

ПРИМЕНЯЕМЫЕ МЕТОДЫ ИНДЕКСАЦИИ СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ ПРИ ФОРМИРОВАНИИ КАТАЛОГОВ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

Мачнева Е.И.

Научный руководитель — аспирант Мальцев Е.А.

Сибирский федеральный университет

Каталогизация (создание и пополнение базы метаданных, описывающих космоснимки) является неотъемлемой частью этапа архивации данных ДЗЗ (физического размещения данных в архиве на определенных носителях). Она позволяет систематизировать данные, обеспечить быстрое извлечение информации о хранимых в архиве данных, организовать их поиск и выборку.

При приёме данных с КА перед этапом каталогизации решаются следующие предварительные задачи:

- предварительная обработка данных ДЗЗ;
- распаковка принятого потока данных;
- извлечение изображений и сопутствующей служебной информации;
- обработка и представление информации в форматах хранения.

Форматы хранения спутниковых данных.

В настоящее время не существует общих форматов хранения и распространения продуктов начальных уровней обработки, что объясняется уникальностью спутников, комплектов съемочной аппаратуры, режимов съемки и т. д. (Д.И. Федоткин, “Технологии предварительной обработки данных ДЗЗ: опыт ИТЦ «СканЭкс»”). Возможно, в будущем компании-операторы систем ДЗЗ договорятся и предложат пользователям стандартизованные форматы, но сейчас каждый оператор использует собственные форматы хранения. Часто структура этих форматов открыта, и компании идут навстречу пользователям, предлагая более приемлемые решения. Например, программа EOS NASA (спутники Terra, Aqua) предполагает хранение и передачу продуктов в формате EOS-HDF, являющемся модификацией общеизвестного формата HDF (Hierarchical Data Format), для представления данных научных исследований произвольного типа и состава. Существуют доступные программные средства, позволяющие работать с этим форматом, и многие современные системы обработки данных ДЗЗ поддерживают hdf-файлы. Другой пример: данные SPOT-4 распространяются в формате DIMAP, который содержит растр в формате GeoTIFF и метаданные (служебная информация) в формате XML, что существенно облегчает дальнейшее использование продуктов.

Обменным форматом спутника IRS-P6 является Fast Format, содержащий заголовочный файл и отдельные файлы для каждого канала изображения. Такой формат поддерживается различными ГИС ПО, в том числе ENVI, ERDAS Imagine и др. Используя информацию в заголовочном файле, пользователь может произвести просмотр и географическую привязку изображения.

Ведущие мировые компании-операторы систем ДЗЗ обычно рекомендуют (реже требуют) соответствия их номенклатуре выходных продуктов и форматов хранения данных, что обеспечивает стандартизацию хранения и распространения результатов ДЗЗ потребителями данных.

Далее представлена сводная таблица форматов метаданных, используемых различными компаниями-операторами систем ДЗЗ.

Спутник (страна)	Сенсор	Формат метаданных	Формат изображения
Terra, Aqua	MODIS	HDF	HDF
SPOT-4	HRVIR MONO, HRVIR XS	XML	TIFF
Resourcesat (IRS-P6)	LISS-III	Fast Format	Fast Format
EROS A, EROS B	PAN	*.PASS file	TIFF
NOAA	AVHRR	HDF	HDF
Quickbird-2	BHRC-60	XML	GeoTIFF
Landsat 7	ETM+	Fast Format	GeoTIFF
Landsat 5	TM	Fast Format	GeoTIFF

Вышеизложенная таблица показывает разнообразие, присущее обменным форматам данных ДЗЗ. Это разнообразие выливается в ряд проблем на разных этапах работы с данными. Стандартизация призвана их решить.

