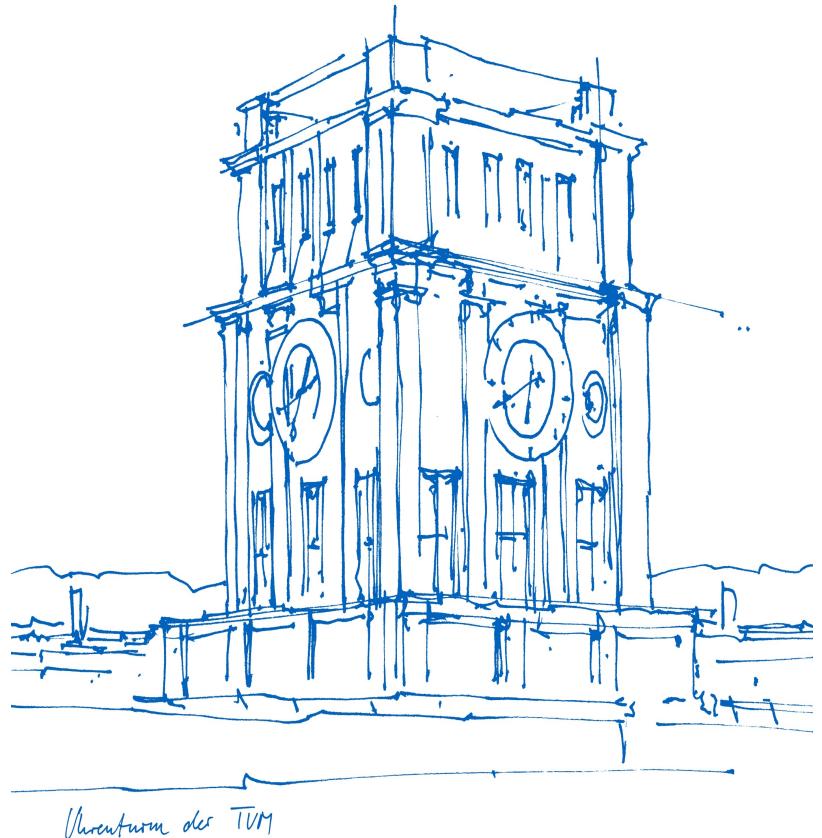


Geoinformation

Prof. Dr.-Ing. Liqiu Meng

Chair of Cartography and Visual Analytics

WS 24/25



Contents of Lectures

L1: Introduction

L2: Spatiotemporal representations and databases

L3-L4: Spatial data analysis

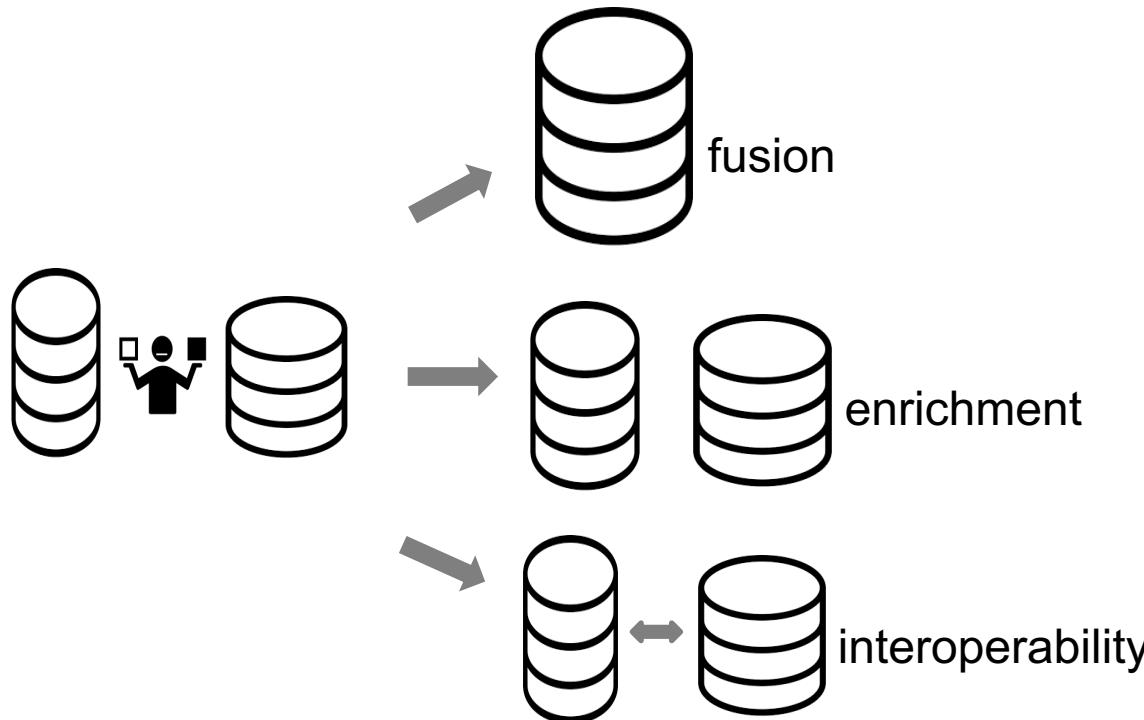
L5: Selected cartographic techniques

Cartographic techniques

I Spatial data integration

II Map generalization algorithms

I. Spatial data integration

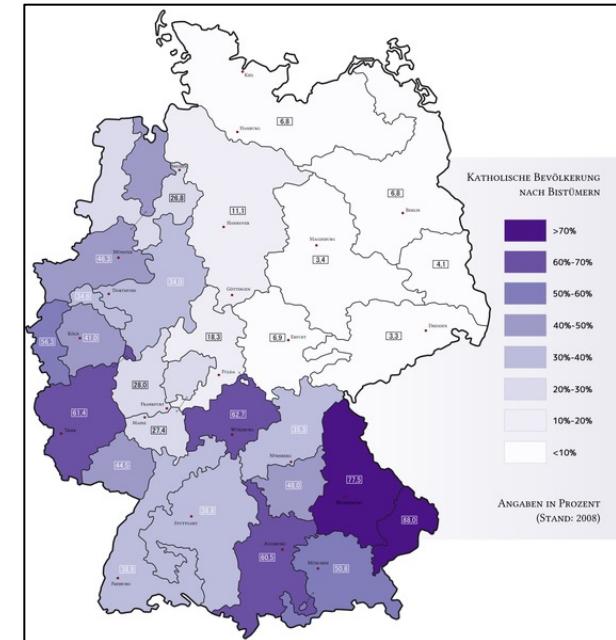
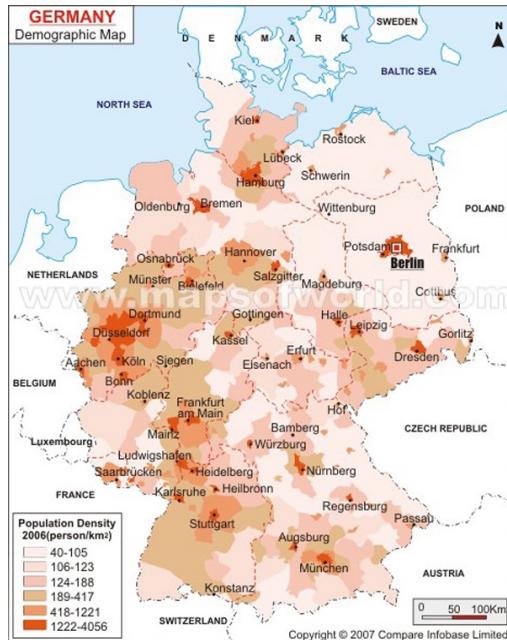


Differences in

- themes
- coordinate systems
- spatiotemporal scales
- data suppliers

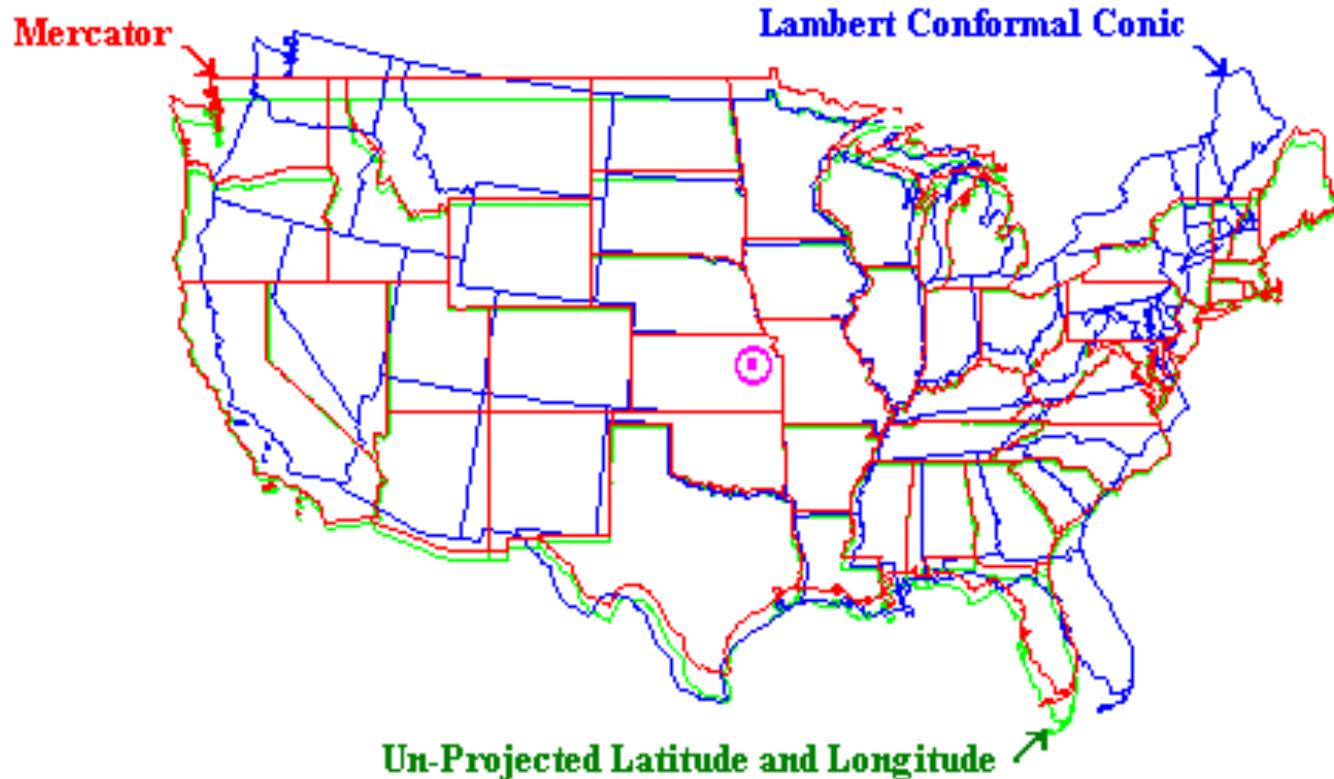
- Bring many data sources describing the same region together (fusion)
- Enhance the richness and quality of the individual datasets (enrichment)
- Link different datasets (interoperability)

Different themes

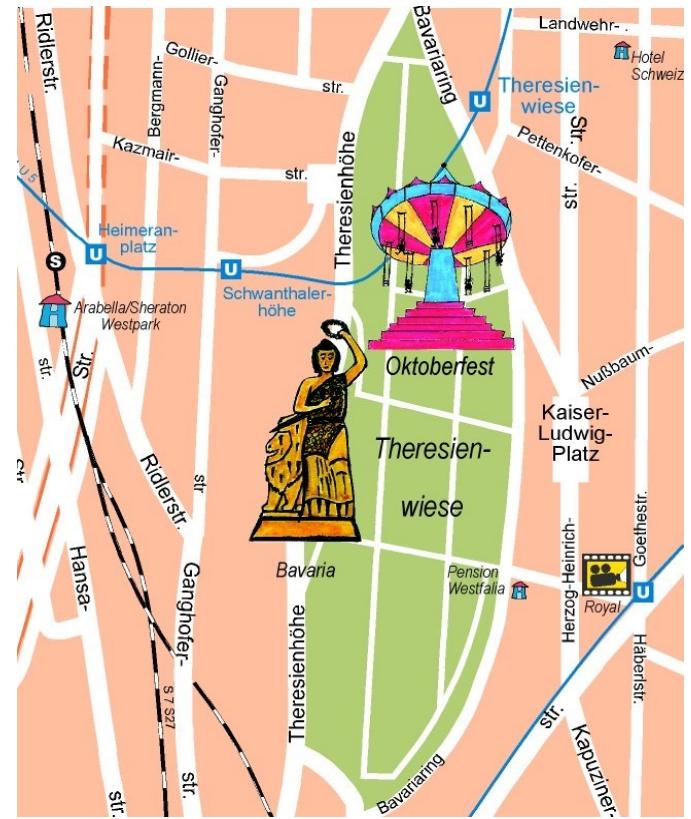


Different map projections

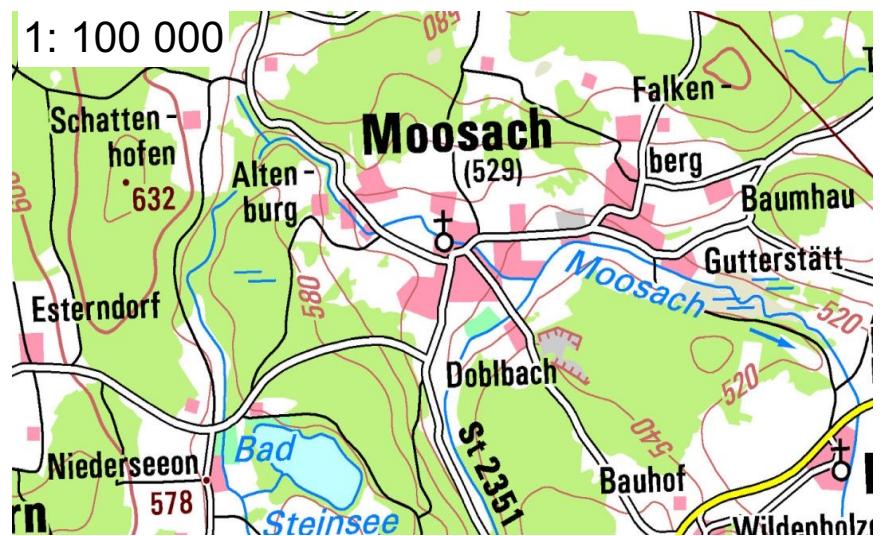
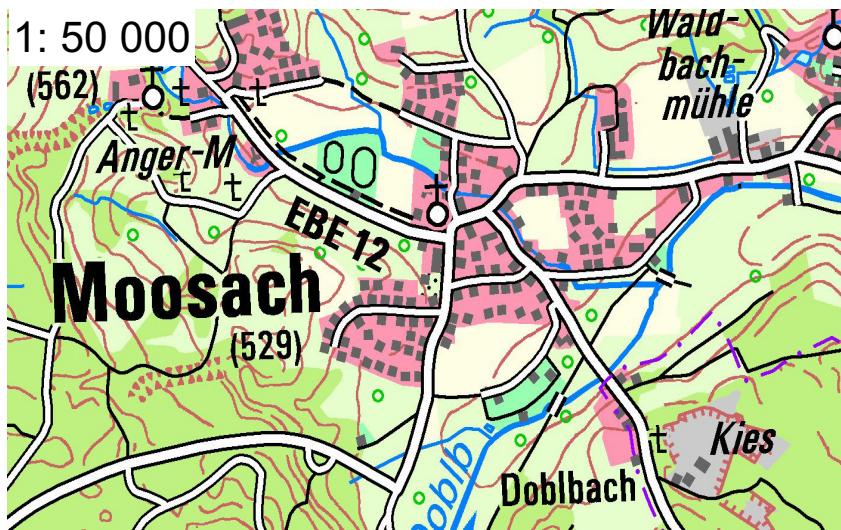
Three Map Projections Centered at 39 N and 96 W



