

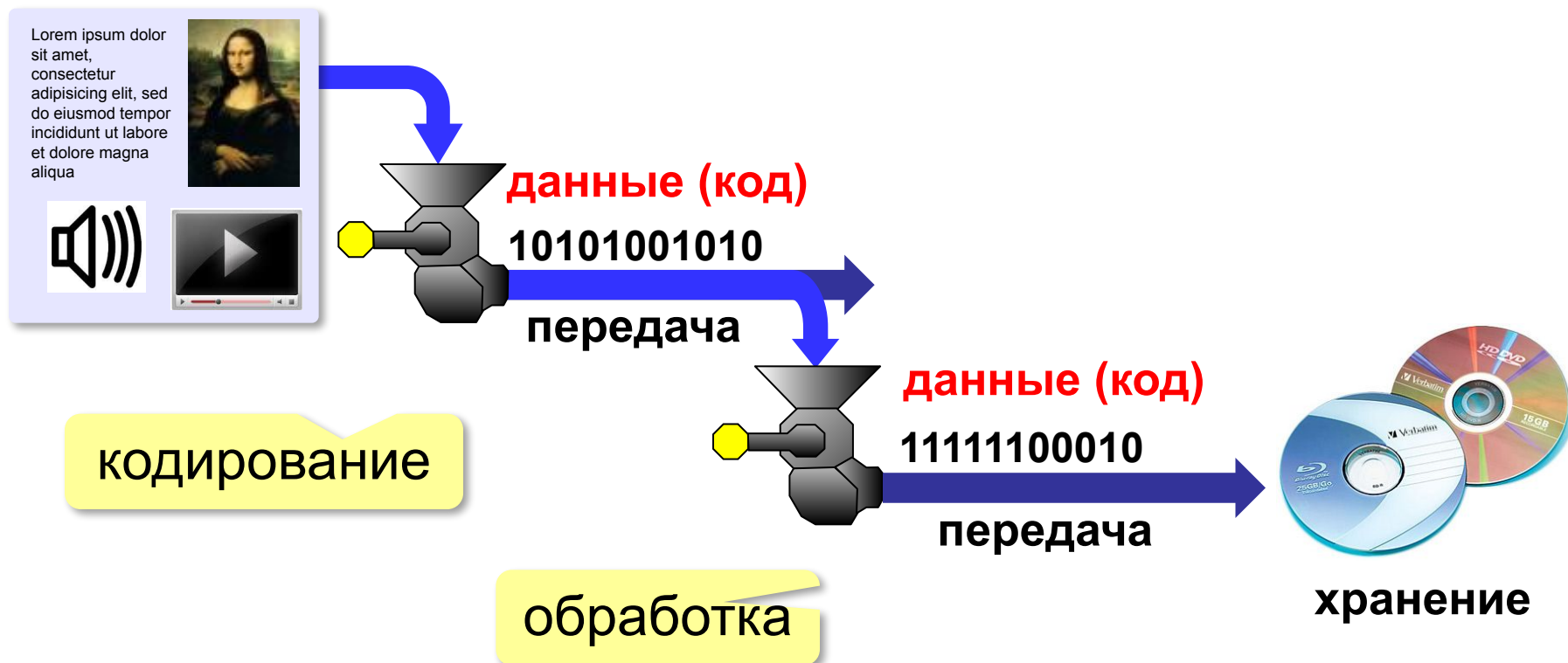
Кодирование информации

- § 15. Кодирование символов
- § 16. Кодирование графической информации
- § 17. Кодирование звуковой и видеоинформации

Зачем кодировать информацию?

Кодирование — это представление информации в форме, удобной для её хранения, передачи и обработки.

В компьютерах используется двоичный код:

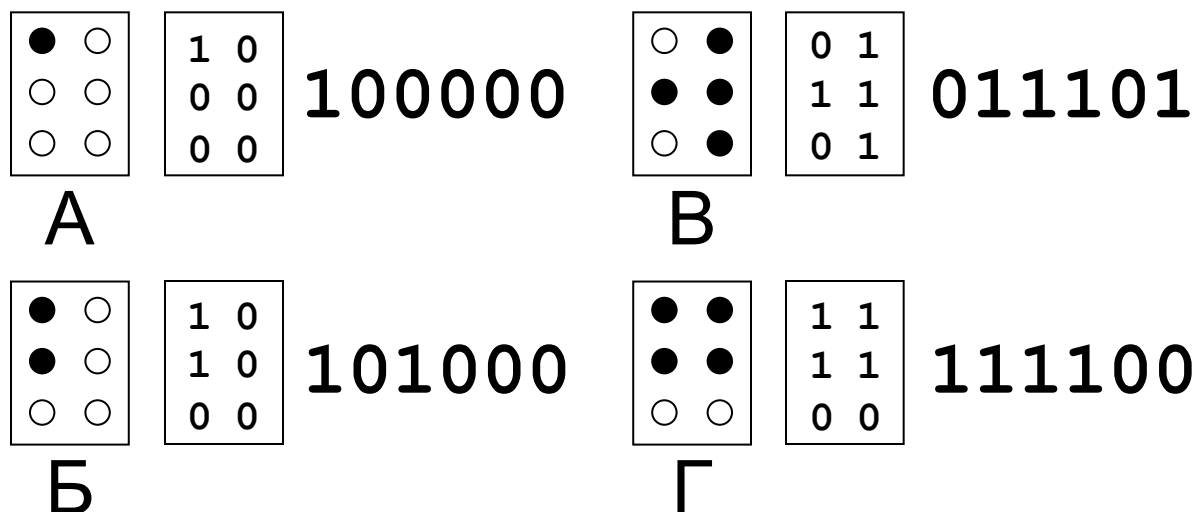


Кодирование информации

§ 15. Кодирование символов

Кодирование символов

Система Брайля:



Общий подход:

- нужно использовать N символов
- выберем число битов k на символ: $2^k \geq N$
- сопоставим каждому символу код – число от 0 до $2^k - 1$
- переведем коды в двоичную систему



Откуда формула?

Кодирование символов

Текстовый файл

- на экране (символы)
- в памяти — коды



1000001 ₂	1000010 ₂	1000011 ₂	1000100 ₂
65	66	67	68



В файле хранятся не изображения символов, а их числовые коды!

