

Преобразование выходного продукта системы дистанционного зондирования Земли SENTINEL-1 в формат HDF5

Conversion of delivery product of Sentinel-1 Earth observing system to HDF5 format

Адамовский / Adamovsky E.

Егор Русланович
(adamovskiy.y@pdu.by)

Полоцкий государственный университет,
аспирант кафедры вычислительных систем и сетей.
г. Новополоцк, Республика Беларусь

Богущ / Bogush R.

Рихард Петрович
(bogushr@mail.ru)

кандидат технических наук, доцент.
Полоцкий государственный университет,
заведующий кафедрой вычислительных систем и сетей.
г. Новополоцк, Республика Беларусь

Ключевые слова: дистанционное зондирование Земли – Earth observing system; радиолокационное изображение – radar image; формат данных – data format.

Разработаны алгоритм и модель преобразования данных системы Sentinel-1, которые представляются в форматах TIFF и XML, в формат HDF5, используемый группировкой спутников COSMO-SkyMed. Модель реализована в пакете MatLab и включает: преобразование исходных данных, формирование структуры целевого файла и последовательную запись значений радиолокационного изображения и метаданных. Приведен обязательный набор метаданных продукта COSMO-SkyMed для обеспечения его дальнейшей обработки. Адекватность модели подтверждается примерами правильной обработки с применением программного обеспечения NEST 5.1.

Algorithm and model are developed for conversion of Sentinel-1 system data, delivered in the TIFF and XML formats, to HDF5 format used by COSMO-SkyMed group of satellites. The model is implemented in MatLab package and includes: conversion of initial data, generation of target file structure and consecutive recording of radar image and metadata values. Required metadata set of COSMO-SkyMed product providing its further processing. Adequacy of the model is confirmed by the examples of proper processing with the use of NEST 5.1 software.

Введение

Современные системы дистанционного зондирования Земли (ДЗЗ) представлены множеством действующих искусственных спутников Земли (ИСЗ), работающих независимо или в группе, снабженных аппаратурой радиолокационного синтезирования апертуры (РСА), которая позволяет получать информацию вне зависимости от времени суток и метеоусловий [1].

Продукты таких систем ДЗЗ содержат радиолокационные данные, которые обработаны до определенного уровня, и вспомогательную информацию, т.е. мета-

данные. Для различных спутниковых систем требования к уровням представления продуктов могут различаться. Международным комитетом по спутниковым наблюдениям (CEOS) принята следующая классификация [2]: 0-й уровень – необработанные данные, 1-й уровень – сфокусированное радиолокационное изображение (РЛИ), снабженное географической привязкой и калибровочными атрибутами, 2-й уровень и выше – реорганизованные, тематически обработанные РЛИ. Современные системы ДЗЗ (TerraSAR, Sentinel, COSMO-SkyMed и др.) используют уровень представления 1А, поскольку он включает минимально обработанные радиолокационные данные и набор необходимых для их дальнейшего преобразования метаданных, что расширяет спектр возможностей их дальнейшей обработки. Продукт уровня 1А содержит амплитуду и фазу отраженного сигнала. Начиная с уровня 1В и выше информация о фазе теряется и не может быть восстановлена на основе имеющихся данных. Также на уровне 1А предоставляется оригинальное и геометрически неискаженное изображение, при этом содержащее достаточный объем метаданных для различных геометрических преобразований, например наклонной дальности в наземную, проекция на эллипсоид или цифровую модель рельефа. Репроекция данных выполняется на уровнях 1В-D.

При разработке системы ДЗЗ на базе РСА осуществляется выбор стандартного формата представления продуктов или разработка нового, например COSAR системы TerraSAR. Формат определяет не только правила хранения радиолокационных данных и метаданных, но также программное обеспечение (ПО), совместимое с ним для дальнейшей обработки. Выбор формата требует рассмотрения особенностей хранения в нем основных радиолокационных данных: РЛИ, атрибуты идентификации, радиометрической калибровки, геопривязки, орбитальных векторов спутника и др.

Таким образом, актуальной является задача представления радиолокационных данных уровня 1А в формате, совместимом со специализированным ПО, применяемым для обработки данных ДЗЗ.

