

МЕТОД РАСПОЗНАВАНИЯ СПУТНИКОВЫХ ИОНОГРАММ НА ОСНОВЕ АЛГОРИТМА НЕЛДЕРА–МИДА

К. Г. Цыбуля

Описывается метод автоматического определения основных ионосферных параметров, включая критическую частоту и высоту главного ионосферного максимума, по ионограммам спутникового зондирования. Метод основан на алгоритме Нелдера–Мида, как и ранее созданная программа автоматической обработки ионограмм наземного ионозонда «Парус». Отработка и отладка алгоритма производились на данных, полученных спутниковым ионозондом, работавшим на борту спутника «Космос-1809» в 1987 г. Эти данные были приняты и записаны на борту ледокола «Сибирь» во время экспедиции к Северному полюсу, поэтому они охватывают, в основном, северную полярную шапку. Поскольку требования к производительности программного обеспечения очень высоки, используется технология параллельных вычислений.

КЛЮЧЕВЫЕ СЛОВА: СПУТНИКОВОЕ РАДИОЗОНДИРОВАНИЕ ИОНОСФЕРЫ, АВТОМАТИЧЕСКАЯ ОБРАБОТКА ИОНОГРАММ

ВВЕДЕНИЕ

В настоящее время ведутся работы над созданием спутникового ионозонда «Лаэрт» [1], предназначенным для глобального оперативного контроля ионосферы Земли методом импульсного многочастотного радиозондирования. Ионозонды этого типа предполагается разместить на спутниках, находящихся на околоземных орбитах, высота которых составляет от 500 до 3000 км. Предполагается, что полная система будет состоять из четырех спутников.

Этот ионозонд будет производить зондирование ионосферы радиоимпульсами длительностью ~100 мкс с частотой следования ~60 Гц. Частота несущей этих импульсов изменяется от минимально заданной частоты диапазона до максимально заданной частоты с шагом 25 кГц. Длительность получения одной ионограммы не превышает 10 с (в зависимости от заданного режима). Для специальных режимов измерений может быть предусмотрен более длительный период снятия ионограммы. Для регистрации интенсивности предусматривается 255 ступеней. Периодичность сеансов зондирования – 10, 30, 60 с, при этом режим зондирования с периодичностью 10 секунд считается основным.

Информация, полученная в ходе зондирования, помещается в промежуточную память, и затем, после бортовой обработки, формируются телеметрические кадры для отправки на Землю. Возможна передача как в режиме реального времени, так и после хранения информации в бортовой памяти.

После приема телеметрии на Земле данные зондирования поступают в центральную базу данных, откуда они берутся для последующей обработки, т. е., для получения критической частоты f_oF_2 и других числовых ионосферных параметров, используемых для решения практических задач. Критически важной при этом является минимальная задержка при доставке ионограммы и ее обработке.

В настоящее время для диагностики состояния ионосферы в Российской Федерации используются наземные ионозонды, установленные на ионосферных станциях. Обработка производится либо операторами (один раз в час), либо автоматически (четыре раза в час). Однако система спутниковых ионозондов будет производить не несколько сот ионограмм в сутки, как наземная система, а десятки тысяч ионограмм ежедневно. Операторы такой объем информации вручную обработать не смогут, а программное обеспечение, позволяющее обрабатывать ионограммы автоматически, отсутствует. Без такого ПО может возникнуть критическая ситуация, когда данные поступают, но не обрабатываются или обрабатываются лишь выборочно.

Создание программного обеспечения для автоматической обработки ионограмм включает в себя решение задач распознавания образов, которые являются специфическим разделом информатики. В ИПГ имеется опыт разработки и эксплуатации программного обеспечения для автоматической обработки ионограмм, получаемых наземным ионозондом «Парус» [2]. Конечно, наземные ионограммы и

спутниковые отличаются друг от друга, но если основные принципы автоматического распознавания ионограмм будут отработаны на наземных ионозондах, то адаптация программного обеспечения для спутниковых данных будет относительно простой задачей. Поскольку реальные данные ионозонда «Лэарт» пока недоступны, в алгоритме предусмотрены возможности доработки, которые будут указываться в описаниях ниже.

Таким образом, задача состоит в обобщении алгоритма обработки на спутниковые ионограммы. В некоторых отношениях распознавание спутниковых ионограмм является более простым по сравнению с обработкой наземных, поскольку нет необходимости выделять отражения от сложных структур нижней ионосферы – слоев F1, E и E_s. Набор параметров, описывающих верхнюю ионосферу, значительно меньше, чем тот, который используется для описания верхней. Однако имеются и дополнительные проблемы, в частности, высокий уровень шумов на спутниковых ионограммах, особенно на частотах выше критической, а также наличие плазменных резонансов и чрезвычайно сложная структура следов на ионограммах, полученных в полярных областях.

К программному продукту предъявляются следующие требования:

1. Обработка данных должна происходить достаточно быстро. Поскольку периодичность зондирования составляет 10 с, а ионограммы поступают с четырех спутников, то время обработки одной ионограммы не должно превышать 2.5 с. Конечно, возможно использование нескольких параллельных серверов для вычислений, но и в это случае желательно, чтобы время обработки не превышало ~10 с.

