

IDCUBE LITE – A FREE INTERACTIVE DISCOVERY CUBE SOFTWARE FOR MULTI AND HYPERSPECTRAL APPLICATIONS

Deependra Mishra¹, Helena Hurbon^{1,4}, John Wang^{1,4}, Steven T. Wang^{1,4}, Tommy Du¹, Qian Wu¹, David Kim¹, Shiva Basir¹, Qian Cao¹, Hairong Zhang¹, Kathleen Xu¹, Andy Yu¹, Yifan Zhang¹, Yunshen Huang¹, Roman Garrett¹, Maria Gerasimchuk-Djordjevic³, Mikhail Y. Berezin^{1,4*}

¹Department of Radiology, Washington University School of Medicine, 4515 McKinley Ave, St. Louis, MO 63110; ²Department of Computer Science and Engineering, Washington University, 1 Brookings Hall, St. Louis, MO 63110; ³Art and Design Department, Missouri State University, 901 S. National Ave, Springfield, MO 65897, ⁴HSpeQ LLC, 4340 Duncan Ave, St. Louis, MO 631

ABSTRACT

Multi- and hyperspectral imaging modalities encompass a growing number of spectral techniques that find many applications in geospatial, biomedical and machine vision fields. The rapidly increasing number of applications requires a convenient easy-to-navigate software that can be used by new and experienced users to analyze data, develop, apply, and deploy novel algorithms. Herein, we present our platform, IDCube that performs essential operations in hyperspectral data analysis to realize the full potential of spectral imaging. The strength of the software lies in its interactive features that enable the users to optimize parameters and obtain visual input for the user. The entire software can be operated without any prior programming skills allowing interactive sessions of raw and processed data. IDCube Lite, a free version of the software described in the paper, has many benefits compared to existing packages and offers structural flexibility to discover new hidden features.

Index Terms— Hyperspectral, Multispectral, Imaging software, Visualization tool, Spectral imaging

1. INTRODUCTION

Multi and hyperspectral imaging (HSI) modalities have emerged as an exciting opportunity to explore optical properties of objects and discover hidden features not accessible by other techniques. In contrast to the traditional spatial images produced by conventional cameras, spectral imaging generates 3D datasets (datacubes), with spatial and spectral dimensions. With each pixel containing information of the entire medium or high-resolution spectrum, spectral imaging provides abundant information about individual chromophores and their interactions that contribute to the location, intensity, and alteration of the optical signal, significantly better than monochromatic or traditional color cameras [1, 2]. This spectral imaging approach leads to a vastly improved ability to classify and differentiate the objects based on their spectral features, enabling even small, otherwise unnoticeable, features to be amplified.

Herein, we present IDCube, Interactive Discovery Cube, a highly versatile software that performs a large number of essential operations in the spectral imaging domain and enables image analysis for users across a range of technical proficiencies. The strength of the proposed software lies in its intuitive design that enables the user to perform high level data analysis as well as develop their own algorithms via a visual, interactive interface. Built around a collection of spectral imaging algorithms, the software facilitates the search of hidden information inside large datasets providing a new experience of data analysis. The entire program can be operated without prior programming skills and allows almost any currently used data formats to be processed. The front end of the software is shown in Figure 1.

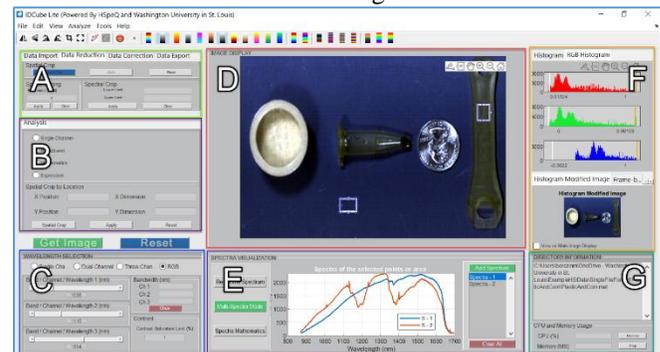


Figure 1 Front end of the IDCube Lite hyperspectral software. **A:** import/export, data reduction, data correction; **B:** image algebra selection with preselected and user-developed functions; **C:** wavelength and bandwidth selection, image optimization; **D:** image visualizer, **E:** spectral analysis of selected ROIs; **F:** image enhancement via histogram manipulation; **G:** information about the image, the consumed computational resources for the performed tasks.

