



Math-Net.Ru

Общероссийский математический портал

С. В. Еремеев, А. В. Абакумов, Д. Е. Андрианов, Д. В. Титов, Метод разложения изображения по топологическим признакам,
Компьютерная оптика, 2022, том 46, выпуск 6, 939–947

<https://www.mathnet.ru/co1089>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<https://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 18.97.9.170

21 апреля 2025 г., 23:01:56



Метод разложения изображения по топологическим признакам

С.В. Еремеев¹, А.В. Абакумов¹, Д.Е. Андрианов¹, Д.В. Титов²

¹ Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»,
602264, Россия, г. Муром, ул. Орловская, д. 23;

² ФГБОУ ВО «Юго-Западный государственный университет»,
305040, Россия, г. Курск, ул. 50 лет Октября, д. 94

Аннотация

В статье предложен новый метод для разложения изображения на отдельные объекты интереса. В основе разработанного метода лежит использование персистентной гомологии. Показан процесс прямого и обратного преобразования изображения. Исходное изображение после прямого преобразования представляется как набор матриц, которые можно разделить на основные и детализирующие. Основные матрицы содержат информацию об основной структуре объектов на изображениях, а детализирующие включают данные о деталях этих объектов, а также о мелких объектах или шумовой составляющей. Показано, что существует определенная аналогия с Wavelet-преобразованием, но в основе предлагаемого метода заложена принципиально другая теоретическая база. Подробно описан численный пример, отражающий основную суть метода. Описаны свойства разложения, а также возможность использования стандартных алгебраических операций над матрицами разложения. Обратное преобразование позволяет учесть измененные свойства отдельных объектов и синтезировать новое изображение. Продемонстрированы перспективы применения разложения для решения практических задач. Разработаны алгоритмы для бинаризации изображений и удаления текста на неоднородном фоне. Анализ и обработка данных ведется с единых позиций в пространстве матриц разложения. Результаты бинаризации показали, что в сравнении с аналогами разработанный алгоритм покажет себя наилучшим образом в ситуациях, когда бинаризация применяется для выделения множества отдельных объектов. Полученные результаты алгоритма удаления текста на неоднородном фоне подтверждают, что информация полностью удаляется, не задевая остальные области на изображениях.

Ключевые слова: топологический анализ данных, персистентная гомология, баркод, топологические признаки, компоненты связности, разложение изображений, низкочастотные и высокочастотные матрицы разложения.

Цитирование: Еремеев, С.В. Метод разложения изображения по топологическим признакам / С.В. Еремеев, А.В. Абакумов, Д.Е. Андрианов, Д.В. Титов // Компьютерная оптика. – 2022. – Т. 46, № 6. – С. 939-947. – DOI: 10.18287/2412-6179-CO-1080.

Citation: Eremeev SV, Abakumov AV, Andrianov DE, Titov DV. Image decomposition method by topological features. Computer Optics 2022; 46(6): 939-947. DOI: 10.18287/2412-6179-CO-1080.

Введение

В последнее время в мировой и отечественной практике стали развиваться методы топологического анализа данных, в частности, персистентная гомология [1]. В своем большинстве эти методы ориентированы на обработку точечных данных, построение симплициальных комплексов и вычисление их топологических особенностей. Одним из ключевых достижений в этой сфере является выделение глобальных структур из топологических признаков на множестве исходных данных.

Отдельное направление для исследований получила область обработки изображений на основе персистентной гомологии. Стали появляться научные лаборатории, которые ведут исследования по разработке новых подходов для решения как классических

задач обработки изображений, так и принципиально новых. Так, например, в работах научной группы Ливерпульского университета под руководством В.А. Курлина получены интересные результаты для решения задач сегментации, скелетизации, поиска линейных объектов с позиций топологического анализа изображений [2]. В основе алгоритмов лежит получение топологических признаков, которые представляются в виде баркодов [1], персистентных ландшафтов [3] и персистентных изображений [4].

Полученные топологические признаки, представленные в таких форматах, используются для поиска засушливости территорий по космическим снимкам [5], для поиска пространственных объектов по шаблону [6], на разных масштабах [7], для сравнения подписей [8] и др. Более сложные дифференциально-топологические инварианты аппроксимирующего

отображения применяются для распознавания гомотопического типа объекта [9].

Несмотря на многочисленные разработки в области топологического анализа данных, все равно остается еще очень много нерешенных проблем. Одной из таких острых проблем является необходимость так представить исходное изображение в виде глобальных и детализирующих структур, чтобы при определенных изменениях в этих структурах была возможность синтезировать новое изображение. По сути, это получение некоторого прямого и обратного преобразования в разрезе топологических признаков, к которым относятся компоненты связности и дыры, формальное определение которых рассмотрено в [7].

Рассмотрим наиболее близкие подходы для разложения изображений на составляющие элементы. Разложение сигнала и изображений на составляющие играет важную роль в области распознавания образов, искусственного интеллекта, обработки медицинских снимков, робототехники, фильтрации данных и т.д. Очень популярным и часто встречаемым на практике является Wavelet-преобразование [10]. Одним из ключевых результатов при выполнении этого преобразования является разложение изображений на низкочастотные и высокочастотные составляющие. При этом низкие частоты содержат общую структуру изображения, а высокие – отражают детали. Общую схему представления изображения после Wavelet-преобразования можно показать следующим образом:

$$I = I_L + I_H,$$

где I – исходное изображение, I_L – изображение, содержащее низкие частоты, I_H – изображение с высокими частотами.

