

ОПРЕДЕЛЕНИЕ СТЕПЕНИ СЛОЖНОСТИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

© 2024 г. В. Б. Бокшанский^a, В. А. Кулин^a, Г. С. Финякин^{a, b},
А. С. Харламов^c, А. А. Шацкий^{a, *}

^aМосковский государственный технический университет имени Н.Э. Баумана

105005 Москва, ул. Бауманская 2-я, д. 5, стр. 1, Россия

^bНациональный исследовательский университет “МЭИ”

111250, Москва, вн. тер. г. муниципальный округ Лефортово, ул. Красноказарменная, д. 14, стр. 1, Россия

^cМосковский государственный технический университет гражданской авиации

125993, Москва, Кронштадтский бульвар, д. 20, Россия

*E-mail: shatskiyalex@gmail.com

Поступила в редакцию: 23.09.2023 г.

После доработки 19.02.2024 г.

Принята к публикации 13.05.2024 г.

Предложен новый метод оценки сложности геометрических фигур (пятен), учитывающий внутреннюю структуру пятен, а не только их внешний контур. Задача по вычислению степени сложности объектов разделена на составляющие: сегментация пятен и оценка сложности изолированных пятен. Новый метод обладает относительно низкой вычислительной сложностью по сравнению с рассмотренными в работе альтернативными методами. С помощью нового метода был создан алгоритм на основе параллельных вычислений языка CUDA для графических ускорителей (видеокарт), что дополнительно повышает быстродействие нашего метода. Проведен качественный и количественный анализ существующих (альтернативных) методов, выявлены их преимущества и недостатки по сравнению с нашим методом и друг с другом. Реализованный на основе нового метода алгоритм апробирован как на искусственных, так и на реальных изображениях.

Ключевые слова: инварианты Ну, сложность изображения, сжатие изображения, контур изображения, сегментация изображения, выделение контуров, классификация изображений, статистические моменты изображений, вычислительная сложность

DOI: 10.31857/S0132347424050036, **EDN:** OLZNYI

1. ВВЕДЕНИЕ

В экспериментальных исследованиях часто встаёт проблема поиска различных аномалий и новых особенностей среди большого количества экспериментальных данных.

В качестве примера можно привести автоматизированный поиск микроорганизмов (амёб, инфузорий и т. п.) в изображениях, полученных с оптических микроскопов. Для решения подобных задач существуют нейросетевые алгоритмы, но обучить нейросети распознавать все возможные типы (известных и неизвестных) микроорганизмов не представляется возможным, поэтому актуальной становится проблема поиска наиболее сложных аномалий в изображениях с микроскопов, что может являться альтернативным методом обнаружения.

Еще одним примером для поиска аномалий может быть фиксация пролета воздушных судов над аэропортами. Как показывают эксперименты, летящий самолет на фоне неба является относительно сложным объектом (аномалией) — в смысле, о котором пойдет речь далее. И на фоне неба самолет можно однозначно выделить как аномалию по сравнению с другими небесными объектами (Луна, Солнце, звёзды, спутники, метеоры и т. п.). Как показывают эксперименты, иногда самолет является даже более сложной аномалией, чем разнообразные наземные строения на фоне горизонта.

Также в качестве примера можно привести поиск аномалий в картах космологического (реликтового) фона. Эксперименты WMAP CMB [1] и Planck CMB [2] получили огромный массив данных таких наблюдений — аномалий

реликтового фона. Астрофизики и астрономы всего мира используют эти данные для самых разнообразных научных целей — от изучения ранней горячей Вселенной от стадии рекомбинации, до поиска топологических аномалий (космологических струн и червоточин). Этих данных настолько много, что их ручная сортировка попросту невозможна, и необходим автоматизированный алгоритм по поиску и отбору аномалий различного типа, по структуризации этих аномалий и разбиению их по типам и классам.

Наша работа предлагает новый метод для автоматизированного поиска и классификации аномалий в любых данных двумерного типа (на двумерных картах-изображениях). Мы разбиваем найденные аномалии по степени их сложности, вводя при этом математическое определение сложности двумерного изображения. При этом наш математический метод определяет сложность текущего двумерного объекта (на изображении) инвариантно по отношению к трем группам геометрических преобразований: к трансляциям, к масштабированию и к вращению.

Подобные исследования не являются принципиально новыми. Например в общепринятом подходе поиска аномалий в космологическом фоне принято разлагать космологический фон по мультиполям и строить спектр (гистограмму) такого разложения. Данный подход (разложения по мультиполям) имеет существенный недостаток — инвариантность мультиполей возможна только если полный монополярный заряд равен нулю. В разложении микроволнового излучения это достигается специальным выбором нулевого уровня яркости изображения, ниже которого яркость считается эффективно отрицательной, так что суммарная эффективная яркость изображения оказывается равной нулю. Такой выбор эффективной яркости возможен только благодаря плотному расположению пятен флуктуаций микроволнового фона друг относительно друга. Если площадь всех пятен на изображении много меньше площади всего изображения, то выбор нулевого уровня яркости приводит к обрезанию краёв пятен. Иногда и сами пятна полностью уходят под уровень фона.

В данной работе мы предлагаем новый метод определения аномалий на изображениях, а главное — вычисление уровня сложности таких аномалий (пятен). Помимо того, что наш метод является инвариантным к сдвигам, к масштабированию и к поворотам изображения, он еще и не зависит от выбора уровня фона на изображении (от выбора уровня нулевой яркости).

Потенциальные применения таких исследований:

- Поиск аномалий в изображениях, полученных с оптических микроскопов — для выявления на снимках микроорганизмов.
- Анализ медицинских снимков для выявления потенциальной онкологии и других заболеваний.
- Применение рентгеновской микротомографии для количественного и структурного анализа фармацевтических многокомпонентных систем.
- Космология (поиск аномалий реликтового фона).

