



Кафедра молекулярных процессов и экстремальных
состояний вещества

Математические основы методов анализа результатов физического эксперимента

8. Методы анализа цифровых изображений.

Коротеева Екатерина Юрьевна, ст. преп.

Спектральный анализ

Спектральный анализ

2D DFT для изображения размером $N_1 \times N_2$

Фурье-анализ:

$$X[k_1, k_2] = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x[n_1, n_2] e^{-j\frac{2\pi}{N_1}n_1k_1} e^{-j\frac{2\pi}{N_2}n_2k_2}$$

Фурье-синтез:

$$x[n_1, n_2] = \frac{1}{N_1 N_2} \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} X[k_1, k_2] e^{j\frac{2\pi}{N_1}n_1k_1} e^{j\frac{2\pi}{N_2}n_2k_2}$$

Спектральный анализ

В пространстве изображения размером $N_1 \times N_2$ существует $N_1 N_2$ ортогональных базисных векторов

$$w_{k_1, k_2}[n_1, n_2] = e^{j \frac{2\pi}{N_1} n_1 k_1} e^{j \frac{2\pi}{N_2} n_2 k_2}$$

при:

$$n_1, k_1 = 0, 1, \dots, N_1 - 1$$

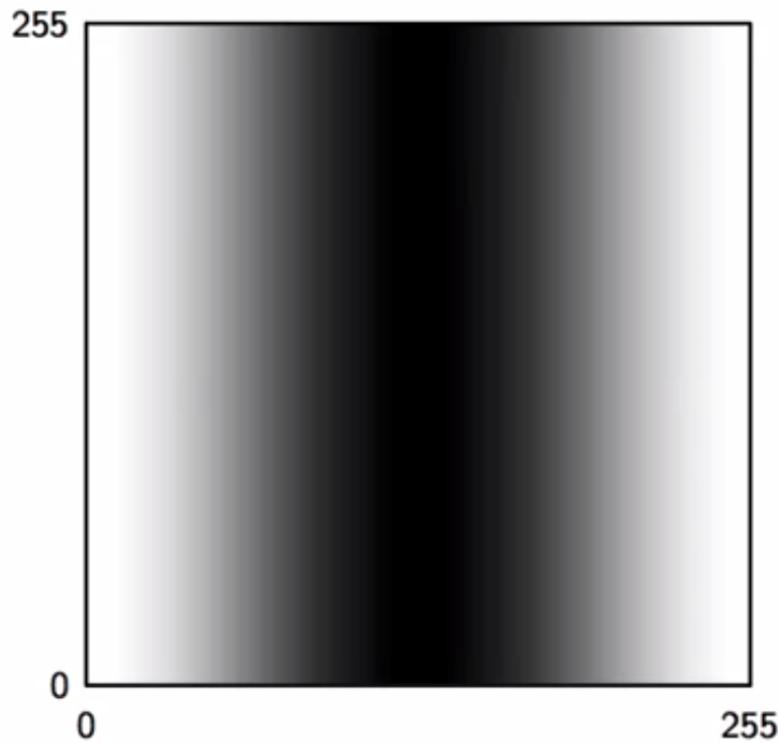
$$n_2, k_2 = 0, 1, \dots, N_2 - 1$$

Пусть $N_1 = N_2 = 256$. Как выглядят базисные функции?

Базис в пространстве 256×256

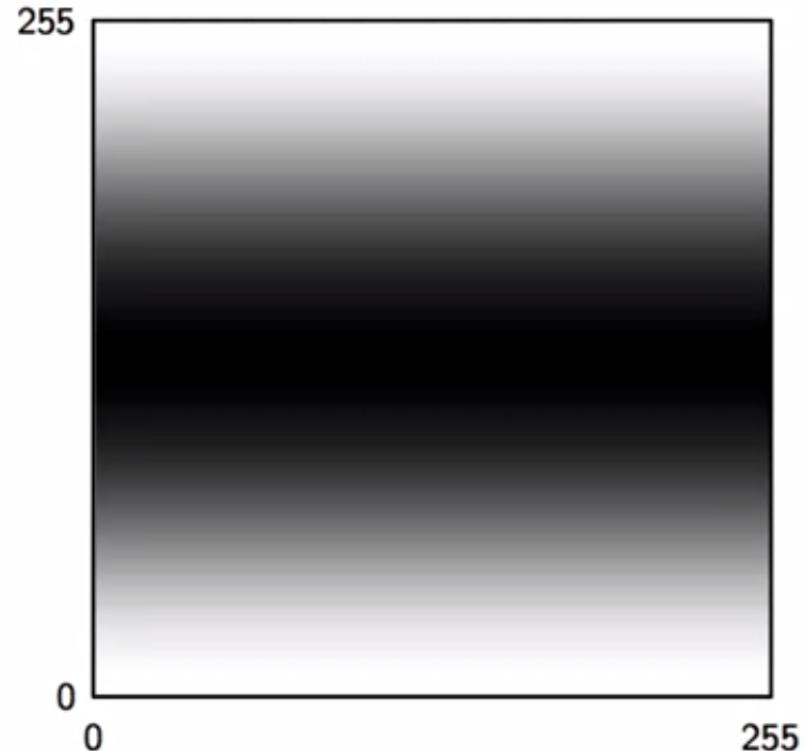
$k_1 = 1, k_2 = 0$

(действительная часть)



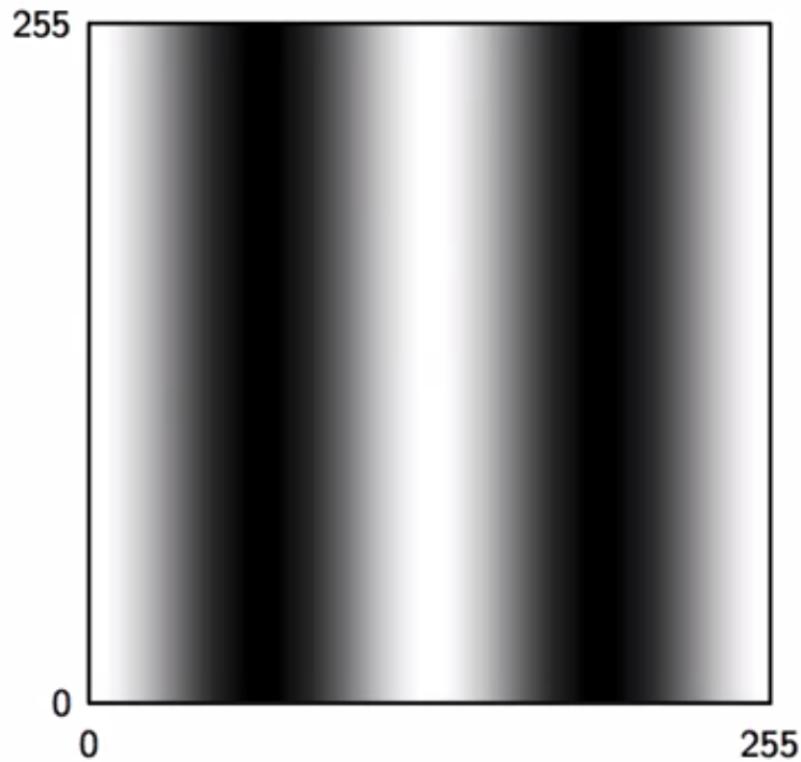
$k_1 = 0, k_2 = 1$

(действительная часть)

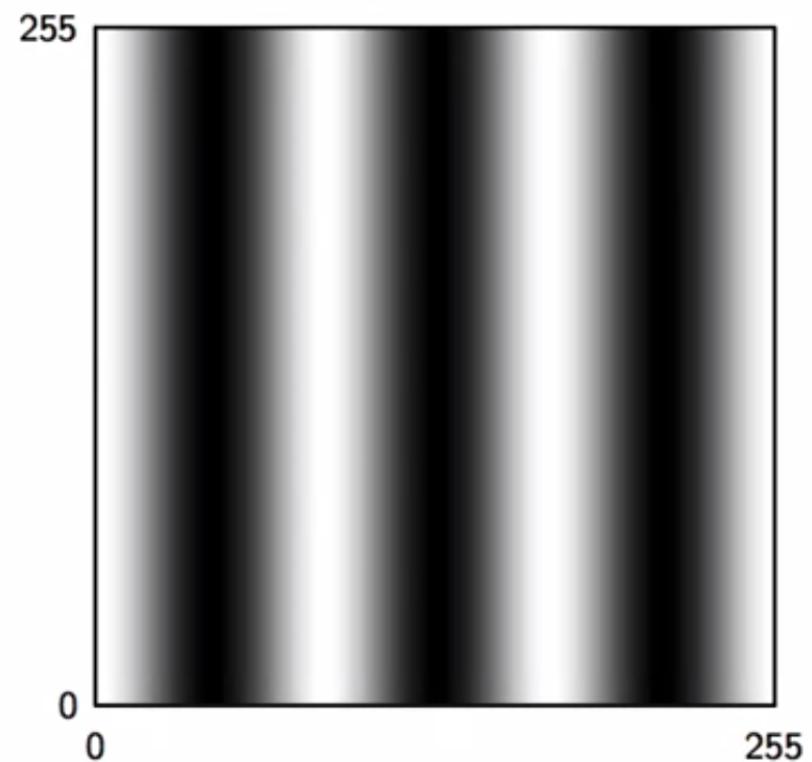


Базис в пространстве 256×256

$k_1 = 2, k_2 = 0$
(действительная часть)



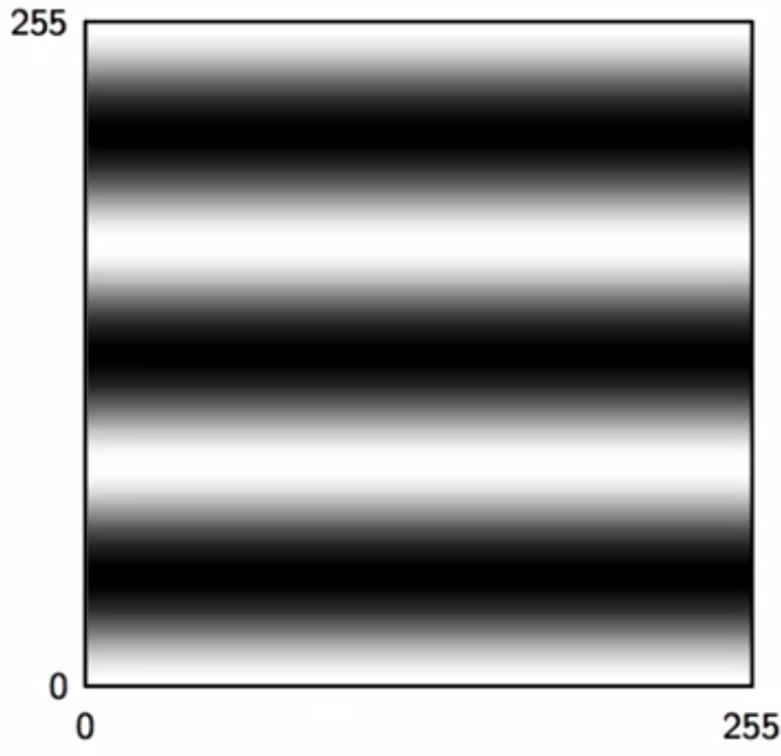
$k_1 = 3, k_2 = 0$
(действительная часть)



