

Поиск структурных различий изображений: алгоритмы и методы исследования*

Ф. А. Корнилов

fakornilov@mail.ru

Институт математики и механики им. Н. Н. Красовского Уральского отделения Российской академии наук (ИММ УрО РАН), 620990, Российская Федерация, г. Екатеринбург, ул. Софьи Ковалевской, 16

В работе рассматривается задача поиска структурных различий на разновременных космических снимках земной поверхности, под которыми понимаются появившиеся или исчезнувшие объекты местности. В качестве отправной точки служит морфологический анализ изображений Ю. П. Пытьева. Вводится общая формулировка понятия структуры изображения и структурных различий. Приводится ряд алгоритмов, решающих поставленную задачу для полутоновых изображений, а также предлагается вариант морфологического проектора для работы с цветными изображениями. Для алгоритма, основанного на морфологическом проекторе, приводятся методика построения распределения яркости выходного изображения в случае возмущения входных изображений аддитивным шумом, и определяются формула оптимального порога и оценки ошибок первого и второго рода результатов работы алгоритма для нее. Формулируется вычислительная методика сравнения алгоритмов поиска структурных различий изображений, для чего вводится специальная математическая модель, позволяющая максимально приблизить эксперимент к реальным задачам. С помощью этой методики определяется наилучший процент верного обнаружения (ложной тревоги), который можно получить при зафиксированном максимально допустимом уровне ложной тревоги (минимально допустимом уровне верного обнаружения), а также приводятся результаты сравнения качества работы предложенных алгоритмов между собой. Эксперименты на реальных данных показывают, что предложенные алгоритмы подходят для прикладного применения.

Ключевые слова: *обработка изображений; структурные различия; математическая морфология; регуляризация; марковские случайные поля*

Detection of structural differences in images: algorithms and methods of research*

F. A. Kornilov

Institute of Mathematics and Mechanics named after N. N. Krasovkiy, Ural Branch of the Russian Academy of Sciences, 16 S. Kovalevskaja Str., Ekaterinburg 620990, Russia

Background: The article is devoted to the problem of structural differences detection at satellite images taken at different times. Here, structural changes mean appeared or disappeared landscape objects. The starting point of the investigation is the morphological analysis of images theory by Yu. P. Pytiev.

Methods: In the work, the general notion of the image's structure and structural difference are introduced. Several algorithms for solving the problem for grayscale images are proposed and also, the morphological projector's modification for dealing with color images is given. For the algorithm based on the morphological projector in the case of image's distortion with additive noise, the method for computing the output image's values distribution is given. Also the formula for optimal threshold and estimation of the I and II type errors are given.

Results: The computational technique for comparison of the algorithms of structure changes detections in images is formulated. For such a purpose, the special mathematical model is introduced. This model provides an opportunity to maximally closer move to applied real problems. The proposed technique gives the way for computing the best value of the true positives rate (false alarm rate), with given fixed maximal value of the false alarm rate (minimal value of the true positives rate). Also, the comparison results of the performance of the proposed algorithms are given.

Concluding Remarks: Experiments with real data demonstrate that the proposed algorithms are good enough for applications.

Keywords: *image processing; structural differences; mathematical morphology; regularization; markov random fields*

Введение

Автоматический анализ данных дистанционного зондирования, в частности космических снимков в оптическом диапазоне, является важной проблемой при решении многих практических задач [1]. Необходимость своевременного обновления топографических карт с наименьшими затратами делает крайне актуальной задачу поиска структурных различий на разновременных снимках одного и того участка земной поверхности. В неформальном, содержательном смысле структурные различия — это существенные изменения местности, при которых объекты интереса появляются, исчезают или изменяют свою форму. При этом изменения освещенности и цвета объектов структурными различиями не считаются. Поскольку процедура создания картографических материалов является достаточно трудоемкой и затратной, то решение об обновлении, в частности, топографических карт принимается в случае, если на местности произошли существенные изменения объектового состава. Таким образом, одной из актуальных задач анализа цифровых космических снимков становится поиск структурных различий на двух разновременных снимках с целью выявления изменений в составе объектов сцены на них. Данная статья является продолжением серии работ, посвященных задаче поиска структурных различий изображений [2, 3].

Задача поиска различий на двух изображениях является классической для анализа изображений и возникает в различных областях компьютерного зрения, таких как сжатие видеоданных, системы видеонаблюдения и других. Основными подходами к ее решению являются методы попиксельного анализа, не учитывающие специфику объектов; методы, основанные на анализе контуров объектов [4] и методы, основанные на текстурной классификации [5]. Фундаментальное отличие этих подходов сводится к различному пониманию структуры изображения, среди многообразия подходов к которому следует выделить морфологический анализ изображений Ю.П. Пытьева [6], который в данной работе является основой для построения алгоритмов поиска структурных различий изображений.

Основные определения

Под *изображением* f будем понимать функцию $f : X \rightarrow \mathbb{R}$, а семейством допустимых изображений будем называть конечномерное евклидово пространство Φ таких функций f ,

наделенное естественным скалярным произведением $(f, g) = \sum_X f(x)g(x)$ и нормой $\|f\| = (\sum_X f^2(x))^{1/2}$. Значение $f(x)$ будем называть *яркостью* изображения f в точке $x \in X$, а множество пикселей X (поле зрения) — подмножество плоскости \mathbb{R}^2 , $S = |X|$ (количество пикселей) — *размером* (или *площадью*) изображения.

Структура изображения — это некоторый способ описания информации о представленной на изображении сцене. Разумеется, для конкретного снимка таких описаний может быть бесчисленное множество и выбор подходящего описания напрямую зависит от требований задачи. Структура изображения должна учитывать существенные изменения сцены и быть инвариантной к несущественным в контексте данной задачи изменениям. Помимо этого, должен существовать способ сравнения структур изображений между собой с целью определения наличия *структурных* различий между ними. Существует множество различных определений структуры изображений, среди которых важное место занимает теория морфологического анализа Ю.П.Пытьева [6], в котором развиты методы, инвариантные относительно преобразований, характеризующих влияние условий регистрации изображений. В его основе лежит понятие *формы* изображения f , под которой понимается оператор проектирования (морфологический проектор) P_f в пространстве изображений Φ на множество изображений, полученных из исходного путем применения всех возможных преобразований яркости:

$$P_f g(x) = \sum_i \frac{\sum_{x' \in X} g(x') \cdot \chi_i^f(x')}{\sum_{x' \in X} \chi_i^f(x')} \cdot \chi_i^f(x), \quad (1)$$

где

$$\chi_i^f(x) = \begin{cases} 1, & \text{если } f(x) = i; \\ 0, & \text{иначе.} \end{cases}$$

Выбор конкретного определения структуры должен быть связан с практическими требованиями задачи. В данной работе поиск структурных различий осуществляется с целью оценки площади произошедших изменений, при этом изображения содержат дискретный набор яркостей, поэтому в качестве носителей структуры будут выбраны множества уровней яркости изображения:

Определение 1. Структурой Ω изображения f называется семейство $\mathcal{L}_f = \{L_f(i)\}_i$ множеств уровня $L_f(i) = \{x \in X \mid f(x) = i\}$ функции f .

