

# Оптимизированная система хранения и представления географически привязанных спутниковых данных

В.Ю. Ефремов, Ю.С. Крашенинникова, Е.А. Лупян, А.А. Мазуров,  
А.А. Прошин, Е.В. Флитман

*Институт космических исследований РАН  
117997 Москва, ул. Профсоюзная, д. 84/32  
E-mail: [andry@iki.rssi.ru](mailto:andry@iki.rssi.ru)*

При построении систем спутникового мониторинга достаточно часто возникает задача предоставления пользователям доступа к спутниковым изображениям и продуктам их тематической обработки по конкретным регионам их интересов в необходимой географической проекции.

Обычно в качестве таких регионов используются:

- единицы административного деления: области, регионы, города и т.п.
- регионы, привязанные к различным географическим объектам: морям, рекам и др.
- регионы, соответствующие зонам ответственности служб и организаций

Наиболее простой подход, обеспечивающий доступ пользователей к данным по регионам наблюдения состоит в том, что соответствующие продукты в нужной географической проекции формируются уже на этапе тематической обработки спутниковых данных, а затем заносятся в архив, построенный по «региональному» принципу [1-3]. При таком подходе, доступ к данным по заданным регионам наблюдения реализуется при помощи интерфейсов, получающих из архива изображения по фиксированному набору регионов наблюдения в нужной проекции. Принципиальная схема построенной таким образом системы доступа к спутниковым данным приведена на рисунке 1.

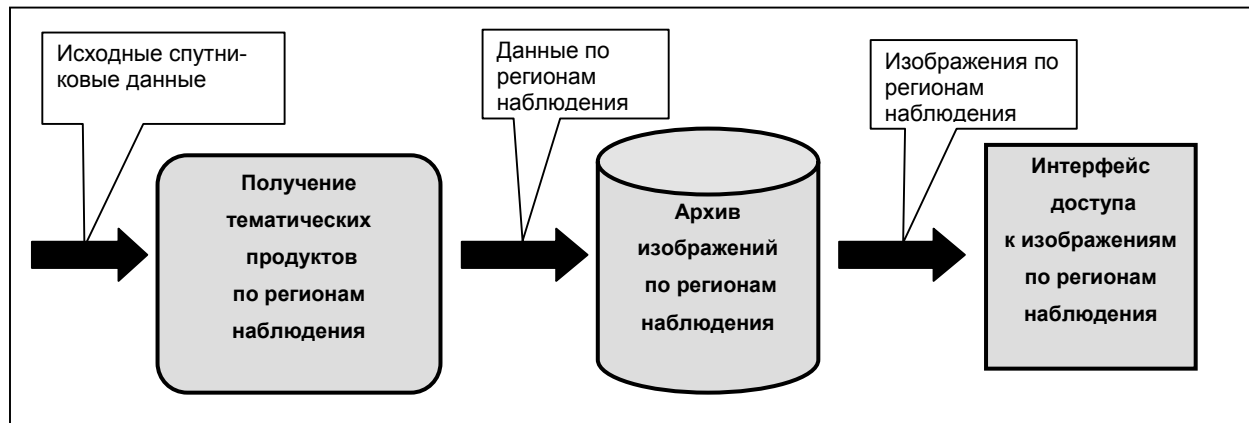


Рис. 1. Принципиальная схема организации доступа к данным по фиксированным регионам наблюдения

Представленный выше подход до недавнего времени с успехом применялся при построении различных систем спутникового мониторинга. Однако с увеличением требований к системам доступа к спутниковым данным и возрастанием степени их интеграции в различные информационные системы, возникли задачи, которые практически невозможно решить при использовании архивов, построенных по «региональному» принципу.

