

Полииерархическая структура изображения и инвариантное детектирование объектов

М.В. Харинов¹

¹ СПб ФИЦ РАН, 14-я линия В.О., д. 39, Санкт-Петербург, 199178, Россия

Аннотация

Статья относится к междисциплинарным исследованиям в области кластерного анализа больших данных и первичного извлечения данных из цветного изображения для детектирования объектов посредством вычисления *оптимальных* кусочно-постоянных приближений изображения с реально минимизированной суммарной квадратичной ошибкой E . Сегментация изображения достигается посредством кластеризации пикселей. В качестве основного метода минимизации E рассматривается метод Уорда. Для метода Уорда устанавливается свойство *вариативности*, которое состоит в выраженной зависимости ошибки E , получаемой при данном числе цветов, от версии алгоритма вычислений или незначительной модификации входных данных. Для преодоления чрезмерной вычислительной сложности без трудоемкого программирования используется параллельное выполнение алгоритмов кластеризации пикселей с одновременным отбором иерархий приближений, достигающих минимумов E для значений числа цветов в заданном диапазоне. Исследуется задача инвариантной иерархической сегментации объектов независимо от контекста изображения. Решение обеспечивается за счет настройки управляющих параметров из условия инвариантности сегментации в контенте данного изображения. Приводятся результаты эксперимента.

Ключевые слова

Уорда кластеризация пикселей, детектирование объектов, цветное изображение, оптимальные приближения.

Polyhierarchical Image Structure and Invariant Object Detection

M.V. Kharinov¹

¹ SPC RAS, 39, 14th Line V.O., St. Petersburg, 199178

Abstract

The paper is related to interdisciplinary research in cluster analysis of big data and primary data acquisition from a color image for object detection using the *optimal* piecewise constant image approximations with a really minimized total squared error or RMSE. Image segmentation is performed via pixel clustering. Ward's clustering is considered as the main method for minimizing RMSE. For Ward's method, the *variability* property is disclosed, which consists in a pronounced dependence of RMSE obtained for a given color number on the calculation algorithm or slight modification of input data. To overcome excessive computational complexity avoiding time-consuming programming, parallel execution of pixel clustering algorithms is used with simultaneous selection of approximation hierarchies that reach RMSE minimums for the color numbers in a given range. The problem of invariant hierarchical object segmentation regardless of the image content is studied. The solution is provided by adjusting the tuning parameters from the condition of segmentation invariance in the given image content. The experiments are presented.

ГрафиКон 2023: 33-я Международная конференция по компьютерной графике и машинному зрению, 19-21 сентября 2023 г., Институт проблем управления им. В.А. Трапезникова Российской академии наук, г. Москва, Россия

EMAIL: khar@iias.spb.su (М.В. Харинов)

ORCID: 0000-0002-5166-1381 (М.В. Харинов)



© 2023 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Keywords

Ward's pixel clustering, invariant object detection, color image, optimal approximations.

1. Введение

Задача вычисления оптимальных кусочно-постоянных приближений множества многомерных векторов данных является NP-трудной ([1,2]). Поэтому она практически не может быть решена для достаточно большой входной выборки, если речь идет о произвольных данных и точном решении. Видимо поэтому, оптимальные приближения почти никто не берется считать ни для некоторых установленных подмножеств данных в кластерном анализе, ни для трехмерных векторов цветных изображений из N пикселей в области компьютерного зрения. Между тем, NP-трудность точной минимизации среднеквадратической ошибки σ или *ошибки аппроксимации* (суммарной квадратичной ошибки) $E=3N\sigma^2$ не исключает приближенного к реальным значениям решения для заведомо ограниченного множества данных, коими являются цветные, в частности, *полутонные* изображения в серых тонах.

В теоретическом кластерном анализе установлено что для решения проблемы оптимизации полезно сначала методом Уорда вычислять бинарную *иерархию* (иерархическую последовательность) начальных приближений рассматриваемого множества векторов последовательным числом 1,2,... кластеров, а затем уточнять приближения методом К-средних.

В компьютерном зрении сегментация изображения посредством кластеризации пикселей, хотя и представлена, но не обеспечивает реальной минимизации E , тогда как кластерный анализ представлен в основном методом К-средних и производных от него методов [3-7]. Кроме того, среднеквадратическая ошибка (RMSE) считается неадекватным критерием оценки качества приближения изображения, что недостаточно корректно с точки зрения кластерного анализа [8,9] обосновывается в популярной работе [10] и до сих пор служит руководством для ряда работ, например, [11].

На самом деле, скоростная сегментация, в процессе которой поддерживается связность сегментов изображения, не эффективна с точки зрения минимизации ошибки E и является вторичной, т.к. проще всего достигается посредством предварительной оптимальной кластеризации пикселей, что иллюстрирует рисунок 1.



