

гарей относится малая надёжность детектирования гарей, которые могут довольно слабо отличаться от водоемов, вспаханной земли или других слабо отражающих поверхностей. Таким образом, два метода распознавания дополняют друг друга.

Следует отметить, что при обработке очередного изображения части поверхности, соответствующие ДАП, аккумулярованным за предыдущий период (с высокой вероятностью, являющиеся горячими на этом изображении), могут быть использованы для обучения классификатора, применяемого для обработки этого же изображения. Модель различения ППО и "не-ППО", построенная по самому изображению, может оказаться более удачной, чем обобщённая модель.

Теоретическая часть. Входными данными служат мультиспектральные изображения поверхности, поставляемые MODIS. Для работы используются 22 спектральных канала, характеризующие отражательные свойства поверхности (а не тепловое излучение). 2 из этих каналов имеют разрешение 250 метров, 5 – разрешение 500 метров, а остальные 1 километр. Изображения более грубых масштабов билинейно интерполируются до 250-метрового. Таким образом, каждая точка изображения, соответствующая приблизительно 250-метровая области земной поверхности, представляется 22-мерным вектором. Обозначим $\{\vec{x}^{(i)}\}_{i=1}^N$ – обучающая

выборка, состоящая из векторов, $\langle \vec{x} \rangle = \frac{1}{N} \sum_{i=1}^N \vec{x}^{(i)}$ – средний вектор, X – матрица,

составленная из нормализованных векторов-колонок $\vec{x}^{(i)} - \langle \vec{x} \rangle$, XX^T – матрица покомпонентной ковариации, размерностью 22x22. Оказывается, что максималь-

ное и минимальное собственные значения этой матрицы $\frac{\lambda_{\max}}{\lambda_{\min}} \gg 1$, т.е. векторы

сильно коррелированы. Для выделения некоррелированного подпространства (уменьшения размерности) был использован метод главных компонент [3]. Число используемых главных компонент выбиралось из расчета, что собственные числа использованных компонент дают не менее 99,9 % нормы матрицы ковариаций, т.е.

выбиралось минимальное $k : \sum_{j=1}^k \lambda_j^2 > 0,999 \sum_{j=1}^{22} \lambda_j^2$. Типичные значения k , полу-

ченные на данных – от 7 до 10. Таким образом, размерность задачи понижается в 2–3 раза.

Множество точек гарей моделировалось многомерным нормальным распределением $p(\vec{x}) = C \exp\left(-(\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu})\right)$, где Σ и $\vec{\mu}$ – ковариационная матрица и среднее подмножества обучающей выборки, состоящее из точек гарей. Таким же образом моделировалось множество "не-гарь". Обучающая выборка для класса «гарь» выбирается согласно карте ДАП, зарегистрированных на предыдущие периоды вплоть до нескольких дней до момента регистрации обрабатываемого изображения, а для класса «не гарь» – из пикселей, не попавших в область ДАП. Так как негоревшая область на несколько порядков больше сгоревшей, точки второго класса берутся со случайным интервалом, равномерно распределенным на отрезке $[0; 100000]$, т.е. в среднем каждая пятидесятитысячная. Для распознавания используется байесовский классификатор, который в случае нормальной плотности для двух классов можно записать как $g(\vec{x}) = (\vec{x} - \vec{\mu}_1)^T \Sigma_1^{-1} (\vec{x} - \vec{\mu}_1) - (\vec{x} - \vec{\mu}_2)^T \Sigma_2^{-1} (\vec{x} - \vec{\mu}_2) + C$ – разделяющая функция, C – константа, выбираемая из желаемого соотношения ошибок первого и второго рода.

Реализационная часть. При всех расчётах не должны использоваться участки поверхности, закрытые облаками, ведь в этом случае их отражающая способность не соответствует ни гарям, ни поверхности земли. Для учёта таких точек использован продукт MOD35 Cloud Mask, который генерируется для каждого набора спутниковых снимков MODIS и содержит информацию о степени закрытости облаками каждого пикселя, построенную с помощью алгоритмов автоматической обработки изображения [2]. Классификатор облачности MODIS выдаёт 4 градации присутствия облаков: 00 – облачно, 01 – сомнительно ясно, 10 – вероятно ясно, и 11 – уверенно ясно. Использование областей с характеристикой облачности "уверенно ясно" и исключение из рассмотрения всех прочих точек приводит к хорошей классификации, однако такие области составляют лишь малую часть изображения и результаты работы мало отличаются от применения только первого подхода (аккумуляция областей ДАП): многие выгоревшие участки не обрабатываются. Поэтому приходится использовать также области с частичной облачностью. Для более точного разделения был введён третий класс, условно называемый "туман". В итоге получается карта участка поверхности с разрешением 250 метров, на которой выделены участки выгоревшей земли, области невыгоревшей земли и облака.

