

УДК 004.932

doi: 10.18101/2304-5728-2016-4-34-42

© *М. В. Харинов, И. Г. Ханьков*

Применение метода Уорда для кластеризации пикселей цифрового изображения

Рассматривается аппроксимация изображения дихотомической последовательностью кусочно-постоянных приближений с различным числом цветов. Иерархия приближений рассчитывается так, что цвета упорядочиваются по убыванию ошибки аппроксимации при разделении надвое множества пикселей данного цвета. Для получения приближений изображения актуального размера предлагается использовать метод Уорда. Скоростные вычисления обеспечиваются сетевой структурой данных на основе динамических деревьев Слейтора-Тарьяна.

Ключевые слова: изображение, приближение, сегментация, суммарная квадратичная ошибка, минимизация.

© *M. V. Kharinov, I. G. Khanykov*

Utilization of Ward's method for clustering of pixels of color image

The approaching of an image by a dichotomous sequence of piecewise constant approximations with incrementing number of colors is considered. The hierarchy of approximations is calculated so that the colors are ordered by decrease of approximating errors accompanying with division into two of the cluster of pixels of a given color. The network data structure for high-speed computing based on Sleator-Tarjan dynamic trees is provided.

Keywords: image, approximation, segmentation, total squared error, minimization.

Введение

Работа относится к области иерархической адаптивной сегментации цифрового изображения с целью предварительного автоматического детектирования на изображении некоторой совокупности объектов для последующего анализа признаков и распознавания. При этом детектируемые компьютером изображения объектов составляют пиксели исходного изображения, помеченные в том или ином приближении одинаковым цветом (рис. 1).



Рис. 1. Стандартное цветное изображение и его приближения с одним, двумя и тремя цветами

1. Постановка задачи

Ставится задача усреднения и упорядочения цветов по убыванию ошибки аппроксимации $|\Delta E|$, которое выражается для разделения кластера $1 \cup 2$ на кластеры 1 и 2 с числом пикселей n_1 , n_2 и трехкомпонентными средними яркостями I_1 и I_2 в виде:

$$|\Delta E| \equiv -\Delta E(1 \cup 2) = \Delta E(1, 2) = E(1 \cup 2) - E(1) - E(2) = \frac{n_1 n_2}{n_1 + n_2} \|I_1 - I_2\|^2, \quad (1)$$

где E — ошибка аппроксимации или суммарная квадратичная ошибка, $\Delta E(1 \cup 2) \leq 0$ — ее приращение при разделении кластера надвое и $\Delta E(1, 2) \geq 0$ — приращение E при обратном слиянии пары кластеров в один.

Полагается, что иерархия требуемых приближений удовлетворяет условиям сохранения порядка при дихотомическом разделении кластеров:

$$|\Delta E(1 \cup 2)| \geq |\Delta E(1)|, \quad |\Delta E(1 \cup 2)| \geq |\Delta E(2)|. \quad (2)$$

Т. е. при слиянии/разделении кластеров $|\Delta E|$ ведет себя подобно числу пикселей n в данном кластере изображения или суммарной квадратичной ошибке E . При этом условия (2) характерны для неиерархической последовательности оптимальных приближений изображения, которая аппроксимируется иерархической последовательностью квазиоптимальных приближений (рис. 1) с числом цветов от 1 до N , где N — число пикселей в изображении.

Визуально величина $|\Delta E|$ характеризует различимость смежных деталей на изображении и поэтому названа делимостью (divisibility).

2. Метод Уорда

Иерархия приближений, которая удовлетворяет условиям (2) сохранения порядка при дихотомическом разделении кластеров проще всего получается методом Уорда [1].

Метод Уорда является классическим методом кластерного анализа [2, 3], который в большинстве приложений сводится к алгоритму слияния пар кластеров с минимальным приращением ΔE суммарной квадратичной ошибки E . В качестве результата вычислений рассматривается ие-

иерархическая последовательность разбиений исходного множества элементарных кластеров. Для изображений характерна повторяемость минимальных значений ΔE на начальных итерациях слияния пикселей. Поэтому на результат вычислений влияет порядок слияния пар кластеров, и иерархическая последовательность приближений, удовлетворяющая (2), строится неоднозначно.

