

МЕТОДЫ ОЦЕНИВАНИЯ КАЧЕСТВА ИЗОБРАЖЕНИЯ

Лекция 2

Пространственное и яркостное разрешение изображений

Пространственным разрешением изображения называется физический размер области, характеризуемой одним пикселем изображения. Например, стандартное пространственное разрешение отсканированных документов – 300 dpi (dots per inch), что означает, что один физический дюйм листа бумаги представлен на изображении 300 пикселями. Пространственное разрешение часто связывают с *четкостью* изображения и возможностью различения объектов на нем.

Яркостное разрешение. Эта характеристика определяется числом различных значений полутона (уровней серого),

Размеры изображения после оцифровки часто делаются равными степеням 2 (например, 1024x1024), что связано с особенностями используемых алгоритмов оцифровки.

Число градаций P серого цвета также обычно имеет степень 2 (например, если изображение имеет $P = 2^8 = 256$ градаций серого, то для хранения интенсивности одного пикселя требуется 8 бит, т.е. 1 байт).

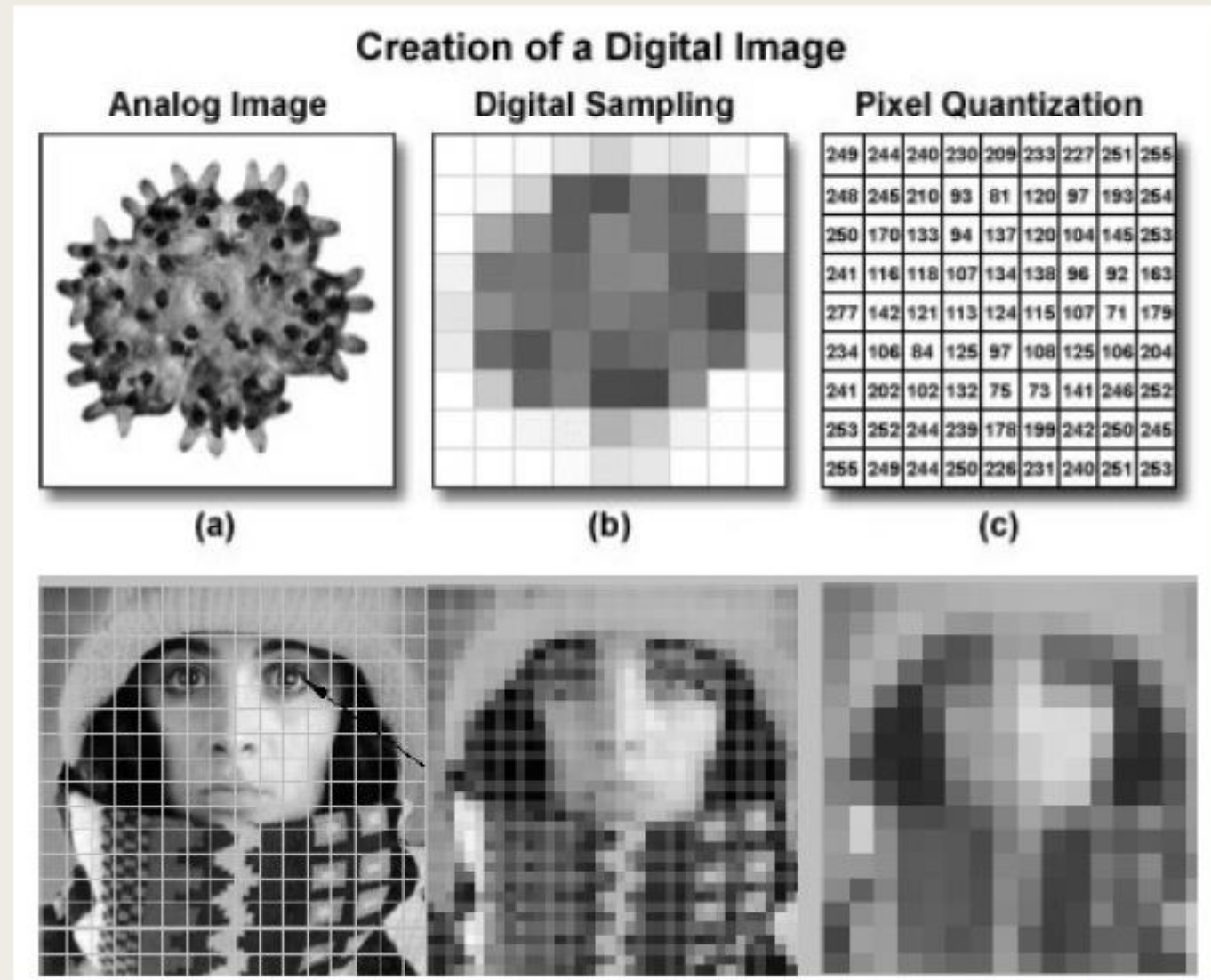
Яркость изображения

Двумерное изображение в реальном мире определяется как функция двух вещественных переменных $a(x, y)$, где a – амплитуда (например, яркости) изображения в точке с координатами (x, y) .

В результате дискретизации двумерное непрерывное изображение $a(x, y)$ разбивается на N строк и M столбцов. На пересечении строки и столбца расположен элемент изображения, называемый *пикселем* (*picture element*). Значение, присвоенное целочисленным координатам (m, n) , где $m = \overline{0, M-1}$, $n = \overline{0, N-1}$, – это соответствующая амплитуда $a(m, n)$ (яркость).

Оцифровка изображения. Иллюстрация

4



Статистические характеристики изображений

Для изображения $a(m, n)$, $m = \overline{0, M-1}$, $n = \overline{0, N-1}$, могут быть рассчитаны следующие статистические характеристики:

- минимальное и максимальное значения яркости:

$$a_{\min} = \min_{\substack{m = \overline{0, M-1} \\ n = \overline{0, N-1}}} a(m, n),$$

$$a_{\max} = \max_{\substack{m = \overline{0, M-1} \\ n = \overline{0, N-1}}} a(m, n);$$

- интервал значений яркости $[a_{\min}; a_{\max}]$, называемый *динамическим диапазоном* изображения;

Статистические характеристики изображений

- среднее значение яркости пикселей, иногда называемое просто *яркостью* изображения:

$$\bar{a} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a(m, n);$$

- дисперсия яркости:

$$D[a] = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (a(m, n) - \bar{a})^2;$$

- среднеквадратическое отклонение яркости:

$$C_{rms} = \sqrt{D[a]} = \sqrt{\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (a(m, n) - \bar{a})^2};$$

Статистические характеристики изображений

- **яркость** изображения
- **контраст** изображения

- **контраст изображения.** Если заметная доля пикселей занимает значительную часть всего диапазона уровней серого, то говорят, что изображение имеет высокий контраст. Наоборот, изображение с малым динамическим диапазоном имеет низкий контраст и обычно выглядит тусклым, размытым и серым. Увеличение динамического диапазона изображения делает его более контрастным.

