

Биологический институт

Томский государственный университет

Информационные технологии

Лекция 2

Кодирование и сжатие графических,
аудио- и видеоданных

**Дмитрий Владимирович
Курбатский**
научный сотрудник НИЛБиЭМ БИ
ТГУ, магистр биологии

- Зоологический музей (к. 123)
- Кафедра зоологии позвоночных (к. 126)

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

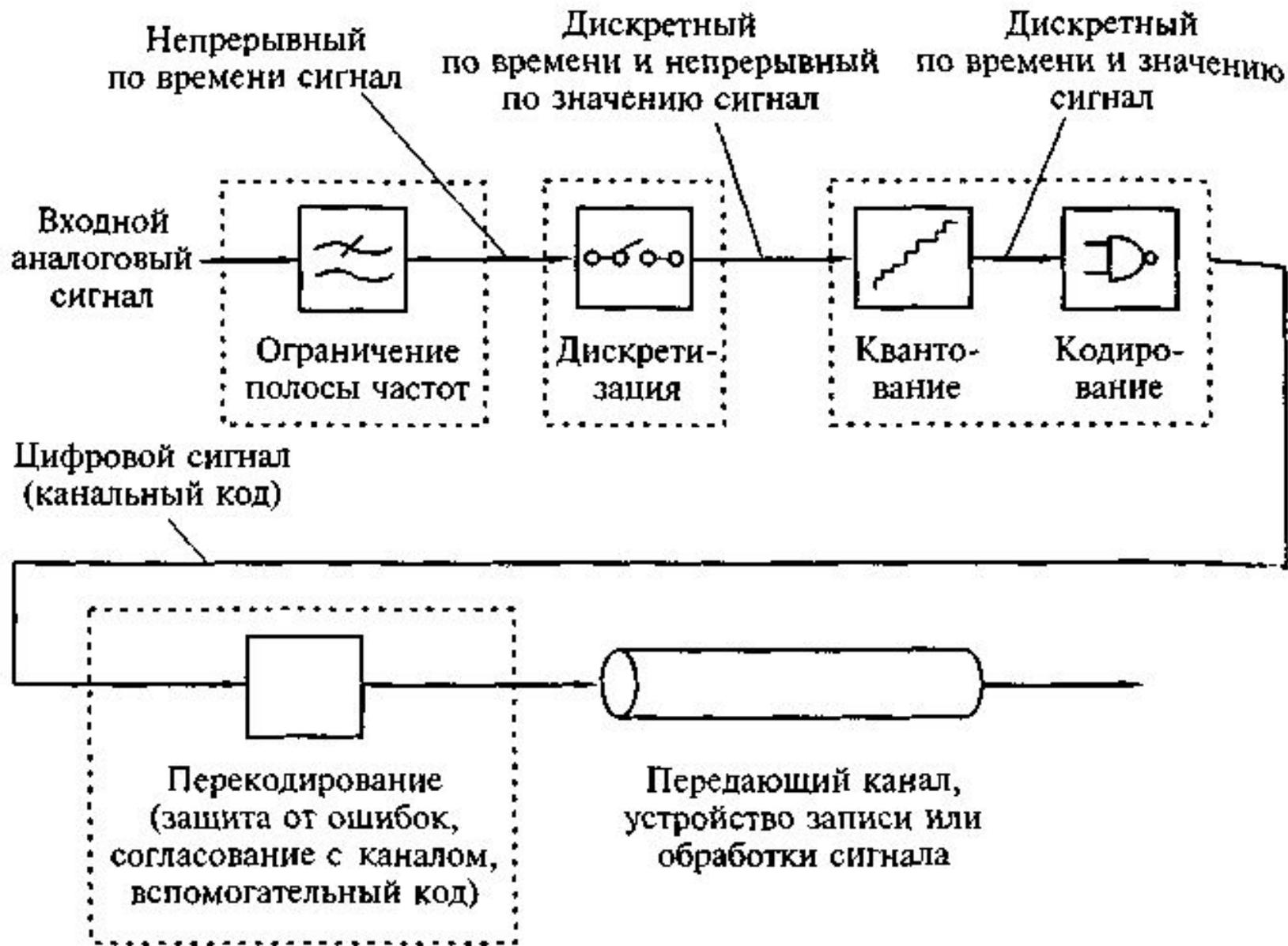
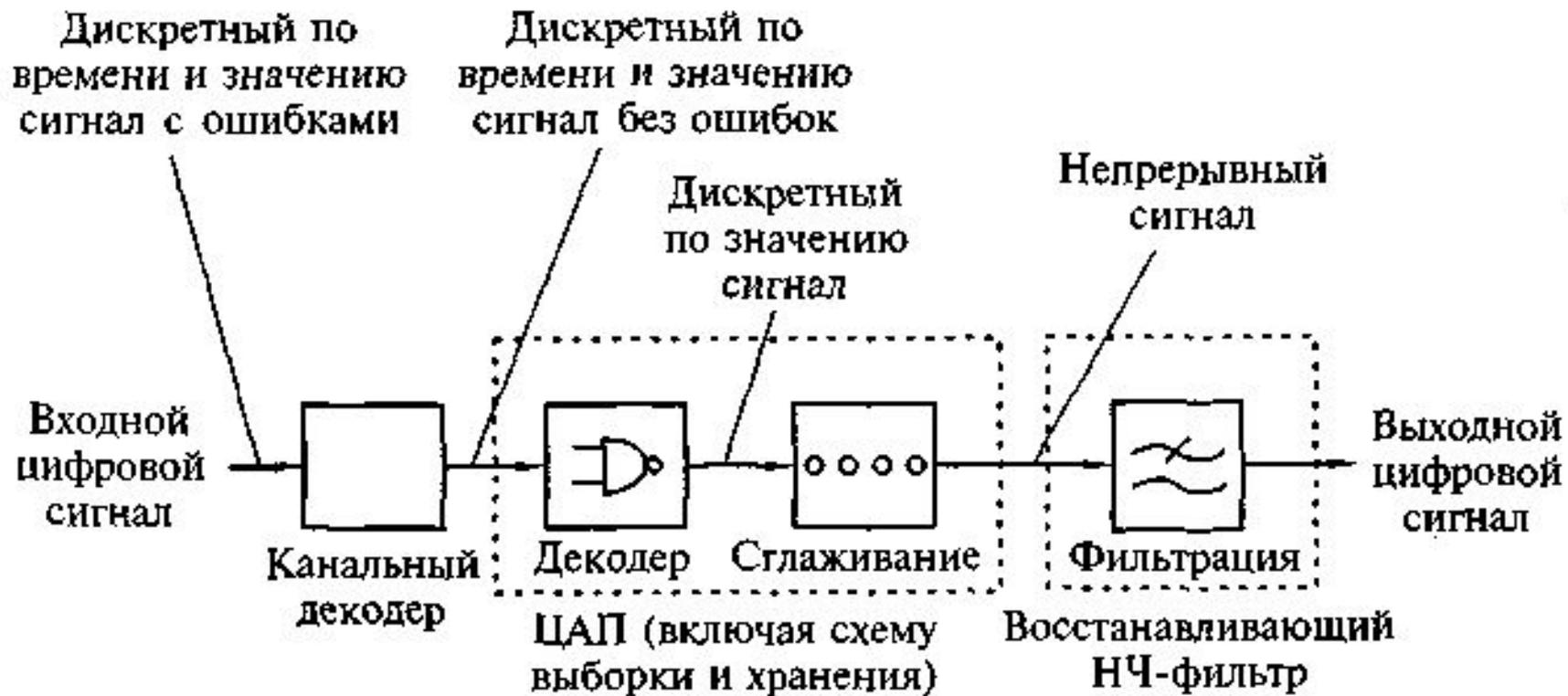
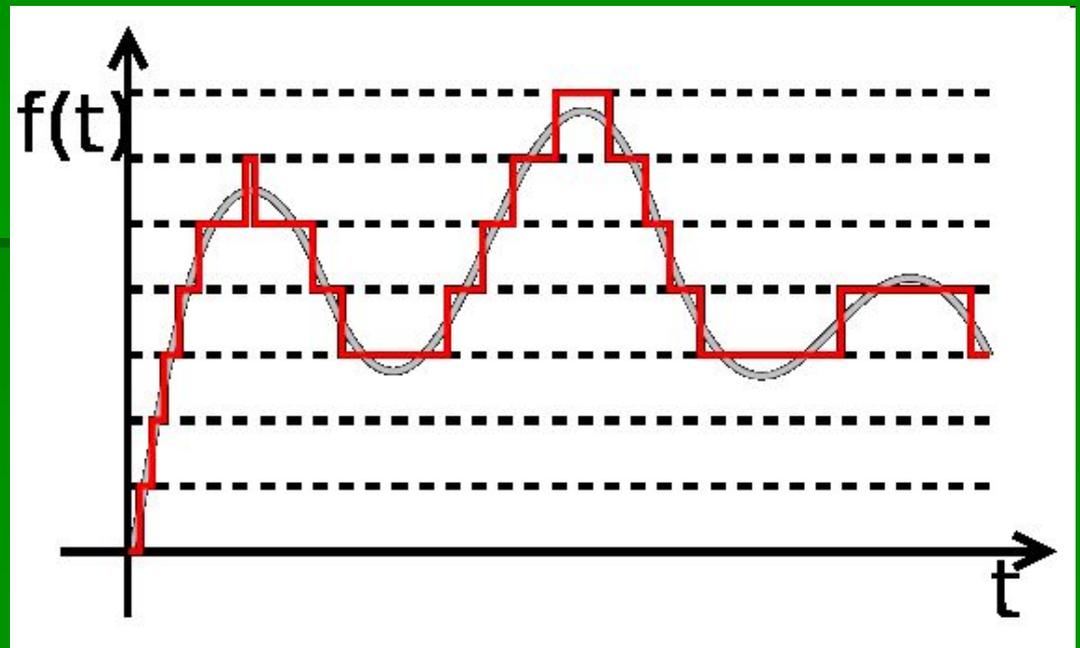


