УДК 519.281.1

В. Л. ГОРОХОВ, В. Н. ПРОКОФЬЕВ

(Ленинград)

УСТОЙЧИВЫЕ АЛГОРИТМЫ ОБНАРУЖЕНИЯ СИГНАЛОВ ДЛЯ СИСТЕМ ПЕРВИЧНОЙ ОБРАБОТКИ ДАННЫХ БОЛЬШИХ ТЕЛЕСКОПОВ

Введение. Исходные предпосылки. При создании крупных радио- и оптических телескопов (таких, как РАТАН-600 или БТА — большой телескоп азимутальный) значительное внимание уделяется задачам автоматизации приема и обработки астрономической информации. Сюда относятся задачи синтеза автоматических решающих модулей для оперативного просмотра данных с целью предварительного выделения «сигнальных зон», для определения графа обработки, обнаружения сбоев, калибровки и т. д. [1, 2]. Соответствующие решающие алгоритмы должны обладать устойчивой по отношению к условиям наблюдения структурой, т. е. не должны, например, зависеть от средних характеристик помех, которые часто неизвестны и поэтому не могут входить в структуру решающих алгоритмов (в их тестовые статистики, пороги).

Наличие случайных помех требует привлечения для синтеза таких устойчивых алгоритмов современных статистических методов [3—6]. Ниже рассматриваются некоторые устойчивые алгоритмы обнаружения, предложенные авторами для решения задачи оперативного контроля процесса обработки в автоматизированной системе первичного анализа астрономической информации.

Конкретно задача формулируется и решается применительно к РАТАН-600, хотя результаты могут использоваться и в других случаях.

Пусть при сканировании участка неба антенной радиотелескопа формируются дискретные массивы наблюдения — сканы. Нужны оперативные решения о наличии сигнальных зон — радиоисточников. Сигнальные зоны могут быть точечными или протяженными; в них мощность принимаемого антенной излучения возрастает по отношению к фону, уровень которого не известен и может меняться по скану. Ниже рассматриваются два варианта задачи — в параметрической и непараметрической формах. Основные предпосылки таковы.

1. Все поле просмотра анализируется последовательным сканированием (решения принимаются в отдельных или парных сканах). В каждом скане анализ проводится последовательными тройками интервалов разрешения, из которых левый считается шумовым, а два других — анализируемыми (в них проверяется наличие сигнала). Этим преодолеваются трудности, вызываемые неопределенным и меняющимся по скану уровнем фона (либо его функции распределения) и возможным появлением сигнала на границе интервалов.

2. Если в анализируемой тройке сигнал обнаружен, ее шумовой интервал используется при формировании и анализе очередной тройки; такая процедура продолжается до тех пор, пока не будет получен ответ об отсутствии сигнала; тогда происходит нужное смещение шумового интервала, и процедура возобновляется. В результате достигается обнаружение как протяженных, так и точечных сигнальных зон. Дополнительные уточнения процедуры анализа обеспечиваются другими программными модулями системы.

Далее рассматривается получение решающих алгоритмов для анализа отдельной тройки интервалов.

Инвариантные алгоритмы. Для исходных наблюдений целесообразна модель χ^2_m-распределения с m степенями свободы. Уровень шума считается одинаковым для интервалов тройки, но неизвестным; сигнал может присутствовать во втором или третьем либо во втором и третьем интервалах; его уровень различен и не известен.

При n независимых наблюдениях в каждом интервале исходной выборкой служит совокупность $(x, y) = (x_1, \ldots, x_n; y_1, \ldots, y_{2n})$, где переменные x_i отвечают шумовому интервалу, а y_i — анализируемым. В соответствии с χ^2-моделью совместное распределение данных

$$g(x, y) = [\sigma^{m/2}\Gamma(m/2)]^{-3n} \exp\left[-\frac{1}{\sigma}\left(\sum_{i=1}^{n} x_i + \sum_{j=1}^{2n} \frac{y_j}{1+q_j}\right)\right] \times$$
$$\times \prod_{j=1}^{n} x_i^{m/2-1} \prod_{j=1}^{2n} (1+q)^{-m/2} y_j^{m/2-1}, \qquad (1)$$

где σ — мощность шума; $q_j = \nu_j/\sigma$ — отношение сигнал/шум (С/Ш) в j-м отсчете; ν_j — мощность сигнала в нем, $j = \overline{1, 2n}$.

