

## Алгоритм совмещения пространственных объектов разномасштабных карт на основе топологического анализа данных

С.В. Еремеев<sup>1</sup>, Д.Е. Андрианов<sup>1</sup>, В.С. Титов<sup>2</sup>

<sup>1</sup> Владимирский государственный университет

имени Александра Григорьевича и Николая Григорьевича Столетовых, Владимир, Россия;

<sup>2</sup> ФГБОУ ВО «Юго-Западный государственный университет», Курск, Россия

### Аннотация

В статье рассматривается проблема автоматического совмещения пространственных объектов на разномасштабных картах одной и той же местности. Для решения поставленной задачи предлагается использовать методы топологического анализа данных. Исходными данными алгоритма являются пространственные объекты, которые могут быть получены с карт разных масштабов и подвержены искажениям. Персистентная гомология позволяет идентифицировать общую структуру таких объектов в виде топологических особенностей. Основными топологическими особенностями в исследовании являются компоненты связности и пустоты объектов. В работе приводится математическое описание метода персистентной гомологии для представления пространственных объектов. Приводится определение баркода для пространственных данных, который содержит описание объекта в виде топологических признаков. Разработан алгоритм сравнения баркодов пространственных данных, который позволяет найти общую структуру объектов. Алгоритм базируется на анализе данных из баркода. Введён показатель схожести объектов по топологическим признакам. Показаны результаты исследований работы алгоритма. Проведённые эксперименты подтвердили высокое качество предложенного алгоритма. Процент схожести при сопоставлении природных объектов с учётом масштаба и деформации получился в пределах от 85 до 92, а для муниципальных при наличии растяжений и искажений частей объектов – от 74 до 87. Отражены преимущества предложенного подхода с аналогами при совмещении объектов, которые подвержены значительной деформации при масштабировании, а также при искажениях.

**Ключевые слова:** персистентная гомология, баркод пространственного объекта, сопоставление объектов, анализ топологических особенностей, разномасштабные карты.

**Цитирование:** Еремеев, С.В. Алгоритм совмещения пространственных объектов разномасштабных карт на основе топологического анализа данных / С.В. Еремеев, Д.Е. Андрианов, В.С. Титов // Компьютерная оптика. – 2019. – Т. 43, № 6. – С. 1021-1029. – DOI: 10.18287/2412-6179-2019-43-6-1021-1029.

### Введение

Сопоставление пространственных объектов на двух картах одной и той же местности является актуальной задачей [1–4]. Она имеет множество различных приложений, среди которых обновление карт местности, сопоставление разномасштабных карт, поиск объектов на карте, выборка похожих объектов по определённым признакам [5–7]. Карты могут хранить разнородную информацию, представленную в растровом или векторном виде. Объём пространственной информации одного и того же участка территории с каждым годом увеличивается. Данные могут быть получены за определённый промежуток времени и сняты с разных углов и разных масштабов. Накопление такого рода пространственных данных влечёт несоответствие между картами и требует автоматизации. Можно выделить следующие подходы для решения данной задачи. Много разработок используют корреляционные методы, основанные на вычислении коэффициента корреляции между сравниваемыми изображениями [8]. Также существует целый ряд алгоритмов и подходов, которые применяют геометрические признаки для сравнения схожих объектов [9]. Вычисляются такие характеристики, как центр масс, площадь и периметр выпуклой оболочки,

а также их отношения и т.д. Данные методы целесообразно использовать, когда объекты идентичны друг другу, и эти характеристики сохраняются при аффинных преобразованиях. Что касается картографических объектов, которые расположены на разных масштабах или изменяются во времени, то они могут иметь другую, но похожую со сравниваемым объектом форму.

Чтобы учитывать подобные свойства, в настоящее время активно применяются методы непрерывного представления бинарных изображений в виде границ, скелета и циркуляров [10, 11], т.е. бинарное изображение представляет собой объединение непрерывных фигур. Отдельную группу составляют методы, направленные на извлечение эскизных характеристик объекта [12]. Важные исследования ведутся при анализе скелета контурных деформаций объекта [13]. Перспективные новые научные исследования направлены на анализ неорганизованных данных и создание персистентного скелета [14].

После генерализации пространственные объекты упрощаются, происходит их деформация, исходные данные на карте могут быть слабо связаны друг с другом, а на другом масштабе представлять единый объект. Все это затрудняет применение стандартных алгоритмов для сопоставления объектов на разных

масштабах. Однако один и тот же объект при генерализации сохраняет свою структуру и глобальные топологические признаки. Таким образом, естественно использовать топологические свойства объектов, которые инвариантны к подобным деформациям и искажениям [15, 16].

Также очень часто для сопоставления разнородных карт применяют ключевые точки [17], например, точку на центральном перекрестке. Для такого сопоставления карт требуется минимум две точки. Однако чаще всего эти ключевые точки необходимо проставить вручную. Кроме того, могут быть большие погрешности при сопоставлении объектов.

Таким образом, в статье рассматривается задача сопоставления пространственных объектов с деформациями на разных масштабах. Для её решения за основу предлагается взять методы персистентной гомологии, которые учитывают топологические свойства набора точек [18–21].

### 1. Представление пространственных объектов на основе метода персистентной гомологии

Персистентная гомология относится к методам топологического анализа данных [22, 23]. Она широко начинает использоваться в разных областях: обработка изображений, сигналов, анализ ДНК, кластерный анализ, анализ текста [24–26]. Суть метода заключается в том, чтобы выявить такие структуры, которые будут устойчиво сохраняться при топологических деформациях и искажениях.

В качестве исходных данных для топологического анализа будем рассматривать контур бинарного изображения. Далее на основе алгоритма SUSAN (*Smallest Univalued Segment Assimilation Nucleus*) из исходного контура формируется набор характерных точек, которые описывают особенности этого контура. Используя найденные точки, анализируем контур и сохраняем только точки с указанным шагом. В результате получим разреженное множество точек.

**Определение 1.** Разреженное множество точек  $V = \{v_1, v_2, \dots, v_n\}$  – это множество характерных точек исходного контура, расположенных друг от друга с заданным шагом, где  $n$  – это количество точек.

Строительными блоками для топологического анализа являются симплексы.