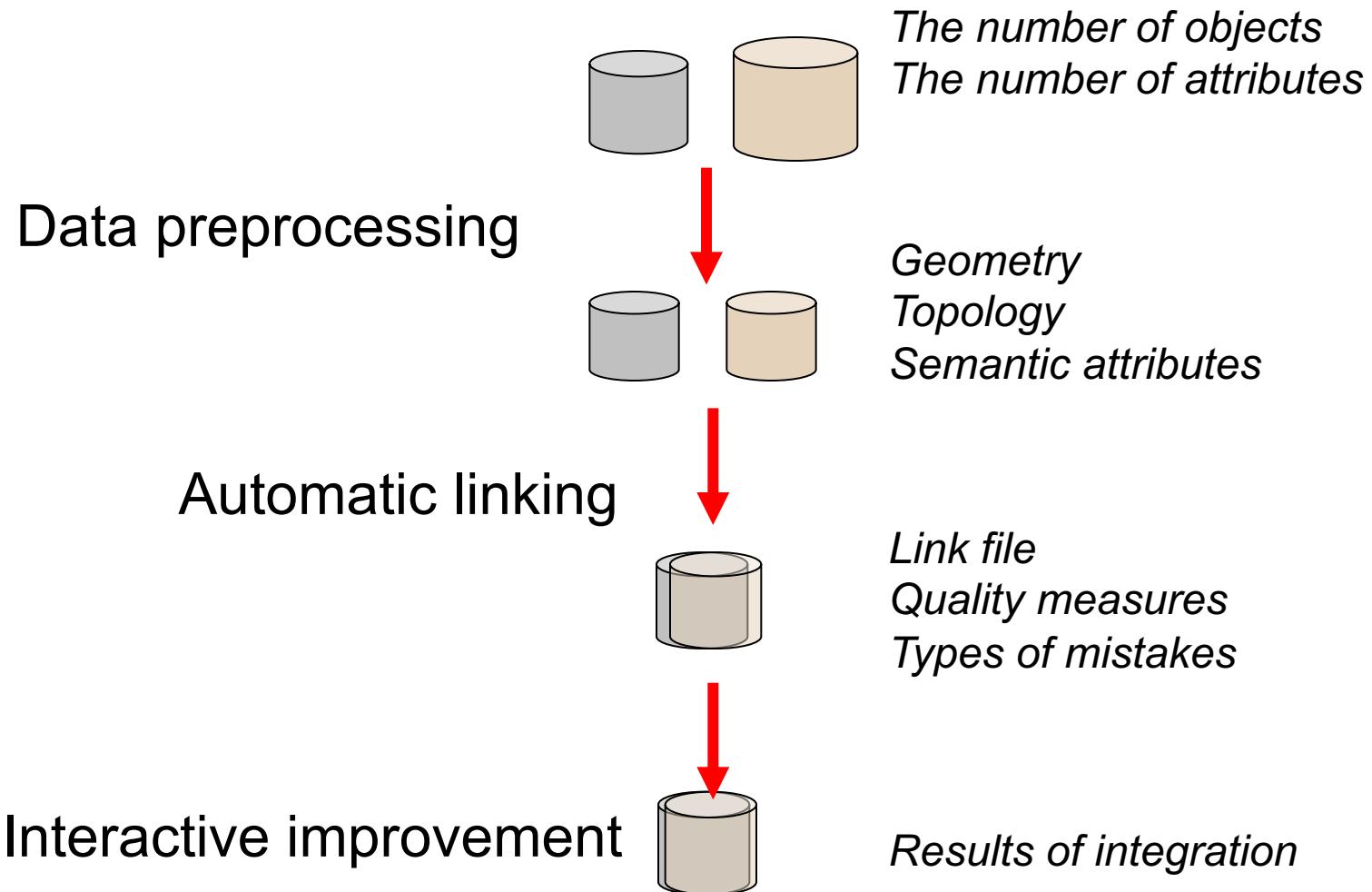
Different geometries driven by purposes



Different map scales



Working principle of spatial data integration



A case study of data integration

Task

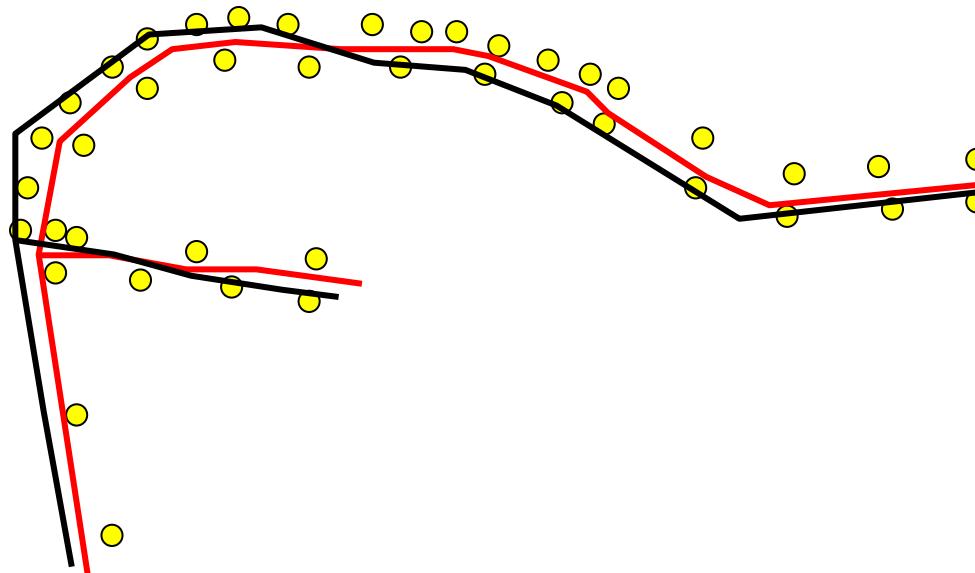
Integration of house number with streets from Basis-DLM

Problem

House number were captured based on TomTom street data

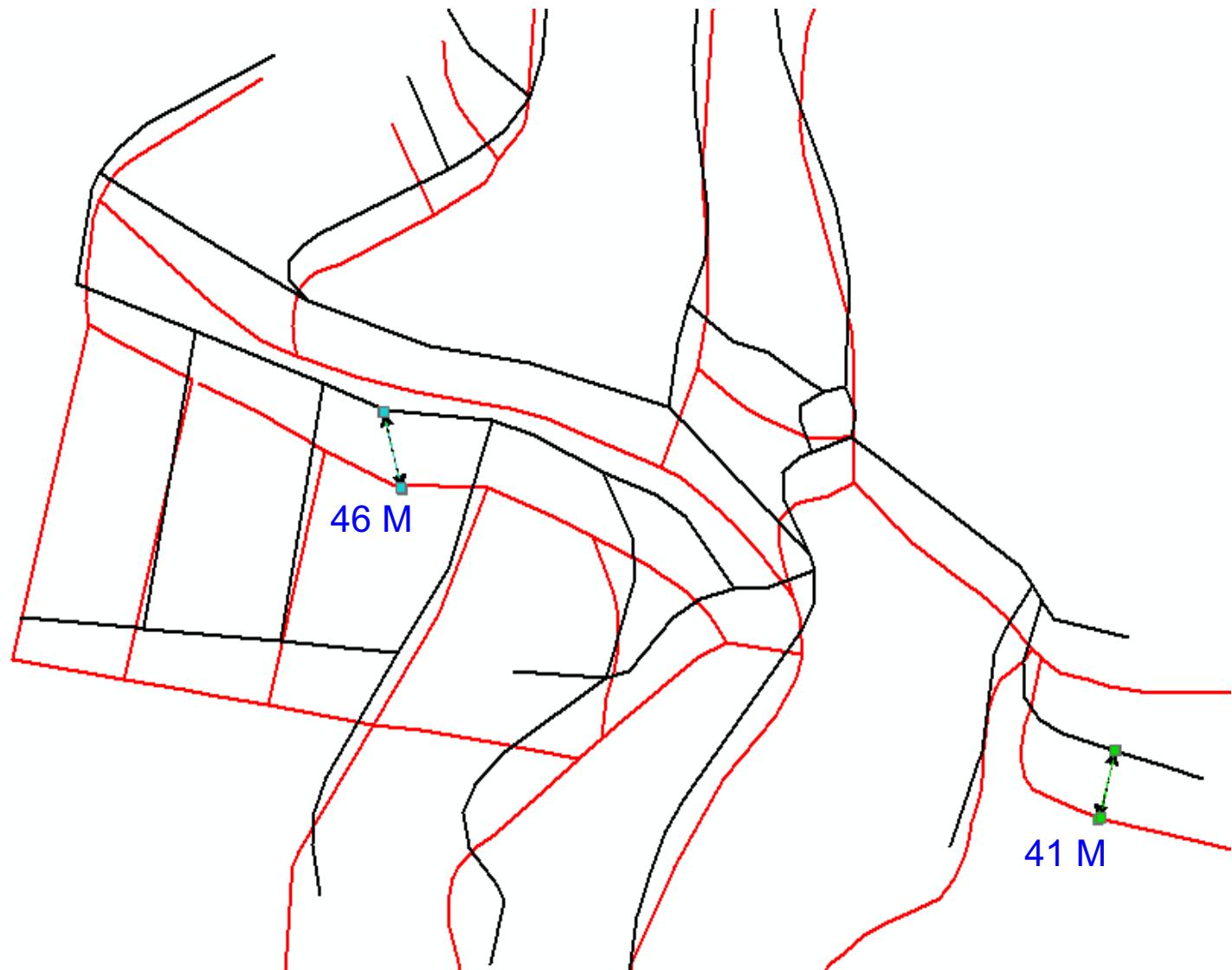
Solution

Matching between TomTom and Basis-DLM

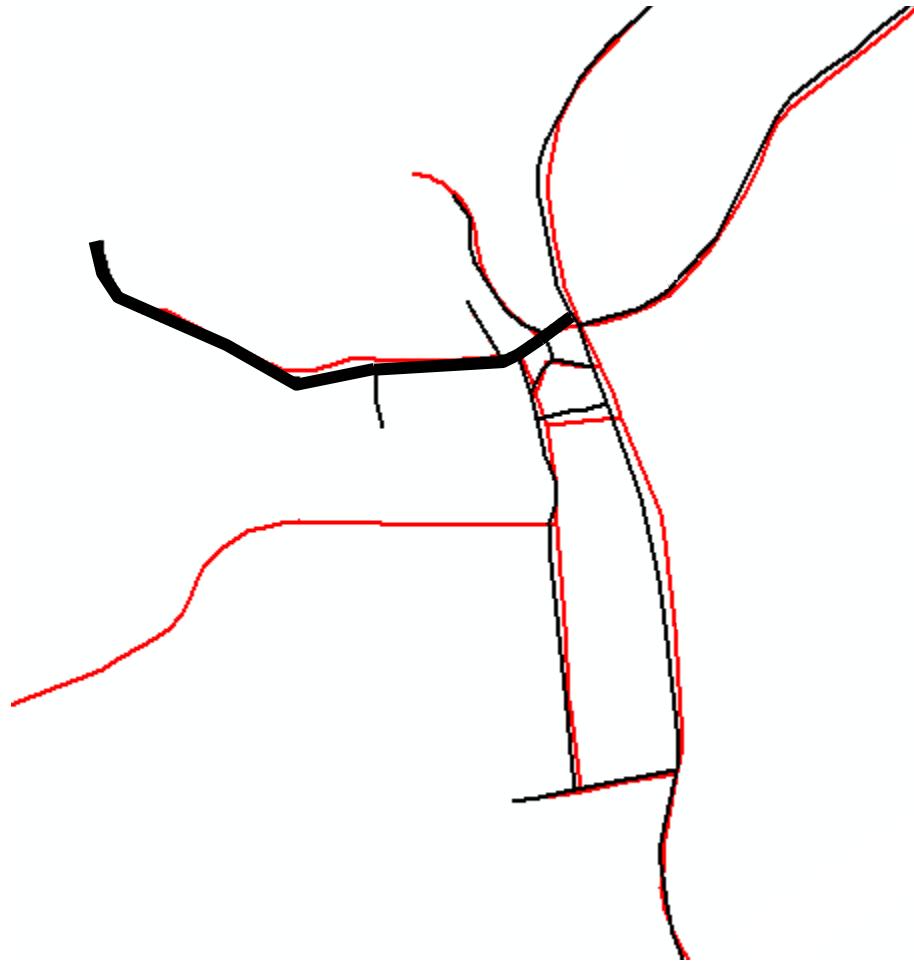




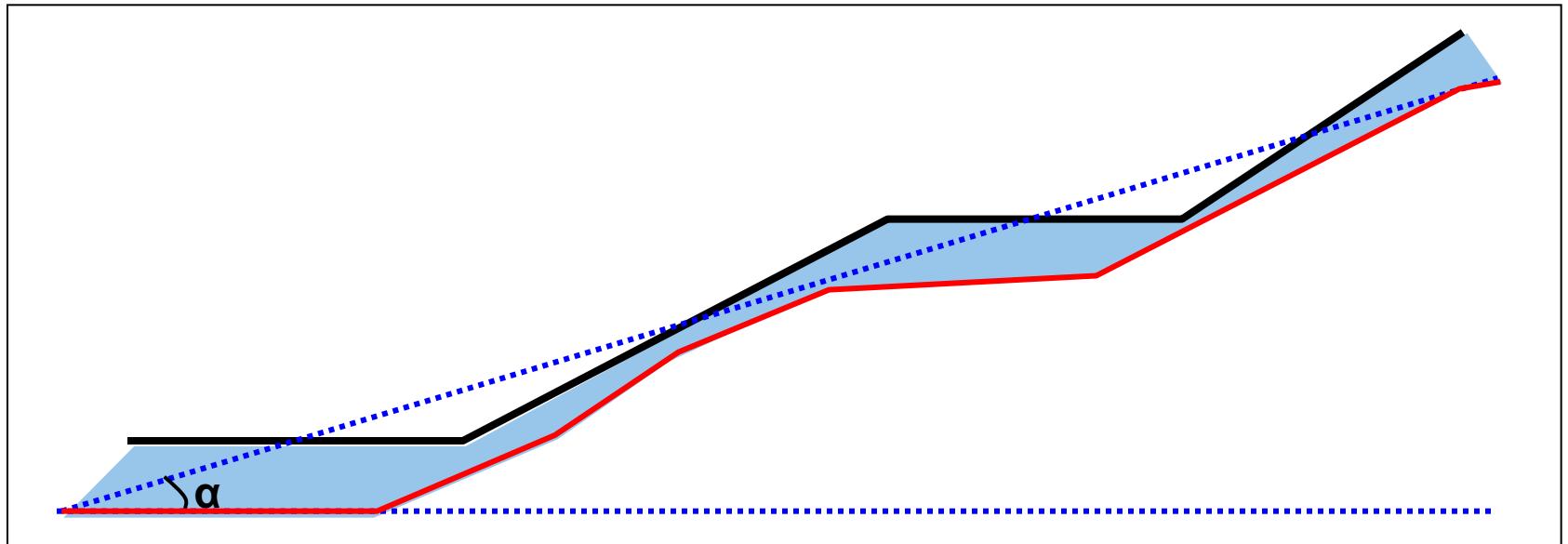
Reference data: Basis-DLM (black)
Candidate data: TomTom (red)