Рисунок 1 – Оптимальная кластеризация пикселей методом Оцу (слева) и оптимизированное приближение изображения двумя связными сегментами (справа). Под представлениями указаны значения среднеквадратической ошибки σ

Рисунок 1 поясняет, что оптимальную кластеризацию проще вычислить, чем оптимальную сегментацию. Так, в случае полутонного изображения, оптимальная кластеризация пикселей в двух цветах получается методом Оцу [12], а оптимальная сегментация получается из

оптимальной кластеризации посредством соединения между собой светлых и, аналогично, темных сегментов отрезками линий шириной в один пиксель, что автоматизировать гораздо сложнее, чем вычислить оптимальное приближение методом Оцу.

Что касается метода К-средних, то этот метод, в оригинале разработанный для вычислений на арифмометрах, не обеспечивает эффективной минимизации ошибки E так как:

- работает по огрубленной формуле приращения ошибки аппроксимации для слияния и обратного разделения кластеров пикселей, которая не учитывает количество пикселей в кластерах [13];
- оперирует с отдельными пикселями, а не с множествами пикселей, реклассифицируемых из одного кластера в другой;
- опирается на вычисление промежуточных средних значений пикселей, из-за устаревания которых возникают вычислительные ошибки в виде пустых кластеров [14].

Для реальной минимизации $E=3N\sigma^2$ вместо К-средних следует применять метод К-meanless (К-средних без средних) [15], но при условии, что исходные приближения получены с ошибками E в окрестности минимальных значений, например, методом Уорда.

Реальной минимизации ошибки аппроксимации E как в теоретическом кластерном анализе [1,2,14], так и в компьютерном зрении препятствует применение стандартных деревьев (дендрограмм), применяемых для вычисления, запоминания и преобразования иерархий кусочно-постоянных приближений. Подобно промежуточным "средним" в методе К-средних, в стандартных деревьях (дендрограммах) продуцируются лишние узлы, которые не совпадают с узлами, взаимно-однозначно сопоставляемым исходным элементам данных, в частности — координатам пикселей цветного изображения, что существенно усложняет расчеты с иерархиями приближений. Незаменимой альтернативой общеупотребительным деревьям (дендрограммам) являются деревья Слейтора-Тарьяна [16], которые, в отличие от стандартных деревьев, строятся на заранее заданном множестве узлов и позволяют оперировать с произвольными множествами пикселей, как с отдельными пикселями. Хотелось бы ошибаться, но в России деревья Слейтора-Тарьяна никем, кроме нас, не освоены в задаче вычисления оптимальных приближений, хотя за рубежом довольно давно начали применяться для сегментации изображений [17]. Однако, при наличии готовых программ это не имеет решающего значения для внедрения, т.к. вычисления в терминах деревьев Слейтора-Тарьяна, не изменяют смысла вычислений, а только ускоряют их и экономят память компьютера.

Получение *оптимальных* приближений с реально минимизированной ошибкой аппроксимации E для полутоновых изображений принципиальной сложности не представляет [18]. Для цветных изображений также имеется обсуждаемый здесь простейший способ генерации оптимальных приближений.

В качестве руководства мы придерживаемся общих требований к целевой сегментации, сформулированных Г. Кёпфлером в [19]:

- наличие унифицированного алгоритма сегментации;
- наличие численного критерия выбора лучшего из нескольких вариантов сегментации;
- инвариантность сегментации относительно масштабирования.

В качестве развития условий инвариантности мы изучаем сохранение сегментации при преобразовании цветного изображения в *полутоновое* представление в серых тонах, а также ее устойчивость при изменении контента изображения.

2. Цветное изображение как полииерархическая структура

В отличие от традиционных работ, которые анализируют изображение как матрицу произвольных пикселей, изображение и объекты, состоящие из меньших объектов, мы описываем посредством последовательности приближений, которые упорядочены по числу цветов g , ошибке аппроксимации E и производной $\frac{\partial E}{\partial g}$ ошибки аппроксимации E по числу

цветов g . Объекты описываются иерархией приближений. Поскольку изображение содержит различные объекты, оно представляется как *полииерархическая* структура, которая сама по себе не является иерархической, но неоднозначно аппроксимируется иерархией приближений, соответствующей тем или иным объектам.

Обнаружение объектов на изображении классическими методами минимизации ошибки аппроксимации E в нашей версии выражается следующими положениями:

- главным предположением о специфике изображения является предположение о выпуклости последовательности ошибок аппроксимации его оптимальных приближений в $1, 2, \dots, N$ цветах, которое служит критерием соответствия входных данных компьютерной модели;
- понятия объектов, изображений и их элементов, называемых *суперпикселями*, выражаются функциями изображения и только изображения;
- поддерживается инвариантность вычислений относительно линейного преобразования числа пикселей и их значений, включая преобразование изображения в негатив. На результаты кластеризации не влияет изменение положения пикселей. Перечисленные свойства обеспечивают коммутацию оптимизации приближения изображения с его масштабированием путем дублирования пикселей;
- вычисления выполняются в соответствии с точной постановкой комбинированной аппроксимационно-оптимизационной задачи двумерного упорядочения приближений изображения для выбора целевой иерархии объектов.

