Индекс структурированности

Александр Абгарян, abgarian.aa@phystech.edu

7 апреля 2024 г.

План презентации

1. Идея решения

2. Основные определения

3. Индекс структурированности

4. Корректность введенного индекса

Ссылки

Репозиторий с проектом

Введение

Задача: дано одноканальное изображение. Нужно определить его *индекс структурированности* — число, низкие значения которого соответствуют изображениям с очевидной структурой, высокие — случайному шуму без явной пространственной структуры. Чтобы определить индекс, нужно ответить на два вопроса:

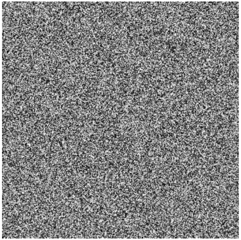
- Что такое шум?
- Чем отличается шум от незашумленного изображения?

Что такое шум?

Цвете пикселя черно-белого изображения I размера $w \times h$ задается функцией цвета пикселя: $X : I \to \{0, 1, \ldots 255\}$. В таком случае, из физических соображений, для шума должно выполняться:

$$\langle X \rangle \approx 255/2 = 127.5$$

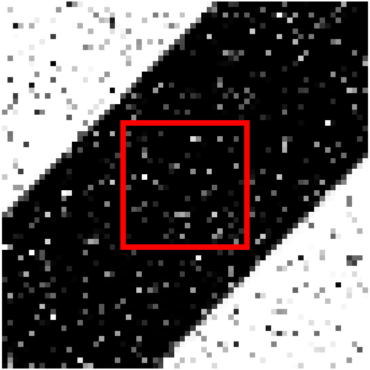
что характеризует его *случайный характер*. При этом дисперсия зашумленного изображения будет порядка $D \approx 5461.25$:

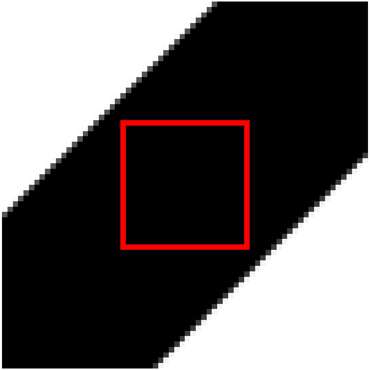


$$\langle X \rangle = 127.8$$

$$D = \langle X^2 \rangle - \langle X \rangle^2 = 5465.5$$

Шум vs незашумленное изображение





На более зашумленном изображении слева дисперсия в выделенном квадрате будет больше, чем дисперсия в том же квадрате на незашумленном изображении

Основные определения

Вместо дисперсии будем использовать конструкцию, похожую на норму в пространстве L^p:

Определение

Пусть I, J — изображения одинакового размера $w \times h$ с функциями цвета пикселя X и Y соответственно. Тогда число:

$$\rho(I, J) = \left(\frac{1}{hw} \sum_{i,j=1}^{w,h} \left| \frac{X(i,j) - Y(i,j)}{255} \right|^p \right)^{1/p} \tag{1}$$

будем называть L^p метрикой для изображений I и J.

Основные определения

Определение

Пусть I — изображение размера $w \times h$ с функцией цвета пикселя X. Положим $\mathcal{S} = \{S_{uv}\}_{u,v=1}^{m,n}$ — семейство непересекающихся прямоугольников, объединение которых дает изображение I. Тогда изображение J размера $w \times h$, в котором функция цвета пикселей:

$$Y(i,j) = \mathrm{agr}(X|_{S_{uv}}), \quad (i,j) \in S_{uv}$$

где agr — некая агрегирующая функция, будем называть *пулингом* изображения I с разбиением $\mathcal{S}$. Обозначение: $J = \mathrm{pool}(I, \mathcal{S}, \mathrm{agr})$.

В качестве агрегирующей функции можно брать среднее значение avg, минимум min или максимум max.

Определение индекса структурированности

Определение

Пусть I — изображение размера $w \times h$, $\mathcal{S}$ — его разбиение на непересекающиеся квадраты размера $k \times k$, дополненные при необходимости прямоугольниками, ρ — L^p метрика. Определим индекс структурированности как:

$$\mathrm{IND}(I, k) = \rho(I, \mathsf{pool}(I, \mathcal{S}, \mathsf{avg}))$$

Алгоритм:

- Разбиваем изображение на квадраты размера $k \times k$

Определение индекса структурированности

Определение

Пусть I — изображение размера $w \times h$, $\mathcal{S}$ — его разбиение на непересекающиеся квадраты размера $k \times k$, дополненные при необходимости прямоугольниками, ρ — L^p метрика. Определим индекс структурированности как:

$$\mathrm{IND}(I, k) = \rho(I, \mathrm{pool}(I, \mathcal{S}, \mathrm{avg}))$$

Алгоритм:

