

Лекция 3

Понятие информации Измерение информации

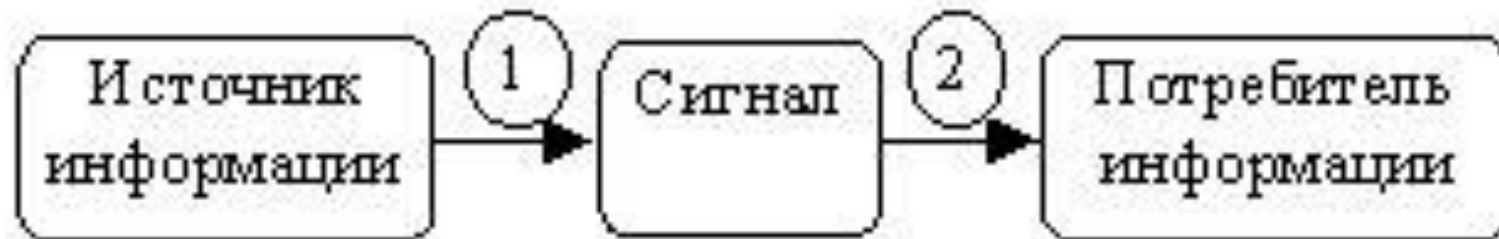
Информация – сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые уменьшают имеющуюся о них степень неопределённости, неполноты знаний.

Данные - признаки или записанные наблюдения, которые по каким – то причинам не используются, а только хранятся.

Информация - используемые данные

Информационные коммуникации

- пути и процессы, обеспечивающие передачу сообщений от источника информации к её потребителю



Адекватность информации

- **это определённый уровень соответствия создаваемого с помощью полученной информации образа реальному объекту, процессу, явлению, и т.д.**

Адекватность информации может выражаться в трех формах:

- ✓ семантической,
- ✓ синтаксической,
- ✓ прагматической.

Синтаксическая адекватность

- **отображает формально-структурные характеристики информации и не затрагивает ее смыслового содержания.**

Информацию, рассматриваемую только с синтаксических позиций, обычно называют **данными, так как при этом не имеет значения смысловая сторона**

На синтаксическом уровне учитываются:

- **тип носителя**
- **способ представления информации**
- **скорость передачи и обработки**
- **размеры кодов представления информации**
- **надёжность и точность преобразования ЭТИХ КОДОВ**
- **и т.п.**

