

# Форматы MrSID, HDF, JPEG 2000.

Выполнила: Казбек  
Мадина, Г-21

# MrSID

- формат графических файлов (с расширением *.sid*), разработанный и запатентованный компанией LizardTech. Используется для сжатия растровой графики, используемой в ГИС, например, для ортографически скорректированной аэрофотосъемки.

# Типичное использование

- **Геоинформационные системы**
- Формат MrSID разрабатывался для использования в ГИС. С этим форматом, огромные растровые изображения, такие как аэрофотосъемка и спутниковые фотографии, могут быть упакованы и быстро просмотрены по частям без необходимости распаковки целого файла. Популярные ГИС поддерживают данный формат, среди них: ESRI, ERDAS, Autodesk, MapInfo, Intergraph.
- Например, в этом формате NASA распространяет обработанные космические снимки со спутников LandSat-5, LandSat-7. В этом формате опубликовано несколько глобальных мозаик сухопутной части планеты.
- **Интернет**
- Поскольку формат файлов MrSID содержит изображения в разных масштабах и уровнях качества, он может использоваться для передачи данных в сети интернет. Только текущий просматриваемый уровень и часть изображения будут переданы в браузер, упрощая и ускоряя отображение, навигацию и просмотр изображения.

# ПО для работы с форматом

- LizardTech предлагает программу GeoExpress для чтения и записи MrSID файлов. Также существует бесплатный плагин браузера для просмотра. Многие коммерческие ГИС имеют поддержку чтения формата. ГИС ERDAS IMAGINE, используемая для обработки сырых растровых данных, также может создавать файлы в этом формате.
- Вьюверы Xnview и IrfanView также умеют читать файлы MrSID.

# Технология

- **Кодирование**
- При создании первичного изображения MrSID использует вейвлет-сжатие без потерь. Затем, изображение делится на несколько уровней детализации (создаются уменьшенные версии изображения), подканалов, подблоков и битовых плоскостей. После кодирования уровней к ним могут быть применены оптимизации и дополнительные режимы сжатия. Уменьшение объема изображения примерно до 2 раз может быть достигнуто без потерь, но большие степени сжатия будут отбрасывать часть информации, подобно алгоритмам сжатия с потерями в JPEG. В отличие от последнего в данном формате используются алгоритмы, дающее меньшее количество артефактов и сохраняющие четкие границы. Одной из особенностей алгоритмов отображения и упаковки является использование пиксельной сетки, повернутой на 45 градусов, и отображение каждого пикселя при большом увеличении с градиентной заливкой, а не со сплошной.

- **Декодирование (просмотр)**
- MrSID использует выборочное декодирование, то есть декодеру (просмотрщику) не требуется декодирование всего файла для отображения какого-то одного уровня детализации или качества изображения.
- **Привязка к координатам**
- Для сохранения метаданных о растре используются заголовки, подобные GeoTIFF GeoKeys. В них хранятся координаты снимка, его размер, использованные спектральные диапазоны, дата и время снимка.

# HDF

- **Hierarchical Data Format, HDF** (Иерархический формат данных) — название формата файлов, разработанного для хранения большого количества цифровой информации. Первоначально был разработан National Center for Supercomputing Applications, но сейчас поддерживаемый некоммерческой организацией HDF Group.
- Библиотеки для работы с форматом и связанные с ним утилиты доступны для использования под свободной лицензией, схожей с лицензией BSD. Формат HDF поддерживается многими коммерческими и некоммерческими программами, существуют библиотеки для работы с ним из Java, Matlab, IDL и Python. Свободно распространяемый пакет HDF состоит из библиотеки, утилиты командной строки, исходных текстов для тестирования, интерфейса для Java и Java-программы для просмотра HDF-файлов.