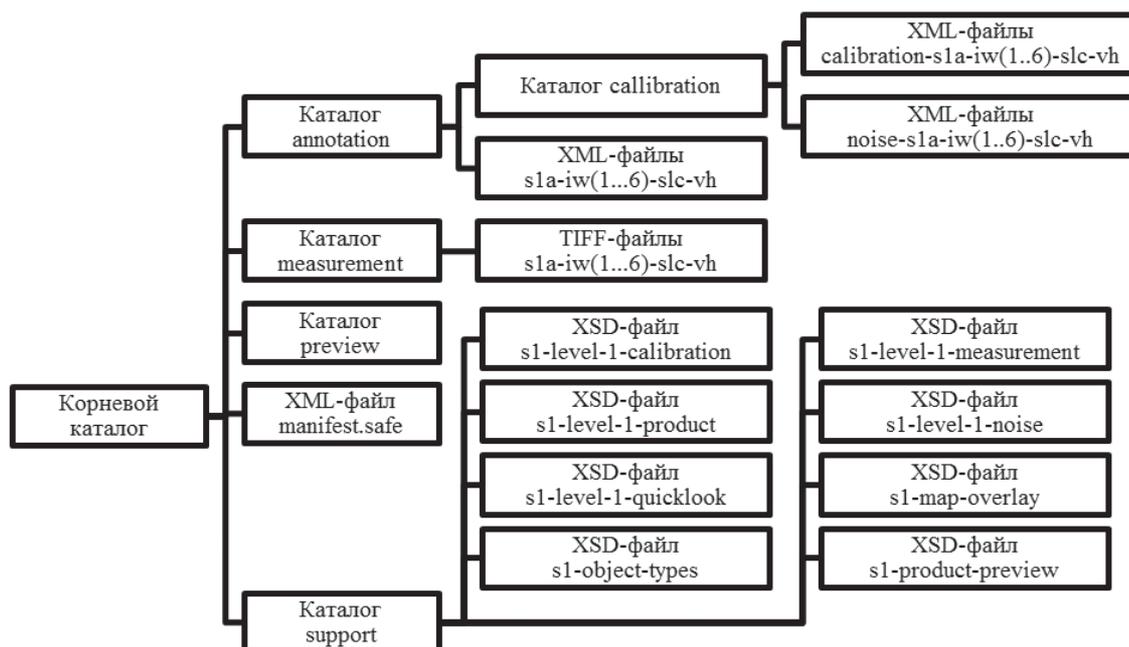


Рис. 1. Структура продукта Sentinel-1 IW SLC

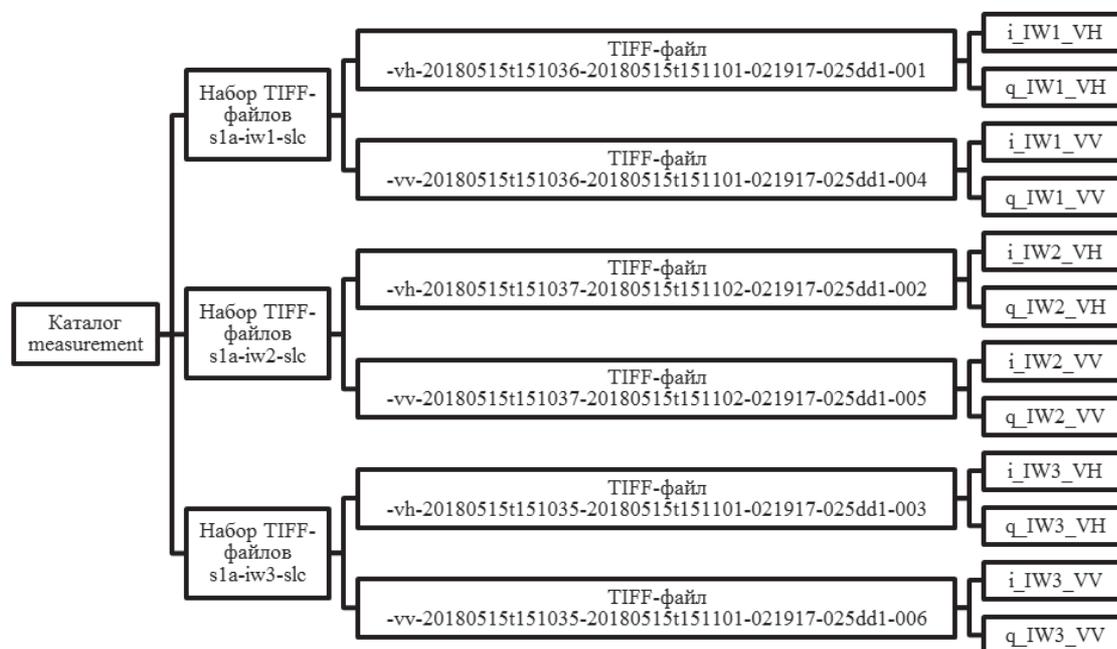


Рис. 2. Структура радиолокационных данных продукта IW SLC Sentinel-1

В первых разделах статьи рассмотрены особенности представления данных системы Sentinel-1 и формата HDF5. Далее рассматривается алгоритм и модель преобразования РЛИ и метаданных, включая описание основных блоков и особенностей программной реализации. И в подтверждение адекватности модели приведены результаты обработки преобразованного в формат HDF5 продукта Sentinel-1 свободно распространяемым ПО.