В обработке изображений метод Уорда применяется редко [4] из-за большой вычислительной сложности, которая квадратично возрастает с увеличением числа N пикселей изображения. Однако получение приближений изображения в упорядоченных цветах (2) можно существенно ускорить, обрабатывая его методом Уорда по частям с ограниченным числом пикселей. В этом случае обработка разделяется на три этапа.

На первом этапе формируется разбиение изображения на некоторое число g_0 кластеров пикселей (в частности сегментов изображения), обработанных по Уорду как самостоятельные изображения. При этом иерархическая последовательность приближений строится для каждого кластера. На втором, промежуточном, этапе при фиксированном g_0 выполняется оптимизация качества разбиения изображения по ошибке аппроксимации, и формируется разбиение изображения на g_0 суперпикселей (элементарных кластеров пикселей), которое характеризуется меньшим значением E . На заключительном третьем этапе выполняется кластеризация по Уорду укрупненных суперпикселей. При этом все g_0 суперпикселей сливаются в один кластер, и вычисляется полная иерархическая последовательность из N приближений изображения.

Суммарная вычислительная сложность первого и третьего этапа обработки по Уорду в зависимости от числа N пикселей изображения по порядку величины оценивается функцией $f(N)$:

$$f(N) \sim \frac{N^2}{g_0} + g_0^2, \quad (3)$$

которая достигает минимума при

$$g_0 = \sqrt[3]{\frac{N^2}{2}}. \quad (4)$$

Тогда, при выборе числа g_0 суперпикселей согласно (4) вычислительная сложность $f(N)$ выражается в виде:

$$f(N) \sim \frac{3}{2} \cdot N \cdot \sqrt[3]{2N} \quad (5)$$

и с изменением числа пикселей меняется как $N^{\frac{4}{3}}$, что позволяет обрабатывать по Уорду изображения актуальных размеров.

Если в результате иерархической кластеризации пикселей по Уорду в пределах каждого из g_0 начальных кластеров пикселей на первом этапе

обработки получается приближение изображения, для которого минимальное приращение ошибки аппроксимации $\Delta E_{merge} \equiv \Delta E(1, 2)$ при слиянии пары кластеров не превосходит максимального значения делимости кластера надвое $|\Delta E_{divide}| \equiv |\Delta E(1 \cup 2)|$:

$$\min \Delta E_{merge} \geq \max |\Delta E_{divide}|, \quad (6)$$

то второй этап генерации упорядоченных цветов опускается, и оценка (5) описывает вычислительную сложность алгоритма в целом.

Критерий (6) является критерием объединения иерархии g_0 разбиений множества суперпикселей с иерархией разбиений множества пикселей каждого из суперпикселей без нарушения упорядоченности цветов (2). Если критерий (6) нарушается для структурированных g_0 пикселей, все нарушения подавляются на промежуточном этапе формирования суперпикселей благодаря, так называемому, SI-методу, в котором итеративно выполняются встречные операции разделения надвое одного кластера пикселей при слиянии пары других [5, 6]. В отличие от скоростной программной реализации, идея SI-метода очевидна. При нарушении условия (6) выполняется разделение надвое кластера с максимальным падением ошибки аппроксимации $|\Delta E_{divide}|$ и слияние пары кластеров с ее минимальным приращением $\min \Delta E_{merge}$. Если вместо простого слияния кластеров вслед за слиянием производится обновление последовательности приближений укрупненного изображения, то эффективность минимизации ошибки аппроксимации возрастает, так как, в этом случае, максимальное падение ошибки аппроксимации $\max |\Delta E_{divide}|$ вычисляется на множестве значений, максимальных для каждого кластера.

