



Institute of Cartography and Geoinformatics | Leibniz Universität Hannover

Automatic Geo-Referencing of Provisional Cadastral Maps: Towards a Survey-Accurate Cadastral Database for a National Spatial Data Infrastructure

David Siriba and Sagi Dalyot

{david.Siriba, sagi.dalyot}@ikg.uni-hannover.de



Introduction

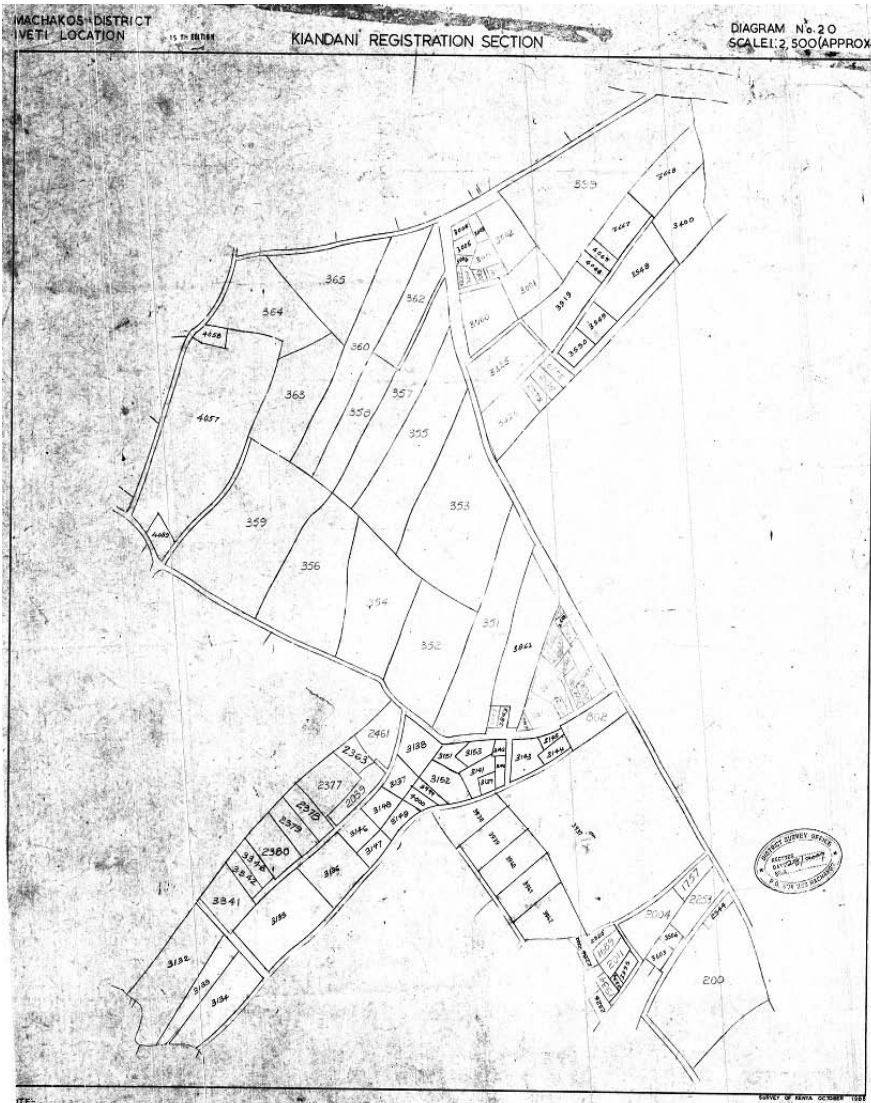
- ▶ Urbanization and urban sprawl rapidly increase in developing countries.
- ▶ Demand for more land for development leads to a better spatial planning.
- ▶ Cadastral maps in many countries are not of uniform geometric quality.
- ▶ Typically, cadastres in the urban areas are survey-accurate, whilst cadastres in the rural neighborhoods are not.