Зарубежные национальные и международные органы стандартизации активно прорабатывают терминологические и пользовательские вопросы данных ДЗЗ. Так, например, в Международной организации по стандартизации ISO находятся в различных стадиях проработки. Стандарты, посвященные ДЗЗ: ISO 19115-2 - Metadata - Part 2: Extensions for imagery and gridded data, ISO 19115-3 - Metadata - Part 3: Raster classes and elements, ISO 19121 - Imagery and gridded data, ISO 19130 - Sensor and data models for imagery and gridded data. В Федеральном комитете по географическим данным США FGDC (Federal Geographic Data Committee) принят стандарт для данных ДЗЗ FGDCSTD-012-2002 Content Standard for Digital Geospatial Metadata: Extensions for Remote Sensing Metadata, в государственных и частных учреждениях активно используется рекомендованный формат пространственных данных HDF-EOS (Hierarchical Data Format - Earth Observing System) Version 5.1.7. В Украине принят национальный стандарт ДСТУ 4220-2003 "Дистанционное зондирование Земли из космоса. Термины и определения понятий", включающий 52 термина. Различные международные организации, документы которых носят рекомендательный характер, также разрабатывают в области ДЗЗ терминологические стандарты и форматы данных ДЗЗ: Открытый консорциум по геоинформационным технологиям Open GIS (Open GIS Consortium Inc.), Международный комитет по ДЗЗ CEOS (Committee on Earth Observation Satellites), Международное общество по дистанционному зондированию и фотограмметрии ISPRS (International Society for Photogrammetry and Remote Sensing).

В России разработка стандартов ведется в рамках технических комитетов стандартизации Федеральной службы по техническому регулированию (Ростехрегулирование).

Основными документами, описывающими формат метаданных космоснимков являются ГОСТ Р 51353-99 «Геоинформационное картографирование. Метаданные электронных карт. Состав и содержание» и ГОСТ Р 52573–2006 «Географическая информация. Метаданные».

В настоящий момент не существует единого стандарта и подхода к предоставлению пользователю метаданных о спутниковых снимках. Это приводит к возникновению проблемы совместимости ПО для работы с данными ДЗЗ. Строгая стандартизация содержимого полей метаданных даёт возможность их применения в автоматических системах обработки спутниковой информации.

Методы индексации спутниковых изображений при формировании каталогов данных ДЗЗ.

Один из основных процессов каталогизации спутниковых изображений является индексация данных, которая включает в себя присвоение уникального номера изображению и сохранение сопутствующей ему метаинформации в БД. Варианты реализации

процесса индексации зависят от конкретной архитектуры построения каталога данных ДЗЗ.

При построении каталога данных ДЗЗ необходимо учитывать ряд факторов использования системы. Таких как:

- возможность распределённой работы пользователей с каталогом;
- возможность интеграции каталога с внешними информационными системами и системами управления (web-портал, настольная ГИС, сторонние механизмы архивации и копирования данных и т.д.);
- количество поддерживаемых спутниковых данных;
- возможность расширения номенклатуры спутниковых данных;
- открытость интерфейсов и структур данных.

На данный момент в мире существует множество каталогов спутниковых изображений. Но, в большинстве случаев, основа построения каждого каталога специфична.

При построении каталогов для обособленной работы пользователя, либо каталога масштаба предприятия, можно воспользоваться готовым решением сторонних производителей, к примеру, ArcCatalog ERSI. В случае, когда задачи, решаемые каталогом, более глобальны и выходят на региональный уровень (каталоги: ИКИТ СФУ Красноярск; СО РАН Новосибирск; ДВО РАН Хабаровск; ИКИ Москва; либо каталоги мирового уровня: USGS, Google и др.), то решения для построения таких каталогов индивидуальны.

Формирование каталога ДЗЗ на базе готовых решений

Стандарты FGDC и ISO в полной мере используются в приложении ERSI ArcGIS Desktop. Один из его модулей, ArcCatalog, предназначен для управления географическими данными. С его помощью можно находить нужные данные, просматривать содержимое и документировать их (создавать метаданные). ArcCatalog предоставляет два редактора – FGDC и ISO – для создания метаданных. При этом создается файл XML, включающий в себя свойства и документацию, описывающую данные. Свойства получаются автоматически из источника данных (например, географическая привязка и проекция из файла в формате GeoTIFF), документация же вводится вручную при помощи вышеуказанных редакторов и предоставляет информацию о том, как были получены данные, какими ключевыми словами описываются данные, определение атрибутов и т.д.

