Дискретное представление текстовой, графической, звуковой информации

Информация — сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые воспринимают информационные системы (живые организмы, управляющие машины и др.) в процессе жизнедеятельности и работы.

Виды информации

(в смысле науки информатики)

- ***Текстовая***
- ***Числовая***
- ***Графическая***
- ***Видеоинформация***

Свойства информации

Информация должна быть

- ***объективной*** *(не зависящей от чьего-либо мнения)*
- ***понятной*** *(английский язык?)*
- ***полезной*** *(получатель решает свои задачи)*
- ***достоверной (правильной)*** *дезинформация, помехи, слухи,*
- ***актуальной*** *– должна быть важна в данный момент (погода, землетрясение) устаревшая, ненужная*
- ***полной*** *(достаточной для принятия правильного решения)*

Дискретные модели данных в компьютере

С текстовой и графической информацией конструкторы ЭВМ «научили» работать машины, начиная с третьего поколения (1970-е гг).

А работу со звуком «освоили» лишь машины ного поколения, современные сональные компьютеры.

С этого момента началось распространение технологии мультимедиа.

Вся информация, которую обрабатывает компьютер, должна быть представлена двоичным кодом с помощью цифр 0 и 1.

Информация

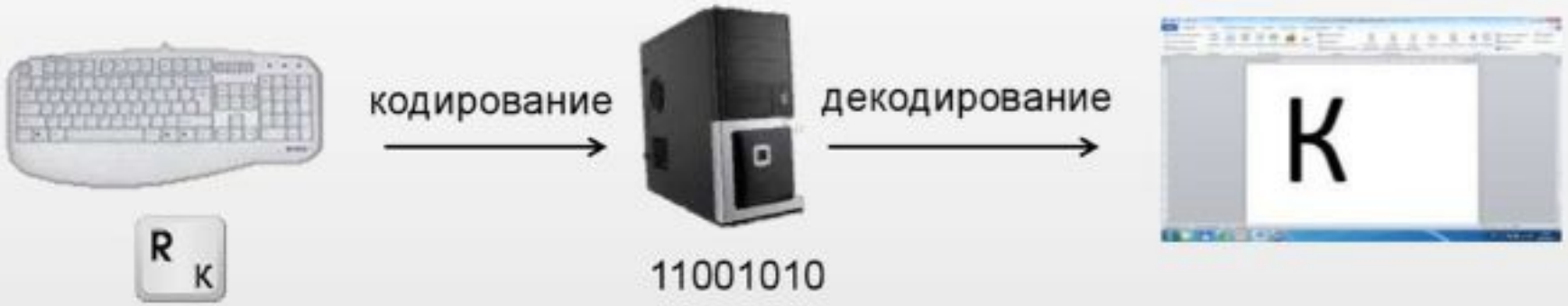
- Текстовая
- Графическая
- Звуковая
- Видео
- Числовая

Двоичный код

101010

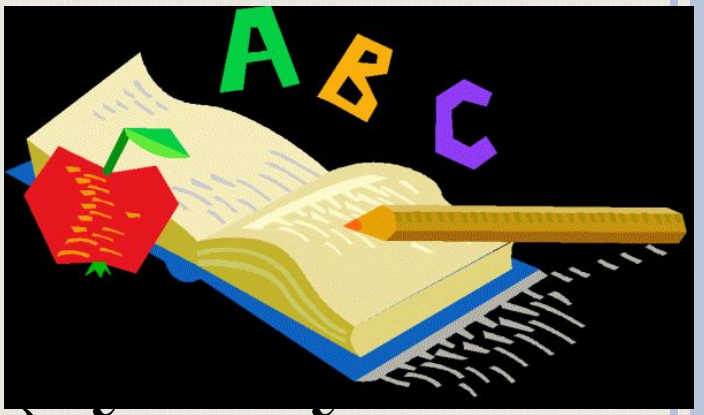
Кодирование – преобразование входной информации в форму, воспринимаемую компьютером, т.е. двоичный код.

Декодирование – преобразование данных из двоичного кода в форму, понятную человеку.



Текстовая информация

Текстовая информация состоит из отдельных знаков (символов).



<u>Символами могут быть</u>:

- ***буквы,***
- ***цифры,***
- ***знаки препинания,***
- ***знаки математических действий,***
- ***скобки и т.д.***

То есть текстовая информация уже дискретна. Поэтому возникает лишь технический вопрос, как разместить ее в памяти компьютера.

