

思路：

Title:

Huffman 编解码的 matlab 实现与探究

Abstract:

这次 project, 用 matlab 对要求的.txt 文件利用 Huffman 技术实现了编解码, which 能产生最佳码, 并计算和呈现相关的参数。并做了探究和创新。探究部分是, 对同样的文件利用 Fano 技术进行编解码并计算和呈现相关参数, 与 the one using Huffman coding technique 进行比较。创新部分是采用 LZ 技术进一步优化 Huffman 编码的效果, 实现更高效率的编码, 并与 the one using Huffman coding technique 进行比较。

Introduction:

信源编码是数据压缩领域中的一种基本技术, 广泛应用于信息传输和存储中。其核心目的是通过有效地表示信息源中的符号, 以减少所需的比特数, 从而实现数据的高效传输和存储。信源编码的基本原理是将符号的不同出现概率映射为不同长度的二进制代码, 通常较高频出现的符号被分配较短的编码, 而低频符号则被分配较长的编码。这样的编码策略不仅减少了冗余数据的传输, 还能够有效节省存储空间。

霍夫曼编码是最著名的信源编码方式之一, 由 David A. Huffman 在 1952 年提出。它是一种基于贪心算法的编码方法, 能够为给定的符号集合生成最优的前缀编码, 从而确保每个符号的编码长度与其出现的概率成反比。霍夫曼编码的优点在于其高效性和优化性: 它可以根据符号的出现频率动态地生成最佳编码, 且生成的编码满足唯一可译性[1,2]。本项目, 我们便主要采用这种编码方式实现对目标文本(我们选取了前三个章节)的编码, 解码和对相关参数包括平均码长, 码率, 效率和压缩率(咱本实验中, 分别达到了 4.4769, 0.9897, 0.9897, 0.5596)的计算。

考虑到霍夫曼编码是最佳编码的产生方法, 尤其是在不同信源符号出现的概率差异较大时, 它明显优于费诺编码的编码效果[2,3], 所以我们设计了一个探究的部分, 也就是用 Fano coding technique 对目标文本进行同样的操作, 并把相关参数的计算结果与 those of the outcome using Huffman coding technique 进行比较。

另外, 考虑到 LZ 的作用, 我们进行了进一步优化……

Results:

1.对于 huffman coding 的实现

(1) 相关参数的计算结果

命令行窗口

```
>> mainHuffman
Decoding is successful!
average code length:
4.4769

code rate:
0.9897

efficiency:
0.9897

zip_rate:
0.5596
```

fx >>

(2) 编码结果 (这里以图片的形式做部分展示, 详细内容见附件中
“Huffman_encoded_output.txt”)



(3) dictionary (这里以图片的形式做部分展示, 详细内容见附件中
“Huffman_dictionary.txt”)

```
dict.txt
文件 编辑 查看

:00101100
:00101011
:11
:1001111100101
:100110
-:0100100110
-:010011
1:10011111101000011
2:1001111101000010
-:1001111101000000
-:010110101100
-:010110100
A:10011111
B:100111000
C:01001001011
D:10011111001
E:10011110011
F:1001111000
G:0100100100
H:01001000
I:001010100
J:010010011
K:100111100100
L:01001001010
M:01011010100
N:0100100011
O:01011010101
P:010110101101
R:100111001
S:100111101
T:01011011
U:100111110101
V:1001111101001
W:10011101
Y:01001000101
a:011
b:0010100
c:100101

行 17, 列 35 1,887 个字符 120% Unix (LF) UTF-8
```

