УДК 004.932
doi: 10.18101/2304-5728-2017-2-61-70

МОДЕЛЬ ЦИФРОВОГО ИЗОБРАЖЕНИЯ НА ОСНОВЕ МОДИФИЦИРОВАННОГО МЕТОДА УОРДА КЛАСТЕРИЗАЦИИ ПИКСЕЛЕЙ

© Ханыков Игорь Георгиевич
аспирант, младший научный сотрудник,
Санкт-Петербургский институт информатики и автоматизации
Российской академии наук (СПИИРАН)
Россия, 199178, Санкт-Петербург, 14 линия, 39
E-mail: igk@iias.spb.su

© Харинов Михаил Вячеславович
кандидат технических наук, доцент, старший научный сотрудник,
Санкт-Петербургский институт информатики и автоматизации
Российской академии наук (СПИИРАН)
Россия, 199178, Санкт-Петербург, 14 линия, 39
E-mail: khar@iias.spb.su

В статье рассматривается модель детектирования дихотомической иерархии объектов на изображении, предназначенная для предобработки изображений сцен любого содержания. В основе модели лежит метод Уорда кластеризации пикселей. Предусматривается тестирование результатов на правдоподобие посредством обработки стереопар. Обсуждаются особенности применения метода Уорда для кластеризации пикселей. Выполняется сравнение с известным решением, принятым в качестве прототипа.
Ключевые слова: сегментация изображения; кластеризация пикселей; суммарная квадратичная ошибка; минимизация; кусочно-постоянное приближение; иерархическая последовательность; выпуклая последовательность значений; двухмасштабная модель; метод Уорда; модель Мамфорда-Шаха.

Введение

В модели [1, 2] любое изображение есть отображение совокупности объектов наблюдаемой сцены на пространство изображения. Поскольку единого подхода к описанию цифрового изображения не выработано, разработчики, как правило, используют априорную информацию о свойствах запечатленных объектов. Согласно [1, 2] формулировка указанных свойств осуществляется в виде математических функций, задающих модель цифрового изображения. Подобная модель считается необходимой для разработки эффективных методов анализа и преобразования изображений, в частности – методов сегментации. Сегментация заключается в разбиении изображения на множество областей, ассоциируемых с объектами сцены.

В обобщенном смысле целью сегментации является перечисление всех объектов, детектируемых на изображении компьютером, что оказывается

проблематичным в отсутствие априорной информации об «объектах интереса».

Общим недостатком большинства алгоритмов сегментации является неустойчивость выделения объектов на изображениях заранее не известного содержания при изменении условий съемки. Множество релевантных алгоритмов сегментации можно существенно сократить, если потребовать, чтобы они давали сходные результаты выделения объектов хотя бы при обработке стереопар, когда каждое из двух изображений одной и той же сцены сегментируется по отдельности, а также когда оба изображения при сегментации объединяются в одно (рис. 1). Если при обработке раздельных стереоизображений моделируется детектирование одних и тех же объектов на различных изображениях, то при объединении стереоизображений (рис. 1) моделируется выделение одинаковых объектов на одном изображении.

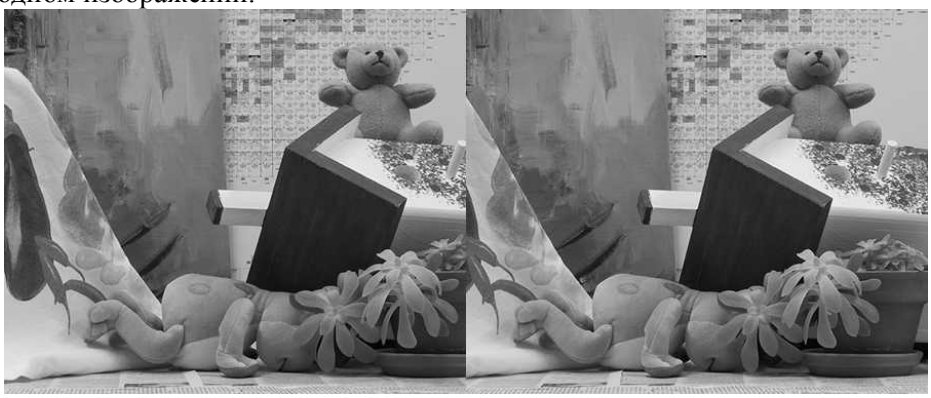

Рис. 1. Стандартные стереоизображения «Teddy», объединенные в единое изображение для тестирования алгоритмов сегментации изображений произвольной сцены.

