

# 屏幕编码中基于二叉树的算术编码二值化方法\*

陶品<sup>1,2</sup>, 姜楠<sup>1</sup>, 杨士强<sup>1,2</sup>



<sup>1</sup>(清华大学 计算机科学与技术系, 北京 100084)

<sup>2</sup>(清华信息科学与技术国家实验室(筹)(清华大学), 北京 100084)

通讯作者: 陶品, E-mail: taopin@tsinghua.edu.cn

**摘要:** 在算术编码研究中,待编码的语法元素需要采用何种二值化方法以及二值化后每个比特的概率模型选择是算术编码算法设计必须面对的问题.提出了一种基于二叉树的熵编码二值化方法.该方法首先获得语法元素的统计概率分布,然后根据不同的二值化方法计算二值化后每个比特的概率分布,再通过熵编码模型的理论计算得到比特二叉树,依据二叉树中节点熵值的计算与排序,从而获得优化的二值化方法和比特概率模型分组方案.实验结果表明,该方法在屏幕编码框架下针对新的语法元素实现了平均 18.06% 的压缩效率提升.

**关键词:** 熵编码;算术编码;二值化

中文引用格式: 陶品,姜楠,杨士强.屏幕编码中基于二叉树的算术编码二值化方法.软件学报,2016,27(Suppl.(2)):113-119. <http://www.jos.org.cn/1000-9825/16026.htm>

英文引用格式: Tao P, Jiang N, Yang SQ. Binarization method based on binary tree for arithmetic coding in screen content coding. Ruan Jian Xue Bao/Journal of Software, 2016,27(Suppl.(2)):113-119 (in Chinese). <http://www.jos.org.cn/1000-9825/16026.htm>

## Binarization Method Based on Binary Tree for Arithmetic Coding in Screen Content Coding

TAO Pin<sup>1,2</sup>, JIANG Nan<sup>1</sup>, YANG Shi-Qiang<sup>1,2</sup>

<sup>1</sup>(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

<sup>2</sup>(Tsinghua National Laboratory for Information Science and Technology (TNList) (Tsinghua University), Beijing 100084, China)

**Abstract:** Binarization, possibility model and the entropy engine are the three stages of the arithmetic coding. In order to deal with the new syntax elements in a screen content codec, the binary tree base binarization decision method is proposed in this paper. The syntax probability distribution is derived, and then the bit probability distribution is calculated with the different binarization method. The binary tree method is used to get the optimal bit grouping. Experimental results show that the method proposed in this paper can improve the compression rate by 18.06% in average.

**Key words:** entropy coding; arithmetic coding; binarization

熵编码是视频编码的最后一个环节<sup>[1,2]</sup>,视频图像经过预测、变换等前期压缩环节,已经去除了视频中大部分的冗余信息,编码器得到了一系列有着确切含义的语法元素.这些语法元素的数值经过二值化,相应的比特为 0 或 1 的概率仍具有一定的统计规律,这为进一步压缩视频编码数据提供了可能,并且这个过程是无损的.熵编码就是利用数据之间的统计特征,在视频数据转换成最终的输出比特流之前进行最后一环节的压缩,算术编码是当前被广泛使用的一种熵编码算法.

熵编码概念在 1986 年国际电信联盟(ITU-T)提出的 H.261 视频压缩标准中就已经存在了,从 H.261 中的 2D-VLC 到 H.263+标准中的 3D-VLC<sup>[3]</sup>,并在 2003 年的 H.264 视频压缩标准中改进为基于算术编码的自适应编

\* 基金项目: 国家自然科学基金(61572275)

Foundation item: National Natural Science Foundation of China (61572275)

收稿时间: 2016-05-01; 采用时间: 2016-10-18

码,其全称为基于上下文的自适应二进制算术编码(context-based adaptive binary arithmetic coding,简称 CABAC)<sup>[4]</sup>,基于算数编码的数学原理,熵编码的性能得到了很大的提升并接近理论极限.2013 年推出的最新 H.265 编码标准<sup>[5]</sup>也采用了算术编码作为最后压缩环节的熵编码算法.

