仿真实验结果表明:6:8 OCWMC比6:8 NCWMC更适宜应用于具有严重的符号间串扰和噪声的全息存储记录通道.

关键词 全息存储;等重调制码;多阶段决策;贪心法;符号间串扰

中图法分类号 TP333

Design of 6:8 Optimal-mapping Constant Weight Modulation Code by Multistage Decision Method

CHEN Duan-Rong^{1),2)} XIE Chang-Sheng²⁾ PEI Xian-Deng²⁾ XIAO Wei-Jun²⁾

¹⁾(School of Mathematics and Computer Science, Hubei University, Wuhan 430062)

²⁾(School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074)

Abstract The 6:8 constant weight modulation code (CWMC) is an effective channel code for the holographic storage channel. However, because of various noise source and inter-symbol interference (ISI), one 2-bit reversing error event (2-BREE) between the two digits of a 1 and a 0 is easy to occur in a detected codeword compared with an original modulation codeword, which causes about a 3-bit error on the average in a 6-bit dataword for the 6:8 natural binary order mapping constant weight modulation code (NCWMC). Therefore, a 6:8 optimal mapping constant weight modulation code (OCWMC) is designed by the multistage decision method and the greedy algorithm. Compared with the 6:8 NCWMC, the average bit-error number for one 2-BREE of the 6:8 OCWMC can be reduced 0.5884 bits. As an example, the authors introduce the design of the 4:6 OCWMC in detail. Simulation results for the bit-error rate performance show that the 6:8 OCWMC is more adapted for the holographic storage channel with severe ISI and noise.

Keywords holographic storage; constant weight modulation code; multistage decision; greedy algorithm; inter-symbol interference

收稿日期:2003-05-17;修改稿收到日期:2004-12-07.本课题得到国家“九七三”重点基础研究发展规划项目基金(1999033006)资助.
陈端荣,男,1960年生,博士,副教授,研究方向为全息数据存储记录通道中的编译码技术. E-mail: duanrongchen@hubu.edu.cn. 谢长生,男,1957年生,教授,博士生导师,研究领域为网络存储和高密度光存储技术. 裴先登,男,1934年生,教授,博士生导师,研究领域为高密度光存储技术. 肖卫军,男,1974年生,博士,讲师,研究方向为计算技术和数据库技术.

1 引言

数字全息数据存储(Digital Holographic Data Storage, DHDS)是一种基于页的光存储技术,由于同时具有高速的数据传输速率(1Gbits/s)、巨大的存储容量(V/λ^3 , V 是全息存储晶体的体积, λ 是光源的波长)和短暂的访问响应时间($<100\mu\text{s}$),极有可能成为人们使用高速度、大容量存储器的一种选择^[1]. 由于 DHDS 记录通道是一个有噪声的数据传输通道,为了使 DHDS 系统能满足商品化的存储系统对误码率(Bit Error Rate, BER)的要求,必须采用纠错码和调制码技术^[2]. 6:8 等重调制码(Constant Weight Modulation Code, CWMC)是 DHDS 记录通道中的一种有效通道码,它是将每 6 比特数据信息(或称为一个数据字)映射为一个相应的 8 比特等重码字(0 和 1 的数量均为 4 个). 由于 6:8 CWMC 具有码率高、译码速度快(使用维特比算法检测一个码字仅需要 13 次加和 13 次比较操作)的特点^[3,4],因此,使用了 6:8 CWMC 的 DHDS 系统的数据传输率可达到 1Gbits/s^[5].

但是,6:8 CWMC 中的 8 比特的等重码的汉明距离为 2,在具有各种各样噪声源的 DHDS 记录通道中,尤其是在系统的信噪比 SNR(Signal-to-Noise Ratio)较低时^[3,6],检测到的一个 8 比特等重码字与原 8 比特等重码字相比较,极易发生一个数字 0 与数字 1 的 2 码元交换错误事件(2-Bit Reversing Error Event, 2-BREE),对于 6 比特数据字和 8 比特码字均按自然二进制顺序排序,然后依次形成映射关系的 6:8 CWMC(Natural binary order mapping CWMC, NCWMC)而言,这将使 6:8 NCWMC 中的一个 6 比特数据字平均产生约 3 比特错误(注:本文所讨论的误比特数均是由 2-BREE 所致),导致在某一信噪比范围不能使系统的最低原始误码率达到小于等于 10^{-3} 的要求. IBM 公司的全息研究人员设计了一种码距大于或等于 4 的 8:12 CWMC 来抑制这种 2-BREE 的发生^[3,4],由于该 8:12 CWMC 是一种有限状态码,其编码和译码的复杂性使得 DHDS 系统难以用硬件实现 1 Gbits/s 的数据传输率. 我们则从降低 2-BREE 所导致的平均误比特数和减少 2-BREE 的发生率这两个方面入手,分别优化了 6:8 CWMC 的编码方法和改进了现有的维特比检测方案^[7].