**Определение 2.**  $k$ -мерным симплексом  $\sigma$  будем называть выпуклую оболочку из  $k+1$  аффинно-независимых точек, т.е.  $\sigma = \langle a_0, a_1, \dots, a_k \rangle$ .

В работе используются симплексы следующих размерностей: 0, 1 и 2. Причём при  $k=0$  симплекс является точкой, при  $k=1$  – отрезком, а при  $k=2$  – треугольником (рис. 1).

Пример разреженного множества точек, полученного из бинарного изображения (рис. 2а), показан на рис. 2б.

**Определение 3.** Симплициальный комплекс  $K$  представляет собой конечное множество симплексов при условии, что граница каждого симплекса принадлежит  $K$  и для двух любых симплексов справедливо:

$\sigma_1 \cap \sigma_2 = \emptyset$  или  $\sigma_1 \cap \sigma_2$  имеют общую грань ( $\sigma_1, \sigma_2 \in K$ ).

Пример симплициального комплекса можно видеть на рис. 2з, который включает множество симплексов в виде точек, отрезков и треугольников.

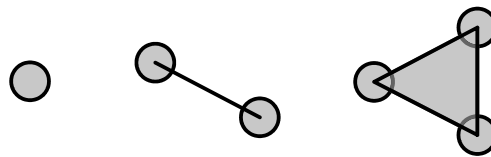


Рис. 1. Симплексы в виде точки, отрезка и треугольника

**Определение 4.**  $k$ -мерной цепью будем называть подмножество  $k$ -мерных симплексов в симплициальном комплексе  $K$ .

Например, на рис. 2б 0-мерные цепи представляют собой произвольное подмножество точек из множества  $V$ . Общее количество подмножеств из этих точек равно  $2^n$ . По аналогии, одномерные цепи – это подмножества из отрезков, а двумерные – из треугольников.

**Определение 5.** Группой цепей  $C_k$  является множество  $k$ -мерных цепей симплициального комплекса  $K$ .

**Определение 6.**  $k$ -мерный цикл – это замкнутая последовательность из цепей размерности  $k$ .

Например, на рис. 2д сформированы циклы, которые образованы замкнутой последовательностью отрезков.

**Определение 7.** Группа циклов  $Z_k$  – это множество циклов размерности  $k$  симплициального комплекса  $K$ .

**Определение 8.**  $k$ -мерная граница – это замкнутая последовательность из  $k$ -мерных цепей вокруг симплексов размерности  $k+1$ .

Например, границей для двумерного симплекса (треугольника) будут три одномерных симплекса (три отрезка).

**Определение 9.** Группа границ  $B_k$  – это множество  $k$ -мерных границ симплициального комплекса  $K$ , при этом  $B_k \subset C_k$ .

**Определение 10.**  $k$ -я гомологическая группа представляет собой фактор-группу  $H_k = Z_k/B_k$  и состоит из гомологических классов.

В работе используются фактор-группы  $H_0$  и  $H_1$ , для которых гомологическими классами являются компоненты связности и пустоты соответственно. Пустоты на плоскости обычно называют дырами, которые ограничены отрезками.

Чтобы определить количество компонент связности и количество дыр, применяются числа Бетти  $\beta_0$  и  $\beta_1$ , которые описывают топологические особенности симплициального комплекса и вычисляются как ранг фактор-группы:  $\beta_k = \text{rank}(H_k)$ . В то же время:  $\text{rank}(H_k) = \text{rank}(Z_k) - \text{rank}(B_k)$ . Ранг группы соответствует мощности минимального подмножества порождающих её элементов.

Например, пусть симплициальный комплекс состоит из трёх точек и трёх отрезков, соединяющих эти точки, т.е.:

$$K = \{ \langle v_1 \rangle, \langle v_2 \rangle, \langle v_3 \rangle, \langle v_1, v_2 \rangle, \langle v_1, v_3 \rangle, \langle v_2, v_3 \rangle \},$$

где  $v_i \in V (i = 1, 2, 3)$ .

Для  $k = 1$  получим:

1) множество одномерных цепей в виде отрезков

$$C_1(K) = \{ \emptyset, \langle v_1, v_2 \rangle, \langle v_1, v_3 \rangle, \langle v_2, v_3 \rangle, \langle v_1, v_2 \rangle + \langle v_1, v_3 \rangle, \langle v_1, v_2 \rangle + \langle v_2, v_3 \rangle, \langle v_1, v_3 \rangle + \langle v_2, v_3 \rangle, \langle v_1, v_2 \rangle + \langle v_1, v_3 \rangle + \langle v_2, v_3 \rangle \};$$

2) множество циклов, образованных отрезками

$$Z_1(K) = \{ \emptyset, \langle v_1, v_2 \rangle + \langle v_1, v_3 \rangle + \langle v_2, v_3 \rangle \};$$

3) пустое множество границ

$$B_1(K) = \{ \emptyset \}.$$

При этом  $\text{rank}(H_k) = 1$ , т.е. количество дыр равно 1. Если же рассматривать треугольник как двумерный симплекс, то количество дыр равно 0. В обоих случаях количество компонент связности равно 1.

Например, на рис. 2г присутствует 5 компонент связности и 0 дыр, а на рис. 2д – 1 компонента связности и 4 дыры.

**Определение 11.** Фильтрация  $K_0 \subset K_1 \subset K_2 \subset \dots \subset K_n$  – это вложенная последовательность симплициальных комплексов.

**Определение 12.** Персистентная гомология  $H_k^{i,j}$  – это множество  $k$ -х гомологических классов для фильтрации  $K_0 \subset K_1 \subset K_2 \subset \dots \subset K_n$ , причём создание класса происходит в симплициальном комплексе  $K_i$  и его существование продолжается до  $K_j (i, j \in 0, 1, 2, \dots, n; i < j)$ .

