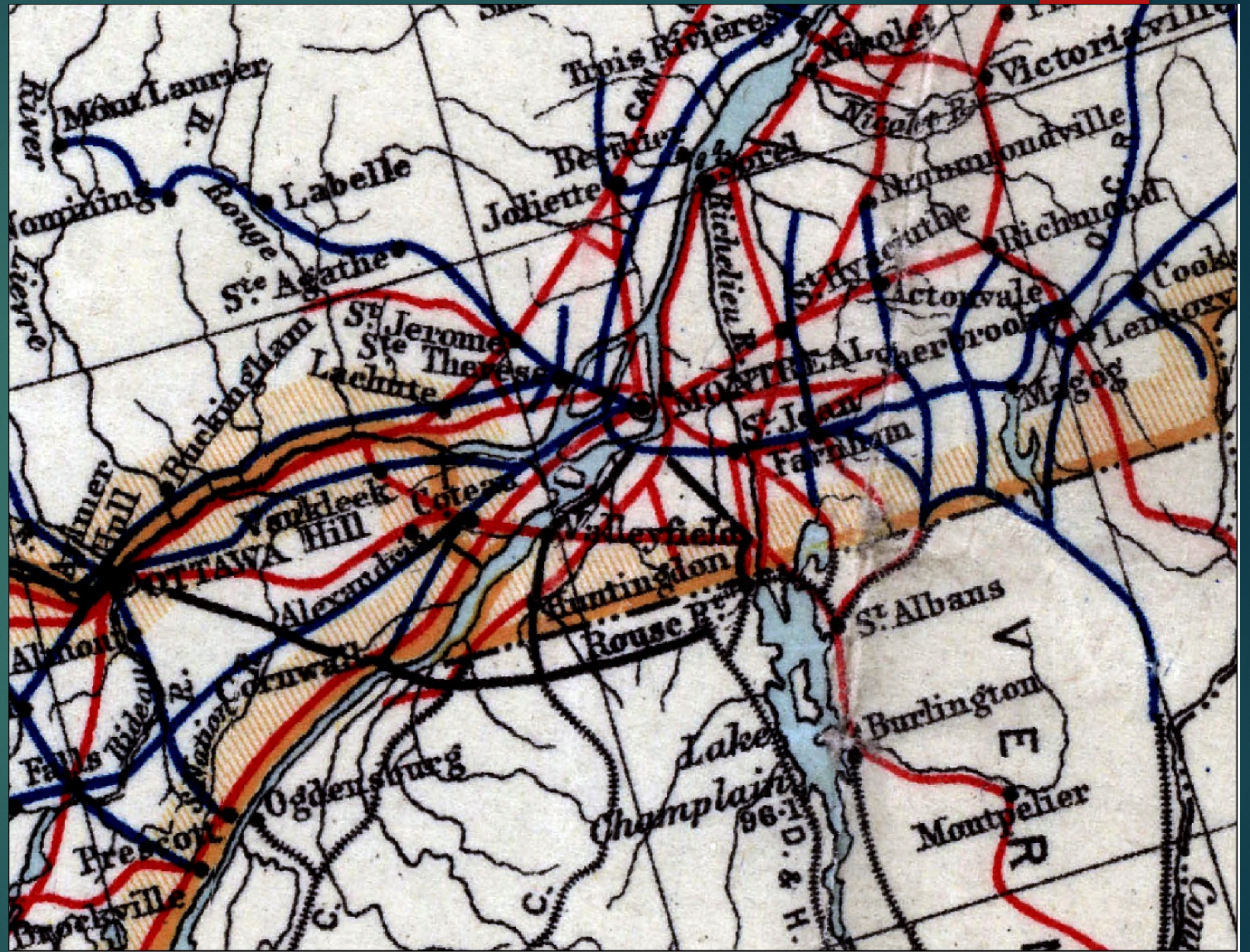


# From Paper Map to Geospatial Vector Layer

Peter Peller

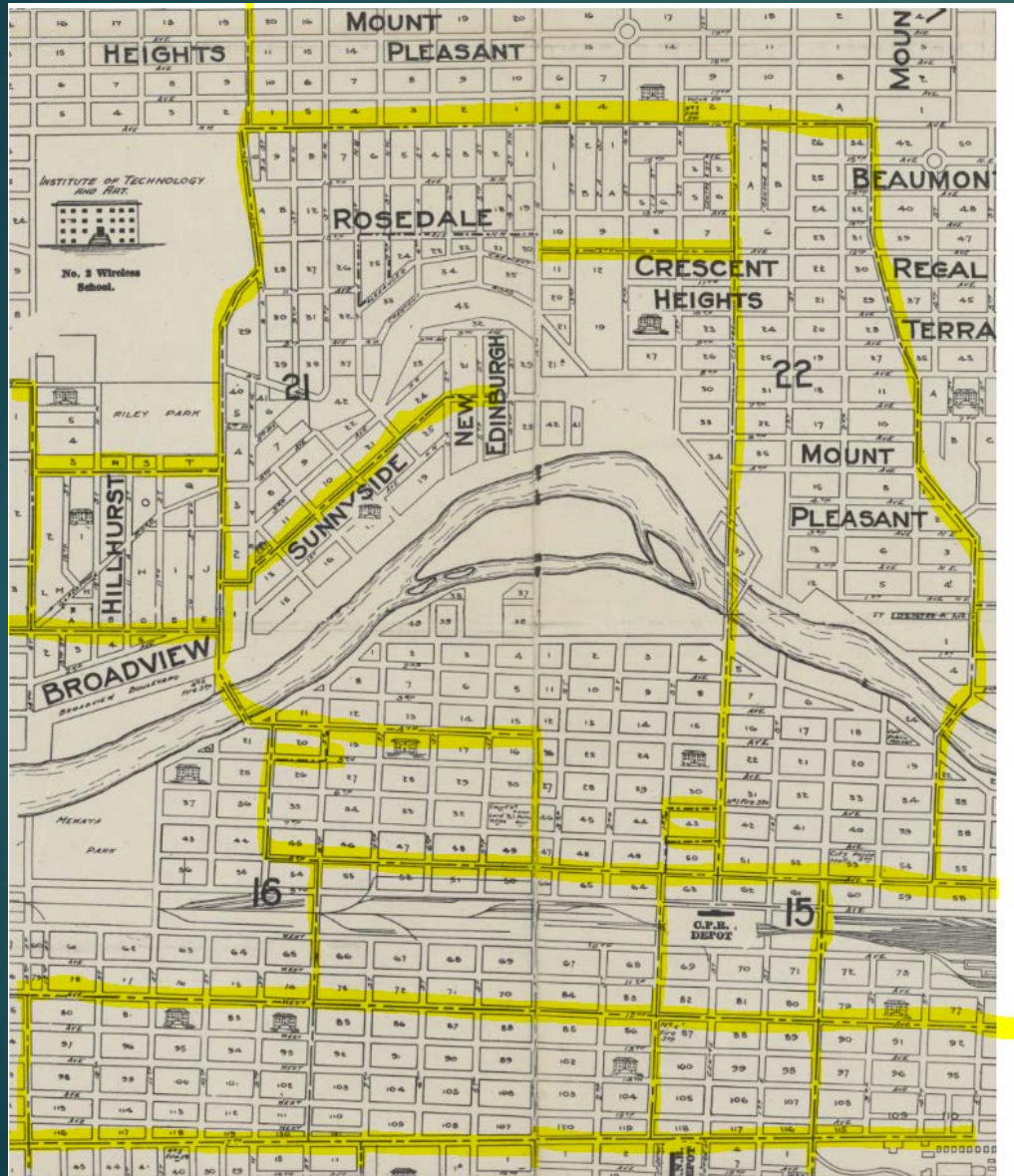
University of Calgary

May 30, 2018

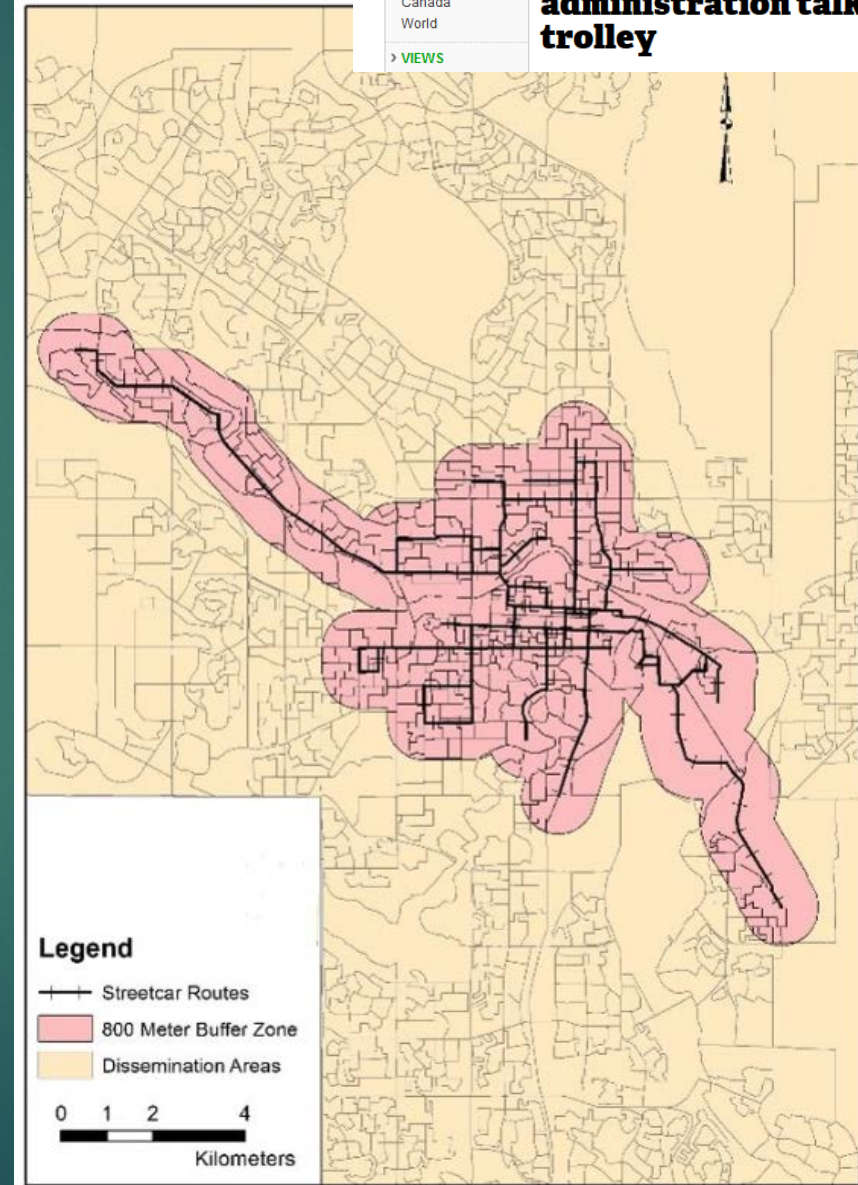




# WHY?



## Street car revival: Calgary administration talks bringing back trolley





# Definitions

- ▶ Raster – data structure is a matrix of cells (or pixels) where each cell contains a value
- ▶ Vector – data is represented as points, lines and polygons with each feature containing information
- ▶ Scan – create a raster image (ex/ scanned map)
- ▶ Vectorize – convert raster image pixels to a vector feature class (or layer)

