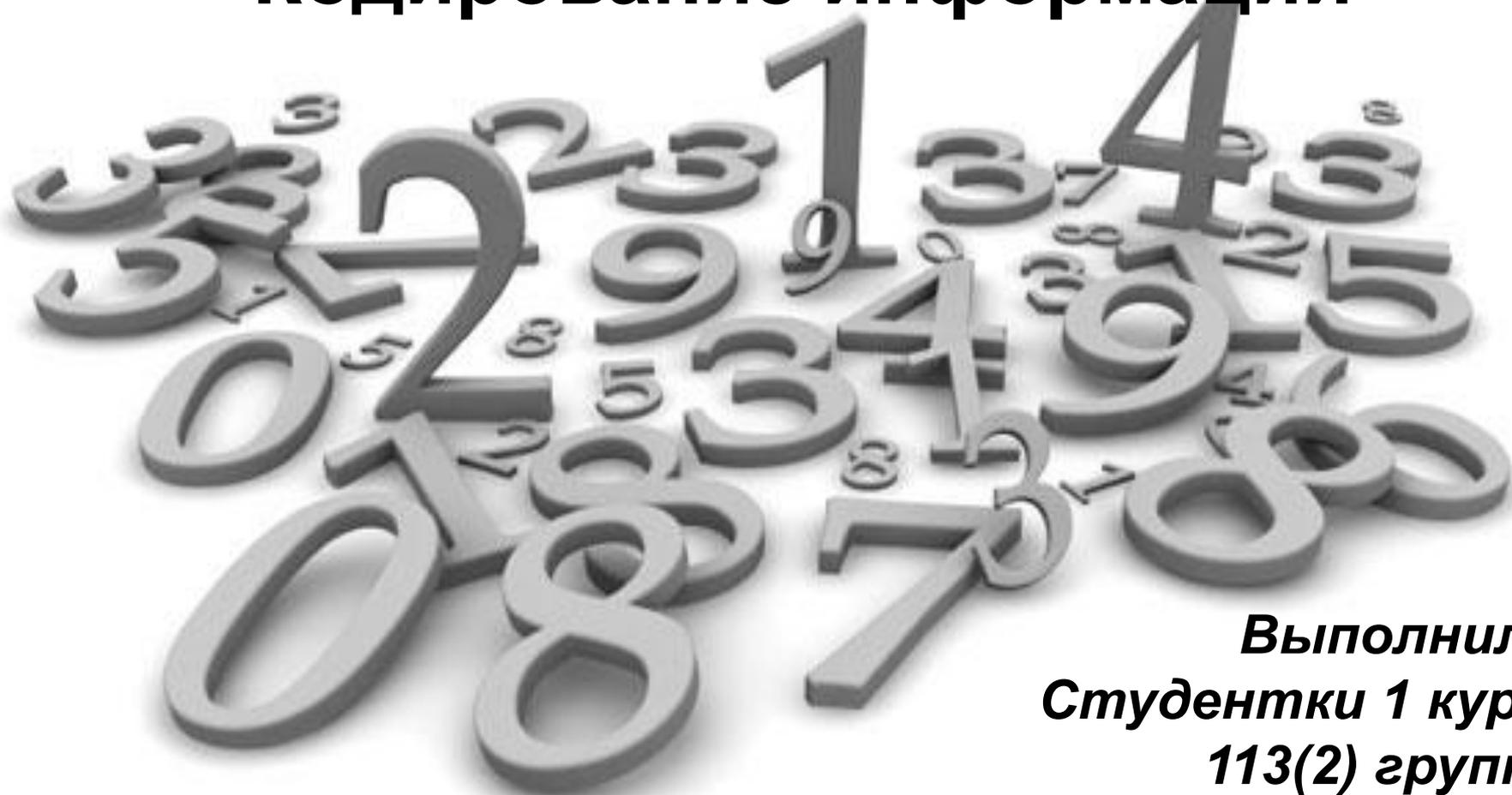


Проект на тему: Кодирование информации



Выполнили:
Студентки 1 курса
113(2) группы
Сердюк Мария
Макарова Анастасия
Потапенко Дарья

Кодирование информации – это процесс формирования определенного представления информации.



