

В. М. Хачумов,  
В. П. Фраленко, Chen Guo Xiang, Zhang Guo Liang

## Перспективы построения высокопроизводительной системы обработки данных дистанционного зондирования Земли

Аннотация. В работе представлены промежуточные результаты исследований, посвященных решению ряда задач обработки полноцветных и мультиспектральных изображений дистанционного зондирования Земли. Приведены результаты экспериментов по обнаружению целевых регионов интереса с помощью метода спектрографической «закраски» и нейросетевой кластеризации; решена задача обнаружения сложных ригидных объектов. Предложена реализация алгоритма преобразования мультиспектральных снимков в полутоновые и цветные. Все программы разработаны для функционирования на гетерогенной вычислительной установке.

*Ключевые слова и фразы:* сотрудничество, дистанционное зондирование, мультиспектральное изображение, параллелизм.

### Введение

Результаты дистанционного зондирования Земли (ДЗЗ) применяются для лучшего распоряжения природными ресурсами, мониторинга чрезвычайных ситуаций, используются в военном деле и т.д. Лидирующую позицию по числу запущенных национальных космических аппаратов (КА) наблюдения с 2007 года занимает Китай. Например, только с сентября 2013 года успешно запущено пятнадцать спутников серии «Яогань». Китай наращивает усилия по созданию многокомпонентной национальной системы ДЗЗ, покупает и продает соответствующие технологии. Лидерство по годовым темпам запусков космических аппаратов ДЗЗ пытаются оспорить лишь США и

- 
- © В. М. Хачумов<sup>(1)</sup>, В. П. Фраленко<sup>(2)</sup>, Chen Guo Xiang<sup>(3)</sup>, Zhang Guo Liang<sup>(4)</sup> 2015
  - © Институт системного анализа РАН<sup>(1)</sup> 2015
  - © Институт программных систем имени А. К. Айламазяна РАН<sup>(2)</sup> 2015
  - © ООО Ханчжоуская компания электронных технологий «AIVISI»<sup>(3)</sup> 2015
  - © Харбинский политехнический университет<sup>(4)</sup> 2015
  - © Программные системы: теория и приложения, 2015

Германия. В России же спутниковая группировка ДЗЗ создается практически заново [1].

Огромная территория Китая, стабильный экономический рост, значительные успехи в освоении космических технологий — все это привело к тому, что сотрудничать с КНР стремятся многие ведущие мировые компании, в том числе российские [2]. Идеи стратегического партнерства России и Китая поддержаны на самом высоком уровне. В принятом 20 мая 2014 года Совместном заявлении Российской Федерации и Китайской Народной Республики о новом этапе отношений всеобъемлющего партнерства и стратегического взаимодействия говорится о необходимости кооперации по программам сотрудничества в сфере фундаментальных космических исследований, дистанционного зондирования Земли, спутниковой навигации, изучения дальнего космоса и пилотируемой космонавтики [3].

Данные ДЗЗ подлежат компьютерной обработке, которая характеризуется необходимостью выполнения сложных вычислительных операций над массивами данных с целью решения задач кластеризации и классификации. Целью проводимого совместного исследования институтов Российской академии наук с компанией “AIVISI” (Китай) является разработка программного обеспечения, предназначенного в том числе для решения задач поиска и классификации объектов на снимках ДЗЗ и спектрографической «закраски» полноцветных и мультиспектральных изображений. Имеющаяся в ИСА РАН гетерогенная вычислительная установка с распределенной памятью и удаленным доступом содержит компоненты, позволяющие в достаточной мере эффективно решать задачи высокопроизводительной потоковой обработки. Общая архитектура системы представлена на рис. 1 [4, 5]. Для организации вычислений используется программно-инструментальный комплекс «ПС НСКиД» [6]. Каждый вычислительный узел оснащен двумя графическими ускорителями Nvidia Tesla M2050, пиковая производительность ускорителя более одного триллиона операций в секунду. Сейчас в составе комплекса функционирует два вычислительных узла (node), однако, в ближайшее время их количество будет удвоено.

В настоящей работе рассмотрен ряд характерных для исследуемой области задач обработки полноцветных и мультиспектральных изображений ДЗЗ. В том числе следующие: «закраска» снимков, поиск и классификация целевых объектов, кластеризация и преобразование изображений.

