Multi-object Spectroscopy: Automatic extraction and previewing of spectra

N. Pirzkal, F. Kerber (ST-ECF/ESO)

With multi-object spectroscopy becoming an integral part of modern astronomical instrumentation, the amount of such data in archives is bound to increase manyfold in the very near future. One challenge is to make individual spectra easily accessible for preview by the archive users.
We have used aXe, a software package developed to perform slitless spectral extraction for the Advanced Camera for Surveys (ACS) on the HST, to extract spectra from FORS MXU data. One key element of routine spectral extraction is that the extraction process must be robust, take place in a reasonable amount of time, and be general enough so that it works equally well independently of the type of spectrum. We believe that the current approach to the extraction of spectra from FORS MXU satisfy these conditions while being more sophisticated than a simple quick and dirty box extraction method. The extraction process follows the general principle of ACS spectral extraction which is outlined in (Pirzkal et al., ST-ECF Newsletter, July 2001, p.5). This allows us to perform tilted spectral extraction and to produce wavelength calibrated spectra which can be readily examined using current spectral analysis tools which are part of IRAF.

For the future we foresee the possibility to add further functionality facilitating scientific research in an AVO context. For a rough classification of the objects the extracted spectra could be cross-correlated with templates from a library of digital spectra helping the archive user to judge the relevance of the spectra for his research interest.

Why is this needed?

Modern telescopes and multi-object spectroscopy instruments are producing increasingly larger volume of data. These data contain a wealth of information but are difficult to extract and calibrate. It is very difficult for a typical archive user to look at one of these datasets and to rapidly pick which objects from the hundred or thousands of objects present in the data should be examined in more details. This is one reason why previously archived MOS data, from the HST/STIS pure parallel grism program for example, have not been used extensively in the past.

As more MOS data start to be ingested into archives around the world, and in order to make these data appealing to users, one needs to make extracted preview spectra available to the users, instead of just the raw datasets. While an automated extraction process might not always perform the extraction in the exact way every user might want, a robust, conservative extraction can nevertheless be set up to produce spectra which are usable for the majority of the science typical users might want to do. At the very least, routine extraction of spectra from these data will offer a wealth of ways for users to select which objects to select to carry out their project.

Robust
The extraction should be designed so that a bad object, or other unforseen problems with part of a dataset does not affect the rest of the extraction process. Extraction parameters should be chosen carefully in order to not be systematically biased toward a single type of objects and/or science since the scientific content of a dataset cannot be necessarily a-priori known.

Fast
The large amount of data to be extracted from current and future multi-object spectroscopic data requires an extraction process which is efficient enough so that it does not take a prohitively large amount of time.

Calibrated
While difficult, it is imperative that the extracted spectra be wavelength and flux calibrated.

Handle both slit and slitless data
Being able to handle both of these spectroscopic data type is essential in order to be able to extract spectra from the majority of the data being currently archived (FORS/VLT, ACS/HST)

Experience generating preview spectra

Our experience extracting spectra from FORS2 MXU data has shown us that simply archiving these images will likely not lead to many users investing the considerable amount of time required before one can start looking at the spectra of objects. A considerable amount of knowledge is required before one can even start to extract information from these data. At the very least, a wavelength calibration of the data is required to identify spectral feature in the data. In this respect, multi object spectroscopic data is significantly more difficult to use right away than imaging data and requires a large amount of "buy-in" time in order to be used even in the most basic form (e.g. to simply select interesting spectra from a large pool of available data). In our case, we wanted to extract all of our spectra in an automatic manner because we had several tens of images, each containing 60 to 70 spectra (Figure 2). But in order to even start selecting interesting objects, we had to combine information obtained from various sources (See Figure 2) so that we could first obtain wavelength and flux calibrated spectra of all the objects. Our goal was not to automatically generate optimally and perfectly calibrated spectra right away but rather to use a relatively straight forward extraction method and apply calibration to the extracted spectra which would be good to within 5 to 10 percent. This, allowed us to examine the content of our images in a much more efficient way.

The following steps were performed:

- Dark and flat-fielding the science and wavelength calibration data
- Finding the positions of the slits in the field of view using the slit acquisition images
- Finding the exact position of the objects in the slits
- Extracting the lamp wavelength calibration spectra
- Extracting the science spectra, using tilted extraction along the direction of the slit
- Wavelength and flux calibrating the extracted spectra

If performed manually, for one spectrum at a time, these steps took a significant amount of time, on the order of several minutes per object. On the other hand, once enough had been learned about these data, mainly in terms of the wavelength and flux calibration of this observing mode, it was relatively easy to set things up so that we could generate hundreds of preview spectra in the course of a few minutes.
Scripting various aXe tasks together (See Figure 1 for a complete list of the available aXe tasks) and adding some Python scripts to pre-compute the location of the slits in our data, we were able to extract and calibrate our data in a manner which is robust (all spectra were extracted successfully), fast (on the order of 1 to 2 minutes per data sets), and produces calibrated spectra in a FITS format which can be readily analysized using IRAF tasks.

