

УДК 551.465

АНАЛИЗ ДАННЫХ КОСМИЧЕСКОГО МОНИТОРИНГА ПРИ ПРОГНОЗИРОВАНИИ РАСПРОСТРАНЕНИЯ ВЫЯВЛЕННЫХ АНОМАЛИЙ

О.С. Бутенко, канд. техн. наук, доц.
(Национальный аэрокосмический университет
им. Н.Е. Жуковского)

С.И. Березина, канд. техн. наук, ст. науч. сотр.
(Национальный аэрокосмический университет
им. Н.Е. Жуковского)

Г.Я. Красовский, д-р техн. наук, проф.
(Институт проблем национальной
безопасности РНБО Украины)

В статье рассматривается методика предварительной обработки космического снимка для обнаружения на нем различного рода аномалий, предварительная идентификация обнаруженного изменения и способ получения его топологических характеристик по изображению. Также рассматривается возможность выдачи предварительного краткосрочного прогноза распространения выявленных аномалий в реальных условиях по полученным топологическим характеристикам обнаруженного изменения.

У статті розглядається методика попередньої обробки космічного знімка для виявлення на ньому різного роду аномалій, попередня ідентифікація виявленої зміни й спосіб одержання його топологічних характеристик по зображенню. Також розглядається можливість видачі попереднього короткострокового прогнозу поширення виявлених аномалій у реальних умовах по отриманих топологічних характеристиках виявленої зміни.

Method of the preprocessing of the cosmic photograph is considered in item for finding on him different sort anomaly, preliminary identification of the discovered change and way of the reception its topological features on image. Also in article is considered possibility of the issue spreading revealed anomaly of the preliminary short-term forecast in real conditions on got topological features of the discovered change.

© О.С. Бутенко, С.И. Березина, Г.Я. Красовский, 2008

Методы дистанционного зондирования Земли (ДЗЗ) из космоса позволяют получать различные виды данных об объектах и явлениях в глобальном масштабе с высоким пространственным и временным разрешением. Космические снимки Земли формируют информацию о физических, химических, биологических, геометрических параметрах объектов наблюдения в различных условиях и с определенной периодичностью. Для решения задачи выявления изменений на одном и том же участке поверхности Земли при экологическом мониторинге, а так же прогнозировании его дальнейшего развития при получении данных ДЗЗ необходимо локализовать однородные пятна, а потом провести их классификацию.

Изображение представляется матрицей чисел (размер матрицы $N \times M$), где значение каждого элемента отвечает определенному уровню квантования его энергетической характеристики (яркости). Снимок может повреждаться шумами и помехами различного происхождения, например шумом видеодатчика, шумом зернистости фотоматериалов и ошибками в канале передачи. Шумы видеодатчиков или ошибки в канале передачи обычно проявляются на изображении как разрозненные изменения изолированных элементов, не обладающие пространственной корреляцией. Искаженные элементы часто весьма заметно отличаются от соседних элементов. Это наблюдение послужило основой для многих алгоритмов, обеспечивающих подавление шума.

Изображение может быть описано как распределение яркости $F(x, y)$, где x, y — пара чисел — координаты точек плоскости изображения. Будем рассматривать далее только дискретизированные изображения с целочисленными координатами.

Поскольку шум пространственно декоррелирован, в его спектре, как правило, содержатся более высокие пространственные частоты, чем в спектре обычного изображения. Следовательно, простая низкочастотная пространственная фильтрация может служить эффективным средством сглаживания шумов.

Рассмотрим массив $F(N, M)$ из $N \times M$ элементов, представляющий исходное (входное) изображение. При воздействии на него обобщенным линейным оператором получается массив из $N_1 \times M_1$ элементов, описывающий преобразованное (выходное) изображение

$$P(N_1, M_1) = \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{M_1} F(n_1, n_2) \cdot O(n_1, n_2; m_1, m_2) \quad (1)$$

где ядро оператора $O(n_1, n_2; m_1, m_2)$ представляет собой набор весовых множителей, зависящих в общем случае от координат элементов как входного, так и выходного изображений.

Применение низкочастотных фильтров влечет за собой некоторое смазывание изображения, интенсивность которого будет зависеть от размера окна в котором ведется обработка. Использование текстурных признаков для локализации пятен на снимке является нецелесообразным, так как повлечет за собой большие временные затраты на обработку снимка. Поэтому это свойство было использовано для локализации пятен однородных по текстурным и цветовым признакам. Характерные пятна проявляются в большей или меньшей степени в той или иной части спектра, что может свидетельствовать об их различной этимологии. Поэтому предварительно изображение раскладывалось на спектральные составляющие, и анализ проводился по каждой составляющей отдельно.