2. ИНВАРИАНТЫ NU

Моменты изображения, или статистические моменты хорошо известны из математической статистики. Для дискретного (пиксельного) изображения в оттенках серого они определяются согласно формулам:

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \quad (1)$$

Здесь x и y — пиксельные координаты на изображении, $I(x, y)$ — яркость пикселя. Моменты M_{ij} не инвариантны ни к каким преобразованиям (за исключением моментов нулевого порядка M_{00}). Мы будем рассматривать моменты M_{ij} с общим порядком степеней координат не выше 3-й степени ($i + j \leq 3$). Таких моментов (далее обозначим их как M_{pq}) существует 10 штук.

С помощью моментов M_{pq} можно получить центр яркости изображения (математические ожидания по x и по y), его координаты определяются формулами:

$$\hat{x} := \frac{M_{10}}{M_{00}}, \quad \hat{y} := \frac{M_{01}}{M_{00}}. \quad (2)$$

С помощью (1) и (2) можно ввести центральные моменты:

$$\mu_{pq} = \sum_x \sum_y (x - \hat{x})^p (y - \hat{y})^q I(x, y) \quad (3)$$

Центральные моменты инвариантны относительно сдвига (трансляций). Ненулевых величин μ_{pq} 8 штук.

Можно также определить масштабные инварианты:

$$\eta_{pq} := \frac{\mu_{pq}}{\mu_{00}^{\frac{1}{1+0.5(p+q)}}}. \quad (4)$$

Величины η_{pq} инвариантны и к трансляциям, и к изменению масштаба — их тоже 8 штук.

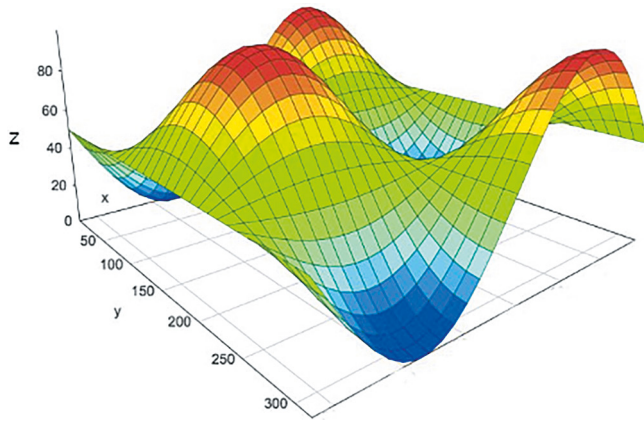
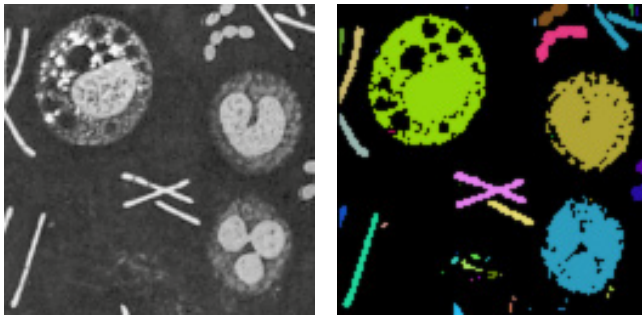
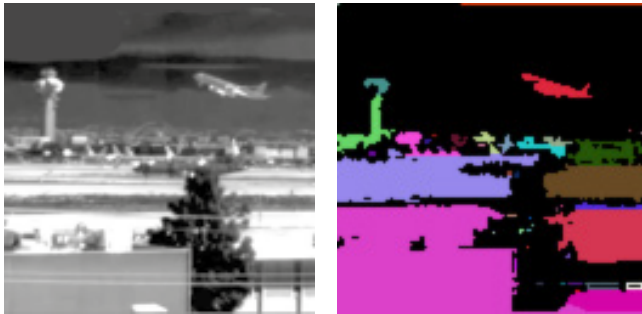


Рис. 1. Интенсивность яркости изображения.

(a) Микроорганизмы:



(b) Взлетающий самолет:



(c) Аномалии космологического фона реликтового излучения:

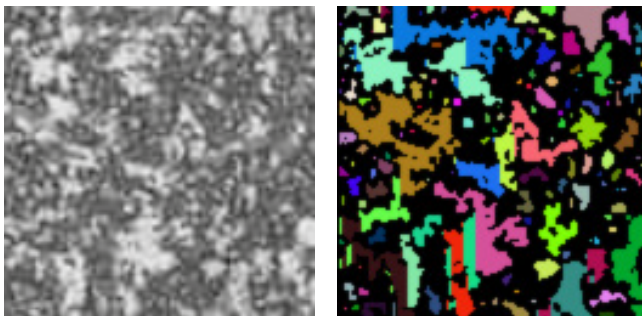


Рис. 2. Изображения в оттенках серого (слева), и сегментация этих же изображений по отдельным сегментам (объединяемых в пятна), выделенных разными цветами (справа).

С помощью выражений (1–4) можно вычислить еще 7 величин, которые будут дополнительно инвариантны еще и к поворотам. Они называются инвариантами Hu [3], [4]. Можно показать, что традиционный набор инвариантов Hu не является ни независимым, ни полным. Третий инвариант Hu_3 не очень полезен, так как он зависит от других инвариантов. В оригинальном наборе инвариантов Hu отсутствует инвариант независимого момента третьего порядка, а вместо него вводится 8-й инвариант Hu_8 .