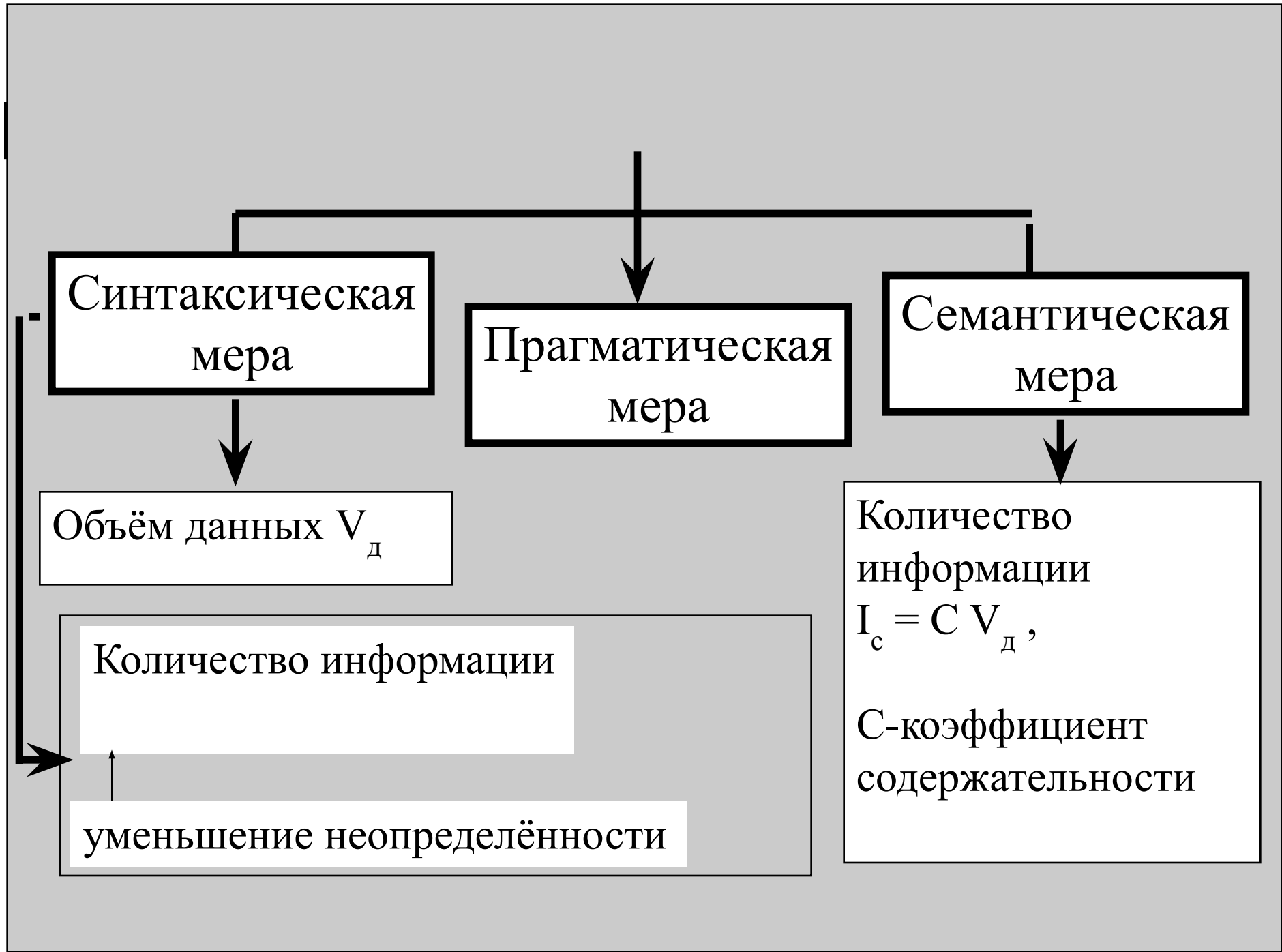
Семантическая (смысловая) адекватность

- определяет степень соответствия образа объекта и самого объекта
- предполагает учет смыслового содержания информации

Прагматическая адекватность

- ✓ отражает отношение информации и ее потребителя
- ✓ связана с практическим использованием информации

Me



Синтаксическая мера

Объём данных (V_D) - информационный объём сообщения или объём памяти, необходимый для хранения сообщения без каких-либо изменений (измеряется в **битах**).

бит - количество информации, которое содержит сообщение, уменьшающее неопределенность в два раза.

формула Шеннона

$$H(\alpha) = - \sum_{i=1}^N p_i \log_2 p_i$$

где p_i — вероятность того, что система находится в i -м состоянии из N возможных состояний

При кодировании информации
числовыми кодами в той или иной
системе счисления

$$N = m^n$$

N — число всевозможных отображаемых состояний;

m — основание системы счисления;

n — число разрядов (символов) в сообщении.

Формула Хартли

- Р. Хартли (1928 г.) процесс получения информации рассматривал как выбор одного сообщения из конечного наперёд заданного множества из N равновероятных сообщений,
- количество информации I , содержащееся в выбранном сообщении:

$$I = \log_2 N$$

Каждая буква русского алфавита (если считать, что е=ё) несет информацию

5 бит $(32 = 2^I)$.

Семантическая мера информации

Тезаурусная мера связывает семантические свойства информации (смысловое содержание) со способностью пользователя принимать поступившее сообщение.

***тезаурус пользователя** - это совокупность сведений, которыми располагает пользователь или система.*

Прагматическая мера информации

Определяет полезность информации (ценность) для достижения пользователем поставленной цели.

Величина относительная, обусловленная особенностями использования этой информации в той или иной системе.

Кодирование данных двоичным кодом

- ❑ бит (bit) - от английского словосочетания **B**inary **di**gi**T** (двоичная цифра).
- ❑ Один бит кодирует два понятия:
0 или **1** (да или нет, черное или белое, истина или ложь и т.п.)
- ❑ два бита - четыре различных значения: **00 01**
10 11
- ❑ Тремя битами можно закодировать восемь различных значений: **000 001 010 011 100 101**
110 111

общая формула имеет вид: $N = 2^m$

где N - количество
независимых
кодируемых значений;
m - разрядность двоичного
кодирования

m	N
1	2
2	4
3	8
4	16
...	...
8	256

Единицы измерения информации:

1 байт = $2^3 = 8$ бит

1 Килобайт = 1024 байт.