Статистические характеристики изображений ⁸

Формально **контраст** определяется как относительный динамический диапазон. В зависимости от того, с чем сравнивается динамический диапазон, выделяют *контраст Михельсона* C_{Mich} и *глобальный контраст* C_g :

$$C_{Mich} = \frac{a_{\max} - a_{\min}}{a_{\max} + a_{\min}},$$

$$C_g = \frac{a_{\max} - a_{\min}}{P_{\max} + P_{\min}} = \frac{a_{\max} - a_{\min}}{P - 1},$$

где P_{\min} и P_{\max} – минимально и максимально возможные значения яркости соответственно. Как правило, $P_{\min} = 0$, $P_{\max} = P - 1$, где P – число возможных градаций яркости.

Контраст Михельсона сравнивает динамический диапазон с удвоенной оценкой средней яркости изображения, а глобальный контраст – с максимально возможной яркостью. Оба контраста принимают значения в интервале от 0 до 1.

Статистические характеристики изображений

Среднеквадратический контраст.

- Величина среднеквадратического отклонения яркости C_{rms} называется также *среднеквадратическим контрастом*.

$$C_{rms} = \sigma_a = \sqrt{D[a]} = \sqrt{\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (a(m, n) - \bar{a})^2}$$

где \bar{a} – средняя яркость пикселей изображения, называемая просто **яркостью изображения**:

$$\bar{a} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a(m, n)$$

Среднеквадратический контраст учитывает яркость **всех** пикселей изображения и, как правило, более информативен, чем глобальный контраст и контраст Михельсона



$$C_{rms} = 0.006$$



$$C_{rms} = 0.018$$

Изображения имеют один и тот же динамический диапазон (и следовательно, глобальный контраст), но разные среднеквадратические контрасты

Замечание

11

- **Яркость** представляет собой характеристику, определяющую то, насколько сильно цвета пикселей отличаются от черного. Например, если на оцифрованной фотографии преобладают светлые полутона, то ее яркость будет высокой. Если же темные, то низкой.
- **Контраст** представляет собой характеристику того, насколько большой разброс имеют яркости пикселей изображения. Чем он больше, тем больший контраст имеет изображение.
- Если проводить аналогию с математической статистикой, то **яркость** характеризуется математическим ожиданием, а **контраст** – дисперсией яркостей пикселей изображения.
- Яркость и контраст могут рассматриваться не только для всего изображения, но и для отдельных его фрагментов. Таким образом, возникают понятия *локальной яркости* и *локального контраста*.

Статистические характеристики изображений 12

■ Гистограмма яркости

Яркость изображения \Leftrightarrow математическое ожидание

Контраст изображения \Leftrightarrow среднеквадратическое отклонение

Гистограмма яркости $h(p)$ характеризует распределение пикселей изображения по значениям яркости:

$$h(p) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [a(m, n) = p], \quad p = 0, P-1,$$

где оператор $[A]$, примененный к логическому выражению A , возвращает значение 1 или 0 в зависимости от того, истинно это выражение или ложно:

$$[A] = \begin{cases} 1, & \text{если } A, \\ 0, & \text{если } \bar{A}. \end{cases}$$

Гистограмма яркости

13

Сумма высот всех столбцов гистограммы: $\sum_{p=0}^{P-1} h(p) = MN$

Нормализованная гистограмма яркости:

$$h_{norm}(p) = \frac{1}{MN} h(p)$$

$h_{norm}(p)$ – оценка вероятности того, что яркость случайно выбранного пикселя будет равна значению p

Гистограмма яркости изображения, часто используемая в методах улучшения качества изображения, определяется как вектор h , содержащий количества пикселей $h(p)$ изображения, обладающих заданными яркостями p , т.е. гистограмма яркости характеризует распределение пикселей по значениям яркости.