# HDF4

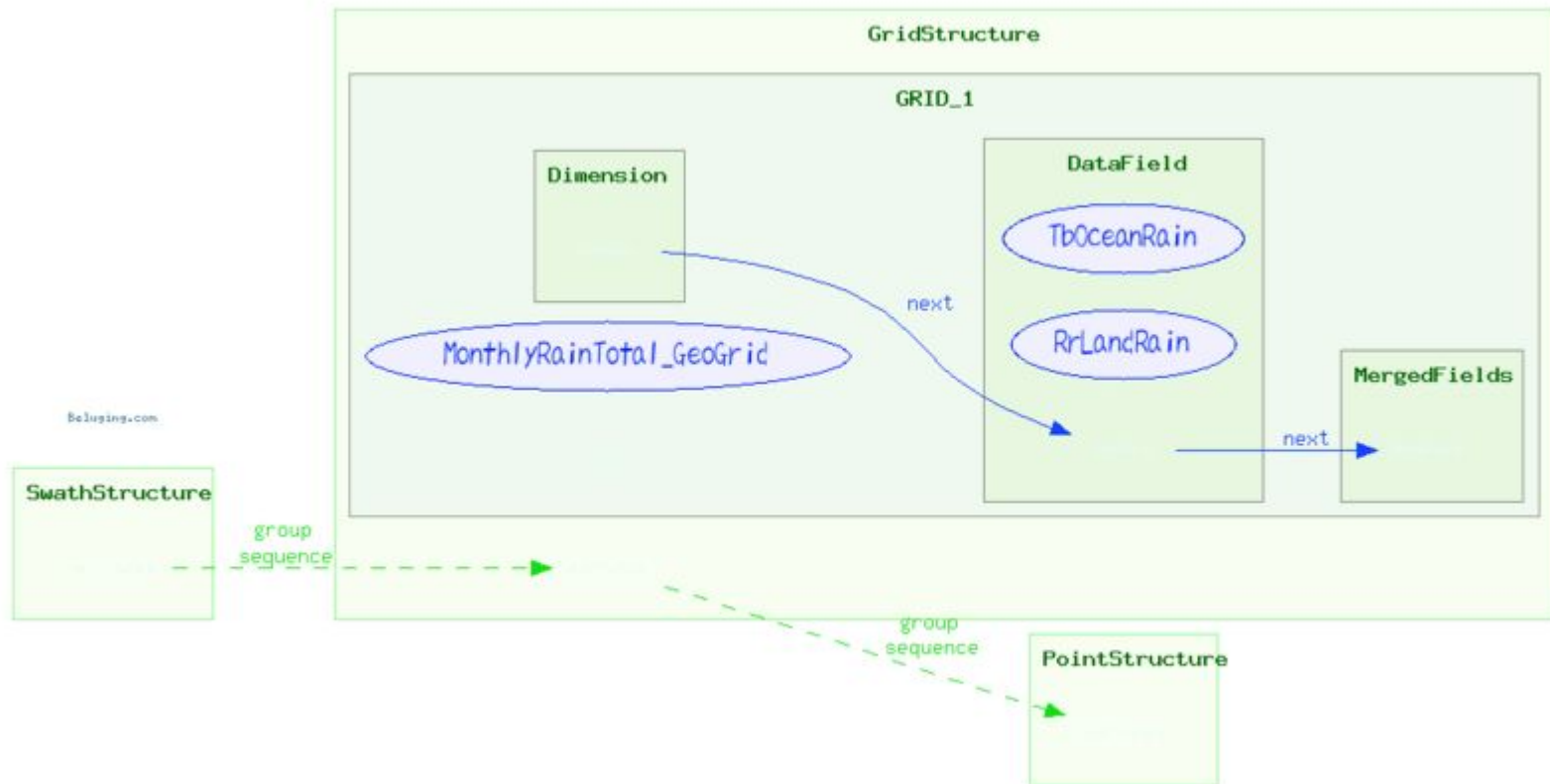
- HDF4 — старая версия формата, однако все еще активно поддерживаемая HDF Group. Он поддерживает различные модели данных, включая многомерные массивы, растровые изображения и таблицы.
- Использует 32-битные целые числа, поэтому имеет проблемы с хранением больших объёмов информации (более нескольких гигабайт).



