

# Hdf5 Reader/Writer Plugin for ImageJ

Matthias Schlachter

July 15, 2010

## 1 Plugin Configuration

The Hdf5 Plugin now offers a configuration dialog which offers setting default values for the plugin behavior. You can open the configuration dialog by “Plugins → HDF5 → HDF5 Preferences” (see figure 1). The configuration dialog is shown in figure 2 and

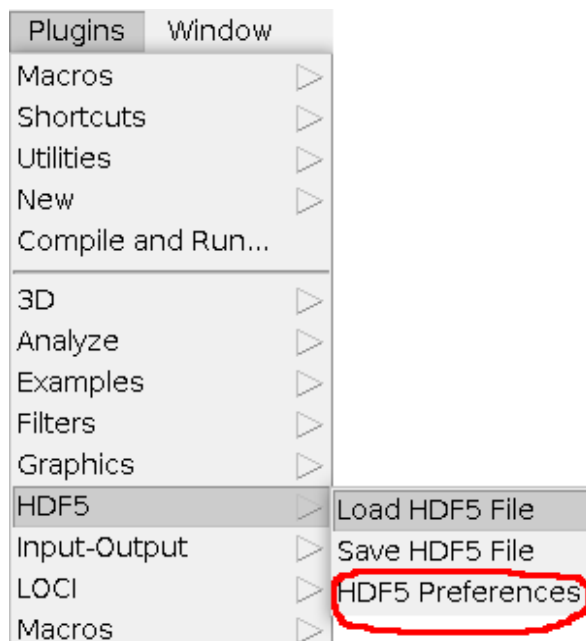


Figure 1: Hdf5 plugin menu

has the following options:

- “Group data set names instead of showing a list of data set names”

- “Show unmatched data set names in a separate list”
- “HDF5 attribute containing pattern for grouping”
- “Pattern for grouping (if no attribute is found)”
- “\$regexp”

which we will explain now in short.

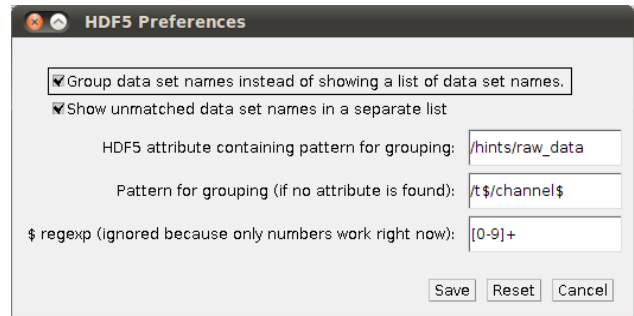


Figure 2: Hdf5 Preferences dialog

“Group data set names instead of showing a list of data set names”: By checking this our plugin tries the group the data set names in the file based on a given for mat string. For example you have recorded 10 timesteps with 3 channels and stored the like this:

```
/t0/channe10  
/t0/channe11  
/t0/channe12  
...
```

```
/t10/channel0  
/t10/channel1  
/t10/channel2
```

You can specify a format string “/t\$T/channel\$C” where “\$T” is the timestep and “\$C” is the channel. The plugin will detect all data set names which match this pattern and allows you to select frame and channel ranges you wish to load (see section 2). The result will be a hyperstack containing all frames and channels you have selected.

**“Show unmatched data set names in a separate list”:** All data set names which do not match the format string are shown in a separate list. By unchecking this all data sets which do not match the format string will be ignored.

**“HDF5 attribute containing pattern for grouping”:** You can specify a global attribute in the file itself which specifies the format string. The default attribute name which holds the format string is “/hints/raw\_data”.

**“Pattern for grouping (if no attribute is found”:** If no attribute is found in the file (see above) the plugin will use the pattern you specify here. The default value is “/t\$T/channel\$C”. The pattern must have “\$T” and “\$C”. “\$T” is the timestep and “\$C” is the channel. Another example for the format string would be “/t\$T/channel\$C\_mip”.

**“\$regex”:** This will replace the “\$” sign in the format string and form a regular expression. For example if “\$ = [0-9]+” the resulting regular expression for the default format string would be “/t[0-9]+/channel[0-9]+”.

