

Автоматизация обработки данных самолетного гиперспектрального зондирования

**Т.В. Кондранин¹, В.В. Козодеров², Е.В. Дмитриев³, В.Д. Егоров³,
В.В. Борзяк², А.А. Николенко¹**

¹ *Московский физико-технический институт (государственный университет)
141700, Московская область, г. Долгопрудный, Институтский переулок, 9
E-mail: kondr@kondr.rector.mipt.ru*

² *Московский государственный университет им. М.В.Ломоносова
119991, Москва, ГСП-1, Ленинские горы
E-mail: vkozod@mes.msu.ru*

³ *Институт вычислительной математики РАН
119333, Москва, ГСП-1, ул. Губкина, 8
E-mail: yegor@inm.ras.ru*

Рассмотрены особенности автоматизации вычислительных процедур обработки данных самолетного гиперспектрального зондирования с акцентом на восстановление характеристик биологической продуктивности лесной растительности. Показаны основы реализации предлагаемого подхода с описанием работы модуля решения данной задачи, пользовательского интерфейса и блока отображения конечных результатов.

Ключевые слова: гиперспектральные изображения, обработка данных, биологическая продуктивность лесной растительности.

Введение

Автоматизация процесса обработки данных аэрокосмического дистанционного зондирования (ДЗ) требует разработки как эффективных методов и алгоритмов решения задач распознавания образов и оценки параметров состояния наблюдаемых объектов, так соответствующих программ, обеспечивающих, с одной стороны, высокую достоверность расчетных данных, а с другой максимальную эффективность всех вычислительных процедур. В настоящее время основные приложения данных аэрокосмического дистанционного зондирования, которые являются мощным инструментом решения научных и прикладных задач в различных областях, связаны с технологиями получения и последующей обработки многоспектральных изображений объектов (обычно 6-7 каналов в видимой и ближней инфракрасной области электромагнитного спектра). Развитие авиакосмического приборостроения в последнее десятилетие обусловило все большее распространение при решении указанных задач наряду с мультиспектральными данными дистанционного зондирования (ДЗ) гиперспектральных технологий ДЗ. В отличие от мультиспектральной гиперспектральная аппаратура обеспечивает получение информации в существенно большем (> 100) количестве узких (от единиц до 10 нм) спектральных диапазонов, что позволяет для каждого пикселя изображения получать детальную зависимость спектральной яркости наблюдаемого объекта от длины волны электромагнитного излучения. Гиперспектральные данные содержат сведения не только об объекте в целом, но также (по ха-

рактерным спектральным линиям) и о веществах, материалах, примесях и т.п., входящих в состав объекта, что обуславливает новые информационные возможности соответствующих систем ДЗ. Имеющийся на сегодня мировой опыт использования гиперспектральной съемки преимущественно основывается на данных, получаемых с авиационных и космических систем NASA: спутника EO-1 (аппаратура Hyperion), и авиационных носителей (аппаратура AVIRIS, CASI).

22 июля 2012 года был запущен на орбиту малый космический аппарат МКА-ФКИ ПН1 (Зонд ПП), на котором в составе дополнительной полезной нагрузки установлен первый российский космический гиперспектрометр разработки базового предприятия кафедры «Системы, устройства и методы геокосмической физики» МФТИ ЗАО НПО «Лептон» (149 спектральных каналов в диапазоне 400 – 1000 нм, пространственное разрешение 52 м при высоте съемки 820 км; полоса захвата при высоте 820 км – 42 км). Прототип этого прибора отработывался в рамках летных экспериментов с использованием средств малой авиации по различным тестовым полигонам в 2007-2011 годах.

