МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)

Индекс структурированности

Авторы:
Абгарян А. А.

Долгопрудный
2023 год

Содержание

1 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 приступить к определению индекса структурированности, нужно ответить на два вопроса: что такое шум, и что отличает зашумленное изображение от незашумленного? Ответ на первый вопрос можно дать из следующих соображений.

Пусть исходное черно-белое изображение I имеет размеры $w \times h$ пикселей. Обозначим через $X : I \to \{0, ...255\}$ функцию, возвращающую цвет каждого пикселя. Исходя из физических соображений, одним из признаков зашумленного изображения является то, что среднее значение:

$$\langle X \rangle \sim 255/2 = 127.5 \tag{1}$$

Это связано с тем, что в зашумленном изображении функция X распределена равномерно — цвет каждого пикселя *случаен* и не зависит от цвета других пикселей.

Теперь перейдем к вопросу о том, как отличить зашумленное изображение от незашумленного. Основная идея предложенного ниже метода заключается в следующем: если рассмотреть маленький кусочек незашумленного изображения, относящийся к четко выделенной структуре (например, линии), то суммарное отклонение цвета каждого пикселя от среднего значения цвета пикселя в этом кусочке будет *мало*. Если же мы рассмотрим маленький кусочек зашумленного изображения, то в виду случайности цвета каждого пикселя суммарное отклонение от среднего будет *велико* (см. рис 1).

Таким образом, основным фактором, отличающим шум от незашумленного изображения, согласно предположению, является *дисперсия цвета на локальном участке изображения*. Стоит сказать, что рассматривать дисперсию глобально нет смысла — она не зависит от расположения пикселей, а только от количества пикселей каждого цвета.

2 Основные определения

Под *функцией цвета пикселя* изображения I будем понимать функцию $X : I \to \{0, 1, \dots 255\}$, возвращающую цвет пикселя.

Определение 2.1. Пусть I, J — изображения одинакового размера $w \times h$ с функциями цвета пикселя X и Y соответственно. Тогда число:

$$\rho(I, J) = \left(\frac{1}{hw} \sum_{i,j=1}^{w,h} \left| \frac{X(i,j) - Y(i,j)}{255} \right|^p \right)^{1/p} \tag{2}$$

будем называть L^p расстоянием изображений I и J.

Мы хотим смотреть на дисперсию локально — для маленьких кусочков изображения, на которых может сохраняться структура. В связи с этим, введем

Определение 2.2. 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: $J = \mathrm{pool}(I, \mathcal{S}, \mathrm{agr})$.

3 Определение индекса структурированности

Сказанные выше идеи можно описать следующим образом:

Определение 3.1. 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:

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

Формально, предложенный алгоритм делает следующее:

1. Разбиваем изображение на квадраты размера $k \times k$
2. В каждом квадрате S считаем среднее значение цвета $\langle X|_S \rangle$
3. В новом изображении цвет каждого пикселя в квадрате $= \langle X|_S \rangle$
4. Считаем расстояние между оригиналом и полученым изображением по формуле (2)

Если воспользоваться определением метрики и пулинга, получим следующее выражение для индекса:

$$\mathrm{IND}(I, k)^p = \frac{1}{hw} \sum_{S \in \mathcal{S}} \sum_{(i,j) \in S} \left| \frac{X(i,j) - \langle X|_S \rangle}{255} \right|^p \tag{3}$$

Можно заметить одно свойство такого индекса — он не зависит от расположения пикселей в окне пулинга, а только от количества пикселей определенного цвета.

4 Свойства индекса структурированности

С выражением в форме (3) трудно работать, поэтому постараемся упростить его. 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 найдем значение индекса:

$$\mathrm{IND}(I, k)^q = \frac{1}{hw} \sum_{S \in \mathcal{S}} \sum_{(i,j) \in S} \left| \frac{X(i,j) - \langle X|_S \rangle}{255} \right|^q = \frac{1}{hw} \sum_{S \in \mathcal{S}} \sum_{r=0}^{255} N(r, S) \left| \frac{r - \langle X|_S \rangle}{255} \right|^q =$$

