

О сходстве задач сглаживания и сегментации изображений

П. А. Чочиа

Институт проблем передачи информации им. А. А. Харкевича РАН, Москва, Россия
Поступила в редколлегию 10.09.2024 г. Принята 1.10.2024 г.

Аннотация—Одними из важнейших направлений обработки и анализа видеoinформации являются задачи сглаживания и сегментации изображений, которые в своей постановке, конечной цели и способе решения имеют много общего. На основании модели изображения, путем анализа и сравнения известных алгоритмов сглаживания и сегментации, показывается сходство и различие указанных задач.

КЛЮЧЕВЫЕ СЛОВА: обработка изображений, сглаживание, сегментация.

DOI: 10.53921/18195822_2024_24_3_175

1. ВВЕДЕНИЕ

Исследования вопросов сглаживания и сегментации изображений приводят к мысли, что эти две достаточно разные проблемы на самом деле весьма близки. Прежде всего, сформулируем те цели и задачи, которые перед каждой из них ставятся.

Сглаживание является одним из самых распространенных способов обработки сигнала. В задачах анализа изображений сглаживание используется для уменьшения шума и других нежелательных артефактов на изображении при сохранении его важных структур и особенностей. Формулировки, приводимые в различных публикациях [1–4] и других, сводятся к тому, что сглаживание изображения является ключевой технологией улучшения сигнала, позволяющей подавить помехи и шум. Также сглаживание используется на этапе предварительной обработки сигнала для удаления мелких деталей и оставления крупных объектов.

Цель сглаживания можно сформулировать как получение изображения, удобного для дальнейшей интерпретации и анализа. При этом важен выбор степени сглаживания согласно требованиям задачи, чтобы избежать потери важной информации. В алгоритмах сглаживания важен выбор множества точек, по которому следует определять значение целевого элемента, для чего требуется решить задачу принадлежности элементов окружения той же области, что и целевой элемент. При этом возникает необходимость сепарации точек, относящихся к области анализа, что перекликается с задачей сегментации изображения.

Сегментация означает разбиение пространства изображения на составляющие его области, соответствующие объектам сцены [5, 6]. Обычно сегментация используется для обнаружения и выделения объектов и их границ. При этом степень детализации, до которой доводится разделение, зависит от решаемой задачи. Сегментация изображений, не являющихся тривиальными, представляет собой одну из самых сложных задач обработки изображений [1].

Цель сегментации — упростить и/или изменить представление изображения на нечто более простое для анализа [7]. В процессе сегментации каждому элементу изображения приписывается индекс обнаруженной области; точнее, каждому пикселю изображения присваивается метка таким образом, чтобы пиксели с одинаковой меткой имели определенные свойства. Результатом является набор сегментов, которые вместе покрывают изображение целиком. В цифровой

обработке изображений и компьютерном зрении сегментация — это процесс разделения изображения на множество сегментов, также называемых *области* или *объекты* изображения.

Нетрудно видеть, что за исключением идентификации, задачи сглаживания и сегментации изображения достаточно близки. Тьюки в [8] рекомендует в задачах анализа “почувствовать особенности данных”. В нашем случае к этому стоит добавить, что требуется также почувствовать *особенности алгоритмов*. Важно, что и в сглаживании, и в сегментации, так или иначе, решаются задачи принадлежности точек интересующим множествам, а также осуществляются объединения или разделения элементов по тем или иным признакам.

Если в сегментации вместо идентификатора присваивать элементу усредненное значение по сегменту, которому он принадлежит, то получим вариант сигнала, который можно визуальнo сравнить с результатом сглаживания. Это позволяет поднять вопрос о близости задач и результатов сглаживания и сегментации изображений.

2. МОДЕЛЬ ИЗОБРАЖЕНИЯ

Для сравнения разных подходов к обработке изображений требуется модель, описывающая свойства сигналов на пространственном протяжении не меньшем, чем размер области анализа. В качестве такой выберем двухмасштабную многокомпонентную модель изображения [9–11], суть которой в следующем. Изображение, как двумерный массив значений $x_{mn} \in [0, x_{\max}]$, $0 \leq m < M$, $0 \leq n < N$, является отображением некоторой реальной сцены. Объектам сцены соответствуют протяженные области изображения с мало меняющимися значениями яркости, разделенные контурными перепадами — участками с высоким значением градиента. Каждая из областей приближается некоторой двумерной поверхностью [12] и может содержать мелкие детали, текстуру и шум. Изображение x_{mn} при этом представимо в виде суммы:

$$x_{mn} = s_{mn} + t_{mn} + \xi_{mn}. \quad (1)$$

где s_{mn} — кусочно-гладкая компонента, задающая приближение яркости в локальной области анализа, t_{mn} — детальная, а ξ_{mn} — шумовая компоненты [9, 13]. Компоненты считаются независимыми, а шум — нормально распределенным. Согласно [14], значительную долю изображения занимают протяженные области, содержащие от 10^2 до 10^4 и более элементов каждая. С каждым элементом x_{mn} связываются окрестность V_{mn} , как множество ближайших элементов из той же области или контура, что и x_{mn} , а также фрагмент W_{mn} — протяженное множество элементов, которое может покрывать несколько областей. Элементы одной области зависимы не только на малых расстояниях (внутри окрестности), но и на больших — по крайней мере, в пределах фрагмента. В то же время даже близко расположенные элементы из разных областей считаются независимыми.

Модель окрестности, содержащей $x_{mn}^v \in V_{mn}$; $v = 1, \dots, V$ элементов, состоит в том, что значения ее пикселей приближаются плоскостью с углом наклона θ_{mn} , существенно различающимся для внутренних точек областей и контурных точек на их границах. Значения элементов окрестности x_{mn}^v при этом связываются формулой:

$$x_{mn}^v = \mu_{mn} + \rho^v g_{mn}^v + \xi_{mn}^v, \quad (2)$$

где μ_{mn} — значение проведенной плоскости в центральной точке окрестности (m, n) , ρ^v — расстояние между элементами x_{mn} и x_{mn}^v , $g_{mn} = \tan \theta_{mn}$, а ξ_{mn}^v — шумовая составляющая. Все элементы любой окрестности $x_{mn}^v \in V_{mn}$ расположены либо в одной области, либо на контурном перепаде между областями.

