

## Storage and Dissemination of SEGY Data in JPEG2000 Format.

Robert Courtney  
Geological Survey of Canada Atlantic  
Bedford Institute of Oceanography  
1 Challenger Drive  
Dartmouth, Nova Scotia B2Y 4A2

The Geological Survey of Canada (GSC) collects over 1 TByte of SEGY (Society of Exploration Geophysicists)-formatted high-resolution seismic and sidescan data each year, and wishes to implement an efficient mechanism for storing, discovering and accessing these data holdings. Currently, scientists often use field-generated paper records in preference to digital data, as accessing and processing existing digital holdings involves excessive effort. High resolution chirp subbottom data is routinely collected during our multibeam survey expeditions and, since hardcopy is not routinely generated during these surveys, these data remain largely unutilized. Much of our digital holdings have never been examined and their content has not been verified. The GSC is exploring JPEG2000 technology to address these challenges.

The JPEG2000 framework, although commonly associated with images, is a wavelet-based compression standard that can accommodate multiplane, signed data arrays with up to 38 bits of resolution. It can generate a multiresolution representation with quality layers and random file access, allowing a quick and quality-progressive peek into file contents. It provides both reversible and lossy compression options. JPEG2000 is a flexible file format optimized for transfer over low-bandwidth internet connections and allows the inclusion of user-designed content and more standardized XML boxes for metadata.

A Windows-based application was developed to transform single channel SEGY data into JPEG2000 format. An XML schema (XSD) was developed to encode SEGY tape and trace header data into JPEG2000 xml boxes. These XML data will also be used as discovery metadata in GSC's relational databases to describe seismic data holdings. Trace data was encoded in a data array at a bit resolution sufficient to preserve data fidelity since JPEG2000 can encode data at fractional word lengths.

A series of trials on existing seismic data sets show that a loss-free conversion of SEGY to JPEG2000 will reduce file size by a factor more than 2:1 but that partial, lossy versions of the same data show no visible artifacts for compression ratios in excess of 20:1. In addition, very large files can be easily and quickly viewed over local networks and the internet using standard JPEG2000 viewers without decoding the entire data stream. Software to convert SEGY to and from JPEG2000 is freely available from the author.