2. Обработка должна быть полностью автоматизирована. В случае, если обработка завершается ошибкой по той или иной причине, или достоверность результата низкая, запись, соответствующая ионограмме в базе данных, должна отмечаться флажком «ошибка» или «низкая достоверность» для привлечения внимания операторов при ежедневном контроле.

3. При обработке должны быть определены следующие параметры:

- а) критическая частота слоя F2 (f_oF2) или критическая частота главного максимума ионосферы, если слой F2 не является наиболее плотным;
- б) высота максимума слоя F2 (h_mF2) или высота главного максимума ионосферы;
- в) характеристика толщины слоя (например, полутолщина $BF2$);
- г) оценка неоднородности ионосферы (например, в результате F-рассеяния)
- д) оценка качества распознавания (например, 0 – распознано, 1 – возможно ошибочное распознавание, 2 – попытка распознавания завершилась ошибкой).

Также, желательно получить последовательности точек, описывающие каждый из следов – от обыкновенной (o), необыкновенной (x) и z-компонент отраженной радиоволны. Каждая точка описывается двумя координатами (частота, действующая глубина). Точки должны быть определены для каждого частного канала или, во всяком случае, лежать достаточно плотно, чтобы описать кривизну следов.

БАЗА ДАННЫХ СПУТНИКА «КОСМОС-1809»

Для отработки и отладки алгоритма использовалась база ионограмм, полученных на ионозонде «ИСС-338», работавшем на спутнике «Космос-1809» [3] в 1986–1987 гг. Этот спутник находился на почти круговой орбите с высотой около 980 км. Ионозонд осуществлял зондирование ионосферы на 338 частотных каналах в диапазоне частот от 0.3 до 16 МГц. В 1987 г. была проведена экспедиция на атомном ледоколе «Сибирь» к Северному полюсу. Среди прочего, на судне была размещена приемная станция, которая принимала и записывала на магнитограф аналоговый сигнал ионозонда, содержащий ионограммы. Кроме того, при каждом зондировании производилась запись уровня естественных радишумов на каждом из частотных каналов. Эти данные также содержались в принятом сигнале. Сброс ионограмм производился в реальном режиме времени, т. е. без записи в память ионозонда. Это позволяло получить ионограммы гораздо лучшего качества, чем те, которые в сильно урезанном виде записывались в ограниченную бортовую память спутника для передачи по штатным телеметрическим каналам. Впоследствии этот аналоговый сигнал был считан с лент магнитографа, оцифрован, раскодирован и занесен в компьютерную базу данных.

Оцифрованный результат каждого зондирования зонда «ИСС-338» представляет последовательность из 338 массивов частотных каналов, каждый массив, в свою очередь, состоит из 640 элементов, соответствующих действующей высоте (точнее, глубине вниз от спутника) от 0 до 2000 км. Каждый однобайтовый элемент представляет собой интенсивность отражения импульса ионозонда, соответствующий данной глубине и частоте. Получившийся результат может быть схематически представлен в виде ионограммы (Рис. 1), где по горизонтальной оси отложена частота, а по вертикальной — расстояние вниз от спутника, вычисленное по радиолокационной задержке между переданным импульсом и принятым эхом. Интенсивность отражения показывается цветом от синего (низкая) до красного (высокая).

Таким образом, исходная таблица базы данных, в которой содержатся результаты зондирования, имеет следующие поля

1. номер ионограммы;
2. номер витка спутника, на котором был произведен сеанс записи;
3. широта точки зондирования;
4. долгота точки зондирования;
5. высота спутника над поверхностью в момент зондирования;
6. запись шумов на всех частотных каналах (массив из 338 элементов);
7. запись сигнала ионограммы (матрица 338x640 байтовых элементов, описывающих интенсивность).

