Korean Journal of Remote Sensing, Vol.18, No.1, 2002, pp.61~69

Development of HDF Browser for the Utilization of EOC Imagery

Hee-Kyung Seo*, Seok-Beom Ahn*, Eun-Chul Park*, Kwang-Soo Hahn*, Joon-Soo Choi*, and Choen Kim**

School of Computer Science*, Department of Forest Resources Kookmin University**

Abstract : The purpose of Electro-Optical Camera (EOC), the primary payload of KOMPSAT-1, is to collect high resolution visible imagery of the Earth including Korean Peninsula. EOC images will be distributed to the public or many user groups including government, public corporations, academic or research institutes. KARI will offer the online service to the users through internet. Some application, e.g., generation of Digital Elevation Model (DEM), needs a secondary data such as satellite ephemeris data, attitude data to process the EOC imagery. EOC imagery with these ancillary information will be distributed in a file of Hierarchical Data Format (HDF) file format. HDF is a physical file format that allows storage of many different types of scientific data including images, multidimensional data arrays, record oriented data, and point data. By the lack of public domain softwares supporting HDF file format, many public users may not access EOC data without difficulty. The purpose of this research is to develop a browsing system of EOC data for the general users not only for scientists who are the main users of HDF. The system is PC-based and has user-friendly interface.

Key Words : KOMPSAT-1, EOC, HDF, file format, Image Processing System.

1. Introduction

KOMPSAT-1 is an earth observing satellite for practical use mainly in remote sensing area which includes the cartography of Korea peninsula, ocean research, environment monitoring, space research, etc. Electro-Optical Camera (EOC) is the main payload of KOMPSAT-1 with the mission of cartography to build up a digital map of Korean territory including Digital Terrain Elevation Map (Choi *et al.*, 1999). EOC is a high spatial resolution, visible imaging sensor which collects visible image data of the earth's sunlit surface. EOC collects image data from sunlit earth scenes in broadband visible wavelength. The collected data along with ancillary spacecraft position, attitude, and the imaging time will be processed by the KOMPSAT ground station to produce standard and value-added image products such as cartographic maps. Korea Aerospace Research Institute (KARI) chose Hierarchical Data Format (HDF) to be the standard file format for storing and distributing EOC data.

The National Center for Supercomputing

Received 5 January 2002; Accepted 4 February 2002.

Applications (NCSA) at the University of Illinois developed HDF as a multi-object file format (NCSA, 1999). It was created to help scientists reduce the time they were spending trying to convert data sets to familiar formats and instead have more time available for actually analyzing data. HDF is a physical file format that allows storage of many different types of scientific data including images, multidimensional data arrays, record oriented data, and point data. HDF was designed to be a machine-independent, self-describing, extendible file format for sharing scientific data in a heterogeneous computing environment, accompanied by a convenient, standardized, public domain I/O library and a comprehensive collection of high quality data manipulation and analysis interfaces and tools.

HDF receives substantial support from its user community in the form of collaborative development, software contributions, and some funding. The community of HDF users consists primarily of scientists, supercomputer users, and developers of advanced scientific visualization software. Disciplines in which HDF is used heavily include the earth sciences, astronomy, biology, materials sciences, chemistry and physics. Partly because of NASA's commitment, many worldwide organizations have also adopted HDF as a common data format and library. Some examples are Landsat 7, NASA EOSDIS, space agencies in many countries, university computer centers, several national laboratories, and a number of private companies.

EOC data were made available for public use in April 2000. These data sets will be produced by KARI and distributed to scientists, academia and the general public over the next few years. HDF file format is selected for the standard data format of EOC data. There are a few software packages which allow the viewing of HDF data files. Many are public domain programs, developed NCSA, NASA Centers, universities. These are generally free to non-commercial users. Other HDF viewing software was developed by commercial software companies. These packages are quite expensive and are not available to the public. The usual public domain softwares are lack of some useful functions.