Код — это набор сигналов для записи или передачи некоторых заранее определенных понятий.

Обычно каждый образ при кодировании представлении отдельным знаком.

Знак - это элемент конечного множества отличных друг от друга элементов

Буква	Код	$p_i \cdot 10^3$	k_i	Буква	Код	$p_i \cdot 10^3$	k_i
пробел	00	174	2	я	----	18	5
о	---	90	4	ы	----	16	5
е	.	72	2	з	---	16	4
а	--	62	3	ь,ъ	----	14	5
и	..	62	3	б	----	14	5
т	-	53	2	г	---	13	4
н	--	53	3	ч	----	12	5
с	...	45	4	й	----	10	5
р	---	40	4	х	----	9	5
в	---	38	4	ж	----	7	5
л	----	35	5	ю	----	6	5
к	---	28	4	ш	----	6	5
м	--	26	2	ц	----	4	5
д	---	25	4	щ	----	3	5
п	---	23	4	э	----	3	6
у	---	21	4	ф	----	2	5

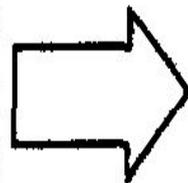
Компьютер может обрабатывать только информацию, представленную в числовой форме.

Соответствие между набором букв и числами называется **кодировкой символов**

sp 32	! 33	" 34	# 35	\$ 36	% 37	& 38	' 39	(40) 41	* 42	+ 43	, 44	- 45	. 46	/ 47
0 48	1 49	2 50	3 51	4 52	5 53	6 54	7 55	8 56	9 57	: 58	; 59	< 60	= 61	> 62	? 63
@ 64	A 65	B 66	C 67	D 68	E 69	F 70	G 71	H 72	I 73	J 74	K 75	L 76	M 77	N 78	O 79
P 80	Q 81	R 82	S 83	T 84	U 85	V 86	W 87	X 88	Y 89	Z 90	[91	\ 92] 93	^ 94	_ 95
` 96	a 97	b 98	c 99	d 100	e 101	f 102	g 103	h 104	i 105	j 106	k 107	l 108	m 109	n 110	o 111
p 112	q 113	r 114	s 115	t 116	u 117	v 118	w 119	x 120	y 121	z 122	{ 123	 124	} 125	~ 126	

Буква	Код	$p_i \cdot 10^3$	k_i	Буква	Код	$p_i \cdot 10^3$	k_i
пробел	00	174	2	я	18	5
о	90	4	ы	16	5
е	72	2	з	16	4
а	62	3	ь, ъ	14	5
и	62	3	б	14	5
т	53	2	г	13	4
н	53	3	ч	12	5
с	45	4	й	10	5
р	40	4	х	9	5
в	38	4	ж	7	5
л	35	5	ю	6	5
к	28	4	ш	6	5
м	26	2	ц	4	5
д	25	4	щ	3	5
п	23	4	э	3	6
у	21	4	ф	2	5

Вид информации
Числовая
Текстовая
Графическая
Звуковая
Видео



Двоичный код

1 0 1 1 0 0 1 1

Способы кодирования информации

Одна и та же информация может быть представлена (закодирована) в нескольких формах. Кодирование информации человек начал задолго до появления компьютеров.



Грандиозные достижения человечества - **письменность и арифметика** - есть не что иное, как система кодирования речи и числовой информации.

