

2. Колесников Е.Г. Введение в численный анализ. М., РУДН, 2002.
3. Жилияков Е.Г. и Фокин Ю.А. Вариационный метод оценивания производных эмпирических функций. – "Журнал вычислительной математики и математической физики", 2002, т. 42, № 8.
4. Жилияков Е.Г., Чудинов С.М. и Созонова Т.Н. Вариационный метод дифференцирования и интерполяции дискретных сигналов. – "Вопросы радиоэлектроники", сер. РЛТ, 2006, вып.1.
5. Рабинер Л. и Голд Б. Теория и применение цифровой обработки сигналов. М., Мир, 1978.
6. Жилияков Е.Г., Белов С.П., Черноморец А.А. Вариационные методы анализа сигналов на основе частотных представлений. – "Вопросы радиоэлектроники", сер. ЭВТ, 2010, вып. 1, с. 10-25.

Статья поступила 12.10.2010

Д.т.н., проф. Е.Г. Жилияков, А.А. Барсук (БелГУ)

E.G. Zhilyakov, A.A. Barsuk

**ПАРАЛЛЕЛЬНЫЕ ВЫЧИСЛЕНИЯ ПРИ АВТОМАТИЧЕСКОЙ
ВАРИАЦИОННОЙ КЛАССИФИКАЦИИ ОБЪЕКТОВ
НА СПУТНИКОВЫХ ФОТОГРАФИЯХ ЗЕМНОЙ ПОВЕРХНОСТИ**

**PARALLEL COMPUTATIONS IN THE METHOD OF AUTOMATIC
VARIATIONAL CLASSIFICATION OF OBJECTS AT THE
SATELLITE EARTH SURFACE PHOTOS**

Изложен подход к реализации параллельных вычислений в методе автоматической вариационной классификации объектов на спутниковых фотографиях земной поверхности, рассмотрены современные технологии разработки параллельных программ.

Key words: parallel computations, MPI, OpenMP, CUDA, object classification, image, variational principle, maximization of partitioning quantity functional, graph splitting, minimal spanning tree.

Одним из основных направлений развития информационных технологий является разработка методов и алгоритмов обработки и анализа изображений, что обусловлено тенденцией использования естественных для человека форм информационного обмена, к которым относятся визуальные отображения реальности. Среди интен-

сивно развивающихся направлений обработки изображений можно выделить анализ космофотоснимков земной поверхности, получаемых со спутников. Важность этого направления исследований определяется задачами обработки данных дистанционного зондирования Земли (ДЗЗ) и, в частности, выделения на изображениях некоторых особенностей (дешифрирование) в целях принятия соответствующих управленческих решений.

Одной из процедур дешифрирования является выделение на изображениях отдельных групп объектов, объединяемых на основе некоторых признаков (классификация), для чего предложены различные подходы. Приписывание объектов в изображениях к определенным классам является, подобно многим другим аспектам обработки изображений и анализа, несомненно, междисциплинарной задачей, которая характерна не только для анализа изображений, а для самого общего типа методов [6].

В настоящее время существует достаточно широкий спектр алгоритмов автоматической классификации объектов на изображениях, использующих в своей основе различные принципы. В современных программных средствах обработки изображений (ENVI, arcGIS, ERMAP Imagine) для решения задачи автоматической классификации объектов используются алгоритмы K-Means (K-внутригрупповых средних) и ISODATA (ИСОМАД). Отдельно следует выделить разработанный авторами метод и алгоритм автоматической вариационной классификации объектов на спутниковых фотографиях земной поверхности, который во многих случаях дает лучшие, более адекватные разбиения с позиции человеческих представлений об адекватности классификации, по сравнению с вышеупомянутыми алгоритмами [5].

Существенной проблемой при использовании многих алгоритмов классификации является их высокая вычислительная сложность в совокупности с большим количеством обрабатываемых объектов. Это относится и к методу автоматической вариационной классификации объектов на спутниковых фотографиях земной поверхности, т.к. в его состав входят алгоритмы, имеющие сложность $O(n^3)$. Основным способом повышения скорости вычислений, и тем самым улучшения как количественных, так и качественных его характеристик, представляется распараллеливание алгоритмов, используемых в методе.

