

Использование преобразования Фурье в обработке изображений и компьютерном зрении

Сергей Колмаков

МФТИ

10.2017

План лекции

- ▶ Алгоритм JPEG
 - ▶ Дискретное косинусное преобразование
 - ▶ Квантование
 - ▶ Кодирование Хаффмана
 - ▶ Цветовая модель YCbCr
- ▶ Преобразование Фурье
 - ▶ Непрерывное и дискретное преобразования
 - ▶ Свойства
 - ▶ Теорема о свёртке
 - ▶ Быстрое преобразование Фурье
 - ▶ Примеры применения
 - ▶ Поиск периодического сигнала на изображении
 - ▶ Оценка резкости изображения

Алгоритм JPEG

Источники

<https://habrahabr.ru/post/206264>

http://sernam.ru/cod_15.php

Алгоритм JPEG

JPEG - Joint Photographic Expert Group

Основная идея

- ▶ Разбить изображение на блоки 8x8 пикселей
- ▶ Рассматривать блоки как векторы в 64-мерном пространстве.
- ▶ Произвести замену базиса так, чтобы наиболее важная информация находилась в первых компонентах векторов
- ▶ Кодировать первые компоненты полученных векторов с меньшими потерями, последние - с большими

Пример: блоки 2x1 вместо 8x8



Рис. 1: Исходное изображение

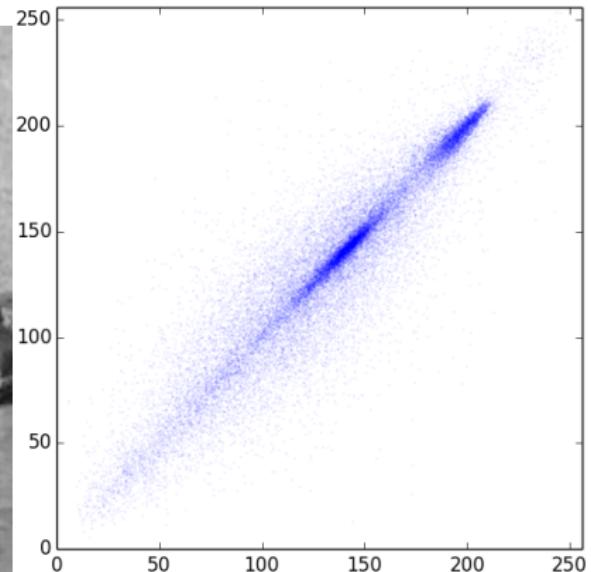


Рис. 2: Корреляция значений соседних пикселей

Новый базис: $e_0 = \frac{(1,1)}{\sqrt{2}}$ $e_1 = \frac{(1,-1)}{\sqrt{2}}$

Напоминание

Базис: $\vec{e}_0 = (e_{00}, e_{01})$ $\vec{e}_1 = (e_{10}, e_{11})$

Рассматриваемый вектор $\vec{p} = (x, y) = a_0 \vec{e}_0 + a_1 \vec{e}_1$

Чтобы найти a_0 , спроектируем \vec{p} на \vec{e}_0 .

$$\langle p, e_0 \rangle = \langle a_0 e_0 + a_1 e_1, e_0 \rangle = a_0 \langle e_0, e_0 \rangle + a_1 \langle e_1, e_0 \rangle = a_0 \|e_0\| = a_0$$

аналогично

$$\langle p, e_1 \rangle = a_1$$

$$\text{В матричной форме } E = \begin{pmatrix} e_{00} & e_{01} \\ e_{10} & e_{11} \end{pmatrix} X = \begin{pmatrix} x \\ y \end{pmatrix} A = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$$

$$A = EX, \quad X = E^T A$$

Пример

Блоки 2x1. Оставлена половина значений.

Оригинал.



Рис. 3: Сжатие в 2 раза для блоков 2x1

Discrete Cosine Transform

Увеличим площадь блока до N пикселей. Как выбрать базис?

Хотим:

- ▶ Ортогональность
- ▶ "Уплотнение энергии": сохранение основной информации в первых компонентах вектора

Дискретизуем с частотой N набор ортогональных функций

$f_n(x) = \cos(nx)$, $n = 0..N - 1$. Получили ортогональный базис.

