

О возможности адаптации алгоритмов объектного анализа цветных изображений трехмерных сцен к параллельной архитектуре

Д. П. Николаев, П. П. Николаев

В работе рассматривается возможность переноса разработанных ранее эффективных алгоритмов цветовой сегментации и константной оценки окраски на параллельную архитектуру. Оцениваются возможная степень параллелизма и выигрыш по времени работы.

Введение

Научное направление, обычно называемое «проблемы технического зрения», за полвека своего развития превратилось в фундаментальную дисциплину со своей структурой, концепциями, приоритетными проблемами и достижениями. В настоящее время развитие этой дисциплины идет в направлении построения общей теории когнитивного визуального процесса и разработки алгоритмов автономного решения различных некорректных задач распознавания зрительных образов. В частности, продолжают попытки построения алгоритмов быстрого и релевантного поиска объектов в базах данных изображений и в видеорядах. При этом специфика двумерной регистрации исходно трехмерного образа (при которой теряется большая часть геометрической информации) делает предпочтительным цветовой (а не пространственный) анализ изображения. Задача инвариантного цветового анализа в техническом зрении фактически состоит в повторении феномена, известного для биологических систем зрения под названием «цветовая константность». Прозрение человека (и большинства животных) твердо установлена его способность различать тела не по их колориметрическому признаку (собственно — «цвету»), а по тому, что является объективным их признаком — по окраске, которая с цветом отраженных телом излучений при цветном его освещении не коррелирует. Поэтому не удивительно, что современные алгоритмы технического зрения зачастую используют те же цветовые инварианты и опираются на тот же набор априорных ограничений, что и зрительная система человека [1].

Однако, на этом пути есть существенные препятствия. Естественный зрительный интеллект пока может быть исследован практически только через феноменологические свои проявления. Поэтому прямой «бионический» подход пока значимых результатов не дал. В области искусственного зрительного интеллекта также еще не предложены сколь угодно полные схемы когнитивного анализа изображений. Следствием недостаточности знаний о работе мозга и оказываются современные постановки в области автоматического объектного распознавания с присутствиями им упрощениями. При разработке подобных схем цветового анализа, именно в поиске «еще относительно простого из уже феноменологически содержательного», авторами было предложено опереться на линейную теорию формирования спектрального стимула [2], выводящую достаточно широкий круг физических явлений в наблюдаемой сцене на уровень формально разрешимых задач. При этом объектный анализ изображения становится возможным разделить на этапы цветовой сегментации и константной оценки окраски. В такой постановке блок цветовой сегментации выполняет функции предобработки, задавая карту границ разноокрашенных объектов для последующих более сложных вычислений, связанных с формированием поправки на цвет освещения этих объектов в регистрируемой сенсором сцене. Линейные свойства кластера, проецирующие в цветовое пространство (ЦП) сенсора спектральные свойства объекта, лежат в основе детектирования объекта на этапе его цветовой сегментации. Они же, как показано рядом авторских работ [3], дают материал для формирования так называемых «ключей константности», как априорных признаков цветности освещения сцены. Ахроматические, глянцевые или складчатые объекты — с их особой геометрией представительства в ЦП — позволяют при их правильной идентификации строить правдоподобные гипотезы о цветности источников освещения. Причем в алгоритмах технического зрения имеется возможность опереться на точные законы аффинной геометрии ЦП, которые для биологических систем выполняются только в особых, колориметрических, условиях (что связано с наличием многочисленных механизмов адаптации). В качестве примеров можно назвать использование свойств ахроматической оси ЦП, однозначную и постоянную связь спектрального и вектор-стимула, применимость в процедурах обработки не только проективных, но и аффинных инвариантов ЦП, и пр.

Современное состояние вычислительных технологий диктует новое требование для систем технического зрения: максимально возможное распараллеливание всех необходимых процедур, позволяющее не только ускорить вычисления, но и отказаться от использования неэффективных универсальных центральных процессоров. По счастью, алгоритмы цветоконстантности — от простейших схем «нормализации» цвета до процедур

с разного типа «ключами константности» (например, алгоритм «аппликация» с его поиском максимального ответа в каждом цветовом канале) в подавляющем своем большинстве являются схемами однородной обработки. То же самое можно сказать и о сегментационных процедурах, предвещающих работу алгоритмов цветовой константности. Однако, это всего лишь означает принципиальную возможность параллельной реализации. Покажем, как можно практически достичь требуемого результата.