При анализе рассматривалось несколько типов размывающих низкочастотных фильтров. В результате установлено, что использование контргармонического среднего, арифметического среднего, геометрического среднего, гармонического среднего, максимума по окрестности, минимума по окрестности, срединной точки не дают приемлемого результата. Наилучшим оказалось использование α -усеченное среднее, медианный фильтр с размером окна 5×5 и фильтр Гаусса с $\sigma \approx 20$. Аргументируем выбор фильтров.

Медианный фильтр представляет собой скользящее окно, охватывающее нечетное число элементов изображения. Центральный элемент заменяется медианой всех элементов изображения в окне. Медианой дискретной последовательности a_1, a_2, \dots, a_N для нечетного N является тот ее элемент, для которого существуют $(N - 1)/2$ элементов, меньших или равных ему по величине, и $(N - 1)/2$ элементов, больших или равных ему по величине. Медианный фильтр не влияет на ступенчатые или пилообразные функции, что обычно является желательным свойством. Однако этот фильтр подавляет импульсные сигналы, длительность которых составляет менее половины ширины окна. Фильтр также вызывает уплощение вершины треугольной функции. Одно из основных преимуществ медианной фильтрации состоит в том, что она приводит к устранению импульсных выбросов, не размывая границ объектов.

Фильтр типа α -усеченное среднее можно представить выражением

$$P(i, j) = \frac{1}{2} \left[\max_{(i, j) \in D} \{F(i, j)\} + \min_{(i, j) \in D} \{F(i, j)\} \right] \quad (2)$$

Суть работы фильтра, представленного выражением (2), можно объяснить так.

Из локальной области D удаляются $\frac{d}{2}$ наибольших и $\frac{d}{2}$ наименьших значений пикселей. Оставшиеся значения усредняются и формируют результат.

При использовании фильтра Гаусса ядро оператора $O(n_1, n_2; m_p, m_2)$ (1) представляет собой набор весовых множителей составленных из формулы (3)

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right) \quad (3)$$

После проведенных операций удалось локализовать однородные участки (рис. 1).

Выявленные участки отмечались на исходном изображении (рис. 2). По полученным данным рассчитывались топологические характеристики пятна. А при необходимости, уже рассматривались более детальные и сложные для вычислений дешифровочные признаки.

После получения топологических характеристик обнаруженного пятна и его предварительной идентификации ставится задача выдачи кратковременного прогноза распространения обнаруженного изменения в реальных условиях на местности только по априорным данным.

Существует несколько стандартных подходов к моделированию динамических процессов. Основой моделирования является исходная информация. Зачастую эта информация представляется в виде временных рядов по некоторым параметрам исследуемого объекта. Адекватность модели определяется качеством исходной информации в соответствии с её количественным представлением. Времен-

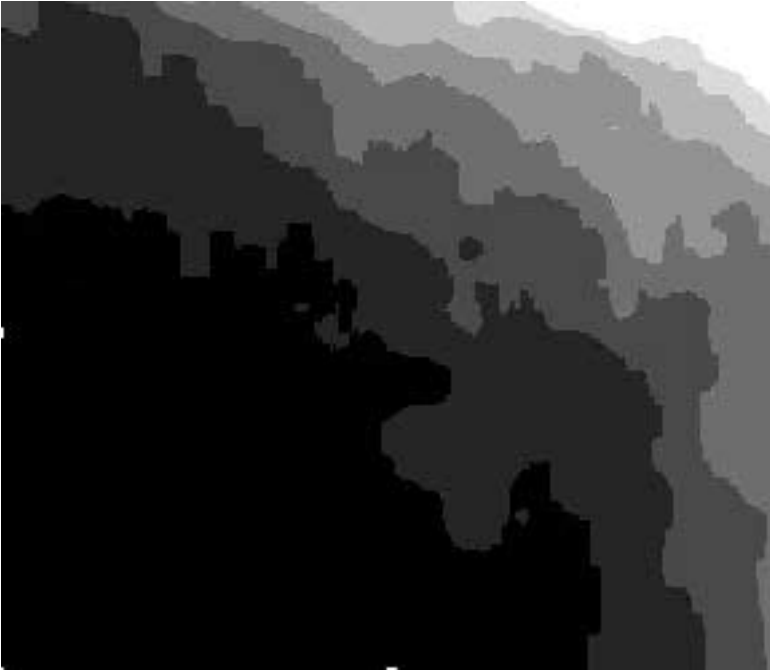


Рис. 1. Результат использования шумоподавляющего и смазывающего фильтров.