В основе метода лежит функционал качества разбиения, который был сконструирован опираясь на результаты психологических тестов [4] и собственные представления об адекватности разбиения исходного множества на классы. В основе функционала лежит мера информационной однородности. Вид данного функционала представлен выражением 1.

$$L = \frac{1}{V} \quad (1)$$

где

$$V = 1 + \sum_{i=1}^K \frac{\frac{S_i}{Norm} * Ln\left(\frac{S_i}{Norm}\right)}{Ln(K)} \quad (2)$$

$$Norm = \sum_{i=1}^K S_i \quad (3)$$

$$S_i = - \sum_{q=1}^{M_i} \rho_{iq} * Ln(\rho_{iq}) \quad (4)$$

$$\rho_{iq} = \frac{r_{iq}}{R_q} \quad (5)$$

$$\sum_{i=1}^{M_q-1} r_{iq} = R_q \quad (6)$$

где: R_q - общая длина внутренних ребер q -го подмножества;

r_{iq} - длина i -ого ребра в q -ом подмножестве ($i=1, \dots, M_q-1$).

Эксперименты показали, что в задачах классификации объектов на изображениях представленный функционал позволяет получить разбиения, характеризующиеся высоким уровнем адекватности [5].

В упрощенном виде алгоритм вариационной автоматической классификации объектов на спутниковых фотографиях земной поверхности состоит из двух основных частей (рис. 1): подготовки изображения к классификации и классификации множества объектов.

На этапе подготовки изображения к классификации производится квантование по уровню каналов интенсивности пикселей изображения и формирование списка объектов, подлежащих классификации. Задается максимально допустимое количество уровней квантования SK_{max} по одному каналу. Максимально возможное количество векторов пиксельных интенсивностей вычисляется с ис-

пользованием соотношения 7.

$$C_{\max} = CK_{\max}^n \quad (7)$$

где n-количество каналов изображения.

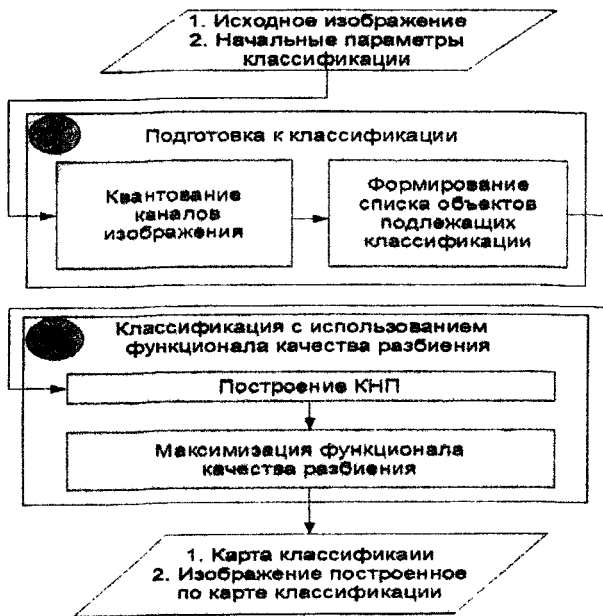


Рис. 1

Упрощенная схема алгоритма вариационной автоматической классификации объектов на изображениях

Новые значения пиксельных интенсивностей рассчитываются согласно соотношения 8.

$$I_{ijk} = I_{\max} \div CK_{\max} \times \left\lfloor r_{ijk} \div (I_{\max} \div CK_{\max}) \right\rfloor, i=1,2,\dots,M; j=1,2,\dots,N \quad (8)$$

Здесь I_{ijk} - интенсивность ij пикселя по k каналу, $I_{\max}$ - максимально возможное значение интенсивности в канале, M, N – ширина и высота изображения в пикселях. После проведения такого преобразования количество различных векторов пиксельных интенсивностей на изображении не превышает $C_{\max}$.