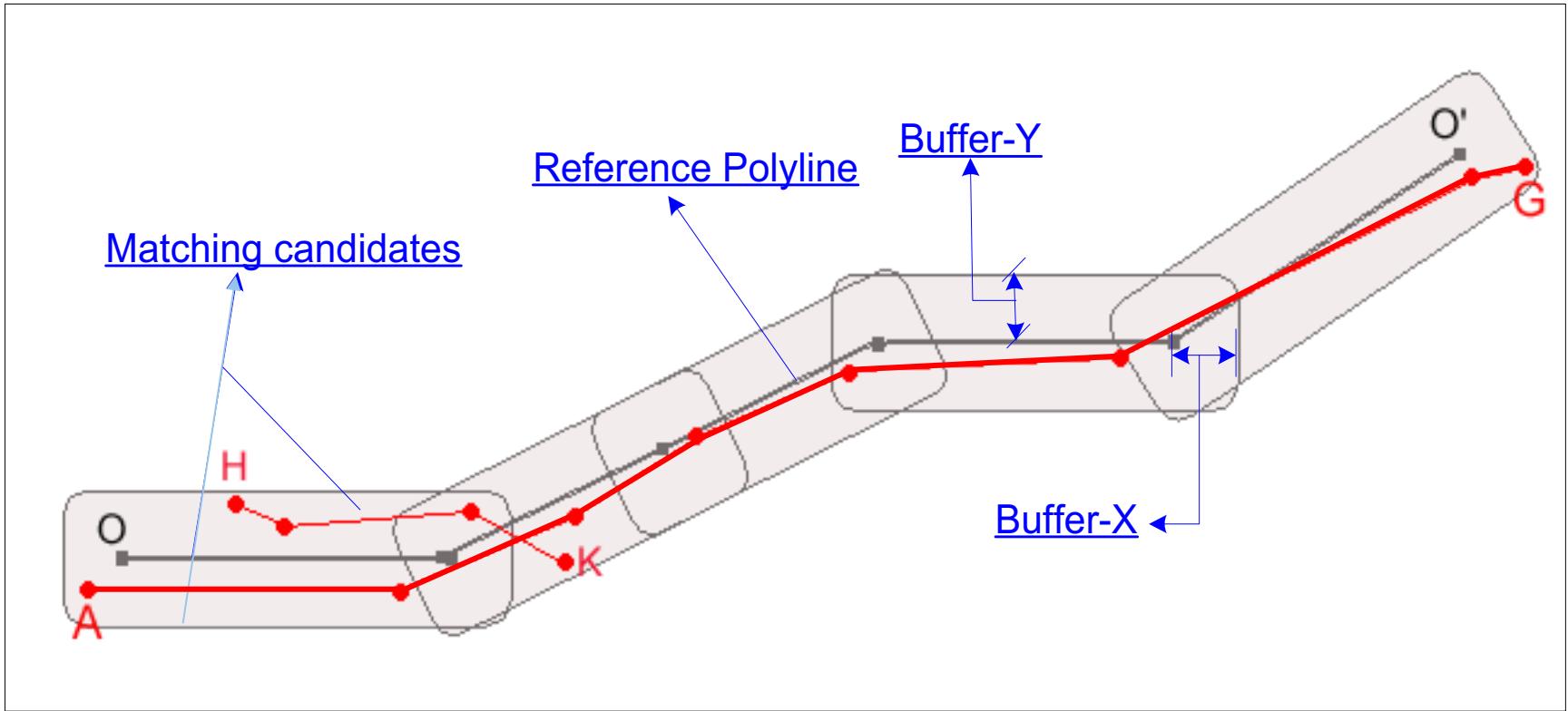
Analysis of geometric discrepancy



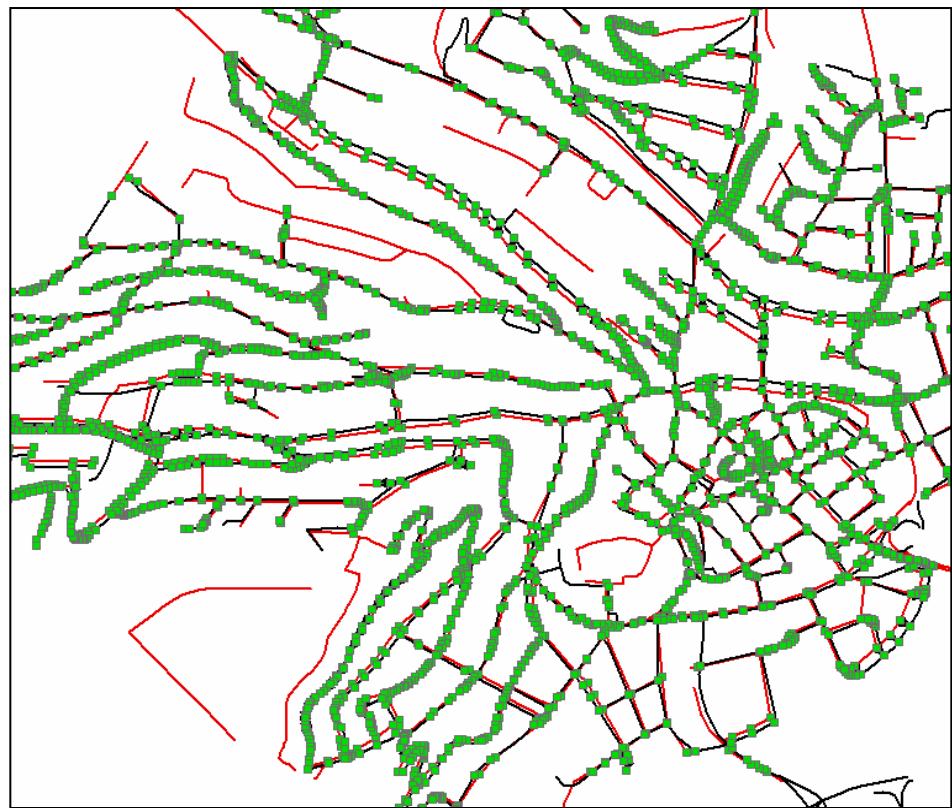
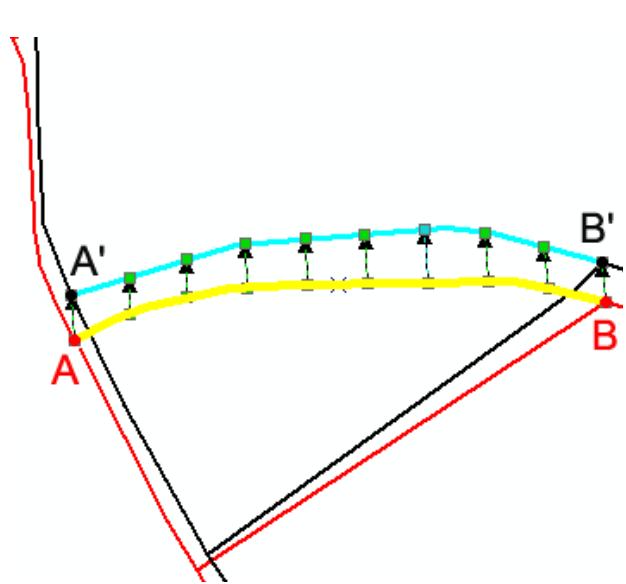
- Chain the segments sharing some common attribute values into long strokes,
- Compare the similarity of corresponding strokes,
- Grow to further strokes in the network



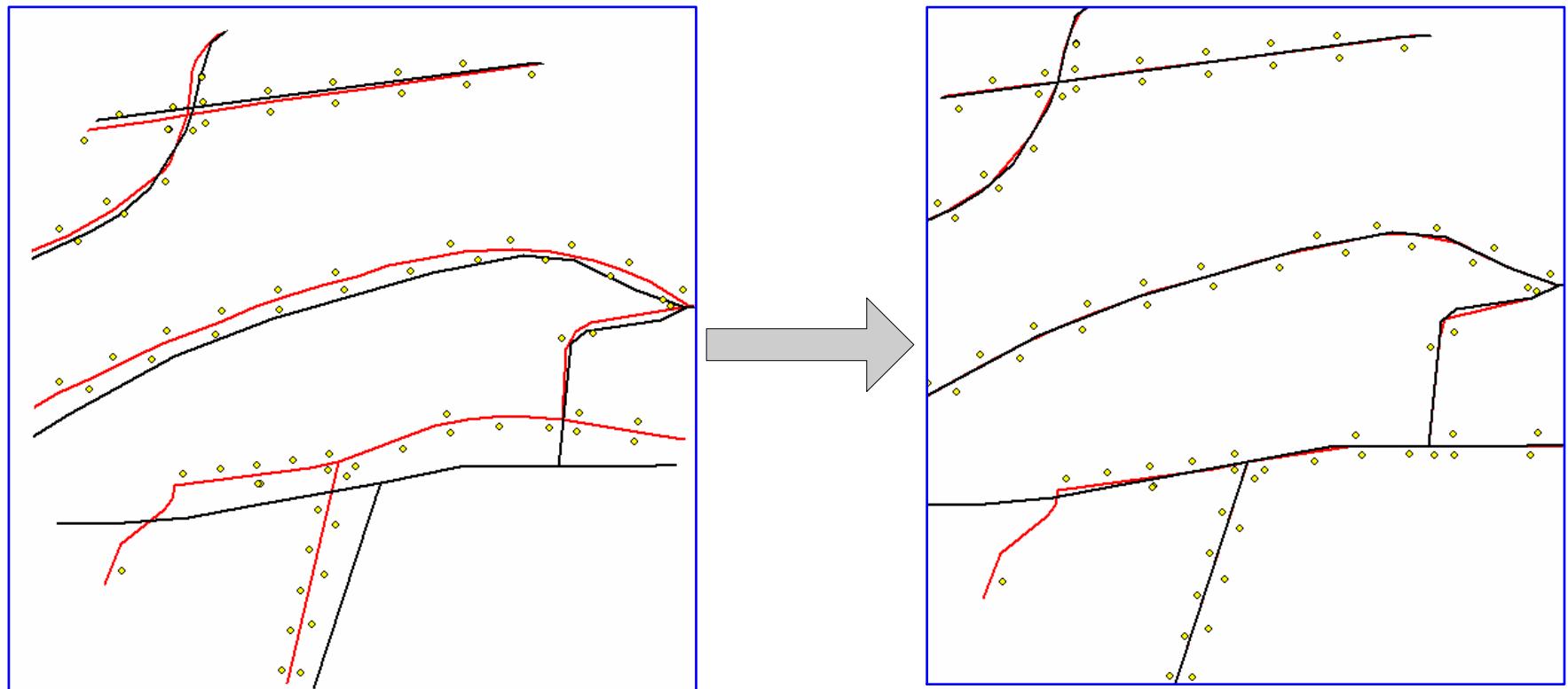
Computing the similarity of each candidate according to a weighted criterion (length, angle, area etc.)



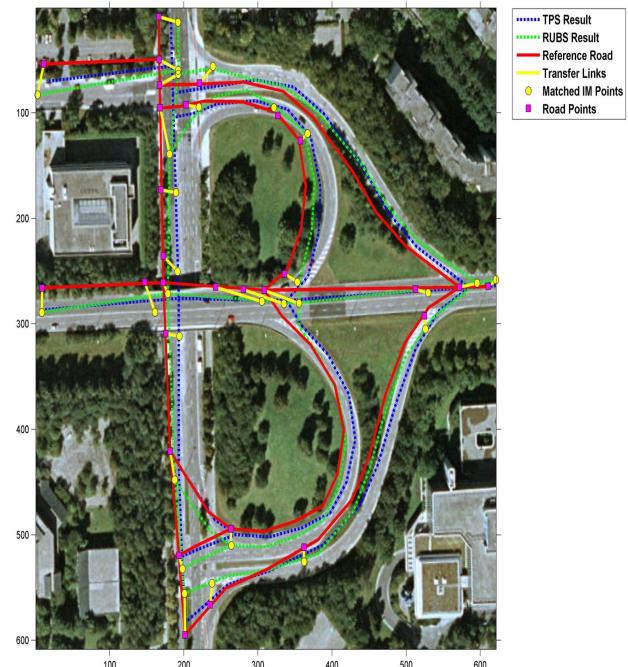
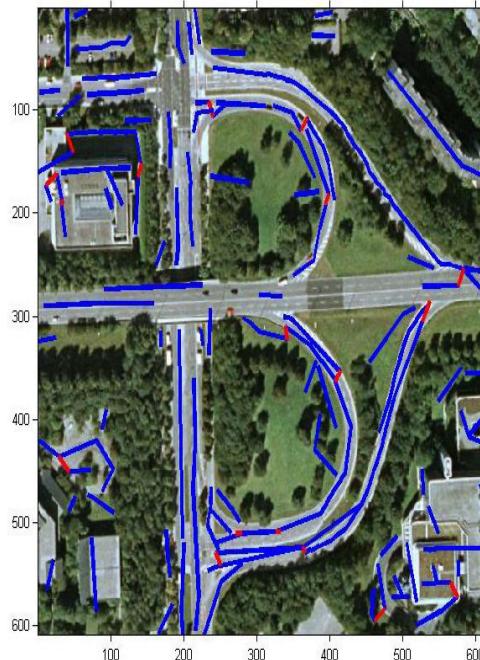
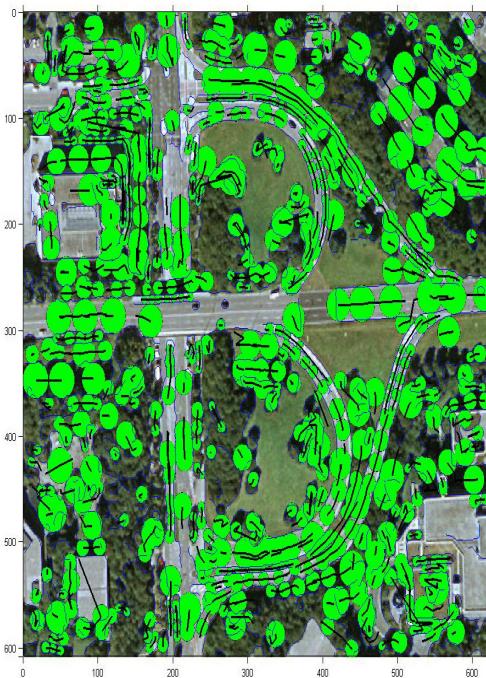
- Constructing buffers around segments and identifying matching candidates from the other database
- Selection of candidates with the largest similarity (\geq threshold)



Establishment of links between line segments



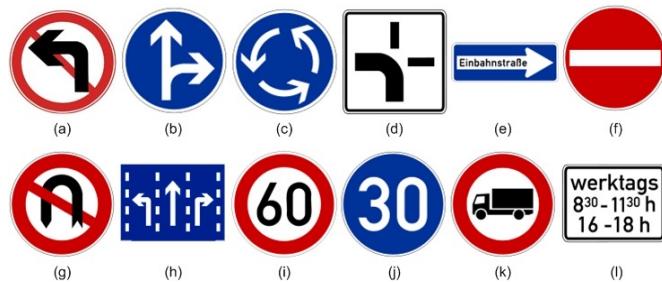
Project house numbers and other POIs to the reference database
based on the rubber-sheeting principle



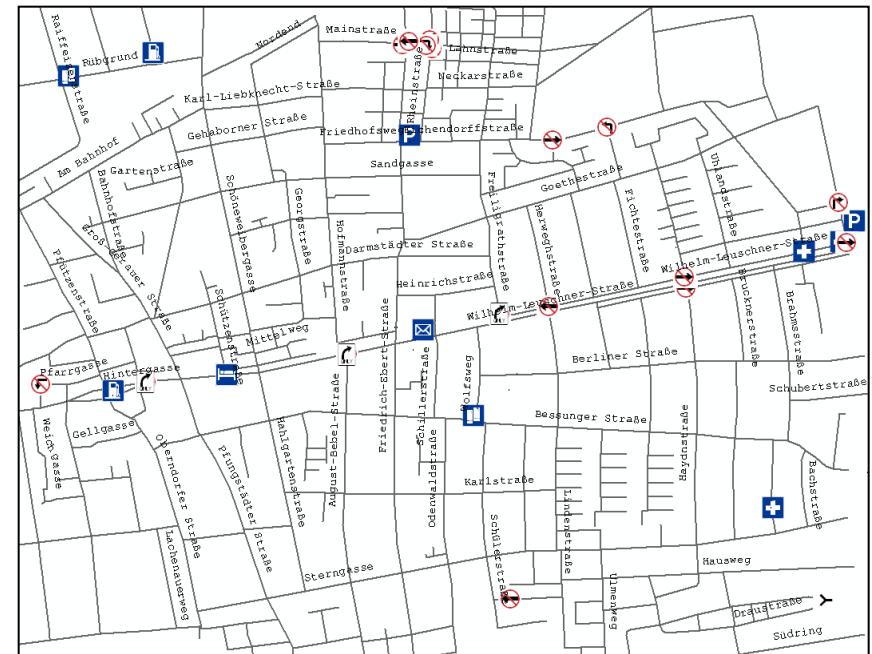
Matching between raster and vector data

- Identify middle axes belonging to road segments in the raster image
- Chain them into longer strokes
- Align the strokes with vector roads

Enrichment with navigation / landmark information



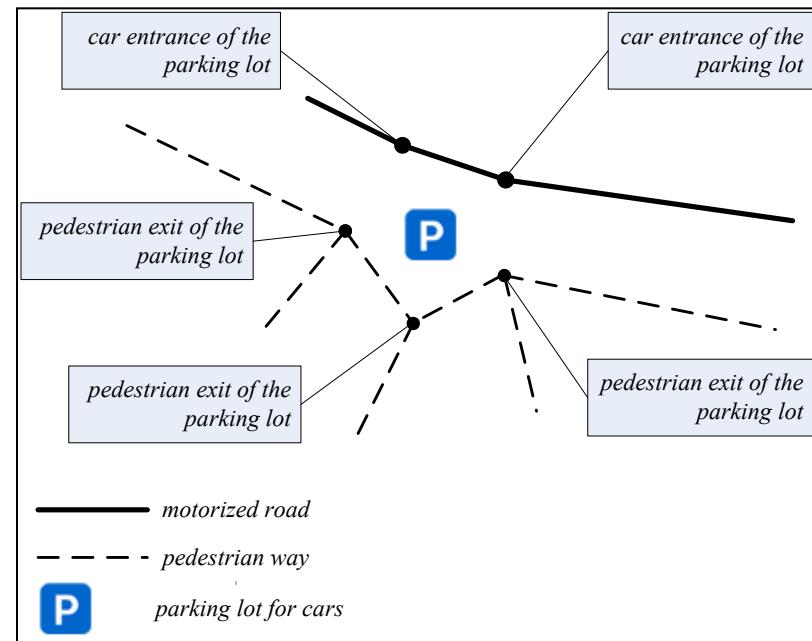
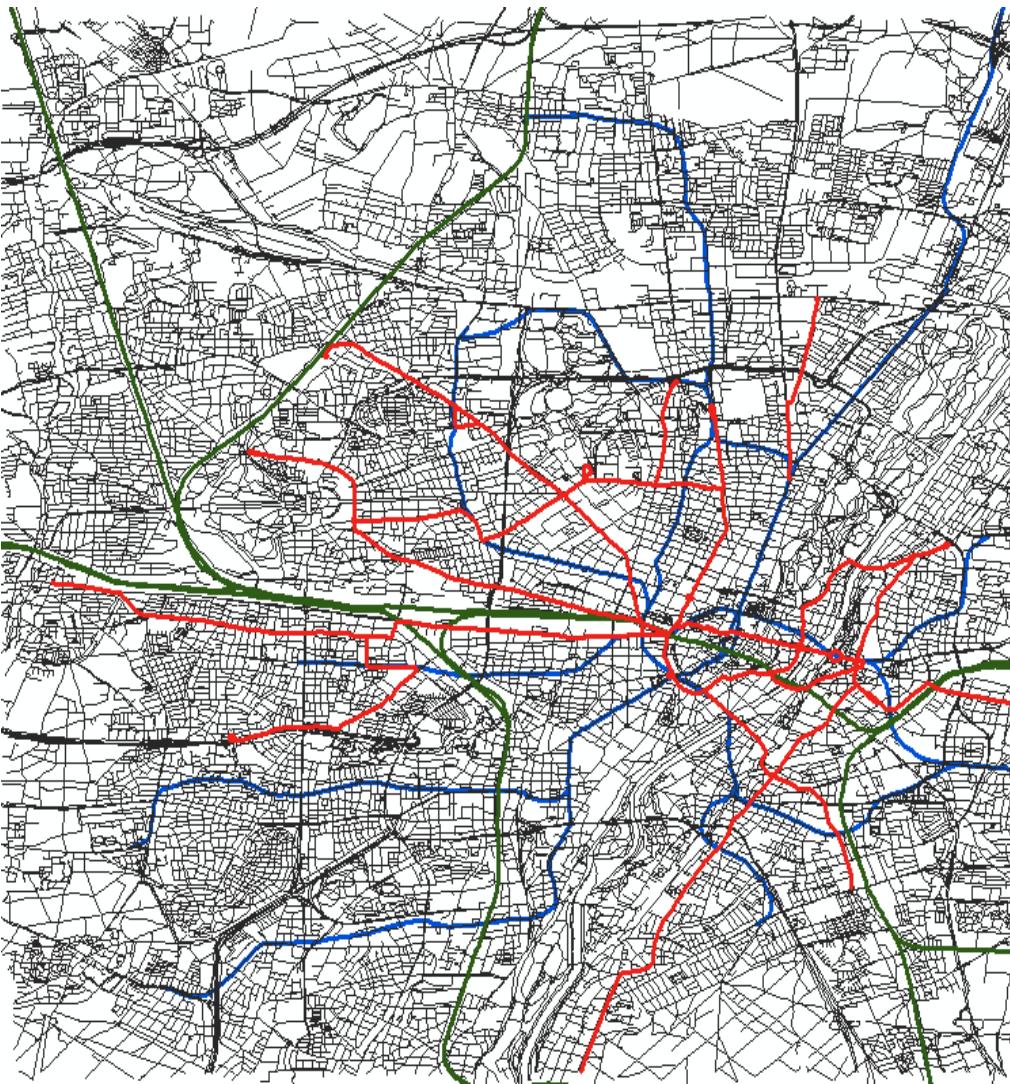
streets from digital landscape model