Процесс формирования симплициальных комплексов в зависимости от радиуса выглядит следующим образом (рис. 2б-ж). Вокруг каждой точки из разреженного множества  $V$  строится круг радиуса  $r$ . Начальным расстоянием  $r$  является минимальное расстояние между точками, максимальное – наибольшее расстояние. Если круги двух различных точек пересекаются, то эти точки соединяются отрезком. Если же пересекаются круги трёх точек, то они образуют треугольник. При увеличении радиуса количество симплексов и топологических свойств изменяется: число симплексов увеличивается, образуются новые компоненты связности и дыры. Дыра существует до тех пор, пока все образующие её точки не будут соединены друг с другом. При увеличении расстояния  $r$  число компонент связности станет равно единице, так как все имеющиеся компоненты объединяются в одну (рис. 2ж). Симплициальный комплекс, построенный по такому принципу, называют комплексом Чеха [27].

**Определение 13.**  $k$ -баркодом для фильтрации симплициальных комплексов множества точек  $V$  будем называть множество пар вида:

$$\text{barcode}_k(V) = \{ (d_i, l_i) \}, (i = 1, 2, \dots),$$

где  $d_i$  – координата создания топологической особенности,  $d_i$  принимает одно из возможных значений радиуса  $r$ ;  $l_i$  – длина существования топологической особенности.

Например, на рис. 2з можно видеть 1 – баркод, показывающий информацию о существовании каждой дыры.

Топологические особенности, длина которых не превышает некоторого значения  $\epsilon$ , будем считать шумом, искажающим представление о расположении объектов. Такие особенности не подлежат рассмотрению. В статье используется значение  $\epsilon$ , равное 1/3 от средней длины всех отрезков баркода. Стойкие особенности проявляются на длинных отрезках баркода, длина которых больше  $\epsilon$ , и являются основой для дальнейшего анализа (рис. 2и).

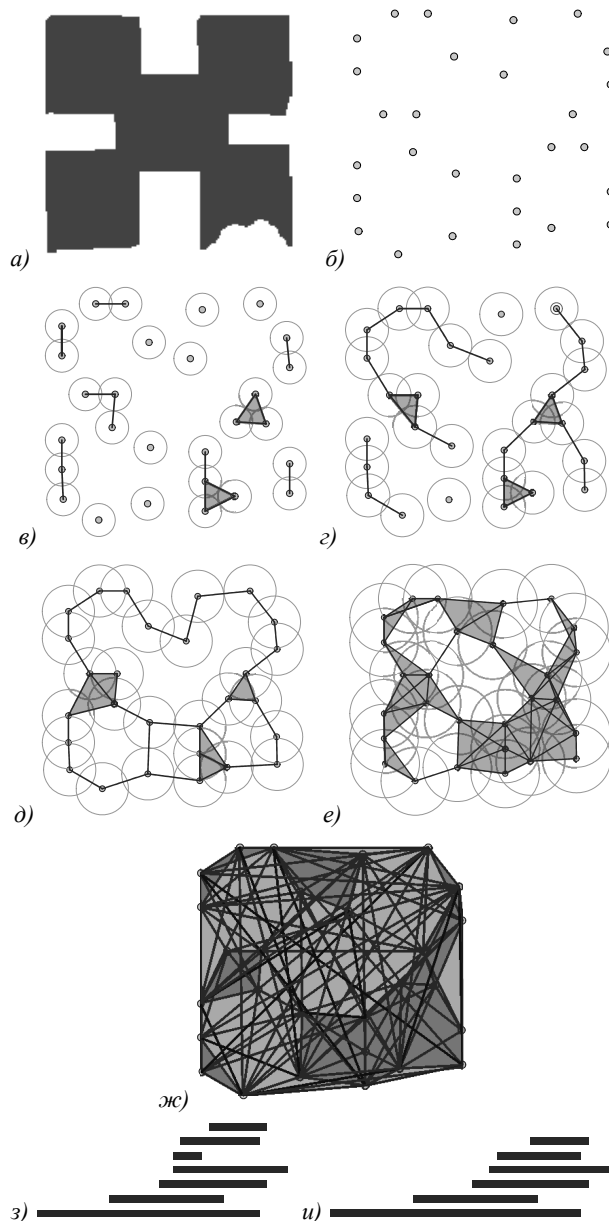


Рис. 2. Исходное бинарное изображение (а); разреженное пространство точек (б); процесс построения симплициальных комплексов при различных значениях  $r$  (0; 11; 15; 19; 27; 80) (б, в, г, д, е, ж); баркод с шумом (з); баркод без шума (и)

## 2. Алгоритм сравнения баркодов пространственных объектов

Исходными данными для сопоставления объектов являются два множества точек контуров:

$$X = \{x_1, x_2, \dots, x_n\} \text{ и } Y = \{y_1, y_2, \dots, y_m\}.$$

По каждому набору точек строятся баркоды. Рассмотрим баркоды

$$\text{barcode}_1(X) = \{(d_X^i, l_X^i)\} \quad (i = 1, 2, \dots, w_X) \text{ и}$$

$$\text{barcode}_1(Y) = \{(d_Y^j, l_Y^j)\} \quad (j = 1, 2, \dots, w_Y)$$

для анализа дыр. Алгоритм работает аналогично для анализа компонентов связности. По сути мы получаем два множества топологических признаков, которые содержат данные по каждой дыре. Причём количество дыр  $w_X$  и  $w_Y$  может быть различным. Требуется на основе информации из баркодов определить степень схожести топологических признаков двух объектов. Длинные отрезки баркода соответствуют устойчивым характеристикам и играют ключевую роль при идентификации объектов. Если длины отрезков баркода меньше некоторого заданного порога, т.е.  $l_X^i, l_Y^j < \varepsilon$  ( $i \in 1, 2, \dots, w_X; j \in 1, 2, \dots, w_Y$ ), то они являются шумом и не включаются в анализ.

Элементы баркодов  $\text{barcode}_1(X)$  и  $\text{barcode}_1(Y)$  отсортируем по значениям  $l_X^i$  и  $l_Y^j$  в порядке убывания, т.е.  $l_X^i \geq l_X^{i+1}$  ( $i \in 1, 2, \dots, w_X-1$ ) и  $l_Y^j \geq l_Y^{j+1}$  ( $j \in 1, 2, \dots, w_Y-1$ ).