Продукт Sentinel-1

В качестве исходных радиолокационных данных использован продукт уровня 1A SLC IW Sentinel-1. Выбор обусловлен тем, что Sentinel-1 является действующей современной системой ДЗЗ, предоставляющей данные достаточной степени обработанности в широко используемых форматах TIFF и XML. Система ДЗЗ Sentinel-1 [3] состоит из спутников Sentinel-1A и Sentinel-1B, снабженных аппаратурой PCA, и предоставляет продукты 1–2 уровней:

- Уровень 1 SLC (Single Look Complex) – комплексные изображения со значениями амплитуды и фазы сигнала в наклонной дальности по азимуту. Снимки имеют географическую привязку и производятся в нулевой доплеровской геометрии. Одним из вариантов уровня 1 является продукт IW SLC (интерферометрический широкополосный продукт), который содержит по одному изображению на полосу и канал поляризации [4]. Структура IW SLC показана на рис. 1.

- Уровень 1 GRD (Ground Range, Multi-look, Detected) – изображения, спроецированные в горизонтальную дальность с использованием модели эллипсоида Земли (геоида) и имеющие равномерное пространственное разрешение с уменьшенным спеклшумом (за счет совмещения).

- Уровень 2 Ocean Product – данные уровня 1, подвергнутые тематической обработке и включающие компоненты: области ветра, радиальная скорость поверхности и спектр волн.

Общий доступ осуществляется через XML-файл manifest.safe, который содержит общую информацию о продукте. Файлы основных метаданных хранятся в каталоге annotation, вспомогательные данные – в support. Каталог preview предоставляет данные Quick Look для быстрого просмотра конечного продукта, представляющего совмещенные и тематически обработанные исходные снимки.

Основные радиолокационные данные записаны в шести TIFF-файлах, представляющих 32-битные изображения в двух вариантах поляризации (VH и VV). При этом каждое из них может быть разделено на два 16-битных канала с синфазными и квадратурными составляющими. На рис. 2 представлена иерархия радиолокационных изображений и их каналов в каталоге measurement.

Общая характеристика формата HDF5

За стандартный формат в статье принят Hierarchical Data Format (версия 5) [5], который в данное время используется группировкой COSMO-SkyMed [6], а его расширение HDF-EOS – программой NASA Earth Observing System [7]. Он позволяет хранить радиолокационные данные и метаданные практически без ограничения в едином файле, поддерживается распространяемыми языками программирования и совместим со специализированным ПО.

Преимущество расположения продукта в одном файле заключается в уменьшении количества используемых инструментов для работы с данными и ее стандартизации за счет единой формы представления информации. Наличие библиотек для работы с множеством языков (Java, Matlab, Python, C, C++ и др.) позволяет выбирать наиболее удобные для работы прикладные приложения из широкого списка вариантов. Также этот формат поддерживается рядом специализированных программных средств, предназначенных для обработки спутниковых данных. Иерархия данных, используемых форматом, а также структура продукта уровня 1A в HDF5 рассмотрены в работе [8].

Алгоритм представления радиолокационных данных в формате HDF5

В настоящее время специализированным ПО (NEST, Geomatica и др.) файлы формата HDF5 при определенных условиях идентифицируются как продукты COSMO-SkyMed, следовательно, для возможности проверки дальнейшей обработки данных после преобразования их, запись в HDF5 должна быть реализована с учетом спецификации этой системы ДЗЗ [9].

Разработан алгоритм представления радиолокационных данных продукта IW SLC Sentinel-1 в формате HDF5, включающий следующие основные шаги:

1. Преобразование входных данных. Метаданные Sentinel-1 содержатся в файлах формата XML и TIFF-изображениях, включая информацию о контрольных точках – наборе координат, поставленных в соответствии определенным пикселям, служащих для геопривязки продукта.

Спецификацией продукта уровня 1A COSMO-SkyMed установлено требование к РЛИ – наличие двух каналов, имеющих тип данных int16. Исходное 32-битное изображение продукта Sentinel-1 должно быть разделено на 16-битные каналы с синфазной и квадратурной составляющими.

