

Standards Change Request

JPEG 2000 as a PDS Archive Format

SCR 3-1003

Steve Hughes (submitted)

Aug. 20, 2003

Joel Wilf (modified)

Apr. 19, 2004

Elizabeth D. Rye (modified)

Sept. 27, 2005

Problem:

The MRO HiRISE (High Resolution Imaging Science Experiment) team is currently predicting the production of 92,500 Gb of RDR products. The maximum sizes of individual RDR products are expected to be about 5,120MB. This large volume of data and the large size of the individual products raises several data management issues involving data storage and the distribution of products over the internet.

The MRO HiRISE team is suggesting that the JPEG 2000 compression scheme be used to compress the HiRISE products for use in both online data repositories and for the PDS archive. The JPEG 2000 compression scheme provides a good compression ratio (~3:1 to 4:1 lossless) and uses a wavelet encoding scheme to allow easy viewing of extremely large images. The use of the JPEG 2000 compression scheme will (in the words of the team) result in lower data volume requirements, smaller storage needs, lower media costs, more easily managed data sets, lower per-unit count of hard media, and faster transfer times. It will also preclude the need for product tiling.

Note that the MRO HiRISE team is proposing JPEG 2000 for HiRISE RDRs only.

Current Urgency:

The MRO HiRISE team as well as the PDS are currently developing plans that include hardware and software procurement, operations center design, product distribution mechanisms, and archive media production. The volume of data and the size of the individual products are the design variables with the most impact at this time. The possible use of JPEG 2000 directly effects these two variables.

Since JPEG 2000 will most likely be used in the online repository in any case the decision on whether or not the PDS will allow JPEG 2000 in the archive impacts the proposed interface and must be made as soon as possible.

Proposed Solution:

Permit the JPEG 2000 compression algorithm to be used for PDS archive products, consistent with PDS policy on compression.

Permit PDS data products to be archived in the JPEG 2000 (or “JP2”) format, consistent with PDS policy on the use of external data formats.

Requested Changes:

1. Amend the Standards Reference to contain a new section describing JPEG 2000 compression. The proposed section is attached.
2. Modify Section 10.2.3 of the Standards Reference, “Reserved Extensions”, to include “J2C – raw JPEG 2000 codestream” and “JP2 – JPEG 2000 (JP2) formatted image”. Update “IMQ” to say “Image data that have been compressed. (Not for use with JPEG 2000 compressed data.)”
3. Add the new keyword, ENCODING_TYPE_VERSION_ID, to the Planetary Science Data Dictionary. The element object definition for this keyword is attached.

Impact Assessment:

1. The PDS Standards Reference will need to be updated as described above.
2. The Object Access Library will need to be modified to include freely available algorithms for the decompression of JPEG 2000 formatted data.

Additional Information:

The JPEG 2000 format is defined in the specification ISO/IEC 15444. Part one of this specification is the part most relevant to consideration of the format for inclusion in PDS archives. This specification may be ordered directly from the ISO by going to the following URL:

<http://www.iso.org/iso/en/StandardsQueryFormHandler.StandardsQueryFormHandler?scope=CATALOGUE&sortOrder=ISO&committee=ALL&isoDocType=ALL&title=true&keyword=15444>

I.3 JPEG 2000

JPEG 2000 is defined as an “image coding system”. The ISO/IEC specification describing it includes not only a syntax for a compressed image codestream, but also a description of the binary “JP2” file format that may be used to wrap the codestream. The bare compressed codestream alone has been assigned a mime type of “J2C”, and thus will be so referred to in the ensuing paragraphs.

Unlike many older compression algorithms, JPEG 2000 provides a great deal of flexibility in the way in which data may be stored in the codestream and retrieved from it. This flexibility allows for the progressive decompression of “layers” of the image with increasing resolution or precision. It also permits the extraction and decompression of only a portion or “tile” of the image. Specific portions of the image of particular interest to the intended audience may also be stored at the beginning of the codestream so that they may be accessed and decompressed first. (This would be of potential interest for approach images where the target of the observation fills only a small portion of the field of view.)

All of the information necessary to successfully decompress a JPEG 2000 image is contained in the J2C codestream. However, the information necessary to take advantage of the additional capabilities of the format is only available with the inclusion of the JP2 header.

The JP2 file format essentially consists of a set of “boxes” that encapsulate both the J2C codestream as well as the meta data that describe it. The first two of these boxes must provide the information that identifies the file as a JP2 formatted file. The following “superbox” is the JP2 header box which contains information about the image size, resolution, colorspace, etc. Following this, in no particular order, are contiguous codestream boxes containing the compressed image data, and optionally, intellectual property rights boxes, XML boxes containing vendor-specific meta data, and UUID boxes containing reference URLs.

Note that the PDS permits the presence of the JP2 binary wrapper so that external software may be utilized to take advantage of the full capabilities of the JPEG 2000 format to decompress only portions of an image. However, the PDS is only obligated to include in the archive algorithms and source code to fully decompress the entire data file.

