Лектор: Лукин Алексей Сергеевич

MSU GRAPHICS & MEDIA LAB

Дискретизация
Свертка
ДПФ

План

- Звуковые сигналы и их восприятие
- Цифровые и аналоговые сигналы. Дискретизация.
- Теорема Котельникова. Алиасинг. Фильтрация звука.
- Линейные системы. Свертка.
- Простейшие двумерные фильтры для изображений
- Дискретное преобразование Фурье
- Спектральный анализ

Звук и слух

- Диапазон звуковых сигналов и пороги восприятия

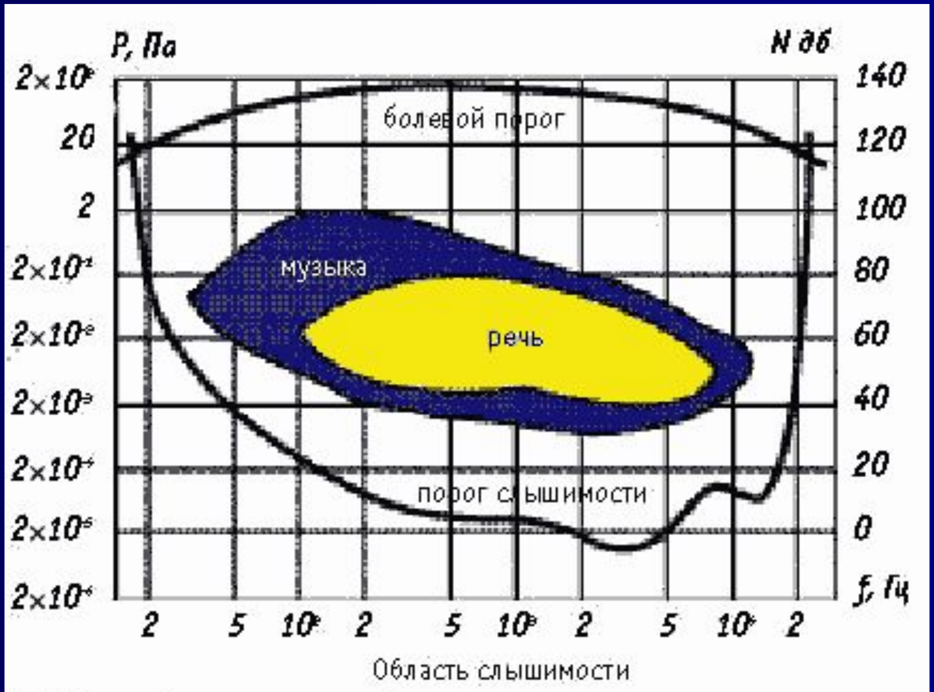


Рис 6 Кривые болевых порогов и область слышимости

Основы слухового восприятия

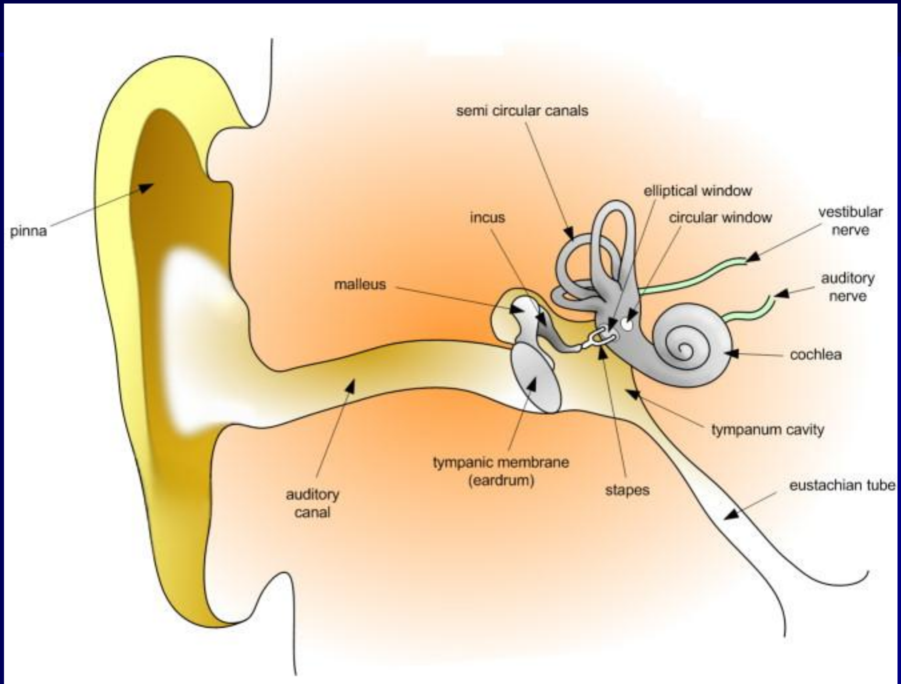


image from Wikipedia

- Звуковые волны поступают на улитку, возбуждая ее колебания
- Жесткость улитки меняется с расстоянием, поэтому каждая часть резонирует в своем частотном диапазоне

Основы слухового восприятия

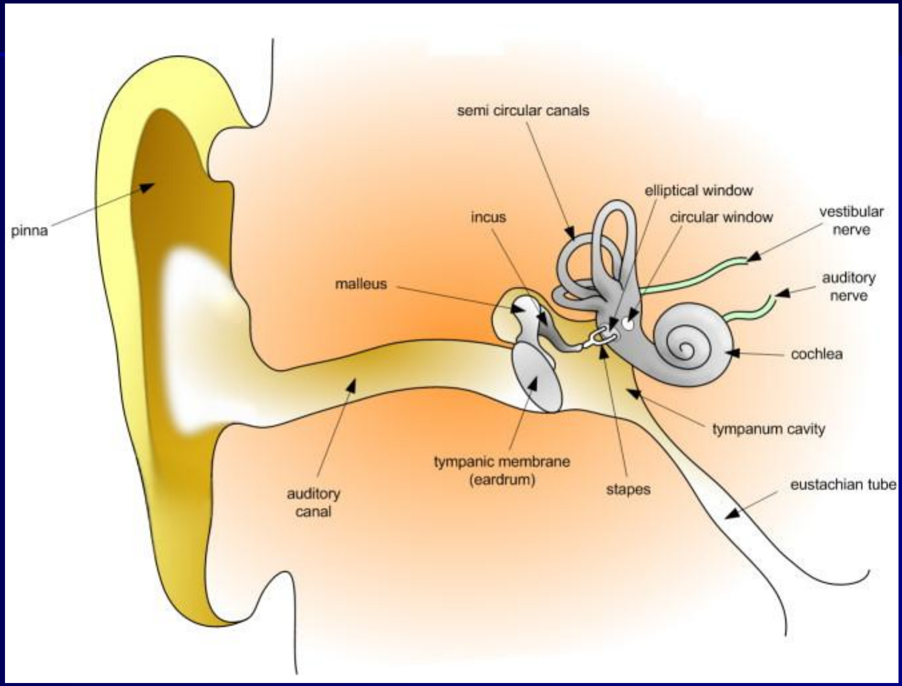


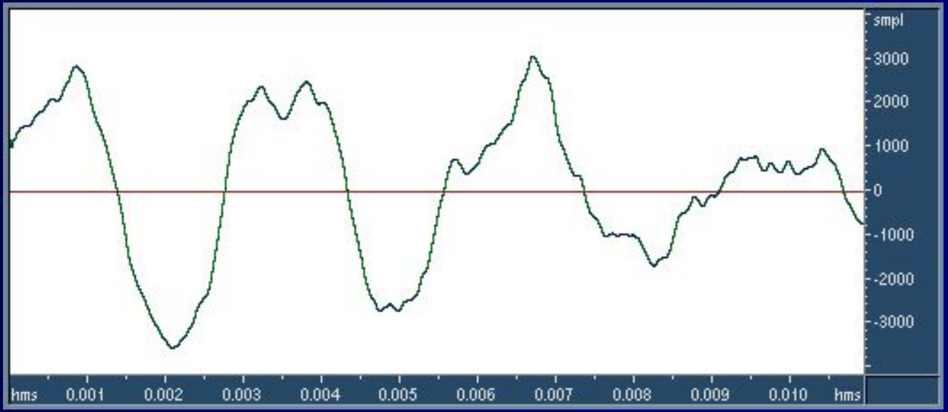
image from Wikipedia

- К разным частям улитки подходят различные группы нервов, передающие в мозг информацию об амплитуде и фазе колебаний
- Таким образом, улитка раскладывает звук на частотные составляющие

Сигналы

- Сигнал – скалярная функция от одного или нескольких аргументов.

