The Raster Data Model

Llano River, Mason Co., TX

Rasters are:
- Tessellations of square cells
- Arrays of values distributed among equal-sized, square cells

<table>
<thead>
<tr>
<th>565</th>
<th>573</th>
<th>582</th>
<th>590</th>
</tr>
</thead>
<tbody>
<tr>
<td>575</td>
<td>580</td>
<td>595</td>
<td>600</td>
</tr>
<tr>
<td>579</td>
<td>581</td>
<td>597</td>
<td>601</td>
</tr>
<tr>
<td>580</td>
<td>600</td>
<td>620</td>
<td>632</td>
</tr>
</tbody>
</table>

Why squares?
- Computer scanners and output devices use square pixels
- Bit-mapping technology/theory can be adapted from computer sciences
- 1-to-1 integer mapping to grid coordinate systems!

Cell Location Specified by:
- Row/column (R/C) address
- Origin is upper left cell (1,1)
- Relative or geographic coordinates can be specified
World File – DRG Example

- 2.43840000000000
- 0.00000000000000
- 0.00000000000000
- 2.43840000000000
- 487888.64154709835000
- 3403923.72301301550000
- /* UTM Zone 14 N with NAD83
- /* This world file shifts the upper left image coordinate to the corresponding
- /* NAD83 location, resulting in an approximated NAD83 image.
- /* Map Name: Art
- /* Map Date: 1982
- /* Map Scale: 24000

Spatial Resolution

- Defined by area or dimension of each cell
- Spatial Resolution = (cell height) X (cell width) e.g. "5m²"
- High resolution: cell represent small area
- Low resolution: cell represent larger area
- Abbreviated by size of one edge of cell (e.g. "5m DEM")
- For fixed area, file size increases with resolution
30 m vs. ~90 m Pixel Size

- Resolution of 30 m data is 9 times better (9X as many pixels) than 90 m data

Resolution Constraint

- Cell size should be less than half of the size of the smallest object to be represented (e.g. the circle below)
- Size of smallest recognizable object at any resolution is the “Minimum mapping unit; MMU”

e.g. DOQQ Resolutions

- Resolution is size of sampled area on ground, not MMU

Raster Dimension:

- Number of rows x columns
- E.g. Monitor with 1900 x 1200 pixels

<table>
<thead>
<tr>
<th>565</th>
<th>573</th>
<th>582</th>
<th>590</th>
</tr>
</thead>
<tbody>
<tr>
<td>575</td>
<td>580</td>
<td>595</td>
<td>600</td>
</tr>
<tr>
<td>579</td>
<td>581</td>
<td>597</td>
<td>601</td>
</tr>
<tr>
<td>580</td>
<td>600</td>
<td>620</td>
<td>632</td>
</tr>
</tbody>
</table>
Raster Attributes – What’s Stored in the Cells

- Two types:
  1. **Integer codes** assigned to raster cells
     - E.g. rock type, land use, vegetation
     - Codes are technically *nominal* or *ordinal* data
  2. Measured “real” values
     - Can be integer or “floating-point” (decimal) values; technically *interval* or *ratio* data
     - E.g. topography, em spectrum, temperature, rainfall, concentration of a chemical element

Integer Code Attributes

- Code is referenced to attribute via a “look-up table” or “value attribute table” – VAT
- Commonly many cells with the same code
- Different attributes must be stored in different raster layers

### Coded Value Raster Types

- **Single-band:** Single raster; *Thematic data*
  - **Black & White:** binary (1 bit) (0 = black, 1 = white)
  - **Panchromatic:** (“Grayscale”) (8 bit): 0 (black) – 255 (white) or graduated color ramps (e.g. blue to red, light to dark red)
  - **Colormaps:** (“Indexed Color”) (8 bit): code cells by values that match prescribed R-G-B combinations in a lookup table

### Single Band Examples – Black & White vs. Grayscale

**Black & White - 1 bit**

**Grayscale – 8 bit:** black, white & 254 shades of gray
Single Band
Example Color Map (Indexed Color)

- Each pixel contains one of 12 unique values, each corresponding to a prescribed color (Red, Green & Blue combination). Could have 254 prescribed colors this way.