Должны быть устранены несоответствия значений, отражающих физические параметры, единицам измерения системы СИ. Атрибуты, являющиеся массивами данных, необходимо перегруппировать согласно спецификации COSMO-SkyMed, поскольку представление информации в форматах HDF5 и XML значительно отличается.

Таблица 3

Соотношение атрибутов орбитальных векторов состояния спутника

Система отсчета	Атрибут COSMO-SkyMed	Атрибут Sentinel-1	Описание	
ECEF	ECEF Satellite Position	x	position x	Положение спутника по оси X
		y	position y	Положение спутника по оси Y
		z	position z	Положение спутника по оси Z
	ECEF Satellite Velocity	x	velocity x	Вектор скорости спутника по оси X
		y	velocity y	Вектор скорости спутника по оси Y
		z	velocity z	Вектор скорости спутника по оси Z
		Number of StateVectors	orbitList count	Количество состояний
	State Vectors Times	time	Отсчеты по времени	
ECI	Inertial Reference Frame ID	frame	Идентификатор инерциальной системы отсчета	
	Attitude Quaternions	[q0 q1 q2 q3]	Кватернионы спутника	
	Attitude Times	time	Отсчеты по времени	
	Inertial Satellite Velocity	[wx wy wz]	Векторы скорости спутника по 3 осям	
	Pitch Rate	pitch	Скорость наклона	
	Roll Rate	roll	Скорость крена	
	Yaw Rate	yaw	Скорость рысканья	

Листинг 2

Преобразование формы записи отсчетов времени продукта Sentinel-1

```

for i = 1:str2num(meta.Children(6).Children(6).Attributes.Value) % цикл по количеству
векторов
[t1, t2] = strtok(t(i).Value, 'T'); % обработка данных UTC: % разделение на дату и время
t2 = t2(2:end); % исключение лишних символов
t3(i) = 3600*str2num(t2(1:2))+ % отсчеты времени переводятся секунды из часов,
+60*str2num(t2(4:5))+ % минут,
+str2num(t2(7:end)); % и секунд
end % завершение цикла
h5writeatt(h5_name, '/', 'State Vectors Times', t3'); % запись отсчетов по времени
    
```

Имя файла:	CSKS<_>_	<YYY_Z>_	<MM>_	<SS>_	<PP>_	<s><o>_	<D>	<G>_	<YYYYMMDDhhmmss>_<YYYYMMDDhhmmss>
Атрибут COSMO-SkyMed:	идентификатор спутника	тип продукта	режим работы	полоса обзора	поляризация	направление съемки и орбиты	режим доставки	доступность во времени сбора данных	время начала и окончания зондирования
	<i>Satellite ID</i>	<i>Product Type</i>	<i>Acquisition Mode</i>	<i>Beam ID</i>	<i>Polarisation</i>	<i>Look Side, Orbit Direction</i>	<i>Delivery Mode</i>	<i>Selective Availability Status</i>	<i>Scene Sensing Start/Stop UTC Time</i>

