

Биологический институт

Томский государственный университет

Информационные технологии

Лекция 1

Кодирование и сжатие графических,
аудио- и видеоданных

Дмитрий Владимирович

Курбатский

старший преподаватель каф. ихтиологии и гидробиологии,
научный сотрудник ЛМБ БИ ТГУ, магистр биологии

- Зоологический музей (к. 123)
- Компьютерный класс (к. 028)
- Группа ВКонтакте «Курсы "Информатика" и "Информационные технологии"»:
vk.com/i_it_bi_tsu
- Персональный раздел:
zoo.tsu.ru/kdv
- [Рейтинг на сайте Professorrating.ru](http://Professorrating.ru)

Блок 1

Кодирование звуковых данных

ЦАП и АЦП

- Аналого-цифровой преобразователь (АЦП, Analog-to-digital converter, ADC) — устройство, преобразующее входной аналоговый сигнал в дискретный код (цифровой сигнал).
- Цифро-аналоговый преобразователь (ЦАП, DAC) — устройство для преобразования цифрового (обычно двоичного) кода в аналоговый сигнал (ток, напряжение или заряд).

Схема АЦП

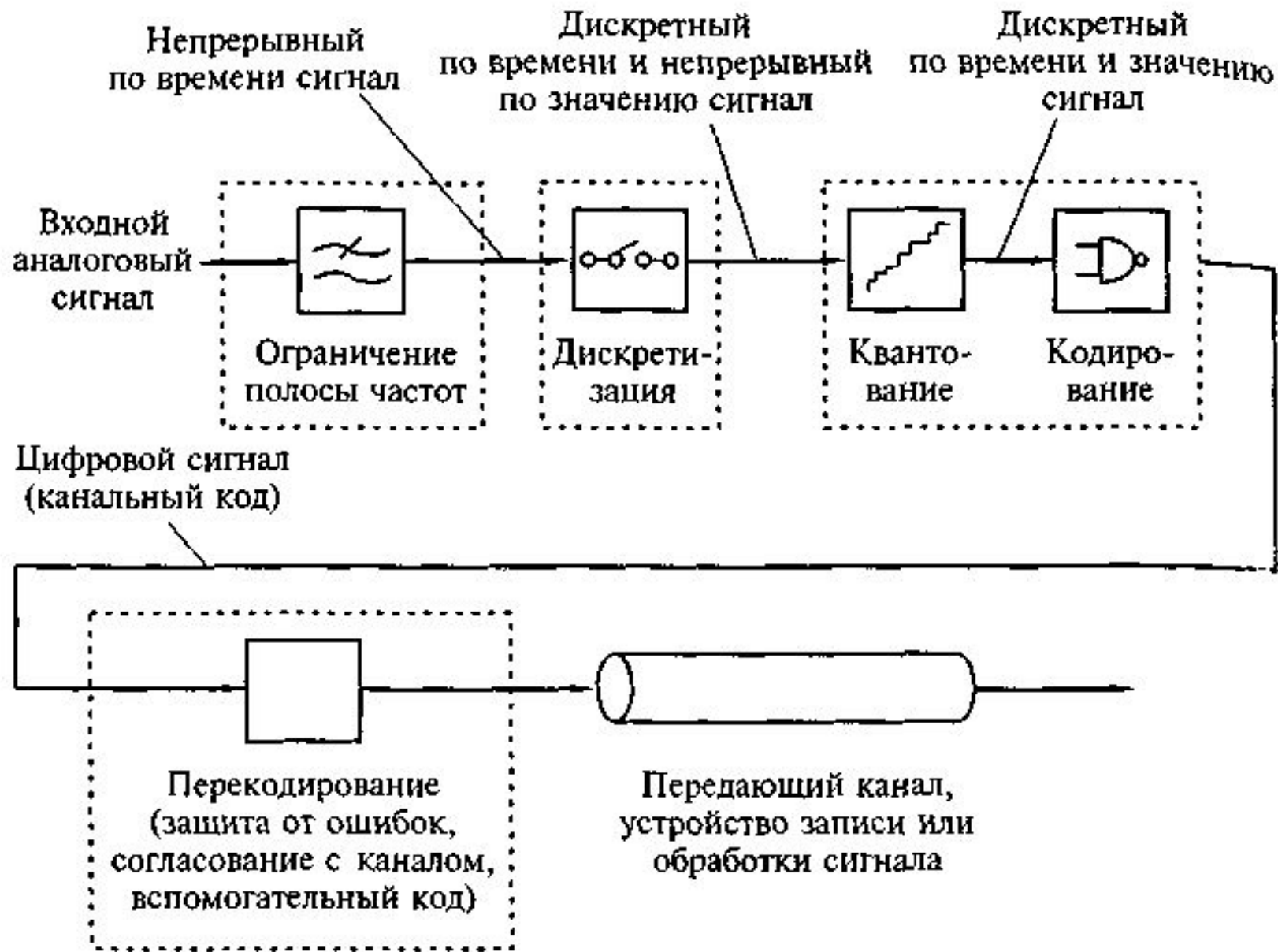
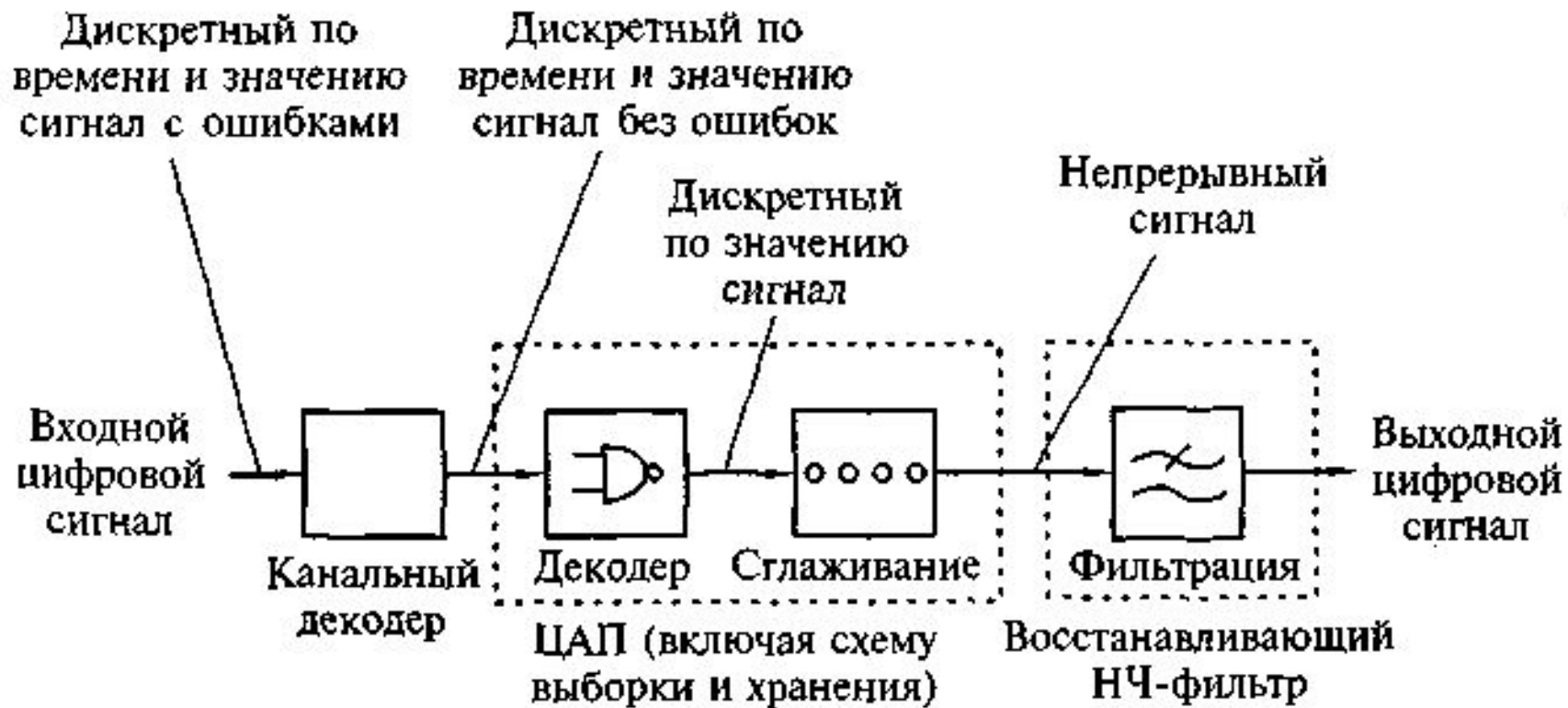
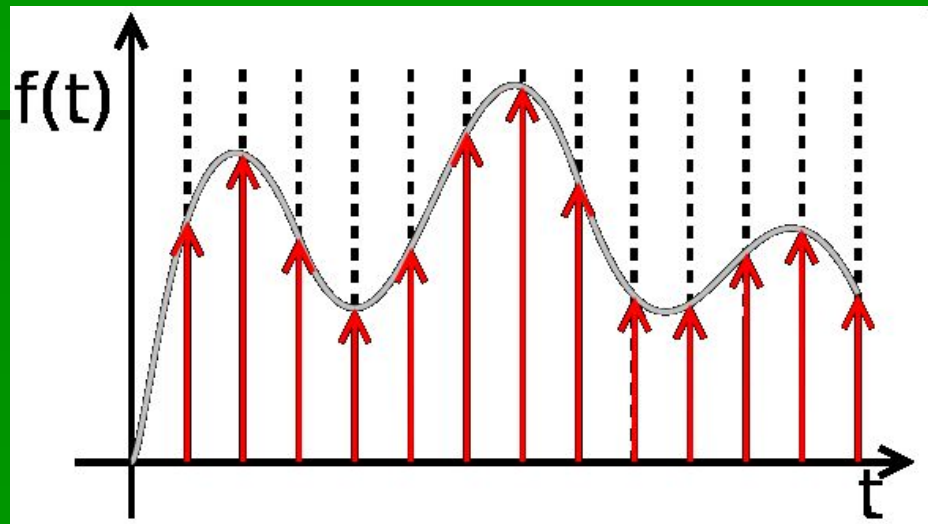


Схема ЦАП



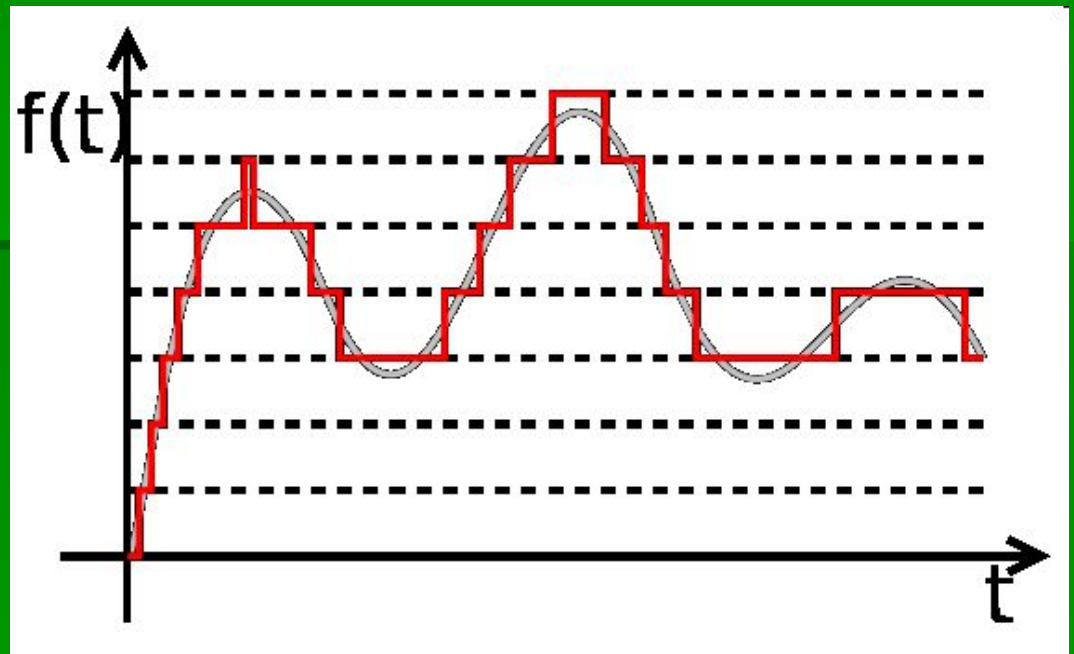
Дискретизация

- При дискретизации изменяющаяся во времени величина (сигнал) замеряется с заданной частотой.
- имеет частоту
 - частота дискретизации
 - частота выборки
 - частота сэмплирования



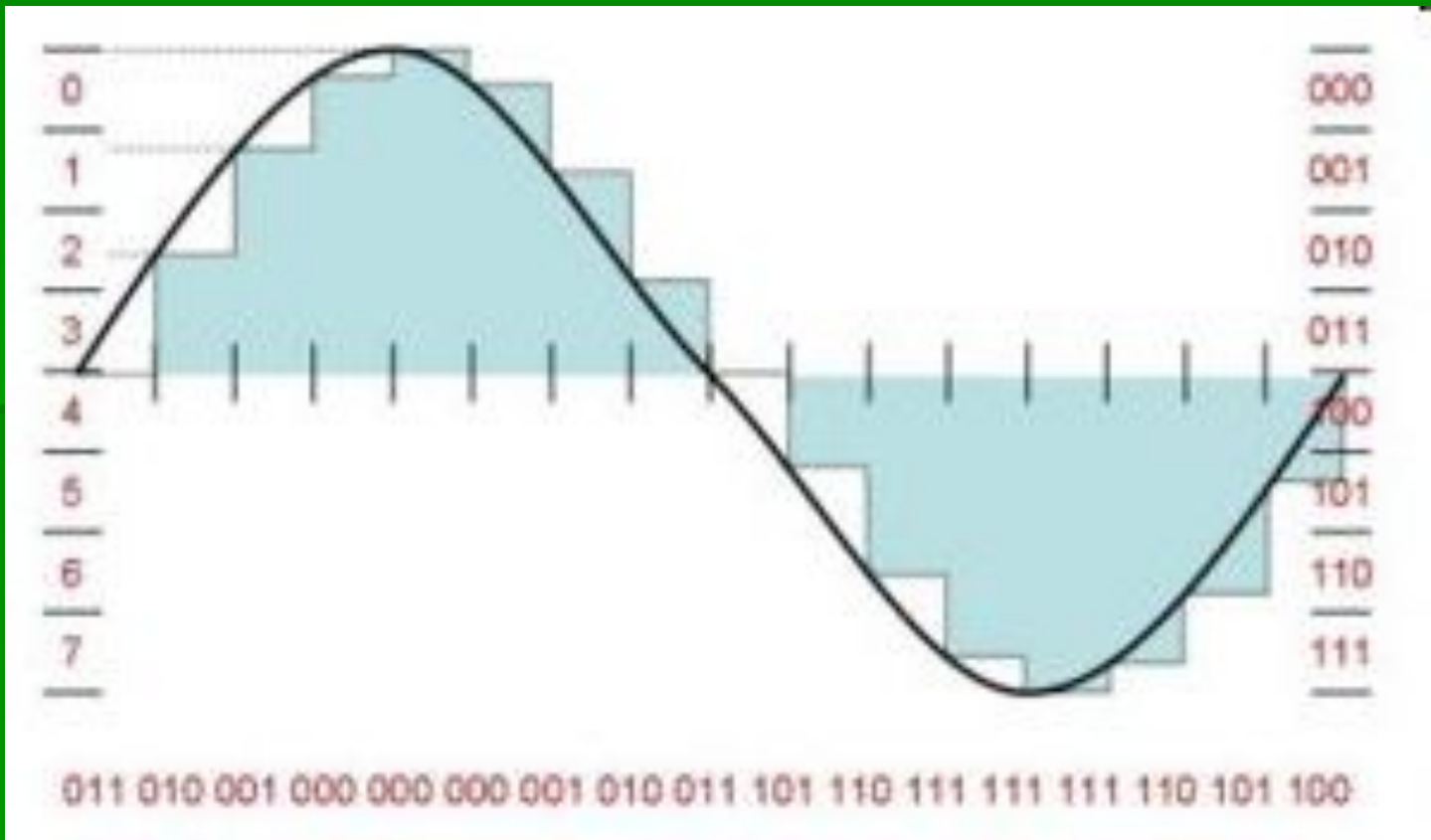
Квантование (quantization)

- — разбиение диапазона значений непрерывной или дискретной величины на конечное число интервалов.
- имеет шаг.
- ~ битности.



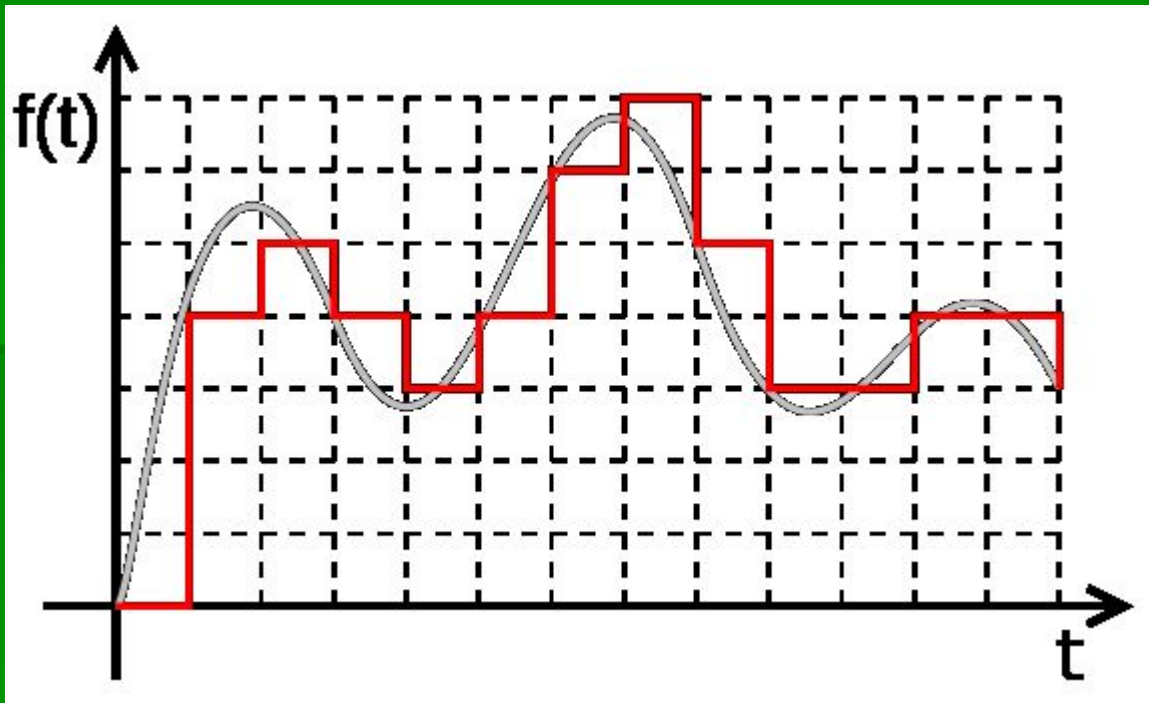
Разрядность квантования

- 2^N уровней
- N – разрядность
- При $N = 16+$ бит – погрешности становятся почти незаметными.