Существующие лицензионные программные комплексы обработки многоспектральных (*применимость этих средств при обработке гиперспектральных изображений требует особого подхода в каждом конкретном случае*) данных ДЗ, поставляемых в основном зарубежными фирмами (ENVI, ERDAS Imagine), задействуются достаточно стандартные операции: цветокодирование изображений; классификация объектов; построение гистограмм распределения разных классов объектов по относительным градиентам исходных яркостей; подсчет площадей, занятых разными классами и т.п. Применяются технологии ГИС, которые позволяют использовать базы данных в различных предметных областях: предшествующие картографические материалы, дополнительные материалы, способствующие улучшению результатов «визуального дешифрирования» – традиционной формы интерпретации информационной продукции обработки данных для конкретных регионов и т.п. Эти технологии также эффективны с точки зрения обновления карт земельных ресурсов, уточнения отдельных нюансов конкретных территорий в плане обоснования потенциала промышленного и сельскохозяйственного производства, наблюдения особенностей дорожно-транспортной инфраструктуры регионов и др. В то же время, практическое использование этих процедур при решении все расширяющегося спектра прикладных задач, выявляет ограничение возможностей, предоставляемых стандартными пакетами: указанного перечня преобразований исходных изображений оказывается недостаточно. В частности возникает необходимость обоснования достоверности соответствующих результатов путем их сравнения с данными наземной наблюдательной сети, других имеющихся материалов по конкретной территории с анализом погрешностей задач распознавания объектов аэрокосмического мониторинга и оценки их состояния. Поэтому на первый план выдвигаются модели описания процессов взаимодействия солнечного излучения с природными средами и методы вычислительной математики при обращении основного функционала интенсивностей уходящего излучения, регистрируемого аппаратурой ДЗ. Требуется решение прямых задач формирования и трансформации излучения в системе «подстилающая поверхность – атмосфера – приемная аппаратура ДЗ», а также обратных задач восстановления параметров состояния наблюдаемых объектов по данным ДЗ. При этом используются результаты исследований в области физики атмосферы, почвенно-геологических и других исследований, современные методы вычислительной математики, компьютерных и геоинформационных технологий с целью разработки адекватных алгоритмов и их программной реализации.

Разработанные ранее вычислительные средства обработки многоспектральных аэрокосмических изображений (Козодеров, Кондранин, 2008) были адаптированы для обработки данных летных испытаний гиперспектральной аппаратуры (Кондранин и др., 2010). Были также показаны широкие возможности аэрокосмической гиперспектрометрии для распознавания природно-техногенных объектов и оценки параметров, характеризующих состояние почвенно-растительного покрова (Козодеров и др., 2010).

Соответствующие приложения обрабатывались в процессе летных экспериментов с использованием разных версий отечественной гиперспектральной камеры (ГСК) разработки ЗАО НПО «Лептон», г. Зеленоград (около 200 спектральных каналов, пространственное разрешение 1-2 м с высот 1.5-2 км) для выбранного тестового региона в 2007-2011 годах. Были проведены оценки информационного содержания данных летных испытаний гиперспектральной аппаратуры, устанавливаемой на одной платформе с аппаратурой аэрофотосъемки и аппаратурой глобального спутникового позиционирования (GPS). Была проведена также валидация (наземное подтверждение) результатов обработки данных гиперспектральной аэросъемки при их сравнении с данными наземных лесотаксационных и других обследований выбранной тестовой территории.

Основы реализации вычислительных процедур обработки данных гиперспектрального зондирования

Теоретической основой используемого метода восстановления объема зеленой фитомассы и других характеристик растительного покрова по данным дистанционных измерений является функционал яркости J_{ml} неоднородного по пространству растительного покрова, непосредственно измеряемый аппаратурой дистанционного зондирования.

Выделяют решение прямой задачи, когда рассчитываются значения функционала яркости J_{ml} для разных значений входящих в него параметров. Такие банки данных результатов расчетов функционала яркости для перечисленных выше типов аппаратуры ДЗ созданы в процессе реализации предлагаемой задачи. Эти данные могут быть представлены в виде электронной «книги», «страницы» которой – плоскости в координатах плотности (сомкнутости) полога лесной растительности C_{II} и ажурности крон деревьев A_{KP} . По каждой паре значений (C_{II} , A_{KP}) определяется значение величины фитомассы M . Остальные входящие в функционал параметры формируют номер страницы «книги». Число используемых каналов и угол визирования аппаратуры считаются заданными.

