SOFTWARE TOOLS FOR RAPID ANALYSIS OF MASTCAM-Z MULTISPECTRAL DATA.

C. C. Million[1], M. St. Clair[2], M. Rice[3], A. Vaughan[4], [1]Million Concepts 1355 Bardstown Rd. #132 Louisville KY 40204 (chase@milliononcepts.com),[2]Million Concepts, [3]Western Washington University, [4]Arizona State University

Introduction: The Mars 2020 Mast Camera Zoom (Mastcam-Z) is a stereoscopic multispectral camera pair mounted on the Perseverance rover. Mastcam-Z observations generally consist of a sequence of stereo-paired 4-megapixel images taken through filters optimized for the Martian surface. [1,2] These data must be scientifically analyzed on short timeframes to inform tactical rover operations.

We have developed an ecosystem of tools for rapid analysis of Mastcam-Z multispectral observations. They generate a variety of views on the data and auto-populate tracking spreadsheets and databases, enabling scientists to easily access and rapidly understand newly-analyzed observations.

asdf: [a]utomated [s]pectral [d]ata [f]unction is a command-line utility that automates last-mile reduction of multispectral (meta)data. The basic command is, as the name suggests, little more than a keysmash: `asdf`. It automatically finds and processes multispectral data files, generating many views on the data: parameter maps, decorrelation stretches, “true” and enhanced color images, etc. If the user provides a "region of interest" (ROI) file (manually-selected areas for spectral extraction), `asdf` generates context images and graphs (e.g., Fig. 1). It also prompts users for ROI descriptors based on classifications generated by the Mastcam-Z Multispectral Working Group (MSWG). Each execution of `asdf` produces an interchange file containing observation/analysis (meta)data in a CSV format shared with our other tools.

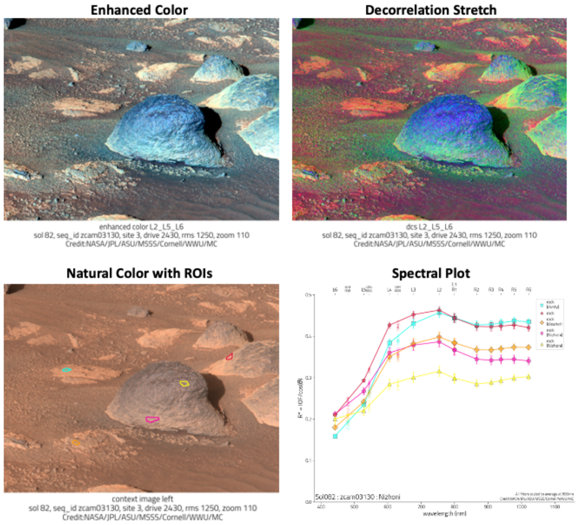
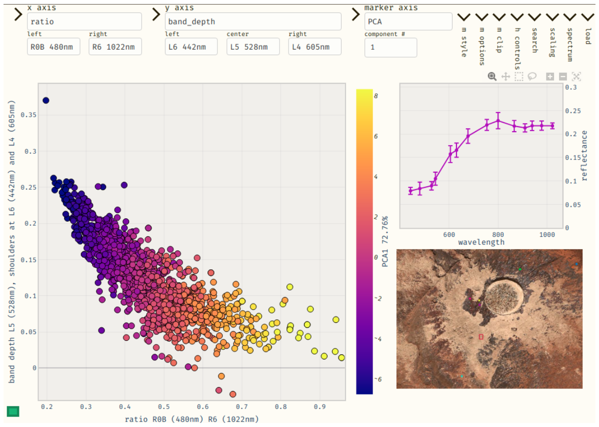

Enhanced Color
enhanced color L2_L5_L6
sol 82, seq_id zcam03130, site 3, drive 2430, rms 1250, zoom 110
Credit:NASA/JPL/ASU/MSSS/Cornell/WWU/MC
Decorrelation Stretch
dcs L2_L5_L6
sol 82, seq_id zcam03130, site 3, drive 2430, rms 1250, zoom 110
Credit:NASA/JPL/ASU/MSSS/Cornell/WWU/MC
Natural Color with ROIs
context image left
sol 82, seq_id zcam03130, site 3, drive 2430, rms 1250, zoom 110
Credit:NASA/JPL/ASU/MSSS/Cornell/WWU/MC
Spectral Plot
L6
L5
L4
L3
L2
L1
R1
R2
R3
R4
R5
R6
0.5
0.4
0.3
0.2
0.1
0.0
R* = IOF/cos(θ)
400
500
600
700
800
900
1000
1100
wavelength (nm)
Sol082 : zcam03130 : Nizhoni

Figure 1. Examples of `asdf` output products for the Nizhoni observation (sol 82, zcam03130).

`asdf` has a secondary mode called `fdsa` that can regenerate previous analyses with modified parameters. Use cases include calibration updates, refined ROI selections, and quality assurance checks on pipeline updates. We estimate that revising all prior analyses given an updated Mastcam-Z calibration will take about 10 minutes per year of rover operations (at current compute capacity). The plotting sub-utility `pretty-plot` is also independently supported in a Jupyter Notebook, allowing users to customize graphics for presentations and publications.


x axis: ratio
left: R0B 480nm
right: R6 1022nm

y axis: band_depth
left: L6 442nm
center: L5 528nm
right: L4 605nm

marker axis: PCA
component #: 1

Figure 2: Compact version of MultiDEx interface.