CABAC 的基本流程如图 1 所示,整个编码流程分为两个阶段.第 1 阶段称为二值化,是把语法元素按一定的规则转化成由 0 和 1 组成的比特串,然后送到第 2 阶段的编码引擎进行编码.第 2 阶段有两个分支,每一个分支都有一个以算术编码为核心的编码引擎,最终输出压缩的比特流.当前,熵编码的核心依然是基于算术编码<sup>[6,7]</sup>,熵编码的相关研究主要集中在 5 个方面.(1) 研究如何提高熵编码的编码效率.熵编码作为一种压缩编码技术,前人的研究方法主要是通过改变 CABAC 编码概率模型的设计结构来达到更高的编码效率<sup>[8,9]</sup>.(2) 研究如何提高熵编码的吞吐率.基于算术编码的熵编码器由于其自身的一些限制和语法元素之间的关联性,很难实现快速的编解码.研究者通过改变编码顺序、优化概率模型、提高预测准确率等方法来提高熵编码器的吞吐率<sup>[10]</sup>.(3) 研究 CABAC 编码器的并行化<sup>[11]</sup>.算术编码的并行编解码对于实际应用有着非常重要的意义,可以极大地提高编码器和解码器的工作效率,特别是对解码器而言,典型的视频解码器有大约 1/3 的解码计算量是用于熵解码.算术编解码器的并行其核心思想是合理分析数据之间的关联,在牺牲尽量少编码效率的情况下,尽量多地提高运算速度与并行度<sup>[12]</sup>.(4) 研究基于熵编码的加密.由于在一些应用场景中存在视频加密的需要,在熵编码的过程中添加加密算法是一种比较方便和安全的选项<sup>[13]</sup>.(5) 研究熵编码的硬件实现<sup>[14]</sup>.视频编解码器常常需要最终在硬件上实现,所以,如何在硬件计算资源、存储资源有限的条件下实现熵编解码器,并且可以高速、高效稳定地运行,也是一个典型的研究方向.