3. Обратимые вычисления

Характерной особенностью развиваемого аппарата формирования и упорядочения цветов в изображении является применение *обратимого* слияния кластеров пикселей изображения, при котором для каждого кластера, содержащего более одного пикселя, запоминаются два кластера, слиянием которых он получен. При этом итеративное слияние пар кластеров выполняется «от пикселей» в некотором установленном порядке, модифицируемом при разделении того или иного кластера надвое. Модификация порядка слияния кластеров поддерживается автоматически. Таким образом, обратимые вычисления не сводятся к простому восстановлению данных на любом шаге [7, 8], а реализуются в обобщенном смысле, при котором оказывается возможным снижать ошибку аппроксимации и улучшать качество приближений изображения за счет комбинации операций слияния и разделения кластеров пикселей надвое.

Реализация обобщенных обратимых вычислений с произвольными кластерами пикселей, в свою очередь, требует эффективной структуры данных, основой которой являются динамические деревья Слейтора-

Тарьяна [9, 10]. В используемой структуре данных деревья (ациклические графы) дополнены циклами (циклическими графами), вместе с которыми составляют сеть, «наброшенную» на пиксели изображения (рис. 2).

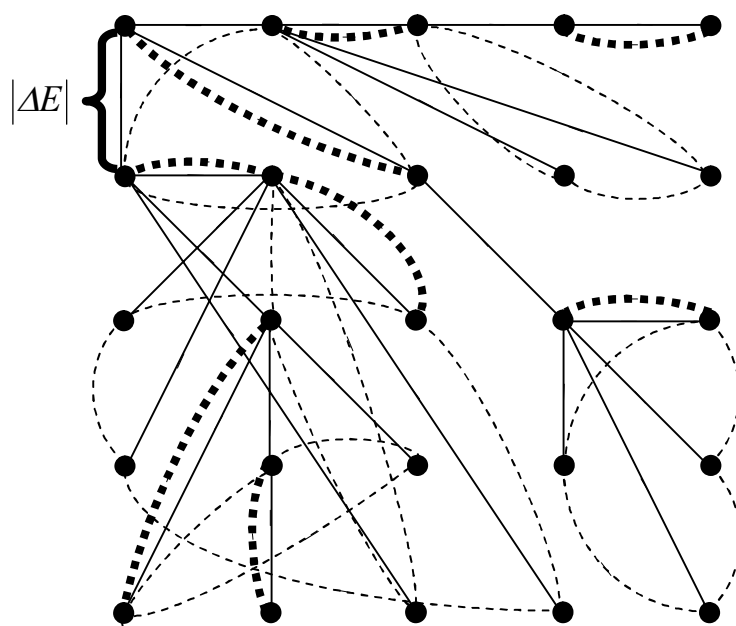


Рис. 2. Схема обратимого слияния кластеров пикселей

Рис. 2 иллюстрирует матрицу пикселей изображения, соединенных между собой дугами дерева Слейтора-Тарьяна, которые показаны сплошными линиями. Дерево имеет единственный корневой узел, который совпадает с первым пикселем изображения и служит идентификатором объединения всех пикселей изображения в единственный кластер. При разрыве дуги, инцидентной корневому узлу, дерево распадается на два дерева, и множество пикселей изображения разделяется на два кластера, которые далее рассматриваются как самостоятельные изображения. Для каждого узла дерева рис. 2 входящие дуги объединены в циклы, показанные тонкими штрихпунктирными линиями. Циклы задают порядок, в котором дуги были установлены, что обеспечивается дополнительным указанием для каждого цикла либо начального, либо конечного узла посредством указателей, обозначенных жирными пунктирными линиями. Слияние кластеров задается установлением дуг между корневыми узлами, а обратная операция разделения кластера надвое обеспечивается разрывом дуг. При обращении процесса слияния кластеров пикселей для данного корневого узла разрыв дуг выполняется в обратном порядке.

Структура данных, поясняемая схемой рис. 2, задается тремя массивами — массивом динамических деревьев, массивом циклов и массивом указателей начальных или конечных узлов в циклах. Реальная структура

данных для скоростных вычислений включает ряд дополнительных масивов, которые описываются менее сложными схемами.