В данной работе основное внимание уделяется относительной радиометрической (яркостной) коррекции исходных снимков. Функции, пригодные для яркостной коррекции изображений связаны с идеей сохранения структуры одного из изображений с приведением яркостей его пикселей максимально близко к яркостям второго изображения. Применительно к задаче поиска структурных различий это означает построение изображений f' и g' , которые были бы по структуре близки к f и g , а по яркости f' близко к g , а g' близко к f ; и дальнейшего вычисления различий уже для построенных изображений с g и f соответственно:

$$\begin{aligned} f' &= \arg \min(\|h - g\|^2 + \chi(h; f) \mid h \in \Phi); \\ g' &= \arg \min(\|h - f\|^2 + \chi(h; g) \mid h \in \Phi), \end{aligned} \quad (2)$$

где под χ понимается характеристическая функция структурной эквивалентности, которая отражает тот факт, что изображение h получено путем изменения значений уровней яркости изображения f с сохранением его структуры. Необходимость построения второго изображения связана с тем, что операция яркостной коррекции несимметрична, т. е. ее применение позволяет найти лишь «появившиеся» изменения (объекты, отсутствующие на первом изображении, но присутствующие на втором). Нахождение «исчезнувших» объектов требует проведения симметризации яркостной коррекции.

В работе будут рассматриваться алгоритмы, в основе которых лежат следующие характеристические функции структурной эквивалентности:

1. характеристическая функция χ_π : $\chi_\pi(h; f) = 0$, если форма h не сложнее f в смысле морфологии Ю. П. Пытьева, и иначе равна бесконечности, т. е. равенство функции χ_π нулю означает, что имеется некоторая функция преобразования яркости $F: \mathbb{R} \rightarrow \mathbb{R}$ такая, что $h(x) = F(f(x)) \forall x \in X$;

2. характеристическая функция χ_1 : $\chi_1(h; f) = 0$, если $h(x) = kf(x) + b \forall x \in X$ и для некоторых вещественных k, b , и иначе равна бесконечности;

3. характеристическая функция χ_2 : $\chi_2(h; f) = 0$, если $h(x) = kf(x)^2 + b \cdot f(x) + c \forall x \in X$ и для некоторых вещественных k, b, c , и иначе равна бесконечности;

4. характеристическая функция χ_m :

$$\chi_m(r; f) = \lambda \sum_{x \in X} \sum_{y \in N_x} G_f(h(x) - h(y), x, y);$$

$$G_f(i, x_1, x_2) = \begin{cases} \min(i^2, \alpha^2), & \text{если } |f(x_1) - f(x_2)| < \alpha; \\ 0, & \text{иначе,} \end{cases}$$

где N_x — окрестность точки x , λ и α — положительные параметры. Параметр λ отвечает за близость приведения изображения f к g : чем он больше, тем больше алгоритм будет стремиться сохранить структуру изображения f , из-за чего яркостное приведение f к g будет менее точным. Если же $\lambda \rightarrow 0$, то структура изображения f будет разрушаться для наиболее полного совпадения яркостей f и g (очевидно, что если $\lambda = 0$, то $f' = g$). Параметр α характеризует структуру изображения f , задавая допустимый перепад яркости между различными линиями уровня: если яркости двух точек отличаются меньше, чем на α , то они принадлежат одному уровню.

Теперь можно сформулировать определение структурного различия:

Определение 2. Структурным различием называется множество пикселей, удовлетворяющее следующему условию:

$$SD(f, g) = \{x \in X \mid |f'(x) - g(x)| > T \text{ или } |f(x) - g'(x)| > T\},$$

где f' и g' определены формулами (2), T — некоторое пороговое значение.

Можно показать, что применение формулы (2) эквивалентно построению морфологической проекции Ю. П. Пытьева при определенном выборе характеристической функции χ_π . Доказательство этого факта и аналогичное утверждение для класса линейных и квадратичных функций приводится в [2].

Введем на пространстве изображений Φ отношение частичного порядка. Для произвольных $f, g \in \Phi$ будем говорить, что структура изображения f не сложнее структуры изображения g , и писать $\mathcal{L}_f \preceq \mathcal{L}_g$, если для каждого значения i функции f найдется разбиение $L_f(i) = \bigcup_{j=1}^n L_g(i_j)$, $n \geq 1$, причем $i_k \neq i_l \forall k, l \in [1, n]$. Если говорить неформально, то

на изображении g сохранены все границы изображения f , но к ним, возможно, добавлены новые. Очевидно, что $|\mathcal{L}_f| \leq |\mathcal{L}_g|$ (здесь $|\cdot|$ — мощность множества). Определение структуры изображения (1) наиболее близко по сути к форме изображения в морфологическом анализе Ю.П. Пытьева, и можно доказать следующее утверждение:

Утверждение 1. $\mathcal{L}_f \preceq \mathcal{L}_g$ тогда и только тогда, когда $f = P_g f$.

Благодаря этому утверждению можно разбить все возможные варианты соотношения структур исходных изображений на следующие три случая.

1. Изображения f и g независимы: $\mathcal{L}_f \not\preceq \mathcal{L}_g$ и $\mathcal{L}_g \not\preceq \mathcal{L}_f$. Тогда $f \neq P_g f$ и $g \neq P_f g$. В этом случае будем говорить, что структура этих изображений различна, т. е. обнаружено структурное различие изображений в целом.

2. Изображения f и g изоморфны: $\mathcal{L}_f \preceq \mathcal{L}_g$ и $\mathcal{L}_g \preceq \mathcal{L}_f$. Тогда $f = P_g f$ и $g = P_f g$. В этом случае будем говорить, что изображения имеют одинаковую структуру, т. е. структурных различий нет.

3. Для изображений f и g выполнено либо $\mathcal{L}_f \preceq \mathcal{L}_g$ и $\mathcal{L}_g \not\preceq \mathcal{L}_f$, либо $\mathcal{L}_f \not\preceq \mathcal{L}_g$ и $\mathcal{L}_g \preceq \mathcal{L}_f$. Поскольку эти случаи симметричны, то, не нарушая общности, далее везде будем считать, что выполняются условия $\mathcal{L}_f \preceq \mathcal{L}_g$ и $\mathcal{L}_g \not\preceq \mathcal{L}_f$ ($f = P_g f$ и $g \neq P_f g$). В отличие от случаев независимых и изоморфных изображений здесь можно говорить не только об обнаружении структурного различия, но и о его локализации, т. е. об определении наличия или отсутствия структурного различия в произвольной зафиксированной точке x_c .

Математическая модель структурных различий

Пусть дана пара изображений $f^0, g^0 \in \Phi$. Первое изображение f^0 (будем называть его фоном) содержит N_f уровней яркости. Значение яркости i -го уровня обозначим f_i^0 , площадь уровня — $S_i = |L_{f^0}(f_i^0)|$ ($i \in [1, N_f]$). Изображение g^0 строится путем добавления к i -ому уровню фона объекта, состоящего из N_g^i уровней яркости. Значение яркостей уровней разбиения $j \in [0, N_g^i]$ обозначим $g_{i,j}^0$, $S_{i,j} = |L_{g^0}(g_{i,j}^0)|$, $S_i = \sum_{j=0}^{N_g^i} S_{i,j}$ (индексы j от 1 до

N_g^i соответствуют объекту, 0 — фону). Таким образом, $L_{f^0}(f_i^0) = \bigcup_{j=0}^{N_g^i} L_{g^0}(g_{i,j}^0)$, $i \in [1, N_f]$.

Чтобы избежать краевых эффектов, добавим условие, что площадь добавленного объекта меньше площади фона. Таким образом, изображение g состоит из $N_f + \sum_{i=1}^{N_f} N_g^i$ уровней яркости, где первое слагаемое соответствует количеству значений функции f^0 , второе слагаемое соответствует добавленному объекту. Будем считать, что в данной точке есть структурное различие, если эта точка принадлежит объекту. Для моделирования реальной ситуации в задаче поиска структурных различий к каждому изображению добавляем дискретный, стационарный и независимый в каждой точке аддитивный шум с известными параметрами: для первого изображения — n_η , для второго — n_ξ (здесь n — означает случайную величину, η и ξ — плотности распределения). Обозначим $f(\cdot) = f^0(\cdot) + n_\eta$ и $g(\cdot) = g^0(\cdot) + n_\xi$. В результате имеем два семейства случайных величин (два случайных поля):

$$\begin{aligned} f &= \{f(x_i)\}_{i=1}^S = \{f^0(x_i) + n_\eta\}_{i=1}^S, \\ g &= \{g(x_i)\}_{i=1}^S = \{g^0(x_i) + n_\xi\}_{i=1}^S, \end{aligned} \quad (3)$$

которые можно интерпретировать как модели случайных изображений, к которым будем применять алгоритм поиска структурных различий.