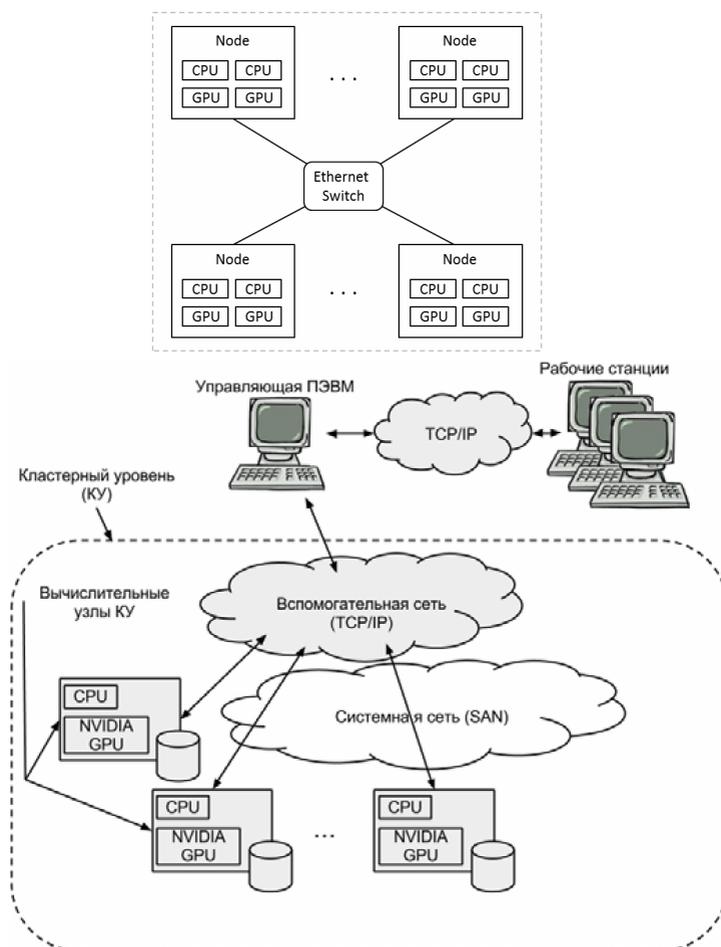


Рис. 1. Архитектура гетерогенной вычислительной системы

## 1. Мультиспектральные снимки

Одно из важнейших направлений исследований в области ДЗЗ связано с обработкой мультиспектральных снимков, которая позволяет выявить многие качественные характеристики как видимых на земной поверхности объектов, так и скрытых от глаз. Эти особенности находят применение в сельском и лесном хозяйстве, в геологии, при

поиске полезных ископаемых и т.д. Современные технологии ДЗЗ позволяют получать высококачественные изображения, которые дают возможность в реальном времени следить за целевыми регионами или отдельными объектами, получать точные трехмерные модели местности и т.п. Мультиспектральные снимки, передаваемые с действующих спутников (например, “Landsat-8”, “TERRA”, “SPOT 6”, “Pleiades-1B”), доступны в виде файлов в форматах “CEOS”, “DIMAP”, “HDF”, “TIFF”, “GeoTIFF” и др. Большую ценность представляют базы накопленных изображений. Так, например, снимки спутников серии “Landsat” являются общественным достоянием и могут быть получены с сайта [7] в рамках совместного проекта NASA, обеспечивающего непрерывную космическую летопись Земли.

Каждому космическому аппарату дистанционного зондирования присущ свой собственный набор сканирующих устройств, наиболее распространены сканеры “ASTER”, “TM” и “ETM+”, не все спектральные каналы того или иного сканера задействованы в конкретных конфигурациях оборудования. Например, снимки на спутнике “Terra” получают с помощью радиометра “ASTER”. Они используются для составления карт температуры поверхности Земли, отражательной способности, коэффициентов излучения, цифровых моделей рельефа. Ряд снимков, используемых в нашем исследовании, получен с помощью сканера “ASTER” и предоставлен сообществом специалистов в области геоинформационных систем и ДЗЗ “GIS-Lab” [8].

