

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ И РЕЧИ

УДК 004.932.2

П.В. Лукашевич, Б.А. Залесский

ОБНАРУЖЕНИЕ ЗДАНИЙ НА КЛАСТЕРИЗОВАННЫХ
АЭРОФОТОСНИМКАХ И КОСМИЧЕСКИХ ИЗОБРАЖЕНИЯХ

Предлагается подход к задаче обнаружения зданий на аэрофотоснимках и космических изображениях, использующий кластерные представления этих изображений. Подход позволяет выделять здания на цветных изображениях с разрешением выше 3-4 м/пиксел. Программная реализация предложенного подхода существенно быстрее известных нам реализованных алгоритмов. Достаточно высокая надежность и быстроедействие реализаций обеспечиваются возможностью выделения границ кластеров и получения их достаточно точного упрощенного представления. Особенностью подхода является использование нескольких критериев, каждый из которых с высокой вероятностью выделяет часть зданий и при этом обладает близкой к нулю вероятностью ошибочного выделения.

Введение

Задача автоматического обнаружения зданий является одним из важнейших этапов анализа аэрофотоснимков и космических изображений. Хотя данное направление активно развивается с 1980-х гг., в настоящее время задача далека от окончательного решения, которое позволило бы создать программные комплексы, надежно обнаруживающие без вмешательства человека большую часть сооружений. И это несмотря на то, что за последние годы разрешение космических снимков значительно возросло, а их качество улучшилось.

По данной тематике опубликовано огромное число работ, известно большое число алгоритмов, использующих различные технические приемы, поэтому описание текущего состояния исследований – непростая задача. Детальный обзор наиболее известных алгоритмов, разработанных в 1990-х гг., и их классификация приведены в работе Х. Мейера [1]. Для выделения признаков зданий в них используются: преобразование Хафа; активные контуры; различные способы вычисления ориентированного градиента и геометрических примитивов (например, образованных крышами зданий линий, углов, прямоугольников и др.); анализ цветового содержания изображений; анализ теней; анализ фотометрических свойств материалов, из которых изготовлены крыши; различные модели зданий и т. д.

За последнее десятилетие алгоритмы существенно усложнились, появились новые подходы к обнаружению зданий. В частности, шире стали применяться подходы, базирующиеся на 3D-анализе сцен: либо одновременно использующие несколько изображений местности, либо применяющие поиск на основе 3D-моделей зданий. Для выделения признаков домов стали использоваться современные методы, включая популярные SURF и SIFT [2–4]. Сформировалось новое направление, основанное на стохастических методах минимизации функций энергии [5–7]. Использование функции энергии позволяет учитывать в модели разнообразные признаки, вплоть до касания и непересечения зданий, и производить поиск глобального минимума достаточно простыми с точки зрения реализации стохастическими методами. Появились подходы к выделению зданий на инфракрасных, мультиспектральных, радарных изображениях. Подробная структурированная по разделам библиография современных алгоритмов выделения зданий приведена в [8, 9]. Заметим, что в силу разнообразия форм и цвета зданий современные алгоритмы, по сути, являются многокритериальными, опирающимися на нетривиальный анализ нескольких признаков выделяемых объектов.

На самом деле, используются несколько формулировок задачи обнаружения зданий, которые определяются типом и объемом используемой априорной информации (например, количеством и качеством изображений, данными о минимальных и максимальных размерах домов, их форме, цвете), а также требованиями к получаемым решениям и т. д.

Задачу обнаружения зданий будем понимать следующим образом: имеется RGB-космическое изображение ландшафта с разрешением от 3-4 м/пиксел и выше, снятое в надири или под небольшим углом к поверхности. Требуется выделить изображения зданий многоугольной формы при соблюдении одного из двух условий: 1) априорная информация о цвете крыш зданий неизвестна; 2) имеются данные о цвете крыш.

Информация о предельных размерах зданий может задаваться при решении задачи как параметр, что сводит задачу к обнаружению зданий, размеры которых лежат в заданном диапазоне.

Следует отметить, что второе условие о цвете крыш облегчает задачу, позволяет получить более точные результаты, но практически никогда не превращает ее в тривиальную сегментацию изображения по цвету, так как в большинстве случаев на изображении присутствует достаточно большое число объектов, цвет которых весьма близок к цвету крыш. Выделение зданий только по цвету во многих случаях дает неточный результат.