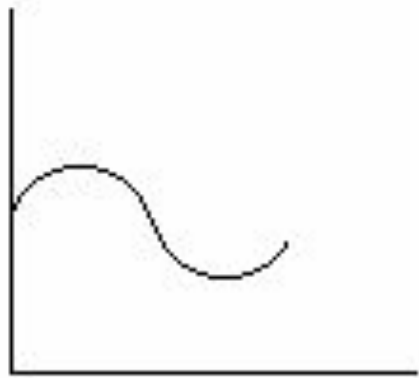


Цифровой сигнал

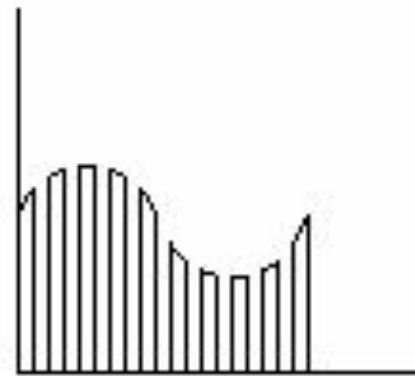
- = квантование + дискретизация



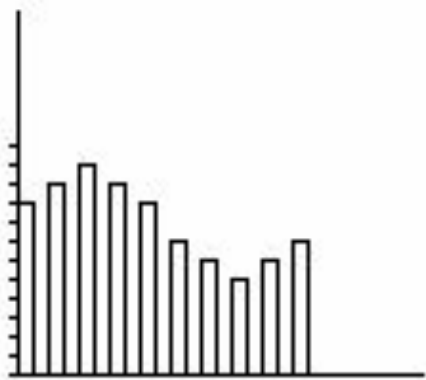
Цифровой сигнал



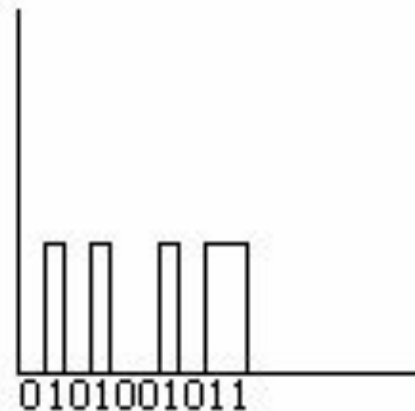
Дискретизация



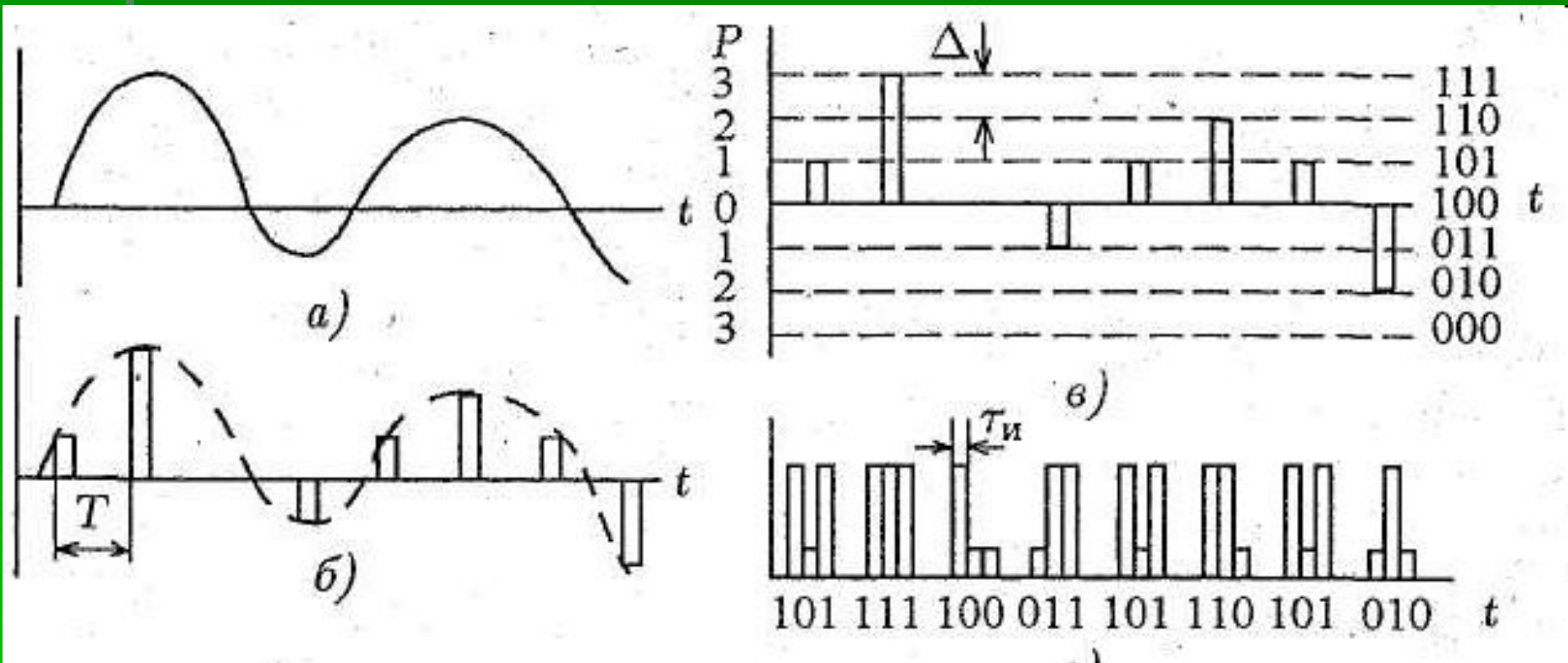
Квантование



Кодирование



Цифровой сигнал



Характеристики АЦП

- Частота дискретизации (Hz, Гц)
 - Поддерживаемый диапазон частот (Hz, Гц)
 - теорема Котельникова
 - $f_d > 2 * f_{\max}$
 - частота Найквиста
- Разрядность (бит)
 - Динамический диапазон (dВ, дБ)
 - Отношение сигнал/шум
 - 1 бит ~ 6 дБ

Импульсно-кодовая модуляция

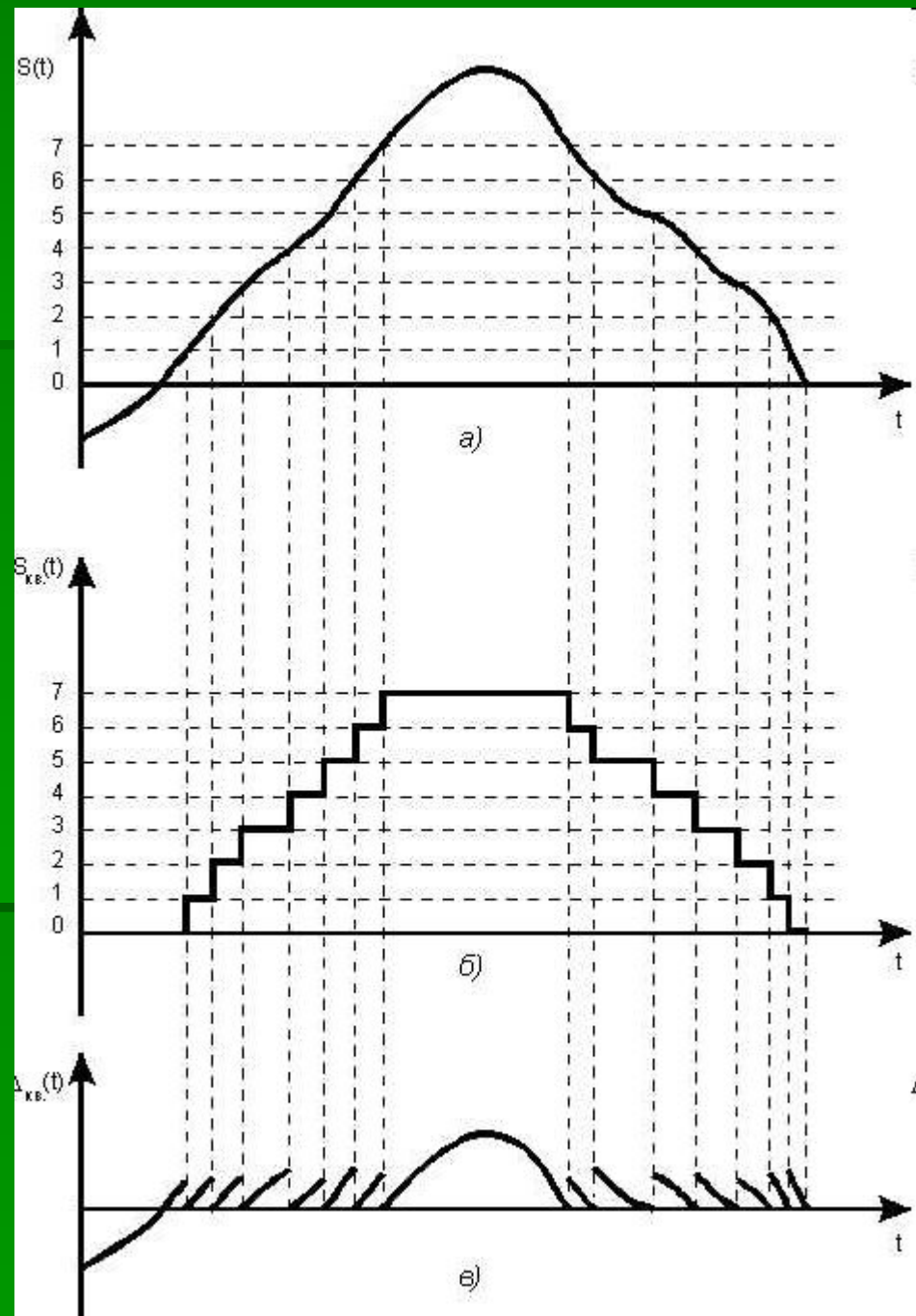
- – дискретизация сигнала во времени в совокупности с методом однородного квантования.
- *Pulse Code Modulation – PCM*
- Записываются **абсолютные значения** амплитуды.
- Пример:
 - 16 бит
 - 44.1 кГц
 - до 20 000 Гц (с запасом)

Цифровой сигнал

- Размеры данных при кодировании:
 - 1 с. стереозвука РСМ:
 - 2 (канала)
 - *
 - 2 (байта = 16 бит)
 - *
 - 44 100 Гц
 - =
 - 176 400 байт ~ 176 кБ
 - Это без сжатия.

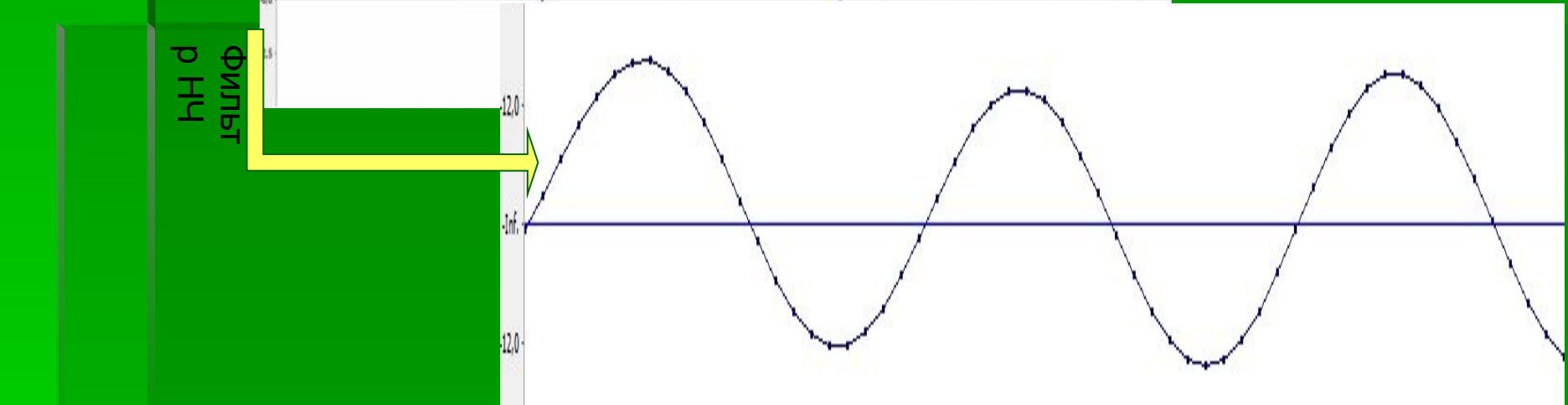
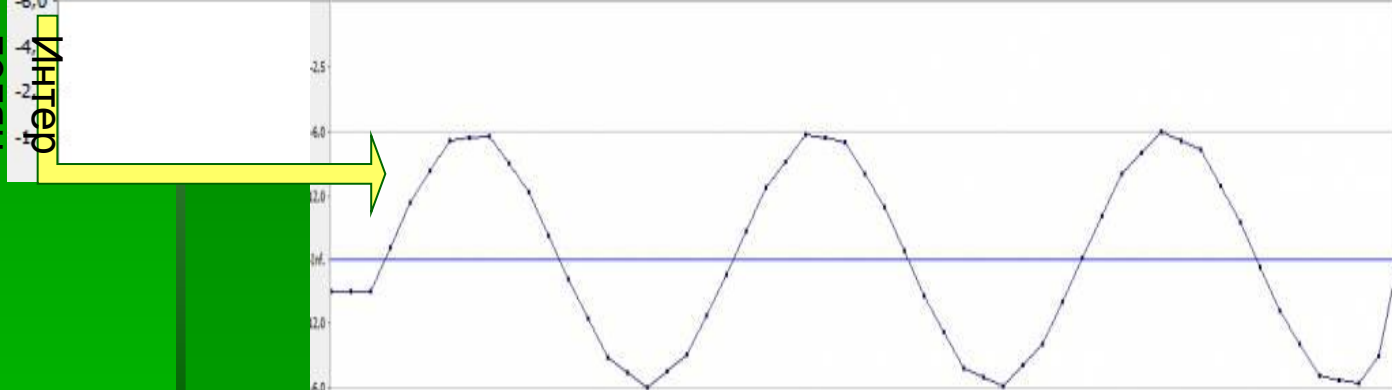
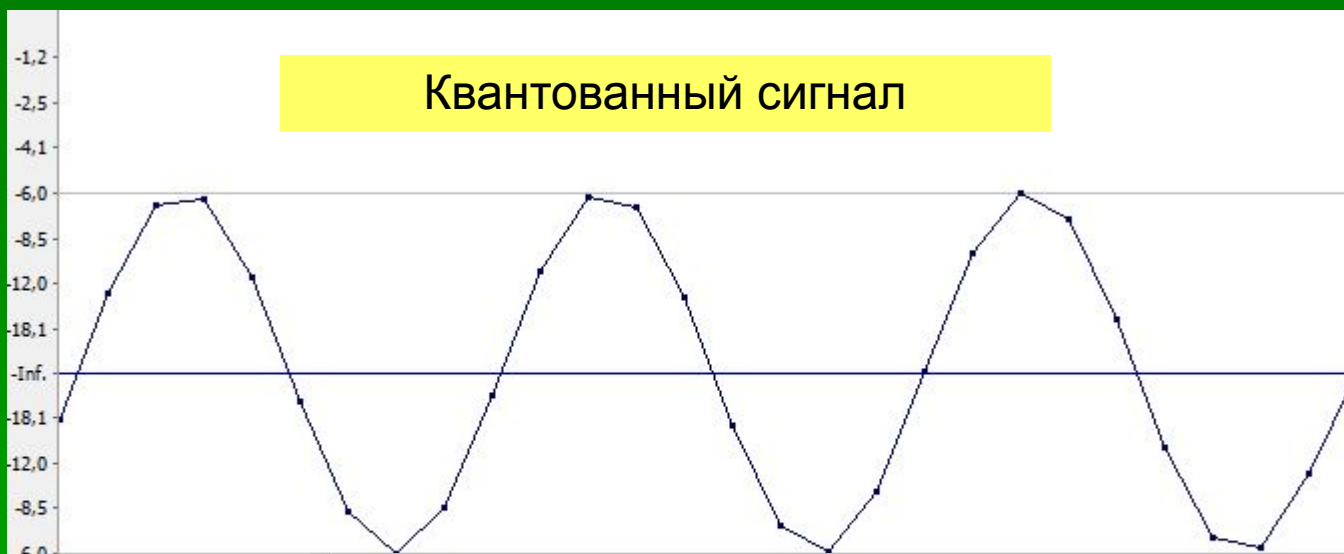
ЦАП