Эквализация гистограмм

14

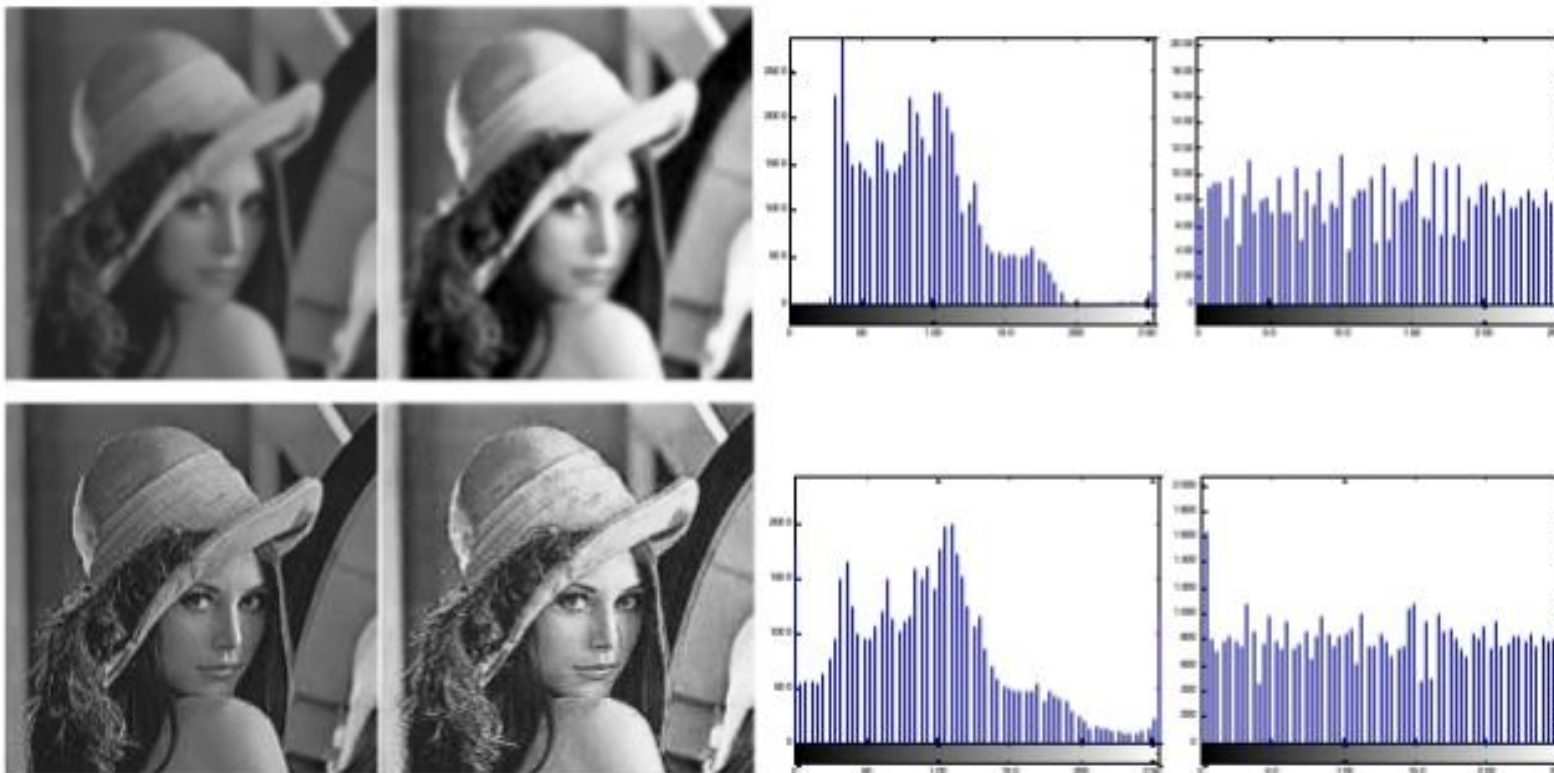


- Интуитивно можно сделать вывод, что наиболее удобным для восприятия человеком будет изображение, у которого гистограмма близка к равномерному распределению. Т.е. для улучшения визуального качества к изображению надо применить такое преобразование, чтобы гистограмма результата содержала все возможные значения яркости и при этом в примерно одинаковом количестве.

Эквализация гистограмм

15

Результат эквализации гистограммы



Сравним!

16



Тусклое изображение



Яркое изображение

Гистограмма яркости

17

Гистограмма (в изображениях) - это график статистического распределения элементов цифрового изображения с различной яркостью, в котором по горизонтальной оси представлена яркость, а по вертикали - относительное число пикселей с конкретным значением яркости.

В черно-белом изображении яркость непосредственно определяется значением пиксела. Чем больше значение пиксела (чем он светлее) - тем он ярче.

В цветных изображениях, обычно, каждый пиксел описывается тремя параметрами: долей красного в цвете, долей зеленого в цвете, долей синего в цвете. Тогда, для вычисления яркости используется взвешенная сумма каждого из значений пикселей.

Эквализация гистограмм

```
import cv2
import numpy as np
import matplotlib.pyplot as plt

# Загрузка изображения
image = cv2.imread('primer.jpg', cv2.IMREAD_GRAYSCALE)

# Применение эквализации гистограммы
equalized_image = cv2.equalizeHist(image)

# Нормализация изображения
normalized_image = cv2.normalize(equalized_image, None, alpha=0, beta=255,
                                norm_type=cv2.NORM_MINMAX, dtype=cv2.CV_8U)

# Вычисление гистограмм для каждого изображения
hist_original = cv2.calcHist([image], [0], None, [256], [0, 256])
hist_equalized = cv2.calcHist([equalized_image], [0], None, [256], [0, 256])
hist_normalized = cv2.calcHist([normalized_image], [0], None, [256], [0, 256])
```

Исходное изображение



Эквализация изображения



Нормализация изображения



Метрики качества изображения

20

Факторы, мешающие визуальному восприятию объектов на изображении:

шум, размытость, артефакты сжатия изображения и пр.

Image Quality Assessment (IQA) - подраздел анализа изображений, изучающий способы оценки качества изображений

Виды метрик качества изображения:

- **Full-reference quality metrics** | Референсные методы (сравнительные)
Основаны на сравнении изображения с "истинным" (незашумленным, референсным) изображением
- **No-reference (or objective-blind) quality metrics** | Не Референсные методы
Основаны на статистических характеристиках изображения

Зачем нужно оценивать качество изображения?

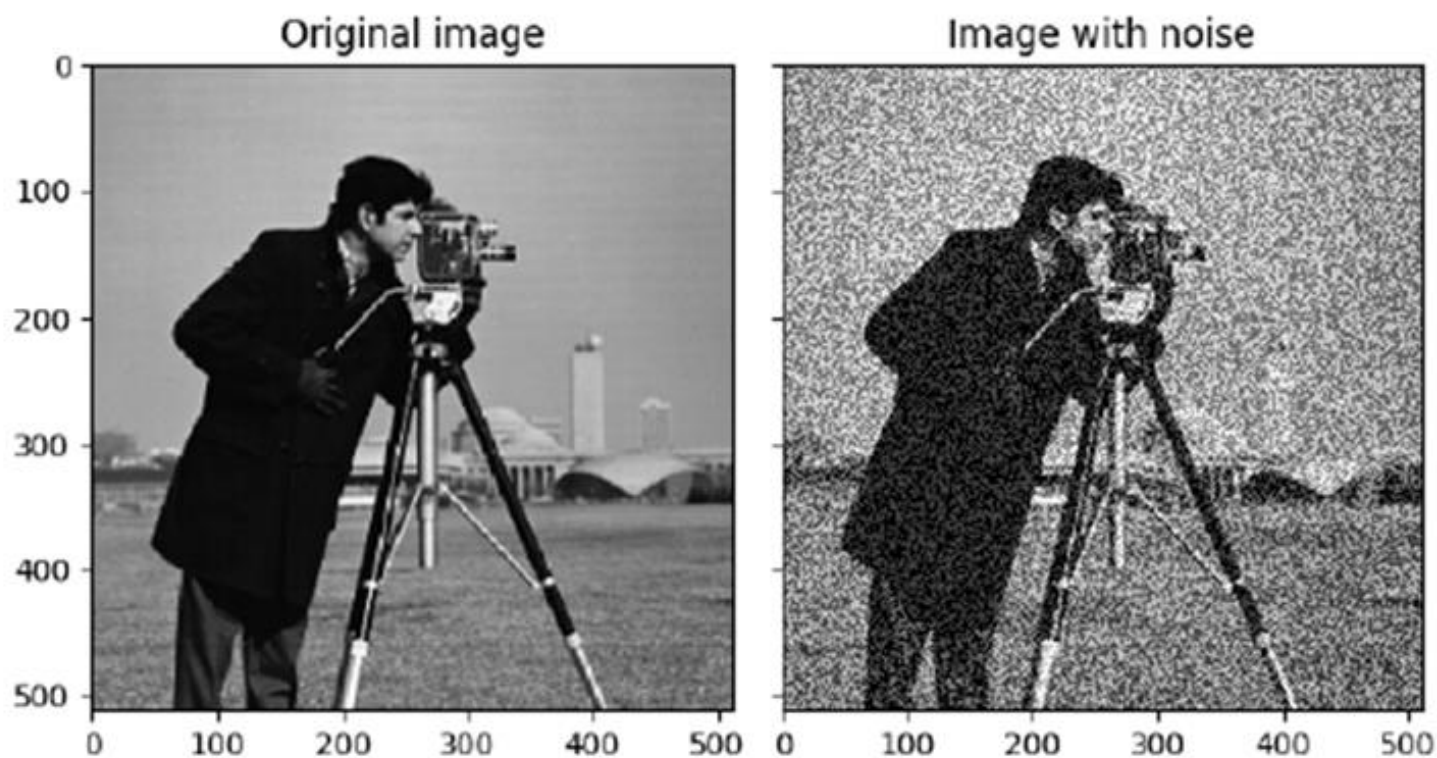
Например, для сравнения работы алгоритмов