Базис в пространстве 256×256

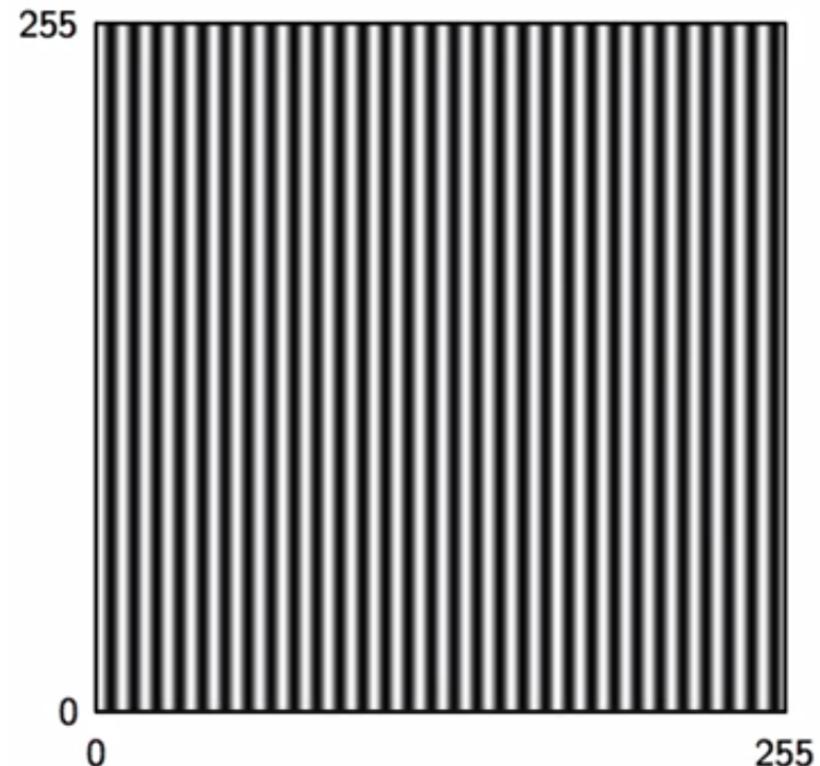
$k_1 = 0, k_2 = 3$

(действительная часть)



$k_1 = 30, k_2 = 0$

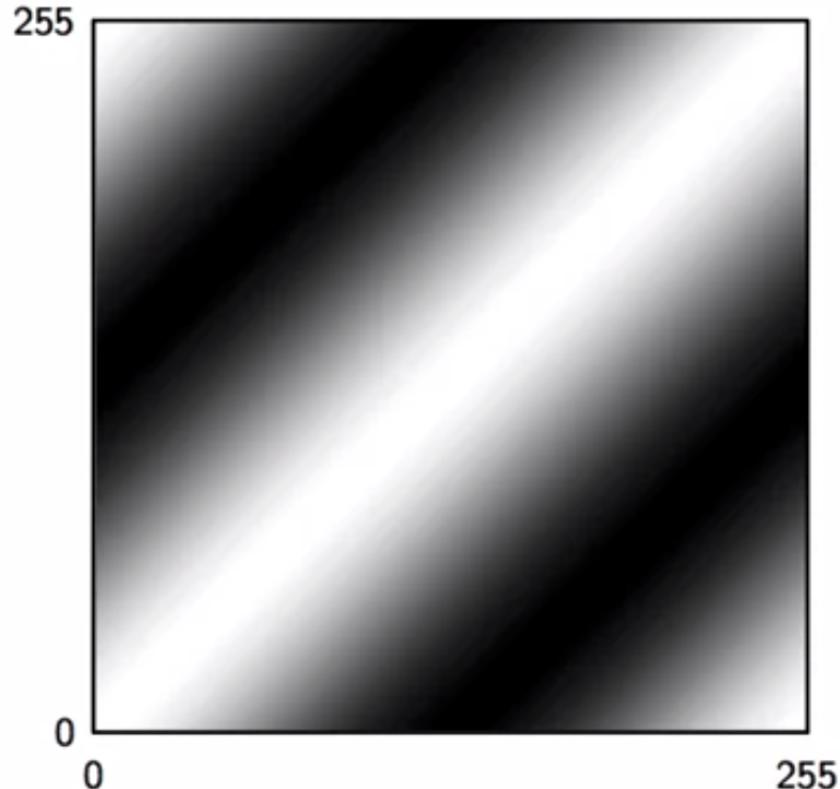
(действительная часть)



Базис в пространстве 256×256

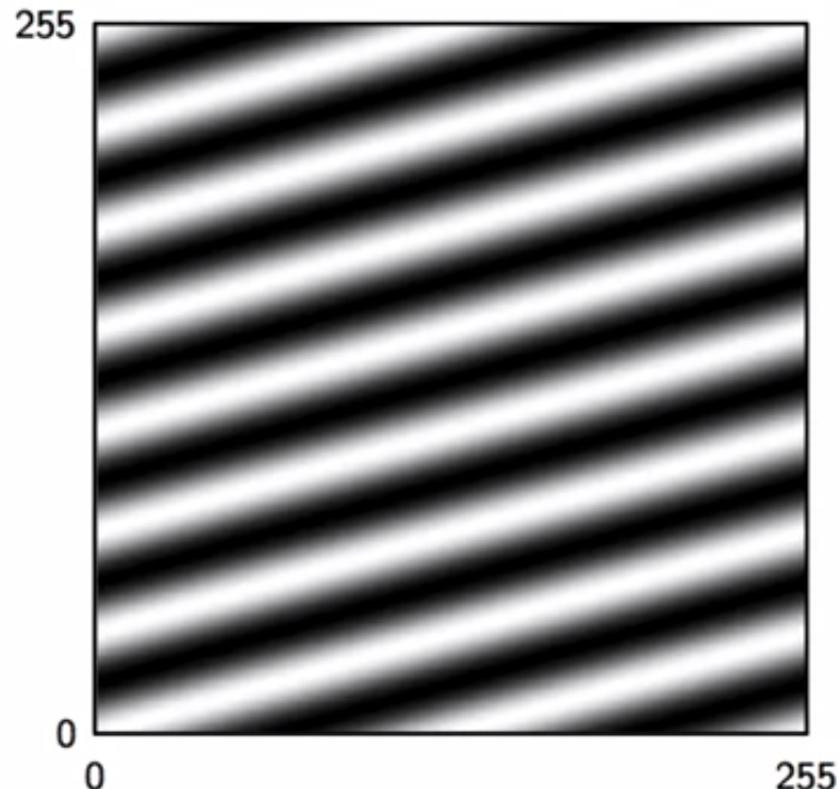
$k_1 = 1, k_2 = 1$

(действительная часть)



$k_1 = 2, k_2 = 7$

(действительная часть)



Спектральный анализ

$$X[k_1, k_2] = \sum_{n_1=0}^{N_1-1} \left[\sum_{n_2=0}^{N_2-1} x[n_1, n_2] e^{-j\frac{2\pi}{N_2} n_2 k_2} \right] e^{-j\frac{2\pi}{N_1} n_1 k_1}$$

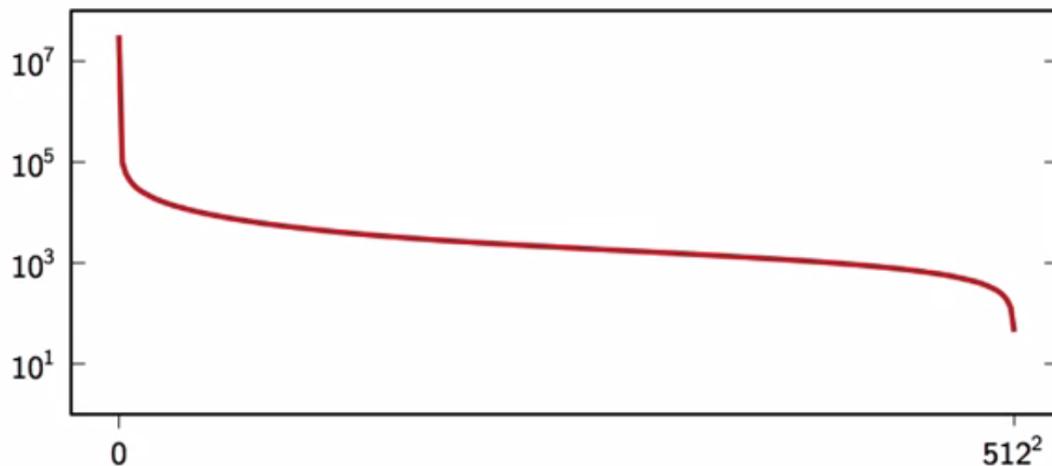
- 1D DFT вдоль столбцов (n_2)
- 1D DFT вдоль строк (n_1)

Количество действий (используя FFT):

$$N_1 N_2 \log_2 N_2 + N_2 N_1 \log_2 N_1 << (N_1 N_2)^2$$

Спектральный анализ

Проблема: диапазон коэффициентов ДПФ слишком большой по сравнению с уровнями градации черно-белых изображений