Поскольку изображения f и g — случайные поля, то значение разностного изображения $R = |f' - g|$ (более подробное описание алгоритма будет дано в следующей главе) в точке $x_c \in X$ также является величиной случайной, и основной задачей будет построение способа ее вычисления. Однако значения случайной величины $R(x_c)$ существенно зависят от того, есть структурное различие в точке x_c или его нет, поэтому необходимо рассматривать соответствующие два условных распределения.

У с л о в и е 1: точка x_c принадлежит объекту, т. е. в этой точке есть структурное различие.

У с л о в и е 2: точка x_c принадлежит фону, т. е. структурного различия в этой точке нет.

Распределение вероятности $r(R(x_c) | x_c \text{ принадлежит объекту})$ обозначим $r^+(R)$. Второе распределение $r(R(x_c) | x_c \text{ принадлежит фону})$ обозначим $r^-(R)$.

З а д а ч а 1. Получить условные распределения вероятностей $r^+(R)$ и $r^-(R)$ величины $R(x_c)$ для модельной задачи.

Алгоритмы поиска структурных различий изображений

Общая схема алгоритма поиска структурных различий изображений

В данном пункте приводится описание разработанных алгоритмов поиска структурных различий изображений. Их общая схема такова:

1. Исходные изображения сканируются с шагом 1 локальным окном заданного размера $d \times d$. Центральную точку этого окна будем обозначать x_c . В случае необходимости производится геометрическое выравнивание фрагментов.

2. Для каждого положения окна по двум фрагментам сравниваемых изображений f и g строятся две функции преобразования яркости F_{fg} и F_{gf} . Использование схемы с двумя функциями позволяет добиться симметризации результата, т. е. становится неважным, сравнивается первое изображение со вторым или наоборот.

3. С помощью функций F_{fg} и F_{gf} строятся преобразованные изображения $f' = F_{fg}(f)$ и $g' = F_{gf}(g)$, при этом яркость изображения f' «выровнена» по яркости изображения g с сохранением структуры изображения f , и аналогично — для изображения g' .

4. Строятся разностные изображения $R_{fg}(x) = |f'(x) - g(x)|$ и $R_{gf}(x) = |g'(x) - f(x)|$. Для этих изображений яркость точки характеризует величину структурного несоответствия исходных изображений, т. е. чем ярче точка, тем вероятнее, что в ней присутствует структурное различие.

5. Для завершения симметризации строится результирующее разностное изображение R , яркость каждой точки которого есть максимум яркостей точек разностных изображений с соответствующими координатами: $R(x) = \max(R_{fg}(x), R_{gf}(x))$.

6. Производится пороговая обработка изображения R , параметр порога будем обозначать T . Если стоит задача локализации структурного различия, то оценивается яркость центральной точки: при выполнении $R(x_c) \geq T$ в данной точке присутствует структурное различие, и в соответствующую точку выходного результирующего изображения $\mathbf{R}(x_c)$ (изображения–результата работы алгоритма) записывается значение 255, если нет, то 0. Для задачи обнаружения структурного различия условие $R(x) \geq T$ проверяется для каждой точки сканирующего окна, и если число таких точек значительно, то считается, что в данном окне обнаружено структурное различие, и в $\mathbf{R}(x_c)$ записывается значение

255, иначе 0. Таким образом, обработка каждого сканирующего окна дает на результирующем изображении одну точку, яркость которой 255 или 0. Далее из этих точек формируются связные области структурных различий — конечный результат. В целях борьбы со случайными выбросами можно фильтровать найденные области, отбрасывая те, площадь которых незначительна.

Функция преобразования яркости — это попиксельно действующий оператор относительной радиометрической коррекции исходных изображений, который строится для каждого положения сканирующего окна на основе наборов яркостей изображений в нем, и применяется для «выравнивания» яркостей фрагментов с сохранением их структуры. Такой подход позволяет игнорировать не считающиеся структурными различиями объекты, изменившие свой цвет, но сохранившие форму. В качестве таких функций в данной работе используются:

1. морфологический проектор (1);
2. регуляризованный вариант морфологического проектора, предназначенный для устранения влияния шума с сохранением точности границ найденных различий, и связанный со сглаживанием характеристических функций уровней яркости:

$$P_f g(x) = \frac{\sum_i \tilde{f}_i(x)}{\sum_j \tilde{\chi}_i^f(x)} = \sum_i \left(\frac{\sum_{x' \in X} g(x') \cdot \tilde{\chi}_i^f(x')}{\sum_{x' \in X} \tilde{\chi}_i^f(x')} \cdot \frac{\tilde{\chi}_i^f(x)}{\sum_j \tilde{\chi}_i^f(x)} \right), \quad (4)$$

где $\tilde{\chi}_i^f$ вычисляется по формуле

$$\tilde{\chi}_i^f(x) = \exp\left(-\frac{(f(x) - i)^2}{\sigma_c^2}\right) \cdot \exp\left(-\frac{(x - x_c)^2}{\sigma_d^2}\right);$$

3. линейная функция преобразования яркости:

$$f'(x) = k \cdot f(x) + b, \quad (5)$$

где коэффициенты находятся с помощью МНК по двум фрагментам изображений, ограниченных сканирующим окном, из соотношения

$$k \cdot f(\cdot) + b = g(\cdot);$$

4. квадратичная функция преобразования яркости:

$$f'(x) = a \cdot f^2(x) + b \cdot f(x) + c, \quad (6)$$

где коэффициенты находятся с помощью МНК по двум фрагментам изображений, ограниченных сканирующим окном, из соотношения

$$a \cdot f^2(\cdot) + b \cdot f(\cdot) + c = g(\cdot).$$

В основе алгоритма поиска структурных различий с непрерывной структурной характеристической функцией χ_m лежит, как и в предыдущих случаях, яркостная коррекция входных снимков, заключающаяся в приведении набора яркостей первого изображения к яркости второго с сохранением структуры первого изображения. Для улучшения результата также производится приведение каждого изображения к самому себе, что позволяет

избавиться от присутствующего шума. Реализация такого алгоритма сводится к задаче глобальной минимизации функционалов (2), что эквивалентно минимизации следующей энергетической функции:

$$\begin{aligned} E(f'; f, g) &= \|f'(x) - g(x)\| + \chi_m(f'; f) = \\ &= \|f'(x) - g(x)\| + \lambda \sum_{x \in X} \sum_{y \in N_x} G_f(f'(x) - f'(y), x, y); \end{aligned} \quad (7)$$

$$G_f(i, x_1, x_2) = \begin{cases} \min(i^2, \alpha^2), & \text{если } |f(x_1) - f(x_2)| < \alpha; \\ 0, & \text{иначе,} \end{cases}$$

где N_x — окрестность точки x , λ и α — положительные параметры (описаны выше). Для нахождения минимума функции $E(f'; f, g)$ будет использоваться аппарат марковских случайных полей (МСП) [7]. Одним из наиболее простых в реализации методов приближенного решения подобной задачи является алгоритм Метрополиса [8]. В качестве энергетической функции для данного метода будем использовать введенную функцию $E(f'; f, g)$.