I.3.1 Table of Compression Ratios

“Busy” Image	“Smooth” Image	ASCII Table
TBD	TBD	N/A

I.3.2 PDS Implementation Rules

The JPEG 2000 compression algorithm may be implemented on PDS archive volumes in one of two ways: 1) as a bare codestream, or 2) as a compressed codestream wrapped by the binary JP2 file format.

In either case, only lossless compression may be used. Furthermore, the compressed codestream must conform to part 1, the “Core coding system”, of the ISO/IEC specification defining JPEG 2000, namely 15444-1.

I.3.3 Labeling

A PDS label used to describe a JPEG 2000 file is describing a whole file compressed as a single entity, rather than a file containing a collection of compressed objects. The label must therefore follow the labeling convention approved for Zip compressed files. This method uses a combined detached label containing two explicit FILE objects, one describing the compressed file and one describing the dynamically generated uncompressed file. (For more details on the specifics of this method, please see the section of this document describing Zip compression.)

I.3.4 Label Example

The following combined detached label describes a JP2 formatted image and the decompressed PDS formatted image derived from it:

```

PDS_VERSION_ID           = PDS3

/* IDENTIFICATION DATA ELEMENTS */

MISSION_NAME              = "MARS RECONNAISSANCE ORBITER"
INSTRUMENT_HOST_NAME     = "MARS RECONNAISSANCE ORBITER"
INSTRUMENT_NAME          = "HIGH RESOLUTION IMAGING SCIENCE
                           EXPERIMENT"
TARGET_NAME              = "MOON"
DATA_SET_ID              = "MRO-M-HIRISE-5-DIM-V1.0"
PRODUCT_ID               = "CRU_000004_1200_RED2_1"
START_TIME               = 2005-09-08T23:16:44.863
STOP_TIME                = 2005-09-08T23:16:51.569
SPACECRAFT_CLOCK_START_COUNT = 810688604:56542
SPACECRAFT_CLOCK_STOP_COUNT  = 810688611:37300
PRODUCT_CREATION_TIME    = 2005-09-09T15:35:45
(etc.)

/* DESCRIPTIVE DATA ELEMENTS */

(etc.)

```

```

OBJECT                = COMPRESSED_FILE
  FILE_NAME           = "filename.jp2"
  RECORD_TYPE         = UNDEFINED
  ENCODING_TYPE       = "JP2"
  ENCODING_TYPE_VERSION_ID = "ISO/IEC15444-1:2004"
  INTERCHANGE_FORMAT  = BINARY
  UNCOMPRESSED_FILE_NAME = "filename.img"
  REQUIRED_STORAGE_BYTES = 2400000000
  ^DESCRIPTION        = "jp2_description.txt"
END_OBJECT            = COMPRESSED_FILE

OBJECT                = UNCOMPRESSED_FILE
  FILE_NAME           = "filename.img"
  RECORD_TYPE         = FIXED_LENGTH
  RECORD_BYTES        = 40000
  FILE_RECORDS        = 60000

/* POINTERS TO DATA OBJECTS */

  ^IMAGE              = "filename.img"

/* DATA OBJECT DEFINITIONS */

OBJECT                = IMAGE
  LINES               = 60000
  LINE_SAMPLES        = 40000
  SAMPLE_TYPE         = UNSIGNED_INTEGER
  SAMPLE_BITS         = 16
  SAMPLE_BIT_MASK     = 2#0011111111111111#
  (etc.)
END_OBJECT            = IMAGE
END_OBJECT            = UNCOMPRESSED_FILE
END

```

PDS_VERSION_ID = PDS3
LABEL_REVISION_NOTE = "Submitted as part of SCR3-1003
(JPEG 2000 as PDS Archive Format) 2005-09-25 EN:E. Rye"

OBJECT = ELEMENT_DEFINITION
ELEMENT_NAME = ENCODING_TYPE_VERSION_ID
BL_NAME = encypeverid
DESCRIPTION = "

The ENCODING_TYPE_VERSION_ID element indicates the version of a standard or specification with which a particular ENCODING_TYPE complies."

GENERAL_DATA_TYPE = CHARACTER
MAXIMUM = "N/A"
MINIMUM = "N/A"
MAXIMUM_LENGTH = 32
MINIMUM_LENGTH = "N/A"
STANDARD_VALUE_TYPE = TEXT
STANDARD_VALUE_SET = "N/A"
STANDARD_VALUE_SET_DESC = "N/A"
KEYWORD_DEFAULT_VALUE = "N/A"
UNIT_ID = "none"
SOURCE_NAME = "PDS EN/E. Rye"
FORMATION_RULE_DESC = "N/A"
SYSTEM_CLASSIFICATION_ID = "COMMON"
GENERAL_CLASSIFICATION_TYPE = "SOFTWARE"
CHANGE_DATE = "2005-09-25"
STATUS_TYPE = "PROPOSED"
STANDARD_VALUE_OUTPUT_FLAG = "N"
TEXT_FLAG = "N"
TERSE_NAME = "encypeverid"
SQL_FORMAT = "CHAR(32)"
BL_SQL_FORMAT = "char(32)"
DISPLAY_FORMAT = "JUSTLEFT"
AVAILABLE_VALUE_TYPE = "N/A"
END_OBJECT = ELEMENT_DEFINITION
END