Перечисленные положения иллюстрируются рисунком 2.

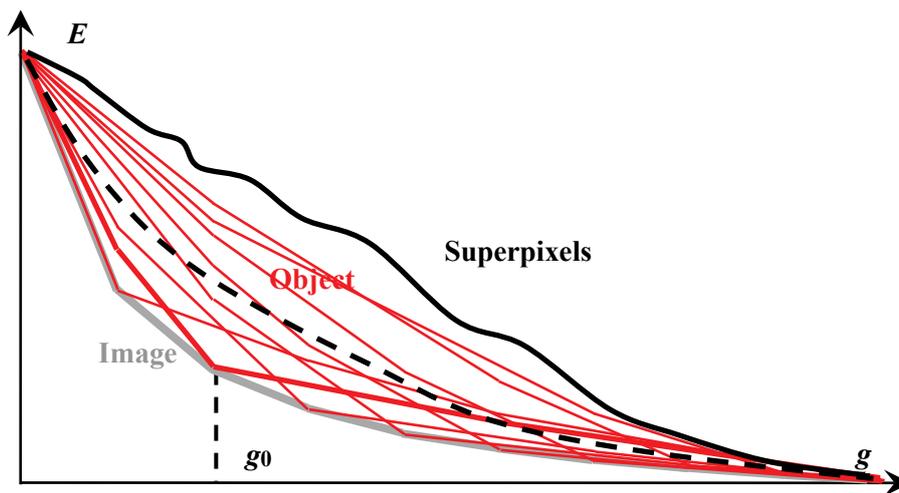


Рисунок 2 – Многочисленное решение задачи иерархического приближения изображения

Нижняя серая выпуклая кривая описывает последовательность ошибок аппроксимации E_g оптимальных приближений изображения в зависимости от числа g цветов. Верхняя черная кривая с множественными точками перегиба описывает ошибки E_g приближений изображения суперпикселями, составляющими некую нерегулярную иерархическую последовательность. Красные выпуклые кривые описывают E_g последовательности иерархических приближений изображения, каждая из которых содержит, по меньшей мере, одно оптимальное приближение в g_0 цветах, где $g_0 = 2, 3, \dots, N-1, N$ – число пикселей в изображении. Пунктиром обозначена нисходящая выпуклая кривая, отвечающая бинарной иерархии приближений, в которую можно конвертировать иерархию суперпикселей.

Рисунок 2 поясняет:

- формулировку комбинированной аппроксимационно-оптимизационной задачи;
- структурные определения изображения, объектов в изображении и суперпикселей;
- обоснование существования решения путем кластеризации пикселей оригинальным методом Уорда;
- введение настроечного параметра g_0 ;

- способ упорядочения N^2 приближений выходного изображения, которые предоставляются пользователю в результате первичной обработки изображения, где N — количество пикселей в изображении.

Множество из n пикселей считается *структурированным*, если для него построена некоторая последовательность кусочно-постоянных приближений из $1, 2, \dots, n$ кластеров, которая описывается выпуклой последовательностью суммарных квадратичных ошибок.

Объекты представляются бинарной иерархией кусочно-постоянных приближений изображения, которая содержит, по меньшей мере, одно оптимальное приближение изображения в g_0 цветах и, в зависимости от числа g цветов, описывается выпуклой последовательностью суммарных квадратичных ошибок $E_g = 3N\sigma_g^2$.

Изображение представляет собой полииерархическую структуру, которая аппроксимируется последовательностью перекрывающихся оптимальных приближений изображения и в зависимости от числа g цветов описывается выпуклой последовательностью суммарных квадратичных ошибок E_g .

Суперпиксели генерируются путем пересечения последовательных оптимальных приближений изображения в ограниченном числе цветов, посредством слияния обеспечивают безошибочное воспроизведение ряда этих оптимальных приближений и при переменном числе учтенных оптимальных приближений образуют нерегулярную иерархию, описываемую невыпуклой зависимостью E_g ошибки E от числа цветов g .

Таким образом, *изображение* трактуется как структурированное множество пикселей, *объекты* — как иерархически структурированные множества пикселей, а *суперпиксели* — как неструктурированные, но иерархически упорядоченные множества пикселей.

