«Введение в компьютерную графику»
лекция 28.10.2010

MSU GRAPHICS & MEDIA LAB

Банки фильтров, шумоподавление

Алексей Лукин
lukin@graphics.cs.msu.ru

План

- Банки фильтров, основанные на STFT
- Психоакустическая компрессия звука
  - Слуховая маскировка
  - Устройство алгоритма mp3
- Подавление стационарных шумов
  - Метод спектрального вычитания
  - STFT как банк фильтров, полосовые гейты
  - Проблема «музыкального шума»
  - Частотно-временное разрешение
- Вейвлеты
- Подавление шума на изображениях

MSU GRAPHICS & MEDIA LAB

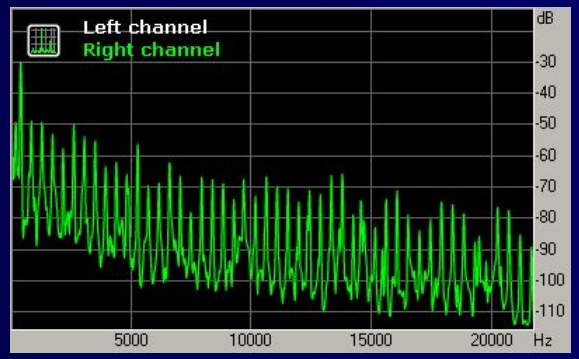
Часть 1

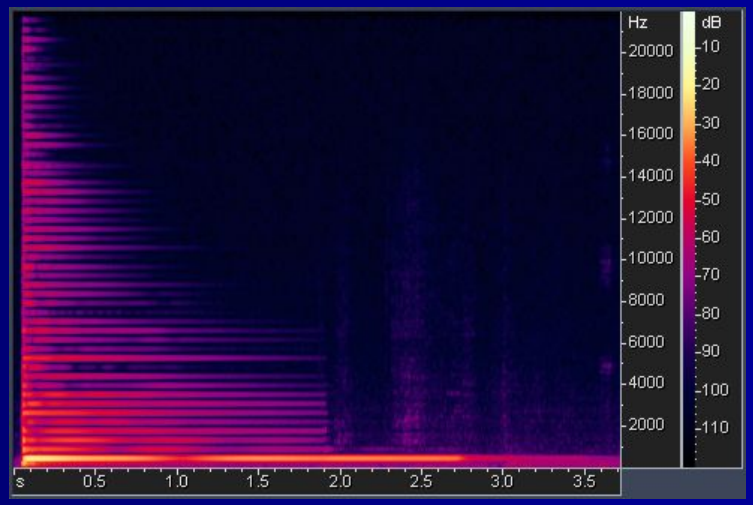
Банки фильтров и их применения

Банки фильтров, основанные на STFT

- Спектрограмма
  - график зависимости амплитуды от частоты и от времени, показывает изменение спектра во времени
  - отображается модуль Short Time Fourier Transform (STFT)

$$STFT[n,\omega] = \sum_{m=-\infty}^{+\infty} x[n+m] \cdot w[m] \cdot e^{-i\omega m}$$

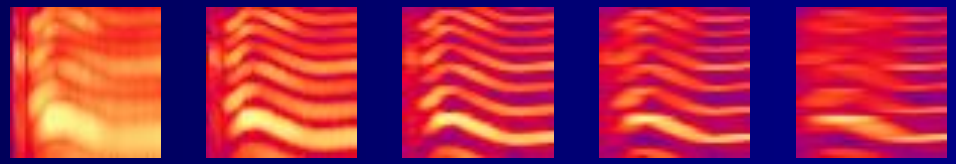




Банки фильтров, основанные на STFT

- Частотно-временное разрешение
  - Способность различать детали по частоте и по времени, «размытость» спектрограммы
  - Для STFT определяется длиной весового окна (а также, отчасти, размером и шагом DFT по времени)
  - Соотношение неопределенностей: разрешение по частоте обратно пропорционально разрешению по времени

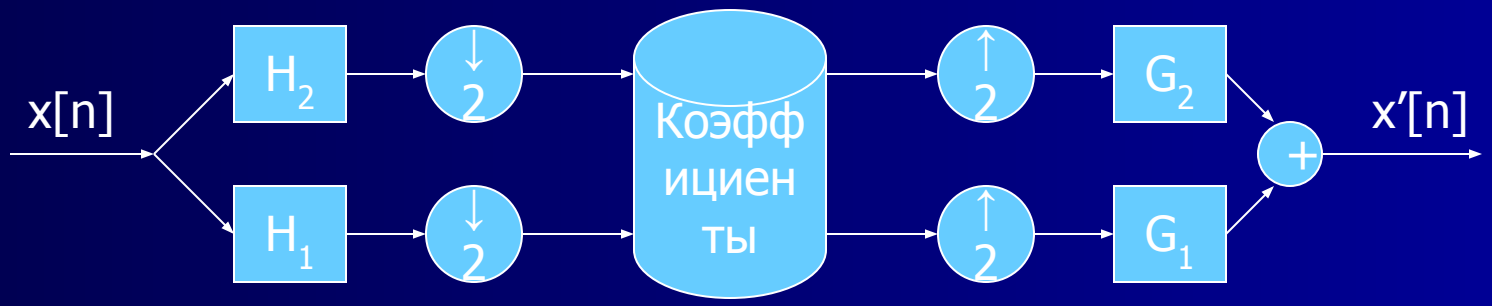
$$\Delta f \cdot \Delta t = const$$



6 ms 12 ms 24 ms 48 ms 96 ms размер окна

Банки фильтров

- Банки фильтров – преобразования, разбивающие сигнал на несколько частотных полос в возможностью обратного синтеза
- Пример: дискретное вейвлет-преобразование



Декомпозиция Реконструкция (синтез)

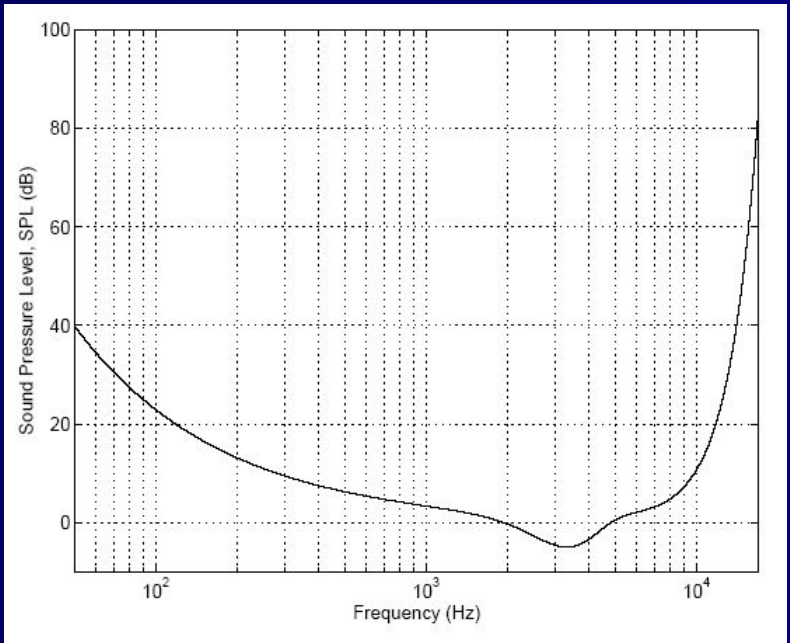
- Возможные свойства БФ: точное восстановление, избыточность

Банки фильтров

- Применения:
  - Раздельная обработка сигнала в разных частотных полосах
  - Компрессия сигналов с независимым квантованием в разных частотных полосах
- Пример банка фильтров, основанного на STFT
  - Декомпозиция: STFT с окном Хана (Hann), и с перекрытием между окнами 75%
  - Синтез: обратное DFT от каждого блока, применение весовых окон Хана и сложение окон с наложением (OLA)
  - Свойства:
    - Точное восстановление
    - Наличие избыточности