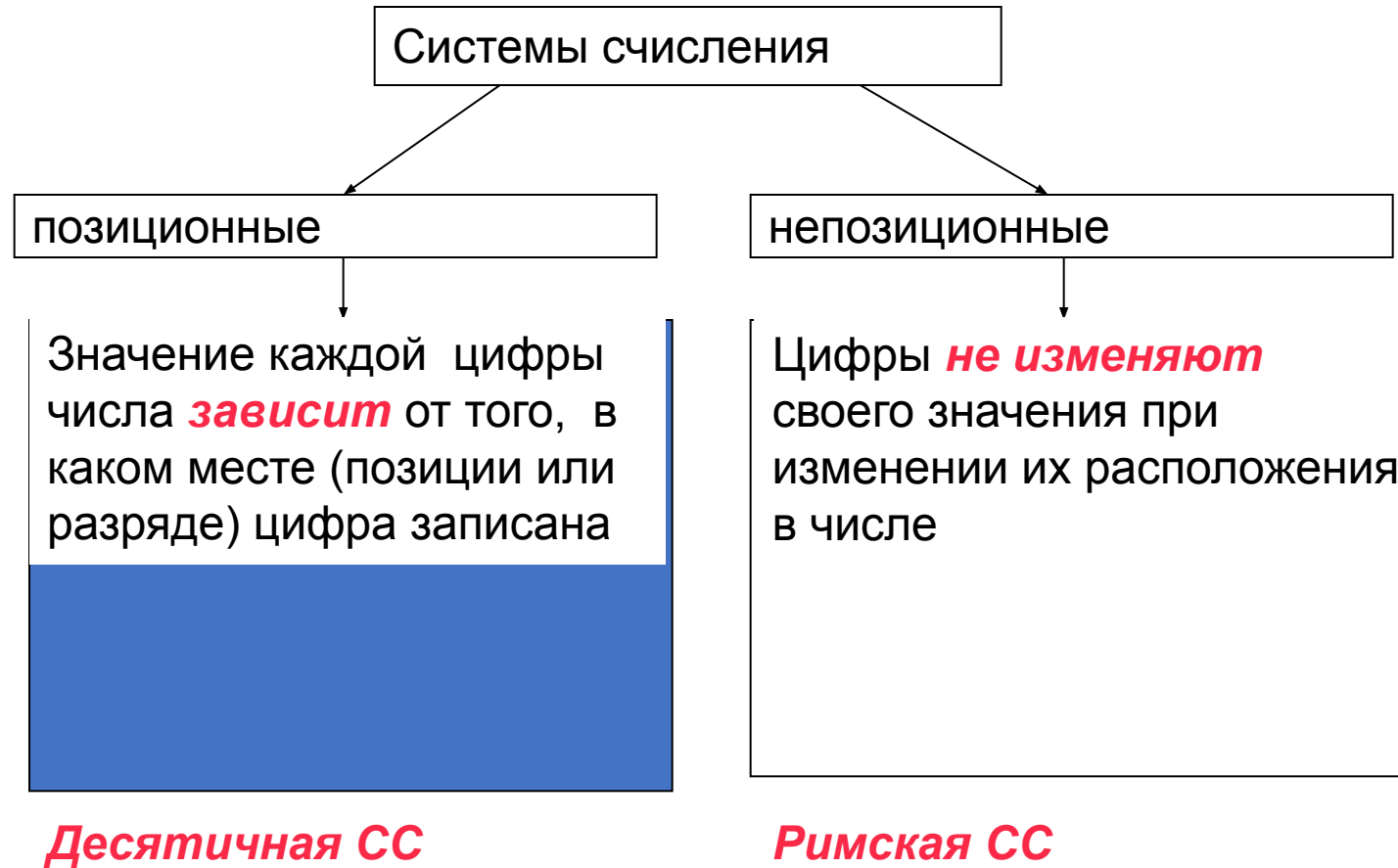
Килобайт	Кбайт = 2^{10} байт
Мегабайт	Мбайт = 2^{10} Кбайт
Гигабайт	Гбайт = 2^{10} Мбайт
Терабайт	Тбайт = 2^{10} Гбайт
Петабайт	Пбайт = 2^{10} Тбайт
Эксобайт	Эбайт = 2^{10} Пбайт

Что такое **система счисления**?

Система счисления – это совокупность правил записи чисел с помощью определенного набора символов.

Для записи чисел могут использоваться не только цифры, но и буквы.

Что такое **система счисления**?