Двоичное кодирование

Двоичное кодирование – один из распространенных способов представления информации. Вся информация кодируется в виде слов двоичного алфавита. Двоичный алфавит состоит из двух цифр **0** и **1**





□ В такой форме можно закодировать **все виды информации**

□ Практически **нет ошибок** при передаче

□ **Компьютеру легче** обрабатывать данные



□ **Длинные** записи двоичных кодов затрудняют работу

**Кодирование символьной
(текстовой) информации.**

Основная операция, производимая над
отдельными символами текста - *сравнение*
СИМВОЛОВ.

Важные аспекты сравнения символов:

□ **Уникальность** кода

□ **Длина** этого кода

Популярные таблицы перекодировки:

- ДКОИ-8
- ASCII
- CP1251
- Unicode

ASCII Code Chart

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2		!	"	#	\$	%	&	'	()	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
80	402 Ъ	403 Ѓ	201A ,	453 ґ	201E „	2026 …	2020 †	2021 ‡	20AC €	2030 ‰	409 Љ	2039 ‹	40A Њ	40C Ќ	40B Ѝ	40F Ў
90	452 ђ	2018 ‘	2019 ’	201C “	201D ”	2022 •	2013 –	2014 —	□	2122 ™	459 Ѝ	203A ›	45A Ў	45C Ў	45B ђ	45F ѓ
A0	A0	40E Ў	45E ъ	408 Ј	A4 #	490 ґ	A6 ;	A7 §	401 Ё	A9 ©	404 €	AB «	AC ¬	AD -	AE ®	407 İ
B0	B0 °	B1 ±	406 İ	456 ï	491 ґ	B5 μ	B6 ¶	B7 ·	451 ё	2116 №	454 €	BB »	458 ј	405 S	455 s	457 İ
C0	410 А	411 Б	412 В	413 Г	414 Д	415 Е	416 Ж	417 З	418 И	419 Й	41A К	41B Л	41C М	41D Н	41E О	41F П
D0	420 Р	421 С	422 Т	423 У	424 Ф	425 Х	426 Ц	427 Ч	428 Ш	429 Щ	42A Ъ	42B Ы	42C Ь	42D Э	42E Ю	42F Я
E0	430 а	431 б	432 в	433 г	434 д	435 е	436 ж	437 з	438 и	439 й	43A к	43B л	43C м	43D н	43E о	43F п
F0	440 р	441 с	442 т	443 у	444 ф	445 х	446 ц	447 ч	448 ш	449 щ	44A ъ	44B ы	44C ь	44D э	44E ю	44F я

Unicode character set

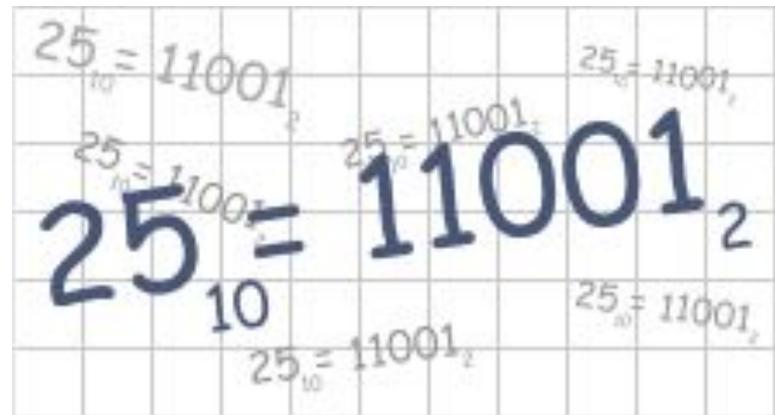
	0x00	0x01	0x02	0x03	0x04	0x05	0x06	0x07	0x08	0x09	0x0a	0x0b	0x0c	0x0d	0x0e	0x0f
0x00000	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0x00010	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0x00020	!	"	#	\$	%	&	'	()	*	+	,	-	.	/	
0x00030	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
0x00040	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0x00050	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
0x00060	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
0x00070	p	q	r	s	t	u	v	w	x	y	z	{		}	~	□
0x00080	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0x00090	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0x000a0	;	€	£	□	¥	!	§	"	©	*	«	¬	™	®	™	
0x000b0	°	±	²	³	´	μ	¶	·	¸	°	»	¼	½	¾	¿	
0x000c0	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ë	Ë	Ë	Ë
0x000d0	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
0x000e0	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ë	ë	ë	ë
0x000f0	ò	ñ	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ü	ý	þ	ÿ

В качестве длины кода для кодирования символов было выбрано **8 бит или 1 байт**.

