

Обратимое слияние структурированных кластеров пикселей*

М.В. Харинов
khar@iiias.spb.su

Россия, Санкт-Петербургский институт информатики и автоматизации Российской академии наук

Рассматривается проблема автоматического обнаружения и единообразного выделения сходных объектов на цифровом изображении посредством вычисления оптимизированных кусочно-постоянных приближений изображения, аппроксимирующих изображение с минимальной суммарной квадратичной ошибкой или среднеквадратичным отклонением приближения от изображения. Развивается модель дихотомических квазиоптимальных приближений, описываемых в зависимости от числа кластеров пикселей выпуклой последовательностью значений суммарной квадратичной ошибки. Обсуждается метод получения иерархии квазиоптимальных приближений на основе того или иного агломеративного метода кластеризации пикселей, в частности, сегментации изображения, что сводится к выполнению след за операцией слияния кластеров или сегментов алгоритма реструктуризации генерируемой иерархии кластеров для компенсации возникающих нарушений выпуклости. Предлагается версия метода, обеспечивающего при сохранении числа кластеров улучшение качества кластеризации пикселей по численному критерию и визуальному восприятию за счет комбинированного слияния/разделения кластеров, вычисляемых из условия максимального падения суммарной квадратичной ошибки.

Ключевые слова: сегментация изображения, кластеризация пикселей, кусочно-постоянное приближение, минимизация суммарной квадратичной ошибки.

Reversible merging of structured pixel clusters*

M. V. Kharinov

Russia, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences

The paper addresses the problem of automatic detection and synchronous allocation of similar objects in an image by calculating of optimized piecewise constant approximations that simulates a digital image with minimal total squared error or standard deviation of the approximation from the image. The subject of the article is the development of the model of dichotomous quasi-optimal image approximations, which correspond to a convex sequence of the total squared errors depending on the number of pixel clusters. To do this, it is proposed to generate the hierarchy of quasi-optimal approximations in some bottom-up method of pixel clustering or image segmentation by merging clusters or segments, followed by restructuring of the hierarchy of clusters to compensate the convexity violations caused by previous merging operation. The version of pixel clustering improvement both by numerical criterion and also by visual perception due to the combined merging/dividing of pixel clusters, calculated from the conditions of the maximal drop of the total squared error and the invariance of cluster number, is proposed.

Keywords: image segmentation, pixel clustering, piecewise constant approximation, total squared error minimizing.

Введение

Относительно недавно в области ИИ появилось новое практическое направление по разработке средств создания интеллектуальных приложений, что планируется выполнить к 2017 г. в рамках американского проекта PRAML (Probabilistic Programming for Advancing Machine Learning)¹. В области машинного зрения речь идет о создании программного инструментария, обеспечивающего решение прикладных задач распознавания конкретных объектов на изображении силами инженеров-программистов. Для создания доступного для программиста программного обеспечения необходима достаточно наглядная модель изобра-

жения и прозрачная интерпретация результатов его предварительной обработки.

В двухмасштабной модели [10] предлагается раскладывать изображение в сумму кусочно-гладкого приближения с протяженными объектам сцены и разностного представления с деталями объектов, текстурой и шумами. Для эффективного двухмасштабного разложения привлекается априорная информация о содержании изображения, что усложняет применение модели [10] на этапе первичного извлечения данных, для которого предназначена обсуждаемая в статье «многомасштабная» модель аппроксимации изображения иерархией кусочно-постоянных приближений [5,6].