Фрагмент W_{mn} имеет размеры значительно большие, чем окрестность и может покрывать R областей: U^1, \dots, U^R . Через x_{mn}^r обозначим элемент области U^r с координатами (i, j) внутри

фрагмента размерами $(2L + 1) \times (2L + 1)$. При этом для анализа доступны лишь элементы и части областей, находящиеся внутри границ фрагмента: $(-L \leq i, j \leq L)$. Основываясь на (1) полагается, что гладкие составляющие тех частей R областей, которые покрываются фрагментом, представимы полиномом степени не выше ω . Тогда составляющая s_{ij} внутри фрагмента W_{mn} описывается формулой

$$s_{ij}^r(W_{mn}) = \sum_{U=1}^R \delta_{U^r} \sum_{p=0}^{\omega} \sum_{q=0}^p a_{pq}^r i^{p-q} j^q, \quad (3)$$

где (i, j) — точка фрагмента W_{mn} ; $\delta_{U^r} = 1$, если $(i, j) \in U^r$ и $\delta_{U^r} = 0$ иначе. Добавляя текстурную t_{mn} и шумовую ξ_{mn} составляющие, получим значение элемента:

$$x_{ij}^r = \sum_{U=1}^R \delta_{U^r} \sum_{p=0}^{\omega} \sum_{q=0}^p a_{pq}^r i^{p-q} j^q + t_{ij}^r + \xi_{ij}. \quad (4)$$

Это основная формула модели фрагмента, связывающая его внутренние элементы.

Как показали эксперименты [9], у реальных изображений значения x_{mn}^r на протяжении фрагмента типичного размера меняются мало, и во многих случаях можно ограничиться кусочно-постоянной моделью с $\omega = 0$. Тогда (4) преобразуется к виду

$$x_{ij}^r = \sum_{U=1}^R \delta_{U^r} s_{mn}^r + t_{ij}^r + \xi_{ij}.$$

Считая t и ξ независимыми и нормально распределенными, их сумма $\tau = t + \xi$ также распределена нормально, и предыдущее выражение упрощается до

$$x_{ij}^r = \sum_{U=1}^R \delta_{U^r} s_{mn}^r + \tau_{ij}^r. \quad (5)$$

Модели (4) и (5) являются разрывными. Последняя из них (5) достаточно проста, близка большинству реальных изображений и удобна для построения на ее основе различных алгоритмов.

3. СГЛАЖИВАНИЕ ИЗОБРАЖЕНИЯ

Сглаживание — одна из важнейших операций в статистике и обработке данных, цель которой — подавление шума, нежелательных артефактов и других несущественных мелких изменений. Обычно это осуществляется отысканием некоторой аппроксимирующей поверхности или функции, прослеживающей основные закономерности изменения данных и сохраняющей важные детали сигнала. Сглаживание может помочь в последующем анализе благодаря возможности извлечь больше информации из данных, которые в дальнейшем легче анализировать, интерпретировать и использовать в приложениях обработки и распознавания [2]. При этом сглаживание следует отличать от сравнительно близкой по концепции задачи аппроксимации кривой [15] или поверхности по точкам, но имеющей принципиально другие цели.

Сглаживание, как операция фильтрации, может производиться как в частотной, так и в пространственной области [1–3, 16]. Частотный анализ во многом предпочтителен при исследовании аналоговых методов; в случае же цифровых методов удобнее рассматривать пространственную область, преобразования в которой представляют для нас наибольший интерес. В сглаживании изображений применяют различные алгоритмы, как правило, базирующиеся на свертках, основными из которых являются фильтры арифметического среднего, медианный, Гауссов [1, 12], сигма-фильтр [17], декомпозиции [18], билатеральный [19].

Алгоритмы сглаживания можно условно разделить на два класса: алгоритмы, основанные на одномерных распределениях и применимые как к одномерным, так и к двумерным сигналам, и существенно двумерно-многомерные алгоритмы, не сводящиеся к одномерному анализу. Как показано ниже, информация о контурных перепадах весьма полезна при решении задачи сглаживания, сохраняющего структуру изображения.

Фильтр арифметического среднего значений элементов по прямоугольному окну W , который аналогичен применению фильтра нижних частот [12], является классической операцией сглаживания в обработке изображений:

$$y_{mn} = \sum_{i=m-k}^{m+k} \sum_{j=n-l}^{n+l} x_{ij} / S_W, \quad \text{где } S_W = (2k+1)(2l+1). \quad (6)$$

Для инвариантного к сдвигу преобразования (6) существуют простые и быстрые алгоритмы вычисления независимо от размера окна W .

Медианный фильтр состоит в замене значения в точке (m, n) изображения на медиану распределения значений элементов в окружающем его окне W_{mn} :

$$y_{mn} = \text{med}\{x_{ij} \in W_{mn}\}. \quad (7)$$

Фильтр (7) является нелинейной и эффективной процедурой сглаживания сигналов, подверженных воздействию импульсных шумов [1]. Во многих случаях результаты медианного фильтра предпочтительны по сравнению с фильтром арифметического среднего, но реализующие его алгоритмы существенно более медленные.

Гауссов фильтр в применении к изображению аналогичен его свёртке с функцией нормального распределения ψ_{ij} с дисперсией σ^2 :

$$y_{mn} = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \psi_{ij} x_{m-i, n-j}, \quad \text{где } \psi_{ij} = \exp\{-(i^2 + j^2)/2\sigma^2\}/2\pi\sigma. \quad (8)$$

В реальных задачах достаточно осуществлять суммирование по формуле (8) в диапазоне $\pm 3\sigma$ с небольшой коррекцией значения ψ . По сравнению с фильтрами (6) и (7), результат Гауссова сглаживания при аналогичных параметрах выглядит более “мягким” и размытым.

Сигма-фильтр Ли [17] является частным случаем парзеновского окна [20]:

$$f(u) = \sum_{k=0}^{K-1} F((x_k - u)/h) / Kh,$$

где $F(x)$ — ядро анализа, h — параметр ширины ядра. На этой идее основаны многие алгоритмы классификации, в частности, метод ближайших соседей. Вид функции F мало влияет на точность классификации, а определяющим является параметр h [21]. Такой подход применим для оценки среднего по множеству элементов $x_k \in W_{mn}$:

$$y_{mn} = \sum_{k=0}^{K-1} x_k F(x_k - x_0) / \sum_{k=0}^{K-1} F(x_k - x_0). \quad (9)$$

Здесь $F(x)$ — весовая функция, x_0 — опорное значение для точки (m, n) . Если $x_0 = x_{mn}$, а $F_i(x) = 1$ для $-\delta \leq x \leq \delta$, и $F_i(x) = 0$ для остальных x , то получим сигма-фильтр Ли [17]. По значениям $x_k \in W_{mn}$ ($k = 1, \dots, K$), попадающим в интервал $(x_{mn} - \delta, x_{mn} + \delta)$, находится среднее, являющееся результатом:

$$y_{mn} = \sum_{k=1}^K x_k / K. \quad x_{mn} - \delta \leq x_k \leq x_{mn} + \delta. \quad (10)$$

Существенное отличие сигма-фильтра (10) от фильтров (6)–(8) состоит в том, что он не размывает границы объектов изображения. Также следует отметить основанный на Парзеновском сглаживании алгоритм сдвига среднего значения [22, 23], использующий итеративную схему вычисления.