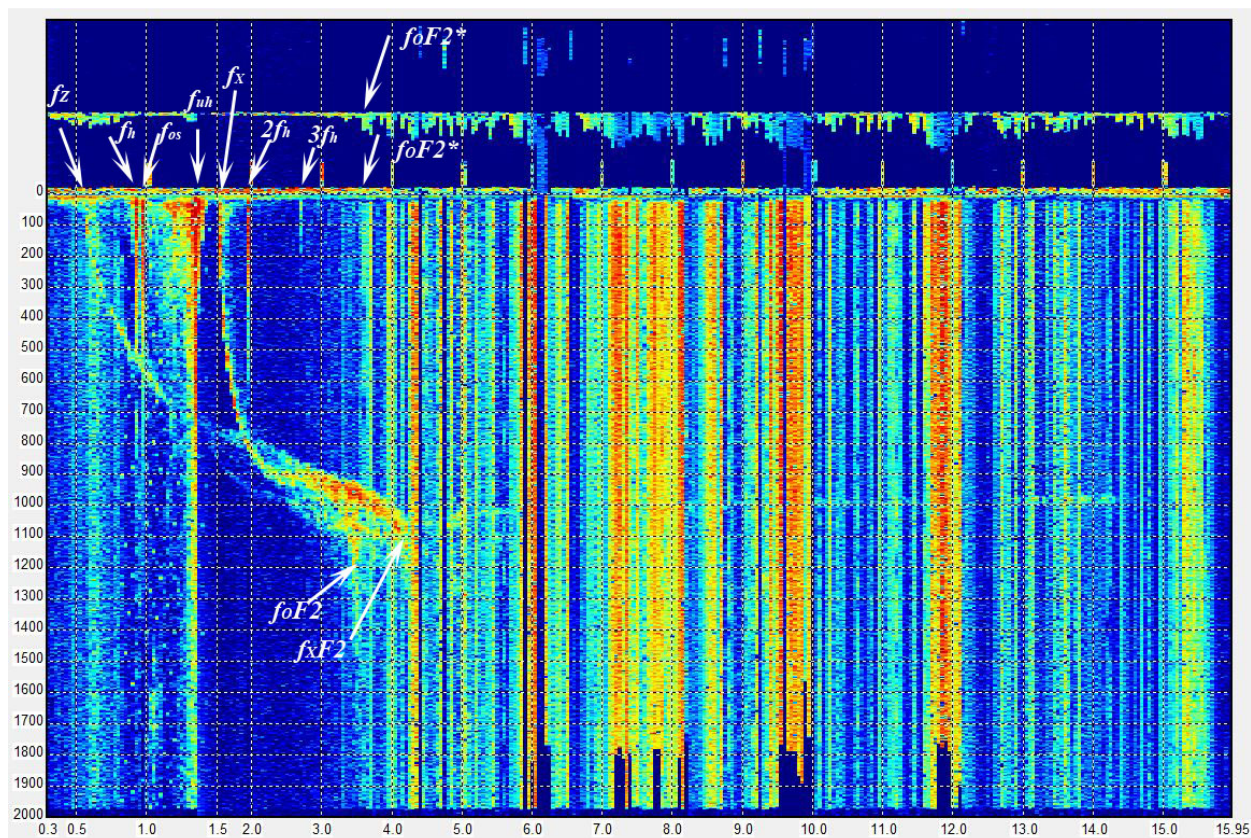


Рис. 1. Ионограмма спутника «Космос-1809» после расшифровки и оцифровки, с отмеченными характерными частотами. Над ионограммой, в виде направленных сверху вниз вертикальных столбиков, приведен записанный уровень радишумов для каждого из частотных каналов.

ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ИОНОГРАММ

Как видно из рисунка 1, в области частот выше f_oF2 ионограмма сильно зашумлена из-за помех работающих на Земле радиостанций, образующих характерную решетку из узких линий. Эти помехи сильно мешают определению точных значений f_oF2 и особенно f_xF2 . Поэтому перед обработкой из

каждого элемента матрицы вычитается среднее (или медианное) значение интенсивности для данного частотного канала. Эта процедура практически не затрагивает свободные от помех каналы, но ослабляет влияние шума на зашумленных каналах. Кроме того, в описаниях последующих алгоритмов предполагается, что никаких операций над матрицей на частотах выше оценочной f_oF2 не производится. Однако из зашумленной части ионограммы можно извлечь и полезную информацию – частота, на которой происходит резкое увеличение уровня шумов на ионограмме (или на спектре радиозумов) является предварительной оценкой f_oF2 . Также можно получить оценку f_oF2 , используя ионосферные модели, такие как International Reference Ionosphere (IRI-2012) [5] или SIMP [6].

Как видно из того же рисунка 1, в области низких частот на ионограмме видны всплески, обусловленные резонансами в плазме, окружающей антенну спутника. Во-первых, резонансы возникают на гиромагнитной частоте

$$f_H = \frac{eB}{2\pi m},$$

и ее гармониках $2f_H, 3f_H, \dots$. Здесь e, m – заряд и масса электрона, B – индукция геомагнитного поля в окрестности спутника. Кроме того, выраженный резонанс существует на верхней гибридной частоте

$$f_{uh} = \sqrt{f_o^2 + f_H^2},$$

где f_o – плазменная частота в окрестности спутника, связанная с электронной концентрацией:

$$f_o = \sqrt{\frac{Ne^2}{\epsilon_0 m}},$$

здесь N – электронная концентрация, ϵ_0 – электрическая постоянная.

Хотя резонансы на гиромагнитной и ее гармониках в большинстве случаев достаточно отчетливы, из них не удастся извлечь дополнительной информации, поскольку они жестко связаны с напряженностью магнитного поля в окрестности спутника, а она хорошо описывается моделью International Geomagnetic Reference Field (IGRF).

Резонанс на верхней гибридной частоте, однако, очень выражен, и след от этого резонанса можно использовать при идентификации следа на ионограммах.

Кроме того, при последующих вычислениях требуются значения магнитного поля как в точке нахождения спутника, так и вдоль всей траектории распространения зондирующего радиоимпульса. Они вычисляются с использованием модели IGRF.