Таким образом, всего можно определить 8 инвариантов Hu :

$$Hu_1 = (\eta_{20} + \eta_{02}). \quad (5)$$

$$Hu_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2. \quad (6)$$

$$Hu_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2. \quad (7)$$

$$Hu_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2. \quad (8)$$

$$Hu_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \times \quad (9)$$

$$\times \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right].$$

$$Hu_6 = (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}). \quad (10)$$

$$Hu_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \times \quad (11)$$

$$\times \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right].$$

$$Hu_8 = \eta_{11} \left[(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2 \right] - (\eta_{20} - \eta_{02})(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21}). \quad (12)$$

3. НАХОЖДЕНИЕ ПЯТЕН НА ИЗОБРАЖЕНИИ

Перед тем как исследовать конкретное пятно изображения на сложность сначала необходимо определить границы этого пятна. Для этого мы используем алгоритм **WaterShed** [5]. Суть алгоритма **WaterShed** – разбиение изображения на множество покрывающих его областей. За основу этого алгоритма был выбран алгоритм сегментации по водоразделам. Если значение яркости изображения откладывать по оси OZ , то в местах

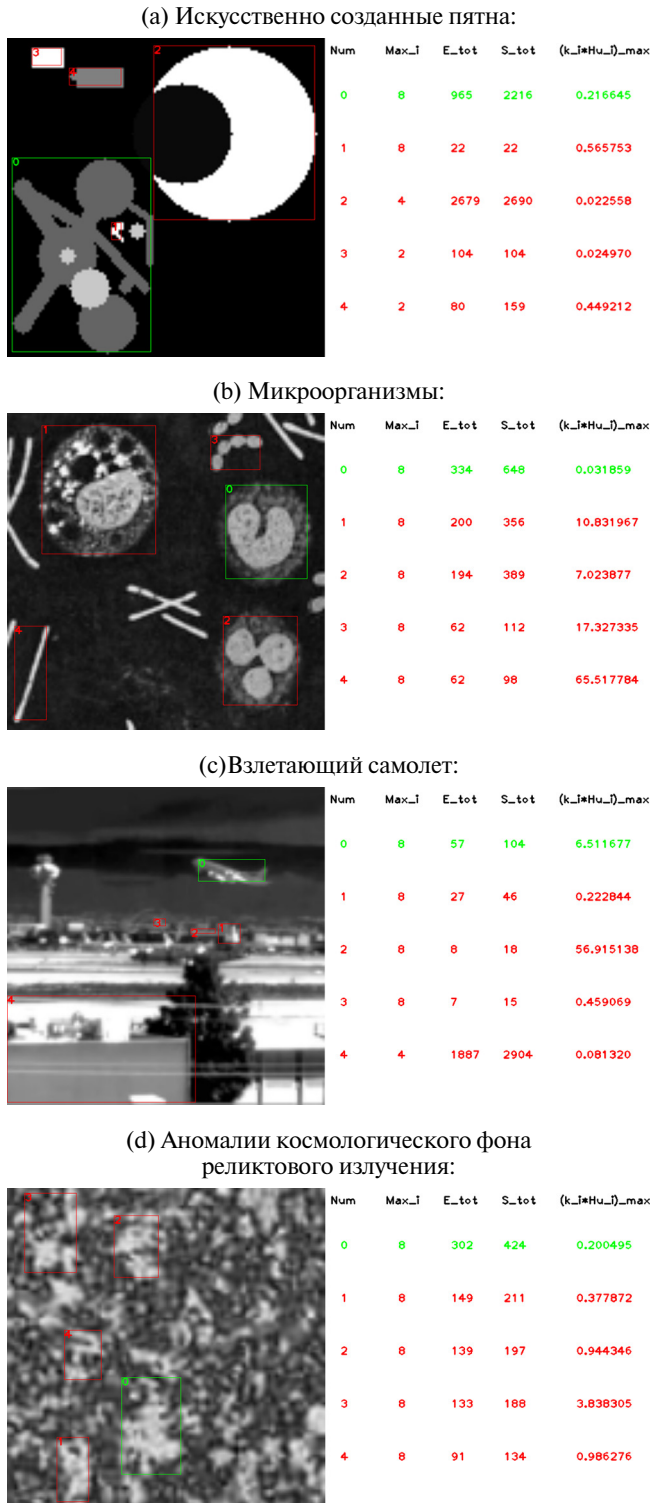


Рис. 3. Сортировка пятен изображения по убыванию уровня сложности и выделение главного пятна зелёной рамкой. Показан вывод топ-5 пятен изображения, ранжированных сначала по величине номера Max_i у величины $|k_i Hu_i|_{max}$, а потом по яркости пятна. Справа приведена таблица характеристик пятен: номер пятна, номер Max_i у максимальной величины $|k_i Hu_i|_{max}$, яркость пятна E_{tot} , площадь пятна S_{tot} и значение $|k_i Hu_i|_{max}$.

наибольшей интенсивности образуются “хребты”, наименьшей — “впадины”, а в однородных регионах — равнины (см. рис. 3).

Объектом будет считаться некоторая область, состоящая в общем случае из пикселей разной интенсивности (отдельное пятно изображения). Обязательное условие для данной области, что интенсивность пикселей, находящихся на краю пятна выше некоторого значения фона и меньше или равно интенсивности пикселей объекта граничащих с ними. Иными словами, выделяемый объект по сути это пятно, выделяющееся на некотором фоне. Причем изменением значения фона можно регулировать уровень чувствительности алгоритма **WaterShed**. Яркость пикселей ниже уровня фона алгоритмом считается нулевой яркостью.

Изображения на левых частях рис. 3 представляют из себя оригинальные кадры, содержащие объекты исследования. На правых частях этого же рисунка разными цветами обозначены сегменты пятен, найденные алгоритмом **WaterShed**.

Для определения сложности объектов следует сегментировать изображение, а затем найти на нем кандидаты на объекты (отдельные пятна) для последующей обработки. Сегменты с разной яркостью, касающиеся друг друга хотя бы в одном пикселе, должны быть объединены в одно пятно. Сегментация отдельных пятен может быть затруднена из-за высокой зашумленности изображения, поэтому первой задачей является определение уровня фона, а уже потом поиск сегментов и отдельных пятен.

4. УРОВЕНЬ СЛОЖНОСТИ ПЯТНА, ОПРЕДЕЛЕННЫЙ ЧЕРЕЗ ИНВАРИАНТЫ Hu

Дискретный уровень сложности пятна, выраженный через инварианты Hu_3 можно определить как номер максимального элемента вектора:

$$Compl = \left\{ |k_1 Hu_1|, |k_2 Hu_2|, |k_3 Hu_3|, |k_4 Hu_4|, |k_5 Hu_5|, |k_6 Hu_6|, |k_7 Hu_7|, |k_8 Hu_8| \right\} \quad (13)$$

Здесь k_i — специально подобранные нормировочные коэффициенты.