由于 6:8 NCWMC 的 64 对映射关系的建立

过程是“同时”进行的,并没有考虑每一对映射关系之间的相互影响,当使用 6:8 NCWMC 时,一个 2-BREE 发生后就会产生约 3 比特错误. 经过分析,我们认为:在 6:8 CWMC 的 64 对 6 比特数据字与 8 比特等重码字之间建立最优映射关系,设计出 6:8 优化映射等重调制码(Optimal mapping CWMC, OCWMC)是降低由 2-BREE 所导致的平均误比特数的关键. 如果使用穷尽法设计 6:8 OCWMC,需要评估 64! 种方案,显而易见目前是无法实现的. 因此,我们将 6:8 CWMC 映射关系的优化问题视为一个多阶段决策问题,将 64 对映射关系的建立过程改为“分时”进行,并使用贪心算法设计出了一种 6:8 OCWMC.

在设计 6:8 OCWMC 时,需要使用 64×64 矩阵,而设计 4:6 OCWMC 只需要使用 16×16 矩阵,鉴于 6:8 OCWMC 的设计过程仅仅是 4:6 OCWMC 的设计过程的推广,因此在本文的第 2 节先介绍了求解多阶段决策问题的方法,然后根据 4:6 NCWMC 给出了其码距分布矩阵和 2 码元交换关联分布矩阵,并作了相应的解释;借助第 2 节给出的几个 16×16 矩阵,在第 3 节提出了设计 OCWMC 的数学模型;在第 4 节详细介绍了使用贪心法设计 4:6 OCWMC 的过程;第 5 节则给出了使用 4:6 OCWMC 和 6:8 OCWMC 的仿真实验结果和性能分析;文章的结语部分则对为设计 6:8 OCWMC 所提出的方法的潜在适用范围给予了说明.

2 多阶段决策问题及 NCWMC 的矩阵

多阶段决策问题^[8]是指决策的全过程可划分为若干个互相联系的阶段,在它的每一个阶段都需要作出决策,上一个阶段的决策常常影响到下一个阶段的决策,即各个阶段决策的选取不是任意的,它依赖于当前面临的状态,又影响以后的决策,图 1 给出了多阶段决策的示意图. 在图 1 中, k 为阶段变量, (x_k, y_k) 为第 k 阶段的决策变量, s_k 为第 k 阶段的状态变量, $v_k(s_k, x_k, y_k)$ 为第 k 阶段的效益函数. 各个阶段的决策组成一个决策序列,也称之为策略. 由于每一个阶段都有若干个决策可共选择,因此就有许多策略可供选取. 由于选取的策略不同,所导致的效果一般也不尽相同. 如何在众多的可供选取的策略中,选取一个最优策略,使其能在一个确定的目标下,获得最好的效果,这类问题就是多阶段决策问题. 对于 6:8 CWMC 的设计问题,64 个 6 比特数

据字与 64 个 8 比特等重码字形成的映射关系不同, 所导致的各种误比特数目的分布也不同, 当把这 64 对映射关系视为是逐一形成时, 就可视为一个多阶段决策问题.

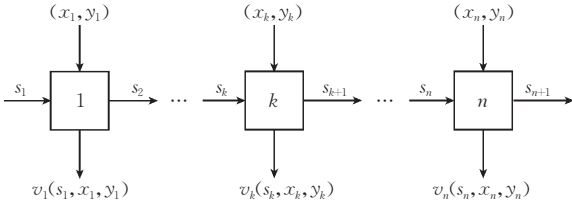


图 1 多阶段决策过程示意图

表 1 列出了 4 : 6 NCWMC 码表和本文利用多阶段决策方法设计的 4 : 6 OCWMC 码表, 表 1 中的第 4 列和第 7 列的十进制数据分别与 4 : 6 NCWMC 和 4 : 6 OCWMC 的 6 比特等重码字相对应.

