

# ОБЪЕКТНАЯ ЛОКАЛИЗАЦИЯ СЕМАНТИЧЕСКИХ БЛОКОВ НА РАСТРОВЫХ ИЗОБРАЖЕНИЯХ

С. В. КУЛЕШОВ, А. А. ЗАЙЦЕВА

Санкт-Петербургский институт информатики и автоматизации РАН

СПИИРАН, 14-я линия ВО, д. 39, Санкт-Петербург, 199178

<http://sial.iias.spb.su>

---

УДК 004.6:004.7

Кулешов С. В., Зайцева А. А. **Объектная локализация семантических блоков на растровых изображениях** // Труды СПИИРАН. Вып. 7. — СПб.: Наука, 2008.

**Аннотация.** Рассматривается метод объектной локализации семантических блоков на растровых изображениях. В основе предлагаемого метода лежит анализ результатов адаптивного пирамидального представления растровых изображений. Особенностью алгоритма, формирующего пирамидальное представление, является связь величины значений последнего уровня иерархии и уровня семантической значимости, что позволяет находить области, представляющие наибольший интерес при восприятии их человеком, и использовать такое выделение для динамического управления компрессией. — Библ. 4 назв.

UDC 004.6:004.7

Kuleshov S. V., Zaitseva A. A. **The object localization of semantic blocks on raster images** // SPIIRAS Proceedings. Issue 7. — SPb.: Nauka, 2008.

**Abstract.** The method of object localization of semantic blocks on raster images is considered. The basis of an offered method are results of analysis of adaptive pyramidal representation of raster images. The feature of the algorithm of forming pyramidal representation is the relation between last level of hierarchy value with the level of semantic value that allows to find the areas representing the greatest interest in human perception and to use such selection for dynamic control of a compression. — Bibl. 4 item.

---

## 1. Введение

Результатом распознавания лиц и идентификации речи часто является идентификация некоторого определенного объекта. Под идентификацией в данном случае понимается присвоение рассматриваемому объекту надлежащего и однозначного названия. Человеку в процессе идентификации совсем необязательно явным образом определять характерные признаки объекта. Имеет значение, в сущности, только окончательный результат процесса наблюдения, восприятия и распознавания. Автоматические системы, решающие проблемы распознавания посредством органов чувств человека, должны осуществлять такую же идентификацию, как и человек, сталкивающийся с определенным объектом. Однако распознающие системы должны явным образом использовать характерные признаки объекта [1].

Задача морфологического анализа является одной из самых общих задач обработки изображений. В общем случае задача морфологического анализа должна выделять все объекты, находящиеся на изображении и давать их описание (принадлежность к классу объектов с указанием их положения на изображении).

Часто задача морфологического анализа сводится к задаче распознавания образа — выделение области интереса, заданного в виде шаблона или определенного по некоторым правилам. На рис. 1 показано выделение области интереса, произведенной человеком-экспертом.



Рис. 1. Пример выделения семантически значимых участков.

Морфологические методы анализа изображений ориентированы на решение задач, которые могут быть сгруппированы в два больших класса. Первый класс включает задачи выделения неизвестных объектов на изображениях одной сцены, полученных при неизвестных условиях. Например, при наличии изображения местности и объектов на ней, полученных в летний день в хорошую погоду, и изображения этой же местности в дождливый осенний день и, возможно, в другом участке электромагнитного спектра необходимо при анализе этих изображений выделить объект, отсутствующий на «летнем» изображении, но различимый на «осеннем». Трудность, с которой традиционно сталкиваются при решении этой задачи, состоит в том, что изменение условий регистрации приводит к изменениям в изображении, часто более существенным, чем появление/исчезновение объектов на местности.

Второй класс задач связан с поиском известных объектов на неизвестной сцене. Например, имеется снимок местности, сделанный с известной точки зрения, причем некоторый участок местности освещен лучше других. На другом изображении, полученном при других условиях и, возможно, в другом спектральном диапазоне, необходимо узнать выделенный участок и оценить ракурс, с которого был сделан этот снимок. Хорошо известно, что так называемые «корреляционные методы», используемые достаточно широко, не справляются с задачами этого типа, поскольку распределение яркости на целевом участке обычно бывает сильно искажено по сравнению с изображением–оригиналом, и это искажение не может быть сведено к простому изменению яркости и контраста фрагмента изображения как целого.

Другим подходом к локализации объектов является использование критерия «увеличение сложности» фрагментов изображения, содержащих семантически значимые элементы. Считается, что наличие семантически значимого объекта сопровождается локальным увеличением некоторой энтропийной характеристики. Такой подход имеет сразу ряд недостатков — сложность прямой реализации математической модели и наличие ложных срабатываний на зашумленных участках изображения.

