

УДК 004.93'1; 004.932

А.А. Дудкин

ПОДСИСТЕМА ПАРАЛЛЕЛЬНОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ ДЛЯ МОНИТОРИНГА СОСТОЯНИЯ ЛЕСНЫХ УГОДИЙ

Предложена подсистема распараллеливания алгоритмов обработки данных дистанционного зондирования Земли для мониторинга лесных угодий на основе многоагентного подхода. Вычислительный процесс представляется в виде направленного ациклического графа. Вершинам графа соответствуют операции обработки, а дуги задают порядок следования операций. Для каждого типа данных формируется свой граф, затем строится обобщенный граф, на котором исходные графы являются подграфами, задающими конкретные схемы (сценарии) обработки. Такое представление вычислительного процесса дает возможность применить для оптимизации вычислений аппарат теории графов и теории сетей как для статической (построение off-line расписаний), так и для динамической оптимизации (построение on-line расписаний). Для проектирования параллельных алгоритмов и организации параллельных вычислений предлагается интегрированный набор средств, включающий в себя визуальный редактор, транслятор и системы оптимизации расписаний и поддержки параллельных вычислений на базе MPI.

Мониторинг леса; дистанционное зондирование Земли; параллельная обработка; многоагентная система; MPI.

A.A. Doudkin

IMAGE PARALLEL PROCESSING SUBSYSTEM FOR MONITORING FOREST LAND

This paper describes the parallel implementation of image processing algorithms in a subsystem for forest land monitoring based on remote sensing data. A multi agent approach is proposed to use for parallelization. The computational process is represented as a directed acyclic graph. Vertices of the graph correspond to the processing operations and the arcs define the order of operations. The graph is formed for each data type, then a generalized graph is constructed as a composition of the original graphs of specific scheme processing (scenarios). Such representation of computational process provides an opportunity to apply for the optimization calculations of the theory of graphs and network theory both for static (off-line scheduling) and for dynamic optimization (on-line scheduling). An integrated set of tools is proposed for the design of parallel algorithms and parallel computing, including visual editor, compiler, schedule optimization system and a system of parallel computing based on MPI.

Forest monitoring; remote sensing; parallel processing; multi agent system; MPI.

Введение. В настоящее время эффективным инструментом мониторинга с целью предотвращения чрезвычайных ситуаций является обработка и анализ мультиспектральных изображений высокого пространственного разрешения лесных районов, получаемых со спутников дистанционного зондирования Земли (далее – ДЗЗ) [1–4]. Основным этапом тематической обработки при решении задачи мониторинга состояния растительности является выделение площадных объектов, представляющих собой связную область элементов пикселей изображения, которые объединяются в одно множество на основании выбранных заранее свойств пикселей. При этом помимо яркостных свойств изображений могут использоваться и другие характеристики снимка, например текстурные. В процессе сегментации мультиспектральных изображений несколько каналов обрабатываются одновременно. Особенностью алгоритмов, дающих качественные результаты, является то, что они могут работать весьма значительное время даже на мощных аппаратных платформах. У этой проблемы есть два решения: разработка алго-

ритмов с низкой вычислительной сложностью или параллельная обработка данных на нескольких вычислительных устройствах (на кластере, многопроцессорной многоядерной архитектуре или графических процессорах).

Существует ряд зарубежных систем обработки данных систем технического зрения, использующих параллельную и распределенную обработку [5–7]. В основе архитектуры этих систем используются различные технологии, например CORBA [5] или многоагентный подход [6]. В данной работе предлагается использовать для проектирования и организации параллельных вычислений интегрированный набор средств, включающий в себя визуальный редактор, транслятор, систему оптимизации и систему поддержки параллельных вычислений на базе MPI [8]. Использование MPI делает нашу систему широко применимой для различных параллельных компьютеров.

В статье представлена графовая модель параллельного алгоритма и средства для разработки параллельных приложений и организации параллельных вычислений.