表 1 4 : 6 NCWMC 和 4 : 6 OCWMC 码表

序号	4 : 6 NCWMC			4 : 6 OCWMC		
1	0000	001101	13	0000	011100	28
2	0001	001110	14	0001	001110	14
3	0010	010101	21	0010	011001	25
4	0011	010110	22	0011	001101	13
5	0100	011001	25	0100	011010	26
6	0101	011010	26	0101	101100	44
7	0110	011100	28	0110	111000	56

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	1	1	2	1	2	2	3	1	2	2	3	2	3	3	4
2	1	0	2	1	2	1	3	2	2	1	3	2	3	2	4	3
3	1	2	0	1	2	3	1	2	2	3	1	2	3	4	2	3
4	2	1	1	0	3	2	2	1	3	2	2	1	4	3	3	2
5	1	2	2	3	0	1	1	2	2	3	3	4	1	2	2	3
6	2	1	3	2	1	0	2	1	3	2	4	3	2	1	3	2
7	2	3	1	2	1	2	0	1	3	4	2	3	2	3	1	2
8	3	2	2	1	2	1	1	0	4	3	3	2	3	2	2	1
9	1	2	2	3	2	3	3	4	0	1	1	2	1	2	2	3
10	2	1	3	2	3	2	4	3	1	0	2	1	2	1	3	2
11	2	3	1	2	3	4	2	3	1	2	0	1	2	3	1	2
12	3	2	2	1	4	3	3	2	2	1	1	0	3	2	2	1
13	2	3	3	4	1	2	2	3	1	2	2	3	0	1	1	2
14	3	2	4	3	2	1	3	2	2	1	3	2	1	0	2	1
15	3	4	2	3	2	3	1	2	2	3	1	2	1	2	0	1
16	4	3	3	2	3	2	2	1	3	2	2	1	2	1	1	0

(a) 4 比特码字的码距分布矩阵 **B**

序号	4 : 6 NCWMC			4 : 6 OCWMC		
8	0111	100101	37	0111	101001	41
9	1000	100110	38	1000	010110	22
10	1001	101001	41	1001	100110	38
11	1010	101010	42	1010	010101	21
12	1011	101100	44	1011	100101	37
13	1100	110001	49	1100	110100	52
14	1101	110010	50	1101	101010	42
15	1110	110100	52	1110	110010	50
16	1111	111000	56	1111	110001	49

(续 表)

图 2(a)是根据表 1 中的 4 : 6 NCWMC 的 16 个 4 比特数据字(以下为叙述方便, 有时也称之为 4 比特码字)之间的码距产生的一个 16 × 16 的码距分布矩阵 **B**; 在 **B** 矩阵中, 主对角线上的 16 个 0 是该 4 比特码字在 4 : 6 NCWMC 码表中所处的位置或者认为是自身的码距(即 $b_{i,j} = 0, 1 \leq i = j \leq 16$); 每一行上的数值 1~4 是对应该行的 4 比特码字与另外 15 个 4 比特码字之间的码距分布. 例如, $b_{1,16} = b_{16,1} = 4$, 表示在 4 : 6 NCWMC 码表中序号为 1 的 4 比特码字与序号为 16 的 4 比特码字之间的码距是 4, 本文将一个码表中具有最大码距的一对码字称之为“对偶”码字对.

图 2(b)则是根据表 1 中的 4 : 6 NCWMC 的

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	1	1	0	1	0	1	1	0	1	0	1	0	0	0	0
2	1	0	0	1	0	1	1	0	1	0	1	1	0	0	0	0
3	1	0	0	1	0	1	1	0	0	0	0	1	0	1	0	0
4	0	1	1	0	0	1	1	0	1	0	0	0	0	1	1	0
5	1	0	1	0	0	1	1	0	0	1	0	0	1	0	0	1
6	0	1	0	1	1	0	1	0	0	0	1	0	0	1	0	1
7	1	1	1	1	1	0	0	0	0	0	1	0	0	1	1	0
8	1	0	1	0	0	0	0	0	1	1	0	1	1	0	1	0
9	0	1	0	1	0	0	0	1	0	0	1	1	0	1	1	0
10	1	0	0	0	1	0	0	1	0	0	1	1	1	0	0	1
11	0	1	0	0	0	1	0	0	1	1	0	1	0	1	0	1
12	1	1	0	0	0	0	1	1	1	1	1	0	0	0	1	1
13	0	0	1	0	1	0	0	1	0	1	0	0	0	1	1	1
14	0	0	0	1	0	1	0	0	1	0	1	0	1	0	1	1
15	0	0	1	1	0	0	1	1	1	0	0	1	1	1	0	1
16	0	0	0	1	1	1	0	0	1	1	1	1	1	1	1	0

