



***Универсальный модуль растеризации
продуктов спутниковых данных UNISAT***

***Руткевич Б.П., Балашов И.В., Кашницкий А.В.,
Прошин А.А.***

***Институт космических исследований
Российской академии наук***

**XIX Всероссийская открытая конференция «Современные проблемы дистанционного зондирования Земли из космоса»
Москва, 2021**

Тезисы

В настоящее время в ИКИ РАН созданы и эксплуатируются целый ряд информационных сервисов дистанционного мониторинга, работающих со сверхбольшими распределенными архивами спутниковых данных. В рамках каждого из сервисов используются специализированные варианты обработки спутниковых данных, и создание такого количества систем было бы невозможным без технологии UNISAT (Прошин и др., 2019), обеспечивающей автоматическое ведение сверхбольших архивов и позволяющей описывать и формировать виртуальные продукты обработки спутниковых данных «на лету» по запросам пользователей и систем обработки. До настоящего времени для различных сценариев использования таких продуктов применялась различные программные блоки, решающие, по сути, одинаковую задачу.

Настоящий доклад посвящен описанию универсального модуля построения продуктов спутниковых данных UNISAT, который будет применен для всех сценариев получения продуктов обработки данных, хранящихся в архивах, включая построение карт, выдачу исходных данных, получение продуктов для задач расчета характеристик объектов. В работе описывается базовый алгоритм получения запрашиваемого продукта по области интереса. Описываются основные этапы обработки запроса на получение данных, включая формирование уточненного запроса с полными путями к данными и правил построения для необходимого продукта. Далее, подробно рассматриваются этапы обработки гео-привязанных данных с использованием библиотеки GDAL, где происходит перепроецирование данных из проекции данных в проекцию карты, обрезка данных по области интереса, формирование данных запрашиваемого размера и этап вычислений продукта, происходящий над числовыми массивами.

В докладе описываются основные задачи и проблемы, возникающие при обработке численных данных в соответствии с требованиями (правилами) для отдельных продуктов, включая задачи обработки различных сценариев работы, такие как попадание области интереса на несколько файлов с данными, необходимые вычисления для каждого канала в соответствии с выбранным продуктом, перенормировку данных, подготовку выходных данных при различных типах формируемого продукта.

В заключении приводится обзор основных нюансов, возникших при реализации модуля для различных версий программного пакета GDAL и языка программирования Python, а так же, результаты опыта эксплуатации в сценарии построения карт для картографических интерфейсов сервисов спутникового мониторинга.

Ключевые слова: ДЗЗ, UNISAT, продукты обработки, GDAL, Python

Отличия старой и новой версии модуля растеризации продуктов

- Визуально в интерфейсе или при вызове сцены изменения не видны.
- Новый модуль универсальный и применяется (может применяться) в разных интерфейсах.
- Объединяет в себе несколько сервисов (построение продуктов, выдача исходных данных).
- Добавлены новые возможности, например, гибкая работа с nodata, что позволяет делать прозрачность в выходных png.
- В новой версии ряд продуктов (например, композиты) работают значительно быстрее, за счет распараллеливания открытия файлов и ресурсоемких вычислений с использованием библиотеки GDAL.
- Новая версия работает без временных GTiff файлов, обработка массивов проходит в оперативной памяти. Работа с Dataset, а не с файлами целиком.
- Решена проблема с отображением данных в полосе от +180 до +200 градусов восточной долготы в различных проекциях.
- Появилась возможность отображать несколько uid в одном слое.

Программный модуль выполнен с использованием программного пакета GDAL 2.2 и выше, и поддерживает версию компилятора Python 2.7 и выше.

Создание модуля построения продуктов дало возможность применить тайловую схему работы интерфейса, что дает возможность сделать его более отзывчивым для пользователя и динамичным.

