

# 基于极化码的分布式信源编码<sup>①</sup>

蔡丽萍, 孙悦

(中国石油大学(华东) 计算机与通信工程学院, 青岛 266555)

**摘要:** 介绍了分布式信源编码的基本思想和经典的分布式信源编码的构造实现方法. 讨论了联合信道-分布式信源编码方法, 通过新兴的极化码, 实现了联合极化码-分布式无损信源编码系统, 并将其应用到图像传输当中, 实验仿真验证了其可行性.

**关键词:** 分布式信源编码; 联合信道-分布式信源编码; 极化码; 联合极化码-分布式无损信源编码; 图像传输

引用格式: 蔡丽萍, 孙悦. 基于极化码的分布式信源编码. 计算机系统应用, 2017, 26(7): 273-277. <http://www.c-s-a.org.cn/1003-3254/5862.html>

## Distributed Source Coding Using Polar Codes

CAI Li-Ping, SUN Yue

(College of Computer and Communication Engineering, China University of Petroleum, Qingdao 266555, China)

**Abstract:** This paper introduces the basic idea of distributed source coding and the construction method of classical distributed source coding. It discusses the joint channel distributed source encoding method with the new polar code. It realizes the joint polar code-distributed lossless source encoding system and finally applies the method to the image transmission and the simulation experiment proves its feasibility.

**Key words:** distributed source coding; joint channel-distributed source coding; polar code; joint polar code-distributed lossless source coding; image transmission

分布式信源编码是信息论中的一个传统问题, 它的基本思想是编码端多信源分别独立编码、解码端联合解码, 适用于编码端资源非常有限的场合. 与传统的信源联合编码相比, 分布式信源编码降低了编码端运算、内存等资源消耗, 并且在解码端利用冗余联合解码, 可以获得和联合编解码方式相近的编码效率.

分布式编码方法的理论基础是Slepian-Wolf理论<sup>[1]</sup>和Wyner-Ziv理论<sup>[2]</sup>, 前者是无损分布式编码, 后者是有损分布式编码. 目前很多种分布式编码的实现框架被提出, 文献[3]中提出了一种基于Slepian-Wolf理论的构造方案. 其思想是使用卷积码进行信道编码得到分布式信源编码器的输入, 并且在每一路信道只传输部分码流, 在解码端利用伴随译码的方法实现重建, 相当于联合了信道编码的方式降低了单路信号的码流, 实现了分布式编码.

由于分布式信源编码的构造实现方式和信道码的特点十分的相似, 目前国内外主要是采用LDPC码<sup>[4]</sup>和Turbo码<sup>[5]</sup>来实现分布式信源编码. 新兴的极化码比起LDPC码和Turbo码, 编译码的复杂度更低, 并且可以无限的接近香农容量, 因此本文采用极化码来实现联合信道-信源编码.

本文的主要工作是介绍了分布式信源编码和极化码的相关原理, 然后实现了联合极化码-分布式无损信源编码系统, 最后提出了一种方案将这一方法应用到图像传输中.

## 1 相关原理

### 1.1 Slepian-Wolf分布式编码原理

Slepian-Wolf分布式编码在编码端对相关信源进

<sup>①</sup> 收稿时间: 2016-11-08; 收到修改稿时间: 2016-12-15

行独立编码,在解码端利用相关性进行联合解码.如图1所示.

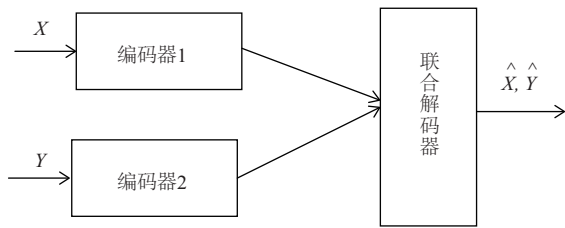


图1 Slepian-Wolf示例框图

根据信息论可知,传统的信源编码当码率符合以下条件时:

$$R_x \geq H(X)$$

$$R_y \geq H(Y)$$

即总码率满足以下条件时:

$$R_x + R_y \geq H(X) + H(Y) > H(X, Y)$$

解码端可以无差错的恢复信息.而分布式信源编码的目标是在总码率低于每个码率的熵的和情况下可以任意小的差错概率无损重建信息.