Схема ЦАП



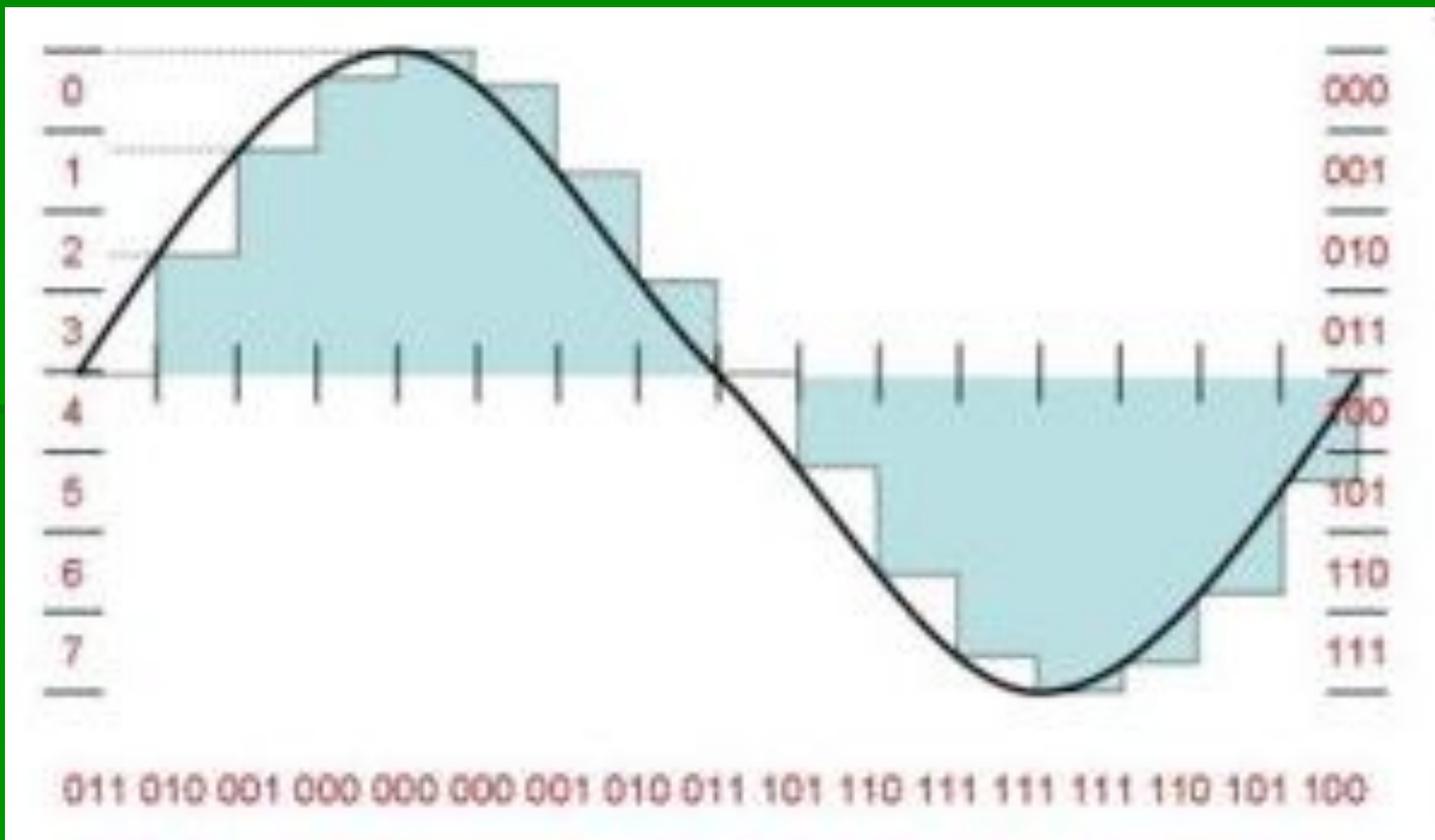
Квантование (quantization)