Примеры сигналов



s(t) – звук

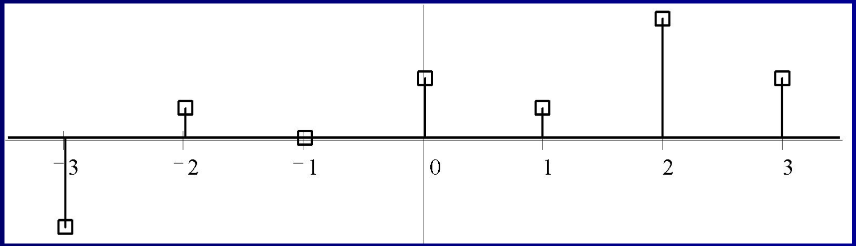
f(x,y) – изображение

Сигналы

MSU GRAPHICS & MEDIA LAB

- Аналоговые (непрерывные)
  - Примеры:
    - звук в воздухе или в проводе, идущем от микрофона
    - изображение (до ввода в компьютер)
    - запись показаний датчика
- Цифровые (дискретные)
  - Примеры:
    - звук в компьютере (одномерный массив чисел)
    - изображение в компьютере (двумерный массив чисел)
    - запись показаний датчика в компьютере (одномерный массив)

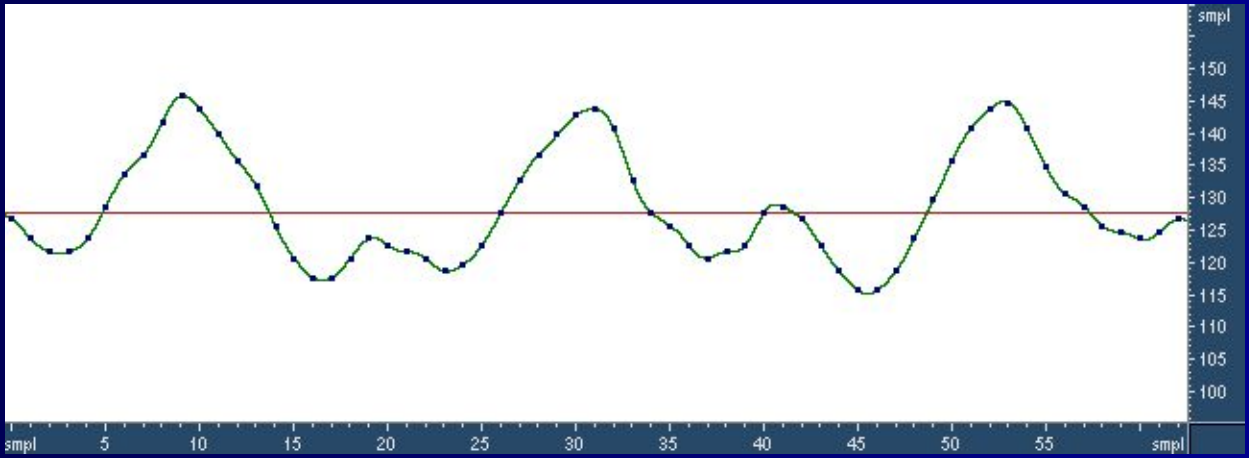
Одномерный цифровой сигнал



Оцифровка сигналов

MSU GRAPHICS & MEDIA LAB

1. Дискретизация по времени (аргумент функции)
2. Квантование по амплитуде (значение функции)



- АЦП (ADC) – аналогово-цифровой преобразователь

  Параметры: частота дискретизации, разрядность квантования (пример: 44.1 кГц, 16 бит – формат Audio CD)

Оцифровка сигналов

MSU GRAPHICS & MEDIA LAB

- При каких условиях по цифровому сигналу можно точно восстановить исходный аналоговый?
- Предположим, что значения амплитуд в цифровом сигнале представлены точно.
- Введем понятие спектра аналогового сигнала:

$$X(\nu) = \int_{-\infty}^{+\infty} x(t) \cdot e^{-2\pi i \nu t} dt \qquad x(t) = \int_{-\infty}^{+\infty} X(\nu) \cdot e^{2\pi i \nu t} d\nu$$

(разложение на синусоиды с различными частотами)

$$e^{-pt} = \cos pt - i \cdot \sin pt$$

x(t) – исходный сигнал

X(ν) – спектр, т.е. коэффициенты при гармониках с частотой ν

Теорема Котельникова

MSU GRAPHICS & MEDIA LAB

- Пусть

1. спектр сигнала x(t) не содержит частот выше F, т.е. X(v)=0 за пределами отрезка [-F, F]
2. дискретизация сигнала x(t) производится с частотой F_s , т.е. в моменты времени nT, здесь T= F_s^{-1}
3. $F_s > 2F$

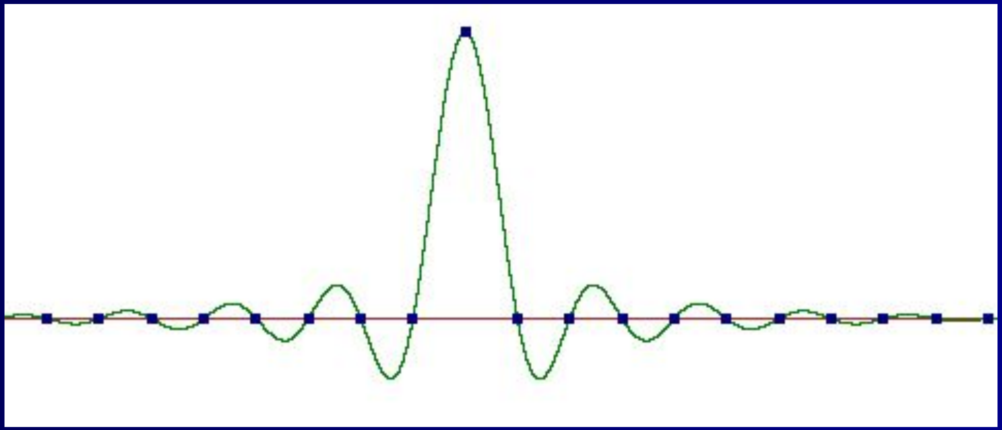
- Тогда исходный аналоговый сигнал x(t) можно точно восстановить из его цифровых отсчетов x(nT), пользуясь интерполяционной формулой

$$x(t) = \sum_{n=-\infty}^{+\infty} x(nT) \cdot Sinc(t - nT) \qquad Sinc(t) = \frac{\sin \pi F_s t}{\pi F_s t}$$

Теорема Котельникова

- Как выглядят интерполирующие sinc-функции?

$$x(t) = \sum_{n=-\infty}^{+\infty} x(nT) \cdot Sinc(t - nT) \qquad Sinc(t) = \frac{\sin \pi F_s t}{\pi F_s t}$$

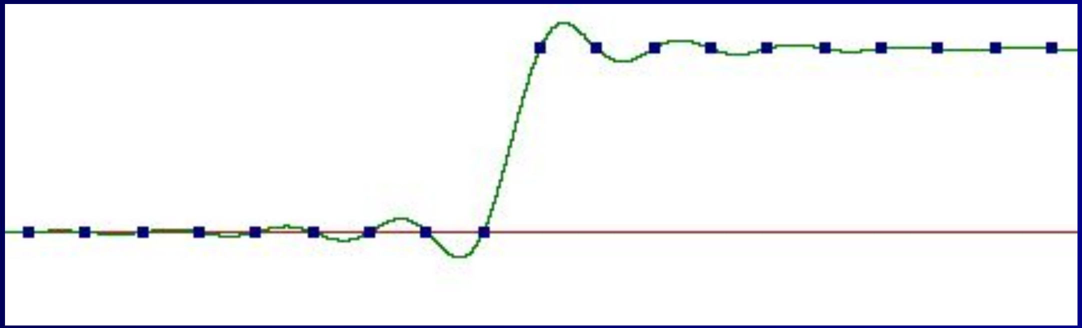


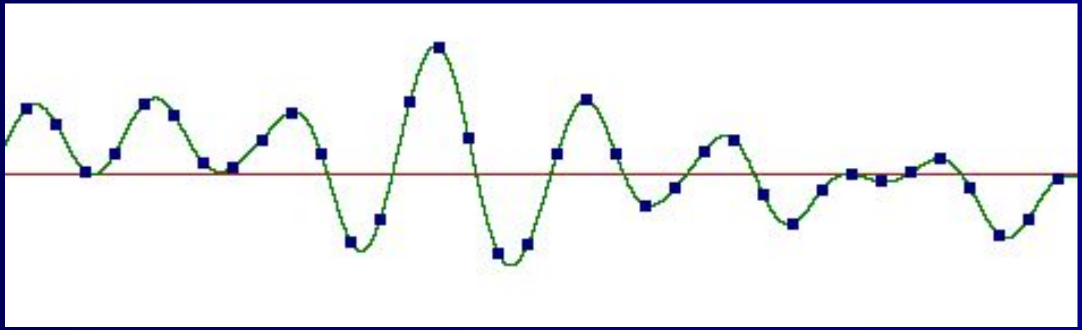
Бесконечно затухающие колебания

Теорема Котельникова

- Реконструкция аналоговых сигналов. Sinc-интерполяция.