## 2. Задача «закраски» изображений ДЗЗ

Современные программные системы обработки данных ДЗЗ обладают широкими возможностями для проведения анализа входной информации. Одной из них является система “ENVI” [9], которая служит средой для обработки и отображения мультиспектральных снимков. Для решения задачи выделения целевых регионов («закраски») используется метрика Махаланобиса (Mahalanobis Distance) [10]. Эта метрика отличается от евклидоваго расстояния тем, что учитывает корреляции между компонентами векторов в заданных экспертом регионах интереса. В то же время она имеет недостаток, поскольку при малых разбросах параметров в выделенных классах-регионах возникают трудности с вычислением обратных матриц ковариаций классов. Этому недостатка лишены две другие метрики: полиномиальная метрика Махаланобиса (Polynomial Mahalanobis Distance) [11, 12] и

обобщенная метрика Евклида–Махаланобиса (Euclidean–Mahalanobis Distance) [13, 14], их использование представляет большой интерес для обработки космических снимков.

Обобщенным расстоянием Евклида–Махаланобиса между двумя точками  $x = (x_1, \dots, x_p)$  и  $y = (y_1, \dots, y_p)$  в пространстве  $R^p$  называется функция вида

$$d_{E-M}(x, y) = \sqrt{(x - y)^T (S + E)^{-1} (x - y)},$$

где  $S$  — матрица ковариаций, а  $E$  — единичная матрица. В работе [15] доказывается, что расстояние Евклида–Махаланобиса является метрикой. Расстояние учитывает распределение точек в заданном классе и применимо при малых значениях дисперсий информативных признаков. Оно устраняет недостаток метрики Махаланобиса, поскольку элементы главной диагонали используемых ковариационных матриц всегда больше нуля. Успешно применяется в задачах обработки изображений [16–18].

В качестве примера спектрографической «закраски» можно привести рис. 2, где соответствующими цветами выделены определенные экспертом регионы интереса. Из-за специфики задачи реализация алгоритма «закраски» демонстрирует близкую к линейной масштабируемость [4].

Для «закраски» снимков могут быть использованы техники кластеризации изображений. На рис. 3 представлен результат автоматической кластеризации мультиспектрального снимка нейронной сетью Кохонена (слева приведен один из спектров). Масштабируемость алгоритма зависит от количества выделенных кластеров. В случае, если их количество незначительно по сравнению с числом задействованных вычислительных ядер, параллельная версия алгоритма может работать менее эффективно.

Методы «закраски» позволяют выделить области, содержащие специфические текстуры, характерные, например, для более сложных ригидных объектов. Используя постобработку и алгоритмы классификации, например, на основе аппарата искусственных нейронных сетей (ИНС) глубокого обучения [19], можно выделять на поверхности Земли необходимые целевые объекты, например, автомобили определенной марки (см. рис. 4 с результатами проведенного эксперимента). Из-за того, что исходное изображение делится на непересекающиеся части, реализация алгоритма обладает линейным масштабированием.