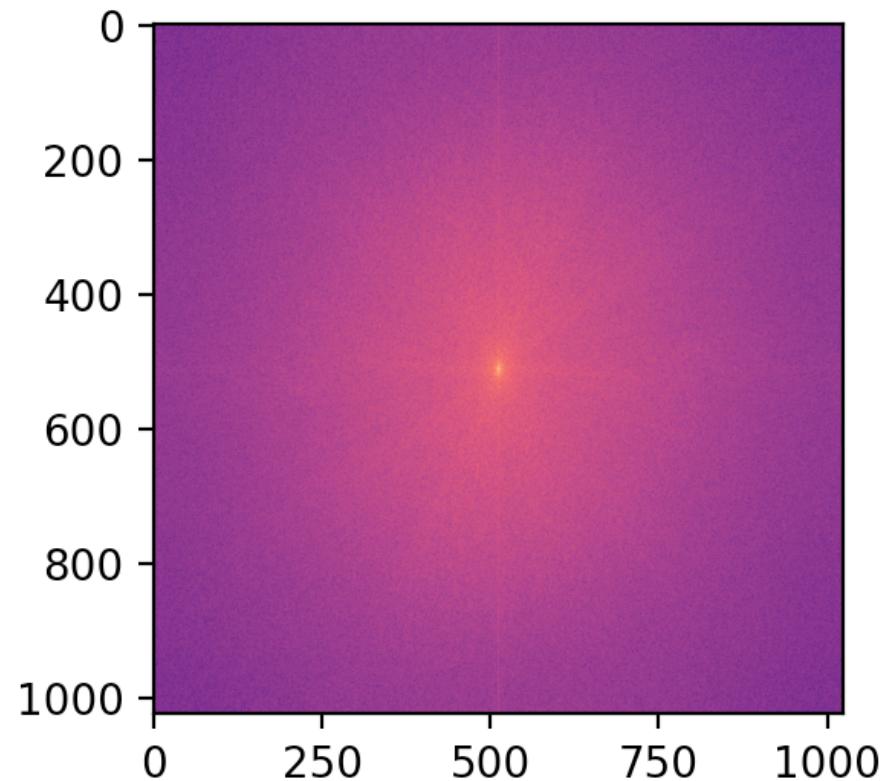
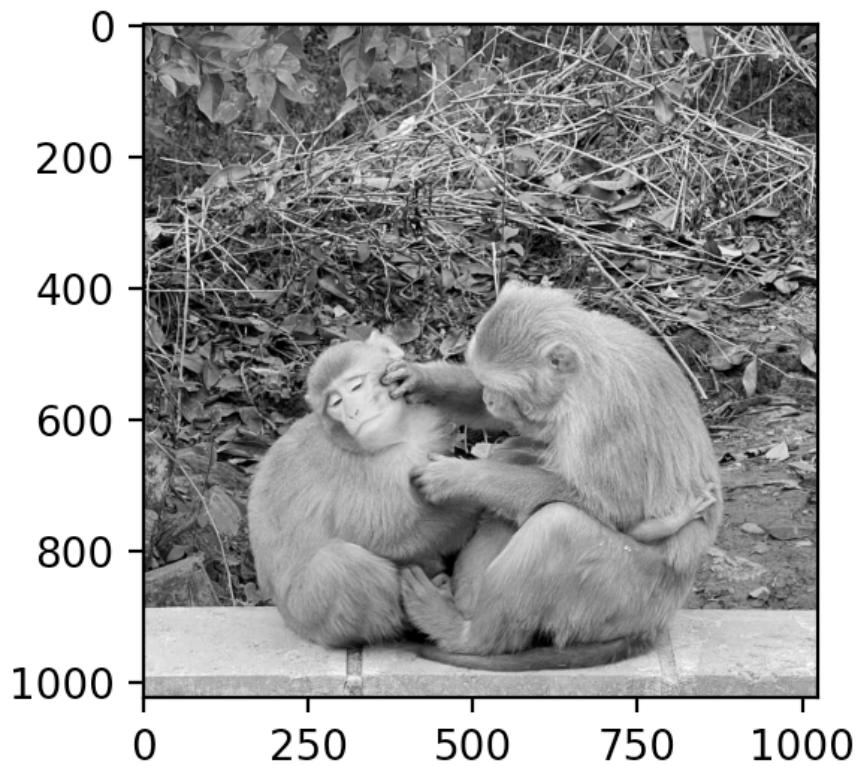
*типичная кривая
распределения
коэффициентов Фурье
по амплитудам*

Решение проблемы HDR (high-dynamic range) изображений:

- удаление “огромных” значений (типа $X[0,0]$)
- использование нелинейной шкалы

Спектральный анализ

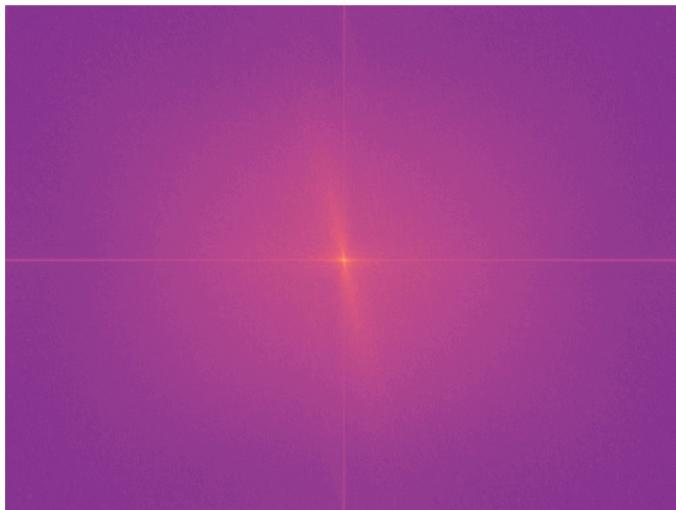
2D FFT для изображения размером 1024 x 1024



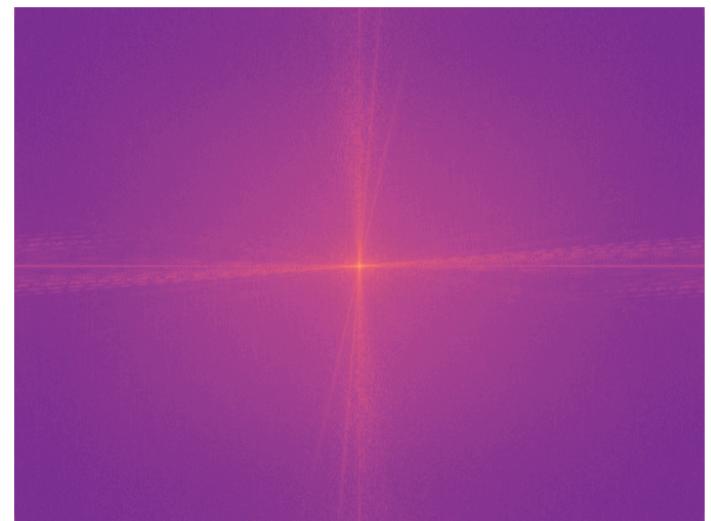
Спектральный анализ



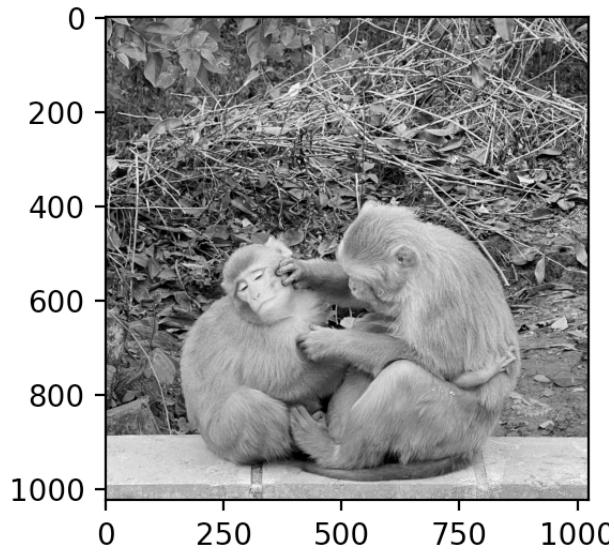
2D DFT



2D DFT



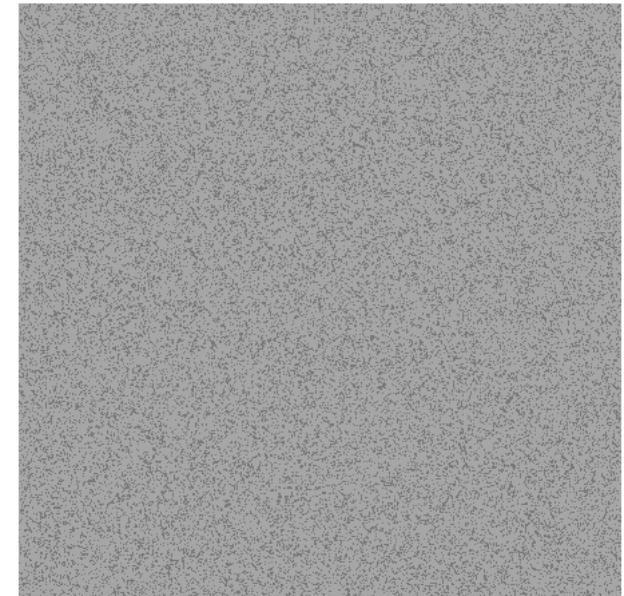
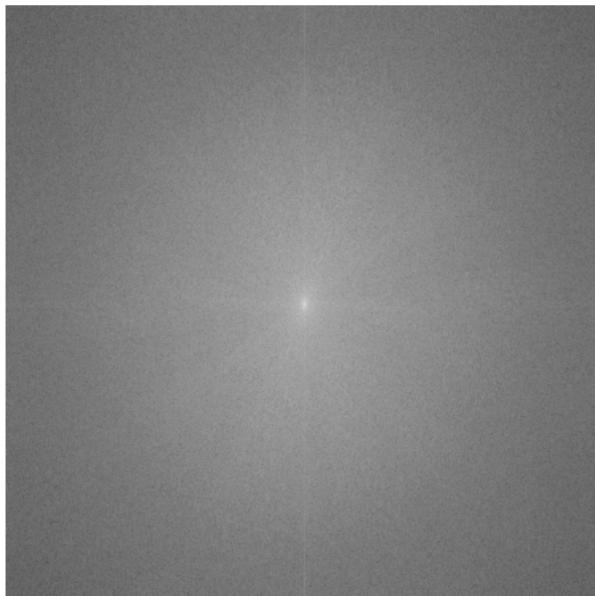
Спектральный анализ