Найдём наибольший отрезок среди всех отрезков двух баркодов:  $\max = \max(l_X^i, l_Y^j)$ . Если  $l_X^i < \max$ , то  $l_X^i = l_X^i \cdot (\max/l_X^i)$  ( $i = 2, 3, \dots, w_X$ ). Если же  $l_Y^j < \max$ , то  $l_Y^j = l_Y^j \cdot (\max/l_Y^j)$  ( $j = 2, 3, \dots, w_Y$ ).

После чего вычислим сумму:

$$S_X = \sum_{i=1}^{w_X} l_X^i, \quad S_Y = \sum_{j=1}^{w_Y} l_Y^j.$$

Исходя из общей суммы, определим вес каждого отрезка в баркодах  $\text{barcode}_1(X)$  и  $\text{barcode}_1(Y)$ :

$$p = \{p_1, p_2, \dots, p_{w_X}\}, \text{ где } p_i = \frac{l_X^i}{S_X} \quad (i = 1, 2, \dots, w_X)$$

$$\text{и } t = \{t_1, t_2, \dots, t_{w_Y}\}, \text{ где } t_j = \frac{l_Y^j}{S_Y} \quad (j = 1, 2, \dots, w_Y).$$

Для того чтобы определить похожесть двух баркодов, вычислим отношение  $l_X^i$  и  $l_Y^j$ , которые имеют одинаковые индексы:

$$z = \{z_1, z_2, \dots, z_{\min(w_X, w_Y)}\},$$

где

$$z_i = \begin{cases} l_X^i / l_Y^i, & \text{если } l_X^i < l_Y^i, \\ l_Y^i / l_X^i, & \text{иначе,} \end{cases}$$

$$(i = 1, 2, \dots, \min(w_X, w_Y)).$$

В местах, где индекс выходит за пределы одного из значений  $w_X$  или  $w_Y$ , примем, что

$$z_{\min(w_X, w_Y)+1} = z_{\min(w_X, w_Y)+2} = \dots = z_{\max(w_X, w_Y)} = 0.$$

Далее умножим значения множества  $z$  на соответствующие значения максимального по мощности множества  $p$  или  $t$  и сложим эти произведения, чтобы вычислить показатель схожести  $Q$ :

$$Q = \begin{cases} \sum_{i=1}^{w_Y} z_i \cdot p_i, & \text{если } w_X > w_Y, \\ \sum_{i=1}^{w_X} z_i \cdot t_i, & \text{иначе.} \end{cases} \quad (1)$$

Имея данные показатели для всех объектов, получаем возможность выделить максимально похожий по топологии объект и принять его за идентичный.

Исследования показали, что баркод по дырам несёт большую часть процента схожести по сравнению с компонентами связности. Это можно выразить так: *HolePart* – доля баркода по дырам, а  $1 - \text{HolePart}$  – по компонентам связности. В статье экспериментально принято, что *HolePart* = 75, т.е. баркод по дырам даёт вклад 75 процентов, а баркод по компонентам связности – 25. Показатель схожести  $Q$  по дырам умножается на 0,75, а по компонентам связности – на 0,25. Таким образом, полностью схожие объекты дадут в сумме 100 процентов.

## 3. Результаты исследований

Исследование проводилось сначала для сопоставления природных объектов на разных масштабах, а затем для муниципальных с учетом деформации и искажений.

В качестве аналога в работе рассмотрено метрическое сопоставление объектов. В этом методе рассчитываются две матрицы расстояний  $\|r_{ij}^X\|$  ( $i, j = 1, 2, \dots, n$ ) и  $\|r_{ij}^Y\|$  ( $i, j = 1, 2, \dots, m$ ) между всеми точками множеств  $X = \{x_1, x_2, \dots, x_n\}$  и  $Y = \{y_1, y_2, \dots, y_m\}$ . Будем считать, что  $n \leq m$ . Для количественной оценки вычисляется параметр  $\delta_{ij} = r_{ij}^X - r_{ij}^Y$ . Далее для каждой точки  $x_i \in X$  определяется возможная сопряжённая ей точка, исходя из выражения:  $\delta_i = \min\{\delta_{1i}, \delta_{2i}, \dots, \delta_{mi}\}$ . Считается, что две точки сопряжённые, если  $\delta_i$  не превышает порога  $\Delta$ . Далее подсчитывается количество сопряжённых точек и оценивается процент схожести относительно общего числа точек.

### 3.1. Сопоставление природных объектов при их масштабировании и деформации

Рассмотрим в качестве исходных данных два снимка разных масштабов одного и того же острова (рис. 3). Как видно из рисунков, они обладают одной и той же структурой, но имеют различия по наличию деталей.

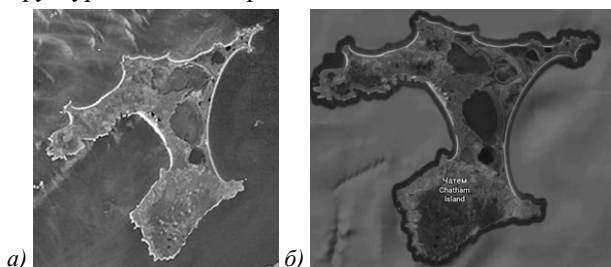


Рис. 3. Исходные снимки острова, сделанные под разным углом и на разных масштабах: 1:30000 (а), 1:10000 (б)

Приведём два изображения к бинарному виду на основе метода Отсу после предварительного сглаживания исходных изображений с параметром размытия по Гауссу, равному 3 (рис. 4а, б). Кроме того, для бинарного изображения (рис. 4б) выполним сглаживающие контуры со значением 10 пикселей (рис. 4в).

Далее каждое бинарное изображение представим в виде разреженного пространства точек и построим баркоды.

На рис. 5 показаны 1 – баркоды объектов.

Те топологические особенности, которые не несут ключевой информации, будем считать шумом. Баркоды без шума имеют более похожую общую структуру и, соответственно, схожесть объектов между собой.

Численные характеристики для сравнения объектов по их баркодам без шума приведены в табл. 1, а в табл. 2 – при метрическом сопоставлении.

