УДК 004.932
DOI 10.18413/2411-3808-2019-46-4-774-783

ПРОБЛЕМА ДЕТЕКТИРОВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ СЦЕНЫ

THE PROBLEM OF OBJECT DETECTION IN THE IMAGES OF GIVEN SCENE

М.В. Харинов, И.М. Толстой
M.V. Kharinov, I.M. Tolstoy

Санкт-Петербургский Институт информатики и автоматизации Российской академии наук,
Россия, 199178, г. Санкт-Петербург, 14 линия В.О., 39

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences,
39 14-th Linia VI, St. Petersburg, 199178, Russia

E-mail: khar@iias.spb.su

Аннотация

В статье обсуждается применение методов минимизации ошибки аппроксимации (суммарной квадратичной ошибки) для детектирования различных объектов на цветовых изображениях. Формулируется проблема детектирования объектов, которая состоит в неустойчивой кластеризации пикселей даже при незначительном изменении изображения. На конкретном примере показывается, что для устойчивого детектирования объектов на изображениях, снятых в различном ракурсе, перспективен метод Уорда, который является основой развиваемой в СПИИРАН модели квазиоптимальных приближений изображения. Новизна модели заключается в том, что в вычислениях поддерживается любая бинарная иерархия разбиений изображения на кластеры, которая строится в режиме попарного слияния кластеров, а анализируется в режиме их разделения надвое. Обратимые вычисления с кластерами пикселей поддерживаются в терминах сети, которая кратко описывается в статье.

Abstract

This paper discusses application of approximation error minimization (aggregate squared error) for detection of various objects on color images. The problem of object detection is formulated in such manner: pixel clustering remains unreliable even by minor image alterations. The considered use case demonstrates, that for reliable detection of objects on images, taken under different angles, Ward method is promising, actually underlying the quasi-optimal image approximation model, currently being developed in the SPIIRAS. The scientific novelty of the proposed model consists in the following: the computational framework enables any binary image clustering approach, which proceeds in pair-wise merge mode and analyzed in bipartition mode. Reversible calculations with pixel clusters are supported in terms of the network, which is briefly described in the paper.

Ключевые слова: детектирование иерархии объектов, кусочно-постоянные приближения изображения, суммарная квадратичная ошибка, оптимизация, аппроксимация, алгебраическая сеть, метод Уорда.

Keywords: hierarchical object detection, piecewise constant image approximation, total squared error, optimization, approximation, algebraic network, Ward’s method.

1. Введение

С практической точки зрения современное распознавание изображений является, по большей части, инженерной наукой, объединяющей множество приемов обнаружения и распознавания типичных или специфических объектов (текста [Shi et al., 2016], лиц [Schroff et al., 2015], дорожных знаков [Berkaya et al., 2016], подстилающей поверхности, проводов

на фоне неба и т. п.) на изображениях заданного типа (графических, полутоновых, дистанционных и т. п.). Тем не менее для того, чтобы выработать требуемые приемы обработки в области машинного зрения, помимо инженеров-программистов, трудятся многочисленные коллективы квалифицированных научных работников, что при интенсивном использовании машинного зрения в промышленности приводит к большим производственным затратам. Для снижения стоимости разработок программ искусственного интеллекта (ИИ) в целом и программ машинного зрения, в частности, в США в 2013-2017 гг. выполнен проект «Probabilistic Programming for Advanced Machine Learning» [PPAML]. В результате выполнения проекта планировалось «совершить революцию» в практическом программировании, подобную той, что произошла с внедрением операционной системы Windows. При этом ставилась задача разработать программный инструментарий, помогающий инженерам-программистам создавать конкретные интеллектуальные продукты подобно тому, как Windows помогает создавать интерфейсы для прикладных программ.

Важной особенностью постановки задачи [PPAML] в приложении к изображениям является то, что в ней обходится спорная проблема использования априорных данных об «объектах интереса» - они известны программисту-инженеру, для которого разрабатывается инструментарий. По нашему опыту [Харинов, 1993], недостатком постановки задачи [PPAML] является упор на машинное обучение, являющееся, по сути, эвристическим приемом обработки, который, с одной стороны, предполагает необходимость коррекции результатов, а с другой стороны, не дает гарантии успешного ее выполнения. Кроме того, обучение, как правило, является трудоемкой рутинной процедурой, причем эффективной только при достаточном объеме обучающих данных. Для эффективного детектирования объектов на изображении вместо обучения логичнее предусмотреть обоснованные настроечные параметры, обеспечивающие предсказуемый результат обработки. Так или иначе, но цель, сформулированная в [PPAML], кажется вполне достижимой. Поэтому, хотя на сегодняшний день обозначенная в [PPAML] революция либо не афишируется, либо откладывается, в ближайшие годы, по всей видимости, можно ожидать появления общедоступного программного обеспечения, снижающего трудоемкость разработки программ обнаружения и распознавания конкретных объектов.