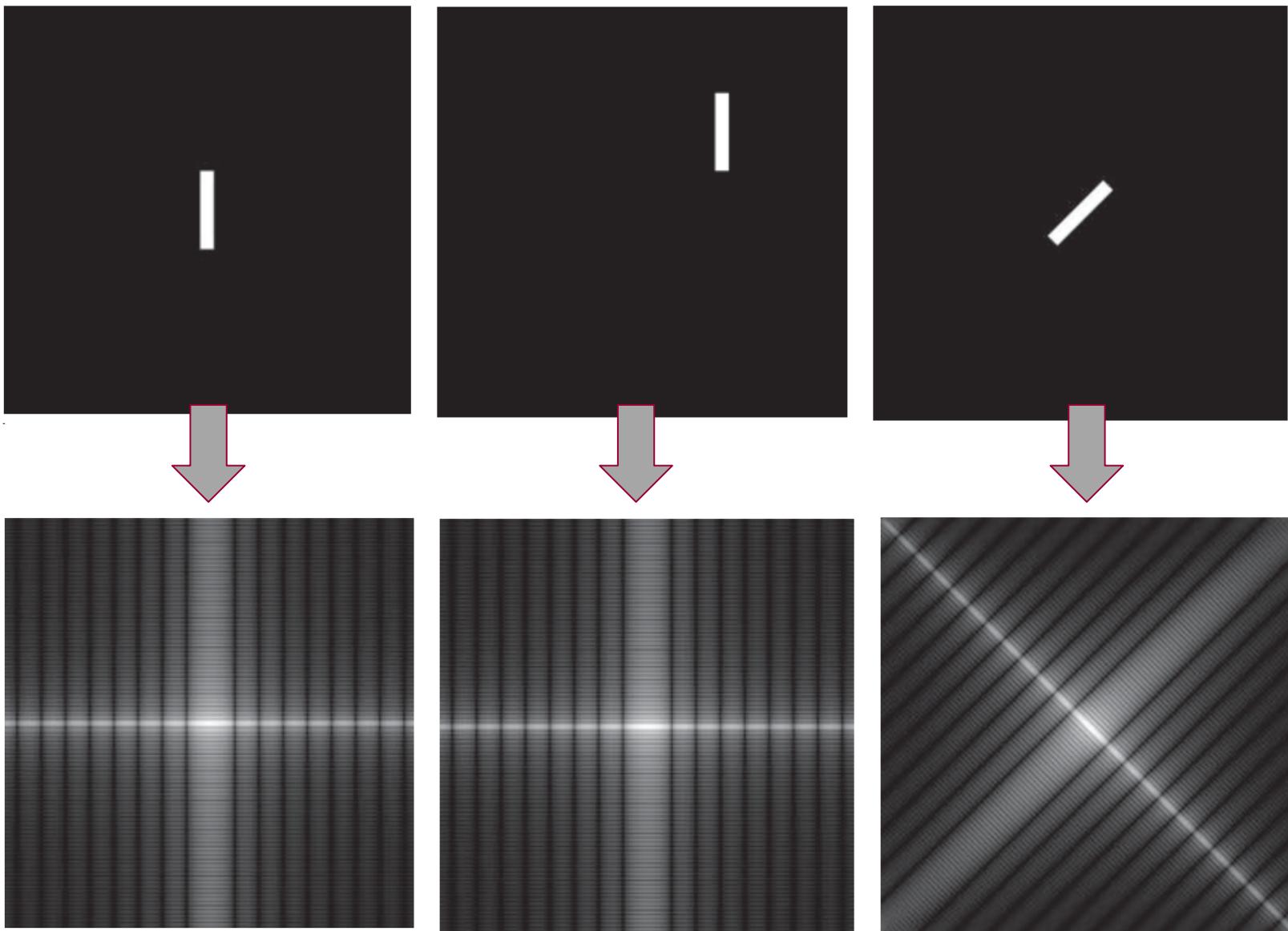
$$X[k_1, k_2] = |X[k_1, k_2]| e^{j\phi(k_1, k_2)}$$

амплитуда ДПФ
($\log(1+|X[k_1, k_2]|)$)

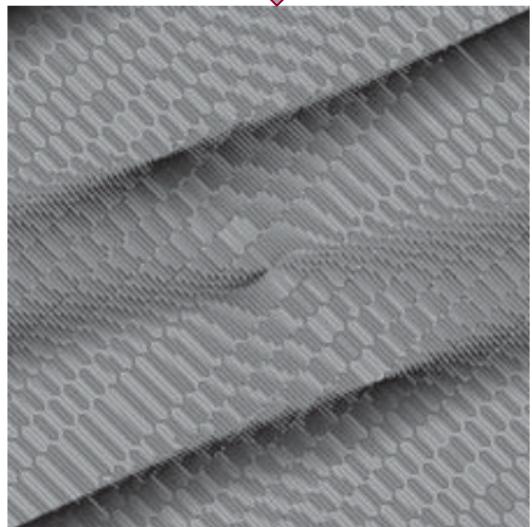
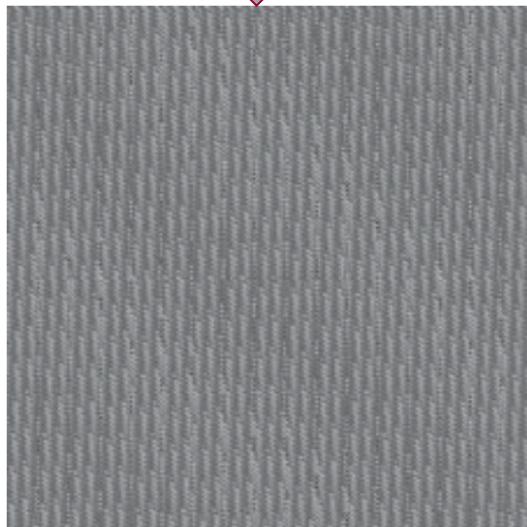
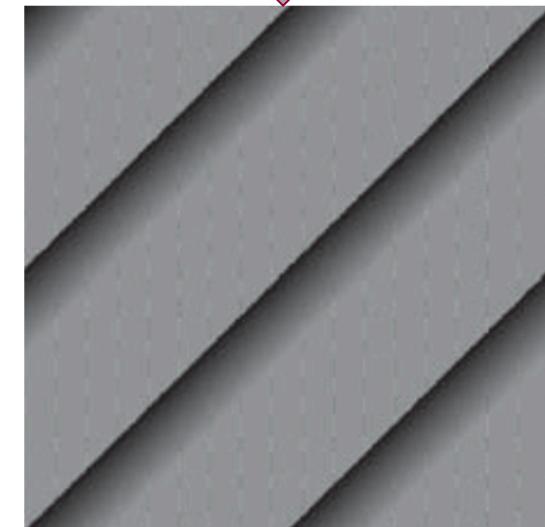
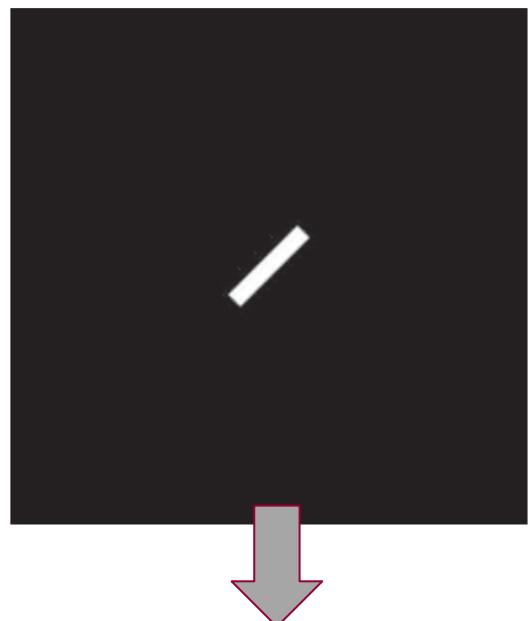
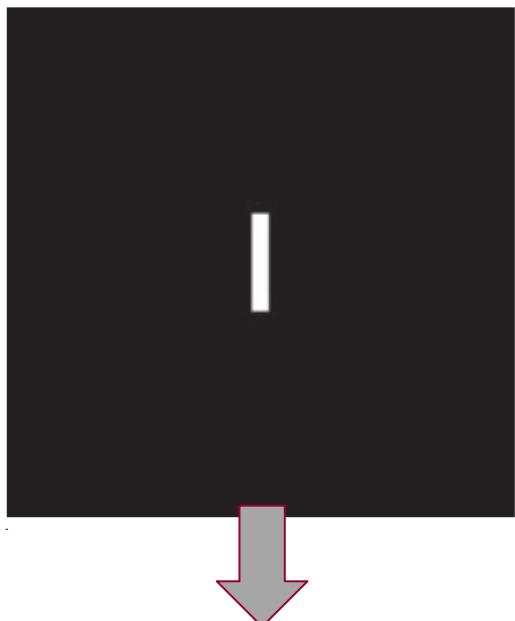
фаза ДПФ



Спектральный анализ



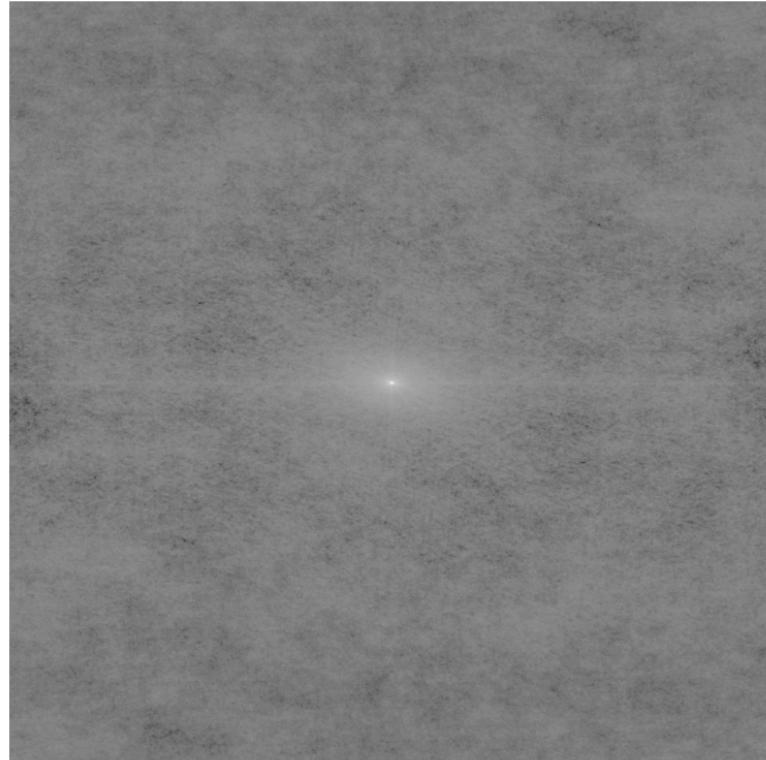
Спектральный анализ



фазы ДПФ

Спектральный анализ

DFT-отбрасывание фазы-
обратное DFT



DFT-отбрасывание
амплитуды-обратное DFT



Базис в 2D пространстве

https://nbviewer.org/github/forsp/dig_proc_lect/blob/main/haar.ipynb

Сжатие изображений

Сжатие изображений

Термин **сжатие данных** означает уменьшение объема данных, используемого для представления определенного количества информации.

Важно: данные ≠ информация

Для представления одного и того же количества информации может быть использовано различное количество данных.

Избыточность данных - ключевое понятие цифрового сжатия сигналов:

- кодовая избыточность
- межэлементная избыточность
- визуальная избыточность

Сжатие изображений

Пусть n_1 и n_2 означают число элементов – носителей информации – в двух наборах данных, представляющих одну и ту же информацию.

