

Российская академия наук
Сибирское отделение
Институт вычислительных технологий

"УТВЕРЖДАЮ"
Директор ИВТ СО РАН

академик _____ Ю. И. Шокин

1 сентября 2009 года

**«Подготовка цифровых батиметрических данных на регуляр-
ной сетке для Дальневосточных акваторий России»**

ВТОРОЙ ПРОМЕЖУТОЧНЫЙ ИНФОРМАЦИОННЫЙ ОТЧЕТ

(контракт № 17/09-1 от 17 июля 2009 г.)

Руководитель работ

г.н.с., д.ф.-м.н., профессор

_____ Л. Б. Чубаров

Новосибирск, 2009 г

Исполнители:

д.ф.-м.н.	В.К.Гусяков
д.ф.-м.н.	Ан.Г.Марчук
к.ф.-м.н.	А.Ю.Бежаев

Введение	3
Основные цели и задачи выполнения работ.....	3
Содержание работ второго этапа.....	3
1. Описание общего сплайнового алгоритма	3
1.1 Описание примера	4
1.2. Локальность.....	4
1.3. Инвариантность.....	5
1.4. Адаптация алгоритма для Курило-Камчатского региона	5
2. Форматы входных и выходных данных.....	5
2.1. Хаотические данные по батиметрии	5
2.2. В алгоритм можно подключить спутниковые данные по рельефу	6
2.3. Выходные данные	6
3. Библиография.....	6

Введение

Основные цели и задачи выполнения работ.

Целью создания батиметрических данных для акваторий, прилегающих к побережью Камчатки и ряда Курильских островов, является обеспечение необходимой точности входных данных для проведения численного моделирования трансформации волн цунами в рассматриваемых акваториях.

Качество батиметрической информации определяет уровень достоверности воспроизведения основных характеристик проявления волн цунами у побережья и, таким образом, определяет уровень достоверности прогнозных параметров, вырабатываемых на основе результатов такого численного моделирования. Справедливость этого утверждения доказана как теоретическими исследованиями, так и практикой модельных и производственных расчетов.

Существующие батиметрические базы на регулярной сетке не способны обеспечить необходимую точность воспроизведения геометрических характеристик береговой линии и распределения глубин вблизи берега. Использование недостоверных входных данных приводит к понижению достоверности результатов вычислительных работ.

Содержание работ второго этапа.

В соответствии со Спецификацией к Контракту на втором этапе, завершившемся 1 сентября 2009 года, задача этапа формулировалась следующим образом: «Разработать пакет демонстрационных компьютерных программ для создания цифровой батиметрии на регулярных сетках с детальностью до 30 географических секунд и документацию к нему с описанием ввода-вывода, инструкцией для пользователя.

Поставленная задача решена. В настоящем отчете указаны методы программной реализации алгоритма, включая детали его функционирования, а также методики построения сплайнового метода функций Грина, на котором базируется алгоритм. В качестве иллюстрации приводится линейный алгоритм восстановления, который может рассматриваться как частный случай общего подхода.

В дополнение приводится описание структур данных, которые могут быть представлены как входные данные алгоритма. Выходные данные представляются на регулярной сетке с разрешением в 30 секунд, в бинарном и текстовом виде.

1. Описание общего сплайнового алгоритма

Представленный алгоритм использует многие алгоритмы теории сплайнов и ее подходы. В основе лежит теория воспроизводящих отображений для интерполяционных сплайнов, функции Грина [1], и локальные способы интерполяции [2]. Детали алгоритма публиковались ранее в [2-3]. В данном документе мы их излагаем более подробно.

Адаптация алгоритмов для реального региона состоит в специализации алгоритмов и выборе небольшого числа параметров на основе визуального сравнения результатов. В частности, в алгоритмах поиска ближайших точек (для восстановления) делается небольшая коррекция на широту. В других местах производилась коррекция параметров функций Грина, которые трудно интерпретировать физически (логически). Впрочем, кажется, этот выбор не должен сильно влиять на результат использования данных в крупно-масштабных задачах, например, при расчете волн цунами в океане и прибрежных районах.

Комбинирование алгоритмов позволяет сделать алгоритм локально-глобальным и пространственно инвариантным. Локальный подход позволяет улучшить качество интерполяции в областях с плотным расположением данных и исключить влияние дальних данных. Глобальный подход позволяет получить визуально хорошую картинку даже в местах отсутствия данных за счет привлечения пространственно удаленных данных.

Использование подобных (практически идентичных) методов как для локальной так и для глобальной интерполяции позволяет увидеть их пространственную инвариантность. Такие методы подходы активно используются в теории фракталов (в применении к сжатию данных). Фракталы еще и аффинно-инвариантны. В предлагаемом алгоритме под пространственной инва-

риантность подразумевается инвариантность относительно вращения и однородного (по координатам растяжения).