Для решения обратной задачи восстановления объема фитомассы и других параметров растительности по значениям яркостей для каждого пикселя многоканального изображения осуществляется поиск в многомерном фазовом пространстве («книга» из «страниц») точек, наиболее близких (в пределах допустимой погрешности) к набору яркостей по каналам для данного пикселя изображения. Фактически речь идет о многомерном фазовом пространстве параметров при заданной высоте Солнца в момент съемки. Если такие точки существуют – решение найдено, если в пределах набора заданных погрешностей таких точек в фазовом пространстве найти не удалось, значит для данного пикселя изображения решение считается не найденным (помечается особым образом). Таким образом для пикселей, где решение найдено, получаем всю совокупность значений параметров, от которых зависит функционал J_{ml} . В итоге по значениям C_{II} и A_{KP} для этого пикселя находятся значения объема фитомассы M , имеем также тип леса, тип межкрупной растительности и все другие, требуемые и заложенные в функционал характеристики.

В существующих программных реализациях этого метода осуществляется поиск решений перебором в указанном многомерном фазовом пространстве («книга» из «страниц»). При дискретном представлении (поскольку каждый параметр, такой как C_{IP} , A_{KP} и другие, заданы на некоторой сетке) это фазовое пространство составляет несколько миллионов точек. Последовательность точек, составляющая решение, ищется из условий удовлетворения некоторым критериям соответствующих значений яркостей по парам каналов. Затем выделяется общая часть этих решений по совокупности всех рассмотренных пар каналов. В итоге, например, для 6-ти канального изображения ЕТМ+ спутника Landsat-7, которое содержит $\sim 10000 \times 10000$ пикселей, обработка одного лишь фрагмента с восстановлением объема биомассы и других параметров занимает существовавшими ранее алгоритмами по времени около 30 часов работы 3 ГГц процессора. Обработка всего такого изображения, в котором содержится от 50 до 60 полностью или частично заполненных данными фрагментов 1000×1000 пикселей, по оценкам потребует от 1500 часов и более процессорного времени такого 3 ГГц процессора. Поэтому предполагается вести такого рода расчеты на больших многопроцессорных комплексах.

Нами разработан новый алгоритм решения рассматриваемой задачи, значительно повышающий вычислительную эффективность расчета объема фитомассы и других параметров растительности при сохранении точностных характеристик, свойственных исходному алгоритму. Для повышения вычислительной эффективности расчета предлагается использовать тот факт, что все современные данные дистанционного зондирования представлены в цифровом виде, причем, как правило, невысокой битности. Это связано, как представляется, с огромными и все возрастающими потоками передаваемых с аппаратуры дистанционного зондирования данных. Точность представления данных даже 8-ми битной структурой составляет десятые доли процента; должны быть, видимо, весомые аргументы в пользу того, чтобы в кратное число раз увеличивать потоки информации для представления этой информации более точно (16-ти битное представление имеет точность 0,002 % и удваивает поток информации).

Особенности автоматизации вычислительных процедур обработки данных

Ранее было показано, что методология построения программного обеспечения по восстановлению биопродукционных характеристик растительности при переходе к высокому пространственному и спектральному разрешению иная в сравнении с более низким разрешением. Этот новый подход может быть представлен в виде схемы, отображенной на рис. 1.

Для модуля восстановления биопродукционных характеристик по гиперспектральным данным был разработан удобный пользовательский интерфейс (рис. 2). Интерфейс включает в себя этапы описания данных ГСК, задание модельных данных, задание классификатора или указание имеющихся данных распознавания, определение рассматриваемых пород деревьев, этап коррекции модельных данных, этап определения/задания прозрачности атмосферы, задание параметров решения обратной задачи и отображения данных расчетов. Интерфейс контролирует диапазоны вводимых параметров и производит автоматическую компоновку, которая предоставляет пользователю возможность задавать только те параметры, которые влияют на конечный результат работы программы.



Рис. 1. Схема работы модуля восстановления биопродукционных характеристик древостоя по гиперспектральным изображениям

На этапе описания данных требуется указать путь к файлам с гиперспектральными изображениями в стандартных форматах BSQ (поканальный порядок) и BIP (по-кадровый порядок). Размеры, и формат записи файлов (поддерживаются single и double) представляются в информационном файле. Также необходимо указать файл с центральными длинами волн. Последнее необходимо, для обеспечения возможности привести модельные и измеряемые данные к единому спектральному разрешению.

