

УДК 621.396.96

В.П. Грозов, В.Е. Носов, Г.А. Ососков

ВОПРОСЫ ОБРАБОТКИ ИЗОБРАЖЕНИЙ ПРИМЕНИТЕЛЬНО К ЗАДАЧАМ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ ИОНОГРАММ

Диагностика ионосферного канала основывается на анализе ионограмм вертикального, наклонного и возвратно-наклонного зондирования. Для проведения анализа необходимо решение двух основных задач: 1) обработка ионограмм, т.е. коррекция и улучшение изображения и выделение точек (моментов прихода) сигнала; 2) связывание выделенных точек в треки для их последующей привязки к конкретным модам распространения. Для решения первой задачи использованы статистические методы обработки изображений. Для решения второй – метод искусственных нейронных сетей (ИНС) Хопфилда. Вследствие сложного характера треков на фоне значительных помех использовалась модифицированная роторная модель. Удачный выбор начальной конфигурации обеспечил быструю сходимость сети. Тестирование модели было проведено на ионограммах, полученных на ЛЧМ-зонде (ИСЗФ, Иркутск, 1987–1996 гг.). Анализ результатов показал, что данный подход дает хорошие результаты и является перспективным в решении задач обработки ионограмм.

Введение

Основной тенденцией в ионосферной информатике является сокращение сроков доступа к текущей диагностической информации о состоянии ионосферы. Оперативность поступления данных приобретает принципиальное значение, так как, во-первых, это существенно облегчает организацию геофизических и радиофизических экспериментов по ис-

следованию ионосферы, а во-вторых, это необходимо для ряда практических задач, в особенности в рамках современной концепции распределения рабочих частот для коротковолновой (КВ) связи и радиолокации.

Оперативность обработки больших потоков данных возможна только в случае автоматической обработки ионограмм при различных режимах зондирования ионосферы.