- Шум квантования



Квантованный сигнал

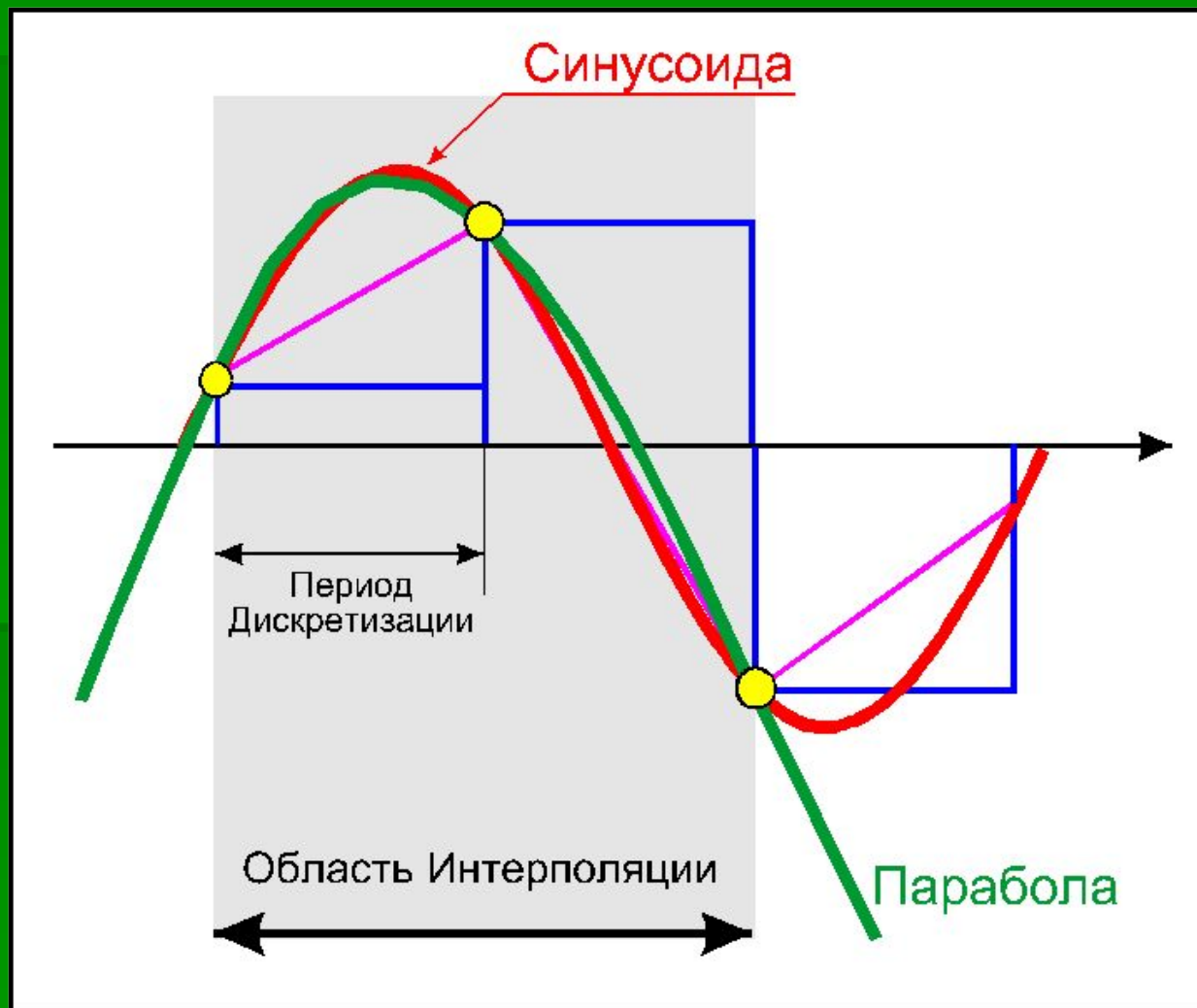
ЦАП



Интер

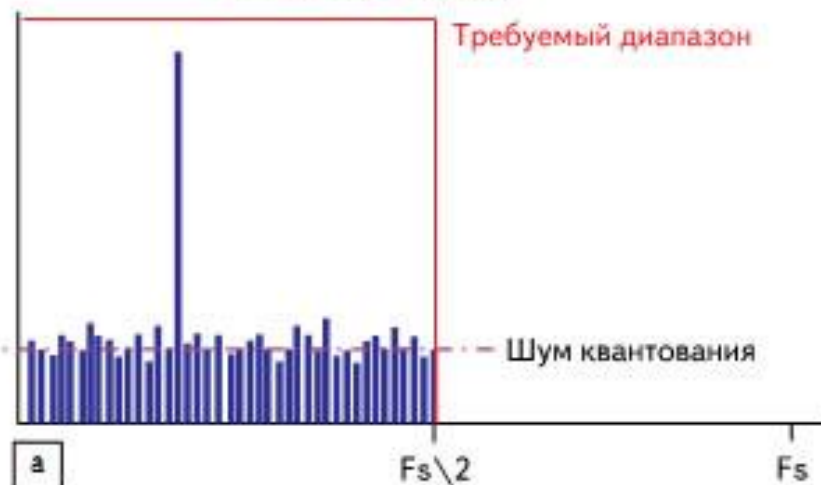
Филт
р НЧ

Интерполяция

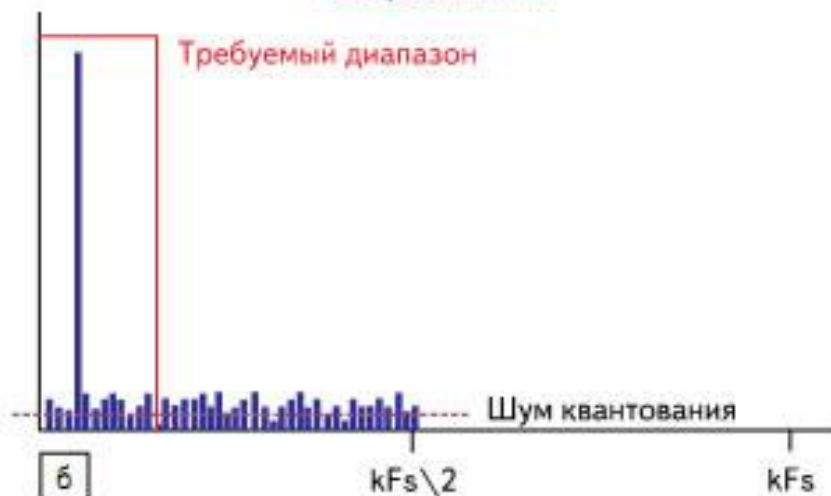


Шумоподавление

Обычные условия



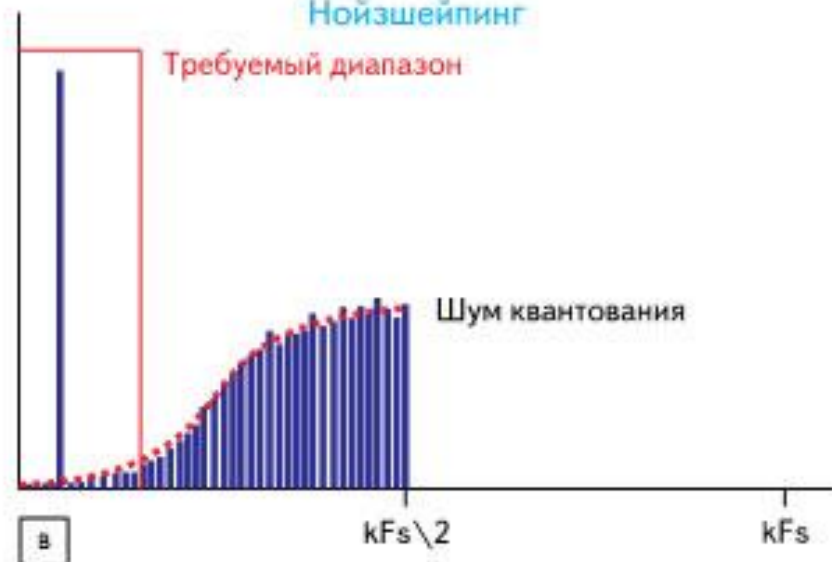
Оверсэмплинг



* $SNR=6,02N+1,6$ дБ
для N-битного ЦАП

* Интегратор выступает
в качестве ФВЧ для шума

Нойзшейпинг

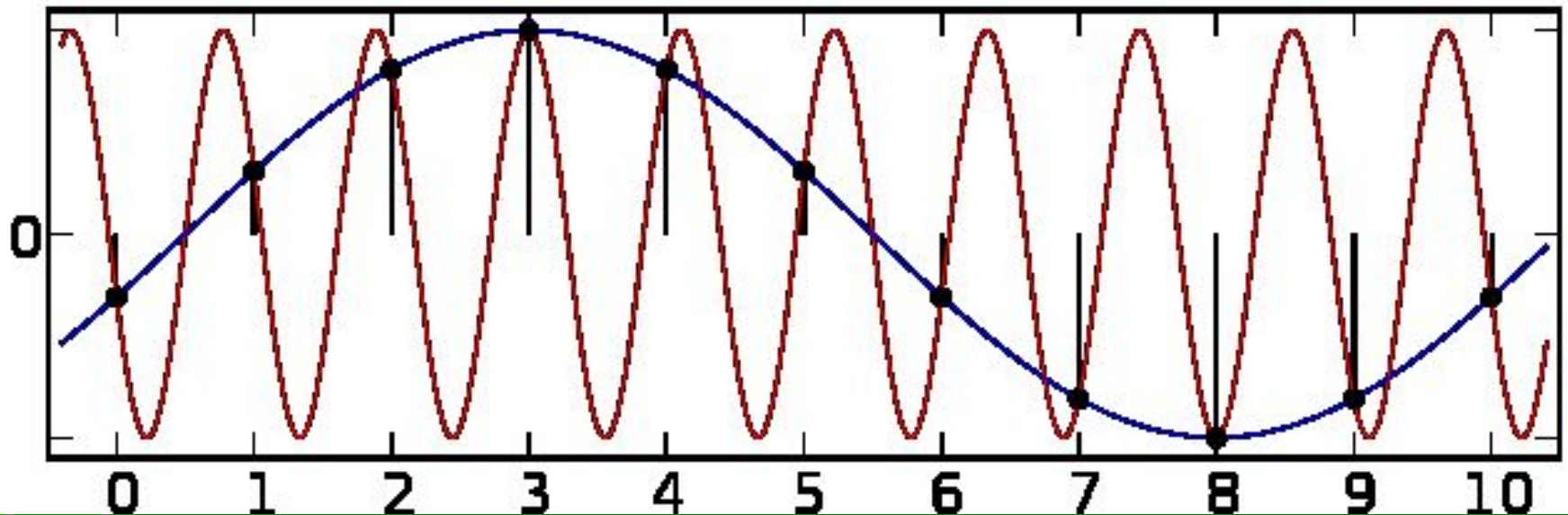


Связанные понятия

- подмешивание псевдослучайного сигнала *dither*
- полоса пропускания

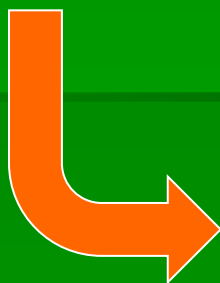
Проблемы

- алиасинг *aliasing*





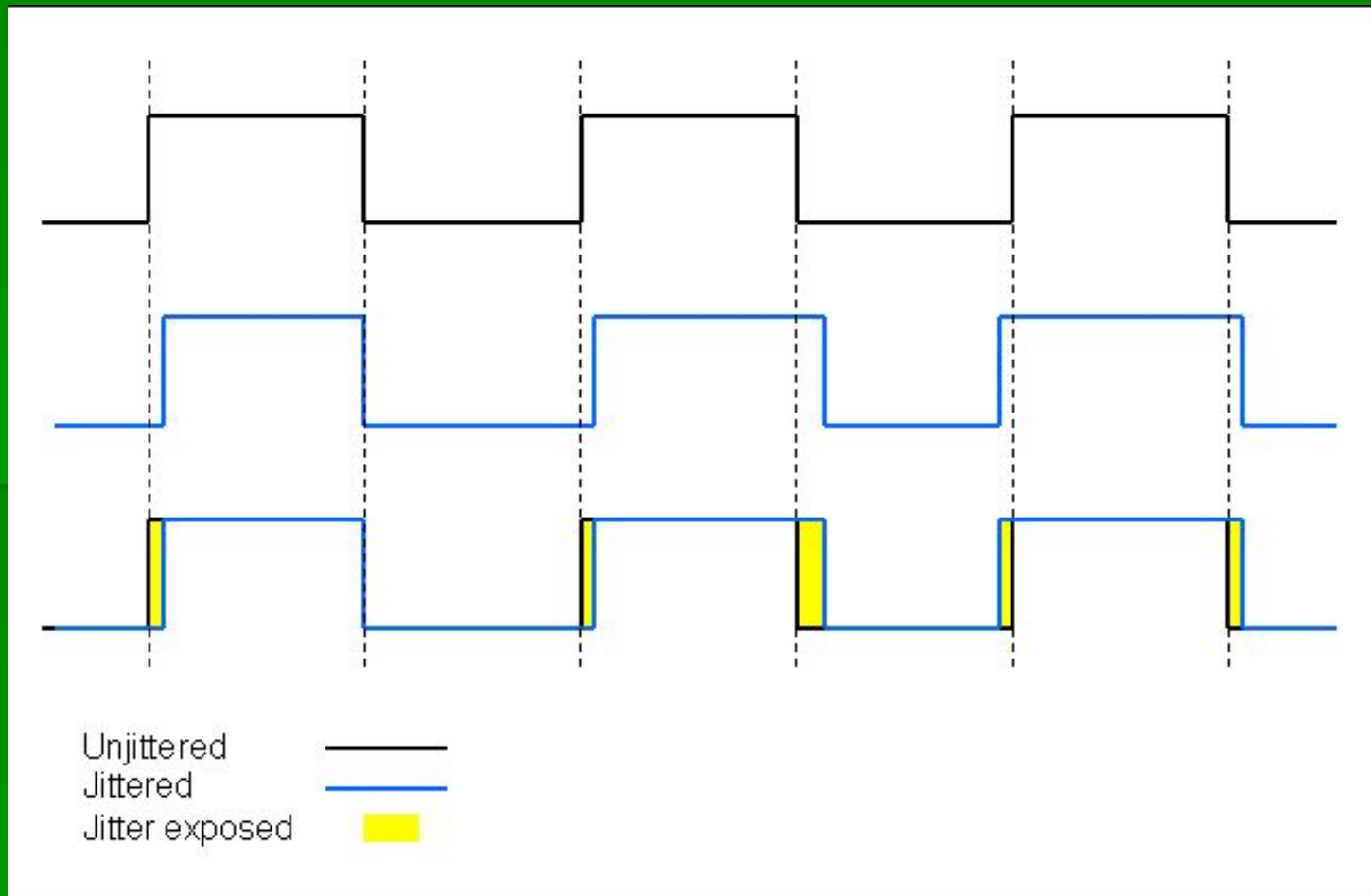
Свёртка спектра



Передискретизация

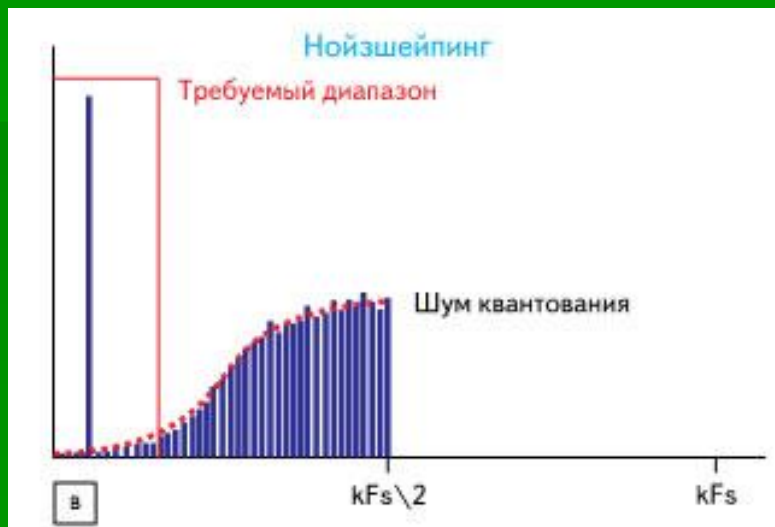
Джиттер *jitter*

- фазовое дрожание цифрового сигнала данных

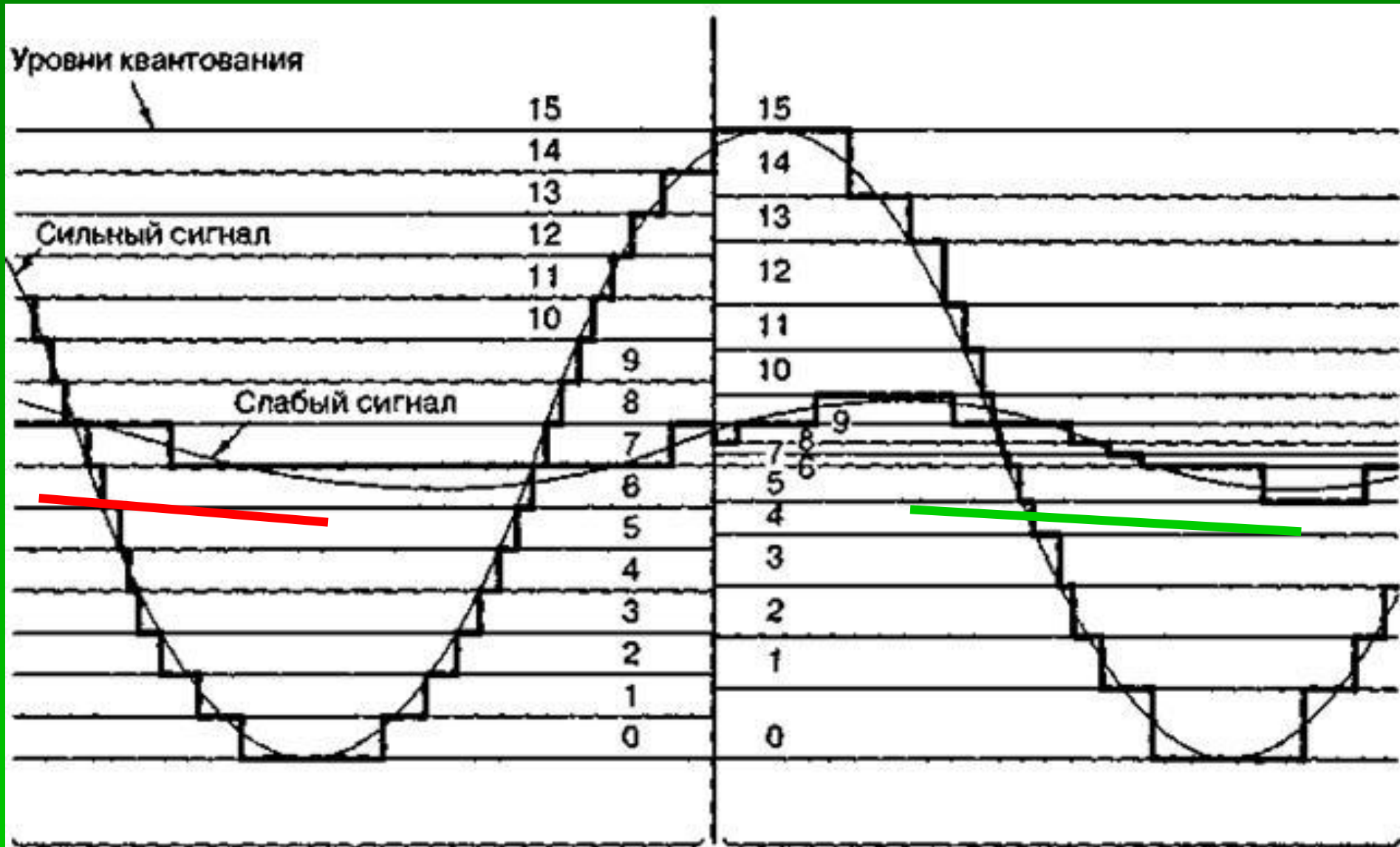


Другие варианты АЦП

- Неравномерное (логарифмическое) квантование
- Дельта-кодирование Дельта-кодирование (дельта-модуляция, дифференциальная импульсно-кодовая модуляция (*Delta PCM*))
 - Адаптивное дельта-кодирование (ADPCM)
- Сигма-дельта-модуляция



