Release Notes for RHSEG version 1.20, September 15, 2006

James C. Tilton, Mail Code 606.3, NASA GSFC
James.C.Tilton@nasa.gov

This is a significant rewrite of the RHSEG and HSEGViewer code. The most significant improvements are the upgrading of RHSEG to process three-dimensional data, and the upgrade of HSEGViewer to view and interact with two-dimensional planes of three-dimensional data, and the implementation in RHSEG of additional dissimilarity criteria. There are additional minor changes as noted below.

Since the three-dimensional code imposes measurable overhead when processing two-dimensional data, separate two-dimensional and three-dimensional versions are provided: rhseg_2d and rhseg_3d. The rhseg_3d version will process two-dimensional data (or one-dimensional) data, but it will run slower than the rhseg_2d version. An additional parameter "nslices" is provided to specify the size of the third dimension for rhseg_3d. This parameter is ignored by rhseg_2d.

Four new dissimilarity criteria are provided in this version of RHSEG (rhseg_3d and rhseg_2d). These new criteria are the "Spectral Angle Mapper" criterion, the "Spectral Information Divergence" criterion, the "Normalized Vector Distance" criterion and the "Entropy" criterion. The "Spectral Angle Mapper" and Spectral Information Divergence" criteria are often used in the analysis of hyperspectral data, while the "Normalized Vector Distance" criterion is useful in the analysis of color (e.g., RGB), multispectral and hyperspectral data. The "Entropy" criterion is original to RHSEG (at least I haven't seen it implemented anywhere else). It is useful in the analysis of grey scale, color and multispectral data - and perhaps hyperspectral data as well.

New input parameters include "scale" and "offset." These new parameters were added to facilitate the input of MODIS data into RHSEG. The MODIS multispectral data are normally stored in scaled short integer format (16 bit), with scale and offset factors provided to rescale the data into calibrated reflectance and/or radiance values.

The "conn_type" parameter has been modified appropriately to properly specify neighborhood connectivity for one-dimensional and three-dimensional data, in addition to two-dimensional data.

HSEGViewer has been augmented to accommodate the "nslices" parameter from the RHSEG analysis of three-dimensional data. However, HSEGViewer can be only used to visualize and interact with two-dimensional planes from the three-dimensional RHSEG analysis results.

RHSEG now orders the region labeling in an approximate order from "darker" to "brighter" regions (instead of the previous ordering of largest to smallest region).

HSEGViewer now attempts to color the "Segmentation Slice View" image with darker colors for regions with smaller region label values and with brighter colors for regions with larger region label values, according to a fix, non-random, color table scheme (previous versions of HSEGViewer used randomized colors).

Release Notes for RHSEG version 1.10, August 15, 2005

James C. Tilton, Mail Code 606.3, NASA GSFC
James.C.Tilton@nasa.gov

This is a significant rewrite of the RHSEG and HSEGViewer code. The most significant improvements are the implementation in RHSEG of a scheme for swapping portions of the data to temporary disk files and an addition of an option for selecting between different convergence criterion in RHSEG.

The swapping scheme is based on the scheme used in the parallel implementation to distribute the data out to processing nodes. This makes it possible to process larger data sets, since the program was previously memory bound for larger images. See the updated User's Guide for instruction on how to use this added RHSEG feature through the new "ionb_levels" parameter. The default value of this parameter should work well for most single CPU computer systems, assuming sufficient space is provided in the specified temporary disk space.

The major impact of this improvement is that large images can now be processed with the sequential version of RHSEG - if one is willing to devote sufficient processing time to the task. Here are the processing times I obtained for a 6912-column, 6528-row, 6-band Landsat TM image using program default values (except as noted):

spclust_wght	Parallel timing	Sequential timing	Parallel speed-up*
0.0	1.4 minutes	8.0 hours	171
0.1	7.5 minutes	43.3 hours	173
1.0	2.5 minutes	15.0 hours	180

*Taking into account differences in clock speed

The recommended optional "rmeanlist," "rconv_critlist," "boundary_npix," and "boundary_map" outputs were created and the "byteswap_out" flag was set to 1 (TRUE). The sequential version was run on a single 1.2 GHz CPU with 1.5 GBytes of RAM. The parallel version was run with 256-2.4 GHz CPUs with 1.0 GBytes of RAM each. Of course, identical hierarchical segmentation results were produced with each version.

