The study and implementation of extraction HJ1A HSI image data based on A HDF5 files

Article *in* Proceedings of SPIE - The International Society for Optical Engineering · October 2009

DOI: 10.1117/12.833065

The Study and Implementation of Extraction HJ1A HSI Image Data Based on A HDF5 File

Yuanfeng Wu[*a,b], Bing Zhang[a], Junsheng Li[a], Hao Zhang[a], Qian Shen[a], Di Wu[a,b]
[a] Laboratory of Digital Earth Sciences, Center for Earth Observation and Digital Earth, Chinese Academy of Sciences, Beijing, 100039, China
[b] Graduate University of Chinese Academy of Sciences, Beijing, 100049, China

ABSTRACT

HJ1A HSI is an interferometric imaging spectrometer (Hyper-spectral Imager, sensor ID: HSI) of HJ-1 small satellite. The hyper-spectral image data are organized and stored in hierarchical data format version 5 (HDF5) files. This paper presents the data model, file structure, library and programming model of HDF5 file format. The adapter design pattern is used for translating hdf5 interface into a compatible interface. Then, we give a detailed analysis of HJ1A hyper-spectral image data. The HJ1A hyper-spectral image data model includes five groups: 'GlobalAttributes', 'ImageAttributes', 'ImageData', 'MapInformation', and 'ProductParameters' under the root group. The 'ImageData' group includes three datasets: 'BandData', 'CalibrationCoefficient', and 'WaveLength'. Based on the relationships between the models and implementations, we give a flow chart of extraction HJ1A hyper-spectral image data from hdf5 files. The level2 product of HJ1A hyper-spectral image data is used for experiment. We present the RGB color composite image and 3D cube of the extracted data. Tests show that the data extraction is correct and rapid with this approach. This work provides a solid foundation for quality evaluation and application of HJ1A hyper-spectral image.

Keywords: HJ1A, Small Satellite, HSI, Hyperspectral Image, HDF5, Interferometric Imaging spectrometer

1. INTRODUCTION

A Micro-satellite Constellation for Environment and Disaster Monitoring was successfully launched in China on September 6th, 2008, which includes two small satellites as Satellite-A (HJ1A) and Satellite-B (HJ1B). It is the first time for China to launch such kind of remote sensing satellites especially for the earth environment and disaster monitoring. The payloads of HJ1A include a CCD camera and an interferometric imaging spectrometer (Hyper-spectral Imager, sensor ID: HSI); the payloads of HJ1B include a CCD camera and an infrared scanner. The HSI installed on HJ1A is the first earth observation imaging spectrometer of China. Its spectrum range covers from 0.45μm to 0.95μm with 115 spectral bands. The average spectral resolution of the 115 bands is about 5 nm. The nominal ground sample distance is 100m with an image swath of about 60km. The hyperspectral image data of HJ1A has been defined by six levels, which are L0A, L0B, L1, L2, L3 and L4, according to different data processing flows and algorithms. These six levels image data are stored in hierarchical data format (HDF5) files. Extraction image data from hdf5 file is the basic step for HJ1A hyper-spectral image quality evaluation and

[*] yuanfengwu@126.com; phone +86 010-64807822

MIPPR 2009: Multispectral Image Acquisition and Processing, edited by
Jayaram K. Udupa, Nong Sang, Laszlo G. Nyul, Hengqing Tong, Proc. of SPIE Vol. 7494
74941L · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.833065

application. In this paper, we analyze the hdf5 file format, and give an object orientated method for extracting HJ1A hyper-spectral image data from a hdf5 file.

Table.1. Product classification and file format of HJ1A hyper-spectral image data

Product Level	File format	Data Processing Steps Description
L0A	HDF5	Interferogram data cube
L0B	HDF5	Spectrum reconstruction (Spectral image cube)
L1	HDF5	Spectrum reconstruction, Radiometric Calibration
L2	HDF5	Spectrum reconstruction, Radiometric Calibration, Geometric correction from satellite
L3	HDF5	Spectrum reconstruction, Radiometric Calibration, Geometric correction from satellite, Geometric precision correction using GCPs
L4	HDF5	Spectrum reconstruction, Radiometric Calibration, Geometric correction from satellite, Geometric precision correction using GCPs , Data product with DEM