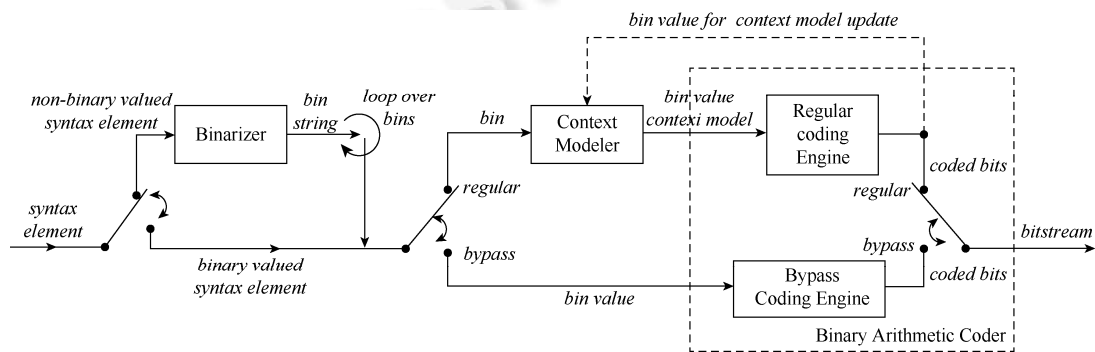


Fig.1 Framework of CABAC arithmetic coding<sup>[5]</sup>

图 1 算术编码 CABAC 的框架<sup>[5]</sup>

上述研究工作大多是针对已经发布的 H.264,AVS 或 H.265 标准开展研究,标准在建立的过程中是通过标准工作组会议讨论确定了各编码技术的细节,对于熵编码而言,各语法元素采用何种二值化方法,以及二值化后哪些比特位采用同一个概率模型,但对于这些技术细节是如何得出的,则相关文献较少.本文针对一种新提出的屏幕编码框架研究算术编码中的二值化和比特概率模型分组问题,该框架产生了多种新的语法元素,不能直接采用已有视频编码标准中的算术编码方案,因此如何实现熵编码中的每一个细节就成为需要解决的问题.针对这些问题,本文研究了算术编码中语法元素二值化方法选择的问题,以及二值化后如何对各比特的概率模型进行分组问题进行了研究,提出了一种基于二叉树的算术编码二值化方法.本文第 1 节简单介绍新的屏幕编码框架及其产生的语法元素,以及常用的二值化方法.第 2 节给出基于二叉树的算术编码二值化方法的细节.第 3 节是实验结果以及对实验结果的分析.

## 1 新的屏幕编码框架及其生产的语法元素

本文的熵编码研究是基于一种新提出的屏幕编码框架<sup>[15]</sup>,该屏幕编码的帧内编码框架如图 2 所示,原始图像首先经过重排序得到重组图像,再对重组图像使用模板匹配方法进行像素粒度的预测,图像原始值与

测值相减得到残差.残差使用调色板等方法经过残差改善后送给熵编码器,同时也送给残差恢复模块用于生成重构图像.

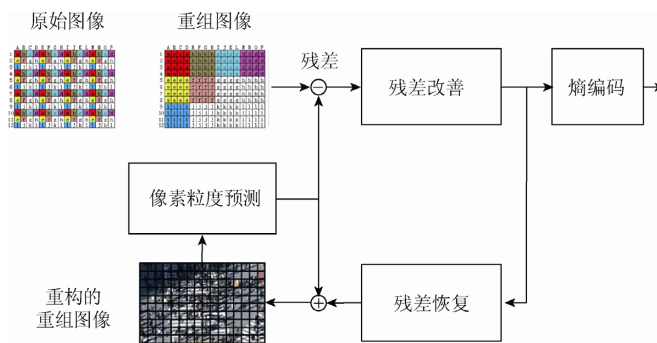


Fig.2 Framework of screen content intra frame coding

图2 屏幕编码帧内编码器框架

由于屏幕图像的无噪声、重复图像块多的特点,经过模板匹配可以对大部分像素得到正确的预测值,因此重组图像块中大部分的参数值为0,只有一小部分残差为非零值.如图3所示,假定一个图像块中只有 $A, B, C, D$ 这4个像素点的预测残差为非零值,因此后续的编码环节就是如何对这些非零值进行描述和记录.

一种直接的方法是采用下面5种语法元素来表示图像块中的非零残差,即 $idx, numindex, x, y$ 和 $runlength$ .其中, $idx$ 是图像块的编码模式,在本文所采用的编码器中只有4种方法可供选择,所以 $idx$ 的值为0~3. $numindex$ 数据表示一个图像块内的非零点数量,最小为1,最大值为 $64 \times 64 = 4096$ ,在每一个图像块中只出现1次,占数据比例较小. $x$ 和 $y$ 是一对坐标,用来表示非零像素值在图像块中的位置,在本文中我们采用了非零像素值之间的相对偏移来记录 $x, y$ 值,在每个图像块中, $x$ 值与 $y$ 值最大为63,最小为-63. $Runlength$ 表示游程,最大可以达到 $64 \times 64$ ,最小为1的正数.

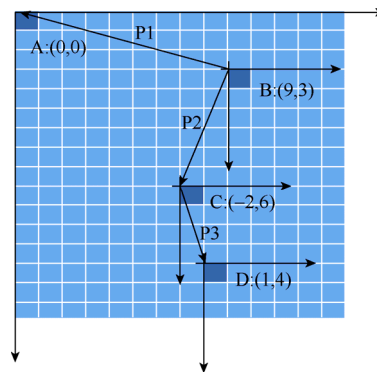


Fig.3 Non zero residual denoted as point  $A, B, C, D$  in the rearranged block

图3 重组图像块中的非零残差,即图中的 $A, B, C, D$ 这4个点

得到上述语法元素和,由于这些语法预算的概率统计规律具有新的特点,因此需要重新考虑这5种语法元素的二值化方法,以及二值化后比特的概率模型分组方案.