Следующим шагом следует задание параметров модельных яркостей измеряемого отраженного солнечного излучения. При использовании стандартных данных необходимо указать зенитный и азимутальный углы Солнца. При использовании внешних данных указываются только углы визирования. Если углы визирования различаются для отдельных пикселей, пользователь имеет возможность загрузить эти углы в виде файла. Такой вариант необходим, если, например, прибор имеет достаточно широкий угол зрения.

При обработке гиперспектральных изображений желательно использовать внешние данные по классификации лесной растительности, поскольку применение стандартных подходов может приводить к существенным ошибкам из-за ограниченности используемой спектральной информации. Однако при отсутствии таких данных пользователь имеет возможность использовать классификацию на основе таблицы значений «вегетационных индексов» NDVI или трехканальный классификатор на основе нормального дискриминантного анализа.

В имеющейся базе данных имеются модельные спектры для древостоев разного породного состава. В случае, если по тем или иным причинам была нарушена калибровка прибора можно сделать тонкую подстройку.

При решении обратной задачи сначала производится поиск всех частных решений. Поиск точек соответствия модельных и измеряемых данных может быть проведен с использованием метода прямоугольников, евклидова расстояния (метод сфер) и метода k наи-

ближайших соседей. В первых двух случаях пользователь имеет возможность симметрично относительно центра масс масштабировать дискриминантную поверхность, а во втором – задавать количество соседей. Поскольку частных решений может быть найдено несколько, возникает задача поиска регуляризованного решения. Описываемый интерфейс предоставляет пользователю возможность выбрать среднее значение, медиану или применить метод максимального правдоподобия поля получения наиболее вероятного значения.