Особенности рассматриваемого подхода заключаются в том, что он предполагает использование кластеризованных изображений. Для кластеризации используются разработанные авторами ранее алгоритмы [10–12], специально модифицированные для решения сформулированной задачи. Кластеры и их границы выделяются метками. Каждая граница имеет метку своего кластера, что позволяет при выделении зданий учитывать только те границы, которые расположены близко к области предполагаемого дома. Граница приближается с заданной точностью ломаной с относительно небольшим числом звеньев, что дает возможность анализировать ее форму, затрачивая на это сравнительно небольшое число операций. Используется информация о геометрических характеристиках границ, о соседстве кластеров, являющихся кандидатами на принадлежность дому, а также информация о тенях там, где они присутствуют.

При решении задачи с учетом любого из двух условий используются несколько критериев для выделения зданий, каждый из которых позволяет выделить только часть зданий, однако при этом имеет маленькую вероятность ошибки второго рода: практически не обнаруживает ложные дома там, где их нет. Выделение домов и их элементов производится на основе анализа кластеров и их границ каждым критерием отдельно, чтобы избежать увеличения ошибки второго рода. Каждый отдельный критерий присваивает бинарные детерминированные метки кластерам, проверяя их на обладание заданного признака. Затем применяется составное решающее правило (которое может быть представлено ациклическим графом), анализирующее расставленные метки и принимающее окончательное решение о том, какие кластеры соответствуют зданиям.

Данный подход позволил получить весьма точный результат (близкий к 90 %) при выделении зданий на тестовых изображениях.

1. Описание алгоритма

Прежде чем начать описание алгоритма, формализуем постановку задачи обнаружения зданий на изображениях. Дан аэрофотоснимок или космическое RGB-изображение ландшафта, снятое в надири или под небольшим углом к поверхности. Требуется найти на нем здания, имеющие форму многоугольников при удовлетворении одного из двух условий о цвете крыш зданий. Предполагается, что известны минимальные и максимальные размеры прямоугольников, ограничивающих величину зданий, задаваемые парами чисел u, v ($u \leq v$) и U, V ($U \leq V$) соответственно.

Алгоритм обнаружения зданий на аэрофотоснимках и космических изображениях состоит из трех независимых блоков. Первый блок алгоритма кластеризует специальным образом исходное цветное изображение. Второй блок находит тени на изображении. Задача обнаружения зданий может решаться без использования выделенных теней, но в этом случае решение имеет меньшую точность. Третий блок выполняет совместный многокритериальный анализ цветного кластеризованного изображения и бинарного изображения с выделенными тенями.

Опишем кратко *первый блок алгоритма*, выполняющий кластеризацию исходного космического изображения. Для выполнения кластеризации могут использоваться разные подходы, в том числе опирающиеся на известный алгоритм Smart Blur, однако применение модификаций разработанных авторами ранее комбинаторных алгоритмов кластеризации, подробно описанных в [10–12], дает лучшие результаты. Модификация известных алгоритмов потребовалась для того, чтобы выделенные на изображении кластеры соответствовали либо крышам (в случае

их однородности), либо реальным частям крыш (скатам, надстройкам, трубам и т. д.). Для этого используется следующая последовательность действий. Исходное цветное космическое изображение (рис. 1, а) кластеризуется одним из методов, предложенных в [10–12]. Использование метода, предложенного в [12], предпочтительнее, так как в нем при кластеризации, выполняемой с помощью методов максимального потока в графе, учитывается градиент изображения. После предварительной кластеризации части крыш и детали почти всех зданий состоят из нескольких кластеров, что снижает качество анализа изображений. В первом блоке к разработанному алгоритму кластеризации добавлена несложная процедура анализа, которая производит слияние/разделение кластеров с помощью алгоритма наращивания областей [13] (рис. 1, в и г). Использовано следующее условие присоединения к текущему кластеру: сосед просматриваемого пиксела текущего кластера присоединяется, если пиксели близки по цвету в RGB- или ином цветовом пространстве и между ними не проходит бинаризованный градиент исходного изображения.



Рис. 1. Этапы обработки изображения: а) исходное изображение; б) выделение теней; в) предварительная кластеризация с наложенным градиентом; г) окончательная кластеризация с наложенным градиентом

Данная простая процедура позволяет преобразовать кластеры таким образом, что на улучшенном изображении новые кластеры в основном соответствуют реальным деталям и элементам крыш (рис. 1, г). Заметим, что использование кластерного представления облегчает процессы исправления ошибок и интерактивного выделения зданий, так как для выделения крыши или ее детали достаточно указать на один из пикселей кластера.