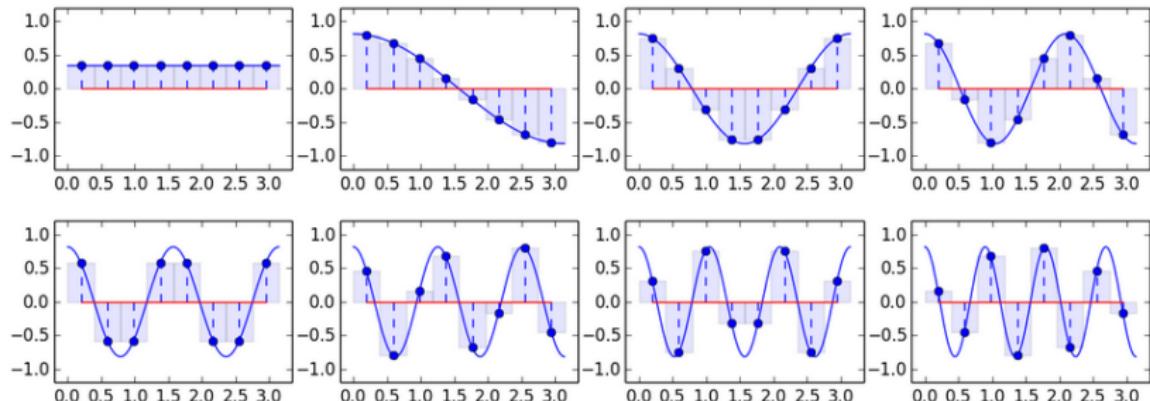


Рис. 4: Набор базисных векторов DCT в 8-мерном пространстве

Discrete Cosine Transform

Выразим произвольный вектор $\vec{X} = (x_0..x_{N-1})$ в новом базисе.
Как и прежде, считаем скалярное произведение:

$$a_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}(n + 0.5)k\right], \quad k = 0..N - 1 \quad (1)$$

или, в матричной форме, $A = EX$

Получили формулу дискретного косинусного преобразования (DCT)