Любая структурированная иерархия приближений изображения, содержащая то или иное оптимальное приближение в g_0 цветах, может быть получена методом Уорда, если оптимальные приближения изображения известны. Для этого достаточно:

- вычислить вложенную иерархию кластеров пикселей в пределах каждого из g_0 кластеров заданного оптимального приближения с помощью метода Уорда;
- пересортировать порядок слияния кластеров по возрастанию приращения ошибки аппроксимации без изменения самих кластеров;
- достроить полученную иерархию до полной методом Уорда.

Существование многозначного решения аппроксимации изображения иерархиями приближений обуславливает вариативность метода Уорда и подсказывает тривиальный способ получения оптимальных приближений изображения. Хотя метод Уорда не учитывает положения пикселей, но на приближения, получаемые в ограниченном числе цветов, влияет порядок перебора пар кластеров. Они меняются даже при повороте изображения на 90° , а также или модификации изображения за счет укрупнения начальных кластеров (pixel enlargement), небольшого сжатия изображения с потерями и при других модификациях алгоритма генерации приближений. Отбор приближений с минимальными ошибками E_g сразу для всех значений числа g цветов в некотором диапазоне, скажем, от 1 до 100 обеспечивает устойчивую минимизацию ошибок аппроксимации, а параллельное выполнение используемых версий алгоритма Уорда обеспечивает ускорение вычислений в 100 и более раз. Также действенным способом ускорения вычислений методом Уорда является его выполнение по подмножествам пикселей, на которые разбиваются N пикселей изображения [20]. В этом случае, сначала методом Уорда структурируются p подмножеств пикселей, затем порядок слияния кластеров меняется так, чтобы получилась структурированная иерархия приближений в диапазоне числа g кластеров от N до p , и перед завершающим получением методом Уорда полной структурированной иерархии приближений в диапазоне g от N до 1 выполняется, так называемый, CI (Clustering Improvement) метод деления/слияния кластеров пикселей.

CI — *жадный* метод итеративного максимально возможного уменьшения ошибки аппроксимации E при сохранении числа цветов в приближении. CI метод выполняется путём деления надвое одного кластера пикселей с последующим слиянием пары кластеров. Это метод деления/слияния, который не изменяет структурированные иерархические приближения, полученные оригинальным методом Уорда. Если метод Уорда обеспечивает генерацию

приближений, описываемых выпуклой кривой во всем диапазоне числа g кластеров от N до 1, то CI метод обеспечивает локальную выпуклость при данном числе кластеров.

Характерно, что CI методом любую иерархию приближений изображения, например, иерархию суперпикселей (черная кривая на рисунке 2), можно преобразовать в структурированную (нисходящая пунктирная кривая на рисунке 2), что является интересным предметом для дальнейшего исследования.

3. Экспериментальные результаты

На рисунке 3 показаны десять изображений, посредством которых иллюстрируются результаты экспериментов. Изображения: "Zebras", "River", "Pyramids", "Girl" и "Boat" взяты из [11].

