5 Методы стегоанализа и противодействие им

Классификация методов стегоанализа

- По известным стегоаналитику данным
  - только с известным носителем информации
  - с известным контейнером
  - с известной встроенной информацией
  - с выбранным контейнером
  - с выбранной встраиваемой информацией
- По решаемой задаче
  - Обнаружение встраивания
  - Оценка объёма встраивания
  - Восстановление встроенной информации

Классификация методов стегоанализа

- По ориентированности на определённый алгоритм
  - Целенаправленный стегоанализ (Targeted)
    - НЗБ, Jsteg, Outguess, F5, EzStego – Histogram Attack, Sample Pair Analysis, QIM steganalysis,…
  - Слепой стегоанализ (Blind)
    - Пространственная область, ДКП
    - Обычно основаны на методах машинного обучения с учителем. Но и методы другой категории зачастую тоже
- Методы на основе обучения включают этапы:
  - выбор информативных признаков
  - классификация векторов признаков с обучением

Но могут и совмещать два этих этапа в одном (на примере свёрточных нейронных сетей)

Идеология модельно-ориентированного стегоанализа

- Всегда можно описать контейнер некоторой моделью (математической).
- Хорошее стеговстраивание должно сохранять модель контейнера
- Тогда если выбранная модель полностью описывает класс контейнеров, то стегосистема является недетектируемой

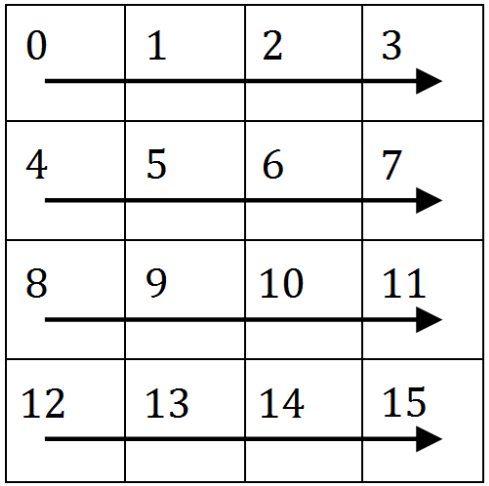
- Типичная простейшая модель – последовательность независимых одинаково распределённых случайных величин
  - Изображение описывается функцией распределения
  - Должна сохраняться гистограмма
  - Можно описать модель контейнера с использованием гистограммы в виде выражения $F(h) = 0$, где h – гистограмма

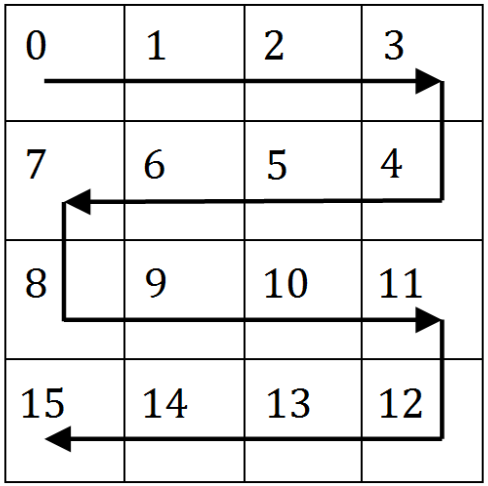
5.1 Методы НЗБ-стегоанализа

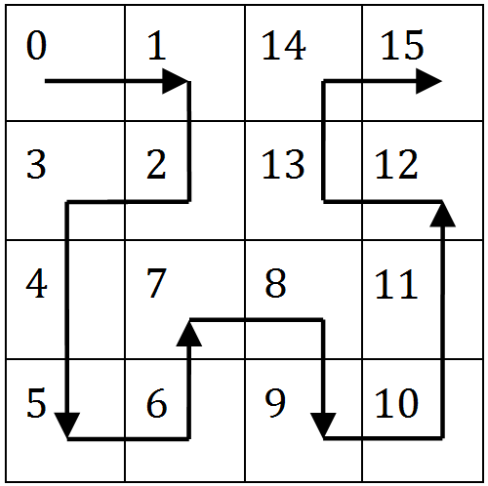
Основные идеи

- При стеганографическом встраивании информация подвергается шифрованию или сжатию без потерь, чтобы повысить защищённость, а также повысить ёмкость встраивания.
- **Следствие:** стеганографическое встраивание разрушает корреляционные связи между соседними отсчётами контейнера.
- Значит, при стегоанализе требуется обнаружить факт декоррелированности отсчётов матрицы признаков