Второй блок алгоритма выделяет тени на космическом изображении. Результатом работы данного блока является бинарное изображение, на котором теньям соответствуют черные области, а незатененным областям – белые. Под теньями понимаются области исходного изображения, не принадлежащие зданиям, на которые не попадают прямые лучи солнечного света. Процедура выделения теней может выполняться классическими пороговыми методами на исходном RGB-изображении. Авторы использовали для этого специально разработанный алго-

ритм, описанный в [14]. Алгоритм выделяет тени, автоматически вычисляя порог бинаризации для каждой локальной области изображения. На использованных изображениях более точный результат давали алгоритмы выделения теней (классические и предложенный авторами) при применении их к кластеризованным изображениям (рис. 1, б). Заметим, что нередко встречаются космические изображения, на которых отсутствуют тени. При работе с ними использование второго блока алгоритма не имеет смысла.

В отличие от первого и второго блоков, опубликованных авторами ранее, *третий блок алгоритма* является новым, поэтому остановимся на его описании подробно.

Сначала опишем признаки кластеров, использованные в алгоритме. Для этого обозначим через $S = \{0, \dots, n_1 - 1\} \times \{0, \dots, n_2 - 1\}$ множество пикселей $\mathbf{j} = (j_1, j_2)$ RGB-изображения \mathbf{I} размером $n_1 \times n_2$ с векторами цвета $I_{\mathbf{j}} = (I_{R,\mathbf{j}}, I_{G,\mathbf{j}}, I_{B,\mathbf{j}})$, $\mathbf{j} \in S$.

Под кластером изображения C будем понимать максимальный по размеру связный в восьмиточечной системе окрестностей набор пикселей, имеющих одно и то же значение цвета, т. е. $I_{\mathbf{j}_1} = I_{\mathbf{j}_2}$ для любых $\mathbf{j}_1, \mathbf{j}_2 \in C$. Кластеризованное изображение будем обозначать через $\tilde{\mathbf{I}}$, набор его кластеров – через \mathcal{E} , а цвет пикселей $\tilde{\mathbf{I}}$ на кластере C – через $\tilde{I}(C)$. Площадь кластера для нас – число пикселей $|C|$ в нем.

Граница $\Gamma(C) = (\gamma_0(C), \gamma_1(C), \dots, \gamma_{M-1}(C))$ кластера C – это упорядоченное множество пикселей, принадлежащих C и образующих замкнутую растровую кривую, такое, что $\gamma_{i-1}(C), \gamma_i(C), \gamma_{i+1}(C)$ – соседи в восьмиточечной системе окрестностей (соседями $\gamma_0(C)$ являются $\gamma_{M-1}(C), \gamma_1(C)$, соседями $\gamma_{M-1}(C)$ – пиксели $\gamma_{M-2}(C), \gamma_0(C)$). Длину границы $\Gamma(C)$ будем понимать как число пикселей в множестве, обозначая ее $|\Gamma(C)|$.

Для повышения быстродействия алгоритма было использовано приближение границы ломаными в sup-норме. Методы приближения границы ломаными хорошо зарекомендовали себя при анализе формы объектов (см., например, [15, 16]). Использовался простой и быстрый итерационный метод приближения, основанный на задании новых узлов ломаной в точках локальных максимумов уклонения участков границы $\Gamma(C)$ между уже найденными соседними узлами от отрезков, соединяющих эти узлы. Обозначим требуемую точность приближения через ε .

Краткое описание *алгоритма приближения границы $\Gamma(C)$ ломаной* выглядит следующим образом. Выбираем пиксели – концы диаметра границы $\Gamma(C)$, т. е. для евклидовой нормы $\|\cdot\|$ пара

$$(d_0, d_1) = \operatorname{argmax}_{(\gamma_0, \gamma_1) \in \Gamma(C)} \|\gamma_0 - \gamma_1\|.$$

