

# Кластеризация пикселей цветового изображения методом Уорда

И.Г. Ханьков, М.В. Харинов  
{igk, khar}@iias.spb.su

Санкт-Петербургский Институт информатики и автоматизации Российской академии наук (СПИИРАН),  
Санкт-Петербург, Россия

*В статье рассматривается кластеризация пикселей, предназначенная для сцен любого содержания. Для эффективного выделения объектов на изображениях применяется метод Уорда и способ тестирования алгоритмов на правдоподобие с использованием стереопар. Обсуждаются особенности применения метода Уорда кластеризации пикселей, и предлагается модель применения метода Уорда. Выполняется сравнение с аналогичными решениями.*

**Ключевые слова:** сегментация изображения, кластеризация пикселей, суммарная квадратичная ошибка, минимизация, иерархия приближений, выпуклая последовательность.

## Ward's pixel clustering for color image

I.G. Khanykov, M.V. Kharinov  
{igk, khar}@iias.spb.su

St.Peterburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS),  
St.Petersburg, Russia

*In the paper pixel clustering, intended for scenes of any content is considered. To effectively detect objects, Ward's method and the method for testing the believability of algorithms using stereo-pairs are proposed. The features of the Ward's pixel clustering are discussed, and a model of the Ward's method application is developed. Comparison with similar solutions is performed.*

**Keywords:** image segmentation, pixel clustering, total squared error, minimization, hierarchy of approximations, convex sequence.

### 1. Введение

На сегодняшний день, для расширяющейся предметной области обработки изображений разработано множество алгоритмов выделения объектов и продолжают разрабатываться новые алгоритмы для новых изображений и возникающих задач. Общим недостатком большинства алгоритмов является неустойчивость выделения объектов на изображениях заранее неизвестного содержания при изменении условий съемки. Число релевантных алгоритмов можно существенно сократить, если потребовать, чтобы они давали сходные результаты выделения объектов хотя бы при обработке стереопар, при которой каждое из двух изображений одной и той же сцены обрабатывается по отдельности, а также когда оба изображения для выделения объектов объединяются в одно (рис. 1). Тем самым, в первом случае (рис.1, сверху), моделируется выделение одних и тех же объектов на различных изображениях, а во втором случае (рис.1, внизу), моделируется выделение одинаковых объектов на одном изображении.

### 2. Двухмасштабная модель изображения

В работах [3,4,6,7] 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 [6,7] изображение представляется в *двухмасштабной* модели суммой кусочно-гладкой компоненты, задающей протяженные объекты с резкими границами между ними, и дополнительной компоненты, которая включает в себя информацию о текстуре, мелких деталях, шумовую компоненту и пр.. В качестве *объектов*, разделенных контурными границами, рассматривается объединение покрывающих изображение связанных сегментов, которые различаются по цвету и др. признакам. В вычислениях существенно используются априорная информация об «объектах интереса».



**Рис.1.** Стандартные стереоизображения «Teddy» для тестирования алгоритмов сегментации изображений произвольной сцены.

### 3. Метод Уорда и двухмасштабная модель дихотомической иерархии объектов

*Модель дихотомической иерархии объектов* развивается с начала 90-х годов [8] и фактически является моделью применения метода Уорда [13,16] для кластеризации пикселей изображения, что при современных изображениях из миллионов пикселей является сложной научно-технической проблемой. Метод Уорда – это классический метод кластеризации [1,5], который в обработке изображений используется крайне редко [13] из-за большой вычислительной сложности, квадратично возрастающей с ростом числа пикселей  $N$  в изображении.

Для изображения из  $N$  пикселей в модели дихотомической иерархии объектов на выходе получается последовательность из  $N$  кусочно-постоянных приближений изображения с последовательно

возрастающим числом  $g = 1, 2, \dots, N$  кластеров пикселей. При этом вычисляется  $2N - 1$  различных множеств пикселей. Основной особенностью указанной последовательности приближений является то, что она описывается выпуклой последовательностью  $E_1, E_2, \dots, E_N \equiv 0$  значений суммарной квадратичной ошибки  $E$ :

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

или монотонно возрастающей последовательностью  $\Delta E_2 \leq \Delta E_3 \leq \dots \leq \Delta E_{N-1} \leq 0$  неположительных приращений  $E$ . Характерно, что по этому признаку обсуждаемая последовательность иерархических приближений имитирует последовательность оптимальных приближений. Последняя отвечает минимально возможным значениям  $E$  и среднеквадратичного отклонения  $\sigma$  приближения от изображения ( $E \equiv 3N\sigma^2$ ), но в общем случае не является иерархической.