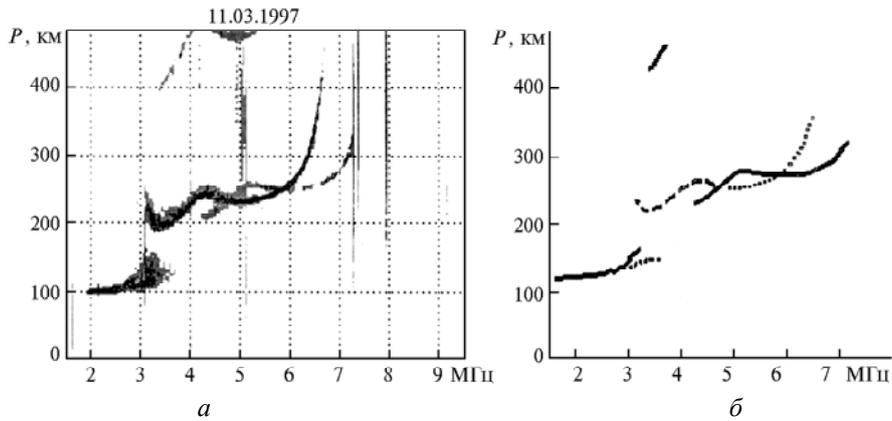


Рис. 1. Ионограмма ВЗ

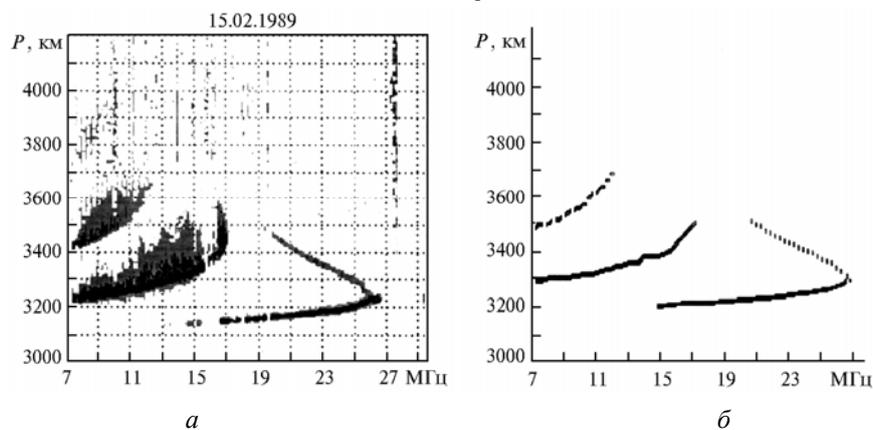


Рис. 2. Ионограмма НЗ

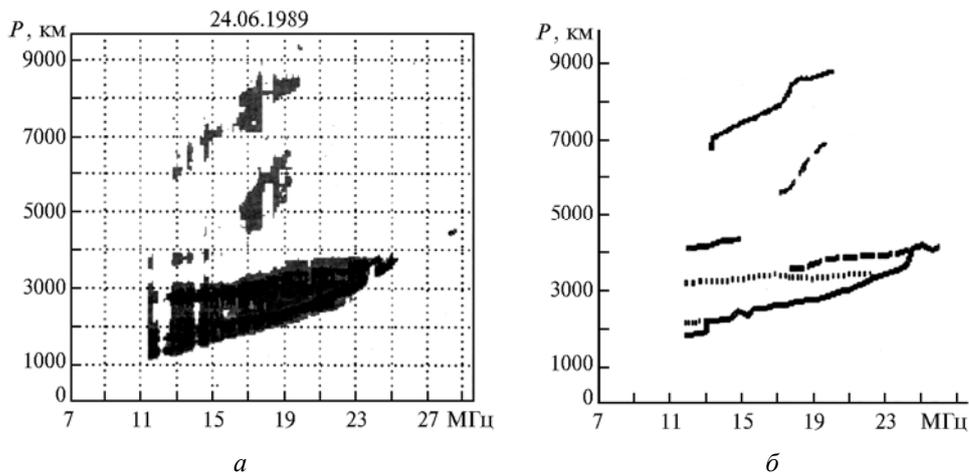


Рис. 3. Ионограмма ВНЗ

В общем виде ионограмму можно рассматривать как растровое изображение дистанционно-частотной характеристики, регистрируемой ионозондом [1]. Каждый элемент изображения (ионограммы) определяется двумя параметрами: координатами (частота и задержка) и яркостью (амплитудой) (рис. 1–3, а). Таким образом, рассматриваемое изображение можно представить в виде матрицы по координатам $A[F_i, D_j]$ размера $N \times M$. В результате обработки необходимо выделить связные линии, образованные характерными точками (определяемыми по какому-либо правилу), т.е. построить трек, который в дальнейшем надо привязать к конкретному моду распространения.

В рамках такой постановки проблемы необходимо провести:

- а) обработку ионограммы, включающую: коррекцию изображения, т.е. фильтрацию импульсных помех, коррекцию амплитудных характеристик, подавления шума и т.д.; улучшение изображения, т.е. повышение контраста, выделение областей, характеризующихся свойством связности, и т.д.; определение характерных точек сигнала, соответствующих физически значимым характеристикам изображения;
- б) построение треков и привязку их к конкретному моду распространения.

1. Обработка ионограммы

Задачи обработки ионограмм тесно связаны между собой по методам их решения. В своей основе эти методы опираются на наличие характерного шума, свойства которого связаны с воздействием среды, параметрами приемопередающей аппаратуры и т.д. Поскольку шум присутствует практически всегда, то изображение можно интерпретировать как случайные процессы двух переменных, т.е. как случайные поля. А следовательно, для обработки изображений можно использовать статистические методы обработки информации [2–6]. Эффективность методов определяется статистическими свойствами шума.

Модель шума охватывает статистическое описание совокупности элементов, образующих изображение, а также вид взаимодействия шума и сигнала.

В качестве модели изображения используется стандартная рэлеевская модель распределения, а в качестве аналитического описания вида взаимодействия между шумом и сигналом используется мультипликативная модель. Выбранная модель позволяет использовать для коррекции изображения локальные методы сглаживания, обладающие большой вычислительной эффективностью и адаптивные к характеру сцены. Данные методы позволяют сформировать оценки поля незашумленного изображения путем анализа и обработки ограниченных и достаточно малых по размерам фрагментов. В качестве фрагмента для ионограмм используется квадрат размером 5×5 точек матрицы $\{A\}$. Выбор данных размеров определяется следующими условиями:

- а) полосой, где сигнал практически не изменяется;
- б) разрешением зонда;
- в) требованием обеспечения достаточности объема выборочных данных и минимума вычислительных затрат.

Обработка изображения ведется в режиме скользящего фрагмента с перекрытием.

На первом этапе проводится корректировка данных с целью исключения вертикальных и точечных помех, а также заполнение случайных пропусков данных. Корректировка проводится с использованием метода линейного предсказания, когда корректированное значение находится как сумма взятых с некоторыми весами значений соседних элементов. В работе была выбрана следующая формула [5]:

$$B_{i,j} = 0,3(A_{i,j-1} + A_{i-1,j}) + 0,2(A_{i-1,j-1} + A_{i+1,j-1}).$$

Удаление импульсных помех проводится с использованием медианного фильтра.