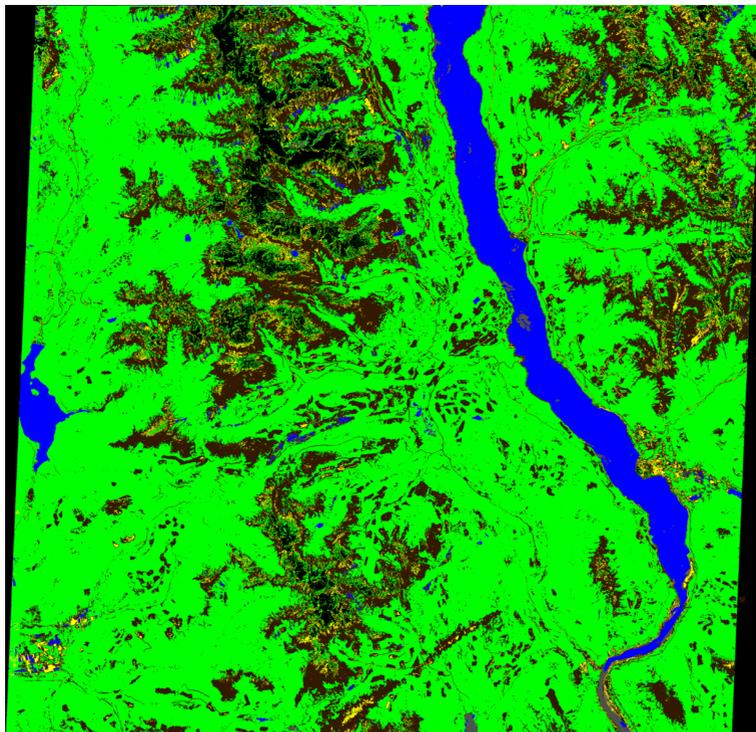


Рис. 2. Результат «закраски» мультиспектрального изображения

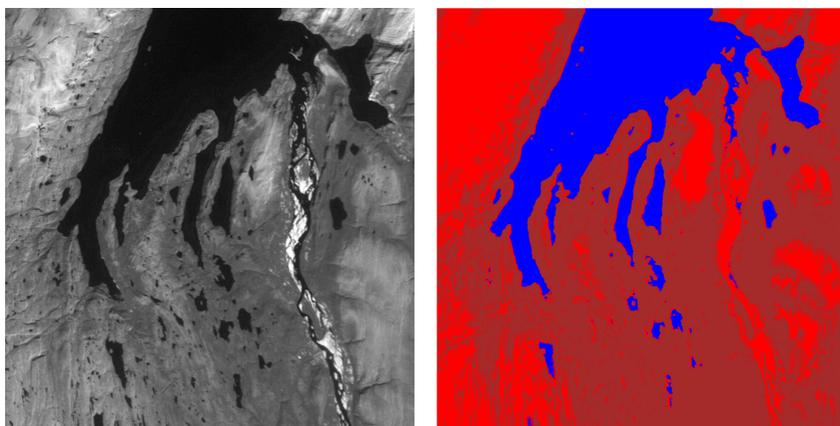


Рис. 3. Эксперимент с автоматической кластеризацией снимка

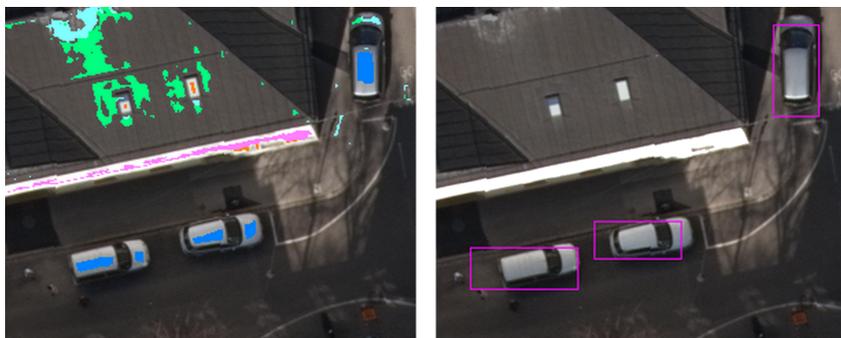


Рис. 4. Результаты «закраски» полноцветного изображения и последующей классификации выделенных зон

### 3. Задача создания полутоновых и цветных изображений из мультиспектральных снимков

Метод главных компонент (МГК) — один из эффективных методов снижения размерности пространства признаков без существенных потерь информации [20]. Использование МГК для преобразования мультиспектральных снимков в псевдоцветные или полутоновые позволяет более наглядно представить исходные данные.

Алгоритм создания полутоновых и псевдоцветных изображений:

- (1) приведение исходных каналов обрабатываемого изображения к единой глубине цвета (от 0 до 255) и к единому масштабу;
- (2) для каждого канала из значений всех его пикселей вычитается среднее значение яркости;
- (3) вычисление ковариационной матрицы;
- (4) вычисление собственных значений и собственных векторов ковариационной матрицы;
- (5) снижение размерности пространства признаков;
- (6) нормализация значений полученных матриц — приведение к диапазону  $[0, 255]$ ;
- (7) слияние результатов в единое изображение.

Примеры создания цветных изображений приведены на рис. 5.