(4) 解码结果（这里以图片的形式做部分展示，详细内容见附件中“Huffman_decode_message.txt”）

```
decode_message.txt
文件 编辑 查看

PROLOGUE
We should start back," Garead urged as the woods began to grow dark around them. "The wildlings are dead."
"Do the dead frighten you?" Ser Waymar Royce asked with just the hint of a smile.
Garead did not rise to the bait. He was an old man, past fifty, and he had seen the lordlings come and go. "Dead is dead," he said. "We have no business with the dead."
"Are they dead?" Royce asked softly. "What proof have we?"
"Will say them," Garead said. "If he says they are dead, that's proof enough for me."
Will had known they would drag him into the quarrel sooner or later. He wished it had been later rather than sooner. "My mother told me that dead men sing no songs," he put in.
"My wet nurse said the same thing, Will," Royce replied. "Never believe anything you hear at a woman's sit. There are things to be learned even from the dead." His voice echoed, too loud in the twilight forest.
"We have a long ride before us," Garead pointed out. "Eight days, maybe nine. And night is falling."
Ser Waymar Royce glanced at the sky with disinterest. "It does that every day about this time. Are you unmanned by the dark, Garead?"
Will could see the tightness around Garead's mouth, the barely suppressed anger in his eyes under the thick black hood of his cloak. Garead had spent forty years in the Night's Watch, man and boy, and he was not accustomed to being made light of. Yet it was more than that. Under the wounded pride, Will could sense something else in the older man. You could taste it; a nervous tension that came perilous close to fear.
Will shared his unease. He had been four years on the Wall. The first time he had been sent beyond, all the old stories had come rushing back, and his bowels had turned to water. He had laughed about it afterward. He was a veteran of a hundred rangings by now, and the endless dark wilderness that the scoutron called the haunted forest had no more terrors for him.
Until tonight. Something was different tonight. There was an edge to this darkness that made his hackles rise. Nine days they had been riding, north and northwest and then north again, farther and farther from the Wall, hard on the track of a band of wildling raiders. Each day had been worse than the day that had come before it. Today was the worst of all. A cold wind was blowing out of the north, and it made the trees rustle like living things. All day, Will had felt as though something were watching him; something cold and implacable that loved him not. Garead had felt it too. Will wanted nothing so much as to ride helibent for the safety of the Wall, but that was not a feeling to share with your commander.
Especially not a commander like this one.
Ser Waymar Royce was the youngest son of an ancient house with too many heirs. He was a handsome youth of eighteen, grey-eyed and graceful and slender as a knife. Mounted on his huge black destrier, the knight towered above Will and Garead on their smaller garrons. He wore black leather boots, black woolen pants, black moleskin gloves, and a fine supple coat of gleaming black ringmail over layers of black wool and boiled leather. Ser Waymar had been a Sworn Brother of the Night's Watch for less than half a year, but no one could say he had not prepared for his vocation. At least insofar as his wardrobe was concerned.
His cloak was his crowning glory; sable, thick and black and soft as sin. "Bet he killed them all himself, he did," Garead told the barracks over wine. "twisted their little heads off, our mighty warrior." They had all shared the laugh. It is hard to take orders from a man you laughed at in your cups, Will reflected as he sat shivering atop his garron. Garead must have felt the same.
"Mormont said as we should track them, and we did," Garead said. "They're dead. They shan't trouble us no more. There's hard riding before us. I don't like this weather. If it snows, we could be a fortnight getting back, and snow's the best we can hope for. Ever seen an ice storm, my lord?"
The lordling seemed not to hear him. He studied the deepening twilight in that half-bored, half-distracted way he had. Will had ridden with the knight long enough to understand that it was best not to interrupt him when he looked like that. "Tell me again what you saw, Will. All the details. Leave nothing out."
Will had been a hunter before he joined the Night's Watch. Well, a poacher in truth. Mallister freeriders had caught him red-handed in the Mallisters' own woods, skinning one of the Mallisters' own bucks, and it had been a choice of putting on the black or losing a hand. No one could move through the woods as silent as Will, and it had not taken the black brothers long to discover his talent.
"The camp is two miles farther on, over that ridge, hard beside a stream," Will said. "I got close as I dared. There's eight of them, men and women both. No children I could see. They put up a lean-to against the rock. The snow's pretty well covered it now, but I could still make it out. No fire burning, but the firepit was still plain as day. No one moving. I watched a long time. No living man ever lay so still."
"Did you see any blood?"
"Well, no," Will admitted.
"Did you see any weapons?"
"Some swords, a few bows. One man had an axe. Heavy-looking, double-bladed, a cruel piece of iron. It was on the ground beside him, right by his hand."
"Did you make note of the position of the bodies?"
Will shrugged. "A couple are sitting up against the rock. Most of them on the ground. Fallen, like."
"Or sleeping," Royce suggested.
"Fallen," Will insisted. "There's one woman up an ironwood, half-hid in the branches. A far-eyes." He smiled thinly. "I took care she never saw me. When I got closer, I saw that she wasn't moving neither." Despite himself, he shivered.
"You have a chill?" Royce asked.
"Same," Will muttered. "The wind, m' lord."
The young knight turned back to his grizzled man-at-arms. Frostfallen leaves whispered past them, and Royce's destrier moved restlessly. "What do you think might have killed these men, Garead?" Ser Waymar asked casually. He adjusted the drape of his long sable cloak.
"It was the cold," Garead said with iron certainty. "I saw men freeze last winter, and the one before, when I was half a boy. Everyone talks about snows forty foot deep, and how the ice wind comes howling out of the north, but the real enemy is the cold. It steals up on you quieter than Will, and at first you shiver and your teeth chatter and you stamp your feet and dream of mulled wine and rice hot fires. It burns, it does. Nothing burns like the cold. But only for a while. Then it gets inside you and starts to fill you up, and after a while you don't have the strength to fight it. It's easier just to sit down or go to sleep. They say you don't feel any pain toward the end. First you go weak and drowsy, and sometimes start to fade, and then it's like sliding into a sea of warm milk. Beautiful like."
```