На примере следующего рисунка мы проиллюстрируем важные идеи алгоритма: локальность, инвариантность и пр.

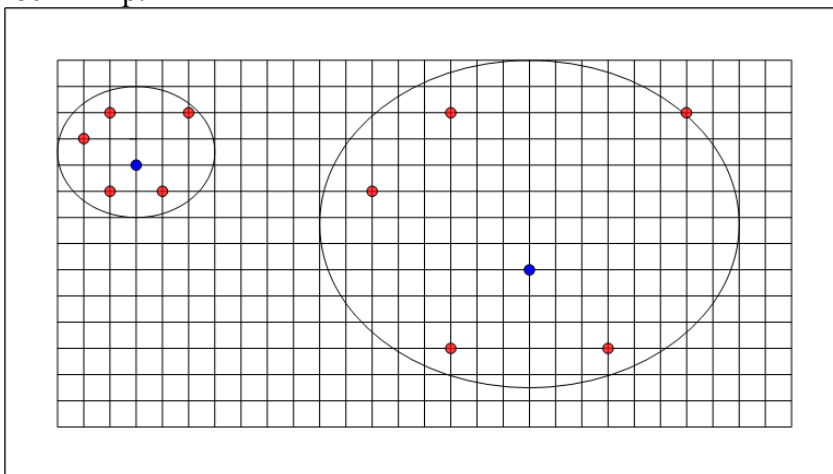


Рис 1. Иллюстрация алгоритма

На этом рисунке видны:

- Прямоугольная сетка с размерами 28x15. Эта сетка, на которой необходимо построить батиметрию.
- 10 красных точек, представляющих исходные данные. Кроме их геометрического (X,Y) расположения, в них заданы глубины Z. На рисунке эти точки расположены точно в узлах расчетной сетки, но это вовсе необязательно.
- 2 синих точки, в которых необходимо рассчитать батиметрию. Конечно, значения Z в этих точках заранее неизвестны.
- 2 эллипса, иллюстрирующие два независимых кластера
- Кластеры специально выбраны похожими (один получается растяжением другого), чтобы проиллюстрировать несколько деталей алгоритма.

1.1 Описание примера

Подобная конфигурация содержит мало входных данных. Она, конечно, недостаточна для получения качественного результата. Исходные данные для Курило-Камчатского региона будут содержать более плотное покрытие, которое будет тщательно отобрано для обеспечения результата в критических прибрежных районах.

Однако, это модельный случай, и можно представить, что имеются другие интерполяционные узлы в регионе, кроме как в этих двух эллипсах.

Расчет значений в каждом узле расчетной сетке 28x15 ведется независимо, сверху вниз и слева направо. В этом случае при расчете первой встретится синяя точка, которая расположена в малом эллипсе, и через какое-то время встретится синяя точка, расположенная в большом эллипсе.

Согласно алгоритму для расчета батиметрии в первой точке будут использованы данные (5 точек) из малого эллипса. Для расчета батиметрии во второй точке будут использованы данные из большого эллипса.

1.2. Локальность

Это очень важное свойство. Глобальный подход обеспечивает идеальную точность, только в случае хорошей гладкости восстанавливаемой функции, и хорошем расположении узлов (соотношении минимального и максимального соотношения расстояний). На практике это встречается редко, и это ведет к неустойчивости или неробастности алгоритма. Кроме того, глобальные методы неприменимы из-за ограничений по памяти и разрядности.

Локальность данного алгоритма обеспечивается выбором ближайших узлов при интерполяции (см. Первый промежуточный информационный отчет по контракту № 17/09-1 от 17 июля 2009 г., п. 4)

1.3. Инвариантность

Данное свойство считается преимуществом алгоритмов интерполяции. Это касается инвариантности относительно растяжения (сжатия) расположения данных и вращения. Инвариантность относительно вращения в условиях дискретности обеспечить невозможно. Однако можно говорить о некоторой инвариантности, которая обеспечена в условиях точности восстановления на сетке.

1.4. Адаптация алгоритма для Курило-Камчатского региона

В целях ускорения работоспособности структуры данных оптимизированы под решение задач батиметрии в Курило-Камчатском регионе. Это не означает, что входные данные нельзя поменять. Такая возможность предоставляется для уточнения/исправления неточных (с точки зрения интерпретатора) данных.

