

New Bayesian Fusion Scheme and Visualization Tool for Astronomical Hyperspectral Data Cubes

Matthieu Petremand^a, Mireille Louys^a, Christophe Collet^a, Vincent Mazet^a
André Jalobeanu^b
Fabien Salzenstein^c

^aLSIIT - UMR CNRS 7005 - University of Strasbourg - France

^bCentro de Geofísica - University of Evora - Portugal

^cINESS - UMR CNRS 7163 - University of Strasbourg - France

Keywords: hyperspectral visualization, hyperspectral data fusion, inverse problem, Bayesian inference

1. INTRODUCTION

Within the DAHLIA project (funded by ANR, France), the PASEO team of the LSIIT laboratory (<http://lsiit-miv.u-strasbg.fr/paseo/>) at the University of Strasbourg aims at providing new methods for the visualization and the statistical fusion of hyperspectral data cubes acquired with the new generation of Integral Field Spectrographs (IFS). In particular, we focus our attention on the Multi Unit Spectroscopic Explorer (MUSE) instrument [1] designed in an European consortium started in 2004 and led by the CRAL laboratory in Lyon (<http://muse.univ-lyon1.fr/>). This new IFS, composed of 24 IFU (Integral Field Unit), shall see its first light in 2012 at the VLT (Very Large Telescope) site in Paranal, Chile. The MUSE consortium brings together more than a hundred scientists and engineers who cover all the required expertise to build and exploit this unique instrument allowing to observe galaxies when the Universe was only a few billion years old. For the first time, high spatial and spectral resolutions should be obtained for extremely distant and faint galaxies. In order to achieve the required sensitivity, the total exposure time must reach 80 hours, divided into 80 exposures of 1 hour each to avoid cosmic ray accumulation. A single deep-field observation will then be composed of 80 raw CCD images covering the same field of view but acquired under varying observation conditions: spatial and spectral sampling lattices, FSF/LSF (Field/Line Spread Function), geometric distortions, cosmic rays, noise, missing or corrupted values... and representing nearly 100 GB of data, which is challenging for modern processing algorithms and technologies.

The expertise domain of PASEO team brings new fusion methods and visualization tools in this project. Indeed, each MUSE acquisition session of the same astronomical scene will lead to a hyperspectral observation (or image) whose size is about 301×301 pixels through 3578 spectral bands. Because of the complexity of a MUSE observation (not less than 80 poses for one final observation), a first step consists in developing *ad hoc* visualization tools allowing to intuitively browse and display this huge amount of data. Such a visualization software we developed is presented in section 2 and provides a convenient way to explore hyperspectral cubes. In a second step, as the amount of data grows while the number of exposures increases, a robust fusion method is required so as to produce a reduced data cube including all the information contained within the original observation set. In addition, hyperspectral images being acquired under different acquisition conditions, the image fusion must be performed according to these varying observation parameters. In this paper, we propose a new Bayesian algorithm (see section 3) allowing the reconstruction and the fusion of the set of raw observations with respect to acquisition parameters. Finally, the Bayesian framework gives access to an estimation of uncertainties over the output image, *i.e.* variances and covariances are available for each fused pixel. Further investigations on the fused image (such as segmentation, source separation, astrometry or photometry computation, etc.) can then use these statistical outputs to improve their performances. This is a challenging but required approach to

Corresponding author: c.collet@unistra.fr.

This work was partially funded by the French Research Agency (ANR) as part of the DAHLIA project (grant #ANR-08-BLAN-0253). Project web site: <http://dahlia.oca.eu>.

correctly manage MUSE observations. Such fusion methods have already been developed and implemented for the reconstruction and the fusion of spectra [2] and 2D images [3].

Before the fusion of raw images, two mandatory tasks must be performed: cosmic ray detection and rejection as well as sky background subtraction. Indeed, the presence of cosmic rays at random locations on CCD chips may influence the fusion result whereas the sky background, different for each observation, must be removed in a pre-processing step. In this paper, we propose a two-step method for the detection/rejection of cosmic rays on raw observations. This algorithm and the sky subtraction one are currently under development and will not be illustrated in fusion results shown in section 4. The management of large datasets and the visualization of uncertainties are also discussed in this section.

2. VISUALIZATION OF HYPERSPECTRAL DATA

2.1 Visualization Needs

Before any processing on hyperspectral observations, a visualization step is often performed so as to provide a first overview of acquired data: object preview, spectrum extraction, data quality estimation, observation parameter displaying, sub-cube extraction, research of areas of interest, basic operations like mean over frames or spectra...

Moreover, additional visualization modes may be involved throughout the data analysis process: visualization of processing algorithm outputs, result comparison or formatting... Thus, during the lifecycle of a hyperspectral observation (from acquisition to result validation), visualization tasks play an important part and require the use of dedicated softwares. However, most of the existing tools are not suitable because of the complexity, heterogeneity and the increasing size of hyperspectral images as well as related observation parameters to be visualized (up to 2.4 GB for a MUSE observation with its associated variance cube). In the framework of the DAHLIA project, we aim at developing a new application that will be used for visualizing, browsing and handling large hyperspectral images (and their associated observation parameters) acquired by the MUSE instrument. Nevertheless, instead of starting the development from scratch, we would prefer to base our application on an existing cube visualization software whose choice depends on its implemented features and scalability. To guide us in this choice, we have determined a list of requirements our final application should meet to be considered as an efficient one:

- basic astronomical visualization features such as 2D (frames) and 1D (spectra) visualizations, spectrum extraction and comparison, zoom capabilities, astronomical unit and calibration management, FITS I/O (2D and 3D files), selection management, support and superposition of astronomical catalogs...
- ability to display or handle observation parameters and metadata like variances or PSF (Point Spread Function) ;
- synchronized browsing in both spatial and spectral calibrated spaces to explore and compare cubes, images and spectra ;
- ability to customize the way spectra are displayed thanks to different visualization modes ;
- implementation of simple visualization algorithms such as colored composition, mean of frames or spectra, sub-cube extraction...
- as the software is mainly intended to only provide visualization features, an interaction with VO (Virtual Observatory) tools (like VOSpec [4, 5] or SPLAT [6]) using SAMP protocol [7] should be available for advanced astronomical processing ;
- optimized memory management to load, handle and view several large data cubes even on desktop computer (32 or 64 bit platforms) ;
- scalability so that users may develop their own algorithms or visualization modes ;

- portability, easy-to-install and user-friendly graphical user interface to quickly be able to use the software on different operating systems without wasting time on compilation and configuration steps.

