

Comments on HiPS Standard (V1.0).

Introduction.

Recently the HEASARC has begun to generate HiPS to distribute overviews of the Swift XRT and UVOT datasets as surveys in SkyView and to enable graphic selection interfaces for data in the HEASARC archive. In this effort we have used the Hipsgen jar developed by CDS and also developed our own software for generating and reading HiPS data, e.g., HiPS is now supported as an archive format within SkyView and we generated our intensity maps for both survey using local developed tools.

Although the HiPS standard is -- by IVOA standards -- relatively clear, there are a few areas where it may not be sufficiently detailed. We hope that this review suggests issues that might be clarified in future versions and could be helpful to others who may be confused by some of the same issues we were. We intended this document -- and hope it is understood -- not as a criticism of either HiPS or the standard, but as a recognition of the value of both with the hope that they can be made even better.

This commentary is broken into three parts: a discussion of areas of significant confusion, a set of more trivial editorial comments, and a brief mention of Aladin's apparent conventions using HiPS. The editorial comments are purely suggestive and reflect my own style preferences. Since Aladin is extensively used in both generating HiPS and viewing them it may establish effective conventions that go beyond the actual standard. Each section is given as best possible in document order.

Significant issues.

- *Hips Directory Structure, first paragraph page 8.* I think this document would do much better to include the formulae here rather than by reference. These aren't too long and are fundamental to the document. Put them in an appendix if you don't want them cluttering up the flow. [I suspect this is true for the MOC document also.] This paragraph may also engender some confusion with the "Even if HEALPix supports" Manifestly HEALPix *does* support both schemes. This needs to be something like: "While HEALPix as defined in [5] supports two distinct numbering schemes for pixels, NESTED and RING, HiPS uses only the NESTED scheme."
- *Hips Directory Structure, page 9, first paragraph.* Is there any way a user can tell that a survey does not cover the entire sky or that it has variable order? The MOC and coverage property can help for the first, but I don't know that there is any way to know about the latter. It would be nice if there were at least optional elements of the properties file to convey that information *and* that they were mentioned here.
- *Hips Tile Formats, first paragraph.* I think this paragraph is very confusing. In principle it's suggesting that we could have HiPS of formats other than are described here, but there is nothing in the document that discusses how we could extend HiPS beyond these formats. I think this would do much better starting with first line (up to the colon) and following that with the second sentence. There might be some motherhood followup that additional formats may be defined in the future and implementors are encouraged to suggest details for such.

- ***Image HiPS tile, first two paragraphs and figure 4.*** The definition of image tiles is both confusing and underspecified. This is the premier area in which I think the document could be made clearer – and it’s a key part of the document. E.g., it might start.

“Each Image HiPS tile of order N must contain a $2^S \times 2^S$ set of pixels arranged as a two dimensional image that exactly covers the region of the order N tile. Here S is a non-negative integer which in combination with the tile order N determines the order of the HEALPix pixels used within the tile. Each of the pixels within the tile MUST correspond to the appropriate HEALPix order N+S pixel. The orientation of pixels within the HiPS tile is shown in Figure 4 and tiles MUST conform to this organization. If the data format used supports them, null pixels may be included but no pixels may be omitted. [Tiles [of a given order] MUST all use the same value of S.]” I’m not sure if the statement in []’s is true, but something should be said one way or the other. If variable values of S are permitted then something about that is needed. “

Effectively this defines a third organization for HEALPix pixels along with the Ring and Nested approaches. It would be extremely nice to explicitly define the translation between x,y,S and N and the HEALPix N+S NESTED pixel number. As it stands the second part of Figure 4 is an absolutely essential but extremely easy to overlook element of the document. Given FITS’ confusing history of backwards X axes and the like, the current document doesn’t really define the order of data within the FITS image and that is absolutely critical. To determine this I basically tried a variety of choices until I found that one that matched what Aladin wants.

I don’t believe there should be any statement about resampling in the normative elements of this section. There should be very limited discussion of how a user might generate HiPS images and none of it should be here.