Тогда **относительная избыточность данных R_D** первого набора по отношению ко второму определяется как

$$R_D = 1 - \frac{1}{C_R},$$

где C_R - **степень сжатия**:

$$C_R = \frac{n_1}{n_2}$$

при степени сжатия 10:1 избыточность данных составляет 90%

Сжатие изображений

Методы сжатия изображений делятся на две группы:

- Сжатие **без потерь** (позволяет точно восстанавливать исходное изображение)

Основано на исключении *статистической избыточности* (алгоритмы RLE, LZW, Хаффмана и пр)

- Сжатие **с потерями** (или сжатие с регулируемой степенью искажений, изображение восстанавливается приближенно)

Основано на снижении *психофизиологической избыточности*

Сжатие изображений

Основные **подходы** к процессу сжатия изображений:

- (1) сжатие на уровне блоков (не пикселей)
- (2) использование подходящего разложения
(смена базиса)
- (3) “умное” квантование
- (4) энтропийное (статистическое) кодирование

Сжатие изображений

(1) сжатие на уровне блоков

Сжатие на уровне пикселей возможно только за счет уменьшения количества бит, выделяемых для хранения значений каждого пикселя.

Максимальное сжатие возможно, только если оставить 1 бит/пиксель



Сжатие изображений

(1) сжатие на уровне блоков

При блочном сжатии используется локальная корреляция между пикселями:

- изображение делится на отдельные блоки
- хранится среднее значение атрибута в каждом блоке (8 бит)

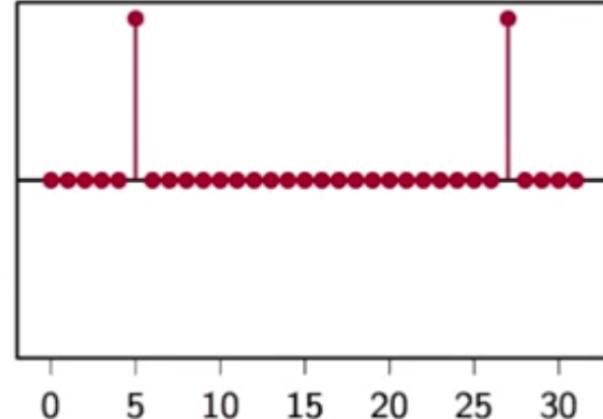
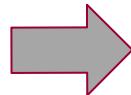
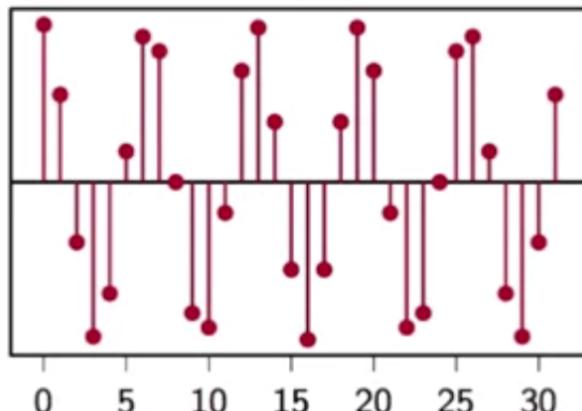
для блоков 3x3 при кодировании
8 бит на блок - меньше 1 бит/
пиксель (!)



Сжатие изображений

(2) смена базиса

- хранение дискретного сигнала размера N , записанного с R бит/отсчет, требует NR бит
- после расчета ДПФ сигнала может оказаться, что большинство коэффициентов разложения равны нулю
- в пределе остается кодировать лишь два ненулевых коэффициента



Сжатие изображений

(2) смена базиса

- нужно разложение, сохраняющее основные черты изображения с минимумом коэффициентов разложения
- разложение должно быть алгоритмически простым



Real - Real

Косинусное преобразование Фурье (DCT)

$$C[k_1, k_2] = \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} x[n_1, n_2] \cos \left[\frac{\pi}{N} \left(n_1 + \frac{1}{2} \right) k_1 \right] \cos \left[\frac{\pi}{N} \left(n_2 + \frac{1}{2} \right) k_2 \right]$$

Сжатие изображений

(2) смена базиса

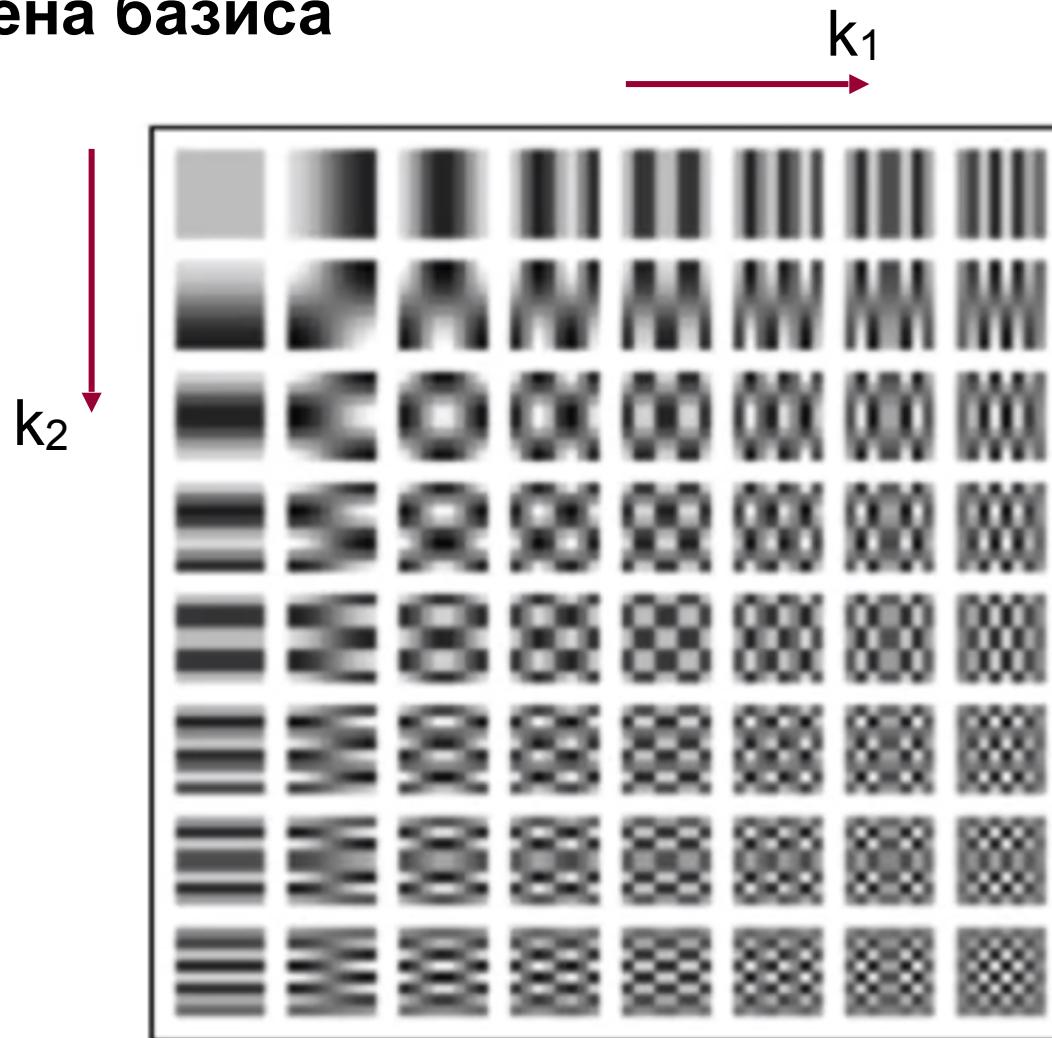
$$C[k_1, k_2] = \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} x[n_1, n_2] \cos \left[\frac{\pi}{N} \left(n_1 + \frac{1}{2} \right) k_1 \right] \cos \left[\frac{\pi}{N} \left(n_2 + \frac{1}{2} \right) k_2 \right]$$

Для блоков 8x8 (N=8) получаем 64 базисные функции

Каждый коэффициент разложения показывает корреляцию между изображением и каждой их базисных функций

Сжатие изображений

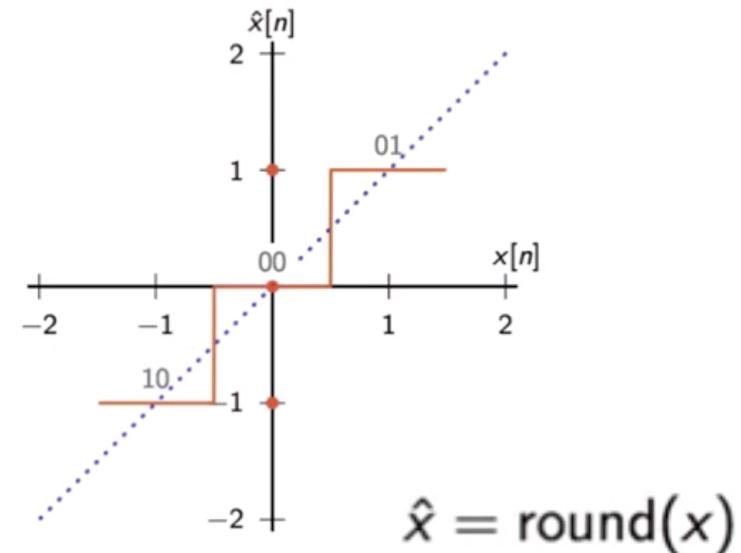
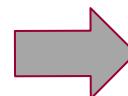
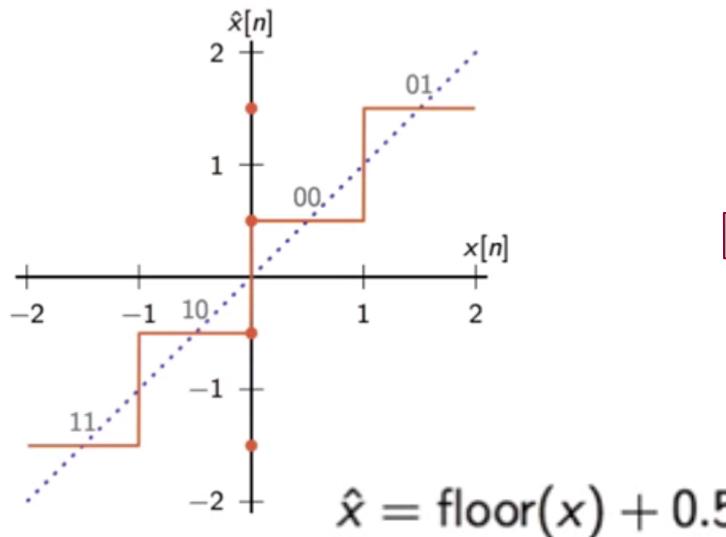
(2) смена базиса



Сжатие изображений

(3) умное квантование

- динамическое изменение интервала квантования (fine to coarse)
- квантование с нулевой зоной (deadzone quantization): изменяется интервал, соответствующий нулевому уровню



Сжатие изображений

(4) энтропийное (статистическое) кодирование

- минимизация затрат на кодирование информации
- сопоставление коротких символов наиболее частым значениям и наоборот