For the time being, there is no existing software that entirely fits our needs. According to our specifications, we found two visualization softwares that could be used as host applications to build our visualization tool (*i.e.*, the chosen software can be easily extended to implement the missing features). The first one, GAIA3D [8], is developed at ESO (European Southern Observatory) as part of the VLT project and corresponds to our prerequisites except, *i.a.*, for the portability and the memory management points (that cannot be added *a posteriori* in the application). Indeed, GAIA3D is developed in various programming languages (C, C++, Tcl/Tk...), cannot handle more than one hyperspectral cube at the same time and requires as much memory as the size of the image to be loaded. Aladin [9, 10], the second one, is developed in Java at the CDS (Centre de Données astronomiques de Strasbourg). It is a cross-platform software applying an optimized memory management: for instance, a MUSE observation of 1.2 GB only uses about 200 MB of memory after loading, meaning one can easily handle several cubes at the same time. It also provides a convenient programming framework allowing developers to easily add plug-ins that can closely interact with it. However, Aladin mainly focuses on 2D visualization and astronomical features (calibration management, catalogs...) even though it is able to load and browse hyperspectral cubes quickly. We then decided to extend the spectral functionalities of Aladin by developing, in Java, a dedicated plug-in named QuickViz [11].

2.2 Aladin/QuickViz Duo

QuickViz provides a set of basic and advanced features designed to ease the exploration of hyperspectral data cubes with the help of Aladin: all astrometrical calibration, WCS coordinates management, registration, etc. are totally delegated to the Aladin browser application whereas all the spectral information is handled by QuickViz. Moreover, QuickViz manages all the interaction for supporting synchronized navigation between spatial and spectral domains. Figure 1 illustrates the main functionalities of QuickViz as mentioned in the previous section and detailed hereafter (circled numbers on the figure 1 are referred to those in italic font in the following):

1. spectrum extraction from hyperspectral cubes can be performed straight at each individual pointed position (1.1) or by computing an averaged spectrum over a user-defined region (1.2). Each extracted spectrum is added to the spectrum list of QuickViz (1.3) and available for display and superposition on spectrum panels (1.4). Selections can then be defined over a spectral range (1.5) and benefit from intuitive zoom features (1.6). To speed up the display, the data extraction uses a multi-thread approach all over the data cube ;
2. navigation coupling between spectral and spatial axes: moving the calibrated cursor (2.1) updates the cube's frame in the Aladin active window (2.2) ;
3. customizable spectrum visualization thanks to multiview (synchronized or not) support (3.1) and many visualization modes (3.2). New modes can easily be added thanks to an implemented plug-in architecture (Java interface) ;
4. spectrum or frame processing algorithms (4.1) can be executed on selected data and provide products such as mean, sum, weighted sum, sub-cube extraction, RGB colored composition (4.2)... Algorithm outputs are automatically added in spectrum (1.3) or cube (4.3) lists. Like visualization modes, new or customized algorithms can be developed in Java and added by users to the QuickViz algorithm list and extend the processing capabilities of the plug-in ;
5. Observation parameters like variances (5.1) can also be visualized (5.2) together with their associated observation.

Additional information can be found on QuickViz's website <http://lsiit-miv.u-strasbg.fr/paseo/cubevisualization.php> such as an exhaustive list of features with corresponding walk-through, documentation, download and installation instructions for Aladin and QuickViz.