Для случая сенсора с известными параметрами (откалиброванного) ранее нами был разработан алгоритм автоматической сегментации изображений, использующий физическую модель формирования изображений [4]. Информация о геометрических особенностях изображений объектов учитывается алгоритмом в минимальной степени, что обеспечивает высокое качество сегментации для объектов произвольной сложной формы. Алгоритм состоит из двух этапов — предварительной сегментации и этапа слияния областей.

Предварительная сегментация

Задача предварительной сегментации — разделить изображение на области площадью хотя бы в несколько пикселей так, чтобы истинные границы объектов не пересекали этих областей. В работе используется обобщенный на случай цветных изображений алгоритм поиска водоразделов [5]. Основная посылка алгоритма поиска водоразделов состоит в том, что каждый сегмент должен содержать точно один минимум значения некоторого детектора границ, и градиентный спуск от любого пикселя сегмента по карте детектора границ должен приводить к этому минимуму. Это условие реализуется проведением градиентного спуска из каждой точки изображения и группированием точек по их принадлежности к минимумам. В результате границы сегментов (групп пикселей) проходят вдоль «хребтов» значений детектора краев. Поскольку классический оператор градиента может быть применен только для скалярного, но не для векторного поля, в цветовой версии алгоритма используется векторный аналог градиента (так называемый цветовой градиент Ди Зензо [6]).

Алгоритм санного спуска позволяет глубокое распараллеливание: возможно разбиение изображения на произвольные области с последующей независимой параллельной обработкой. Для вычисления поля градиентов в каждой области достаточно, чтобы области перекрывались на 1 пиксель по контуру. После сегментации в каждой области будет два типа сегментов: бассейны, целиком лежащие в одной области, и бассейны, частично принадлежащие другим областям. В параллельном

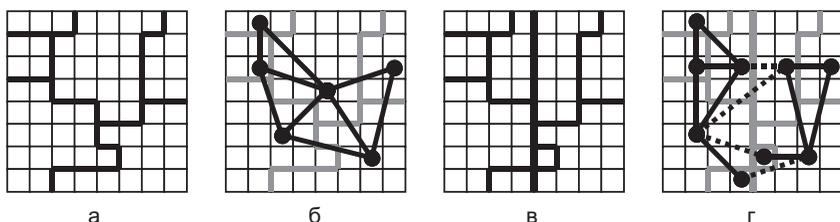


Рис. 1. Сегментация водоразделами вблизи границы области разделения. а) карта сегментации, получающаяся без разбиения на области; б) соответствующий граф соседства областей; в) карта сегментации на границе области разделения (граница выделена толстой линией); г) соответствующий граф соседства областей (пунктиром показаны ребра с весом 0)

варианте алгоритма не обязательно производить немедленную релаксацию таких бассейнов, достаточно проследить, чтобы вес ложных ребер (соединяющих бассейны второго типа через границы областей разбиения) на втором этапе был равен 0 (см. рис. 1).

Сегментация второго уровня

На втором этапе окончательная сегментация производится слиянием областей с помощью графа соседства областей (ГСО). Каждая вершина ГСО представляет сегмент изображения. Две вершины связаны ребром тогда и только тогда, когда соответствующие сегменты имеют общую границу (см. рис. 1). Каждому ребру приписывается вес, характеризующий вероятность того, что эти сегменты имеют несовпадающие параметры модели. Граф последовательно модифицируется слиянием пар вершин, соединенных ребром с наименьшим весом. Слияние производится до тех пор, пока наилучшее ребро не будет иметь вес выше некоторого порога. Качество сегментации, очевидно, определяется выбором весовой функции и уровнем порога.

