JavaSeis

SegyZip

JPEG-like Seismic Data Compression
For SEG-Y files

Dave Diller, Weinman GeoScience

SegyZip is a utility program that compresses SEG-Y files using the JavaSeis SeisPEG 2D compression algorithm. The SeisPEG compression algorithm is a lossy technique, similar to the widely know JPEG standard, but with adaptations for seismic data. For access to the SeisPEG code and documentation, see:

https://javaseis.svn.sourceforge.net/svnroot/javaseis/javaseis/trunk/src/org/javaseis/seiszip/

which includes an overview of SeisPEG and the underlying lapped orthogonal transform in the files SeisPEG_Overview.pdf and LOT_ZouAndPearlman.pdf, respectively. The reader is encouraged to refer to those documents because a discussion of SeisPEG and data compression is not included here.

SegyZip is invoked from the command line, as follows:

> java org.javaseis.seiszip.SegyZip segyInputFile zippedOutputFile distortion ftGainExponent verticalBlockSize horizontalBlockSize frameKeyOffset frameKeyLength

Note that the order of the arguments is significant, and all arguments must be specified. The meaning of the arguments is as follows:

segyInputFile is the input SEG-Y file, which is assumed to contain one or more frames. *Only IEEE format traces are supported.*

zippedOutputFile is the compressed output file. By convention the names of SeisZip output files end in .syz, so that for an input file named SomeData.sgy the output file would be SomeData.sgy.syz.

distortion is the allowed distortion. The value 0.01 is a good default.

ftGainExponent is a function-of-time gain exponent. This parameter is useful for gaining raw un-gained shot records before compression so that the low amplitudes in the deep data are well preserved. The gain is removed after de-compression. A value of 1.5-2.0 is a good default for un-gained data. Use 0.0 for no gain.

verticalBlockSize the vertical block size. Normally this is the number of samples per trace, *increased to the next multiple of 16*.

horizontalBlockSize the horizontal block size. Normally this is the maximum number of traces per frame, *increased to the next multiple of 16*.

frameKeyOffset is the byte offset in the SEGY header of a header entry that is the same for each frame and different between frames, which is used to recognize the boundary between frames. *Note that this is the offset (which starts at 0), not the byte number (which starts at 1).* For example this value would be 16 for source and 20 for CDP.

frameKeyLength the length (in bytes) in the SEGY header of a header entry that is the same for each frame and different between frames. This value is 2 for a short and 4 for an int. Floating point values are not supported.

The property **org.javaseis.seiszip.nthreads** can be used to set the number of threads that SeisPEG uses during compression. On the command line an example of the syntax is:

java **-Dorg.javaseis.seiszip.nthreads=8** org.javaseis.seiszip.SegyZip …

Further Improvements

The current implementation of SegyZip would benefit from support of trace formats other than IEEE.

Known Problems

See SeisPEG_Overview.pdf.