常用的二值化方法有4种,分别是定长码、一元码、截断一元码、 $k$ 阶哥伦布编码<sup>[16]</sup>.

(1) 定长码.定长码是用固定长度的比特串来表示数据,例如,长度为8比特的定长码用00101000来表示十进制数字40.定长码的特点是容易识别,一次读写都是8位数字,编码与解码过程简单易行.缺点是,一旦长度确定,则表示的范围也随之确定,对表示范围有一定的局限性.此外,最为重要的是,定长码产生的信息熵比较大,不利于提高压缩效率.

(2) 一元码和截断一元码.一元码用1的个数来表示数值,用0作结束符.例如,用1110表示3,用111110来表示5.一元码的生成算法简单,信息熵产生的少,特别是后文中分组计算信息熵后,一元码会产生较小的理论熵.缺点是编码后,码流很长,不利于编码取值范围较大的数据.由于理论熵与实际编码产生的比特数之间会有一定的差距,虽然一元码在理论上熵很低,但实际编码中要想达到理论熵相近的效果,要设置与编码数据值相等数量的

概率模型,这在实际编码器的实现中是很难做到的,所以一元码只能用于较小的数据编码。

(3) 截断一元码是只取一元码的前  $n$  位,舍去其余部分,为保证数据正确,只能舍去最末一位 0,且在数据范围存在最大值时使用,其特点与一元码相似。

(4)  $k$  阶哥伦布编码.如表 1 所示,哥伦布编码分为前缀和后缀,前缀为一元码,后缀为定长码,后缀的长度由前缀决定.例如,0 阶哥伦布编码从 0 开始,即用“0”表示 0,用“100”表示 1,中间 0 为分隔符,前部的 1 为前缀,也可以把前 2 位的 10 看成是一元码,后部的 0 为后缀,前部的“1”的个数决定了后缀长度为 1,且从 0 开始编码.1 阶哥伦布编码默认有一位长度的后缀,因此,当从 2 开始有了 1 比特前缀时,后缀已经有 2 比特. $k$  阶哥伦布编码是一种变长的编码方式<sup>[17]</sup>,可以用相对较短的码字表示较大的数据.缺点是,计算相对比较复杂,在二值化之后的算术编码部分,哥伦布编码后半部分的数据 0 和 1 分布比较均匀,难以达到很高的压缩效率。

Table 1 Example of four binarization method

表 1 4 种二值化示例

数值	定长码	一元码	0 阶哥伦布编码	1 阶哥伦布编码
0	0000	0	0	00
1	0001	10	100	01
2	0010	110	101	1000
5	0101	111110	11010	1011
10	1010	1111111110	1110011	110100

## 2 基于二叉树的算术编码二值化方法

熵是衡量信息量的量,在视频压缩领域,熵可以认为是由 0/1 组成的比特串来表示信息量,每一个比特表示信息的多少,由它在比特串中出现的频率决定<sup>[18]</sup>.在某个比特串  $X$  中,设  $n_0$  为比特串中 0 的个数, $n_1$  是结果中 1 的个数,则比特 0 发生的概率  $p_0 = \frac{n_0}{n_1 + n_0}$ ,同理,对应的  $p_1 = \frac{n_1}{n_1 + n_0}$ ,计算比特串的信息熵为<sup>[19]</sup>