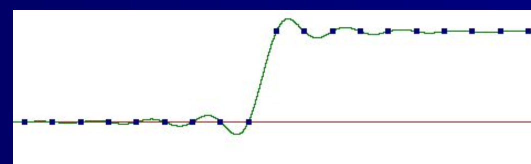
$$x(t) = \sum_{n=-\infty}^{+\infty} x(nT) \cdot Sinc(t - nT)$$





Эффект Гиббса

- Применимость sinc-интерполяции для изображений
- Эффект Гиббса



Цифровые отсчеты

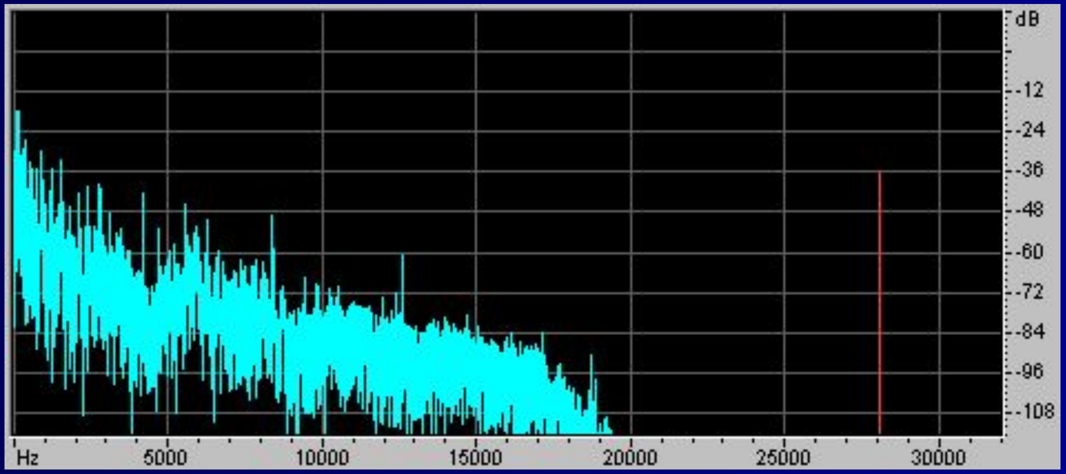
sinc-интерполяция

другая интерполяция

Наложение спектров

MSU GRAPHICS & MEDIA LAB

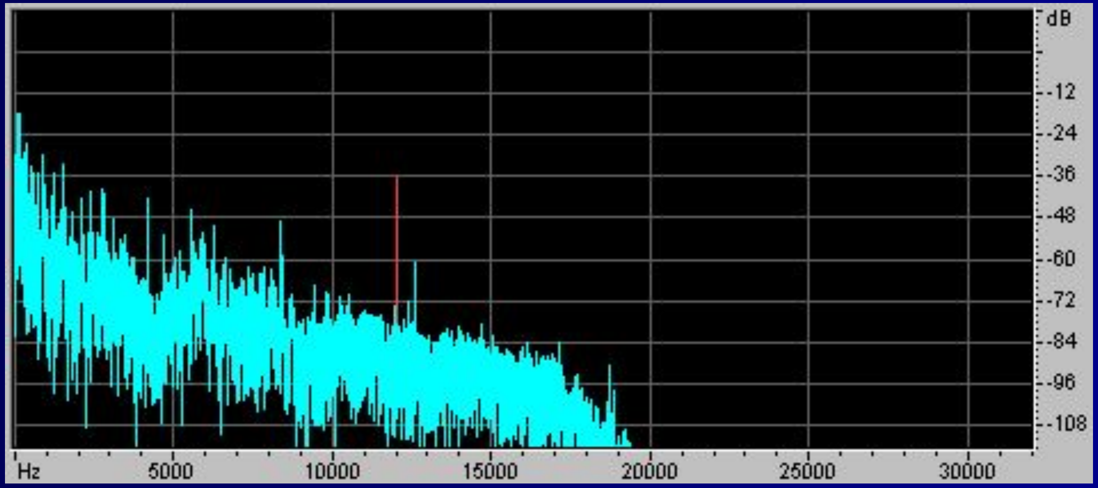
- Что будет, если условия теоремы Котельникова не выполнены?
- Пусть звук не содержит частот выше 20 кГц. Тогда, по теореме Котельникова, можно выбрать частоту дискретизации 40 кГц.
- Пусть в звуке появилась помеха с частотой 28 кГц. Условия теоремы Котельникова перестали выполняться.



Наложение спектров

MSU GRAPHICS & MEDIA LAB

- Проведем дискретизацию с частотой 40 кГц, а затем – восстановим аналоговый сигнал sinc-интерполяцией.



- Помеха отразилась от половины частоты дискретизации в нижнюю часть спектра и наложилась на звук. Помеха переместилась в слышимый диапазон. Это называется наложением спектров (алиасинг).

Наложение спектров

- Как избежать наложения спектров?
- Применить перед оцифровкой анти-алиасинговый фильтр
  - Он подавит все помехи выше половины частоты дискретизации (выше 20 кГц) и пропустит весь сигнал ниже 20 кГц.
  - После этого условия теоремы Котельникова будут выполняться и алиасинга не возникнет.
  - Следовательно, по цифровому сигналу можно будет восстановить исходный аналоговый сигнал.

Линейные системы

- Система – преобразователь сигнала.

$$y(t) = H(x(t))$$

- Линейность:

$$H(\alpha \cdot x(t)) = \alpha \cdot H(x(t))$$

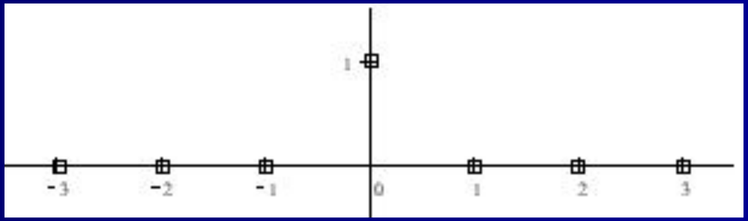
$$H(x(t) + z(t)) = H(x(t)) + H(z(t))$$

- Инвариантность к сдвигу:

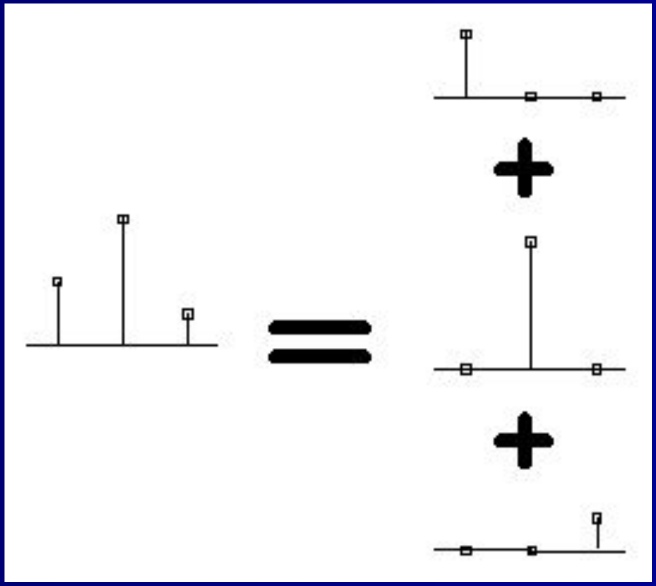
$$H(x(t - t_0)) = y(t - t_0)$$

Импульсная характеристика

- Единичный импульс δ[n]

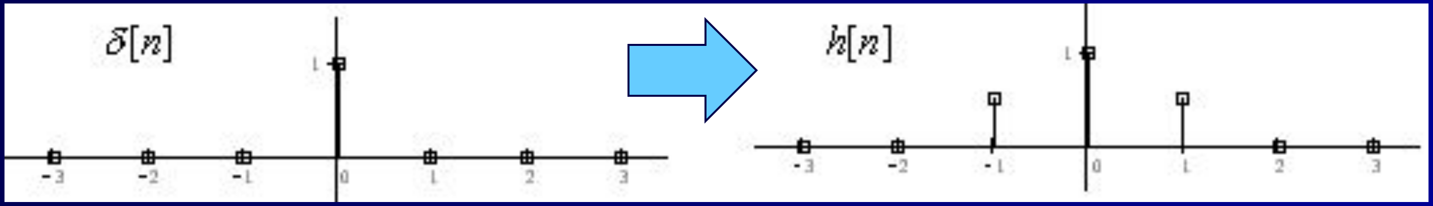


- Разложение произвольного сигнала на взвешенную сумму единичных импульсов



Импульсная характеристика

- Отклик системы на единичный импульс



- h[n] – импульсная характеристика системы (импульсный отклик системы)