Для определенности полагаем, что d_0 встречается в $\Gamma(C)$ раньше, чем d_1 . Считаем, что на первом шаге граница приближена диаметром d_0, d_1 . Просматриваем часть пикселей границы $\Gamma(C)$ между d_0 и d_1 , находим пиксел, наиболее удаленный от отрезка прямой с концами d_0, d_1 . Если он удален от отрезка на расстояние, большее ε , добавляем его в приближающую ломаную. Повторяем аналогичные вычисления для второй части пикселей $\Gamma(C)$, лежащих между d_1 и d_0 . Если находится пиксел из второй части $\Gamma(C)$, удаленный от отрезка на расстояние, большее ε , также добавляем его в приближающую ломаную. Далее повторяем вычисления для каждого звена построенной ломаной. Вычисления прекращаются, когда на текущем шаге не нашлось ни одного пиксела $\Gamma(C)$, удаленного от построенной ломаной на расстояние, большее ε . Для приближения границ кластеров использованных изображений $\tilde{\mathbf{I}}$ требовалось не более четырех-пяти итераций построения ломаных. Узлы полученных ломаных, приближающих границы $\Gamma(C)$ кластеров $C \in \mathcal{E}$ с заданной точностью ε , обозначим как $\mathfrak{B}(C) = (\beta_1(C), \beta_2(C), \dots, \beta_K(C))$ ($K = K(\varepsilon, C)$ и $K \leq M$).

Специфика предложенного алгоритма состоит в том, что решение о принадлежности кластера зданию принимается по совокупности признаков. Каждый из признаков конструируется так, чтобы с наименьшими ошибками второго рода (ложного обнаружения) выделить часть зданий. Все признаки имеют числовое выражение в своей шкале, поэтому для простоты работы и исключения трудностей оценки их совместного применения на основании значения числового признака вычисляется значение бинарной переменной (принимающей только значения ноль или единица), отражающей отсутствие или наличие признака у кластера. Окончательное решение о принадлежности кластера зданию производится на основе разработанного классификатора, который является логическим выражением от бинарных переменных и также может быть представлен в виде ациклического графа.

Опишем количественные признаки и правила задания на их основе значений бинарных переменных, использованных в третьем блоке алгоритма для обнаружения зданий.

Замечание. При решении сформулированной задачи вычислялось более 30 различных признаков зданий, получаемых на основе кластерных представлений изображений. Исследовались различные виды решающих правил, представимых в виде ациклических графов. Ниже описывается решающее правило, дающее наилучшие результаты для аэрокосмических изображений плотной европейской городской застройки. Очевидно, что при изменении характеристик изображений и типов зданий система признаков и вид решающего правила могут измениться.

Начнем с самого простого признака $\rho_{size}(C) = |C|$, который характеризует площадь кластера, и связанной с ним бинарной переменной $\varphi_{size}(C)$, равной единице, если $u \cdot v \leq \rho_{size}(C) \leq U \cdot V$, и равной нулю в противном случае.

Одним из наиболее простых признаков, позволяющих отбросить группу кластеров, не являющихся частями крыш зданий, является изопериметрическое отношение $\rho_{iso}(C)$, задаваемое дробью $\rho_{iso}(C) = |\Gamma(C)| / \sqrt{|C|}$. Известно, что наименьшее значение изопериметрического отношения, равное $2\sqrt{\pi} (\approx 3,545)$, достигается для круга. Для квадрата оно равно четырем. Чем больше вытянут объект и чем менее гладкая у него граница, тем больше значение $\rho_{iso}(C)$. Значение связанной с ним бинарной переменной $\varphi_{iso}(C) = \varphi_{iso}(\rho_{iso}(C))$ задается по порогу τ_{iso} : $\varphi_{iso}(C) = 1$, если $\rho_{iso}(C) \geq \tau_{iso}$, и $\varphi_{iso}(C) = 0$ в противном случае.

Следующий признак $\rho_{tetr}(C)$ несет существенно бóльшую информацию о форме кластера. Он задает количественную меру близости формы кластера к форме четырехугольника с заданными углами, наилучшим образом приближающего кластер. В разработанных ранее алгоритмах признак такого типа часто используется в качестве основного или даже единственного. В качестве меры близости обычно используются либо площадь симметрической разности четырехугольника и кластера, либо метрические характеристики близости границ четырехугольника и кластера. Сложность вычисления $\rho_{tetr}(C)$ заключается в большом объеме операций, а следовательно, и времени счета, требуемых для построения оптимального приближения. Существенное ускорение вычисления приближающих четырехугольников было получено за счет использования вместо границы кластера $\Gamma(C)$ приближающей ее ломаной $\mathcal{B}(C)$ и задания меры близости границы $\Gamma(C)$ к сторонам приближающего четырехугольника как суммы расстояний узлов $\mathcal{B}(C)$ от его сторон. Обозначим множество четырехугольников с наперед заданными углами $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ через $\mathcal{Q} = \mathcal{Q}(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, а его элементы – через Q . В качестве меры точности наилучшего приближения будем использовать функцию

$$\rho_{tetr}(C) = \min_{Q \in \mathcal{Q}} \sum_{\ell=0}^M \text{dist}(\beta_{\ell}(C), Q),$$