Слуховая маскировка

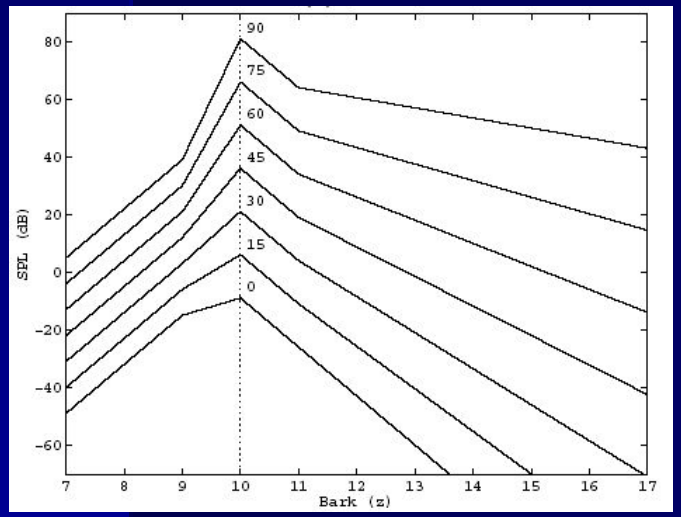
- Абсолютный порог слышимости

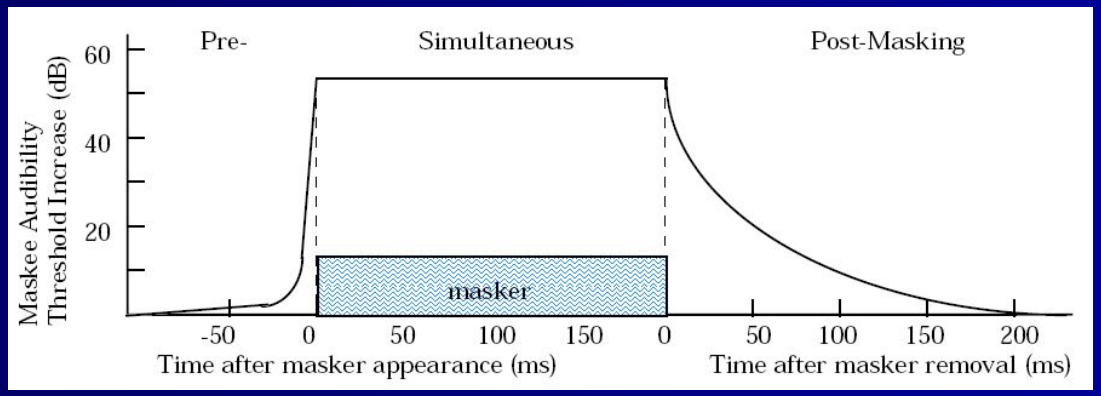


- Как соотнести уровни в звуковом файле с абсолютными уровнями звука?

Слуховая маскировка

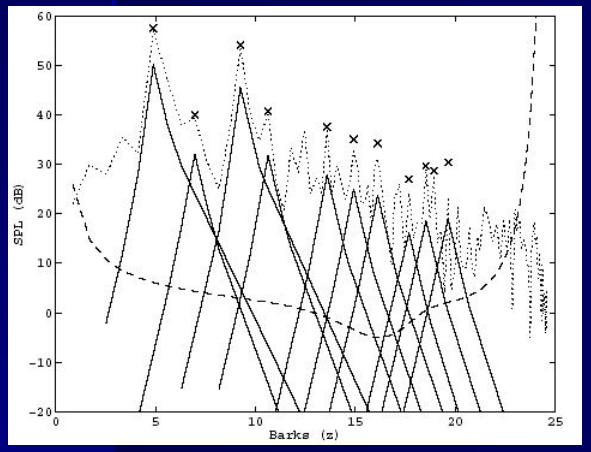
- Сильные звуки *(masker)* маскируют более слабые *(maskee)*
  - Одновременная маскировка
  - Временн**а**я маскировка (прямая и обратная)

