

Отчёт по лабораторной 2

Наумов Семен Дмитриевич, М4138

BWT + MTF + RLE + A

Описание работы алгоритмов

Алгоритм кодирования

Burrows Wheeler transformation, Move To Front transformation

Реализации алгоритмов преобразования взяты из лабораторной работы 1.

Run Length Encoding

Вместо `n` подряд идущих одинаковых байт `a` записываем `[a, a, n]`.

Arithmetic Encoding

Арифметическое кодирование – отображение последовательности байт на отрезок. Можно брать вместо вещественной оси отрезок от 0 до $2^{16} - 1$. По вероятностному распределению частей (появления определённого байта) выбирать рекурсивно сужая нужный отрезок. И как только перестанет хватать точности, можно выводить этот отрезок в выходной поток и расширять отрезок.

Алгоритм декодирования

Reverse Arithmetic Encoding

Действуя по тем же правилам, что и при кодировании, можно восстановить изначальные символы: так же рассматривая отрезки и учитывая те же самые вероятностные распределения.

Вероятностными распределениями тут называются отношения появления битов в просмотренном потоке на `n`-ый момент времени.

Reverse Run Length Encoding

Если подряд идут два одинаковых байта, то следующий байт будет отвечать за количество повторений этого байта.

Reverse Move To Front transformation, Reverse Burrows-Wheeler transformation

Реализации алгоритмов преобразования взяты из лабораторной работы 1.

Результаты

QF = 30

file name	размер сжатого файла (байты)	размер исходного файла (байты)	сжатие (проценты)
airplane	18892	19207	1.6400
arctichare	11467	12080	5.0745
baboon	36037	36113	0.2105
cat	35873	35902	0.0808
fruits	17516	17646	0.7367
frymire	194813	200860	3.0106
girl	23645	24310	2.7355
lena	17480	17578	0.5575
monarch	28269	28501	0.8140
peppers	18671	18746	0.4001

file name	размер сжатого файла (байты)	размер исходного файла (байты)	сжатие (проценты)
pool	6216	7467	16.7537
sails	45233	45512	0.6130
serrano	57501	58288	1.3502
tulips	37870	37972	0.2686
watch	44497	46741	4.8009
SUMMARY	593 980	606 923	2.1325604731

Суммарное значение сжатого размера всех файлов в байтах: 593 980

QF = 80

file name	размер сжатого файла (байты)	размер исходного файла (байты)	сжатие (проценты)
airplane	44098	44079	-0.04310
arctichare	26287	26569	1.06139
baboon	88526	88465	-0.06895
cat	82037	81974	-0.07685
fruits	45344	45303	-0.09050
frymire	433379	440857	1.69624
girl	59691	59979	0.48017
lena	43860	43872	0.02735
monarch	64126	64055	-0.11084
peppers	47935	47929	-0.01259
pool	13845	14492	4.46453
sails	105878	105830	-0.04536
serrano	137058	138167	0.80265
tulips	85867	85764	-0.12010
watch	100028	101074	1.03489
SUMMARY	1 377 959	1 388 409	0.7526600591

Суммарное значение сжатого размера всех файлов в байтах: 1 377 959