MultiDEx: The [Multi]spectral [D]ata [Ex]plorer is a graphical user interface (GUI) application allowing users to visualize arbitrary numbers of multispectral observations in enormously varied parameter spaces. It runs on a database built from CSV files generated by `asdf`. Users can generate two-axis scatter plots from a huge variety of qualitative and quantitative parameters; each point corresponds to a multispectral measurement [e.g., 3-7]. Many of these parameters are computed on the fly, including database-level operations such as principal component analysis (PCA). Color can also be arbitrarily mapped, acting as a third axis for the parameter space. Users can subsample the database using almost arbitrary search criteria; these subsamples can be independently highlighted or exported (in the CSV interchange format) for further analysis.

VISOR: The [V]isible and [I]nfrared [S]pectroscopy br[o]wse[r] is an online data portal (https://westernreflectancelab.com/visor/) for search, display, and retrieval of an ever-accumulating collection of laboratory and in-situ observations of materials in visible and infrared wavelengths. Any spectra can be quickly convolved to the bandpasses of a number of remote-sensing spectrometers (including Mastcam-Z) for direct comparison. Our standardized CSV interchange format permits users to add data exported from `VISOR` to any `MultiDEX` database instance, enabling them to rapidly identify and contextualize laboratory analogues.

Design Principles:

Flexible: Both `MultiDEX` and `VISOR` are extensible to other multi- or hyper-spectral data sets. The instrument-specific elements of `MultiDEX` are defined in external reference files; it is already in use for Mastcam [e.g., 8] and being explored for other instruments. Given bandpass specifications, a new instrument can be added to `VISOR's` convolution functionality in minutes.

While `asdf` is tailor-made for Mastcam-Z team practices, observational conventions, and compute environments, its support libraries (e.g. `marslab` and `dustgoggles`) are almost trivially extensible to any multispectral imager. These libraries instantiate best practices for working with multispectral image data—band parameter calculations, decorrelation stretches, radiometric scaling, etc.—some of which did not previously have well-tested open-source or Python-language equivalents. An "`asdf` for Mastcam" is also in development. Furthermore, `asdf` reads (meta)data and view definitions from user-modifiable reference files that can be readily updated given emerging mission requirements. These tools either currently are or will soon be fully open-sourced under a BSD 3-clause license. Please use them.

Reproducible: `asdf` records analyses in archive-ready formats (this is to say: in compliance with Planetary Data System v4 policies, absent valid labels). If a user passes `asdf` the `--upload` flag (which is required to add observations to the shared tracking table as part of the tactical workflow), it also uploads its records to multiple private repositories. These repositories are fully auditable, and analyses are fully reproducible (which can be verified with `fdsa`). We therefore not only have records of the "final" analysis for each observation (the uniqueness of which is dominated by selection and description of ROIs), but also a time-series of analyses by the operations team, providing thorough records of how scientific practices changed as the mission progressed.

Fast and Expressive: No particular view on data is optimal for all use cases, and the optimal view is not always obvious at the outset. On the other hand, the total space of possibly useful views on the data is finite, and humans—especially experts—are quick at parsing visual signals for utility. However, human time and attention are limited and should not be wasted. Generating views on the data was previously manual and time-consuming; we have made it fast and automated. The "exploratory" stage now merely requires flipping between views—which can be done at a rate of a few per second—to find those most applicable to the task at hand. Experienced users develop intuition for which views are most likely to be interesting and useful, which further cuts search time.

The software has also been heavily optimized for speed and leverages multiprocessing when appropriate, so it is very fast. A single run of `asdf` typically takes seconds. There is almost no perceptual lag to user requests to `MultiDEx` and `VISOR` (aided by pre-computation of some derived parameters at ingest).

Durable: The software needs to last a long time. Planetary mission support software has historically achieved exceptionally long useful lifetimes. (For example, `VICAR` might be the oldest continuously maintained image processing software in the world.) We expect the Mars 2020 mission to continue for years and *hope* it will continue for decades. The original software authors are unlikely to still be working on the project late in the mission, and scarce late-mission resources should not be burned maintaining bitrotted ground software. We therefore placed heavy emphasis on producing clean code with good test coverage, few dependencies, thorough documentation, and stable virtual environments. Simple user interfaces and multiple backups minimize chances for handoff problems as operations staff churns.

Assume Expert Users: Design and development benefited from critical observations of / assumptions about our users. Specifically: these tools are designed for computer-literate domain experts who are highly trainable and motivated to generate correct analyses. For instance, users are assumed to be comfortable on a Linux command line, which eliminates the need for superfluous GUIs. We are therefore sometimes able to offload software complexity (which generally reduces speed, flexibility, or durability) to procedures, documentation, or user-editable configuration files.

Acknowledgments: This work was supported by the NASA Mars2020 Rover Project.

References: [1] Bell, J. F., et al. *Space Science Reviews* 217.1 (2021): 1-40. [2] Hayes, A. G., et al. *Space science reviews* 217.2 (2021): 1-95. [3] Rice, M. S. et al., LPSC, *this issue*; [4] Horgan, B. et al., LSPC, *this issue*; [5] Nuñez, J. et al., LSPC, *this issue*; [6] Garczynski, B. et al., LPSC, *this issue*; [7] E Cardarelli, et al., LPSC, *this issue*; [8] Eng, A. et al., LSPC, *this issue*.