Таким образом, предварительная обработка ионограмм состоит из

- а) фильтрации ионограммы с целью ослабления шумов;
- б) вычисления оценочных значений f_oF2 либо по спектру шумов, либо по ионосферной модели;
- в) вычисления высотного профиля магнитного поля над точкой земной поверхности, заданной географическими координатами зондирования в базе данных.

АЛГОРИТМ НЕЛДЕРА–МИДА

Основой обработки ионограмм является оптимизационный алгоритм Нелдера–Мида [4]. Основной рабочий цикл алгоритма распознавания приведен на Рис. 2. Алгоритм подбирает набор параметров, описывающих ионосферный профиль, таким образом, чтобы порожденная таким набором ионограмма наилучшим образом совпадала с реальной. В качестве параметров используются, например, критическая частота ионосферного максимума f_oF2 , его высота h_mF2 и полутолщина $BF2$. Для каждого частотного канала решается прямая задача распространения радиоволны – вычисление времени между посылкой радиоимпульса и приемом отраженного от ионосферных слоев радиоэха. По набору задержек на всех частотах составляется симулированная ионограмма-шаблон (Рис. 3). Толщина следов определяется в соответствии с известной длиной

импульса ионозонда. Затем производится сравнение следов, то есть шаблон накладывается на обрабатываемую ионограмму и определяется интегральная интенсивность сигнала внутри области, покрытой шаблоном. Если совпадение хорошее, суммарная интенсивность будет высокой. Теперь можно изменить значения ионосферных параметров, провести новый расчет шаблона и сравнение с обрабатываемой ионограммой. Если результат сравнения будет большим по значению, то новые ионосферные параметры описывают ионосферу лучше.

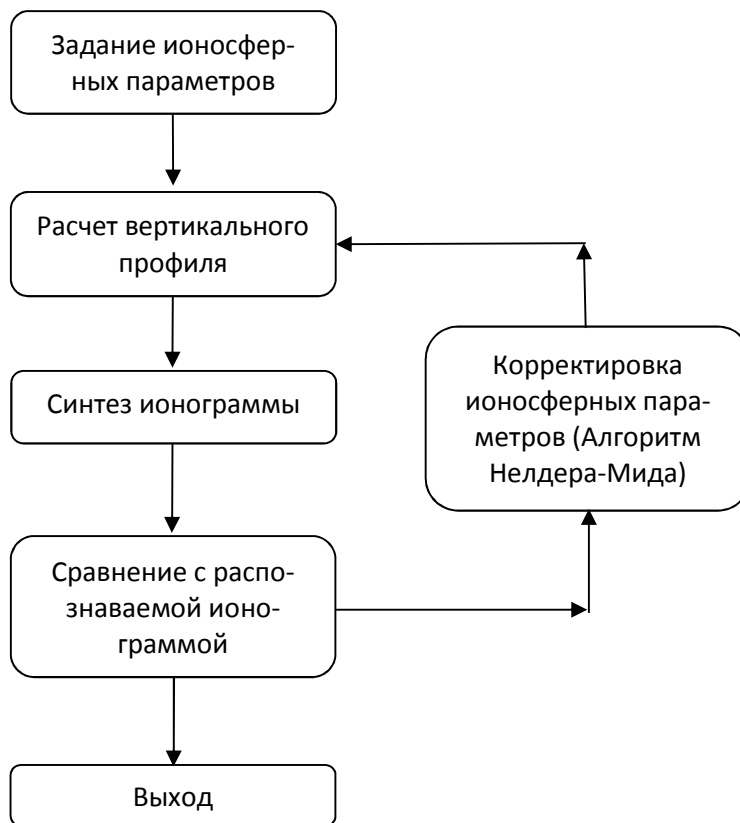


Рис. 2. Принципиальная схема алгоритма распознавания ионограммы.

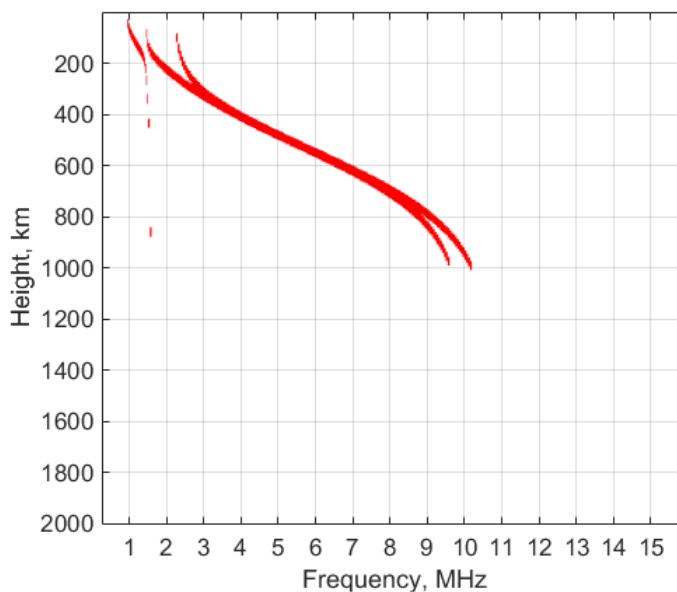


Рис. 3. Синтезированная ионограмма-шаблон.