Файлы со шрифтами: ***.fon**, ***.ttf**, ***.otf**

Кодировка ASCII (7-битная)

ASCII = *American Standard Code for Information Interchange*

Коды 0-127:

0-31 **управляющие символы:**

7 – звонок, 10 – новая строка,

13 – возврат каретки, 27 – Esc.

32 пробел

знаки препинания: . , : ; ! ?

специальные знаки: + – * / () { } []

48-57 цифры **0..9**

65-90 заглавные латинские буквы **A-Z**

97-122 строчные латинские буквы **a-z**



Где русские буквы?

8-битные кодировки

Кодовые страницы (расширения ASCII):

0	1	127	128	254	255
		таблица ASCII		национальный алфавит	

Для русского языка:

CP-866 для *MS DOS*

CP-1251 для *Windows* (Интернет)

KOI8-R для *UNIX* (Интернет)

MacCyrillic для компьютеров *Apple*

Проблема:

Windows-1251

Привет, Вася!

рТЙЧЕФ,

чБУС!



KOI8-R

оПХБЕР,

бЮЯЪ!

Привет, Вася!

8-битные кодировки

- 
 - 1 байт на символ – файлы небольшого размера!
 - просто обрабатывать в программах
- 
 - нельзя использовать символы разных кодовых страниц одновременно (русские и французские буквы, и т.п.)
 - неясно, в какой кодировке текст (перебор вариантов!)
 - для каждой кодировки нужен свой шрифт (изображения символов)

Стандарт UNICODE

1 112 064 знаков, используются около **100 000**

Windows: **UTF-16**

16 битов на распространённые символы,
32 бита на редко встречающиеся

Linux: **UTF-8**

8 битов на символ для ASCII,
от 16 до 48 бита на остальные



- совместимость с ASCII
- более экономична, чем UTF-16, если много символов ASCII



2010 г. – 50% сайтов использовали UTF-8!

Знания для решения задач

- русский алфавит
- принципы работы с числами, записанными в позиционных системах счисления
- если слово состоит из L букв, причем есть n_1 вариантов выбора первой буквы, n_2 вариантов выбора второй буквы и т.д., то число возможных слов вычисляется как произведение

$$N = n_1 \cdot n_2 \cdot \dots \cdot n_L$$

- если слово состоит из L букв, причем каждая буква может быть выбрана n способами, то число возможных слов вычисляется как

$$N = n^L$$

ЕГЭ 10 задание

Вася составляет 3-буквенные слова, в которых есть только буквы В, Е, С, Н, А, причём буква А используется в каждом слове хотя бы 1 раз. Каждая из других допустимых букв может встречаться в слове любое количество раз или не встречаться совсем. Словом считается любая допустимая последовательность букв, не обязательно осмысленная.

Сколько существует таких слов, которые может написать Вася?

Решение (способ 1):

1. буква А может стоять на одном из трёх мест: А**, *А*, **А, где * обозначает любой

Решение (способ 2):

1. количество слов с буквой А можно вычислить как разность между количеством всех возможных слов и количеством слов, в которых нет буквы А
2. для шаблона А* получаем (перемножая количество вариантов для каждой позиции) $1 \cdot 5 \cdot 5 = 25$ слов
3. для шаблона *А* тоже получим 25 слов, но нужно учесть, что все слова, в которых первая буква А мы уже подсчитали, поэтому считаем только слова, где первая буква не А
4. для шаблона **А получим 25 слов, но нужно учесть, что все слова, в которых первая буква А мы уже подсчитали, поэтому считаем только слова, где первая буква не А
5. количество слов, в которых нет буквы А равно $4^3 = 64$ (на любой из 3-х позиций может стоять любая из 4 букв кроме А)
6. получается $25 + 25 + 25 - 64 = 61$ слово, в котором есть буква А (она или несколько)
7. рассматривая шаблон А*, не учитывая уже подсчитанные слова, в которых буква А есть на первом или втором местах, количество новых слов – $4 \cdot 4 \cdot 1 = 16$
8. всего получается $25 + 20 + 16 = 61$ слово Ответ: 61.

ЕГЭ 10 задание

Р-06. Вася составляет 5-буквенные слова, в которых есть только буквы С, Л, О, Н, причём буква С используется в каждом слове ровно 1 раз. Каждая из других допустимых букв может встречаться в слове любое количество раз или не встречаться совсем. Словом считается любая допустимая последовательность букв, не обязательно осмысленная. Сколько существует таких слов, которые может написать Вася?

Решение:

- 1.буква С может стоять на одном из пяти мест: $C****$, $*C***$, $**C**$, $***C*$ и $****C$, где * обозначает любой из оставшихся трёх символов
- 2.в каждом случае в остальных четырёх позициях может быть любая из трёх букв Л, О, Н, поэтому при заданном расположении буквы С имеем $3^4 = 81$ вариант
- 3.всего вариантов $5 \cdot 81 = 405$.
Ответ: 405.

ЕГЭ 10 задание

Р-05. Сколько существует различных символьных последовательностей длины 5 в четырёхбуквенном алфавите {А, С, G, Т}, которые содержат ровно две буквы А?

Решение :

1. рассмотрим различные варианты слов из 5 букв, которые содержат две буквы А и начинаются с А: **АА*** А*А** А**А* А***А**

Здесь звёздочка обозначает любой символ из набора {С, G, Т}, то есть один из трёх символов. итак, в каждом шаблоне есть 3 позиции, каждую из которых можно заполнить тремя способами, поэтому общее число комбинаций (для каждого шаблона!) равно $3^3 = 27$ всего 4 шаблона, они дают $4 \cdot 27 = 108$ комбинаций

2. теперь рассматриваем шаблоны, где первая по счёту буква А стоит на второй позиции, их всего три: ***АА** *А*А* *А**А**
они дают $3 \cdot 27 = 81$ комбинацию

3. два шаблона, где первая по счёту буква А стоит на третьей позиции:
****АА* **А*А** они дают $2 \cdot 27 = 54$ комбинации

4. и один шаблон, где сочетание АА стоит в конце *****АА**
они дают 27 комбинаций, всего получаем $(4 + 3 + 2 + 1) \cdot 27 = 270$ комбинаций
Ответ: 270.

ЕГЭ 10 задание (самостоятельно 2 б)

Р-04. Сколько слов длины 5, начинающихся с гласной буквы, можно составить из букв Е, Г, Э? Каждая буква может входить в слово несколько раз. Слова не обязательно должны быть осмысленными словами русского языка.

Решение:

1. первая буква слова может быть выбрана двумя способами (Е или Э), остальные – тремя
2. общее число различных слов равно
 $2 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 162$
ответ: 162.