Теоретическое исследование алгоритма

Описав схему алгоритма и используемые в нем функции преобразования яркости, можно перейти к решению задачи локализации структурных различий для модельной задачи 1 для алгоритма поиска структурных различий, основанного на морфологическом проекторе. Поскольку алгоритм поиска структурных различий оценивает только одну зафиксированную точку $x_c \in X$, то при использовании морфологического проектора важно только множество $L_f(f(x_c))$, которому эта точка принадлежит. Пусть $k = |L_f(f(x_c))|$. Очевидно, что $k \in [1, S]$. Введем обозначения: $L_f^{i,j}(f(x_c)) = \{x \in L_f(f(x_c)) \mid f^0(x) = f_i^0 \text{ и } g^0(x) = g_{i,j}^0\}$; под $B(n, m, p)$ будем понимать биномиальное распределение вероятностей: вероятность появления события ровно n раз в серии из m испытаний при условии, что в единичном испытании вероятность его появления равна p ; конфигурация $a = (a_{1,0}, a_{1,1}, \dots, a_{1,N_g^1}, \dots, a_{N_f,0}, a_{N_f,1}, \dots, a_{N_f,N_g^{N_f}})$, где $a_{q,j}$ соответствует мощности пересечения множества $L_f(f(x_c))$ и уровня яркости объекта с номером j , полученного из разбиения уровня фона с номером q (индекс 0 соответствует фону на втором изображении). Из-за разницы яркостей незашумленного изображения f^0 в точках разных уровней после добавления шума распределения вероятностей яркостей в этих уровнях также будут различными, поэтому далее будем обозначать $p_q(f = i) = p(f(x) = i \mid f^0(x) = f_q^0)$ — вероятность появления яркости i в точке, которая на изображении f^0 принадлежит уровню яркости с номером q . Теперь можно доказать следующую теорему:

Теорема 1. Пусть дано случайное поле f , заданное по формуле (3).

1. В задаче 1 при выполнении условий 1 и $x_c \in L_f^{n,m}(f(x_c))$ ($n \in [1, N_f]$ и $m \in [1, N_g^n]$) вероятность появления реализации поля f такой, что $|L_f(f(x_c))| = k$, $|L_f^{n,m}(f(x_c))| = a_{n,m} + 1$ и $|L_f^{q,j}(f(x_c))| = a_{q,j}$ для $(q, j): q \in [1, N_f]$, $j \in [0, N_g^q]$, $(q, j) \neq (n, m)$, равна

$$\begin{aligned} p_{k,a}^{n,m} &= \sum_i \left(p_n(f = i) \prod_{q=1}^{n-1} \prod_{j=0}^{N_g^q} B(a_{q,j}, S_{q,j}, p_q(f = i)) \times \right. \\ &\times \left. \prod_{j=0}^{m-1} B(a_{n,j}, S_{n,j}, p_n(f = i)) B(a_{n,m}, S_{n,m} - 1, p_n(f = i)) \times \right. \end{aligned}$$

$$\times \prod_{j=m+1}^{N_g^n} B(a_{n,j}, S_{n,j}, p_n(f=i)) \prod_{q=n+1}^{N_f} \prod_{j=0}^{N_g^q} B(a_{q,j}, S_{q,j}, p_q(f=i)),$$

где

$$\begin{aligned} \sum_{q=1}^{N_f} \sum_{j=1}^{N_g^q} a_{q,j} &= k-1; \\ 0 \leq a_{q,j} &\leq S_{q,j} \quad \forall (q,j): q \in [1, N_f], j \in [0, N_g^q], (q,j) \neq (n,m); \\ 0 &\leq a_{n,m} \leq S_{n,m} - 1. \end{aligned}$$

2. В задаче 1 при выполнении условий 2 и $x_c \in L_f^{n,0}(f(x_c))$ ($n \in [1, N_f]$) вероятность появления реализации поля f такой, что $|L_f(f(x_c))| = k$, $|L_f^{n,0}(f(x_c))| = a_{n,0} + 1$ и $|L_f^{q,j}(f(x_c))| = a_{q,j}$ для $(q,j): q \in [1, N_f], j \in [0, N_g^q], (q,j) \neq (n,0)$, равна

$$\begin{aligned} p_{k,a}^{n,0} &= \sum_i \left(p_n(f=i) \prod_{q=1}^{n-1} \prod_{j=0}^{N_g^q} B(a_{q,j}, S_{q,j}, p_q(f=i)) \times \right. \\ &\times B(a_{n,0}, S_{n,0} - 1, p_n(f=i)) \prod_{j=1}^{N_g^n} B(a_{n,j}, S_{n,j}, p_n(f=i)) \times \\ &\left. \times \prod_{q=n+1}^{N_f} \prod_{j=0}^{N_g^q} B(a_{q,j}, S_{q,j}, p_q(f=i)) \right), \end{aligned}$$

где

$$\begin{aligned} \sum_{q=1}^{N_f} \sum_{j=1}^{N_g^q} a_{q,j} &= k-1; \\ 0 &\leq a_{n,0} \leq S_{n,0} - 1; \\ 0 \leq a_{q,j} &\leq S_{q,j}, \quad \forall (q,j): q \in [1, N_f], j \in [0, N_g^q] \quad (q,j) \neq (n,0). \end{aligned}$$

Таким образом, пространство всех изображений разбивается на классы в зависимости от параметров (k, a) путем перебора всех возможных вариантов реализации поля (изображения) f , а теорема 1 позволяет определить вероятности для классов из этого разбиения. Можно доказать основную теорему:

Теорема 2. 1. В задаче 1 при условии 1 распределение вероятностей случайной величины $R(x_c)$ имеет вид:

$$r^+(i) = \frac{\sum_{n=1}^{N_f} \sum_{m=1}^{N_g^n} (S_{n,m} r_{n,m}(i))}{\sum_{u=1}^{N_f} \sum_{v=1}^{N_g^u} S_{u,v}},$$

где

$$r_{n,m}(i) = \sum_{k,a} (p_{k,a}^{n,m} p(|G_{k,a}^{n,m} + E_k| = i));$$

величина $p_{k,a}^{n,m}$ определяется теоремой 1:

$$G_{k,a}^{n,m} = \frac{1}{k} \left(\sum_{q=1}^{N_f} \sum_{j=0}^{N_g^q} (a_{q,j} g_{q,j}^0) + (1-k) g_{n,m}^0 \right); \quad E_k = \frac{1}{k} \sum_{j=1}^{k-1} \varepsilon_j + \frac{1-k}{k} \varepsilon_c;$$

ε_j — распределение вероятностей случайной величины $g(x_j)$, соответствующей точке $x_j \in L_f(f(x_c))$, ε_c — распределение вероятностей случайной величины $g(x_c)$.

2. В задаче 1 при условии 2 распределение вероятностей случайной величины $R(x_c)$ имеет вид:

$$r^-(i) = \frac{\sum_{n=1}^{N_f} (S_{n,0} r_{n,0}(i))}{\sum_{u=1}^{N_f} S_{u,0}},$$

где

$$r_{n,0}(i) = \sum_{k,a} (p_{k,a}^{n,0} P(|G_{k,a}^{n,0} + E_k| = i)),$$

величина $p_{k,a}^{n,0}$ определяется теоремой 1:

$$G_{k,a}^{n,0} = \frac{1}{k} \left(\sum_{q=1}^{N_f} \sum_{j=0}^{N_g^q} (a_{q,j} g_{q,j}^0) + (1-k) g_{n,0}^0 \right);$$

определения E_k , ε_j и ε_c такие же, как в п. 1.

Опираясь на теорему 2, можно определить способ вычисления оптимального порога алгоритма. В качестве критерия оптимальности будет использоваться *percentage correct classification*, т. е. минимум суммы ошибок первого и второго рода. Имеет место следующая теорема.