As mentioned above, this version of RHSEG includes an option to select between different convergence criterion. See the User's Guide for instructions on how to utilize this added RHSEG feature through the new "conv_criterion" parameter. Note that along with the addition of this feature, the old "convfact" parameter has been renamed to "conv_factor."

In addition to the above major changes to RHSEG, previously known bugs in RHSEG related to image masking and connected component labeling have been corrected, as well as a newly discovered bug in the calculation of the global dissimilarity criterion values. Another minor change to the RHSEG code makes it more efficient in processing images with large constant areas. Some minor bugs were also corrected in HSEGViewer. In addition, improvements were made in the HSEGViewer histogram equalization code used to create the "RGB Image" and "Region Mean Image" views of the data (for 16-bit input data).

**Special note for users of earlier versions of Java: The current default release of HSEGViewer can only be run with version 1.5 (or later) of Java. However, included in this release of HSEGViewer is a version compatible with Java version 1.4. To invoke this version run the "HSEGViewer1_4" MS-DOS Batch File instead of HSEGViewer.

Release Notes for RHSEG version 1.03, April 1, 2005

James C. Tilton, Mail Code 606.3, NASA GSFC
James.C.Tilton@nasa.gov

This version of RHSEG includes code to prevent the overwriting of any input file with any output file, along with a minor correction of the on-line help for RHSEG. It also includes a new version of HSEGViewer with bug fixes for the rthreshlist input file and for the reference1, reference2 and reference3 input files. The new version of HSEGViewer also uses histogram equalization when the input data is unsigned short in order to create viewable images for the "RGBImage" and "Region Mean Image."

Release Notes for RHSEG version 1.02, March 1, 2005

James C. Tilton, Mail Code 606.3, NASA GSFC
James.C.Tilton@nasa.gov

This version of RHSEG provides a reorganization and minor correction of the on-line help for RHSEG. It also includes a more complete user's guide.

Release Notes for RHSEG version 1.01, February 18, 2005

James C. Tilton, Mail Code 606.3, NASA GSFC
James.C.Tilton@nasa.gov

This version of RHSEG includes a correction of an implementation oversight in the processing window artifact elimination process for spclust < 1.0. This correction should optimize the segmentation results by a minor amount versus the previous version.

Release Notes for RHSEG version 1.0, February 10, 2005

James C. Tilton, Mail Code 606.3, NASA GSFC
James.C.Tilton@nasa.gov

This version of RHSEG incorporates a significant fundamental improvement in the underlying implementation of the RHSEG algorithm. It also includes important bug fixes that correct earlier implementation errors. These errors resulted in slightly suboptimal results for cases with spclust_wght < 1.0. Further, the definition of the parameters "dtype," "dissim_crit" and "gdissim_crit" were changed, the rnb_levels parameter was made optional, and some cosmetic changes where made in the screen and log file outputs.

Included with this release are the executables and source code for a couple utility programs for extracting particular information from the RHSEG output files (see below).

As noted in the on-line help (available with the "rhseg -help" command), the definition of the "dtype" parameter is now:

-dtype (short unsigned int) Data type of input image data
dtype = 8 => "unsigned char (byte)"
dtype = 16 => "short unsigned int"
(otherwise undefined)

This makes the definition more rational: the designator now corresponds to the number of bits per pixel.

NOTE: Because of this change, versions of HSEGViewer with release dates prior to February 8, 2005 will no longer work with RHSEG. The newest release of HSEGViewer is contained in the "jar" file named hsv20050208.jar.

The definition of "dissim_crit" and "gdissim_crit" were changed to include a new "SAR Speckle Noise" criterion. This criterion was taken from a recent paper by J.-M. Beaulieu (Utilisation of contour criteria in micro-segmentation of SAR images, Int. J. Remote Sensing, 10, September, 2004, Vol. 25, No. 17, 3497-3512).