Кластеризация пикселей по Уорду выполняется посредством итеративного слияния кластеров 1, 2, а именно тех, что обеспечивают минимальное приращение суммарной квадратичной ошибки  $\Delta E_{merge}$ :

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

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

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

Если при минимизации приращения  $\Delta E_{merge}$  вместо всевозможных пар кластеров анализируются только пары укрупняемых смежных множеств пикселей, то кластеризация переходит в сегментацию, как в модели Мамфорда-Шаха [2,14,15], позволяющей повысить скорость вычислений по Уорду за счет предварительного уменьшения числа обрабатываемых множеств пикселей.

Для операции *merge* при вычислениях поддерживается обратная операция *split* разделения кластера, в частности, сегмента на два, слиянием которых данный кластер или сегмент был получен. При этом приращение  $\Delta E_{split}$  суммарной квадратичной ошибки  $E$  получается с обратным знаком:

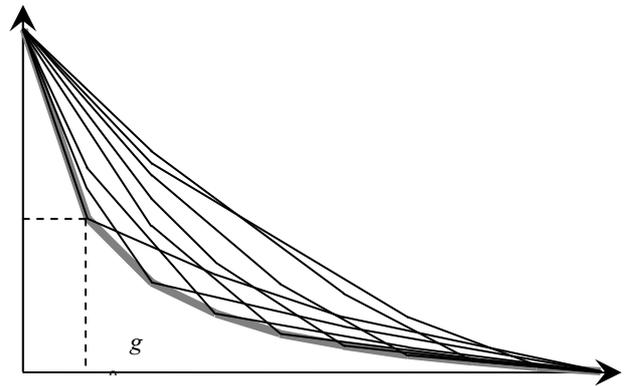
$$\Delta E_{split}(1 \cup 2) = -\Delta E_{merge}(1, 2) \leq 0.$$

В модели предусмотрено также преобразование приближения изображения посредством операции *correct* реклассификации пикселей. В этом случае, при переносе  $n_3$  пикселей кластера  $3 \subset 1$ , имеющих среднюю яркость  $I_3$ , из кластера 1 в кластер 2 приращение  $\Delta E_{correct}$  суммарной квадратичной ошибки  $E$  выражается в виде:

$$\Delta E_{correct} = \frac{n_2 n_3}{n_2 + n_3} \|I_3 - I_2\|^2 - \frac{n_1 n_3}{n_1 - n_3} \|I_3 - I_1\|^2.$$

Немаловажной особенностью дихотомической иерархии, получаемой методом Уорда или в модели Мамфорда-Шаха, является ее неоднозначность, которая связана с выраженной повторяемостью минимальных значений приращения суммарной квадратичной ошибки, особенно на начальных шагах укрупнения кластеров (или сегментов). При этом результат итеративных вычислений определяется выбором на каждом шаге очередной пары объединяемых множеств пикселей из набора равноправных вариантов и зависит от порядка их слияния, а изображение аппроксимируется многочисленными иерархическими последовательностями квазиоптимальных приближений,

которые описываются переплетающимися выпуклыми кривыми (рис. 2).



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

На рис.2 изображены выпуклые графики суммарной квадратичной ошибки  $E$  в зависимости от числа кластеров  $g$ . Из-за неоднозначной аппроксимации оптимальных приближений изображения иерархическими квазиоптимальными приближениями возникает параметризация последних значением числа кластеров  $g_0$ , при которой кривая для квазиоптимальных приближений максимально приближается к кривой для оптимальных приближений. Приближение с  $g_0$  кластерами пикселей разделяет иерархию объектов с числом кластеров от 1 до  $g_0 - 1$  от иерархии частей внутри объектов с числом кластеров от  $g_0$  до  $N$ . Тем самым изображение иерархически структурируется в двух масштабах — объектов и частей объектов.

