

К.Т. Протасов

ОБНАРУЖЕНИЕ ТЕПЛОВЫХ АНОМАЛИЙ (ПОЖАРОВ) ПО ДАННЫМ ПРИБОРОВ AVHRR И МЕТЕОСЛУЖБ В УСЛОВИЯХ ВЛИЯНИЯ АТМОСФЕРЫ

Рассмотрена двухэтапная процедура обнаружения тепловых аномалий типа пожаров на территории региона по спутниковым данным прибора AVHRR. На первом этапе восстанавливается поле термодинамической температуры подстилающей поверхности с помощью нелинейного непараметрического уравнения регрессии, адаптируемого к конкретным условиям наблюдения по согласованным данным метеослужб в пунктах расположения метеостанций и данным прибора AVHRR спутника NOAA. На втором этапе, на основе восстановленного поля температуры, строится адаптивное байесово правило обнаружения тепловых аномалий, основанное на принципе идентификации составляющих смешанного распределения и аппроксимации условных функций плотности кривыми Джонсона. Приведен пример выделения тепловых аномалий по спутниковым видеоданным для территории Томской области.

Проблема раннего обнаружения пожаров, размеры которых невелики, чрезвычайно актуальна, особенно для труднодоступных и малонаселенных районов. В связи с этим возникает необходимость в использовании спутниковой информации для решения задачи раннего и оперативного обнаружения источников тепловых аномалий (возможных пожаров). Среди ИСЗ, наиболее часто появляющихся над данной местностью и позволяющих решать задачи мониторинга подстилающей поверхности Земли (ППЗ), естественно ориентироваться на спутники серии NOAA и использовать данные прибора AVHRR, который позволяет фиксировать радиояркости в пяти спектральных диапазонах, включая тепловые, в виде изображений. К сожалению, низкая разрешающая способность прибора AVHRR и сравнительно небольшой диапазон фиксируемых им радиояркостей не позволяют эффективно решать задачу раннего обнаружения малоразмерных пожаров.

Рассмотрим подход, основанный на методах предсказания ненаблюдаемых величин (значений термодинамической температуры) по информации, содержащейся в косвенных измерениях, которыми в данном случае будут пятиканальные наблюдения прибора AVHRR. Заметим, что для эффективного решения упомянутой задачи требуется высокая точность картографической привязки сканерных снимков. Привязка видеоданных по орбитальным данным не приводит к желаемой точности, и прежде всего из-за ошибок определения орбитальных элементов, дрейфа углов ориентации ИСЗ. В связи с этим орбитальная привязка является лишь предварительной, которая уточнялась по опорным точкам, назначаемым оператором.

Как известно, собственное излучение объекта является функцией температуры, физических свойств и структурных характеристик излучающей

поверхности. Содержательной характеристикой собственного некогерентного излучения следует считать спектральную энергетическую яркость нагретого тела, которая может быть определена по формуле Планка. Энергетическая яркость зависит от температуры, длины волны и некоторого множителя $\epsilon_\lambda(T)$ (коэффициента излучения), называемого спектральной степенью черноты излучающей поверхности при заданной температуре и определенном направлении визирования. Так, абсолютно черное тело имеет степень черноты $\epsilon_\lambda = 1$ во всем диапазоне длин волн. Для серого тела в определенном диапазоне длин волн $0 < \epsilon_\lambda < 1$. Для некоторых типов подстилающей поверхности параметр $\epsilon_\lambda(T)$ известен, но реально прибор регистрирует радиояркости их комбинаций, кроме того, свое искажающее влияние оказывает состояние атмосферы. Это создает проблему в оценивании температуры поверхности Земли по радиояркостям, наблюдаемым в условиях влияния атмосферы.

Будем полагать, что проведена предварительная обработка изображений: устранены геометрические искажения данных; осуществлена географическая (координатная) привязка; «вырезан» фрагмент видеоданных, содержащий территорию Томской области (ТО) и ее окрестности; произведены радиационная коррекция и калибровка данных прибора AVHRR переходом к альбедо для 1-го и 2-го каналов и к термодинамической температуре в 3, 4, 5-м каналах (по обращенной формуле Планка).