Во-первых, с развитием компьютерных сетей, возник спрос на WEB интерфейсы доступа к произвольной области спутникового изображения в нужном масштабе. В частности, большой популярностью в настоящее время пользуются интерфейс доступа к спутниковым изображениям

«Google Maps» и различные его аналоги. При этом такие картографические интерфейсы обладают функциональностью, ранее присутствовавшей только в ГИС системах, такой как возможность выделения области изображения и изменения его масштаба «на лету». В то же время, для WEB интерфейсов особенно остро стоит вопрос минимизации используемых ресурсов компьютера и трафика, передаваемого по сети Интернет. По этим причинам эффективно реализовать такой интерфейс на основе данных по отдельным регионам наблюдения на наш взгляд не представляется возможным.

Во-вторых, при построении современных систем спутникового мониторинга возникает необходимость в динамическом формировании спутниковых изображений. Это нужно как для предоставления доступа к одним и тем же спутниковым изображениям в различных географических проекциях, так и для получения изображений «на лету» на основе проведения различных операций над имеющимися данными. Таким образом, в частности, могут быть получены различные композитные изображения, усредненные данные и т.п. Использование динамического формирования спутниковых изображений позволяет не только обеспечить пользователю доступ к изображениям в архиве, но и предоставить ему различные инструменты для их анализа. В первую очередь это касается работы с различными тематическими продуктами, представляющими из себя значения того или иного индекса, например различных индексов отражающих состояние растительности.

В-третьих, при построении систем спутникового мониторинга остро стоит задача минимизации объемов хранящихся в архиве данных, что приводит к нежелательности дублирования информации и хранения большого числа производных продуктов. В частности, часто возникает ситуация, когда центрам приема, обработки и распространения данных приходится предоставлять одни и те же продукты, привязанные к различным, порою сильно пересекающимся регионам, что при использовании архивов построенных по «региональному» принципу приводит к нежелательному дублированию информации. При этом такие проблемы возникают в еще большей степени при интеграции продуктов обработки спутниковых данных одновременно в различные информационные системы.

В-четвертых, подсистемы обработки и архивации спутниковых изображений не должны зависеть от текущего набора регионов наблюдения. Хранение спутниковых изображений в архиве, построенном по «региональному» принципу, приводит к необходимости изменения настроек этих подсистем при каждом изменении регионов, что существенно понижает гибкость информационной системы в целом. При этом без проведения, и порой достаточно трудоемкой, обработки пользователь не может получить доступ к данным по региону за тот период времени, когда этот регион не был определен, либо имел другие границы.

Для решения вышеперечисленных задач был выработан новый подход к построению систем обработки, хранения и представления спутниковых продуктов, реализуемых в виде растровых изображений. Он основан на том, что данные в архиве хранятся в виде непересекающихся гранул небольшого размера, а требуемые изображения динамически формируются на сервере архивации. Это позволяет отделить задачу обработки и хранения данных в архиве от задачи обеспечения доступа к данным по конкретному региону наблюдения.

Принципиальная схема системы построенной по представленному подходу приведена на рисунке 2.

В рамках представленного подхода, территория, соответствующая максимальной зоне покрытия того или иного спутникового продукта (например, вся Россия или весь мир), разбивается на фиксированный набор непересекающихся гранул. Идея хранения спутниковых изображений в виде непересекающихся гранул, естественно, не нова. Такой подход, в частности, используется в архивах центра LP DAAC, входящего в состав проекта NASA EOSDIS, для хранения данных прибора MODIS, установленного на спутниках TERRA и AQUA. Данные в этом архиве хранятся в виде гранул в синусоидальной проекции большого размера (примерно 1000км. на 1000км.), что позволяет оптимизировать задачу хранения и предоставления пользователям под обработку

больших объемов данных. Однако такой вариант организации хранения данных не подходит для построения интерфейсов доступа к произвольной области спутникового продукта.