Таким образом, варьируя параметры, в большинстве случаев можно подобрать достаточно хорошие для практического использования результаты. Если, однако, ионограмма имеет нетривиальный характер из-за сильного рассеяния на мелких неоднородностях или сложных плазменных структур в полярной шапке, то подобрать удачные значения параметров не удастся. В таких случаях ионограмма отмечается как непрошедшая автоматическую обработку.

Надо отметить, что используемые в практических целях наборы ионосферных параметров предполагают, что ионосфера является плоскостной или близкой к ней, и вертикальный профиль электронной концентрации достаточно хорошо определен. Поэтому в случаях, когда алгоритм не дает результата, ручная обработка также не привела бы к четкому определению стандартных ионосферных параметров.

Набор параметров для будущей системы, как и многие другие ее черты, невозможно определить четко в настоящее время, поэтому ниже следующее описание алгоритма приводится для произвольного числа параметров. Когда реальные данные нового ионозонда станут доступны, можно будет расширить этот набор, добавив характеристику уширения следа из-за неоднородностей ионосферы или другие характеристики, которые позволят лучше описать ионограмму.

С математической точки зрения, алгоритм является способом решения оптимизационной задачи по нахождению максимума функции $M(f_oF2, h_mF2, BF2, \dots)$ в пространстве, размерность которого равна числу ионосферных параметров, используемых для описания вертикального профиля плазменной частоты в ионосфере. Для решения этой задачи используется алгоритм Нелдера–Мида (или симплекс-метод), который позволяет найти максимум функции без вычисления ее производных. Достоинством этого метода является также его относительно низкая чувствительность к шумам.

Сам алгоритм Нелдера–Мида для подбора оптимальных параметров состоит в следующем:

1. Подготовка. Будем считать набор ионосферных параметров $(f_oF2, h_mF2, BF2, \dots)$ вектором в линейном пространстве, размерность которого равна числу этих параметров и обозначать x . Вначале выберем $n + 1$ точку $x_i, i = 1, \dots, n + 1$. Будем считать эти точки вершинами симплекса (тетраэдра в n -мерном пространстве). Вычислим в этих точках значения функции $M(x_i)$, т.е. рассчитаем для каждого набора ионосферных параметров ионограмму-шаблон, и произведем сравнение с обрабатываемой ионограммой, получив скалярную метрику сравнения $M(x_i)$ для каждой вершины симплекса.

2. Сортировка. Отсортируем все вершины симплекса по увеличению значения функции $M(x_i)$ и выберем три точки: x_l с наименьшим значением функции $M_l = M(x_l)$, x_g с вторым наименьшим значением $M_g = M(x_g)$ и x_h с наибольшим $M_h = M(x_h)$.

3. Центр тяжести. Найдём центр тяжести всех точек, за исключением оптимальной точки x_l :

$$x_c = \frac{1}{n} \sum_{i \neq h} x_i.$$

4. Отражение. Отразим точку x_l относительно x_c , т.е. определим

$$x_r = 2x_c - x_l$$

и вычислим в ней значение функции $M_r = M(x_r)$.

5. Ветвление. Сравним M_r с M_h, M_g, M_l и предпримем следующие действия.

Если $M_r > M_h$, то найдена новая, более близкая к максимуму точка и следует попробовать увеличить шаг. Определим новую точку

$$x_e = 2x_r - x_c$$

и вычислим значение функции в ней $M_e = M(x_e)$. Если значение функции в ней еще больше, т.е. $M_e > M_r$, то модифицируем симплекс, заменяя точку x_r на точку x_e и закончим итерацию (перейдем на шаг 8). Если $M_e \leq M_r$, то заменим точку M_l на x_r .

Если $M_h > M_r \geq M_g$, то выбор новой точки достаточно хороший, т. к. результат лучше, чем во всех точках, кроме одной. Присвоим точке x_l значение x_r и закончим итерацию.

Если $M_g > M_r \geq M_l$, то изменим значения x_r и x_l и перейдем к следующему шагу.

Если ни одно из условий не выполняется (т.е., $M_l > M_r$), то просто перейдем к следующему шагу.

6. Сортировка. Заново отсортируем вершины так, чтобы x_l была точкой с наименьшим значением функции $M(x_i)$, x_g – со вторым наименьшим, а x_h с наибольшим.

7. Сжатие. Определим точку

$$x_s = \frac{x_l + x_g}{2}.$$

и вычислим в ней значение функции $M_s = M(x_s)$.

Если $M_s > M_l$, то присвоим точке x_l значение x_s и завершим итерацию.

Если $M_s \leq M_l$, то максимум, по-видимому, лежит внутри симплекса. В этом случае надо произвести пропорциональное сжатие симплекса к точке с оптимальным значением x_h .

$$x_i \rightarrow x_h + \frac{x_i - x_h}{2}, i \neq h.$$

8. Проверка условия выхода. Если размер симплекса, т.е., разброс значений в наихудшей и наилучшей точках меньше заданного по всем измерениям, произведем выход из цикла. Если ни одно из условий выхода не выполняется, повторим цикл, т. е. перейдем опять на шаг 3.