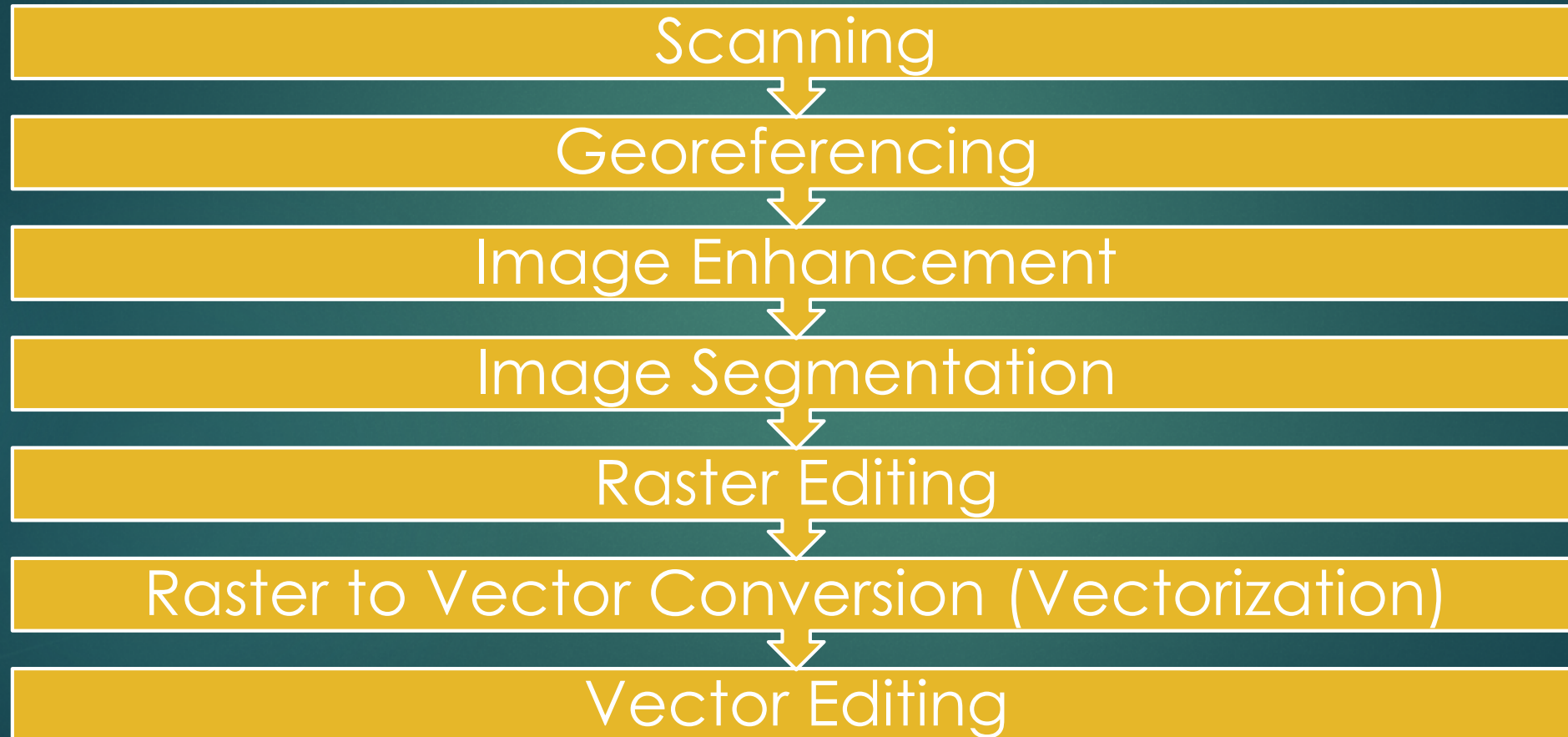
# Study Methodology

- ▶ Multiple databases searched
- ▶ 70 studies identified and examined
- ▶ 11 specific cases selected based on the following criteria:
  - ❖ Software available openly or commercially
  - ❖ Semi-automated process
  - ❖ Scalable to multiple maps and map types
  - ❖ Appropriate for a library and archive setting

★	●	📄	Authors	Title	Year	Published In	Added
★	●	📄	Chiang, Yao-Yi	Harvesting Geographic Features From Heterogeneous Raster Maps	2010	the 1st ACM SI...	4/18/17
★	●	📄	Sharma, O.	A Methodology for Raster to Vector Conversion of Colour Scanned Maps	2006		4/18/17
★	●	📄	Ansault, M.; Soille, P.	Mathematical Morphology: A Tool for Automated GIS Data Acquisition from Scanned Thematic Maps	1990	Photogrammetric Engineering an...	4/18/17
★	●	📄	Leyk, Stefan; Boesch, Ruedi	Colors of the past: color image segmentation in historical topographic maps based on homogeneity	2009	GeoInformatica	4/18/17
★	●	📄	Dunshee, Ian	Towards Standards for the Manual Digitization and Vectorization of Analog Historical Maps and Records	2016		4/12/17
★	●	📄	Chiang, Yao-Yi; Leyk, Stefan; Honarvar Nazari, Narges; M...	Assessing the impact of graphical quality on automatic text recognition in digital maps	2016	Computers & Geosciences	4/12/17
★	●	📄	Dhar, Deeptendu Bikash; Chanda, Bhabatosh	Extraction and recognition of geographical features from paper maps	2006	International Journal of Docu...	3/29/17
★	●	📄	Gibson, L; Lucas, D	Vectorization of raster images using hierarchical methods	1982	Computer Graphics and I...	3/24/17
★	●	📄	Lacroix, V	Raster-to-vector conversion: problems and tools towards a solution a map segmentation application	2009	Proceedings of the 7th Interna...	3/24/17
★	●	📄	Levachkine, Serguei	Raster to Vector Conversion of Color Cartographic Maps	2004	Lecture Notes in Computer Scien...	3/24/17
★	●	📄	Dedic, Abdulah; Murn, Rudolf; Pecek, Dusan	Digitalization of large area drawings and maps	1991	Melecon	3/24/17
★	●	📄	Naouai, M; Narjess, M; Hamouda, A	Line recognition algorithm using constrained Delaunay Triangulation	2010	Proceedings Elmar - Internat...	3/24/17
★	●	📄	Miyoshi, T; Li, W; Kaneda, K; Yamashita, H; Nakamae, E	Automatic extraction of buildings utilizing geometric features of a scanned topographic map	2004	Proceedings - International C...	3/24/17
★	●	📄	Velázquez, A; Levachkine, Serguei	Text/Graphics Separation and Recognition in Raster-Scanned Color Cartographic Maps	2004	Lecture Notes in Computer Scien...	3/24/17
★	●	📄	Knoblock, Craig; Chen, C.-C.; Chiang, Y.-Y.; Goel, A; Mich...	A general approach to discovering, registering, and extracting features from raster maps	2010	Proceedings of SPIE - The Inte...	3/24/17
★	●	📄	Tarabek, P	General vectorization of line objects in drawn maps	2009	Komunikacie	3/24/17
★	●	📄	Junhua, T; Fahui, W	Improved algorithm for data conversion from raster to vector	2007	Proceedings of SPIE - The Inte...	3/24/17
★	●	📄	Chiang, Y.-Y.; Leyk, S.; Knoblock, Craig	Efficient and robust graphics recognition from historical maps	2013	Lecture Notes in Computer Scien...	3/24/17
★	●	📄	Taie, Shereen; ElDeeb, Hesham; Atiya, Diyaa	A new model for automatic raster-to-vector conversion	2011	International Journal of Engi...	3/24/17
★	●	📄	Bodansky, E; Gribov, A; Pilouk, M	Smoothing and compression of lines obtained by raster-to-vector conversion	2002	Lecture Notes in Computer Scien...	3/24/17
★	●	📄	Chiang, Y.-Y.; Knoblock, Craig	Extracting road vector data from raster maps	2010	Lecture Notes in Computer Scien...	3/24/17
★	●	📄	Baily, Brian; Riley, Mark; Aucott, Paula; Southall, Hu...	Extracting digital data from the First Land Utilisation Survey of Great Britain – Methods, issues and potential	2011	Applied Geography	3/24/17
★	●	📄	Du, J; Zhang, Y	Automatic extraction of contour lines from scanned topographic map	2004	International Geoscience and...	3/24/17
★	●	📄	Southall, Humphrey; Brown, Nigel; Burton, Nick	Digitising the Inter-War Land Use Survey of Great Britain: A Pilot Project	2003		3/24/17

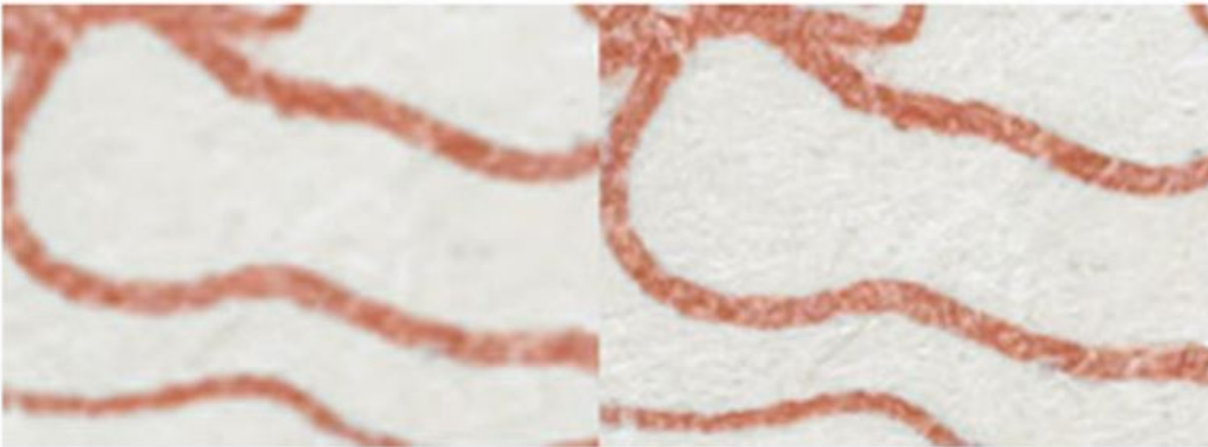


# The Process



# Scanning

- ▶ Poor original maps – try to borrow it from someone else
- ▶ Higher resolutions are not always better: 300-400 dpi sufficient
- ▶ 8-bit colour for grayscale maps; 24 bit colour for colour maps
- ▶ Output file format: TIFF