Табл. 1. Численные характеристики топологической схожести природных объектов на разных масштабах

	Острова					
	б			в		
а	63,35	24,55	<b>87,90</b>	63,30	22,02	<b>85,32</b>
б				67,14	24,75	<b>91,89</b>

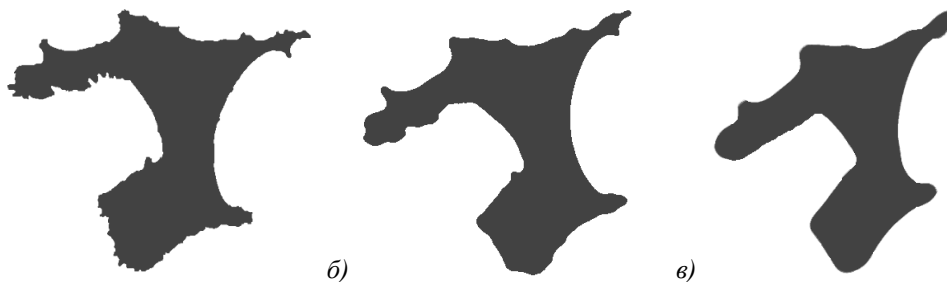


Рис. 4. Бинарные изображения острова с разных масштабов (а, б); деформация бинарного изображения (в)

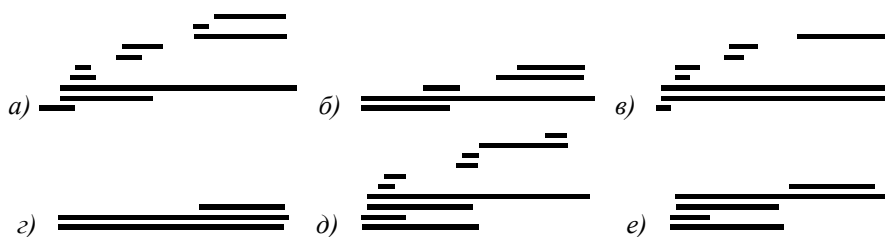


Рис. 5. Баркоды объектов с шумом (а, в, д); без шума (б, г, е) для бинарных изображений из рис. 4



Рис. 6. Бинарные изображения муниципальных объектов: исходный объект (а); здание с растяжением (б); здание с растяжением и искажениями (в)

Для каждой пары объектов указана схожесть по дырам, компонентам связности и итоговая.

Табл. 2. Численные характеристики при метрическом сопоставлении природных объектов на разных масштабах

	Острова	
	б	в
а	<b>89</b>	<b>67</b>
б		<b>72</b>

### 3.2. Сопоставление муниципальных объектов при их деформации

Сопоставление объектов в муниципальных ГИС осложняется выделением контуров в связи с наличием перекрывающихся объектов на карте. Над исходным зданием (рис. 6а) выполним деформацию за счёт вертикального сжатия на 20% и горизонтального растяжения также на 20% (рис. 6б). Чтобы получить изображение на рис. 6в, сделаем искажения вручную в отдельных угловых частях бинарного изображения на рис. 6б. Далее проводим аналогичные исследования, как в подпараграфе 3.1, результаты которых показаны на рис. 7, 8.

Численные характеристики схожести, вычисленные по выражению (1) и при метрическом совмещении точек, представлены в табл. 3 и 4.

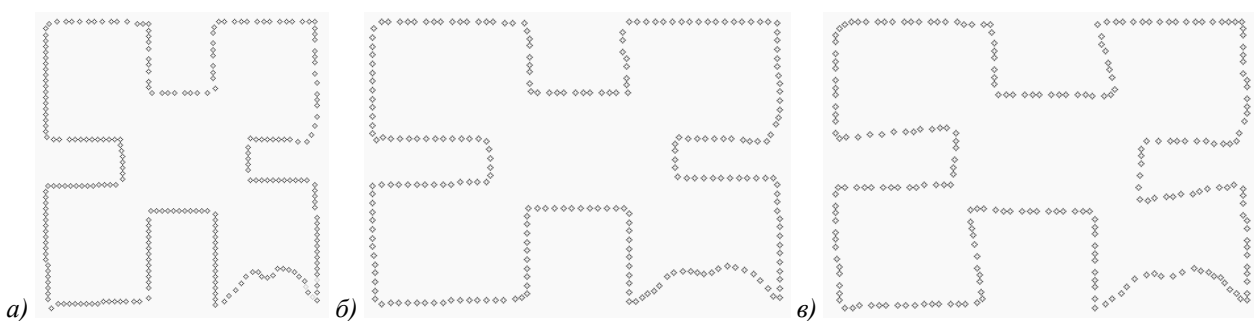


Рис. 7. Разреженное пространство точек муниципальных объектов

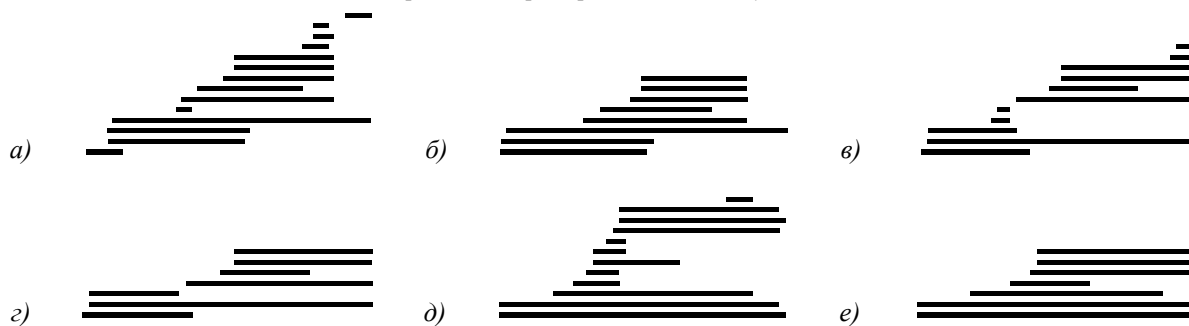


Рис. 8. Баркоды муниципальных объектов с учётом деформации и искажений с шумом (а, в, д.); без шума (б, г, е)

Табл. 3. Численные характеристики топологической схожести муниципальных объектов при деформациях и искажениях

	Муниципальные объекты					
	б			в		
а	50,17	23,87	<b>74,04</b>	63,25	24,35	<b>87,60</b>
б				60,33	24,50	<b>84,83</b>

Табл. 4. Численные характеристики при метрическом совмещении муниципальных объектов с учётом деформации и искажений

	Муниципальные объекты	
	б	в
а	<b>73</b>	<b>54</b>
б		<b>75</b>

Анализ результатов показывает близкое совмещение баркодов при незначительных удалениях деталей после деформации. Даже при сильном масштабировании структура объекта сохраняется. Т.е. глобальные характеристики исходного объекта совпадают с глобальными характеристиками объекта после генерализации или масштабирования, несмотря на то что детальная информация на одном из объектов карты отсутствует. При анализе муниципальных объектов также были выявлены общие топологические особенности между объектами после деформации и искажений. При сравнении с метрическим сопоставлением предложенный подход показал преимущества при анализе объектов, которые подвержены значительной деформации при масштабировании, а также при искажениях.