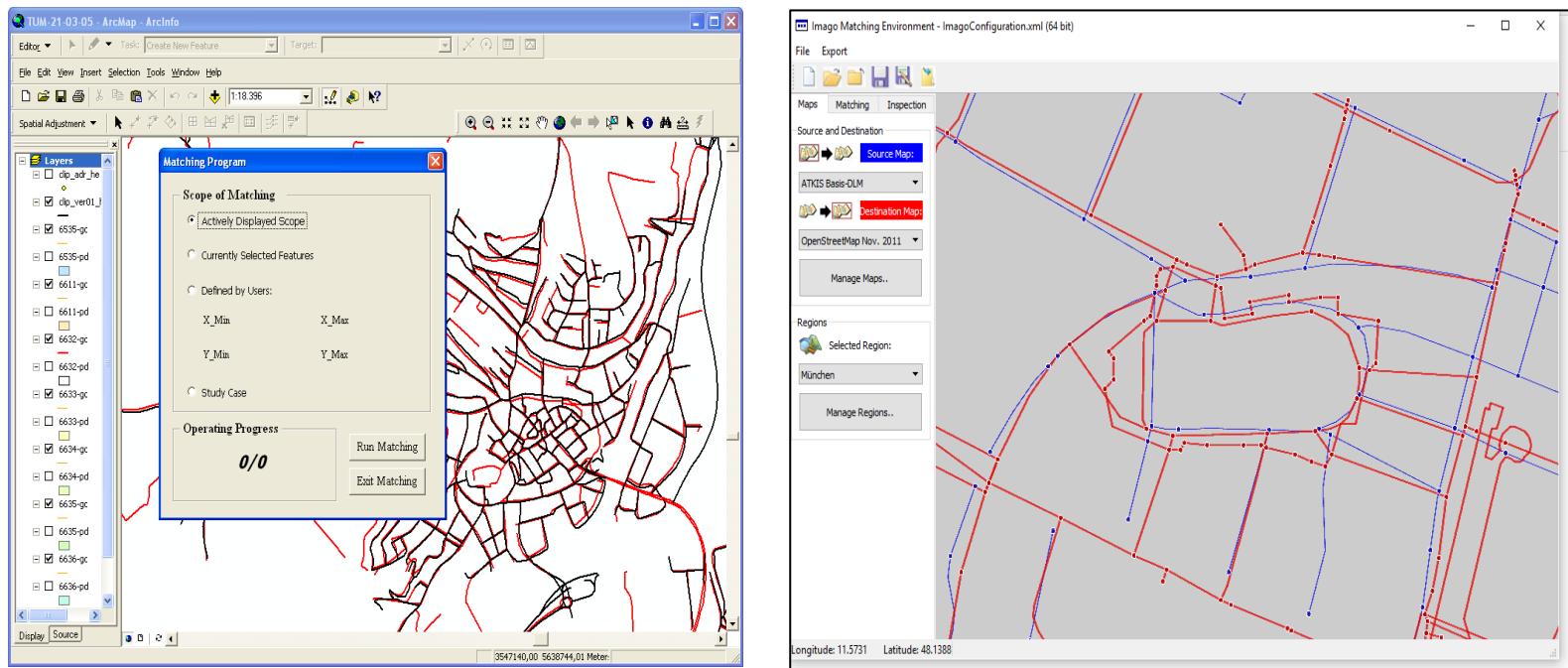
after integration



Cell-based docking of pedestrian roads (grey) to motor ways (brown)



Fusion of motorways, pedestrian roads and public transport network in Munich (Tram - red, S-Bahn – green, U-Bahn – blue)



Algorithms

- Delimited stroke
- Buffer growing
- Iterative Hierarchical Conflation
- Sparse matching
- Conditional Random Field
- Switch-point docking
- ...



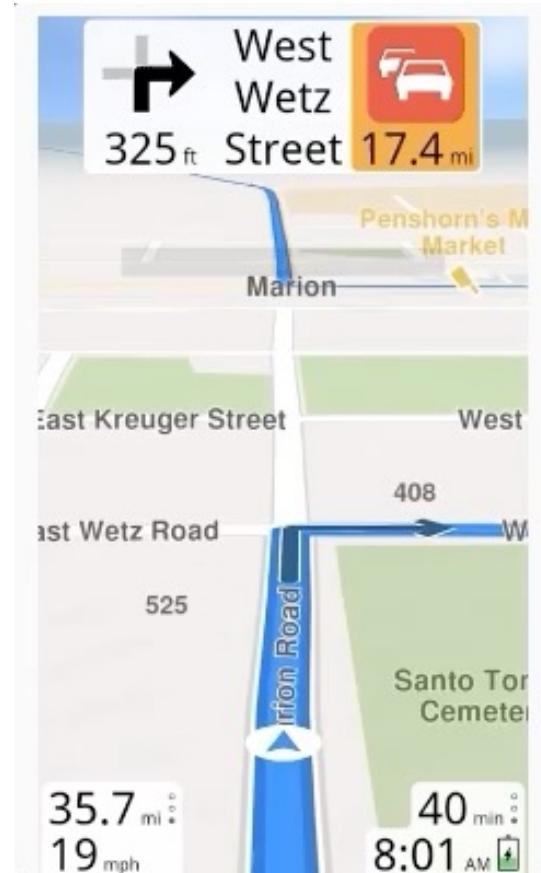
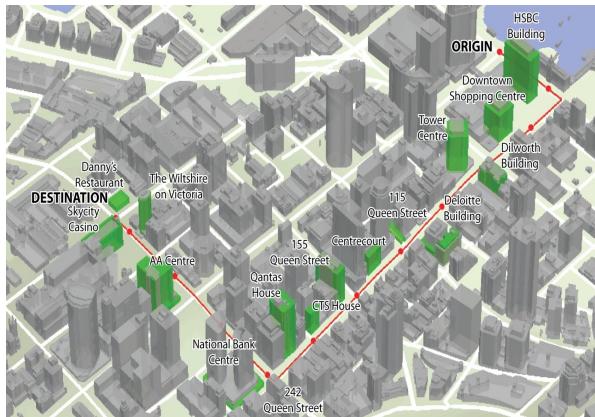
Quality indicators



- Connectivity
- Computing speed
- Robustness
- Repeatability
- Transferability
- Transparency
- ...

Applications of data integration

Landmark-based navigation



Multimodal navigation



Augmented reality

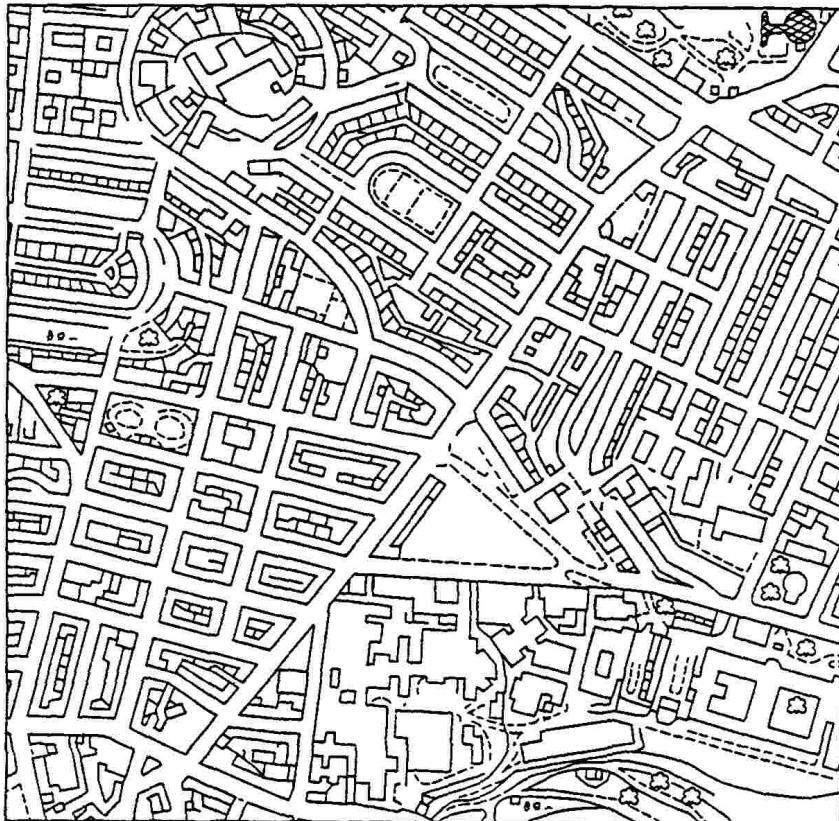
II. Map generalization algorithms

Goals

- Reduce the data amount
- Enhance the abstraction level
- Harmonize the level of details of different data sources
- Improve the legibility of the map

Generalization leads to

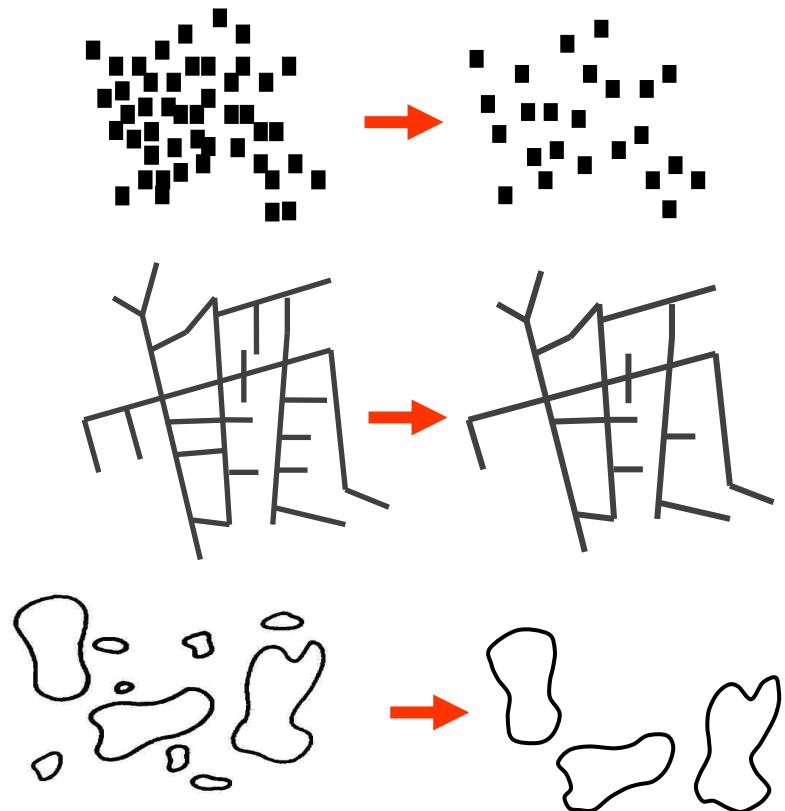
- continuous simplification and
- abrupt changes



Automatic generalization operations

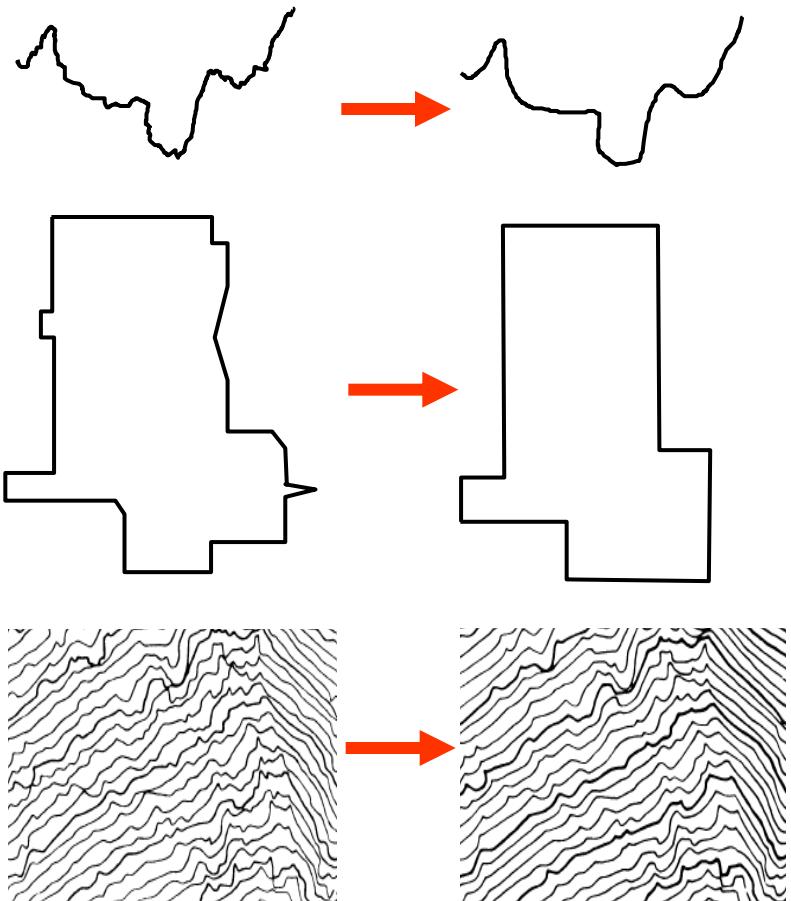
Selection / Elimination

- It works on groups of discrete objects.
- Important objects are preserved, unimportant objects are eliminated.
- It results in a reduced number of objects and distribution density.



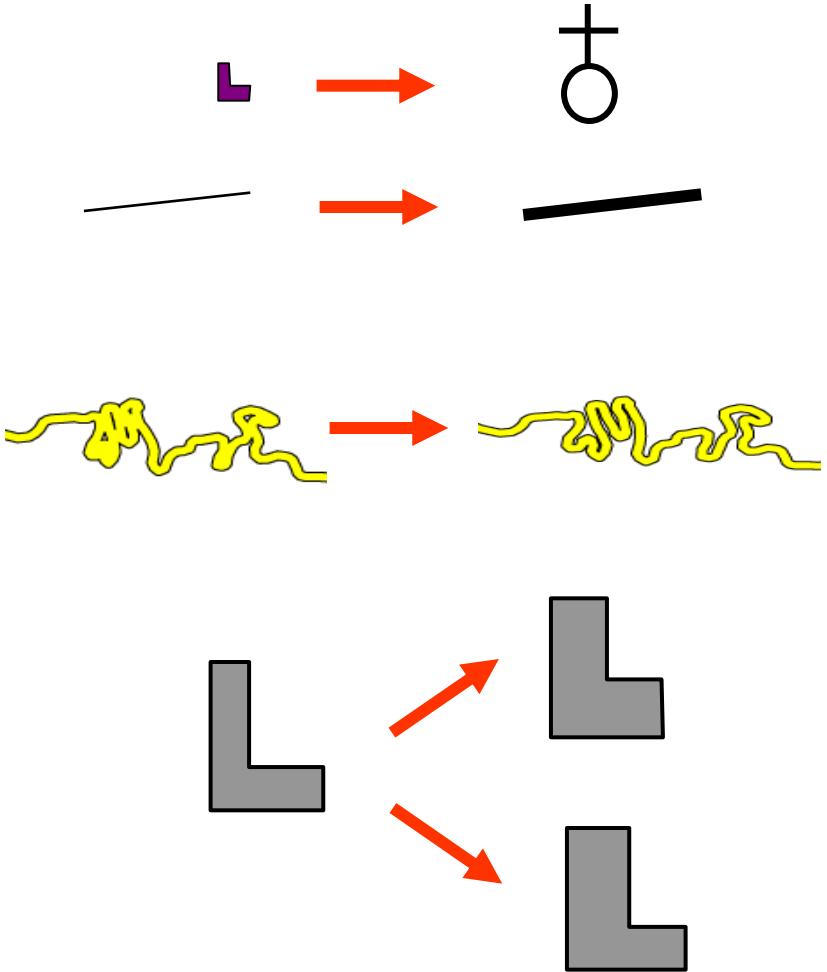
Graphic simplification

- It works on individual line and polygon objects (or object groups).
- Important details are preserved, unimportant details are eliminated.
- It results in a reduced graphic complexity of the concerned object.