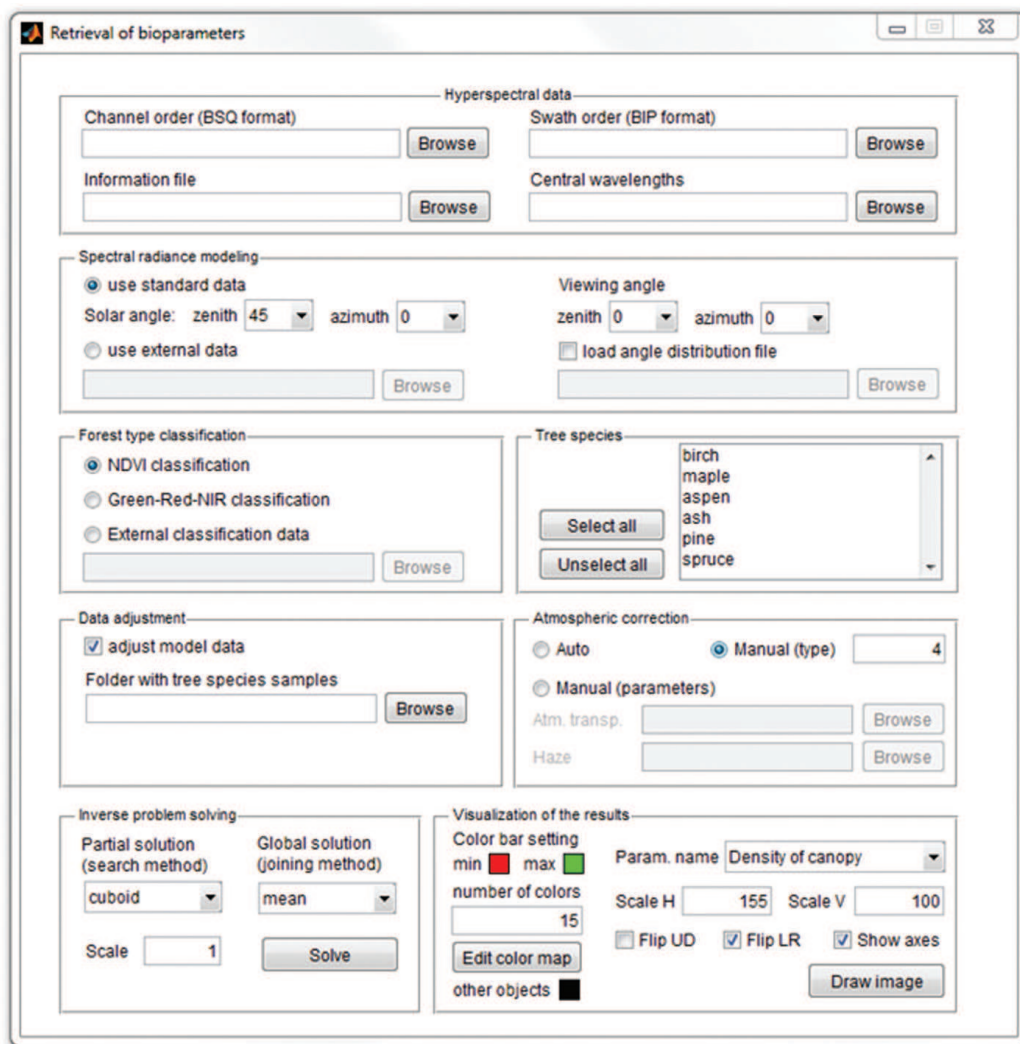


Рис. 2. Пользовательский интерфейс модуля восстановления биопродукционных характеристик лесного полога по гиперспектральным изображениям

После расчета проекционных характеристик лесного полога производится оценка связанных с ними биопродукционными характеристиками на основе статистических параметризаций. Полученные результаты могут быть отображены в виде цветных карт. Данный интерфейс позволяет выбрать необходимый параметр:

- плотность полога;
- плотность крон;
- тип межкрупной растительности,
- изрезанность верхней границы полога;
- биомасса листьев,

- площадь проективного покрытия,
- доля поглощенной фотосинтетически-активной радиации;
- чистая первичная продукция;
- породный состав древостоя;
- индекс нормализованных разностей

и задать цветовую карту, которая может быть в дальнейшем скорректирована в специальном редакторе. Это позволяет сделать карты биопродукционных параметров максимально контрастными и интуитивно понятными (рис. 3). Также пользователю предоставляется возможность изменять пропорции изображения, направление и отображение осей (координаты рисунка).

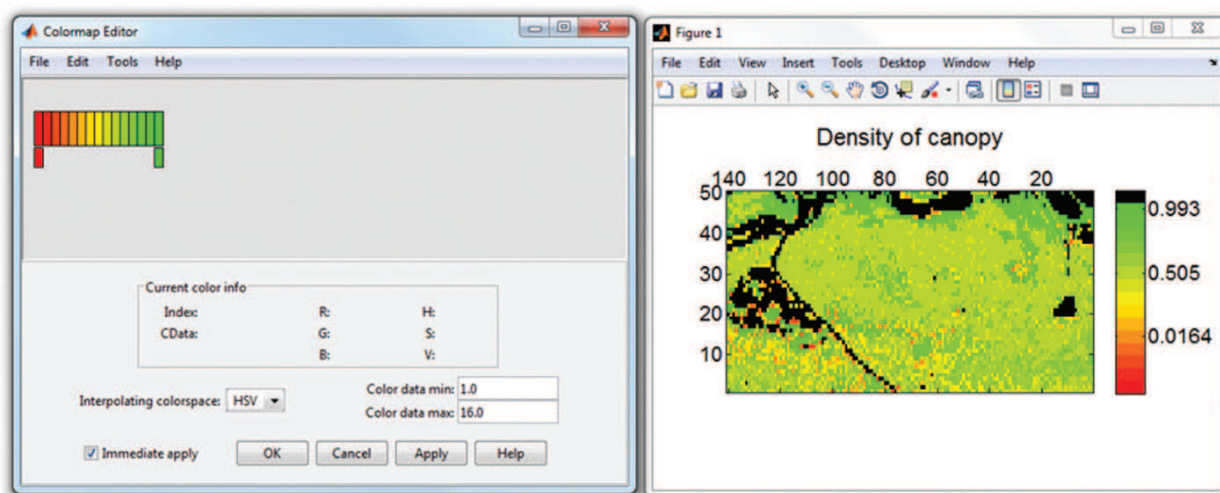


Рис. 3. Блок отображение результатов восстановления биопродукционных характеристик лесного полога по гиперспектральным изображениям