1836 г -
Азбука Морзе

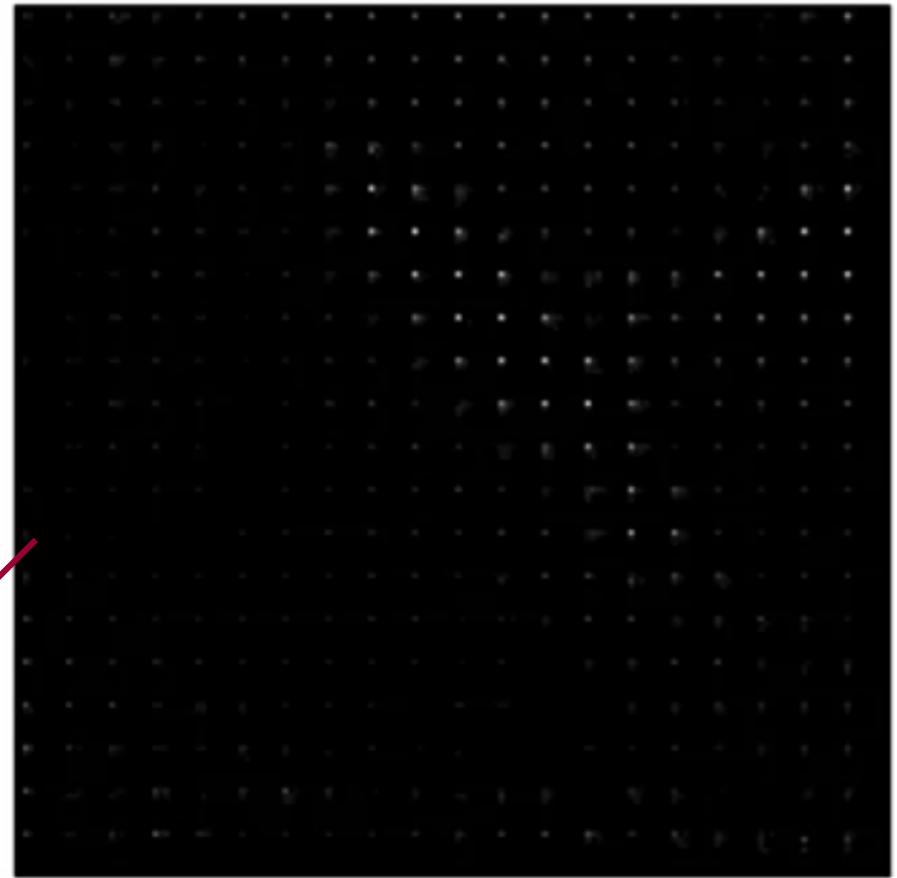


JPEG сжатие

JPEG сжатие

JPEG-алгоритм (стандарт
The Joint Photographic
Experts Group, 1993 г)

сжатие с потерями



DCT-коэффициенты (блоки 8x8)

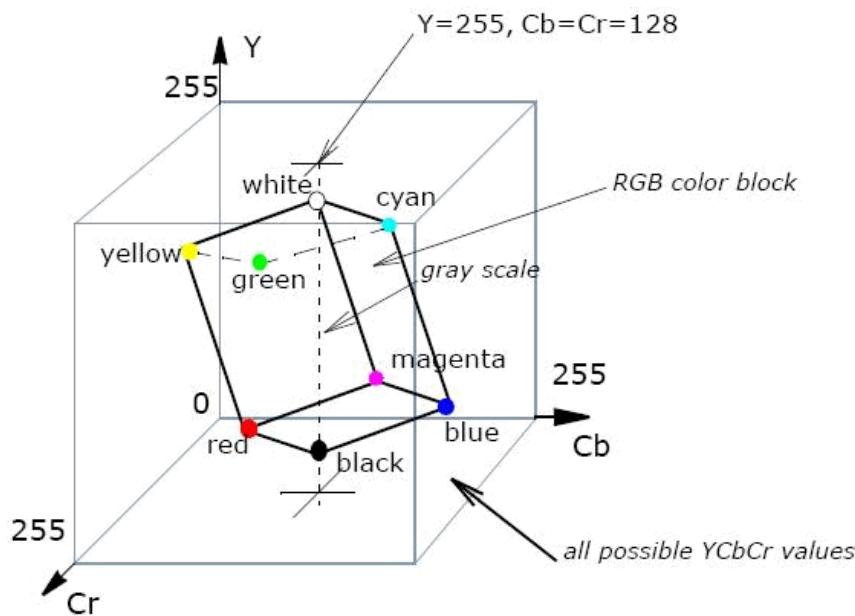
JPEG сжатие: этап 1-3

Этап 1: конвертация из RGB в цветовое пространство YCbCr

$$Y = 0,299 \cdot R + 0,587 \cdot G + 0,114 \cdot B$$

$$Cb = -0,1687 \cdot R - 0,3313 \cdot G + 0,5 \cdot B + 128$$

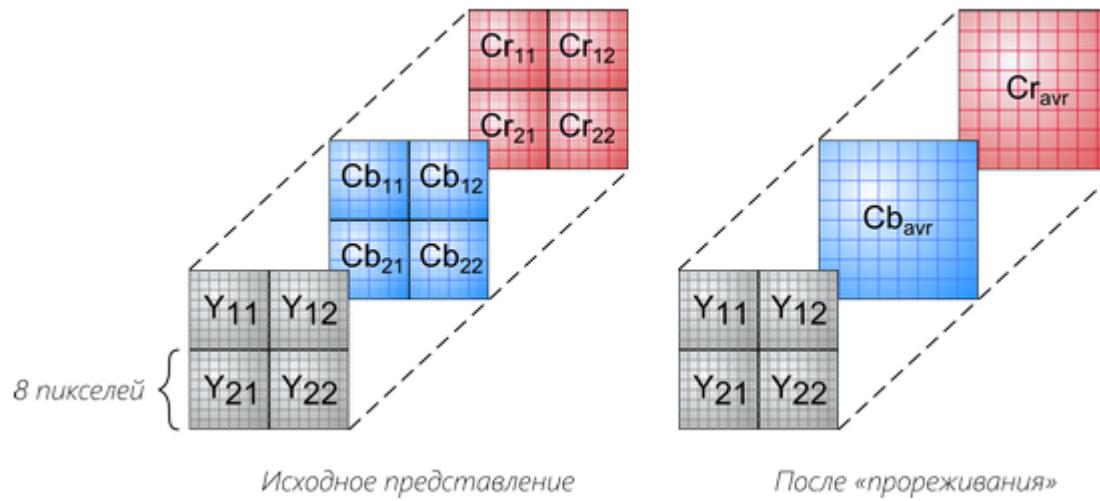
$$Cr = 0,5 \cdot R - 0,4187 \cdot G + 0,0813 \cdot B + 128$$



Y – яркостная составляющая;
Cr, Cb – компоненты,
отвечающие за цвет
(хроматический красный и
хроматический синий)

JPEG сжатие: этап 1-3

Этап 2: разбивка изображения на блоки 8x8



Этап 3: расчет ДКП на каждом блоке 8x8,
на каждом блоке – по три матрицы ДКП для коэффициентов Y, Cr, Cb

JPEG сжатие: этап 4

“Умное” квантование DCT-коэффициентов проводится на основе результатов психо-визуальных экспериментов

$$\hat{c}[k_1, k_2] = \text{round}(c[k_1, k_2]/Q[k_1, k_2])$$

$$Q_l = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix} \quad Q_c = \begin{bmatrix} 17 & 18 & 24 & 47 & 99 & 99 & 99 & 99 \\ 18 & 21 & 26 & 66 & 99 & 99 & 99 & 99 \\ 24 & 26 & 56 & 99 & 99 & 99 & 99 & 99 \\ 47 & 66 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \end{bmatrix}$$

для Y компоненты

для Cr/Cb компонент

JPEG сжатие: этап 4

Значения пикселей

DCT

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|--|-------|-------|-------|------|------|------|------|------|
| 139 | 144 | 149 | 153 | 155 | 155 | 155 | 155 | | 235,6 | -1,0 | -12,1 | -5,2 | 2,1 | -1,7 | -2,7 | 1,3 |
| 144 | 151 | 153 | 156 | 159 | 156 | 156 | 156 | | -22,6 | -17,5 | -6,2 | -3,2 | -2,9 | -0,1 | 0,4 | -1,2 |
| 150 | 155 | 160 | 163 | 158 | 156 | 156 | 156 | | -10,9 | -9,3 | -1,6 | 1,5 | 0,2 | -0,9 | -0,6 | -0,1 |
| 159 | 161 | 162 | 160 | 160 | 159 | 159 | 159 | | -7,1 | -1,9 | 0,2 | 1,5 | 0,9 | -0,1 | 0,0 | 0,3 |
| 159 | 160 | 161 | 162 | 162 | 155 | 155 | 155 | | -0,6 | -0,8 | 1,5 | 1,6 | -0,1 | -0,7 | 0,6 | 1,3 |
| 161 | 161 | 161 | 161 | 160 | 157 | 157 | 157 | | 1,8 | -0,2 | 1,6 | -0,3 | -0,8 | 1,5 | 1,0 | -1,0 |
| 162 | 162 | 161 | 163 | 162 | 157 | 157 | 157 | | -1,3 | -0,4 | -0,3 | -1,5 | -0,5 | 1,7 | 1,1 | -0,8 |
| 162 | 162 | 161 | 161 | 163 | 158 | 158 | 158 | | -2,6 | 1,6 | -3,8 | -1,8 | 1,9 | 1,2 | -0,6 | -0,4 |

DCT+Q matrix

JPEG сжатие: этап 4

| <i>Значения пикселей</i> | <i>Восстановленные DCT</i> |
|--|---------------------------------|
| 139 144 149 153 155 155 155 155 | 240 0 -10 0 0 0 0 0 |
| 144 151 153 156 159 156 156 156 | -24 -12 0 0 0 0 0 0 |
| 150 155 160 163 158 156 156 156 | -14 -13 0 0 0 0 0 0 |
| 159 161 162 160 160 159 159 159 | 0 0 0 0 0 0 0 0 |
| 159 160 161 162 162 155 155 155 | 0 0 0 0 0 0 0 0 |
| 161 161 161 161 160 157 157 157 | 0 0 0 0 0 0 0 0 |
| 162 162 161 163 162 157 157 157 | 0 0 0 0 0 0 0 0 |
| 162 162 161 161 163 158 158 158 | 0 0 0 0 0 0 0 0 |
| | |
| <i>Восстановленные значения пикселей</i> | 144 146 149 152 154 156 156 156 |
| | 148 150 152 154 156 156 156 156 |
| | 155 156 157 158 158 157 156 155 |
| | 160 161 161 162 161 159 157 155 |
| | 163 163 164 163 162 160 158 156 |
| | 163 164 164 164 162 160 158 157 |
| | 160 161 162 162 162 161 159 158 |
| | 158 159 161 161 162 161 159 158 |

