MultiSpec© – A MULTISPECTRAL IMAGE DATA ANALYSIS SYSTEM

MultiSpec (©Purdue Research Foundation) is a processing system for interactively analyzing Earth observational multispectral image data such as that produced by the Landsat series of Earth satellites and hyperspectral image data from current and future airborne and spaceborne systems such as AVIRIS. The primary objective of MultiSpec is as an aid to export the results of our research into devising good methods for analyzing such hyperspectral image data. It has also found significant use in other applications such as multiband medical imagery and in K-12 and university level educational activities. There are currently in excess of several thousand known, registered users.

MultiSpec satisfies the following design goals:
- The implementation should be on a readily available platform which has adequate processing power, but which is financially within the reach of any remote sensing user (i.e., ≤ $5000).
- The system should be easy to learn and easy to use, even for the infrequent user, using the most modern of software environments.
- The system should provide for easy import of data in a variety of formats, and easy export of results, both in thematic map and in tabular form.

The work of building the current capability began by implementing an upgraded version of the LARSYS multispectral image data analysis system. LARSYS is one of the first remote sensing multispectral data processing systems, originally created during the 1960's. A number of systems in government laboratories, university research labs, and several commercially offered products are descendants of LARSYS. The current system, called MultiSpec, has been implemented for the Apple Macintosh personal workstation. Versions are available for PowerPC based machines in native code and for 680x0 based machines both with and without a math coprocessor. A PC-Windows version is also available, although it does not yet have all the capabilities of the Macintosh version. A color display capability with at least an 8 bit color range is highly desirable but not necessary; a 24 bit color system provides even more attractive performance. A reasonably current generation, middle range Macintosh and color display, would have a street price of less than $2000 at the present time. Such a system is capable of classifying in excess of 6 million pixel-classes per minute using 12 bands and a Gaussian maximum likelihood scheme. Given the current cost/performance trends, even more cost-effective systems are likely to be available in the future.

New capabilities are continually added to MultiSpec as they emerge from our research on hyperspectral processing. Capabilities of the current version of MultiSpec (released 4/2/99) include the following.

- **Import data** in either Binary or ASCII format with or without a header, and in Band Interleaved by Line (BIL), Band Sequential (BSQ), or Band Interleaved by Sample (BIS) formats. The data may have either one or two bytes per data value, and may have 4 to 16 bits per data value. In the case of two bytes per sample, the two bytes may be in either order.

- **Display multispectral images** in a variety of B/W or color formats using linear or equal area gray scales; display (internally generated) thematic images also in B/W or color, with an ability to control the color used for each theme.

- **Histogram** data for use in determining the gray scale regime for a display or for listing and graphing.

- **Reformat** the data file in a number of ways, e.g., by adding a standard header, changing from any one of the three interleave formats to either of the other two, editing out channels, combining files, adding or modifying channel descriptions, mosaicing data sets, changing the geometry of a data set, and a number of other changes.

- **Create new channels** of data from existing channels. The new channels may be the result of a principal components or feature extraction transformation of the existing ones, or they may result from the ratio of a linear combination of existing bands divided by a different linear combination of bands.

- **Cluster** data using either a single pass or an iterative (isodata) clustering algorithm. Save the results for display as a thematic map. Cluster statistics can also be saved as class statistics. Initial cluster centers may be pre-specified or established by one of several algorithms. Use of clustering followed by ECHO spectral/spatial classification provides an effective multivariate scene segmentation scheme.

- **Define classes** via designating rectangular, polygonal training fields or mask image files, compute field and class statistics, and define test fields for use in evaluating classification results quantitatively. A feature called "Enhance Statistics" also allows one to improve the extent to which the defined class statistics fit the composite of all data in the data set.

- **Determine** the **best spectral features** to use for a given classification using (a) searching for the best subset of features using any of five statistical distance measures, (b) a method based directly upon decision boundaries defined by training samples, or (c) a second method based directly upon the discriminant functions. Also included are methods especially designed to search for narrow spectral features such as spectroscopic characteristics, and for use of projection pursuit as a means of further improving the features extracted.

- **Classify** a designated area in the data file. Six different classification algorithms are available: use of minimum distance to means, correlation (SAM), matched filter (CEM), Fisher linear discriminant, the Gaussian maximum likelihood pixel scheme, or the ECHO spectral/spatial classifier. Save the results for display as a thematic map, with or without training and test fields being shown. Apply a threshold to a classification, and generate a probability or threshold map showing the degree of membership of each pixel to the class to which it was assigned.

- **List classification results** of training or test areas in tabular form on a per field, per class, or groups of classes basis.

- Show a **graph** of the **spectral values** of a currently selected pixel or the mean ± one standard deviation for a selected area. Show **scatter diagrams** of data from pairs of bands and ellipses of concentration for training sets and selected areas. Show a graph of the histograms of the class or field data values used for training. Show the coordinates of a currently selected area.

- Show a **color presentation** of the **correlation matrix** for a field or class as a visualization tool especially for hyperspectral data.

- Several additional **utility functions** including listing out a subset of the data e.g., for use externally, conducting principal component analysis, etc.

- **Transfer** intermediate or final **results**, be they text, B/W image or color image, to other application programs such as word processors, spreadsheet, or graphics program by copying and pasting or by saving and then opening the saved file within another application.

The MultiSpec implementation is carried out in such a way that the primary limit on the number of lines or columns of the data, the number of spectral bands, etc., are those determined by the available disk and memory space. Taken together, these capabilities provide a state-of-the-art capability to analyze moderate and high dimensional multispectral data sets of practical size. All versions, along with a 80+ page document listing its capabilities in more detail and providing tutorial exercises in its use, is available, along with substantial additional documentation, via the World Wide Web at the following URL:

http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/