В предлагаемом методе тоже рассматривается разложение изображения на общую структуру объектов и их детали, но с точки зрения персистентной гомологии, т.е. разложение в разрезе топологических признаков. Таким образом, есть определенная аналогия с Wavelet-преобразованием, но в основе предлагаемого метода заложена принципиально другая теоретическая база.

Для разделения изображения на отдельные объекты также используется метод независимых компонент (Independent Component Analysis), который позволяет представить исходный сигнал как линейную комбинацию содержащихся в нем компонент [11]. Данный подход показал свое применение в ряде приложений не только в области обработки сигналов и изображений, но и в машинном обучении. Для поиска компонент применяется ковариационная матрица и требуется расчет собственных значений исходного сигнала, что влечет большие вычислительные ресурсы. В литературе предлагаются методы для разделения наложенных друг на друга изображений на основе FastICA [12].

Кроме того, существуют смежные с топологическим анализом данных подходы в области обработки изображений. Один из таких близких подходов рассмотрен в [13] для решения задачи сегментации мелких объектов, в основе которого лежит метод перколяции данных и анализ бинарных изображений на каждом из 256 порогов бинаризации изображения в градациях серого. В этом методе рассчитываются геометрические характеристики объекта на каждом уровне яркости, что позволяет сделать более качественную селекцию объектов. В работе [14] используется формирование суперпикселей на изображении с минимальной потерей исходной информации. Относительно полученных суперпикселей вычисляются различные признаки объекта, что позволяет использовать данный метод для решения широкого круга задач в области обработки изображений и геоинформатики. Авторы работы [15] применяют мозаичное разложение изображения как систему непересекающихся множеств и успешно используют этот подход для анализа формы объекта.

Технологии, рассмотренные выше, а именно, многопороговая бинаризация, слияние компонент связности, алгебраические операции над элементами изображения, также используются в предлагаемом методе, но с позиции топологического анализа данных.

Таким образом, в статье предлагается разработать новый метод для разложения изображения на отдельные матрицы по топологическим особенностям. Матрицы можно будет условно разделить на основные, которые содержат информацию об основной структуре объектов, и детализирующие, которые включают данные о деталях этих объектов, а также о мелких объектах или шумовой составляющей. Также предлагается показать практическое применение этого метода в сравнении с другими для решения конкретных практических задач. Для примера были выбраны две задачи: бинаризация изображений и удаление текстовой информации на неоднородном фоне. При этом решение практических задач будем рассматривать с единых позиций путем разложения изображений на составляющие их матрицы, которые формируются по определенным алгоритмам топологической фильтрации. Покажем, как эти матрицы будут использоваться для дальнейшего анализа изображений путем выполнения над ними классических алгебраических операций, а также для синтеза новых изображений в виде обратного преобразования из этих матриц.

1. Предлагаемый метод разложения изображения

Предпосылками для создания теории разложения изображений на матрицы послужили методы топологического анализа данных и непосредственно персистентная гомология. На текущий момент подход с использованием персистентной гомологии извлекает полезную информацию в виде топологических особенностей из слабоструктурированных данных. Эти

особенности позволяют выявить общие глобальные структуры из набора данных, причем эти структуры сохраняются даже при определенных искажениях и деформациях. Также в персистентной гомологии есть менее значимые характеристики баркодов, обычно называемые шумом.

В ходе проведенных исследований выявлено, что можно не только извлекать информацию из изображений, а разложить исходное изображение на составляющие в виде некоторых матриц, которые условно разделим на основные и детализирующие. Кроме того, появляется возможность сделать обратное преобразование из матриц разложения и получить новое изображение после обработки отдельных матриц. Основные матрицы содержат информацию об основных структурах исходных данных, а детализирующие матрицы – их детали. Если проводить аналогию с Wavelet-преобразованием, то можно сказать, что основные матрицы – это низкочастотные матрицы, а детализирующие – высокочастотные.

Рассмотрим на примере классический подход для получения топологических признаков на основе персистентной гомологии и покажем отличия с предлагаемым подходом разложения исходного изображения на основную и детализирующую матрицы. В качестве топологической особенности будем использовать компоненты связности.

На рис. 1 продемонстрирован процесс построения топологических признаков в виде баркода. Для данного численного примера будем рассматривать трехбитное изображение со значениями яркости в диапазоне от 0 до 7. На вход подается исходное изображение (на рис. 1 слева), и на каждой интенсивности яркости (начиная с 1

и до 7) пиксели изображения постепенно объединяются в отдельные компоненты связности.