2. Форматы входных и выходных данных

2.1. Хаотические данные по батиметрии

Они собраны из различных источников, и представлены для программ в простом формате:

Это текстовый файл в формате X Y Z, например,

```
79.79000 51.13500 -3725.0
79.77200 51.13900 -3715.0
79.73500 51.14600 -3770.0
79.71900 51.14900 -3865.0
79.68300 51.15700 -3975.0
79.68300 51.15700 -3975.0
79.67300 51.15600 -4010.0
79.65400 51.15600 -4040.0
79.63200 51.15500 -4005.0
79.62600 51.15400 -3530.0
79.61200 51.15400 -3470.0
79.60100 51.15400 -3285.0
79.59700 51.15300 -3290.0
```

Первая координата X представляет широту, вторая координата X представляет широту долготу, третья координата Z представляет глубину океана в точке (X,Y). Координаты (X,Y) должны покрывать территорию Курило-Камчатского региона [138,170]x[40-65] (в географических градусах) и должны быть предоставлены с достаточной точностью. Такие данные, достаточные для расчета батиметрии на регулярной сетке с точностью до 30 географических секунд, будут представлены для 4-х регионов.

- Координата X должна представлять долготу. Пример: 79.4523. Это значение представляет 79 (и 4523 десятитысячных градуса). Мы отмечаем этот факт, чтобы пользователь не перепутал с координатами 79 градусов 45 минут и 23 секунды.
- Координата Y представляет широту, имеет положительное значение в рамках Курило-Камчатского региона.

Данные должны достаточно плотно покрывать регион. В случае недостаточности данных возможны погрешности в расчетах (в зависимости от плотности).

Предоставленные данные алгоритмом обрабатываются в случае незначительных изменений (10-20% дадут практически ту же точность). Этот запас гарантирует возможность перерасчета, в случае если пользователь захочет исключить часть неправильных данных, которые дают локальные возмущения. С другой стороны пользователь может добавить свои данные в мало-описанные районы, чтобы уточнить локальное поведение батиметрии.

2.2. В алгоритм можно подключить спутниковые данные по рельефу.

Они должны быть представлены на сетке с пространственным шагом в 0,5 географической минуты (30 географических секунд) и иметь вид регулярного массива. Этот массив должен совпадать с сеткой, на которую должен быть произведен расчет.

Этот массив должен представлять таблицу, записанную в текстовом файле. Строчка представляет данные по одной широте. Количество данных в каждой строчке одинаково.

Следующий пример дает простейший (нереальный) образец 3x4:

```
1 10 10 10 1
1 10 100 10 1
1 10 10 10 1
```

Отметим, что географическая привязка отсутствует и дается во входных параметрах алгоритма.

2.3. Выходные данные

Выходные данные представляют глубины (возможно рельеф) в заданном районе с точностью до 1 метра. В текстовом виде они будут выглядеть примерно следующим образом.

```
Number of columns: 1077
Number of rows : 995
Left boundary: 151E 53' 30"
Right boundary: 160E 52' 00"
Low latitude: 46N 40' 30"
High latitude: 54N 58' 00"
-355 -358 -359 -359 -357 -354 -351 -348
-354 -356 -357 -357 -356 -353 -350 -347
-353 -355 -356 -356 -354 -352 -349 -346
-352 -354 -354 -354 -353 -350 -347 -345
-351 -353 -353 -353 -351 -349 -346 -344
-351 -352 -352 -351 -350 -348 -345 -343
-351 -351 -351 -350 -348 -346 -344 -342
-350 -351 -350 -349 -347 -345 -344 -342
-350 -350 -349 -348 -347 -345 -343 -341
-350 -350 -349 -348 -346 -345 -343 -342
-350 -350 -349 -348 -346 -345 -343 -342
-350 -349 -349 -348 -347 -345 -343 -343
-350 -350 -349 -348 -346 -345 -343 -343
-351 -350 -349 -348 -346 -345 -344 -343
-351 -351 -349 -348 -346 -345 -344 -343
```

В бинарном формате заголовок отсутствует, а данные представляются слитно и последовательно по 2 байта на отсчет.

3. Библиография

1. Bezhaev A.Yu., Vasilenko V.A. Variational Theory of Splines. Kluwer Academic/Plenum Publishers, 2001, 280pp
2. Bezhaev A.Yu. Spline interpolation of huge multivariate data. Сибирский журнал вычислительной математики. РАН, Сиб. отд., Новосибирск, 2003, Т. 6, № 3. с.249-261
3. Бежаев А.Ю., Марчук А.Г. Метод создания цифровых массивов глубин океана на регулярных сетках. Труды Международной конференции Математические Методы в Геофизике. ММГ-2003, Новосибирск 2003., ИВМиМГ (ВЦ) СО РАН, г. Новосибирск, 2003, ч.2, с.561-567.