- **Замечание.** Референсные методы - оценка качества изображения при наличии эталона.

Референсные методы

21

Full-Reference Quality Metrics



- Среднеквадратическая ошибка (**Mean-squared error, MSE**)
- Отношение сигнал-шум (**Signal-to-noise ratio, SNR**)
- Индекс структурной схожести (**Structural similarity (SSIM) index**)

Референсные методы

Среднеквадратическая ошибка

$$MSE = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (a(m, n) - a_0(m, n))^2$$

где $a_0(m, n)$ – незашумленное (референсное) изображение

Если изображение $a(m, n)$ наблюдается на фоне **аддитивного центрированного независимого шума** $\varepsilon(m, n)$ с интенсивностью σ^2 (например, $\varepsilon \sim N(0, \sigma)$):

$$a(m, n) = a_0(m, n) + \varepsilon(m, n)$$

то MSE – оценка **интенсивности** шума σ^2

Референсные методы.

$D[a]$ – дисперсия яркости

Отношение сигнал-шум

$$SNR = \frac{S^2}{\sigma^2}$$

где S^2 — величина, характеризующая контраст изображения или объекта по отношению к фону (например, $D[a]$),
 σ^2 — интенсивность шума на изображении

В роли S также могут быть:

- Средняя яркость изображения \bar{a}
- Максимально возможная яркость изображения P_{\max} (например, 1, 255 или 65535) (**peak SNR**) PSNR

Референсные методы

Логарифмическое отношение сигнал-шум

$$SNR = 10 \log_{10} \frac{D[a]}{\sigma^2} = 10 \log_{10} \frac{\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (a(m, n) - \bar{a})^2}{\sigma^2}$$

Пиковое логарифмическое отношение сигнал-шум:

$$pSNR = 10 \log_{10} \frac{P_{\max}^2}{\sigma^2}$$

Способы оценки интенсивности шума σ^2 :

- Как MSE между исходным и референсным изображением
- Как дисперсия яркостей пикселей участков изображения, принадлежащих фону
- С использованием **разностного изображения** (если имеется несколько изображений одного и того же объекта)

Референсные методы

25

Индекс структурного сходства (Structural similarity index):

Оценивается некоторая степень сходства между исходным изображением и его искаженной копией. Несет важную информацию о структуре объектов и о сцене в целом.

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)},$$

где

- μ_x — среднее x ,
- μ_y — среднее y ,
- σ_x^2 — дисперсия x ,
- σ_y^2 — дисперсия y ,
- σ_{xy} — ковариация x и y ,
- $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$ — две переменных:
 - L — динамический диапазон пикселей (обычно $2^{(\text{bits per pixel})} - 1$),
 - $k_1 = 0,01$ и $k_2 = 0,03$ — константы.

Ковариация является мерой совместной изменчивости двух случайных величин

Для случайных величин X и Y ковариация вычисляется по формуле:
 $\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$, где E — математическое ожидание. По аналогии с дисперсией ковариацию часто обозначают σ_{xy} .

Часто называется метрикой качества изображения, хотя в математически не является метрикой (4 аксиомы не выдерживает).

Референсные методы. Вариации SSIN

26

Индекс структурного сходства (Structural similarity index):

$$SSIM(a, b) = l(a, b)^\alpha c(a, b)^\beta s(a, b)^\gamma$$
$$l(a, b) = \frac{2\bar{a}\bar{b} + C_1}{\bar{a}^2 + \bar{b}^2 + C_1}, \quad c(a, b) = \frac{2\sigma_a\sigma_b + C_2}{\sigma_a^2 + \sigma_b^2 + C_2}$$
$$s(a, b) = \frac{\sigma_{ab} + C_3}{\sigma_a\sigma_b + C_3}$$

где $\bar{a}, \bar{b}, \sigma_a, \sigma_b$ – средние яркости и с.к.о. яркостей изображений a и b , σ_{ab} – ковариация яркостей, α, β, γ – параметры, C_1, C_2, C_3 – регуляризующие коэффициенты,

$$\sigma_{ab} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (a(m, n) - \bar{a})(b(m, n) - \bar{b})$$

l – яркостная компонента, c – контрастная компонента, s – структурная компонента.

Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. Image Quality Assessment: from Error Visibility to Structural Similarity // IEEE Transactions on Image Processing, 13, 4 (2004), 600-612
J.Nilsson, T.Akenine-Moller. Understanding SSIM. 2020 // <https://arxiv.org/abs/2006.13846>

Референсные методы. Вариации SSIN

27

- **SSIM** (Structural Similarity Index Measure) — метод, приближённый к человеческому восприятию, оценивающий структуру, яркость и контраст **изображений**.