The purpose of this work is to develop a software package browsing HDF file so that the public users can access EOC data without difficulty. The main features of this software are to read/write EOC images, view the images, and export to many other standard image file formats. One of the important goal of this software is to make EOC data readily available and usable to the public so that it may contribute to a broad understanding of the earth observation by KOMPSAT-1.

2. EOC Data

The Electro-Optical Camera (EOC), the primary payload of KOMPSAT-1, is a high spatial resolution, visible imaging sensor which collects visible image data of the earth's sunlit surface. EOC collects panchromatic imagery with the ground sample distance of 6.6m and the swath of 17km at nadir through the visible spectral band of 510nm ~ 730 nm. EOC scans the ground track of 800 km per orbit by push-broom and body pointing method. The main usage of EOC image is the making of Digital Elevation Model (DEM) data of entire Korean peninsula. The cross-track rolling ability of KOMPSAT-1, as large as 45 deg., enable EOC to take stereo pictures. DEM data can be produced through processing the images by image processing software packages (Ye *et al.*, 2000). This DEM could be used as a basic material form of Geographic Information System (GIS), and also the DEM itself is useful for land development program and military purposes.

The KOMPSAT Receiving and Processing System (KRPS) in the KOMPSAT ground station accepts the real-time data stream, formats the data and stores the data on a data backup system (Kim *et al.*, 1999). The data can be archived or processed immediately by the

Table 1. Data product level definitions.

Prod. Level	Description
Level 0	Frame formatted, unprocessed instrument/payload data at full resolution
Level 1A	Unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information
Level 1R	Level 1A data that have been radiometrically corrected
Level 1GR	Level 1R data that have been geometically corrected and geo-refenced
Level 1GC	Level 1R data that have been geometically corrected and geo-coded

Table 2. Ancillary parameters of EOC products.

Groups	Parameters
General Parameters	Satellite Name, Instrument Name, Scene ID, GRS Designator, Acquisition Data & Time, Processing Level, Production Date, Number of Pixels per Line, Number of Lines per Image, Inter-pixel Distance in Meters, Number of Bits per Pixel, Latitude & Longitude of Scene, Image Coordinate of Scene, etc.
General Imaging Data	Image Orientation Angle, Angle of Incidence, Sun Angle, Satellite Altitude, Detector Signal Integration Time
Sensor Modeling Parameters	Satellite Position / Velocity Data, Scene Center Time, Transformation Data, Sensor Look Vectors of EndDetectors, Detector Gain and Bias Data
Orbital Elements	Field of View, Viewing Angle, Radial Speed, Eccentricity, Height of Satellite, Inclination, Angular Speed, Earth Satellite Distance, etc.

Data Processing Facility (DPF) where standard image products are generated. DPF performs the radiometic and geometric correction on the received image data, and performs several image processing steps with ancillary data. Spacecraft provides ancillary data such as image time, position and attitude of spacecraft for EOC image processing.

The KOMPSAT data product level definitions are based on the standard of the Committee on Earth Observation Satellites which is identical to NASA EOSDIS data convention. The following table lists the production levels of EOC products.

Level-0 data is a local-area coverage image data and it is stored as one byte per pixel. Level-1A products contain all the Level-0 data and appended ancillary information. The ancillary information annotated in Level 1A includes radiometric and geometric calibration coefficients and georeferencing parameters. Some ancillary parameters stored in Level-1A products are listed in the following table:

Each product of Level-1A to Level-1GC is stored as one physical HDF file. The usual size of each product is 4–11M bytes.

3. HDF File Format

HDF is a file format developed at NCSA (National Center for Supercomputing Applications) at the University of Illinois. It was created to aid scientists and programmers in the storing, transfer and distribution of data sets and products created on various machines and with different software. HDF is a multi-object file format for the sharing and storing of scientific data. Some of the most important features of HDF are the following (NCSA 1999):

1) Self-describing: For each data object in an HDF file, there is also information (or metadata) about the data type, size, dimensions and location found within the file itself.

2) Extensibility: HDF is designed to accommodate future (new) data types and data models.

3) Versatility: Currently, HDF supports six different data types and provides software and applications to read and write these data types in HDF. These types are commonly used by scientists.