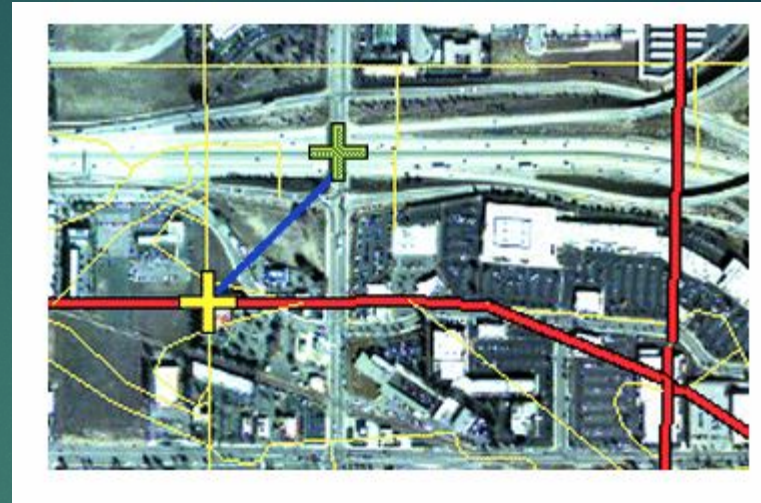


**FIGURE 10** Texture as noise in 300PPI (left) and 400PPI (right) variations.

Source: Pearson et al, 2013, 327.



# Georeferencing



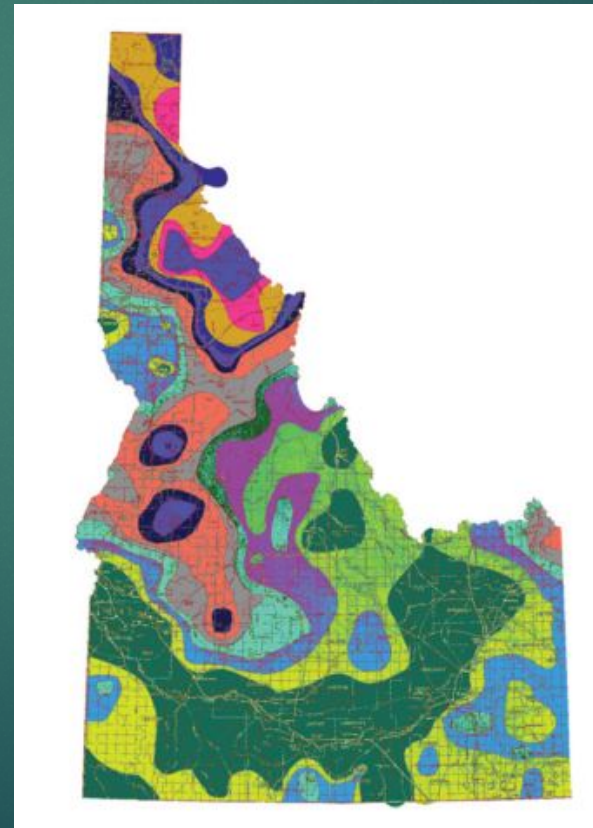
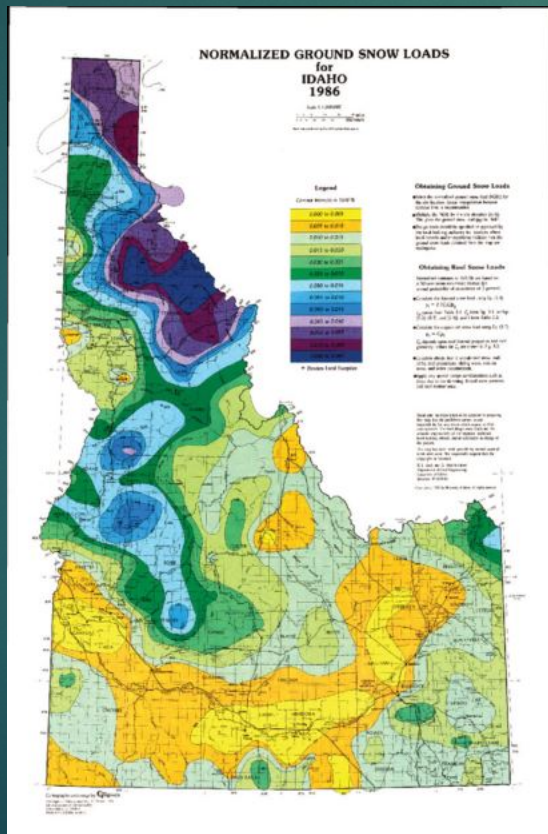
Source: ESRI, 2016

- ▶ Don't overdo the number of control points
- ▶ Root Mean Square of all the residual errors  $\leq$  cell size of the raster
- ▶ First Order Polynomial transformation usually sufficient in majority of cases for georectification
- ▶ Possibly employ crowdsourcing (ex/NYPL: MapWarper)



# Image Enhancement

- ▶ Resample image from archival quality scanned map
- ▶ Use a “mask” to exclude extraneous parts of the map