Метод расчёта частоты переходов







- $\{\beta_k\}_{k=0..N-1}$, где $N = N_1 N_2$ - последовательность двоичных значений битовой плоскости, в которой мы хотим обнаружить следы встраивания

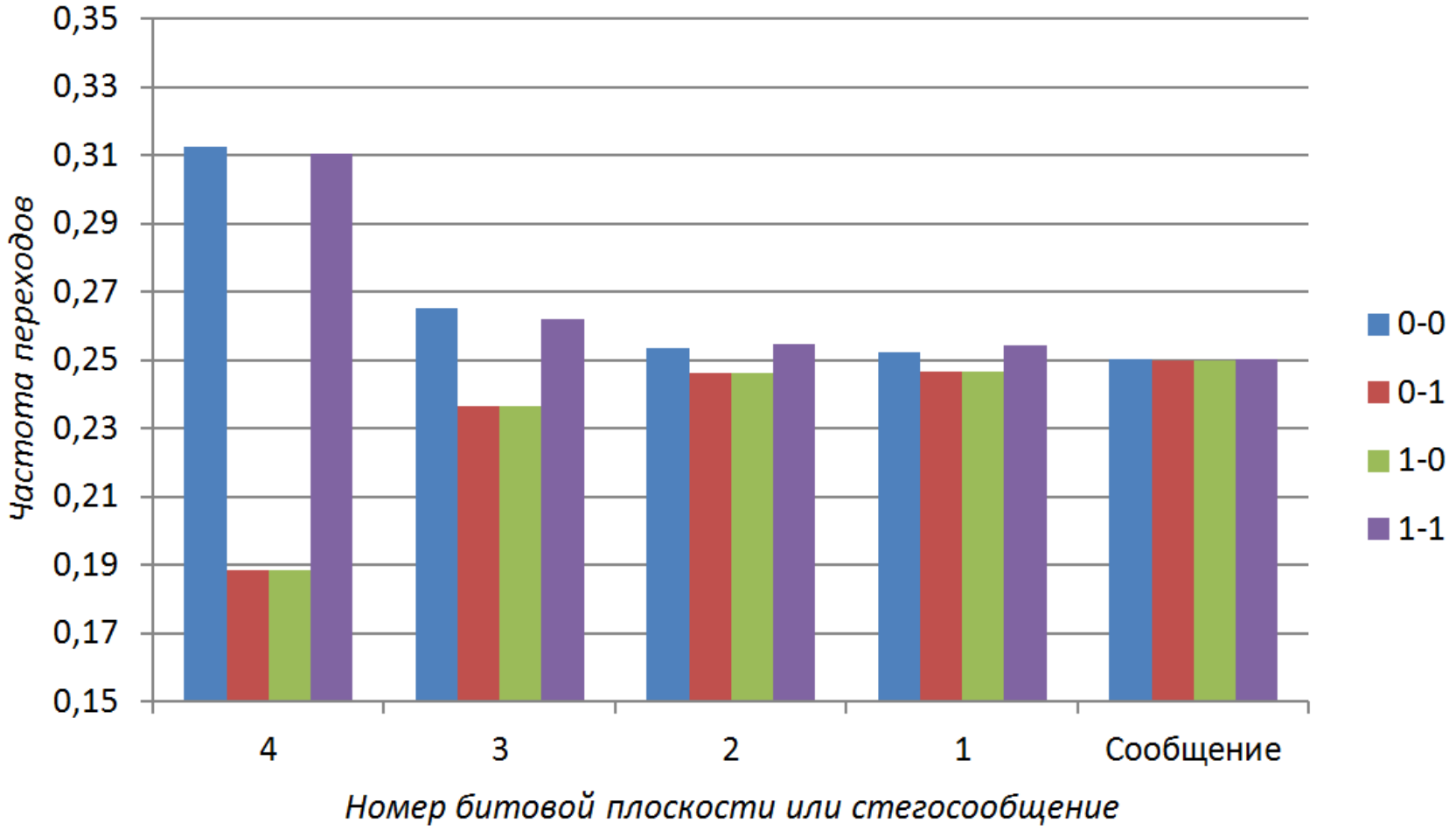