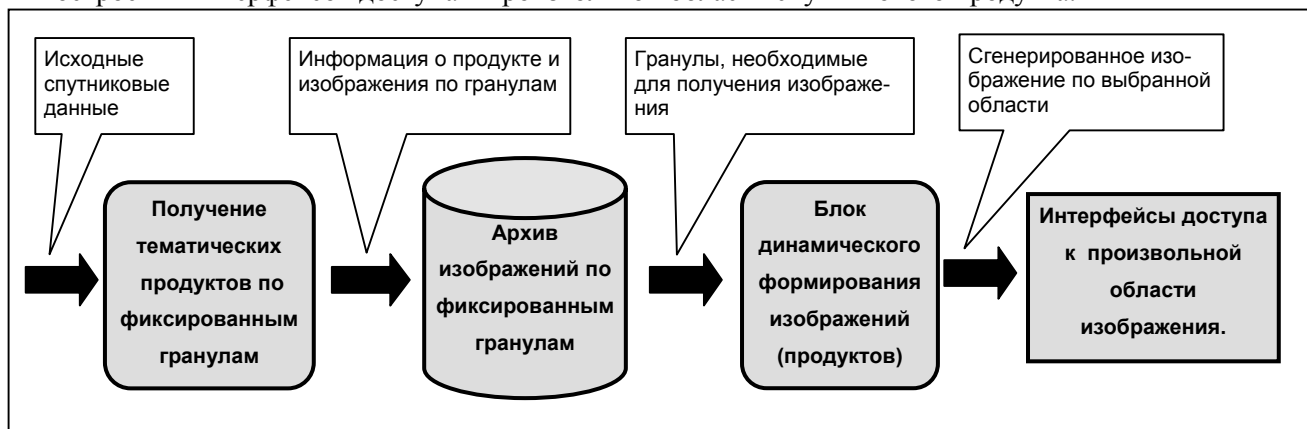


Рис. 2. Принципиальная схема организации доступа к произвольной области спутникового изображения

Для того чтобы эффективно реализовать интерфейс доступа к произвольной области спутникового продукта, гранулы должны иметь небольшой размер, а именно быть существенно меньше, чем размер предоставляемого пользователю изображения. Это позволяет минимизировать объем «лишних» данных, т.е. размер фрагментов гранул, которые не входят в результирующее изображение. С другой стороны, чрезмерное уменьшение размеров гранул приводит к накладным расходам, связанным с необходимостью оперировать большим количеством файлов, кроме того, при использовании различных форматов, использующих сжатие с потерей данных, в «склеенных» изображениях могут появиться граничные эффекты. Для эффективного отображения данных разного масштаба, наряду с исходным разбиением изображения на гранулы, должны быть реализованы дополнительные разбиения с большим шагом по координатам для «разреженных» гранул. Использование большего количества дополнительных разбиений с одной стороны увеличивает производительность интерфейса, а с другой - приводит к увеличению объема архива и накладных расходов, связанных с их получением.

После проведения всестороннего анализа различных вариантов разбиения спутниковых изображений на гранулы, нами было выбрано за основу разбиение, реализованное в системе «Google Maps». При использовании этого разбиения все гранулы в независимости от масштаба имеют одинаковый размер 256 на 256 пикселей, а для дополнительных разбиений используется ряд масштабов, каждый из которых вдвое больше предыдущего, т.е. гранула второго масштаба соответствует 4 исходным гранулам, третьего масштаба – 16 исходным гранулам и т.д. Количество дополнительных разбиений зависит от разрешения спутниковых данных и максимальной территории, которая может отображаться в интерфейсе доступа. Использование «кратных» масштабов позволяет получать каждый следующий масштаб на основе предыдущего, а не на основе исходного, что существенно сокращает накладные расходы на их создание. Суммарный объем всех дополнительных масштабов вычисляется по формуле суммы геометрической прогрессии и не превышает  $4/3$  от объема исходных гранул, что на наш взгляд является приемлемым вариантом.

