

# GEOSPATIAL IMAGE FORMATS SHOOTOUT – PART ONE

CHRIS HANSON INVESTIGATES PROS AND CONS OF DIGITAL FORMATS IN PART ONE OF TWO PART SERIES

Since the introduction of the JPEG standard in 1992, the geospatial industry has been moving towards more and more sophisticated compression methods. Originally using industry or vendor-specific (typically uncompressed) formats, the focus moved through adopting and extending industry standards (GeoTIFF, JPEG, PNG), finally to devising or extending very advanced new formats (ECW, SID, GeoJP2) with features specifically for geospatial use.

Image file formats are typically characterized as “lossy” or “lossless” depending on whether their output is exactly the same as their input. Lossy formats discard some degree of exact reproduction in pursuit of more aggressive compression. While it would seem that lossy compression would be unacceptable in geoscience, it turns out that in many situations focused on visual interpretation, some loss is permissible in exchange for better storage convenience. However, image data for use in quantitative analysis should be preserved in a lossless format to avoid contaminating the results of the calculations. This article will focus on and compare both lossy and lossless formats.

## JPEG

JPEG is a “lossy” file format — some information is unavoidably lost in the compression process though the algorithm tries to discard the least important and noticeable details first. JPEG uses a Discrete Cosine Transform to construct a mathematical representation of each 8 pixel by 8 pixel block. This change of structure doesn’t inherently lose any data, it just prepares it for the next stage. To actually compress the image, a quantizing encoder “rounds off” a certain quantity of the least significant factors of that approximation.

## PNG

PNG was devised in 1996 as a replacement for the limited and patent-encumbered GIF file format. While not specifically a geospatial file format, it is one of the most recent general-purpose lossless compressed image file formats and compresses well. Geospatial positioning data can only be provided by an accompanying (PGW) Worldfile.

## GeoTIFF

GeoTIFF is an outgrowth of the TIFF format, originally devised in the mid-80’s as a common format for scanned images. Supporting a huge variety of options as well as complete embedded positioning information, GeoTIFF is a very popular geospatial file format. TIFF is usually lossless, also offering an uncompressed variant and one lossy method that effectively inserts a JPEG image into a TIFF wrapper.

## ECW (.ECW)

ER Mapper introduced their ECW (Enhanced Compressed Wavelet) file format in the late 90s, driven by the same needs as MrSID (below). As a geospatial company, ER Mapper included region-of-interest decoding, geospatial metadata and efficient server streaming right from the start. ECW is always “lossy” to some small degree, even at the lowest compression ratios.

## JPEG 2000 (.JP2)

JPEG 2000 began in the late 90s, was ratified as a standard in 2000 and is steadily becoming a major player for advanced geospatial image storage. JP2 offers a comprehensive list of features with nearly every

desirable option: lossy and lossless variants, efficient progressive resolution and region-of-interest decoding, as well as extensions for geospatial metadata. JP2 is defined by the same JPEG group as the original beloved JPEG standard and is an open format and an ISO standard. However, despite being from the same standards organization, JPEG and JP2 are not related at all, technologically.

The JP2 JPIP extension provides for efficient server streaming, but is not yet widely supported. LizardTech and ER Mapper both offer JPEG 2000 server streaming capability in their respective Express Server and Image Web Server products, though neither employs the JPIP protocol.