Рис. 3. Структура имени продукта COSMO-SkyMed

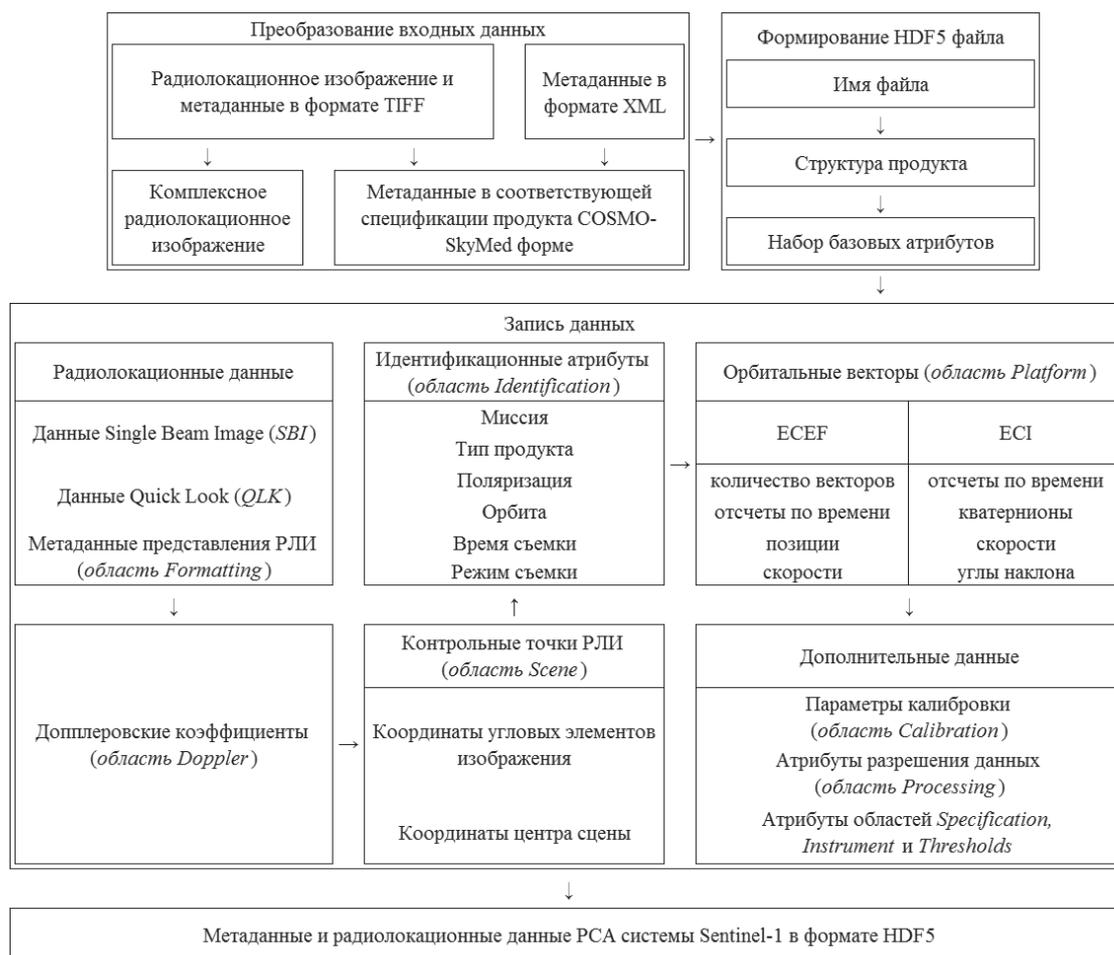


Рис. 4. Модель представления радиолокационных данных системы Sentinel-1 в формате HDF5

Таблица 4

Соответствие атрибутов TIFF-файла Sentinel-1 и COSMO-SkyMed

Атрибут TIFF-файла	Атрибут COSMO-SkyMed	Описание
<i>BitsPerSample</i>	<i>Bits per Sample</i>	бит на пиксель
<i>SamplesPerPixel</i>	<i>Samples per Pixel</i>	отсчетов на пиксель
<i>MaxSampleValue</i>	<i>Image Max</i>	максимальное значение РЛИ
<i>MinSampleValue</i>	<i>Image Min</i>	минимальное значение
<i>SampleFormat</i>	<i>Sample Format</i>	тип данных РЛИ
<i>ModelTiepointTag</i>	<i>Bottom/Top Left/Right Geodetic Coordinates</i>	координаты крайних угловых элементов РЛИ
<i>GeoDoubleParamsTag</i>	<i>Ellipsoid Semimajor/Semiminor Axis</i>	размеры осей эллипсоида

Листинг 3

Чтение, обработка и запись атрибутов контрольных точек продукта Sentinel-1

```

info = imfinfo('D:\...\image.tiff'); % чтение метаданных TIFF-файла
N = numel(info.ModelTiepointTag)/6; % определение количества точек
Coordinates=reshape(info.ModelTiepointTag,6,N)'; % перевод данных в матрицу
Coordinates = Coordinates(:,4:end); % удаление координат пикселей
Columns = N/swathNumber; % количества столбцов точек
c_topleft = Coordinates(1,:); % чтение контрольных точек
c_topright = Coordinates(Columns,:); % чтение контрольных точек
c_bottomleft = Coordinates(N-Columns,:); % чтение контрольных точек
c_bottomright = Coordinates(N,:); % чтение контрольных точек
h5_name = 'D:\...\CSK_Sentinel.h5'; % имя целевого файла HDF5
h5writeatt(h5_name,'/S01/SBI','Top Left...',[c1(2) c1(1) c1(3)]);
h5writeatt(h5_name,'/S01/SBI','Top Right...',[c2(2) c2(1) c2(3)]);
h5writeatt(h5_name,'/S01/SBI','Bottom Left...',[c3(2) c3(1) c3(3)]);
h5writeatt(h5_name,'/S01/SBI','Bottom Right...',[c4(2) c4(1) c4(3)]); % запись данных в файл
HDF5
    
```


не задействованные в других блоках, например данные области Instrument, среди которых информация о поляризации и длине антенны, частоты радара и импульсов.