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

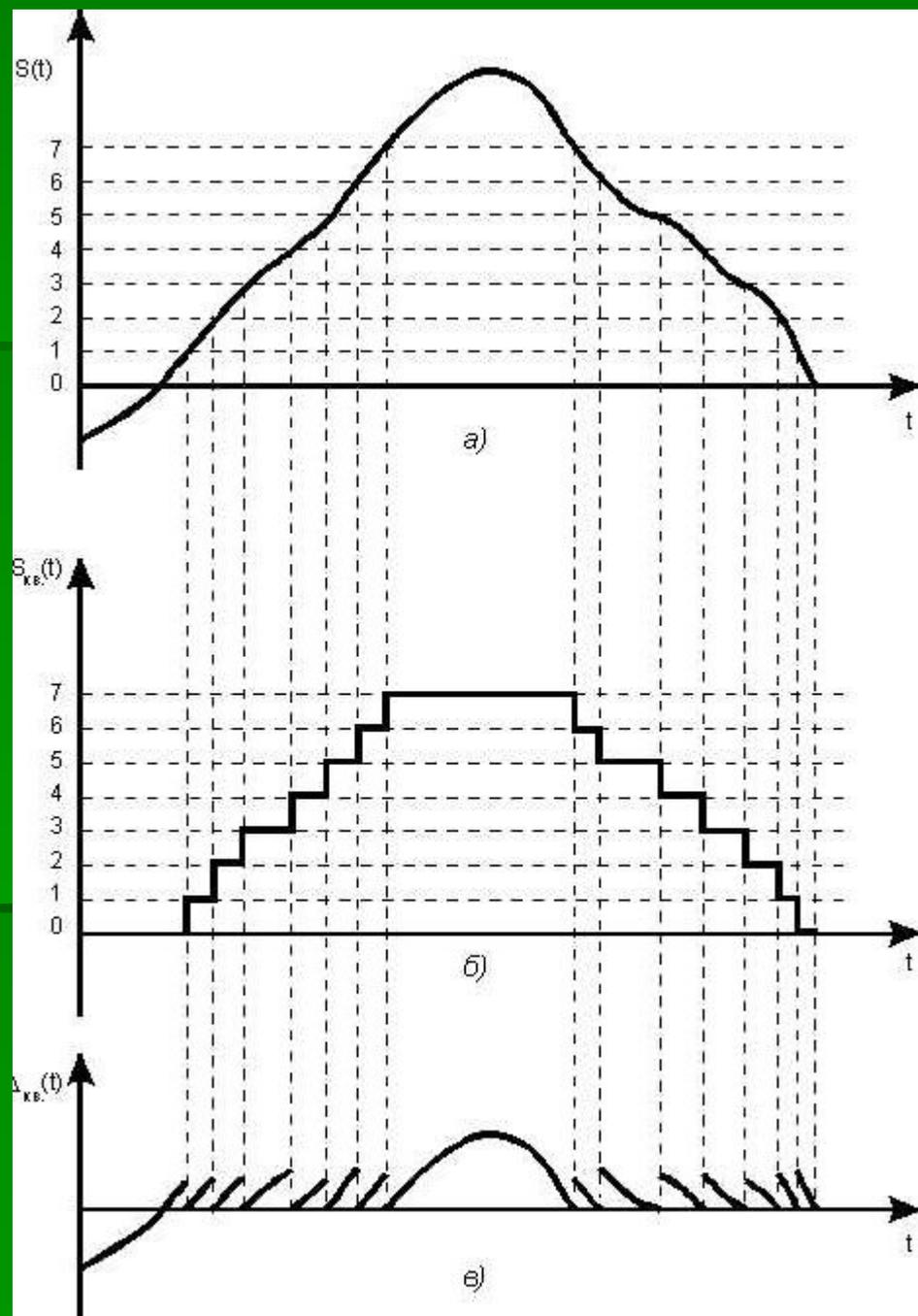


Разрядность квантования

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

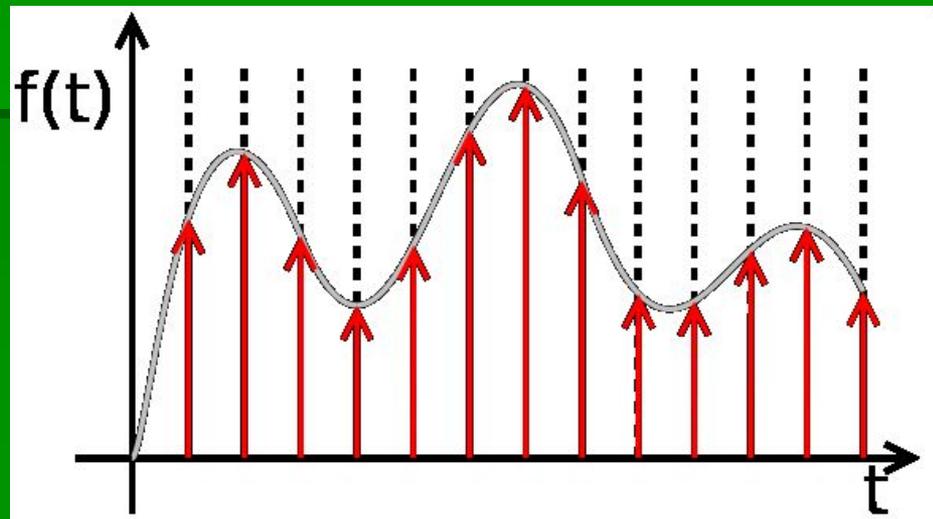


Шум квантования



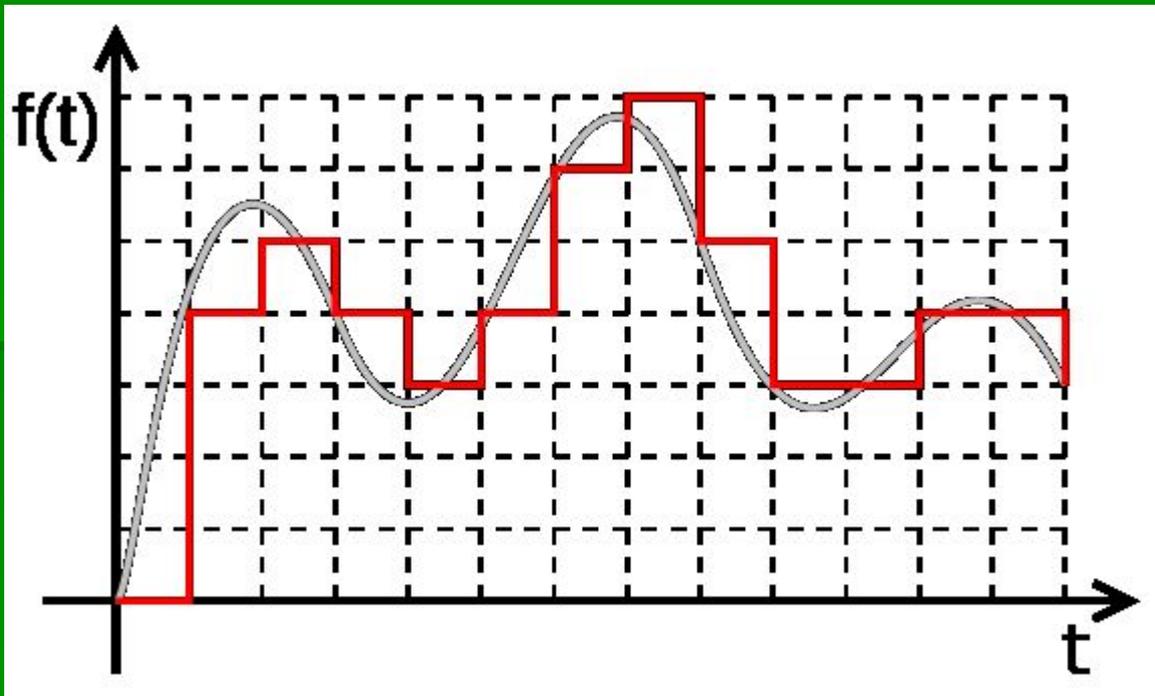
Дискретизация

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

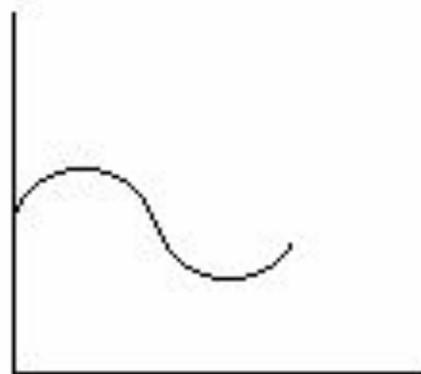


Цифровой сигнал

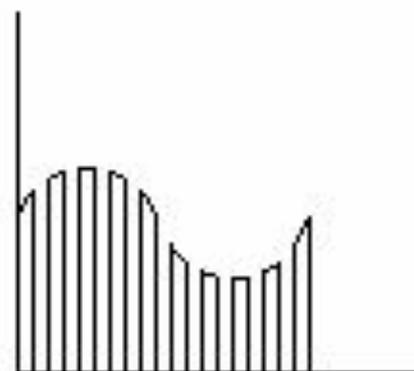
- = квантование + дискретизация



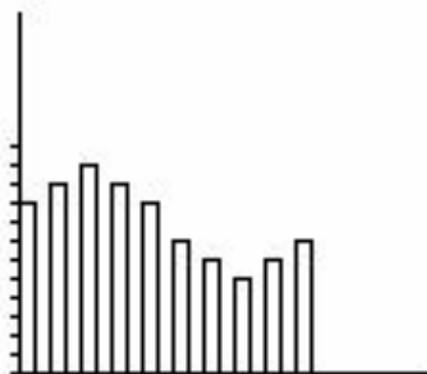
Цифровой сигнал



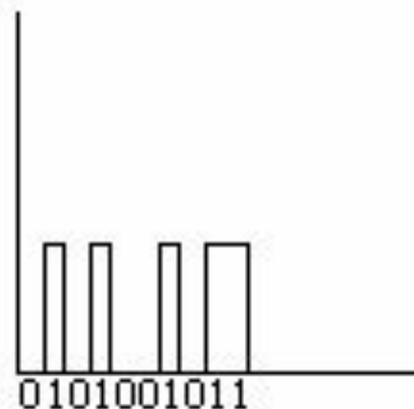
Дискретизация



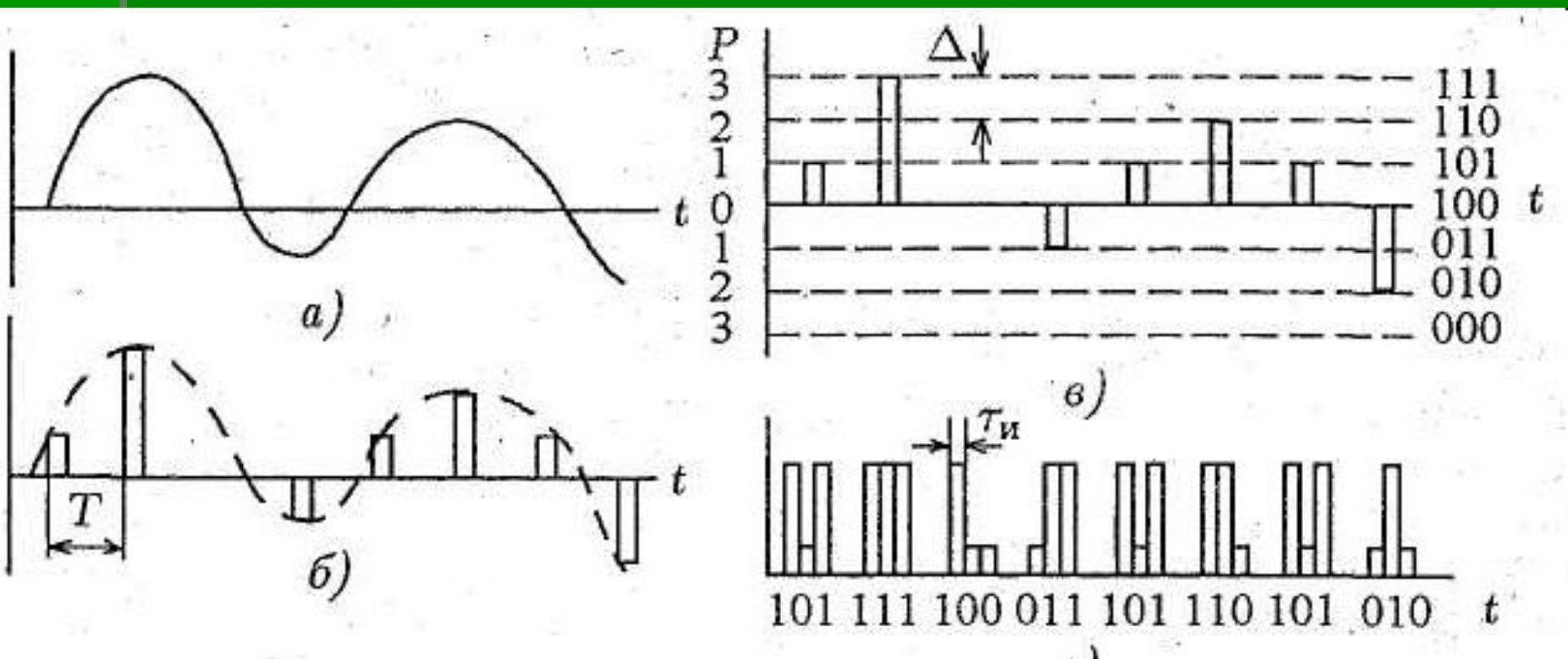
Квантование



Кодирование



Цифровой сигнал



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.

Характеристики АЦП

- Поддерживаемый диапазон частот (Hz, Гц)
- Динамический диапазон (dB, дБ)
- Разрядность (бит)
- Частота дискретизации (Hz, Гц)
 - теорема Котельникова
 - $f_d > 2 * f_{\max}$
 - частота Найквиста