При проведении сглаживания шума наибольший интерес представляют операции, снижающие уровень шума без размывания перепадов яркости. Наи-

более эффективными являются методы, основанные на локальных порядковых статистиках. Так, для удаления шума реализована методика, основанная на исключении вкладов в усредненное значение отсчетов, не удовлетворяющих некоторой модели однородности. Идея методики заключается в том, что в пределах окна большего фрагмента (например, 5×5) выбираются фрагменты меньшего размера (3×3). В каждом вычисляются среднее значение и дисперсия. Среднее значение яркости окрестности с минимальной дисперсией присваивается центральному элементу фрагмента [7]. Для увеличения быстродействия вместо дисперсии в качестве показателя однородности используется выражение

$$S(k) = \sum_{n=1}^9 |(A_{ij} - A_n(k))|, \quad k = \overline{1, 9}.$$

Здесь A_{ij} – яркость центральной точки большого фрагмента; $A_n(k)$ – яркость n -й точки k -го малого фрагмента. При этом A_{ij} заменяется на средний уровень яркости в пределах фрагмента с минимальным значением $S(k)$. Необходимо отметить, что данный подход не требует никаких априорных знаний и полностью определяется характером сцены.

На следующем этапе проводится контрастирование деталей изображения с использованием нелинейного адаптивного амплитудного преобразования, которое строится на основе измерения и анализа гистограммы распределения значений сигнала. Функцию преобразования сигнала подбирают таким образом, чтобы преобразованный сигнал имел требуемую гистограмму распределения. В рассматриваемых задачах хорошие результаты были получены, когда в качестве выходного распределения использовалась равномерная плотность вероятности. Функция преобразования при этом имеет вид:

$$D = (\overline{D}_{\max} - D_{\min})P_A(A) + D_{\min},$$

где $P_A(A)$ – функция распределения вероятностей исходного изображения, а D_{\max} и D_{\min} – максимальный и минимальный уровни преобразованного сигнала. Результатом такого преобразования является увеличение контрастов на участках изображения с наиболее часто встречающимися значениями сигнала.

Результаты выполнения контрастирования изображения позволяют приступить к сегментации с использованием методов пороговой обработки. Поскольку можно считать, что значения порогов для рассматриваемых задач постоянны в некоторой окрестности точки изображения и зависят от локальных признаков изображения, а также от ее координат, то можно использовать методы обработки с переменным порогом. Идея обработки заключается в том, что всегда имеется неравномерность яркости в пределах изображения. А поэтому выбором соответ-

ствующего порога можно разделить области со значимой информацией и область фона.

Поскольку неравномерность яркости сложно описать какой-либо известной функцией, то реализована оценка порога в пределах ограниченного фрагмента. Если фрагмент содержит как объект, так и фон, то по идее его гистограмма яркостей должна быть (по крайней мере) бимодальной. Минимальное значение гистограммы на участке между областями дает локальный порог для выделения объекта на фоне в данном фрагменте изображения. Когда фрагмент содержит только объект или только фон, то гистограмма унимодальная и локальный порог для фрагмента не может быть определен. В этом случае он назначается путем интерполяции локальных порогов, найденных для ближайших бимодальных фрагментов. Результатом этапа является выделение на ионограмме квазиоднородных областей, соответствующих модам сигнала.

Таким образом, в результате использования описанных выше алгоритмов удается в значительной мере удалить помехи, сгладить шум и сжать данные, оставив точки максимальной амплитуды. На последнем этапе в выделенных областях определяются моменты прихода сигнала, в качестве которых выбираются локальные максимумы для каждой из выделенных областей. В результате получаем изображение в виде облаков (матрицы) точек со значимой амплитудой в координатах: групповой путь – рабочая частота (рис. 1–3, в). Полученная матрица точек используется для построения треков и их интерпретации.

2. Построение треков

Для построения треков был использован метод искусственных нейронных сетей (ИНС), успешно работающий в физике высоких энергий [8, 12]. Эффективность метода основывается на способности нейронной сети эволюционировать к некоторым равновесным состояниям, которым соответствует минимум энергетической функции. При надлежащем выборе весовой функции минимум энергетической функции должен соответствовать оптимальному распределению точек по трекам. В силу схожести проблемы метод нейронных сетей был использован для распознавания треков на ионограммах вертикального зондирования [9], а затем и для других типов ионограмм [10].