Все этапы алгоритма, за исключением пятого, достаточно хорошо распараллеливаются на графическом ускорителе. Ковариационная матрица имеет небольшой размер, поэтому выгоднее проводить вы-

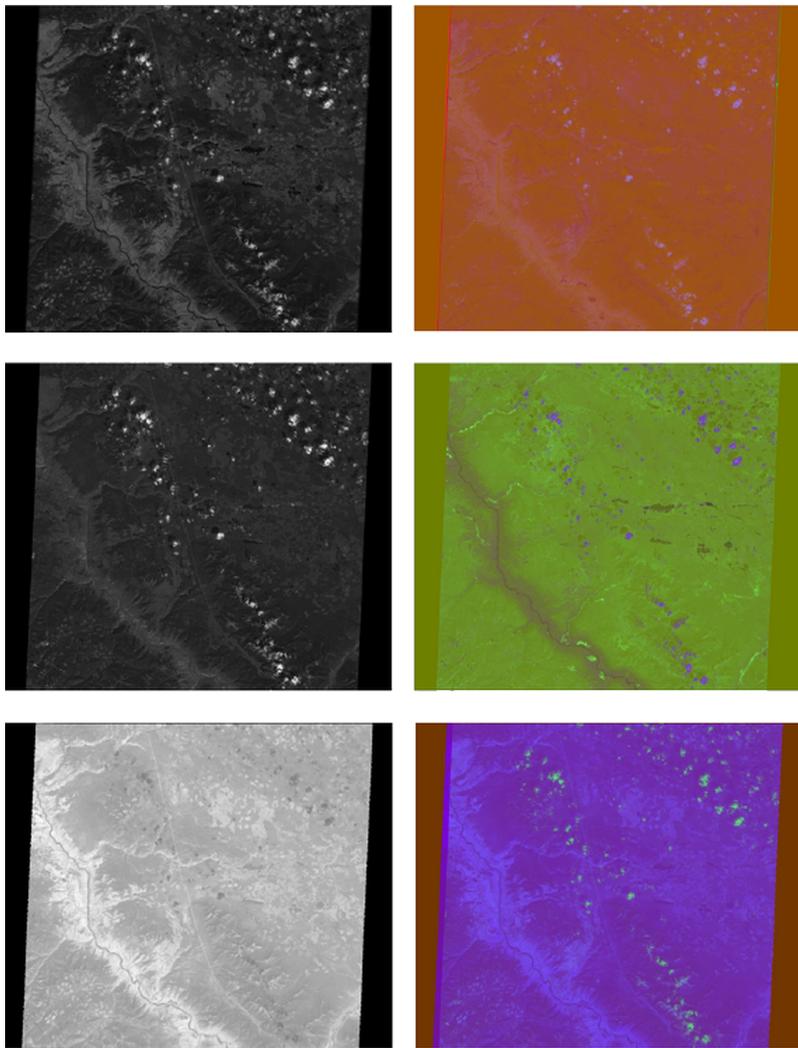


Рис. 5. Полутоновые и псевдоцветные изображения, созданные из мультиспектрального снимка с помощью МГК

числения на процессоре общего назначения. В таблице 1 приведены результаты сравнения производительности реализаций алгоритма.

Результаты экспериментов свидетельствуют о том, что использо-

Таблица 1. Основные текстурные признаки, рассчитанные по матрице зависимости

Каналы мультиспек- трального снимка	Количество каналов в новом изображении	Время обработки на процессоре общего назначения, сек.	Время обработки на графическом ускорителе, сек.
5–10	1	5.57	0.24
5–10	3	6.14	0.32
1–3	1	1.51	0.12
1–3	3	2.05	0.19
10–15	1	0.97	0.06
10–15	3	1.12	0.07

вание графических ускорителей для создания псевдоцветных и полутоновых снимков из мультиспектральных с помощью МГК позволяет добиться ускорения вычислений в 10–20 раз.

### Заключение

Выполненные исследования направлены на проведение совместных российско-китайских работ по созданию базового программного обеспечения высокопроизводительной обработки снимков ДЗЗ. Решением служит автоматическое или автоматизированное распараллеливание программ на основе универсальных и специализированных библиотек/программных модулей. Экспериментальные реализации отдельных алгоритмов («закраска», кластеризация, классификация) показали достаточную эффективность обработки данных ДЗЗ предлагаемыми методами и наличие масштабируемости параллельного счета.