где $\text{dist}(p, Q)$ – евклидово расстояние от точки p до границы четырехугольника Q . Четырехугольник, минимизирующий $\rho_{tetr}(C)$, обозначим через $Q_{opt} = Q_{opt}(C)$.

Определение бинарной переменной $\varphi_{tetra}(C)$ только по значению $\rho_{tetra}(C)$ дает далеко не самые лучшие результаты. Для получения надежных результатов выделения зданий использовалось несколько условий, при одновременном выполнении которых выполняется соотношение $\varphi_{tetra}(C) = 1$, в противном случае – $\varphi_{tetra}(C) = 0$:

1. Условие, ограничивающее периметр кластера:

$$2(u + v) \leq |\Gamma(C)|, \quad (1)$$

где u, v – длины сторон минимального ограничивающего прямоугольника.

2. Условие, ограничивающее снизу длину наименьшей u_1 и второй по величине u_2 сторон четырехугольника:

$$u \leq u_1; v \leq u_2. \quad (2)$$

3. Условие, ограничивающее отклонение площади Q_{opt} от площади кластера C :

$$|Q_{opt}| \geq \tau_{square} |C|, \quad 0 < \tau_{square} < 1. \quad (3)$$

4. Условие, гарантирующее заданную точность аппроксимации для трех сторон u_2, u_3, u_4 четырехугольника Q_{opt} из четырех, для которых количество $n_1 \leq n_2 \leq n_3 \leq n_4$ пикселей границы $\Gamma(C)$, лежащих на расстоянии меньше ε от них, наибольшее:

$$n_\ell \geq \tau_{appr} |u_\ell|, \quad \ell = 2, 3, 4, \quad 0 < \tau_{appr} < 1. \quad (4)$$

Несмотря на несколько громоздкую формулировку данных четырех условий, они легко программируются, их проверка занимает доли миллисекунды. Условия 1 и 2 проверяют соответствие геометрических размеров Q_{opt} допустимым размерам крыш. Условие 3 оценивает близость формы кластера к форме четырехугольника. Чем больше форма кластера C похожа на форму четырехугольника, тем меньше отличаются площади $|Q_{opt}|$ и $|C|$. Условие 4 проверяет близость трех сторон Q_{opt} к C . Оно работает лучше, чем его версия для четырех сторон, так как позволяет выделять частично невидимые здания. Переменная $\varphi_{tetra}(C) = 1$, если одновременно выполняются все вышеперечисленные условия, в противном случае $\varphi_{tetra}(C) = 0$.

Признак $\rho_{tetra}(C)$ позволяет выделить только те здания, крыши (или части крыш) которых имеют форму, близкую к форме четырехугольника. Наряду с $\rho_{tetra}(C)$ имеет смысл использовать еще один признак $\rho_{par\&ort}(C)$, который анализирует аппроксимирующую ломаную $\mathfrak{B}(C)$, находит в ней длинные параллельные и перпендикулярные отрезки, обладающие наибольшей суммарной длиной $s_{par\&ort}(C)$, и сравнивает ее с длиной границы $\Gamma(C)$. Если для наперед заданного числа $0 < \tau_{par\&ort} < 1$

$$s_{par\&ort}(C) \geq \tau_{par\&ort} |\Gamma(C)|,$$

значение сопряженной бинарной переменной полагается $\varphi_{par\&ort}(C) = 1$, в противном случае $\varphi_{par\&ort}(C) = 0$.

При наличии на изображении теней точность выделения зданий повысится, если использовать еще и признак принадлежности кластера тени $\rho_{shad}(C)$. Значение признака может вы-

числяться, например, как доля кластера, попадающего в тень $\rho_{shad}(C) = |C_{shad}|/|C|$. Значение связанной с ним бинарной переменной $\varphi_{shad}(C)$ полагается равным единице, если кластер имеет малое пересечение с областью тени, и равным нулю в противном случае.