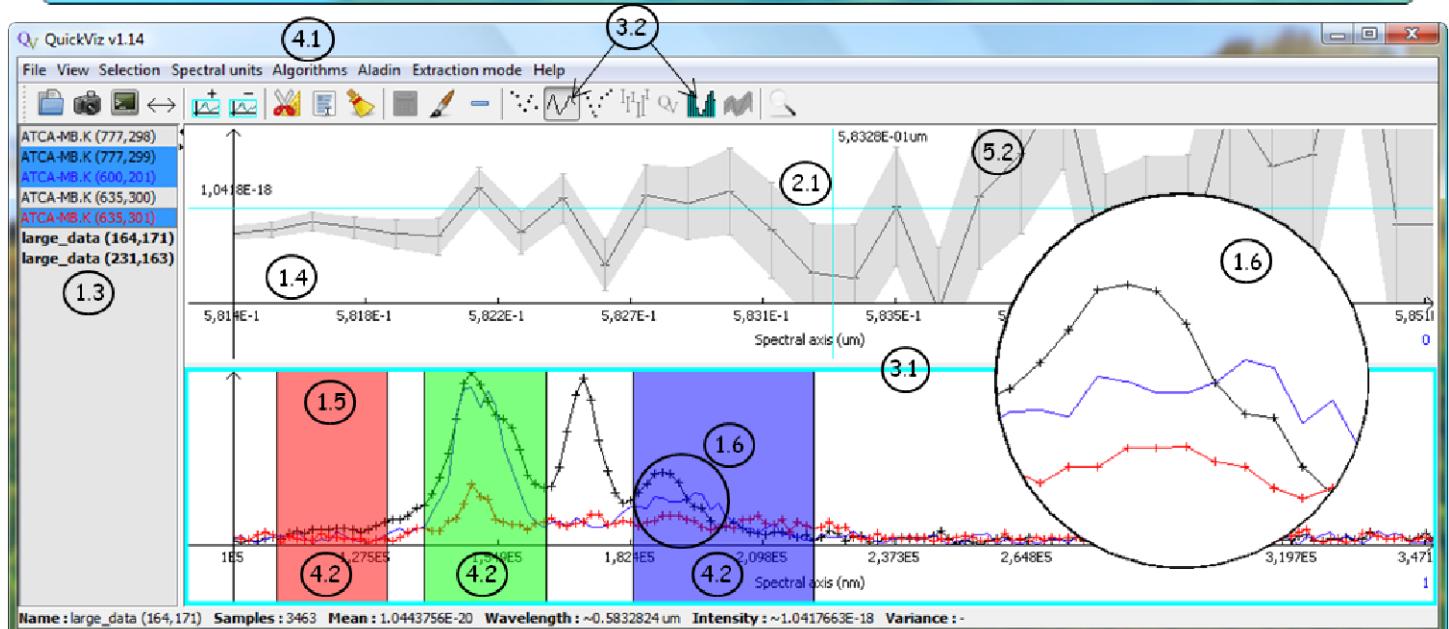
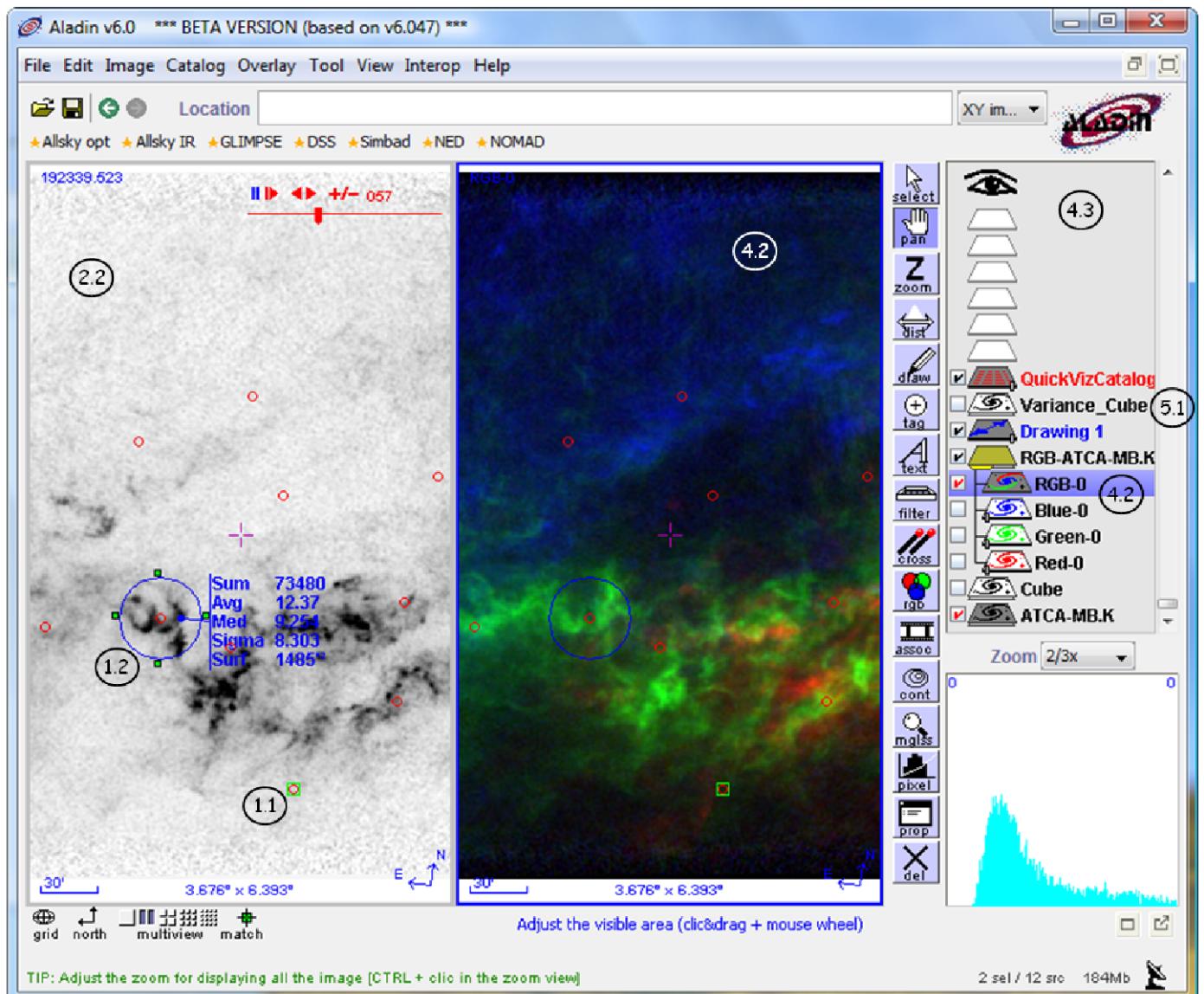


Figure 1. QuickViz plug-in (bottom frame) and its host application Aladin (top frame). Spectral data are fully tackled by the plug-in, in connection with the spatial domain managed by Aladin functionalities. Circled numbers refer to the required features listed in section 2.2

2.3 Development Perspectives

The development of QuickViz is almost completed. We plan to extend the first release with several functionalities. The implementation of the SAMP protocol is one of our first priorities. This VO standard messaging protocol allows QuickViz to communicate with a set of softwares dedicated to a finer astronomical analysis (like VOSpec or SPLAT). In the case of MUSE hyperspectral data, extra datasets useful for interpretation will be associated to an observation like variances, local PSF information, data quality flags... We first focused our developments on the visualization of the variance along the spectral axis but we are currently working on 2D visualization of the variance (see section 4.2) and PSF information.