Метод расчёта частоты переходов

$$\pi_{00} = \frac{1}{N-1}\sum_{k=0}^{N-2}\gamma_k^{00},$$

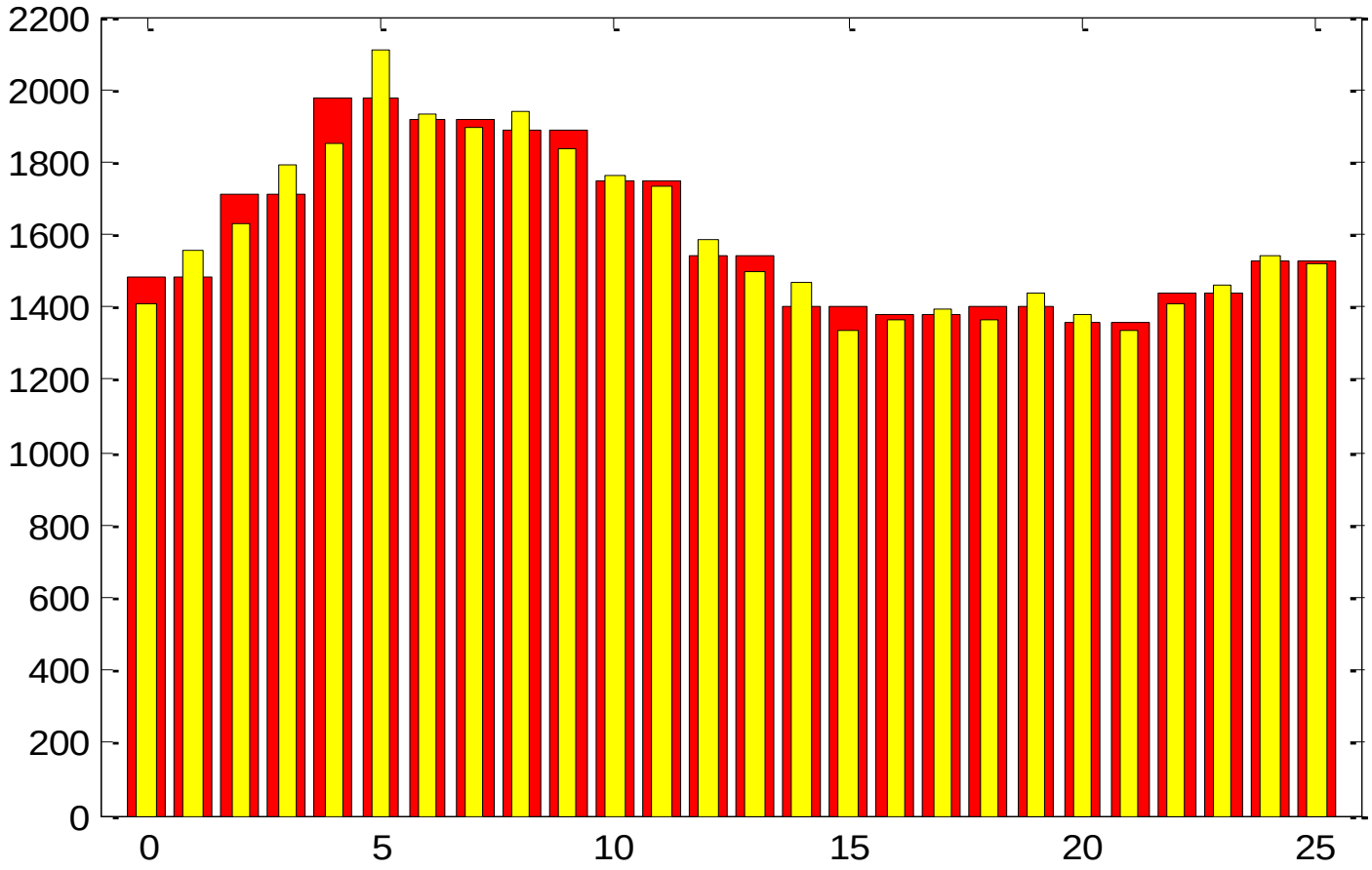
$$\gamma_k^{00} = \begin{cases} 1, & (\beta_k = 0) \wedge (\beta_{k+1} = 0), \\ 0, & \text{иначе.} \end{cases}$$