1. Постановка задачи

Метод Уорда [3] – это классический метод кластеризации пикселей изображения [4, 5]. Кластеризация пикселей по Уорду выполняется итеративным слиянием кластеров 1, 2, обеспечивающих минимальное приращение суммарной квадратичной ошибки ΔE_{merge}:

$$1,2: \quad \Delta E_{merge}(1,2) = \min ,$$

где ΔE_{merge} выражается через число пикселей n_1, n_2 и средние яркости I_1, I_2 в кластерах 1, 2 в виде:

$$\Delta E_{merge}(1,2) = \frac{n_1 n_2}{n_1 + n_2} \left\| I_1 - I_2 \right\|^2 \geq 0$$

Для обработки современных изображений из миллионов пикселей метод Уорда применяется крайне редко из-за большой вычислительной сложности [6], которая возрастает с ростом числа пикселей N в изображении кубично, или квадратично. При этом квадратичная зависимость в

традиционном методе Уорда обеспечивается только при нетривиальной оптимизации алгоритма слияния кластеров, а дальнейшее снижение вычислительной сложности до $N\sqrt[3]{N}$ обеспечивается за счет кластеризации пикселей изображения по частям [7].

Если при минимизации приращения ΔE_{merge} вместо всевозможных пар кластеров анализируются только пары укрупняемых смежных множеств пикселей, то кластеризация пикселей на начальном этапе переходит в сегментацию изображения подобно модели Мамфорда-Шаха [8–10], что позволяет повысить скорость вычислений по Уорду за счет предварительного уменьшения числа обрабатываемых множеств пикселей, называемых *суперпикселями*.

2. Двухмасштабная модель многокомпонентного изображения

В работах [1, 2, 13 ,14] выдвигается идея двухэтапной сегментации изображения, при которой «объекты интереса» сначала локализуются в некотором фрагменте изображения, а затем детектируются в масштабе предварительно выделенного фрагмента.

Согласно фундаментальной работе [1, 2] изображение представляется в *двухмасштабной* модели суммой кусочно-гладкой компоненты, задающей протяженные объекты с резкими границами между ними (главной компоненты), и дополнительной компоненты, которая включает в себя информацию о текстуре, мелких деталях и пр., а также шумовой компоненты с нормальным распределением пикселей по разностной яркости. В качестве *объектов,* разделенных контурными границами, рассматриваются покрывающие изображение связные сегменты, которые различаются по цвету и другим признакам. Для повышения скорости обработка в [1, 2] выполняется в паре скользящих окон прямоугольной формы, называемых «областями анализа». При этом неадаптивность обработки компенсируется специальными приемами сглаживания границ.

В двухмасштабной многокомпонентной модели [1, 2] задача декомпозиции изображения на компоненты с различным информационным содержанием учитывает семантику объектов и решается с учетом априорной информации, что усложняет модель и является ее основным недостатком. При этом определение главной компоненты в модели [1, 2] остается недостаточно формализованным.

2. Модель дихотомической иерархии объектов

В предлагаемой модели дихотомической иерархии объектов исходное изображение разбивается на иерархически структурированные максимальные объекты. *Максимальные объекты* – это сегменты изображения или кластеры пикселей, которые при последующей обработке анализируются как самостоятельные изображения. При этом *максимальным объектом* считается либо исходное изображение, либо любая его часть, для которой вычисляется дихотомическая последовательность кусочно-

постоянных приближений. Термин «максимальный объект» в модели дихотомической иерархии является синонимом термина «изображение» и употребляется для обозначения составной части исходного изображения.

Итеративное деление максимальных объектов надвое порождает вложенные кластеры пикселей или сегментов изображения. Они представляют собой меньшие объекты, которые в отличие от максимальных объектов взаимозависимы и не могут быть любыми фрагментами изображения.

Обработка входного изображения из N пикселей в модели дихотомической иерархии объектов состоит в получении на выходе последовательности из N приближений изображения. Выходная последовательность из N приближений характеризуется последовательным возрастанием числа $g = 1, 2, \ldots N$ кластеров пикселей (или сегментов изображения). Различные $2N-1$ множества пикселей рассматриваются как *объекты*, детектируемые компьютером.