2. THE HDF5 DATA MODEL

2.1 HDF5 Data Model

An HDF5 file is organized as a rooted, directed graph. The Named Data Objects are the nodes of the graph, and the links are the directed arcs. Each arc of the graph has a name. The root group has the name "/". Objects are created and then inserted into the graph with the link operation, which creates a named link from a Group to the object. An object can be the target of more than one link. An HDF5 group contains zero or more objects and every object must be a member of at least one group. The root group may not belong to any group. An HDF5 dataset is an object composed of a collection of data elements, or raw data, and metadata that stores a description of the data elements, data layout, and all other information necessary to write, read, and interpret the stored data. An HDF5 attribute is a small metadata object describing the nature and/or intended usage of a primary data object, which may be a dataset, group, or named data type.

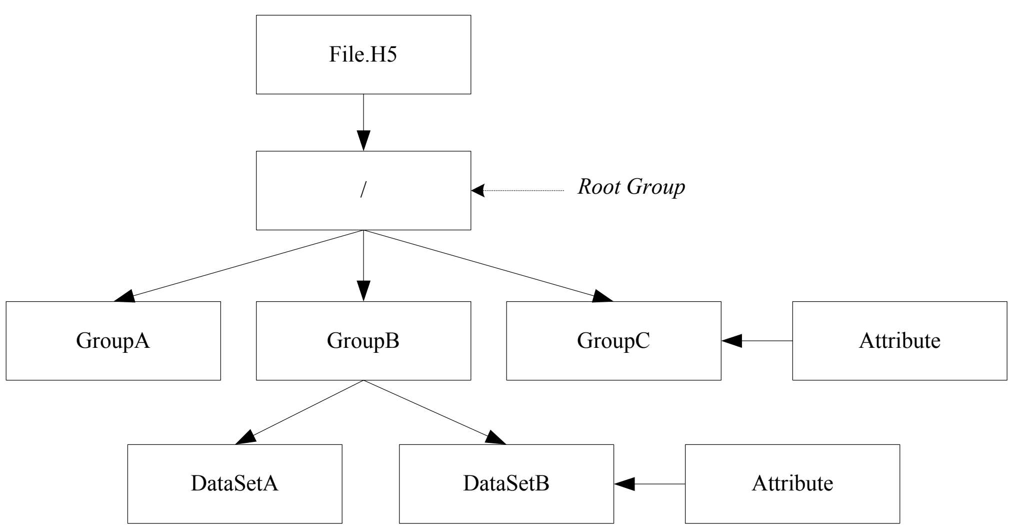

Fig.1. Data Model of HDF5 File

2.2 HDF5 Library API

The HDF5 library provides a programming interface to a concrete implementation of the abstract data models. The library exports a set of application programming interfaces, APIs, as its external interface. These APIs perform several categories of operations. HDF5 APIs have naming scheme, for example, APIs operates on files have prefix 'H5F'. APIs operates on groups have prefix 'H5G'. APIs operates on datasets have prefix 'H5D'. APIs operates on attributes have prefix 'H5A'. Table.2. lists useful APIs provided by HDF5 library.

Table.2. APIs provided by HDF5 Library

API	Description
H5Fopen	Opening an existing HDF5 file
H5Fclose	Closing a HDF5 file
H5Gopen	Opening a group
H5Gclose	Closing a group
H5Dcreate	Create a dataset at the specified location
H5Dopen	Open an existing dataset
H5Dclose	Close the specified dataset
H5Dget_offset	Returns dataset address in file
H5Dget_storage_size	Returns the amount of storage required for a dataset
H5Dread	Reads raw data from a dataset into a buffer
H5Dwrite	Writes raw data from a buffer to a dataset
H5Aopen_name	Opens an attribute specified by its name
H5Aread	Reads an attribute
H5Aclose	Closes the specified attribute

3. HJ1A HSI HYPER-SPECTAL IMAGE DATA EXTRACTION

3.1 Encapsulation of HDF5 Library APIs

We use the adapter design pattern (often referred to as the wrapper pattern or simply a wrapper) translates hdf5 interface for a class into a compatible interface. An adapter allows classes to work together that normally could not because of incompatible interfaces, by providing its interface to clients while using the original interface. The adapter translates calls to its interface into calls to the original interface, and the amount of code necessary to do this is typically small. The adapter is also responsible for transforming data into appropriate forms.