Индекс структурного сходства. Замечания•

- Факторы $l(a, b)$, $c(a, b)$, $s(a, b)$ отвечают за **схожесть** в яркостях, контрасте и структурную схожесть изображений соответственно:

$$1 - l(a, b) = 1 - \frac{2\bar{a}\bar{b} + C_1}{\bar{a}^2 + \bar{b}^2 + C_1} = \frac{(\bar{a} - \bar{b})^2}{\bar{a}^2 + \bar{b}^2 + C_1}$$

- Структурная схожесть связана с **корреляцией** яркостей пикселей изображений
- Параметры α , β , γ определяют **важность** этих факторов
- При $C_1 = 0$, $C_2 = 0$, $C_3 = 0$:

$$0 \leq l(a, b) \leq 1, \quad 0 \leq c(a, b) \leq 1, \quad -1 \leq s(a, b) \leq 1$$

- $SSIM(a, b) = 1 \Leftrightarrow a = b$
- $SSIM(a, b) = 0 \Rightarrow$ отсутствие структурной схожести
- $SSIM(a, b) < 0 \Rightarrow$ изображения имеют инвертированную структуру

Вариации SSIN -> MSSIM

28

Карта индексов структурного сходства

Для $\alpha = \beta = \gamma = 1$, $C_3 = C_2/2$:

$$SSIM(a, b) = \frac{(2\bar{a}\bar{b} + C_1)(2\sigma_{ab} + C_2)}{(\bar{a}^2 + \bar{b}^2 + C_1)(\sigma_a^2 + \sigma_b^2 + C_2)}$$

По умолчанию:

$$C_1 = (0.01P)^2, \quad C_2 = (0.03P)^2, \quad C_3 = C_2/2$$

Рекомендуется рассчитывать $SSIM$ в локальных окнах \Rightarrow
карта $SSIM(m, n)$, $m = \overline{0, M-1}$, $n = \overline{0, N-1}$

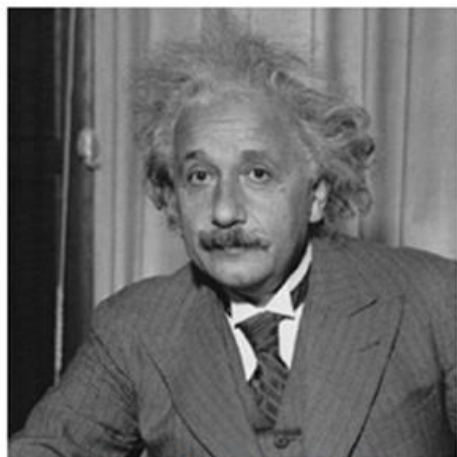
Mean Structural Similarity Index (MSSIM) – среднее значение карты $SSIM$:

$$MSSIM = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} SSIM(m, n)$$

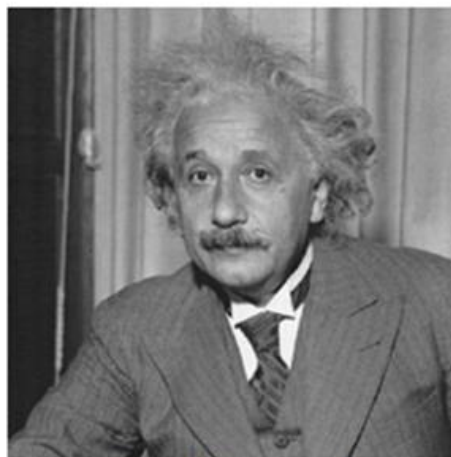
SSIN -> MSSIM

29

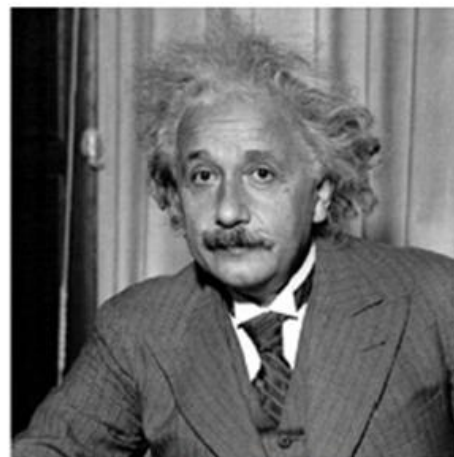
Индекс структурного сходства. Иллюстрации