Алгоритм декомпозиции [18] основан на модели (5), согласно которой сглаживание есть нахождение компоненты s_{mn} . Он использует два окна анализа: малое V_{mn} — окрестность и большое W_{mn} — фрагмент. Сначала по элементам окрестности V_{mn} определяется усеченное значение x_1 , подставляемое в качестве x_{mn} в формулу (10) и вычисляется сглаженное значение y_1 по окрестности. Затем анализ повторяется по элементам фрагментам W_{mn} , где по y_1 аналогично определяется значение x_2 , подставляемое также в (10), и вычисляется окончательное сглаженное значение y_{mn} :

$$y_{mn} = f(x_{mn}, h(V_{mn}), h(W_{mn}), \delta_V, \delta_W), \quad (11)$$

где $h(V_{mn})$ и $h(W_{mn})$ — распределения значений элементов в окружающих окрестности и фрагменте. Полученный фильтр (11) качественно улучшает сигма-фильтр (10) и позволяет управлять разделением объектов по площади с отнесением их в сглаженную (s) или в разностную (τ) компоненту модели (5).

Билатеральный фильтр [19] описывается выражением:

$$y_{mn} = \sum_{i,j \in W_{mn}} x_{ij} \psi((i, j), (m, n,)) \varphi(x_{ij} - x_{mn}) / \sum_{i,j \in W_{mn}} \psi((i, j), (m, n,)) \varphi(x_{ij} - x_{mn}). \quad (12)$$

Здесь $\psi((i, j), (m, n,))$ — Гауссова функция по расстоянию в пространстве между точками (i, j) и (m, n) , а $\varphi(x)$ — Гауссова функция по разности значений интенсивности в этих точках. Как нетрудно видеть, билатеральный фильтр (12) отличается от сигма-фильтра (10) лишь тем, что весовые функции в пространстве и по значениям являются Гауссовыми, а не прямоугольными. Фильтр (12) обеспечивает более “мягкое” сглаживание по сравнению с фильтрами (9)–(11).

Нижеследующие алгоритмы принципиально отличаются от предыдущих способом выбора множества усреднения. Они основаны на нахождении сглаженного значения по области анализа, форма которой может изменяться в зависимости от самого сигнала. Из них наиболее известен алгоритм сглаживания по геодезическим линиям, предложенный в [24, 25] для бинарных, а затем расширенный и модифицированный для полутоновых изображений [26–28]. Он основан на рассмотрении сглаживаемого сигнала $f(z)$ как поверхности на двумерной области задания Z . Для целевой точки z из окна анализа Ω_x , $z \in \Omega_x \subset Z$, на $f(z)$ отыскивается геодезическая линия $G(x, z)$ из точки x , длина которой $L(G(x, z))$ минимальна. Веса точек при суммировании задаются функцией $\psi(t)$, $t \geq 0$, быстро убывающей при возрастании t , например, $\psi(t) = \exp\{-kt\}$. Сглаживающий фильтр при этом выглядит:

$$s(x) = \sum_{z \in \Omega_x} f(z) \psi(L(G(x, z))) / \sum_{z \in \Omega_x} \psi(L(G(x, z))). \quad (13)$$

Сложность алгоритма (13) в отыскании геодезических для каждой возможной пары точек $(x, z) \in \Omega_x$. Наличие шума существенно затрудняет решение задачи. Для бинарных изображений геодезическая совпадает с кратчайшей “резиновой нитью”, огибающей объекты противоположного знака [25]. Для полутоновых изображений задача требует сочетания расстояний в пространственной и яркостной областях. В [26] предложен паллиативный подход, предполагающий предварительное обнаружение масок объектов и частично сводящий задачу к бинарному случаю. В [27] значение $L(G(x, z))$ определяется как интеграл произведения

$|f(z_1) - f(z_2)| \cdot |z_1 - z_2|$, а в [28] длина проекции геодезической на координатную плоскость суммируется с вариацией яркостей вдоль нее. И в том и в другом случае не приводятся достаточных обоснований операций произведения или суммирования пространственных и яркостных компонент.

Алгоритм сглаживания изображения с сохранением структуры [29] лишен указанного недостатка. Согласно модели (4), критерием для ограничения области усреднения должна быть не разница в значениях точек, а их принадлежность одной и той же связной области. В качестве критерия “связности” предлагается коэффициент сходства $A[x, y]$ пары точек x и y , определяемый на соединяющем их отрезке $[x, y]$:

$$A[x, z] = a(z)A[x, z - 1] = a_1 \cdot a_2 \cdot \dots \cdot a_N, \quad z \in [x, y].$$

При этом значение $a(z)$ задается как $a(z) = 1 - g(z)/g_{\max}$, где $g(z)$ — градиент $f(z)$ в точке z . Вклад точки z составляет $w(z) = A(x, z)\psi(d(x, z))$, где $\psi(d) = \exp\{-d^2/2\delta^2\}$, а $d(x, z)$ — расстояние между точками x и y . Сглаженное значение $s(x)$ вычисляется как:

$$s(x) = \sum_{z \in \Omega_x} f(z)w(x) / \sum_{z \in \Omega_x} w(z). \quad (14)$$

Эксперименты показывают [29], что эффективность алгоритма (14) выше, чем у других алгоритмов сглаживания.

Известны и другие алгоритмы сглаживания, ориентированные на сохранение контурных перепадов. Интересен подход [30], где отыскивается преобразование, сокращающее размерность пространства данных, но сохраняющее расстояния между признаками деталей изображения. В [31] предлагается использовать контуры для построения бинарных индикаторных функций, ограничивающих диапазон сглаживания.

Основная проблема сглаживания изображения — выбор того множества точек, по которому производить усреднение. Большинство методов основываются на косвенной информации о характеристиках области, например, распределении значений сигнала, и не используют данные о пространственной структуре изображения, в то время как именно она несет информацию о положении и характеристиках контурных перепадов. Отметим особенности алгоритмов сглаживания, позволяющие сформировать следующие группы.

Алгоритмы (6),(8) и в меньшей степени (7) не используют тот факт, что изображение состоит из множества областей с резкими границами, характеристики которых практически не связаны. Эти алгоритмы позволяют снизить шум и удалять мелкие детали, но их результатом является размытое изображение, поэтому они требуют тщательного выбора размера окна сглаживания, чтобы исключить потерю существенной информации. Такие методы, как сигма-фильтр (10), декомпозиция (11), билатеральный фильтр (12) позволяют вычленивать из локальной области анализа множество элементов, близких по значениям обрабатываемой точке. Они ориентированы на сохранение контурных перепадов, но контурные ограничения подменяются ограничениями в области значений.

Подавляющее большинство фильтров, применяемых для сглаживания изображения, основаны на статистическом анализе и не используют реальную структуру изображения. Общая черта алгоритмов (6)–(12), а также [22, 23] в том, что они основаны на анализе распределения значений элементов вокруг обрабатываемого элемента, и поэтому применимы как к одномерному, так и к двумерному сигналу. С одной стороны это является их преимуществом, но с другой стороны это означает, что часть доступной информации об изображении в алгоритмах сглаживания не используется.