Пример картографического интерфейса

ВЕГА-PRO картографический и

← → ↻ 🏠 ⚠ Не защищено

☑ **Данные высокого разрешения**

Дата с **2021-06-28**
по **2021-09-01** - только за эту дату

Отобразить на таймлайне
 Только для выбранного периода в году

Фильтр
Качество данных
Выбранные данные
2021-09-01 08:06:58 MSI (ESA)

Список сцен

Сцены 1-150, всего 686

- 2021-09-01 08:08:24 MSI (ESA)
- 2021-09-01 08:07:55 MSI (ESA) 8%, с9
- 2021-09-01 08:06:58 MSI (ESA) 9%, с10
- 2021-09-01 08:06:58 MSI_BOA(ESA) 8%, с4
- 2021-09-01 08:06:29 MSI (ESA) 0%, с1
- 2021-09-01 08:06:29 MSI_BOA(ESA) 1%, с1
- 2021-09-01 06:36:53 ETM+ (USGS)
- 2021-09-01 06:36:29 ETM+ (USGS) 7%
- 2021-09-01 06:36:05 ETM+ (USGS) 7%
- 2021-08-31 08:36:46 MSI (ESA) 5%, с8
- 2021-08-31 08:36:46 MSI_BOA(ESA) 9%, с9
- 2021-08-31 07:48:12 MSI (ESA) 4%, с6
- 2021-08-31 07:47:14 MSI (ESA) 9%, с12

↑ Пред.порция В корзину
↓ След.порция Снять выделение

Только для выбранного продукта

Продукты:

0 37 30 140 210 km 0 00 42 30 45 00 47 30 40 00 100 Leaflet

Блок-схема модуля растеризации

- Вызов модуля построения продуктов:

- параметры карты
- правила построения продуктов;

- Вычисление:

- Чтение данных
- GDAL обработка
- мозаика фрагментов
- вычисление продуктов
- нормировка данных

- Выходные данные:

- png (в файл или stdout)
- GTiff (вариант выдачи исходных данных)

Вызов модуля построения продуктов:

Для вызова модуля необходимо задать ряд параметров:

- Параметры вызова (ширина и высота выходного изображения в пикселях, координаты области интереса, проекция выходной карты, тип выходных данных, uid -сцены).
- Правила построения продуктов json - содержат всю необходимую информацию для построения продукта. Правила формируются отдельным модулем с использованием БД UNISAT “на лету”. Свои правила формируются для каждого продукта, uid сцены, даты, области интересов. Могут содержать формулу вычисления, нормировку данных, и т.д

Выходные данные

В зависимости от правил построения продукта предусмотрены следующие режимы построения png файла:

- трехканальный RGB продукт,
- одноканальный палитровый (палитра может быть как в самом тиф файле, так и на внешнем хранилище),
- одноканальный серый,
- пустой прозрачный png в случае отсутствия данных в области интереса.

Модуль возвращает запрашиваемый продукт в формате png в файл или в stdout сразу для отрисовки фрагмента. Также для режима выдачи исходных данных есть возможность вернуть геопривязанный GTiff файл с необходимым разрешением обрезанный по области интереса или по полигону.

Блок-схема вычисления

Прочитать входные параметры и правила

Из правил поступает набор данных для формирования продукта. В частности приходит набор путей к файлам в хранилище UNISAT. Чтение данных организовано так, что к файлу в архиве обращение только один раз.

UNISAT
хранилище

Набор dataset с указанием каналов

Распараллеливание обработки входных tif файлов (до 8 потоков).
GDAL => перепроецирование, bbox, ширина, высота, dpi

Обработка набор NumPy массивов

- мозаика, сшивка данных
- вычисление в соответствии с правилами продукта
- нормировка данных
- формирование маски прозрачности по каждому каналу