Предварительно все пиксели на изображении сортируются по возрастанию значений их яркостей и записываются в отдельный список, сохраняя информацию о своем месторасположении в исходном изображении. Далее последовательно анализируется каждый пиксель из этого списка. Если он соприкасается с уже существующей компонентой, то этот пиксель присоединяется к ней. Если он ни с кем не соприкасается, то создается новая компонента и сохраняется информация, на каком значении яркости она была сформирована. А если пиксель соприкасается сразу с несколькими компонентами, то происходит процесс их объединения. Компоненты, образованные на яркости с меньшим значением, поглощают другие компоненты, созданные на яркости с большим значением. На практике это сводится в основном к тому, что более крупные компоненты поглощают более мелкие. Если на одной яркости созданы одновременно две и более компонент, то приоритет при объединении получает та компонента, которая имеет меньший порядковый номер. Компонента, которая была поглощена другой компонентой, завершает свое существование с сохранением информации о значении яркости, на которой она прекратила существовать.

В данном примере получаем две компоненты, причем первая из них была образована на яркости 1 (включает точки со значением 1 и соответствует сплошной линии), вторая – на яркости 3 (состоит из точек со значением 3 и соответствует пунктирной линии), а далее первая компонента поглотила вторую на яркости 5.

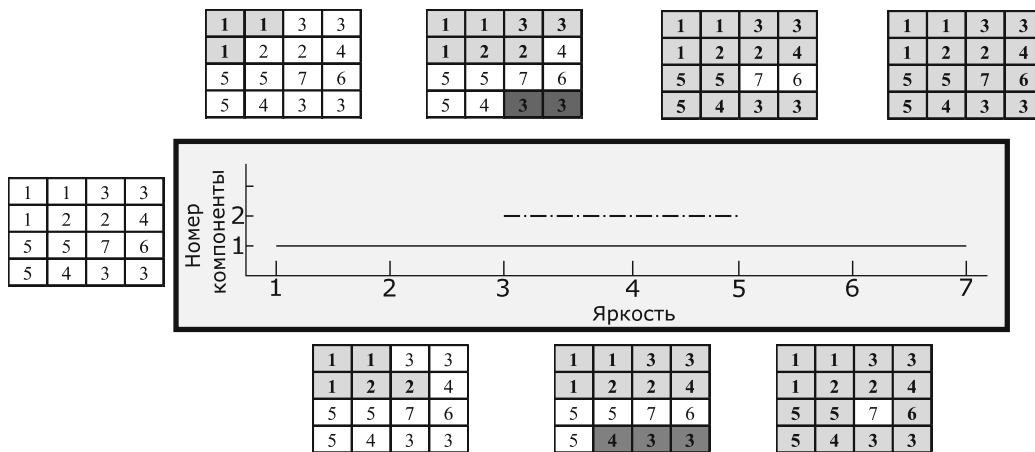


Рис. 1. Классический подход для построения баркода на основе персистентной гомологии

Характеристику в виде начала появления каждой компоненты и окончания ее существования принято называть баркодом, построение которого является классическим подходом для анализа данных в современной персистентной гомологии.

В работе предлагается для исходного изображения сначала получить последовательность бинарных изображений для каждого значения порога T (рис. 2).

Стрелки вниз между соответствующими элементами изображения на соседних уровнях порога показывают, какие именно единицы с предыдущего уровня переходят на текущий.

Так же как и в классическом подходе, образуются компоненты, которые создаются и завершают существование на определенной яркости. Однако при этом каждой такой компоненте будет соответствовать своя

матрица. Эти матрицы являются обобщением информации по всем бинарным изображениям на всех порогах T . Каждый элемент такой матрицы содержит число, равное количеству единиц в соответствующей ячейке каждого бинарного изображения, которое попадает в диапазон существования компоненты.

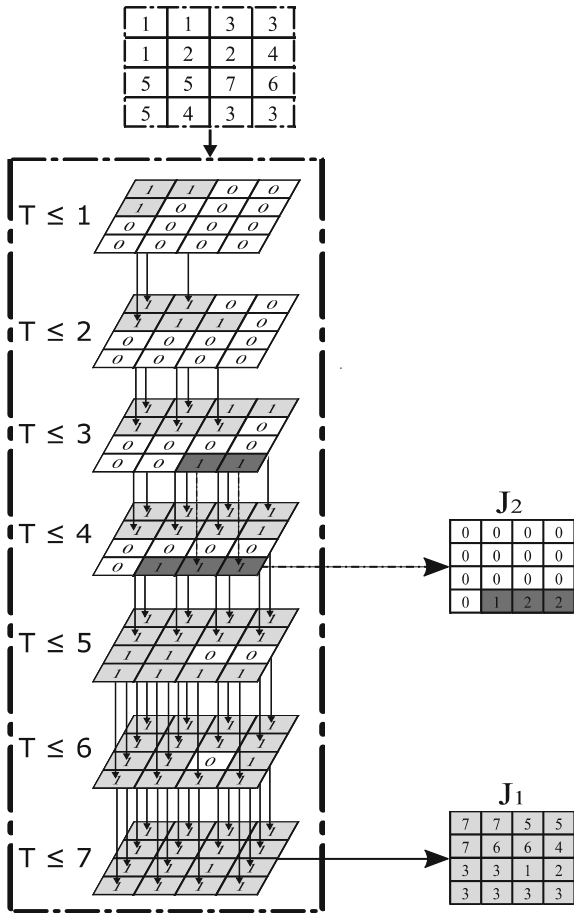


Рис. 2. Пример разложения изображения на основную J_1 и детализирующую J_2 матрицы с использованием предлагаемого подхода

Рассмотрим формально этот процесс. Пусть $I(x, y)$ – это исходное изображение размером $m \times n$, содержащее значения яркости в градациях серого от 0 до 255, а x, y – координаты точки изображения. После бинаризации изображения I для каждого значения порога $T=0, 1, 2, \dots, 255$ получим семейство бинарных изображений $V=\{I_0, I_1, \dots, I_{255}\}$, где I_k – это бинарное изображение с порогом k . Очевидно, что последовательность изображений из множества V имеет вложенную структуру, т.е. $I_0 \subseteq I_1 \subseteq \dots \subseteq I_{255}$.