Представление текстовой информации

! *За каждой буквой* ***алфавита, цифрой, знаком препинания*** *закрепляется определенный двоичный код*, ***длина которого фиксирована.***

Символ	Двоичный код	Десятичный код	Символ	Двоичный код	Десятичный код
A	01000001	65	N	01001110	78
B	01000010	66	O	01001111	79
C	01000011	67	P	01010000	80
D	01000100	68	Q	01010001	81
E	01000101	69	R	01010010	82
F	01000110	70	S	01010011	83
G	01000111	71	T	01010100	84
H	01001000	72	U	01010101	85
I	01001001	73	V	01010110	86
J	01001010	74	W	01010111	87
K	01001011	75	X	01011000	88
L	01001100	76	Y	01011001	89
M	01001101	77	Z	01011010	90

ASCII

В системе кодировки ASCII (American Standard Code for Information Interchange - американский стандартный код для обмена информации) каждый символ заменяется на 8-разрядное целое положительное двоичное число; оно хранится в одном байте памяти.

Это число является порядковым номером символа в кодовой таблице.

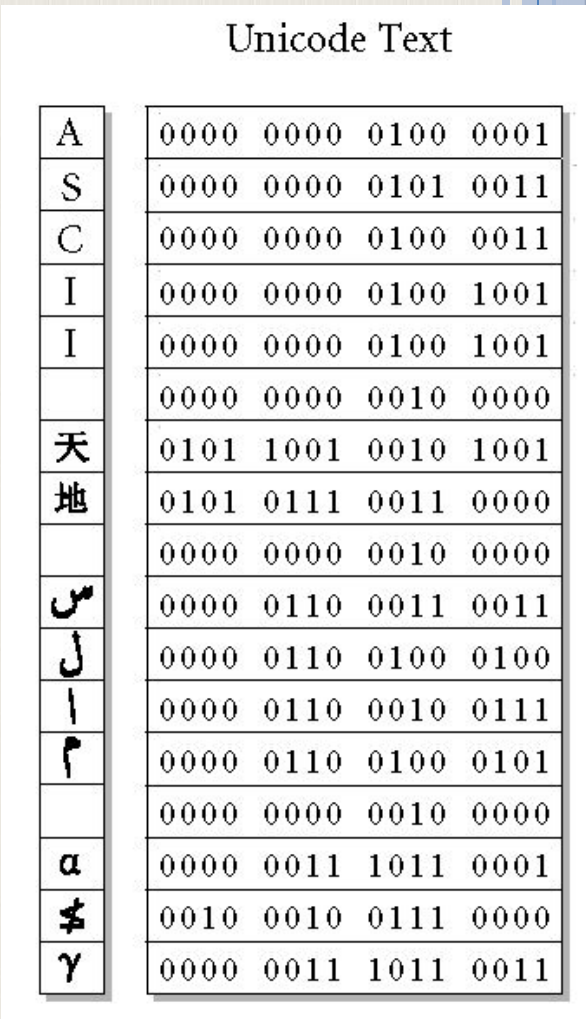
Таблица 1.1. Базовая таблица кодировки ASCII

32	пробел	48	0	64	@	80	P	96	`	112	p
33	!	49	1	65	A	81	Q	97	a	113	q
34	"	50	2	66	B	82	R	98	b	114	r
35	#	51	3	67	C	83	S	99	c	115	s
36	$	52	4	68	D	84	T	100	d	116	t
37	%	53	5	69	E	85	U	101	e	117	u
38	&	54	6	70	F	86	V	102	f	118	v
39	'	55	7	71	G	87	W	103	g	119	w
40	(	56	8	72	H	88	X	104	h	120	x
41	)	57	9	73	I	89	Y	105	i	121	y
42	*	58	:	74	J	90	Z	106	j	122	z
43	+	59	;	75	K	91	[	107	k	123	{
44	,	60	<	76	L	92	\	108	l	124	|
45	-	61	=	77	M	93	]	109	m	125	}
46	.	62	>	78	N	94	^	110	n	126	~
47	/	63	?	79	O	95	_	111	o	127	

UNICODE

Поскольку в мире много языков и много алфавитов, то постепенно совершается переход на международную 16-битовую систему кодировки Unicode.