The core of the software is written in MATLAB with an extensive number of built-in image processing functions. The compiled software can be run on any computer with a Windows or Mac Operating System, without having a MATLAB license. An entirely free version of the software

IDCube Lite is available for download from www.idcubes.com from a secure cloud location. An extensive library of video tutorials is available from the learning module of the software website. Below we will focus on the technical capabilities of IDCube Lite and demonstrate its performance through several experiments in the geospatial, machine vision, and biomedical fields.

2. FEATURES

IDCube Lite enables visual analysis of many datasets from a variety of formats. The platform is constantly improving since its first launch in September 2020 and includes more than fifty integrated algorithms for data processing and data visualization. These algorithms are grouped into the following categories: Input/Output, Data Reduction, Image Visualization, Spectral Analysis, PCA/MNF, Segmentation, and Spectral Matching. The selected algorithms are optimized for processing time and do not exceed several seconds when processing a 1 GB file on a standard desktop computer. The descriptions of the major features are given below.

2.1 Input/Output module

IDCube Lite supports a wide range of image formats, including Raw/HDR, DAT, TIFF, JPEG2k, BMP, PNG, and several others. The imported files (datafile+header file) are first converted then saved in the same directory as a single MATLAB-type file. The single file that combines raw data and the channel assignment information (wavelength vector) automatically opens in the IDCube interface after the conversion is complete. IDCube Lite supports imaging data from a variety of bench-type HSI systems that utilize a Raw/HDR format as well as from satellites offering data in JP2 and TIFF. In addition to HSI datasets, the software can also treat standard RGB images and thus presents an interesting opportunity to apply sophisticated computational HSI techniques to common pictures.

The software can open up to 10 hyperspectral datasets simultaneously for all datasets with the same dimensions. If the images have different spatial dimensions, they can be first cropped using the spatial cropping function. Analyzing multiple datasets in one setting enables data comparison and the same processing for data from longitudinal studies.

2.2 Preprocessing module

The preprocessing module includes the ability to perform binning in the spatial domain to decrease the size of the file and reduce noise. The module offers interactivity through spatial cropping, spectral cropping, and flipping/transposition/rotation. All preprocessing functions are applied globally to the entire dataset. DATA CORRECTION tool removes undesirable artifacts caused by scattering of light that produces undesirable artifacts. The correction algorithms include multiplicative scattering

correction (MSC)[3] and standard normal variation (SNV)[4] functions. This feature is especially useful for biomedical imaging, for example, to decrease scattering artifacts from skin. Other preprocessing steps to decrease the size of the image file via removal of highly correlating spectral bands and extracting endmember signatures from hyperspectral data[5] will be implemented in future versions.

2.3 Visualization module

This module is central to image processing and presents three-dimensional (3D) datasets through a set of two dimensional (2D) images. The 2D images are generated through the selection up to three wavelength channels. The wavelength selection can also be appended with a preselected bandwidth rather than the default bandwidth of 1. The produced monochromatic image can be color optimized by applying an appropriate lookup table (LUT) from more than 20 available LUTs. The visualization is further adjusted via the histogram tool. BROADBAND function enables visualizing the image from a range of wavelengths. MATHEMATICS mode in conjunction with TWO-CHANNEL modes enables the user to conduct image algebra. EXPRESSION module allows the user to type custom mathematical function or select from the list of implemented functions specified in the Mathematics Sheet for Expressions available from the software website <https://www.idcubes.com/tutorials>.

SINGLE CHANNEL, TWO-CHANNEL and THREE-CHANNEL tools allow users to scroll the images through the individual wavelengths. RGB mode enables the user to combine up to three wavelengths into a pseudo RGB image. Each wavelength can be used with the specific bandwidth. HISTOGRAM optimization and CONTRAST ADJUSTMENT fields enable the user to improve the image. One of the unique features of IDCube is to present the dataset as a movie where each frame represents an image at a specific wavelength or with a formula applied. This approach radically minimizes the amount of user interaction and prevents image processing fatigue. The image with the entire dataset can be flipped, transposed, and rotated. The produced 2D image can be also copied, zoomed, panned, and saved.