Модель

В модели [5,6], подобно методам К-средних и Уорда [1,4,7,12], для цветового изображения решается задача минимизации среднеквадратичного отклоне-

Работа опубликована при финансовой поддержке РФФИ, грант 16-07-20482

¹<http://www.darpa.mil/program/probabilistic-programming-foradvancingmachine-learning>

ния σ приближения от изображения или суммарной квадратичной ошибки $E = 3N\sigma^2$. Для изображения из N пикселей вычисляется иерархическая последовательность из N приближений. Приближения генерируются для каждого числа кластеров от 1 до N в алгоритмах итеративного слияния и образуются $2N - 1$ различными кластерами пикселей. Операция *merge* слияния кластеров программируется как обратимая [9,11]. При этом для каждого кластера, содержащего более одного пикселя, поддерживается операция *divide* разделения кластера надвое, в результате которой он разделяется на пару тех кластеров, слиянием которых был получен. Тогда приращение суммарной квадратичной ошибки $\Delta E_{merge}(1, 2)$ при слиянии кластеров 1 и 2 совпадает с величиной $\Delta E_{divide}(1 \cup 2)$ падения E в результате разделения надвое кластера $1 \subset 2$:

$$\Delta E_{divide}(1 \cup 2) \equiv -\Delta E_{merge}(1, 2) \equiv \Delta E_{merge}(1, 2) \quad (1)$$

Для неделимых кластеров, содержащих единственный пиксель, величина падения E полагается равной нулю.

Поскольку оптимальные приближения изображения не составляют иерархическую последовательность, аппроксимирующая последовательность иерархических (дихотомических) приближений строится не однозначно (рис. 1).

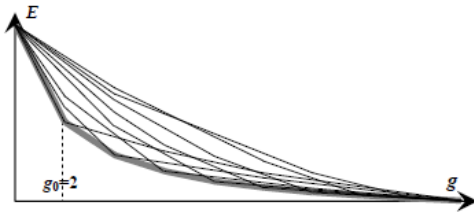


Рис. 1: Аппроксимация оптимальных приближений

Рис. 1 иллюстрирует зависимость суммарной квадратичной ошибки E от числа g кластеров пикселей в изображении.

Граничная нижняя жирная серая кривая описывает оптимальные приближения изображения. Остальные кривые отвечают целевым последовательностям приближений изображения, которые совпадают с оптимальными, по крайней мере, при трех значениях числа кластеров — минимальном $g = 1$, максимальном $g = N$ и некотором промежуточном $g = g_0$. Все кривые исходят из одной точки с координатой $g = 1$ и при $g = N$ сливаются в другую точку, отвечающей разбиению изображения на отдельные пиксели. В зависимости от контекста изображения кривые, переплетаясь друг с другом, проходят через различное число точек оптимальной кривой.

В докладе в качестве аналогов целевых иерархических приближений обсуждаются *квазиоптимальные* приближения изображения, которые удовлетворяют условиям:

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

$$\begin{aligned} |\Delta E_{divide}(1 \cup 2)| &\geq |\Delta E_{divide}(1)|, \\ |\Delta E_{divide}(1 \cup 2)| &\geq |\Delta E_{divide}(2)|, \end{aligned} \quad (3)$$

$$\min E_{merge} \geq \max E_{merge}, \quad \text{при } g = g_0. \quad (4)$$

Здесь (2) выражает свойство выпуклости последовательности значений E , которое характерно для оптимальных значений и переносится на квазиоптимальные приближения. Условие (3) является необходимым условием получения выпуклой последовательности E_1, E_2, \dots , которое состоит в том, что при разделении кластера надвое на данном шаге суммарная квадратичная ошибка приближается к нулю больше, чем на последующих шагах. В (4) требуется, чтобы максимальное падение E при разделении надвое любого из g_0 кластеров не превышало минимального приращения E при слиянии пары кластеров в один. Выполнение (4) обеспечивает для приближения с g_0 кластерами минимизацию E посредством, так называемых, *SI* (Segmentation Improvement) методов применения встречных операций разделения надвое и парного слияния для кластеризации пикселей и, в частности, сегментации изображения [5,6].