Поэтому чаще всего одному символу текста, хранимому в компьютере, соответствует **один байт** памяти.

Кодирование числовой информации

Система счисления — это знаковая система, в которой числа записываются по определенным правилам с помощью символов некоторого алфавита, называемых цифрами.



The image shows a grid with the equation $25_{10} = 11001_2$ repeated multiple times. The central equation is the largest and most prominent, with the number 25 in base 10 on the left, an equals sign in the middle, and the binary number 11001 with a subscript 2 on the right. Several smaller, semi-transparent versions of the same equation are scattered around it, some rotated and some overlapping.

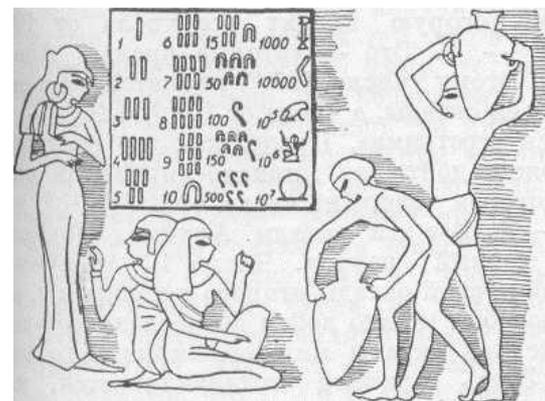
Системы счисления



ПОЗИЦИОННЫЕ



НЕПОЗИЦИОННЫЕ



Непозиционная система счисления - это система счисления, в которой от положения цифры в записи числа не зависит величина, которую она обозначает.

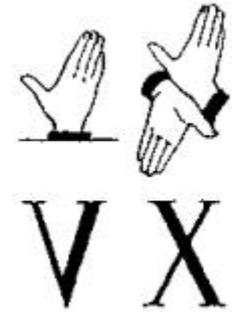


Римская система счисления

I (один палец) для числа 1

V (раскрытая ладонь) для числа 5

X (две сложенные ладони) для числа 10



В римской системе счисления количественное значение цифры не зависит от ее положения в числе.

Например, в римском числе **XXX** (**30**) цифра **X** встречается **трижды** и в каждом случае обозначает одну и ту же величину – число **10**, **три раза по 10 в сумме дают 30**.

Позиционная система счисления

Каждая позиционная система счисления имеет определенный алфавит цифр и основание.

В позиционных системах счисления количественное значение цифры зависит от **ее позиции в числе**. Позиция цифры в числе называется **разрядом**. Разряды числа возрастают справа налево, от младших разрядов к старшим, причем значение одинаковых цифр, стоящих в соседних разрядах числа, различаются на величину основания.



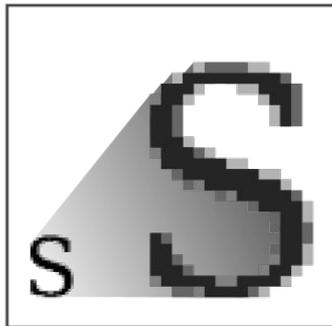
В настоящее время наиболее распространены позиционными системами счисления являются **десятичная и двоичная**.

Основной системой счисления для **представления чисел в компьютере** является **двоичная позиционная система счисления**.