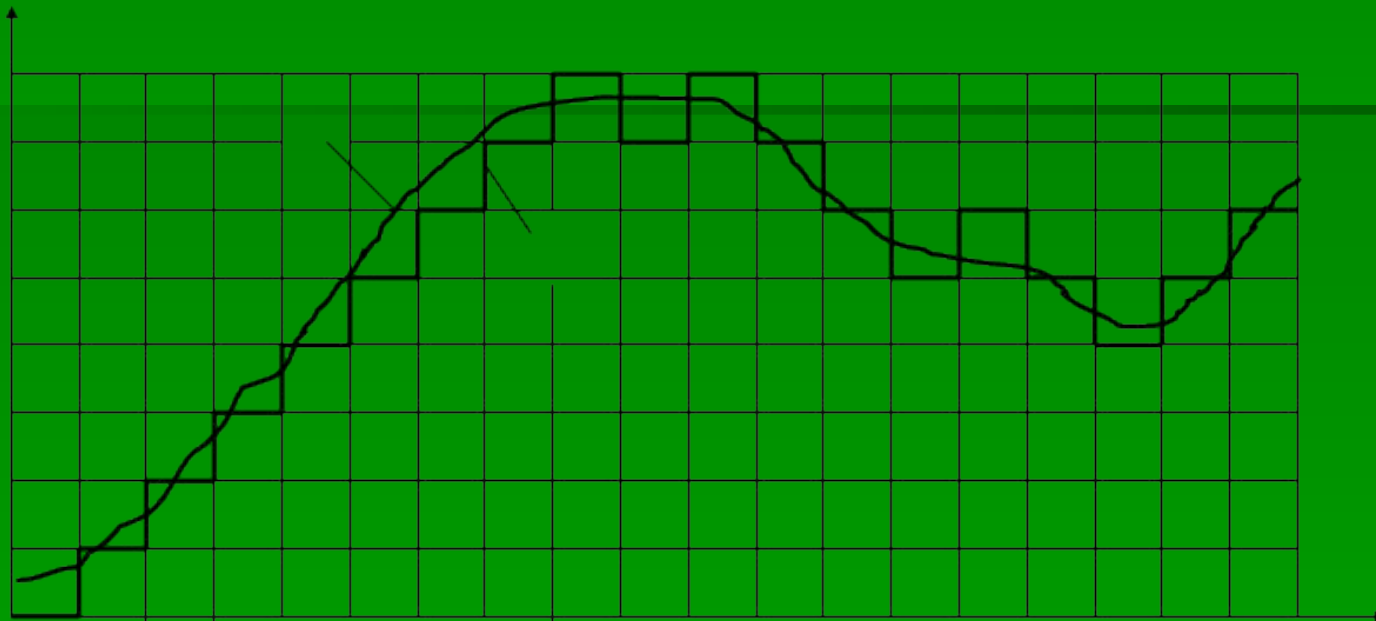
Неравномерное квантование



Равномерное квантование

Неравномерное квантование

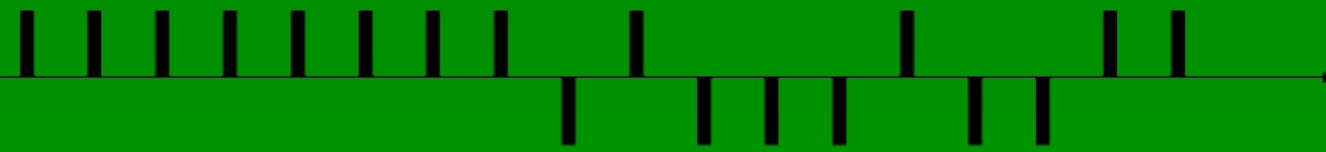
Дельта-кодирование (**DPCM**)



1 2 3 4 5 6 7 8 7 8 7 6 5 6 5 4 5 6

PCM

+
-

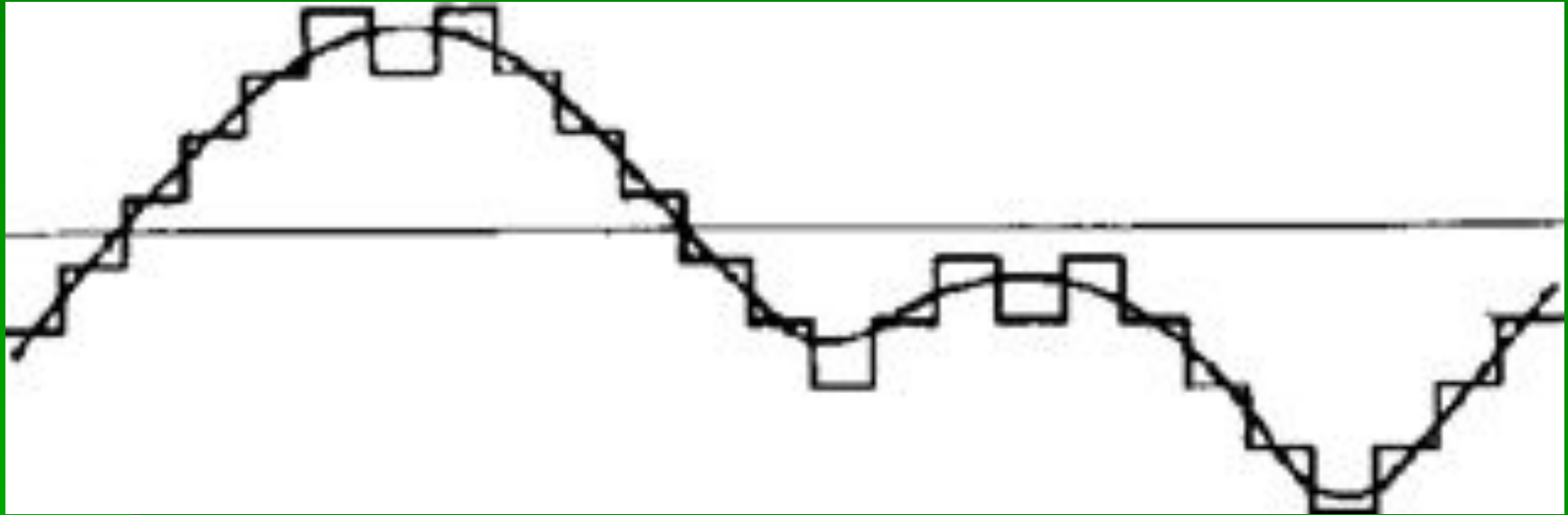


1 1 1 1 1 1 1 0 1 0 0 0 1 0 0 1 1

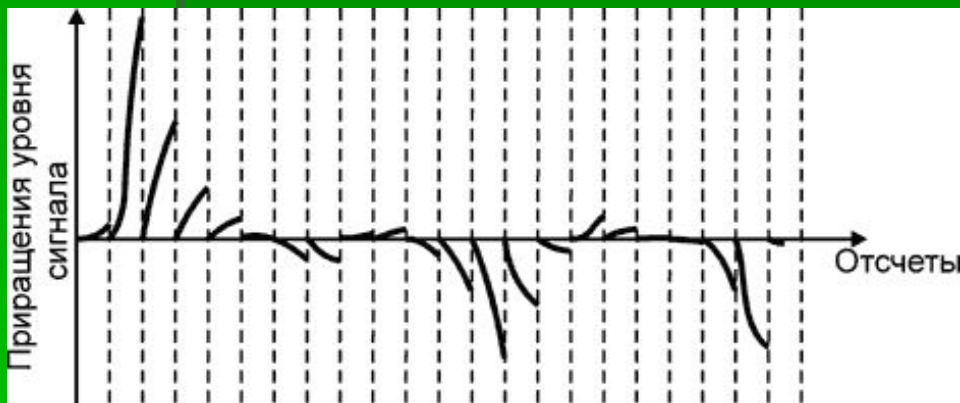
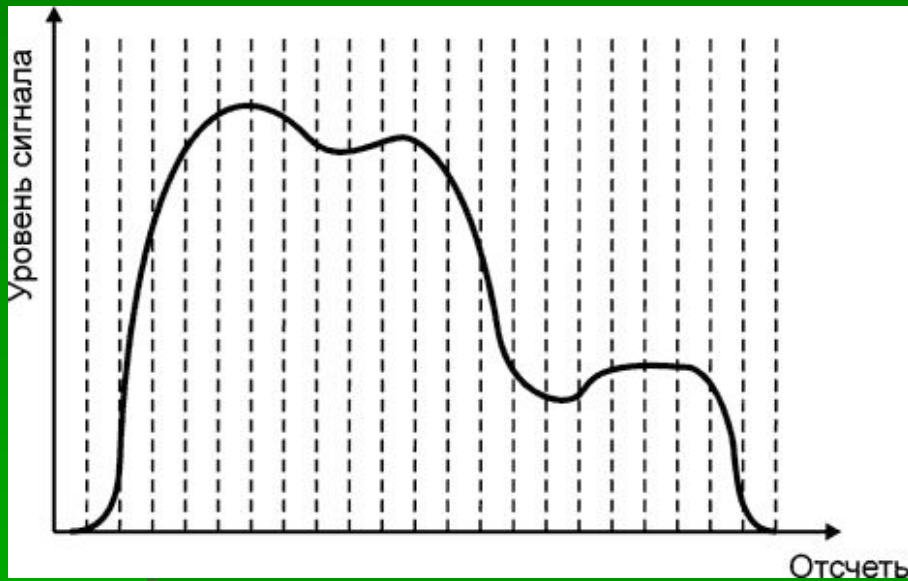
27

DPCM

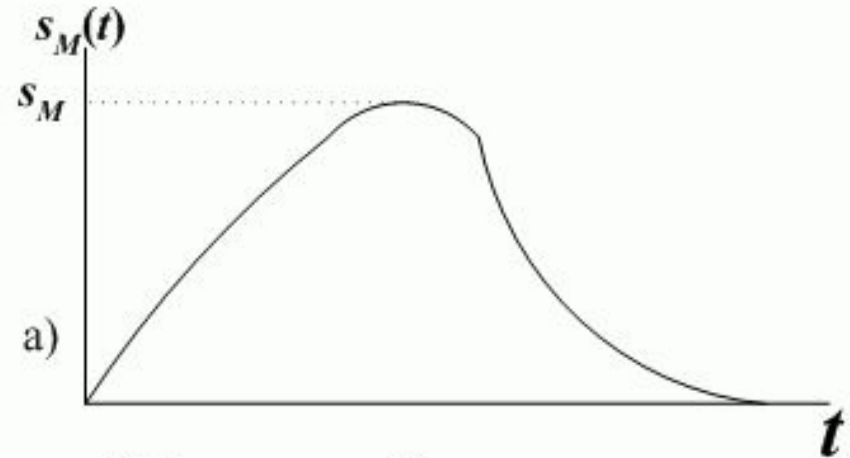
Проблемы **DPSCM**



ADPCM

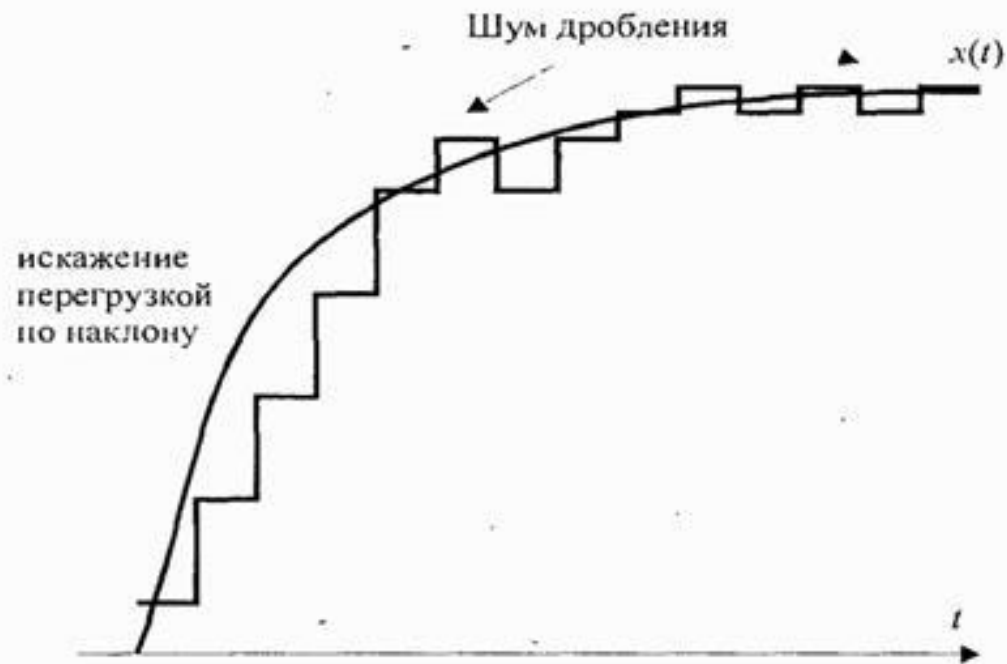
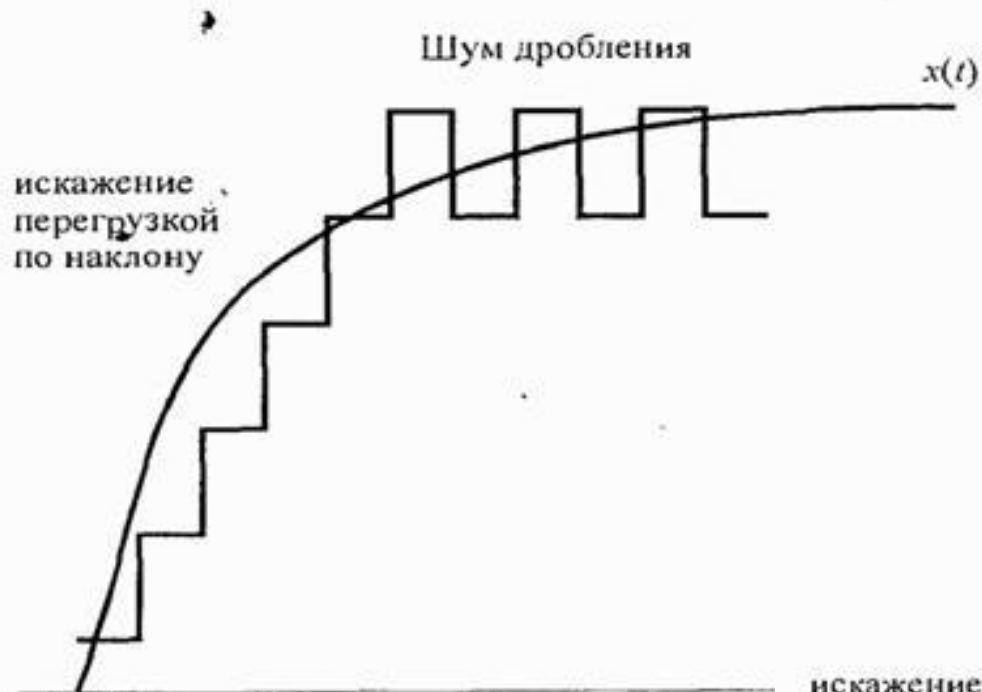


- Квантование приращений сигнала



ADPCM

- Квантование с переменным шагом



Дополнительные действия

- Канальное кодирование
- Помехоустойчивое кодирование
 - Перемежение



ThisIsAnExampleOfInterleaving...

TIEpfeaghsxlIrv.iAenli.snmOten.

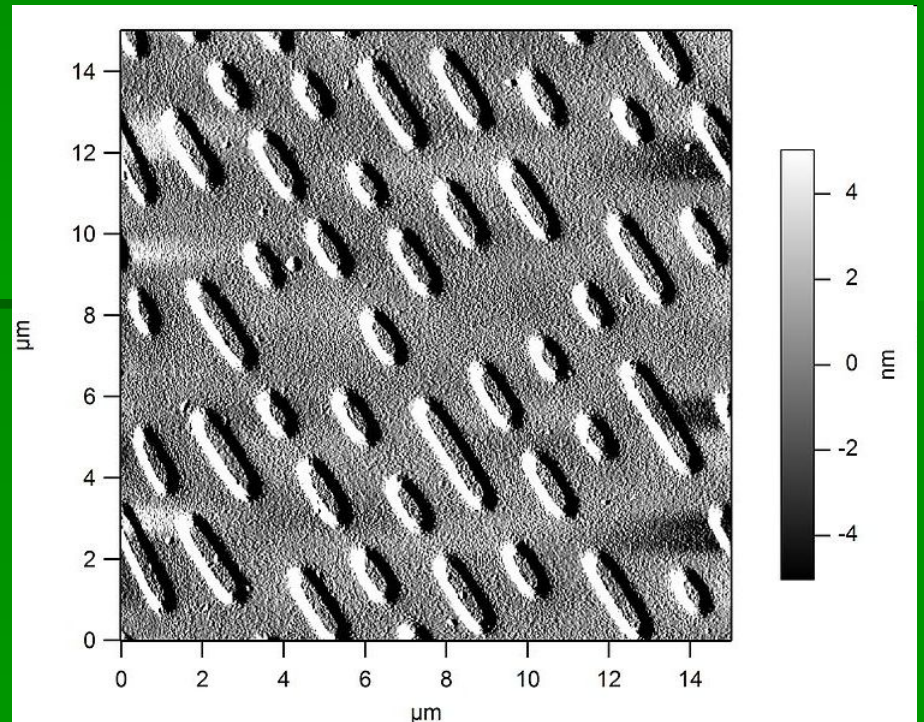
ThisIs _____pleOfInterleaving...

TIEpfe _____Irv.iAenli.snmOten.

T_isI_AnE_amp_eOfInterle_vin_...

ЗВУКОВОЙ КОМПАКТ-ДИСК

- он же *CDDA*, он же *Compact Disc Digital Audio*, он же *Audio CD*, он же *Red Book*
- 1980 год, Philips и Sony



Звуковой компакт-диск

- Максимальное время всех записей — 74 (79,8) мин.
- Минимальное время трека — 4 секунды (включая 2-секундную паузу)
- Максимальное количество треков — 99
- Максимальное число точек отсчёта (разделов трека) — 99 без ограничений по времени
- Должен присутствовать *International Standard Recording Code (ISRC)*
- Диаметр диска — 120 мм
- Диаметр центрального отверстия — 15 мм
- Толщина диска — 1,2 мм
- Материал — поликарбонат
- Воспроизведение информации — постоянная линейная скорость 1,2-1,4 м/с

Звуковой компакт-диск

- Шаг дорожки — $1,6 \pm 0,1$ мкм
- Ширина пиков — $0,4$ мкм
- Глубина пиков — $0,12$ мкм
- Длина пиков — $0,83\text{—}3,1$ мкм
- Длина волны лазера — 780 нм
- Частота дискретизации — $44,1$ кГц
- Разрядность — 16 бит (линейное квантование)
- Скорость считывания звуковой информации — $1,4112$ Мбит/с
- Общая скорость считывания информации — $1,9404$ Мбит/с
- Коррекция ошибок — CIRC (двойной код Рида-Соломона с тройным перемежением)
- Избыточность — 25%
- Максимальный объём информации, записанной на диске — 650 ($700, 800, 900$) МБ
- Скорость чтения/записи CD указывается кратной 150 Кб/с

Сравнение аудиоформатов

Название	Расширение	Квант., бит	Частота д., кГц	Число каналов	Битрейт на канал, Мбит/с	Степень сжатия /упаковки
WAVE (WAV)	.wav	8 - 32	11,025 (44,1) 192	1- 6	до 6,144	1:1
Digital eXtreme Definition (DXD)	-	24; 32	352,8	2, 5.1	8,4672; 11,2896	1:1
FLAC	.flac	4 — 32	1Гц — 655.350 кГц	1 — 8	-	1.4:1 — 4:1
Windows Media Audio 9 Lossless	.wma	16; 24	8 — 96	до 6	-	1.7:1 — 3:1
AC-3	.ac3, .vob и др.	16-24	32; 44.1; 48	6 — 13.1	32 — 640, до 6144	-
MP3 (MPEG-1, 2, 2.5 Audio Layer III)	.mp3	16-24	8 — 48	до 2	8 — 320	~11:1
Vorbis (Ogg)	.ogg	до 32	8 — 192	до 255	до 1000	-

MIDI

- *Musical Instrument Digital Interface* — цифровой интерфейс музыкальных инструментов
- Секвенсоры

Блок 2

Сжатие данных

Медиаконтейнер

- – формат файла или потока данных, определяющий только способ сохранения (т. е. внутренней структуры этого файла) данных, но НЕ алгоритм кодирования.