(b) 6 比特等重码字的码字关联分布矩阵 **M**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	13	12	13	13	13	15	17	12	17	15	16	18	17	16	21	20
2	12	13	16	14	14	14	16	15	14	16	11	19	16	17	20	21
3	12	11	12	14	14	14	16	13	14	14	19	17	16	15	18	21
4	11	12	15	15	15	13	15	16	11	15	14	18	15	16	17	22
5	14	15	10	12	12	14	18	15	18	14	17	19	14	15	20	17
6	13	16	13	13	13	13	17	18	15	15	12	20	13	16	19	18
7	13	14	9	13	13	13	17	16	15	13	20	18	13	14	17	18
8	12	15	12	14	14	12	16	19	12	14	15	19	12	15	16	19
9	16	13	16	14	14	16	20	9	16	14	13	17	16	13	20	17
10	15	14	19	15	15	15	19	12	13	15	8	18	15	14	19	18
11	15	12	15	15	15	15	19	10	13	13	16	16	15	12	17	18
12	14	13	18	16	16	14	18	13	10	14	11	17	14	13	16	19
13	17	16	13	13	13	15	21	12	17	13	14	18	13	12	19	14
14	16	17	16	14	14	14	20	15	14	14	9	19	12	13	18	15
15	16	15	12	14	14	14	20	13	14	12	17	17	12	11	16	15
16	15	16	15	15	15	13	19	16	11	13	12	18	11	12	15	16

(c) 矩阵 **Z** 为矩阵 **B** 与 **M'** 之积

16个6比特等重码字之间是否存在2码元交换关系而产生的一个 16×16 的码字关联分布矩阵 \mathbf{M} .在 \mathbf{M} 矩阵中,主对角线上的0(即 $m_{i,j}=0, 1 \leq i=j \leq 16$)只表示该码字在码表中的位置(或者理解为第 i 个码字自身的2码元交换关系);不在主对角线上的0则表示与该0元素有关的2个等重码字间不存在2码元交换关系,例如 $m_{1,14}=m_{14,1}=0$,表示在4:6 NCWMC码表中序号为1的等重码字与序号为14的等重码字之间不存在2码元交换关系;而 $m_{1,7}=m_{7,1}=1$ 则表示序号为1的等重码字与序号为7的等重码字之间存在2码元交换关系.

图2(c)为矩阵 \mathbf{B} 与矩阵 \mathbf{M} (或者是 \mathbf{M} 的转置矩阵 \mathbf{M}' ,因为 \mathbf{B} 和 \mathbf{M} 都是对称矩阵)之积得到的矩阵 $\mathbf{Z}=\mathbf{B} \times \mathbf{M}'$.在矩阵 \mathbf{Z} 的主对角线上的一个 $z_{i,j}$ ($1 \leq i=j \leq 16$)是4:6 NCWMC的第 i 个等重码字的7次或9次有效的2-BREE所导致的误比特数之和,所谓有效的2-BREE是因为该码字集中还存在禁用的等重码字,此外由 \mathbf{M} 矩阵可见在12行中有7个 $m_{i,j}=1$ 的元素,另外的4行包含有9个 $m_{i,j}=1$ 的元素;而矩阵 \mathbf{Z} 的主对角线上的16个元素 $z_{i,i}$ 的和($1 \leq i=j \leq 16$)是第3节的表2所列出的在4:6 NCWMC中由有效的2-BREE所导致的总误比特数TBEN(Total Bit Error Number).

3 设计优化映射等重调制码的数学模型

在介绍设计OCWMC的数学模型之前,用表2一并列出了4:6和6:8 NCWMC,OCWMC和ICWMC关于2-BREE的几项统计数据,表2中的ICWMC是理想映射的CWMC(ideal mapping CWMC),总误比特数TBEN是相应的CWMC对于所有有效2-BREE的误比特数之和;表中的 N, O 和 I 分别表示NCWMC,OCWMC和ICWMC.

表2 4:6 CWMC和6:8 CWMC的2-BREE的统计数据

等重调制码	1~4(6)误比特数分布						总误比特数	平均误比特数
	1	2	3	4	5	6		
4:6 N	40	48	28	4			236	1.9667
4:6 O	54	56	10	0			196	1.6333
4:6 I	64	56	0	0			176	1.4667
6:8 N	126	278	312	164	64	4	2618	2.7616
6:8 O	252	376	236	68	16	0	2064	2.1772
6:8 I	384	564	0	0	0	0	1512	1.5949

本文中的NCWMC,OCWMC和ICWMC的码字集合是完全一样的,而选择码字集的原则是为了便于使用维特比检测方法.因此,对于3种CWMC

而言,它们各自的等重码字之间具有2码距关系的总数是相同的,即能发生有效的2-BREE的总数是相同的.在本文的4:6 CWMC和6:8 CWMC中,有效的2-BREE的总数分别为120次和948次.

此外,对于4:6 CWMC而言,由组合公式可以计算出,也可以从图2(a)的 \mathbf{B} 矩阵的任意一行的数据统计出一个4比特码字与另外15个4比特码字之间的码距为1,2,3,4的数目分布,它们分别是4,6,4和1个;而16个4比特码字之间的码距为1至4的数目分布的总数则是64,96,64和16个.因此,对于4:6 ICWMC的120次2-BREE而言,它是由64个1bits误码次数和56个2bits的误码次数组成,由此得到理想的总误比特数TBEN是176bits,它是由12个10bits(即 $1 \times 4 + 2 \times 3$)和4个14bits(即 $1 \times 4 + 2 \times 5$)组成.尽管我们现在还无法知道4:6 ICWMC的映射关系(也许它实际上根本不存在),但是从上述的推导过程中可以得出4:6 ICWMC的120次2-BREE的总误比特数目及误比特数目的理想分布,这是我们设计4:6 OCWMC所追求的目标.同理,表2列出了6:8 ICWMC的948次2-BREE的总误比特数目及误比特数目的分布.