DCT для двумерного случая

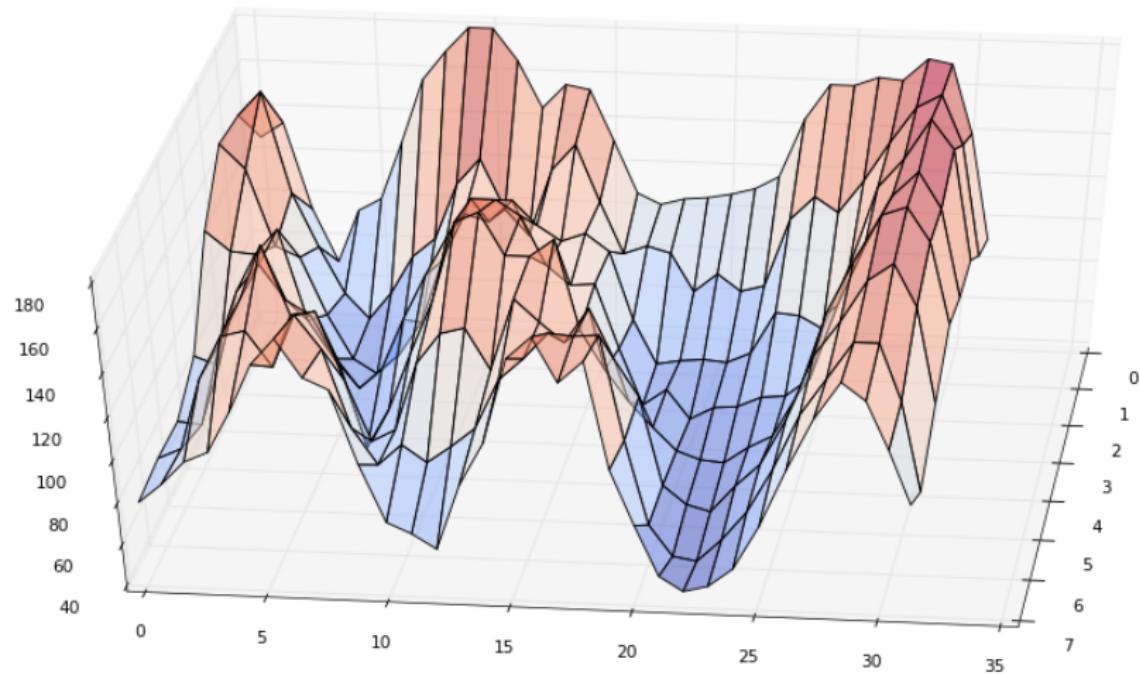


Рис. 5: Фрагмент исходного изображения

DCT для двумерного случая

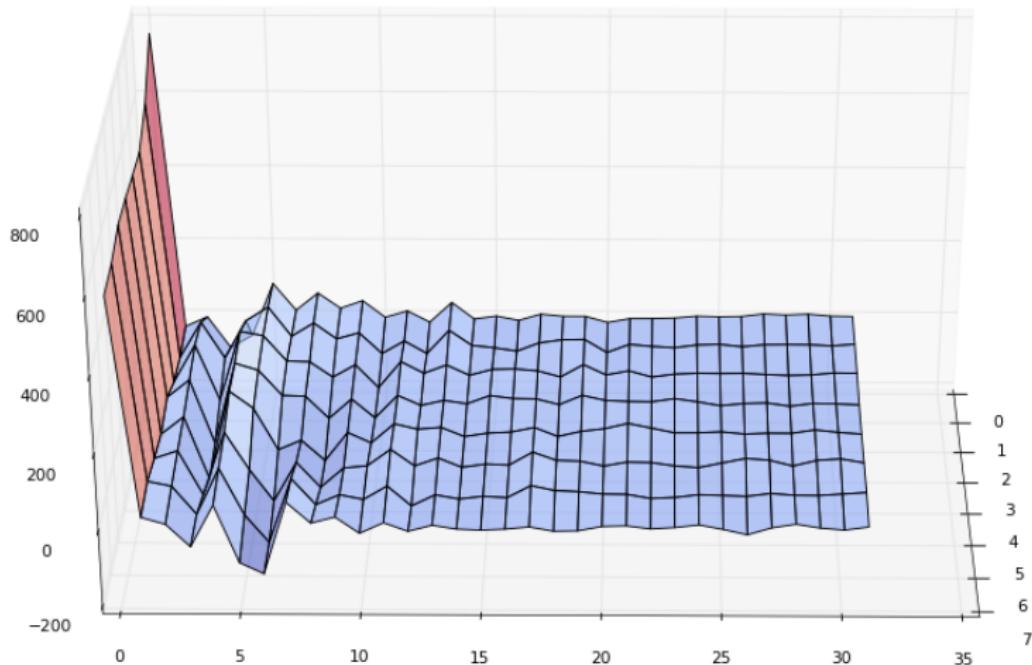


Рис. 6: После применения DCT построчно

DCT для двумерного случая

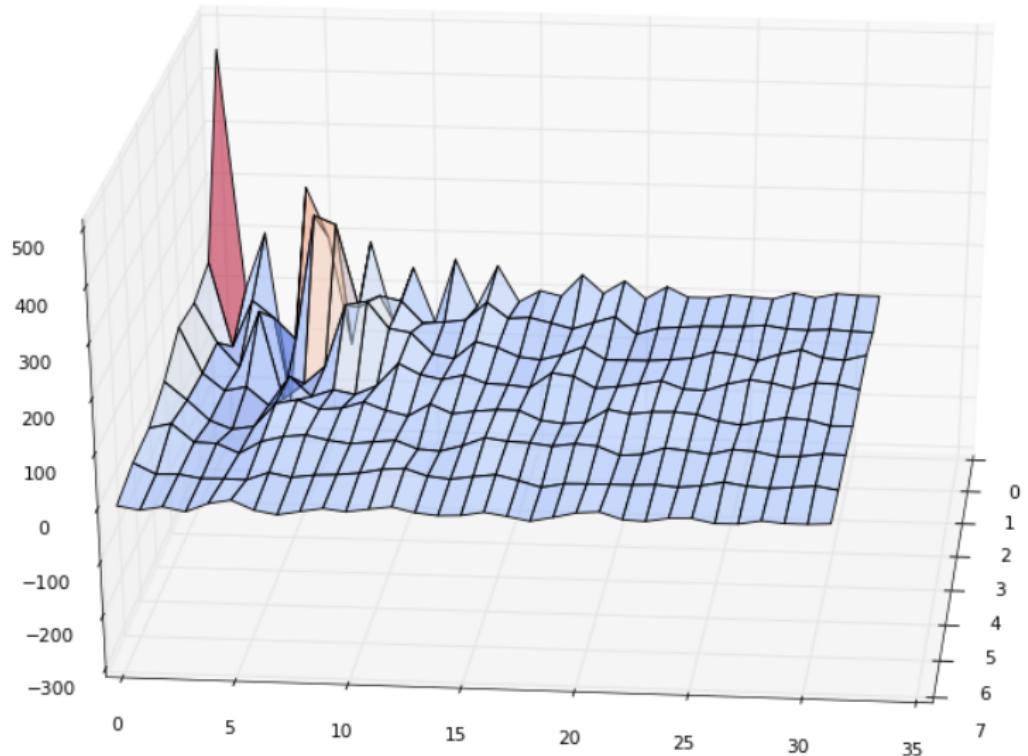
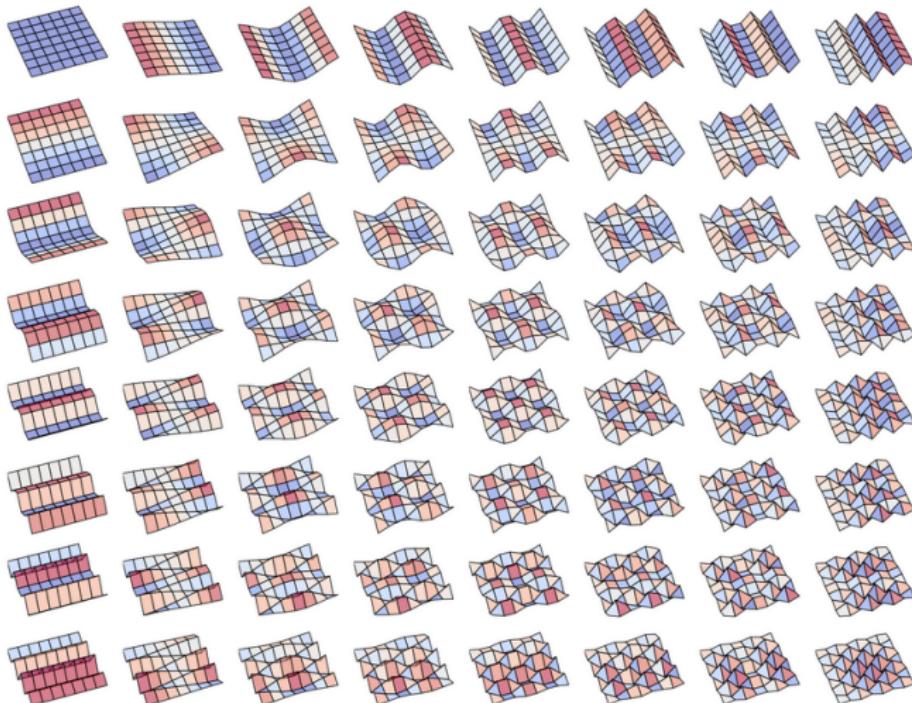