2. Проблема выделения объектов

Проблема распознавания изображений, особенно изображений произвольного содержания, заключается в неустойчивом детектировании объектов, которое наблюдается даже при незначительном изменении содержания изображения. Это нетрудно проиллюстрировать на простом примере (рис. 1).

На рис. 1 вверху показано исходное изображение, которое составлено из 9 стандартных изображений сцены («Teddy»), снятой в различном ракурсе. Под ним показаны результаты выделения объектов, демонстрируемые кусочно-постоянными приближениями в трех цветах. Центральное представление изображения составлено из 9 изображений, каждое из которых получено методом Уорда [Ward, 1963; Мандель, 1988; Айвазян, 1989] из соответствующего фрагмента исходного изображения. Нижнее представление изображения получено тем же самым методом Уорда, но составное изображение обработано, как целое.

Задача предельно проста. Для выделения объектов требуется обработать изображения одной и той же сцены, снятые в одних и тех же условиях так, чтобы обработанные представления выглядели одинаково, - приблизительно как на нижней картинке.

Однако, как видно на картинке справа, минимальная модификация входных данных заметным образом сказывается на результатах детектирования объектов.

Прием совмещения изображений сцены в различном ракурсе для детектирования объектов на составном изображении (нижнем на рис. 1) является хорошим тестом при сравнении алгоритмов детектирования объектов. Метод Уорда выдерживает указанный тест, но это является, скорее, исключением, чем правилом. В противном случае, в программах вычисления расстояний до объектов сцены не приходилось бы прибегать к эвристическому предварительному поиску характерных точек, устойчивых к модификации контекста при изменении ракурса съемки [Малашин, 2014; Фаворская, Проскурин, 2015].

Устойчивое детектирование объектов на совмещенных стереопарах методом Уорда, как на нижней картинке рис. 1, может оказаться достаточным для продолжения решения задач обработки стереопар силами инженеров-программистов. Однако в общем случае распознавания цифрового изображения дополнительное изображение в момент обработки недоступно, а наблюдается только неустойчивое детектирования объектов, как на картинке справа рис. 1.