Примеры:

- Графика: JPEG, TIFF
- Аудио: WAV
- Многоцелевые: 3GP, Matroska, AVI

Кодек

- (*codec*, от *coder/decoder* — шифратор/дешифратор — кодировщик/декодировщик или *compressor/decompressor*) — устройство или программа, способная выполнять преобразование данных или сигнала.
- видеокодек
- аудиокодек
- сжимающие без потерь (*lossless codec*)



Студенту на заметку

FFmpeg



- - набор утилит и библиотек для записи, конвертации и проигрывания практически любого видео и аудио.
- Бесплатно (лицензия GNU LGPL Бесплатно (лицензия GNU LGPL или GNU GPL), без регистрации, работает под Linux, Windows, и не только.
- Консольное приложение.
- Включает видеопроигрыватель и http-сервер для потокового вещания (можно сделать свой Youtube).
- Статья «19 команд ffmpeg для любых⁴⁰ нужд».

Сжатие данных без потерь

- *Lossless data compression*
- Информационная энтропия
- Энтропийное кодирование
- Теоремы Шеннона для источника общего вида

$$E_{U^w}(U) < \frac{H(U)}{\log_2 D} + 1$$

Префиксный код

- Условие Фано: если в код входит слово a , то для любой непустой строки b слова ab в коде не существует.
- ~~0, 1, 00, 01~~, 10, 11 =>
 - 01001101110 ~ 0 10 0 11 0 11 10
- Пример сжатия:
 - 00 → 0, 01 → 10, 10 → 110, 11 → 111
=>
 - 00 01 00 00 11 10 00 00 (16 бит)
 - ~
 - 0 10 0 0 111 110 0 0 (13 бит)

Код Хаффмана

- Принцип:

Частоты:

15 7 6 6 5

А Б В Г Д

~

Коды

А	Б	В	Г	Д
0	100	101	110	111

- Применение

- сжатие JPEG, MPEG
- архиваторы PKZIP, LZH и др.
- протоколы передачи данных MNP5 и MNP7

Кодирование длин серий

- Кодирование повторов, *Run-length encoding*, *RLE*

- Принцип:

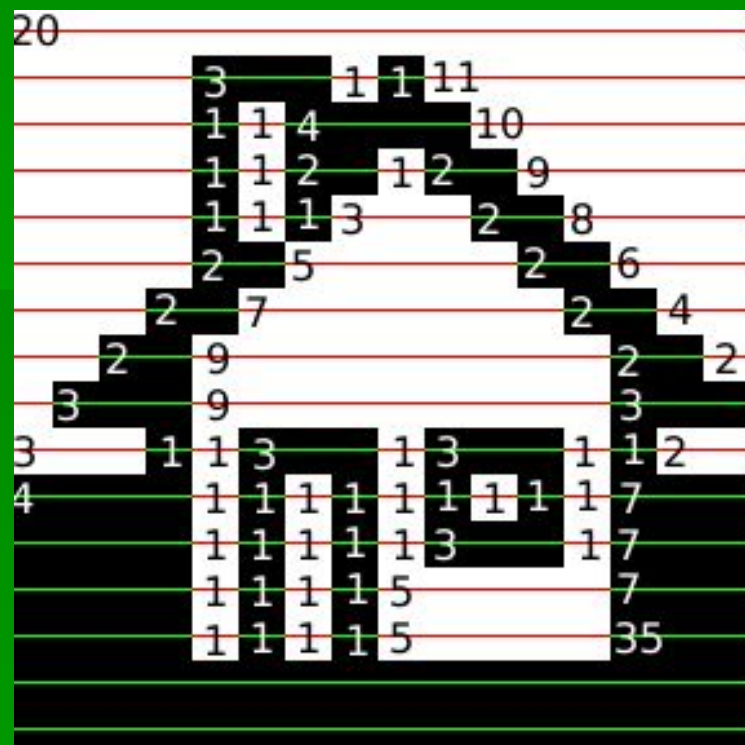
- ```
WWWWWWWWWWWWWWWWWWBWWWWWWWWWWWWWWWWWWBWB
BWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWBWB
WWWWWWWWWWWWWWWWWW
```

~

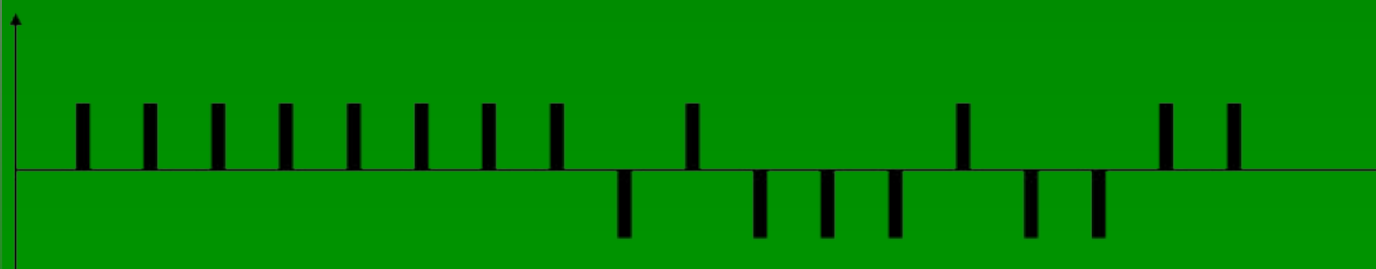
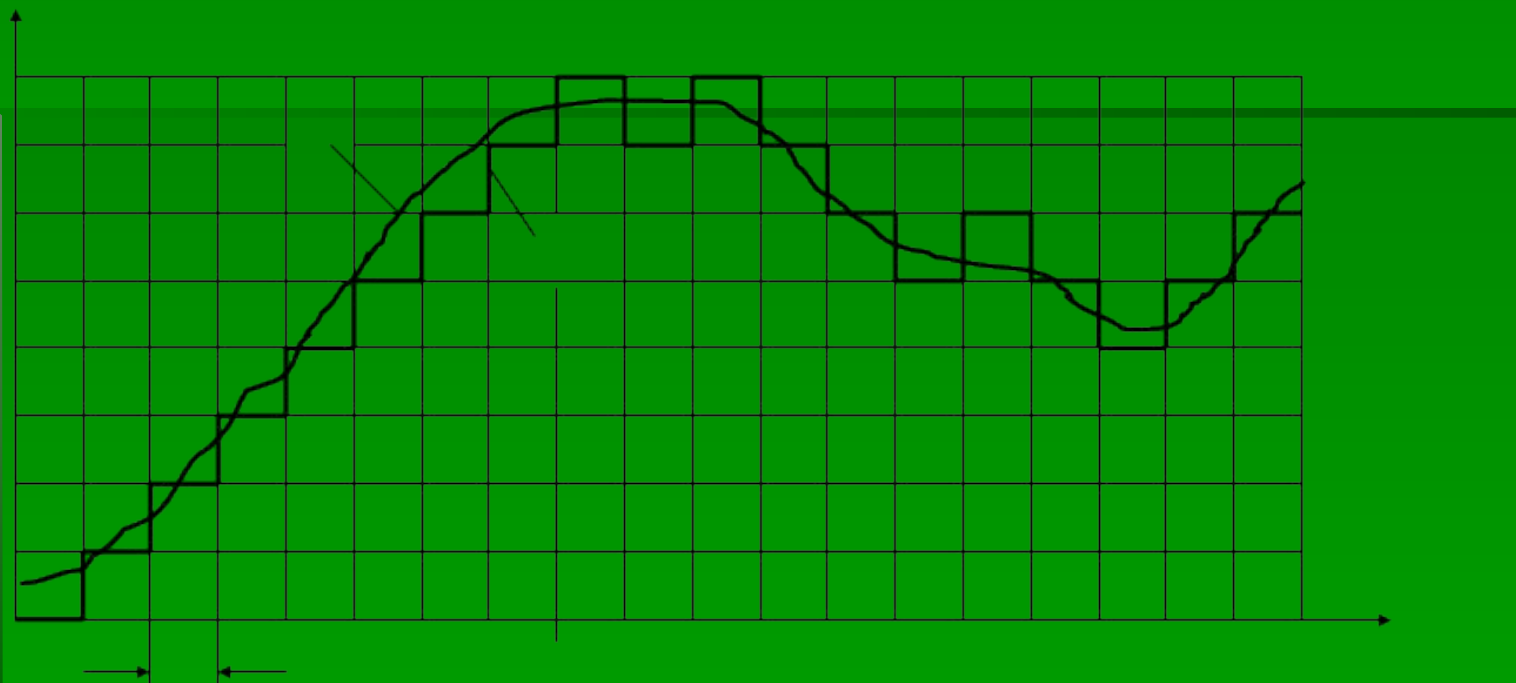
- ```
12W1B12W3B24W1B14W
```

- Применение:

- графика: BMP, PCX
- звук (после Дельта-кодирования)



RLE и DPCM



1 1 1 1 1 1 1 0 1 0 0 0 1 0 0 1 1 45

Алгоритм Лемпеля — Зива — Велча

- *Lempel-Ziv-Welch, LZW*
- Применение: GIF, TIFF, PDF
 - *PNG's Not GIF*

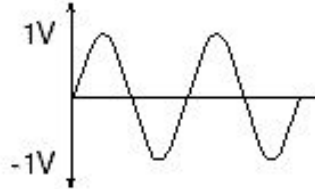

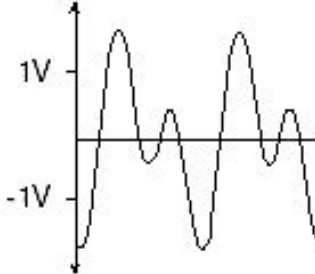
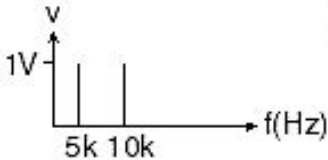
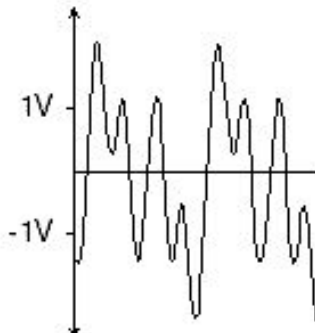
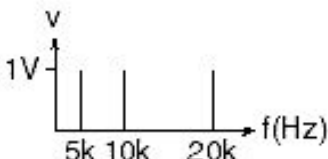
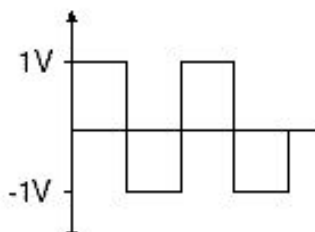
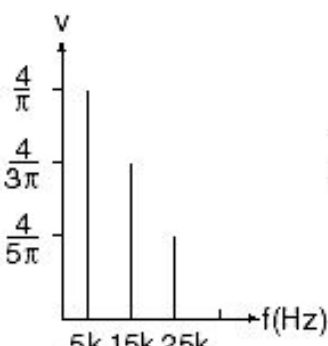
Сжатие данных с потерями

- *lossy compression*
- Варианты:
 - Трансформирующее
 - фрейм целиком
 - поблочно
 - Предсказывающее (предиктивное)

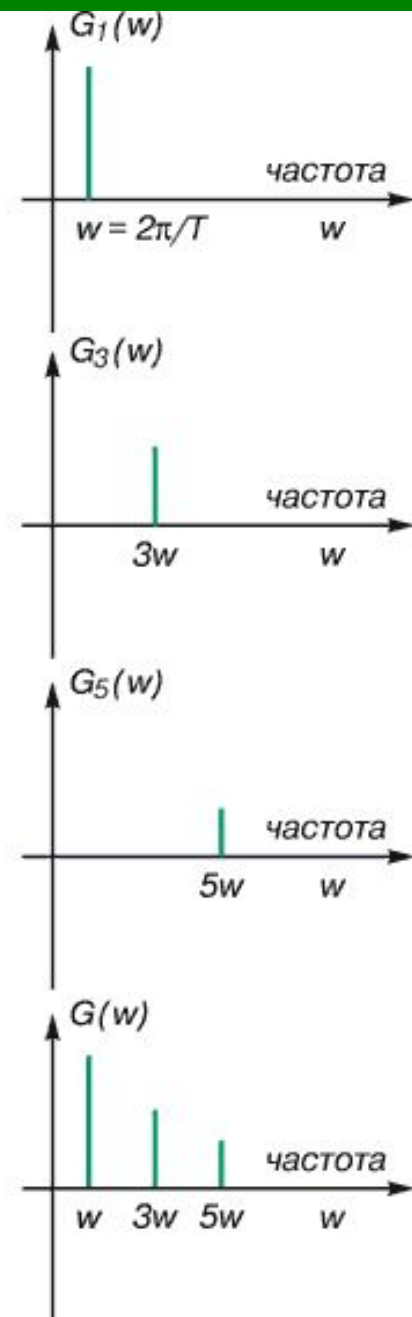
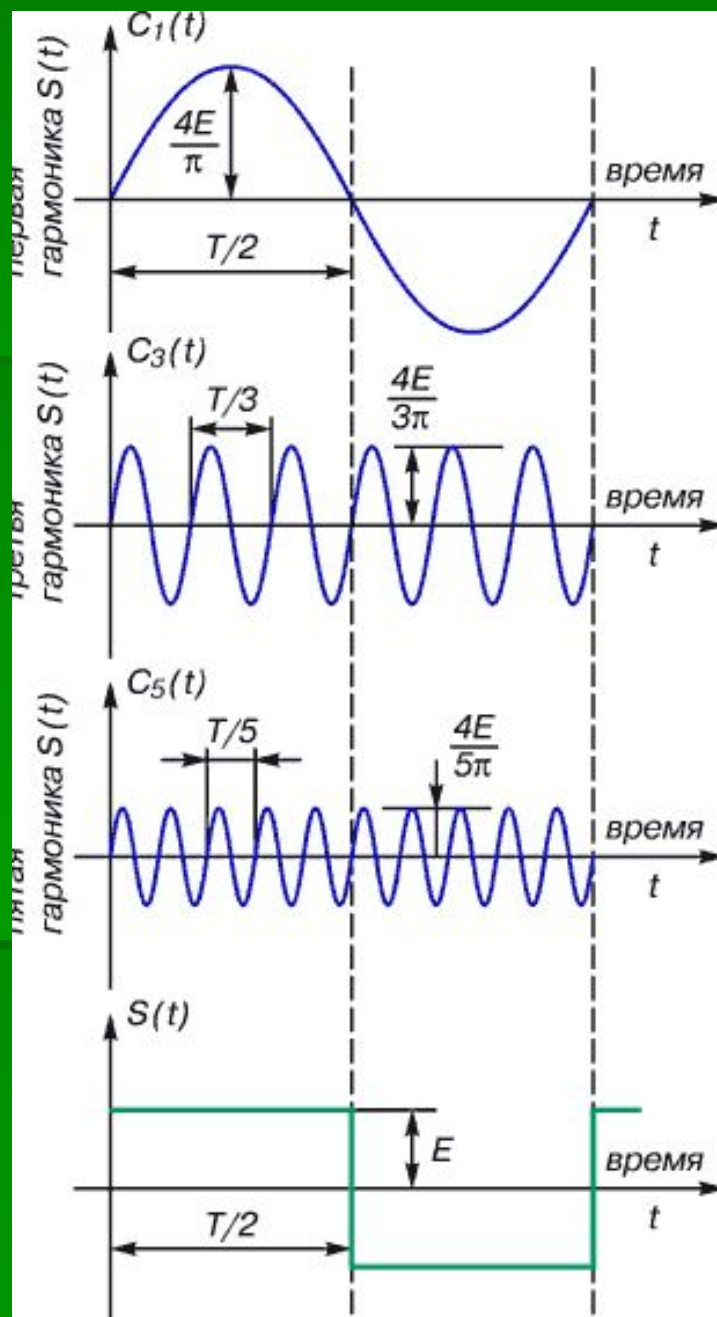
Преобразование Фурье

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ix\omega} dx.$$

Преобразование Фурье

Description	Time Series	Fourier Expansion	Power Spectrum
A pure 5kHz sine wave measuring 1 volt peak		$v(t) = 1\sin(\omega_1)t$ $\omega_1 = 2\pi(5\text{kHz})$	
A pure 5kHz and 10kHz sine wave, each measuring 1 volt peak, added together		$v(t) = 1\sin(\omega_1)t + 1\sin(\omega_2)t$ $\omega_1 = 2\pi(5\text{kHz})$ $\omega_2 = 2\pi(10\text{kHz})$	
A pure 5kHz, 10kHz, and 20kHz sine wave, each measuring 1 volt peak, added together		$v(t) = 1\sin(\omega_1)t + 1\sin(\omega_2)t + 1\sin(\omega_3)t$ $\omega_1 = 2\pi(5\text{kHz})$ $\omega_2 = 2\pi(10\text{kHz})$ $\omega_3 = 2\pi(20\text{kHz})$	
A pure 5kHz square wave measuring 1 volt		$v(t) = \frac{4}{\pi}\sin(\omega_1)t + \frac{4}{3\pi}\sin(\omega_2)t + \frac{4}{5\pi}\sin(\omega_3)t \dots$ $\omega_1 = 2\pi(5\text{kHz})$ $\omega_2 = 2\pi(15\text{kHz})$ $\omega_3 = 2\pi(25\text{kHz}) \dots$	

Преобразование Фурье



MP3



(a)

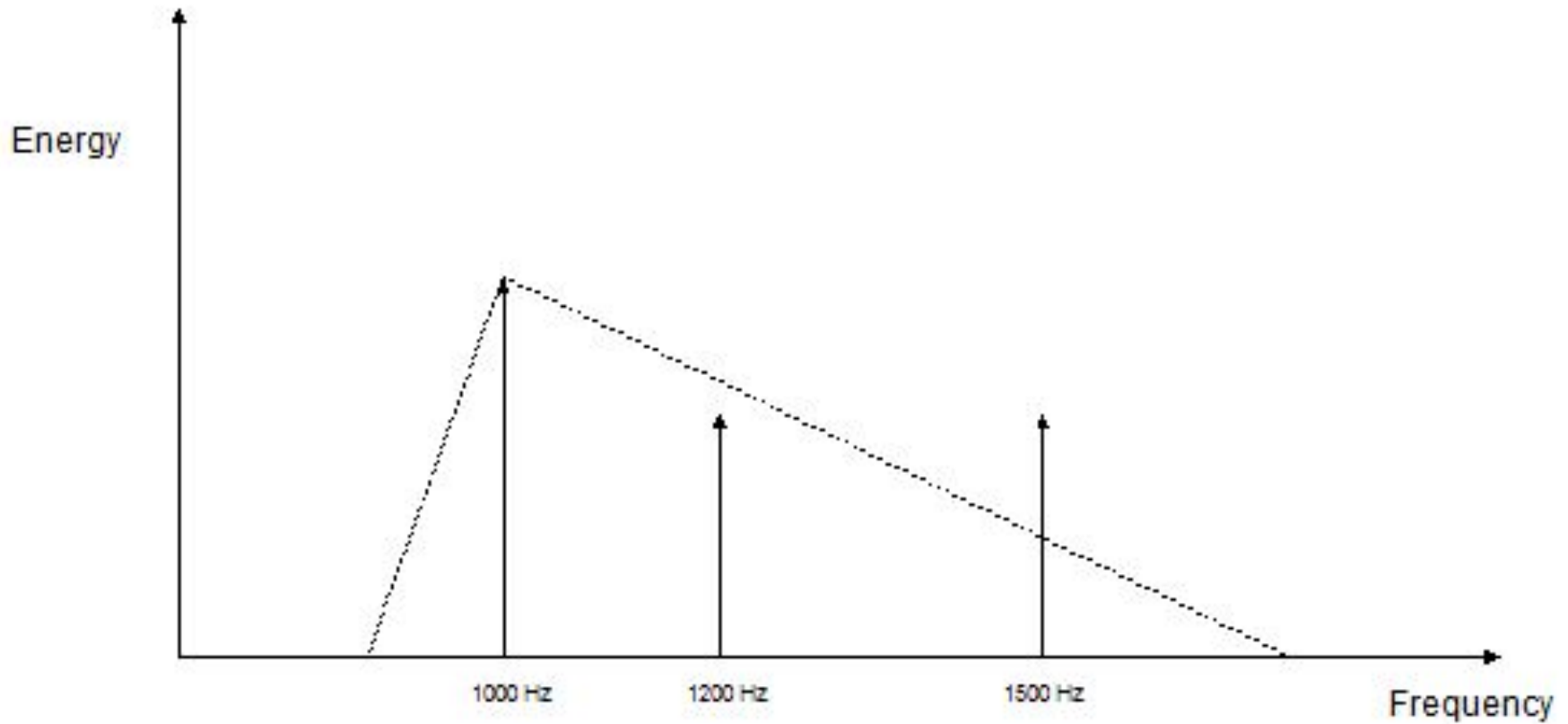


(b)

Эффект Хааса (*Haas*)

- - неспособность человека выявлять искажения в импульсах длительностью менее 20 мс

Эффект маскировки



Битрейт

- Здесь – степень сжатия потока аудио- или видеоданных.
- Ширина потока: определяет, сколько бит необходимо для кодирования 1 секунды звука.