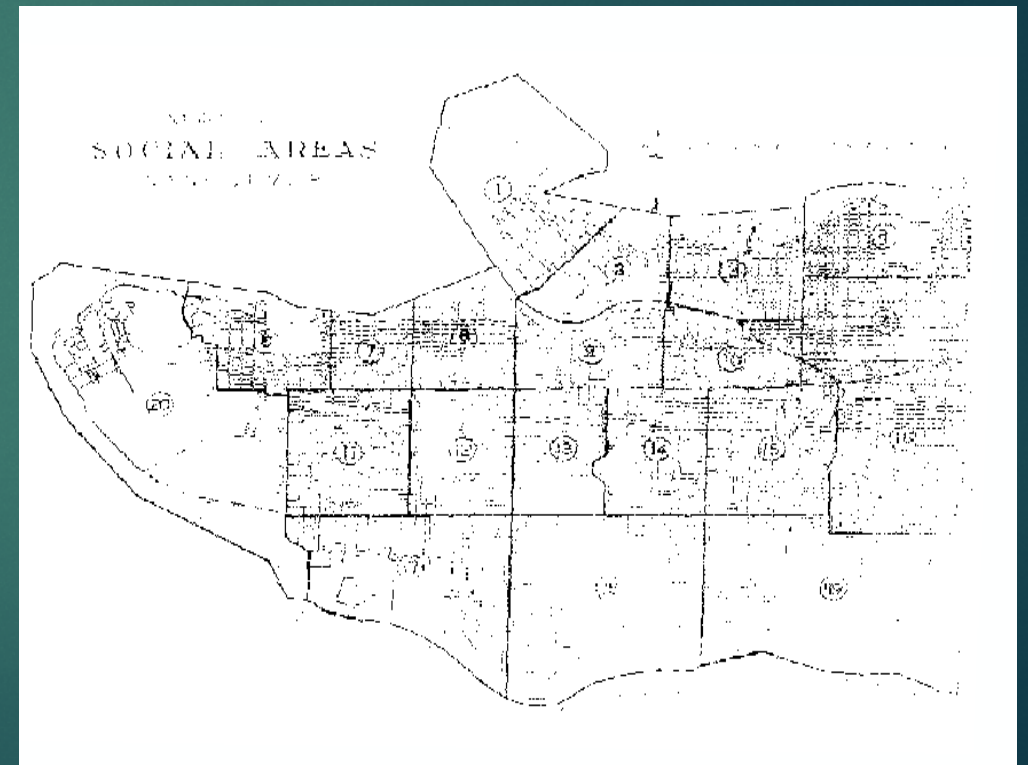
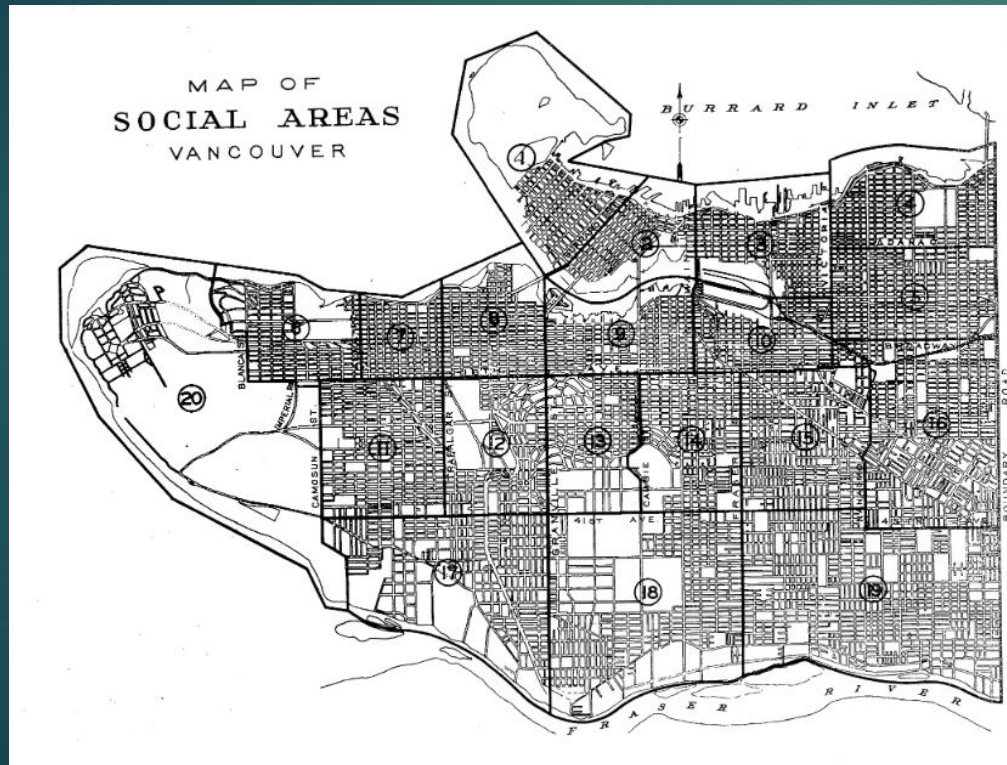
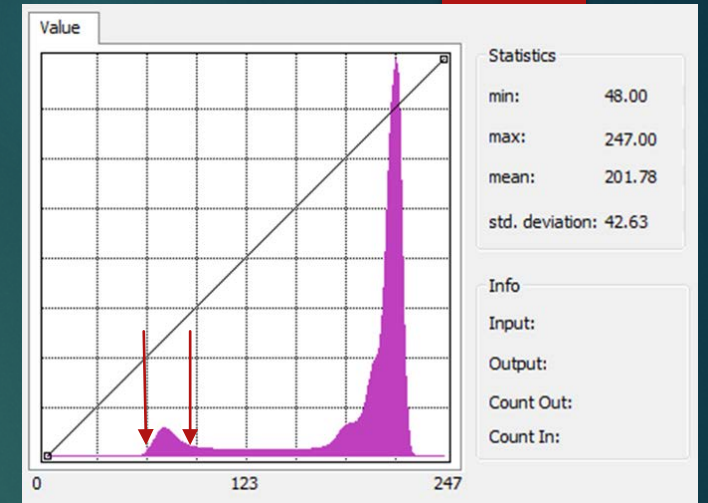


Source: Godfrey and Eveleth, 2015



# Image Segmentation

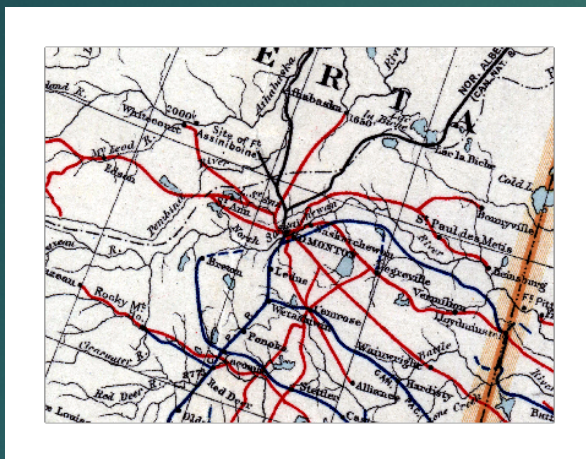
- Simple thresholding used with greyscale images



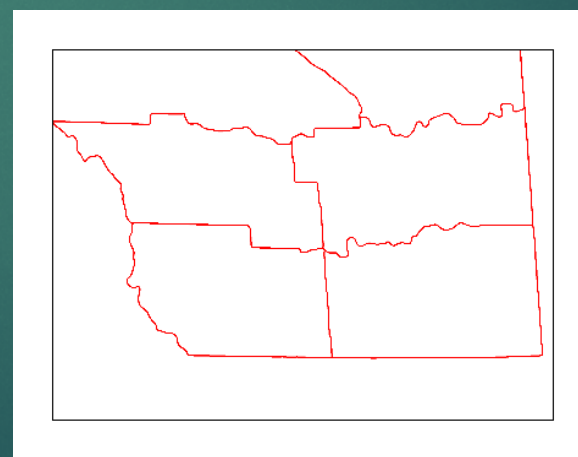
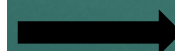
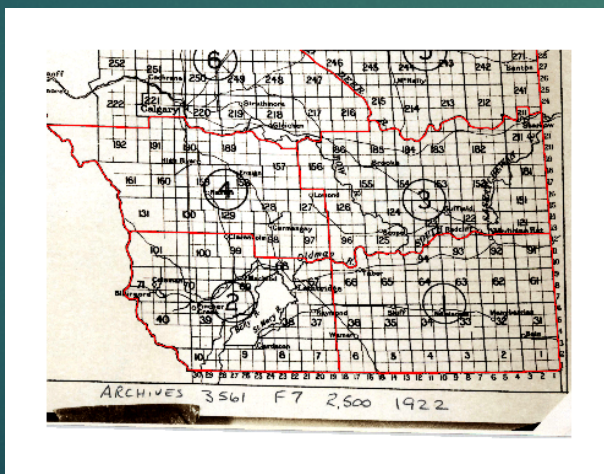