Задача обнаружения заключается в проверке гипотез

$$H_0 : q = (q_1, \ldots, q_{2n}) = 0 \text{ (нулевой вектор) — сигнала нет;}$$
$$H_1 : q \neq 0 \text{ — сигнал есть, } \sigma \text{ — неизвестно} \qquad (2)$$

на базе семейства (1). Это семейство содержит мешающий (неизвестный) масштабный параметр σ, и решение задачи (2) следует искать среди правил, инвариантных к масштабу наблюдений [5, 6].

Используя методику [6] (как, например, в [7]), можно найти, что наиболее мощное инвариантное правило обнаружения имеет критическую область вида

$$\left(\sum_{i=1}^{n} x_i + \sum_{j=1}^{2n} y_j\right) \Bigg/ \left(\sum_{i=1}^{n} x_i + \sum_{j=1}^{2n} \frac{y_j}{1+q_j}\right) > C, \qquad (3)$$

C — пороговая константа. Будучи инвариантным к масштабу наблюдений, алгоритм (3) не зависит от неизвестной мощности шума (порог C также не зависит от σ); однако правило (3) содержит в своей структуре параметры q_j, $j = \overline{1, 2n}$, и этим неудобно для применения. Получение более практичного инвариантного решения основано на использовании методики [8] нахождения максиминного инвариантного решения (максимизирующего минимум мощности решения на некотором классе альтернатив H_1).

Пусть q_* — некоторое заданное граничное значение С/Ш (указываемое по физическим соображениям); тогда максиминная критическая область имеет вид

$$\sum_{j=1}^{2n} (1 - a_* z_j)^{-3mn/2} > C,\ z_j = y_j \Bigg/ \left(\sum_{i=1}^{n} x_i + \sum_{k=1}^{2n} y_k\right),\ a_* = q_*/(1+q_*). \qquad (4)$$

В типичном случае $q_* \ll 1$ (слабые сигналы) из (4) получается следующая инвариантная локально-максиминная (ЛММ) область:

$$v = \frac{1}{2} \sum_{j=1}^{2n} y_j \Bigg/ \sum_{i=1}^{n} x_i > C, \qquad (5)$$

которая и предлагается для практического использования (множитель 1/2 связан с нормировкой). Порог C в (5) может быть найден (независимо от уровня шума) с помощью таблиц центрального F-распределения [9] по заданной вероятности α ложных тревог из обычного условия

$$\alpha = \text{Вер}\{v > C/H_0\} = \int_{C}^{\infty} F_{2mn, mn}(v)\, dv, \qquad (6)$$

где $F_{2mn, mn}$ — центральное F-распределение с $2mn$, mn степенями свободы статистики в (5) при H_0. Мощность β правил (5), (6) (вероятность правильного обнаружения) легко рассчитать также с помощью таблиц для случая $q_j = q$, $j = \overline{1, 2n}$; при этом

$$\beta = \int\limits_{C_1}^{\infty} F_{2mn,\,mn}(v)\,dv, \quad C_1 = C/(1+q). \tag{7}$$

При проведении обзоров часто целесообразно учитывать тот факт, что если сигнал присутствует в интервалах одного скана, то он есть и в соответствующих интервалах смежного скана. При этом желательно вести анализ парами сканов. В этом случае сохраняются предыдущие предпосылки; дополнительно считается, что неизвестный уровень шума различен в двух смежных одновременно анализируемых сканах, что отвечает практике. Аналогично предыдущему ЛММ инвариантная критическая область получается здесь в виде

$$\sum_{k=1}^{2} \frac{v_k}{1+v_k} > C,\ v_k = \sum_{j=1}^{2n} y_{kj} \Big/ \sum_{i=1}^{n} x_{ki},\ k = 1, 2. \tag{8}$$

При $mn \gg 1$ можно использовать нормальную аппроксимацию распределения статистики в (8) (величины $v_k/(1+v_k)$, $k = 1, 2$ — первый и второй сканы — при H_0 имеют бета-распределение с mn, $mn/2$ степенями свободы). Тогда порог в (8) $C = 4/3 + (2/3)\sqrt{2/(3mn)}\Phi^{-1}(1-\alpha)$, где $\Phi^{-1}(\cdot)$ — функция, обратная функции $\Phi(\cdot)$ распределения стандартного нормального закона. Оценкой мощности правила (8) служит формула

$$\beta \sim 1 - \Phi[(1/(1+q))\Phi^{-1}(1-\alpha) - (q/(1+q))\sqrt{2mn(mn-4)/(3mn-2)}].$$

