Astronomical Data Analysis Software and Systems XVI
ASP Conference Series, Vol. XXX, 2006
R. Shaw, F. Hill and D. Bell, eds.

The VIRUS Emission Line Detection Recipe

Claus A. Gössl, Ulrich Hopp, Ralf Köhler, Frank Grupp

Universitäts-Sternwarte München

Helena Relke, Niv Drory

Max-Planck-Institut für extraterrestrische Physik

Karl Gebhardt, Gary Hill, Phillip MacQueen

The University of Texas, Department of Astronomy

Abstract. HETDEX, the Hobby-Eberly Telescope Dark Energy Experiment, will measure the imprint of the baryonic acoustic oscillations on the galaxy population at redshifts of $1.8 < z < 3.7$ to constrain the nature of dark energy. The survey will be performed over at least 200deg^2. The tracer population for this blind search will be Ly-α emitting galaxies through their most prominent emission line. The data reduction pipeline will extract these emission line objects from $\sim$ 35.000 spectra per exposure (5 million per night, i.e. 500 million in total) while performing astrometric, photometric, and wavelength calibration fully automatically. Here we will present our ideas how to find and classify objects even at low signal-to-noise ratios.

1. Introduction

HETDEX (Hill et al. 2004), a new survey of the distribution of high redshift galaxies, aims at understanding the evolution history of our universe. Observations of supernovae of type SN Ia show that about 70% of the total energy of the Universe consists of dark energy. Very little is known about this mysterious dark energy so HETDEX has been designed to describe precisely the expansion history of the universe based on observations within about 110 clear nights. The distribution of one million galaxies in a volume several times larger than those investigated so far (in the redshift z regime $1.8 < z < 3.7$) will be mapped by HETDEX. VIRUS (Hill et al. 2004), a new instrument for the Hobby-Eberly Telescope (HET, Ramsey et al. 1998, Booth et al. 2003). is specifically designed to meet the challenge. Industrial replication of 145 copies of a simple integral field spectrograph will allow not only an efficient, but also much less expensive instrument to be built, compared to classical large prototype designs where in most cases only a single copy is built (Hill & MacQueen 2002). The highly parallel design of the instrument results in a massive parallel data output, and a challenging series of calibration steps because of the IFU technique. So far the latter has been only accomplished in interactive pipelines for existing single-spectrograph IFU systems. The VIRUS data rate will require a robust automatic pipeline from basic removal of the instrument profile up to the extraction of the final catalog of galaxies (Gössl et al. 2006). A prototype has already been installed at the McDonald 2.7m telescope to test the whole concept.

The distribution of matter will be traced with galaxies emitting Ly-α in emission (LAE). Because a large fraction of all galaxies is still actively forming stars at these high redshifts (e.g. Gabasch et al. 2006), many show this line in relatively strong emission and are therefore relatively easy to find and their redshifts are straightforward to extract.

2. Emission Line Detection

2.1. The VIRUS data challenge

The VIRUS instrument will cover a 20 arcmin diameter field-of-view (FoV) of the new HET corrector with 1/9 fill factor employing a set of 145 identical IFU spectrographs. Every IFU is built of 247 fibers and covers 0.2 arcmin^2 (1 arcsec^2/fiber) which again results in 35.815 fibers (i.e. spectra) and therefore about 14 million resolution elements per exposure. The current observational strategy foresees a data output of 20 GB raw data per hour or (including calibration) about 200 GB raw data per night.

2.2. Simulating spectra

An image simulation package was developed to help in assessing the instrument performance during the design phase as well as to help develop the VIRUS reduction pipeline. This simulation software `VSIM` was built using the output of the optical modeling and ray-tracing software (ZEMAX) used in the optical design process and simulates CCD images resembling the final instrument as closely as possible by: Sampling of the sky with the IFU head, followed by mapping of the fiber position along the pseudo-slit and wavelength onto the CCD, succeeded by convolution of the spectra with the output fiber profile and the PSF of the spectrograph and camera, and finally, resampling the output of the fibers onto CCD pixels.

2.3. Reduction pipeline

The reduction pipeline first takes care of splitting exposure data cubes into single per detector frames which can be processed in parallel. The "raw" reduction part of the pipeline takes care of bias correction, photon noise statistics, bad pixel mask, flatfield calibration, cosmic ray rejection, and straylight illumination correction with every step including full per pixel error propagation following the recipes of Gössl & Riffeser (2002, 2003). Raw reduction will produce data that resembles an ideal or the "simulated" instrument as closely as possible. The pipeline continues with deriving first guess WCS data, adding spectral trace information to the frames, performing a wavelength calibration, subtracting night sky emission lines, refining the astrometric solution, and applying a photometric calibration.