Таким образом, имеем поле 1024×1024 5-мерных векторов и для каждого элемента-пиксела этого изображения (поля видеоданных) заданы географические координаты. Кроме того, имеются данные метеослужб о температуре (это может быть температура почвы или приземного слоя воздуха) ППЗ в некоторых достаточно равномерно расположенных

пунктах с координатной привязкой на момент пролета спутника. В идеальном случае необходимо знать температуру в контрольных точках для каждого класса (типа) поверхности (вода, поле, пашня, лес и т.п.).

Восстановление поля термодинамической температуры

Термодинамическую температуру ППЗ прогнозируемого поля будем описывать случайной величиной $Y \in R^1$, а радиояркости полей, являющихся источниками прогнозирующей информации, будем описывать случайным вектором $\mathbf{X} \in R^k$, где R^k – 5-мерное евклидово пространство; $\mathbf{X} = (X^1, \dots, X^k)^T$, X^i – радиояркости i -го канала прибора AVHRR, $i = 1, \dots, k$, T – знак транспонирования.

Взаимосвязь прогнозируемой переменной Y и вектора \mathbf{X} будем описывать функционалом регрессии следующего вида:

$$m(\mathbf{x}) = E(Y|\mathbf{X} = \mathbf{x}), \quad (1)$$

где $E(\cdot)$ – оператор математического ожидания, причем $E(|Y|) < \infty$.

Если существуют нижеследующие плотности вероятностей случайных величин \mathbf{X} и Y , то с учетом (1) имеем

$$y = m(\mathbf{x}) = \int_{R^1} y \frac{f(\mathbf{x}, y)}{f(\mathbf{x})f(y)} dF(y), \quad (2)$$

где $\mathbf{x} \in R^k$, $y \in R^1$, $f(\mathbf{x}, y)$ – совместная плотность вероятностей случайных вектора \mathbf{X} и величины Y ; $f(\mathbf{x})$ – плотность вероятности случайного вектора \mathbf{X} ; $f(y)$ – плотность вероятности случайной величины Y , а $F(y)$ – интегральная функция распределения Y .

Пусть в нашем распоряжении имеется выборка попарно независимых одинаково распределенных случайных величин $\{(\mathbf{X}_i, Y_i)\}_{i=1}^n$, где n – количество контрольных отсчетов температур по данным метеослужб на момент фиксирования спутниковых наблюдений. В этом случае для вычисления выражения (2) естественно воспользоваться непараметрическими оценками неизвестных распределений по выборочным данным [1–3]. Заменим неизвестные распределения их непараметрическими оценками ядерного типа, а $F(y)$ – эмпирической функцией $F_n(y)$, тогда оценка уравнения регрессии (2) примет вид [3]:

$$\hat{m}(\mathbf{x}) = \frac{Y_i \sum_{j=1}^n K_h(Y_i - Y_j) \prod_{i=1}^k K_h(x_i - X_j^i)}{\sum_{i=1}^n \sum_{j=1}^n \prod_{i=1}^k K_h(x_i - X_j^i) \sum_{i=1}^n K_h(Y_i - Y_j)}, \quad (3)$$

где h – ширина окна (параметр сглаживания), описываемого функцией $K_h(u) = h^{-1} K(u/h)$; в качестве $K(\cdot)$ может быть взято ядро Епанечникова [1–3] или гауссово ядро.

Накопленный практический опыт использования подобных оценок показывает, что точностные характеристики уравнения регрессии $\hat{m}_h(\mathbf{x})$ в большей степени определяются не формой ядра, а параметром масштаба h . Ввиду важности параметра h естественно перейти к векторному параметру $\mathbf{h} = (h^1, \dots, h^k)^T$ и в выражении (3) использовать модифицированное ядро $K'_h(u) = (h^i)^{-1} K(u/h^i)$, $i = 1, \dots, k$, при этом возникает задача оценивания \mathbf{h} с учетом наблюдений $\{(\mathbf{X}_i, Y_i)\}_{i=1}^n$. Для оценивания \mathbf{h} воспользуемся методом скользящего контроля, смысл которого заключается в том, что строится модифицированная оценка регрессии $\hat{m}_{h,j}(\mathbf{X}_j)$, в которой пропускается (исключается) j -е наблюдение. Критерий качества оценивания \mathbf{h} зависит от способности предсказывать набор значений $\{Y_j\}_{j=1}^n$ по наборам подвыборок $\{(\mathbf{X}_i, Y_i)\}_{i \neq j}$:

$$J(\mathbf{h}) = n^{-1} \sum_{j=1}^n [Y_j - \hat{m}_{h,j}(\mathbf{X}_j)]^2 w(\mathbf{X}_j), \quad (4)$$