通过分析无损编码器传输码率可知码率如下:

$$R_x \geq H(X|Y)$$

$$R_y \geq H(Y|X)$$

$$R_x + R_y \geq H(X, Y)$$

因此,分布式无损信源编码理论可以获得和传统的信源编码相同的编码效率.

### 1.2 Wyner-Ziv有损分布式编码原理

与Slepian-Wolf理论不同的是, Wyner-Ziv的主要研究是在边信息存在的情况下进行分布式信源编码,即在解码端边信息Y存在的情况下, Wyner-Ziv给出了如何在可接受失真d的情况下实现分布式有损信源编码.图2是一个Wyner-Ziv有损编码器.

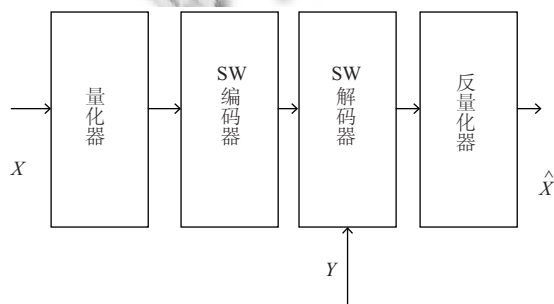


图2 Wyner-Ziv有损编码框图

Wyner-Ziv有损编码器可以看做是矢量量化器和

无损信源编码器的结合.可以看出,在整个Wyner-Ziv编码过程中,量化器的选择影响了整个编码器的性能和编码速率.

### 1.3 分布式信源编码的实现方法

分布式信源编码原理虽然早已提出,但理论上并没有给出具体的码字构造,直到最近十几年才出现具体的方案,其中最主要是利用陪集,即Pradhan和Ramchandran提出的DISCUS<sup>[6]</sup>的方法来构造实现的,编码器只需要通过边信息和陪集的索引即可恢复信源.由于这种陪集的划分与信道码的特性十分的相似,因此可以利用信道码来实现陪集的构造和信源的分割.

假定一个线性分组码(n, k),它的生成矩阵为G,校验矩阵为H.对于每一组信道码的陪集都可以用它的伴随式(Syndrome)来表示,定义伴随式为S=XH<sup>T</sup>,X是该信道码中的码字,每一组信道码对应一个伴随式,该伴随式也就是这个陪集的索引,在编码过程中传输该信道码的索引到解码端,在解码端利用参考信息在伴随式所指示的信道码中找到与距离最近的码字,即解码结果.

近年来对于分布式信源编码的研究主要是基于DISCUS方法利用LDPC码和Turbo码来实现的.而本文采用的极化码,因为编译码的复杂度低,并且可以无限接近香农容量,因此在分布式信源编码中也有很好的应用.

### 1.4 极化码

极化码最早是由Arikan在2009年提出的<sup>[7]</sup>,它是基于信道极化原理<sup>[7]</sup>实现的一种新型线性分组码.极化码的编译码复杂度低,并且可以无限的接近信道容量,因此一经提出便成了研究热点.

#### 1.4.1 信道极化

信道极化是对于任意的N个独立二进制无记忆信道,其输入比特经过一系列的变换后,输入信道传输,除了一小部分信道外,其余的信道都会表现为信道容量趋向于1或者是0的现象.

信道极化包含信道组合和信道分裂两个过程.信道组合是将N个相同的信道W经过一系列变换之后,产生一个向量信道W<sub>N</sub>,而信道分裂则是将已经合成的信道分成N个一维并且信道容量不同的信道.

#### 1.4.2 极化码编码

极化码的编码方式很简单.对于一个线性陪集码P(N, k, A, u<sub>A<sup>c</sup></sub>), N表示码长, k表示信息位的长度, A表示

信息位位置的集合,即信道容量接近1的信道索引集合,  $u_{A^c}$ 表示冻结位置上的信息比特,则极化码的编码过程可以表示为:

$$x^N = u^N G_N$$

### 1.4.3 极化码译码

极化码的译码方法主要有连续删除译码(SC)<sup>[8]</sup>和置信译码(BP)<sup>[9]</sup>等.其中,SC译码方法是在已知接收端接收到的信息、固定信息位置和冻结位置的比特的情况下,通过计算似然比进行判决,最后得到输入的比特估计.过程如下文所述.