Коэффициенты k_i подбираются эмпирическим путем, главным критерием для такого подбора является вид гистограммы величин $|k_i Hu_i|$, состоящей из восьми бинов. При выборе коэффициентов k_i мы исходили из того, что наиболее простые геометрические фигуры (круг, или окружность) должны иметь максимум гистограммы, смещенный максимально влево (к первому

бину), а наиболее сложные геометрические фигуры — наоборот: максимум гистограммы должен соответствовать восьмому бину.

Стоит отметить, что пятна с центрально-симметричным распределением яркости имеют только первый ненулевой инвариант Hu_1 . Пятна с осевой симметрией имеют ненулевые инварианты Hu с номерами не выше четвертого. Максимально асимметричные пятна содержат все ненулевые инварианты Hu .

Один из возможных, подобранных по этому принципу, вариантов набора коэффициентов k_i может быть следующим:

$$k_i = \{1, 3.5, 12.25, 43, 150, 525, 1840, 6430\}.$$

Если несколько пятен будут иметь одинаковую степень сложности, то эти пятна между собой можно будет ранжировать по яркости пятна, либо по площади пятна, либо по номеру Max_i бина гистограммы, соответствующего её максимуму: $|k_i Hu_i|_{max}$, либо по комбинации этих величин.

5. РЕЗУЛЬТАТ РАБОТЫ АЛГОРИТМА

В результате работы алгоритма **WaterShed** мы получаем маску с сегментированным изображением (правая часть рис. 3), где пиксели одного сегмента помечены одинаковой меткой и образуют связную область. Основным недостатком данного алгоритма является использование процедуры предварительной обработки для картинок с большим количеством локальных минимумов (изображения со сложной текстурой и с обилием различных цветов). Для ускорения работы, изображение обрабатывается на графическом ускорителе (видеокарте). При этом, каждый пиксель обрабатывается своим собственным потоком с использованием архитектуры CUDA, а информация о пикселях сохраняется в память видеокарты. В итоге после обработки изображения мы получаем массив структур с данными обо всех объектах (пятнах) на изображении. После обработки этого массива структур с информацией о сегментированном объекте далее могут быть вычислены центральные статистические моменты всех пятен изображения, необходимые для вычисления инвариантов Hu .

На рис. 4 показаны результаты работы алгоритма по вычислению сложности на искусственно-созданных пятнах, а также на изображениях реальных объектов: микроорганизмов, взлетающего самолета и аномалии космологического фона реликтового излучения.

Исходя из этого, для выделения наиболее приоритетного объекта, следует в первую очередь

выбрать объекты с максимальным значением номера Max_i у величины $|k_i Hu_i|_{max}$, а затем отдать предпочтение объекту с наибольшим значением одного из параметров¹: E_{tot} , S_{tot} , $|k_i Hu_i|_{max}$. Какой из них выбрать, можно решить экспериментальным путем, в зависимости от поставленной задачи. Также можно использовать все три параметра E_{tot} , S_{tot} , $|k_i Hu_i|_{max}$ путем голосования по большинству.

6. АЛЬТЕРНАТИВНЫЕ МЕТОДЫ ОПРЕДЕЛЕНИЯ СЛОЖНОСТИ

Про метод максимального мультиполя было написано во введении. Здесь мы хотим привести еще несколько методов определения сложности пятен и сравнить их:

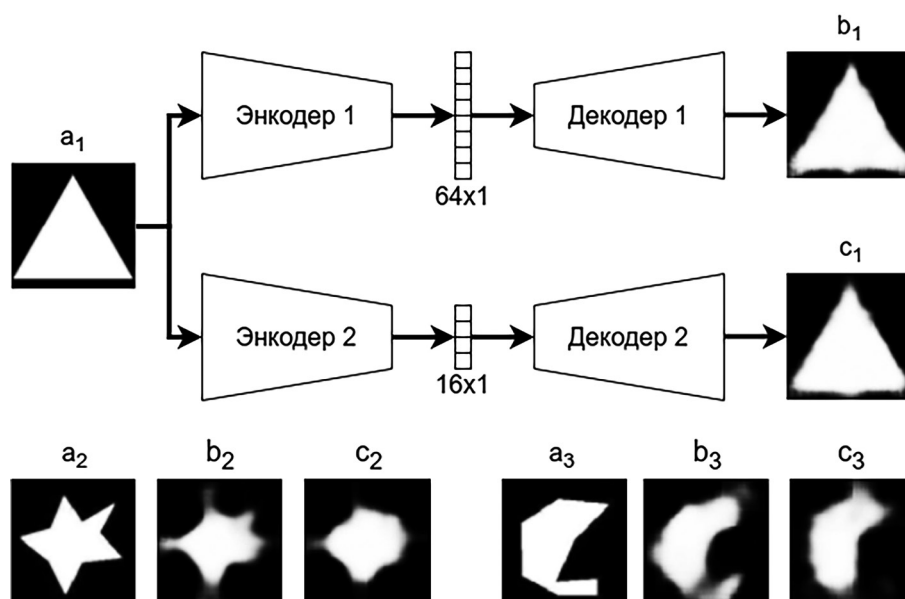
- 1) Метод Шайдука-Останина [6];
- 2) Метод быстрого преобразования Фурье (БПФ) в координатной сетке Декарта [7];
- 3) Метод вычисления коэффициента сжатия изображения [7], [8];
- 4) Нейросетевой метод с использованием вариационного автоэнкодера (ВАЭ) [7].