В предложенном алгоритме используется семейство весовых функций, соответствующих моделям кластеров различных рангов (согласно определению в [4]). В общем виде эти весовые функции записываются следующим образом:

$$d_j[k, l] = \sqrt{\frac{1}{n_l} \sum_{i=1}^{n_k} \rho_j^2(\vec{p}_{k,i}) + \frac{1}{n_k} \sum_{i=1}^{n_l} \rho_j^2(\vec{p}_{l,i})},$$

где $d_j[k, l]$ — вес ребра, связывающего узлы, соответствующие сегментам k и l , j — ранг модельного распределения, $\rho_j(\vec{p})$ — отклонение

пикселей \vec{p} от идеальной модели для области, объединяющей сегменты k и l , $\vec{p}_{k,i}$ — пиксель i сегмента k , а n_k — количество пикселей в сегменте k .

Собственно алгоритм включает следующие этапы:

1. Применение техники слияния областей с весовой функцией $d_0[k, l]$ и порогом σ_0 , соответствующим по порядку уровню шума на изображении.
2. Маркирование изолированных сегментов ранга 0, то есть тех, для которых минимальное d_{\min} для всех прилегающих ребер сильно больше порога: $d_{\min} > \sigma_G$, где $\sigma_G \gg \sigma_0$.
3. Применение техники слияния областей с весовой функцией $d_1[k, l]$ и порогом σ_0 , с исключением из рассмотрения ребер, ведущих к изолированным сегментам.
4. Применение техники слияния областей с весовой функцией $d_1[k, l]$ и порогом σ_0 , с исключением из рассмотрения ребер, соединяющих два изолированных сегмента.
5. Дополнительное маркирование изолированных сегментов ранга 1 с порогом σ_G .
6. Применение техники слияния областей с весовой функцией $d_2[k, l]$ и порогом σ_0 , с исключением из рассмотрения ребер, ведущих к изолированным сегментам.
7. Применение техники слияния областей с весовой функцией $d_2[k, l]$ и порогом σ_0 , с исключением из рассмотрения ребер, соединяющих два изолированных сегмента.

Ранее было показано, что подобная схема при правильно выбранных порогах σ_0 и σ_G позволяет получить верную сегментацию, не решая задачу определения ранга конкретного сегмента изображения.

Алгоритм слияния областей позволяет параллельную реализацию: ГСО разбивается на подграфы, каждый из которых обрабатывается почти независимо. Каждый процессор релаксирует лучшее ребро в своем подграфе до тех пор, пока ребро не окажется «несобственным» (связывающим данный подграф с соседним). В случае, если в процессе релаксации некоторое ребро было переоценено, соседнему процессору отправляется соответствующее сообщение. Несобственное ребро не удаляется до тех пор, пока оно не будет переоценено соседним процессором или не станет также лучшим и в соседнем подграфе. В последнем случае образовавшаяся при релаксации вершина заносится в меньший из подграфов. Порядок слияний в таком алгоритме может (несущественно) отличаться от исходного последовательного алгоритма. Однако, как показывают эксперименты, это практически не влияет на результат.

Обобщенное преобразование Хафа

В ходе работ было проведено сравнение точности алгоритмов определения параметров цветовых распределений сегментов — обобщенного преобразования Хафа (ОПХ) и метода собственных векторов. В качестве оптимального выбора предлагается использовать преобразование Хафа, если это позволяют вычислительные мощности системы.

Метод наименьших квадратов решает задачу аппроксимации формы одного сегмента оптимально в квадратичной метрике при наличии нормального шума:

$$\int_{A_i} \left\| \vec{c}(\vec{r}) - \left(g_{i,0} \cdot \vec{c}_{i,0} + \sum_{j=1}^R g_{i,j}(\vec{r}) \cdot \vec{c}_{i,j} \right) \right\|^2 d\vec{r} \equiv \int_{A_i} \rho_R^2(\vec{r}) d\vec{r} \rightarrow \min_{\vec{c}_{i,j}, g_{i,j}}$$

где $(\vec{c}_{i,j} \cdot \vec{c}_{i,k}) = \delta_{j,k}$, а R — ранг линейного подмногочисления. В свою очередь, метод ОПХ решает ту же задачу, оптимизируя другой функционал, что обеспечивает методу устойчивость по отношению к импульсному (выбросовому) шуму:

$$\int_{A_i} \frac{\exp \{ -\rho_R^2(\vec{r}) / (2\sigma^2) \}}{\sqrt{2\pi}\sigma} d\vec{r} \rightarrow \max_{\vec{c}_{i,j}, g_{i,j}}$$

где σ — параметр сглаживания.