首先,定义似然比(LR):

$$L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) = \frac{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} | 0)}{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} | 1)}$$

然后令:

$$h_i(y_1^N, \hat{u}_1^{i-1}) = \begin{cases} 0 & \frac{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} | 0)}{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} | 1)} \geq 1 \\ 1 & \frac{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} | 0)}{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} | 1)} < 1 \end{cases}$$

最后进行判决得到比特的估计值:

$$\hat{u}_i = \begin{cases} u_i & i \in A^c \\ h_i(y_1^N, \hat{u}_1^{i-1}) & i \in A \end{cases}$$

通过计算得到极化码的运算复杂度为  $N(1 + \log_2 N)$ .可以看出,相比起传统的Turbo码和LDPC码,极化码的运算复杂度只有线性级,并且随着输入码字长度的增加,极化码可以无限的接近香农容量,因此极化码已被应用到通信保密等众多领域中.

## 2 联合极化码-分布式无损信源编码的实现

本文根据分布式无损信源编码原理,极化码原理和DISCUS算法实现了一种联合极化码-分布式无损信源编码<sup>[10]</sup>的方法.

### 2.1 分布式无损编码系统实现

根据分布式无损信源编码的原理,实现了一种分布式无损信源编码器.过程如下:

(1) 随机产生一组长度为  $N$  的序列  $X$  和一组长度为  $N$  的与  $X$  相关的序列  $Y$ .

(2) 对  $X$  和  $Y$  进行分割: 令  $X=[X_a, X_b]$ ,  $Y=[Y_a, Y_b]$ , 长度分别为  $k$  和  $N-k$ .

(3) 对长度为  $k$  的  $X_a$  和  $X_b$  进行分割, 令  $X_a=[X_a^1, X_a^2]$ ,  $Y_a=[Y_a^1, Y_a^2]$ , 长度分别为  $k_1$  和  $k-k_1$ .

(4) 计算  $X$  和  $Y$  的伴随式  $S_x$  和  $S_y$ , 长度为  $N-k$ .

(5) 将  $[X_a^1, S_x]$  和  $[Y_a^2, S_y]$  作为两个独立信道编码器的输入.

(6) 在解码端利用以下公式进行解码:

$$X_2^a = Y_2^a \oplus E_{k_1+1}^k$$

$$Y_1^a = X_1^a \oplus E_0^{k_1}$$

$$x^b = (s_x \oplus x^a H_a^T)(H_b^T)^{-1}$$

$$y^b = (s_y \oplus y^a H_a^T)(H_b^T)^{-1}$$

过程如图3所示.

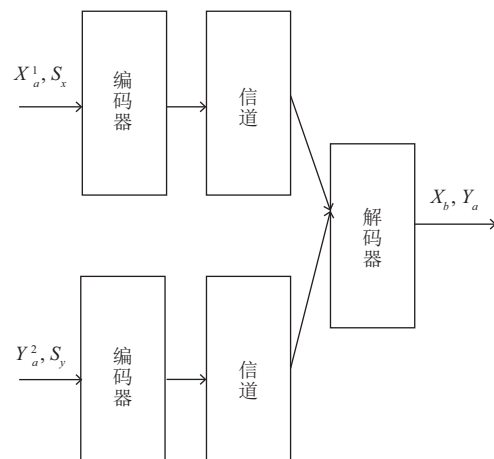


图3 分布式无损信源编码框图

其中,解码器是利用伴随式进行伴随式译码得到分布式信源编码器输入端口的信息比特,最后利用解码公式进行解码得到全部输入序列.

### 2.2 对称极化码的实现

标准的极化码是不对称的极化码,是无法用上文设计的解码方法实现无损分布式信源解码的.因此,首先要将标准的不对称极化码转换成对称的极化码<sup>[11]</sup>.过程如下:

(1) 随机产生一组长度为  $N$  的0,1序列  $u$ .