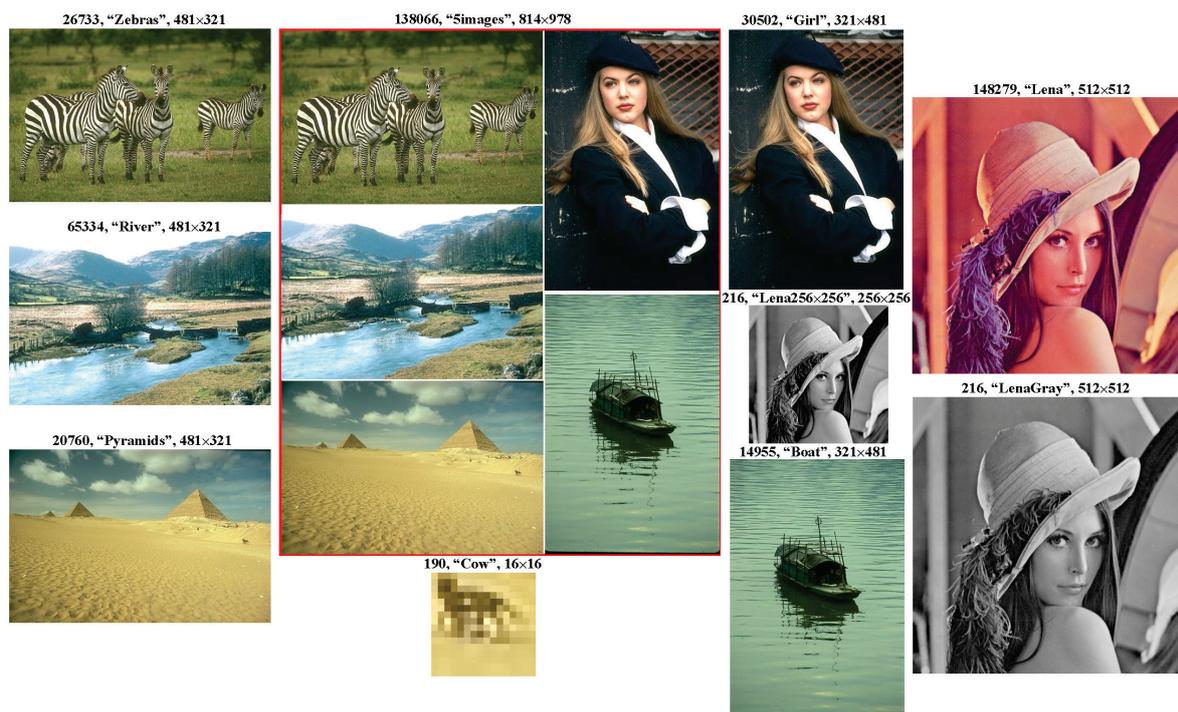


Рисунок 3 – Исходные изображения со своими характеристиками: над каждым изображением слева — число g цветов или тонов, следующее — название, справа — ширина и длина.

Изображения, слитые воедино, обведены красной рамкой. Изображения показаны в одинаковом масштабе, за исключением "Cow" (фрагмент "Pyramids"), увеличенного в 12 раз

3.1. Динамическая таблица приближений изображения

Следует отметить, что при детектировании объектов часто используют обучение, которое предполагает участие человека и не представляется необходимым при моделировании зрения более простых систем, скажем, насекомых. Поэтому оно не используется при рассматриваемом здесь *первичном упорядочении входных данных* по количеству цветов g , ошибке аппроксимации

E и параметру неоднородности $H = \left| \frac{\partial E}{\partial g} \right|$, которое трактуется как общая стадия восприятия для

различных естественных зрительных систем. Структуризация и упорядочение данных могут быть представлены в виде программы, которая автоматически генерирует, так называемую, Динамическую таблицу приближений изображения, введенную в [20].

Динамическая таблица для изображения "Girl" представлена на рисунке 4.



Рисунок 4 – Фрагмент Динамической таблицы из 23840 млн. приближений для изображения "Girl", содержащем 321×481 пикселей. Первая строка и столбец опущены. Бинарные иерархии приближений размещены по столбцам. В строках показаны приближения в $g=2,3,4$ и 5 цветах. Оптимальные приближения, помеченные красным цветом, занимают диагональ. Для каждого приближения указана пара g_0, g и среднеквадратическая ошибка σ

На рисунке 4 показан фрагмент Динамической таблицы, который иллюстрирует рисунок 2. В столбцах Динамической таблицы размещены иерархии приближений изображения. При увеличении номера строки на единицу один из цветов в текущем приближении изображения разделяется на два. На диагонали располагаются оптимальные приближения, которые вычисляются независимо от масштаба изображения и улучшаются по ошибке E при применении метода Уорда с различными значениями управляющего параметра g_0 . Параметр g_0 равен числу цветов в оптимальном приближении изображения и отсчитывается по диагонали.

В представлении пользователя вся динамическая таблица, содержащая $N \times N$ приближений изображения, размещается в ОЗУ. Фактически, она кодируется в ОЗУ в терминах специальной структуры данных [20]. При этом приближения генерируются в режиме on-line по мере необходимости. Именно поэтому таблица 1 называется Динамической таблицей.

При визуальном управлении сегментацией задача пользователя состоит в выборе столбца приближений, в котором наилучшим образом отображаются иерархически структурированные объекты интереса. В соответствии с выбором пользователя устанавливается параметр g_0 , и объекты аппроксимируются либо объединениями, либо частями кластеров пикселей оптимального приближения изображения в g_0 цветах. Например, судя по рисунку 5, в задаче идентификации человека по относительному положению глаз, уголков губ и кончика носа первая и предпоследняя иерархия приближений изображения и соответствующие параметры $g_0=2$ и $g_0=4$ являются предпочтительными.

Таблица 1 иллюстрирует Динамическую таблицу для изображения “5images”, в которой сами приближения изображения “5images” заменены на значения среднеквадратической ошибки σ (RMSE в англоязычных статьях).

Таблица 1 – Динамическая таблица значений σ для изображения "5images"

$g_0 \backslash g$	2	3	4	5	6	7	8	9	10
1	62,11036	62,11036	62,11036	62,11036	62,11036	62,11036	62,11036	62,11036	62,11036
2	38,38280	38,72923	38,52619	38,64821	38,78163	41,09046	39,43764	42,30092	42,37132
3	30,42113	29,05928	30,72957	30,80608	29,28113	29,90124	30,20335	31,37998	31,22463
4	23,97177	24,59723	23,25888	23,34421	24,57672	24,17970	25,14314	24,15173	24,49769
5	21,50731	21,51242	20,45330	20,30430	21,17090	21,98895	21,62782	21,52593	21,84682
6	19,62089	19,61104	19,34542	18,87722	18,71443	19,63935	18,91667	19,85759	20,14883
7	18,49336	17,90441	18,26962	17,78181	17,33875	17,26866	17,33373	18,12586	18,37371
8	17,38270	16,78000	17,13836	16,79245	16,40023	16,38050	16,32671	16,69588	16,47044
9	16,20349	15,66579	16,01561	15,77196	15,69258	15,63975	15,61573	15,33083	15,33397
10	14,98624	14,87754	15,10671	14,95546	15,0303	14,94219	14,98109	14,75379	14,49886