На основе описанных признаков $\rho_{size}(C)$, $\rho_{iso}(C)$, $\rho_{tetr}(C)$, $\rho_{par\&ort}(C)$, $\rho_{shad}(C)$ в большинстве случаев удастся с высокой точностью выделить большую часть (но не все) зданий. На изображениях могут присутствовать здания нестандартной формы или образующие кластеры с недостаточно гладкой границей, поэтому имеет смысл использовать цвет найденных с помощью вышеуказанных признаков зданий для поиска оставшихся независимо от наличия априорной информации о цвете крыш зданий.

Обозначим через $\rho_{col}(C)$ средний RGB-цвет кластера. Этот признак используется для сравнения по цвету кластеров, не распознанных на основе признаков $\rho_{size}(C)$, $\rho_{iso}(C)$, $\rho_{tetr}(C)$, $\rho_{par\&ort}(C)$, $\rho_{shad}(C)$, с кластерами, которые были распознаны с их помощью как части зданий. Значение бинарной переменной $\varphi_{col}(C)$ полагается равным единице, если значение $\rho_{col}(C)$ близко к значению цвета $\rho_{col}(C')$ одного из кластеров C' , распознанных как часть здания, в противном случае $\varphi_{col}(C) = 0$. В качестве меры близости можно выбрать расстояние в каком-либо цветовом пространстве, нечувствительном к изменению освещенности. Удовлетворительный результат дало сравнение признаков (цветов) $\rho_{col}(C)$ с помощью коэффициента корреляции Пирсона, применение которого оправдано его инвариантностью к добавлению к вектору константы и умножению вектора на число. По аналогии с другими переменными значение $\varphi_{col}(C)$ устанавливается по порогу: $\varphi_{col}(C) = 1$, если $\rho_{col}(C) \geq \tau_{col}$, $0 < \tau_{col} < 1$, и $\varphi_{col}(C) = 0$ в противном случае.

Построение классификатора $\mathbf{D} = \mathbf{D}(C)$, принимающего решение о принадлежности кластера зданию для поставленной задачи при выполнении условия 1, проводится на основе значений вектора наличия признаков

$$\Phi(C) = (\varphi_{size}(C), \varphi_{iso}(C), \varphi_{tetr}(C), \varphi_{par\&ort}(C), \varphi_{shad}(C), \varphi_{col}(C)).$$

Значение классификатора \mathbf{D} равно единице, если кластер принадлежит зданию, и нулю в противном случае. Как упоминалось выше, признаки ρ и соответствующие им бинарные переменные φ строились так, чтобы каждый признак выделял часть зданий с как можно меньшей ошибкой второго рода (ложного обнаружения). При этом переменные $\varphi_{size}(C)$, $\varphi_{iso}(C)$, $\varphi_{tetr}(C)$, $\varphi_{par\&ort}(C)$, $\varphi_{shad}(C)$ – базовые в том смысле, что каждая из них вычисляется независимо от других. Переменная $\varphi_{col}(C)$ – производная от базовых признаков, так как ее значение определяется на основе цвета уже классифицированных кластеров. Значения классификатора задаются несложной логической формулой. Введем вспомогательный классификатор

$$\mathbf{D}_1(C) = \varphi_{size}(C) \wedge \varphi_{iso}(C) \wedge \varphi_{shad}(C) \wedge [\varphi_{tetr}(C) \vee \varphi_{par\&ort}(C)],$$

где $a \wedge b = \min(a, b)$ и $a \vee b = \max(a, b)$, который выделяет часть зданий на основе базовых бинарных переменных (без анализа цвета) и который может быть использован как самостоятельное решающее правило. Тогда классификатор

$$\mathbf{D}(C) = \mathbf{D}_1(C) \vee \{[1 - \mathbf{D}_1(C)] \wedge \varphi_{size}(C) \wedge \varphi_{iso}(C) \wedge \varphi_{shad}(C) \wedge \varphi_{col}(C)\}.$$

Нетрудно видеть, что классификатор $\mathbf{D}_1(C)$, выделяющий часть зданий без анализа цвета, может быть реализован в виде дерева, а $\mathbf{D}(C)$ – в виде ациклического графа.