Импульсно-кодовая модуляция

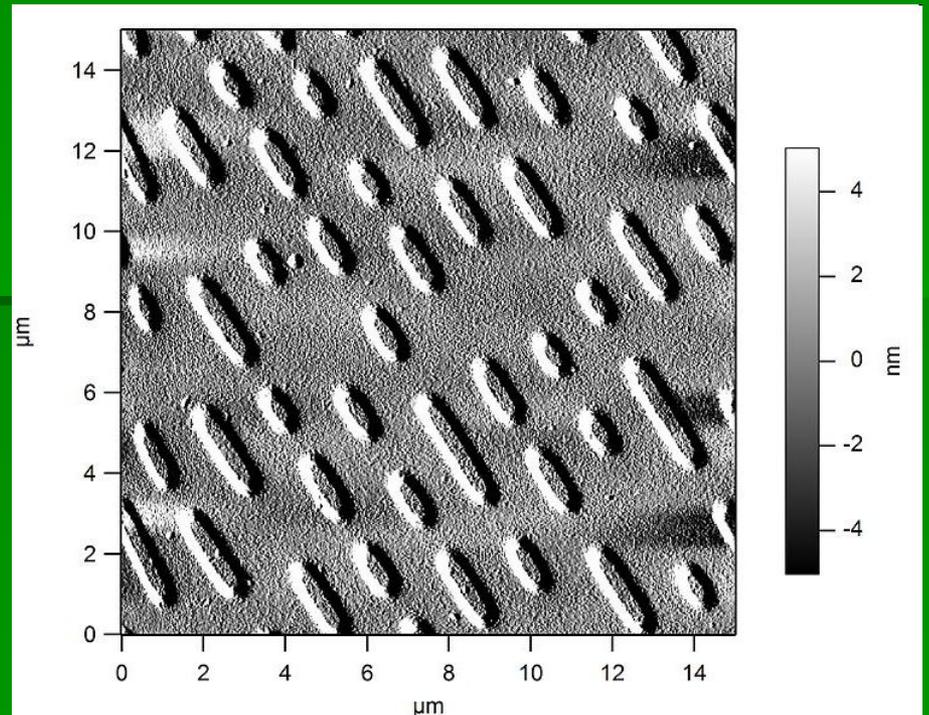
- – дискретизация сигнала во времени в совокупности с методом однородного квантования.
- *Pulse Code Modulation – PCM*
- Записываются абсолютные значения амплитуды.
- Пример:
 - 16 бит
 - 44.1 кГц
 - до 20 000 Гц (с запасом)

Дополнительные действия

- Помехоустойчивое кодирование
 - Переमेжитель
- Канальное кодирование

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

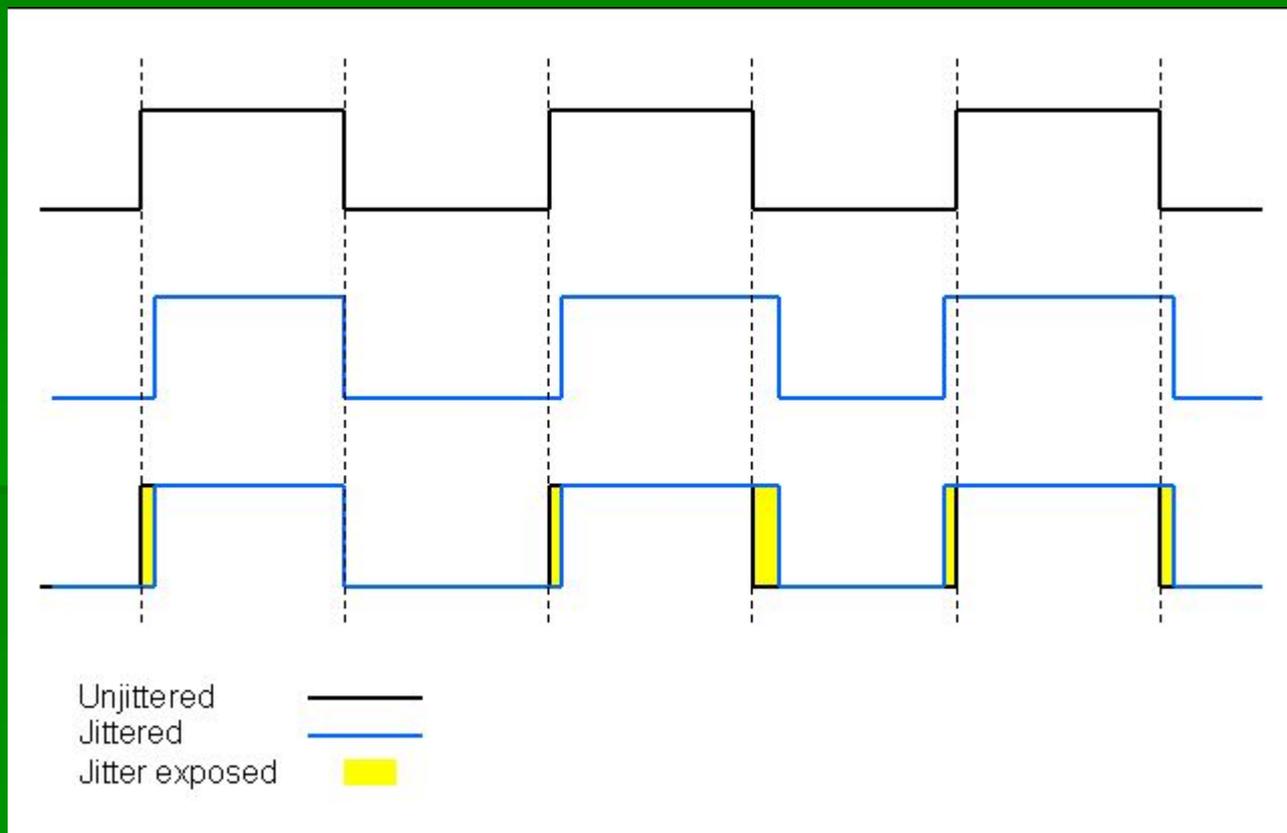
- Шаг дорожки — $1,6 \pm 0,1$ мкм
- Ширина пиков — $0,4$ мкм
- Глубина пиков — $0,12$ мкм
- Длина пиков — $0,83$ — $3,1$ мкм
- Длина волны лазера — 780 нм
- Частота дискретизации — $44,1$ кГц
- Разрядность — 16 бит (линейное квантование)
- Скорость считывания звуковой информации — $1,4112$ Мбит/с
- Общая 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

Другие варианты

- Логарифмическое квантование
- Дельта-кодирование (*Adaptive Delta PCM*)
- Дифференциальная импульсно-кодовая модуляция
- Сигма-дельта-модуляция

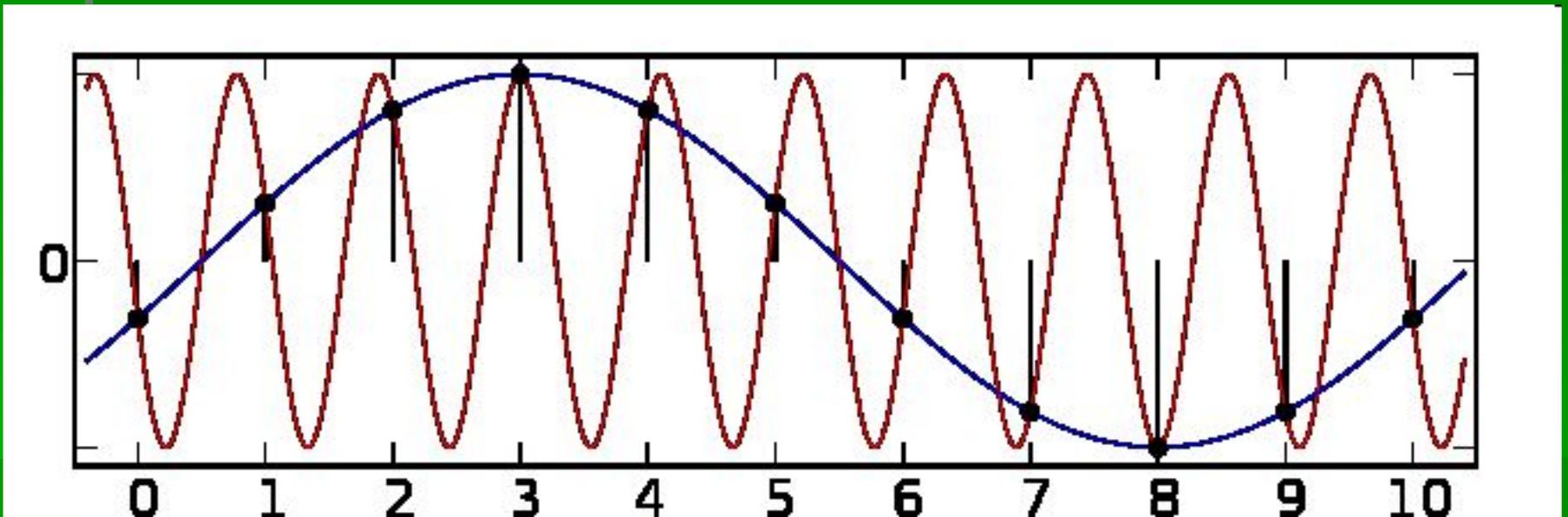
Связанные понятия

- джиттер jitter



Связанные понятия

- алиасинг aliasing



Связанные понятия

- интерполяция
- подмешивание псевдослучайного сигнала dither
- передискретизация
- отношение сигнал/шум
- полоса пропускания

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)

Сжатие данных без потерь

- *Lossless data compression*
- Информационная энтропия
- Энтропийное кодирование

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:
 - 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

Алгоритм Лемпеля — Зива — Велча

- *Lempel-Ziv-Welch, LZW*
- Применение: GIF, TIFF, PDF,

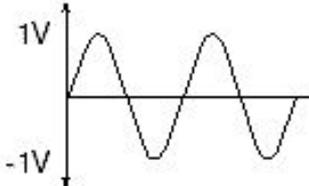
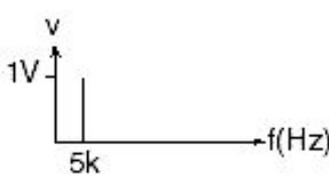
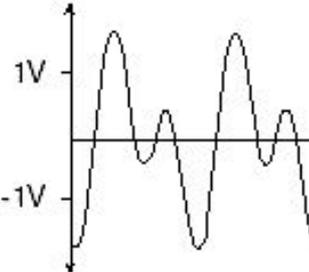
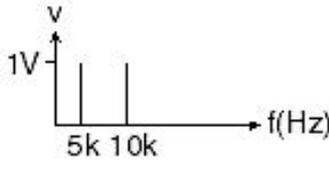
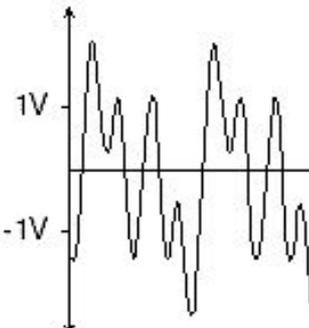
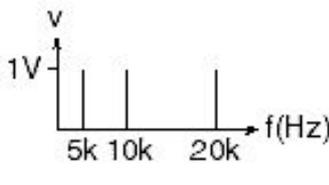
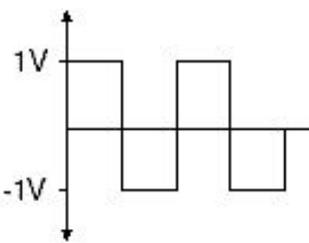
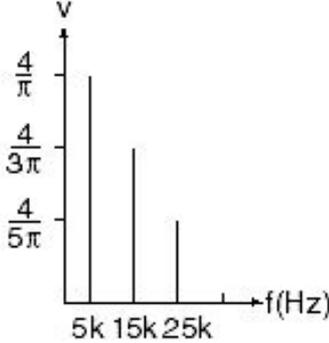
Сжатие данных с потерями