Свойство 1. Основной особенностью указанной последовательности приближений является то, что она описывается выпуклой последовательностью $E_1, E_2, \ldots, E_N \equiv 0$ значений суммарной квадратичной ошибки E:

$$E_g \le \frac{E_{g-1} + E_{g+1}}{2}, \quad g = 2, 3, \ldots, N-1$$

или эквивалентной монотонно возрастающей последовательностью $\Delta E_2 \le \Delta E_3 \le \ldots \le \Delta E_{N-1} \le 0$ неположительных приращений E. Характерно, что по этому признаку обсуждаемая последовательность иерархических приближений имитирует последовательность оптимальных приближений, которая отвечает минимально возможным значениям E, но в общем случае не является иерархической.

Свойство 2. Другой особенностью дихотомической иерархической последовательности приближений является ее неоднозначность, возникающая при использовании метода Уорда. Она обусловлена выраженной повторяемостью минимальных значений приращения суммарной квадратичной ошибки на начальных шагах укрупнения кластеров или сегментов. Результат итеративных вычислений определяется выбором на каждом шаге очередной пары объединяемых множеств пикселей из набора равноправных вариантов. Результирующая кластеризация зависит от порядка слияния пар кластеров, а изображение аппроксимируется многочисленными иерархическими последовательностями *квазиоптимальных* приближений, которые описываются переплетающимися выпуклыми кривыми (рис. 2).

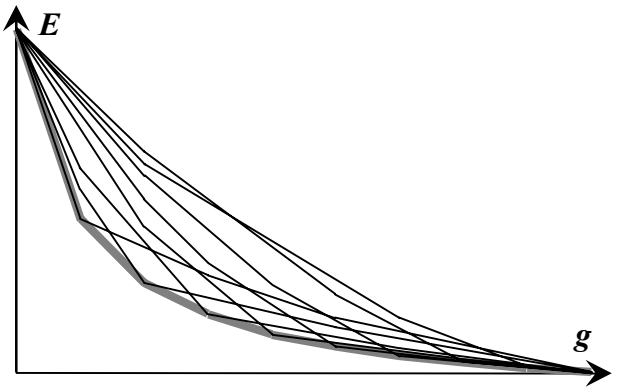

Рис. 2. Неоднозначная аппроксимация последовательности оптимальных приближений (граничная серая кривая) выпуклыми последовательностями квазиоптимальных приближений (верхние кривые)

Сопоставление двухмасштабной модели [1, 2] с предлагаемой моделью дихотомической иерархии объектов иллюстрируется в табл. 1.

Таблица 1.
Сопоставление моделей

модель / признак	Двухмасштабная модель П. А.Чочиа	Двухмасштабная модель дихотомической иерархии объектов
Количество разбиений на выходе	Единственное разбиение	Множество из N разбиений
Число объектов на выходе	–	$2N-1$
Представление изображения	Аддитивная сумма нескольких компонент	Набор максимальных структурированных объектов
Определение компонент	Сумма кусочно-гладкой и аддитивной компонент	Иерархия из N компонент, определяемых приближениями изображениями
Метод анализа изображения	Анализ множеств пикселей в паре квадратных окошек различного масштаба	Анализ изображения в адаптивных алгоритмах обратимого слияния кластеров пикселей или сегментов изображения
Описание объектов	Посредством связных сегментов, не объединяемых в кластеры	Посредством кластеров пикселей, в частности, сегментов изображения
Определение компонент	Кусочно-гладкая и разностная компоненты, подразделяемые на	Кусочно-постоянные приближения изображения, описываемые установленным

	мелкие объекты, шумы и пр.	свойством выпуклости соответствующей последовательности значений E
Генерация разбиений	Сложный алгоритм декомпозиции	Слияние кластеров пикселей по Уорду, или смежных сегментов изображения в модели Мамфорда-Шаха

3. Экспериментальные результаты

Рис. 3 иллюстрирует преимущества кластеризации пикселей по Уорду [3–6] в сравнении со стандартной сегментацией изображения по Мамфорду-Шаху на примере обработки стереопары.