Motivation

- ▶ Survey-accurate cadastral maps are considered as core components of Spatial Data Infrastructures (SDI).
- ▶ Replacing existing provisional (temporary) cadastral maps is lengthy and costly.
- ▶ There exist a widespread more accurate topographic and image datasets that can be used to enhance the geometric quality of the temporary cadastral maps.
- ▶ Introducing an automatic approach for geometric registration (geo-referencing) as a first stage in the quality enhancement.
- ▶ Replacing the need for preliminary manual geo-referencing, thus avoiding ambiguities and user-errors.

Problem Definition



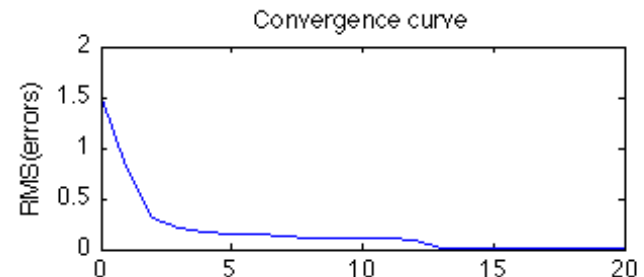
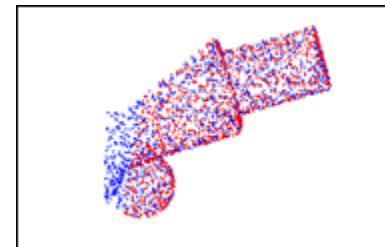
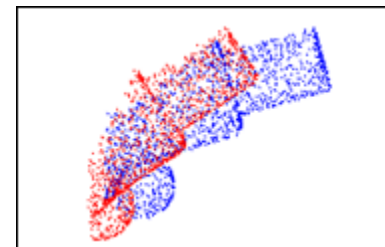
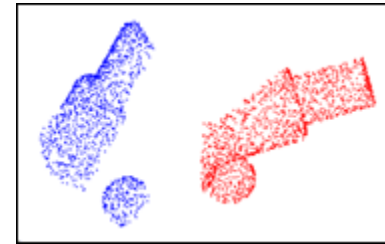
- ▶ Requirement – cadastral maps to be geo-referenced to the national coordinate system and serve as reliable SDIs.
- ▶ Temporary cadastral maps that have no coordinate grid and have distortions.
- ▶ Introducing a qualitative automatic geometric registration of temporary cadastral maps.

Contents

- ▶ Geometric registration (short review)
- ▶ Ranked Hausdorff distance approach
- ▶ Experimental results and analysis
- ▶ Discussion and conclusions

Geometric Registration

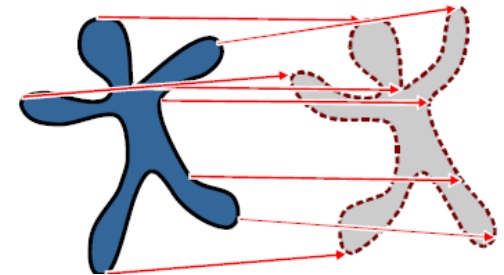
- ▶ Geometric registration entails two complementary problems:
 - determining mutual correspondences (based on features or object similarities)
 - establishing the appropriate geometric transformation (derived from the geometry at hand, i.e., rigid or non-rigid)
- ▶ Simultaneous solution of the correspondence and transformation problem is usually done in an alternating manner until convergence (threshold scheme).



Geometric Registration

▶ Common methods/algorithms:

- Iterative Closet Point (ICP) algorithm:
 - Coupling pairs of counterpart points, considered the nearest ones; estimating transformation that aligns best both datasets.
 - Algorithm breaks down quickly in the presence of local distortions and absence of preliminary mutual correspondence; one-to-one correspondences are not guaranteed.
- Thin Plate Splines (TPS) Rapid Point Matching (RPM) and the Coherent Point Drift (CPD) algorithms:
 - Feature-based non-rigid registration method.
 - The algorithms only work well for point sets that form a solid figure with a contour like outline.