Методы и результаты эксперимента

Проще всего построить последовательность квазиоптимальных приближений изображения можно методом Уорда [12] слияния кластеров пикселей по критерию:

$$1, 2 : \Delta E_{merge}(1, 2) = \min, \quad (5)$$

согласно которому на каждом шаге выбирается пара кластеров 1,2 так, чтобы при слиянии получить минимальное приращение суммарной квадратичной ошибки E . Однако, метод Уорда имеет относительно большую вычислительную сложность и, в случае серых изображений, не учитывает геометрического распределения пикселей, т.к. сводится к гистограммной обработке. Преодолению указанных и др. ограничений способствует метод использования любого агломеративного алгоритма генерации иерархии приближений для получения приближений, описываемых выпуклой последовательностью значений E и удовлетворяющих (2), (3). Для этого достаточно в компьютерной программе, вслед за слиянием кластеров, исправлять нарушения выпуклости за счет реструктуризации укрупняемой иерархии. Алгоритм реструктуризации сводится к разделению очередного укрупненного кластера пикселей на вложенные кластеры,

до подавления нарушений выпуклости, и последующему обратному слиянию вложенных кластеров в один кластер методом Уорда.

В качестве примера исходной агломеративной кластеризации взята иерархическая сегментация изображения посредством итеративного слияния смежных сегментов, как в модели Мамфорда–Шаха [2,3,5,6,8].

Для эффективной минимизации E при установленном числе кластеров $g = g_0$ в докладе предлагается версия SI-метода, обозначаемая аббревиатурой ASI (Advanced Segmentation Improvement). ASI-метод улучшения кластеризации пикселей, в котором (4) служит критерием останова. Если (4) нарушается, то: а) разделяется надвое кластер, обеспечивающий максимальное падение E ; б) находится пара кластеров, которые сливаются между собой с минимальным приращением E , и производится реструктуризация иерархии для обеспечения (2), (3); в) процедура повторяется до выполнения (4).

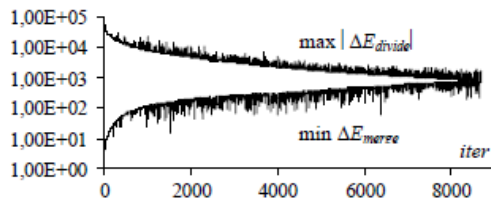


Рис. 2: Сходимость ASImax

Сходимость метода в логарифмическом масштабе по вертикальной оси иллюстрируется графиками на рис. 2, построенными для $g_0 = 2000$. Характерно, что при возрастании номера итерации $iter$ и выраженном немонотонном изменении величин $\min \Delta E_{merge}$ и $\max \Delta E_{divide}$ обеспечивается монотонное уменьшение E и σ . Дополнительные рабочие характеристики ASI-метода отражены в таблице.

В таблице для значений g_0 в диапазоне от 100 до 10000 приведено число итераций, начальное и результирующее значения среднеквадратичного отклонения σ , а также значение σ , получаемое при $g = 2$ кластеров пикселей. В паре последних строк указаны значения σ , полученные для $g = 2$ при генерации иерархии кластеров по Уорду и методом преобразования сегментации в кластеризацию посредством комбинированной операции сегментации/реструктуризации без промежуточного улучшения g_0 приближений ASI-методом.

Рабочие характеристики ASI-метода

g_0	Число итераций	σ		σ при $g=2$
		до	после	
100	275	13,06437	5,856266	28,57039
200	648	11,12987	4,657292	30,28822
300	995	10,18135	4,248172	28,18590
400	1392	9,559305	3,871545	27,86088
500	1795	9,127210	3,625355	30,51271
1000	3976	7,967459	2,963685	28,33971
1500	6281	7,350920	2,616557	28,19810
2000	8667	6,924968	2,377199	28,12330
4000	18272	5,946816	1,845133	28,20215
6000	28034	5,397584	1,576422	28,09626
8000	37286	5,017345	1,390714	27,89052
10000	46373	4,725770	1,258438	29,05978
Кластеризация по Уорду				28,06958
Сегментация/реструктуризация				27,92410



Рис. 3: Приближения изображения в 2 цветах

Сами перечисленные в таблице приближения изображения $g = 2$ кластерами пикселей показаны на рис. 3 вместе с исходным изображением, размещенным в левом верхнем углу. Рядом с ним показаны приближения, полученные без использования ASI-метода. Остальные упорядочены по возрастанию значений σ , выписанных под приближениями.