6.1. Метод Шайдука-Останина

Суть метода Шайдука-Останина заключается в вычислении энтропии спектральной плотности мощности (СПМ) центрированной сигнатуры радиуса, проведенного из центра тяжести контура изображения к каждой из точек этого контура. По сравнению с нашим методом (на основе инвариантов Hu) метод Шайдука-Останина обладает как преимуществами, так и недостатками. Оба метода дают оценку изображения (или участка изображения), инвариантную к смещению, к вращению и к масштабированию.

Оценивая сложность пятна по нашему методу, мы получаем оценку сложности целым числом от 1 до 8 (по числу инвариантов Hu), и поэтому метод оценки сложности, предложенный нами, требует дополнительного сравнения для пятен. В свою очередь, по методу Шайдука-Останина, оценка сложности является числом, лежащим в диапазоне от 0 до бесконечности, то есть для любого из двух пятен мы можем однозначно сказать, какое из них сложнее (тогда как по нашему методу мы можем распределить в одну из 8 категорий). Более “подробного” ранжирования сложности нашим методом можно добиться, используя значение $|k_i Hu_i|_{max}$ пятна, либо яркость пятна, либо его

¹ Величина E_{tot} является суммой яркостей всех пикселей пятна изображения, а величина S_{tot} — его полной площадью (в пикселях).



Здесь:

1. a_1, a_2, a_3 – оригинальные изображения (далее обозначенные как $S_{\{i\}}$);
2. b_1, b_2, b_3 – результат трассировки через слой скрытого представления размером 64×1 (далее обозначенные как $R64_{\{i\}}$);
3. c_1, c_2, c_3 – результат трассировки через слой скрытого представления размером 16×1 (далее обозначенные как $R16_{\{i\}}$);
4. $C_{\{i\}} = \sum |R64_{\{i\}} - R16_{\{i\}}| / \sum |S_{\{i\}}|$ – численно рассчитанная сложность для исходного изображения $S_{\{i\}}$:
 $C_{\{1\}} = 137.7/1661 = 0.083$;
 $C_{\{2\}} = 204.0/1064 = 0.192$;
 $C_{\{3\}} = 433.5/1293 = 0.335$.

Рис. 4. Схема архитектуры нейронной сети – вариационного автоэнкодера, пример трассировки тестового бинарного изображения и оценки его степени сложности.

площадь (что, конечно, является недостатком нашего метода).

К недостаткам метода Шайдука-Останина можно отнести оценку сложности пятна только по его внешней форме (только по контуру вокруг пятна). Таким образом, сложность, полученная методом Шайдука-Останина, никак не зависит от внутренней структуры пятна и распределения яркости в пятне – как в нашем методе.

Также к недостаткам метода Шайдука-Останина можно отнести вычислительную сложность. Для получения результата необходимо выполнить следующие действия:

- 1) Выделить пятна на общем фоне.
- 2) Найти контуры вокруг выделенных пятен.
- 3) Вычислить усреднённую сигнатуру радиуса.
- 4) Найти СПМ сигнатуры радиуса.
- 5) Вычислить модифицированную энтропию Шеннона [9].

Видим, что помимо более сложной предобработки (необходимо выделить контуры рассматриваемых объектов), метод Шайдука-Останина требует вычисления корней при нахождении ра-

диусов, преобразования Фурье для нахождения СПМ и взятия логарифмов при вычислении энтропии. В свою очередь наш метод требует только более быстрых с точки зрения вычислений на ЭВМ операций сложения, вычитания, деления, умножения и сравнения. Таким образом, несмотря на меньший динамический диапазон оценки сложности изображения, предложенный нами метод обладает более низкой вычислительной сложностью, а потому может быть использован в системах реального времени. Более того, наш метод хорошо поддаётся распараллеливанию вычислений (на физические потоки), чего нельзя сказать о методе Шайдука-Останина, в котором необходимо вычислять контуры, что требует последовательных вычислений.

6.2. Метод БПФ в координатной сетке Декарта

Метод БПФ в декартовых координатах заключается в предположении, что присутствие в спектре интенсивных высокочастотных гармоник является показателем сложности формы, что в то

же время может указывать на присутствие шума на изображениях. При отсутствии методов нахождения отличий воздействия шума от сложности формы в качестве меры может быть использовано некоторое среднее или интегральное значение мощности спектра амплитуд изображения. Присутствие на изображении относительно сложных элементов выражается в увеличении относительной интенсивности высокочастотных гармоник.

6.3. Метод вычисления коэффициента сжатия изображения

Метод сжатия заключается в расчете отношения байтового объема занимаемого пространства на информационном накопителе сжатого и несжатого (исходного) изображений при использовании алгоритма сжатия без потерь. В настоящей работе использовался python-модуль `zlib` [8] для сжатия изображений в представлении массива байтов.

6.4. Нейросетевой метод с использованием вариационного автоэнкодера (ВАЭ)

Нейронной сетью выступает предварительно обученный вариационный автоэнкодер (ВАЭ), предложенный в работе [7]. Нейронная сеть решает задачу воспроизведения на выходном слое изображения, поступившего на входной. Входное изображение рассматривается с позиции тензорного представления и трассируется через нейронную сеть, включающую в себя энкодер, слой скрытого представления и декодер. Тензор в энкодере последовательно сжимается по ширине и высоте слоями свёртки (Conv2d), активации (ReLU) и прореживания (MaxPooling2d), при этом увеличивается количество слоев (глубина) тензора. Слой скрытого представления (latent layer) является “бутылочным горлышком” системы, в нем сосредоточена сжатая информация о входном изображении. Декодер последовательно выполняет развертку данных. В настоящем (частном) случае выходное изображение, поступающее на дальнейшую обработку из декодера, согласовано по высоте и ширине со входным изображением. Размер скрытого слоя влияет на способность нейронной сети воспроизводить (реконструировать) поступающие на входной слой изображения. В экспериментах работы [7] размер скрытых слоев ограничен 16 и 64 нейронами для двух ВАЭ. Для оценки сложности используется сумма абсолютной попиксельной разницы между продуктами трассировки входного изображения через два ВАЭ с разным размером слоя скрытого представления. Сложность

вычисляется как частное суммы абсолютных отклонений и интегральной яркости исходного изображения. Нужно отметить, что авторами работы [7] не рассматриваются альтернативные методы оценки схожести продуктов трассировки ветвей модели. На рис. 4 представлена схема архитектуры нейронной сети, состоящей из двух независимых ветвей ВАЭ, а также представлен пример трассировки некоторого бинарного изображения и расчета его сложности. Схема взята из оригинальной публикации [7].