Кроме того, предложена простая быстрая схема вычисления преобразования Хафа ($O(n^2 \log n)$ операций для двумерного случая), не использующая быстрого преобразование Фурье (БПФ). Отказ от преобразования Фурье позволяет производить все вычисления в целых числах и не использовать операцию умножения.

В массиве с линейными размерами $n \times n = 2^D \times 2^D$ суммирование вдоль прямой, соединяющей точки (x, y) и $(x + \text{shift}, y + 2^{\text{deg}} - 1)$ в горизонтальную полосу $n \times 2^{\text{deg}}$ ($0 < \text{deg} \leq D$) с вертикальным смещением $y = k2^{\text{deg}}$ ($0 \leq k < 2^{D-\text{deg}}$), проводится по рекуррентной формуле

$$\begin{aligned} [(x, y), (x + \text{shift}, y + 2^{\text{deg}} - 1)] &= \left[(x, y), \left(x + \left\lfloor \frac{\text{shift}}{2} \right\rfloor, y + 2^{\text{deg}-1} - 1 \right) \right] \cup \\ &\cup \left[\left(x + \left\lfloor \frac{\text{shift}}{2} \right\rfloor, y + 2^{\text{deg}-1} \right), (x + \text{shift}, y + 2^{\text{deg}} - 1) \right], \end{aligned}$$

что и обеспечивает (при порядке суммирования по deg от 0 до $D - 1$) увеличение быстродействия за счет отсутствия повторного суммирования для фрагментов дискретного представления прямых, входящих в несколько сумм.

Топология вычислений в предложенном алгоритме аналогична топологии БПФ: сеть вычислений можно представить в виде последовательности из $O(\log n)$ слоев, в каждом из которых независимо вычисляется $O(n^2)$ сумм [7]. Это позволяет без труда распараллеливать алгоритм на практически произвольное число процессоров.

Случай некалиброванного сенсора

Для случая нелинейного сенсора с неизвестными характеристиками предложен метод сегментации изображения, не использующий линейного вырождения цветовых распределений. Анализ сцен, в которых присутствуют лишь ахроматические источники (источники с относительным спектральным распределением излучения $S(\lambda) \approx \text{const}$ в видимой области спектра), показал, что каждый однородно окрашенный объект при таком освещении, в безрефлексном приближении проецируется в нелинейное ЦП как точечный кластер или кластер в форме отрезка кривой с преимущественной ориентацией вдоль оси яркости. Это подтверждается экспериментальным исследованием изображений, полученных с реальных камер. Для таких случаев был разработан простой и эффективный алгоритм сегментации, основывающийся на детектировании в цветовом пространстве сенсора кластеров типа криволинейных отрезков [8].

Предлагаемый алгоритм состоит из этапа разбиения цветового пространства на слои семейством плоскостей, нормальных к главной диагонали цветового пространства, послойной кластеризации с помощью алгоритма поиска водоразделов, сборки кластеров и обратного проецирования на плоскость изображения с целью получения карты сегментации. Такой алгоритм показывает хорошие результаты при обработке сцен с малым количеством объектов, так как в каждом слое сечение кластера хорошо локализовано на поверхности сечения. Однако, на сложных изображениях при использовании приведенного выше алгоритма не исключена вероятность ошибок сегментации, если два объекта, не имеющие общей границы, порождают в цветовом пространстве перекрывающиеся распределения. Для таких изображений предложена (основная) версия алгоритма, учитывающая геометрическое соседство точек. В ней используется разбиение изображения на небольшие, плотно покрывающие изображение, прямоугольные области, каждая из которых сегментируется независимо. При этом при сегментации текущей области анализируется цветовое распределение, взятое по большей области. Тогда объект, далекий от текущей области, уже не зашумляет ее локальное цветовое распределение, а близкие объекты вносят существенный вклад, даже если их вклад по площади в сегментируемую область мал (см. рис. 2).