Область PCD (Precise Orbit Determination) содержит параметры для точного определения орбиты и статистическую информацию о данных RAW: минимальное, среднее и максимальное значение, а также процент перенасыщенных и недонасыщенных пикселей относительно пороговых значений, установленных в области Thresholds. Атрибуты области Specification представляют основные технические характеристики системы, заявленные производителем.

Атрибуты, отображающие информацию о разрешении данных, представлены в области Processing и дают представление о минимальном размере объектов, доступных для визуального анализа с помощью радиолокационного изображения, а также используются при геометрических преобразованиях РЛИ:

- Column Spacing – интервал между соседними столбцами РЛИ;
- Line Spacing – интервал между соседними строками РЛИ.

Калибровочные данные, в частности атрибут Calibration Constant, отображающий пропорцию между целевой энергией в исходном продукте и фактическим обратным

рассеянием, представлены в области Calibration и могут быть использованы для радиометрической коррекции РЛИ.

Результаты моделирования

На основе разработанной программной модели выполнено преобразование метаданных и радиолокационных данных PCA уровня 1A спутниковой системы Sentinel-1 в файл формата HDF5.

Для работы с продуктами системы ДЗЗ Sentinel-1 используется специализированное ПО Sentinel Application Platform (SNAP) [10] и ПО Next ESA SAR Toolbox (NEST 5.1) [11]. Поэтому адекватность реализованной модели может быть подтверждена возможностью работы с полученными данными с применением данного ПО. В работе использовано ПО NEST 5.1. На рис. 5 представлены изображения каналов Intensity и Phase продукта Sentinel-1 в формате HDF5, автоматически сформированные на основе каналов оригинального комплексного изображения по формулам (2–3).

$$Intensity = i^2 + q^2 \quad (2)$$

$$Phase = \arctan 2(q, i) \quad (3)$$

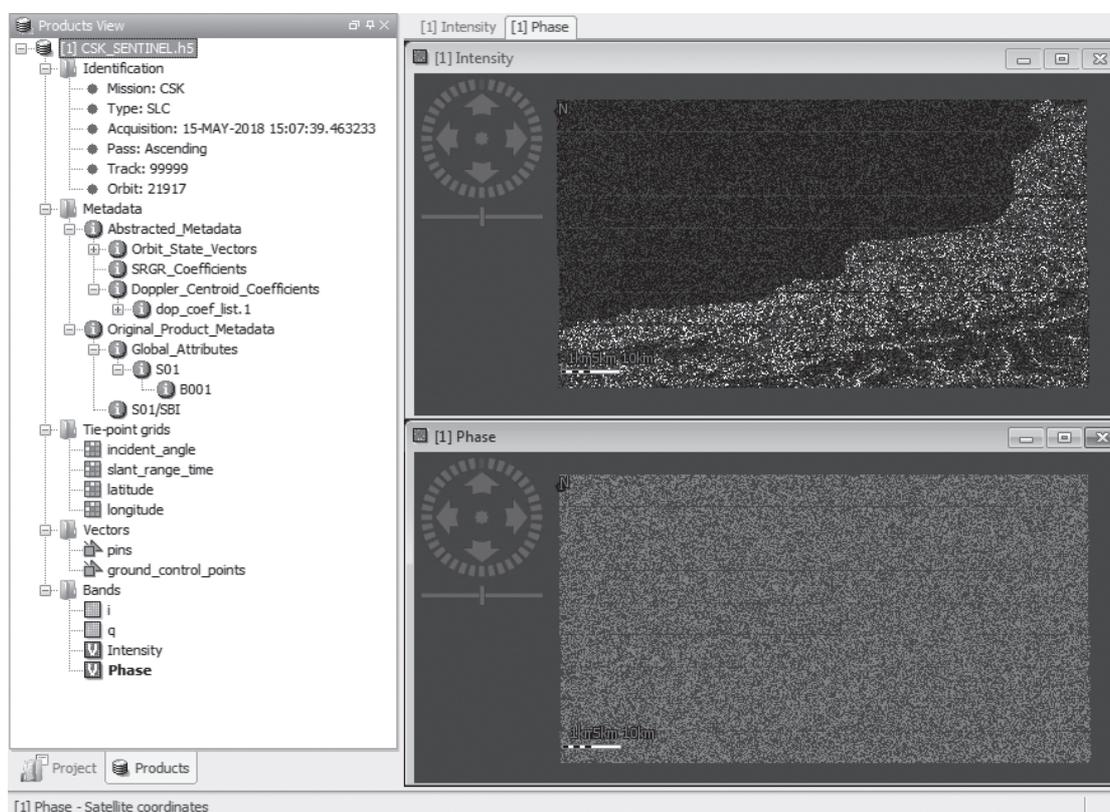
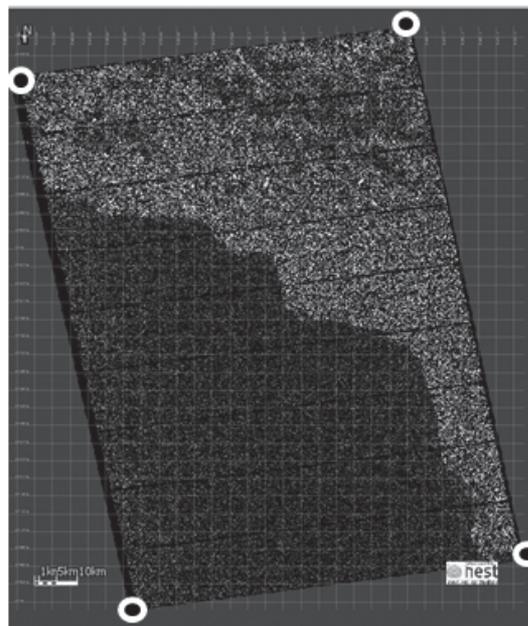