Динамическое дерево Слейтора-Тарьяна и циклы рис. 2 составляют «динамическую» сеть, в которой входящие дуги для данного узла индексируются значениями убывания ошибки аппроксимации $|\Delta E_{divide}|$, сопровождающей разделение надвое задаваемого этим узлом кластера пикселей. При этом условие (2) обеспечивает получение иерархической последовательности разбиений изображения, которая описывается выпуклой последовательностью значений ошибки аппроксимации и выражается в том, что веса дуг $|\Delta E_{divide}|$ нестрого монотонно уменьшаются при обходе циклов от конечного к начальному элементу в обратном направлении и от корня к периферии дерева (рис. 2). Для получения иерархической последовательности разбиений изображения, отвечающей выпуклой последовательностью значений ошибки аппроксимации, объекты вычисляются посредством разрыва на каждом шаге дуги с максимальным значением $|\Delta E_{divide}|$.

С другой стороны, если исходное изображение разделено на g самостоятельных изображений, содержащих более одного пикселя, то имеется g вариантов построения его разбиения на $g + 1$ самостоятельных изображений. Тем самым, помимо единственной бинарной иерархической последовательности разбиений изображения с выделением объектов в порядке убывания $|\Delta E_{divide}|$, обеспечивается получение множества различных последовательностей разбиений исходного изображения на самостоятельные структурированные изображения, которые представляют объекты в различных сочетаниях.

4. Заключение

Судя по проведенным экспериментам, аппарат усреднения и упорядочения цветов обеспечивает улучшение качества традиционной сегментации изображения, а также исправление ошибочной сегментации [5]. Помимо традиционной сегментации, оказывается возможным улучшать иерархическую сегментацию изображения [6]. Конкретный практический интерес представляют эксперименты по синхронному выделению объектов на составных изображениях, скомпонованных из снимков одной и той же сцены, снятой в различном ракурсе [11], что оказывается проблематичным для большинства алгоритмов сегментации. Именно поэтому в задачах стереозрения при оценке расстояний посредством анализа стереопар вместо сопоставления объектов прибегают к поиску и сопоставлению на изображении отдельных ключевых точек [12, 13], что ограничивает решение уже на постановочном уровне. Усреднение и упорядочение цветов в изображении успешно решает проблему.

Несмотря на обнадеживающие результаты, внедрение аппарата усреднения и упорядочения цветов в практику обработки изображений требует

от программистов значительных усилий по освоению сетевой технологии вычислений на основе деревьев Слейтора-Тарьяна [9, 10], которые употребляются значительно реже традиционных деревьев, особенно в России. В отличие от традиционных деревьев, динамические деревья Слейтора-Тарьяна строятся непосредственно на множестве пикселей изображения, без задания дополнительных узлов. При этом дихотомическая иерархия кластеров пикселей задается нерегулярным деревом, но сохраняется наглядная интерпретация вычислений (рис. 2). Тем не менее, как видно на примере предыдущего раздела, перевод наглядно очевидных операций с множествами пикселей на язык программы и, тем более, воспроизведение программы по ее словесному описанию, оказывается непростой задачей. Поэтому для внедрения обсуждаемого аппарата гораздо полезнее разместить готовые эффективные программные реализации в свободном доступе. Особенностью программной реализации формирования упорядоченных цветов в приближениях изображения является многократное вычисление экстремальных значений элементов массивов модифицируемых данных, что требует трудоемкого ускорения алгоритмов рутинными и специальными приемами программирования. Можно надеяться, что по завершению этой работы аппарат усреднения и упорядочения цветов будет активно применяться для распознавания объектов на цифровых изображениях посредством готового программного обеспечения.

Литература

1. Ward J.H., Jr. Hierarchical grouping to optimize an objective function // J. Am. Stat. Assoc. — 1963. — V. 58. — Issue 301. — P. 236 – 244.
2. Мандель И.Д. Кластерный анализ. — М.: Финансы и статистика, 1988. — 176 с.
3. 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.
4. 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. — V. 5. — №. 3. — P. 35 – 856.
5. Харинов М.В., Ханыков И.Г. Оптимизация кусочно-постоянного приближения сегментированного изображения. // Труды СПИИРАН. — 2015. — Вып. 3(40). — С. 183 – 202.
6. Kharinov M.V. Reversible Image Merging for Low-level Machine Vision. URL: <http://arxiv.org/abs/1604.03832>.
7. Toffoli T. Reversible computing, In International Colloquium on Automata, Languages, and Programming,–Springer Berlin Heidelberg. — 1980. — 632 – 644.
8. Zongxiang Yan Reversible Three-Dimensional Image Segmentation. US Patent № 20110158503 A1. 2009. — 10 p.
9. Sleator D.D., Tarjan R.E. Self-Adjusting Binary Search Trees // Journal

of the ACM. 1985. Vol. 32 , № 3. — 652 –6 86.