Структура аппаратно-программного комплекса (АПК). АПК мониторинга состояния лесных угодий предназначен для осуществления контроля развития заболеваний растительности, основными компонентами которого являются:

- 1) файловая система / ГИС – предназначена для хранения исходных данных;
- 2) подсистема распараллеливания обработки – для выполнения обработки растровых данных;
- 3) подсистема взаимодействия с пользователем – позволяет осуществлять контроль и управление обработкой;
- 4) подсистема принятия решений – для принятия решения о необходимости внесения химикатов и их количестве, исходя из результатов обработки, данных системы управления внесением химикатов и параметров работы, задаваемых пользователем (выбор алгоритма обработки и их параметров).

Методика мониторинга включает:

- ◆ формирование информативных признаков и подбор значений их параметров;
- ◆ обучение нейросетевого классификатора системы с использованием экспертных данных о здоровых и пораженных заболеваниях участках растительности;
- ◆ вычисление карты специальных областей (например участков с развивающимся заболеванием растений) с использованием алгоритма совместной сегментации;
- ◆ получение географической привязки построенных карт и сохранение их в базе данных ГИС;
- ◆ использование полученных карт при принятии решения о необходимом количестве средств защиты растений на том либо ином участке.

В качестве данных, получаемых в реальном времени, могут использоваться:

- ◆ данные глобальной навигационной спутниковой системы;
- ◆ данные от цветной камеры видимого диапазона. Система управления может в реальном времени корректировать данные карты особых областей, принимая более точные решения, что увеличивает эффективность решения задачи.

Модель представления параллельного алгоритма. Параллельная обработка информации выполняется на кластере, в общем случае гетерогенном. Он представляется множеством $P = \{p_1, p_2, \dots, p_m\}$, где p_i – отдельный компьютер (узел кластера). Каждый компьютер p_i имеет характеристику производительности s_i , которая определяет время выполнения одной условной единицы вычислений. Узлы кластера соединены коммуникационными каналами. Каждый канал между узлами p_i и p_j , обозначенный l_{ij} , имеет характеристику b_{ij} , которая определяет ширину канала и скорость передачи данных между p_i и p_j .

При оптимизации расписания для конвейера следует учитывать не только характеристики алгоритма обработки, но и количество поступивших объектов. Сам по себе поток объектов может не быть однородным, т.е. отдельные объекты обрабатываются по различным алгоритмам. Мы называем алгоритм обработки объекта определенного типа сценарием. Сценарии обработки в общем случае содержат множество операций, каждая из которых может избирательно применяться к объектам различных типов либо применяться с различными параметрами в зависимости от типа обрабатываемого объекта. В этом случае поток может быть детерминированным, когда его структура и характеристики объектов известны заранее, или стохастическим, когда количество и порядок следования типов объектов случайны. Средства организации и оптимизации вычислений позволяют управлять обработкой как детерминированного, так и стохастического потока.

Все сценарии обработки представляются в форме направленных ациклических графов (НАГ). Построение приложения для параллельной обработки потока данных состоит из нескольких шагов: создание НАГ сценариев, описывающего логическую структуру программы; назначение операций обработки вершинам графа и определение параметров вызова для каждого типа данных; отображение графа сценариев на топологию вычислительного кластера. Целью обработки потока является минимизация времени обработки N объектов данных.

Созданный НАГ сценариев обработки представляется в форме документа XML, который используется как для визуального представления, так и для выполнения программы средствами организации вычислительного процесса.

Вычислительные операции взаимодействуют со специальным механизмом хранения данных, входящим в архитектуру платформы организации параллельных вычислений (рис. 1). Этот механизм реализует интерфейс доступа к разделяемой памяти, в которой находятся данные и результаты операций.



Рис. 1. Взаимодействие вычислительной гранулы с хранилищем данных

Вычислительные гранулы объединяются в специализированные библиотеки, которые динамически подключаются при выполнении параллельного приложения. Гранула загружается из библиотеки при необходимости ее выполнения и идентифицируется по имени операции. Реализация различных классов алгоритмов обработки информации в виде библиотек вычислительных гранул позволяет расширить сферу применения системы организации параллельных вычислений и дает возможность создавать новые классы приложений.

Распараллеливание обработки изображений. Для решения задачи распараллеливания вычислений в системе используется понятие мультиагентных систем. Архитектура вычислительной системы представляет собой множество агентов, объединенных общими средствами обмена данными. Каждый агент является равноправным участником процесса обработки и способен выполнять любую операцию из множества отдельных операций обработки, применяемых для решения задач различных предметных областей. Пример схемы графа сценария мультиагентной системы для решения задачи параллельной обработки аэрокосмических снимков приведен на рис. 2. Исходные данные включают совокупность кадров и параметры обработки. В результате обработки формируются карты специальных областей, на которых отмечены здоровые и пораженные участки растительности, а также оценка погрешности выполнения классификации участков растительности.