Ниже рассматривается альтернативный подход к локализации семантически значимых блоков на изображениях.

## 2. Описание метода

В основе предлагаемого метода объектной локализации семантических блоков лежит анализ результатов адаптивного пирамидального представления растровых изображений [2].

Традиционный способ представления растровых изображений основан на передаче абсолютных значений интенсивности цветовых компонент, т. е. по сути отличия передаваемого изображения от черного (в некоторых форматах — белого) однородного прямоугольника.

Основной принцип пирамидального представления приведен на рис. 2 и заключается в том, что для минимизации битовой размерности передаваемых элементов имеет смысл передавать разность не относительно пустого изображения, а относительно похожего изображения, так как в этом случае отличия будут минимальны.

Таким шаблоном (подложкой, канвой) может служить уменьшенный вариант изображения, растянутый до требуемого размера. Интуитивно понятно, что многократное повторение уточнения областей относительно фона является наиболее эффективным способом представления, в смысле минимизации, разрядности передаваемых данных. При этом уровни передаются последовательно группами.

Применение адаптивного пирамидального представления для задач компрессии изображений подробно рассмотрено в [3].

Особенностью алгоритма, формирующего пирамидальное представление, является связь величины значений последнего уровня иерархии (для элементов с размером  $1 \times 1$ ) и уровня семантической значимости, что позволяет находить области, представляющие наибольший интерес при восприятии их человеком.

Элементом разложения (на заданном уровне разложения) назовем неделимый фрагмент изображения, воспринимаемый алгоритмом как одно целое, имеющее яркостные и цветовые характеристики, определяемые как средние значения соответствующих характеристик входящих в него пикселей. Исходное изображение (первый уровень разложения) воспринимается как один элемент размером  $W \times H$ , где  $W$  — ширина изображения,  $H$  — высота изображения в пикселях.

На каждом последующем уровне разложения происходит разбиение каждого элемента на элементы, имеющие линейный размер в 2 раза меньший предыдущего, пока не достигается размер элементов в  $1 \times 1$  пиксел. Такое последовательное разложение назовем пирамидальным представлением изображения.

Алгоритм, показанный на рис. 2, одновременно с формированием пирамидального представления определяет уровни, на которых находятся терминальные зоны. Выделению подлежат участки, соответствующие терминальным зонам с наибольшим номером уровня.

Назовем терминальной зоной элемент уровня разбиения, последующая детализация которого не дает дополнительной информации.

Терминальная зона отвечает критерию

$$\sum_i d_i > \varepsilon,$$

где  $d_i$  — величина изменения элемента на следующем иерархическом уровне,  $\varepsilon$  — величина порога неразличимости (понятие  $\varepsilon$ -неразличимости описано в [4]).

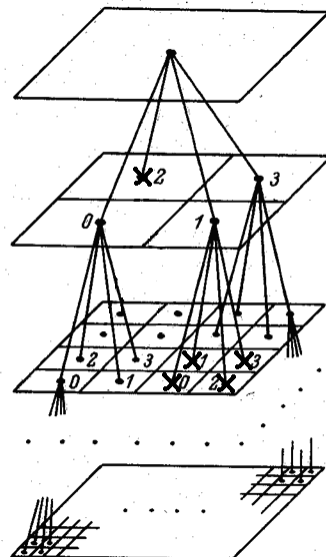


Рис. 2. Иллюстрация алгоритма локализации областей при пирамидальном представлении (крестиком отмечены терминальные зоны).

Недостатком алгоритма является привязанность границ выделения к сетке разбиения изображения. Для компенсации этого недостатка возможно многократное повторение алгоритма со сдвигом центра первого уровня разбиения и формирование границ выделения по совокупности результатов каждого из разбиений.

Основным применением такого алгоритма могут стать алгоритмы предобработки, компрессии с потерями и передачи по линиям с фиксированной пропускной способностью с динамическим изменением качества фрагментов изображения для компенсации психофизиологического фактора восприятия этих изображений.

### 3. Оценка результатов

Для оценки результатов работы предлагаемого метода были проведены эксперименты по выделению семантически значимых участков (области интереса) на различных типах растровых изображений с фиксированным значением параметра порога неразличимости.

Примеры работы, приведенные на рис. 3, 4, показывают устойчивое выделение области интереса (фрагменты, имеющие повышенную сложность) на фотографиях с достаточно сложным фоном. В левой части приведены исходные изображения, в правой — сумма исходного изображения и маски, определяющей область интереса.

По сравнению с примером на рис. 3, на котором объект, задающий область интереса, находится на однородном фоне, пример на рис. 4

показывает наличие дополнительных объектов на заднем фоне, усложняющих задачу локализации семантически значимых областей.