- Разбиваем изображение на квадраты размера $k \times k$
- В каждом квадрате S считаем среднее значение цвета $\langle X|_S \rangle$
- В новом изображении цвет каждого пикселя в квадрате $= \langle X|_S \rangle$

Определение индекса структурированности

Определение

Пусть I — изображение размера $w \times h$, $\mathcal{S}$ — его разбиение на непересекающиеся квадраты размера $k \times k$, дополненные при необходимости прямоугольниками, ρ — L^p метрика. Определим индекс структурированности как:

$$\mathrm{IND}(I, k) = \rho(I, \mathrm{pool}(I, \mathcal{S}, \mathrm{avg}))$$

Алгоритм:

- Разбиваем изображение на квадраты размера $k \times k$
- В каждом квадрате S считаем среднее значение цвета $\langle X|_S \rangle$
- В новом изображении цвет каждого пикселя в квадрате $= \langle X|_S \rangle$
- Считаем расстояние между оригиналом и полученым изображением по формуле (1)

Свойства введенного индекса

Определим $N(q, S) = \#\{(i,j) \in S \mid X(i,j) = q\}$ — количество пикселей цвета q в квадрате $S \in \mathcal{S}$. Тогда среднее значение пикселя в квадрате S:

$$\langle X|_S \rangle = \frac{1}{|S|} \sum_{(i,j) \in S} X(i,j) = \frac{1}{|S|} \sum_{q=0}^{255} \left[\sum_{(i,j):\, X(i,j)=q} q \right] = \frac{1}{|S|} \sum_{q=0}^{255} N(q, S) q$$

Пользуясь полученным выражением можно найти значение индекса:

$$\mathrm{IND}(I, k)^p = \frac{1}{hw} \sum_{S \in \mathcal{S}} \sum_{r=0}^{255} N(r, S) \left| \sum_{q=0}^{255} \frac{N(q, S)}{|S|} \frac{r - q}{255} \right|^p \tag{2}$$

Свойства введенного индекса

Пусть $(x_0, x_1, ... x_{255}) \in [0,1]^{256}$ такие, что $x_0 + ... + x_{255} = 1$. Введем функцию:

$$f(x_0, ... x_{255}) = \sum_{r=0}^{255} x_r \left| \sum_{q=0}^{255} x_q \frac{r-q}{255} \right|^p$$

Из определения $N(q, S)$ получаем:

$$\sum_{p=0}^{255} N(q, S) = |S| \Rightarrow \sum_{p=0}^{255} \frac{N(q, s)}{|S|} = 1$$

Свойства введенного индекса

Взяв в качестве $x_r = N(r,S)/|S|$ получим:

$$f(S) = \sum_{r=0}^{255} \frac{N(r,s)}{|S|} \left| \sum_{q=0}^{255} \frac{N(q,s)}{|S|} \frac{r-q}{255} \right|^p$$

Тогда выражение (2) можно записать следующим образом:

$$\mathrm{IND}(I,k) = \left(\sum_{S \in \mathcal{S}} f(S) \frac{|S|}{hw} \right)^{1/p}$$

Двухцветное изображение

Рассмотрим простейший случай: исходное изображение двухцветное, т.е. возможны лишь два цвета — 0 и 255. Тогда $\forall S \in \mathcal{S}$ и $\forall q \neq 0, 255: \ N(q, S) = 0$, и функция f принимает вид:

$$f(S) = \frac{N(0,s)}{|S|}\left(\frac{N(255,s)}{|S|}\right)^p + \frac{N(255,s)}{|S|}\left(\frac{N(0,s)}{|S|}\right)^p$$

Т.к. $N(0,S) + N(255,S) = 1$ то можно записать:

$$f(S) = g\left(\frac{N(0,s)}{|S|}\right), \quad g(t) = t(1-t)^p + (1-t)t^p$$

Двухцветное изображение

В дальнейшем нам понадобится

Утверждение

$g(t)$ выпукла при $p = 1, 2, 3$. Также при этих значениях p:

$$\max_{t\in[0,1]} g(t) = 2^{-p}$$

Из этого утверждения следует:

$$\mathrm{IND}(I, k) = \left(\sum_{S\in\mathcal{S}} g\left(\frac{N(0, s)}{|S|}\right)\frac{|S|}{hw}\right)^{1/p} \leq \left(\frac{1}{hw}\sum_{S\in\mathcal{S}}\frac{|S|}{2^p}\right)^{1/p} =$$

$$= \frac{1}{2}\left(\frac{1}{hw}\sum_{S\in\mathcal{S}}|S|\right)^{1/p} = \frac{1}{2}$$