$$H_X = -p_1 \log_2 p_1 - p_0 \log_2 p_0 \tag{1}$$

$$p_1 + p_0 = 1 \tag{2}$$

其中,  $-\log_2 p_1$  是每一个值为 1 的比特所包含的信息量,同理,  $-\log_2 p_0$  是一个 0 比特的信息量, $H_X$  是比特串  $X$  的熵,它表示在这个比特串中每一个字符所包含的信息量。

$$N_{\text{理论}} = -\sum(p_1 \log_2 p_1 + p_0 \log_2 p_0) \tag{3}$$

$$p_0 + p_1 = 1 \tag{4}$$

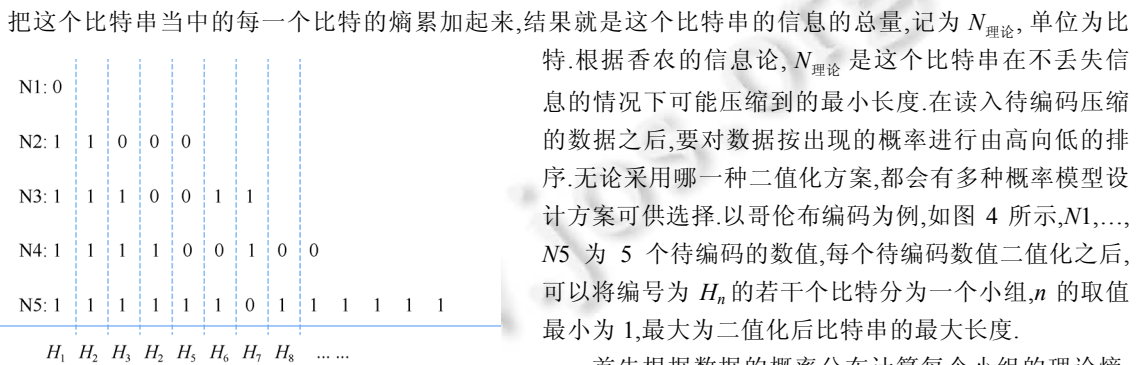


Fig.4 Example of group entropy calculating

图 4 小组熵计算示例

101 个概率值及其信息熵构成图中虚线所示的熵曲线.最初仅包含一个比特的分组信息熵记为  $H_1, H_2, H_3, \dots, H_n$ .为了确定哪几个比特小组可以进一步进行合并,从而减少总的分组个数,需要尝试:如果其中的两个比特合并分

组,新的分组可能的信息熵数值.假定两个比特对应的信息熵分别为  $p_a$  和  $p_b$ ,记每个比特为仅包含 1 个比特的小组,其信息熵为  $H_a$  和  $H_b$ ,则合并后的比特分组  $m$  信息熵为  $H_m$ .

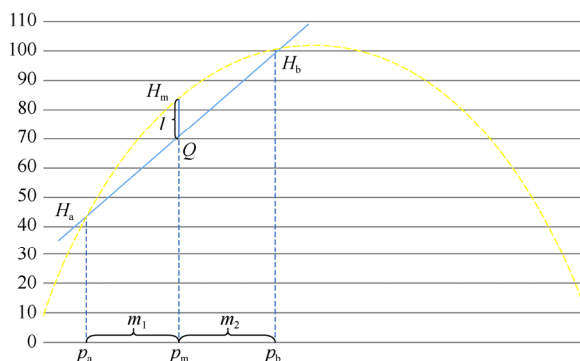


Fig.5 Diagram of  $m$  value

图 5  $m$  值示意图

如图 5 所示,比特分组  $m$  的概率  $p_m$  位于  $p_a$  和  $p_b$  之间,其在横轴上的位置满足如下关系式:

$$\frac{m_1}{m_2} = \frac{\text{(小组 } b \text{ 的总比特数)}}{\text{(小组 } a \text{ 的总比特数)}} \tag{5}$$