На рис. 4 в координатах $\sigma \times g$, в логарифмическом масштабе по g и диапазоне от 1 до 1000 кластеров показаны фактические результаты аппроксимации оптимальных приближений квазиоптимальными, которые, согласно таблице, описываются 14 переплетающимися кривыми (ср. с рис. 1). Верхняя кривая описывает дихотомическую сегментацию изображения посредством слияния смежных сегментов [5,6], которая дает существенно большую ошибку аппроксимации E . Рис. 5 демонстрирует преобразование дихотомической сегментации изображения в кластеризацию пикселей на примере первых приближений из 2-5 сегментов, преобразуемых в приближения изображения пикселями 2-5 цветов.

Анализируя наглядные результаты кластеризации пикселей на рис. 3 и рис. 5, можно заметить, что парные объекты (глаза, зрачки и пр.) проявляются в приближениях одновременно. Эффект не случаен и может использоваться, например, для сопоставления точек на стереоизображениях посредством предварительного сопоставления объектов, что планируется проиллюстрировать в докладе на примерах обработки тестовых изображений.

Заключение

Основное преимущество ASI-метода по сравнению, например, с методом К-средних, является то, что он эффективно справляется с улучшением грубых приближений изображения и эффективен для минимизации ошибки E при достаточно большом числе кластеров g_0 . Для минимизации E при любом наперед заданном g_0 в модели квазиоптимальных приближений планируется применить метод K-meanless [4], который является уточненной версией метода К-средних и допускает существенную дальнейшую модернизацию.

Программирование в модели квазиоптимальных приближений [5,6] предполагает использование развитой иерархической структуры данных. Особенностью практической реализации 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 модель квазиоптимальных приближений будет активно применяться для распознавания объектов на цифровых изображениях посредством готового программного обеспечения.

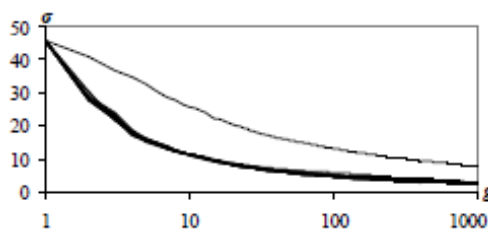


Рис. 4: Квазиоптимальная кластеризация пикселей



Рис. 5: Сегментация (слева) и кластеризация (справа)

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] 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
- [4] Dvoenko S.D. Meanless k-means as k-meanless clustering with the bi-partial approach // Pattern Recognition and Information Processing (PRIP'2014) / Proc. of the 12th Int. Conf., Minsk, 2014. – pp. 50–54.
- [5] Kharinov M.V. Pixel Clustering for Color Image Segmentation. Programming and Computer Software, 2015. Vol. 41, №. 5, – pp.258-266.
- [6] Kharinov M.V. Model of the quasi-optimal hierarchical segmentation of a color image. Journal of Optical Technology, 2015. Vol. 82, Issue 7, – pp.425-429.
- [7] Jain A.K. Data Clustering: 50 Years Beyond K-Means. Pattern Recognition Letters, 2010. Vol. 31, № 8, – pp. 651–666.
- [8] Mumford D., Shah J. Boundary detection by minimizing functionals, I // Proc. IEEE Comput. Vision Patt. Recogn. Conf., San Francisco, 1985. – pp. 22–26.
- [9] Toffoli T. Reversible computing. Springer Berlin Heidelberg, 1980. – pp. 632-644.
- [10] Чочиа П.А. Теория и методы обработки видеoinформации на основе двухмасштабной модели изображения / Дис. докт. технич. наук. М.:ИППИ РАН, 2016. – 302 с.
- [11] Zongxiang Yan. Reversible Three-Dimensional Image Segmentation. US Patent № 20110158503 A1. 2009. –10 p.
- [12] 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

Михаил Харинов — старший научный сотрудник СПИИРАН. khar@iias.spb.su.