Furthermore, we use the facade pattern to provide a unified interface to hdf5 interfaces. Façade defines a higher-level interface that makes the hdf5 libraries much easier to use. A facade is an object that provides a simplified interface to a larger body of code, such as a class library. It can make the hdf5 libraries easier to use and understand, since the facade has convenient methods for common tasks. This also has some advantages as follows: make HJ1A hyper-spectral image data extraction code that uses the hdf5 libraries more readable; reduce dependencies of outside code on the inner workings of a library, since most code uses the facade, thus allowing more flexibility in developing the system; wrap a poorly-designed collection of APIs with a single well-designed API (as per task needs).

3.2 Analysis of HJ1A Hyper-spectral Image Data Format

The HJ1A hyper-spectral image data model has a root group named by HDF5 file name. The root group includes five groups: ‘GlobalAttributes’, ‘ImageAttributes’, ‘ImageData’, ‘MapInformation’, and ‘ProductParameters’. The ‘ImageData’ group includes three datasets: ‘BandData’, ‘CalibrationCoefficient’, and ‘WaveLength’.

Table.3. Data Model of HJ1A Hyper-spectral Image in HDF5 file format

(G: HDF5 Group, D: HDF5 Dataset, A: HDF5 Attribute)

(Root Group) FileName	**(G)**GlobalAttributes	**(A)**HDFVersion	
		(A) Provider	
		(A) SceneID	
		(A) ProductId	
		(A)ProductLevel	
		(A)InterleaveFormat	
		(A)ByteOrder	
		… …	
	(G) ImageAttributes	**(A)**SatelliteId	
		(A) SensorId	
		(A)SceneDate	
		(A)Bands	
		(A) ScenePath	
		(A) SceneRow	
		(A) SceneCenterTime	
		(A) SunAzimuthElevation	
		(A)ProductLines	
		(A)ProductSamples	
		… …	
	(G) ImageData	**(D)**BandData	**(A)**Units
			(A) ValidRange
			(A) FillValue
			(A) LongName
		(D) CalibrationCofficient	**(A)**Units
			(A) ValidRange
			(A) FillValue
			(A) LongName
			(A) GenerationDate
		(D) WaveLength	**(A)**Units
			(A) ValidRange
			(A) FillValue
			(A) LongName

	(G) MapInfomation	**(A)Projection**	
		(A) ProjectionParameters	
		(A)Utm-zone	
		(A)PixelSpacing	
		(A)Datum	
		(A)SceneCenterLat	
		(A)SceneCenterLong	
		… …	
	(G) ProductParameters	**(A)**ProductDate	
		(A) OperatorName	
		… …	

3.3 Data Extraction of HJ1A Hyper-spectral Image

HJ1A hyper-spectral image data extraction include four parts: get global attribute, get image attribute, get image data, and get map information. Fig.2. shows the detailed flow chart. First, get ‘productlevel’, ‘InterleaveFormat’, ‘ByteOrder’ attributes from ‘GlobalAttribute’ group. Then, get ‘Bands’, ‘productLines’, ‘productSamples’ attribus from ‘ImageAttribute’ group. These attributes are used for determine the memory space size for image data. The ‘BandData’ dataset is a [‘productSamples’, ‘productLines’, ‘Bands’] dimensions H5T_INTERGER data.