- *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		$v(t) = \frac{4}{\pi}\sin(\omega_1)t + \frac{4}{3\pi}\sin(\omega_2)t + \frac{4}{5\pi}\sin(\omega_3)t \dots$ $\omega_1 = 2\pi(5\text{kHz})$ $\omega_2 = 2\pi(15\text{kHz})$ $\omega_3 = 2\pi(25\text{kHz}) \dots$	

MP3



(a)

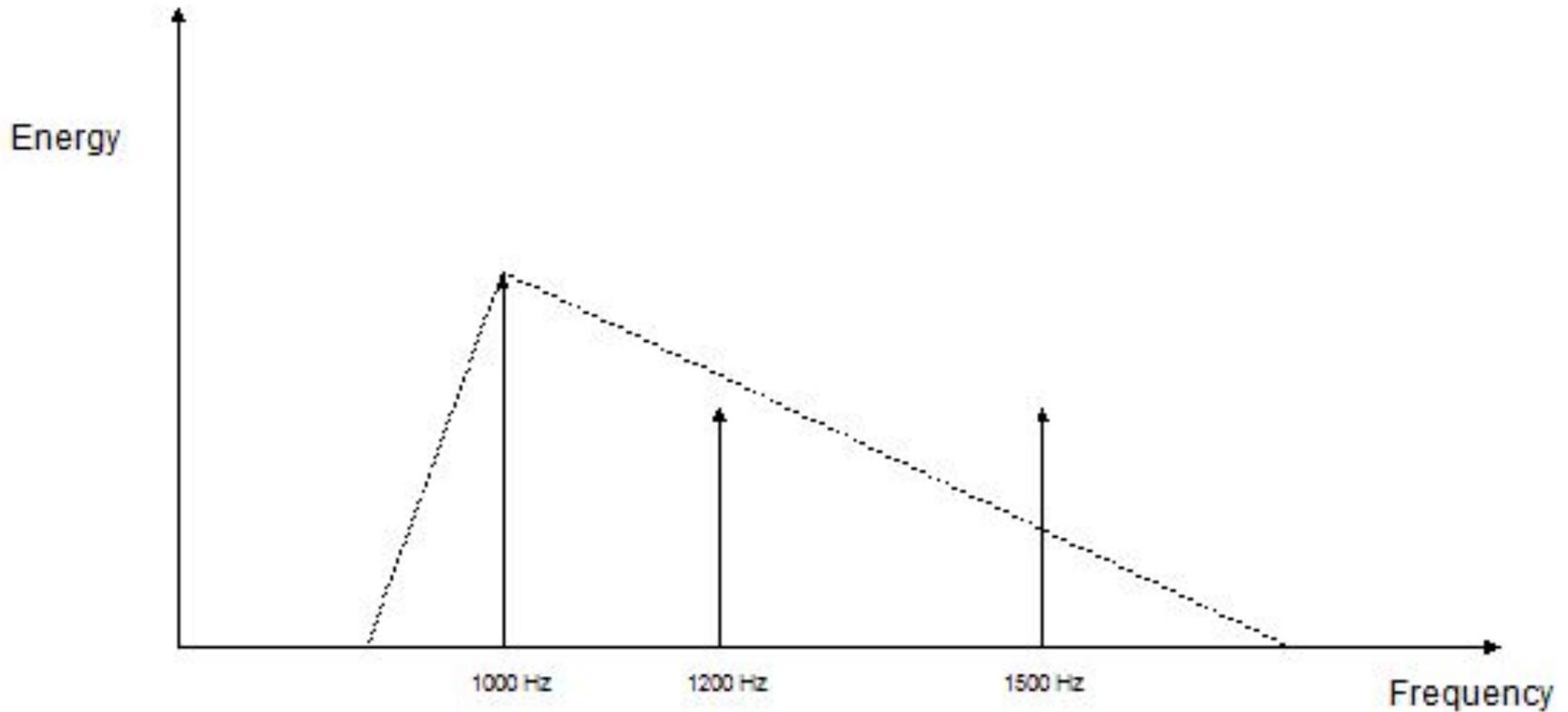


(b)

Эффект Хааса (**Haas**)

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

Эффект маскировки



Битрейт

- здесь – степень сжатия потока аудио- или видеоданных.

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

Алгоритмы

- Дискретное вейвлет-преобразование

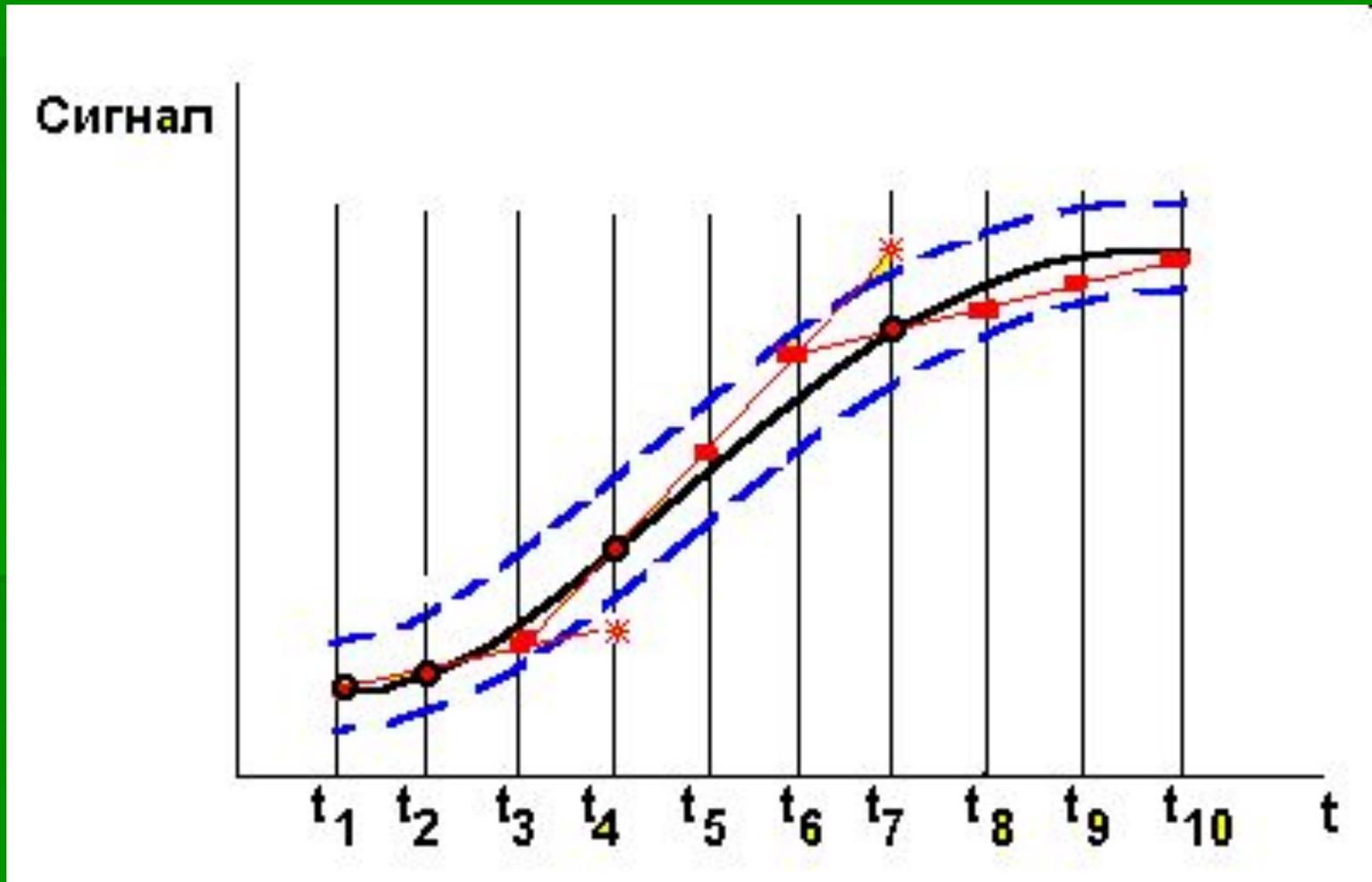
JPEG

- изображение преобразуется из цветового пространства RGB в YCbCr
- «прореживание» (*subsampling*)
- дискретное косинусное преобразование