ТЕХНИЧЕСКИЕ ДЕТАЛИ РЕАЛИЗАЦИИ АЛГОРИТМА

Решение прямой задачи распространения радиоимпульса

Для синтеза ионограммы необходимо произвести расчет действующих высот отражения сигнала для всех частотных каналов. Расчет производится как интегрирование группового показателя преломления $\mu'(h)$, начиная от высоты отражения h_r , где этот показатель становится из мнимого действительным, вверх до высоты спутника h_s :

$$h' = \int_{h_r}^{h_s} \mu'(h) dh.$$

Групповые показатели преломления для обыкновенной (μ'_+) и необыкновенной (μ'_-) волн определяются по следующим формулам:

$$\mu'_\pm = \mu_\pm + \frac{2}{\mu_\pm D} \left[1 - \mu_\pm^2 - X^2 \pm \frac{(1 - \mu_\pm^2)(1 - X^2)Y_L^2}{A} \right];$$

$$\mu_\pm = \sqrt{1 - \frac{2X(1 - X)}{D}};$$

$$A = \sqrt{Y_T^4 + 4Y_L^2(1 - X)^2};$$

$$D = 2(1 - X) - Y_T^2 \pm A.$$

Частота зондирующего импульса f и наклонение магнитного поля θ входят в следующие параметры:

$$X = \left(\frac{f_0}{f}\right)^2;$$

$$Y = f_H/f;$$

$$Y_L = Y \cos \theta;$$

$$Y_T = Y \sin \theta.$$

Этих определений достаточно для полного определения ионограммы по заданному профилю ионосферной электронной концентрации и связанной с ней частоты f_o , если учесть также ширину следа, которая определяется продолжительностью импульса ионозонда и, возможно, сильные резонансы на частотах f_h и f_{uh} , которые проявляются в виде вертикальных полос.

Форма профиля

Для описания высотного профиля плазменной частоты модельного слоя F2 используются функции Эпштейна:

$$f(h) = F(h - h_m F2, B2);$$

$$F(x, B) = \frac{4e^{x/B}}{(1 + e^{x/B})^2}.$$

Для более точной подгонки могут использоваться и другие модельные слои, такие как:

гауссовский $F(x, B) = e^{-(x/B)^2};$

Аньези $F(x, B) = \frac{1}{1+(x/B)^2};$

параболический $F(x, B) = 1 - \left(\frac{x}{B}\right)^2$, если $|x| < B$, иначе 0;

косинусный $F(x, B) = \left(1 + \cos \pi \frac{x}{B}\right)/2$, если $|x| < B$, иначе 0.

Метрика сравнения

Введем следующие понятия:

Метрика реальной (т.е., обрабатываемой) ионограммы M_R – сумма всех элементов (интенсивностей) матрицы ионограммы.

Метрика синтезированной ионограммы M_S – число всех пикселей шаблона.

Перекрестная метрика M_C – сумма всех элементов обрабатываемой ионограммы, попадающих внутрь шаблона.

Казалось бы, для оценки совпадения шаблона с обрабатываемой ионограммой можно использовать перекрестную метрику M_C . Однако во многих случаях алгоритм стремится увеличить шаблон по площади, так что в него попадает как можно больше точек-данных, включая шумовые. Чтобы исключить преимущество шаблонов с большой площадью, лучше использовать отношение M_C/M_S . Если бы интенсивность сигнала везде была равной единице, то это отношение было бы просто процентом заполнения шаблона реальными данными. Но и этот вариант часто оказывается неработающим, так как в этом случае преимущество имеют шаблоны с малой площадью, которые стремятся наложиться на области с высокой заполненностью и интенсивностью – процент заполнения в этом случае может быть близок к 100%, хотя большие части следов на обрабатываемой ионограмме лежат вне шаблона. Поэтому оптимальной является выбор метрики в виде

$$M = \frac{M_C^2}{M_S M_R}$$

С практической точки зрения, наиболее важным из ионосферных параметров является критическая частота f_oF2 , которая определяется наиболее точно по тонкому ответвлению следа обыкновенной волны в нижней части ионограммы. Это ответвление, однако, занимает небольшую площадь и часто алгоритм предпочитает лучшую подгонку других частей следа. Это особенно резко проявляется в случаях, когда форма профиля плазменной частоты отклоняется от избранной модельной (например, эпштейновской). Для того, чтобы исправить эту проблему можно при расчете перекрестной метрики добавлять нижние (более удаленные от передатчика) элементы с весовым коэффициентом, большим единицы. Введем, например, характерную высоту h_{lim} , равную разности высоты спутника h_s и высоты максимума модельного профиля $h_m F2$. Тогда весовой коэффициент для высоты h можно положить равным

$$w = \left(\frac{h_s - h}{h_{lim}} \right)^3,$$

в случае если действующая глубина $h_s - h$ больше h_{lim} , или единице, если $h_s - h < h_{lim}$.