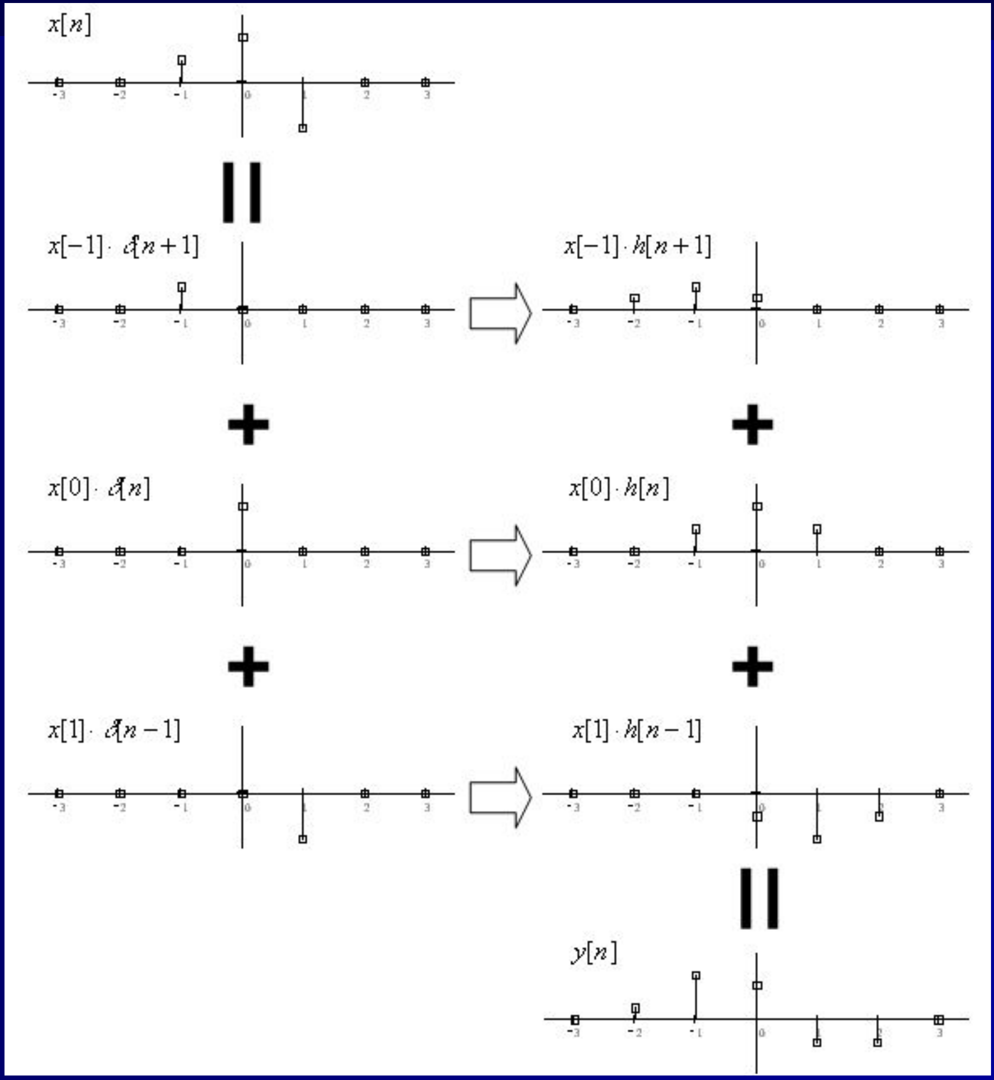
Импульсная характеристика

- Вычисление отклика линейной системы на произвольный входной сигнал
- Свертка

$$y[n] = h[n] * x[n]$$

$$y[n] = \sum_{k=-\infty}^{+\infty} x[n-k]h[k]$$

h[n] – ядро свертки



Линейные системы

MSU GRAPHICS & MEDIA LAB

- Итак, любая линейная инвариантная к сдвигу система производит операцию свертки входного сигнала со своей импульсной характеристикой.
- Важное свойство линейных систем:

  При подаче на любую линейную систему синусоиды, на выходе получается синусоида той же частоты, что и на входе. Измениться могут только ее амплитуда или фаза.
- Следствие: линейные системы удобно анализировать, раскладывая любые входные сигналы на синусоиды.

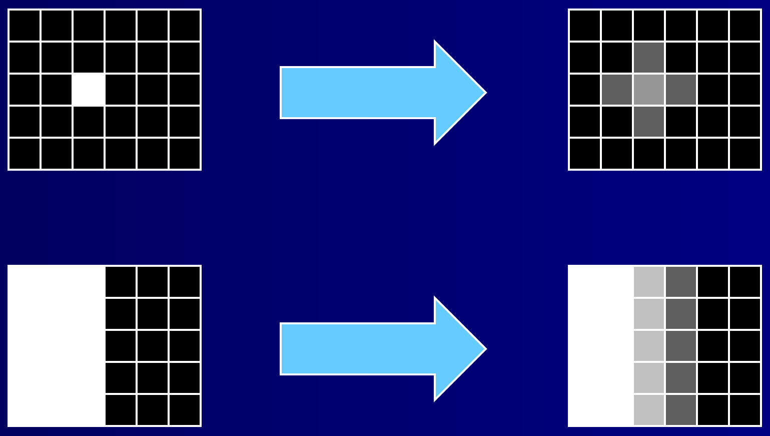
Двумерные фильтры

- Как работают фильтры

Коэффициенты фильтра,
ядро свертки 3x3,
«функция размытия точки»

$$Ker[k,p] = \frac{1}{6} \cdot \begin{Vmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{Vmatrix}$$

$-1 \le k \le 1$,
$-1 \le p \le 1$



Двумерные фильтры

MSU GRAPHICS & MEDIA LAB

- Свертка

$$Dst[i,j] = Src[i,j] * Ker[k,p]$$

$$Dst[i,j] = \sum_{k,p} Src[i-k, j-p] \cdot Ker[k,p]$$

```
// Обнулить изображение Dest[i][j]
...
// Выполнить свертку
for (i=0; i<Height; i++)        // Для каждого пикс. Dest[i][j]...
  for (j=0; j<Width; j++)
    for (k=-1; k<=1; k++)       // ...превратить его в ядро свертки
      for (p=-1; p<=1; p++)
        Dest[i+k][j+p] += Src[i][j] * Ker[k][p];    // и сложить
```

Подводные камни:

► Выход за границы массива

► Выход за пределы допустимого диапазона яркости пикселей

► Обработка краев.

Двумерные фильтры

- Свойства фильтров
  1. Результат фильтрации однотонного (константного) изображения – константное изображение. Его цвет равен
  $$Dest = Src \cdot \sum_{k,p} Ker[k,p]$$
  2. Следствие: чтобы фильтр сохранял цвет однотонных областей, нужно чтобы
  $$\sum_{k,p} Ker[k,p] = 1$$
  3. Следствие: если сумма коэффициентов фильтра равна нулю, то он переводит однотонные области в нулевые.

Примеры фильтров

- Размытие (blur)

Примеры фильтров

- Повышение четкости (sharpen)

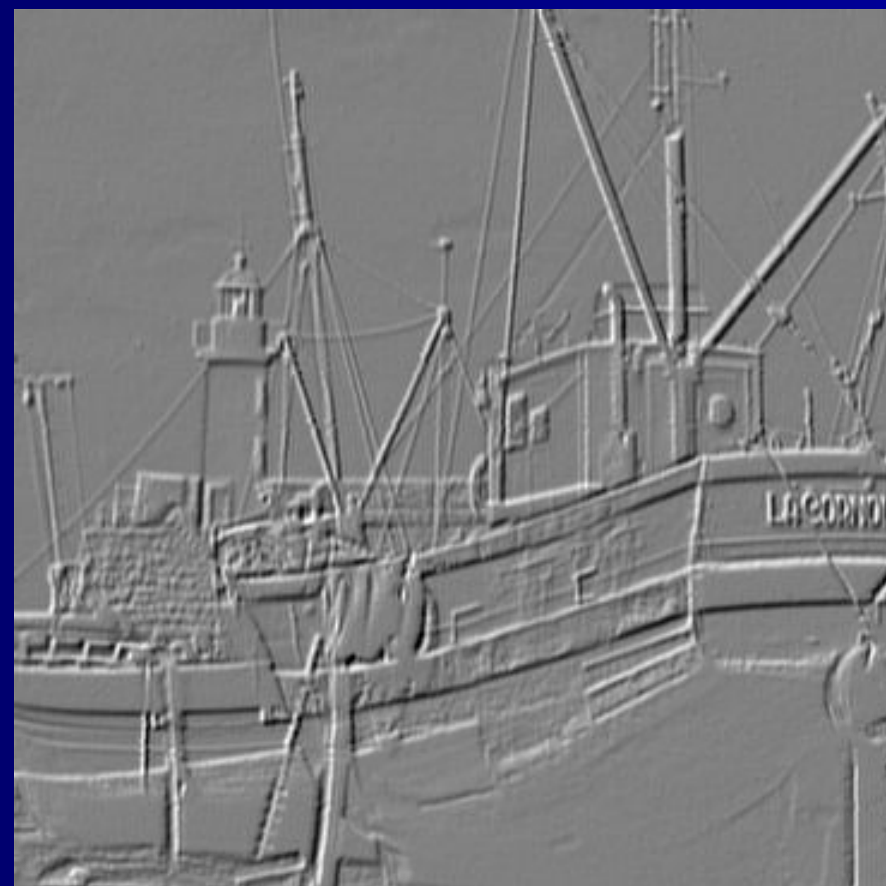
Примеры фильтров

- Нахождение границ (edges)