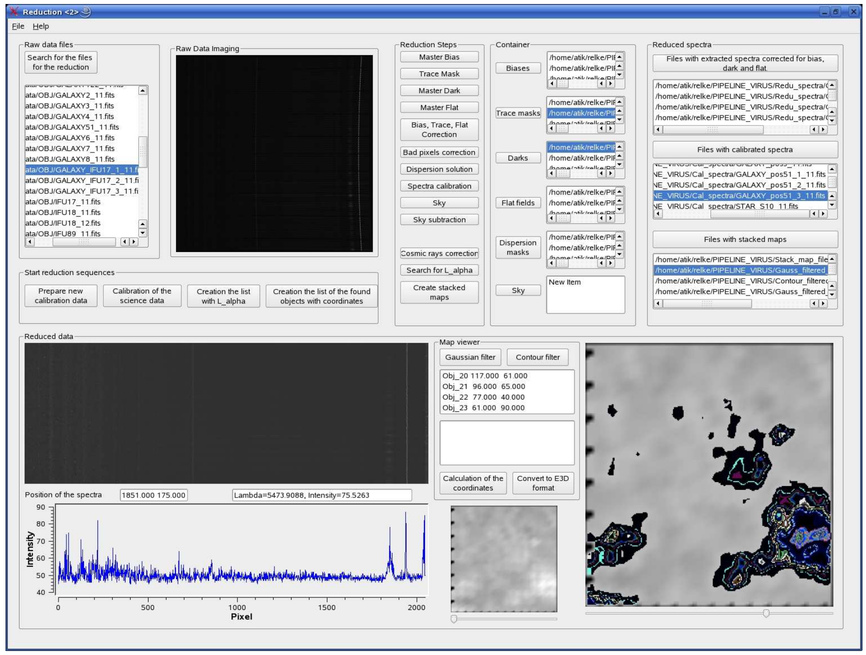

Figure 1. For evaluation of the reduction procedures and as a quick-look interface for the full VIRUS instrument we developed a GUI (QT) which can operate all reduction steps (i.e. scripts).

2.4. Line detection

To reach the science goal it is crucial to find as many as possible and therefore also the very weak LAE galaxies. And, of course, having a quantified detection qualifier is highly desirable to help for probing the powerspectrum of the baryonic acoustic oscillations. We used simulated data to test blind line profile searching approaches as well as noise suppression techniques by wavelet transforms (Starck et al. 1997) but with no convincing results. We therefore decided to implement the following recipe: All spectra will be convolved with a spectral line model using the simulator after improving the optical design parameters with real data. Successive rebinning to a uniform wavelength grid gives correlated noise within size of the convolution kernel. Finally, spatial convolution of a complete set of dithers with a stellar PSF yields a “spaxel” based probability function in terms of S/N for an emission line in the $\lambda - RA - Dec$ space. While the noise of close spaxels is correlated within the limits of the convolution kernels it is still a meaningful indicator. The calculation of this probability function has no iterative steps, requires only multiplication and addition, and therefore will be very fast. The ray-tracing simulator can now be used to create artificial frames with the candidates indicated by the probability function. To determine the final λ, RA, Dec, and flux of emission line objects a least-χ^2 solution for the comparison of the simulated frame with the reduced science frame will be calculated (e.g. by adaptive simulated annealing). That way it is also possible to quantify the error budget of every parameter (λ, position and flux). This approach can also be expanded to search for specific wavelength ratios or even flux ratios in the continuum on both sides of the line adding further LAE galaxy (dis-/)qualifiers.

3. Project Status

We recently added PMAS-PPAK observations as well as the first VIRUS-Prototype observations to the simulated data. A graphical user interface (Fig. 1) has been designed to interactively test parts of the pipeline as well as to inspect every step. Different configuration setups for PPAK, VIRUS-P, and the simulated VIRUS are used. We are currently evaluating the observations as well as further implementing and improving the pipeline.

Acknowledgments. Our thanks are due to M. Roth, A. Kelz and P. Böhm, AIP, for discussion and software support as well as a collaborative effort in hardware development for the VIRUS IFU. We further acknowledge discussions with P. Schuecker, MPE (who unexpectedly passed away Nov. 12, 2006) and E. Komatsu, Austin. We like to thank BMBF/DESY (05AV5WMB/6) for their financial support. The HETDEX team acknowledges generous support from the George and Cynthia Mitchell Foundation, and Harold C. Simmons.

References

Booth, J. A., Wolf, M. J., Fowler, J. R., Adams, M. T., Good, J. M., Kelton, P. W., Barker, E. S., Palunas, P., Bash, F. N., Ramsey, L. W., Hill, G. J., MacQueen, P. J., Cornell, M. E., & Robinson, E. L. 2003, Proceedings of the SPIE, 4837, 919

Gabasch, A., Hopp, U., Feulner, G., Bender, R., Seitz, S., Saglia, R. P., Snigula, J., Drory, N., Appenzeller, I., Heidt, J., Mehlert, D., Noll, S., Böhm, A., Jäger, K., & Ziegler, B. 2006, A&A, 448, 101

Gössl, C. A., Drory, N., Relke, H., Gebhardt, K., Grupp, F., Hill, G., Hopp, U., Köhler, R., & MacQueen, P. 2006, Proceedings of the SPIE, 6270, 66

Gössl, C. A., & Riffeser, A. 2002, A&A, 381, 1095

Gössl, C. A., & Riffeser, A. 2003, in ASP Conf. Ser., Vol. 295, ADASS XII, ed. H. E. Payne, R. I. Jedrzejewski, & R. N. Hook (San Francisco: ASP), 229

Hill, G. J., Gebhardt, K., Komatsu, E., & MacQueen, P. J. 2004, AIP Conference Proceedings, 743, 224

Hill, G. J., & MacQueen, P. J. 2002, Proceedings of the SPIE, 4836, 306

Hill, G. J., MacQueen, P. J., Tejada, C., Cobos, F. J., Palunas, P., Gebhardt, K. & Drory, N. 2004, Proceedings of the SPIE 5492, 94

Ramsey, L. W., Adams, M. T., Barnes, T. G., Booth, J. A., Cornell, M. E., Fowler, J. R., Gaffney, N. I., Glaspey, J. W., Good, J. M., Hill, G. J., Kelton, P. W., Krabbendam, V. L., Long, L., MacQueen, P. J., Ray, F. B., Ricklefs, R. L., Sage, J., Sebring, T. A., Spiesman, W. J., & Steiner, M. 1998, Proceedings of the SPIE, 3352, 34

Starck, J.-L., Siebenmorgen, R., & Gredel, R. 1997, ApJ, 482, 1011