4) Flexibility: HDF lets the user group, store, and read/write different data types in the same file or in more than one file.

5) Portability: HDF software is mainly platform independent and can be shared across most computer platforms (all platforms have not been tested).

6) Standardization: HDF standardizes the formats and descriptions.

7) HDF is available in the public domain.

As of the latest release, the primary data types, called HDF data structures, that HDF library supports are (NCSA 1999):

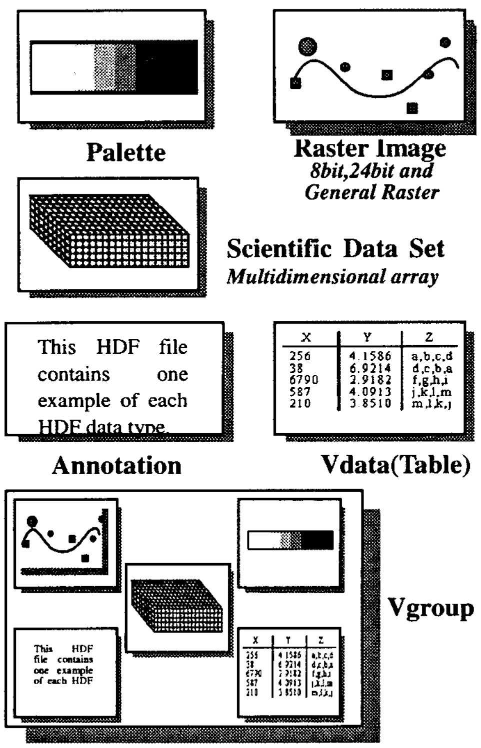

Fig. 1. HDF data structures.

1) Raster image sets: 8-bit and 24-bit raster images and optional color palettes. Three forms of image compression are supported: JPEG, run-length encoding, and IMCOMP compression.

2) Palettes : Color tables for 8-bit raster image data.

3) Scientific data sets: Multidimensional integer/floating point arrays with specific attribute data. The latest version supports 8, 16, and 32-bit integers and 64-bit floating point data.

4) Annotations: Labels and unstructured text associated with any HDF object or globally to the HDF file as a whole.

5) Vdata (Vertex data): Multivariate data such as meshes, polygonal data with connection information, instrument data, and sparse matrices stored as records in a table.

6) Vgroups : A logical grouping of data objects (e.g., scientific data sets, Vdatas, annotations, rasters, palettes, and another Vgroups). The general structure of Vgroup is similar to that of UNIX file system.

The HDF data structures are illustrated in Fig. 1. Many types of data can be included within an HDF file. For example, it is possible to store symbolic, numerical and graphical data within an HDF file by using appropriate HDF data structures.

In addition, HDF also refers to the collection of software, application interfaces, and utilities that comprise the HDF library and allows users to work with HDF files. HDF library contains interfaces for storing and retrieving compressed or uncompressed raster images with palettes; an interface for storing and retrieving n-Dimensional scientific datasets together with information about the data, such as labels, units, formats, and scales for all dimensions.

HDF has been selected by the NASA ESDIS project as the format of choice for the standard product distribution that will be part of the Earth Observing System Data and Information System (EOSDIS) (NASA, 1999). Partly because of NASA's commitment, many worldwide organizations have also adopted HDF as a common data format and library. Some examples are Landsat 7, space agencies in many countries, university computer centers, several national laboratories, and a number of private companies.

Graphical file formats like HDF can be categorized into one of the three categories (Brown and Shepherd 1995). The formats are either "official" standards, de facto standards because of their widespread use, or simply a format that is widely (or not so widely) used. Official standards include JPEG, MPEG file formats, de facto standards include DXF, GIF, TIFF file formats. Other formats including BMP, HDF are not categorized into a standard format. A standard is needed when a lot of people use it. Because HDF is created to support a wide range of scientific data by a group of scientists using supercomputers, it is not included in one of the standard categories yet.

