

ЦЕЛЬ РАБОТЫ

Углубление теоретических знаний в области оптимального кодирования данных в информационных системах, исследование способов построения таблиц кодирования первичных кодов и простейших методов сжатия символьных последовательностей, приобретение практических навыков исследования процессов кодирования информационных сообщений.

ПРОГРАММА ЛАБОРАТОРНОЙ РАБОТЫ

1. Набрать в текстовом редакторе (Блокноте) строку произвольного сообщения размером 10-15 символов и сохранить ее в файле.
2. Открыть сохраненный файл в Total Commander в режиме просмотра (F3) и найти кодировки, в которых происходит правильное отображение текста.
3. Записать закодированную строку в 16-ричном коде.
4. Найти символы кодируемой строки в таблице CP-1251, выписать их десятичные коды и представить их в двоичном виде. Сравнить эти коды с представлением символов в 16-ричном коде.
5. Посмотреть кодируемую строку при кодировке ASCII/DOS, выписать 16-коды символов и сравнить их с кодами соответствующих символов при использовании кодовой страницы CP-1251.
6. Выполнить пункт 4.5 при кодовой странице KOI8-R и пояснить причину неверного отображения закодированной строки.
7. Вычислить объем видеофайла, содержащего данные для отображения на экране дисплея с разрешающей способностью 800×600 изображения, в котором на синем фоне в центре экрана располагается красный прямоугольник размером 20×20 пикселей.

8. Закодировать содержимое видеофайла методом RLE и определить объем сжатого файла и рассчитать коэффициент компрессии.

ХОД РАБОТЫ

1. Наберем в текстовом редакторе строку произвольного сообщения размером 10-15 символов и сохраним ее в файле. Файл верно открывается в кодировке UTF-8, а в кодировке KOIR-R и ASCII выглядит неверно.

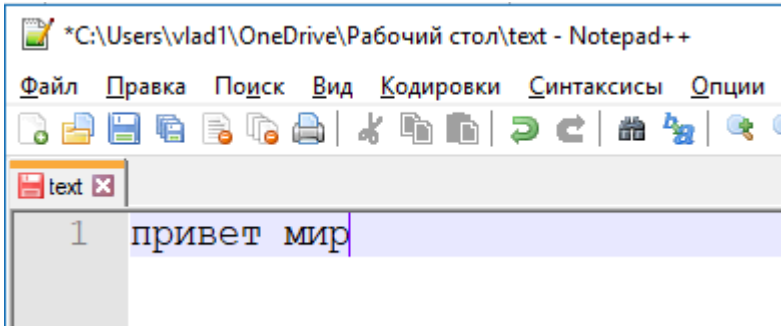


Рисунок 1 – Создание файла

2. Записать закодированную строку в 16-ричном коде.

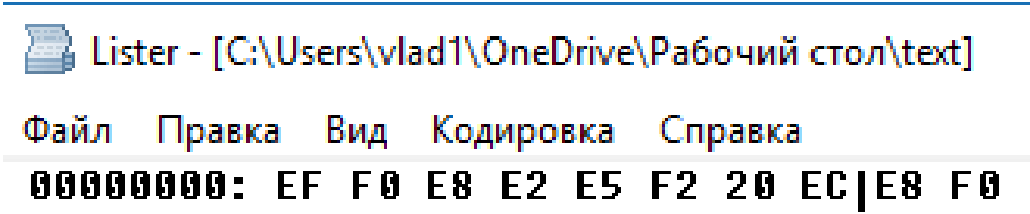


Рисунок 2 – Запись строки в шестнадцатеричном формате

4. Таблица символов в кодировке CP-1251:

Символ	HEX	DEC	BIN
п	EF	239	11101111
р	F0	240	11110000
и	E8	232	11101000
в	E2	226	11100010
е	E5	229	11100101
т	F2	242	11110010
Пробел	20	119	1110111
м	EC	236	11101100
и	E8	232	11101000
р	F0	240	11110000

5. Таблица символов в кодировке ANSI:

Символ	HEX	DEC	BIN
п	EF	239	11101111
р	F0	240	11110000
и	E8	232	11101000
в	E2	226	11100010
е	E5	229	11100101
т	F2	242	11110010
Пробел	20	119	1110111
м	EC	236	11101100
и	E8	232	11101000
р	F0	240	11110000

6. Отображение строки в кодировке KOI8-R.

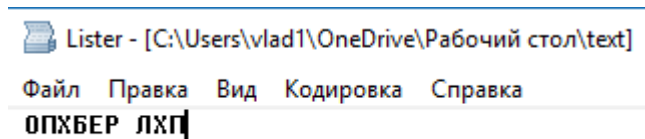


Рисунок 3 – Вид строки в кодировке KOI8-R

7. Вычисление объема видеофайла.

Без кодировки объем видеофайла будет равен:

Всего точек на экране (разрешающая способность): $600 * 800 = 480\,000$

$24 * 480\,000 = 11\,520\,000$ бит = $1\,440\,000$ байт = 1,44 Мбайт

8. Объём видеофайла методом RLE.

390 строк: C800 0000FF

20 строк: C390 0000FF C20 FF0000 C390 0000FF

390 строк: C800 0000FF

Получаем объем: $780 * 10 \text{ байт} + 20 * 29 \text{ байт} = 8380 \text{ байт}$

Коэффициент сжатия: 171 раз.

ВЫВОДЫ

В ходе лабораторной работы были углублены теоретические знания в области оптимального кодирования данных в информационных системах. Также были исследованы способы построения таблиц кодирования первичных кодов и простейших методов сжатия символьных последовательностей.