Measured, “Real Value” Attributes

- Commonly stored as floating point values
- Different attributes must be stored in different layers, e.g. spectral bands in satellite imagery
- Compression techniques for rasters of integer-valued cells, but not floating point (see below)

Multiband Image Raster Attributes

- Multi-band Spectral Data
  - Band = segment of Em spectrum
  - Map intensities of each band as red, green or blue
  - Display alone or as composite

- Band 1
- Band 2
- Band 3

Attribute values
0 - 255

Figures from: Modeling our World, ESRI press

Multiband Image
8 bits/Band, 3 Band RGB

- E.g. Austin East 7.5’ Color Infrared Digital Orthophotograph (“CIR DOQ")
Cell Values Apply To:

- Middle of cell, e.g. Digital Elevation Models (DEM)
- Whole cell, e.g. most other data

Source: Modeling our World, ESRI Press

Digital Elevation Model

Airborn Magnetic (TFI) Map

TFI Pixel Values

How Are Rasters Projected?

- Problem: Square cells must remain square after projection.
- Solution: Resampling (interpolation); add, remove, reassign cells to conform to new spatial reference.
Raster File Size

- fixed by dimension, not information

At 1 bit/cell, file size = 8 x 8 x 1 = 64 bits (8 bytes)

File Structure

Header: (dimension, max. cell value) & resolution, coordinate of one corner pixel, etc.

Data File (linear array)

File Compression

- E.g. Run-length encoding

<table>
<thead>
<tr>
<th>Row</th>
<th>Run</th>
<th>Freq., Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,1</td>
<td>8,5</td>
<td>8,5</td>
</tr>
<tr>
<td>2,3</td>
<td>4,5</td>
<td>2,2, 2,5</td>
</tr>
<tr>
<td>3,3</td>
<td>4,5</td>
<td>2,2, 2,8</td>
</tr>
<tr>
<td>4,3</td>
<td>4,5</td>
<td>2,2, 2,8</td>
</tr>
<tr>
<td>5,2</td>
<td>6,2</td>
<td>2,8</td>
</tr>
<tr>
<td>6,2</td>
<td>6,2</td>
<td>2,2, 6,5</td>
</tr>
<tr>
<td>7,2</td>
<td>4,2</td>
<td>4,5</td>
</tr>
<tr>
<td>8,3</td>
<td>1,5</td>
<td>3,2, 4,5</td>
</tr>
</tbody>
</table>

Before: 64 characters
After: 46 characters (28% reduction; ratio of 1.4:1)
File Compression

- E.g. Block encoding

<table>
<thead>
<tr>
<th>Block Size</th>
<th>Value</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>5,1 6,1 3,6 4,6 8,6 8,7</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>7,5 6,5</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3,5 4,5 1,7 2,7 2,8 1,8 8,8</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>7,1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>5,2 5,4 1,5 3,7</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>7,3</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>5,6</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
<td>1,1</td>
</tr>
</tbody>
</table>

Before: 64 characters

After: 61 characters (5% reduction ratio of 1.05:1)

MrSID or ECW (Wavelet) Compression

- Multi-resolution Seamless Image Database – commercialized by LizardTech
- Compression ratios of 15-20:1 for single band 8-bit images
- Ratios of 2-100:1 (!) for multiband color images
- Also ECW by ER Mapper Ltd. (now Intergraph/ERDAS)

*** Enormous raster data sets now manageable on PCs and across web with this technology ***

“Lossy” vs. Lossless Compression

- Techniques that combine similar attribute information to reduce file size are “lossy” e.g. JPEG, GIF, PNG, MrSID
- Lossless formats; TIFF, BMP, GRID

Raster Pyramids

- Store reduced-resolution copies of a raster for rapid display – e.g. ArcGIS, Google, many others
- Often combined with image tiling and compression for rapid rendering of images
Image “Tiling”
- Split raster into small contiguous rectangles or squares = Tiles
- Display only the tile required upon zooming

Level 0 = 100% of image = 16 low res. tiles
Level 1 = higher res. (parts of 4 med. res. tiles)
Level 2 = highest res. (1+ high res. tiles), res. of image

Supported Raster Formats
- See ArcCatalog>Tools>Options
- Each explained in Help
- 24 supported formats

Vector or Raster?
- Spatially continuous data = raster
- Modeling of data with high degree of variability = raster
- Objects with well defined boundaries = vector
- Geographic precision & accuracy = vector
- Topological dependencies = vector or raster

Raster or Vector?
- Raster
  - Simple data structure
  - Ease of analytical operation
  - Format for scanned or sensed data – easy, cheap data entry
  - But……
    - Less compact
    - Query-based analysis difficult
    - Coarser graphics
    - More difficult to transform & project
- Vector
  - Compact data structure
  - Efficient topology
  - Sharper graphics
  - Object-orientation better for some modeling
  - But……
    - More complex data structure
    - Overlay operations computationally intensive
    - Not good for data with high degree of spatial variability
    - Slow data entry
Raster Spatial Analysis, e.g. “Map Algebra”

Fires (1 or 4)
Slope (1-4)
+ Geology (1-4)
Erosion Ranking (3-12)