Принципиально ориентированными на двумерные свойства сигнала являются методы сглаживания по геодезическим линиям [25–28], контурно-ограниченное сглаживание [29], сохраняющее структуру изображения, а также алгоритм [31]. Сравнительные характеристики большинства из них приведены в [29].

Исследования показывают, что задача сглаживания существенно связана с необходимостью сохранения структуры изображения и требует для каждой целевой точки тем или иным образом лимитировать область анализа границами объекта ее принадлежности. Такая постановка близка проблеме сегментации изображения.

4. СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЯ

Цель сегментации изображения — упрощение его представления, состоящее в разбиении исходного сигнала на множество непересекающихся связных областей, однородных по каким-то критериям и ассоциируемых с объектами наблюдаемой сцены или их частями. Результатом является получение карты, т.е. отображение точек изображения в набор индексов $1, \dots, Q$, где Q — общее число сегментов, объединение которых составляет полное изображение. В некоторых случаях индекс означает какую-то характеристику, например, цвет или текстуру. Соседние же сегменты существенно различаются по выбранной характеристике [7, 32, 34]. Такое разбиение может служить этапом для последующего анализа сцены или обнаружения объектов.

К результатам сегментации обычно предъявляются следующие требования [35]: а) выбранные характеристики изображения должны быть однородны внутри обнаруженных областей; б) смежные области должны существенно различаться по характеристикам; в) формы областей должны быть максимально простыми, а границы между ними точными. Отметим, что в данном случае нас интересует несемантическая сегментация, которая объединяет пиксели в связные объекты только на основе выбранных характеристик. Она отличается от семантической сегментации, задача которой — выделение на изображении участков, соответствующих тем или иным признакам [36], и которая, по сути, является задачей классификации.

Подходов к исследованию и решению задачи сегментации предложено много [34–43], но универсального метода, одинаково хорошо работающего применительно ко всем задачам, не существует, и сегментация по-прежнему остается одной из важнейших проблем анализа видеоинформации. Большинство алгоритмов строятся либо на использовании базовых свойств изображений, таких как однородность характеристик на связных областях и разрывность на их границах [12, 33, 37], либо на преобразовании сигнала в пространство признаков и классификации в этом пространстве [42, 44]. Из основных подходов к построению алгоритмов сегментации изображения можно выделить следующие.

1. Сегментация проведением границ. Заключается в обнаружении границ объектов с помощью тех или иных контурных операторов, в их прослеживании, связывании, и составлении из них замкнутых границ областей. Один из распространенных алгоритмов — метод активного контура, который проводит границы сегментов вдоль контуров [45]. Такие методы, как правило, применяются для специфических классов анализируемых изображений.

2. Морфологический подход. Из подобных методов наиболее известна сегментация по водоразделам [46, 47], идея которой состоит в интерпретации гладких областей с низким значением градиента яркости как локальных бассейнов, а контуров между ними, характеризующихся высоким градиентом, как водоразделов. В большинстве случаев результатом является излишне подробное разбиение изображения, что требует дополнительного его анализа.

3. Разбиение изображения на однородные области. Из них наиболее распространены пороговые методы, использующие как глобальные, так и адаптивные подходы [1, 48], а также методы выращивания областей [35, 38], заключающиеся в группировке элементов и мелких областей в более крупные. Важным является выбор подходящих мер близости и критериев

остановки процесса выращивания [49]. В этот же раздел относится и сегментация на основе графов, в которой пиксели интерпретируются как вершины, а вес ребра является мерой сходства, например близости значений яркости и цвета [50–52].

4. Классификация в пространстве признаков. Состоит в отображении набора входных данных в некоторое пространство признаков, в котором в дальнейшем и происходит решение задачи кластеризации — разбиение данного пространства на классы, базируясь на плотности распределения в нем [44, 53]. Традиционным является метод K -средних [54–56]. Трудности здесь — необходимость заранее задать число искомых классов K , а также нахождение поверхностей их разделения, особенно в случаях окружения одного класса другим. Еще одним распространенным подходом является использование алгоритма сдвига среднего значения [22], который находит центры кластеров в позициях с наибольшей локальной плотностью точек.

5. На основе нейронных сетей (НС). НС — граф со структурой, похожей на связи нейронов в мозге [57, 58], в вершинах которого расположены простые элементы (персептроны), реализующие сверточные и дискриминантные функции [59]. В анализе изображений успешными оказались сверточные нейронные сети конфигурации U-net, реализующие архитектуру “кодер-декодер” [60, 61] и применяемые в задачах классификации [62] и сегментации [63] изображений.

Несмотря на различия алгоритмов сегментации в подходах и качестве, принципы формирования окончательного результата у них одни и те же. Каждому элементу изображения присваивается некоторый индекс, соответствующий той области, к которой отнесен данный элемент в результате сегментации. Такой индекс формально никак не связан со значениями элементов области, и в дальнейшем его будет трудно сравнивать с сигналом, формируемым алгоритмом сглаживания. Чтобы операция сравнения по значениям стала возможной, вместо индекса можно записывать в каждый элемент сегментированного изображения среднее значение по той области, которой он принадлежит. Именно в таком виде мы и будем в дальнейшем рассматривать результат сегментации. Отметим также, что данный вариант является сравнительно частым способом представления данных.

5. СРАВНЕНИЕ ОПЕРАЦИЙ СГЛАЖИВАНИЯ И СЕГМЕНТАЦИИ

Рассмотренные в разделе 3 алгоритмы сглаживания с достаточной степенью точности описываются следующей общей формулой:

$$y_{mn} = f(W_{mn}(\mathbf{X})). \quad (15)$$

То есть значение изображения-результата y_{mn} в точке (m, n) является некоторой функцией от локальной (возможно, достаточно большой) области анализа $x_{ij} \in W_{mn}$ исходного изображения \mathbf{X} , центрированной в той же точке. Поскольку для каждой точки область анализа W_{mn} своя, то значения в соседних точках сглаженного изображения \mathbf{Y} , вообще говоря, разные. В случае же сегментации, как было сказано выше, все элементы каждой из найденных областей получают среднее значение сигнала по всей найденной области.

Результаты сглаживания и сегментации изображений можно сравнивать по двум критериям: по совпадению границ сглаженных/обнаруженных областей и по совпадению значений сигнала в соответствующих точках изображений. У сегментированного изображения границы областей уже изначально определены — как результат процедуры сегментации. У сглаженного же изображения нахождение границ областей требует дополнительного анализа. По сути, необходимо осуществить обнаружение контурных линий. В нашей работе для обнаружения контуров мы использовали алгоритм Кэнни [64], имеющий одни из наилучших показателей среди известных методов.