JPEG сжатие: этап 4

“Умное” квантование DCT-коэффициентов с матрицей Q

0.2 bpp



uniform



tuned

JPEG сжатие: этап 4

0.2 bpp



uniform

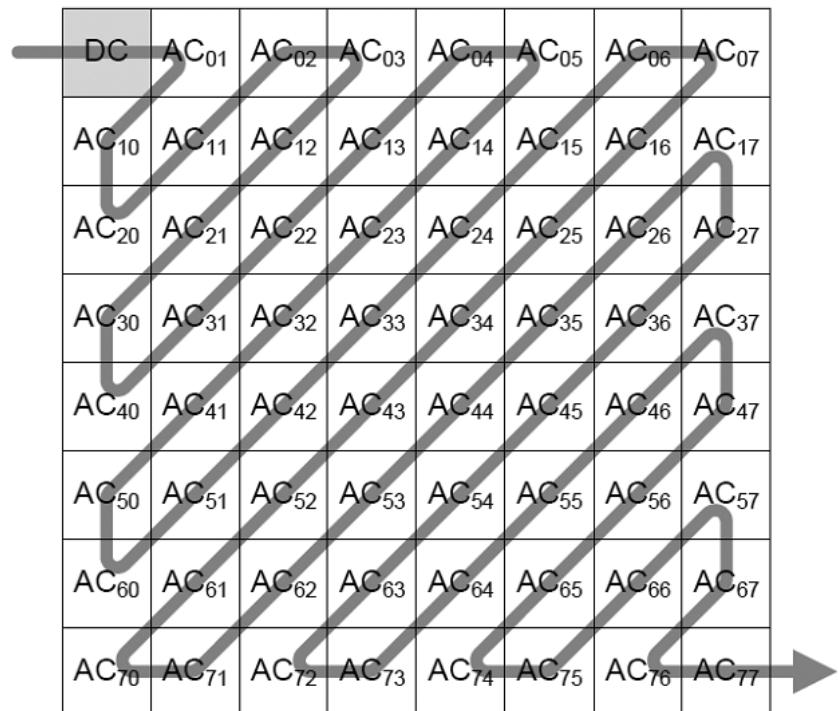


tuned

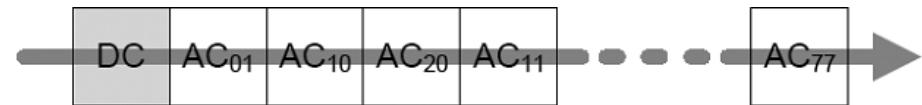
JPEG сжатие: этап 5

«умное» кодирование:

матрица квантованных коэффициентов 8x8 переводится в 64-элементный вектор при помощи **«зигзаг»-сканирования**



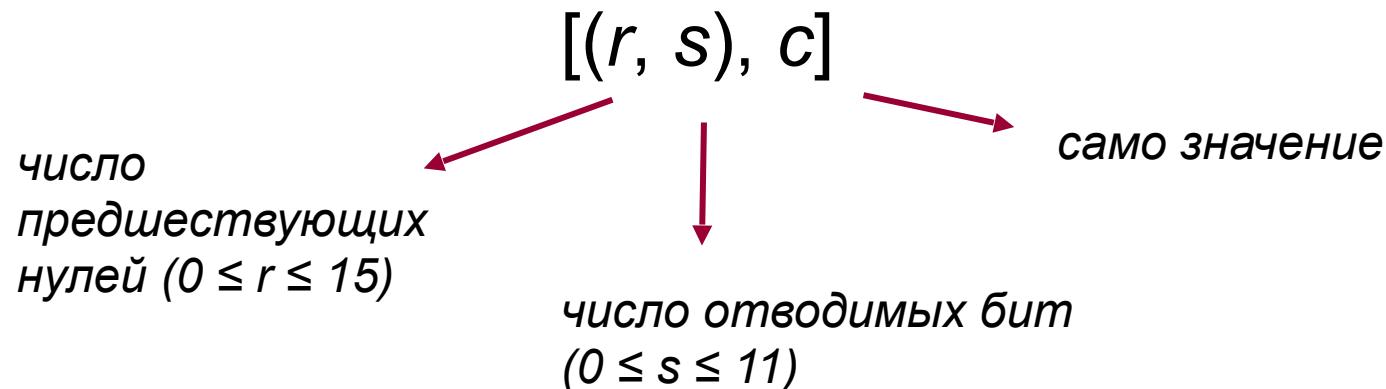
в начале вектора будут записаны коэффициенты, соответствующие низким частотам, а в конце – высоким (почти всегда равны нулю)



JPEG сжатие: этап 6

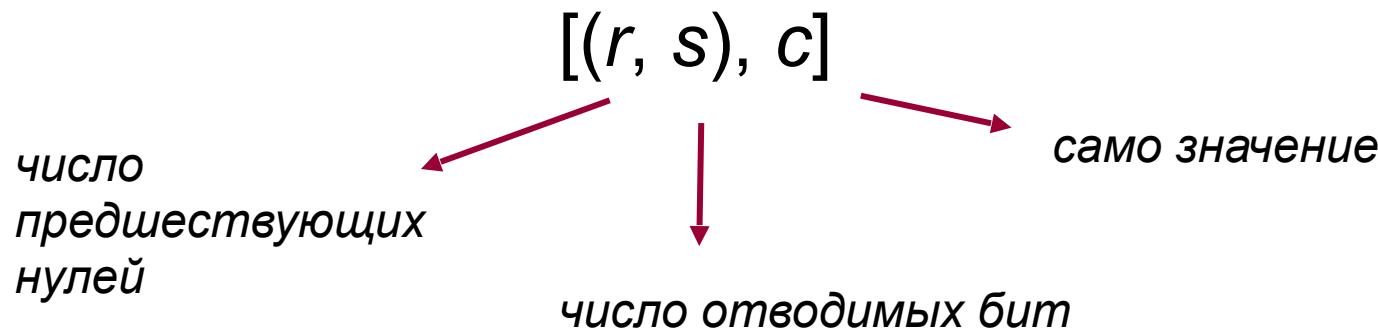
Run-Length Encoding (**RLE**) – алгоритм **группового кодирования**

Каждое ненулевое значение кодируется тройкой:

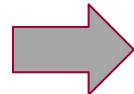


(0,0) – конец блока (дальше нули)

JPEG сжатие: этап 6



| | | | | | | | |
|-----|-----|---|---|---|---|---|---|
| 100 | -60 | 0 | 6 | 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 |
| 13 | -1 | 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 |

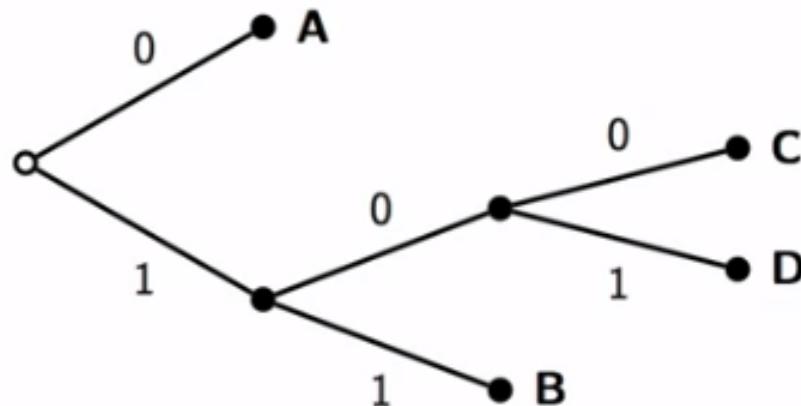


$[(0, 7), 100], [(0, 6), -60],$
 $[(4, 3), 6], [(3, 4), 13],$
 $[(8, 1), -1], [(0, 0)]$

JPEG сжатие: этап 7

алгоритм Хаффмана

- префиксное правило: ни один код не является префиксом другого



001100110101100



AABAABADC

JPEG сжатие: этап 7

алгоритм Хаффмана

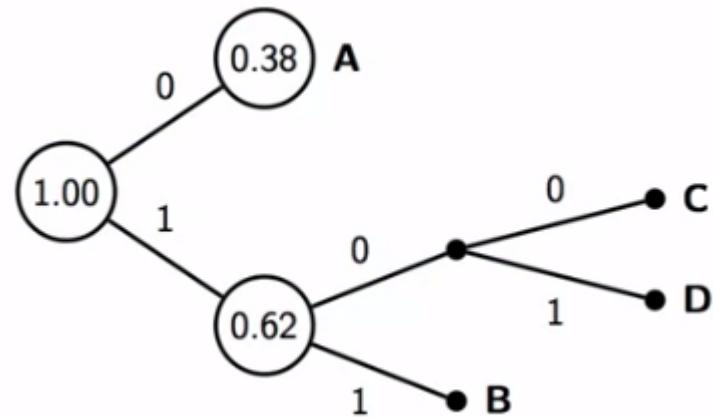
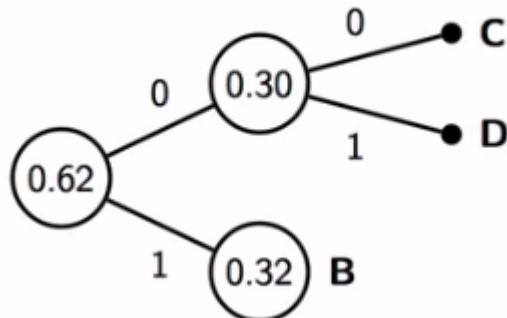
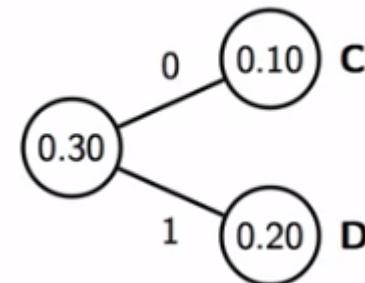
- построение дерева Хаффмана

$$p(A) = 0.38$$

$$p(B) = 0.32$$

$$p(C) = 0.1$$

$$p(D) = 0.2$$



JPEG сжатие: этап 7

таблицы Хаффмана

Таблица Хаффмана для кодирования коэффициентов в алгоритме JPEG

| Диапазоны коэффициентов | Кол-во битов (категория) | Значения кодов |
|--------------------------------------|--------------------------|-------------------------------------|
| [-1, 1] | 1 | 0, 1 |
| [-3, -2], [2, 3] | 2 | 00, 01, 10, 11 |
| [-7, -4], [4, 7] | 3 | 000, 001, 010, ..., 111 |
| [-15, -8], [8, 15] | 4 | 0000, 0001, ..., 1111 |
| [-31, -16], [16, 31] | 5 | 00000, ..., 11111 |
| [-63, -32], [32, 63] | 6 | 000000, ..., 111111 |
| [-127, -64], [64, 127] | 7 | 0000000, ..., 1111111 |
| [-255, -128], [128, 255] | 8 | 00000000, ..., 11111111 |
| [-511, -256], [256, 511] | 9 | 000000000, ..., 111111111 |
| [-1023, -512], [512, 1023] | 10 | 0000000000, ..., 1111111111 |
| ... | ... | ... |
| [-32 767, -16 384], [16 384, 32 767] | 15 | 000000000000000, ..., 1111111111111 |

