МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное образовательное учреждение высшего образования

«Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»

Мегафакультет трансляционных информационных технологий

Факультет информационных технологий и программирования

**Лабораторная работа №2**

**По дисциплине «Введение в цифровую культуру и программирование»**

**M3104 Триголос Алексей Павлович Работа2**

**Выполнил студент группы №М3104**

***Триголос Алексей Павлович***

**Проверил:**

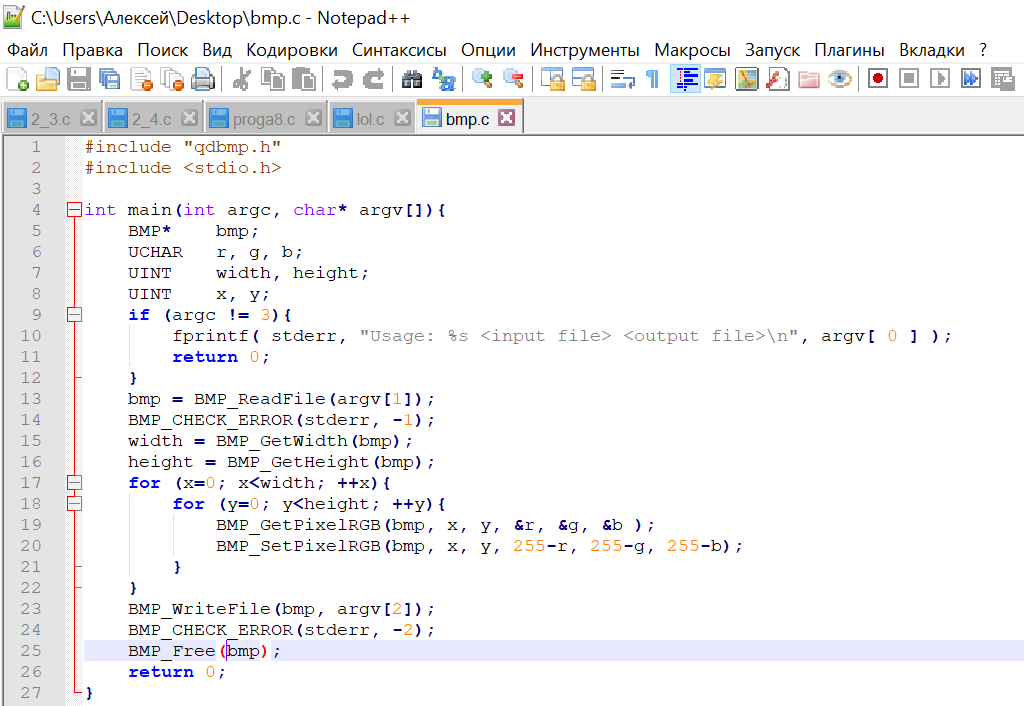
***Хлопотов Максим Валерьевич***

***САНКТ-ПЕТЕРБУРГ***

***2019***

Исходная картинка в формате bmp:



Делаем квантование центральной строки пикселей и вычисляем частоту появления уникального символа в цифровой последовательности. 

Полученная числовая последовательность: 200 200 200 200 200 200 200 200 200 200 200 200 200 180 180 180 160 160 160 180 180 180 140 140 140 160 160 160 140 140 140 140 140 140 120 120 120 100 100 100 100 100 100 140 140 140 160 160 160 160 160 160 140 140 140 140 140 140 160 160 160 140 140 140 160 160 160 160 160 160 180 180 180 200 200 200 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 220 200 200 200 200 200 200 200 200 200 200 200 200 200 200 200 200 200 200 200 200 200 200 200

Упорядоченная последовательность, полученная из преобразованной средней строки исходного изображения

|  |  |  |
| --- | --- | --- |
| Value: | Count: | Appearance count: |
| 200 | 38 | 0.296875 |
| 220 | 28 | 0.21875 |
| 160 | 22 | 0.171875 |
| 140 | 22 | 0.171875 |
| 180 | 9 | 0.0703125 |
| 100 | 6 | 0.046875 |
| 120 | 3 | 0.0234375 |

Вычисляем значение энтропии по формуле:

H - этропия

I - количество получаемой в результате снятия неопределенности информации

N - количества рассматриваемых вариантов

Количество элементов: 7  
Значение энтропии: 2.4762799838438947

Таблица равномерного кода

|  |  |
| --- | --- |
| Value: | Code: |
| 200 | 000 |
| 220 | 001 |
| 160 | 010 |
| 140 | 011 |
| 180 | 100 |
| 100 | 101 |
| 120 | 110 |

Таблица кодов Шеннона-Фано

|  |  |
| --- | --- |
| Value: | Code: |
| 200 | 00 |
| 220 | 01 |
| 160 | 10 |
| 140 | 110 |
| 180 | 1110 |
| 100 | 11110 |
| 120 | 11111 |

Таблица кодов Хаффмана

|  |  |
| --- | --- |
| Value: | Code: |
| 200 | 01 |
| 220 | 10 |
| 160 | 11 |
| 140 | 000 |
| 180 | 0010 |
| 100 | 00110 |
| 120 | 00111 |

Пояснение к дереву Хаффмана: в прямоугольниках числа, состоящие из 0 и 1 это путь до данного прямоугольника, а в скобочках находятся значения пикселей, которые кодируются таким (число, которое в этом же прямоугольнике, но сверху) двоичным кодом. Если нет числа в скобочках, следовательно, данным двоичным кодом ничего не закодировано.

Закодированное равномерным кодом сообщение, длина кодового слова, количество переданной информации при передаче сообщения

000000000000000000000000000000000000000100100100010010010100100100011011011010010010011011011011011011110110110101101101101101101011011011010010010010010010011011011011011011010010010011011011010010010010010010100100100000000000001001001001001001001001001001001001001001001001001001001001001001001001001001001001001000000000000000000000000000000000000000000000000000000000000000000000  
Длина сообщения: 384  
Размер сообщения: 384 бит = 48 байт  
Средняя длина кодового слова: 3

Закодированная полученным кодом Шеннона–Фано последовательность, длина кодового слова, количество переданной информации при передаче сообщения

0000000000000000000000000011101110111010101011101110111011011011010101011011011011011011011111111111111111110111101111011110111101111011011011010101010101011011011011011011010101011011011010101010101011101110111000000001010101010101010101010101010101010101010101010101010101010000000000000000000000000000000000000000000000

Длина сообщения: 322

Размер сообщения: 322 бита

Средняя длина кодового слова: 2.515625

Закодированная полученным кодом Хаффмана последовательность, длина кодового слова, количество переданной информации при передаче сообщения

0101010101010101010101010100100010001011111100100010001000000000011111100000000000000000000111001110011100110001100011000110001100011000000000011111111111100000000000000000011111100000000011111111111100100010001001010110101010101010101010101010101010101010101010101010101010100101010101010101010101010101010101010101010101  
Длина сообщения: 322  
Размер сообщения: 322 бита  
Средняя длина кодового слова: 2.515625

Степень сжатия сообщения кодом Шеннона-Фано: 0.83854167  
0.83854167= число знаков в исходном сообщении, которое закодировано кодом Шеннона-Фано / число знаков в исходном сообщении, которое закодировано равномерным кодом.

Степень сжатия сообщения кодом Хаффмана: 0.83854167  
0.83854167 = число знаков в исходном сообщении, которое закодировано кодом Хаффмана / число знаков в исходном сообщении, которое закодировано равномерным кодом.

Относительная избыточность = 1 – (энтропия / средняя длина кода).

Относительная избыточность (Шеннон): 1 - 2.4762799838438947 / 2.515625 = 0,015640254869508

Относительная избыточность (Хаффман): 1 - 2.4762799838438947 / 2.515625 = 0,015640254869508

Вывод

Благодаря этой работе я научился кодировать информацию такими способами: равномерное двоичное кодирование, кодирование методом Шеннона-Фано и методом Хаффмана. После сжатия заданного сообщения данными тремя способами, были посчитаны длины закодированных сообщений (размер сообщения). По измерениям мы можем заметить, что при равномерном кодировании мы можем занимать больше памяти так как, не учитываем частоту появления символа, что для нас не выгодно. Так как в картинки больше преобладают светлые тона, следовательно кодирование методами Шеннона-Фано и Хаффмана дало возможность сжать фотографию примерно на 16%, что мне кажется достаточно хорошим результатом.