# HDF5

- HDF5 — современная версия формата. Получил премию RD 2002.
- Содержит иерархию из двух основных типов объектов:
- HDF-Structure-Example
- Datasets — наборы данных, многомерные массивы объектов одного типа
- Groups — группы, являются контейнерами для наборов данных и других групп
- Содержимое файлов HDF5 организовано подобно иерархической файловой системе, и для доступа к данным применяются пути, сходные с POSIX-синтаксисом, например, */path/to/resource*. Метаданные хранятся в виде набора именованных атрибутов объектов.
- Формат NetCDF 4 основан на HDF5.
- Для индексации используются Б-деревья.

# HDF-структура-пример



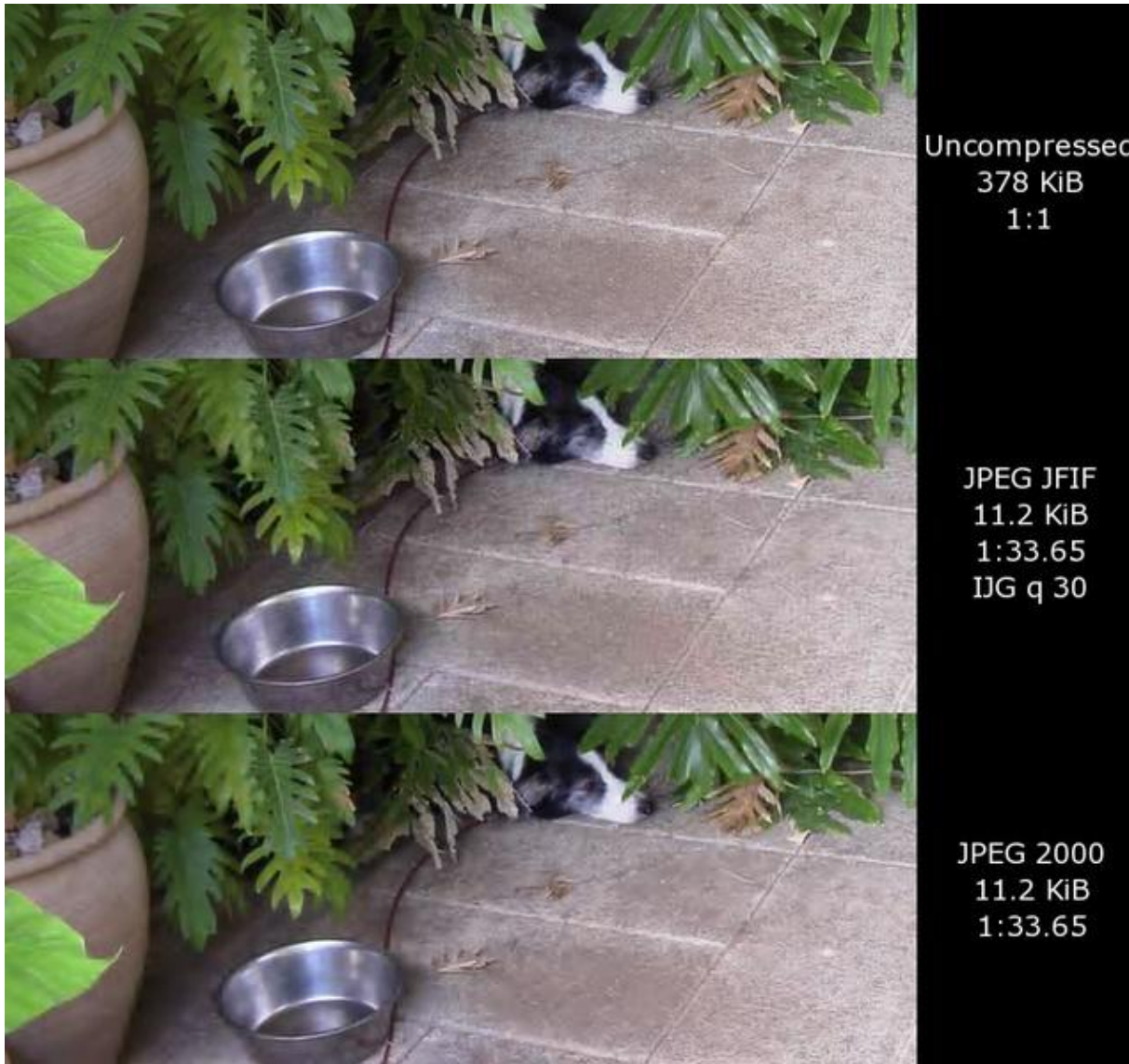
# JPEG 2000

- **JPEG 2000** (или **jp2**) — графический формат, который вместо дискретного косинусного преобразования, применяемого в формате JPEG, использует технологию вейвлет-преобразования, основывающуюся на представлении сигнала в виде суперпозиции базовых функций — волновых пакетов.
- В результате такой компрессии изображение получается более гладким и чётким, а размер файла по сравнению с JPEG при одинаковом качестве оказывается меньшим. JPEG 2000 полностью свободен от главного недостатка своего предшественника: благодаря использованию вейвлетов, изображения, сохранённые в этом формате, при высоких степенях сжатия не содержат артефактов в виде «решётки» из блоков размером 8x8 пикселей. Формат JPEG 2000 так же, как и JPEG, поддерживает так называемое «прогрессивное сжатие», позволяющее по мере загрузки видеть сначала размытое, но затем всё более чёткое изображение.
- Пока этот формат мало распространён и поддерживается не всеми современными браузерами.

# Преимущества

- Основные преимущества JPEG 2000 по сравнению с JPEG:
- Большая степень сжатия: на высоких битрейтах, где артефакты незначительны, JPEG 2000 имеет степень сжатия в среднем на 20 %, больше, чем JPEG. На низких битрейтах JPEG 2000 также имеет преимущество над основными режимами JPEG. Большая степень сжатия достигается благодаря использованию дискретного вейвлет-преобразования и более сложного энтропийного кодирования.