Пример сжатия **JPEG** с разными коэффициентами



Предиктивное сжатие

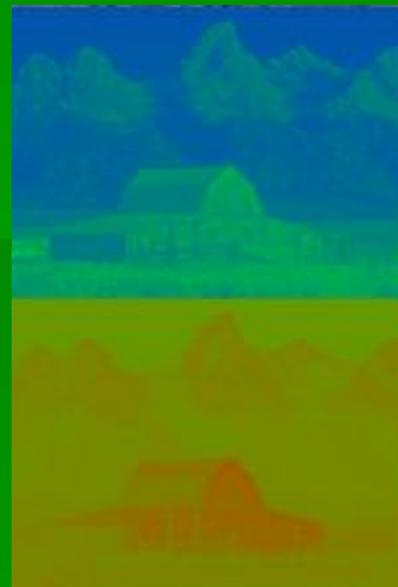


Блок 3

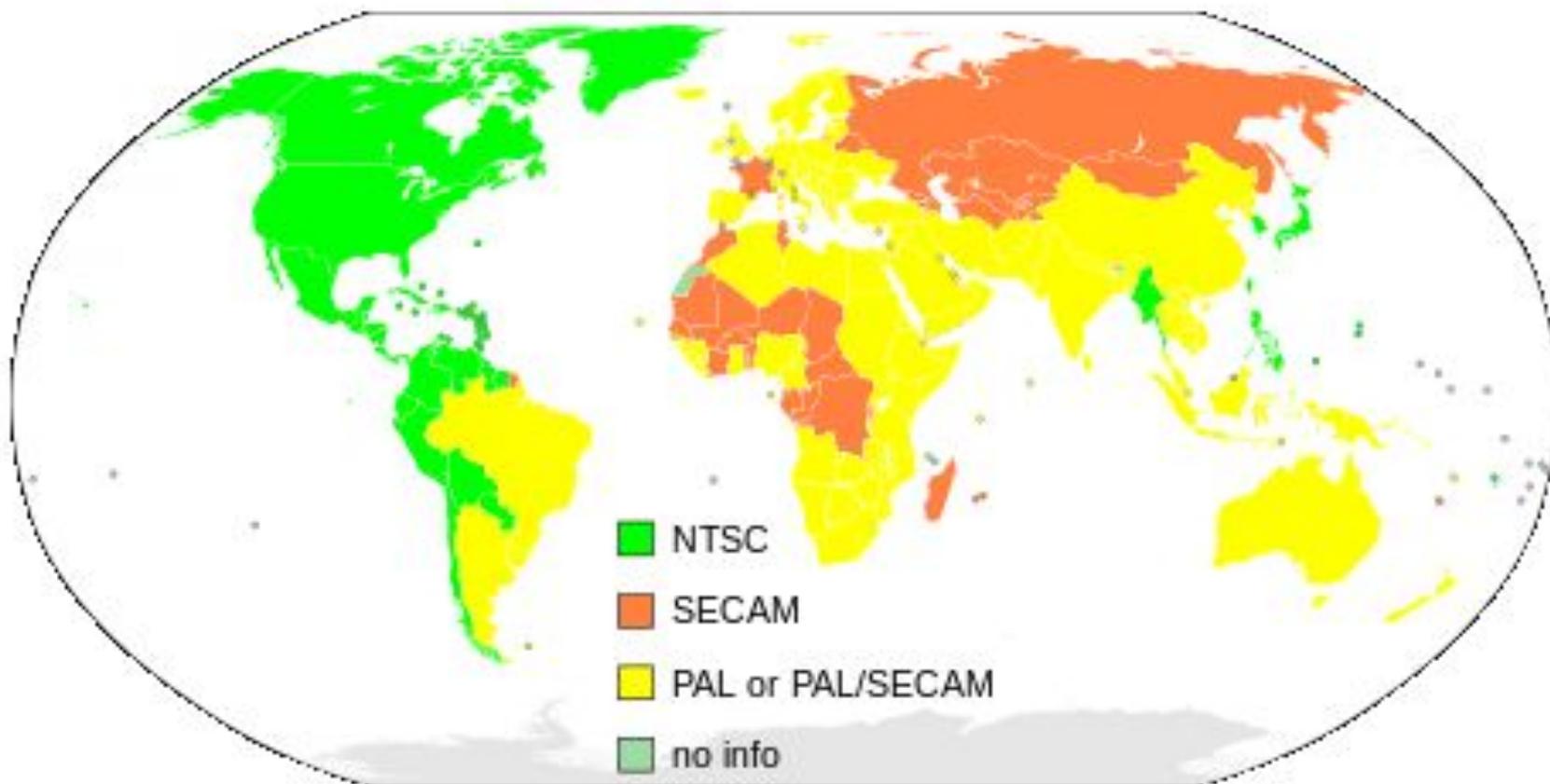
Особенности кодирования и сжатия видеоданных

YUV

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



Стандарты телевидения



Стандарты телевидения

- NTSC
 - 525 строк
 - 59,94 полей в секунду
 - в Америке и Японии
- PAL/SECAM
 - 625 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



Компенсация движения

- *Motion Compensation*

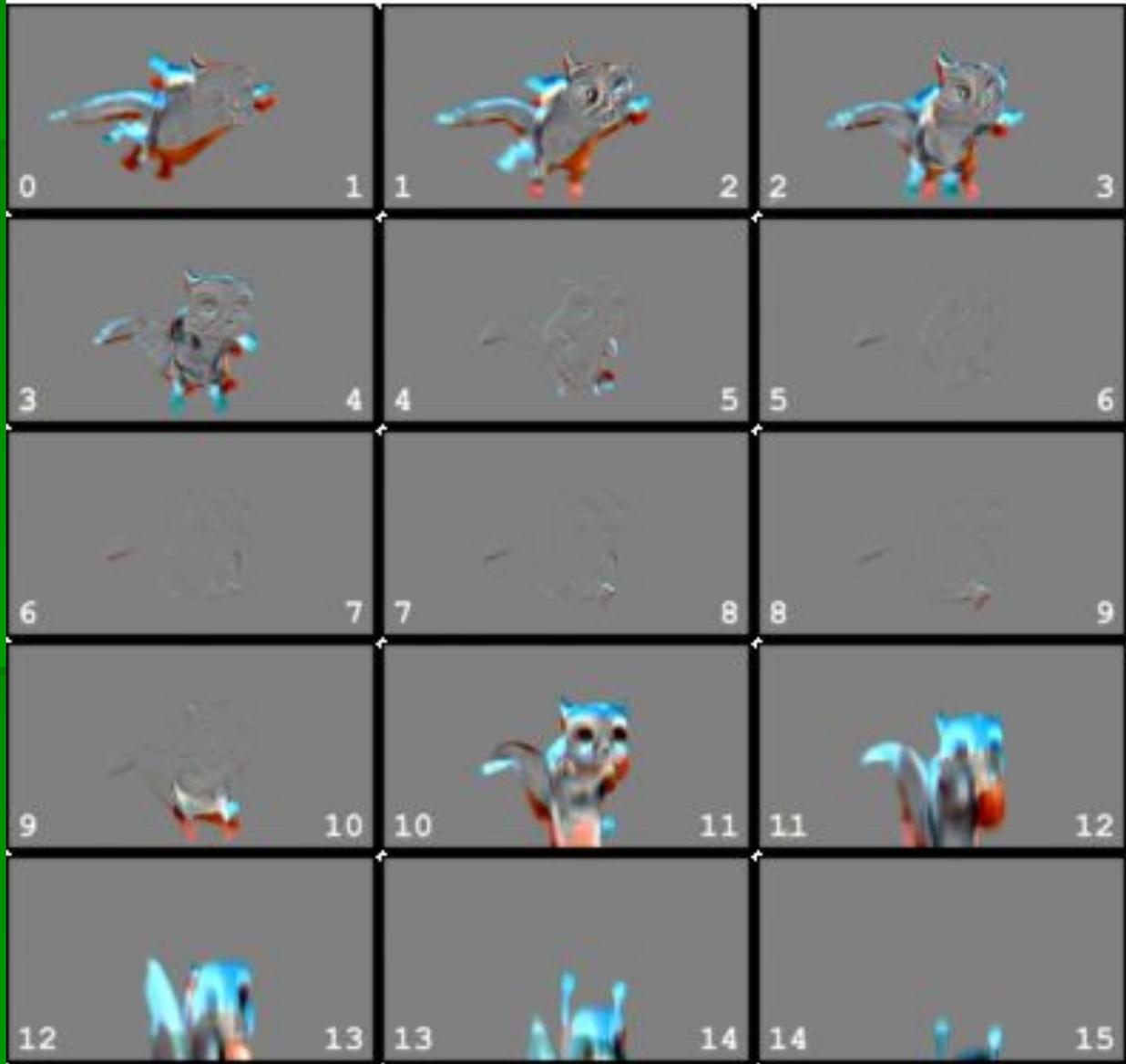
Компенсация движения

Телесная восприимчивость (Big Buck Bunny)