Для исследования при сопоставлении разных пространственных объектов были взяты бинарные изображения острова (рис. 4б) и муниципального объекта (рис. 6а). В результате работы алгоритма схожесть

объектов составила 56,23%, включая схожесть по дырам (32,68% из 75% возможных) и схожесть по компонентам связности (23,55% из 25% допустимых). Анализ показывает, что основное различие достигается за счёт несоответствия баркодов по дырам, в то время как по компонентам связности баркоды имеют достаточно близкое значение.

**Заключение**

В статье рассмотрена задача сопоставления объектов на разнородных картах. Основу алгоритма составляет анализ формы объектов, которые подвергаются масштабированию, генерализации, деформациям и искажениям, однако общая структура объекта сохраняется. К таким деформациям инвариантны методы персистентной гомологии. Для сопоставления объектов из разных карт анализируются их баркоды. Приведены результаты исследований сопоставления пространственных объектов для природных и муниципальных карт. Проведённые эксперименты показали, что основное влияние на схожесть объектов при их совмещении оказывают топологические особенности в виде дыр. Чем больше у исходных объектов устойчивых дыр, тем больше будет качество при их совмещении. Разработанный подход может быть использован, например, для решения задачи автоматического заполнения атрибутивных данных карты одного масштаба на основе данных карты другого масштаба. Это позволит осуществить более быструю интеграцию пространственных и семантических данных многомасштабных карт.

**Благодарности**

Исследование выполнено при финансовой поддержке РФФИ и администрации Владимирской области в рамках научного проекта № 17-47-330387.

*Литература*

1. **Wallgrün, J.** Qualitative matching of spatial information / J.O. Wallgrün, D. Wolter, K. Richter // GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems. – 2010. – P. 300-309.
2. **Mustière, S.** Matching networks with different levels of detail / S. Mustière, T. Devogele // GeoInformatica. – 2008. – Vol. 12, Issue 4. – P. 435-453.
3. **Biedl, T.** Planar matchings for weighted straight skeletons / T. Biedl, S. Huber, P. Palfrader // International Journal of Computational Geometry and Applications. – 2016. Vol. 26, Issues 3-4. – P. 211-229.
4. **Ефимов, А.И.** Алгоритм поэтапного уточнения проективного преобразования для совмещения изображений / А.И. Ефимов, А.И. Новиков // Компьютерная оптика. – 2016. – Т. 40, № 2. – С. 258-265. – DOI: 10.18287/2412-6179-2016-40-2-258-265.
5. **Zhao, L.** Shape matching algorithm based on shape contexts / L. Zhao, Q. Peng, B. Huang // IET Computer Vision. – 2015. – Vol. 9, Issue 5. – P. 681-690.
6. **Еремеев, С.В.** Comparison of urban areas based on database of topological relationships in geoinformational systems / S.V. Eremeev, D.E. Andrianov, V.A. Komkov // Pattern Recognition and Image Analysis. – 2015. – Vol. 25, No 2. – P. 314-320.
7. **Еремеев, С.** An approach to establishing the correspondence of spatial objects on heterogeneous maps based on methods of computational topology / S. Eremeev, K. Kuptsov, S. Romanov. – In: Analysis of images, social networks and texts (AIST 2017) / ed. by W. van der Aalst [et al.]. – Cham: Springer, 2018. – P. 172-182.
8. **Zhang, T.** Multi-task correlation particle filter for robust object tracking / T. Zhang, C. Xu, M. Yang // IEEE Conference on Computer Vision and Pattern Recognition. – 2017. Vol. 1(2). – P. 4819-4827.
9. **Садыков, С.С.** Алгоритм построения выпуклой оболочки бинарного изображения и формирование его безразмерных признаков / С.С. Садыков // Алгоритмы, методы и системы обработки данных. – 2015. – № 2(31). – С. 77-85.
10. **Ломов, Н.А.** Площадь дискового покрытия – дескриптор формы изображения / Н.А. Ломов, Л.М. Местецкий // Компьютерная оптика. – 2016. – Т. 40, № 4. – С. 516-525. – DOI: 10.18287/2412-6179-2016-40-4-516-525.
11. **Ломов, Н.А.** Классификация двумерных фигур с использованием скелетно-геодезических гистограмм толщин-расстояний / Н.А. Ломов, С.В. Сидякин, Ю.В. Визильтер // Компьютерная оптика. – 2017. – Т. 41, № 2. – С. 227-236. – DOI: 10.18287/2412-6179-2017-41-2-227-236.
12. **Eitz, M.** Sketch-based shape retrieval / M. Eitz, R. Richter, T. Boubekeur, K. Hildebrand, M. Alexa // ACM Transactions on Graphics. – 2012. – Vol. 31, Issue 4. – P. 31.
13. **Bai, X.** Path similarity skeleton graph matching / X. Bai, L.J. Latecki // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2008. – Vol. 30, Issue 7. – P. 1282-1292.
14. **Kališnik, S.** A higher-dimensional homologically persistent skeleton / S. Kališnik, V. Kurlin, D. Lesnik // Advances in Applied Mathematics. – 2019. – Vol. 102. – P. 113-142.
15. **Carlsson, G.** Persistence barcodes for shapes / G. Carlsson, A. Zomorodian, A. Collins, L. Guibas // Proceedings of the 2004 Eurographics, ACM SIGGRAPH Symposium on Geometry Processing. – 2004. – P. 124-135.
16. **Skraba, P.** Persistence-based segmentation of deformable shapes / P. Skraba, M. Ovsjanikov, F. Chazal, L. Guibas // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition – Workshops. – 2010. – P. 45-52.
17. **Förstner, W.** Detecting interpretable and accurate scale-invariant keypoints / W. Förstner, T. Dickscheid, F. Schindler // IEEE 12<sup>th</sup> International Conference on Computer Vision. – 2009. – P. 2256-2263.
18. **Su, Y.** Contour guided hierarchical model for shape matching / Y. Su, Y. Liu, B. Cuan, N. Zheng // IEEE International Conference on Computer Vision (ICCV). – 2015. – P. 1609-1617.
19. **Ahmed, M.** Local persistent homology based distance between maps / M. Ahmed, B. Fasy, C. Wenk // Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. – 2014. – P. 43-52.
20. **Bendich, P.** Homology and robustness of level and interlevel sets / P. Bendich, H. Edelsbrunner, D. Morozov, A. Patel // Homology, Homotopy and Applications. – 2013. Vol. 15. – P. 51-72.
21. **Collins, A.** A barcode shape descriptor for curve point cloud data / A. Collins, A. Zomorodian, G. Carlsson, L. Guibas // Computers and Graphics. – 2004. – Vol. 28. – P. 881-894.
22. **Carlsson, E.** An algebraic topological method for feature identification / E. Carlsson, G. Carlsson, V. de Silva, S. Fortune // International Journal of Computational Geometry and Applications. – 2006. Vol. 16(4). – P. 291-314.
23. **Carlsson, G.** Topological pattern recognition for point cloud data / G. Carlsson // Acta Numerica. – 2014. – Vol. 23. – P. 289-368.
24. **Lum, P.Y.** Extracting insights from the shape of complex data using topology / P.Y. Lum, G. Singh, A. Lehman, T. Ishkanov, M. Vejdemo-Johansson, M. Alagappan, J. Carlsson, G. Carlsson // Scientific Reports. – 2013. – Vol. 3. – P. 12-36.
25. **Макаренко, Н.Г.** Распознавание текстур на цифровых изображениях методами вычислительной топологии / Н.Г. Макаренко, Ф.А. Уртьев, И.С. Князева, Д.Б. Малкова, И.Т. Пак, Л.М. Каримова // Современные проблемы дистанционного зондирования Земли из космоса. – 2015. – Т. 12, № 1. – С. 131-144.
26. **Zhu, X.** Persistent homology: An introduction and a new text representation for natural language processing / X. Zhu // Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence. – 2013. – P. 1953-1959.
27. **Edelsbrunner, H.** Computational topology: An introduction / H. Edelsbrunner. – American Mathematical Society, 2009.