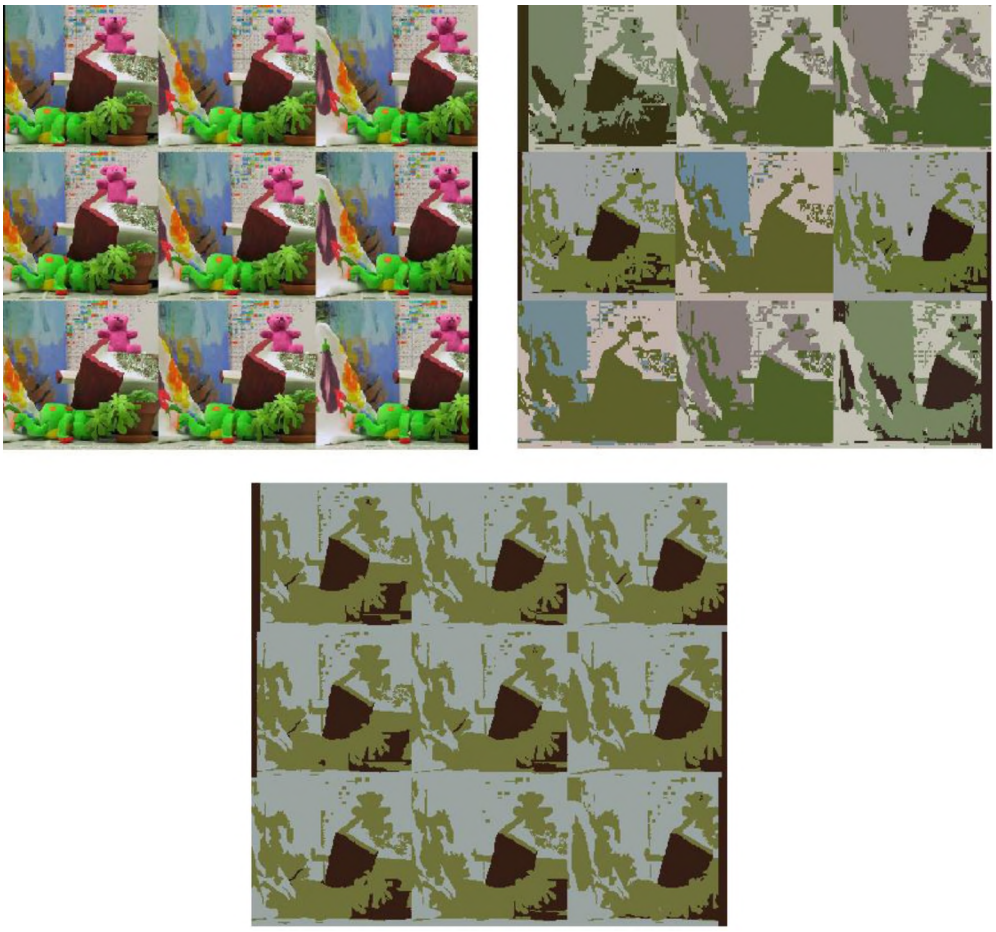

Рис. 1. Неустойчивое выделение объектов
Fig. 1. The unstable object selection

Проблему устойчивого детектирования объектов еще предстоит решить, возможно, сначала для незначительного изменения контента, как в случае стереопар, а затем и для более сложных случаев обработки изменчивых изображений. Для уверенного решения полезно воспользоваться той или иной моделью изображения [Kharinov, 2015; Kharinov, 2015; Чочиа, 2016; Харинов, 2018]. В настоящей работе решение строится в рамках модели квазиоптимальной кластеризации пикселей изображения [Kharinov, 2015; Kharinov, 2015; Харинов, 2018], разрабатываемой в СПИИРАН.

3. Модель

Модель [Kharinov, 2015; Kharinov, 2015; Харинов, 2018] разработана для полноразмерных (full HD) изображений, состоящих из миллионов пикселей. Без значительных изменений он переносится на случай аудиосигналов, поскольку использует линейное упорядочение пикселей в порядке прогрессивной развертки. При этом пиксели изображения размещаются в линейные массивы точно так же, как упорядоченные по времени семплы аудиосигнала.

Модель поддерживает первоначальную обработку изображения, в процессе которой выполняется иерархическое упорядочение кластеров пикселей. При этом ставится и решается проблема разделения изображения на вложенные изображения. Вложенные изображения обрабатываются как «объекты» различных размеров, которые обнаруживаются компьютером для распознавания, скажем, лиц и элементов лица на изображении, или предложений, слов и букв в речевом сигнале. По сравнению с обычной кластеризацией пикселей, иерархия генерируемых в модели кластеров пикселей позволяет конструктивно рассматривать детектирование всех объектов, доступных на этапе первоначальной обработки.

Модель опирается на то, что оцифрованное изображение характеризуется выраженной повторяемостью значений пикселей в цветовых компонентах. Основные положения модели [Kharinov, 2015; Kharinov, 2015; Харинов, 2018] заключаются в следующем:

1. Изображение без существенного искажения контента может быть аппроксимировано кусочно-постоянным представлением с ограниченным числом цветов.

2. Целевая последовательность кусочно-постоянных приближений изображения, рассчитанная для обнаружения всех доступных объектов, является иерархической.

3. Для изображения из N пикселей целевая иерархическая последовательность приближений изображения оптимизирована по среднеквадратичному отклонению cm или

$ошибке$ $аппрок\wedge\wedge мац\wedge и$ $E = 3Nc\wedge$, т. е. суммарной квадратичной ошибке.

Последовательность оптимальных приближений не является иерархической. Поэтому задача оптимизации состоит в приближении последовательности оптимальных приближений другой иерархической последовательностью $квазиоптимальных$ приближений, которые близки к оптимальным, по ошибке аппроксимации.

4. Целевая иерархическая последовательность квазиоптимальных приближений описывается выпуклой последовательностью значений ошибки аппроксимации.