ной ряд рассматривается как случайный стационарный процесс. Появляется необходимость развивать различные математические методы количественного анализа с целью извлечения из временных рядов достоверной информации.

В качестве исходных данных в рассматриваемом случае при построении временного ряда использовались площадь пятна и его периметр. В соответствии с выбранным центром инерции фигуры из исходного изображения были получены данные по яркости пятна в центре инерции, средняя яркость на краях и за пределами обнаруженного пятна.

К настоящему моменту разработано большое число методов прогнозирования временных рядов. Достаточно часто используется полиномиальная модель, но она условно описывает динамические процессы. Наиболее точных результатов в анализе временных

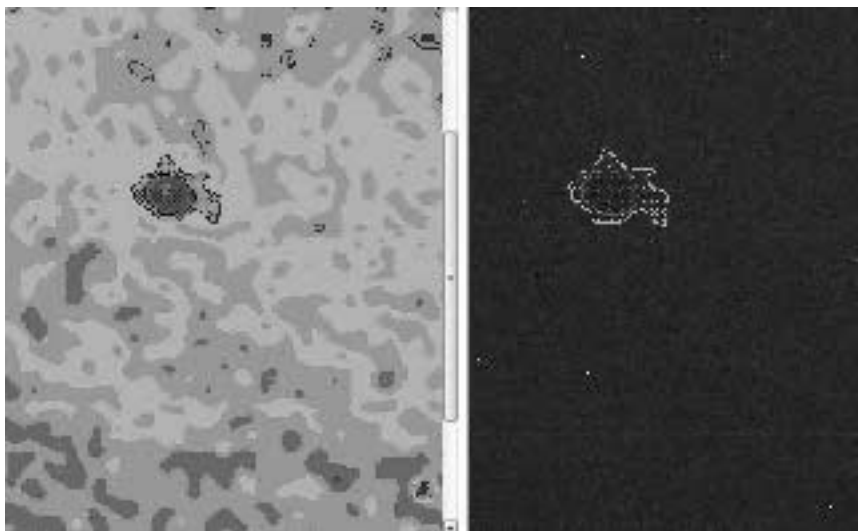


Рис. 2. Локалізація пляна на вихідному зображенні.

рядов можно добиться, рассматривая построенный ряд с точки зрения теории динамического хаоса.

Однако для построения краткосрочного прогноза временного ряда необходимо осуществить переход от хаотического состояния ряда, о котором свидетельствует его случайный характер, к регулярному состоянию. Для этого необходимо детально проанализировать все исходные параметры и влияние на них дополнительных случайных характеристик. Изменяя дополнительно введенные параметры и рассматривая их влияние, как влияние внешних условий, можно перестроить характеристики системы, свидетельствующие о состоянии динамической системы. Оценив полученный временной ряд и соответствующий ему аттрактор, мы сможем найти необходимый горизонт прогноза.

Изначально, при проведении анализа необходимо оценить класс и вид рассматриваемого процесса. Установить взаимосвязь между информативными признаками, полученными из изображения с реальными условиями, которые могут повлиять на прогнозируемую ситуацию.

Установление априори класса процесса во многом предопределяет алгоритм обработки результатов измерений и аппаратные средства.

Итак, классификационными признаками в данном случае могут быть стационарность или нестационарность, вид процесса — аддитивный, мультипликативный, аддитивно-мультипликативный и тип детерминированной составляющей. Поскольку речь идет о построении прогноза только по априорным данным, то рассматривается алгоритм классификации случайных процессов по одной реализации, основанный на использовании непараметрических критериев, показателя Херста, байесовской процедуре классификации и нечеткой логике.

Случайный характер воздействия введенных дополнительных факторов на характер обнаруженного явления предполагает применение процедуры статистической обработки результатов измерений, что обуславливает наличие таких составляющих погрешности, как статистическая погрешность и погрешность, вызванная неадекватностью алгоритма обработки реальному случайному процессу. Здесь необходимо учитывать, что, классифицируя нестационарный процесс как стационарный, можно увеличить методическую погрешность при оценке математического ожидания за счет увеличения интервала сглаживания.