Разработка описанного выше интерфейса позволила существенно упростить работу пользователя с модулем восстановления биопродукционных характеристик по гиперспектральным изображениям. Работа с внутренними файлами и запуск процедур производится в автоматическом режиме. Обработчик ошибок позволит вводить только допустимые значения параметров и анализировать, насколько это возможно, корректность внешних файлов данных. Для дальнейшего совершенствования пользовательского интерфейса необходимо добавить контекстные подсказки для всех заполняемых полей.

Заключение

Автоматизация процесса обработки гиперспектральных изображений рассмотрена на примере восстановления характеристик биологической продуктивности лесной растительности с использованием данных, полученных гиперспектральной камерой видимого и ближнего инфракрасного диапазонов спектра, установленной на борту авиационного носителя. Данные ДЗ получены в ходе летных экспериментов, сопровождающихся наземными обследованиями в различные сезоны на тестовых полигонах северо-западных регионов Центральной России в период 2007-2011 годах.

Дано описание модуля решения рассматриваемой задачи и пользовательского интерфейса отображения полученных результатов. Перспективы развития соответствующих приложений – в увязке предлагаемых вычислительных процедур с распознаванием породного состава и возраста лесной растительности с построением оптимальных классификаторов.

Исследования проводятся при финансовой поддержке грантов РФФИ 11-07-00382, 11-07-12006_офи_м и ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы, г/к № ПЗ49 и № 14.740.11.1091.

Литература

1. Козодеров В.В., Кондранин Т.В. Методы оценки состояния почвенно-растительного покрова по данным оптических систем дистанционного аэрокосмического зондирования. М.: изд. МФТИ, 2010. 222 с.
2. Кондранин Т.В., Козодеров В.В., Дмитриев Е.В., Казанцев О.Ю., Бобылев В.И., Борзяк В.В., Егоров В.Д., Каменцев В.П., Беляков А.Ю., Логинов С.Б. Основы технологии восстановления количественных характеристик лесных экосистем по многоспектральным и гиперспектральным данным аэрокосмического зондирования // Современные проблемы дистанционного зондирования Земли из космоса, 2010. Вып.7. Т.1. С.299-309.
3. Козодеров В.В., Кондранин Т.В., Райкунов Г.Г., Казанцев О.Ю., Белоцерковский А.В., Асташкин А.А., Бобылев В.И., Дмитриев Е.В., Каменцев В.П., Борзяк В.В., Щербаков М.В., Лесуновский А.А. Аэрокосмическая гиперспектрометрия: летные испытания аппаратуры, программно-алгоритмическое обеспечение обработки данных // Исследование Земли из космоса, 2010. №5. С.59-68.

Automated airborne hyperspectral remote sensing data processing

**T.V. Kondranin¹, V.V. Kozoderov², E.V. Dmitriev³, V.D. Egorov³,
V.V. Borzyak², A.A. Nikolenko¹**

¹ *Moscow Institute for Physics and Technology (State University)
141700, Moscow region, Dolgoprudny, Institutsky per., 9
E-mail: kondr@kondr.rector.mipt.ru*

² *M.V. Lomonosov Moscow State University
119992 Moscow, GSP-2, Leninskie Gory
E-mail: vkozod@mes.msu.ru*

³ *Institute of Numerical Mathematics of Russian Academy of Sciences
119991, Moscow, ul. Gubkina, 8
E-mail: yegor@inm.ras.ru*

Problems are considered of the automated computational procedures of airborne hyperspectral data processing accented on the retrieval of the biological productivity of forest vegetation. Bases are shown of the proposed approach realization with description of the module that enables to solve the retrieval problem as well as to employ the users interface together with mapping of the final results.

Keywords: hyperspectral remote sensing imagery, data processing, biological productivity of forested vegetation.