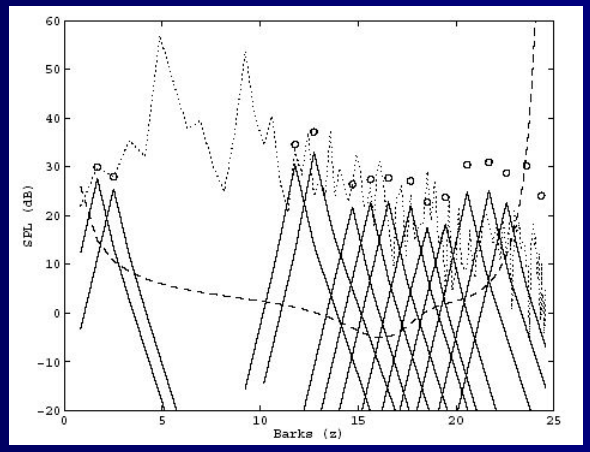


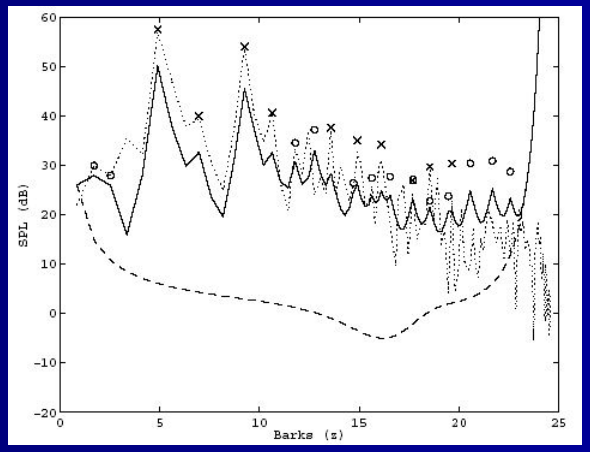


Слуховая маскировка

- Маскировка тонами, шумами и общий порог маскировки







- Шаг квантования выбирается пропорциональным порогу маскировки

Алгоритм mp3

- Кодирование аудиоданных с потерями

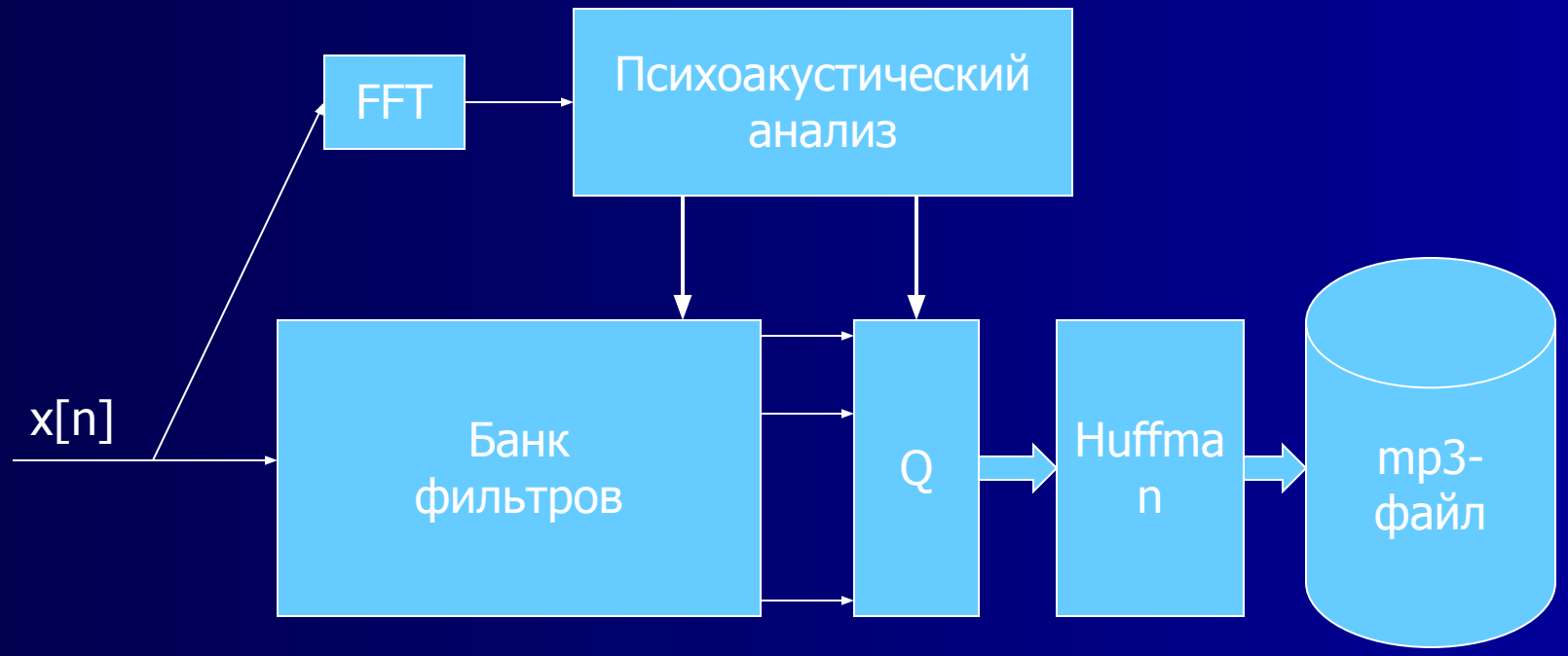


Схема кодера mp3

Пред-эхо

- Pre-echo (pre-ringing)

Размытие ошибки квантования по времени на всю длину окна

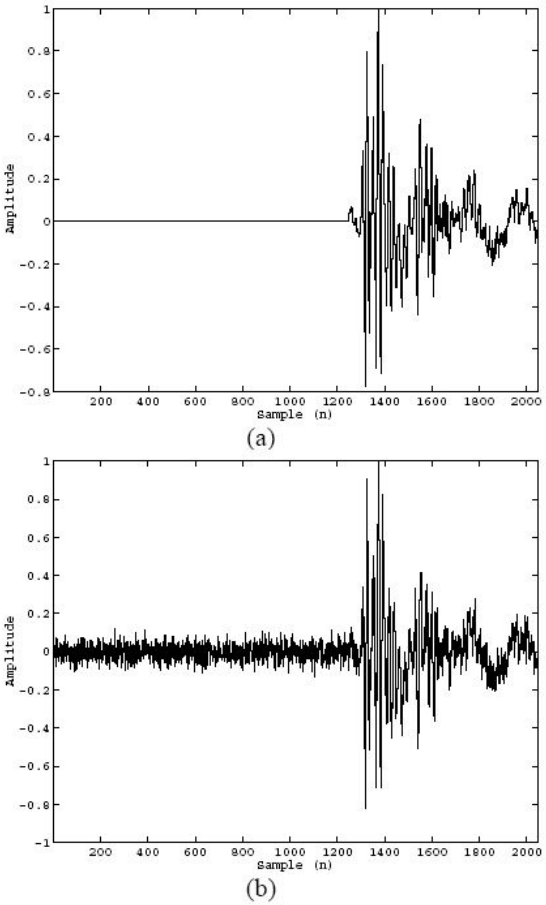
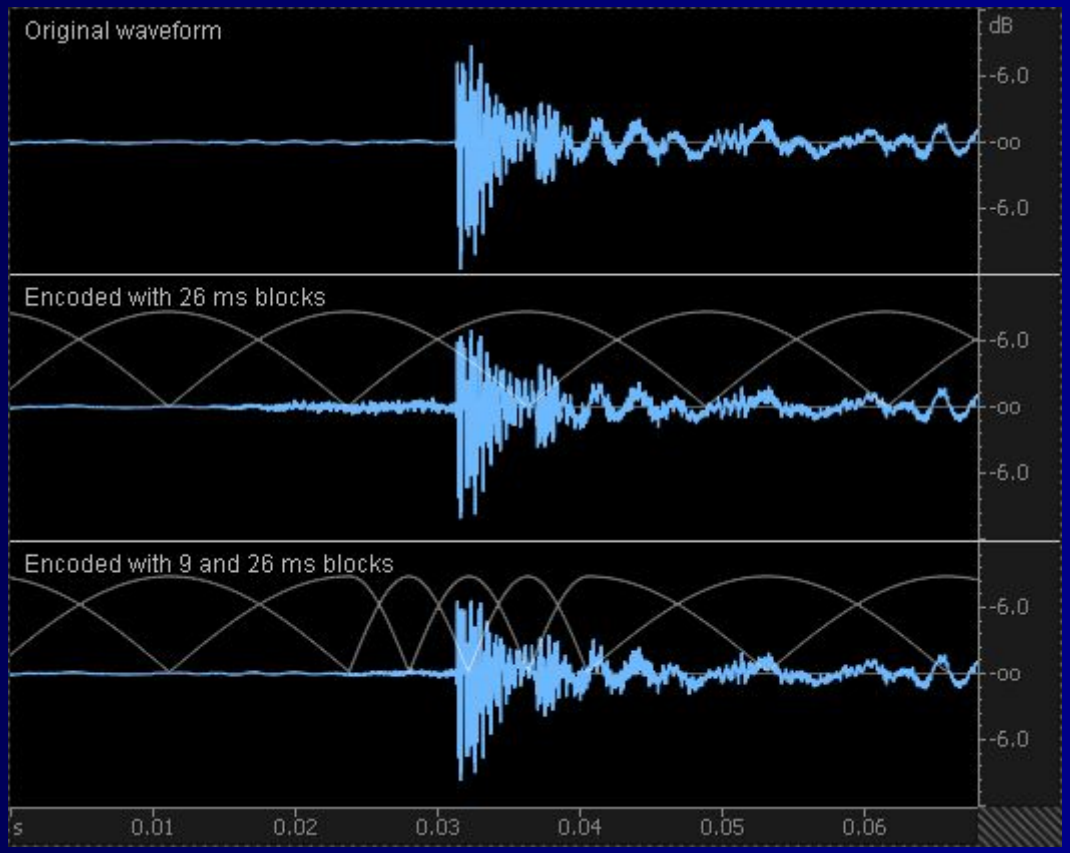


Fig. 8. Pre-Echo Example: (a) Uncoded Castanets. (b) Transform Coded Castanets, 2048-Point Block Size

Пред-эхо

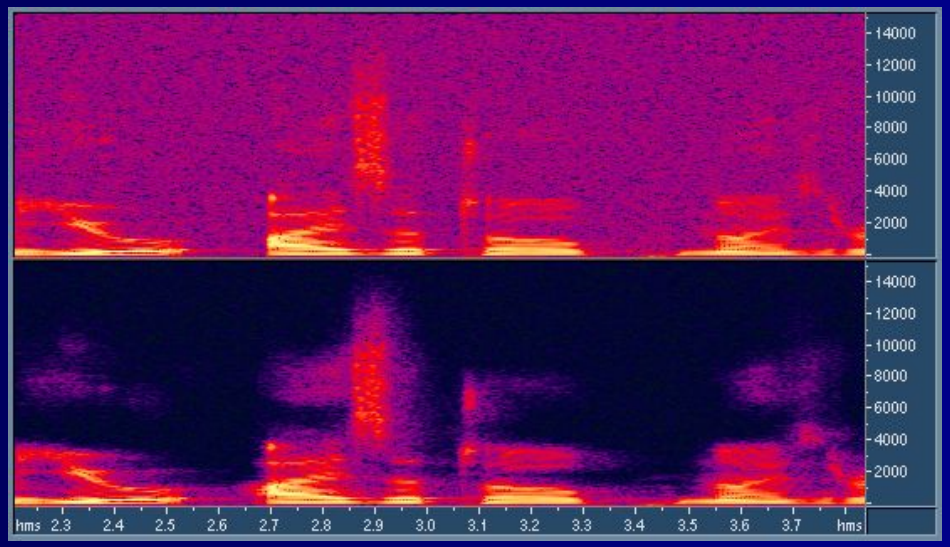
- Переключение размера окон в банке фильтров