Кодирование графической информации

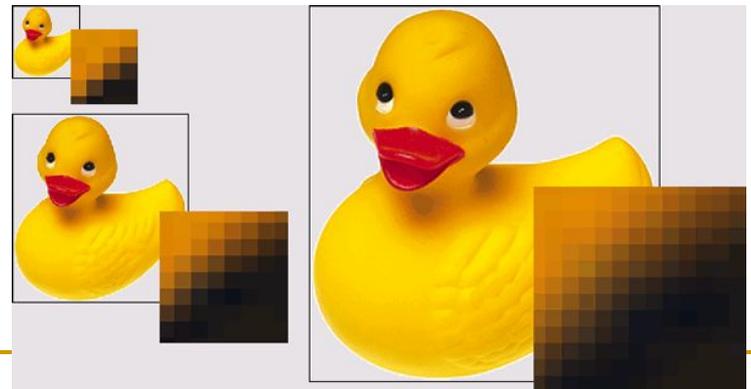
Важным этапом кодирования графического изображения является разбиение его на дискретные элементы (**дискретизация**). Основными способами представления графики для ее хранения и обработки с помощью компьютера являются **растровые и векторные** изображения.



РАСТР
.jpeg .gif .png



ВЕКТОР
.svg



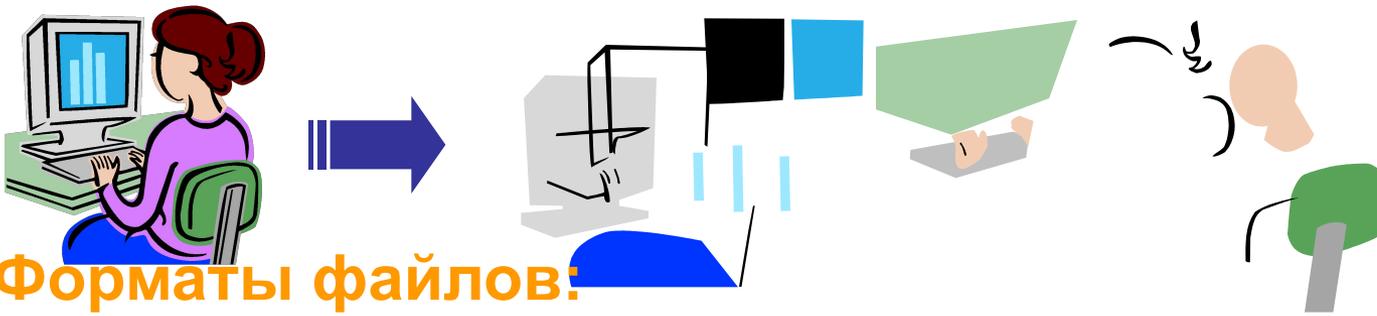
Векторные рисунки

Строятся 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)



Форматы файлов:

- **WMF** (*Windows Metafile*)
- **AI** (*Adobe Illustrator*)
- **CDR** (*CorelDraw*)
- **FH** (*FreeHand*)



- лучший способ для хранения **чертежей, схем, карт;**

- при кодировании **нет потери информации;**

- при изменении размера **нет искажений;**

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



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

Растровые рисунки

Растровое изображение представляет собой совокупность точек (пикселей), полученных в результате дискретизации изображения в соответствии с матричным принципом.

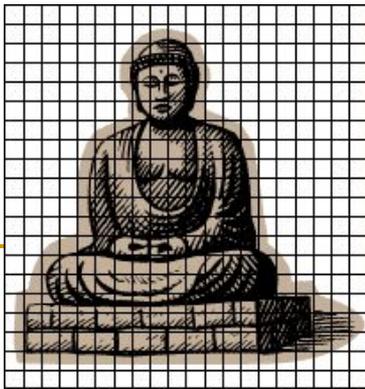
Форматы файлов (растровые рисунки):

BMP

GIF

JPG

PNG



-
-  лучший способ для хранения **фотографий** и изображений без четких границ
 - **спецэффекты** (тени, ореолы, и т.д.)

