

Отчет по лабораторной работе №2. Сжатие JPEG

Мясников Владислав, М4106

Общий принцип работы алгоритма

1. Модифицированная версия декодера JPEG-файла извлекает и сохраняет в отдельный файл следующие данные о jpg-файле: width, height, comp, quality factor и блоки AC/DC коэффициентов.
2. Архиватор получает эти данные из файла и обрабатывает следующим образом:
 - a. width, height, comp и quality factor записываются в выходную последовательность прямым кодом без сжатия
 - b. для DC-коэффициентов вычисляется diff, длины (length) diff-коэффициентов сжимаются при помощи РРМА (на одну длину выделяется 4 бита), сами значения diff записываются компактно и без сжатия друг за другом (на каждый diff выделяется length бит, на нулевые diff выделяется 0 бит)
Note: отрицательные значения преобразуются по формуле $2^{\text{length}} - |x| - 1$, при этом получается бинарная последовательность с ведущими нулями и длиной length
 - c. для ненулевых AC-коэффициентов выполняются похожие действия, за несколькими исключениями: помимо длин при помощи РРМА сжимаются еще и run-значения (кол-во подряд идущих нулей, предшествующих данному ненулевому коэффициенту AC), а также добавляются маркеры End Of Block (1 байт со значением 0)
Note: одному ненулевому AC может соответствовать один байт (4 бита на run и 4 бита на length) или два байта (8 бит на run, 4 бита на length, 4 бита-маркера)
3. Деархиватор получает сжатые данные и обрабатывает следующим образом:
 - a. извлекает width, height, comp и quality factor
 - b. при помощи деархиватора РРМА получает последовательность длин diff-DC/AC и последовательность run
 - c. последовательно считывает коэффициенты побитово и заполняет массив блоков DC/AC (DC вычисляется преобразованием diff, также происходит обработка отрицательных коэффициентов - если первый бит в бинарной последовательности числа равен нулю)
 - d. массив блоков DC/AC сохраняется в файле для выполнения следующего шага
4. Запускается модифицированная версия кодера JPEG-файла, на вход которому передаются width, height, comp, quality factor и имя файла, содержащего блоки AC/DC коэффициентов

Результаты работы алгоритма на тестовом стенде

Изображение	QF = 30	QF = 80
airplane	17841	41880
arctichare	11256	25992
baboon	33948	82709
cat	34970	77351
fruits	16851	43341
frymire	187021	402790
girl	22960	57607
lena	16821	41525
monarch	27484	60686
peppers	17703	44731
pool	6369	13817
sails	42867	100147
serrano	54455	126293
tulips	35762	80187
watch	44165	98526
Исходный размер	606923	1388409
Сжатый размер	570473	1297582
Процент сжатия	6,005704183	6,541804324