Не позиционные системы счисления

Римская система счисления

- Является непозиционной, т.к. каждый символ обозначает всегда одно и тоже число;
- Цифры обозначаются латинскими буквами:

I, V, X, L, C, D, M

(1, 5, 10, 50, 100, 500, 1000)

Например: XXX – 30; LXI - 61

Позиционные системы счисления

Основание ПСС – это количество цифр, используемое для представления чисел;

Алфавит – набор символов, используемый для обозначения цифр.

- Значение цифры зависит от ее позиции, т.е. одна и та же цифра соответствует разным значениям в зависимости от того, в какой позиции числа она стоит;

Например: 888: 800; 80; 8

- Любое позиционное число можно представить в виде суммы степеней основания системы.

Позиционные системы счисления

Десятичная СС

- Основание системы – число 10;
- Алфавит (10 цифр): 0, 1, 2, 3, 4, 5, 6, 7, 8, 9;
- Любое десятичное число можно представить в виде суммы степеней числа 10 – основания системы;

$$2345_{10} = 2 \cdot 10^3 + 3 \cdot 10^2 + 4 \cdot 10^1 + 5 \cdot 10^0$$

Позиционные системы счисления

Двоичная СС

- Основание системы – 2;
- Алфавит (2 цифры): 0; 1;
- Любое двоичное число можно представить в виде суммы степеней числа 2 – основания системы;

$$10101_2 = 1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$$

Кодирование Графической информации

формы представления графической информации

Аналоговая



живописное полотно, цвет
которого изменяется непрерывно

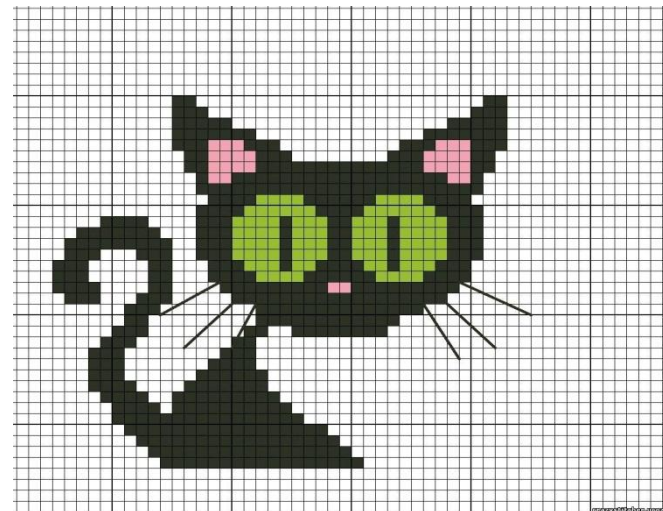
Дискретная



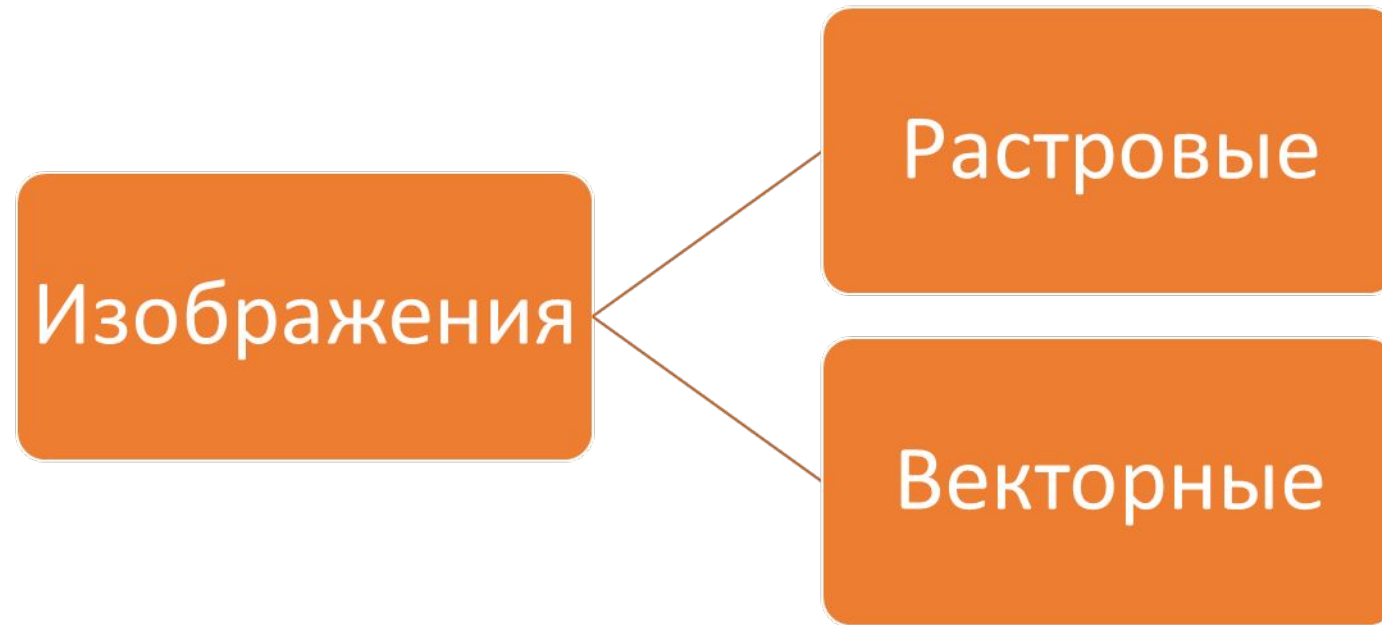
изображение, напечатанное с помощью
струйного принтера и состоящее из
отдельных точек разного цвета