- Масштабируемость фрагментов изображений: JPEG 2000 обеспечивает бесшовное сжатие разных компонентов изображения, с каждым компонентом хранится от 1 до 16 бит на семпл. Благодаря разбиению на блоки можно хранить изображения разных разрешений в одном кодовом потоке.
- Произвольный доступ к кодовому потоку, также иногда называемый доступом к областям интереса кодовый поток JPEG 2000 обеспечивает несколько механизмов для поддержки произвольного доступа, также поддерживается несколько степеней разбиения на части (области интереса).
- Гибкий формат файла: форматы файлов JP2 и JPX обеспечивают хранение информации о цветовых пространствах, метаданных и информации для согласованного доступа в сетевых приложениях, взаимодействующих с помощью протокола JPEG Part 9 JPIP.
- Другие преимущества JPEG 2000 представлены на официальной странице.

# Сравнение JPEG и JPEG 2000



# Родственные стандарты

- К стандарту JPEG 2000 имеется несколько дополнений, некоторые из них ещё не реализованы, среди которых:
- ISO/IEC 15444-2:2000 — дополнительные функциональные возможности, такие как сеточное квантование, расширенный формат файла и дополнительные преобразования цветовых пространств;
- ISO/IEC 15444-4:2000 — тестирование ссылок;
- ISO/IEC 15444-6:2000 — составной формат файлов изображения, позволяет сжимать одновременно текст и изображение;
- JPSEC — расширение для безопасной передачи изображений (обсуждается в ISO);
- JPIP — основанный на соединении просмотр изображений (обсуждается в ISO).

# Приложения JPEG 2000

- Основные области применения этого стандарта:
- цифровой кинематограф;
- охранные системы (для сжатия изображений, получаемых с цифровых видеокамер, преимущественно как motion JPEG 2000, цифровые факсы, принтеры, сканеры);
- клиент-серверные взаимодействия (Интернет, базы данных изображений, видеосерверы);