-dissim_crit (short unsigned int) Dissimilarity criterion
1. "1-Norm,"
2. "2-Norm,"
3. "Infinity Norm,"
4. "(undefined),"
5. "(undefined),"
6. "Square Root of Band Sum Mean Squared Error,"
7. "Square Root of Band Maximum Mean Squared Error."
8. "(undefined),"
9. "SAR Speckle Noise Criterion."

		(default: 6. "Square Root of Band Sum Mean Squared Error")
-gdissim_crit	(short unsigned int)	Global dissimilarity criterion 1. "1-Norm," 2. "2-Norm," 3. "Infinity Norm," 4. "(undefined)," 5. "(undefined)," 6. "Square Root of Band Sum Mean Squared Error," 7. "Square Root of Band Maximum Mean Squared Error." 8. "(undefined)," 9. "SAR Speckle Noise Criterion." (default = {none})

The "rnb_levels" parameter has been made optional. When not specified, the value for rnb_levels is calculated as the smallest value that would make the number of pixels <= 2048 in the processing window at the deepest level of recursion. Experience has shown this to be a good choice for the value of rnb_levels vis-a-vis processing time.

The cosmetic changes in screen and log file output are as follows:

"rhseg -v" now gives the version number and date.

With the "debug" parameter set to 0, no log file is generated, and the parameter settings and information about the hierarchical levels output are printed to the screen.

With the "debug" parameter set to 1, a log file is generated, and the parameter settings and information about the hierarchical levels output are printed to this log file. Only the program start and end time are output to the screen.

With the "debug" parameter set to 2, in addition to the same outputs generated for debug = 1, some program progress statements are output to the log file.

With the "debug" parameter set to 3, additional detailed information concerning region merging is output to the log file.

Finally, significant fundamental improvement in the underlying implementation of the RHSEG algorithm is included in this release of RHSEG. This has resulted in improvements in processing throughput, which are most significant for spclust_wght = 0.0. This improvement was achieved by using a "data heap" to determine the optimal regions to merge at each iteration instead of a fully sorted list. For information on data heaps, see Chapter 11 of "Algorithms" by Robert Sedgewick (1983, Addison-Wesley Publishing Company, Inc.) or other algorithms textbooks.

Included with this release are the executables and source code for a couple utility programs for extracting particular information from the RHSEG output files. The source code for these programs can be found in the "Sample Source Code" directory under the RHSEG installation directory.

The RHSEG programs outputs files that do not have obvious interpretations. The "rlblmap" file contains just the region label map from hierarchical level 0 (the finest level of detail) and the "regmerges" file contains information on how to translate the "rlblmap" at hierarchical level 0 to the region label map at other hierarchical levels. The "rlblmap_extract" program demonstrates how the "regmerges" file is used to determine the region label map at other hierarchical levels.

The "rlblmap_extract" program is designed to use the "oparam" output file from RHSEG to obtain most of its input parameter information. The preferred usage of "rlblmap_extract" is:

rlblmap_extract parameter_file_name rlblmap_extract_file_name hlevel_extract

where parameter_file_name is the name of the input parameter file name (normally the output "oparam" file from RHSEG), rlblmap_extract_file_name is the name of the output file to contain the region label map at the selected hierarchical level, and hlevel_extract is the hierarchical level desired. More detailed help can be obtained from the "rlblmap_extract -help" command.

An additional sample utility program, regmean_extract, demonstrates how to generate a region mean image for a particular hierarchical level from the RHSEG outputs. Like the "rlblmap_extract" program, the "regmean_extract" program is designed to use the "oparam" output file from RHSEG to obtain most of its input parameter information. The preferred usage of "regmean_extract" is:

regmean_extract parameter_file_name regmean_extract_file_name hlevel_extract

where parameter_file_name is the name of the input parameter file name (normally the output "oparam" file from RHSEG), regmean_extract_file_name is the name of the output file to contain the region mean image at the selected hierarchical level, and hlevel_extract is the hierarchical level desired. More detailed help can be obtained from the "regmean_extract -help" command.

—James C. Tilton