ЕГЭ 10 задание

Р-01. Все 5-буквенные слова, составленные из 5 букв А, К, Л, О, Ш, записаны в алфавитном порядке.

Вот начало списка:

1. ААААА
2. ААААК
3. ААААЛ
4. ААААО
5. ААААШ
6. АААКА

.....

На каком месте от начала списка стоит слово ШКОЛА?

Решение:

1. Будем использовать пятеричную систему счисления с заменой $A \rightarrow 0$, $K \rightarrow 1$, $L \rightarrow 2$, $O \rightarrow 3$ и $Sh \rightarrow 4$
2. слово ШКОЛА запишется в новом коде так: 41320_5
3. переводим это число в десятичную систему:
4. $41320_5 = 4 \cdot 5^4 + 1 \cdot 5^3 + 3 \cdot 5^2 + 2 \cdot 5^1 = 2710$
5. поскольку нумерация элементов списка начинается с 1, а числа в пятеричной системе – с нуля, к полученному результату нужно прибавить 1, тогда... Ответ: 2711.

ЕГЭ 10 задание (самостоятельно 26)

Р-00. Все 5-буквенные слова, составленные из букв А, О, У, записаны в **обратном алфавитном порядке**. Вот начало списка:

1. УУУУУ
2. УУУУО
3. УУУУА
4. УУУОУ

.....

Запишите слово, которое стоит на 240-м месте от начала списка.

1. выпишем начало списка, заменив буквы на цифры так, чтобы **порядок символов был обратный алфавитный** ($У \rightarrow 0$, $О \rightarrow 1$, $А \rightarrow 2$):
2. переведем 239 в троичную систему: $239 = 22212_3$
3. заменяем обратно цифры на буквы, **учитывая обратный алфавитный порядок** ($0 \rightarrow У$, $1 \rightarrow О$, $2 \rightarrow А$): $22212 \rightarrow АААОА$

Ответ: АААОА.

Все четырёхбуквенные слова, составленные из букв Р, Е, К, А, записаны в алфавитном порядке и пронумерованы, начиная с 1. Начало списка выглядит так:

1. АААА
2. АААЕ
3. АААК
4. АААР
5. ААЕА

...

Под каким номером в списке идёт первое слово, в котором нет буквы А?

1. Заменяем буквы на цифры следующим образом: А — 0, Е — 1, К — 2, Р — 3.
2. Для нахождения первого слова, в котором нет буквы А, нужно найти номер ЕЕЕЕ, то есть 1111. Используется четырёхбуквенный алфавит, значит, нужно найти значение числа 1111_4 в десятичном виде. Им является число 85_{10} , но поскольку номер в списке на единицу больше самого числа, то остаётся добавить к получившемуся числу единицу. Получаем 86.
3. Ответ: 86.

Выполнить к следующему уроку

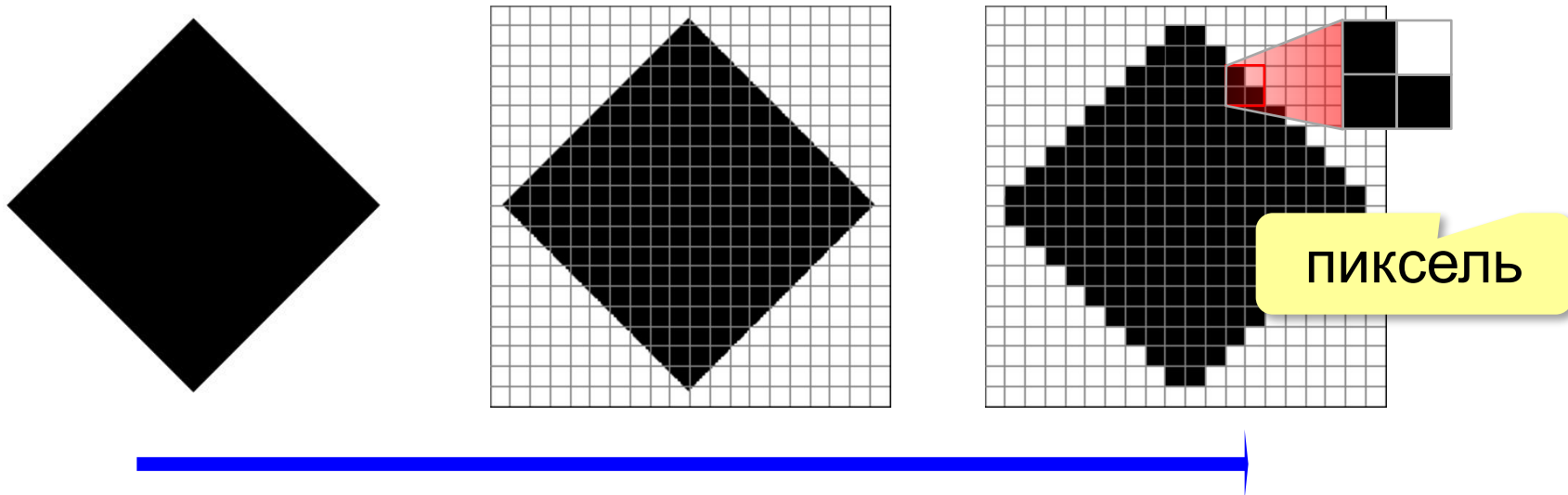
**Выполнить работу № 7785472 на Сайте
решу ЕГЭ**