We also plan to suggest VO compliant description of metadata for observed data cubes in collaboration with the IVOA Datamodel working group [12]. Finally, we think that the support of new file formats, especially those dedicated to the storage of large scientific datasets like the HDF format [13], is crucial for handling advanced hyperspectral data products and we will enable the use of such formats in QuickViz.

3. FUSION SCHEME

The proposed fusion scheme consists in computing a fused hyperspectral image from n raw images Y^i acquired under varying observation parameters Θ^i such as PSF, spatial and spectral shifts, atypical pixels, acquisition noise... Thus, each Y^i corresponds to a sampled, band-limited, noisy and disturbed version of the observed scene T whose spatial and spectral resolutions are infinite. The set of disturbances affecting T is summarized in the following forward model.

3.1 Forward Model

The continuous scene $T(x, y, \lambda)$ lies in the model space whereas discrete raw observations $Y^i(s, t, k)$ lies in the sensor space where s corresponds to spatial indices, t to spectral indices and k to IFU indices. Thus, a MUSE raw observation is composed of 24 CCD images whose size is about 4000×4000 pixels each. Mappings between sensor positions $p = (s, t, k)$ and model space locations $u = (x, y, \lambda)$ are given by tables called “pixtables” and associated to observations. Additional information about optical image formation inside MUSE can be found in [1]. One can now write the forward model in the form of an observation equation linking each \tilde{Y}_p^i , i.e. Y_p^i after radiometric correction, to T (detailed explanations are available in [14]):

$$\tilde{Y}_p^i = (T \star h_{u_p^i}^i)(u_p^i) + B_p^i \quad (1)$$

where:

- u_p^i stands for the 3D sampling grid including all the sampling geometry (shifts, orientation parameters...) and linking sensor space to model space locations ;
- $h_{u_p^i}^i$ corresponds to a 3D separable convolution kernel ($h = LSF \times FSF$) whose width depends on the model space location ;
- B_p^i is an additive noise such as $B_p^i \sim \mathcal{N}(0, \sigma_p^i)$ where σ_p^i is a signal-dependent standard deviation. σ_p^i is set to $+\infty$ for cosmic rays (once they have been detected) or outlier pixels (known locations).

As the high frequency content of T has been definitely lost during the acquisition process, one cannot recover T from the set of band-limited observations. Thus, our fusion method rather aims at recovering a band-limited version of the scene $F = T \star \varphi$ from the set $\{\tilde{Y}^i, \Theta^i\}_i$ where φ corresponds to the constant PSF of an ideal instrument whose resolution may be higher than the MUSE instrument. This ideal PSF is modeled by a nearly band-limiting 3D separable B-spline of degree 3 [15]. F can then be interpolated for each model space location u :

$$F(u) \simeq \sum_m L_m \varphi(u - m), \quad m \in \mathbb{Z}^3 \quad (2)$$

where L is a discrete set of interpolation coefficients.

Instead of recovering a continuous function F , we aim at its discrete version $X = s \star L$ ($X = \mathbf{S}L$ in matrix notation) where s is the discrete version of φ . Under the assumption of band-limited PSF (basically, they are wider than φ as for the MUSE instrument), one can rewrite Eqn. (1) so as to express each \tilde{Y}_p^i as a linear combination of interpolation coefficients L [14]:

$$\tilde{Y}_p^i = \underbrace{\sum_m L_m \alpha_{pm}^i}_{I_p^i} + B_p^i \quad \text{with} \quad \alpha_{pm}^i = h_{u_p^i}^i (u_p^i - m) \quad (3)$$

where α_p^i is the set of *rendering coefficients* associated to \tilde{Y}_p^i which are totally determined by Θ^i (in matrix notation, $I^i = \boldsymbol{\alpha}^i L$). Thus, with the knowledge of $\{\tilde{Y}^i, \Theta^i\}_i$, one can infer \hat{L} from $\{\tilde{Y}^i, \alpha^i\}_i$ (which corresponds to an inverse problem) and then deduce $\hat{X} = \mathbf{S}\hat{L}$.

3.2 Hyperspectral Data Fusion

In a Bayesian framework, the fusion of raw observations all comes down to maximize the following posterior probability:

$$P(L|\{\tilde{Y}^i\}_i, \omega) \propto \prod_i P(\tilde{Y}^i|L) \times P(L|\omega) \quad (4)$$

where, from Eqn. (3), $P(\tilde{Y}^i|L) \sim \mathcal{N}(\boldsymbol{\alpha}^i L, \mathbf{P}^{i-1})$ with \mathbf{P}^i being the diagonal inverse covariance matrix associated to \tilde{Y}^i and $P(L|\omega) \propto \exp(-\omega \|\mathbf{D}SL\|^2)$ a smoothness prior (first derivative operator \mathbf{D}) defined on $X = \mathbf{S}L$ and governed by the hyperparameter ω . Maximize the posterior probability is equivalent to minimize the quadratic form $U(L) = -\log P(L|\{\tilde{Y}^i\}_i, \omega)$ whose gradient is given by [14]:

$$\nabla_L U(L) = \underbrace{\left(\sum_i \boldsymbol{\alpha}^{iT} \mathbf{P}^i \boldsymbol{\alpha}^i \right)}_{\boldsymbol{\alpha}^f} L - \underbrace{\sum_i \overbrace{\boldsymbol{\alpha}^{iT} \mathbf{P}^i \tilde{Y}^i}^{\Lambda^i} + 2\omega \mathbf{S}^T \mathbf{D}^T \mathbf{D}SL}_{\Lambda^f} \quad (5)$$

In Eqn. (5), Λ^f reminds us of the reconstruction step performed in the Drizzling method [16] except the normalization. Indeed, applying $\boldsymbol{\alpha}^{iT}$ leads to the reconstruction of weighted raw observations $\mathbf{P}^i \tilde{Y}^i$ in the model space to produce a fuzzy but geometrically consistent cube Λ^i . The final reconstructed fused image Λ^f is affected by the blur operator $\boldsymbol{\alpha}^f$.