Известно, что последовательность оптимальных кусочно-постоянных приближений изображения в зависимости от количества g кластеров пикселей описывается монотонно

возрастающей последовательностью $ЛЕ$ ^ AE <... < AEn _ J < 0 неположительных

приращений ошибки аппроксимации $_{ае}$ < о или выпуклой последовательностью самих значений E:

$$ {}^{E}g < \frac{{}^{E}g_{-1} + {}^{E}g_{1}}{2} \quad g = 2,3,\ldots, N-1. \qquad (1) $$

Такой же самой (выпуклой) последовательностью описывается и последовательность квазиоптимальных приближений (рис. 2).

Рис. 2 иллюстрирует аппроксимацию сигнала кусочно-постоянными приближениями. Показана зависимость ошибки аппроксимации E от числа кластеров g . Серая кривая описывает оптимальные приближения сигнала. Сплошная черная кривая описывает иерархию квазиоптимальных приближений. Обе кривые являются выпуклыми и имеют общее начало и конец. При определенном значении g = g кривая для квазиоптимальных приближений максимально приближается или совпадает с кривой для

оптимальных приближений. Это достигается путем минимизации ошибки аппроксимации для данного значения = 2,3,....

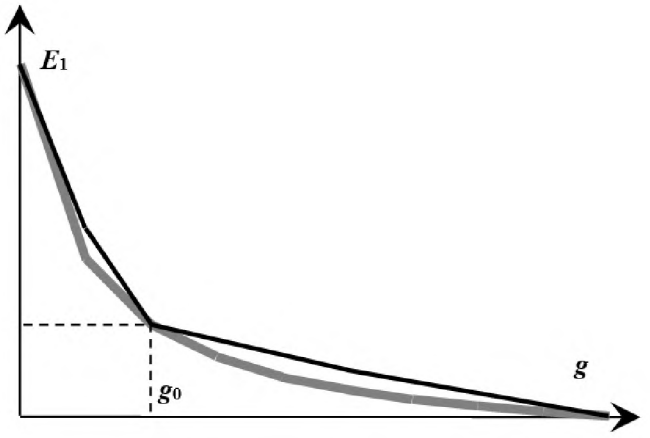

Рис. 2. Аппроксимация оптимальных приближений (серая кривая) квазиоптимальными приближениями (сплошная черная кривая)

Fig. 2. Approximation of optimal approximations (gray curve) by quasioptimal approximations (solid black curve)

Значение g является параметром настройки, который предусматривается в модели для управления online-обработкой изображения. Смысл его кажется очевидным - это некоторое число объектов, на которое настраивается компьютер для наилучшего их детектирования.

Для характеристики неоднородности, «сложности» [Чочиа, 2016], «заметности» [Achanta et al., 2009; Cheng et al., 2015] множеств пикселей изображения в модели вводится параметр гетерогенности |, который присваивается каждому кластеру. Этот параметр приравнивается к абсолютному значению величины приращения ошибки аппроксимации $Л E$, вызванной делением кластера на два. Важно, что параметр |л^^.^| является мерой неоднородности кластеров пикселей, т. к. подобно числу пикселей он не убывает при слиянии кластеров i , j в один кластер $i \wedge j$:

$$ЛЕ. \quad (\mathrm{i} \wedge \mathrm{j}) \backslash Ц^{Л}{}_{\varepsilon split}\{ i \wedge, \backslash Л_{\varepsilon split}\{ J \wedge . \qquad (2)$$

Иерархия приближений, описываемая выпуклой последовательностью значений ошибки аппроксимации E , E , ..., E и обозначенная на рис. 2 сплошной черной кривой, получается укрупнением пикселей методом Уорда [Ward, 1963; Мандель, 1988; Айвазян и др., 1989], в котором вначале каждый пиксель составляет самостоятельный кластер. Затем на каждом шаге сливается пара кластеров i , j , отвечающая минимальному приращению ошибки аппроксимации i, j):

$$\mathrm{i,j} \wedge {}^{i} \wedge j : \mathrm{i,j} = \arg\min_{\mathrm{i,\ j=0,2,..g\ -1}} {}^{ЛЕ}m \wedge erge\ ({}^{i,} J) \wedge \qquad (3)$$

где число кластеров снижается om N до 1, а приращение ошибки аппроксимации $Л^{\wedge}{}^{\wedge}rge(i, j)$ выражается через число писелей n, $П$ и трехкомпонентные средние значения I, I пикселей внутри кластеров i , j в виде:

$$Л_{Еwerge}\ (\mathrm{i,j}) = \sim Л_{ЕрШ}\ ({}^{i} \wedge j) = \frac{{}^{n}{}_{n}{}^{nj}}{+\ {}^{n}j} \quad \mathrm{I} - \mathrm{и} > 0. \qquad (4)$$

где $\{i \wedge j) \wedge 0$ неположительная величина приращения ошибки аппроксимации при разделении надвое кластера $i \wedge j$.