<https://inf-ege.sdamgia.ru/test?id=7785472>

Кодирование информации

§ 16. Кодирование графической информации

Растровое кодирование



дискретизация

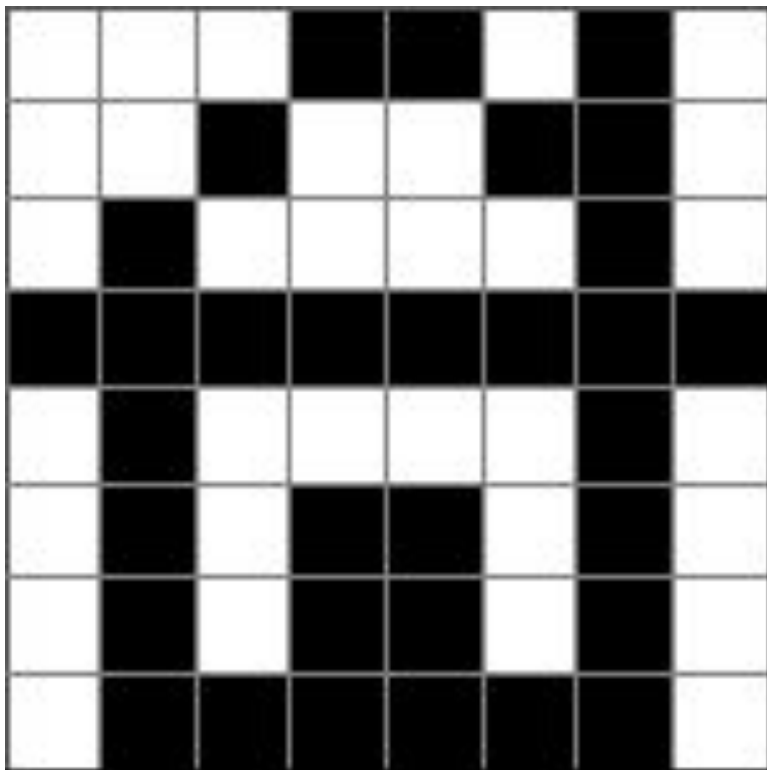


Рисунок искажается!

Пиксель – это наименьший элемент рисунка, для которого можно задать свой цвет.

Растровое изображение – это изображение, которое кодируется как множество пикселей.

Растровое кодирование

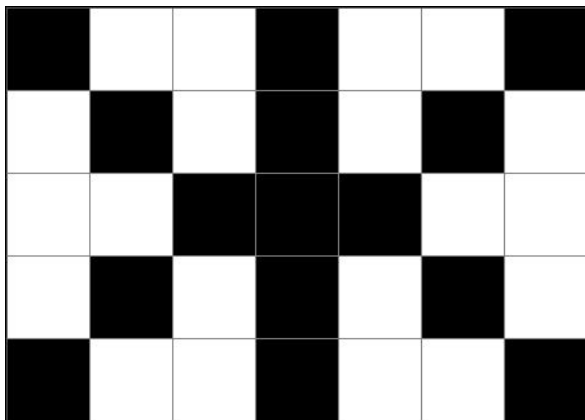


0	0	0	1	1	0	1	0	1A
0	0	1	0	0	1	1	0	26
0	1	0	0	0	0	1	0	42
1	1	1	1	1	1	1	1	FF
0	1	0	0	0	0	1	0	42
0	1	0	1	1	0	1	0	5A
0	1	0	1	1	0	1	0	5A
0	1	1	1	1	1	1	0	7E

1A2642FF425A5A7E₁₆

Задача

Закодируйте рисунок с помощью шестнадцатеричного кода:



Разрешение

Разрешение – это количество пикселей, приходящихся на дюйм размера изображения.

ppi = *pixels per inch*, пикселей на дюйм

1 дюйм = 2,54 см



300 ppi

печать

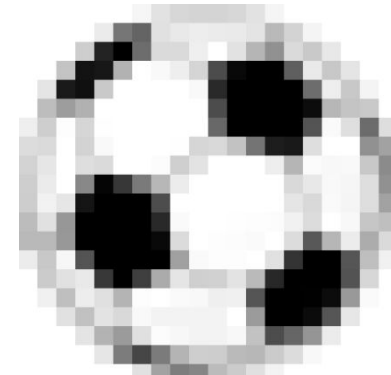


96 ppi

экран



48 ppi



24 ppi

Разрешение

Задача 1. Какой размер в пикселях должен иметь закодированный рисунок с разрешением **300 ppi**, чтобы с него можно было сделать отпечаток размером **10×15 см**?

$$\text{высота} \quad \frac{10 \text{ см} \times 300 \text{ пикселей}}{2,54 \text{ см}} \approx \mathbf{1181 \text{ пиксель}}$$

$$\text{ширина} \quad \frac{15 \text{ см} \times 300 \text{ пикселей}}{2,54 \text{ см}} \approx \mathbf{1771 \text{ пиксель}}$$

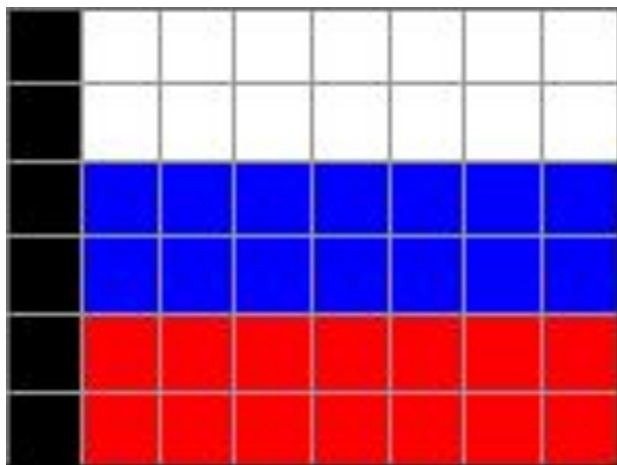
Разрешение

Задача 2. Закодированный рисунок имеет размеры 5760×3840 пикселей и разрешение 600 ppi . Какой размер будет у изображения, отпечатанного на принтере?

$$\text{ширина} \quad \frac{5760 \text{ пикселей} \times 2,54 \text{ см}}{600 \text{ пикселей}} \approx \mathbf{24,4 \text{ см}}$$

$$\text{высота} \quad \frac{3840 \text{ пикселей} \times 2,54 \text{ см}}{600 \text{ пикселей}} \approx \mathbf{16,3 \text{ см}}$$