Рис. 2. Схема графа сценария системы

Подсистема распараллеливания обработки данных состоит из трех основных блоков.

Блок работы с изображениями содержит элементы, работающие непосредственно с файловой системой либо ГИС, а также модули, которые позволяют читать, сохранять растровые изображения, поддерживая, если необходимо, целостность геоданных.

Блок управления выполняет функции вычисления координат для разделения (каждый растр разделяется на части, каждая из которых обрабатывается отдельным вычислительным модулем) и последующей сборки растров, а также собственно разделение и сборку.

Вычислительный кластер осуществляет параллельную обработку частей исходного изображения и содержит несколько экземпляров вычислительных модулей, осуществляющих собственно обработку данных. В их качестве выступает набор функций, предназначенных для обработки изображений, позволяющих решить поставленную задачу (например, набор функций для сегментации изображения).

В качестве базы для параллельной обработки выбрана технология MPI из-за возможности реализации простого взаимодействия между вычислительными модулями и синхронизации.

Для передачи данных между вычислительными и управляющим модулями используется TCP-соединение, которое позволяет выполнять передачу данных по сети без потерь. Взаимодействие реализуется единообразно для каждого вычислительного модуля. Однако подобное взаимодействие может негативным образом сказываться на быстродействии системы. Чтобы снизить затраты на взаимодействие, обмен данными управляющего модуля с соответствующими вычислительными модулями реализован в виде отдельного потока выполнения [9–11].

Обработка изображения проходит пошаговым образом.

1. Исходное изображение разделяется на части, каждая из которых содержит в себе участок исходного мультиспектрального изображения. Пример простого разделения изображения приведен на рис. 3. Некоторые алгоритмы обработки изо-

бражений при работе используют не только данные пикселя, но и данные о его окрестности. В таком случае необходимо, чтобы окрестность пикселей также была в наличии, как это показано на рис. 4.

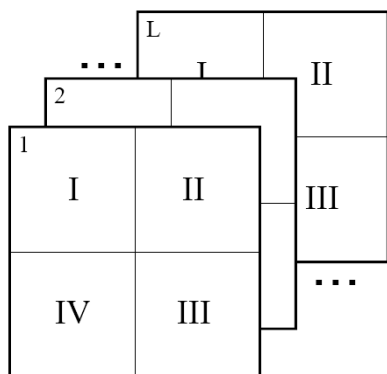


Рис. 3. Простое разделение

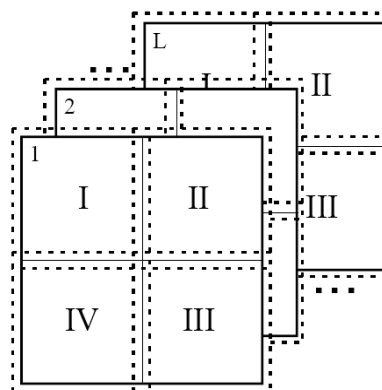


Рис. 4. Разделение с перекрытием

2. После разделения выполняется обработка, при которой части исходного мультиспектрального изображения обрабатываются параллельно.

3. Результаты обработки частей изображения собираются вместе, формируя единый растр.

Организация вычислений на базе многоагентной архитектуры. Общая архитектура вычислительной платформы показана на рис. 5. Интерфейс пользователя (блок *A*) представляет собой средство визуального проектирования и анализа параллельной программы обработки потока данных. Это специализированный графический редактор, использующий графовую модель представления программы. Редактор позволяет создать логический граф программы, определить соответствие вершин графа операциям из библиотек вычислительных гранул (блок *B*), задать параметры вызова гранул для различных типов обрабатываемых данных.