Принципиальное отличие предлагаемого нами подхода к организации доступа к произвольной области изображения от варианта, используемого в системе «Google Maps», заключается в том, что формирование результирующего изображения производится на сервере, а не на клиенте. Это позволяет решить задачу динамического формирования спутникового изображения и представления его в нужной географической проекции, в то время как интерфейс системы «Google Maps» предназначен только для визуализации имеющихся в архиве данных. Существенным отличием также является возможность получения изображения в произвольном масштабе, в то время

как в системе «Google Maps» используется фиксированный набор масштабов. Это приводит к некоторому уменьшению скорости работы интерфейса, что связано с увеличением суммарного объема гранул, используемых для построения результирующего изображения. Так при использовании фиксированных масштабов для построения результирующего изображения могут потребоваться максимум  $(\lceil X/x \rceil + 1) * (\lceil Y/y \rceil + 1)$  гранул соответствующего масштаба, где  $X, Y$  и  $x, y$  соответственно размеры изображения и гранул в пикселях. В то время как для организации доступа к изображению в произвольном масштабе без потери разрешения может потребоваться почти вдвое больше гранул  $(\lceil X^2/x \rceil + 1) * (\lceil Y^2/y \rceil + 1)$  ближайшего разрешения. По этой причине для уменьшения накладных расходов, связанных с открытием большого количества файлов было решено увеличить размер гранул вдвое до 512 на 512 пикселей. В этом случае максимальное количество используемых для получения результирующего изображения гранул для размера изображения 800 на 800 пикселей уменьшается с 64 до 25. Заметим, что указанный размер гранул не является строго обязательным и в конкретных системах хранения может несколько отличаться от рекомендованного. Важно отметить, что перечисленные выше отличия от системы «Google Maps», в первую очередь связаны с различием решаемых ими задач.

Как уже было сказано выше, представленный подход призван обеспечить доступ к спутниковым изображениям в заданной пользователем проекции. Если эта проекция отличается от той, в которой хранятся исходные гранулы, то на этапе динамического формирования требуемого изображения производится необходимая конвертация проекции. Так как эта процедура является довольно затратной, то важно правильно выбрать проекцию для хранения гранул изображения и проекции, в которых будут строиться результирующие изображения. С одной стороны, для минимизации объема данных, т.е. устранения издержек, связанных с хранением данных по разным широтам целесообразно использовать для гранул синусоидальную проекцию. С другой стороны, эта проекция является довольно сложной и практически непригодной непосредственно для показа, что приводит к существенным затратам производительности на ее конвертацию. Накладные расходы на конвертацию изображений могут быть существенно уменьшены при использовании широтно-долготной проекции, однако при этом объем данных существенно возрастает. В качестве компромиссных вариантов могут быть использованы цилиндрическая или азимутальная эквидистантная проекции, которые позволяют хранить данные в более компактном виде и в то же время несущественно увеличивают накладные расходы на конвертацию изображений.

Как уже было отмечено, представленный подход предназначен для организации хранения продуктов обработки спутниковых данных, реализуемых в виде растровых изображений. Применение простых графических форматов в совокупности с метаданными, содержащимися в БД, позволяет в большинстве случаев избежать трудоемких операций, связанных с использованием специальных форматов для хранения спутниковых данных, таких как HDF и GeoTif. В то же время, при необходимости данные в этих форматах могут быть получены «на лету» или в пакетном режиме.

Важно заметить, что рассматриваемый подход к организации хранения спутниковых изображений является не только более сложным, но и существенно более требовательным к производительности сервера архивации спутниковых данных по сравнению с организацией архива по «региональному» принципу. Это вполне естественно, так как он позволяет реализовать гораздо более функциональные интерфейсы доступа к спутниковым изображениям. Поэтому на этапе разработки системы необходимо сопоставить предъявляемые к системе требования и имеющиеся в распоряжении компьютерные мощности с тем, чтобы выработать приемлемое по функциональности и производительности решение.