Обратим внимание на одну фундаментальную особенность всех процедур фильтрации помех, состоящую в том, что любая из них, построенная на основе какой-либо модели сигнала, приближает обрабатываемый сигнал по своим свойствам к этой самой модели. Соответственно, алгоритмы сглаживания, основанные на разрывной модели, в частности — кусочно-постоянной (5), имеют эффект повышения резкости контурных перепадов, что было отмечено в [11, 18]. Этот эффект несколько улучшает результаты последующего детектирования контуров.

Рассматривая второй критерий, отметим, что согласно модели (4) сглаживание изображения соответствует нахождению гладкой составляющей $s_{ij}^r(W_{mn})$ (3), приближаемой полиномом степени ω . Сегментация же более близка нахождению значения s_{mn}^r кусочно-постоянной модели (5), согласно которой значение s_{mn}^r находится по элементам области изображения r , ограниченной пересечением с областью анализа W_{mn} . Однако размеры объектов наблюдаемой сцены зачастую значительно превышают размеры используемых при сглаживании областей анализа, и на протяжении объекта значение s^r может заметно меняться.

Для визуального сравнения результатов сглаживания и сегментации изображений на Рис.1–4 приведены иллюстрации, которые построены по единой схеме: “а” — исходное изображение; “б” — результат сглаживания; “в” — результат сегментации; “г” — разность сглаженного “б” и сегментированного “в” изображений; “д” — контуры сглаженного изображения “б”; е) контуры сегментированного изображения “в”. Используемый в экспериментах алгоритм цветотекстурной сегментации [51, 52] был в свое время разработан и настроен для разбиения изображения на области по следующим характеристикам: объединение элементов по близости признаков цвета и текстуры при небольшом подавлении яркостных различий. В алгоритме [29] производилось сглаживание по цвето-яркостным компонентам без учета текстурной составляющей.

Визуальное сравнение изображений позволяет отметить следующее:

- результаты операций сглаживания и сегментации, показанные на изображениях “б” и “в”, весьма похожи, практически все объекты сохранены;
- этот вывод также подтверждается сравнением соответствующих контурных линий на рисунках “д” и “е”, у которых несовпадения небольшие;
- наиболее заметны различия результатов на Рис.3, где сегментация “склеила” небольшие объекты, а также на Рис.4, где возникли ложные границы;
- содержательными являются разностные изображения “г”, которые при точном совпадении изображений “б” и “в” должны быть однотонными серыми. На них можно обнаружить пропадания малых объектов при сегментации, несовпадения контурных линий, появление ложных границ, а также различия яркостей/цветностей из-за усреднения этих значений по обнаруженным объектам при сегментации;
- визуальные различия операций сглаживания и сегментации изображения по существу вызваны способами отображения результата: усреднением по локальной области сегмента или усреднением по всему сегменту.

Проведенные экспериментальные исследования показывают, что задачи сегментации и проведения контурных линий достаточно близки, и решать одну из них можно через решение другой.

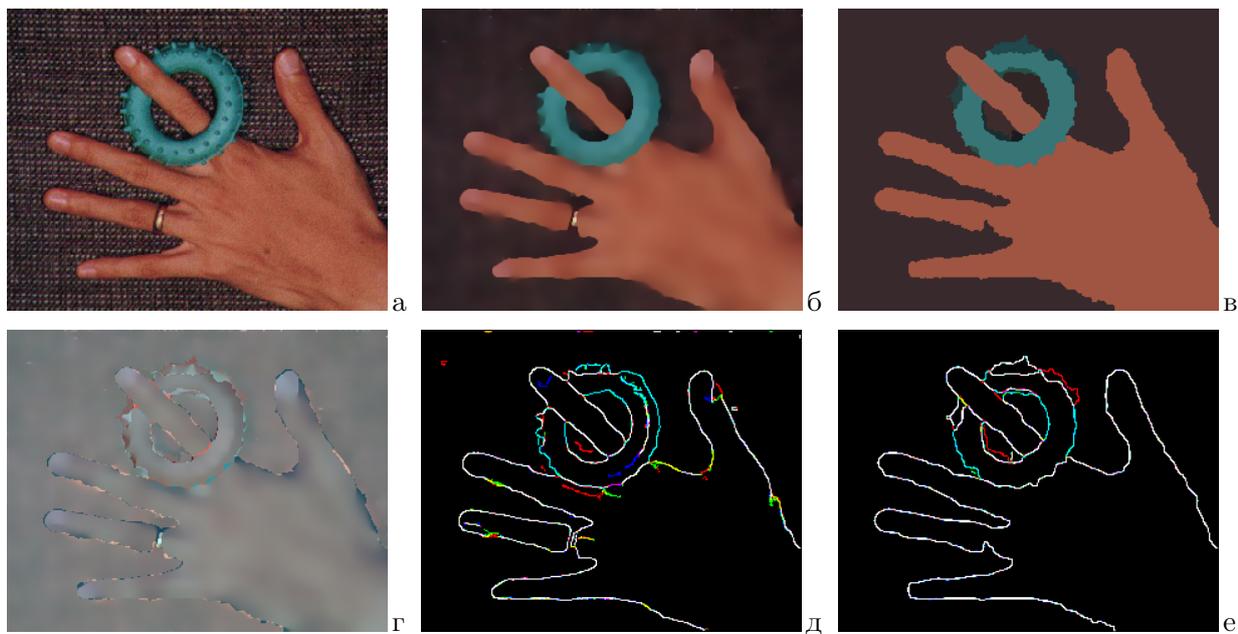


Рис. 1. а) исходное изображение “Рука”; б) сглаживание; в) сегментация; г) разность изображений (“б”–“в”); д) контуры изображения “б”; е) контуры изображения “в”.