Нефизические значения параметров

Алгоритм Нелдера–Мида требует определения оптимизируемой функции для всех действительных значений параметров, в то время как физический смысл имеют ионосферные параметры только в достаточно жестко заданной области. Например, все параметры должны быть положительными, кроме того, на каждый из них накладываются ограничения сверху. Но в ходе работы алгоритма значения параметров могут легко выйти из этой области. Особенно часто это происходит при первых итерациях, когда параметры изменяются большими шагами. Допустим, значение параметра f_oF2 на обрабатываемой ионограмме является небольшим, в то время как значения в точках симплекса – в полтора-два раза большими. В этом случае на этапе отражения может возникнуть запрос на синтез ионограммы с небольшим отрицательным значением f_oF2 . Просто прекратить вычисления как заведомо ошибочные также нельзя – искомое оптимальное значение может лежать внутри или вблизи симплекса и прекращение обработки будет означать, что информативную ионограмму не удастся обработать. Поэтому в алгоритм вводится система штрафов: если модификация параметров приводит к выходу хотя бы одного из них из физической области, то вместо синтеза ионограммы и вычисления положительной метрики сравнения выдается отрицательная метрика, заведомо меньшая, чем при любых физических значениях параметров. Модуль метрики тем больше, чем больше отклонение параметра от предельного значения, таким образом, оптимизационный алгоритм «подталкивается» к возвращению симплекса в физически осмысленную область.

Условия выхода

Выход из цикла Нелдера–Мида происходит тогда, когда достигается желаемая точность, т.е., когда разность параметров между наилучшей и наихудшей вершиной симплекса становится меньше заданных значений. Таким образом, для каждого из параметров, наряду с диапазоном физических значений, необходимо задать и приемлемую точность.

Однако для сильно зашумленных и «пустых» ионограмм, а также для ионограмм со сложной структурой, алгоритм сходится медленно, поскольку метрика мало изменяется при модификации параметров. Чтобы избежать потери процессорного времени при обработке ионограмм с «плоской» оптимизационной функцией без выраженных максимумов, следует добавить дополнительное условие выхода по достижении минимальной разницы метрики между вершинами симплекса с максимальным и минимальным ее значением.

Также, чтобы гарантировать невозможность заикливания алгоритма, добавляется условия выхода по счетчику итераций, т.е., происходит аварийный выход, если число итераций превысило разумный предел (~100–200 циклов).

Локальные максимумы

Недостатком метода Нелдера–Мида является невозможность отличить локальный максимум от глобального. Иными словами, может возникнуть такая ситуация, что модификация любого ионосферного параметра на небольшую величину приводит к ослаблению суммарной интенсивности внутри шаблона, тем не менее, шаблон совершенно не совпадает с реальными следами. Чтобы преодолеть эту проблему, создается набор начальных симплексов, покрывающих всю область физических значений параметров. Для каждого из них производится прогон алгоритма до выхода. Глобальным максимумом считается та из точек сходимости алгоритма, для которой значение оптимизационной функции наибольшее. Хотя различные соседние симплексы часто сходятся к одной и той же точке, покрытие области должно быть достаточно плотным для надежной работы алгоритма распознавания. Поэтому такой подход резко увеличивает необходимое для распознавания ионограммы процессорное время. Однако в этой ситуации есть выгодная сторона – обработка каждого из начальных симплексов не зависит от остальных, поэтому вычисления могут производиться одновременно (параллельно).

Применение графического процессора

Технически решить эту проблему можно разными способами, например, используя мощный многопроцессорный вычислительный сервер или создав компьютерный кластер для параллельных вычислений. Для отработки алгоритма использовался подход, не требующий больших затрат на аппаратное обеспечение – технология GPGPU (General-purpose computing for graphics processing units, общецелевые вычисления на графических процессорах) [7]. Современные видеокарты включают в себя достаточно мощные процессоры для осуществления сложных вычислений с плавающей точкой, необходимых для современной компьютерной графики. Сейчас имеется возможность использовать эти процессоры и для неграфических вычислений. Специфика ядер графического процессора состоит в том, что возможна быстрая параллельная обработка независимых или мало зависящих друг от друга блоков данных – именно такие вычисления характерны для графических операций. Таким образом, если удастся разбить алгоритм на одновременно исполняемые потоки, можно добиться прироста производительности, хотя каждое из ядер графического процессора обладает меньшей мощностью, чем центральный процессор компьютера.

Для реализации алгоритма была выбрана библиотека C++ AMP, входящая в состав компилятора Microsoft Visual C++ 2012, которая позволяет перенести алгоритмы на графический процессор без больших изменений. Графический процессор используется для расчета действующей высоты отражения – параллельно для нескольких частотных каналов, а также для вычисления метрик сравнения – параллельно для разных высотных уровней, причем окончательное суммирование производится на центральном процессоре. Также возможно применение параллельных вычислений для предварительной обработки ионограммы, например, фильтрации шумов.

Инициализация симплексов

Для создания набора начальных симплексов область параметров, в которой ищется максимальное значение оптимизационной функции, разбивается n -мерной сеткой с заданным шагом. Однако каждая ячейка такой сетки представляет собой n -мерный параллелепипед, а не симплекс.