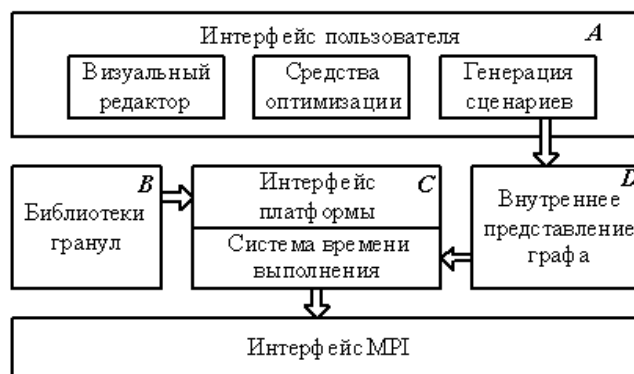


Рис. 5. Архитектура платформы организации параллельных вычислений

Средства оптимизации состоят из алгоритма оптимизации расписания на основе виртуальных ассоциативных сетей и имитационной модели. Имитационная модель используется для оценки вариантов расписания в процессе оптимизации, а также для визуализации расписания. Расписание представляется в виде диаграмм Ганта. Алгоритм оптимизации реализует технологию эволюционной оптимизации и относится к

гибридным генетическим алгоритмам, которые имеют преимущества перед эвристическими алгоритмами назначения задач на вычислительные узлы и алгоритмами статической оптимизации вычислений, принимая во внимание специфику обрабатываемого потока (неоднородность данных и задержек при их передаче, включая нулевые, изменение состава кластера в процессе обработки и др.) [9, 12, 13].

Созданная в визуальном редакторе программа преобразуется в текстовое представление на языке XML. Это представление содержит информацию для системы времени выполнения, а также отображение графа на топологию вычислительной системы. Файл XML при выполнении программы транслируется во внутреннее представление (блок *D*), позволяющее каждому элементу системы времени выполнения (блок *C*) получать необходимую информацию во время работы.

Система времени выполнения является многоагентной системой, использующей средства MPI для осуществления координации взаимодействия элементов. Она содержит два типа программных агентов: координатор и исполнитель. Количество агентов-исполнителей при выполнении параллельного приложения равно количеству физических процессоров вычислительной системы плюс один агент-координатор. Взаимодействие между агентами осуществляется с помощью передачи сообщений, реализующих простой протокол взаимодействия (рис. 6).

Координатор является главным процессом и управляет выполнением алгоритма, используя очередь дескрипторов обрабатываемых объектов, готовых к выполнению. Основная задача координатора состоит в передаче дескрипторов операций свободным исполнителям в соответствии с отображением операций графа на топологию системы. Кроме того, координатор следит за завершением операций и помещает в очередь новые дескрипторы в соответствии с топологией графа приложения.

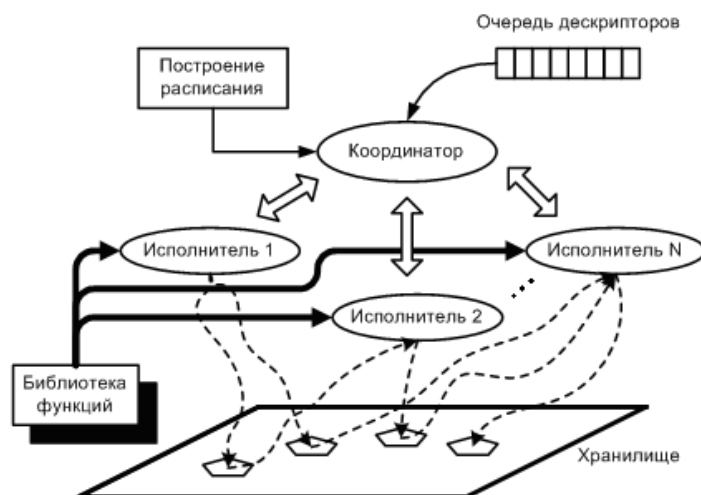


Рис. 6. Архитектура системы времени выполнения

Агенты-исполнители являются абстракцией физических процессоров системы. Их основная задача – выполнение вычислительных гранул, которые передаются координатором. Исполнители взаимодействуют с библиотекой гранул, загружая требуемые гранулы и выполняя их на подчиненном процессоре. Взаимодействие между агентами происходит через механизмы разделяемой памяти. После завершения операции исполнитель посылает сообщение координатору, который отмечает момент готовности результатов операции для дальнейшего использования. Синхронизация через координатор гарантирует детерминированное выполнение алгоритма.