6.5. Сравнение методов оценки сложности

Метод Шайдука-Останина базируется на инвариантной к вращению математической модели. Вращение влияет только на фазовый спектр сигнатуры контура. Спектр мощности не зависит от поворота контура относительно центра тяжести. Метод не анализирует внутреннюю структуру пятен. Следовательно, метод Шайдука-Останина не может рассматриваться для достоверного анализа пятен на предмет двойкой сложности: сложность окаймляющего контура и сложность внутреннего распределения интенсивности (структура пятна).

В настоящей работе были поставлены эксперименты, в которых сравнивались методы: БПФ, нейросетевой, сжатие и предлагаемый нами (на основе моментов H_u). Анализировалось изменение оценки сложности модели при добавлении преобразований сдвига, масштабирования и вращения, а также при добавлении синтетического шума к тестовым изображениям. Тестовые изображения приведены на рис. 5. Устойчивость модели к преобразованиям типа сдвиг, масштабирование и вращение при оценке сложности важна, так как одна и та же геометрическая форма может быть зарегистрирована на изображении с произвольным положением в плоскости кадра и с различным масштабом. Шум является трудно устранимым явлением, в особенности, при недостаточном значении светосилы оптической системы, низкой чувствительности матричного приемника излучения и при недостаточной



Рис. 5. Тестовые изображения, используемые для сравнения методов оценки сложности на предмет инвариантности к некоторым из аффинных преобразований и к аддитивному шуму: яблоко (слева), птица (в центре) и муха (справа).

ширине диапазона варьирования времени экспозиции.

Из группы преобразований, используемых при сравнении методов, наибольший интерес представляет именно устойчивость метода к вращению. Устойчивость к сдвигу обеспечивается выделением области интереса на изображении (ROI), которая получена с использованием алгоритма сегментации. Устойчивость к масштабу в нашем методе обеспечивается автоматически (инварианты Ну масштабно-инвариантны). В нейросетевом методе все изображения (пятна с объектами) масштабируются перед применением свёрточных слоёв, поэтому нейросетевой метод также гарантирует инвариантность к масштабу. Что касается метода БПФ и метода сжатия, то эти методы не содержат в явном виде инвариантности к масштабу, но её в них можно добавить — также масштабируя пятно с объектом к заданному размеру перед применением этих методов.

В первом эксперименте анализировалось поведение оценки сложности при добавлении шума на изображения. Сложность оценивалась и усреднялась для трех предварительно отобранных исходно бинарных (черно-белых) тестовых изображений при наложении синтетического шума. Выбранная модель шума — шум Гаусса. Среднеквадратичное отклонение (СКО) шума рассматривается из дискретного набора значений $[0, 0.5, 1, 2, 4, 8, 16, 32, 64]$ (у.е.). Нулевое значение СКО было добавлено в ряд с целью последующих графических построений усредненной вычисленной сложности в зависимости от СКО шума из “нуля”. Глубина кодирования интенсивности (яркости) пикселя изображения равняется 8 битам, возможные значения находятся в диапазоне $BrighnessRange = [0...255]$ (у.е.). При добавлении реализации шума на изображение получаемое значение оценки в общем случае изменяется. Положительным качеством метода является способность стабильно удерживать числовую оценку.

Анализ результатов полученной оценки проводится не на одном изображении, а на статистической выборке. Объем статистической выборки $N_{МК} = 1000$. По этой выборке могут быть рассчитаны статистические моменты (например, среднее или СКО). Это используется для достижения репрезентативности результатов анализа. В рамках одной статистической выборки СКО шума неизменно. Для просмотра динамики оцениваемой сложности рассматриваются значения СКО шума из дискретного набора (показан выше). Выборка для конкретного значения СКО шума составляется как массив изображений, на

которых добавлена реализация шума с текущим значением СКО. Далее в рамках полученной выборки каждым методом по порядку производится оценка сложности всех изображений. Полученные оценки усредняются. Результатом является значение средней сложности, полученной каждым методом для каждого значения СКО шума из набора.

Результаты математически представляются так, что средняя сложность является функцией двух переменных: метода оценки и СКО шума. Каждый метод представлен отдельной кривой на графике, по оси абсцисс которого откладывается СКО шума, а по оси ординат — нормированная средняя сложность. Нормировка средней сложности позволяет уравнивать порядки получаемых величин разными методами и оценивать относительную динамику изменения.

По определению считаем, что стабильность метода при воздействии шума $SN(Stability to Noise)$ является описанной выше нормированной средней сложностью. Стабильность метода при воздействии шума зависит от СКО шума (аргумент функции), параметром выступает конкретный метод. Ниже представлена формула, по которой рассчитывается стабильность метода при воздействии шума.

$$SN_{id}(\sigma_{IN}) = \frac{M[C_{id}(\sigma_{IN})]}{C_{id}^{ref}}, \quad (14)$$

где: SN — стабильность метода при воздействии шума; id — идентификатор метода; σ_{IN} — СКО шума; M — оператор вычисления среднего значения по статистической выборке случайной величины; $C_{id}(\sigma_{IN})$ — сложность, оцененная методом id при добавлении шума с СКО σ_{IN} ; C_{id}^{ref} — сложность, оцененная методом id на исходном (оригинальном) изображении;

На рис. 6 представлено семейство графиков, показывающих зависимость стабильности метода от СКО шума для различных методов. На рис. 7 представлен пример одного из тестовых изображений (муха) с реализацией аддитивного шума при $СКО = 64$ (у.е.).