На рис. 3 вверху показано представление тестового стереоизображения рис. 1, которое получено сегментацией по Мамфорду-Шаху и содержит 256 сегментов различного цвета. Под ним показано представление того же изображения в 256 цветах, полученное кластеризацией пикселей по Уорду.

Как можно заметить, на рис. 3 при сегментации по Мамфорду-Шаху одни и те же объекты сцены раскрашиваются по-разному, тогда как при кластеризации пикселей по Уорду достигается сходная раскраска объектов.

Рис. 4 иллюстрирует влияние объединения изображений в единое целое на детектирование объектов. При сегментации по Мамфорду-Шаху наблюдается выраженная разница результатов обработки отдельных и объединенных изображений. При этом в случае обработки по Уорду разностное изображение получается более сглаженным.

Представленные результаты наглядно демонстрируют преимущество кластеризации писклей по Уорду в сравнении с сегментацией изображений по Мамфорду-Шаху. Об этом свидетельствует как субъективная оценка качества визуального восприятия, так и объективная — соотношение численных значений среднеквадратичного отклонения σ и числа цветов. Сам по себе предложенный прием обработки стереопар может служить для проверки адекватности результатов, получаемых различными методами сегментации изображений или кластеризации пикселей.

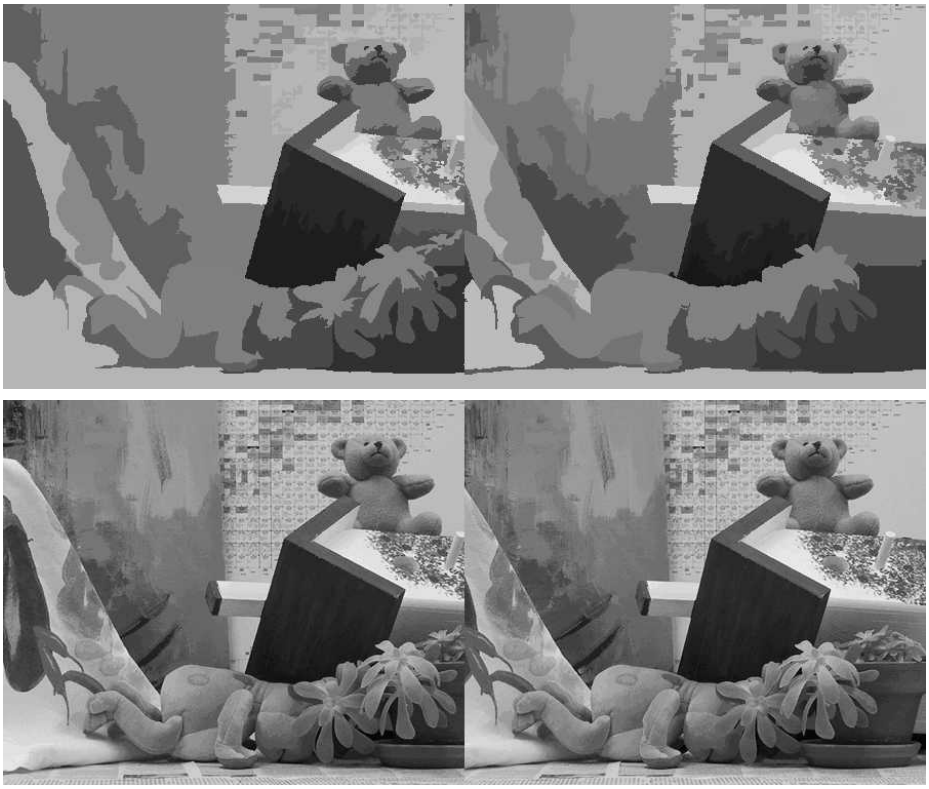

Рис. 3. Результаты совместной обработки стереопары как единого изображения по Мамфорду-Шаху (вверху, σ=24,07426) и Уорду (внизу, σ=5,48068) при числе кластеров g=256.

Заключение

Таким образом, в настоящей работе рассмотрена модель дихотомической иерархии объектов на изображении, в основе которой лежит *модифицированный* метод Уорда, в котором скоростная кластеризация и сегментация достигается за счет:

1) обработки ограниченного числа *суперпикселей* (укрупненных пикселей);

2) обработки пикселей изображения по частям [7];

3) ограничения числа обрабатываемых пар смежных сегментов изображения при формировании суперпикселей в модели Мамфорда-Шаха.