В основе предлагаемого метода лежит формирование p матриц J_1, J_2, \dots, J_p . Рассмотрим сначала бинарное изображение I_0 . При его анализе создается одна или несколько новых компонент, которые на следующих шагах будут расширяться. При формировании только одной компоненты создается соответствующая ей матрица J_1 , которая изначально будет содержать единицы и нули. Ненулевые значения в

ней обозначают информацию об объекте интереса на исходном изображении. Например, для рис. 2 матрица J_1 инициализируется следующими значениями:

$$J_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

В случае появления нескольких компонент формируется последовательность матриц J_1, J_2, \dots, J_q ($q \leq p$), каждая из которых заполняется по аналогии нулями и единицами из бинарного изображения. Параметр q означает номер последней созданной матрицы на текущий момент. При создании очередной матрицы на следующих шагах ее номер будет увеличиваться на 1. Новая матрица создается в том случае, если появляется компонента, которая не пересекается ни с одной из предыдущих. Изображение обрабатывается построчно слева направо.

Далее анализируются данные на бинарном изображении I_1 с учетом матриц J_1, J_2, \dots, J_q и изображения I_0 на предыдущем шаге. Существующая компонента может расширяться за счет присоединения к ней соприкасающихся новых единиц из I_1 . Тогда к соответствующей матрице прибавляются новые значения из единиц, а каждый прежний ненулевой элемент также увеличивается на 1. Для рис. 2 на шаге $T \leq 3$ будем иметь следующие текущие значения матриц:

$$J_1 = \begin{bmatrix} 3 & 3 & 1 & 1 \\ 3 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad J_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Кроме того, компоненты на некотором шаге могут объединяться. Тогда происходит поглощение в одну из существующих матриц J_1, J_2, \dots, J_q . Основным критерием для определения поглощающей матрицы является меньшая яркость, на которой она появилась. Если на одной яркости появляются сразу несколько матриц, то приоритет при поглощении получает матрица с наименьшим порядковым номером. В предложенном примере матрица J_1 поглощает J_2 на яркости 5, тем самым, вторая компонента завершает свое существование.

Такой процесс анализа бинарных изображений выполняется для всех I_k ($k=0, 1, 2, \dots, 255$) и позволяет постепенно заполнить итоговые значения набора матриц J_1, J_2, \dots, J_p .

Такие матрицы будем называть спектр-матрицами разложения по топологическим признакам, которые соответствуют определенным структурам на исходном изображении.

Определение 1. Под спектр-матрицей разложения по топологическим признакам будем понимать матрицу J_k ($k \in \{1, 2, 3, \dots\}$), суммирующую только те зна-

чения бинарных изображений I_0, I_1, \dots, I_{255} , которые относятся к определенной компоненте с номером k , причем суммирование происходит в пределах порога $l_k \leq T \leq r_k$, где l_k и r_k – яркости, на которых соответственно появляется и завершает свое существование k -я компонента.

Так, для примера из рис. 2 получаем две итоговые матрицы J_1 и J_2 . Матрица J_1 является основной и содержит глобальную информацию исходной матрицы, а детализирующая матрица J_2 показывает детали. Причем яркость, на которой появляется и прекращает существование каждая из них, равны соответственно: $l_1 = 1, r_1 = 7; l_2 = 3, r_2 = 4$. Более того, вся информация по баркодам, которая есть в классическом подходе (на рис. 1), может быть получена из матриц J_1 и J_2 , но при этом матрицы J_1 и J_2 содержат больше информации для анализа.

Таким образом, любое изображение I в градациях серого можно разложить на сумму из p спектр-матриц разложения J_k следующим образом:

$$I = 256 \times \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix} - \sum_{k=1}^p J_k, \quad (1)$$

где \times – операция умножения числа на матрицу.

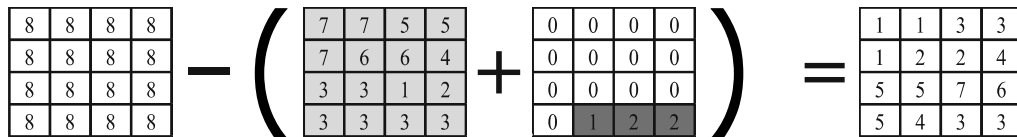


Рис. 3. Пример обратного процесса восстановления по основной и детализирующей матрицам (значение 8 – это максимальное значение, которое было в исходном изображении, плюс один; для обычных 8-битных цифровых изображений в градациях серого необходимо использовать значение 256)

Используя эти свойства, введем понятия низкочастотных и высокочастотных спектр-матриц разложения по топологическим признакам.