2.探究：比较 Fano Coding 和 Huffman Coding

下面呈现 Fano Coding 的结果

(1) 相关参数的计算结果

命令行窗口

```
>> mainFano
Decoding is successful!
average code length:
    8.6370

code rate:
    0.5117

efficiency:
    0.5117

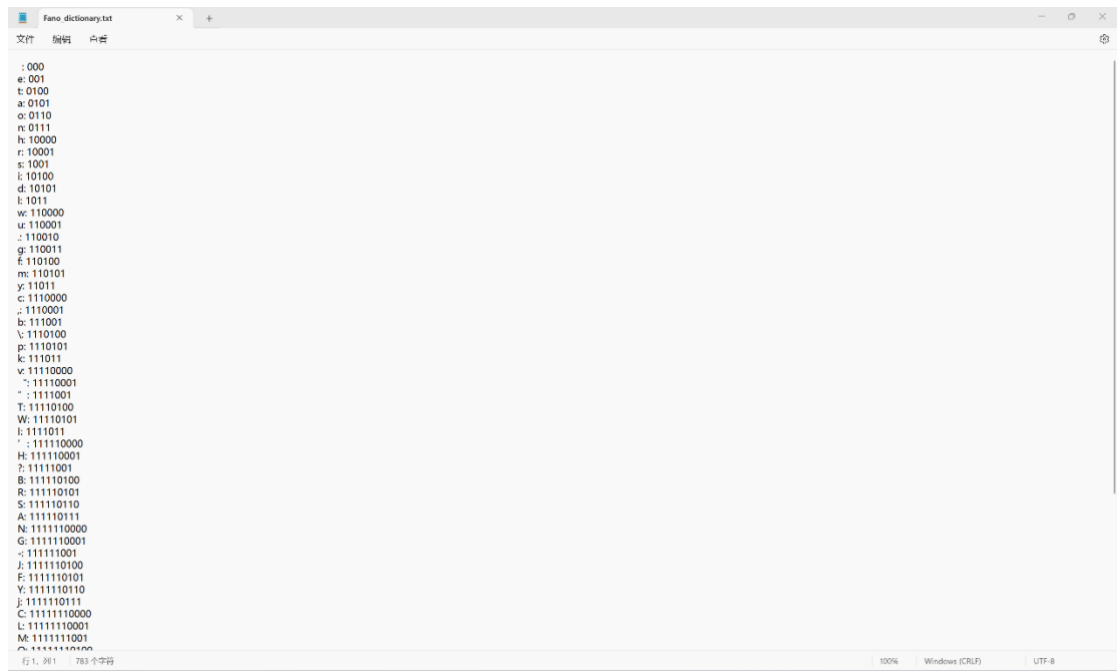
zip_rate:
    1.0796
```