Целесообразно продолжить совместную работу над развитием механизмов высокопроизводительной обработки сверхбольших полноцветных и мультиспектральных снимков на гетерогенных установках. Дальнейшие исследования предполагается сфокусировать на развитии методов параллельного программирования и улучшении технических характеристик программно-аппаратного комплекса.

## Список литературы

- [1] А. Кучейко. «Текущее состояние дистанционного зондирования Земли из космоса в России и за рубежом», *Газета.Ру*, 2013, URL [http://www.gazeta.ru/science/2013/02/27\\_a\\_4990217.shtml](http://www.gazeta.ru/science/2013/02/27_a_4990217.shtml) ↑ 122.
- [2] «Текущее состояние дистанционного зондирования Земли из космоса в России и за рубежом», *Ракурс*, URL <http://www.racurs.ru/?page=801> ↑ 122.
- [3] *Совместное заявление Российской Федерации и Китайской Народной Республики о новом этапе отношений всеобъемлющего партнерства и стратегического взаимодействия*, 2014, URL [http://news.kremlin.ru/ref\\_notes/1642](http://news.kremlin.ru/ref_notes/1642) ↑ 122.
- [4] А. Ю. Попков, И. В. Соченков, В. М. Хачумов. «Гетерогенный вычислительный комплекс для решения инженерных и научных задач», Третий Национальный Суперкомпьютерный Форум НСКФ-2014 (Переславль-Залесский, 2014), URL <https://docs.google.com/uc?export=download&id=0B-Qay3kEFxqfZTJLeEgYbVnKOWc> ↑ 122, 125.
- [5] A. A. Kondratyev, V. M. Khachumov. “High-performance stream processing on multicore and graphics processors”, Second International Conference “Cluster Computing” CC 2013 (Lviv, 2013), pp. 105–110 ↑ 122.
- [6] В. М. Хачумов, И. П. Тищенко, А. А. Талалаев, К. А. Константинов, В. П. Фраленко, Ю. Г. Емельянова. *Нейросетевая система контроля телеметрической информации, диагностики подсистем космических аппаратов, обработки космических снимков (ПС НСКиД)*, Свидетельство о государственной регистрации программы для ЭВМ № 2012613261, дата приоритета: 18.11.2011, дата регистрации: 06.04.2012 ↑ 122.
- [7] *Landsat Image Gallery*, URL <http://landsat.visibleearth.nasa.gov/> ↑ 124.
- [8] *GIS LAB. Геоданные. Новый источник доступных данных ASTER*, URL <http://gis-lab.info/projects/aster/about.html> ↑ 124.
- [9] О. Н. Колесникова. «Использование программного комплекса ENVI для обработки данных дистанционного зондирования Земли», *Геоматика*, 2009, №1, с. 38–41, URL [http://geomatika.ru/pdf/2009\\_01/2009\\_01\\_005.pdf](http://geomatika.ru/pdf/2009_01/2009_01_005.pdf) ↑ 124.
- [10] S. V. Aelst. “Robust Estimation, Inference and Prediction”, COMPSTAT 2014 (University of Leuven, 2014), URL <http://compstat2014.org/Stefan%20Van%20Aelst%20-%20TutorialCompstat2014.pdf> ↑ 124.
- [11] G. Grudic, J. Mulligan, “Outdoor path labeling using polynomial Mahalanobis distance”, *Proc. of Robotics: Science and Systems*, 2006, pp. 16–19 (english) ↑ 124.
- [12] M. R. Ackerman. *Algorithms for the Bregman k-Median Problem*, A dissertation submitted to the Department of Computer Science University of Paderborn, 2009, 220 pp. ↑ 124.