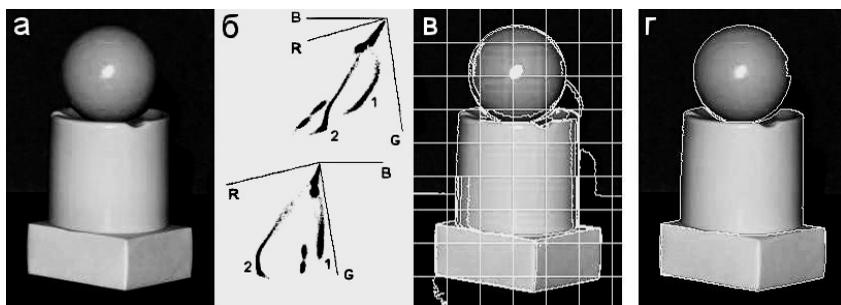


Рис. 2. Сегментация изображения с неизвестными характеристиками. а) исходное изображение; б) две проекции его цветового распределения; в) карта сегментации с разбиением на зоны; г) карта сегментации после релаксации

Для устранения границ между «областями внимания» в алгоритме используется процедура слияния областей с помощью ГСО. В качестве меры несходства $d_{i,j}$ при этом взята следующая функция:

$$d_{i,j} = \sqrt{\sum_{k=1}^3 \frac{(\mu_{i,k} - \mu_{j,k})^2 n_i n_j}{n_i + n_j}},$$

где n_i и n_j — количество пикселей в сравниваемых сегментах, а $\mu_{i,k}$ и $\mu_{j,k}$ — средние значения k -той цветовой компоненты по соответствующим сегментам.

Этот алгоритм допускает параллелизм двух уровней: каждая область изображения может обрабатываться независимо, как это было предложено и для линейного алгоритма, а, кроме того, первый (самый трудоемкий) этап кластеризации цветового распределения каждой области может производиться независимо для каждого слоя ЦП.

Заключение

Таким образом, на каждом из этапов сегментации возможно глубокое распараллеливание вычислительных процедур, позволяющее (при неограниченном числе процессоров) получать результат за время, не зависящее от размера изображения.

Подобная реализация позволит усложнение иерархии первоначальной двухэтапной схемы обработки до итеративной схемы цветового объектного различения, где возврат от гипотез об окраске «нулевого цикла» к уточняемой карте сегментов позволил бы на втором цикле улучшить и оценки окраски и/или цветности освещения. Это позволило бы при

известном дефиците надежных признаков освещения интерпретируемой сцены уменьшить влияние ошибок этапа сегментации. Такова наша концепция, эффективность которой еще предстоит показать в модельных численных экспериментах.

Литература

1. *Николаев П. П.* Модели и механизмы константного зрительного восприятия // Автореферат докторской диссертации. М., 1992. 60 с.
2. *Николаев П. П.* Монокулярное цветоразличение объемных предметов в разных условиях освещения // *Биофизика*. 1988. Т. 33. № 1. С. 140–144.
3. *Николаев Д. П., Николаев П. П.* Спектральные модели и алгоритмы цветовой константности // В сборнике: «Интеллектуальные системы (IEEE AIS'07)», «Интеллектуальные САПР (CAD–2007)»: Труды международной конференции. М.: Изд-во Физматлит, 2007. Т. 2. С. 117–125.
4. *Nikolaev D. P., Nikolayev P. P.* Linear color segmentation and its implementation. *Color Vision and Image Understanding*. 2004. V. 94. P. 115–139.
5. *Harris K., Efstratiadis N., Maglaveras N. et al.* Hybrid image segmentation using watersheds and fast region merging // *IEEE Trans. Image Proces.* 1998. V. 7. № 12. P. 1684–1699.
6. *Di Zenzo S.* A note on the gradient of multi-image // *Comput. Vision Graphics Image Process.* 1986. V. 33. P. 116–125.
7. *Карпенко С. М., Николаев Д. П., Николаев П. П., Постников В. В.* Быстрое преобразование Хафа с управляемой робастностью // *Искусственные интеллектуальные системы и Интеллектуальные САПР. Труды международной конференции IEEE AIS'04 и CAD–2004*. М.: Изд-во Физматлит. 2004. Т. 2. С. 303–309.
8. *Kim S. G., Nikolayev D. P.* Method and apparatus for sectioning image into plurality of regions. US patent application publication. Application No. 09/983032, Publication No. US 2002/0102017 A1. 2002. 25 p.