При изменении параметра минимизации выпуклая кривая рис. 2 преобразуется в новую выпуклую кривую, которую, очевидно, можно получить тем же самым методом Уорда, но с измененным порядком слияния кластеров. Обсуждаемая неоднозначность кластеризации пикселей по Уорду связана с необходимостью выбора между вариантами слияния пары кластеров, обеспечивающих минимальное приращение ошибки аппроксимации в процессе выполнения итеративного алгоритма. При этом эффективная минимизация ошибки аппроксимации при установленном числе кластеров g [Харинов, 2018] способствует подавлению неустойчивого детектирования объектов (см. рис. 1).

Другим предусмотренным в модели способом подавления неустойчивого детектирования объектов является алгоритм получения результирующих представлений изображения, генерируемых системой.

Результатом иерархической кластеризации является бинарная иерархия кластеров, сформированная в алгоритме итеративного слияния множеств пикселей. Указанная иерархия кластеров содержит N приближений изображения. Приближения содержат всего $2N$ -1 различных кластеров пикселей. Если N равно, скажем, миллиону, то для того, чтобы анализ такого числа кластеров пикселей (объектов) был доступен для человека, необходимо параметризовать и автоматизировать выделение относительно небольшого количества детектируемых объектов под управлением оператора. С целью детектирования объектов на упрощенных представлениях изображения в модели предусмотрена генерация, так называемых, «карт рейтинга объектов».

4. Карты рейтинга объектов

Детектирование последовательности геометрически непересекающихся объектов выполняется по пороговому значению параметра гетерогенности или площади, т. е. числу пикселей в кластере. Для этого из полной иерархической последовательности приближений выбирается подпоследовательность приближений, в которых встречаются кластеры пикселей со значениями параметра гетерогенности как выше, так и ниже установленного порога. В этом случае, поле изображения разделяется на области объектов и фона. Последовательность объектов, обнаруживаемых в сигнале, кодируется в рейтинговой карте объектов последовательными номерами в порядке их обнаружения (рис. 3).

На рис. 3 вверху слева показано исходное изображение «Лена». Справа показана рейтинговая карта объектов в 5 тонах с 14132 сегментами. Эта карта рассчитана для порога I, равного 1 % от максимального значения E , отмеченного на рис. 2. Внизу показано представление изображения, которое получается путем усреднения пикселей внутри сегментов карты рейтинга объектов и содержит 13151 цветов.

На карте рейтинга объектов, объекты, обнаруженные первыми, помечены черным цветом. Последними выявлены объекты, помеченные белым цветом. Значение параметра неоднородности, а также число пикселей обнаруженных объектов возрастает с увеличением яркости пикселей на карте рейтинга объектов.

Следует обратить внимание, что для получения экземпляра серой карты рейтинга объектов в нескольких тонах достаточно указать единственное пороговое значение параметра гетерогенности или числа пикселей, а не пару значений, как при детектировании объектов по диапазону значений числа пикселей в кластерах.

Полученные для данного порога значения пикселей карты рейтинга объектов можно рассматривать в качестве автоматически рассчитанных идентификаторов объектов. Карты рейтинга объектов, полученные для нескольких порогов параметра гетерогенности или числа пикселей в кластере, дают точки изображения, помеченные *векторами* порядковых номеров множеств пикселей, которые можно трактовать как автоматически генерируемые

имена объектов, предназначенные для дальнейшего распознавания объектов на изображении.