Набор NumPy массивов

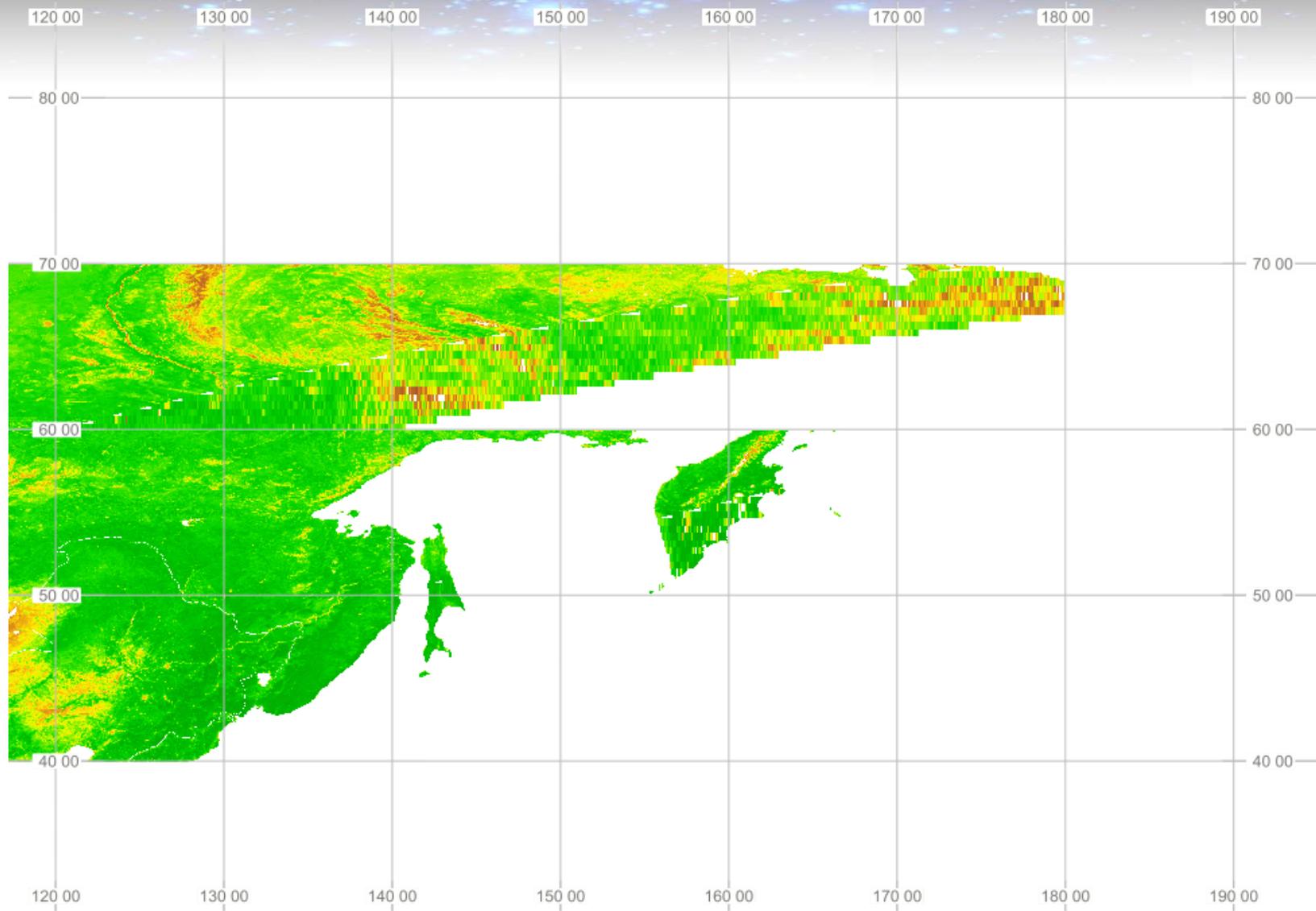
Особенности обработки композитов

При обработке некоторых специфических данных возникли сложности:

- Широкий тиф файл - перепроецирование широкого файл данных (360 на 10 градусов) в пакете GDAL приводит к пикселизации данных – т.е. искажению и значительному изменению размера пикселя, что становится заметным при достаточном увеличении. Решение – принудительно использовать базовый масштаб. (см пример далее)

- Файл данных охватывает долготу от -160 до +200. При обработке теряется полоса +180 +200. Решение – «на лету» преобразовать файл к долготам -180 + 180.