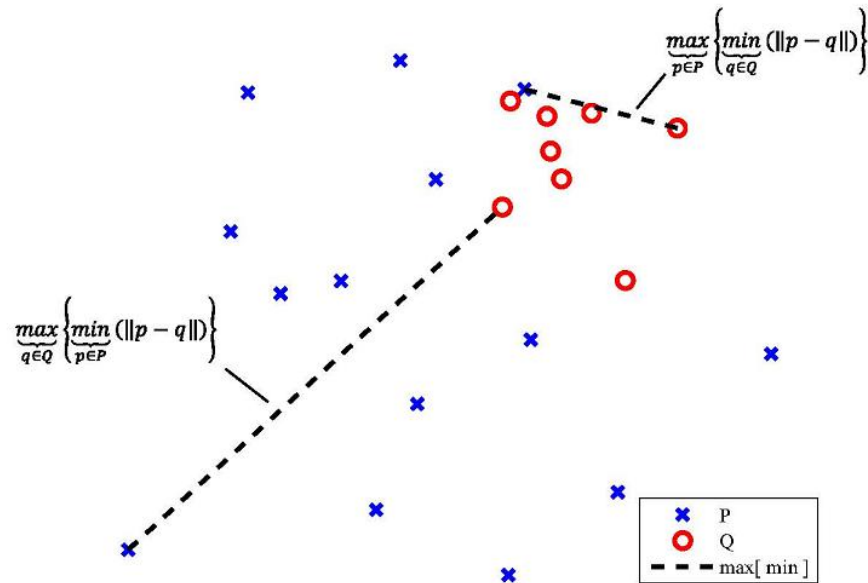


Geometric Registration

- ▶ Main aspects that should be considered during geometric registration:
 - **Data coverage** – the spatial extent of the datasets may not be the same (outliers).
 - **Time-frame** – both datasets may have been produced at different times (outliers, distortions).
 - **Accuracy** – the subject dataset may contain errors and noise.
- ▶ Research solution:
 - Ranked Hausdorff Distance approach:
 - An automatic geo-referencing approach that considers the above challenges.

Ranked Hausdorff Distance Approach

- ▶ The forward Hausdorff distance H measures the extent to which each point in the subject dataset $P \{p_1, p_2, \dots, p_n\}$ lies "near" other points in the reference dataset $Q \{q_1, q_2, \dots, q_m\}$.



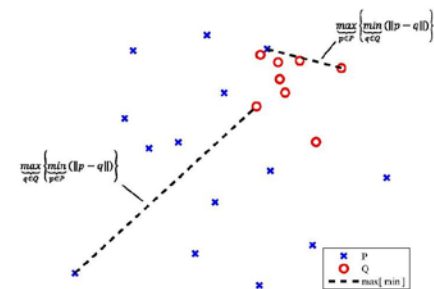
$$H(P, Q) = \max \left(\max_{p \in P} \left\{ \min_{q \in Q} \|p - q\| \right\}, \max_{q \in Q} \left\{ \min_{p \in P} \|p - q\| \right\} \right)$$

Ranked Hausdorff Distance Approach

- ▶ Approximating the degree of similarity (resemblance) - or minimum mismatch - between two datasets.
- ▶ Consists of qualitative measures that evaluate the correspondence of two datasets in metric space by assigning distance measure.

$$H(P, Q) = \max \left(\max_{p \in P} \{ \min_{q \in Q} \| p - q \| \}, \max_{q \in Q} \{ \min_{p \in P} \| p - q \| \} \right)$$

- ▶ The expression in $\| \|$ defines a norm on the points existing in A and B , such as distance (or displacement), consisting of translation(s), rotation(s), and scaling factor(s).