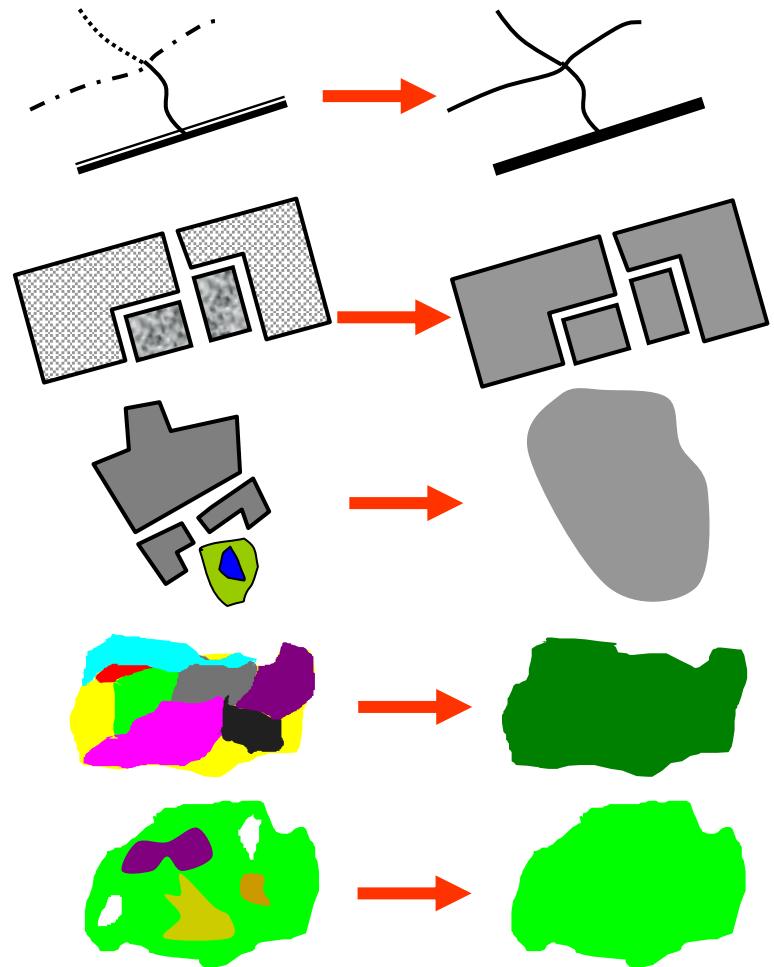
Graphic exaggeration

- It works on any kind of individual objects.
- Important objects (or object parts) are enlarged.
- It leads to an enhanced clarity and legibility of important contents.



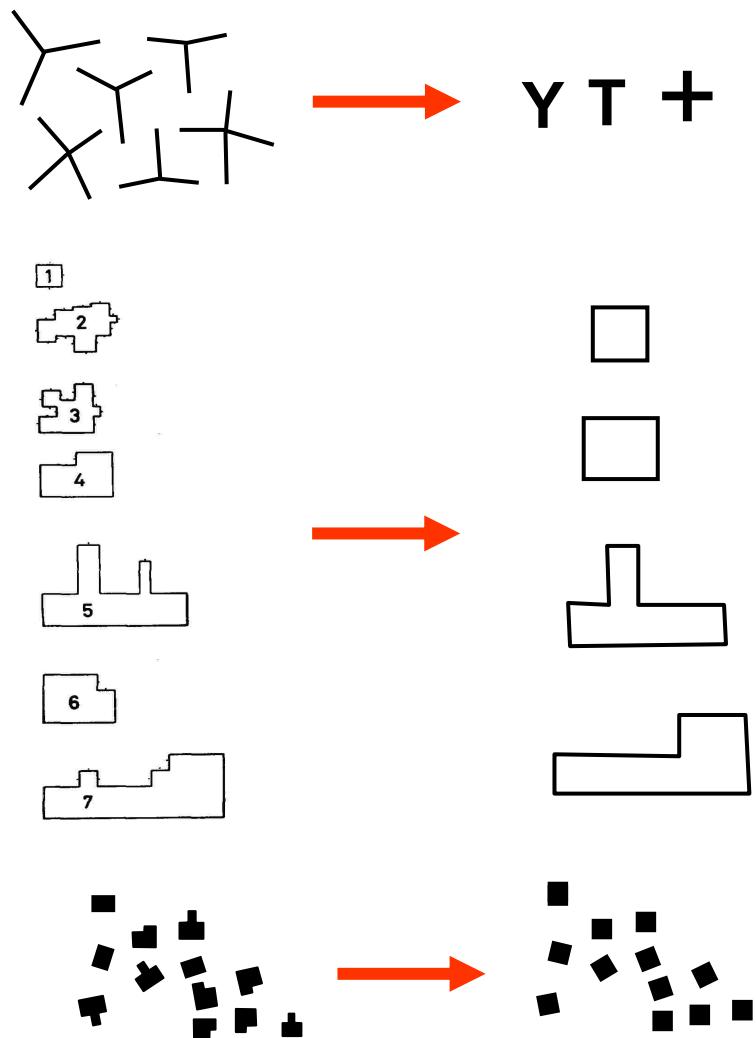
Classification

- It works on discrete objects.
- Common or dominant characteristics of the objects are preserved. Subtle distinctions are ignored.
- It results in reduced semantic variations and a higher level of semantic abstraction.



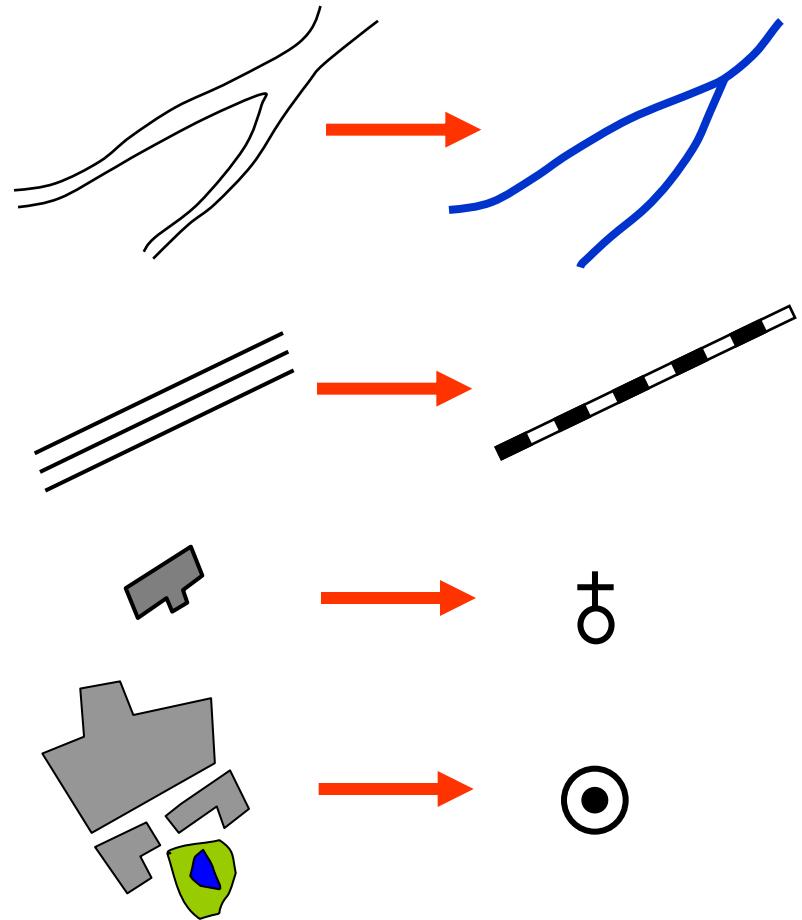
Graphic typification

- It works on discrete objects.
- Various forms are replaced by a few prototypical forms.
- It results in reduced form variations and an improved visual perception of spatial distributions.



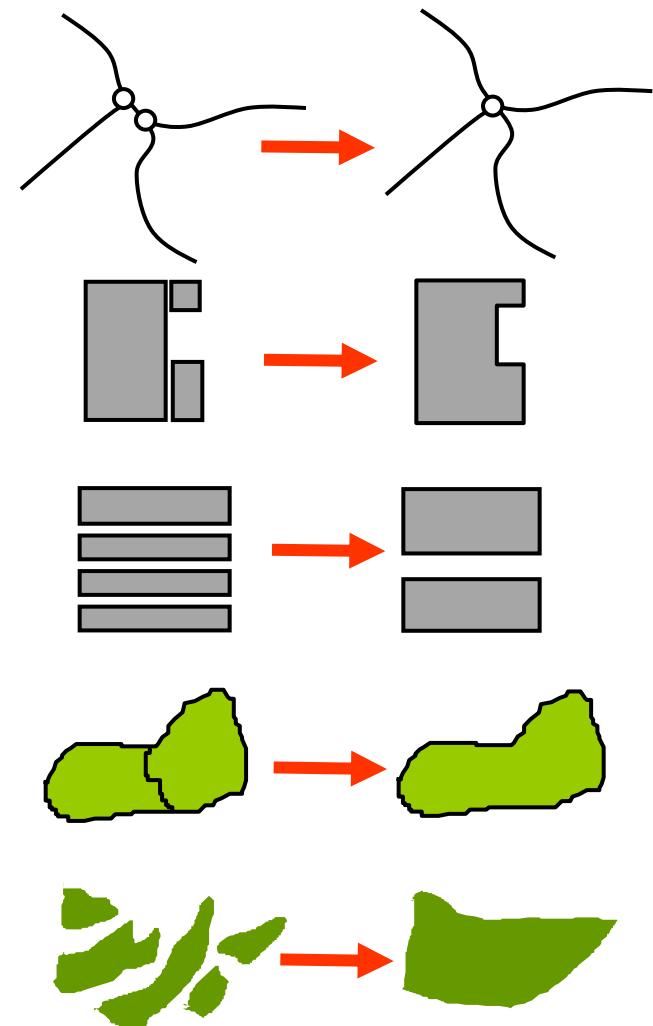
Dimensional collapse

- It works on discrete objects.
- Primary characteristics of the objects are preserved.
Secondary characteristics are ignored.
- It results in an improved clarity and a higher level of semantic abstraction.



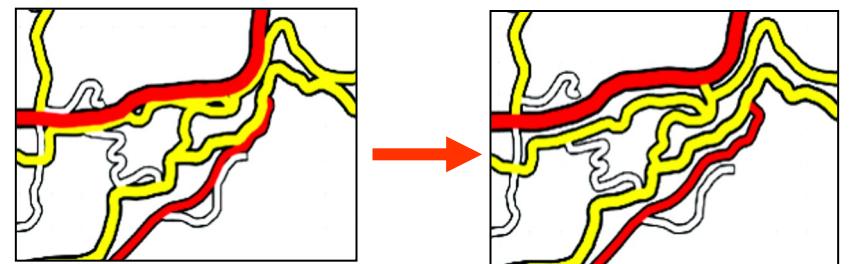
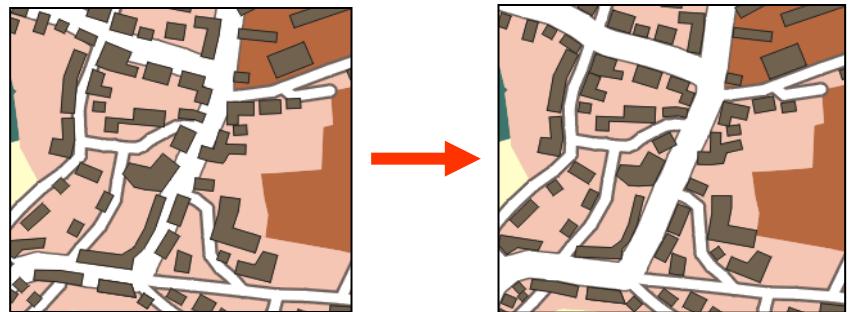
Aggregation / Amalgamation

- It works on proximity objects.
- Common boundaries, small gaps or bridges between proximity objects are eliminated. Universal forms are preserved.
- It results in a reduced number of objects and an improved clarity.



Displacement

- It works on spatially congested objects.
- Important objects remain on their locations. Unimportant objects are moved away.
- It results in an improved clarity and correct map graphics.





© IGN

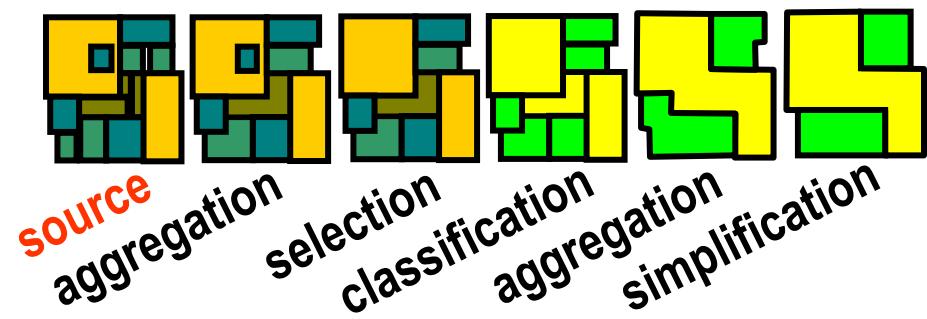
Often many generalization operations have to work iteratively to solve the conflicts

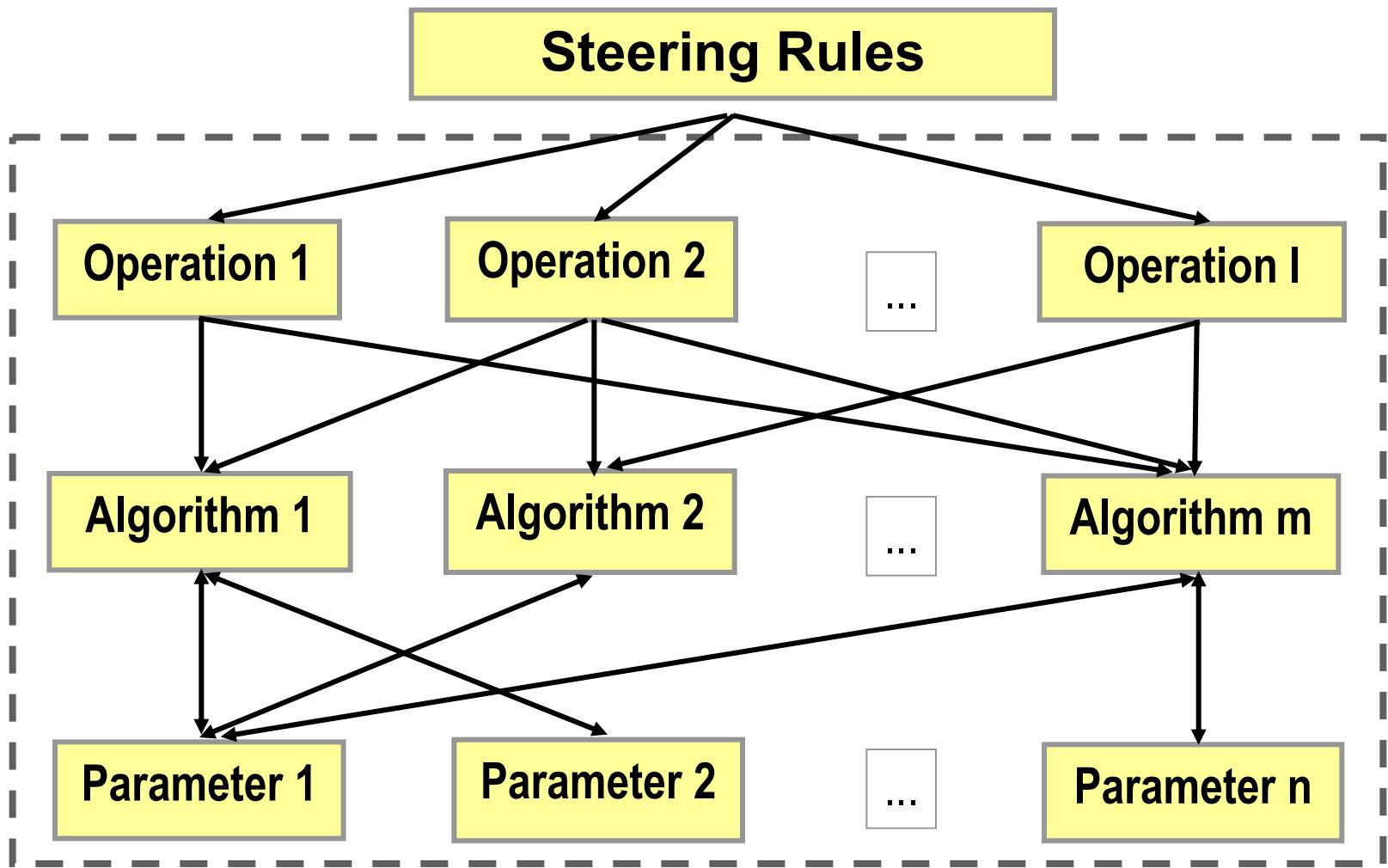
Example of a generalisation workflow

Dividing the generalization space into small chunks and working on chunks

Beginning with the most important feature types and using their generalization results as anchor points for other feature types