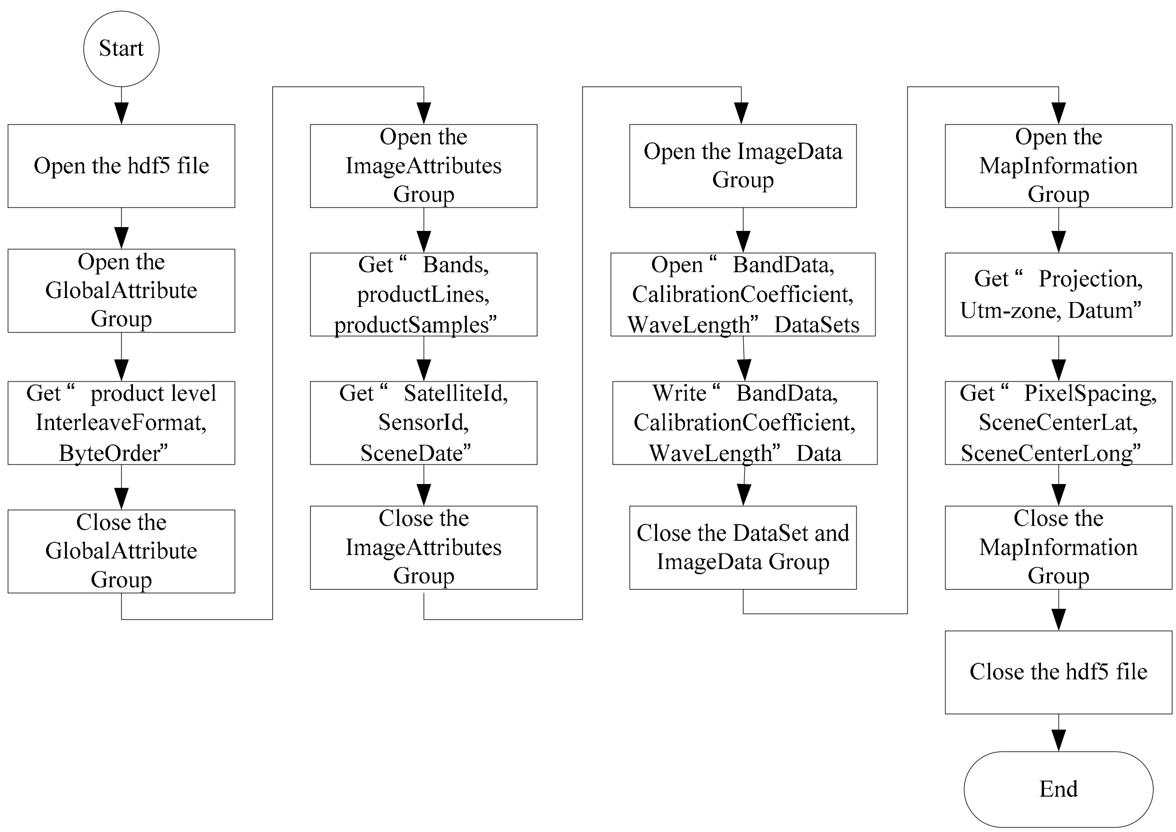

Fig.2. Flow chart of extraction HJ1A hyper-spectral image data from hdf5 file

The ‘CalibrationCoefficient’ dataset is a [‘Bands’, 2] dimensions H5T_FLOAT data. The ‘WaveLength’ dataset is a [‘Bands’, 3] dimensions H5T_FLOAT data which stores the wavelength of HJ1A hyper-spectral image.

We use a Level2 product data acquired by HJ1A HSI on Step 16th, 2008 for experiment. The ‘ProductLevel’ attribute is ‘LEVEL2’. The ‘InterleaveFormat’ attribute is ‘BSQ’. The ‘ByteOrder’ attribute is ‘Host(Intel)’. The ‘Bands’ is ‘115’. The ‘ProductLines’ attribute is ‘585’. The ‘ProductSamples’ attribute is ‘621’. The ‘BandData’ dataset is [621, 585, 115] dimensions data. The ‘CalibrationCoefficient’ dataset is [115, 2] dimensions data. The ‘WaveLength’ dataset is [115, 3] dimensions data. Fig.3. presents the RGB color composite image and 3D cube of the extracted data. The RGB face wavelength is R: 834.280nm, G: 668.620nm, B: 557.850nm.