Примеры фильтров

MSU GRAPHICS & MEDIA LAB

- Тиснение (embossing)



Примеры фильтров

MSU GRAPHICS & MEDIA LAB

- Простейшее размытие

$$Ker[k,p]=\frac{1}{15}\cdot\begin{Vmatrix}1 & 2 & 1\\ 2 & 3 & 2\\ 1 & 2 & 1\end{Vmatrix}$$

- Константное размытие "box-фильтр"

  (любой размер фильтра)

$$Ker[k,p]=\frac{1}{Sum}$$

- Гауссово размытие

  (любой размер фильтра)

$$Ker[k,p]=\frac{1}{Sum}\cdot\exp\frac{k^2+p^2}{-2\sigma^2}$$

Примеры фильтров

- Повышение резкости $\frac{1}{10} \cdot \begin{Vmatrix} -1 & -2 & -1 \\ -2 & 22 & -2 \\ -1 & -2 & -1 \end{Vmatrix}$

- Нахождение границ $\begin{Vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{Vmatrix}$ + модуль, нормировка, применение порога…

- Тиснение $\begin{Vmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{Vmatrix}$ + сдвиг яркости, нормировка…

Двумерные фильтры

- Свойства двумерной свертки (повторение)

  Пусть *X* и *Y* – изображения, *H* – ядро свертки

1. Линейность

$$(const \cdot X) * H = const \cdot (X * H)$$

$$(X + Y) * H = (X * H) + (Y * H)$$

2. Инвариантность к сдвигу

$$(X[i - i_0, j - j_0] * H = (X * H)[i - i_0, j - j_0]$$

Двумерные фильтры

- Сепарабельные (разделимые) фильтры

$$Ker[k,p] = F(k) \cdot G(p)$$

Если фильтр сепарабельный, то фильтрацию можно производить быстрее:

1. Отфильтровать все столбцы одномерным фильтром F(k)
2. Отфильтровать все строки одномерным фильтром G(p)

Гауссиан – сепарабельный фильтр, т.к.

$$Gauss[k,p] = \frac{1}{Sum} \cdot \exp\frac{k^2}{-2\sigma^2} \cdot \exp\frac{p^2}{-2\sigma^2}$$

Еще один сепарабельный фильтр – box-фильтр

Двумерные фильтры

- Unsharp Mask
  - Параметры: радиус, сила эффекта, порог срабатывания
  - Идея: вычесть из изображения его размытую копию, скомпенсировав уменьшение яркости

$$R[i,j]=(1+\alpha)X[i,j]-\alpha GX[i,j]$$

α контролирует силу эффекта,
GX – размытая копия изображения (обычно фильтр Гаусса)

  - Переменная сила эффекта α помогает избежать усиления шума. Обычно α уменьшают при малых значениях разности X – GX (меньше порога срабатывания)

Двумерные фильтры

- Медианный фильтр
  - Каждый пиксель принимает значение, являющееся медианой значений пикселей в окрестности
  - Медиана – средний элемент в отсортированном массиве
  - Позволяет подавить шум (особенно, единичные «выпадающие» пиксели), не размывая границ
  - Медианный фильтр нелинейный (как доказать?)

  - Векторная медиана – такой элемент массива, для которого сумма L1-расстояний до остальных элементов минимальна (для одномерного случая – совпадает с предыдущим определением)

Двумерные фильтры

- Медианный фильтр 5x5

Двумерные фильтры

- Понятие о частотах в изображении и звуке
  - Частоты и гармонические колебания (звук)
  - Частоты и детали (изображение)
  - Постоянная составляющая
  - Действие фильтров
    - Фильтр размытия – НЧ-фильтр
    - Фильтр повышения четкости – ВЧ-фильтр
    - Фильтр нахождения границ – ВЧ-фильтр
    - Фильтры и обработка звука

Преобразование Фурье

- Зачем раскладывать сигналы на синусоиды?
  - Анализ линейных систем
  - Слух и синусоиды
  - Хорошо разработана теория и практика
- Дискретное преобразование Фурье (ДПФ)

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X_k \exp\left(\frac{2\pi i k n}{N}\right)$$

$$\exp i\varphi = \cos\varphi + i\sin\varphi$$

  - Для вещественного сигнала

$$x[n] = \sum_{k=0}^{N/2} C_k \cos\frac{2\pi k(n+\varphi_k)}{N} = \sum_{k=0}^{N/2} A_k \cos\frac{2\pi k n}{N} + \sum_{k=0}^{N/2} B_k \sin\frac{2\pi k n}{N}$$

- Прямое и обратное преобразования Фурье

Преобразование Фурье

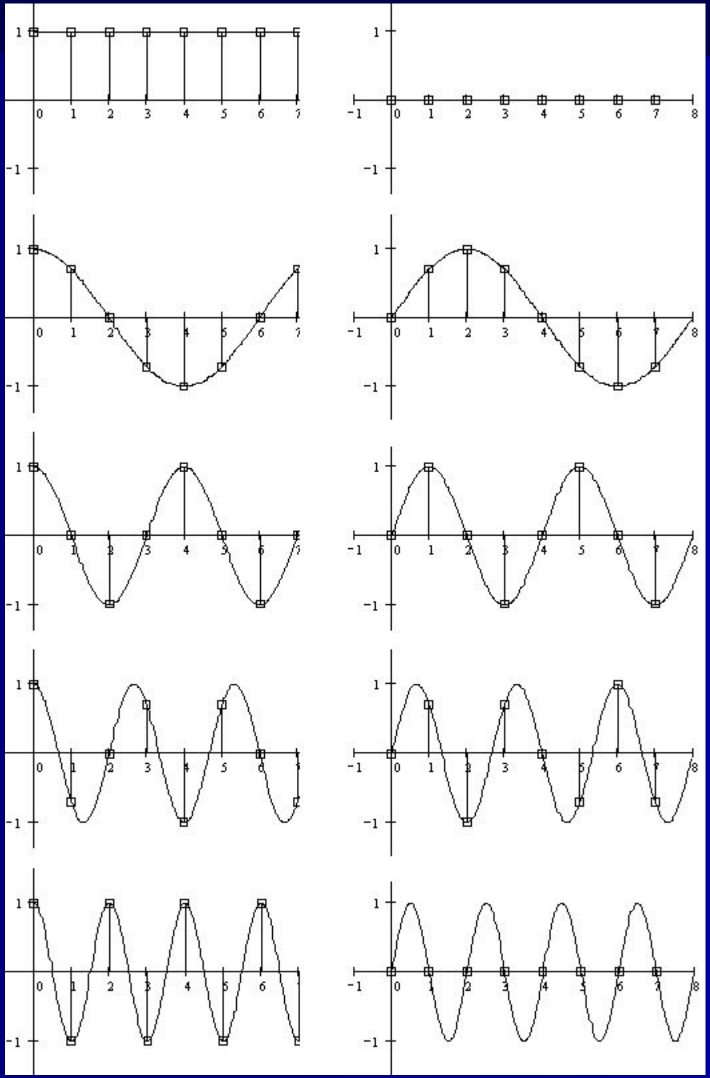
MSU GRAPHICS & MEDIA LAB

- Базисные функции дискретного преобразования Фурье для сигнала длины N = 8.

  Имеем N/2 + 1 = 5 различных базисных частот.

  Имеем N+2 базисные функции, 2 из которых тождественно равны нулю.

  Количество информации не изменяется: N чисел



Преобразование Фурье

- Базисные функции образуют N-мерный ортогональный базис в пространстве N-мерных векторов исходных сигналов.
- Следовательно, разложение обратимо, т.е. по коэффициентам разложения (A_k, B_k) можно точно восстановить исходный дискретный сигнал.
- Обратное преобразование Фурье – вычисление суммы конечного ряда Фурье (сложить N штук N-точечных синусоид со своими коэффициентами).

Преобразование Фурье

- Прямое преобразование Фурье – вычисление скалярных произведений сигнала на базисные функции:

$$A_k = \frac{2}{N}\sum_{i=0}^{N-1} x[i]\cos\frac{2\pi ki}{N} \qquad k = 1,\ldots,\frac{N}{2}-1$$

$$A_k = \frac{1}{N}\sum_{i=0}^{N-1} x[i]\cos\frac{2\pi ki}{N} \qquad k = 0,\frac{N}{2}$$

$$B_k = \frac{2}{N}\sum_{i=0}^{N-1} x[i]\sin\frac{2\pi ki}{N} \qquad k = 0,\ldots,\frac{N}{2}$$

- Для вычисления всех коэффициентов по этому алгоритму требуется примерно N^2 умножений: очень много при больших длинах сигнала N.