Разделение изображения на  $g_0$  изображений объектов, задаваемых кластерами пикселей, определяется тем, что:

- 1) последовательность приближений внутри каждого объекта как самостоятельного изображения описывается выпуклой последовательностью значений суммарной квадратичной ошибки  $E$ ;
- 2) максимальное падение  $|\Delta E_{split}|$  суммарной квадратичной ошибки при разделении надвое любого объекта не превышает минимального приращения  $\Delta E_{merge}$  при слиянии любой пары объектов в один объект:  $\min \Delta E_{merge} \geq \max |\Delta E_{split}|$ ;
- 3) реклассификация любой части объекта из одного объекта в другой не повышает суммарной квадратичной ошибки  $E$ :  $\min \Delta E_{correct} \geq 0$ .

Выполнение первого условия достигается расчетом иерархии частей по Уорду и обеспечивает максимальное падение  $|\Delta E_{split}|$  суммарной квадратичной ошибки при разделении каждого объекта надвое.

Выполнение второго условия поддерживается встречным применением операций *split* и *merge* в так называемом *SI*-методе [10] и обеспечивает кластеризацию объектов при достраивании иерархии объектов до полной метод Уорда без нарушения условия выпуклости при начальном числе кластеров  $g_0$ .

Третье условие обеспечивается итеративным применением операции *correct*, обеспечивающей

максимальное падение суммарной квадратичной ошибки  $|\Delta E_{split}|$ . Указанный метод является уточненной версией метода  $K$ -средних [1,5] и его обобщением на обработку иерархически структурированных кластеров пикселей.

#### 4. Экспериментальные результаты

Рисунки 3-6 показывают результаты обработки тестовых изображений рис. 1 посредством итеративного слияния множеств пикселей.

На рис. 3-4 приведены результаты сегментации в модели сегментации Мамфорда-Шаха для результирующих изображений из  $g=2$  и  $g=7$  сегментов различных цветов.

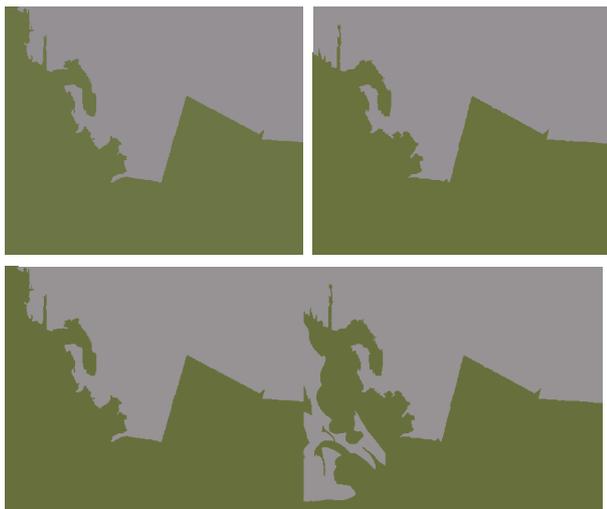


Рис. 3. Сегментация в модели Мамфорда-Шаха при  $g = 2$ .

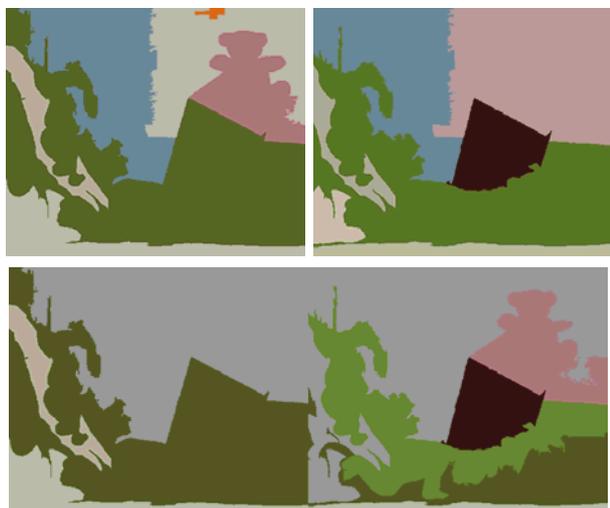


Рис. 4. Сегментация в модели Мамфорда-Шаха при  $g = 7$ .