2.4. Spectral Analysis Module

The spectral analysis module enables the user to visualize and process spectral information from individual pixels and interactively select regions of interests (ROIs). In REAL-TIME spectral mode, the module enables visualization of spectra from individual pixels by moving the cursor over the image. The module automatically recognizes a spectral range and scales the dimensions of the spectral plot. In MULTI-SPECTRA mode, the user is prompted to select one or more ROIs. The spectrum for each ROI reflects the average spectrum across the selected area. SPECTRA MATHEMATICS enables the user to perform basic math

functions, i.e., subtraction and division of the spectra, spectra normalization, and calculations of the first and second derivatives. The user can compare the spectral output from the selected ROIs using spectral correlation, spectral information divergence [6], or spectral angle [7] functions. Relatively high spectral correlation values, low divergence, and low spectral angles suggest regions with similar spectral properties indicating the objects belong to the same class of objects with similar optical profiles or same materials. All spectra can be zoomed, panned, and exported to Excel or other data analysis software.

2.5. Principal Component Analysis (PCA) module

The PCA module (**Figure 2**) computes associations between data points and converts a dataset of potentially correlated variables into a new set of linearly uncorrelated principal components [8, 9]. Up to three principal components can be selected by the user to generate a pseudo-color RGB image, where the selected components are assigned to three colors, red, green, and blue. Objects with the same color indicate high similarity between two subjects. For example, a centrifuge tube and a wrench shown in **Figure 2** are apparently made from the same material, since their PCA based colors are almost identical. The pseudo-color RGB image can be further adjusted through changing the WEIGHT of the individual component, adjusting the CONTRAST to the whole image, and applying GAMMA CORRECTION. In general, the total number of principal components is equal to the number of wavelength channels. Since most of the information is in the first few components, the IDCube Lite version limits the number of stored principal components to twenty. COMPONENT SPECTRA window enables the user to examine these components visually to check if relevant features may be extracted. The module also presents the cumulative fraction of variance. For example, the first two components carry >96% of the variance. PCA transforms the original datacube into a new datacube with the same spatial dimensions and changes the Z-axis from wavelengths to principal component scores. In that case, PCA can be also considered as a function that decreases the size of the data, since only very few principal components (first 20, for example) can be used. The new, smaller datacube can be exported back to the VISUALIZATION panel and analyzed using implemented algorithms.

A variation of the PCA method is a Maximum Noise Fraction (MNF) [10]. The MNF transform is advantageous over the PCA transform because it takes the noise information in the spatial domain into consideration. For example, the shadows seen in **Figure 2A** can be removed to some extent using MNF function implemented in IDCube Lite as shown in **Figure 2B**.

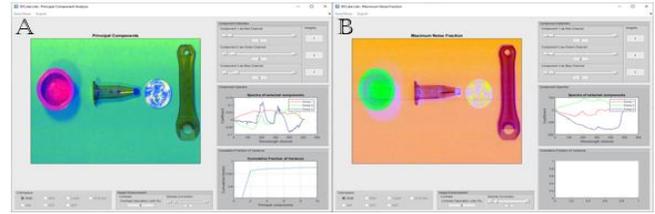


Figure 2 Preprocessing tools. **A.** Principal Component Analysis of a hyperspectral dataset. First, second, and third principal components are combined in a pseudo-color RGB image. The module also presents the spectra of selected components and a cumulative fraction of variance. The image can be improved by adjusting the contrast and gamma correction. **B.** Maximum Noise Fraction of the same data shows the partial removal of the shadow.

2.6. Segmentation module

The segmentation module in IDCube Lite (**Figure 3**) enables minimally supervised classification of a dataset from any of the pre-processing algorithms (spatial, spectral cropping, binning, PCA, etc.). In a typical workflow, the user selects one of the classification algorithms (i.e., based on the improved Spectral Angular Mapper (SAM) currently implemented in IDCube Lite) and the metrics (i.e., area, perimeter) then draws an area (class) on the image passed from the Visualization module. The areas with similar spectral properties (that have low values of spectral angle) are represented by the same color and quantified according to the selected metrics. The example shown in **Figure 3** illustrates this method for classification. One can notice that the method is highly sensitive to even small spectral changes, where the vial and the wrench can be separated even though their spectral correlation value is 0.99 (as measured using SPECTRAL ANALYSIS module, see above).