Преобразование Фурье

- Быстрое преобразование Фурье (БПФ, FFT) – ускоренный алгоритм вычисления ДПФ
  - ► Основан на периодичности базисных функций (много одинаковых множителей)
  - ► Математически точен (ошибки округления даже меньше, т.к. меньше число операций)
  - ► Число умножений порядка $N \cdot \log_2 N$, намного меньше, чем N^2
  - ► Ограничение: большинство реализаций FFT принимают только массивы длиной $N = 2^m$

- Существует и обратное БПФ (IFFT) – такой же быстрый алгоритм вычисления обратного ДПФ.

Преобразование Фурье

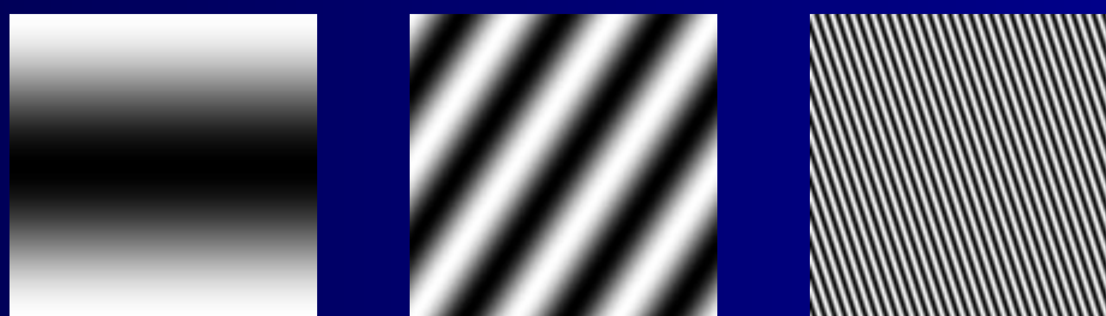
- Входные данные FFT
  - $N = 2^m$, размер FFT
  - Входной вектор длины N, иногда в комплексном представлении
- Выходные данные FFT
  - Коэффициенты A_k и B_k, иногда записанные в комплексном представлении

$$A_k + iB_k$$

Преобразование Фурье

- Двумерное ДПФ
  - Базисные функции имеют вид двумерных синусоид с разными углами наклона и фазами

$$\sin\left(\frac{2\pi ni}{N} + \frac{2\pi mj}{M}\right)$$



- Вычисление двумерного ДПФ

1. Прямой способ – скалярные произведения со всеми базисными функциями. Очень много операций.
2. Быстрый способ – декомпозиция на одномерные ДПФ

Преобразование Фурье

- Быстрое вычисление двумерного ДПФ

1. Вычислить одномерные комплексные ДПФ от каждой строки изображения. Результаты записать в виде комплексных массивов «обратно» в промежуточное «комплексное» изображение.
2. Вычислить одномерные комплексные ДПФ от каждого столбца промежуточного комплексного изображения. Комплексные результаты записать «обратно». Это и есть коэффициенты двумерного ДПФ.

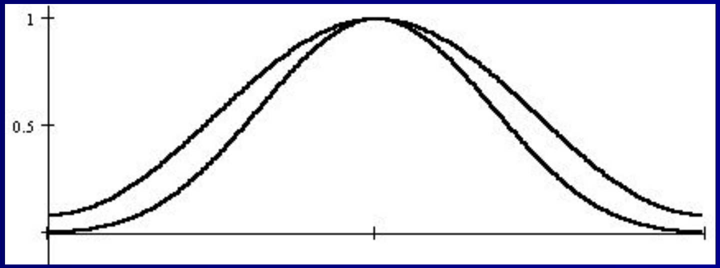
Одномерные ДПФ можно считать с помощью FFT.

MSU GRAPHICS & MEDIA LAB

Спектральный анализ

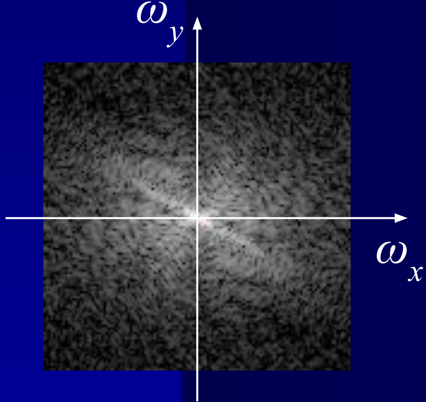
- Как вычислить и отобразить спектр сигнала?
  1. Взять нужный отрезок сигнала длины 2^m; если нужный отрезок короче – дополнить его нулями.
  2. Если нужно – домножить сигнал на весовое окно, плавно спадающее к краям (для уменьшения размытия спектра).
  3. Вычислить FFT.
  4. Перевести комплексные коэффициенты в полярную форму: получить амплитуды.
  5. Отобразить график зависимости амплитуды от частоты.

Примеры весовых окон



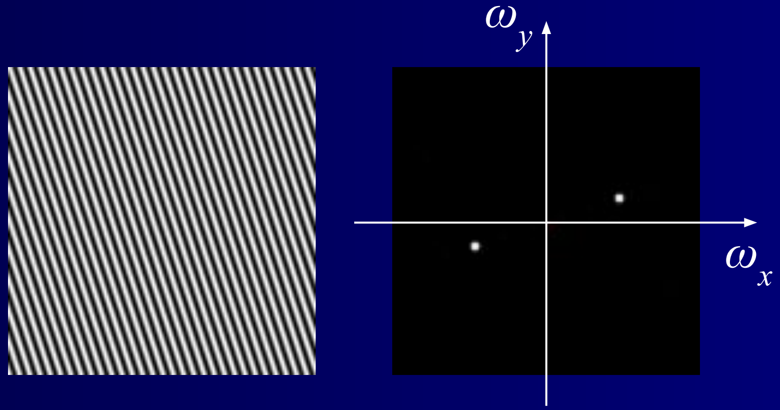
Спектральный анализ

- Отображение спектров изображений
  - Спектр – это картинка, показывающая зависимость амплитуды от частоты и от направления синусоиды.
  - Амплитуды отображаются в виде яркостей.
  - Нулевая частота – в центре спектра, низкие частоты вокруг центра, высокие – дальше от центра.
  - Спектр обычно продублирован отражением от нулевой частоты.
  - В реальных изображениях чаще всего гораздо большие амплитуды имеют низкие частоты (и постоянная составляющая). Поэтому постоянную составляющую иногда удаляют, или применяют логарифмический масштаб отображения амплитуд, чтобы пара самый мощных гармоник не скрыла остальные, менее мощные, но тоже существенные гармоники.

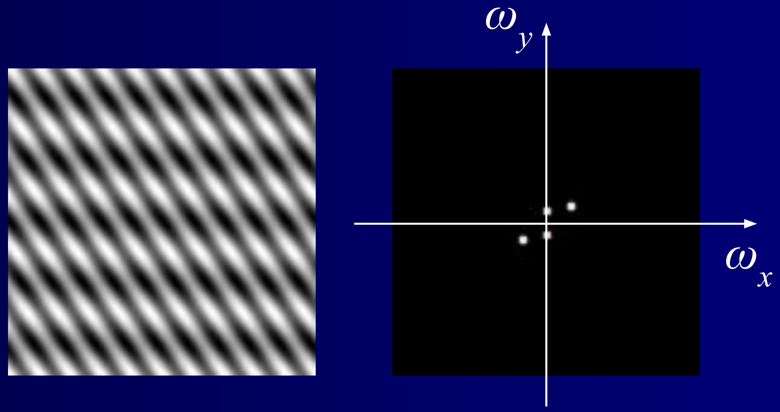


Спектральный анализ

- Примеры изображений и их спектров



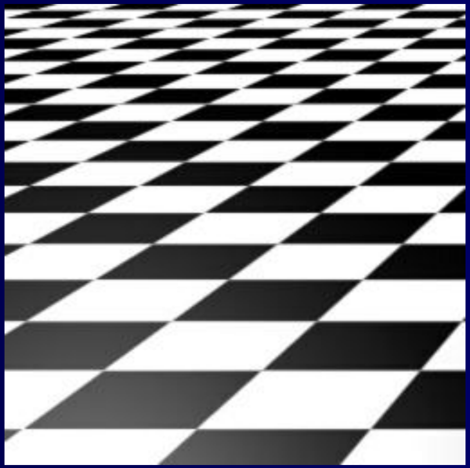
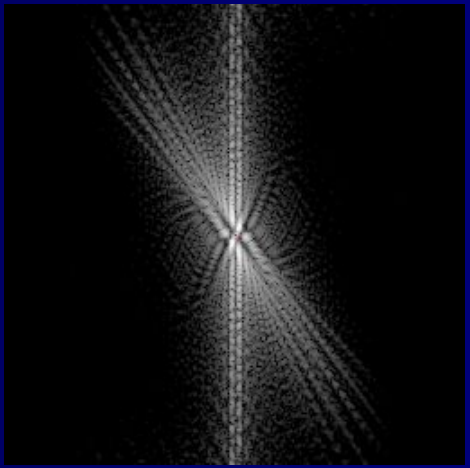
Видно, что спектр одной синусоиды – это точка
(не забываем про симметричное отражение спектра)