### MrSID

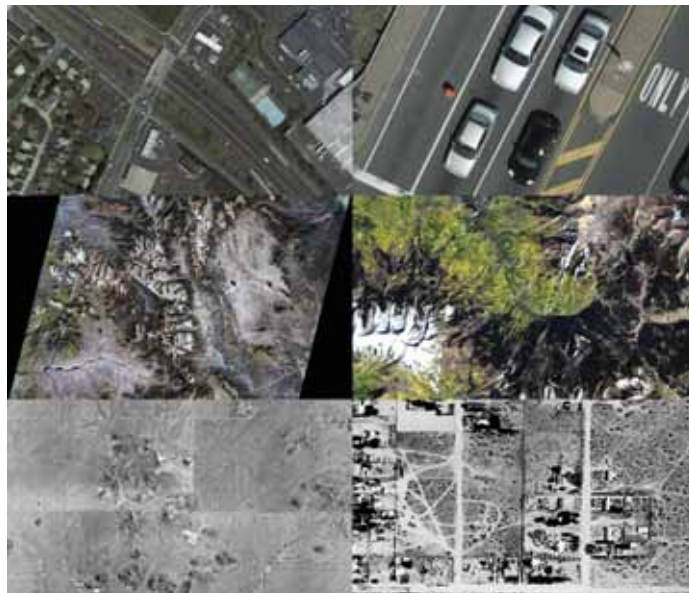
LizardTech's MrSID (Multi-resolution Seamless Image Database) was originally developed at the Los Alamos National Laboratory (LANL), and is now used extensively as a geospatial raster image format. MrSID images can be lossy or lossless and offer progressive and region-of-interest decoding, geospatial metadata and efficient server streaming.

### Data Overview

We've specifically chosen several data sets to exercise the abilities of these file formats. In order to ensure the validity of the compression tests, we must ensure the samples have never been lossy compressed before. Lossy compression leaves an indelible mark on an image that can greatly affect the quality and size of subsequent compressions. We have verified the provenance of each dataset and inspected the visual and statistical properties of the images to ensure they are clean, uncompressed imagery and are representative of typical geospatial imagery.

For size/quality tests, we have one monochrome/greyscale (panchromatic) image, one full-scene natural color satellite image and one natural color aerial orthophoto. The panchromatic image is a USGS DOQQ aerial orthophoto (doq02435117a7.TIF, "North Edwards", 153Mb from the Mojave Desert Ecosystem Project

<http://www.mojavedata.gov/datasets.php?&qclass=doq>). The satellite image is a 30m Landsat 3-2-1 image (p034r033\_7t20010924\_z13\_nnN0.tif, 206Mb) from Landcover.org processed with AlphaPixel's PixelSense. The



color orthophoto is (0008.TIF, UltraCam-X imagery, 388Mb) courtesy of Microsoft Corporation.

While not massive images, their overall size is not important. We will be testing overall quality respective to the final compressed size, which should be the same no matter how little or how much source data we process.

### Software

We'll be testing the just-released ER Mapper Image Compressor (7.2) and LizardTech's GeoExpress 6.1.1.1513). Both programs are capable of compressing either a single input image as well as balancing and mosaicking a set of photos into a seamless super-mosaic. Comparative tests with other image formats (JPEG, PNG, BMP, TIFF) may use Adobe Photoshop CS2. In






THIRD INTERNATIONAL CONFERENCE  
**EARTH FROM SPACE**  
 THE MOST EFFECTIVE SOLUTIONS

**December 4-6, 2007**  
 Moscow, Russia

— is the perfect opportunity to demonstrate the latest innovations, achievements and practical results in sphere of space information technologies

**Main conference topics:**

- Russian and foreign remote sensing programs.
- Remote sensing for territory management and decision-making support.
- Remote sensing for disaster mitigation and assessment of environmental consequences of natural and technological disasters.
- Technologies and software for data processing.
- Education for sustainable development: New information technologies.

**Important Deadlines:**

May 20, 2007 — End of Early-bird registration  
 September 20, 2007 — Deadline for abstract submission

**Organizing Committee:**

Build 1, 5/22 Rossolimo Street, 119021 Moscow, Russia  
 Phone: +7(495)246-38-53 • Fax: +7(495)246-25-93  
 E-mail: [conference@scanex.ru](mailto:conference@scanex.ru)  
[www.transparentworld.ru/conference](http://www.transparentworld.ru/conference)