Рис. 3. Результат детектирования объектов
Fig. 3. The object detection result

5. Алгебраическая сеть

Скоростные операции с иерархически структурированными кластерами пикселей изображения поддерживаются в терминах сети, «наброшенной» на пиксели изображения. Сеть образуется деревьями (ациклическими графами) и циклами (циклическими графами). При этом вместо традиционных деревьев используются динамические деревья Слейтора-Тарьяна [Kharinov, 2015; Харинов, 2018]. Динамические деревья и циклы составляют динамическую сеть, которая называется алгебраической, поскольку получается путем итеративного объединения начальных элементарных сетей (рис. 4).

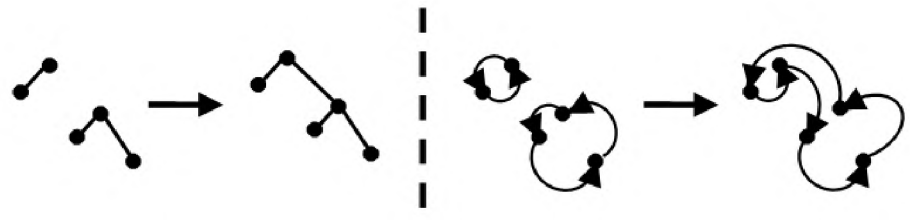

Рис. 4. Схема слияния деревьев и циклов
Fig. 4. The tree and loop merging scheme

Рис. 4 иллюстрирует слияние деревьев, при котором корневые узлы связываются дугой, и слияние циклов, при котором начальные узлы, соответствующие корневым узлам деревьев, обмениваются своими ссылками. Пиксели изображения на рис. 4, обозначенные черными точками, расположены вертикально и для наглядности разнесены горизонтально.

Характерной особенностью развиваемого программного инструментария является активное использование обратимых вычислений, которые поддерживаются в разных вариантах в терминах сети. В простейшем случае это обеспечивается благодаря формированию дерева Слейтора-Тарьяна, для которого посредством цикла чередования дуг запоминается порядок их установления (рис. 5).

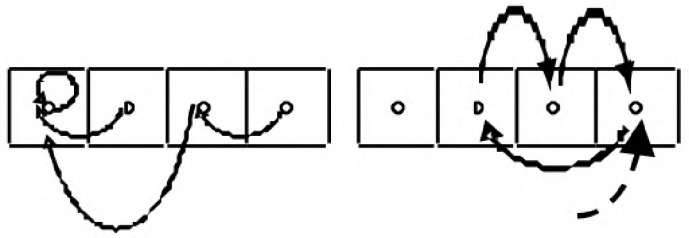

Рис. 5. Сеть, образованная деревом Слейтора-Тарьяна (слева) и циклом чередования дуг (справа)

Fig. 5. Kernel network of Sleator-Tarjan tree (left) and a cycle of setting order (right)

Рис. 5 иллюстрирует формирование простейшей сети на примере изображения из четырех пикселей. Пиксели показаны в виде клеток и размещены горизонтально.

На рис. 5 слева показано дерево пикселей, разрыв дуг в котором задает разделение множества пикселей на два подмножества. Справа на рис. 5 показан соответствующий цикл пикселей, который задает прямой порядок установления дуг и обратный порядок их разрыва. Тем самым обеспечивается воспроизведение шагов кластеризации в прямом и обратном порядке. Дополнительным параметром является указатель на начало или конец цикла. Упомянутый указатель показан пунктирной линией.

В обсуждаемой модели в дополнение к простому обращению порядка слияния кластеров в процессе разделения кластеров обеспечивается произвольный выбор того или иного кластера, который следует разделить на две части. Одновременно выполняется эквивалентная модификация исходного порядка слияния кластеров. Таким образом, обратимые вычисления не ограничиваются простым восстановлением данных на любом этапе, как в [Toffoli, 1980; Zongxiang, 2009], а реализуются в обобщенном смысле. При этом становится возможным уменьшить погрешность и улучшить качество аппроксимации сигналов за счет встречных операций слияния и разделения надвое кластеров пикселей.

6. Заключение

В заключение необходимо отметить следующее. Для развития внедрения машинного зрения актуальна разработка программного инструментария, обеспечивающего решение конкретных задач силами инженеров-программистов. Требуемый инструментарий возможно реализовать на основе модели квазиоптимальных приближений изображения [Харинов, 1993; Kharinov, 2015; Kharinov, 2015; Харинов, 2018; Kharinov, Buslavsky, 2019]. Предварительно протестировать модель можно в двух упрощенных вариантах:

1. Создать инструментарий для детектирования объектов на цветовых полномасштабных изображениях стереопар (см. рис. 1).