Последовательности значений σ^2 для столбцов таблицы — выпуклые. В строках таблицы перечислены значения σ для приближений в 1,2,...,10 цветах. Диагональ занимают минимальные по строкам значения σ (выделены жирным шрифтом). На первом месте таблицы может располагаться любой из остальных столбцов. Поэтому первый столбец опущен. Квадраты диагональных значений $\sigma^2=62,11036^2$, $\sigma^2=38,38280^2$, $\sigma^2=29,05928^2$ и т.д. составляют выпуклую последовательность.

3.2. Инвариантность оптимизации приближения цветного изображения при масштабировании и преобразовании в полутоновое представление

Продемонстрируем, как работают условия инвариантности при масштабировании и преобразовании цветного изображения в полутоновое представление на примере стандартного цветного изображения "Lena". Что касается инвариантности сегментации относительно изменения масштаба изображения, то она обосновывается логикой алгоритмов оптимизации. С другой стороны, согласованность оптимальных приближений изображения в цветном и полутоновом представлениях естественно ожидать в предположении адекватного моделирования зрительного восприятия, так как оба представления воспринимаются как изображения одних и тех же объектов.

На рисунке 5 показаны оптимальные приближения для цветного изображения “Лена” и двух его полутоновых представлений в оригинальном и уменьшенном виде, — “LenaGray” и “Lena256×256”, соответственно.

Представления изображения в оттенках серого в 2, 4 и 5 тонах из центрального ряда выглядят так, как будто они получены путем преобразования в полутоновые оптимальных цветных представлений изображения в 2, 4 и 5 цветах верхнего ряда. В оптимальных приближениях в трех тонах некоторые светлые блики по-разному проявляются в цветном и полутоновом представлениях. Последнее связано именно с преобразованием из цветного представления в полутоновое, так как при масштабировании оптимальные приближения заметно не меняются (сравните центральный и нижний ряд полутоновых представлений на рисунке 5).



Рисунок 5 – Сравнение оптимальных приближений для цветного изображения (верхний ряд), его полутонового представления (средний ряд) и уменьшенной копии представления. Над каждым изображением указаны: число g_0 цветов или тонов и среднеквадратическая ошибка σ

3.3. Инвариантность объектов на оптимальных приближениях цветного изображения

Рисунок 6 иллюстрирует инвариантность объектов на оптимальных приближениях изображения "5images".

На рисунке 6 в красных рамках показаны оптимальные приближения композиционного изображения "5images" в 2-7 цветах. Рядом с ними демонстрируются оптимальные приближения для компонент — изображений "Zebras", "River", "Pyramids" (слева) и "Girl", "Boat" (справа), полученные для них, как самостоятельных изображений. Цвета на различных изображениях рисунка 6 не повторяются. Так, для воспроизведения шестерки приближений в $g_0=7$ цветах в нижнем правом углу рисунка 6 требуется 42 цвета.

Палитра 2-7 цветов в оптимальных приближениях композиционного изображения "5images" относительно палитры цветов остальных пяти изображений смещается в сторону серых тонов. Тем не менее, нетрудно заметить, что при слиянии самостоятельных изображений в единое композиционное изображение наблюдается выраженная тенденция к сохранению оптимальной сегментации. Во всяком случае, многие объекты, на оптимальных приближениях изображения "5images" занимают близкие по размерам и форме сегменты, а также разделяются на близкие, если не на те же самые, вложенные сегменты, что и на самостоятельных изображениях. При этом, скажем, сегментация изображения "Girl" лучше других сохраняется при $g_0=2$ и $g_0=4$ значениях числа цветов в оптимальных приближениях.

Для того, чтобы лучше представить влияние контекста изображения на его сегментацию посредством оптимальной кластеризации, имеет смысл рассмотреть детектирование довольно сложного объекта "Cow" минимального размера на изображениях "Pyramids" и "5images" (рисунок 7).