Свертка и корреляция

Свертка и корреляция

- Свертка фильтра (изображения) $w(x, y)$ размерами $m \times n$ с изображением $f(x, y)$

$$w(x, y) \star f(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x-s, y-t)$$

↓

фильтр свертки, маска свертки или ядро свертки

- Корреляция (кросс-корреляция, перекрестная корреляция) фильтра (изображения) $w(x, y)$ размерами $m \times n$ с изображением $f(x, y)$

$$w(x, y) \star f(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x+s, y+t)$$

$$a = (m - 1)/2, b = (n - 1)/2$$

Свертка и корреляция

Корреляция

Начало

↙ координат f
0 0 0 1 0 0 0 0

↙ координат w
1 2 3 2 8

Свертка

Начало

↙ координат f
0 0 0 1 0 0 0 0

w повернута на 180°
8 2 3 2 1

↓
0 0 0 1 0 0 0 0
1 2 3 2 8
↖ Выравнивание начальной позиции

0 0 0 1 0 0 0 0
8 2 3 2 1

↖ Расширение нулями

0 0 0 0 0 0 0 1 0 0 0 0 0
1 2 3 2 8

0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
8 2 3 2 1

Свертка и корреляция

Корреляция

| Начало | | | | | | | w | | | | |
|-----------|---|---|---|---|---|---|-----|---|---|---|---|
| координат | | | | | | | f | | | | |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 2 | 3 | 2 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 3 | 2 | 8 |

Свертка

| Начало | | | | | | | w повернута на 180° | | | | |
|-----------|---|---|---|---|---|---|---------------------|---|---|---|---|
| координат | | | | | | | f | | | | |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 8 | 2 | 3 | 2 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 3 | 2 | 1 |

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
1 2 3 2 8

↑ Позиция после одного сдвига

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 2 3 2 1

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
1 2 3 2 8

↑ Позиция после четырех сдвигов

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 2 3 2 1

0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
1 2 3 2 8

↑ Конечная позиция

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8 2 3 2 1

Свертка и корреляция

Корреляция

| Начало координат | | | | | | | f | w | | | |
|------------------|---|---|---|---|---|---|-----|-----|---|---|---|
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 2 | 3 | 2 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 8 | | |

Полный результат корреляции

0 0 0 8 2 3 2 1 0 0 0 0

Обрезанный результат корреляции

0 8 2 3 2 1 0 0

Свертка

| Начало координат | | | | | | | f | w повернута на 180° | | | |
|------------------|---|---|---|---|---|---|-----|-----------------------|---|---|---|
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 8 | 2 | 3 | 2 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 1 | | |

Полный результат свертки

0 0 0 1 2 3 2 8 0 0 0 0

Обрезанный результат свертки

0 1 2 3 2 8 0 0

корреляция и свертка являются функциями смещения

Свертка и корреляция

Нормированный коэффициент корреляции двух изображений

$$K(f, g) = \frac{\sum_{\langle x,y \rangle \in X} (f(x, y) - f^0)(g(x, y) - g^0)}{\sqrt{\sum_{\langle x,y \rangle \in X} (f(x, y) - f^0)^2} \sqrt{\sum_{\langle x,y \rangle \in X} (g(x, y) - g^0)^2}},$$

где f_0 и g_0 - средние значения интенсивности (яркости) пикселей на изображениях f и g

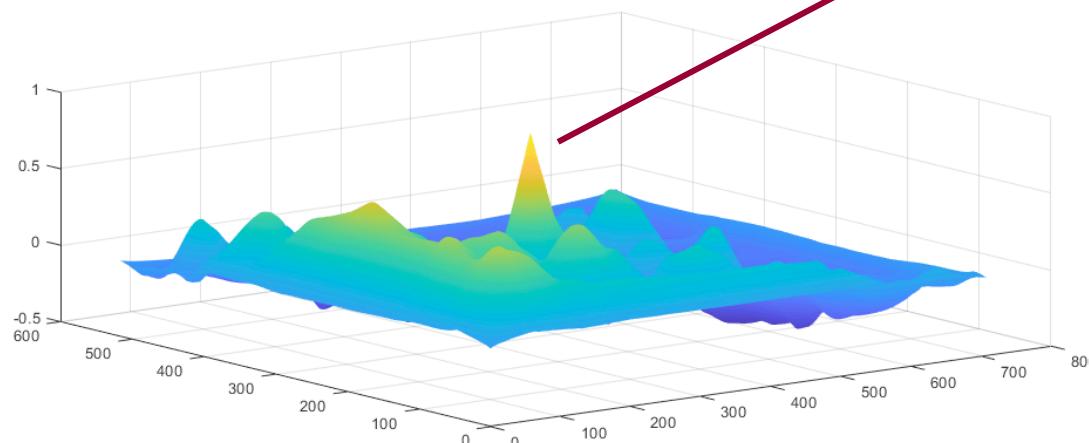


<https://www.mathworks.com/help/images/ref/normxcorr2.html>

Свертка и корреляция

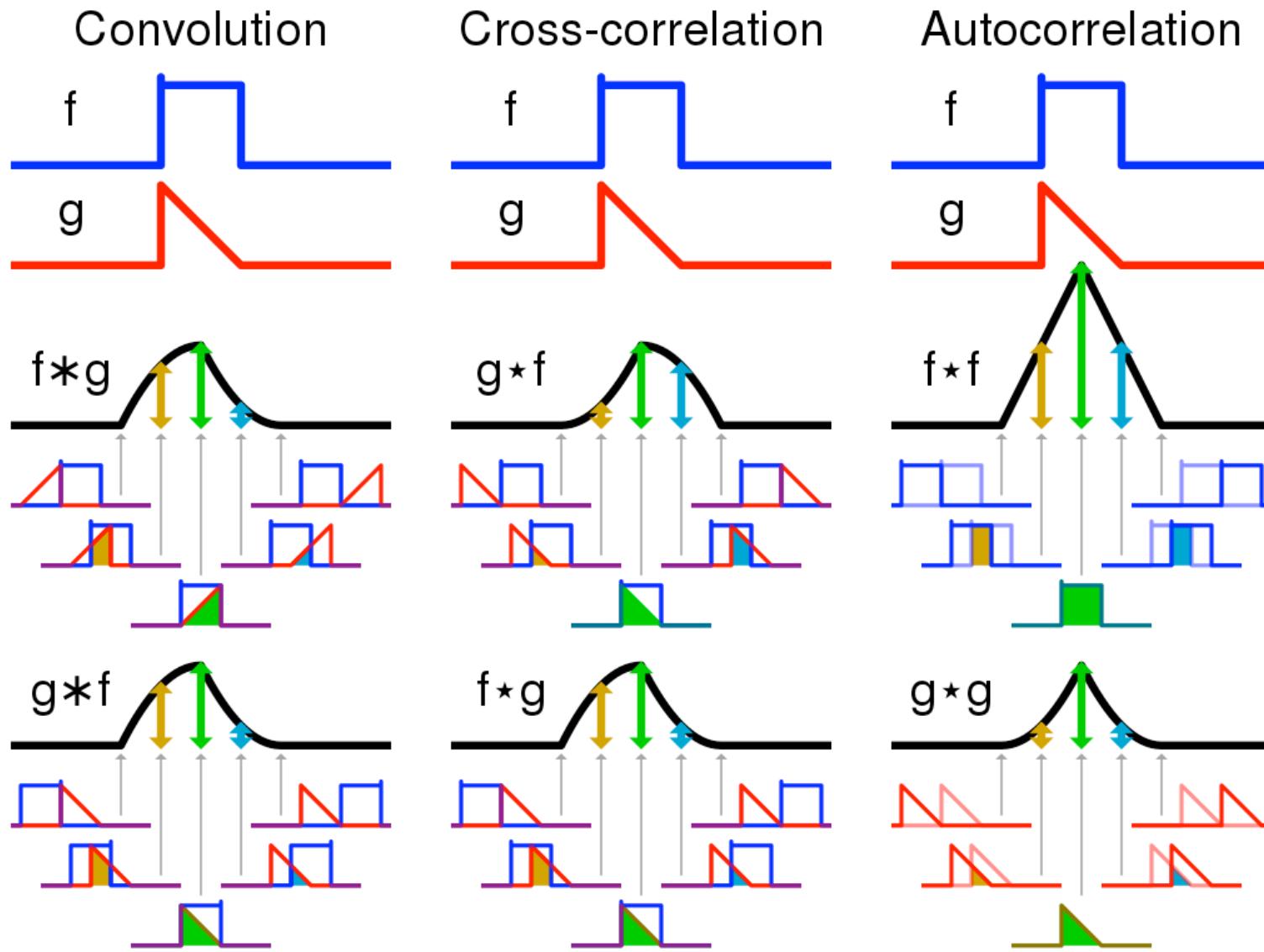


максимум кросс-корреляции



<https://www.mathworks.com/help/images/ref/normxcorr2.html>

Свертка и корреляция



Фрактальный анализ

Фрактальный анализ

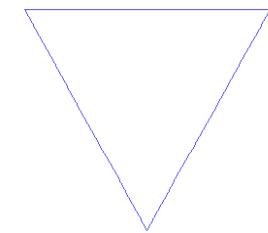
Под **фракталом** (от лат. *fractus* — дробленый, разбитый на куски) понимают математическое множество, обладающее свойством самоподобия, т. е. масштабной инвариантности (Мандельброт, 1977).



Треугольник Серпинского
задается 3 афинными преобразованиями



Снежинка Коха



Фрактальный анализ

Применение фракталов в обработке / анализе изображений:

- *фрактальное сжатие изображений* (Барнсли, 1988)

Фракталом в изображении называется структура, выделенная при анализе изображения и обладающая схожей формой независимо от ее размеров.

Постулат фрактального подхода:

существует набор аффинных коэффициентов, описывающих вращение, сжатие, расширение, искажение формы и сдвиг объектов, составляющих изображение.

Фрактальное сжатие - алгоритм сжатия изображений с потерями, основанный на применении систем итерируемых функций (как правило являющимися аффинными преобразованиями) к изображениям.

Фрактальный анализ

Применение фракталов в обработке изображений:

- генерация фрактальных структур при получении синтетических изображений



Пример фрактального изображения - изображение папоротника. Оно состоит из четырех аффинных преобразований. Каждое преобразование кодируется считанными байтами, хотя исходное изображение может быть любого размера.

- Signal Processing for Communications, by P.Prandoni and M. Vetterli,
© 2008, EPFL Press
- Гонсалес Р., Вудс Р. Цифровая обработка изображений. 2012
- В.А. Сойфер. Компьютерная обработка изображений. Часть 1.
1996