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

В полученном выражении прослеживается определенная симметрия. Чтобы еще более упростить его, сделаем следующее. $(x_0, x_1, \ldots x_{255}) \in [0, 1]^{256}$ такие, что $x_0 + \ldots + x_{255} = 1$. Введем функцию:

$$f(x_0, ...x_{255}) = \sum_{r=0}^{255} x_r \left| \sum_{q=0}^{255} x_q \frac{r-q}{255} \right|^p \tag{4}$$

Из определения $N(q,S)$ получаем:

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

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

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

Откуда окончательное выражение для индекса:

$$\mathrm{IND}(I,k) = \left(\sum_{S \in \mathcal{S}} f(S) \frac{|S|}{hw} \right)^{1/p} \tag{5}$$

5 Двухцветное изображение

Рассмотрим простейший случай: исходное изображение двухцветное, т.е. возможны лишь два цвета — белый 0 и черный 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:

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

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

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

Доказательство. Так как $g(t) = t(1-t)(t^{p-1} + (1-t)^{p-1})$ то при соответствующих значениях $p = 1,2,3$ получим:

$$g(t) = 2(t - t^2); \quad g(t) = t - t^2; \quad g(t) = -2t^4 + 4t^3 - 3t^2 + t$$

Найдем производную g:

$$g'(t) = 2(1-2t); \quad g'(t) = 1 - 2t; \quad g'(t) = (1-2t)^3$$

Получаем, что максимум g достигается при $t = 1/2 \Rightarrow \max_{t \in [0,1]} g(t) = 2^{-q}$. Найдем вторую производную g:

$$g''(t) = -4 \leq 0; \quad g''(t) = -2 \leq 0; \quad g''(t) = -6(1-2t)^3 \leq 0$$

Значит, g выпуклая для всех t. □

Из этого утверждения следует:

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

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

Причем максимум достигается при $N(0,S)/|S| = 1/2$ для всех S, т.е. когда количество белых и черных пикселей в каждом квадрате одинаково — это один из признаков шума.

6 Корректность индекса структурированности

Будем определять шум исходя из его случайного характера. Для этого введем:

Определение 6.1. 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 $p_0(i,j) + p_1(i,j) + ... + p_{255}(i,j) = 1$.

Теперь можно определить понятие шума:

Определение 6.2. Случайное изображение I будем называть *шумом*, если для любого пикселя (i,j) и цвета $k:\ p_k(i,j) = 1/256$.

Данное определение согласуется с реальностью. Действительно, для произвольного шума среднее значение цвета пикселя (а точнее, его матожидание):

$$\mathbb{E}\langle X\rangle = \frac{1}{hw}\sum_{(i,j)\in I}\mathbb{E}X(i,j) = \frac{1}{hw}\sum_{(i,j)\in I}\sum_{q=0}^{255} q\mathbb{P}(X(i,j)=q) =$$

$$= \frac{1}{hw}\sum_{(i,j)\in I}\sum_{q=0}^{255} qp_q(i,j) = \frac{1}{hw}\sum_{(i,j)\in I}\sum_{q=0}^{255}\frac{q}{256} = \frac{1}{hw}hw\cdot\frac{255\cdot 256}{2\cdot 256} = \frac{255}{2} = 127.5$$

Что согласуется с предположением (1). Согласно определению, найдем индекс шума:

Теорема 6.1. *Пусть I — шум. Тогда в среднем, при размерности метрики p значение его индекса:*

$$\mathrm{IND}(I,k) = \mathrm{N\text{–}IND}(p) = \frac{1}{255}\left(\frac{1}{128}\sum_{n=0}^{127}(127.5-n)^p\right)^{1/p} \tag{6}$$

Доказательство. Найдем матожидание $N(k,S)$ в произвольном квадрате $S \in \mathcal{S}$:

$$\mathbb{E}N(k,S) = \sum_{(i,j)\in S} \mathbb{P}(X(i,j)=k) = \sum_{(i,j)\in S} p_k(i,j) = \sum_{(i,j)\in S} \frac{1}{256} = \frac{|S|}{256}$$