Корректность введенного индекса

Дадим формальное определение шума

Определение

Будем говорить, что изображение I является *случайным*, если при генерации изображения цвет пикселя определяется случайно:

$$X(i,j) = \begin{cases} 0, & p_0(i,j) \\ 1, & p_1(i,j) \\ \vdots & \\ 255, & p_{255}(i,j) \end{cases}$$

где $p_k(i,j)$ — вероятность того, что пиксель (i,j) будет покрашен в цвет k. При этом выполняется $p_0(i,j) + p_1(i,j) + \ldots + p_{255}(i,j) = 1$.

Корректность введенного индекса

Определение

Случайное изображение I будем называть *шумом*, если для любого пикселя (i, j) и цвета k : $p_k(i, j) = 1/256$.

Утверждение

Пусть I — шум. Тогда в среднем, при размерности метрики p значение его индекса:

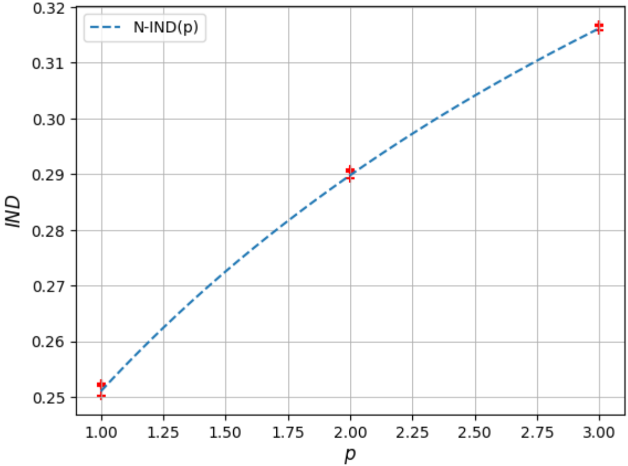
$$\mathrm{IND}(I, k) = \mathrm{N\text{–}IND}(p) = \frac{1}{255}\left(\frac{1}{128}\sum_{n=0}^{127}(127.5 - n)^p\right)^{1/p}$$

Получаем, что индекс шума не зависит ни от размеров изображения, ни от размера пулинга. Для $p = 1, 2, 3$ получаем:

$$\mathrm{N\text{–}IND}(1) \approx 0.25; \quad \mathrm{N\text{–}IND}(2) \approx 0.29; \quad \mathrm{N\text{–}IND}(3) \approx 0.32$$

Проверка утверждения об индексе шума

Для случайно сгенерированного шума проверим, что $\mathrm{IND}(I, k) = \mathrm{N\text{–}IND}(p)$ не зависит от размеров окна пулинга k.



Корректность введенного индекса

Для проверки индекса структурированности введем определение зашумленного изображения:

Определение

Пусть I — изображение размера $w \times h$ с функцией цвета пикселя X. Тогда случайное изображение J такого же размера с функцией цвета:

$$Y(i,j) = \begin{cases} X(i,j), & p \\ \text{другой цвет}, & (1-p)/255 \end{cases}$$

где p — вероятность того, что пиксель (i,j) не поменяет цвет, будем называть *зашумлением изображения* I.

Корректность введенного индекса

Корректность введенного индекса заключается в следующем: индекс зашумленного изображения будет больше индекса оригинала:

Теорема

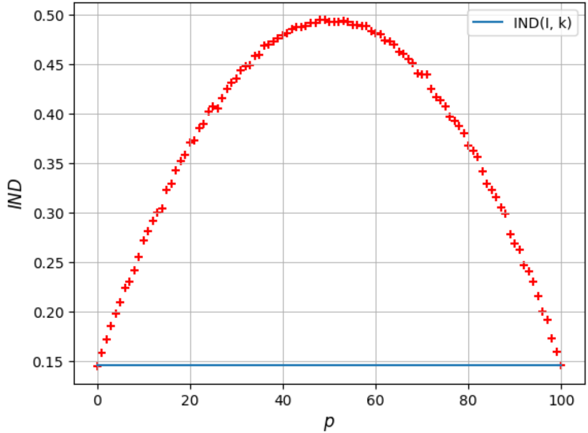
Пусть I — двухцветное изображение размера $w \times h$ с функцией цвета пикселя X. Пусть J — его зашумление с вероятностью $1 - p$, т.е. функция цвета пикселя изображения J:

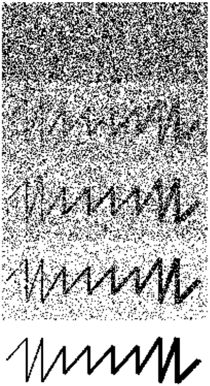
$$Y(i,j) = \begin{cases} X(i,j), & p \\ 1 - X(i,j), & 1 - p \end{cases}$$

Тогда в среднем, при размерности метрики $q = 1, 2, 3$, выполняется неравенство:

$$\mathrm{IND}(I, k) \leq \mathrm{IND}(J, k)$$

Проверка теоремы о двухцветном изображении

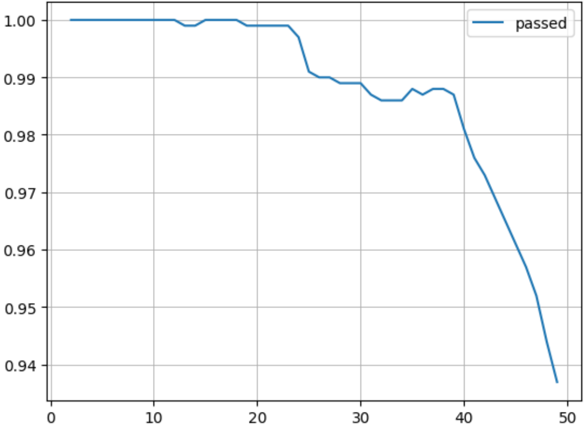




При значениях p близких к $1/2$ значение индекса $\mathrm{IND} \approx 1/2$, что также согласуется с теорией.

Проверка корректности индекса

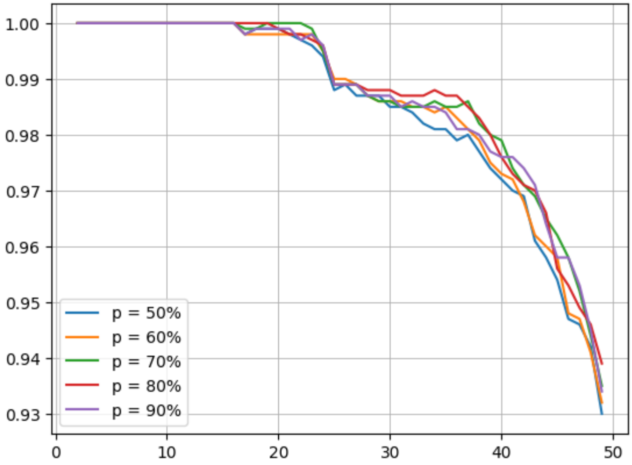
Датасет: 1000 черно-белых изображений размера 400×400
Первый тест: сравниваем индекс изображения с индексом его зашумления.



Построим зависимость процента верных ответов от размера окна пулинга.

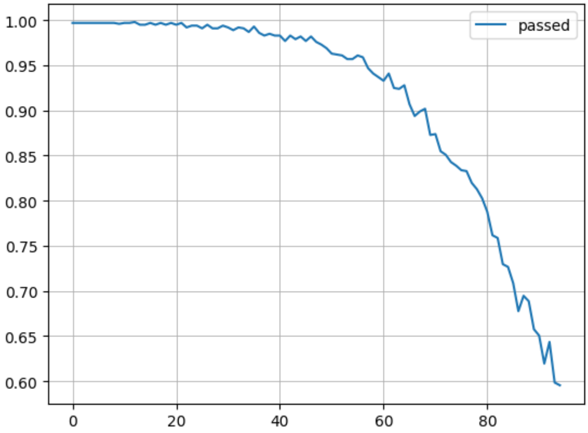
Проверка корректности индекса

На графике четко виден "провал" при размере окна пуллинга $\approx 25 = 50/2$.
При различных вероятностях зашумления он также наблюдается:



Проверка корректности индекса

Второй тест: сравниваем индекс изображения с индексом другого, случайно выбранного зашумленного изображения.



Процент угадываний в 90% достигается при вероятности зашумления $\geq 36\%$

Заключение

В результате проведенного исследования был предложен метод оценки зашумленности изображения (индекс структурированности). В ходе тестов было проверено:

1. При небольших размерах окна пулинга (меньше четверти размеров исходного изображения) индекс верно отличает изображение от его зашумления в 99% случаев.
2. Было доказано, что выполняется теорема о двухцветном изображении

Также можно отметить следующие недостатки:

1. Долгое время работы
2. Плохо работает при сравнении изображения с зашумлением другого изображения (хорошая вероятность верного ответа лишь при вероятность шума не менее 36%)