Во втором эксперименте отслеживалось изменение оценки сложности разными методами при вращении тестовых изображений (рис. 5). Вращение производилось без “подрезания” содержимого изображения. Под “относительной сложностью” $C^N(\alpha)$ при повороте оцениваемого объекта на некоторый угол α понимается отношение значения оцененной сложности при повороте на определенный угол $C^{rot}(\alpha)$ и значения

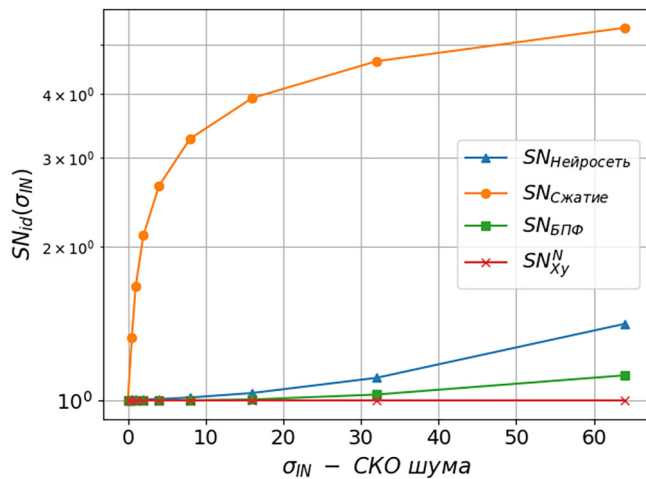


Рис. 6. Стабильность при воздействии шума, полученная разными методами.

сложности исходного изображения C^{ref} . Параметрами выступают идентификаторы метода id и тест-изображения I_i из рисунка 5.

$$C_{id,I_i}^N(\alpha) = \frac{C_{id,I_i}^{rot}(\alpha)}{C_{id,I_i}^{ref}} \quad (15)$$

Далее объект выделялся по содержанию. Шаг поворота составлял $\delta\alpha = 1^\circ$, при визуализации сглаживались высокочастотные колебания показаний одномерным усредняющим фильтром с размером ядра равным $\Delta\alpha = 5^\circ$. Использовались три обозначенных ранее тестовых изображения рис. 5. На рис. 8 представлены три блока графиков, каждый блок соответствует своему тест-изображению из рис. 5. Графики показывают изменение относительной сложности для разных методов при вращении тест-объектов.

6.6. Выводы по сравнению методов

По результатам поставленного эксперимента метод сжатия показал наименьшую “стабиль-

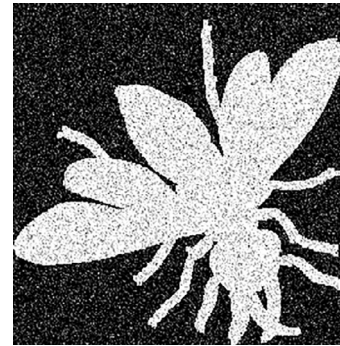


Рис. 7. Пример тестового изображения мухи с реализацией шума при СКО = 64(у. е.). Для бинарных изображений, у которых значения яркости равны либо 0, либо 255 (при 8-битной глубине кодирования) добавление шума с СКО = 64 не является фактором, в следствие которого исходные геометрические формы зрительно не могут быть распознаны.

ность” при воздействии шума, в то же время наш метод наиболее стабилен.

Изменчивость нормированной оценки сложности по нашему методу на каждом тестовом изображении минимальна в диапазоне вращения изображения от нуля до 360° .

Нейросетевой метод наиболее изменчив при вращении всех рассмотренных тест-изображений рис. 5.

Метод сжатия является наиболее сложным с вычислительной точки зрения и в исходном виде не имеет параметров для для настройки.

Из недостатков метода БПФ можно выделить отсутствие параметров для калибровки метода и потенциальная чувствительность к геометрическому шуму (коррелированному), в том числе к периодическому.

Методы сжатия, БПФ и наш метод обладают таким достоинством как “интерпретируемость”. В настоящем контексте под интерпретируемостью понимается возможность прямого обоснования полученных результатов, данное свойство

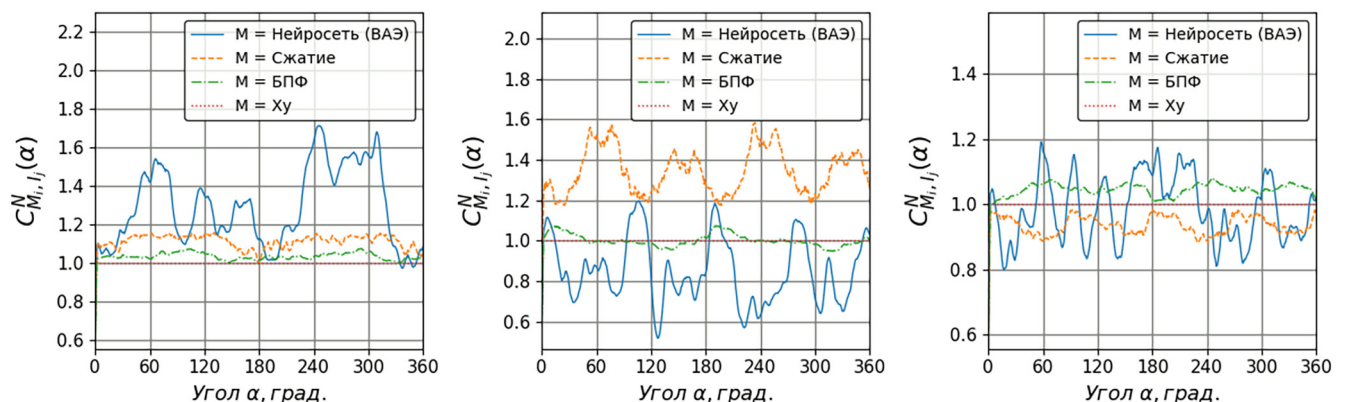


Рис. 8. Изменение относительной сложности для разных методов при вращении тест-объектов: яблоко (слева), птица (в центре) и муха (справа).