где $w(\cdot)$ – весовая функция, которую, в простейших случаях, можно и не использовать. Задачу оптимизации (4) удобно осуществить поисковым методом [4], с помощью двухэтапной процедуры оценивания глобального экстремума функционала (4). На первом этапе в области поиска, представляющей

собой многомерный квадрат $\prod_{i=1}^k [h_{\min}^i, h_{\max}^i]$, где h_{\min} и

h_{\max} – нижняя и верхняя оценочные границы параметра сглаживания соответственно, случайно с равномерным распределением фокусируется точка, затем из этой точки осуществляется градиентный спуск, при этом используются поисковые методы адаптации [4]. С этой целью осуществляется вариация функционала качества (4) по параметрам сглаживания следующим образом. Вычисляются значения приращений функционала (4)

$$J_+(\mathbf{h}, a) = (J(\mathbf{h} + a\mathbf{e}_1), \dots, J(\mathbf{h} + a\mathbf{e}_k)),$$

$$J_-(\mathbf{h}, a) = (J(\mathbf{h} - a\mathbf{e}_1), \dots, J(\mathbf{h} - a\mathbf{e}_k)),$$

где k – количество параметров h , соответствующих количеству компонент вектора, собранных в вектор параметров $\mathbf{h} = (h^1, \dots, h^k)^T$; a – скалярный параметр, определяющий величину поискового шага; $\mathbf{e}_i = \left(\underbrace{0, \dots, 1, \dots, 0}_i \right)^T$, $i = 1, \dots, k$, – базисные векторы поисковых направлений. Оценочное значение градиента вычислим следующим образом:

$$J_+(\mathbf{h}, a) J_-(\mathbf{h}, a)/(2a) = \nabla_{h_{\pm}} J(\mathbf{h}, a),$$

где $\nabla_{h_{\pm}}$ – обозначение градиента. Поисковый алгоритм адаптации в рекуррентной форме имеет вид

$$\mathbf{h}[j] = \mathbf{h}[j-1] - \gamma[j] \nabla_{h_{\pm}} J(\mathbf{h}[j-1], a[j]), \quad (5)$$

выбор поискового $a[\cdot]$ и рабочего $\gamma[\cdot]$ шага рассмотрен в [4], причем $\gamma[\cdot] < a[\cdot]$.

После того как параметр \mathbf{h} в выражении (3) для $\hat{m}_h(\mathbf{x})$ конкретизирован, уравнение регрессии можно использовать для восстановления значений Y по наблюдаемым \mathbf{X} и для всего поля видеоданных. Следует заметить, что регрессионная модель прогнозирования ненаблюдаемых значений будет работать лишь при соблюдении статистической однородности данных, по которым производилось восстановление зависимости и которые восстанавливаются. С этой целью необходим предварительный анализ «полного» изображения алгоритмом сегментации, который позволяет выделить статистически однородные участки и исключить поля облачности. Далее на каждом из этих участков необходимо по выборочным данным восстановить локальное выражение (3) со своими значениями \mathbf{h} . С помощью восстановленной зависимости (3) оценивается термодинамическая температура каждого пиксела поля 1024×1024 , включающего изображение ТО. На рис. 1 приведен пример распределения восстановленной таким образом термодинамической

температуры на территории ТО по данным метеосводок в 16 опорных пунктах наблюдений. Влияние неоднородности атмосферы учитывается в модели (3) использованием конкретно измеренных значений температур участков ППЗ, радиояркости которых одновременно фиксируются прибором AVHRR.

Решающее правило обнаружения тепловых аномалий

Учитывая тот факт, что регистраторы теплового излучения прибора AVHRR способны зарегистрировать температуру ППЗ лишь до 45°C , мы можем располагать статистикой, описывающей наблюдения в ситуации, когда тепловые аномалии отсутствуют: класс «фон» (Ф), а наблюдения в ситуации, когда тепловые аномалии присутствуют: класс «тепловая аномалия» (ТА) будут «урезаны» уровнем регистрации прибора. Поэтому необходимо тщательно восстановить распределение фона, а порог t решающего правила определить уровнем ошибки

ложной тревоги $\alpha = \int_t^{\infty} f(y) dy$, где $f(y)$ – функция

плотности вероятностей распределения фона. Обозначим значения полученных на первом этапе температур Y_i , $i = 1, \dots, n$, и восстановим их вероятностное распределение, воспользовавшись параметрическим семейством Джонсона [5], выбрав аппроксимацию S_B .