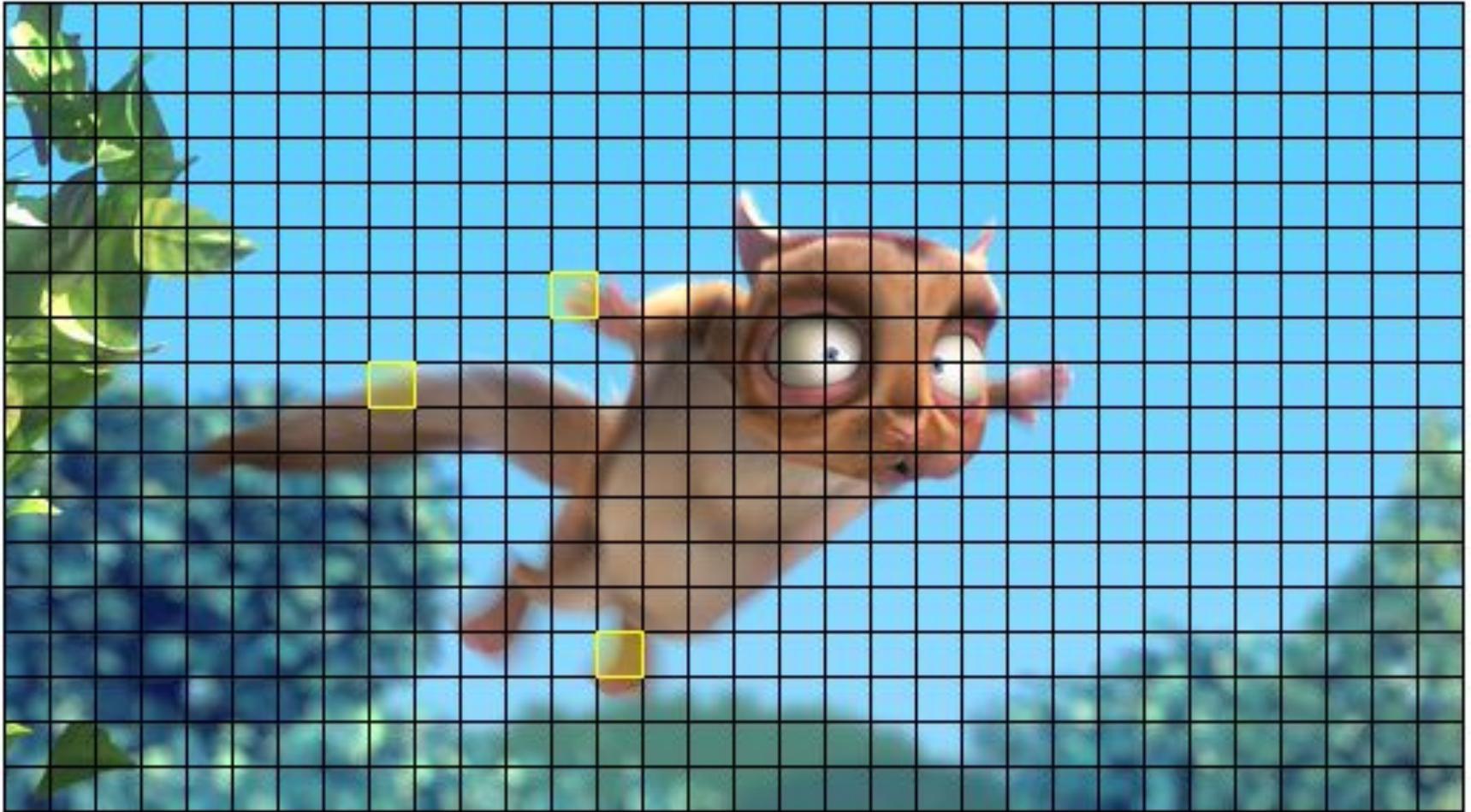


Компенсация движения

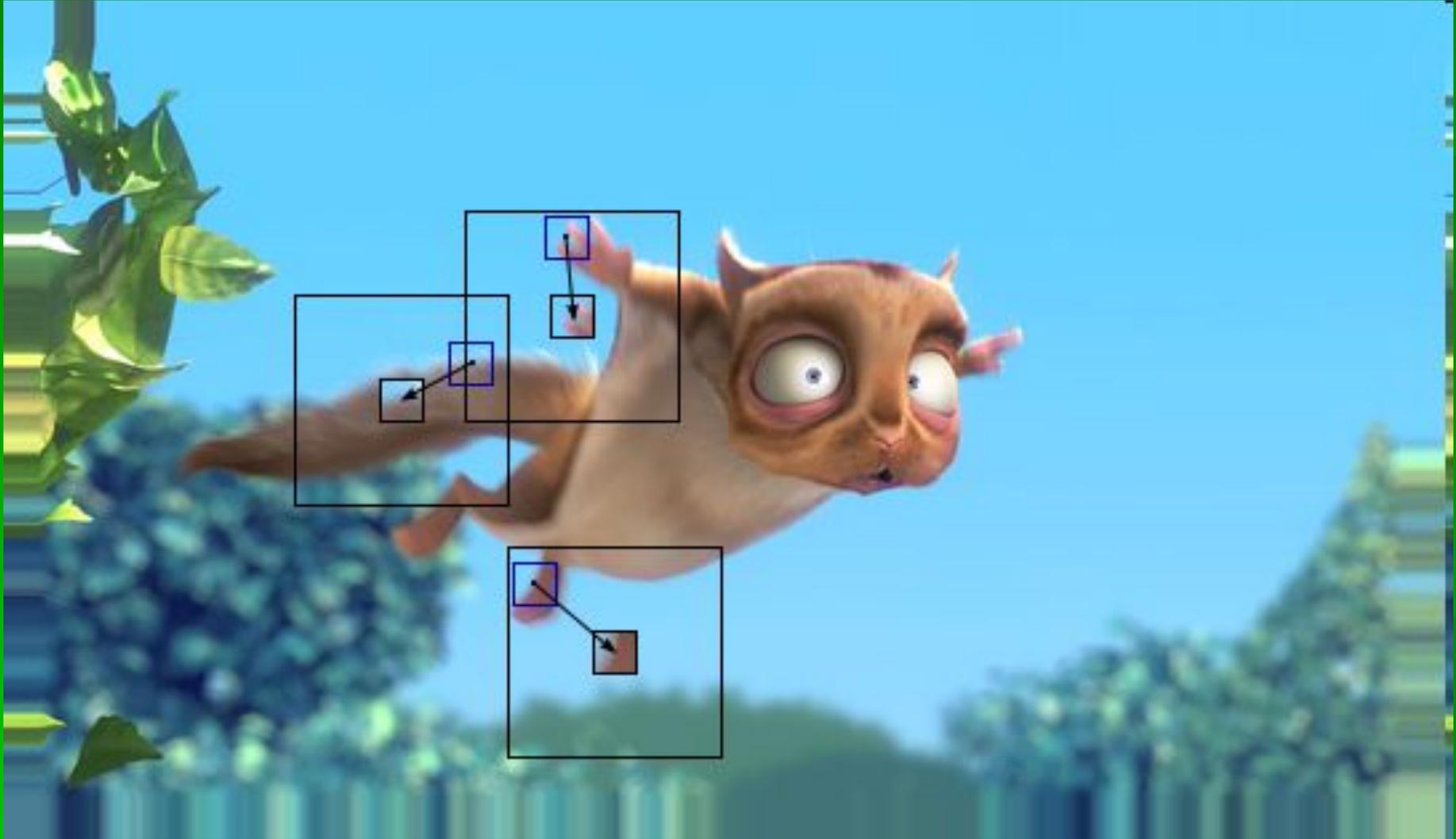
Миниатюры кадров без применения алгоритма компенсации движения