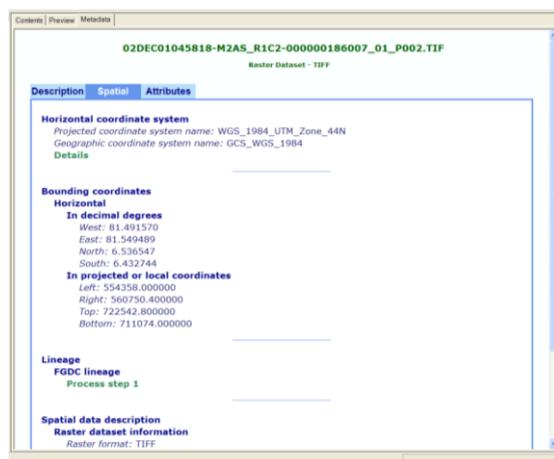


Рисунок 1 – Отображение XML файла метаданных в формате FGDC

При помощи тех же редакторов можно импортировать существующие файлы метаданных, но это должны быть текстовые или XML файлы, структурированные согласно форматам, поддерживаемым парсером метаданных FGDC, mp.

ArcCatalog успешно используется в геоинформационном портале <http://maps.dataplus.ru/Portal>. Неполный список функциональных возможностей геопортала включает в себя: интерактивную онлайн-подготовку пространственных метаданных с помощью встроенной в геопортал HTML-формы, ArcGIS ArcCatalog или любого XML-редактора; их публикацию средствами ArcIMS Metadata Server напрямую из интерфейса геопортала; поиск метаданных по карте или их атрибутам.

Другим примером использования стороннего ПО является региональный геопортал Самарской области. Здесь редактирование, просмотр и поиск метаданных осуществляются с использованием инструментальной системы «ИнМета»-разработка ЦСИ «Интегро», которая предназначена для создания реестровых приложений.

Формирование каталога ДЗЗ на базе собственных разработок

В некоторых случаях специфичность решаемых задач заставляет разработчиков самим создавать решения, не прибегая к помощи стороннего ПО. Примерами являются распределенные геоинформационные системы, разработанные в СО РАН и ДВО РАН. При реализации базы метаданных, лежащей в основе каталога спутниковых данных, разработан простой и легко расширяемый интерфейс для добавления в нее новых записей. В качестве интерфейса между программами сбора и базой метаданных используются текстовые файлы со строками вида «параметр = значение». Каждый такой файл должен содержать определение идентификатора и набора полей метаданных. Файлы генерируются скриптами, предназначенными для данных конкретного типа. Параметры, представление которых в текстовом виде затруднительно (например, обзорное изображение), хранятся как ссылки на внешние файлы. При поступлении нового файла в архив происходит запуск соответствующего скрипта, который формирует и добавляет в каталог метафайл и связанные с ним файлы. После этого метафайл дополняется полями, определяющими размещение файла в архиве.

В системе АСПКС, созданной в ИКИТ СФУ механизм обработки данных, поступающих с приемного оборудования, реализован следующим образом: данные спутника SPOT-4, Landsat-5, Landsat-7 и MODIS, прошедшие предварительную обработку, поступают на вход подсистемы обработки метаинформации. Затем выполняются PHP скрипты, производящие анализ метафайлов с выделением необходимой атрибутивной информации, которая заносится в базу данных, так же в базу данных поступают пространственные данные: файл предварительного просмотра изображения и файл иконки изображения. Параллельно работает модуль загрузки картографических данных в архив спутниковых данных. Достоинством такого подхода является простота реализации и быстрое действие. Однако при увеличении номенклатуры принимаемых данных появляется необходимость в написании дополнительных скриптов для каждого нового формата данных.

Заключение

На данный момент существует разнообразие форматов метаинформации, сопутствующей спутниковым изображениям. При построении каталогов данных ДЗЗ это приводит к необходимости использования большого количества ПО разных поставщиков, либо создания множества модулей, поддерживающих разнообразные форматы метаинформации. Значительно упростить процесс каталогизации данных сможет стандартизация обменных форматов в будущем. Несмотря на существование нескольких международных стандартов и российского ГОСТа, единого решения их применения пока не существует.