Сложность алгоритма квантования линейна, каждый пиксель изображения обрабатывается независимо. Распараллеливание вычислительного процесса на K вычислительных устройствах позволит достичь выигрыша в производительности в K раз.

В результате квантования на изображении появляется большое количество пикселей, описанных одинаковыми векторами значений интенсивностей. Поэтому следующим этапом будет подсчет и выделение различных векторов интенсивностей, присутствующих на изображении. В основе процедуры подсчета лежит алгоритм пирамидальной сортировки, сложность которого составляет $O(n \cdot \ln(n))$. Использование параллельных алгоритмов сортировки позволит получить ускорение в K раз.

Следующим этапом является построение минимального остового дерева или кратчайшего незамкнутого (КНП). Для этого используется модифицированный алгоритм Краскала, сложность которого составляет $O(n^3)$. Столь высокая сложность обуславливается невозможностью хранения в памяти всех ребер исходного графа. Параллельный алгоритм построения КНП будет иметь временную сложность $O(n^3/K)$ [3]. Несмотря на неизменный порядок временной сложности алгоритма реальный выигрыш во времени его выполнения может составить K раз.

Следующим этапом работы алгоритма автоматической вариационной классификации будет разрезание максимального ребра КНП. На основе полученного разбиения вычисляется функционал качества L (рис. 2).

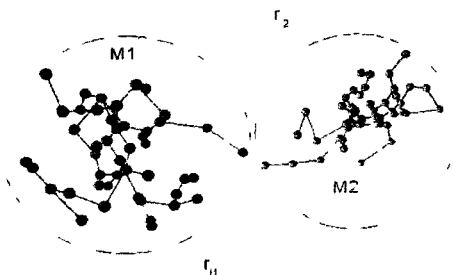


Рис. 2

Пример разрезания КНП на два поддерева соответствующих двум классам

Расчет компонент функционала качества разбиения производится в процессе обхода ребер поддеревьев, полученных в результате разрезания исходного дерева. Очевидно, что такой обход-

можно производить параллельно. Таким образом, возможно получить ускорение в K раз.

В настоящее время самыми широко распространенными технологиями параллельного программирования являются OpenMP, MPI, и технология параллельных вычислений на графических процессорах CUDA.

Одним из наиболее популярных средств программирования для компьютеров с общей памятью, базирующихся на традиционных языках программирования и использовании специальных комментариев, в настоящее время является технология OpenMP. За основу берётся последовательная программа, а для создания её параллельной версии пользователю предоставляется набор директив, функций и переменных окружения. Интерфейс OpenMP задуман как стандарт для программирования на масштабируемых SMP-системах (SSMP, ccNUMA и др.) в модели общей памяти. В стандарт OpenMP входят спецификации набора директив компилятора, вспомогательных функций и переменных среды. OpenMP реализует параллельные вычисления с помощью многопоточности, в которой «главный» поток создает набор «подчиненных» потоков, и задача распределяется между ними. Предполагается, что потоки выполняются параллельно на машине с несколькими процессорами [1]. OpenMP максимально подходит для эффективной реализации параллельных программ для современных многоядерных процессоров.

Наиболее распространенной технологией программирования для параллельных компьютеров с распределенной памятью в настоящее время является *MPI Message Passing Interface*. Основным способом взаимодействия параллельных процессов в таких системах является передача сообщений друг другу. Стандарт MPI фиксирует интерфейс, который должен соблюдаться как системой программирования на каждой вычислительной платформе, так и пользователем при создании своих программ.

Интерфейс MPI поддерживает создание параллельных программ в стиле MIMD (Multiple Instruction Multiple Data), что подразумевает объединение процессов с различными исходными текстами. Однако писать и отлаживать такие программы очень сложно, поэтому на практике программисты гораздо чаще используют SPMD-модель (Single Program Multiple Data) параллельного программирования, в рамках которой для всех параллельных процессов

используется один и тот же код [2]. MPI предназначена для разработки параллельных программ, исполняемых на вычислительных кластерах и суперкомпьютерах.