В ней каждый символ занимает 2 байта, что обеспечивает 2^{16} = 65 536 кодов для различных символов, этого поля вполне достаточно для размещения в одной таблице символов большинства языков планеты.



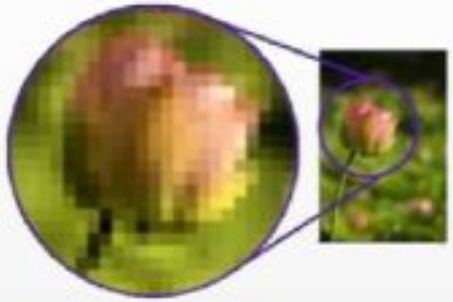
Unicode Text

A	0000 0000 0100 0001
S	0000 0000 0101 0011
C	0000 0000 0100 0011
I	0000 0000 0100 1001
I	0000 0000 0100 1001
	0000 0000 0010 0000
天	0101 1001 0010 1001
地	0101 0111 0011 0000
	0000 0000 0010 0000
س	0000 0110 0011 0011
ل	0000 0110 0100 0100
ا	0000 0110 0010 0111
م	0000 0110 0100 0101
	0000 0000 0010 0000
α	0000 0011 1011 0001
≤	0010 0010 0111 0000
γ	0000 0011 1011 0011

Графическая информация

Растровое кодирование

Рисунок, состоящий из отдельных точек



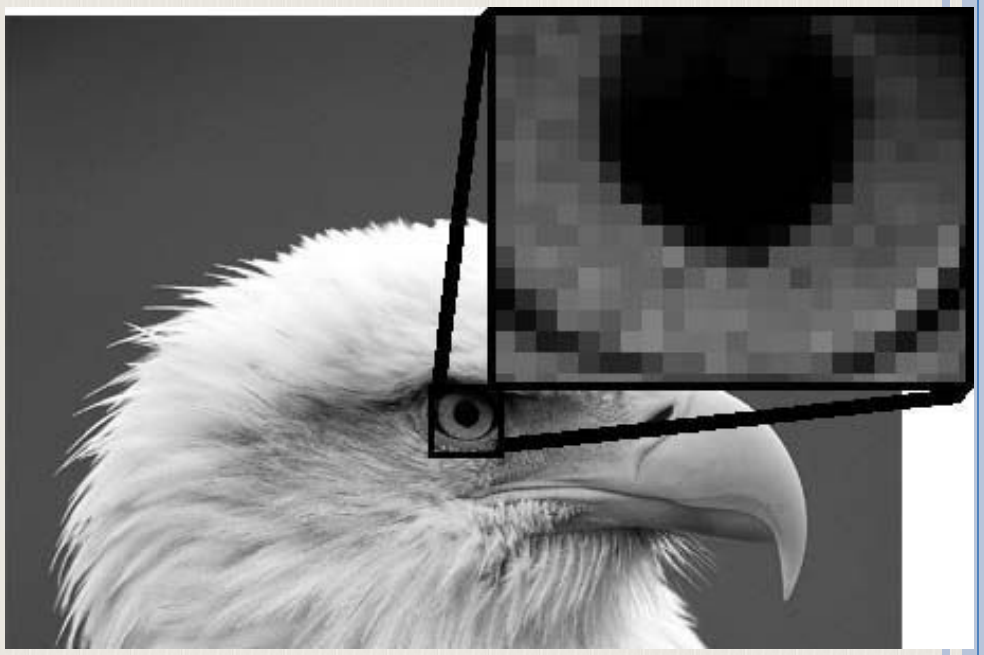
Векторное кодирование

Рисунок, состоящий из точек, прямых, дуг, овалов и т.д.

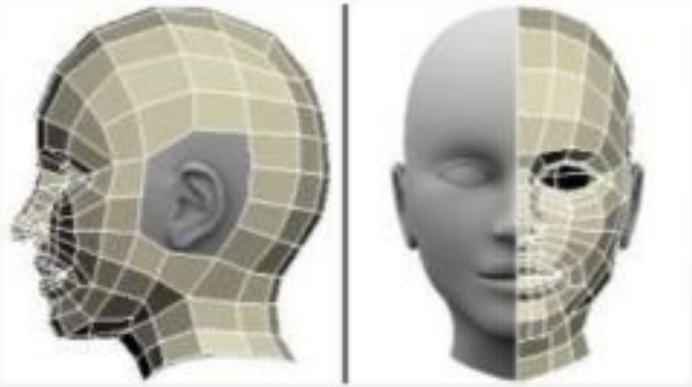
чертежи, схемы, карты

Изображение на экране монитора дискретно. Оно составляется из отдельных точек, которые называются пикселями.