Для выбора координат $n + 1$ вершины симплекса используется следующая схема. Для первой вершины все значения параметров выбираются минимальными. Для второй вершины первый параметр устанавливается в максимальное значение. Для вершины m параметр с номером $m - 1$ устанавливается в максимальное значение, параметры с меньшими номерами – в средние $[(\text{макс} + \text{мин}) / 2]$, а с большими – остаются минимальными. Это иллюстрируется следующей таблицей:

	Параметр 1	Параметр 2	...	Параметр n
Вершина 1	мин	мин	...	мин
Вершина 2	макс	мин	...	мин
Вершина 3	сред	макс	...	мин
...
Вершина $n + 1$	сред	сред	...	макс

Чтобы уменьшить затраты процессорного времени, минимальное и максимальное значение каждого параметра можно выбрать на основе величин, характерные для данной широты, местного времени, сезона и солнечной активности. Для этого можно использовать климатические ионосферные модели (IRI или SIMP).

Окончательная подгонка

Реальный высотный профиль плазменной частоты всегда будет отличаться от модельного, поэтому шаблон следа никогда не ляжет точно на след. Если имеется необходимость воспроизвести профиль по следу, решив обратную задачу, то следует определить очертания следа более точно, чем позволяет шаблон. Задача решается за счет поиска максимума интенсивности сигнала вблизи части, вырезанной шаблоном. Поиск проводится на каждом частотном канале в отдельности. После выделения точек максимума, через них проводится сглаженная кривая, являющаяся сплайном 5–7 порядка. Последовательность действующих высот на этой кривой и является входным блоком данных для алгоритма восстановления высотных профилей плазменной частоты и электронной концентрации. Предполагается использовать один из существующих вариантов алгоритма (см., например, [8]).

ЗАКЛЮЧЕНИЕ

В результате проведенной работы создан алгоритм автоматической обработки ионограмм спутникового зондирования. Алгоритм использует матрицу спутниковой ионограммы как входные данные и после обработки выдает как основные числовые параметры, описывающие ионосферу, так и последовательности действующих высот, описывающие следы отражений. Отладка производилась на базе данных ионозонда, работавшего на борту спутника «Космос-1809». Хотя общие принципы работы алгоритма можно считать установленными, для уверенной работы с данными ионозонда «Лаэрт» потребуется полная переработка блока ввода информации, поскольку ионозонд «Лаэрт» будет использовать другой формат данных. Также, вероятно, понадобится некоторая доработка алгоритма для более уверенного распознавания следов, поскольку новый ионозонд будет иметь другие шумовые характеристики и другую контрастность следа. Приняты меры по эффективному распараллеливанию алгоритма для максимально эффективной обработки поступающих данных.

Работа выполнена при поддержке гранта РФФИ 15-29-06052.

ЛИТЕРАТУРА

1. Пулинец С.А. Перспективы спутникового зондирования ионосферы в рамках проекта «Ионозонд». Гелиогеофизические исследования, 2013, № 6, с. 1–7.
2. Гивишвили Г.В., Крашенинников И.В., Лещенко Л.Н., Власов Ю.М., Кузьмин А.В. Ионозонд «Парус-А»: функциональные возможности и перспективы развития. Гелиогеофизические исследования, 2013, № 2, с. 68–74.
3. Pulnits S. Kosmos 1809 Topside Sounder As A Tool For Remote and In-situ Sensing of The Earth's Ionosphere. EGS XXVII General Assembly, Nice, 21–26 April 2002, abstract 2663.
4. Nelder J.A. and Mead R. A Simplex Method For Function Minimization. Computer Journal, 1965, V. 7, p. 308–313.
5. Bilitza D., McKinnell L.-A., Reinisch B., Fuller-Rowell T., The International Reference Ionosphere (IRI) Today And In The Future, J. Geodesy, V.85, p.909–920, 2011.
6. Mikhailov A.V., Perrone L. A Method For foF2 Short-Term (1–24 h) Forecast Using Both Historical And Real-Time foF2 Observations Over European Stations: Euromap Model. Radio Science, 2014, V. 49, p. 253–270.
7. Gregory K., Miller A. C++ AMP: Accelerated Massive Parallelism With Microsoft Visual C++. Microsoft, 2012, 326 p.
8. Titheridge J E. The real height analysis of ionograms: a generalized formulation. Radio Sci., 1988, V. 23, p. 831–837.

A TECHNIQUE OF AUTOMATIC SCALING OF THE SATELLITE IONOGRAMS BASED ON THE NELDER–MEAD ALGORITHM

K.G. Tsybulya

This article describes an approach to automatic computing of the main ionospheric parameters, including the critical frequency and height of the main ionospheric maximum, from the top-sounding ionograms. The technique is based on the Nelder–Mead algorithm which was used in the past to create a software to scale automatically the ionograms of the ground-based *Parus* ionosonde. The ionosonde data from the *Kosmos-1809* satellite were used to work out and adjust the algorithm. These data were received and recorded on board of *Sibir* icebreaker in 1987 and cover mostly the Northern polar cap.

KEYWORDS: IONOSPHERIC TOP-SOUNDING, AUTOMATIC IONOGRAM SCALING