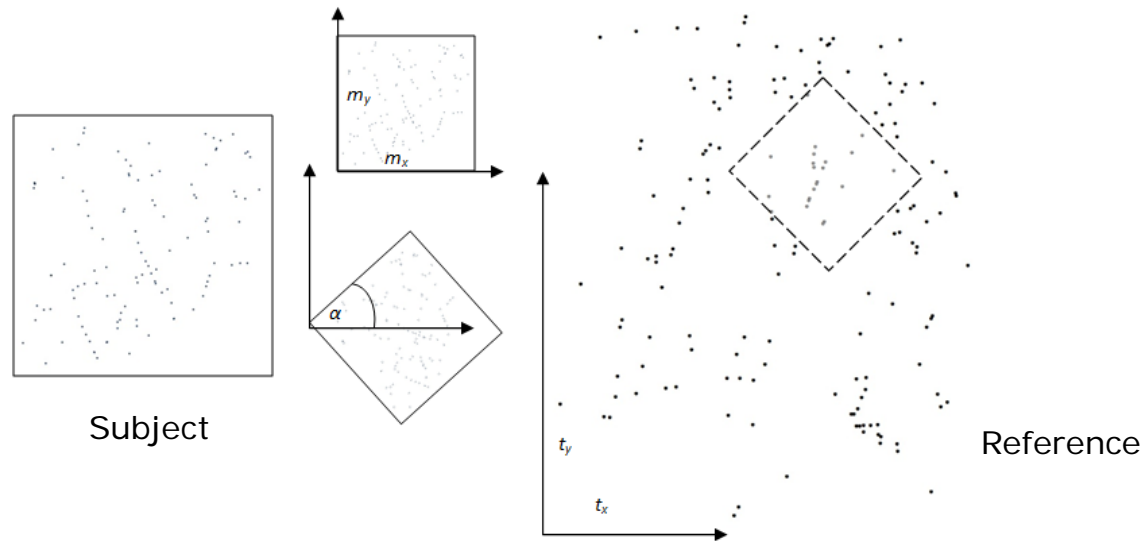


Ranked Hausdorff Distance Approach

- ▶ Datasets do not exist on the same metric space – pixel vs. metric, thus finding a specific rigid transformation $T_s(P)$.
- ▶ A minimization problem is formulated that searches for the transformation parameters.

$$f_s = \min_{s \in S} H(T_s(P), Q)$$

- ▶ $s \in S$ encapsulates transformation parameters: translation (t_x, t_y) , rotation (α) , and, scaling (m_x, m_y) .



Ranked Hausdorff Distance Approach

- ▶ Preliminary knowledge (might suggest using a series of sets):
 - **Scaling** – DPI derive rough scaling of the map itself (m_x, m_y)
 - **Rotation** – relation between the true north and map northing (a)
- ▶ Iterative voting stage until certain criteria are satisfied and validated (converging to global minimum) by ranking:
 - Standard deviation of Euclidean distance in x and y for all corresponding points identified (Q_i vs. $T_s(P_j)$) - relative correspondence.
 - Number of correspondences (points).
 - Euclidean distance threshold is derived by datasets accuracies.