Рис. 7: После применения DCT к получившимся столбцам

DCT для двумерного случая

$$A_{kl,k_2} = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} X_{n_1,n_2} \cos\left[\frac{\pi}{N_1}\left(n_1 + \frac{I}{2}\right)k_1\right] \cos\left[\frac{\pi}{N_2}\left(n_2 + \frac{I}{2}\right)k_2\right]$$

Рис. 8: Вычисление коэффициентов для двумерного DCT



Квантование

1104	-6	-24	15	-14	0	0	0
-108	-8	4	0	8	0	0	0
8	16	20	7	1	0	0	0
24	10	0	9	0	1	0	0
10	35	0	0	0	0	0	0
7	-11	17	0	24	0	0	1
15	0	0	0	0	0	0	0
0	0	0	1	0	1	0	0

÷

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

=

69	-1	-2	1	-1	0	0	0
-9	-1	0	0	0	0	0	0
1	1	1	0	0	0	0	0
2	1	0	0	0	0	0	0
1	2	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Становится возможным кодирование длин серий (run-length encoding, RLE) для нулей

Схемы упорядочения

Alternating Current - AC, Direct Current - DC

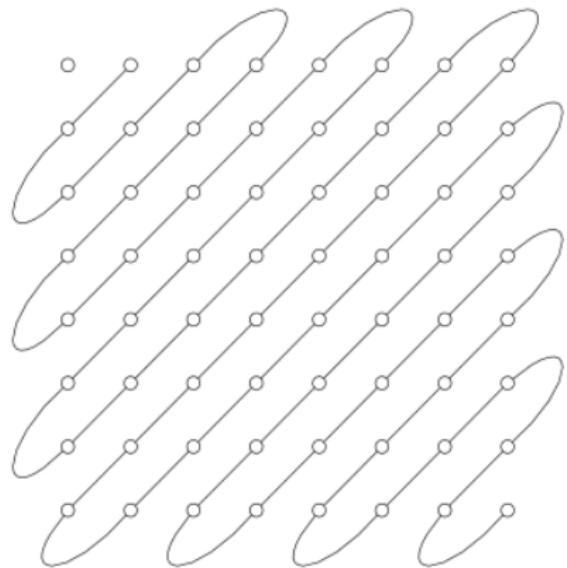


Рис. 9: Для АС-коэффициентов
внутри блока (Zig-zag)

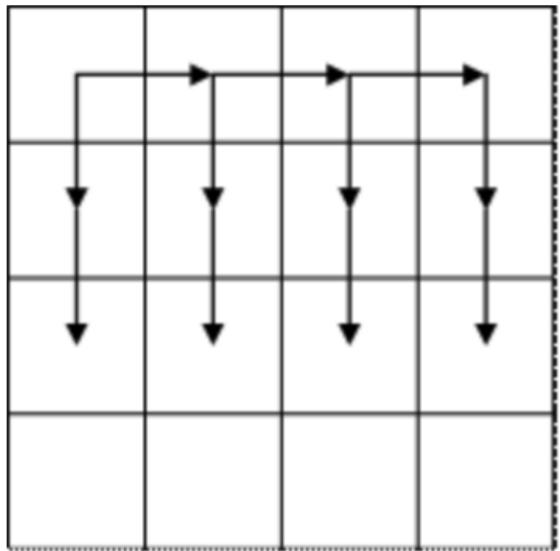


Рис. 10: Для DC-коэффициентов
между блоками

Кодирование Хаффмана

- ▶ Больше частота - меньше длина кодирующей последовательности
- ▶ Свойство префиксности

Алгоритм:

- 1) Все символы сортируются в порядке убывания весов.
- 2) Два последних символа объединяются в группу. Ей присваивается вес, равный сумме весов этих элементов. Эта группа участвует в алгоритме наравне с символами и другими группами.

Шаги повторяются, пока не останется только одна группа. В каждой группе одному символу (или подгруппе) присваивается бит 0, а другому бит 1.