比特分组  $m$  的概率  $p_m$  计算如下:

$$m = \frac{p_a \text{ 中 } 1 \text{ 的个数} + p_b \text{ 中 } 1 \text{ 的个数}}{p_a \text{ 的比特数} + p_b \text{ 的比特数}} \tag{6}$$

图 5 中的  $l$  值表示比特组  $a$  与  $b$  合并后的概率平均点  $m$  的熵减去比特组  $a$  与  $b$  分别编码的熵之和的差值。 $l$  值的实际意义是,如果把比特组  $a$  与  $b$  分别编码,则实际的总熵值为图 5 中的  $Q$  点,如果把比特组  $a$  与  $b$  合并为一组,则实际编码结果熵为  $H_m$ ,用它们的差与对应小组比特数的积可以较为准确地估计两个比特小组合成以后,会有多少信息熵的增加,即理论上编码后比特的增量.通过计算任意两个小组之合并后  $l$  值与比特数的积,以积最小为寻优目标合并两个小组生成一个新的组,通过这个方法即可生成分组合并顺序的最优二叉树.如图 6 所示, $R_1, R_2, R_3, \dots$  表示各小组的编码结果理论估计值;1,2,3,4 表示对应小组合并后新生成的组的编码结果估计值.根据编码器的设计需求,如果希望把所有比特最终合并成 2 个组,则应在 1 处断开, $R_1$  为一组, $R_2, R_3, R_4, R_5$  为一组.如果要合并成 3 组,则断开 1 和 2;如果要分成 4 组,则比较 3 与 4 中各自的  $l$  值与对应的比特数的积,断开  $l$  值大的组.

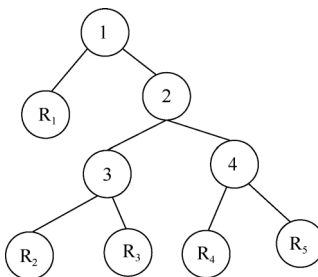


Fig.6 Optimal binary tree generated by product of  $l$  value and bits number

图 6 根据  $l$  值与比特数的积生成分组最优二叉树

在实际的编码器设计中,为了避免编解码器过于复杂,每种语法元素二值化后的比特形成的分组个数根据具体情况而定,利用本文提出的二叉树方法,可以根据编码器设计的需要,将语法元素的所有比特以最佳的方式分解为所需的分组个数.

### 3 实验结果

HM16.4-pgr 是以标准参考软件 HM-16.4 为基础,集成了按像素点的模板匹配、调色板方法编码等视频压缩

方法配置的实验环境.它以  $64 \times 64$  为基础的图像块进行编码,已经完成了预测编码等工作.其熵编码模块是用 0 阶哥伦布编码,然后以旁路算术编码的形式将结果写入了二进制文件中.本文工作以此为基础,重新实现了二值化模块,并按第 2 节所述,实现了编码概率模型和算术编码器.实验用计算机配置为 4G 内存,CPU 型号为 i5-4210U,64 位 Windows 10 专业版操作系统.HM16.4-pgr 运行环境为 Visual Studio 2012 版本.

屏幕编码的测试序列包括 desktop,flyinggraphics,map,robot,web\_browsing,wordedting 共计 6 个测试序列,反映了屏幕编码的各种典型图像类型.以 desktop 为例,其屏幕图像如图 7 所示.

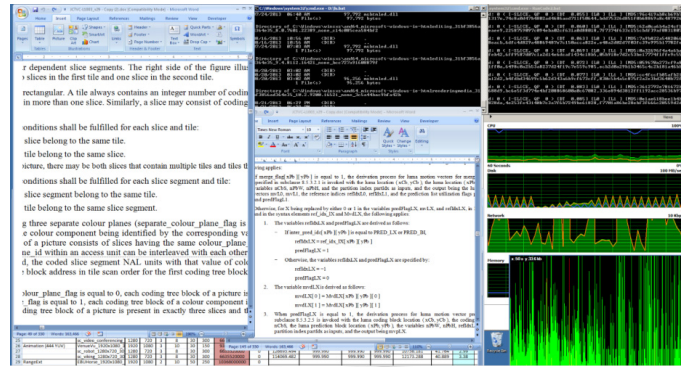


Fig.7 Test sequence of screen content image, “desktop”