There are several public domain software packages available to view HDF files on local machines. Specially, NCSA released JHV (Java HDF Viewer) as an addition to the latest version of HDF (NCSA, 2000). Some other packages developed at NCSA and some packages released by private companies can browse HDF files. There are a few commercial tools available to read HDF files. Many of them are digital image processing and remote sensing system to help earth scientists integrate, enhance, visualize, and interpret their geographic data. Descriptions of these programs are generally provided at the developer's web sites. One of the most comprehensive web site is the home page HDF file format, http://hdf.ncsa.uiuc.edu, developed by NCSA. The latest version of HDF is 4.3. HDF 5 is another version of HDF and it is designed to address some of the deficiencies of HDF 4.x particularly the need to store large files and more objects per file. But HDF 5 is not compatible with HDF 4.x.

Some commercial software packages that can browse HDF files are quite expensive and are not available to the public. The usual public domain softwares are lack of some useful functions such as saving an EOC image to popular image file formats like JPEG or BMP. The purpose of this work is to develop a software package browsing HDF file so that the public users can access EOC data without difficulty. The features included in this software are saving EOC images to other file formats, image-processing utilities like image enhancement, drawing geometric feature like lines, circles, texts and overlapping the features with images, etc.

4. HDF Browser for EOC Imagery

The Hierarchical Data Format (HDF) is a library and platform independent data format for the storage and exchange of scientific data. HDF files may contain many types of information that a scientist might need. Here we developed and HDF Browser for EOC imagery. The file to store EOC imagery uses many data structures supported by HDF format except scientific data sets. Therefore, we developed a browser to visualize almost all data structures supported by HDF. In the following, we explain the features included in the browser we developed to visualize the data structures of HDF format. We developed the browser based on the latest version HDF 4.3.

The HDF browser displays the structure of an HDF file, showing the objects and their groupings. The browser also displays a description of each object in the file. The browser allows users to browse through any HDF file. When opening a new HDF file, it shows a view of all top-level objects in an HDF file's hierarchy, and allows users to descend through the hierarchy and navigate among the file's data objects. The content of a data object is loaded only when the object is selected, providing interactive and efficient access to HDF files.

The user interface is designed to be easy to use and intuitive. The internal structure of HDF file is similar to that of the UNIX file system in that a group of HDF objects may contain references to other groups or HDF data objects just as the UNIX directory may contain subdirectories or files. Therefore it is natural to use similar user interface of the windows systems for PC's or workstations in which the hierarchy of files and directories are used for the file system. The structure of the file is displayed in the HDF tree window. Data objects are represented as icons, groups of objects are represented by folders. Users can easily expand or collapse folders to navigate the hierarchical structure of an HDF file as users navigate the file structures in a UNIX file system or Windows system in PC. Each object may be selected to view the data it contains.

The browser can be used to view several types of HDF data objects. The information in each object is displayed appropriately. In the following, we describe the HDF data objects and the way of HDF browser displaying the related information.

Vgroups An HDF vgroup object is a structure designed to associate related objects. The general structure of vgroups is similar to that of the UNIX file system. Any HDF object can be included within a Vgroup. There are many ways to organize vgroups through the use of the Vgroup interface. Vgroups may contain any number of vgroups and data objects, including data objects and vgroups that are members of other vgroups.

A Vgroup data object is usually a Vdata or an SDS data set. It also can be a raster image. Clicking on a Vgroup object node opens the Vgroup object in the appropriate Vdata, SDS and image mode. The HDF browser shows each Vgroup as a "folder" icon. When the Vgroup folder is selected, the objects in the group are displayed. When the folder is closed, the objects in the group are not shown. In this way, the hierarchical structure of the HDF file can be navigated. Fig. 2 shows the HDF tree window for a file, with some of the folders open, and several objects visible. In the figure, an image object has been selected, and the image is displayed.