Маска облачности содержит для каждого пикселя 48-битную запись, из которой для наших целей извлекаются следующие данные: флаг вычисления маски облачности для данной точки (1 бит); облачность (2 бита): 00 – облачно, 01 – сомнительно ясно, 10 – вероятно ясно, и 11 – уверенно ясно; флаги «день/ночь», наличия солнечного блика, наличия снега/льда – (по 1 биту); флаг «земля/вода» (2 бита): 00 – вода, 01 – побережье, 10 – пустыня, 11 – земля; флаги необлачной помехи (плотный аэрозоль), тонких перистых облаков, облачной тени. Берутся в рассмотрение только те пиксели, для которых маска облачности показывает, что: для данного пикселя она вычислена; на снимке день, нет льда, солнечного блика; флаг «земля/вода» принимает значение «11», «10» или «01»; нет аэрозоля, перистых облаков, облачной тени; значение облачности «11», то есть «ясно»; индивидуальный флаг данного 250-метрового пикселя показывает, что он не закрыт облаками. Этот подход предполагает, что не рассматриваются точки, на которых маска облачности MODIS показывает хотя бы минимальную помеху видимости поверхности. В данный момент описанный алгоритм проходит экспериментальную проверку на сериях изображений MODIS, полученных в лето 2010 года, в качестве обучающей выборки использована серия шейпфайлов, содержащих активные пожары за этот период.

Первоначально планировалось не проверять маску облачности на флаги тени, перистых облаков и аэрозоля, а также использовать те пиксели, значение облачности для которых равно «10», т.е. «вероятно, ясно». Но данный подход уже в ходе предварительных тестов показал свою неприемлемость, так как значительная площадь поверхности, закрытая дымкой, попала в обучающую выборку классификатора, чего не произошло в случае более жесткой отбраковки.

Вывод. На основании двух классификаторов: детектора активных пожаров MODIS и классификатора гарей (представляющего собой байесовский с многомерным нормальным распределением) построена комбинированная система распознавания гарей, в которой неполные данные первого детектора дополняются результатами работы второго, причём обучение второго происходит на данных работы первого детектора для текущего обрабатываемого изображения. Это позволяет адаптивно настраивать классификатор гарей и компенсировать различия в условиях съёмки. Помимо классов «гарь» и «не гарь», вводится класс «туман», обучающую выборку для которого можно строить на зонах, отбракованных маской облачности.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Бондур В.Г.* Актуальность и необходимость космического мониторинга природных пожаров в России // Вестник ОНЗ РАН, NZ11001. – 2010 (а). – Т. 2.
2. *Ackerman S.A., Strabala K.I., Menzel W.P., Frey R.A., Moeller C.C., Gumley L.E., Baum B.A., Schaaf C., Riggs G.* 1997: Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35). EOS ATBD web site. – 125 p.
3. *Айвазян С.А., Бухитабер В.М., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика: классификация и снижение размерности. – М.: Финансы и статистика, 1989.
4. *Журавлёв Ю.И.* Распознавание. Классификация. Прогноз. Математические методы и их применение. Вып. 2. – М.: Наука, 1989.

Статью рекомендовал к опубликованию д.ф.-м.н. В.И. Цурков.

Бондур Валерий Григорьевич – Научно-исследовательский институт аэрокосмического мониторинга «Аэрокосмос»; e-mail: office@aerocosmos.info; 105064, г. Москва, Гороховский пер. 4; тел.: +74956321654; директор; академик РАН; д.т.н.; профессор.

Матвеев Иван Алексеевич – Федеральное государственное бюджетное учреждение науки Вычислительный центр им. А.А. Дородницына Российской академии наук; e-mail: matveev@ccas.ru; 119333, г. Москва, ул. Вавилова, 40; тел.: 84991352228; отдел сложных систем; зав. сектором; к.ф.-м.н.

Мурынин Александр Борисович – e-mail: AMurynin@bk.ru; тел.: +79266902722; старший научный сотрудник; к.т.н.

Трёкин Алексей Николаевич – Московский физико-технический институт (государственный университет); e-mail: jagdpanzerIV@yandex.ru; 141700, Московская область, г. Долгопрудный, Институтский переулок, 9; тел.: 84954084554; отдел сложных систем; факультет управления и прикладной математики; студент.

Bondur Valery Grigoryevich – Research Institute for Aerospace Monitoring “Aerocosmos”; e-mail: office@aerocosmos.info; 4, Gorokhovskij, Moscow, 105064, Russia; phone: +74956321654; director, academician of Russian Academy of Sciences; dr. of eng. sc.; professor.

Matveev Ivan Alexeevich – Institution of Russian Academy of Sciences Dorodnicyn Computing Centre of RAS; e-mail: matveev@ccas.ru; 40, Vavilov street, Moscow, 119333, Russia; phone: +74991352228; the department of complex systems; head of sector; cand. of phis.-math. sc.

Murynin Alexander Borisovich – e-mail: AMurynin@bk.ru; 40 phone: +79266902722; senior researcher; cand. of eng. sc.

Trekin Alexej Nikolaevich – Moscow Institute of Physics and Technology; e-mail: jagdpanzerIV@yandex.ru; 9, Institutskii per., Dolgoprudny, Moscow Region, 141700, Russia; phone: +74954084554; the department of complex systems; student.

УДК 621.391.25

М.В. Стремоухов, С.В. Чистяков

**МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ФОРМИРУЮЩЕЙ СИСТЕМЫ,
УЧИТЫВАЮЩАЯ НЕЛИНЕЙНЫЕ СТАТИСТИЧЕСКИЕ ЗАВИСИМОСТИ
ИЗОБРАЖЕНИЙ**

Рассматривается проблема повышения качества при эффективном кодировании изображений с учетом нелинейных статистических зависимостей между их элементами. Предлагается способ параметрического кодирования изображений на основе математической модели нелинейной формирующей системы при заданной возбуждающей последовательности, обеспечивающей наилучшее по выбранному критерию совпадение синтезированного изображения с исходным. Авторы указывают на достоинства предлагаемого под-