 -  есть **потеря информации** (почему?)
 - при изменении размеров рисунка он **искажается**
 - **размер файла** не зависит от сложности рисунка (а от чего зависит?)
-

Матричный принцип кодирования

графических изображений заключается в том, что изображение разбивается на заданное количество строк и столбцов. Затем каждый элемент полученной сетки кодируется по выбранному правилу.

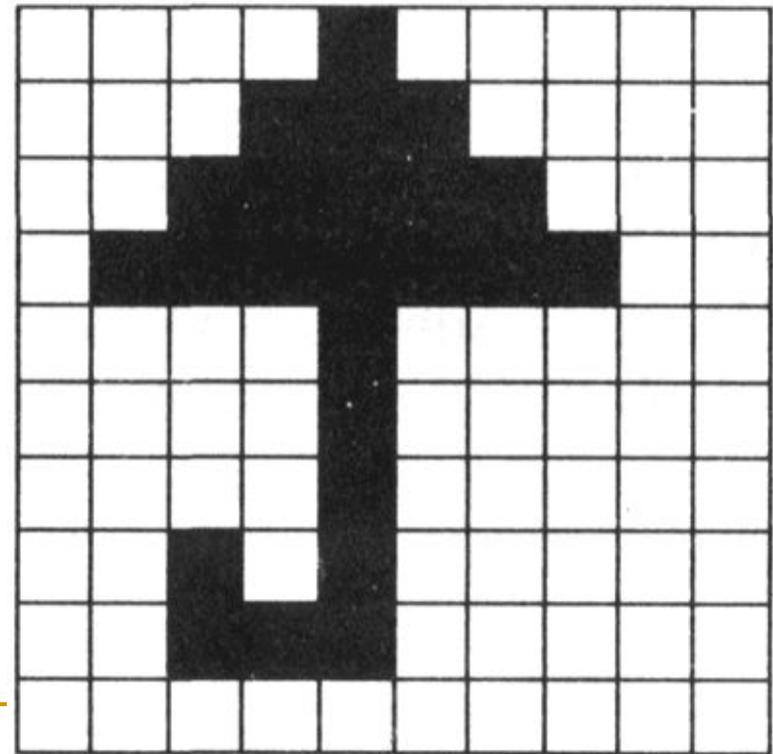
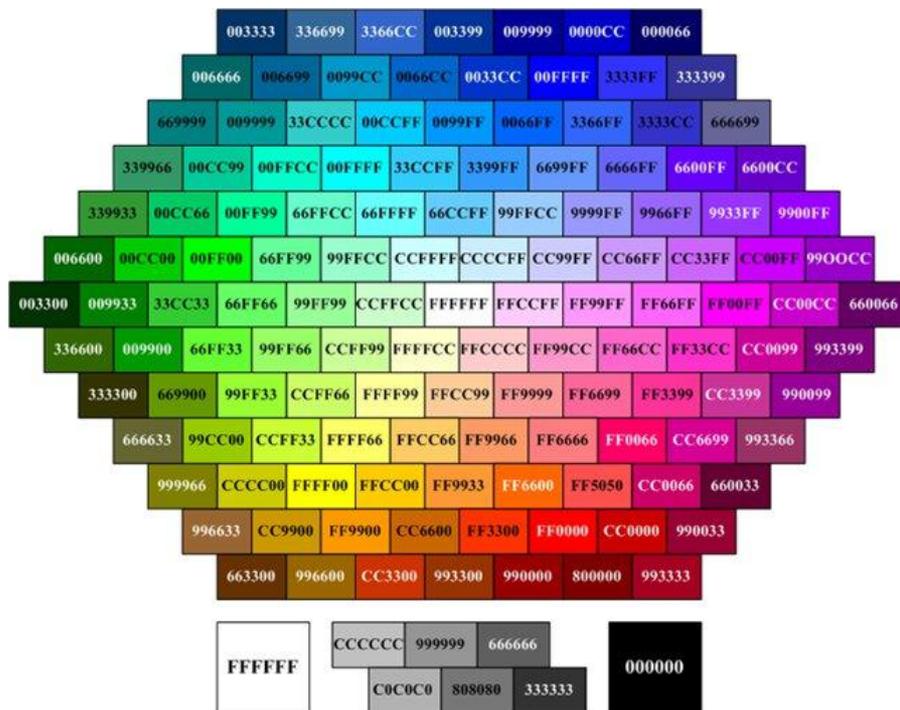
Pixel (picture element - элемент рисунка) - минимальная единица изображения, цвет и яркость которой можно задать независимо от остального изображения.