Кодирование цвета



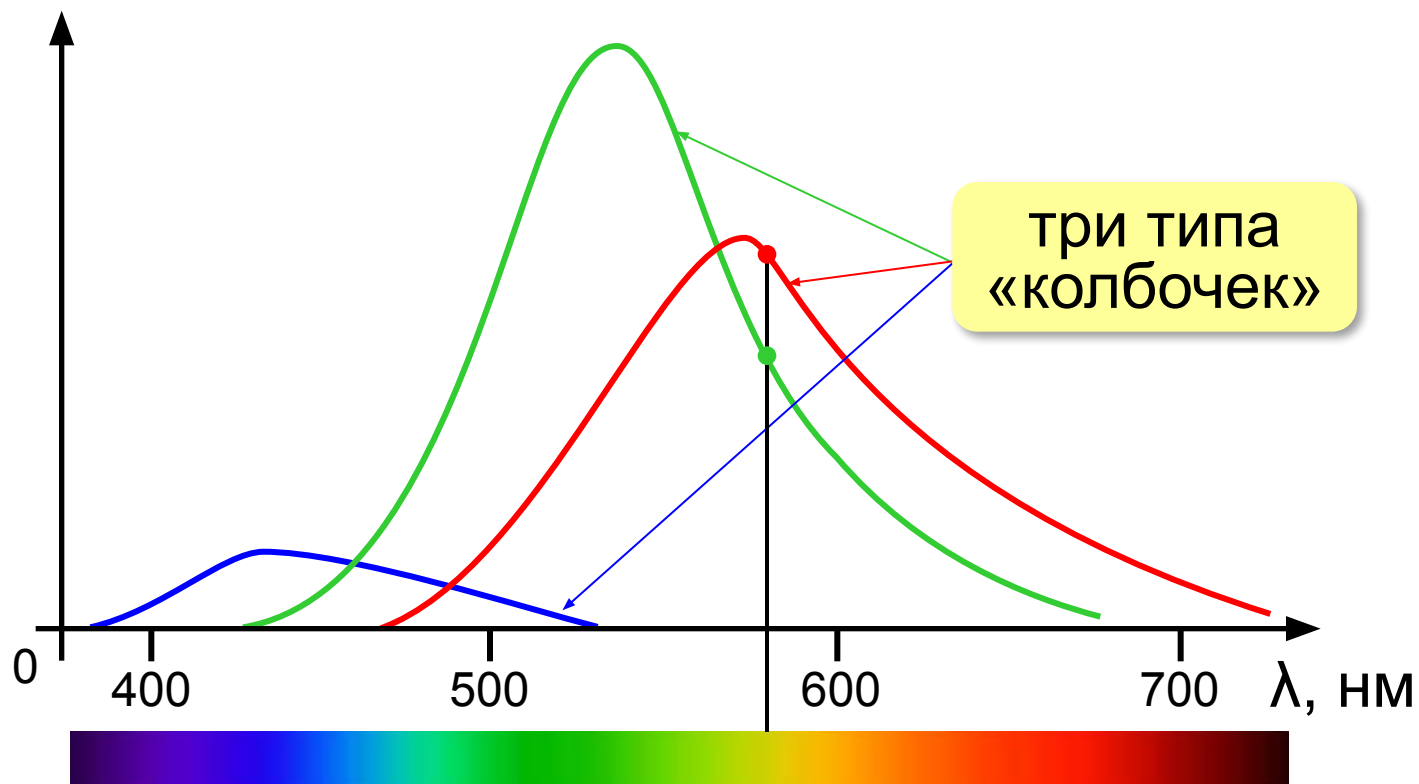
00	11	11	11	11	11	11	11
00	11	11	11	11	11	11	11
00	01	01	01	01	01	01	01
00	01	01	01	01	01	01	01
00	10	10	10	10	10	10	10
00	10	10	10	10	10	10	10

? Как выводить на монитор цвет с кодом 00?

? Как закодировать цвет в виде чисел?

Теория цвета Юнга-Гельмгольца

чувствительность



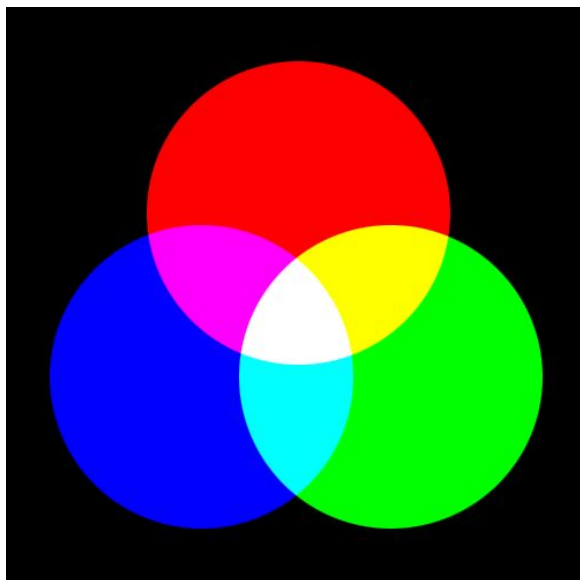
Свет любой длины волны можно заменить на красный, зелёный и синий лучи!

Цветовая модель RGB

Д. Максвелл, 1860

цвет = (**R**, **G**, **B**)

red *green* *blue*
красный зеленый синий
0..255 0..255 0..255



■ (0, 0, 0)

□ (255, 255, 255)

■ (255, 0, 0)

■ (255, 150, 150)

■ (0, 255, 0)

■ (255, 255, 0)

■ (0, 0, 255)

■ (100, 0, 0)



Сколько разных цветов можно кодировать?

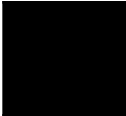


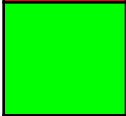



$256 \cdot 256 \cdot 256 = \mathbf{16\ 777\ 216}$ (*True Color*, «истинный цвет»)



RGB – цветовая модель для устройств, излучающих свет (мониторов)!

Цветовая модель RGB

(**255**, **255**, **0**) → **#FFFF00**

	RGB	Веб-страница
	(0, 0, 0)	#000000
	(255,255,255)	#FFFFFF
	(255, 0, 0)	#FF0000
	(0, 255, 0)	#00FF00
	(0, 0, 255)	#0000FF
	(255, 255, 0)	#FFFF00
	(204,204,204)	#CCCCCC

Задачи

Постройте шестнадцатеричные коды:

RGB (100, 200, 200) →

RGB (30, 50, 200) →

RGB (60, 180, 20) →

RGB (220, 150, 30) →

Глубина цвета

Глубина цвета — это количество битов, используемое для кодирования цвета пикселя.



Сколько памяти нужно для хранения цвета 1 пикселя в режиме *True Color*?

R (0..255) 256 = 2^8 вариантов 8 битов = 1 байт

R G B: 24 бита = 3 байта

True Color
(истинный цвет)

Задача. Определите размер файла, в котором закодирован растровый рисунок размером **20×30 пикселей** в режиме истинного цвета (*True Color*)?