Beginning with operations that result in the reduced number of objects but no graphic conflicts

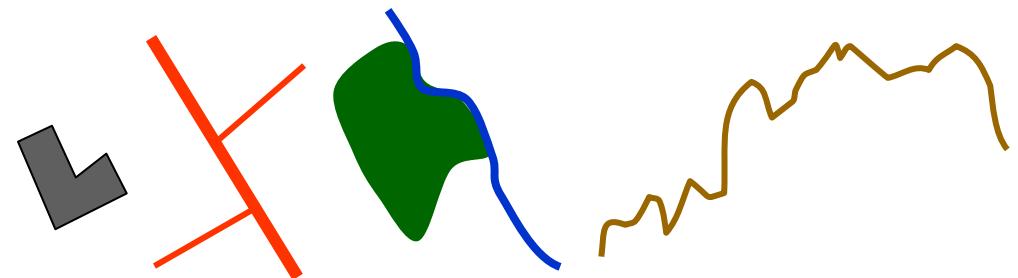




Algorithms – Töpfer's radical law

It decides how many objects (object parts) should be included in the derived map or database according to:

$$n_t = n_s k_i \sqrt{\frac{s_t}{s_s}}$$



n_t the number of objects on the target map

n_s the number of objects on the source map

k_i weight for the relative importance of object type i ($i = 1, 2, 3, \dots$)

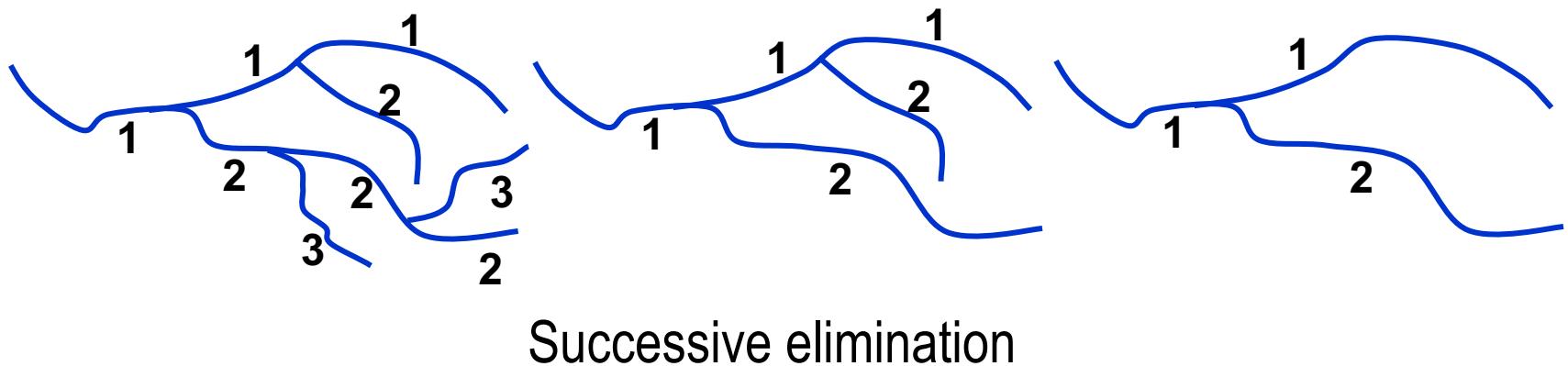
s_t scale of the target map

s_s scale of the source map

Algorithms – Horton's orders

It decides the relative importance of individual streams.

A downstream that intersects an upstream at the smallest angle will be identified as the continuation of the upstream and is assigned the same order class. The smaller the order number, the more important is the stream segment.



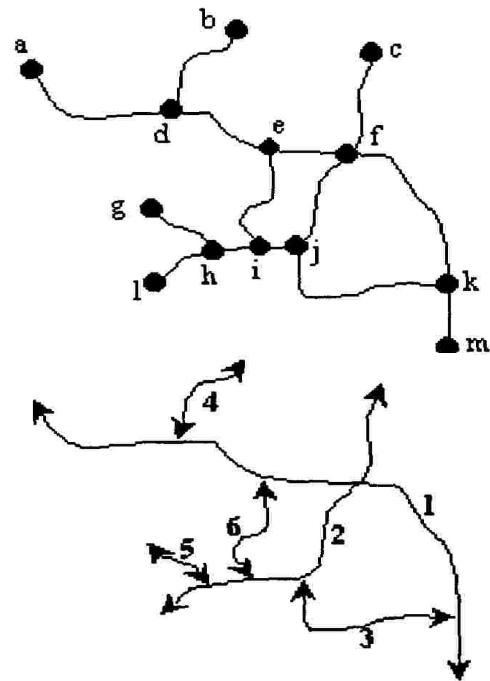
Algorithms – perception-based orders

It decides the relative importance of line objects.

Line objects are partitioned into strokes according to the “good-continuity” principle. The strokes are then ranked based on criteria such as relative length, the number of line segments, or the number of intersections.

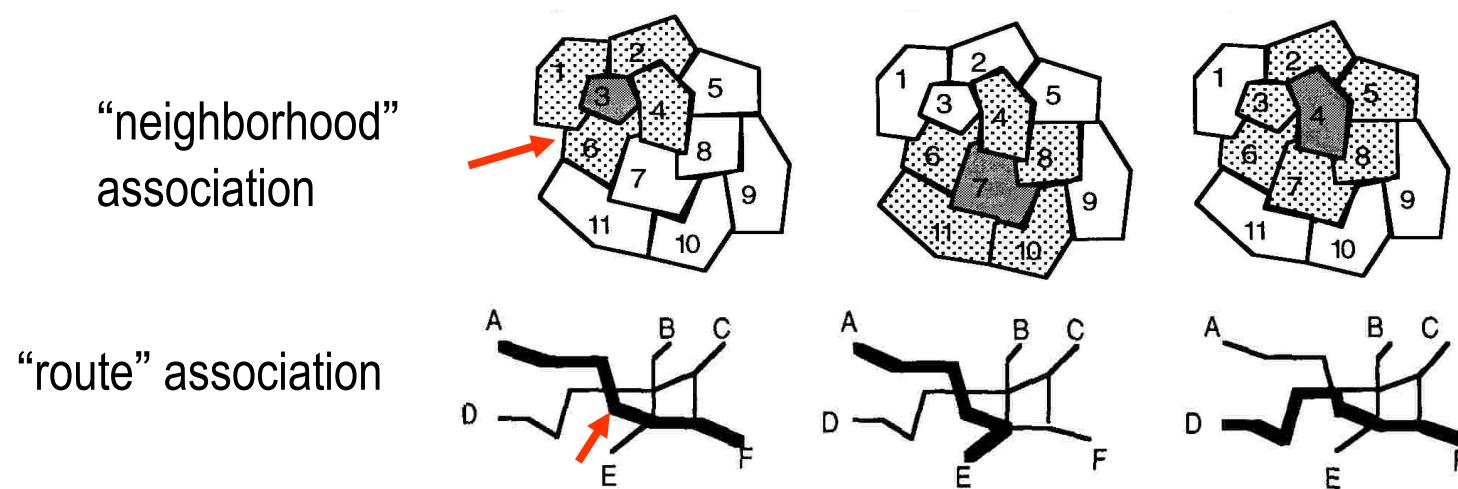
stroke	end nodes		number of line segments
1	a	m	5
2	c	l	5
3	j	k	1
4	b	d	1
5	g	h	1
6	e	i	1

©Thomson & Richardson



Algorithms - set-relationship orders

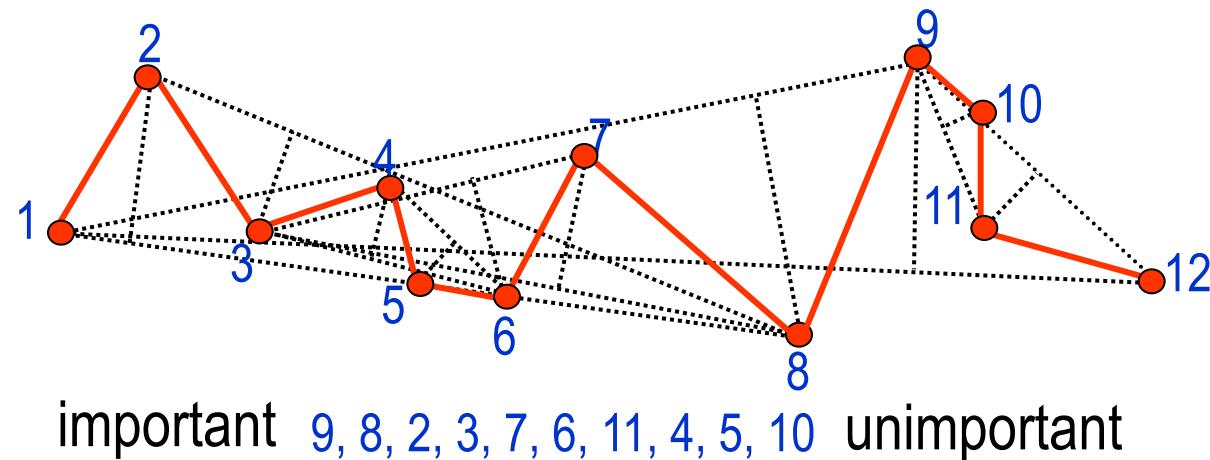
The number of associations an object belongs to reflects its relative importance.



Algorithms - *Douglas-Peucker*

It selects characteristic points on a line, P = tolerance length

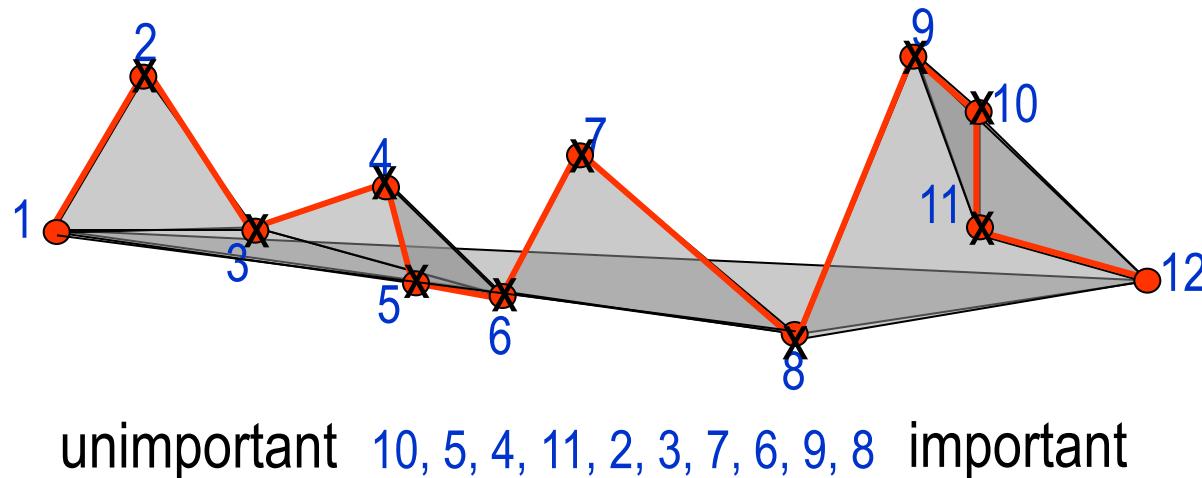
1. define a straight line segment between the terminating points
2. measure the perpendiculars between the line segment and the intermediate points
3. split the line into two parts at the furthest point beyond P
4. repeat 1-3 on the two resulting parts until all perpendiculars fall within the tolerance



<http://129.187.45.33/CartographyPlayground/playgrounds/douglas-peucker-algorithm/>

It eliminates unimportant points on a line, P = tolerance area

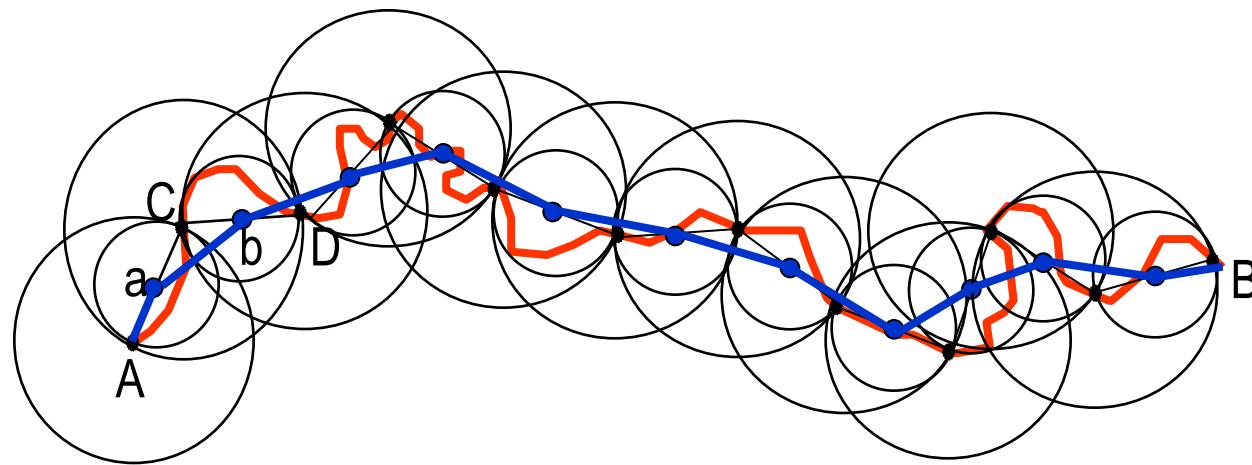
1. measure the relative importance of the intermediate points
2. eliminates the least important point, e.g. the point with the smallest area of the triangular feature ($< P$) formed by connecting the point with its two neighbors
3. repeat 1-2 until all remaining points are beyond the importance threshold



Algorithms - Li-Openshaw

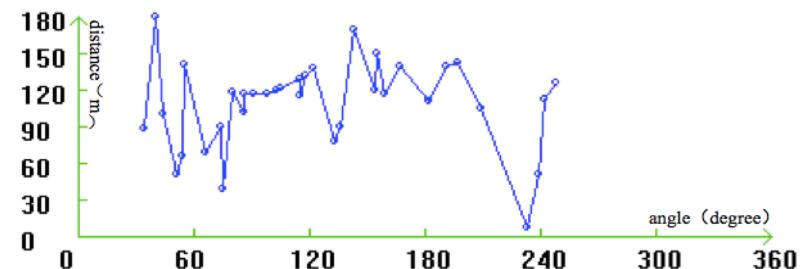
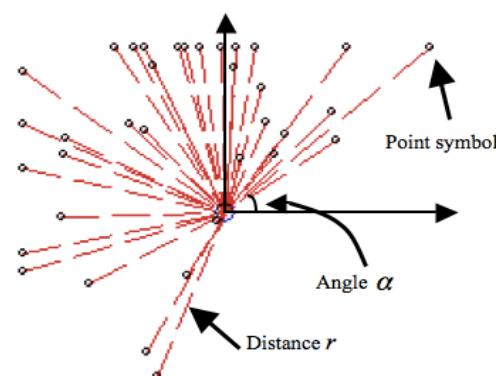
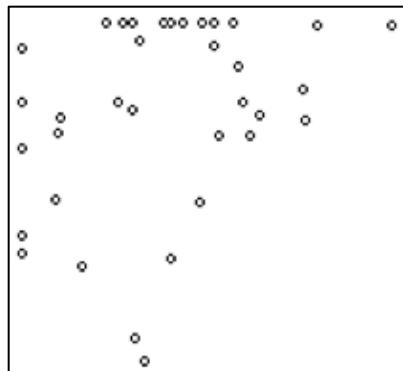
It smoothens a line, P = tolerance length

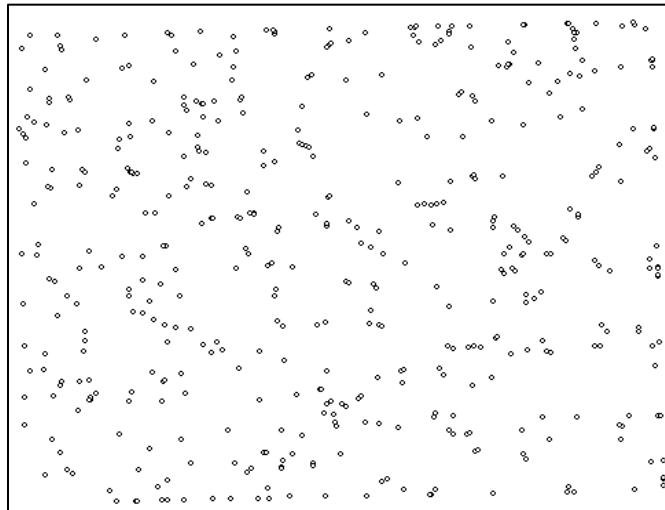
1. draw a circle at starting point A with diameter double so much as P , which intersects the point C on the line
2. draw a circle with AC as diameter. The center ‘a’ along the straight line section AC is selected as a point along the generalized line
3. set C as the new starting point and repeat 1-2