- По аналогии рассчитываются π_{01}, π_{10},π_{11}
- Итог – вектор из 4 признаков: $(\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$.

Метод расчёта частоты переходов



Метод гистограмм пар значений



Пример эмпирической гистограммы изображения (жёлтый цвет) и соответствующей ей теоретической гистограммы (красный цвет)

Метод гистограмм пар значений

- $h_i^e, i = 0..255$ - эмпирическая гистограмма анализируемого изображения
- Теоретическая гистограмма:

$$h_i^t = \frac{h_{2\cdot\lfloor i/2 \rfloor}^e + h_{2\cdot\lfloor i/2 \rfloor+1}^e}{2}.$$

- Расчёт статистики хи-квадрат для чётных отсчётов гистограммы:

$$\chi^2 = \sum_{i=0}^{127} \frac{\left(h_{2i}^t - h_{2i}^e\right)^2}{h_{2i}^t}$$

- Если условие ниже выполняется, значит информация встроена

$$\chi^2 < \chi_\alpha^2(k-1),$$

  - где α – уровень значимости,
  - $k - 1 = 127$ – число степеней свободы.

Противодействие методу ГПЗ: +-1-встраивание

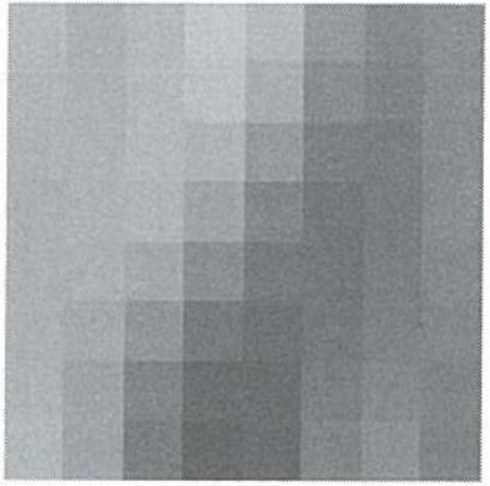
- $\{\xi_i\}$ – последовательность псевдослучайных чисел, с равной вероятностью принимающих положительные и отрицательные значения

$$sign(\xi_i) = \begin{cases} 1, & \xi_i \geq 0, \\ -1, & \xi_i < 0, \end{cases}$$

- $$C^W(n_1, n_2) = \begin{cases} C(n_1, n_2), & C_1(n_1, n_2) = b_i, \\ C(n_1, n_2) + sign(\xi_i), & C_1(n_1, n_2) \neq b_i, \end{cases}$$
- С учётом необходимости использования диапазона значений 0..255:
- $$C^W(n_1, n_2) = \begin{cases} C(n_1, n_2), & C_1(n_1, n_2) = b_i, \\ C(n_1, n_2) + 1, & C_1(n_1, n_2) \neq b_i \wedge C(n_1, n_2) = 0, \\ C(n_1, n_2) - 1, & C_1(n_1, n_2) \neq b_i \wedge C(n_1, n_2) = 255, \\ C(n_1, n_2) + sign(\xi_i), & \text{иначе.} \end{cases}$$