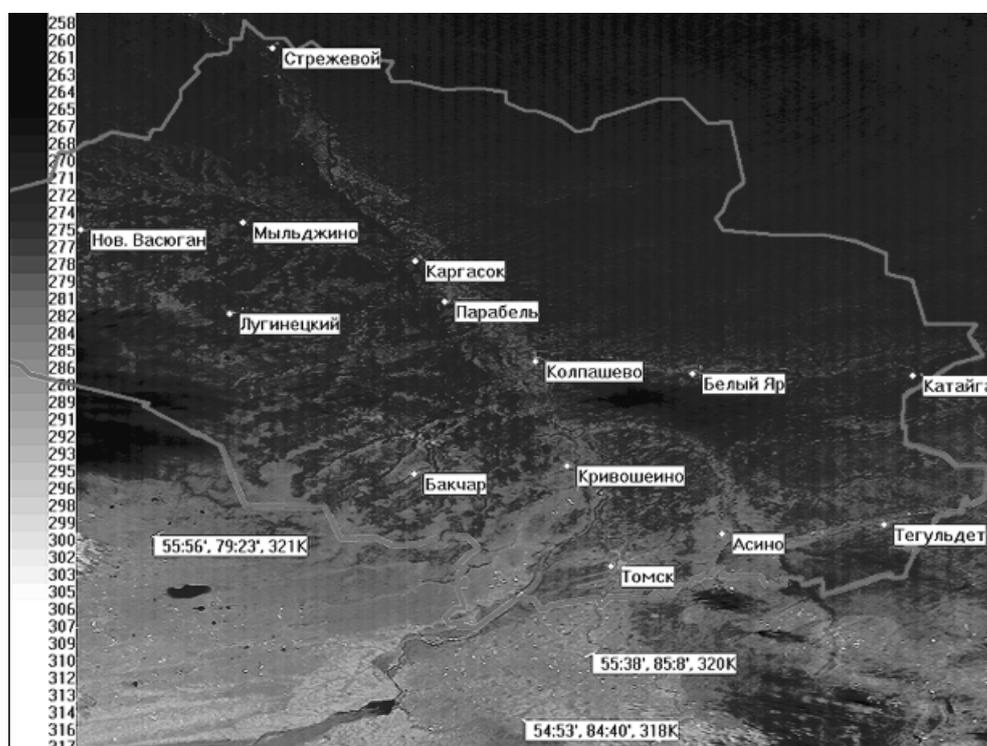


Рис. 1. Фрагмент температурной карты Томской области. Градации температуры выражены оттенками цветовой палитры (кельвины), представленной слева

Для построения решающего правила воспользуемся следующим выражением [5] для $f(y)$, а именно:

$$f_b(y) = \frac{\eta}{\sqrt{2\pi(y-\varepsilon)(\lambda-y+\varepsilon)}} \exp\left\{-\frac{1}{2}\left[\gamma + \eta \ln\left(\frac{y-\varepsilon}{\lambda-y+\varepsilon}\right)\right]^2\right\}, \quad (6)$$

$$\varepsilon \leq y \leq \varepsilon + \lambda, \quad \eta > 0, \quad -\infty < \gamma < \infty, \\ \lambda > 0, \quad -\infty < \varepsilon < \infty,$$

где ε (нижняя граница значения) и y , λ (размах выборки) могут быть оценены по наблюдениям $\{Y_j\}_{j=1}^n$, а параметры формы η , γ необходимо определить, например, методом максимального правдоподобия [5]:

$$\hat{\delta} = \left\{ \sqrt{\frac{1}{n} \sum_{i=1}^n \tau^2(Y_i; \varepsilon, \lambda) - \left(\frac{1}{n} \sum_{i=1}^n \tau(Y_i; \varepsilon, \lambda)\right)^2} \right\}^{-1}, \\ \hat{\gamma} = - \left[\frac{1}{n} \sum_{i=1}^n \tau(Y_i; \varepsilon, \lambda) \right] \hat{\delta}$$

где $\tau(y; \varepsilon, \lambda) = \ln[(y - \varepsilon)/(\lambda + \varepsilon - y)]$. Используя выражение (6) и зафиксировав уровень ошибки α , находим порог t решающего правила обнаружения ТА.