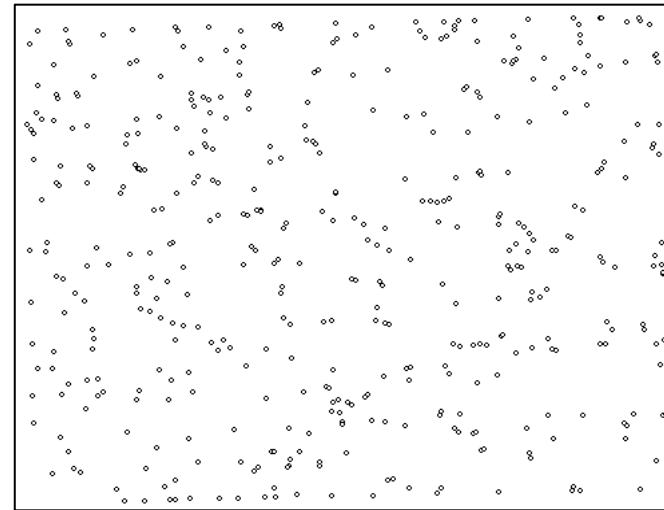
It simplifies a point cluster with a distance tolerance Δr and an angle tolerance $\Delta\alpha$

1. determine the polar center p_c of a cluster, which has the minimum average distance to all other points
2. determine the polar radius r_i and the polar angle α_i of each point p_i , $i = 1, 2, \dots, n-1$ in the cluster to p_c
3. convert the cluster into a polar line ranging from 1° to 360° ,
4. delete the points on the polar line that fall below Δr , $\Delta\alpha$ and convert it back to a simplified cluster

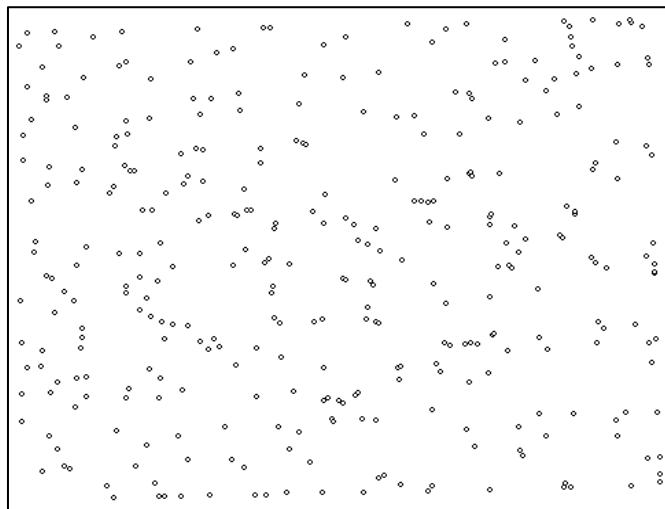




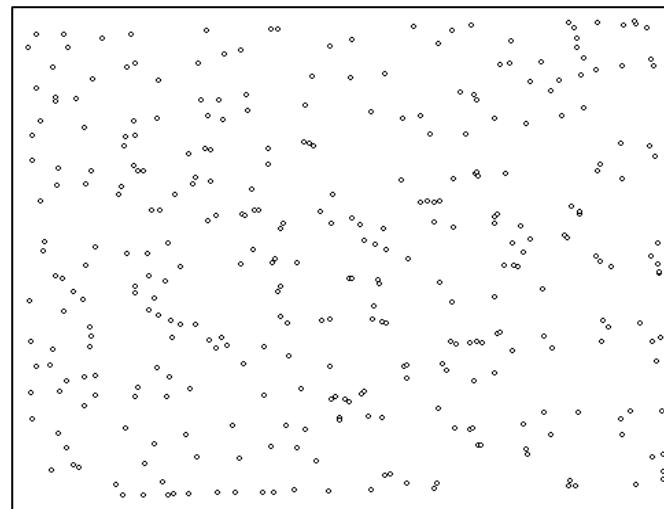
$n = 461$



$\Delta\alpha, \Delta r = (10, 10)$ 437



$\Delta\alpha, \Delta r = (10, 200)$ 332



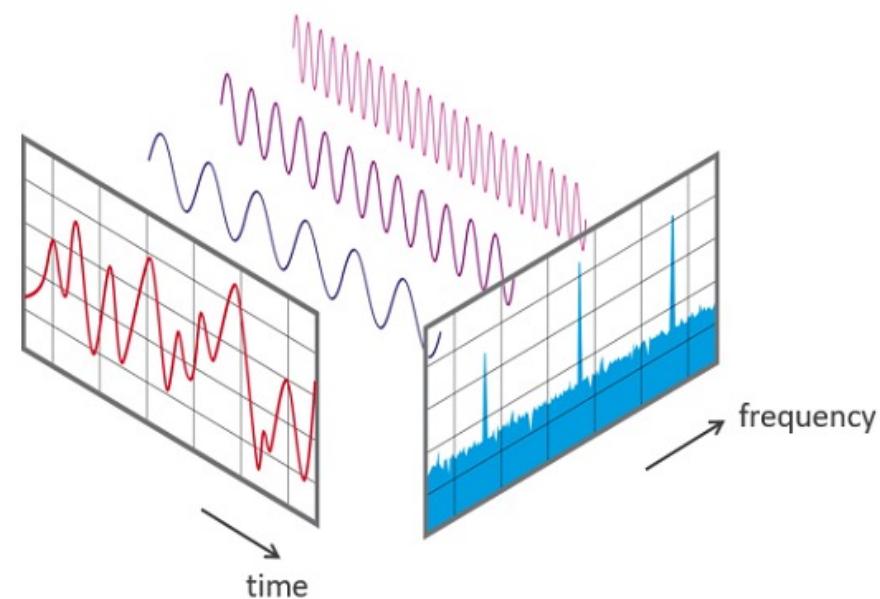
$\Delta\alpha, \Delta r = (20, 100)$ 337

Algorithms – Fourier Transform

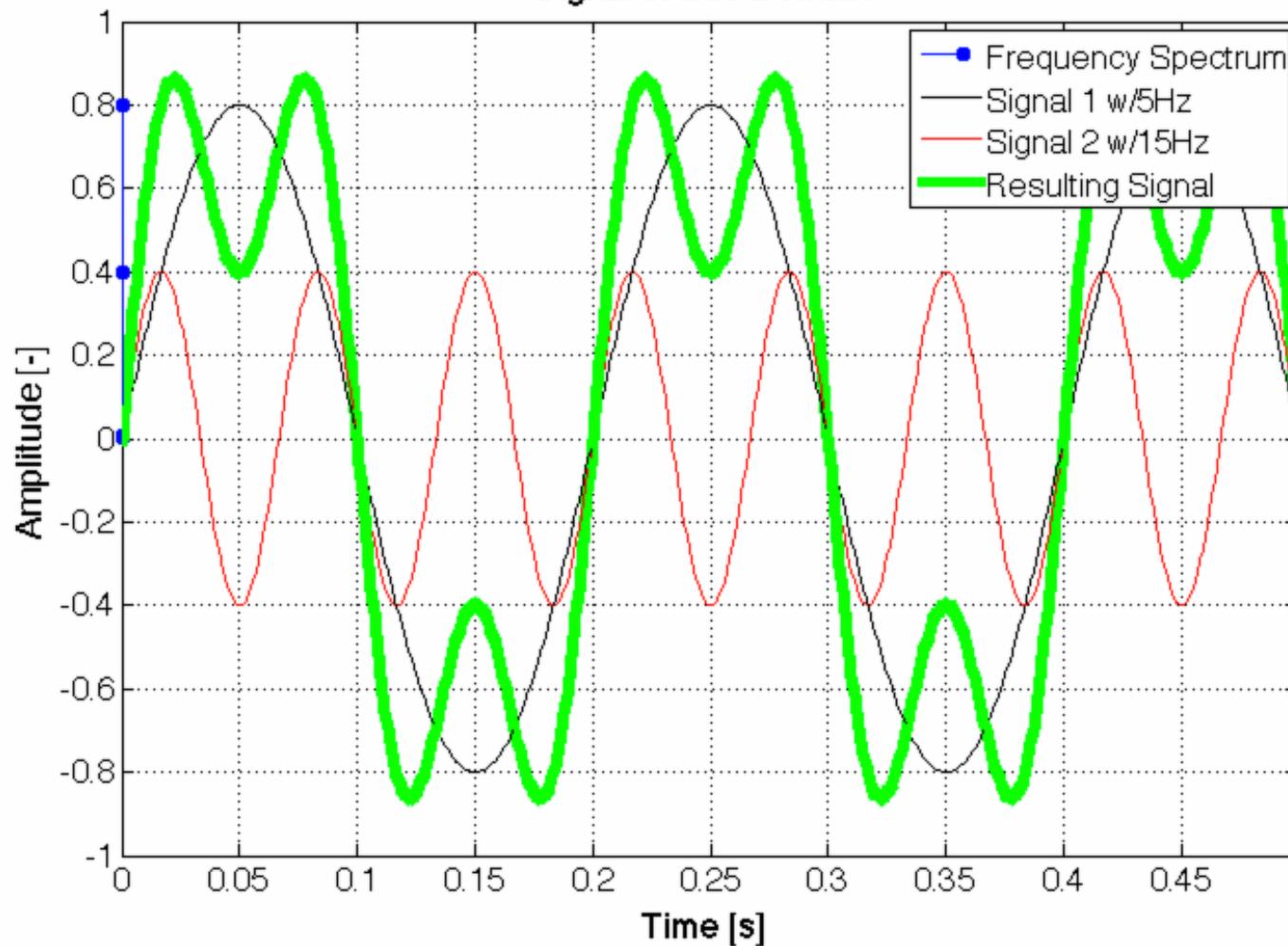
It smoothens a spectrum line globally, $P =$ the number of Fourier coefficient pairs

Any spectrum can be approximated by the summation of a series of sine and cosine waves. The determination of sine and cosine components of a given spectrum is known as FT.

FT converts the time domain to frequency domain. The useful information is found at lower frequencies and the higher frequencies can be ignored.



Signal in time Domain



If a spectrum is recorded at n equally spaced wavelength intervals, $\gamma(1)$, $\gamma(2)$, $\gamma(3), \dots, \gamma(n)$, then it can be approximated by:

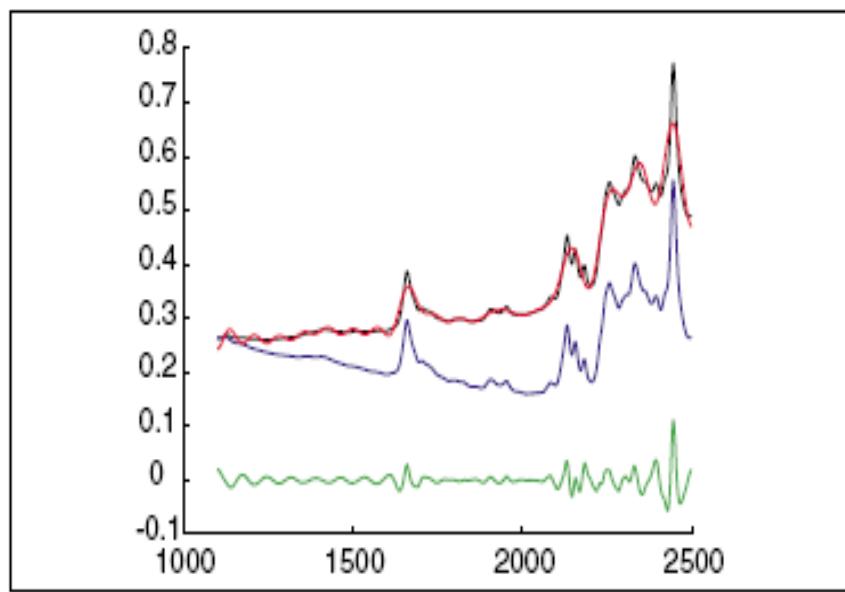
$$\hat{\gamma}(h) = a_0 + \sum_{j=1}^k a_j \cos\left(\frac{2\pi j h}{n}\right) + \sum_{j=1}^{k-1} b_j \sin\left(\frac{2\pi j h}{n}\right)$$

$$a_0 = \frac{1}{n} \sum_{h=1}^n \gamma(h)$$

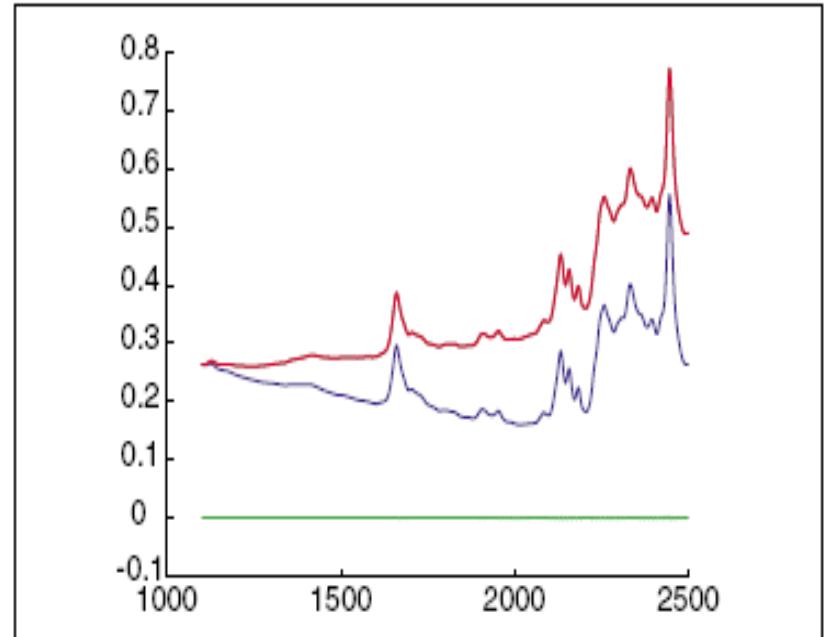
$$a_j = \frac{2}{n} \sum_{h=1}^n \gamma(h) \cos\left(\frac{2\pi j h}{n}\right) \quad \text{for } j = 1, 2, 3, \dots, k-1$$

$$a_k = \frac{1}{n} \sum_{h=1}^n \gamma(h) \cos \pi h$$

$$b_j = \frac{2}{n} \sum_{h=1}^n \gamma(h) \sin\left(\frac{2\pi j h}{n}\right) \quad \text{for } j = 1, 2, 3, \dots, k-1$$



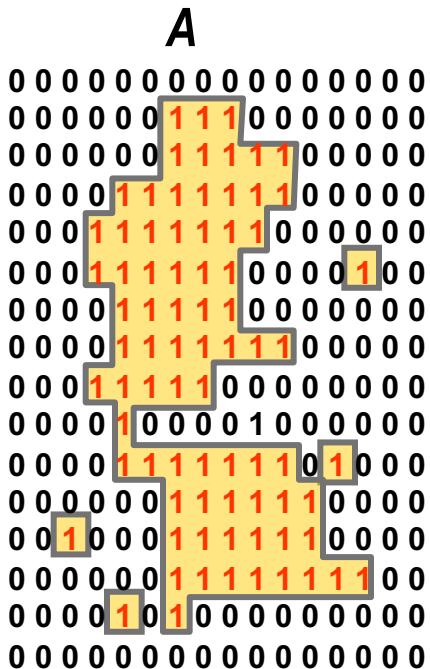
The original curve (black),
reconstructed (red) using 20 pairs of
Fourier coefficients, tilted (blue),
difference (green)



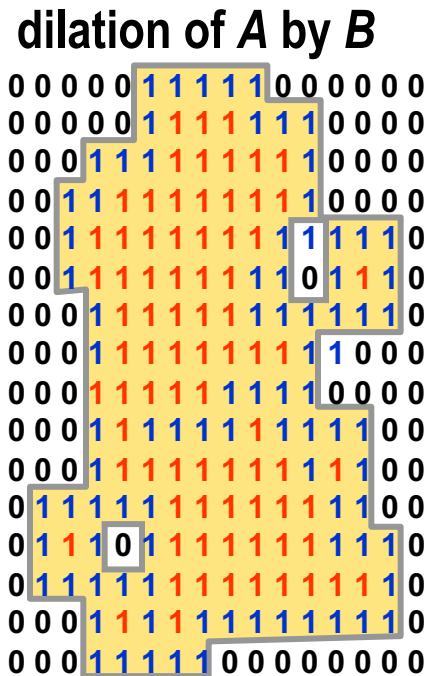
The reconstructed (red) using 100
pairs of Fourier coefficients, tilted
(blue), difference (green)

Algorithms - morphological transformations

Dilation $A \oplus B = \bigcup_{b \in B} A_b$



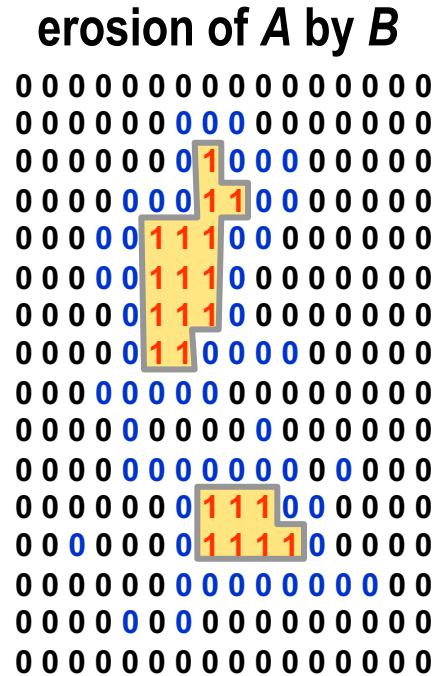
$$B = \begin{matrix} 1 & 1 & 1 \\ 1 & \textcircled{1} & 1 \\ 1 & 1 & 1 \end{matrix}$$