Рассматриваемый подход к организации архивов спутниковых изображений открывает новые перспективы для создания распределенных архивов спутниковых данных, с возможностью получения требуемых изображений «на лету» на основе гранул, находящихся на разных серверах. В первую очередь это применимо для получения различных композитных изображений на основе данных разных центров приема и обработки спутниковых данных. Кроме этого, даже в рамках одного информационного центра, может быть реализовано хранение спутниковых продуктов на

нескольких серверах, что позволяет повысить скорость работы интерфейса доступа к данным. Принципиальная схема построения таких интерфейсов приведена на рисунке 3. В случае использования во всех центрах единого разбиения на гранулы и одинаковой структуры архивов для получения итогового изображения достаточно запросить из центров необходимые гранулы. При использовании в центрах различного разбиения на гранулы, для построения интерфейса доступа к распределенному архиву в каждом центре должен быть установлен специальный шлюз для получения запрашиваемой области спутникового изображения.

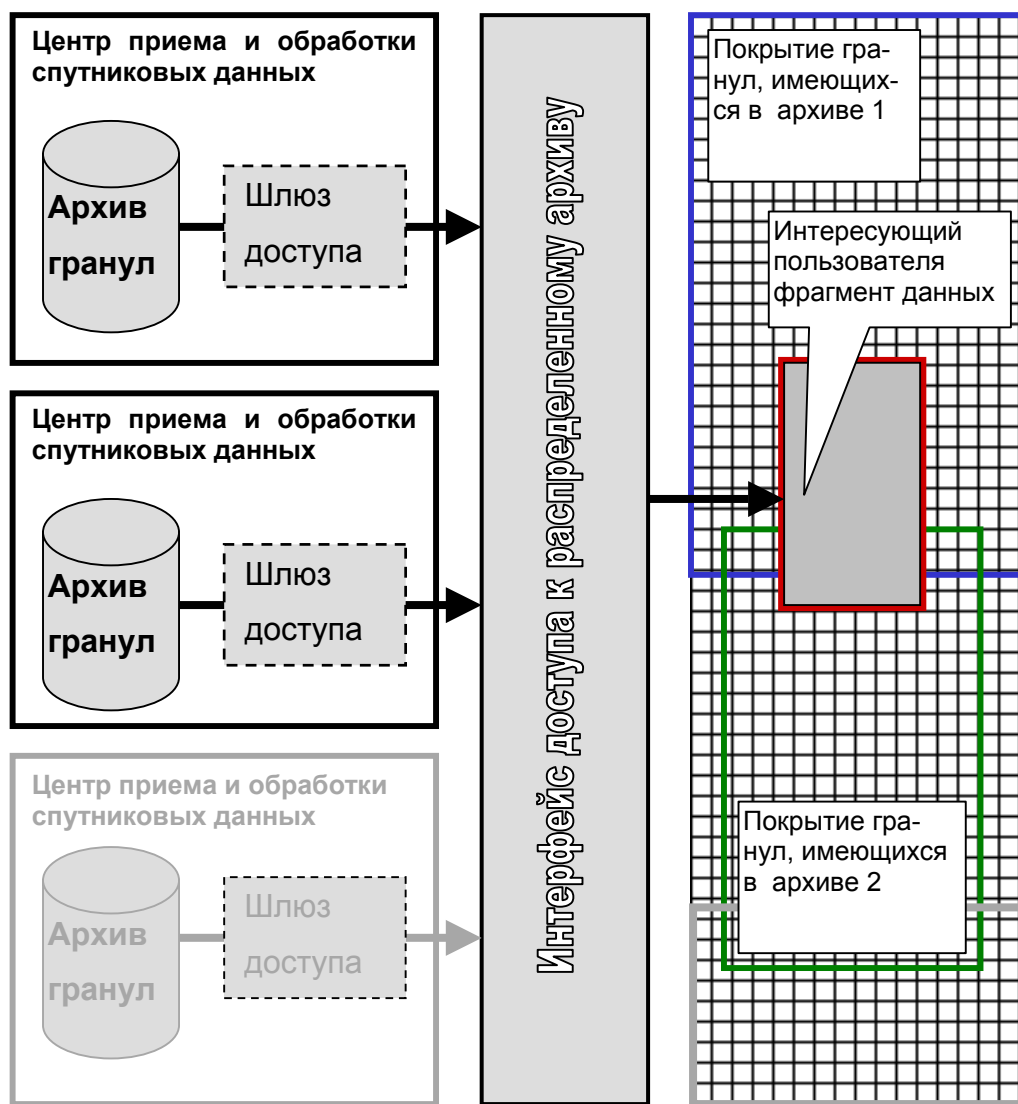


Рис. 3. Принципиальная схема построения интерфейсов доступа к распределенным архивам спутниковых данных

В настоящее время представленная технология используется нами в ряде систем спутникового мониторинга различного назначения. В качестве примера рассмотрим систему спутникового мониторинга сельскохозяйственных земель МСХ РФ [4]. На рисунке 4 приведен пример работы с интерфейсом спутникового мониторинга состояния растительности, входящим в состав этой системы (адрес интерфейса [http://193.232.9.72/geocover/cgi/geocover\\_proj.pl](http://193.232.9.72/geocover/cgi/geocover_proj.pl)). Интерфейс обеспечивает доступ к недельным композитам индекса NDVI по территории России, полученным по данным прибора MODIS, а также к динамически формируемым информационным продуктам, получаемым как разница с предыдущим композитом, разница с годом аналогом и другие. Также в интерфейсе реализован доступ к следующим продуктам, получаемым на основе отдельных спутнико-

вых сеансов: индекс NDVI, псевдоцветное изображение, изображение облачности. На выводимое изображение могут быть наложены различные картографические слои, маски озимых, пахотных земель и т.п. Интерфейс позволяет выбрать интересующую пользователя область изображения, увеличить или уменьшить масштаб изображения, сдвинуть изображение и т.п.