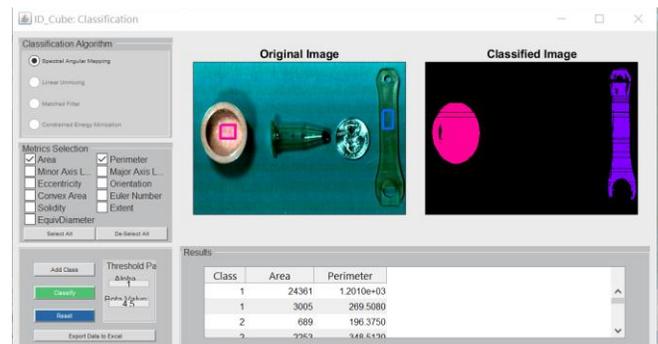


Figure 3 Image Classification Module. In the Spectral Angular mapping, the user interactively selects different areas (classes) from the 2D image. Although there is no limit of the number of classes, the module works best for one or two classes.

3. EXAMPLES

3.1. Biomedical Applications

The example shown in **Figure 4** illustrates the utility of IDCube to better visualize blood vessels. First, the acquired dataset was spectrally cropped to eliminate the noisy wavelength channels. Since InGaAs detectors used in this study typically have lower sensitivity below 950 nm and above 1700 nm, these wavelengths were excluded. The RGB mode was then selected and the wavelengths were manually adjusted to visualize the blood vessels. The image made with a conventional color camera does not show the blood vessels (**Figure 4A**). The visualization of the resulting image was further improved by adjusting corresponding color band histograms. The resulting image shown in **Figure 4B** presents blood vessels in a greater contrast than the visible image. The dataset was then treated with the MNF function selected from the ANALYSE tab. Similar to the PCA module described above, the software enables the user to select individual components and presents them in the pseudo RGB format (**Figure 4C**). High contrast for blood vessels was achieved by using components #2 and #3.

In addition to HSI, IDCube can handle other spectral imaging modalities commonly used in preclinical and clinical studies, such as Raman, FTIR, and FLIM.

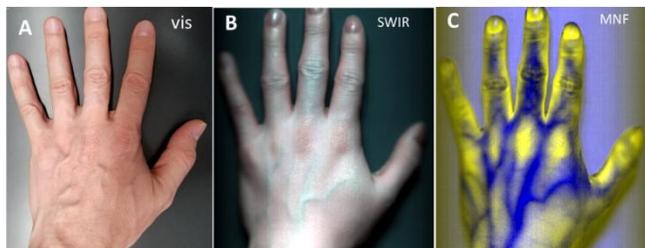


Figure 4 Hyperspectral imaging of a hand. **A**: made by a conventional visible camera; **B**: using a hyperspectral SWIR imager. Pseudo RGB image at 1070 nm (red), 1260 nm (green), 1320 nm (blue); **C**: MNF function applied to the HSI dataset. Pseudo RGB image at #3 (red), #3 (green), #2 (blue) components.

3.2. Environmental Applications

HSI of plants provides solutions to a large number of challenges from identifying environmental issues to monitoring crops yield and diseases [11] and even detection of contaminations [12] and minerals. The optical signature of plants and especially leaves are an important monitoring and predictive parameter for a variety of biotic and abiotic stresses. **Figure 6** illustrates an application of IDCube on a dataset from leaves with different moisture levels were analyzed. The right leaf on each image belongs to a plant grown under normal conditions, and the left leaf was exposed to a drying element. This treatment was used to mimic a