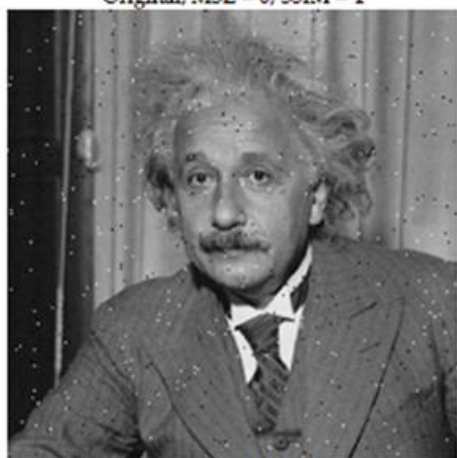
Original, MSE = 0; SSIM = 1



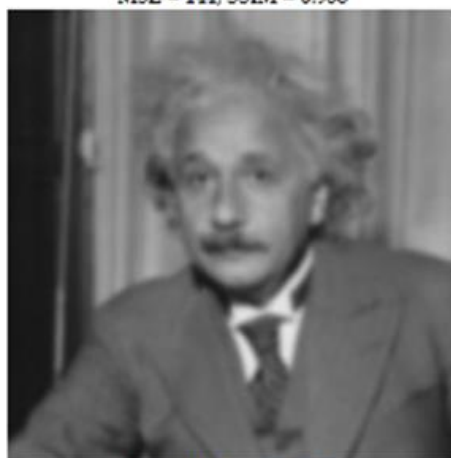
MSE = 144, SSIM = 0.988



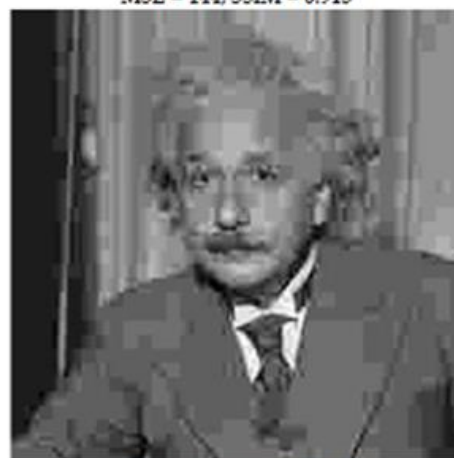
MSE = 144, SSIM = 0.913



MSE = 144, SSIM = 0.840

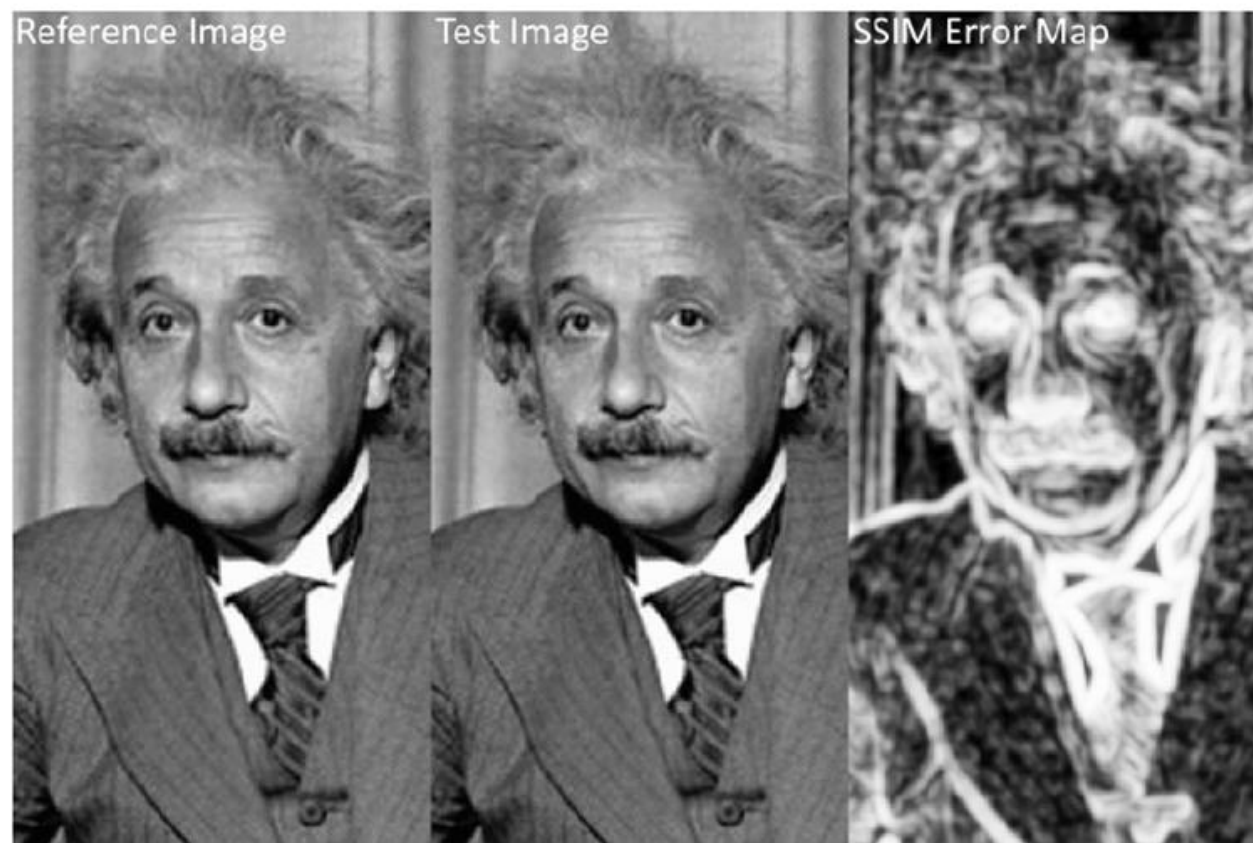


MSE = 144, SSIM = 0.694



MSE = 142, SSIM = 0.662

Карта индексов структурного сходства. Иллюстрации



Безреференсные метрики качества

No-reference (objective-blind) quality metrics основаны на использовании статистических характеристик изображения и **не требуют референсного изображения**

Виды no-reference quality metrics:

- Natural Image Quality Evaluator (NIQE)
- Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)
- Perception based Image Quality Evaluator (PIQE)

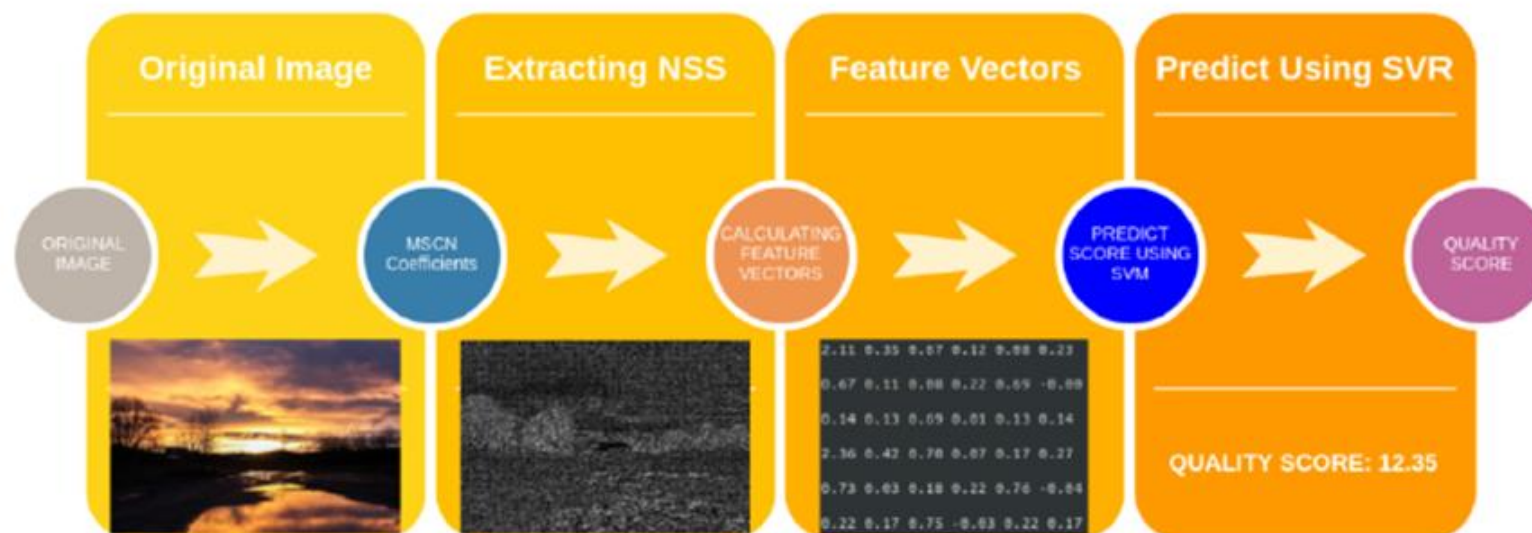
Метрики NIQE и BRISQUE используют **предобученные модели** для оценки качества изображения

Метрика PIQE использует **только статистику изображения**

Безреференсные методы

32

Схема NIQE и BRISQUE



Шаги методов NIQE и BRISQUE:

Шаг 1. Расчет статистических характеристик естественных сцен (Natural scene statistics, NSS)

Шаг 2. Расчет признаков изображения (calculating features)

Шаг 3. Расчет показателя качества (quality score)

Безреференсные методы

Natural Scene Statistics (NSS)

Статистические характеристики естественных сцен (Natural scene statistics, NSS) - низкоуровневые характеристики изображения, позволяющие отличить естественные изображения от искаженных и искусственных.