$20 \cdot 30 \cdot 3 \text{ байта} = \mathbf{1800}$
байт

Кодирование с палитрой



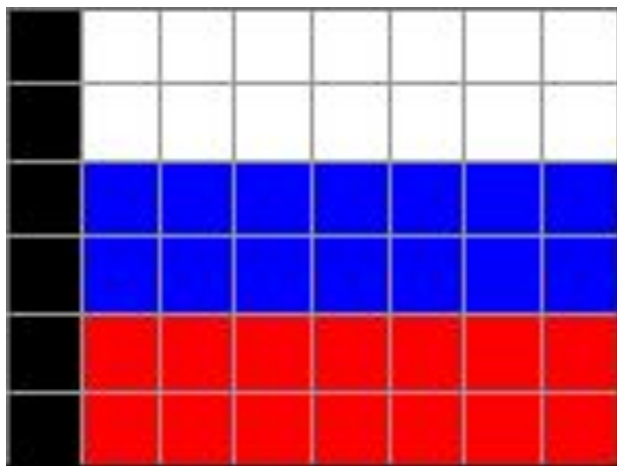
Как уменьшить размер файла?

- уменьшить разрешение
- уменьшить глубину цвета

снижается
качество

Цветовая палитра – это таблица, в которой каждому цвету, заданному в виде составляющих в модели RGB, сопоставляется числовой код.

Кодирование с палитрой



00	11	11	11	11	11	11	11
00	11	11	11	11	11	11	11
00	01	01	01	01	01	01	01
00	01	01	01	01	01	01	01
00	10	10	10	10	10	10	10
00	10	10	10	10	10	10	10

Палитра:

0	0	0	0	0	255	255	0	0	255	255	255
цвет 00 ₂			цвет 01 ₂			цвет 10 ₂			цвет 11 ₂		



Какая глубина цвета?

2 бита на пиксель



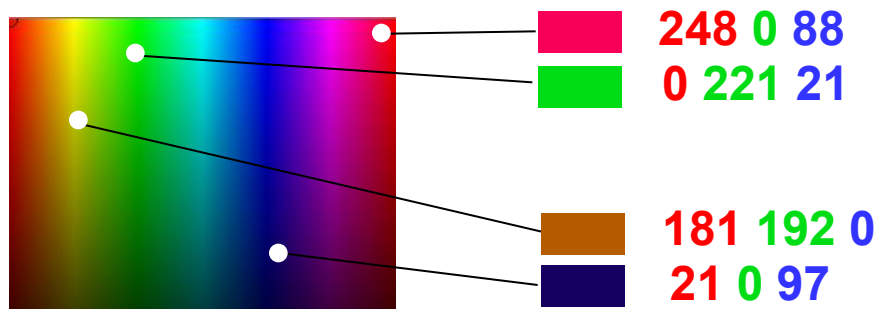
Сколько занимает палитра?

$3 \cdot 4 = 12$
байтов

Кодирование с палитрой

Шаг 1. Выбрать количество цветов: 2, 4, ... 256.

Шаг 2. Выбрать 256 цветов из палитры:



Шаг 3. Составить палитру (каждому цвету – номер 0..255)
палитра хранится в начале файла

0	1		254	255
248 0 88	0 221 21	...	181 192 0	21 0 97

Шаг 4. Код пикселя = номеру его цвета в палитре

2	45	65	14	...	12	23
---	----	----	----	-----	----	----

Кодирование с палитрой

Файл с палитрой:



коды пикселей

Количество цветов	Размер палитры (байтов)	Глубина цвета (битов на пиксель)
2	6	1
4	12	2
16	48	4
256	768	8

Что нужно знать:

1. для хранения растрового изображения нужно выделить в памяти $I = K \cdot i$ битов, где K – количество пикселей и i – глубина цвета (разрядность кодирования)
2. количество пикселей изображения K вычисляется как произведение ширины рисунка на высоту (в пикселях)
3. глубина кодирования – это количество бит, которые выделяются на хранение цвета одного пикселя
4. при глубине кодирования i битов на пиксель код каждого пикселя выбирается из 2^i возможных вариантов, поэтому можно использовать не более 2^i различных цветов
5. нужно помнить, что
 - 1 Мбайт = 2^{20} байт = 2^{23} бит,
 - 1 Кбайт = 2^{10} байт = 2^{13} бит

Выполнить и выслать на почту учителю

Задачи (ЕГЭ № 9)

















Задача 1. В процессе преобразования растрового графического файла количество цветов уменьшилось с 512 до 8. Во сколько раз уменьшился информационный объем файла?

Задача 2. Разрешение экрана монитора – 1024 x 768 точек, глубина цвета – 16 бит. Каков необходимый объем видеопамяти для данного графического режима?

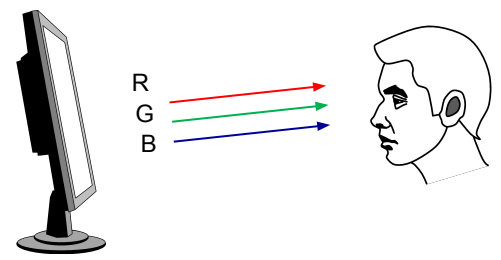
Задача 3. Сколько байт будет занимать код рисунка размером 40×50 пикселей в режиме истинного цвета? при кодировании с палитрой 256 цветов?

Задача 4. Для хранения растрового изображения размером 128 x 128 пикселей отвели 4 килобайта памяти. Каково максимально возможное число цветов в палитре изображения?

Растровые рисунки: форматы файлов

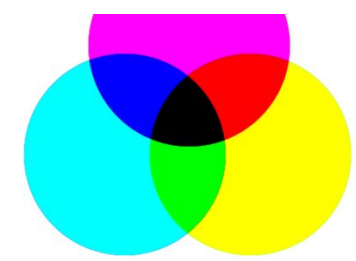
Формат	True Color	Палитра	Прозрачность	Анимация
BMP				
JPG				
GIF				
PNG				

Кодирование цвета при печати (СМУК)