drought condition in order to showcase the effectiveness of the index. **Figure 5A** shows an image recorded by a conventional visible camera. **Figure 5B-D** show processed images recorded using a SWIR hyperspectral imager in the reflection mode. Low signal/noise ratio bands were removed using a spectral cropping function. A pseudo RGB image composed from three wavelengths shows the difference between the two leaves (**Figure 5B**). The contrast between two leaves can be further enhanced with PCA. Three selected principal components #1/2/3 were used in a pseudo RGB format as red, green, and blue (**Figure 5C**). Combined in a single image, these components highlight the difference between the two leaves. Even higher contrast can be achieved using a previously developed index of drought using the formula: $I = (1529 \text{ nm} - 1416 \text{ nm}) / (1519 \text{ nm} + 1416 \text{ nm})$ [13] (**Figure 5D**). This can be achieved by selecting Mathematics from the ANALYSIS section, then selecting Michelson Ratio and finally selecting two wavelengths 1416 nm and 1529 nm.

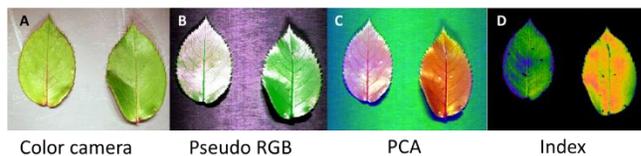


Figure 5 Hyperspectral imaging of leaves with IDCube Lite. **A**: image of two similar leaves obtained with a conventional color camera; **B**: contrast improvement with a three band approach in a pseudo RGB image: 1421 nm (red), 1351 nm (blue), and 1476 nm (green); **C**: PCA with three components #1 (red), #2 (green), #3 (blue) in a pseudo RGB image; **D**: Monochromatic image reflecting a drought index $(1529 \text{ nm} - 1416 \text{ nm}) / (1519 \text{ nm} + 1416 \text{ nm})$.

3.3. Geospatial Applications

Geospatial applications are one of the mature applications of hyperspectral imaging due to its relatively long history beginning in the late 1980s. Since then, a great number of platforms based on satellites, planes and, recently, drones have been developed. IDCube can be used on any of these platforms. The current version of the software enables data processing from hyperspectral and multispectral satellites such as ER-01 Hyperion [14], Sentinel-2 [15], Orbital satellites [16], airborne systems such as AVIRIS [17, 18], and other platforms.

An example of this workflow is shown in **Figure 6**, where the original dataset was first downloaded from a commercial vendor (Apollo Hunter), converted to the IDCube format, and processed to produce an RGB image using the first three bands (**Figure 6A**). The dataset was then processed by PCA. For better visualization of the objects of interest three principal components were used to construct a pseudo-RGB image (**Figure 6B**). A PCA-based datacube was further classified using a SAM method by selecting road as

endmember spectra to generate an image of roads and streets (**Figure 6C**).

The IDCube Lite version does not provide correction of the data due to atmospheric and other conditions. These functions will be implemented in IDCube Geo.

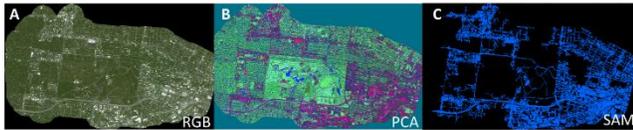


Figure 6 Multispectral image of the St. Louis area with four bands: RGB + NIR: 430-550 nm (blue), 490-610 nm (green); 600-720 nm (red); 750-950 nm (NIR). **A:** RGB image; **B:** Principal Component Analysis (PCA); **C:** Spectral Angular Mapping (SAM) with one class selected. Satellite Sensor *Pleiades-1B (0.5 m)* operated by AIRBUS Defense & Space. The data were acquired from *Apollo Hunter*.

5. ACKNOWLEDGEMENT

The team would like to acknowledge funding from National Science Foundation, NSF 1827656 (MB) and NSF 1355406 (MB, RG) and Mallinckrodt Institute of Radiology (HH). The original standalone package was licensed by HSpEQ LLC from Washington University where the software was previously developed and modified by HSpEQ LLC.