Свободный от распределения (непараметрический) алгоритм. Помехи с неизвестным распределением приводят к непараметрической формулировке задачи обнаружения, решение которой должно быть поэтому свободным от распределения (СР), т. е. не должны использоваться сведения о форме распределения данных. Пусть распределение $F(\cdot)$ с плотностью $f(\cdot)$ одинаково для всех x_i, $i = \overline{1, n}$, и неизвестно; распределения G_j с плотностями $g_j(\cdot)$, $j = \overline{1, 2n}$, смеси сигнала и шума могут быть различны и также неизвестны. В отсутствие сигнала $G_j = F$, $j = \overline{1, 2n}$, а при его наличии каждое из распределений G_j стохастически возрастает по отношению к F ($G_j \succ\succ F$), $j = \overline{1, 2n}$ [4], т. е. задача обнаружения сводится к проверке непараметрических гипотез

$$H_0 : G_j = F;\ H_1 : G_j \succ\succ F,\ j = \overline{1, 2n}. \tag{9}$$

Предлагается получить решение задачи в два этапа: на первом этапе принимаются отдельные i-е решения по каждой тройке наблюдений (x_i, y_i, y_{i+n}), $i = \overline{1, n}$; на втором — объединяются (обычным суммированием) все n таких решений, $i = \overline{1, n}$. СР-решение первого этапа на базе величин (x, y_1, y_2) (индекс «i» опущен, y_1, y_2 — анализируемые отсчеты) удобно найти, используя вместо y_1, y_2 экстремальное наблюдение, например $z = \max(y_1, y_2)$.

По методике [10] можно получить критическую область (нерандомизированную) вида $z > x$. Этот алгоритм свободен от распределения: при всех F он имеет вероятность ложного решения $p_0 = 2/3$; его мощность для рэлеевских распределений (случай $m = 2$, $n = 1$ в (1)) равна

$$\begin{aligned} p = {} & (1+q_1)/(2+q_1) + (1+q_2)/(2+q_2) - (1+q_1)\times \\ & \times(1+q_2)/(2+q_1+q_2+(1+q_1)(1+q_2)), \end{aligned} \tag{10}$$

где q_i — отношения С/Ш в анализируемых интервалах, $i = 1, 2$.

Если t_i, $i = \overline{1, n}$, — отмеченные отдельные решения первого этапа, то алгоритм накопления второго этапа есть алгоритма (11) надлежит использовать (вместе с формулой типа (10)) таблицы биномиального распределения [9]; при больших n возможна нормальная аппроксимация статистики в (11), и тогда $K \sim np_0 + \sqrt{np_0(1-p_0)}\Phi^{-1}(1-\alpha)$, а мощность $\beta \sim 1 - \Phi[\sqrt{p_0(1-p_0)/p(1-p)} \times \Phi^{-1}(1-\alpha) - \sqrt{n}(p-p_0)/\sqrt{p(1-p)}]$.

Результирующий алгоритм (11) на базе СР-правила предварительного квантования является также свободным от распределения, т. е. обладает высокими свойствами устойчивости по отношению к условиям наблюдения и пригоден для реализации в автоматизированных устройствах.

Устойчивость данного алгоритма проверялась методом статистического моделирования на ЦВМ. Находилась оценка вероятности ложных тревог при четырех типах распределения фона (показательное, рэлеевское, нормальное, равномерное) согласно методике из [11]. Результаты показали, что СР-алгоритм действительно выдерживает заданный уровень ложных тревог (были приняты заданные значения $\alpha = 0{,}2$; $0{,}5$; погрешность статистических оценок не более 5%).

ВЫВОДЫ

1. Предложены инвариантные (к масштабу наблюдений) алгоритмы обнаружения сигналов неизвестной протяженности при «просмотре» одиночными и парными сканами для модели χ^2-распределения радиоастрономических наблюдений. Тестовые статистики и пороги алгоритмов не зависят от неизвестных уровней шума и сигнала; при этом вероятность ложных тревог выдерживается неизменной при любой мощности фона, а вероятность правильного обнаружения подчиняется максиминному свойству. Расчеты эффективности алгоритмов можно выполнить с помощью таблиц нормального и центрального F-распределений.

2. В отсутствие сведений о форме распределения данных предложен свободный от распределения алгоритм, выдерживающий назначенный уровень ложных тревог для широкого класса распределений фона; устойчивость алгоритма подтверждена моделированием на ЦВМ.

3. Отмеченная устойчивость, а также простота решающих алгоритмов позволяют рекомендовать их при создании автоматизированных систем обработки астрономических наблюдений; такое применение алгоритмов начато в системах РАТАН-600.