Качество изображения будет тем выше, чем "плотнее" расположены пиксели, то есть чем больше разрешающая способность устройства, и чем точнее закодирован цвет каждого из них.

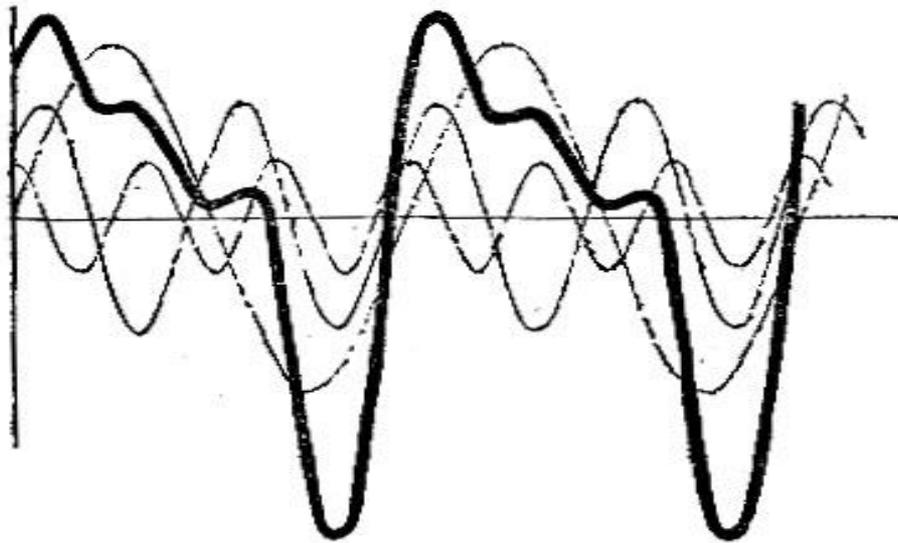
Для черно-белого изображения код цвета каждого пикселя задается **одним битом**.

Если рисунок цветной, то для каждой точки задается **двоичный код ее цвета**.



Кодирование звуковой информации

Звук - это колебания воздуха. По своей природе звук является непрерывным сигналом. Если преобразовать звук в электрический сигнал (например, с помощью микрофона), мы увидим плавно изменяющееся с течением времени напряжение.



Для компьютерной обработки аналоговый сигнал нужно каким-то образом преобразовать в последовательность двоичных чисел, а для этого его необходимо **дискретизировать и оцифровать**.

Плюс оцифровки:

- можно закодировать любой звук (в т.ч. голос, свист, шорох, ...)

Минусы оцифровки:

- есть потеря информации
- большой объем файлов



Форматы файлов:

- **WAV** (*Waveform audio format*), часто без сжатия (размер!)
 - **MP3** (*MPEG-1 Audio Layer 3*, сжатие с потерями)
 - **WMA** (*Windows Media Audio*, потоковый звук, сжатие)
-

Инструментальное кодирование

MIDI (*Musical Instrument Digital Interface*), файлы ***.MID**

в файле:

- нота (высота, длительность)
- музыкальный инструмент
- параметры звука (громкость, тембр)
- может быть несколько каналов



- **нет потери информации** при кодировании инструментальной музыки
- маленький **размер файлов**



невозможно закодировать нестандартный звук, голос

MIDI-клавиатура:



Спасибо за внимание