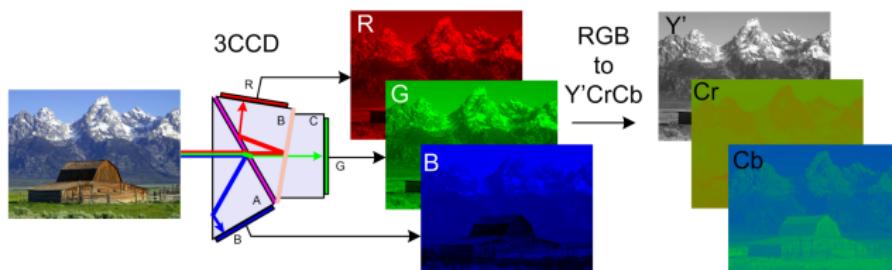
RGBtoYCbCr

Y - средневзвешенная яркость

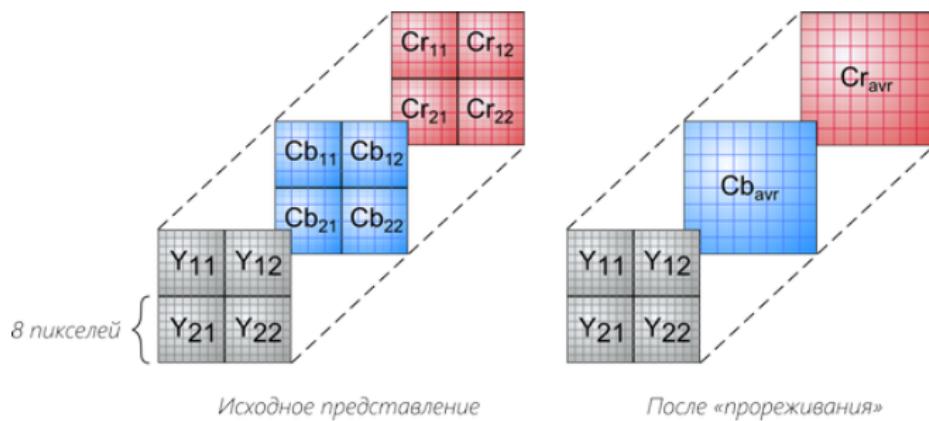
Cb - синяя цветоразностная компонента

Cr - красная цветоразностная компонента

$$\begin{vmatrix} Y \\ Cb \\ Cr \end{vmatrix} = \begin{vmatrix} 0.299 & 0.587 & 0.114 \\ 0.1687 & -0.3313 & 0.5 \\ 0.5 & -0.4187 & -0.0813 \end{vmatrix} \begin{vmatrix} R \\ G \\ B \end{vmatrix} + \begin{vmatrix} 0 \\ 128 \\ 128 \end{vmatrix} \quad (2)$$



Прореживание каналов Cb и Cr



Получившийся алгоритм

1. RGB to YCbCr
2. Прореживание Cb, Cr. Разбиение на блоки 8x8
3. Применение DCT к блокам
4. Квантование
5. Применение кодирования Хаффмана отдельно для DC и AC коэффициентов

Преобразование Фурье

Источники

«Основы цифровой обработки сигналов» - Алексей Лукин

«Алгоритм детектирования и определения фазы решетки
периодических защитных элементов изображения документа» -
Тимофей Чернов, Сергей Колмаков, Дмитрий Николаев

Преобразование Фурье

Дана функция вещественной переменной. Хотим узнать:

- ▶ Наличие периодичности
- ▶ Какие частоты преобладают
- ▶ Фазовый сдвиг

Попробуем использовать косинусное преобразование:

$$\cos(x) = 0 \cdot \cos(0) + 1 \cdot \cos(x) + 0 \cdot \cos(2x) + \dots$$

$$\begin{aligned}\cos(x - \Pi/2) &= \frac{2}{\Pi} \cdot \cos(0) - \frac{4}{3\Pi} \cos(x) - \dots - \frac{4}{\Pi(4k^2-1)} \cdot \cos(2kx) - \dots \\ &= \sin(x)\end{aligned}$$

Преобразование Фурье

Прямое непрерывное

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \omega} dx \quad (3)$$

Обратное непрерывное

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\omega) e^{2\pi i x \omega} d\omega \quad (4)$$

Прямое дискретное

$$X_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-2\pi i \frac{kn}{N}} \quad (5)$$

Обратное дискретное

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k e^{2\pi i \frac{kn}{N}} \quad (6)$$

Преобразование Фурье

Прямое многомерное дискретное

$$Fx(K_1, K_2, \dots, K_n) = \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_m=0}^{N_m-1} fx(n_1, n_2, \dots, n_N) e^{-i \frac{2\pi}{N_1} n_1 K_1 - i \frac{2\pi}{N_2} n_2 K_2 - \cdots - i \frac{2\pi}{N_m} n_m K_m}$$

Обратное многомерное дискретное

$$fx(n_1, n_2, \dots, n_m) = \frac{1}{N_1 \cdots N_m} \sum_{K_1=0}^{N_1-1} \cdots \sum_{K_m=0}^{N_m-1} Fx(K_1, K_2, \dots, K_m) e^{i \frac{2\pi}{N_1} n_1 K_1 + i \frac{2\pi}{N_2} n_2 K_2 + \cdots + i \frac{2\pi}{N_m} n_m K_m}$$