(2) 对于信息位位置集合  $A$  进行比特翻转得到集合  $B$ .

(3) 将生成矩阵  $G$  进行矩阵分割:

$$\begin{aligned} G_{AB} &= G_{i \in A, j \in B} \\ G_{AB^c} &= G_{i \in A, j \in B^c} \\ G_{A^c B} &= G_{i \in A^c, j \in B} \\ G_{A^c B^c} &= G_{i \in A^c, j \in B^c} \end{aligned}$$

(4) 利用以下公式将不对称的极化码转换成对称的极化码:

$$\begin{aligned} X_B &= u_A G_{AB} + u_{A^c} G_{A^c B} \\ X_{B^c} &= u_A G_{AB^c} + u_{A^c} G_{A^c B^c} \end{aligned}$$

### 2.3 联合极化码-分布式信源编码系统实现

将转换后的对称极化码和无损分布式信源编码系统进行联合. 设置初始值如下:

$$\begin{aligned} X_B &= X_a \\ X_{B^c} &= X_b \\ S_x &= u_{A^c} \end{aligned}$$

最后利用下式进行解码:

$$\begin{aligned} u_A &= x_B - u_{A^c} G_{A^c B} (G_{AB})^{-1} \\ x_{B^c} &= u_A G_{AB} + u_{A^c} G_{A^c B^c} \end{aligned}$$

### 2.4 方法可行性分析

对上文设计的编码系统的码率和熵进行计算得到:

$$\begin{aligned} R_x &= N - K + k_1; \\ H(X|Y) &= N - K; \\ R_y &= N - K + k_2; \\ H(Y|X) &= N - K; \\ H(X, Y) &= 2N - K; \end{aligned}$$

可以看出, 该编码系统采用了独立编码, 利用相关性进行联合解码的方法, 码率和熵符合分布式无损信源编码的码率界限. 因此该方法是可行有效的. 由于极化码的编译码复杂度低, 因此相比于传统的利用LDPC等实现的分布式信源编码, 该方法对于传输相关多信源数据的效率更高.

## 3 联合极化码-分布式无损信源编码在灰度图像传输中的应用

### 3.1 方法设计

本文提出了一种利用联合极化码-分布式无损信源编码的系统来实现无损图像传输. 这里的图像我们采用的是灰度图像.

- (1) 随机选取一张灰度图像.
- (2) 读入图像, 得到像素矩阵集合X.
- (3) 将X作为极化码编码器的输入, 利用不对称极化码转换成对称极化码的方法转换成X'.
- (4) 将X'经过一组虚拟的噪声信道产生于X'相关的序列Y.

(5) 将X'和Y作为分布式无损信源编码的输入X和Y进行分布式无损信源编码.

(6) 将分布式信源编码的输出经过运算得到原始图像矩阵集合的估计X\_estimate.

(7) 将估计值X\_estimate恢复成灰度图像, 并与初始的灰度图像作对比.

### 3.2 方法可行性分析

通过实验仿真分析, 利用该联合极化码-分布式无损信源编码系统得到了与原始图像相同的图像, 验证了该方法的可行性.

## 4 实验结果与分析

### 4.1 联合极化码-分布式无损信源编码系统的实验结果与分析

通过Matlab仿真, 分别设置了不同的码长N和不同的信息位传输码字数k<sub>1</sub>, 结果如表1和表2所示.

通过表1可以分析出: 利用任意码长的极化码均可以实现分布式无损信源编码, 即误码比特为0的目标, 即联合极化码-分布式无损信源编码系统是可行有效的.

表1 不同码长(N)的信息位个数(K)和传输误码比特数(ERR)

N(bit)	64	128	256	512	1024
K(bit)	13	26	51	102	205
ERR(bit)	0	0	0	0	0

通过表2可以分析出: 只要传输码率在SW分布式信源编码理论码率界限之内, 传输任意的信息位比特都可获得和传统的信源编码相同的编码效率.