Предложенный алгоритм способен успешно решать как задачи выделения неизвестных объектов на изображениях одной сцены, полученных при неизвестных условиях (задачи первого класса при морфологическом анализе), так и задачи поиска известных объектов на неизвестном фоне (задачи второго класса).

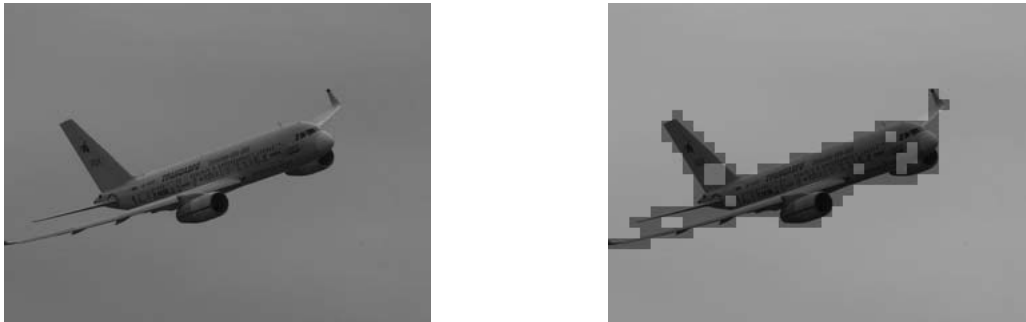


Рис. 3. Пример работы алгоритма на простом изображении.



Рис. 4. Пример работы алгоритма на сложном изображении.

Особенностью предложенного алгоритма в отличие от алгоритмов поиска границ является его нечувствительность к шуму, т.е. при его использовании выделения шума не наблюдается (примеры использования алгоритма поиска границ к тестовым изображениям приведены для сравнения на рис. 5).

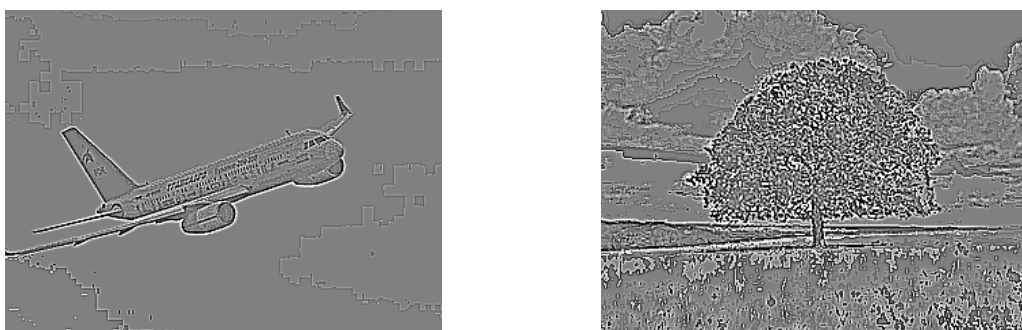


Рис. 5. Пример работы алгоритма выделения границ.

Эффективность динамического управления степенью компрессии можно оценить по рис. 6, 7. Примем размер исходного изображения (рис. 6, а) за 100%, тогда размер компрессированного изображения без динамического управления степенью компрессии (рис. 7) — за 35%, а с динамическим управлением

степенью компрессии (рис. 6, б) — за 38%. Результат эксперимента показывает, что визуально изображения с динамическим управлением степенью компрессии с потерями (lossy compression) лучше изображения с фиксированным значением компрессии при сравнимом битовом объеме.



Рис. 6. Пример работы алгоритма на изображениях в задаче динамического управления степенью компрессии с потерями: а — исходное изображение, б — изображение после обработки.



Рис. 7. Пример применения алгоритма сжатия с потерями без динамического управления степенью компрессии.

Совокупность полученных экспериментальных результатов показывает применимость описанного метода для этапа объектной локализации в задачах морфологического анализа.

Основным применением такого алгоритма могут стать алгоритмы предобработки в задачах узнавания и распознавания, динамического управления степенью компрессии с потерями и в ряде других задач.

## Литература

1. *Максимов В. В.* Система, обучающаяся классификации геометрических изображений. / Сборник «Моделирование обучения и поведения». М.: Наука. 1975. С. 29–120
2. *Александров В. В., Горский Н. Д.* Представление изображений. Рекурсивный подход. — Л.: Наука. 1985. 190 с.
3. *Кулешов С. В., Зайцева А. А., Аксенов А. Ю.* Ассоциативно-пирамидальное представление данных. // Информационно-измерительные и управляющие системы, №4, т. 6, 2008. — С. 14–17
4. *Александров В. В., Кулешов С. В., Цветков О. В.* Цифровая технология инфокоммуникации. Передача, хранение и семантический анализ текста, звука, видео. — СПб.: Наука. 2008. 244 с.