Тогда:

$$f(S) = \sum_{r=0}^{255} \frac{1}{256} \left| \sum_{q=0}^{255} \frac{1}{256} \frac{r-q}{255} \right|^p = \sum_{r=0}^{255} \frac{1}{256} \left| \frac{r}{255} - \frac{1}{256} \sum_{q=0}^{255} \frac{q}{255} \right|^p =$$

$$= \sum_{r=0}^{255} \frac{1}{256} \left| \frac{r}{255} - \frac{1}{256} \cdot \frac{255 \cdot 256}{2 \cdot 255} \right|^p = \sum_{r=0}^{255} \frac{1}{256} \left| \frac{r}{255} - \frac{1}{2} \right|^p = \sum_{r=0}^{255} \frac{1}{256} \left| \frac{r - 255/2}{255} \right|^p =$$

$$= \sum_{r=0}^{255} \frac{1}{256 \cdot 255^p} |r - 127.5|^p = \frac{1}{256 \cdot 255^p} \left[\sum_{r=0}^{127} |r-127.5|^p + \sum_{r=128}^{255} |r-127.5|^p \right] =$$

$$= \frac{1}{128 \cdot 255^p} \sum_{r=0}^{127} (127.5 - r)^p$$

Откуда по формуле (5) получаем:

$$\mathrm{IND}(I,k) = \left(\sum_{S \in \mathcal{S}} f(S) \frac{|S|}{hw} \right)^{1/p} = \frac{1}{255} \left(\frac{1}{128} \sum_{n=0}^{127} (127.5 - n)^p \right)^{1/p}$$

□

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

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

Осталось доказать корректность определенного индекса, а именно то что для зашумленных изображений индекс будет больше, чем для менее зашумленных. Введем определение зашумленного изображения:

Определение 6.3. 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.

Пользуясь этим определением, докажем, что индекс зашумленного двухцветного изображения всегда больше индекса оригинала:

Теорема 6.2. *Пусть I — двухцветное изображение размера $w \times h$. Пусть J — его зашумление с вероятностью p, т.е. функция цвета изображения J:*

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

где $X(i,j)$ — функция цвета пикселя изображения I. Тогда в среднем, при размерности метрики $\hat{p} = 1, 2, 3$, выполняется неравенство:

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

Доказательство. Найдем среднее значение числа пикселей цвета 255 в произвольном квадрате S изображения J:

$$\mathbb{E}N_J(255,S) = \sum_{(i,j)\in S} \mathbb{P}(Y(i,j) = 255) = \sum_{X(i,j)=255} p + \sum_{X(i,j)=0} (1-p) =$$

$$= pN_I(255,S) + (1-p)N_I(0,S) = pN_I(255,S) + (1-p)(|S| - N_I(255,S))$$

Согласно формуле (5) индекс изображения I:

$$\mathrm{IND}(I,k)^{\hat{p}} = \frac{1}{hw}\sum_{S\in\mathcal{S}} |S| g(x(S)), \quad x(S) = \frac{N_I(255,S)}{|S|}$$

Так как:

$$y(S) = \frac{pN_I(255,S) + (1-p)(|S| - N_I(255,S))}{|S|} = x(S)p + (1-p)(1-x(S))$$

То индекс изображения J:

$$\mathrm{IND}(I,k)^{\hat{p}} = \frac{1}{hw}\sum_{S\in\mathcal{S}} |S| g(y(S)) = \frac{1}{hw}\sum_{S\in\mathcal{S}} |S| g(x(S)p + (1-p)(1-x(S)))$$

В силу утверждения (1) функция g выпукла вниз, а т.к. $p \in [0,1]$ то для нее выполняется:

$$f(x(S))p + (1-p)(1-x(S))) \geq pg(x(S)) + (1-p)g(1-x(S))$$

По определению $g(1-t) = g(t)$, откуда:

$$g(y(S)) = g(x(S)p + (1-p)(1-x(S))) \geq pg(x(S)) + (1-p)g(x(S)) = g(x(S))$$

А значит выполняется:

$$\mathrm{IND}(I,k)^{\hat{p}} = \frac{1}{hw}\sum_{S\in\mathcal{S}} |S| g(x(S)) \leq \frac{1}{hw}\sum_{S\in\mathcal{S}} |S| g(y(S)) = \mathrm{IND}(J,k)^{\hat{p}}$$

□