Графические изображения из аналоговой (непрерывной) формы в цифровую (дискретную) преобразуются путем пространственной дискретизации.

Пиксель - минимальный участок изображения, для которого независимым образом можно задать цвет.



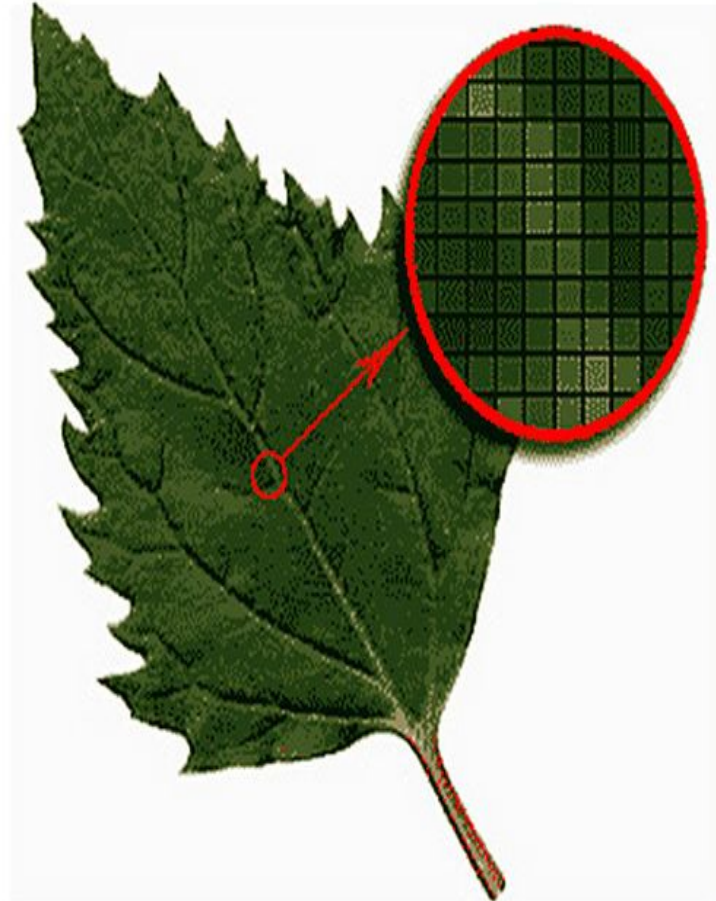
Все компьютерные изображения разделяют на два основных типа: растровые и векторные



Векторное изображение представляет собой совокупность графических примитивов (точка, отрезок, эллипс и т.д.). Каждый примитив описывается математическими формулами.



растровое изображение – это набор пикселей, расположенных на прямоугольной сетке.



Качество изображения зависит от **разрешающей способности**. Разрешающая способность растрового изображения определяется количеством точек по горизонтали и количеством точек по вертикали на единицу длины изображения.

- Чем меньше размер точки, тем больше разрешающая способность (больше строк растра и точек в строке) и, соответственно, выше качество изображения. Величина разрешающей способности выражается в dpi (dot per inch - точек на дюйм), т.е. в количестве точек в полоске изображения длиной в 1 дюйм (1 дюйм = 2,54 см).

Количество информации, которое используется для кодирования цвета точки изображения, называется *глубиной* цвета.