图 7 屏幕图像测试序列“desktop”

所有图像块生成 5 种语法元素,每种语法元素的具体含义已在第 1 节加以说明.首先用 HM16.4-pgr 代码进行编码,记录编码结果,命名为原始结果,该编码结果表示没有采用算术编码的编码结果.然后将  $idx$  采用 2 比特定长码、 $x$  和  $runlength$  采用 0 阶哥伦布编码、 $y$  采用 2 阶哥伦布编码、 $numindex$  用 9 阶哥伦布编码.概率模型划分如下: $x$  的第 1 个比特用概率模型  $X_0$ ,其他用模型  $X_1$ , $y$  的前 2 比特用概率模型  $Y_0$ ,其他用模型  $Y_1$ , $idx$ , $runlength$  和  $numindex$  各使用一个概率模型,该编码方法记为压缩方法 1.压缩方法 2 基于第 2 节介绍的二叉树方法,采用优化后的二值化方法和概率模型分组,具体而言, $x$  和  $y$  分别采用 0 阶哥伦布二值化方法, $idx$  采用 2 比特定长码, $runlength$  采用 1 阶哥伦布编码, $numindex$  采用 9 阶哥伦布编码.对于  $x$  语法元素,第 0 比特、第 12 比特采用概率模型  $X_0$ ,其他比特用模型  $X_1$ , $y$  语法元素的第 1 比特、第 2 比特、第 3 比特、第 13 比特合并为 1 个分组,采用概率模型  $Y_0$ ,其他比特用模型  $Y_1$ , $idx$ , $runlength$  和  $numindex$  各使用一个概率模型,该编码方法记为压缩方法 2.上述 3 种熵编码的实验结果见表 2.

Table 2 Experimental result of entropy coding

表 2 熵编码结果

序列	原始结果	压缩方法 1		压缩方法 2	
		压缩结果大小	压缩率提升(%)	压缩结果大小	压缩率提升(%)
desktop	1 586 835	1 384 122	12.77	1 271 606	19.87
flyinggraphics	1 849 679	1 652 079	10.68	1 539 816	16.75
map	1 423 529	1 262 228	11.33	1 152 869	19.01
robot	939 994	874 844	6.93	844 611	10.15
web_browsing	639 034	539 697	15.54	473 507	25.90
wordedting	1 121 297	1 006 441	10.24	934 142	16.69
平均提升	-	-	11.25	-	18.06

实验结果表明,采用算术编码可以有效地根据语法元素的概率分布特性对数据进行压缩,本文使用的压缩方法 1 可以达到 11.25%的压缩率提升.但是,是否精心采用二值化方法,以及是否精心对二值化后的比特根据概率模型进行分组,对压缩率的提升具有明显效果.在本文的实验中,采用第 2 节的基于二叉树的二值化方法,达到了 18.06%的压缩效率提升.

## 4 结 论

算术编码是视频压缩编码中的一种重要熵编码方法,算术编码包括二值化、概率模型选择和编码输出这 3 个阶段.本文提出了一种基于二叉树的熵编码二值化方法,针对屏幕编码框架中产生的新语法元素,实验结果表明,该方法可以从多种二值化方法中为每个语法元素选择最佳的二值化方法,并给出二值化后比特的最佳比特分组方案,从而有效地提升了熵编码阶段的数据压缩率.