Сравнительный анализ алгоритмов сжатия карт глубин без потерь*

А.В. Бойцов, В.Ю. Гудков, А.Г. Цыбулевский

antonboytsov@gmail.com | diana@sonda.ru | andrew.tsibulevsky@3divi.com

Челябинск, Россия, Южно-Уральский Государственный Университет

В статье представлен анализ методов сжатия без потерь статических изображений, представляющих собой карты глубин. Приведено описание основных методов сжатия данных без потерь и осуществлена модификация методов в соответствии с особенностями карт глубин. Собрана статистика по выполнению алгоритмов на тестовой выборке в различных разрешениях, определены алгоритмы, показывающие наилучшие характеристики как по коэффициенту, так и по времени сжатия.

Ключевые слова: карта глубины, сжатие без потерь, алгоритм Хаффмана, арифметическое сжатие, алгоритм Лемпеля-Зива-Велча, алгоритм RLE.

Введение

На протяжении своего развития и роста научно-технического прогресса, теория информации встречается с потребностью эффективно обрабатывать как привычные типы данных, так и информацию новой природы, например, снимки спутников и цифровых микроскопов. В настоящее время как в потребительской среде, так и в сфере исследований набирают популярность сенсоры глубины, позволяющие фиксировать трёхмерное состояние окружения. Информация о глубине объектов сцены может быть использована в различных целях как сама по себе, так и в сочетании с RGB-изображением обычной камеры. Доступ к информации о сцене предоставляется сенсором через генерацию последовательности карт глубин — изображений, яркость каждого пикселя в которых соответствует расстоянию данной точки реального пространства до камеры сенсора глубины (рис. 1). К областям приме-



Рис. 1: Карта глубины

нения карт глубин относятся выделение фона и переднего плана сцены, определение областей активности, трекинг пользователей, их поз и т. д. Также

Работа опубликована при финансовой поддержке РФФИ, грант 16-07-20482.

современные технологии позволяют использование сенсоров глубины в сочетании с технологиями виртуальной реальности, что позволяет пользователю, надевшему маску, видеть положение своего тела и взаимодействовать с объектами виртуального мира. Подобная технология подразумевает быструю передачу по проводным или беспроводным сетям карты глубины на маску либо смартфон пользователя, что, в свою очередь, касается задачи сжатия карты глубины.

Постановка задачи

Задача кодирования изображения подразумевает выбор метода кодирования. Среди алгоритмов кодирования можно выделить семейства методов сжатия без потерь и с потерями. Сжатие без потерь производится в том случае, когда восстановленная информация должна полностью соответствовать исходной. Сжатие с потерями используется, если малые потери информации (например, о заднем плане сцены) не критичны для восприятия изображения после распаковки. Оба метода кодирования активно применяются по отношению к разным типам изображений [1].

Также если говорить о кодировании последовательности изображений, то алгоритмы можно разделить на межкадровые и внутрикадровые. Межкадровое сжатие использует факт, что изображение на двух или нескольких соседних кадрах практически не меняется. В таком случае задача кодирования кадра может сводиться к кодированию разности последовательности кадров. Если метод сжатия избегает этой особенности, ему необходимо кодировать кадр целиком, что подразумевает внутрикадровое сжатие. У обоих методов есть свои преимущества — при межкадровом сжатии зачастую достигается максимальная эффективность кодирования видео, однако оно невозможно, если, например, при трансляции по сетям могут происходить потери кадров. В данной статье рассмотрены наиболее известные внутрикадровые алгоритмы сжа-