Виды NSS:

- Нормализованные локальные яркости (**Mean subtracted contrast normalized (MSCN) coefficients**)
- Коэффициенты дискретного косинусного преобразования (**Discrete cosine transform (DCT) coefficients**)
- Коэффициенты вейвлет-преобразования (**Wavelet transform coefficients**)

Известно, что MSCN-коэффициенты имеют нормальное распределение для естественных сцен.

Безреференсные методы

Mean Subtracted Contrast Normalized (MSCN) Coefficients

Карта MSCN:

$$MSCN(i, j) = \frac{a(i, j) - \bar{a}(i, j)}{\sigma(i, j) + C}$$

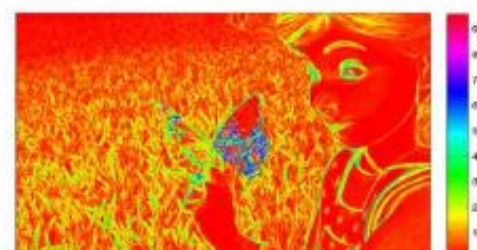
где $\bar{a}(i, j)$ и $\sigma(i, j)$ – среднее и с.к.о. яркости в окрестности пикселя (i, j) (например, размера 7×7), C – регуляризующий коэффициент



Original Image



MSCN Image



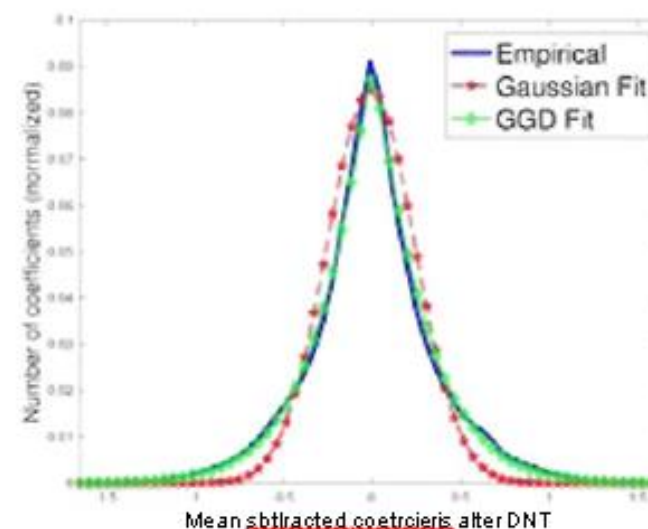
Standard Deviation Image

Безреференсные методы

35

Mean Subtracted Contrast Normalized (MSCN) Coefficients

Распределение MSCN-коэффициентов. Иллюстрация

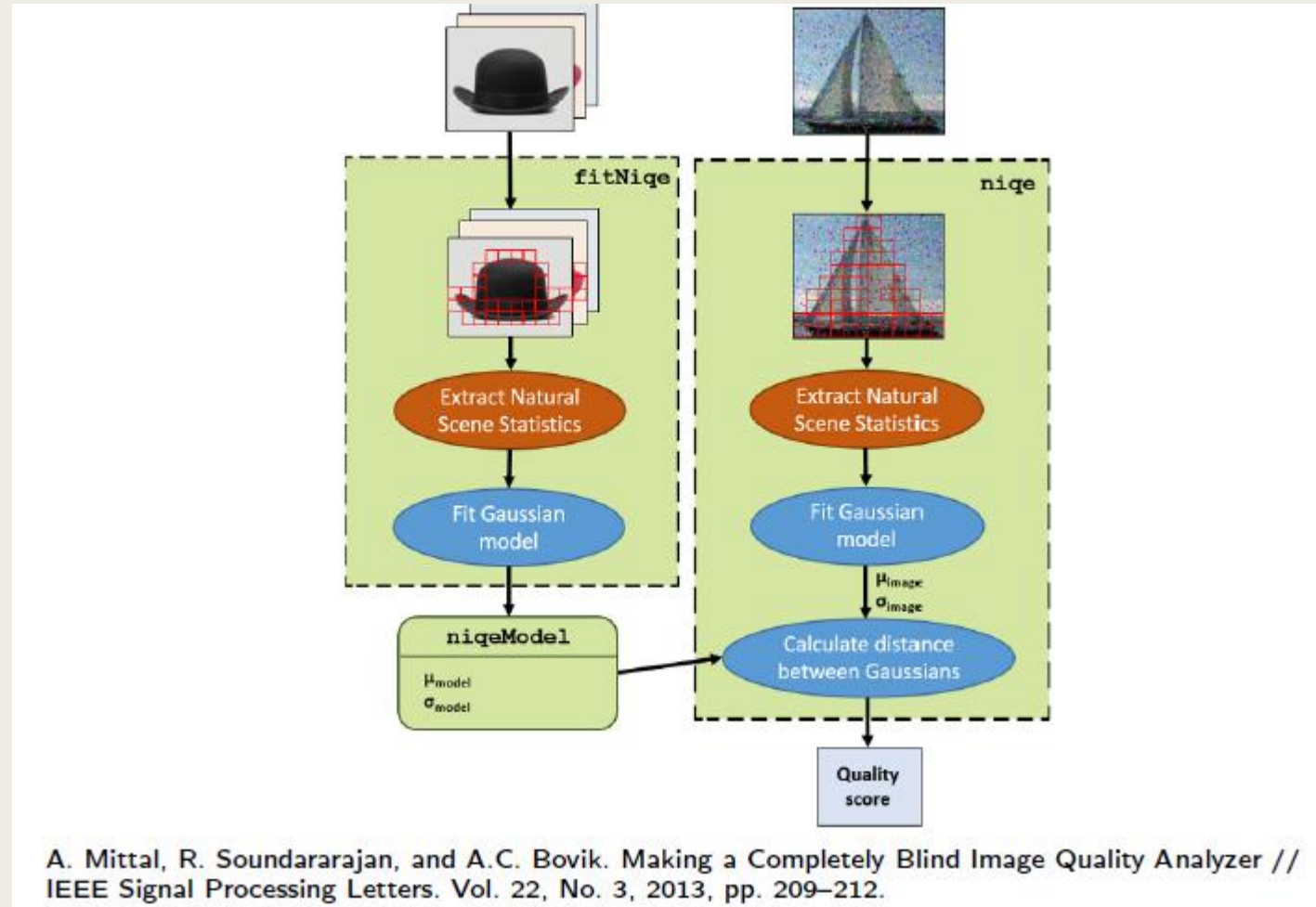


Признаки изображения (features) связаны с отклонением распределения NSS от нормального распределения