Ranked Hausdorff Distance Approach

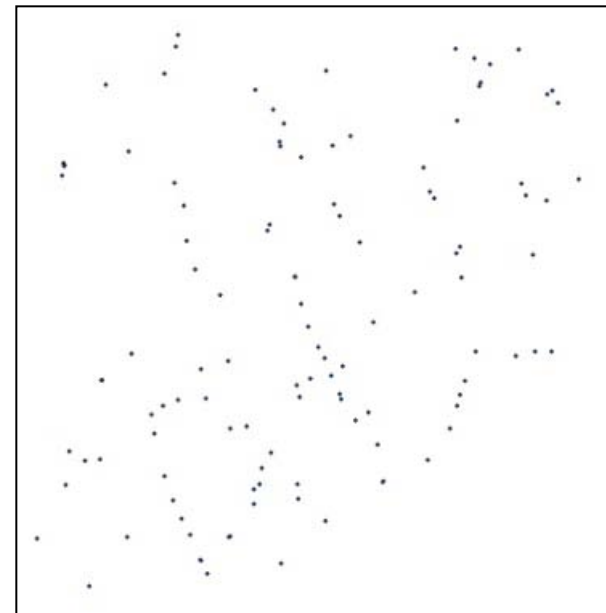
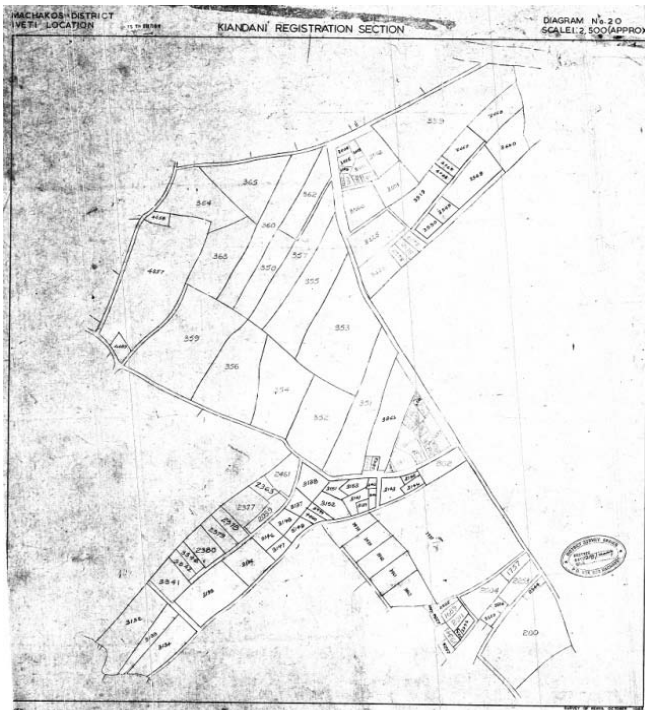
- ▶ A voting mechanism is introduced based on partial distances voting criteria relying on ranking.
- ▶ For l of the L model points ($1 \leq l \leq L$) is given by taking the K^{th} ranked point of Q – rather than the largest ranked one.

$$h_L(Q, P) = K^{th}_{q \in Q} \min_{p \in P} \|p - q\|$$

- ▶ K^{th} denotes the ranked value in the set of existing distances - partial distances.
- ▶ Ranking is achieved by the perspective values of this distance → “best matching” by identifying the subset of the model that minimizes the direct Hausdorff distance.
- ▶ Ranking enables to identify the corresponding subset that minimizes the directed Hausdorff distance.
- ▶ The subject dataset is “moved” iteratively toward its “position” in the reference dataset.

Experimental Results and Analysis

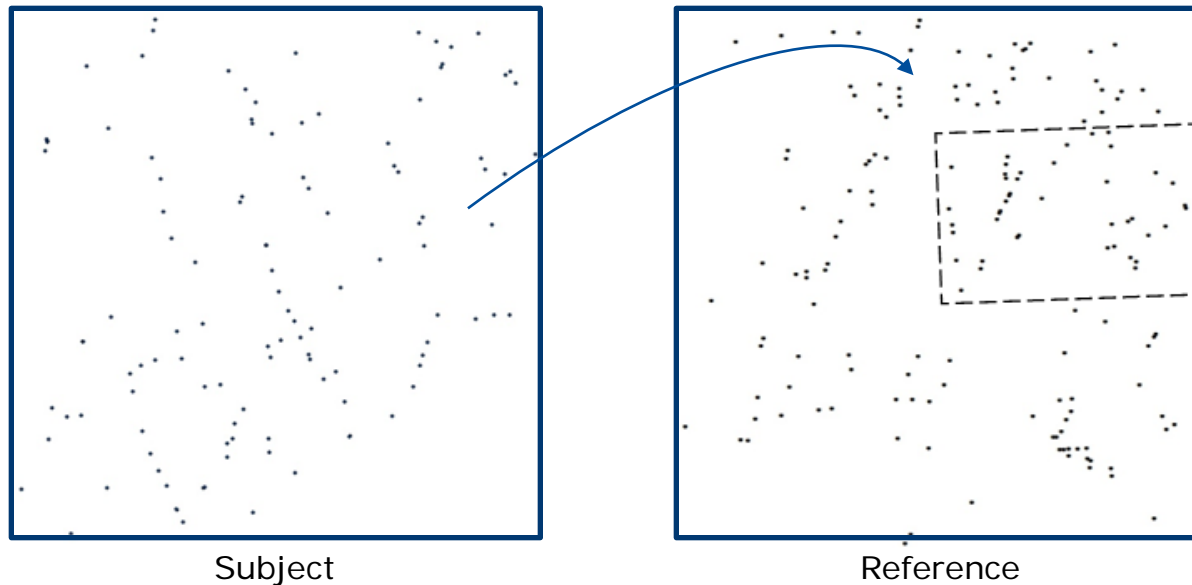
- ▶ Provisional cadastral maps are used in Kenya in the rural areas in the immediate neighborhood of urban areas.
- ▶ Maps were originally created by tracing the approximated parcel boundaries from un-rectified aerial photographs.
- ▶ Road junctions are extracted automatically – subject dataset.