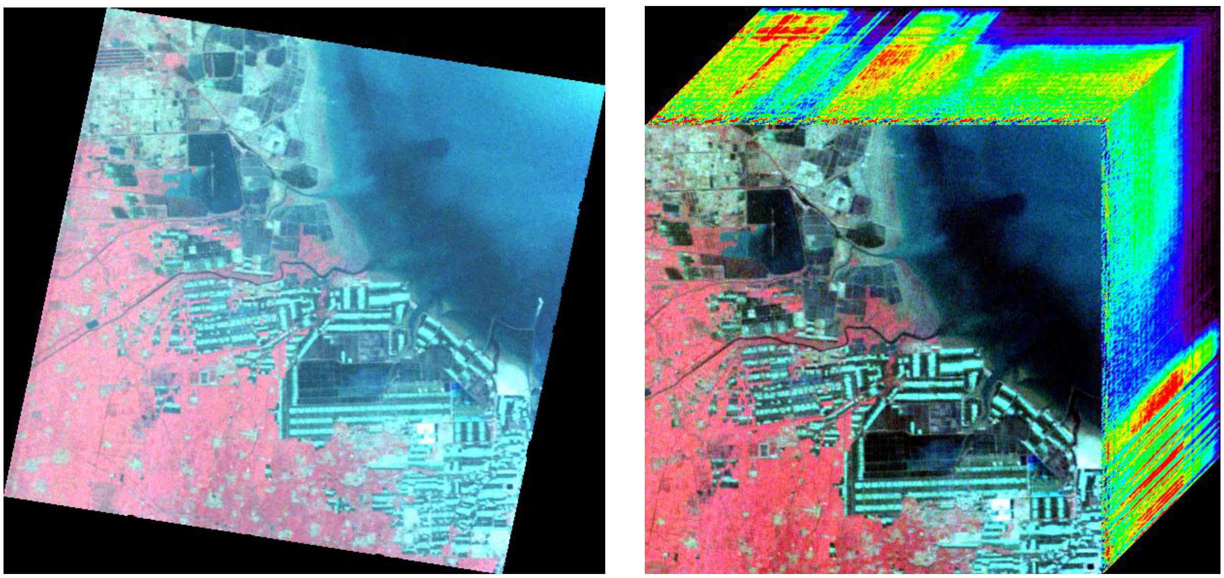

Fig.3. RGB color composite image (Left) and 3D cube (Right) of the extracted data from HDF5 file

4. CONCLUSIONS

HDF technologies are relevant when the data challenges being faced push the limits of what can be addressed by traditional database systems, XML documents, or other data formats. With the standardization of remote sensing data, HDF5 file format will be widely used in hyper-spectral image data storage. This paper analyzes the characteristics of HDF5 data model and useful HDF5 library APIs. Also, presents a detailed analysis of data format of HJ1A HSI hyper-spectral images. After that, we gave an object-oriented method for HJ1A hyper-spectral image data extraction with encapsulated HDF5 library APIs. Tests show that the data extraction is correct and rapid with this approach. This provides a solid foundation for HJ1A hyper-spectral image quality evaluation and application.

ACKNOWLEDGMENTS

This work has been supported by the civil spatial application projects of Commission of Science Technology and Industry for National Defense (Grant No. 5.3.1), the Key Projects of the Knowledge Innovation Program of Chinese Academy of Sciences, China (Grant No. KZCX1-YW-14-2), and the National Science & Technology Supporting Project (Grant No. 2008BAC34B05-2).

REFERENCES

1. http://www.hdfgroup.org/HDF5/Tutor.
2. http://gdal.org.
3. Erich Gamma, Richard Helm, Ralph Johnson, John Vlissides, et al. *Design Patterns-Elements of Reusable Object-Oriented Software*, China Machine Press, 2005.
4. Wu yuanfneg, Zhang bing, Li junsheng, et al. "A HJ-1 satellite imagery oriented software system for water environmental monitoring", International Conference on Earth Observation Data Processing and Analysis (ICEODPA). Proc. of SPIE Vol. 7285 72854W-1-7, 2008.
5. Junsheng Li, Bing Zhang, Qian Shen, Xia Zhang, Zhengchao Chen. "Future Generation of Inland Water Environmental Monitoring Missions". The 3rd International Symposium on Future Intelligent Earth Observing Satellites (FIEOS), 2006.
6. Yuanfeng Wu, Bing Zhang, Junsheng Li, et al. "A first approach to automatic monitoring procedure of case II water quality from HJ-1 satellite images". IEEE international Geoscience and Remote Sensing Symposium (IGARSS), 2009.
7. Kruse F A, et al. The Spectral Image Processing System (SIPS)- Interactive Visualization and Analysis of Imaging Spectrometer Data, Remote Sensing of Environment, 1993, 44:145-163.
8. Mazer A S, Martin M, Lee M, Solomon J E. Image Processing Software for Imaging Spectrometry Analysis, Remote Sensing of Environment, 1998, 24(1):201-210.
9. Zhang B, Wang X J, Liu J G, et al. "Hyperspectral Image-Processing and Analysis System and Its Application," Photogrammetric Engineering and Remote Sensing, 66(5):605-609, 2000.