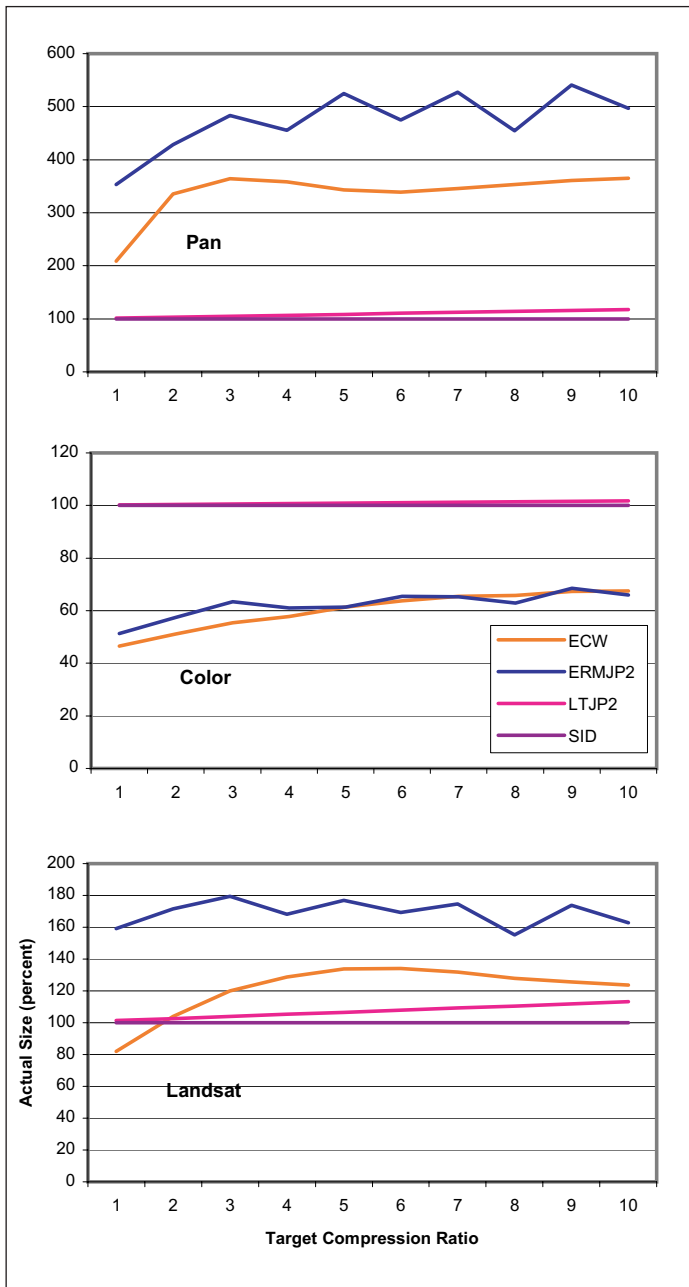
Supported by GIS-Association



Media Partners





part two we will delve into speed testing, and we'll outline our test hardware at that point.

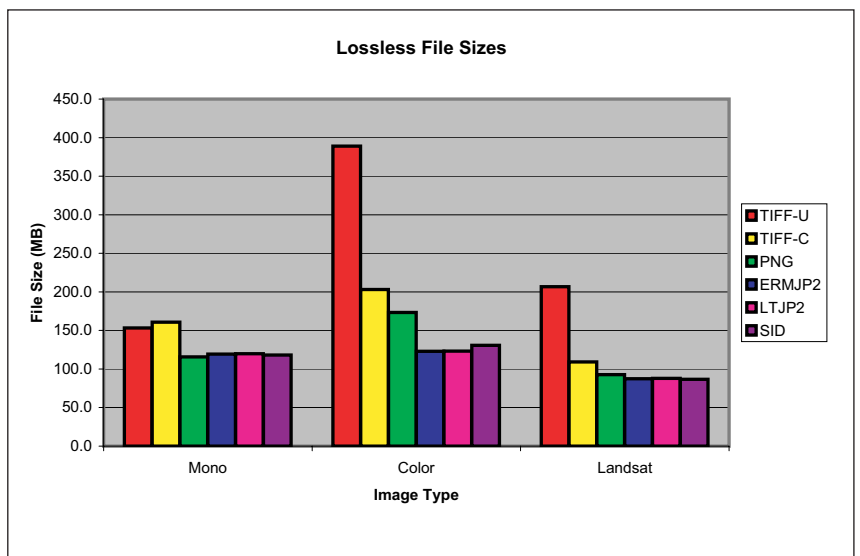
### Test Methodology

There are four main tests we will perform. Part one will focus on lossless size comparisons, and part two will focus on quality and speed.