Subject

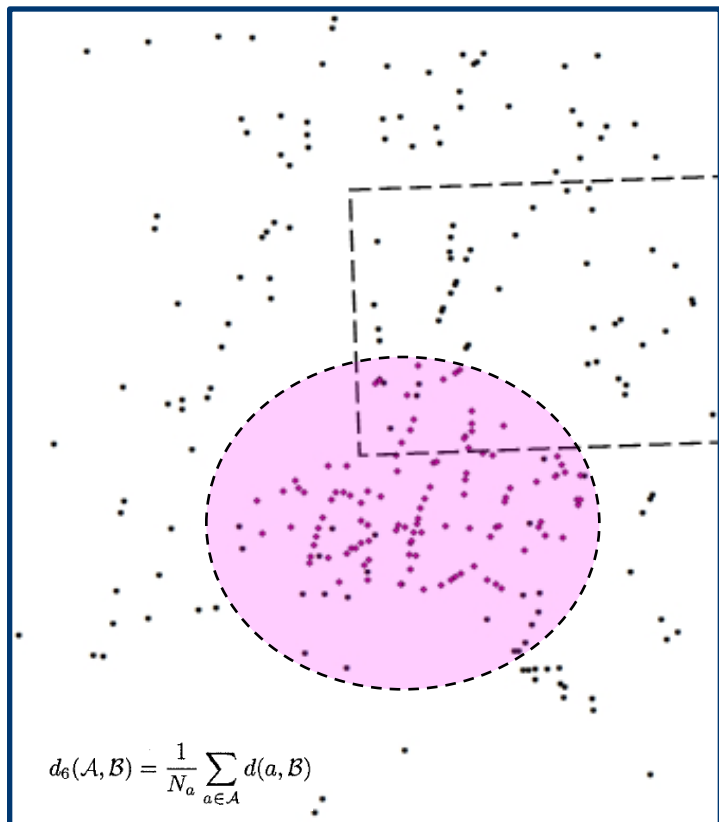
Experimental Results and Analysis

- ▶ Reference dataset used - an accurate topographic road-network dataset.
- ▶ Both datasets consist of different datums: pixel-space vs. metric geo-referenced (in UTM zone 37S coordinate system), different scaling both in x and y directions, as well as rotation.
- ▶ Points from cadastral dataset are a subset of the reference topographic dataset.