2. Создать инструментарий для обработки серых изображений с ограниченным числом, скажем, до десятка объектов, на основе оригинального мультипорогового метода Оцу [Liao, 2001] в сочетании с иерархическим методом Оцу [Kharinov, 2015; Kharinov, 2015], позволяющими выполнить вычисления согласно рис. 2.

Вероятно, со временем развитие системного программного обеспечения вычислений в терминах деревьев и алгебраических сетей на языках высокого уровня, в среде разработки MatLab и др. переведет в разряд инженерных саму задачу разработки инструментария для создания интеллектуальных программ согласно конкретной модели. Это позволит многочисленным коллективам научных работников сосредоточиться на решении фундаментальных задач информатики и теории информации.

Список литературы

1. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. 1989. Прикладная статистика: Классификация и снижение размерности. Москва, Финансы и статистика, 607.

2. Малашин Р.О. 2014. Методы структурного анализа изображений трехмерных сцен. Автореф. дис. канд. технич. наук. Санкт-Петербург, 22.

3. Мандель И.Д. 1988. Кластерный анализ. Москва, Финансы и статистика, 176.

4. Фаворская М.Н., Проскурин А.В. 2015. Категоризация сцен на основе расширенных цветовых дескрипторов. Труды СПИИРАН, 3(40): 203-220.

5. Харинов М.В. 1993. Разработка динамических структур данн^хх системы автоматизированного распознавания изображений. Автореф. дис. канд. технич. наук. Санкт-Петербург, 20.

6. Харинов М.В.2018. Об СКО, методе К-средних и деревьях в приложении к обработке изображений. Труды 28-й международной конференции по компьютерной графике и зрению (Графикон'2018), 180-184.

7. Чочиа П.А. 2016. Теория и методы обработки видеоинформации на основе двухмасштабной модели изображения. Дис. докт. технич. наук. Москва, 302.

8. Achanta R., Hemami S., Estrada F., Susstrunk S. 2009. Frequency-tuned salient region detection. IEEE conference on computer vision and pattern recognition (CVPR), 1597-1604.

9. Berkaya S.K., Gunduz H., Ozsen O., Akinlar C., Gunal S. 2016. On circular traffic sign detection and recognition. Expert Systems with Applications, 48: 67-75.

10. Cheng M.M., Mitra N.J., Huang X., Torr P.H., Hu S.M. 2015. Global contrast based salient region detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(3): 569-582.

11. Kharinov M.V. 2015. Model of the quasi-optimal hierarchical segmentation of a color image. Journal of Optical Technology, 82(7): 425-429.

12. Kharinov M.V. 2015. Pixel Clustering for Color Image Segmentation. Programming and Computer Software, 41(5): 258-266.

13. Kharinov M.V., Buslavsky M.V. 2019. Object Detection in Color Image. Pattern Recognition and Information Processing (PRIP'2019). Proceedings of the 14th International Conference. Minsk: Publishing Center of BSU, 43-47.

14. Liao P.S., Chen T.S., Chung P.C. 2001. A fast algorithm for multilevel thresholding . J. Inf. Sci. Eng, 17(5): 713-727.

15. PPAML (Probabilistic Programming for Advanced Machi-ne Learning), DARPA project, 2013-2017. Available at: https://galois.com/project/probabilistic-programming-for-advanced-machine-learning/.(accessed 7 November 2019)

16. Schroff F., Kalenichenko D., Philbin J. Facenet. 2015. A unified embedding for face recognition and clustering. Proceedings of the IEEE conference on computer vision and pattern recognition, 815-823.

17. Shi B., Bai X., Yao C. 2016. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. IEEE transactions on pattern analysis and machine intelligence, 39(11): 2298-2304.

18. Toffoli T. 1980. Reversible computing. In International Colloquium on Automata, Languages, and Programming, 632-644.

19. Ward Jr J.H. 1963. Hierarchical grouping to optimize an objective function. Journal of the American statistical association, 58(301): 236-244.

20. Zongxiang Y. 2009. Reversible Three-Dimensional Image Segmentation. US Patent № 20110158503 A1.

References

1. Ayvazyan S.A., Bukhshtaber V.M., Enyukov I.S., Meshalkin L.D. 1988. Prikladnaya statistika: Klassifikatsiya i snizhenie razmernosti [Applied Statistics: Classification and Dimension Reduction]. Moscow, Finansy i statistika, 607.