Решение задачи при условии 2 имеет смысл выполнять с помощью того же классификатора \mathbf{D} с тем лишь отличием, что при вычислении значения переменной $\varphi_{col}(C)$ нужно исполь-

зывать имеющийся набор цветов зданий. Если производить классификацию только по заданному набору цветов, с большой вероятностью появляются ложно обнаруженные здания.

2. Эксперименты

Для экспериментов использовались цветные космические изображения европейского городского ландшафта с разрешением примерно 2 м/пиксел. Помимо кластеризованного изображения, которое вычислялось с помощью разработанных авторами алгоритмов максимального потока сети [12], в качестве входных данных для алгоритма использовались бинарные изображения теней, построенные на основе алгоритмов [14]. Выделение зданий на цветном изображении размером 1500×1500 пикселов с помощью неоптимизированной реализации алгоритма занимает около 4 с.

В данной версии алгоритма присутствуют шесть параметров: τ_{iso} , τ_{square} , τ_{appr} , $\tau_{par\&ort}$, τ_{col} и точность аппроксимации ε . Значения параметров выбирались на первых тестовых изображениях так, чтобы получить наименьшее число ошибок первого и второго рода. Например, на исходном изображении (см. рис. 1, а) при оптимальном выборе значений параметров были обнаружены 100 % зданий из 61 при двух кластерах, ошибочно выделенных в качестве элементов крыш (рис. 2).

Результаты тестирования на четырех аэрокосмических изображениях показало, что среднее число пропущенных зданий составляет менее 5 % от общего их количества. Средний процент ложно обнаруженных кластеров (которые распознаны как принадлежащие зданиям) не превосходит 15 % от общего числа кластеров. Большая часть ложно распознанных кластеров может быть отброшена при последующей обработке путем оценки их геометрических характеристик.

Замечание. Полученные кластерные представления также дают дополнительную возможность удобного выделения здания вручную. Для этого достаточно выделить (например, с помощью мыши) по одному пикселу в каждом кластере, принадлежащем крыше здания. При этом нередко крыша представляется одним или, самое большое, четырьмя-пятью кластерами.



Рис. 2. Результат автоматического выделения зданий для исходного изображения (см. рис. 1, а)

Очевидно, что необходимость задавать значения параметров затрудняет работу с алгоритмом. Однако заметим, что, во-первых, большинство известных нам алгоритмов, обрабатывающих изображения, используют параметры, значения которых приходится подбирать эмпирически, и, во-вторых, алгоритм достаточно быстродействующий, поэтому процесс выбора параметров можно организовать в виде обучающей процедуры.

Для исследования устойчивости алгоритма от значений параметров было проведено тестирование космических изображений при значениях параметров, выбранных для других космических изображений того же типа. Например, значения параметров, найденных при выделении зданий на космическом изображении части города, использовались для решения этой же задачи на другом космическом изображении другой части этого же города. Точность классификации в таких случаях достигала 90 %. В будущем планируется провести исследования, позволяющие облегчить выбор значений параметров, а также повысить точность выделения зданий.

Заключение

В статье предложен новый подход к задаче автоматического обнаружения зданий на цветных аэрофотоснимках и космических изображениях. Его специфика заключается в применении кластерных представлений изображений городского ландшафта, дающих возможность выделять границы кластеров, а затем строить приближающие их ломаные. Использование границ кластеров и приближающих их ломаных повышает точность решения задачи и сокращает время, необходимое для ее решения. Задача обнаружения зданий на изображении размером 1500×1500 пикселей решается на одном процессоре Intel Core 2 2.13GHz менее чем за 4 с с точностью более 90 %. Помимо этого, кластерное представление облегчает процесс интерактивного исправления ошибок и интерактивного выделения зданий.

Второй особенностью подхода является использование нескольких критериев, каждый из которых позволяет выделить часть зданий с близкой к нулю вероятностью ошибки ложного обнаружения.

К недостаткам программной реализации подхода следует отнести необходимость задавать значения параметров, которые, однако, не требуют значительных изменений на однотипных изображениях.

В будущем планируется усовершенствовать алгоритм обнаружения зданий, чтобы облегчить и по возможности исключить процесс ручного задания параметров, ускорить работу программной реализации, а также усовершенствовать решающее правило для повышения надежности получаемых результатов.