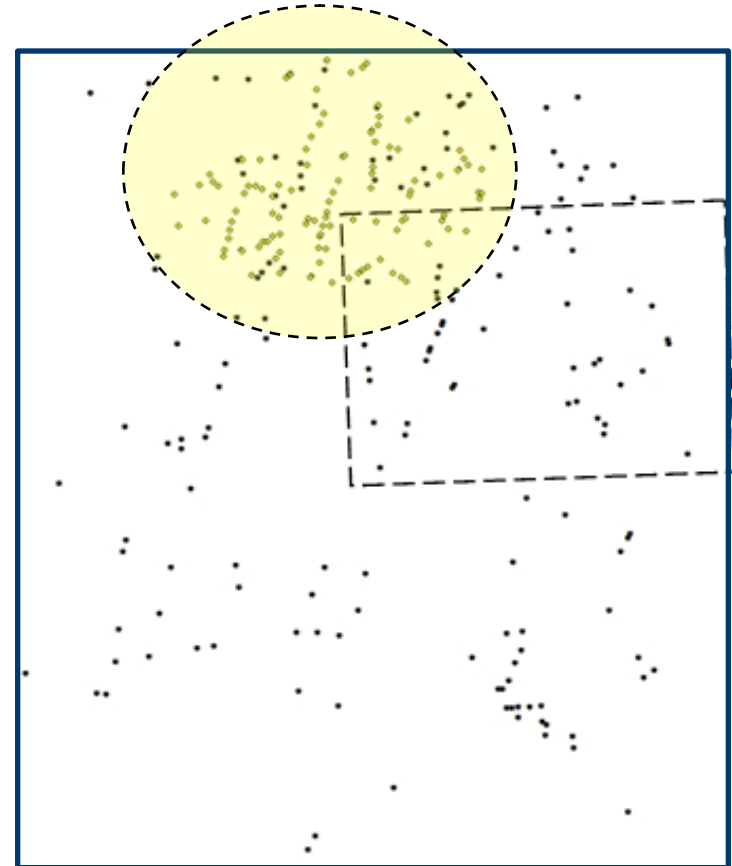


Experimental Results and Analysis

- ▶ Modified Hausdorff Distance (left) and the Forward Hausdorff distance (right): a displacement of more than 1,200 m in position.