Для проведения анализа одной реализации необходимо выполнить классификацию по типам нестационарности и при этом рассматривать поведение во времени оценок статистических характеристик.

Используя классическую теорию, случайные процессы можно представить как результат совместного действия детерминированного полезного сигнала и стационарной помехи.

В качестве полезного сигнала в рассматриваемом случае выступает обнаруженное изменение после проведения операции фильтрации, обнаружения и предварительного дешифрирования. В качестве помехи — подстилающая поверхность.

В общем случае влияние помехи на полезный сигнал может быть выражено оператором $X(t)=V(\varphi(t), \varepsilon(t))$, где $\varphi(t)$ — полезный сигнал (сигналы), $\varepsilon(t)$ — стационарная помеха. В зависимости от вида оператора V различают следующие модели сигналов

аддитивная модель $X(t)=\varphi_1(t)+\varepsilon(t)$;

мультипликативная модель $X(t) = \varphi_2(t)\varepsilon(t)$;

аддитивно-мультипликативная модель $X(t) = \varphi_1(t) + \varphi_2(t)\varepsilon(t)$, где $\varphi_1(t)$, $\varphi_2(t)$ — детерминированные функции времени, $\varepsilon(t)$ —

стационарный случайный процесс с нулевым математическим ожиданием, $m_\epsilon = 0$ и постоянной дисперсией $D\epsilon$.

Обычно рассматриваются стационарные процессы, когда оценивается во времени поведение математического ожидания, дисперсии и корреляционной функции. Поэтому и при классификации нестационарных процессов следует исходить из анализа этих же характеристик.

С учетом принятых допущений математическое ожидание mX , дисперсия DX и корреляционная функция RX случайных процессов, имеют следующий вид:

- аддитивная $mX(t) = \varphi_1(t)$; $D^{\wedge}t = D\epsilon$; $RX(t_1, t_2) = R\varphi(t_1, t_2)$;
- мультипликативная $mX(t) = 0$; $DX(t) = \varphi_2^2(t)D\epsilon$;
- аддитивно-мультипликативная $mX(t) = \varphi(t)$; $DX(t) = \varphi^2(t)D\epsilon$;
- $RX(t_1, t_2) = \varphi_2(t_1)\varphi_2(t_2)R\epsilon(t_1, t_2)$.

Из приведенных соотношений следует, что математическое ожидание для аддитивной и аддитивно-мультипликативной моделей зависит от детерминированной составляющей $\varphi_1(t)$. Дисперсия и корреляционная функция аддитивной модели полностью характеризуются свойствами стационарной помехи. А для мультипликативной и аддитивно-мультипликативной моделей эти вероятностные характеристики определяются также и детерминированной составляющей $\varphi_2(t)$.

Анализируя представленные соотношения можно сделать вывод, о том что для процессов, представленных аддитивной и аддитивно-мультипликативной моделями, математическое ожидание можно оценить по одной реализации с помощью той или иной операции, эквивалентной фильтрации низких частот.

Если дисперсия помехи $\epsilon(t)$ постоянная, то определить средний квадрат мультипликативного и аддитивно-мультипликативного процессов (и тем самым получить оценку дисперсии) также можно по одной реализации.

Таким образом, рассматривая анализируемые процессы, представленных любыми из этих моделей, нет необходимости проверять эргодические свойства нестационарного случайного процесса.

Точность оценки статистических характеристик будет зависеть от типа и параметров детерминированных процессов $\varphi_1(t)$ и $\varphi_2(t)$.

Классификацию по нестационарности — это необходимый предварительный этап исследования случайного процесса с целью

выявления его свойств до проведения основной статистической обработки.

В качестве классификационных признаков при наличии одной реализации исследуемого процесса рассматриваются класс процесса и вид нестационарности: нестационарность по математическому ожиданию, нестационарность по дисперсии, нестационарность по корреляционной функции, а также законы изменения математического ожидания и дисперсии.

Согласно классическому определению классификации, данная задача была решена еще на этапе построения «критериального дерева» на каждом из исходных изображений. В качестве одного из критериев рассматривалось разделение рассматриваемой совокупности объектов и явлений на однородные, в определенном смысле, группы, либо отнесение каждого из заданного множества объектов к одному из заранее известных классов. Это позволит произвести анализ разнородной информации: выделить объекты на изображении, произвести их привязку по географическим координатам и определить их качественные и количественные характеристики, а также идентифицировать объекты в соответствии с имеющейся базой знаний, основанной на нечёткой логике.