Свойства преобразования Фурье

1. $\widehat{\alpha f + \beta g} = \alpha \hat{f} + \beta \hat{g}$
2. $\widehat{f(x-a)}(\omega) = e^{-2\pi i a \omega} \hat{f}(\omega)$
3. $\widehat{f(ax)}(\omega) = \frac{1}{|a|} \hat{f}\left(\frac{\omega}{a}\right)$
4. Теорема свертки
 $(\widehat{f * g})(\omega) = \hat{f}(\omega) \hat{g}(\omega)$
 $\widehat{(fg)}(\omega) = (\hat{f} * \hat{g})(\omega),$
где $(f * g)(x) = \int_{-\infty}^{\infty} f(y)g(x-y)dy$
5. $\widehat{\overline{f(x)}} = \overline{\hat{f}(-\omega)}$
6. $\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |f(x)|^2 dx$

Пример

$$f(x) = 0, x < 0$$

$$f(x) = e^{-x}, x \geq 0$$

$$\hat{f}(\omega) = ?$$

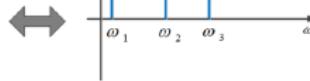
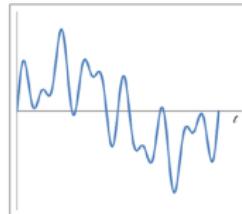
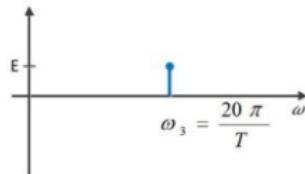
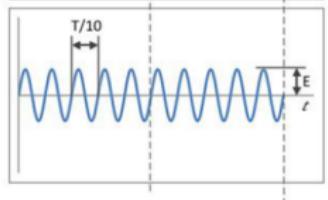
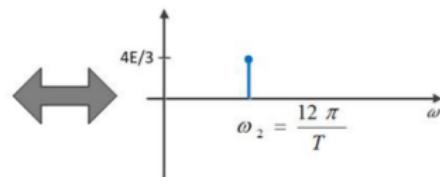
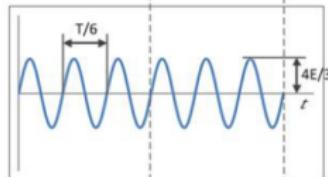
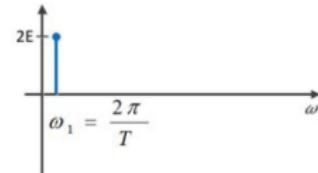
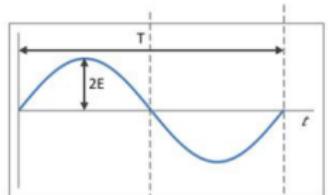
Дельта-функция

- $\delta(x) = \begin{cases} +\infty, & x = 0, \\ 0, & x \neq 0; \end{cases}$
- $\int_{-\infty}^{+\infty} \delta(x) dx = 1.$

$$(\delta(x) * f(x))(x) = \int_{-\infty}^{\infty} \delta(y)f(x - y)dy = f(x)$$

$$(\delta(x - a) * f(x))(x) = \int_{-\infty}^{\infty} \delta(y - a)f(x - y)dy = f(x - a)$$

Пример



FFT

Сложность тривиальной реализации - $O(N^2)$

Сложность FFT - $O(N \log(N))$

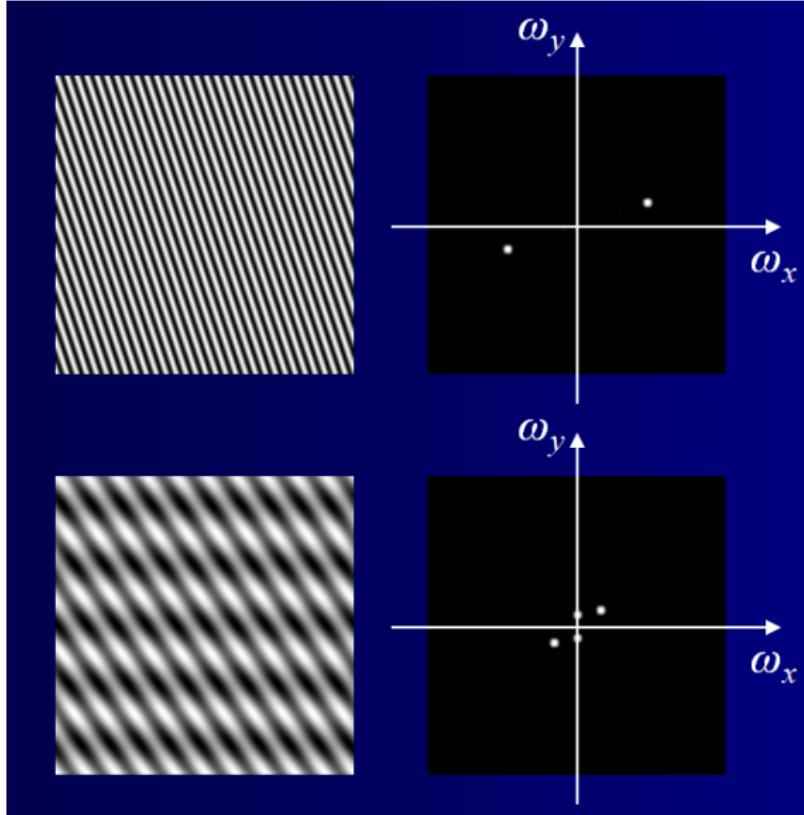
Алгоритм Кули и Таки(Cooley-Tukey)