Определение 2. Низкочастотной спектр-матрицей разложения по топологическим признакам будем называть сложение матриц с продолжительным периодом существования, т.е.

$$J_n = \sum_{k=1}^p J_k, \text{ если } r_k - l_k \geq c,$$

где r_k, l_k – яркости, на которых соответственно появляется и заканчивает существование k -я компонента; c – некоторый порог, например, среднее значение продолжительностей для всех матриц J_k .

Определение 3. Высокочастотной спектр-матрицей разложения по топологическим признакам J_e будем называть сложение матриц с коротким периодом существования, т.е.

$$J_e = \sum_{k=1}^p J_k, \text{ если } r_k - l_k < c.$$

Заметим, что каждая из спектр-матриц разложения $J_1, J_2, \dots, J_k, \dots, J_p$ имеет и притом единственную связную область.

Ключевой особенностью является возможность обработки матриц J_k стандартными алгебраическими методами, т.е. складывать, обнулять отдельные матрицы, изменять значения относительно конкретных компонент, а потом получить новую матрицу, сделав обратный процесс восстановления. Фактически, мы приходим к прямому и обратному преобразованию изображений по топологическим особенностям.

Пример процесса восстановления исходного изображения показан на рис. 3. Как можно увидеть, если не изменять значения в матрицах J_1 , и J_2 , то получим исходное изображение без изменений.

Обнаружено, что матрицы обладают определенными свойствами. Покажем некоторые из них:

Свойство 1. Изменение значений любой из матриц J_k оказывает влияние на новое изображение после восстановления.

Свойство 2. Матрицы J_k можно складывать, что позволяет объединить сразу несколько спектр-матриц разложения.

Свойство 3. Спектр-матрицы J_k , которые имеют продолжительный период существования, содержат больше информации, чем матрицы с меньшим периодом существования.

Это приводит к разложению изображения на низкочастотные и высокочастотные спектр-матрицы. Низкочастотные матрицы содержат основные топологические структуры на изображении и наиболее устойчивы к деформациям, а высокочастотные включают детали изображения и шум.

На основе этого получаем, что любое изображение I в градациях серого можно разложить на низкочастотные J_n и высокочастотные J_e спектр-матрицы, т.е.

$$I = 256 \times \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix} - (J_n + J_e). \quad (2)$$

Пример низкочастотной и высокочастотной спектр-матриц для изображений в градациях серого приведен на рис. 4.

Подобного эффекта можно достичь через уже упомянутое Wavelet-преобразование, но это потребует долгого подбора нужных коэффициентов. Есть

определенные сходства, но также есть принципиальные отличия. В предлагаемом подходе мы можем регулировать итоговое количество матриц разложения для последующей обработки. Например, сделать две, три и более матриц разложения с уменьшением их

степени детализации, т.е. последняя матрица будет содержать информацию о самых мелких деталях или шумовой составляющей. Т.е. можно выделить не только низкие и высокие частоты, но и среднечастотные J_c спектр-матрицы разложения.



Рис. 4. Пример разложения изображения «Лодки» на две матрицы: исходное изображение (а), низкочастотная спектр-матрица разложения (б), высокочастотная спектр-матрица разложения (в)

Для повышения или понижения значений матриц применим следующий подход:

$$I = 256 \times \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix} - (\alpha_1 \times J_n + \alpha_2 \times J_c + \alpha_3 \times J_g), \tag{3}$$

где $\alpha_1, \alpha_2, \alpha_3$ – числовые коэффициенты, воздействующие на значения матриц J_n, J_c, J_g и в итоге на результирующее изображение.

Кроме того, можно воздействовать на отдельные элементы любой из матриц J_n, J_c, J_g . Обработав некоторые матрицы, мы получим новое изображение. Например, при обнулении высокочастотных матриц при $\alpha_3 = 0$ получим изображение без шума или, наоборот, можем выделить только мелкие объекты на изображении.

2. Практическое использование

Рассмотрим применение предложенного метода для решения конкретных задач, а именно, для бинаризации изображений и удаления текста на изображениях.

Для анализа каждой матрицы J_1, J_2, \dots, J_p введем функцию $f(M)$, которая возвращает одно из двух возможных значений: 0 или 1. Если матрица M удовлетворяет условиям отбора, то $f(M) = 1$, иначе $f(M) = 0$. Примером условия отбора может являться заданное количество ненулевых элементов в матрице; яркость, на которой была создана матрица и т.д.

2.1. Использование предложенного разложения для решения задачи бинаризации изображений

Алгоритм для бинаризации будет состоять из следующих шагов:

Шаг 1. На вход подается исходное изображение $I(x, y)$.

Шаг 2. Получаем разложение изображения $I(x, y)$ на множество матриц J_1, J_2, \dots, J_p .

Шаг 3. В цикле перебираем все матрицы и отмечаем только те, которые соответствуют отборочной функции $f(M)$.

Шаг 4. Суммируем только отмеченные матрицы J_k и записываем результат в аккумулирующую матрицу A , т.е.:

$$A = \sum_{k=1}^p J_k, \text{ если } f(J_k) = 1.$$

Шаг 5. Применяем пороговую бинаризацию для аккумулирующей матрицы A следующим образом:

$$BI(x, y) = \begin{cases} 255, & \text{если } A(x, y) \neq 0 \\ 0, & \text{иначе} \end{cases},$$

где $BI(x, y)$ – бинаризованное изображение.