A conjugate gradient method is used to solve $\nabla_L U(L) = 0$ which corresponds to the deconvolution of Λ^f by $\boldsymbol{\alpha}^f$, leads to an estimation of interpolation coefficients \hat{L} and then to the fused image \hat{X} . Uncertainties on \hat{X} are directly deduced from the Hessian of $U(L)$ as shown in [14].

3.3 Cosmic Ray Detection and Deletion

We propose a two-step method for the rejection of cosmic rays. The first step consists in detecting cosmic rays while computing Λ^f . For each model space location u , we consider the set $\{\tilde{\Lambda}_u^i\}_i$ where $\tilde{\Lambda}^i$ stands for the normalized version of Λ^i . Most of the cosmic rays (*e.g.* impact centers) are firstly detected in each $\{\tilde{\Lambda}_u^i\}_i$ with a simple thresholding method (deviation from the median) and thus ignored during the building of a clean $\bar{\Lambda}^f$ which leads to a first estimate \bar{L} through deconvolution.

The second step comes down to “simulate” the acquisition process from \bar{L} to generate near-cosmic free raw observations $\bar{I}^i = \boldsymbol{\alpha}^i \bar{L}$. Thus, each \bar{I}^i is compared, in the sensor space, to its associated original observation \tilde{Y}^i and neighborhood priors can then be defined (this part is still in progress) to improve the cosmic detection. Finally, elements of \mathbf{P}^i are set to 0 for each flagged pixels on \tilde{Y}^i and the reconstruction/deconvolution steps are performed again (see figure 2).

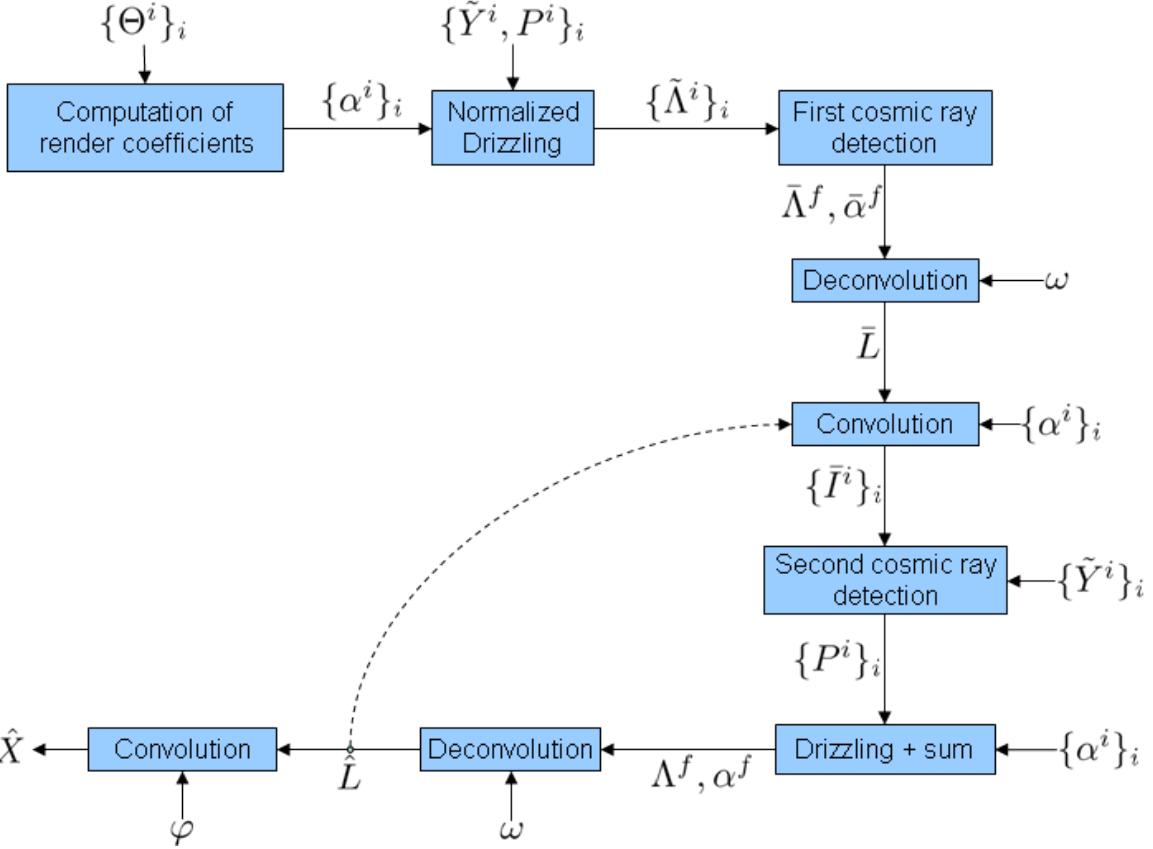


Figure 2. Fusion pipeline. The second cosmic ray step can be recursively performed if needed so as to refine the detection (dashed line).