*Сведения об авторах*

**Еремеев Сергей Владимирович**, 1980 года рождения, кандидат технических наук, доцент, работает на кафедре информационных систем Муромского института (филиала) Владимирского государственного университета. Область научных интересов: обработка пространственных данных, геоинформационные системы, топологический анализ данных. E-mail: [sv-eremeev@yandex.ru](mailto:sv-eremeev@yandex.ru).

**Андрианов Дмитрий Евгеньевич**, 1973 года рождения, доктор технических наук, доцент, заведующий кафедрой информационных систем Муромского института (филиала) Владимирского государственного университета. Область научных интересов: исследование теоретических основ обработки пространственных данных, разработка методов и моделей обработки и анализа данных в геоинформационных системах, топология пространственно-распределенных объектов. E-mail: [AndrianovDE@inbox.ru](mailto:AndrianovDE@inbox.ru).

**Титов Виталий Семенович**, 1943 года рождения, доктор технических наук, профессор, заведующий кафедрой вычислительной техники Юго-западного государственного университета. Область научных интересов: распознавание образов и цифровая обработка изображений. E-mail: [titov-kstu@rambler.ru](mailto:titov-kstu@rambler.ru).

ГРНТИ: 28.23.15.

Поступила в редакцию 14 января 2019 г. Окончательный вариант – 29 июля 2019 г.

## An algorithm for matching spatial objects of different-scale maps based on topological data analysis

S.V. Eremeev<sup>1</sup>, D.E. Andrianov<sup>1</sup>, V.S. Titov<sup>2</sup>

<sup>1</sup> Vladimir State University, Vladimir, Russia,

<sup>2</sup> Southwest State University, Kursk, Russia

### Abstract

A problem of automatic comparison of spatial objects on maps with different scales for the same locality is considered in the article. It is proposed that this problem should be solved using methods of topological data analysis. The initial data of the algorithm are spatial objects that can be obtained from maps with different scales and subjected to deformations and distortions. Persistent homology allows us to identify the general structure of such objects in the form of topological features. The main topological features in the study are the connectivity components and holes in objects. The paper gives a mathematical description of the persistent homology method for representing spatial objects. A definition of a barcode for spatial data, which contains a description of the object in the form of topological features is given. An algorithm for comparing feature barcodes was developed. It allows us to find the general structure of objects. The algorithm is based on the analysis of data from the barcode. An index of objects similarity in terms of topological features is introduced. Results of the research of the algorithm for comparing maps of natural and municipal objects with different scales, generalization and deformation are shown. The experiments confirm the high quality of the proposed algorithm. The percentage of similarity in the comparison of natural objects, while taking into account the scale and deformation, is in the range from 85 to 92, and for municipal objects, after stretching and distortion of their parts, was from 74 to 87. Advantages of the proposed approach over analogues for the comparison of objects with significant deformation at different scales and after distortion are demonstrated.

**Keywords:** persistent homology, barcode of spatial object, comparison of objects, analysis of topological features, multi-scale maps.

**Citation:** Eremeev SV, Andrianov DE, Titov VS. An algorithm for matching spatial objects of different-scale maps based on topological data analysis. *Computer Optics* 2019; 43(6): 1021-1029. DOI: 10.18287/2412-6179-2019-43-6-1021-1029.

**Acknowledgements:** The work was funded by the Russian Foundation for Basic Research (RFBR) and Vladimir region authorities under the research project No. 17-47-330387.