Приближения изображений одинаковых объектов при  $g=2$  сегментах недостаточно детальны (рис. 3), а при  $g=7$  сегментах выделение объектов плохо предсказуемо (рис.4).

На рис. 5-6 приведены результаты кластеризации по Уорду для приближений с кластерами пикселей  $g_0=2$  и  $g_0=7$  различных цветов. Под приближениями указаны достигнутые значения среднеквадратичного отклонения  $\sigma$ , связанного с  $E$  соотношением  $E \approx 3N\sigma^2$ .

По сравнению с моделью сегментации Мамфорда-Шаха, метод Уорда кластеризации пикселей обеспечивает единообразное детектирование гораздо большего числа объектов. Устойчивое выделение объектов достигается при различных  $E$  на одном и том же изображении (приближения внизу на рис. 5, 6). Это проверено в экспериментах с многочисленными совмещенными изображениями стереопар.

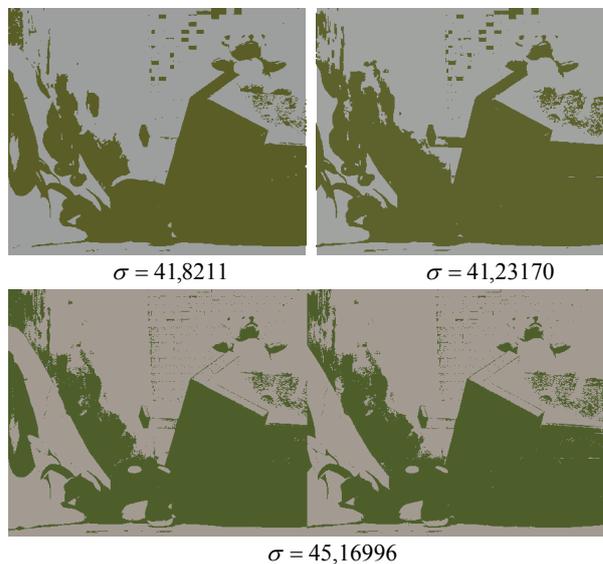


Рис. 5. Результаты кластеризации по Уорду при  $g_0 = 2$ .

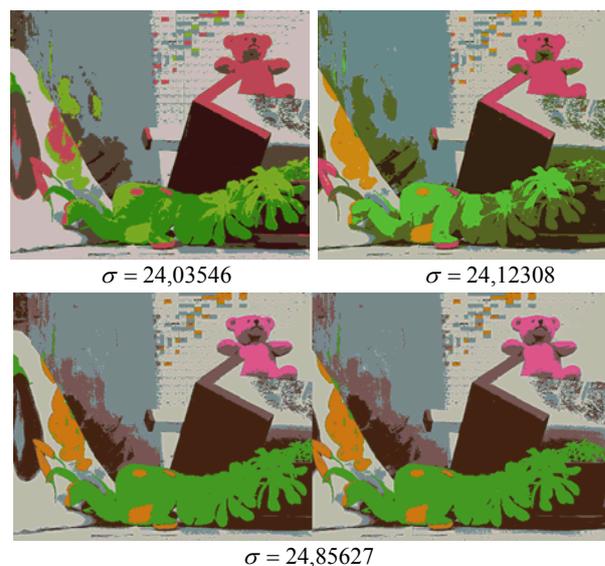


Рис. 6. Результаты кластеризации по Уорду при  $g_0 = 7$ .

Для улучшения выделения объектов на отдельных изображениях за счет получения близких по  $E$  приближений (вверху на рис. 5, 6), необходимо развитие программной реализации минимизации  $E$  по пунктам 1)-3) предыдущего раздела.

#### 5. Заключение

Судя по нашему опыту, метод Уорда обеспечивает устойчивое выделение на изображении объектов в виде близких по усредненным цветам кластеров пикселей, которые «видит» компьютер. При этом внедрение метода Уорда в практику обработки изображений приводит к двухмасштабной модели дихотомической иерархии

объектов [8-11], которая поддерживает скоростные вычисления с кластерами пикселей и сегментами изображения, а также обеспечивает эффект запоминания миллионов приближений изображения в RAM.