# Сходства с компрессором ICER

- JPEG 2000 во многом сходен с форматом сжатия изображений ICER, используемым NASA для сжатия изображений в космических экспериментах.
- ICER также основан на вейвлетах и обеспечивает:
- прогрессивное кодирование,
- сжатие без потерь (в отличие от JPEG 2000, ICER в этом режиме использует модифицированный компрессор LOCO (англ. *Low Complexity Lossless Compression*)),
- сжатие с потерями,
- коррекцию ошибок, позволяющую ограничить эффект потери данных в канале связи.
- ICER в целом обеспечивает сжатие, сравнимое с JPEG 2000.
- Функциональные возможности ICER, сходные с JPEG 2000, состоят в том, что оба компрессора:
- обеспечивают разбиение изображения на блоки для увеличения эффективности сжатия, позволяя более эффективно использовать канал связи, оперативную память и процессорное время;
- позволяют варьировать степень сжатия в зависимости от размера изображения (в байтах);
- позволяют варьировать степень сжатия в зависимости от качества (хотя ICER варьирует степень сжатия с 1%-й погрешностью).
- Отличия между ICER и JPEG 2000:
- JPEG 2000 использует арифметику с плавающей запятой, ICER — только целочисленную арифметику;
- ICER использует модифицированный LOCO-компрессор для сжатия без потерь;
- JPEG 2000 использует несколько разных моделей сжатия без потерь, с помощью переключения вейвлет-компрессора в режим сжатия без потерь;
- ICER и JPEG 2000 используют разные цветовые пространства;
- ICER в своей нынешней форме лучше сжимает монохромные изображения, чем цветные



# Правовые споры

- JPEG 2000 не является свободным от патентованных алгоритмов компрессии, но усилиями комитета JPEG достигнуто согласие, что в составе этого формата они могут использоваться бесплатно:
- Всегда одним из самых больших преимуществ стандартов, выпущенных комитетом JPEG, было то, что они могут быть реализованы в базовой конфигурации без каких-либо лицензионных выплат. Новый стандарт JPEG 2000 был подготовлен с учетом этой возможности, согласие было достигнуто между 20 большими организациями-держателями большинства патентов в области сжатия, что позволило использовать их интеллектуальную собственность в составе реализации стандарта без лицензионных выплат.
- Разумеется, неопределённые и скрытые патенты могут всё ещё представлять опасность. Тем не менее, JPEG 2000 стоит рассматривать как более защищённый от притязаний формат, чем, например, MP3, для которого подобная работа велась на гораздо более низком уровне.
- Однако, не обращая внимания на свободу лицензирования патентов, JPEG 2000 всё равно не может соответствовать Debian Free Software Guidelines (тест на свободу программного обеспечения). Это может затруднить адаптацию JPEG 2000 к требованиям веба, так как это исключит свободные веб-браузеры (особо отметим браузеры, основанные на Gecko) и популярные веб-приложения LAMP.

# Артефакты, возникающие при сжатии

- Артефакты, возникающие при сжатии по алгоритму JPEG 2000 с высокой степенью компрессии (потерь), качественно мало отличаются от артефактов, возникающих при сжатии компрессором JPEG — в тех местах, где оригинальное изображение имело плавные цвето-переходы, — они становятся ещё более плавными (размытыми); в тех же местах где были резкие переходы яркости или цвета (участки изображения с высокой контрастностью) — возникает характерный артефакт в виде яркого контура, обрамляющего границу перехода, незначительно (на пару пикселей) отступающего от более тёмного участка перехода. Различия в артефактах — нет сетки в 8 на 8 пикселей; не искажаются цвета мелких деталей, сильно отличающихся по цвету от фона; артефакты, характерные для JPEG 2000, становятся заметными при больших, чем в случае JPEG, степенях сжатия.
- Часто фотографическое изображение может быть сжато в отношении 1/20 к оригинальному размеру без появления значительных искажений. Изображение справа демонстрируют различные искажения JPEG 2000 при различных степенях сжатия (верхнее изображение — это оригинал без сжатия). Цифры указывают соотношение размера оригинального Bitmap файла к размеру файла Jpeg2000 .

# Сравнение с PNG

- Хотя формат JPEG 2000 поддерживает сжатие без потерь, он не предназначен для усовершенствования наилучшего формата сжатия без потерь.
- Формат PNG (Portable Network Graphics) более эффективен для изображений, содержащих одноцветные области (при небольшом количестве цветов — например, < 1000), и поддерживает специальные функциональные возможности, которых нет у JPEG 2000.
- Считается, что в текущей реализации стандартов применение PNG более эффективно для сжатия диаграмм, а JPEG 2000 — для сжатия фотографических изображений.

**Спасибо за внимание!**