CUDA (Compute Unified Device Architecture) — программно аппаратная вычислительная архитектура, позволяющая программистам реализовывать алгоритмы, выполнимые на графических процессорах ускорителей GeForce восьмого поколения и старше. CUDA даёт разработчику возможность по своему усмотрению организовывать доступ к набору инструкций графического ускорителя и управлять его памятью, организовывать на нём сложные параллельные вычисления. Графический ускоритель с поддержкой CUDA становится мощной программируемой открытой архитектурой, приближаясь к сегодняшним центральным процессорам, а в массово параллельных задачах существенно превосходя их. В этой технологии реализован SIMD принцип компьютерных вычислений, что накладывает определенные ограничения на эффективное использование CUDA.

Выбор конкретной технологии зависит от масштаба решаемых задач и наличия необходимого оборудования. Программная реализация алгоритма автоматической вариационной классификации объектов на спутниковых фотографиях земной поверхности для ЭВМ класса IBM-PC совместимые может быть разработана с использованием технологий OpenMP и CUDA (при наличии соответствующего графического ускорителя). Это позволит за приемлемое время (минуты) производить обработку 3-канальных снимков с $СК_{\max}=20$. Использование же технологии MPI даст возможность производить вычисления на кластерах и суперкомпьютерах, что позволит изменить порядок ограничений на входные данные.

Работа выполнена при поддержке ФЦП «Научные и научно-педагогические кадры для инновационной России» на 2009-2013 годы, гос. контракт № 14.740.11.0390.

Литература

1. Антонов А.С. Параллельное программирование с использованием технологии OpenMP: Учеб. пособ. М., МГУ, 2009. 77 с.
2. Антонов А.С. Параллельное программирование с использованием технологии MPI: Учеб. пособ. М., МГУ, 2004. 71 с.

3. Макконнелл Дж. Основы современных алгоритмов. 2-е доп. изд. М.; Техносфера, 2004. 368 с.

4. Елкина В.Н. и Загоруйко Н.Г. Количественные критерии качества таксономии и их использование в процессе принятия решений. – "Вычислительные системы", 1969, вып. 36, с.29 – 46.

5. Жилияков Е.Г. и Барсук А.А. О компьютерной реализации автоматической вариационной классификации объектов на спутниковых фотографиях земной поверхности. – "Вопросы радиоэлектроники", сер. ЭВТ, 2010, вып. 1, с. 166-171.

6. Яне Б. Цифровая обработка изображений. М.; Техносфера, 2007. 584 с.

Статья поступила 12 10 2010

Д.т.н., проф. Е.Г. Жилияков, А.Ю. Лихошерстный (БелГУ)

E.G. Zhilyakov , A.J. Lihosherstnyj

ПАРАЛЛЕЛЬНАЯ ОБРАБОТКА АЭРОКОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ НА ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ

PARALLEL PROCESSING OF SPACE IMAGES FOR HIGH- PERFORMANCE COMPUTING SYSTEMS

Проведено распараллеливание алгоритма фильтрации изображений на основе частотных представлений с помощью технологии MPI на суперкомпьютере IBM Blue Gene/P

Key words image, parallel processing, filtering, acceleration

Развитие суперкомпьютерных технологий определяет прогресс в области обработки информации. Такие технологии являются базовыми критическими технологиями, поскольку сегодня они важнейшие в спектре тех, которыми владеет человечество. Именно на их основе решаются наиболее трудные и ресурсоемкие междисциплинарные задачи современной науки, техники, промышленности и бизнеса.

Одной из важнейших задач, решаемых при помощи суперкомпьютерных технологий, является исследование поверхности Земли, а именно: мониторинг влияния изменения климата, оценка загрязнения атмосферы, моделирование наводнений и оползней,