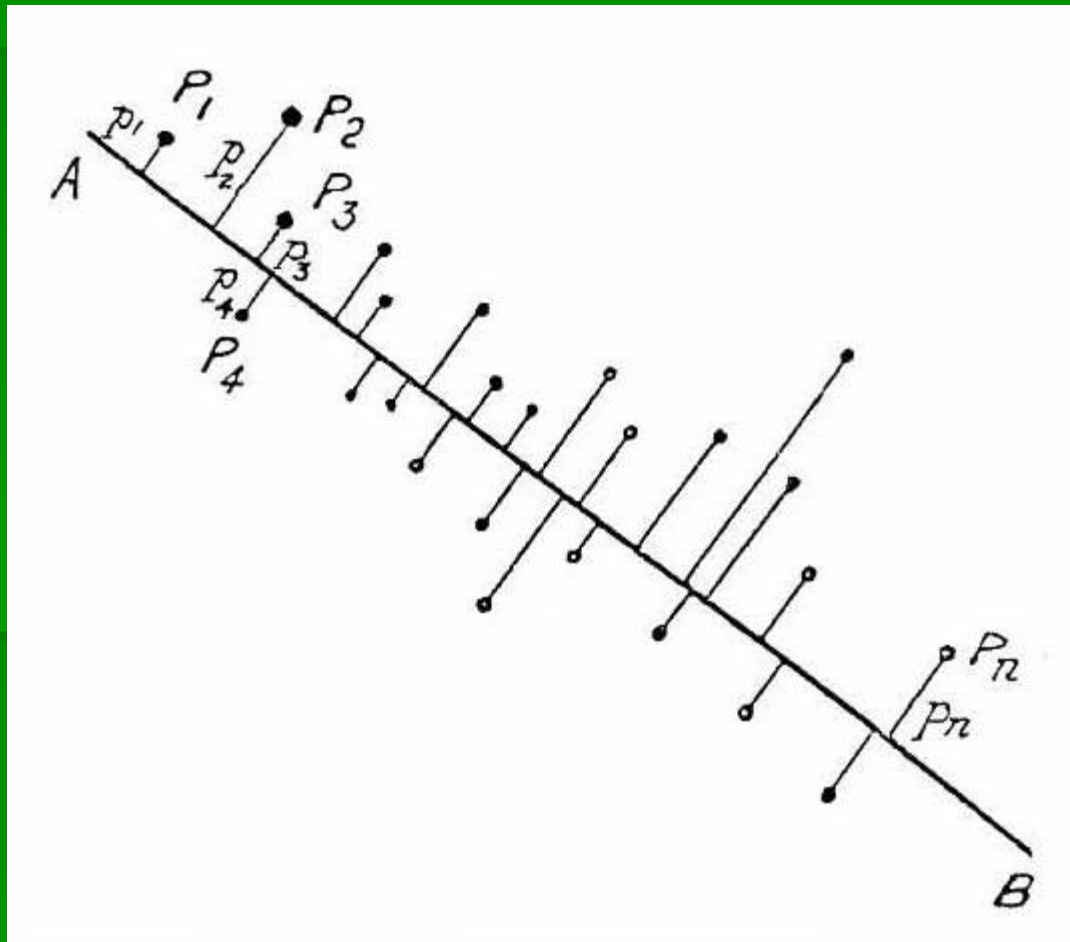
Бывает:

- постоянным (англ. *Constant bitrate*, CBR)
 - переменным (англ. *Variable bitrate*, VBR)
 - усреднённым (англ. *Average bitrate*, ABR)
-
- чаще всего измеряют в килобитах (мегабитах) в секунду (*kilobit per second*, kbps; Mbit/s, Mbps)

Компрессия изображений

- Снижение глубины цвета
- Метод главных компонент
- Фрактальное сжатие
- Сжатие на основе предсказателей
 - JPEG-LS
 - ДИКМ
 - Иерархическая сеточная интерполяция
 - CALIC
- JPEG
- Вэйвлетная компрессия
 - JPEG 2000
 - DjVu

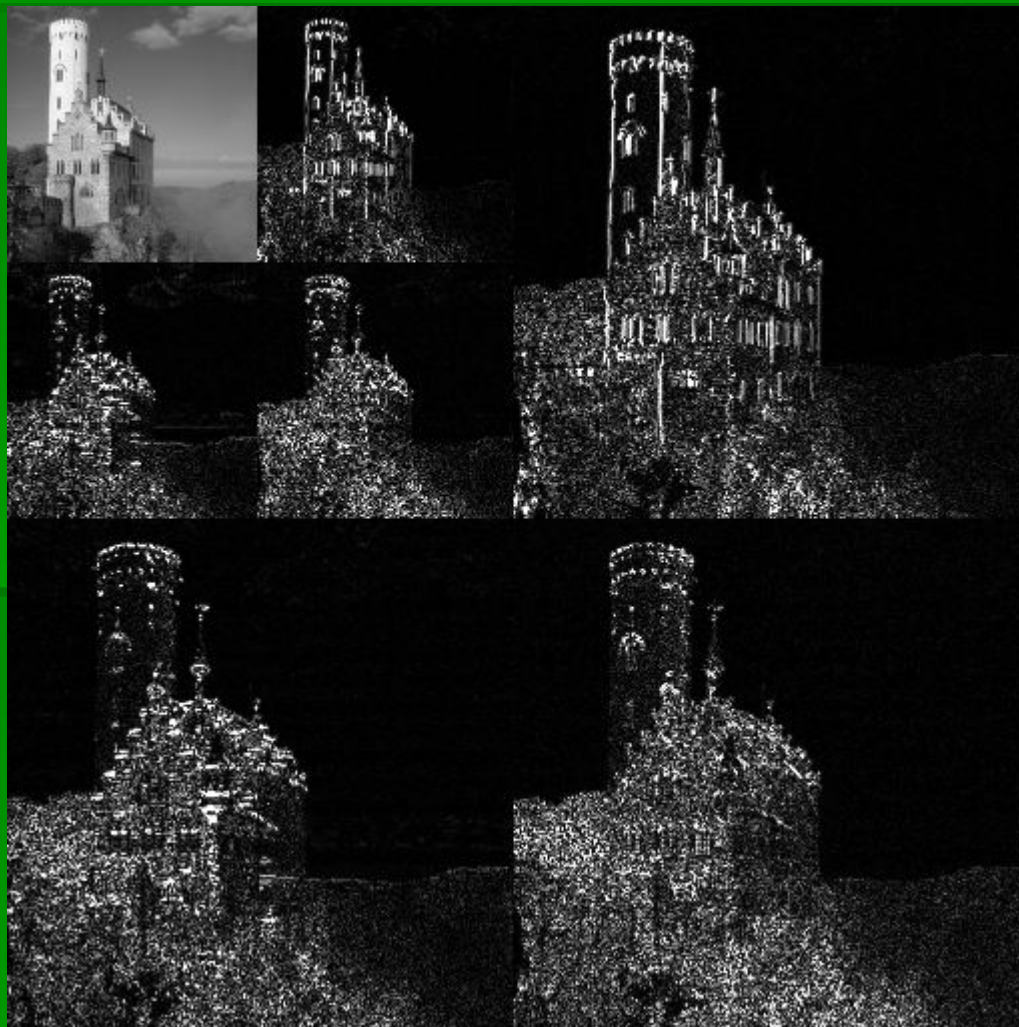
Метод главных компонент



Фрактальное сжатие



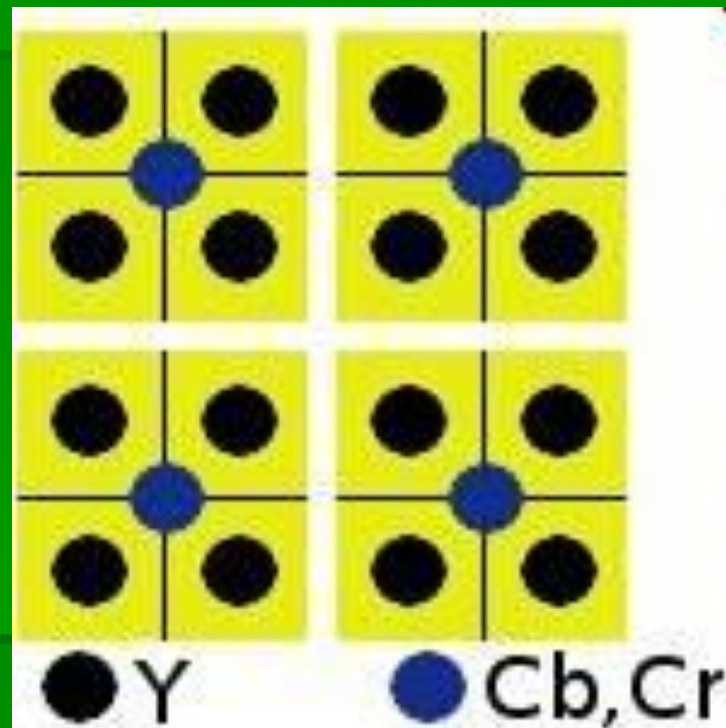
Дискретное вейвлет-преобразование



JPEG

1. преобразование из цветового пространства RGB в YCbCr
2. «прореживание» (*subsampling*)
3. дискретное косинусное преобразование
4. квантование коэффициентов ДКП
5. сжатие с использованием кодирования серий и кодов Хаффмана

Прореживание



Дискретное косинусное преобразование

75	76	75	75	69	66	77	71
73	74	73	74	63	64	68	69
69	68	71	72	67	58	48	41
59	55	56	52	47	40	24	9
51	50	45	41	33	22	7	-5
43	37	32	24	15	5	-6	-25
29	21	9	-2	-10	-21	-44	-69
9	-4	-17	-35	-52	-61	-57	-35

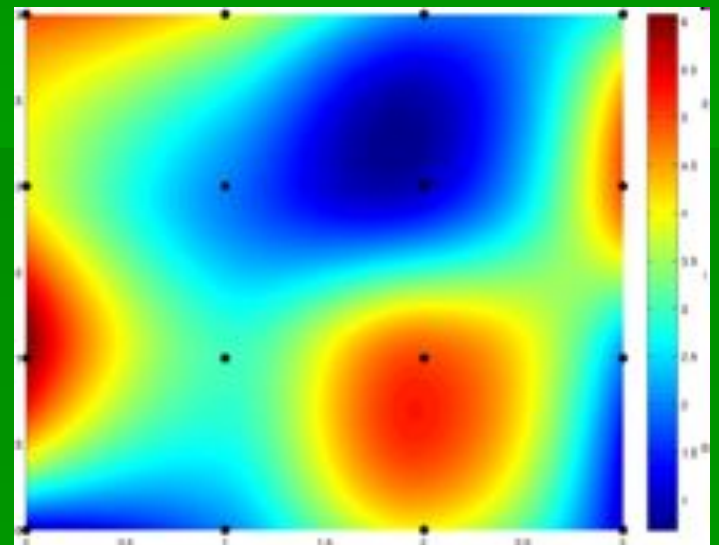
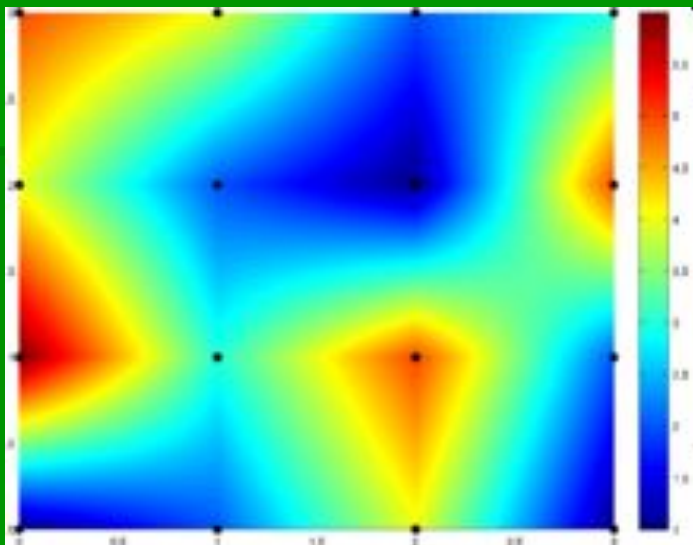
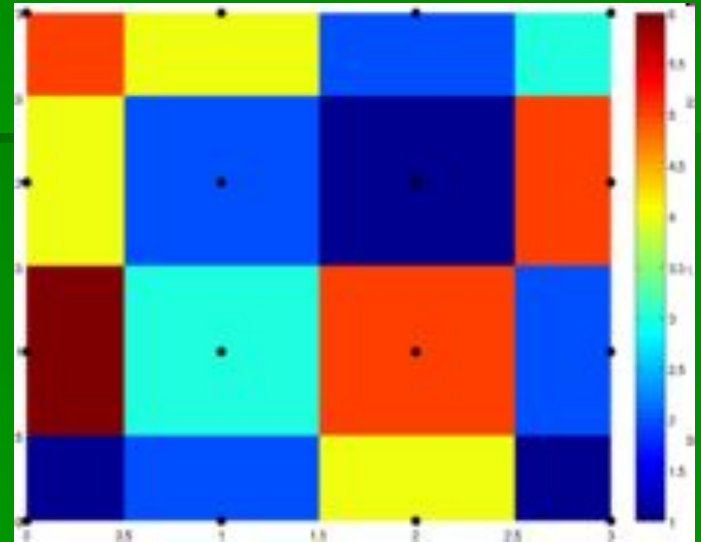
251	118	-13	6	-2	6	-1	0
279	-68	-8	-7	-1	4	-4	-1
-51	-14	34	-14	5	0	-1	0
27	5	-10	8	-7	4	-5	1
-22	-7	14	-9	4	-2	1	1
-3	15	-18	15	-6	2	-1	2
7	-9	6	-6	4	0	0	2
3	7	-9	3	0	-2	-1	0