Другой подход построения байесова решающего правила обнаружения тепловых аномалий основан на том факте, что гистограмма восстановленных термодинамических температур $\hat{f}(y)$ является смешанным распределением. Она включает распределение температур в ситуации, когда тепловые аномалии отсутствуют, и распределение температур в ситуации, когда тепловые аномалии наблюдаются на территории региона. В этом случае естественно воспользоваться подходом, связанным с декомпозицией смешанного распределения наблюдений на составляющие, входящие в смесь со своим весом. Если предположить, что состояние природы в ситуации Φ описывается функцией плотности вероятностей $f_0(y)$, а состояние природы для ситуации ТА описывается функцией плотности вероятностей $f_1(y)$ с неизвестными параметрами, тогда возникает следующая задача: по смешанной гистограмме $\hat{f}(y)$ распределения термодинамической температуры восстановить параметры составляющих смесь распределений $Pf_0(y) + Qf_1(y)$, где P , Q – априорные вероятности состояний природы Φ и ТА соответственно, причем $P + Q = 1$. Следует заметить, что на территории ТО всегда есть тепловые аномалии, связанные с наличием десятка факелов газовых месторождений, координаты которых известны. Это говорит о том, что класс ТА не пуст.

В качестве $f_0(y)$, $f_1(y)$ нами были вновь использованы распределения из семейства S_B Джонсона. При этом возникает задача оптимизации

квадратичного критерия качества идентификации смеси следующего вида:

$$J(\theta) = \frac{1}{m} \sum_{j=1}^m \{ \hat{f}(y_j) - [Pf_0(y_j) + Qf_1(y_j)] \}^2, \quad (7)$$

где m – количество отсчетов по координате y , а $\theta = (P, \varepsilon_0, \lambda_0, \gamma_0, \eta_0, \varepsilon_1, \lambda_1, \gamma_1, \eta_1)^T$ – вектор неизвестных параметров функций плотности $f_0(y)$, $f_1(y)$ из семейства S_B соответственно. Для поиска минимума в выражении (7) использовалась та же процедура адаптивного поиска (5). После того как идентификация смеси произведена и распределения $f_0(y)$ и $f_1(y)$ восстановлены, байесово решающее правило обнаружения тепловых аномалий имеет следующий вид [6]:

$$u(y) = \arg \max \{ Pf_0(y), Qf_1(y) \}, \quad (8)$$

где $u(y)$ – принимаемое решение о наличии (гипотеза H_1) или отсутствии ТА (гипотеза H_0).

Выбор порога с использованием аппроксимаций Джонсона для восстановленного температурного поля, изображенного на рис. 1, иллюстрирует рис. 2, при этом оптимальное начение критерия качества аппроксимации (7) было равно 0,065.

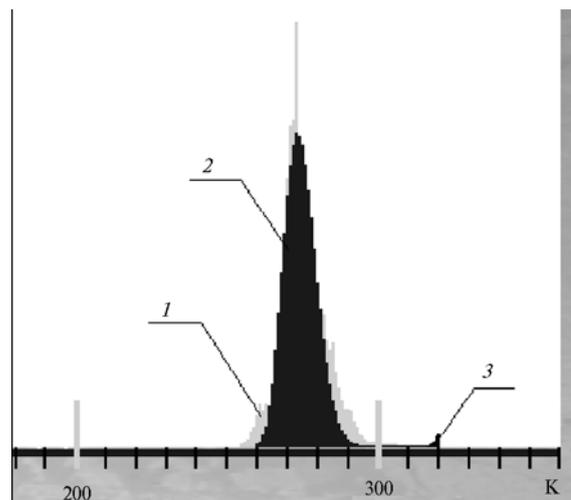


Рис. 2. Гистограмма восстановленных температур (1) и аппроксимации Джонсона S_B : 2 – при условии, когда наблюдениями является класс Φ ; 3 – при условии, когда наблюдениями является класс ТА

На рис. 3 приведены результаты обнаружения тепловых аномалий в одном сеансе NOAA, принятом 26 мая 1998 г., на котором фрагменты изображений с выделенными алгоритмом тепловыми аномалиями связаны «горячей линией» с картографической основой ТО (пакет ArcView).