fx >>

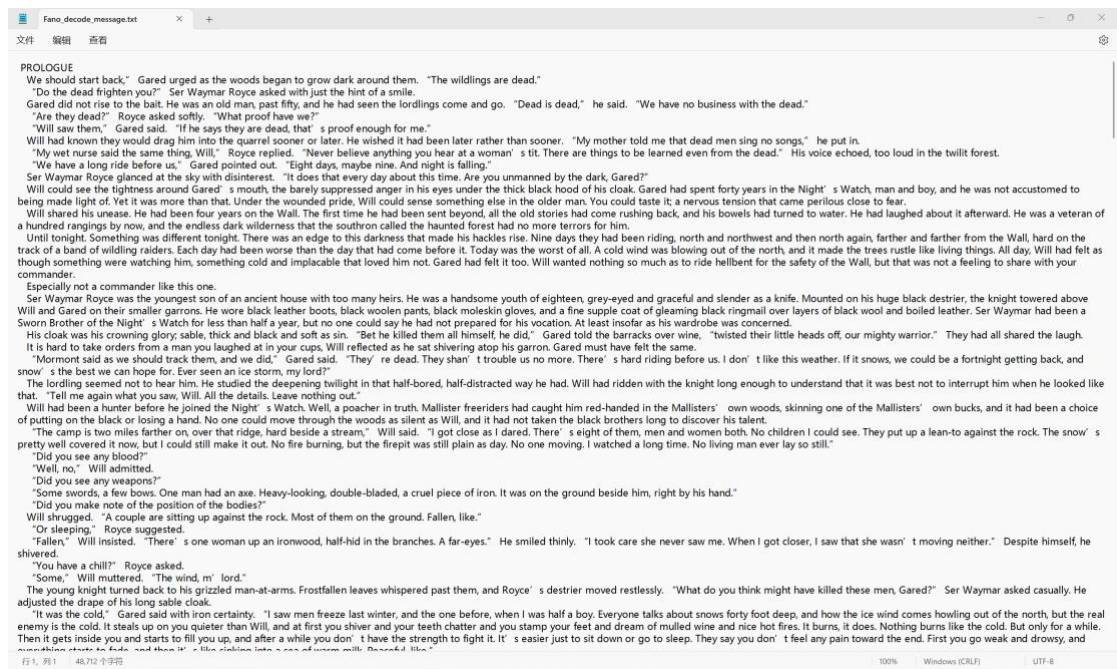
(2) 编码结果 (这里以图片的形式做部分展示, 详细内容见附件中“Fano_encoded_output.txt”)



(3) dictionary (这里以图片的形式做部分展示, 详细内容见附件中“Fano_dictionary.txt”)



(4) 解码结果（这里以图片的形式做部分展示，详细内容见附件中“Fano_decode_message.txt”）



3.创新：利用 LZ，进一步优化 Huffman Coding 的编码效果

.....

Discussion:

1.对于 Huffman Coding 的实现

通过利用 MATLAB 对初始文本实现编码，解码和计算相关参数，我们得到 Huffman 编码的平均码长是 4.4769，码率是 .

0.9897，效率是 0.9897，压缩率是 0.5596，公式：……………。

这说明我们成功的利用 Huffman Coding 技术对目标文本实现了高效的压缩，which 允许后续进行高效的压缩和存储。

2.探究：比较 Fano Coding 和 Huffman Coding

(1) Huffman Coding 明显优于 Fano Coding

经过对两种编码方式结果的分析，我们发现，霍夫曼编码的效果明显优于费诺编码的效果，并且达到了 4.4769 的平均码长，0.9897 的码率，0.9897 的效率，和 0.5596 的压缩率，这是一个非常好的压缩效果。并且这也符合霍夫曼编码的结果是最佳码的事实。

(2) Fano Coding 不适合用于编码这个文本

费诺编码的结果对应的压缩率大于 1，这是不合理的，经过分析，我们认为这篇文章不适合使用费诺编码。与其他编码方式比较，Fano 编码最有效的前提是符号的频率分布不均匀。即某些符号出现得频繁，而其他符号则相对较少。这种频率差异使得可以为频繁出现的符号分配较短的编码，从而实现压缩。但是在我们作为输入并进行压缩的这个文本中，符号的频率分布比较均匀。频繁出现的符号（如空格、字母）和稀有符号（如标点符号）的概率差异并不十分显著。在这种符号分布非常均匀的情况下，Fano 编码可能无法显著减小文件大小，甚至可能因为编码的冗余和长码字而导致文件增大，从而出现压缩率大于 1 的情况。并且我们计算了此文本各符号出现频率的方差如下：

```
K>> var(probabilities)

ans =

    8.9864e-04
```

可见，这些字符对应频率的方差非常小，这也验证了这篇文章中符号的频率分布比较均匀这一观点。

而这一切，更进一步说明了霍夫曼编码的优越性。

3.创新：利用 LZ，进一步优化 Huffman Coding 的编码效果

.....

Conclusion:

信源编码是数据压缩领域中的一种基本技术，广泛应用于信息传输和存储中。其核心目的是通过有效地表示信息源中的符号，以减少所需的比特数，从而实现数据的高效传输和存储（和开头一样，转述一下即可）。在这个 project 中，我们成功地用 MATLAB 实现了 Huffman Coding technique，并形成了压缩后的便于传输和存储的编码文件。除此之外，还通过和 Fano 编码的结果进行比较，探究了 Huffman Coding 高效性和优化性。更重要的是，我们还通过配合使用 LZ,,,,,,,,,,