Шаг 6. Выводим результат бинаризации.

Приведем теперь результат работы рассмотренного алгоритма бинаризации в сравнении с некоторыми аналогами. В качестве аналогов были взяты следующие: метод Оцу [16], K-Means [17] и адаптивная бинаризация [18]. На рис. 5 и 6 показаны результаты бинаризации алгоритмов.

Для получения результатов на рис. 5 и 6 у метода Оцу и K-Means были использованы определенные параметры. Для K-Means: количество попыток – 100; заданная точность – 85%; критерий завершения – по достижении заданной точности либо по достижении максимального числа итераций; способ выбора начальных центров – случайно. Для адаптивной бинаризации были выбраны такие параметры: тип бинаризации – «бинаризованное значение минус константа», константа – 10. Параметры предложенного алгоритма показаны в подписях к рис. 5 и 6.

Визуально результаты предложенного алгоритма наиболее схожи с адаптивной бинаризацией. При этом можно заметить, что предложенный алгоритм наиболее четко выделяет буквы на изображении «Лодки», а на изображении «Иван» не происходит явного слияния объектов, как в методе Оцу и K-Means.



Рис. 5. Результаты бинаризации на изображении «Лодки»: исходное изображение (а), метод Оцу (б), K-Means (в), адаптивная бинаризация (г), предложенный алгоритм (д), в котором функция $f(M)$ – это разница между минимальным и максимальным значениями на матрице M в пределах от 15 до 142 и количество ненулевых точек на каждой матрице разложения больше 50



Рис. 6. Результаты бинаризации на изображении «Иван»: исходное изображение (а), метод Оцу (б), K-Means (в), адаптивная бинаризация (г), предложенный алгоритм (д), в котором функция $f(M)$ – это разница между минимальным и максимальным значениями на матрице M в пределах от 30 до 150 и количество ненулевых точек на каждой матрице разложения больше 30

Чтобы дать численную оценку качества бинаризации, выделим основные критерии. Бинаризацию будем считать успешной, если:

- 1) были выделены все наиболее яркие объекты;
- 2) выделенные объекты чётко различимы и не сливаются в один;
- 3) на изображении присутствует минимальный объём шума.

Из этого следует, что в идеальном случае все объекты должны отделиться от фона. При подобных условиях получается, что наибольшая яркость зон на исходном изображении должна быть покрыта бинарным изображением. Условно этот параметр можно назвать яркостным отношением суммы яркостей тех точек изображения, для которых $BI(x, y) = 255$, к общей сумме всех яркостей исходного изображения, т.е.:

$$\delta = \frac{\sum_{x=1}^m \sum_{y=1}^n I(x, y), \text{ если } BI(x, y) = 255}{\sum_{x=1}^m \sum_{y=1}^n I(x, y)},$$

где $BI(x, y)$ – бинаризованное изображение.

Таким образом, оценка δ показывает, насколько много было покрыто ярких пикселей, что соответствует критерию 1. Чтобы проверить критерии 2 и 3, будем подсчитывать количество отдельных объектов на бинаризованном изображении BI .

Эти две характеристики довольно слабы сами по себе, но хорошо подходят при сравнении разных подходов, что в данном случае и происходит.

Численные параметры работы подходов сведены в табл. 1.

Корреляция оценки δ и количества объектов в совокупности с анализом объектов на исходном изображении говорит о том, что в сравнении с аналогами разработанный алгоритм покажет себя наилучшим образом в ситуациях, когда бинаризация применяется для выделения множества отдельных объектов.

2.2. Использование предложенного разложения для решения задачи удаления текста на изображениях уличных и автомобильных номеров

В заключении приведем еще одно назначение, имеющее меньше известных аналогов. Речь идет об удалении текста с растровых изображений.

Табл. 1. Численные параметры работы подходов для бинаризации изображений

| Изображение | Подход | δ | Количество объектов |
|-------------|-----------------------|----------|---------------------|
| «Лодки» | Оцу | 0,91 | 150 |
| | К-Means | 0,86 | 339 |
| | Адаптивный | 0,89 | 190 |
| | Предложенный алгоритм | 0,88 | 216 |
| «Иван» | Оцу | 0,79 | 1319 |
| | К-Means | 0,82 | 1457 |
| | Адаптивный | 0,89 | 2949 |
| | Предложенный алгоритм | 0,90 | 2308 |

Обычно на общедоступных фотографиях его предпочитают размывать, но не исключено, что уже с текущим уровнем технологий вскоре размытый текст удастся полностью восстановить, тем более что нейронная сеть DeepDeblure уже сильно преуспела в этом [19].

Эта проблема особенно актуальна для автомобильных номеров, учитывая ограниченный набор знаков на каждой позиции номера и известный формат.

В данном случае предложенный подход может обеспечить корректное удаление исключительно текста. Опять же, всё сводится к определению матриц из набора J_1, J_2, \dots, J_p через отборочную функцию $f(M)$ с последующим их обнулением. После этого необходимо выполнить обратное преобразование по формуле (1) и синтезировать новое изображение.