Теорема 3. В условиях задачи 1 порог алгоритма поиска структурных различий на основе морфологического проектора, вычисленный по следующей формуле:

$$T_{opt} = \arg \min_T \left(\sum_{i < T} r^+(i) + \sum_{i > T} r^-(i) \right),$$

является оптимальным в смысле критерия *percentage correct classification*.

Доказательство теорем 1–3 приведено в [3]. Результаты численного эксперимента по применению теоремы 3 изображены на рис. 1 для случая объекта, состоящего из двух уровней яркости, графики распределений вероятностей $r^+(R)$ и $r^-(R)$, полученных с помощью теоремы 2. Параметры изображений $S = 441$, $S_1 = 60$, $S_2 = 60$, $g_1 = 120$, $g_2^0 = 80$, $g_0^0 = 0$. Оба изображения содержат дискретный шум, распределение которого аппроксимируется нормальным распределением с параметрами $N(0, 10)$. График распределения $r^+(R)$ состоит из двух пиков, соответствующих двум уровням объекта.

Морфологический проектор для цветных изображений

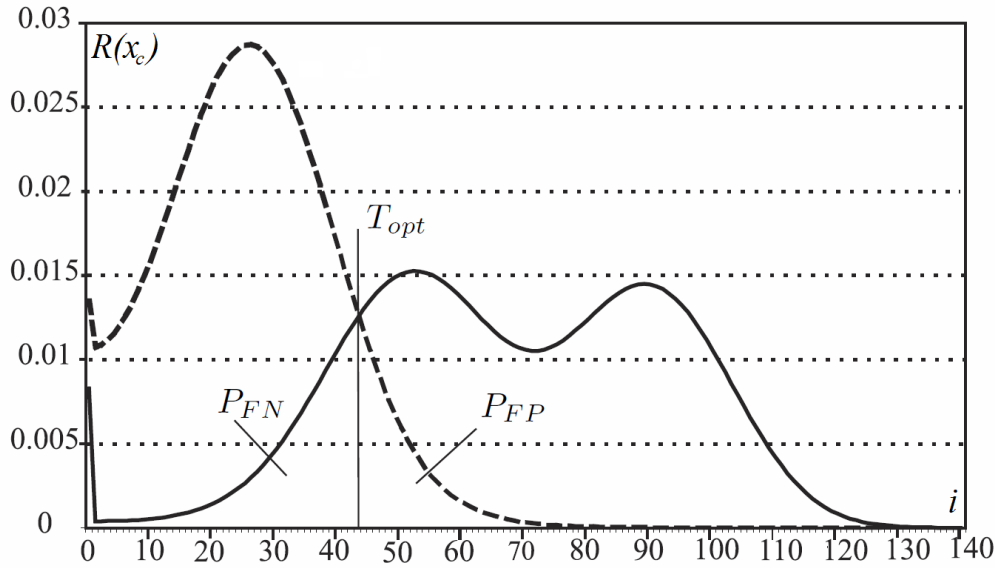


Рис. 1. Графики распределений $r^+(R)$ (сплошная линия) и $r^-(R)$ (пунктирная линия) для изображений с параметрами $S = 441$, $S_1 = 60$, $S_2 = 60$, $g_1^0 = 120$, $g_2^0 = 80$, $g_0^0 = 0$ (добавление двухуровневого объекта). Оптимальный порог $T_{opt} = 43$. Оценки ошибок: $P_{FN} = 0.15$, $P_{FP} = 0.1$

Представленные в данной работе алгоритмы предназначены для поиска структурных различий на полутоновых снимках земной поверхности. Тем не менее определенный интерес представляет решение данной задачи для цветных изображений. Простейшим вариантом является поканальная обработка цветных снимков предложенными методами. Однако, такой подход не дает ничего нового по сравнению со схемой обесцвечивания и применения стандартных методов, в то время как цветное изображение содержит гораздо больше полезной информации. Идея предлагаемого метода идентична вышеописанным подходам: первоначально требуется произвести радиометрическую коррекцию входных изображений. Ю. П. Пытьевым в работе [6] тема устранения влияния условий регистрации для цветных изображений раскрыта достаточно широко. Здесь будет предложен один из возможных вариантов проектора.

Поскольку теперь речь будет идти о цветных изображениях, то придется внести изменения во введенное определение изображения. А именно, в данном пункте под *изображением* \mathbf{f} будем понимать функцию $\mathbf{f}: X \rightarrow \mathbb{R}^3$, а значение $\mathbf{f}(x) = (f_R(x), f_G(x), f_B(x))$ будем называть *яркостью* изображения \mathbf{f} в точке $x \in X$, где f_R , f_G и f_B — яркости по красному, зеленому и синему каналу, соответственно.

Суть морфологического проектора заключается в усреднении яркостей второго изображения по уровням яркости первого (1). Однако представляется довольно затруднительным задать уровень яркости как некоторое множество в пространстве \mathbb{R}^3 : любое разбиение является и подходящим, и неудачным, что делает задачу такого разбиения неоднозначной. Однако известно, что предлагаемый в данной работе алгоритм оценивает центральную точку окна, поэтому можно задавать уровень яркости как точки, близкие по яркости к центральной x_c :

$$\hat{L}_{\mathbf{f}}(x_c) = \{x \in X \mid \max(|f_R(x) - f_R(x_c)|, |f_G(x) - f_G(x_c)|, |f_B(x) - f_B(x_c)|) \leq t_c\},$$

где t_c — параметр яркостной близости. В таком случае каждой точке первого изображения соответствует свой набор точек и свой уровень яркости, но эти наборы точек могут иметь непустое пересечение. Тогда формула морфологического проектора для цветных изображений выглядит следующим образом:

$$P_f g(x) = \sum_i \frac{\sum_{x' \in X} g(x') \hat{\chi}^f(x', x)}{\sum_{x' \in X} \hat{\chi}^f(x', x)} \cdot \chi_i^f(x), \quad (8)$$

где

$$\mathbf{i} = (i_R, i_G, i_B); \quad \chi_i^f(x) = \begin{cases} 1, & \text{если } \mathbf{f}(x) = \mathbf{i}; \\ 0, & \text{иначе;} \end{cases} \quad \hat{\chi}^f(x', x) = \begin{cases} 1, & \text{если } x' \in \hat{L}_f(x); \\ 0, & \text{иначе.} \end{cases}$$

Алгоритм поиска структурных различий с функцией преобразования яркости (8) полностью совпадает с описанным выше. На рис. 2 представлены результаты его работы для двух цветных изображений городской застройки:

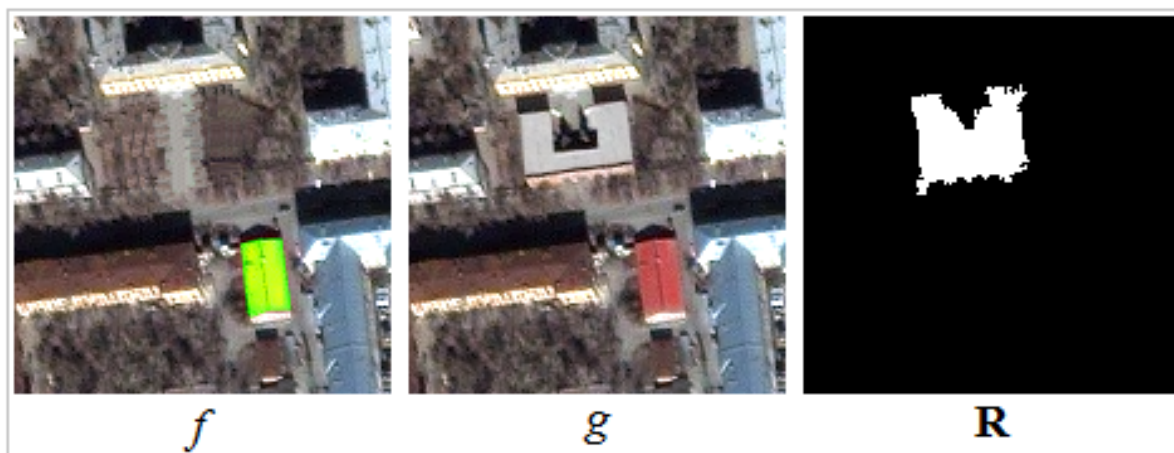


Рис. 2. Результаты работы алгоритма поиска структурных различий для цветных изображений: исходные изображения f и g и найденные области структурных различий R