10. Nock R., Nielsen F. Statistical Region Merging // IEEE Trans. Pattern Anal. Mach. Intell. — 2004. — V. 26(11). — 1452 – 1458.

11. Харинов М. В., Ханыков И. Г. Комбинированный метод улучшения сегментации изображения // Вестник Бурятского государственного университета. — 2015. — №9. — С. 118 – 124.

12. Малашин Р. О. Методы структурного анализа изображений трехмерных сцен. — Автореф. дис. ... канд. техн. наук. — СПб, 2014. — 22 с.

13. Фаворская М. Н., Прокурин А. В. Категоризация сцен на основе расширенных цветовых дескрипторов // Труды СПИИРАН. — 2015. — Т. 3. — №. 40. — С. 203 – 220.

References

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

2. Mandel' I.D. Klasternyj analiz. — M.: Finansy i statistika, 1988.— 176 s.

3. Ajvazjan S. A., Buhstaber V. M., Enjukov I. S., Meshalkin L. D. Prikladnaja statistika: Klassifikacija i snizhenie razmernosti. — M.: Finansy i statistika, 1989. — 607 s.

4. 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. — V. 5. — №. 3. — P. 35 – 856.

5. Harinov M.V., Hanykov I.G. Optimizacija kusochno-postojannogo priblizhenija segmentirovannogo izobrazhenija. // Trudy SPIIRAN. — 2015. — Vyp. 3(40). — S. 183 – 202.

6. Kharinov M.V. Reversible Image Merging for Low-level Machine Vision. URL: <http://arxiv.org/abs/1604.03832>.

7. Toffoli T. Reversible computing, In International Colloquium on Automata, Languages, and Programming,–Springer Berlin Heidelberg. — 1980. — 632 – 644.

8. Zongxiang Yan Reversible Three-Dimensional Image Segmentation. US Patent № 20110158503 A1. 2009. — 10 p.

9. Sleator D.D., Tarjan R.E. Self-Adjusting Binary Search Trees // Journal of the ACM. 1985. Vol. 32 , № 3. — 652 –6 86.

10. Nock R., Nielsen F. Statistical Region Merging // IEEE Trans. Pattern Anal. Mach. Intell. — 2004. — V. 26(11). — 1452 – 1458.

11. Harinov M. V., Hanykov I. G. Kombinirovannyj metod uluchshenija segmentacii izobrazhenija // Vestnik Burjatskogo gosudarstvennogo universiteta. — 2015. — №9. — S. 118 – 124.

12. Malashin R. O. Metody strukturnogo analiza izobrazhenij treh-mernyh scen. — Avtoref. dis. ... kand. tehn. nauk. — SPb, 2014. — 22 s.

13. Favorskaja M. N., Proskurin A. V. Kategorizacija scen na osnove rasshirenyh cvetovyh deskriptorov // Trudy SPIIRAN. — 2015. — Т. 3. — №. 40. — S. 203 – 220.

Харинов Михаил Вячеславович, кандидат технических наук, старший научный сотрудник лаборатории Прикладной информатики Санкт-Петербургского института информатики и автоматизации, e-mail: khar@iiias.spb.su.

Ханыков Игорь Георгиевич, младший научный сотрудник лаборатории Прикладной информатики Санкт-Петербургского института информатики и автоматизации, e-mail: igk@iiias.spb.su.

Kharinov Mikhail Vyacheslavovich, candidate of technical sciences, senior researcher of Laboratory of Applied Informatics of St. Petersburg Institute for Informatics and Automation of RAS.

Khanykov Igor Georgievich, junior researcher of Laboratory of Applied Informatics of St. Petersburg Institute for Informatics and Automation of RAS.