Белый – красный = голубой
 Белый – зелёный = пурпурный
 Белый – синий = желтый

C = Cyan
M = Magenta
Y = Yellow



Модель CMY

C	M	Y
---	---	---

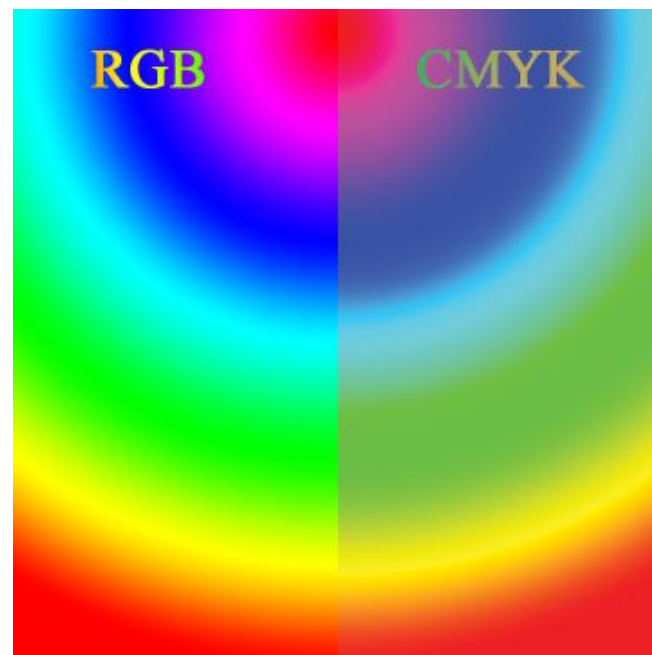
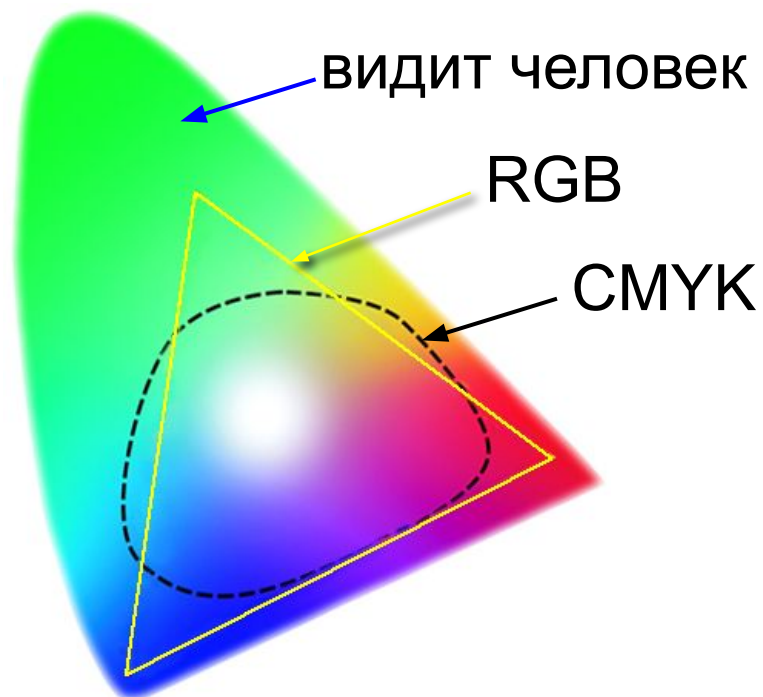
0	0	0	
255	255	0	
255	0	255	
0	255	255	
255	255	255	

Модель СМУК: + **Key color**



- меньший расход краски и лучшее качество для чёрного и серого цветов

RGB и CMYK



- не все цвета, которые показывает монитор (RGB), можно напечатать (CMYK)
- при переводе кода цвета из RGB в CMYK цвет искажается

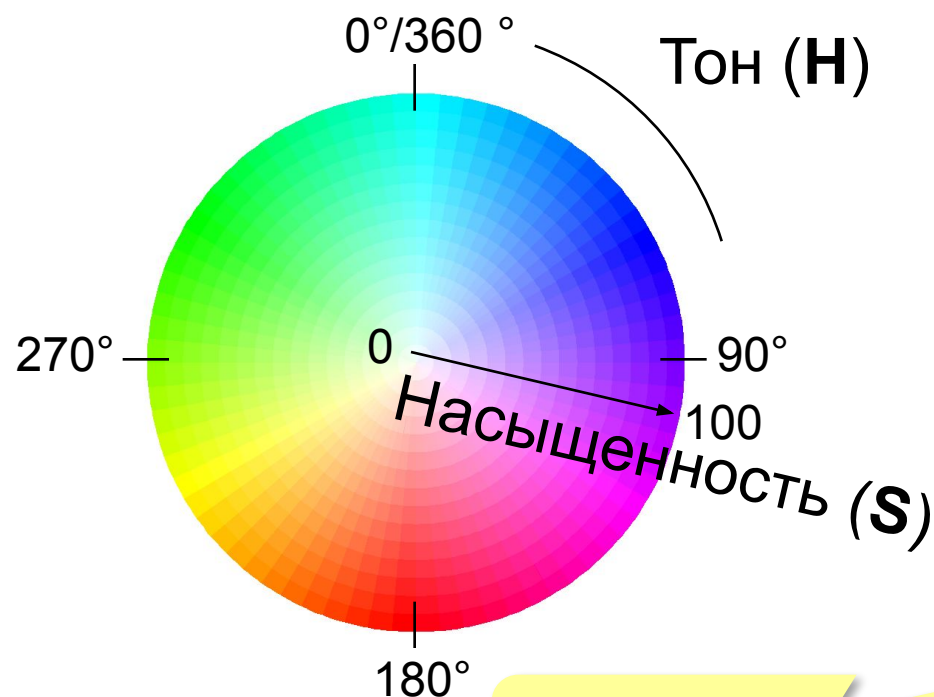
 **RGB(0,255,0)**
 **CMYK(65,0,100,0)**
→ **RGB(104,175,35)**

Цветовая модель HSB (HSV)

HSB = *Hue* (тон, оттенок)

Saturation (насыщенность)

Brightness (яркость) или *Value* (величина)



↓ яркость —
добавить чёрного

↓ насыщенность —
добавить белого

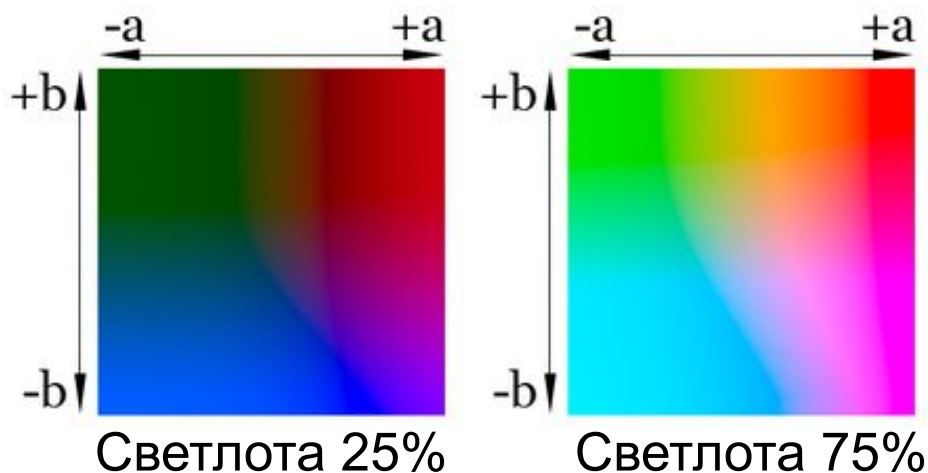
Цветовая модель Lab

Международный стандарт кодирования цвета, независимого от устройства (1976 г.)