4. PRELIMINARY RESULTS

4.1 Fusion Results

As neither real MUSE raw observations nor accurate simulations based on MUSE specifications are yet available, we developed a simple simulation tool derived from the image formation model described in the previous section for the validation of our fusion method. Spatial profiles of simulated astronomical objects contained in the scene are Gaussian (galaxies) and Dirac (stars) functions whereas spectral profiles are made up of a mixture of both (simulation of background and emission lines). This parametric astronomical scene can then be “observed” under various observation conditions (Gaussian FSF and LSF, sampling lattices, noise, shifts, missing values, cosmic rays, IFU number...) and the fusion of the resulting raw images is performed thanks to the knowledge of corresponding observation parameters.

To validate our fusion algorithm, we simulated 4 raw observations from an astronomical scene composed of 2 stars and 2 elliptical galaxies (see Fig. 3). The used acquisition parameters for each observation are detailed in table 1. The reconstruction size is $32 \times 32 \times 32$ pixels and the model space has been tiled with 4 equal-sized IFU, *i.e.* each raw observation is composed of four 1024×32 pixels CCD images. However, raw data can be easily rearranged in a (x, y, λ) cube for testing and display purposes because there is no spectral shift induced by IFU.

Figure 3 shows a part of the results for several selected frames (1, 13 and 32), star spatial profiles and spectra. The ideal image has been obtained with constant ideal LSF and FSF φ and is compared with our Bayesian fusion method and a simple B-Spline interpolation. Unlike interpolation techniques that only takes the geometry of the system into account without compensating the blur introduced by the acquisition step, the Bayesian inference

Obs.	FSF (λ_0)	FSF (λ_1)	LSF (λ_0)	LSF (λ_1)	Spatial shifts (x, y)	SNR (Star, Galaxy, Total)
1	1.4	1.96	1.8	1.9	(0, 0)	(57, 38, 44)
2	1.6	2.24	1.4	1.46	(1.2, 1.4)	(56, 38, 43)
3	1.4	1.96	1.4	1.46	(0.4, 0.5)	(57, 38, 44)
4	1.7	2.38	1.7	1.8	(0.2, 0.3)	(55, 38, 43)

Table 1. Acquisition parameters for each observation (spatial shifts in pixels, SNR in dB). FSF and LSF values correspond to FWHM in pixels whereas λ_0 (resp. λ_1) stands for the wavelength at the beginning (resp. end) of the model space. FSF and LSF size increases linearly along wavelength axis but remains spatially constant. Noise variances, pixel sizes and sampling steps remain constant for all sensors and neither spectral shift nor cosmic ray or sky background has been introduced.

takes care of both spectral and spatial blur. However, in practice, higher frequency content (*e.g.* stars) may be lost and cannot be recovered with the chosen prior (smooth objects like galaxies are less concerned by this issue). Compared to cubic interpolation, the restored objects with our Bayesian method are close to those obtained in the ideal case, especially for high band numbers (where LSF and FSF are the widest), and most of the blur effects have been correctly removed during the deconvolution step. The pixel size of the fused image may be chosen according to astronomer requirements so as to maximize both spectral and spatial resolutions or signal-to-noise ratio.