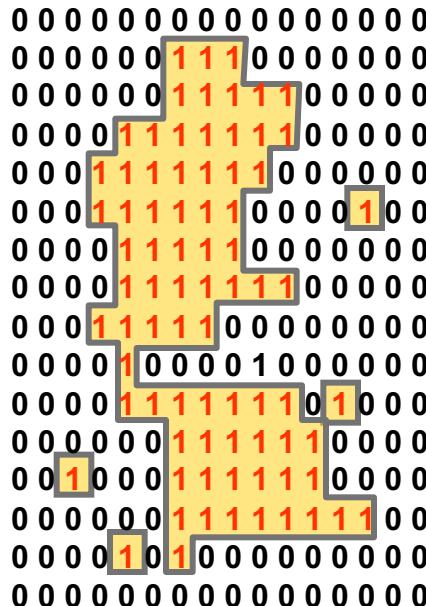
Blue "1" pixels form
the expanded area

Erosion

$$A \ominus B = \bigcap_{b \in B} A_b$$

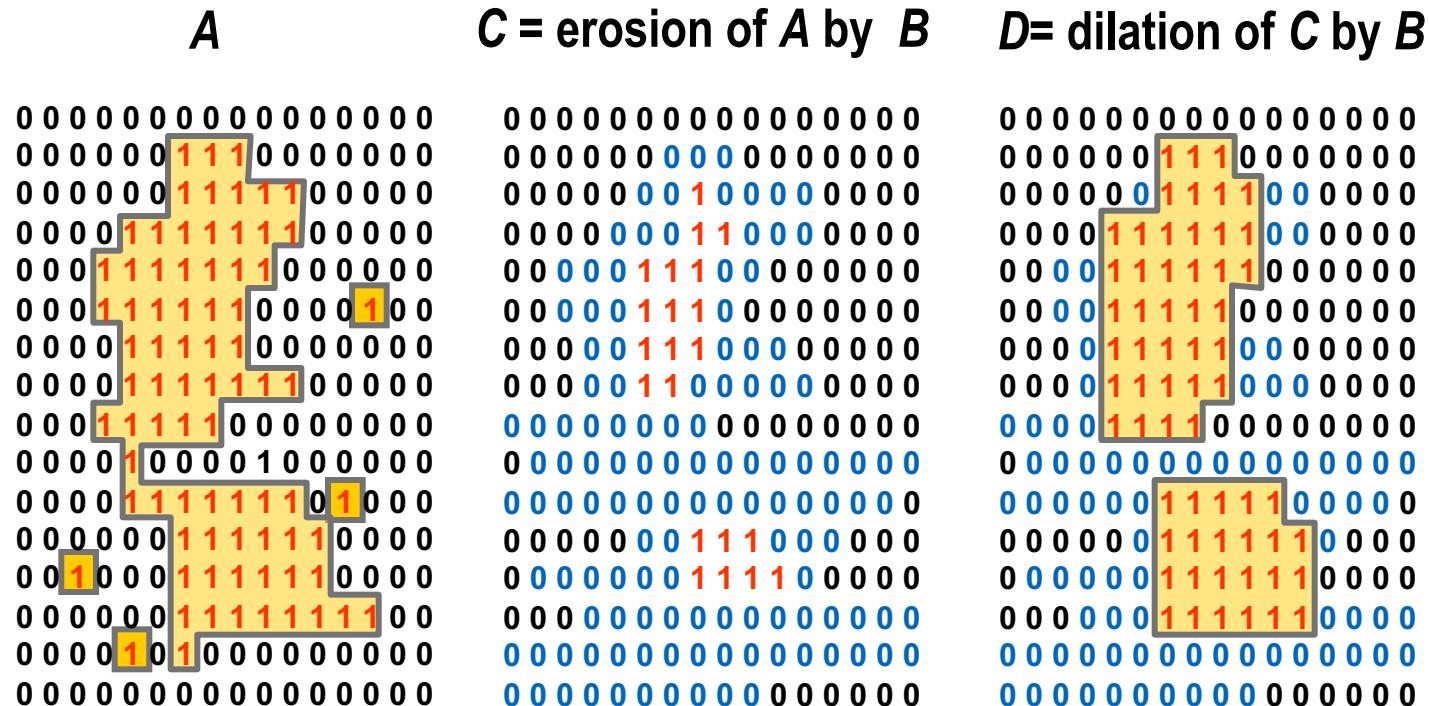


Blue "0" pixels form
the eroded area

A**C = dilation of A by B****D= erosion of C by B**

$$B = \begin{matrix} 1 & 1 & 1 \\ 1 & \textcircled{1} & 1 \\ 1 & 1 & 1 \end{matrix}$$

First dilation, then erosion results in amalgamation, boundary simplification and elimination of small details.

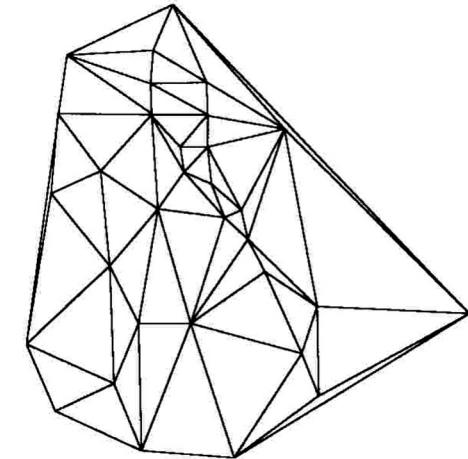


$$B = \begin{matrix} 1 & 1 & 1 \\ 1 & \textcircled{1} & 1 \\ 1 & 1 & 1 \end{matrix}$$

First erosion, then dilation results in boundary simplification, elimination of small details and bridges.

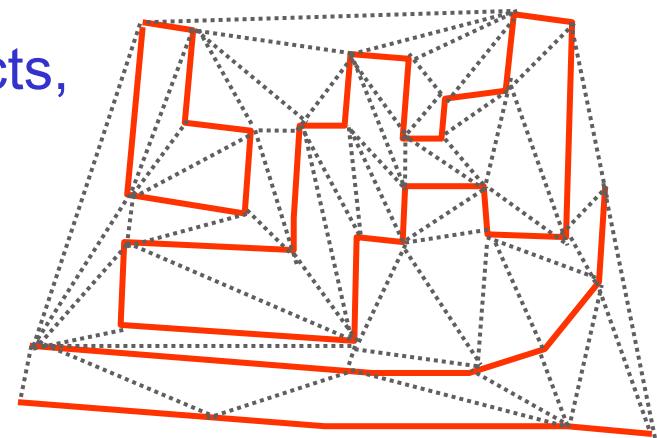
Algorithms - Graph-based data structures

Delaunay Triangulation (DT)



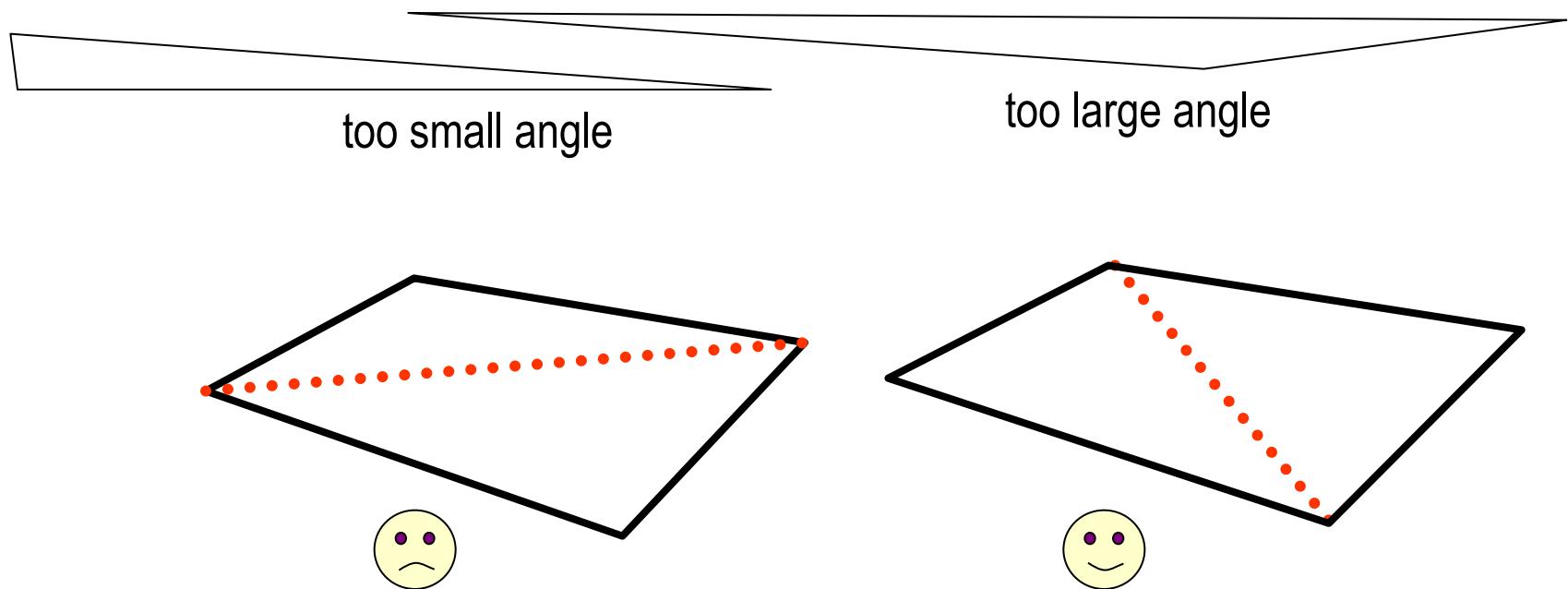
Constraint Delaunay Triangulation (CDT)

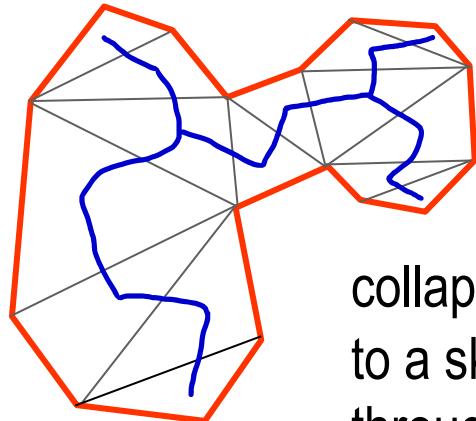
It is a modification of a Delaunay triangulation, in which edges, belonging to polygon and line objects, are enforced as edges within the triangulation.



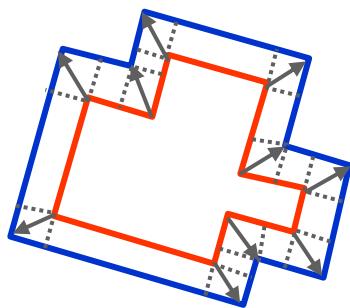
Quality criteria for Delaunay triangulation

- The edges of triangles should have similar length.
- The angle difference of each triangle should be kept as small as possible.
- The area difference between the neighboring triangles should be kept as small as possible.

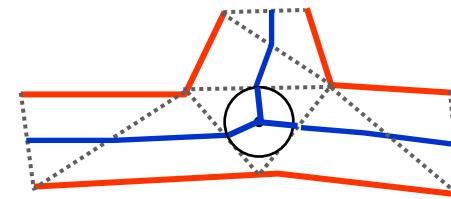




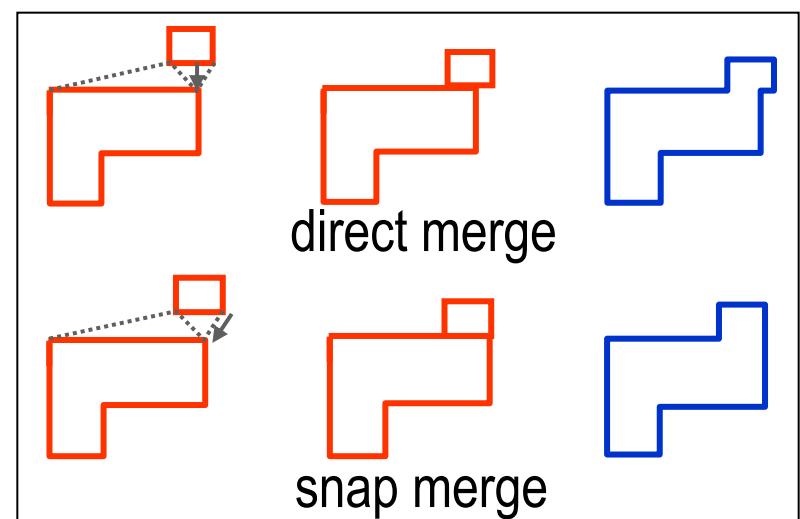
collapse of a polygon
to a skeleton, e.g.
through centers of
triangle circumcircles

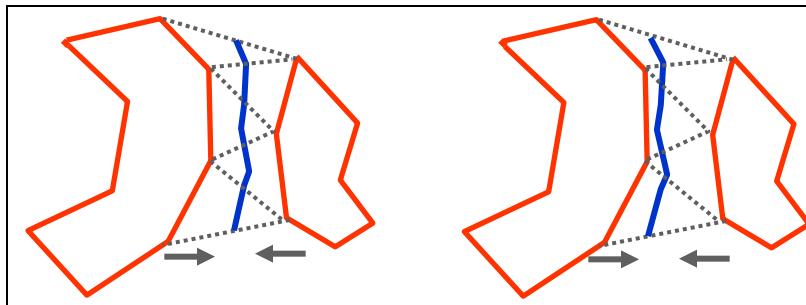


graphic exaggeration of a concave
polygon based on displacement
vector of triangle nodes

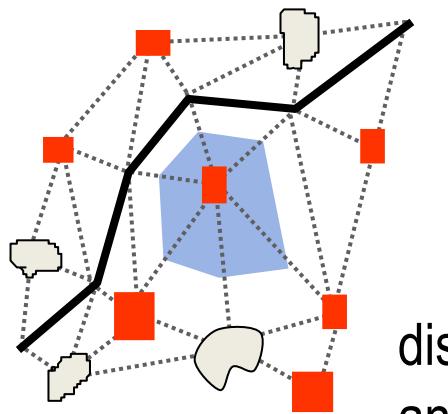
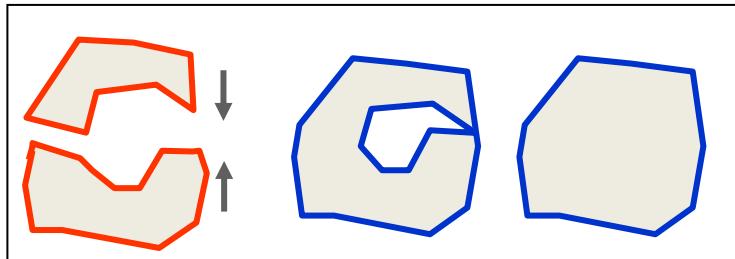


collapse of a ribbon-shaped object
to a skeleton, e.g. through the
centers of triangle edges
connecting opposite sides

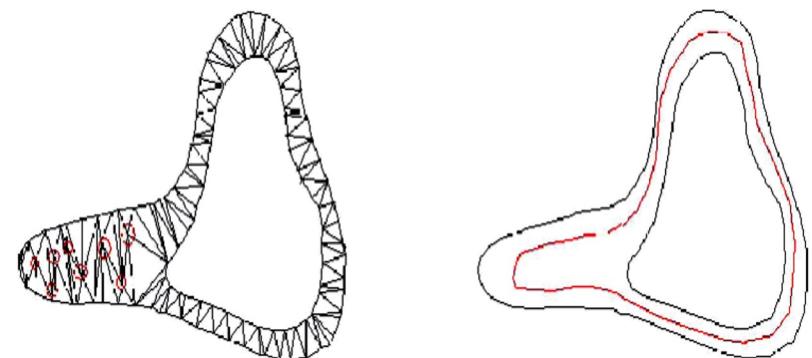




middle axis and weighted
middle axis as merging
boundary



displacement buffer around
an arbitrary object



Derivation of a new contour line
based on CDT of existing contour
lines and terrain points

Questions

1. Which differences can two data sources of the same region reveal?
2. Why do we need to integrate different data sources?
3. Describe the general workflow of geodata integration.
4. How can you define the similarity between objects to be matched?
5. How can you determine the tolerance threshold for similar objects?
6. Give some examples for the applications of data integration.
7. Why do we have to generalize maps?
8. Which automatic generalization operations do you know and how do they work?
9. What is the Töpfer's radical law?
10. Which methods of ranking the relative importance of individual objects do you know and how do they work?
11. Which automatic algorithms of line simplification do you know and how do they work?
12. Explain the working principle of Fourier Transform for line generalization.
13. How can the morphological transformations be used as generalization means?
14. How can the graph-based data structures be used to generalize maps?