Рис. 5. Радиолокационные данные продукта Sentinel-1, каналы Intensity и Phase в NEST 5.1



- - координаты крайних угловых элементов РЛИ

Рис. 6. Результирующее изображение, обработанное инструментом Reprojection в NEST 5

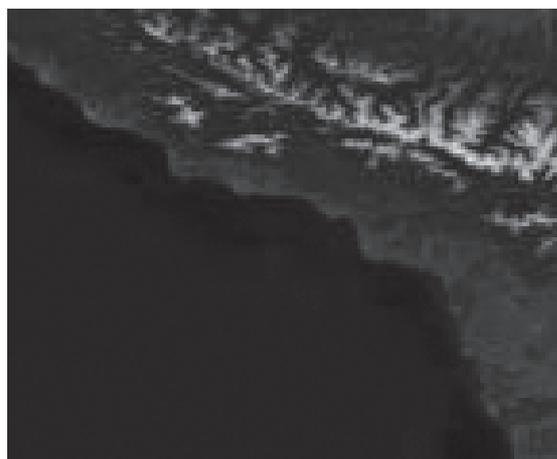


Рис. 7. Проекция РЛИ на прямоугольную координатную сетку а) участок карты NASA Blue Marble

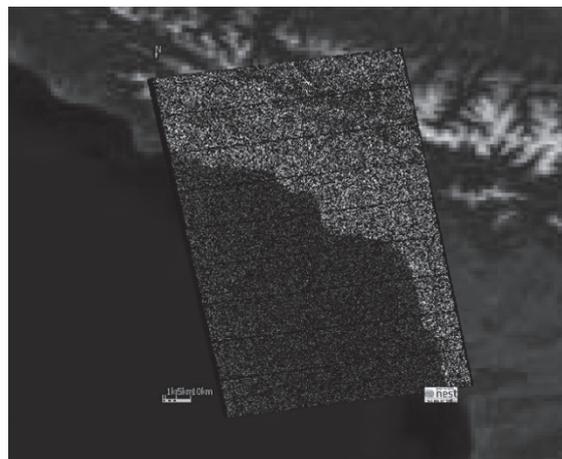


Рис. 7. Проекция РЛИ на прямоугольную координатную сетку б) аналогичный участок карты с проекцией результирующего изображения

где i – канал синфазных составляющих;
 q – канал квадратурных составляющих.

В качестве примера обработки радиолокационных данных в формате HDF5 применен инструмент преобразования РЛИ в ортогональную проекцию (вкладка Geometry > инструмент Reprojection). Используя географические координаты контрольных точек (крайних угловых элементов) изображения, приложение геометрически искажает его так, чтобы они соответствовали прямоугольной земной системе координат. Преобразование выполняется при необходимости совмещения разновременных или перекрывающихся изображений, а результирующие данные доступны для дальнейшей обработки.

Инструмент использует информацию из исходного файла HDF5 о типе геоида (поверхность, на которую проецируются радиолокационные данные), координаты контрольных точек, а также межпиксельный и межстрочный интервалы. По умолчанию приложением применяется модель геоида WGS84 (World Geodetic System – 1984) – мировая система геоцентрических координат, разработанная Военно-картографическим агентством Министерства обороны США [12]. Используемый системой ДЗЗ геоид ПО NEST 5.1 отображает в собственной переменной geo_ref_system (раздел Abstracted_Metadata).