- [13] С. А. Амелькин, В. М. Хачумов, «Обобщенное расстояние Евклида–Махаланобиса и его применение в задачах распознавания образов», *12-я Всероссийская конференция «Математические методы распознавания образов»*, МАКС Пресс, М., 2005, с. 7–9 ↑ 125.
- [14] С. А. Амелькин, А. В. Захаров, В. М. Хачумов. «Обобщенное расстояние Евклида–Махаланобиса и его свойства», *Информационные технологии и вычислительные системы*, 2006, №4, с. 40–44 ↑ 125.
- [15] M. V. Khachumov. “Distances, Metrics and Cluster Analysis”, *Scientific and Technical Information Processing*, **39:6** (2012), pp. 1–7 (english) ↑ 125.
- [16] В. П. Фраленко. «Анализ спектрографических текстур данных дистанционного зондирования Земли», *Искусственный интеллект и принятие решений*, 2010, №2, с. 11–15 ↑ 125.
- [17] А. И. Смирнова, В. М. Хачумов. «Метод обработки мультиспектральных снимков дистанционного зондирования Земли», *Авиакосмическое приборостроение*, 2013, №2, с. 50–56 ↑ 125.
- [18] В. П. Фраленко. «Методы текстурного анализа изображений, обработка данных дистанционного зондирования Земли», *Программные системы: теория и приложения*, **5:4(22)** (2014), с. 19–39, URL [http://psta.psir.ru/read/psta2014\\_4\\_19-39.pdf](http://psta.psir.ru/read/psta2014_4_19-39.pdf) ↑ 125.
- [19] *Convolutional Neural Networks (LeNet)*, DeepLearning 0.1 documentation (english), URL <http://deeplearning.net/tutorial/lenet.html> ↑ 125.
- [20] K. Pearson. “On lines and planes of closest fit to systems of points in space”, *Philosophical Magazine*, **2** (1901), pp. 559–572 (english) ↑ 127.

Рекомендовал к публикации

к.т.н. С. А. Амелькин

*Об авторах:*



**Вячеслав Михайлович Хачумов**

Д.т.н., профессор, заведующий лабораторией методов интеллектуального управления ИСА РАН, автор более 180 научных публикаций.

*e-mail:*

[vmh48@mail.ru](mailto:vmh48@mail.ru)



**Виталий Петрович Фраленко**

К.т.н., старший научный сотрудник ИПС им. А.К. Айламазяна РАН, автор более 65 публикаций.

*e-mail:*

[alarmod@pereslavl.ru](mailto:alarmod@pereslavl.ru)



**Chen Guo Xiang**

Генеральный директор ООО Ханчжоуская компания электронных технологий «AIVISI»

*e-mail:*

[chengx@ketuoda.com](mailto:chengx@ketuoda.com)



**Zhang Guo Liang**

Старший научный сотрудник Технологического института Харбинского политехнического университета

*e-mail:*

[0451\\_zgl@163.com](mailto:0451_zgl@163.com)

*Пример ссылки на эту публикацию:*

В. М. Хачумов, В. П. Фраленко, Chen Guo Xiang, Zhang Guo Liang. «Перспективы построения высокопроизводительной системы обработки данных дистанционного зондирования Земли», *Программные системы: теория и приложения*, 2015, **6**:1(24), с. 121–133.

URL

[http://psta.psiras.ru/read/psta2015\\_1\\_121-133.pdf](http://psta.psiras.ru/read/psta2015_1_121-133.pdf)

Vyacheslav Khachumov, Vitaly Fralenko, Chen Guo Xiang, Zhang Guo Liang. *Construction perspectives of the remote sensing data high-performance processing system.*

ABSTRACT. The paper presents the intermediate results of studies on the challenges of full color and multispectral remote sensing images processing. Presented the experiments results of target regions detection with spectrographic fill and neural network clustering; solved the problem of complex rigid objects detection. Proposed implementation of the multispectral images to grayscale and color algorithm conversion. All programs are designed for operation on a heterogeneous computing machine. (*In Russian*).

*Key Words and Phrases:* collaboration, remote sensing, multispectral images, parallelism.

*Sample citation of this publication*

Vyacheslav Khachumov, Vitaly Fralenko, Chen Guo Xiang, Zhang Guo Liang. "Construction perspectives of the remote sensing data high-performance processing system", *Program systems: theory and applications*, 2015, 6:1(24), pp. 121–133. (*In Russian*.) URL [http://psta.psiras.ru/read/psta2015\\_1\\_121-133.pdf](http://psta.psiras.ru/read/psta2015_1_121-133.pdf)