Шумоподавление

- Аддитивный шум $$dirty[n] = clean[n] + noise[n]$$

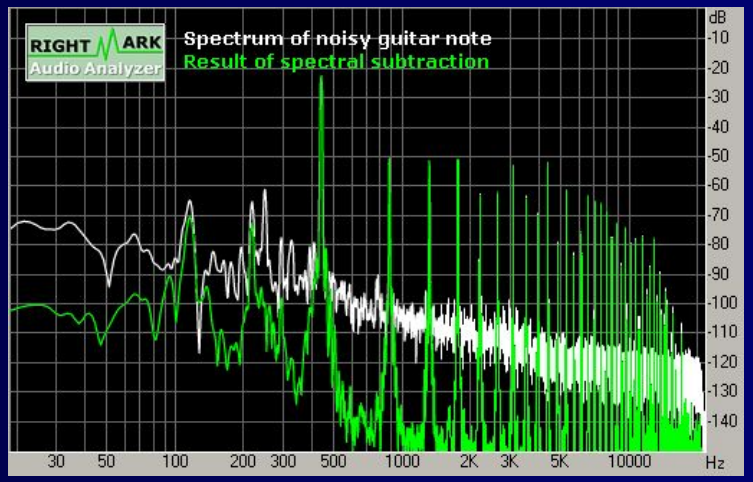
Шум предполагается стационарным,
т.е. не меняющимся во времени (средняя мощность, спектр)

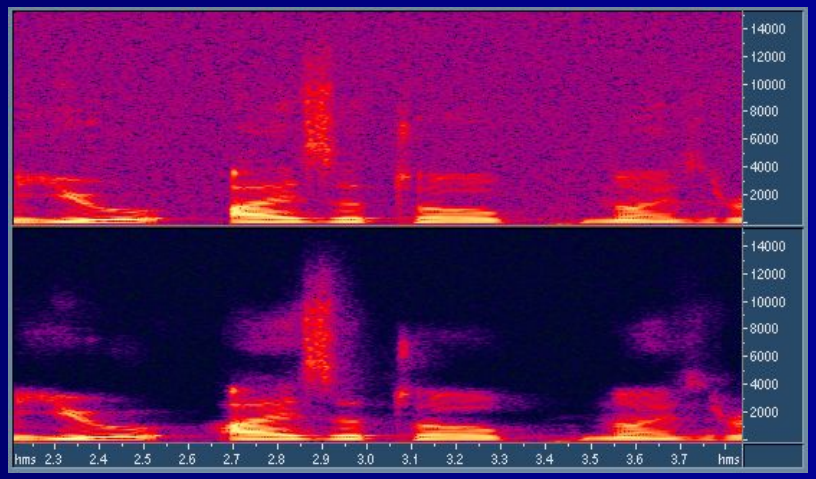


Метод спектрального вычитания

Стационарные шумы

- Общий принцип подавления
  1. Преобразование, компактно локализующее энергию (energy compaction)
  2. Модификация коэффициентов преобразования (подавление коэффициентов, соответствующих шуму)
  3. Обратное преобразование (восстановление очищенного сигнала)





Спектральное вычитание

Spectral Subtraction,
Short-Time Spectral Attenuation

- Спектральное вычитание для аудиосигналов
  1. STFT
  2. Оценка спектра шума по участку без полезного сигнала
  3. «Вычитание» спектра шума из спектра сигнала
  4. Обратное STFT

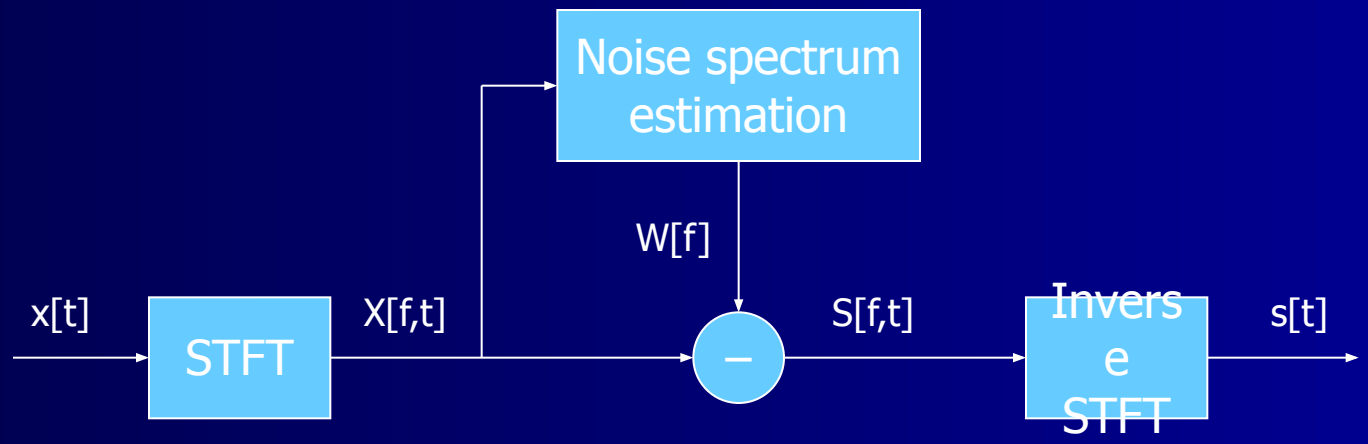


Схема алгоритма спектрального вычитания

Спектральное вычитание

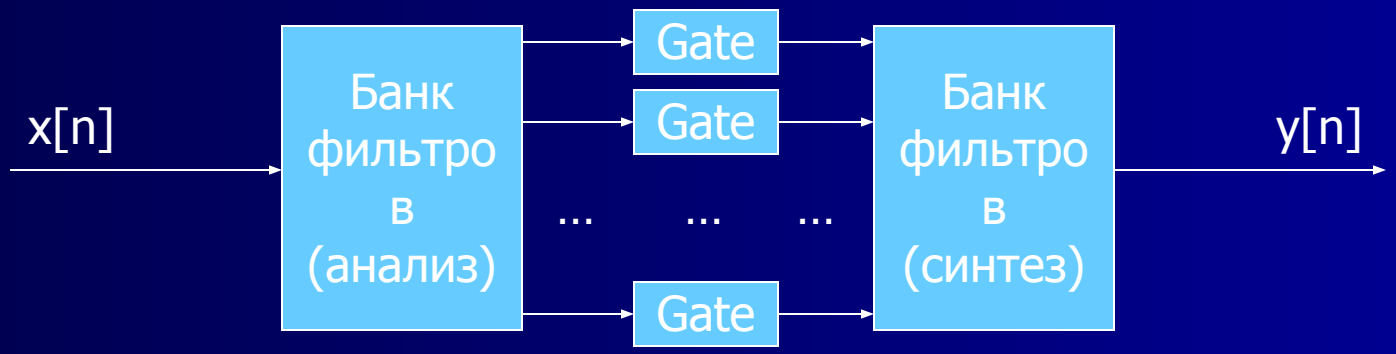
- Требования к банку фильтров
  - Точное (или почти точное) восстановление
  - Отсутствие «эффекта блочности» (перекрытие, окна)
  - Хорошая частотная локализация
  - Не требуется: сохранение количества информации

Выбираем банк фильтров,
основанный на STFT

Шумоподавление

- Многополосная интерпретация

Гейт (gate) – устройство, подавляющее тихие сигналы (громкие пропускаются без изменения)



Пороги срабатывания гейтов зависят от уровня шума в каждой частотной полосе

Спектральное вычитание

MSU GRAPHICS & MEDIA LAB

- Конструкция гейтов
  - Порог срабатывания зависит от шума → нужно знать параметры шума → обучение
  - Мягкое или жесткое срабатывание

Пример подавления:

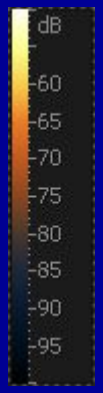
$$G[f,t] = \max\left\{1 - \frac{W[f,t]}{X[f,t]},\ 0\right\}$$