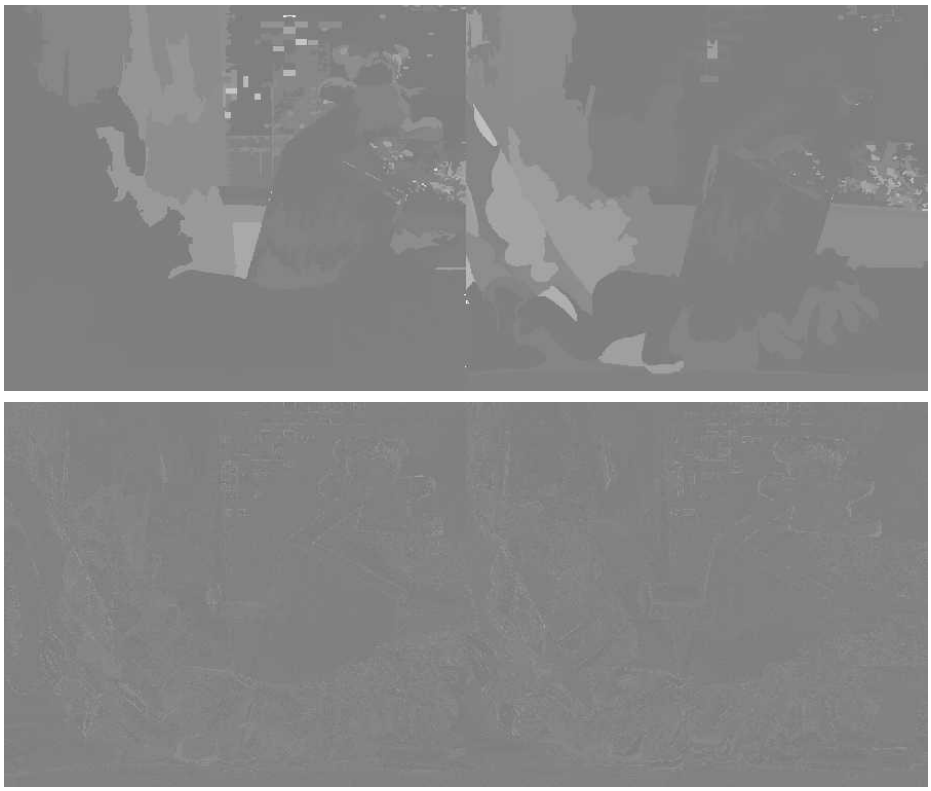

Рис. 4. Разностное представление результатов обработки единого изображения и обработки каждого из двух изображений по отдельности по Мамфорду-Шаху (вверху, 75 цветов) и Уорду (внизу, 42 цвета).

В качестве прототипа модели дихотомической иерархии объектов принята двухмасштабная модель изображения [1, 2], основными ограничениями которой являются:

1. На выходе вычисляется единственное разбиение исходного изображения.

2. Объекты описываются посредством связных сегментов, не объединяемых в кластеры.

3. Анализ множеств пикселей выполняется в паре квадратных окошек различного масштаба.

В модели дихотомической иерархии объектов указанные ограничения преодолеваются:

1. На выходе вычисляются варианты иерархических последовательностей квазиоптимальных приближений, с числом классов от 1 до N.

2. Объекты описываются посредством кластеров пикселей и сегментов изображения.

3. Анализ изображения выполняется в адаптивных алгоритмах слияния/разделения множеств пикселей.

Смысл модели дихотомической иерархии объектов прозрачен и сво-

дится к тому, что исходное изображение разбивается на самостоятельные иерархически структурированные изображения объектов давным-давно известным методом Уорда [3]. При этом нетривиальный формализм модели состоит в организации вычислений, которые предполагается оформить в виде стандартных утилит обработки, чтобы ими пользовались программисты-инженеры, не вдаваясь в детали математического описания.

Судя по экспериментальным результатам, метод Уорда обеспечивает устойчивое детектирование на изображении объектов в виде близких по усредненным цветам кластеров пикселей, которые «видит» компьютер. Внедрение метода Уорда в практику обработки изображений приводит к двухмасштабной модели, которая согласуется с моделью [1, 2], но для эффективной реализации требует разработки специальной структуры данных [7, 15], являющейся основой вычислительной части двухмасштабной модели дихотомической иерархии объектов.