To begin with we will compare final output sizes for all of the lossless formats (BMP, compressed and uncompressed TIFF, PNG, JP2 and Lossless MrSID). Speed and architecture will not be compared, and quality will be identical for all since by definition, all are lossless. To finish part one, we will compare target compressed size versus actual compressed size for all size-predictive lossy formats (ECW, JP2, SID). Part two will be a followup article in the next issue.

### Lossless: File Sizes

Choice of an image format often boils down to either the supported featureset, or the ability to store your image data compactly. Assuming you require lossless data storage, which file format is the most efficient at packing the data into the fewest bytes? (In this graph, smaller is better.)



Clearly, Uncompressed TIFF (TIFF-U) is the loser in this comparison. Compressed TIFF (TIFF-C) is quite a bit better than uncompressed in some cases (Color and Landsat) but is actually larger than uncompressed for the Pan image. PNG is always better than TIFF, and in fact is the best of all for the Pan image. Jumping to the end of the chart, MrSID puts up good numbers as well, being best for the Landsat image, second place for the Pan image and third place for the Color image.

Both JPEG 2000 candidates did very well finishing in the top three for every test except LizardTech's JP2 take on the Pan image. LizardTech's JP2 didn't take first place in anything, but was always very close. ER Mapper's JP2 took first place for the Color image, second for the Landsat and third for the Pan image.

Remember that quality doesn't play any part in this test – every contender is exactly the same quality as the others, since they are all by definition, lossless.

Who is the winner? I'd call it a dead even tie between ER Mapper's JP2 and LizardTech's MrSID, with LizardTech's JP2 clomping in a nose behind them. LizardTech's JP2 is slightly larger than one of the front-runners in all three tests.

### Lossy: Target Size versus Actual Size

Lossy formats like ECW, MrSID and JPEG 2000 all allow you to indicate to the compressor a desired output size, or compression ratio. Basic JPEG only offers a vague "quality" setting which indicates little about the actual resulting size or ratio, so it is not tested here. How closely this goal is met depends on the compression tool you are using. (In this graph, closer to 100% is better.)

Clearly, the LizardTech GeoExpress encoder is much more careful to hit the user-specified desired compression ratio or size than the ER Mapper Image Compressor encoder. The graphs for both MrSID and GeoExpress' JP2 are so close to the "optimal" size as to be indistinguishable. ER Mapper's ECW and JP2 encoders sometimes undershoot and sometimes overshoot the mark depending on the input data. How important this criteria is to you in your applications will determine whether this is significant or not. ER Mapper explains that they aim for a target Quality rather than an exact ratio or size.

### Part One Conclusion

So far all the wavelet contenders are both putting in a good effort. In part two we'll dive into the hard numbers of image quality and compression speed to see if any one pulls definitively ahead in the race.

**Chris "Xenon" Hanson** is the co-founder of 3D Nature and AlphaPixel, innovators in the 3D Landscape Visualization and Remote Sensed Image Processing fields respectively. He has been working with computers for 27 years, computer graphics for 20 years and GIS for 15 years. He can be contacted at [Xenon@AlphaPixel.com](mailto:Xenon@AlphaPixel.com).

Whatever  
Whenever  
Wherever...



## Your imagery. Delivered.

**Express Server** enables you to serve more users in less time using widely adopted standards and familiar tools—all without adding hardware or buying more bandwidth.

To learn how your company can benefit from instant online viewing of massive geospatial imagery with Express Server, visit:

[www.lizardtech.com/products/exp](http://www.lizardtech.com/products/exp)

Or contact us at 206.652.5211 or [geosales@lizardtech.com](mailto:geosales@lizardtech.com)

© 2007 LizardTech, a division of Celartem, Inc. All rights reserved. LizardTech, LizardTech logo and Express Server logo are trademarks or registered trademarks of LizardTech in the United States, and all are the property of LizardTech. All other trademarks are property of their respective owners.

Earth photography courtesy of NASA