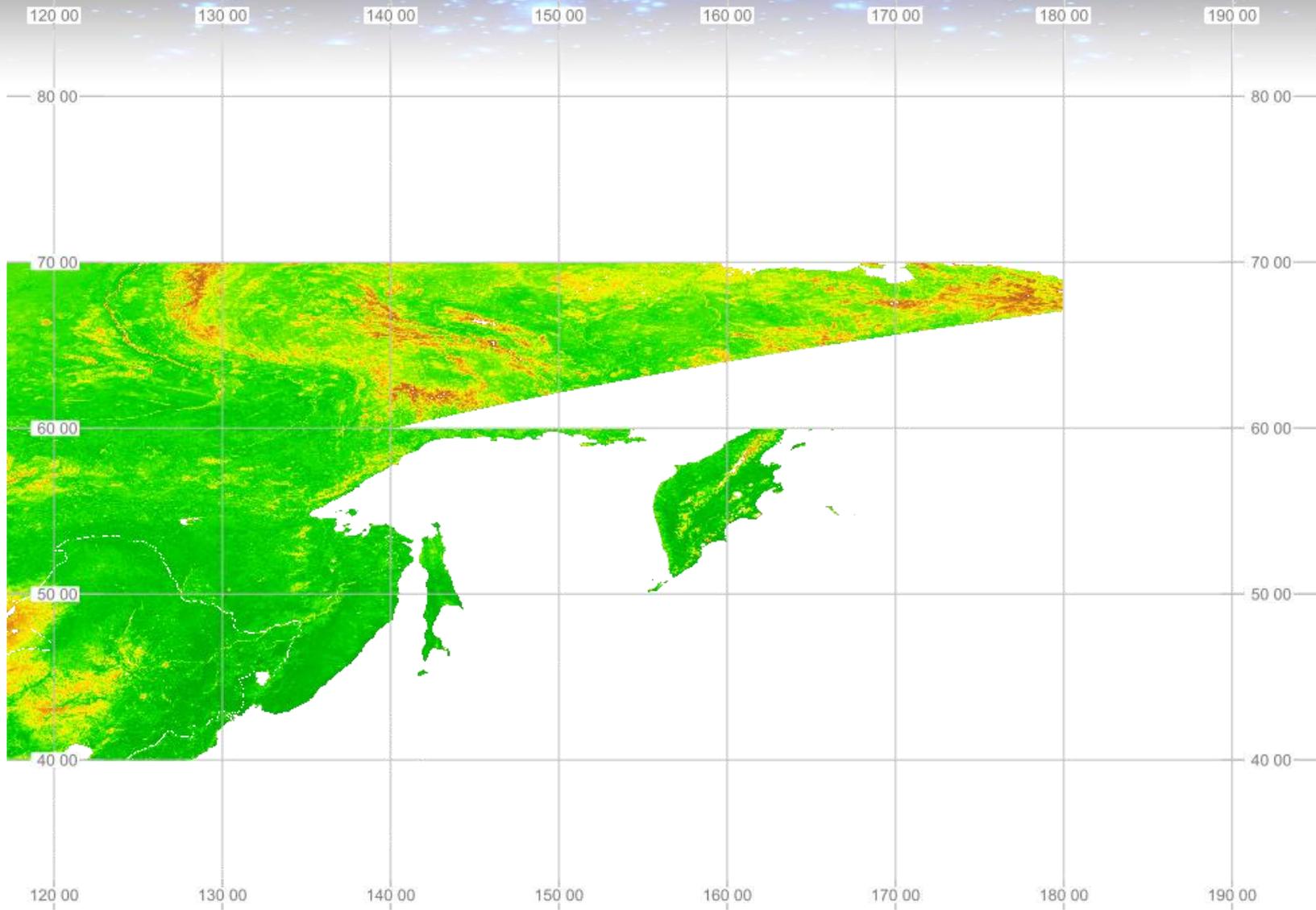
Пример пикселизации данных в композитах. Несколько входных файлов 360 на 10 градусов



октябрь 2021

ИКИ РАН

Устранили пикселизацию данных



октябрь 2021

ИКИ РАН

GDAL 2 vs GDAL 3

В последнее время программный пакет GDAL получил интенсивное развитие и вышла его третья версия. Поэтому ряд наших серверов имеет связку GDAL 2 & Python 2, а ряд сервером имеет GDAL 3 & Python 3. Здесь мы столкнулись с различием обработки ряда специфических случаев в программном пакете GDAL 2 и GDAL 3. Например:

- Если область интересе пересекает долготу +180 градусов в GDAL2 возникает проблема при одновременно перепроецировании и обрезании области интереса в одном `gdal.Warp`.

Литература

•Прошин А.А., Лупян Е.А., Кашницкий А.В., Балашов И.В. UNISAT. Технология построения унифицированных систем ведения сверхбольших распределенных архивов разнородных спутниковых данных // Обработка пространственных данных в задачах мониторинга природных и антропогенных процессов (SDM-2019) (Электронный ресурс): Сборник трудов всероссийской конференции (26-30 августа 2019 г., г. Бердск). Новосибирск: ИВТ СО РАН, 2019. С. 153-158.



***Спасибо за
внимание!***