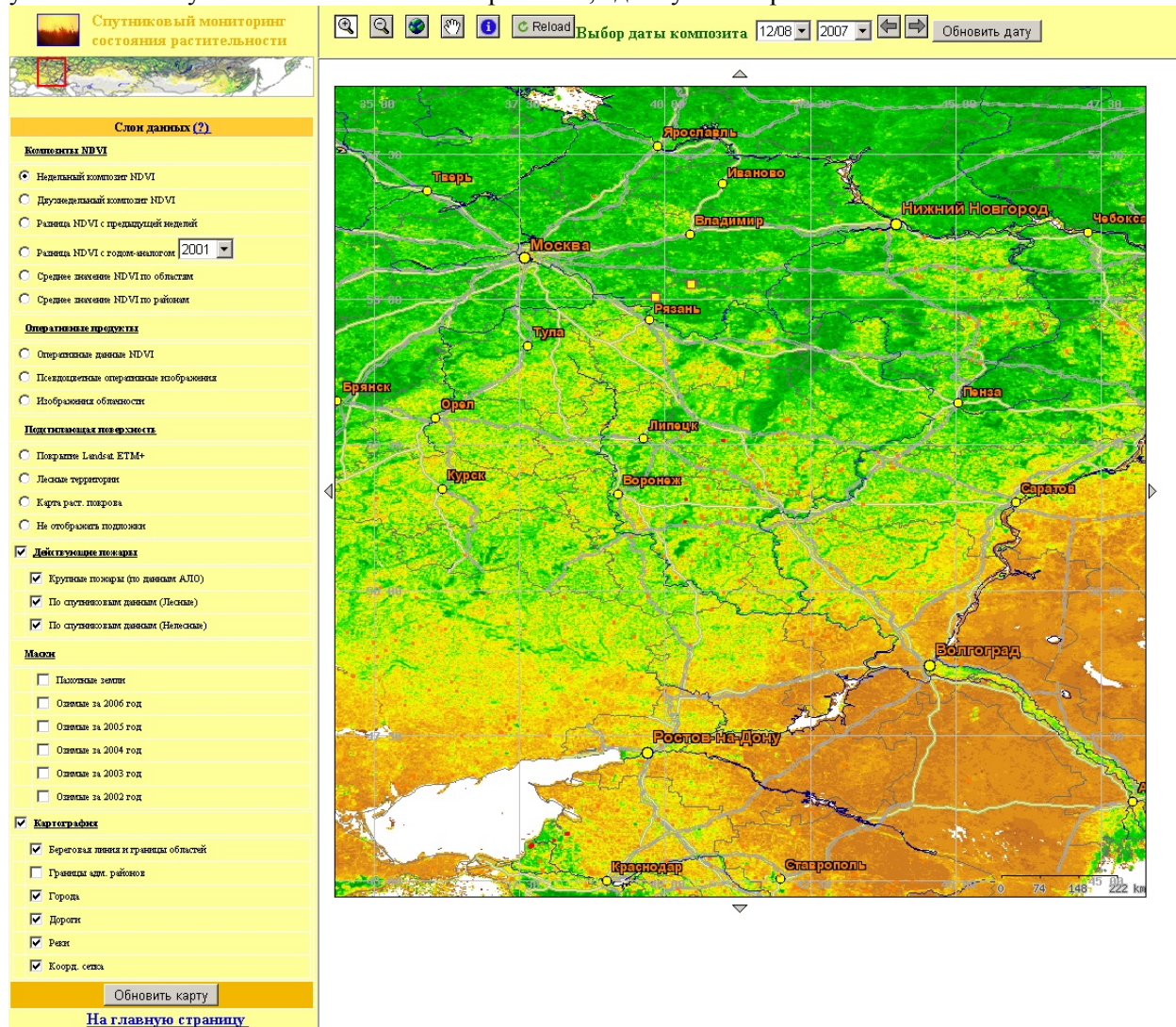


Рис. 4. Интерфейс спутникового мониторинга состояния растительности

Хранение недельных композитов индекса NDVI и информационных продуктов, полученных по отдельным сеансам, реализовано в виде гранул в цилиндрической проекции. При этом для вывода изображений в интерфейсе доступа они переводятся в широтно-долготную проекцию. Для упрощения системы обработки и представления данных для исходного разбиения изображения на гранулы был выбран шаг ровно в один градус по широте и долготе. Размер гранул по долготе и широте соответственно равен 380 и 500 пикселей, что соответствует соотношению расстояний соответствующих одному градусу для широты и долготы для средних широт, по которым в основном производится мониторинг растительности. Реализованы дополнительные разбиения на гранулы такого же размера, соответствующие 4, 16, 64 и 256 гранулам исходного разрешения. Генерация гранул дополнительных разбиений производится на этапе пополнения архива. Для того чтобы избежать эффекта «пестроты» на этапе генерации производных гранул используется усреднение по четырем соседним точкам. Отметим, что отличие описанного разбиения от описанного выше носит непринципиальный характер и связано, в первую очередь, с упрощением реализации.



Приведем некоторые сведения о реализации системы хранения и представления в рассмотренном примере. В качестве СУБД используется сервер MySQL версии 5. Для хранения изображений в архиве используется ПО FDB (File Data Base), разработанное в ИКИ РАН на языке программирования Си. База данных *granules\_products* содержит два аналогичных набора из четырех таблиц для недельных композитов и для данных по сеансам соответственно. Таблицы *composites*, *seances* описывают соответственно композитные продукты и продукты по сеансам, таблицы *composites\_product\_types*, *seances\_product\_types* – содержат описания типов продуктов, таблицы *composites\_granules*, *seances\_granules* – содержат описания самих гранул, а таблицы *composites\_granules\_types*, *seances\_granules\_types* – описывают реализованные типы гранул (разбиений). Система архивации спутниковых данных построена на основе по технологии построения автоматизированных систем хранения спутниковых данных, разработанной в ИКИ РАН [5]. В качестве основного языка реализации используется Perl версии 5.