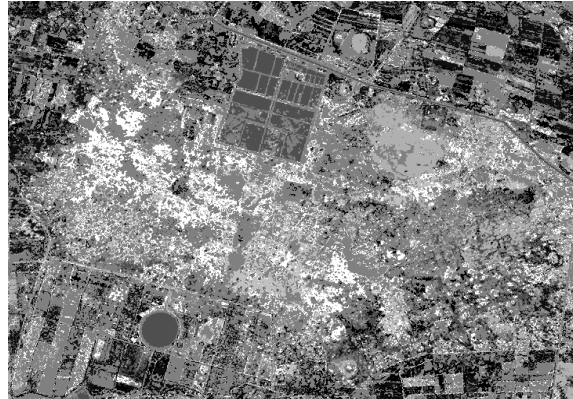


Рис. 7. Результат сегментации с помощью FCM на 15 кластеров без фильтрации

Тестирование системы обработки изображений производилось на суперкомпьютере СКИФ К-1000 [14] на изображениях, полученных с помощью аппаратуры Landsat 7 ETM+ (использовались каналы 3,4,7,8). На рис. 7 представлены результаты сегментации достаточно сложного участка, характерного для Белорусского Полесья. Объект представляет собой окрестности болота Званец в Брестской области, где присутствуют различные типы покрытий, включая болота, открытую воду, заливные луга, пахотные угодья, хвойный, лиственный и смешанный лес различных пород.

В экспериментах было задействовано от 1 до 64 вычислительных узлов, в качестве исходных данных были использованы цветные изображения размером $2\,000 \times 2\,000$ и $1\,000 \times 1\,000$ пикселей. Из графиков (рис. 8) видно, что зависимость времени работы от количества задействованных узлов пропорциональна $1/x$. Так, при обработке изображения размером $2\,000 \times 2\,000$ пикселей время работы модулей изменяется от 159 с для одного узла до 6 с на 64 узлах, а для изображения размером $1\,000 \times 1\,000$ пикселей – с 34 с на одном вычислительном узле до 1 с на 64 узлах.

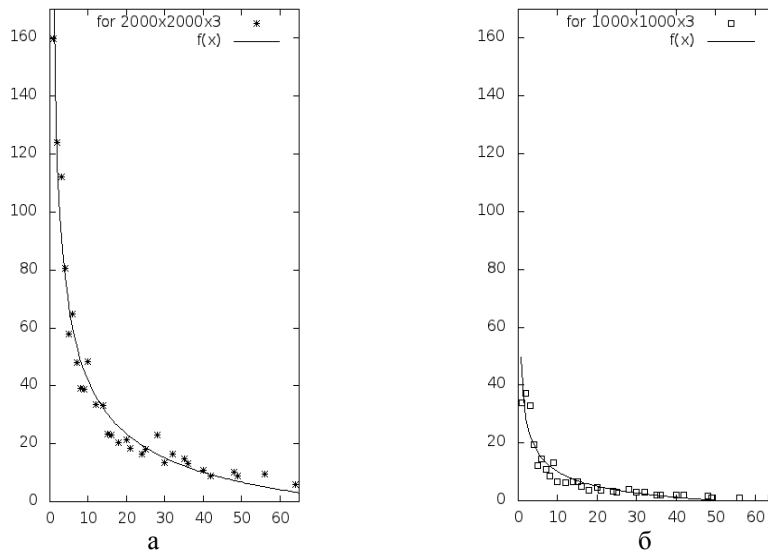


Рис. 8. График зависимости времени обработки от количества задействованных вычислительных узлов для цветных изображений размером пикселей:
 а – $2\,000 \times 2\,000$; б – $1\,000 \times 1\,000$