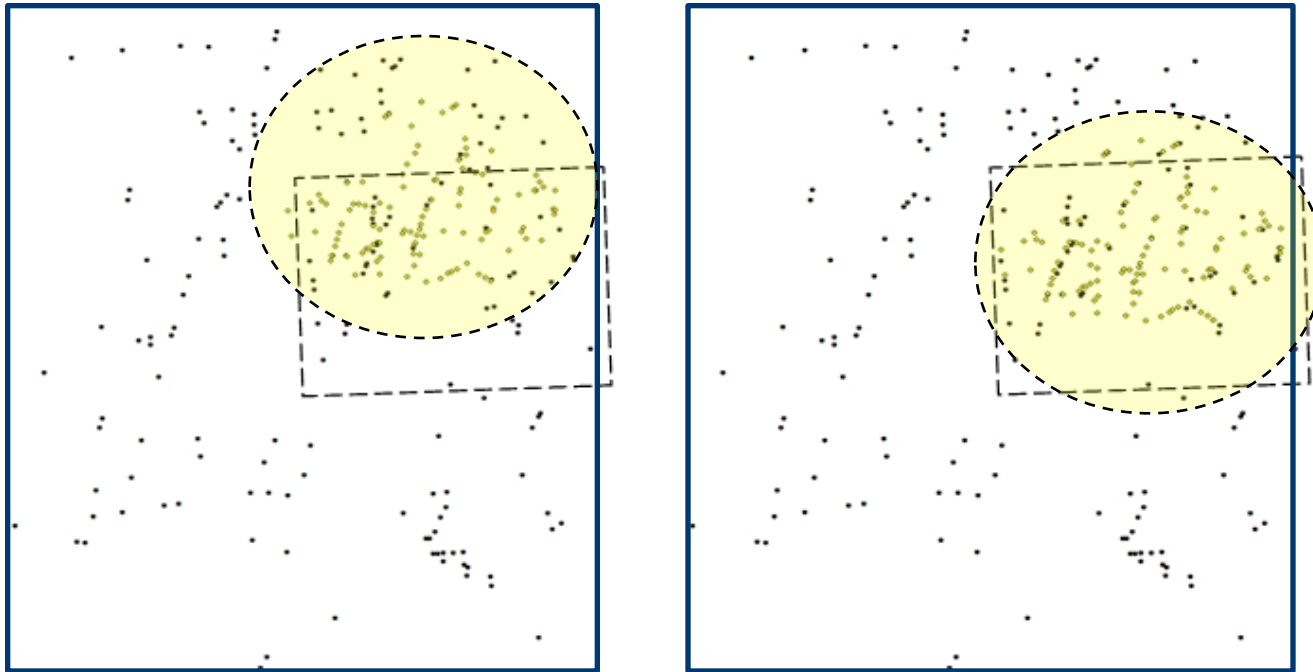


Dubuisson and Jain, 1994



Experimental Results and Analysis

- ▶ Proposed algorithm after the first (left) and second (right) iteration.
- ▶ After the first iteration the positioning is better than both previous processes - a displacement of less than 300 m.
- ▶ After the second iteration the displacement in position is less 30 m (in most areas).



Experimental Results and Analysis

- ▶ A third-party independent manual registration, i.e., geo-referencing, was carried out on the same datasets.
- ▶ Total number of identified corresponding points:
 - Manual: 14.
 - Automatic algorithm: 24.

$$[X \ Y] = [x \ y \ 1] \cdot \begin{bmatrix} m_x c & -m_y s \\ m_x s & m_y c \\ t_x & t_y \end{bmatrix}$$

Parameter	Manual	Automatic	Value Difference
$m_x c$ [-]	0.152	0.155	0.003
$m_y s$ [-]	0.137	0.131	-0.006
$m_x s$ [-]	0.124	0.126	0.001
$m_y c$ [-]	0.151	0.157	0.006
t_x [m]	307,522.72	307,533.69	10.96
t_y [m]	9,835,112.65	9,835,128.66	16.01

Experimental Results and Analysis

- ▶ Subject road dataset transformed using the approximate affine transformation parameters obtained for both processes.
- ▶ Results achieved by using the values extracted in the manual process (left) give an inferior geometric and topologic alignment in most areas than the automatic one (right).



Experimental Results and Analysis

- ▶ Number of corresponding points that were identified in the final automatic matching process together with TPS transformation is higher.
- ▶ Automatic registration is more qualitative than the manual one, which enabled the extraction of a more reliable correspondence of the two datasets.

	Manual	Automatic
Number of corresponding points	14	24
RMSE [m]	17.54	22.84
Number of corresponding points after matching	25	28

- ▶ Arbitrary selection of 10 line features was carried out (after alignment). Length of each one was measured in the two datasets, and compared to the reference one.

	Manual Dataset	Automatic Dataset
Mean Length Difference [%]	4.75	3.77

Discussion and Conclusions

- ▶ During geometric registration, identification of corresponding features is usually done manually.
- ▶ An alternative automatic process was presented, which relies on ranking process of Hausdorff distances measure.
- ▶ The values extracted via this process were statistically qualitative and reliable.
- ▶ A-priori knowledge - obtained from the datasets used.
- ▶ The automatic approach produces a more qualitative geometric registration than the manual approach.
- ▶ A further research is required to decrease the significance of the initial approximate values (using a series of sets, GA), and also to make the approach more tolerable to errors and outliers.

Discussion and Conclusions

- ▶ This approach is a preliminary step toward the establishment of an automatic computational basis required for the geometric quality enhancement of legacy analogue data.
- ▶ The approach is a step towards the integration of the legacy data while saving financial costs and avoiding technical challenges that would otherwise be involved.
- ▶ The enhanced data can therefore be used and included as a more precise and reliable SDI that is required for supporting governance as well as private sectors legislation processes.



Institute of Cartography and Geoinformatics | Leibniz Universität Hannover

Thank you



Leibniz
Universität
Hannover