### References

- [1] Wallgrün J, Wolter D, Richter K. Qualitative matching of spatial information. *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems* 2010: 300-309.
- [2] Mustière S, Devogele T. Matching networks with different levels of detail. *GeoInformatica* 2008; 12(4): 435-453.
- [3] Biedl T, Huber S, Palfrader P. Planar matchings for weighted straight skeletons. *International Journal of Computational Geometry and Applications* 2016; 26(3-4): 211-229.
- [4] Efimov AI, Novikov AI. An algorithm for multistage projective transformation adjustment for image superimposition [In Russian]. *Computer Optics* 2016; 40(2): 258-265. DOI: 10.18287/2412-6179-2016-40-2-258-265.
- [5] Zhao L, Peng Q, Huang B. Shape matching algorithm based on shape contexts. *IET Comp Vis* 2015; 9(5): 681-690.
- [6] Eremeev SV, Andrianov DE, Komkov VA. Comparison of urban areas based on database of topological relationships in geoinformational systems. *Pattern Recognition and Image Analysis* 2015; 25(2): 314-320.
- [7] Eremeev S, Kuptsov K, Romanov S. An approach to establishing the correspondence of spatial objects on heterogeneous maps based on methods of computational topology. In Book: van der Aalst W, et al, eds. *Analysis of images, social networks and texts (AIST 2017)*. Cham: Springer; 2018: 172-182.
- [8] Zhang T, Xu C, Yang M. Multi-task correlation particle filter for robust object tracking. *IEEE Conf Comp Vis Pattern Recog* 2017; 1(2): 4819-4827.



- [9] Sadykov SS. Algorithm for the construction of a convex hull of a binary image and the formation of its dimensionless features [In Russian]. Algorithms, Methods and Data Processing Systems 2015; 2(31): 77-85.
- [10] Lomov NA, Mestetskiy LM. Area of the disk cover as an image shape descriptor [In Russian]. Computer Optics 2016; 40(4): 516-525. DOI: 10.18287/2412-6179-2016-40-4-516-525.
- [11] Lomov NA, Sidiyakin SV, Visilter YuV. Classification of two-dimensional figures using skeleton-geodesic histograms of thicknesses and distances. Computer Optics 2017; 41(2): 227-236. DOI: 10.18287/2412-6179-2017-41-2-227-236.
- [12] Eitz M, Richter R, Boubekur T, Hildebrand K, Alexa M. Sketch-based shape retrieval. ACM Transactions on Graphics 2012; 31(4): 31.
- [13] Bai X, Latecki LJ. Path similarity skeleton graph matching. IEEE Transactions on Pattern Analysis and Machine Intelligence 2008; 30(7): 1282-1292.
- [14] Kališnik S, Kurlin V, Lesnik D. A higher-dimensional homologically persistent skeleton. Advances in Applied Mathematics 2019; 102: 113-142.
- [15] Carlsson G, Zomorodian A, Collins A, Guibas L. Persistence barcodes for shapes. Proc 2004 Eurographics, ACM SIGGRAPH Symposium on Geometry Processing 2004: 124-135.
- [16] Skraba P, Ovsjanikov M, Chazal F, Guibas L. Persistence-based segmentation of deformable shapes. Proc CVPRW 2010: 45-52.
- [17] Förstner W, Dickscheid T, Schindler F. Detecting interpretable and accurate scale-invariant keypoints. IEEE 12<sup>th</sup> ICCV 2009: 2256-2263.
- [18] Su Y, Liu Y, Cuan B, Zheng N. Contour guided hierarchical model for shape matching. IEEE ICCV 2015: 1609-1617.
- [19] Ahmed M, Fasy B, Wenk C. Local persistent homology based distance between maps. Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems 2014: 43-52.
- [20] Bendich P, Edelsbrunner H, Morozov D, Patel A. Homology and robustness of level and interlevel sets. Homology, Homotopy and Applications 2013; 15: 51-72.
- [21] Collins A, Zomorodian A, Carlsson G, Guibas L. A barcode shape descriptor for curve point cloud data. Computers and Graphics 2004; 28: 881-894.
- [22] Carlsson E., Carlsson G., de Silva V., and Fortune S. An algebraic topological method for feature identification. International Journal of Computational Geometry and Applications 2006; 16(4): 291-314.
- [23] Carlsson G. Topological pattern recognition for point cloud data. Acta Numerica 2014; 23: 289-368.
- [24] Lum PY, Singh G, Lehman A, Ishkanov T, Vejdemo-Johansson M, Alagappan M, Carlsson J, Carlsson G. Extracting insights from the shape of complex data using topology. Sci Rep 2013; 3: 12-36.
- [25] Makarenko ND, Yuriev FA, Knyazeva IS, Malkova DB, Pak IT, Karimova LM. Texture recognition on digital images using computational topology methods [In Russian]. Modern problems of remote sensing the Earth from space 2015; 12(1): 131-144.
- [26] Zhu X. Persistent homology: An introduction and a new text representation for natural language processing. In: Proceedings of the 23rd International Joint Conference on Artificial Intelligence 2013: 1953-1959.
- [27] Edelsbrunner H. Computational topology: An introduction. American Mathematical Society; 2009.

#### *Author's information*

**Sergey Vladimirovich Eremeev** (b. 1980), candidate of Technical Sciences, Associate Professor. Currently he works at Information Systems department, Murom Institute (branch) of Vladimir State University. Research interests are spatial data processing, geographic information systems and topological data analysis. E-mail: [sv-eremeev@yandex.ru](mailto:sv-eremeev@yandex.ru).

**Dmitry Evgenyevich Andrianov** (b. 1973), doctor of Engineering, Associate Professor, Head of Information Systems department of MI VISU. Research interests are theoretical bases of spatial data processing, development of methods and models of data processing, data analysis in geoinformation systems and topology of spatially distributed objects. E-mail: [AndrianovDE@inbox.ru](mailto:AndrianovDE@inbox.ru).

**Vitaliy Semenovich Titov** (b. 1943), doctor of Engineering, professor, Head of Computer Technology department of SWSU. Research interests are image recognition and digital image processing. E-mail: [titov-kstu@rambler.ru](mailto:titov-kstu@rambler.ru).

*Received January 14, 2019. The final version – July 29, 2019.*