Параметры алгоритма были выбраны вручную: яркостная близость для линии уровня $t_c = 16$, порог $T = 55$. Можно видеть, что алгоритм успешно нашел появившийся дом, проигнорировав строение, крыша которого изменила свой цвет. Таким образом, можно говорить о применимости данного метода для поиска структурных различий на разновременных цветных снимках земной поверхности.

Эмпирическое исследование алгоритмов и сравнение результатов их работы

Работа каждого алгоритма напрямую зависит от выбранного набора его параметров. Ключевое место среди них занимает оптимальный порог, однако оценка требуется и для всех остальных: например, для размера сканирующего окна. Среди всех возможных значений параметров требуется выбрать те, при которых алгоритм показывает наилучшее качество работы в зависимости от уровня шума, присутствующего на изображениях. Подходящим критерием качества работы является анализ ROC-кривых [9] (ROC — *Receiver*

Таблица 1. Максимальные значения TP (минимальные FP) при зафиксированном уровне FP = 0,1 (TP = 0,9) для алгоритмов поиска структурных различий, и оптимальные параметры, при которых они получены

Алгоритм	TP _{max}	FP _{min}	Параметры
Рег. морфологический проектор	0.892091	0.113738	$d = 29, \sigma_c = 2$
Квадратичная функция	0.890949	0.113725	$d = 23$
Морфологический проектор	0.885229	0.120545	$d = 27$
Линейная функция	0.884299	0.122274	$d = 23$
Глобальная оптимизация	0.80634	0.332012	$d = 19, \lambda = 1816, \alpha = 5$

Operating Characteristic, операционная характеристика приемника), появившийся в теории обработки сигналов в середине прошлого века. Количественную интерпретацию ROC-кривой дает показатель AUC (англ. *Area Under the Curve*, площадь под ROC-кривой) — площадь, ограниченная ROC-кривой и осью доли ложных положительных классификаций. Чем выше показатель AUC, тем качественнее классификатор, при этом значение 0,5 соответствует случайному решению.

В первой части данной статьи была предложена математическая модель структурных различий, подходящая для теоретического исследования алгоритмов. Однако, несмотря на ее близость к реальным изображениям, спутниковые снимки все же намного сложнее. Тем не менее для проведения численного эксперимента можно расширить эту модель путем усложнения «фона» исходных изображений и добавляемых «объектов», в качестве которых теперь будут рассматриваться не уровни яркости, а фрагменты спутниковых снимков, что позволит учесть следующие соображения:

1. Сравнение алгоритмов необходимо проводить на реальных, а не на синтезированных изображениях. Работа каждого алгоритма будет оцениваться путем сравнения (с помощью анализа ROC-кривых) полученных результатов с ручной разметкой исходных изображений. Однако получение ручной разметки реальных снимков представляет собой сложную и трудоемкую задачу. Причиной является сложность дешифрирования таких снимков, даже в ручном режиме. Иными словами, на снимках могут (и будут) присутствовать области, которые нельзя однозначно отнести к областям структурных различий или областям их отсутствия.

2. Возможность варьировать и строго контролировать тип и уровень шума, добавляемого к изображениям. То есть необходимо, что бы исходные изображения имели «нулевой» уровень шума.

3. Возможность выбора размера и типа объектов, присутствующих на изображениях.

Построения пары изображений, которые станут входными данными для алгоритма поиска структурных различий, и изображения с ручной разметкой осуществляется следующим образом. Выбирается космический снимок подходящего участка земной поверхности, который разделяется на две части: первый фрагмент и его копия становятся исходными изображениями, второй фрагмент будет донором для генерации структурных различий. Поскольку исходные изображения изначально идентичны, то одинаковы и присутствующие на них уровни яркости, вызванные влиянием шума. Таким образом, эти уровни яркости становятся в некотором смысле «содержательными», т. е. определяют «фон» на обоих изображениях. Применение алгоритмов поиска структурных различий к таким изображениям, очевидно, дает нулевое результирующее разностное изображение R . Структур-

ными различиями становятся объекты, вырезанные из второго фрагмента снимка и добавленные к исходным изображениям. При этом добавление объекта ко второму изображению эквивалентно его появлению; добавление к первому изображению — исчезновению объекта; добавление двух разных объектов к одной и той же области обоих снимков — изменению формы. Типы переносимых объектов также определяются требованиями задачи.

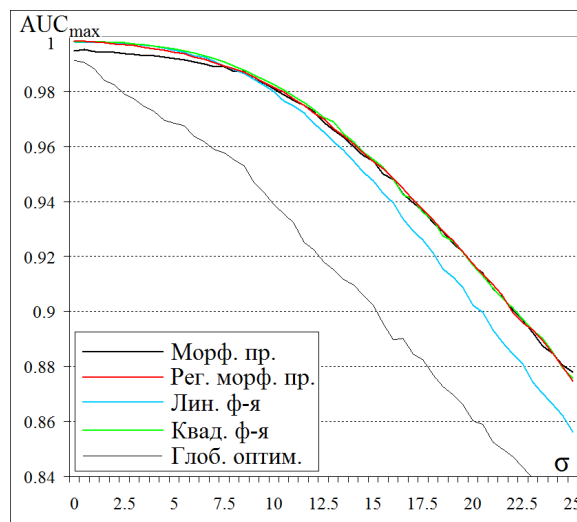


Рис. 3. Графики зависимости максимальной по всем параметрам величины площади под ROC-кривой AUC_{max} от величины стандартного отклонения шума σ для алгоритмов поиска структурных различий изображений

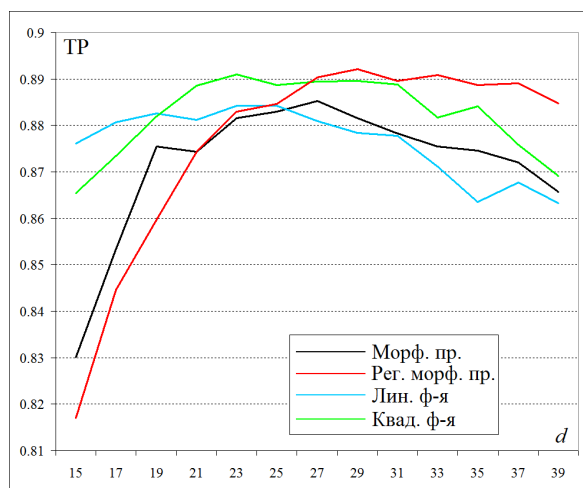


Рис. 4. Графики наибольшего уровня верного обнаружения TP алгоритмов для заданного максимально допустимого уровня ложной тревоги $FP = 0,1$ при различных значениях размера сканирующего окна d