The Fastest Fourier Transform in the West (FFTW)

C++: FFTW

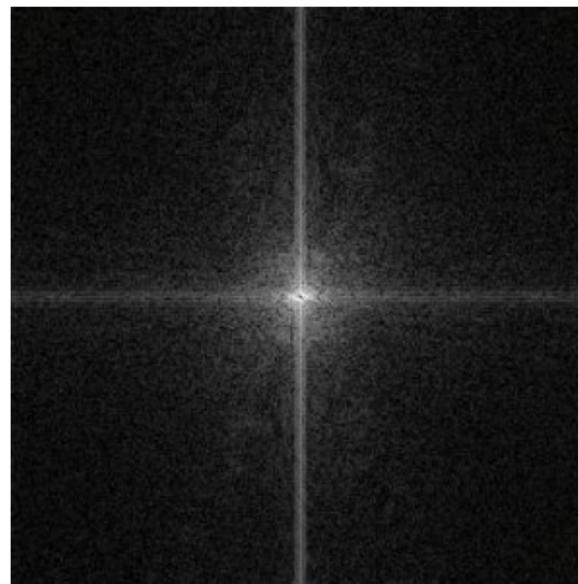
Python: pyFFTW

Пример 2d



Примеры применения FFT

Оценка резкости изображения



Примеры применения FFT. Задача классификации

Поиск периодического сигнала. Сканы с голограммой



Задача классификации

Скан без голограммы



Математическая модель (одномерный случай)

$I(x)$ - изображение, $h(x)$ - фоновое изображение, $g(x)$ - периодический шаблон. $I(x) = h(x) + g(x)$

Пусть N - длина сигнала. M - число шаблонов. Тогда период $T = N/M$. Если $f(x)$ - сигнал единичного периодического элемента, то

$$g(x) = \sum_{m=0}^{M-1} f(x - mT). \quad (7)$$

Пусть

$$c(x) = \sum_{m=0}^{M-1} \delta(x - mT). \quad (8)$$

Периодический сигнал $g(x)$ - свёртка $f(x)$ и $c(x)$:

$$g(x) = f(x) * c(x). \quad (9)$$

Тогда:

$$I(x) = h(x) + f(x) * c(x). \quad (10)$$

Детектирование – оценка мощности $f(x)$, а локализация – как оценка фазы $c(x)$.

Математическая модель (одномерный случай)

Преобразование Фурье импульсного сигнала $c(x)$ также представляет собой импульсный сигнал с периодом M :

$$\hat{c}(k) = \sum_{m=0}^{M-1} e^{-2\pi i k \frac{m}{M}} = M \cdot \sum_m \delta(k - mM). \quad (11)$$

Например, на рисунке 11 приведено изображение (б) амплитуды импульсного сигнала (а) с $T = 32$ и $N = 256$, который вписывается 8 раз, следовательно, его ДПФ имеет период, равный 8.

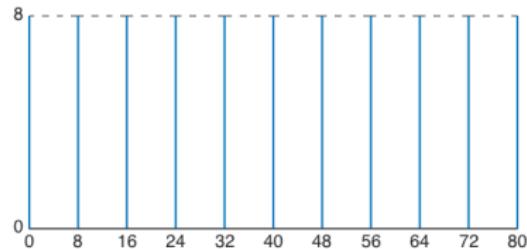
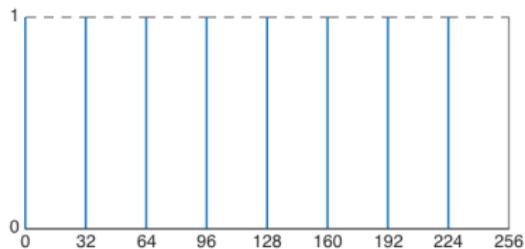


Рис. 11: Импульсный сигнал и его ДПФ амплитуда

Математическая модель (одномерный случай)

Из уравнений для c и g и теоремы о свертке получается окончательное выражение для $\hat{g}(k)$:

$$\hat{g}(k) = \begin{cases} M \cdot \hat{f}(k), & \text{если } k = 0, M, 2M, \dots \\ 0, & \text{иначе.} \end{cases} \quad (12)$$

Работа с модельным изображением: ДПФ решётки гауссиан

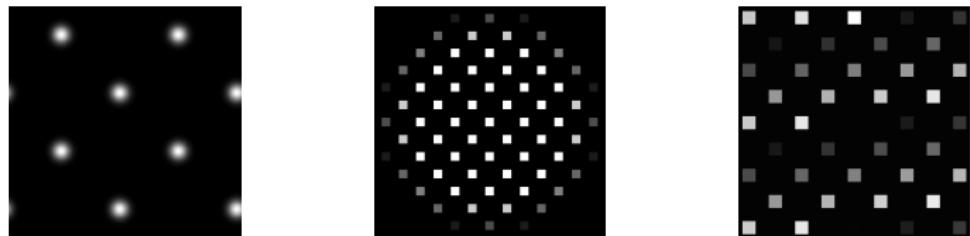


Рис. 12: Решетка Гауссиан, подобная шахматной доске (а), ее ДПФ амплитуда (б) и фаза (в)