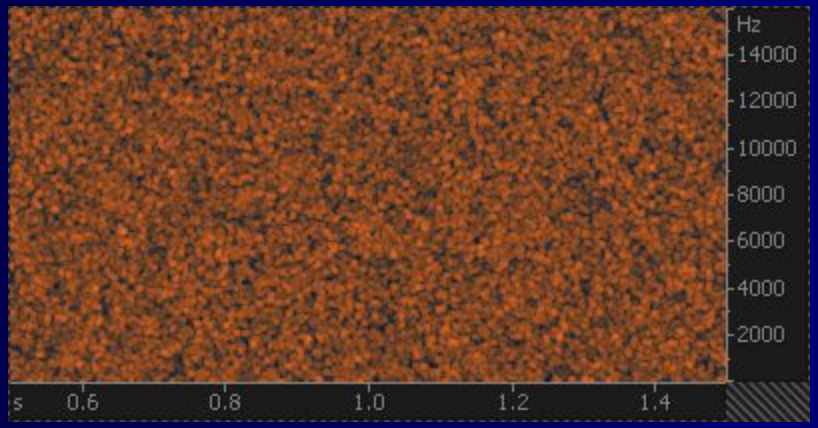
Здесь G – коэффициент усиления,
W – оценка амплитуды шума,
X – амплитуда сигнала.

  - Время срабатывания (attack/release time)
  - Ограничение степени подавления

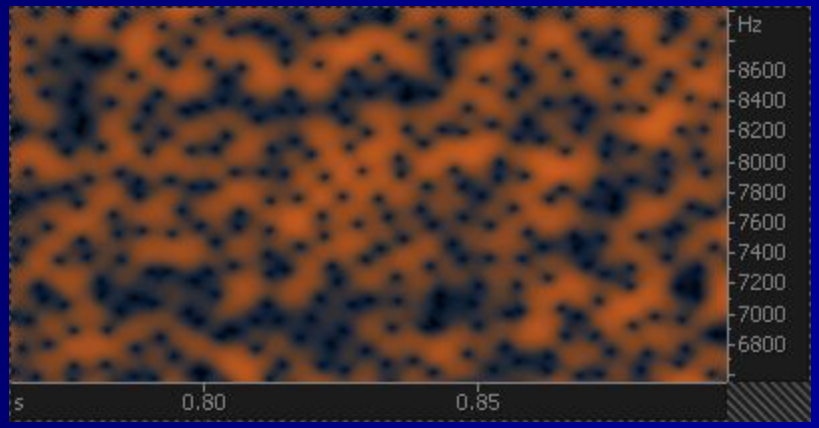
Шумоподавление

- Шум случаен → его спектр тоже случаен
- Пример спектра белого шума:





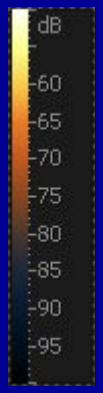
общая спектрограмма

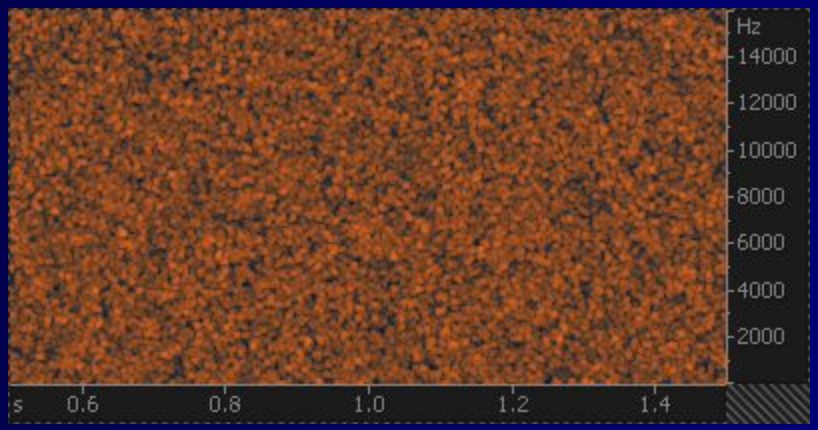


приближенный фрагмент

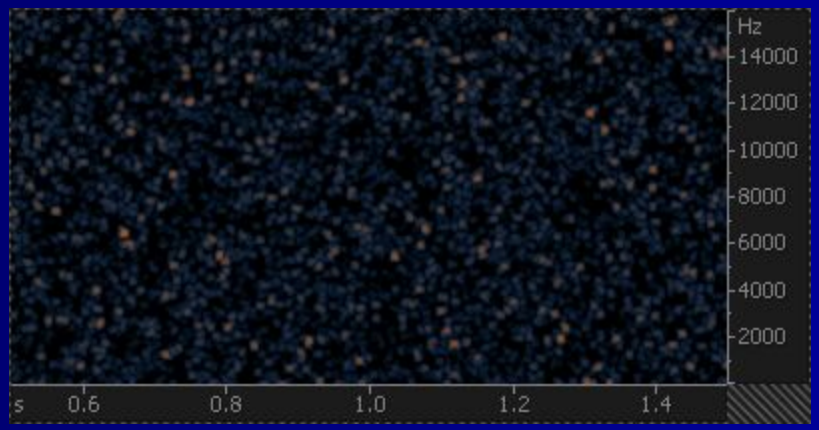
Шумоподавление

- После спектрального вычитания появляются случайно расположенные всплески энергии – артефакт «музыкальный шум» (musical noise)





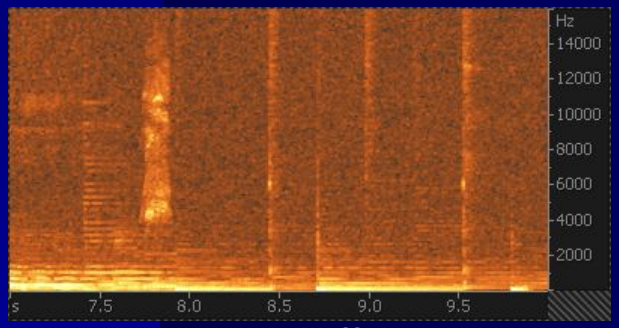
общая спектрограмма



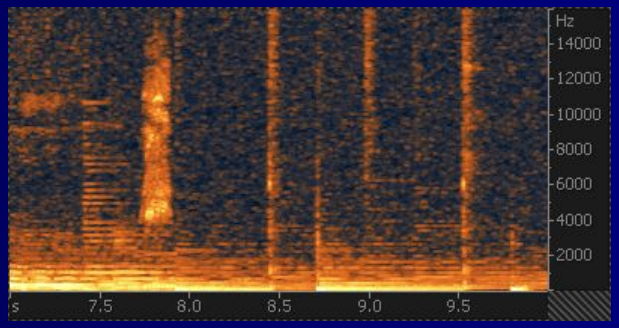
«музыкальный шум»

Шумоподавление

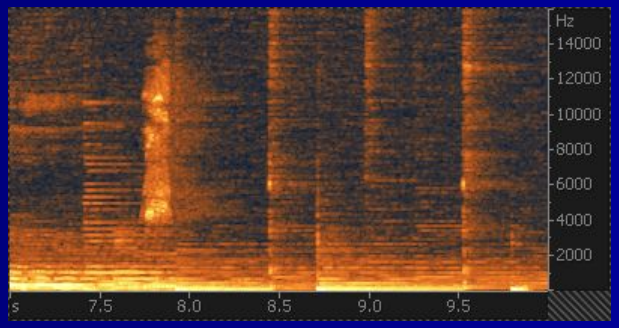
- Музыкальный шум: методы борьбы
  - Завышение порога *(недостаток – теряем больше сигнала)*
  - Ограничение G(f, t) снизу *(чтобы музыкальный шум маскировался естественным шумом)*
  - Увеличение времени срабатывания гейтов *(при слишком сильном увеличении получается шумовое эхо и смазываются атаки в сигнале)*
  - Сглаживание G(f, t) по времени и/или по частоте