Пример сжатия **JPEG** с разными коэффициентами

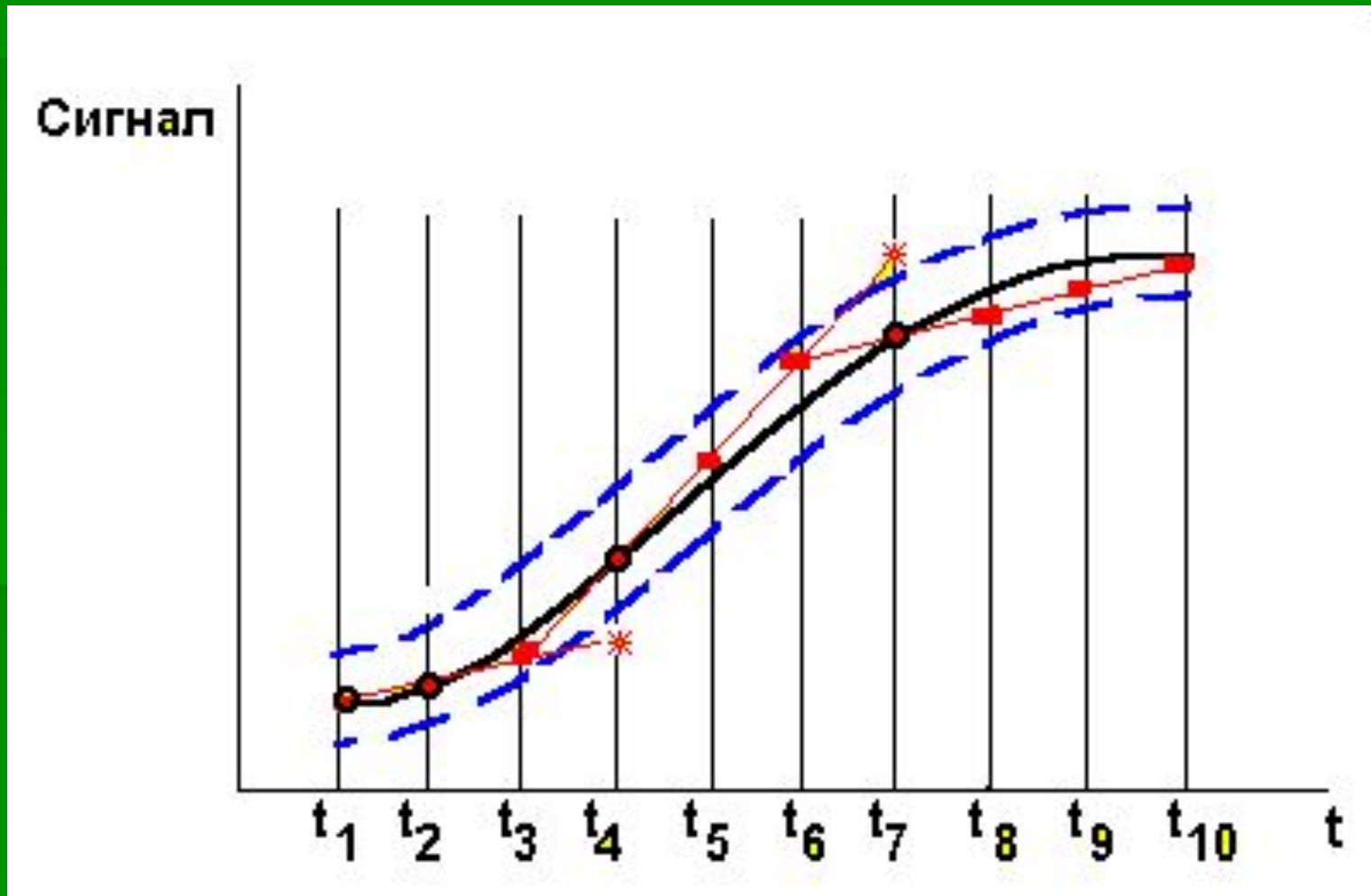


Сглаживание и интерполяция

- Ближайший сосед
- Билинейная
- Бикубическая



Предиктивное сжатие



Блок 3

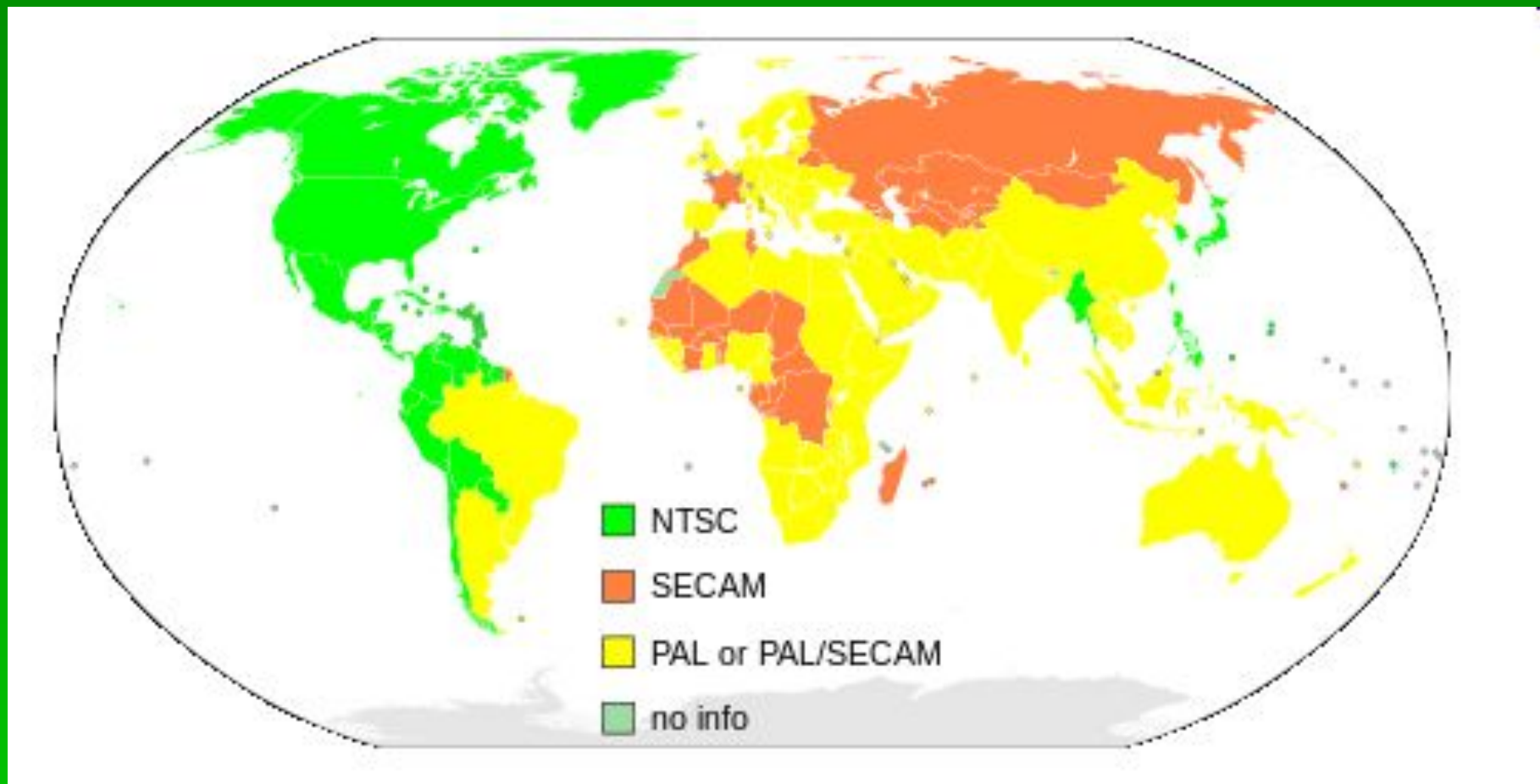
Особенности кодирования и сжатия видеоданных

YUV

- 3 компоненты — яркость (Y) и две цветоразностных (U и V)
- YPbPr
- YCbCr



Стандарты телевидения



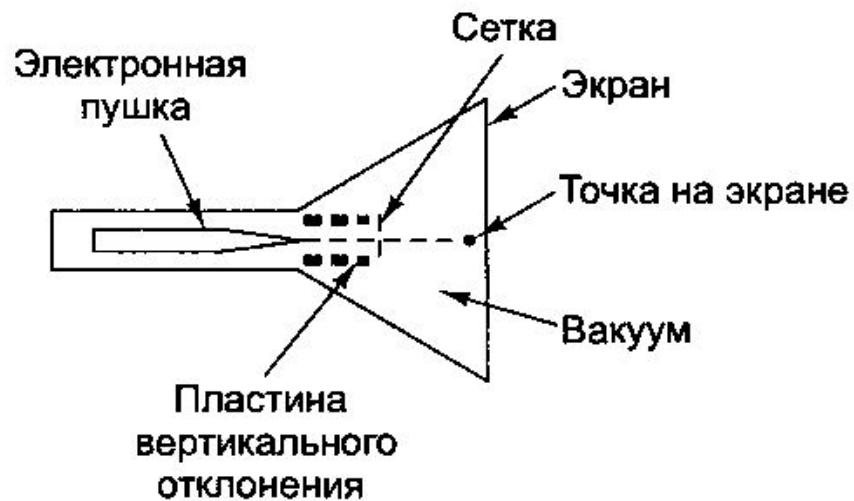
Стандарты телевидения

- NTSC
 - 525 (480) строк
 - 59,94 полей в секунду
 - в Америке и Японии
- PAL/SECAM
 - 625 (576) строк
 - 50 полей в секунду
 - в Европе (PAL)
 - во Франции, России, Китае и некоторых странах Ближнего Востока (SECAM)
- HDTV
 - 720 строк / 50 полей, 60 полей / 30 кадров, 25 кадров, 24 кадра
- Full HD
 - 1080 строк / 50 полей, 60 полей / 30 кадров, 25 кадров, 24 кадра

Разрешение видео



Развёртка



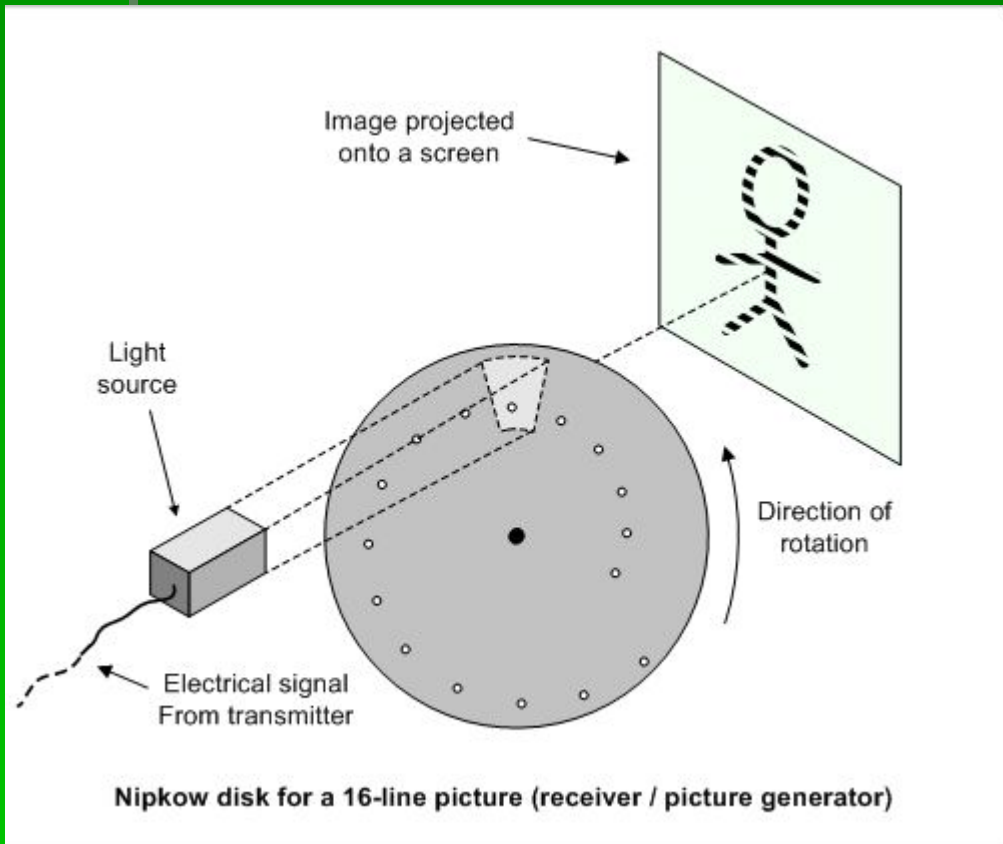
а



б

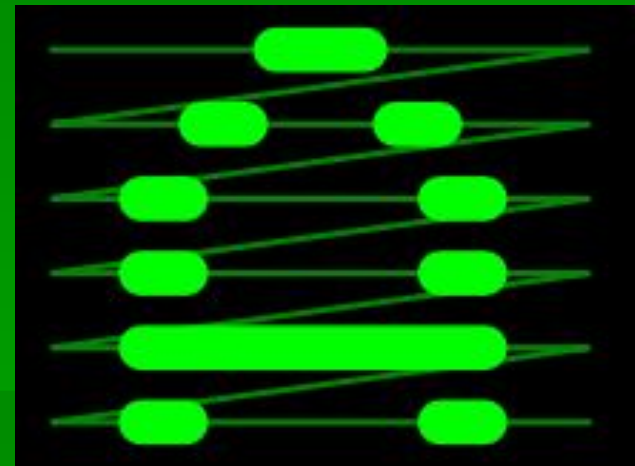
Механическая развёртка

■ Диск Нипкова



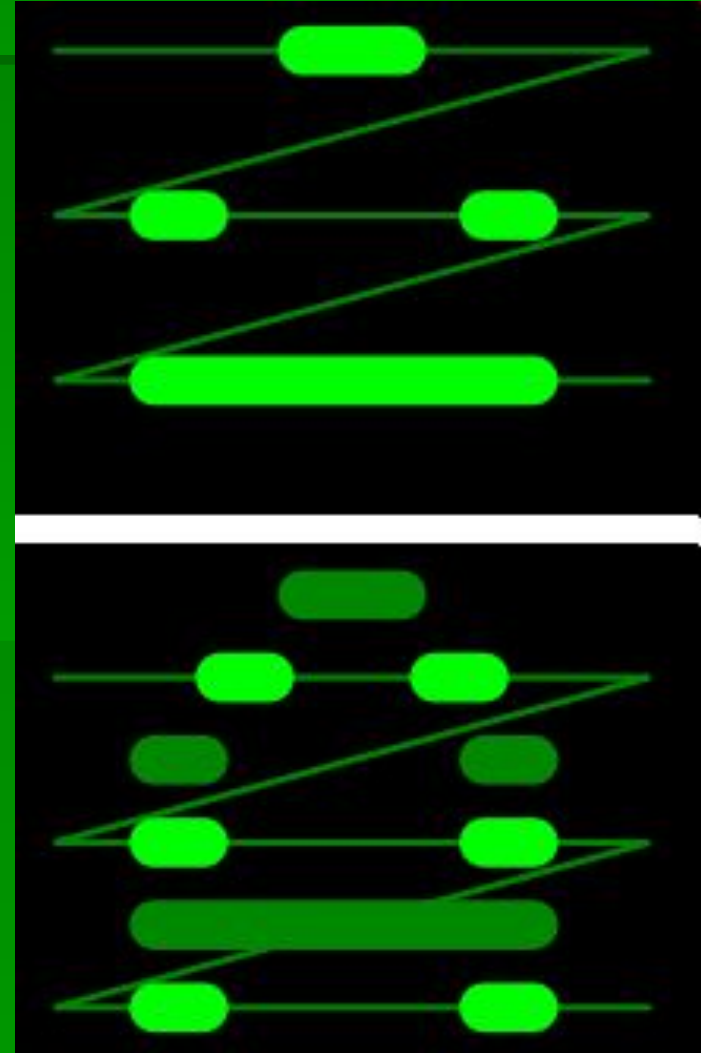
Прогрессивная развёртка

- Отсутствие визуальных искажений.
- Нет необходимости применять сглаживание.
- Видеоизображение можно масштабировать до большего разрешения быстрее и качественнее.
- Кадр может быть сохранен как отдельная фотография.



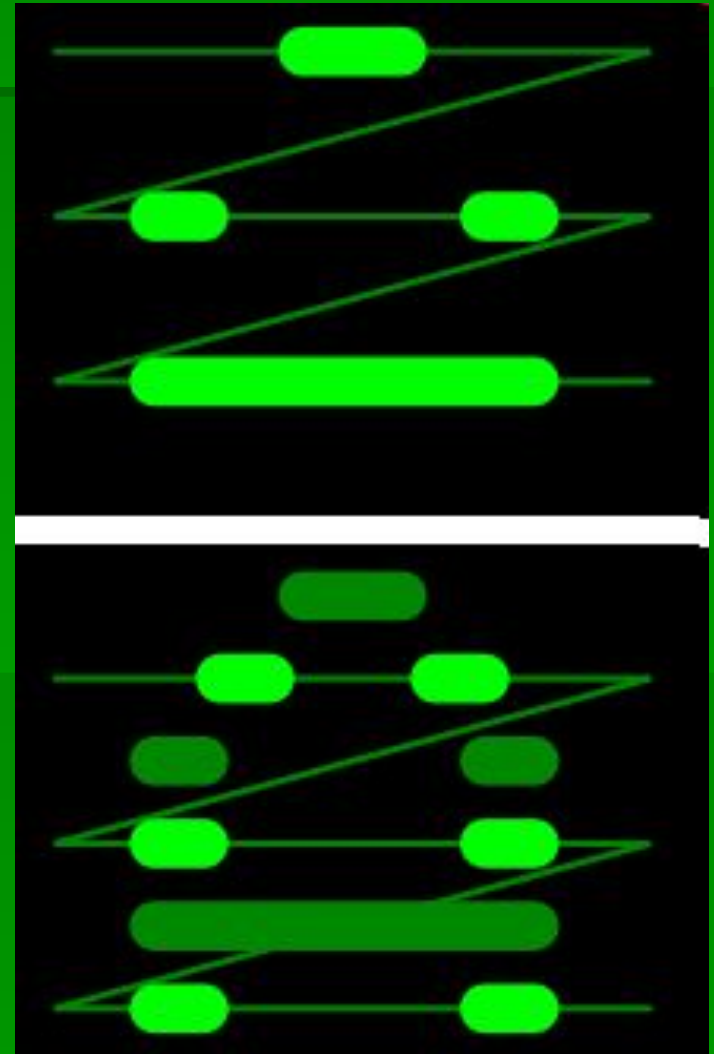
Чересстрочная развёртка

- Уменьшение полосы частот вдвое.
- Упрощение оборудования.