В данной задаче важно, чтобы исходный текст невозможно было восстановить, поэтому может потребоваться несколько проходов.

На рис. 7 и 8 показаны результаты удаления текста на неоднородном фоне. Если одного прохода недостаточно, то к полученному изображению применяется алгоритм повторно. Это нужно в случаях, когда тёмный текст расположен на светлом фоне.

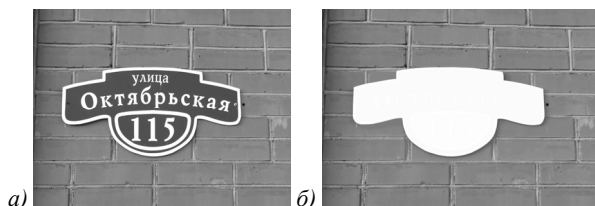


Рис. 7. Изображение «Номер улицы»: до обработки (а) и после обработки (б)

Результаты показывают, что информация полностью удаляется, не задевая остальные области.

Заключение

В работе предложен новый метод для разложения исходного изображения на отдельные компоненты. Частным случаем метода является разложение изображения на низкочастотные и высокочастотные матрицы, что позволяет выделить на изображении ос-

новную информацию и детали. Показано, что можно управлять процессом выделения объектов, что дает возможность разделять изображение на компоненты с учетом точности детализации. В основе метода лежит использование персистентной гомологии для обработки изображений. Разработанный метод имеет прямое и обратное преобразование. Обратное преобразование позволяет использовать полученные матрицы разложения для получения нового преобразованного изображения. На основе метода разложения созданы алгоритмы для бинаризации изображений и удаления текста на неоднородном фоне, которые показали качественные результаты. Также данный метод в перспективе можно использовать для решения других задач, включая сегментацию изображений, анализ текстурных признаков и другие.

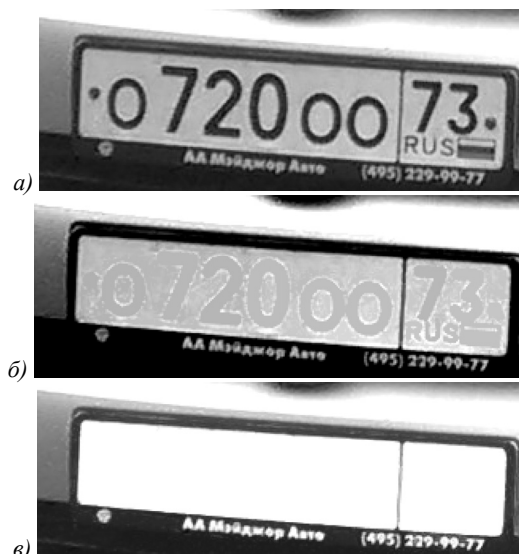


Рис. 8. Изображение «Автомобильный номер»: исходное изображение (а), после первого прохода (б), после второго прохода (в)

Благодарности

Исследование выполнено в рамках Программы развития ЯргУ, проект № П2-ГМ3-2021.

References

- [1] Edelsbrunner H, Letscher H, Zomorodian A. Topological persistence and simplification. *Discrete Comput Geom* 2002; 28: 511-533. DOI: 10.1007/s00454-002-2885-2.
- [2] Kurlin V, Muszynski G. Persistence-based resolution-independent meshes of superpixels. *Pattern Recognition Letters* 2020; 131: 300-306.
- [3] Bubenik P, Dlotko P. A persistence landscapes toolbox for topological statistics. *J Symb Comput* 2017; 78: 91-114.
- [4] Adams H, Emerson T, Kirby M, Neville R, Peterson C, Shipman P, Chepushtanova D, Hanson E, Motta F, Ziegelmeier L. Persistence images: A stable vector representation of persistent homology. *J Mach Learn Res* 2017; 18(8): 1-35.
- [5] Karimova L, Terekhov A, Makarenko N, Rybintsev A. Methods of computational topology and discrete Riemannian geometry for the analysis of arid territories. *Cogent Eng* 2020; 7: 1808340. DOI: 10.1080/23311916.2020.1808340.

