

HOW TO COMPRESS IMAGE DATA

Raster data is comprised of grid cells. Unlike its vector counterpart that is comprised of vertices and paths. These are the two primary spatial data types in GIS. Accuracy depends on cell size. Cell size must be small enough to capture the required detail. And as resolution increases, the size of the cell decreases. But this comes at a cost for speed of processing and **data storage**.

Raster storage in its raw state is inefficient because it is normally stores values row-by-row from the top left corner.

SOLUTION:

The way to improve raster size efficiency is through **image compression**. And there are several ways to use GIS compression to reduce file size and still maintain quality of data.

Let's take a look at some examples how rasters images are compressed:

RUN LENGTH ENCODING – GROUPING ROWS OF DATA

Run length encoding stores cells on a **row-by-row** basis. Instead of recording each individual cell's values, run length encoding groups cell values by row.

Take this line of data:

AAAAAABBBBCCCCCCCC

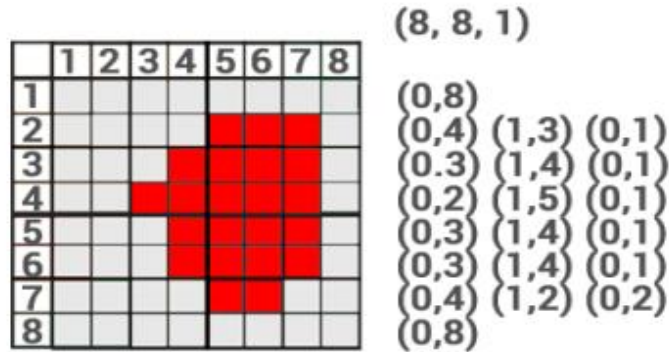
It can be rendered as:

6A4B9C

This image encoding method reduces data volumes because each line is recorded more efficiently. Even though the same information is being held, values that are the same are stored as a string.

In the example above, the first row is blank and is stored as (0,8). This means there are 8 cells and they are all zeros. In the second row, there are 4 consecutive zeros so it gets a value of (0,4).

After this, we have three consecutive cells with the value 1 so it gets a value of (1,3). This continues until it reaches the bottom-right cell.

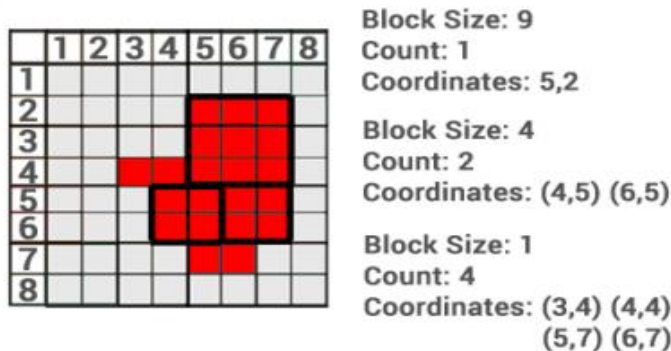


Block Coding – Grouping Blocks of Data

The block coding raster storage technique assigns areas that are blocks to reduce redundancy.

The block coding raster image compression method subdivides an entire raster image into **hierarchical blocks**. It's an extension of the run length encoding technique, but extends it to two dimensions. In the example above:

Instead of storing 64 grid cells, all it takes is just 7 blocks. Using block coding, it requires one 3x3 block, two 2x2 blocks and four 1x1 cell blocks to encode this raster image. In this block coding example, the top-left corner is used as a reference for each



block.

CHAIN CODING – DEFINING THE EXTERIOR BOUNDARY

Chain coding defines the outer boundary using **relative positions** from a start point. The sequence of the exterior is stored where the endpoint finishes at the start point.

During the encoding, the direction is stored as an integer. However, in this example we use cardinal directions for simplicity. For example, the value 0 is north and 1 is east.

In the example, we start at position (5,2). From here we define the border using cardinal directions and number of movements. We move east 3 positions until we hit the edge. At this location, we move south 4 positions. This process continues until the end point hits the start point.

Note: Only for the purpose of this exercise, we used north, east, south and west as alphabetical values. When encoded, it is a numerical value.

	1	2	3	4	5	6	7	8
1								
2								
3								
4								
5								
6								
7								
8								

(5,2)

(E3, S4, W1,
S1, W1, N1,
W1, N3, E1,
N1)

QUADTREE ENCODING – SUBDIVIDING DATA INTO QUARTERS

Quadtrees are raster data structures based on the successive reduction of homogeneous cells. It **recursively subdivides** a raster image into quarters. The subdivision process continues until each cell is classed.

It reduces raster storage requirements. It also is dependent on the complexity of the feature and the resolution of the smallest grid cell.

In the example, the top-left and bottom-right 8×8 grids do not need to be subdivided further because they are homogeneous. The top-right 8×8 grid is subdivided into three 4×4 grid. The remaining 4×4 grid is separated into 4 individual classes.

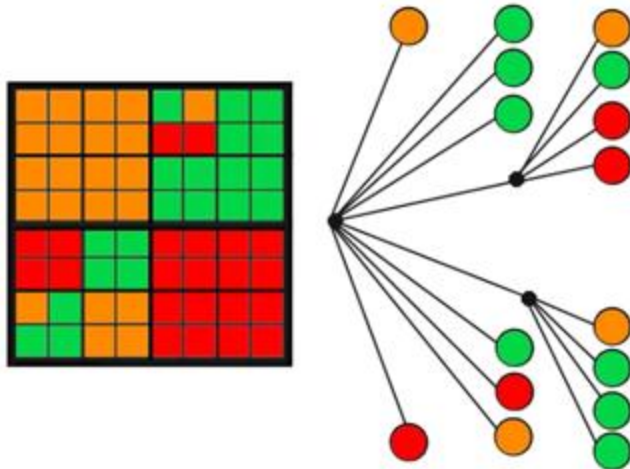


IMAGE COMPRESSION REDUCES FILE SIZE

GIS data is abundant. With satellites acquiring images each day, raster data is the spatial model of choice. But which format will you use?

Deploying efficient raster image compression techniques means reducing storage space. This is the primary benefit of compressing your data.

It can **save money and time**. You can also improve your **network performance** because you are working with a reduced amount of data.