Рисунок 6 – Оптимальные приближения изображения "5images" в 2-7 цветах, размещенные в лексикографическом порядке в двух строках. Над приближениями указано число цветов g_0

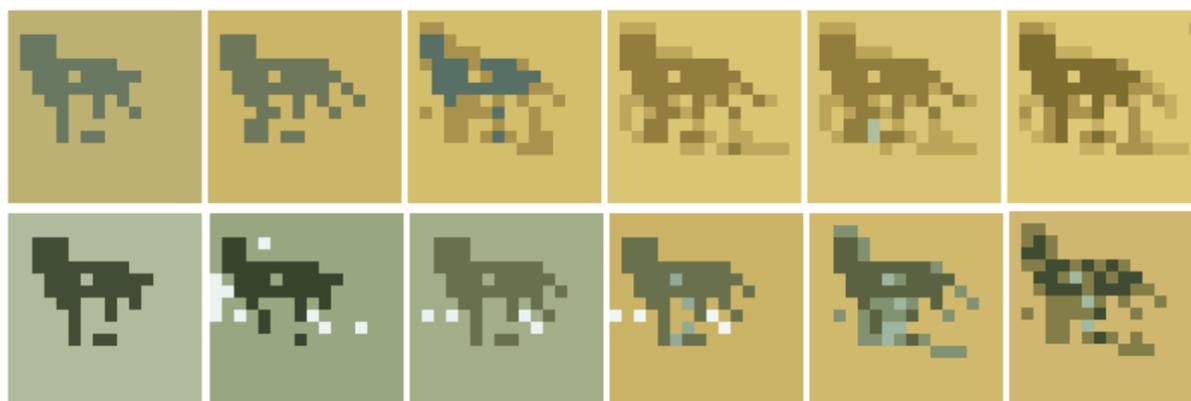


Рисунок 7 – Приближения в 2-7 цветах фрагмента "Cow" самостоятельного изображения "Pyramids" (вверху) и в составе композиционного изображения "5images" (внизу)

Рисунок 7 иллюстрирует детектирование объекта "Cow" в контексте изображения "Pyramids" и в общем контексте изображения "5images". Очевидно, в данном случае, лучше всего сегментация сохраняется при числе цветов $g_0=2$. Дополнительным вариантом является сегментация при значении настроечного параметра $g_0=4$.

Поскольку параметр g_0 определяет вариант иерархической кластеризации пикселей (столбец приближений в Динамической таблице рисунка 4), условия инвариантного детектирования объектов могут использоваться для автоматической настройки программы на тот или иной объект интереса, если образцы объектов интереса присоединяются к входному изображению и сегментируются как часть композитного входного изображения [21]. При этом специфика обработки состоит в необходимости вычисления одного или нескольких кластеров пикселей, отвечающих образцу объекта интереса (рисунок 7), а сегменты из пикселей того же кластера или аналогичным образом сгруппированные пиксели нескольких кластеров на входном изображении детектируются как новые объекты искомого типа.

4. Заключение

В статье мы постарались на конкретных примерах продемонстрировать возможности сегментации полноразмерных изображений через оптимальную кластеризацию пикселей.

Представленные приемы обработки изображений разрабатывались с 80-х годов. Они согласуются со сравнимыми по длительности разработки результатами отечественных исследований [22,23]. По сравнению с более общими подходами [22,23] мы рассматриваем только первичный этап структурирования данных для обнаружения объектов на изображении, обеспечивающий в терминологии [23] "приведение изображения к виду, удобному для дальнейшего распознавания". При этом параметр "сложности изображения", введенный в [22], отождествляем с неоднородностью H множества пикселей. В отличие от подходов [22,23] мы полагаем что при первичной структуризации данных на входе принимается изображение s , возможно, прикрепленными или встроенными в него образцами объектов интереса, а на выходе получается двумерная Динамическая таблица из N^2 приближений изображения или, при автоматической настройке на объекты интереса, — конкретная иерархия из N приближений изображения.

По всей видимости, вычисление оптимальных приближений изображения и анализ свойств инвариантности объектов окажется полезным не только для традиционного детектирования объектов, но также и для постановки и решения новых задач, например, настройки съемки для наилучшего отображения объектов интереса, автоматизации художественной фотографии, расшифровки смысла картин художников и др.

Для дальнейшего изучения свойств оптимальных приближений имеет смысл добавить готовые оптимальные приближения к той или иной базе изображений, например, к Berkeley Segmentation Data Set, а для обеспечения доступности вычисления оптимальных приближений программистам полезно модернизировать методы минимизации ошибки аппроксимации в составе общеупотребительных инструментариев, например, в Матлаб.

5. Благодарности

Исследование выполнено при поддержке государственного бюджета, номер проекта FFZF-2022-0006.