表2 当N=512时, 不同K<sub>1</sub>的误码比特数

K <sub>1</sub> (bit)	5	20	45	80	90
ERR(bit)	0	0	0	0	0

### 4.2 联合极化码-分布式无损信源编码系统在灰度图像传输应用中的实验结果与分析

基于上述联合极化码-分布式无损信源编码系统应用到灰度图像传输中, 并利用软件进行了仿真. 这里我们的原始图像采用的是一张大脑内部灰度图像.

通过实验仿真得到了和原始图像像素矩阵相同的重建图像像素矩阵, 并通过图像观察可以发现该方法实现了灰度图像无损传输. 相比于传统的独立解码的方式, 该方法减少了编码器的运算量, 对于传输数据量大的图像更有优势. 相比于联合LDPC-分布式无损信源编码方法<sup>[12]</sup>, 该方法的编译码复杂度更低. 这一



结果的实现对于其他更加复杂的图像(如多光谱图像, 高光谱图像等)的压缩传输提供了新的研究思路.

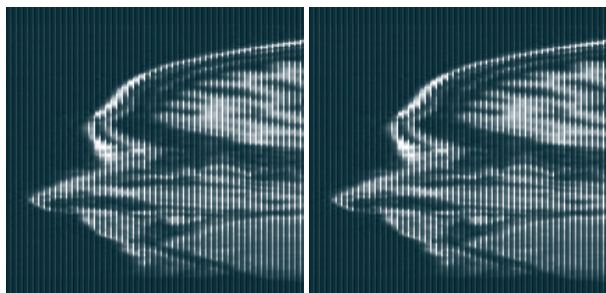


图4 原始图像和重建图像对比图

#### 4 结语

本文实现了一种联合极化码-分布式无损信源编码系统, 并提出了一种方法将这一联合信源-信道码应用到灰度图像传输技术中. 文中详细阐述了实验的具体实现方式, 内容详细, 步骤完整, 并通过仿真工具进行了实验的验证. 下一步的主要工作是将该方案应用到高光谱图像传输中.

#### 参考文献

- Slepian D, Wolf J. Noiseless coding of correlated information sources. *IEEE Trans. on Information Theory*, 1973, 19(4): 471–480. [doi: [10.1109/TIT.1973.1055037](https://doi.org/10.1109/TIT.1973.1055037)]
- Wyner A, Ziv J. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. on Information Theory*, 1976, 22(1): 1–10. [doi: [10.1109/TIT.1976.1055508](https://doi.org/10.1109/TIT.1976.1055508)]
- Gehrig N, Dragotti PL. Symmetric and asymmetric Slepian-Wolf codes with systematic and nonsystematic linear codes. *IEEE Communications Letters*, 2005, 9(1): 61–63. [doi: [10.1109/LCOMM.2005.1375242](https://doi.org/10.1109/LCOMM.2005.1375242)]
- 许迪佳. 基于LDPC的分布式信源编码关键技术研究[硕士学位论文]. 北京: 北京邮电大学, 2015.
- 吴宪云. 分布式信源编码关键技术研究[博士学位论文]. 西安: 西安电子科技大学, 2012.
- Pradhan SS, Ramchandran K. Distributed source coding using syndromes (DISCUS): Design and construction. *IEEE Trans. on Information Theory*, 2003, 49(3): 626–643. [doi: [10.1109/TIT.2002.808103](https://doi.org/10.1109/TIT.2002.808103)]
- Arikan E. Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *IEEE Trans. on Information Theory*, 2009, 55(7): 3051–3073. [doi: [10.1109/TIT.2009.2021379](https://doi.org/10.1109/TIT.2009.2021379)]
- 宋雷. 极化码SC译码算法研究[硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2015.
- 陈国泰, 张朝阳, 张亮, 等. 系统极化码的置信传播译码性能分析. *电讯技术*, 2016, 56(8): 839–843.
- Önay S. Polar codes for distributed source coding. Turkey: Bilkent University, 2014.
- Arikan E. Systematic polar coding. *IEEE Communications Letters*, 2011, 15(8): 860–862. [doi: [10.1109/LCOMM.2011.061611.110862](https://doi.org/10.1109/LCOMM.2011.061611.110862)]
- 马丕明, 袁东风, 杨秀梅, 等. 低密度校验码及其在图像传输中的应用. *电子与信息学报*, 2004, 26(8): 1269–1275.