Для решения задачи идентификации треков на ионограммах была выбрана ИНС Хопфилда (ХНС) [11]. Преимущество ХНС заключается в способности находить оптимальное решение некоторой задачи без предварительного обучения путем минимизации соответствующего функционала, заранее заданного и содержащего информацию о таком решении. ХНС – это полностью связанная сеть, т.е. имеющая пути передачи сигнала от выхода нейрона к их входам. Отклик таких сетей на информацию извне является динамическим, так как вычисленный выход, передава-

ясь по обратным связям, модифицирует вход; затем выход повторно вычисляется, и весь процесс повторяется снова. Для устойчивой сети последовательные операции приводят к достижению сетью стабильного состояния.

Проведенный анализ ионограмм (см. рис. 1–3, *a*) показал, что линии, аппроксимирующие треки даже одного типа, сильно отличаются по форме и протяженности. Так, треки на ионограммах нельзя целиком описать прямой или участком окружности, что затрудняет их аналитическое описание. Поэтому речь может идти лишь о локальной аппроксимации. Для этой цели наиболее подходит дуга окружности; в этом случае трек будет образован дугами разной длины и кривизны. Столь же локальным должно быть и взаимодействие нейронов. Поскольку ионограммы характеризуются наличием сложных треков на фоне значительных помех, то в этих условиях была выбрана модифицированная роторная модель Хопфилда.

В рамках ХНС предполагается, что нейроны представлены в виде роторов (рис. 4). Динамической переменной является угол. Эти роторы взаимодействуют друг с другом и с вектором \mathbf{L}_{ij} , соединяющим их, т.е. нейроны характеризуются значением, координатой и наклоном. При этом предполагается, что сигналы достаточно хорошо ложатся на некоторую окружность. А нейронная сеть должна расположить векторы по касательной к этой окружности и сильно уменьшить векторы, которые не лежат на каком-либо истинном треке.

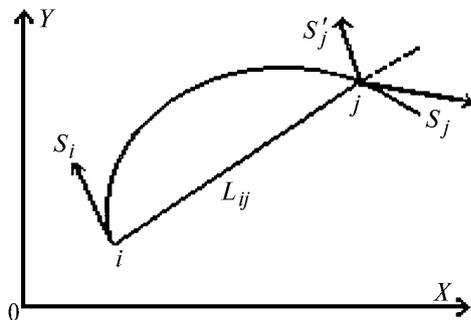


Рис. 4. Модифицированная роторная модель

Таким образом, каждому нейрону ставится в соответствие вектор \mathbf{s}_i , модуль которого характеризует интенсивность воздействия данного нейрона на остальные, а направление в идеале должно совпадать с касательной к треку. Поскольку трек имеет непостоянную кривизну и изломы, то взаимодействовать могут лишь нейроны некоторого локального участка. Это достигается введением робастного множителя в весовую функцию T_{ij} , играющую роль жесткого фильтра. Переходя к полевой терминологии, введем векторную величину \mathbf{h}_{ij} как поле, созданное в точке i нейроном j :

$$\mathbf{h}_{ij} = T_{ij} \mathbf{s}_i.$$

Тогда общее поле \mathbf{H}_i в произвольной точке определится как векторная сумма полей от всех нейронов:

$$\mathbf{H}_i = \sum_j \mathbf{h}_{ij} = \sum_j (T_{ij} \mathbf{s}_i).$$

С учетом замечаний о локальности предполагаем, что трек образуют нейроны, собственный вектор которых близок к полемому вектору в данной точке по направлению. За меру близости удобно взять скалярное произведение векторов \mathbf{s}_i и \mathbf{H}_i . Тогда энергетическая функция системы запишется в виде

$$E = -\frac{1}{2} \sum_i \mathbf{H}_i \mathbf{s}_i = -\frac{1}{2} \sum_{ij} \mathbf{s}_i T_{ij} \mathbf{s}_j.$$

Ее минимум соответствует оптимальному распределению точек по трекам при надлежащем выборе весовой функции.

Задача идентификации треков сводится к организации такой эволюции состояний нейронной сети, которая обеспечивает ее сходимость к конфигурации, минимизирующей энергетическую функцию системы. Итеративная процедура поиска такой конфигурации ХНС строится как последовательное вычисление поля в точках, соответствующих нейронам, определение каждым нейроном своего состояния и нового значения вектора для следующего шага итерации. Новое значение модуля вектора i -го нейрона определяется по формуле [12]:

$$|\mathbf{s}_i| = \tanh(|\mathbf{H}_i|/T), \text{ где } T = 1,5.$$

Новое направление вектора принимается совпадающим с направлением поля в точке. Вычисленный таким образом вектор участвует в определении векторов других нейронов уже на данном шаге эволюции ХНС. На следующем шаге вся процедура повторяется. После ряда итераций (достаточность их числа определяется по малости изменений энергетической функции с новой итерацией) векторы нейронов с некоторой погрешностью выстраиваются по направлению касательных к треку в точках локализации нейронов.