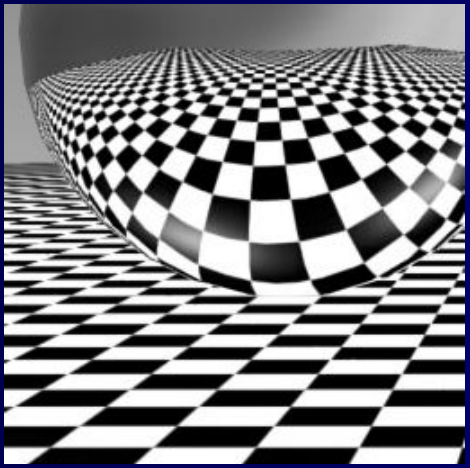
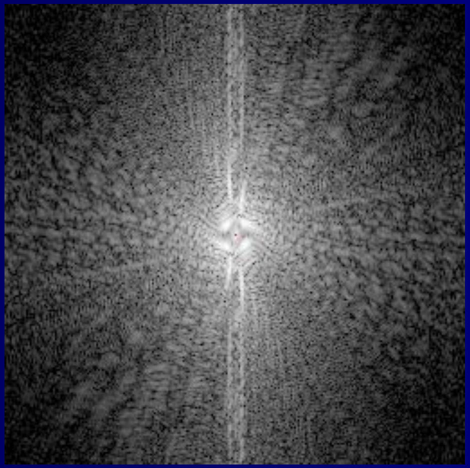
Две синусоиды – две точки

Спектральный анализ

- Примеры изображений и их спектров

По спектру прослеживаются преобладающие направления в исходной картинке

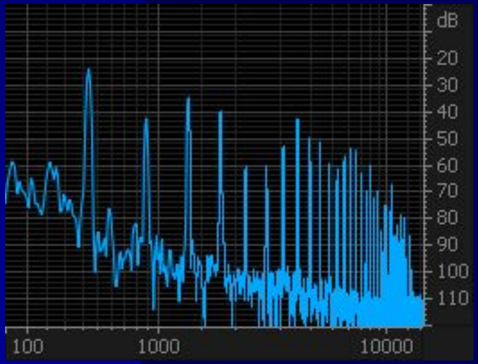
Много высоких частот в спектре – много мелких деталей в исходном изображении

Спектральный анализ

- Отображение спектра звука: спектр
  - Спектр – график зависимости амплитуды от частоты
  - Низкие частоты – слева, высокие – справа
  - Часто применяется логарифмический масштаб частот и амплитуд: “log-log-спектр”
  - Временное и частотное разрешение спектра

Децибелы: $$D = 20\lg\frac{A_1}{A_0}$$

A_1 – амплитуда измеряемого сигнала, A_0 – амплитуда сигнала, принятого за начало отсчета (0 дБ)

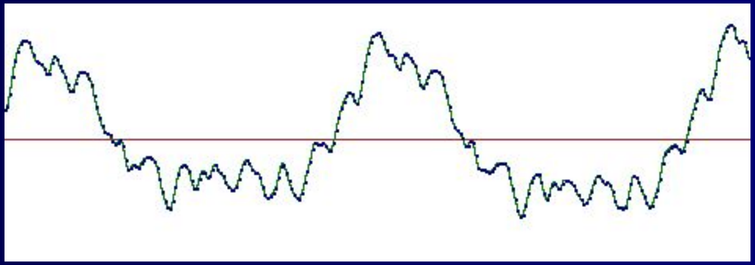


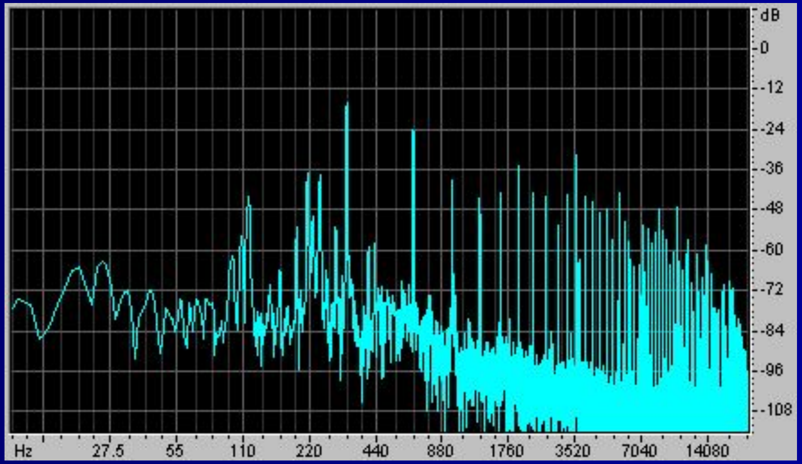
Разница на 6 дБ – разница по амплитуде в 2 раза,
разница на 12 дБ – разница по амплитуде в 4 раза.

Часто за 0 дБ принимается либо самый тихий слышимый звук, либо самый громкий звук, который может воспроизвести аудио-устройство.

Спектральный анализ

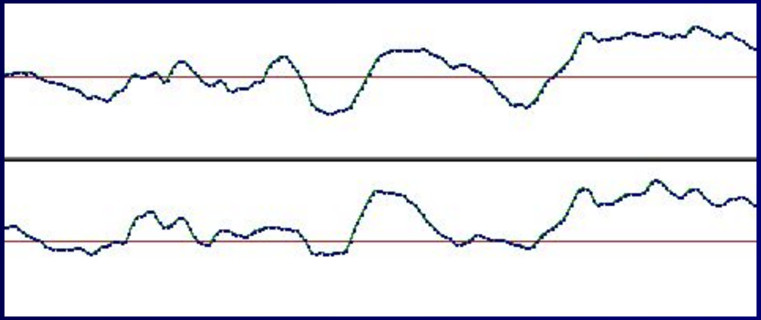
- Примеры звуков и их спектров

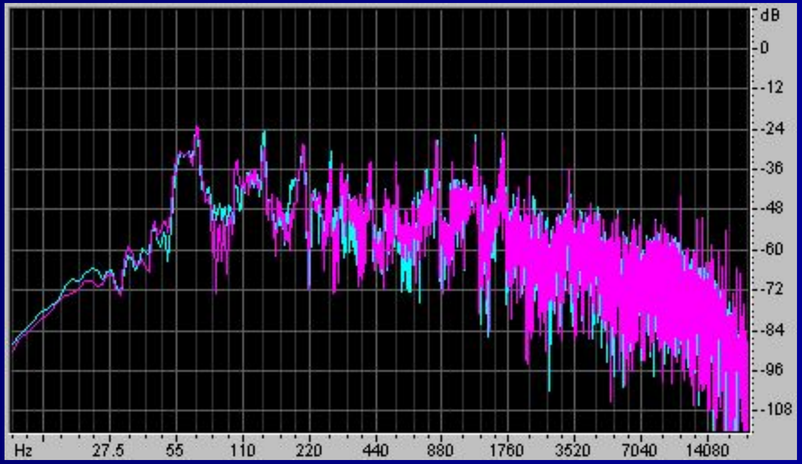




Нота на гитаре

сигнал близок к периодическому → его спектр линейчатый



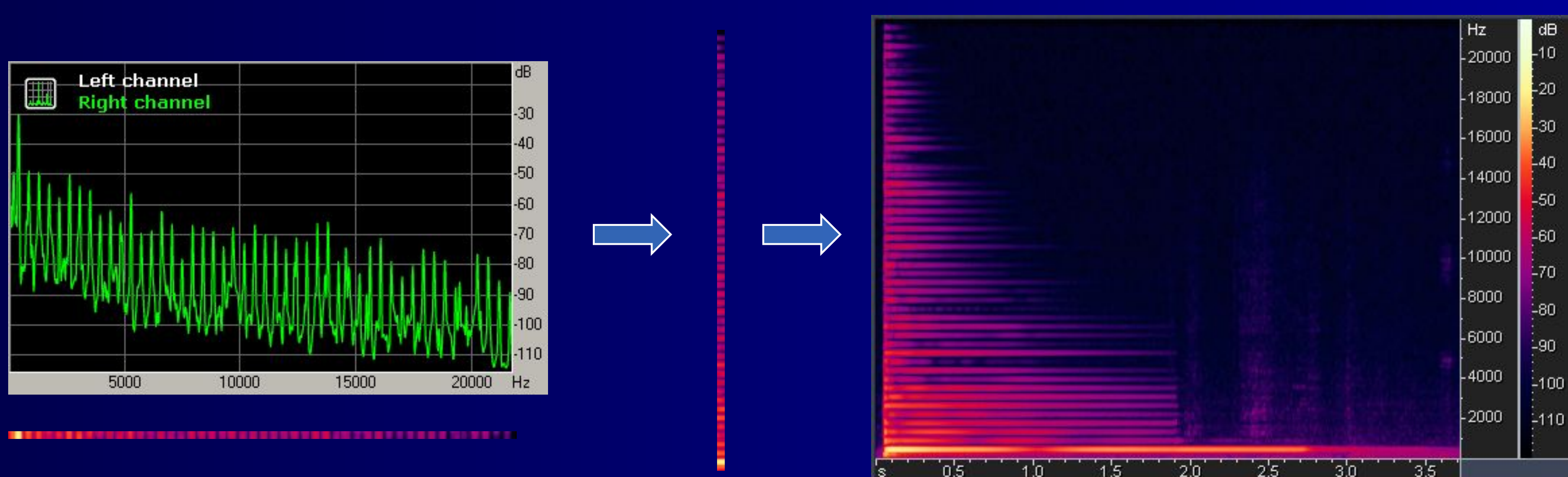


Фрагмент песни (стерео запись)