Эти «точки» столь близки друг другу, что глаз не различает промежутков между ними, поэтому Изображение воспринимается как непрерывное, сплошное.



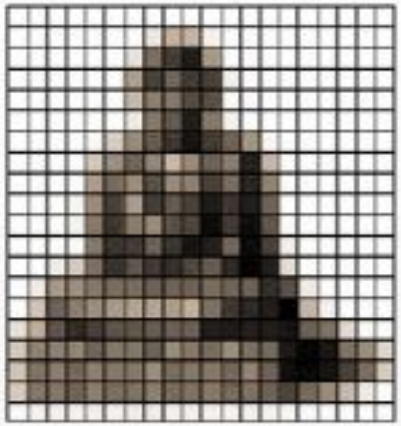
Векторные рисунки формируются из графических примитивов.



Для каждого примитива задаются:

- вид (точка, прямая, кривая, окружность и т.д)
- координаты опорных точек
- цвет заливки
- цвет, толщина и стиль линии контура

Используются двумерные и трехмерные координаты.

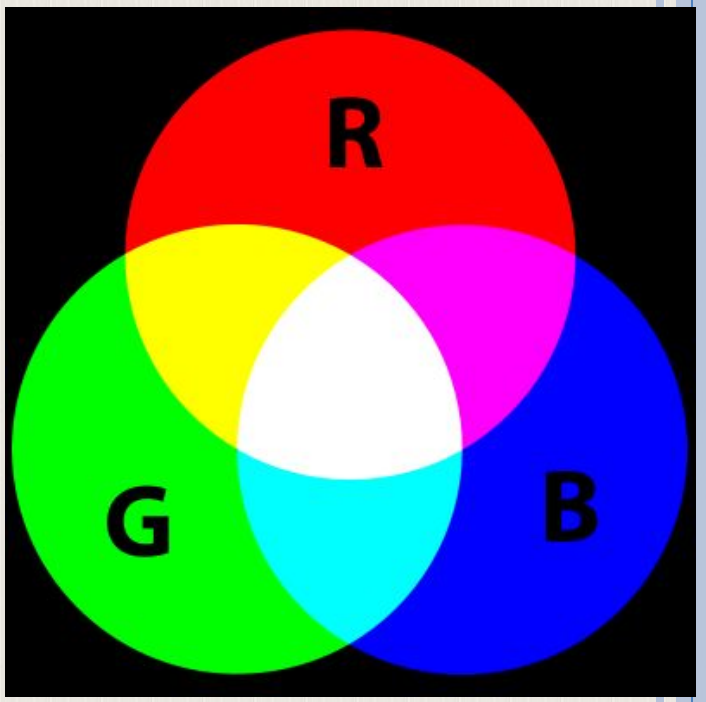


Шаг 1. Дискретизация: разбивка на *пиксели*.

Шаг 2. Для каждого пикселя определяется **цвет**.

Пиксель – это наименьший элемент рисунка, для которого можно независимо установить цвет.