### References:

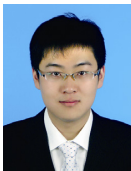
- [1] Sze V, Budagavi M, Sullivan GJ. High Efficiency Video Coding (HEVC). Springer-Verlag, 2014.
- [2] Joint Video Team (JVT) of ITU-T VCEG and ISO/IEC MPEG. Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification. ITU-T Rec. H.264 and ISO/IEC 14496-10 AVC, 2003.
- [3] Gao W, Ma SW, Zhang L, Su L, Zhao DB. AVS video coding standard. In: Intelligent Multimedia Communication: Techniques and Applications. Berlin, Heidelberg: Springer-Verlag, 2010. 125–166.
- [4] Wu X, Memon N. Context-Based, adaptive, lossless image coding. IEEE Trans. on Communications, 1997,45(4):437–444.
- [5] Sullivan GJ, Ohm JR, Han WJ, Wiegand T. Overview of the high efficiency video coding (HEVC) standard. IEEE Trans. on Circuits and Systems for Video Technology, 2012,22(12):1649–1668.
- [6] 毕厚杰,王健.新一代视频压缩编码标准-H.264/AVC.北京:清华大学出版社,2005
- [7] Richardson IEG, 著;欧阳合,韩军,译.H.264 和 MPEG-4 视频压缩:新一代多媒体视频编码技术.长沙:国防科技大学出版社,2004.
- [8] Yu W, He Y. A high performance CABAC decoding architecture. IEEE Trans. on Consumer Electronics, 2005,51(4):1352–1359.
- [9] Zhang P, Xie D, Gao W. Variable-Bin-Rate CABAC engine for H.264/AVC high definition real-time decoding. IEEE Trans. on Very Large Scale Integration Systems, 2009,17(3):417–426.
- [10] Sze V, Budagavi M. High throughput CABAC entropy coding in HEVC. IEEE Trans. on Circuits & Systems for Video Technology, 2012,22(12):1778–1791.
- [11] 祁跻.HEVC 熵编码技术研究及并行算法设计[硕士学位论文].北京:北京邮电大学,2013.
- [12] Khurana G, Kassim AA, Chua TP, Mi MB. A pipelined hardware implementation of in-loop deblocking filter in H.264/AVC. IEEE Trans. on Consumer Electronics, 2006,52(2):536–540.
- [13] 谢强,郑世宝,于晓静.一种结合 H.264/AVCCABAC 熵编码器特征的视频选择性内容加密算法.上海交通大学学报,2008,42(10):1762–1766.
- [14] Zhou DJ, Zhou JJ, Fei W, Goto S. Ultra-High-Throughput VLSI Architecture of H.265/HEVC CABAC encoder for UHD TV Applications. IEEE Trans. on Circuits & Systems for Video Technology, 2015,25(3):497–507.
- [15] Feng LX, Tao P, Guang DX, Wen JT, Yang SQ. Pixel granularity template matching method for screen content lossless intra picture. In: Advances in Multimedia Information Processing—PCM 2014. Springer Int'l Publishing, 2014. 143–152.
- [16] Marpe D, Schwarz H, Wiegand T, *et al.* Entropy coding. US 20130027230 A1, 2013.
- [17] Wu D, Gao W, Hu MZ, Ji ZZ. An Exp-Golomb encoder and decoder architecture for JVT/AVS. In: Proc. of the Int'l Conf. on ASIC, Vol.2. 2003. 910–913.
- [18] Marpe D, Schwarz H, Wiegand T. Context-Based adaptive binary arithmetic coding in the H.264/AVC video compression standard. IEEE Trans. on Circuits & Systems for Video Technology, 2003,13(7):620–636.
- [19] Rubinstein RY, Kroese DP. The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning. New York: Springer-Verlag, 2004.



陶品(1974—),男,安徽天长人,博士,副研究员,CCF 专业会员,主要研究领域为嵌入式媒体处理,视频编码。



杨士强(1952—),男,教授,博士生导师,CCF 会士,主要研究领域为多媒体信息处理,Web 社会网络分析。



姜楠(1986—),男,硕士,主要研究领域为熵编码与算术编码。