Ссылка на код программы: `https://github.com/tanyarubtsova/Homework/blob/master/Huffman/Huffman.py`

Для корректной работы программы требуются пакеты time, sys, bitarray, pickle.

Программа поддерживает работу в двух режимах (предусмотрен как интерактивный метод выбора режима, так и через аргумент программы) :

- enc – шифрование (сжатие) файла.

  На вход подаётся путь к файлу (предусмотрен как интерактивный метод ввод имени файла, так и через аргументы программы).

  В процессе работы программы вычисляется количество встречаний каждого байта в файле, по этим значениям строится код Хаффмана. Далее создаётся файл encoded_file.mzh, содержащий таблицу-словарь, состоящий из пар (байт, код для этого байта), и зашифрованное кодом Хаффмана содержимое исходного файла. Также в стандартный поток вывода выводится время работы программы.

- dec – расшифрование файла.

  На вход подаётся путь к файлу с зашифрованным кодом Хаффмана файлом (предусмотрен как интерактивный метод ввод имени файла, так и через аргументы программы).

  В результате работы создаётся файл decoded_file.ext, содержащий расшифрованное содержимое рассматриваемого файла. Также в стандартный поток вывода выводится время работы программы.

Анализ:

Файл	Размер файла	Размер zmh	Время сжатия zmh	Размер zip	Размер rar
text1.txt	26б	272б	0,002с	138б	100б
text2.txt	18132Кб	3468Кб	3,014с	85Кб	45Кб
text3.txt	10452Кб	6040Кб	1,893с	82Кб	9Кб
pic1.png	59Кб	64Кб	0,047с	49Кб	100б
pic2.png	307Кб	312Кб	0,047с	307Кб	307Кб
video.mov	13984Кб	13989Кб	2,031с	13,7Мб	17,9Мб
text4.pdf	218Кб	222Кб	0,051с	210Кб	210Кб

В основном картинки, поскольку в них довольно равномерно встречаются все биты и нельзя выделить наиболее частые или редкие биты, сжимаются плохо. Также это происходит потому, что многие форматы уже ориентированы на занятие как можно меньшего размера (например png или mp4). Аналогичная ситуация с pdf-файлами, аудиофайлами и видеофайлами. К тому же в zmh есть накладные расходы за счет хранения словаря. Таким образом, данная реализация zmh сжимает данные, но работает не лучше существующих решений.