Дискретное представление цвета

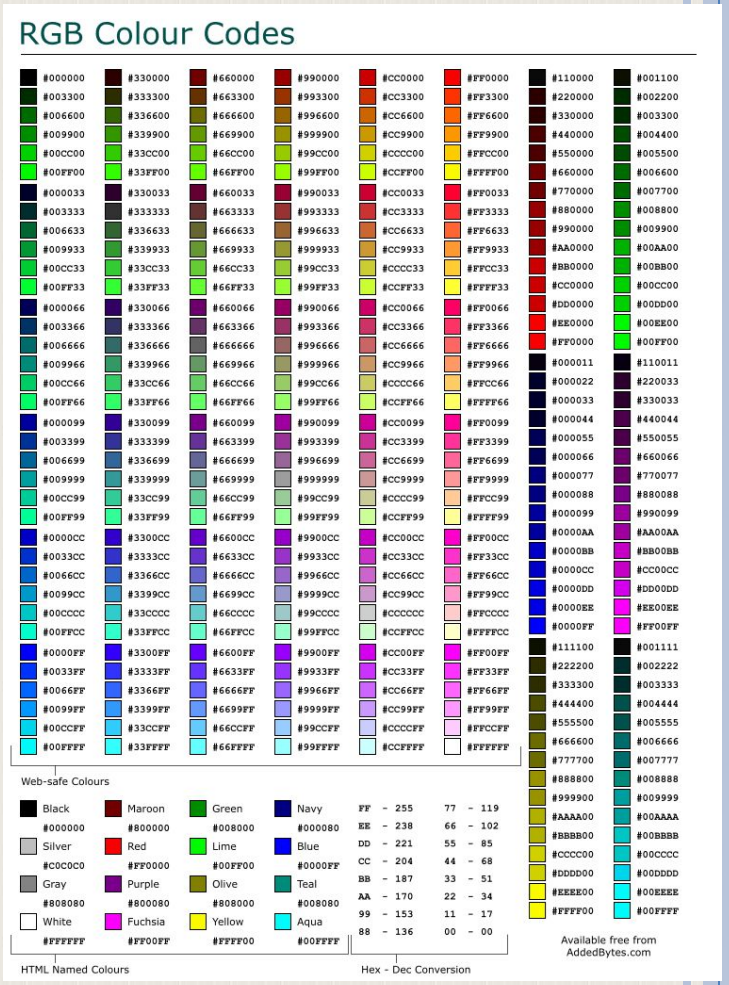


Любой цвет точки на экране компьютера получается путем смешивания трех базовых цветов: красного, зеленого, синего.

Этот принцип называется цветовой моделью RGB (Red, Green, Blue).

Дискретное представление цвета

Двоичный код цвета определяет, в каком соотношении находятся интенсивности трех базовых цветов. Если все они смешиваются в одинаковых долях, то в итоге получается белый цвет. Если все три компоненты «выключены», то цвет пикселя — черный. Все остальные цвета лежат между белым и черным.

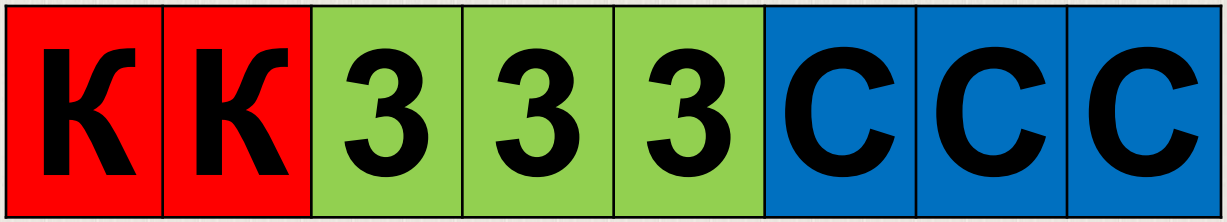


Дискретность цвета

Состоит в том, что интенсивности базовых цветов могут принимать конечное число дискретных значений.

Пусть, например, размер кода цвета пикселя равен 8 битам — 1 байту. Между базовыми цветами они могут быть распределены так:

2 бита под красный цвет, 3 бита — под зеленый и 3 бита — под синий.



Интенсивность красного цвета может принимать $2^2 = 4$ значения, интенсивности зеленого и синего цветов — по $2^3 = 8$ значений. Полное число цветов, которые кодируются 8-разрядными кодами, равно: $4x8x8 = 256 = 2^8$.

Из описанного правила, в частности, следует:

красный		зеленый			синий		
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
0	1	0	0	1	0	0	1
0	0	1	1	1	0	0	0
0	0	0	0	1	0	0	0

- код черного цвета
- код белого цвета
- код бледно-серого цвета
- код ярко-зеленого цвета
- код бледно-зеленого цвета

Цветовая модель CMYK

Цветовая модель для изображения на бумаге, формируемого с помощью отражения света.