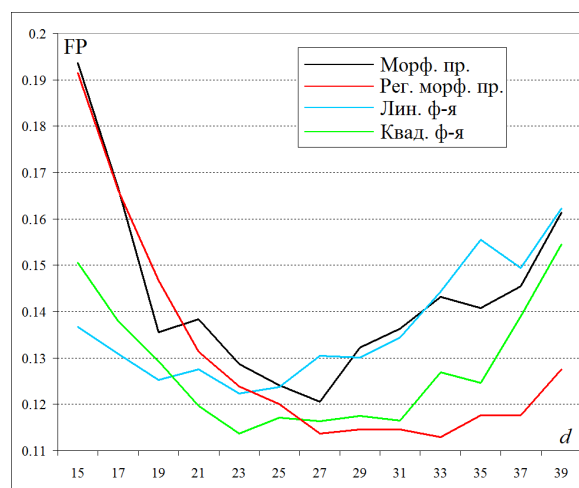


Рис. 5. Графики наименьшего уровня ложной тревоги FP алгоритмов для заданного минимально допустимого уровня верного обнаружения $TP = 0,9$ при различных значениях размера сканирующего окна d

Ввиду своей крайне высокой вычислительной сложности, численный эксперимент проводился с использованием технологии CUDA [10] и суперкомпьютера «Уран» ИММ УрО

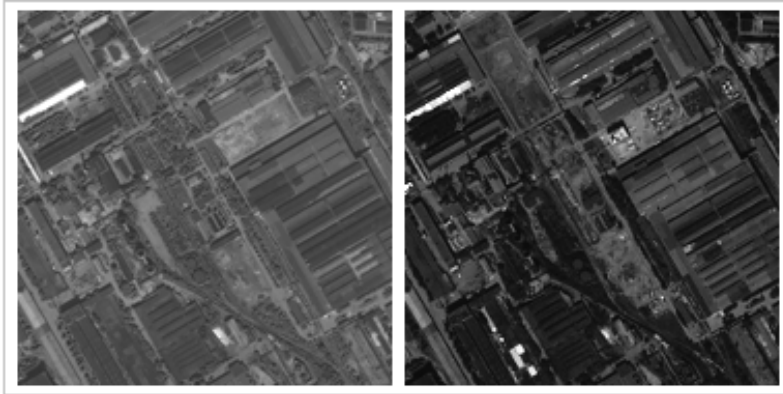


Рис. 6. Исходные полутоновые снимки городской застройки, сделанные с разницей в 4 года

РАН [11]. Схемы серий испытаний для различных алгоритмов отличаются только набором исследуемых параметров, и состоят из следующих шагов:

1. Перебираются значения параметров алгоритма. Интервал варьирования параметров установлен на основе опыта работы с алгоритмом так, чтобы гарантированно содержать оптимальные значения.

2. Генерируется пара изображений (в соответствии с моделью) с добавлением шума, стандартное отклонение которого выбирается из интервала $\sigma \in [0, 25]$ с шагом 0.5.

3. К полученной паре изображений применяется алгоритм с выбранными значениями параметров. Результат работы алгоритма сравнивается с ручной разметкой для нахождения величины площади под ROC-кривой AUC. Шаги 2–3 повторяются 10 раз и полученные результаты усредняются. Максимальное значение AUC находится для каждого значения σ по всем параметрам алгоритмов, т. е. это наилучшее качество работы алгоритма при заданном уровне шума. Полученные графики изображены на рис. 3.

В реальной задаче может потребоваться знание следующих показателей: какой наилучший процент верного обнаружения (ложной тревоги) можно получить при зафиксированном максимально допустимом уровне ложной тревоги (минимально допустимом уровне верного обнаружения)? Для этого зададим допустимый уровень ложной тревоги $FP = 0.1$ и для каждого алгоритма определим наибольшее значение верного обнаружения TP . Полученные графики в зависимости от размера сканирующего окна d представлены на рис. 4. Аналогично, на рис. 5 представлены графики наименьшего уровня ложной тревоги FP алгоритмов для заданного уровня верного обнаружения $TP = 0.9$. В обоих случаях можно видеть, что до значения $d = 19$ лучшие результаты показывает алгоритм, основанный на линейной функции преобразования яркости, от $d = 21$ до $d = 25$ — алгоритм, основанный на квадратичной функции преобразования яркости, а далее — алгоритм, основанный на регуляризованном морфологическом проекторе, что подтверждает представленные на рис. 3 результаты. В табл. 1 представлены максимальные значения верного обнаружения TP (ложной тревоги FP) для всех алгоритмов и оптимальные параметры, при которых они получены.

Наконец, на рис. 6 изображены два фрагмента реальной пары космоснимков городской застройки, а на рис. 7 — фрагменты созданного вручную эталона (ручной разметки присутствующих на снимках структурных различий) и результатов работы алгоритмов. Черным цветом отмечены верно классифицированные положительные примеры, белым — верно классифицированные отрицательные случаи, красным — пропуск (ошибка I рода),

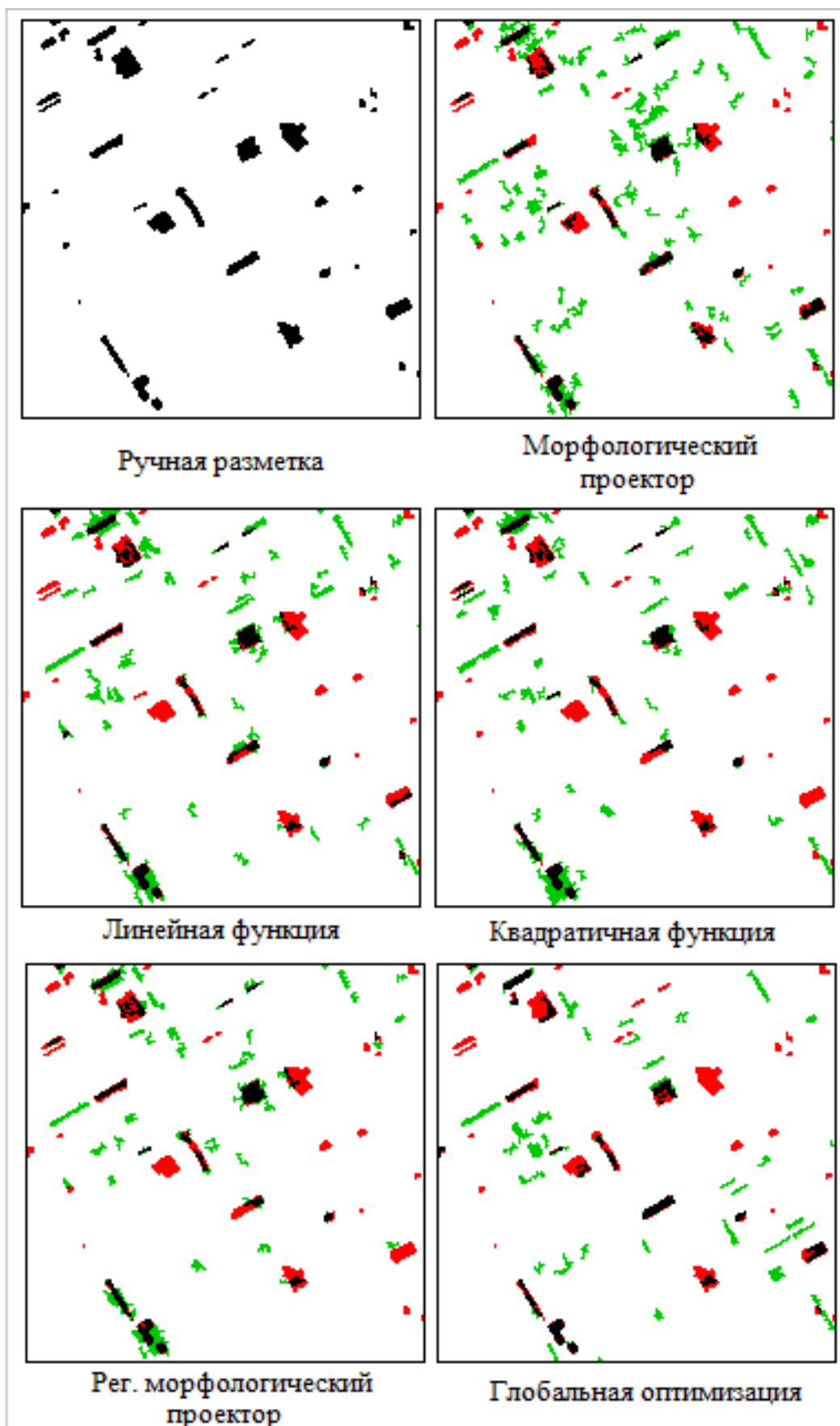


Рис. 7. Ручная разметка и результаты работы алгоритмов поиска структурных различий с оптимальными параметрами: черным цветом отмечено верное обнаружение, белым — верно классифицированные отрицательные случаи, красным — пропуск, зеленым — ложная тревога