# Image Segmentation – Linear Features

a)

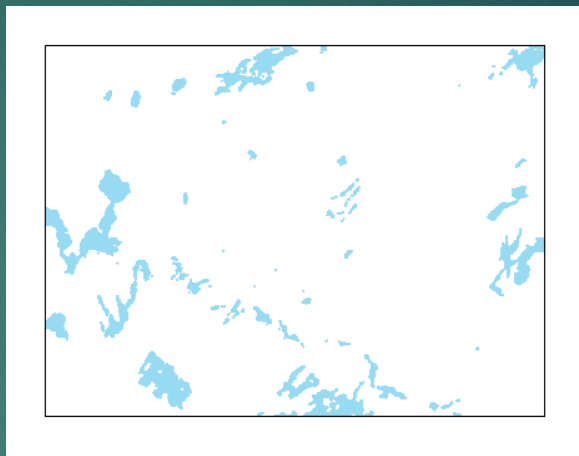
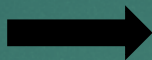


b)

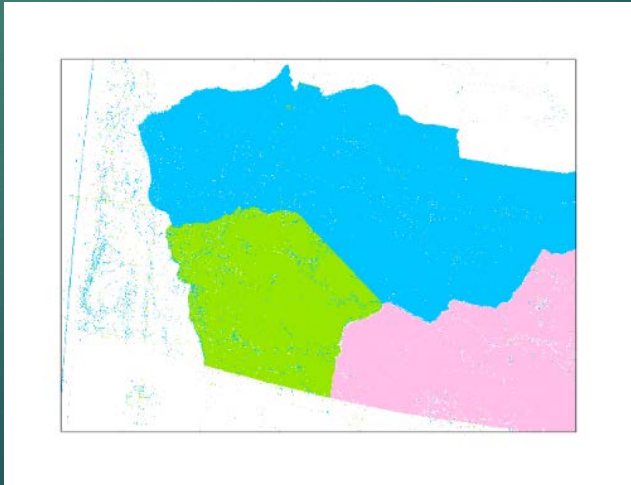
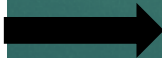




This is a historical map of the Yukon River region, showing the Yukon River, Klondike River, and various settlements and landmarks. The map is titled "Yukon River" and includes a scale bar. Key locations marked include Fairbanks, Fort Reliance, and various Indian reservations. The map is dated 1890.



A historical map of the Balkan region, showing the borders of the Ottoman Empire (shaded in blue) and the Kingdom of Serbia (shaded in green). The map includes geographical features like mountains and rivers, and a small inset map in the bottom left corner.



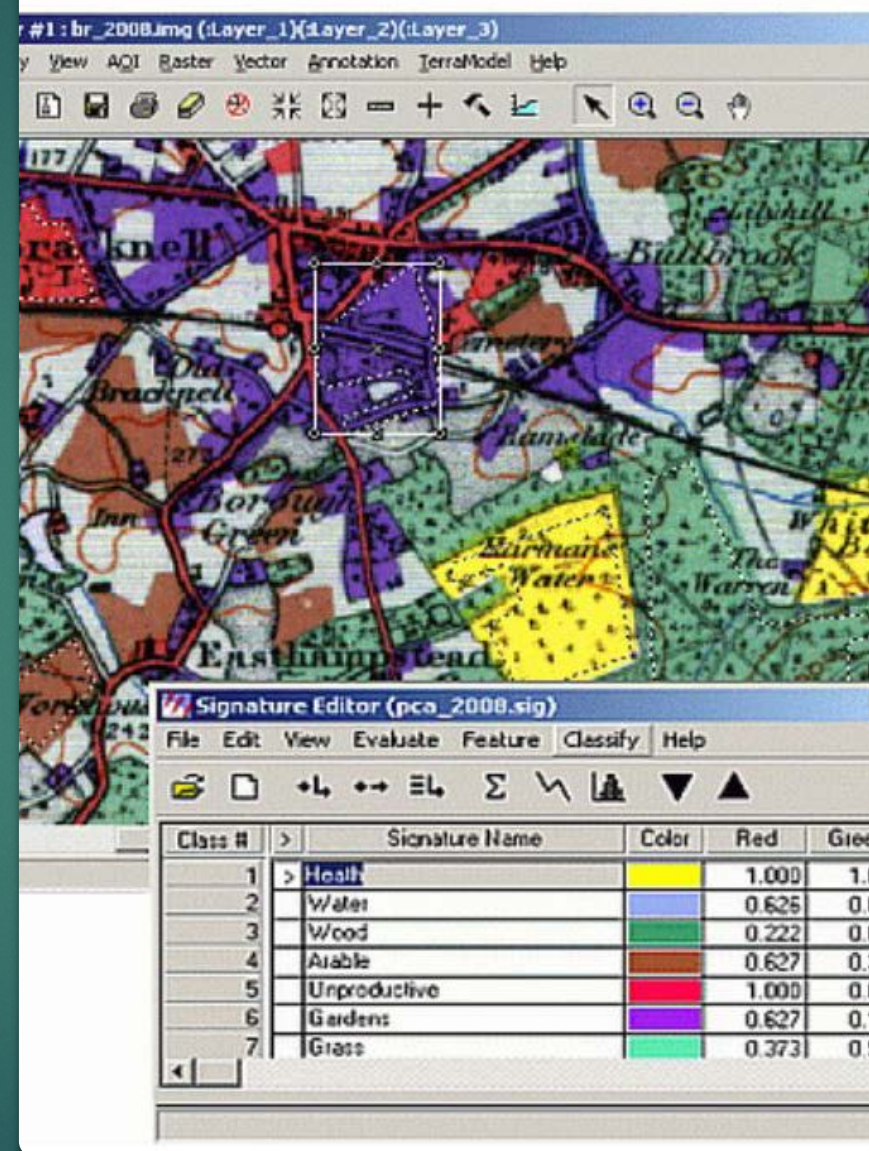


# Image Segmentation: Supervised Classification

## Training Sample Selection

- Select from cluttered areas
- Select a number of samples
- Can be reused with maps in same series