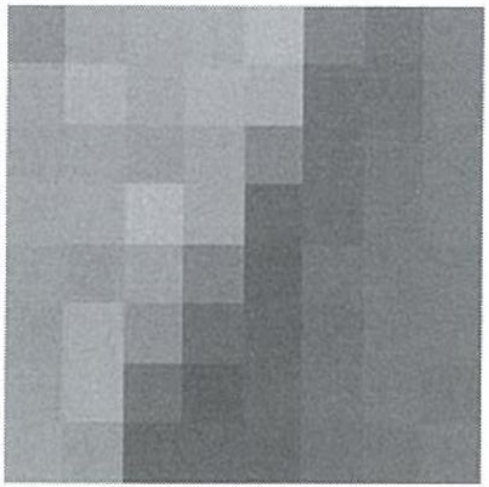
5.2 JPEG-стеганография и JPEG-стегоанализ

Статистика ДКП-коэффициентов



-1	6	3	-1	0	0	0	0
4	1	-4	-1	0	0	0	0
0	-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	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

QF=20

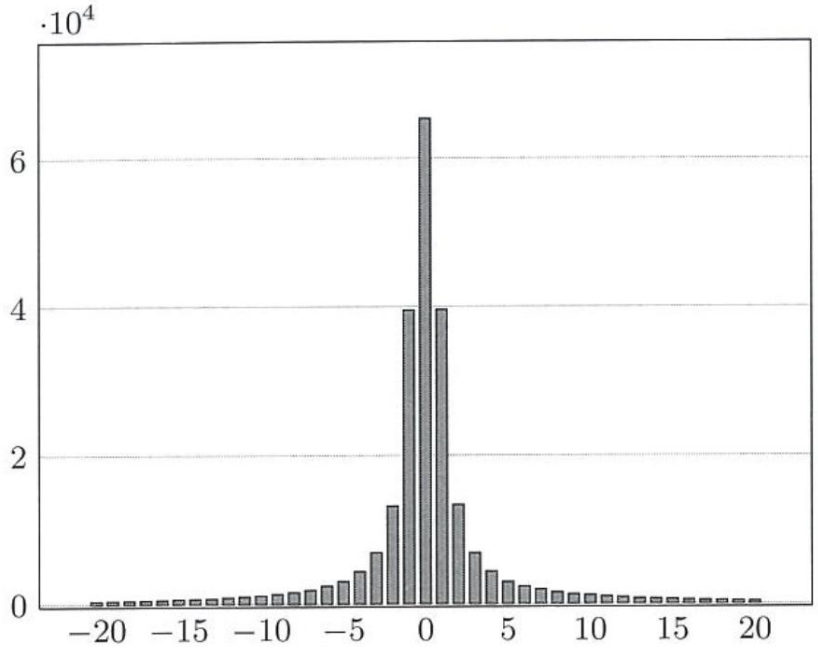


-16	90	37	-17	-1	-2	-2	-1
63	10	-46	-14	12	0	0	2
-2	-9	-5	12	4	-5	-2	1
1	-3	-2	0	-3	-1	1	1
0	-2	-1	-1	0	1	1	-1
0	0	0	0	-1	0	0	0
0	-1	0	0	0	0	0	0
0	-1	0	1	0	0	0	0