Заключение. Разработанная подсистема распараллеливания обработки изображений на основе интерфейса передачи сообщений MPI позволяет осуществлять прозрачное для программиста взаимодействие вычислительных узлов кластера и обмен данными между ними. Она предназначена для сокращения времени обработки данных ДЗЗ и построена на основе параллельной обработки данных на вычислительных кластерах суперкомпьютеров. Практическая значимость данной разработки заключается в возможности повышения производительности вычислений в задачах обработки данных при выполнении совместной тематической обработки данных ДЗЗ и данных наземных наблюдений, ориентированных на решение задач устойчивого лесопользования, анализ и отображение результатов тематической обработки в виде тематических карт.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Беляев Б.И., Катковский Л.В.* Оптическое дистанционное зондирование. – Минск: БГУ, 2006. – 455 с.
2. Мониторинг биологического разнообразия лесов России: методология и методы / Под ред. А.С. Исаева. ЦЭПЛРАН. – М.: Наука, 2008. – 453 с.
3. *Григорьева О.В.* Наблюдение деградации лесов по данным гиперспектрального аэро- и космического зондирования // Исследование Земли из космоса. – 2013. – № 1. – С. 43-47.
4. *Марков А.В. и др.* Автоматизированные методы оценки состояния окружающей среды по данным мульти- и гиперспектральной съемки // Геоматика. – 2012. – № 4. – С. 102-106.
5. *Argiro D. et al.* Khoros: An integrated development environment for scientific computing and visualization // Khoros Research, Inc. – Режим доступа: <https://www.cs.purdue.edu/homes/enh/PSEbook/PARTII/papersii/khoros.pdf> (дата обращения: 17.10.2014).
6. *Wickel J. et al.* Axiom - a modular visual object retrieval system / M. Jarke, J. Koehler, G. Lakemeyer, editors, KI 2002: Advances in Artificial Intelligence. – LNAI 2479. – Springer, 2002. – P. 253-267.
7. *Dean J., Ghemawat S.* MapReduce: Simplified Data Processing on Large Clusters. – Режим доступа: <http://research.google.com/archive/mapreduce.html> (дата обращения: 17.10.2014).
8. MPI: A Message-Passing Interface Standard Version 3.0. Message Passing Interface Forum, September 21, 2012. – Режим доступа: <http://www.mpi-forum.org/docs/mpi-3.0/mpi30-report.pdf> (дата обращения: 17.10.2014).
9. *Otwagin A.* Multiagent System for Reliable and Efficient Parallel Computing // Proc. of the 5th International Conf. on Neural Networks and Artificial Intelligence (ICNNAI'2008) (Minsk, 27–30 May, 2008). – Minsk, 2008. – P. 46-50.
10. *Sadykhov R.Kh., Ganchenko V.V., Podenok L.P.* Fuzzy clustering methods in multispectral satellite image segmentation // International Journal of Computing. – 2009. – Vol. 8 (1). – P. 87-94.
11. *Ганченко В.В. и др.* Обработка данных дистанционного зондирования земли для задач землепользования и мониторинга земельных ресурсов // Системы наблюдения, мониторинга и дистанционного зондирования Земли (Системы ДЗЗ'2010): тр. VII науч.-техн. конф. (Сочи, Адлерский р-н, пос. Веселое, 13–17 сент. 2010 г.). – М., 2010. – С. 181-186.
12. *Beaumont O., Legrand A., Robert Y.* Static scheduling strategies for heterogeneous systems // Computing and Informatics. – 2002. – № 21. – P. 413-430.
13. *Nagras T., Janeczek J.A.* Fast Compile-Time Task Scheduling Heuristic for Homogeneous Computing Environments // International Journal of Computers and Their Applications. – 2005. – № 12 (2). – P. 76-82.
14. *Анищенко В.В. и др.* Принципы создания базовых конфигураций суперкомпьютерных систем отраслевого назначения // Информатика. – 2012. – № 1 (33). – С. 97-105.
15. *Hwang K., Xu Z.* Scalable Parallel Computing Technology, Architecture, Programming. – USA: McGraw-Hill, 1998. – 832 p.
16. *Ганченко В.В., Дудкин А.А.* Решение задач лесопользования и мониторинга земельных ресурсов на основе обработки данных дистанционного зондирования Земли // Тезисы докл. Второй Междунар. науч.-техн. конф. «Актуальные проблемы создания космических систем дистанционного зондирования Земли» (Москва, 16 мая 2014 г.). – М.: ОАО «Корпорация «ВНИИЭМ»», 2014. – С. 11-13.