В классическом виде решение задачи классификации рассматривается как отображение вида:

$$R = \{r_1, r_2, \dots, r_n\} \rightarrow ye\{d_1, d_2, \dots, d_m\},$$

т. е. отнесение объекта, заданного вектором информативных признаков $R = \{r_1, r_2, \dots, r_n\}$, к одному из заранее определенных классов $\{d_1, d_2, \dots, d_m\}$.

Процессы, представленные рассмотренными выше моделями, относятся к классу нестационарных случайных процессов. Для выявления нестационарных свойств предлагается использовать непараметрические критерии, показатель Херста и коррелограммы, по результатам применения которых будет формироваться вектор информативных признаков R .

Значительное большинство непараметрических критериев реагируют на изменение оценки математического ожидания. Таким образом, непараметрические критерии без предварительной обработки наблюдаемого ряда позволяют выделить два класса процес-

сов «стационарные по математическому ожиданию» и «нестационарные по математическому ожиданию».

Для определения динамического состояния системы используются показатели Херста. Показатель Херста характеризует направление развития системы; если его значение превышает число 0.5, то временной ряд в будущем промежутке времени стремится к возрастанию; если показатель Херста равен 0.5, то временной ряд находится в неопределенном состоянии, это говорит о том, что мы не можем сказать о направлении его развития; если показатель меньше 0.5, то временной ряд стремится к убыванию. Если приращения были положительными в течение некоторого времени в прошлом, то есть происходило увеличение, то и впредь в среднем будет происходить увеличение. Таким образом, для процесса с $H > 0.5$ тенденция к увеличению в прошлом означает тенденцию к увеличению в будущем. И наоборот, тенденция к уменьшению в прошлом означает, в среднем, продолжение уменьшения в будущем. Чем больше H , тем сильнее тенденция.

По значению показателя Херста можно судить как о стационарности процесса по математическому ожиданию, так и о виде детерминированной составляющей.

Однако, наибольший интерес представляет случай, когда показатель Херста принимает значение равное индифферентности, т. е. равен 0.5. В данном случае появляется возможность, используя алгебру нечеткой логики учесть дополнительные возможности, влияющие на точность прогноза.

В основе классификации по непараметрическим критериям лежит байесовская процедура для бинарных признаков. Полученные таким образом оценки рассматриваются как обобщенный результат применения непараметрических критериев, а апостериорная вероятность — как классификационный признак. При этом шкала измерений становится такой же, что и для показателя Херста.

С помощью метода нормированного размаха Херста есть возможность рассчитать показатели Херста не только в областях выявленных изменений или пятен, но и в прилегающих к ним зонам. Это даст возможность, установив зависимость между полученными показателями Херста в различных зонах, получить некоторую оценку интервалов предсказуемости поведения данного ряда. Границы будут характеризоваться точками излома статистики

Хёрста, на которых оценка показателя Хёрста будет существенно изменяться.

Однако, все полученные результаты требуют дополнительного анализа. Точность анализа и дальнейшего прогноза во многом зависит от количества априорных данных, а это в свою очередь требует создания значительной базы данных, содержащих снимки, полученные различными системами и в различных диапазонах.

* * *

Алексеев А.В. Применение нечеткой математики в задачах принятия решений. — В сб.: Методы и системы принятия решений. — Рига: РПИ, 1983. — С. 38—42.

Аоки М. Введение в методы оптимизации. — М.: Наука, 1977. — 344с.

Мусеев Н.Н. Элементы теории оптимальных систем. М: Наука, 1975, 528с.

Fractal geometry of information space as represented by cocitation clustering / Van Raan A. F. J. // Scientometrics. — 1991. — Vol. 20, 3. — P. 439—449.

Федер Е. Фракталы / М.: Мир, 1991. — 254 с.

Дистанционное зондирование: количественный подход / Ш.М. Дейвис, Д.Я. Ландгребе, Т.Л. Филлипс и др. — М.: Недра, 1983. — 415 с.

Красовський Г.Я., Петросов В.А. Інформаційні технології космічного моніторингу водних екосистем і прогнозу водоспоживання міст. — К.: Наук. думка, 2003.

Отримано: 09.09.2008 р.