

JavaSeis
SegyUnzip
JPEG-like Seismic Data Compression
For SEG-Y files

Dave Diller, Weinman GeoScience

SegyUnzip is a utility program that uncompresses SEG-Y files that were compressed using SegyZip 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.

SegyUnzip has two modes of use, “historic” and “range-limited”. The *historic* usage is invoked from the command line, as follows:

```
> java org.javaseis.seiszip.SegyUnzip zippedInputFile segyOutputFile
```

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

zippedInputFile is a file that was compressed by SegyZip. By convention the names of SeisZip files end in .syz, so that for a SEG-Y named SomeData.sgy the compressed file would be SomeData.sgy.syz.

segyOutputFile is the output SEG-Y file.

The *range-limited* usage of SegyUnzip is invoked from the command line, as follows:

```
> java org.javaseis.seiszip.SegyUnzip inputDir=some_directory  
outputDir=some_directory volumeName=SOURCE_ segyExtension=.sgy  
seispegExtension=.syz minVolume=first_available maxVolume=last_available  
volumeInc=1 minFrame=1 maxFrame=last_available frameInc=1 minTrace=1  
maxTrace=last_available traceInc=1 maxSample=last_available ignoreErrors=false
```

The order of the arguments is not significant, and all arguments except `inputDir` and `outputDir` are optional. The arguments listed above show the default values. The meaning of the arguments is as follows:

inputDir=`some_directory` where *some_directory* is the path to a directory that contains one or more files that were compressed with SegyZip.

outputDir=`some_directory` where *some_directory* is the path to a directory to place the output files. The input and output directories are specified separately so that compressed data can be decompressed from an external portable disk directly to an internal disk in one step.

volumeName=`SOURCE_` where *SOURCE_* is the default naming convention for shots. An example of a typical compressed file name is “SOURCE_004514.sgy.syz”, for a shot number of 4514. The naming convention is used to recognize compressed files in the input directory.

segyExtension=`.sgy` where *.sgy* is the default naming convention for SEG-Y files. An example of a typical compressed file name is “SOURCE_004514.sgy.syz”, for a shot number of 4514. The naming convention is used to recognize compressed files in the input directory.

seispegExtension=`.syz` where *.syz* is the default naming convention for SeisPEG files. An example of a typical compressed file name is “SOURCE_004514.sgy.syz”, for a shot number of 4514. The naming convention is used to recognize compressed files in the input directory.

minVolume=`first_available` where *first_available* is the default minimum volume to uncompress. The volume number must be part of the compressed file name. An example of a typical compressed file name is “SOURCE_004514.sgy.syz”, for a volume (shot) number of 4514.

maxVolume=`last_available` where *last_available* is the default maximum volume to uncompress. The volume number must be part of the compressed file name. An example of a typical compressed file name is “SOURCE_004514.sgy.syz”, for a volume (shot) number of 4514.

volumeInc=1 where 1 is the default increment between volumes to uncompress. For example, if in the case where `minVolume`=104 `maxVolume`=110 `volumeInc`=2 then SegyUnzip would uncompress the files SOURCE_000104.sgy.syz, SOURCE_000106.sgy.syz, SOURCE_000108.sgy.syz and SOURCE_000110.sgy.syz if these files existed in the input directory.

minFrame=1 where 1 is the default minimum sequential frame number to uncompress within each volume/file. Frame numbers begin at 1 and ascend monotonically in each file – they are not related to any trace header values.

maxFrame=last_available where last_available is the default maximum sequential frame number to uncompress within each volume/file. Frame numbers begin at 1 and ascend monotonically in each file – they are not related to any trace header values.

frameInc=1 where 1 is the default increment between frames to uncompress. For example, in the case where minFrame=1 maxFrame=101 frameInc=10, then SegyUnzip would uncompress frames 1, 11, 21 ... 101 within each file.

minTrace=1 where 1 is the default minimum sequential trace number to uncompress within each frame. Trace numbers begin at 1 and ascend monotonically in each frame – they are not related to any trace header values.

maxTrace=last_available where last_available is the default maximum sequential trace number to uncompress within each frame. Trace numbers begin at 1 and ascend monotonically in each frame – they are not related to any trace header values.

traceInc=1 where 1 is the default increment between traces to uncompress. For example, in the case where minTrace=1 maxTrace=101 traceInc=10, then SegyUnzip would uncompress traces 1, 11, 21 ... 101 within each frame.

maxSample=last_available where last_available is the default maximum sample number to uncompress within each trace. Sample numbers begin at 1. We limit the length of traces with the sample *number* rather than a length of *time* because the sample interval (and hence the apparent total recording time) is not always handled correctly in SEG-Y files.

ignoreErrors=false where false is the default value for the policy regarding ignoring errors. If ignoreErrors=true then SegyUnzip will print a warning but continue if it encounters an error uncompressing any individual file. This policy is useful when uncompressing large numbers of files from inexpensive portable disks where the likelihood of a corrupted file is high.

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

```
java -Dorg.javaseis.seiszip.nthreads=8 org.javaseis.seiszip.SegyUnzip ...
```

Known Problems

See SeisPEG_Overview.pdf.