Список литературы

1. Mayer, H. Automatic Object Extraction from Aerial Imagery – A Survey Focusing on Buildings / H. Mayer // *Computer Vision and Image Understanding*. – 1999. – Vol. 74, № 2. – P. 138–149.
2. SURF: Speeded Up Robust Features / H. Bay [et al.] // *Computer Vision and Image Understanding (CVIU)*. – 2008. – Vol. 110, № 3. – P. 346–359.
3. Lowe, D. Object recognition from local scale-invariant features / D. Lowe // *Proc. of the Intern. Conf. on Computer Vision*. – Washington, DC, USA, 1999. – Vol. 2. – P. 1150–1157.
4. Scovanner, P. A 3-dimensional sift descriptor and its application to action recognition / P. Scovanner, S. Ali, M. Shah // *Proc. of the 15th Intern. Conf. on Multimedia*. – N.Y., USA, 2007. – P. 357–360.
5. Lacoste, C. Unsupervised line network extraction in remote sensing using a polyline process / C. Lacoste, X. Descombes, J. Zerubia // *Pattern Recognition*. – 2010. – Vol. 43 (4). – P. 1631–1641.
6. Lafarge, F. Geometric Feature Extraction by a Multi-Marked Point Process / F. Lafarge, G. Gimel'farb, X. Descombes // *IEEE Trans. Pattern Analysis and Machine Intelligence*. – 2010. – Vol. 32 (9). – P. 1597–1609.
7. Benedek, C. Building Development Monitoring in Multitemporal Remotely Sensed Image Pairs with Stochastic Birth-Death Dynamics / C. Benedek, X. Descombes, J. Zerubia // *IEEE Trans. Pattern Analysis and Machine Intelligence*. – 2011. – Vol. 99.
8. Annotated Computer Vision Bibliography: Table of Contents [Electronic resource]. – Mode of access : [http://www.visionbib.com/bibliography/keyword/b.html#Building Recognition](http://www.visionbib.com/bibliography/keyword/b.html#Building%20Recognition). – Date of access : 01.11.2011.
9. Yang, X. *Urban Remote Sensing: Monitoring, Synthesis and Modeling in the Urban Environment* / X. Yang. – Blackwell : John Wiley & Sons, 2011. – 408 p.

10. Zalesky, B.A. Integer Programming Methods in Image Processing and Bayes Estimation / B.A. Zalesky // *Studies in Fuzziness and Soft Computing*. – 2007. – Vol. 210. – P. 415–444.
11. Zalesky, B.A. Gibbs Classifiers / B.A. Zalesky // *Probability Theory and Mathematical Statistics*. – 2004. – Vol. 70. – P. 36–46.
12. Прадун, Д.В. Фильтрация и кластеризация мультиспектральных изображений с помощью алгоритма максимального потока в сети на основе вычисления градиента / Д.В. Прадун, Б.А. Залесский // *Информатика*. – 2010. – № 3 (27). – С. 73–80.
13. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М. : Техносфера, 2005. – 1070 с.
14. Лукашевич, П.В. Выделение теней на изображениях с помощью анализа гистограмм / П.В. Лукашевич, Б.А. Залесский // *Информатика*. – 2011. – № 1 (29). – С. 118–128.
15. Latecki, L.J. Application of Planar Shape Comparison to Object Retrieval in Image Databases / L.J. Latecki, R. Lakaemper // *Pattern Recognition*. – 2002. – Vol. 35 (1). – P. 15–29.
16. Bai, X. Detection and Recognition of Contour Parts Based on Shape Similarity / X. Bai, X. Yang, L.J. Latecki // *Pattern Recognition*. – 2008. – Vol. 41, № 7. – P. 2189–2199.

Поступила 28.07.11

*Объединенный институт проблем
информатики НАН Беларуси,
Минск, Сурганова, 6
e-mail: pavel.lukashevich@newman.bas-net.by,
zalesky@newman.bas-net.by*

P.V. Lukashevich, B.A. Zalesky

BUILDING DETECTION ON CLUSTERED AERIAL AND SPACE IMAGES

An approach to the problem of building detection on aerial and space images is presented. It is based on cluster representation of the images. The approach allows detecting buildings on color images with resolution higher than 3–4 meter per pixel. Its program implementation is faster than algorithms known to authors. High reliability and speed of the realization resulted from the use of cluster boundaries and their simplified representations. Another distinction of the approach is simultaneous application of several criteria, each of which detects a part of buildings.