OCWMC的设计是一个多目标线性规划问题,设计4:6 OCWMC的多阶段决策问题的数学模型为(注:设计6:8 OCWMC时,图1中的 $n=64$):

$$\min f(\mathbf{Z}) = \sum_{k=1}^n Z_k = \sum_{k=1}^{16} \sum_{u=1}^{16} b_{x_k, u} \cdot m_{u, y_k}$$

$$\text{s. t.} \begin{cases} 1 \leq x_k \leq 16, & 1 \leq k \leq 16 \\ 1 \leq y_k \leq 16, \\ x_{k_i} \neq x_{k_j}, & k_i \in k, k_j \in k, k_i \neq k_j \\ y_{k_i} \neq y_{k_j}, & 1 \leq i \leq 16, 1 \leq j \leq 16 \end{cases} \quad (1)$$

$\min f(\mathbf{Z})$ 表示使目标函数 $f(\mathbf{Z})$ 极小化,也就是使设计出的4:6 OCWMC所给出的矩阵 \mathbf{Z} 的主对角线上的16个元素 Z_k ($1 \leq k \leq 16$)的和取极小值,对于4:6 ICWMC而言,理想的16个 Z_k 的值的分布是12个 Z_k 的值为10,另外4个 Z_k 的值是14.因此,理想的总误比特数TBEN的数值为176.值得说明的是,图2中的矩阵 \mathbf{Z} 的主对角线上的16元素 $z_{i,j}$ ($1 \leq i=j \leq 16$)与式(1)中的16个 Z_k ($1 \leq k \leq 16$)所表示的含义相同,而式(1)中的 \mathbf{Z} 是16个 Z_k 的一个抽象表达;另外,为了简洁,图2中的矩阵 \mathbf{Z} 与式(1)中的 \mathbf{Z} 使用了相同的符号.

在式(1)中的 b 和 m 分别是矩阵 \mathbf{B} 和矩阵 \mathbf{M}

的元素, 它们的下标 x_k 和 y_k 是第 k 个阶段时的 2 个决策变量, (x_k, y_k) 在此表示一个决策对, 其含义为原来在 4:6 NCWMC 码表中序号为 x_k 的 4 比特码字与序号为 y_k 的 6 比特等重码字形成映射关系。

另外, 值得说明的是, 式(1)中的数学模型只是用来寻求 4:6 OCWMC 的形式表述, 而各个阶段的效益函数 $v_k(s_k, x_k, y_k)$ 才是 4:6 OCWMC 设计过程中的真正实施者, 实际上 $v_k(s_k, x_k, y_k)$ 与 Z_k 的表述有关. 关于 s_k 和 $v_k(s_k, x_k, y_k)$ 在设计 4:6 OCWMC 时所表示的意义, 将在第 4 节中介绍。

简而言之, 对于实现目标函数 $f(\mathbf{Z})$ 的具体要求是在式(1)中用 s. t. (即 subject to 的缩写) 表示的限制条件下, 使总数为 120 次 2-BREE 所导致的总误比特数最小, 这只有通过每一次 2-BREE 所导致的误比特数最小来实现该目标。

4 贪心算法设计 OCWMC 的过程

贪心算法(Greedy Algorithm)^[9] 也称之为登山法, 是一种最简单、最直接的算法设计技术, 可以求出很多组合最优化问题的最优解. 用贪心算法求解某一问题时, 往往面对多个输入, 要求根据约束条件(选择攀登原则)从对问题的某一初始推测或初始解出发, 一步一步地“攀登”给定的目标, 尽可能快地去逼近更好的解. 由此可见, 贪心算法是一个多步决策过程(即多阶段决策过程), 每一步的选择都是为了能构成问题的一个可行解, 而且当前的选择依赖于已经作出的所有选择所产生的状态, 对于最小化问题, 其选择的目的是要使目标函数的值增加最少. 选择过程是以某些最优化量度为依据的, 最优化的量度就是图 1 中各个阶段的效益函数. 用贪心算法求解最小化问题时, 往往总是先选择耗费最小的元素加入解集合, 每一步都试图取得局部最优解. 尽管贪心算法具有求解速度快, 时间复杂性具有较低阶的优点, 但是它的缺点是一般找不到最优解, 因为它的致命弱点是后进入解向量的元素一般都有较大的耗费。