зашумленный сигнал



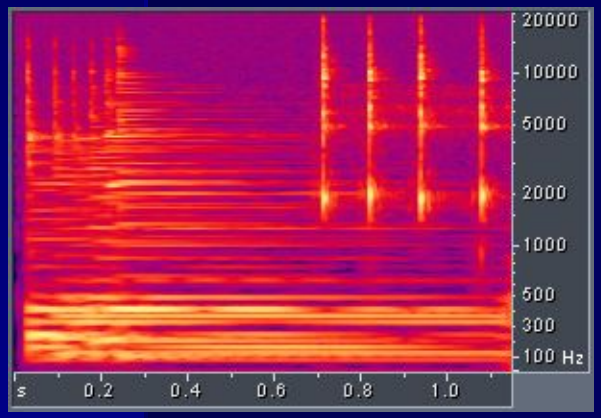
простое спектр. вычитание



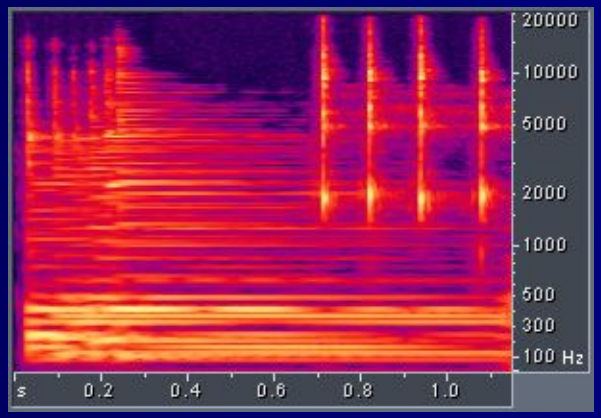
сглаживание по времени

Спектральное вычитание

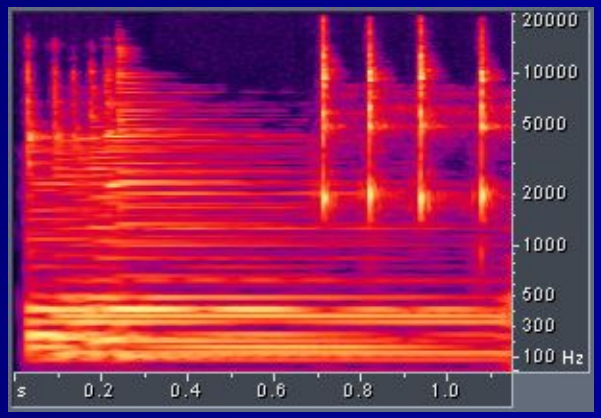
- Эффекты фиксированного частотно-временного разрешения
  - Эффект Гиббса (размытие транзиентов)
  - Недостаточное частотное разрешение



Зашумленный сигнал



Размер окна 50 мс



Адаптивное разрешение

Понятие вейвлета

- Вейвлеты – это сдвинутые и масштабированные копии $\psi_{a,b}(t)$ («дочерние вейвлеты») некоторой быстро затухающей осциллирующей функции $\psi(t)$ («материнского вейвлета»)

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$$

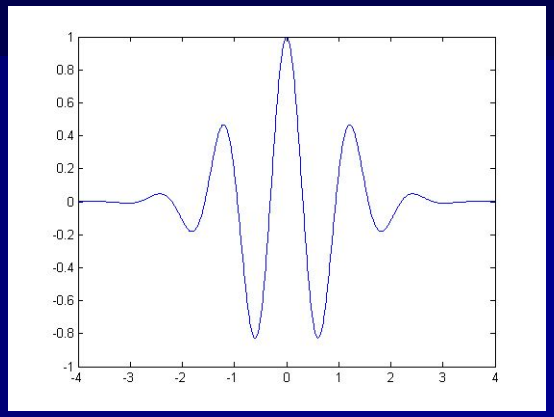
- Используются для изучения частотного состава функций в различных масштабах и для разложения/синтеза функций в компрессии и обработке сигналов

Понятие вейвлета

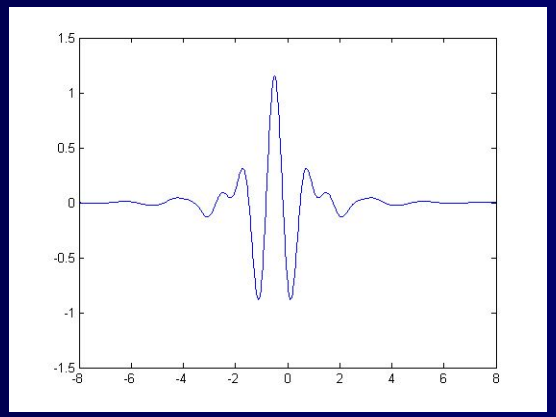
- Обычно накладываемые условия на ψ(t):
  - Интегрируемость $\int_{-\infty}^{\infty} |\psi(t)| dt < \infty \quad \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty$
  - Нулевое среднее, нормировка $\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad \int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1$
  - Нулевые моменты (vanishing moments) $\int_{-\infty}^{\infty} t^m \psi(t) dt = 0$

Понятие вейвлета

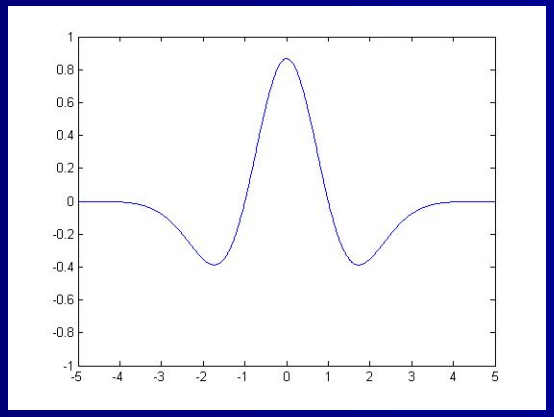
- Примеры вейвлетов



Mortlet



Meyer

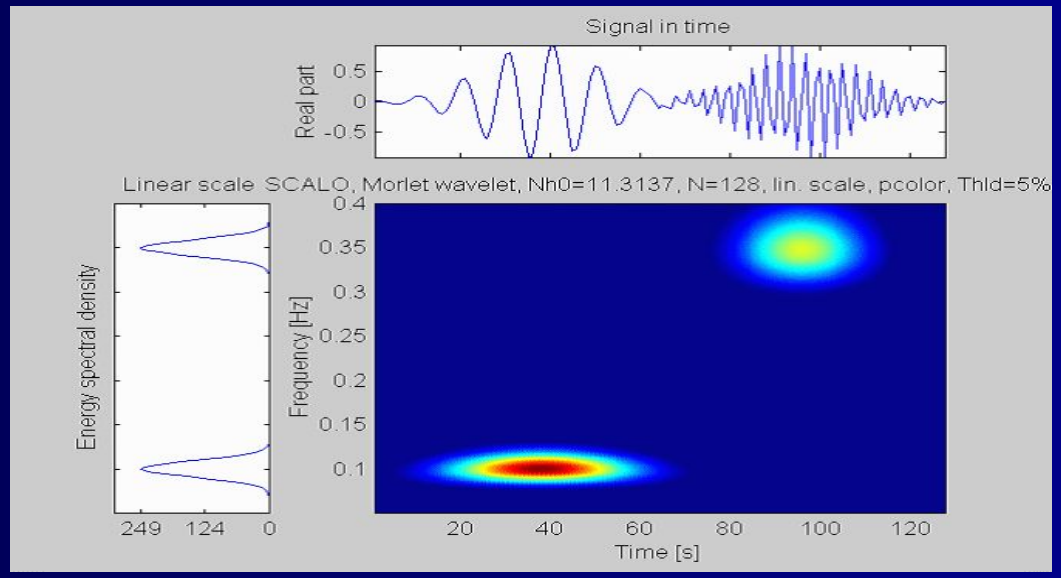


Mexican hat

Непрерывное вейвлет-преобразование (CWT)

- Скалярные произведения исследуемой функции f(t) с вейвлетами $\psi_{a,b}(t)$

$$W_{\psi}\{x\}(a,b) = \left\langle x, \psi_{a,b} \right\rangle = \int_{-\infty}^{+\infty} x(t)\psi_{a,b}(t)dt$$