Спектральный анализ

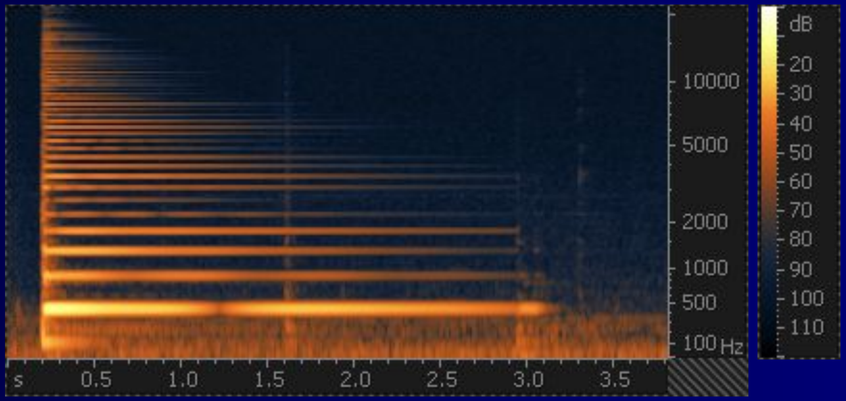
- Отображение спектра звука: спектрограмма (сонограмма)
  - Спектрограмма – график зависимости амплитуды от частоты и от времени, показывает изменение спектра во времени
  - Short Time Fourier Transform (STFT)

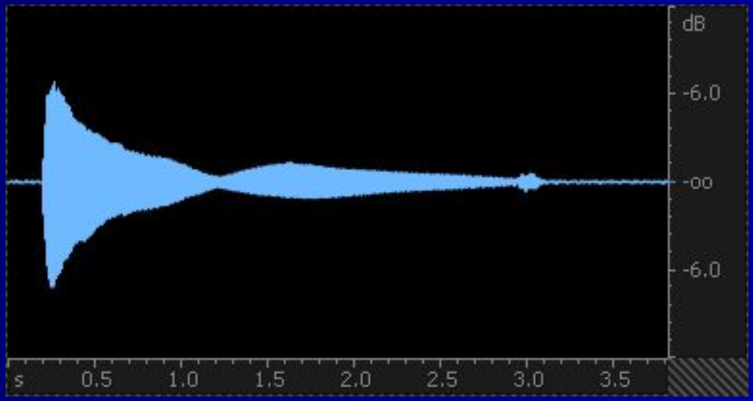
$$STFT[n,\omega] = \sum_{m=-\infty}^{+\infty} x[n+m] \cdot w[m] \cdot e^{-i\omega m}$$



Спектральный анализ

- Отображение спектра звука: спектрограмма (сонограмма)
  - Спектрограмма – график зависимости амплитуды от частоты и от времени, показывает изменение спектра во времени
  - Низкие частоты – снизу, высокие – сверху
  - Время идет справа налево
  - Амплитуда – яркость или цвет
  - Частотное и временное разрешение
  - Short Time Fourier Transform (STFT)

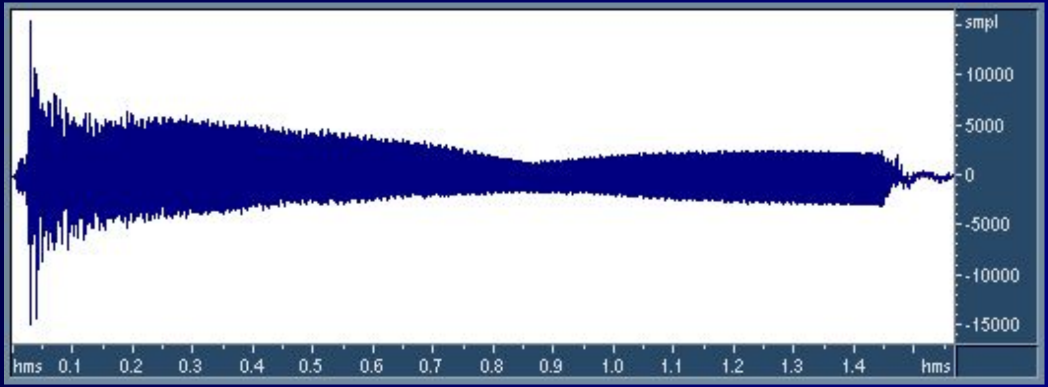


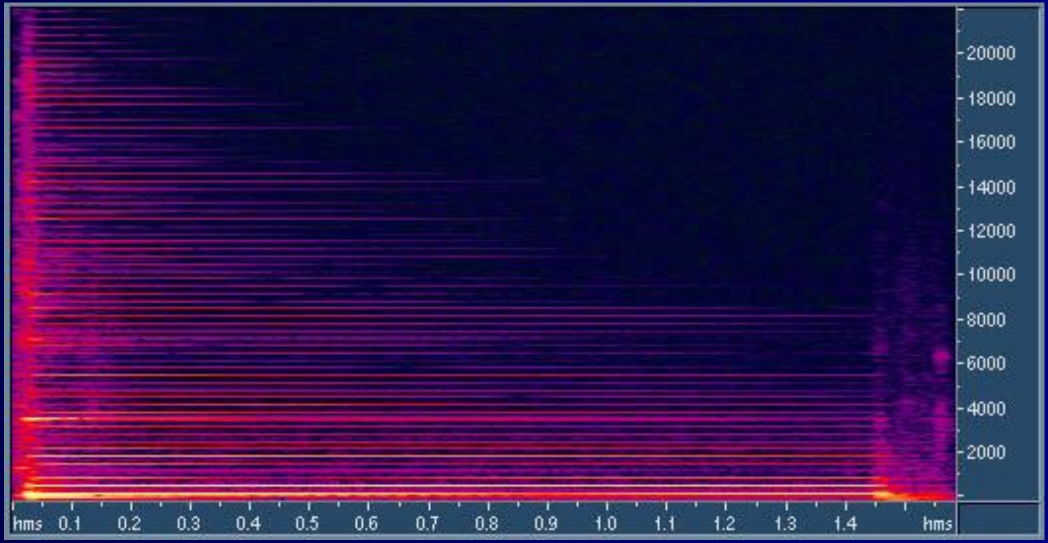


Показывает изменение спектра во времени

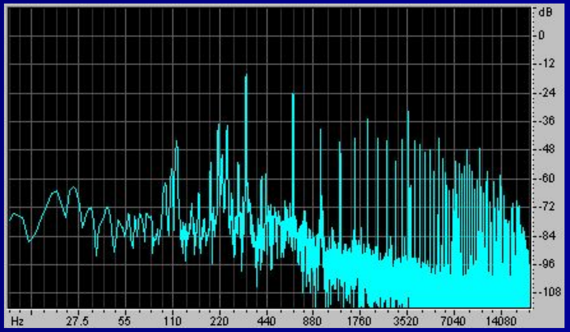
Спектральный анализ

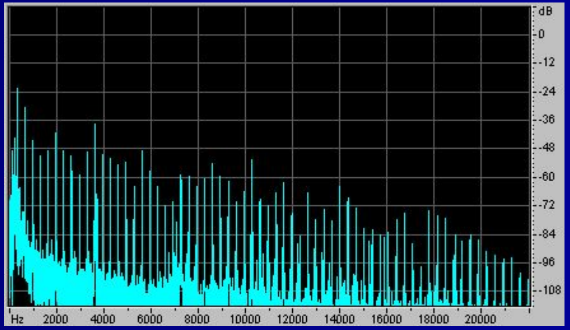
- Примеры звуков и их спектрограмм





логарифмический масштаб частот





линейный масштаб частот

Нота на гитаре