因此, 在设计 OCWMC 时, 规定有“对偶”关系的一对 4 比特码字必须相继完成“配对”(即建立与 6 比特等重码字的映射关系); 另外随着攀登的高度增加, 逐渐降低贪心度(即适度贪心), 避免用登山法(贪心法)在快要接近山的顶峰时代价甚高(4 比特误码数目增加). 这样既发挥了贪心法的优点, 又尽量避免了它的弱点. 用贪心法设计 4:6 OCWMC 的步骤如下。

1. 确定在表 1 中的 4:6 NCWMC 的 8 对“对偶”码字的序号: $x_k \in \{1, 16; 2, 15; 3, 14; 4, 13; 5, 12; 6, 11; 7, 10; 8, 9\}$, 以便根据所选定的 x_k 对 y_k 作出决策, $y_k \in \{1, 2, 3, \dots, 14, 15, 16\}$, $1 \leq k \leq 16$;

2. 当 $k=1$ 时, 给出一个初始决策: $(1, 7)$, 即 $x_1=1$ 时, $y_1=7$;

3. 当 $1 < k \leq 16$, 对于剩余的而且是有效的 $n-k+1$ 个 x_k 与 $n-k+1$ 个有效的 y_k 进行组合后, 分别计算如下的效益函数 $v_k(s_k, x_k, y_k)$ (注: 此处的 n 为 16, “有效”的含义是指还未建立映射关系, 即作出“配对”选择):

$$v_k(s_k, x_k, y_k) = \sum_{j=1}^{k-1} b_{x_k, x_j} \cdot m_{y_k, y_j}, \quad 1 < k \leq 16 \quad (2)$$

4. 从 $(n-k+1) \times (n-k+1)$ 个效益函数 $v_k(s_k, x_k, y_k)$ 的计算值中选取最优值(但并不一定是最小值), 此时的 x_k 和 y_k 就为该阶段的一个决策对, 若 $k < 16$, 则转步 3, 而 $k=16$, 则是最后一个决策对。

式(2)中的效益函数 $v_k(s_k, x_k, y_k)$ 的物理意义是在第 k 个阶段作出决策 (x_k, y_k) 之后的序号为 x_k 的 4 比特码字与 $k-1$ 个已作出映射的 4 比特码字在 2 码元交换发生的条件下的误比特数之和; 显而易见, $v_k(s_k, x_k, y_k)$ 是越小越好, 当然其和的各个分量也要求越小越好, 即 1 比特误码的个数越多越好. 由于 $k=1$ 之前没有对任何 4 比特码字与 6 比特等重码字建立映射关系, 因此, $v_1(s_1, x_1, y_1) = 0$ 。

表 3 列出了实施上述算法的 16 对决策下的误比特数目分布(实际上是阶段效益函数 $v_k(s_k, x_k, y_k)$) 及阶段平均误比特数(Stage Average Bit-Error Number, SABEN), 同时给出了由此决策序列设计的 4:6 OCWMC 的各个等重码字的误比特数目分布及码字平均误比特数(XCodeword Average Bit-Error Number, CABEN)。

表 3 设计 4:6 OCWMC 时的各个阶段决策的数据

阶段变量 K	决策 X_k, Y_k		阶段误比特数目分布 及阶段平均误比特数					最终误比特数目分布 及码字平均误比特数				
	1	2	3	4	SABEN	1	2	3	4	CABEN		
1	1	7	0	0	0	0	0.000	4	5	0	0	1.556
2	16	13	0	0	0	0	0.000	3	3	1	0	1.714
3	2	2	1	0	0	0	1.000	4	3	0	0	1.429
4	15	14	1	0	0	0	1.000	3	3	1	0	1.714
5	8	10	1	0	0	0	1.000	4	3	0	0	1.429
6	9	4	1	2	0	0	1.667	4	3	0	0	1.429
7	4	1	2	1	0	0	1.333	4	3	0	0	1.429
8	13	15	2	2	0	0	1.500	2	6	1	0	1.889
9	7	16	2	3	0	0	1.600	4	4	1	0	1.667
10	10	9	2	1	1	0	1.750	4	2	1	0	1.571
11	5	6	2	3	0	0	1.600	2	5	0	0	1.714
12	12	8	3	1	1	0	1.600	4	1	2	0	1.714
13	3	5	3	2	1	0	1.667	4	2	1	0	1.571
14	14	11	1	4	1	0	2.000	2	4	1	0	1.857
15	6	12	3	5	1	0	1.778	3	5	1	0	1.778
16	11	3	3	4	0	0	1.571	3	4	0	0	1.571
小计			27	28	5	0	1.633	54	56	10	0	1.633