Литература

1. Чочиа П. А. Некоторые алгоритмы обнаружения объектов на основе двухмасштабной модели изображения // Информационные процессы. 2014. Т.14, №2. С.117–136.

2. Чочиа П.А. Теория и методы обработки видеоинформации на основе двухмасштабной модели изображения: дис. … д-ра техн. наук. М.:ИППИ РАН, 2016. 302 с.

3. Ward J. H., Jr. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 1963. Vol. 58, Issue 301. P. 236–244.

4. Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: Классификация и снижение размерности. М.: Финансы и статистика, 1989. 607 с.

5. Мандель И. Д. Кластерный анализ. М.: Финансы и статистика, 1988. 176 с.

6. Jain A. K., Murty M.N., Flynn P.J. Data clustering: a review // ACM computing surveys (CSUR). 1999. Vol. 31, №3. P. 264–323.

7. Харинов М. В., Ханыков И. Г. Применение метода Уорда для кластеризации пикселей цифрового изображения // Вестник Бурятского государственного университета. Математика, информатика. 2016. № 4. С. 34–42.

8. Бугаев А. С., Хельвас А. В. Поисковые исследования и разработка методов и средств анализа и автоматического распознавания потоковой информации в глобальных информационных системах. Шифр «Лацкан» // Отчет по НИР. М.: Изд-во МФТИ, 2001. Т. 1. 140 с.

9. Mumford D., Shah J. Boundary detection by minimizing functionals, I // Proc. IEEE Comput. Vision Patt. Recogn. Conf., San Francisco, 1985. P. 22–26.

10. Jenatton R., Gramfort A., Michel V., Obozinski G., Eger E., Bach F., Thirion B. Multiscale mining of fMRI data with hierarchical structured sparsity // SIAM Journal on Imaging Sciences. 2012. Vol. 5, №. 3. P. 35–856.

11. Mumford D., Shah J. Boundary detection by minimizing functionals, I // Proc. IEEE Comput. Vision Patt. Recogn. Conf., San Francisco, 1985. P. 22–26.

12. Bar L., Chan T.F., Chung G., Jung M., Vese L.A., Kiryati N., Sochen N. Mumford and Shah Model and Its Applications to Image Segmentation and Image Restoration. Handbook of Mathematical Methods in Imaging. 2015. P. 1539–1597.

13. Визильтер Ю. В., Желтов С. Ю. Проблемы технического зрения в современных авиационных системах // Механика, управление и информатика. 2011. №6. С. 11–44.

14. Луцив В. Р. Объектно-независимый подход к структурному анализу изображений: дис. … д-ра техн. наук. СПб.: ГУАП, 2011. 318 с.

15. Philipp Galiano, Mikhail Kharinov, Victor Kuzenny. Remote Sensing Data Analysis Based on Hierarchical Image Approximation with Previously Computed Segments // Information Fusion and Geographic Information Systems: Towards the Digital Ocean (IF&GIS'2011) / Proceedings of the Fifth International Workshop 10-11 May, 2011, Brest (France): Springer, LNG&C. P. 105–115.

A MODEL OF DIGITAL IMAGE PREPROCESSING BASED ON PIXEL CLUSTERING BY MODIFIED WARD'S METHOD

Igor G. Khanykov
Research Assistant, Junior Researcher,
St. Petersburg Institute for Informatics and Automation, RAS,
39, 14th Line, St. Petersburg 199178, Russia
E-mail: igk@iias.spb.su

Mikhail V. Kharinov
Cand. Sci. (Engineering), A/Prof., Senior Researcher,
St. Petersburg Institute for Informatics and Automation, RAS,
39, 14th Line, St. Petersburg 199178, Russia
E-mail: khar@iias.spb.su

The article deals with the model of detecting the dichotomous hierarchy of objects in the image. The model based on Ward's method of pixel clustering is intended for preprocessing images of scenes of any content. It is proposed to test the results for credibility by processing of stereopairs. We discuss the features of pixel clustering by Ward's method and compared with the known solution accepted as a prototype.
Keywords: image segmentation; pixel clustering; integral square error; minimization; piecewise constant approximation; hierarchical sequence; convex sequence of values; double scaled model; Ward's method; Mamford-Shah segmentation model.