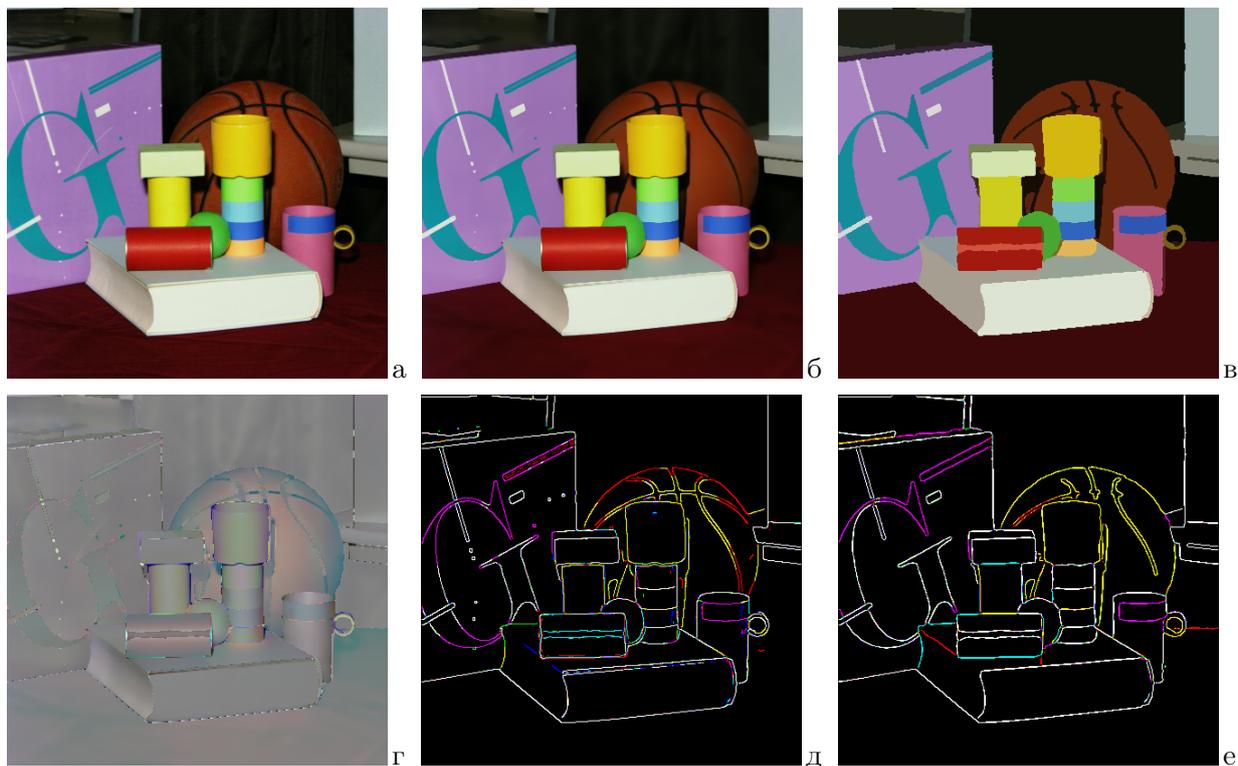


Рис. 2. а) исходное изображение “Сцена”; б) сглаживание; в) сегментация; г) разность изображений (“б”–“в”); д) контуры изображения “б”; е) контуры изображения “в”.

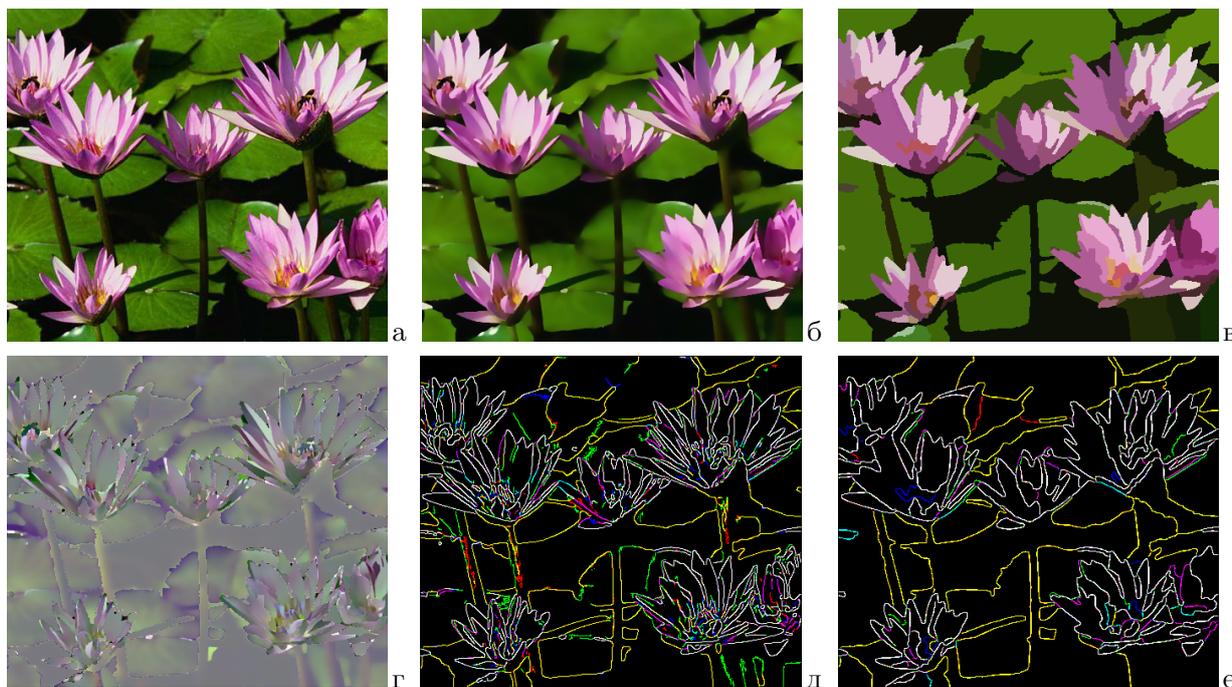


Рис. 3. а) исходное изображение “Водяные лилии”; б) сглаживание; в) сегментация; г) разность изображений (“б”–“в”); д) контуры изображения “б”; е) контуры изображения “в”.

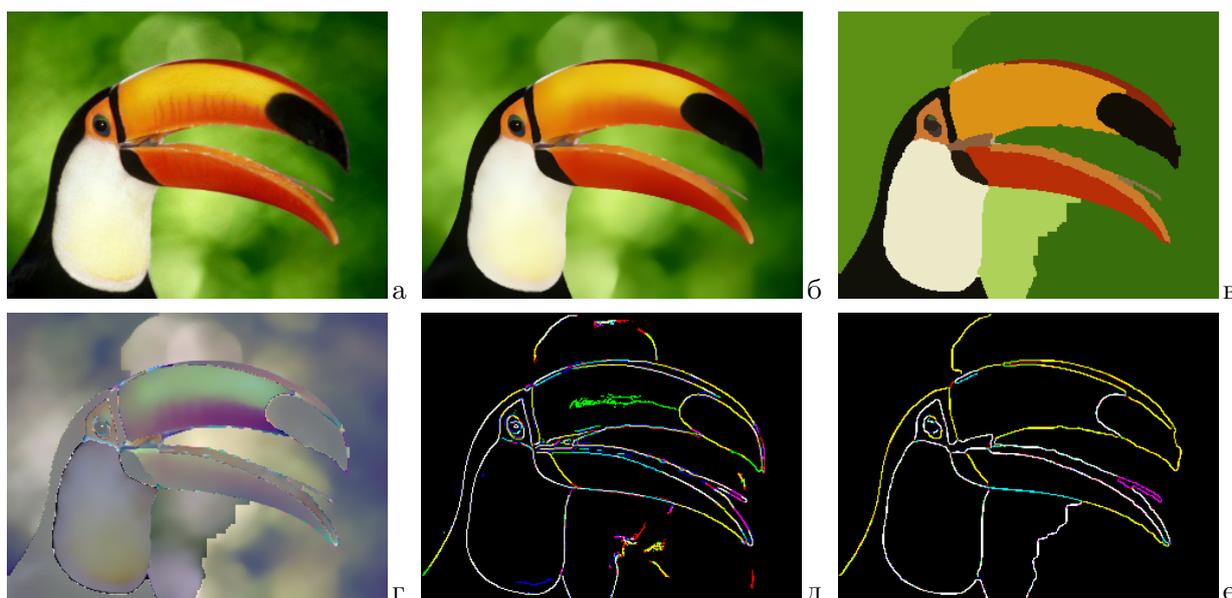


Рис. 4. а) исходное изображение “Тукан”; б) сглаживание; в) сегментация; г) разность изображений (“б”–“в”); д) контуры изображения “б”; е) контуры изображения “в”.