На рис. 6 показано радиолокационное изображение, обработанное инструментом Reprojection в NEST 5.1 с наложением виртуальной сетки координат (опция Show Graticule Overlay). Визуально отмечено положение контрольных точек преобразованного РЛИ, демонстрирующих характер геометрических искажений, вносимых инструментом в исходные данные: переориентация и изменение формы изображения.

Для обработанных данных в NEST 5.1 доступен режим проекции на карту Земли NASA Blue Marble (опция Show World Map Overlay). Участок этой карты, содержащий спутниковое изображение береговой линии, разделяющий участок суши (серый цвет) с горным ландшафтом (белый цвет) и море (черный цвет), приведен на рис. 7а, и может быть использован для контроля геометрической обработки изображения, представленного на рис. 6 (рис. 7б).

Визуальный анализ рисунков 7а и 7б демонстрирует корректность работы инструмента и записи метаданных продукта в формат HDF5, поскольку результирующее изображение получает геопривязку и искажается таким образом, что очертания объектов исходной карты и проекции совпадают, это подтверждает правильное выполнение данной операции.

Заключение

Разработана модель представления набора метаданных и радиолокационных данных PCA системы ДЗЗ уровня 1А в стандартном формате HDF5. В качестве исходных данных использован продукт Single Look Complex Interferometric Wide системы ДЗЗ Sentinel-1 уровня 1А, рассмотрены особенности представления и

обработки радиолокационной информации в форматах TIFF и XML.

Дальнейшие исследования планируется направить на разработку универсального алгоритма представления метаданных и радиолокационных данных, обработанных до уровня представления 1А, в формат HDF5. Решение указанной задачи позволит обеспечить совместимость продукта разрабатываемой системы ДЗЗ на базе PCA с существующим свободно распространяемым программным обеспечением, а также предоставит возможность более эффективной обработки данных, полученных предшествующими системами ДЗЗ.

Литература

1. Радиолокационные системы землеобзора космического базирования / В.С. Вербя [и др.]. – М.: Радиотехника, 2010. – 680 с.
2. Лупян, Е. А. Базовые продукты обработки данных дистанционного зондирования Земли / Е.А. Лупян, В.П. Саворский // Современные проблемы дистанционного зондирования Земли из космоса. – 2012. – Т. 9, № 2. – С. 87–97.
3. Copernicus Open Access Hub [Электронный ресурс] – Режим доступа: <https://scihub.copernicus.eu/dhus/#/home> [дата обращения: 22.10.2019], свободный. – Загл. с экрана.
4. Sentinel-1 Product Specification [Электронный ресурс] – Режим доступа: <https://earth.esa.int/documents/247904/349449/Sentinel-1-Product-Specification.pdf> [дата обращения: 24.07.2019], свободный. – Загл. с экрана.
5. HDF5 Documentation [Электронный ресурс] – Режим доступа: <https://portal.hdfgroup.org/display/HDF5/HDF5> [дата обращения: 09.07.2019], свободный. – Загл. с экрана.
6. COSMO-Sky Med Web Site [Электронный ресурс] – Режим доступа: <http://www.cosmo-skymed.it/en/index.htm> [дата обращения: 28.06.2019], свободный. – Загл. с экрана.
7. NASA Earth Observing System [Электронный ресурс] – Режим доступа: <https://eosps.nasa.gov/> [дата обращения: 22.10.2019], свободный. – Загл. с экрана.
8. Адамовский, Е. Р. Особенности представления радиолокационных данных ДЗЗ в формате HDF5 / Е.Р. Адамовский, Р.П. Богуш // Материалы XXIV Международной научно-технической конференции «Современные средства связи». – Минск: Белорусская государственная академия связи, 2019. – С. 117–118.
9. COSMO-SkyMed Mission [Электронный ресурс] – Режим доступа: <http://www.cosmo-skymed.it/docs/ASI-CSM-ENG-RS-092-A-CSKSARProductsHandbook.pdf> [дата обращения: 19.07.2019], свободный. – Загл. с экрана.
10. Sentinel Application Platform [Электронный ресурс] – Режим доступа: <https://step.esa.int/main/toolboxes/snap/> [дата обращения: 12.08.2019], свободный. – Загл. с экрана.
11. ESA NEST 5.1 [Электронный ресурс] – Режим доступа: <https://earth.esa.int/web/nest/news/-/article/nest-5-1-released> [дата обращения: 19.07.2019], свободный. – Загл. с экрана.
12. Антонович, К. М. Использование спутниковых радионавигационных систем в геодезии. Монография / К.М. Антонович. – М.: ФГУП «Картгеоцентр», 2005. – 334 с.