Основана на модели восприятия цвета человеком.

Lab = *Lightness* (светлота)

a, b (задают цветовой тон)



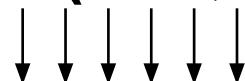
- для перевода между цветовыми моделями: RGB → Lab → CMYK
- для цветокоррекции фотографий

Профили устройств



Какой цвет увидим?

RGB(255,0,0)



RGB(255,0,0)



как $\lambda \approx 680\text{nm}$

профиль
монитора

$\lambda \approx 680\text{nm}$



RGB(225,10,20)

профиль
сканера

CMYK(0,100,100,0)



профиль
принтера

Растровое кодирование: итоги



- универсальный метод (можно закодировать любое изображение)
- единственный метод для кодирования и обработки размытых изображений, не имеющих чётких границ (фотографий)



- есть **потеря информации** (почему?)
- при изменении размеров цвет и форма объектов на рисунке **искажается**
- **размер файла** не зависит от сложности рисунка (а от чего зависит?)

Векторное кодирование

Рисунки из геометрических фигур:

- отрезки, ломаные, прямоугольники
- окружности, эллипсы, дуги
- сглаженные линии (кривые Безье)

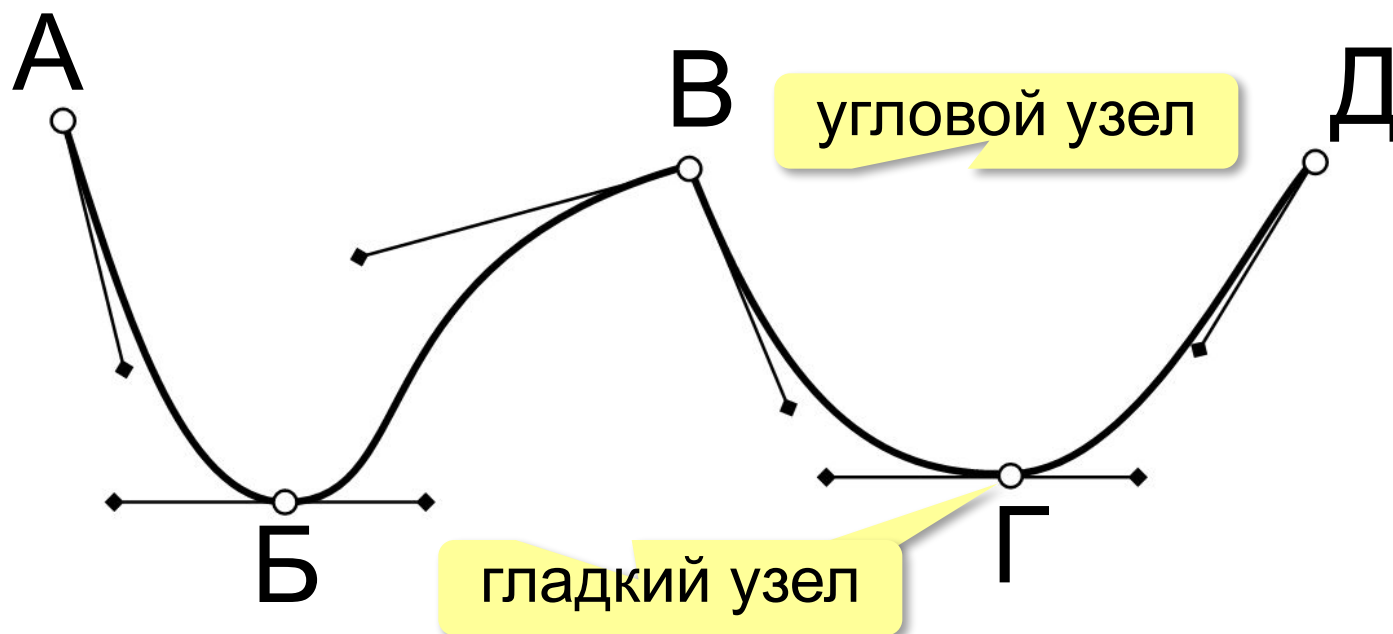
Для каждой фигуры в памяти хранятся:

- размеры и координаты на рисунке
- цвет и стиль границы
- цвет и стиль заливки (для замкнутых фигур)



Векторное кодирование

Кривые Безье:



Хранятся координаты узлов и концов «рычагов»
(3 точки для каждого узла, кривые 3-го порядка).

Векторное кодирование (итоги)



- лучший способ для хранения **чертежей, схем, карт**
- при кодировании **нет потери информации**
- при изменении размера **нет искажений**



растровый
рисунок



векторный
рисунок

- меньше **размер файла**, зависит от сложности рисунка



- неэффективно использовать для **фотографий** и размытых изображений

Векторное кодирование: форматы файлов

- **WMF** (*Windows Metafile*)
- **EMF** (*Windows Metafile*)
- **CDR** (программа *CorelDraw*)
- **AI** (программа *Adobe Illustrator*)
- **SVG** (*Scalable Vector Graphics*, масштабируемые векторные изображения)

для веб-страниц

Векторные рисунки: SVG

<svg>

прямоугольник

размеры

<rect width="135" height="30"

x="0" y="10"

координаты

stroke-width="1" stroke="rgb(0,0,0)"

контур

fill="rgb(255,255,255)"/>

заливка

<rect width="135" height="30" x="0" y="40"

stroke-width="1" stroke="rgb(0,0,0)"

fill="rgb(0,0,255)"/>

<rect width="135" height="30" x="0" y="70"

stroke-width="1" stroke="rgb(0,0,0)"

fill="rgb(255,0,0)"/>

<line x1="0" y1="0"

x2="0" y2="150"

stroke-width="15" stroke="rgb(0,0,0)"/>

</svg>