Дискретное вейвлет-преобразование (DWT)

- Используются лишь целочисленные сдвиги вейвлета и масштабирование в 2 раза
- Возможность построения ортогонального преобразования
- Дискретный вейвлет
  1. Последовательность чисел $h_2[m]$
  2. Ортогональна своим сдвигам на четное число точек

$$\sum_{m=-\infty}^{+\infty} h_2[m]h_2[m+2k]=0, \quad \forall k \in Z, \quad k \neq 0$$

  3. Существует скейлинг-функция (НЧ-фильтр), ортогональная вейвлету

$$\sum_{m=-\infty}^{+\infty} h_1[m]h_2[m]=0$$

Преобразование Хаара

- Простейший случай вейвлет-преобразования
  - Дан входной сигнал x[n]
  - Образуем от него последовательности полусумм и полуразностей:

$$x_1^*[n] = \frac{x[n] + x[n+1]}{2} \qquad x_2^*[n] = \frac{x[n] - x[n+1]}{2}$$

  - Легко видеть, что сигнал x[n] можно восстановить:

$$x[n] = x_1^*[n] + x_2^*[n]$$

  - Такое кодирование избыточно: из одной последовательности получаем две

Преобразование Хаара

- Устранение избыточности

  Проредим полученные последовательности в 2 раза:

$$x_1[n] = x_1^*[2n] \qquad x_2[n] = x_2^*[2n]$$

  Легко видеть, что справедлив алгоритм восстановления:

$$y_i[n] = \begin{cases} x_i\left[\frac{n}{2}\right], & n - \text{четное} \\ 0, & n - \text{нечетное} \end{cases} \qquad i = 1,\ 2$$

(интерполяция нулями)

$$x_1^{**}[n] = y_1[n] + y_1[n-1] \qquad x_2^{**}[n] = y_2[n] - y_2[n-1]$$

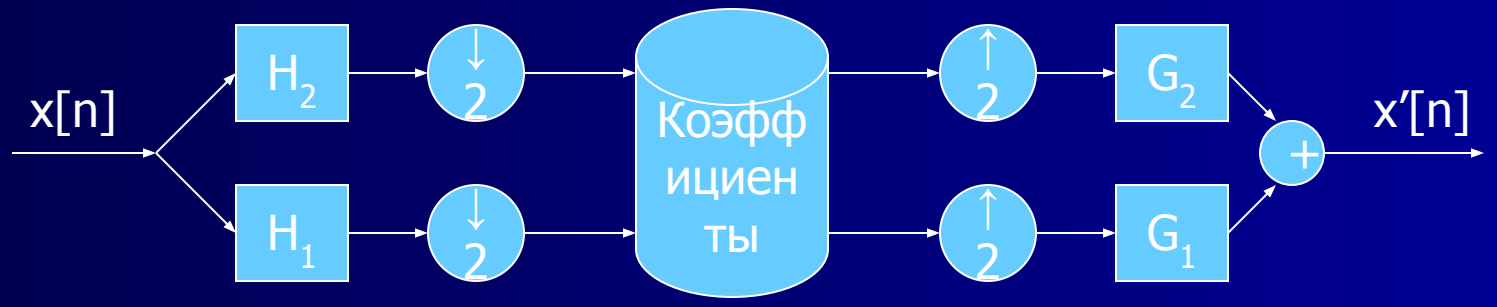
(фильтрация)

$$x[n] = x_1^{**}[n] + x_2^{**}[n]$$

(суммирование)

Дискретное вейвлет-преобразование

- Обобщение преобразования Хаара



Декомпозиция Реконструкция

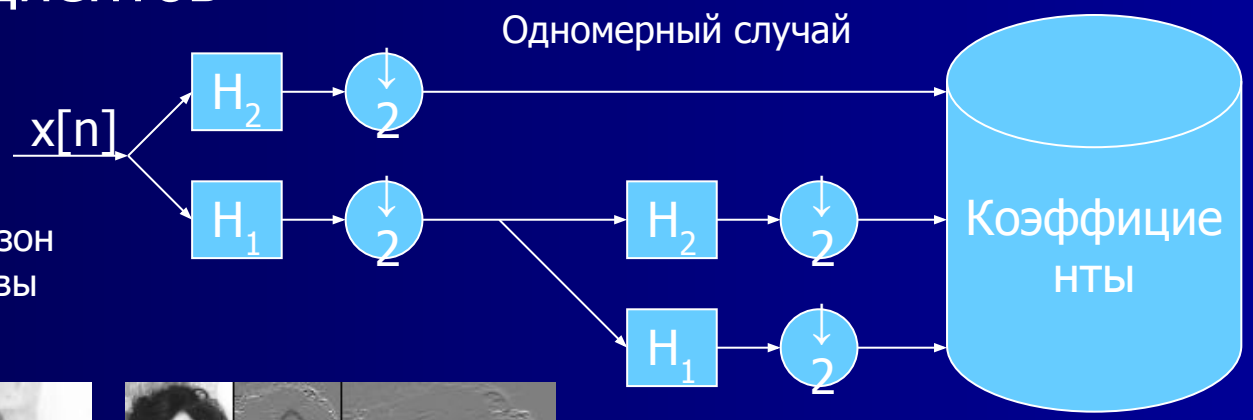
Свойство точного восстановления (PR): $x[n] \equiv x'[n]$

Количество информации не изменяется.

Нужно найти хорошие фильтры, обеспечивающие точное восстановление.

Пирамидальное представление

- Продолжаем вейвлет-разложение для НЧ-коэффициентов



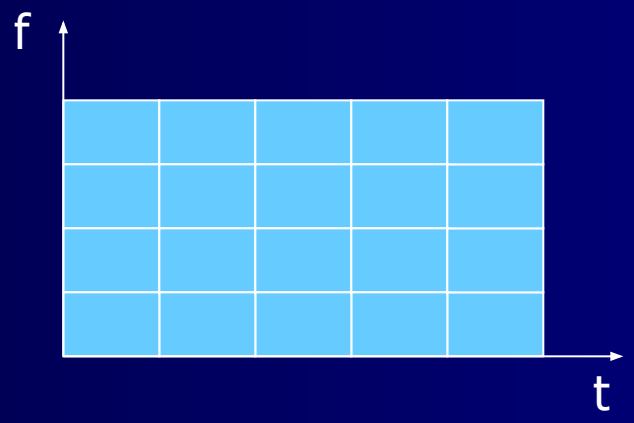
Частотный диапазон делится на октавы

Двумерное вейвлет-преобразование

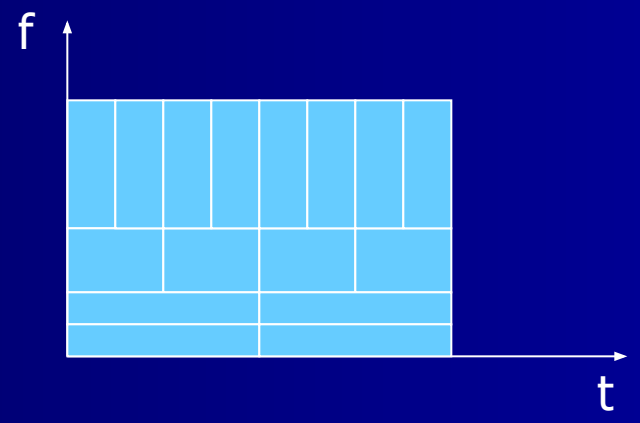
на каждом шаге получаем 4 набора коэффициентов: НЧ («основные») и ВЧ («детализирующие»)

Банки фильтров

- Как банки фильтров разбивают частотно-временную плоскость?