1) Annotations

HDF annotation data structure is designed to accommodate a wide variety of information including titles, comments, variable names, parameters, formulas, and source code. In fact, HDF annotations can encompass any textual information regarding the

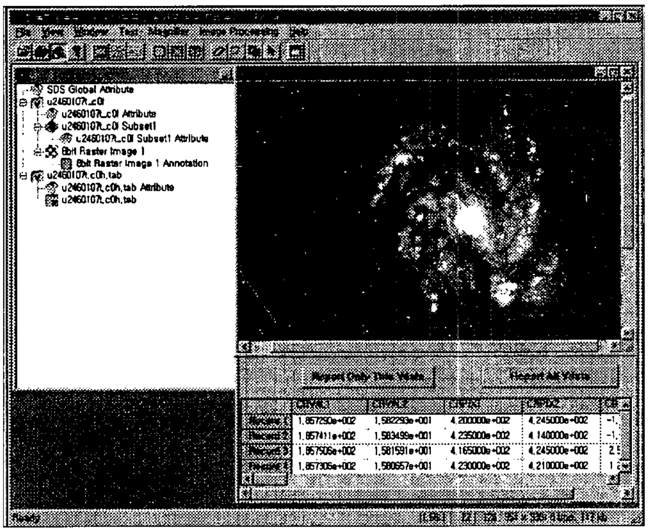

Fig. 2. Subwindows viewing hierarchical structure of HDF objects, a general raster image and a Vdata object.

collection, meaning, or intended use of the data. Annotations come in two forms; labels and descriptions. Labels are short annotations used for assigning things like titles or time stamps to a file or its data objects. Descriptions are longer annotations; they typically may contain more extensive information, such as a source code module or mathematical formulae.

There are also two types of annotations: file annotations and data object annotations. A file annotation describes an entire HDF file. A data object annotation describes a particular data object within the file. File and data object annotations can each be either labels or descriptions. Thus an annotation might be a file label, a file description, an object label, or an object description.

Although an attribute set is associated to its data, the attribute set is displayed as a separate object in the HDF tree. The browser displays attribute sets in the same way as annotations. The browser shows annotations attached to the file or object. When an annotation is selected to view, the browser displays the text of the annotation in the subwindow.

2) Raster Image Sets

The general raster, or GR, data model is designed to provide a flexible means of storing raster image data, including 8-bit raster or 24-bit raster images. Raster image data is stored in a two-dimensional array and image array dimensions specify the size of the image array. Each image array has a name consisting of a string of characters and each element in an image array corresponds to one pixel and each pixel can consist of a number of color component values. Attributes may be associated with the image, the file, or both. Palettes (also known as color look-up tables) and compression method information can also be associated with an image.

The image-displaying window provides the tools for processing images and for saving data. When a GR image is selected to be viewed, the browser will display the full size image in the subwindow. An image can be enlarged or contracted by "zooming", which decreases or increases the subsampling. The zoom-in selection displays a selected part the image at higher resolution and zoom-out displays a larger area of the image with lower resolution. The zooming scale factor rangers from 16:1 to 1:16, which can be changed with the mouse or menu. To examine only a small part of the image at greater resolution, users may use the "magnifier" tool. The tool shows the region of interest which the user selected in a subwindow with a predefined resolution. The subwindow display the regions centered by the mouse as the user moves mouse around the original image. HDF image can be exported to other graphics file format, e.g., BMP, TIFF and JPEG format for use in other analysis program, or sent directly to a printer.

There are several utility tools to help users to manipulate images. For example, there is a tool to determine the length of a linear feature within the raster image data. The distance tool supports both single and multiple-segment lines. The area tool allows users to determine the area of a simple polygonal figure within the raster image. Users can draw geometric shapes, e.g., rectangles, ellipses on an image, where the width or color of the shape borders can be changed. Text tools can be used to add texts to an image. Text is created in an editing window and added to the image as a new floating selection, which can be moved around. Users can change the fonts and colors of text.

3) Palette

A palette is the means by which color is applied to an image and is also referred to as a color lookup table. It is a table in which every row contains the numerical representation of a particular color. Palettes can be many different sizes, but HDF only supports palettes with 256 colors, corresponding to the 256 different possible pixel values (0 to 255) in 8-bit raster images. In the browser, the information in the palette is used to display the color image. Users may not change the information in the palette.