Деинтерлейсинг

- *Deinterlacing*
- — устранение
чересстрочности



Варианты устранения



а) Чересстрочная развёртка



б) Сглаживание

Варианты устранения



в) Интерполяция



г) Адаптивный метод

Рис. 7. Применение деинтерлейсинга

Компенсация движения

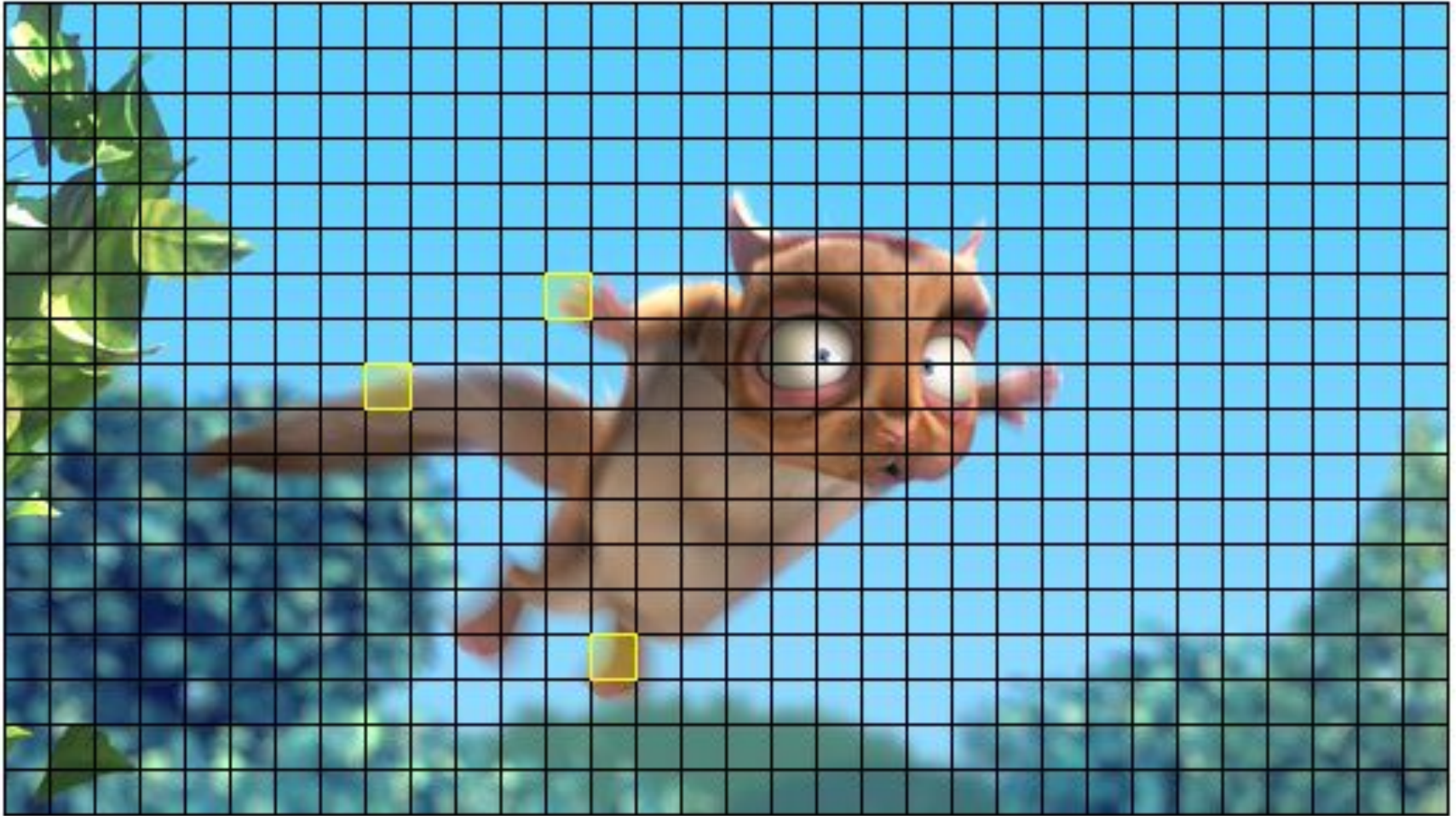
- *Motion Compensation*

Компенсация движения

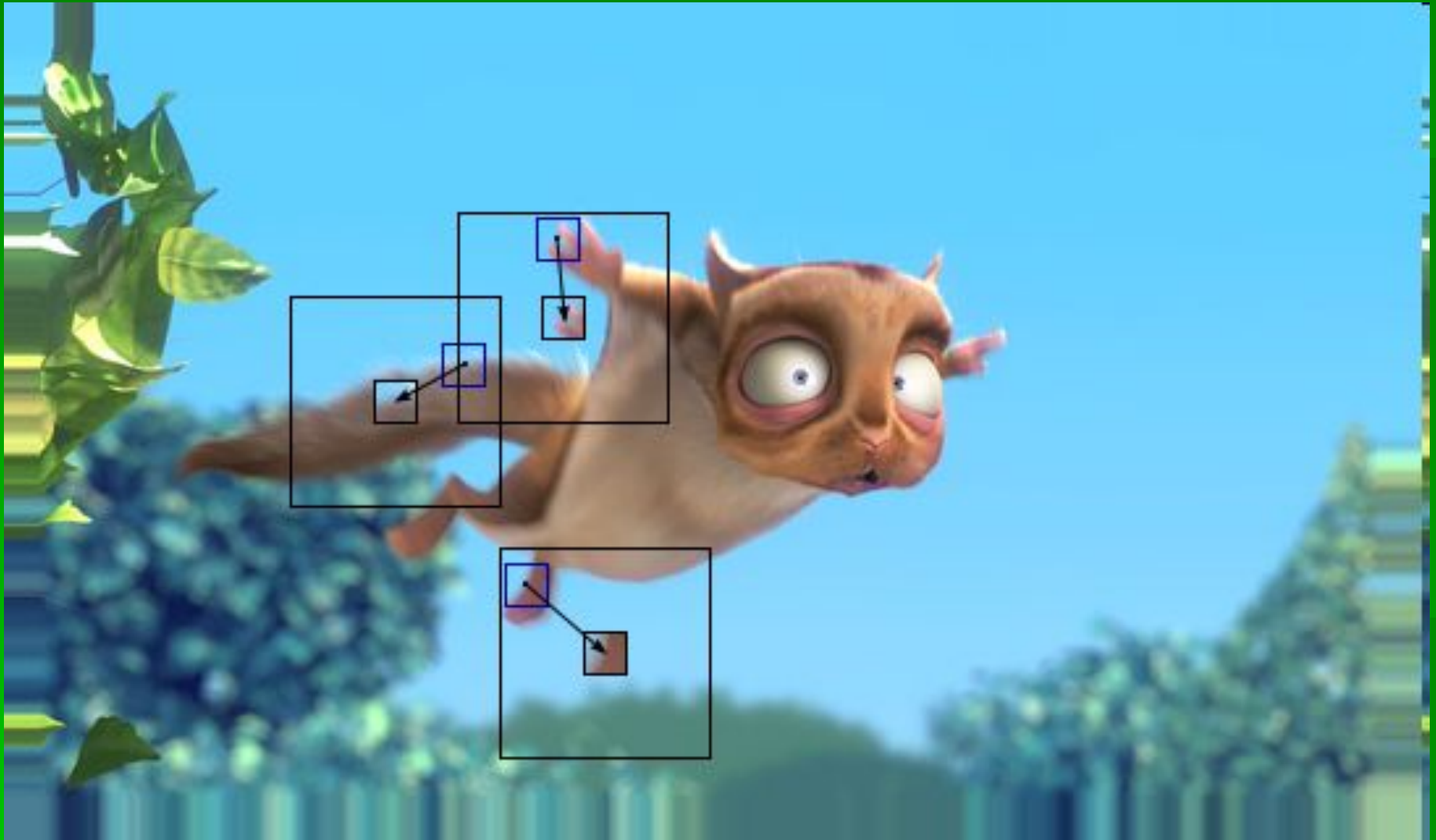
Телесная восприимчивость (Big Buck Bunny)



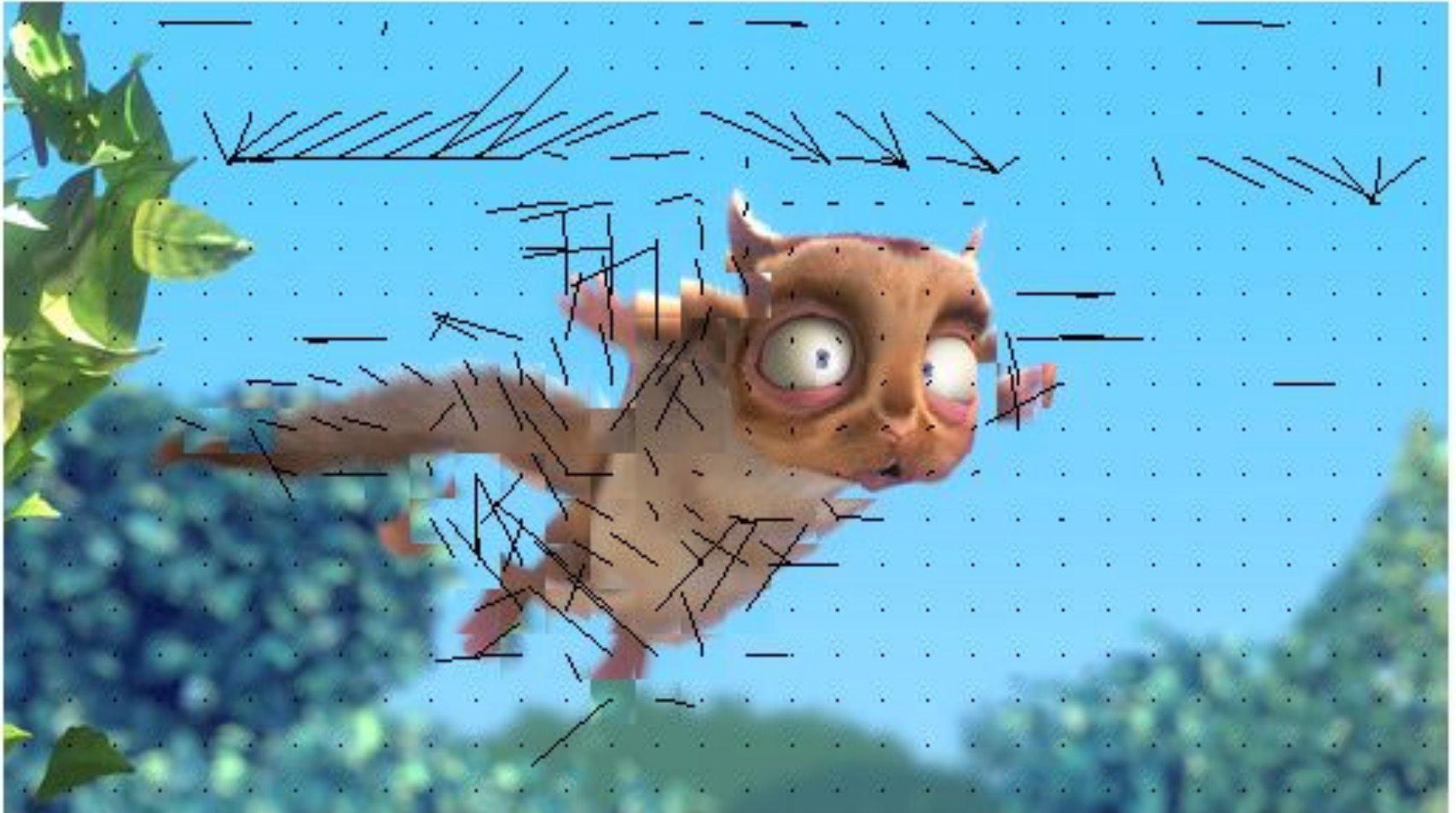
Компенсация движения



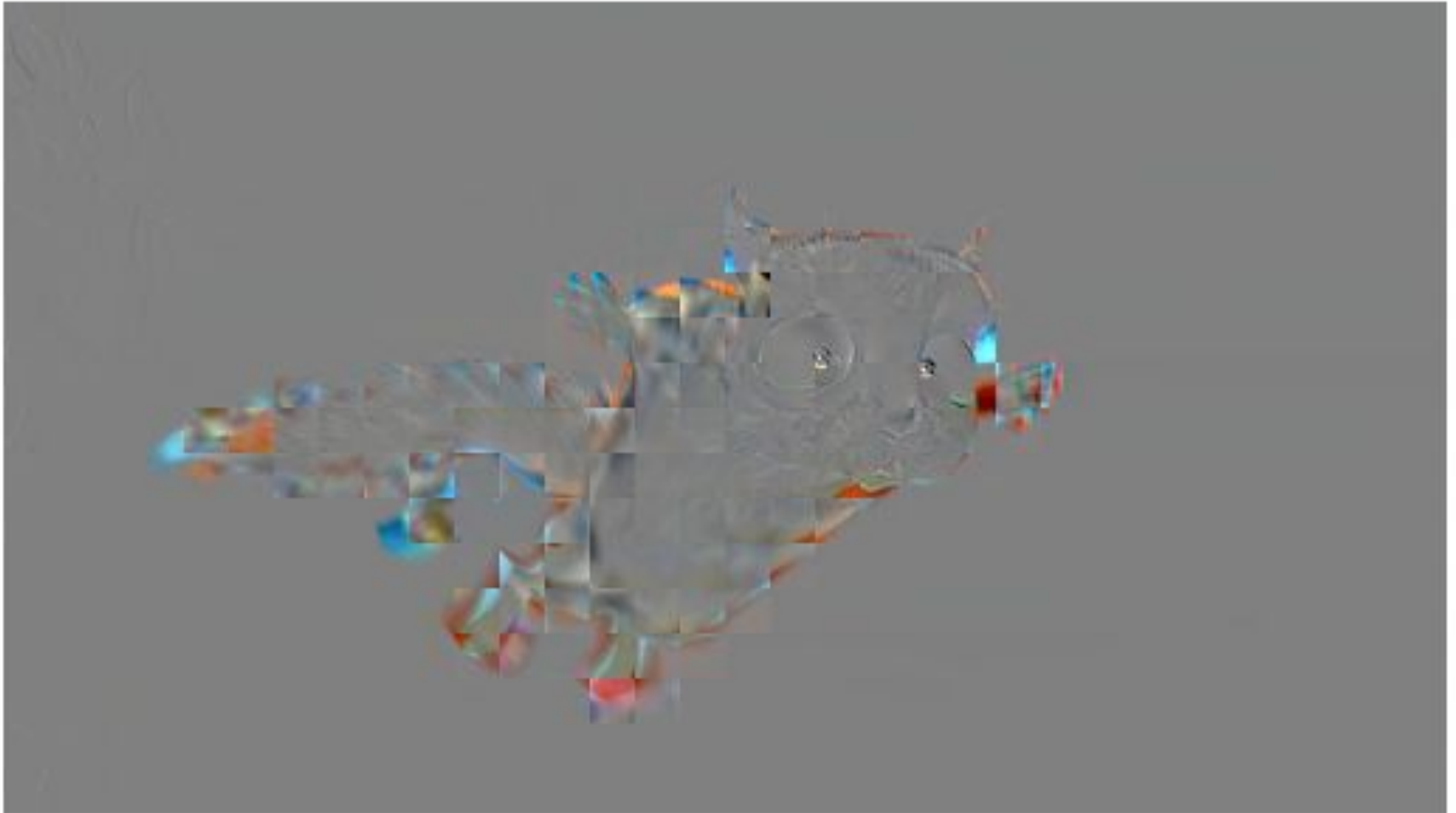
Компенсация движения



Компенсация движения

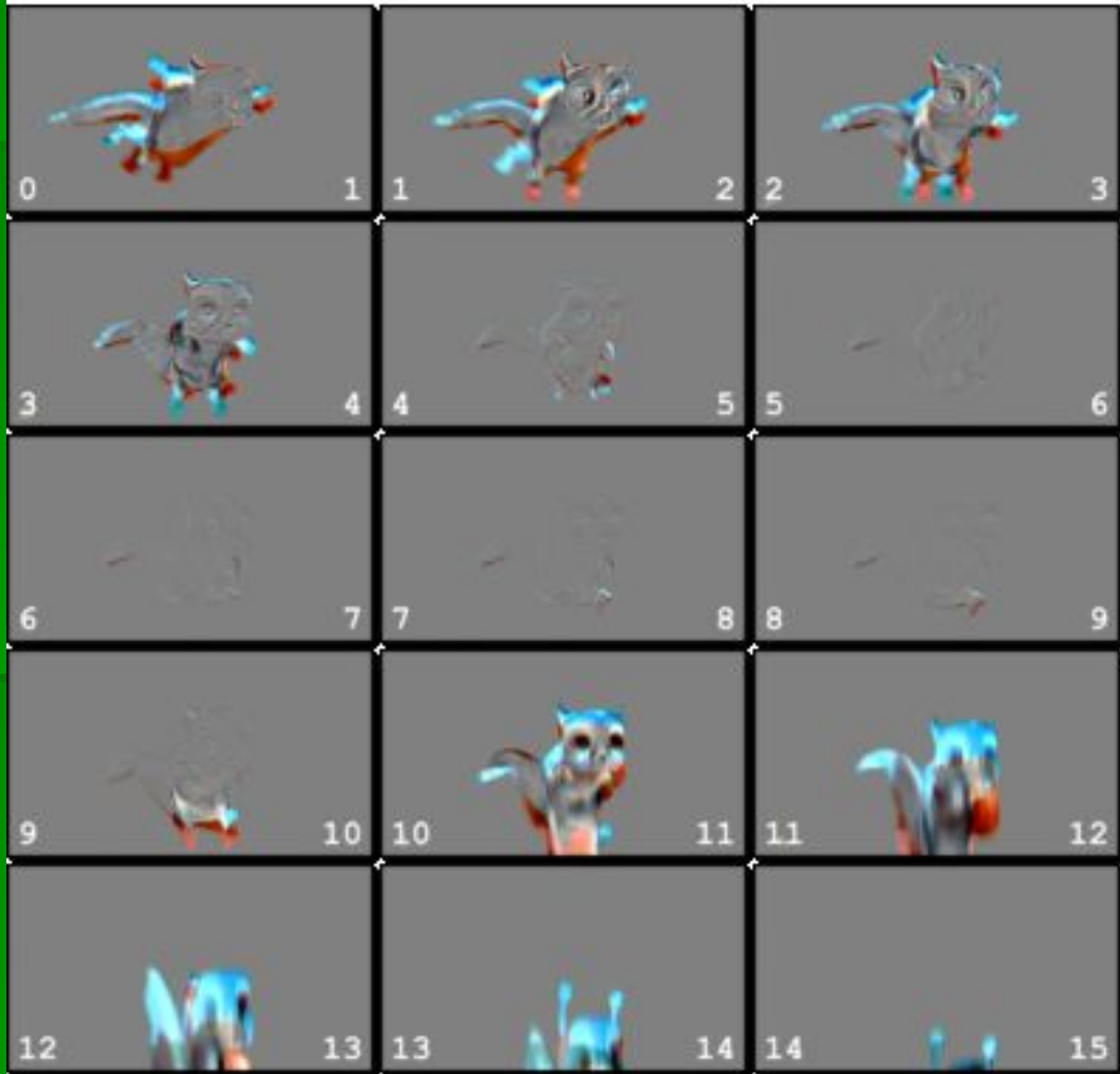


Компенсация движения

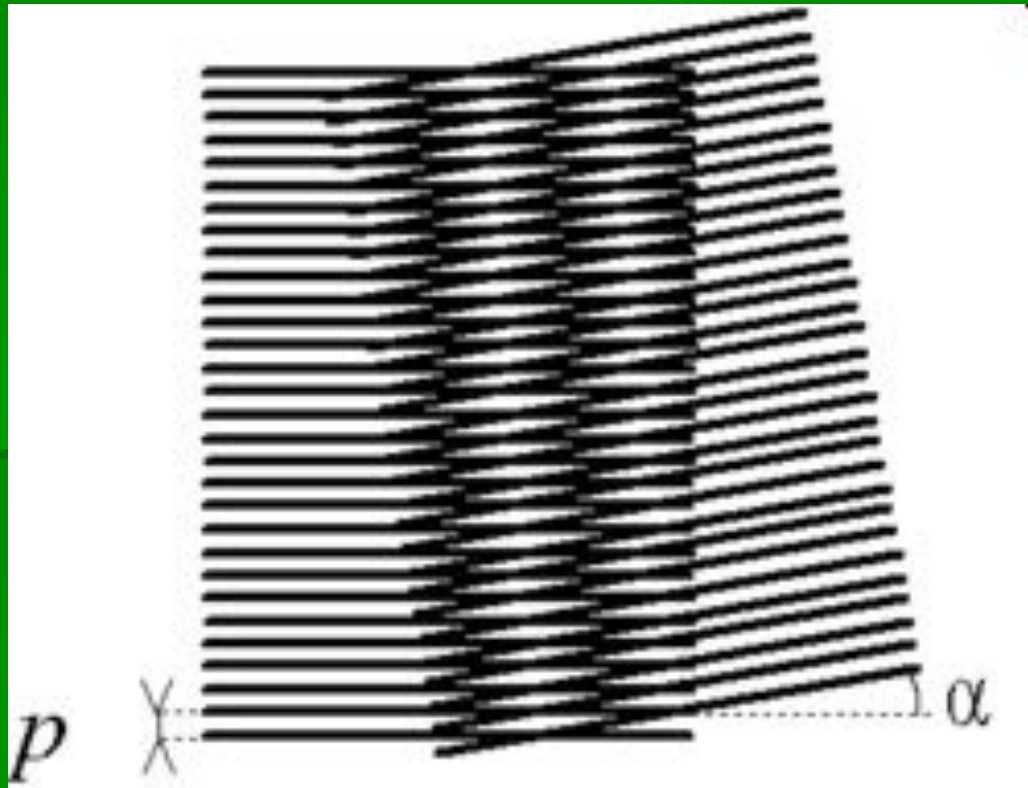


Компенсация движения

Миниатюры кадров без применения алгоритма компенсации движения



Myap



ХИТИНГ

abcfгор АО *abcfгор*

abcfгор АО *abcfгор*

abcfгор

abcfгор