Image Compression

INFORMATION SHEET
June 2010

- **What is compression?**

The term compression in the computer world means to make file sizes smaller by reorganizing the data within the file. Data that is duplicated or that has no value is eliminated or saved in a shorter format, greatly reducing a file's size.

Probably the most common compression format is a zip file. Files within the zip file return to their original state when unzipped and viewed.

- **What is imagery compression?**

Imagery compression is a bit different from zipping files. Imagery compression changes the content and organization of the data within a file, not necessarily to be restored to its original condition upon opening. An imagery compression routine reorganizes the data and sacrifices data where necessary to achieve the desired level of compression. Depending on the compression ratio, this sacrifice of data may or may not be apparent to the user.

An example of image compression: large areas of the sky are the same shade of blue - only the value for one pixel needs to be saved along with the locations of the other pixels with the same color.

- **Why do people compress files?**

Compression is done for a number of reasons:

1. Make files easier to send via email
2. Fit data on media
3. Save storage space
4. Archive data
5. Organize projects
6. Make files easier to open and use

For example, many TIFF images can be mosaicked together and compressed into one image to make the product more usable with image viewing software. Image file sizes without compression can be so large, that for the common home computer system they may become almost unusable.

Conversely, compression can make a file size so small, that image quality becomes extremely poor. If the user doesn't mind poor quality, this technique is useful to save storage space or for applications such as indexing image archives.

- **Can imagery become better with compression?**

If by "better", one means more user friendly, then yes. However, imagery cannot be made "higher quality" or "higher resolution" through compression.

- **So what's the downside of compression?**

Files that have been compressed can lose data, even though it may not be apparent to the user. Sometimes this is not a desirable side effect, especially if done incorrectly on high quality raster data.

- **What's a compression ratio?**

The easiest way to understand a compression ratio (shown as 1:50 or 1:15, etc.) is to think of it in terms of file size and quality. The higher the compression ratio, the smaller the file size (more data is packed into a smaller space), but the lower the quality of the compressed product.

- **Is prep work required to compress imagery?**

Before compressing imagery, image processing may be desirable to prepare images for compression, including: orthorectify or georeference the source images, mosaic and color balance, combine multi-resolution imagery, add vector data, etc.

- **What's Lossless and Lossy Compression?**

Lossless compression reduces an image so its quality matches the original source - mathematically nothing is lost. Although lossless sounds ideal, at times it doesn't provide enough compression and files could remain quite large.

Lossy means that the compressed image isn't quite the same as the one you started with. Lossy compression degrades images to some degree and the more they're compressed, the more degraded they become. In many

situations, such as posting images on the web or making small to medium sized prints, the image degradation isn't obvious. However, compress an image too much or enlarge an image enough and it will show.

- **Are there many compression formats?**

Yes. However, in the geospatial arena two main proprietary formats continue to dominate the market for image compression. MrSID (Multiresolution Seamless Image Database) from LizardTech and ECW (Enhanced Compressed Wavelet) from ERDAS ER Mapper. Both support the JPEG 2000 compression format.

Compression file extensions:
1. MrSID MG2 and MG3 (.sid)
2. JPEG and JPEG 2000 (.jpg, .jpeg, or .jp2)
3. ERDAS ER Mapper ECW (.ecw)

- **What formats does the Aerial Photography Field Office (APFO) use?**

The USDA Farm Service Agency (FSA) and Natural Resources Conservation Service (NRCS) have used image compression for many years in order to save disk space as well as to distribute and/or package data in a more convenient form.

The MrSID format is used by APFO, in large part because a number of other government agencies were already using this format. The earlier versions of the MrSID software not only could compress the imagery but could also mosaic multiple images into a single seamless, tone-matched, compressed county mosaic. This made for an easy-to-use and pleasing to the eye base layer for GIS or hardcopy uses.

APFO has used several compression formats for the NAIP program.

1. All Compressed County Mosaics (CCMs) before 2005 NAIP have the following specifications:
    a. MrSID MG2 format
    b. Compression ratio of 1:50
2. All CCMs from 2005 NAIP through 2007 NAIP have the following specifications:
    a. MrSID MG3 format
    b. Compression ratio of 1:15
3. For 2008 NAIP, CCMs with 4-band spectral resolution are:
    a. JPEG 2000 format
    b. Compression ratio of 1:15

    All other CCMs are MrSID MG3 and a compression ratio of 1:15
4. 2009 and 2010 NAIP CCMs specifications are:
    a. MrSID MG3 format
    b. Compression ratio of 1:15

    These will be the specifications for the foreseeable future.

Note: Some compression software can do more than compress data; they can reproject imagery into different coordinate systems or allow images to be added to already compressed files without having to start over from scratch.

- **What viewers view compressed imagery?**

Most image viewers and GIS software can view most compressed imagery formats.

APFO uses MG2, MG3, and JPEG 2000 for imagery. These two free viewers can view different formats:

1. **Tatuk GIS Viewer**, which can be downloaded via the APFO website:
http://www.fsa.usda.gov/FSA/apfoapp?area=home&subject=prod&topic=vdi

    Note: Tatuk cannot view a 4-band image.

2. **Global Mapper 9**, which also can be downloaded via the APFO website.

ArcGIS 9.x can view MG2, MG3, and JPEG 2000.

ArcGIS 8.3 can view MG2, MG3, and JPEG 2000 with the ArcGIS 8.3 Raster Update, which can be downloaded from the ESRI website here:
http://resources.arcgis.com/content/patches-and-service-packs?fa=viewPatch&PID=43&MetaID=579

ArcView 3.x can view the MG3 format with the appropriate plug in, which can be downloaded from the LizardTech website here:
http://www.lizardtech.com/download/dl_download.php?detail=geo_avmrsid_dll&platform=win

Installing the following plug in will allow **ArcView 3.x** to view JPEG2000 compressed images.
http://forums.esri.com/Attachments/12333.zip

- **Who do I contact for more information?**

For more information on compression, contact GIS Specialist, Zack Adkins, at (801) 844-2925, or GIS Specialist, David Davis, at (801) 844-2933.