Безреференсные методы

36

■ АЛГОРИТМ NIQE

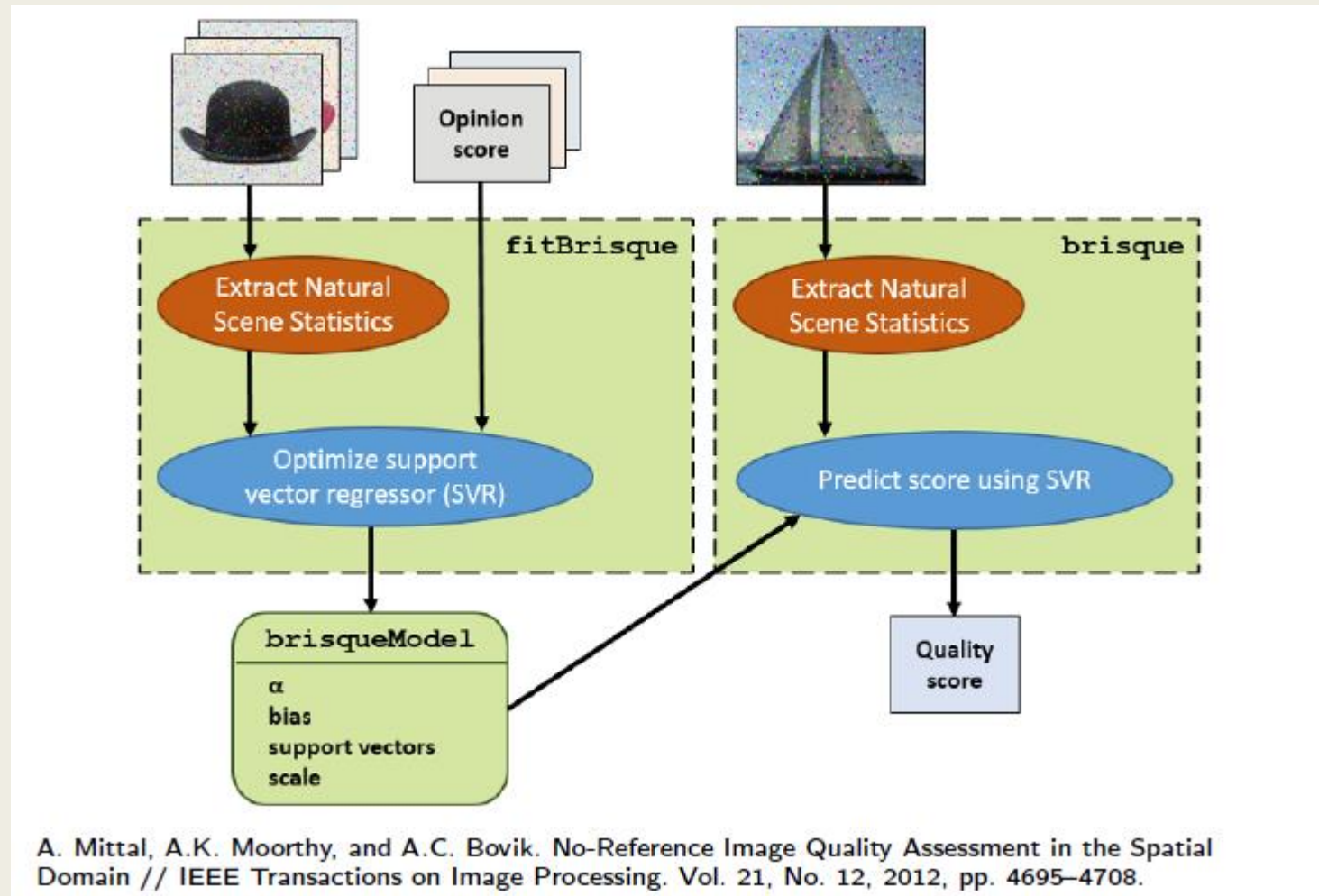


A. Mittal, R. Soundararajan, and A.C. Bovik. Making a Completely Blind Image Quality Analyzer // IEEE Signal Processing Letters. Vol. 22, No. 3, 2013, pp. 209–212.

Безреференсные методы

37

■ АЛГОРИТМ BRISQUE



Безреференсные методы

NIQE vs BRISQUE

- Алгоритмы BRISQUE и NIQUE отличаются тем, как они используют NSS для обучения модели и расчета показателя качества
- Значения метрики **NIQE** обычно неотрицательные числа. Меньшие значения означают лучшее качество изображения
- Для обучения модели **NIQE** требуется выборка естественных изображений
- Для обучения модели **BRISQUE** требуется размеченная выборка (изображение-показатель качества), как естественных, так и зашумленных изображений (например, датасеты TID2008, LIVE IQA и др.)
- Значения метрики **BRISQUE** обычно в интервале от 0 до 100. Меньшие значения означают лучшее качество изображения



(a) Reference Image
BRISQUE: 29.9521
SSIM: 1
PSNR: ∞



(b) Image w/ JPEG Compression
BRISQUE: 49.4060
SSIM: 0.5136
PSNR: 23.9622



(c) SRCNN Reconstruction
BRISQUE: 18.1413
SSIM: 0.3343
PSNR: 18.3930



(d) Reference Image
BRISQUE: 19.3352
SSIM: 1
PSNR: ∞



(e) Image w/ JPEG Compression
BRISQUE: 42.1659
SSIM: 0.6187
PSNR: 28.2818



(f) SRCNN Reconstruction
BRISQUE: 16.9035
SSIM: 0.4777
PSNR: 23.3253

Безреференсные методы

■ Алгоритм PIQE

Шаг 1. Расчет статистических характеристик естественных сцен (NSS) для каждого пикселя изображения (коэффициентов MSCN)

Шаг 2. Расчет статистики коэффициентов MSCN в локальных областях (например, 16×16 пикселей)

Шаг 3. Расчет показателей **активности**, **загрязненности артефактами** и **зашумленности** на основе статистики коэффициентов MSCN в каждой локальной области (формирование пространственных карт)

Шаг 4. Расчет показателя качества изображения как среднее значение показателей активности, загрязненности артефактами и зашумленности

N. Venkatanath, D. Praneeth, Bh.M. Chandrasekhar, et.al. Blind Image Quality Evaluation Using Perception Based Features// In Proceedings of the 21 st National Conference on Communications (NCC), 2015.

Original Image: PIQE score = 24.8481 | Noisy Image: PIQE score = 72.3643 | Blurred Image: PIQE score = 85.7362



Значения метрики PIQE обычно в интервале от 0 до 100.
Меньшие значения означают лучшее качество изображения

Пространственные карты качества

Алгоритм PIQE позволяет получить как глобальный показатель качества, так и **пространственные карты качества**



Недостаток: высокая вычислительная сложность

Как сравниваются изображения...