Работа с реальным изображением

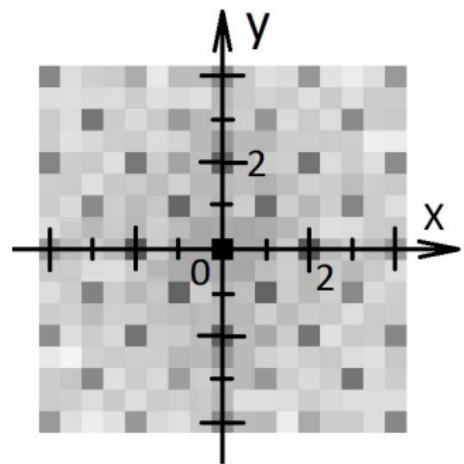


Рис. 13: Вырезанный, предобработанный 2×2 регион паспорта РФ и амплитуда его ДПФ

Локализация голограмм

Фазовый угол $\phi_{i,j}$ на (i,j) пике для двумерного импульсного сигнала равен:

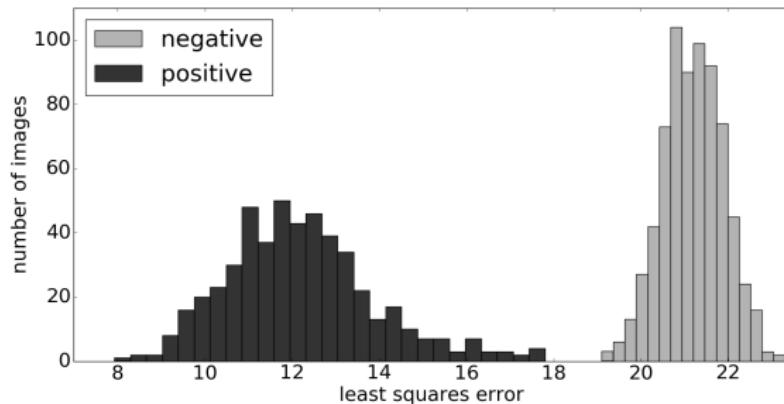
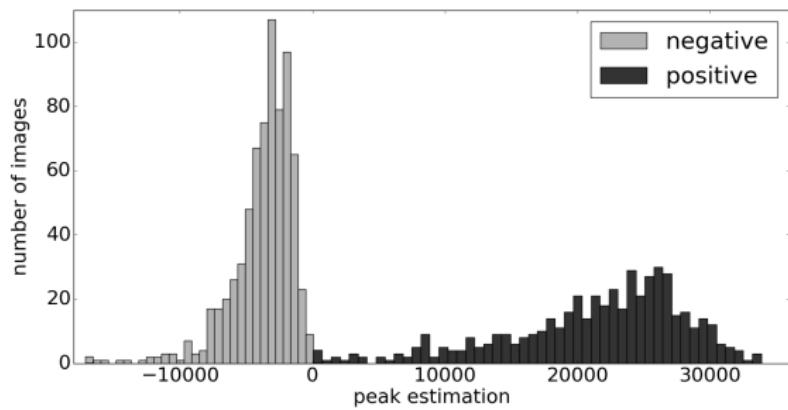
$$\phi_{i,j} = i \cdot \phi_x + j \cdot \phi_y. \quad (13)$$

Решаем переопределённую систему, в которую включены уравнения для всех пиков, находим ϕ_x , ϕ_y . Вычисляем линейное смещение, зная фазовое



Рис. 14: Результат локализации

Разделяющая способность в задаче классификации



Замещение искаженных областей



Замещение искаженных областей

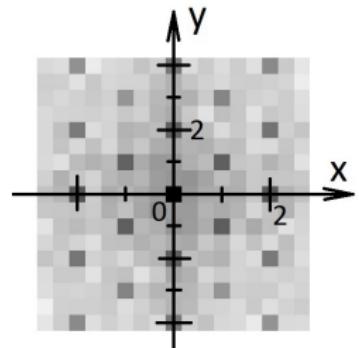
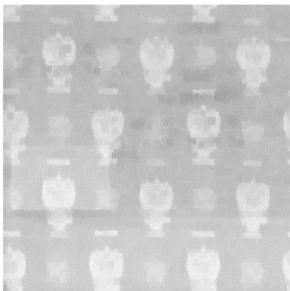
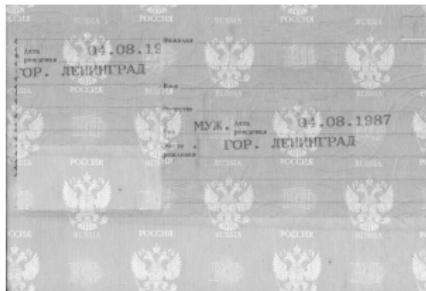


Рис. 15: Вырезанный, предобработанный 3×2 регион паспорта РФ и амплитуда его ДПФ