Окончательное восстановление треков проводится с помощью алгоритма, считывающего информацию с ИНС. Две точки считаются связанными, если их взаимодействие вносит минимальный вклад в энергетическую функцию. Основными проблемами на этом этапе являются: 1) формулировка условий обрыва одного трека и начало нового, если треки расположены близко; 2) корректность соединения треков при их пересечении. От этого зависит адекватность интерпретации ионограммы. Основной процент ошибок делается на этом этапе.

После построения треков необходима привязка их к конкретным модам распространения. Эта достаточно сложная задача, поскольку методика привязки существенно зависит от режима зондирования и

требует отдельного рассмотрения. Один из вариантов методики идентификации треков в случае наклонного зондирования рассмотрен в [13].

Заключение

Рассмотренный подход был реализован в виде пакета программ и применялся для обработки ионограмм вертикального, наклонного и возвратно-наклонного зондирования ионосферы, полученных в ИСЗФ зондом с непрерывной линейной частотной модуляцией [1]. В качестве примера на рис. 1–3, с представлены результаты для ионограмм каждого вида. На обработку одной ионограммы требуется 15–20 с на РС IBM типа 486DX2 в зависимости от типа ионограмм.

1. Brynko I.G. et al. // Adv. Space Res. 1980. V. 8. N 4. P(4)121–(4)124.
2. Pratt W.K. Digital image processing. N.Y. 1978. V. 1. 368 p., V. 2. 368 p.
3. Дуда Р., Харм П. Распознавание образов и анализ сцен. М.: Мир, 1976. 147 с.

Институт солнечно-земной физики СО РАН, Иркутск
Объединенный институт ядерных исследований, Дубна

4. Ramm A.G. Random Estimation Theory. Harlow: Longman, 1990. 271 p.
5. Ярославский Л.П. Введение в цифровую обработку изображений. М.: Сов. радио, 1979. 312 с.
6. Fukunaga K. Introduction to statistical pattern recognition. Academic Press, New York and London, 1972. 360 p.
7. De Albuquerque Araujo A. // Electron. Lett. 1985. V. 21. N 25–26. P. 1219–1221.
8. Baginyan S. et al. // Comput. Phys. Commun. N 79. 1994. P. 165–178.
9. Ososkov G.A., Zaznobina E.G. Neural Network Application for Data Analysis of Vertical Ionosphere Locating // Preprint Communication of the Joint Institute for Nuclear Research. Dubna, 1993. 12 p.
10. Grozov V.P., Nosov V.E., Ososkov G.A., Zaznobina E.G. Research results on the artificial neural network method in computer-aided ionogram analysis // XXVth General Assembly of URSI, Abstract (1996). P. 395.
11. Hopfield J.J. Neural networks and physical system with emergent collective computational abilities // Proc. Natl. Acad. Sci. USA. 1982. V. 79. P. 2254–2258.
12. Ososkov G.A. Robust tracing by cellular automata and neural network with nonlocal weights, in Applications and Science of Artificial Neural Networks, S.K. Rogers, D.W. Ruck, Editors // Proc. SPIE (1995). V. 2492. P. 1180–91.
13. Grozov V.P. et al. An interpretation of data oblique-incidence sounding using the chirp-signal // Proceedings of ISAP'96. Chiba, Japan. P. 693–696.

Поступила в редакцию
6 августа 1997 г.

V.P. Grozov, V.E. Nosov, G.A. Ososkov. Image Processing as Applied to Automatic Ionogram Processing Problems.

Diagnostics of the ionospheric channel is based upon analyzing vertical/oblique-incidence and backscatter sounding ionograms. To carry out such analysis, it is required to solve two main problems: 1) to process ionograms, i.e. to update and improve the image, and to identify signal points (arrival times); and 2) to combine points identified into traces to be subsequently referenced to particular propagation modes. In this paper, to solve problem 1, statistical methods of image processing are used. With regard to problem 2, the use of the Hopfield artificial neural networks (ANN) is considered. Because of the complicated character of traces in the presence of considerable interference, a modified rotor model was employed. By aptly choosing the initial configuration, it was possible to ensure a fast convergence of the network. The model was tested against ionograms from the chirp-soude (ISTP, Irkutsk, 1987–1996). The analysis of results has shown that this approach performs well and offers considerable promise as a tool for solving ionogram processing problems.