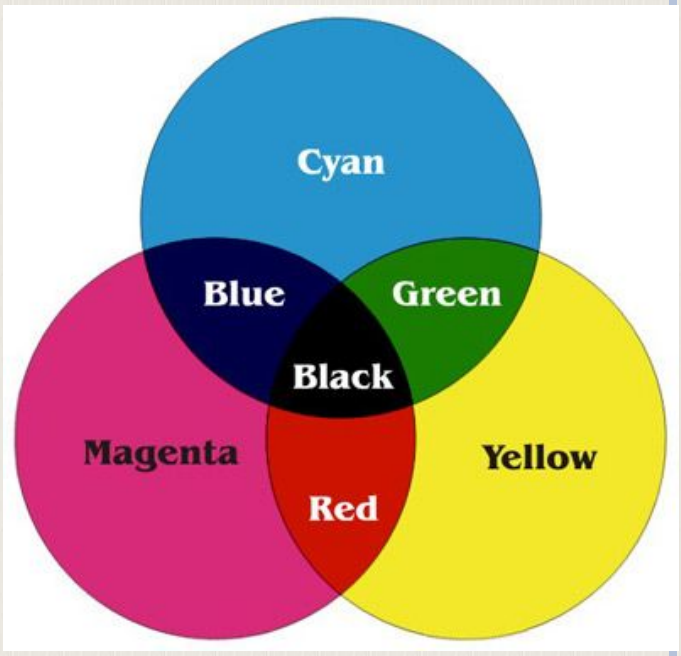
CMYK:

Cyan — голубой,

Magenta — пурпурный,

Yellow — желтый,

blacK — черный.



Цветовая модель

Цвет, который мы видим на листе бумаги, — это отражение белого (солнечного) света. Нанесенная на бумагу краска поглощает часть палитры, составляющей белый цвет, а другую часть отражает.

Таким образом, нужный цвет на бумаге получают путем «вычитания» из белого цвета «ненужных красок». Поэтому в цветной полиграфии действует не правило сложения цветов (как на экране компьютера), а правило вычитания.

Растровая и векторная графика

- В **растровой графике** графическая информация — это *совокупность данных о цвете каждого пикселя на экране*.

- В **векторной графике** графическая информация — это *данные, математически описывающие графические примитивы, составляющие рисунок*: прямые, дуги, прямоугольники, овалы и пр.

Растровая графика

Достоинство растровой графики — эффективное представление изображений фотографического качества.

Недостатки:

- большой объем занимаемой памяти,
- искажение изображения при его масштабировании.

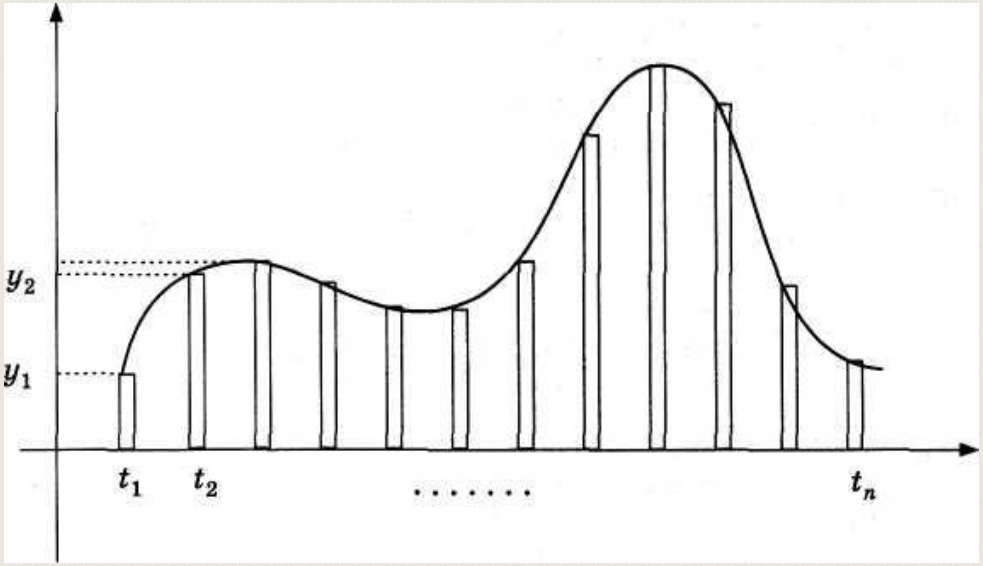
Векторная графика

Достоинства векторной графики —

- сравнительно небольшой объем памяти, занимаемой векторными файлами,
- масштабирование изображения без потери качества.

Недостаток: средствами векторной графики проблематично получить высококачественное художественное изображение.