结合上述算法步骤和表3,再作如下补充说明:

(1) 状态变量 s_k 反映的是 \mathbf{B} 矩阵中的元素取值的个数(维数变化),随着阶段变量 k 的不同, s_k 的取值在发生变化. 对于 $2 \leq k \leq 16$, s_k 在 \mathbf{B} 中的元素取值范围是行数为 $17-k$, 列数为 $k-1$; 对于某一个阶段变量 k 而言, 当决策变量 x_k 确定之后, 状态变量 s_k 的取值则是 \mathbf{B} 中的第 x_k 行的 $k-1$ 个列的元素, 至于说是哪 $k-1$ 个列的元素, 是由已作出的决策 y_1, y_2, \dots, y_{k-1} 来决定;

(2) 在步骤3中, 也可以采用一种探索方法, 即下棋中所采用的“看几步走一步”的方法;

(3) $f(\mathbf{Z}) = 2 \sum v_k(s_k, x_k, y_k), 1 \leq k \leq 16$, 即目标函数的值等于各阶段效益函数的值的2倍. 这从表3的最后一行小计中可以看出此结论, 这也可由式(2)得到解释, 即第 k 阶段的决策未能考虑未来第 $K+1, K+2, \dots$, 等阶段对当前决策的影响; 另外在矩阵 \mathbf{B} 和 \mathbf{M} 中有 $b_{i,j} = b_{j,i}, m_{i,j} = m_{j,i}$ 的关系说明2码元交换是一个双向事件;

(4) 随着 k 的增加, 要注意 y_k 的选择对全局的影响了. 例如, 在 $k=12$ 时就要注意对 y_{12} 的取值, 能否使 $y_{13}, y_{14}, y_{15}, y_{16}$, 组成2对有 $m_{y_i, y_j} = 0$ 的关系成立 ($i, j \in \{13, 14, 15, 16\}$), 由此保证剩余的2对“对偶”码字在2-BREE条件下无法关联(即阻止4比特误码的发生); 否则, 从 $k=12$ 开始就要考虑降低贪心度, 即效益函数 $v_k(s_k, x_k, y_k)$ 要取次最小值. 此外, 用贪心法设计4:6 OCWMC的计算时间复杂度为 $O(n^4)$.

5 仿真实验结果及性能分析

我们在仿真实验过程中使用了蒙特卡罗方法, 以便较为准确地比较用 OCWMC 和 NCWMC 编码的数据在 DHDS 记录通道中的误码性能. 仿真过程是先随机地产生5页大小为 720×720 的二值化数据页, 分别用2种 OCWMC 和 NCWMC 编码后, 经过一个具有严重的2维符号间串扰(two-dimensional inter-symbol interference, 2-D ISI)并伴有不同方差的加性高斯白噪声(additive white Gaussian noise, AWGN)的记录通道传输, 所使用的表示具有严重的符号间串扰的离散通道模型是参考文献[10]中的一个 5×5 矩阵, 再利用参考文献[4]中的维特比检测器对恢复的数据进行检测, 最后对估计的码字译码后分别统计出它们各自的误比特数, 由此画出它们的 BER 性能曲线如图3所示.

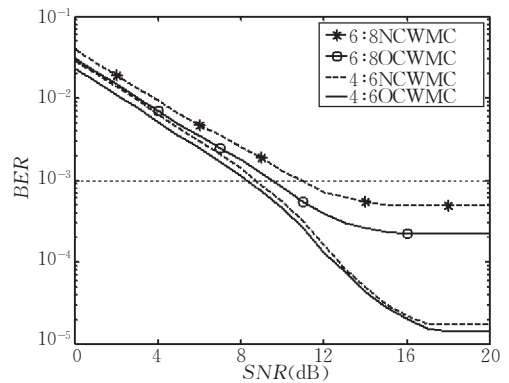


图3 2种NCWMC和OCWMC在2-D ISI/AWGN全息存储的记录通道中的BER性能

由图3可见, 在 $BER = 10^{-3}$ 时, 6:8 OCWMC 比 6:8 NCWMC 能获取约 1.362dB 的编码增益, 根据在全息存储中信噪比 SNR 与容量的关系式: $SNR = \gamma/N^2$ [11], 其中常数 γ 反映了给定的光记录通道的特性, N 为存储的页数, 因此, 相应的容量增益为 6.89%. 在仿真实验中曾使用过 RS(15, 11, 2) 纠错码, 与参考文献[6]的实验结果一样能确保将 BER 从 10^{-3} 降至 10^{-12} 以下.