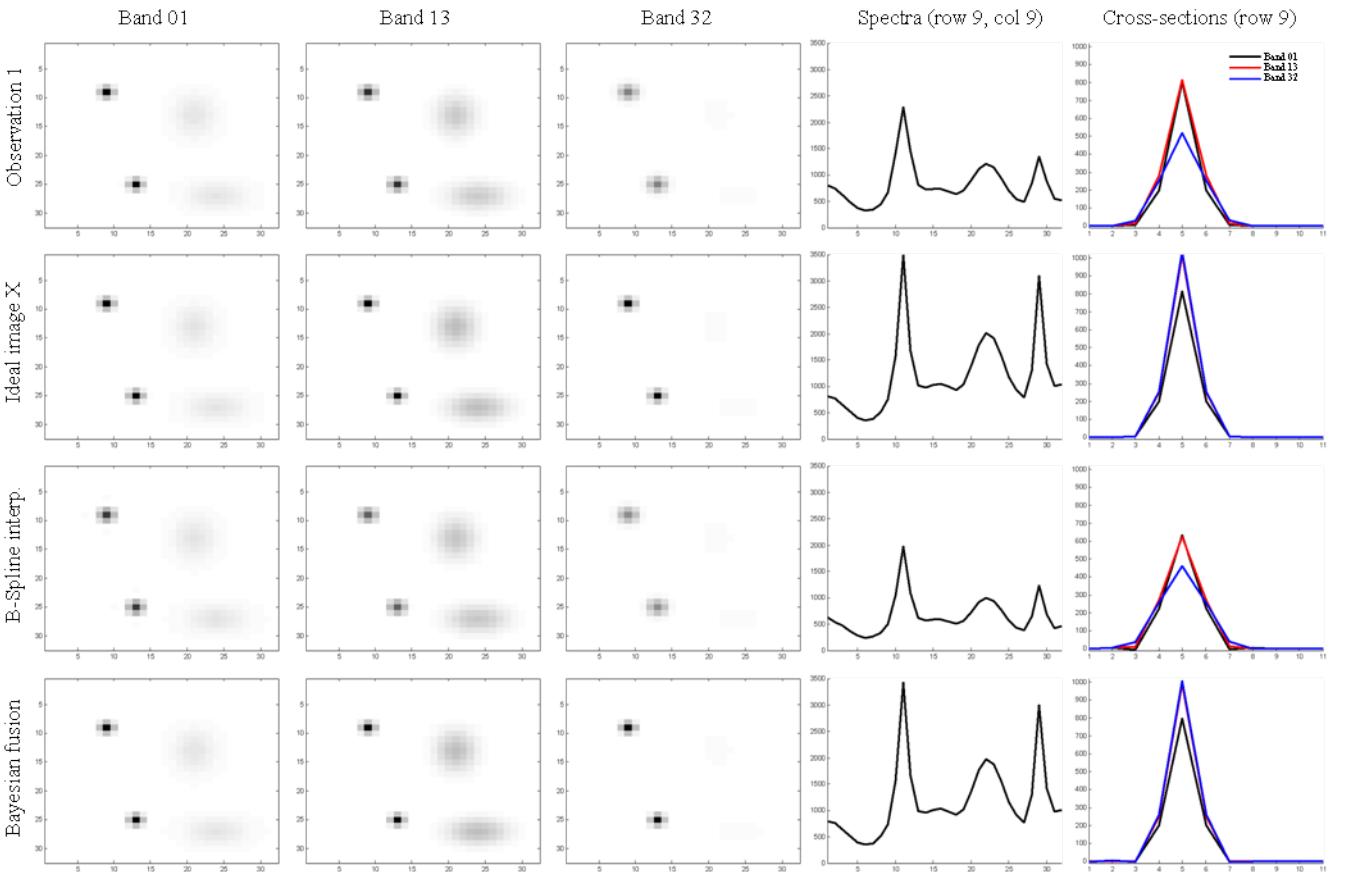


Figure 3. Results from simulated data (32 × 32 × 32 pixels): (x, y, λ) version of the raw observation 1, ideal image (blurred with φ), B-Spline interpolation and proposed Bayesian fusion results. Three bands (1, 13, 32) together with star profiles and spectra (same y-axis scale per column) are shown for each case.

Uncertainties associated to the Bayesian reconstruction are presented in the figure 4. In our results, they are

closely related to LSF and FSF and then only depend on λ (PSF are spatially constant) as the noise level is fixed. Indeed, variances are spatially invariant except for border effects due to truncations of blur kernels (left of fig. 4) whereas, spectrally, the covariance increase mainly reflects the enlargement of the spatial blur (right of fig. 4). Variances are then an estimation of the data quality and covariances encode the dependence between model coefficients. This strong correlation should be avoided by using a model with a wavelength-dependent spatial resolution but resulting fused images would not be easily handled.

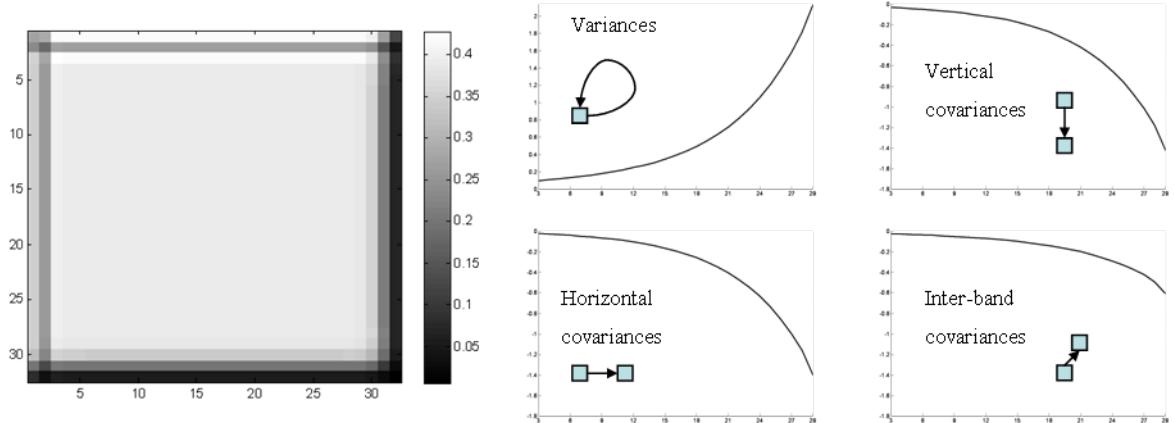


Figure 4. Covariances associated to the Bayesian fusion result. Left: variances for the band 16 (middle of the cube). Right: covariances along spectral axis for the location (16, 16).

4.2 Uncertainty Representation

Uncertainties are a significant and valuable output of the fusion algorithm and are available at each location in the reconstructed cube. In addition, variances induced by the acquisition process are also provided for each observation session.

The visualization of the behavior of uncertainties along spatial and spectral axes is essential as it may guide further investigations towards image areas or pixels having low uncertainties (*e.g.* for astrometry and photometry computations). Both coarse and fine views are interesting: a general overview allows to quickly localize high varying regions whereas pixel by pixel representation offers high resolution interpretation. The joint visualization variance/intensity data has been investigated along the spectral axis thanks to visualization modes available in QuickViz (see section 2.2) but the simultaneous spatial representation remains unsolved and problematic. Indeed, straightforward static visualizations such as colored or transparent compositions fail and instead tend to entangle both information. We have then selected two approaches based on a dynamical viewing process which is the most striking to the human eye.

The first strategy consists in using the third dimension z so that pixels (represented by 3D objects like spheres or cubes) oscillate more or less quickly around the (x, y) plane according to their variance: the highest the variance, the quickest the oscillation. The number of video frames to end a cycle is thus inversely proportional to the variance value. Such a 3D movie is relevant to instantaneously visualize high variance areas in spectral bands but requires the choice of an appropriate projection angle for an efficient perspective view. In addition, when the image size increases, it becomes more difficult to clearly distinguish swinging pixels. Thus, this representation should be preferred for limited regions or for the visualization of covariance values around a neighborhood. Two video examples are available at <http://lsit-miv.u-strasbg.fr/paseo/cubevisualization.php>.