6. СХОДСТВО И РАЗЛИЧИЕ ЗАДАЧ СГЛАЖИВАНИЯ И СЕГМЕНТАЦИИ

Анализируя алгоритмы сглаживания и сегментации, можно сделать следующие заключения о сходстве или различии задач сглаживания и сегментации.

1. В обеих задачах изображение рассматривается как плотное множество непересекающихся областей, соответствующих объектам сцены.

2. Суть задач сегментации и сглаживания изображений в объединении элементов, близких по тем или иным признакам, в связные множества.

3. С позиции модели изображения, целью сглаживания является нахождение значения кусочно-гладкой компоненты s_{mn} в (5), а целью сегментации — нахождение границ зон непрерывности данной компоненты и объединение множества точек (пикселей) изображения в образующие области.

4. Точности обнаружения границ объектов в алгоритмах сглаживания и сегментации находятся приблизительно на одном уровне.

5. В сегментации при обнаружении границ объектов обязательным является замыкание контурных линий, что не требуется при сглаживании.

6. Операция заполнения обнаруженных областей средним значением при сегментации принципиально отличается от присваивания пикселям результирующего значения при сглаживании. В сглаживании — нахождение значения по локальной области сегмента, содержащего точку, а в сегментации — определение среднего значения по всему сегменту.

7. Если при сегментации формировать результат не заполнением сегмента одним числом, а каждой точке присваивать усредненное значение по локальной области сегмента, окружающей эту точку, то получим алгоритм сглаживания с резкими краями областей.

7. ВЫВОДЫ

Сравнение подходов к решению задач сглаживания и сегментации показывает, что обе они имеют общую идею, состоящую в том, что исходный сигнал — изображение — рассматривается как объединение связанных, но независимых областей. Центральным вопросом всех используемых для этого алгоритмов является проблема объединения множества элементов той области, которой принадлежит обрабатываемый элемент, и отделения их от элементов остальных областей.

Исследования также показали, что с точки зрения конечных результатов, операции сглаживания и сегментации изображения (при условии закраски сегментов средними значениями), являются сходными операциями. Основные визуальные различия результатов указанных операций вызваны особенностями формирования сигнала на сегментированном изображении.

СПИСОК ЛИТЕРАТУРЫ

1. Gonzalez R.C., Woods R.E. Digital Image Processing. — Prentice Hall, Upper Saddle River, New Jersey: Pearson Education, 2010. 976 p.
2. Simonoff J.S. Smoothing Methods in Statistics. N.-Y.: Springer-Verlag, 1996. 340 p.
3. O'Haver T. Smoothing. 2012. <<http://terpconnect.umd.edu/~toh/spectrum/Smoothing.html>>
4. Jahne B. Digital Image Processing. Berlin, Heidelberg: Springer-Verlag, 2005.
5. Skarbek W., Koschan A. Color Image Segmentation — A Survey // Technischer Bericht, Technical University of Berlin, 1994, pp. 94–32.
6. Lucchese L., Mitra S. Color Image Segmentation: A State-of-the-Art Survey // Image Processing, Vision, and Pattern Recognition. Proc. of the Indian National Science Academy (INSA-A), 2001, New Delhi, India, pp. 207–221.
7. Shapiro L.G., Stockman G.C. Computer Vision. New Jersey: Prentice-Hall, 2001. 580 p.
8. Tukey J.W. Exploratory Data Analysis // Reading, Mass.: Addison-Wesley, 1971.
9. Чочиа П.А. Двухмасштабная модель изображения // Кодирование и обработка изображений. М.: Наука, 1988, С. 69–87.
10. Chochia P.A. Image Enhancement Using Sliding Histograms // Computer Vision, Graphics, and Image Processing. 1988, vol. 44, no. 2, pp. 211–229.

11. Чочиа П.А. Методы обработки видеоинформации на основе двухмасштабной модели изображения. Saarbrücken: LAP Lambert Academic Publishing, 2017. 308 с.
12. Pratt W.K. Digital Image Processing. N.-Y.: Wiley & Sons, 2001. 738 p.
13. Yan J.K., Sakrison D.J. Encoding of Images Based on a Two-Component Source Model // IEEE Trans. COM-25, 1977, no. 11, pp. 1315–1322.
14. Nishikawa S., Massa R.J., Mott-Smith J.C. Area Properties of Television Pictures // IEEE Trans. IT-11, 1965, no. 3, pp. 348–352.
15. Arlinghaus S.L. Practical Handbook of Curve Fitting. Boca Raton, Florida: CRC Press, 1994. 272 p.
16. Li P., Wang H., Yu M., Li Y. Overview of Image Smoothing Algorithms // Journal of Physics: Conference Series, 2021, vol. 1883, no. 1, id. 012024, pp. 1–8.
17. Lee J.S. Digital Image Smoothing and the Sigma Filter // Computer Vision, Graphics, Image Processing, 1983, vol. 24, no. 2, pp. 255–269.
18. Чочиа П.А. Сглаживание изображения при сохранении контуров // Кодирование и обработка изображений. М.: Наука, 1988, С. 87–98.
19. Tomasi C; Manduchi R. Bilateral filtering for gray and color images // Proc. IEEE 6th Int. Conf. on Computer Vision. Bombay, India, 1998, pp. 839–846.
20. Parzen E. On Estimation of a Probability Density Function and Model // Annals of Mathematical Statistics, 1962, vol. 33, pp. 1065–1076.
21. Vincent P., Bengio Y. Manifold Parzen Windows // Advances in Neural Information Processing Systems, 2002, v. 15, pp. 825–832
22. Comaniciu D., Meer P. Mean shift: A robust approach toward feature space analysis // IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002, vol. 24, no. 5, pp. 603–619.
23. Paris S. Edge-Preserving Smoothing and Mean-Shift Segmentation of Video Streams // Computer Vision – ECCV 2008. Lecture Notes in Computer Science, vol. 5303, Berlin: Springer, pp. 460–473.
24. Lantuéjoul C., Beucher S. On the use of the geodesic metric in image analysis // Journal of Microscopy, 1981, vol. 121, no. 1, pp. 39–49.
25. Lantuéjoul C., Maisonneuve F. Geodesic methods in quantitative image analysis // Pattern Recognition, 1984, vol. 17, no. 2, pp. 177–187.
26. Criminisi A., Sharp T., Blake A. Geos: Geodesic image segmentation // Computer Vision – ECCV 2008, Springer, 2008, pp. 99–112.
27. Grazzini J., Soille P. Edge-preserving smoothing using a similarity measure in adaptive geodesic neighbourhoods // Pattern Recognition, 2009, vol. 42, pp. 2306–2316.
28. Mozerov M., van de Weijer J. Improved recursive geodesic distance computation for edge preserving filter // IEEE Trans. on Image Processing, 2017, vol. 26, no. 8, pp. 3696–3706.
29. Chochia P.A. Contour-Constrained Image Smoothing Preserving Its Structure // Journal of Communications Technology and Electronics, 2021, vol. 66, no. 6, pp. 769–777.
30. Gastal E.S.L., Oliveira M.M. Domain transform for edge-aware image and video processing // Proceedings of SIGGRAPH 2011, ACM Trans. Graph., 2011, vol. 30, no. 4, pp. 69:1–69:12.
31. Abiko R., Ikehara M. Fast edge preserving 2D smoothing filter using indicator function // IEICE Trans. on Information and Systems, 2019, vol. E102D, no. 10, pp. 2025–2032.
32. Hartigan J.A. Clustering algorithms. John Wiley & Sons, Inc., 1975. 351 p.
33. Rosenfeld A., Kak A.C., Digital Picture Processing, vol. 1, 2. N.-Y.: Academic press, 1982.
34. Zhang Y., ed. Advances in Image and Video Segmentation. Hershey: IRM Press. 2006. 472 p.