4) Scientific Data Sets

The scientific data set, or SDS, is a group of data structures used to store and describe multidimensional arrays of scientific data. The HDF scientific data set model has four primary data objects: arrays, dimensions, dimension scales, and dimension attributes. The fundamental object of the model is the SDS array; the remaining objects describe the array. An SDS array is a multidimensional data structure and is the primary data component of the SDS model.

The browser displays the data of an SDS array in form similar to a spreadsheet which shows a two-way table of values in the table. The arrangement of the data is controlled by the selection panel. A one dimensional array is displayed as a single row. A two dimensional array is displayed as a two-way table, and the user may select which dimension should be the columns and which should be the rows. Higher than two dimensions can also be displayed. Three dimensional data can be viewed as a series of two dimensional arrays tables. The user may select which dimension is the columns, which is the rows, and which is the 'depth'. Each 2D array is a slice of the 3D data taken along an axis selected by the user as a depth. For arrays with more than three dimensions, any three dimensions may be selected to be displayed. These 2D array can be change by controlling the value along the user selected axis. The selection panel also allows the selection of a range of rows and columns to be displayed. This is useful for viewing a part of a very large array.

5) Vdata

The HDF Vdata model provides a framework for storing customized tables, or vdatas, in HDF files. The term "vdata" is an abbreviation of "vertex data", alluding to the fact that the object was first implemented in HDF to store the vertex and edge information of polygon sets. The vdata design has since been generalized to apply to a broader variety of applications. A vdata is like a table that consists of a collection of records whose values are stored in fixed-length fields. All records have the same structure and all values in each field have the same data type. The information associated with Vdata includes the number of records, and the names, types of the fields (columns) of the table, record size, etc.

The browser displays the Vdata in a subwindow as a table with columns and rows representing the Vdata fields and records respectively. The Vdata display shows the title (field name) of each column. A Vdata subwindow is created when a Vdata object is selected from in the HDF hierarchy subwindow. The whole table can be exported to a text file in a certain form. Currently, each record can be written one by one in a way that each field data of the record is printed in a separate line. We can also export the table in a two-dimensional form as it is displayed in the subwindow.

The browser is developed for PC platforms running Windows 98/2000/NT. Microsoft Visual C^{++} 6.0 and MFC library are used to build the entire software including user-interface. Precompiled HDF library of Release-4.1r3 from NCSA is used for the interface to access objects and associated information in an HDF file.

5. Conclusions

Korea Aerospace Research Institute (KARI) selected the Hierarchical Data Format (HDF) as the common data format of choice for standard product exchange and distribution of EOC imagery. HDF is a physical file format that allows storage of many different types of scientific data including images, multidimensional data arrays, record oriented data, and point data. In this paper, we develop an HDF file browser viewing several types of HDF data objects. The information in each object is displayed appropriately. The user interface of the browser is designed to be easy to use and intuitive so that the public users can access the EOC data without any difficulty.

Acknowledgements

This work was supported by the grant from Korea Aerospace Research Institute and Ministry of Science and Technology of the Republic of Korea.

References

C.W. Brown, and B.J. Shepherd, 1995. *Graphics File Format: reference and guide*, Manning Pub. Co., Greenwich, CT, USA.

G.H. Choi, J.H. Lee, and H.Y. Paik, 1999. Establishing application system of KOMPSAT-1, *J. Korean Soc. of Remote Sensing*, 15(4): 349-356.

Y.S. Kim, Y.S. Kim, H.S. Lin, D.H. Lee, and C.H. Kang, 1999. KOMPSAT data processing system: An overview and preliminary acceptance test results, *J. Korean Soc. of Remote Sensing*, 15(4): 357-365.

NASA, 1999. Earth Observing System (EOS) Data and Information System (EOSDIS).

NCSA, 1999. HDF User's Guide.

NCSA, 2000. Java-based HDF Viewer (JHV).

C.S.Ye, B.M. Jeon, and K.H. Lee, 2000. A wide DEM generation based on orthorectification and DEM data fusion, *J. Korean Soc. of Remote Sensing*, 16(1): 99-108.