In the second visualization, a map is created where the intensity of the pixel varies with the corresponding variance measurement. Thus, high variance positions will change their intensity quicker than low variance ones. Contrary to the first strategy, oscillation amplitudes are not spatially constant but are taken in the interval $[Y_p - x\sigma_p; Y_p + x\sigma_p]$ where Y_p is a pixel intensity with its associated standard deviation σ_p and $x \in \mathbb{R}_+$. Compared to the 3D method, this one is particularly well-suited for large images because an intensity variation is less confusing

than a height one and more attractive to the human eye, *i.e.* one can easily identify general trends in the image as illustrated in the movie available at <http://lsiit-miv.u-strasbg.fr/paseo/cubevisualization.php>.

The evaluation and validation of both methods with the help of the astronomical community is still in progress.

4.3 Dealing With Large Datasets

As mentioned previously, the dimensionality of MUSE hyperspectral cubes increases the complexity of visualizing, handling and processing of such data. In the Bayesian fusion framework, this becomes problematic as we deal with many additional and intermediate datasets throughout the fusion pipeline: observation parameters, rendering coefficients (up to 1TB per MUSE observation), α^f and covariance matrices... Algorithms have then to be designed and implemented in order to optimize memory usage while minimizing computationally expensive file access. Moreover, the development of such algorithms requires to define a high number of dedicated and optimized data structures (compound datatypes, variable-length and multidimensional arrays...) that need to be quickly saved and restored. Instead of choosing the FITS format (widely used by the astronomical community) for solving such problems, we have selected the hierarchical file format HDF5 [13] as the best candidate in terms of implementation, storage, optimized file access and complex datatype support. Finally, future QuickViz releases will support the HDF5 file management (see section 2.3) so as to avoid the use of FITS buffer files for the exploration of hyperspectral datasets.

5. CONCLUSION

In this paper, we presented a visualization software named QuickViz that extends the spectral features of Aladin and provides a set of functionalities designed to ease the exploration of hyperspectral data cubes, *e.g.* spectrum extraction, navigation coupling between spectral and spatial domains, customizable spectrum visualization and processing, synchronized multiview... We also presented a new method for the optimal fusion of hyperspectral raw observations in the Bayesian framework. This innovative approach takes all the observation parameters into account thanks to rendering coefficients and thus leads to a consistent fused image. The main steps of the fusion pipeline have been drawn and preliminary results validating the method are shown. This global fusion process opens the way to the detection of cosmic rays on CCD images where neighborhood priors can be defined. Finally, uncertainties over the resulting image are available thanks to the rigorous Bayesian approach and can be used as a prior for further investigations on the fused data. The visualization software has been firstly designed to apply to MUSE observations but can also tackle any kind of hyperspectral images if they follow the FITS extension data format. The fusion technique can also apply to most of hyperspectral datasets where raw images at the sensor level as well as acquisition parameters are given.

References

- [1] F. Laurent, F. Henault, E. Renault, R. Bacon, and J.P. Dubois. Design of an integral field unit for MUSE, and results from prototyping. *Publications of the Astronomical Society of the Pacific*, 118, 2006.
- [2] A. Jalobeanu and J.A. Gutiérrez. Multisource data fusion for bandlimited signals : a Bayesian perspective. *Proc. of 25th workshop on Bayesian Inference and Maximum Entropy methods (MaxEnt'06), Paris, France*, 2006.
- [3] A. Jalobeanu, J.A. Gutiérrez, and E. Slezak. Multisource data fusion and super-resolution from astronomical images. *Statistical Methodology*, 5(4), Jul 2008.
- [4] VO group ESAC Science Operations Department. VOSpec - VO Spectral analysis tool. <http://www.sclops.esa.int/index.php?project=ESAV0&page=vospec>.
- [5] A. Laruelo et al. VOSpec Spectral analysis tool. *ASP Conference Series, Proceedings of Astronomical Data Analysis Software and Systems (ADASS) XVII*, 394:513–516, 2008.
- [6] Starlink Project. SPLAT - Spectral Analysis Tool. <http://star-www.dur.ac.uk/~pdraper/splat/splat-vo/>.

- [7] M. Taylor et al. SAMP - Simple Application Messaging Protocol v1.11. <http://www.ivoa.net/Documents/cover/SAMP-20090421.html>, 2009.
- [8] Starlink project group. GAIA3D - Graphical Astronomy and Image Analysis Tool. <http://star-www.dur.ac.uk/~pdraper/gaia/gaia3d/index.html>.
- [9] F. Bonnarel et al. The Aladin interactive sky atlas, a reference tool for identification of astronomical sources. *Astron. Astrophys., Suppl. Ser.*, 143:33–40, 2000.
- [10] P. Fernique et al. Website of Aladin. <http://aladin.u-strasbg.fr/>.
- [11] M. Petremand et al. Website of Quickviz. <http://lsiit-miv.u-strasbg.fr/paseo/cubevisualization.php>.
- [12] Data model Working Group. International Virtual Observatory Alliance. <http://www.ivoa.net/cgi-bin/twiki/bin/view/IVOA/IvoaDataModel>.
- [13] The HDF group. Hierarchical data format. <http://www.hdfgroup.org/>.
- [14] A. Jalobeanu, M. Petremand, and C. Collet. Bayesian fusion of hyperspectral astronomical images. *Proc. of 30th workshop on Bayesian Inference and Maximum Entropy methods (MaxEnt'10), Chamonix, France*, 2010.
- [15] M. Unser. Splines : a perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16, 1999.
- [16] A.S. Fruchter and R.N. Hook. Drizzle : a method for linear reconstruction of undersampled images. *Publications of the Astronomical Society of the Pacific*, 114, 2001.