WEB интерфейс реализован на основе CGI скриптов, разработанных на языке программирования perl. В качестве WEB сервера используется ПО Apache. Для вывода изображений по заданному региону в требуемой проекции используются программный пакет *Mapserver*, разрабатываемый в университете штата Миннесота и использующая его библиотека *MapImage*, разработанная в ИКИ РАН. Для высокопроизводительного склеивания отдельных гранул в изображение на сервере применяется графический программный пакет GD. На клиентской стороне используются возможности языка сценариев JavaScript и разработанные на нем библиотеки для обработки событий от мыши и проведения запроса к БД без перезагрузки страницы (технология Ajax), а также свободно распространяемая библиотека *wz\_dragdrop* (*Walterzorn Drag'nDrop DHTML Library*) для работы с графическими слоями.

В настоящее время нами активно ведутся работы по совершенствованию представленной в настоящей статье технологии, в частности:

- прорабатывается вопрос использования различных географических проекций для хранения гранул и для предоставления информации пользователям;
- разрабатывается механизм получения спутниковых продуктов на основе данных, находящихся на нескольких серверах;
- проводится оптимизация программного обеспечения для повышения скорости его работы;

В перспективе планируется внедрение рассматриваемой технологии в центрах приема спутниковых данных, сотрудничающих с ИКИ РАН, с целью построения распределенных систем хранения и представления спутниковых изображений. Также планируется разработка WEB интерфейсов для доступа к распределенным архивам спутниковых данных, позволяющих получать «на лету» композитные изображения по данным разных центров приема и обработки спутниковых данных.

## Литература

1. Андреев М.В., А.А. Галеев, В.Ю. Ефремов, В.О. Ильин, Ю.С. Крашенинникова, Е.А. Лупян, А.А. Мазуров, Р.Р. Назиров, А.А. Прошин, Е.В. Флитман. Построение автоматизированных систем сбора, хранения, обработки и представления спутниковых данных для задач мониторинга окружающей среды // Солнечно-земная физика, 2004. Вып. 5. С. 8-11.
2. Андреев М.В., Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Наглин Ю.Ф., Прошин А.А., Флитман Е.В. Построение интерфейсов для организации работы с архивами спутниковых данных удаленных пользователей // Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных объектов и явлений. Сборник научных статей. Москва, 10-12 ноября 2003 г. М.: Полиграф сервис, 2004. С. 514-520.
3. Асмус В.В., Лупян Е.А., Мазуров А.А., Милехин О.Е., Прошин А.А., Флитман Е.В. Возможности и технологии эффективного использования спутниковой информации в региональных систе-

мах мониторинга // Тезисы докладов Четвертой всероссийской конференции «Современные проблемы дистанционного зондирования Земли из космоса». Москва, ИКИ РАН, 13-17 ноября 2006 г.

4. *Акаткин Ю.М., Барталев С.А., Мельник Н.Н., Лупян Е.А., Нейштадт И.А., Ляпинков Д.В., Столпаков А.В., Темников В.Н., Толпин В.А., Флитман Е.В.* Опыт использования и перспективы развития системы спутникового мониторинга сельскохозяйственных земель МСХ РФ // Тезисы докладов Четвертой всероссийской конференции «Современные проблемы дистанционного зондирования Земли из космоса». Москва, ИКИ РАН, 13-17 ноября 2006 г.
5. *Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В.* Технология построения автоматизированных систем хранения спутниковых данных // Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных объектов и явлений. Сборник научных статей. М.: Полиграф сервис, 2004. С. 437-443.