2. Malashin R.O. 2014. Metody strukturnogo analiza izobrazheniy trekhmernykh stsen [Methods of structural analysis of images of three-dimensional scenes]. Abstract. dis. cand. tech. sciences. Saint Petersburg, 22.

3. Mandel' I.D. 1988. Klasternyy analiz [Cluster analysis]. Moscow, Finansy i statistika, 176.

4. Favorskaya M.N., Proskurin A.V. 2015. Scene Categorization Based on Extended Color Descriptors. SPIIRAS Proceedings, 3(40): 203-220. (in Russian)

5. Kharinov M.V. 1993. Razrabotka dinamicheskikh struktur dannykh sistemy avtomatizirovannogo raspoznavaniya izobrazheniy [Development of dynamic data structures of an automated image recognition system]. Abstract. dis. cand. tech. sciences. Saint Petersburg, 20.

6. Kharinov M.V. 2018. Ob SKO, metode K-srednikh i derev'yakh v prilozhenii k obrabotke izobrazheniy [About the standard deviation, the K-means method and trees in the application to image processing]. 28th International Conference on Computer Graphics and Vision (GraphiCon 2018), 180-184.

7. Chochia P.A. Teoriya i metody obrabotki videoinformatsii na osnove dvukhmasshtabnoy modeli izobrazheniya [Theory and methods of video processing based on a two-scale image model]. Dis. cand. tech. sciences. Moscow, 302.

8. Achanta R., Hemami S., Estrada F., Susstrunk S. 2009. Frequency-tuned salient region detection. IEEE conference on computer vision and pattern recognition (CVPR), 1597-1604.

9. Berkaya S.K., Gunduz H., Ozsen O., Akinlar C., Gunal S. 2016. On circular traffic sign detection and recognition. Expert Systems with Applications, 48: 67-75.

10. Cheng M.M., Mitra N.J., Huang X., Torr P.H., Hu S.M. 2015. Global contrast based salient region detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(3): 569-582.

11. Kharinov M.V. 2015. Model of the quasi-optimal hierarchical segmentation of a color image. Journal of Optical Technology, 82(7): 425-429.

12. Kharinov M.V. 2015. Pixel Clustering for Color Image Segmentation. Programming and Computer Software, 41(5): 258-266.

13. Kharinov M.V., Buslavsky M.V. 2019. Object Detection in Color Image. Pattern Recognition and Information Processing (PRIP'2019). Proceedings of the 14th International Conference. Minsk: Publishing Center of BSU, 43-47.

14. Liao P.S., Chen T.S., Chung P.C. 2001. A fast algorithm for multilevel thresholding . J. Inf. Sci. Eng, 17(5): 713-727.

15. PPAML (Probabilistic Programming for Advanced Machi-ne Learning), DARPA project, 2013-2017. Available at: https://galois.com/project/probabilistic-programming-for-advanced-machine-learning/.(accessed 7 November 2019)

16. Schroff F., Kalenichenko D., Philbin J. Facenet. 2015. A unified embedding for face recognition and clustering. Proceedings of the IEEE conference on computer vision and pattern recognition, 815-823.

17. Shi B., Bai X., Yao C. 2016. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. IEEE transactions on pattern analysis and machine intelligence, 39(11): 2298-2304.

18. Toffoli T. 1980. Reversible computing. In International Colloquium on Automata, Languages, and Programming, 632-644.

19. Ward Jr J.H. 1963. Hierarchical grouping to optimize an objective function. Journal of the American statistical association, 58(301): 236-244.

20. Zongxiang Y. 2009. Reversible Three-Dimensional Image Segmentation. US Patent № 20110158503 A1.

Ссылка для цитирования статьи
For citation

Харинов М.В., Толстой И.М. 2019. Проблема детектирования объектов на изображениях сцены. Научные ведомости Белгородского государственного университета. Серия: Экономика. Информатика. 46 (4): 774-783. DOI 10.18413/2411-3808-2019-46-4-774-783

Kharinov M.V., Tolstoy I.M. 2019. The problem of object detection in the images of given scene. Belgorod State University Scientific Bulletin. Economics. Information technologies. 46 (4): 774-783 (in Russian). DOI 10.18413/2411-3808-2019-46-4-774-783