根据表2的统计数据, 相对于 6:8 NCWMC, 6:8 OCWMC 与 6:8 ICWMC 的理想总误比特数偏差为 49.91%; 而相对于 4:6 NCWMC, 4:6 OCWMC 与 4:6 ICWMC 的理想总误比特数偏差为 33.33%. 我们从图2的矩阵 \mathbf{M} 的一行中可以看出, “1”的比重为 7/15 和 9/15, 它反映了 4:6 NCWMC 中一个等重码字与其它等重码字的关联程度, 该比重越大优化越困难; 而 6:8 NCWMC 中一个等重码字与其它等重码字的关联程度为 3种: 14/63, 15/63 和 16/63. 由此可以得出这样的结论, 4:6 OCWMC 进一步优化的余地较小, 而 6:8 OCWMC 还有优化的余地, 不过也许要使用其它的优化算法, 例如, 如果选择了合适的状态变量, 并确定了状态转移方程, 则有可能使用动态规划方法.

6 结 语

6:8 等重调制码是数字全息数据存储系统记录通道中的一种有效通道码, 本文借助多阶段决策方法和贪心算法, 设计出一种 6:8 优化映射等重调制码. 使用该 6:8 优化映射等重调制码能够降低数字全息数据存储系统的误码率, 加速数字全息数据存储系统的商品化进程. 我们相信为设计 6:8 优化映射等重调制码所提出的设计方法, 既适用于最小

化问题的求解,也可以用于最大化问题的求解;不仅可以用于编码设计,而且还可以用于对具有复杂关系的任务分配问题和生物中的种群优化配对等问题的求解。

参 考 文 献

- 1 Heanue J. F., Bashaw M. C., Hesselink B. L.. Volume holographic storage and retrieval of digital data. *Science*, 1994, 265: 749~752
- 2 Burr G. W., Ashley J., Coufal H., *et al.*. Optimizing the holographic digital data storage channel. In: Proceedings of the SPIE, San Diego, CA, USA, 1998, 3468: 64~75
- 3 Burr G. W., Ashley J., Coufal H., *et al.*. Modulation coding for pixel-matched holographic data storage. *Optics Letters*, 1997, 22 (9): 639~641
- 4 Ashley J., Marcus B. H.. Encoding and detection of balanced codes. United States, Electrical, 6016330, 2000
- 5 Orlov S. S., Phillips W., Bjornson E., *et al.*. Ultra-high transfer rate high capacity holographic disk digital data storage system. In: IEEE Proceedings of the 29th Applied Imagery Pattern Recognition Workshop, New York, 2000, 71~77
- 6 Burr G. W., Chou Wu-Chun, Neifeld M. A., *et al.*. Experimental evaluation of user capacity in holographic data-storage systems. *Applied Optics*, 1998, 37 (23): 5431~5443
- 7 Chen Duan-Rong, Xie Chang-Sheng, Pei Xian-Deng. Combined Viterbi detector for a balanced code in page memories. *Wuhan Univeristy Journal of Natural Sciences*, 2004, 9(1): 37~40
- 8 Deng Cheng-Liang. Theories and Methods of Operational Research. Wuhan: Huazhong University of Science and Technology, 1996, 215~217(in Chinese)
(邓成梁. 运筹学的原理和方法. 武汉: 华中理工大学出版社, 1996)
- 9 Zhou Pei-De. Algorithm Design and Analysis. Beijing: Machine Industry Publishing House, 1996, 41~95(in Chinese)
(周培德. 算法设计与分析. 北京: 机械工业出版社, 1996)
- 10 King B. M., Neifeld M. A.. Parallel detection algorithm for page—Oriented optical memories. *Applied Optics*, 1998, 37 (26): 6275~6298
- 11 Mok F. H., Burr G. W., Psaltis D.. System metric for holographic memory systems. *Optics Letters*, 1996, 21 (12): 896~898



CHEN Duan-Rong, born in 1960, Ph. D., associate professor. His research interests include coding and signal processing for the recording channel of holographic data storage.

XIE Chang-Sheng, born in 1957, professor and Ph. D. supervisor. His research interests include computer network

storage, computer architecture and high density optical data storage technologies.

PEI Xian-Deng, born in 1934, professor and Ph. D. supervisor. His research interests include computer storage system and high density optical data storage technologies.

XIAO Wei-Jun, born in 1974, Ph. D., lecturer. His research interests include compute and database application technologies.

Background

The work is supported by the National Basic Research Program of China (973 Program) under No. G1999033006. The project is named as Study of Fast and Parallel Processing for Data Storage, whose goal is to explore the theories and applications of holographic storage. Coding techniques, signal processing, and holographic storage prototype are all the directions of the project group. More than 50 papers supported by the project have been published.

ted by the project have been published.

In coding techniques, the study of an effective channel code is an important task for the holographic storage. In this paper, a 6 : 8 optimal mapping constant weight modulation code is designed by the multistage decision method and the greedy algorithm.