Figure 19: Typical training areas are chosen and digitised

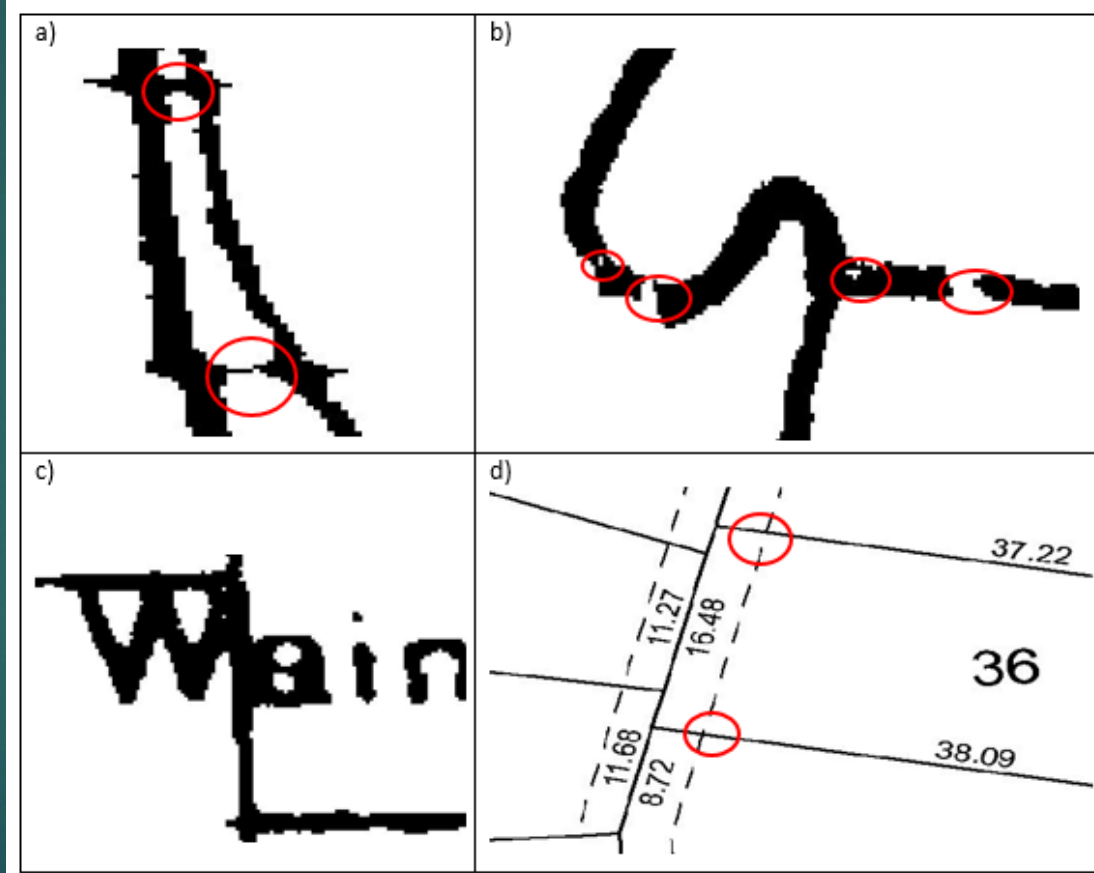


Source: Southall et al., 2003, 24



# Raster Editing – Problems

Bridges



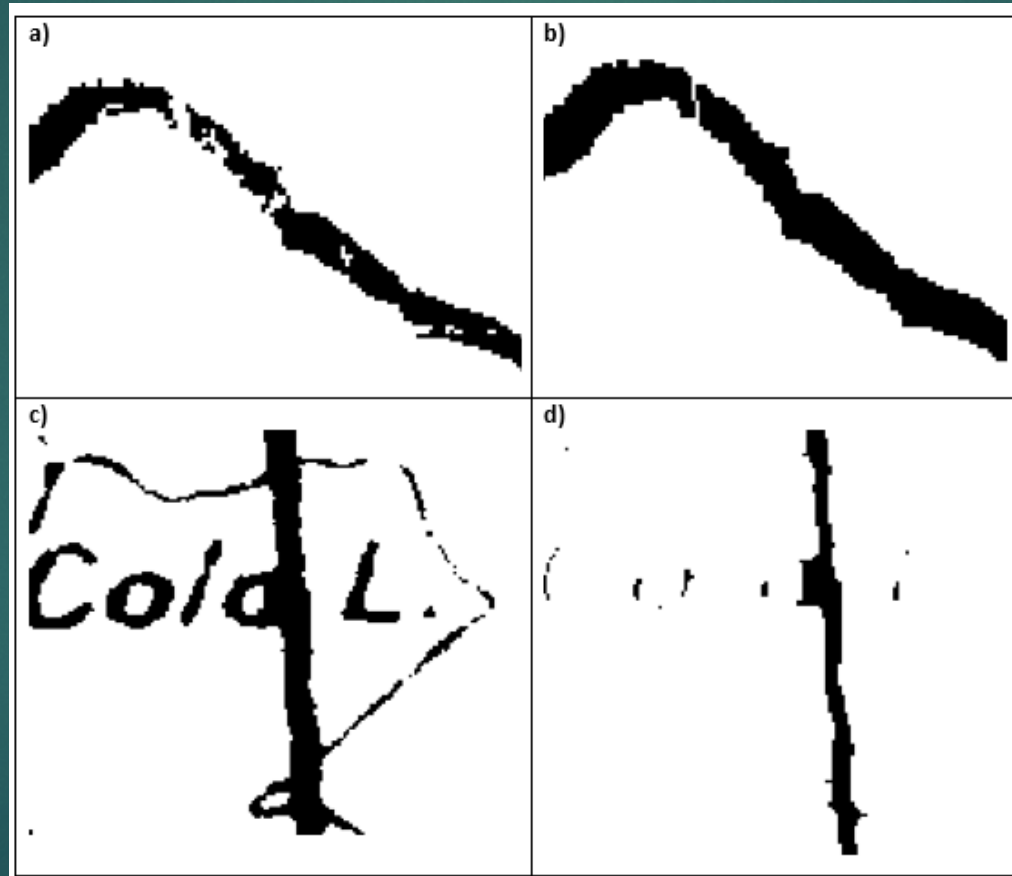
Text overlapping  
feature

Gaps and holes

Intersecting features



# Raster Editing – Morphological Operators

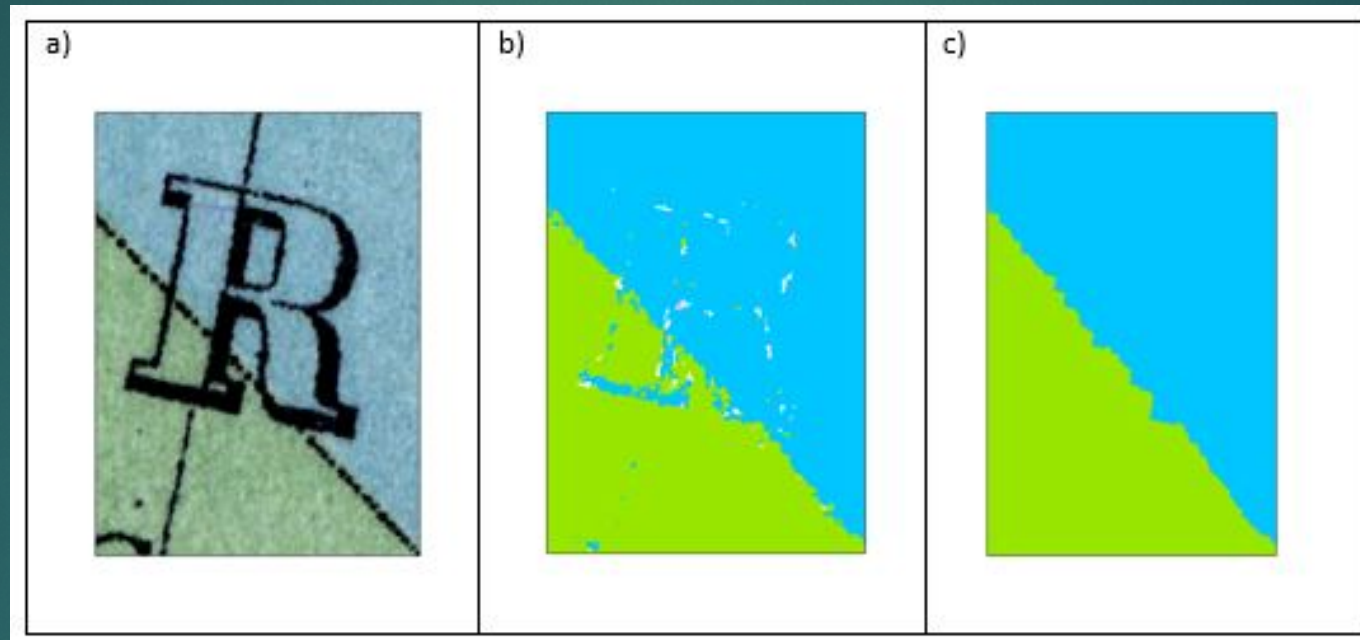


Dilate

Erode



# Raster Editing – GIS Functions



Original

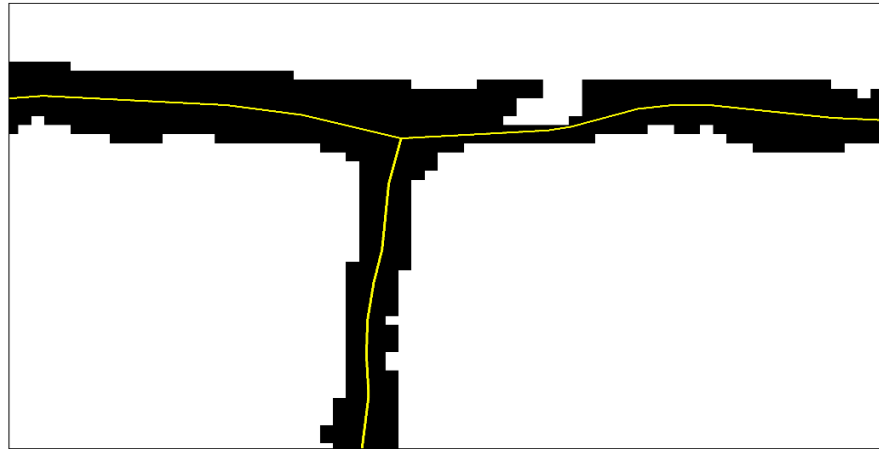
After  
Segmentation

Applying  
Majority Filter,  
Expand & Shrink



# Vectorization

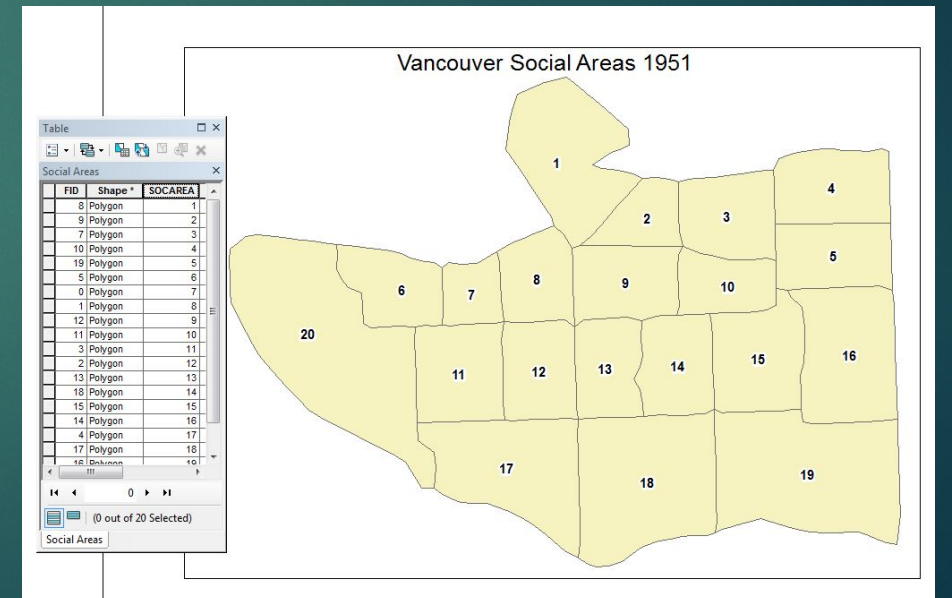
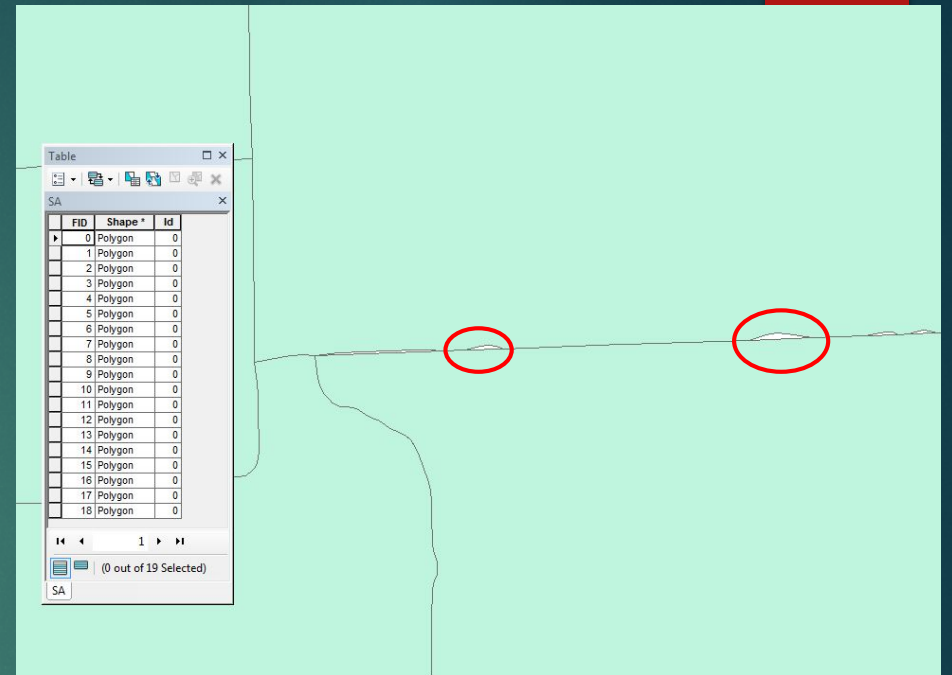
- ▶ ArcScan and R2V – lines and outlines
  - ❖ Automated vectorization and interactive raster tracing
- ▶ ArcGIS' "RasterToPolygon" – polygons





# Vector Editing

- ▶ Use ArcGIS Eliminate function to dissolve areas of unwanted map elements (below a minimum threshold) and No Data gaps
- ▶ Use Smooth Line and Smooth Polygon tools to make these features look better
- ▶ Add attribute data to vector feature classes through the use of lookup tables to transfer multi-variable data to features





# Conclusions

## ► Bad News:

- ❖ There is no easy-to-use one-size-fits-all automated solution that works for all map types and features.

## ► Good News:

- ❖ There is a semi-automated process that can be adopted by libraries and archives to vectorize scanned maps with a few caveats:
  - The process does require some GIS skills.
  - Some experimentation required to determine optimal process.
  - Although significantly quicker than heads-up digitizing, it still takes time and effort.
- ❖ Vectorization enables all kinds of new possibilities for map features.