REFERENCES

1. *Belyaev B.I., Katkovskiy L.V.* Opticheskoe distantsionnoe zondirovanie [Optical remote sensing]. Minsk: BGU, 2006, 455 p.
2. Monitoring biologicheskogo raznoobraziya lesov Rossii: metodologiya i metody [Monitoring of the biological diversity of forests in Russia: methodology and techniques], Ed. A.S. Isaeva. CEPF RAS. Moscow: Nauka, 2008, 453 p.
3. *Grigor'eva O.V.* Nablyudenie degradatsii lesov po dannym giperspektral'nogo aero- i kosmicheskogo zondirovaniya [Observation of forest degradation according hyperspectral aerial and satellite sensing], *Issledovanie Zemli iz kosmosa* [Study of Earth from space], 2013, No. 1, pp. 43-47.
4. *Markov A.V. i dr.* Avtomatizirovannye metody otsenki sostoyaniya okruzhayushchey sredy po dannym multi- i giperspektral'noy s'emki [Automated methods for environmental assessment according to multi- and hyperspectral survey], *Geomatika* [Geomatics], 2012, No. 4, pp. 102-106.
5. *Argiro D. et al.* Khoros: An integrated development environment for scientific computing and visualization, *Khoros Research, Inc.* Available at: <https://www.cs.purdue.edu/homes/enh/PSEbook/PARTII/papersii/khoros.pdf> (Accessed 17 October 2014).
6. *Wickel J. et al.* Axiom - a modular visual object retrieval system / M. Jarke, J. Koehler, G. Lakemeyer, editors, *KI 2002: Advances in Artificial Intelligence. LNAI 2479.* Springer, 2002, pp. 253-267.
7. *Dean J., Ghemawat S.* MapReduce: Simplified Data Processing on Large Clusters. Available at: <http://research.google.com/archive/mapreduce.html> (Accessed 17 October 2014).
8. MPI: A Message-Passing Interface Standard Version 3.0. Message Passing Interface Forum, September 21, 2012. Available at: <http://www.mpi-forum.org/docs/mpi-3.0/mpi30-report.pdf> (Accessed 17 October 2014).
9. *Orwagin A.* Multiagent System for Reliable and Efficient Parallel Computing // Proc. of the 5th International Conf. on Neural Networks and Artificial Intelligence (ICNNAI'2008) (Minsk, 27–30 May, 2008). Minsk, 2008, pp. 46-50.
10. *Sadykhov R.Kh., Ganchenko V.V., Podenok L.P.* Fuzzy clustering methods in multispectral satellite image segmentation, *International Journal of Computing*, 2009, Vol. 8 (1), pp. 87-94.
11. *Ganchenko V.V. i dr.* Obrabotka dannykh distantsionnogo zondirovaniya zemli dlya zadach zemlepol'zovaniya i monitoringa zemel'nykh resursov [Processing of Remote Sensing Data for the problems of land use and land resources monitoring], *Sistemy nablyudeniya, monitoringa i distantsionnogo zondirovaniya Zemli (Sistemy D33'2010): tr. VII nauch.-tekhn. konf. (Sochi, Adlerskiy r-n, pos. Veseloe, 13–17 sent. 2010 g.)* [Observation systems, monitoring, and remote sensing (Systems of RSD'2010): proc. of the VII scientific-tech. conf. (Sochi, Veseloe, 13-17 September, 2010)]. Moscow, 2010, pp. 181-186.
12. *Baumont O., Legrand A., Robert Y.* Static scheduling strategies for heterogeneous systems, *Computing and Informatics*, 2002, No. 21, pp. 413-430.
13. *Hagras T., Janecek J.A.* Fast Compile-Time Task Scheduling Heuristic for Homogeneous Computing Environments, *International Journal of Computers and Their Applications*, 2005, No. 12 (2), pp. 76-82.
14. *Anishchenko V.V. i dr.* Printsipy sozdaniya bazovykh konfiguratsiy superkomp'yuternykh sistem otraslevogo naznacheniya [Principles for creating basic configurations supercomputer systems branch for destination], *Informatika* [Informatics], 2012, No. 1 (33), pp. 97-105.
15. *Hwang K., Xu Z.* Scalable Parallel Computing Technology, Architecture, Programming. USA: McGraw-Hill, 1998, 832 p.
16. *Ganchenko V.V., Dudkin A.A.* Reshenie zadach lesopol'zovaniya i monitoringa zemel'nykh resursov na osnove obrabotki dannykh distantsionnogo zondirovaniya Zemli [Problem solving forest management and monitoring of land resources based on the processing of remote sensing data], *Tezisy dokl. Vtoroy Mezhdunar. nauch.-tekhn. konf. «Aktual'nye problemy sozdaniya kosmicheskikh sistem distantsionnogo zondirovaniya Zemli» (Moskva, 16 maya 2014 g.)* [Abstracts of the Second International Scientific Conference «Actual problems of creation of space systems for remote sensing of the Earth» (Moscow, 16 May 2014)]. Moscow: OAO «Korporatsiya «VNIIEМ», 2014, pp. 11-13.