зеленым — ложная тревога (ошибка II рода). Результаты таковы: морфологический проектор достаточно неплохо нашел имеющиеся структурные различия, но и выдал больше всех ложных срабатываний; линейная и квадратичная функции оказались близки — среднее количество найденных различий и умеренное ложных тревог; регуляризованный морфологический проектор нашел практически те же различия, что и предыдущие алгоритмы, но зато выдал наименьшее количество ложных срабатываний; глобальная оптимизация показала наилучший результат по обнаружению и локализации различий, но уровень ложных тревог оказался чуть выше среднего.

Заключение

В данной работе был подведен итог серии работ, посвященных задаче поиска структурных различий изображений, а также предложены варианты морфологического проектора для цветных изображений и методика исследования алгоритмов с целью определения оптимальных параметров работы. В последней была сформулирована особая математическая модель, подходящая для проведения численного эксперимента, максимально приближенного к реальным задачам. На основе всех полученных данных был разработан программный комплекс, решающий задачу поиска структурных различий изображений. Данный комплекс был интегрирован в среду визуализации и обработки данных дистанционного зондирования Земли ENVI, и внедрен в работу ОАО «НИИП центр Природа».

Резюмируя, можно отметить, что в ходе исследовательской работы удалось получить удовлетворительное решение поставленной задачи с помощью различных алгоритмов, но в тоже время на основе единого теоретического подхода. Можно сформулировать следующие направления дальнейшей работы.

1. Получить теоретические формулы для оптимального порога алгоритмов поиска структурных различий. В настоящий момент такая формула есть лишь для алгоритма, основанного на морфологическом проекторе. Определив эти формулы для остальных алгоритмов, можно будет сформулировать утверждение о теоретическом преимуществе одного из них. Эмпирическое преимущество осталось за алгоритмом, основанном на регуляризованном морфологическом проекторе.

2. Использование глобальной оптимизации дает очень неплохой и перспективный результат, поэтому важным направлением является усовершенствование данного метода и его дальнейшее применение в прикладных задачах.

3. Разработать версии алгоритмов, работающие с многоканальными изображениями, а также использующими информацию о тестурах и границах поверхностей. Также стоит использовать также данные различных геофизических полей, например, поле высот строений и сооружений, которое можно получить, используя стереоснимки.

Литература

- [1] Kostousov V. B., Kandoba I. N., Skripnuk V. V., Shabanov G. A. The system for automated deciphering of cosmic Earth surface photographs // *Int. Archives Photogrammetry Remote Sensing*. Amsterdam, 2000. Vol. XXXIII. Part B4. P. 425–432.
- [2] Корнилов Ф. А., Перевалов Д. С. Задача обнаружения структурных различий изображений // *Алгоритмы и программ. средства парал. вычислений: Сб. науч. тр.* — Екатеринбург: УрО РАН, 2011. Вып. 11. С. 37–56.
- [3] Корнилов Ф. А. Исследование алгоритма поиска структурных различий изображений // *Тр. Ин-та математики и механики УрО РАН.* — Екатеринбург: ИММ УрО РАН, 2013. Т. 19. № 1. С. 144–159.

- [4] *Li W., Li X., Wu Y., Hu Z.* A novel framework for urban change detection using VHR satellite images // IEEE The 18th Conference (International) on Pattern Recognition (ICPR'06). — Hong Kong, 2006. P. 312–315.
- [5] *Li L., Leung M. K. H.* Integrating intensity and texture differences for robust change detection // *IEEE Trans. Image Processing*, February 2002. Vol. 11. No. 2. P. 105–112.
- [6] *Пытьев Ю. П., Чуличков А. И.* Методы морфологического анализа изображений. — М.: Физ-МатЛит, 2010. 336 с.
- [7] *Li S. Z.* Markov random field modeling in image analysis. — 3rd ed. — Berlin–Heidelberg: Springer, 2009. 362 p.
- [8] *Винклер Г.* Анализ изображений, случайные поля и динамические методы Монте-Карло. Математические основы. — Новосибирск: Изд-во СО РАН, филиал «Гео», 2002. 343 с.
- [9] *Green D. M., Swets J. M.* Signal detection theory and psychophysics. — New York: John Wiley and Sons Ltd., 1966. 455 p.
- [10] *Боресков А. В. и др.* Параллельные вычисления на GPU. Архитектурная и программная модель CUDA. — М.: Изд-во Моск. ун-та. Сер. «Суперкомпьютерное образование», 2012. 336 с.
- [11] Суперкомпьютерный центр «Уран» ИММ УрО РАН. <http://parallel.uran.ru/>.

References

- [1] *Koustousov V. B., Kandoba I. N., Skripnuk V. V., Shabanov G. A.* 2000. The system for automated deciphering of cosmic Earth surface photographs. *Int. Archives Photogrammetry Remote Sensing*. Amsterdam. XXXIII(part B4):425–432.
- [2] *Kornilov F. A., Perevalov D. S.* 2011. The problem of detecting structural differences on the images. *Algorithms Software Parallel Computing*. Ekaterinburg: IMM UB RAS. 11:37–56.
- [3] *Kornilov F. A.* 2013. Study algorithm detection structural differences on the images. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*. Ekaterinburg: IMM UB RAS. 19(1):144–159.
- [4] *Li W., Li X., Wu Y., Hu Z.* 2006. A novel framework for urban change detection using VHR satellite images. *IEEE The 18th Conference (International) on Pattern Recognition (ICPR'06)*. Hong Kong. 312–315.
- [5] *Li L., Leung M. K. H.* 2002. Integrating intensity and texture differences for robust change detection. *IEEE Trans. Image Processing* 11(2):105–112.
- [6] *Pitev Y. P., Chulichkov A. I.* 2010. *The methods of morphological analysis of images*. Moscow: FizMatLit. 336 p.
- [7] *Li S. Z.* 2009. *Markov random field modeling in image analysis*. 3rd ed. Berlin–Heidelberg: Springer. 362 p.
- [8] *Wincler G.* 2002. *Image analysis, random fields and markov chain Monte Carlo methods: A mathematical introduction*. Novosibirsk: SB RAS press, “Geo.” 343 p.
- [9] *Green D. M., Swets J. M.* 1966. *Signal detection theory and psychophysics*. New York: John Wiley and Sons Ltd. 455 p.
- [10] *Boreskov A. V., et al.* 2012. *Parallel computing with GPU. Architecture and programming model CUDA*. Moscow: Moscow University Press. Ser. Supercomputer education. 336 p.
- [11] Supercomputer center “Uran” IMM UB RAS. Available at: <http://parallel.uran.ru/>.