STFT



Вейвлеты

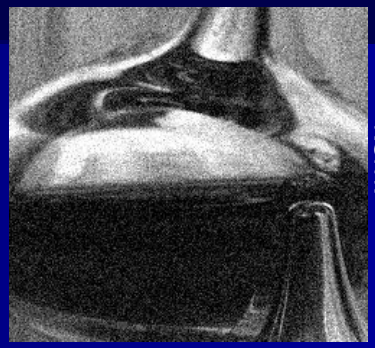
MSU GRAPHICS & MEDIA LAB

Часть 2

Подавление шума на изображениях

Виды и примеры шумов

Salt and pepper



AWGN

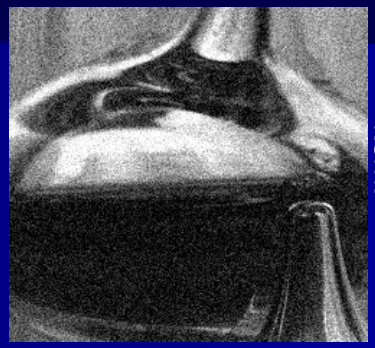
- Шумы
  - Импульсные
    - Salt and pepper
    - Помехи в видео
  - Стационарные
    - Аддитивный белый
    - Зерно пленки
  - Смешанные
    - Шум + помехи в видео

Белый шум – пиксели случайны, не коррелированны друг с другом.
Гауссов/равномерный/треугольный шум – закон распределения амплитуд пикселей.
Аддитивный шум – прибавляется к «чистому» изображению и не зависит от него.

Методы шумоподавления

MSU GRAPHICS & MEDIA LAB

Salt and pepper



AWGN

Шумы

- Импульсные
  - Медианный фильтр
  - Взвешенная медиана
  - Ранговые фильтры
- Стационарные
  - Bilateral filter
  - Non-Local Means
  - Wavelet thresholding
  - DCT, PCA, ICA
  - Анизотропная диффузия
  - Алгоритм BM3D
- Смешанные
  - Ранговые фильтры
  - Комбинированные методы

Простейшие методы

- Простейшие методы
  - Размытие изображения – вместе с шумом размывает детали

  - Размытие в гладких областях – остается шум вблизи границ
  - Медианная фильтрация – хорошо подавляет импульсный шум, но удаляет мелкие детали

Bilateral filter

- Адаптивные алгоритмы
  - Bilateral filter
    усреднение окружающих пикселей с весами

$$y_{i,j} = \sum_{k,m \in Q} x_{i+k,j+m} \cdot W(i,j,k,m)$$

$$W(i,j,k,m) \approx \exp\left(-\frac{(x_{i,j} - x_{i+k,j+m})^2}{h^2}\right) \cdot \exp\left(-\frac{k^2+m^2}{\rho^2}\right)$$

фотометрическая близость пространственная близость

Bilateral filter

- Bilateral filter: художественное применение

(при слишком сильном действии)

Non-Local Means

- Адаптивные алгоритмы
  - Non-local means (NL-means) – веса зависят от близости целых блоков, а не отдельных пикселей

$$W(i,j,k,m) \approx \exp\left(-\frac{\left\|\upsilon(x_{i,j}) - \upsilon(x_{i+k,j+m})\right\|^2}{h^2}\right)$$

$v(x_{i,j})$ – блок вокруг пикселя $x_{i,j}$

Non-Local Means

- Вычисление весов

Веса высоки для q1 и q2,
но не для q3

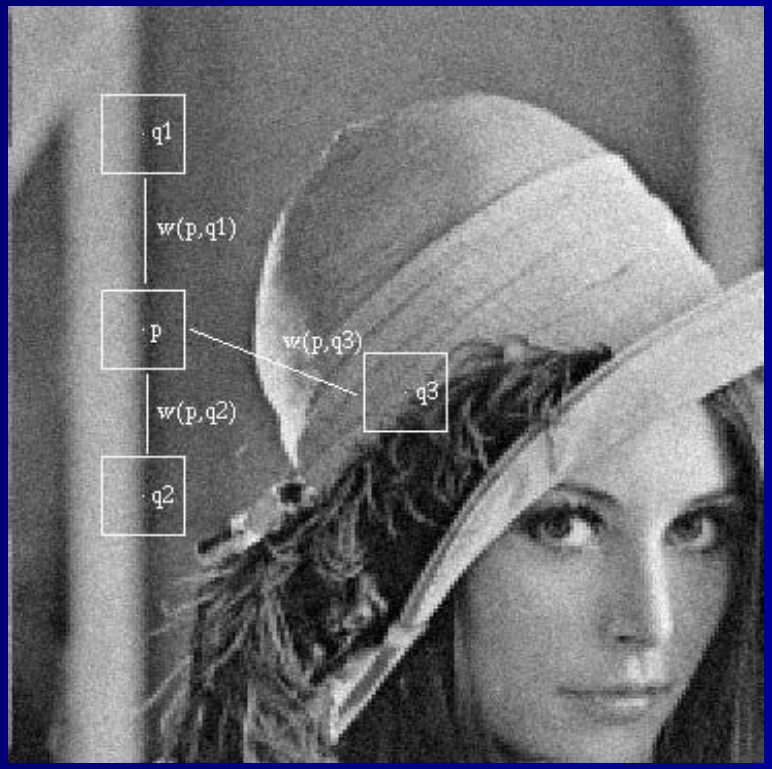


Иллюстрация из
Buades et al 2005

+ ► Способен сохранять текстуру изображения лучше, чем bilateral filter

Non-Local Means

MSU GRAPHICS & MEDIA LAB

- Достоинства и недостатки:
  - ► Высокое качество результирующего изображения
  - ► В исходном варианте – очень высокая вычислительная сложность
- Ускоряющие расчет оптимизации:
  - ► Использование команд MMX/SSE для сравнения блоков
  - ► Разбиение изображения на блоки и усреднение целых блоков, а не отдельных пикселей
  - ► Промежуточный вариант между усреднением блоков и усреднением пикселей: усреднение маленьких блоков

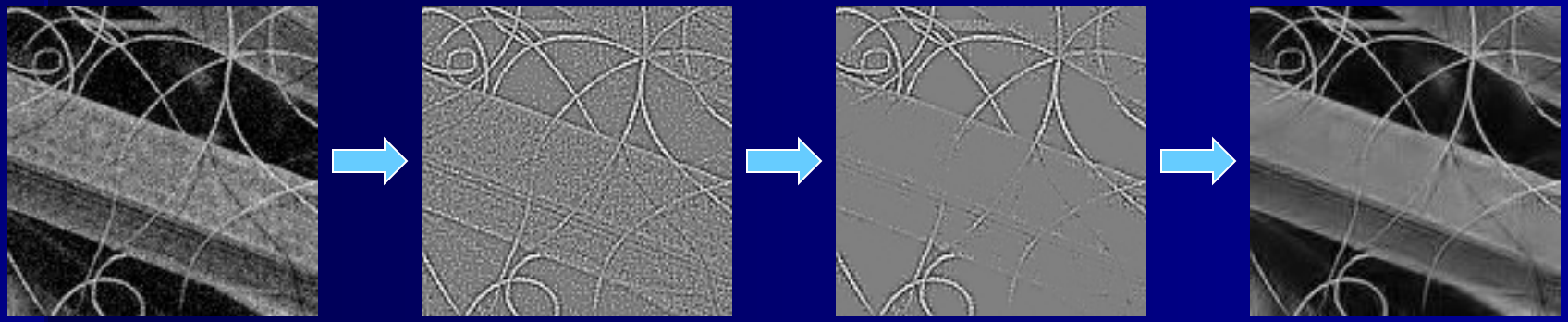
Non-Local Means

MSU GRAPHICS & MEDIA LAB

- Применение к видео
  - Область поиска блоков можно расширить на соседние кадры (сделать ее пространственно-временной)
  - Для ускорения просчета можно применять сравнение блоков только по Y-каналу в цветовой модели YCrCb (YUV)

Вейвлетный метод

- Вейвлетное шумоподавление для изображений
  1. DWT
  2. Оценка уровня и спектра шума
  3. Подавление вейвлет-коэффициентов (thresholding, shrinkage)
  4. Обратное DWT



\+ Подавление шума различного масштаба

\– Отсутствие инвариантности к сдвигу
Плохая локализация энергии для наклонных границ