6. REFERENCES

[1] Lu, G., and Fei, B.: 'Medical hyperspectral imaging: a review', *Journal of biomedical optics*, 2014, 19, (1), pp. 10901

[2] Jahr, W., Schmid, B., Schmied, C., Fahrbach, F.O., and Huisken, J.: 'Hyperspectral light sheet microscopy', *Nature communications*, 2015, 6, pp. 7990

[3] Candolfi, A., De Maesschalck, R., Jouan-Rimbaud, D., Hailey, P., and Massart, D.: 'The influence of data pre-processing in the pattern recognition of excipients near-infrared spectra', *Journal of Pharmaceutical and Biomedical Analysis*, 1999, 21, (1), pp. 115-132

[4] Grisanti, E., Totska, M., Huber, S., Krick Calderon, C., Hohmann, M., Lingenfeller, D., and Otto, M.: 'Dynamic localized snv, peak snv, and partial peak snv: Novel standardization methods for preprocessing of spectroscopic data used in predictive modeling', *Journal of Spectroscopy*, 2018, 2018

[5] Chang, C.-I., and Plaza, A.: 'A fast iterative algorithm for implementation of pixel purity index', *IEEE Geoscience and Remote Sensing Letters*, 2006, 3, (1), pp. 63-67

[6] Chang, C.-I.: 'An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis', *IEEE Transactions on information theory*, 2000, 46, (5), pp. 1927-1932

[7] Kruse, F.A., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P., and Goetz, A.: 'The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data', in Editor (Ed.) (Eds.): 'Book The spectral image processing system (SIPS)-interactive visualization

and analysis of imaging spectrometer data' (American Institute of Physics, 1993, edn.), pp. 192-201

[8] Zuendorf, G., Kerrouche, N., Herholz, K., and Baron, J.C.: 'Efficient principal component analysis for multivariate 3D voxel-based mapping of brain functional imaging data sets as applied to FDG-PET and normal aging', *Human brain mapping*, 2003, 18, (1), pp. 13-21

[9] Rodarmel, C., and Shan, J.: 'Principal component analysis for hyperspectral image classification', *Surveying and Land Information Science*, 2002, 62, (2), pp. 115

[10] Green, A.A., Berman, M., Switzer, P., and Craig, M.D.: 'A transformation for ordering multispectral data in terms of image quality with implications for noise removal', *IEEE Transactions on geoscience and remote sensing*, 1988, 26, (1), pp. 65-74

[11] Lowe, A., Harrison, N., and French, A.P.: 'Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress', *Plant Methods*, 2017, 13, (1), pp. 80

[12] Manley, P.V., Sagan, V., Fritschi, F.B., and Burken, J.G.: 'Remote Sensing of Explosives-Induced Stress in Plants: Hyperspectral Imaging Analysis for Remote Detection of Unexploded Threats', *Remote Sensing*, 2019, 11, (15), pp. 1827

[13] Kim, D.M., Zhang, H., Zhou, H., Du, T., Wu, Q., Mockler, T.C., and Berezin, M.Y.: 'Highly sensitive image-derived indices of water-stressed plants using hyperspectral imaging in SWIR and histogram analysis', *Scientific reports*, 2015, 5, pp. 15919

[14] Pearlman, J.S., Barry, P.S., Segal, C.C., Shepanski, J., Beiso, D., and Carman, S.L.: 'Hyperion, a space-based imaging spectrometer', *IEEE Transactions on Geoscience and Remote Sensing*, 2003, 41, (6), pp. 1160-1173

[15] Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., and Martimort, P.: 'Sentinel-2: ESA's optical high-resolution mission for GMES operational services', *Remote sensing of Environment*, 2012, 120, pp. 25-36

[16] <https://www.obtdata.com/en/>

[17] Green, R.O., Eastwood, M.L., Sarture, C.M., Chrien, T.G., Aronsson, M., Chippendale, B.J., Faust, J.A., Pavri, B.E., Chovit, C.J., and Solis, M.: 'Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)', *Remote sensing of environment*, 1998, 65, (3), pp. 227-248

[18] Hamlin, L., Green, R., Mouroulis, P., Eastwood, M., Wilson, D., Dudik, M., and Paine, C.: 'Imaging spectrometer science measurements for terrestrial ecology: AVIRIS and new developments', in Editor (Ed.) (Eds.): 'Book Imaging spectrometer science measurements for terrestrial ecology: AVIRIS and new developments' (IEEE, 2011, edn.), pp. 1-7