The same methodology, also based on an automatic processing using aXe, was recently used to analyze new HST/ACS grism observations of the HDF-N. Using this method, 1700 objects were extracted in a matter of minutes, and each extracted spectra could then be quickly examined. This resulted in two supernova being discovered within a matter of hours from when the data was first archived. This, together with our experience with FORS MXU data illustrates the strength of being able to archive spectra which have been generated in an automatic manner using relatively robust extraction parameters.

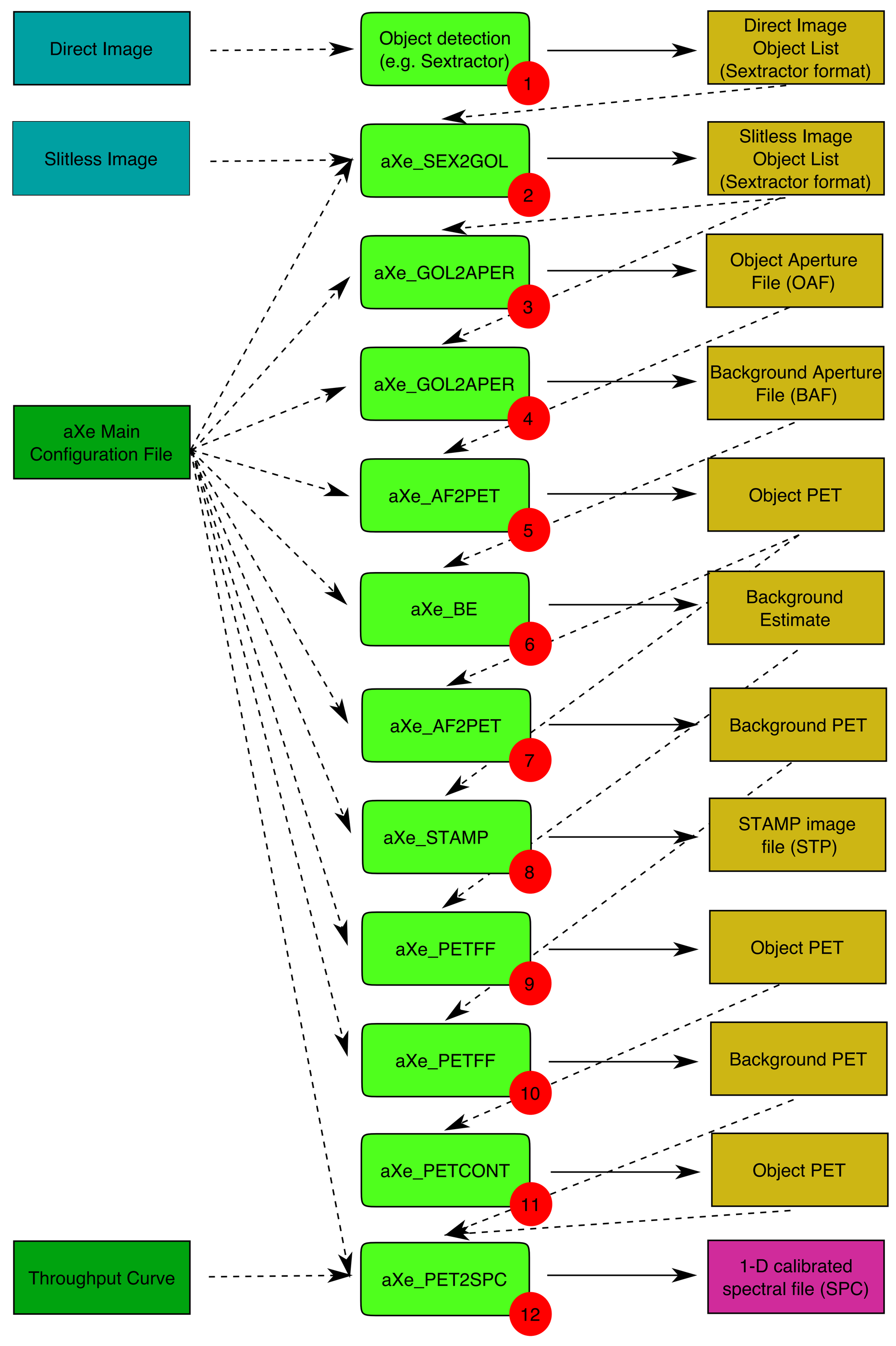

Figure 1: aXe tasks and a typical extraction process

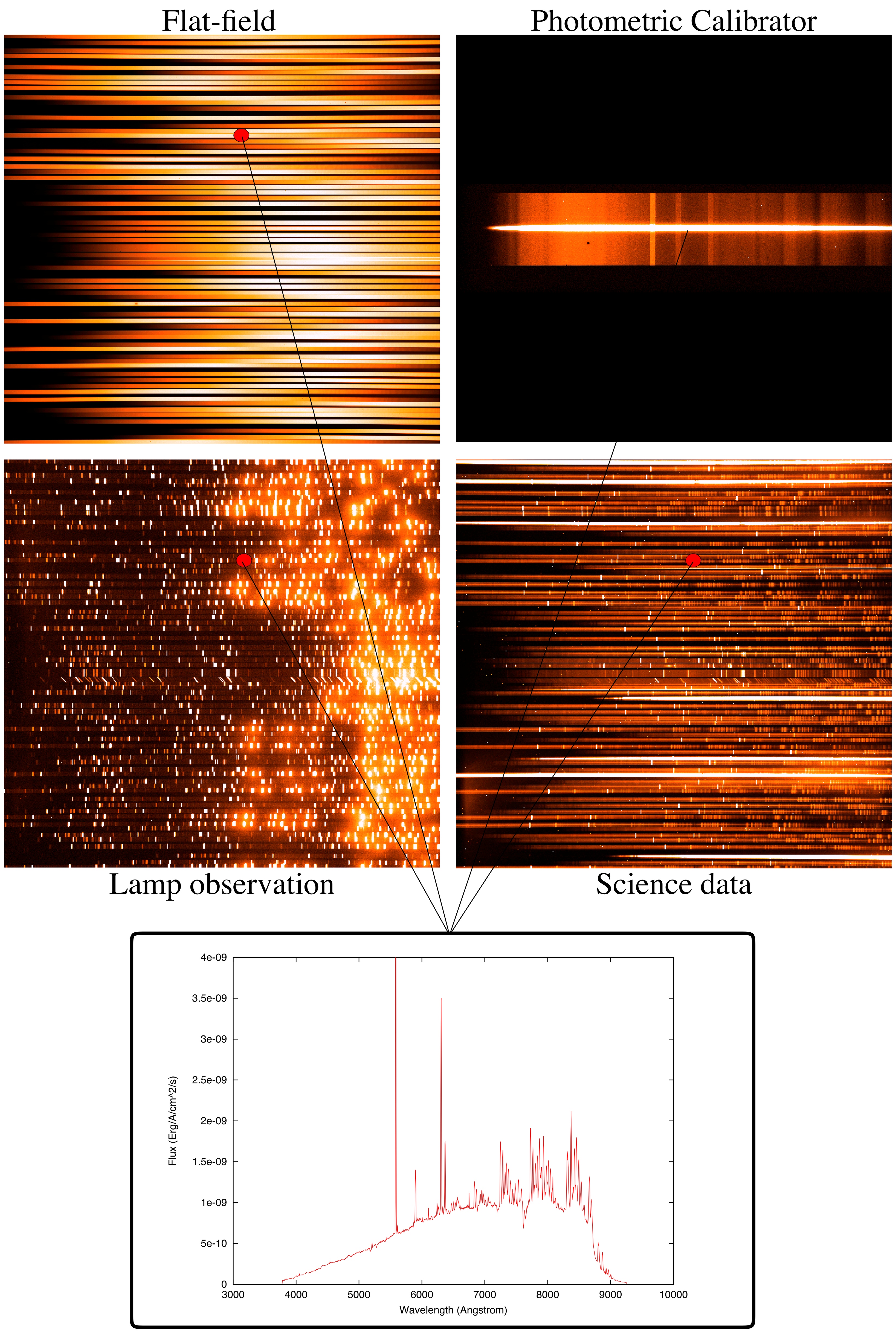

Figure 2: A meaningful preview spectra which is wavelength calibrated requires a bit more processing than simple direct imaging. Information must be gathered from a variety of source and spectra must be extracted in such a way that most spectra are successfully extracted and automatically calibrated. Note the sky lines in the sample spectrum which is shown.

Further use of preview spectra

While we have shown the strength of spectra extracted automatically from MOS data and how examining preview spectra is an efficient way for users to determine the scientific content of these data, we have so far made only basic use of these spectra. More useful use of these data can be foreseen. For example, routinely extracted spectra can be cross-correlated with templates of stellar or galactic spectral template to automatically estimate their stellar or galactic types.