ЛИТЕРАТУРА

1. **Westerhout G.** Some Remarks on the Ideal Automated Observatory.— Astron. Astrophys. Suppl., 1974, vol. 15, p. 327.
2. **Haslam C. G. T.** NOD2 General System of Analysis for Radioastronomy.— Ibid, 1974, vol. 15, p. 333.
3. **Левин Б. Р.** Теоретические основы статистической радиотехники. М.: Сов. радио, 1976, кн. 3.
4. **Репин В. Г., Тартаковский Г. П.** Статистический синтез при априорной неопределенности и адаптации информационных систем. М.: Сов. радио, 1977.
5. **Леман Э.** Проверка статистических гипотез. М.: Наука, 1964.
6. **Гаек Я., Шидак З.** Теория ранговых критериев. М.: Наука, 1971.
7. **Прокофьев В. Н.** Инвариантное обнаружение флуктуирующего сигнала по наблюдениям из гамма-распределения с неизвестными параметрами.— Изв. высш. учебн. заведений СССР. Сер. Радиоэлектроника, 1977, т. 20, № 3, с. 3.
8. **Прокофьев В. Н.** Максиминное решение задачи обнаружения с векторным информативным параметром.— Изв. АН СССР. Сер. Техн. кибернетика, 1978, № 5, с. 145.
9. **Большев Л. Н., Смирнов Н. В.** Таблицы математической статистики. М.: Наука, 1965.

10. **Прокофьев В. Н.** Двухвыборочные ранговые правила обнаружения сигнала с использованием крайних наблюдений.— Радиотехника и электроника, 1976, т. 21, № 8, с. 1782.
11. **Сосулин Ю. Г.** Теория обнаружения и оценивания стохастических сигналов. М.: Сов. радио, 1978.

Поступила в редакцию 25 января 1981 г.

УДК 535. 317.25 : 535.8

А. А. ВАСИЛЬЕВ, Б. В. КЛИМКОВИЧ,
И. Н. КОМПАНЕЦ, С. П. КОТОВА
(Москва)

ЛОГАРИФМИЧЕСКОЕ ПРЕОБРАЗОВАНИЕ КООРДИНАТ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ ЖИДКОКРИСТАЛЛИЧЕСКОГО МОДУЛЯТОРА

Введение. Преобразования координат или геометрические преобразования сигналов [1, 2] относятся к классу линейных пространственно-зависимых преобразований. Интерес к геометрическим преобразованиям обусловлен возможностью решения с их помощью таких практически важных задач, как коррекция изображений, подвергшихся пространственно-зависимым искажениям, и корреляционное опознавание оптических сигналов с разным масштабом и различной ориентацией. Преобразования координат могут быть выполнены электронными (с помощью специальных телевизионных разверток) и оптическими методами [1, 2]. Последние из них являются наиболее привлекательными, поскольку не требуют многократных преобразований оптических сигналов в электрические и обратно.

Известно несколько оптических методов выполнения геометрических преобразований. Один из них, предложенный в [3], предназначен для обработки одномерных сигналов. В нем световой пучок, пространственно промодулированный вдоль одной координаты в соответствии с амплитудой сигнала, проходит сквозь щелевую диафрагму, форма которой зависит от выполняемого геометрического преобразования сигнала. Недостатком такого метода является несовместимость большой светосилы оптической системы с высокой пропускной способностью. С ростом числа N разрешаемых системой элементов сигнала светосила падает как $1/N^2$. Отмеченный недостаток присущ и методам, развитым в работах [4, 5].

В другом способе используется голографический фильтр, полученный путем записи наложенных голограмм, каждая из которых представляет результат интерференции световых волн от двух точечных источников, являющихся дискретными отсчетами преобразуемого и результирующего сигналов [6, 7]. Каждая точка входной плоскости восстанавливает только одну соответствующую ей точку в результирующем сигнале. При использовании объемных регистрирующих сред это достигается за счет избирательности фильтра по углу падения волны, а для тонких голограмм — за счет модуляции опорной волны разными участками диффузно рассеивающей пластины. Пропускная способность таких схем недостаточно высока, поскольку с ростом числа наложенных голограмм растут шумы в преобразованном сигнале и падает дифракционная эффективность. Метод экспериментально продемонстрирован для сигналов, содержащих 3—4 отсчета [6, 7].

Наиболее эффективным, по-видимому, является метод фазовых масок, предложенный в [8] (рис. 1). Во входной плоскости системы (передняя фокальная плоскость линзы) вплотную с преобразуемым сигналом