- [6] Eremeev SV, Abakumov AV. Software complex for detection and classification of natural objects based on topological analysis [In Russian]. *Software & Systems* 2021; 34(1): 201-208. DOI: 10.15827/0236-235X.133.201-208.
- [7] 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.
- [8] Chukanov SN. Comparison of objects' images based on computational topology methods [In Russian]. *SPIIRAS Proceedings* 2019; 18(5): 1043-1065. DOI: 10.15622/sp.2019.18.5.1043-1065.
- [9] Kurochkin SV. Detection of the homotopy type of an object using differential invariants of an approximating map. *Computer Optics* 2019; 43(4): 611-617. DOI: 10.18287/2412-6179-2019-43-4-611-617.
- [10] Vimala C, Aruna Priya P, Subramani C. Wavelet transform approach for image processing – A research motivation for engineering graduates. *Int J Electr Eng Educ* 2021; 58(2): 373-384. DOI: 10.1177/0020720919825815.
- [11] Tharwat A. Independent component analysis: An introduction. *Appl Comput Inform* 2018; 14: 1-15. DOI: 10.1016/j.aci.2018.08.006.
- [12] Huang X. An improved FastICA algorithm for blind signal separation and its application. *Int Conf on Image Analysis and Signal Processing* 2012: 1-4. DOI: 10.1109/IASP.2012.6425039.
- [13] Volkov VY. Adaptive multi-threshold object selection in remote sensing images. [In Russian]. *Information and Control Systems* 2020; 3: 12-24. DOI: 10.31799/1684-8853-2020-3-12-24.
- [14] Egorova AA, Sergeev VV. Extended set of superpixel features. *Computer Optics* 2021; 45(4): 562-574. DOI: 10.18287/2412-6179-CO-876.
- [15] Vizilter YuV, Vygolov OV, Zheltov SYu. Morphological analysis of mosaic shapes with directed relationships based on attribute and relational model representations. *Computer Optics* 2021; 45(5): 756-766. DOI: 10.18287/2412-6179-CO-843.
- [16] Sekertekin A. Potential of global thresholding methods for the identification of surface water resources using Sentinel-2 satellite imagery and normalized difference water index. *J Appl Remote Sens* 2019; 13(4): 044507. DOI: 10.1117/1.JRS.13.044507.
- [17] Yuan C, G Yang. Research on k-value selection method of k-means clustering algorithm. *J* 2019; 2(2): 226-235. DOI: 10.3390/j2020016.
- [18] Roy P, Dutta S, Dey N, Dey G, Chakraborty S, Ray R. Adaptive thresholding: A comparative study. *Int Conf on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)* 2014: 1182-1186. DOI: 10.1109/ICCICCT.2014.6993140.
- [19] Mei J, Wu Z, Chen X, Qiao Y, Ding H, Jiang X. DeepDe-blur: text image recovery from blur to sharp. *Multimed Tools Appl* 2019; 78(13): 18869-18885. DOI: 10.1007/s11042-019-7251-y.

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@vandex.ru.

Абакумов Артём Владимирович, 1998 года рождения, магистрант Муромского института (филиала) Владимирского государственного университета. Область научных интересов: обработка пространственных данных, геоинформационные системы, топологический анализ данных. E-mail: artem210966@vandex.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.

Титов Дмитрий Витальевич, 1986 года рождения, д.т.н., доцент, проректор по науке и международной деятельности Юго-Западного государственного университета. Область научных интересов: обработка изображений, компьютерное зрение. E-mail: umsswsu@gmail.com.

ГРНТИ: 28.23.15

Поступила в редакцию 6 декабря 2021 г. Окончательный вариант – 30 июня 2022 г.

Image decomposition method by topological features

S.V. Ereemeev¹, A.V. Abakumov¹, D.E. Andrianov¹, D.V. Titov²

¹ Murom Institute (branch), Vladimir State University named after Alexander and Nikolay Stoletovs, 602264, Murom, Russia, Orlovskaya 23,

² Southwest State University, 305040, Kursk, Russia, 50 Let Oktyabrya 94

Abstract

A new method for decomposing an image into separate objects of interest is proposed in the article. The developed method is based on the use of persistent homology. A process of direct and reverse image transformation is shown. Following direct transformation, the original image is represented as a set of matrices that can be divided into basic and detailing ones. The basic matrices contain information about the basic structure of objects in the images, and the detailing ones include data about the details of these objects, about small objects or the noise component. It is shown that there is a certain analogy with the Wavelet transformation, but the proposed method is based on a fundamentally different theoretical basis. A numerical example reflecting the basic essence of the method is described in detail. Properties of the decomposition and the possibility of using standard algebraic operations on decomposition matrices are described. The reverse transformation allows us to take into account the changed properties of individual objects and synthesize a new image. Prospects of using the proposed decomposition for solving practical problems are demonstrated. Algorithms have been developed for binarization of images and removal of text on a non-uniform background. Data analysis and processing is carried out using a unified approach in the space of decomposition matrices. The results of binarization have shown that, when compared with analogues, the developed algorithm will perform better when the binarization is used to isolate a multitude of individual objects. The obtained results of the algorithm for deleting text on a non-uniform background confirm that the information is completely deleted without affecting the rest image areas.

Keywords: topological data analysis, persistent homology, barcode, topological features, connectivity components, image decomposition, low-frequency and high-frequency decomposition matrices.

Citation: Ereemeev SV, Abakumov AV, Andrianov DE, Titov DV. Image decomposition method by topological features. *Computer Optics* 2022; 46(6): 939-947. DOI: 10.18287/2412-6179-CO-1080.

Acknowledgements: The reported study was funded by the YSU Programme under research project No. P2-GM3-2021.

Authors' information

Sergey Vladimirovich Ereemeev (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.

Artyom Vladimirovich Abakumov (b. 1998), Master of Science, Murom Institute (branch) of Vladimir State University. Research interests are spatial data processing, geographic information systems and topological data analysis. E-mail: artem210966@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.

Dmitry Vitalievich Titov (b. 1986), Doctor of Engineering, Associate Professor, vice-rector for Science and International Affairs of SWSU. Research interests are image processing, computer vision. E-mail: umsswsu@gmail.com.

Received December 06, 2021. The final version – June 30, 2022.