It might help the user to expand upon the example and explicitly note the relationship between the HiPS tile order and the HEALPix pixel order.

“The XMM PN survey has 4” angular resolution. A provider might wish to resample this into a HEALPix grid of order 16 (which has roughly 3.2” pixels [since the pixels are not square’s using 4 digits of precision in the size seems overkill]). If they use 512x512 tiles with S=9 then the highest resolution tiles would have tile order $N = 16-9 = 7$. The HiPS order 6 tiles have HEALPix order 15 pixels, and order 3 tiles would have HEALPix order 12 pixels.”

- ***Deepest HiPS order, entire section***

This entire section should be deleted and replaced by something like:

“The mechanisms by which a data provider resamples some original survey data into the HEALPix tiles are beyond the scope of this document. For data that is already in HEALPix this involves only rearrangement of existing pixel data. Typically a user will generate the highest resolution HEALPix tiles from the original data and then build up lower resolution data by combining higher resolution pixels and tiles using methods appropriate to the data, but this is not required.”

This replacement paragraph, the formula and table (figure 5) should be retained as part of 4.2.1. I'd add a column that gives the HEALPix pixel order (i.e., column 1 plus 9) to the table.

- *Formula before Figure 5.* This is a bad formula to give here. How does the user know the tileWidth? For quite some while I assumed it referred to the tileWidth in angular size, not the pixels width. We might also indicate where this size comes from which gives the user a better understanding of it and its limitations.

This formula is in radians, but we use degrees and arcseconds in the table below. User could be quite confused. We should at least tell them what's going on or we could put in 41253 sq degrees below rather than 4π so that the formula is more consistent with the table.

E.g., “

A nominal HEALPix pixel size can be computed as the square root of the area of each pixel of a given order. Dividing the area of the sky by the number of pixels

$$\text{Tile Pixel Angular Size (radians)} = \sqrt{\frac{4\pi}{12 \times 2^{2(N+S)}}$$

where N is the tile order and S determines the number of pixels in the tiles as discussed above. Note that HEALPix pixels are not generally square so this size should be treated as a typical value.”

The caption and the table are poorly aligned.

- *Other HiPS orders entire section.*
As with the previous section I believe this is not needed in the document and certainly should not be in the normative sections (though it makes no normative statement). If these two sections are retained they should be moved somewhere else.
- *Format of Tiles, paragraph 1.*
Strictly speaking this paragraph is violated by allowing compressed FITS files. A compressed FITS file is not a FITS file. E.g, it doesn't begin with SIMPLE. Support for compression needs to be made clearly stated here. We shouldn't require the user to find this elsewhere. In fact since compression is mentioned in the context of PNG and JPEG's there is a hint here that compressed FITS files are not allowed.
- *Cube tile format.*
I'm not enamored of the use of the lack of suffix having to correspond with the 0 plane. Seems much better if the user got to define that. E.g., using the middle frame which would likely be much more representative. Note also that the title of this section deviates from the format from its two siblings.
- *Catalogue HiPS file*

I've unclear about the relationship between the N and N+1 level tiles in this format (which I have not used). The text seems to say something like:

- Suppose I set my limit for rows in a tile to 100.
- If I have 500 sources in one tile, I pick 100 (somehow, the format doesn't and shouldn't care) and put them in that tile.
- I then take the remaining 400 and divvy them up among the appropriate subtiles

However if this is true then if I'm interested in sources in say a level 8 tile, I need to read the data for that tile, every tile it is contained in since apparently I don't repeat data at lower levels, and every tile it contains (assuming I'm looking for all of the data in this tile). That seems very weird, since if I need to go up the tile hierarchy, I need to check every row I get from the less resolved tiles since they extend beyond my tile. I would have thought that I should never have to look at the lower resolution tiles and that every row appropriate for a given tile is given in that tile directly or one of its subtiles (so that rows must be repeated at multiple levels). I think this needs considerable clarification either way.