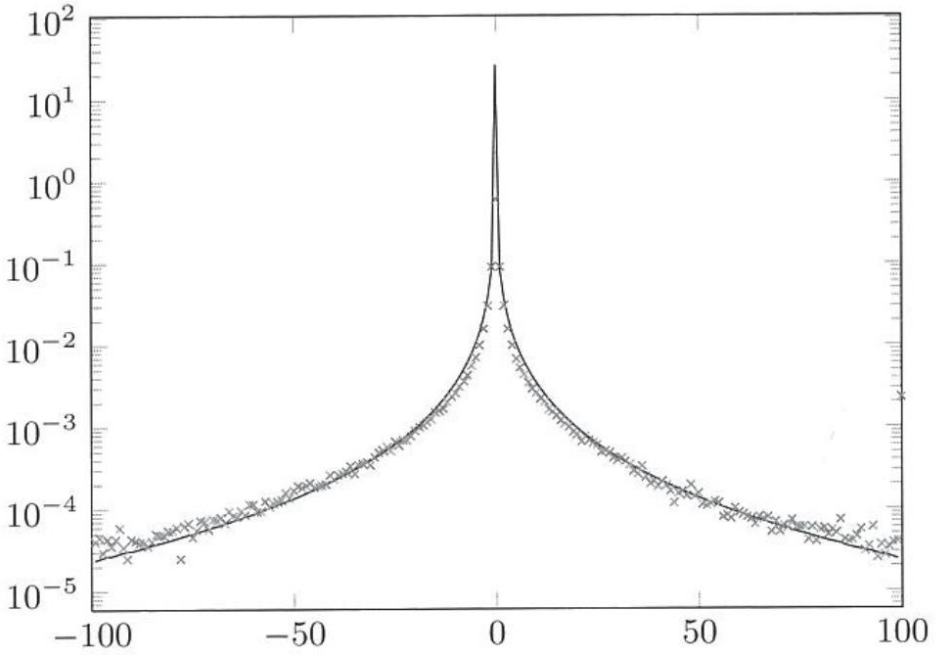
QF=90

Статистика ДКП-коэффициентов

Значения JPEG-коэффициентов в архиве могут лежать в диапазоне [-1023;1024]



Гистограмма ДКП-коэффициентов для QF=95



Гистограмма ДКП-коэффициентов для QF=95 и её аппроксимация функцией Гаусса (логарифмическая шкала)

JSteg

- НЗБ-встраивание для ДКП-коэффициентов
- Пара коэффициентов {0, 1} не используется для встраивания

Особенности гистограммы ДКП-коэффициентов и их использование

- $h_i \approx h_{-i}$
- $h_i > h_{i+1}$ для $i \geq 0$
- $h_{i-1} < h_i$ для $i \leq 0$

- Статистики НЗБ-пар: {-4,-3}, {-2,-1}, {2,3} и др. существенно различаются по статистике, но НЗБ-встраивание приводит к их выравниванию
- Можно использовать функцию следующего вида, близость значения которой к нулю говорит о вероятности отсутствия встраивания

$$F(h) = \sum_{k>0} h_{2k} + \sum_{k<0} h_{2k+1} - \sum_{k\geq 0} h_{2k+1} - \sum_{k<0} h_{2k}$$

Outguess

- НЗБ-встраивание для ДКП-коэффициентов
- Пара коэффициентов {0, 1} не используется для встраивания
- Далее двухпроходная процедура встраивания:
  - По ключу формируется подмножество коэффициентов D, в которые осуществляется встраивание информации
  - В остальные коэффициенты, не входящие в множество D, вносятся коррективы для выравнивания гистограммы
- Допустимая ёмкость встраивания 0.2 bpnc (bits per non-zero DCT coefficient). Это довольно мало

F5: основные идеи

- Цель – повысить ёмкость встраивания относительно Outguess
- [Забегая вперёд] В итоге алгоритм позволяет повысить ёмкость встраивания до 0.75 bpnc
- Другая форма НЗБ-встраивания:

$$LSB_{F5}(x) = \begin{cases} 1 - x\ (mod\ 2), & x < 0, \\ x\ (mod\ 2), & x \geq 0. \end{cases}$$

- При встраивании исключаем 0 и исключаем DC-отсчёт

F5: процедура встраивания

- Инициализируем $f^W = f$, j – индекс ДКП-коэффициента, i – индекс очередного бита сообщения
- Если $f(j) \neq 0$ и это не DC
  - Если $LSB_{F5}(f(j)) = b_i$, то $inc(i)$
  - Иначе
    - $f^W(j) = f(j) - sign\big(f(j)\big)$
    - Если $f^W(j) \neq 0$, то $inc(i)$

F5: свойства

- F5 не сохраняет гистограмму, но сохраняет ключевые её характеристики
- Почему происходит перекос?