модели является особенно ценным в процессе отладки и калибровки под конкретную предметную область. Нейросетевой метод включает в себя блоки последовательной обработки (в том числе нелинейные операции), в которых, к сожалению, обоснование получаемых результатов только косвенное.

Нейросетевой (ВАЭ) метод показал среднюю стабильность оценки при воздействии шума, а также наибольшую изменчивость оценки к вращению на тестовом наборе данных. Кроме этого результаты работы нейросетевого метода зависят от тренировочной выборки, настроек обучения, параметров нейронной сети и способа сравнения изображений.

Среди потенциальных преимуществ нейросетевого метода может быть отмечена возможность получения высокой производительности при вычислении на графических ускорителях (видеокартах). Еще одним преимуществом нейросетевого метода можно считать возможность тренировки сети на базовую фильтрацию шумов. Вероятно, при более оптимальном подборе разрешения слоя скрытого представления или тренировочного набора данных, а также при использовании методов аугментации нейросетевым методом возможно удастся получить лучшие результаты.

Резюмируя, можно отметить, что рассмотренные методы сжатия, БПФ, нейросетевой метод и наш метод при наличии определенных преимуществ также обладают и недостатками.

7. ВЫВОДЫ

- Нами предложен новый метод оценки сложности геометрических фигур (пятен), учитывающий внутреннюю структуру пятен, а не только их внешний контур.

- Задача по вычислению степени сложности объектов разделена на составляющие: сегмен-

тация пятен + оценка сложности изолированных пятен.

- Предлагаемый метод обладает относительно низкой вычислительной сложностью по сравнению с вышеуказанными альтернативными методами.

- Нами проведен качественный и количественный анализ существующих (альтернативных) методов, выявлены их преимущества и недостатки по сравнению с нашим методом и друг с другом.

- Реализованный на основе нашего метода алгоритм апробирован как на искусственных, так и на реальных изображениях.

СПИСОК ЛИТЕРАТУРЫ

1. <https://ru.wikipedia.org/wiki/WMAP>
2. [https://ru.wikipedia.org/wiki/Планк_\(космическая_обсерватория\)](https://ru.wikipedia.org/wiki/Планк_(космическая_обсерватория))
3. *Hu M.K.* Visual pattern recognition by moment invariants. IRE Trans. Info. Theory, IT-8:179–187, 1962.
4. Alastair Florence Frederik Doerr. A micro-xrt image analysis and machine learning methodology for the characterisation of multi-particulate capsule formulations. International Journal of Pharmaceutics, 2, 2020.
5. *Kornilov A.S., Safonov I.V.* An overview of watershed algorithm implementations in open source libraries. Journal of Imaging, 4(10):123, 2018.
6. *Шайдук А.М., Останин С.А.* Количественная оценка сложности контура медицинских изображений // Журнал радиоэлектроники. 2013. № 2.
7. *Ritter H., Rothgänger M., Melnik A.* Shape complexity estimation using vae. arXiv preprint arXiv:2304.02766., 2023.
8. Gilchrist J. Parallel data compression with bzip2. Proceedings of the 16th IASTED international conference on parallel and distributed computing and systems, 16:559–564, 2004.
9. Shannon C.E. A mathematical theory of communication. The Bell system technical journal, 27 (3):379–423, 1948.

ESTIMATING THE COMPLEXITY OF OBJECTS IN IMAGES

© 2024 V. B. Bokshanskiy^a, V. A. Kulin^a, G. S. Finiakin^{a,b},
A. S. Kharlamov^c, A. A. Shatskiy^a

^a*Bauman Moscow State Technical University, Moscow, Russia*

^b*National Research University "Moscow Power Engineering Institute", Moscow, Russia*

^c*Moscow State Technical University of Civil Aviation, Moscow, Russia*

A new method for estimating the complexity of geometric shapes (spots) is proposed, taking into account the internal structure of the spots, and not only their external contour. The task of calculating the degree of complexity of objects is divided into components: segmentation of spots and estimation of the complexity of isolated spots. The new method has a relatively low computational complexity compared to the alternative methods considered in the work. Using the new method, an algorithm based on parallel computing of the CUDA programming language for graphics accelerators (video cards) was created, which further increases the performance of our method. A qualitative and quantitative analysis of existing (alternative) methods has been carried out, their advantages and disadvantages in comparison with our method and with each other have been revealed. The algorithm implemented on the basis of the new method has been tested on both artificial and real digital images.

Keywords: Hu invariants, image complexity, image compression, image contour, image segmentation, contour selection, image classification, statistical moments of images, computational complexity

REFERENCES

1. <https://ru.wikipedia.org/wiki/WMAP/>
2. [https://en.wikipedia.org/wiki/Planck_\(spacecraft\)](https://en.wikipedia.org/wiki/Planck_(spacecraft))
3. Hu M.K. Visual pattern recognition by moment invariants, IRE Trans. Info. Theory, 1962. V. IT-8, P. 179–187 (1962).
4. Doerr F.J.S., Florence, A.J. A micro-xrt image analysis and machine learning methodology for the characterisation of multi-particulate capsule formulations, Int. J. Pharm., 2020. V. 2.
5. Kornilov A.S., Safonov I.V. An overview of watershed algorithm implementations in open source libraries, J. Imaging, 2018. V. 4. No. 10. P. 123.
6. Shaiduk A.M., Ostanin S.A. Quantitative estimation of shape complexity in medical images // Zh. Radioelektron. 2013. V. 2.
7. Rothganger M., Melnik A., Ritter H. Shape complexity estimation using VAE. ArXiv:2304.02766, 2023.
8. Gilchrist J. Parallel data compression with bzip2, Proceedings of the 16th IASTED International Conference on Parallel and Distributed Computing and Systems, 2004. V. 16. P. 559–564.
9. Shannon C.E. A mathematical theory of communication. Bell Syst. Techn. J., 1948. V 27, No. 3. P. 379–423.