- *Order omission, paragraph 1.* It would be nice to understand how a user is to determine if an order has been omitted. As far as I can tell there is no obvious way for a remote HiPS since there is no requirement that a user be able to access the order directories, only the actual tiles. There should be a property (optional) for this.
- *Allsky preview file.*
This is a horrible format. Why? If we want something efficient then just use a NxNx12 cube with the 12 base pixels. That would save space and look better!
- *Properties issues:*
 - If a HiPS is both 'live' and 'color' how is that specified? Two dataproduct_subtype entries or two fields in a single entry?
 - Obs_regime: What about other regimes, gravitational wave, cosmic ray, ...?
 - Hips_status: should refer to where the field values are explained (are all of them?)
 - Hips_estsize: compressed or uncompressed?
 - Hips_pixel_cut: explanation is meaningless to me.
 - Hips_data_range: I understand an actual data range, but what does "data range taken into account" mean?
 - Hips_sampling: Values suggested are far too limited. There are lots of resampling algorithms out there. May even use different methods in different areas. Probably should be free text. Sounds like these are just Hipsgen choices.
 - Hips_overlay: ditto
 - Hips_skyval: ditto
 - Dataproduct_subtype: slightly peculiar to say it's required when it's only so in certain circumstances.

- *5.2 Hips List: In practice...:* This suggests that we can just concatenate properties files, however it was unclear to me how code reading this can separate tell when one entry ends and another begins since there is no required order in these files.
- 6.3. I'd delete this section or make it an appendix. It has nothing to do with the format.

Editorial Comments.

- *The Abstract* uses the abbreviation HiPS before defining it.
- The last clause of the last sentence in *the abstract* is very unclear and could perhaps be omitted. If the intent is to say that this is a simple and easy to implement standard then that can be said, but "leads to practical implementations" means nothing to me.
- *Introduction, paragraph 2.* Typo. "... can manage any type[] of data" (since I gather than only one type of data is handled in a given HiPS)
- *Introduction, paragraph 3.* Run-on sentence. "... into account. [E]mphasis is placed"
- *Introduction, paragraph 4.* Ditto. "... of the Aladin project. [I]ts ..."
- *Introduction, figure at top of page 4.* VOTable is not shown. Although the dependency on VOTable is pretty small this should presumably be consistent with the text in the preceding paragraph.
- *Usage examples, last paragraph.* Run-on sentence. "... into a common architecture. [A]s such..."
- *HiPS principle. Title.* I think this would sound better as HiPS Principle[s].
- *HiPS principle, paragraph 4.* The statement that "Each HiPS tile is **fully** identified..." is a bit confusing in light of the later distinction among different types of HiPS tiles so that a tile file may be a .fits or .tsv extension. While I believe here the intent is to understand the tile purely as a geometric region on the sky, I think the statement would work better if 'fully' were omitted. This is implicit in the following discussion (4.1) where the document notes that an extension of some type will be needed.
- *HiPS directory structure, paragraph 1.* I think this is written in a confusing way. First we say that we have a hierarchy, order->tile, then we insert a new level in that hierarchy. It would be clearer to simply start with a three level hierarchy in the first place:
E.g.,
"The structure of these directories must follow the hierarchy:
Order -> GroupDirectory -> Tile
To avoid very large numbers of tiles in an order directory, each order is broken into Group directories which contain up to 10,000 tiles. For orders lower than 5, all tiles will fit into a single group directory, *Dir0*, but for higher orders contain other group directories given as *DirN* where N is the largest multiple of 10,000 less than or equal to the tile numbers contained. If no tiles for a particular group directory are included in the HiPS then the group directory MAY (or MUST?) be omitted.
Formally, ..."
- *HiPS Directory Structure, page 8 last paragraph.* This paragraph seems out of place. This topic is covered properly in section 5. I think it confuses the reader here.

should be treated a typical value.”

- *Other HiPS orders, first paragraph.* “... four times [fewer] tiles ...”

Aladin Conventions.

My playing with building HiPS suggests that Aladin really wants:

Skip order 0-2.

Order 3-4 required to be uncompressed.

Order 5+ allowed to be compressed or uncompressed

It would be nice if the properties file had some way of indicating which orders were compressed.