35. Haralick R.M., Shapiro L.G. Image Segmentation Techniques // Computer Vision, Graphics, Image Processing, 1985, vol. 29, no. 2. pp. 100–132.
36. Thoma M. A Survey of Semantic Segmentation // arXiv:1602.06541v2 [cs.CV], 2016, pp. 1–16.
37. Pal N.K., Pal S.K. A Review on Image Segmentation Techniques // Pattern Recognition, 1993, vol. 26, no. 9. pp. 1277–1293.
38. Fu K.S., Mui J.K. A Survey of Image Segmentation // Pattern Recognition, 1981, vol. 13, no. 1. pp. 3–16.
39. Skarbek W., Koschan A. Color image segmentation — A survey: Technischer Bericht // Technical Univ. of Berlin, 1994, pp 94–32.
40. Luccheseysz L., Mitray S.K. Color image segmentation: A state-of-the-art survey // Proc. of the Indian National Science Academy (INSA-A), 2001, vol. 67, no. 2, pp. 207–221.
41. Rosenfeld A. Image pattern recognition // Proc. of IEEE, 1981, vol. 69, no. 5, pp. 596–605.
42. Pantofaru C., Hebert M. A comparison of image segmentation algorithms // Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, 2005. 31 p.
43. Unnikrishnan R., Pantofaru C., Hebert M. Toward Objective Evaluation of Image Segmentation Algorithms // IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, vol. 29, no. 6, pp. 929–944.
44. Duda R.O., Hart P.E. Pattern Classification and Scene Analysis. N.–Y., London: Wiley, 1973. 512 p.
45. Kass M., Witkin A., Terzopoulos D. Snakes: Active contour models // International journal of computer vision, 1988, vol. 1, no. 4, pp. 321–331.
46. Image Analysis and Mathematical Morphology. Vol. 2: Theoretical Advances / Ed. J.Serra. N.–Y.: Academic Press, 1988. 411 p
47. Special issue on Mathematical Morphology and Nonlinear Image Processing // Pattern Recognition, 2000, vol. 33, no. 6. pp. 875–1117.
48. Jain R., Kasturi R., Schunk B. Machine Vision. New York: McGraw-Hill, 2003, pp. 63–69.
49. Haralick R.M., Shapiro L.G. Computer and Robot Vision, vols. 1, 2. Reading, MA: Addison-Wesley, 1992.
50. Felzenszwalb P.F., Huttenlocher D.P. Efficient graph-based image segmentation // International Journal of Computer Vision, 2004, vol. 59, no. 2, pp. 167–181.
51. Chochia P.A. A Pyramidal Image Segmentation Algorithm // Journal of Communications Technology and Electronics, 2010, vol. 55, no. 12, pp. 1550–1560.
52. Chochia P.A. Image Segmentation Based on the Analysis of Distances in an Attribute Space. // Optoelectronics, Instrumentation and Data Processing, 2014, vol. 50, no. 6, pp. 1–12.
53. Jain A.K., Dubes R.C. Algorithms for Clustering Data. Prentice Hall, 1988. 320 p.
54. Steinhaus H. Sur la division des corps materiels en parties // Bulletin L'Académie Polonaise des Science, 1956, Cl.III, vol. IV, no. 12. pp. 801–804.
55. Lloyd S.P., McMillan D. Linear least squares filtering and prediction of sampled signals // Proc. Symp. On modern Network Synthesis, Brooklyn, N.-Y.: Polytechnic Institute of Brooklyn, 1956, vol. 5, pp. 221–247.
56. Lloyd S.P. Least square quantization in PCM // IEEE Trans. Information Theory, 1982, vol. IT-28, no. 2, pp. 129–137.
57. LeCun Y., Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W., Jackel L.D. Backpropagation Applied to Handwritten Zip Code Recognition // Neural Computation, 1989, vol. 1, no. 4, pp. 541–551.
58. Long J., Shelhamer E., Darrell T. Fully Convolutional Networks for Semantic Segmentation // The IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431–3440.

59. Годунов А.И., Баламян С.Т., Егоров П.С. Сегментация изображений и распознавание объектов на основе технологии сверточных нейронных сетей // Надежность и качество сложных систем, 2021, № 3, С. 62–73.
60. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // International Conf. on Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015. Lecture Notes in Computer Science, 2015, vol. 9351, pp. 234–241.
61. Venkatesan R., Li B. Convolutional Neural Networks in Visual Computing: A Concise Guide. Boca Raton, FL: CRC Press, 2018. 168 p.
62. Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks // Communications of the ACM, 2017, vol. 60, no. 6, pp. 84–90.
63. Badrinarayanan V., Kendall A., Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation // IEEE Trans. on pattern analysis and machine intelligence, 2017, vol. 39, no. 12, pp. 2481–2495
64. Canny J. A computational approach to edge detection // IEEE Trans. on Pattern Analysis and Machine Intelligence, 1986, vol. 8, no. 6, pp. 679–698.

On the affinity of image smoothing and segmentation problems

P. A. Chochia

One of the most important directions of video information processing and analysis are the tasks of image smoothing and segmentation, which have much in common in their formulation, final goals and solution methods. Based on the image model, by analyzing and comparing known smoothing and segmentation algorithms, the similarities and differences of the mentioned tasks are shown.

KEYWORDS: image processing, smoothing, segmentation