При анализе видеоданных следует иметь в виду наличие артефактов, связанных с бликами на воде и облаках при низком положении Солнца, которые несложно исключить с помощью пороговых ограничителей по 1-му и 2-му каналам, где измеряется альbedo воды и облаков.

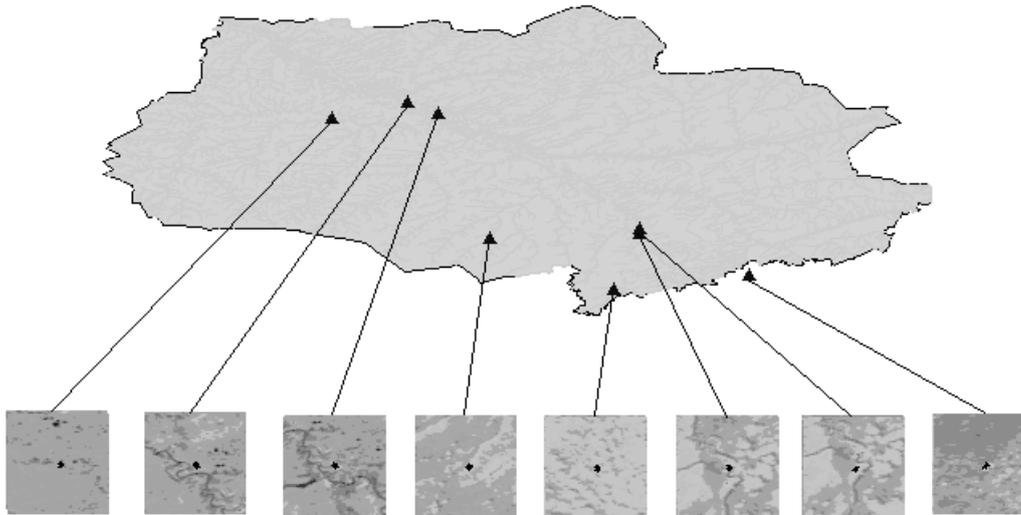


Рис. 3. Фрагменты спутниковых видеоданных с выделенными программой тепловыми аномалиями по данным прибора AVHRR спутника NOAA-14 26 мая 1998 г. и их картографическая привязка к карте ТО

1. Хардле В. Прикладная непараметрическая регрессия / Пер. с англ. М.: Мир, 1993. 349 с.
2. Прикладная статистика: Основы моделирования и первичная обработка данных: Спр. изд. / С.А. Айвазян и др. М.: Финансы и статистика, 1983. 471 с.
3. Тарасенко Ф.П. Непараметрическая статистика. Томск: Изд-во ТГУ, 1976. 294 с.
4. Цыпкин Я.З. Адаптация и обучение в автоматических системах. М.: Наука, 1968. 400 с.
5. Хан Г., Шапиро С. Статистические модели в инженерных задачах. М.: Мир, 1969. 396 с.
6. Ван Трис Г. Теория обнаружения, оценок и модуляции. Т. 1 / Пер. с англ. М.: Сов. радио, 1972. 744 с.

Институт оптики атмосферы СО РАН,
Томск

Поступила в редакцию
2 июля 1998 г.

K.T. Protasov. Detection of Thermal Anomalies (Fires) from the Data of AVHRR Instruments and Meteorological Service under Atmospheric Influence.

Two-stage procedure for detection of thermal anomalies of fire type on a region territory from the data of space-born device AVHRR is treated in the paper. The first stage assumes a reconstruction of the field of thermodynamical temperature of underlying surface using nonlinear nonparametric regression equation adaptable to particular observing conditions by coordinated data of meteorological service and the AVHRR device established in NOAA satellite. At the second stage, based on the reconstructed temperature field, the adaptive Bayes rule is constructed for detection of the thermal anomalies based on the principle of identification of the mixed distribution components and on approximation of conditional density functions by Johnson curves. An example of the thermal anomalies discernment from the space videodata on the territory of Tomsk region is presented.