Каждый цвет можно рассматривать как возможное состояние точки. Количество цветов N в палитре и количество информации для кодирования цвета каждой точки связаны между собой формулой: $N=2^I$.

Глубина цвета I	Количество отображаемых цветов N
4	$2^4 = 16$
8	$2^8 = 256$
16 (hige color)	$2^{16} = 65\ 536$
24 (true color)	$2^{24} = 16\ 777\ 216$
32 (true color)	$2^{32} = 4\ 294\ 967\ 296$

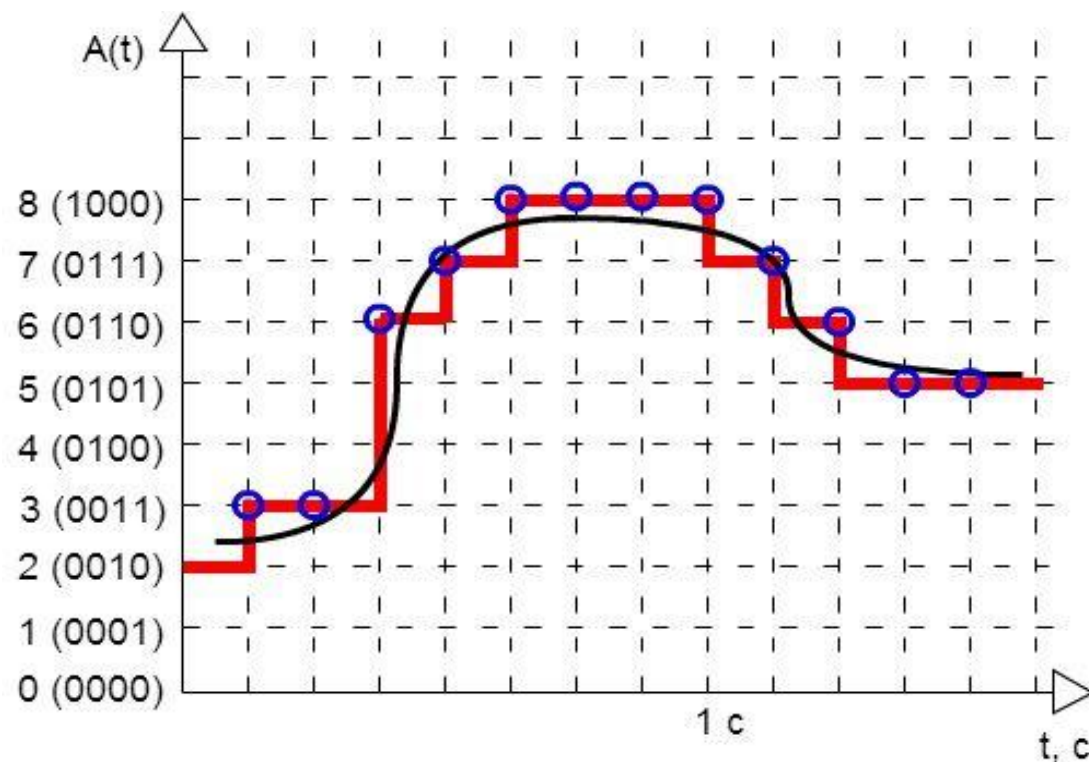
Звук - это физическое явление, представляющее собой распространение механических колебаний в упругой среде в виде упругих волн .



ЦИФРОВОЕ ПРЕДСТАВЛЕНИЕ ЗВУКА

Для того чтобы компьютер мог обрабатывать звук, непрерывный звуковой сигнал должен быть превращен в последовательность электрических импульсов (двоичных нулей и единиц).

Каждой «ступеньке» присваивается значение уровня громкости звука, его код (1, 2, 3 и так далее).



Информационный объем оцифрованного звука

$$2^b = N$$

$$I = H \times t \times b$$

N – количество уровней громкости

b – глубина кодирования звука, разрядность квантования

H – частота дискретизации

I – информационный объем звукового файла

Кодирование текстовых данных

СИМВОЛЫ	Число кодов
Цифры 0-9	10
Знаки препинания, математические символы, специальные символы (\$, # и т.д.)	34
Прописные и строчные латинские буквы	26+26=52
Управляющие символы	32
Итого:	128

$(2^7 = 128)$

каждый текстовый символ

(из 128 + нац. алфавит и псевдографические символы) кодируется 8-ми битовым двоичным числом.

$2^8 = 256$ СИМВОЛОВ.