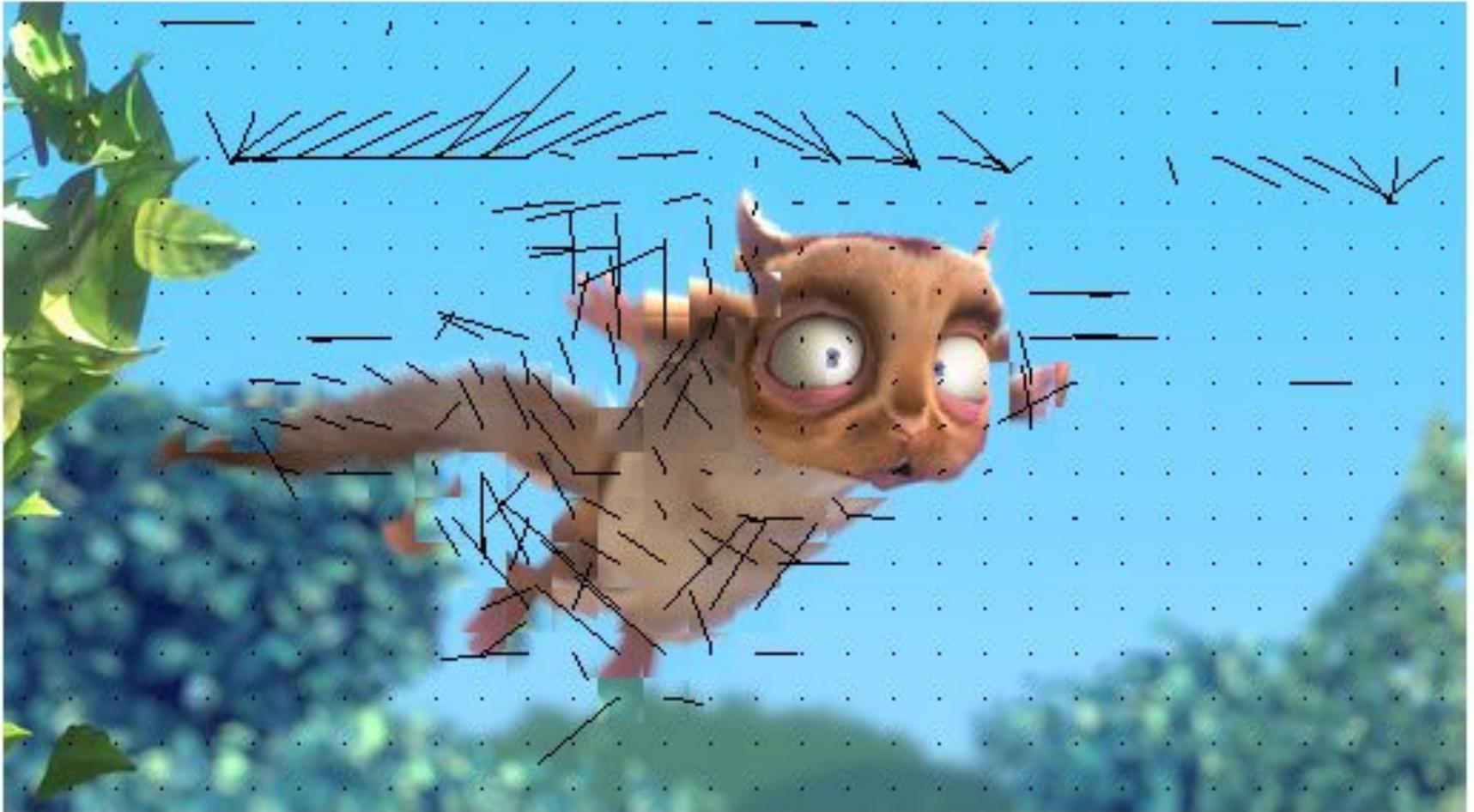
Компенсация движения



Компенсация движения



Компенсация движения

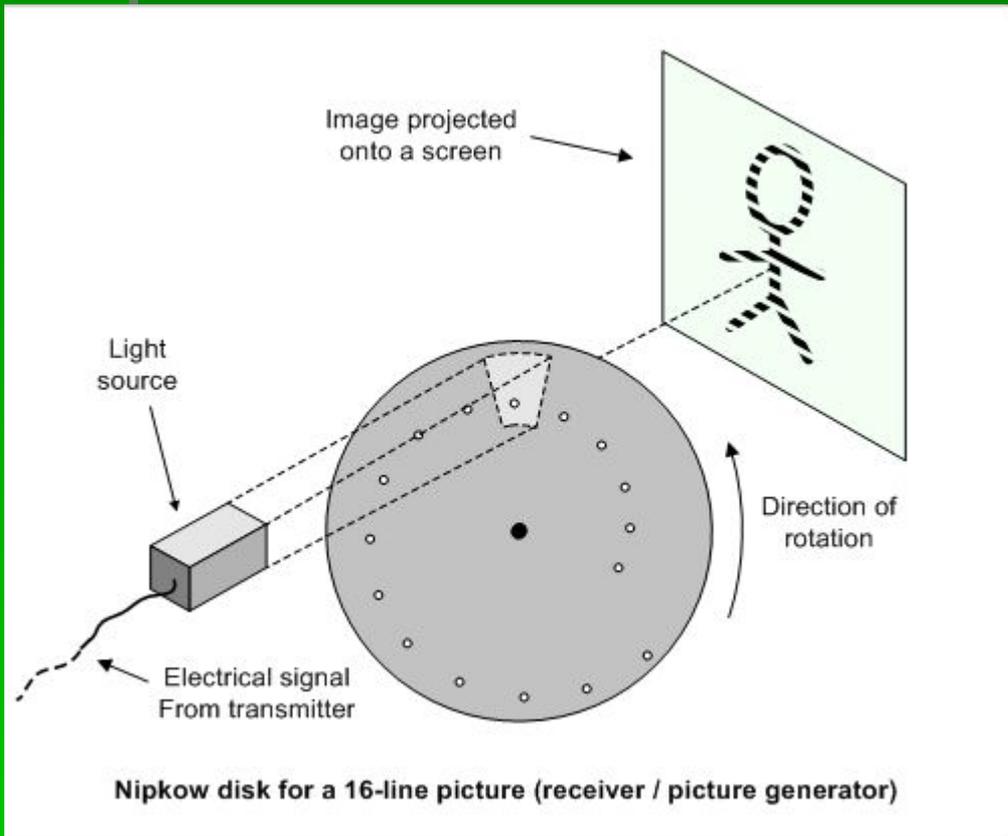


Компенсация движения



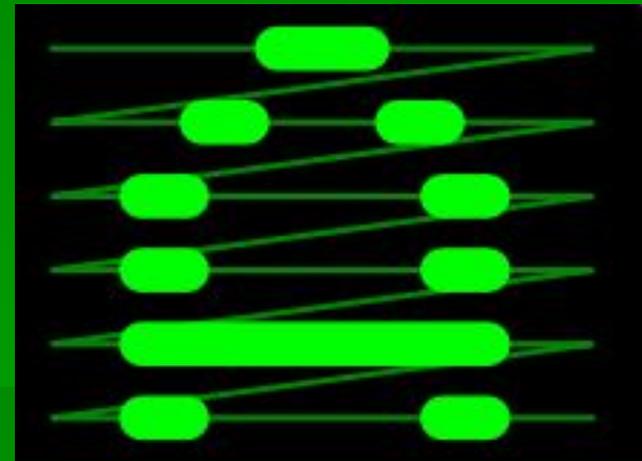
Механическая развёртка

- Диск Нипкова



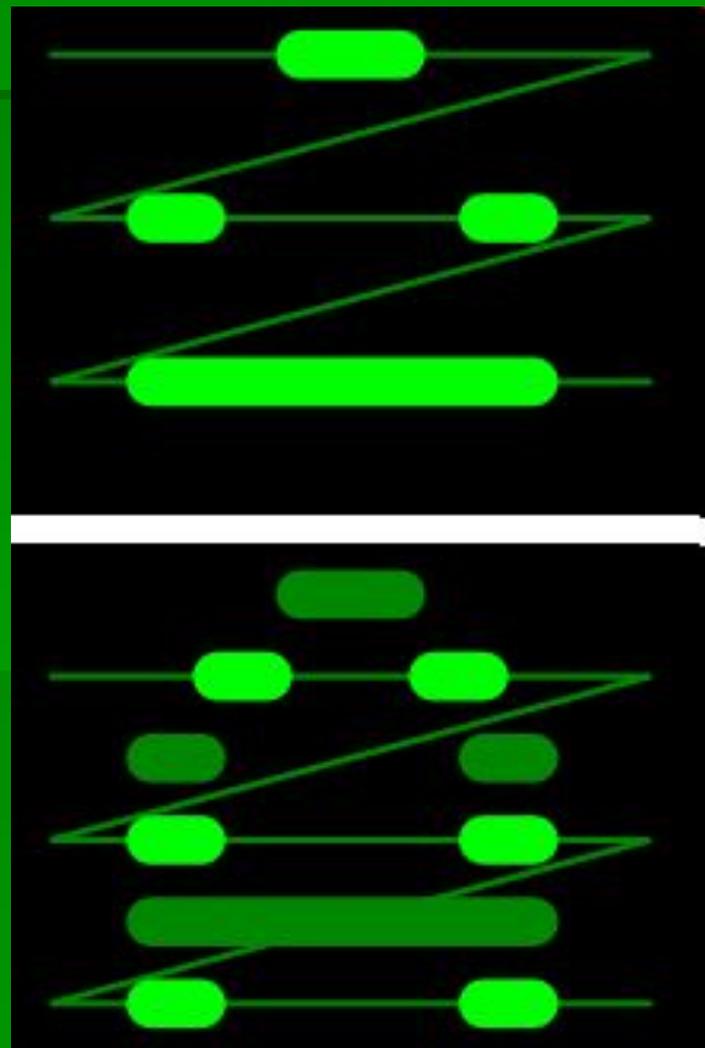
Прогрессивная развёртка

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



Деинтерлейсинг

- *Deinterlacing*
- — устранение
чересстрочности



Варианты устранения



а) Чересстрочная развёртка

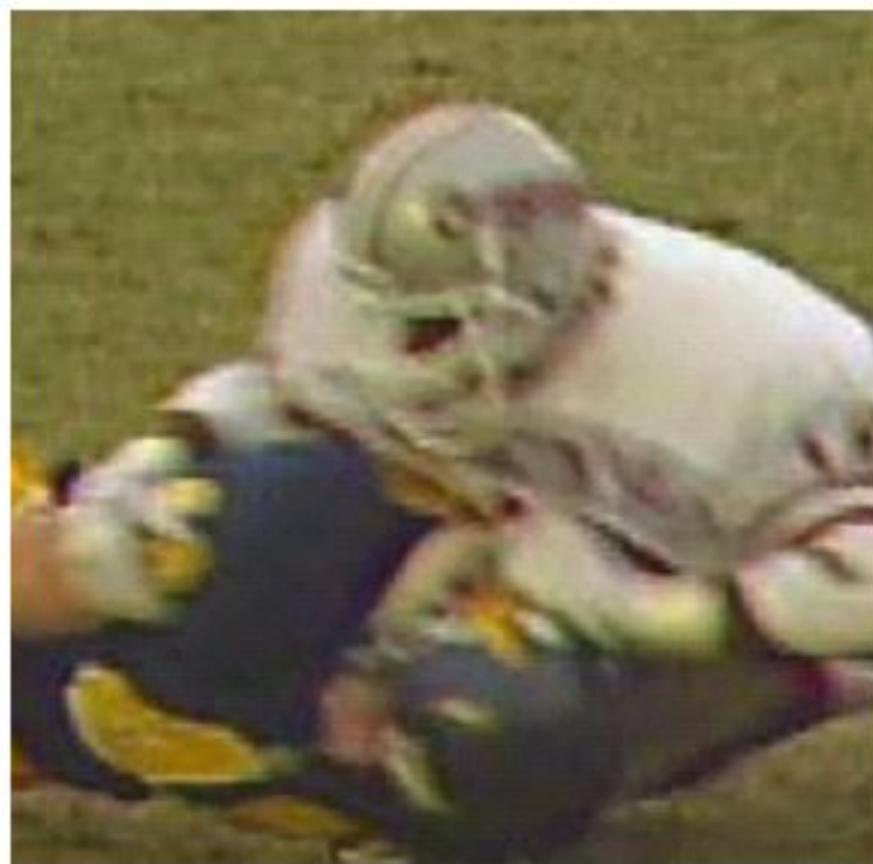


б) Сглаживание

Варианты устранения



в) Интерполяция



г) Адаптивный метод

Рис. 7. Применение деинтерлейсинга