However, right now we need numbers to sort the frames and channels and only support “\$ = [0-9]+”. In future releases this may change.

## 2 Reader

If you use the data set grouping feature of the plugin you will get a dialog which lets you specify the ranges you want to load. For example figure 3 shows the range selection dialog for a file containing two timesteps (frames) and three channels. You can now

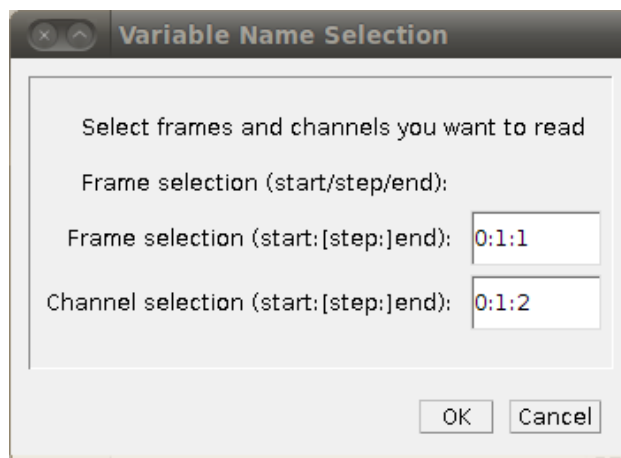


Figure 3: Hdf5 Reader dialog with range selection

specify a start frame/channel, a skip frame/channel and an end frame/channel number, where the skip frame/channel number is optional. Here are some examples:

- You want to read every third timestep/channel starting from 5 ending at 15: “5:3:15”
- You want to read every timestep/channel from 5 to 7: “5:7” (Note: this is equivalent to “5:1:7”)
- You want to read a single frame/channel: specify a single number

If you enter something wrong during the range selection dialog the dialog will start over until you specify correct ranges or press “Cancel”.

If your file contains data set names which do not match the format string and you selected in the configuration to show all unmatched data sets in a list the dialog looks like the one shown in figure 4. You can select the ranges as before and additionally select the (unmatched) data sets you want to load.

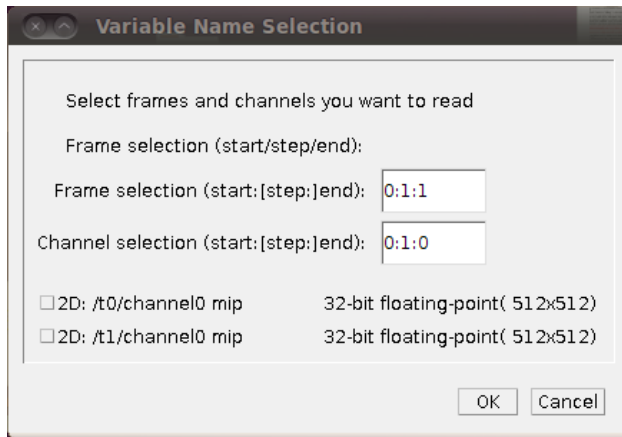


Figure 4: Hdf5 Reader dialog with range selection and unmatched vars

### 3 Writer

No changes made right now. Old behaviour. Means you cannot save the new formats. Will come in ca. 2 weeks

#### 2.1 Supported data sets and storage orders

For the grouping option all single data sets have to be either two- or three-dimensional (greyscale, because frames and channels are stored separately). All other data sets can be either two-,three-,four- or five-dimensional. The only storage order supported right now is time-level-row-col-channel (tzyxc) but may be extended in future releases. This means in detail:

- greyscale image: dim=2D, order=yx
- color image: dim=3D, order=yxc (with last dim “c” = 3)
- greyscale image series: dim=3D, order=tyx
- greyscale volume: dim=3D, order=zyx
- color volume: dim=4D, order=zyxc (with last dim “c” = 3)
- greyscale volume time series: dim=4D, order=tzyx (with last dim “x” > 3)
- color volume time series: dim=5D, order=tzyxc (with last dim “c” = 3)

All data sets which have channels or frame now get opened as a hyperstack which means you can use the channel tool to enable/disable channels.