6. Список источников

- [1] Murtagh, F., Pierre, L. Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? // J. Classif. 2014. № 31, С. 274–295.
- [2] Torrente, A., Romo, J. Initializing k-means Clustering by Bootstrap and Data Depth // J. Classif., 2021. № 38, С. 232–256.
- [3] A Novel Type-2 Fuzzy C-Means Clustering for Brain MR Image Segmentation / P.K. Mishro, S. Agrawal, R. Panda, A. Abraham // IEEE Transactions on Cybernetics. 2021. № 51(8), P. 3901–3912, DOI: 10.1109/TCYB.2020.2994235.
- [4] Bora, D.J., Gupta, A.K. Clustering approach towards image segmentation: an analytical study // arXiv preprint arXiv:1407.8121 2014.

- [5] Fuzzy c-means clustering with spatial information for image segmentation / Chuang, K.S., Tzeng, H.L., Chen, S., Wu, J., Chen, T.J. // *Computerized Medical Imaging and Graphics*, 2006. № 30(1), P. 9–15, DOI: 10.1016/j.compmedimag.2005.10.001.
- [6] Pappas, T.N., Jayant, N.S. An adaptive clustering algorithm for image segmentation // *International Conference on Acoustics, Speech, and Signal Processing*. 1989. № 3, P. 1667–1670, DOI: 10.1109/ICASSP.1989.266767.
- [7] Конушин А.С. Компьютерное зрение [Электронный ресурс]: Конспект лекций ВМК МГУ, 11 лекций, <https://teach-in.ru/file/synopsis/pdf/computer-vision-M.pdf> (дата обращения 30.06.2023).
- [8] Прикладная статистика: классификация и снижение размерности / Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. М: Финансы и статистика. 1989. 607 с.
- [9] Мандель И. Д. Кластерный анализ. М: Финансы и статистика. 1988. 176 с.
- [10] Image quality assessment: from error visibility to structural similarity /Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P. // *IEEE transactions on image processing*. 2004. № 13(4), P. 600–612.
- [11] Efficient Approach to Color Image Segmentation Based on Multilevel Thresholding Using EMO Algorithm by Considering Spatial Contextual Information / Rangu, S., Veramalla, R., Salkuti, S.R., Kalagadda, B. // *J. Imaging*. 2023. № 9, 74, DOI: 10.3390/jimaging9040074
- [12] Liao Ping-Sung, Chen Tse-Sheng, Chung Pau-Choo Algorithm for Multilevel Thresholding // *J. Inf. Sci. Eng*. 2001. № 17, P. 713–727.
- [13] Kharinov, M.V. Reclassification formula that provides to surpass K-means method // *arXiv preprint arXiv:1209.01940*.
- [14] On strategies to fix degenerate k-means solutions /Aloise, D., Damasceno, N.C., Mladenovic, N., and Pinheiro, D.N. // *Journal of Classification*. 2017. № 34, P. 165–190, DOI: 10.1007/s00357-017-9231-0
- [15] Dvoenko, S.D. Meanless k-means as k-meanless clustering with the bi-partial approach // *12th International Conference on Pattern Recognition and Image Processing (PRIP'2014)*, UIIP NASB, Minsk, Belarus, 24–27 September 2014. P. 50–54.
- [16] Sleator, D.D., Tarjan, R.E. Self-Adjusting Binary Search Trees // *Journal of the ACM*. 1985. № 32(3), P. 652–686.
- [17] Nock, R., Nielsen, F. Statistical Region Merging // *IEEE Trans. Pattern Anal. Mach. Intell*. 2004. № 26(11), P. 1452–1458.
- [18] Kharinov, M.V. Image Segmentation Method by Merging and Correction of Sets of Pixels // *Pat. Recog. and Image Analysis: Advances in Mathematical Theory and Applications*. 2013. № 23(3), P. 393–401, DOI:10.1134/s1054661813030061.
- [19] Koepfler, G. Segmentation by minimizing functionals and the merging methods // *World Congress of Nonlinear Analysts'92: Proceedings of the First World Congress of Nonlinear Analysts, 1933–1942, Tampa, Florida, August P. 19–26, 1992, Berlin, Boston: De Gruyter , 1996. DOI:10.1515/9783110883237.1933.*
- [20] Nenashev, V.,A., Khanykov, I.,G., Kharinov, M.,V. A Model of Pixel and Superpixel Clustering for Object Detection // *J. Imaging*. 2022. № 8(10), 274.
- [21] Kharinov, M. Example-Based Object Detection in the Attached Image // *Graphicon-Conference on Computer Graphics and Vision*. 2022. № 32, P. 490–501, https://www.graphicon.ru/html/2022/papers/paper_050.pdf (дата обращения 30.06.2023).
- [22] Чочиа П.А. История исследований в лаборатории обработки изображений ИППИ РАН // *Информационные процессы*. Том 23, № 1, 2023. P. 11–112.
- [23] Gurevich, I., Yashina, V. Basic models of descriptive image analysis // *ICPR International Workshops and Challenges: Virtual Event, January 10–15, 2021. Proceedings*, Cham: Springer International Publishing, Part V, P. 275–288.