В качестве прототипа модели дихотомической иерархии объектов рассматривается двухмасштабная модель изображения [6,7], основными ограничениями которой являются:

- 1) на выходе вычисляется единственное разбиение исходного изображения;
- 2) объекты описываются посредством связанных сегментов, не объединяемых в кластеры;
- 3) анализ множеств пикселей выполняется в паре квадратных окошек различного масштаба.

Указанные ограничения преодолеваются в нашей модели [8-11], в которой:

- 1) на выходе вычисляются варианты иерархических последовательностей квазиоптимальных приближений с числом кластеров 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 [8-11] прозрачен и сводится к тому, что исходное изображение разбивается на самостоятельные иерархически структурированные изображения объектов классическим методом Уорда [13,16], но с учетом особенностей изображения. При этом нетривиальный формализм модели состоит в определении разбиения изображения на  $g_0$  объектов и организации расчетов в терминах специальной структуры данных динамической сети, которая «набрасывается» на пиксели для ускорения вычислений и экономии памяти [11]. Для эффективного внедрения двухмасштабной модели дихотомической иерархии объектов в практику обработки изображений, вычисления в терминах динамической сети предполагается оформить в виде стандартных утилит обработки, чтобы ими пользовались программисты-инженеры, не вдаваясь в детали математического описания.

Для современной теории и практики обработки изображений проблема выделения объектов особенно актуальна. Возможно, в обозримом будущем ее удастся решить на основе отечественных моделей [3,4], [6,7] и [8-11].

## 6. Литература

- [1] 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.
- [2] 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.
- [3] Визильтер Ю.В., Желтов С.Ю. Проблемы технического зрения в современных авиационных системах // Механика, управление и информатика, 2011. №6. — С. 11–44.
- [4] Луцив В.Р. Объектно-независимый подход к структурному анализу изображений / Дис. докт. технич. наук. СПб.: ГУАП, 2011. — 318 с.
- [5] Мандель И.Д. Кластерный анализ // М.: Финансы и статистика. 1988. — 176 с.

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

[7] Чочиа П.А. Теория и методы обработки видеoinформации на основе двухмасштабной модели изображения / Дис. докт. технич. наук. М.: ИППИ РАН, 2016. — 302 с.

[8] Харинов М.В. Разработка динамических структур данных системы автоматизированного распознавания изображений / руков. В.В. Александров / Автореф. Дис. канд. технич. наук. — С.П. 1993. — 20с.

[9] Харинов М.В., Ханьков И. Г. Применение метода Уорда для кластеризации пикселей цифрового изображения // Вестник Бурятского государственного университета.

[10] Харинов М.В. Обратимое слияние структурированных кластеров пикселей // Компьютерная графика и зрение (Графикон'2016) / Труды 26-й международной конференции по компьютерной графике и зрению: Москва, МГУ им. М. В. Ломоносова, 19 – 23 сентября 2016 г., Нижний Новгород, — С. 298–302.

[11] Харинов М.В. Основы модели квазиоптимальных приближений изображения // Вестник Бурятского государственного университета, Математика и информатика, Улан-Удэ: Изд-во Бурятского госуниверситета, 2016. №1 — С. 60-72.

[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. — pp. 1539-1597.

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

[14] 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. — pp. 35-856.

[15] Mumford D., Shah J. Boundary detection by minimizing functionals, I // Proc. IEEE Comput. Vision Patt. Recogn. Conf., San Francisco, 1985. — pp. 22–26.

[16] Ward J.H., Jr. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 1963. Vol. 58, Issue 301, — pp. 236-244.

## Об авторах

Ханьков Игорь Георгиевич, аспирант, м.н.с. Лаборатории прикладной информатики СПИИРАН Его e-mail igk@iias.spb.su.

Харинов Михаил Вячеславович, доцент, с.н.с. Лаборатории прикладной информатики СПИИРАН. Его e-mail khar@iias.spb.su.