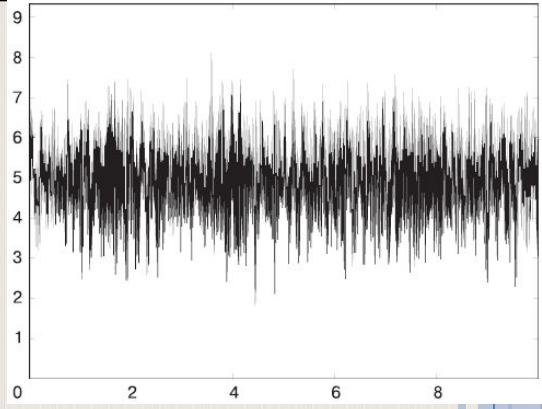
Звуковая информация



В процессе кодирования звукового сигнала непрерывная волна разбивается на отдельные маленькие временные участки и для каждого такого участка устанавливается определенная величина амплитуды.

Таким образом, непрерывная зависимость амплитуды сигнала от времени заменяется на дискретную последовательность уровней громкости.

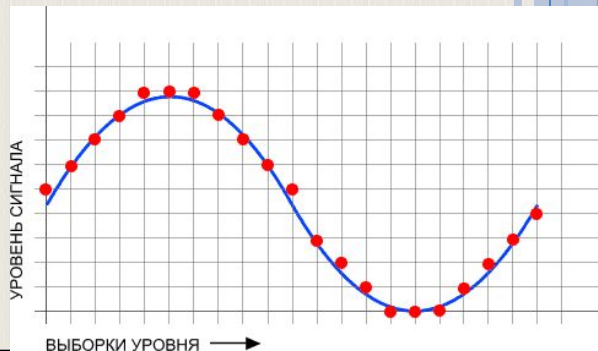
Оцифровка звука



- Ввод звука в компьютер производится с помощью звукового устройства (микрофона, радио и др.), выход которого подключается к порту звуковой карты.
- **Задача звуковой карты** — с определенной частотой производить измерения уровня звукового сигнала (преобразованного в электрические колебания) и результаты измерения записывать в память компьютера.
- Этот процесс называют ***оцифровкой звука***.

Дискретизация

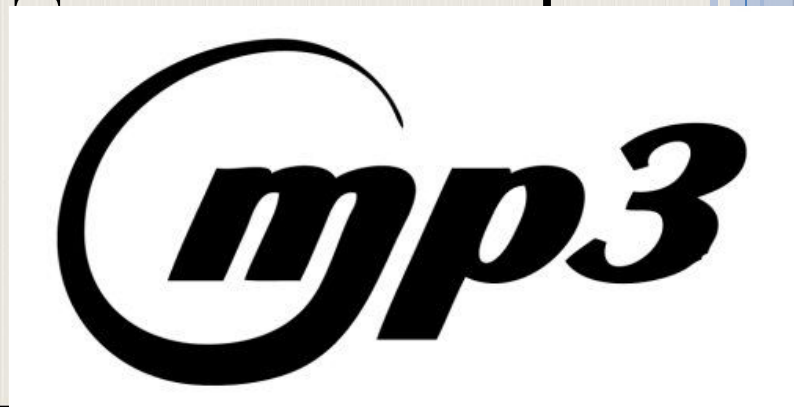
- **Дискретизация** - это преобразование непрерывных изображений и звука в набор дискретных значений в форме кодов.
- Промежуток времени между двумя измерениями называется **периодом измерений**.
- Обратная величина называется **частотой дискретизации** - количество измерений уровня сигнала в единицу времени — 1/т (герц).
- ***Чем выше частота измерений, тем выше качество цифрового звука.***



Форматы файлов

Для сохранения звука без потерь используется универсальный звуковой формат файлов **WAV**.

Наиболее известный *формат «сжатого» звука (с потерями)* — **MP3**. Он обеспечивает сжатие данных в 10 раз и более.



Представление видеоинформации

- ***Что представляет собой фильм с точки зрения информатики? Прежде всего, это сочетание звуковой и графической информации.***
- ***Исследования показали, что если за одну секунду сменяется более 10-12 кадров, то человеческий глаз воспринимает изменения на них как непрерывные.***

ФОРМАТЫ ПРЕДСТАВЛЕНИЯ ВИДЕОДАННЫХ

- В среде Windows применяется формат Video for Windows, базирующийся на универсальных файлах с расширением **AVI** (Audio Video Interleave – чередование аудио и видео).
- Более универсальным является мультимедийный формат **Quick Time**
- Наиболее известным стандартом систем сжатия видеоизображений служит **MPEG**(Motion Picture Expert Group).