Статью рекомендовал к опубликованию д.т.н., профессор М.М. Татур.

Дудкин Александр Арсентьевич – Государственное научное учреждение «Объединенный институт проблем информатики Национальной академии наук Беларуси»; e-mail: doudkin@newman.bas-net.by; 220012, Беларусь, г. Минск, ул. Сурганова, 6; тел.: 375172842148, факс: 375172842175; д.т.н.; с.н.с.; доцент; заместитель заведующего лабораторией идентификации систем.

Doudkin Alexandr Arsent'evich – The State Scientific Institution «The United Institute of Informatics Problems of the National Academy of Sciences of Belarus»; e-mail: doudkin@newman.bas-net.by; 6, Surganova street, Minsk, 220012, Belarus; phone: 375172842148, fax: 375172842175; dr. of eng. sc.; senior researcher; associate professor; deputy head of the laboratory of system identification.

УДК 004.382.2

О.В. Ершова, А.В. Чкан, Е.В. Кириченко, Е.А. Семерников

ОШИБКИ УСЕЧЕНИЯ РЕЗУЛЬТАТОВ АРИФМЕТИЧЕСКИХ ОПЕРАЦИЙ С ФИКСИРОВАННОЙ ТОЧКОЙ В АЛГОРИТМАХ БЫСТРОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Рассматриваются ошибки вычисления быстрого преобразования Фурье (БПФ) с прореживанием по частоте в отсутствие масштабирования, обусловленные ограничением разрядности результатов арифметических операций умножения при обработке данных с фиксированной точкой в зависимости от размера преобразования, номера спектральной составляющей, а также от способа приведения результата к системной длине слова (усечение или округление данных). Получены аналитические выражения среднеквадратического значения ошибок вычисления БПФ белого гауссовского шума как при усечении, так и при округлении результатов операций умножения в алгоритме БПФ с прореживанием по частоте в отсутствие масштабирования. Показано, что в области низких частот для приближенной оценки уровня шумов, вызванных усечением разрядности результатов арифметических операций, можно использовать зависимость модуля математического ожидания ошибок усечения от номера спектрального отсчета. Поставлен математический эксперимент по вычислению погрешности арифметических операций БПФ для входной последовательности в виде белого гауссовского шума. Приведены графики среднеквадратических значений ошибок усечения и округления в зависимости от номера спектральной составляющей. Показано, что зависимость среднеквадратического значения ошибок усечения арифметических операций от номера спектральной составляющей носит асимметричный характер и приводит к значительным искажениям спектра в области низких частот. Моделирование показало совпадение уровней ошибок, полученных экспериментальным и аналитическим путём.

Эффект накопления ошибок; быстрое преобразование Фурье; анализ погрешностей; численное моделирование; цифровая обработка сигналов.

O.V. Yershova, A.V. Chkan, E.V. Kirichenko, E.A. Semernikov

TRUNCATION ERRORS OF THE RESULTS OF FIXED-POINT ARITHMETIC OPERATIONS IN THE FFT ALGORITHMS

The errors of decimation-in-frequency fast Fourier transform (DIF FFT) calculation (without scaling) caused by limitations of the length fixed-point results of arithmetic operations depending on the FFT size, the number of the spectrum component and the method of bringing the result to the system word length (truncation or rounding) are considered. Analytic expressions of root-mean-square (RMS) errors of spectrum calculation both for truncation and rounding of the results of multiplication in the DIF FFT algorithm (without scaling) depending on the FFT size and the number of the spectrum component were obtained. It is shown that in the range of low frequency it